



Name_____

Section_____

DEPARTMENT OF STATISTICS
The Wharton School
University of Pennsylvania

Statistics 405/705

Spring 2019

QUIZ #2

Instructions

1. Place your name and section at the top of this page.
2. Keep this side up until you are told to begin.
3. This is a closed-book quiz. You may use a calculator.
4. Circle the single best answer to each question.
5. You will have exactly 10 minutes to complete the quiz.

1. The command `read.csv` reads data from a file into R. What type of data structure is it read into?
 - A. A matrix.
 - B. A vector.
 - C. A list.
 - D. A data frame.**

2. If the variable “a” was a 4 x 3 matrix, then what type of data structure would the command `a[, 2]` return?
 - A. It would depend on whether “a” contained character or numeric elements.
 - B. A matrix.
 - C. A vector.**
 - D. A data frame.

3. In a linear regression model, what type of model would the following R model syntax fit?
$$Y \sim X1 + X2$$
 - A. Two simple regressions, with variables X1 and X2.
 - B. A multiple regression, with variables X1 and X2.**
 - C. A multiple regression with variables X1 and X2 as well as their interaction.
 - D. None of the above.

4. If you include a multilevel categorical variable as a predictor in a regression, what is R’s default way of choosing the “baseline” category?
 - A. It chooses whichever level has the fewest observations as the baseline.
 - B. It chooses whichever level is first in the alphabet as the baseline.**
 - C. It chooses whichever level is last in the alphabet as the baseline.
 - D. By default, R doesn’t use a baseline category, rather it compares each level to the average of all the levels.

5. If you wanted to do a partial-F test on a multilevel categorical variable in R, which command would you use?
 - A. anova**
 - B. lm
 - C. predict
 - D. glm