AP Statistics Summer Assignment 2024-2025

Welcome to Advanced Placement Statistics. This is a yearlong course that completes a semester and a half of college level statistics. Your performance in this course is evaluated on the AP exam, which is to be administered in early May, exact dates have not been released yet. A score of 3 or higher out of 5 on this exam can lead to college credit at many colleges and universities.

For your first assignment, you will be learning about how to collect data and run an experiment. In order to do this, you will first need to log into our Google Classroom, the code is **gcsgcqr**.

Once you have logged into the Google classroom:

(1) Log into the google classroom by August 1, 2024.

(2) Read chapter 4 of our textbook (On Google Classroom) pages 207 – 229, 234 – 259, and 266 – 276. Be sure to read the green summary of each section.

(3) Complete the reading guide, **Designing Studies**, for chapter 4 (attached). **This is due Wednesday, September 4th, 2024. No late assignments accepted.**

(4) Watch all of the videos under Mr. G's Mathematics Emporium Chapter 4 and complete the fill – in class notes (attached). **Class notes are due on Wednesday, September 4th, 2024. No late assignments accepted.**

Videos are uploaded on the Google Classroom.

(5) Complete the attached take home test. **It is due Friday, September 6th, 2024. No late assignments will be accepted.**

This summer assignment *(Reading guide, Fill- in class notes, and Take home)* will count as your first 3 grades for the 2024-2025 school year.

I look forward to a productive year!

CHAPTER 4 READING GUIDE
DESIGNING STUDIES

Name _____

Key Vocabulary:

- voluntary response sample
- confounded
- population
- sample
- design
- convenience sampling
- biased
- simple random sample
- table of random digits
- probability sample
- stratified random sample
- strata
- undercoverage
- nonresponse
- response bias
- sampling frame
- systematic random sample
- observational study
- experimental units
- subjects
- treatment
- factor
- level
- placebo effect
- control group
- randomization
- completely randomized experiment
- statistically significant
- replication
- hidden bias
- double-blind experiment
- block design
- matched pairs design

4.1 Sampling and Surveys

1. How does a population differ from a sample?

2. What are the steps to planning a sample survey?

3. Why are voluntary response samples unreliable?

4. Why might convenience sampling be unreliable?

5. What is a biased study?

6. Define simple random sample.

7. What two properties of a table of random digits make it a good choice for creating a simple random sample?

8. State the two steps in choosing an SRS?
   -
   -

9. How do you select a stratified random sample?

10. What is cluster sampling?

11. What is the difference between a strata and cluster? (see page 221)

12. Give an example of undercoverage in a sample.

13. Give an example of non-response bias in a sample.

14. What factors can cause response bias in a sample.

15. How can the wording of questions cause bias in a sample?

16. What is the difference between nonresponse and voluntary response?

4.2 Experiments

1. How does an experiment differ to an observational study?

2. What is confounding?

3. Explain the difference between experimental units and subjects.

4. Define treatment.

5. What is the difference between factor and level in an experiment? Example on page 235.

6. Explain how to perform a completely randomized design.

7. What is the significance of using a control group?

8. The basic principles of statistical design of experiments are:
    -
    -
    -

9. Describe the placebo effect.

10. Define randomization.

11. Define statistically significant.

12. Describe a block design.

13. When does randomization take place in a block design, and how does this differ to a completely randomized design?

14. What is the goal of a matched pairs design?

15. State the two most common ways in which matched pairs experiments are designed.

-
-

16. What are the advantages of a double-blind study?

4.3 Using Studies Wisely

1. What are the criteria for establishing causation when we can't experiment?

2. What is meant by inference about cause and effect?

Chapter 4: [Video #1] – Intro to sampling

"The U.S. **census** is mandated by the Constitution and provides vital data for the nation. Census data are used, among other purposes, to apportion the seats of the U.S. House of Representatives; redraw congressional districts in each state; and allocate billions of dollars each year in federal financial assistance." - ***gao.gov***

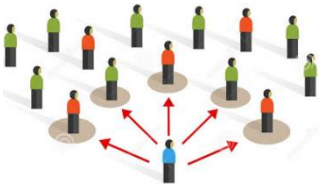
- *What is a census*?
  _____
- *What's the purpose of a census*? → It helps answer a specific question about the _____.
- *Why don't we always use a census*? → Too much _____ and _____!
- *How can we answer a question about the population instead*? → We take a _____!

Population: _____
vs.
Sample:_____

Questioning a population → A _____ helps answer a specific question about the population.
vs.
Questioning a sample → We can use the data from a sample to _____ (or _____ _____) about the population.

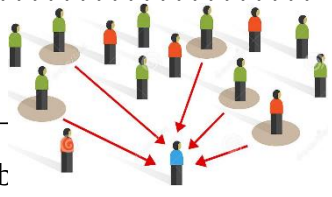
Can you ask whoever you want to be in a sample? ____ , _____ , _____ , _____ , _____ , _____ ,_____!

(BAD) Sampling Techniques



1) _____ *sample*
   a. Whoever is most _____ to you!
   b. This will likely produce _____ (error) in your conclusions!
      i. Those closest to you _____ _____ represent the _____ well.

