

# Package: vivainsights (via r-universe)

September 9, 2024

**Type** Package

**Title** Analyze and Visualize Data from 'Microsoft Viva Insights'

**Version** 0.5.4

**Maintainer** Martin Chan <martin.chan@microsoft.com>

**Description** Provides a versatile range of functions, including exploratory data analysis, time-series analysis, organizational network analysis, and data validation, whilst at the same time implements a set of best practices in analyzing and visualizing data specific to 'Microsoft Viva Insights'.

**RoxygenNote** 7.3.2

**Roxygen** list(markdown = TRUE)

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**Language** en-US

**Depends** R (>= 3.1.2)

**Imports** dplyr, stats, utils, tidyr, tidyselect (>= 1.0.0), magrittr, ggplot2, reshape2, scales, ggrepel, purrr, data.table, methods, htmltools, markdown, networkD3, rmarkdown, wpa, ggraph, igraph, tidytext, ggwordcloud, lifecycle

**Suggests** flexdashboard, testthat (>= 3.0.0)

**URL** <https://microsoft.github.io/vivainsights/>

**BugReports** <https://github.com/microsoft/vivainsights/issues/>

**Config/testthat/edition** 3

**Repository** <https://microsoft.r-universe.dev>

**RemoteUrl** <https://github.com/microsoft/vivainsights>

**RemoteRef** HEAD

**RemoteSha** fd40e856cb0ff687cddb25026269fae6890774e

## Contents

afterhours_dist . . . . .	4
afterhours_fizz . . . . .	6
afterhours_line . . . . .	7
afterhours_rank . . . . .	9
afterhours_summary . . . . .	10
afterhours_trend . . . . .	12
anonymise . . . . .	13
any_idate . . . . .	14
camel_clean . . . . .	15
check_inputs . . . . .	15
check_query . . . . .	16
collaboration_area . . . . .	18
collaboration_dist . . . . .	19
collaboration_fizz . . . . .	21
collaboration_line . . . . .	22
collaboration_rank . . . . .	24
collaboration_sum . . . . .	26
collaboration_trend . . . . .	27
comma . . . . .	29
copy_df . . . . .	29
create_bar . . . . .	30
create_bar_asis . . . . .	32
create_boxplot . . . . .	34
create_bubble . . . . .	36
create_density . . . . .	37
create_dist . . . . .	39
create_dt . . . . .	41
create_fizz . . . . .	42
create_hist . . . . .	43
create_inc . . . . .	45
create_IV . . . . .	47
create_line . . . . .	49
create_line_asis . . . . .	51
create_lorenz . . . . .	52
create_period_scatter . . . . .	53
create_rank . . . . .	55
create_rank_combine . . . . .	57
create_sankey . . . . .	58
create_scatter . . . . .	60
create_stacked . . . . .	61
create_tracking . . . . .	64
create_trend . . . . .	65
cut_hour . . . . .	67
email_dist . . . . .	68
email_fizz . . . . .	69
email_line . . . . .	70

email_rank . . . . .	72
email_summary . . . . .	73
email_trend . . . . .	75
export . . . . .	76
external_dist . . . . .	77
external_fizz . . . . .	79
external_line . . . . .	80
external_rank . . . . .	81
external_sum . . . . .	83
extract_date_range . . . . .	84
extract_hr . . . . .	85
flag_ch_ratio . . . . .	86
flag_em_ratio . . . . .	87
flag_extreme . . . . .	88
flag_outlooktime . . . . .	90
g2g_data . . . . .	91
generate_report . . . . .	92
generate_report2 . . . . .	93
heat_colours . . . . .	94
hrvar_count . . . . .	95
hrvar_count_all . . . . .	96
hrvar_trend . . . . .	98
hr_trend . . . . .	99
identify_churn . . . . .	100
identify_datefreq . . . . .	102
identify_holidayweeks . . . . .	103
identify_inactiveweeks . . . . .	105
identify_nkw . . . . .	106
identify_outlier . . . . .	107
identify_privacythreshold . . . . .	108
identify_shifts . . . . .	109
identify_tenure . . . . .	110
import_query . . . . .	112
is_date_format . . . . .	113
IV_report . . . . .	113
jitter_metrics . . . . .	115
keymetrics_scan . . . . .	116
keymetrics_scan_asis . . . . .	117
maxmin . . . . .	119
meeting_dist . . . . .	120
meeting_fizz . . . . .	122
meeting_line . . . . .	123
meeting_rank . . . . .	124
meeting_summary . . . . .	126
meeting_tm_report . . . . .	127
meeting_trend . . . . .	128
mt_data . . . . .	130
network_g2g . . . . .	131

network_p2p . . . . .	134
network_summary . . . . .	138
one2one_dist . . . . .	139
one2one_fizz . . . . .	141
one2one_freq . . . . .	142
one2one_line . . . . .	144
one2one_rank . . . . .	145
one2one_sum . . . . .	147
one2one_trend . . . . .	148
p2p_data . . . . .	150
p2p_data_sim . . . . .	151
pad2 . . . . .	152
pairwise_count . . . . .	152
pq_data . . . . .	153
read_preamble . . . . .	158
rgb2hex . . . . .	158
theme_wpa . . . . .	159
theme_wpa_basic . . . . .	159
tm_clean . . . . .	160
tm_cooc . . . . .	161
tm_freq . . . . .	162
tm_wordcloud . . . . .	163
totals_bind . . . . .	165
totals_col . . . . .	166
track_HR_change . . . . .	167
tstamp . . . . .	168
us_to_space . . . . .	169
validation_report . . . . .	169
wrap . . . . .	171
wrap_text . . . . .	172

**Index****173**


---

afterhours_dist	<i>Distribution of After-hours Collaboration Hours as a 100% stacked bar</i>
-----------------	--

---

**Description**

Analyse the distribution of weekly after-hours collaboration time. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
afterhours_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(1, 2, 3)
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
cut	A vector specifying the cuts to use for the data, accepting "default" or "range-cut" as character vector, or a numeric value of length three to specify the exact breaks to use. e.g. c(1, 3, 5)

**Details**

Uses the metric `After_hours_collaboration_hours`. See `create_dist()` for applying the same analysis to a different metric.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#),

```
external_sum(), hr_trend(), hrvar_count(), hrvar_trend(), keymetrics_scan(), meeting_dist(),
meeting_fizz(), meeting_line(), meeting_rank(), meeting_summary(), meeting_trend(),
one2one_dist(), one2one_fizz(), one2one_freq(), one2one_line(), one2one_rank(), one2one_sum(),
one2one_trend()
```

```
Other After-hours Collaboration: afterhours_fizz(), afterhours_line(), afterhours_rank(),
afterhours_summary(), afterhours_trend(), external_rank()
```

## Examples

```
# Return plot
afterhours_dist(pq_data, hrvar = "Organization")

# Return summary table
afterhours_dist(pq_data, hrvar = "Organization", return = "table")

# Return result with a custom specified breaks
afterhours_dist(pq_data, hrvar = "LevelDesignation", cut = c(4, 7, 9))
```

---

afterhours\_fizz      *Distribution of After-hours Collaboration Hours (Fizzy Drink plot)*

---

## Description

Analyze weekly after-hours collaboration hours distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

## Usage

```
afterhours_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

## Details

Uses the metric After\_hours\_collaboration\_hours. See create\_fizz() for applying the same analysis to a different metric.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other After-hours Collaboration: [afterhours\\_dist\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [external\\_rank\(\)](#)

**Examples**

```
# Return plot
afterhours_fizz(pq_data, hrvar = "LevelDesignation", return = "plot")

# Return summary table
afterhours_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

afterhours\_line

*After-hours Collaboration Time Trend - Line Chart*

---

**Description**

Provides a week by week view of after-hours collaboration time, visualized as line charts. By default returns a line chart for after-hours collaboration hours, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

**Usage**

```
afterhours_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Details**

Uses the metric `After_hours_collaboration_hours`.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

[create\\_line\(\)](#) for applying the same analysis to a different metric.

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other After-hours Collaboration: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [external\\_rank\(\)](#)

**Examples**

```
# Return a line plot
afterhours_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
```

```
afterhours_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

afterhours\_rank      *Rank groups with high After-Hours Collaboration Hours*

---

### Description

This function scans a Standard Person Query for groups with high levels of After-Hours Collaboration. Returns a plot by default, with an option to return a table with all groups (across multiple HR attributes) ranked by hours of After-Hours Collaboration Hours.

### Usage

```
afterhours_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> <li>• "table"</li> </ul> <p>See Value for more information.</p>

**Details**

Uses the metric `After_hours_collaboration_hours`. See `create_rank()` for applying the same analysis to a different metric.

**Value**

When 'table' is passed in return, a summary table is returned as a data frame.

**See Also**

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other After-hours Collaboration: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_summary()`, `afterhours_trend()`, `external_rank()`

**Examples**

```
# Return plot
afterhours_rank(pq_data, return = "plot")

# Return summary table
afterhours_rank(pq_data, return = "table")
```

---

afterhours\_summary      *Summary of After-Hours Collaboration Hours*

---

**Description**

Provides an overview analysis of after-hours collaboration time. Returns a bar plot showing average weekly after-hours collaboration hours by default. Additional options available to return a summary table.

**Usage**

```
afterhours_summary(data, hrvar = "Organization", mingroup = 5, return = "plot")

afterhours_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Details**

Uses the metric After\_hours\_collaboration\_hours.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other After-hours Collaboration: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_trend\(\)](#), [external\\_rank\(\)](#)

**Examples**

```
# Return a ggplot bar chart
afterhours_summary(pq_data, hrvar = "LevelDesignation")

# Return a summary table
afterhours_summary(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

afterhours_trend	<i>After-Hours Time Trend</i>
------------------	-------------------------------

---

### Description

Provides a week by week view of after-hours collaboration time. By default returns a week by week heatmap, highlighting the points in time with most activity. Additional options available to return a summary table.

### Usage

```
afterhours_trend(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

### Details

Uses the metric `After_hours_collaboration_hours`.

### Value

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other After-hours Collaboration: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [external\\_rank\(\)](#)

### Examples

```
# Run plot
afterhours_trend(pq_data)

# Run table
afterhours_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

anonymise	<i>Anonymise a categorical variable by replacing values</i>
-----------	---

---

### Description

Anonymize categorical variables such as HR variables by replacing values with dummy team names such as 'Team A'. The behaviour is to make 1 to 1 replacements by default, but there is an option to completely randomise values in the categorical variable.

### Usage

```
anonymise(x, scramble = FALSE, replacement = NULL)

anonymize(x, scramble = FALSE, replacement = NULL)
```

### Arguments

x	Character vector to be passed through.
scramble	Logical value determining whether to randomise values in the categorical variable.
replacement	Character vector containing the values to replace original values in the categorical variable. The length of the vector must be at least as great as the number of unique values in the original variable. Defaults to NULL, where the replacement would consist of "Team A", "Team B", etc.

### Value

Character vector with the same length as input x, replaced with values provided in replacement.

### See Also

[jitter](#)

### Examples

```
unique(anonymise(pq_data$Organization))

rep <- c("Manager+", "Manager", "IC")
unique(anonymise(pq_data$Layer), replacement = rep)
```

---

any\_idate

*Identify whether variable is an IDate class.*

---

### Description

This function checks whether the variable is an IDate class.

### Usage

```
any_idate(x)
```

### Arguments

x Variable to test whether an IDate class.

### Value

logical value indicating whether the string is of an IDate class.

### See Also

Other Support: [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

### Examples

```
any_idate("2023-12-15")
```

---

camel_clean	<i>Convert "CamelCase" to "Camel Case"</i>
-------------	--

---

**Description**

Convert a text string from the format "CamelCase" to "Camel Case". This is used for converting variable names such as "LevelDesignation" to "Level Designation" for the purpose of prettifying plot labels.

**Usage**

```
camel_clean(string)
```

**Arguments**

string            A string vector in 'CamelCase' format to format

**Value**

Returns a formatted string.

**See Also**

Other Support: [any\\_idate\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

**Examples**

```
camel_clean("NoteHowTheStringIsFormatted")
```

---

check_inputs	<i>Check whether a data frame contains all the required variable</i>
--------------	--

---

**Description**

Checks whether a data frame contains all the required variables. Matching works via variable names, and used to support individual functions in the package. Not used directly.

**Usage**

```
check_inputs(input, requirements, return = "stop")
```

**Arguments**

input	Pass a data frame for checking
requirements	A character vector specifying the required variable names
return	A character string specifying what to return. The default value is "stop". Also accepts "names" and "warning".

**Value**

The default behaviour is to return an error message, informing the user what variables are not included. When return is set to "names", a character vector containing the unmatched variable names is returned.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

**Examples**

```
# Return error message
## Not run:
check_inputs(iris, c("Sepal.Length", "mpg"))

## End(Not run)

#' # Return warning message
check_inputs(iris, c("Sepal.Length", "mpg"), return = "warning")

# Return variable names
check_inputs(iris, c("Sepal.Length", "Sepal.Width", "RandomVariable"), return = "names")
```

---

 check\_query

---

*Check a query to ensure that it is suitable for analysis*


---

**Description**

Prints diagnostic data about the data query to the R console, with information such as date range, number of employees, HR attributes identified, etc.

**Usage**

```
check_query(data, return = "message", validation = FALSE)
```

## Arguments

data	<p>A person-level query in the form of a data frame. This includes:</p> <ul style="list-style-type: none"><li>• Standard Person Query</li><li>• Ways of Working Assessment Query</li><li>• Hourly Collaboration Query</li></ul> <p>All person-level query have a <code>PersonId</code> column and a <code>MetricDate</code> column.</p>
return	<p>String specifying what to return. This must be one of the following strings:</p> <ul style="list-style-type: none"><li>• "message" (default)</li><li>• "text"</li></ul> <p>See Value for more information.</p>
validation	<p>Logical value to specify whether to show summarized version. Defaults to FALSE. To hide checks on variable names, set <code>validation</code> to TRUE.</p>

## Details

This can be used with any person-level query, such as the standard person query, Ways of Working assessment query, and the hourly collaboration query. When run, this prints diagnostic data to the R console.

## Value

A different output is returned depending on the value passed to the `return` argument:

- "message": a message is returned to the console.
- "text": string containing the diagnostic message.

## See Also

Other Data Validation: [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

## Examples

```
check_query(pq_data)
```

---

collaboration\_area      *Collaboration - Stacked Area Plot*

---

### Description

Provides an overview analysis of Weekly Digital Collaboration. Returns an stacked area plot of Email and Meeting Hours by default. Additional options available to return a summary table.

### Usage

```
collaboration_area(data, hrvar = NULL, mingroup = 5, return = "plot")
```

```
collab_area(data, hrvar = NULL, mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame. A Ways of Working assessment dataset may also be provided, in which Unscheduled call hours would be included in the output.
hrvar	HR Variable by which to split metrics, defaults to NULL, but accepts any character vector, e.g. "LevelDesignation". If NULL is passed, the organizational attribute is automatically populated as "Total".
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Details

Uses the metrics Meeting\_hours, Email\_hours, Unscheduled\_Call\_hours, and Instant\_Message\_hours.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked area plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Collaboration: [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#)

**Examples**

```
# Return plot with total (default)
collaboration_area(pq_data)

# Return plot with hrvar split
collaboration_area(pq_data, hrvar = "Organization")

# Return summary table
collaboration_area(pq_data, return = "table")
```

---

`collaboration_dist`      *Distribution of Collaboration Hours as a 100% stacked bar*

---

**Description**

Analyze the distribution of Collaboration Hours. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
collaboration_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(15, 20, 25)
)

collab_dist(
```

```

data,
hrvar = "Organization",
mingroup = 5,
return = "plot",
cut = c(15, 20, 25)
)

```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
cut	A numeric vector of length three to specify the breaks for the distribution, e.g. c(10, 15, 20)

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

### Metrics used

The metric Collaboration\_hours is used in the calculations. Please ensure that your query contains a metric with the exact same name.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Collaboration: [collaboration\\_area\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#)

### Examples

```
# Return plot
collaboration_dist(pq_data, hrvar = "Organization")

# Return summary table
collaboration_dist(pq_data, hrvar = "Organization", return = "table")
```

---

collaboration\_fizz      *Distribution of Collaboration Hours (Fizzy Drink plot)*

---

### Description

Analyze weekly collaboration hours distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

### Usage

```
collaboration_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")

collab_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

**Metrics used**

The metric `Collaboration_hours` is used in the calculations. Please ensure that your query contains a metric with the exact same name.

**See Also**

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other Collaboration: `collaboration_area()`, `collaboration_dist()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`

**Examples**

```
# Return plot
collaboration_fizz(pq_data, hrvar = "Organization", return = "plot")

# Return summary table
collaboration_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

`collaboration_line`      *Collaboration Time Trend - Line Chart*

---

**Description**

Provides a week by week view of collaboration time, visualised as line charts. By default returns a line chart for collaboration hours, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

**Usage**

```
collaboration_line(data, hrvar = "Organization", mingroup = 5, return = "plot")

collab_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**Metrics used**

The metric Collaboration\_hours is used in the calculations. Please ensure that your query contains a metric with the exact same name.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Collaboration: [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#)

**Examples**

```
# Return a line plot
collaboration_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
```

```
collaboration_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

collaboration\_rank      *Collaboration Ranking*

---

## Description

This function scans a standard query output for groups with high levels of 'Weekly Digital Collaboration'. Returns a plot by default, with an option to return a table with a all of groups (across multiple HR attributes) ranked by hours of digital collaboration.

## Usage

```
collaboration_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

```
collab_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot".

- 1: Top and bottom five groups across the data population are highlighted
- 2: Top and bottom groups *per* organizational attribute are highlighted

return String specifying what to return. This must be one of the following strings:

- "plot" (default)
- "table"

See Value for more information.

### Details

Uses the metric `Collaboration_hours`. See `create_rank()` for applying the same analysis to a different metric.

### Value

A different output is returned depending on the value passed to the `return` argument:

- "plot": 'ggplot' object. A bubble plot where the x-axis represents the metric, the y-axis represents the HR attributes, and the size of the bubbles represent the size of the organizations. Note that there is no plot output if mode is set to "combine".
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other Collaboration: `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_sum()`, `collaboration_trend()`

### Examples

```
# Return rank table
collaboration_rank(
  data = pq_data,
  return = "table"
)
```

```
# Return plot
collaboration_rank(
  data = pq_data,
```

```

    return = "plot"
  )

```

---

collaboration\_sum      *Collaboration Summary*

---

### Description

Provides an overview analysis of 'Weekly Digital Collaboration'. Returns a stacked bar plot of Email and Meeting Hours by default. Additional options available to return a summary table.

### Usage

```
collaboration_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

```
collab_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

```
collaboration_summary(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot"
)
```

```
collab_summary(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

### Details

Uses the metrics Meeting\_hours, Email\_hours, Unscheduled\_Call\_hours, and Instant\_Message\_hours.

### Value

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other Collaboration: `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_trend()`

**Examples**

```
# Return a ggplot bar chart
collaboration_sum(pq_data, hrvar = "LevelDesignation")

# Return a summary table
collaboration_sum(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

`collaboration_trend`     *Collaboration Time Trend*

---

**Description**

Provides a week by week view of collaboration time. By default returns a week by week heatmap, highlighting the points in time with most activity. Additional options available to return a summary table.

**Usage**

```
collaboration_trend(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

**Value**

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**Metrics used**

The metric Collaboration\_hours is used in the calculations. Please ensure that your query contains a metric with the exact same name.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Collaboration: [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#)

**Examples**

```
# Run plot
collaboration_trend(pq_data)

# Run table
collaboration_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

comma	<i>Add comma separator for thousands</i>
-------	--

---

### Description

Takes a numeric value and returns a character value which is rounded to the whole number, and adds a comma separator at the thousands. A convenient wrapper function around `round()` and `format()`.

### Usage

```
comma(x)
```

### Arguments

x	A numeric value
---	-----------------

### Value

Returns a formatted string.

---

copy_df	<i>Copy a data frame to clipboard for pasting in Excel</i>
---------	--

---

### Description

This is a pipe-optimised function, that feeds into `vivainsights::export()`, but can be used as a stand-alone function.

Based on the original function from <https://github.com/martinctc/surveytoolbox>.

### Usage

```
copy_df(x, row.names = FALSE, col.names = TRUE, quietly = FALSE, ...)
```

### Arguments

x	Data frame to be passed through. Cannot contain list-columns or nested data frames.
row.names	A logical vector for specifying whether to allow row names. Defaults to FALSE.
col.names	A logical vector for specifying whether to allow column names. Defaults to FALSE.
quietly	Set this to TRUE to not print data frame on console
...	Additional arguments for <code>write.table()</code> .

**Value**

Copies a data frame to the clipboard with no return value.

**See Also**

Other Import and Export: [create\\_dt\(\)](#), [export\(\)](#), [import\\_query\(\)](#)

---

 create\_bar

*Mean Bar Plot for any metric*


---

**Description**

Provides an overview analysis of a selected metric by calculating a mean per metric. Returns a bar plot showing the average of a selected metric by default. Additional options available to return a summary table.

**Usage**

```
create_bar(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  bar_colour = "default",
  na.rm = FALSE,
  percent = FALSE,
  plot_title = us_to_space(metric),
  plot_subtitle = paste("Average by", tolower(camel_clean(hrvar))),
  legend_lab = NULL,
  rank = "descending",
  xlim = NULL,
  text_just = 0.5,
  text_colour = "#FFFFFF"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings:

	<ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul>
	See Value for more information.
bar_colour	String to specify colour to use for bars. In-built accepted values include "default" (default), "alert" (red), and "darkblue". Otherwise, hex codes are also accepted. You can also supply RGB values via <code>rgb2hex()</code> .
na.rm	A logical value indicating whether NA should be stripped before the computation proceeds. Defaults to FALSE.
percent	Logical value to determine whether to show labels as percentage signs. Defaults to FALSE.
plot_title	An option to override plot title.
plot_subtitle	An option to override plot subtitle.
legend_lab	String. Option to override legend title/label. Defaults to NULL, where the metric name will be populated instead.
rank	String specifying how to rank the bars. Valid inputs are: <ul style="list-style-type: none"> <li>• "descending" - ranked highest to lowest from top to bottom (default).</li> <li>• "ascending" - ranked lowest to highest from top to bottom.</li> <li>• NULL - uses the original levels of the HR attribute.</li> </ul>
xlim	An option to set max value in x axis.
text_just	<b>[Experimental]</b> A numeric value controlling for the horizontal position of the text labels. Defaults to 0.5.
text_colour	<b>[Experimental]</b> String to specify colour to use for the text labels. Defaults to "#FFFFFF".

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bar plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

## Examples

```
# Return a ggplot bar chart
create_bar(pq_data, metric = "Collaboration_hours", hrvar = "LevelDesignation")

# Change bar colour
create_bar(pq_data,
          metric = "After_hours_collaboration_hours",
          bar_colour = "alert")

# Custom data label positions and formatting
pq_data %>%
  create_bar(
    metric = "Meetings",
    text_just = 1.1,
    text_colour = "black",
    xlim = 20)

# Return a summary table
create_bar(pq_data,
          metric = "Collaboration_hours",
          hrvar = "LevelDesignation",
          return = "table")
```

---

create\_bar\_asis

*Create a bar chart without aggregation for any metric*

---

## Description

This function creates a bar chart directly from the aggregated / summarised data. Unlike `create_bar()` which performs a person-level aggregation, there is no calculation for `create_bar_asis()` and the values are rendered as they are passed into the function.

## Usage

```
create_bar_asis(
  data,
  group_var,
  bar_var,
  title = NULL,
  subtitle = NULL,
  caption = NULL,
  ylab = group_var,
  xlab = bar_var,
```

```

    percent = FALSE,
    bar_colour = "default",
    rounding = 1
  )

```

### Arguments

data	Plotting data as a data frame.
group_var	String containing name of variable for the group.
bar_var	String containing name of variable representing the value of the bars.
title	Title of the plot.
subtitle	Subtitle of the plot.
caption	Caption of the plot.
ylab	Y-axis label for the plot (group axis)
xlab	X-axis label of the plot (bar axis).
percent	Logical value to determine whether to show labels as percentage signs. Defaults to FALSE.
bar_colour	String to specify colour to use for bars. In-built accepted values include "default" (default), "alert" (red), and "darkblue". Otherwise, hex codes are also accepted. You can also supply RGB values via <code>rgb2hex()</code> .
rounding	Numeric value to specify number of digits to show in data labels

### Value

'ggplot' object. A horizontal bar plot.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```

# Creating a custom bar plot without mean aggregation
library(dplyr)

pq_data %>%
  group_by(Organization) %>%
  summarise(across(.cols = Meeting_hours,
                  .fns = ~sum(., na.rm = TRUE))) %>%
  create_bar_asis(group_var = "Organization",
                 bar_var = "Meeting_hours",
                 title = "Total Meeting Hours over period",
                 subtitle = "By Organization",
                 caption = extract_date_range(pq_data, return = "text"),
                 bar_colour = "darkblue",
                 rounding = 0)

library(dplyr)

# Summarise Non-person-average median `Emails_sent`
med_df <-
  pq_data %>%
  group_by(Organization) %>%
  summarise(Emails_sent_median = median(Emails_sent))

med_df %>%
  create_bar_asis(
    group_var = "Organization",
    bar_var = "Emails_sent_median",
    title = "Emails sent by organization",
    subtitle = "Median values",
    bar_colour = "darkblue",
    caption = extract_date_range(pq_data, return = "text")
  )

```

---

create\_boxplot

*Box Plot for any metric*


---

**Description**

Analyzes a selected metric and returns a box plot by default. Additional options available to return a table with distribution elements.

**Usage**

```

create_boxplot(
  data,
  metric,

```

```

    hrvar = "Organization",
    mingroup = 5,
    return = "plot"
  )

```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Details

This is a general purpose function that powers all the functions in the package that produce box plots.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A box plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
# Create a box plot for Collaboration_hours by Level Designation
create_boxplot(pq_data, metric = "Collaboration_hours", hrvar = "LevelDesignation", return = "plot")

# Create a box plot for Collaboration_hours by Organization
create_boxplot(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "plot")

# Create a summary statistics table for Collaboration_hoursby Organization
create_boxplot(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "table")
```

---

create_bubble	<i>Create a bubble plot with two selected Viva Insights metrics (General Purpose), with size representing the number of employees in the group.</i>
---------------	---

---

**Description**

Returns a bubble plot of two selected metrics, using size to map the number of employees.

**Usage**

```
create_bubble(
  data,
  metric_x,
  metric_y,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  bubble_size = c(1, 10)
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric_x	Character string containing the name of the metric, e.g. "Collaboration_hours"
metric_y	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	HR Variable by which to split metrics, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation"
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: - "plot" - "table"
bubble_size	A numeric vector of length two to specify the size range of the bubbles

**Details**

This is a general purpose function that powers all the functions in the package that produce bubble plots.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bubble plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
create_bubble(pq_data, "Collaboration_hours", "Multitasking_hours", hrvar ="Organization")
```

---

create\_density

*Create a density plot for any metric*

---

**Description**

Provides an analysis of the distribution of a selected metric. Returns a faceted density plot by default. Additional options available to return the underlying frequency table.

**Usage**

```
create_density(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  ncol = NULL,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	String containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
ncol	Numeric value setting the number of columns on the plot. Defaults to NULL (automatic).
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> <li>• "data"</li> <li>• "frequency"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted density plot for the metric.
- "table": data frame. A summary table for the metric.
- "data": data frame. Data with calculated person averages.
- "frequency": list of data frames. Each data frame contains the frequencies used in each panel of the plotted histogram.

**See Also**

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
# Return plot for whole organization
create_density(pq_data, metric = "Collaboration_hours", hrvar = NULL)

# Return plot
create_density(pq_data, metric = "Collaboration_hours", hrvar = "Organization")

# Return plot but coerce plot to three columns
create_density(pq_data, metric = "Collaboration_hours", hrvar = "Organization", ncol = 3)

# Return summary table
create_density(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "table")
```

---

create\_dist

*Horizontal 100 percent stacked bar plot for any metric*


---

**Description**

Provides an analysis of the distribution of a selected metric. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
create_dist(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(15, 20, 25),
  dist_colours = c("#facebc", "#fcf0eb", "#b4d5dd", "#bfe5ee"),
  unit = "hours",
  lbound = 0,
  ubound = 200,
  sort_by = NULL,
  labels = NULL
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	String containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.

return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
cut	A numeric vector of length three to specify the breaks for the distribution, e.g. <code>c(10, 15, 20)</code>
dist_colours	A character vector of length four to specify colour codes for the stacked bars.
unit	String to specify what unit to use. This defaults to "hours" but can accept any custom string. See <code>cut_hour()</code> for more details.
lbound	Numeric. Specifies the lower bound (inclusive) value for the minimum label. Defaults to 0.
ubound	Numeric. Specifies the upper bound (inclusive) value for the maximum label. Defaults to 100.
sort_by	String to specify the bucket label to sort by. Defaults to NULL (no sorting).
labels	Character vector to override labels for the created categorical variables. Must be a named vector - see examples.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

## Examples

```
# Return plot
create_dist(pq_data, metric = "Collaboration_hours", hrvar = "Organization")

# Return summary table
create_dist(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "table")

# Use custom labels by providing a label vector
eh_labels <- c(
  "Fewer than fifteen" = "< 15 hours",
  "Between fifteen and twenty" = "15 - 20 hours",
  "Between twenty and twenty-five" = "20 - 25 hours",
  "More than twenty-five" = "25+ hours"
)

pq_data %>% create_dist(metric = "Meeting_hours", labels = eh_labels, return = "plot")

# Sort by a category
pq_data %>% create_dist(metric = "Collaboration_hours", sort_by = "25+ hours")
```

---

create\_dt

*Create interactive tables in HTML with 'download' buttons.*

---

## Description

See <https://martinctc.github.io/blog/vignette-downloadable-tables-in-rmarkdown-with-the-dt-package/> for more.

## Usage

```
create_dt(x, rounding = 1, freeze = 2, percent = FALSE)
```

## Arguments

x	Data frame to be passed through.
rounding	Numeric vector to specify the number of decimal points to display
freeze	Number of columns from the left to 'freeze'. Defaults to 2, which includes the row number column.
percent	Logical value specifying whether to display numeric columns as percentages.

## Details

This is exported from wpa: :create\_dt().

## Value

Returns an HTML widget displaying rectangular data.

**See Also**

Other Import and Export: [copy\\_df\(\)](#), [export\(\)](#), [import\\_query\(\)](#)

**Examples**

```
output <- hrvar_count(pq_data, return = "table")
create_dt(output)
```

---

 create\_fizz

*Fizzy Drink / Jittered Scatter Plot for any metric*


---

**Description**

Analyzes a selected metric and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
create_fizz(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Details**

This is a general purpose function that powers all the functions in the package that produce 'fizzy drink' / jittered scatter plots.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
# Create a fizzy plot for Collaboration hours by Level Designation
create_fizz(pq_data, metric = "Collaboration_hours", hrvar = "LevelDesignation", return = "plot")

# Create a summary statistics table for Collaboration hours by Organization
create_fizz(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "table")
```

---

create\_hist

*Create a histogram plot for any metric*

---

**Description**

Provides an analysis of the distribution of a selected metric. Returns a faceted histogram by default. Additional options available to return the underlying frequency table.

**Usage**

```
create_hist(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  binwidth = 1,
  ncol = NULL,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	String containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
binwidth	Numeric value for setting binwidth argument within <code>ggplot2::geom_histogram()</code> . Defaults to 1.
ncol	Numeric value setting the number of columns on the plot. Defaults to NULL (automatic).
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> <li>• "data"</li> <li>• "frequency"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted histogram for the metric.
- "table": data frame. A summary table for the metric.
- "data": data frame. Data with calculated person averages.
- "frequency": list of data frames. Each data frame contains the frequencies used in each panel of the plotted histogram.

**See Also**

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```

# Return plot for whole organization
create_hist(pq_data, metric = "Collaboration_hours", hrvar = NULL)

# Return plot
create_hist(pq_data, metric = "Collaboration_hours", hrvar = "Organization")

# Return plot but coerce plot to 3 columns
create_hist(pq_data, metric = "Collaboration_hours", hrvar = "Organization", ncol = 3)

# Return summary table
create_hist(pq_data, metric = "Collaboration_hours", hrvar = "Organization", return = "table")

```

---

create_inc	<i>Create an incidence analysis reflecting proportion of population scoring above or below a threshold for a metric</i>
------------	---

---

**Description**

An incidence analysis is generated, with each value in the table reflecting the proportion of the population that is above or below a threshold for a specified metric. There is an option to only provide a single hrvar in which a bar plot is generated, or two hrvar values where an incidence table (heatmap) is generated.

**Usage**

```

create_inc(
  data,
  metric,
  hrvar,
  mingroup = 5,
  threshold,
  position,
  return = "plot"
)

create_incidence(
  data,
  metric,
  hrvar,
  mingroup = 5,
  threshold,
  position,
  return = "plot"
)

```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	Character vector of at most length 2 containing the name of the HR Variable by which to split metrics.
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
threshold	Numeric value specifying the threshold.
position	String containing the below valid values: <ul style="list-style-type: none"> <li>• "above": show incidence of those equal to or above the threshold</li> <li>• "below": show incidence of those equal to or below the threshold</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> <p>See Value for more information.</p>

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A heat map.
- "table": data frame. A summary table.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

## Examples

```
# Only a single HR attribute
create_inc(
  data = pq_data,
  metric = "After_hours_collaboration_hours",
  hrvar = "Organization",
  threshold = 4,
  position = "above"
)

# Two HR attributes
create_inc(
  data = pq_data,
  metric = "Collaboration_hours",
  hrvar = c("LevelDesignation", "Organization"),
  threshold = 20,
  position = "below"
)
```

---

create\_IV

*Compute Information Value for Predictive Variables*

---

## Description

This function calculates the Information Value (IV) for the selected numeric predictor variables in the dataset, given a specified outcome variable. The Information Value provides a measure of the predictive power of each variable in relation to the outcome variable, which can be useful in feature selection for predictive modeling.

## Usage

```
create_IV(
  data,
  predictors = NULL,
  outcome,
  bins = 5,
  siglevel = 0.05,
  exc_sig = FALSE,
  return = "plot"
)
```

## Arguments

data	A Person Query dataset in the form of a data frame.
predictors	A character vector specifying the columns to be used as predictors. Defaults to NULL, where all numeric vectors in the data will be used as predictors.

outcome	String specifying the column name for a binary variable, containing only the values 1 or 0.
bins	Number of bins to use, defaults to 5.
siglevel	Significance level to use in comparing populations for the outcomes, defaults to 0.05
exc_sig	Logical value determining whether to exclude values where the p-value lies below what is set at siglevel. Defaults to FALSE, where p-value calculation does not happen altogether.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "summary"</li> <li>• "list"</li> <li>• "plot-WOE"</li> <li>• "IV"</li> </ul> See Value for more information.

### Details

This is a wrapper around `wpa::create_IV()`.

### Value

A different output is returned depending on the value passed to the `return` argument:

- "plot": 'ggplot' object. A bar plot showing the IV value of the top (maximum 12) variables.
- "summary": data frame. A summary table for the metric.
- "list": list. A list of outputs for all the input variables.
- "plot-WOE": A list of 'ggplot' objects that show the WOE for each predictor used in the model.
- "IV" returns a list object which mirrors the return in `Information::create_infotables()`.

### See Also

Other Variable Association: [IV\\_report\(\)](#)

Other Information Value: [IV\\_report\(\)](#)

### Examples

```
# Return a summary table of IV
pq_data %>%
  dplyr::mutate(X = ifelse(Internal_network_size > 40, 1, 0)) %>%
  create_IV(outcome = "X",
            predictors = c("Email_hours",
                          "Meeting_hours",
                          "Chat_hours"),
            return = "plot")
```

```
# Return summary
pq_data %>%
  dplyr::mutate(X = ifelse(Internal_network_size > 40, 1, 0)) %>%
  create_IV(outcome = "X",
            predictors = c("Email_hours", "Meeting_hours"),
            return = "summary")
```

---

create\_line

*Time Trend - Line Chart for any metric*


---

### Description

Provides a week by week view of a selected metric, visualised as line charts. By default returns a line chart for the defined metric, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

### Usage

```
create_line(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  ncol = NULL,
  return = "plot"
)
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
ncol	Numeric value setting the number of columns on the plot. Defaults to NULL (automatic).
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul>

See Value for more information.

## Details

This is a general purpose function that powers all the functions in the package that produce faceted line plots.

## Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

## See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

Other Time-series: [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_trend\(\)](#)

## Examples

```
# Return plot of Email Hours
pq_data %>% create_line(metric = "Email_hours", return = "plot")

# Return plot of Collaboration Hours
pq_data %>% create_line(metric = "Collaboration_hours", return = "plot")

# Return plot but coerce plot to two columns
pq_data %>%
  create_line(
    metric = "Collaboration_hours",
    hrvar = "Organization",
    ncol = 2
  )

# Return plot of email hours and cut by `LevelDesignation`
pq_data %>% create_line(metric = "Email_hours", hrvar = "LevelDesignation")
```

---

create_line_asis	<i>Create a line chart without aggregation for any metric</i>
------------------	---

---

## Description

This function creates a line chart directly from the aggregated / summarised data. Unlike `create_line()` which performs a person-level aggregation, there is no calculation for `create_line_asis()` and the values are rendered as they are passed into the function. The only requirement is that a `date_var` is provided for the x-axis.

## Usage

```
create_line_asis(  
  data,  
  date_var = "MetricDate",  
  metric,  
  title = NULL,  
  subtitle = NULL,  
  caption = NULL,  
  ylab = date_var,  
  xlab = metric,  
  line_colour = rgb2hex(0, 120, 212)  
)
```

## Arguments

<code>data</code>	Plotting data as a data frame.
<code>date_var</code>	String containing name of variable for the horizontal axis.
<code>metric</code>	String containing name of variable representing the line.
<code>title</code>	Title of the plot.
<code>subtitle</code>	Subtitle of the plot.
<code>caption</code>	Caption of the plot.
<code>ylab</code>	Y-axis label for the plot (group axis)
<code>xlab</code>	X-axis label of the plot (bar axis).
<code>line_colour</code>	String to specify colour to use for the line. Hex codes are accepted. You can also supply RGB values via <code>rgb2hex()</code> .

## Value

Returns a 'ggplot' object representing a line plot.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

Other Time-series: [create\\_line\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
library(dplyr)

# Median `Emails_sent` grouped by `MetricDate`
# Without Person Averaging
med_df <-
  pq_data %>%
    group_by(MetricDate) %>%
    summarise(Emails_sent_median = median(Emails_sent))

med_df %>%
  create_line_asis(
    date_var = "MetricDate",
    metric = "Emails_sent_median",
    title = "Median Emails Sent",
    subtitle = "Person Averaging Not Applied",
    caption = extract_date_range(pq_data, return = "text")
  )
```

---

 create\_lorenz

---

*Calculate the Lorenz Curve and Gini Coefficient in a Person Query*


---

**Description**

This function computes the Gini coefficient and plots the Lorenz curve based on a selected metric from a Person Query data frame. It provides a way to measure inequality in the distribution of the selected metric. This function can be integrated into a larger analysis pipeline to assess inequality in metric distribution.

**Usage**

```
create_lorenz(data, metric, return = "plot")
```

**Arguments**

data	Data frame containing a Person Query.
metric	Character string identifying the metric to be used for the Lorenz curve and Gini coefficient calculation.
return	Character string identifying the return type. Options are: <ul style="list-style-type: none"> <li>• "gini" - Numeric value representing the Gini coefficient.</li> <li>• "table" - Data frame containing a summary table of population share and value share.</li> <li>• "plot" (default) - ggplot object representing a plot of the Lorenz curve.</li> </ul>

**Examples**

```
create_lorenz(data = pq_data, metric = "Emails_sent", return = "gini")
create_lorenz(data = pq_data, metric = "Emails_sent", return = "plot")
create_lorenz(data = pq_data, metric = "Emails_sent", return = "table")
```

---

```
create_period_scatter Period comparison scatter plot for any two metrics
```

---

**Description**

Returns two side-by-side scatter plots representing two selected metrics, using colour to map an HR attribute and size to represent number of employees. Returns a faceted scatter plot by default, with additional options to return a summary table.

**Usage**

```
create_period_scatter(
  data,
  hrvar = "Organization",
  metric_x = "Large_and_long_meeting_hours",
  metric_y = "Meeting_hours",
  before_start = min(as.Date(data$MetricDate, "%m/%d/%Y")),
  before_end,
  after_start = as.Date(before_end) + 1,
  after_end = max(as.Date(data$MetricDate, "%m/%d/%Y")),
  before_label = "Period 1",
  after_label = "Period 2",
  mingroup = 5,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	HR Variable by which to split metrics. Accepts a character vector, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation"
metric_x	Character string containing the name of the metric, e.g. "Collaboration_hours"
metric_y	Character string containing the name of the metric, e.g. "Collaboration_hours"
before_start	Start date of "before" time period in YYYY-MM-DD
before_end	End date of "before" time period in YYYY-MM-DD
after_start	Start date of "after" time period in YYYY-MM-DD
after_end	End date of "after" time period in YYYY-MM-DD
before_label	String to specify a label for the "before" period. Defaults to "Period 1".
after_label	String to specify a label for the "after" period. Defaults to "Period 2".
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

**Details**

This is a general purpose function that powers all the functions in the package that produce faceted scatter plots.

**Value**

Returns a 'ggplot' object showing two scatter plots side by side representing the two periods.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

Other Time-series: [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
# Return plot
create_period_scatter(pq_data,
                      hrvar = "LevelDesignation",
                      before_start = "2022-05-01",
                      before_end = "2022-05-31",
                      after_start = "2022-06-01",
                      after_end = "2022-07-03")

# Return a summary table
create_period_scatter(pq_data, before_end = "2022-05-31", return = "table")
```

---

create_rank	<i>Rank all groups across HR attributes on a selected Viva Insights metric</i>
-------------	--

---

**Description**

This function scans a standard Person query output for groups with high levels of a given Viva Insights Metric. Returns a plot by default, with an option to return a table with all groups (across multiple HR attributes) ranked by the specified metric.

**Usage**

```
create_rank(
  data,
  metric,
  hrvar = extract_hr(data, exclude_constants = TRUE),
  mingroup = 5,
  return = "table",
  mode = "simple",
  plot_mode = 1
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> </ul>

	<ul style="list-style-type: none"> <li>• "table"</li> </ul>
	See Value for more information.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bubble plot where the x-axis represents the metric, the y-axis represents the HR attributes, and the size of the bubbles represent the size of the organizations. Note that there is no plot output if mode is set to "combine".
- "table": data frame. A summary table for the metric.

### Author(s)

Carlos Morales Torrado [carlos.morales@microsoft.com](mailto:carlos.morales@microsoft.com)

Martin Chan [martin.chan@microsoft.com](mailto:martin.chan@microsoft.com)

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

### Examples

```
pq_data_small <- dplyr::slice_sample(pq_data, prop = 0.1)

# Plot mode 1 - show top and bottom five groups
```

```
create_rank(  
  data = pq_data_small,  
  hrvar = c("FunctionType", "LevelDesignation"),  
  metric = "Emails_sent",  
  return = "plot",  
  plot_mode = 1  
)  
  
# Plot mode 2 - show top and bottom groups per HR variable  
create_rank(  
  data = pq_data_small,  
  hrvar = c("FunctionType", "LevelDesignation"),  
  metric = "Emails_sent",  
  return = "plot",  
  plot_mode = 2  
)  
  
# Return a table  
create_rank(  
  data = pq_data_small,  
  metric = "Emails_sent",  
  return = "table"  
)  
  
# Return a table - combination mode  
create_rank(  
  data = pq_data_small,  
  metric = "Emails_sent",  
  mode = "combine",  
  return = "table"  
)
```

---

create\_rank\_combine     *Create combination pairs of HR variables and run 'create\_rank()'*

---

## Description

Create pairwise combinations of HR variables and compute an average of a specified advanced insights metric.

## Usage

```
create_rank_combine(data, hrvar = extract_hr(data), metric, mingroup = 5)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.

**Details**

This function is called when the mode argument in create\_rank() is specified as "combine".

**Value**

Data frame containing the following variables:

- hrvar: placeholder column that denotes the output as "Combined".
- group: pairwise combinations of HR attributes with the HR attribute in square brackets followed by the value of the HR attribute.
- Name of the metric (as passed to metric)
- n

**Examples**

```
# Use a small sample for faster runtime
pq_data_small <- dplyr::slice_sample(pq_data, prop = 0.1)

create_rank_combine(
  data = pq_data_small,
  metric = "Email_hours",
  hrvar = c("Organization", "FunctionType", "LevelDesignation")
)
```

---

create\_sankey

---

*Create a sankey chart from a two-column count table*


---

**Description**

Create a 'networkD3' style sankey chart based on a long count table with two variables. The input data should have three columns, where each row is a unique group:

1. Variable 1
2. Variable 2
3. Count

**Usage**

```
create_sankey(data, var1, var2, count = "n")
```

**Arguments**

data	Data frame of the long count table.
var1	String containing the name of the variable to be shown on the left.
var2	String containing the name of the variable to be shown on the right.
count	String containing the name of the count variable.

**Value**

A 'sankeyNetwork' and 'htmlwidget' object containing a two-tier sankey plot. The output can be saved locally with `htmlwidgets::saveWidget()`.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
pq_data %>%
  dplyr::count(Organization, FunctionType) %>%
  create_sankey(var1 = "Organization", var2 = "FunctionType")
```

---

create_scatter	<i>Create a Scatter plot with two selected Viva Insights metrics (General Purpose)</i>
----------------	--

---

### Description

Returns a scatter plot of two selected metrics, using colour to map an HR attribute. Returns a scatter plot by default, with additional options to return a summary table.

### Usage

```
create_scatter(  
  data,  
  metric_x,  
  metric_y,  
  hrvar = "Organization",  
  mingroup = 5,  
  return = "plot"  
)
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
metric_x	Character string containing the name of the metric, e.g. "Collaboration_hours"
metric_y	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	HR Variable by which to split metrics, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation"
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

### Details

This is a general purpose function that powers all the functions in the package that produce scatter plots.

### Value

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

**Examples**

```
create_scatter(
  pq_data,
  metric_x = "Collaboration_hours",
  metric_y = "Multitasking_hours",
  hrvar = "Organization"
)

create_scatter(
  pq_data,
  metric_x = "Collaboration_hours",
  metric_y = "Multitasking_hours",
  hrvar = "Organization",
  mingroup = 100,
  return = "plot"
)
```

---

create\_stacked

*Horizontal stacked bar plot for any metric*

---

**Description**

Creates either a single bar plot, of a stacked bar using selected metrics (where the typical use case is to create different definitions of collaboration hours). Returns a plot by default. Additional options available to return a summary table.

**Usage**

```

create_stacked(
  data,
  hrvar = "Organization",
  metrics = c("Meeting_hours", "Email_hours"),
  mingroup = 5,
  return = "plot",
  stack_colours = c("#1d627e", "#34b1e2", "#b4d5dd", "#adc0cb"),
  percent = FALSE,
  plot_title = "Collaboration Hours",
  plot_subtitle = paste("Average by", tolower(camel_clean(hrvar))),
  legend_lab = NULL,
  rank = "descending",
  xlim = NULL,
  text_just = 0.5,
  text_colour = "#FFFFFF"
)

```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>hrvar</code>	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
<code>metrics</code>	A character vector to specify variables to be used in calculating the "Total" value, e.g. c("Meeting_hours", "Email_hours"). The order of the variable names supplied determine the order in which they appear on the stacked plot.
<code>mingroup</code>	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
<code>return</code>	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".
<code>stack_colours</code>	A character vector to specify the colour codes for the stacked bar charts.
<code>percent</code>	Logical value to determine whether to show labels as percentage signs. Defaults to FALSE.
<code>plot_title</code>	String. Option to override plot title.
<code>plot_subtitle</code>	String. Option to override plot subtitle.
<code>legend_lab</code>	String. Option to override legend title/label. Defaults to NULL, where the metric name will be populated instead.
<code>rank</code>	String specifying how to rank the bars. Valid inputs are: <ul style="list-style-type: none"> <li>"descending" - ranked highest to lowest from top to bottom (default).</li> <li>"ascending" - ranked lowest to highest from top to bottom.</li> <li>NULL - uses the original levels of the HR attribute.</li> </ul>
<code>xlim</code>	An option to set max value in x axis.

text_just	<b>[Experimental]</b> A numeric value controlling for the horizontal position of the text labels. Defaults to 0.5.
text_colour	<b>[Experimental]</b> String to specify colour to use for the text labels. Defaults to "#FFFFFF".

### Value

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#)

### Examples

```
pq_data %>%
  create_stacked(hrvar = "LevelDesignation",
                metrics = c("Meeting_hours", "Email_hours"),
                return = "plot")
```

```
pq_data %>%
  create_stacked(hrvar = "FunctionType",
                metrics = c("Meeting_hours",
                          "Email_hours",
                          "Call_hours",
                          "Chat_hours"),
                return = "plot",
                rank = "ascending")
```

```
pq_data %>%
  create_stacked(hrvar = "FunctionType",
                metrics = c("Meeting_hours",
                          "Email_hours",
                          "Call_hours",
                          "Chat_hours"),
```

```
return = "table")
```

---

create_tracking	<i>Create a line chart that tracks metrics over time with a 4-week rolling average</i>
-----------------	--

---

## Description

### [Experimental]

Create a two-series line chart that visualizes a set of metric over time for the selected population, with one of the series being a four-week rolling average.

## Usage

```
create_tracking(
  data,
  metric,
  plot_title = us_to_space(metric),
  plot_subtitle = "Measure over time",
  percent = FALSE
)
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours" percentage signs. Defaults to FALSE.
plot_title	An option to override plot title.
plot_subtitle	An option to override plot subtitle.
percent	Logical value to determine whether to show labels as percentage signs. Defaults to FALSE.

## Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A time-series plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_trend\(\)](#)

**Examples**

```

pq_data %>%
  create_tracking(
    metric = "Collaboration_hours",
    percent = FALSE
  )

```

---

create\_trend

*Heat mapped horizontal bar plot over time for any metric*

---

**Description**

Provides a week by week view of a selected Viva Insights metric. By default returns a week by week heatmap bar plot, highlighting the points in time with most activity. Additional options available to return a summary table.

**Usage**

```

create_trend(
  data,
  metric,
  hrvar = "Organization",
  mingroup = 5,
  palette = c("steelblue4", "aliceblue", "white", "mistyrose1", "tomato1"),
  return = "plot",
  legend_title = "Hours"
)

```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
metric	Character string containing the name of the metric, e.g. "Collaboration_hours"
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
palette	Character vector containing colour codes, ranked from the lowest value to the highest value. This is passed directly to <code>ggplot2::scale_fill_gradientn()</code> .
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".
legend_title	String to be used as the title of the legend. Defaults to "Hours".

**Value**

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Flexible: [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_density\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_hist\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#)

Other Time-series: [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#)

**Examples**

```
create_trend(pq_data, metric = "Collaboration_hours", hrvar = "LevelDesignation")

# custom colours
create_trend(
  pq_data,
  metric = "Collaboration_hours",
  hrvar = "LevelDesignation",
```

```
palette = c(
  "#FB6107",
  "#F3DE2C",
  "#7CB518",
  "#5C8001"
)
)
```

---

cut_hour	<i>Convert a numeric variable for hours into categorical</i>
----------	--

---

### Description

Supply a numeric variable, e.g. `Collaboration_hours`, and return a character vector.

### Usage

```
cut_hour(metric, cuts, unit = "hours", lbound = 0, ubound = 100)
```

### Arguments

<code>metric</code>	A numeric variable representing hours.
<code>cuts</code>	A numeric vector of minimum length 3 to represent the cut points required. The minimum and maximum values provided in the vector are inclusive.
<code>unit</code>	String to specify the unit of the labels. Defaults to "hours".
<code>lbound</code>	Numeric. Specifies the lower bound (inclusive) value for the minimum label. Defaults to 0.
<code>ubound</code>	Numeric. Specifies the upper bound (inclusive) value for the maximum label. Defaults to 100.

### Details

This is used within `create_dist()` for numeric to categorical conversion.

### Value

Character vector representing a converted categorical variable, appended with the label of the unit. See examples for more information.

### See Also

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

**Examples**

```
# Direct use
cut_hour(1:30, cuts = c(15, 20, 25))

# Use on a query
cut_hour(pq_data$Collaboration_hours, cuts = c(10, 15, 20), ubound = 150)
```

---

email\_dist

*Distribution of Email Hours as a 100% stacked bar*


---

**Description**

Analyze Email Hours distribution. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
email_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(0.5, 1, 1.5)
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.
cut	A numeric vector of length three to specify the breaks for the distribution, e.g. c(10, 15, 20)

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#)

**Examples**

```
# Return plot
email_dist(pq_data, hrvar = "Organization")

# Return summary table
email_dist(pq_data, hrvar = "Organization", return = "table")

# Return result with a custom specified breaks
email_dist(pq_data, hrvar = "LevelDesignation", cut = c(1, 2, 3))
```

---

email\_fizz

*Distribution of Email Hours (Fizzy Drink plot)*


---

**Description**

Analyze weekly email hours distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
email_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.

return           String specifying what to return. This must be one of the following strings:

- "plot"
- "table"

See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_dist\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#)

### Examples

```
# Return plot
email_fizz(pq_data, hrvar = "Organization", return = "plot")

# Return summary table
email_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

email\_line

*Email Time Trend - Line Chart*

---

### Description

Provides a week by week view of email time, visualised as line charts. By default returns a line chart for email hours, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

### Usage

```
email_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#)

**Examples**

```
# Return a line plot
email_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
email_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

email_rank	<i>Email Hours Ranking</i>
------------	----------------------------

---

### Description

This function scans a standard query output for groups with high levels of 'Weekly Email Collaboration'. Returns a plot by default, with an option to return a table with a all of groups (across multiple HR attributes) ranked by hours of digital collaboration.

### Usage

```
email_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> <li>• "table"</li> </ul> See Value for more information.

### Details

Uses the metric Email\_hours. See create\_rank() for applying the same analysis to a different metric.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bubble plot where the x-axis represents the metric, the y-axis represents the HR attributes, and the size of the bubbles represent the size of the organizations. Note that there is no plot output if mode is set to "combine".
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#)

**Examples**

```
# Return rank table
email_rank(
  data = pq_data,
  return = "table"
)

# Return plot
email_rank(
  data = pq_data,
  return = "plot"
)
```

---

email\_summary

*Email Summary*

---

**Description**

Provides an overview analysis of weekly email hours. Returns a bar plot showing average weekly email hours by default. Additional options available to return a summary table.

**Usage**

```
email_summary(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

```
email_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_trend\(\)](#)

**Examples**

```
# Return a ggplot bar chart
email_summary(pq_data, hrvar = "LevelDesignation")

# Return a summary table
email_summary(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

email_trend	<i>Email Hours Time Trend</i>
-------------	-------------------------------

---

### Description

Provides a week by week view of email time. By default returns a week by week heatmap, highlighting the points in time with most activity. Additional options available to return a summary table.

### Usage

```
email_trend(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

### Details

Uses the metric Email\_hours.

### Value

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Emails: [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#)

**Examples**

```
# Run plot
email_trend(pq_data)

# Run table
email_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

 export
 

---



---

*Export 'vivainsights' outputs to CSV, clipboard, or save as images*


---

**Description**

A general use function to export 'vivainsights' outputs to CSV, clipboard, or save as images. By default, `export()` copies a data frame to the clipboard. If the input is a 'ggplot' object, the default behaviour is to export a PNG.

**Usage**

```
export(
  x,
  method = "clipboard",
  path = "insights export",
  timestamp = TRUE,
  width = 12,
  height = 9
)
```

**Arguments**

<code>x</code>	Data frame or 'ggplot' object to be passed through.
<code>method</code>	Character string specifying the method of export. Valid inputs include: <ul style="list-style-type: none"> <li>• "clipboard" (default if input is data frame)</li> <li>• "csv"</li> <li>• "png" (default if input is 'ggplot' object)</li> <li>• "svg"</li> <li>• "jpeg"</li> <li>• "pdf"</li> </ul>
<code>path</code>	If exporting a file, enter the path and the desired file name, <i>excluding the file extension</i> . For example, "Analysis/SQ Overview".
<code>timestamp</code>	Logical vector specifying whether to include a timestamp in the file name. Defaults to TRUE.
<code>width</code>	Width of the plot
<code>height</code>	Height of the plot

**Value**

A different output is returned depending on the value passed to the method argument:

- "clipboard": no return - data frame is saved to clipboard.
- "csv": CSV file containing data frame is saved to specified path.
- "png": PNG file containing 'ggplot' object is saved to specified path.
- "svg": SVG file containing 'ggplot' object is saved to specified path.
- "jpeg": JPEG file containing 'ggplot' object is saved to specified path.
- "pdf": PDF file containing 'ggplot' object is saved to specified path.

**Author(s)**

Martin Chan [martin.chan@microsoft.com](mailto:martin.chan@microsoft.com)

**See Also**

Other Import and Export: [copy\\_df\(\)](#), [create\\_dt\(\)](#), [import\\_query\(\)](#)

---

external\_dist

*Distribution of External Collaboration Hours as a 100% stacked bar*

---

**Description**

Analyze the distribution of External Collaboration Hours. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

**Usage**

```
external_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(5, 10, 15)
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings:

- "plot"
- "table"

See Value for more information.

cut A numeric vector of length three to specify the breaks for the distribution, e.g. `c(10, 15, 20)`

## Details

Uses the metric `External_collaboration_hours`. See `create_dist()` for applying the same analysis to a different metric.

## Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

## See Also

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other External Collaboration: `external_fizz()`, `external_line()`, `external_sum()`

## Examples

```
# Return plot
external_dist(pq_data, hrvar = "Organization")

# Return summary table
external_dist(pq_data, hrvar = "Organization", return = "table")

# Return result with a custom specified breaks
external_dist(pq_data, hrvar = "LevelDesignation", cut = c(2, 4, 6))
```

---

external_fizz	<i>Distribution of External Collaboration Hours (Fizzy Drink plot)</i>
---------------	--

---

### Description

Analyze weekly External Collaboration hours distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

### Usage

```
external_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Details

Uses the metric `Collaboration_hours_external`. See `create_fizz()` for applying the same analysis to a different metric.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#),

email\_trend(), external\_dist(), external\_line(), external\_rank(), external\_sum(), hr\_trend(), hrvar\_count(), hrvar\_trend(), keymetrics\_scan(), meeting\_dist(), meeting\_fizz(), meeting\_line(), meeting\_rank(), meeting\_summary(), meeting\_trend(), one2one\_dist(), one2one\_fizz(), one2one\_freq(), one2one\_line(), one2one\_rank(), one2one\_sum(), one2one\_trend()

Other External Collaboration: external\_dist(), external\_line(), external\_sum()

## Examples

```
# Return plot
external_fizz(pq_data, hrvar = "LevelDesignation", return = "plot")

# Return summary table
external_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

external\_line

*External Collaboration Hours Time Trend - Line Chart*

---

## Description

Provides a week by week view of External collaboration time, visualized as line chart. By default returns a separate panel per value in the HR attribute. Additional options available to return a summary table.

## Usage

```
external_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

## Details

Uses the metric Collaboration\_hours\_external.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

[create\\_line\(\)](#) for applying the same analysis to a different metric.

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other External Collaboration: [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_sum\(\)](#)

**Examples**

```
# Return a line plot
external_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
external_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

external_rank	<i>Rank groups with high External Collaboration Hours</i>
---------------	---

---

**Description**

This function scans a Standard Person Query for groups with high levels of External Collaboration. Returns a plot by default, with an option to return a table with all groups (across multiple HR attributes) ranked by hours of External Collaboration.

**Usage**

```
external_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
```

```

    plot_mode = 1,
    return = "plot"
  )

```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> <li>• "table"</li> </ul> See Value for more information.

### Details

Uses the metric `Collaboration_hours_external`. See `create_rank()` for applying the same analysis to a different metric.

### Value

When 'table' is passed in return, a summary table is returned as a data frame.

### See Also

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other After-hours Collaboration: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`

**Examples**

```
# Return rank table
external_rank(data = pq_data, return = "table")

# Return plot
external_rank(data = pq_data, return = "plot")
```

---

external\_sum

*External Collaboration Summary*


---

**Description**

Provides an overview analysis of 'External Collaboration'. Returns a stacked bar plot of internal and external collaboration. Additional options available to return a summary table.

**Usage**

```
external_sum(
  data,
  hrvar = "Organization",
  mingroup = 5,
  stack_colours = c("#1d327e", "#1d7e6a"),
  return = "plot"
)

external_summary(
  data,
  hrvar = "Organization",
  mingroup = 5,
  stack_colours = c("#1d327e", "#1d7e6a"),
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
stack_colours	A character vector to specify the colour codes for the stacked bar charts.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

**Value**

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other External Collaboration: [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#)

**Examples**

```
# Return a plot
external_sum(pq_data, hrvar = "LevelDesignation")

# Return summary table
external_sum(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

extract\_date\_range      *Extract date period*

---

**Description**

Return a data frame with the start and end date of the query data by default. There are options to return a descriptive string, which is used in the caption of plots in this package.

**Usage**

```
extract_date_range(data, return = "table")
```

**Arguments**

data	Data frame containing a query to pass through. The data frame must contain a Date column. Accepts a Person query or a Meeting query.
return	String specifying what output to return. Returns a table by default ("table"), but allows returning a descriptive string ("text").

**Value**

A different output is returned depending on the value passed to the return argument:

- "table": data frame. A summary table containing the start and end date for the dataset.
- "text": string. Contains a descriptive string on the start and end date for the dataset.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

---

 extract\_hr

---

*Extract HR attribute variables*


---

**Description**

This function uses a combination of variable class, number of unique values, and regular expression matching to extract HR / organisational attributes from a data frame.

**Usage**

```
extract_hr(data, max_unique = 50, exclude_constants = TRUE, return = "names")
```

**Arguments**

data	A data frame to be passed through.
max_unique	A numeric value representing the maximum number of unique values to accept for an HR attribute. Defaults to 50.
exclude_constants	Logical value to specify whether single-value HR attributes are to be excluded. Defaults to TRUE.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "names"</li> <li>• "vars"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "names": character vector identifying all the names of HR variables present in the data.
- "vars": data frame containing all the columns of HR variables present in the data.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

Other Data Validation: [check\\_query\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
pq_data %>% extract_hr(return = "names")
```

```
pq_data %>% extract_hr(return = "vars")
```

---

flag_ch_ratio	<i>Flag unusual high collaboration hours to after-hours collaboration hours ratio</i>
---------------	---

---

**Description**

This function flags persons who have an unusual ratio of collaboration hours to after-hours collaboration hours. Returns a character string by default.

**Usage**

```
flag_ch_ratio(data, threshold = c(1, 30), return = "message")
```

**Arguments**

data	A data frame containing a Person Query.
threshold	Numeric value specifying the threshold for flagging. Defaults to 30.
return	String to specify what to return. Options include: <ul style="list-style-type: none"> <li>• "message"</li> <li>• "text"</li> <li>• "data"</li> </ul>

**Value**

A different output is returned depending on the value passed to the return argument:

- "message": message in the console containing diagnostic summary
- "text": string containing diagnostic summary
- "data": data frame. Person-level data with flags on unusually high or low ratios

**Metrics used**

The metric Collaboration\_hours is used in the calculations. Please ensure that your query contains a metric with the exact same name.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
flag_ch_ratio(pq_data)
```

```
data.frame(PersonId = c("Alice", "Bob"),
           Collaboration_hours = c(30, 0.5),
           After_hours_collaboration_hours = c(0.5, 30)) %>%
  flag_ch_ratio()
```

---

flag\_em\_ratio

*Flag Persons with unusually high Email Hours to Emails Sent ratio*

---

**Description**

This function flags persons who have an unusual ratio of email hours to emails sent. If the ratio between Email Hours and Emails Sent is greater than the threshold, then observations tied to a PersonId is flagged as unusual.

**Usage**

```
flag_em_ratio(data, threshold = 1, return = "text")
```

**Arguments**

data	A data frame containing a Person Query.
threshold	Numeric value specifying the threshold for flagging. Defaults to 1.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "text"</li> <li>• "data"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "text": string. A diagnostic message.
- "data": data frame. Person-level data with those flagged with unusual ratios.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
flag_em_ratio(pq_data)
```

---

flag\_extreme

*Warn for extreme values by checking against a threshold*

---

**Description**

This is used as part of data validation to check if there are extreme values in the dataset.

**Usage**

```
flag_extreme(
  data,
  metric,
  person = TRUE,
  threshold,
  mode = "above",
  return = "message"
)
```

**Arguments**

- |           |  |
|-----------|--|
| data      | A Standard Person Query dataset in the form of a data frame.   |
| metric    | A character string specifying the metric to test.  |
| person    | A logical value to specify whether to calculate person-averages. Defaults to TRUE (person-averages calculated).  |
| threshold | Numeric value specifying the threshold for flagging.   |
| mode      | String determining mode to use for identifying extreme values. <ul style="list-style-type: none"> <li>• "above": checks whether value is great than the threshold (default)</li> </ul> |

- "equal": checks whether value is equal to the threshold
- "below": checks whether value is below the threshold

return      String specifying what to return. This must be one of the following strings:

- "text"
- "message"
- "table"

See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "text": string. A diagnostic message.
- "message": message on console. A diagnostic message.
- "table": data frame. A person-level table with PersonId and the extreme values of the selected metric.

### See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

### Examples

```
# The threshold values are intentionally set low to trigger messages.
flag_extreme(pq_data, "Email_hours", threshold = 15)

# Return a summary table
flag_extreme(pq_data, "Email_hours", threshold = 15, return = "table")

# Person-week level
flag_extreme(pq_data, "Email_hours", person = FALSE, threshold = 15)

# Check for values equal to threshold
flag_extreme(pq_data, "Email_hours", person = TRUE, mode = "equal", threshold = 0)

# Check for values below threshold
flag_extreme(pq_data, "Email_hours", person = TRUE, mode = "below", threshold = 5)
```

---

flag_outlooktime	<i>Flag unusual outlook time settings for work day start and end time</i>
------------------	---

---

### Description

This function flags unusual outlook calendar settings for start and end time of work day.

### Usage

```
flag_outlooktime(data, threshold = c(4, 15), return = "message")
```

### Arguments

data	A data frame containing a Person Query.
threshold	A numeric vector of length two, specifying the hour threshold for flagging. Defaults to c(4, 15).
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "text" (default)</li> <li>• "message"</li> <li>• "data"</li> </ul>

### Value

A different output is returned depending on the value passed to the return argument:

- "text": string. A diagnostic message.
- "message": message on console. A diagnostic message.
- "data": data frame. Data where flag is present.

See Value for more information.

### See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

### Examples

```
# Demo with `pq_data` example where Outlook Start and End times are imputed
spq_df <- pq_data

spq_df$WorkingStartTimeSetInOutlook <- "6:30"

spq_df$WorkingEndTimeSetInOutlook <- "23:30"
```

```
# Return a message
flag_outlooktime(spq_df, threshold = c(5, 13))

# Return data
flag_outlooktime(spq_df, threshold = c(5, 13), return = "data")
```

---

g2g\_data

*Sample Group-to-Group dataset*

---

## Description

A demo dataset representing a Group-to-Group Query. The grouping organizational attribute used here is `Organization`, where the variable have been prefixed with `PrimaryCollaborator_` and `SecondaryCollaborator_` to represent the direction of collaboration.

## Usage

```
g2g_data
```

## Format

A data frame with 150 rows and 11 variables:

**PrimaryCollaborator\_Organization**  
**PrimaryCollaborator\_GroupSize**  
**SecondaryCollaborator\_Organization**  
**SecondaryCollaborator\_GroupSize**  
**MetricDate**  
**Percent\_Group\_collaboration\_time\_invested**  
**Group\_collaboration\_time\_invested**  
**Group\_email\_sent\_count**  
**Group\_email\_time\_invested**  
**Group\_meeting\_count**  
**Group\_meeting\_time\_invested ...**

## Value

data frame.

## Source

<https://analysis.insights.viva.office.com/analyst/analysis/>

**See Also**

Other Data: [mt\\_data](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#), [pq\\_data](#)

Other Network: [network\\_g2g\(\)](#), [network\\_p2p\(\)](#), [network\\_summary\(\)](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#)

---

generate_report	<i>Generate HTML report with list inputs</i>
-----------------	--

---

**Description**

This is a support function using a list-pmap workflow to create a HTML document, using RMarkdown as the engine.

**Usage**

```
generate_report(
  title = "My minimal HTML generator",
  filename = "minimal_html",
  outputs = output_list,
  titles,
  subheaders,
  echos,
  levels,
  theme = "united",
  preamble = ""
)
```

**Arguments**

title	Character string to specify the title of the chunk.
filename	File name to be used in the exported HTML.
outputs	A list of outputs to be added to the HTML report. Note that outputs, titles, echos, and levels must have the same length
titles	A list/vector of character strings to specify the title of the chunks.
subheaders	A list/vector of character strings to specify the subheaders for each chunk.
echos	A list/vector of logical values to specify whether to display code.
levels	A list/vector of numeric value to specify the header level of the chunk.
theme	Character vector to specify theme to be used for the report. E.g. "united", "default".
preamble	A preamble to appear at the beginning of the report, passed as a text string.

**Value**

An HTML report with the same file name as specified in the arguments is generated in the working directory. No outputs are directly returned by the function.

### Creating a custom report

Below is an example on how to set up a custom report.

The first step is to define the content that will go into a report and assign the outputs to a list.

