

How do you calculate the mean of columns in Pandas?

Authored by
stats writer

April 17, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you calculate the mean of columns in Pandas?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136380>

Calculating the mean of columns in Pandas involves using the built-in function "mean()" which calculates the average value of a specific column in a Pandas DataFrame. This function can be applied to a single column or multiple columns at once, providing a quick and efficient way to calculate the mean of large datasets. The mean can also be calculated for specific rows by using conditional statements. Overall, using the "mean()" function in Pandas allows for easy and accurate calculation of column means, which is a crucial step in data analysis and statistical modeling.

Calculate the Mean of Columns in Pandas

Often you may be interested in calculating the mean of one or more columns in a pandas DataFrame. Fortunately you can do this easily in pandas using the function.

This tutorial shows several examples of how to use this function.

Example 1: Find the Mean of a Single Column

Suppose we have the following pandas DataFrame:

```
import pandas as pd
import numpy as np

#create DataFrame
df = pd.DataFrame({'player': ,
'points': ,
'assists': ,
```

```
'rebounds': })
```

```
#view DataFrame
```

```
df
```

```
player points assists rebounds
```

```
0 A 25 5 NaN
```

```
1 B 20 7 8.0
```

```
2 C 14 7 10.0
```

```
3 D 16 8 6.0
```

```
4 E 27 5 6.0
```

```
5 F 20 7 9.0
```

```
6 G 12 6 6.0
```

```
7 H 15 9 10.0
```

```
8 I 14 9 10.0
```

```
9 J 19 5 7.0
```

We can find the mean of the column titled "points" by using the following syntax:

```
df.mean()
```

```
18.2
```

The mean() function will also exclude NA's by default.

For example, if we find the mean of the "rebounds" column, the first value of "NaN" will simply be excluded from the calculation:

```
df.mean()
```

```
8.0
```

If you attempt to find the mean of a column that is not numeric, you will receive an error:

```
df.mean()
```

```
TypeError: Could not convert ABCDEFGHIJ to numeric
```

Example 2: Find the Mean of Multiple Columns

We can find the mean of multiple columns by using the following syntax:

```
#find mean of points and rebounds columns
```

```
df].mean()
```

```
rebounds 8.0
```

```
points 18.2
```

```
dtype: float64
```

Example 3: Find the Mean of All Columns

We can find also find the mean of all numeric columns by using the following syntax:

```
#find mean of all numeric columns in DataFrame  
df.mean()
```

```
points 18.2
```

```
assists 6.8
```

```
rebounds 8.0
```

```
dtype: float64
```

ARABPSYCHOLOGY.COM