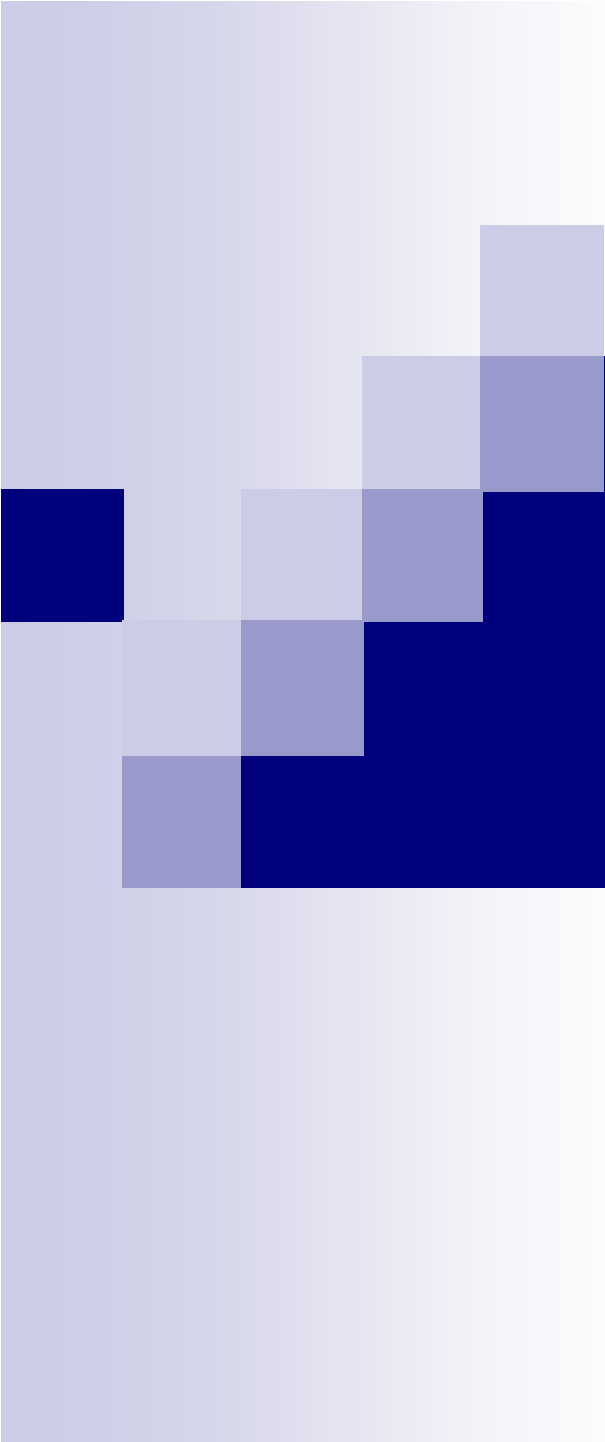
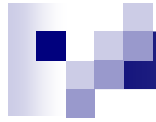




Chapter 1

Statistics



Chapter 1 Overview

Introduction

1-1 What is Statistics?

1-2 Measurability & Variability

1-3 Data Collection

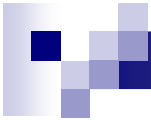
1-4 Statistics & Technology



Introduction

Examples of Statistics

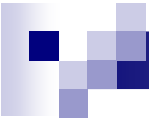
- About 60% of RN jobs are in hospitals.
- Median annual wages of registered nurses were \$62,450 in May 2008.
- The majority of college students have been women since 1979, and 30 percent of women held a bachelor's degree or higher in 2001 compared to 28 percent of men.
- After five consecutive years of annual increases, real median household income did not change significantly between 1999 and 2000.



Introduction

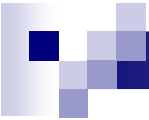
Why do we need to study Statistics?

- Statistics are everywhere – it is important to be able to interpret, understand, and analyze them (especially when they have been misused)
- You may need to conduct statistics research or analyze statistics in your field of study or job – they are very commonly used in science & business/economics



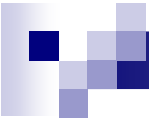
1-1 What is Statistics?

- **Statistics** is the science of conducting studies to
 - collect,
 - organize,
 - summarize,
 - analyze, and
 - draw conclusions from data.
- The field of statistics can be roughly subdivided into two areas: descriptive statistics and inferential statistics.



1-1 What is Statistics?

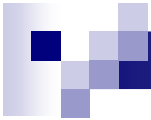
- **Descriptive statistics** consists of the collection, organization, summarization, and presentation of data.
 - The number of people under 20 living in a region.
 - A chart that shows the percent of people employed by different industries.
 - A table showing the GPA breakdown of students by year.



1-1 What is Statistics?

■ **Inferential statistics** refers to the technique of interpreting the values resulting from the descriptive techniques and making decisions and drawing conclusions about the population.

- After conducting a study, it was found that more females like Coke over Pepsi.
- A link between gingivitis and heart disease exists.
- There is a 40% chance of rain tomorrow.



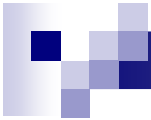
Class Activity

- Identify the following as either descriptive or inferential statistics.
 - ☐ The average age of 500 surveyed students in your institution is 23 years.
 - ☐ Based on a sample of 10,000 Americans, it is fair to say that 70% of all Americans are overweight.
 - ☐ Based on a study of 100,000 people, the drug Tylenol appears to have no adverse side effects.
 - ☐ A correlation has been found between lack of exercise and heart disease.



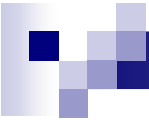
1-1 What is Statistics?

- Before we can study statistics in more detail, we need to be familiar with the vocabulary.



1-1 Descriptive and Inferential Statistics

- A **population** consists of all subjects (human or otherwise) that are studied
 - If we are studying the number of people employed in the U.S., then the population is everyone in the U.S.
 - If we are studying the survival rate of cancer patients, then the population is all cancer patients
 - If we are studying the average failure rate of engines, the population is all engines



1-1 Descriptive and Inferential Statistics

- A **sample** is a subset of the population.
 - When selecting a sample, we want to select a subset that is representative of the population (avoid *bias*).
 - We usually do this by selecting a *random* sample – one in which every unit in the population has an equal chance of being selected for the sample.
 - Ex.) drawing names out of a hat, flipping a coin, etc.



1-1 What is Statistics?

- A **variable** is a characteristic or attribute that can assume different values.
 - From algebra, the variable x as in the equation $2x + 3 = 6$
 - For a car driving up and down a hill, a variable could be the speed of the car
 - In a study of how different doses of a drug affect the severity of symptoms, the variables are the frequency and intensity of symptoms and the different doses administered



1-1 What is Statistics?

- The values that a variable can assume are called **data**.
 - The data for the speed of a car driving from 0 to 20 seconds is:

Time (s)	Speed (mph)
0	0
10	30
20	40

- This table would be called a **data set** since it a collection of data values.



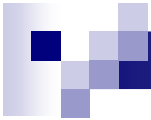
1-1 What is Statistics?

- An **Experiment** is a planned activity whose results yield a set of data.
 - An experiment includes the activities for both selecting the elements and obtaining the data values.



1-1 What is Statistics?

- A **Parameter** is a numerical value summarizing all the data of an entire population.
 - The “average” age at time of admission for all students who have ever attended our college and the “proportion” of students who were older than 21 years of age when they entered college are examples of two population parameters.
 - A parameter is a value that describes the entire population. Often a Greek letter is used to symbolize the name of a parameter.

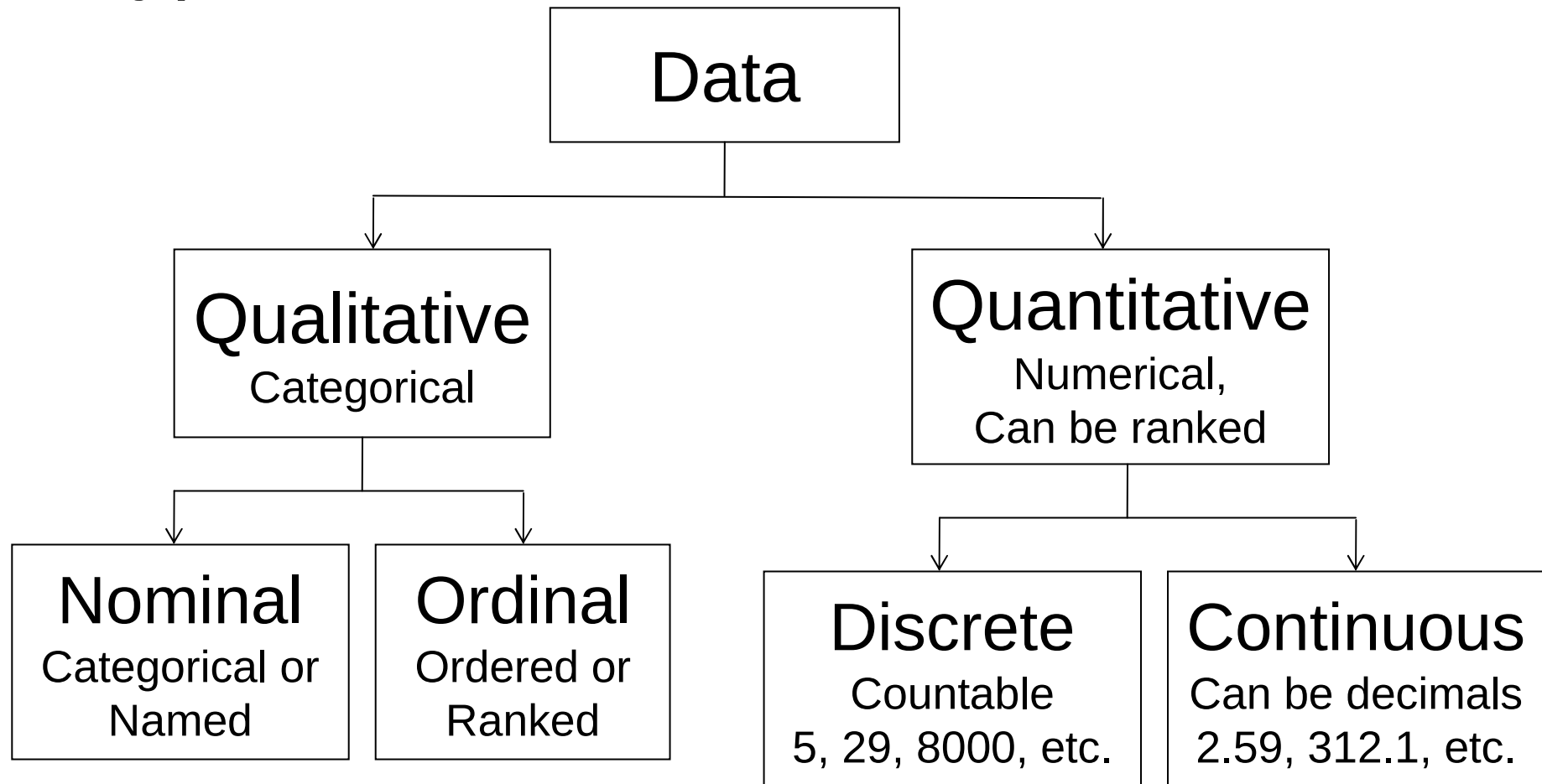


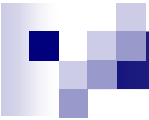
Class Activity

- Consider the following: As a pediatric nurse you are promoting routine childhood vaccinations, so you are studying the number of children vaccinated against certain diseases. In 2010, 91.5% of children in the United States aged 19-35 months received the MMR vaccine.
 - ☐ What is the population of interest?
 - ☐ What is the sample?
 - ☐ Identify the parameter of interest.
 - ☐ Identify the statistic and give it's value.
 - ☐ Do we know the value of the parameter?

1-1 What is Statistics?

Types of Data

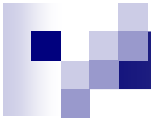




1-1 What is Statistics?

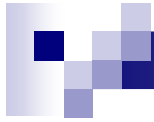
Types of Data

	Qualitative		Quantitative	
Variable	Nominal	Ordinal	Discrete	Continuous
Hair Color	X			
Zip Code	X			
Letter Grade		X		
ACT Score			X	
Height				X
Age				X
Temperature (F)				X



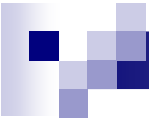
Class Activity

- Identify the following as either qualitative or quantitative. If it is quantitative, determine if it is discrete or continuous. If it is qualitative, determine if it is nominal or ordinal.
 - ☐ The number of children vaccinated against certain diseases.
 - ☐ The race of children being vaccinated.
 - ☐ The amount of time symptoms last after a child is vaccinated.
 - ☐ The birth order of the child being vaccinated (ie oldest, middle, youngest).



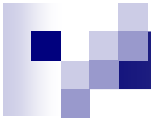
1-2 Measurability and Variability

- Within a set of measurement data, we always expect variation. If little or no variation is found, we would guess that the measuring device is not calibrated with a small enough unit.
- For example, we take a carton of a favorite candy bar and weigh each bar individually. We observe that each of the 24 candy bars weighs $\frac{7}{8}$ ounce, to the nearest $\frac{1}{8}$ ounce.
- Does this mean that the bars are all identical in weight?



