Package ‘rfm’

April 23, 2020

Type Package

Title Recency, Frequency and Monetary Value Analysis

Version 0.2.1

Description Tools for RFM (recency, frequency and monetary value) analysis.
Generate RFM score from both transaction and customer level data. Visualize the relationship between recency, frequency and monetary value using heatmap, histograms, bar charts and scatter plots. Includes a ‘shiny’ app for interactive segmentation. References:

License MIT + file LICENSE

URL https://github.com/rsquaredacademy/rfm,
    https://rfm.rsquaredacademy.com/

BugReports https://github.com/rsquaredacademy/rfm/issues

Depends R (>= 3.2)

Imports dplyr, ggplot2, ggthemes, lubridate, magrittr, RColorBrewer, rlang, stats, tibble, tidyr, utils, xplorerr

Suggests covr, DT, kableExtra, knitr, rmarkdown, testthat, vdiffr

VignetteBuilder knitr

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Author Aravind Hebbali [aut, cre] (<https://orcid.org/0000-0001-9220-9669>)

Maintainer Aravind Hebbali <hebbali.aravind@gmail.com>

Repository CRAN

Date/Publication 2020-04-23 12:50:02 UTC
R topics documented:

- rfm ................................................................. 2
- rfm_barchart_data ............................................ 2
- rfm_bar_chart .................................................. 3
- rfm_data_customer ............................................ 4
- rfm_data_orders .............................................. 4
- rfm_heatmap .................................................. 5
- rfm_heatmap_data ............................................ 6
- rfm_histograms .............................................. 7
- rfm_hist_data ................................................ 8
- rfm_launch_app ............................................... 9
- rfm_order_dist ............................................... 9
- rfm_plot_median_recency .................................... 10
- rfm_rm_plot .................................................. 11
- rfm_segment .................................................. 12
- rfm_table_customer ......................................... 13
- rfm_table_customer_2 ....................................... 15
- rfm_table_order ............................................. 16

Index 18

---

rfm rfm package

description

Tools for customer segmentation analysis

---

rfm_barchart_data Bar chart data

description

Data for generating bar charts.

usage

rfm_barchart_data(rfm_table)

arguments

rfm_table An object of class rfm_table.
**rfm_bar_chart**

**Examples**

```r
# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date,
revenue, analysis_date)

# bar chart data
rfm_barchart_data(rfm_order)

# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id,
number_of_orders, recency_days, revenue, analysis_date)

# bar chart data
rfm_barchart_data(rfm_customer)
```

---

**rfm_bar_chart**  
**RFM bar chart**

**Description**

Examine the distribution of monetary scores for the different combinations of frequency and recency scores.

**Usage**

```r
rfm_bar_chart(rfm_table, bar_color = "blue",
  xaxis_title = "Monetary Score", sec_xaxis_title = "Frequency Score",
  yaxis_title = "", sec_yaxis_title = "Recency Score",
  print_plot = TRUE)
```

**Arguments**

- **rfm_table**: An object of class rfm_table.
- **bar_color**: Color of the bars.
- **xaxis_title**: X axis title.
- **sec_xaxis_title**: Secondary x axis title.
- **yaxis_title**: Y axis title.
- **sec_yaxis_title**: Secondary y axis title.
- **print_plot**: logical; if TRUE, prints the plot else returns a plot object.

**Value**

Bar chart.
Examples

# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# bar chart
rfm_bar_chart(rfm_order)

rfm_data_customer  
**RFM customer data**

**Description**

A dataset containing customer level data.

**Usage**

rfm_data_customer

**Format**

A tibble with 39,999 rows and 5 variables:

- **customer_id**  Customer id.
- **total_amount**  Total amount of all orders.
- **most_recent_visit**  Date of the most recent transaction.
- **number_of_purchases**  Total number of transactions/orders.
- **purchase_interval**  Number of days since last transaction/order.

rfm_data_orders  
**RFM transaction data**

**Description**

A dataset containing transactions of different customers.

**Usage**

rfm_data_orders
rfm_heatmap

Format

A tibble with 49.6 rows and 3 variables:

order_date  order date
customer_id  customer id
revenue     transaction amount

Description

The heat map shows the average monetary value for different categories of recency and frequency scores. Higher scores of frequency and recency are characterized by higher average monetary value as indicated by the darker areas in the heatmap.

Usage

rfm_heatmap(data, plot_title = "RFM Heat Map",
            plot_title_justify = 0.5, xaxis_title = "Frequency",
            yaxis_title = "Recency", legend_title = "Mean Monetary Value",
            brewer_n = 5, brewer_name = "PuBu", print_plot = TRUE)

Arguments

data                An object of class rfm_table.
plot_title          Title of the plot.
plot_title_justify  Horizontal justification of the plot title; 0 for left justified and 1 for right justified.
xaxis_title         X axis title.
yaxis_title         Y axis title.
legend_title        Legend title.
brewer_n            Indicates the number of colors in the palette; RColorBrewer is used for the color palette of the heatmap; check the documentation of brewer_pal.
brewer_name         Palette name; check the documentation of brewer_pal.
print_plot          logical; if TRUE, prints the plot else returns a plot object.
Examples

# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date,
revenue, analysis_date)

# heat map
df <- rfm_table(rfm_order)
rfm_heatmap(df)

# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id,
number_of_orders, recency_days, revenue, analysis_date)

# heat map
df <- rfm_table(rfm_customer)
rfm_heatmap(df)

rfm_heatmap_data  Heatmap data

Description

Data for generating heatmap.

Usage

rfm_heatmap_data(rfm_table)

Arguments

rfm_table  An object of class rfm_table.

Examples

# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date,
revenue, analysis_date)

# heat map data
df <- rfm_table(rfm_order)
rfm_heatmap_data(df)

# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id,
number_of_orders, recency_days, revenue, analysis_date)

# heat map data
df <- rfm_table(rfm_customer)
rfm_heatmap_data(df)
**rfm_histograms**

**RFM histograms**

---

**Description**

Histograms of recency, frequency and monetary value.

**Usage**

```r
rfm_histograms(rfm_table, hist_bins = 9, hist_color = "blue",
plot_title = "RFM Histograms", xaxis_title = "",
yaxis_title = "Count", hist_m_label = "Monetary",
hist_r_label = "Recency", hist_f_label = "Frequency",
plot_title_justify = 0.5, print_plot = TRUE)
```

**Arguments**

- `rfm_table` An object of class `rfm_table`.
- `hist_bins` Number of bins of the histograms.
- `hist_color` Color of the histogram.
- `plot_title` Title of the plot.
- `xaxis_title` X axis title.
- `yaxis_title` Y axis title.
- `hist_m_label` Label of the monetary value histogram.
- `hist_r_label` Label of the recency histogram.
- `hist_f_label` Label of the frequency histogram.
- `plot_title_justify` Horizontal justification of the plot title; 0 for left justified and 1 for right justified.
- `print_plot` logical; if TRUE, prints the plot else returns a plot object.

**Value**

Histograms

**Examples**

```r
# using transaction data
analysis_date <- lubridate::as_date("2006-12-31", tz = "UTC")
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# histogram
rfm_histograms(rfm_order)
```
# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id, number_of_orders, recency_days, revenue, analysis_date)

# histogram
rfm_histograms(rfm_customer)

---

### rfm_hist_data

**Histogram data**

*Description*

Data for generating histograms.

*Usage*

```r
rfm_hist_data(rfm_table)
```

*Arguments*

- `rfm_table`: An object of class `rfm_table`.

*Examples*

```r
# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# histogram data
rfm_hist_data(rfm_order)

# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id, number_of_orders, recency_days, revenue, analysis_date)

# histogram data
rfm_hist_data(rfm_customer)
```
Launch shiny app

Description
Launches shiny app.

Usage
rfm_launch_app()

Examples
## Not run:
rfm_launch_app()
## End(Not run)

Customers by orders

Description
Visualize the distribution of customers across orders.

Usage
rfm_order_dist(rfm_table, bar_color = "blue", xaxis_title = "Orders",
yaxis_title = "Customers", plot_title = "Customers by Orders",
plot_title_justify = 0.5, print_plot = TRUE)

Arguments
rfm_table  An object of class rfm_table.
bar_color   Color of the bars.
xaxis_title X axis title.
yaxis_title Y axis title.
plot_title  Title of the plot.
plot_title_justify
Horizontal justification of the plot title; 0 for left justified and 1 for right justified.
print_plot  logical; if TRUE, prints the plot else returns a plot object.
Value

Bar chart.

Examples

```r
# using transaction data
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_order <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# order distribution
rfm_order_dist(rfm_order)

# using customer data
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_customer <- rfm_table_customer(rfm_data_customer, customer_id, number_of_orders, recency_days, revenue, analysis_date)

# order distribution
rfm_order_dist(rfm_customer)
```

rfm_plot_median_recency

Segmentation plots

Description

Segment wise median recency, frequency & monetary value plot.

