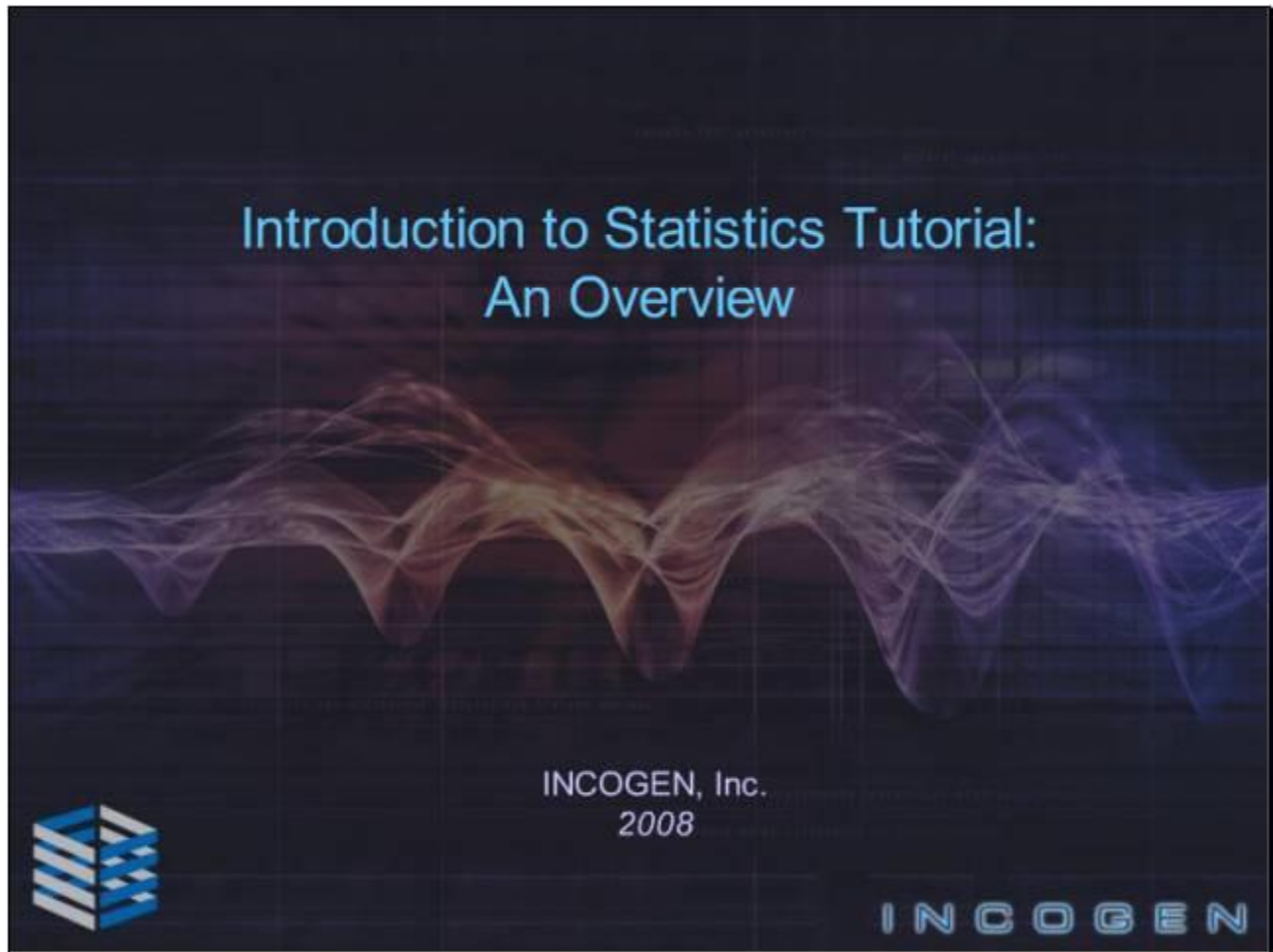


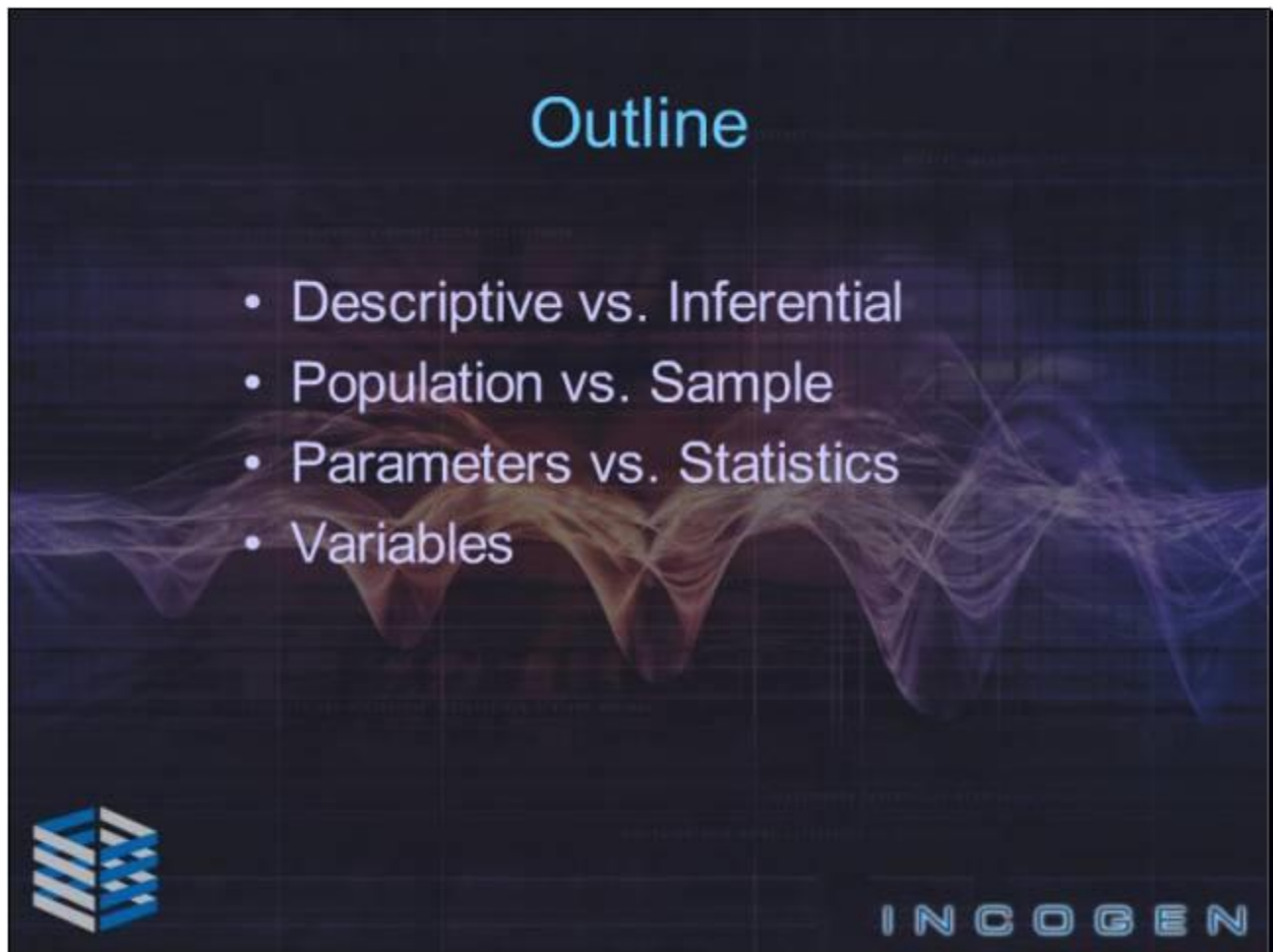
**Slide 1 - Introduction to Statistics Tutorial: An Overview**



**Slide notes**

Introduction to Statistics Tutorial: An Overview. This tutorial is the first in a series of several tutorials that introduce probability and statistics. Here we will concentrate on what is statistics and give a foundation of definitions that will be useful throughout this series of tutorials.


## Slide 2 - Outline



The slide features a dark blue background with a grid pattern and abstract, glowing wave-like patterns in shades of purple and blue. The word "Outline" is centered at the top in a light blue, sans-serif font. Below it, a bulleted list contains four items: "Descriptive vs. Inferential", "Population vs. Sample", "Parameters vs. Statistics", and "Variables". In the bottom left corner, there is a logo consisting of a 3D cube-like structure made of blue and white lines. In the bottom right corner, the word "INCOGEN" is written in a light blue, spaced-out, sans-serif font.

# Outline

- Descriptive vs. Inferential
- Population vs. Sample
- Parameters vs. Statistics
- Variables



INCOGEN



### Slide notes

Throughout this short overview, we will present definitions that will be important for all of the statistics tutorials and for a general practical understanding. We will discuss Descriptive vs. Inferential statistics, and definitions explaining Population vs. Sample, Parameters vs. Statistics, and Variables.

# Statistics

## Two Meanings

1. Specific numbers - numerical measurement determined by a set of data
2. Method of analysis - a collection of methods for planning experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions based on the data



### Slide notes

The word statistics has two meanings. One, statistics can be specific numbers such as numerical measurements determined from a set of data. For example, 2 out of 3 dentists recommend sugarless gum for their patients that chew gum. It can also be a method of analysis, or a collection of methods for planning experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions based on the data.



# Statistics

Study Question: Are there any differences in math scores between males and females?

Statistical Question: Is there a *significant mean difference* in math scores between males and females?

Statistical Conclusion: *Reject the null hypothesis* that there are no differences between males and females.

Study Conclusion: Evidence suggests that females tend to score higher than males in math.





## Slide notes

Statistics provides us with a mathematical technique for drawing conclusions from collected evidence. There may be a researcher who has the question: Are there any differences in math scores between males and females? So in statistical terms that question is phrased: Is there a significant mean difference in math scores between males and females? The researcher then collects the data: randomly selects individuals (males and females) and their test scores. The researcher takes the data and does a statistical analysis and makes the statistical conclusion that we Reject the null hypothesis that there are no differences between males and females. And with further analysis may conclude that the Evidence suggests that females tend to score higher than males in math.

# Statistics

Goals:

1. **Descriptive** – to organize and summarize data so that they are more readily understood
2. **Inferential** – to take into account the chance factors associated with sampling and subsequently draw valid conclusions





**Slide notes**

There are two branches of statistics: Descriptive and Inferential. The goal of descriptive statistics is to organize and summarize data so that they are more readily understood. Examples of descriptive statistics are mean, median, standard deviations, and plots. The goal of inferential statistics is to take into account the chance factors associated with sampling and subsequently draw valid conclusions. Inferential statistics includes significance testing such as the t-test.

## Definitions

**Population** - the complete collection of all elements to be studied. The collection is complete in that it represents all measurement of interest to the investigator.

**Sample** - a subset of elements drawn from a population.





### Slide notes

Understanding the difference between population and sample is a key importance to statistics. A Population is the complete collection of all elements to be studied. The collection is complete in that it represents all measurement of interest to the investigator. The Sample is a subset of elements drawn from a population. Using statistics we try to describe or predict the behavior of the population on the basis of information obtained from a representative sample from that population. Imagine trying to get the heights of all US women over the age of 18. You can't measure all of them (which is the population), but you can randomly look at a sample of them to describe the population.

## Definitions

- **Parameter** - a numerical measurement describing some characteristic of a population
- **Statistic** - a numerical measurement describing some characteristic of a sample





**Slide notes**

Understanding the difference between a parameter and statistic is also very important. A Parameter is a numerical measurement describing some characteristic of a population. A Statistic is a numerical measurement describing some characteristic of a sample. In our discussion of US Women's heights, a parameter (such as the population mean) would be the actual number calculated if you could collect all US women's heights. The Statistic is the value you get when you take information just from the sample (such as a sample mean). The statistic estimates the parameter.

# Definitions: Parameter and Statistics

Parameter ↔ population  
Statistic ↔ sample

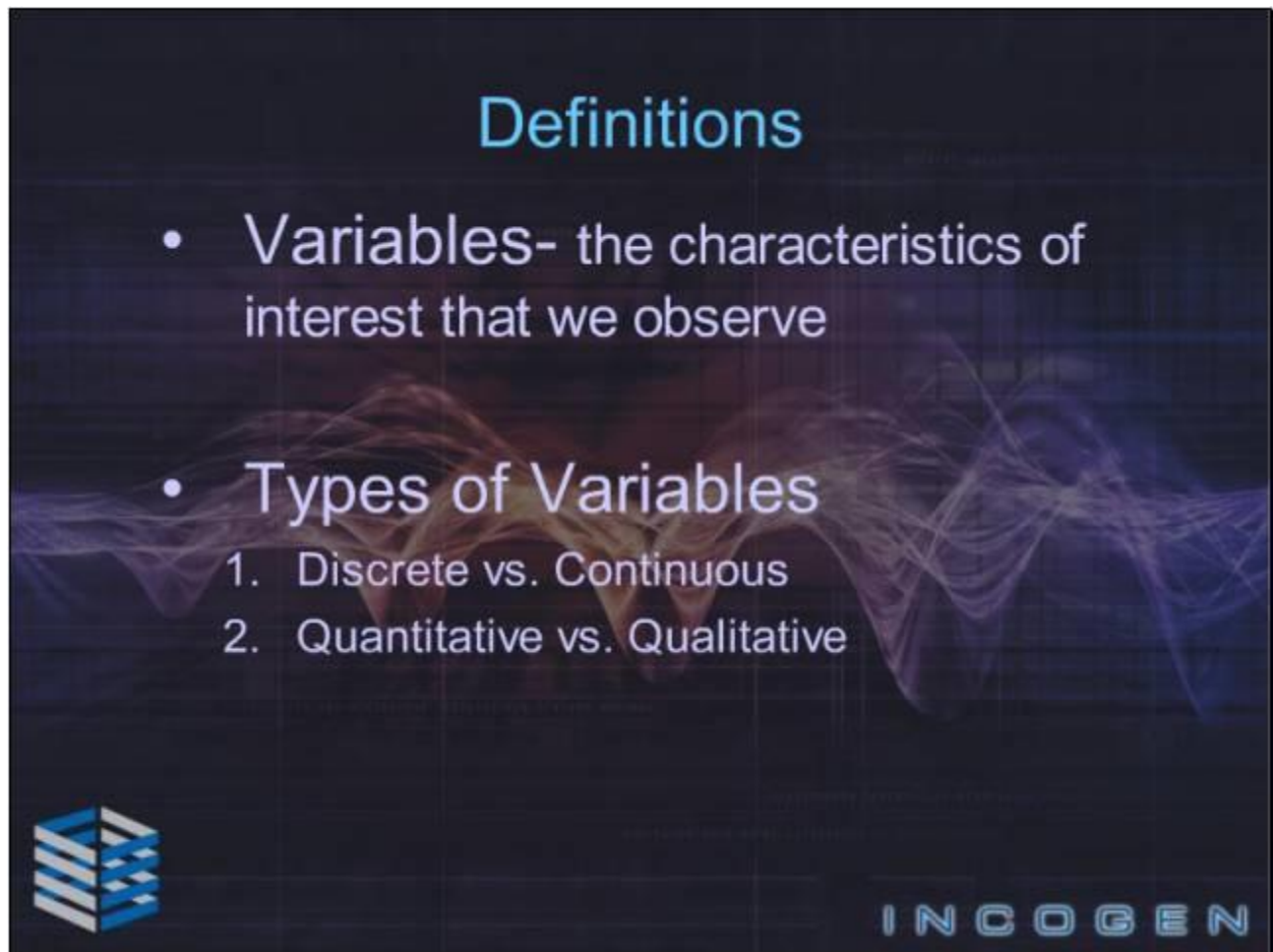
For example,  
Population Mean,  $\mu$  = average of the entire population of interest.  
Sample Mean,  $\bar{x}$  = average of the sample observations.



**Slide notes**



Again, a parameter describes the population and a statistic describes the sample. For example, the population mean (represented by the Greek letter "mu") is the average of the entire population of interest. The sample mean, (represented by "x-bar") is the average of the sample observations. So, the sample mean estimates the population mean. In other words, the statistic estimates the parameter.





# Definitions

- Variables- the characteristics of interest that we observe
- Types of Variables
  1. Discrete vs. Continuous
  2. Quantitative vs. Qualitative



## Slide notes



A variable is a characteristic that we wish to observe; it changes or varies over time or different individuals or objects. An example is measuring body temperature – it varies over time for a single person and it also varies from person to person. We do not say that every person alive has a temperature of 98.6 degrees Fahrenheit...we usually mean that 98.6 is the average healthy temperature.

There are different types of measurements that can be variables. There are discrete and continuous variables. Discrete variables are countable values such as the number of family members or number of defective toys returned to a store. Continuous variables are the infinitely many values corresponding to the number line such as height, weight or time. You can also categorize variables as quantitative or qualitative. Quantitative variables are numerical observations that represent an amount or quantity. An example of quantitative data would be cost of a gallon of gas. Qualitative variables, on the other hand, are observations that represent a non-numerical quality or characteristic such as marital status or religious affiliation.

# References

This tutorial is comprised of materials from the following sources:

- Introduction to Probability and Statistics by Mendenhall and Beaver. ITP/Duxbury.
- The Cartoon Guide to Statistics by Gonick and Smith. HarperCollins.
- Elementary Statistics by Triola. Addison-Wesley-Longman
- Basic Statistics: an abbreviated overview by Ackerman, Bartz, and Deville. 2006 Accountability Conference



**Slide notes**

This concludes our overview tutorial, which was comprised of materials from the listed resources.