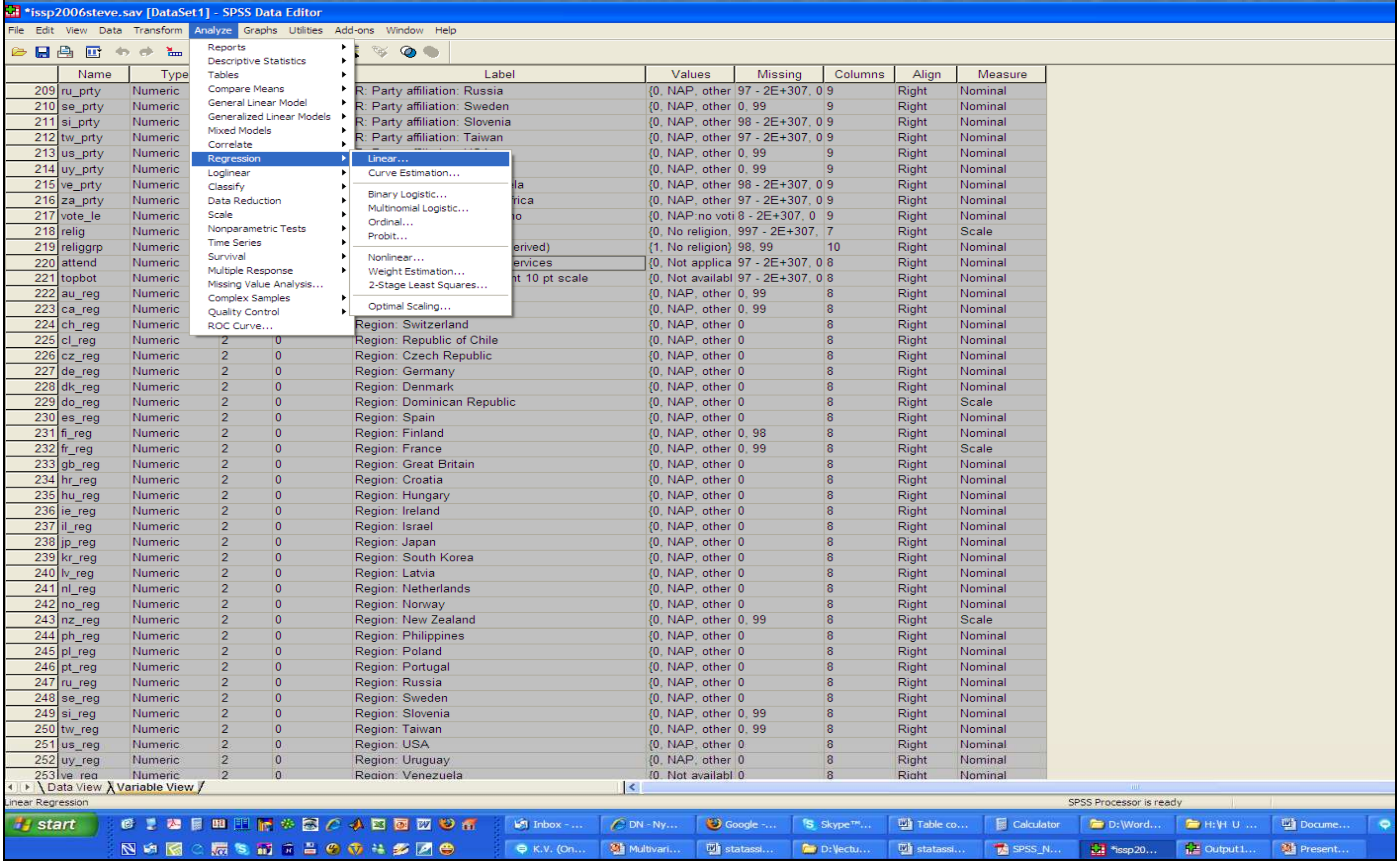
Multivariate Lab Exercises

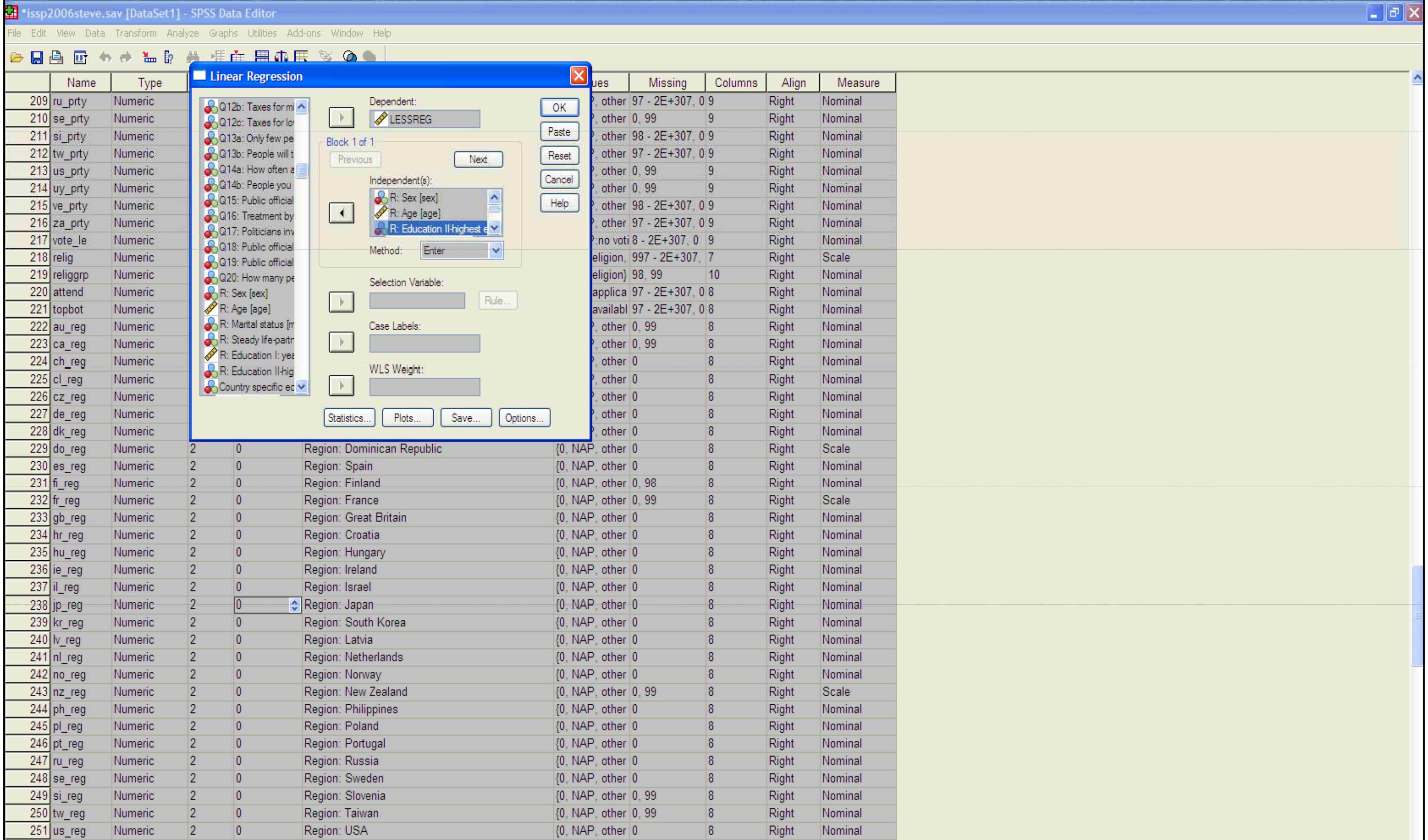
Getting Started

- Remember again to put on the filter
- Think about which independent variables you want to use
- For today we will only look at one dependent variable, so choose the one that you think does the best job in measuring the issue you are interested in.
- For example, if you are interested in gender equality, choose the one question that you think most clearly shows gender attitudes. If you are choosing market-liberalism, think whether less regulation really is the best question or should you choose a different question.

