How can I create a boxplot from a Pandas DataFrame?

Authored by

stats writer

May 3, 2024

A Pandas DataFrame is a data structure used for storing and manipulating tabular data in Python. To create a boxplot from a Pandas DataFrame, you can use the built-in plot function and specify the 'kind' parameter as 'box'. This will generate a boxplot visualizing the distribution of the data within the DataFrame. Additionally, you can customize the appearance of the boxplot by specifying the 'color' and 'whis' (whisker length) parameters. By creating a boxplot from a Pandas DataFrame, you can easily identify outliers and the overall spread of the data.

Create Boxplot from Pandas DataFrame

You can use the following syntax to create boxplots from a pandas DataFrame:

#create boxplot of one column
df.boxplot(column=)

#create boxplot of multiple columns
df.boxplot(column=)

#create boxplot grouped by one column
df.boxplot(column=, by='col2')

The following examples show how to use this syntax in practice with the following DataFrame:

import pandas as pd

#create DataFrame
df = pd.DataFrame({'conference': ,

'points': ,

'assists': ,

'rebounds': ,})

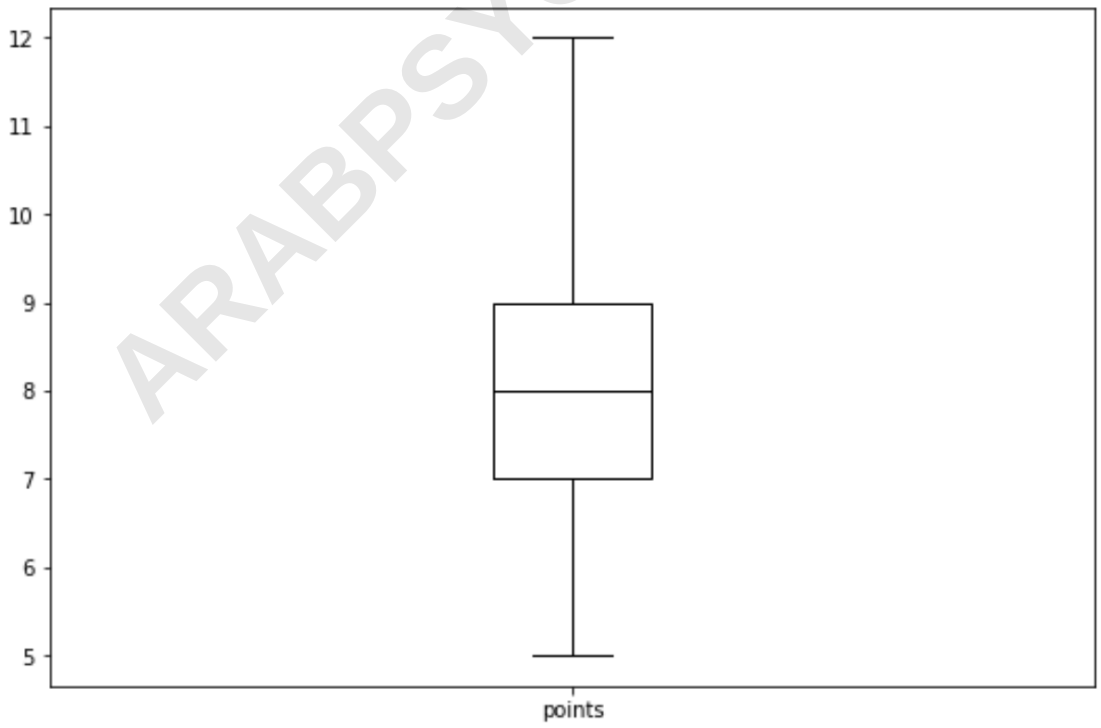
#view DataFrame

df

Example 1: Boxplot of One Column

The following code shows how to create a boxplot for one column in a pandas DataFrame:

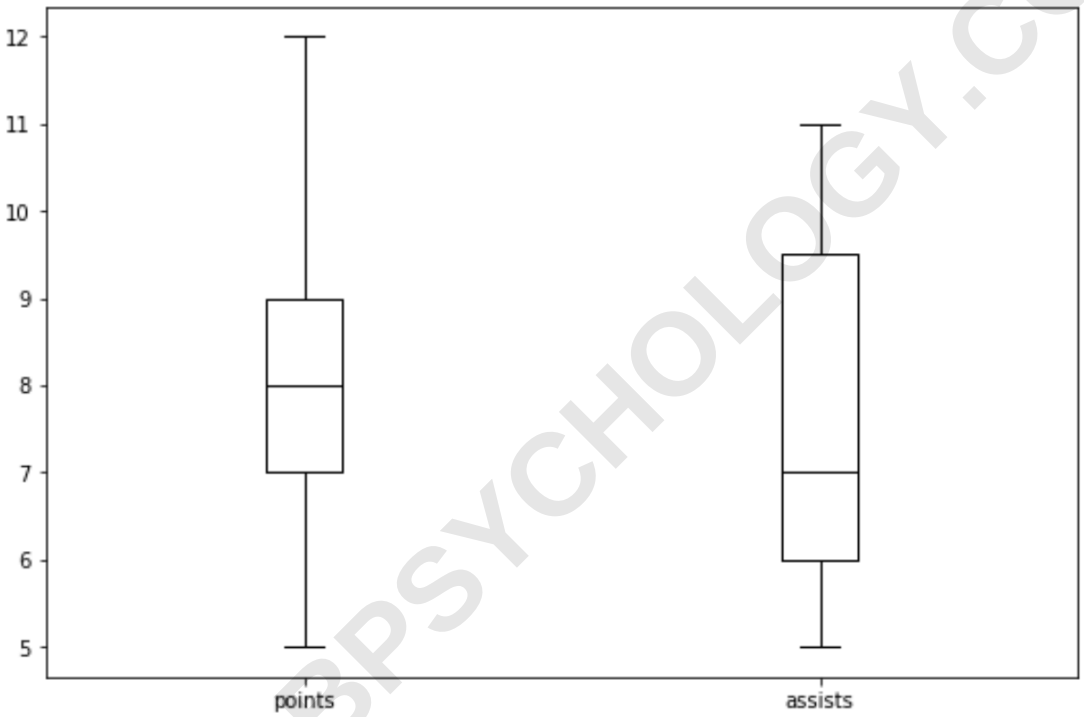
df.boxplot(column=, grid=False, color='black')



Example 2: Boxplot of Multiple Columns

The following code shows how to create a boxplot for multiple columns in a pandas DataFrame:

df.boxplot(column=, grid=False, color='black')



Example 3: Boxplot Grouped by One Column

The following code shows how to create a boxplot grouped by one column in a pandas DataFrame:

df.boxplot(column=, by='conference', grid=False, color='black')