```
# Step 1: Define Content
output_list <-
  list(pq_data %>% workloads_summary(return = "plot"),
       pq_data %>% workloads_summary(return = "table")) %>%
  purrr::map_if(is.data.frame, create_dt)
```

The next step is to add a list of titles for each of the objects on the list:

```
# Step 2: Add Corresponding Titles
title_list <- c("Workloads Summary - Plot", "Workloads Summary - Table")
n_title <- length(title_list)
```

The final step is to run `generate_report()`. This can all be wrapped within a function such that the function can be used to generate a HTML report.

```
# Step 3: Generate Report
generate_report(title = "My First Report",
               filename = "My First Report",
               outputs = output_list,
               titles = title_list,
               subheaders = rep("", n_title),
               echos = rep(FALSE, n_title))
```

### Author(s)

Martin Chan [martin.chan@microsoft.com](mailto:martin.chan@microsoft.com)

### See Also

Other Reports: [IV\\_report\(\)](#), [meeting\\_tm\\_report\(\)](#), [read\\_preamble\(\)](#), [validation\\_report\(\)](#)

---

generate\_report2

*Generate HTML report based on existing RMarkdown documents*

---

### Description

This is a support function that accepts parameters and creates a HTML document based on an RMarkdown template. This is an alternative to `generate_report()` which instead creates an RMarkdown document from scratch using individual code chunks.

**Usage**

```
generate_report2(
  output_format = rmarkdown::html_document(toc = TRUE, toc_depth = 6, theme = "cosmo"),
  output_file = "report.html",
  output_dir = getwd(),
  report_title = "Report",
  rmd_dir = system.file("rmd_template/minimal.rmd", package = "vivainsights"),
  ...
)
```

**Arguments**

output_format	output format in <code>rmarkdown::render()</code> . Default is <code>rmarkdown::html_document(toc = TRUE, toc_depth = 6, theme = "cosmo")</code> .
output_file	output file name in <code>rmarkdown::render()</code> . Default is <code>"report.html"</code> .
output_dir	output directory for report in <code>rmarkdown::render()</code> . Default is user's current directory.
report_title	report title. Default is <code>"Report"</code> .
rmd_dir	string specifying the path to the directory containing the RMarkdown template files.
...	other arguments to be passed to <code>params</code> . For instance, pass <code>hrvar</code> if the RMarkdown document requires a <code>'hrvar'</code> parameter.

**Note**

The implementation of this function was inspired by the 'DataExplorer' package by boxuancui, with credits due to the original author.

---

heat_colours	<i>Generate a vector of n contiguous colours, as a red-yellow-green palette.</i>
--------------	--

---

**Description**

Takes a numeric value `n` and returns a character vector of colour HEX codes corresponding to the heat map palette.

**Usage**

```
heat_colours(n, alpha, rev = FALSE)

heat_colors(n, alpha, rev = FALSE)
```

**Arguments**

n	the number of colors ( $\geq 1$ ) to be in the palette.
alpha	an alpha-transparency level in the range of 0 to 1 (0 means transparent and 1 means opaque)
rev	logical indicating whether the ordering of the colors should be reversed.

**Value**

A character vector containing the HEX codes and the same length as n is returned.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

**Examples**

```
barplot(rep(10, 50), col = heat_colours(n = 50), border = NA)

barplot(rep(10, 50), col = heat_colours(n = 50, alpha = 0.5, rev = TRUE),
border = NA)
```

---

hrvar\_count

---

*Create a count of distinct people in a specified HR variable*


---

**Description**

This function enables you to create a count of the distinct people by the specified HR attribute. The default behaviour is to return a bar chart as typically seen in 'Analysis Scope'.

**Usage**

```
hrvar_count(data, hrvar = "Organization", return = "plot")

analysis_scope(data, hrvar = "Organization", return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	HR Variable by which to split metrics, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation". If a vector with more than one value is provided, the HR attributes are automatically concatenated.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object containing a bar plot.
- "table": data frame containing a count table.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
# Return a bar plot
hrvar_count(pq_data, hrvar = "LevelDesignation")

# Return a summary table
hrvar_count(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

hrvar_count_all	<i>Create count of distinct fields and percentage of employees with missing values for all HR variables</i>
-----------------	---

---

**Description****[Experimental]**

This function enables you to create a summary table to validate organizational data. This table will provide a summary of the data found in the Viva Insights *Data sources* page. This function will return a summary table with the count of distinct fields per HR attribute and the percentage of employees with missing values for that attribute. See [hrvar\\_count\(\)](#) function for more detail on the specific HR attribute of interest.

**Usage**

```
hrvar_count_all(  
  data,  
  n_var = 50,  
  return = "message",  
  threshold = 100,  
  maxna = 20  
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
n_var	number of HR variables to include in report as rows. Default is set to 50 HR variables.
return	String to specify what to return
threshold	The max number of unique values allowed for any attribute. Default is 100.
maxna	The max percentage of NAs allowable for any column. Default is 20.

**Value**

Returns an error message by default, where 'text' is passed in return.

- 'table': data frame. A summary table listing the number of distinct fields and percentage of missing values for the specified number of HR attributes will be returned.
- 'message': outputs a message indicating which values are beyond the specified thresholds.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
# Return a summary table of all HR attributes  
hrvar_count_all(pq_data, return = "table")
```

---

hrvar_trend	<i>Track count of distinct people over time in a specified HR variable</i>
-------------	--

---

### Description

This function provides a week by week view of the count of the distinct people by the specified HR attribute. The default behaviour is to return a week by week heatmap bar plot.

### Usage

```
hrvar_trend(data, hrvar = "Organization", return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	HR Variable by which to split metrics, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation". If a vector with more than one value is provided, the HR attributes are automatically concatenated.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object containing a bar plot.
- "table": data frame containing a count table.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#),

```
identify_privacythreshold(), identify_shifts(), identify_tenure(), track_HR_change(),
validation_report()
```

### Examples

```
# Return a bar plot
hrvar_trend(pq_data, hrvar = "LevelDesignation")

# Return a summary table
hrvar_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

hr_trend	<i>Employee count over time</i>
----------	---------------------------------

---

### Description

Returns a line chart showing the change in employee count over time. Part of a data validation process to check for unusual license growth / declines over time.

### Usage

```
hr_trend(data, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": ggplot object. A line plot showing employee count over time.
- "table": data frame containing a summary table.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#),

```
create_trend(), email_dist(), email_fizz(), email_line(), email_rank(), email_summary(),
email_trend(), external_dist(), external_fizz(), external_line(), external_rank(),
external_sum(), hrvar_count(), hrvar_trend(), keymetrics_scan(), meeting_dist(), meeting_fizz(),
meeting_line(), meeting_rank(), meeting_summary(), meeting_trend(), one2one_dist(),
one2one_fizz(), one2one_freq(), one2one_line(), one2one_rank(), one2one_sum(), one2one_trend()
```

Other Data Validation: `check_query()`, `extract_hr()`, `flag_ch_ratio()`, `flag_em_ratio()`, `flag_extreme()`, `flag_outlooktime()`, `hrvar_count()`, `hrvar_count_all()`, `hrvar_trend()`, `identify_churn()`, `identify_holidayweeks()`, `identify_inactiveweeks()`, `identify_nkw()`, `identify_outlier()`, `identify_privacythreshold()`, `identify_shifts()`, `identify_tenure()`, `track_HR_change()`, `validation_report()`

## Examples

```
# Return plot
hr_trend(pq_data)

# Return summary table
hr_trend(pq_data, return = "table")
```

---

identify_churn	<i>Identify employees who have churned from the dataset</i>
----------------	---

---

## Description

This function identifies and counts the number of employees who have churned from the dataset by measuring whether an employee who is present in the first  $n_1$  weeks of the data is present in the last  $n_2$  weeks of the data.

## Usage

```
identify_churn(data, n1 = 6, n2 = 6, return = "message", flip = FALSE)
```

## Arguments

data	A Person Query as a data frame. Must contain a PersonId.
n1	A numeric value specifying the number of weeks at the beginning of the period that defines the measured employee set. Defaults to 6.
n2	A numeric value specifying the number of weeks at the end of the period to calculate whether employees have churned from the data. Defaults to 6.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"message" (default)</li> <li>"text"</li> <li>"data"</li> </ul>

See Value for more information.

`flip` Logical, defaults to FALSE. This determines whether to reverse the logic of identifying the non-overlapping set. If set to TRUE, this effectively identifies new-joiners, or those who were not present in the first `n` weeks of the data but were present in the final `n` weeks.

## Details

An additional use case of this function is the ability to identify "new-joiners" by using the argument `flip`.

If an employee is present in the first `n` weeks of the data but not present in the last `n` weeks of the data, the function considers the employee as churned. As the measurement period is defined by the number of weeks from the start and the end of the passed data frame, you may consider filtering the dates accordingly before running this function.

Another assumption that is in place is that any employee whose `PersonId` is not available in the data has churned. Note that there may be other reasons why an employee's `PersonId` may not be present, e.g. maternity/paternity leave, Viva Insights license has been removed, shift to a low-collaboration role (to the extent that he/she becomes inactive).

## Value

A different output is returned depending on the value passed to the return argument:

- `"message"`: Message on console. A diagnostic message.
- `"text"`: String. A diagnostic message.
- `"data"`: Character vector containing the the `PersonId` of employees who have been identified as churned.

## See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

## Examples

```
pq_data %>% identify_churn(n1 = 3, n2 = 3, return = "message")
```

---

identify\_datefreq      *Identify date frequency based on a series of dates*

---

## Description

### [Experimental]

Takes a vector of dates and identify whether the frequency is 'daily', 'weekly', or 'monthly'. The primary use case for this function is to provide an accurate description of the query type used and for raising errors should a wrong date grouping be used in the data input.

## Usage

```
identify_datefreq(x)
```

## Arguments

x                      Vector containing a series of dates.

## Details

Date frequency detection works as follows:

- If at least three days of the week are present (e.g., Monday, Wednesday, Thursday) in the series, then the series is classified as 'daily'
- If the total number of months in the series is equal to the length, then the series is classified as 'monthly'
- If the total number of sundays in the series is equal to the length of the series, then the series is classified as 'weekly'

## Value

String describing the detected date frequency, i.e.:

- 'daily'
- 'weekly'
- 'monthly'

## Limitations

One of the assumptions made behind the classification is that weeks are denoted with Sundays, hence the count of sundays to measure the number of weeks. In this case, weeks where a Sunday is missing would result in an 'unable to classify' error.

Another assumption made is that dates are evenly distributed, i.e. that the gap between dates are equal. If dates are unevenly distributed, e.g. only two days of the week are available for a given week, then the algorithm will fail to identify the frequency as 'daily'.

**Examples**

```
start_date <- as.Date("2022/06/26")
end_date <- as.Date("2022/11/27")

# Daily
day_seq <-
  seq.Date(
    from = start_date,
    to = end_date,
    by = "day"
  )

identify_datefreq(day_seq)

# Weekly
week_seq <-
  seq.Date(
    from = start_date,
    to = end_date,
    by = "week"
  )

identify_datefreq(week_seq)

# Monthly
month_seq <-
  seq.Date(
    from = start_date,
    to = end_date,
    by = "month"
  )
identify_datefreq(month_seq)
```

---

identify\_holidayweeks *Identify Holiday Weeks based on outliers*

---

**Description**

This function scans a standard query output for weeks where collaboration hours is far outside the mean. Returns a list of weeks that appear to be holiday weeks and optionally an edited dataframe with outliers removed. By default, missing values are excluded.

As best practice, run this function prior to any analysis to remove atypical collaboration weeks from your dataset.

**Usage**

```
identify_holidayweeks(data, sd = 1, return = "message")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
sd	The standard deviation below the mean for collaboration hours that should define an outlier week. Enter a positive number. Default is 1 standard deviation.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "message" (default)</li> <li>• "data"</li> <li>• "data_cleaned"</li> <li>• "data_dirty"</li> <li>• "plot"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "message": message on console. a message is printed identifying holiday weeks.
- "data": data frame. A dataset with outlier weeks flagged in a new column is returned as a dataframe.
- "data\_cleaned": data frame. A dataset with outlier weeks removed is returned.
- "data\_dirty": data frame. A dataset with only outlier weeks is returned.
- "plot": ggplot object. A line plot of Collaboration Hours with holiday weeks highlighted.

**Metrics used**

The metric Collaboration\_hours is used in the calculations. Please ensure that your query contains a metric with the exact same name.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
# Return a message by default
identify_holidayweeks(pq_data)

# Return plot
identify_holidayweeks(pq_data, return = "plot")
```

---

`identify_inactiveweeks`*Identify Inactive Weeks*

---

### Description

This function scans a standard query output for weeks where collaboration hours is far outside the mean for any individual person in the dataset. Returns a list of weeks that appear to be inactive weeks and optionally an edited dataframe with outliers removed.

As best practice, run this function prior to any analysis to remove atypical collaboration weeks from your dataset.

### Usage

```
identify_inactiveweeks(data, sd = 2, return = "text")
```

### Arguments

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>sd</code>	The standard deviation below the mean for collaboration hours that should define an outlier week. Enter a positive number. Default is 1 standard deviation.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"><li>• "text"</li><li>• "data_cleaned"</li><li>• "data_dirty"</li></ul> See Value for more information.

### Value

Returns an error message by default, where 'text' is returned. When 'data\_cleaned' is passed, a dataset with outlier weeks removed is returned as a dataframe. When 'data\_dirty' is passed, a dataset with outlier weeks is returned as a dataframe.

### See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

---

identify_nkw	<i>Identify Non-Knowledge workers in a Person Query using Collaboration Hours</i>
--------------	---

---

## Description

This function scans a standard query output to identify employees with consistently low collaboration signals. Returns the % of non-knowledge workers identified by Organization, and optionally an edited data frame with non-knowledge workers removed, or the full data frame with the kw/nkw flag added.

## Usage

```
identify_nkw(data, collab_threshold = 5, return = "data_summary")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
collab_threshold	Positive numeric value representing the collaboration hours threshold that should be exceeded as an average for the entire analysis period for the employee to be categorized as a knowledge worker ("kw"). Default is set to 5 collaboration hours. Any versions after v1.4.3, this uses a "greater than or equal to" logic ( $\geq$ ), in which case persons with exactly 5 collaboration hours will pass.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "text"</li> <li>• "data_with_flag"</li> <li>• "data_clean"</li> <li>• "data_summary"</li> </ul> See Value for more information.

## Value

A different output is returned depending on the value passed to the return argument:

- "text": string. Returns a diagnostic message.
- "data\_with\_flag": data frame. Original input data with an additional column containing the kw/nkw flag.
- "data\_clean": data frame. Data frame with non-knowledge workers excluded.
- "data\_summary": data frame. A summary table by organization listing the number and % of non-knowledge workers.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

---

identify_outlier	<i>Identify metric outliers over a date interval</i>
------------------	--

---

**Description**

This function takes in a selected metric and uses z-score (number of standard deviations) to identify outliers across time. There are applications in this for identifying weeks with abnormally low collaboration activity, e.g. holidays. Time as a grouping variable can be overridden with the `group_var` argument.

**Usage**

```
identify_outlier(  
  data,  
  group_var = "MetricDate",  
  metric = "Collaboration_hours"  
)
```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>group_var</code>	A string with the name of the grouping variable. Defaults to Date.
<code>metric</code>	Character string containing the name of the metric, e.g. "Collaboration_hours"

**Value**

Returns a data frame with MetricDate (if grouping variable is not set), the metric, and the corresponding z-score.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
identify_outlier(pq_data, metric = "Collaboration_hours")
```

---

`identify_privacythreshold`*Identify groups under privacy threshold*

---

### Description

This function scans a standard query output for groups with of employees under the privacy threshold. The method consists in reviewing each individual HR attribute, and count the distinct people within each group.

### Usage

```
identify_privacythreshold(  
  data,  
  hrvar = extract_hr(data),  
  mingroup = 5,  
  return = "table"  
)
```

### Arguments

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>hrvar</code>	A list of HR Variables to consider in the scan. Defaults to all HR attributes identified.
<code>mingroup</code>	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"><li>• "table"</li><li>• "text"</li></ul> See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "table": data frame. A summary table of groups that fall below the privacy threshold.
- "text": string. A diagnostic message.

Returns a ggplot object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

### See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
# Return a summary table
pq_data %>% identify_privacythreshold(return = "table")

# Return a diagnostic message
pq_data %>% identify_privacythreshold(return = "text")
```

---

identify_shifts	<i>Identify shifts based on outlook time settings for work day start and end time</i>
-----------------	---

---

**Description**

This function uses outlook calendar settings for start and end time of work day to identify work shifts. The relevant variables are WorkingStartTimeSetInOutlook and WorkingEndTimeSetInOutlook.

**Usage**

```
identify_shifts(data, return = "plot")
```

**Arguments**

data	A data frame containing data from the Hourly Collaboration query.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> <li>• "data"</li> </ul> See Value for more information.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": ggplot object. A bar plot for the weekly count of shifts.
- "table": data frame. A summary table for the count of shifts.
- "data": data frame. Input data appended with the Shifts columns.

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
# Demo with `pq_data` example where Outlook Start and End times are imputed
spq_df <- pq_data

spq_df$WorkingStartTimeSetInOutlook <- "6:30"

spq_df$WorkingEndTimeSetInOutlook <- "23:30"

# Return plot
spq_df %>% identify_shifts()

# Return summary table
spq_df %>% identify_shifts(return = "table")
```

---

identify_tenure	<i>Tenure calculation based on different input dates, returns data summary table or histogram</i>
-----------------	---

---

**Description**

This function calculates employee tenure based on different input dates. `identify_tenure` uses the latest Date available if user selects "MetricDate", but also have flexibility to select a specific date, e.g. "1/1/2020".

**Usage**

```
identify_tenure(
  data,
  end_date = "MetricDate",
  beg_date = "HireDate",
  maxten = 40,
  return = "message"
)
```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>end_date</code>	A string specifying the name of the date variable representing the latest date. Defaults to "MetricDate".
<code>beg_date</code>	A string specifying the name of the date variable representing the hire date. Defaults to "HireDate".
<code>maxten</code>	A numeric value representing the maximum tenure. If the tenure exceeds this threshold, it would be accounted for in the flag message.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"message"</li> </ul>

- "text"
- "plot"
- "data\_cleaned"
- "data\_dirty"
- "data"

See Value for more information.

## Value

A different output is returned depending on the value passed to the return argument:

- "message": message on console with a diagnostic message.
- "text": string containing a diagnostic message.
- "plot": 'ggplot' object. A line plot showing tenure.
- "data\_cleaned": data frame filtered only by rows with tenure values lying within the threshold.
- "data\_dirty": data frame filtered only by rows with tenure values lying outside the threshold.
- "data": data frame with the PersonId and a calculated variable called TenureYear is returned.

## See Also

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [track\\_HR\\_change\(\)](#), [validation\\_report\(\)](#)

## Examples

```
library(dplyr)
# Add HireDate to `pq_data`
pq_data2 <-
  pq_data %>%
  mutate(HireDate = as.Date("1/1/2015", format = "%m/%d/%Y"))

identify_tenure(pq_data2)
```

import\_query

*Import a query from Viva Insights Analyst Experience***Description**

Import a Viva Insights Query from a .csv file, with variable classifications optimised for other functions in the package.

**Usage**

```
import_query(
  x,
  pid = NULL,
  dateid = NULL,
  date_format = "%m/%d/%Y",
  convert_date = TRUE,
  encoding = "UTF-8"
)
```

**Arguments**

x	String containing the path to the Viva Insights query to be imported. The input file must be a .csv file, and the file extension must be explicitly entered, e.g. "/files/standard query.csv"
pid	String specifying the unique person or individual identifier variable. import_query renames this to PersonId so that this is compatible with other functions in the package. Defaults to NULL, where no action is taken.
dateid	String specifying the date variable. import_query renames this to MetricDate so that this is compatible with other functions in the package. Defaults to NULL, where no action is taken.
date_format	String specifying the date format for converting any variable that may be a date to a Date variable. Defaults to "%m/%d/%Y".
convert_date	Logical. Defaults to TRUE. When set to TRUE, any variable that matches true with is_date_format() gets converted to a Date variable. When set to FALSE, this step is skipped.
encoding	String to specify encoding to be used within data.table::fread(). See data.table::fread() documentation for more information. Defaults to 'UTF-8'.

**Details**

import\_query() uses data.table::fread() to import .csv files for speed, and by default stringsAsFactors is set to FALSE. A data frame is returned by the function (not a data.table). Column names are automatically cleaned, replacing spaces and special characters with underscores.

**Value**

A tibble is returned.

**See Also**

Other Import and Export: [copy\\_df\(\)](#), [create\\_dt\(\)](#), [export\(\)](#)

---

is\_date\_format      *Identify whether string is a date format*

---

**Description**

This function uses regular expression to determine whether a string is of the format "mdy", separated by "-", "/", or ".", returning a logical vector.

**Usage**

```
is_date_format(string)
```

**Arguments**

string      Character string to test whether is a date format.

**Value**

logical value indicating whether the string is a date format.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

**Examples**

```
is_date_format("1/5/2020")
```

---

IV\_report      *Generate a Information Value HTML Report*

---

**Description**

The function generates an interactive HTML report using Standard Person Query data as an input. The report contains a full Information Value analysis, a data exploration technique that helps determine which columns in a data set have predictive power or influence on the value of a specified dependent variable.

**Usage**

```
IV_report(
  data,
  predictors = NULL,
  outcome,
  bins = 5,
  max_var = 9,
  path = "IV report",
  timestamp = TRUE
)
```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>predictors</code>	A character vector specifying the columns to be used as predictors. Defaults to <code>NULL</code> , where all numeric vectors in the data will be used as predictors.
<code>outcome</code>	A string specifying a binary variable, i.e. can only contain the values 1 or 0.
<code>bins</code>	Number of bins to use in <code>Information::create_infotables()</code> , defaults to 10.
<code>max_var</code>	Numeric value to represent the maximum number of variables to show on plots.
<code>path</code>	Pass the file path and the desired file name, <i>excluding the file extension</i> . For example, "IV report".
<code>timestamp</code>	Logical vector specifying whether to include a timestamp in the file name. Defaults to <code>TRUE</code> .

**Value**

An HTML report with the same file name as specified in the arguments is generated in the working directory. No outputs are directly returned by the function.

**Creating a report**

Below is an example on how to run the report.

