Exploratory Multivariate Data Analysis

Homework 2

Due 9/23/2024

1. The qnorm function returns the percentiles (quantiles) of a normal distribution. Use the qnorm function to find the 95th percentile of the standard normal distribution. Then, use the qnorm function to find the quartiles of the standard normal distribution (the quartiles are the 25, 50, and 75 percentiles). Hint: Use c(.25, .5, .75) as the first argument to qnorm.
2. (iris data). The iris data gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of three species of iris. There are four numeric variables corresponding to the sepal and petal measurements and one factor, Species. Display a table of means by Species (means should be computed separately for each of the three Species).
3. (mtcars data). Display the mtcars data included with R and read the documentation using ?mtcars. Display parallel boxplots of the quantitative variables. Display a pairs plot of the quantitative variables. Does the pairs plot reveal any possible relations between the variables?
4. Use the bivariate boxplot on the scatterplot of each pair of variables in the air pollution data to identify any outliers. Calculate the correlation between each pair of variables using all the data and the data with any identified outliers removed. Comment on the results. (USpollution data in MVA package)