Package 'edm1.df'

July 7, 2024

Title Set of functions for dataframe manipulation
Version 2.0.0.0
Description Provides multiple functions to manipulate data in dataframe according to different algorythms for different goals.
License GPL (==3)
Encoding UTF-8
Roxygen list(markdown = TRUE)
RoxygenNote 7.3.1
Imports stringr, stringi, dplyr, openxlsx
Contents
colins_datf

2 colins_datf

Description

Allow to insert vectors into a dataframe.

Usage

```
colins_datf(inpt_datf, target_col = list(), target_pos = list())
```

Arguments

```
inpt_datf is the dataframe where vectors will be inserted
target_col is a list containing all the vectors to be inserted
target_pos is a list containing the vectors made of the columns names or numbers where
the associated vectors from target_col will be inserted after
```

```
datf1 <- data.frame("frst_col"=c(1:5), "scd_col"=c(5:1))</pre>
print(colins_datf(inpt_datf=datf1, target_col=list(c("oui", "oui", "oui", "non", "non"),
            c("u", "z", "z", "z", "u")),
              target_pos=list(c("frst_col", "scd_col"), c("scd_col"))))
# frst_col cur_col scd_col cur_col.1 cur_col
       1 oui 5 oui
#1
#2
         2
              oui
                       4
                              oui
                                        7.
        3
                       3
#3
              oui
                              oui
                                       Z
                       2
#4
        4
                              non
                                        Z
              non
#5
              non
                               non
print(colins_datf(inpt_datf=datf1, target_col=list(c("oui", "oui", "oui", "non", "non"),
            c("u", "z", "z", "z", "u")),
              \verb|target_pos=list(c(1, 2), c("frst_col")))|
# frst_col cur_col scd_col cur_col cur_col
#1
        1 oui 5 u oui
                       4
#2
         2
              oui
                              Z
                                    oui
                       3
                              Z
#3
        3
              oui
                                    oui
#4
        4
                       2
              non
                               Z
                                    non
#5
        5
                       1
              non
                                     non
```

cumulated_rows 3

```
cumulated_rows cumulated_rows
```

Description

Output a vector of size that equals to the rows number of the input dataframe, with TRUE value at the indices corresponding to the row where at least a cell of any column is equal to one of the values inputed in values_v

Usage

```
cumulated_rows(inpt_datf, values_v = c())
```

Arguments

Examples

```
datf_teste <- data.frame(c(1:10), c(10:1))</pre>
print(datf_teste)
   c.1.10. c.10.1.
1
        1
                10
2
         2
                 9
3
         3
                 8
4
         4
                 7
5
         5
                 6
6
         6
                 5
7
         7
8
         8
                 3
9
         9
                 2
10
        10
                 1
print(cumulated_rows(inpt_datf = datf_teste, values_v = c(2, 3)))
      FALSE TRUE TRUE
                                        FALSE FALSE TRUE TRUE
[1]
                         FALSE
                                FALSE
                                                                     FALSE
```

Description

Output a vector of size that equals to the rows number of the input dataframe, with TRUE value at the indices corresponding to the row where at least a cell of any column is equal to NA.

4 cut_v

Usage

```
cumulated_rows_na(inpt_datf)
```

Arguments

```
inpt_datf is the input data.frame
```

Examples

```
datf_teste <- data.frame(c(1, 2, 3, 4, 5, NA, 7), c(10, 9, 8, NA, 7, 6, NA))
print(datf_teste)
 c.1..2..3..4..5..NA..7. c.10..9..8..NA..7..6..NA.
1
                       1
2
                        2
3
                        3
4
                                                 NA
5
                       5
                                                  7
6
                       NA
                                                  6
                                                 NA
print(cumulated_rows_na(inpt_datf = datf_teste))
[1] FALSE FALSE FALSE TRUE FALSE TRUE TRUE
```

```
cut_v cut_v
```

Description

Allow to convert a vector to a dataframe according to a separator.

Usage

```
cut_v(inpt_v, sep_ = "")
```

Arguments

```
inpt_v is the input vector
sep_ is the separator of the elements in inpt_v, defaults to ""
```

diff_datf 5

diff_datf

 $diff_datf$

Description

Returns a vector with the coordinates of the cell that are not equal between 2 dataframes (row, column).

Usage

```
diff_datf(datf1, datf2)
```

Arguments

```
datf1 is an an input dataframe datf2 is an an input dataframe
```

Examples

```
datf1 <- data.frame(c(1:6), c("oui", "oui", "oui", "oui", "oui", "oui", "oui"), c(6:1))
datf2 <- data.frame(c(1:7), c("oui", "oui", "oui", "oui", "non", "oui", "zz"))
print(diff_datf(datf1=datf1, datf2=datf2))
#[1] 5 1 5 2</pre>
```

groupr_datf

groupr_datf

Description

Allow to create groups from a dataframe. Indeed, you can create conditions that lead to a flag value for each cell of the input dataframeaccording to the cell value. This function is based on see_datf and nestr_datf2 functions.

6 id_keepr

Usage

```
groupr_datf(
  inpt_datf,
  condition_lst,
  val_lst,
  conjunction_lst,
  rtn_val_pos = c()
)
```

Arguments

Examples

```
interactive()
datf1 <- data.frame(c(1, 2, 1), c(45, 22, 88), c(44, 88, 33))
val_lst <- list(list(c(1), c(1)), list(c(2)), list(c(44, 88)))
condition_lst <- list(c(">", "<"), c("%%"), c("==", "=="))
conjunction_lst <- list(c("|"), c(), c("|"))
rtn_val_pos <- c("+", "++", "+++")
print(groupr_datf(inpt_datf=datf1, val_lst=val_lst, condition_lst=condition_lst, conjunction_lst=conjunction_lst, rtn_val_pos=rtn_val_pos))
# X1 X2 X3
#1 <NA> + +++
#2 ++ ++++++
#3 <NA> ++++ ++
```

id_keepr

id_keepr_datf

Description

Allow to get the original indexes after multiple equality comparaison according to the original number of row

insert_datf 7

Usage

```
id_keepr(inpt_datf, col_v = c(), el_v = c(), rstr_l = NA)
```

Arguments

inpt_datf	is the input dataframe
col_v	is the vector containing the column numbers or names to be compared to their respective elements in "el_v" $$
el_v	is a vector containing the elements that may be contained in their respective column described in " col_v "
rstr_l	is a list containing the vector composed of the indexes of the elements chosen for each comparison. If the length of the list is inferior to the length of comparisons, so the last vector of rstr_l will be the same as the last one to fill make rstr_l equal in term of length to col_v and el_v

Examples

insert_datf

edm1 insert_datf

Description

Allow to insert dataframe into another dataframe according to coordinates (row, column) from the dataframe that will be inserted

Usage

```
insert_datf(datf_in, datf_ins, ins_loc)
```

8 intersect_mod

Arguments

datf_in is the dataframe that will be inserted
datf_ins is the dataset to be inserted
ins_loc is a vector containg two parameters (row, column) of the begining for the insertion

Examples

```
datf1 \leftarrow data.frame(c(1, 4), c(5, 3))
datf2 \leftarrow data.frame(c(1, 3, 5, 6), c(1:4), c(5, 4, 5, "ereer"))
print(insert_datf(datf_in=datf2, datf_ins=datf1, ins_loc=c(4, 2)))
    c.1..3..5..6. c.1.4. c.5..4..5...ereer..
# 1
                1
                       1
# 2
                 3
                        2
                                             4
                                             5
# 3
                5
                        3
# 4
                 6
                        1
                                             5
print(insert_datf(datf_in=datf2, datf_ins=datf1, ins_loc=c(3, 2)))
#
    c.1..3..5..6. c.1.4. c.5..4..5...ereer..
# 1
                1
                        1
# 2
                 3
                        2
                                             4
# 3
                5
                        1
                                             5
# 4
                 6
                        4
                                             3
print(insert_datf(datf_in=datf2, datf_ins=datf1, ins_loc=c(2, 2)))
    c.1..3..5..6. c.1.4. c.5..4..5...ereer..
# 1
                1
                       1
# 2
                3
                        1
                                             5
# 3
                 5
                        4
                                             3
                        4
# 4
                                         ereer
```

