# Statistics Basics

Author: Steffen Knödler
Date: 03.07.2019

## 1 Basic Terms

### 1.1 **Average/ Arithmetic Mean/ Mean**

The *arithmetic mean* (or simply *mean*) of a sample $x_1 + x_2 + \cdots + x_n$, usually denoted by $\bar{x}$ , is the sum of the sampled values divided by the number of items in the example

$$\bar{x} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

For example, the arithmetic mean of five values: 4, 36, 45, 50, 75 is:

$$\frac{4 + 36 + 45 + 50 + 75}{5} = \frac{210}{5} = 42.$$

### 1.2 Median

The median is the value separating the higher half from the lower half of the data sample. We can refer to the "middle" value. Therefore, the median is less influenced by distorting high or low numbers.

For example, if there is an odd number of numbers, the middle one is picked:
1, 3, 3, 6, 7, 8, 9 – median = 6

If there is an even number of observations, the median is then usually defined as the mean of the two middle numbers
1, 2, 3, 4, 5, 6, 8, 9 – median = 4.5

### 1.3 Mode

The mode type describes the most frequent value in a data set.

For example, 1, 2, 2, 3, 4, 7, 9 – mode = 2

# 2 Boxplot

Boxplots enable us to study the distributional characteristics of a group of scores as well as the level of the scores.