**Dataset overview**

What is dataset?

Dataset is the basic or unclean data for or machine learning project.it is basically in the form of the .csv file (excel file) . we need to clean our data set for using some basic algorithm

Ex-

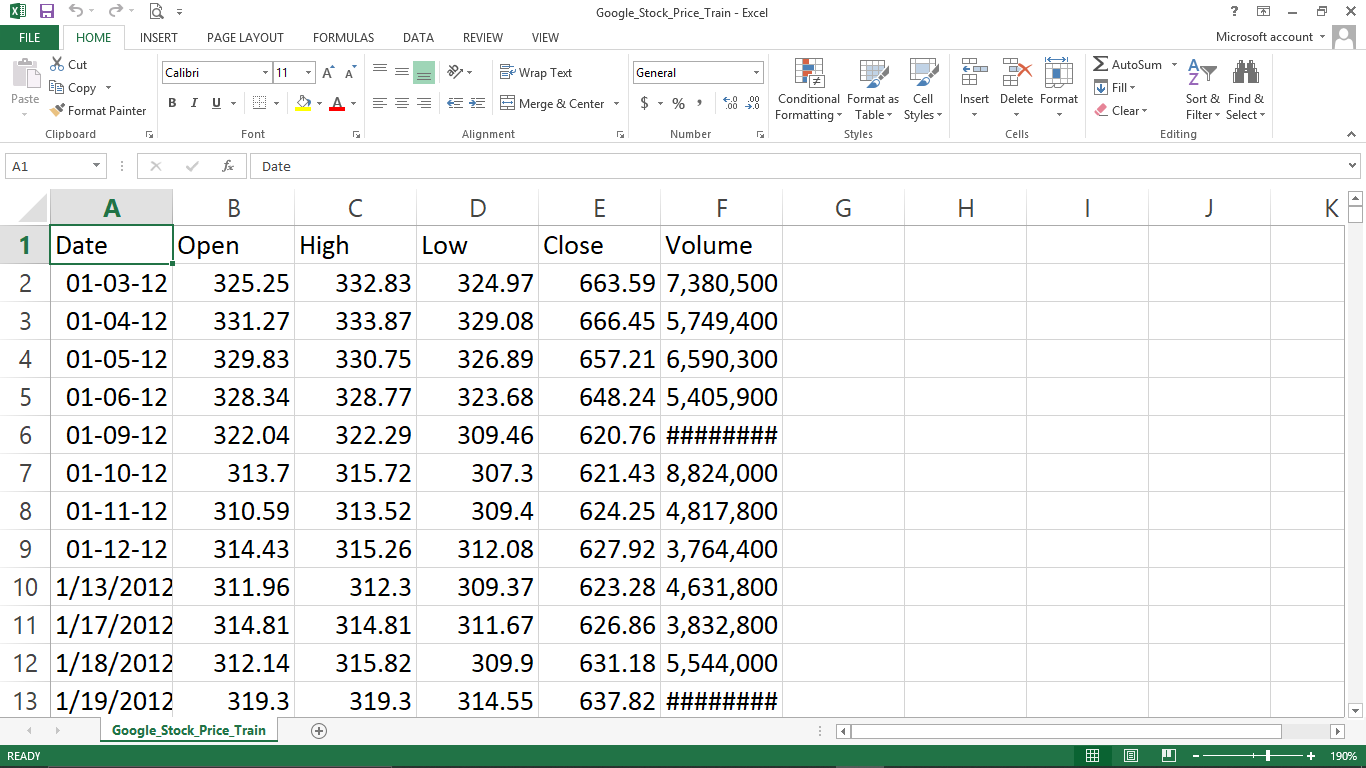
* like we need to do fill missing value in the data set
* we need to find mean and then fill that missing values
* or else we can delete that particular entry too

so for our stock market prediction project we need some basic data

like-

* open price
* closing price
* volume
* high for the day
* low for the day

so we can find many raw data for this project but I use Google stock price dataset from kaggel.as well as I am attaching the one photo of our dataset