Description

Returns the mods that have elements in common

Usage

```
intersect_mod(datf, inter_col, mod_col, n_min, descendly_ordered = NA)
```

Arguments

is the input dataframe
inter_col is the column name or the column number of the values that may be commun
betwee the different mods

intersect_mod 9

 $\begin{array}{ll} \verb|mod_col| & is the column name or the column number of the mods in the data frame \\ \verb|n_min| & is the minimum elements in common a mod should have to be taken in count \\ \verb|ordered_descendly| & \\ \hline| & &$

in case that the elements in commun are numeric, this option can be enabled by giving a value of TRUE or FALSE see examples

Examples

5

oui

5

```
datf <- data.frame("col1"=c("oui", "oui", "oui", "oui", "oui", "oui",</pre>
                    "non", "non", "non", "ee", "ee", "ee"), "col2"=c(1:6, 2:5, 1:
print(intersect_mod(datf=datf, inter_col=2, mod_col=1, n_min=2))
  col1 col2
  oui
          3
3
   oui
7
  non
          2
8
          3
   non
          2
12
   ee
13
          3
   ee
print(intersect_mod(datf=datf, inter_col=2, mod_col=1, n_min=3))
  col1 col2
2
  oui 2
          3
3
  oui
4
  oui
          4
5
          5
  oui
7
          2
   non
8
          3
   non
9
   non
          4
10 non
          5
print(intersect_mod(datf=datf, inter_col=2, mod_col=1, n_min=5))
  col1 col2
1 oui
       1
2 oui
3 oui
         3
4 oui
        4
5 oui
6 oui
datf <- data.frame("col1"=c("non", "non", "oui", "oui", "oui", "oui",</pre>
                     "non", "non", "non", "ee", "ee", "ee"), "col2"=c(1:6, 2:5, 1
print(intersect_mod(datf=datf, inter_col=2, mod_col=1, n_min=3))
  col1 col2
8
  non 3
9
   non
          4
10 non
          5
          3
3
   oui
4
          4
   oui
```

10 nestr_datf1

nestr_datf1

 $nestr_datf1$

Description

Allow to write a value (1a) to a dataframe (1b) to its cells that have the same coordinates (row and column) than the cells whose value is equal to a another special value (2a), from another another dataframe (2b). The value (1a) depends of the cell value coordinates of the third dataframe (3b). If a cell coordinates (1c) of the first dataframe (1b) does not correspond to the coordinates of a good returning cell value (2a) from the dataframe (2b), so this cell (1c) can have its value changed to the same cell coordinates value (3a) of a third dataframe (4b), if (4b) is not set to NA.

Usage

```
nestr_datf1(
  inptf_datf,
  inptt_pos_datf,
  nestr_datf,
  yes_val = TRUE,
  inptt_neg_datf = NA
)
```

