

How can I create a scatterplot from a Pandas DataFrame?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I create a scatterplot from a Pandas DataFrame?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138656>

A scatterplot is a visual representation of the relationship between two numerical variables. To create a scatterplot from a Pandas DataFrame, you can use the "plot.scatter()" function. This function takes in the two columns from the DataFrame that you want to plot and generates a scatterplot with those variables as the x and y axes. By specifying the desired columns and adding any necessary labels and titles, you can easily create a scatterplot from a Pandas DataFrame to visualize the relationship between your data.

Make a Scatterplot From a Pandas DataFrame

There are two ways to create a scatterplot using data from a pandas DataFrame:

1. Use pandas.DataFrame.plot.scatter

One way to create a scatterplot is to use the built-in pandas function:

```
import pandas as pd

df.plot.scatter(x = 'x_column_name', y = 'y_column_name')
```

2. Use matplotlib.pyplot.scatter

Another way to create a scatterplot is to use the Matplotlib function:

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

`plt.scatter(df.x, df.y)`

This tutorial provides an example of how to use each of these methods.

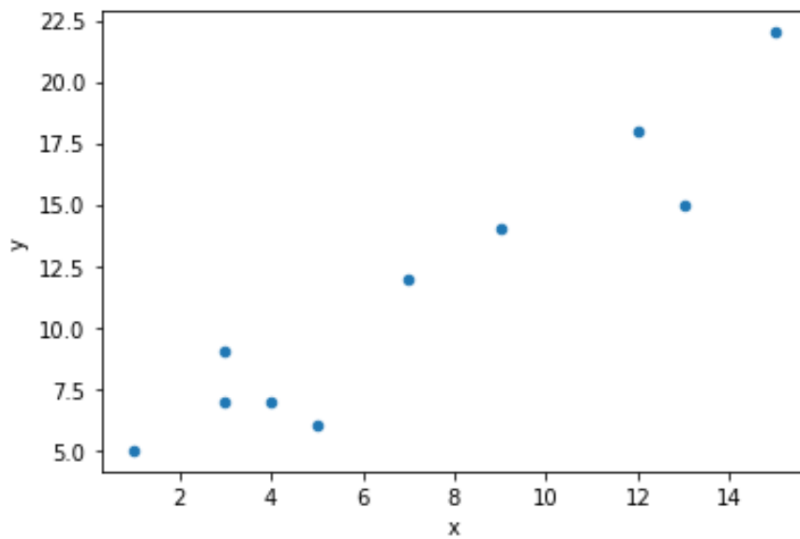
Example 1: Use Pandas

The following code shows how to use the function to create a simple scatterplot:

```
import pandas as pd

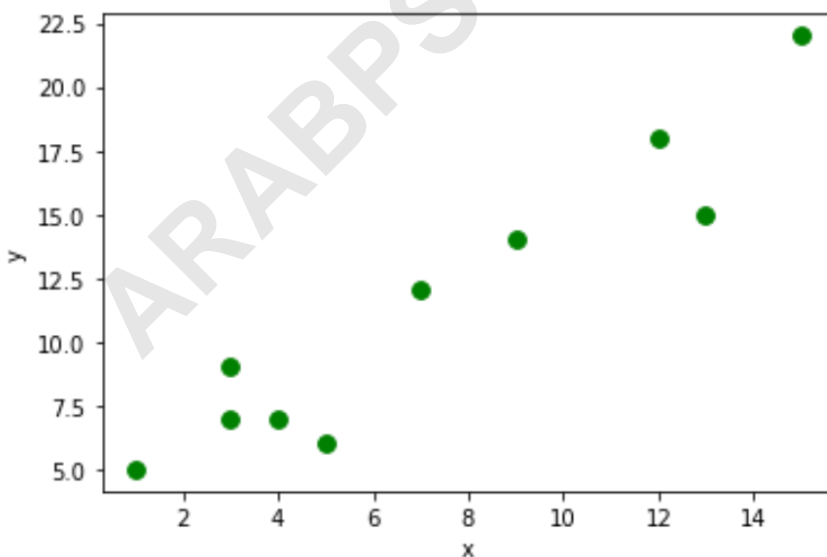
#create DataFrame
df = pd.DataFrame({'x': ,
'y': })

#create scatterplot
df.plot.scatter(x='x', y='y')
```



