Terms to Know

Population	entire collection of individuals or objects that information is gathered from
Sample	subsets of population that is randomly selected
Census	performed to gather information about an entire population
Descriptive Statistics	method of organizing and summarizing data
Inferential Statistics	making generalizations from a sample to a population
Variable	characteristics that make change
Data	actual observations collected
Categorical Variable	no meaningful numerical value
Numerical Variable	measurements, can be averaged
Discrete Numerical	can only be specific values
Continuous Numerical	can be any value
Univariate	describes single characteristics
Symmetrical Distribution	both sides of data are approximately the same when folded down the middle
Normal Distribution	bell-shaped, occurs a lot in real life
Uniform Distribution	rectangular shaped
Skewed Distribution	one side is longer than the other
Bimodal	has two distinct maximums
Parameter	a fixed value about a population; typically unkown
Statistic	a value calculated from a sample that is used to estimate the parameter
Median	the middle of the data; 50th percentile
Mean	the arithmetic average
Mode	the observation that occurs the most often
Resistant	when a statistic is not affected by outliers
Variability	how spread out the data is
Standard Deviation	a measure of the average distance from the mean

Terms to Know (cont)

Interquartile Range	the range of the middle 50% of the data
Variance	the average of the squared deviations

SOCCS

Shape	the distribution of the graph (symmetrical, uniform, normal, skewed, bimodal)
Outliers	values that are away from the rest of the data
Context	explain what the graph is showing
Center	describe the mean or median of the data
Spread	describe the range, IQR, or standard deviation of the data

Graph Types

Bar Graph	categorical, bars do not touch
Pie Graph	categorical, percentages, can be approximated
Dotplot	numerical, putting dots on a number line
Stem Plot	univariate numerical, must have a key
Discrete Histograms	numerical, bars touch, centered over discrete values
Continuous Histograms	numerical, bars touch, bars cover an intervals of values
Boxplot	displays percentages of data, construction is subjective, useful for comparisons
Ogives	shows relative frequency, or percentiles for data

Formulas

Z-Score	Z = (value - mean)/(standard deviation)
IQR	Q3 - Q1
Lower Fence	Q1 - 1.5(IQR)
Upper Fence	Q3 - 1.5(IQR)

Empirical Rule