HOMEWORK 1.1

An Overview of Statistics

Problem 1 A researcher measures the cholesterol levels of 250 patients at a clinic to learn about the cholesterol levels of all adults in Ohio.

a.) Population cholesterol levels of all adults in Ohio

b.) Sample cholesterol levels of the 250 patients

Problem 2 To estimate the average price of regular gas in a state, an agency records the price at 120 gas stations.

a.) Population gas prices at all stations in the state

b.) Sample gas prices at the 120 stations

Problem 3 Decide whether each number is a parameter or a statistic. Explain.

a.) A poll of 900 likely voters shows that 47% favor a new school levy. statistic

47% comes from a sample of 900 voters.

b.) Of the 30 Major League Baseball teams, 15 play in the American League. parameter

15 of 30 describes ALL MLB teams (population).

c.) A survey of 300 restaurants found an average wait time of 12 minutes. statistic

12 minutes comes from a sample of 300 restaurants.

d.) The 2020 U.S. Census counted about 11.8 million people living in Ohio. parameter

A census measures the entire population.

Problem 4 Decide whether each statement uses descriptive or inferential statistics. Explain.

a.) The mean score on yesterday's quiz in Room 214 was 84. descriptive

It only summarizes the scores in Room 214.

b.) After testing 1,000 phones, a company estimates that 2% of all its phones have a defective screen. inferential

Uses a sample of 1,000 phones to draw a conclusion about ALL of the company's phones.

c.) Of the 30 students in our class, 12 have a part-time job. descriptive

It only describes our class of 30.

Problem 5 A study of 500 high school seniors found that seniors who ate breakfast every day had a mean GPA of 3.3. Seniors who usually skipped breakfast had a mean GPA of 2.9.

a.) Identify the population. GPAs of all high school seniors

b.) Identify the sample. GPAs of the 500 seniors studied

c.) Which part of the study is descriptive statistics?

The mean GPAs: 3.3 for daily breakfast eaters and 2.9 for those who usually skip breakfast.

d.) What inference might be drawn from the study?

High school seniors who eat breakfast every day tend to have higher GPAs.

e.) Does this study prove that eating breakfast raises a student's GPA? Explain.

No. It shows a link, not a cause. Other factors (sleep, home routine, motivation) could explain both.

Problem 6 Decide whether each statement is true or false. If it is false, rewrite it so it is true.

a.) A statistic is a number that describes a population characteristic. false

A statistic describes a SAMPLE characteristic. (A parameter describes a population.)

b.) A sample is a subset of a population. true

A sample is part of the population.

c.) Inferential statistics involves only organizing, summarizing, and displaying data. false

That is DESCRIPTIVE statistics. Inferential statistics uses a sample to draw conclusions about a population.

Problem 7 Decide whether each symbol describes a parameter or a statistic.

a.) $\bar{x}$ statistic

b.) σ parameter

c.) s statistic

d.) μ parameter

Greek letters (μ, σ) → parameters.
Roman letters ($\bar{x}$, s) → statistics.

HOMEWORK 1.2

Data Classification

Problem 1 Decide whether each variable is qualitative or quantitative.

a.) Favorite music genre qualitative

b.) Number of pets in a household quantitative

c.) Jersey numbers of the Cleveland Browns players qualitative

d.) Heights of plants, in centimeters quantitative

e.) Brand of running shoes qualitative

f.) Gallons of gas bought at a station quantitative

c.) Jersey numbers are labels, not amounts → qualitative.

Problem 2 Identify the level of measurement of each data set. Explain.

a.) The types of trees in a city park nominal

Tree types are names only, with no order.

b.) The final standings of the four AFC North teams (1st, 2nd, 3rd, 4th) ordinal

Places have an order, but the gap between places is not a measured amount.

c.) T-shirt sizes: S, M, L, XL ordinal

Sizes can be ordered, but differences between sizes are not measured amounts.

d.) Water temperatures of a lake, in degrees Celsius interval

0°C is not "no temperature" → no inherent zero.

e.) Monthly rent amounts, in dollars ratio

$0 means no rent: inherent zero.
$1,200 is twice $600.

f.) Seasons in which Ohio State won a national championship in football since 2000: 2002, 2014, 2024 interval

Differences in years make sense, but year 0 is not "no time," so ratios do not.

g.) The number of points a team scored in each game ratio

0 points means no points: inherent zero.

Problem 3 A student says, "It is 30°C in Miami and 15°C in Chicago, so Miami is twice as warm as Chicago." Is the student correct? Explain.

No. °C is interval level. 0°C does not mean "no heat," so ratios such as "twice as warm" are not meaningful.

Problem 4 One employee earns $60,000 per year and another earns $30,000 per year. Is it meaningful to say the first employee earns twice as much? Explain.

Yes. Salary is ratio level: $0 means no pay, so ratios are meaningful. 60,000 = 2(30,000).

