Package ‘tangram.pipe’

October 12, 2021

Title Row-by-Row Table Building
Version 1.1.0
Description Builds tables with customizable rows. Users can specify the type of data to use for each row, as well as how to handle missing data and the types of comparison tests to run on the table columns.
License MIT + file LICENSE
Encoding UTF-8
RoxygenNote 7.1.1.9001
Depends dplyr
Suggests knitr, kableExtra, rmarkdown
VignetteBuilder knitr
URL https://github.com/thomasgstewart/tangram.pipe
BugReports https://github.com/thomasgstewart/tangram.pipe/issues
NeedsCompilation no
Author Andrew Guide [aut, cre], Thomas Stewart [aut]
Maintainer Andrew Guide <andrew.guide@vumc.org>
Repository CRAN
Date/Publication 2021-10-11 22:00:02 UTC

R topics documented:

  binary_default .................................................. 2
  binary_diff .................................................... 3
  binary_jama ..................................................... 3
  binary_or ........................................................ 4
  binary_pct ....................................................... 5
  binary_row ....................................................... 5
  binary_rr ........................................................ 7
binary_default

Description
Summarizes a binary row using counts and column proportions.

Usage
binary_default(dt, ...)

Arguments
dt the name of the dataframe object.
... Additional arguments supplied within the package row functions.

Details
This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.
reference: the name of the row category to use as the reference. Default will use alphabetical first category
rowlabel: the label for the table row name, if different from row_var.
compact: if TRUE, data displayed in one row.
missing: if TRUE, missing data is considered; FALSE only uses complete cases.
digits: significant digits to use.
**binary_diff**

**Value**

A dataframe with summary statistics for a binary variable.

---

**binary_diff**  
*Binary Difference in Proportions*

**Description**

Default comparison function for binary data.

**Usage**

binary_diff(dt, num_col, reference, digits)

**Arguments**

dt: the name of the dataframe object.
num_col: the number of categorical columns in the data.
reference: the name of the reference row category to use.
digits: significant digits to use.

**Value**

A dataframe with difference in proportions test results between pairs of columns for binary data, as well as an overall chi-squared test across all groups.

---

**binary_jama**  
*JAMA-style summary for a Binary Row*

**Description**

Summarizes a binary row using column percentages and the total number in each cell divided by the column total. This is the style used by the Journal of the American Medical Association.

**Usage**

binary_jama(dt, ...)

**Arguments**

dt: the name of the dataframe object.
... Additional arguments supplied within the package row functions.
Details

This is an internal function of `tangram.pipe`. Additional arguments should be supplied for this function to work properly.

- `reference`: the name of the row category to use as the reference. Default will use alphabetical first category.
- `rowlabel`: the label for the table row name, if different from `row_var`.
- `compact`: if TRUE, data displayed in one row.
- `missing`: if TRUE, missing data is considered; FALSE only uses complete cases.
- `digits`: significant digits to use.

Value

A dataframe with summary statistics for a binary variable.

```
  binary_or         Binary Odds Ratio
```

Description

Calculates odds ratio across categories for binary data.

Usage

`binary_or(dt, num_col, reference, digits)`

Arguments

- `dt`: the name of the dataframe object.
- `num_col`: the number of categorical columns in the data.
- `reference`: the name of the reference row category to use.
- `digits`: significant digits to use.

Value

A dataframe with computed odds ratios between pairs of columns for binary data, as well as an overall chi-squared test across all groups.
**binary_pct**

*Percentage summary for a Binary Row*

---

**Description**

Summarizes a binary row using counts and column percentages.

**Usage**

`binary_pct(dt, ...)`

**Arguments**

- **dt**: the name of the dataframe object.
- **...**: Additional arguments supplied within the package row functions.

**Details**

This is an internal function of `tangram.pipe`. Additional arguments should be supplied for this function to work properly.

- **reference**: the name of the row category to use as the reference. Default will use alphabetical first category.
- **rowlabel**: the label for the table row name, if different from row_var.
- **compact**: if TRUE, data displayed in one row.
- **missing**: if TRUE, missing data is considered; FALSE only uses complete cases.
- **digits**: significant digits to use.

**Value**

A dataframe with summary statistics for a binary variable.