Usage

```r
rfm_plot_median_recency(rfm_segment_table, print_plot = TRUE)
rfm_plot_median_frequency(rfm_segment_table, print_plot = TRUE)
rfm_plot_median_monetary(rfm_segment_table, print_plot = TRUE)
```

Arguments

- `rfm_segment_table`: Output from `rfm_segment`.
- `print_plot`: logical; if TRUE, prints the plot else returns a plot object.
Examples

```r
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_result <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)


recency_lower <- c(4, 2, 3, 4, 3, 2, 2, 1, 1, 1)
recency_upper <- c(5, 5, 5, 5, 4, 3, 3, 2, 1, 2)
frequency_lower <- c(4, 3, 1, 1, 1, 2, 1, 2, 4, 1)
frequency_upper <- c(5, 5, 3, 1, 1, 3, 2, 5, 5, 2)
monetary_lower <- c(4, 3, 1, 1, 2, 1, 2, 4, 1)
monetary_upper <- c(5, 5, 3, 1, 1, 3, 2, 5, 5, 2)

segments <- rfm_segment(rfm_result, segment_names, recency_lower, recency_upper, frequency_lower, frequency_upper, monetary_lower, monetary_upper)

rfm_plot_median_recency(segments)
rfm_plot_median_frequency(segments)
rfm_plot_median_monetary(segments)
```

**Description**

Examine the relationship between recency, frequency and monetary values.

**Usage**

```r
rfm_rm_plot(rfm_table, point_color = "blue", xaxis_title = "Monetary", yaxis_title = "Recency", plot_title = "Recency vs Monetary", print_plot = TRUE)

rfm_fm_plot(rfm_table, point_color = "blue", xaxis_title = "Monetary", yaxis_title = "Frequency", plot_title = "Frequency vs Monetary", print_plot = TRUE)

rfm_rf_plot(rfm_table, point_color = "blue", xaxis_title = "Frequency", yaxis_title = "Recency", plot_title = "Recency vs Frequency", print_plot = TRUE)
```
Arguments

rfm_table  An object of class rfm_table.
point_color  Color of the scatter points.
xaxis_title  X axis title.
yaxis_title  Y axis title.
plot_title  Title of the plot.
print_plot  logical; if TRUE, prints the plot else returns a plot object.

Value

Scatter plot.

Examples

# rfm table
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_result <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# monetary value vs recency
rfm_rm_plot(rfm_result)

# frequency vs monetary value
rfm_fm_plot(rfm_result)

# frequency vs recency
rfm_rf_plot(rfm_result)

rfm_segment

Segmentation

Description

Create segments based on recency, frequency and monetary scores.

Usage

rfm_segment(data, segment_names = NULL, recency_lower = NULL, recency_upper = NULL, frequency_lower = NULL, frequency_upper = NULL, monetary_lower = NULL, monetary_upper = NULL)
**rfm_table_customer**

**RFM table (customer data)**

**Description**

Recency, frequency, monetary and RFM score.

**Usage**

```r
rfm_table_customer(data = NULL, customer_id = NULL, n_transactions = NULL, recency_days = NULL, total_revenue = NULL, analysis_date = NULL, recency_bins = 5, frequency_bins = 5, monetary_bins = 5, ...)```
**Arguments**

- **data**: A `data.frame` or `tibble`.
- **customer_id**: Unique id of the customer.
- **n_transactions**: Number of transactions/orders.
- **recency_days**: Number of days since the last transaction.
- **total_revenue**: Total revenue from the customer.
- **analysis_date**: Date of analysis.
- **recency_bins**: Number of bins for recency or custom threshold.
- **frequency_bins**: Number of bins for frequency or custom threshold.
- **monetary_bins**: Number of bins for monetary or custom threshold.
- **...**: Other arguments.

**Value**

`rfm_table_customer` returns a list with the following:

- **rfm**: RFM table.
- **analysis_date**: Date of analysis.
- **frequency_bins**: Number of bins used for frequency score.
- **recency_bins**: Number of bins used for recency score.
- **monetary_bins**: Number of bins used for monetary score.
- **threshold**: `tibble` with thresholds used for generating RFM scores.

**Examples**

```r
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_table_customer(rfm_data_customer, customer_id, number_of_orders,
                   recency_days, revenue, analysis_date)

# access rfm table
result <- rfm_table_customer(rfm_data_customer, customer_id, number_of_orders,
                             recency_days, revenue, analysis_date)
result$rfm

# using custom threshold
rfm_table_customer(rfm_data_customer, customer_id, number_of_orders,
                    recency_days, revenue, analysis_date, recency_bins = c(115, 181, 297, 482),
                    frequency_bins = c(4, 5, 6, 8), monetary_bins = c(256, 382, 506, 666))
```
RFM table 2 (customer data)

**Description**

Recency, frequency, monetary and RFM score.

**Usage**

