

Az R adatelemzési nyelv alapjai II.

Gyakorlat

Függvények

1. Írjon egy olyan függvényt, amely kiszámítja az argumentumban kapott szám faktoriálisát! Figyeljen az argumentum ellenőrzésére is.

Gyakorlandó ismeretek:

- függvény létrehozása
- vezérlési szerkezet: if (elágazás), for ciklus
- adattípus ellenőrzése (pl. `is.numeric`)

2. Írjon egy olyan függvényt, amely egy listában a hónapok neveivel indexelve visszaadja az argumentumban kapott évszámnak megfelelően az adott hónapba tartozó napok számát! Figyeljen a szökőévekre (4-el osztható, de 100-al nem osztható, kivéve a 400-al osztható évek).

Példa kimenet:

```
> m(2000)
$January
[1] 31

$February
[1] 29

$March
[1] 31
...
```

Gyakorlandó ismeretek:

- függvény létrehozása
- vezérlési szerkezet: if (elágazás)
- listák kezelése

3. Írjon egy olyan függvényt, amely a bemenetéről kapott adattábla (`data.frame`) alapján visszaad egy olyan új adattáblát, amely csak azokat a sorokat tartalmazza, amelyek nem fordul elő NA érték, illetve az argumentumban kapott oszlop szerint van sorrendezve.

Függvény deklaráció:

```
complete.ordered.dataframe <- function( data = data, sort.by = sort.by )
```

Gyakorlandó ismeretek:

- függvény létrehozása, nevesített argumentumok használata
- adattáblák kezelése
- hiányzó adatok kezelése

Grafikonok létrehozása

A feladat megoldása során egy olyan adattáblával fogunk dolgozni, amely hipotetikus személyek testmagasság és testsúly adatait tartalmazza.

A táblázat elállításának R-kódja:

```
df <- rbind( data.frame( Magassag = rnorm( n = 50, mean = 175, sd = 10 ),
                        Suly = rnorm( n = 50, mean = 80, sd = 10 ),
                        Nem = "Ferfi" ),
            data.frame( Magassag = rnorm( n = 50, mean = 165, sd = 8 ),
                        Suly = rnorm( n = 50, mean = 65, sd = 10 ),
                        Nem = "No" ) )

df$Nem <- as.factor( df$Nem )

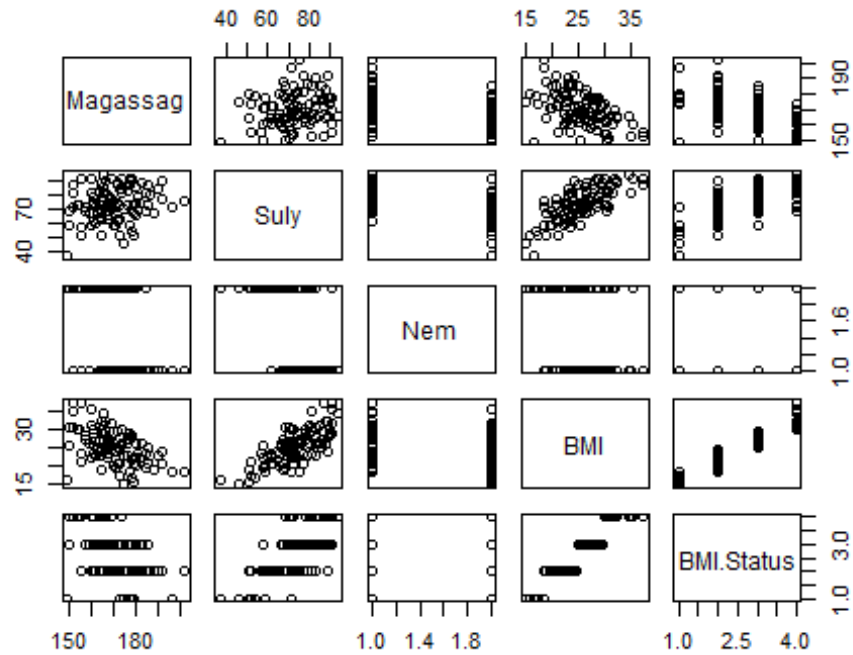
df$BMI <- df$Suly / ( ( df$Magassag / 100 ) * ( df$Magassag / 100 ) )

df$BMI.Status <- cut( x = df$BMI,
                     breaks = c( 0, 18.5, 25, 30, Inf ),
                     labels = c( "Alultaplalt",
                                "Normalis",
                                "Tulsulyos",
                                "Elhizott" ) )
```

1. Ábrázolja az adattábla összes változóját az összes változó függvényében.

Parancsok: `plot`

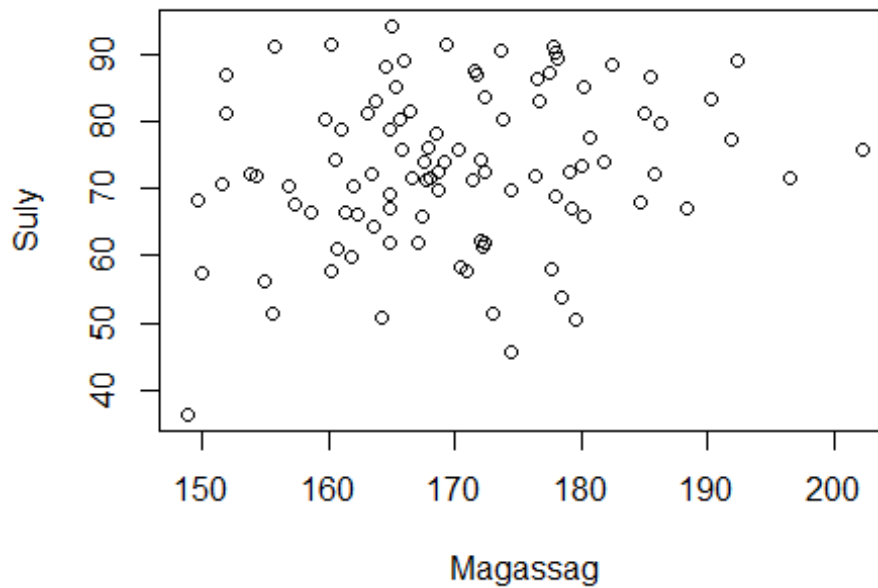
Várt eredmény:



2. Ábrázolja a testsúlyt a testmagasság függvényében.

Parancsok: `plot`

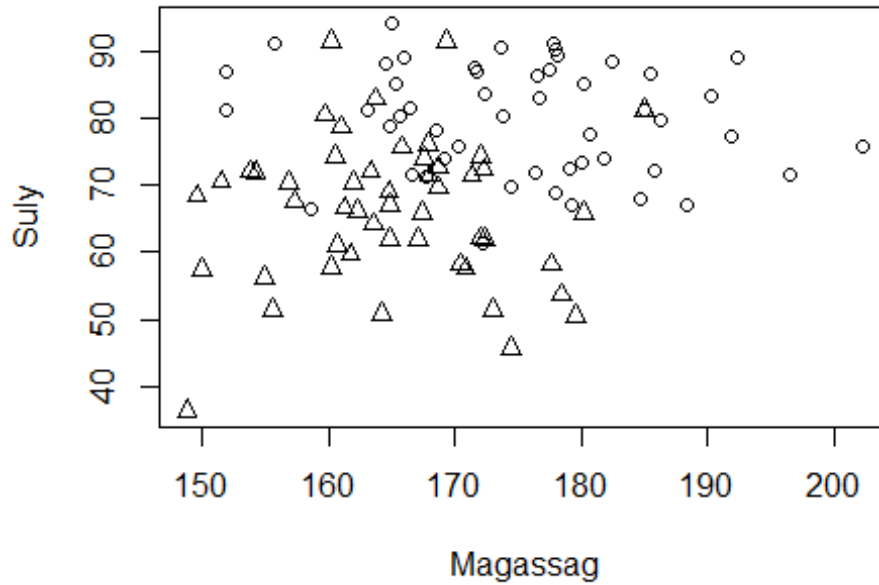
Várt eredmény:



3. Ábrázolja a testsúlyt a testmagasság függvényében úgy, hogy nemenként különböző pont típussal ábrázolja az értékeket.

