

S3 MCA (2023-25 Batch)

20MCA241 Data Science Lab

Lab Exercise #04

Date of Submission: 24 Aug 2024

1. Learn **Exploratory Data Analysis** using the following links and prepare your own worksheet.

- <https://www.kaggle.com/code/harshjain123/breast-cancer-exploratory-data-analysis>
- [https://github.com/sara-kassani/Python-Machine-Learning-Codes/blob/master/Visualization - Breast Cancer Wisconsin \(Diagnostic\) Data Set.ipynb](https://github.com/sara-kassani/Python-Machine-Learning-Codes/blob/master/Visualization%20-%20Breast%20Cancer%20Wisconsin%20(Diagnostic)%20Data%20Set.ipynb)

2. Download the **breast cancer dataset** from the following link

- <https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data>
- Using Python, visualize the various characteristics of the dataset.
- Use Google Colab for implementation.
- Define your own conditions to learn the characteristics
- After each visualization write your findings
- Use the following types of charts for **visualization**:
 - i. Histogram
 - ii. Quartile (Box) plot
 - iii. Distribution chart
 - iv. Scatterplot
 - v. Scatter multiple
 - vi. Scatter matrix
 - vii. Bubble chart
 - viii. Density chart
 - ix. Parallel chart
 - x. Deviation chart
 - xi. Andrews curves