Simulation Lab(MC503)

Assignment-3

Try to solve all the problems

---

1. The information below shows the population (in millions) of the top ten most populous cities in the US in 2019. (Source: `moving.com`) New York - 8.60, Los Angeles - 4.06, Chicago - 2.68, Houston - 2.40, Phoenix - 1.71, Philadelphia - 1.58, San Antonio - 1.57, San Diego - 1.45, Dallas - 1.40, San Jose - 1.03, Using the given information,

   i. Letue create a data frame.

   ii. Draw the pie chart in ascending or descending order, rearrange the dataset and rename the object first. Then draw the pie chart of the new object.

   iii. To add percentages to the label, calculate the percentage of each entry as compared to the total. Then round the result to the nearest whole number in order to avoid decimals.

   iv. Using *ggplot2* create a pie charts.

2. Check to see if you have the mtcars dataset

   i. How many observations (rows) and variables (columns) are in the mtcars dataset?

   ii. Convert the column names of cars to all upper case.

   iii. Subset the columns from cars that end in "p" and call it pvars.

   iv. Create a subset of the data that only contains the columns: wt, qsec, and hp and assign this object to carsSub - what are the dimensions of this dataset?

   v. Subset the rows of cars that get more than 20 miles per gallon (mpg) of fuel efficiency - how many are there?

   vi. Subset the rows that get less than 16 miles per gallon (mpg) of fuel efficiency and have more than 100 horsepower (hp) - how many are there?

   vii. Create a subset from the cars data that only contains the columns: wt, qsec, and hp for only the cars with 8 cylinders and reassign this object to carsSub - what are the dimensions of this dataset?

   viii. Re-order the rows of carsSub by weight in increasing order.

   ix. Create a new variable in carsSub called wt2, which is equal to wt^2.

3. Bike Lanes Dataset: BikeBaltimore is the Department of Transportation's bike program. The data is from `http://data.baltimorecity.gov` You can download it as a CSV in your current working directory. Note its also available at: `http://johnmuschelli.com/intro_to_r/data/Bike_Lanes.csv`

i. How many bike "lanes" are currently in Baltimore? You can assume each observation/row is a different bike "lane". (hint: how to get the number of rows of a data set)

ii. How many (a) feet and (b) miles of bike "lanes" are currently in Baltimore? (5280 feet in a mile).

iii. How many types of bike lanes are there? (Hints: unique, table, or bike % > % count). Which type has (a) the most number of and (b) the longest average bike lane length? (Hint: group_by and summarize)

iv. Which project category has the longest average bike length?

. . . . . end . . . . . .