HYPOTHESIS TESTING: CONFIDENCE INTERVALS, T-TESTS, ANOVAS, AND REGRESSION

HOD 2990 • 10 November 2010

Lecture Background

- This is a lightning speed summary of introductory statistical methods for senior undergraduate students in the honors program. They had designed their own research studies and had collected quantitative data.
- The previous lecture had covered data entry, data cleaning, and univariate and bivariate analyses.
- The lecture was occurring in a computer lab. The class had access to a common dataset collected from undergraduate students. After I introduced each new analytic method, we took time to do the procedure and explore the output thoroughly.
- Lecture time is four hours.
- Students were not expected to be experts on statistical analysis themselves. Rather, they should have been able to know possible avenues of analysis and begin to analyze their data with significant support

Inferential Statistics

- The only way to know the true mean (or any other parameter) of a population is by surveying the entire population
- When we take only a sample of the population, we are *inferring* or generalizing known characteristics of the sample to the population
- Unfortunately, there is always uncertainty about whether the characteristics of our sample reflect the characteristics of the population
  - This is true even if we have a representative sample and followed the best sampling methods available

Confidence Intervals

Confidence Intervals

- The confidence interval is an expression of that uncertainty, expressing an area around the sample mean we think the population mean is likely to fall

- The larger our sample, the smaller the confidence interval
  - The larger the sample, the more sure we are that the sample mean approximates the population mean

Confidence Intervals

- There are different confidence intervals, depending on how sure we want to be that the population mean falls within the CI
  - The more sure we want to be, the wider the range of the confidence interval
- $CI = \dot{X} \pm {}_{1-\alpha/2}z(\sigma/\sqrt{n})$
  - 90% CI: $\dot{X} \pm 1.645(\sigma/\sqrt{n})$
  - 95% CI: $\dot{X} \pm 1.960(\sigma/\sqrt{n})$
  - 99% CI: $\dot{X} \pm 2.576(\sigma/\sqrt{n})$

Confidence Interval

- Interpretation: if we took the a sample from the population repeatedly, in [90, 95, 99, or whatever confidence level we specify] percent of the samples the true population parameter will fall within the range generated by the CI equation
- Or, we are X% sure that the population mean falls within the range of the CI

CAUTION

Before we go on to examine hypothesis testing, there are several operative assumptions:

- Your data is properly entered and cleaned
- You know the univariate and bivariate distributions of your data

T-tests

T-tests

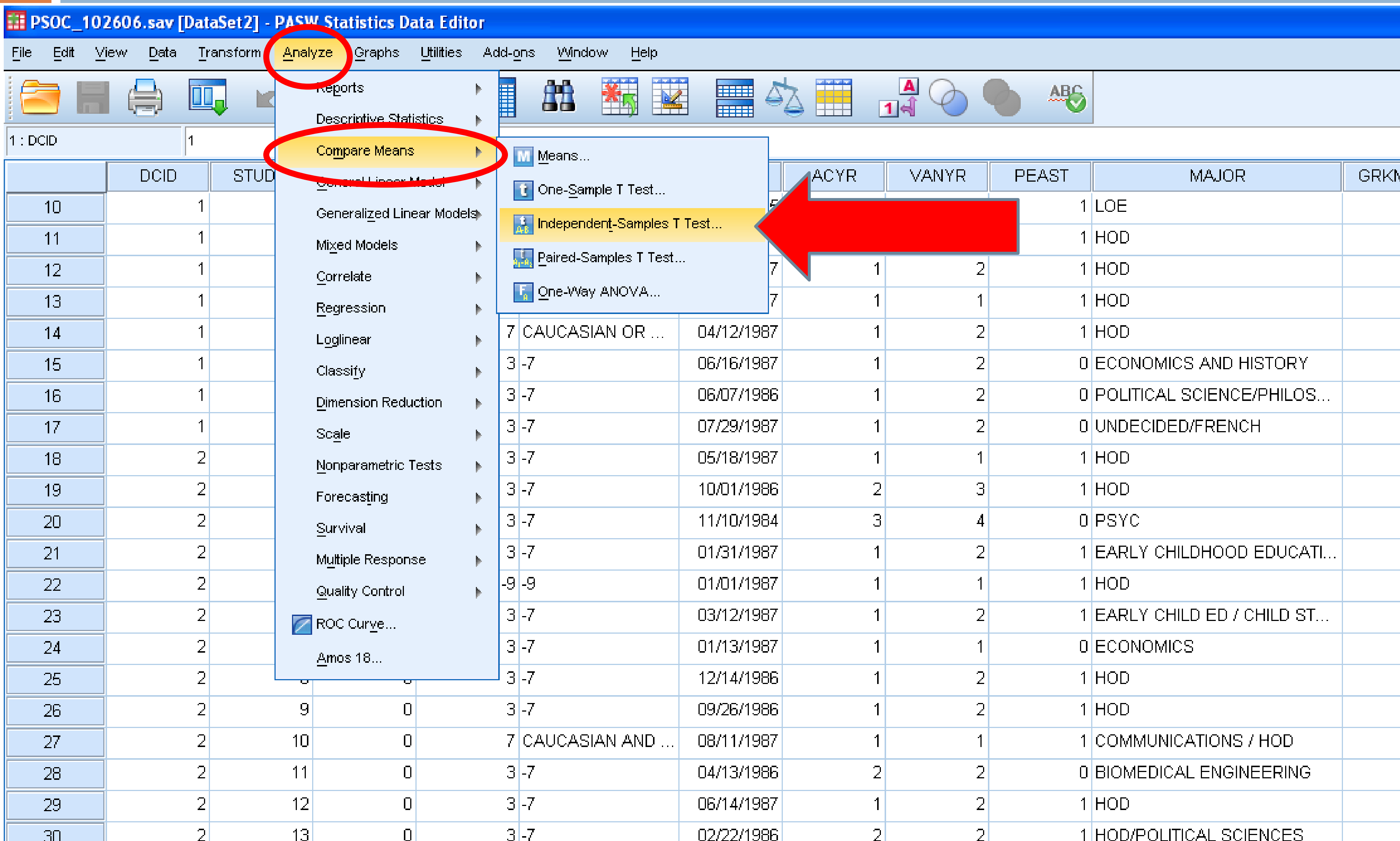
- When we have a continuous variable and we want to know if its mean differs in value between two groups, we use a *t*-test
  - It is extremely rare that the groups will have exactly the same mean, but the t-test compensates for the uncertainty involved in measuring a population using a sample—basically, it measures whether or not the group means are far enough apart that we can be confident that they come from different distributions
  - Example: Do male and female undergraduate students at Vanderbilt have different scores on the Satisfaction with Life scale?

T-test assumptions

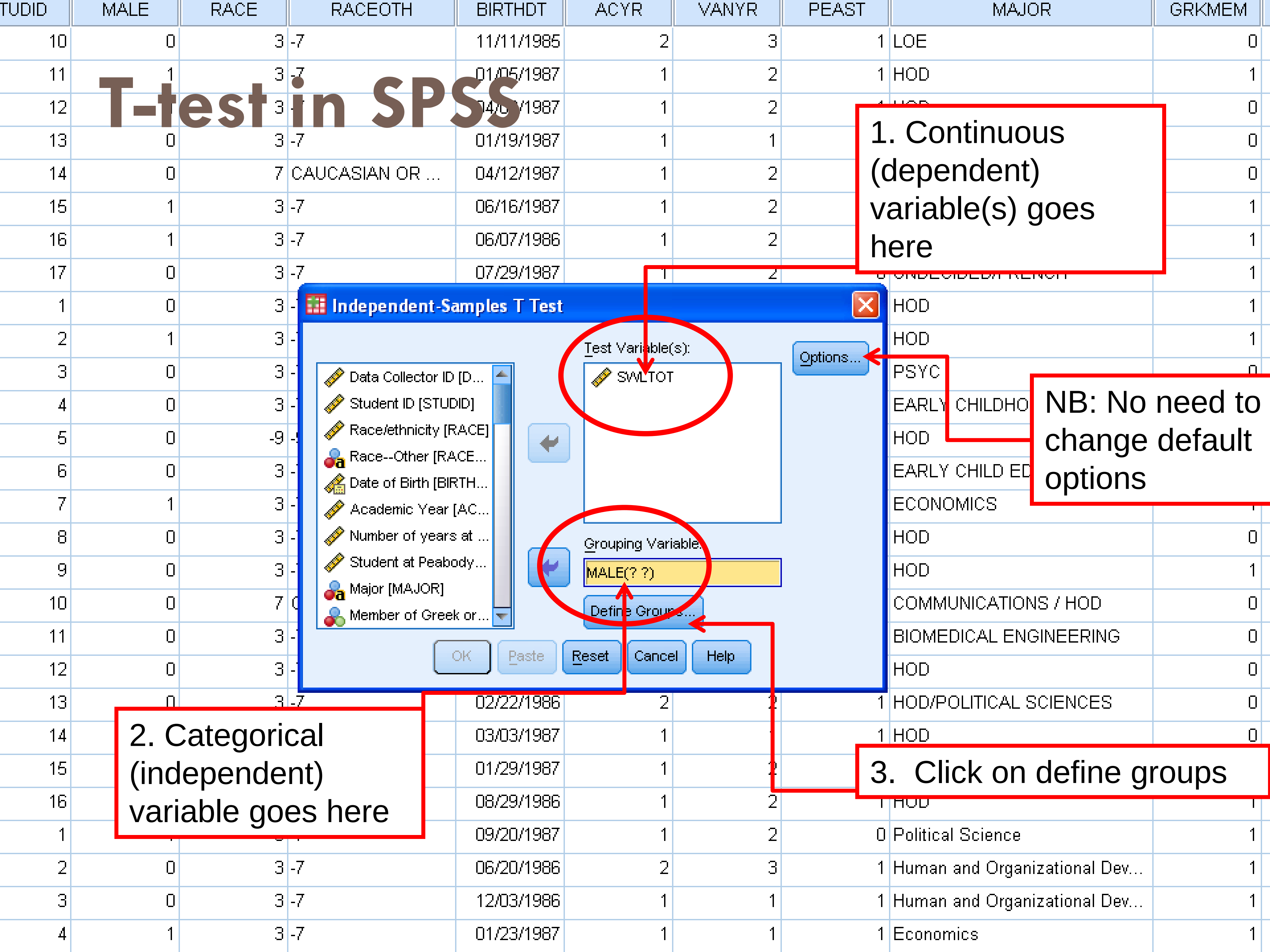
1. The dependent—continuous—variable (Satisfaction with Life) is normally distributed when it's examined on both categories
2. The dependent variable has a more-or-less equal variance / standard deviation when broken up into the two categories
3. The cases are independent from each other

