**EXERCISE:02**

**Reg.no:953622243037**

**Date:02-10-2023**

**Objective :**

**Working with Numpy arrays , Pandas Data frames , Basic plots using Matplotib**

**Dataset Description :**

**The dataset is obtained by daily analysis of the page of the site selling macbooks. A CSV file stores data about the cost of a particular macbook model (row) on a particular day (column). The data is updated approximately once a week.**

import **pandas** as **pd**import **matplotlib.pyplot** as **plt**import **seaborn** as **sns**import **numpy** as **np  
df=pd.read\_csv("/content/macbooks.csv")  
df**

**name \  
0 Apple MacBook Air 13 Retina MGND3 Gold (M1 8-C...   
1 Apple MacBook Air 13 Retina MGN93 Silver (M1 8...   
2 Apple MacBook Air 13 Retina MGN63 Space Gray (...   
3 Apple MacBook Pro 13 Retina Touch Bar MYD82 Sp...   
4 Apple MacBook Pro 13 Retina Touch Bar MYDA2 Si...   
.. ...   
128 Apple MacBook Pro 16 Z14Y0008J Silver (M1 Pro ...   
129 Apple MacBook Pro 14 Z15G000DW Space Gray (M1 ...   
130 Apple MacBook Pro 14 Z15H0007V Space Gray (M1 ...   
131 Apple MacBook Pro 16 Z1500004J Silver (M1 Max ...   
132 Apple MacBook Pro 16 Z14V00093 Space Gray (M1 ...   
  
 link 2022-06-07 \  
0 https://biggeek.ru/products/apple-macbook-air-... 99990.0   
1 https://biggeek.ru/products/apple-macbook-air-... 103990.0   
2 https://biggeek.ru/products/apple-macbook-air-... 103990.0   
3 https://biggeek.ru/products/apple-macbook-pro-... 129990.0   
4 https://biggeek.ru/products/apple-macbook-pro-... 129990.0   
.. ... ...   
128 https://biggeek.ru/products/apple-macbook-pro-... 699990.0   
129 https://biggeek.ru/products/apple-macbook-pro-... 739990.0   
130 https://biggeek.ru/products/apple-macbook-pro-... 789990.0   
131 https://biggeek.ru/products/apple-macbook-pro-... 849990.0   
132 https://biggeek.ru/products/apple-macbook-pro-... 919990.0   
  
 2022-06-08 2022-06-09 2022-06-10 2022-06-11 2022-06-13 2022-06-14 \  
0 97990.0 97990.0 97990.0 97990.0 97990.0 97990.0   
1 97990.0 97990.0 97990.0 97990.0 97990.0 97990.0   
2 97990.0 97990.0 97990.0 97990.0 97990.0 97990.0   
3 129990.0 129990.0 129990.0 129990.0 129990.0 129990.0   
4 129990.0 129990.0 129990.0 129990.0 129990.0 129990.0   
.. ... ... ... ... ... ...   
128 699990.0 699990.0 699990.0 699990.0 699990.0 699990.0   
129 739990.0 739990.0 739990.0 739990.0 739990.0 739990.0   
130 789990.0 789990.0 789990.0 789990.0 789990.0 789990.0   
131 849990.0 849990.0 849990.0 849990.0 849990.0 849990.0   
132 919990.0 919990.0 919990.0 919990.0 919990.0 919990.0   
  
 2022-06-15 ... 2022-09-24 2022-09-25 2022-09-26 2022-09-27 \  
0 97990.0 ... 75990.0 75990.0 75990.0 75990   
1 97990.0 ... 87990.0 87990.0 87990.0 78990   
2 97990.0 ... 79990.0 79990.0 79990.0 79990   
3 129990.0 ... 109990.0 109990.0 109990.0 89990   
4 129990.0 ... 109990.0 109990.0 109990.0 89990   
.. ... ... ... ... ... ...   
128 699990.0 ... 719990.0 719990.0 719990.0 719990   
129 739990.0 ... 739990.0 739990.0 739990.0 739990   
130 789990.0 ... 789990.0 789990.0 789990.0 789990   
131 849990.0 ... 859990.0 859990.0 859990.0 859990   
132 919990.0 ... 929990.0 929990.0 929990.0 929990   
  
 2022-09-28 2022-09-29 2022-09-30 2022-10-01 2022-10-02 2022-10-03   
0 75990 75990 75990 75990 75990 75990   
1 78990 78990 78990 78990 78990 78990   
2 79990 79990 79990 79990 79990 79990   
3 89990 89990 89990 89990 89990 89990   
4 89990 89990 89990 89990 89990 89990   
.. ... ... ... ... ... ...   
128 719990 719990 719990 719990 719990 719990   
129 739990 739990 739990 739990 739990 739990   
130 789990 789990 789990 789990 789990 789990   
131 859990 859990 859990 859990 859990 859990   
132 929990 929990 929990 929990 929990 929990   
  
[133 rows x 114 columns]**

**df.isnull().sum()**

**name 0  
link 0  
2022-06-07 13  
2022-06-08 13  
2022-06-09 13  
 ..  
2022-09-29 0  
2022-09-30 0  
2022-10-01 0  
2022-10-02 0  
2022-10-03 0  
Length: 114, dtype: int64**

**Observation:**

**From the above we can understand that the null values present in the dataset in every columns using the null operation. There are 13 not filled space in the 2022-06-07,2022-06-08,2022-06-09. It is showed by the function name as is null.**

**df.describe().T**

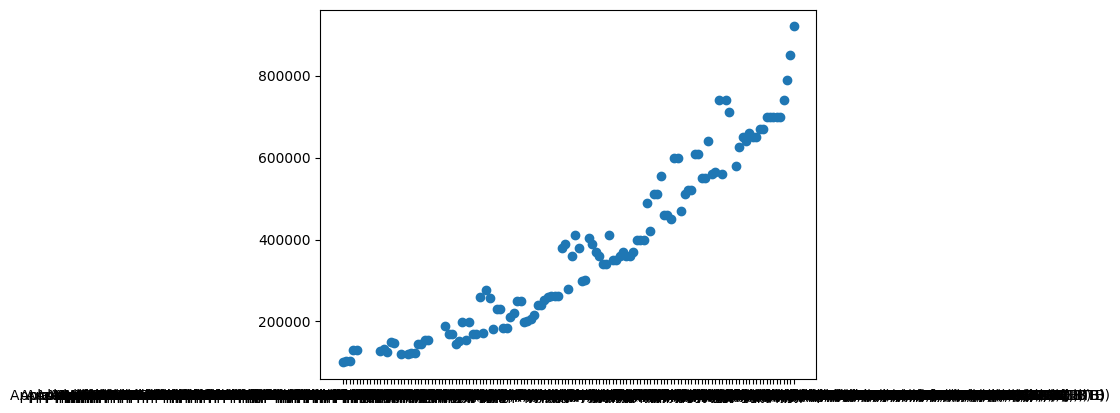
**count mean std min 25% 50% \  
2022-06-07 120.0 376731.666667 206539.983141 99990.0 197490.0 359990.0   
2022-06-08 120.0 376515.000000 206819.409818 97990.0 197490.0 359990.0   
2022-06-09 120.0 376515.000000 206819.409818 97990.0 197490.0 359990.0   
2022-06-10 120.0 376515.000000 206819.409818 97990.0 197490.0 359990.0   
2022-06-11 120.0 375098.333333 207857.934509 97990.0 184990.0 359990.0   
... ... ... ... ... ... ...   
2022-09-29 133.0 332350.902256 208825.225544 75990.0 144990.0 289990.0   
2022-09-30 133.0 332350.902256 208825.225544 75990.0 144990.0 289990.0   
2022-10-01 133.0 332350.902256 208825.225544 75990.0 144990.0 289990.0   
2022-10-02 133.0 332350.902256 208825.225544 75990.0 144990.0 289990.0   
2022-10-03 133.0 332350.902256 208825.225544 75990.0 144990.0 289990.0   
  
 75% max   
2022-06-07 551240.0 919990.0   
2022-06-08 551240.0 919990.0   
2022-06-09 551240.0 919990.0   
2022-06-10 551240.0 919990.0   
2022-06-11 551240.0 919990.0   
... ... ...   
2022-09-29 489990.0 929990.0   
2022-09-30 489990.0 929990.0   
2022-10-01 489990.0 929990.0   
2022-10-02 489990.0 929990.0   
2022-10-03 489990.0 929990.0   
  
[112 rows x 8 columns]**

**Observation:**

**From the above we can understand and know the value of mean, average, percentage by without calculating for many times. The program calculate the mean ,average ,percentage, max value ,min value for the macbook sold in the week.**

**plt.scatter(x=df["name"],y=df["2022-06-07"])**

**<matplotlib.collections.PathCollection at 0x79595232b940>**

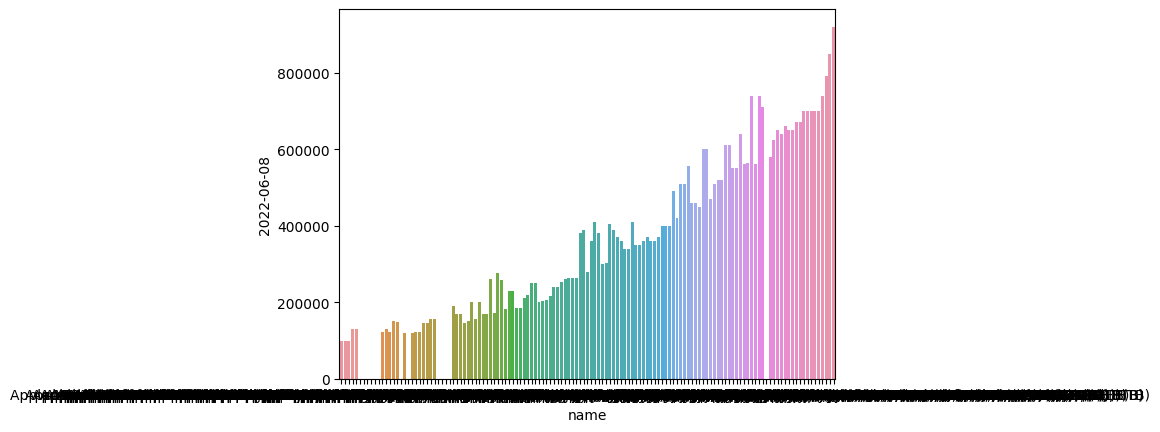
****

**Observation:**

**From the above scatter plot we defined the name of the macbook and who they sold in date:2022-06-07 . The Apple MacBook Pro 16 Z14V00093 Space Gray (M1 Max 10-Core, GPU 32-Core, 64GB, 8TB), this macbook which has the highest sale price than any other macbook version in 2022-06-07. The Apple MacBook Air 13 Retina MGND3 Gold (M1 8-Core, GPU 7-Core, 8GB, 256Gb) which had been introduced in begiening which has the lowest sale price than the other macbook version.**

**sns.barplot(data=df,y="2022-06-08",x="name")**

**<Axes: xlabel='name', ylabel='2022-06-08'>**

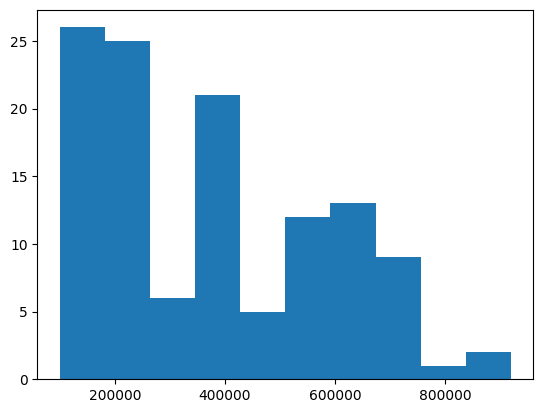
****

**Observation:**

**From the above bar plot graph the value of the mac book sold how higher and which mac book is sold higher and lower in the date:2022-06-08 is displayed in the bar graph. The Apple MacBook Pro 16 Z14V00093 Space Gray (M1 Max 10-Core, GPU 32-Core, 64GB, 8TB), this macbook which has the highest sale price than any other macbook version in 2022-06-08. The Apple MacBook Air 13 Retina MGND3 Gold (M1 8-Core, GPU 7-Core, 8GB, 256Gb) which had been introduced in begiening which has the lowest sale price than the other macbook version.** **Apple MacBook Pro 13 Retina Touch Bar Z11B0004Q Space Gray (M1 8-Core 8GB, 2TB) and Apple MacBook Air 13 Retina Z1250007H Space Gray (M1 8-Core, GPU 8-Core, 8GB, 1Tb) which has been sold equally in this day and the amount is rupees:239990.**

