

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4
5 dataset=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science
6
7
8
9
```

temp.py x untitled0.py*

Name	Type	Size	Value
dataset	DataFrame	(30, 2)	Column names: YearsExperience, Salary

dataset - DataFrame

Index	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2	43525
4	2.2	39891
5	2.9	56642
6	3	60150
7	3.2	54445
8	3.2	64445

Format Resize Background color Column min/max Save and Close Close

IPython Console History

conda: base (Python 3.10.9) Completions: conda LSP: Python Line 6, Col 1 UTF-8 CRLF RW Mem 87

Type here to search

20°C 22:24 02-11-2023

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4
5 dataset=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Scienc
6
7 X = dataset.iloc[:, :-1].values
8 y = dataset.iloc[:,1].values
9
10
11
```

temp.py X untitled0.py X

Name	Type	Size	Value
dataset	DataFrame	(30, 2)	Column names: YearsExperience, Salary
X	Array of float64	(30, 1)	[[1.1] [1.3]
y	Array of int64	(30,)	[39343 46205 37731 ... 112635 122391 121872]

y - NumPy object array

	0
0	39343
1	46205
2	37731
3	43525
4	39891
5	56642
6	60150
7	54445
8	64445

X - NumPy object array

	0
0	1.1
1	1.3
2	1.5
3	2
4	2.2
5	2.9
6	3
7	3.2
8	3.2

Format Resize Background color Save and Close Close

Python Console History

base (Python 3.10.9) Completions: conda LSP: Python Line 9, Col 1 UTF-8 CRLF RW Mem 83%

22:25

Python 3.10

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4
5 dataset=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Scienc
6
7 X = dataset.iloc[:, :-1].values
8 y = dataset.iloc[:,1].values
9
10 # split the dataset to 80-20%
11 from sklearn.model_selection
12
13 X_train,X_test,y_train,y_te
14
```

temp.py X untitled0.py X

X_train - NumPy object array

	0
15	6
16	3.7
17	3.2
18	9
19	2
20	1.1
21	7.1
22	4.9
23	4

X_test - NumPy object array

	0
0	1.5
1	10.3
2	4.1
3	3.9
4	9.5
5	8.7

y_test - NumPy object array

	0
0	37731
1	122391
2	57081
3	63218
4	116969
5	109431

y_train - NumPy object array

	0
13	39891
14	81363
15	93940
16	57189
17	54445
18	105582
19	43525
20	39343
21	98273
22	67938
23	66057

dataset DataFrame (30, 2) Column names: YearsExperience, Salary

X Array of float64 (30, 1)

X_test Array of float64 (6, 1)

X_train Array of float64

y Array of int64

y_test Array of int64

y_train Array of int64

Format Resize Background ve and Clc Close

Completions: conda LSP: Python Line 14, Col 1 UTF-8 CRLF RW Mem 83

Type here to search

20°C

22:29

The screenshot displays a Python IDE with the following code in the editor:

```

1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4
5 dataset = pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Scienc
6
7 X = dataset.iloc[:, :-1].values
8 y = dataset.iloc[:, 1].values
9
10 # split the dataset to 80-20%
11 from sklearn.model_selection import train_test_split
12
13 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, rand
14
15 from sklearn.linear_model import LinearRegression
16 regressor = LinearRegression()
17 regressor.fit(X_train, y_train)
18
19 y_pred = regressor.predict(X_test)
20
21