T-tests are, however, reasonably robust even if these assumptions are violated

T-test in SPSS



T-test in SPSS



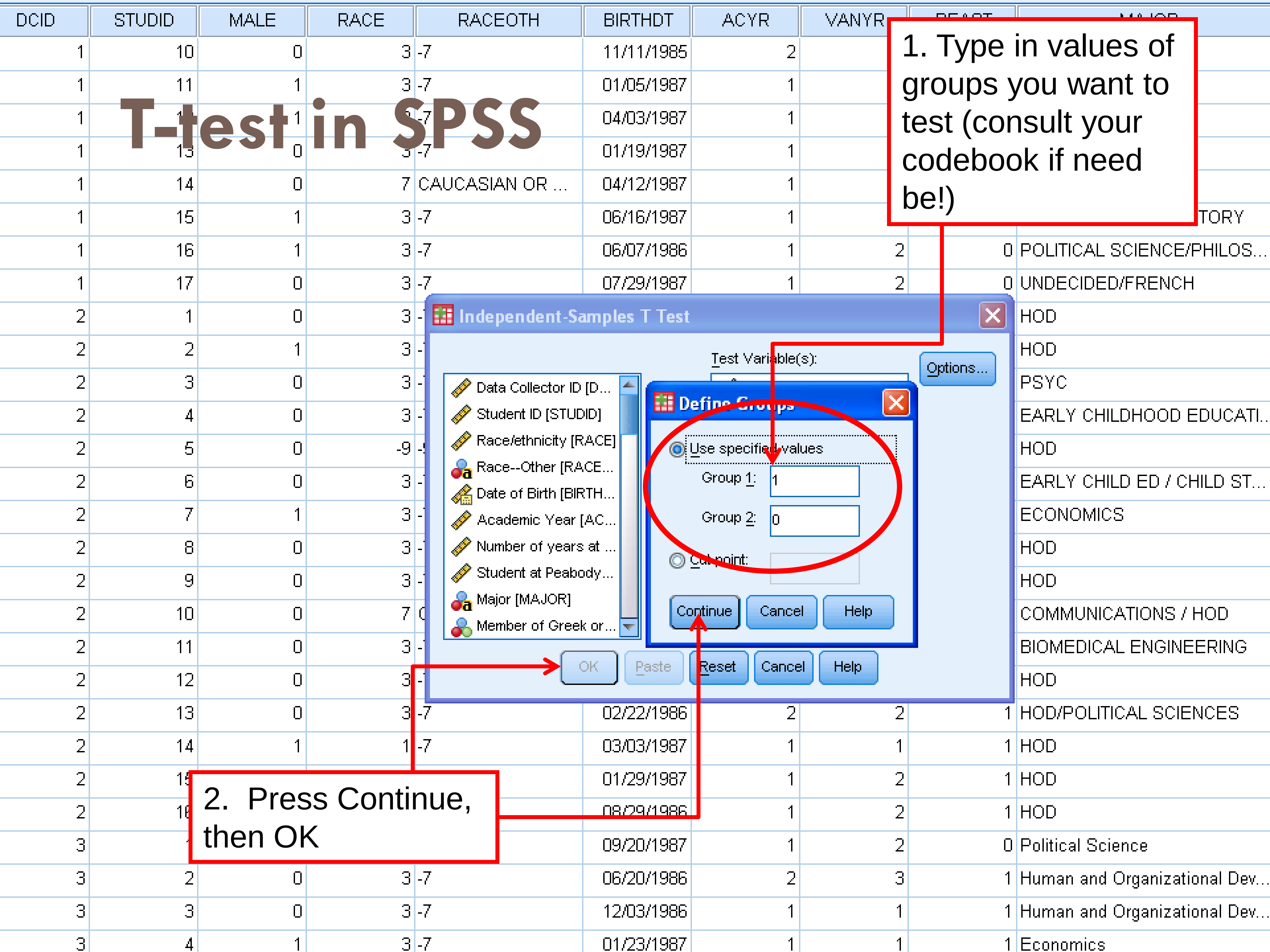
1. Continuous (dependent) variable(s) goes here

NB: No need to change default options

2. Categorical (independent) variable goes here

3. Click on define groups

T-test in SPSS



1. Type in values of groups you want to test (consult your codebook if need be!)

2. Press Continue, then OK

T-test: Interpreting the output

Group Statistics

	Male gender	N	Mean	Std. Deviation	Std. Error Mean
SWLTOT	Male	46	3.8087	.59473	.08769
	Female	79	3.9456	.61092	.06873

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SWLTOT	Equal variances assumed	.088	.767	-1.220	123	.225	-.13687	.11221	-.35900	.08525
	Equal variances not assumed			-1.229	96.309	.222	-.13687	.11142	-.35802	.08427

T-test: Interpreting the output

Group Statistics

	Male gender	N	Mean	Std. Deviation
SWLTOT	Male	46	3.8087	.59473
	Female	79	3.9456	.61092

This information tells you whether or not the second assumption of t-tests, equality of variances, has been violated. If the number under Sig is less than 0.05, there is no violation. If the Sig < .05, then there is a problem—BUT SPSS has automatically made corrections for you.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SWLTOT	Equal variances assumed	.088	.767	-1.220	123	.225	-.13687	.11221	-.35900	.08525
	Equal variances not assumed			-1.229	96.309	.222	-.13687	.11142	-.35802	.08427

If the variances are not equal, use the second line

If the variances are equal, use this first line

T-test: Interpreting the output

Group Statistics

	Male gender	N	Mean	Std. Deviation	Std. Error Mean
SWLTOT	Male	46	3.8087	.59473	.08769
	Female	79	3.9456	.61092	.06873

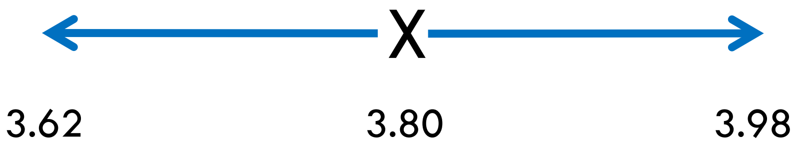
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SWLTOT	Equal variances assumed	.088	.767	-1.220	123	.225	-.13687	.11221	-.35900	.08525
	Equal variances not assumed			-1.229	96.309	.222	-.13687	.11142	-.35802	.08427

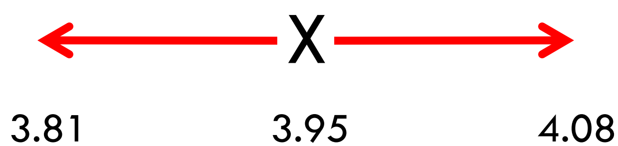
This significance result tells you the probability that we'd be making an error if we said there was a difference in the scores between male and female students. By convention, there needs to be less than a 5% (.05) chance of error if we are to declare that there is a difference. Alternately, we need to be 95% sure that there is a difference in the population means based on our sample means.

Understanding t-tests

Male



Female



T-tests and confidence intervals

- If the value of one of the group means falls within the confidence interval (degree of uncertainty) of the other, then the risk of making an error when saying the two have different means is unacceptably high

- SPSS reports the confidence intervals of the mean difference (mean of group one minus mean of group two)
  - If there is no significant difference, the confidence interval will contain zero

T-tests: Reporting results

- Unless you have a whole bunch of t-tests, there is no need for a table

- The means & standard deviations for both groups, *t* value, degrees of freedom, and *p* value all need to be reported.

- Text: There was no difference in satisfaction with life scores in male and female students (M_{male} = 3.81, SD_{male} = .59, M_{female} = 3.95, SD_{female} = .61; $t(123) = -1.22$, $p = .23$)

ANOVA

ANOVA

- When we have a continuous variable and we want to know if it differs in value between <u>three or more</u> groups, we use ANOVA (ANalysis Of VAriance)
- ANOVA asks whether the mean of all groups are equal to each other or not
  - Example: Do students at different racial groups at Vanderbilt University differ in Sense of Community (Emotional Connection)?

ANOVA assumptions

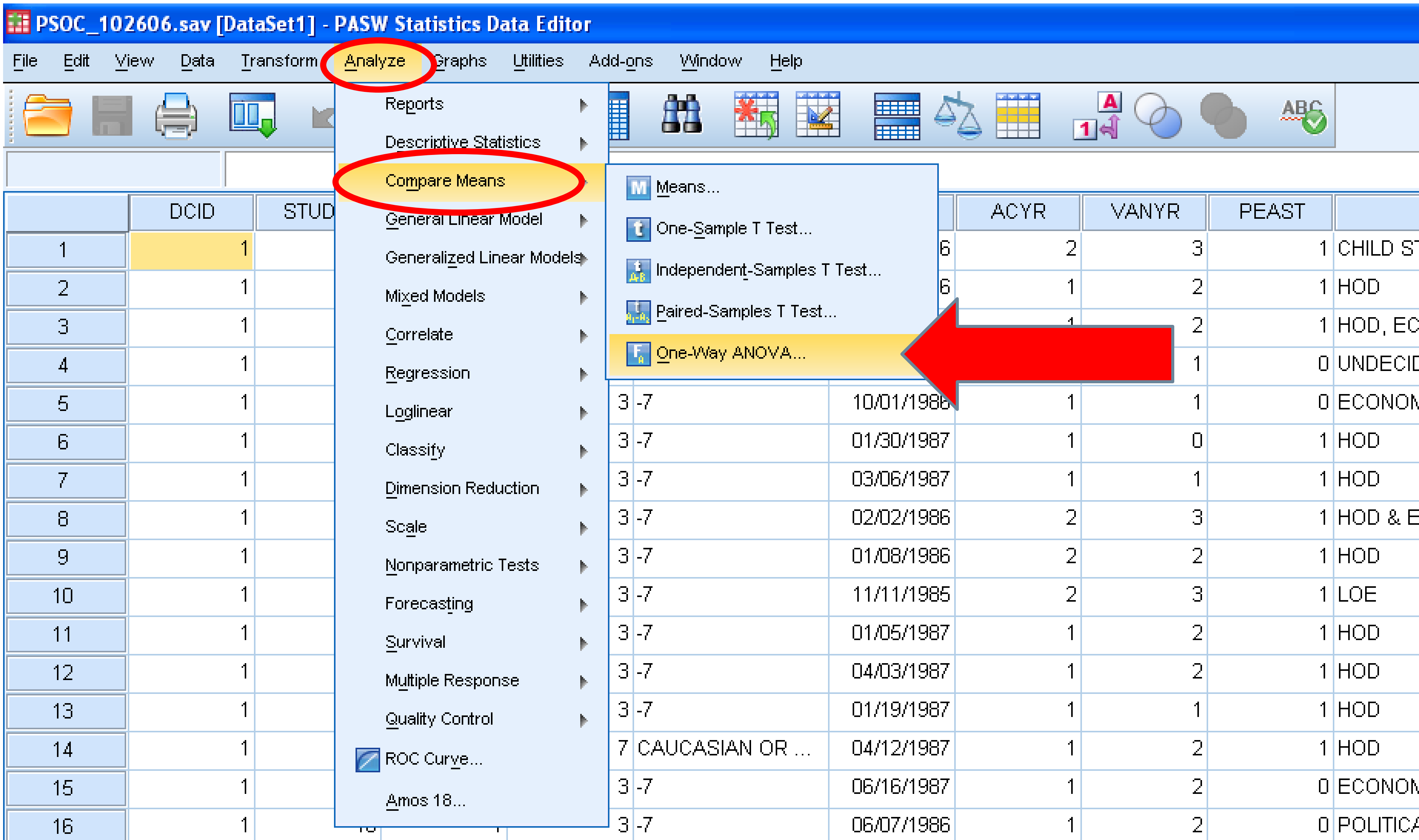
- The dependent variable (Satisfaction with Life) is normally distributed when it's examined individually for all categories

- The errors are normally distributed

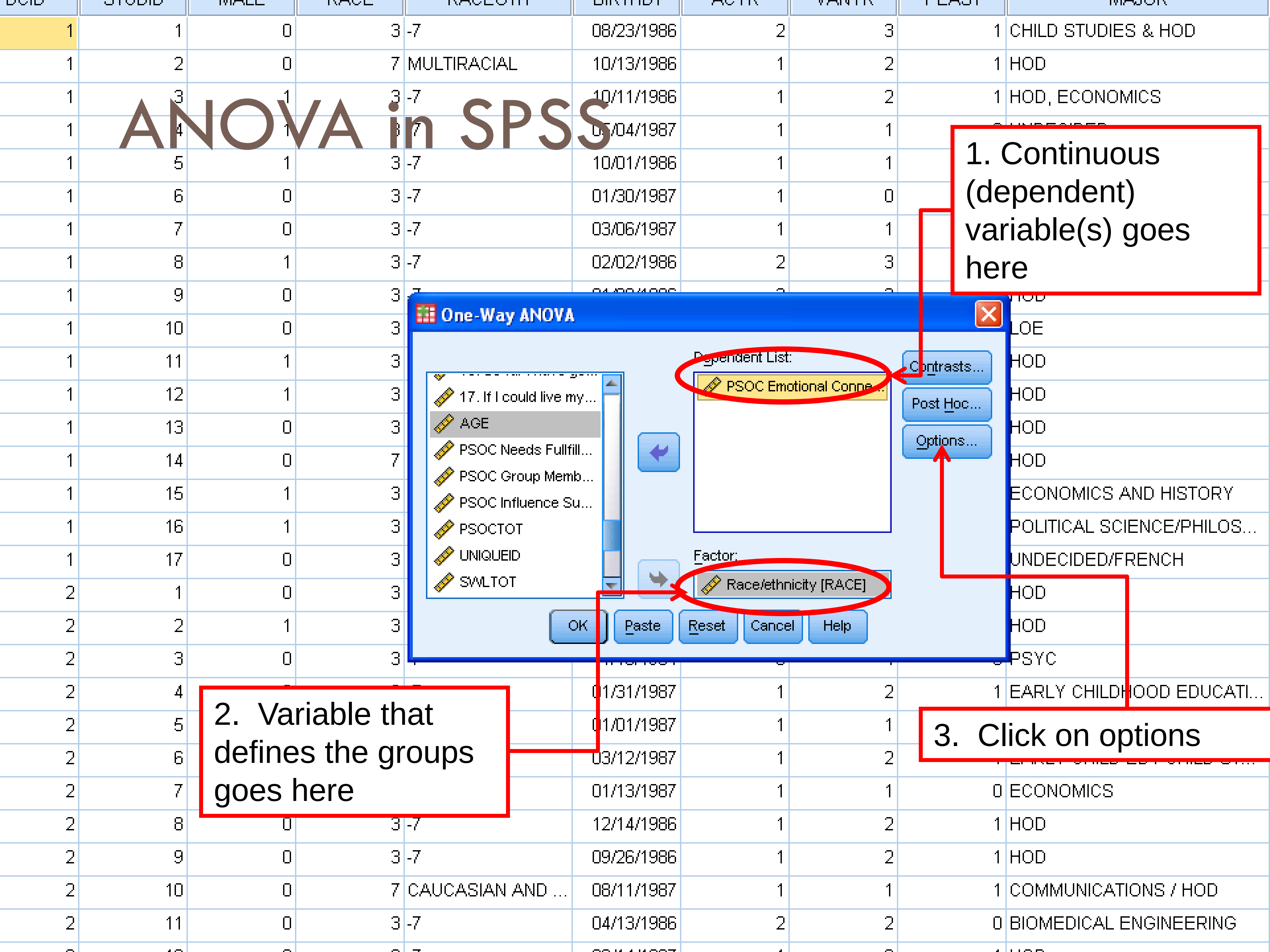
- The cases are independent from each other

Like the t-test, ANOVA is robust in the face of relatively minor violations of these assumptions

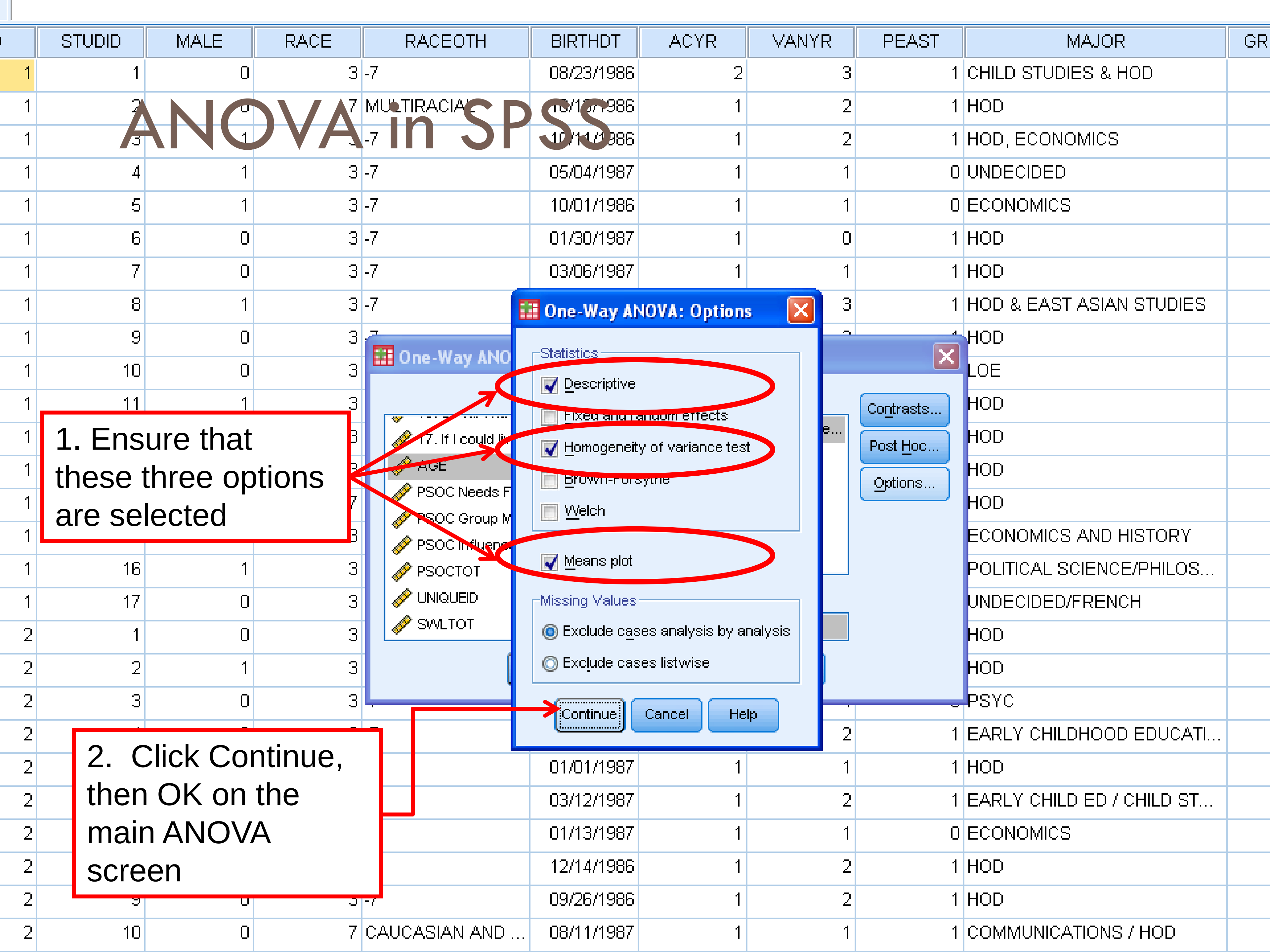
ANOVA in SPSS



ANOVA in SPSS



ANOVA in SPSS



1. Ensure that these three options are selected

2. Click Continue, then OK on the main ANOVA screen

ANOVA: Interpreting the Output

Descriptive statistics, including the confidence interval for each mean

Descriptives

PSOC Emotional Connection Subscale

	N	Mean	Std. Deviation	Std. Error	95% Co[illegible] Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Africa-American, Black	3	4.4444	.19245	.11111	3.9664	4.9225	4.33	4.67
Asian-American, Asian, including Indian Subcontinent	3	3.0000	1.52753	.88192	-.7946	6.7946	1.67	4.67
Caucasian or white	110	4.2091	.43275	.04126	4.1273	4.2909	2.67	5.00
Mexican-American, Chicao/a	2	4.3333	.00000	.00000	4.3333	4.3333	4.33	4.33
Other (describe)	6	3.8333	.34960	.14272	3.4664	4.2002	3.33	4.33
Total	124	4.1694	.50202	.04508	4.0801	4.2586	1.67	5.00

Test of Homogeneity of Variances

PSOC Emotional Connection Subscale

Levene Statistic	df1	df2	Sig.
7.873	4	119	.000

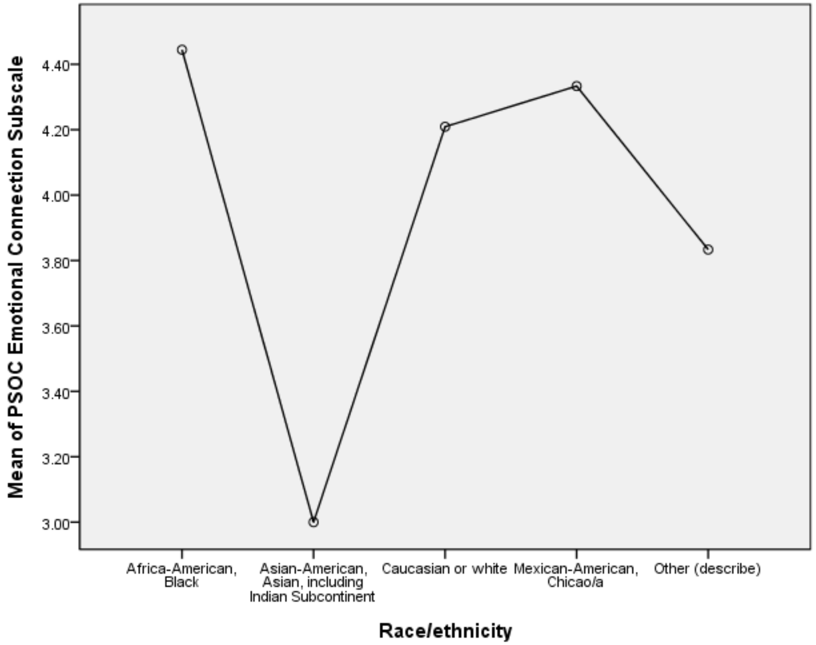
This information tells you whether or not the second assumption of ANOVA, equality of variances, has been violated. If the number under Sig is more than 0.05, there is no violation. If the Sig < .05, then there is a problem, especially if the number of cases in each group is unbalanced.

ANOVA: Interpreting the Output

ANOVA

PSOC Emotional Connection Subscale

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.234	4	1.309	6.044	.000
Within Groups	25.765	119	.217		
Total	30.999	123			



This significance result tells you the probability that we'd be making an error if we said there was a difference in the mean scores of the different groups. Again, there (usually) needs to be less than a 5% (.05) chance of error if we are to declare that there is a difference.

ANOVA: A caution

- A positive result in an ANOVA ($p < .05$) ***only indicates that one or more of the means is not equal to the other means in the group. It does not tell us which group(s) is the one that doesn't belong.***
  - Sometimes the probable group can be identified just by looking at the descriptive statistics or means plot
  - Otherwise, there are post-hoc tests to help identify the group(s)

Writing up the results

- Report means and standard deviations for all groups, as well as the *F* value, degrees of freedom, and *p* value.

- Text for this example→ Analysis of variance indicated that the different racial groups present in our sample report unequal levels of Sense of Community: $F(4, 119) = 6.04, p < .001$.

Regression

Caution: If you're planning to use a regression in your analysis, Josh will help you through the process. It's a lot more complicated than we're showing here.

Regression

- Regression involves using multiple known characteristics of a dataset to try and build a mathematical model which predicts an outcome variable (dependent) using predictor (independent) variables
- The resulting model is then tested against the real observations, and the degree to which the predicted values of the dependent variable matches the real observations is called the model fit
- The advantage of regression is that it allows us to measure the relative, independent contributions of multiple predictors to the outcome

Regression Types

- The type of regression one uses depends on the nature of the outcome variable:
  - **Ordinary Least Squares (OLS)**: when predicting a continuous, normally distributed outcome variable
  - **Logistic**: when predicting a dichotomous outcome
  - **Poisson/Negative Binomial**: when predicting a count variable
  - Many, many more

Regression: Predictor variables

- Continuous variables should be normally distributed
  - Significantly skewed variables need to be transformed

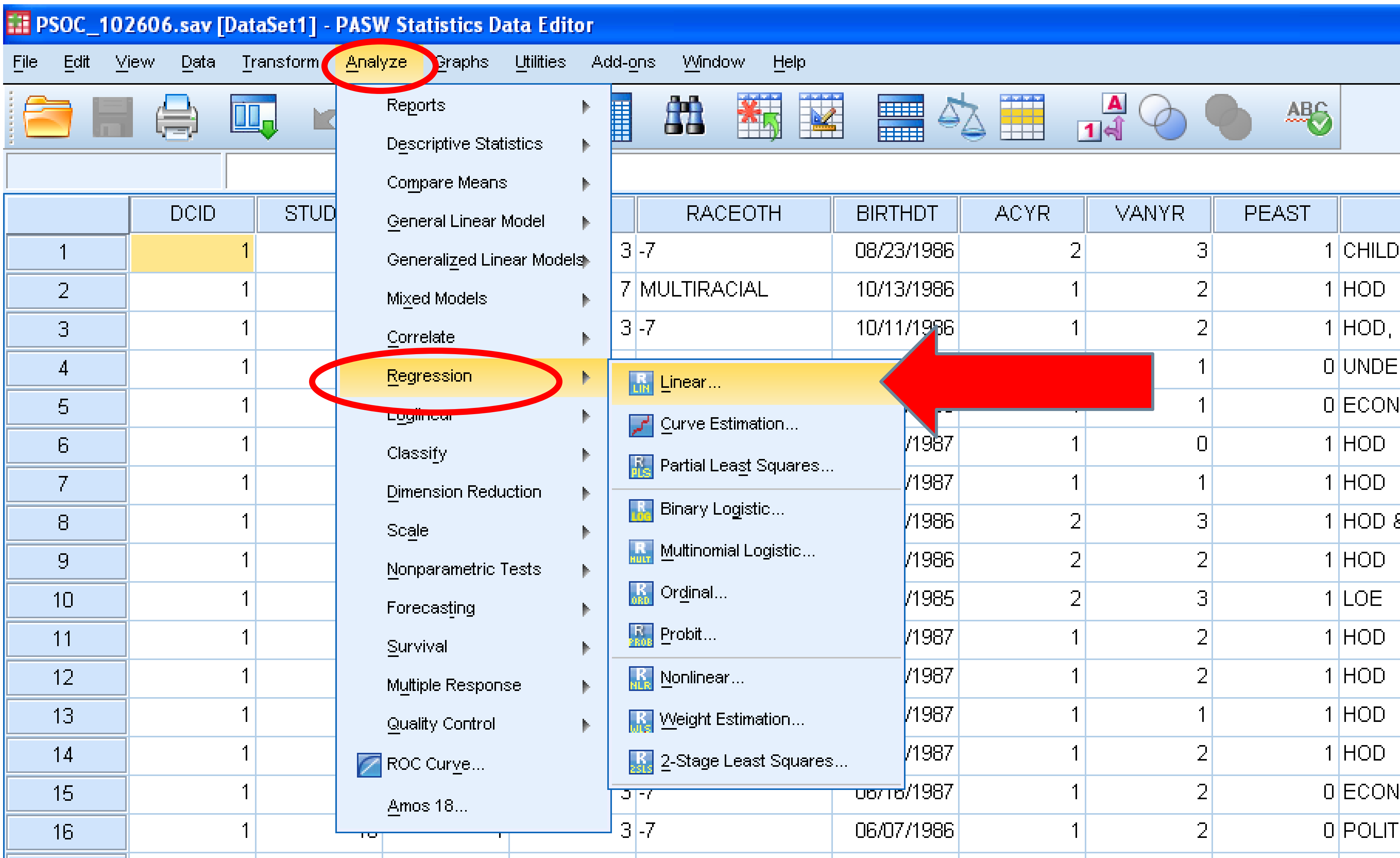
- Dichotomous categorical variables can be entered straight into the model, as long as it is coded zero and one; other codings need to be transformed

- Categorical variables with more than two values need to be transformed into a series of dummy variables

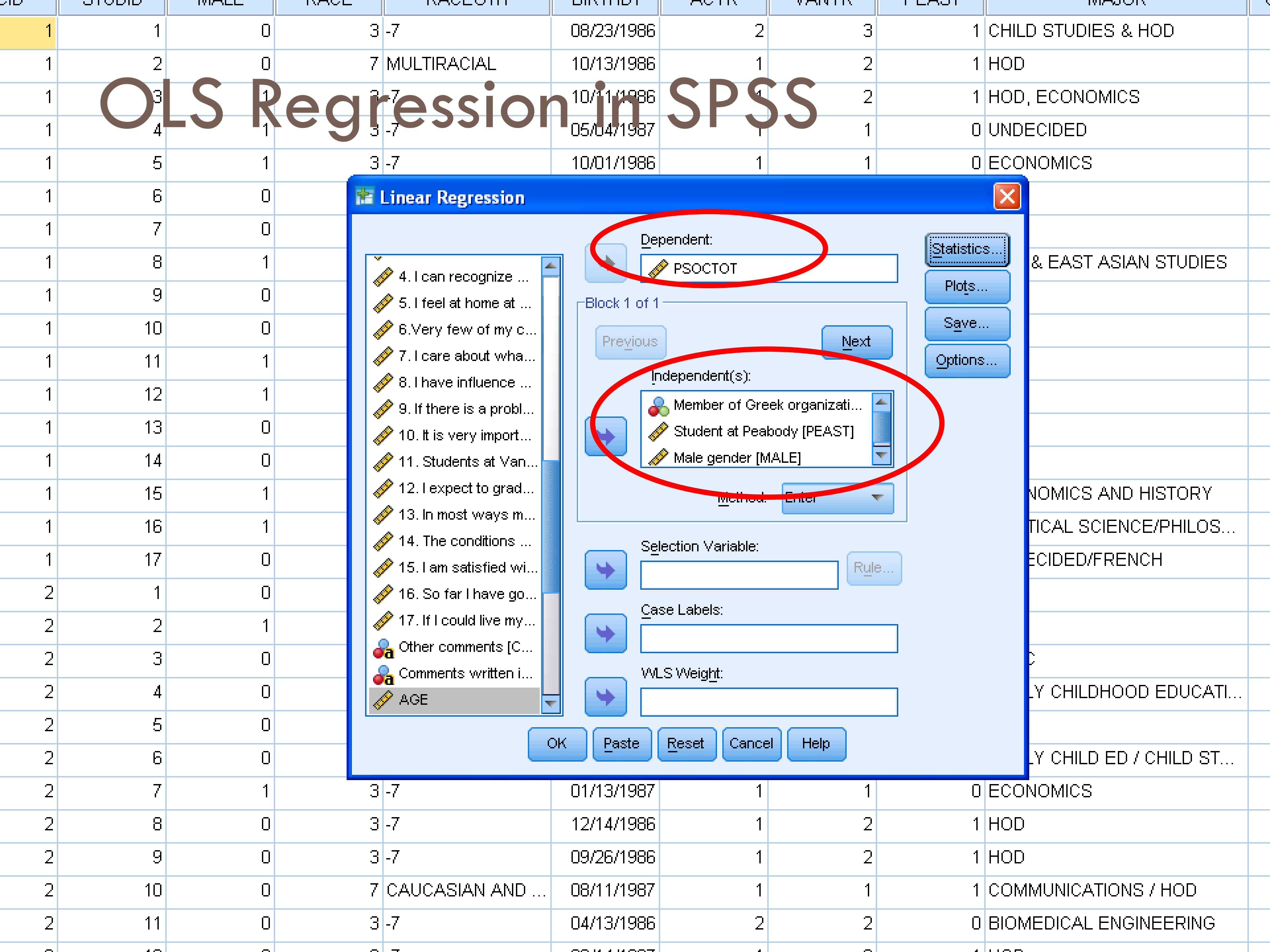
Regression assumptions

- You've chosen the right kind of regression to suit the dependent variable
- There is a more-or-less linear relationship between each predictor and the outcome variable
- None of the predictor variables are too highly correlated with each other
- Many, many more

OLS Regression in SPSS



OLS Regression in SPSS



Interpreting the output

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	3.614	.097		37.358	.000	3.423	3.806
	Member of Greek organization at Vanderbilt	.223	.072	.269	3.096	.002	.080	.365
	Student at Peabody	.053	.084	.058	.631	.529	-.113	.218
	Male gender	-.076	.075	-.093	-1.007	.316	-.224	.073

a. Dependent Variable: PSOCTOT

The constant or intercept represents the value of the outcome if all the predictor variables had zero values (which may or may not be meaningful).

The significance level indicates whether that variable is or is not a statistically significant predictor of the outcome. Again, we generally need to be 95% certain ($p < .05$) that we're not making an error when we declare something to be significant.

Interpreting the output

R-squared is the model fit index, which ranges from 0 to 1, with values closer to one indicating a better fit.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.300[a]	.090	.068	.38145

a. Predictors: (Constant), Male gender, Member of Greek organization at Vanderbilt, Student at Peabody

b. Dependent Variable: PSOCTOT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	3.614	.097		37.358	.000	3.423	3.806
	Member of Greek organization at Vanderbilt	.223	.072	.269	3.096	.002	.080	.365
	Student at Peabody	.053	.084	.058	.631	.529	-.113	.218
	Male gender	-.076	.075	-.093	-1.007	.316	-.224	.073

a. Dependent Variable: PSOCTOT

The unstandardized beta represents the coefficients of our regression equation which uses the independent variables to predict the dependent variable

The standardized beta represents the relative contribution of each predictor variable in influencing the outcome

Understanding what it's all about

- The results tell us that the best linear model for predicting PSOC is as follows:

PSOC = 3.61 + .22(GRK) + .05(PBDY)* - .09(MALE)*

- But the model fit is quite poor ($R^2 = .09$)

* Non-significant coefficients and variables can be deleted from the final model

Interpreting the results

- If all the variables were significant, a table like this is one way to make the results more meaningful:

Greek member	No		Yes	
Peabody student	No	Yes	No	Yes
Female	3.61	3.66	3.83	3.92
Male	3.52	3.57	3.74	3.79

Writing it up

- Tables are very appropriate when presenting regression results:
  - http://owl.english.purdue.edu/owl/resource/560/19/
- Don't include any equations
- Include at minimum the intercept, all unstandardized betas, an indication of whether each beta was significant, and the model fit

Table One
Student PSOC Regression Results

	B	SE(B)	β	B CI
Intercept	3.61**	.10		[3.42, 3.81]
Greek member	.22**	.07	.27	[.08, .37]
Peabody student	.05	.08	.06	[-.11, .22]
Male	-.08	.08	-.09	[-.22, .08]

*Significant at p < .05 **Significant at p < .01