```r
rfm_table_customer_2(data = NULL, customer_id = NULL, n_transactions = NULL, latest_visit_date = NULL, total_revenue = NULL, analysis_date = NULL, recency_bins = 5, frequency_bins = 5, monetary_bins = 5, ...)
```

**Arguments**

- `data` A data.frame or tibble.
- `customer_id` Unique id of the customer.
- `n_transactions` Number of transactions/orders.
- `latest_visit_date` Date of the latest visit.
- `total_revenue` Total revenue from the customer.
- `analysis_date` Date of analysis.
- `recency_bins` Number of bins for recency or custom threshold.
- `frequency_bins` Number of bins for frequency or custom threshold.
- `monetary_bins` Number of bins for monetary or custom threshold.
- `...` Other arguments.

**Value**

`rfm_table_order` returns a list with the following:

- `rfm` RFM table.
- `analysis_date` Date of analysis.
- `frequency_bins` Number of bins used for frequency score.
- `recency_bins` Number of bins used for recency score.
- `monetary_bins` Number of bins used for monetary score.
- `threshold` tibble with thresholds used for generating RFM scores.
Examples

```r
analysis_date <- lubridate::as_date('2007-01-01', tz = 'UTC')
rfm_table_customer_2(rfm_data_customer, customer_id, number_of_orders,
most_recent_visit, revenue, analysis_date)

# access rfm table
result <- rfm_table_customer_2(rfm_data_customer, customer_id, number_of_orders,
most_recent_visit, revenue, analysis_date)
result$rfm

# using custom threshold
rfm_table_customer_2(rfm_data_customer, customer_id, number_of_orders,
most_recent_visit, revenue, analysis_date, recency_bins = c(115, 181, 297, 482),
frequency_bins = c(4, 5, 6, 8), monetary_bins = c(256, 382, 506, 666))
```

---

### rfm_table_order

**RFM table (transaction data)**

**Description**

Recency, frequency, monetary and RFM score.

**Usage**

```r
rfm_table_order(data = NULL, customer_id = NULL, order_date = NULL,
revenue = NULL, analysis_date = NULL, recency_bins = 5,
frequency_bins = 5, monetary_bins = 5, ...)
```

**Arguments**

- `data`: A data frame or tibble.
- `customer_id`: Unique id of the customer.
- `order_date`: Date of the transaction.
- `revenue`: Revenue from the customer.
- `analysis_date`: Date of analysis.
- `recency_bins`: Number of bins for recency or custom threshold.
- `frequency_bins`: Number of bins for frequency or custom threshold.
- `monetary_bins`: Number of bins for monetary or custom threshold.
- `...`: Other arguments.
rfm_table_order

Value

rfm_table_order returns a list with the following:

- **rfm**: RFM table.
- **analysis_date**: Date of analysis.
- **frequency_bins**: Number of bins used for frequency score.
- **recency_bins**: Number of bins used for recency score.
- **monetary_bins**: Number of bins used for monetary score.
- **threshold**: tibble with thresholds used for generating RFM scores.

Examples

```r
analysis_date <- lubridate::as_date('2006-12-31', tz = 'UTC')
rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)

# access rfm table
result <- rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date)
result$rfm

# using custom threshold
rfm_table_order(rfm_data_orders, customer_id, order_date, revenue, analysis_date, 
recency_bins = c(115, 181, 297, 482), frequency_bins = c(4, 5, 6, 8),
monetary_bins = c(256, 382, 506, 666))
```
Index

*Topic datasets
  rfm_data_customer, 4
  rfm_data_orders, 4

rfm, 2
rfm-package (rfm), 2
rfm_bar_chart, 3
rfm_barchart_data, 2
rfm_data_customer, 4
rfm_data_orders, 4
rfm_fm_plot (rfm_rm_plot), 11
rfm_heatmap, 5
rfm_heatmap_data, 6
rfm_hist_data, 8
rfm_histograms, 7
rfm_launch_app, 9
rfm_order_dist, 9
rfm_plot_median_frequency
  (rfm_plot_median_recency), 10
rfm_plot_median_monetary
  (rfm_plot_median_recency), 10
rfm_plot_median_recency, 10
rfm_rf_plot (rfm_rm_plot), 11
rfm_rm_plot, 11
rfm_segment, 12
rfm_table_customer, 13
rfm_table_customer_2, 15
rfm_table_order, 16