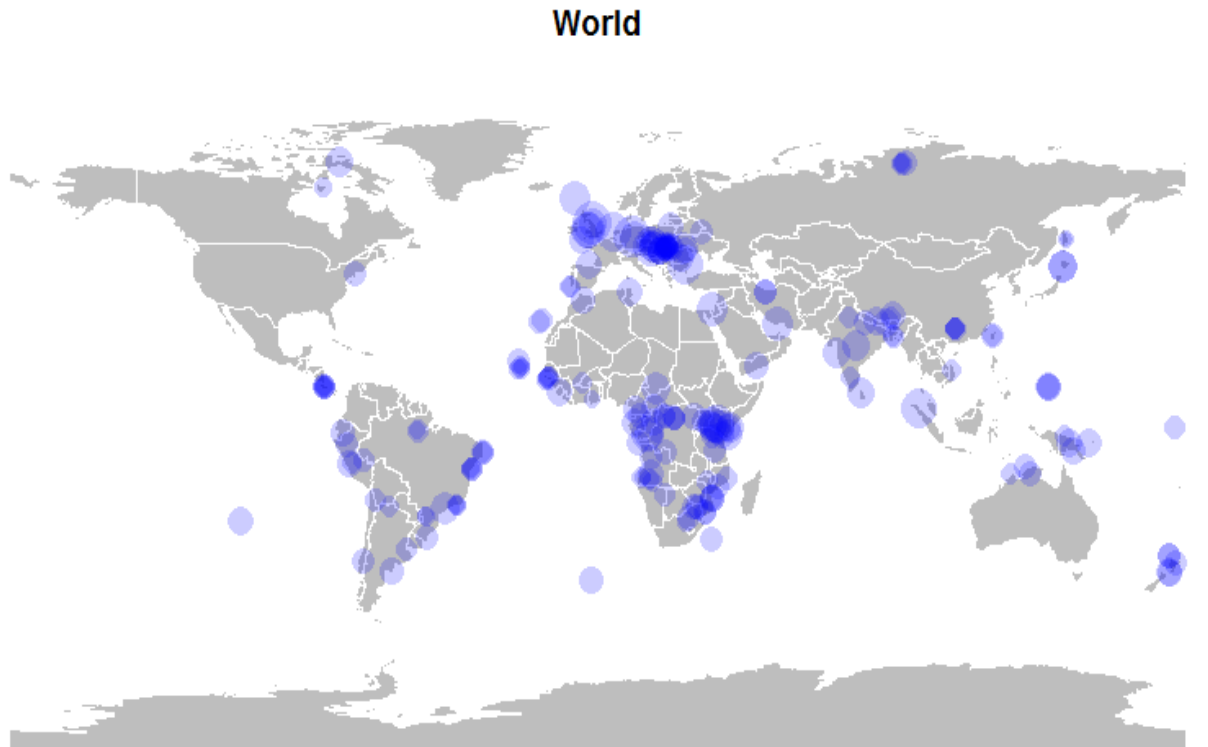


Načtení dat - souřadnic

- `"11_mapy_data.txt"`
- Tři sloupce:
 - `lat` - latitude
 - `long` – longitude
 - Souřadnice zkoumaných populací – celý svět
 - `log` – specifická proměnná určující velikost bodů
 - v tomto případě se jedná o logaritmus z velikosti zkoumaných populací na daných souřadnicích
- Data po načtení zkontrolujeme a nezapomeneme na funkci:
 - `attach(data)`

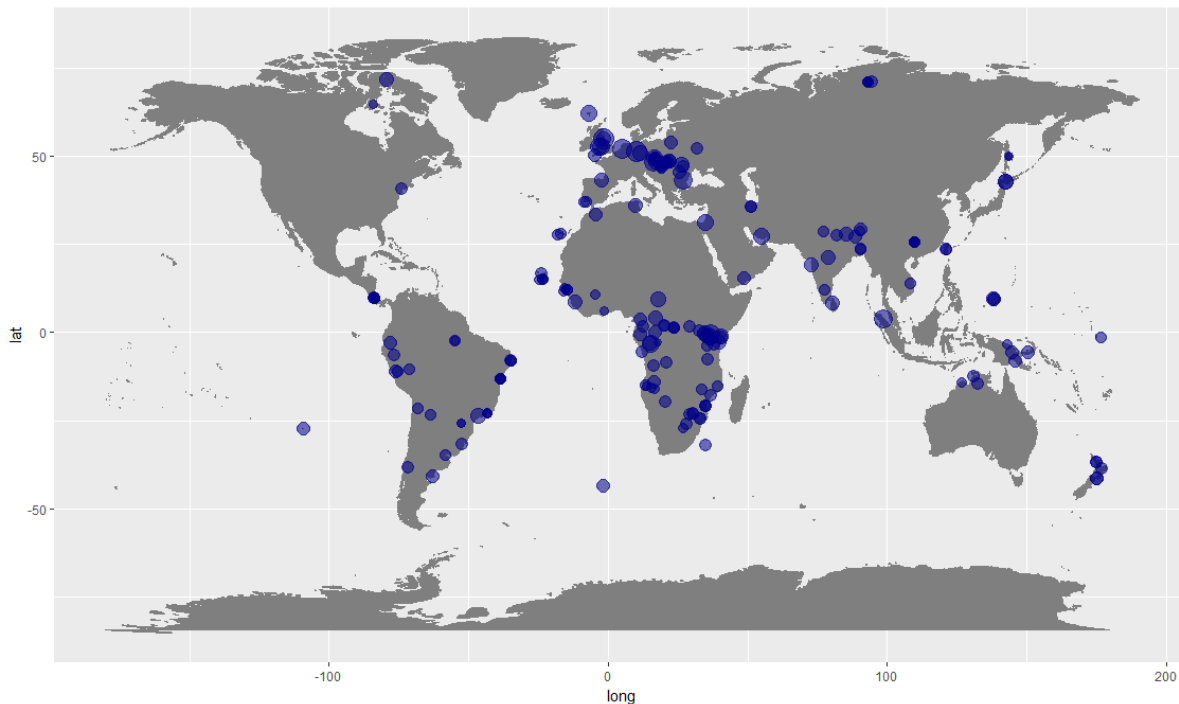
Zobrazení v `rworldmap`

- `newmap <- getMap(resolution = "low")`
- `plot(newmap, col="grey", border="white", main = "World")`
- `points(data$long, data$lat, col = alpha(4,0.2), cex = data$log, pch = 16)`



Zobrazení v ggmap

- `mp <- NULL`
- `mapWorld <- borders("world", colour="gray50", fill="gray50")`
- `mp <- ggplot() + mapWorld`
- `windows()`
- `mp + geom_point(data = data, aes(x = long, y = lat, alpha = 0.8), size = c(data$log*2), shape = 19, col = "darkblue") + guides(fill=FALSE, alpha=FALSE, size=FALSE)`

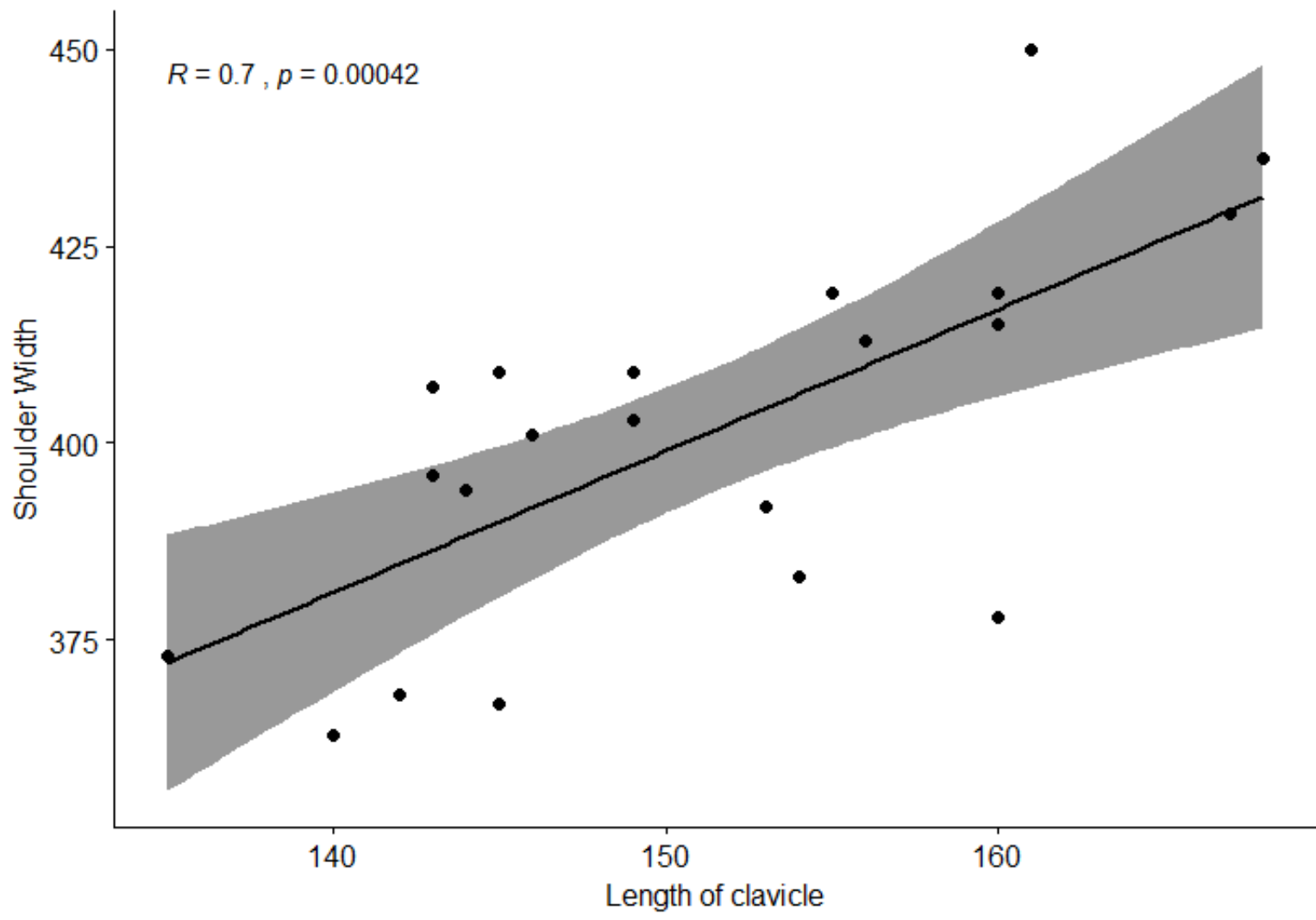


Korelační analýza v R

- vztah/ závislost mezi dvěma nebo více proměnnými
- metody korelační analýzy:
 - **Pearsonova korelace** (lineární vztah mezi x a y) – parametrická (=> záleží na rozložení dat a může být použita pouze pokud mají data normální rozložení) $y = f(x)$ resp. lineární regrese
