Nauka dla społeczeństwa

Katedra i Zakład Informatyki i Statystyki

dr Justyna Marcinkowska

data przygotowania: 2024-05-05

Zalecenia dotyczące warsztatów – część 4 – kod w R cd. :

*Ilustracja graficzna*

*w publikacji naukowej*

*(język R)*

# Ćwiczenia praktyczne na warsztatach:

• Omówienie podstawowej składni języka R,

• Przygotowanie do pracy z danymi: import, tworzenie data frame,

• Wstęp do analizy danych i graficznej wizualizacji:

* przetwarzanie i analiza danych (biblioteka dplyr),
* tworzenie bogatych graficznie, typowych i nietypowych wykresów na podstawie warstw (biblioteka ggplot2)

• Przygotowanie wykresów pod wymogi czasopism naukowych, eksport wykresów wysokiej rozdzielczości



Wstęp do R cd.

# przykład - wektory

R jest mocno nastawiony na wektory - wypróbuj

#wszystko jest wektorem   
x[1]

[1] -12

napis<-"to jest opis mojego projektu"   
napis[1]

[1] "to jest opis mojego projektu"

# przykład -zakładamy bazę

Tworzymy pierwszą ramkę danych. Baza będzie składać się z 3 wektorów - czyli inaczej zmiennych: imiona, wzrost, waga

1. Tworzymy wektory
2. Łączymy je w jedną ramkę danych (musimy mieć wektory równej długości, co oznacza żadnych braków danych)

imiona<-c('Paulina', 'Dawid', 'Bruno', 'Jakub', 'Michał', 'Weronika','Michał','Dominik','Maciej','Jakub','Jan','Krzysztof')   
wzrost<-c(156 ,160 ,164 ,168, 172 ,176 ,180 ,184, 188 ,192 ,190,200)   
waga<- c(45 ,66 ,88 ,94, 72 ,54 ,93 ,90, 74 ,87 ,120 ,97)   
#gdyby połączyć te trzy w jeden wektor to dostaniemy je wrzucone do jednego dużego wektora:   
c(imiona,wzrost,waga)

[1] "Paulina" "Dawid" "Bruno" "Jakub" "Michał" "Weronika"   
 [7] "Michał" "Dominik" "Maciej" "Jakub" "Jan" "Krzysztof"  
[13] "156" "160" "164" "168" "172" "176"   
[19] "180" "184" "188" "192" "190" "200"   
[25] "45" "66" "88" "94" "72" "54"   
[31] "93" "90" "74" "87" "120" "97"

#chcemy uzyskać ramkę danych używając funkcji data.frame():   
baza\_grupy<-data.frame(imiona,wzrost,waga)   
baza\_grupy

imiona wzrost waga  
1 Paulina 156 45  
2 Dawid 160 66  
3 Bruno 164 88  
4 Jakub 168 94  
5 Michał 172 72  
6 Weronika 176 54  
7 Michał 180 93  
8 Dominik 184 90  
9 Maciej 188 74  
10 Jakub 192 87  
11 Jan 190 120  
12 Krzysztof 200 97

# Przykład - używamy funkcji:

1. Zliczamy kolumny itp.
2. Przeglądamy naszą bazę

#ile mamy kolumn w naszej ramce (bazie) grupa\_fakultetu?,   
print("W bazie mamy wierszy.....")

[1] "W bazie mamy wierszy....."

nrow(baza\_grupy)

[1] 12

#wyświetl nazwy kolumn  
print("W bazie mamy następujące kolumny.....")

[1] "W bazie mamy następujące kolumny....."

colnames(baza\_grupy)

[1] "imiona" "wzrost" "waga"

#wyświetl nazwy wierszy   
print("wiersze mają nazwy...")

[1] "wiersze mają nazwy..."

row.names(baza\_grupy)

[1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" "11" "12"

#wyświetl kolumnę waga:  
print("cała kolumna waga.....")

[1] "cała kolumna waga....."

baza\_grupy$waga

[1] 45 66 88 94 72 54 93 90 74 87 120 97

# Przykład - filtrujemy:

#wyświetlamy wzrost dla Michałów:   
baza\_grupy[imiona=="Michał","wzrost"]

[1] 172 180

#wszystkie kolumny dla wiersazy opisanych warunkiem: wzrostu>185   
baza\_grupy[baza\_grupy$wzrost>185,]

imiona wzrost waga  
9 Maciej 188 74  
10 Jakub 192 87  
11 Jan 190 120  
12 Krzysztof 200 97

# zad 3.

Do naszej bazy baza\_grupa dodajemy nową kolumnę BMI

## Jak uczyć się samodzielnie…

Do naszej bazy baza\_grupa dodajemy nową kolumnę BMI przy pomocy funkcji **mutate().**

Wskazówki/podpowiedzi dotyczące zadania:

## wskazówka 1: skąd jest ta funkcja?

Wyłajmy pomoc (help(mutate) lub klawisz “F1” lub ?mutate lub example(mutate)) Sprawdźmy, jaki pakiet będzie nam potrzebny.?

W przeglądarce poszukajmy pomocy o całym pakiecie (tzw. cheat-seet), tu: dplyr

<https://nyu-cdsc.github.io/learningr/assets/data-transformation.pdf>

## wskazówka 2: kodujemy…

Trzeba nadpisać bazę

baza\_grupy<-

mutate(baza\_grupy,BMI=…….***TU napiszemy konkretny wzór !!!*** ….)

## jak wyświetlić tą nową kolumnę?

baza\_grupy$BMI

## jak wyświetlić całą bazę?

View(baza\_grupy)

#wpisz kod: