Package: themask (via r-universe)

October 31, 2024

Title Masks and houses the PEPFAR MSD-style training dataset for testing and training

Version 1.1.5

Description This package creates and hosts a masked, dummy dataset that should be used for testing, training, and demoing instead of using actual PEPFAR data.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Imports base, cli, dplyr, googlesheets4, magrittr, piggyback, readr, stats, stringr, tibble, utils

Remotes USAID-OHA-SI/gagglr, USAID-OHA-SI/glamr, USAID-OHA-SI/gophr

Depends R (>= 2.10)

LazyData true

Suggests glue, googledrive, ggplot2, keyring, knitr, pak, rlang (>= 1.0.0), rmarkdown, testthat (>= 3.0.0), withr

Config/testthat/edition 3

Config/testthat/parallel true

VignetteBuilder knitr

Repository https://usaid-oha-si.r-universe.dev

RemoteUrl https://github.com/USAID-OHA-SI/themask

RemoteRef HEAD

RemoteSha 094cfa2e48906690c9ed49153383aba0a6a31e84

Contents

`milb` *MiLB Information Table*

Description

This table has been extracted from Wikipedia's Minor League Baseball (MiLB) page and is used for masking geographic and partner information in PEPFAR's MER Structured Datasets (MSD).

Usage

```
milb
```

Format

`milb`**:**

A data frame with 120 rows and 6 columns:

division MLB Division (East, Central, West)

mlb_team Affiliated MLB

level MilB team level (High A, A, AA, AAA)

league MilB team level (varies by level)

city MilB team city/location

name MilB team name

Source

https://en.wikipedia.org/wiki/Minor_League_Baseball

minoria_geo *Minoria Geography Table*

Description

This table mimics the structure of PEPFAR's MER Structured Datasets (MSD). The dummy dataset for the (Kingdom of) Minoria has four region (snu1), each containing four districts (psnu). Derived from milb.

Usage

```
minoria_geo
```

Format

minoria_geo:
A data frame with 16 rows and 7 columns:

operatingunit Operating Unit name (Minoria)
operatingunituid OU unique ID (Minoria)
country Country (Minoria)
snu1 Sub-National Unit 1 level below national (from MiLB league)
snu1uid SNU1 unique ID
cop22_psnu Priority SNU (from MiLB city)
cop22_psnuuid PSNU unique ID

minoria_mechs *Minoria Mechanism Table*

Description

This table mimics the structure of PEPFAR's MER Structured Datasets (MSD) for mechanims, which originally originate from FACTInfo NextGen. The dummy dataset for the (Kingdom of) Minoria has 120 (plus two dedup) mechanisms available for use. Derived from milb.

Usage

```
minoria_mechs
```

Format

minoria_mechs:
A data frame with 122 rows and 7 columns:

mech_code Unique implementing mechanism code
mech_name implementing mechanism code (from MiLB team name)
prime_partner_name implementing mechanism partner (from MiLB team name)

minoria_shp_ou *Minoria PSNU Shape File*

Description

This dataset is the sf file for mapping (Kingdom of) Minoria at the national level

Usage

```
minoria_shp_ou
```

Format

`minoria_shp_ou`:
A data frame with 16 rows and 7 columns:
operatingunituid OU unique ID (Minoria)
operatingunit Operating Unit name (Minoria)
country Country (Minoria)
geometry Country level polygon shape for mapping

See Also

Other shp: `minoria_shp_psnu`, `minoria_shp_snu1`

minoria_shp_psnu *Minoria PSNU Shape File*

Description

This dataset is the sf file for mapping (Kingdom of) Minoria by PSNU.

Usage

```
minoria_shp_psnu
```

Format

`minoria_shp_psnu`:
A data frame with 16 rows and 8 columns:
psnu Priority SNU (from MiLB city)
psnuuid PSNU unique ID
snu1 Sub-National Unit 1 level below national (from MiLB league)
snu1uid SNU1 unique ID
operatingunit Operating Unit name (Minoria)
operatingunituid OU unique ID (Minoria)
country Country (Minoria)
geometry PSNU level polygon shape for mapping

See Also

Other shp: minoria_shp_ou, minoria_shp_snu1

`minoria_shp_snu1` *Minoria PSNU Shape File*

Description

This dataset is the sf file for mapping (Kingdom of) Minoria by SNU1.

