

# Probability and Statistics

**MA2103 - 2023**

Rajesh Kumble Nayak

# **Statistics:**

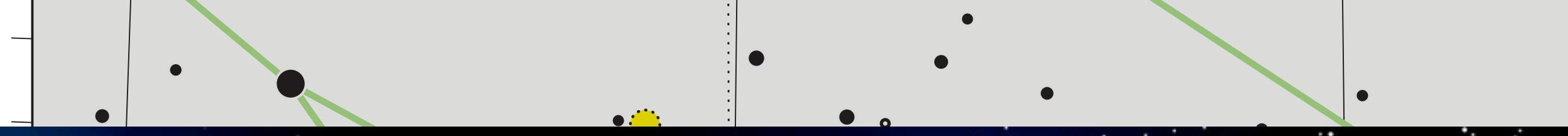
The basic goal of statistics is to analyse, model or to understand and summarise a set of Data

# **Data:**

Data is an outcome of experiments or an observations

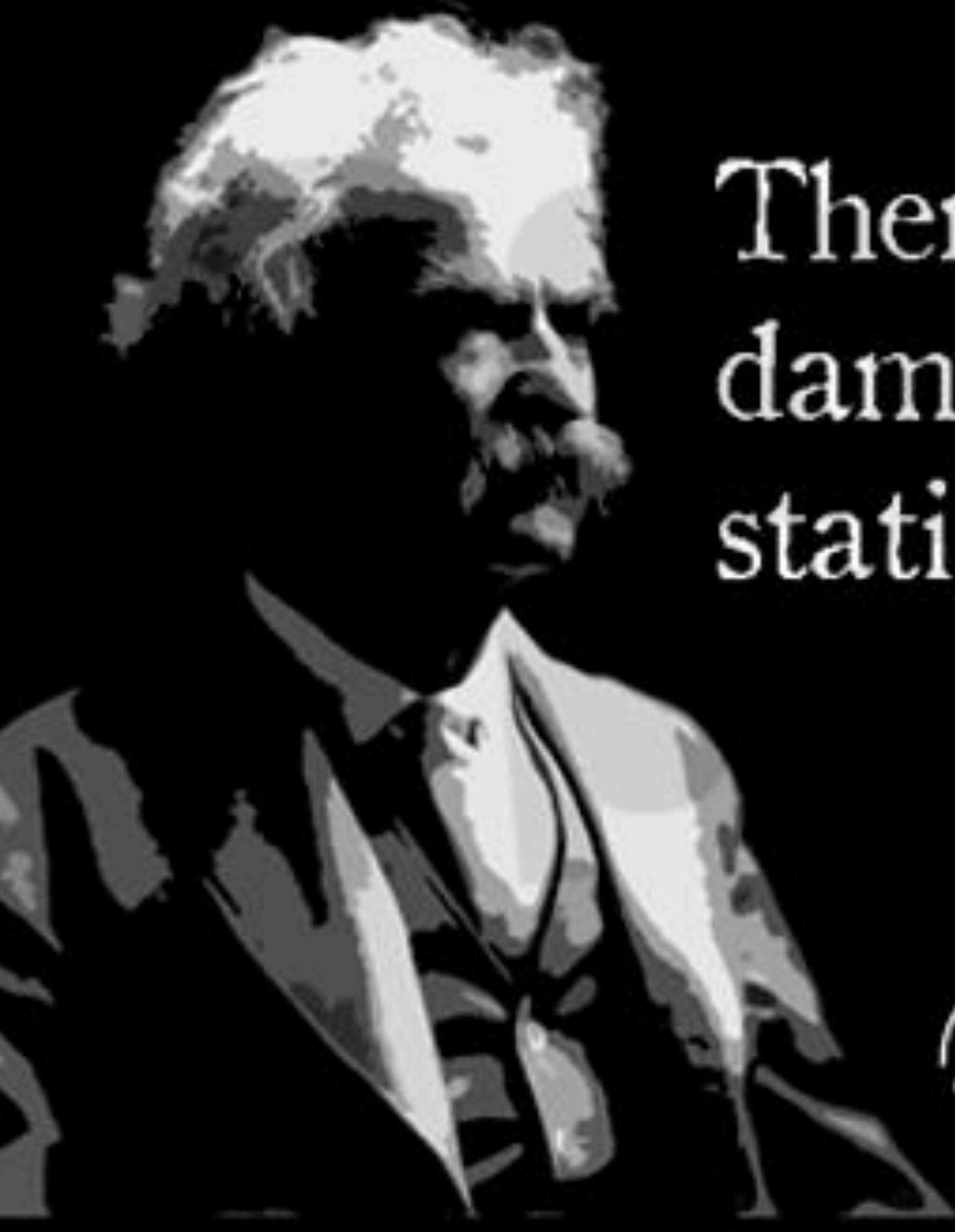
We apply tools of statistics on data to understand and interpret the data.

