**Project guide**

1. **Data exploration**
2. Choose and load the R dataset corresponding to your group subject’s and Identify which variables your data set are numeric, and which are categorical (factors) if applicable

You can use the command data() to list all the datasets,

You can create a table by using the data(name of your data set). to get more information on your data set tape its name in the help window.

You can use the command mode()

1. Generate summary level descriptive statistics: Show the mean, median, 25th and 75th quartiles, min, and max for each of the applicable variables in your data set
2. Determine the frequency for each of one of the categorical variables.
3. Determine the frequency for each of the one of the categorical variables, by a different categorical variable.

You can use the command levels to discover the different modalities of the categorical variables. Use also Summary() or table() for question 4.

1. **Graphic data representation**
   1. **Univariate study**
2. Use the commands pie() barplot and dotchart() to represent the categorical data. Comment
3. Create a graph for each single numeric variable. (histogram)
4. Create a graph for each single numeric corresponding to different categorical variable (histogram)
   1. **Bivariate study**
5. Use the command plot or sunflowerplot to plot the scatterplot of the dependent and independent variables. What is the difference between these two commands? Comment your results.
   1. **Graphic representation for the different data categories**
6. Represent the scatter plot for the dependent and independent variables for each data category. Comment
7. **Regression Analysis**
   1. **Testing hypothesis**
8. Use the Khi-deux test to verify the independent of each data category.
9. Testing the Standard Assumptions of linear regression
   1. **Build the model the regression model**
   2. **Verify model significance (Model validation)**

**Datasets**

|  |  |  |
| --- | --- | --- |
| **dataset** | **Dataset number** | **group** |
| **airquality** | **1** |  |
| **ChickWeight** | **2** |  |
| **quakes** | **3** |  |
| **iris** | **4** |  |
| **longley** | **5** |  |
| **esoph** | **6** |  |
| **occupationalStatus** | **7** |  |
| **mtcars** | **8** |  |