**plt.hist(df["2022-06-07"])**

**(array([26., 25., 6., 21., 5., 12., 13., 9., 1., 2.]),  
 array([ 99990., 181990., 263990., 345990., 427990., 509990., 591990.,  
 673990., 755990., 837990., 919990.]),  
 <BarContainer object of 10 artists>)**

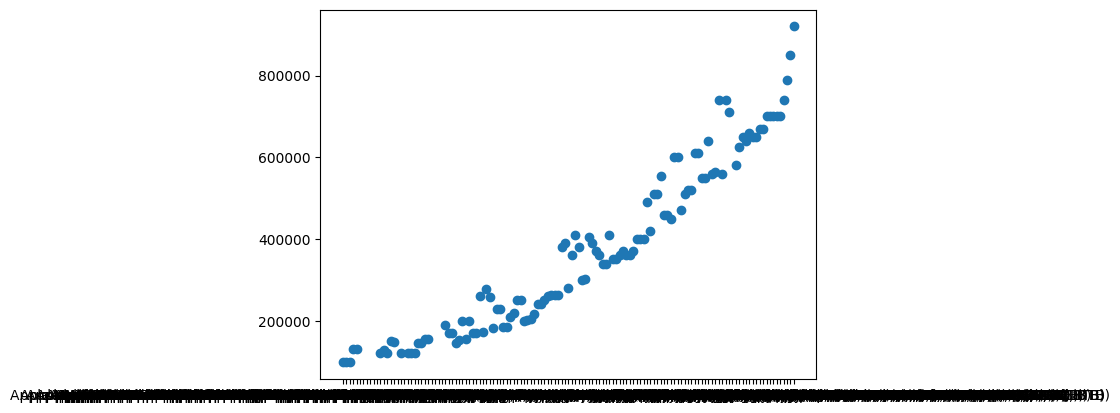
****

**Observation:**

**From the above plot the mac book sold in that date 2022-06-07. The maximum number of macbook that has been sold is 206539 with this sale price and the minimum number of macbook that has been sold is 551240.0 .**

**plt.scatter(df["name"],df["2022-06-09"])**

**<matplotlib.collections.PathCollection at 0x795952b85bd0>**

****

**Observations :**

**From the above graph we can the data that has visulised that has been taken from the dataset. From the above scatter plot we defined the name of the macbook and who they sold in date:2022-06-09 . The Apple MacBook Pro 16 Z14V00093 Space Gray (M1 Max 10-Core, GPU 32-Core, 64GB, 8TB), this macbook which has the highest sale price than any other macbook version in 2022-06-09. The Apple MacBook Air 13 Retina MGND3 Gold (M1 8-Core, GPU 7-Core, 8GB, 256Gb) which had been introduced in begiening which has the lowest sale price than the other macbook version.**

mini = macbooks.iloc[:7].reset\_index(drop=True)

mini.drop(columns='name').T.plot(figsize=(12,12))

list(mini['name'])

['Apple MacBook Air 13 Retina MGN93 Silver (M1 8-Core, GPU 7-Core, 8GB, 256Gb)',

'Apple MacBook Air 13 Retina MGND3 Gold (M1 8-Core, GPU 7-Core, 8GB, 256Gb)',

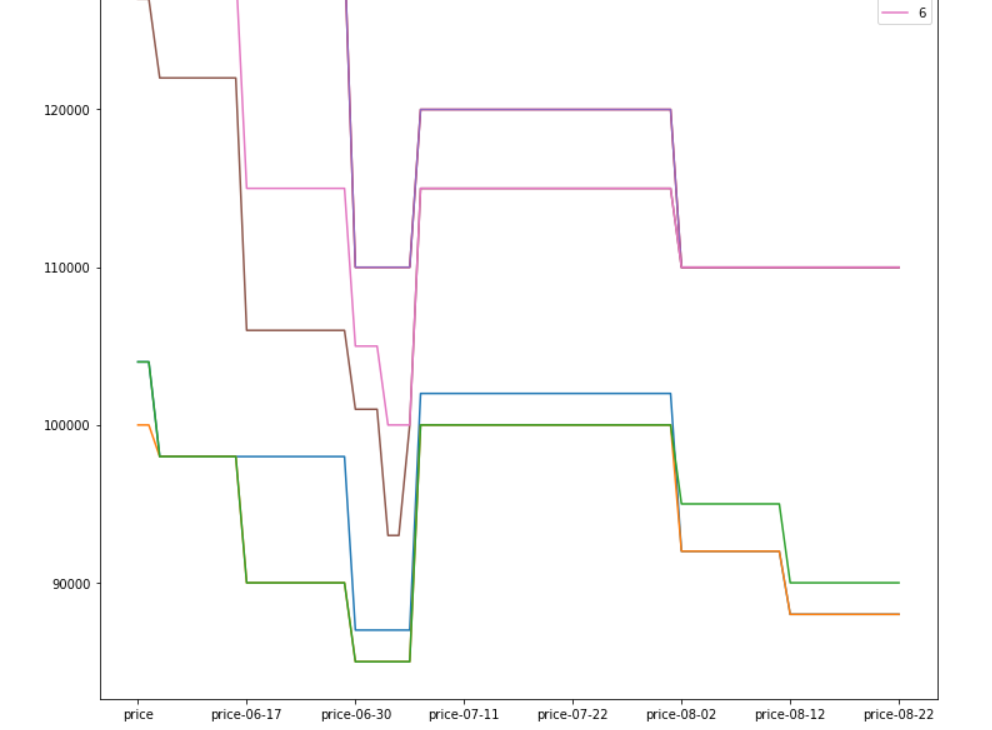
'Apple MacBook Air 13 Retina MGN63 Space Gray (M1 8-Core, GPU 7-Core, 8GB, 256Gb)',

'Apple MacBook Pro 13 Retina Touch Bar MYD82 Space Gray (M1 8-Core, 8GB, 256Gb)',

'Apple MacBook Pro 13 Retina Touch Bar MYDA2 Silver (M1 8-Core, 8GB, 256Gb)',

'Apple MacBook Air 13 Retina MGN73 Space Gray (M1 8-Core, GPU 8-Core, 8GB, 512Gb)',

'Apple MacBook Air 13 Retina MGNA3 Silver (M1 8-Core, GPU 8-Core, 8GB, 512Gb)']

****

**Observation:**

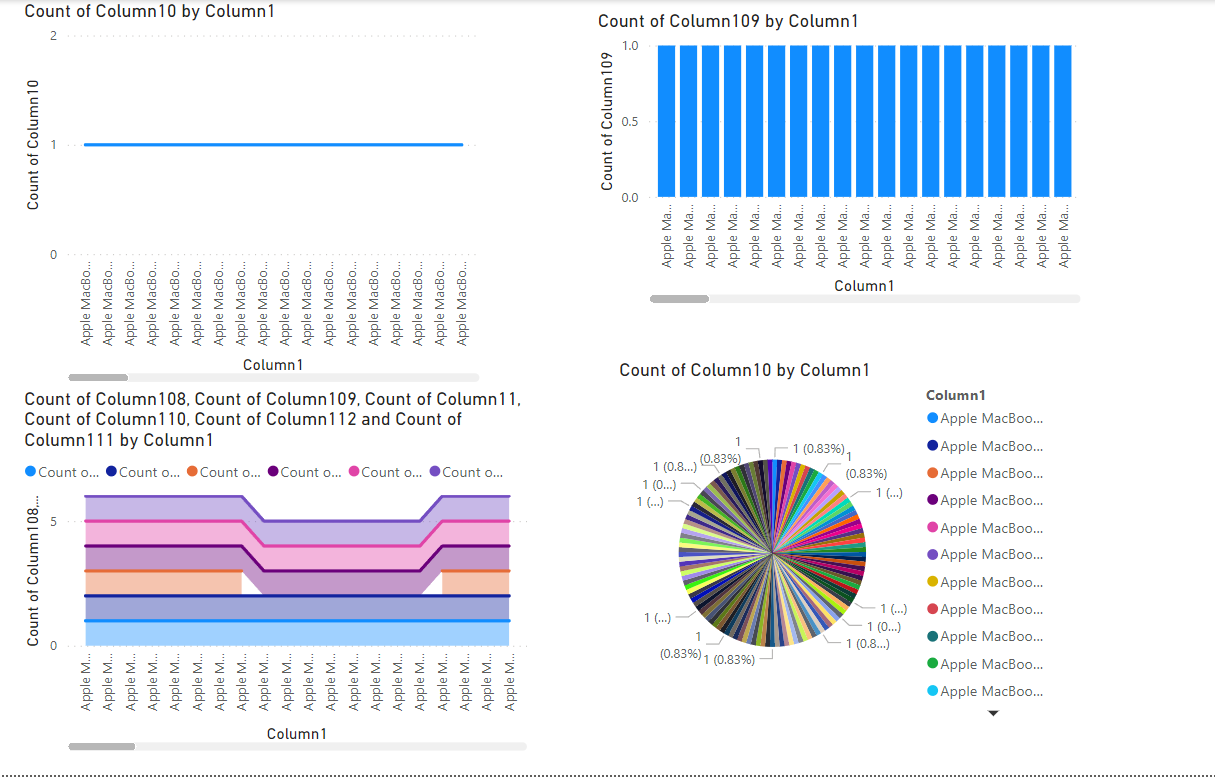
**The above graph shows that the different kind of macbook and their sale price in the different days of the week. The line colour which in pink colour**

