How can the regression equation be displayed in a Seaborn regplot?

Authored by

stats writer

June 25, 2024

The Seaborn regplot is a visualization tool that allows users to display the relationship between two variables using a scatter plot and a regression line. By simply passing the two variables of interest to the regplot function, users can generate a scatter plot with the regression line displayed. The equation for the regression line, which represents the mathematical relationship between the two variables, is also displayed on the plot, providing a clear and concise visual representation of the relationship between the variables. This allows for easy interpretation and understanding of the data, making the Seaborn regplot a useful tool for data analysis.

Display Regression Equation in Seaborn Regplot

You can use the seaborn regplot function to plot a linear regression model fit to a dataset.

Unfortunately there is no built-in feature in seaborn to extract the regression equation of the line, but you can use the scipy.stats.linregress function to quickly find the regression coefficients:

```
import scipy
import seaborn as sns

#create regplot
p = sns.regplot(data=df, x=df.x, y=df.y)

#calculate slope and intercept of regression equation
slope, intercept, r, p, sterr = scipy.stats.linregress(x=p.get_lines().get_xdata(), y=p.get_lines().get_ydata())
```

The following example shows how to use this syntax in practice.

Example: Display Regression Equation in Seaborn Regplot

Suppose we have the following pandas DataFrame that contains information about the hours studied and final exam score of various students:

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'hours': ,
'score': })

#view DataFrame
print(df)
```

hours score
0 1 77
1 2 79
2 3 84
3 4 80
4 5 81
5 6 89
6 7 95

7 8 90

8 9 83

9 10 89

Suppose we would like to plot the data points and add a fitted regression line to the data.

We can use the following syntax to do so:

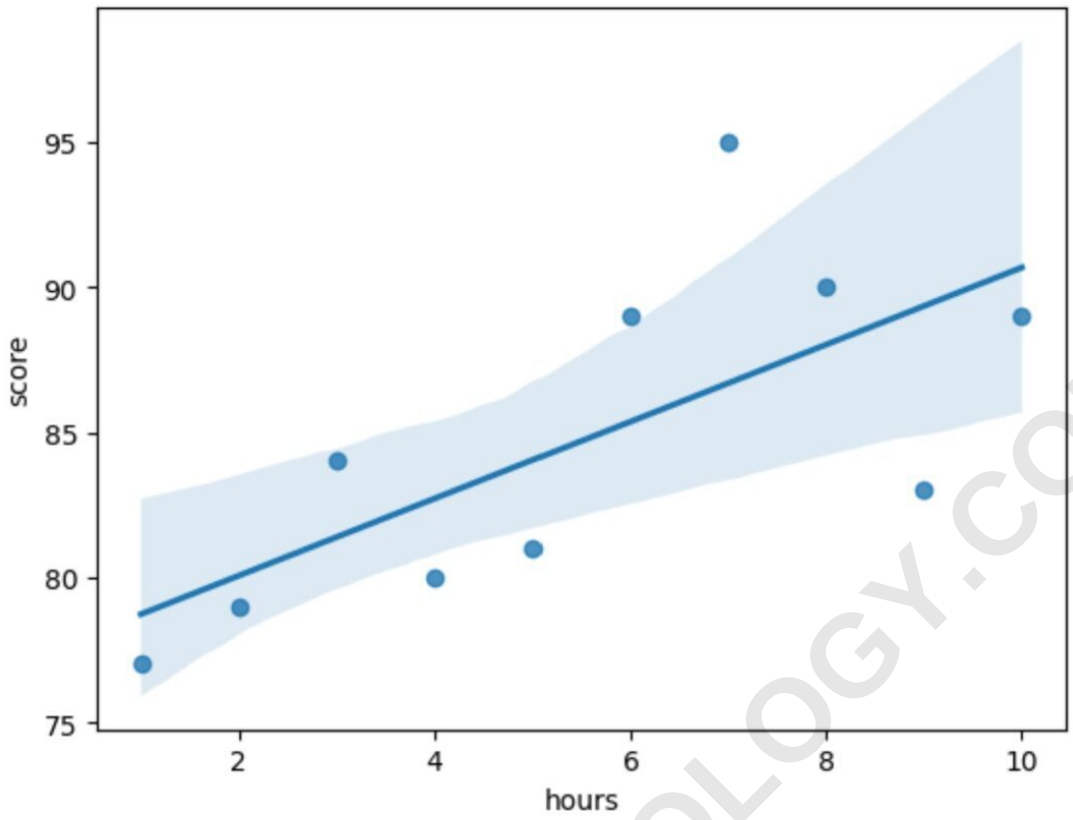
```
import scipy
import seaborn as sns

