**Data Mining Assignment 1 (Statistical Representation of Data)**

Note: Follow the deadline. Time will not be extended. The assignment should be submitted in the zip file with python code and report. The zip file name should contain your name and roll number

1. Import CEREAL dataset from the given link <https://www.kaggle.com/code/cameronbujaucius/starter-80-cereals-28561345-7/data>
2. Calculate Statistical Measures: Find all the statistical measures(Mean Median, Mode, Five number summary) for each of the attributes. If the attribute is categorical just find the mode. Write your report with screenshot of code and output.
3. Graphical Representation: For each question, answer and write your analysis with the code and output screen shots to support your answer)
   1. Is there a relation between sugars, calories, carbs, and fat? (Scatter plot for the above-mentioned attributes)
   2. How are calories and potassium distributed? (Seperate histograms of these attributes with their frequencies)
   3. Which manufacturers produce cereal with highest calories? (Bar Graph of the concerned attributes)
   4. How does rating compare to calorie count? (Scatter plot of the two attributes)
   5. Which nutrients are essential for a good rating for a cereal? (Scatter matrix for nutrients and ratings of the attributes)
   6. Is there a relation between manufacturer and rating? (Box plots all manufacturers and ratings)
   7. Is there a relation between shelf number and rating? (Box plots all shelves and ratings)