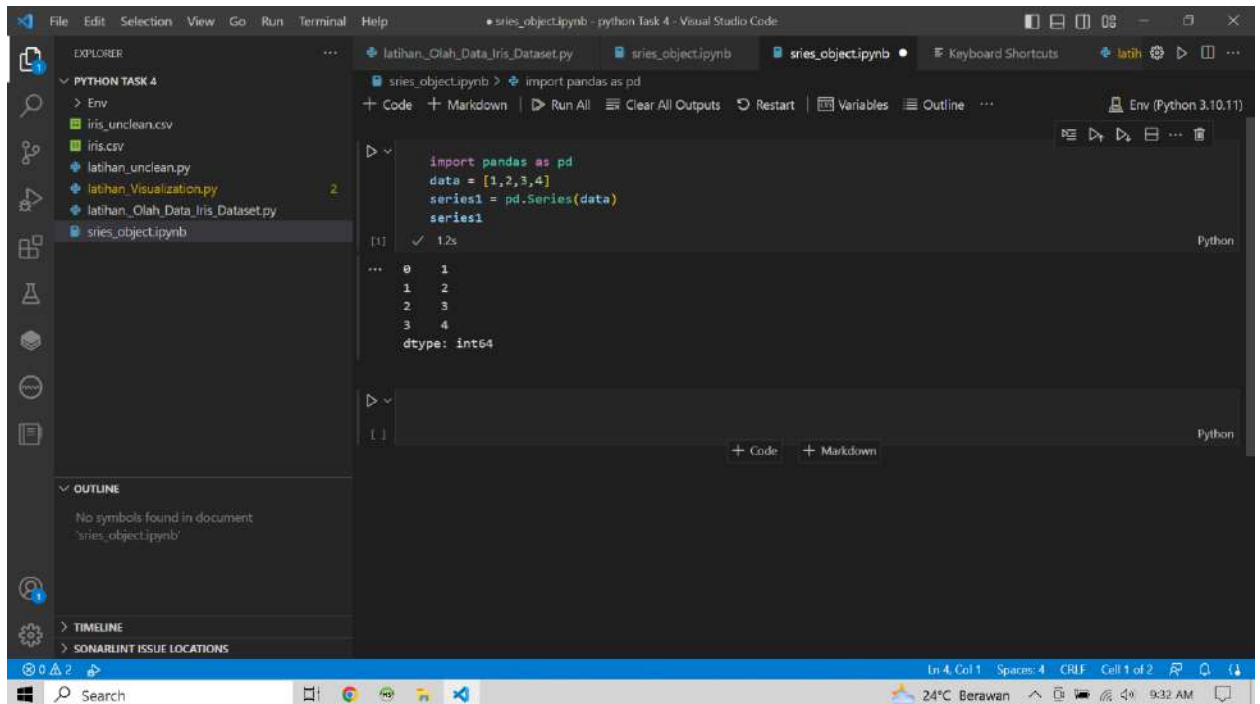


Nama : Ambar Wati  
NIM : 20.01.013.001  
Kelas : C  
Mata Kuliah : Perograman Python

## Task 6

### 4-Data Visualisasi Data-Pandas DataFrame

#### 1. membuat Series



The screenshot shows the Visual Studio Code interface with a Python file named `sries_object.ipynb` open. The code in the first cell is:

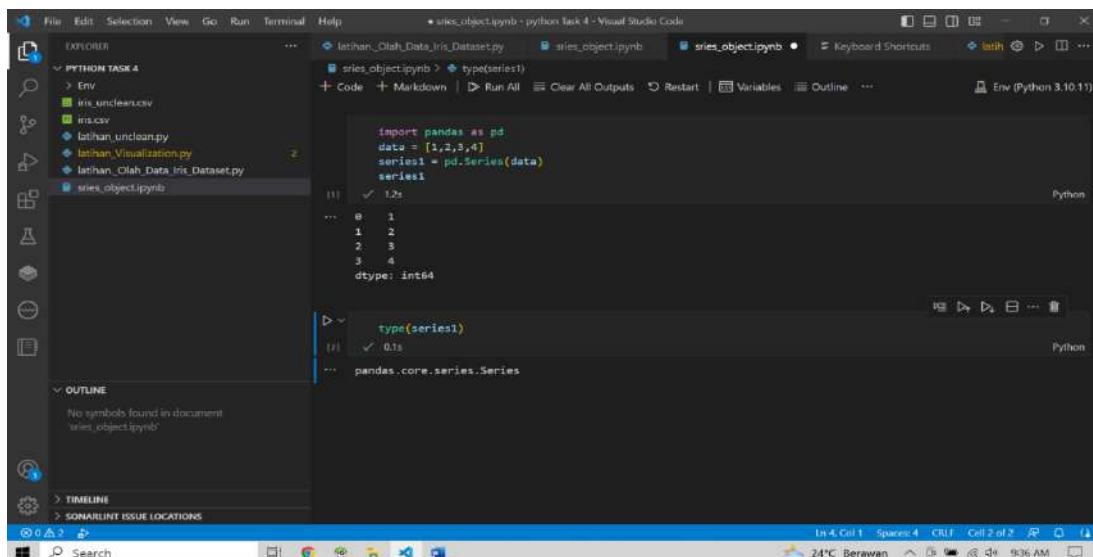
```
import pandas as pd
data = [1,2,3,4]
series1 = pd.Series(data)
```

The output of the cell is displayed below the code:

```
[1] ✓ 12s
...
0 1
1 2
2 3
3 4
dtype: int64
```

The Explorer sidebar on the left shows the file structure for 'PYTHON TASK 4', including `Env`, `iris_unclean.csv`, `iris.csv`, `latihan_unclean.py`, `latihan_Visualization.py`, `latihan_Olah_Data_Iris_Dataset.py`, and `sries_object.ipynb`. The Outline sidebar shows 'No symbols found in document 'sries\_object.ipynb''. The status bar at the bottom indicates 'In 4, Col 1, Spaces: 4, CRLF, Cell 1 of 2' and the system clock shows '24°C Berawan' and '9:32 AM'.

#### 2. Cek Tipe Struktur Data Series atau Bukan



The screenshot shows the Visual Studio Code interface with the same Python file `sries_object.ipynb`. The code in the first cell is the same as in the previous screenshot. The output is the same. A second cell of code is added below the first:

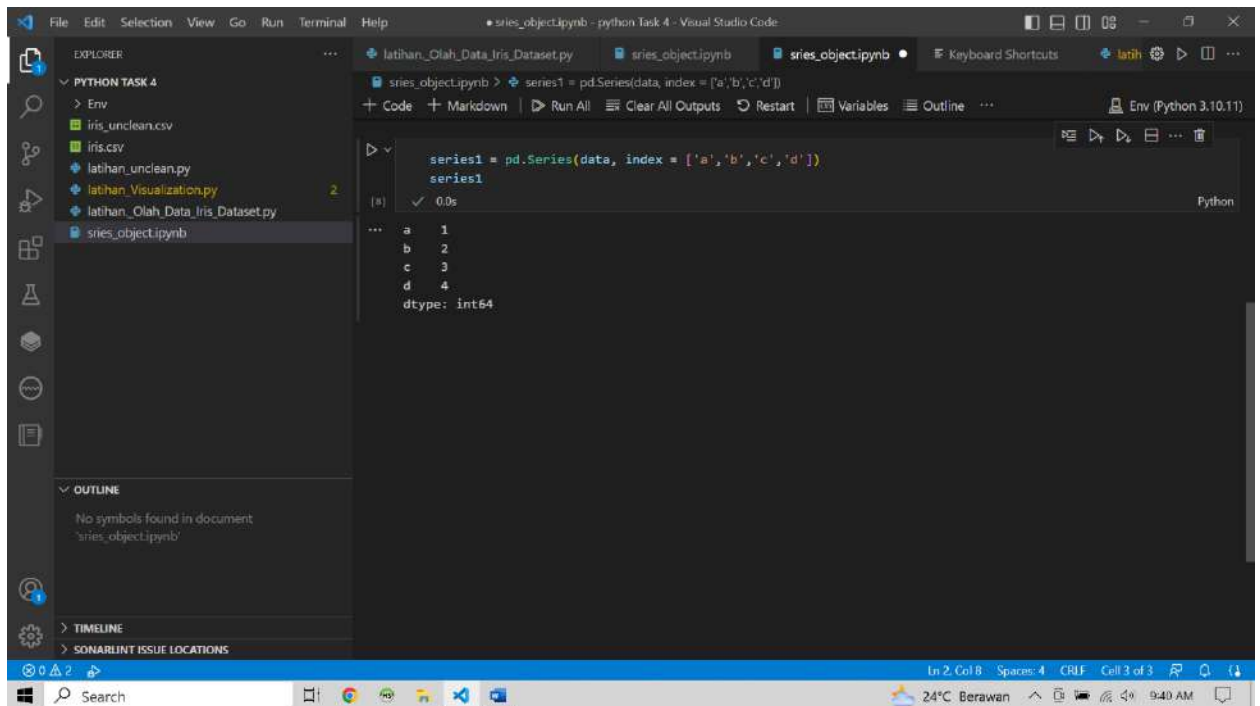
```
type(series1)
```

The output of the second cell is displayed below the code:

```
[1] ✓ 0.1s
...
pandas.core.series.Series
```

The Explorer sidebar and Outline sidebar are the same as in the previous screenshot. The status bar at the bottom indicates 'In 4, Col 1, Spaces: 4, CRLF, Cell 2 of 2' and the system clock shows '24°C Berawan' and '9:36 AM'.

### 3. Ubah Nama Index



The screenshot shows a Visual Studio Code window with a Jupyter Notebook. The Explorer sidebar on the left shows a project named 'PYTHON TASK 4' with files like 'iris\_unclean.csv', 'iris.csv', and 'sries\_object.ipynb'. The main editor displays a code cell in 'sries\_object.ipynb' with the following code:

```
series1 = pd.Series(data, index = ['a','b','c','d'])
series1
```

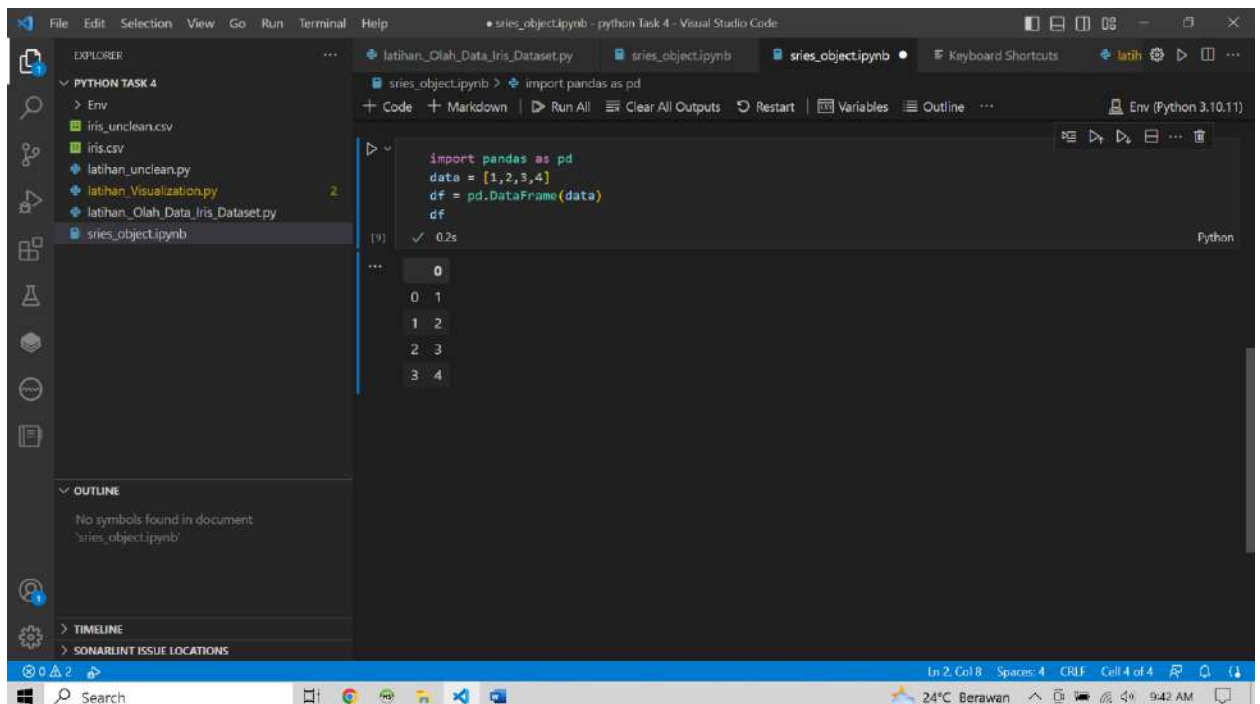
The output of the cell is a pandas Series:

```
a    1
b    2
c    3
d    4
dtype: int64
```

The status bar at the bottom indicates 'Ln 2, Col 8', 'Spaces: 4', 'CR LF', 'Cell 2 of 3', and the system clock shows '24°C Berawan' and '9:40 AM'.

