Problem Set 3

Mathematics and Statistics for Economists

Due date: 11:59pm Wednesday 18th September 2024

For this problem set, please send me your **R-script** file in addition to your PDF file.

If there is a command we didn't cover in class, you could read the R tutorial (on Moodle), use "help" command, or use our powerful ChatGPT to quickly get yourself familiar with it.

Question 1 The "mtcars" dataset (prestored in R) contains various features about cars, such as miles per gallon (mpg), number of cylinders (cyl), horsepower (hp), and other characteristics. This dataset will be used for all the exercises below.

1.1 Load the "mtcars" dataset.

1.2 Find out how many observations and how many variables are in the dataset.

1.3 Use the "select()" function from "dplyr" package to select the columns mpg, cyl, and hp from the mtcars dataset.

1.4 Keep only the rows where the number of cylinders (cyl) is 6 and horsepower (hp) is greater than 100 from the mtcars dataset.

1.5 Create a new column hp_per_cyl, which is the ratio of horsepower (hp) to the number of cylinders (cyl).

1.6 Use "group_by()" and "summarize()" to find the average miles per gallon (mpg) for each number of cylinders (cyl).

Question 2 Keep using the same R-script and add the commands to reproduce the following exercises. To creat graphs, we have to install and include package "ggplot2".

2.1 Create a scatter plot of mpg versus hp, colored by the number of cylinders (cyl).

2.2 Create a boxplot showing the distribution of mpg for each number of cylinders (cyl).

2.3 Create a histogram of the mpg variable, with bin sizes of 2.

Question 3 Keep using the same R-script and add the commands to reproduce the following exercises.

3.1 Calculate the mean of mpg and hp in the mtcars dataset.

3.2 Use the var() function to calculate the variance of mpg and hp.

3.3 Use the cov() function to calculate the covariance between mpg and hp.

3.4 Use the cor() function to calculate the correlation between mpg and hp.

```
car_names <- data.frame(
  car_model = rownames(mtcars),
  origin = c(rep('USA', 10), rep('Europe', 10), rep('Japan', 2))
)
```

Question 4 Keep using the same R-script and add the commands to reproduce the following exercises. You are given another dataset, car_names, which contains the names of the car models:

4.1 Merge the mtcars dataset with car_names by the car model names.

4.2 Reshape the mtcars dataset into long format.

4.3 Reshape the long format data back to short format.

Question 5 Keep using the same R-script and add the commands to reproduce the following exercises.

4.1 Use the "lm()" function to fit an OLS regression model predicting mpg based on hp and wt.

4.2 Use the "summary()" function to display the summary of the regression model.

4.3 Use the "predict()" function to predict mpg for a car with 150 horsepower (hp) and a weight (wt) of 3.0.