---

**binary_row**

*Binary Row*

---

**Description**

Adds in a binary row to the table.
Usage

```r
binary_row(
  list_obj,
  row_var,
  col_var = NULL,
  newdata = FALSE,
  rowlabel = NULL,
  summary = NULL,
  reference = NULL,
  compact = TRUE,
  missing = NULL,
  overall = NULL,
  comparison = NULL,
  digits = 2,
  indent = 5
)
```

Arguments

- **list_obj** the name of the tbl_start object previously initialized.
- **row_var** the name of the variable to be used in the rows.
- **col_var** the variable to be used in the table columns. Default is from initialized tbl_start object.
- **newdata** enter new dataset name if different from that initialized in tbl_start.
- **rowlabel** the label for the table row name, if different from row_var.
- **summary** summary function for the data, if different from the one supplied in tbl_start.
- **reference** the name of the row category to use as the reference. Default will use alphabetical first category.
- **compact** logical: if TRUE, data displayed in one row.
- **missing** logical: if TRUE, missing data is considered; FALSE only uses complete cases.
- **overall** logical: if TRUE, an overall column is included.
- **comparison** the name of the comparison test to use, if different from that initialized in tbl_start.
- **digits** significant digits to use.
- **indent** number of spaces to indent category names.

Value

A list with the binary row’s table information added as a new element to list_obj.

Examples

```r
iris$color <- sample(c("Blue", "Purple"), size=150, replace=TRUE)
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
  binary_row("color", rowlabel="Color")
```
### Binary Risk Ratio

**Description**

 Calculates risk ratio across categories for binary data.

**Usage**

```r
call región(rr, dt, num_col, reference, digits)
```

**Arguments**

- `dt` : the name of the dataframe object.
- `num_col` : the number of categorical columns in the data.
- `reference` : the name of the reference row category to use.
- `digits` : significant digits to use.

**Value**

A dataframe with computed risk ratios between pairs of columns for binary data, as well as an overall chi-squared test across all groups.

### Chi-Squared Test for Categorical Variables

**Description**

Default comparison function for categorical data.

**Usage**

```r
cat_comp_default(dt, digits)
```

**Arguments**

- `dt` : the name of the dataframe object.
- `digits` : significant digits to use.

**Value**

A dataframe calculating relative entropy between column pairs, as well as an overall chi-squared test across all groups.
**cat_default**

*Default summary for a Categorical Row*

**Description**
Summarizes a categorical row using counts and column proportions.

**Usage**
cat_default(dt, ...)

**Arguments**

- **dt**: the name of the dataframe object.
- **...**: Additional arguments supplied within the package row functions.

**Details**
This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.

- **rowlabel**: the label for the table row name, if different from row_var.
- **missing**: if TRUE, missing data is considered; FALSE only uses complete cases.
- **digits**: significant digits to use.

**Value**
A dataframe with summary statistics for a categorical variable.

---

**cat_jama**

*JAMA-style summary for a Categorical Row*

**Description**
Summarizes a categorical row using column percentages and the total number in each cell divided by the column total. This is the style used by the Journal of the American Medical Association.

**Usage**
cat_jama(dt, ...)

**Arguments**

- **dt**: the name of the dataframe object.
- **...**: Additional arguments supplied within the package row functions.
Details

This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.

rowlabel : the label for the table row name, if different from row_var.
missing : if TRUE, missing data is considered; FALSE only uses complete cases.
digits : significant digits to use.

Value

A dataframe with summary statistics for a categorical variable.

---

cat_pct | Percentage summary for a Categorical Row

Description

Summarizes a categorical row using counts and column percentages.

Usage

```r
cat_pct(dt, ...)
```

Arguments

dt the name of the dataframe object.
... Additional arguments supplied within the package row functions.

Details

This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.

rowlabel : the label for the table row name, if different from row_var.
missing : if TRUE, missing data is considered; FALSE only uses complete cases.
digits : significant digits to use.

Value

A dataframe with summary statistics for a categorical variable.
Description

Adds in a categorical row to the table.

Usage

cat_row(
  list_obj,
  row_var,
  col_var = NULL,
  newdata = FALSE,
  rowlabel = NULL,
  summary = NULL,
  missing = NULL,
  overall = NULL,
  comparison = NULL,
  digits = 2,
  indent = 5
)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>list_obj</td>
<td>the name of the tbl_start object previously initialized.</td>
</tr>
<tr>
<td>row_var</td>
<td>the name of the variable to be used in the rows.</td>
</tr>
<tr>
<td>col_var</td>
<td>the variable to be used in the table columns. Default is from initialized tbl_start object.</td>
</tr>
<tr>
<td>newdata</td>
<td>enter new dataset name if different from that initialized in tbl_start.</td>
</tr>
<tr>
<td>rowlabel</td>
<td>the label for the table row name, if different from row_var.</td>
</tr>
<tr>
<td>summary</td>
<td>summary function for the data, if different from the one supplied in tbl_start.</td>
</tr>
<tr>
<td>missing</td>
<td>logical: if TRUE, missing data is considered; FALSE only uses complete cases.</td>
</tr>
<tr>
<td>overall</td>
<td>logical: if TRUE, an overall column is included.</td>
</tr>
<tr>
<td>comparison</td>
<td>the name of the comparison test to use, if different from that initialized in tbl_start.</td>
</tr>
<tr>
<td>digits</td>
<td>significant digits to use.</td>
</tr>
<tr>
<td>indent</td>
<td>number of spaces to indent category names.</td>
</tr>
</tbody>
</table>

Value

A list with the categorical row’s table information added as a new element to list_obj.
empty_row

Examples

```r
iris$Stem.Size <- sample(c("Small", "Medium", "Medium", "Large"), size=150, replace=TRUE)
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
cat_row("Stem.Size", rowlabel="Stem Size")
```

<table>
<thead>
<tr>
<th>empty_row</th>
<th>Empty Row</th>
</tr>
</thead>
</table>

Description

Produces a empty dividing row in the table. May have a row header.

Usage

```r
empty_row(list_obj, header = NULL)
```

Arguments

- `list_obj` the name of the tbl_start object previously initialized.
- `header` a header to include for the empty row.

Value

If a header is included, a list object is returned with a one-element dataframe containing the header as the most recent entry to `list_obj`. Otherwise, a list is returned containing a blank character as the last element of `list_obj`.

---

num_default

| Default summary for a Numeric Row |

Description

Summarizes a numeric row using the five-number summary, mean, and standard deviation.

Usage

```r
num_default(dt, ...)
```

Arguments

- `dt` the name of the dataframe object.
- `...` Additional arguments supplied within the package row functions.
Details
This is an internal function of `tangram.pipe`. Additional arguments should be supplied for this function to work properly.

`rowlabel`: the label for the table row name, if different from `row_var`.

`missing`: if TRUE, missing data is considered; FALSE only uses complete cases.

`digits`: significant digits to use.

Value
A dataframe with summary statistics for a numeric variable.

<table>
<thead>
<tr>
<th><code>num_diff</code></th>
<th>Numeric Difference in Means</th>
</tr>
</thead>
</table>

Description
Default comparison function for numeric data.

Usage

```
num_diff(dt, num_col, row_var, digits)
```

Arguments

- `dt`: the name of the dataframe object.
- `num_col`: the number of categorical columns in the data.
- `row_var`: the name of the row variable in the data.
- `digits`: significant digits to use.

Value
A dataframe calculating the difference in means between column pairs, as well as an overall one-way ANOVA across all groups.
**num_mean_sd**

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarizes a numeric row using the mean and standard deviation.</td>
</tr>
</tbody>
</table>

**Usage**

```
num_mean_sd(dt, ...)  
```

**Arguments**

- `dt` the name of the dataframe object.
- `...` Additional arguments supplied within the package row functions.

**Details**

This is an internal function of `tangram.pipe`. Additional arguments should be supplied for this function to work properly.

- `rowlabel` : the label for the table row name, if different from `row_var`.
- `missing` : if TRUE, missing data is considered; FALSE only uses complete cases.
- `digits` : significant digits to use.

**Value**

A dataframe with summary statistics for a numeric variable.

