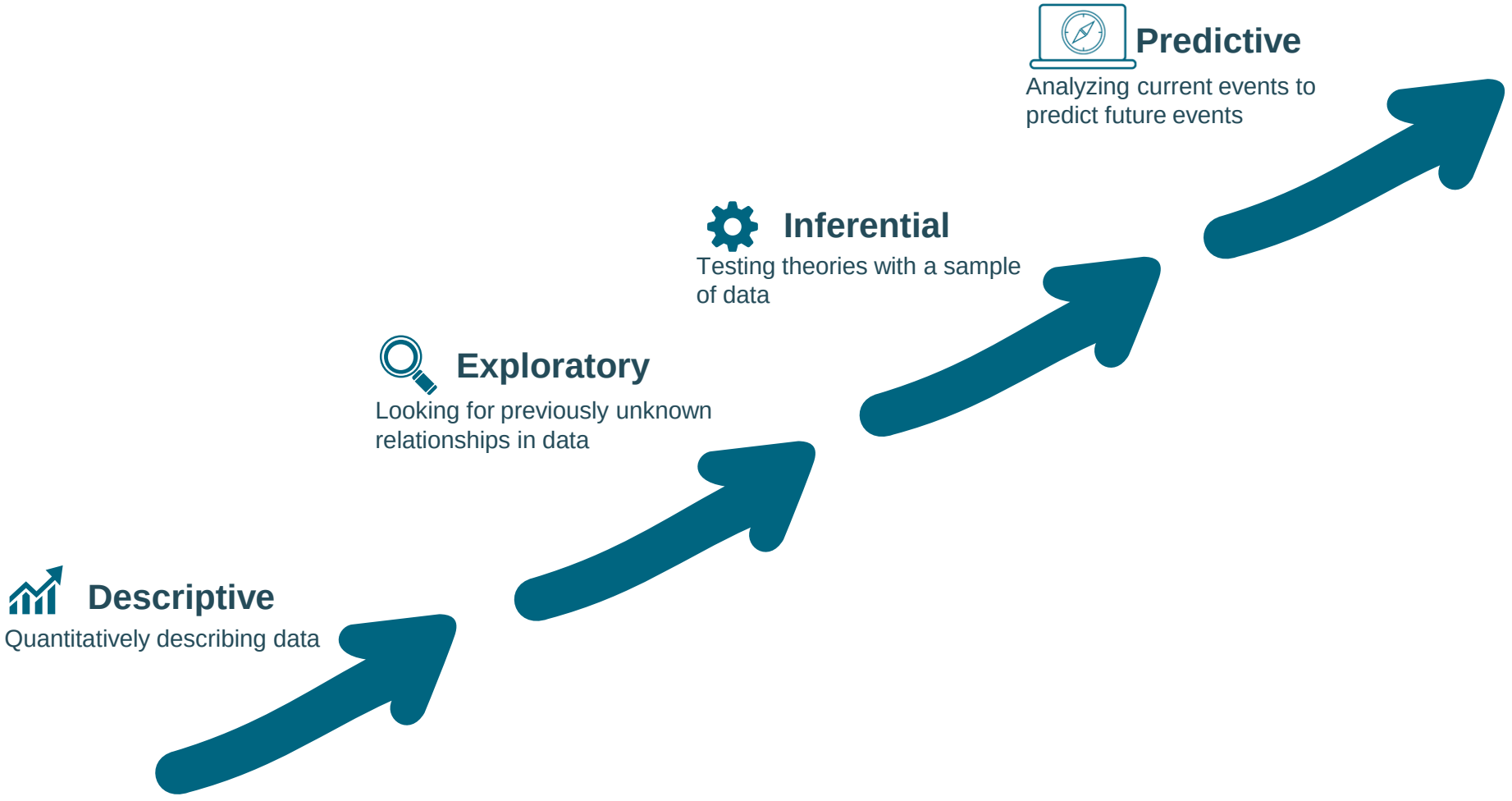
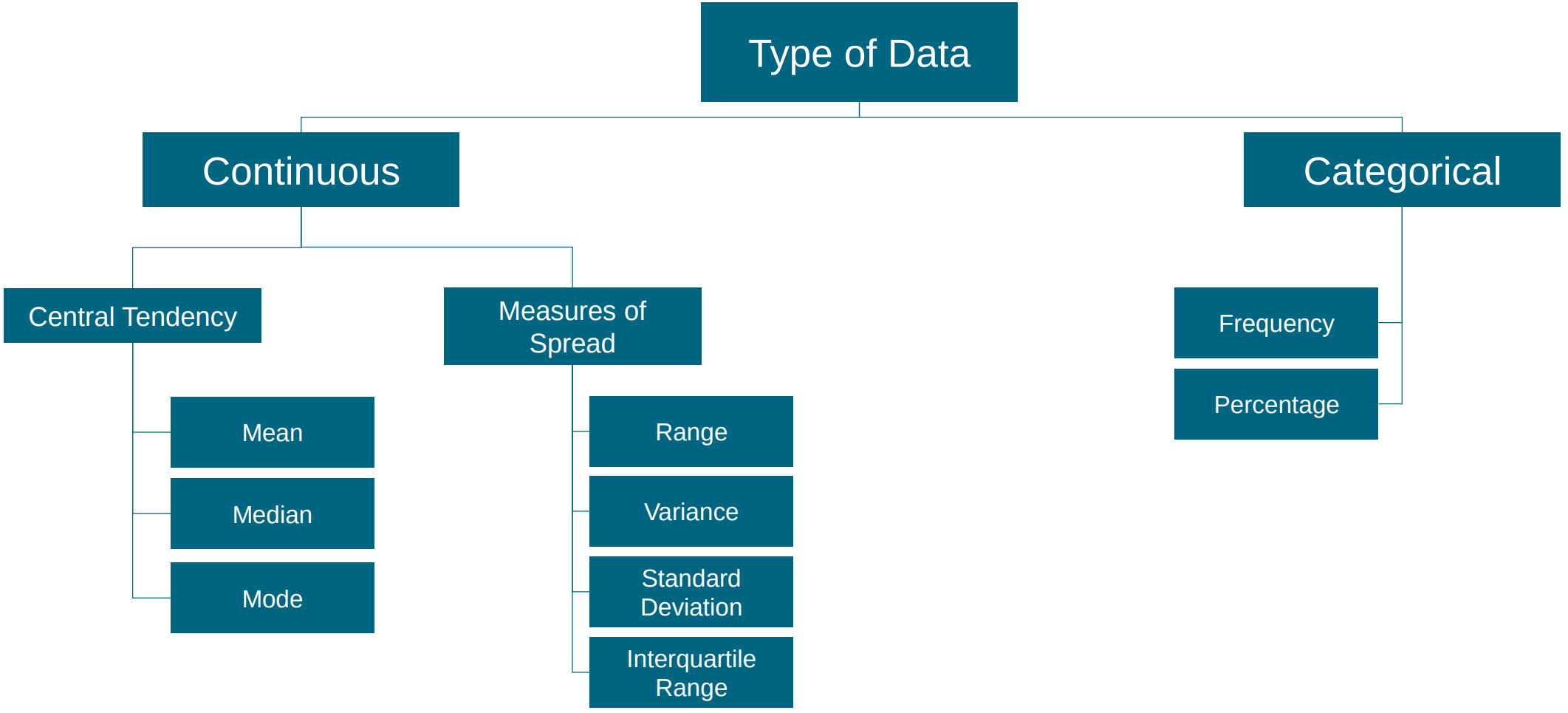
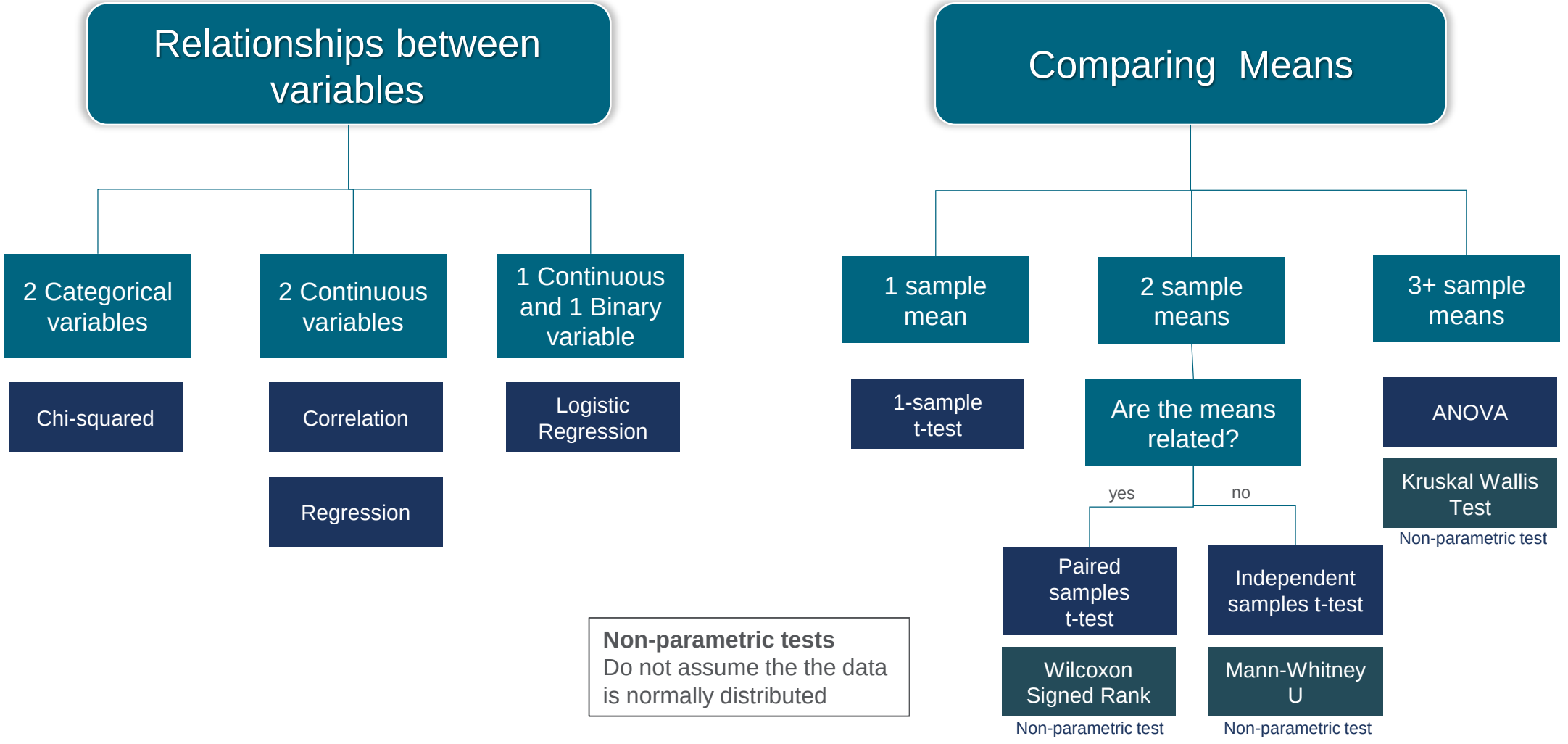
Analytics Framework



