

# Pakiet vdmR

Generating Web-Based Visual Data Mining Tools with R

Eliza Kaczorek

9 maja 2018

# Przykład

```
1 > install.packages("vdmR")
2 > library("vdmR")
3
4 > data("vsfuk2012", package = "vdmR")
5 > head(vsfuk2012[,1:5])
6
7   CityCode CityName Type FertilityRate MortalityRate
8 1    40133  Chuouou Ward           8.9           6.0
9 2    40132   Hakata Ward          10.4           6.9
10 3    40134  Minami Ward           9.9           7.3
11 4    40137   Sawara Ward           9.9           7.0
12 5    40131  Higashi Ward          10.5           7.1
13 6    40135    Nishi Ward          10.2           7.4
14
15 # KROK 1
16 > vscat(MortalityRate, FertilityRate, vsfuk2012, "scat01", "vsfuk2012")
17 > vhist(MarriageRate, vsfuk2012, "hist01", "vsfuk2012")
18
19 # KROK 2
20 > vlaunch(vsfuk2012, "main", "vsfuk2012")
```

# vlaunch()

```
1 vlaunch(data, name, tag, path = tempdir(), iframe =  
    FALSE, browse = TRUE)
```

- ❶ **data** - data frame lub inny zbiór danych
- ❷ **name** - pierwszy człon nazwy dla głównego okna
- ❸ **tag** - drugi człon nazwy dla głównego okna, tylko wykresy z tym tagiem zostaną zawarte w tym oknie
- ❹ **path** - ścieżka dla plików
- ❺ **iframe** - jeśli TRUE, to wszystkie wykresy zostaną wyświetlone w oknie głównym
- ❻ **browse** - jeśli TRUE, po wywołaniu funkcji wyświetla automatycznie wynik w przeglądarce

# vscat()

```
1 vscat(x, y, data, name, tag, path = tempdir(), ...)
```

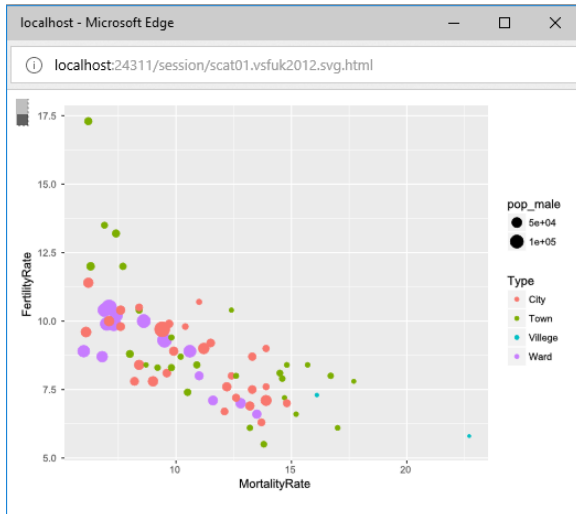
- ❶ **x, y** - nazwy kolumn ze zbioru **data**
- ❷ **data** - data frame lub inny zbiór danych
- ❸ ... - argumenty przekazywane do funkcji `aes()` z pakietu `ggplot2`

## Przykład

```
1 vscat(MortalityRate, FertilityRate, vsfuk2012, "scat01", "vsfuk2012",  
2     color = Type, size = pop_male)  
3 vlaunch(vsfuk2012, "main", "vsfuk2012")
```

vlaunch  
vscat  
vhist  
vpvp  
vcmmap

# vscat()



# vhist()

```
1 vhist(x, data, name, tag, path = tempdir(), ...)
```

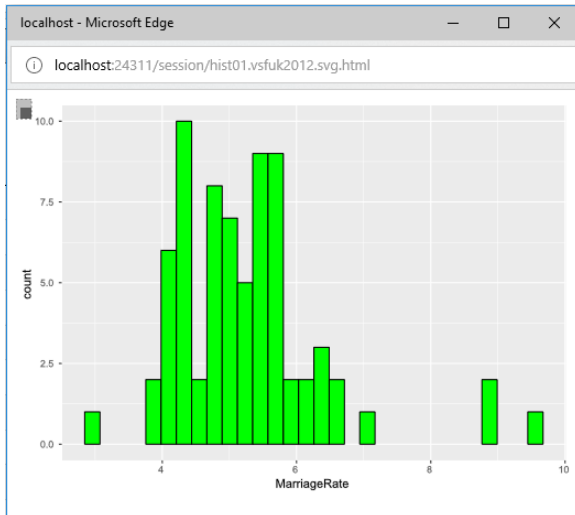
- ❶ **x** - nazwa kolumny ze zbioru **data**
- ❷ **data** - data frame lub inny zbiór danych
- ❸ ... - parametry graficzne do narysowania histogramu

