

The Mean: Takeaways

by Dataquest Labs, Inc. - All rights reserved © 2020

Syntax

- Computing the mean of any numerical array:

```
### Pure Python ###  
mean = sum(array) / len(array)  
  
### Using numpy ###  
from numpy import mean  
mean_numpy = mean(array)
```

- Computing the mean of a `Series` :

```
mean = Series.mean()
```

Concepts

- We can summarize the distribution of a numerical variable by computing its **mean**.
- The mean is a single value and is the result of taking into account **equally** each value in the distribution.
- The mean is **the balance point** of a distribution — the total distance of the values below the mean is equal to the total distance of the values above the mean.
- The mean μ of a population can be defined algebraically in several equivalent ways:

$$\mu = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N} = \frac{1}{N} \left(\sum_{i=1}^N x_i \right)$$

- The mean \bar{x} of a sample can be defined algebraically in several equivalent ways:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n} = \frac{1}{n} \left(\sum_{i=1}^n X_i \right)$$

- The sample mean \bar{X} is an unbiased estimator for the population mean μ .

Resources

- [The Wikipedia entry](#) on the mean.
- Useful documentation:
 - [numpy.mean\(\)](#)
 - [Series.mean\(\)](#)



Takeaways by Dataquest Labs, Inc. - All rights reserved © 2020