

# Introduction to Statistics

The Five-number Summary, the IQR,  
the Midhinge and the Trimean

[www.Stats-Lab.com](http://www.Stats-Lab.com)

Twitter: @StatsLabDublin

# The Five-number Summary

The five-number summary is a descriptive statistic that provides information about a set of observations.

It consists of the five most important sample percentiles:

- ▶ the sample minimum (smallest observation)
- ▶ the lower quartile or first quartile ( $Q_1$ )
- ▶ the median (middle value)
- ▶ the upper quartile or third quartile ( $Q_3$ )
- ▶ the sample maximum (largest observation)

## The Five-number Summary : Sample Data

Suppose a sample has the following five-number summary:

- ▶ the sample minimum = 11
- ▶ the lower quartile  $Q_1 = 25$
- ▶ the median = 27
- ▶ the upper quartile  $Q_3 = 31$
- ▶ the sample maximum = 38

We will use these values for later examples.

## Interquartile Range

The interquartile range (IQR) is a measure of statistical dispersion, being equal to the difference between the first and third quartiles,

$$IQR = Q_3 - Q_1.$$

For our sample data, the interquartile range is

$$IQR = 31 - 25 = 6$$

(The median is the corresponding measure of location or central tendency.)

## Interquartile Range and Outliers

The interquartile range is often used to find outliers in data.

Using this approach, outliers are observations that fall below the **lower fence**

$$\text{Lower fence} = Q1 - (1.5 \times IQR)$$

or above the **upper fence**

$$\text{Upper fence} = Q3 + (1.5 \times IQR)$$

.

# Interquartile Range and Outliers

## Lower Fence

- ▶ Lower fence =  $Q1 - (1.5 \times IQR)$
- ▶ Lower fence =  $25 - (1.5 \times 6) = 25 - 9 = 16$

Any value less than 16 (for example, the minimum value 11) is considered an outlier.

# Interquartile Range and Outliers

## Upper Fence

- ▶ Upper fence =  $Q3 + (1.5 \times IQR)$
- ▶ Upper fence =  $31 + (1.5 \times 6) = 31 + 9 = 40$

Any value greater than 40 would be considered an outlier. However, as the maximum value is 38, there are no high-value outliers.

## Midhinge

The midhinge is a measure of central location, determined as the average of the first and third quartiles.

$$\text{midhinge} = \frac{Q_1 + Q_3}{2}$$

For the sample data, the midhinge is computed as

$$\text{midhinge} = \frac{25 + 31}{2} = \frac{56}{2} = 28$$



## Trimean

The **trimean** (TM) is a measure of central location defined as a weighted average of the distribution's median and its two quartiles:

$$TM = \frac{Q_1 + 2Q_2 + Q_3}{4}$$

This is equivalent to the average of the median and the midhinge:

$$TM = \frac{1}{2} \left( Q_2 + \frac{Q_1 + Q_3}{2} \right)$$

## Trimean

$$TM = \frac{Q_1 + 2Q_2 + Q_3}{4}$$

For the sample data, the trimean is computed as

$$TM = \frac{25 + (2 \times 27) + 31}{4} = \frac{110}{4} = 27.5$$