## Data Frame

### 1. DataFrame menggunakan List



The screenshot shows a Visual Studio Code window with a Jupyter Notebook. The Explorer sidebar on the left shows the same project as the previous image. The main editor displays a code cell in 'sries\_object.ipynb' with the following code:

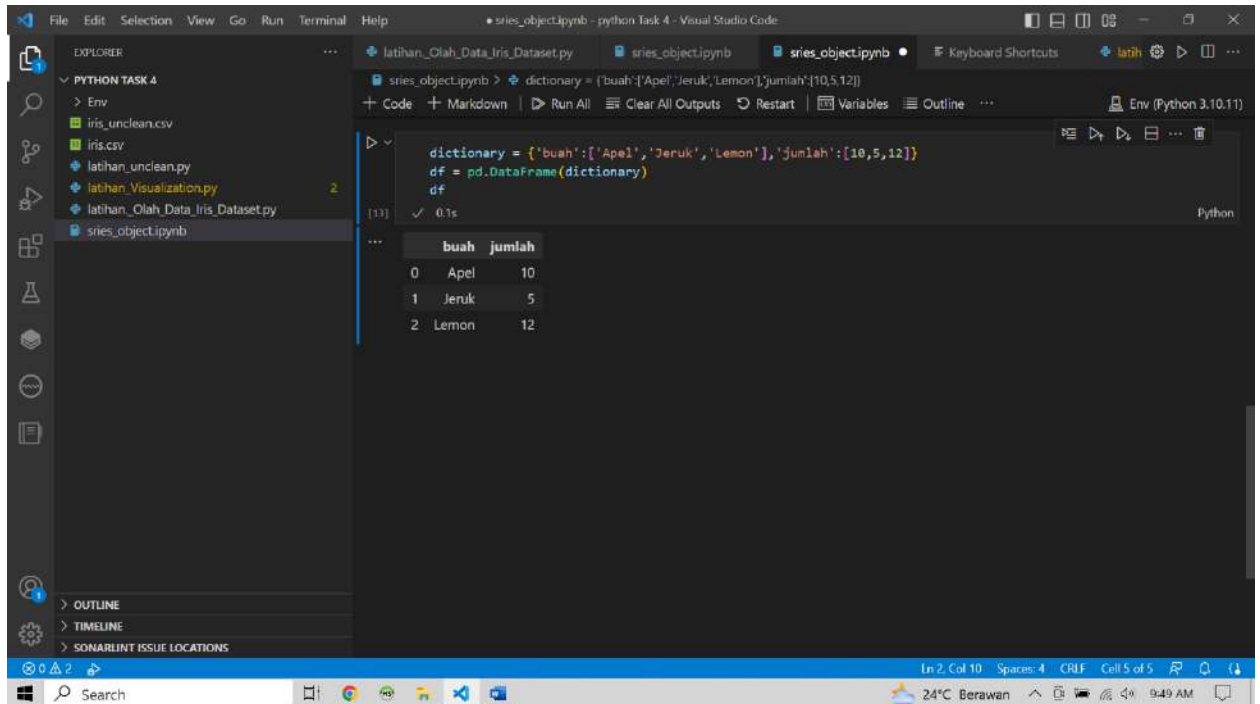
```
import pandas as pd
data = [1,2,3,4]
df = pd.DataFrame(data)
df
```

The output of the cell is a pandas DataFrame:

```
0
0  1
1  2
2  3
3  4
```

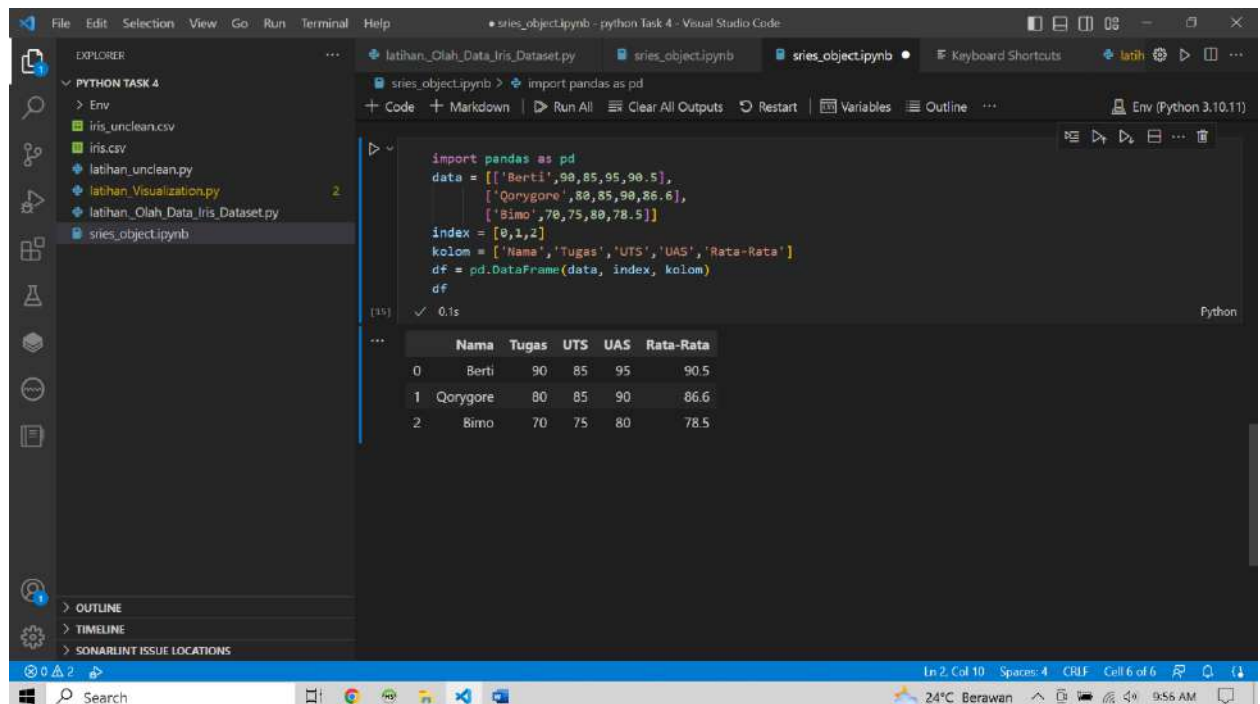
The status bar at the bottom indicates 'Ln 2, Col 8', 'Spaces: 4', 'CR LF', 'Cell 4 of 4', and the system clock shows '24°C Berawan' and '9:42 AM'.

## 2. DataFrame Menggunakan Dictionary



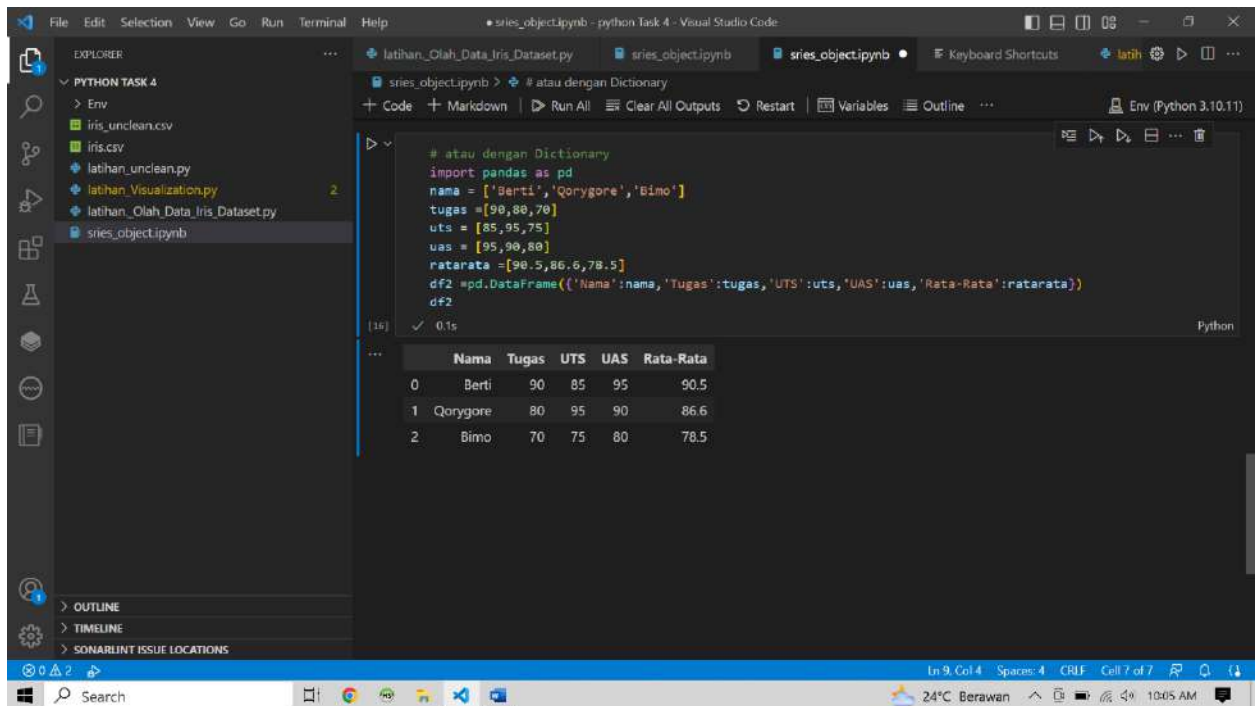
```
File Edit Selection View Go Run Terminal Help
sries_object.ipynb - python Task 4 - Visual Studio Code
latihan_Olah_Data_Iris_Dataset.py sries_object.ipynb sries_object.ipynb Keyboard Shortcuts latih
sries_object.ipynb > dictionary = {'buah': ['Apel', 'Jeruk', 'Lemon'], 'jumlah': [10, 5, 12]}
+ Code + Markdown | Run All | Clear All Outputs | Restart | Variables | Outline ...
df = pd.DataFrame(dictionary)
[13] ✓ 0.1s Python
...
    buah  jumlah
0   Apel     10
1   Jeruk     5
2   Lemon    12
```

## 3. DataFrame Menggunakan List dengan Tipe Data Campuran



```
File Edit Selection View Go Run Terminal Help
sries_object.ipynb - python Task 4 - Visual Studio Code
latihan_Olah_Data_Iris_Dataset.py sries_object.ipynb sries_object.ipynb Keyboard Shortcuts latih
sries_object.ipynb > import pandas as pd
+ Code + Markdown | Run All | Clear All Outputs | Restart | Variables | Outline ...
import pandas as pd
data = [['Berti', 90, 85, 95, 90.5],
        ['Qorygore', 80, 85, 90, 86.6],
        ['Bimo', 70, 75, 80, 78.5]]
index = [0, 1, 2]
kolom = ['Nama', 'Tugas', 'UTS', 'UAS', 'Rata-Rata']
df = pd.DataFrame(data, index, kolom)
df
[15] ✓ 0.1s Python
...
    Nama  Tugas  UTS  UAS  Rata-Rata
0   Berti    90   85   95    90.5
1  Qorygore   80   85   90    86.6
2    Bimo    70   75   80    78.5
```

#### 4. DataFrame Menggunakan List & Dictionary dengan Tipe Data Campuran



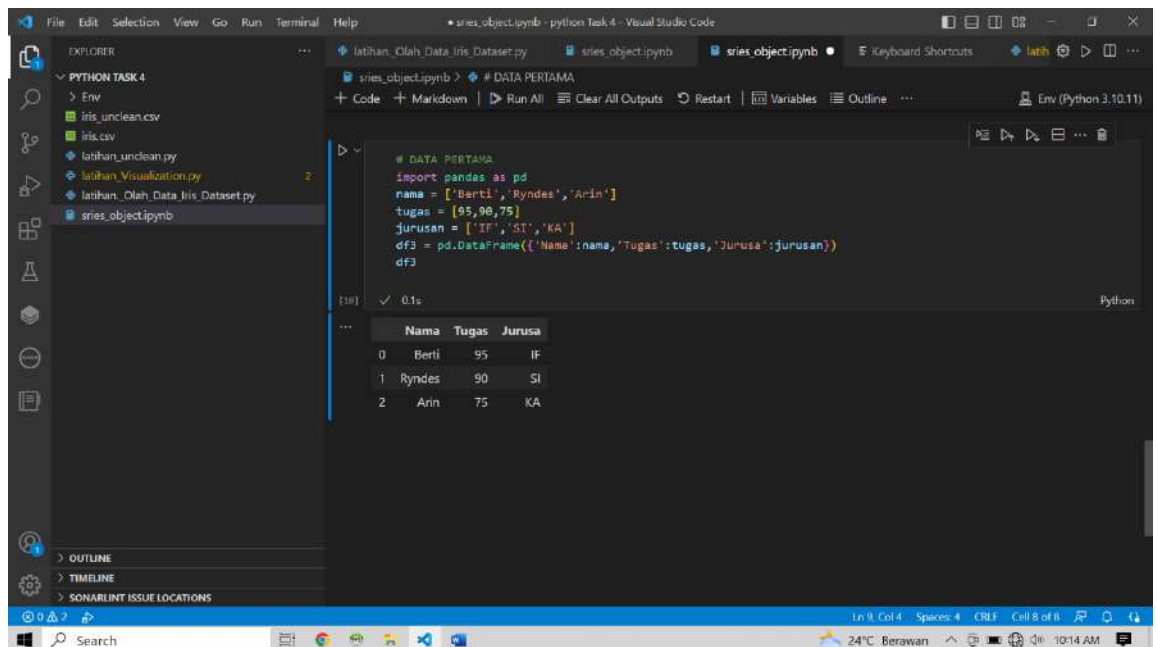
```
# atau dengan Dictionary
import pandas as pd
nama = ['Berti', 'Qorygore', 'Bimo']
tugas = [90, 80, 70]
uts = [85, 95, 75]
uas = [95, 90, 80]
ratarata = [90.5, 86.6, 78.5]
df2 = pd.DataFrame({'Nama': nama, 'Tugas': tugas, 'UTS': uts, 'UAS': uas, 'Rata-Rata': ratarata})
df2
```

