Package 'identifyr'

December 23, 2015

Type Package

Title Clean X-Numbers

Version 0.1

Date 2015-12-21

Author George Miranda

Maintainer George Miranda <georgemirandajr@gmail.com>

Description This package provides methods for standardizing identifiers commonly used by Probation Department.

License GPL

LazyData TRUE

Depends stringr (>= 1.0.0), magrittr (>= 1.5)

Imports stringr (>= 1.0.0), magrittr (>= 1.5)

RoxygenNote 5.0.1

Suggests testthat

NeedsCompilation no

R topics documented:

---

clean_id   *Clean ID*

Description

Provides a method for applying both cleaning functions (clean_x and clean_pb) to a dataframe by index.

Usage

```
clean_id(data, cols = list(...), FUN = list(...))
```

Arguments

data	is a data.frame containing any number of columns.
cols	is a list of integers indicating the column numbers that are to be cleaned up.
FUN	is a list of functions to pass to the specified columns (must be in same order as the column index)

Details

This function is meant to be used on multiple columns at once. It makes use of the magrittr pipe operator to provide neat code.

Value

Returns a dataframe with replacements in the specified columns

See Also

clean_x, clean_pb

Examples

```
df <- data.frame(pbnum = c("PB123", "PB 0034", "  5678 ", "None"),
status = c("Active", "Closed", "Closed", "Active"),
xnum = c("X00123", "9512", "X789", "NOT IN APS"))

The magrittr way:
df %>% clean_id(cols = c(1,3), FUN = c("clean_pb", "clean_x"))

The usual way:
clean_id(data = df, cols = c(1,3), FUN = c("clean_pb", "clean_x"))
```

---

clean_pb	*Clean PB Numbers*

Description

Useful for correcting manually entered PB numbers that may be abbreviated or contain non-standard methods for handling null values such as "No #"

Usage

```
clean_pb(x)
```

Arguments

x	is a character vector containing PB numbers

Value

Character vector of standardized PB numbers. NA values are introduced where there appears to be non-X characters present.

See Also

clean_x, clean_id

Examples

```
clean_pb("PB123") # PB000123
clean_x("None") # NA
clean_x("5678") # PB005678
```

---

clean_x	*Clean X Numbers*

Description

Useful for correcting manually entered X-numbers that may be abbreviated or contain non-standard methods for handling null values such as "No #"

Usage

```
clean_x(x)
```

Arguments

x   is a character vector containing X numbers

Value

Character vector of standardized X numbers. NA values are introduced where there appears to be non-X characters present.

See Also

clean_pb, clean_id

Examples

```
clean_x("X6789") # X00006789
clean_x("No #") # NA
clean_x("1234") # X00001234
```

Index