

DATA SCIENCE

SP19-BSE-003

SP19-BSE-007

SP19-BSE-013

Dr. Hikmat Ullah

Wajahat Masood

Abdullah Rauf

Saif Ur Rehman

Course Code : CSC 461



TOPICS WE CHOOSE



We Gather data of Stock Exchange And
than we visualize the data using python.

STOCK EXCHANGE

This Document Includes

Data Set

Origin of Data

Language: Python

IDE : PYCHARM

Libraries

SCREEN SHOTS

Stock Exchange of Pakistan (2019-2022)

Dataset Date

Attributes/ Records

Format : CSV

Why We Choose This Data

Stock Exchange

This dataset is taken from the website “World Stock Exchange”. They provide trusted source of information and tools that inspire and empower us understand and share the Stock Exchange Information.

And the purpose of this data set is to get the knowledge of World Stock market and The current trends that is being followed and few other things like stock rates, Current prices, And Unit sales etc.



2019-2022



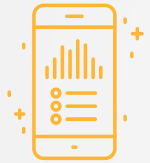
8 Attributes



40 Record

About Data Set And Visualization

Stock Exchange



Attributes

Our Data set Has
8 Attributes



Data Records

Our Data set Has
40 Records



File Type

CSV



Libraries

Panda, MAT-PLOT, NUMPY,
SEABORN

How we Get The Data (Acknowledgment)

Stock Exchange

We have used EDA to analyze the data of the of the Stock Exchange. We download the raw form data which has approximately 8000+ record and most of the records are missing and many duplication. To improve our data first we reduce a lot of data than we watch a lot of YouTube (Code with HARRY), (Garry Smasher) videos for data cleaning and buy applying those techniques we clean a little bit of data and after that we manually clean the data and after cleaning all the data, we left with approximately 40 records and 8 attributes. After all of this we use python to visualize the data in the form of charts.

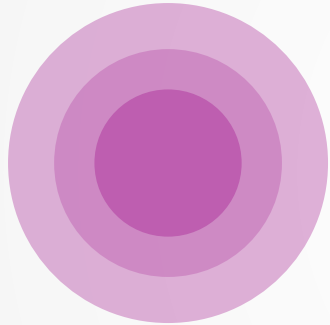
Attributes And their Types

Stock Market



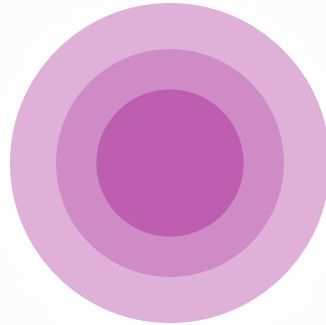
Invoice Number

Numerical {Interval}



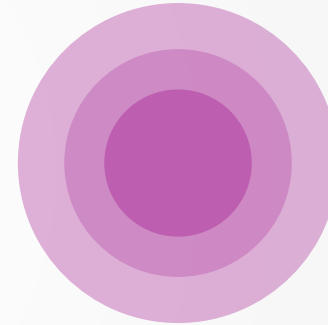
Stock Code

Numerical {Interval}



Description

Textual data



Quantity

Numerical



Date

Categorical {Nominal}

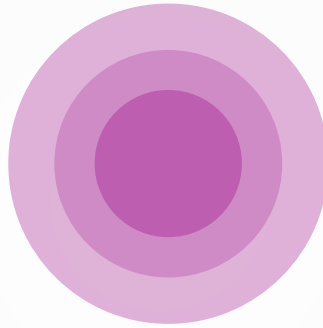
Attributes And their Types

Stock Market



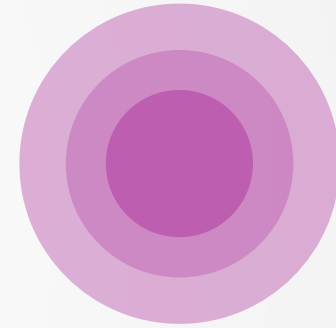
Unit Price

Numerical {Ratio}



Customer Id

Numerical



Country

Categorical {Nominal}

Screen Shot (Data)

Stock Exchange

AutoSave Off | data | Search (Alt+Q) | ABDULLAH BIN ABDUR RAUF

File Home Insert Page Layout Formulas Data Review View Help

Clipboard Font Alignment Number Styles Cells Editing Analysis

H61

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	Customer	Country								
2	536365	85123	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2022 8:26	2.55	17850	PAKISTAN								
3	536370	22728	ALARM CLOCK BAKELIKE PINK	24	12/1/2022 8:45	3.75	12583	France								
4	536389	35004	SET OF 3 GOLD FLYING DUCKS	4	12/1/2022 10:03	6.35	12431	Australia								
5	536527	22243	5 HOOK HANGER RED	12	12/1/2022 13:04	1.65	12662	Germany								
6	536532	22556	PLASTERS IN TIN	24	12/1/2022 13:24	1.65	12433	Norway								
7	536540	85136	YELLOW SHARK HELICOPTER	2	12/1/2022 14:05	7.95	14911	KSA								
8	536861	22634	CHILDS BREAKFAST SET SPACEBOY	8	12/2/2022 10:44	8.5	12427	Germany								
9	536863	22384	LUNCH BAG PINK POLKADOT	1	12/2/2022 11:19	1.65	17967	PAKISTAN								
10	536863	22383	LUNCH BAG SUKI DESIGN	1	12/2/2022 11:19	1.65	17967	PAKISTAN								
11	536863	22147	FELTCRAFT BUTTERFLY HEARTS	3	12/2/2022 11:19	1.45	17967	United Kingdom								
12	536863	20961	STRAWBERRY BATH SPONGE	2	12/2/2022 11:19	1.25	17967	United Kingdom								
13	536863	20963	APPLE BATH SPONGE	1	12/3/2021 11:19	1.25	17967	United Kingdom								
14	536863	22750	FELTCRAFT PRINCESS LOLA DOLL	1	12/3/2021 11:19	3.75	17967	United Kingdom								
15	536890	17084	ASSORTED INCENSE PACK	1440	12/3/2021 11:48	0.16	14156	KSA								
16	536938	21931	JUMBO STORAGE BAG SUKI	20	12/3/2021 12:05	1.95	14680	USA								
17	536938	20712	JUMBO BAG WOODLAND ANIMALS	20	12/3/2021 12:05	1.95	14680	USA								
18	536938	20713	JUMBO BAG OWLS	20	12/3/2021 12:05	1.95	14680									
19	536938	20724	RED RETROSPOT CHARLOTTE BAG	20	12/3/2021 12:05	0.85	14680	PAKISTAN								
20	536938	22356	CHARLOTTE BAG PINK POLKADOT	20	12/3/2021 12:05	0.85	14680	PAKISTAN								
21	543113	22766	PHOTO FRAME CORNICE	8	2/3/2021 13:06	2.95	13047	United Kingdom								
22	543113	22708	WRAP DOLLY GIRL	25	2/3/2021 13:06	0.42	13047	United Kingdom								
23	543113	22721	WRAP DOLLY DESIGN	25	2/3/2021 13:06	0.42	13047	United Kingdom								

Select destination and press ENTER or choose Paste

Type here to search | 21°C | 2:23 PM 3/2/2022

Records

Attributes

Screen Shot (Data)