	Nama	Tugas	UTS	UAS	Rata-Rata
0	Berti	90	85	95	90.5
1	Qorygore	80	95	90	86.6
2	Bimo	70	75	80	78.5

#### Merge, Join, dan Concatenate DataFrame

##### a. Merge

##### 1. Siapkan 2 Data



```
# DATA PERTAMA
import pandas as pd
nama = ['Berti', 'Ryndes', 'Arin']
tugas = [95, 90, 75]
jurusan = ['IF', 'SI', 'KA']
df3 = pd.DataFrame({'Nama': nama, 'Tugas': tugas, 'Jurusa': jurusan})
df3
```

	Nama	Tugas	Jurusa
0	Berti	95	IF
1	Ryndes	90	SI
2	Arin	75	KA

The screenshot shows a Visual Studio Code window with a Jupyter Notebook. The Explorer sidebar on the left shows a project named 'PYTHON TASK 4' with files including 'iris\_unclean.csv', 'iris.csv', 'latihan\_unclean.py', 'latihan\_Visualization.py', 'latihan\_Olah\_Data\_Iris\_Dataset.py', and 'sries\_object.ipynb'. The active cell in the notebook is titled '# DATA KEDUA' and contains the following Python code:

```
# DATA KEDUA
import pandas as pd
nama = ['Berti', 'Ryndes', 'Rylo']
uts = [85, 84, 70]
jurusan = ['IF', 'SI', 'SI']
df4 = pd.DataFrame({'Nama':nama, 'UTS':uts, 'Jurusa':jurusan})
df4
```

The output of the cell is a DataFrame with 3 rows and 3 columns:

	Nama	UTS	Jurusa
0	Berti	85	IF
1	Ryndes	84	SI
2	Rylo	70	SI

The status bar at the bottom indicates 'Ln 5, Col 25', 'Spaces: 4', 'CRLF', 'Cell 9 of 9', and the system clock shows '24°C Berawan' and '10:18 AM'.

## 2. Inner Merge

The screenshot shows the same Visual Studio Code window, but the active cell in the notebook is titled '# INNER MERGE' and contains the following Python code:

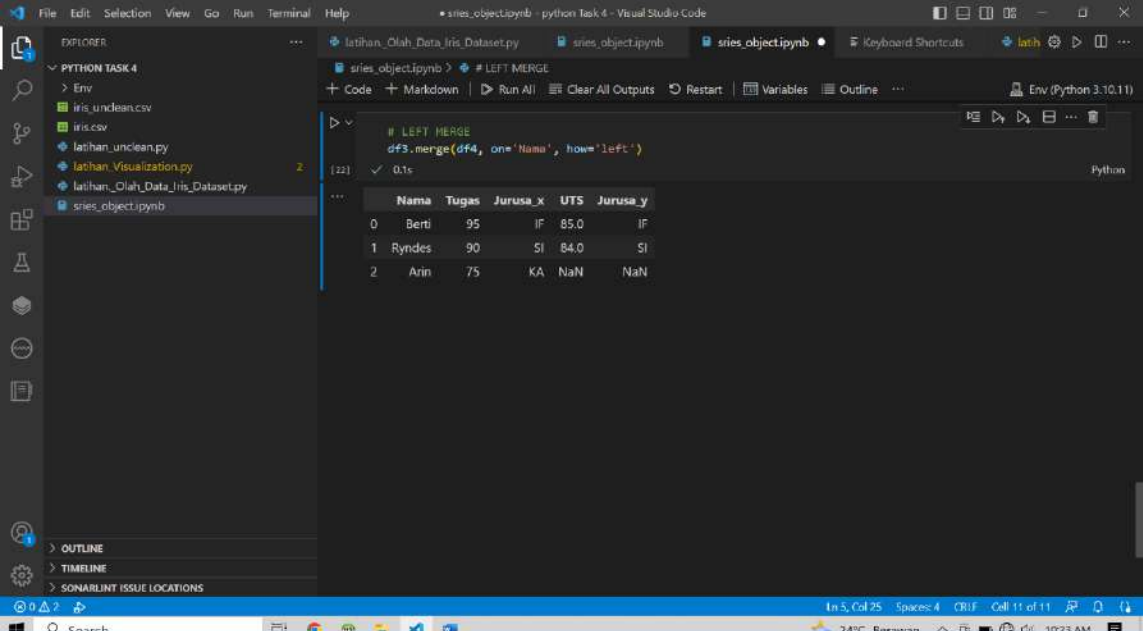
```
# INNER MERGE
df3.merge(df4)
```

The output of the cell is a DataFrame with 2 rows and 4 columns:

	Nama	Tugas	Jurusa	UTS
0	Berti	95	IF	85
1	Ryndes	90	SI	84

The status bar at the bottom indicates 'Ln 5, Col 25', 'Spaces: 4', 'CRLF', 'Cell 10 of 10', and the system clock shows '24°C Berawan' and '10:20 AM'.

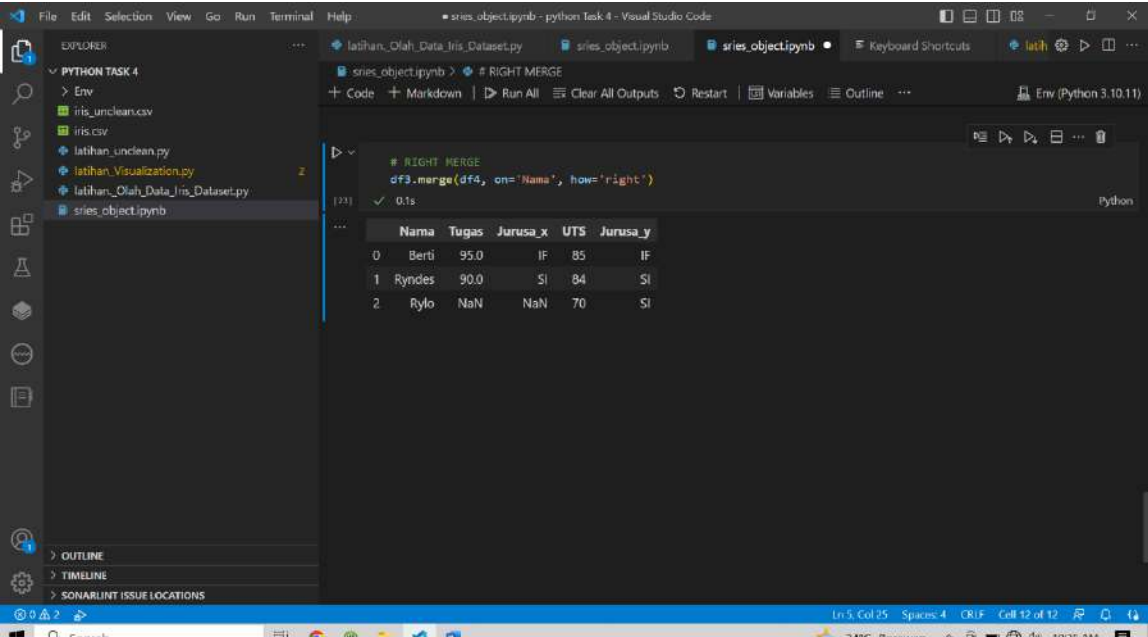
### 3. Left Merge



```
# LEFT MERGE
df3.merge(df4, on='Nama', how='left')
```

	Nama	Tugas	Jurusa_x	UTS	Jurusa_y
0	Berti	95	IF	85.0	IF
1	Ryndes	90	SI	84.0	SI
2	Arin	75	KA	NaN	NaN

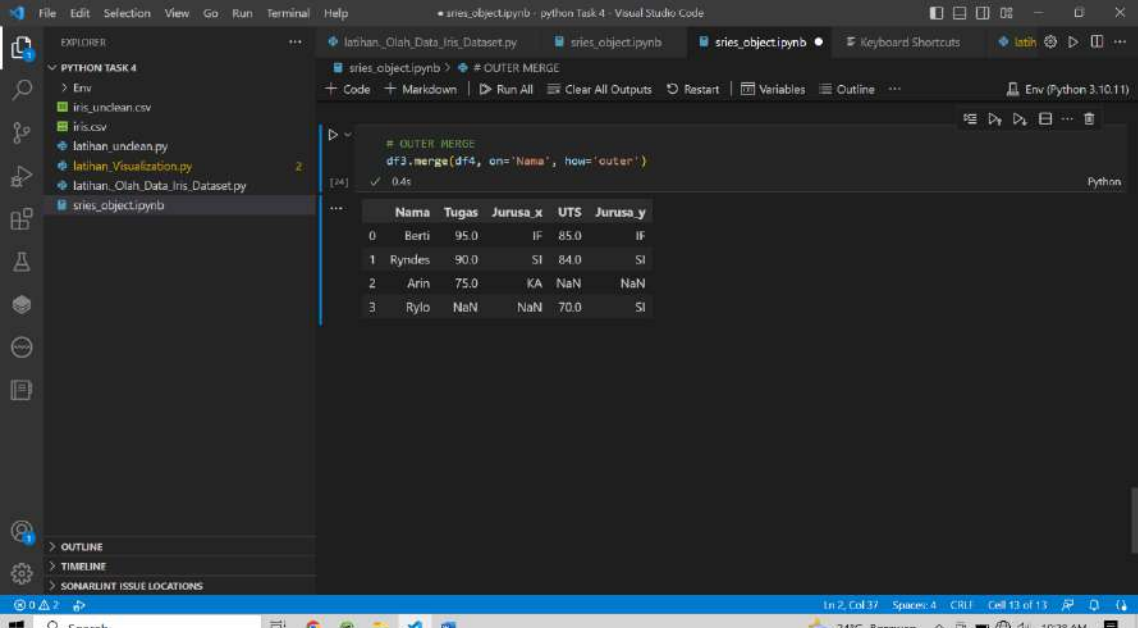
### 4. Right Merge



```
# RIGHT MERGE
df3.merge(df4, on='Nama', how='right')
```

	Nama	Tugas	Jurusa_x	UTS	Jurusa_y
0	Berti	95.0	IF	85	IF
1	Ryndes	90.0	SI	84	SI
2	Rylo	NaN	NaN	70	SI

## 5. Outer Merge



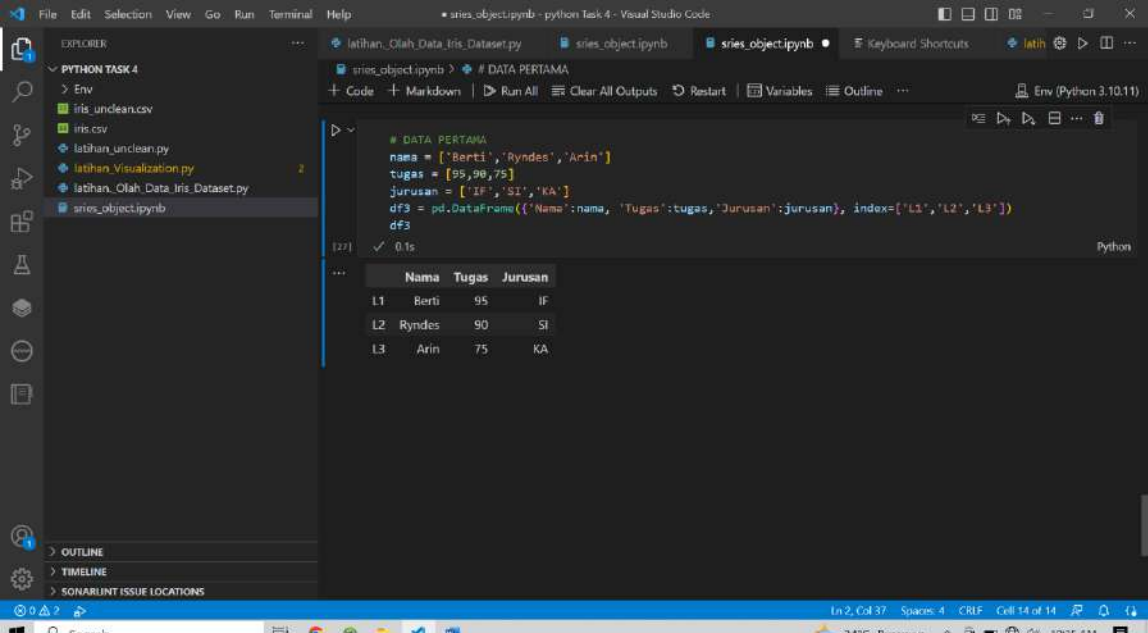
The screenshot shows a Jupyter Notebook cell in Visual Studio Code. The code performs an outer merge of two DataFrames, df3 and df4, on the 'Nama' column. The resulting DataFrame is displayed below the code.

```
# OUTER MERGE
df3.merge(df4, on='Nama', how='outer')
```

	Nama	Tugas	Jurusa_x	UTS	Jurusa_y
0	Berti	95.0	IF	85.0	IF
1	Ryndes	90.0	SI	84.0	SI
2	Arin	75.0	KA	NaN	NaN
3	Rylo	NaN	NaN	70.0	SI

### b. Join

#### 1. Siapkan 2 Data

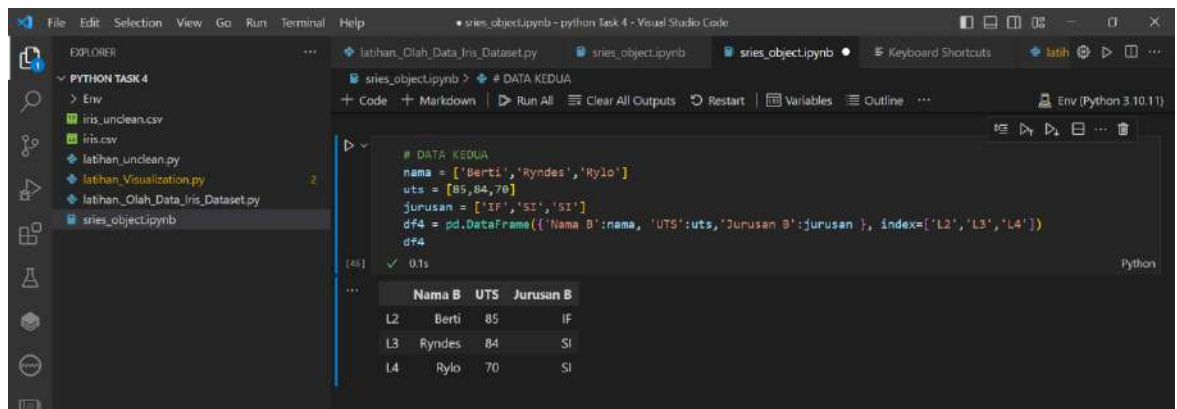


The screenshot shows a Jupyter Notebook cell in Visual Studio Code. The code creates two lists, 'nama' and 'tugas', and a list 'jurusan'. These are then used to create a DataFrame 'df3' with columns 'Nama', 'Tugas', and 'Jurusan'. The resulting DataFrame is displayed below the code.

```
# DATA PERTAMA
nama = ['Berti', 'Ryndes', 'Arin']
tugas = [95, 90, 75]
jurusan = ['IF', 'SI', 'KA']
df3 = pd.DataFrame({'Nama':nama, 'Tugas':tugas, 'Jurusan':jurusan}, index=['L1', 'L2', 'L3'])
df3
```

	Nama	Tugas	Jurusan
L1	Berti	95	IF
L2	Ryndes	90	SI
L3	Arin	75	KA





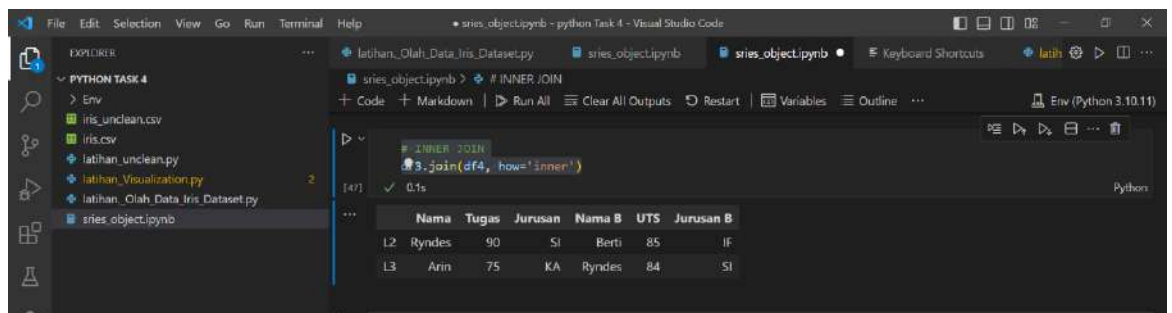
The screenshot shows a Visual Studio Code window with a Python Jupyter notebook. The Explorer sidebar on the left shows a project named 'PYTHON TASK 4' with files like 'iris\_unclean.csv', 'iris.csv', 'latihan\_unclean.py', 'latihan\_Visualization.py', 'latihan\_Olah Data Iris Dataset.py', and 'sries\_object.ipynb'. The active notebook is 'sries\_object.ipynb'. The code cell contains the following Python code:

```
# DATA KEDUA
nama = ['Berti', 'Ryndes', 'Rylo']
uts = [85, 84, 70]
jurusan = ['IF', 'SI', 'SI']
df4 = pd.DataFrame({'Nama B':nama, 'UTS':uts, 'Jurusan B':jurusan }, index=['L2', 'L3', 'L4'])
df4
```

The output of the code is a DataFrame with the following data:

	Nama B	UTS	Jurusan B
L2	Berti	85	IF
L3	Ryndes	84	SI
L4	Rylo	70	SI

## 2. Inner Join



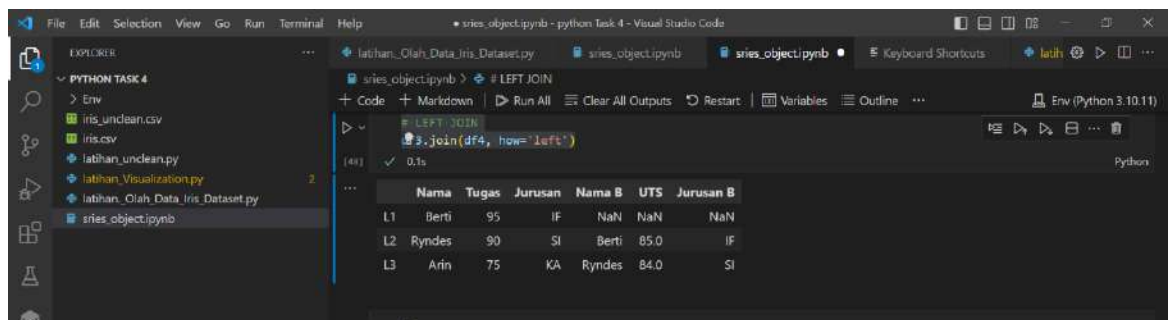
The screenshot shows a Visual Studio Code window with a Python Jupyter notebook. The Explorer sidebar on the left shows the same project as the first screenshot. The active notebook is 'sries\_object.ipynb'. The code cell contains the following Python code:

```
# INNER JOIN
df3.join(df4, how='inner')
```

The output of the code is a DataFrame with the following data:

	Nama	Tugas	Jurusan	Nama B	UTS	Jurusan B
L2	Ryndes	90	SI	Berti	85	IF
L3	Arin	75	KA	Ryndes	84	SI

## 3. Left join



The screenshot shows a Visual Studio Code window with a Python Jupyter notebook. The Explorer sidebar on the left shows the same project as the first screenshot. The active notebook is 'sries\_object.ipynb'. The code cell contains the following Python code:

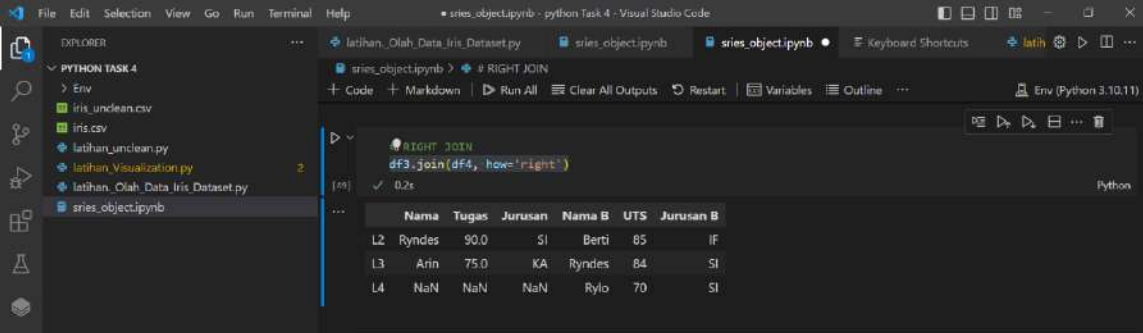
```
# LEFT JOIN
df3.join(df4, how='left')
```

The output of the code is a DataFrame with the following data:

	Nama	Tugas	Jurusan	Nama B	UTS	Jurusan B
L1	Berti	95	IF	NaN	NaN	NaN
L2	Ryndes	90	SI	Berti	85.0	IF
L3	Arin	75	KA	Ryndes	84.0	SI



#### 4. Right join

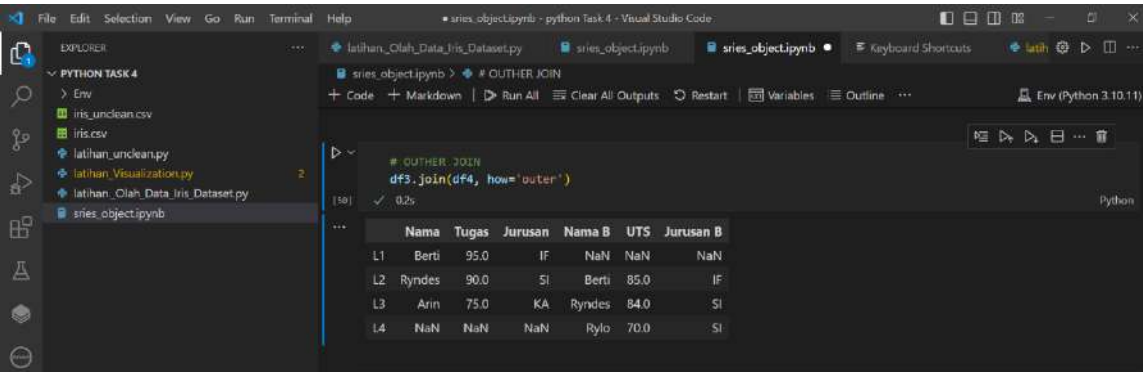


The screenshot shows a Jupyter Notebook cell in Visual Studio Code. The code performs a right join between two DataFrames, df3 and df4, using the 'right' how parameter. The resulting DataFrame is displayed below the code.

```
# RIGHT JOIN
df3.join(df4, how='right')
```

	Nama	Tugas	Jurusan	Nama B	UTS	Jurusan B
L2	Ryndes	90.0	SI	Berti	85	IF
L3	Arin	75.0	KA	Ryndes	84	SI
L4	NaN	NaN	NaN	Rylo	70	SI

#### 5. Outer Join

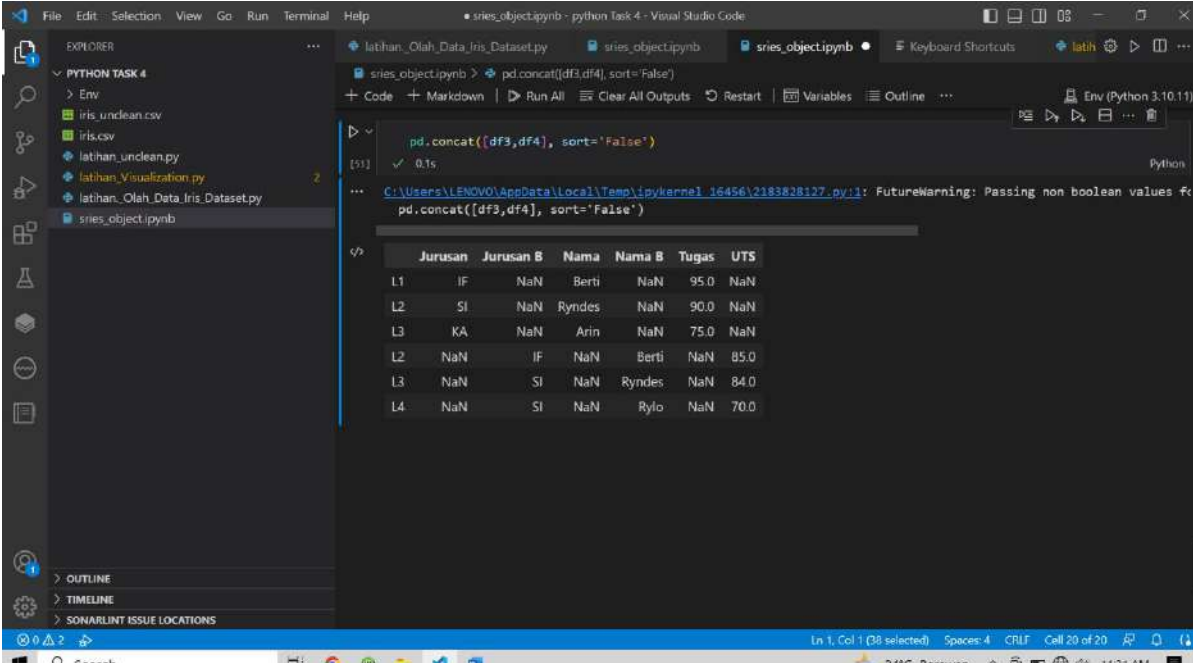


The screenshot shows a Jupyter Notebook cell in Visual Studio Code. The code performs an outer join between two DataFrames, df3 and df4, using the 'outer' how parameter. The resulting DataFrame is displayed below the code.

```
# OUTER JOIN
df3.join(df4, how='outer')
```

	Nama	Tugas	Jurusan	Nama B	UTS	Jurusan B
L1	Berti	95.0	IF	NaN	NaN	NaN
L2	Ryndes	90.0	SI	Berti	85.0	IF
L3	Arin	75.0	KA	Ryndes	84.0	SI
L4	NaN	NaN	NaN	Rylo	70.0	SI