Parancsok: `plot` (`pch` argumentum)

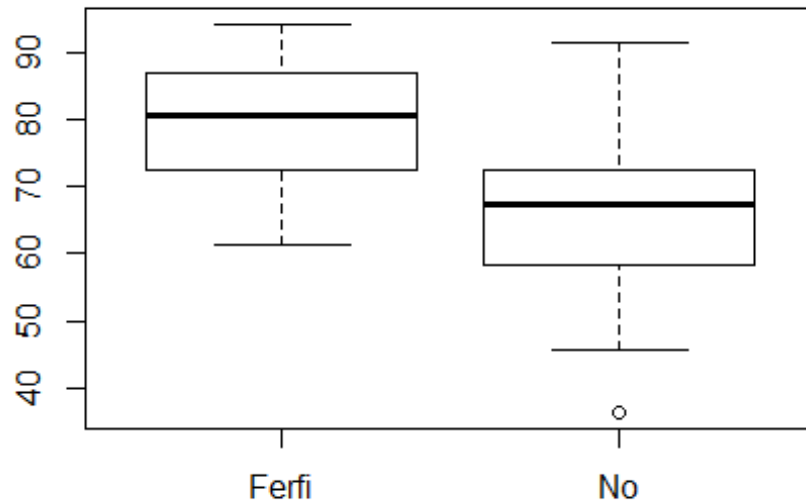
Várt eredmény:



4. Ábrázolja dobozdiagramon a testsúly eloszlását nemenként.

Parancsok: `boxplot` (formula használatával)

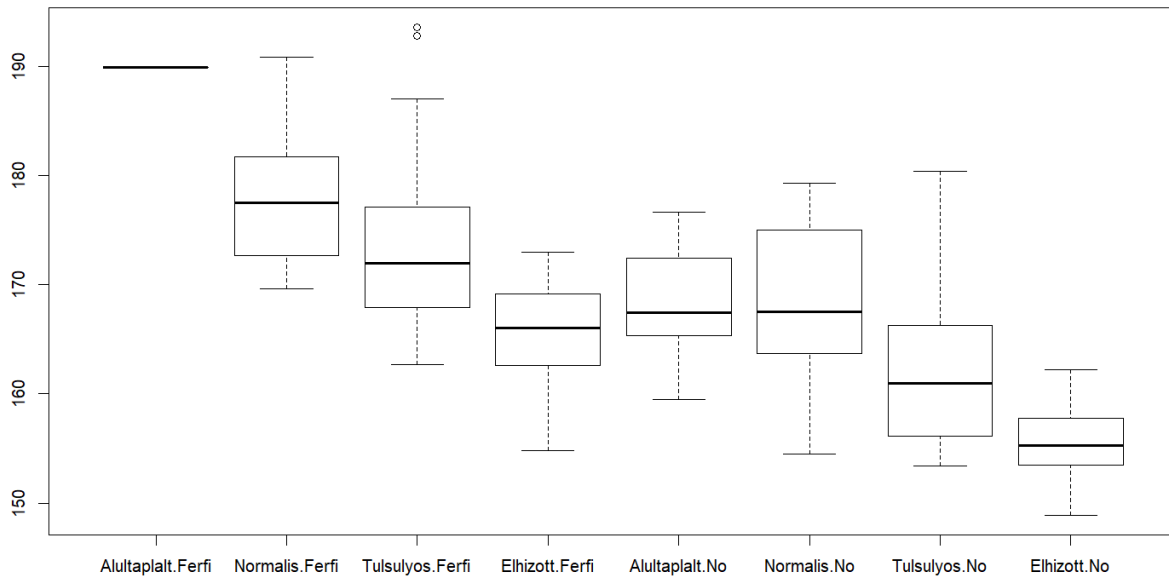
Várt eredmény:



5. Ábrázolja dobozdiagramon a magasság eloszlását nemenként és BMI státuszonkénti bontásban.