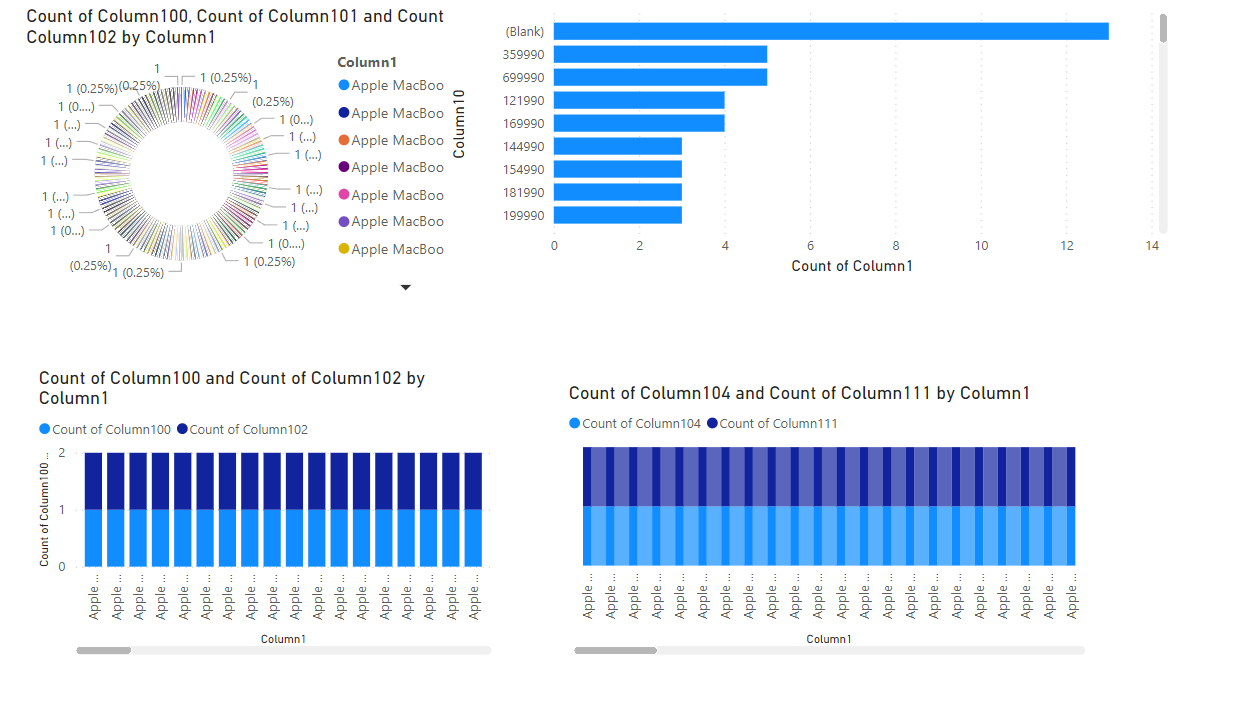
**maximum sale price in the macbook. The line colour which in orange colour minimum sale price of the macbook.**

Result:

Thus Working with Numpy arrays , Pandas Data Frames , basic plots using Matplotlib be executed and successfully completed.

**Power BI:**

****

****