```
library(dplyr)

pq_data %>%
  mutate(CH_binary = ifelse(Collaboration_hours > 12, 1, 0)) %>% # Simulate binary variable
  IV_report(outcome = "CH_binary",
            predictors = c("Email_hours", "Meeting_hours"))
```

**See Also**

Other Reports: [generate\\_report\(\)](#), [meeting\\_tm\\_report\(\)](#), [read\\_preamble\(\)](#), [validation\\_report\(\)](#)

Other Variable Association: [create\\_IV\(\)](#)

Other Information Value: [create\\_IV\(\)](#)

---

jitter_metrics	<i>Jitter metrics in a data frame</i>
----------------	---------------------------------------

---

### Description

Convenience wrapper around `jitter()` to add a layer of anonymity to a query. This can be used in combination with `anonymise()` to produce a demo dataset from real data.

### Usage

```
jitter_metrics(data, cols = NULL, ...)
```

### Arguments

<code>data</code>	Data frame containing a query.
<code>cols</code>	Character vector containing the metrics to jitter. When set to <code>NULL</code> (default), all numeric columns in the data frame are jittered.
<code>...</code>	Additional arguments to pass to <code>jitter()</code> .

### Value

data frame where numeric columns specified by `cols` are jittered using the function `jitter()`.

### See Also

`anonymise`

### Examples

```
jittered <- jitter_metrics(pq_data, cols = "Collaboration_hours")

# compare jittered vs original results of top rows
head(
  data.frame(
    original = pq_data$Collaboration_hours,
    jittered = jittered$Collaboration_hours
  )
)
```

---

keymetrics\_scan      *Run a summary of Key Metrics from the Standard Person Query data*

---

### Description

Returns a heatmapped table by default, with options to return a table.

### Usage

```
keymetrics_scan(
  data,
  hrvar = "Organization",
  mingroup = 5,
  metrics = c("Workweek_span", "Collaboration_hours", "After_hours_collaboration_hours",
    "Meetings", "Meeting_hours", "After_hours_meeting_hours",
    "Low_quality_meeting_hours", "Meeting_hours_with_manager_1_on_1",
    "Meeting_hours_with_manager", "Emails_sent", "Email_hours",
    "After_hours_email_hours", "Generated_workload_email_hours", "Total_focus_hours",
    "Internal_network_size", "Networking_outside_organization", "External_network_size",
    "Networking_outside_company"),
  return = "plot",
  low = rgb2hex(7, 111, 161),
  mid = rgb2hex(241, 204, 158),
  high = rgb2hex(216, 24, 42),
  textsize = 2
)
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
metrics	A character vector containing the variable names to calculate averages of.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".
low	String specifying colour code to use for low-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
mid	String specifying colour code to use for mid-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
high	String specifying colour code to use for high-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
textsize	A numeric value specifying the text size to show in the plot.

**Value**

Returns a ggplot object by default, when 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

**Examples**

```
# Heatmap plot is returned by default
keymetrics_scan(pq_data)

# Heatmap plot with custom colours
keymetrics_scan(pq_data, low = "purple", high = "yellow")

# Return summary table
keymetrics_scan(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

`keymetrics_scan_asis` *Run a summary of Key Metrics without aggregation*

---

**Description**

Return a heatmapped table directly from the aggregated / summarised data. Unlike `keymetrics_scan()` which performs a person-level aggregation, there is no calculation for `keymetrics_scan_asis()` and the values are rendered as they are passed into the function.

**Usage**

```
keymetrics_scan_asis(
  data,
  row_var,
  col_var,
  group_var = col_var,
  value_var = "value",
  title = NULL,
```

```

  subtitle = NULL,
  caption = NULL,
  ylab = row_var,
  xlab = "Metrics",
  rounding = 1,
  low = rgb2hex(7, 111, 161),
  mid = rgb2hex(241, 204, 158),
  high = rgb2hex(216, 24, 42),
  textsize = 2
)

```

### Arguments

<code>data</code>	data frame containing data to plot. It is recommended to provide data in a 'long' table format where one grouping column forms the rows, a second column forms the columns, and a third numeric columns forms the
<code>row_var</code>	String containing name of the grouping variable that will form the rows of the heatmapped table.
<code>col_var</code>	String containing name of the grouping variable that will form the columns of the heatmapped table.
<code>group_var</code>	String containing name of the grouping variable by which heatmapping would apply. Defaults to <code>col_var</code> .
<code>value_var</code>	String containing name of the value variable that will form the values of the heatmapped table. Defaults to "value".
<code>title</code>	Title of the plot.
<code>subtitle</code>	Subtitle of the plot.
<code>caption</code>	Caption of the plot.
<code>ylab</code>	Y-axis label for the plot (group axis)
<code>xlab</code>	X-axis label of the plot (bar axis).
<code>rounding</code>	Numeric value to specify number of digits to show in data labels
<code>low</code>	String specifying colour code to use for low-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
<code>mid</code>	String specifying colour code to use for mid-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
<code>high</code>	String specifying colour code to use for high-value metrics. Arguments are passed directly to <code>ggplot2::scale_fill_gradient2()</code> .
<code>textsize</code>	A numeric value specifying the text size to show in the plot.

### Value

ggplot object for a heatmap table.

## Examples

```
library(dplyr)

# Compute summary table
out_df <-
  pq_data %>%
  group_by(Organization) %>%
  summarise(
    across(
      .cols = c(
        Email_hours,
        Collaboration_hours
      ),
      .fns = ~median(., na.rm = TRUE)
    ),
    .groups = "drop"
  ) %>%
  tidyr::pivot_longer(
    cols = c("Email_hours", "Collaboration_hours"),
    names_to = "metrics"
  )

keymetrics_scan_asis(
  data = out_df,
  col_var = "metrics",
  row_var = "Organization"
)

# Show data the other way round
keymetrics_scan_asis(
  data = out_df,
  col_var = "Organization",
  row_var = "metrics",
  group_var = "metrics"
)
```

---

maxmin

*Max-Min Scaling Function*

---

## Description

This function allows you to scale vectors or an entire data frame using the max-min scaling method. A numeric vector is always returned.

## Usage

```
maxmin(x)
```

## Arguments

x                      Pass a vector or the required columns of a data frame through this argument.

## Details

This is used within `keymetrics_scan()` to enable row-wise heatmapping. Originally implemented in <https://github.com/martinctc/surveytoolbox>.

## Value

Returns a numeric vector with the input rescaled.

## See Also

Other Support: `any_idate()`, `camel_clean()`, `check_inputs()`, `cut_hour()`, `extract_date_range()`, `extract_hr()`, `heat_colours()`, `is_date_format()`, `pairwise_count()`, `read_preamble()`, `rgb2hex()`, `totals_bind()`, `totals_col()`, `tstamp()`, `us_to_space()`, `wrap()`

## Examples

```
numbers <- c(15, 40, 10, 2)
maxmin(numbers)
```

---

meeting\_dist

*Distribution of Meeting Hours as a 100% stacked bar*

---

## Description

Analyze Meeting Hours distribution. Returns a stacked bar plot by default. Additional options available to return a table with distribution elements.

## Usage

```
meeting_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  cut = c(5, 10, 15)
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
cut	A numeric vector of length three to specify the breaks for the distribution, e.g. c(10, 15, 20)

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Meetings: [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_tm\\_report\(\)](#), [meeting\\_trend\(\)](#)

**Examples**

```
# Return plot
meeting_dist(pq_data, hrvar = "Organization")

# Return summary table
meeting_dist(pq_data, hrvar = "Organization", return = "table")

# Return result with a custom specified breaks
meeting_dist(pq_data, hrvar = "LevelDesignation", cut = c(4, 7, 9))
```

---

meeting\_fizz                      *Distribution of Meeting Hours (Fizzy Drink plot)*

---

### Description

Analyze weekly meeting hours distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

### Usage

```
meeting_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.

### Details

Uses the metric Meeting\_hours.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#),

```
external_sum(), hr_trend(), hrvar_count(), hrvar_trend(), keymetrics_scan(), meeting_dist(),
meeting_line(), meeting_rank(), meeting_summary(), meeting_trend(), one2one_dist(),
one2one_fizz(), one2one_freq(), one2one_line(), one2one_rank(), one2one_sum(), one2one_trend()
Other Meetings: meeting_dist(), meeting_line(), meeting_rank(), meeting_summary(), meeting_tm_report(),
meeting_trend()
```

## Examples

```
# Return plot
meeting_fizz(pq_data, hrvar = "Organization", return = "plot")

# Return summary table
meeting_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

meeting_line	<i>Meeting Time Trend - Line Chart</i>
--------------	--

---

## Description

Provides a week by week view of meeting time, visualised as line charts. By default returns a line chart for meeting hours, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

## Usage

```
meeting_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

## Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_rank()`, `meeting_summary()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other Meetings: `meeting_dist()`, `meeting_fizz()`, `meeting_rank()`, `meeting_summary()`, `meeting_tm_report()`, `meeting_trend()`

**Examples**

```
# Return a line plot
meeting_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
meeting_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

meeting\_rank

*Meeting Hours Ranking*

---

**Description**

This function scans a standard query output for groups with high levels of Weekly Meeting Collaboration. Returns a plot by default, with an option to return a table with a all of groups (across multiple HR attributes) ranked by hours of digital collaboration.

**Usage**

```
meeting_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> <li>• "table"</li> </ul> See Value for more information.

**Details**

Uses the metric Meeting\_hours. See create\_rank() for applying the same analysis to a different metric.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bubble plot where the x-axis represents the metric, the y-axis represents the HR attributes, and the size of the bubbles represent the size of the organizations. Note that there is no plot output if mode is set to "combine".
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Meetings: [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_tm\\_report\(\)](#), [meeting\\_trend\(\)](#)

## Examples

```
# Return rank table
meeting_rank(data = pq_data, return = "table")

# Return plot
meeting_rank(data = pq_data, return = "plot")
```

---

meeting_summary	<i>Meeting Summary</i>
-----------------	------------------------

---

## Description

Provides an overview analysis of weekly meeting hours. Returns a bar plot showing average weekly meeting hours by default. Additional options available to return a summary table.

## Usage

```
meeting_summary(data, hrvar = "Organization", mingroup = 5, return = "plot")

meeting_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

## Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: `afterhours_dist()`, `afterhours_fizz()`, `afterhours_line()`, `afterhours_rank()`, `afterhours_summary()`, `afterhours_trend()`, `collaboration_area()`, `collaboration_dist()`, `collaboration_fizz()`, `collaboration_line()`, `collaboration_rank()`, `collaboration_sum()`, `collaboration_trend()`, `create_bar()`, `create_bar_asis()`, `create_boxplot()`, `create_bubble()`, `create_dist()`, `create_fizz()`, `create_inc()`, `create_line()`, `create_line_asis()`, `create_period_scatter()`, `create_rank()`, `create_sankey()`, `create_scatter()`, `create_stacked()`, `create_tracking()`, `create_trend()`, `email_dist()`, `email_fizz()`, `email_line()`, `email_rank()`, `email_summary()`, `email_trend()`, `external_dist()`, `external_fizz()`, `external_line()`, `external_rank()`, `external_sum()`, `hr_trend()`, `hrvar_count()`, `hrvar_trend()`, `keymetrics_scan()`, `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_trend()`, `one2one_dist()`, `one2one_fizz()`, `one2one_freq()`, `one2one_line()`, `one2one_rank()`, `one2one_sum()`, `one2one_trend()`

Other Meetings: `meeting_dist()`, `meeting_fizz()`, `meeting_line()`, `meeting_rank()`, `meeting_tm_report()`, `meeting_trend()`

**Examples**

```
# Return a ggplot bar chart
meeting_summary(pq_data, hrvar = "LevelDesignation")

# Return a summary table
meeting_summary(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

meeting_tm_report	<i>Generate a Meeting Text Mining report in HTML</i>
-------------------	--

---

**Description**

Create a text mining report in HTML based on Meeting Subject Lines

**Usage**

```
meeting_tm_report(
  data,
  path = "meeting text mining report",
  stopwords = NULL,
  timestamp = TRUE,
  keep = 100,
  seed = 100
)
```

**Arguments**

<code>data</code>	A Meeting Query dataset in the form of a data frame.
<code>path</code>	Pass the file path and the desired file name, <i>excluding the file extension</i> . For example, "meeting text mining report".

stopwords	A character vector OR a single-column data frame labelled 'word' containing custom stopwords to remove.
timestamp	Logical vector specifying whether to include a timestamp in the file name. Defaults to TRUE.
keep	A numeric vector specifying maximum number of words to keep.
seed	A numeric vector to set seed for random generation.

### Details

Note that the column Subject must be available within the input data frame in order to run.

### Value

An HTML report with the same file name as specified in the arguments is generated in the working directory. No outputs are directly returned by the function.

### How to run

```
meeting_tm_report(mt_data)
```

This will generate a HTML report as specified in path.

### See Also

Other Reports: [IV\\_report\(\)](#), [generate\\_report\(\)](#), [read\\_preamble\(\)](#), [validation\\_report\(\)](#)

Other Meetings: [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#)

Other Text-mining: [pairwise\\_count\(\)](#), [tm\\_clean\(\)](#), [tm\\_cooc\(\)](#), [tm\\_freq\(\)](#), [tm\\_wordcloud\(\)](#)

---

meeting\_trend

*Meeting Hours Time Trend*

---

### Description

Provides a week by week view of meeting time. By default returns a week by week heatmap, highlighting the points in time with most activity. Additional options available to return a summary table.

### Usage

```
meeting_trend(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

**Details**

Uses the metric Meeting\_hours.

**Value**

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Meetings: [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_tm\\_report\(\)](#)

**Examples**

```
# Run plot
meeting_trend(pq_data)

# Run table
meeting_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

mt\_data

*Sample Meeting Query dataset*

---

**Description**

A dataset generated from a Meeting Query from Viva Insights.

**Usage**

mt\_data

**Format**

A data frame with 612 rows and 41 variables:

**MeetingId**

**Attendee\_meeting\_hours**

**Number\_of\_attendees**

**Number\_of\_attendees\_multitasking**

**Number\_of\_attendees\_who\_didn\_t\_end\_the\_meeting\_on\_time**

**Number\_of\_attendees\_who\_didn\_t\_join\_the\_meeting\_on\_time**

**Number\_of\_attendees\_who\_ended\_the\_meeting\_on\_time**

**Number\_of\_attendees\_who\_joined\_the\_meeting\_on\_time**

**Number\_of\_chats\_sent\_during\_the\_meeting**

**Number\_of\_emails\_sent\_during\_the\_meeting**

**Number\_of\_redundant\_attendees**

**Subject**

**All\_Day\_Meeting**

**Cancelled**

**Recurring**

**Accept\_count**

**No\_response\_count**

**Decline\_count**

**Tentatively\_accepted\_count**

**Intended\_participant\_count**

**Collaboration\_start\_time**

**Organizer**

**zId**

**attainment**

**TimeZone**

**SupervisorIndicator**  
**Region**  
**Population\_Type**  
**Organization**  
**OnsiteDays**  
**Number\_of\_directs**  
**LevelDesignation**  
**Layer**  
**HireDate**  
**GroupNum**  
**GroupName**  
**FunctionType**  
**Domain**  
**ADO\_PersonSK**  
**ADO\_PersonIndicator**  
**Duration**

**Value**

data frame.

**Source**

<https://learn.microsoft.com/en-us/viva/insights/advanced/analyst/meeting-query/>

**See Also**

Other Data: [g2g\\_data](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#), [pq\\_data](#)

---

network\_g2g

*Create a network plot with the group-to-group query*

---

**Description**

Pass a data frame containing a group-to-group query and return a network plot. Automatically handles "Within Group" and "Other\_collaborators" values within query data.

**Usage**

```
network_g2g(
  data,
  primary = NULL,
  secondary = NULL,
  metric = "Group_collaboration_time_invested",
  algorithm = "fr",
  node_colour = "lightblue",
  exc_threshold = 0.1,
  org_count = NULL,
  subtitle = "Collaboration Across Organizations",
  return = "plot"
)
```

**Arguments**

data	Data frame containing a group-to-group query.
primary	String containing the variable name for the Primary Collaborator column.
secondary	String containing the variable name for the Secondary Collaborator column.
metric	String containing the variable name for metric. Defaults to Group_collaboration_time_invested.
algorithm	String to specify the node placement algorithm to be used. Defaults to "fr" for the force-directed algorithm of Fruchterman and Reingold. See <a href="https://rdrr.io/cran/ggraph/man/layout_tbl_graph_igraph.html">https://rdrr.io/cran/ggraph/man/layout_tbl_graph_igraph.html</a> for a full list of options.
node_colour	String or named vector to specify the colour to be used for displaying nodes. Defaults to "lightblue". <ul style="list-style-type: none"> <li>• If "vary" is supplied, a different colour is shown for each node at random.</li> <li>• If a named vector is supplied, the names must match the values of the variable provided for the primary and secondary columns. See example section for details.</li> </ul>
exc_threshold	Numeric value between 0 and 1 specifying the exclusion threshold to apply. Defaults to 0.1, which means that the plot will only display collaboration above 10% of a node's total collaboration. This argument has no impact on "data" or "table" return.
org_count	Optional data frame to provide the size of each organization in the secondary attribute. The data frame should contain only two columns: <ul style="list-style-type: none"> <li>• Name of the secondary attribute excluding any prefixes, e.g. "Organization". Must be of character or factor type.</li> <li>• "n". Must be of numeric type. Defaults to NULL, where node sizes will be fixed.</li> </ul>
subtitle	String to override default plot subtitle.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul>

- "network"
- "data"

See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A group-to-group network plot.
- "table": data frame. An interactive matrix of the network.
- "network": 'igraph' object used for creating the network plot.
- "data": data frame. A long table of the underlying data.

### See Also

Other Network: [g2g\\_data](#), [network\\_p2p\(\)](#), [network\\_summary\(\)](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#)

### Examples

```
# Return a network plot
g2g_data %>% network_g2g()

# Return a network plot - Meeting hours and 5% threshold
network_g2g(
  data = g2g_data,
  primary = "PrimaryCollaborator_Organization",
  secondary = "SecondaryCollaborator_Organization",
  exc_threshold = 0.05
)

# Return a network plot - custom-specific colours
# Get labels of orgs and assign random colours
org_str <- unique(g2g_data$PrimaryCollaborator_Organization)

col_str <-
  sample(
    x = heat_colours(n = length(org_str)), # generate colour codes for each one
    size = length(org_str),
    replace = TRUE
  )

# Create and supply a named vector to `node_colour`
names(col_str) <- org_str

g2g_data %>%
  network_g2g(node_colour = col_str)

# Return a network plot with circle layout
# Vary node colours and add org sizes
org_tb <-
```

```

data.frame(
  Organization = c(
    "G&A East",
    "G&A West",
    "G&A North",
    "South Sales",
    "North Sales",
    "G&A South"
  ),
  n = sample(30:1000, size = 6)
)

g2g_data %>%
  network_g2g(algorithm = "circle",
             node_colour = "vary",
             org_count = org_tb)

# Return an interaction matrix
# Minimum arguments specified
g2g_data %>%
  network_g2g(return = "table")

```

---

network\_p2p

---

*Perform network analysis with the person-to-person query*


---

## Description

### [Experimental]

Analyse a person-to-person (P2P) network query, with multiple visualisation and analysis output options. Pass a data frame containing a person-to-person query and return a network visualization. Options are available for community detection using either the Louvain or the Leiden algorithms.

## Usage

```

network_p2p(
  data,
  hrvar = "Organization",
  return = "plot",
  centrality = NULL,
  community = NULL,
  weight = NULL,
  comm_args = NULL,
  layout = "mds",
  path = paste("p2p", community, sep = "_"),
  style = "igraph",
  bg_fill = "#FFFFFF",
  font_col = "grey20",

```