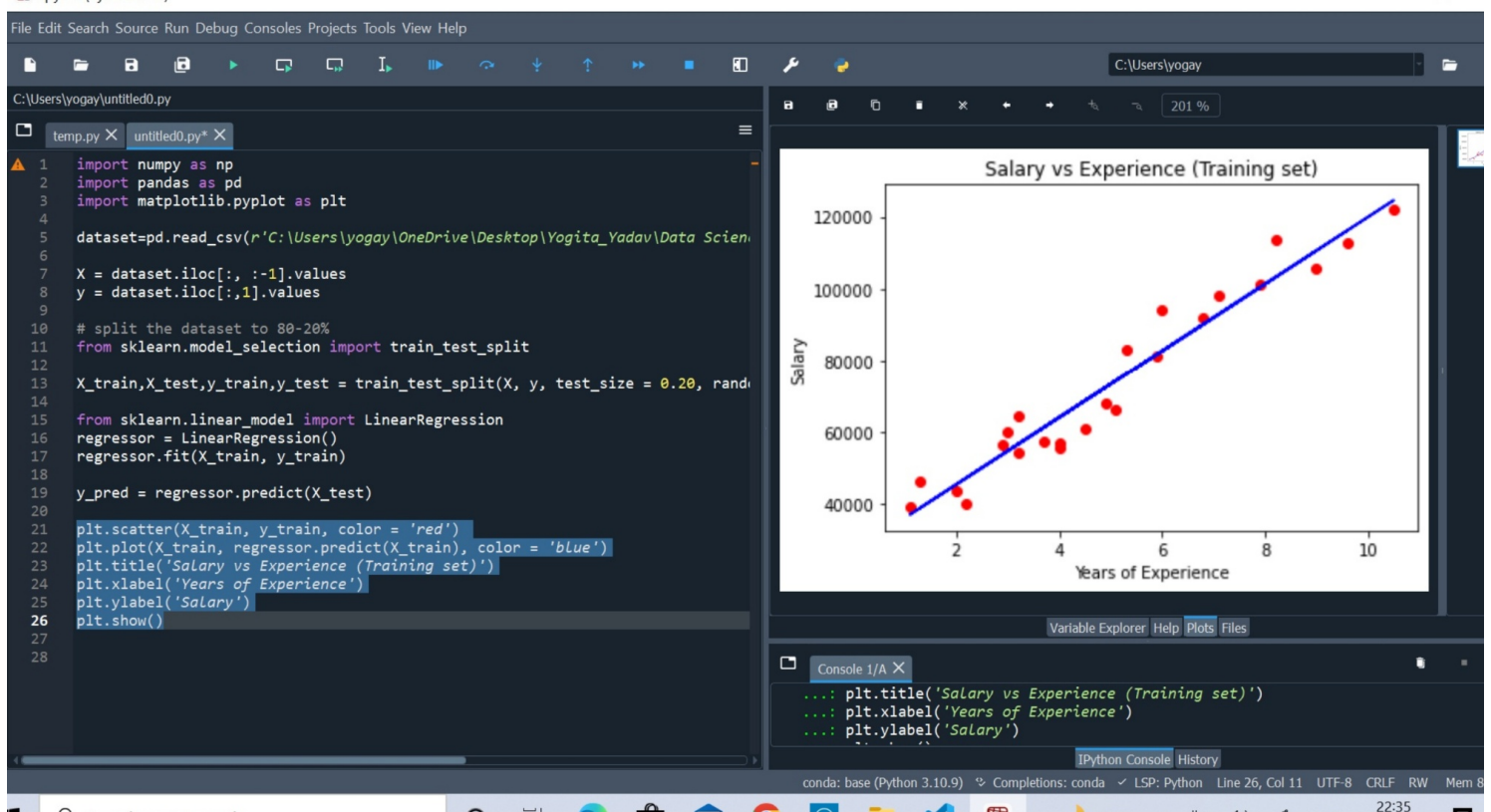
```

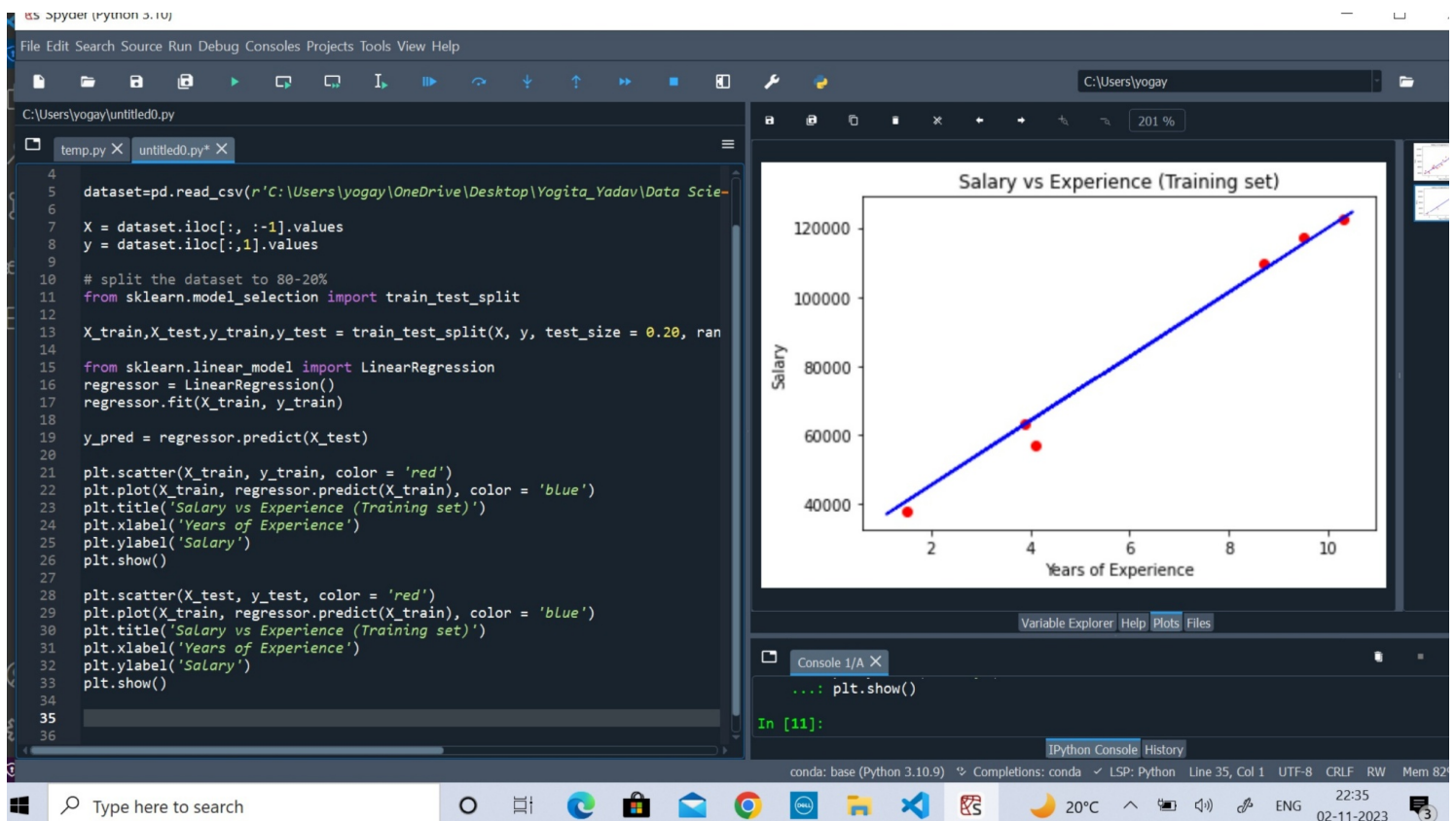
Two variable inspectors are open:

- X_test - NumPy object array:** Shows a 1D array of floats with 6 elements: [1.5, 10.3, 4.1, 3.9, 9.5, 8.7].
- y_pred - NumPy object array:** Shows a 1D array of floats with 6 elements: [40749, 122700, 64961.7, 63099.1, 115250, 107800].

The background variable inspector shows the state of the entire script:

Name	Type	Size	Value
regressor	linear_model_base.LinearRegression	1	LinearRegression object of sklearn.linear_model_base module
X	Array of float64	(30, 1)	[[1.1] [1.3] [1.5] [10.3] [9.6] [4.]
X_test	Array of float64	(6, 1)	[[1.1] [1.3] [1.5] [10.3] [9.6] [4.]
X_train	Array of float64	(24, 1)	[[1.1] [1.3] [1.5] [10.3] [9.6] [4.]
y	Array of int64	(30,)	[39343 46205 37731 ... 112635 122391 121872] [40748.96184072 122699.62295594 6 1152 ...
y_pred	Array of float64	(6,)	[39343 46205 37731 ... 112635 122391 121872] [40748.96184072 122699.62295594 6 1152 ...
y_test	Array of int64	(6,)	[39343 46205 37731 ... 112635 122391 121872] [40748.96184072 122699.62295594 6 1152 ...





The screenshot displays a Jupyter Notebook environment with the following components:

- Code Editor:** Contains Python code for loading a dataset, splitting it into training and testing sets, fitting a linear regression model, and plotting the results. The code is as follows:


```

4 dataset=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Scie-
5
6 X = dataset.iloc[:, :-1].values
7 y = dataset.iloc[:,1].values
8
9 # split the dataset to 80-20%
10 from sklearn.model_selection import train_test_split
11
12 X_train,X_test,y_train,y_test = train_test_split(X,y,train_size=0.8,random_state=42)
13
14 from sklearn.linear_model import LinearRegression
15 regressor = LinearRegression()
16 regressor.fit(X_train, y_train)
17
18 y_pred = regressor.predict(X_test)
19
20 plt.scatter(X_train, y_train, color='blue')
21 plt.plot(X_train, regressor.predict(X_train), color='red')
22 plt.title('Salary vs Experience (Training set)')
23 plt.xlabel('Years of Experience')
24 plt.ylabel('Salary')
25 plt.show()
26
27 plt.scatter(X_test, y_test, color='green')
28 plt.plot(X_train, regressor.predict(X_train), color='red')
29 plt.title('Salary vs Experience (Test set)')
30 plt.xlabel('Years of Experience')
31 plt.ylabel('Salary')
32 plt.show()
33
34 m = regressor.coef_[0] #slope
35
36 
```
- Variable Inspector:** Shows the 'dataset' variable as a DataFrame with 30 rows and 2 columns. The 'Value' column lists the column names: 'YearsExperience' and 'Salary'.
- IPython Console:** Displays the output of the code execution:


```

In [12]: m
Out[12]: array([9312.57512673])

In [13]:

```
- Figure:** A scatter plot titled 'Salary vs Experience (Training set)' showing a positive correlation between Years of Experience (X-axis) and Salary (Y-axis). The plot includes a red line representing the linear regression fit.

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```

6 X = dataset.iloc[:, :-1].values
7 y = dataset.iloc[:, 1].values
8
9
10 # split the dataset to 80-20%
11 from sklearn.model_selection import train_test_split
12
13 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state = 0)
14
15 from sklearn.linear_model import LinearRegression
16 regressor = LinearRegression()
17 regressor.fit(X_train, y_train)
18
19 y_pred = regressor.predict(X_test)
20
21 plt.scatter(X_train, y_train, color = 'red')
22 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
23 plt.title('Salary vs Experience (Training set)')
24 plt.xlabel('Years of Experience')
25 plt.ylabel('Salary')
26 plt.show()
27
28 plt.scatter(X_test, y_test, color = 'red')
29 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
30 plt.title('Salary vs Experience (Training set)')
31 plt.xlabel('Years of Experience')
32 plt.ylabel('Salary')
33 plt.show()
34
35 m = regressor.coef_ #slope
36
37 c = regressor.intercept_ #intercept
38

```

Name	Type	Size	Value
c	float64	1	26780.099150628186
dataset	DataFrame	(30, 2)	Column names: YearsE...
m	Array of float64	(1,)	[9312.57512673]
regressor	linear_model.base.LinearRegression	1	LinearRegression obj...
X	Array of float64	(30, 1)	[[1.1] [1.3] [1.5] [10.3]
X_test	Array of float64	(6, 1)	[[1.5] [10.3]
X_train	Array of float64	(24, 1)	[[9.6] [4.]
y	Array of int64	(30,)	[39343 46205 3773...
y_pred	Array of float64	(6,)	[40748.96184072 122...

Variable Explorer Help Plots Files

Console 1/A

```

In [11]: m = regressor.coef_ #slope
In [12]: m
Out[12]: array([9312.57512673])

In [13]: c = regressor.intercept_ #intercept
In [14]: c
Out[14]: 26780.099150628186

In [15]:

```

IPython Console History

conda: base (Python 3.10.9) Completions: conda LSP: Python Line 38, Col 1 UTF-8 CRLF RW Mem 82%

22:38

Spyder (Python 3.10)

