Experiment – 2

Student Name: Vivek Kumar **UID: 21BCS8129**
Branch: BE-CSE(LEET) **Section/Group: WM-20BCS-616/A**
Semester: 5th **Date of Performance: 16/08/2022**
Subject Name: Machine Learning Lab **Subject Code: 20CSP-317**

1. Aim/Overview of the practical:

Implement Data Visualization.

2. Task to be done/ Which logistics used:

Data Visualization using matplotlib, seaborn, plotly

3. Algorithm/Flowchart (For programming-based labs):

4. Steps for experiment/practical/Code:

```
from google.colab import drive
drive.mount('/content/drive')

import pandas as pd

data = pd.read_csv("/content/drive/MyDrive/Data/Students_data.csv")

data.head(10)

data.tail()

import matplotlib.pyplot as plt

plt.scatter(data['race'],data['GPA'])
plt.title('Scatter Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.show()

plt.scatter(data['race'],data['GPA'],c=data['Probability'],s=data['Statistics'])
plt.title('Scatter Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.colorbar()
plt.show()
```

```
plt.bar(data['race'],data['GPA'])
plt.title('Bar Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.show()

plt.hist(data['race'])
plt.title('Histogram Plot')
plt.show()

import seaborn as sb

sb.scatterplot(x='race',y='GPA',data=data)

sb.scatterplot(x='race',y='GPA',data=data,hue='gender')

sb.lineplot(x='race',y='GPA',data=data)

sb.lineplot(x='race',y='GPA',data=data,hue='gender')

sb.barplot(x='race',y='GPA',data=data,hue='gender')

sb.histplot(x='GPA', data=data, kde=True, hue='gender')

import plotly.express as px

#plotting the scatter chart
fig = px.scatter(data, x="GPA", y='Algebra', color="gender")
#showing the plot
fig.show();

#plotting the line chart
fig = px.line(data, y='Algebra', color="gender")
#showing the plot
fig.show();

#plotting the line chart
fig = px.line(data, x="Algebra", y='GPA', color="race")
#showing the plot
fig.show();

#plotting the bar chart
fig = px.bar(data, x="Algebra", y='GPA', color="race")
#showing the plot
```

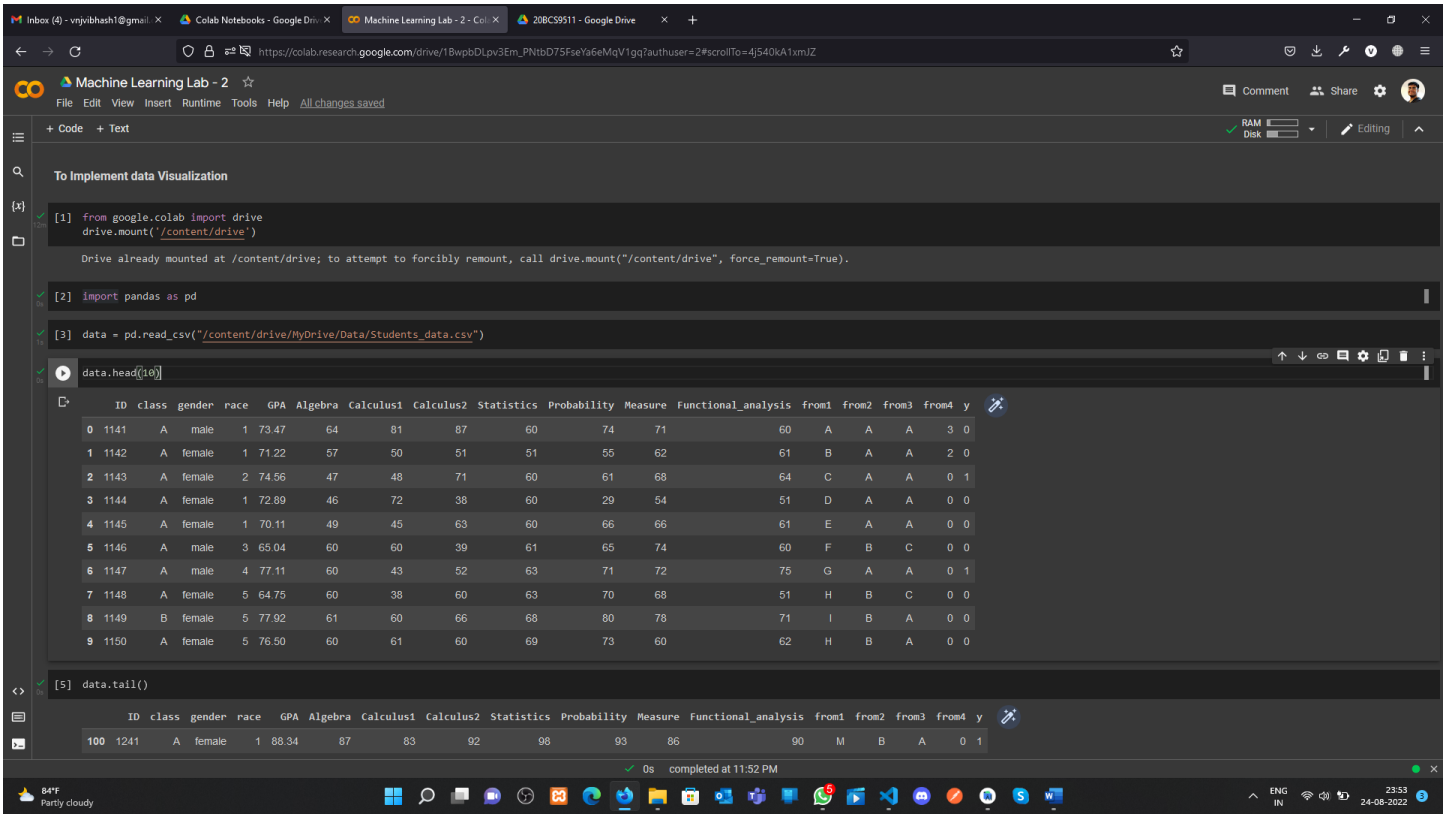
```
fig.show();

#plotting the histogram chart
fig = px.histogram(data, x="Algebra", y='GPA', color='GPA')
#showing the plot
fig.show();
```

