**Lab#02**

Data Science Lab Dated: Feb. 19, 2024

List of Lab Activities:

1. Identify the outliers visually using box plot and make a subset of outliers using various columns.
2. Replace the outliers either it is missing value or any other type of outlier with median value of the said column.
3. Do previous two steps on all columns.
4. Statistically, it is proven that sample mean is always equals to the population mean. To prove this concept, you need to start to select a subset of data for a numeric column randomly and increase it sequence and keep computing mean and compare it with the mean of the selected column. Stop the loop once you reach the condition of sample mean and population mean equality.
5. As you are aware from area under the curve. Consider the numeric columns of the dataset, apply the Z-scale formula and apply it on the numeric columns of the dataset. Find the mean and standard deviation of the numeric columns and compute the percentage of data within 1, 2, and 3 Standard deviation. [hint: Mean+/-1SD]