Problem 1.

Many factors affect the price of a commodity. These factors fall into two broad categories: fixed costs and variable costs. As an example, let's consider the price charged by a jewelry merchant for a diamond. The variable cost of a diamond depends on its size. A variable cost is the product of the quantity being sold times the cost per unit of quantity. For a diamond, the variable cost is the product of its weight (in carats) times the cost in dollars per carat. (A carat is a unit of weight; one carat is 0.2 gram.) Fixed costs are present regardless of the size of the diamond. Fixed costs include overhead expenses, such as the cost of maintaining the store where diamonds are shown or hosting a Web site to advertise the gems online. The ratio of the cost of a diamond to its weight mixes fixed and variable costs. A one-carat diamond might cost, say, $2,500. That's not the variable cost unless this jeweler has no fixed costs. Fixed and variable costs can be separated by comparing the prices of diamonds of varying sizes. The relationship between the price and weight in a collection of diamonds of varying weights allows us to separate these costs and come to a better understanding of what determines the final cost of a gem.

Questions:

(i) Draw the scatter plot. What kind of model would be appropriate? (ii) Find the ordinary least squares estimates and interpret it. (iii) Find the 95% confidence intervals of the intercept and slope. (iv) Interpret the value of R squared and root mean squared error. (v) Check the assumptions of the model using regression diagnostics.

Problem 2.

Utility companies in many older communities still rely on “meter readers” who visit homes to read meters that measure consumption of electricity and gas. Unless someone is home to let the meter reader inside; the utility company estimates the amounts used. The utility company in this example sells natural gas to homes in the Philadelphia area. Many of these are older homes that have the gas meter in the basement. We can estimate the use of gas with a regression equation

The explanatory variable is the average number of degrees below 65^0 during the billing period, and the response is the number of hundred cubic feet of natural gas (CCF) consumed during the billing period (about a month). The explanatory variable is 0 if the average temperature is above 65^0 Fahrenheit (assuming a homeowner won’t need heating in this case). The intercept estimates the amount of gas consumed for activity unrelated to temperature (such as cooking). The slope estimates the average amount of gas used per 1^0 decrease in temperature. For this experiment, the local utility has 4 years of data (n = 48) months for an owner-occupied, detached home.

Questions:

(vi) Draw the scatter plot. What kind of model would you assume? (vii) Find the ordinary least squares estimates and interpret it. (viii) Find the 95% confidence intervals of the intercept and slope. (ix) Interpret the value of R squared and root mean squared error. (x) Check the assumptions of the model using regression diagnostics.

Problem 3.

THE PRICE OF GASOLINE CAN BE A PAINFUL REMINDER OF THE LAWS OF SUPPLY AND DEMAND.

The first big increase struck in 1973 when the Organization of Petroleum Exporting Countries (OPEC) introduced production quotas. Gas prices had varied so little that accurate records had not been kept. The data shown in Figure 1 begin in 1975, in time to capture a second jump in 1979. After selling for 60 to 70 cents per gallon during the 1970s, the average price rose above $1.40 per gallon in 1981. That jump seems small, though, compared to recent increases. In response to rising prices, Congress passed the Energy Policy and Conservation Act of 1975. This Act established the corporate average fuel economy (CAFE) standards. The CAFE standards set mileage requirements for the cars sold in the United States. The current standard for cars is 30.2 miles per gallon and is slated to increase to 37.8 miles per gallon for 2016 models and to more than 50 miles per gallon by 2025. One way to improve mileage is to reduce the weight of the car. Lighter materials such as carbon fiber and aluminum, however, cost more than the steel they replace. Before investing in exotic materials, manufacturers want evidence of the benefit. What sort of improvements in mileage should a manufacturer expect from reducing the weight of a car by, say, 200 pounds?

The dependent variable is MPG (miles per gallon) and the independent variable is the weight of the car.

Questions:

(i) Draw the scatter plot. What kind of model would be appropriate? (ii) Find the ordinary least squares estimates using a linear regression model and interpret it.

(iii) Find the 95% confidence intervals of the intercept and slope. (iv) Interpret the value of R squared and root mean squared error. (v) Check the assumptions of the model using regression diagnostics. (vi) Consider a transformation of MPG as100/MPG. In other words, run the

regression of 100/MPG (dependent variable) on car weight (independent variable) and carry out all the above steps. What do you think? Should you prefer the latter model? How would you interpret the regression coefficient in this case?