Assignment32: Session 32 Machine Learning 11

**2. Problem Statement**

In this assignment students have to make ARIMA model over shampoo sales data and

check the MSE between predicted and actual value.

Student can download data in .csv format from the following link:

https://datamarket.com/data/set/22r0/sales-of-shampoo-over-a-three-yearperiod#!

ds=22r0&display=line

Hint:

Following is the command import packages and data

from pandas import read\_csv

from pandas import datetime

from matplotlib import pyplot

from statsmodels.tsa.arima\_model import ARIMA

from sklearn.metrics import mean\_squared\_error

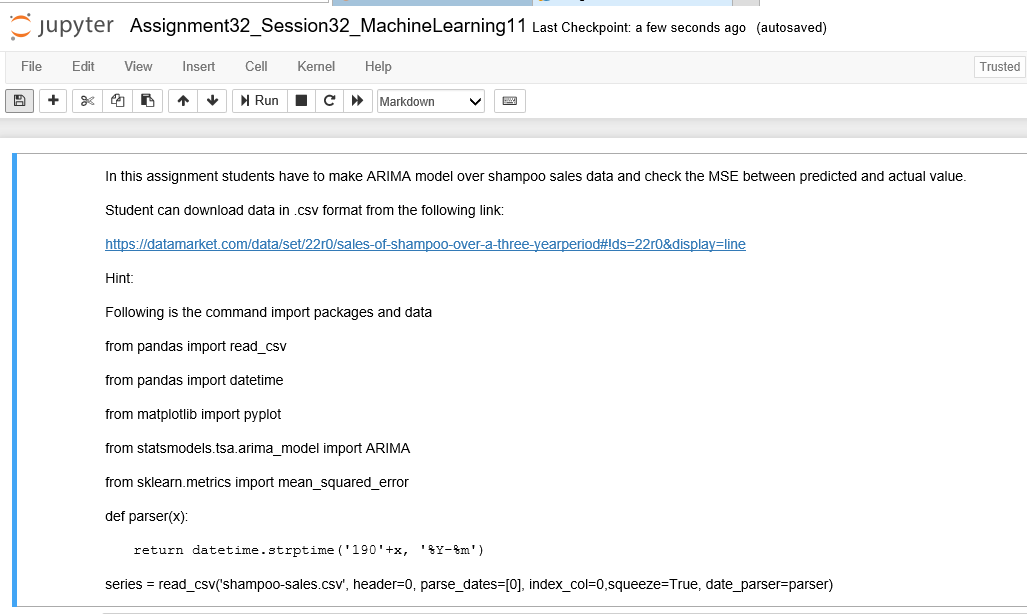
def parser(x):

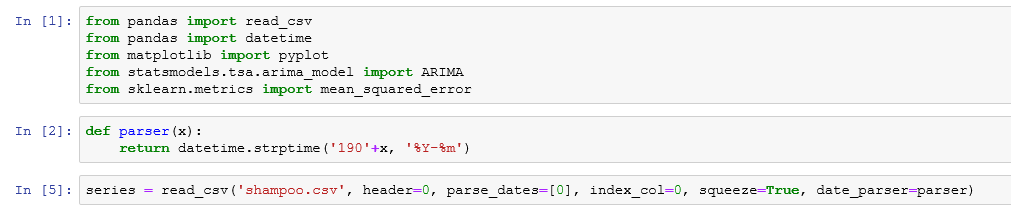
return datetime.strptime('190'+x, '%Y-%m')

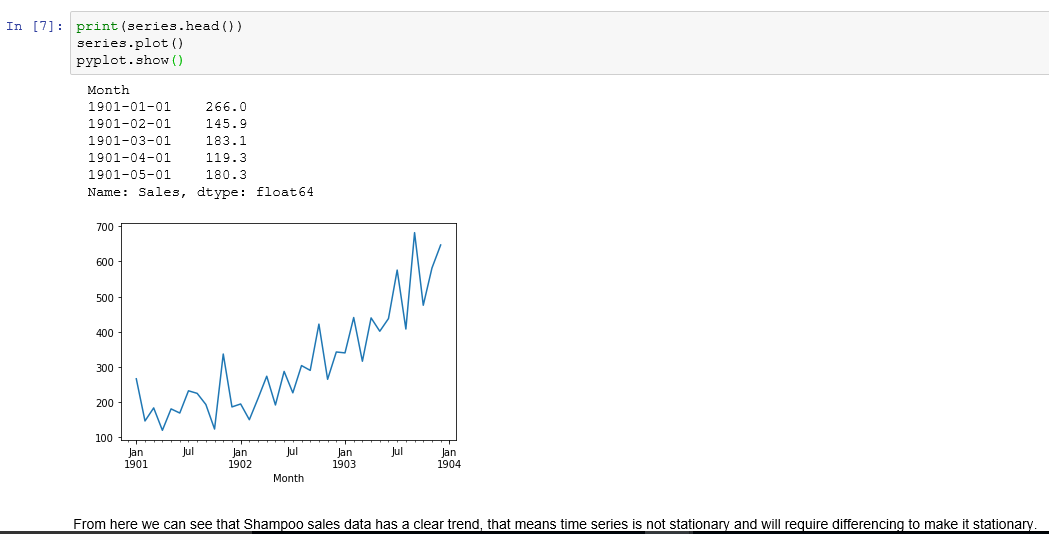
series = read\_csv('shampoo-sales.csv', header=0, parse\_dates=[0], index\_col=0,

