Package 'migrate'

July 22, 2025

Type Package

Title Create Credit State Migration (Transition) Matrices

Version 0.5.0

Description Tools to help convert credit risk data at two timepoints
    into traditional credit state migration (aka, ``transition") matrices.
    At a higher level, 'migrate' is intended to help an analyst understand
    how risk moved in their credit portfolio over a time interval.
    References to this methodology include:
    1. Schuermann, T. (2008) <doi:10.1002/9780470061596.risk0409>.
    2. Perederiy, V. (2017) <doi:10.48550/arXiv.1708.00062>.

License MIT + file LICENSE

URL https://github.com/ketchbrookanalytics/migrate,
    https://ketchbrookanalytics.github.io/migrate/

BugReports https://github.com/ketchbrookanalytics/migrate/issues

Encoding UTF-8

LazyData true

Suggests testthat (>= 2.1.0), knitr, rmarkdown

Depends R (>= 4.1)

Imports dplyr (>= 1.0.7), tidyr (>= 1.1.0), tibble (>= 3.0.1), rlang,
    utils, cli, glue

VignetteBuilder knitr

Language en-US

RoxygenNote 7.3.1

NeedsCompilation no

Author Michael Thomas [aut, cre],
    Brad Lindblad [ctb],
    Ivan Millanes [ctb]

Maintainer Michael Thomas <mthomas@ketchbrookanalytics.com>

Repository CRAN

Date/Publication 2024-07-10 13:20:02 UTC

Contents

---

build_matrix   *Build a state migration (transition) matrix*

Description

build_matrix() creates a state transition matrix from summarized data (i.e., a data frame returned by migrate()) representing each unique combination of beginning & ending states and a numeric metric.

Usage

```
build_matrix(data, state_start = NULL, state_end = NULL, metric = NULL)
```

Arguments

data   A data frame or data frame extension (e.g., a tibble or data.table) containing a minimum of three (3) column variables representing a starting credit risk state, an ending credit risk state, and a metric containing values representing the movement (i.e., "transition) in that metric between the starting credit risk state point in time and the ending credit risk state point in time. This style of data frame is output by the migrate() function within this package.

state_start   (Optional) A symbol or string, representing the column variable of the data data frame argument that contains the starting credit risk state values. If left null, function will attempt to find it for you.

state_end   (Optional) A symbol or string, representing the column variable of the data data frame argument that contains the starting credit risk state values. If left null, function will attempt to find it for you.

metric   (Optional) A symbol or string, representing the column variable of the data data frame argument that contains the metric for which the grouped difference in value between the starting credit risk state period and ending credit risk state period was computed.

Value

A matrix object, where the first (row) dimension represents the starting credit risk state, the second (column) dimension represents the ending credit risk state, and the values within the matrix represent the transitioned amount based upon the values in the metric numeric column variable from the data data frame.

Note: A matrix object can be coerced to a data frame using as.data.frame().

Examples

```
# Let `build_matrix()` guess which column variables represent `state_start`,
# `state_end` and `metric`
mock_credit |>
  migrate(
    time = date,
    state = risk_rating,
    id = customer_id,
    metric = principal_balance
  ) |>
  build_matrix()

# Specify which column variables represent `state_start`, `state_end` and
# `metric`
mock_credit |>
  migrate(
    id = customer_id,
    time = date,
    state = risk_rating,
    percent = FALSE
  ) |>
  build_matrix(
    state_start = risk_rating_start,
    state_end = risk_rating_end,
    metric = count
  )
```

---

migrate | *Summarize the migration of a data frame*

Description

migrate() summarizes the transition amount (or percentage) of a numeric variable from each beginning credit risk state category to each ending credit risk state, given a data frame input.

Usage

```
migrate(
  data,
  id,
  time,
  state,
  metric = NULL,
  percent = TRUE,
  fill_state = NULL,
  verbose = TRUE
)
```

Arguments

`data`	A data frame or data frame extension (e.g., a tibble or data.table) containing a minimum of three (3) column variables representing a time, a credit risk state, and an ID identifying the credit facility (we would expect to see most unique values in this column variable appear twice in the dataset; once at the first unique `time` value and again at the second unique `time` value, unless the ID only existed at one of those two times).
`id`	The column variable of the `data` data frame argument that contains the unique identifier to track where a particular credit facility migrated to/from. If left null, `migrate()` will attempt to use the first column variable from the data frame provided in the `data` argument.
`time`	The column variable of in the `data` data frame representing the timepoint (e.g., a Date) of each observation; this column should contain two unique values (migration from Time A to Time B)
`state`	The column variable of the `data` data frame argument that contains the credit risk state values.
`metric`	(Optional) The column variable of type "numeric" in the `data` data frame argument that contains the continuous metric values to weight the state migration by
`percent`	If `FALSE`, will calculate the migration on an absolute basis (rather than a percentage basis, which is the default)
`fill_state`	(Optional) A value (e.g., a character string such as "No Rating" or "NR") to be used as the *filler* state for any `id` values that only exist at one timepoint in the data.
`verbose`	If `TRUE`, the function returns an informational message about the transition period

Value

A data frame containing three (3) column variables representing the unique combinations of starting & ending credit risk states and the calculated migration observed during the period.

Examples

```
# Return the percent migration of the number of credit facilities
migrate(
  data = mock_credit,
  id = customer_id,
  time = date,
  state = risk_rating
)

# Return the absolute migration in `principal_balance`
migrate(
  data = mock_credit,
  id = customer_id,
  time = date,
```

```
    state = risk_rating,
    metric = principal_balance,
    percent = FALSE
  )

  # Provide a filler `state` value when a unique `id` is missing a timepoint
  migrate(
    data = head(mock_credit, n = 995),   # drop the last 5 observations
    id = customer_id,
    time = date,
    state = risk_rating,
    fill_state = "NR",
    percent = FALSE
  )
```

mock_credit	*Mock dataset containing credit statistics by customer at two time intervals. Some customers only exist in one time interval (they either became a customer after the first time interval, or discontinued being a customer before the second time interval).*

Description

Mock dataset containing credit statistics by customer at two time intervals. Some customers only exist in one time interval (they either became a customer after the first time interval, or discontinued being a customer before the second time interval).

Usage

```
mock_credit
```

Format

A data frame with columns:

customer_id Customer ID for 497 unique customers.

date Date of the observation; there are two unique dates in the dataset.

risk_rating Factor representing risk level on a 1-14 scale.

principal_balance Principal balance outstanding on the loan.

Source

Developed by Ketchbrook Analytics

Examples

```
## Not run:
 mock_credit

## End(Not run)
```

Index