**Credit Card Fraud Detection**

**PHASE-3**

**PROGRAM:**

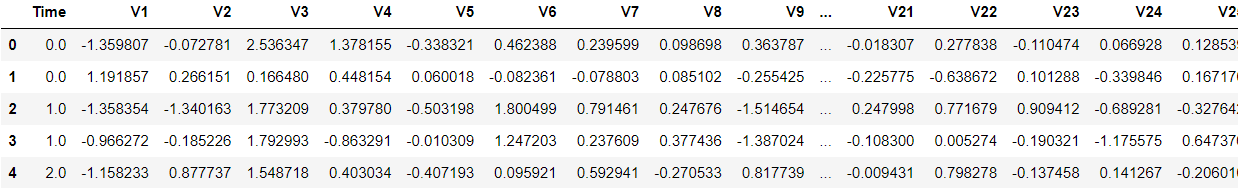
**import pandas as pd**

**from collections import Counter**

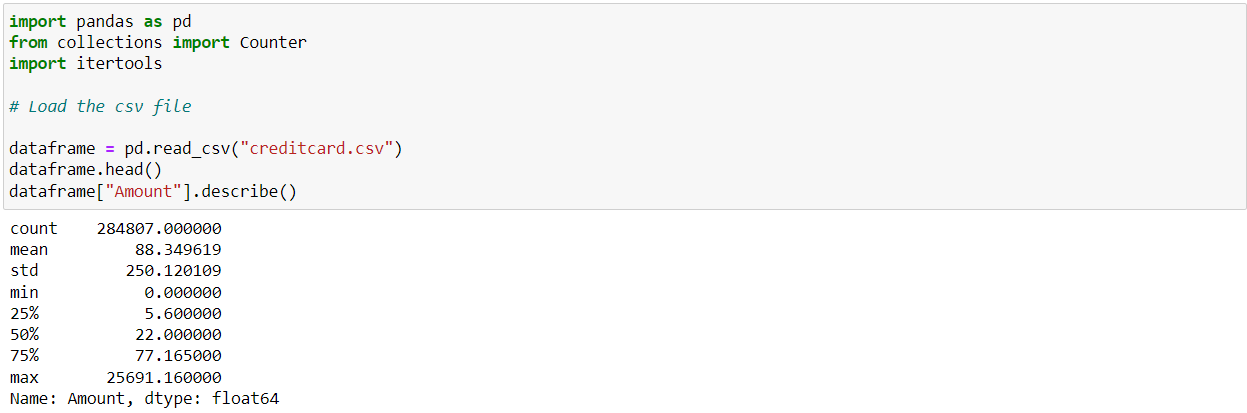
**import itertools**

**dataframe = pd.read\_csv("creditcard.csv")**

**dataframe.head()**

****

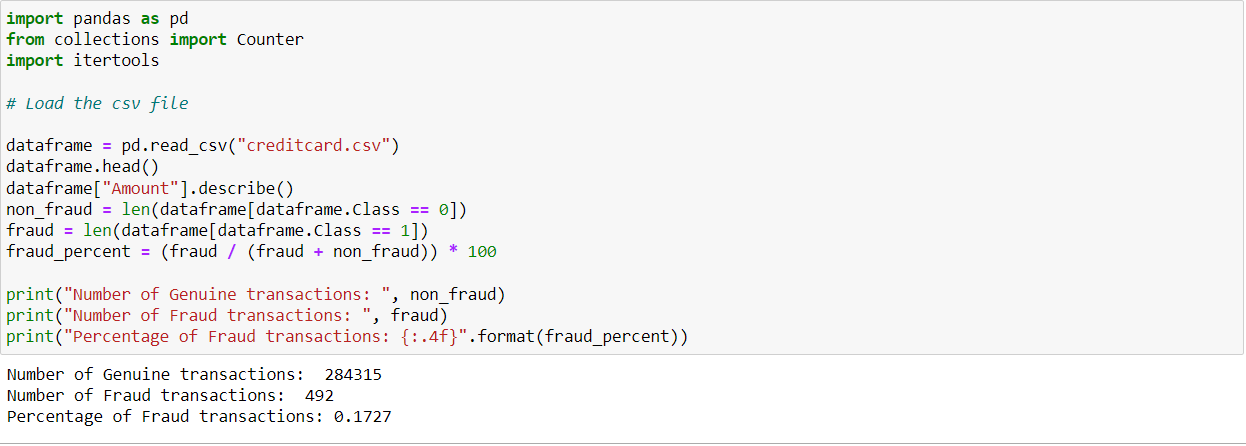
