

Rodion Kovalenko

Matriculation number 3009393

(Studiengang: Wirtschaftsinformatik)

Exercise 4.

4. a) Least variance of petal length among all species was by “setosa”, which can be seen at the following figure 1. For execution of this task I calculated the variances of petal length in all species and plotted it in graph with y-axis for variance and the name of species for the x-axis.

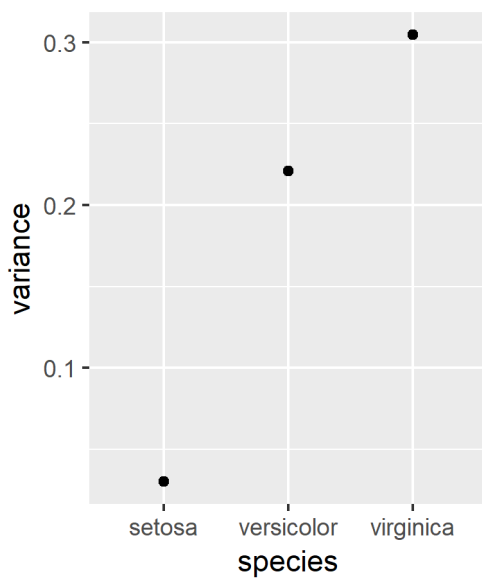


Figure 1. Variance in terms of sepal length in all species

4. b) Comparing the attributes sepal length and sepal width we come the conclusion that there are outliers, what can be clearly seen in the figure 2. There are some points stay separately from the others.

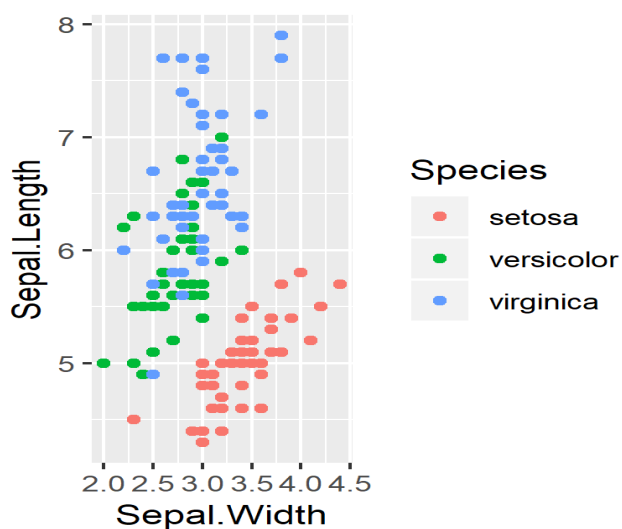


Figure 2. Outliers in species according to sepal length and sepal width

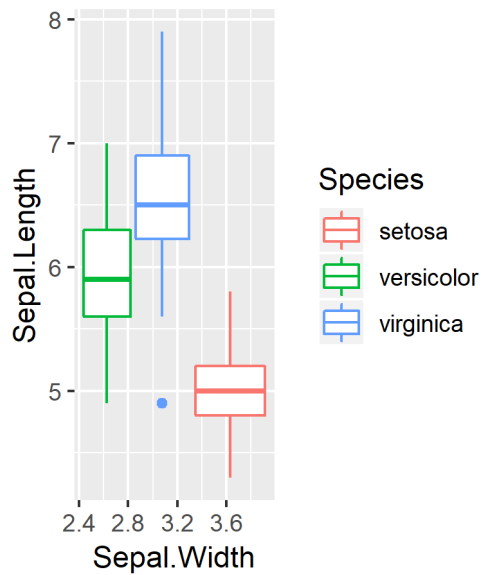


Figure 3. Outliers with plot box

Figure 3 shows plot boxes and again outliers. We can see at the figure, that “virginica” has outlier.

4. c) Pairwise correlation can be done with standard library of R. In this case it will plot all the attribute of observation in matrix form, enabling graphical comparison. Figure 4 shows correlation of pairwise attributes sepal length, sepal width, petal length and petal width.

Figure 4. Pairwise correlation of species attributes