Note that you can use the `s` and `c` arguments to modify the size and color of the points, respectively:

```
df.plot.scatter(x='x', y='y', s=60, c='green')
```



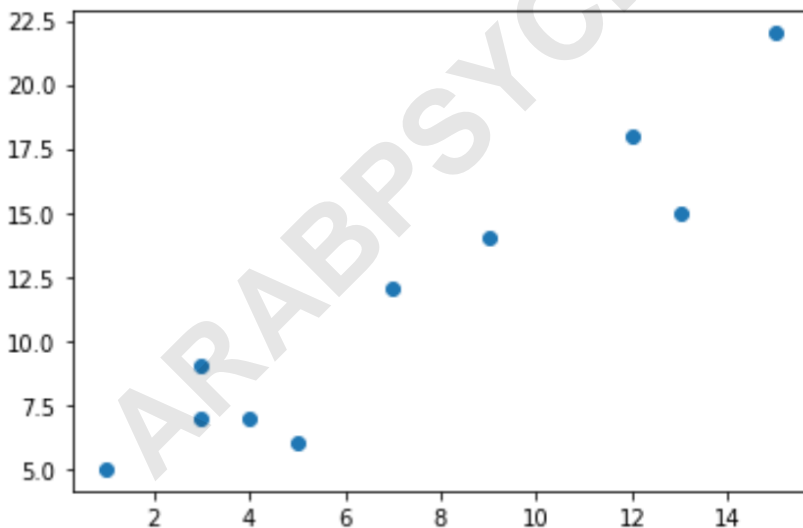
Example 2: Use Matplotlib

The following code shows how to use the function to create a scatterplot:

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

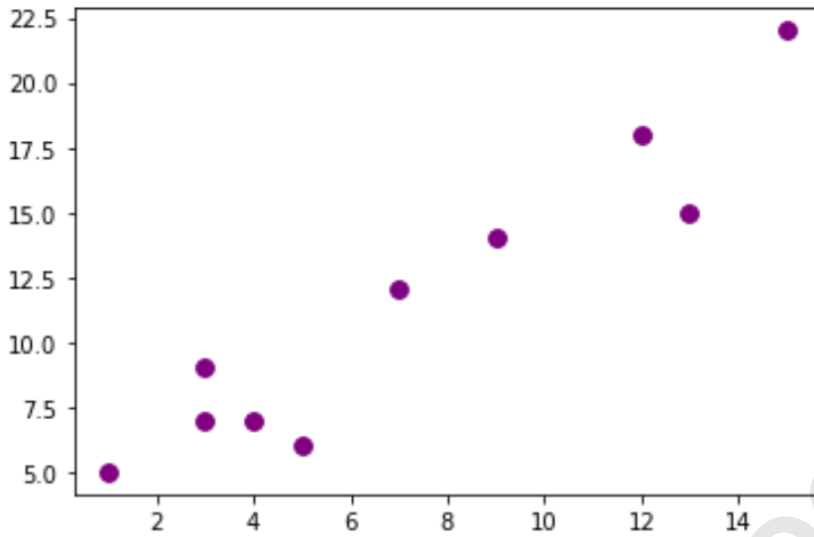
#create DataFrame
df = pd.DataFrame({'x': ,
'y': })

#create scatterplot
plt.scatter(df.x, df.y)
```



Note that you can use the `s` and `c` arguments to modify the size and color of the points, respectively:

```
plt.scatter(df.x, df.y, s=60, c='purple')
```



You can find more Python tutorials .