```

    legend_pos = "right",
    palette = "rainbow",
    node_alpha = 0.7,
    edge_alpha = 1,
    edge_col = "#777777",
    node_sizes = c(1, 20),
    seed = 1
)

```

### Arguments

<code>data</code>	Data frame containing a person-to-person query.
<code>hrvar</code>	String containing the label for the HR attribute.
<code>return</code>	A different output is returned depending on the value passed to the return argument: <ul style="list-style-type: none"> <li>• 'plot' (default)</li> <li>• 'plot-pdf'</li> <li>• 'sankey'</li> <li>• 'table'</li> <li>• 'data'</li> <li>• 'network'</li> </ul>
<code>centrality</code>	string to determines which centrality measure is used to scale the size of the nodes. All centrality measures are automatically calculated when it is set to one of the below values, and reflected in the 'network' and 'data' outputs. Measures include: <ul style="list-style-type: none"> <li>• betweenness</li> <li>• closeness</li> <li>• degree</li> <li>• eigenvector</li> <li>• pagerank</li> </ul> When <code>centrality</code> is set to NULL, no centrality is calculated in the outputs and all the nodes would have the same size.
<code>community</code>	String determining which community detection algorithms to apply. Valid values include: <ul style="list-style-type: none"> <li>• NULL (default): compute analysis or visuals without computing communities.</li> <li>• "louvain"</li> <li>• "leiden"</li> <li>• "edge_betweenness"</li> <li>• "fast_greedy"</li> <li>• "fluid_communities"</li> <li>• "infomap"</li> <li>• "label_prop"</li> <li>• "leading_eigen"</li> </ul>

- "optimal"
- "spinglass"
- "walk\_trap"

These values map to the community detection algorithms offered by igraph. For instance, "leiden" is based on `igraph::cluster_leiden()`. Please see the bottom of [https://igraph.org/r/html/1.3.0/cluster\\_leiden.html](https://igraph.org/r/html/1.3.0/cluster_leiden.html) on all applications and parameters of these algorithms. .

weight	String to specify which column to use as weights for the network. To create a graph without weights, supply NULL to this argument.
comm_args	list containing the arguments to be passed through to igraph's clustering algorithms. Arguments must be named. See examples section on how to supply arguments in a named list.
layout	String to specify the node placement algorithm to be used. Defaults to "mds" for the deterministic multi-dimensional scaling of nodes. See <a href="https://rdr.io/cran/ggraph/man/layout_tbl_graph_igraph.html">https://rdr.io/cran/ggraph/man/layout_tbl_graph_igraph.html</a> for a full list of options.
path	File path for saving the PDF output. Defaults to a timestamped path based on current parameters.
style	String to specify which plotting style to use for the network plot. Valid values include: <ul style="list-style-type: none"> <li>• "igraph"</li> <li>• "ggraph"</li> </ul>
bg_fill	String to specify background fill colour.
font_col	String to specify font colour.
legend_pos	String to specify position of legend. Defaults to "right". See <code>ggplot2::theme()</code> . This is applicable for both the 'ggraph' and the fast plotting method. Valid inputs include: <ul style="list-style-type: none"> <li>• "bottom"</li> <li>• "top"</li> <li>• "left" -"right"</li> </ul>
palette	String specifying the function to generate a colour palette with a single argument n. Uses "rainbow" by default.
node_alpha	A numeric value between 0 and 1 to specify the transparency of the nodes. Defaults to 0.7.
edge_alpha	A numeric value between 0 and 1 to specify the transparency of the edges (only for 'ggraph' mode). Defaults to 1.
edge_col	String to specify edge link colour.
node_sizes	Numeric vector of length two to specify the range of node sizes to rescale to, when centrality is set to a non-null value.
seed	Seed for the random number generator passed to either <code>set.seed()</code> when the louvain or leiden community detection algorithm is used, to ensure consistency. Only applicable when community is set to one of the valid non-null values.

**Value**

A different output is returned depending on the value passed to the return argument:

- 'plot': return a network plot, interactively within R.
- 'plot-pdf': save a network plot as PDF. This option is recommended when the graph is large, which make take a long time to run if return = 'plot' is selected. Use this together with path to control the save location.
- 'sankey': return a sankey plot combining communities and HR attribute. This is only valid if a community detection method is selected at community.
- 'table': return a vertex summary table with counts in communities and HR attribute. When centrality is non-NULL, the average centrality values are calculated per group.
- 'data': return a vertex data file that matches vertices with communities and HR attributes.
- 'network': return 'igraph' object.

**See Also**

Other Network: [g2g\\_data](#), [network\\_g2g\(\)](#), [network\\_summary\(\)](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#)

**Examples**

```
p2p_df <- p2p_data_sim(dim = 1, size = 100)

# default - ggraph visual
network_p2p(data = p2p_df, style = "ggraph")

# return vertex table
network_p2p(data = p2p_df, return = "table")

# return vertex table with community detection
network_p2p(data = p2p_df, community = "leiden", return = "table")

# leiden - igraph style with custom resolution parameters
network_p2p(data = p2p_df, community = "leiden", comm_args = list("resolution" = 0.1))

# louvain - ggraph style, using custom palette
network_p2p(
  data = p2p_df,
  style = "ggraph",
  community = "louvain",
  palette = "heat_colors"
)

# leiden - return a sankey visual with custom resolution parameters
network_p2p(
  data = p2p_df,
  community = "leiden",
  return = "sankey",
  comm_args = list("resolution" = 0.1)
)
```

```

# using `fluid_communities` algorithm with custom parameters
network_p2p(
  data = p2p_df,
  community = "fluid_communities",
  comm_args = list("no.of.communities" = 5)
)

# Calculate centrality measures and leiden communities, return at node level
network_p2p(
  data = p2p_df,
  centrality = "betweenness",
  community = "leiden",
  return = "data"
) %>%
  dplyr::glimpse()

```

---

network\_summary

*Summarise node centrality statistics with an igraph object*


---

### Description

Pass an igraph object to the function and obtain centrality statistics for each node in the object as a data frame. This function works as a wrapper of the centralization functions in 'igraph'.

### Usage

```
network_summary(graph, hrvar = NULL, return = "table")
```

### Arguments

graph	'igraph' object that can be returned from network_g2g() or network_p2p() when the return argument is set to "network".
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to NULL.
return	String specifying what output to return. Valid inputs include: <ul style="list-style-type: none"> <li>• "table"</li> <li>• "network"</li> <li>• "plot"</li> </ul>

See Value for more information.

**Value**

By default, a data frame containing centrality statistics. Available statistics include:

- `betweenness`: number of shortest paths going through a node.
- `closeness`: number of steps required to access every other node from a given node.
- `degree`: number of connections linked to a node.
- `eigenvector`: a measure of the influence a node has on a network.
- `pagerank`: calculates the PageRank for the specified vertices. Please refer to the `igraph` package documentation for the detailed technical definition.

When `"network"` is passed to `"return"`, an `'igraph'` object is returned with additional node attributes containing centrality scores.

When `"plot"` is passed to `"return"`, a summary table is returned showing the average centrality scores by HR attribute. This is currently available if there is a valid HR attribute.

**See Also**

Other Network: [g2g\\_data](#), [network\\_g2g\(\)](#), [network\\_p2p\(\)](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#)

**Examples**

```
# Simulate a p2p network
p2p_data <- p2p_data_sim(size = 100)
g <- network_p2p(data = p2p_data, return = "network")

# Return summary table
network_summary(graph = g, return = "table")

# Return network with node centrality statistics
network_summary(graph = g, return = "network")

# Return summary plot
network_summary(graph = g, return = "plot", hrvar = "Organization")

# Simulate a g2g network and return table
g2 <- g2g_data %>% network_g2g(return = "network")
network_summary(graph = g2, return = "table")
```

---

one2one\_dist

*Distribution of Manager 1:1 Time as a 100% stacked bar*

---

**Description**

Analyze Manager 1:1 Time distribution. Returns a stacked bar plot of different buckets of 1:1 time. Additional options available to return a table with distribution elements.

**Usage**

```

one2one_dist(
  data,
  hrvar = "Organization",
  mingroup = 5,
  dist_colours = c("#facebc", "#fcf0eb", "#b4d5dd", "#bfe5ee"),
  return = "plot",
  cut = c(5, 15, 30)
)

```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>hrvar</code>	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
<code>mingroup</code>	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
<code>dist_colours</code>	A character vector of length four to specify colour codes for the stacked bars.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
<code>cut</code>	A numeric vector of length three to specify the breaks for the distribution, e.g. <code>c(10, 15, 20)</code>

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Managerial Relations: [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

### Examples

```
# Return plot
one2one_dist(pq_data, hrvar = "Organization", return = "plot")

# Return summary table
one2one_dist(pq_data, hrvar = "Organization", return = "table")
```

---

one2one_fizz	<i>Distribution of Manager 1:1 Time (Fizzy Drink plot)</i>
--------------	--

---

### Description

Analyze weekly Manager 1:1 Time distribution, and returns a 'fizzy' scatter plot by default. Additional options available to return a table with distribution elements.

### Usage

```
one2one_fizz(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

### Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A jittered scatter plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Managerial Relations: [one2one\\_dist\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

**Examples**

```
# Return plot
one2one_fizz(pq_data, hrvar = "Organization", return = "plot")

# Return a summary table
one2one_fizz(pq_data, hrvar = "Organization", return = "table")
```

---

one2one\_freq

*Frequency of Manager 1:1 Meetings as bar or 100% stacked bar chart*

---

**Description****[Experimental]**

This function calculates the average number of weeks (cadence) between of 1:1 meetings between an employee and their manager. Returns a distribution plot for typical cadence of 1:1 meetings. Additional options available to return a bar plot, tables, or a data frame with a cadence of 1 on 1 meetings metric.

**Usage**

```
one2one_freq(
  data,
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  mode = "dist",
  sort_by = NULL
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul>
mode	String specifying what method to use. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "dist"</li> <li>• "sum"</li> </ul>
sort_by	String to specify the bucket label to sort by. Defaults to NULL (no sorting).

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A stacked bar plot for the metric.
- "table": data frame. A summary table for the metric.

**Distribution view**

For this view, there are four categories of cadence:

- Weekly (once per week)
- Twice monthly or more (up to 3 weeks)
- Monthly (3 - 6 weeks)
- Every two months (6 - 10 weeks)
- Quarterly or less (> 10 weeks)

In the occasion there are zero 1:1 meetings with managers, this is included into the last category, i.e. 'Quarterly or less'. Note that when mode is set to "sum", these rows are simply excluded from the calculation.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#),

external\_sum(), hr\_trend(), hrvar\_count(), hrvar\_trend(), keymetrics\_scan(), meeting\_dist(), meeting\_fizz(), meeting\_line(), meeting\_rank(), meeting\_summary(), meeting\_trend(), one2one\_dist(), one2one\_fizz(), one2one\_line(), one2one\_rank(), one2one\_sum(), one2one\_trend()

Other Managerial Relations: one2one\_dist(), one2one\_fizz(), one2one\_line(), one2one\_rank(), one2one\_sum(), one2one\_trend()

## Examples

```
# Return plot, mode dist
one2one_freq(pq_data, hrvar = "Organization", return = "plot", mode = "dist")

# Return plot, mode sum
one2one_freq(pq_data,
             hrvar = "Organization",
             return = "plot",
             mode = "sum")

# Return summary table
one2one_freq(pq_data, hrvar = "Organization", return = "table")
```

---

one2one\_line

*Manager 1:1 Time Trend - Line Chart*

---

## Description

Provides a week by week view of 1:1 time with managers, visualised as line charts. By default returns a line chart for 1:1 meeting hours, with a separate panel per value in the HR attribute. Additional options available to return a summary table.

## Usage

```
one2one_line(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

## Arguments

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>"plot"</li> <li>"table"</li> </ul> See Value for more information.

**Details**

Uses the metric Meeting\_hours\_with\_manager\_1\_1.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A faceted line plot for the metric.
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Managerial Relations: [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

**Examples**

```
# Return a line plot
one2one_line(pq_data, hrvar = "LevelDesignation")

# Return summary table
one2one_line(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

one2one\_rank

*Manager 1:1 Time Ranking*

---

**Description**

This function scans a standard query output for groups with high levels of 'Manager 1:1 Time'. Returns a plot by default, with an option to return a table with a all of groups (across multiple HR attributes) ranked by manager 1:1 time.

**Usage**

```
one2one_rank(
  data,
  hrvar = extract_hr(data),
  mingroup = 5,
  mode = "simple",
  plot_mode = 1,
  return = "plot"
)
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
mode	String to specify calculation mode. Must be either: <ul style="list-style-type: none"> <li>• "simple"</li> <li>• "combine"</li> </ul>
plot_mode	Numeric vector to determine which plot mode to return. Must be either 1 or 2, and is only used when return = "plot". <ul style="list-style-type: none"> <li>• 1: Top and bottom five groups across the data population are highlighted</li> <li>• 2: Top and bottom groups <i>per</i> organizational attribute are highlighted</li> </ul>
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot" (default)</li> <li>• "table"</li> </ul> See Value for more information.

**Details**

Uses the metric Meeting\_hours\_with\_manager\_1\_1. See create\_rank() for applying the same analysis to a different metric.

**Value**

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bubble plot where the x-axis represents the metric, the y-axis represents the HR attributes, and the size of the bubbles represent the size of the organizations. Note that there is no plot output if mode is set to "combine".
- "table": data frame. A summary table for the metric.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

Other Managerial Relations: [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_sum\(\)](#), [one2one\\_trend\(\)](#)

**Examples**

```
# Return rank table
one2one_rank(data = pq_data, return = "table")

# Return plot
one2one_rank(data = pq_data, return = "plot")
```

---

one2one\_sum

*Manager 1:1 Time Summary*


---

**Description**

Provides an overview analysis of Manager 1:1 Time. Returns a bar plot showing average weekly minutes of Manager 1:1 Time by default. Additional options available to return a summary table.

**Usage**

```
one2one_sum(data, hrvar = "Organization", mingroup = 5, return = "plot")

one2one_summary(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.

return           String specifying what to return. This must be one of the following strings:

- "plot"
- "table"

See Value for more information.

### Value

A different output is returned depending on the value passed to the return argument:

- "plot": 'ggplot' object. A bar plot for the metric.
- "table": data frame. A summary table for the metric.

### See Also

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_trend\(\)](#)

Other Managerial Relations: [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_trend\(\)](#)

### Examples

```
# Return a ggplot bar chart
one2one_sum(pq_data, hrvar = "LevelDesignation")

# Return a summary table
one2one_sum(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

one2one\_trend

*Manager 1:1 Time Trend*

---

### Description

Provides a week by week view of scheduled manager 1:1 Time. By default returns a week by week heatmap, highlighting the points in time with most activity. Additional options available to return a summary table.

**Usage**

```
one2one_trend(data, hrvar = "Organization", mingroup = 5, return = "plot")
```

**Arguments**

data	A Standard Person Query dataset in the form of a data frame.
hrvar	String containing the name of the HR Variable by which to split metrics. Defaults to "Organization". To run the analysis on the total instead of splitting by an HR attribute, supply NULL (without quotes).
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".

**Details**

Uses the metric Meeting\_hours\_with\_manager\_1\_1.

**Value**

Returns a 'ggplot' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**See Also**

Other Visualization: [afterhours\\_dist\(\)](#), [afterhours\\_fizz\(\)](#), [afterhours\\_line\(\)](#), [afterhours\\_rank\(\)](#), [afterhours\\_summary\(\)](#), [afterhours\\_trend\(\)](#), [collaboration\\_area\(\)](#), [collaboration\\_dist\(\)](#), [collaboration\\_fizz\(\)](#), [collaboration\\_line\(\)](#), [collaboration\\_rank\(\)](#), [collaboration\\_sum\(\)](#), [collaboration\\_trend\(\)](#), [create\\_bar\(\)](#), [create\\_bar\\_asis\(\)](#), [create\\_boxplot\(\)](#), [create\\_bubble\(\)](#), [create\\_dist\(\)](#), [create\\_fizz\(\)](#), [create\\_inc\(\)](#), [create\\_line\(\)](#), [create\\_line\\_asis\(\)](#), [create\\_period\\_scatter\(\)](#), [create\\_rank\(\)](#), [create\\_sankey\(\)](#), [create\\_scatter\(\)](#), [create\\_stacked\(\)](#), [create\\_tracking\(\)](#), [create\\_trend\(\)](#), [email\\_dist\(\)](#), [email\\_fizz\(\)](#), [email\\_line\(\)](#), [email\\_rank\(\)](#), [email\\_summary\(\)](#), [email\\_trend\(\)](#), [external\\_dist\(\)](#), [external\\_fizz\(\)](#), [external\\_line\(\)](#), [external\\_rank\(\)](#), [external\\_sum\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_trend\(\)](#), [keymetrics\\_scan\(\)](#), [meeting\\_dist\(\)](#), [meeting\\_fizz\(\)](#), [meeting\\_line\(\)](#), [meeting\\_rank\(\)](#), [meeting\\_summary\(\)](#), [meeting\\_trend\(\)](#), [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#)

Other Managerial Relations: [one2one\\_dist\(\)](#), [one2one\\_fizz\(\)](#), [one2one\\_freq\(\)](#), [one2one\\_line\(\)](#), [one2one\\_rank\(\)](#), [one2one\\_sum\(\)](#)

**Examples**

```
# Run plot
one2one_trend(pq_data)

# Run table
one2one_trend(pq_data, hrvar = "LevelDesignation", return = "table")
```

---

p2p\_data

*Sample person-to-person dataset*

---

### Description

A demo dataset representing a person-to-person query, structured as an edgelist. The identifier variable for each person is PersonId, where the variables have been prefixed with PrimaryCollaborator\_ and SecondaryCollaborator\_ to represent the direction of collaboration.

### Usage

```
p2p_data
```

### Format

A data frame with 11550 rows and 13 variables:

**PrimaryCollaborator\_PersonId**  
**SecondaryCollaborator\_PersonId**  
**MetricDate**  
**Diverse\_tie\_score**  
**Diverse\_tie\_type**  
**Strong\_tie\_score**  
**Strong\_tie\_type**  
**PrimaryCollaborator\_Organization**  
**SecondaryCollaborator\_Organization**  
**PrimaryCollaborator\_LevelDesignation**  
**SecondaryCollaborator\_LevelDesignation**  
**PrimaryCollaborator\_FunctionType**  
**SecondaryCollaborator\_FunctionType** ...

### Value

data frame.

### Source

<https://analysis.insights.viva.office.com/analyst/analysis/>

### See Also

Other Data: [g2g\\_data](#), [mt\\_data](#), [p2p\\_data\\_sim\(\)](#), [pq\\_data](#)

Other Network: [g2g\\_data](#), [network\\_g2g\(\)](#), [network\\_p2p\(\)](#), [network\\_summary\(\)](#), [p2p\\_data\\_sim\(\)](#)

---

p2p\_data\_sim

*Simulate a person-to-person query using a Watts-Strogatz model*

---

### Description

Generate an person-to-person query / edgelist based on the graph according to the Watts-Strogatz small-world network model. Organizational data fields are also simulated for Organization, LevelDesignation, and City.

### Usage

```
p2p_data_sim(dim = 1, size = 300, nei = 5, p = 0.05)
```

### Arguments

dim	Integer constant, the dimension of the starting lattice.
size	Integer constant, the size of the lattice along each dimension.
nei	Integer constant, the neighborhood within which the vertices of the lattice will be connected.
p	Real constant between zero and one, the rewiring probability.

### Details

This is a wrapper around `igraph::watts.strogatz.game()`. See `igraph` documentation for details on methodology. Loop edges and multiple edges are disabled. Size of the network can be changing the arguments `size` and `nei`.

### Value

data frame with the same column structure as a person-to-person flexible query. This has an edgelist structure and can be used directly as an input to `network_p2p()`.

### See Also

Other Data: [g2g\\_data](#), [mt\\_data](#), [p2p\\_data](#), [pq\\_data](#)

Other Network: [g2g\\_data](#), [network\\_g2g\(\)](#), [network\\_p2p\(\)](#), [network\\_summary\(\)](#), [p2p\\_data](#)

### Examples

```
# Simulate a p2p dataset with 800 edges
p2p_data_sim(size = 200, nei = 4)
```

---

pad2	<i>Create the two-digit zero-padded format</i>
------	--

---

**Description**

Create the two-digit zero-padded format

**Usage**

```
pad2(x)
```

**Arguments**

x                    numeric value or vector with maximum two characters.

**Value**

Numeric value containing two-digit zero-padded values.

---

pairwise_count	<i>Perform a pairwise count of words by id</i>
----------------	--

---

**Description**

This is a 'data.table' implementation that mimics the output of pairwise\_count() from 'widy' to reduce package dependency. This is used internally within tm\_cooc().

**Usage**

```
pairwise_count(data, id = "line", word = "word")
```

**Arguments**

data                Data frame output from tm\_clean().  
id                   String to represent the id variable. Defaults to "line".  
word                String to represent the word variable. Defaults to "word".

**Value**

data frame with the following columns representing a pairwise count:

- "item1"
- "item2"
- "n"

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

Other Text-mining: [meeting\\_tm\\_report\(\)](#), [tm\\_clean\(\)](#), [tm\\_cooc\(\)](#), [tm\\_freq\(\)](#), [tm\\_wordcloud\(\)](#)

**Examples**

```
td <- data.frame(line = c(1, 1, 2, 2),
                 word = c("work", "meeting", "catch", "up"))

pairwise_count(td, id = "line", word = "word")
```

---

pq\_data

*Sample Person Query dataset*

---

**Description**

A dataset generated from a Person Query from Viva Insights.

