|  |  |
| --- | --- |
| **Branch:** | BE ETRX |
| **Batch:** | A |
| **Course:** | Minor – Computer Science |
| **Subject:** | Machine Learning |
| **Student Name:** | Dhananjay Joshi |
| **Experiment No.** | 01 |
| **UID:** | 2019110021 |

Aim:  
Import the dataset and perform EDA such as number of data samples, number of features, number of classes, number of data samples per class, removing missing values, conversion to numbers, explore dimensionality, type the mean or average value, and using seaborn library to plot different graphs. Consider one of the datasets given below.

1. [NASA](https://nasa.github.io/data-nasa-gov-frontpage/#_blank)**:** If you’re interested in space and earth science, see what you can find among the tens of thousands of public datasets made available by NASA.

### Software Used: Google Colab Notebook

### Dataset Description: Twentieth Century Crop Statistics 1900-2017 This data consists of national and subnational maize and wheat production, yield and harvested area statistics for all available years. It combines statistics from Italy, Spain, Indonesia, Chine, Mexico, Uruguay, Chile, Sweden, Morocco, India, Australia, USA, Canada and many more countries. All units are in hectares for harvested areas, tonnes for production and tonnes/ha for yield.

### Code:

### Conclusion: