Department of Computing & Technology, Iqra University, Islamabad Campus
Data Science/Special Topics in Software Engineering

Assignment 1 (Max. Marks: 100)

Max Group Size: 3 members (each member will be evaluated separately)

Deadline: 26.10.2021 (Only one member will submit on Blackboard)

Submissions: Code + Document

It is usually hard to see patterns in a dataset when viewing it in its raw form. Computational tools like summary tables and visualizations make it easier to find interesting trends and relationships within the data.

Note: A Summary table is a table of aggregate information about a dataset (e.g., the average, sum, count of some values).

The given dataset is a simple dataset about Cereals. In general, the data set is self-explanatory and does not need any further explanation.

For the given dataset, discover interesting "*data stories*" :

- Read the columns of the dataset to understand the data that was collected.
- Use visualization to help find an interesting pattern in your dataset.
- You can also make summary tables
- You should find some interesting trends or relationships within your data. Next, you'll want to communicate your findings in a way that others can easily understand. Make sure you are making good choices for presenting the pattern or trend you have found.

Answer these questions in your answer document:

A. Briefly describe the data set you worked with (limit 50 words).
B. Make the summary tables and visualizations which you found to be the most interesting or insightful. Paste it in your document. The types of the charts are left to your creativity. State why you think it is most interesting and insightful.
C. Describe the trend, pattern, or relationship you found within your data. How is it shown with your visualization? (50 - 100 words)
D. Come up with a possible story or explanation for the trend you described. Make sure to note any assumptions you are making in this interpretation of the data.