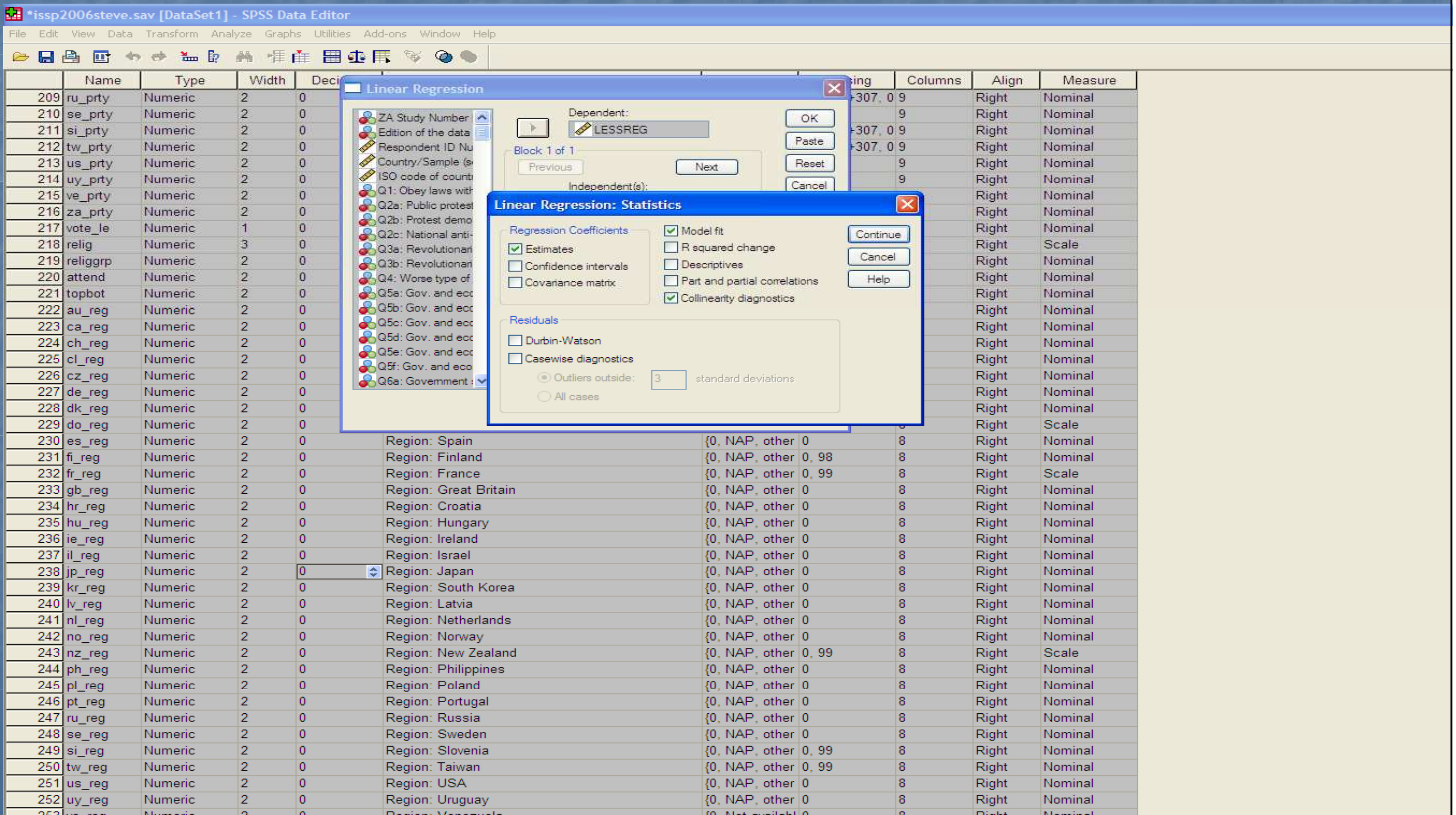
After adding the filter, go to linear regression



