

How can I create a scatterplot with a regression line in Python?

Authored by
stats writer

April 18, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I create a scatterplot with a regression line in Python?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136628>

Creating a scatterplot with a regression line in Python involves using the Matplotlib library to plot the data points as a scatterplot and then using the NumPy and SciPy libraries to calculate and add the regression line to the plot. The regression line is a visual representation of the relationship between the two variables in the scatterplot, allowing for the identification of any linear trend in the data. By following a few simple steps, one can easily create a scatterplot with a regression line in Python, providing a clear and concise visualization of the data.

Create a Scatterplot with a Regression Line in Python

Often when you perform simple linear regression, you may be interested in creating a scatterplot to visualize the various combinations of x and y values along with the estimation regression line.

Fortunately there are two easy ways to create this type of plot in Python. This tutorial explains both methods using the following data:

```
import numpy as np
```

```
#create data
```

```
x = np.array()
```

```
y = np.array()
```

Method 1: Using Matplotlib

The following code shows how to create a scatterplot with an estimated regression line for this data using

Matplotlib:

```
import matplotlib.pyplot as plt
```

```
#create basic scatterplot
```

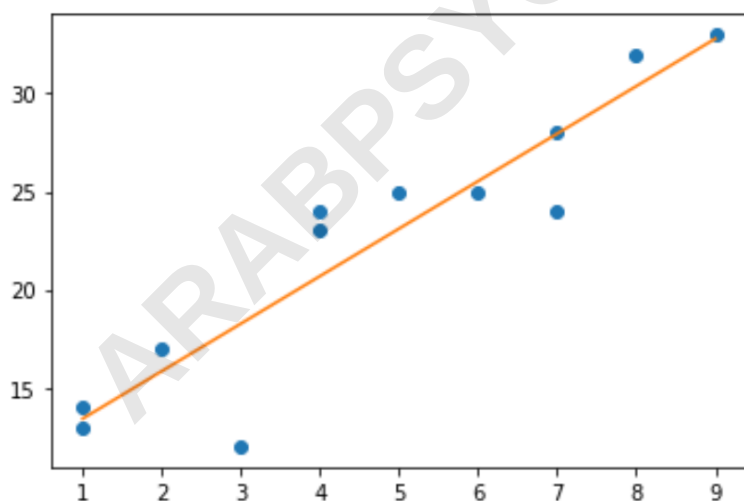
```
plt.plot(x, y, 'o')
```

```
#obtain m (slope) and b(intercept) of linear regression  
line
```

```
m, b = np.polyfit(x, y, 1)
```

```
#add linear regression line to scatterplot
```

```
plt.plot(x, m*x+b)
```



Feel free to modify the colors of the graph as you'd like. For example, here's how to change the individual points

to green and the line to red:

#use green as color for individual points

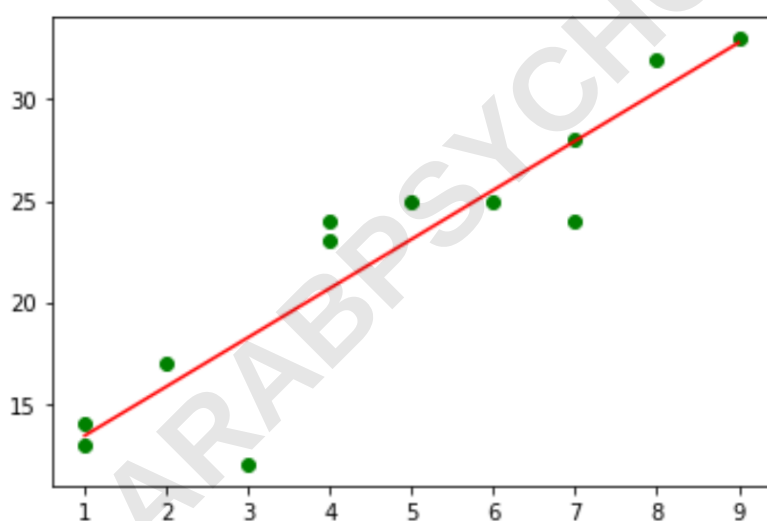
```
plt.plot(x, y, 'o', color='green')
```

#obtain m (slope) and b(intercept) of linear regression line

```
m, b = np.polyfit(x, y, 1)
```

#use red as color for regression line

```
plt.plot(x, m*x+b, color='red')
```



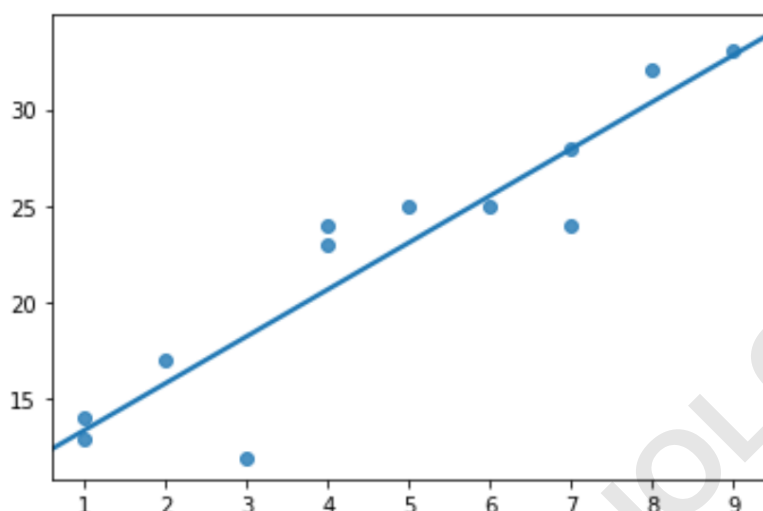
Method 2: Using Seaborn

You can also use the `regplot()` function from the Seaborn visualization library to create a scatterplot with a regression line:

```
import seaborn as sns
```

```
#create scatterplot with regression line
```

```
sns.regplot(x, y, ci=None)
```



Note that `ci=None` tells Seaborn to hide the confidence interval bands on the plot. You can choose to show them if you'd like, though:

```
import seaborn as sns
```

```
#create scatterplot with regression line and confidence interval lines
```

```
sns.regplot(x, y)
```

