

Wprowadzenie do R-clustering

Aleksandra Stachniak, Agata Słonka, grupa 1

Zadanie 1

```
library(magrittr)
library("ggplot2")
library("GGally")

#zad1

list10 <- 1:10
print(list10)

list10 %<>% log2() %<>% sin() %<>% sum() %<>% sqrt()
print(list10)

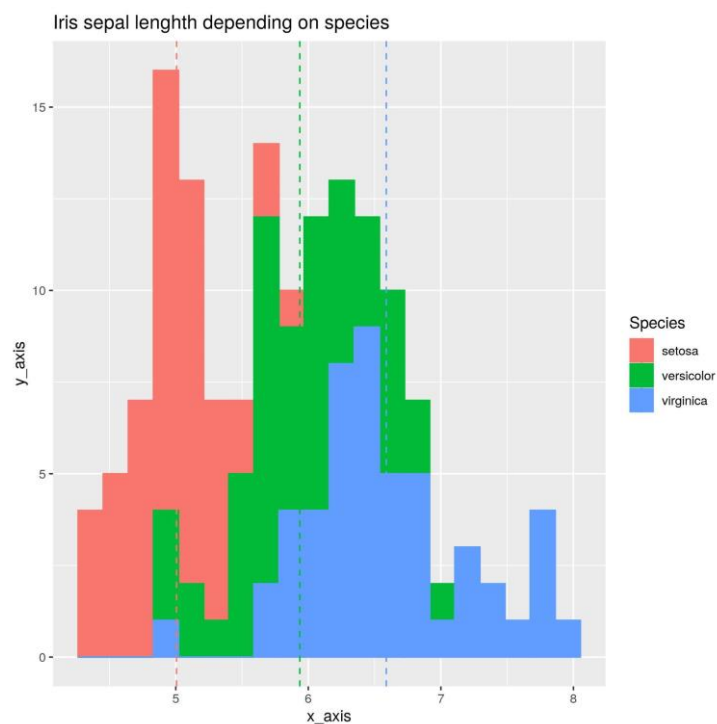
data(iris)
print(head(iris))
agg <- iris %>% aggregate (. ~ Species, ., mean)
print(agg)
```

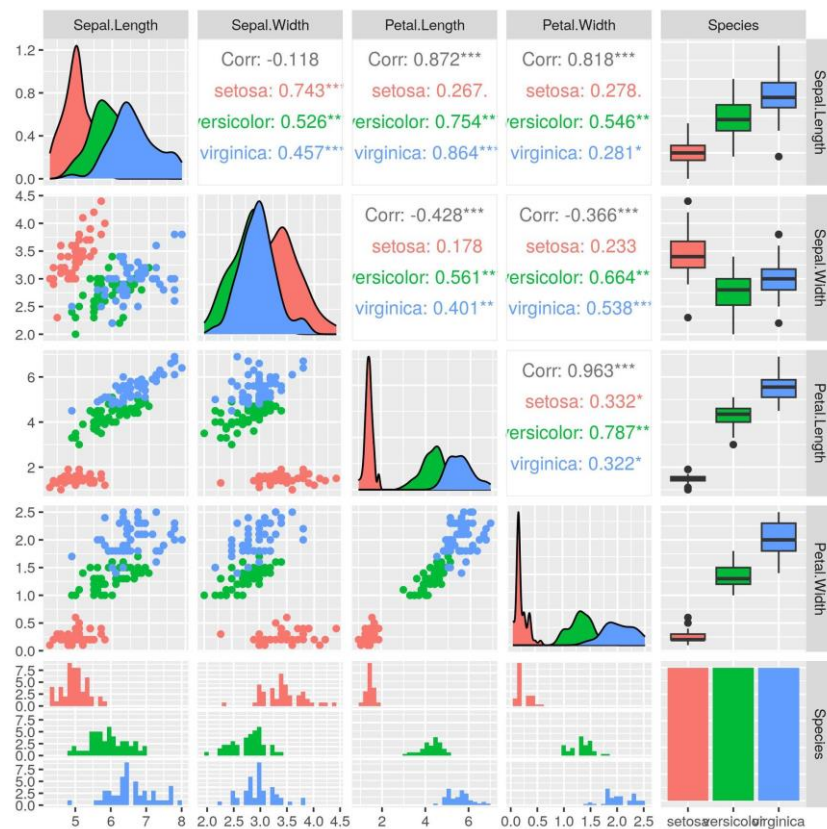
Zadanie 2

```
a <- ggplot(iris, aes(x=Sepal.Length)) + geom_histogram(aes(fill=Species, color=Species), bins=20) +
  geom_vline(data=agg, aes(xintercept=Sepal.Length, color = Species), linetype = "dashed") +
  labs(x='x_axis', y='y_axis', title='Iris sepal length depending on species')

ggsave("/home/podpunkt34.jpg", plot=a)

fig2 <- ggpairs(data = iris, aes(color=Species))
ggsave("/home/podpunkt6.jpg", plot=fig2)
```





Zadanie 3

```
x<- iris[,1:4]
y<- iris[,5]

sum_sqr <- c()

for (i in 1:10){
  kmeans_result <- kmeans(x, i)
  sum_sqr <- append(sum_sqr, kmeans_result$tot.withinss)
}

fig3 <- ggplot(data.frame(iteration = 1:length(sum_sqr), value = sum_sqr), aes(x = iteration, y=sum_sqr))+geom_line()
ggsave("/home/zad3.jpg", plot=fig3)

kmeans_result <- kmeans(x,3)
fig4 <- ggplot(iris, aes(x = Sepal.Width, y = Petal.Width, color = kmeans_result$cluster)) + geom_point()
ggsave("/home/zad3_width.jpg", plot=fig4)

fig5 <- ggplot(iris, aes(x = Sepal.Width, y = Petal.Width, color = Species)) + geom_point()
ggsave("/home/zad3_width_sp.jpg", plot=fig5)
```