Parancsok: `boxplot` (formula használatával)

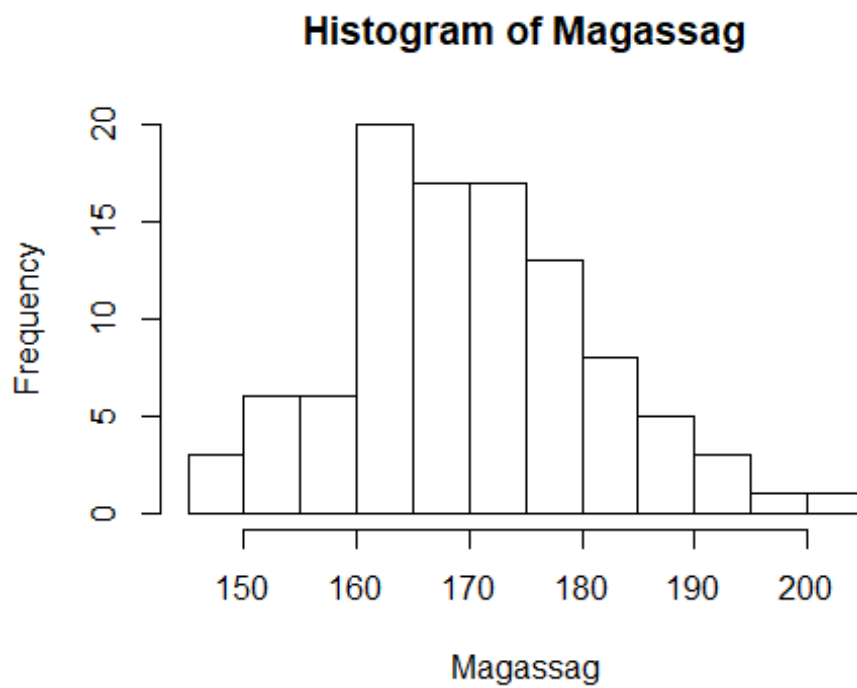
Várt eredmény:



6. Ábrázolja a magasság eloszlását hisztogram segítségével.

Parancsok: `hist`

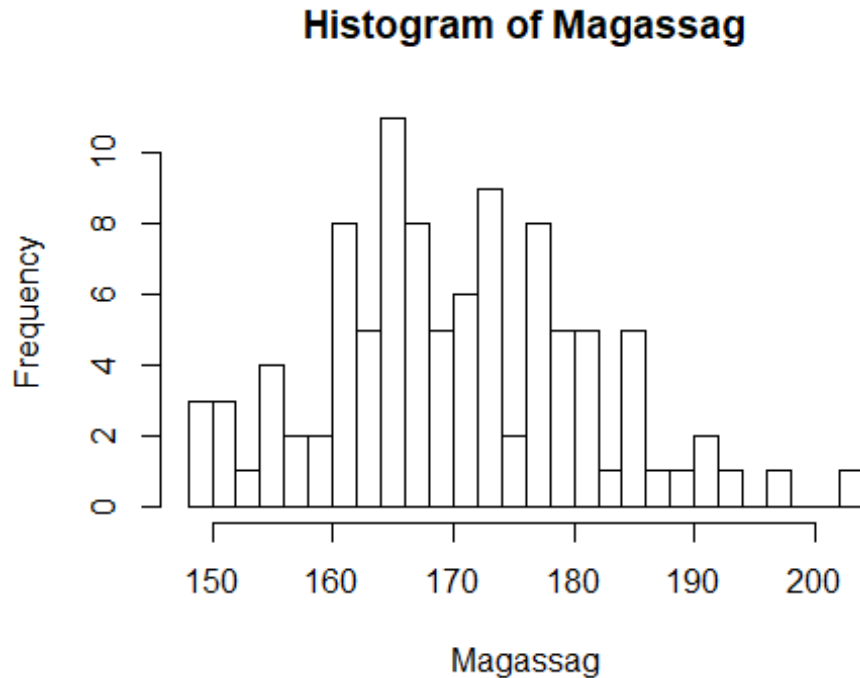
Várt eredmény:



7. Ábrázolja a magasság eloszlását hisztogram segítségével úgy, hogy az oszlopok száma 20 legyen.

Parancsok: `hist`

Várt eredmény:



8. Ábrázolja a magasság és a testsúly eloszlását nemenként hisztogram segítségével úgy, hogy a 4 ábrát rendezze el egy 2*2-es táblázatban. Figyeljen arra, hogy változónként azonos megjelenítési határokat alkalmazzon az x tengelyen.

Parancsok: `hist`, `par (mfrow argumentum)`

Várt eredmény:

