

How can I change the number of bins used in a Pandas histogram?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I change the number of bins used in a Pandas histogram?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=151907>

To change the number of bins used in a Pandas histogram, the "bins" parameter in the "hist" function can be adjusted. This parameter allows for specifying the number of bins or the bin edges to be used in the histogram. By increasing or decreasing the number of bins, the granularity of the histogram can be adjusted, providing a more detailed or broader view of the data distribution. This is useful for better understanding the data and identifying any potential patterns or outliers present.

Change Number of Bins Used in Pandas Histogram

You can use the bins argument to modify the number of bins used in a pandas histogram:

```
df.plot.hist(columns=, bins=10)
```

The default number of bins used in a pandas histogram is 10.

The following example shows how to use the bins argument in practice.

Example: How to Change Number of Bins in Pandas Histogram

Suppose we have the following pandas DataFrame that contains information about points scored by basketball players on various teams:

```
import pandas as pd
import numpy as np
```

```
#make this example reproducible
```

```
np.random.seed(1)
```

```
#create DataFrame
```

```
df = pd.DataFrame({'team': np.repeat('A', 100),
```

```
'points': np.random.normal(loc=20, scale=2, size=300)})
```

```
#view head of DataFrame
```

```
print(df.head())
```

```
team points
```

```
0 A 23.248691
```

```
1 A 18.776487
```

```
2 A 18.943656
```

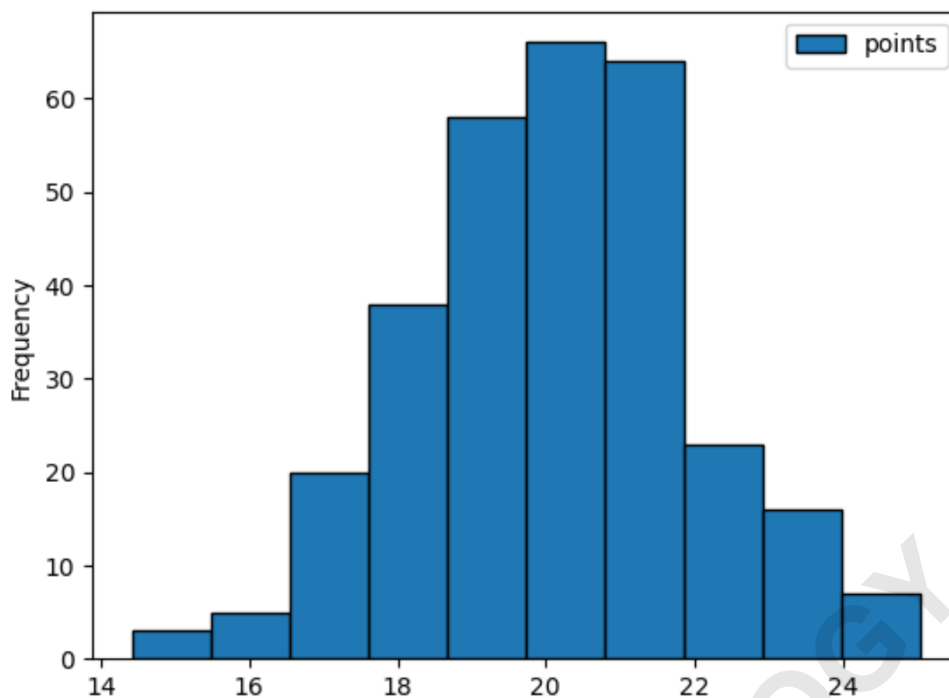
```
3 A 17.854063
```

```
4 A 21.730815
```

If we create a histogram to visualize the distribution of values for the points variable, pandas will use 10 bins in the histogram by default:

```
#create histogram to visualize distribution of points
```

```
df.plot.hist(column='points', edgecolor='black')
```

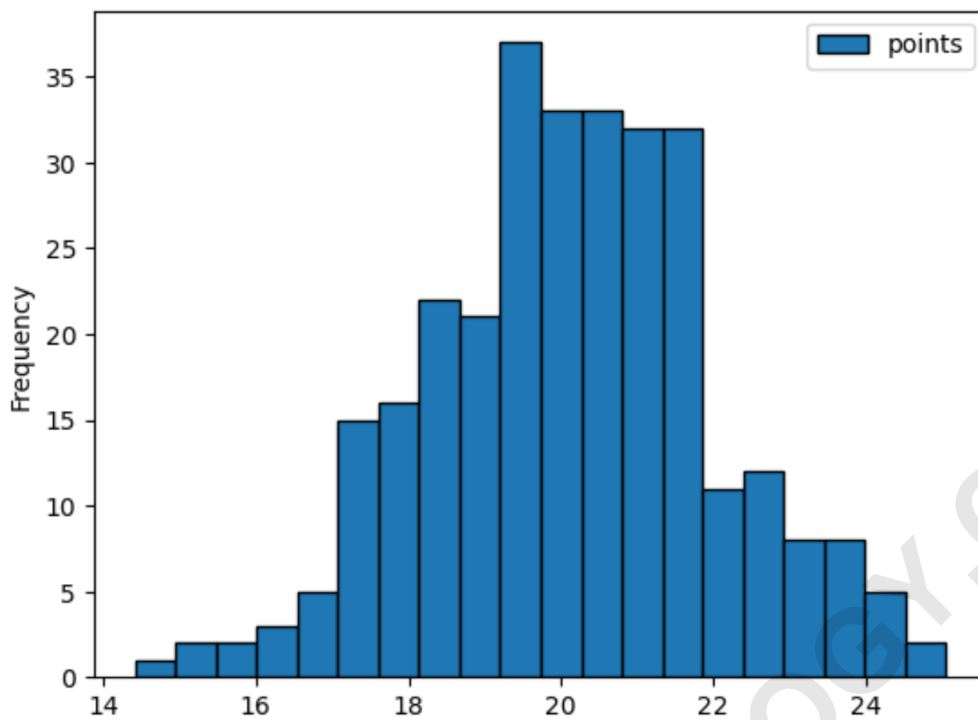


Notice how there are 10 bars in the histogram.

However, suppose we use the bins argument to change the total number of bins to 20:

#create histogram with 20 bins

df.plot.hist(column=, edgecolor='black', bins=20)

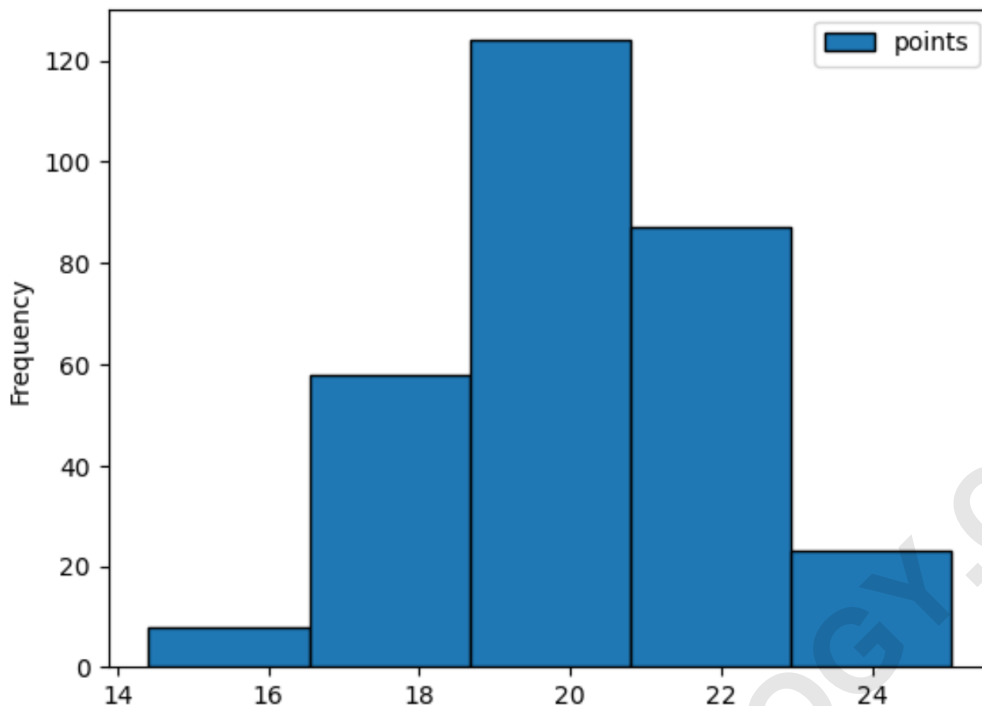


Notice how there are now 20 bars in the histogram.

We could also decrease the number of bins to 5:

#create histogram with 5 bins

df.plot.hist(column=, edgecolor='black', bins=5)



There are now 5 total bars in the histogram.

Feel free to adjust the value for the bins argument to create the exact histogram you'd like.

However, keep the following in mind when choosing the number of bins:

If you choose too few bins, the true underlying pattern in the data may be hidden. If you choose too many bins, you may just be visualizing the noise in the data.

One helpful way to determine the optimal number of bins to use in a histogram is to use .

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

ARABPSYCHOLOGY.COM