Problem 5 A car dealership keeps the data below for every car it sells. Identify the level of measurement of each variable.

a.) Body style (sedan, SUV, pickup truck, …) nominal

category names only, no order

b.) Customer satisfaction: poor, fair, good, excellent ordinal

ordered categories, no measured differences

c.) Price, in dollars ratio

$0 means nothing paid: inherent zero

d.) Model year (2023, 2024, 2025, …) interval

differences make sense; year 0 is not "none"

Problem 6 Which data set is at the **nominal** level of measurement?

(A) Finishing places in a 5K race
(B) The blood types of donors at a blood drive
(C) The weights of packages, in pounds
(D) Daily high temperatures, in degrees Fahrenheit

Blood types (A, B, AB, O) are labels only.

HOMEWORK 1.3

Data Collection and Experimental Design

Problem 1 Decide which method of data collection fits best: observational study, experiment, simulation, or survey.

a.) Asking 500 voters which candidate they prefer survey

b.) Recording the number of birds that visit a feeder each morning observational study

Birds are watched, nothing is changed.

c.) Giving one group of plants a new plant food and comparing their growth to plants without it experiment

A treatment (plant food) is applied.

d.) Using a computer model to predict how a flu spreads through a city simulation

Problem 2 A researcher tests whether a new app helps students learn vocabulary. 120 students volunteer. 60 are randomly assigned to use the app for 2 weeks, and 60 study with paper flash cards. Then both groups take the same vocabulary test.

a.) Describe the treatment group. the 60 students who use the app

b.) Describe the control group. the 60 students who use flash cards

c.) Can the students be blinded? Explain. How could blinding still be used?

No. Students know whether they use an app or cards. The people who grade the tests could be blinded: they do not know which group a student is in.

d.) The researcher thinks grade level may affect the results. Describe a randomized block design.

Divide the 120 students into blocks by grade level. Within each grade, randomly assign half to the app and half to the flash cards.

Problem 3 To test a new tire, a company puts the new tires on trucks that drive in the city and its old tires on trucks that drive on highways. The new tires wore out faster.

a.) Identify a problem with this experiment.

Confounding variable: type of road. City driving may wear tires faster, so we cannot tell whether the tire or the road caused the wear.

b.) Suggest a way to fix it.

Randomly assign new and old tires to trucks of each route (block by city or highway), so both tires face the same roads.

Problem 4 Researchers apply a new sunscreen to one arm of each of 50 volunteers and the old sunscreen to the other arm. Which experimental design is this? Explain why it is a good choice.

matched-pairs design

Each person is paired with themselves, so skin type, sun exposure, and age are the same for both treatments.

Problem 5 Identify the sampling technique. Then tell whether the sample could be biased.

a.) A pollster picks a random starting point on a list of phone numbers and calls every 50th number after it. systematic

every 50th → systematic. Not likely biased.

b.) A researcher divides NFL players into offense, defense, and special teams, then randomly selects 20 players from each group. stratified

random from EACH group → stratified.

c.) A teacher surveys the first 10 students who arrive to class. convenience

Easy to reach → convenience. Biased: early students may differ from late students.

d.) A factory randomly selects 5 of its 40 shipping crates and inspects every item in those crates. cluster

ALL items in some crates → cluster.

e.) The ID numbers of all 2,400 students are entered into a random number generator, and 100 numbers are chosen. simple random

Every group of 100 has the same chance.

Problem 6 A club numbers its 45 members from 1 to 45 and wants to choose 6 of them at random.

a.) What should be entered on the TI-84? randInt(1, 45, 6)

randInt(lowest, highest, how many)
randInt(1, 45, 6)

b.) The calculator shows 17, 4, 38, 17, 22, 9. Which members are in the sample? 17, 4, 38, 22, 9, and one more

17 repeats, so skip the second 17.
Press ENTER again to get one more new number.

Problem 7 Decide whether you would take a census or use a sample. Explain.

a.) The mean age of the 9 members of a school board census

Small population: easy to ask all 9.

b.) The mean sugar content of all boxes of cereal a company makes sample

Millions of boxes, and testing opens them.

Problem 8 A sports website asks visitors to vote on whether their city needs a new stadium. Of those who voted, 85% said yes. Explain why this result may be biased.

Only people who visit the site AND choose to vote are counted. Sports fans with strong opinions are more likely to vote yes.

Problem 9 To test a new pain reliever, a company gives it to 6 patients and gives a placebo to 6 other patients. Five of the 6 patients who took the new drug felt better.

a.) Identify a problem with this experiment.

The sample size is too small. With 6 patients per group, the result could be due to chance.

b.) Suggest a way to fix it.

Replication: repeat with a large number of patients, randomly assigned to drug or placebo.

Problem 10 A principal wants the opinions of 40 students about a new schedule. The school has grades 9 through 12. Explain why a stratified sample might be better than a simple random sample.

A simple random sample could, by chance, include very few students from some grade. Randomly choosing 10 from *EACH* grade guarantees every grade is represented.