5. Observations/Discussions/ Complexity Analysis:

In this have done Data visualization with matplotlib and used various function such as scatter, scatter with colorbar, bar with x-y label and hist. Then seaborn and plotted various graph such as scatterplot, lineplot, barplot and histplot. Another library which I have used plotly and plotted scatter line, bar and histogram.

6. Result/Output/Writing Summary:



	ID	class	gender	race	GPA	Algebra	Calculus1	Calculus2	Statistics	Probability	Measure	Functional_analysis	from1	from2	from3	from4	y
0	1141	A	male	1	73.47	64	81	87	60	74	71	60	A	A	A	3	0
1	1142	A	female	1	71.22	57	50	51	51	55	62	61	B	A	A	2	0
2	1143	A	female	2	74.56	47	48	71	60	61	68	64	C	A	A	0	1
3	1144	A	female	1	72.89	46	72	38	60	29	54	51	D	A	A	0	0
4	1145	A	female	1	70.11	49	45	63	60	66	66	61	E	A	A	0	0
5	1146	A	male	3	65.04	60	60	39	61	65	74	60	F	B	C	0	0
6	1147	A	male	4	77.11	60	43	52	63	71	72	75	G	A	A	0	1
7	1148	A	female	5	64.75	60	38	60	63	70	68	51	H	B	C	0	0
8	1149	B	female	5	77.92	61	60	66	68	80	78	71	I	B	A	0	0
9	1150	A	female	5	76.50	60	61	60	69	73	60	62	H	B	A	0	0

	ID	class	gender	race	GPA	Algebra	Calculus1	Calculus2	Statistics	Probability	Measure	Functional_analysis	from1	from2	from3	from4	y
100	1241	A	female	1	88.34	87	83	92	98	93	86	90	M	B	A	0	1

Machine Learning Lab - 2

[5] data.tail()

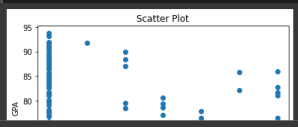
	ID	class	gender	race	GPA	Algebra	Calculus1	Calculus2	Statistics	Probability	Measure	Functional_analysis	from1	from2	from3	from4	y
100	1241	A	female	1	88.34	87	83	92	98	93	86	90	M	B	A	0	1
101	1242	B	male	1	89.84	98	77	95	98	96	88	100	A	B	A	0	1
102	1243	B	male	1	88.82	83	80	91	98	93	95	71	T	B	A	0	2
103	1244	A	male	1	86.60	92	82	91	99	94	82	78	S	B	A	0	2
104	1245	A	male	1	93.71	93	97	99	100	97	90	90	K	B	A	0	2

