Basic Statistics

### Descriptive Analytics and Data Preprocessing on Sales & Discounts Dataset

#### Introduction

* To perform descriptive analytics, visualize data distributions, and preprocess the dataset for further analysis.

#### Descriptive Analytics for Numerical Columns

* Objective: To compute and analyze basic statistical measures for numerical columns in the dataset.
* Steps:
  + Load the dataset into a data analysis tool or programming environment (e.g., Python with pandas library).
  + Identify numerical columns in the dataset.
  + Calculate the mean, median, mode, and standard deviation for these columns.
  + Provide a brief interpretation of these statistics

Solution ->

1. Dataset is loaded in jupyter. The data is analyzed using the pandas Library.
2. The following are the numerical columns identified using pandas’ info()

* Avg Price
* Total Sales Value
* Discount Rate (%)
* Discount Amount
* Net Sales Value

1. The mean, median, and mode values of the numerical column are as follows->
2. Avg. Price->

* Mean of Avg Price = 10453.433333333332
* Median of Avg Price = 1450.0
* Mode of Avg Price = 400,450,500,1300,8100. (Occurrence Count = 30)

1. Total Sales Value->

* Mean of Total Sales Value = 33812.83555555555
* Median of Total Sales Value = 5700.0
* Mode of Total Sales Value = 24300

1. Discount Rate (%)

* Mean of Discount Rate (%) = 15.155241895330914
* Median of Discount Rate (%) = 16.57776564929597
* Mode of Discount Rate (%) = As each discount Rate is different, the mode should contain all

the values, as the frequency of each value is one.