**Data Analysis and Statistics:**

Fundamentals of Statistics

|  |  |
| --- | --- |
| **Symbol** | **Meaning** |
| **x** | A variable |
| **StartAbsoluteValue x EndAbsoluteValue** | Absolute value of a variable The absolute value of a number is always positive, so StartAbsoluteValue negative 5 EndAbsoluteValue equals 5. |
| **upper Sigma** | Summation For example, sigma-summation x Subscript i denotes adding all observations of a variable together. |
| **N** | Population size |
| **mu** | Population mean |
| **sigma squared** | Population variance |
| **sigma** | Population standard deviation |
| **n** | Sample size |
| **x overbar** | Sample mean |
| **x overTilde** | Sample median |
| **s squared** | Sample variance |
| **s** | Sample standard deviation |
| **upper C** | Confidence level value |
| **upper Z** | Standardized score |
| **alpha** | Significance level |
| **upper Z Subscript StartFraction alpha Over 2 EndFraction** | Critical value for a confidence interval level |
| **r** | Pearson correlation coefficient |

**Descriptive statistics** is a branch of statistics that summarizes and describes data.

* descriptive statistics measures to help you understand the characteristics of your dataset.

Measures of frequency:

* helps you understand how often something happens
* When encountering a dataset for the first time, you want to determine how much data you are working with to help guide your analysis

**Example:** suppose you are working with human performance data. One of the first things to understand is the size of the dataset. One way to accomplish this quickly is to count the number of observations.

1. Count
2. **Percentage**

- The percentage is a frequency measure that identifies the proportion of a given value for a variable with respect to the total number of rows in the dataset.

- To calculate a percentage, you need the total number of observations and the total number of observations for a specific value of a variable.

1. Frequency

* Frequency describes how often a specific value for a variable occurs in a dataset.

**Measures of Central Tendency**