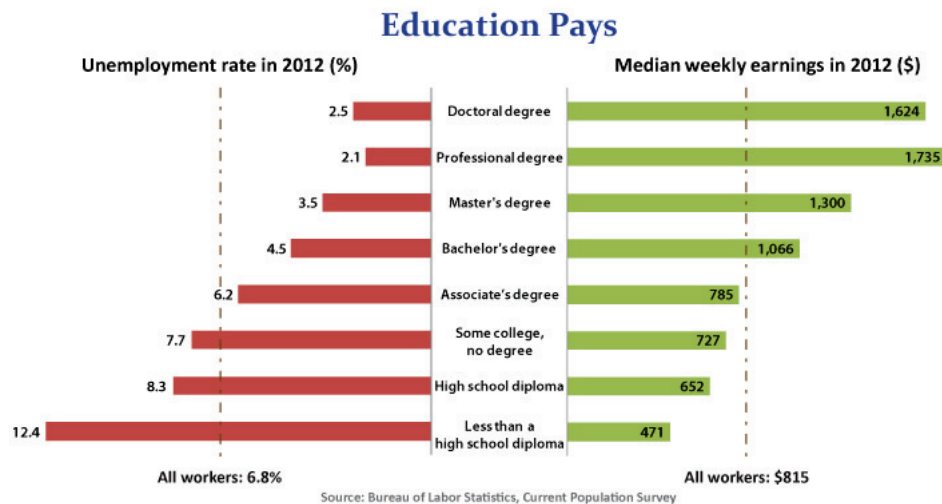# 1 Introduction

Here are some *statistics* that students have found. I don't know for certain if any are true. Without formally fact-checking these numbers, consider what is required to make these claims. Are they believable?

1. The average pineapple plant is 3.47 feet tall.

2. Married men live, on average, ten years longer than unmarried men.

3. Right-handed people live, on average, nine years longer than left-handed people.

4. 1 in 112,000,000 people will die from a vending machine accident in a year.

5. 1 in 289,200,000 people will die from a shark attack per year.

6. Girls have more taste-buds than boys.

7. The proportion of people who can roll their tongue is between 65 and 81 percent.

8. The average person spends 2 weeks of life waiting for traffic lights to change.

9. Americans throw out 27% of the 350,000,000 pounds of food they buy each year.

10. 12% of lightning strikes occur at golf courses.

11. There are about 45 million tattooed Americans. About 17% of them will come to regret it.

12. 100,000 dogs are killed each year by falling out of pick-up trucks.

13. Based on a 2012 public opinion poll, about one in two people believe that colleges are not affordable.

14. There is a strong correlation between education and income & unemployment. *

Education Pays

Unemployment rate in 2012 (%)		Median weekly earnings in 2012 ($)
2.5	Doctoral degree	1,624
2.1	Professional degree	1,735
3.5	Master's degree	1,300
4.5	Bachelor's degree	1,066
6.2	Associate's degree	785
7.7	Some college, no degree	727
8.3	High school diploma	652
12.4	Less than a high school diploma	471

All workers: 6.8% All workers: $815

Source: Bureau of Labor Statistics, Current Population Survey

*Published by the U.S. Bureau of Labor Statistics. http://www.bls.gov/emp/ep_chart_001.htm.

1.1   Statistics and Data

- **Definitions**

  - **Data** is a collection of observations about the members of a group - human or otherwise.
  - A **population** is the complete collection of all members in a group.
  - A **sample** is a sub-group of the population.
  - A **parameter** is a numerical description of some characteristic of a **population**.
  - A **statistic** is a numerical description of some characteristic of a **sample**.
  - **Statistics** (The Practice) comes in two different flavors
    - **Descriptive Statistics** is the collection, organization, and presentation of data.
    - **Inferential Statistics** is the art/science of making inferences (estimates, predictions) about population parameters from sample statistics.

    The link between these types of statistics is **Probability**.

- **Example:** Suppose I am in charge of lunches at Scooter's Summer Super-Fun Camp. I am looking to save a buck by serving a cheaper brand of macaroni and cheese than I currently serve because I suspect most of the kids won't be able to tell the difference. Before acting on this plan I decide to do a survey on a subgroup of kids at the camp. In a sample of 40 kids, it turns out that 26 were not able to tell the difference between the cheap stuff and the good stuff. I then, through the process described in Chapter 8, conclude that most kids at camp can not tell the difference but there is a 3% chance that I am wrong.

  - The **data** here consists of the results from the 40 kids in the survey.
  - The **population** I am concerned with is all of the kids at the camp.
  - The **sample** is the subgroup of 40 kids I select for my survey.
  - The **parameter** I seek is the percentage of all kids at camp who can tell the difference.
  - The **statistic** is that 65% (26/40) of kids in the sample can't tell the difference. This is a **descriptive statistic**.
  - The process of estimating the population parameter from the sample statistic is a form of **inferential statistics**. The 3% chance that I have reached the wrong conclusion is a **probability**.

- **Your Turn:** In a sample of 230 college students, the average number of hours slept per night is 6.2 hours. After analyzing the data, I am 90% confident that the average number of hours of sleep for all college students is between 5.9 and 6.5 hours per night. Determine the following:

  - What is the population I am studying?
  - What is the sample?
  - What is the statistic?
  - Describe the parameter we are seeking to find?
  - What is the estimate on this parameter?
  - What is the probability that my conclusion about the parameter is correct?

- **Qualitative and Quantitative Data**

  - **Qualitative (categorical) data** consist of non-numerical categories such as name, eye color, gender, college attended. Some appear to be numerical, such as ID numbers.

  - **Quantitative data** consist of numerical counts or measurements such as age, test-scores, rainfall, weight or the *number* of red cars. Quantitative Data can be subdivided as follows:

    - **Discrete data** can only take certain values within a given range - usually counts. There are gaps between possible data values. Examples include the number of cars sold by color, the number of children in a family, or calendar years (2011, 2012, ...).

    - **Continuous data** can take on any value in a given range - usually measurements such as time, length, volume, or weight. In between any two measurements exists another. Continuous data often appear to be discrete because of the measuring device.

- **Levels of measurements** (given here lowest to highest)

  1. **Nominal measurements** consist of categories, names, labels, etc., which can not be ordered, added, or averaged. Examples generally come from qualitative data but might be disguised numerically by an identification number of some sort.

  2. **Ordinal measurements** can be ordered (ranked) but the difference between measurements are not well defined. Examples: grades (A,B,C,D,F), hotel ratings, levels of pain.

  3. **Interval measurements** are like ordinal but better because differences are meaningful. However, zero is arbitrary and ratios do not make sense. Examples: dates, non-Kelvin temperatures.

  4. **Ratio measurements** are the best for numerical comparisons. Differences are meaningful, zero is not arbitrary, and ratios make sense. Examples: time, temperature in degrees Kelvin, counts, prices, weights, lengths, etc.

- **Examples:** Determine whether the given data is qualitative or quantitative. If it is quantitative, categorize it as discrete or continuous. Then, determine the level of measurement of the data collected.

  (a) The finishing times of the top 100 runners in the Boston Marathon.
      Answer: Quantitative, Continuous, Ratio

  (b) The ISBN numbers for all of your textbooks.
      Answer: While these are numerical, the numbers are acting as names so the data is qualitative and the level of measurement is nominal.

  (c) Each of 12 hotels are ranked by the number of stars.
      Answer: Quantitative, Discrete, Ordinal

  Your Turn

  (a) Each of 30 people in a stats class must categorize themselves as conservative, moderate, or liberal.

  (b) The number of donuts made by a baker on Sunday morning.

  (c) The high temperatures for each day this week in degrees Celsius.

1.2   Sampling

- **Census and Parameters -vs- Samples and Statistics**

  - When you gather data from every member of a population it is called a **census** and the results are called **parameters**.
  - When you gather data from a subgroup of a population it is called a **sample** and the results are called **statistics**.

- **Random and Simple Random Samples**

  - In a **random sample** every member of the population has an equal chance of being selected.
  - A **simple random sample** is a random sample where every sample of the same size has the same chance of being selected. There can be no sub-grouping of the population.
  - A simple random sample is always random but the reverse is not necessarily true.

- **Examples:**
  Classify each sampling method as simple random, random but not simple random, or neither.

  1. In a class with 200 males and 300 females, I want to select 50 students for a survey.
     (a) I randomly select 20 males and 30 females for the survey.
         Answer: This is random because all students have a 1/10 chance of being picked, but it is not simple random because I can't have a sample of say 25 males and 25 females.
     (b) I put all 500 students in a list and randomly select 50 students.
         Answer: This is simple random sample because everyone has an equal (1/10) chance of being picked **and** no sample of 50 has a better chance of being picked.
     (c) I randomly select 25 males and 25 females.
         Answer: Neither. Males have a 25/200 chance of being picked and females have a 25/300 chance of being picked. Not all students have the same probability of being picked so it is not even a random sample.

  2. **Your Turn:** In my class I have 30 students, there are 5 rows of 6 students each. I want to select 12 students for a survey. Classify each sampling method as simple random, random but not simple random, or neither.
     (a) I put all 30 students names in a basket and randomly select 12 students.

     (b) I randomly select two of the five rows and choose all students in each of these rows.

     (c) I randomly select one odd row and one even row and take all students in the chosen rows.

- **Sampling Strategies**

  - In a **systematic sample**, every n^{th} member of the population is selected.

  - In a **convenience sample**, the most convenient subgroup is selected.

  - A **stratified sample** is one in which the population is divided into two or more sub-groups, called strata, that share similar characteristics. We then draw a random sample from each strata. This is good if you want to make sure that some members from all strata are present in the sample. Note: A stratified sample may be random but it won't be simple random.

  - In **cluster sampling**, we divide the population into groups (or clusters), then randomly select some of those clusters. Once a cluster is selected all the members of that cluster are included in the sample. This method is often used for convenience purposes. Note: Again, a cluster sample may be random but it won't be simple random.

- **Examples:** Classify each sampling method as **systematic**, **convenience**, **stratified**, **cluster**, or **none of these**. Does the method constitute a random sample? If it is random, is it simple random?

  1. You're considering a lunch-delivery business and want to gather lunch-break data on a sample of employees from your target population of 20 local businesses.

     (a) You randomly select 3 of the businesses and interview all the employees from those businesses.
         Answer: This is a cluster sample (the employees are clustered by the business employing them). It is random because all members of your target population have a 3/20 chance of being selected. It is not simple random because you have grouped your subjects prior to sampling.

     (b) You get a sample of 100 by randomly selecting 50 employees from labor and 50 employees from management.
         Answer: This is a stratified sample (the strata are labor and management). It is unlikely to be random unless there are an equal number of labor and management employees in your target population. Either way, it is not simple random because not all samples of 100 have the same probability of being selected.

  2. **Your Turn:** Suppose you want to gather household income information from a sample of 10 houses on North Street. The house numbers start at 1 and end at 100 without any missing house numbers.

     (a) You randomly select 5 even numbered houses and 5 odd numbered houses.

     (b) You take every 10 $^{\text{th}}$ house starting at number 7.

     (c) You group the houses as #'s 1 - 10, 11 - 20, ..., 91 - 100. You then randomly select one of these groups to be in the sample.

     (d) You randomly select 10 numbers between 1 and 100 and select those 10 houses.

     (e) You include the first 10 houses where someone answers the door.

1.3   Lying with Statistics and Percentages

There are three kinds of lies: lies, damned lies, and statistics.

This phrase was popularized by Mark Twain and is an acknowledgement of the persuasive power of statistics. You may often encounter what appears to be contradictory statistical evidence provided by opposing parties in a debate. The problem is that, quite often, no one is actually lying. There is a subtle difference between clever and deceitful. There is also a fine line between carelessness and intentional distortion of data. I'll describe a few issues here but books have been written on the topic.

- **Loaded Questions:** (Schuman and Presser[†])

  Do you think the United States should forbid public speeches against democracy? (21% said yes).

  Do you think the United States should allow public speeches against democracy? (48% said no).

- **Polls:** Those most motivated respond and the others do not. Even *scientific* polls can can vary by preference. In April, 2012, the presidential approval rating was quoted at 42% (Fox News), 49% (MSNBC), and 47% (Gallup). Non-scientific polls may be entertaining but the results should not be used to make generalizations about a population. Examples: internet polls, mail-in polls, and *call-in* polls.

- **Small Samples** can introduce a greater chance of missing an important segment of the population.

- **Sample versus Population:** Should the results of a poll taken at a small private college be used to make generalizations about all college students?

- **Correlation and Causality**. A correlation does not prove (or disprove) cause and effect.

  - A scientist demonstrates a positive correlation between carbon dioxide and global temperatures. When asked by a reporter: *Does this prove that carbon dioxide causes global warming?*, the scientist says *No*. The headline in the paper then states *Top scientist concludes carbon dioxide does not cause global warming*. The headline should state *Top scientist agrees a significant correlation does not prove that carbon dioxide causes global warming*. There is an important difference between the two headlines. One is false, and the other is hard to read.

  - There is a known negative correlation between campaign spending by an incumbent candidate and his/her results in an election (the more they spend, the worse they do). Does this suggest that an incumbent candidate would be well-advised to spend less for an upcoming election? Not likely.

  - There is a positive correlation between the daytime high temperature and the number of admissions at the local community pool. Does one cause the other? Most definitely. Does the correlation prove this? No.

- **Sponsors:** Often a third party will sponsor (pay for) a survey which directly effects that party. It is always good to know who performed the study and who sponsored it. Pharmaceutical companies often sponsor studies which involve their products.

- **Percentages and Proportions:** Some folks don't understand percentages as well as they probably should. Percentages can be trickier than you may think. As such, this topic gets it's own couple pages.

[†]Schuman, H. & Presser, S. (1996). Questions & Answers in Attitude Surveys. Thousand Oaks, CA: Sage Publications.

- **The Basics of Percentages and Proportions:**

  - **Percentages** and **proportions** are essentially the same. A proportion (in this text) is usually expressed as a fraction or decimal with a value less than one. A percentage is a fraction of 100. Quite often, a percentage results in a proportion that is greater than one.

    - $82\% = \dfrac{82}{100} = 0.82$ and $134\% = \dfrac{134}{100} = 1.34$

  - **Converting a percentage to a proportion:** Drop the % symbol, and divide by 100.
    This is equivalent to dropping the % symbol and moving the decimal point two places to the left.

    - $20\% = \dfrac{20}{100} = \dfrac{1}{5} = 0.2$ and $4.7\% = \dfrac{4.7}{100} = 0.047.$

  - **Converting a proportion to a percent:** Perform the division with a calculator if necessary, then multiply by 100 and insert the % symbol.

    - $\dfrac{21}{40} = 0.525 = 52.5\%$ and $\dfrac{2}{3} = 0.667 = 66.7\%$

  - **Percentage of a value:** Drop the percent sign, divide by 100 and multiply by the value.

    - 15% of $242 = (0.15) \cdot (242) = 36.3$ and 115% of $242 = (1.15) \cdot (242) = 278.3$

  - **Percent Changes:** If a number (x) changes by a given percentage, you first multiply x by the percentage and then add or subtract it to/from the original number (x).

    - If 235 is increased by 20% the result is $235 + (.20)(235) = 282.$
    - If 235 is decreased by 20% the result is $235 - (.20)(235) = 188.$

- **Your Turn**

  (a) Suppose there are 2200 students at the college and 1243 of them are female.

    - What is the proportion of females in the college?
    - What is the percentage of females in the college?

  (b) What is 6% of 80?   What is 250% of 80?

  (c) You invest $100 in a given stock. After one year it decreases by 20% then during the following year there is a 20% increase. After two years are you back to the original $100?

  (d) US Air has a flight to New Orleans for $372 and Continental has a flight to New Orleans for $496. Fill in the blanks. **Round your answer to the nearest whole percentage point**.

    - With respect to the price of the US Air flight, the difference in price is ⬚ %.
    - With respect to the price of the Continental flight, the difference in price is ⬚ %.
    - The Continental flight costs ⬚ % more than US Air.
    - The US Air flight costs ⬚ % less than Continental.

- **Percentages of Percentages:** These can be extra tricky.
  Example: Suppose a smoker's risk of death by heart attack is 60% and a non-smoker's risk is 20%.

  - **Actual Quote (Wrong):** *Quitting smoking reduces your risk of heart attack by 300%.*
  - **Better but still wrong:** *Smoking increases your risk of heart attack by 300%.*
  - **Correct:** *Smoking triples your risk of heart attack.*
  - **Correct:** *Smoking increases your risk of heart attack by 200%.*
  - **Correct:** *Quitting smoking reduces your risk of heart attack by 66.7%.*
  - **Correct:** *Smoking increases your risk of heart attack by 40 percentage points.*

- **Your Turn:** At a given bank, the interest you **earn** in a savings account is 1.5% and the interest you **pay** on a loan is 4.5%. Categorize the following statements as True or False.

  (a) The interest you pay is 3 times what you earn.

  (b) The interest you pay is 200% more than you earn.

  (c) The interest you pay is 300% more than you earn.

  (d) You pay 3 percentage points more than you earn.

  (e) The interest you earn is 33.3% of what you pay.

  (f) The interest you earn is 66.7% less than what you pay.

- **Even Trickier - Simpson's Paradox:** Sometimes proportions across categories directly contradict proportions within categories. When this happens, it is called Simpson's paradox.[‡] Here we investigate the proportion of on-time flight arrivals between two different airlines at two different cities during a one-month period.

| | Proportion of on-time arrivals (%) | | |
Airline	Seattle	San Diego	Combined
National Air	360/400 (90%)	48/50 (96%)	408/450 (90.6%)
Regional Air	44/50 (88%)	376/400 (94%)	420/450 (93.3%)

The Paradox:
If you are in Seattle, National Air (90%) is better than Regional Air (88%).
If you are in San Diego, National Air (96%) is better than Regional Air (94%).
But across categories, Regional Air (93.3%) is better than National Air (90.6%).

What happened?
Regional Air had more flights to San Diego which has a greater on-time percentage for both airlines.

[‡]Technically speaking, this happens when $a/b > A/B$ and $c/d > C/D$, but $(a + c)/(b + d) < (A + C)/(B + D)$.

Chapter 1: Discussions

1. **Qualitative -vs- Quantitative Data:** Later in the text we will conduct various statistical tests about qualitative and quantitative characteristics of various populations. For qualitative data we usually use proportions (percentages) and for quantitative data we usually use means (averages).

   - **Qualitative Data and Proportions:** Suppose you want to compare the popularity of red cars versus black cars at various car dealers. How would you use percentages to make such a comparison?

   - **Quantitative Data and Means:** Suppose you want to compare the gas mileage between American-made cars and German-made cars. How would you use averages to make such a comparison?

2. **Lowest Prices in Town:** Two grocery stores claim to have the lowest prices in town. How would you verify which one actually has the lowest prices?

3. **Percentages of Percentages:** According to a study by the U.S. Center for Disease Control (CDC) the flu rate in people who do not get the flu shot is 2.73% and the rate for people who do get the flu shot is 1.18%. They claim the flu shot is 60% effective. How did they arrive at this number?

4. **Simpson's Paradox - A Classic Example:** In 1973 the University of California, Berkeley was sued for bias against women because 44% of the men applicants were admitted but only 35% of the women applicants were admitted. The data from the three largest departments is sufficient to demonstrate why the lawsuit was dismissed.

Department	Men Applicants	Men % admitted	Women Applicants	Women % Admitted
A	825	62%	108	**82%**
B	560	63%	25	**68%**
C	325	**37%**	593	34%

Why did the lawsuit fail?

Chapter 1: Summary Worksheet

1. Determine whether the given piece of information is most likely a statistic or parameter.

   (a) Campus squirrels weigh 20% more than squirrels living in the woods.

   (b) Less than half of all registered voters in Orange County are democrats.

   (c) Most people who own a yacht make more than $100,000 per year.

2. Determine whether the given value came from discrete or continuous data.

   (a) Larry's highest bowling score is 287.

   (b) Scott's fastest mile time was 4 minutes and 35 seconds.

   (c) The average number of children in American households is 2.4.

3. Determine which of the four levels of measurement (nominal, ordinal, interval, ratio) is most appropriate.

   (a) The birth state of each student in class.

   (b) The birth year of each student in class.

   (c) The number of siblings of each student in class.

   (d) The letter grades given to each student in class.

4. Identify the type of sampling used: systematic, convenience, stratified, cluster, or none of them. Does the sample constitute a random sample?