[6] pip install matplotlib

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: matplotlib in /usr/local/lib/python3.7/dist-packages (3.2.2)
Requirement already satisfied: python-dateutil>=2.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (2.8.2)
Requirement already satisfied: numpy>=1.11 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (1.21.6)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (3.0.9)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (1.4.4)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (0.11.0)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-packages (from kiwisolver>=1.0.1->matplotlib) (4.1.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from python-dateutil>=2.1->matplotlib) (1.15.0)
```

[7] import matplotlib.pyplot as plt

```
plt.scatter(data['race'],data['GPA'])
plt.title('Scatter Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.show()
```



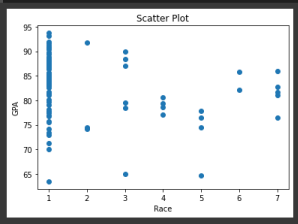
0s completed at 11:52 PM

84°F Partly cloudy

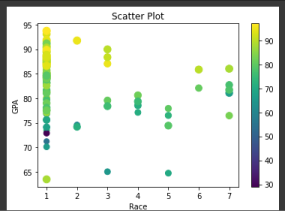
ENG IN 23:54 24-08-2022

Machine Learning Lab - 2

```
plt.scatter(data['race'],data['GPA'])
plt.title('Scatter Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.show()
```




```
plt.scatter(data['race'],data['GPA'],c=data['Probability'],s=data['Statistics'])
plt.title('Scatter Plot')
plt.xlabel('Race')
plt.ylabel('GPA')
plt.colorbar()
plt.show()
```



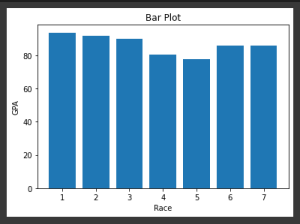
0s completed at 11:52 PM

84°F Partly cloudy

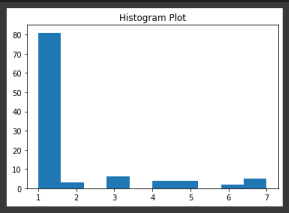
ENG IN 23:55 24-08-2022

Machine Learning Lab - 2

```
[10] plt.bar(data['race'],data['GPA'])
     plt.title('Bar Plot')
     plt.xlabel('Race')
     plt.ylabel('GPA')
     plt.show()
```



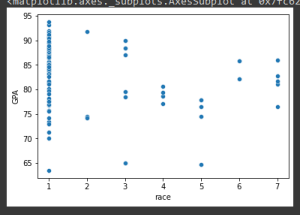

```
[11] plt.hist(data['race'])
     plt.title('Histogram Plot')
     plt.show()
```




```
[12] import seaborn as sb
```

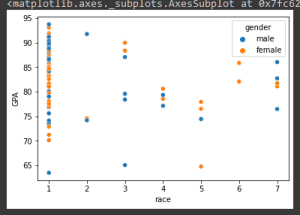
```
sb.scatterplot(x='race',y='GPA',data=data)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6273354d0>
```




```
sb.scatterplot(x='race',y='GPA',data=data,hue='gender')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6272bd410>
```



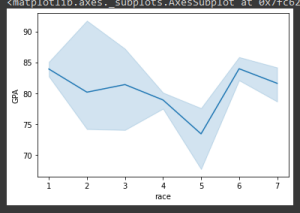

```
[15] sb.lineplot(x='race',y='GPA',data=data)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6272e9850>
```

Machine Learning Lab - 2

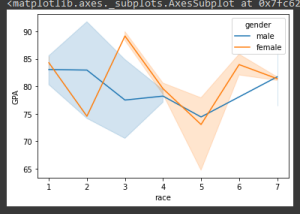
```
[15] sb.lineplot(x='race',y='GPA',data=data)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6272e9850>
```



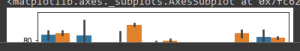

```
[16] sb.lineplot(x='race',y='GPA',data=data,hue='gender')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc62732581o>
```




