

# Pakiet filtr

Anita\_Kamińska

16 02 2021

## Witam w pakiecie **filtr**

Pakiet filtr zawiera dwie funkcje: **filtr1** i **filtr2**.  
Funkcje te działają na ramce danych zawierającej 3 kolumny:

- data (rok-miesiąc-dzień),
- cena,
- ilość,

Pakiet zawiera przykładową ramkę danych **dane**:

- data z każdego dnia od 2014-11-01 do 2021-02-05,
- cena z rozkładu normalnego o średniej 1000 i odchylenie standardowe 100,
- ilość zbudowana losowo spośród liczb od 1 do 2289 - losowanie liczb ze zwracaniem.

Są to dane transakcyjne zawierające 2289 obserwacji.

Pakiet wczytujemy następująco:

```
library(filtr)
```

```
## Witamy w pakiecie filtr. Przygotowanym w celu zaliczenia przedmiotu Analiza danych w srodowisku R.
```

Pojawienie się przywitania informuje nas, że pakiet został wczytany.

## Filtr 1 i Filtr 2

Pierwsza funkcja (**filtr1(dane, kolumna, limit\_dolny, limit\_górny, flaga=FALSE)**) dokonuje selekcji tych rekordów, dla których wskazana kolumna (np. cena) jest między **limit\_dolny** a **limit\_górny**.

Druga funkcja (**filtr2(dane, kolumna, kwantyl\_dolny, kwantyl\_górny, flaga=FALSE)**) oblicza kwantyle rozkładu wartości ze wskazanej kolumny (odpowiednio rzędu **kwantyl\_dolny** i **kwantyl\_górny**) i dokonuje selekcji tych wartości z zadanej kolumny, które się mieszczą pomiędzy.

Funkcje te zwracają

- dla **flaga=FALSE** zawężona ramka (tylko rekordy spełniające warunki filtru),
- dla **flaga=TRUE** z pełną ramką z dodatkową kolumną **FLAGA** (1, gdy wiersz/rekord spełnia kryterium filtru, 0 gdy nie spełnia).

## Zastosowanie

```
filtr1(dane, "cena", 600, 730)
```

```
##          data    cena ilosc
## 1 2016-06-08 680.46  1021
## 2 2017-03-02 725.80  1390
## 3 2018-01-05 717.25   104
## 4 2018-05-31 716.13  1411
## 5 2019-01-05 704.57   280
## 6 2019-04-02 659.00   241
## 7 2019-11-05 709.52  1682
## 8 2020-05-10 700.73  1503
## 9 2020-10-25 718.37  2165
```

```
filtr1(dane,"ilosc", 400, 410)
```

```
##          data    cena ilosc
## 1 2015-09-29 967.07   409
## 2 2015-10-05 1076.78  410
## 3 2016-08-01 953.08  405
## 4 2016-09-16 857.12  402
## 5 2017-11-24 1056.98  408
## 6 2017-12-02 953.26  402
## 7 2019-04-30 1122.30  403
## 8 2019-08-26 951.21  402
## 9 2020-09-15 885.10  404
## 10 2020-10-06 1129.07 406
```

```
head(filtr1(dane,"ilosc", 800, 1750, TRUE))
```

```
##          data    cena ilosc flaga
## 1 2014-11-01 1115.87   103     0
## 2 2014-11-02 925.71  1495     1
## 3 2014-11-03 916.00  1735     1
## 4 2014-11-04 783.34   890     1
## 5 2014-11-05 999.90  1514     1
## 6 2014-11-06 956.23  1000     1
```

```
filtr2(dane,"ilosc", 0.5, 0.5)
```

```
##          data    cena ilosc
## 1 2015-07-07 1122.30  1135
## 2 2018-11-22 1027.50  1135
## 3 2019-09-26 937.31  1135
## 4 2020-07-24 1062.81  1135
```

```
head(filtr2(dane,"cena", 0.25, 0.26))
```

```
##          data    cena ilosc
## 1 2014-11-28 933.01   326
## 2 2015-08-16 933.40  1253
## 3 2015-09-02 930.43  2273
## 4 2015-12-22 933.35  1263
## 5 2016-03-30 932.27  1716
## 6 2016-04-16 930.04   45
```

## Błędy

Przykładowe reakcje programu na błędne/nieprawidłowe dane:

```
filtr1(dane,"cena", 700, 600)
```

```
## Error in .filtr(dane, kolumna, limit_dolny, limit_gorny, flaga, 1):  
## Ograniczenie dolne musi byc mniejsze od ograniczenia gornego
```

```
filtr2(dane,"cena", 0.5, 2)
```

```
## Error in .filtr(dane, kolumna, kwantyl_dolny, kwantyl_gorny, flaga, 2):  
## Kwantyle z przedzialu [0,1]
```

```
filtr2(dane,"ilosc", c(0.5,0.6), 0.7)
```

```
## Error in .filtr(dane, kolumna, kwantyl_dolny, kwantyl_gorny, flaga, 2):  
## Podaj tylko jedna wartosc - ograniczenie dolne
```

```
filtr2(dane,"okno", 0.5, 1)
```

```
## Error in .filtr(dane, kolumna, kwantyl_dolny, kwantyl_gorny, flaga, 2):  
## Nie ma takiej kolumny
```

```
filtr1(iris,"cena", 700, 900)
```

```
## Error in .filtr(dane, kolumna, limit_dolny, limit_gorny, flaga, 1):  
## Dane w zlym formacie.
```

```
filtr1("dane","cena", 700, 900)
```

```
## Error in .filtr(dane, kolumna, limit_dolny, limit_gorny, flaga, 1):  
## Klasa danych to nie data frame!
```

```
filtr1(dane, "cena", 700, 900, 5)
```

```
## Error in .filtr(dane, kolumna, limit_dolny, limit_gorny, flaga, 1):  
## Flaga nie jest wartoscia logiczna
```