Types of Descriptive Analytics



Types of Inferential Analytics



Inferential Analytics - Terminology

Population

An entire pool from which a sample is drawn. Samples are used in statistics because of how difficult it can be to study an entire population.

Sample

A subset of the population we care about.

Null Hypothesis

A statement that suggests that (statistically) there is no difference between two things.

Alternate Hypothesis

A statement that suggests that there is a difference between two things and that this difference is not due to random chance.

Significance Level

A decision criterion that specifies the degree of certainty with which you want to make your judgment to accept or reject the null hypothesis.

Statistically Significant

A result that is unlikely to be caused by random variation or errors.

Predictive Analytic Techniques

Linear Programming

Linear programming are models to help optimize a given variable. For example, figure out an investment strategy to maximize profit.

Breakeven Analysis

A modeling technique used for profitability analysis to choose the point for both units of something produced and sold where total cost and total revenue are equal.

Crossover Analysis

A modeling technique where you choose an option from multiple options that will influence some output, like minimizing cost or maximizing profit.

Cluster Analysis

The task of grouping a set of objects in such a way that objects in the same group, called a cluster, are more similar to each other than to those in other clusters.

Simulation

A mathematical imitation of a real-world system. A simulation can also be considered to be an experimental process.

Decision Tree

A pictorial description of a well-defined decision problem. It is a graphical representation consisting of nodes (where decisions are made or chance events occur) and arcs (which connect nodes).

Markov Analysis

Determines the probability of future occurrences of an event by analyzing presently known probabilities of those events. It involves defining the likelihood of a future action given the current state of a variable.

Sentiment Analysis

Organizes unstructured data into various categories which can be analyzed to show positive or negative sentiment. Commonly uses natural language processing to data mine the unstructured content