**dataframe["Amount"].describe()**

****

**non\_fraud = len(dataframe[dataframe.Class == 0])**

**fraud = len(dataframe[dataframe.Class == 1])**

**fraud\_percent = (fraud / (fraud + non\_fraud)) \* 100**

****

**import matplotlib.pyplot as plt**

**labels = ["Genuine", "Fraud"]**

**count\_classes = dataframe.value\_counts(dataframe['Class'], sort= True)**

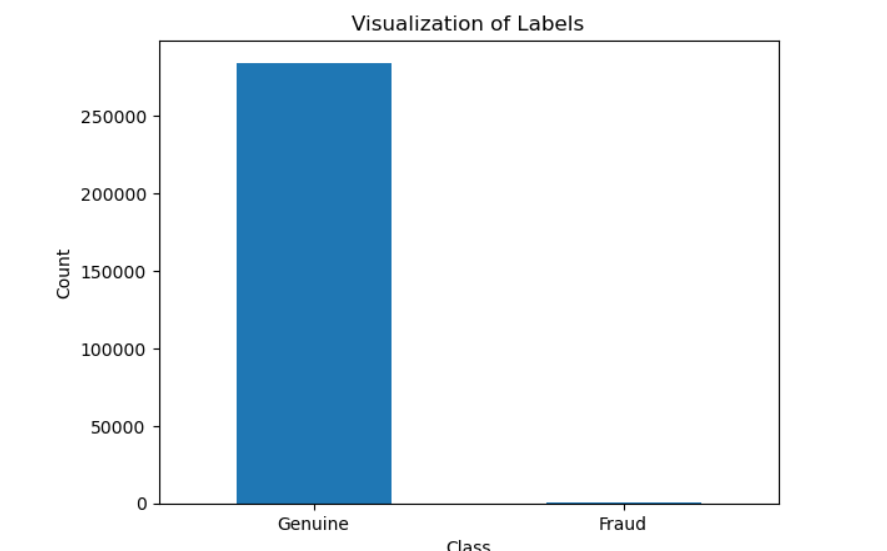
**count\_classes.plot(kind = "bar", rot = 0)**

