DEPARTMENT OF POLITICAL SCIENCE
AND
INTERNATIONAL RELATIONS

Posc/Uapp 816

Assignment 4
TWO VARIABLE REGRESSION

Name______________________________
(Printed)
Student Number____________________
(Social Security Number)
E-mail____________________________

The first problem comes from Agresti and Finlay, ***Statistical Methods for the Social Sciences***, 3rd edition, although you don't need the book to work on it.

1. The authors write "In an article in ***USA Today*** (December 28, 1884), sociologists N. Glenn and B. A. Shelton are quoted as showing a strong link between residential mobility and divorce rates. In [the table below], divorce rate is the annual number of divorces and annulments per 1000 population, and mobility rate is the percentage of people living in a different house from five years ago." (Page 350) The data pertain to Census Bureau regions of the country.

Region	**Mobility rate**	**Divorce rate**
New England	41	4.0
Middle Atlantic	37	3.4
East North Central	44	5.1
West North Central	46	4.6
South Atlantic	47	5.6
East South Central	44	6.0
West South Central	50	6.5
Mountain	57	7.6
Pacific	56	5.9

A. Although the direction of causation is not clear, let's treat divorce as the dependent variable and see how much variation in it can be "explained" by mobility.

B. Attach a neatly labeled plot that treats the mobility rate as the independent variable.
   i. You can use the annotation features of MINITAB to add titles and labels your plot.

C. Now estimate an OLS squares model:___________________________________
   i. Print clearly print the equation and do not use column numbers for variable names.

D. What is R^2____________

E. Give a substantive interpretation of the results..

____

____

____

2. Speaking of mobility, it might be interesting to see if population shifts have any impact on a community's quality of life. In particular, we hear that as people move to the suburbs they leave urban areas with crime and poverty. Here are some data taken from the ***Statistical Abstract of the United States 1997***. They pertain to cities with more than 100,000 population. "Change" refers to the percent change in population from 1990 to 1994; the total crime rate is the total crime index and murder is the index for homicides per 100,000 population.

   A. A negative percent means, of course, that the city lost population.

   B. Here are the data (next page).

City	Percent change	Total crime	Murder
Los Angeles, CA	-1.1	7680.0	24.5
Dallas, TX	1.5	9464.1	26.5
Los Vegas, NV	27	7584.5	14.9
Baltimore, MD	-4.5	13318.4	45.6
Milwaukee, WS	-1.8	8462.9	22.2
El Paso, TX	12.4	7063.9	6.3
Washington, D.C.	-6.6	12166.4	65.2
Denver, Co	5.5	6873.5	16.0
Oklahoma City, OK	4.2	11501.8	48.7
Virginia Beach, VA	9.5	4651.8	3.7
St. Louis, MO	-7.2	16082.9	54.9
Colorado Springs, CO	.6	6765.2	5.5
Buffalo, NY	-4.6	9205.3	19.8
Louisville, KY	.3	7149.0	18.3
Jersey City, NJ	-1.1	7946.0	11.0

C. **Note: as you enter these into MINITAB (or SPSS) make sure that you include the minus signs but do not use commas.**

3. Attach a neatly labeled plot that treats "Total crime rate" as the dependent variable and "Percent change" as the explanatory (independent) factor.

A. Very neatly sketch by hand an approximate least squares line through the data.

B. Now use the regression procedure to find the regression constant. ___________

What does it mean in substantive terms?

____

____

C. What is the regression coefficient? _____________

D. What does it mean?

____

____

E. What is R^2? ________________

4. Let's finally analyze the impact (if any) of population change on murder rates.
   A. First attach a neatly labeled plot. (Murder is the dependent variable.)
      i. Do you see an "outlying" value? Circle it.
   B. What is the estimated regression coefficient? ______________
   C. What is the R^2? _________________
   D. Use the manipulate commands (e.g., copy columns) to eliminate case 3 from the analysis.
      i. For example, you can tell MINITAB to copy columns to new columns **while omitting** row 3.
   E. Now attach another appropriately labeled plot. (The title or note should indicate that one case has been removed.)
   F. Re-estimate the model on the "reduced" data set (i.e., N = 14).
   G. What is the R^2? ______________________
   H. What lesson does this little example illustrate?

____

____

Go to Assignment page

Go to Statistics page