#create regplot
p = sns.regplot(data=df, x=df.hours, y=df.score)

#calculate slope and intercept of regression equation
slope, intercept, r, p, sterr = scipy.stats.linregress(x=p.get_lines().get_xdata(), y=p.get_lines().get_ydata())

#display slope and intercept of regression equation
print(intercept, slope)
```

77.39999999999995 1.3272727272727356



From the output we can see that the regression line has the following equation:

y = 77.4 + 1.327

If we would like to display this equation on the seaborn regplot, we can use the text() function from matplotlib:

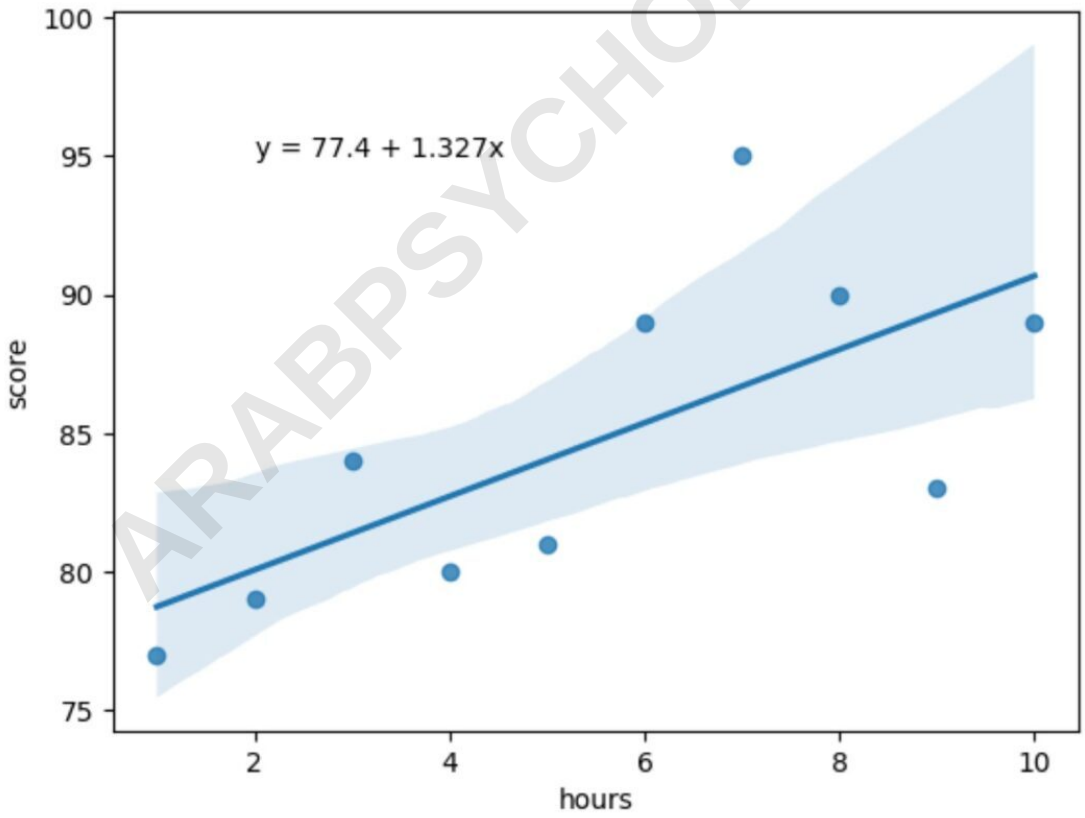
```
import matplotlib.pyplotas plt
import scipy
import seaborn as sns

#create regplot
```

```
p = sns.regplot(data=df, x=df.hours, y=df.score)

#calculate slope and intercept of regression equation
slope, intercept, r, p, sterr = scipy.stats.linregress(x=p.get_lines().get_xdata(), y=p.get_lines().get_ydata())

#add regression equation to plot
plt.text(2, 95, 'y = ' + str(round(intercept,3)) + ' + ' + str(round(slope,3)) + 'x')
```



Notice that the regression equation is now displayed in

the top left corner of the plot.

Feel free to modify these coordinates to display the regression equation where you'd like in your own plot.

Note: You can find the complete documentation for the seaborn regplot function .

The following tutorials explain how to perform other common tasks in seaborn: