

How can I calculate deciles in Python with examples?

Authored by
stats writer

April 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate deciles in Python with examples?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140067>

The process of calculating deciles in Python involves dividing a dataset into ten equal parts, where each part contains an equal number of data points. This allows for the analysis of distribution and patterns within the dataset. To calculate deciles, the data must first be sorted in ascending order. Then, the desired decile can be found by multiplying the total number of data points by the decimal equivalent of the desired decile (e.g. for the 5th decile, multiply by 0.5) and rounding up to the nearest whole number. This value represents the index of the data point in the sorted dataset that corresponds to the desired decile. Examples of decile calculations in Python can be found in various libraries such as Pandas, NumPy, and SciPy, and can be implemented using built-in functions or custom code. These calculations can provide valuable insights into the distribution and characteristics of a dataset, allowing for informed decision making and analysis.

Calculate Deciles in Python (With Examples)

In statistics, deciles are numbers that split a dataset into ten groups of equal frequency.

The first decile is the point where 10% of all data values lie below it. The second decile is the point where 20% of all data values lie below it, and so on.

We can use the following syntax to calculate the deciles for a dataset in Python:

```
import numpy as np
```

```
np.percentile(var, np.arange(0, 100, 10))
```

The following example shows how to use this function in practice.

Example: Calculate Deciles in Python

The following code shows how to create a fake dataset with 20 values and then calculate the values for the deciles of the dataset:

```
import numpy as np

#create data
data = np.array()

#calculate deciles of data
np.percentile(data, np.arange(0, 100, 10))

array()
```

The way to interpret the deciles is as follows:

10% of all data values lie below 63.420% of all data values lie below 67.8.30% of all data values lie below 76.5.40% of all data values lie below 83.6.50% of all data values lie below 88.5.60% of all data values lie below 90.4.70% of all data values lie below 92.3.80% of all data values lie below 93.2.90% of all data values lie below 95.2.

Note that the first value in the output (56) simply denotes the minimum value in the dataset.

Example: Place Values into Deciles in Python

To place each data value into a decile, we can use the `qcut` pandas function.

Here's how to use this function for the dataset we created in the previous example:

```
import pandas as pd

#create data frame
df = pd.DataFrame({'values': })

#calculate decile of each value in data frame
df = pd.qcut(df, 10, labels=False)

#display data frame
df

values Decile
0 56 0
1 58 0
2 64 1
3 67 1
```

4 68 2

5 73 2

6 78 3

7 83 3

8 84 4

9 88 4

10 89 5

11 90 5

12 91 6

13 92 6

14 93 7

15 93 7

16 94 8

17 95 8

18 97 9

19 99 9

The way to interpret the output is as follows:

The data value 56 falls between the percentile 0% and 10%, thus it falls in decile 0. The data value 58 falls between the percentile 0% and 10%, thus it falls in decile 0. The data value 64 falls between the percentile 10% and 20%, thus it falls in decile 1. The data value 67

falls between the percentile 10% and 20%, thus it falls decile 1. The data value 68 falls between the percentile 20% and 30%, thus it falls decile 2.

ARABPSYCHOLOGY.COM