TAURUS



# Statistics:

Systematic understanding or modelling a complex data and summarise the interpretation and conclusions in a reproducible manner!



There are lies,  
damned lies and  
statistics.

Mark Twain

# The French paradox

# French people display the lowest mortality rate from heart disease and cardiovascular disease

# French drinks

# French



A good measure or sometimes also called statistics is essential part of the understanding data

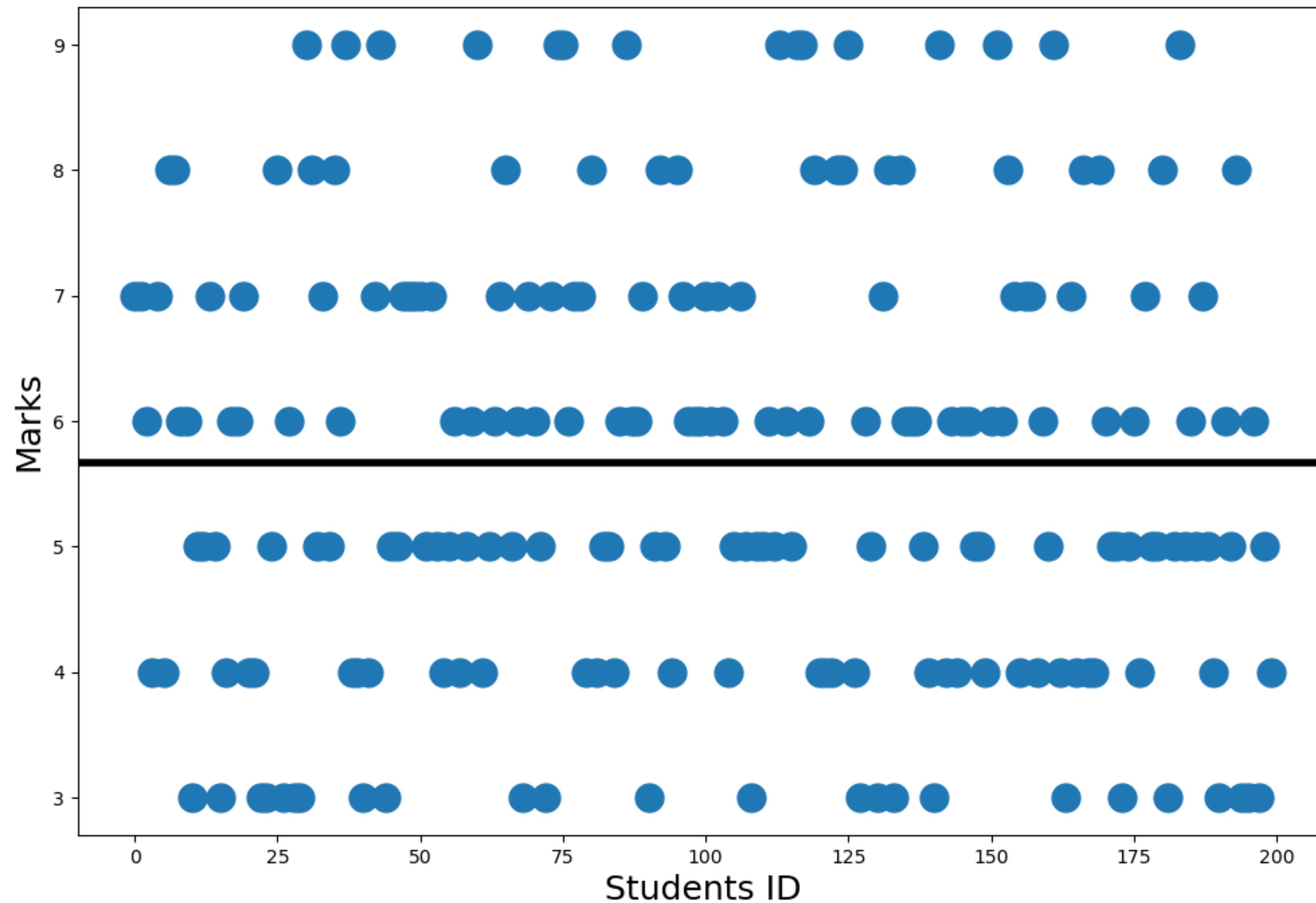
**Example:**

We have 200 students in a class and in an class test marks are given in range of 0 to 10. We have to come up with various measure to understand the results

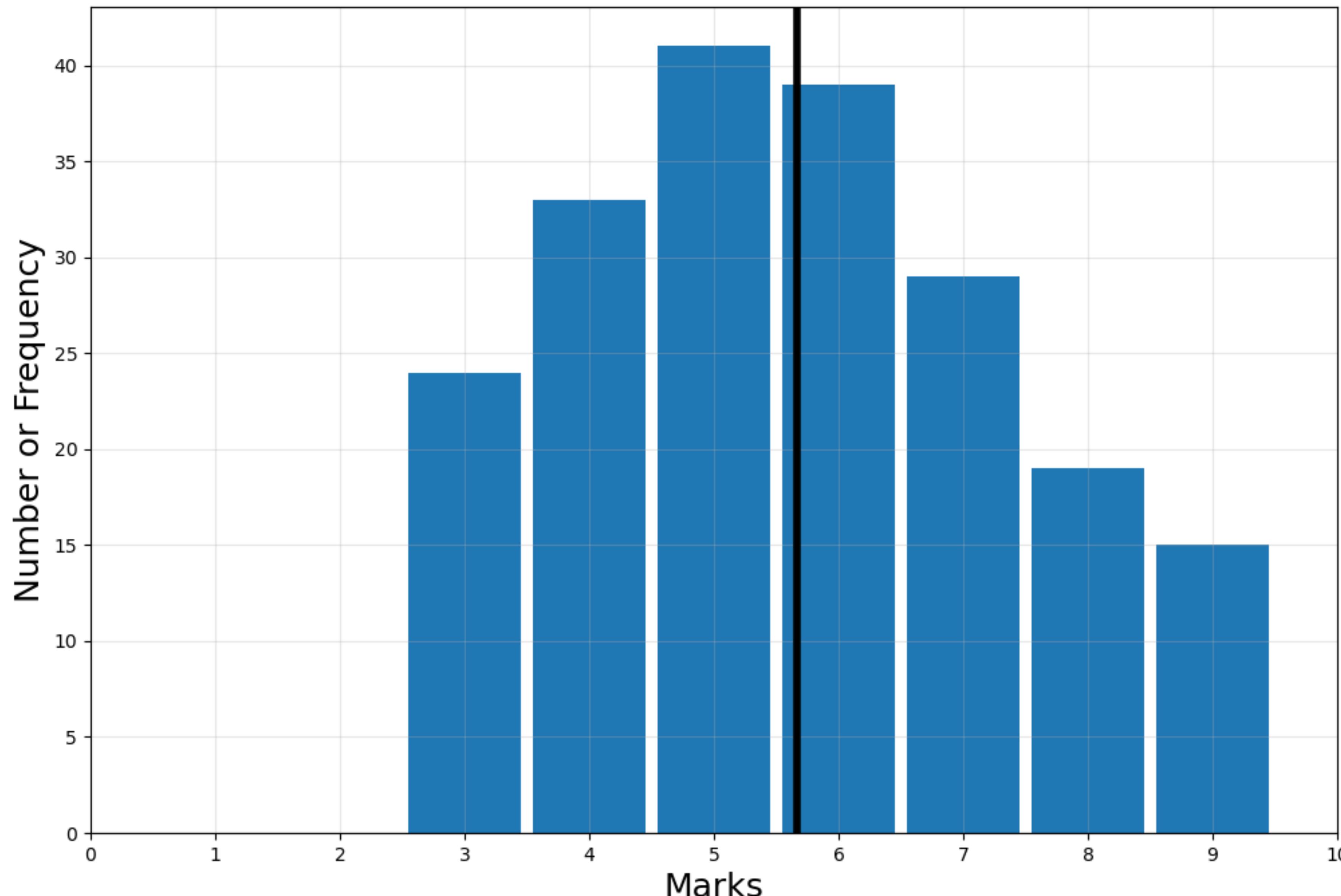
Let  $\{x_n\}$  for  $n = 1$  to 200 be the marks a number between 0 to 10

