

How can I create boxplots by group using Matplotlib?

Authored by
stats writer

April 29, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I create boxplots by group using Matplotlib?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140772>

Boxplots are a type of statistical graph used to display the distribution of numerical data through quartiles. In order to create boxplots by group using Matplotlib, one must first import the necessary libraries and data. Then, the data needs to be grouped according to the desired categories. Next, the boxplot function from Matplotlib can be used to plot the data for each group, with the option to customize the appearance of the boxplots. This process allows for easy comparison of the data distribution between different groups, making it a useful tool for data analysis and visualization.

Matplotlib: Create Boxplots by Group

You can use the following syntax to create boxplots by group in Matplotlib:

```
import matplotlib as plt
```

```
import seaborn as sns
```

```
sns.boxplot(x='group', y='values', data=df)
```

The following examples show how to use this syntax to create boxplots by group for datasets in both long-form and wide-form.

Example 1: Boxplots by Group for Long-Form Data

The following code shows how to create boxplots by group for a pandas DataFrame in a long format:

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib as plt
import seaborn as sns

#create long-form data
df = pd.DataFrame({'points': ,
'team': np.repeat(, 5)})
```

```
#view data
print(df)
```

```
points team
```

```
0 7 A
```

```
1 8 A
```

```
2 9 A
```

```
3 12 A
```

```
4 14 A
```

```
5 5 B
```

```
6 6 B
```

```
7 6 B
```

```
8 8 B
```

```
9 11 B
```

```
10 8 C
```

```
11 9 C
```

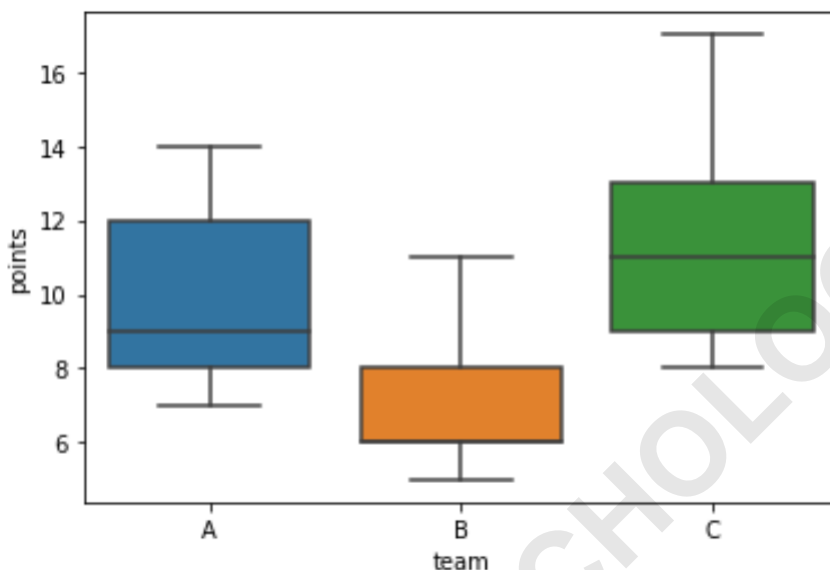
```
12 11 C
```

```
13 13 C
```

14 17 C

#create boxplot by group

```
sns.boxplot(x='team', y='points', data=df)
```



Example 2: Boxplots by Group for Wide-Form Data

The following code shows how to create boxplots by group for a pandas DataFrame in a wide format:

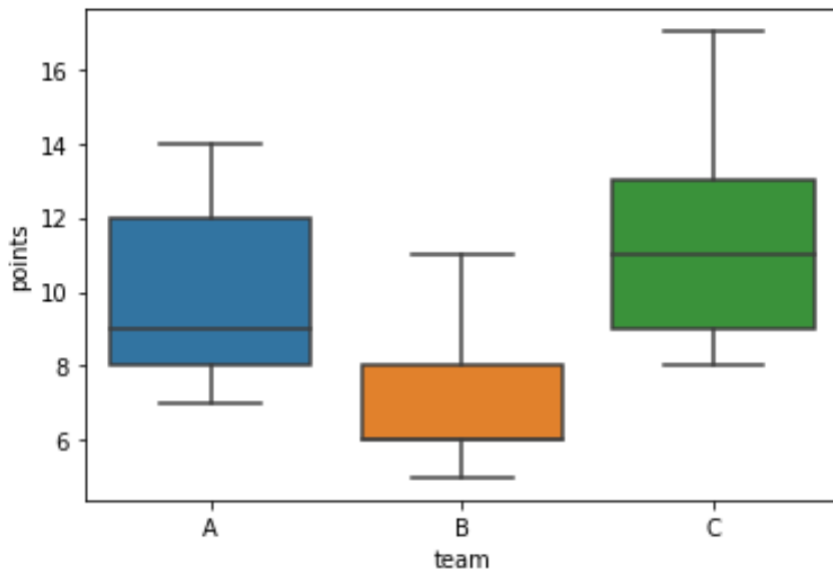
```
import pandas as pd  
import numpy as np  
import matplotlib as plt  
import seaborn as sns  
  
#create wide-form data
```

```
df = pd.DataFrame({'A': ,  
'B': ,  
'C': })
```

```
#view data  
print(df)
```

```
A B C  
0 7 5 8  
1 8 6 9  
2 9 6 11  
3 12 8 13  
4 14 11 17
```

```
#create boxplot by group  
sns.boxplot(x='variable', y='value',  
data=pd.melt(df)).set(  
xlabel='team',  
ylabel='points')
```



ARABPSYCHOLOGY.COM