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```

13 X_train,X_test,y_train,y_test = train_test_split(X, y, test_size = 0.20, random_state = 0
14
15 from sklearn.linear_model import LinearRegression
16 regressor = LinearRegression()
17 regressor.fit(X_train, y_train)
18
19 y_pred = regressor.predict(X_test)
20
21 plt.scatter(X_train, y_train, color = 'red')
22 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
23 plt.title('Salary vs Experience (Training set)')
24 plt.xlabel('Years of Experience')
25 plt.ylabel('Salary')
26 plt.show()
27
28 plt.scatter(X_test, y_test, color = 'red')
29 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
30 plt.title('Salary vs Experience (Training set)')
31 plt.xlabel('Years of Experience')
32 plt.ylabel('Salary')
33 plt.show()
34
35 m = regressor.coef_ #slope
36
37 c = regressor.intercept_ #intercept
38
39 y_12 = 9312 * 12 + 26780
40 y_15 = 9312 * 20 + 26780
41 y_20 = 9312 * 12 + 26780
42 y_21 = 9312 * 20 + 26780
43
44 bias = regressor.score(X_train, y_train)
45 bias

```

Name	Type	Size	Value
y_train	Array of int64	(24,)	[112635 55794 8308...
y_test	Array of int64	(6,)	[37731 122391 5708...
y_pred	Array of float64	(6,)	[40748.96184072 122...
y_21	int	1	213020
y_20	int	1	138524
y_15	int	1	213020
y_12	int	1	138524
y	Array of int64	(30,)	[39343 46205 3773...
X_train	Array of float64	(24,)	[[9.6

Variable Explorer Help Plots Files

Console I/A X

```

... variance
Out[16]: 0.988169515729126

In [17]: y_12 = 9312 * 12 + 26780
...: y_15 = 9312 * 20 + 26780
...: y_20 = 9312 * 12 + 26780
...: y_21 = 9312 * 20 + 26780

In [18]: y_12
Out[18]: 138524

In [19]:

```

IPython Console History

conda: base (Python 3.10.9) Completions: conda LSP: Python Line 40, Col 25 UTF-8 CRLF RW Mem 84%

Type here to search

20°C 22:43 02-11-2023

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\yogay\untitled0.py

```

18 y_pred = regressor.predict(X_test)
19
20 plt.scatter(X_train, y_train, color = 'red')
21 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
22 plt.title('Salary vs Experience (Training set)')
23 plt.xlabel('Years of Experience')
24 plt.ylabel('Salary')
25 plt.show()
26
27 plt.scatter(X_test, y_test, color = 'red')
28 plt.plot(X_train, regressor.predict(X_train), color = 'blue')
29 plt.title('Salary vs Experience (Training set)')
30 plt.xlabel('Years of Experience')
31 plt.ylabel('Salary')
32 plt.show()
33
34
35 m = regressor.coef_ #slope
36
37 c = regressor.intercept_ #intercept
38
39 y_12 = 9312 * 12 + 26780
40 y_15 = 9312 * 20 + 26780
41 y_20 = 9312 * 12 + 26780
42 y_21 = 9312 * 20 + 26780
43
44 bias = regressor.score(X_train, y_train)
45 bias
46
47 variance = regressor.score(X_test, y_test)
48 variance
49
50

```

Variable Explorer

Name	Type	Size	Value
bias	float64	1	0.9411949620562126
c	float64	1	26780.099150628186
dataset	DataFrame	(30, 2)	Column names: Year
m	Array of float64	(1,)	[9312.57512673]
regressor	linear_model_base.LinearRegression	1	LinearRegression d
variance	float64	1	0.988169515729126
X	Array of float64	(30, 1)	[[1.1] [1.3]
X_test	Array of float64	(6, 1)	[[1.5] [10.3]
y_train	Array of float64	(30, 1)	[[9.6]

Console 1/A X

```

Out[18]: 150524

In [19]:
...: bias = regressor.score(X_train, y_train)
...: bias
Out[19]: 0.9411949620562126

In [20]: variance = regressor.score(X_test, y_test)
...: variance
Out[20]: 0.988169515729126

In [21]:

```

IPython Console History

conda: base (Python 3.10.9) Completions: conda LSP: Python Line 48, Col 9 UTF-8 CRLF RW Mem 84

Type here to search

20°C 22:43 02-11-2023