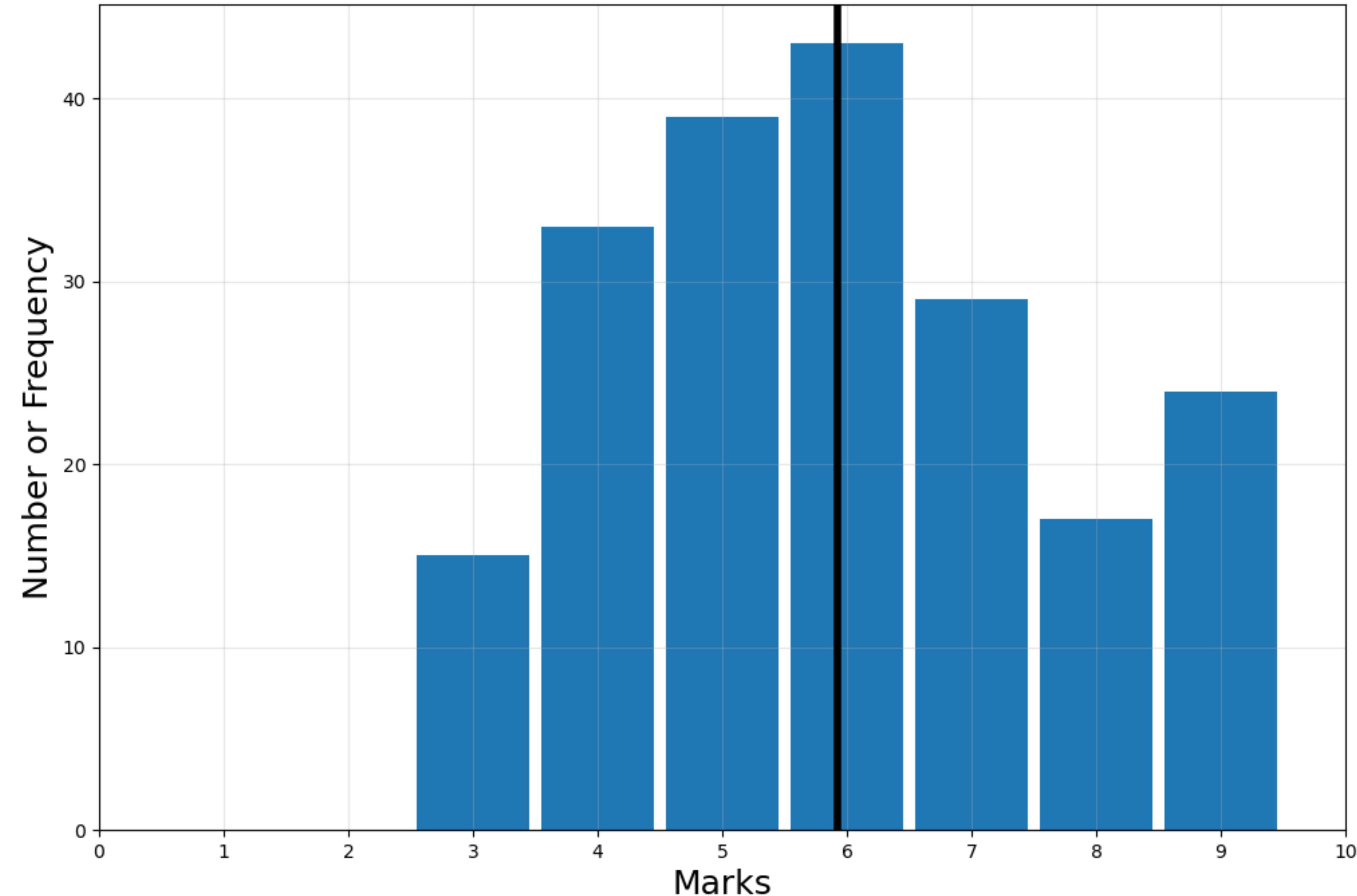
## Mean or Arithmetic average

One of the simplest measure is mean  $\mu$  or  $\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n$

