**Handout #1: Introduction of Statistics**  
  
1. What is Statistics?  
Statistics is the science of

2. Statistical process in learning from data: Scientific method scheme

Decision:  
written conclusions,  
oral presentations

Formulate research goal:  
research hypothesis, models

Inferences:  
estimation, hypothesis testing,

Summarize data:  
 Graphs, numerical summary

Plan study:  
sample size, variables, experimental units,   
sampling mechanism

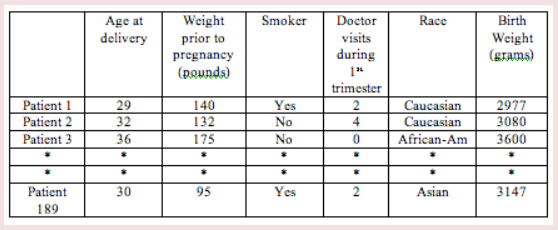
Formulate new research goals:  
new models,   
new hypothesis

Collect sample data and perform data management

**Variable:**Collecting data means collecting information from sampled individual. The characteristic (s) of an individual on which data is collected is called . Two types of variables:   
  
Quantitative –  
  
Qualitative –

**Units:** Subject or objects on which the data is collected.

**Example (Work together):** This dataset is from a medical study. In this study, researchers wanted to identify variables connected to low birth weights.

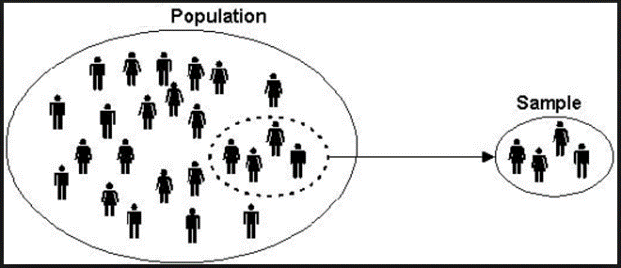


a. Units:   
  
b. Identify the variables and their type:

**Population vs Sample**

Population: A large group consisting of of a defined group that we are studying. Populations are not just people. Populations may consist of, but are not limited to, people, animals, businesses, buildings, motor vehicles, farms, objects or events.   
 **Examples:**ALL people living in the US   
All registered voters in Spartanburg county  
All Americans who played golf at least once in the past year  
All widgets produced last Tuesday by the Acme Widget Company  
All daily maximum temperatures in July for major U.S. cities

A sample is a of units in a population, selected to represent all units in a population of interest. It is a partial enumeration.



**Example:** Identify the population and the sample.

a. A survey of 1015 U.S. adults found that 32% have had to put off medical care for themselves or their family in the past year due to the cost.

b. A survey of 55 U.S. law firms found that the average hourly billing rate was $425.

c. A survey of 496 students at a college found that 10% planned on traveling out of the country during spring break.

There are various sampling techniques to obtain a sample. All we want is a random sample, which is representative sample.

**Parameter and Statistic:**  
Parameters are numbers that summarize data for an .

Statistics are numbers that summarize data from a   
  
Example 1: A national organization of personnel managers has estimated that about 25% of all resumes contain a major fabrication. Is 25 the value of a parameter or a statistic?

Example 2: Consider the problem of estimating the average grade point average (GPA) of the 750 seniors at a college.  
  
a. What is the population? How many data values are in the population?

b. What is the parameter of interest?

c. Suppose that a sample of 10 seniors is selected, and their GPAs are 2.72, 2.81, 2.65, 2.69, 3.17, 2.74, 2.57, 2.17, 3.48, 3.10. Calculate a statistic that you would use to estimate the parameter.

**Practice:** For each study, identify both the parameter and the statistic in the study.

 1. A researcher wants to estimate the average height of women aged 20 years or older. From a simple random sample of 45 women, the researcher obtains a sample mean height of 63.9 inches.

2. A nutritionist wants to estimate the mean amount of sodium consumed by children under the age of 10. From a random sample of 75 children under the age of 10, the nutritionist obtains a sample mean of 2993 milligrams of sodium consumed.

**Note: Sample data is used to make an inference about population**. **Example:** How do you describe the following sample results and what inference could you make about the population.  
  
a. Of 350 randomly selected people in the town of Luserna, Italy, 280 people had the last name Nicolussi.

b. On the last 3 Sundays, Henry D. Carsalesman sold 2, 1, and 0 new cars respectively.