****



2) _____ _____ *sample*
   a. *If you build it* (a _____), *they will come* (_____
   b. Still has the potential for high _____, so also avoid if possib
   c. Often times, only those with _____ _____ will respond which, again, _____ _____ represent the entire _____ feelings.

Ex: KSDK posts a daily question on its website. Why should we NOT trust the results???

- People with strong opinions are _____ _____ to respond which will **NOT** be representative of the entire population.
- The results will *likely* _____ or _____ the truth about the pop.

Sampling vs. Nonsampling Errors

- ***Sampling error*s**: mistakes made in the process of taking a sample that could lead to _____ information about the population.
- ***Nonsampling errors***: mistakes _____ made in the process of taking a sample that could lead to _____ information about the population.

Types of sampling errors:

- ***Undercoverage****:* When some groups in the population are _____ _____ of the sampling process
  - *If surveys are sent to households, then who might we be leaving out???*

_____

  - *Why is this considered an "error"?* → Leaving out the answers from these groups of people _____ the rest of the population considered.

Types of nonsampling errors:

- ***Nonresponse****:* When someone selected to be in a sample either cannot be _____ or _____ to participate.
- ***Response bias***: When someone answers questions _____ or _____ because the question is written to produce a desired response or because the question is _____.
  - *Why is this considered an "error"?* → The data from the survey may not accurately _____ about how the whole population feels.

YOU DO!!! **********

Would you consider each of the following a sampling or nonsampling error?

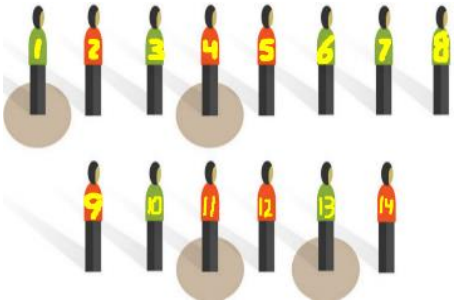
1) *You accidentally record the wrong answer to a question from someone's survey.*

2) *You ask your IG followers to comment about your most recent selfie taken at Starbucks.*

***Chapter 4:* [Video #2] – Why, hello there SRS!**

(GOOD) Sampling Techniques

3) *Simple random sample* (_____)



- Every group of individuals have the _____of being selected.
- By using _____ to select our sample, we will reduce *most* bias.

- **3 ways to take a Simple Random Sample (SRS):**
  1) Use an actual, physical routine like drawing names out of a hat.
     - Make sure everything is "_____" to ensure _____chance.

  2) Use a random number table (back of formula sheet) to select numbers tied to people.

1. It is a long string of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 with these two properties. a. Each entry in the table is **equally likely** to be any of the ten digits 0 through 9. b. The entries are **independent** of each other. That means that knowledge of one part of the table gives no information about any other part. 2. Q: Why are the numbers in groups of 5? A: To make it easier to read 3. Q: Why are the Lines numbered from 101 to 150? A: To make it easier to read

  3) Use technology to randomly generate numbers that are tied to people.

Using the **randInt** function on your calculator → **[Math], PRB Menu, #5: randInt** The inputs are: **randInt(min #, max #, number of numbers you want displayed at a time).**

We will devise a "scheme" for how we will accomplish our random sample which has 3 parts:

1) ***Label the individuals in the population.*** With the "hat method", this step could just be writing everyone's name on equal sized piece of paper. For the technology or random number table method, this means assigning numbers to each individual. You must be ***specific*** about the range of numbers that will be used. You cannot just say "assign numbers to individuals".

2) ***How you will select the first individual.*** With the "hat method", this step is simply saying you would mix up the papers and blindly select a piece of paper. With technology, this means describing what function of you would use to select your random number. With the random number table, this means stating what line will you use and *how many digits will you select* at a time to represent the first individual.

3) ***Repeating the process and conditions to watch out for.*** With the "hat method", this step just means how many names will you select out of the hat, will you put a selected name back in the hat before you select the next, etc. With technology and the random number table, this means using the function or the table from part 2 however many times you need before your whole sample is selected and discussing what to do with repeated numbers.

****

Ex #1: The names of 10 CHS baseball players are given. At practice, every 15 minutes the coach randomly selects two players to run sprints up the hill. Practice is two hours long and starts at 3:00. Write a scheme using the random number table and use the given Line 314 to determine who runs at what time.

Joe	Ron	Dave	Matt	Bryan
Mark	Pat	Jimmy	Adam	Gary

Line 314: **35476 65972 39421 65850 04266 35435**

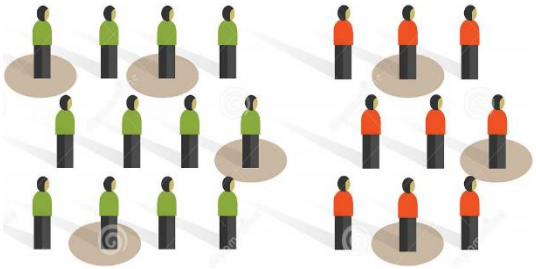
****

YOU DO!!! The principal has a list of 50 students who completed a recent SAT prep course. He needs to select an SRS of 10 students to receive a prize. The principal has asked you to **write a scheme for how he should randomly select the winners**.