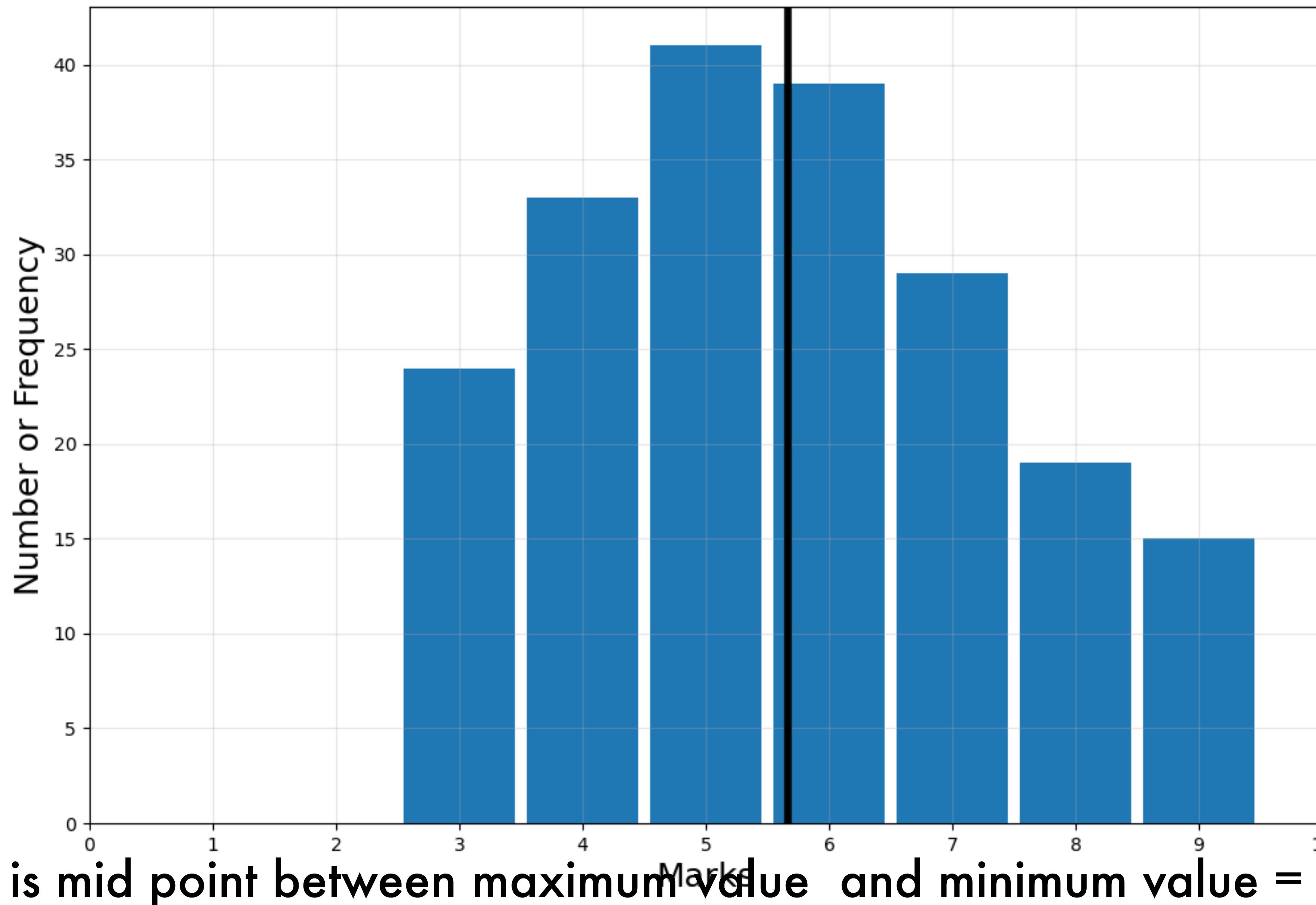


In this case histogram is better representation of data

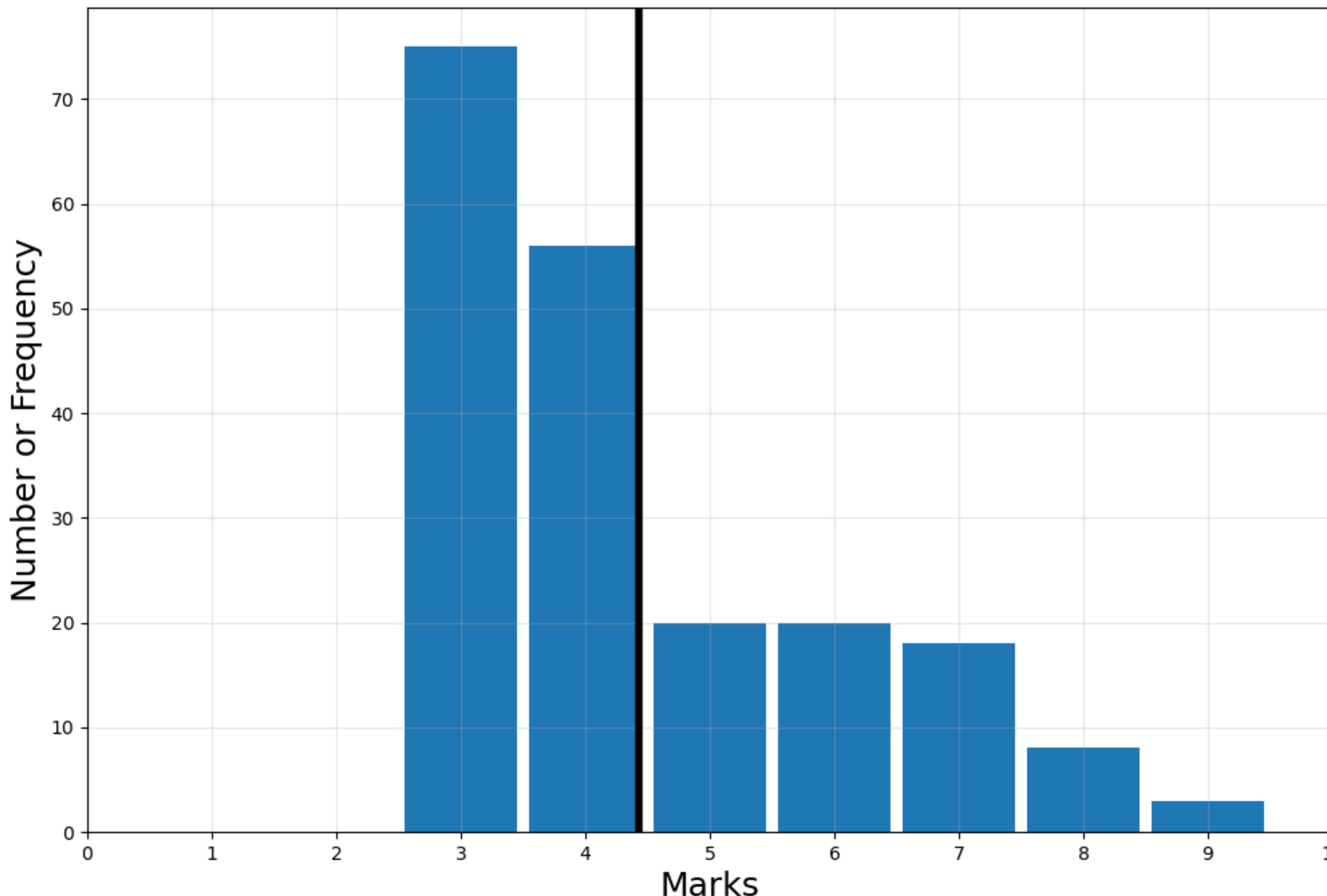




**One of the other measure is median, usually referenced as mid point**



In this example, we have mean and median very different number!



# What if Data is not a number

Data need not be numbers in general

Ex: Phone number

Favourite colour

etc

We can map them into numbers, but mean and median might not have any meaning

Another example: what is your favourite colour?