**Usage**

```
pq_data
```

**Format**

A data frame with 1000 rows and 154 variables:

**PersonId**

**MetricDate**

**After\_hours\_call\_hours**

**After\_hours\_chat\_hours**

**After\_hours\_collaboration\_hours**

**After\_hours\_email\_hours**

**After\_hours\_meeting\_hours**

**After\_hours\_scheduled\_call\_hours**

**After\_hours\_unscheduled\_call\_hours**

**Call\_hours**

**Calls**

**Chat\_hours**

**Chats\_sent**

**Collaboration\_hours**

**Email\_hours**  
**Emails\_sent**  
**Meeting\_and\_call\_hours**  
**Meeting\_hours**  
**Meetings**  
**Multitasking\_hours**  
**Scheduled\_call\_hours**  
**Unscheduled\_call\_hours**  
**Urgent\_email\_hours**  
**Urgent\_meeting\_hours**  
**Collaboration\_hours\_on\_Friday**  
**Collaboration\_hours\_on\_Monday**  
**Collaboration\_hours\_on\_Saturday**  
**Collaboration\_hours\_on\_Sunday**  
**Collaboration\_hours\_on\_Thursday**  
**Collaboration\_hours\_on\_Tuesday**  
**Collaboration\_hours\_on\_Wednesday**  
**Meeting\_hours\_on\_Friday**  
**Meeting\_hours\_on\_Monday**  
**Meeting\_hours\_on\_Saturday**  
**Meeting\_hours\_on\_Sunday**  
**Meeting\_hours\_on\_Thursday**  
**Meeting\_hours\_on\_Tuesday**  
**Meeting\_hours\_on\_Wednesday**  
**Unscheduled\_weekend\_calls**  
**Weekend\_chats\_sent**  
**Weekend\_emails\_sent**  
**Weekend\_meetings**  
**Chats\_sent\_00\_01**  
**Chats\_sent\_01\_02**  
**Chats\_sent\_02\_03**  
**Chats\_sent\_03\_04**  
**Chats\_sent\_04\_05**  
**Chats\_sent\_05\_06**  
**Chats\_sent\_06\_07**  
**Chats\_sent\_07\_08**  
**Chats\_sent\_08\_09**

**Chats\_sent\_09\_10**  
**Chats\_sent\_10\_11**  
**Chats\_sent\_11\_12**  
**Chats\_sent\_12\_13**  
**Chats\_sent\_13\_14**  
**Chats\_sent\_14\_15**  
**Chats\_sent\_15\_16**  
**Chats\_sent\_16\_17**  
**Chats\_sent\_17\_18**  
**Chats\_sent\_18\_19**  
**Chats\_sent\_19\_20**  
**Chats\_sent\_20\_21**  
**Chats\_sent\_21\_22**  
**Chats\_sent\_22\_23**  
**Chats\_sent\_23\_24**  
**Emails\_sent\_00\_01**  
**Emails\_sent\_01\_02**  
**Emails\_sent\_02\_03**  
**Emails\_sent\_03\_04**  
**Emails\_sent\_04\_05**  
**Emails\_sent\_05\_06**  
**Emails\_sent\_06\_07**  
**Emails\_sent\_07\_08**  
**Emails\_sent\_08\_09**  
**Emails\_sent\_09\_10**  
**Emails\_sent\_10\_11**  
**Emails\_sent\_11\_12**  
**Emails\_sent\_12\_13**  
**Emails\_sent\_13\_14**  
**Emails\_sent\_14\_15**  
**Emails\_sent\_15\_16**  
**Emails\_sent\_16\_17**  
**Emails\_sent\_17\_18**  
**Emails\_sent\_18\_19**  
**Emails\_sent\_19\_20**  
**Emails\_sent\_20\_21**  
**Emails\_sent\_21\_22**

**Emails\_sent\_22\_23**  
**Emails\_sent\_23\_24**  
**Meetings\_00\_01**  
**Meetings\_01\_02**  
**Internal\_meeting\_hours\_with\_3\_to\_8\_attendees**  
**Internal\_meeting\_hours\_without\_manager\_1\_1**  
**Small\_group\_chats\_sent\_excluding\_manager**  
**Small\_group\_emails\_sent\_excluding\_manager**  
**Small\_group\_meeting\_call\_and\_chat\_hours**  
**Collaboration\_hours\_with\_direct\_reports**  
**Manager\_coaching\_hours\_1\_1**  
**Meeting\_and\_call\_hours\_with\_manager**  
**Meeting\_and\_call\_hours\_with\_manager\_1\_1**  
**Meeting\_and\_call\_hours\_with\_skip\_level**  
**Meeting\_hours\_with\_manager**  
**Meeting\_hours\_with\_manager\_1\_1**  
**Meeting\_hours\_with\_skip\_level**  
**Meetings\_with\_manager**  
**Meetings\_with\_manager\_1\_1**  
**Meetings\_with\_skip\_level**  
**Small\_group\_chats\_sent\_including\_manager**  
**Small\_group\_emails\_sent\_including\_manager**  
**Unscheduled\_call\_hours\_with\_manager**  
**Unscheduled\_call\_hours\_with\_manager\_1\_1**  
**Unscheduled\_call\_hours\_with\_skip\_level**  
**Internal\_network\_size**  
**Conflicting\_meeting\_hours**  
**Large\_and\_long\_meeting\_hours**  
**Large\_and\_long\_recurring\_meeting\_hours**  
**Large\_and\_short\_meeting\_hours**  
**Large\_and\_short\_recurring\_meeting\_hours**  
**Meeting\_hours\_ended\_on\_time**  
**Meeting\_hours\_joined\_on\_time**  
**Meeting\_hours\_not\_ended\_on\_time**  
**Meeting\_hours\_not\_joined\_on\_time**  
**Meeting\_hours\_with\_12\_to\_24\_hours\_of\_advanced\_notice**  
**Meeting\_hours\_with\_24\_or\_more\_hours\_of\_advanced\_notice**

**Meeting\_hours\_with\_six\_or\_fewer\_hours\_of\_advanced\_notice**  
**Meeting\_hours\_with\_six\_to\_12\_hours\_of\_advanced\_notice**  
**Recurring\_meeting\_hours**  
**Small\_and\_long\_meeting\_hours**  
**Small\_and\_long\_recurring\_meeting\_hours**  
**Small\_and\_short\_meeting\_hours**  
**Small\_and\_short\_recurring\_meeting\_hours**  
**Available\_to\_focus\_hours**  
**Interrupted\_hours**  
**Uninterrupted\_hours**  
**External\_chat\_hours**  
**External\_collaboration\_hours**  
**External\_email\_hours**  
**External\_meeting\_hours**  
**External\_unscheduled\_call\_hours**  
**Working\_hours\_call\_hours**  
**Working\_hours\_chat\_hours**  
**Working\_hours\_collaboration\_hours**  
**Working\_hours\_email\_hours**  
**Working\_hours\_meeting\_hours**  
**Working\_hours\_scheduled\_call\_hours**  
**Working\_hours\_unscheduled\_call\_hours**  
**LevelDesignation**  
**Layer**  
**SupervisorIndicator**  
**Organization**  
**FunctionType**  
**WeekendDays**  
**IsActive**

**Value**

data frame.

**Source**

<https://learn.microsoft.com/en-us/viva/insights/advanced/analyst/person-query/>

**See Also**

Other Data: [g2g\\_data](#), [mt\\_data](#), [p2p\\_data](#), [p2p\\_data\\_sim\(\)](#)

---

read_preamble	<i>Read preamble</i>
---------------	----------------------

---

**Description**

Read in a preamble to be used within each individual reporting function. Reads from the Markdown file installed with the package.

**Usage**

```
read_preamble(path)
```

**Arguments**

path	Text string containing the path for the appropriate Markdown file.
------	--

**Value**

String containing the text read in from the specified Markdown file.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

Other Reports: [IV\\_report\(\)](#), [generate\\_report\(\)](#), [meeting\\_tm\\_report\(\)](#), [validation\\_report\(\)](#)

---

rgb2hex	<i>Convert rgb to HEX code</i>
---------	--------------------------------

---

**Description**

Convert rgb to HEX code

**Usage**

```
rgb2hex(r, g, b)
```

**Arguments**

r, g, b	Values that correspond to the three RGB parameters
---------	--

**Value**

Returns a string containing a HEX code.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

---

theme_wpa	<i>Main theme for 'vivainsights' visualisations</i>
-----------	---

---

**Description**

A theme function applied to 'ggplot' visualisations in 'vivainsights'. Install and load 'extrafont' to use custom fonts for plotting.

**Usage**

```
theme_wpa(font_size = 12, font_family = "Segoe UI")
```

**Arguments**

font_size	Numeric value that prescribes the base font size for the plot. The text elements are defined relatively to this base font size. Defaults to 12.
font_family	Character value specifying the font family to be used in the plot. The default value is "Segoe UI". To ensure you can use this font, install and load 'extrafont' prior to plotting. There is an initialisation process that is described by: <a href="https://stackoverflow.com/questions/34522732/changing-fonts-in-ggplot2">https://stackoverflow.com/questions/34522732/changing-fonts-in-ggplot2</a>

**Value**

Returns a ggplot object with the applied theme.

**See Also**

Other Themes: [theme\\_wpa\\_basic\(\)](#)

---

theme_wpa_basic	<i>Basic theme for 'vivainsights' visualisations</i>
-----------------	--

---

**Description**

A theme function applied to 'ggplot' visualisations in 'vivainsights'. Based on theme\_wpa() but has no font requirements.

**Usage**

```
theme_wpa_basic(font_size = 12)
```

**Arguments**

font\_size      Numeric value that prescribes the base font size for the plot. The text elements are defined relatively to this base font size. Defaults to 12.

**Value**

Returns a ggplot object with the applied theme.

**See Also**

Other Themes: [theme\\_wpa\(\)](#)

---

tm_clean	<i>Clean subject line text prior to analysis</i>
----------	--

---

**Description**

This function processes the Subject column in a Meeting Query by applying tokenisation using `tidytext::unnest_tokens()`, and removing any stopwords supplied in a data frame (using the argument `stopwords`). This is a sub-function that feeds into `tm_freq()`, `tm_cooc()`, and `tm_wordcloud()`. The default is to return a data frame with tokenised counts of words or ngrams.

**Usage**

```
tm_clean(data, token = "words", stopwords = NULL, ...)
```

**Arguments**

data            A Meeting Query dataset in the form of a data frame.

token          A character vector accepting either "words" or "ngrams", determining type of tokenisation to return.

stopwords     A character vector OR a single-column data frame labelled 'word' containing custom stopwords to remove.

...            Additional parameters to pass to `tidytext::unnest_tokens()`.

**Value**

data frame with two columns:

- line
- word

**See Also**

Other Text-mining: [meeting\\_tm\\_report\(\)](#), [pairwise\\_count\(\)](#), [tm\\_cooc\(\)](#), [tm\\_freq\(\)](#), [tm\\_wordcloud\(\)](#)

**Examples**

```
# words
tm_clean(mt_data)

# ngrams
tm_clean(mt_data, token = "ngrams")
```

tm\_cooc

*Analyse word co-occurrence in subject lines and return a network plot***Description**

This function generates a word co-occurrence network plot, with options to return a table. This function is used within `meeting_tm_report()`.

**Usage**

```
tm_cooc(data, stopwords = NULL, seed = 100, return = "plot", lmult = 0.05)
```

**Arguments**

<code>data</code>	A Meeting Query dataset in the form of a data frame.
<code>stopwords</code>	A character vector OR a single-column data frame labelled 'word' containing custom stopwords to remove.
<code>seed</code>	A numeric vector to set seed for random generation.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
<code>lmult</code>	A multiplier to adjust the line width in the output plot. Defaults to 0.05.

**Details**

This function uses `tm_clean()` as the underlying data wrangling function. There is an option to remove stopwords by passing a data frame into the `stopwords` argument.

**Value**

A different output is returned depending on the value passed to the `return` argument:

- "plot": 'ggplot' and 'ggraph' object. A network plot.
- "table": data frame. A summary table.

**Example**

The function can be run with subject lines from `mt_data`, as per below.

```
mt_data %>%
  tm_cooc(lmult = 0.01)
```

**Author(s)**

Carlos Morales [carlos.morales@microsoft.com](mailto:carlos.morales@microsoft.com)

**See Also**

Other Text-mining: [meeting\\_tm\\_report\(\)](#), [pairwise\\_count\(\)](#), [tm\\_clean\(\)](#), [tm\\_freq\(\)](#), [tm\\_wordcloud\(\)](#)

**Examples**

```
# Demo using a subset of `mt_data`
```

---

<code>tm_freq</code>	<i>Perform a Word or Ngram Frequency Analysis and return a Circular Bar Plot</i>
----------------------	--

---

**Description**

Generate a circular bar plot with frequency of words / ngrams. This function is used within `meeting_tm_report()`.

**Usage**

```
tm_freq(data, token = "words", stopwords = NULL, keep = 100, return = "plot")
```

**Arguments**

<code>data</code>	A Meeting Query dataset in the form of a data frame.
<code>token</code>	A character vector accepting either "words" or "ngram", determining type of tokenisation to return.
<code>stopwords</code>	A character vector OR a single-column data frame labelled 'word' containing custom stopwords to remove.
<code>keep</code>	A numeric vector specifying maximum number of words to keep.
<code>return</code>	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul>

See Value for more information.

## Details

This function uses `tm_clean()` as the underlying data wrangling function. There is an option to remove stopwords by passing a data frame into the `stopwords` argument.

## Value

A different output is returned depending on the value passed to the `return` argument:

- "plot": 'ggplot' object. A circular bar plot.
- "table": data frame. A summary table.

## See Also

Other Text-mining: [meeting\\_tm\\_report\(\)](#), [pairwise\\_count\(\)](#), [tm\\_clean\(\)](#), [tm\\_cooc\(\)](#), [tm\\_wordcloud\(\)](#)

## Examples

```
# circular network plot with words
tm_freq(mt_data, token = "words")

# circular network plot with ngrams
tm_freq(mt_data, token = "ngrams")

# summary table of text frequency
tm_freq(mt_data, token = "words", return = "table")
```

---

tm\_wordcloud

*Generate a wordcloud with meeting subject lines*

---

## Description

Generate a wordcloud with the meeting query. This is a sub-function that feeds into `meeting_tm_report()`.

## Usage

```
tm_wordcloud(
  data,
  stopwords = NULL,
  seed = 100,
  keep = 100,
  return = "plot",
  ...
)
```

**Arguments**

data	A Meeting Query dataset in the form of a data frame.
stopwords	A character vector OR a single-column data frame labelled 'word' containing custom stopwords to remove.
seed	A numeric vector to set seed for random generation.
keep	A numeric vector specifying maximum number of words to keep.
return	String specifying what to return. This must be one of the following strings: <ul style="list-style-type: none"> <li>• "plot"</li> <li>• "table"</li> </ul> See Value for more information.
...	Additional parameters to be passed to <code>ggwordcloud::geom_text_wordcloud()</code>

**Details**

Uses the 'ggwordcloud' package for the underlying implementation, thus returning a 'ggplot' object. Additional layers can be added onto the plot using a `ggplot + syntax`. The recommendation is not to return over 100 words in a word cloud.

This function uses `tm_clean()` as the underlying data wrangling function. There is an option to remove stopwords by passing a data frame into the `stopwords` argument.

**Value**

A different output is returned depending on the value passed to the `return` argument:

- "plot": 'ggplot' object containing a word cloud.
- "table": data frame returning the data used to generate the word cloud.

**See Also**

Other Text-mining: [meeting\\_tm\\_report\(\)](#), [pairwise\\_count\(\)](#), [tm\\_clean\(\)](#), [tm\\_cooc\(\)](#), [tm\\_freq\(\)](#)

**Examples**

```
tm_wordcloud(mt_data, keep = 30)

# Removing stopwords
tm_wordcloud(mt_data, keep = 30, stopwords = c("weekly", "update"))
```

---

totals_bind	<i>Row-bind an identical data frame for computing grouped totals</i>
-------------	--

---

## Description

Row-bind an identical data frame and impute a specific column with the `target_value`, which defaults as "Total". The purpose of this is to enable to creation of summary tables with a calculated "Total" row. See example below on usage.

## Usage

```
totals_bind(data, target_col, target_value = "Total")
```

## Arguments

<code>data</code>	data frame
<code>target_col</code>	Character value of the column in which to impute "Total". This is usually the intended grouping column.
<code>target_value</code>	Character value to impute in the new data frame to row-bind. Defaults to "Total".

## Value

data frame with twice the number of rows of the input data frame, where half of those rows will have the `target_col` column imputed with the value from `target_value`.

## See Also

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

## Examples

```
pq_data %>%
  totals_bind(target_col = "LevelDesignation", target_value = "Total") %>%
  create_bar(hrvar = "LevelDesignation", metric = "Email_hours", return = "table")
```

---

totals_col	<i>Fabricate a 'Total' HR variable</i>
------------	--

---

### Description

Create a 'Total' column of character type comprising exactly of one unique value. This is a convenience function for returning a no-HR attribute view when NULL is supplied to the hrvar argument in functions.

### Usage

```
totals_col(data, total_value = "Total")
```

### Arguments

data	data frame
total_value	Character value defining the name and the value of the "Total" column. Defaults to "Total". An error is returned if an existing variable has the same name as the supplied value.

### Value

data frame containing an additional 'Total' column on top of the input data frame.

### See Also

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

### Examples

```
# Create a visual without HR attribute breaks
pq_data %>%
  totals_col() %>%
  create_fizz(hrvar = "Total", metric = "Email_hours")
```

---

track_HR_change	<i>Sankey chart of organizational movement between HR attributes and missing values (outside company move) (Data Overview)</i>
-----------------	--

---

### Description

Creates a list of everyone at a specified start date and a specified end date then aggregates up people who have moved between organizations between this to points of time and visualizes the move through a sankey chart.

Through this chart you can see:

- The HR attribute/orgs that have the highest move out
- The HR attribute/orgs that have the highest move in
- The number of people that do not have that HR attribute or if they are no longer in the system

### Usage

```
track_HR_change(
  data,
  start_date = min(data$MetricDate),
  end_date = max(data$MetricDate),
  hrvar = "Organization",
  mingroup = 5,
  return = "plot",
  NA_replacement = "Out of Company"
)
```

### Arguments

data	A Person Query dataset in the form of a data frame.
start_date	A start date to compare changes. See end_date.
end_date	An end date to compare changes. See start_date.
hrvar	HR Variable by which to compare changes between, defaults to "Organization" but accepts any character vector, e.g. "LevelDesignation"
mingroup	Numeric value setting the privacy threshold / minimum group size. Defaults to 5.
return	Character vector specifying what to return, defaults to "plot". Valid inputs are "plot" and "table".
NA_replacement	Character replacement for NA defaults to "out of company"

### Value

Returns a 'NetworkD3' object by default, where 'plot' is passed in return. When 'table' is passed, a summary table is returned as a data frame.

**Author(s)**

Tannaz Sattari Tabrizi [Tannaz.Sattari@microsoft.com](mailto:Tannaz.Sattari@microsoft.com)

**See Also**

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [validation\\_report\(\)](#)

**Examples**

```
pq_data %>% track_HR_change()
```

---

tstamp

*Generate a time stamp*

---

**Description**

This function generates a time stamp of the format 'yymmdd\_hhmmss'. This is a support function and is not intended for direct use.

**Usage**

```
tstamp()
```

**Value**

String containing the timestamp in the format 'yymmdd\_hhmmss'.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [us\\_to\\_space\(\)](#), [wrap\(\)](#)

---

us_to_space	<i>Replace underscore with space</i>
-------------	--------------------------------------

---

**Description**

Convenience function to convert underscores to space

**Usage**

```
us_to_space(x)
```

**Arguments**

x                      String to replace all occurrences of \_ with a single space

**Value**

Character vector containing the modified string.

**See Also**

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [wrap\(\)](#)

**Examples**

```
us_to_space("Meeting_hours_with_manager_1_on_1")
```

---

validation_report	<i>Generate a Data Validation report in HTML</i>
-------------------	--

---

**Description**

The function generates an interactive HTML report using Standard Person Query data as an input. The report contains checks on Viva Insights query outputs to provide diagnostic information for the Analyst prior to analysis.

An additional Standard Meeting Query can be provided to perform meeting subject line related checks. This is optional and the validation report can be run without it.

**Usage**

```
validation_report(
  data,
  meeting_data = NULL,
  hrvar = "Organization",
  path = "validation report",
  hrvar_threshold = 150,
  timestamp = TRUE
)
```

**Arguments**

<code>data</code>	A Standard Person Query dataset in the form of a data frame.
<code>meeting_data</code>	An optional Meeting Query dataset in the form of a data frame.
<code>hrvar</code>	HR Variable by which to split metrics, defaults to "Organization" but accepts any character vector, e.g. "Organization"
<code>path</code>	Pass the file path and the desired file name, <i>excluding the file extension</i> .
<code>hrvar_threshold</code>	Numeric value determining the maximum number of unique values to be allowed to qualify as a HR variable. This is passed directly to the threshold argument within <code>hrvar_count_all()</code> .
<code>timestamp</code>	Logical vector specifying whether to include a timestamp in the file name. Defaults to TRUE.

**Details**

For your input to `data` or `meeting_data`, please use the function `vivainsights::import_query()` to import your csv query files into R. This function will standardize format and prepare the data as input for this report.

For most variables, a note is returned in-line instead of an error if the variable is not available.

**Value**

An HTML report with the same file name as specified in the arguments is generated in the working directory. No outputs are directly returned by the function.

**Checking functions within `validation_report()`**

- `check_query()`
- `flag_ch_ratio()`
- `hrvar_count_all()`
- `identify_privacythreshold()`
- `identify_nkw()`
- `identify_holidayweeks()`
- `subject_validate()` (available in 'wpa')

- [identify\\_tenure\(\)](#)
- [flag\\_outlooktime\(\)](#)
- [identify\\_shifts\(\)](#)
- [track\\_HR\\_change\(\)](#)

You can browse each individual function for details on calculations.

### Creating a report

Below is an example on how to run the report.

```
validation_report(pq_data,
                 hrvar = "Organization")
```

### See Also

Other Reports: [IV\\_report\(\)](#), [generate\\_report\(\)](#), [meeting\\_tm\\_report\(\)](#), [read\\_preamble\(\)](#)

Other Data Validation: [check\\_query\(\)](#), [extract\\_hr\(\)](#), [flag\\_ch\\_ratio\(\)](#), [flag\\_em\\_ratio\(\)](#), [flag\\_extreme\(\)](#), [flag\\_outlooktime\(\)](#), [hr\\_trend\(\)](#), [hrvar\\_count\(\)](#), [hrvar\\_count\\_all\(\)](#), [hrvar\\_trend\(\)](#), [identify\\_churn\(\)](#), [identify\\_holidayweeks\(\)](#), [identify\\_inactiveweeks\(\)](#), [identify\\_nkw\(\)](#), [identify\\_outlier\(\)](#), [identify\\_privacythreshold\(\)](#), [identify\\_shifts\(\)](#), [identify\\_tenure\(\)](#), [track\\_HR\\_change\(\)](#)

---

wrap

*Add a character at the start and end of a character string*

---

### Description

This function adds a character at the start and end of a character string, where the default behaviour is to add a double quote.

### Usage

```
wrap(string, wrapper = "\"")
```

### Arguments

string	Character string to be wrapped around
wrapper	Character to wrap around string

### Value

Character vector containing the modified string.

### See Also

Other Support: [any\\_idate\(\)](#), [camel\\_clean\(\)](#), [check\\_inputs\(\)](#), [cut\\_hour\(\)](#), [extract\\_date\\_range\(\)](#), [extract\\_hr\(\)](#), [heat\\_colours\(\)](#), [is\\_date\\_format\(\)](#), [maxmin\(\)](#), [pairwise\\_count\(\)](#), [read\\_preamble\(\)](#), [rgb2hex\(\)](#), [totals\\_bind\(\)](#), [totals\\_col\(\)](#), [tstamp\(\)](#), [us\\_to\\_space\(\)](#)

---

`wrap_text`*Wrap text based on character threshold*

---

**Description**

Wrap text in visualizations according to a preset character threshold. The next space in the string is replaced with `\n`, which will render as next line in plots and messages.

**Usage**

```
wrap_text(x, threshold = 15)
```

**Arguments**

<code>x</code>	String to wrap text
<code>threshold</code>	Numeric, defaults to 15. Number of character units by which the next space would be replaced with <code>\n</code> to move text to next line.

**Value**

String output representing a processed version of `x`, with spaces replaced by `\n`.