## Przykład

```
1 vhist(MarriageRate, vsfuk2012, "hist01", "vsfuk2012",  
2       fill = I("green"), color = I("black"))  
3 vlaunch(vsfuk2012, "main", "vsfuk2012")
```

vlaunch  
vscat  
vhist  
vpvp  
vmap

# vhist()



## vpvp - parallel coordinate plot

```
1 vpvp(data, columns, name, tag, path = tempdir(),  
      groupColumn = NULL,  
2 scale = "std", alphaLines = 0.5, missing = "exclude")
```

- ❶ **data** - data frame lub inny zbiór danych
- ❷ **columns** - kolumny uwzględniane na wykresie
- ❸ **groupColumn** - kolorowanie w podgrupach wyznaczonych przez daną kolumnę
- ❹ **scale** - metoda wykorzystana do skalowania zmiennych (std, uniminmax, ...)
- ❺ **alphaLines** - stopień przezroczystości linii
- ❻ **missing** - metoda wypełniania braków danych (exclude, mean, median, ...)

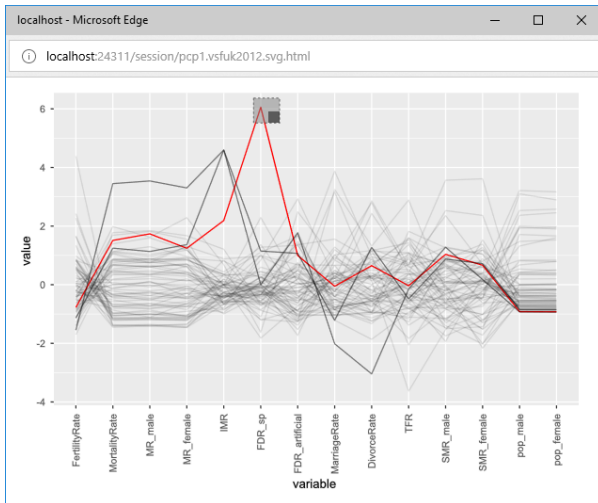


## vpvp

```
1 vpcp(vsfuk2012, 4:17, "pcp1", "vsfuk2012",  
2     alphaLines = 0.1, scale = "std",  
3     missing = "mean")  
4  
5 vlaunch(vsfuk2012, "main", "vsfuk2012")
```

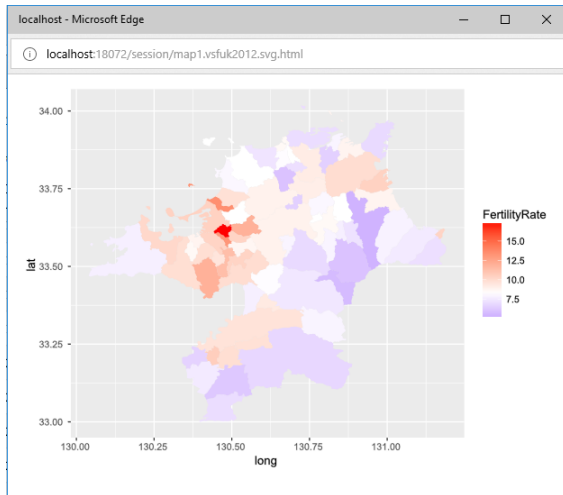
vlaunch  
vscat  
vhist  
vpvp  
vcmap

vpcp



vlaunch  
vscat  
vhist  
vpvp  
vcmap

## vcmap - choropleth map / kartogram



- + łatwość i szybkość użycia
- + interaktywność
- + estetyczna prezentacja wyników analiz
- nie poradzi sobie z dużymi zbiorami danych
- małe możliwości analiz

## Bibliografia

- [1] Tomokazu Fujino, *vdmR: Generating Web-Based Visual Data Mining Tools with R*, 2017.

Dostęp:

<https://www.jstatsoft.org/article/view/v082i06>.