Arguments

```
print(nestr_datf1(inptf_datf=data.frame(c(1, 2, 1), c(1, 5, 7)))
inptt_pos_datf=data.frame(c(4, 4, 3), c(2, 1, 2)),
inptt_neg_datf=data.frame(c(44, 44, 33), c(12, 12, 12)),
nestr_datf=data.frame(c(TRUE, FALSE, TRUE), c(FALSE, FALSE, TRUE)), yes_val=TRUE))
  c.1..2..1. c.1..5..7.
#1
           4
                      12
#2
           44
                      12
#3
            3
                       2
print(nestr_datf1(inptf_datf=data.frame(c(1, 2, 1), c(1, 5, 7)),
inptt_pos_datf=data.frame(c(4, 4, 3), c(2, 1, 2)),
inptt_neg_datf=NA,
nestr_datf=data.frame(c(TRUE, FALSE, TRUE), c(FALSE, FALSE, TRUE)), yes_val=TRUE))
#
    c.1..2..1. c.1..5..7.
#1
            4
            2
#2
                       5
```

nestr_datf2

#3 3 2

Description

Allow to write a special value (1a) in the cells of a dataframe (1b) that correspond (row and column) to whose of another dataframe (2b) that return another special value (2a). The cells whose coordinates do not match the coordinates of the dataframe (2b), another special value can be written (3a) if not set to NA.

Usage

```
nestr_datf2(inptf_datf, rtn_pos, rtn_neg = NA, nestr_datf, yes_val = T)
```

Arguments

```
inptf_datf is the input dataframe (1b)
rtn_pos is the special value (1a)
rtn_neg is the special value (3a)
nestr_datf is the dataframe (2b)
yes_val is the special value (2a)
```

Examples

```
print(nestr_datf2(inptf_datf=data.frame(c(1, 2, 1), c(1, 5, 7)), rtn_pos="yes",
rtn_neg="no", nestr_datf=data.frame(c(TRUE, FALSE, TRUE), c(FALSE, FALSE, TRUE)), yes_val
# c.1..2..1. c.1..5..7.
#1
```

```
#1 yes no
#2 no no
#3 yes yes
```

Description

Return a vector composed of pasted elements from the input dataframe at the same index.

Usage

```
paste_datf(inpt_datf, sep = "")
```

12 see_datf

Arguments

```
inpt_datf is the input dataframe
sep is the separator between pasted elements, defaults to ""
```

Examples

```
print(paste_datf(inpt_datf=data.frame(c(1, 2, 1), c(33, 22, 55))))
#[1] "133" "222" "155"
```

see_datf

see_datf

Description

Allow to return a dataframe with special value cells (ex: TRUE) where the condition entered are respected and another special value cell (ex: FALSE) where these are not

Usage

```
see_datf(
  datf,
  condition_l,
  val_l,
  conjunction_l = c(),
  rt_val = TRUE,
  f_val = FALSE
)
```

Arguments

f_val

datf is the input dataframe is the vector of the possible conditions ("==", ">", "<", "!=", "%%", "reg", condition 1 "not_reg", "sup_nchar", "inf_nchar", "nchar") (equal to some elements in a vector, greater than, lower than, not equal to, is divisible by, the regex condition returns TRUE, the regex condition returns FALSE, the length of the elements is strictly superior to X, the length of the element is strictly inferior to X, the length of the element is equal to one element in a vector), you can put the same condition n times. is the list of vectors containing the values or vector of values related to condival l tion_l (so the vector of values has to be placed in the same order) conjunction_1 contains the and or conjunctions, so if the length of condition_1 is equal to 3, there will be 2 conjunctions. If the length of conjunction_l is inferior to the length of condition_l minus 1, conjunction_l will match its goal length value with its last argument as the last arguments. For example, c("&","|","&") with a goal length value of 5 \rightarrow c("&", "|", "&", "&", "&") is a special value cell returned when the conditions are respected rt_val

is a special value cell returned when the conditions are not respected

swipr 13

Details

This function will return an error if number only comparative conditions are given in addition to having character values in the input dataframe.

Examples

```
datf1 <- data.frame(c(1, 2, 4), c("a", "a", "zu"))</pre>
print(see_datf(datf=datf1, condition_l=c("nchar"), val_l=list(c(1))))
    X1
           X2
#1 TRUE TRUE
#2 TRUE TRUE
#3 TRUE FALSE
print(see_datf(datf=datf1, condition_l=c("=="), val_l=list(c("a", 1))))
    Х1
          X2
#1 TRUE TRUE
#2 FALSE TRUE
#3 FALSE FALSE
print(see_datf(datf=datf1, condition_l=c("nchar"), val_l=list(c(1, 2))))
    Х1
          X2
#1 TRUE TRUE
#2 TRUE TRUE
#3 TRUE TRUE
print(see_datf(datf=datf1, condition_l=c("not_reg"), val_l=list("[a-z]")))
    Х1
          X2
#1 TRUE FALSE
#2 TRUE FALSE
#3 TRUE FALSE
```

swipr swipr

Description

Returns an ordered dataframes according to the elements order given. The input datafram has two columns, one with the ids which can be bonded to multiple elements in the other column.

Usage

```
swipr(inpt_datf, how_to = c(), id_w = 2, id_ids = 1)
```

14 unique_datf

Arguments

inpt_datf	is the input dataframe
how_to	is a vector containing the elements in the order wanted
id_w	is the column number or the column name of the elements
id_ids	is the column number or the column name of the ids

Examples

```
datf <- data.frame("col1"=c("Af", "Al", "Al", "Al", "Arg", "Arg", "Arg", "Arm", "Arm", "A
       "col2"=c("B", "B", "G", "S", "B", "S", "G", "B", "G", "B"))
print(swipr(inpt_datf=datf, how_to=c("G", "S", "B")))
  coll col2
1
    Af B
    Al
2
        G
3
    Al
         S
4
    Al
         В
5
   Arg
         G
6
   Arg
          S
   Arg
          В
8
   Arm
          G
9
   Arm
          В
10
   Al
          В
```

unique_datf

unique_datf

Description

Returns the input dataframe with the unique columns or rows.

Usage

```
unique_datf(inpt_datf, col = FALSE)
```

Arguments

```
inpt_datf is the input dataframe

col is a parameter that specifies if the dataframe returned should have unique columns or rows, defaults to F, so the dataframe returned by default has unique rows
```

```
datf1 <- data.frame(c(1, 2, 1, 3), c("a", "z", "a", "p"))
print(datf1)

c.1..2..1..3. c..a...z...a...p.. c.1..2..1..3..1
1 1 1 1 1</pre>
```

val_replacer 15

```
Z
                                                    2
3
                                    а
              3
print (unique_datf(inpt_datf=datf1))
   c.1..2..1..3. c..a...z...a...p..
#1
              1
#2
               2
               3
#4
                                    р
datf1 <- data.frame(c(1, 2, 1, 3), c("a", "z", "a", "p"), c(1, 2, 1, 3))</pre>
print(datf1)
 c.1..2..1..3. c..a...z...a...p..
1
         1
2
             2
3
             1
             3
print(unique_datf(inpt_datf=datf1, col=TRUE))
  cur_v cur_v
#1
      1
      2
#2
            Z
#3
      1
            а
#4
      3
```

val_replacer

Description

Allow to replace value from dataframe to another one.

Usage

```
val_replacer(datf, val_replaced, val_replacor = TRUE)
```

Arguments

```
datf is the input dataframe

val_replaced is a vector of the value(s) to be replaced

val_replacor is the value that will replace val_replaced
```

16 vec_in_datf

#1	1	NA
#2	004	FALSE
#3	NA	NA
#4	FALSE	NA

Description

Allow to get if a vector is in a dataframe. Returns the row and column of the vector in the dataframe if the vector is contained in the dataframe.

Usage

```
vec_in_datf(
  inpt_datf,
  inpt_vec = c(),
  coeff = 0,
  stop_untl = 1,
  conventional = FALSE
)
```

Arguments

```
inpt_datf is the input dataframe
inpt_vec is the vector that may be in the input dataframe
coeff is the "slope coefficient" of inpt_vec
stop_untl is the maximum number of the input vector the function returns, if in the dataframe
conventional is if a positive slope coefficient means that the vector goes upward or downward
```

```
datf1 \leftarrow data.frame(c(1:5), c(5:1), c("a", "z", "z", "z", "a"))
print(datf1)
  c.1.5. c.5.1. c..a...z...z...z...a..
#1
       1
              5
#2
        2
               4
#3
        3
               3
             2
#4
        4
                                           Z
             1
        5
#5
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(5, 4, "z"), coeff=1))
#NULL
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(5, 2, "z"), coeff=1))
#[1] 5 1
```

vlookup_datf 17

```
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(3, "z"), coeff=1))
#[1] 3 2
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(4, "z"), coeff=-1))
#[1] 2 2
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(2, 3, "z"), coeff=-1))
#[1] 2 1
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(5, 2, "z"), coeff=-1, conventional=TRUE))
#[1] 5 1
datf1[4, 2] <- 1
print(vec_in_datf(inpt_datf=datf1, inpt_vec=c(1, "z"), coeff=-1, conventional=TRUE, stop_#[1] 4 2 5 2</pre>
```

vlookup_datf

vlookup_datf

Description

Alow to perform a vlookup on a dataframe

Usage

```
vlookup_datf(datf, v_id, col_id = 1, included_col_id = "yes")
```

Arguments

```
datf is the input dataframe

v_id is a vector containing the ids

col_id is the column that contains the ids (default is equal to 1)

included_col_id

is if the result should return the col_id (default set to yes)
```

```
datf1 \leftarrow data.frame(c("az1", "az3", "az4", "az2"), c(1:4), c(4:1))
print(vlookup_datf(datf=datf1, v_id=c("az1", "az2", "az3", "az4")))
    c..az1....az3....az4....az2.. c.1.4. c.4.1.
#2
                                az1
                                         1
#4
                                                 1
                                az2
                                         4
#21
                                         2
                                                 3
                                az3
#3
                                az4
```

18 wider_datf

Description

Takes a dataframe as an input and the column to split according to a seprator.

Usage

```
wider_datf(inpt_datf, col_to_splt = c(), sep_ = "-")
```

Arguments

```
inpt_datf is the input dataframe
col_to_splt is a vector containing the number or the colnames of the columns to split according to a separator
sep_ is the separator of the elements to split to new columns in the input dataframe
```

```
datf1 \leftarrow data.frame(c(1:5), c("o-y", "hj-yy", "er-y", "k-ll", "ooo-mm"), c(5:1))
datf2 <- data.frame("col1"=c(1:5), "col2"=c("o-y", "hj-yy", "er-y", "k-ll", "ooo-mm"))</pre>
print(wider_datf(inpt_datf=datf1, col_to_splt=c(2), sep_="-"))
        pre_datf X.o. X.y.
              "o" "y" 5
#о-у
        1
#hj-yy 2
               "hj" "yy" 4
               "er" "y" 3
"k" "11" 2
#er-y 3
#k-11 4
#000-mm 5
               "000" "mm" 1
print(wider_datf(inpt_datf=datf2, col_to_splt=c("col2"), sep_="-"))
        pre_datf X.o. X.y.
               "о" "у"
#о-у
       1
               "hj" "yy"
"er" "y"
"k" "ll"
#hj-yy 2
#er-y 3
#k-11 4
               "000" "mm"
#000-mm 5
```

Index

```
\verb|colins_datf|, 2
\verb|cumulated_rows|, 3
cumulated_rows_na, 3
cut_v,4
diff_datf, 5
groupr_datf,5
id_keepr,6
insert_datf, 7
intersect\_mod, 8
nestr_datf1, 10
nestr_datf2, 11
paste_datf, 11
see\_datf, 12
swipr, 13
{\tt unique\_datf}, 14
val\_replacer, 15
vec_in_datf, 16
vlookup_datf, 17
{\tt wider\_datf}, 18
```