Chapter 4: [Video #3] **– Other good sampling techniques**

4. **Stratified Random Sample**

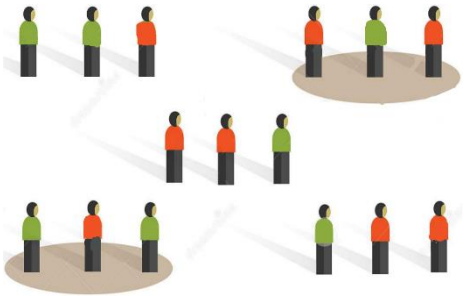
- Divide the population into _____ groups (***strata***) and select an _____ from _____ strata.



- This allows us to obtain a sample that best represents the population of interest.
    - *We want to survey 100 high school students about food preferences in the cafeteria.*
    - *In a survey about social justice issues, we want a sample that reflects the local population that is 60% white, 25% Hispanic, and 15% black.*

5. **Cluster Sample**

- Divide the population into _____ groups (*clusters*) and select an SRS of _____ clusters.

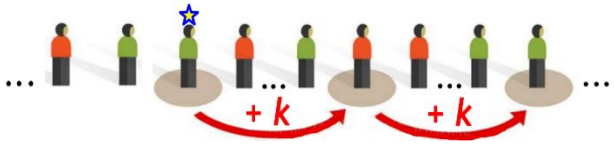


- This allows us to obtain a sample that best represents the population of interest.
    - *We want to survey 100 high school students about food preferences in the cafeteria.*
    - *In a survey about social justice issues, we want a sample that reflects the local population that is 60% white, 25% Hispanic, and 15% black.*

Stratified	**vs.**	**Cluster**
_____ groupings		_____ groupings
randomy select _____ people from _____ strata.		randomly select _____ clusters and use _____ from those clusters.

6. **Systematic Random Sample**

- Arrange the population. Select a random starting point*. Then, select every k^{th} person to obtain your sample. **(*k* = _____ size / desired _____ size)**
  - Great to use with databases/spreadsheets.
  - Usually easier to perform than an _____.
  - Potentially has more _____ than SRS.



7. **Multistage Random Sample**

- Any _____ sampling techniques previously mentioned done consecutively.
- Great if you have a _____ population that you can wittle away into a good sample.

****

YOU DO!!! The manager of a sports arena wants to learn more about the financial status of people who are attending a basketball game. He plans to give a survey to 500 fans in attendance. There are 30 sections going around the arena (*section 101 to 130*) each containing 250 seats. There are 25 rows per section (*rows A – Y*). Rows A – E tickets cost $200, rows F – J tickets cost $150, rows K – O tickets cost $75, rows P – T tickets cost $50, and rows U – Y tickets cost $25.

- Write a scheme to take a stratified random sample using the letter rows.

- Write a scheme to take a cluster sample using the numbered sections.

- Which of the above two samples do you think is better? Explain!

Chapter 4: [Video #4] – Intro to experimental design (part 1)

The school wants to increase SAT scores by offering an SAT prep course. 100 juniors signed up. The free prep course is offered for two hours per week for 6 weeks. A student is randomly selected each week to receive $50. The mean is 1230 for the prep course and 1080 for the non-prep course people. Taking the SAT prep course *will cause* your SAT score to go up!..._____

Observational Study

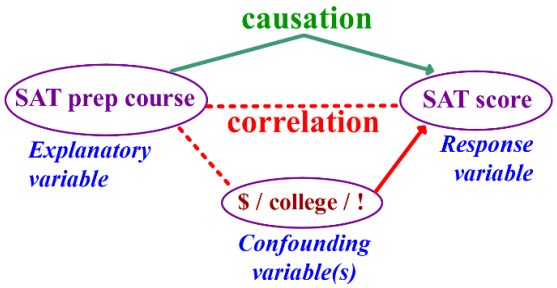
- Observes individuals and measures variables of interest _____ attempt to influence the responses.
    - The students who took the prep course were merely _____.
- Observational studies can find a _____ between an explanatory and response variable **BUT** this correlation _____ _____ imply causation!

****

Confounding variables

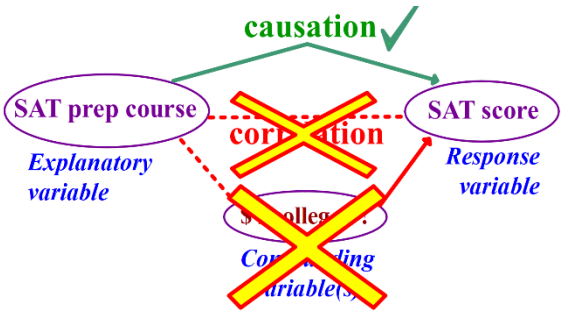
- _____ variables/factors that might explain a _____ in SAT scores (*the response variable*).

Confounding: when it's _____ if the explanatory variable(s) and/or the confounding variable(s) led to a change in the response variable.