**plt.title("Visualization of Labels")**

**plt.ylabel("Count")**

**plt.xticks(range(2), labels)**

**plt.show()**

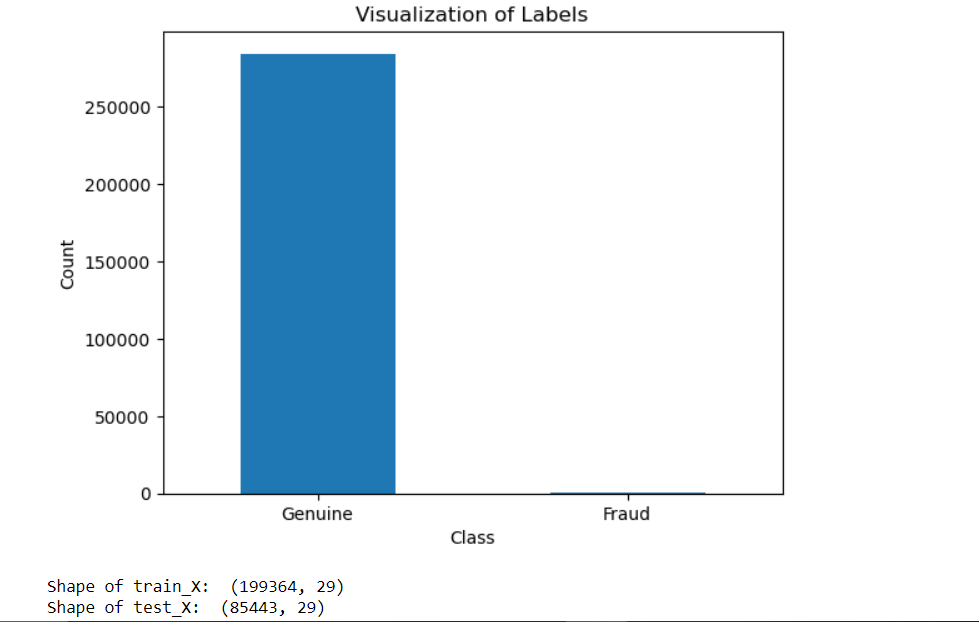
****

**from sklearn.model\_selection import train\_test\_split**

**(train\_X, test\_X, train\_Y, test\_Y) = train\_test\_split(X, Y, test\_size= 0.3, random\_state= 42)**

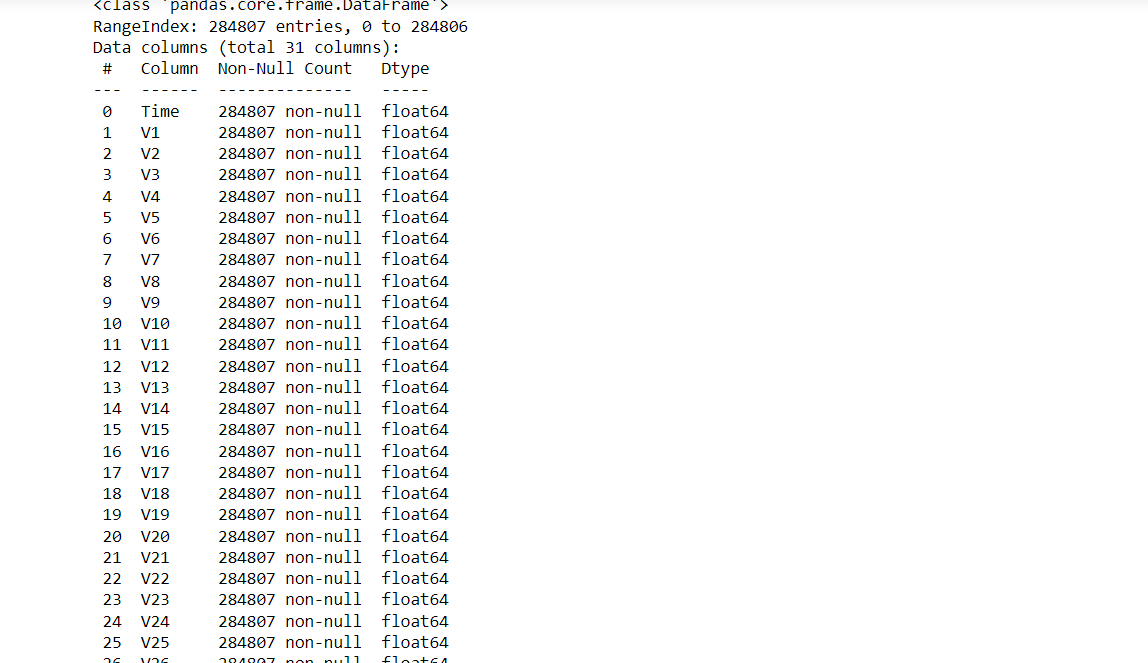
**print("Shape of train\_X: ", train\_X.shape)**