```
sb.barplot(x='race',y='GPA',data=data,hue='gender')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6270f48d0>
```



0s completed at 11:52 PM

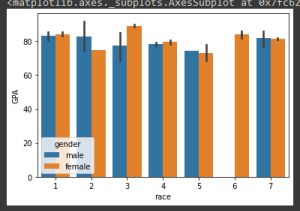
84°F Partly cloudy

ENG IN 23:56 24-08-2022

Machine Learning Lab - 2

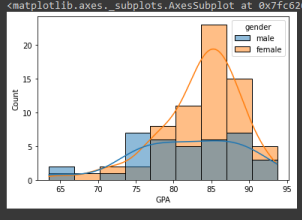
```
sb.barplot(x='race',y='GPA',data=data,hue='gender')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6270f48d0>
```




```
sb.histplot(x='GPA', data=data, kde=True, hue='gender')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc626fc9cd0>
```




```
[19] import plotly.express as px
```

```
[20] #plotting the scatter chart
     fig = px.scatter(data, x="GPA", y='Algebra', color="gender")
     #showing the plot
```

0s completed at 11:52 PM

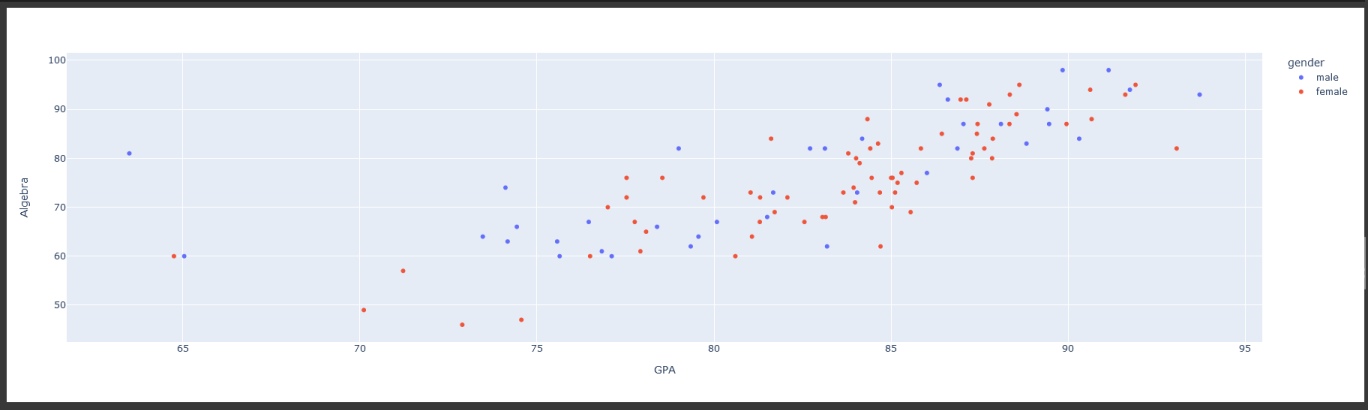
84°F Partly cloudy

ENG IN 23:56 24-08-2022

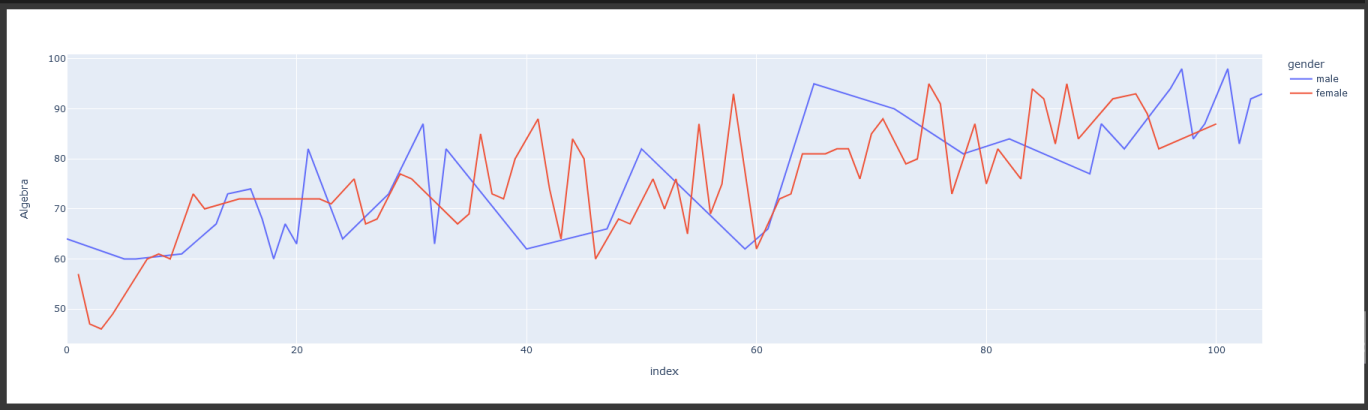
Machine Learning Lab - 2

```
[19] import plotly.express as px

[20] #plotting the scatter chart
     fig = px.scatter(data, x="GPA", y='Algebra', color="gender")
     #showing the plot
     fig.show();
```



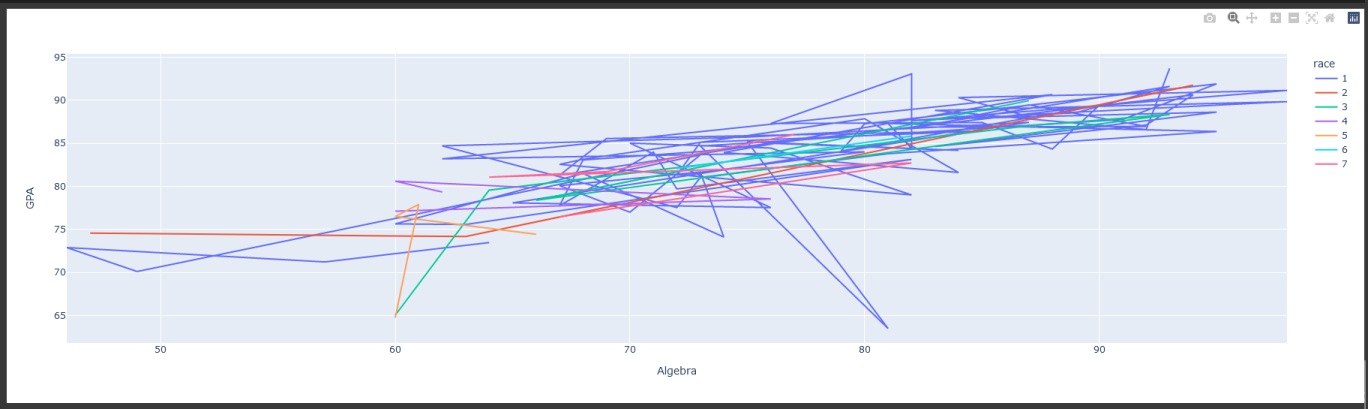

```
[21] #plotting the line chart
     fig = px.line(data, y='Algebra', color="gender")
     #showing the plot
     fig.show();
```




```
[22] #plotting the line chart
     fig = px.line(data, x="Algebra", y='GPA', color="race")
     #showing the plot
     fig.show();
```

Machine Learning Lab - 2

```
[22] #plotting the line chart
     fig = px.line(data, x="Algebra", y='GPA', color="race")
     #showing the plot
     fig.show();
```

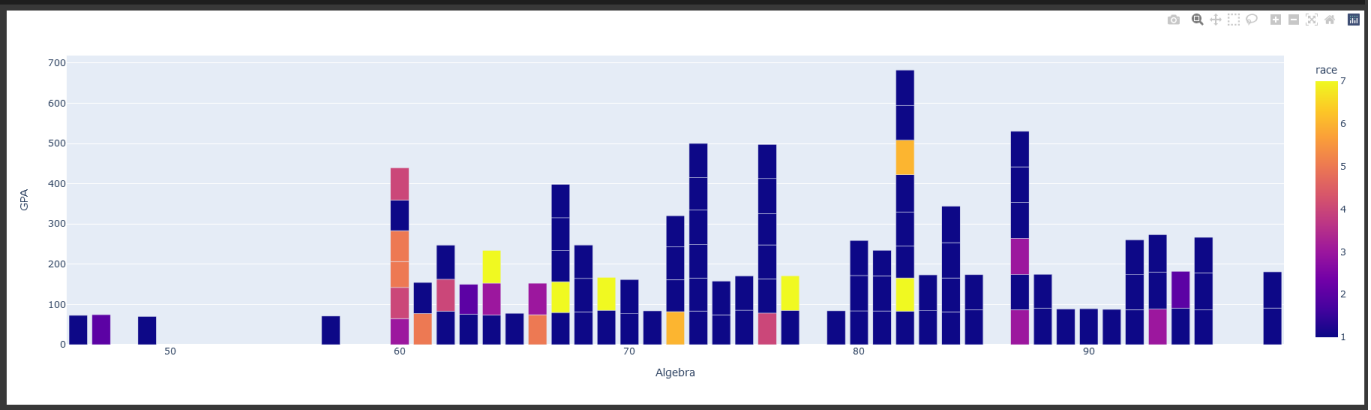



```
#plotting the bar chart
fig = px.bar(data, x="Algebra", y='GPA', color="race")
#showing the plot
fig.show();
```

0s completed at 11:52 PM

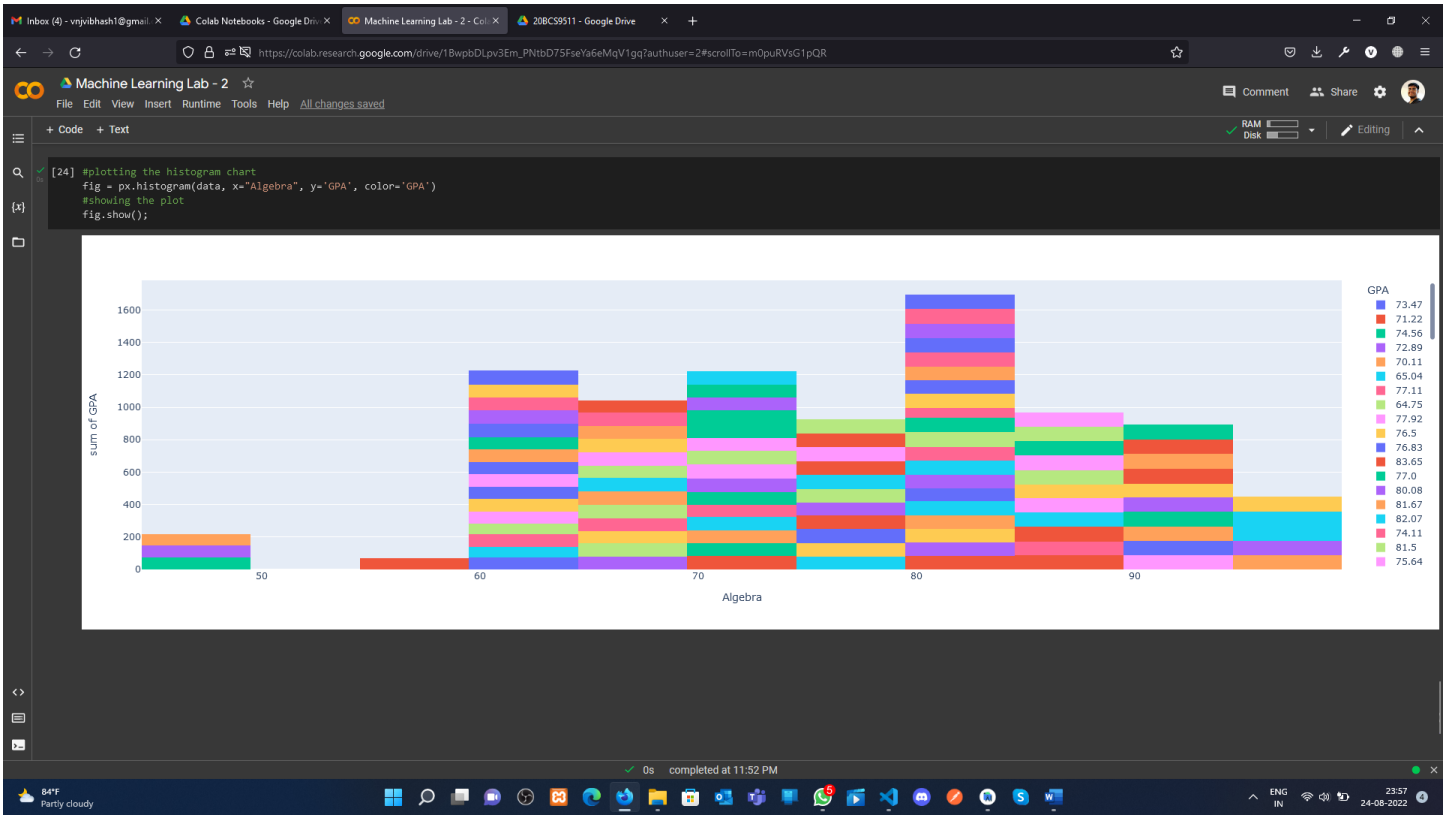
Machine Learning Lab - 2

