

Package ‘rrefine’

April 9, 2016

Type Package

Title R Client for OpenRefine API

Version 1.0

Date 2016-03-26

Maintainer VP Nagraj <vpnagraj@virginia.edu>

Description 'OpenRefine' (formerly 'Google Refine') is a popular, open source data cleaning software. This package enables users to programmatically trigger data transfer between R and 'OpenRefine'. Available functionality includes project import, export and deletion.

License GPL-3

URL <http://openrefine.org/>

LazyData TRUE

RoxygenNote 5.0.1

Imports httr (>= 1.1.0), rlist, readr

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author VP Nagraj [aut, cre]

Repository CRAN

Date/Publication 2016-04-09 00:55:18

R topics documented:

lateformeeting	2
lfn_clean	2
refine_check	3
refine_delete	3
refine_export	4
refine_id	5
refine_metadata	5
refine_path	6
refine_upload	6

Index[7](#)

lateformeeting	<i>a "dirty" data set to demonstrate rrefine features</i>
----------------	---

Description

This data is a simulated collection of dates, days of the week, numbers of hours slept and indicators of whether or not the subject was on time for work. All observations appearing in this data set are fictitious, and any resemblance to actual arrival times for work is purely coincidental.

Usage

```
lateformeeting
```

Format

A data frame with 63 rows and 4 variables

- theDate date of observation in varying formats
- what.day.whas.it day of the week in varying formats
- sleephours number of hours slept
- was.i.on.time.for.work indicator of on-time arrival to work

Examples

```
head(lateformeeting)
```

lfm_clean	<i>a "clean" version of the lateformeeting sample data set</i>
-----------	--

Description

This data is a simulated collection of dates, days of the week, numbers of hours slept and indicators of whether or not the subject was on time for work. All observations appearing in this data set are fictitious, and any resemblance to actual arrival times for work is purely coincidental.

Usage

```
lfm_clean
```

Format

A data frame with 63 rows and 4 variables

- date date of observation in POSIXct format
- dotw day of the week in consistent format
- hours.slept number of hours slept
- on.time indicator of on-time arrival to work

Examples

```
head(lfm_clean)
```

refine_check	<i>helper function to check if rrefine can connect to OpenRefine</i>
--------------	--

Description

helper function to check if rrefine can connect to OpenRefine

Usage

```
refine_check()
```

Value

error message if rrefine is unable to connect to OpenRefine, otherwise is invisible

refine_delete	<i>delete project from OpenRefine</i>
---------------	---------------------------------------

Description

delete project from OpenRefine

Usage

```
refine_delete(project.name = NULL, project.id = NULL)
```

Arguments

project.name	name of project to be deleted
project.id	unique identifier for open refine project to be deleted

Examples

```
## Not run:
refine_delete(project.name = "foo")
refine_delete(project.id = 1901017388690)
refine_delete(project.name = "Untitled", project.id = 1901018888332)

## End(Not run)
```

refine_export	<i>export data from OpenRefine</i>
---------------	------------------------------------

Description

export data from OpenRefine

Usage

```
refine_export(project.name = NULL, project.id = NULL, format = "csv",
  col.names = TRUE, encoding = "UTF-8")
```

Arguments

project.name	name of project to be exported
project.id	unique identifier for project to be exported
format	file format of project to be exported, default is 'csv'
col.names	logical indicator for whether column names should be included
encoding	character encoding for exported data, default is 'UTF-8'

Examples

```
## Not run:
refine_export("purple_rain")
refine_export(project.id = 1901017388690, format = "csv")
refine_export(project.name = "Untitled", project.id = 1901017388888)

## End(Not run)
```

refine_id	<i>helper function to get OpenRefine project.id by project.name</i>
-----------	---

Description

helper function to get OpenRefine project.id by project.name

Usage

```
refine_id(project.name, project.id)
```

Arguments

project.name	name of project to be exported or deleted
project.id	unique identifier for project to be exported or deleted

Value

unique id of project to be exported or deleted

refine_metadata	<i>get all project metadata from OpenRefine</i>
-----------------	---

Description

get all project metadata from OpenRefine

Usage

```
refine_metadata()
```

Examples

```
## Not run:  
refine_metadata()  
  
## End(Not run)
```

refine_path	<i>helper function to configure and call path to OpenRefine</i>
-------------	---

Description

helper function to configure and call path to OpenRefine

Usage

```
refine_path(host = "127.0.0.1", port = "3333")
```

Arguments

host	host path for your OpenRefine application
port	port number for your OpenRefine application

Value

path to be executed

refine_upload	<i>upload a csv file to OpenRefine</i>
---------------	--

Description

upload a csv file to OpenRefine

Usage

```
refine_upload(file, project.name = NULL, open.browser = FALSE)
```

Arguments

file	Name of csv file to be uploaded
project.name	optional parameter to specify name of the project to be created upon upload, default is NULL and project will be named 'Untitled' in OpenRefine
open.browser	logical for whether or not the browser should open on successful upload

Examples

```
## Not run:
write.csv(x = mtcars, file = "mtcars.csv")
refine_upload(file = "mtcars.csv", project.name = "mtcars_clean_up")

## End(Not run)
```

Index

*Topic **datasets**

lateformeeting, [2](#)

lfm_clean, [2](#)

lateformeeting, [2](#)

lfm_clean, [2](#)

refine_check, [3](#)

refine_delete, [3](#)

refine_export, [4](#)

refine_id, [5](#)

refine_metadata, [5](#)

refine_path, [6](#)

refine_upload, [6](#)