Self-Assessment
Weeks 9: Multiple Regression with Both Qualitative and Quantitative Predictors; Multiple Comparisons

1. What is an adjusted mean? What potential benefit does it offer? How could it potentially be misleading?

2. Below is linked blood pressure data that was presented in previous self-assessments. Perform comparisons of systolic blood pressure among drugs (Ziac, Losartan, Lisinopril 40mg, Lisinopril 12.5mg) while controlling for weight, i.e.,

Systolic Blood Pressure = b0 + bj Drug + bi Weight

Where bj represents several drug coefficients and bi is an unnumbered coefficient for Weight.

Use the Bonferroni adjustment and set familywise alpha to .05. Present results in APA style. Also, present a table showing the predicted systolic blood pressure for someone who weighs 220, 200, and 180 pounds for each of the four treatments.

SPSS
http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week09/BloodPressureDrugs.sav

Excel
http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week09/BloodPressureDrugs.xlsx

3. Below is a data file containing the following variables for cars taken between 1970 and 1982:

mpg:	miles per gallon
engine:	engine displacement in cubic inches
horse:	horsepower
weight:	vehicle weight in pounds
accel:	time to accelerate from 0 to 60 mph in seconds
year:	model year (70 = 1970, to 82 = 1982)
origin:	country of origin (1=American, 2=Europe, 3=Japan)
cylinder:	number of cylinders

SPSS Data: http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week04/cars_missing_deleted.sav
(Note: There are underscore marks between words in the SPSS data file name.)
Other Data Format: If you prefer a data file format other than SPSS, let me know.

For this problem we wish to know whether MPG differs among car origins and number of cylinders while controlling for the weight of the car. The regression model for this study follows:

Predicted MPG = b0 + origin of car + number of cylinders + car weight

Origin of car is categorical. Number of cylinders may appear to be ratio, but since observed categories of this variable are limited, it is best to treat this variable as categorical. Note the following number of cylinders reported:

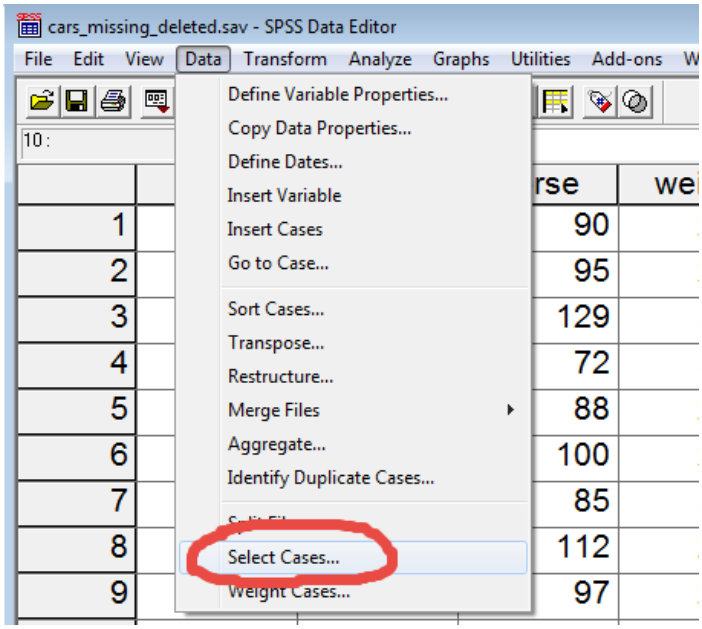
Number of Cylinders

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3 Cylinders	4	1.0	1.0	1.0
	4 Cylinders	199	50.9	50.9	51.9
	5 Cylinders	3	.8	.8	52.7
	6 Cylinders	83	21.2	21.2	73.9
	8 Cylinders	102	26.1	26.1	100.0
	Total	391	100.0	100.0	

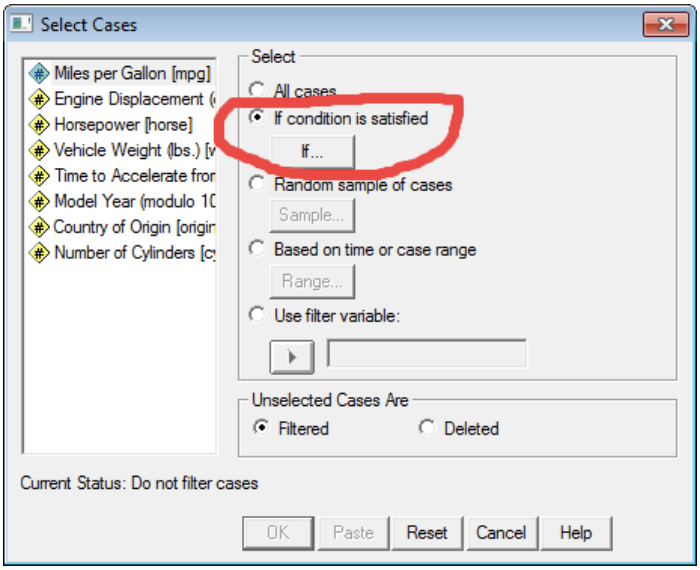
As the frequency display above shows, the number of cylinders include 3, 4, 5, 6, and 8. However, only 4 cars had 3 cylinders and only 3 cars had 5 cylinders. Given the small sample sizes for these categories, it is best to remove these cases from the regression analysis. There are several ways to accomplish this. Four approaches are (a) manually delete these cases after sorting all cases on number of cylinders, (b) telling SPSS to treat these 7 cases as missing values so they will not be included in any analysis (use Recode into Same Variable and set 3 Cylinders and 5 Cylinders as system missing), (c) defining 3 and 5 Cylinders as missing values in the variable missing values, or (d) using the Select Cases command to filter these cases from all analyses. Other possibilities also exist.

Of these four, option (d) works well and does not require deletion of any cases. This option is explained below.

Step 1: Open the Select Case window



Step 2: Choose the select If option



Step 3: Define the filter so SPSS can determine which cases NOT to select.

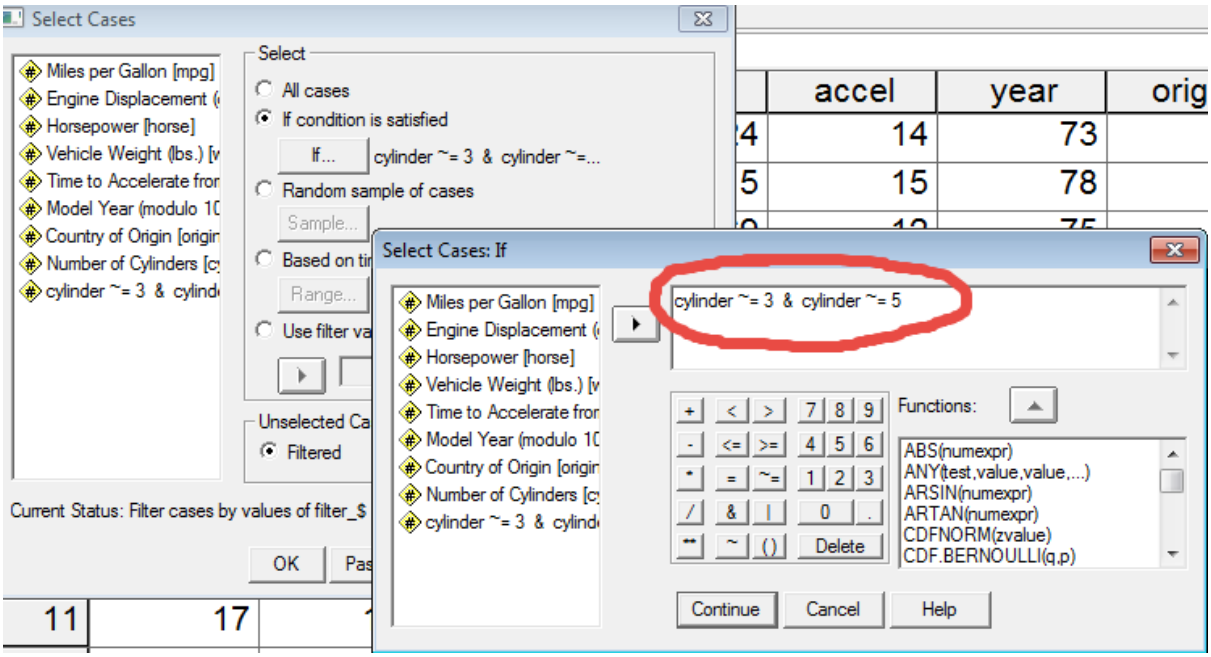
We do not want cylinders of 3 or 5, so in the Select Cases IF box, write

cylinders ~= 3

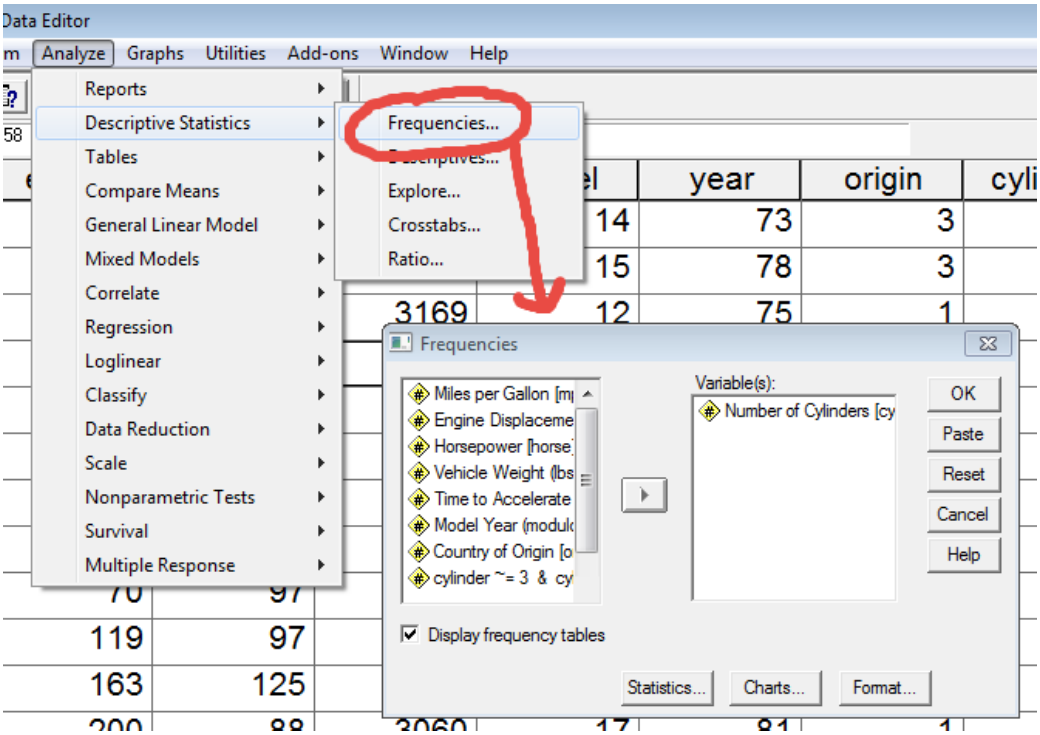
The symbol ~= means “not equal”; this tells SPSS not to select any cases in which cylinders are 3. Also, write

cylinders ~= 5

so SPSS knows not to select cases when cylinders are 5. To combine these two, we use the ampersand symbol, &, which means select all cases which are not 3 and 5 cylinders. See image below.



Once these cases are defined, click Continue then OK to process this command. Next, check that the appropriate cases were selected by running the Frequency command for cylinders as shown below.



And the results should look like this:

Frequencies

Statistics

Number of Cylinders

N	Valid	384
	Missing	0

Number of Cylinders

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4 Cylinders	199	51.8	51.8	51.8
	6 Cylinders	83	21.6	21.6	73.4
	8 Cylinders	102	26.6	26.6	100.0
	Total	384	100.0	100.0	

Note that no cases of cylinders equal to 3 or 5 were selected.

Present an APA styled regression analysis with DV = MPG, IV = origin, IV = Cylinders (4, 6, and 8 only), and IV = vehicle weight. Set alpha = .01. You will have to create the dummy variables for origins and cylinders. Also present Scheffé confidence intervals comparisons among origins and among cylinders.

In addition to APA styled results, present literal interpretations for each regression coefficient.