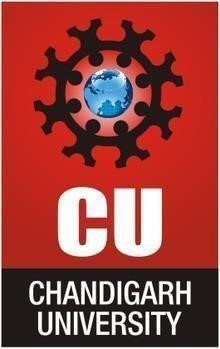
**CHANDIGARH UNIVERSITY**

**UNIVERSITY INSTITUTE OF ENGINEERING**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



|  |  |
| --- | --- |
| **Submitted by:** | **Submitted To:** Ajay Kumar (E13141) |
| **Subject Name:** | Machine Learning Lab |
| **Subject Code:** | 20CSP-317 |
| **Branch:** | CSE |
| **Semester:** | 5th |

LAB INDEX

1. **Aim/Overview of the practical:** Implement Exploratory Data Analysis on any data set.
2. **Requirements:** Breast Cancer Dataset, Jupyter notebook or Google collab**.**

1. **Description/ Code:**

**# importing all necessary libraries**

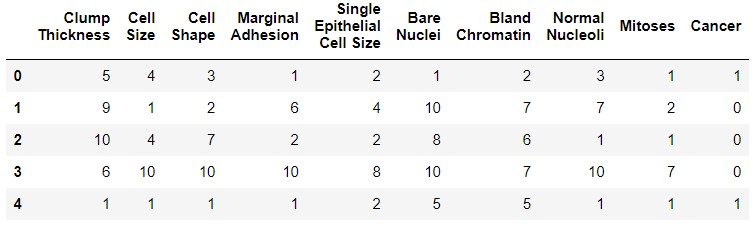
import numpy as np

import pandas as pd

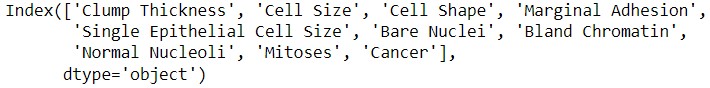
**# Reading the data**

breast\_data=pd.read\_csv("Sample5 - Brest Cancer.csv")

breast\_data.head()



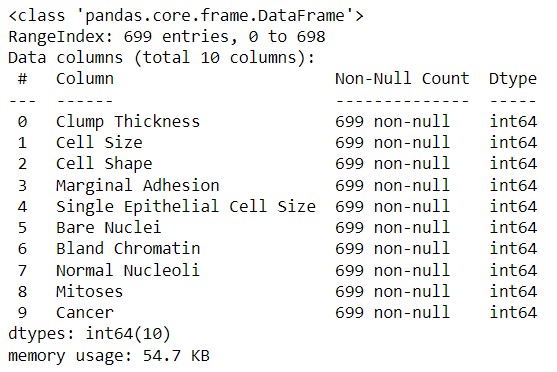
breast\_data.columns



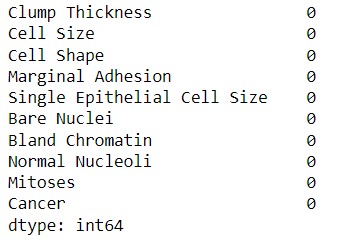
breast\_data.shape



breast\_data.info()

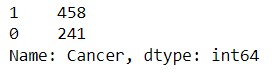


breast\_data.isnull().sum()



**# Checking the wheather the data is imbalance or not**

breast\_data["Cancer"].value\_counts()



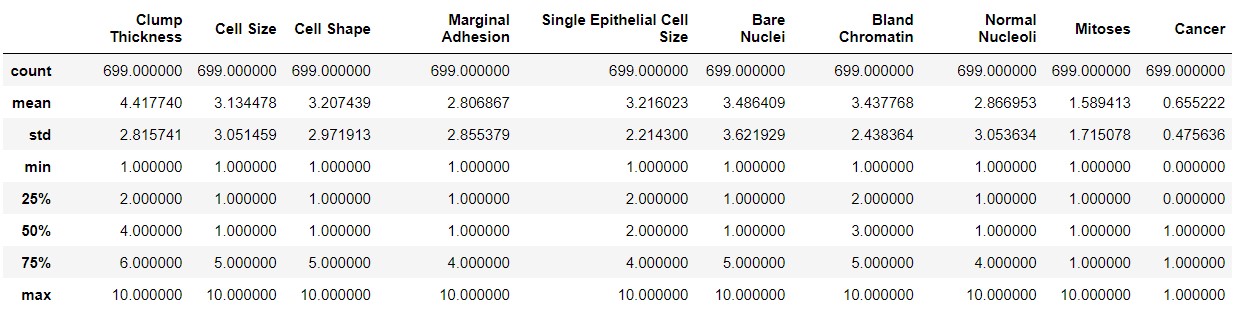
**#### Our dataset is imbalance**

print("Percentage of 0's in dependent variable is",241/(241+458))

print("Percentage of 1's in dependent variable is",458/(241+458))



breast\_data.describe()



**Learning outcomes (What I have learnt):**

**1. Exploratory Data Analysis**

**Evaluation Grid (To be created as per the SOP and Assessment guidelines by the faculty):**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No. | Parameters | Marks Obtained | Maximum Marks |
| 1. |  |  |  |
| 2. |  |  |  |
| 3. |  |  |  |
|  |  |  |  |