Usage

```
minoria_shp_snu1
```

Format

`minoria_shp_snu1`:
A data frame with 4 rows and 6 columns:

snu1uid SNU1 unique ID
snu1 Sub-National Unit 1 level below national (from MiLB league)
operatingunit Operating Unit name (Minoria)
operatingunituid OU unique ID (Minoria)
country Country (Minoria)
geometry SNU1 level polygon shape for mapping

See Also

Other shp: minoria_shp_ou, minoria_shp_psnu

`msk_available` *Check the latest version available*

Description

This function is used to check what masked version is currently available and will flag if there it is up to date or you should run `msk_create` yourself. It will also list all available historic releases that can be downloaded in `msk_download` by specifying the version in the tag param.

Usage

```
msk_available()
```

See Also

Other download: msk_download()

msk_create *Create the Masked Dataset*

Description

This function is used to create a masked dataset for use in testing and training. The data are subset to 16 PSNUs and masked across geographic and mechanism variables. Either the PSNUxIM or NAT_SUBNAT Structured Datasets can be masked.

Usage

```
msk_create(filepath, output_folder)
```

Arguments

filepath	path to the PSD file (PSNUxIM or NAT_SUBNAT)
output_folder	location where you want to store the new file (default = does not export the data)

Value

dataframe with converted geography + mech info

Examples

```
## Not run:
#create a masked dataset
library(glamr)
library(themask)

#store path to latest MSD
path <- si_path() %>% return_latest("PSNU_IM")

#create a masked dataset from the PSNUxIM MSD
msk_create(path, "project1/data")

## End(Not run)
```

msk_download *Download Masked Dataset*

Description

This function download a masked dataset from GitHub (USAID-OHA-SI/themask) for use in training or testing. Recommend running msk_available to see what version is available to download.

Usage

```
msk_download(folderpath, tag = "latest", launch = FALSE)
```

Arguments

folderpath	where should the file be downloaded to?
tag	version tag, default = "latest"
launch	whether to launch Windows Explorer to the location after the download completes (default = FALSE)

References

msk_available

See Also

Other download: msk_available()

Examples

```
## Not run:
#check available version
msk_available()

#download to your downloads folder
msk_download("~/Downloads")

## End(Not run)
## Not run:
#download an older version
msk_download("~/Downloads", tag = "2023.06.27c")

## End(Not run)
```

---

msk_gen_uid *Generate a Unique ID*

Description

This function is used to create a unique ID (UID) to mimic the UIDs created and used by DATIM for the MER and other PEPFAR Structured Datasets.

Usage

```
msk_gen_uid(codeSize = 11)
```

Arguments

codeSize	character length for UID output (default = 11)

Value

random alphanumeric string

Examples

```
msk_gen_uid()
```

`msk_release` *Upload a New Masked Dataset to GitHub*

Description

This function is used to upload the new masked dataset to GitHub (USAID-OHA-SI/themask) so that others can use it. This function is for package developers use only. The data from `msk_create()` are subset to 16 PSNUs and masked across geographic and mechanism variables. Either the PSNUxIM or NAT_SUBNAT Structured Datasets can be masked.

Usage

```
msk_release(filepath, output_folder)
```

Arguments

`filepath`	path to the PSD file (PSNUxIM or NAT_SUBNAT) or masked file
`output_folder`	location where you want to store the new file (default does not save the file)

Value

dataframe with converted geography + mech info

References

msk_create

Examples

```
## Not run:
#create and upload a new release
library(glamr)
library(themask)

#store path to latest MSD
path <- si_path() %>% return_latest("PSNU_IM")

#create a masked dataset from the PSNUxIM MSD
msk_release(path, "project1/data")

## End(Not run)
```

```
## Not run:
#upload release from an existing masked dataset
library(glamr)
library(themask)

#store path to masked dataset
path_msk <- return_latest("project1/data","TRAINING")

#create a masked dataset from the PSNUxIM MSD
msk_release(path_msk)

## End(Not run)
```

Index