Statistics – A Full University Course on Data Science Basics

What is Statistics?

Statistics in the study of how to collect, organize, analyze, and interpret numerical information and data. It is both the science of uncertainty and the technology of extracting information from data. Statistics is used to help use make decisions.

Individuals – are people or objects included in a study.

Variable – is a characteristic of the individual to be measured or observed.

Population (N) – is a group of people or objects with a common theme. When every member of that thematic group is considered, it is a population. A parameter is a measure that describes the entire population.

In a population, if variable data is measured and gathered for every individual in that population then this set of data is called a census.

Sample (n) – is a small portion of the population that can be either representative or biased. A statistic is a measure that describes only a sample of a population.

The distinction between population parameter and sample statistic is important because the types of data used in different depending on context.

In sample data, data is measured and gathered for some of the individuals in the population.

Descriptive Statistics – involve methods of organizing, picturing, and summarizing information from samples and populations.

Inferential Statistics – involves methods of using information from a sample to draw conclusions regarding the population.

Classifying Variables

Quantitative versus Qualitative

Interval versus Ratio

Nominal versus Ordinal