Experiment 17

Aim: To Write a Pandas program to split the following dataframe by school code and get mean, min, and max value of age for each school.

Code

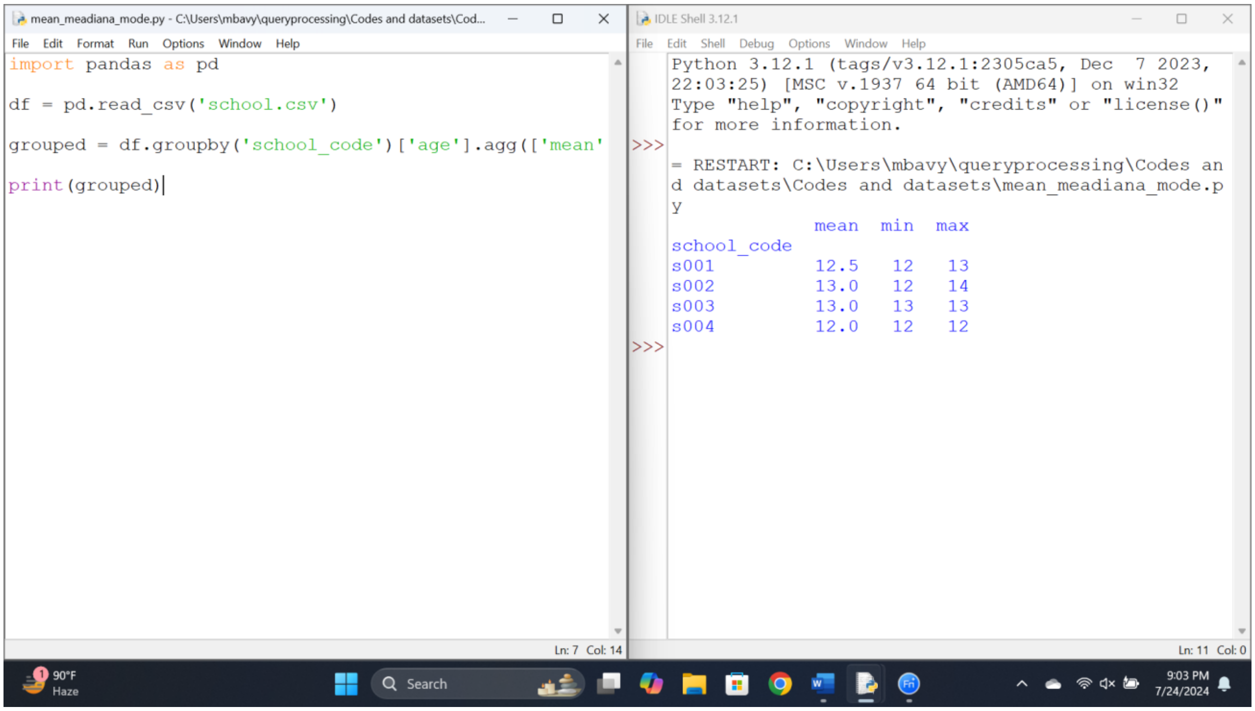
import pandas as pd

df = pd.read_csv('school.csv')

grouped = df.groupby('school_code')['age'].agg(['mean', 'min', 'max'])

print(grouped)

Output



mean_meadiana_mode.py - C:\Users\mbavy\queryprocessing\Codes and datasets\Cod...
File Edit Format Run Options Window Help
Ln: 7 Col: 14
IDLE Shell 3.12.1
File Edit Shell Debug Options Window Help
Ln: 11 Col: 0
90°F
Haze
Search
9:03 PM
7/24/2024

```
import pandas as pd

df = pd.read_csv('school.csv')

grouped = df.groupby('school_code')['age'].agg(['mean'

print(grouped)
```

```
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023,
22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
for more information.
>>>
= RESTART: C:\Users\mbavy\queryprocessing\Codes an
d datasets\Codes and datasets\mean_meadiana_mode.p
y
             mean  min  max
school_code
s001         12.5   12   13
s002         13.0   12   14
s003         13.0   13   13
s004         12.0   12   12
>>>
```

Results:

Thus a Pandas program to split the following dataframe by school code and get mean, min, and max value of age for each school.

