Package: diffdfs (via r-universe)

September 8, 2024

Title Compute the Difference Between Data Frames

Version 0.9

Description Shows you which rows have changed between two data frames
with the same column structure. Useful for diffing slowly
mutating data.

License MIT + file LICENSE

Imports arrow, dplyr, janitor, rlang

BugReports <https://github.com/riazarbi/diffdfs>

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

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

RemoteUrl https://github.com/riazarbi/diffdfs

RemoteRef HEAD

RemoteSha 750cf074b965583a1ef8ce05de7b5b1cdd859eb2

Contents

---

checkkey   *Check That A Dataframe Key Col Set Is Unique*

Description

Checks that a provided vector of column names constitue a unique key (that is, no rows are duplicated) for a dataframe.

Usage

```
checkkey(df, key_cols, verbose = FALSE)
```

Arguments

df	a dataframe
key_cols	vector of column names
verbose	TRUE/FALSE should we print a message?

Value

TRUE if key cols have unique rows; FALSE if not

Examples

```
irisint = iris
irisint$rownum = 1:nrow(irisint)
key_cols = c("rownum")
checkkey(irisint, key_cols, TRUE)
checkkey(irisint, "Species", TRUE)
```

---

diffdfs   *Compute the Difference Between Dataframes*

Description

Returns a dataframe describing the modifications required to transform old_df into new_df. The dataframes needBugReports: https://github.com/tidyverse/dplyr/issues to have identical columns and column types and share unique index columns.

Usage

```
diffdfs(new_df, old_df = NA, key_cols = NA, verbose = FALSE)
```

Arguments

new_df	A dataframe of new data.
old_df	A dataframe of old data. new_df and old_df can (and usually do) have overlapping data.
key_cols	optional vector of column names that constitute a unique table key. If NA, colnames(old_df) will be used.
verbose	logical, default FALSE. Should the processing be chatty?

Value

a dataframe.

Examples

```
iris$key <- 1:nrow(iris)

old_df <- iris[1:100,]
old_df[75,1] <- 100
new_df <- iris[50:150,]
diffdfs(new_df, old_df, key_cols = "key")
```

Index