

Procvičování 3 s řešením

Vektory

1. V RStudiu vytvořte nový projekt, umístěte ho do adresáře cv03.
2. Do adresáře cv03 si ze studijních materiálů stáhněte pracovní prostředí cv03.RData a načtěte ho do R pomocí příkazu `load()`. Pracovní prostředí obsahuje vektor `vyska` s výškami [cm] jedinců v náhodném vzorku naší Rkové populace a jejich jména `jmena`.

```
load(url("http://www.sci.muni.cz/~syrovat/cv03.RData"))
ls()

## [1] "jmena" "vyska"
```

3. Vytvořte vektor `noha` obsahující délky nohou jedinců našeho vzorku Rkové populace (viz tabule).

```
noha<- c(23,25,NA,27,NA,25.5,NA,27,NA,29.6,29,31)
```

4. Vytvořte vektor `pomer` obsahující podíly délky nohou (`noha`) ku výšce těla (`vyska`).

```
pomer<- noha/vyska
pomer

## [1] 0.1446541 0.1461988          NA 0.1542857          NA 0.1517857          NA
## [8] 0.1516854          NA 0.1541667 0.1494845 0.1597938
```

5. Vytvořte textový vektor `sex` odlišující pohlaví měřených jedinců (m/z). (Pohlaví odvoďte ze jmen jedinců).

```
sex<- c("z","z","z","z","z","z","m","m","m","m","m","m")
sex

## [1] "z" "z" "z" "z" "z" "z" "m" "m" "m" "m" "m" "m"
```

6. Vytvořte vektor jedniček (1) a dvojek (2) `barva`, v němž jedničky budou odpovídat pohlaví 'm' a dvojky pohlaví 'z'. Nikoliv však manuálně, ale pomocí logického výrazu.

```
barva<- (sex == "z") + 1
barva

## [1] 2 2 2 2 2 2 1 1 1 1 1 1
```

Logické výrazy

7. Zjistěte, kolik jedinců má nadprůměrnou výšku.

```
prum.vyska<- mean(vyska, na.rm= T)
sum(vyska > prum.vyska, na.rm= T)

## [1] 4
```

8. Zjistěte, kolik jedinců má podprůměrnou nohu.

```
prum.noha<- mean(noha, na.rm= T)
sum(noha < prum.noha, na.rm= T)

## [1] 5
```

9. Zjistěte, kolik žen a kolik mužů obsahuje náš dataset.

```
# ženy
sum(sex == "z")

## [1] 6

# muži
sum(sex == "m")

## [1] 6

# funkce table() zjistí počet vyskytu jedinečných hodnot v objektu:
table(sex)

## sex
## m z
## 6 6
```

10. Zjistěte, kolik mužů má nadprůměrnou nohu a kolik žen má podprůměrnou nohu.

```
# muži
sum(sex == "m" & noha > prum.noha, na.rm= T)

## [1] 3

# ženy
sum(sex == "z" & noha < prum.noha, na.rm= T)

## [1] 4
```

11. Zjistěte, kolik mužů a kolik žen je aspoň v něčem naprůměrných.

```
# muzi
sum(sex == "m" & (vyska > prum.vyska | noha > prum.noha), na.rm= T)

## [1] 4

# zeny
sum(sex == "z" & (vyska > prum.vyska | noha > prum.noha), na.rm= T)

## [1] 0
```