---

**num_medianiqr**

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarizes a numeric row using the median and interquartile range.</td>
</tr>
</tbody>
</table>

**Usage**

```
num_medianiqr(dt, ...)  
```

**Arguments**

- `dt` the name of the dataframe object.
- `...` Additional arguments supplied within the package row functions.
Details
This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.
rowlabel: the label for the table row name, if different from row_var.
missing: if TRUE, missing data is considered; FALSE only uses complete cases.
digits: significant digits to use.

Value
A dataframe with summary statistics for a numeric variable.

num_minmax Min-Max summary for a Numeric Row

Description
Summarizes a numeric row using the minimum and maximum values.

Usage
num_minmax(dt, ...)

Arguments
dt the name of the dataframe object.
... Additional arguments supplied within the package row functions.

Details
This is an internal function of tangram.pipe. Additional arguments should be supplied for this function to work properly.
rowlabel: the label for the table row name, if different from row_var.
missing: if TRUE, missing data is considered; FALSE only uses complete cases.
digits: significant digits to use.

Value
A dataframe with summary statistics for a numeric variable.
**num_row**

### Description

Adds in a numeric row to the table.

### Usage

```r
num_row(
  list_obj,
  row_var,
  col_var = NULL,
  newdata = FALSE,
  rowlabel = NULL,
  summary = NULL,
  missing = NULL,
  overall = NULL,
  comparison = NULL,
  digits = 2
)
```

### Arguments

- **list_obj**: the name of the tbl_start object previously initialized.
- **row_var**: the name of the variable to be used in the rows.
- **col_var**: the variable to be used in the table columns. Default is from initialized tbl_start object.
- **newdata**: enter new dataset name if different from that initialized in tbl_start.
- **rowlabel**: the label for the table row name, if different from row_var.
- **summary**: summary function for the data, if different from the one supplied in tbl_start.
- **missing**: logical: if TRUE, missing data is considered; FALSE only uses complete cases.
- **overall**: logical: if TRUE, an overall column is included.
- **comparison**: the name of the comparison test to use, if different from that initialized in tbl_start.
- **digits**: significant digits to use.

### Value

A list with the numeric row’s table information added as a new element to list_obj.

### Examples

```r
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
  num_row("Sepal.Length", rowlabel="Sepal Length")
```
Description

Counts the instances of each column variable of the dataframe to be used in the table (if applicable), and gives an overall row count.

Usage

```r
def n_row(
  list_obj, 
  col_var = NULL, 
  newdata = FALSE, 
  missing = NULL, 
  overall = NULL 
)
```

Arguments

- **list_obj**: the name of the tbl_start object previously initialized.
- **col_var**: the variable to be used in the table columns. Default is from initialized tbl_start object.
- **newdata**: enter new dataset name if different from that initialized in tbl_start.
- **missing**: logical: if TRUE, missing data in the column variable is considered; FALSE only uses complete cases.
- **overall**: logical: if TRUE, an overall column is included.

Value

A list with the row counts added as a new element to list_obj.

Examples

```r
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
n_row()
```
print.tangram.pipe

**Printing a Table**

**Description**

Prints a finished table created from tangram.pipe.

**Usage**

```r
## S3 method for class 'tangram.pipe'
print(x, ...)
```

**Arguments**

- `x` the name of the tbl_start object previously initialized.
- `...` further arguments passed to or from other methods.

**Value**

A dataframe object containing the information from the last element of a tangram.pipe class object created using `tbl_out()`. This is the finalized table object.

**Examples**

```r
iris$color <- sample(c("Blue", "Purple"), size=150, replace=TRUE)
iris$Stem.Size <- sample(c("Small", "Medium", "Medium", "Large"), size=150, replace=TRUE)
iris$Leaf.Color <- "Green"
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
  num_row("Sepal.Length", rowlabel="Sepal Length") %>%
  empty_row() %>%
  num_row("Sepal.Width", rowlabel="Sepal Width") %>%
  empty_row() %>%
  num_row("Petal.Length", rowlabel="Petal Length") %>%
  empty_row() %>%
  num_row("Petal.Width", rowlabel="Petal Width") %>%
  empty_row() %>%
  cat_row("Stem.Size", rowlabel="Stem Size") %>%
  empty_row() %>%
  binary_row("color", rowlabel="Color") %>%
  tbl_out() %>%
  print()
```
### tangram_styling

**Tangram Styling**

**Description**

Used to preprocess a tangram.pipe table for HTML formatting.

**Usage**

`tangram_styling(df)`

**Arguments**

- **df**
  
  The output data frame object to be printed in HTML form.

**Value**

A dataframe containing HTML formatting code where applicable.

---

### tbl_out

**Output Table**

**Description**

Produces a finalized table.

**Usage**

`tbl_out(list_obj)`

**Arguments**

- **list_obj**

  the name of the tbl_start object previously initialized.

**Value**

A tangram.pipe class object with the finalized table as a dataframe added as the most recent element of `list_obj`. 
Examples
iris$color <- sample(c("Blue", "Purple"), size=150, replace=TRUE)
iris$Stem.Size <- sample(c("Small", "Medium", "Medium", "Large"), size=150, replace=TRUE)
iris$Leaf.Color <- "Green"
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE) %>%
  num_row("Sepal.Length", rowlabel="Sepal Length") %>%
  empty_row() %>%
  num_row("Sepal.Width", rowlabel="Sepal Width") %>%
  empty_row() %>%
  num_row("Petal.Length", rowlabel="Petal Length") %>%
  empty_row() %>%
  num_row("Petal.Width", rowlabel="Petal Width") %>%
  empty_row() %>%
  cat_row("Stem.Size", rowlabel="Stem Size") %>%
  empty_row() %>%
  binary_row("color", rowlabel="Color") %>%
  tbl_out()

Table Initialization

Description
Initializes the table by specifying the desired elements and data components.

Usage
tbl_start(
data,
  col_var,
  missing = FALSE,
  overall = TRUE,
  comparison = FALSE,
  default_num_summary = num_default,
  default_cat_summary = cat_default,
  default_binary_summary = binary_default
)

Arguments
data The dataset to be used in the table.
col_var The variable to be used in the table columns. NULL if single summary column desired.
missing logical: if TRUE, missing data is considered; FALSE only uses complete cases.
overall logical: if TRUE, an overall column is included.
comparison logical: if TRUE, a comparison test is conducted between columns.
default_num_summary
The default summary function to use for numerical rows. By default, the package will use num_default(), but the user can also choose num_minmax, num_medianiqr, num_mean_sd, or write a custom function to use for the rows.

default_cat_summary
The default summary function to use for categorical rows. By default, the package will use cat_default(), but the user can also choose cat_pct or write a custom function to use for the rows.

default_binary_summary
The default summary function to use for binary rows. By default, the package will use binary_default(), but the user can also choose binary_pct or write a custom function to use for the rows.

Value
A list containing separate entries holding information provided in the function’s arguments, as well as a calculated number of column categories to include for the initialized table.

Examples
x <- tbl_start(iris, "Species", missing=TRUE, overall=TRUE, comparison=TRUE)
Index

* `tangram.pipe`
  - `binary_default`, 2
  - `binary_diff`, 3
  - `binary_jama`, 3
  - `binary_or`, 4
  - `binary_pct`, 5
  - `binary_row`, 5
  - `binary_rr`, 7
  - `cat_default`, 8
  - `cat_jama`, 8
  - `cat_pct`, 9
  - `cat_row`, 10
  - `empty_row`, 11
  - `n_row`, 16
  - `num_default`, 11
  - `num_diff`, 12
  - `num_mean_sd`, 13
  - `num_median_iqr`, 13
  - `num_minmax`, 14
  - `num_row`, 15
  - `print.tangram.pipe`, 17
  - `tangram_styling`, 18
  - `tbl_out`, 18
  - `tbl_start`, 19

  - `binary_default`, 2
  - `binary_diff`, 3
  - `binary_jama`, 3
  - `binary_or`, 4
  - `binary_pct`, 5
  - `binary_row`, 5
  - `binary_rr`, 7
  - `cat_comp_default`, 7
  - `cat_default`, 8
  - `cat_jama`, 8
  - `cat_pct`, 9
  - `cat_row`, 10
  - `empty_row`, 11
  - `n_row`, 16

  num_default, 11
  num_diff, 12
  num_mean_sd, 13
  num_median_iqr, 13
  num_minmax, 14
  num_row, 15
  print.tangram.pipe, 17
  tangram_styling, 18
  tbl_out, 18
  tbl_start, 19