

How can I create a histogram with a log scale using Pandas?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I create a histogram with a log scale using Pandas?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154146>

To create a histogram with a log scale using Pandas, follow these steps:

1. Import the Pandas library into your Python code.
2. Load your dataset into a Pandas DataFrame.
3. Use the Pandas plot() function to create a histogram of your data.
4. Set the x-axis to display a logarithmic scale by using the logx=True parameter in the plot() function.
5. Adjust the other parameters of the plot, such as the number of bins and the color, to suit your preferences.
6. Add labels and a title to the histogram for better understanding.
7. Finally, use the show() function to display the histogram with a log scale.

By following these steps, you can easily create a histogram with a log scale using Pandas, which can be helpful in visualizing data that has a large range of values.

Pandas: Create a Histogram with Log Scale

You can use the logx and logy arguments to create histograms with log scales on the x-axis and y-axis, respectively, in pandas:

```
#create histogram with log scale on x-axis  
df.plot(kind='hist', logx=True)
```

```
#create histogram with log scale on y-axis  
df.plot(kind='hist', logy=True)
```

The following example shows how to use these arguments to create histograms with log scales in pandas.

Related:

Example: Create Histogram with Log Scale in Pandas

Suppose we have the following pandas DataFrame with 5,000 rows:

```
import pandas as pd
import numpy as np

#make this example reproducible
np.random.seed(1)

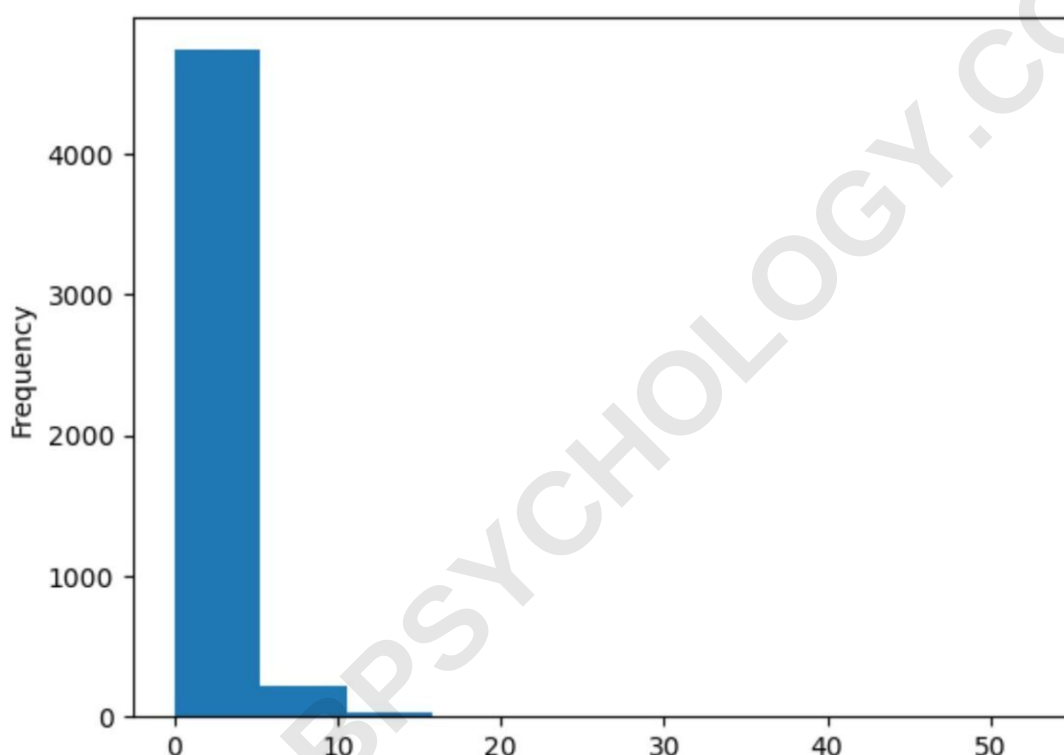
#create DataFrame
df = pd.DataFrame({'values':
np.random.lognormal(size=5000)})

#view first five rows of DataFrame
print(df.head())

values
0 5.075096
1 0.542397
2 0.589682
3 0.341992
4 2.375974
```

We can use the following syntax to create a histogram with a linear scale on both the x-axis and y-axis:

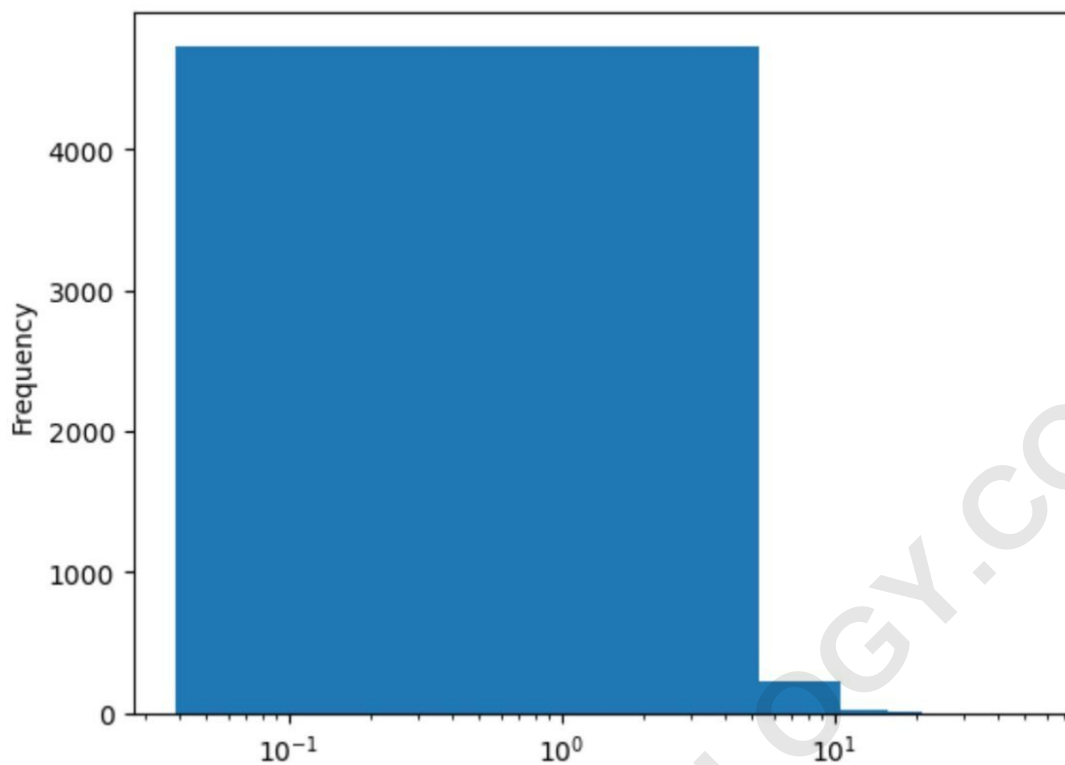
```
#create histogram  
df.plot(kind='hist')
```



The x-axis and y-axis both currently have a linear scale.

We can use the `logx=True` argument to convert the x-axis to a log scale:

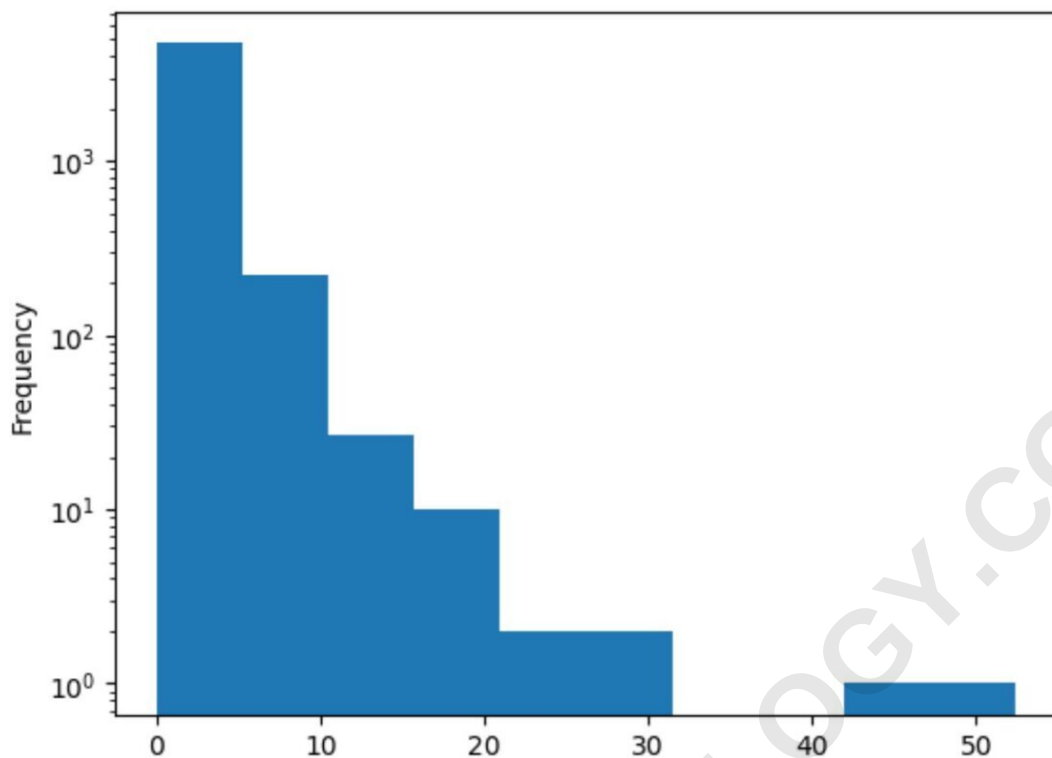
```
#create histogram with log scale on x-axis  
df.plot(kind='hist', logx=True)
```



The values on the x-axis now follow a log scale.

And we can use the `logy=True` argument to convert the y-axis to a log scale:

```
#create histogram with log scale on y-axis  
df.plot(kind='hist', logy=True)
```



The values on the y-axis now follow a log scale.

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