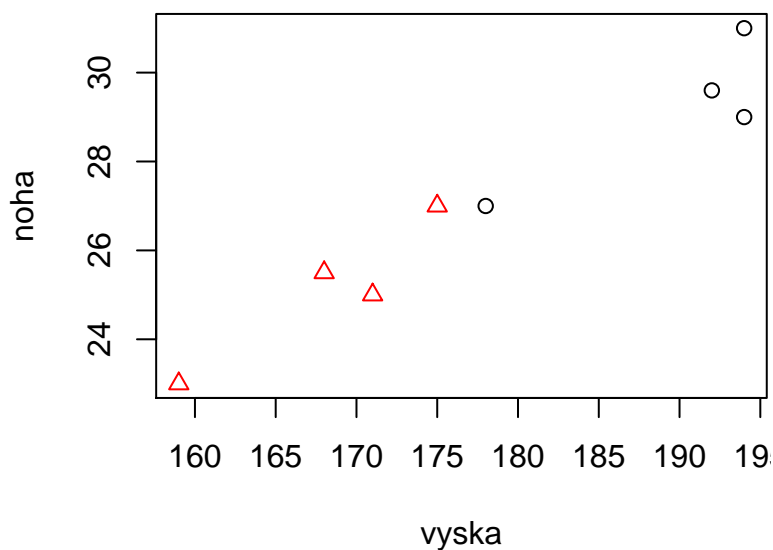
12. Zjistěte, zda všichni se známou výškou jsou vyšší než 160 cm.

```
all(vyska > 160, na.rm= T)

## [1] FALSE
```

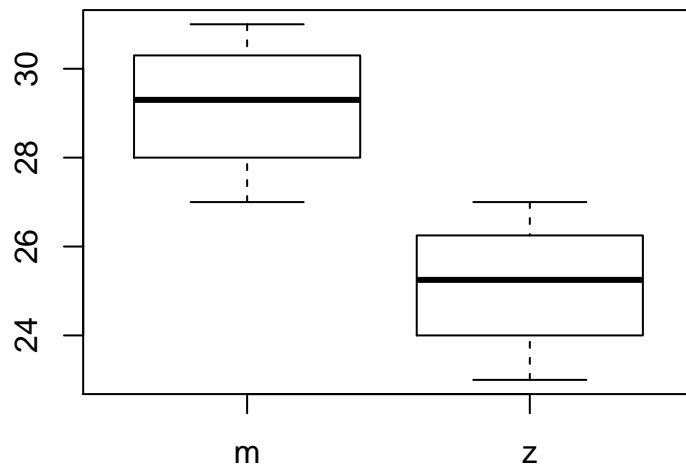
13. Zobrazte bodový graf `plot()` závislosti délky nohy (osa y) na výšce (osa x). V grafu odlište muže a ženy barvou `col=` a tvarem `pch=` symbolů pomocí vektoru `barva`. (Pro definování barev existuje několik možností, jednou z nich je pouhé číslo odkazující na pořadí barvy v barevné paletě `palette()`). A právě jednička a dvojka odkazují na černou a červenou barvu (při defaultním nastavení palety). Symboly se specifikují primárně čísly.

```
palette("default") # abych mel stejnou (default) paletu jako ostatni
par(mar=c(4,4,1,0.1)) # aby obrazek nemel velke okraje
plot(noha ~ vyska, col= barva, pch= barva)
```



14. Zobrazte krabicový graf `boxplot()` závislosti poměru `pomer` na pohlavi `sex`. Použijte tvar formula, tedy `pomer ~ sex`.

```
par(mar=c(4,4,1,0.1)) # aby obrazek nemel velke okraje
boxplot(noha ~ sex)
```



Podvýběry - subscripty

15. Nechejte si vypsát 5. hodnotu vektoru *vyska*.

```
vyska[5]
## [1] NA
```

16. Nechejte si vypsát 8., 2., a 5. hodnotu vektoru *vyska*.

```
vyska[c(8,2,5)]
## [1] 178 171 NA
```

17. Nechejte si vypsát 1., 2., 1., 2., 1. a 1. hodnotu vektoru *noha*.

```
vyska[c(1,2,1,2,1,1)]
## [1] 159 171 159 171 159 159
```

18. Nechejte si vypsát vektor *jmena* v nahodnem poradi.

```
jmena[sample(1:length(jmena))]
```

```
## [1] "Lenka" "Vita" "Klara" "Dominik" "Karolina" "Ondra"
## [7] "Lukas" "Marketa" "Jitka" "Dorota" "Jarda" "Jirka"

sample(jmena)

## [1] "Jitka" "Jarda" "Jirka" "Vita" "Dorota" "Karolina"
## [7] "Lenka" "Dominik" "Ondra" "Lukas" "Klara" "Marketa"
```

19. Nechejte si vypsat vektor výšek, v němž bude chybět 6. a 7. hodnota.

```
vyska[-(6:7)]

## [1] 159 171 NA 175 NA 178 NA 192 194 194
```

20. Zjistěte jméno/a jedince/ů, jehož/jejichž výška je neznámá.

```
jmena[is.na(vyska)]

## [1] "Dorota" "Marketa" "Ondra"
```

21. Zjistěte, kdo je nejnižší.

```
jmena[vyska == min(vyska, na.rm= T)]

## [1] "Klara" NA NA NA
```

22. Zjistěte pohlaví nejnohatějšího jedince.

```
sex[noha == max(noha, na.rm= T)]

## [1] NA NA NA NA "m"
```

23. Vypište jména všech mužů.

```
jmena[sex == "m"]

## [1] "Lukas" "Dominik" "Ondra" "Jarda" "Jirka" "Vita"
```

24. Zjistěte průměrnou délku nohy zvlášť mužů a žen.

```
# muzi
mean(noha[sex == "m"], na.rm= T)

## [1] 29.15

# zeny
mean(noha[sex == "z"], na.rm= T)

## [1] 25.125
```

25. Zjistěte, která žena je nejvyšší.

```
# tady je vic moznosti, muzeme si pomoct vytazenim zen:
vyska.z<- vyska[sex == "z"]
jmena.z<- jmena[sex == "z"]
jmena.z[vyska.z == max(vyska.z, na.rm= T)]

## [1] NA      "Lenka" NA

# nebo pouzit slozitejsi logicky vyraz:
jmena[sex == "z" & vyska == max(vyska[sex == "z"], na.rm= T)]

## [1] NA      "Lenka" NA

# nebo treba pouzit postupne 2 vybery za sebou:
jmena[sex == "z"][vyska[sex == "z"] == max(vyska[sex == "z"], na.rm= T)]

## [1] NA      "Lenka" NA
```

26. Zjistěte, zda je nejmenší mužská noha větší než největší ženská noha.

```
min(noha[sex == "m"], na.rm= T) > max(noha[sex == "z"], na.rm= T)

## [1] FALSE
```

27. Pojmenujte vektor *vyska* jmény jedinců, stejně tak vektor *noha* a *sex*. Použijte `names()`.

```
names(vyska)<- jmena
names(noha)<- jmena
names(sex)<- jmena
```

28. Zjistěte délku nohy Jardy.

```
vyska["Jarda"]

## Jarda
## 192
```

29. Zjistěte výšku Lenky, Kláry a Jitky.

```
vyska[c("Lenka", "Klara", "Jitka")]

## Lenka Klara Jitka
## 175 159 171
```

30. Zjistěte, zda noha Dominika je větší než noha Karolíny.

```
noha["Dominik"] > noha["Karolina"]
```

```
## Dominik
## TRUE
```

31. Doplňte neznámou hodnotu výšky a velikosti nohy Markěty. Výška 176 cm a noha 25.6 cm.

```
vyska["Marketa"] <- 176
```

```
vyska
```

```
## Klara Jitka Dorota Lenka Marketa Karolina Lukas Dominik
## 159 171 NA 175 176 168 181 178
## Ondra Jarda Jirka Vita
## NA 192 194 194
```

```
noha["Marketa"] <- 25.6
```

```
noha
```

```
## Klara Jitka Dorota Lenka Marketa Karolina Lukas Dominik
## 23.0 25.0 NA 27.0 25.6 25.5 NA 27.0
## Ondra Jarda Jirka Vita
## NA 29.6 29.0 31.0
```

32. Odstraňte z vektorů *jmena*, *vyska*, *noha* a *sex* záznamy nezměřených jedinců (těch, jejichž *vyska* je NA).

```
# Za normalnich okolnosti je dobre si pro informaci ponechat kopie puvodnich vektoru.
# Logicky vektor oznacujici NA vysky:
```

```
na.vyska <- is.na(vyska)
```

```
na.vyska
```

```
## Klara Jitka Dorota Lenka Marketa Karolina Lukas Dominik
## FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## Ondra Jarda Jirka Vita
## TRUE FALSE FALSE FALSE
```

```
# Vyber jen hodnot, kde vysky nebyly NA:
```

```
jmena <- jmena[!na.vyska]
```

```
jmena
```

```
## [1] "Klara" "Jitka" "Lenka" "Marketa" "Karolina" "Lukas"
## [7] "Dominik" "Jarda" "Jirka" "Vita"
```

```
vyska <- vyska[!na.vyska]
```

```
vyska
```

```
## Klara Jitka Lenka Marketa Karolina Lukas Dominik Jarda
## 159 171 175 176 168 181 178 192
## Jirka Vita
## 194 194
```

```

noha<- noha[!na.vyska]
noha

## Klara Jitka Lenka Marketa Karolina Lukas Dominik Jarda
## 23.0 25.0 27.0 25.6 25.5 NA 27.0 29.6
## Jirka Vita
## 29.0 31.0

sex<- sex[!na.vyska]
sex

## Klara Jitka Lenka Marketa Karolina Lukas Dominik Jarda
## "z" "z" "z" "z" "z" "m" "m" "m"
## Jirka Vita
## "m" "m"

```

33. Vraťte se k bodům 5. a 9. a vytvořte vektor libovolných předdefinovaných barev *barva2*, které budou odlišovat mezi muži a ženami našich jedinců. Barvy vyberte ze seznamu `colours()` a pro vytvoření vektoru použijte vektor *barva* jako subscript.

```

# Protoze jsme odstranili nektere NA hodnoty, potrebujeme vektor barva udelat znovu
# (nebo z nej odstranit hodnoty stejneho poradí)
barva<- (sex == "z") + 1
barva

## Klara Jitka Lenka Marketa Karolina Lukas Dominik Jarda
## 2 2 2 2 2 1 1 1
## Jirka Vita
## 1 1

barva2<- c("slateblue1", "honeydew4")[barva]
barva2

## [1] "honeydew4" "honeydew4" "honeydew4" "honeydew4" "honeydew4"
## [6] "slateblue1" "slateblue1" "slateblue1" "slateblue1" "slateblue1"

```

34. Obdobně vytvořte vektor *puntik*, který bude rovněž rozlišovat mezi muži a ženami a bude obsahovat kódové označení dvou vybraných symbolů (dvě číselné hodnoty mezi 1 a 25).

```

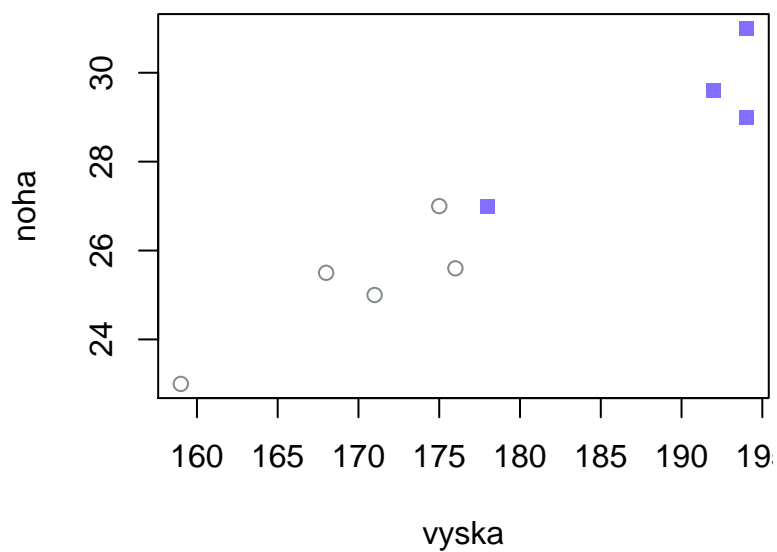
puntik<- c(15,1)[barva]
puntik

## [1] 1 1 1 1 1 15 15 15 15 15

```

35. Následně překreslete obrázek bodu 11., pro definici barvy použijte vektor *barva2* a pro definici symbolu *puntik*.


```
par(mar=c(4,4,1,0.1))  
plot(noha ~ vyska, col= barva2, pch= puntik)
```



Logické operátory:

>	větší
<	menší
==	roven
<=	menší nebo rovno
>=	větší nebo rovno
!	opak
!=	není rovno
&	a
	nebo