

FREC 408

Assignment 8

Issued: Dec. 3, 2003

Due: At the last exam, December 18, 2003

Be sure to:

- Put your name and the Assignment # on the front
- Answer as completely as you can. All I can go on is what you give me, so show your work.
- Be as neat as possible. You can write it out, but please be neat.

1. Single variable regression. On the web site is an Excel file called TAMPALMS.xls 60 pts
The data are from a study of residential property sales to appraisals. A detailed description of the data can be found on page 536 in the book in the Statistics in the Real World box (the graph is on page 541). We will focus on sale price as the dependent variable (Y), and appraised price as the independent variable (X).

- a. Run descriptive statistics on Sale Price and Appraised Value and briefly describe each one.
- b. Use Excel to create a correlation matrix of all the variables in the data set (including S to App Ratio, App Land and App Improve). Briefly describe the matrix, with an emphasis on the relationship between Sale Price and Appraised Value, but also mention any noticeable (meaning high positive or negative) correlations with the other variables.
- c. Use Excel to make a scatter-gram of Sale Price (Y) to Appraised Value (X). Briefly describe the relationship.
- d. Run the regression of Sale Price on Appraised Value. Include the output with you homework. Briefly describe the output noting such things as R^2 and the coefficients from the model.
- e. Conduct the formal hypothesis test that the coefficient for Appraised Value is equal to zero. Make it a two-tailed test with an alpha level of .01. You can use a z-value for the critical value for the test.
- f. The data were in \$1,000 dollars. There is another worksheet that has the data re-expressed in dollars. Re-run the regression of Sale Price on Appraised Value on the transformed data and note any differences in the output. Note: the only difference is that the data have been multiplied through by a constant (i.e., * \$1,000).

2. Dummy Variable Regression. On the web site is an Excel file called police.xls 40 pts
The data are from a study done by a police department to determine which cars to purchase for their fleet. Before changing the whole fleet they buy 21 cars of three different models. They test the cars for a few months and record the average cost per mile (the dependent variable). We looked at this data with ANOVA. You will run a regression on the data. The worksheet labeled regression has the data reorganized for regression, with three CAR variables that are coded as 0/1 dichotomies. You must choose to omit one of the CARs from the model to serve as the reference category.

I want you to run two regression models. The first one has CAR1 and CAR3 as independent variables; the second one has CAR3 and CAR2 as independent variables.

- a. Run the regressions and show the output as part of your answer.
- b. Briefly compare the two regression models - what is similar and what is different between the two models.
- b. Pick one model and conduct the formal F-test of the overall model (that the means differ)
- c. For the model you choose, solve the equation and predict the average cost for each car model. Show that the same prediction would come from the other model (i.e., it doesn't matter which CAR you leave out of the model, the predictions would be the same).
- d. In the same file is the ANOVA worksheet. Confirm that the results (Sums of Squares, F-test, and R-square) are the same as for the regression.