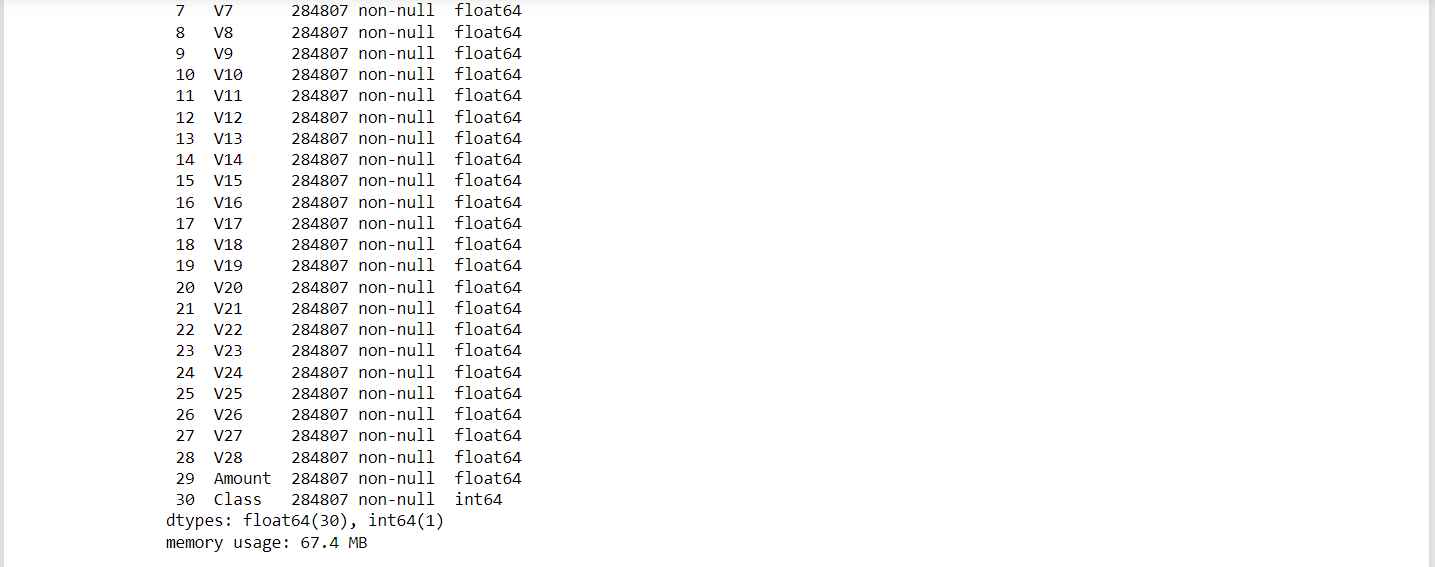
**print("Shape of test\_X: ", test\_X.shape)**

****

**dataframes.head()**

**dataframes.info()**

****

****

### About Credit Card Fraud Detection:

**In this machine learning project, we solve the problem of detecting credit card fraud transactions using machine numpy, scikit learn, and few other python libraries. We overcome the problem by creating a binary classifier and experimenting with various machine learning techniques to see which fits better.**

### Credit Card Fraud Dataset

**The dataset consists of 31 parameters. Due to confidentiality issues, 28 of the features are the result of the PCA transformation. “Time’ and “Amount” are the only aspects that were not modified with PCA.**

**There are a total of 284,807 transactions with only 492 of them being fraud. So, the label distribution suffers from imbalance issues.**

**Tools and Libraries used**

**We use the following libraries and frameworks in credit card fraud detection project.**

* **Python – 3.x**
* **Numpy – 1.19.2**
* **Scikit-learn – 0.24.1**
* **Matplotlib – 3.3.4**
* **Imblearn – 0.8.0**
* **Collections, Itertools**

### Credit Card Fraud Project Code

**Please download the source code of the credit card fraud detection project (which is explained below):**[**Credit Card Fraud Detection Machine Learning Code**](https://data-flair.s3.ap-south-1.amazonaws.com/machine-learning-projects/credit-card-fraud-python-ml-code.zip)**.**

### Steps to Develop Credit Card Fraud Classifier in Machine Learning

**Our approach to building the classifier is discussed in the steps:**

1. **Perform Exploratory Data Analysis (EDA) on our dataset**
2. **Apply different Machine Learning algorithms to our dataset**
3. **Train and Evaluate our models on the dataset and pick the best one**.

#### Perform Exploratory Data Analysis (EDA)

**There are a total of 284,807 transactions with only 492 of them being fraud. Let’s import the necessary modules, load our dataset, and perform EDA on our dataset. Here is a peek at our dataset.**

**CONCLUSION:**

**Building the project by loading and preprocessing the dataset. Begin building the credit card fraud detection project by loading and preprocessing the dataset.Load the credit card transaction dataset and preprocess the data for analysis**.