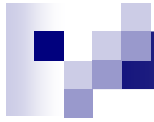
1-2 Measurability and Variability

- Not really! Suppose we were to weigh them on an analytical balance that weighs to the nearest ten-thousandth of an ounce. Now the 24 weights will most likely show **variability**.
- It does not matter what the response variable is; there will most likely be variability in the data if the tool of measurement is precise enough.
- One of the primary objectives of statistical analysis is measuring variability.



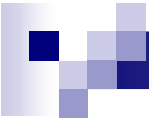
1-2 Measurability and Variability

- For example, in the study of quality control, measuring variability is absolutely essential.
- Controlling (or reducing) the variability in a manufacturing process is a field all its own—namely, statistical process control.



Class Activity

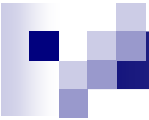
- Consider the following: As a pediatric nurse you are recording length and severity of symptoms of children being vaccinated against MMR. How much variability do you expect to see in these variables?



1-3 Data Collection

Data Collection

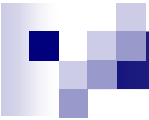
- **Surveys** – be careful using surveys, it is very easy to be biased.
 - Write the survey carefully
 - Don't be biased in the way you ask questions
 - Make sure to collect surveys from a good representation of the population
 - Voluntary surveys usually have the disadvantage of being very biased – ex.) customer satisfaction surveys



1-3 Data Collection

Some Sampling Techniques

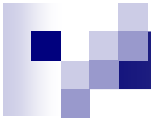
- **Random** – random number generator
 - Drawing a name out of a hat
 - Having a computer program randomly select
- **Systematic** – every k^{th} subject
 - Randomly select the 1st subject. Number the population from 1 to n . Choose every k^{th} subject from there. (careful about bias again)



1-3 Data Collection

Some Sampling Techniques

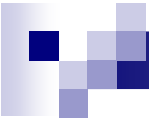
- **Stratified** – divide population into “layers”
 - If we are interested in certain aspects of students who get different letter grades, we would divide the students into A students, B students, C students, D students, and F students and randomly sample from each of these subgroups



1-3 Data Collection

Some Sampling Techniques

- **Cluster** – use intact groups
 - If we were studying people in Bourbonnais using this technique, we would select n neighborhoods & interview everyone in each of these neighborhoods.
 - Be careful about bias again – the neighborhoods we select may not be a good representation of all of Bourbonnais.



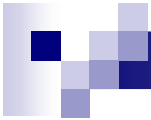
1-3 Data Collection

Some Sampling Techniques

■ **Convenient** – mall surveys

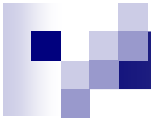
- Remember this method is probably easier & cheaper, but is most likely biased.

- Time of day, time of year, where are you standing in the mall, etc.



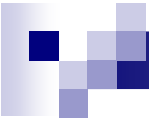
1-3 Data Collection

Remember: If our sampling method is not good, then our study is open to a large amount of criticism. It is also very difficult to draw conclusions from the study.



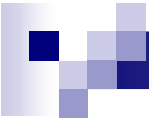
1-3 Data Collection

- In an **observational study**, the researcher merely observes and tries to draw conclusions based on the observations.
 - Ex.) Give surveys and draw conclusions.
 - The researcher does not interfere in any way.
 - It is often more difficult to draw strong conclusions – these are mainly used to find descriptive statistics (they can still be used for inferential)



1-3 Data Collection

- The researcher manipulates the **independent (explanatory) variable** and tries to determine how the manipulation influences the **dependent (outcome) variable** in an **experimental study**.
 - Ex.) Set up & study 2 groups – 1 which takes a particular drug and the other which takes a placebo.
 - Researcher has complete control over the study.
 - It is easier to draw strong conclusions here – we can use this kind of study to find descriptive & inferential statistics.



Class Activity

As a pediatric nurse you are recording length and severity of symptoms of children being vaccinated against MMR. You would like a gauge of what is normal so that severe symptoms can be immediately treated.

You decide to do the following:

On your shifts this week at ABC pediatrics, you take the information of each child that is vaccinated and call them 3 days later to record length and severity of symptoms.

- ☐ What kind of sampling did you use?
- ☐ Is there likely to be bias?
- ☐ Can you think of a better way to sample?



1-4 Statistics & Technology

- **Microsoft Excel**
- **Microsoft Excel with MegaStat**
- **TI-83/84**
- **Minitab**
- **SAS**
- **SPSS**