Experiment

- Deliberately _____ a treatment(s) on individuals to measure their response(s) while hopefully reducing or eliminating any _____ variables.



****

100 juniors → ***experimental units*** **(** _____ **)** → randomly selected from the JR population.

Was the sample randomly taken?

- *Yes!* → results apply to the _____ that the _____ was taken from.
- *No!* → results only apply to those in the _____ sample (or applies to people _____ to those in the _____ sample.

SAT prep course → ***explanatory variable(s) / treatment(s) /*** _____ → _____.

- *One factor* = _____
- *Two levels* = _____ prep course **vs.** _____ prep course

  _____ for confounding variables *control group* = provides a _____

Were treatments randomly assigned?

- *Yes!* → results imply _____.
- *No!* → results do not imply _____.
  - *If NO, then we introduce back all of those confounding variables that may contribute to _____ in SAT scores _____ related to taking the prep course!*

****

SAT scores → _____ _____→ compare SAT scores of the two groups.

YOU DO!!! **Ex:** A group of researchers in Africa found a creative way to protect cattle from lion attacks. They painted eyes on the cows' rear ends. To determine if this treatment is effective, they randomly assign 2100 cattle to one of three treatments: eyes on rear ends, cross-marks on rear ends, or no markings on rear ends. After four years of living a normal cow life, the cows with eyes had zero deaths, cows with cross-marks had 4 deaths, & cows with no markings had 15 deaths.

- Is this an observational study or an experiment? Explain!

- What are the experimental units?

- How many factors are there? _____

- How many levels are there for each factor? _____
- What purpose does the “no markings” group provide? _____
- What is the response variable? _____
- To whom do the results apply? _____
- Does this show correlation or causation? _____

Chapter 4: [Video #5] **– Intro to experimental design (part 2)**

Let's test out a new headache medication on 500 volunteers suffering from frequent headaches.

- _____ → 500 volunteers

- **Treatment(s) / factor(s)** = _____
    - **Level(s)** = _____
        - *New or _____ medication*
        - *New, _____ , or old medication*

- _____ → a fake or dummy treatment
- _____ effect → when a placebo produces positive results when none is expected.
- _____ effect → when a placebo produces negative results when none is expected.

- _____ → patients do not know which treatment they are receiving.
- _____ → patients AND experimenter do not know who receives what

- _____ → effectiveness (*eliminate / time to*)

****

Key Principles of Experimental Designs

1) _____ _____

    - Allows impartial chance to assign subjects to treatment groups that create _____ groupings.
    - This helps to naturally reduce the effects of _____ or confounding variables.

2) _____

- Use enough experimental units / subjects in each treatment group to show _____ in response variable was not by chance alone.

3) _____

- An attempt to reduce / eliminate any effects from lurking (_____) or confounding (_____) variables.
- Any differences in the response variable can then be primarily attributed to the difference(s) in the treatment(s).

****

YOU DO!!! **Ex**: A pet pharmaceutical company wants to test two new heartworm treatments for dogs. They will randomly assign one of the two new treatments to dogs that have been recently diagnosed with heartworms until 50 dogs are assigned each of the two treatments.

- How many experimental units are there? _____
- How many factors are there? _____
- How many levels are there? _____
- Should a control group be used? Why?

- Does the experiment need to be:
  - single-blinded? _____
  - double-blinded? _____
- Should a placebo be used? Why?

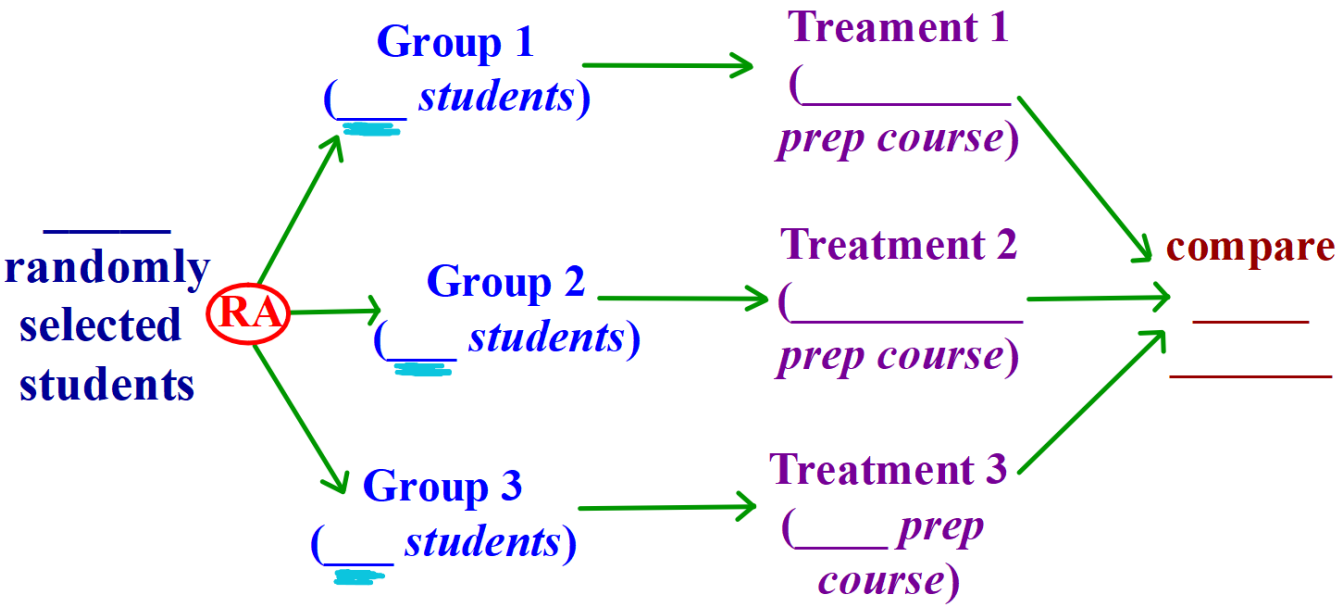
- To whom do the results of the study apply?