**Examples**

```
wrapped <- wrap_text(  
  "The total entropy of an isolated system can never decrease."  
)  
message(wrapped)
```

# Index

## \* **After-hours Collaboration**

- afterhours\_dist, 4
- afterhours\_fizz, 6
- afterhours\_line, 7
- afterhours\_rank, 9
- afterhours\_summary, 10
- afterhours\_trend, 12
- external\_rank, 81

## \* **Collaboration**

- collaboration\_area, 18
- collaboration\_dist, 19
- collaboration\_fizz, 21
- collaboration\_line, 22
- collaboration\_rank, 24
- collaboration\_sum, 26
- collaboration\_trend, 27

## \* **Data Validation**

- check\_query, 16
- extract\_hr, 85
- flag\_ch\_ratio, 86
- flag\_em\_ratio, 87
- flag\_extreme, 88
- flag\_outlooktime, 90
- hr\_trend, 99
- hrvar\_count, 95
- hrvar\_count\_all, 96
- hrvar\_trend, 98
- identify\_churn, 100
- identify\_holidayweeks, 103
- identify\_inactiveweeks, 105
- identify\_nkw, 106
- identify\_outlier, 107
- identify\_privacythreshold, 108
- identify\_shifts, 109
- identify\_tenure, 110
- track\_HR\_change, 167
- validation\_report, 169

## \* **Data**

- g2g\_data, 91

- mt\_data, 130
- p2p\_data, 150
- p2p\_data\_sim, 151
- pq\_data, 153

## \* **Emails**

- email\_dist, 68
- email\_fizz, 69
- email\_line, 70
- email\_rank, 72
- email\_summary, 73
- email\_trend, 75

## \* **External Collaboration**

- external\_dist, 77
- external\_fizz, 79
- external\_line, 80
- external\_sum, 83

## \* **Flexible**

- create\_bar, 30
- create\_bar\_asis, 32
- create\_boxplot, 34
- create\_bubble, 36
- create\_density, 37
- create\_dist, 39
- create\_fizz, 42
- create\_hist, 43
- create\_inc, 45
- create\_line, 49
- create\_line\_asis, 51
- create\_period\_scatter, 53
- create\_rank, 55
- create\_sankey, 58
- create\_scatter, 60
- create\_stacked, 61
- create\_tracking, 64
- create\_trend, 65

## \* **Import and Export**

- copy\_df, 29
- create\_dt, 41
- export, 76

- import\_query, 112
- \* **Information Value**
  - create\_IV, 47
  - IV\_report, 113
- \* **Managerial Relations**
  - one2one\_dist, 139
  - one2one\_fizz, 141
  - one2one\_freq, 142
  - one2one\_line, 144
  - one2one\_rank, 145
  - one2one\_sum, 147
  - one2one\_trend, 148
- \* **Meetings**
  - meeting\_dist, 120
  - meeting\_fizz, 122
  - meeting\_line, 123
  - meeting\_rank, 124
  - meeting\_summary, 126
  - meeting\_tm\_report, 127
  - meeting\_trend, 128
- \* **Network**
  - g2g\_data, 91
  - network\_g2g, 131
  - network\_p2p, 134
  - network\_summary, 138
  - p2p\_data, 150
  - p2p\_data\_sim, 151
- \* **Reports**
  - generate\_report, 92
  - IV\_report, 113
  - meeting\_tm\_report, 127
  - read\_preamble, 158
  - validation\_report, 169
- \* **Support**
  - any\_idate, 14
  - camel\_clean, 15
  - check\_inputs, 15
  - cut\_hour, 67
  - extract\_date\_range, 84
  - extract\_hr, 85
  - heat\_colours, 94
  - is\_date\_format, 113
  - maxmin, 119
  - pairwise\_count, 152
  - read\_preamble, 158
  - rgb2hex, 158
  - totals\_bind, 165
  - totals\_col, 166
  - tstamp, 168
  - us\_to\_space, 169
  - wrap, 171
- \* **Text-mining**
  - meeting\_tm\_report, 127
  - pairwise\_count, 152
  - tm\_clean, 160
  - tm\_cooc, 161
  - tm\_freq, 162
  - tm\_wordcloud, 163
- \* **Themes**
  - theme\_wpa, 159
  - theme\_wpa\_basic, 159
- \* **Time-series**
  - create\_line, 49
  - create\_line\_asis, 51
  - create\_period\_scatter, 53
  - create\_trend, 65
- \* **Variable Association**
  - create\_IV, 47
  - IV\_report, 113
- \* **Visualization**
  - afterhours\_dist, 4
  - afterhours\_fizz, 6
  - afterhours\_line, 7
  - afterhours\_rank, 9
  - afterhours\_summary, 10
  - afterhours\_trend, 12
  - collaboration\_area, 18
  - collaboration\_dist, 19
  - collaboration\_fizz, 21
  - collaboration\_line, 22
  - collaboration\_rank, 24
  - collaboration\_sum, 26
  - collaboration\_trend, 27
  - create\_bar, 30
  - create\_bar\_asis, 32
  - create\_boxplot, 34
  - create\_bubble, 36
  - create\_dist, 39
  - create\_fizz, 42
  - create\_inc, 45
  - create\_line, 49
  - create\_line\_asis, 51
  - create\_period\_scatter, 53
  - create\_rank, 55
  - create\_sankey, 58
  - create\_scatter, 60

- create\_stacked, 61
- create\_tracking, 64
- create\_trend, 65
- email\_dist, 68
- email\_fizz, 69
- email\_line, 70
- email\_rank, 72
- email\_summary, 73
- email\_trend, 75
- external\_dist, 77
- external\_fizz, 79
- external\_line, 80
- external\_rank, 81
- external\_sum, 83
- hr\_trend, 99
- hrvar\_count, 95
- hrvar\_trend, 98
- keymetrics\_scan, 116
- meeting\_dist, 120
- meeting\_fizz, 122
- meeting\_line, 123
- meeting\_rank, 124
- meeting\_summary, 126
- meeting\_trend, 128
- one2one\_dist, 139
- one2one\_fizz, 141
- one2one\_freq, 142
- one2one\_line, 144
- one2one\_rank, 145
- one2one\_sum, 147
- one2one\_trend, 148
- \* **Working Patterns**
  - identify\_shifts, 109
- \* **datasets**
  - g2g\_data, 91
  - mt\_data, 130
  - p2p\_data, 150
  - pq\_data, 153
- \* **max-min**
  - maxmin, 119
- afterhours\_dist, 4, 7, 8, 10–13, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- afterhours\_fizz, 5, 6, 6, 8, 10–13, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- afterhours\_line, 5–7, 7, 10–13, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- afterhours\_rank, 5–8, 9, 11–13, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- afterhours\_sum (afterhours\_summary), 10
- afterhours\_summary, 5–8, 10, 10, 12, 13, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- afterhours\_trend, 5–8, 10, 11, 12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- analysis\_scope (hrvar\_count), 95
- anonymise, 13
- anonymize (anonymise), 13
- any\_idate, 14, 15, 16, 67, 85, 86, 95, 113, 120, 153, 158, 159, 165, 166, 168, 169, 171
- camel\_clean, 14, 15, 16, 67, 85, 86, 95, 113, 120, 153, 158, 159, 165, 166, 168, 169, 171
- check\_inputs, 14, 15, 15, 67, 85, 86, 95, 113, 120, 153, 158, 159, 165, 166, 168, 169, 171
- check\_query, 16, 86–90, 96–98, 100, 101, 104, 105, 107–109, 111, 168, 171
- collab\_area (collaboration\_area), 18
- collab\_dist (collaboration\_dist), 19

- collab\_fizz (collaboration\_fizz), 21
- collab\_line (collaboration\_line), 22
- collab\_rank (collaboration\_rank), 24
- collab\_sum (collaboration\_sum), 26
- collab\_summary (collaboration\_sum), 26
- collaboration\_area, 5, 7, 8, 10–12, 18, 20–23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_dist, 5, 7, 8, 10–12, 19, 19, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_fizz, 5, 7, 8, 10–12, 19–21, 21, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_line, 5, 7, 8, 10–12, 19–22, 22, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_rank, 5, 7, 8, 10–12, 19–23, 24, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_sum, 5, 7, 8, 10–12, 19–23, 25, 26, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- collaboration\_summary (collaboration\_sum), 26
- collaboration\_trend, 5, 7, 8, 10–12, 19–23, 25, 27, 27, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- comma, 29
- copy\_df, 29, 42, 77, 113
- create\_bar, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 30, 33, 35, 37, 38, 40, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- create\_bar\_asis, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 32, 32, 35, 37, 38, 40, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- create\_boxplot, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31–33, 34, 37, 38, 40, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- create\_bubble, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31–33, 35, 36, 38, 40, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- create\_density, 32, 33, 35, 37, 37, 40, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66
- create\_dist, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31–33, 35, 37, 38, 39, 43, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84, 96, 98, 99, 117, 121, 122, 124, 125, 127, 129, 140, 142, 143, 145, 147–149
- create\_dt, 30, 41, 77, 113
- create\_fizz, 5, 7, 8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31–33, 35, 37, 38, 40, 42, 44, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 79, 81, 82, 84,

- 96, 98, 99, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- create\_hist, 32, 33, 35, 37, 38, 40, 43, 43,  
46, 50, 52, 54, 56, 59, 61, 63, 65, 66
- create\_inc, 5, 7, 8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31–33, 35, 37, 38, 40, 43, 44,  
45, 50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73–75, 78, 79, 81, 82, 84, 96,  
98, 99, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- create\_incidence(create\_inc), 45
- create\_IV, 47, 114
- create\_line, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31–33, 35, 37, 38, 40, 43,  
44, 46, 49, 52, 54, 56, 59, 61, 63, 65,  
66, 69–71, 73–75, 78, 79, 81, 82, 84,  
96, 98, 99, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- create\_line(), 8, 81
- create\_line\_asis, 5, 7, 8, 10–12, 19, 20, 22,  
23, 25, 27, 28, 31–33, 35, 37, 38, 40,  
43, 44, 46, 50, 51, 54, 56, 59, 61, 63,  
65, 66, 69–71, 73–75, 78, 79, 81, 82,  
84, 96, 98, 99, 117, 121, 122, 124,  
125, 127, 129, 140, 142, 143, 145,  
147–149
- create\_lorenz, 52
- create\_period\_scatter, 5, 7, 8, 10–12, 19,  
20, 22, 23, 25, 27, 28, 31–33, 35, 37,  
38, 40, 43, 44, 46, 50, 52, 53, 56, 59,  
61, 63, 65, 66, 69–71, 73–75, 78, 79,  
81, 82, 84, 96, 98, 99, 117, 121, 122,  
124, 125, 127, 129, 140, 142, 143,  
145, 147–149
- create\_rank, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31–33, 35, 37, 38, 40, 43,  
44, 46, 50, 52, 54, 55, 59, 61, 63, 65,  
66, 69–71, 73–75, 78, 79, 81, 82, 84,  
96, 98, 99, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- create\_rank\_combine, 57
- create\_sankey, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31–33, 35, 37, 38, 40, 43,  
44, 46, 50, 52, 54, 56, 58, 61, 63, 65,  
66, 69–71, 73–75, 78, 79, 81, 82, 84,  
96, 98, 99, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- create\_scatter, 5, 7, 8, 10–12, 19, 20, 22,  
23, 25, 27, 28, 31–33, 35, 37, 38, 40,  
43, 44, 46, 50, 52, 54, 56, 59, 60, 63,  
65, 66, 69–71, 73–75, 78, 79, 81, 82,  
84, 96, 98, 99, 117, 121, 122, 124,  
125, 127, 129, 140, 142, 143, 145,  
147–149
- create\_stacked, 5, 7, 8, 10–12, 19, 20, 22,  
23, 25, 27, 28, 31–33, 35, 37, 38, 40,  
43, 44, 46, 50, 52, 54, 56, 59, 61, 61,  
65, 66, 69–71, 73–75, 78, 79, 81, 82,  
84, 96, 98, 99, 117, 121, 122, 124,  
125, 127, 129, 140, 142, 143, 145,  
147–149
- create\_tracking, 5, 7, 8, 10–12, 19, 20, 22,  
23, 25, 27, 28, 31–33, 35, 37, 38, 40,  
43, 44, 46, 50, 52, 54, 56, 59, 61, 63,  
64, 66, 69–71, 73–75, 78, 79, 81, 82,  
84, 96, 98, 99, 117, 121, 122, 124,  
125, 127, 129, 140, 142, 143, 145,  
147–149
- create\_trend, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31–33, 35, 37, 38, 40, 43,  
44, 46, 50, 52, 54, 56, 59, 61, 63, 65,  
65, 69–71, 73–75, 78, 79, 81, 82, 84,  
96, 98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- cut\_hour, 14–16, 67, 85, 86, 95, 113, 120,  
153, 158, 159, 165, 166, 168, 169,  
171
- email\_dist, 5, 7, 8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 68, 70,  
71, 73–75, 78, 79, 81, 82, 84, 96, 98,  
100, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- email\_fizz, 5, 7, 8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69, 69,  
71, 73–75, 78, 79, 81, 82, 84, 96, 98,  
100, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- email\_line, 5, 7, 8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,

- 52, 54, 56, 59, 61, 63, 65, 66, 69, 70,  
70, 73–75, 78, 79, 81, 82, 84, 96, 98,  
100, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- email\_rank, 5, 7, 8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
72, 74, 75, 78, 79, 81, 82, 84, 96, 98,  
100, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- email\_sum (email\_summary), 73
- email\_summary, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73, 73, 75, 78, 79, 81, 82, 84,  
96, 98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- email\_trend, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73, 74, 75, 78, 80–82, 84, 96,  
98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- export, 30, 42, 76, 113
- external\_dist, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73–75, 77, 80–82, 84, 96, 98,  
100, 117, 121, 122, 124, 125, 127,  
129, 140, 142, 143, 145, 147–149
- external\_fizz, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73–75, 78, 79, 81, 82, 84, 96,  
98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- external\_line, 5, 7, 8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73–75, 78, 80, 80, 82, 84, 96,  
98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- external\_rank, 5–8, 10–13, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,  
69–71, 73–75, 78, 80, 81, 81, 84, 96,  
98, 100, 117, 121, 122, 124, 125,  
127, 129, 140, 142, 143, 145,  
147–149
- external\_sum, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 83, 96, 98, 100,  
117, 121, 123–125, 127, 129, 140,  
142, 144, 145, 147–149
- external\_summary (external\_sum), 83
- extract\_date\_range, 14–16, 67, 84, 86, 95,  
113, 120, 153, 158, 159, 165, 166,  
168, 169, 171
- extract\_hr, 14–17, 67, 85, 85, 87–90, 95–98,  
100, 101, 104, 105, 107–109, 111,  
113, 120, 153, 158, 159, 165, 166,  
168, 169, 171
- flag\_ch\_ratio, 17, 86, 86, 88–90, 96–98,  
100, 101, 104, 105, 107–109, 111,  
168, 171
- flag\_em\_ratio, 17, 86, 87, 87, 89, 90, 96–98,  
100, 101, 104, 105, 107–109, 111,  
168, 171
- flag\_extreme, 17, 86–88, 88, 90, 96–98, 100,  
101, 104, 105, 107–109, 111, 168,  
171
- flag\_outlooktime, 17, 86–89, 90, 96–98,  
100, 101, 104, 105, 107–109, 111,  
168, 171
- g2g\_data, 91, 131, 133, 137, 139, 150, 151,  
157
- generate\_report, 92, 114, 128, 158, 171
- generate\_report2, 93
- heat\_colors (heat\_colours), 94
- heat\_colours, 14–16, 67, 85, 86, 94, 113,  
120, 153, 158, 159, 165, 166, 168,  
169, 171
- hr\_trend, 6–8, 10–12, 17, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 86–90, 96–98,  
99, 101, 104, 105, 107–109, 111,  
117, 121, 123–125, 127, 129, 140,  
142, 144, 145, 147–149, 168, 171

- hrvar\_count*, 6–8, 10–12, 17, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 86–90, 95, 97, 98, 100, 101, 104, 105, 107–109, 111, 117, 121, 123–125, 127, 129, 140, 142, 144, 145, 147–149, 168, 171  
*hrvar\_count\_all*, 17, 86–90, 96, 96, 98, 100, 101, 104, 105, 107–109, 111, 168, 171  
*hrvar\_trend*, 6–8, 10–12, 17, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 86–90, 96, 97, 98, 100, 101, 104, 105, 107–109, 111, 117, 121, 123–125, 127, 129, 140, 142, 144, 145, 147–149, 168, 171  
*identify\_churn*, 17, 86–90, 96–98, 100, 100, 104, 105, 107–109, 111, 168, 171  
*identify\_datefreq*, 102  
*identify\_holidayweeks*, 17, 86–90, 96–98, 100, 101, 103, 105, 107–109, 111, 168, 171  
*identify\_inactiveweeks*, 17, 86–90, 96–98, 100, 101, 104, 105, 107–109, 111, 168, 171  
*identify\_nkw*, 17, 86–90, 96–98, 100, 101, 104, 105, 106, 107–109, 111, 168, 171  
*identify\_outlier*, 17, 86–90, 96–98, 100, 101, 104, 105, 107, 107, 108, 109, 111, 168, 171  
*identify\_privacythreshold*, 17, 86–90, 96, 97, 99–101, 104, 105, 107, 108, 109, 111, 168, 171  
*identify\_shifts*, 17, 86–90, 96, 97, 99–101, 104, 105, 107, 108, 109, 111, 168, 171  
*identify\_tenure*, 17, 86–90, 96, 97, 99–101, 104, 105, 107–109, 110, 168, 171  
*import\_query*, 30, 42, 77, 112  
*is\_date\_format*, 14–16, 67, 85, 86, 95, 113, 120, 153, 158, 159, 165, 166, 168, 169, 171  
*IV\_report*, 48, 93, 113, 128, 158, 171  
*jitter\_metrics*, 115  
*keymetrics\_scan*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 116, 121, 123–125, 127, 129, 140, 142, 144, 145, 147–149  
*keymetrics\_scan\_asis*, 117  
*maxmin*, 14–16, 67, 85, 86, 95, 113, 119, 153, 158, 159, 165, 166, 168, 169, 171  
*meeting\_dist*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 117, 120, 123–129, 140, 142, 144, 145, 147–149  
*meeting\_fizz*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 117, 121, 122, 124–129, 140, 142, 144, 145, 147–149  
*meeting\_line*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 117, 121, 123, 123, 125–129, 140, 142, 144, 145, 147–149  
*meeting\_rank*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 117, 121, 123, 124, 124, 127–129, 140, 142, 144, 145, 147–149  
*meeting\_sum* (*meeting\_summary*), 126  
*meeting\_summary*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66, 69–71, 73–75, 78, 80–82, 84, 96, 98, 100, 117, 121, 123–126, 126, 128, 129, 140, 142, 144, 145, 147–149  
*meeting\_tm\_report*, 93, 114, 121, 123, 124, 126, 127, 127, 129, 153, 158, 160, 162–164, 171  
*meeting\_trend*, 6–8, 10–12, 19, 20, 22, 23, 25, 27, 28, 31, 33, 35, 37, 40, 43, 46, 50, 52, 54, 56, 59, 61, 63, 65, 66,

- 69–71, 73–75, 78, 80–82, 84, 96, 98,  
100, 117, 121, 123–128, 128, 140,  
142, 144, 145, 147–149
- mt\_data, 92, 130, 150, 151, 157
- network\_g2g, 92, 131, 137, 139, 150, 151
- network\_p2p, 92, 133, 134, 139, 150, 151
- network\_summary, 92, 133, 137, 138, 150, 151
- one2one\_dist, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129, 139,  
142, 144, 145, 147–149
- one2one\_fizz, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129, 140,  
141, 141, 144, 145, 147–149
- one2one\_freq, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129,  
140–142, 142, 145, 147–149
- one2one\_line, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129,  
140–142, 144, 144, 147–149
- one2one\_rank, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129,  
140–142, 144, 145, 145, 148, 149
- one2one\_sum, 6–8, 10–12, 19, 20, 22, 23, 25,  
27, 28, 31, 33, 35, 37, 40, 43, 46, 50,  
52, 54, 56, 59, 61, 63, 65, 66, 69–71,  
73–75, 78, 80–82, 84, 96, 98, 100,  
117, 121, 123–125, 127, 129,  
140–142, 144, 145, 147, 147, 149
- one2one\_summary (one2one\_sum), 147
- one2one\_trend, 6–8, 10–12, 19, 20, 22, 23,  
25, 27, 28, 31, 33, 35, 37, 40, 43, 46,  
50, 52, 54, 56, 59, 61, 63, 65, 66,
- 69–71, 73–75, 78, 80–82, 84, 96, 98,  
100, 117, 121, 123–125, 127, 129,  
140–142, 144, 145, 147, 148, 148
- p2p\_data, 92, 131, 133, 137, 139, 150, 151,  
157
- p2p\_data\_sim, 92, 131, 133, 137, 139, 150,  
151, 157
- pad2, 152
- pairwise\_count, 14–16, 67, 85, 86, 95, 113,  
120, 128, 152, 158–160, 162–166,  
168, 169, 171
- pq\_data, 92, 131, 150, 151, 153
- read\_preamble, 14–16, 67, 85, 86, 93, 95,  
113, 114, 120, 128, 153, 158, 159,  
165, 166, 168, 169, 171
- rgb2hex, 14–16, 67, 85, 86, 95, 113, 120, 153,  
158, 158, 165, 166, 168, 169, 171
- theme\_wpa, 159, 160
- theme\_wpa\_basic, 159, 159
- tm\_clean, 128, 153, 160, 162–164
- tm\_cooc, 128, 153, 160, 161, 163, 164
- tm\_freq, 128, 153, 160, 162, 162, 164
- tm\_wordcloud, 128, 153, 160, 162, 163, 163
- totals\_bind, 14–16, 67, 85, 86, 95, 113, 120,  
153, 158, 159, 165, 166, 168, 169,  
171
- totals\_col, 14–16, 67, 85, 86, 95, 113, 120,  
153, 158, 159, 165, 166, 168, 169,  
171
- track\_HR\_change, 17, 86–90, 96, 97, 99–101,  
104, 105, 107–109, 111, 167, 171
- tstamp, 14–16, 67, 85, 86, 95, 113, 120, 153,  
158, 159, 165, 166, 168, 169, 171
- us\_to\_space, 14–16, 67, 85, 86, 95, 113, 120,  
153, 158, 159, 165, 166, 168, 169,  
171
- validation\_report, 17, 86–90, 93, 96, 97,  
99–101, 104, 105, 107–109, 111,  
114, 128, 158, 168, 169
- wrap, 14–16, 67, 85, 86, 95, 113, 120, 153,  
158, 159, 165, 166, 168, 169, 171
- wrap\_text, 172