



SPSS Regression and Importing Data

Fall 2012

Regression

EarthQuakeSmall.sav

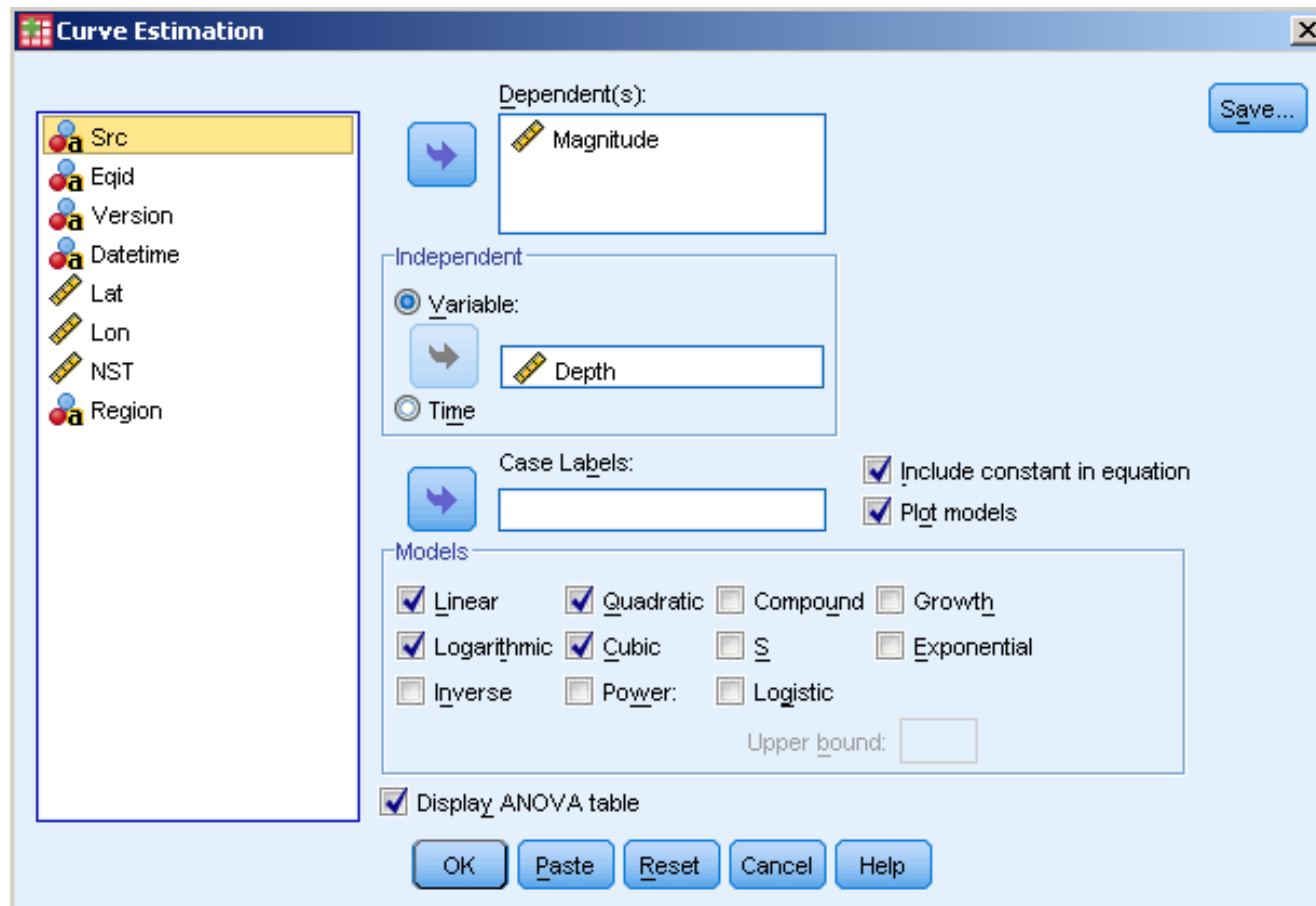
in CS 130 Public

- Analyze |
- Regress |
- Curve Estimation

The screenshot shows the PASW Statistics Data Editor interface. The title bar reads "EarthQuake7DayM25_orig.sav [DataSet1] - PASW Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The Analyze menu is open, showing options like Reports, Descriptive Statistics, Compare Means, General Linear Model, Generalized Linear Models, Mixed Models, Correlate, Regression (highlighted), Loglinear, Classify, Dimension Reduction, Scale, Nonparametric Tests, Forecasting, Survival, Multiple Response, and Quality Control. The Regression submenu is also open, showing options like Linear..., Curve Estimation..., Partial Least Squares..., Binary Logistic..., Multinomial Logistic..., Ordinal..., Probit..., Nonlinear..., Weight Estimation..., and 2-Stage Least Squares... The data grid shows columns for Src, Eqid, and other variables. The status bar at the bottom indicates "Monday, February 28, 2011 13:15:53 UTC".

	Src	Eqid	
1	us	b0001mgd	4
2	us	b0001mg2	5
3	nc	71530295	1
4	nc	71530230	6
5	ak	10182565	1
6	us	b0001mfa	9
7	us	b0001mej	5
8	us	b0001m9j	6
9	us	b0001m4j	7
10	ak	10182353	2
11	us	b0001lzv	5
12	us	2011hxbb	3
13	ak	10182270	1
14	nm	022811f	A
15	nm	022811e	A

