Package: shadowpop (via r-universe)

September 8, 2024

Title Creates a ``shadow population'' for comparative data analysis

Version 0.0.1

License MIT + file LICENSE

Description Creates a ``shadow population'' for comparative data analysis.

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Depends R (>= 4.4.0)

Config/testthat/edition 3

Imports assertthat, cli, dplyr, glue, purrr, rlang, stringr, tidyr, tidyselect, withr

Suggests gt, palmerpenguins, quarto, testthat (>= 3.0.0), tibble

VignetteBuilder quarto

Repository https://francisbarton.r-universe.dev

RemoteUrl https://codeberg.org/francisbarton/shadowpop

RemoteRef HEAD

RemoteSha ddf43b7b30851fad7616823a9b59f4923a4105f4

Contents

---

shadowpop	*Create a shadow population that closely resembles a sample population*

Description

This is a bit like creating a control group, or a training/testing set. However, the aim is to create the best possible match between the sample population (provided by the user) and the shadow population (generated by the shadowpop() function).

Usage

```
shadowpop(
  sample_data,
  source_data,
  match_cols = NULL,
  id_col = NULL,
  shadow_size = NULL
)
```

Arguments

sample_data	data frame Your sample population
source_data	data frame A larger population from which to extract a shadow sample
match_cols	A character vector of column names shared between sample_data and source_data, to be used as variables to match against. Put the variable you care about most at the start of the list; "nice to have" matches towards the end
id_col	character A column name not included in match_cols, which acts as a unique ID or key for each row in sample_data and in source_data. An ID column named "id" will be automatically created if no id_col is specified
shadow_size	integer The number of rows of shadow population to create. If not specified this will be set to the same nrow as sample_data. Should be (meaningfully) less than nrow(source_data)!

Index

shadowpop,