```
#plotting the bar chart
fig = px.bar(data, x="Algebra", y='GPA', color="race")
#showing the plot
fig.show();
```

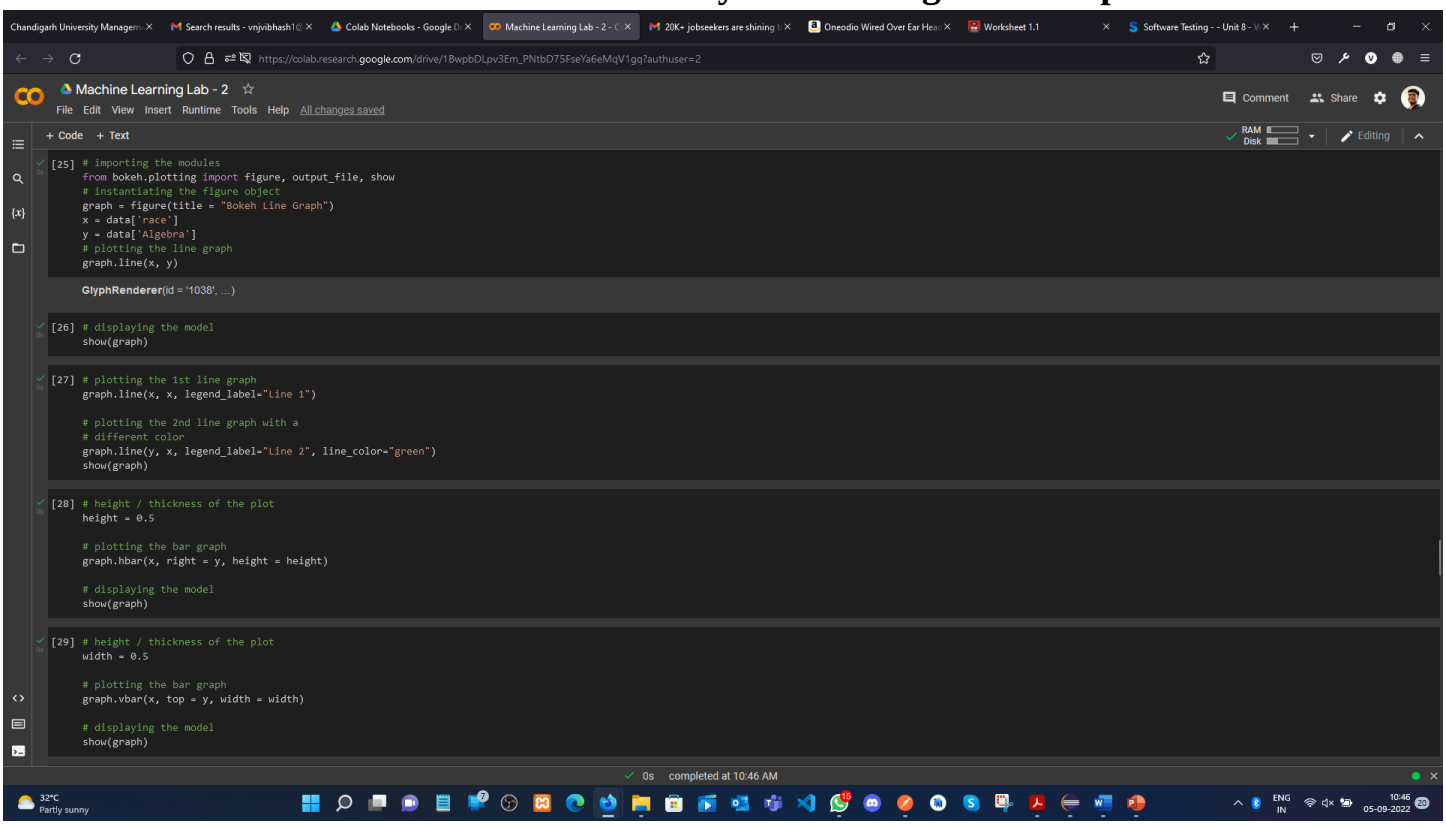


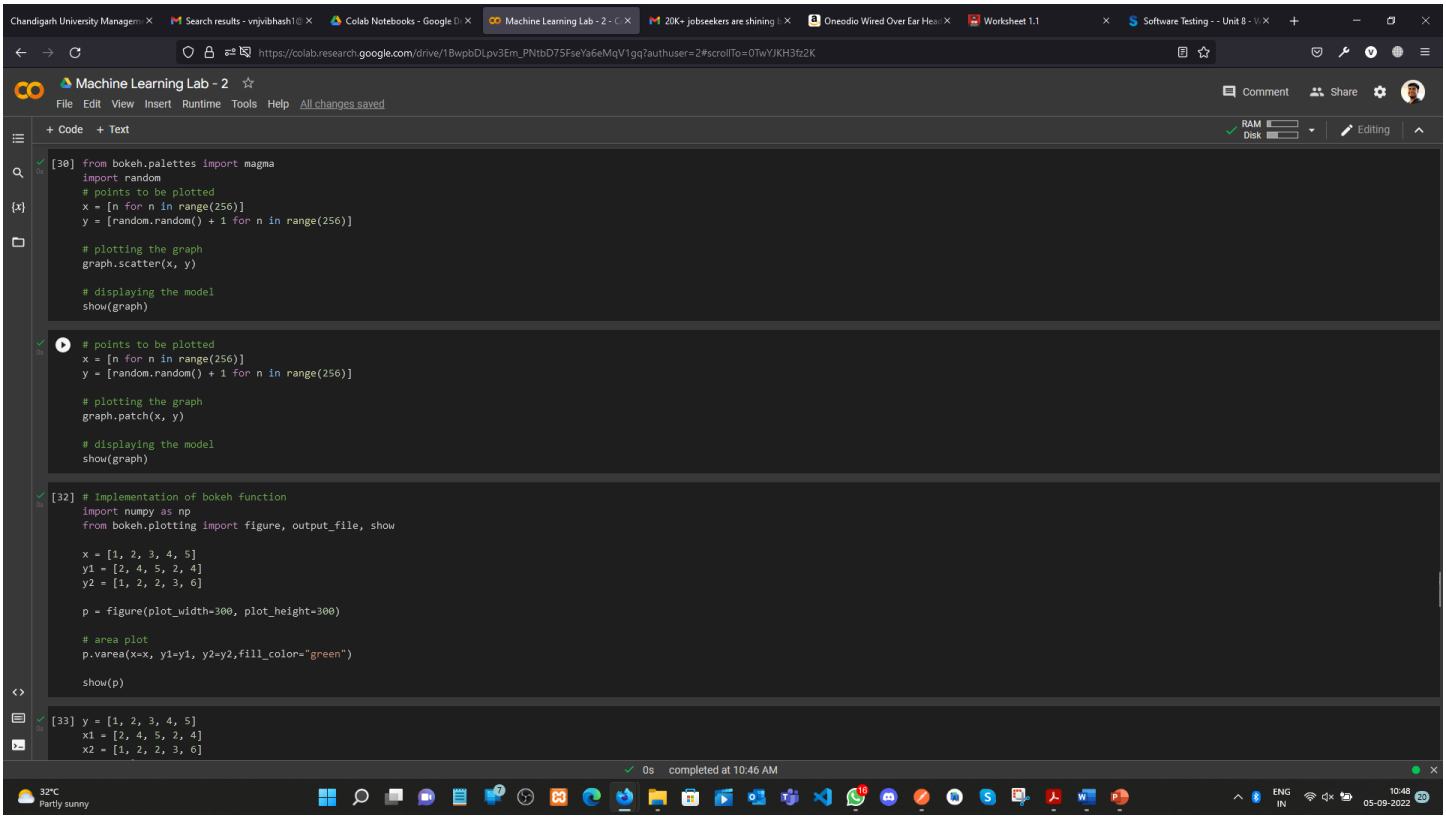

```
[24] #plotting the histogram chart
     fig = px.histogram(data, x="Algebra", y='GPA', color='GPA')
     #showing the plot
     fig.show();
```

0s completed at 11:52 PM



Executed some of the Code for Bokeh Library but Didn't get the Output.





Learning outcomes (What I have learnt):

1. Data Visualization using matplotlib
2. Data visualization using seaborn
3. Data Visualization using plotlib

Evaluation Grid (To be created as per the SOP and Assessment guidelines by the faculty):

Sr. No.	Parameters	Marks Obtained	Maximum Marks
1.			
2.			
3.			