#### c. Concatenate



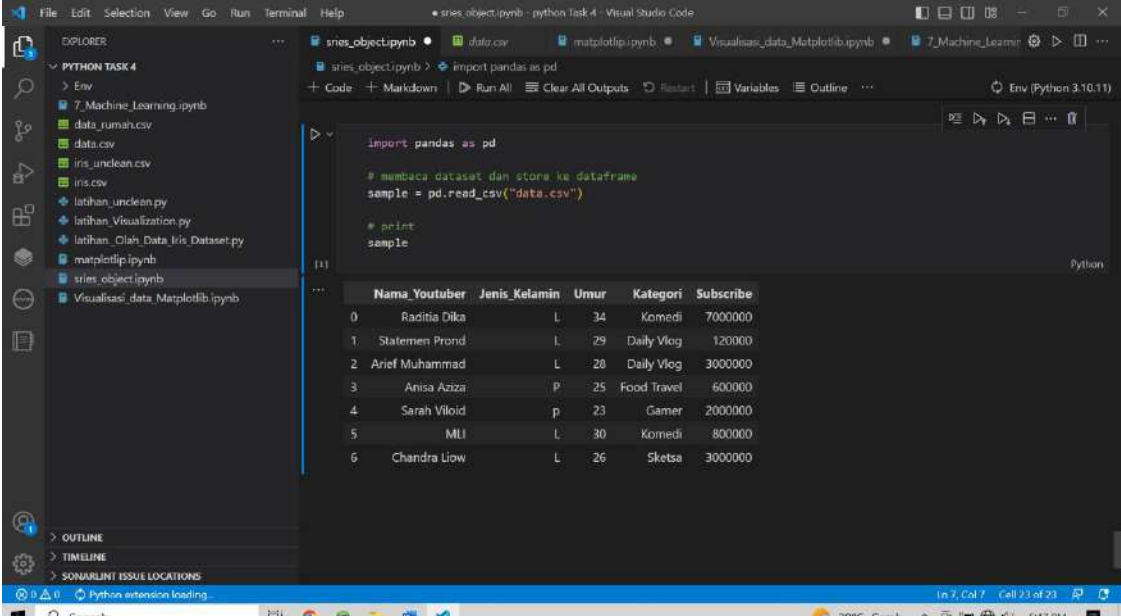
The screenshot shows a Jupyter Notebook cell in Visual Studio Code. The code concatenates two DataFrames, df3 and df4, using the pd.concat function with sort=False. The resulting DataFrame is displayed below the code.

```
pd.concat([df3, df4], sort=False)
```

	Jurusan	Jurusan B	Nama	Nama B	Tugas	UTS
L1	IF	NaN	Berti	NaN	95.0	NaN
L2	SI	NaN	Ryndes	NaN	90.0	NaN
L3	KA	NaN	Arin	NaN	75.0	NaN
L2	NaN	IF	NaN	Berti	NaN	85.0
L3	NaN	SI	NaN	Ryndes	NaN	84.0
L4	NaN	SI	NaN	Rylo	NaN	70.0

## d. Pandas DataFrame – Import Data CSV

### 1. Import Data CSV



```
import pandas as pd

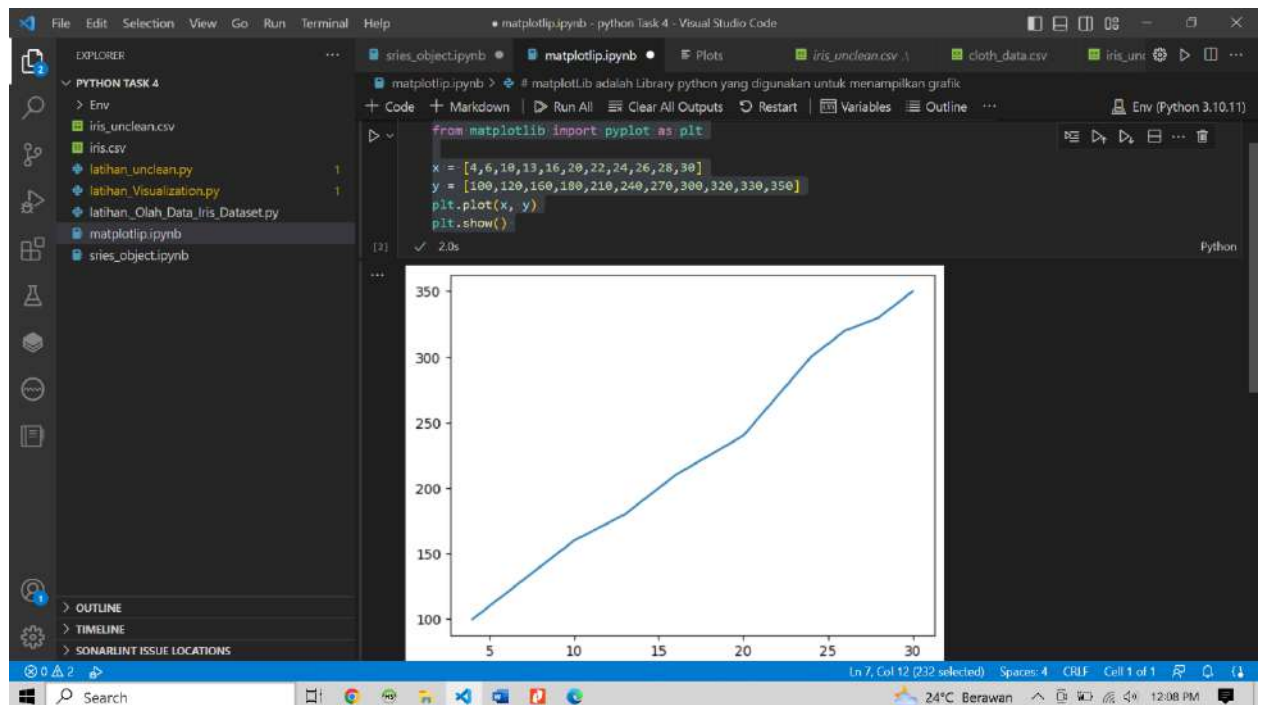
# membaca dataset dan store ke dataframe
sample = pd.read_csv("data.csv")

# print
sample
```

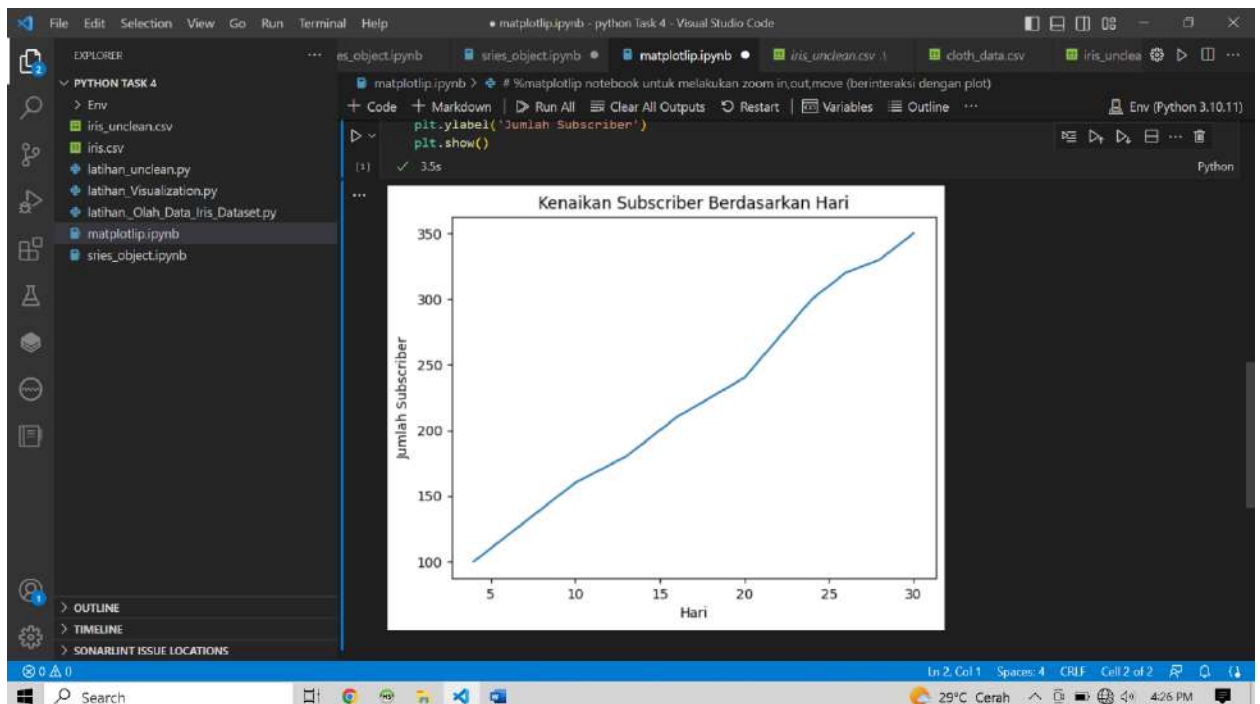
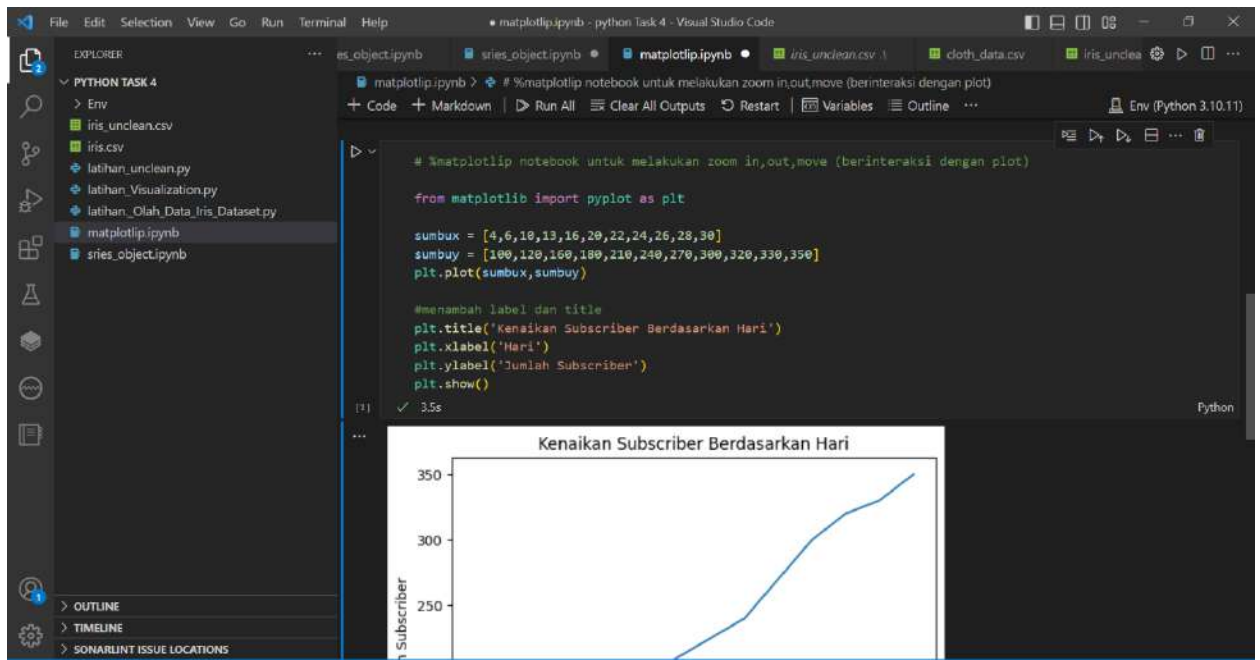
	Nama_Youtuber	Jenis_Kelamin	Umur	Kategori	Subscribe
0	Raditia Dika	L	34	Komedi	7000000
1	Statemen Prond	L	29	Daily Vlog	120000
2	Arief Muhammad	L	28	Daily Vlog	3000000
3	Anisa Aziza	P	25	Food Travel	600000
4	Sarah Viloid	p	23	Gamer	2000000
5	MLI	L	30	Komedi	800000
6	Chandra Liow	L	26	Sketsa	3000000

## 5-Visualisasi Data-Matplotlib

### 1. Plot Garis

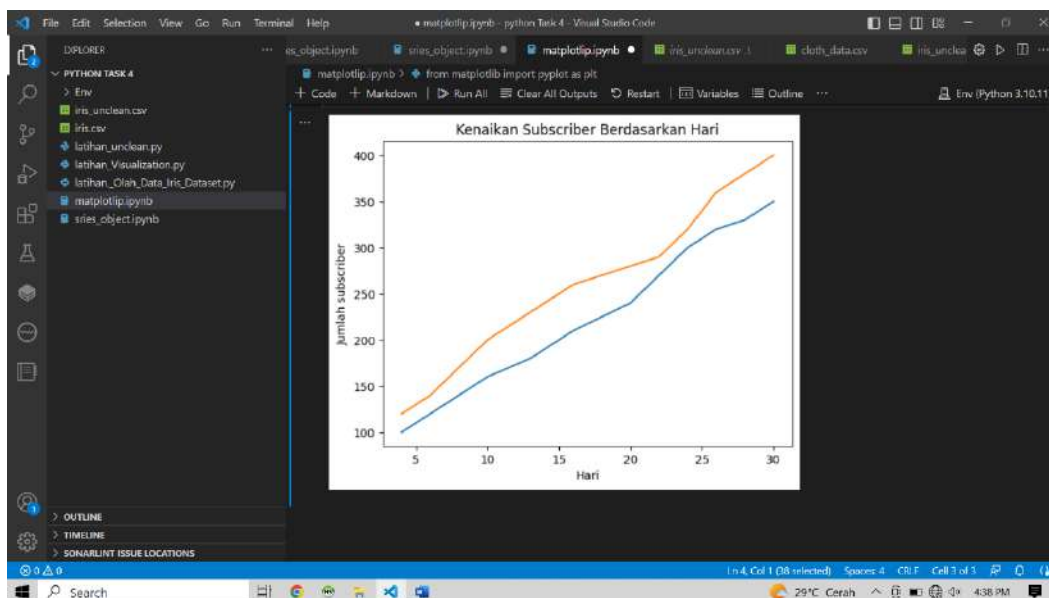
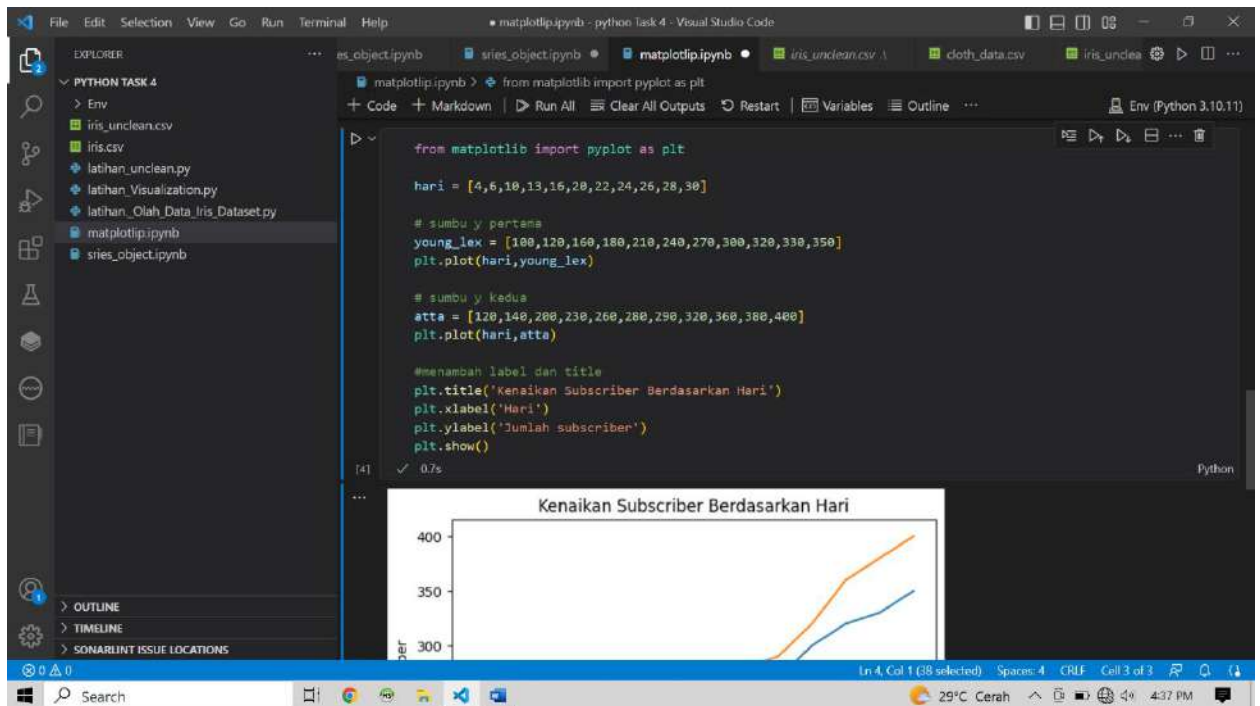


## 2. Menambah Label dan Title



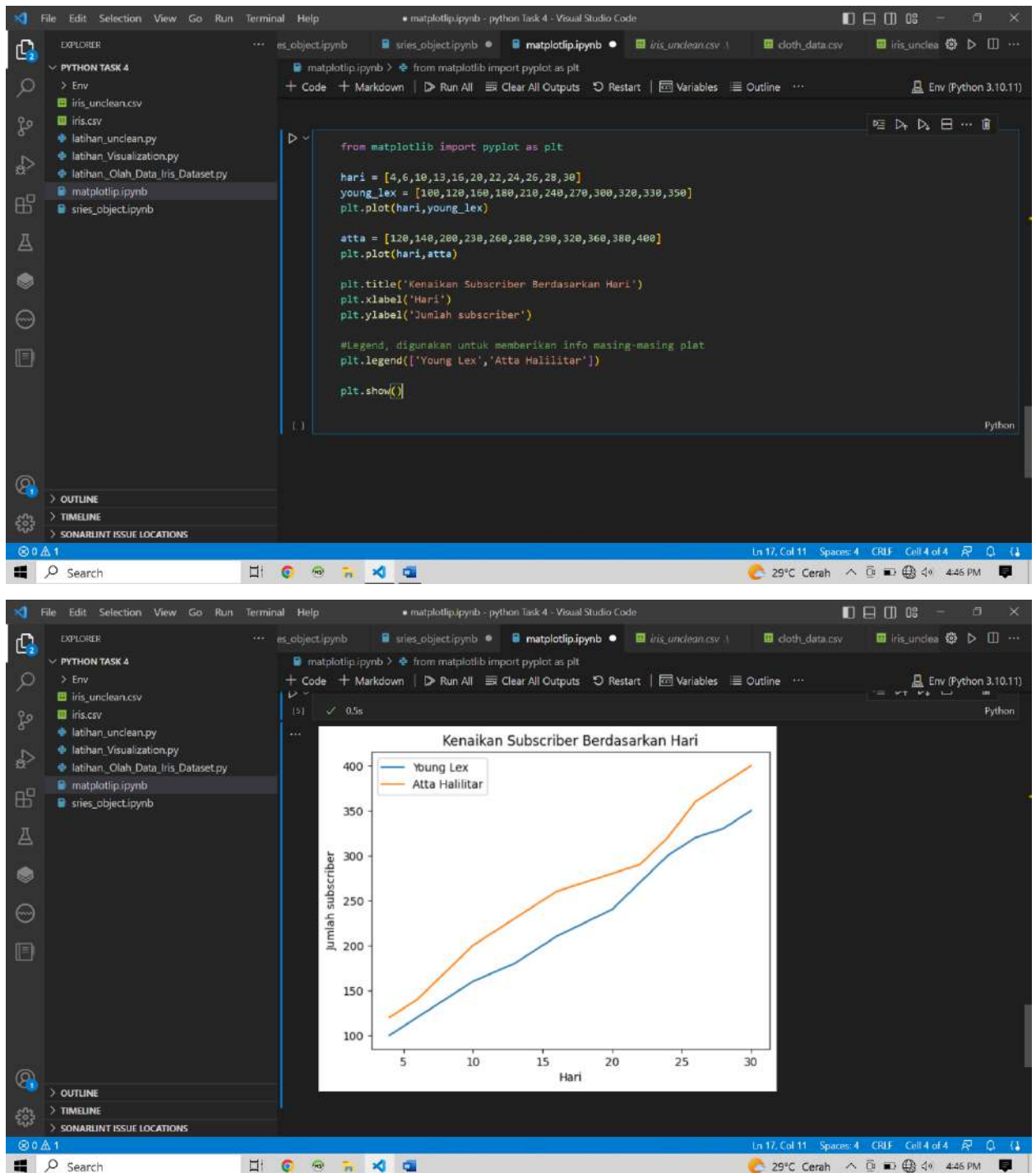
Gambar. Label dan Title

### 3. Plot Garis Multiline



Gambar. Plot Garis multiline

#### 4. Legend



Gambar. Legend



## 5. Color, Linestyle, Marker

```
File Edit Selection View Go Run Terminal Help
matplotlib.ipynb - python Task 4 - Visual Studio Code

EXPLORER
PYTHON TASK 4
  Env
  iris_unclean.csv
  iris.csv
  latihan_unclean.py
  latihan_Visualization.py
  latihan_Olah_Data_Iris_Dataset.py
  matplotlib.ipynb
  sries_object.ipynb

OUTLINE
TIMELINE
SONARLINT ISSUE LOCATIONS

matplotlib.ipynb > from matplotlib import pyplot as plt
+ Code + Markdown | Run All | Clear All Outputs | Restart | Variables | Outline ...
Env (Python 3.10.11)

from matplotlib import pyplot as plt

hari = [4,6,10,13,16,20,22,24,26,28,30]
young_lex = [100,120,160,180,210,240,270,300,320,330,350]

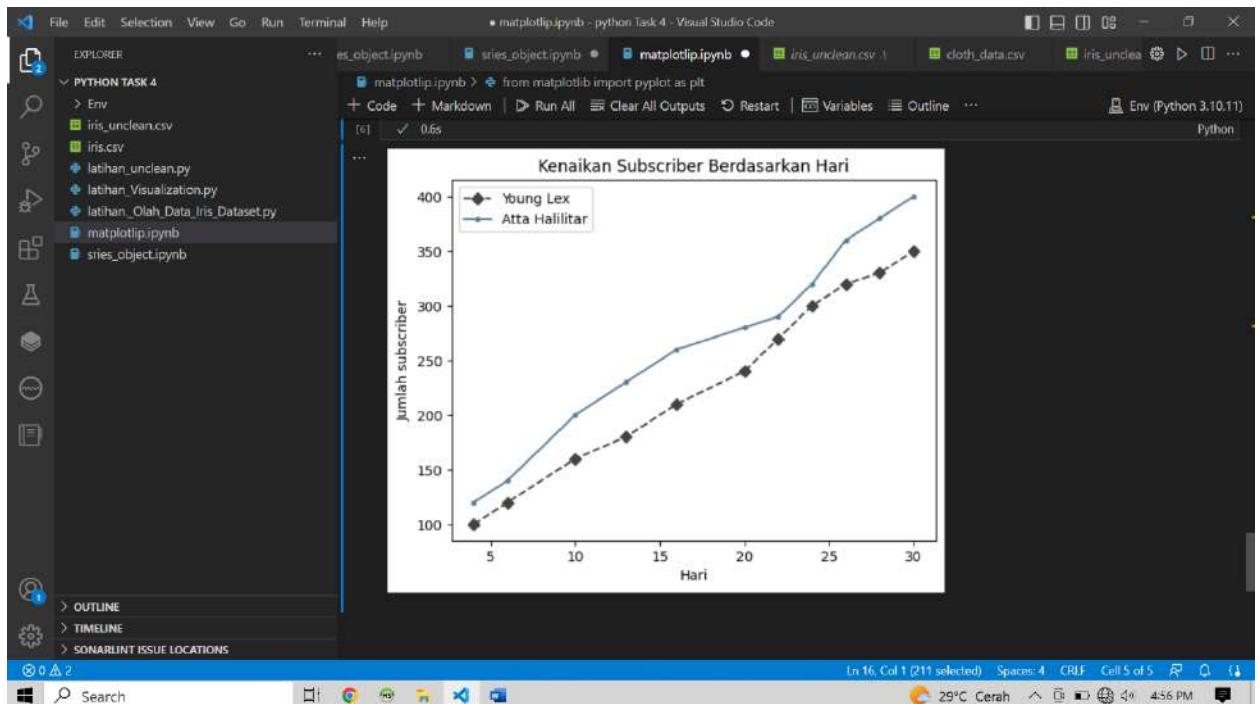
# color, untuk warna garis
# linestyle, untuk jenis garis
# --
# --
# marker, untuk penanda pertemuan sumbu x dan y
# p, pentagon
# D, diamond
plt.plot(hari, young_lex, color='044444', linestyle='--', marker='D')

atta = [120,140,200,230,260,280,290,320,360,380,400]
plt.plot(hari, atta, color='5A7D9A', markers='.')

plt.title('Kenaikan Subscriber Berdasarkan Hari')
plt.xlabel('Hari')
plt.ylabel('Jumlah subscriber')

plt.legend(['Young Lex', 'Atta Halililar'])

plt.show()
```



