1. Coefficients in simple linear regression vs multiple linear regression.
Suppose that there are two covariates, X_1 and X_2, which are generated from a bivariate normal distribution with correlation ρ. Assume that a normal linear model holds for Y, so that our observations $i = 1, \ldots, n$ follow the distribution
$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \text{ where } \epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2).$$
   (a) Run a simulation of this problem, and find choices of the parameters ρ, β_0, β_1, β_2, σ^2 such that:
      - If you fit a linear model of Y on covariate X_1 only, then the fitted slope is generally positive,
      - But if you fit a linear model of Y on both covariates X_1 and X_2, then the coefficient $\hat{\beta}_1$ on X_1 is generally negative.

   (b) Give a concrete example of three variables X_1, X_2, Y where you might plausibly expect to see this kind of trend, and explain. Your variables should be intuitive and common quantities, such as income, height, test score, etc.

2. Faraway chapter 3 problem 1

3. Faraway chapter 3 problem 3

4. Suppose that we have a data set following the multiple linear regression model with normal noise,
$$Y_i = \beta_1 X_{i1} + \cdots + \beta_p X_{ip} + \epsilon_i,$$
where $\epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$. (For simplicity, there's no additional intercept term—as in class, if an intercept is needed then it can be one of the p covariates.) Let $\hat{\beta}$ and $\hat{\sigma}^2$ be the usual estimates of β and σ^2 computed via least squares.
Now let $x^{(0)} \in \mathbb{R}^p$ and $x^{(1)} \in \mathbb{R}^p$ be two new covariate vectors, i.e. you have two new points in your data set, with covariate values $x_1^{(0)}, \ldots, x_p^{(0)}$ for the first new data point and similarly $x_1^{(1)}, \ldots, x_p^{(1)}$ for the second. Let $y^{(0)}$ and $y^{(1)}$ denote the response values for these two data points, which follow the same model, but are unobserved.
   (a) What is your estimate for the difference in response values, i.e. for $y^{(0)} - y^{(1)}$?
   (b) Construct a confidence interval around this estimate with coverage level $1 - \alpha$ (e.g. $\alpha = 0.05$ for 95% confidence).
   (c) Construct a prediction interval for the actual difference $y^{(0)} - y^{(1)}$ with coverage level $1 - \alpha$.