   (a) At a police sobriety checkpoint, every 5th driver is stopped and interviewed.

   (b) I randomly select 40 freshmen, 40 sophomores, 40 juniors, and 40 seniors from Champlain College for a survey on eating habits of our students.

   (c) We randomly select one of the 4 sections of statistics and interview all of the students in that class.

   (d) In a Gallup poll of 1059 adults, the interview subjects were selected by using a computer to randomly generate telephone numbers that were then called.

   (e) I select the first 5 men in suits and the first 5 men in T-shirts walking down Church Street for a survey on income.

5. In a college of 10,000 students (4,500 males and 5,500 females) a researcher wants 100 students in a sample. Determine if the sampling method is simple random, random but not simple random, or neither.

   (a) The students are given numbers between 1 and 10,000. Then, 100 numbers are randomly selected (by computer) and those 100 persons are in the sample.

   (b) The researcher randomly selects 45 males and 55 females for the sample.

   (c) The researcher randomly selects 50 males and 50 females.

6. At *Nuts Online*, Walnuts are $7.99 per pound and Roasted Pistachios are $8.99 per pound. Complete the following statements. Round your answer (in percent form) to one decimal place.

   Pistachios are ⬚ % more expensive than walnuts.
   Walnuts are ⬚ % cheaper than pistachios.

7. In January 2012, the unemployment rate in North Dakota was 3.2% and the unemployment rate in Florida was 8.0%. Categorize each statement as True or False.

   (a) The unemployment rate in Florida is 2.5 times the rate in North Dakota.
   (b) The unemployment rate in North Dakota is 250% less than the rate in Florida.
   (c) The unemployment rate in Florida is 250% greater than the rate in North Dakota.
   (d) The unemployment rate in Florida is 150% greater than the rate in North Dakota.
   (e) The unemployment rate in Florida is 4.8 percentage points higher than North Dakota.
   (f) The unemployment rate in North Dakota is 60% less than the rate in Florida.

Chapter 1: Problem Set

Numbers with an asterisk[*] have solutions in the back of the book.

Statistics and Data (1.1)

1.[*] **Nielsen Ratings:** Nielsen ratings are audience measurement systems developed by Nielsen Media Research in an effort to determine the audience size and composition of television programming in the United States. These ratings are determined by way of viewer *diaries* and *Set Meters* which are devices connected to televisions in selected homes. The most commonly cited Nielsen results are reported in two measurements: ratings points and share. For example, Nielsen may report a show as receiving a 9.2/15 during its broadcast, meaning that on average 9.2 percent (rating points) of all television-equipped households were tuned in to that program at any given moment, while 15 percent (share) of households watching TV were tuned into the program during this time slot. Describe the two different populations being investigated with these two rating systems.

2. **Average Income:** Do a little research and investigate the difference between the economic indicators of a state's *median household income* and and a state's *income per capita*. What are the populations being studied here? Is it possible for one indicator to increase while the other decreases?

3. **Sample -vs- Population:** In these problems determine the population and the sample.

   (a)[*] In a survey of 2000 U.S. high school seniors less than one in five are current marijuana users.

   (b) In a sample of 10 bags of Doug's Super Green grass seed only 70% of the seeds were actually grass seeds.

   (c) The Gallup Organization contacts 1500 men over 50 years of age and live in the United States and asks them whether or not they smoke.

4.[*] **Parameters -vs- Statistics:** Determine whether the given piece of information is most likely a parameter or a statistic. Ie. Determine if the quoted value came from population or sample data.

   (a) The average GPA for the graduating class of State College was 2.87.

   (b) 8 out 10 Vermont teenagers do not smoke.

   (c) In a sample of 100 students the average GPA was 2.87.

   (d) According to the U.S. Bureau of Labor and Statistics, the U.S. unemployment rate for January 2011, was 9.2%.

5. **Parameters -vs- Statistics:** Determine whether the given piece of information is most likely a parameter or a statistic.

   (a) The 2008 Toyota Prius gets 46 miles to a gallon of gas.

   (b) In a long-term observational study of 1000 adult males, 10% of the smokers developed throat or lung cancer while only 1.5% of the non-smokers developed throat or lung cancer.

   (c) Women have more taste buds then men.

6. **Types of Data:** For each of the following variables, classify them as quantitative or qualitative. If the variable is quantitative determine whether the variable is discrete or continuous. In addition, determine the level of measurement for these variables (nominal, ordinal, interval, or ratio).

   (a)* The gas mileage from 30 different types of cars.

   (b) The number of taste-buds on the tongues of 20 males and 20 females.

   (c)* The colors of all models of a certain type of car.

   (d) Movie ratings of (number of stars) from a certain magazine.

   (e)* The years of major tsunami events.

   (f) The student ID numbers from 100 randomly selected college students.

7.* **Money:** Explain why money might be considered a discrete or continuous variable.

Sampling Techniques (1.2)

8. **Random and Simple Random:** Determine whether the sampling method is random, simple random, or neither.

   (a)* In a group of 100 males and 200 females, 40 participants are selected for a survey: 20 males and 20 females.

   (b) In a group of 100 males and 100 females, 40 participants are selected for a survey: 20 males and 20 females.

   (c)* Student ID numbers are put into a computer and 20 of these students are randomly selected for a survey.

   (d) There are 10 bags of spinach per box coming from Green Acres Spinach Company. The health inspector wants to sample 40 bags of spinach. She randomly selects 4 boxes from the 100 boxes being shipped that day and tests each bag in the selected boxes for the presence of the bacteria E. coli.

9. **Sampling Methods:** Identify the type of sampling used; systematic, convenience, stratified, or cluster, or none of them. Then comment on whether the sample is random or not.

   (a)* The commissioner randomly selects 2 teams from all 30 MLB teams and has everyone on that team drug tested.

   (b) At a border crossing, every 15th car is searched.

   (c)* I randomly select 10 male and 10 female students to be in a survey on whether or not they like my statistics class.

   (d)* In a survey about a given movie, a surveyor stands at the exit and interviews people as they come out of the theater. When one survey is completed the surveyor picks the next person coming out of the theater.

   (e) A quality assurance person randomly selects one box of CD's coming off the production line and tests all of the CD's in that box for defects.

   (f) In a survey of TV viewing preferences 100 people are chosen from the age brackets of 15-20, 21-25, 26-30, ...

Tricks and Percentages (1.3)

10. **Survey Questions:** Comment on the next three problems as a potential sample survey question. Is the question clear? Is it slanted toward a desired response?

    (a)* Some cell phone users have developed brain cancer. Should all cell phones come with a warning label explaining the danger of cell phone use?

    (b) Do you agree that a national system of health insurance should be favored because it would provide health insurance for everyone and would reduce administrative costs?

    (c) In view of the negative externalities in parent labor force participation and pediatric evidence associating increased group size with morbidity of children in day care, do you support government subsidies for day care programs?

11.* **Tax Rate Change:** Suppose a tax rate increases from 5% to 10%. Which statement is **not** true.

    (a) The tax rate doubled.

    (b) The tax rate increased by 200%.

    (c) The tax rate increased by 5 percentage points.

    (d) The tax rate increased by 100%.

12. **Attrition Rates:** The rate at which students do not return to college is called the attrition rate. At State College 120 of the 2000 Freshmen did not return for their Sophomore year. At State University, 360 of the 12000 Freshman did not return for their Sophomore year. Use percentages to describe the difference in attrition rates between the two schools. Make the difference sound big, and make it sound small.

13.* **Jeans -vs- Khakis:** A pair of jeans costs \$100 and a pair of khakis costs \$80. Complete the following statements. Round your answer to the nearest whole percentage point.

    The jeans cost ⬚ % more than the khakis.

    The khakis cost ⬚ % less than the jeans.

14. **Tuition:** The tuition at State College is about \$8,000 per year. The tuition at State University is about \$12,000 per year. Complete the following statements. Round your answer to the nearest whole percentage point.

    State University costs ⬚ % more than State College.

    State College costs ⬚ % less than the State University.

15.* **Grass Seed:** In a 10 pound bag of Doug's grass seed, 95% of the actual seeds are grass while the other seeds are weeds. In a 10 pound bag of generic grass seed, only 80% of the seeds are actually grass seeds. Complete the following statement. Round your answer (as a percent) to two decimal places.

    Doug's grass seed contains ⬚ % more actual grass seeds than the generic brand.