Gambar. Color, Linestyle, Marker



## 6. Linewidth

```
File Edit Selection View Go Run Terminal Help • matplotlib.ipynb - python Task 4 - Visual Studio Code

EXPLORER
PYTHON TASK 4
  Env
  7_Machine_Learning.ipynb
  data_rumah.csv
  data.csv
  iris_unclean.csv
  iris.csv
  latihan_unclean.py
  latihan_Visualization.py
  latihan_Olah_Data_Iris_Dataset.py
  matplotlib.ipynb
  sries_object.ipynb
  Visualisasi_data_Matplotlib.ipynb

OUTLINE
TIMELINE
SONARLINT ISSUE LOCATIONS

matplotlib.ipynb > from matplotlib import pyplot as plt
+ Code + Markdown | Run All | Clear All Outputs | Restart | Variables | Outline ...
Env (Python 3.10.11)

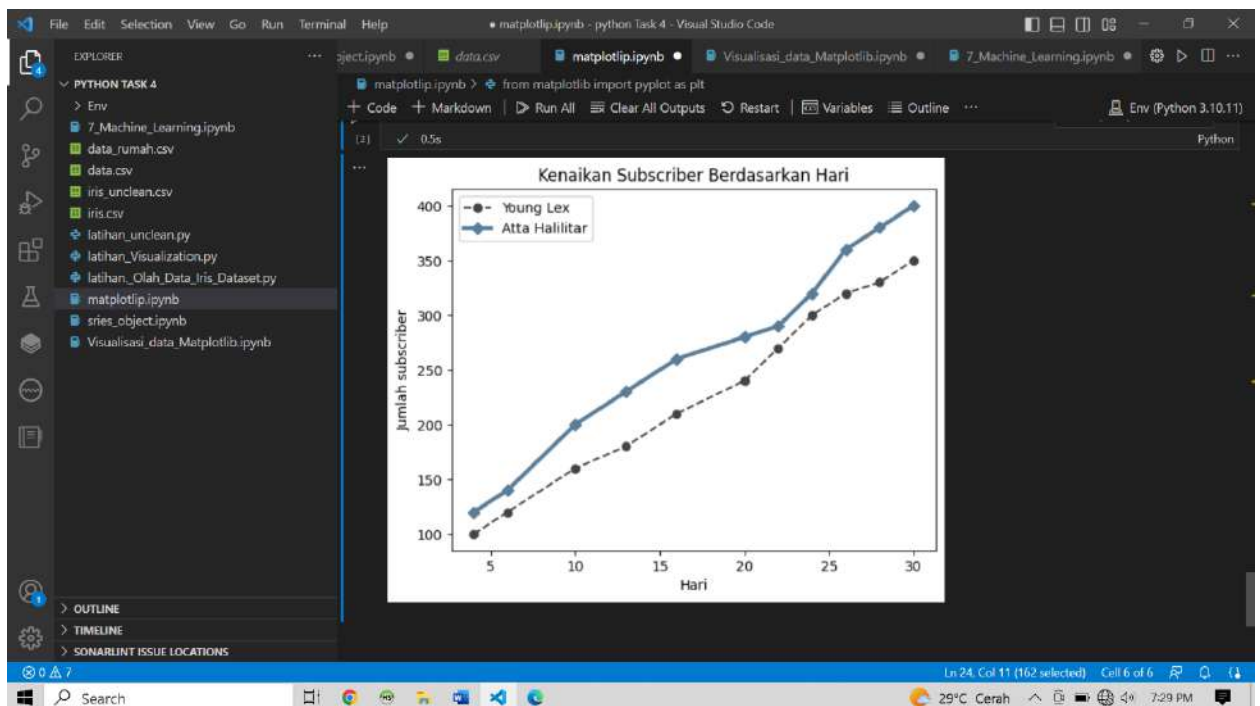
from matplotlib import pyplot as plt

hari = [4,6,10,13,16,20,22,24,26,28,30]
young_lex = [100,120,160,180,210,240,270,300,320,330,350]
plt.plot(hari,young_lex, color='#444444', linestyle='--', marker='o')

# linewidth, untuk menebalkan garis
atta = [120,140,200,230,260,280,290,320,360,380,400]
plt.plot(hari,atta,color='#5A7D9A', marker='D', linewidth=3)
plt.title('Kenaikan Subscriber Berdasarkan Hari')
plt.xlabel('Hari')
plt.ylabel('Jumlah subscriber')

plt.legend(['Young Lex','Atta Halililar'])

plt.show()
```



Gambar Linewidth

## 7. Grid

```
from matplotlib import pyplot as plt

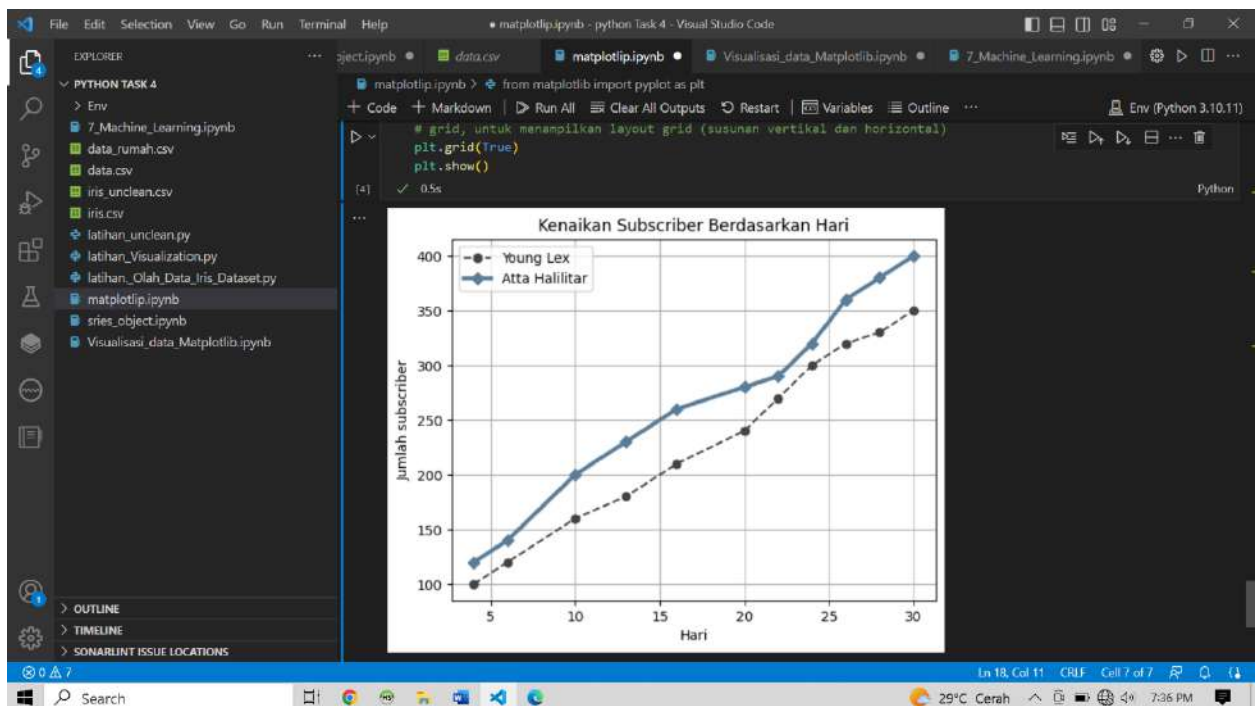
hari = [4,6,10,13,16,20,22,24,26,28,30]
young_lex = [100,120,160,180,210,240,270,300,320,330,350]
plt.plot(hari,young_lex, color='d', linestyle='--', marker='o')

atta = [120,140,200,230,260,280,290,320,360,380,400]
plt.plot(hari,atta,color='b', marker='D', linewidth='3')

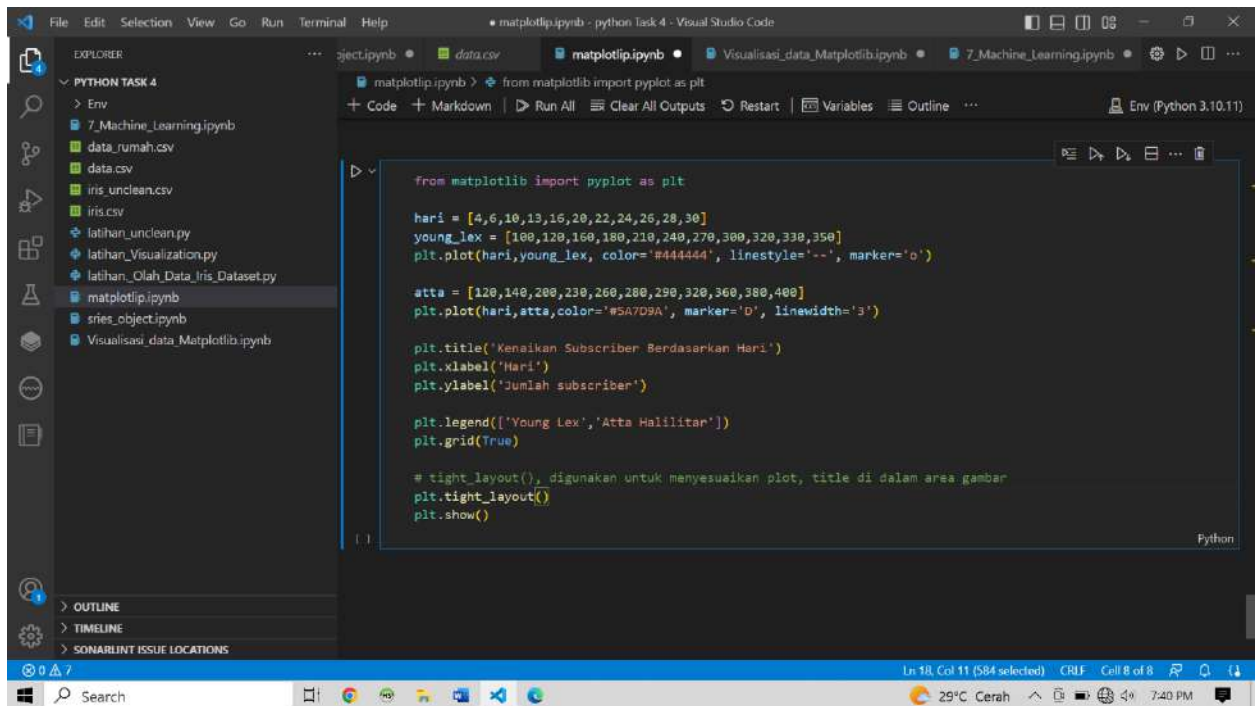
plt.title('Kenaikan Subscriber Berdasarkan Hari')
plt.xlabel('Hari')
plt.ylabel('Jumlah subscriber')

plt.legend(['Young Lex','Atta Halilitar'])

# grid, untuk menampilkan layout grid (susunan vertikal dan horizontal)
plt.grid(True)
plt.show()
```



## 8. Menggunakan Style (fivethirtyeight) dan Tight Layout



```
from matplotlib import pyplot as plt

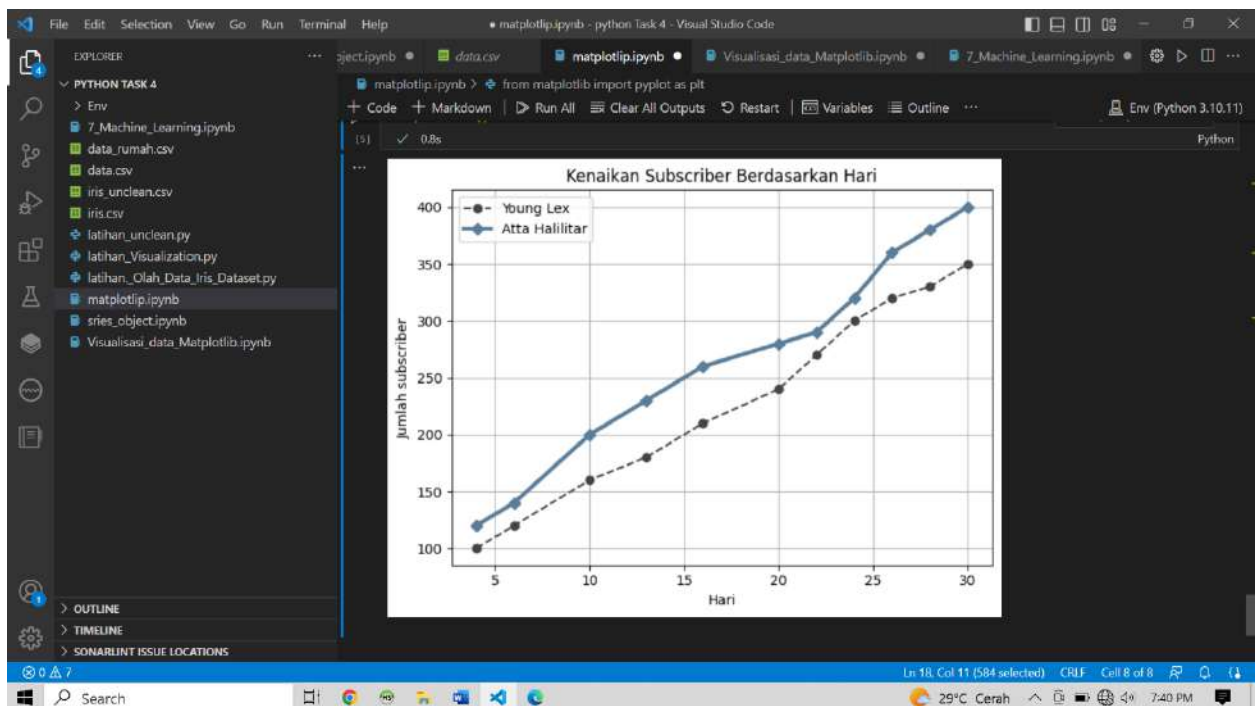
hari = [4,6,10,13,16,20,22,24,26,28,30]
young_lex = [100,120,160,180,210,240,270,300,320,330,350]
plt.plot(hari,young_lex, color='#444444', linestyle='--', marker='o')

atta = [120,140,200,230,260,280,290,320,360,380,400]
plt.plot(hari,atta,color='#5A7D9A', marker='D', linewidth='3')

plt.title('Kenaikan Subscriber Berdasarkan Hari')
plt.xlabel('Hari')
plt.ylabel('Jumlah subscriber')

plt.legend(['Young Lex','Atta Halilitar'])
plt.grid(True)

# tight_layout(), digunakan untuk menyesuaikan plot, title di dalam area gambar
plt.tight_layout()
plt.show()
```



Gambar. Style dan Tight Layout

## 9. Menggunakan Style ggplot

```
from matplotlib import pyplot as plt

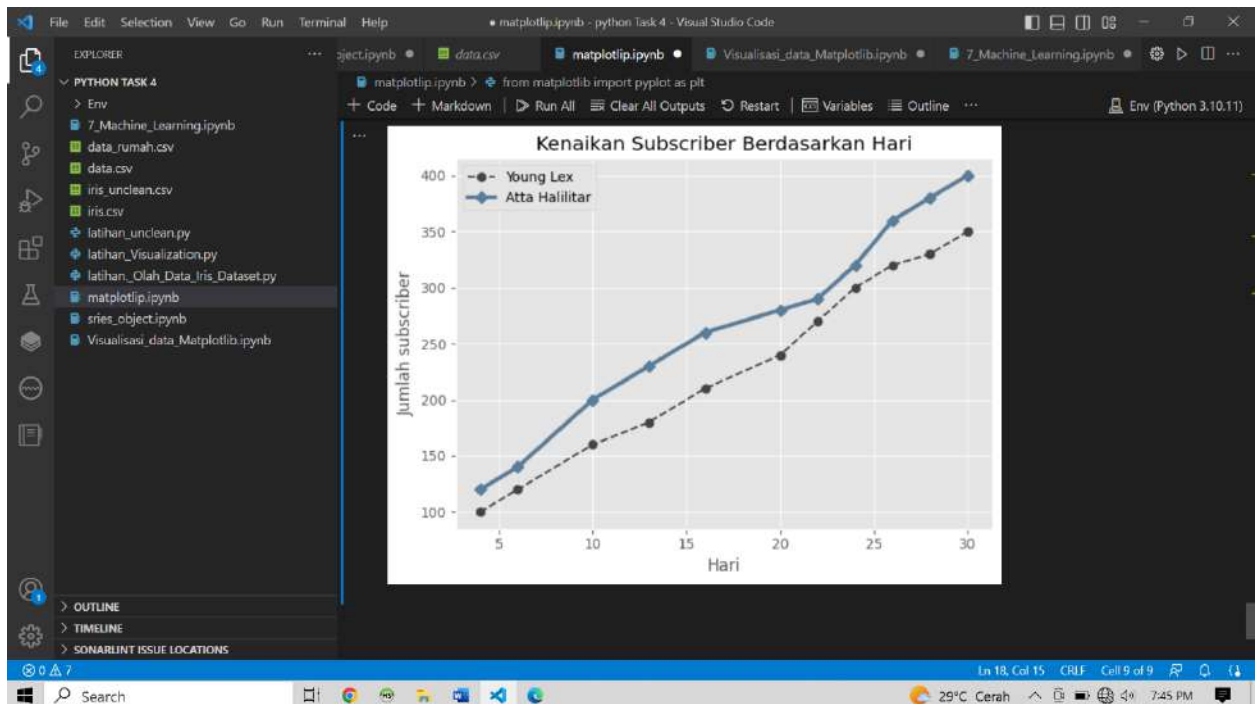
# menggunakan style ggplot
plt.style.use('ggplot')

hari = [4,6,10,13,16,20,22,24,26,28,30]
young_lex = [100,120,160,180,210,240,270,300,320,330,350]
plt.plot(hari,young_lex, color='#444444', linestyle='--', marker='o')

atta = [120,140,200,230,260,280,290,320,360,380,400]
plt.plot(hari,atta,color='#5A7D9A', marker='D', linewidth='3')

plt.title('Kenaikan Subscriber Berdasarkan Hari')
plt.xlabel('Hari')
plt.ylabel('Jumlah subscriber')

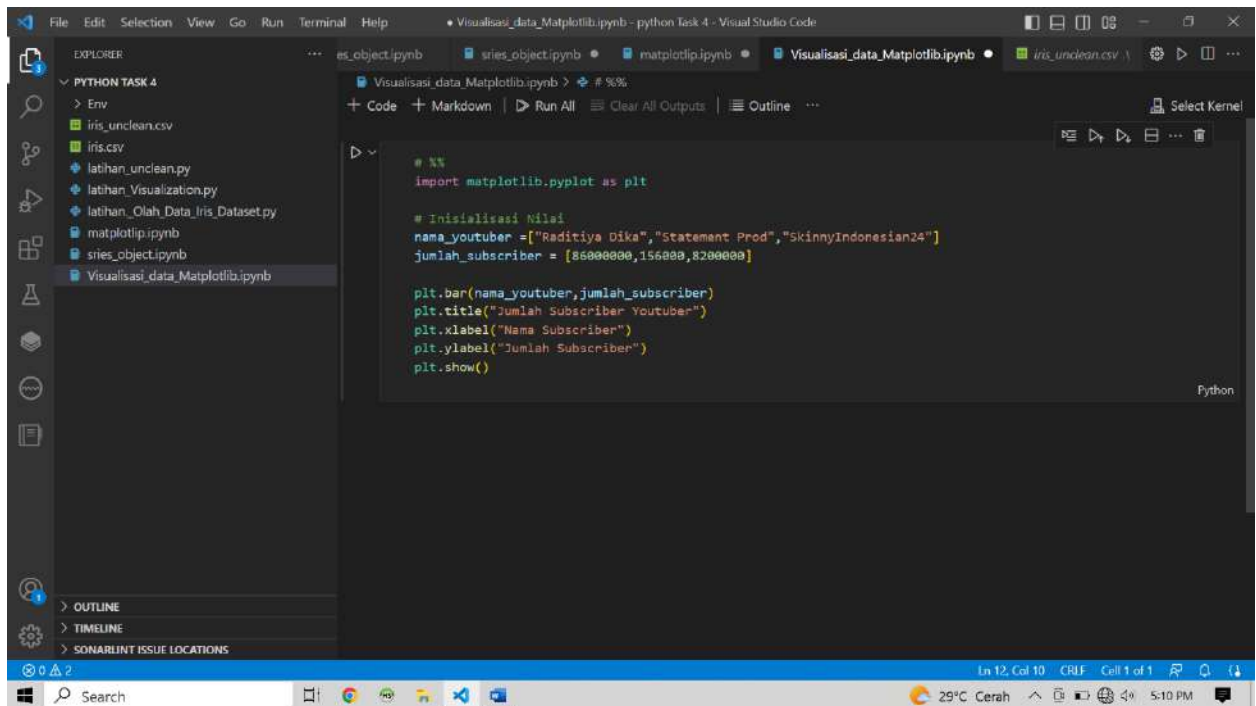
plt.legend(['Young Lex','Atta Halilitar'])
plt.grid(True)
plt.tight_layout()
plt.show()
```



Gambar. Style ggplot

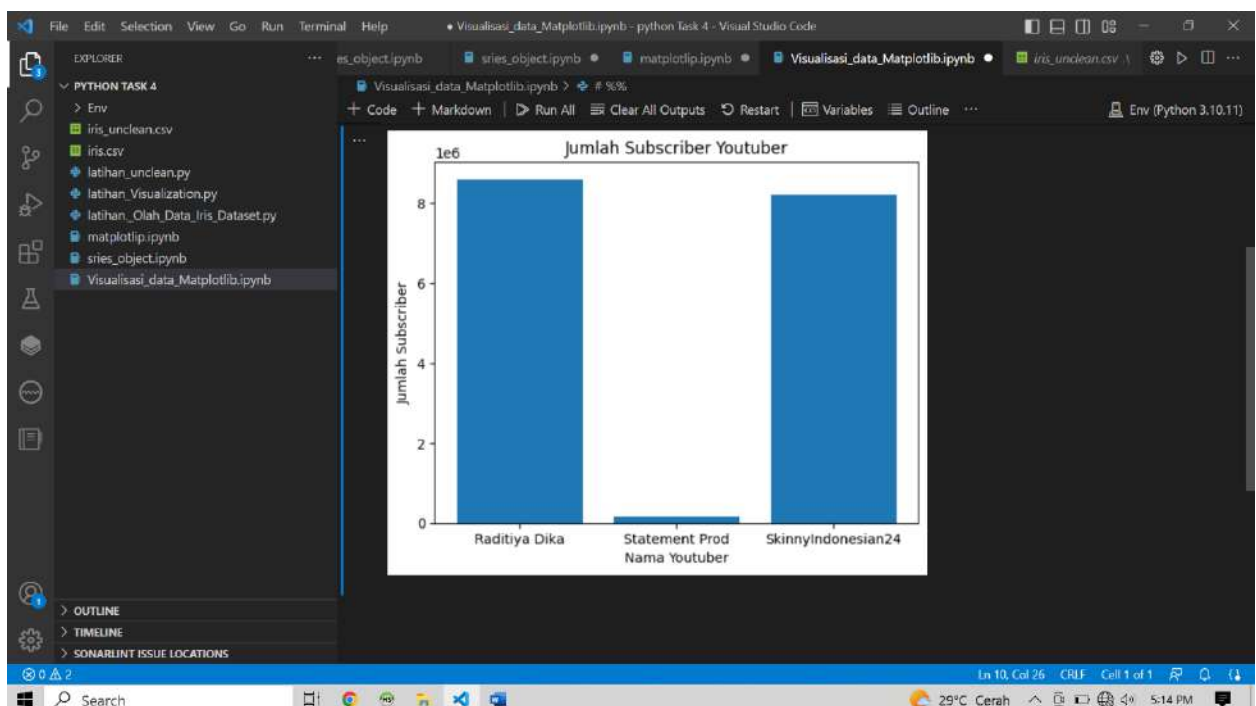
## 6-Visualisasi Data\_Matplotlib Import CSV (contoh 2)

### 1. Contoh Plot Dasar



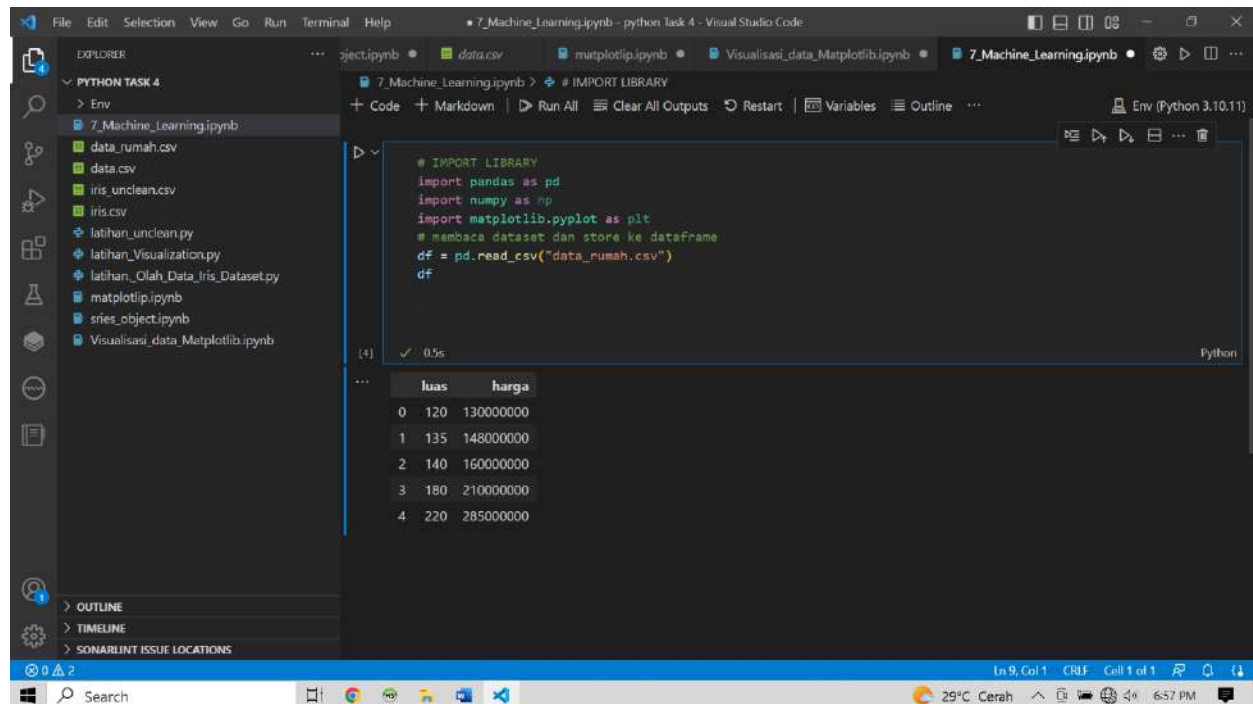
The screenshot shows the Visual Studio Code interface with a Python file named 'Visualisasi\_data\_Matplotlib.ipynb'. The code in the file is as follows:

```
# %%  
import matplotlib.pyplot as plt  
  
# Inisialisasi Nilai  
nama_youtuber = ["Raditiya Dika", "Statement Prod", "SkinnyIndonesian24"]  
jumlah_subscriber = [86000000, 156000, 82000000]  
  
plt.bar(nama_youtuber, jumlah_subscriber)  
plt.title("Jumlah Subscriber Youtuber")  
plt.xlabel("Nama Subscriber")  
plt.ylabel("Jumlah Subscriber")  
plt.show()
```





## 7-Machine Learning \_Prediksi Harga Rumah dengan Regresi Linier (Satu Variable)



Visual Studio Code interface showing the first cell of a Jupyter notebook. The code imports pandas, numpy, and matplotlib, and reads a CSV file named 'data\_rumah.csv' into a DataFrame. The output shows a table with columns 'luas' and 'harga'.

```
# IMPORT LIBRARY
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
# membaca dataset dan store ke dataframe
df = pd.read_csv("data_rumah.csv")
df
```

	luas	harga
0	120	130000000
1	135	148000000
2	140	160000000
3	180	210000000
4	220	285000000

