Problem 1. Good forecasting and control of reconstruction activities leads to more efficient use of time and resources in highway construction projects. Data on construction costs (in \$1,000s) and person-hours of labor required on several projects are presented in the following table and are taken from the article, "Forecasting Engineering Manpower Requirements for Highway Reconstruction Activities" (Persad et al., Journal of Management Engineering, 1995). Each value represents an average of several projects, and two outliers have been deleted.

Person-Hours (X)	Cost (Y)
939	251
5796	4690
289	124
283	294
138	138
2698	1385
663	345
1069	355
6945	5253
4159	1177
1266	802
1481	945
4716	2327

(a) Make a scatterplot (with a regression line) of cost versus person-hours. Present the least-squares line for predicting cost (y) from person-hours (x).

(b) Plot the residuals versus the fitted values. Does the model seem appropriate?

(c) Compute the least-squares line for predicting $\ln y$ from $\ln x$, together with a new set of regression plots.

(d) Plot the residuals versus the fitted values. Does the model seem appropriate?

(e) Using the more appropriate model, construct a 95% prediction interval for the cost of a project that requires 1000 person-hours of labor.

Problem 2. The file "bodytemp.csv" contains normal body temperature readings (degrees Fahrenheit) and heart rates (beats per minute) of 65 males (coded by 1) and 65 females (coded by 2).

(a) For both males and females, make scatterplots of heart rate versus body temperature. Comment on the relationship or lack thereof.

(b) Does the relationship for males appear to be the same as that for females? Examine this question graphically, by making a scatterplot showing both females and males and identifying females and males by different plotting symbols.

(c) For the males, fit a linear regression to predict heart rate from temperature. Plot the residuals versus temperature and comment on whether the relationship is linear. Find the estimated slope and its standard error.

(d) Repeat the above for females.

Problem 3. A company which markets and repairs small computers needs to forecast the number of service engineers required over the next few years. This requires consideration of the length of service calls, which in turn depends on the number of components that need to be repaired or replaced. The data given ("repair.csv") consists of the number of components repaired and the length of the service call (in minutes) for a random sample of 24 calls. We will use a simple regression model to explain the relationship between the length of service call (response variable, Y) and the number of repaired units (predictor variable, X).

(a) Make a scatterplot (with a regression line) of Y versus X. Present the least-squares line for predicting Y from X.

(b) Plot the residuals versus the fitted values. Does the model seem appropriate?

(c) Now, consider a multiple linear regression model to fit the data with X and X^2 as predictors. Compare R^2's from SLR and MLR.

(d) Plot the residuals versus the fitted values. Which model seems more appropriate?