Choose your independent variables



Ask for collinearity diagnostics under "test" (you can also choose Durbin-Watson and casewise diagnostics if you want to test for heteroscedasticity, but consult the textbooks for advise)



The F-statistic shows that the model as a whole is significant, but this is almost ALWAYS the case, so it tells you when the model is bad, but not when it is good

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.125	3	3.042	3.053	.028[a]
	Residual	1112.752	1117	.996		
	Total	1121.877	1120			

a. Predictors: (Constant), R: Education II-highest education level, R: Sex, R: Age

b. Dependent Variable: LESSREG

Which variable(s) is/are significant? Which Variable explainces the greatest amount of variance?

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.639	.151		24.082	.000		
	R: Sex	-.090	.060	-.045	-1.491	.136	.997	1.003
	R: Age	.000	.002	-.005	-.164	.869	.990	1.010
	R: Education II-highest education level	.071	.027	.077	2.576	.010	.993	1.007

a. Dependent Variable: LESSREG

Which variable(s) is/are significant?

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.639	.151		24.082	.000		
	R: Sex	-.090	.060	-.045	-1.491	.136	.997	1.003
	R: Age	.000	.002	-.005	-.164	.869	.990	1.010
	R: Education II-highest education level	.071	.027	.077	2.576	.010	.993	1.007

a. Dependent Variable: LESSREG

Is there a problem with Collinearity?

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.639	.151		24.082	.000		
	R: Sex	-.090	.060	-.045	-1.491	.136	.997	1.003
	R: Age	.000	.002	-.005	-.164	.869	.990	1.010
	R: Education II-highest education level	.071	.027	.077	2.576	.010	.993	1.007

a. Dependent Variable: LESSREG

Nevertheless, SEX and AGE are rather equally distributed on dimensions 3 and 4 which indicates there could be a problem between them

Collinearity Diagnostics[a]

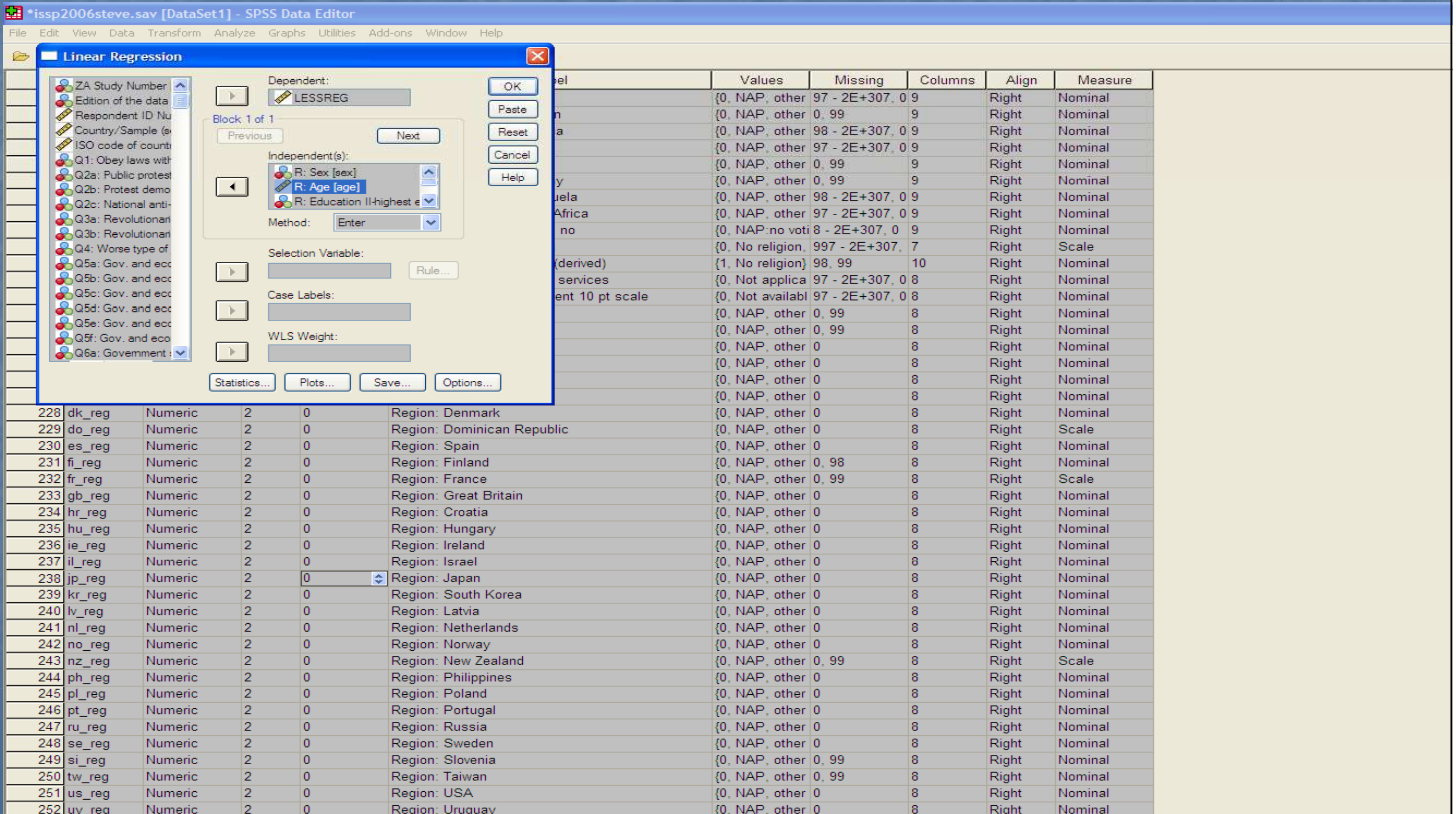
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	R: Sex	R: Age	R: Education II-highest education level
1	1	3.735	1.000	.00	.01	.01	.01
	2	.144	5.100	.00	.03	.21	.67
	3	.093	6.352	.00	.55	.46	.06
	4	.029	11.427	.99	.42	.32	.26

a. Dependent Variable: LESSREG

Next step

- Probably only educational level is a good predictor, so it would be best to use a bivariate regression or replace SEX and AGE with some other variables
- But since there might have been a problem of collinearity between AGE and SEX and since AGE had a much lower std. Coefficient, we will eliminate AGE now and see if SEX become significant

Elminate AGE



Sex is still insignificant

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	3.618	.124		29.265	.000		
	R: Sex	-.089	.060	-.044	-1.483	.138	1.000	1.000
	R: Education II-hig education level	.073	.027	.079	2.667	.008	1.000	1.000

a. Dependent Variable: LESSREG

Adjusted R-square

- Increased now from .005 to .007
- (Not shown here)
- So eliminating a variable made the model better
- But still it cannot explain even 1% of the change in LESSREG

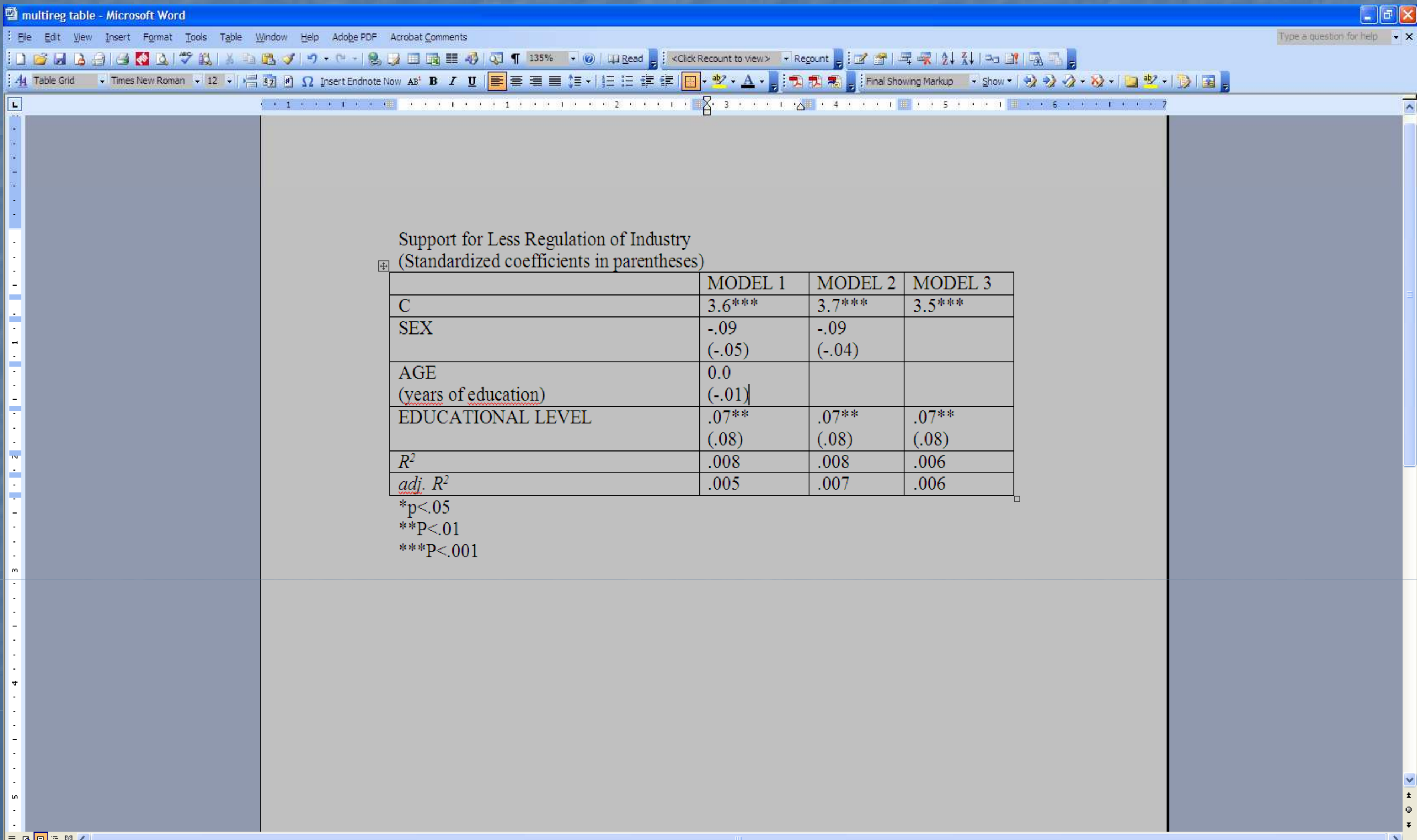
With only EDUCATIONAL LEVEL: it is significant, but the adjusted r-square is only .06, so again it shows we should look for different variables

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.475	.078		44.504	.000		
	R: Education II-highe education level	.073	.027	.080	2.689	.007	1.000	1.000

a. Dependent Variable: LESSREG

Make table in Word



Support for Less Regulation of Industry
(Standardized coefficients in parentheses)

	MODEL 1	MODEL 2	MODEL 3
C	3.6***	3.7***	3.5***
SEX	-.09 (-.05)	-.09 (-.04)	
AGE (years of education)	0.0 (-.01)		
EDUCATIONAL LEVEL	.07** (.08)	.07** (.08)	.07** (.08)
R^2	.008	.008	.006
adj. R^2	.005	.007	.006

*p<.05
**P<.01
***P<.001

Your Task

- Run a multiple regession with the dependent variable you have chosen and at least 5 independent variables
- Eliminate all the variables that are not significant
- Make sure there is no-collinearity
- Try to make the best model
- Make a table in Word