 - **Kendall tau** a **Spearman rho** – neparametrické; ordinální data

R funkce:

- `cor()` výpočet korelačního koeficientu
- `cor.test()` test vztahu dvou párových vzorků, který vrací jak korelační koeficient tak p hodnotu korelace



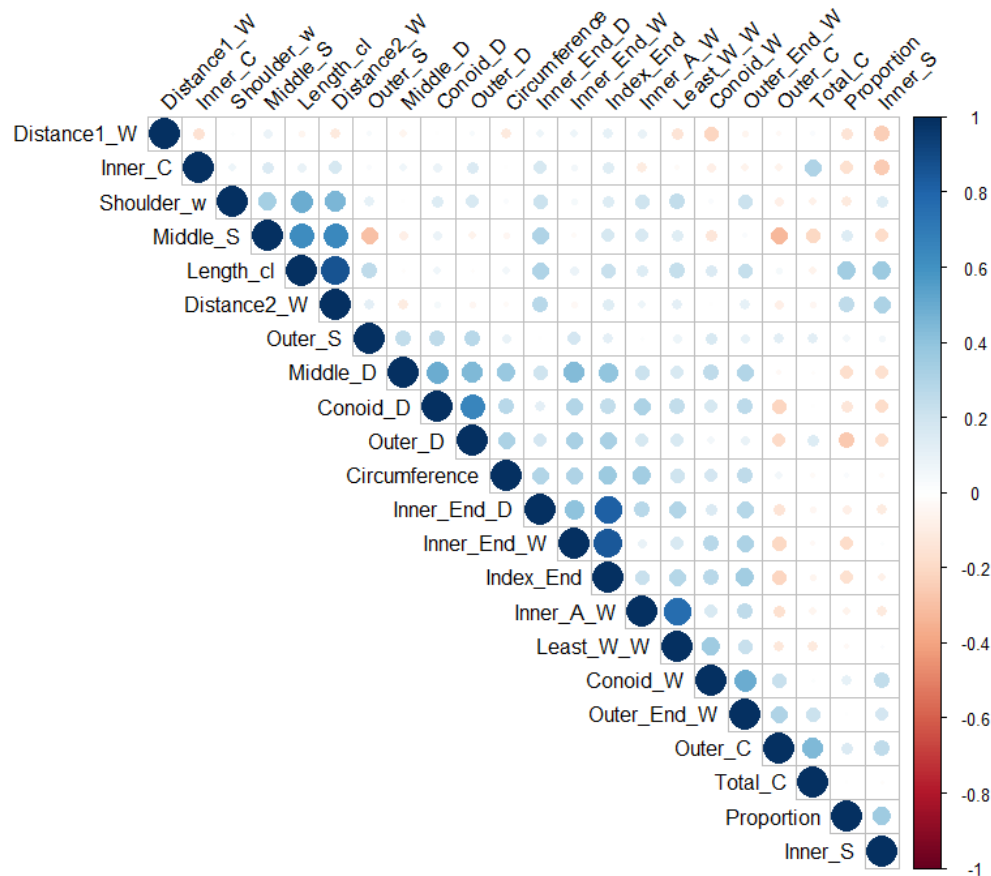
```
cor(x, y, method = c("pearson", "kendall", "spearman"))
cor.test(x, y, method=c("pearson", "kendall", "spearman"))
```

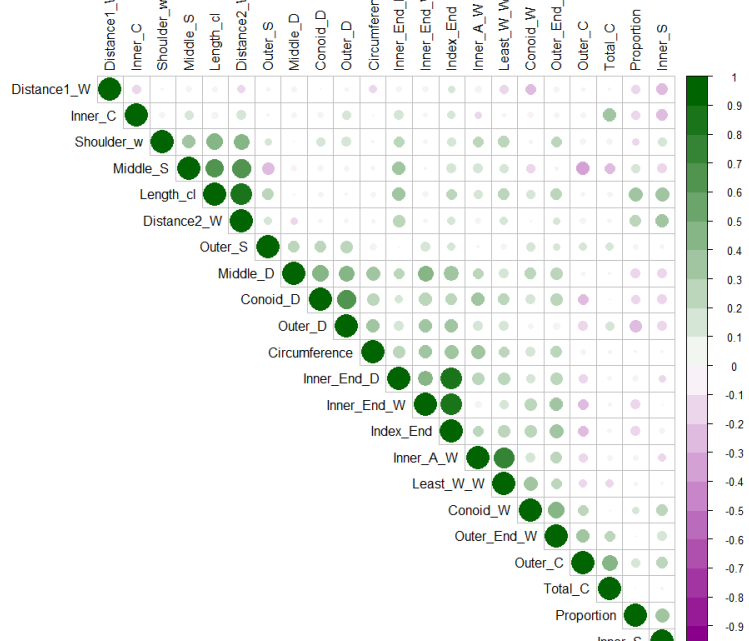
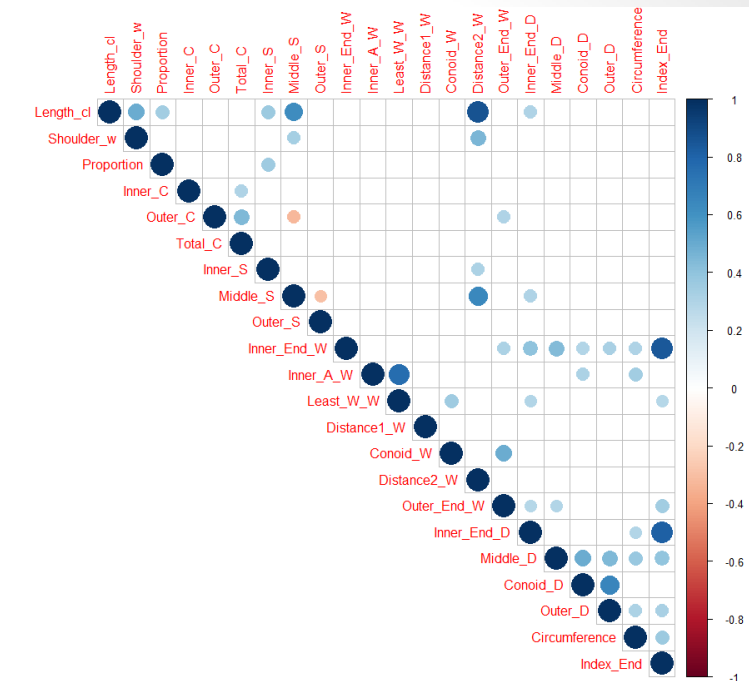
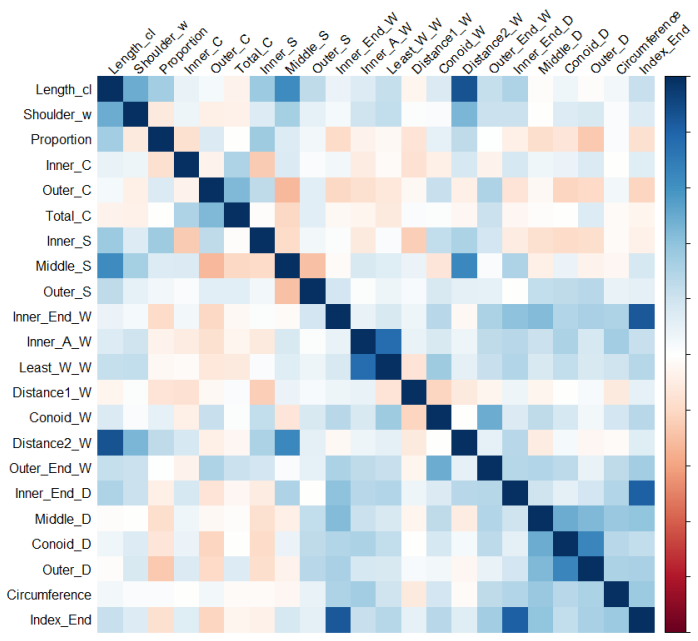
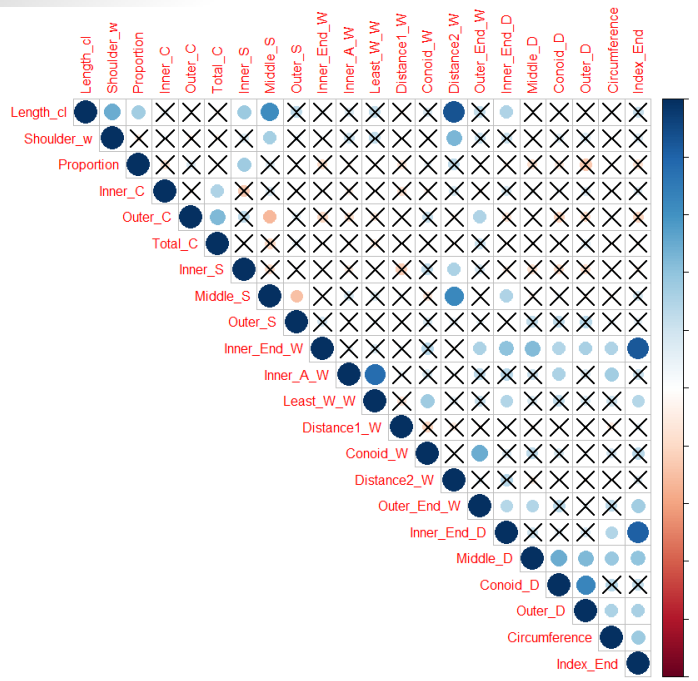
- pozor na NA hodnoty!
- Pearson's product-moment correlation data: SINDX[Sex.x == "male", "Length_cl.x"] and SINDX[Sex.x == "male", "Shoulder_w.x"]
t = 3.4634, df = 19, p-value = **0.002602**
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.2603227 0.8306999
sample estimates:
cor
0.6220993

res1\$p.value
res\$estimate

KORELAČNÍ MATICE

- Výpočet: balíček `Hmisc`, který umožňuje jak korelační matici tak matici příslušných p -hodnot
- Přehledné zobrazení pomocí funkce `corrplot()`
 - Balíček: `corrplot`





DOMÁCÍ ÚKOL 11

- A) Načtěte data "11_data.csv" a vytvořte correlogram z proměnných: Length_cl, Shoulder_w, Proportion, Inner_C, Outer_C, Total_C (7 až 12) – pouze pro ženy a pro pravou stranu.
 - Nezapomeňte vytvořit matici p -hodnot.
 - Nezobrazujte diagonálu.
 - Barvy correlogramu volte libovolně, stejně tak typ zobrazení.

Výsledek může vypadat např. takto:

