

Full name (in CAPITAL LETTERS) _____ **SUNet ID** (the part before @stanford.edu) _____ **Section** (e.g., TUES 9:30) _____

Please write legibly so the teaching team can log your attendance: if we can't read your name, we can't give you credit.

PSYCH 10 · Fall 2026 · Worksheet for Lecture 8: Modeling relationships

Work through these with the people sitting next to you; the teaching team will be walking around to help. Credit is for a good-faith effort, not for right answers. Hand it in before you leave.

Part 1 · A mean for each group

In the Week 2 survey, 148 students reported their commute time. The mean of all 148 is 19 minutes. The 120 students who walked or biked have a mean of 13 minutes; the 28 who drove or took the shuttle have a mean of 43 minutes. (All three means are rounded to whole minutes.)

- One student walked and took 22 minutes. The empty model predicts 19 minutes for every student. Compute this student's residual under the empty model, in minutes, with its sign.
- The group model predicts each student's time with their own group's mean. Compute the same student's residual under the group model, in minutes, with its sign.
- For this one student, which model has the smaller residual? Write *empty model* or *group model*.
- Part c compared the two models for one student. To compare them for all 148 students, you would compute one quantity for each model. Name that quantity.

a. _____ b. _____

c. _____

d. _____

Part 2 · How much error did the groups remove?

Eight students from one discussion section reported their commute times. The four who walked: **8, 12, 12, and 16** minutes. The four who took the shuttle: **16, 20, 20, and 24** minutes.

- Compute three means: the walkers' mean, the shuttle riders' mean, and the mean of all eight students. Write all three, in minutes. Then fill in the rest of the table on your sheet; the first row is done.

student	minutes	deviation from the mean of all eight	squared	group mean	deviation from the group mean	squared
walked	8	$8 - 16 = -8$	64	12	$8 - 12 = -4$	16
walked	12					
walked	12					
walked	16					
shuttle	16					
shuttle	20					
shuttle	20					
shuttle	24					
			SS Total =			SS Error =

- b. Compute SS Total: add up the eight squared deviations from the mean of all eight students (the table's fourth column).
- c. Compute SS Error: add up the eight squared deviations from each student's own group mean (the table's last column).
- d. Compute SS Model = SS Total - SS Error.
- e. Compute the proportional reduction in error, $PRE = SS \text{ Model} / SS \text{ Total}$. Write it as a fraction and as a decimal.
- f. Suppose the four shuttle riders had taken 36, 40, 40, and 44 minutes instead (each time 20 minutes longer). For each of SS Total, SS Error, and SS Model, write *larger*, *smaller*, or *unchanged*.

d. _____ e. _____

f. _____

Part 3 • Reading the model's two numbers

In the NHANES survey, 7,182 adults had their height measured. A group model of height by sex, fitted in R, gives these two numbers (in inches, rounded):

(Intercept)	64
Gendermale	5

and these sums of squares: **SS Model 52,000 • SS Error 61,000 • SS Total 113,000.**

- a. What is the mean height of the men, in inches?
- b. Write the model's prediction for a man in the form $b_0 + b_1X_i$, with numbers in place of b_0 , b_1 , and X_i .
- c. What proportion of the variation in height does sex account for? Write the division and its result as a decimal.
- d. An AI assistant summarized this output in three sentences: "The intercept of 64 is the mean height of the women. The coefficient of 5 for Gendermale is the mean height of the men. $PRE = 0.46$ means that the model predicts 46% of adults' heights correctly." Two of the three sentences are wrong. Say which two, and write a corrected version of each.
- e. Sex accounts for about 46% of the variation in height. Name one other variable that could account for some of the other 54%.

a. _____ b. _____

c. _____

d. _____

e. _____

Something new you learned today: _____