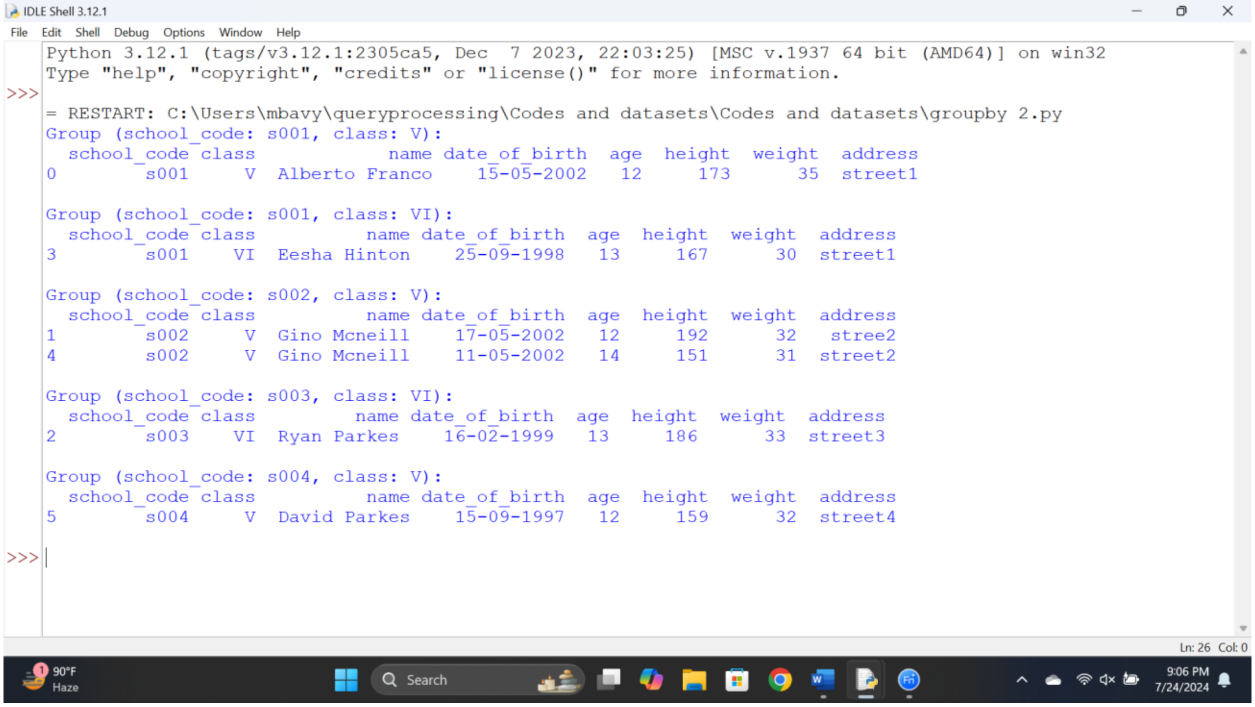
Experiment 18

Aim: To Write a Pandas program to split the following given dataframe into groups based on school code and class.

Code

```
import pandas as pd

df = pd.read_csv('school.csv')

grouped = df.groupby(['school_code', 'class'])

for (school_code, class_number), group in grouped:

    print(f"Group (school_code: {school_code}, class: {class_number}):")

    print(group)

    print()
```

Output:



```
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023, 22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:\Users\mbavy\queryprocessing\Codes and datasets\Codes and datasets\groupby 2.py
Group (school_code: s001, class: V):
  school_code class            name date_of_birth  age  height  weight  address
0        s001     V  Alberto Franco    15-05-2002   12     173      35  street1

Group (school_code: s001, class: VI):
  school_code class          name date_of_birth  age  height  weight  address
3        s001    VI  Eesha Hinton    25-09-1998   13     167      30  street1

Group (school_code: s002, class: V):
  school_code class          name date_of_birth  age  height  weight  address
1        s002     V  Gino Mcneill    17-05-2002   12     192      32   stree2
4        s002     V  Gino Mcneill    11-05-2002   14     151      31  street2

Group (school_code: s003, class: VI):
  school_code class         name date_of_birth  age  height  weight  address
2        s003    VI  Ryan Parkes    16-02-1999   13     186      33  street3

Group (school_code: s004, class: V):
  school_code class          name date_of_birth  age  height  weight  address
5        s004     V  David Parkes    15-09-1997   12     159      32  street4

>>>
```

Results:

Thus a Pandas program to split the following given dataframe into groups based on school code and class.

Experiment 19

Aim: To Write a Pandas program to display the dimensions or shape of the World alcohol consumption dataset.

Code:

import pandas as pd

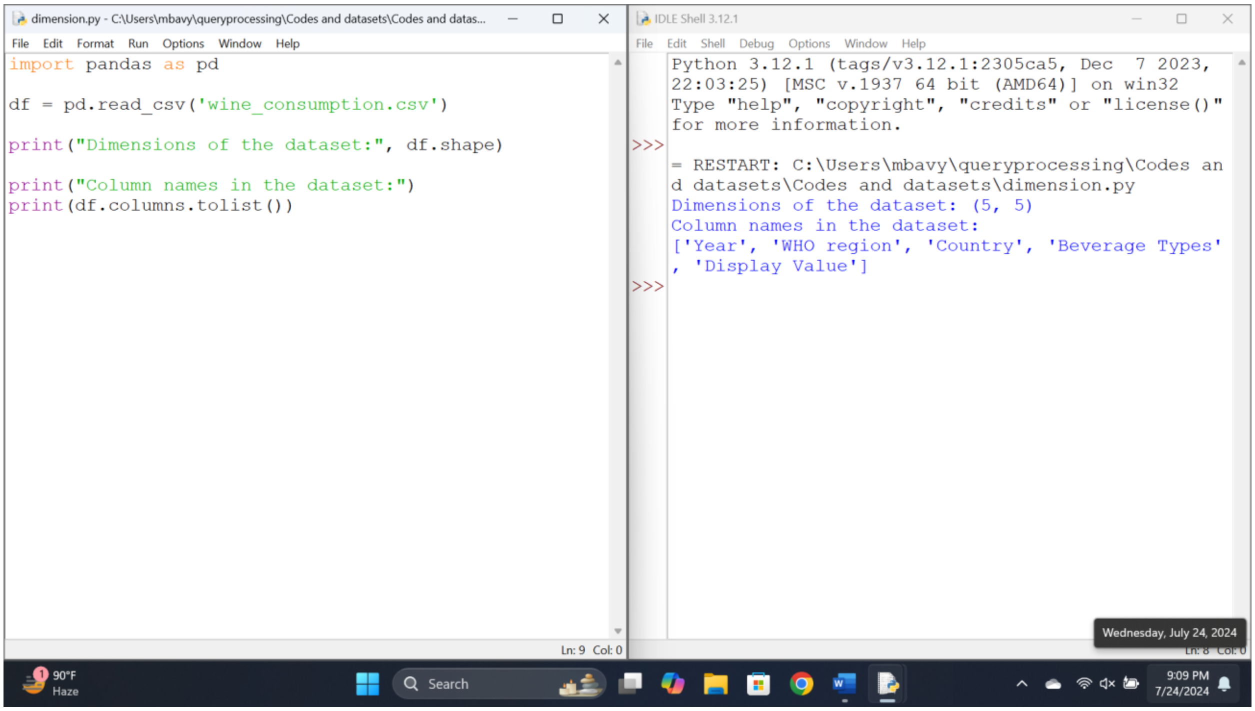
df = pd.read_csv('wine_consumption.csv')

print("Dimensions of the dataset:", df.shape)

print("Column names in the dataset:")

print(df.columns.tolist())

Output



dimension.py - C:\Users\mbavy\queryprocessing\Codes and datasets\Codes and datas...
File Edit Format Run Options Window Help
Ln: 9 Col: 0
IDLE Shell 3.12.1
File Edit Shell Debug Options Window Help
Wednesday, July 24, 2024
Ln: 8 Col: 0
90°F
Haze
Search
9:09 PM
7/24/2024

```python
import pandas as pd

df = pd.read_csv('wine_consumption.csv')

print("Dimensions of the dataset:", df.shape)

print("Column names in the dataset:")
print(df.columns.tolist())
```

```
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023,
22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
for more information.
>>>
= RESTART: C:\Users\mbavy\queryprocessing\Codes an
d datasets\Codes and datasets\dimension.py
Dimensions of the dataset: (5, 5)
Column names in the dataset:
['Year', 'WHO region', 'Country', 'Beverage Types'
, 'Display Value']
>>>
```

Results:

Thus Write a Pandas program to display the dimensions or shape of the World alcohol consumption dataset.

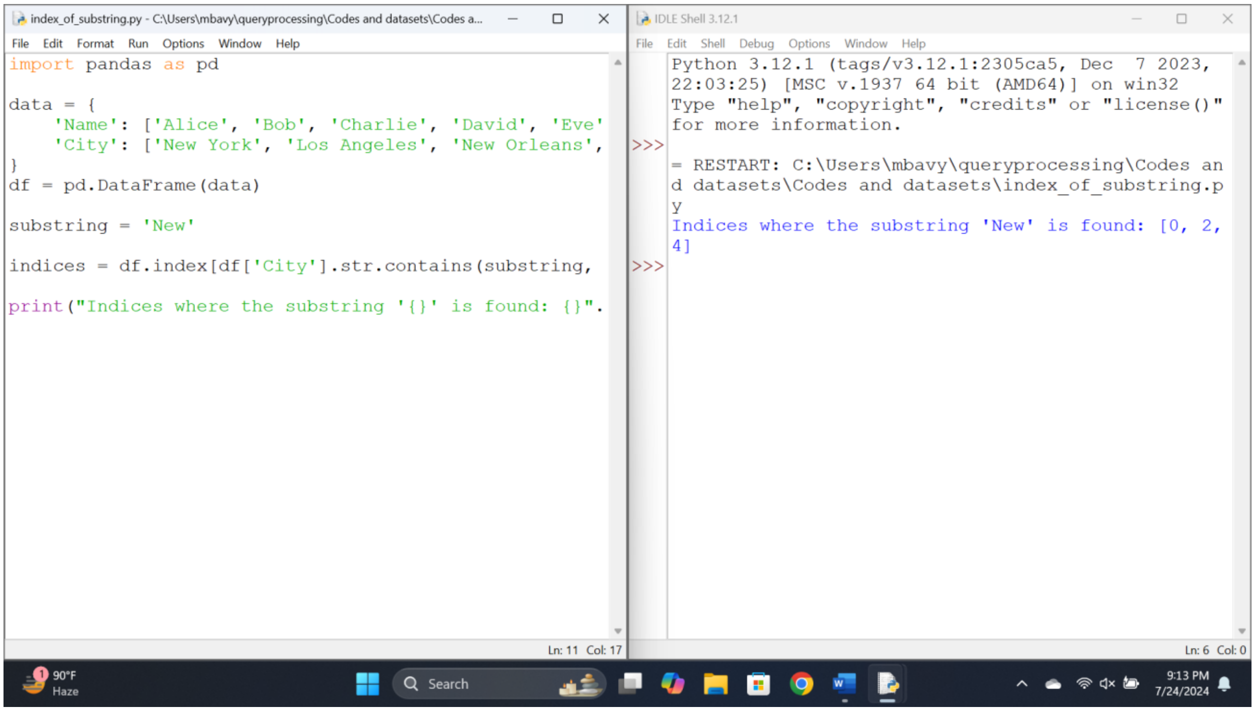
Experiment 20

Aim: To Write a Pandas program to find the index of a given substring of a DataFrame column.

Code

```
import pandas as pd

data = {

    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Eve'],

    'City': ['New York', 'Los Angeles', 'New Orleans', 'Chicago', 'New Haven']

}

df = pd.DataFrame(data)

substring = 'New'

indices = df.index[df['City'].str.contains(substring, case=False, na=False)].tolist()

print("Indices where the substring '{}' is found: {}".format(substring, indices))
```

Output



index_of_substring.py - C:\Users\mbavy\queryprocessing\Codes and datasets\Codes a...
File Edit Format Run Options Window Help
IDLE Shell 3.12.1
File Edit Shell Debug Options Window Help
Ln: 11 Col: 17
Ln: 6 Col: 0
90°F
Haze
Search
9:13 PM
7/24/2024

```
import pandas as pd

data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Eve'
    'City': ['New York', 'Los Angeles', 'New Orleans',
}
df = pd.DataFrame(data)

substring = 'New'

indices = df.index[df['City'].str.contains(substring,

print("Indices where the substring '{}' is found: {}".
```

```
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023,
22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
for more information.
>>>
= RESTART: C:\Users\mbavy\queryprocessing\Codes an
d datasets\Codes and datasets\index_of_substring.p
y
Indices where the substring 'New' is found: [0, 2,
4]
>>>
```

Results:

Thus a Pandas program to find the index of a given substring of a DataFrame column.

Experiment 21

Aim: Write a Pandas program to swap the cases of a specified character column in a given DataFrame.

Code:

import pandas as pd

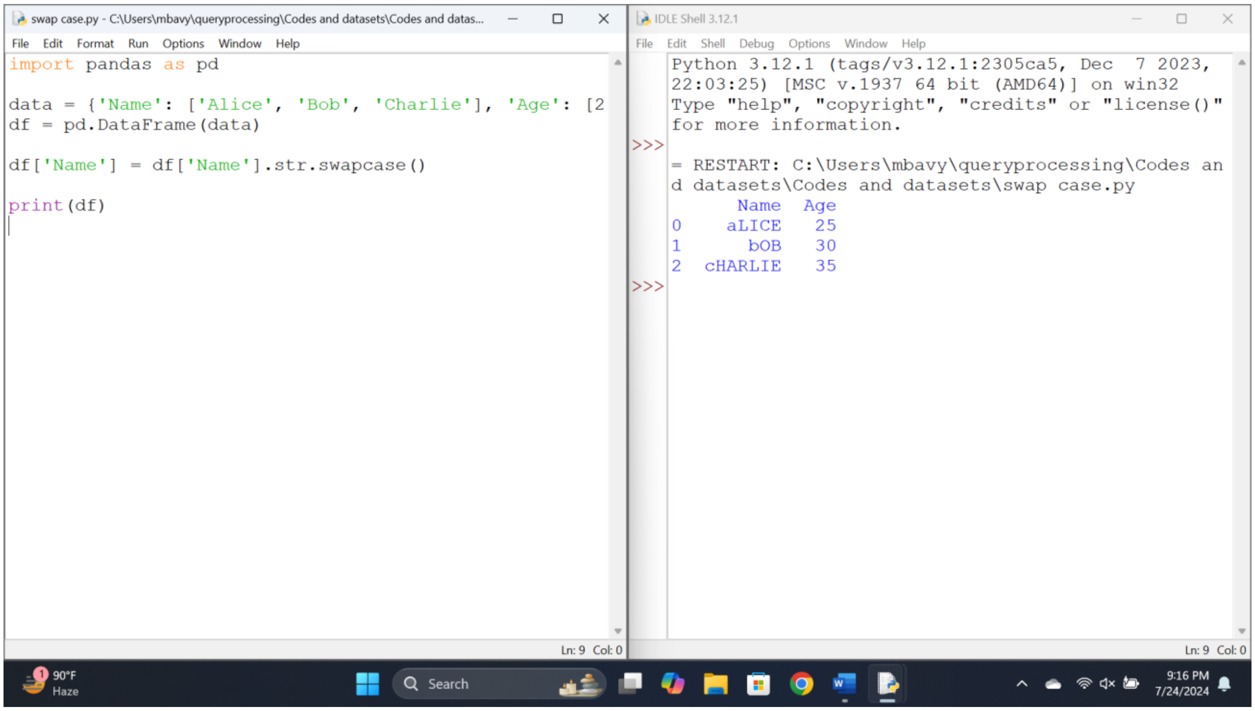
data = {'Name': ['Alice', 'Bob', 'Charlie'], 'Age': [25, 30, 35]}

df = pd.DataFrame(data)

df['Name'] = df['Name'].str.swapcase()

print(df)

Output:



swap case.py - C:\Users\mbavy\queryprocessing\Codes and datasets\Codes and datas...
File Edit Format Run Options Window Help
IDLE Shell 3.12.1
File Edit Shell Debug Options Window Help
Ln: 9 Col: 0
Ln: 9 Col: 0
90°F
Haze
Search
9:16 PM
7/24/2024

```
import pandas as pd

data = {'Name': ['Alice', 'Bob', 'Charlie'], 'Age': [2
df = pd.DataFrame(data)

df['Name'] = df['Name'].str.swapcase()

print(df)
```

```
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023,
22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
for more information.
>>>
= RESTART: C:\Users\mbavy\queryprocessing\Codes an
d datasets\Codes and datasets\swap case.py
      Name  Age
0    aLICE   25
1      bOB   30
2  cHARLIE   35
>>>
```

Results:

Thus a Pandas program to swap the cases of a specified character column in a given DataFrame.

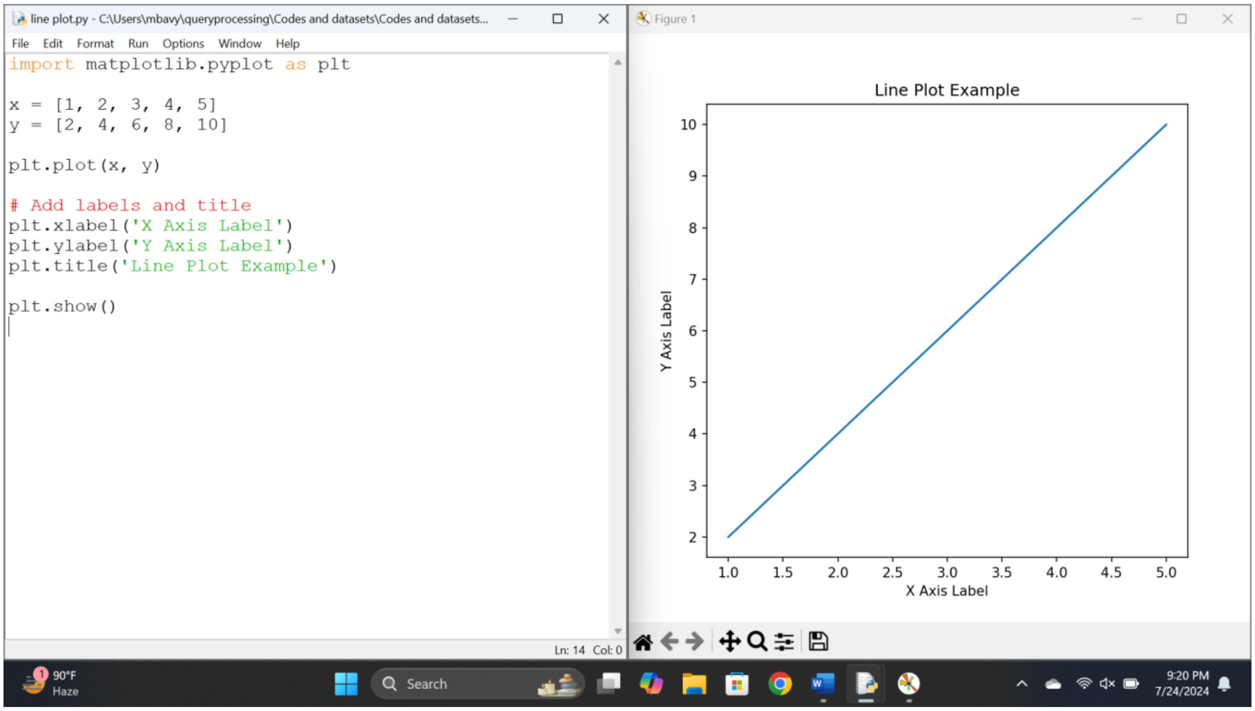
Experiment 22

Aim:To Write a Python program to draw a line with suitable label in the x axis, y axis and a title.

Code:

```
import matplotlib.pyplot as plt

x = [1, 2, 3, 4, 5]

y = [2, 4, 6, 8, 10]

plt.plot(x, y)

# Add labels and title

plt.xlabel('X Axis Label')

plt.ylabel('Y Axis Label')

plt.title('Line Plot Example')

plt.show()
```

Output:



```
import matplotlib.pyplot as plt

x = [1, 2, 3, 4, 5]
y = [2, 4, 6, 8, 10]

plt.plot(x, y)

# Add labels and title
plt.xlabel('X Axis Label')
plt.ylabel('Y Axis Label')
plt.title('Line Plot Example')

plt.show()
```

Results:

Thus a Python program to draw a line with suitable label in the x axis, y axis and a title.

Experiment 23

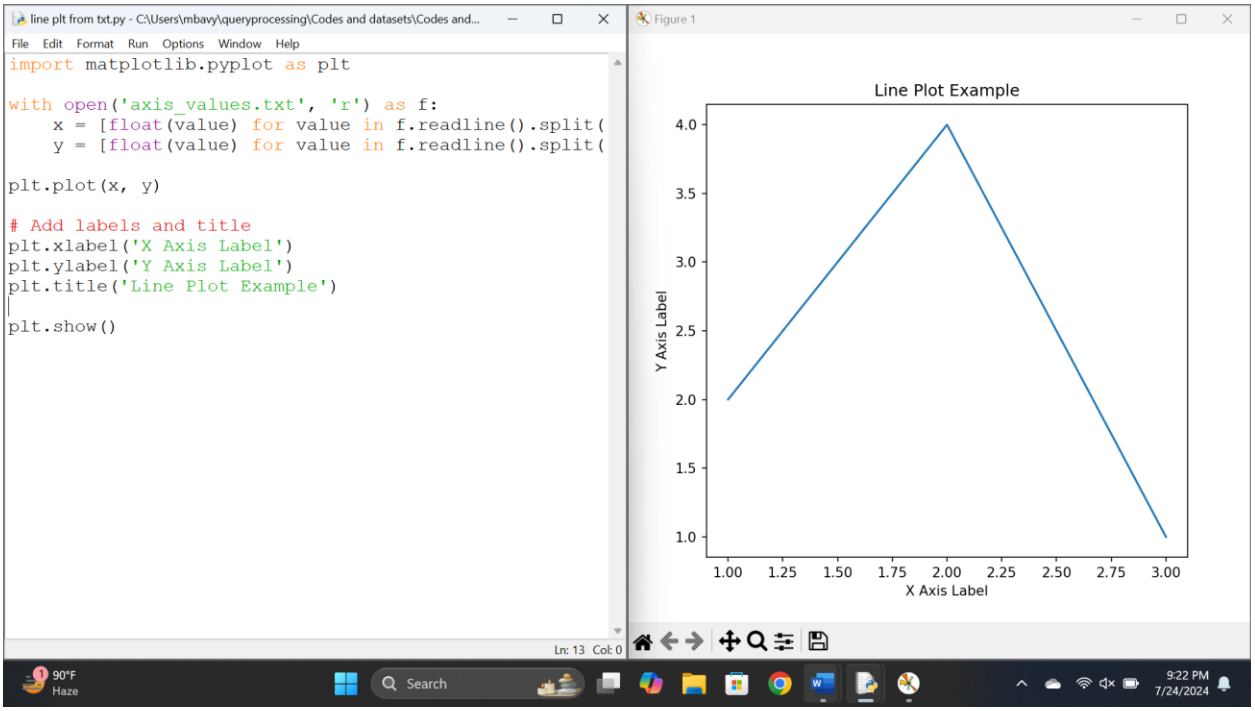
Aim:

To Write a Python program to draw a line using given axis values taken from a text file, with suitable label in the x axis, y axis and a title.

Code:

```
import matplotlib.pyplot as plt

with open('axis_values.txt', 'r') as f:

    x = [float(value) for value in f.readline().split()]

    y = [float(value) for value in f.readline().split()]

plt.plot(x, y)

# Add labels and title

plt.xlabel('X Axis Label')

plt.ylabel('Y Axis Label')

plt.title('Line Plot Example')

plt.show()
```

Output:



```
import matplotlib.pyplot as plt

with open('axis_values.txt', 'r') as f:
    x = [float(value) for value in f.readline().split(
    y = [float(value) for value in f.readline().split(

plt.plot(x, y)

# Add labels and title
plt.xlabel('X Axis Label')
plt.ylabel('Y Axis Label')
plt.title('Line Plot Example')

plt.show()
```

Results:

Thus a Python program to draw a line using given axis values taken from a text file, with suitable label in the x axis, y axis and a title.

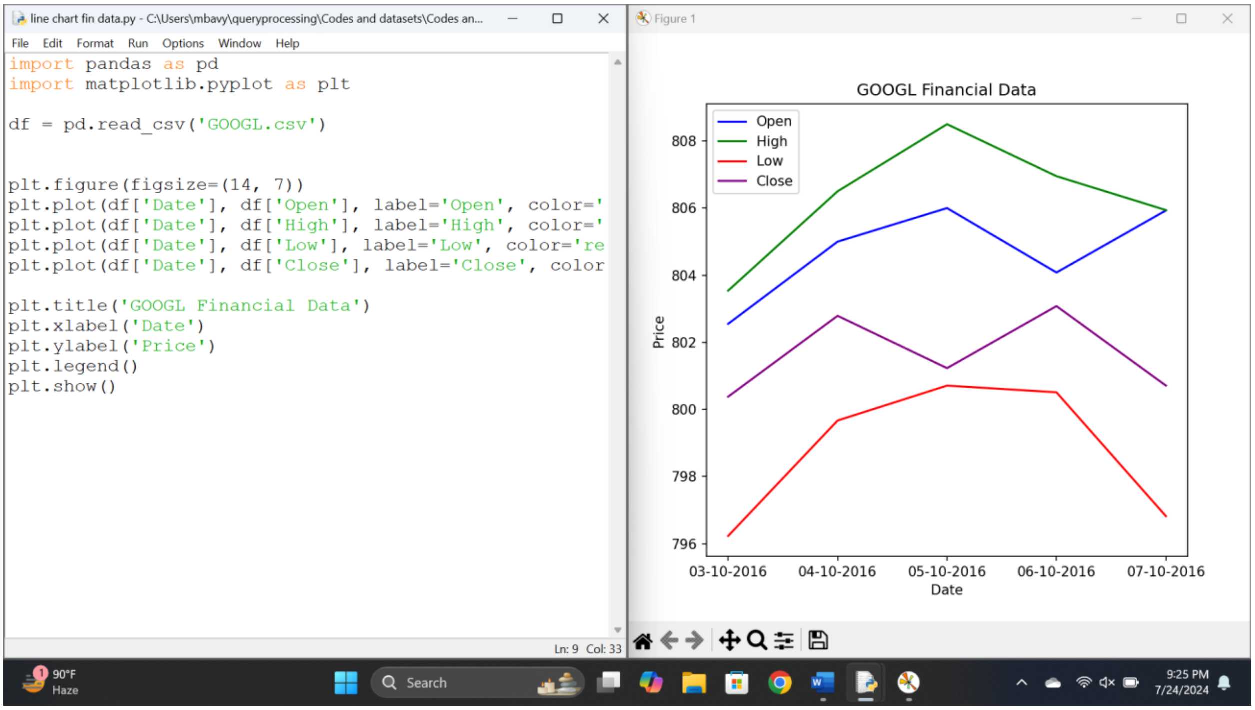
Experiment 24

Aim: To Write a Python program to draw line charts of the financial data of Alphabet Inc. between October 3, 2016 to October 7, 2016.

Code:

```
import pandas as pd

import matplotlib.pyplot as plt

df = pd.read_csv('GOOGL.csv')

plt.figure(figsize=(14, 7))

plt.plot(df['Date'], df['Open'], label='Open', color='blue')

plt.plot(df['Date'], df['High'], label='High', color='green')

plt.plot(df['Date'], df['Low'], label='Low', color='red')

plt.plot(df['Date'], df['Close'], label='Close', color='purple')

plt.title('GOOGL Financial Data')

plt.xlabel('Date')

plt.ylabel('Price')

plt.legend()

plt.show()
```

Output:



```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('GOOGL.csv')


plt.figure(figsize=(14, 7))
plt.plot(df['Date'], df['Open'], label='Open', color='
plt.plot(df['Date'], df['High'], label='High', color='
plt.plot(df['Date'], df['Low'], label='Low', color='re
plt.plot(df['Date'], df['Close'], label='Close', color

plt.title('GOOGL Financial Data')
plt.xlabel('Date')
plt.ylabel('Price')
plt.legend()
plt.show()
```

Results:

Thus a Python program to draw line charts of the financial data of Alphabet Inc. between October 3, 2016 to October 7, 2016.