

Assignment 2

You will need to R and RStudio for this assignment.

You can download both, R and RStudio at <https://rstudio-education.github.io/hopr/starting.html>

Task

Choose your own dataset from Kaggle (<https://www.kaggle.com/datasets>). With that dataset, we will perform some EDA (Exploratory Data Analysis).

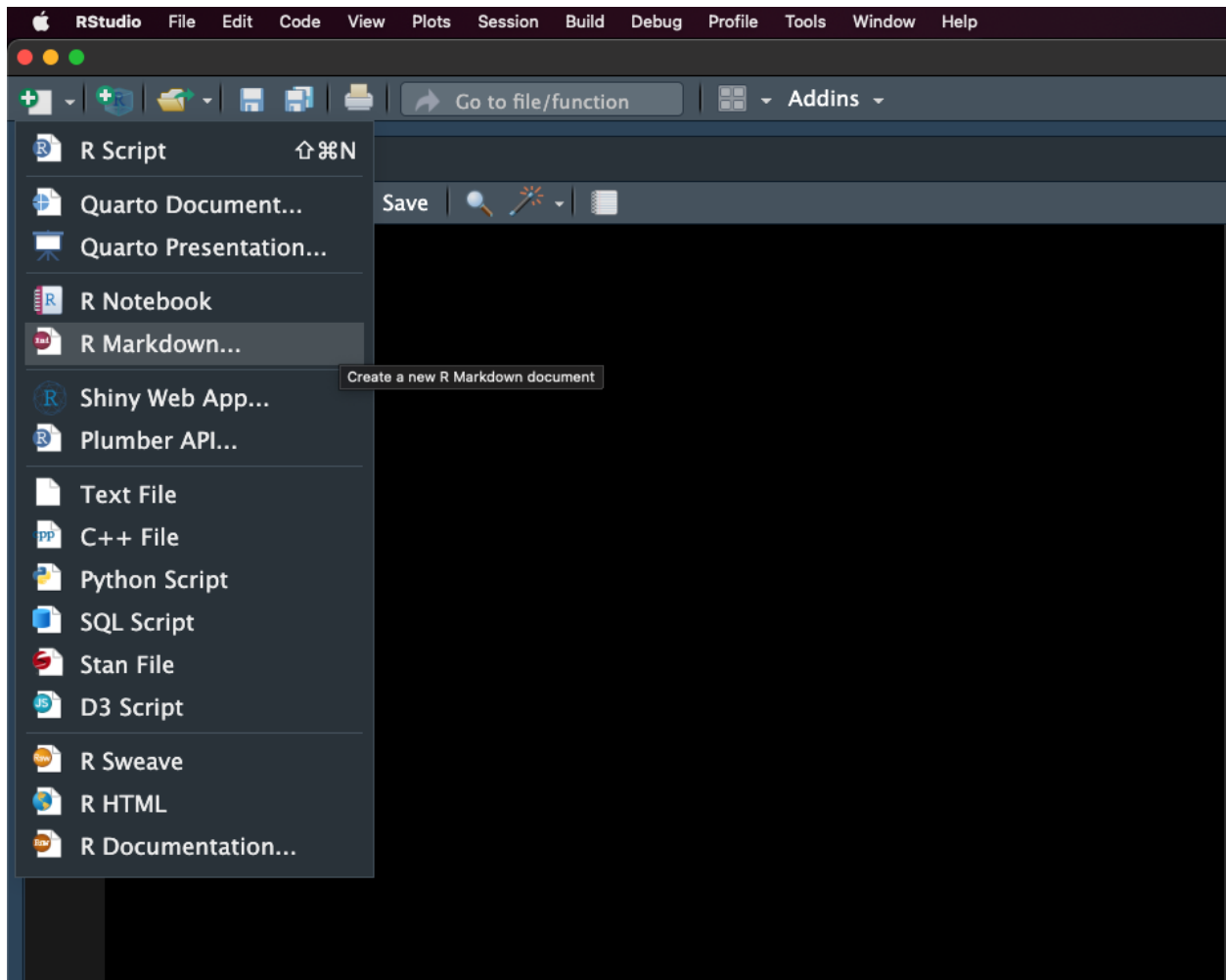
This assignment is fairly straight forward. You are required to perform three (3) data visualization techniques on your data. After each data visualization technique, please elaborate on your findings and explain what conclusion did you come to from the chart/plots/etc..

Examples of data visualization techniques:

- Plots (box plots, scatter plots, etc.)
- Histograms
- Heat maps
- Charts
- Tree Maps
- Graphs (bar graphs, line graphs, pie graph, etc.)

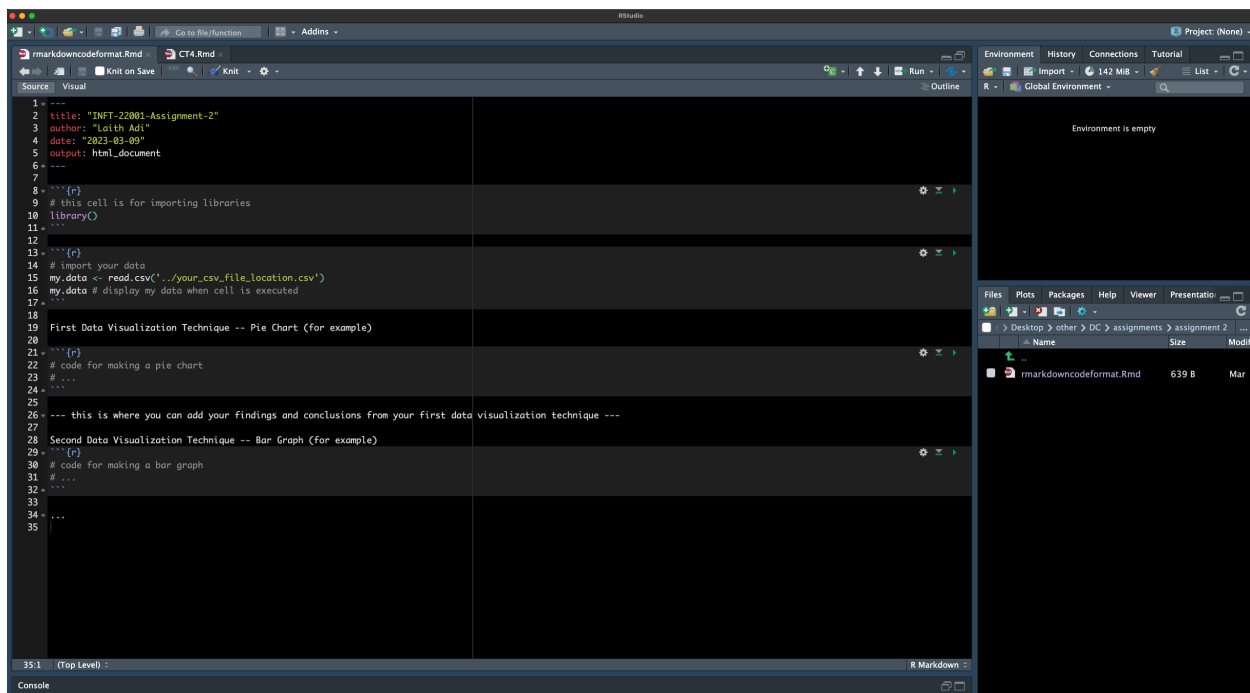
Setting up RStudio

Please create a R Markdown file.



An R Markdown file acts like Jupyter Notebook. You can create a cell to write and run your code in, and have sections where you can add plain text.

R Markdown Example



Where to get started

1. Please let me know which dataset you decided to go with BEFORE Saturday March 11.
 - I would like to make sure you get started on the assignment as soon as possible because you only have 1 week to complete it. Also, I need to make sure that not all of you picked the same dataset.
 - Pick something you are interested in. Kaggle has a lot of data on almost anything you can think of.
2. Set up your R Markdown file. Import your data.
 - Look through your data (either in RStudio or Excel) to see what sort of EDA you can perform on the data.
3. Decide on the three data visualization techniques you will be performing on your data.
4. Get started on your first data visualization technique!

Good Luck!