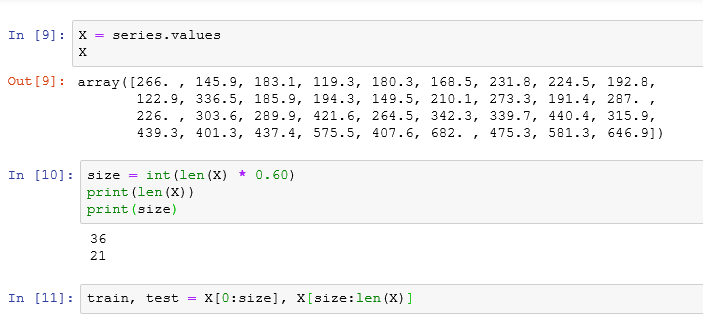
squeeze=True, date\_parser=parser)

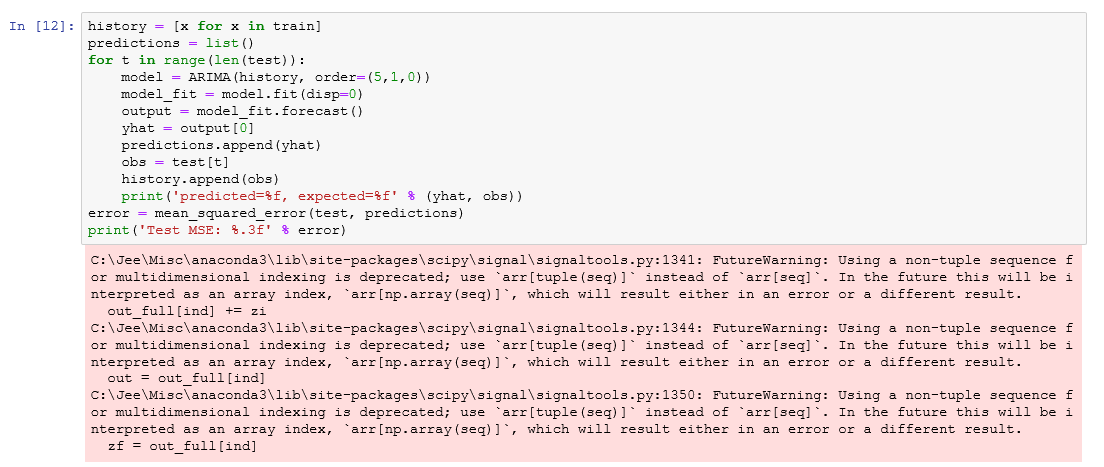
**3. Output**

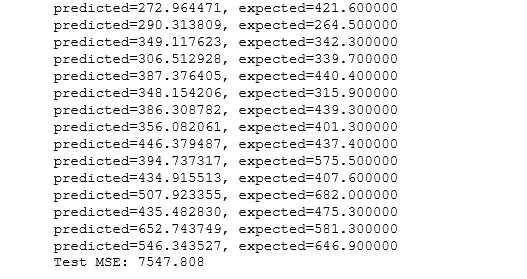


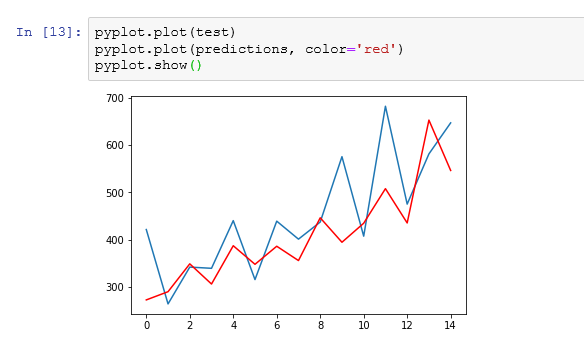


****

****

****

****

****