**Implemented Functionality**

In this project we created many functionality

* rolling mean for 7days
* comparing original value and predicted value
* graph plotting of particular stock
* dataset information
* Not applicable value return
* Cleaning dataset
* Converting object value in to float
* Model training

So this project is based on stock market prediction using machine learning(stocker) in this project we need to import first basic library of python to do operation on our dataset so using “pandas” library we can import our dataset in our project code then we need to clean our data like filling missing value and deleting the data entry and calculating rolling mean so this is basic functionality

in next stage we can plot the graph of that data entries so for this particular project we can plot graph of opening price of the stock then next and very important functionality is we can observe that if there is any not applicable value in the data set column then change our dataset or we need to clean our dataset more so it is main functionality then we need to convert our object value in float so we can easily monitor the value

in the third stage of the project we need to calculate 7 day rolling mean so using that we can apply proper algorithm for market prediction after that we will also comparing the original value and 7 day rolling mean in one graph so this is the basic but very important functionality of our project

after this the vey most important functionality is to train our machine learning model so for that our machine learning model take hundred input from our dataset after that prediction part we need to check once again if there if any not applicable value then we need to clean it in this part our algorithms create data frame and it take 0 to 60 days as input and predict 61th day output and it will go on

after that this is last stage of our project and it is to plot graph of that predicted out vs. original dataset using so we can measure successful ratio of our predicted output

this is basic and advanced functionality of our project

**Results and Reports**

**Basic code and its results**

1. **Basic imports**

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

import datetime

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Dense

from tensorflow.keras.layers import LSTM

from tensorflow.keras.layers import Dropout

from sklearn.preprocessing import MinMaxScaler

1. **dataset import part**

dataset =pd.read\_csv (r"C:\Users\HP\data set\Google\_Stock\_Price\_Train.csv",index\_col="Date", parse\_dates=True)

1. **dataset review**

dataset.head()

dataset.tail()

print(dataset)

**4) not applicable values & dataset information**

dataset.info()

dataset.isna().any()

**5)graph plotting**

dataset['Open'].plot(figsize=(16,6))

dataset['Open'].plot(figsize=(16,6))

dataset.rolling(window=30).mean()['Close'].plot()

dataset['Close: 30 Day Mean'] = dataset['Close'].rolling(window=30).mean()

dataset[['Close', 'Close: 30 Day Mean']].plot(figsize=(16,6))

**6) converting objects into float**

dataset["Close"] = dataset["Close"].str.replace(',', '').astype(float)

dataset["Volume"] = dataset["Volume"].str.replace(',', '').astype(float)

**7) rolling mean**

dataset.rolling(7).mean().head(20)

**8)data frame**

training\_set=dataset["Open"]

training\_set=pd.DataFrame(training\_set)

**9)model training**

epoch 1/100

1198/1198 [==============================] - 12s 10ms/step - loss: 0.0519

/ /

/ /

epoch 100/100

1198/1198 [==============================] - 6s 5ms/step - loss: 0.0014

**10) original graph vs predicted**

plt.plot(real\_stock\_price, color = 'red', label = 'Real Google Stock Price')

plt.plot(predicted\_stock\_price, color = 'blue', label = 'Predicted Google Stock Price')

plt.title('Google Stock Price Prediction')

plt.xlabel('Time')

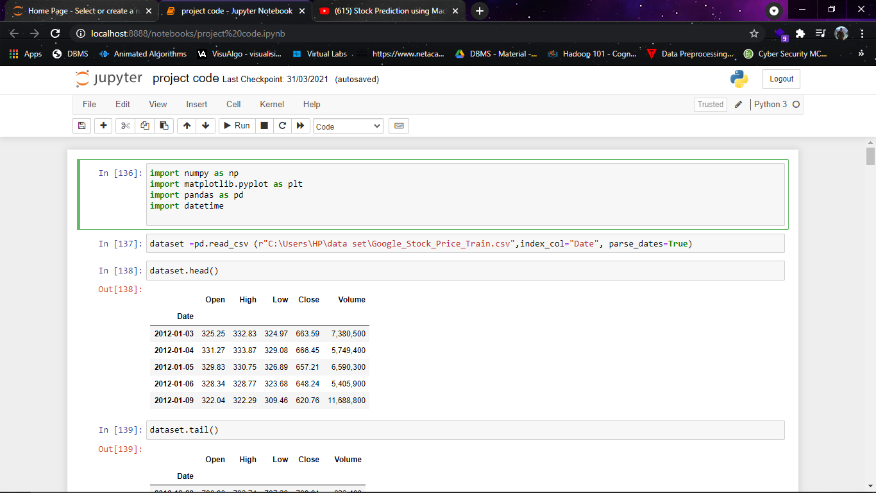
plt.ylabel('Google Stock Price')

plt.legend()

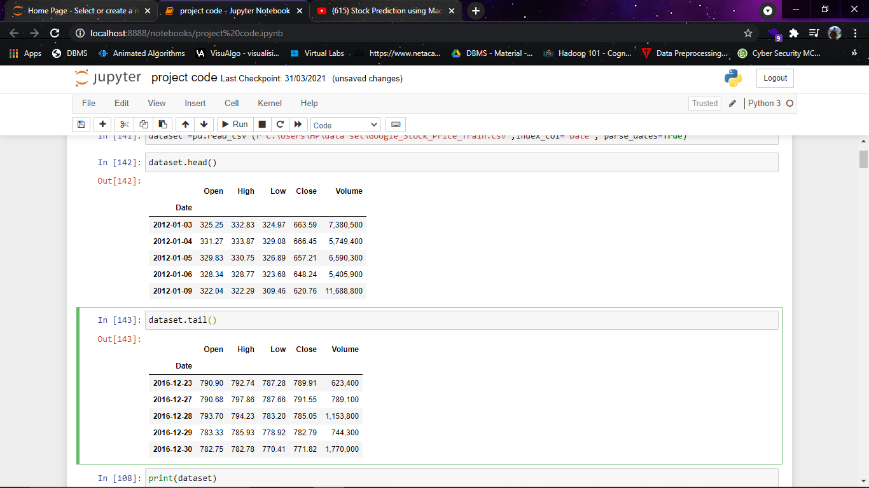
plt.show()

**Snapshots**

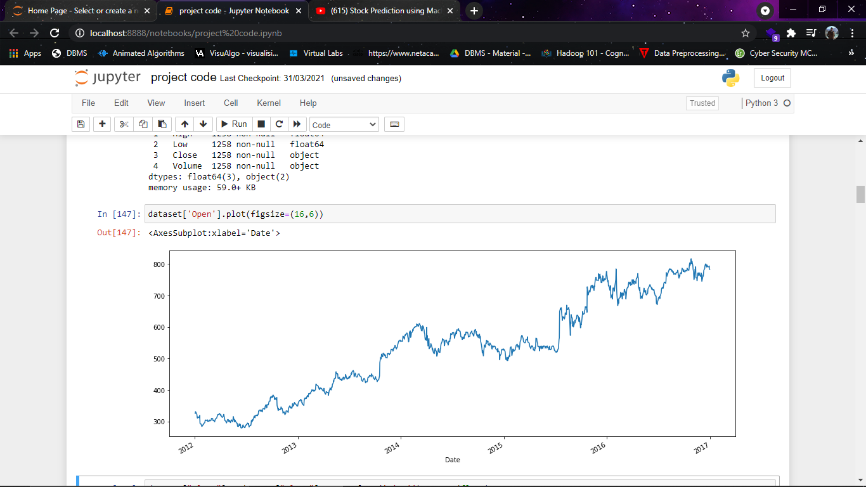
1)basis import

****

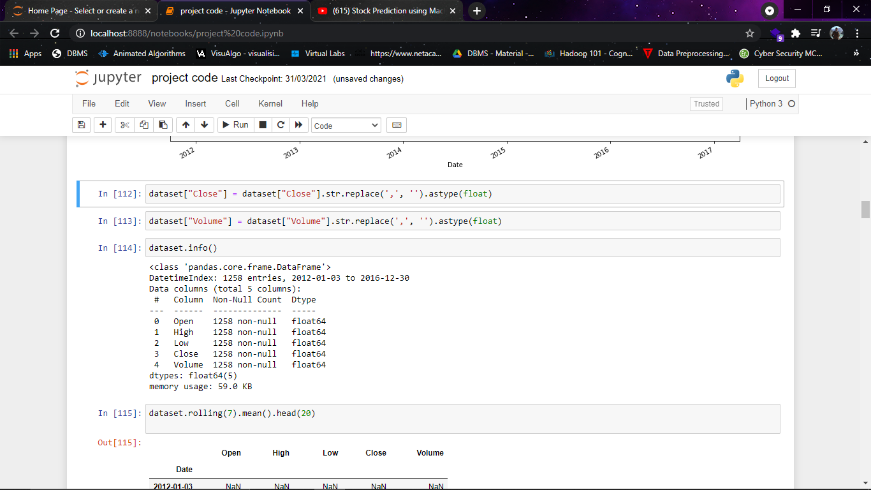
2)dataset load



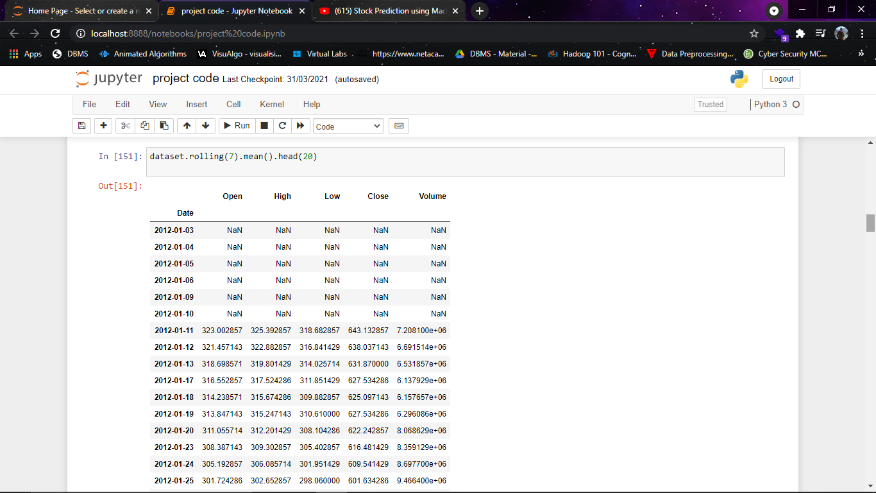
3)graph plotting of open price of stock(daily basis)



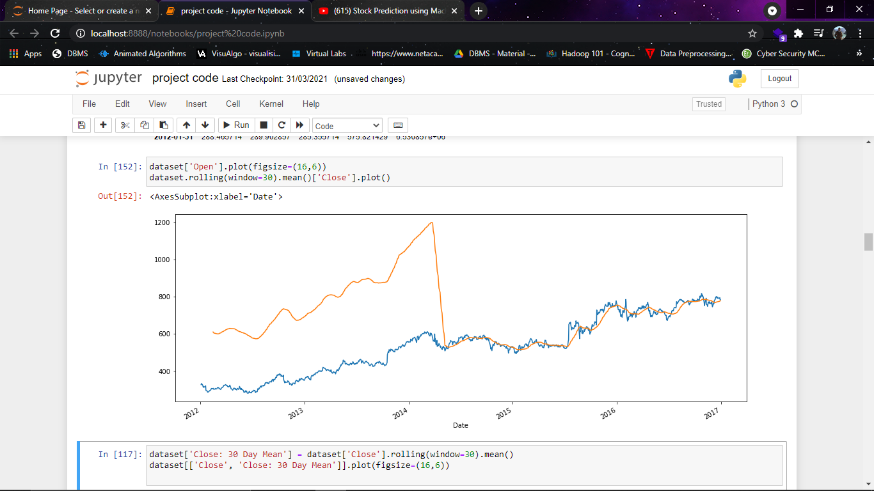
4)dataset information



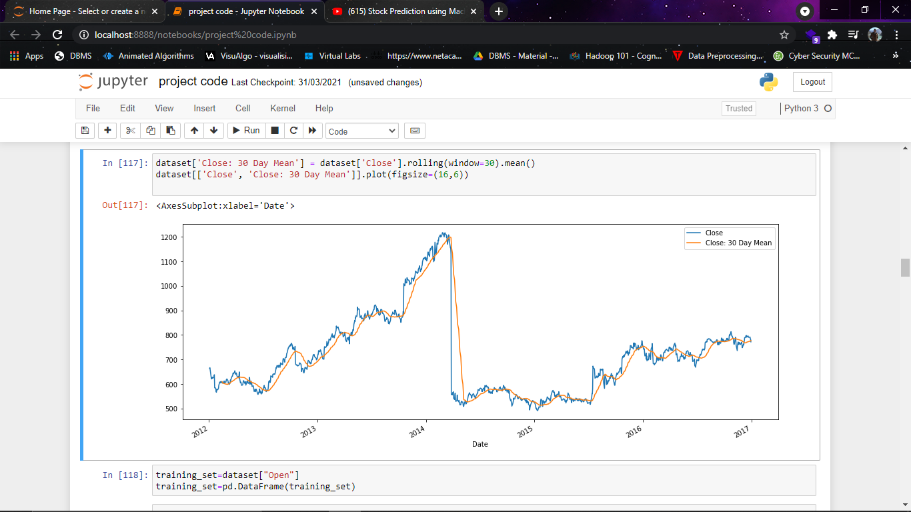
5)calculating rolling mean



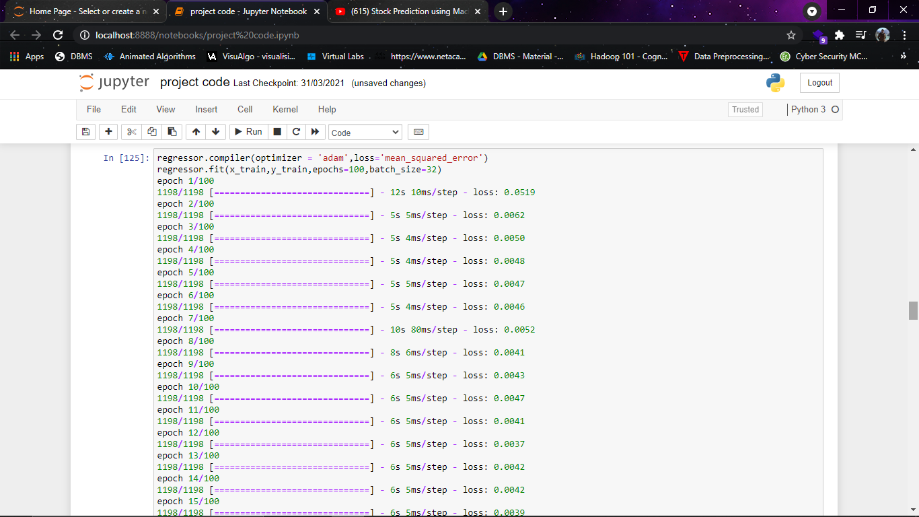
6)7 days rolling mean vs. original



7)30days rolling mean vs. original



8)machine model turning



9)final result