Chapter 4: [Video #6] **– Experimental designs**

Going back to a familiar example…

- 120 juniors + 60 seniors = 180 students (all randomly selected)
- Teacher-led SAT prep vs. online self-guided SAT prep vs. no SAT prep
- Compare SAT scores

Experimental Design #1: *Completely Randomized Design (CRD)*



How to randomly assign to groups:

1) *Label 3 pieces of paper with* _____ *and mix them up in a hat. Have each person* _____ *select a slip of paper. That person is assigned to that group. Continue until one group reaches* _____. *Continue until another group has* ____. *All others go into* _____ *group.*

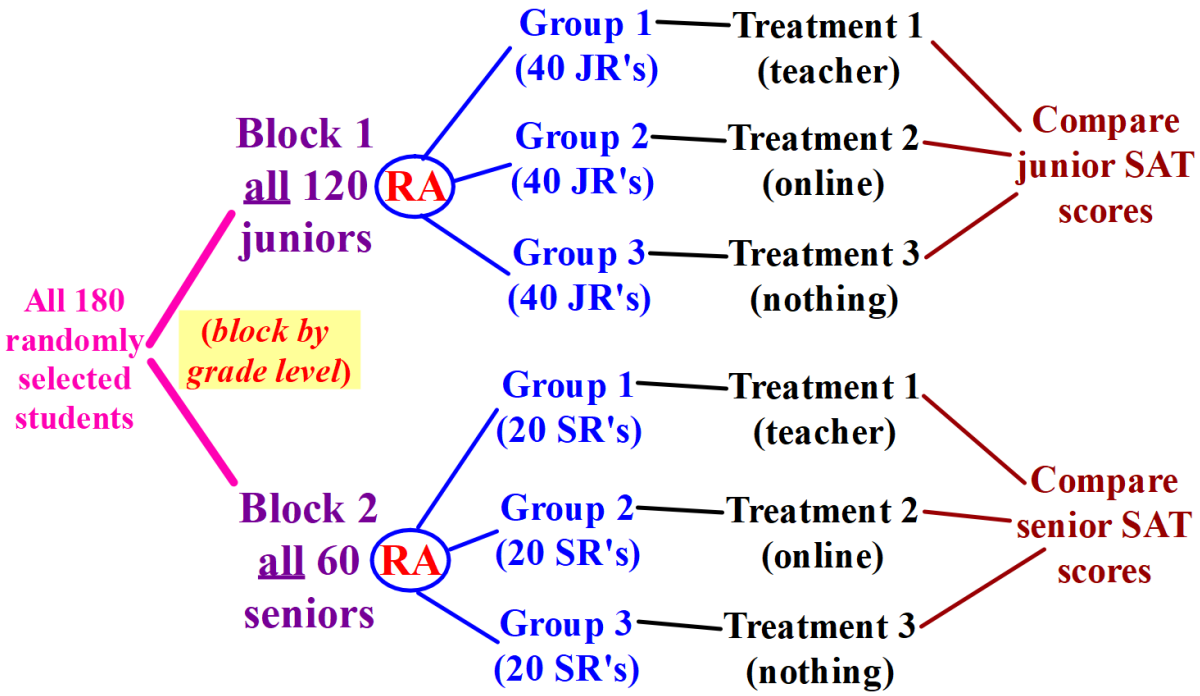
2) Label each student _____. Use a random number _____
to select _____ _____numbers between _____. Assign them to group ____. Select ____ _____ unique numbers that have not been chosen. Assign them to group ____. The rest go into group ____.

3) Label each student _____. Use any line of a random number _____ to select _____ unique _____ numbers between _____.

Assign them to group ____. Continue in the random number table to select 60 new unique three-digit numbers that have not been chosen. Assign them to group ____. The rest go into group ____. The 60 seniors have already taken the SAT with no prior prep course experience. Seniors tend to score higher than juniors by having more educational experience. Therefore, we should experiment on seniors and juniors _____.

One grade may do better with one prep course than the other. _____ is a confounding variable, so we will control for it by using a _____. This is essentially doing a _____ for ***each*** block (*grade level*).

Experimental Design #2: *Block Design*

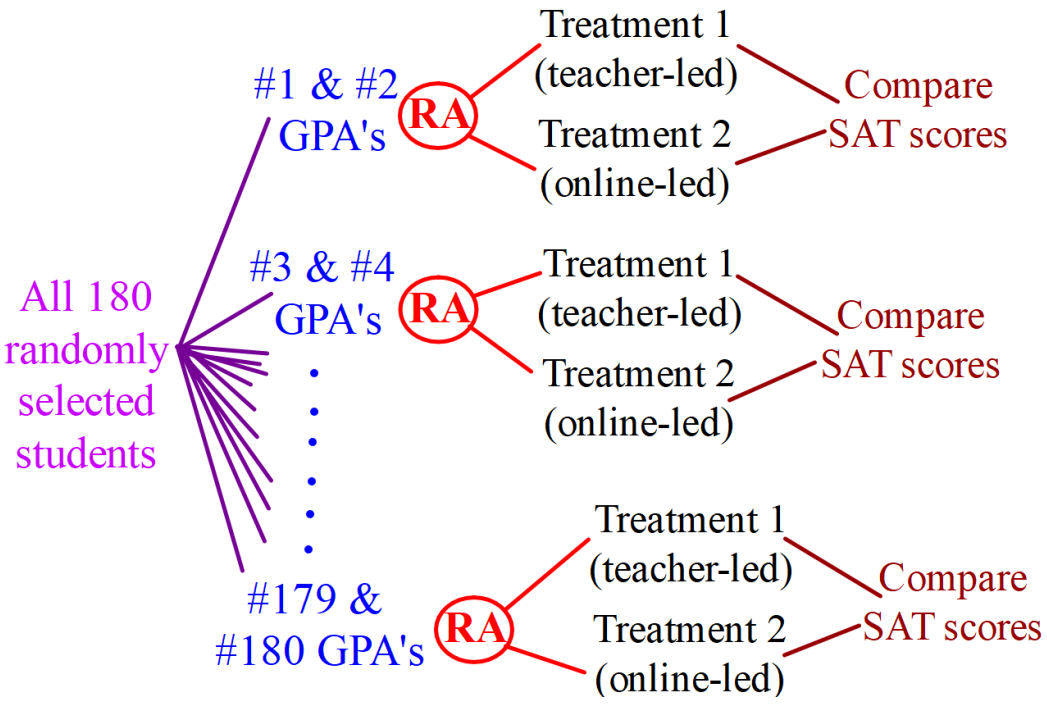


****

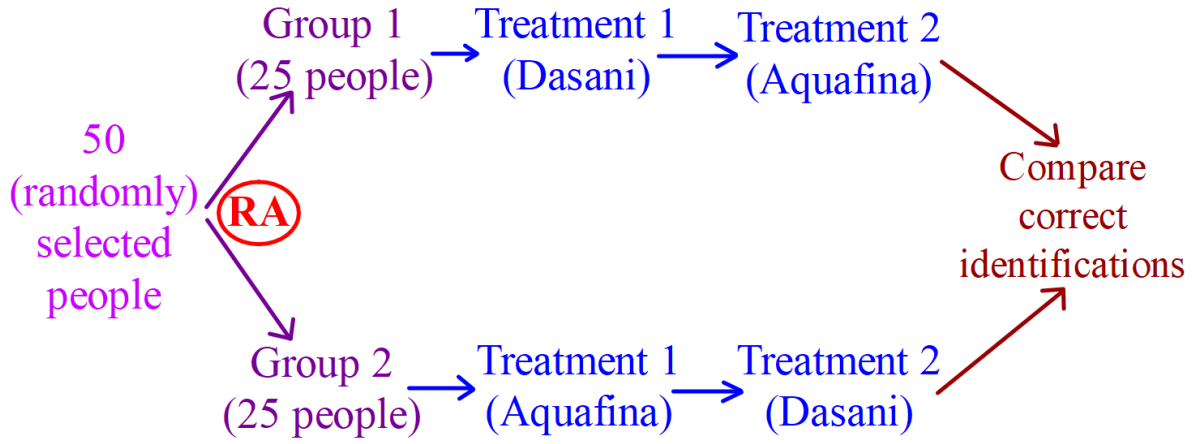
Another confounding variable is a student's _____ which may have a greater effect on SAT scores more than _____. We should compare students with similar _____ to know which prep course method works best for students with varying _____. We will ignore the no prep course option this time around. We will use a _____.

In a **matched pairs design**, we will manually _____ (or block!) _____ students at a time starting with the two _____ GPA's, then the next two _____ GPA's, and so forth until the l_____ two GPA's are paired. For each pair, it will be randomly determined which student will receive the teacher-led vs. online-led prep course.

Experimental Design #3: *Matched Pairs Design*



The "pairs" in a matched pairs design could be _____ person, so each person would receive each treatment but in a _____. Our prep course example will not work with this scenario, so...***can you tell the difference between Dasani and Aquafina bottled water?***



YOU DO!!! **Ex**: After the success of determining painted eyes on cow's rears helps deter lion attacks, researchers want to test different color combination of eyes. They randomly assign 1000 cattle to one of two treatments: yellow eyes outlined in red or white eyes outlined in black.
Outline a completely randomized design.

Researchers believe the color of the cow itself may play a role in whether it is attacked. They have 350 brown cows, 400 white cows, and 250 black cows. ***Outline a block design addressing the cow colors.***

Other researchers believe the size (weight) of the cow plays more of a role in whether it is attacked compared to the cow's natural color. ***Outline a matched pairs design addressing this.***

Preparing for your Chapter 4 Test

You Should Know:

- How to identify the sample and the population
- How to identify sources of bias such as voluntary response, convenience sampling, undercoverage, nonresponse, response bias, lurking variables, confounding, wording of questions, etc.
- How to explain and describe the differences between SRS, convenience sampling, stratified random sampling, and cluster random sampling
- How to use the table of random numbers and the randInt function on the calculator to select random samples of various sizes.
- The difference between observational studies, surveys, and experiments.
- How to identify the factors (explanatory variables), treatments, response variables, experimental units or subjects in an experimental design.
- The purpose for using control groups in an experiment
- The three principles of an experimental design: control, randomization, replication
- How to outline the design of a completely randomized experiment and a randomized block design experiment
- The correct usage for a matched pairs experimental design
- That a larger sample size will provide more accurate results than a small sample size

Chapter 4 – Important Terms

- Voluntary Response Sample
- Experiment
- Confounding
- Population
- Sample
- Convenience Sampling
- Bias
- Simple Random Sample
- Stratified Random Sample
- Cluster Random Sample
- Systematic Random Sample
- Multistage Random Sample
- Sampling Error
- Nonsampling error
- Undercoverage
- Nonresponse
- Response Bias
- Observational Study
- Experimental Units
- Subject
- Treatment
- Factors
- Levels
- Placebo Effect
- Control Group
- Completely Randomized Design
- 3 Principles of Experimental Design
- Double Blind Experiment
- Block Design
- Matched Pairs Design

Name: _____ **Per: _____ Date: _____**

AP Statistics **Summer Assignment TAKE HOME**

DUE: Friday September 6th, 2024

Part 1: Multiple Choice. *Circle the letter corresponding to the best answer.*

1. A new headache remedy was given to a group of 25 subjects who had headaches. Four hours after taking the new remedy, 20 of the subjects reported that their headaches had disappeared. From this information you conclude
(a) that the remedy is effective for the treatment of headaches.
(b) nothing, because the sample size is too small.
(c) nothing, because there is no control group for comparison.
(d) that the new treatment is better than aspirin.
(e) that the remedy is not effective for the treatment of headaches.

2. We wish to draw a sample of 5 without replacement from a population of 50 households. Suppose the households are numbered 01, 02, . . . , 50, and suppose that the relevant line of the random number table is `11362 35692 96237 90842 46843 62719 64049 17823`.
Then the households selected are
(a) households 11 13 36 62 73
(b) households 11 36 23 08 42
(c) households 11 36 23 23 08
(d) households 11 36 23 56 92
(e) households 11 35 96 90 46

3. A maple sugar manufacturer wants to estimate the average trunk diameter of Sugar Maples trees in a large forest. There are too many trees to list them all and take a SRS, so he divides the forest into several hundred 10 meter by 10 meter plots, selects 25 plots at random, and measures the diameter of every Sugar Maple in each one. This is an example of a
(a) multistage sample.
(b) stratified sample.
(c) simple random sample.
(d) cluster sample.
(e) convenience sample.

4. A researcher for a consumer products company is field testing a new formula for laundry detergent. He has contracted with 60 families, each with two children, who have agreed to test the product. He randomly assigns 30 families to the group that will use the new formula and 30 to the group that will use the company's current detergent formula. The most important reason for this random assignment is that
(a) randomization makes the analysis easier since the data can be collected and entered into the computer in any order.
(b) randomization eliminates the impact of any confounding variables.
(c) randomization is a good way to create two groups of 30 families that are as similar as possible, so that comparisons can be made between the two groups.
(d) randomization ensures that the study is double-blind.
(e) randomization reduces the impact of outliers.

5. To test the effect of music on productivity, a group of assembly line workers are given portable mp3 players to play whatever music they choose while working for one month. For another month, they work without music. The order of the two treatments for each worker is determined randomly. This is
(a) an observational study.
(b) a matched pairs experiment.
(c) a completely randomized experiment.
(d) a block design, but not a matched pairs experiment.
(e) impossible to classify unless more details of the study are provided.

6. A survey was done in the town of Mechanicsville to estimate the proportion of cars that are red and made by companies based in Japan. A simple random sample of 25 cars from a parking lot at Lee-Davis High School was taken. Which of the following statements is correct?
(a) Since this is a simple random sample, it should be representative of all the cars in Mechanicsville.
(b) If a simple random sample of 15 cars were taken, we would expect the same amount of variability in the proportion of red cars as we would with a sample of 25 cars.
(c) An alternative method for getting a representative sample would be to select the 25 cars closest to a specified location, such as the entrance to the gymnasium.
(d) A different team doing the sampling independently would probably obtain a slightly different answer for their sample proportion.
(e) The results would be the same regardless of the time of day that the sample is taken.

