Algebra 2
CHAPTER 11
Statistics

Name _____

Period _____ Date _____

Reading: Different Types of Sampling

Have you ever had a piece of cake? How many bites did it take to decide if the cake was good or not? Did you have to eat the whole cake before you knew what it tasted like? Of course not! Using a small part of something to learn about the whole is called *sampling.* To take a *sample*, you choose a few individuals from a clearly defined larger group, called the *population.* When done correctly, sampling allows you to use a small group to draw conclusions about the entire population. Sometimes, it is simply easier to take a sample. For example, if you wanted to know what students at your school thought about a new cafeteria menu, you could learn what everyone thinks by asking only a few people. At other times, it is simply impossible to measure the whole population. Suppose you want to know how long car tires last. If you test every tire, you won't have any tires left to use! Sampling is a very powerful tool, saving time and money, but you have to be very careful about how you choose your sample.

Representative vs. non-representative samples

If you are not careful about how you choose your sample, you can introduce *bias* into your data. *Bias* is a systematic favoring of one outcome. If you only taste the frosting of the cake, you would conclude the cake is too sweet. This would be a *sampling bias* toward sweetness. To avoid bias, you want to have a sample that is *representative* of the whole population. A representative sample is a small group from the population that is like the population in some important way. A representative sample of cake would include more than just frosting; it would have bites from the middle, bottom and edges as well. There are many ways to choose a representative sample; however, there are also many ways to draw a non-representative sample. It is very important to think about how the sample was chosen. If a sample does not represent the population, it will give incorrect information. How the data was collected can be just as important as what the data say.

You are probably more familiar with sampling than you think. A popular TV program lets viewers call in to vote on a favorite singer. This is a type of sampling, because they use the people who called to decide which singer the viewers like. Is this sample representative of all viewers or are some viewers more likely than others to call in? Of course people who really like one particular singer, or really dislike a singer, will take the time to call in. So singers with a small, but passionate fan-base might get more votes than a singer with more fans. When

people decide to participate in the survey or not, it is called a *self-selected* or *voluntary response sample.* Many people and programs use it, like news shows, newspapers, and websites, even though it always results in non-representative samples.

Have you ever collected information about the students in your class? Are the students in your class representative of all people? Or are they all the same in some way? Selecting a sample by choosing subject that are close-by or easy to ask is called *convenience sampling*. Another example of convenience sampling is listening only to your friends. Friends tend to agree with each other about major topics, so if you ask only your friends about politics, you might only get one side of the story.

Statisticians' favorite sampling method is *random sampling.* In random sampling, the sample is chosen by rolling a dice, or choosing names out of a hat, or using a random number generator on a table. Remember all those "picking marbles" and "matching socks" problems from math class? Those were probability problems! When a sample is drawn at random, we can solve probability problems and calculate the chances that a sample is non-representative. Random sampling is difficult to do, because you need a list of the population, which can be difficult to make. On the other hand, the randomness introduces probabilities, and we can apply all sorts of mathematics to learn more about the sample.

When looking at a data set, it is important to think about how the data was collected. Some data collection methods result in non-representative sample, which results in data that does not accurately reflect the population. Always ask how the data was collected, and if the sample was representative of the population.

Different types of samples

Self-selected (or Volunteer) sample — subjects from the chosen population volunteer to be in the sample.	Convenience sample – easy-to-reach subjects from the population are selected to be in the sample.
Example: send out invitation to female high school volleyball players to join the study and only selecting those that respond to the invitation.	Example: select the female volleyball players at your high school.
Systematic sample – a rule is used to select subjects from the population to be in the sample (such as every 5^{th} person).	Random sample – each subject from the population has an equal chance of being chosen for the sample.
Example: You get list of all the female high school volleyball players in the country and selected every 30^{th} person on the list.	Example: You get a list of all the female high school volleyball players in the country and you assign each of them a number. Then you use a random number generator to create however many numbers you need. You then include the female volleyball players associated with the random numbers that were created.