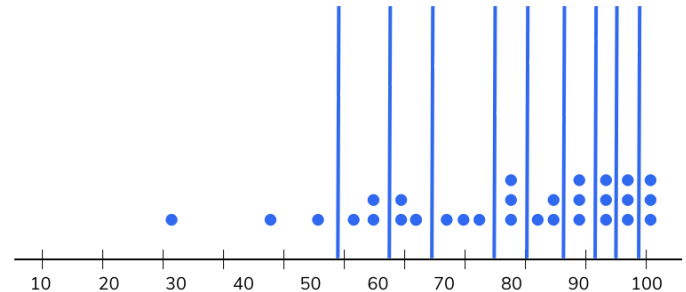


Quartiles, Quantiles, and Interquartile Range

Quantiles

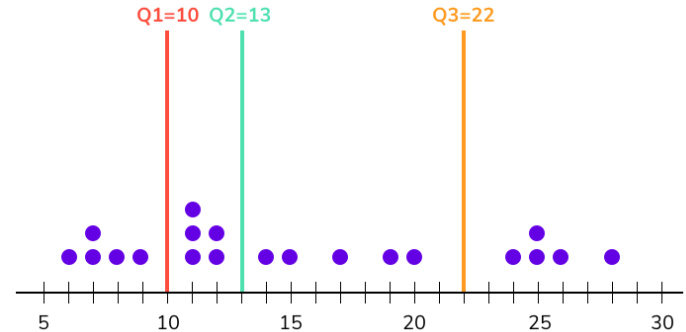
Quantiles are the set of values/points that divides the dataset into groups of equal size. For example, in the figure, there are nine values that splits the dataset.

Those nine values are quantiles.



Quartiles

The three dividing points (or quartiles) that split data into four equally sized groups are called quartiles. For example, in the figure, the three dividing points Q1, Q2, Q3 are quartiles.



Numpy's Quantile() Function

In Python, the `numpy.quantile()` function takes an array and a number say `q` between 0 and 1. It returns the value at the `q` th quantile. For example, `numpy.quantile(data, 0.25)` returns the value at the first quartile of the dataset `data`.

Quantiles and Groups

If the number of quantiles is n , then the number of equally sized groups in a dataset is $n+1$.

```
import numpy as np
data = [1,2,3,4,5]
first_quartile = np.quantile(data, 0.25)
```

Median in Quantiles

The median is the divider between the upper and lower halves of a dataset. It is the 50%, 0.5 quantile, also known as the 2-quantile.

```
# The value 5 is both the median and the  
2-quantile
```

```
data = [1, 3, 5, 9, 20]  
Second_quantile = 5
```

Interquartile Range Definition

The interquartile range is the difference between the first(Q1) and third quartiles(Q3). It can be mathematically represented as $IQR = Q3 - Q1$.

Interquartile Range and Outliers

The interquartile range is considered to be a robust statistic because it is not distorted by outliers like the average (or mean).

```
# Eventhough d_2 has an outlier, the IQR  
is identical for the 2 datasets
```

```
d_1 = [1,2,3,4,5,6,7,8,9]  
d_2 = [-100,2,3,4,5,6,7,8,9]
```