**About Dataset**

**Context**

If you like to eat cereal, do yourself a favor and avoid this dataset at all costs. After seeing these data it will never be the same for me to eat Fruity Pebbles again.

**Dataset Name: cereal.csv**

**Content:**

Fields in the dataset:

* Name: Name of cereal
* mfr: Manufacturer of cereal
  + A = American Home Food Products;
  + G = General Mills
  + K = Kelloggs
  + N = Nabisco
  + P = Post
  + Q = Quaker Oats
  + R = Ralston Purina
* type:
  + cold
  + hot
* calories: calories per serving
* protein: grams of protein
* fat: grams of fat
* sodium: milligrams of sodium
* fiber: grams of dietary fiber
* carbo: grams of complex carbohydrates
* sugars: grams of sugars
* potass: milligrams of potassium
* vitamins: vitamins and minerals - 0, 25, or 100, indicating the typical percentage of FDA recommended
* shelf: display shelf (1, 2, or 3, counting from the floor)
* weight: weight in ounces of one serving
* cups: number of cups in one serving
* rating: a rating of the cereals (Possibly from Consumer Reports?)

**Task 1:**

* Load the following libraries :

1. Pandas
2. Matplotlib
3. Seaborn

**Task 2:**

* Read the data file and give it suitable name

**Task 3:**

* Inspect the data frame using frequently used functions.

**Task 4:**

* Using mfr: Manufacturer of cereal columns find the number of cereal per manufacturer
* Visualize it using count plot and pie chart.

**Task 5:**

* Create a Histogram for the distribution of calories

**Task 6:**

* Visualize the distribution of 'vitamins' content in cereals. You can use a histogram for this

**Task 7:**

* Find the correlation matrix
* Create a correlation heatmap.

**Task 8:**

* Find the correlation between Calories vs Protein.
* Visualize this using scatter plot.

**Task 9: Challenge Task:**

Create Boxplot for the following columns:

1. Calories
2. Fiber
3. Subars
4. Carbohydrates

Comment on the above boxplots, are there any outliers? Can you remove outliers?