Package 'edm1.sequence'

July 29, 2024

Title	Set	of	tools	to	manipul	ate	time	series
-------	-----	----	-------	----	---------	-----	------	--------

Version 2.0.0.0

Description Provides set of functions to manipulate time series dataset; create variables that designates the value of the individual at n-x, handle missing values...

License GPL (==3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Imports stringr

Contents

Index		9
	sequence_na_med2	7
	sequence_na_med1	
	sequence_na_mean2	5
	sequence_na_mean1	4
	historic_sequence2	2
	historic_sequence1	

historic_sequence1 historic_sequence1

Description

Allow to perform a pivot wider on a sequencial dataset (here the type is dataframe), each variable will be dupplicated in a column to show the value to this variable at n-1 for each individual, see examples.

Usage

```
historic_sequence1(inpt_datf, bf_ = 1)
```

Arguments

inpt_datf is the input dataframe
bf_ is the number of previous value of the individual it will search for, see examples

2 historic_sequence2

```
set.seed(123)
var1 < - round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 \leftarrow round(runif(n = 14, min = 14, max = 20))
datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 18,
                       17, 17, 17),
               "individual" = c("oui", "non", "peut1", "peut2", "oui", "peut1", "peut2"),
               "var1" = var1,
                "var2" = var2)
print(datf)
  ids individual var1 var2
  20 oui 106 16
           non 117
  20
2
                     19
         peut1 109 16
3
  20
         peut2 119 19
4
  20
           oui 121
5
  19
                     20
6 19
         peut1 101
                    14
7 19
         peut2 112
                    17
8 18
          oui 120 19
9 18
           non 112 17
10 18
         peut1 110
                    17
11 18
         peut2 121
                     20
12 17
          oui 110
                    17
13 17
          peut1 115
                     18
14 17
         peut2 113
                     17
historic_sequence1(inpt_datf = datf, bf_ = 2)
 id_seq individual var1-1 var1-2 var2-1 var2-2
        oui 121 120 20 19
1
    20
                                NA
2
                   NA
                         112
     20
             non
                                       17
                   101
                                14
3
     20
                         110
                                      17
           peut1
                                17
4
     20
                  112
                         121
                                      20
           peut2
5
                   120
                         110
                                19
     19
                                      17
            oui
                   110
                         115
                                17
                                      18
6
    19
           peut1
           peut2
    19
                  121
                         113
                                20
                                      17
historic_sequence1(inpt_datf = datf, bf_ = 3)
 id_seq individual var1-1 var1-2 var1-3 var2-1 var2-2 var2-3
1
    20
            oui 121 120 110 20 19 17
2
                   NA
                         112
                                      NA
                                             17
                                                  NA
     20
             non
                                NA
                                     14 17
17 20
3
                   101 110 115
    20
          peut1
                                                  18
4
    20
          peut2 112
                         121 113
                                                   17
```

historic_sequence2 3

Description

Allow to perform a pivot wider on a sequencial dataset (here the type is dataframe), each variable will be dupplicated in a column to show the value to this variable at n - 1 for each individual, see examples.

Usage

```
historic_sequence2(inpt_datf, bf_ = 1)
```

Arguments

```
is the input dataframe
inpt_datf
bf_
                  is the number of previous value of the individual it will search for, see examples
```

Examples

19

```
set.seed(123)
var1 < - round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 \leftarrow round(runif(n = 14, min = 14, max = 20))
datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 18,
                            17, 17, 17),
                  "individual" = c("oui", "non", "peut1", "peut2",
                                    "oui", "peut1", "peut2"),
                  "var1" = var1,
                  "var2" = var2)
print(datf)
   ids individual var1 var2
1
   20 oui 106 16
2
   20
             non 117 19
          peut1 109 16
3
   20
          peut2 119 19
   20
5
  19
            oui 121
         peut1 101
peut2 112
6
  19
                         14
7
   19
                         17
           oui 120
8
   18
                         19
9
             non 112
                         17
   18
10 18
          peut1 110
                         17
11 18
          peut2 121
                         2.0
12
   17
             oui 110
                         17
13
   17
            peut1
                   115
                         18
   17
            peut2
                   113
                         17
print(historic_sequence2(inpt_datf = datf, bf_ = 2))
  id_seq individual var1-0 var1-1 var1-2 var2-0 var2-1 var2-2
1
              oui 106 121 120
                                           16 20
                                                          19
      20
2
      20
                       117
                                      112
                                                            17
                               NA
                                              19
                                                     NA
                non
      20
                      109
                            101
                                     110
                                                            17
3
                                             16
                                                     14
             peut1
4
      20
              peut2 119 112 121
                                             19
                                                    17
                                                            20

    oui
    121
    120
    110
    20
    19

    peut1
    101
    110
    115
    14
    17

    peut2
    112
    121
    113
    17
    20

5
      19
                                                            17
6
     19
            peut1
                                                            18
```

17

4 sequence_na_mean1

```
print(historic_sequence2(inpt_datf = datf, bf_ = 3))
 id_seq individual var1-0 var1-1 var1-2 var1-3 var2-0 var2-1 var2-2 var2-3
1
                       121 120
                                   110 16 20 19 17
     2.0
            oui
                 106
2
     20
                   117
                               112
                                            19
                                                  NA
                                                        17
             non
                         NA
                                     NA
                                                              NA
3
                  109
                              110
                                            16
                                                        17
     20
                         101
                                     115
                                                 14
                                                             18
           peut1
4
     20
                  119
                       112 121
                                    113
                                           19
                                                 17
                                                        20
                                                             17
           peut2
```

Description

In a dataframe generated by the function historic_sequence1, convert all NA to the mean of the values at the same variable for the individual at the id where the NA occurs, see examples (only accepts numeric variables)

Usage

```
sequence_na_mean1(inpt_datf, bf_, step = 1)
```

Arguments

inpt_datf is the input dataframe

```
set.seed(123)
var1 < - round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 \leftarrow round(runif(n = 14, min = 14, max = 20))
datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 18,
17, 17, 17),
"individual" = c("oui", "non", "peut1", "peut2",
"oui", "peut1", "peut2"),
"var1" = var1,
"var2" = var2)
datf <- historic_sequence1(inpt_datf = datf, bf_ = 2)</pre>
datf[3, 4] <- NA
datf[6, 4] <- NA
datf[1, 3] <- NA
print(datf)
  id_seq individual var1-1 var1-2 var2-1 var2-2
1
                                  20
     20
               oui
                    NA
                           120
                                           19
2
      20
                             112
                                            17
                       NA
                                     NA
               non
     20
                                            17
3
                      101
                                     14
             peut1
                              NA
4
     20
             peut2
                     112
                             121
                                     17
                                            20
5
     19
              oui
                      120
                             110
                                    19
                                            17
6
     19
             peut1
                      110
                             NA
                                    17
                                            18
7
     19
             peut2
                     121
                             113
                                    20
                                            17
```

sequence_na_mean2 5

```
print(sequence_na_mean1(inpt_datf = datf, bf_ = 2))
 id_seq individual var1-1 var1-2 var2-1 var2-2
            oui 115 120.0 20
1
     2.0
                   112 112.0
                                17
2
     20
             non
                                       17
                   101 105.5
3
                                14
     2.0
                                       17
          peut1
4
     2.0
           peut2
                  112 121.0
                                17
                                       2.0
                  120 110.0
5
     19
                               19
                                      17
            oui
    19 peut1
                  110 105.5
6
                               17
                                      18
     19
           peut2 121 113.0
                               20
                                      17
```

```
sequence_na_mean2 sequence_na_mean2
```

Description

In a dataframe generated by the function historic_sequence1, convert all NA to the mean of the values at the same variable for the individual at the id where the NA occurs, see examples (only accepts numeric variables)

Usage

```
sequence_na_mean2(inpt_datf, bf_, step = 1)
```

Arguments

```
inpt_datf is the input dataframe
```

```
set.seed(123)
var1 < - round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 \leftarrow round(runif(n = 14, min = 14, max = 20))
datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 18,
17, 17, 17),
"individual" = c("oui", "non", "peut1", "peut2",
"oui", "peut1", "peut2"),
"var1" = var1,
"var2" = var2)
datf <- historic_sequence2(inpt_datf = datf, bf_ = 2)</pre>
datf[3, 4] <- NA
datf[6, 4] \leftarrow NA
datf[1, 3] <- NA
print(datf)
 id_seq individual var1-0 var1-1 var1-2 var2-0 var2-1 var2-2
1
     20
             oui NA 121 120
                                       16 NA
2
     20
                     117
                            NA 112
                                         19
                                                NA
                                                       17
              non
3
     20
            peut1
                    109
                           NA 110
                                         16
                                               14
                                                       17
            peut2 119 112 121
4
     20
                                         19
                                               17
                                                       20
5
     19
             oui 121 120 110
                                         20
                                                19
                                                       17
```

6 sequence_na_med1

```
peut1
                     101
                            NA
                                  115
                                          14
                                                 17
                                                       18
7
     19
             peut2
                     112
                            121
                                   113
                                          17
                                                 20
                                                       17
print(sequence_na_mean2(inpt_datf = datf, bf_ = 2))
 id_seq individual var1-0
                           var1-1 var1-2 var2-0 var2-1 var2-2
1
     20
                   117 121.0000
                                   120
                                           16
                                                  18
              oui
                    117 114.5000
2
     20
                                    112
                                           19
                                                  18
                                                         17
              non
3
     20
            peut1
                    109 108.3333
                                    110
                                           16
                                                  14
                                                         17
                    119 112.0000
4
     20
            peut2
                                   121
                                           19
                                                  17
                                                         20
5
     19
                    121 120.0000
                                   110
                                          20
                                                         17
             oui
                                                  19
6
     19
             peut1
                    101 108.3333
                                   115
                                           14
                                                  17
                                                         18
7
     19
            peut2 112 121.0000
                                   113
                                           17
                                                  20
                                                         17
```

```
sequence_na_med1 sequence_na_med1
```

Description

In a dataframe generated by the function historic_sequence1, convert all NA to the median of the values at the same variable for the individual at the id where the NA occurs, see examples (only accepts numeric variables)

Usage

```
sequence_na_med1(inpt_datf, bf_, step = 1)
```

Arguments

inpt_datf	is the input dataframe
bf_	is how at how many n 1 we look for the value of the variables for the individual at time index \boldsymbol{n}
step	is the base step for the time indexes, step of one year (or any time unit), two year (or any time unit)?

```
set.seed(123)
var1 <- round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 <- round(runif(n = 14, min = 14, max = 20))

datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 18, 17, 17, 17),
  "individual" = c("oui", "non", "peut1", "peut2",
  "oui", "peut1", "peut2"),
  "var1" = var1,
  "var2" = var2)
datf <- historic_sequence1(inpt_datf = datf, bf_ = 2)
datf[3, 4] <- NA
datf[6, 4] <- NA
datf[1, 3] <- NA</pre>
```

sequence_na_med2 7

```
print(datf)
 id_seq individual var1-1 var1-2 var2-1 var2-2
       oui NA 120 20 19
    2.0
                  NA
                        112
                              NA
2
    20
                                     17
            non
3
                  101
                              14
    2.0
                        NA
                                     17
          peut1
           peut2 112
                      121
                              17
4
    20
                                    2.0
                 120 110
                              19
5
    19
           oui
                                    17
           peut1
6
    19
                  110
                        NA
                              17
                                    18
                              20
    19
           peut2
                 121
                        113
                                    17
print(sequence_na_med1(inpt_datf = datf, bf_ = 2))
 id_seq individual var1-1 var1-2 var2-1 var2-2
1
    20 oui 115 120.0 20
            non 112 112.0
                              17
2
    2.0
                                     17
                              14
3
    20
         peut1 101 105.5
                                     17
4
    20
          peut2 112 121.0
                              17
                                     20
                              19
5
    19
                  120 110.0
                                     17
           oui
        peut1
                  110 105.5
121 113.0
                              17
6
    19
                                     18
7
    19
           peut2
                               20
                                     17
```

```
sequence_na_med2 sequence_na_med2
```

Description

In a dataframe generated by the function historic_sequence2, convert all NA to the median of the values at the same variable for the individual at the id where the NA occurs, see examples (only accepts numeric variables)

Usage

```
sequence_na_med2(inpt_datf, bf_, step = 1)
```

Arguments

inpt_datf	is the input dataframe
bf_	is how at how many n 1 we look for the value of the variables for the individual at time index \boldsymbol{n}
step	is the base step for the time indexes, step of one year (or any time unit), two year (or any time unit)?

```
set.seed(123)
var1 <- round(runif(n = 14, min = 100, max = 122))
set.seed(123)
var2 <- round(runif(n = 14, min = 14, max = 20))
datf <- data.frame("ids" = c(20, 20, 20, 20, 19, 19, 19, 18, 18, 18, 17, 17, 17),
"individual" = c("oui", "non", "peut1", "peut2",</pre>
```

8 sequence_na_med2

```
"oui", "peut1", "peut2"),
"var1" = var1,
"var2" = var2)
datf <- historic_sequence2(inpt_datf = datf, bf_ = 2)</pre>
datf[3, 4] <- NA
datf[6, 4] <- NA
datf[1, 3] <- NA
print(datf)
  id_seq individual var1-0 var1-1 var1-2 var2-0 var2-1 var2-2
             oui NA 121 120 16 20 19
2
      20
              non 117
                            NA 112
                                          19
                                                 NA
                                                       17
3
                    109
                            NA 110
     20
            peut1
                                          16
                                                14
                                                       17
            peut2 119
     20
                          112 121
                                         19 17
                                                       20
4
                   121 120 110 20 19
101 NA 115 14 17
5
                                                       17
     19
             oui
     19
                                                       18
6
           peut1
     19
           peut2 112
                            121 113
                                          17
                                                  20
                                                       17
print(sequence_na_med2(inpt_datf = datf, bf_ = 2))
  id_seq individual var1-0 var1-1 var1-2 var2-0 var2-1 var2-2

    oui
    120
    121.0
    120

    non
    117
    114.5
    112

    peut1
    109
    109.0
    110

     20
                                        16
                                               20
2
      20
                                           19
                                                  18
                                                        17
                                                       17
3
     20
                                           16
                                                 14
                     119 112.0
                                  121
                                                       20
4
     20
                                          19
                                                 17
            peut2
                     121 120.0
5
                                 110
                                           20
                                                       17
     19
                                                 19
             oui
                     101 109.0 115
                                          14
                                                 17
                                                       18
    19
6
             peut1
             peut2 112 121.0 113
    19
                                          17
                                                 20
                                                       17
```

Index

```
historic_sequence1, 1
historic_sequence2, 2
sequence_na_mean1, 4
sequence_na_mean2, 5
sequence_na_med1, 6
sequence_na_med2, 7
```