STAT 3008
Exercise 7

Problems refer to the problem sets in the textbook: Applied Linear Regression, 3rd edition by Weisberg.

1. Problem 4.3. For the transactions data described in Section 4.6.1, define $A = (T_1 + T_2)/2$ to be the average transaction time, and $D = T_1 - T_2$, and fit the following four mean functions:

   $M1 : E(Y|T_1, T_2) = \beta_{01} + \beta_{11}T_1 + \beta_{21}T_2$
   $M2 : E(Y|T_1, T_2) = \beta_{02} + \beta_{32}A + \beta_{42}D$
   $M3 : E(Y|T_1, T_2) = \beta_{03} + \beta_{23}T_2 + \beta_{43}D$
   $M4 : E(Y|T_1, T_2) = \beta_{04} + \beta_{14}T_1 + \beta_{24}T_2 + \beta_{34}A + \beta_{44}D$

   (a) In the fit of $M4$, some of the coefficients estimates are labelled as either "aliased" or as missing. Explain what this means.

   (b) What aspects of the fitted regressions are the same? What is different?

   (c) Why is the estimate for T_2 different in $M1$ and $M3$?

2. Problem 4.4. **Interpreting coefficients with logarithms**

   (a) For the simple regression with mean function $E(log(Y)|X = x) = \beta_0 + \beta_1 \log(x)$, provide an interpretation for β_1 as a rate of change in Y for a small change in x.

   (b) Show that the results of Section 4.1.7 do not depend on the base of the logarithms.

3. Problem 4.7. Suppose we fit a regression with the true mean function

   $$E(Y|X_1 = x_1, X_2 = x_2) = 3 + 4x_1 + 2x_2$$

   Provide conditions under which the mean function for $E(Y|X_1 = x_1)$ is linear but has a negative coefficient for x_1.

4. Problem 4.8. In a study of faculty salaries in a small college in the Midwest, a linear regression model was fit, giving the fitted mean function

   $$E(\widehat{Salary}|Sex) = 24697 - 3340Sex$$

where Sex equals one if the faculty member was female and zero if male. The response Salary is measured in dollars (the data are from the 1970s).

(a) Give a sentence that describes the meaning of the two estimated coefficients.

(b) An alternative mean function fit to these data with an additional term, Years, the number of years employed at this college, gives the estimated mean function

$$\widehat{E(Salary|Sex, Years)} = 18065 + 201Sex + 759Years$$

The important difference between these two mean functions is that the coefficient for Sex has changed signs. Using the results of this chapter, explain how this could happen. (Data consistent with these equations are presented in Problem 6.13).