Name: Key
8A; Algebra 1

Date _____
Period _____

What Do We Know About The Collection of Data?

Statistics: the study of numerical data

Data: facts given, numerical information

The typical steps in a statistical study include:

Step1 : The collection of data

Step 2: The organization of this data into tables, charts, and graphs

Step 3: The drawing of conclusions from an analysis of this data.

* These three steps, which describe and summarize a set of data are often called descriptive statistic

st 3 ways that data may be collected:

1) Questionnaire

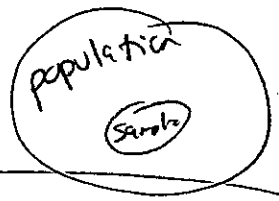
2) Interview - either in person or by phone - answers given verbally and recorded by person asking questio
(email / social network)

3) Log or Diary - person records information on a regular basis.
ex: hospital chart, hourly recording of the outdoor temperatu

* A portion of items to be counted in statistical studies is called a sample

entire thing = population

Techniques of Sampling



1) The sample must be fair to reflect the entire population being studied.

2) The sample must contain a reasonable number of items being tested or counted

3) Patterns of sampling or random selection should be employed in a study.

Biased → NOT fair

unbiased → fair

Possibly Biased Data

It is often possible that different research studies concerning the same issue can produce very dissimilar or contradictory results. How can this happen? Aren't all statistics "true"?

Statistics are influenced by a multitude of factors. It is even possible that statistics can be manipulated so that they tell the story that the person using them wants to tell.

Because of these influencing factors, it is important to understand how to evaluate statistical information.

When viewing statistics, you should consider:

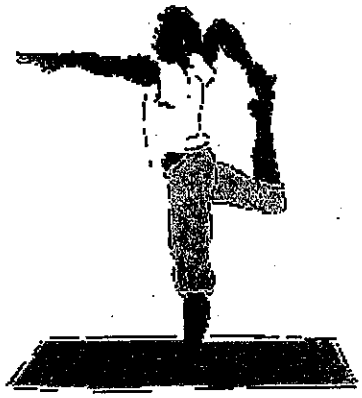
• Who collected the data? Does the group collecting the data have an interest in how the results turn out?	• **Example:** A study on the hazards of cigarette smoking being done by a tobacco company. (may not be reliable findings - conflict of interest)
• Is the study a recent study, or did it occur decades ago? Could recent developments have changed the findings?	• **Example:** Decades past, second-hand cigarette smoke was found to not be hazardous. More recent findings prove that this is not true. (findings should be current)
• What is the sample size of the study? How many people/items were studied?	• **Example:** A study is done on the favorite color of 14 year olds. The sample group for the study is Mrs. Smith's third period class containing 20 students. (too few participants to generalize a finding to all 14 year olds)
• Is the data from a primary source? Or has the data been "condensed" by another group?	• **Example:** The US Census Bureau collects data on US populations. A tabloid magazine publishes a synopsis of the findings. (the most reliable information comes from the original source - avoid the "Reader's Digest" condensed version by another publisher who may be interpreting the findings)
• Do the statistics show any bias?	• **Example:** The study of how many people can walk a balance beam is conducted with students from a gymnastics class. (the results are biased due to the very specific selection of the participants)

Sources of bias:

Selection bias:

In a statistical study, it is important that the smaller group used for the study (the sample) be truly representative of the larger group to whom the findings will be directed (the population). Preferably the sample group should be chosen at random.

Dexterity Study: 500 people are invited to a research center for an experiment in dexterity and flexibility. 100 of the people show up. The researchers document the number of people who can clasp their right foot with their right hand behind their backs, by reaching over their right shoulders (as seen at the right). They conclude that an amazing 62% of people can perform this act of dexterity and flexibility.



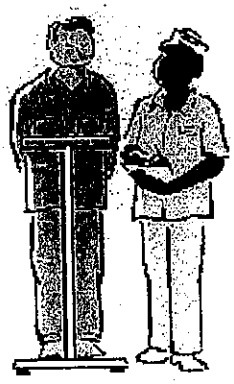
Are these finding reliable? What is wrong with this study?

Since only 100 of the 500 invited participants showed up for this study, this is not a representative sample. It may also be the case that the people who were confident about their dexterity and flexibility showed up and the results are biased in their favor.

Measurement bias:

In a statistical study, it is important that the means of gathering measured data be reliable, accurate and appropriate for the study.

Dexterity Study Weights: The people participating in the dexterity study mentioned above were weighed to determine if participants who were not overweight were more flexible. It was discovered after the study that the scale used for weighing had a tendency to list weights over 150 pounds inaccurately.



What affect did this have on the results of the study?

This situation presents two problems for the study. First, the weights are unreliable for participants weighing over 150 pounds. Second, the question should be asked if the participants' heights were also measured. If not, what was the definition of an "overweight" participant?

Examples: Biased - NOT Fair UnBiased - Fair

1) To determine which television shows are the most popular in a large city, a poll is conducted by selecting people at random at a street corner and interviewing them. Outside of which location would we find the most fair sample?

(a) a ball park (b) a concert hall (c) a supermarket

2) A sample of students is to be selected and the weight of each student is to be measured to determine the average weight of a student in high school. Tell if the sample is fair or unfair and if it is unfair explain why.

(a) the freshman class – unfair (biased) (unbiased / biased)
(b) all 15-year-old students unfair (biased)
(c) all girls – unfair (Biased)
(d) the wrestling team – unfair (Biased)
(e) the girls dance class – unfair (Biased)
(f) every fifth person selected from a list of students ordered by their social security number Fair (unbiased) (Random)
(g) the first five students to enter school Monday morning. – unfair (Biased)
(h) The first five students who enter each classroom Monday morning.

Fair (unbiased) variety (random)

Not large enough sample

Name _____ Date _____

8A; Algebra 1 Period _____

What is Qualitative vs Quantitative Data?
How Can Data Be Biased?

Quantitative Data

- Deals with numbers.
- Also referred to as Numerical Data.
- Data which can be measured.
- Height, weight, area, volume, length, time, temperature, speed, cost, etc.
- Quantitative → Quantity

Example 1:

Candy Bar

Quantitative Data:

- weight 1.83 ounces
- 280 calories
- length 10 cm
- width 3 cm
- height 1.8 cm

Example 2:

Spanish Club

Quantitative Data:

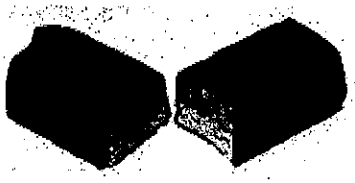
- 38 students
- 3 field trips per year
- average GPA 3.5
- 20 girls, 18 boys
- 3 foreign exchange students

Qualitative Data

- Deals with names, labels, descriptions.
- Also referred to as Categorical Data.
- Data which can not measured.
- Eye color, smells, car models, textures, tastes, favorites, candy bars, etc.
- Qualitative → Quality

Example 1:

Candy Bar



Qualitative Data:

- dark chocolate
- contains peanuts
- caramel smell
- brown wrapper
- nougat center

Example 2:

Spanish Club

Qualitative Data:

- charity work
- friendly atmosphere
- vocal concerts
- produce a Spanish Play
- enjoy Spanish food

Example 3:

Cocker Spaniel Puppy

Quantitative Data:

- adult weight 28 pounds
- life span 15 years
- height 15 inches
- hip dysplasia ranking 115 *good
- shelter price $200

Example 3:

Cocker Spaniel Puppy

Qualitative Data:

- color black
- trusting
- fluffy
- baby smell
- likes to be held

• Number of Variables in Data:

Univariate data means "one variable" (one type of data).
Bivariate data means "two variables" (two types of data).

1

Univariate Data	Bivariate Data
• Deals with one variable. • Major purpose is to describe. • No relationships or causes.	• Deals with two variables. • Major purpose is to explain. • Relationships and causes.
Statistical Analysis: ✱ Frequency ✱ • measures of central tendency - mean, mode, median • outliers and interquartile range • range, maximum, minimum, variance, quartiles, mean absolute deviation, standard deviation • shape, center, spread or distributions	*Statistical Analysis:* • correlations • comparison, causes, relationships, explanations • analysis of 2 variables simultaneously • tables showing one variable depending upon the other variable • independent and dependent variables
Displays: • Dot Plots • Histograms • Box Plots • Quartiles • MAD, Standard Deviation	*Displays:* • Two-Way Frequency Tables • Scatter Plots • Line of Best Fit • Linear/Quadratic Regressions • Residuals
Example: How many students in the freshman class own a skateboard?	*Example:* Is there a relationship between the number of skateboards a freshman owns and his/her final test score in Algebra 1?

✱ a comparison of 2 things

Chapter 16-1 Collecting Data

PART I

Answer all questions in this part.

1. Which of the variables below represents a qualitative variable?
   (1) The number of students in a school
   (2) The types of pets a family owns
   (3) The number of pets a family owns
   (4) The weight of a pet

2. A manager of a grocery store wants to know if his customers are happy with the bakery department. Which of these samples is most likely to give an unbiased result?
   (1) Customers in line at the bakery
   (2) Every fifth customer at the store checkout
   (3) Children who are in the store
   (4) Customers in line at the deli

3. At a sports training camp, participants run three miles a day. Before running, they record their resting heart rate. After running, they record their time, heart rate, and describe how they feel about their time. Which of the following is *not* a quantitative variable?
   (1) Resting heart rate
   (2) Time
   (3) Heart rate after running
   (4) How they feel about their running time

4. A study wants to determine the average number of calories eaten daily by boys on high school football teams. Which sample is most likely to give an unbiased result?
   (1) The seniors on one team
   (2) The quarterbacks on one team
   (3) The entire team at a school
   (4) The 10 smallest players from each team

5. A medical study was done to determine the effects of a new drug. Group A received the new drug for 6 weeks. Group B received a placebo. Group B is called the
   (1) control group
   (2) treatment group
   (3) placebo
   (4) placebo effect

6. To determine which type of music students in a particular school prefer, a poll is conducted by selecting students at random and interviewing them. Which location would the interviewer be most likely to find an unbiased sample?
   (1) The chorus room
   (2) The band room
   (3) The art room
   (4) The school cafeteria

7. Which of the variables below represents a quantitative variable?
   (1) Weight of a bag of apples
   (2) Types of apples in a bag
   (3) Color of apples in a bag
   (4) Types of bags offered at the grocery store