Stock Exchange

AutoSave Off | data | Search (Alt+Q) | ABDULLAH BIN ABDUR RAUF

File Home Insert Page Layout Formulas Data Review View Help

Clipboard | Font | Alignment | Number | Styles | Cells | Editing | Analysis

H61

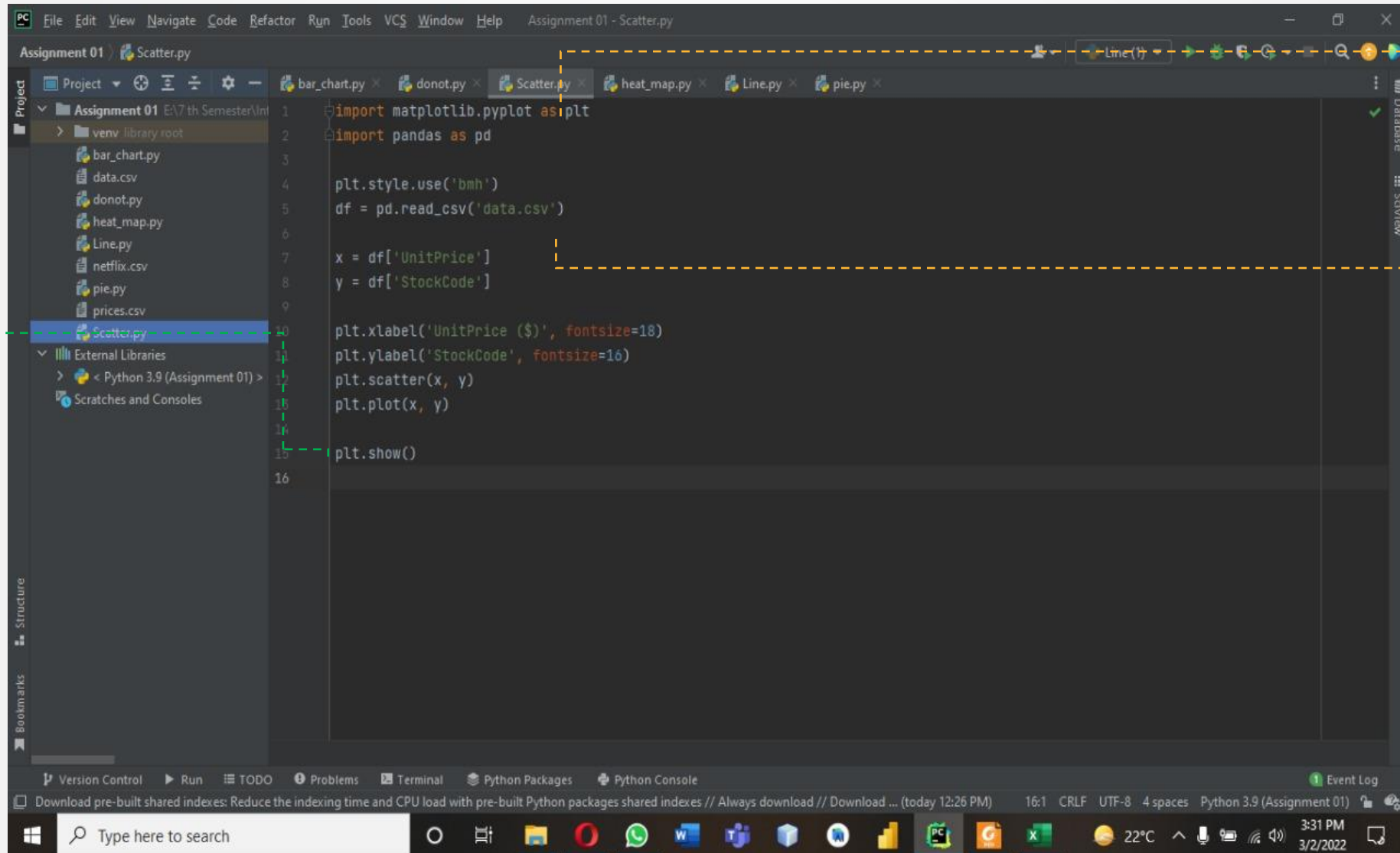
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
19	536938	20724	RED RETROSPOT CHARLOTTE BAG	20	12/3/2021 12:05	0.85	14680	PAKISTAN								
20	536938	22356	CHARLOTTE BAG PINK POLKADOT	20	12/3/2021 12:05	0.85	14680	PAKISTAN								
21	543113	22766	PHOTO FRAME CORNICE	8	2/3/2021 13:06	2.95	13047	United Kingdom								
22	543113	22708	WRAP DOLLY GIRL	25	2/3/2021 13:06	0.42	13047	United Kingdom								
23	543113	23231	WRAP DOILEY DESIGN	25	2/3/2021 13:06	0.42	13047	United Kingdom								
24	543113	85015	SET OF 12 VINTAGE POSTCARD SET	6	2/3/2021 13:06	2.55	13047	United Kingdom								
25	543113	85016	SET OF 6 VINTAGE NOTELETS KIT	6	2/3/2021 13:06	2.55	13047	United Kingdom								
26	543113	22829	SWEETHEART WIRE WALL TIDY	2	2/3/2021 13:06	9.95	13047	United Kingdom								
27	543113	22722	SET OF 6 SPICE TINS PANTRY DESIGN	4	2/3/2021 13:06	3.95	13047	United Kingdom								
28	543113	22960	JAM MAKING SET WITH JARS	6	2/3/2021 13:06	4.25	13047	United Kingdom								
29	543113	22961	JAM MAKING SET PRINTED	12	2/3/2020 13:06	1.45	13047	United Kingdom								
30	543113	22969	HOMEMADE JAM SCENTED CANDLES	12	2/3/2020 13:06	1.45	13047	United Kingdom								
31	543113	22720	SET OF 3 CAKE TINS PANTRY DESIGN	3	2/3/2020 13:06	4.95	13047	United Kingdom								
32	543113	22549	PICTURE DOMINOES	12	2/3/2020 13:06	1.45	13047	United Kingdom								
33	543113	20972	PINK CREAM FELT CRAFT TRINKET BOX	12	2/3/2020 13:06	1.25	13047	United Kingdom								
34	543113	22743	MAKE YOUR OWN FLOWERPOWER CARD KIT	6	2/3/2020 13:06	2.95	13047	United Kingdom								
35	543113	84969	BOX OF 6 ASSORTED COLOUR TEASPOONS	12	2/3/2020 13:06	4.25	13047	PAKISTAN								
36	543113	22622	BOX OF VINTAGE ALPHABET BLOCKS	2	2/3/2019 13:06	9.95	13047	PAKISTAN								
37	543113	21658	GLASS BEURRE DISH	4	2/3/2019 13:06	3.95	13047	United Kingdom								
38	543113	21756	BATH BUILDING BLOCK WORD	3	2/3/2019 13:06	5.95	13047	United Kingdom								
39	543113	21754	HOME BUILDING BLOCK WORD	3	2/3/2019 13:06	5.95	13047	United Kingdom								
40	543113	21755	LOVE BUILDING BLOCK WORD	3	2/3/2019 13:06	5.95	13047	United Kingdom								

Ready | data | 100%

Type here to search | 21°C | 2:23 PM 3/2/2022

Screen Shot (Code Python)

Scatter Graph CODE



The screenshot shows a Python IDE with a project named 'Assignment 01'. The file explorer on the left shows a directory structure with files like 'bar_chart.py', 'data.csv', 'donot.py', 'heat_map.py', 'Line.py', 'netflix.csv', 'pie.py', and 'prices.csv'. The 'Scatter.py' file is selected. The code editor displays the following Python code:

```
1 import matplotlib.pyplot as plt
2 import pandas as pd
3
4 plt.style.use('bmh')
5 df = pd.read_csv('data.csv')
6
7 x = df['UnitPrice']
8 y = df['StockCode']
9
10 plt.xlabel('UnitPrice ($)', fontsize=18)
11 plt.ylabel('StockCode', fontsize=16)
12 plt.scatter(x, y)
13 plt.plot(x, y)
14
15 plt.show()
16
```

The status bar at the bottom indicates the file encoding is CRLF, the line length is 16:1, and the Python version is 3.9 (Assignment 01).

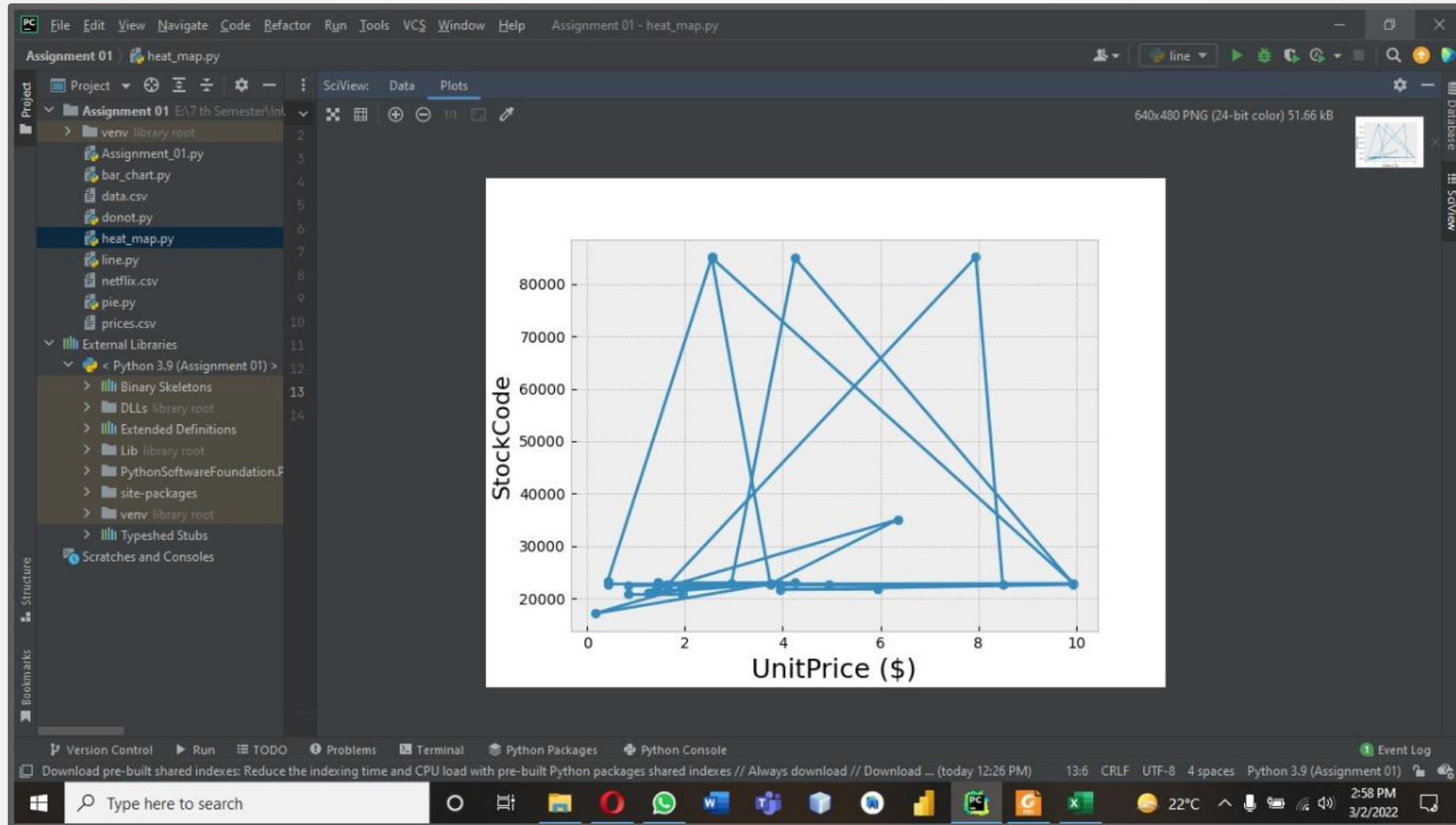
Libraries

This will read
the CSV file

This Will Plot
the Scatter
Graph

Screen Shot (Output Python)

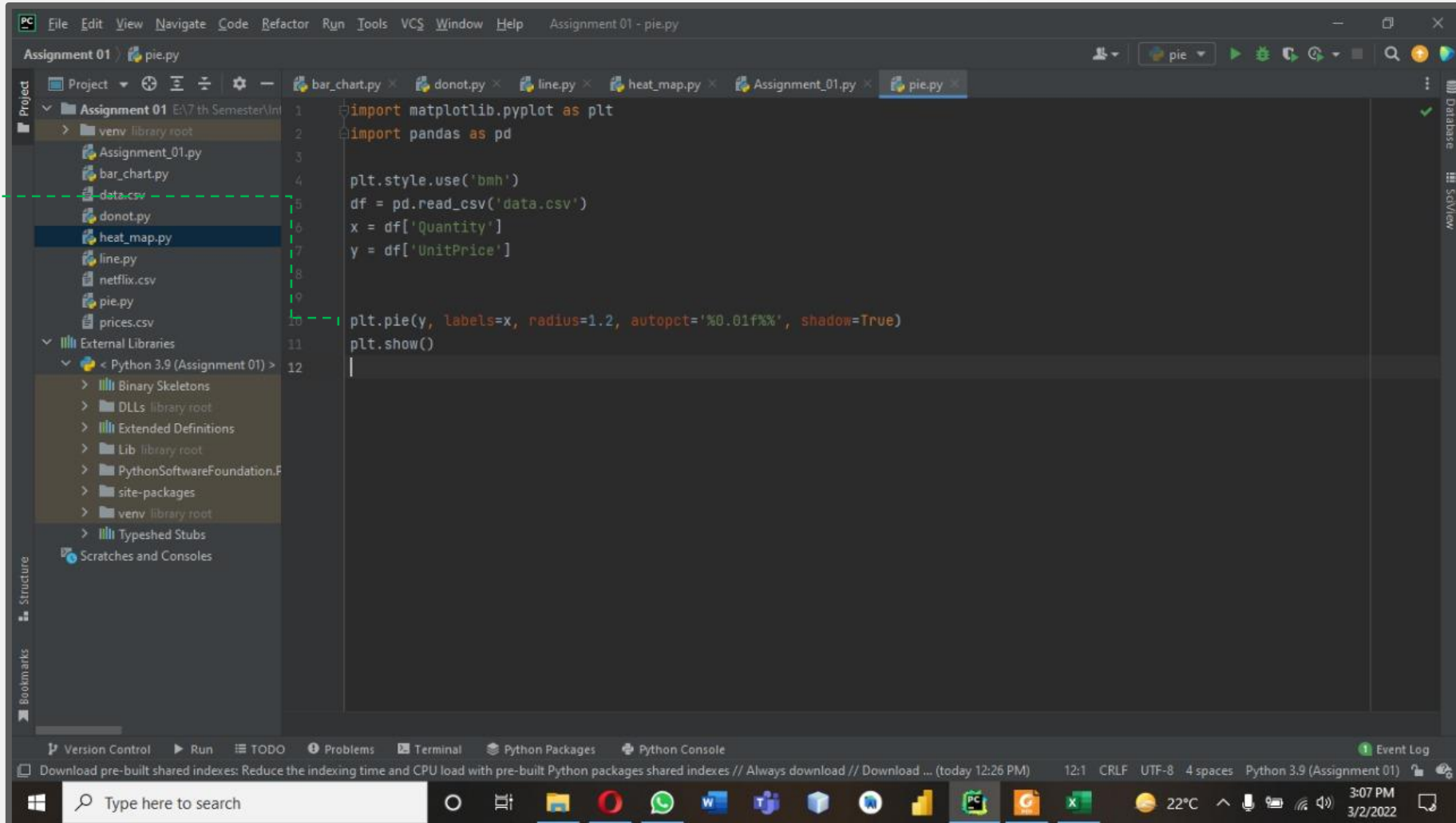
Scatter CHART OUTPUT



Screen Shot (Code Python)

PI CHART CODE

Pi Chart Code
line

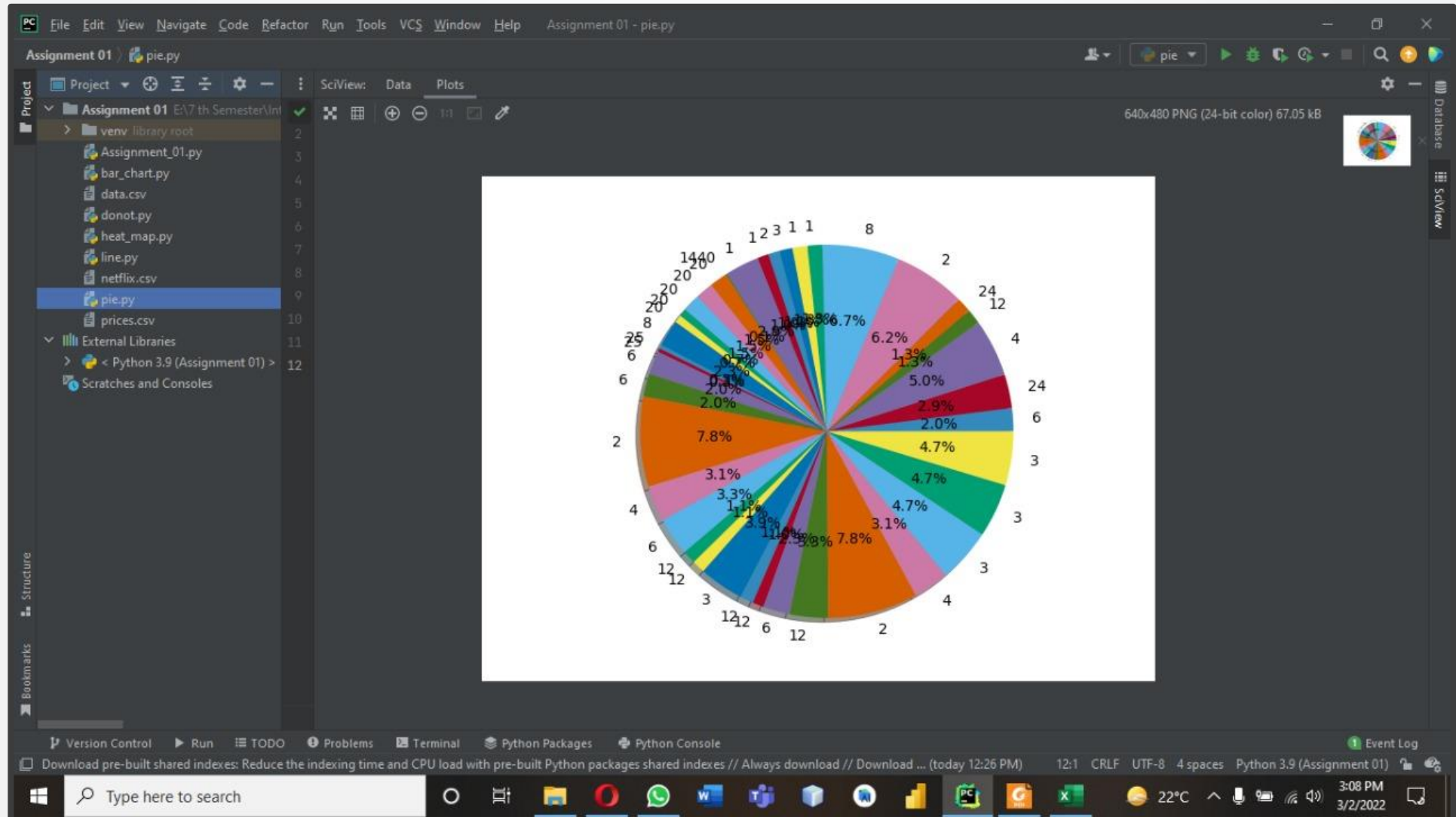


```
1 import matplotlib.pyplot as plt
2 import pandas as pd
3
4 plt.style.use('bmh')
5 df = pd.read_csv('data.csv')
6 x = df['Quantity']
7 y = df['UnitPrice']
8
9
10 plt.pie(y, labels=x, radius=1.2, autopct='%0.01f%%', shadow=True)
11 plt.show()
12
```

The screenshot shows an IDE window titled "Assignment 01 - pie.py". The left sidebar displays a project structure for "Assignment 01" with files like "bar_chart.py", "donot.py", "line.py", "heat_map.py", "pie.py", "data.csv", "netflix.csv", and "prices.csv". The "pie.py" file is selected. The main editor shows the Python code for generating a pie chart. A green circle and dashed line highlight the line `plt.pie(y, labels=x, radius=1.2, autopct='%0.01f%%', shadow=True)` in the code editor.

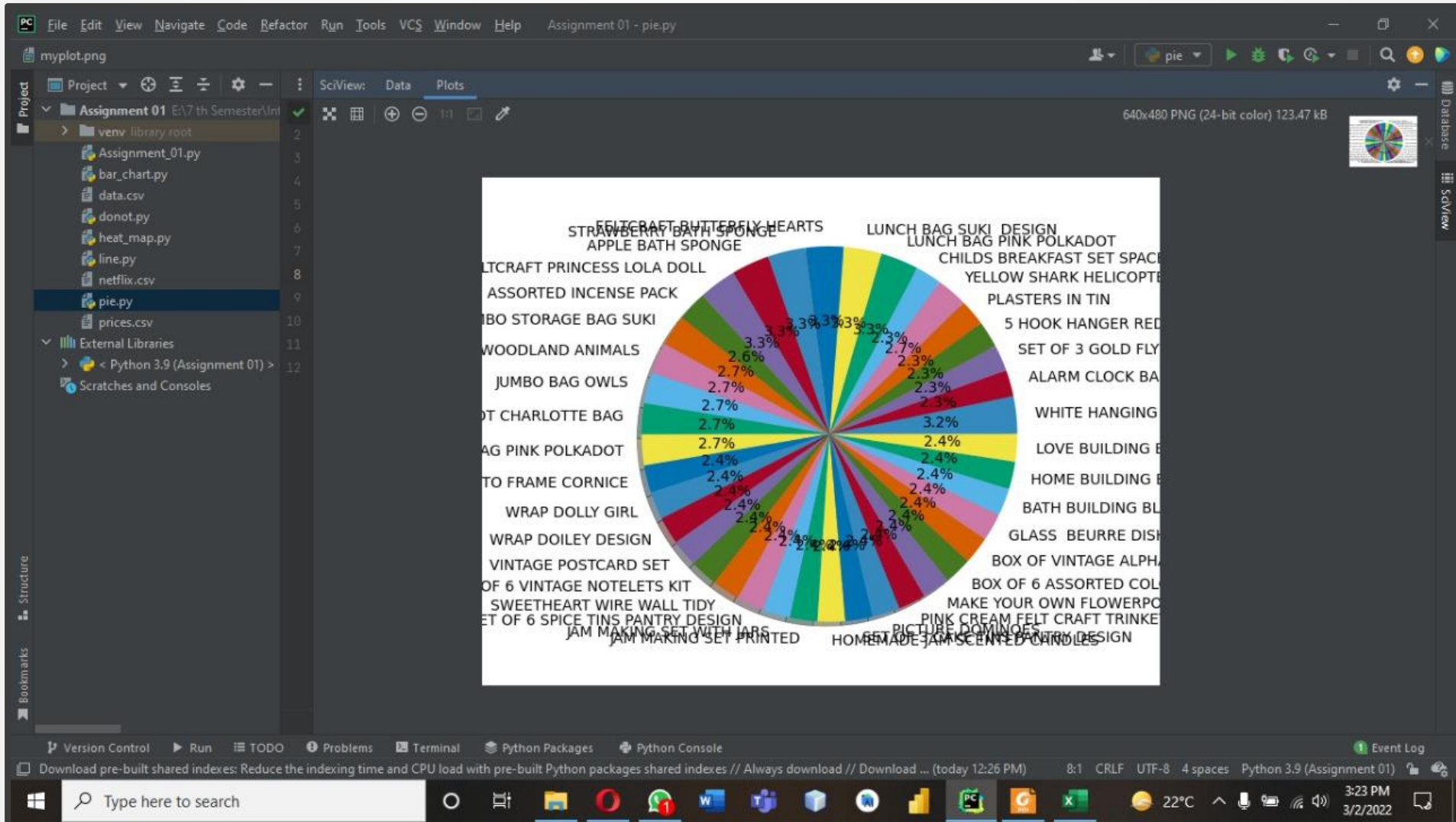
Screen Shot (Output Python)

LINE CHART OUTPUT



Screen Shot (Output Python)

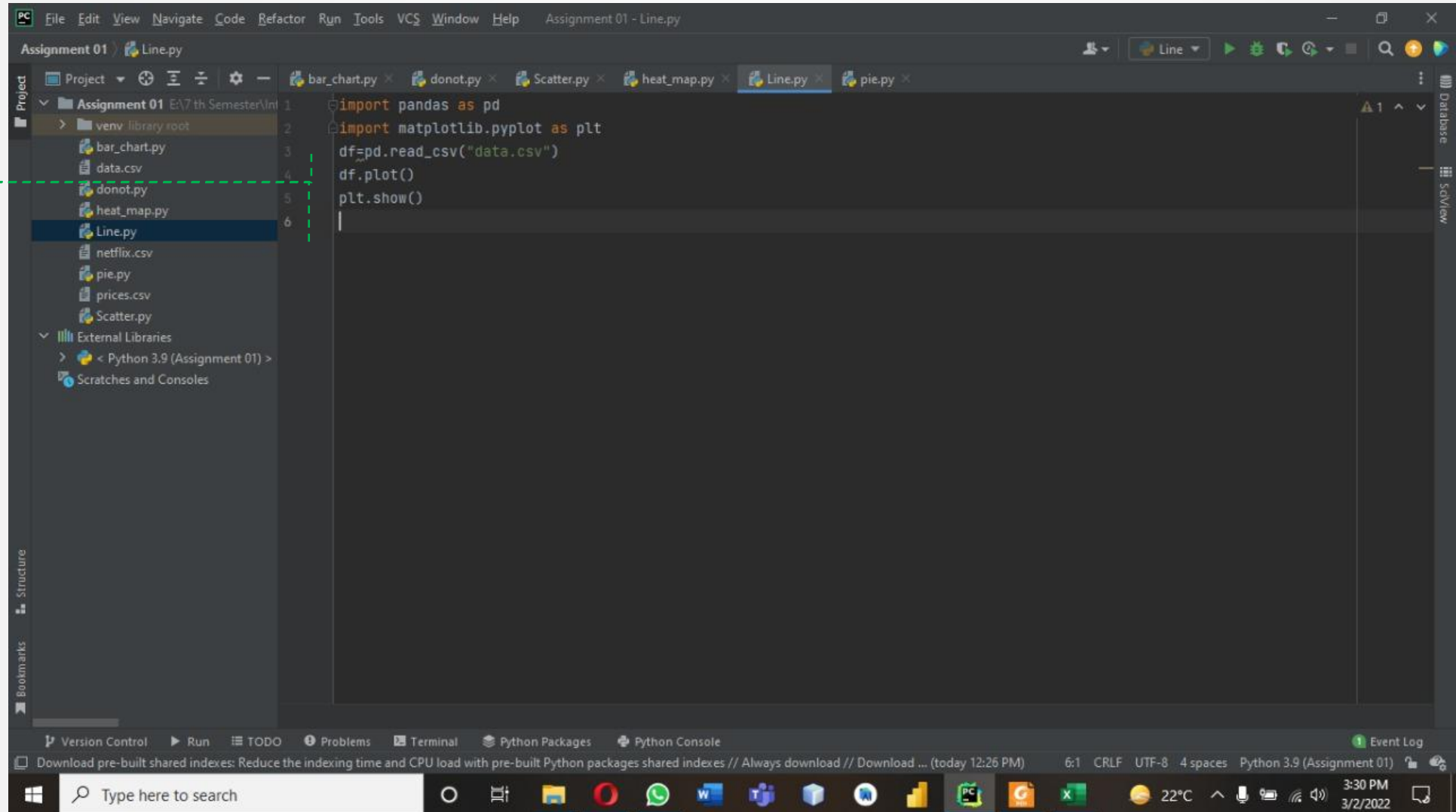
LINE CHART OUTPUT



Screen Shot (Code Python)

LINE Graph CODE

Line Graph



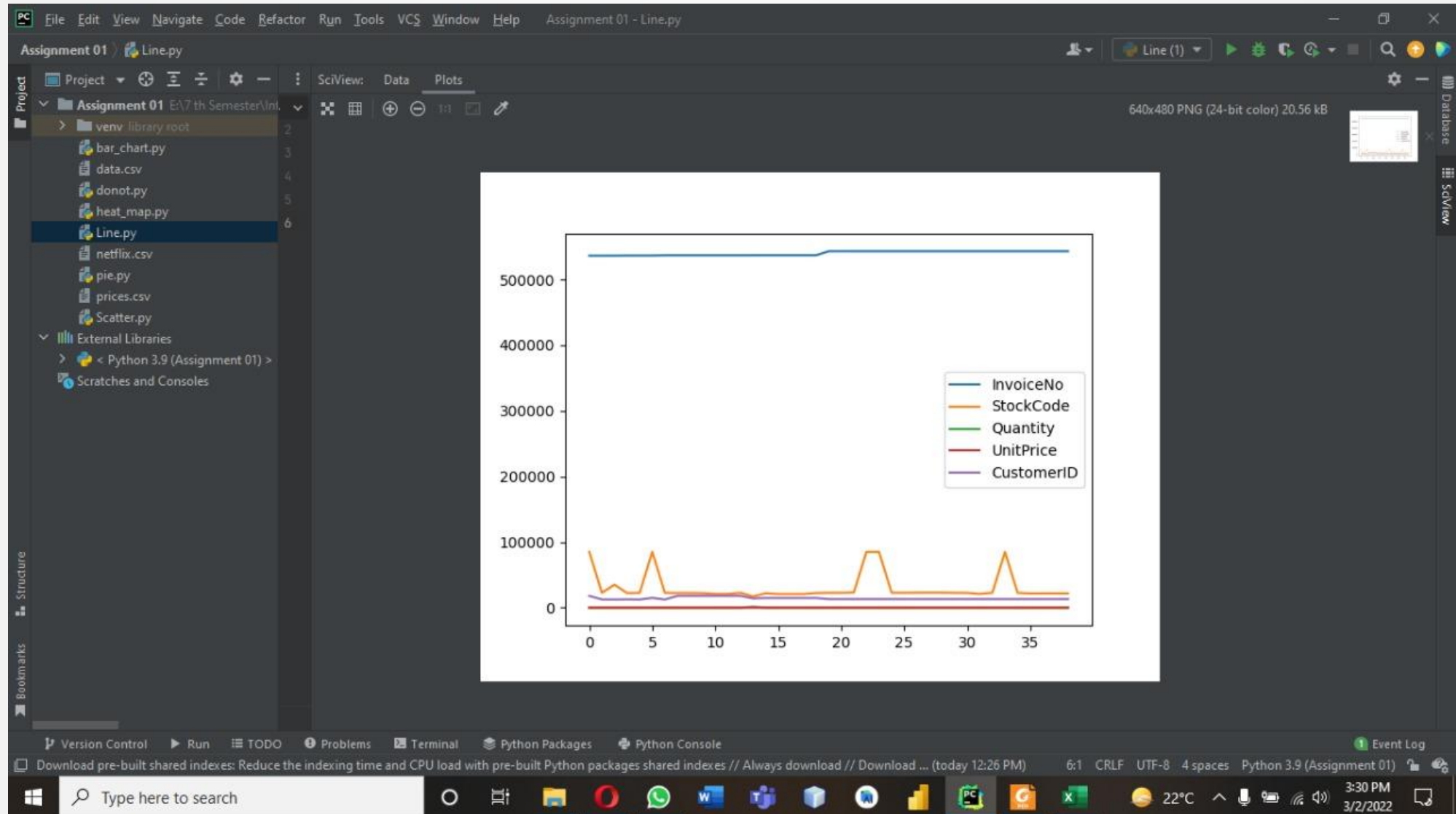
The screenshot shows an IDE window titled "Assignment 01 - Line.py". The left sidebar displays a project structure for "Assignment 01" with files like bar_chart.py, data.csv, donot.py, heat_map.py, Line.py (selected), netflix.csv, pie.py, prices.csv, and Scatter.py. The main editor area shows the following Python code:

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 df=pd.read_csv("data.csv")
4 df.plot()
5 plt.show()
6
```

The bottom status bar indicates the file is encoded in UTF-8 with 4 spaces, using Python 3.9 (Assignment 01). The Windows taskbar at the bottom shows the time as 3:30 PM on 3/2/2022.

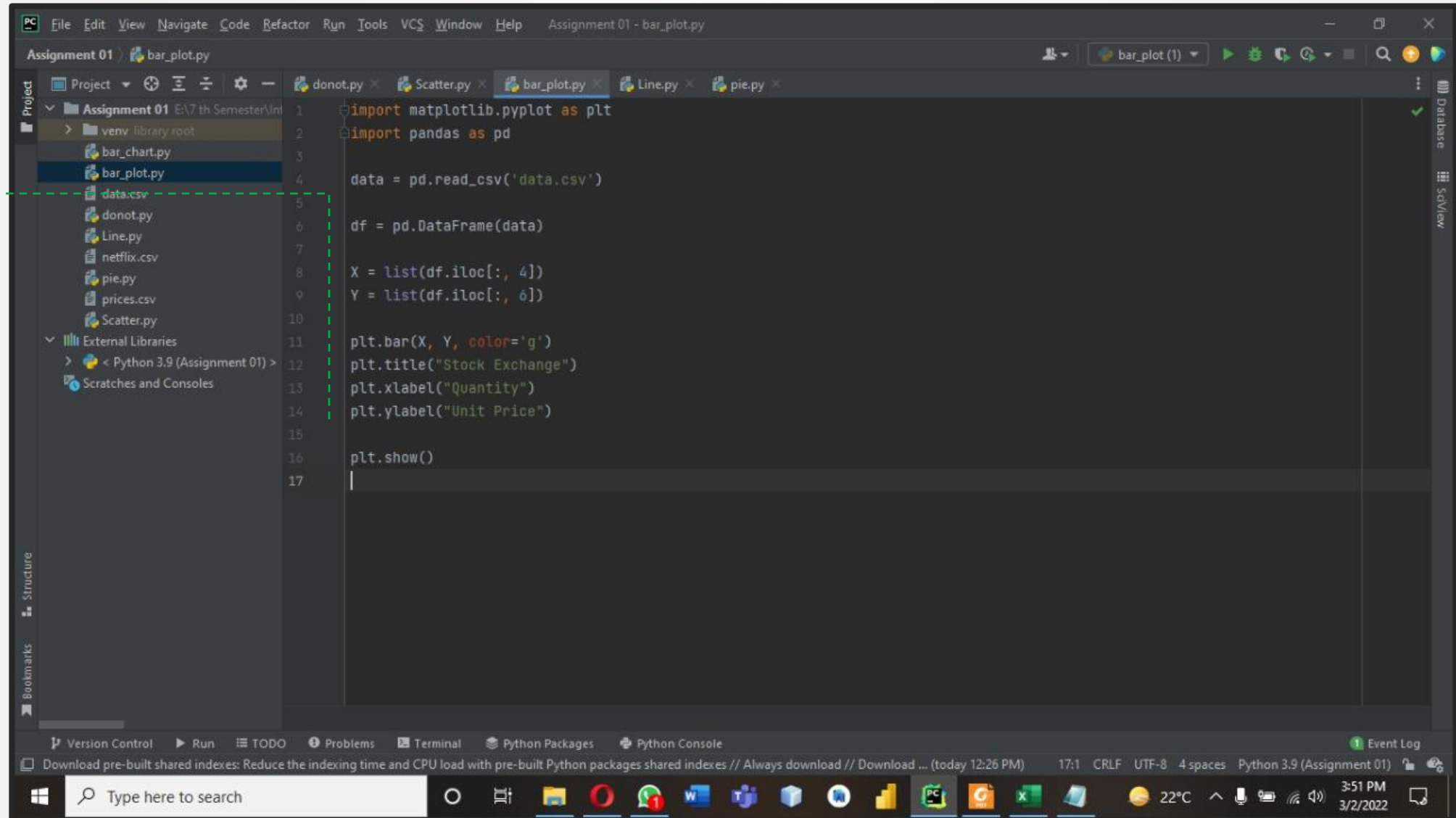
Screen Shot (Output Python)

LINE GRAPH OUTPUT



Screen Shot (Code Python)

Bar Plot Graph

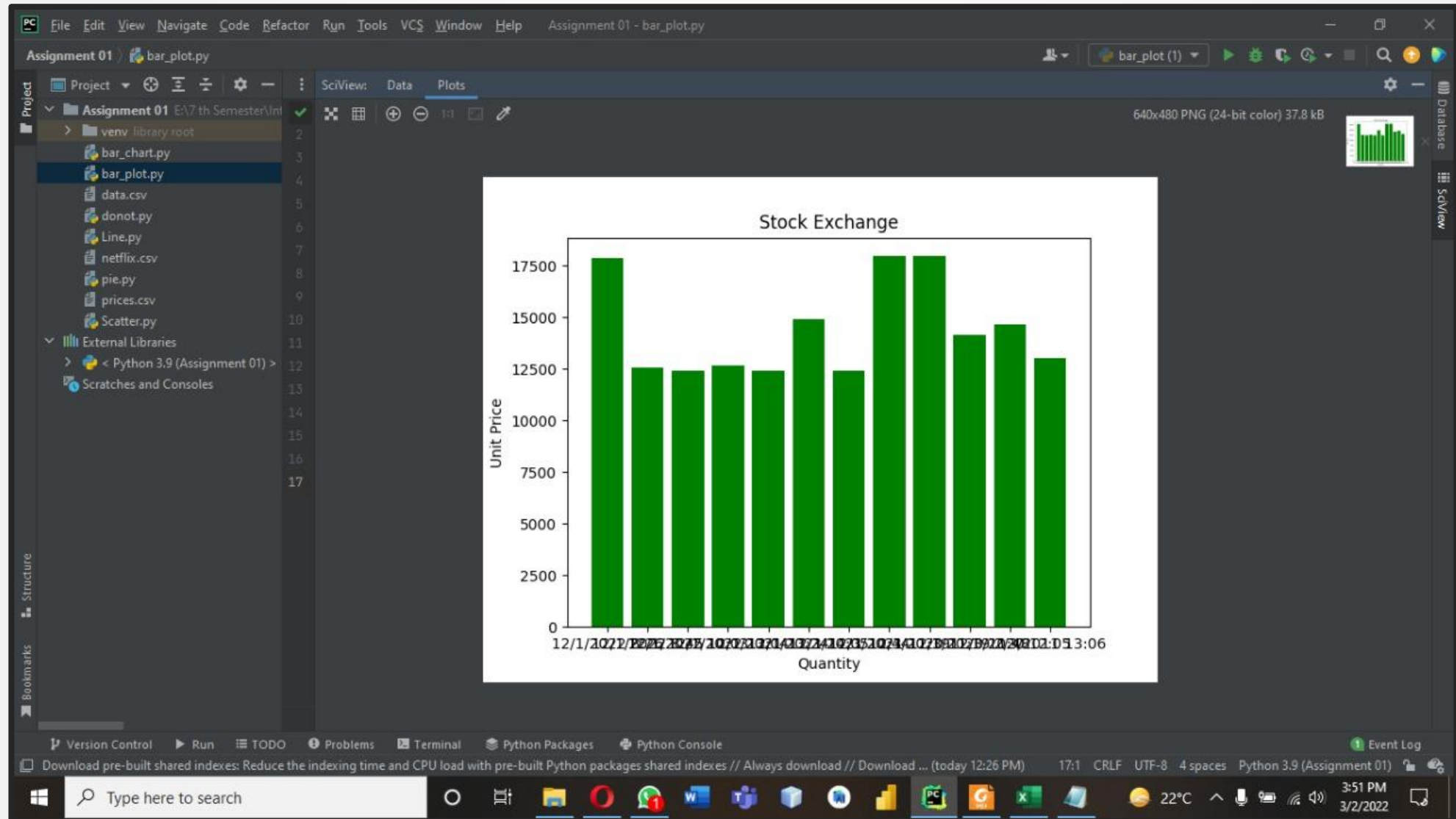


```
1 import matplotlib.pyplot as plt
2 import pandas as pd
3
4 data = pd.read_csv('data.csv')
5
6 df = pd.DataFrame(data)
7
8 X = list(df.iloc[:, 4])
9 Y = list(df.iloc[:, 6])
10
11 plt.bar(X, Y, color='g')
12 plt.title("Stock Exchange")
13 plt.xlabel("Quantity")
14 plt.ylabel("Unit Price")
15
16 plt.show()
17
```

Bar Plot

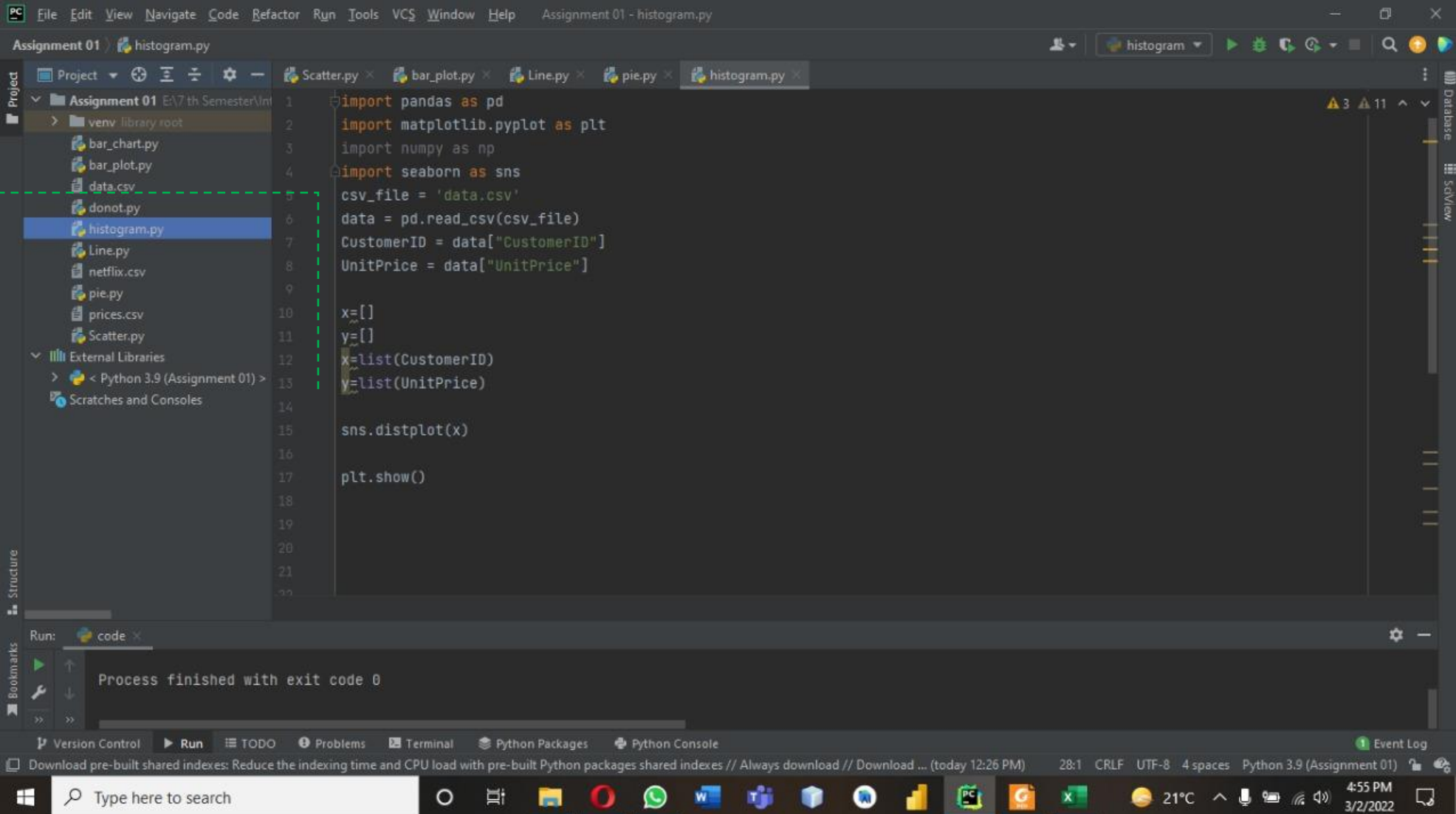
Screen Shot (Output Python)

LINE GRAPH OUTPUT



Screen Shot (Code Python)

Histogram



The screenshot shows a Python IDE with a project named 'Assignment 01'. The file explorer on the left shows a directory structure with files like 'bar_chart.py', 'bar_plot.py', 'data.csv', 'donot.py', 'histogram.py', 'Line.py', 'netflix.csv', 'pie.py', 'prices.csv', and 'Scatter.py'. The 'histogram.py' file is selected and open in the editor. The code in the editor is as follows:

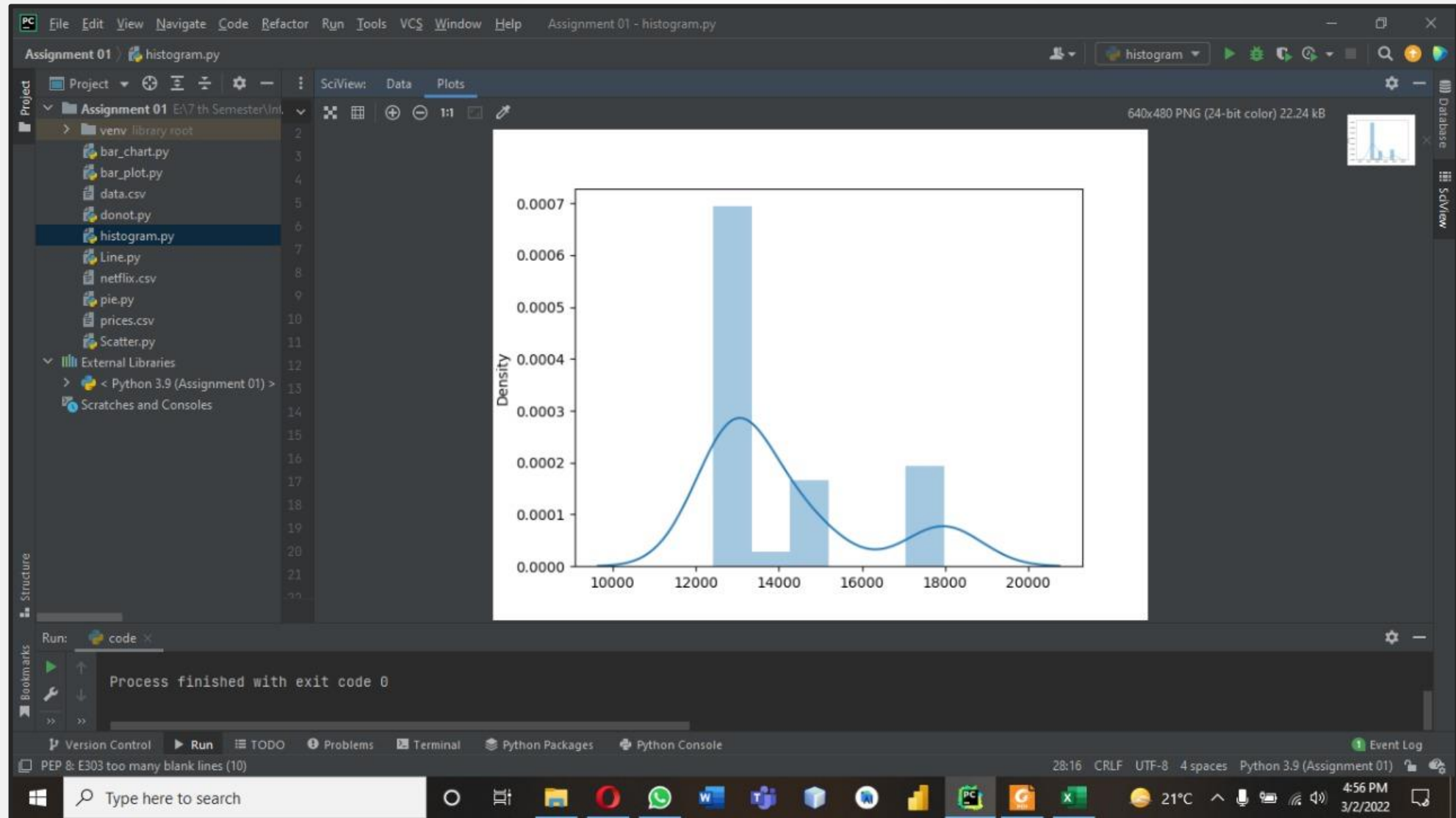
```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import numpy as np
4 import seaborn as sns
5 csv_file = 'data.csv'
6 data = pd.read_csv(csv_file)
7 CustomerID = data["CustomerID"]
8 UnitPrice = data["UnitPrice"]
9
10 x=[]
11 y=[]
12 x=list(CustomerID)
13 y=list(UnitPrice)
14
15 sns.distplot(x)
16
17 plt.show()
```

The IDE also shows a 'Run' tab at the bottom with the message 'Process finished with exit code 0'. The Windows taskbar at the bottom shows the time as 4:55 PM on 3/2/2022.

Histogram

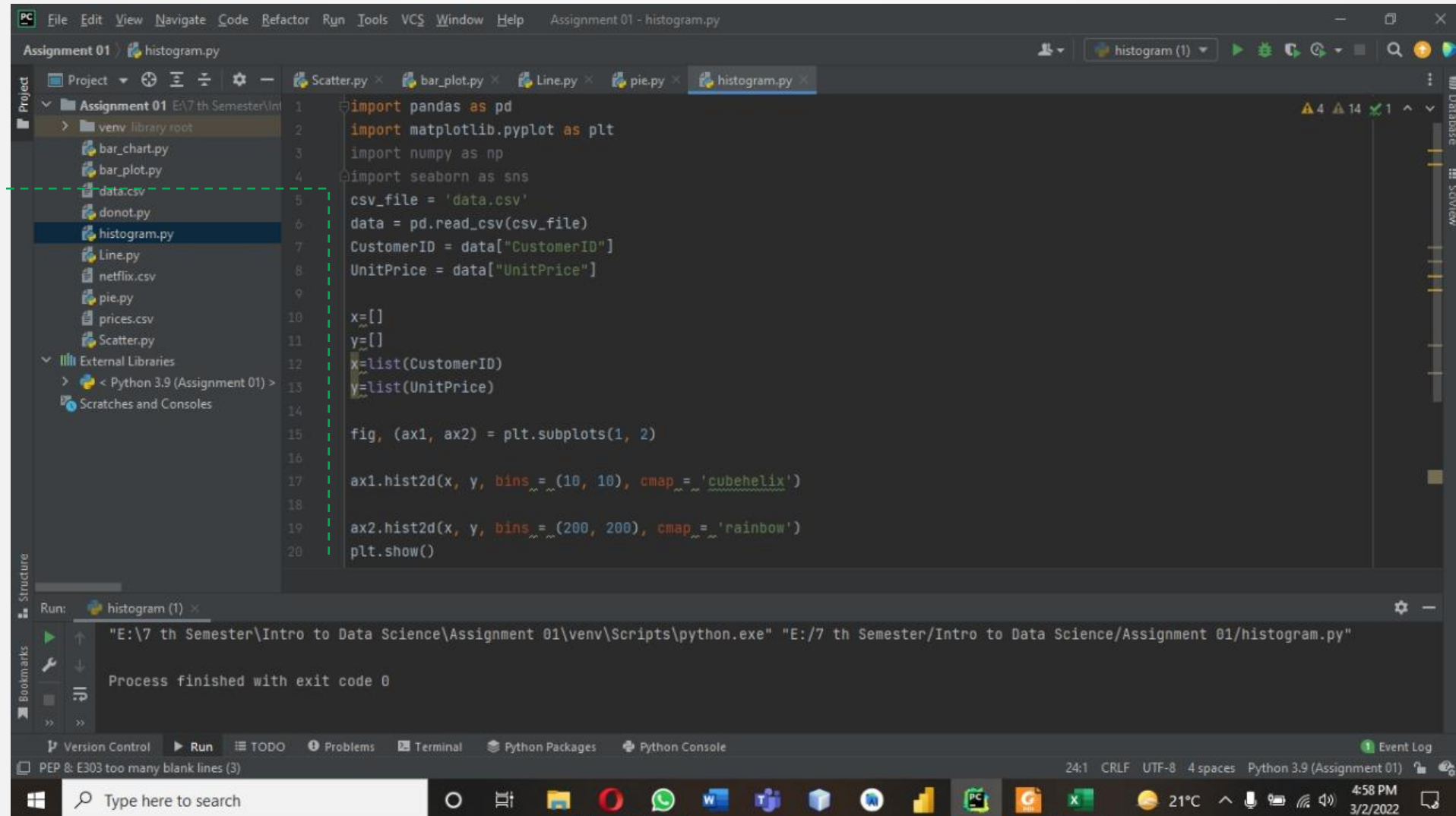
Screen Shot (Output Python)

Histogram



Screen Shot (Code Python)

Seaborn Histogram



The screenshot displays a Python IDE with a project named 'Assignment 01'. The file explorer on the left shows a directory structure with files like 'bar_chart.py', 'bar_plot.py', 'data.csv', 'donot.py', 'histogram.py', 'Line.py', 'netflix.csv', 'pie.py', 'prices.csv', and 'Scatter.py'. The 'histogram.py' file is selected and open in the editor. The code in the editor is as follows:

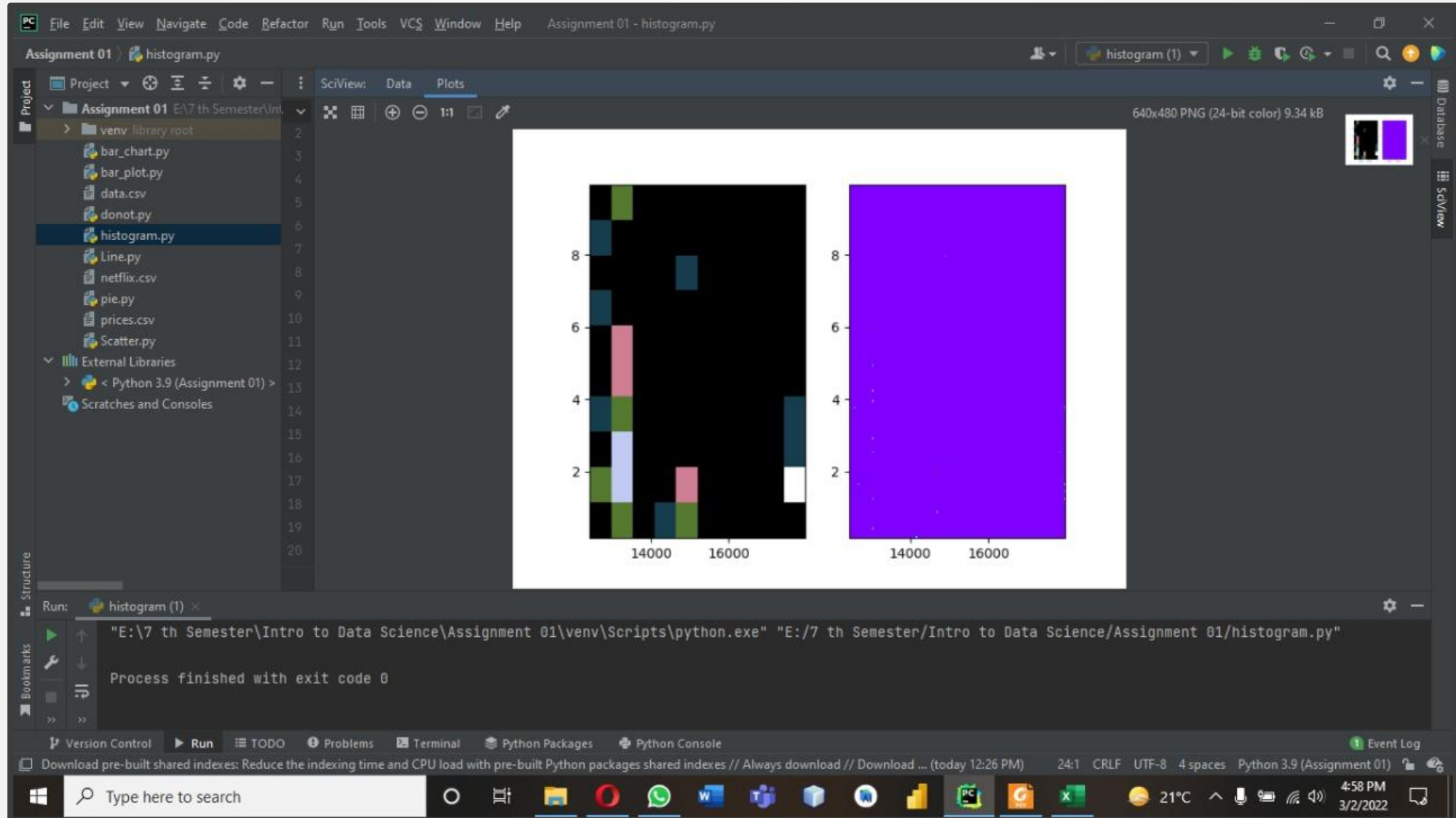
```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import numpy as np
4 import seaborn as sns
5 csv_file = 'data.csv'
6 data = pd.read_csv(csv_file)
7 CustomerID = data["CustomerID"]
8 UnitPrice = data["UnitPrice"]
9
10 x=[]
11 y=[]
12 x=list(CustomerID)
13 y=list(UnitPrice)
14
15 fig, (ax1, ax2) = plt.subplots(1, 2)
16
17 ax1.hist2d(x, y, bins_=(10, 10), cmap_='cubehelix')
18
19 ax2.hist2d(x, y, bins_=(200, 200), cmap_='rainbow')
20 plt.show()
```

The bottom of the IDE shows the 'Run' panel with the command: `"E:\7 th Semester\Intro to Data Science\Assignment 01\venv\Scripts\python.exe" "E:\7 th Semester\Intro to Data Science\Assignment 01\histogram.py"` and the message: `Process finished with exit code 0`. The status bar at the bottom indicates the file encoding is UTF-8, the editor has 4 spaces, and the Python version is 3.9 (Assignment 01). The system tray shows the date and time as 4:58 PM on 3/2/2022.

Seaborn
Histogram

Screen Shot (Output Python)

Sea Histogram



Group Details

Group 1

SP19-BSE-003

Wajahat Masood

SP19-BSE-007

Abdullah Rauf

SP19-BSE-013

Saif Ur Rehman

Email

Sp19-bse-003@cuiwah.edu.pk

Sp19-bse-007@cuiwah.edu.pk

Sp19-bse-013@cuiwah.edu.pk



Wah Campus
COMSATS ISLAMABAD

ENDING

STOCK EXCHANGE



References

<https://www.youtube.com/watch?v=9m4n2xVzk9o&t=265s>

<https://youtu.be/uufDGjTuq34>

<https://www.kaggle.com/carrie1/ecommerce-data>