Regressions continued



Model Summary and Parameter Estimates

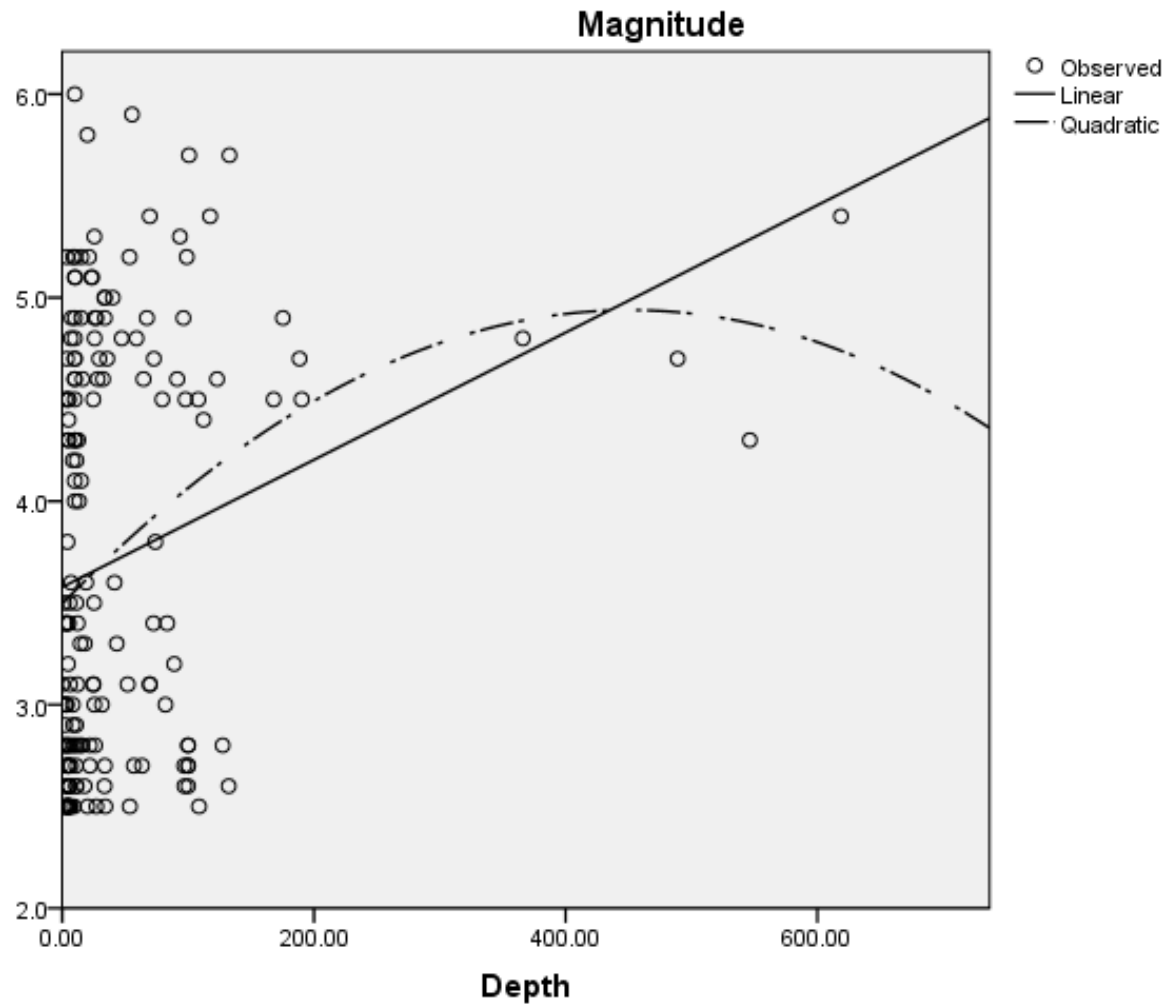
Dependent Variable: Magnitude

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	Constant	b1	b2
Linear	.061	11.765	1	181	.001	3.578	.003	
Quadratic	.074	7.173	2	180	.001	3.495	.006	-7.097E-6

The independent variable is Depth.

Equations?

R²?



Importing Data

- Let's import some data from a table on a web page. The following steps will get us weather data for Forest Grove for September 2012.
- In a web browser, go to:
<http://www.wunderground.com/history/>
- Location: 97116
September 1, 2012
Submit
Custom
September 1, 2012 to September 25, 2012
Go

Importing Data

- Scroll down to the Daily Observations table and click on Comma Delimited File at the bottom. This converts the table to a text file with the items separated by commas. Select the text and copy it.
- Open Excel and paste the data into cell A1.

Convert to Columns

- Now we need to convert the text to columns. Go to the Data tab and click on Text to Columns. Select Delimited, and on the next screen select commas.
- Your data should now be in the correct format.

Preparing Data for SPSS

- Since SPSS needs the data in a particular format, we need to edit the spreadsheet. The first row needs to contain the variable names that SPSS will use during the import. Remember, spaces are not your friend.
- Edit row 1 to remove all spaces. In the names.

Preparing Data for SPSS

- Add a DayOfWeek column before the MaxTemperatureF column.
- Insert a column, add a heading DayOfWeek, and type Saturday in the second row (Sep 1, 2012 was a Saturday). Click and drag Thursday to fill in the remaining rows. Excel should fill out the days of the week correctly.

Opening Data in SPSS

- Save this as an Excel file (Wunderground-Sep2012) and close the workbook.
- Open the new Excel file in SPSS. Check the definition of each variable and make any changes you deem necessary.
- Add values to DayOfWeek (1=Sunday, 2=Monday....)
- Transform | Recode into Same Variables...
 - Map: Sunday=1, etc for DayOfWeek.

Charts

- Build a chart to show the mean temperature over time, from September 1 to September 25.
- What type of chart should you use? Why?
- What is on the X-Axis? Why?
- What is on the Y-Axis? Why?

Charts

- Build a chart to show the mean high temperature per day of the week, from September 1 to September 25. (This should show the mean temperature of Sunday, Monday, Tuesday, etc. in the correct order).
- What type of chart should you use? Why?
- What is on the X-Axis? Why?
- What is on the Y-Axis? Why?

Charts

- How well does average humidity predict low dew point?

Tree Data

- Using the sample data set entitled "TreeData.txt" found in the CS130 Public folder and SPSS create a dataset called **TreeData.sav** and report in the Word document PUNetIDAnswers.doc the answers to each of the following questions.
- When asked for, place a graph into your document with the appropriate explanation.

Tree Data: Descriptive Stats

- List each variable in the dataset TreeData.sav. Further, list the type and measure that each variable should be and briefly explain why.
- What is the mean, median, mode, and standard deviation for each of the variables: (a) Trunk Girth and (b) Weight.
- Construct a single bar chart that shows the Mean Weight of each root category.

Tree Data: Regression

- Using SPSS, perform the correct linear regression on weight and trunk girth. Make sure you properly identify the Dependent and Independent variable. Paste in the **Coefficients** table and the scatter chart.