7. A nutritionist wants to study the effect of storage time (6, 12, and 18 months) on the amount of vitamin C present in freeze dried fruit when stored for these lengths of time. Six fruit packs were randomly assigned to each of the three storage times. The treatment, experimental unit, and response are respectively:
(a) A specific storage time, amount of vitamin C, a fruit pack
(b) A fruit pack, amount of vitamin C, a specific storage time
(c) Random assignment, a fruit pack, amount of vitamin C
(d) A specific storage time, a fruit pack, amount of vitamin C
(e) A specific storage time, six fruit packs, amount of vitamin C

8. A researcher observes that, on average, the number of divorces in cities with Major League Baseball teams is larger than in cities without Major League Baseball teams. Which of the following is the most plausible explanation for this observed association?
(a) The presence of a Major League Baseball team causes the number of divorces to rise (perhaps husbands are spending too much time at the ballpark).
(b) The high number of divorces is responsible for the presence of Major League Baseball teams (more single men means potentially more fans at the ballpark, making it attractive for an owner to relocate to such cities).
(c) The association is due to confounding (Major League teams tend to be in large cities with more people, hence a greater number of divorces).
(d) The association makes no sense, since many married couples go to the ballpark together.
(e) The association is purely coincidental. It is implausible to believe the observed association could be anything other than accidental.

9. Control groups are used in experiments in order to accomplish which one of the following?
(a) Limit the effects of variables other than the explanatory variable on the outcome.
(b) Control the subjects of a study to ensure that all participate equally.
(c) Guarantee that someone other than the investigators, who have a vested interest in the outcome, controls how the experiment is conducted.
(d) Achieve a proper and uniform level of randomization.
(e) Reduce variability in results.

10. A survey is to be administered to recent graduates of a certain nursing school in order to compare the starting salaries of women and men. For a random sample of graduates, three variables are to be recorded: sex, starting salary, and area of specialization. Which of the following best describes a conclusion that can be drawn from this study?
(a) Whether being female causes graduates of this nursing school to have lower (or higher) starting salaries than males.
(b) Whether being female causes graduates in this sample to have lower (or higher) starting salaries than males.
(c) Whether choosing a certain area of specialization causes females graduates of this nursing school to have lower (or higher) starting salaries than males.
(d) Whether there is an association between sex and starting salary among graduates of this nursing school.
(e) Whether there is an association between sex and starting salary at all nursing schools.

Part 2: Free Response

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

11. Read the following brief article about aspirin and alcohol.

> **Aspirin may enhance impairment by alcohol**
> Aspirin, a long time antidote for the side effects of drinking, may actually enhance alcohol's effect, researchers at the Bronx Veterans' Affairs Medical Center say. In a report on a study published in the *Journal of the American Medical Association*, the researchers said they found that aspirin significantly lowered the body's ability to break down alcohol in the stomach. As a result, five volunteers who had a standard breakfast and two extra-strength aspirin tablets an hour before drinking had blood alcohol levels 30 percent higher than each had when they drank alcohol alone. Each volunteer consumed the equivalent of a glass and a half of wine.
> That 30 percent could make the difference between sobriety and impairment, said Dr. Charles S. Lieber, medical director of the Alcohol Research and Treatment Center at the Bronx center, who was co-author of the report with Dr. Risto Roine.

(a) Explain why this is an experiment and not an observational study.

(b) Identify the explanatory and response variables.

(c) Identify the experimental design used in this study. Justify your answer.

(d) In the second sentence above is the phrase, "…researchers said they found that aspirin significantly lowered the body's ability to break down alcohol…" What is the statistical meaning of the word "significantly" in the context of this study?

(e) This was a controlled experiment. Describe how it was controlled and explain the purpose of doing so.

12. High blood pressure adds to the workload of the heart and arteries and may increase the risk of heart attacks. If not treated, this condition can also lead to heart failure, kidney failure, or stroke. We wish to test the effectiveness of Angiotensin-converting enzyme (ACE) inhibitors as a treatment for high blood pressure.

(a) It is well known that men and women may react differently to common cardiovascular drug treatments. What sort of experimental design would you choose for this study, and why?

(b) Explain why an experiment involving 600 men and 500 women is preferable to one involving 60 men and 50 women.

(c) Assume that 600 men and 500 women suffering from high blood pressure are available for the study. Describe a design for this experiment. Be sure to include a description of how you assign individuals to the treatment groups.

13. Bias is present in each of the following sampling designs. In each case, identify the type of bias involved and state whether you think the sample result obtained is lower or higher than the actual value for the population.

(a) A political pollster seeks information about the proportion of American adults who oppose gun controls. He asks an SRS of 1000 American adults: "Do you agree or disagree with the following statement: Americans should preserve their constitutional right to keep and bear arms." A total of 910, or 91%, said, "Agree" (that is, 910 out of the 1000 oppose gun controls).

(b) A flour company in Minneapolis wants to know what percent of local households bake at least twice a week. A company representative calls 500 randomly-selected households during the daytime and finds that 50% of those who responded bake at least twice a week.