Data Analysis in SPSS

Syntax

For all analyses, save your syntax! It is best to choose Paste in the dialog boxes instead of Ok. Then, in the syntax window, select what you want to run and press Run (green triangle/arrow/play symbol).

Save output

This is a separate file from your data and has a different file type extension (.spv). The output is the log of the steps you have completed and the results you generated included tables or graphs you created.

You can also export the output to Word or a PDF from the file menu.

File > Export

In this Document:

Questions?

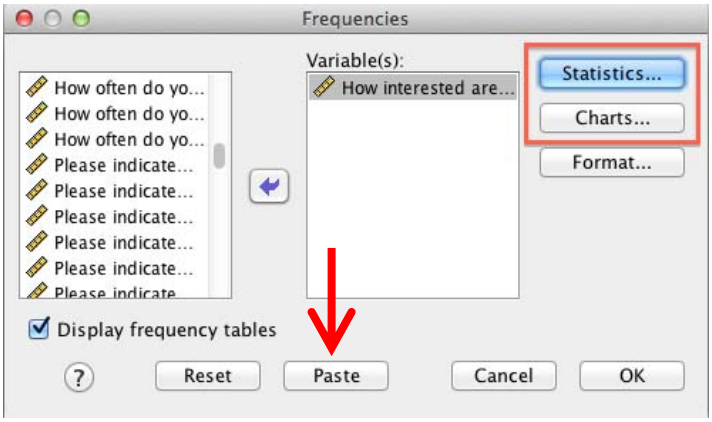
Please see *Using SPSS* (Green and Salkind) or *Statistics for People who Think They Hate Statistics* (Salkind, N) for further direction.

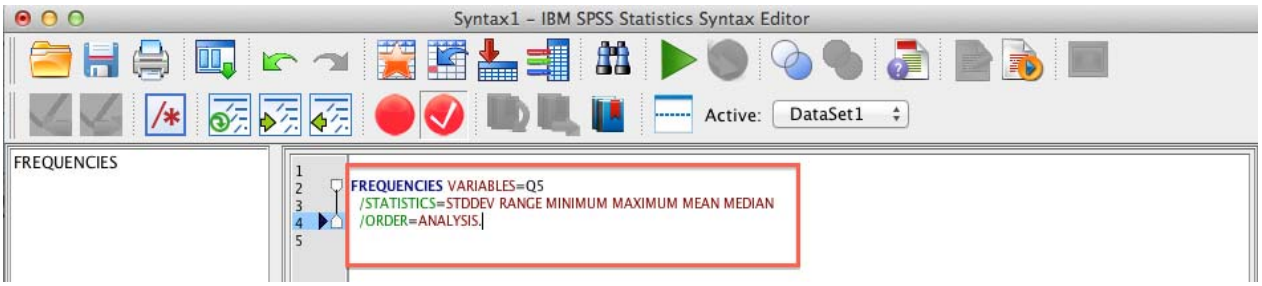
Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

Descriptive Statistics

Analyze > Descriptive Statistics > Frequencies

- Select one or more variables (use Ctrl or Shift to select multiple at once)
- Click the "**Statistics**" button
- Check the appropriate central tendency and distribution statistics for that type of variable:
  - **Nominal** – mode (absolute and relative frequencies are produced automatically)
  - **Ordinal** – quintile range
  - **Interval** – mean and standard deviation
- Click the "**Charts**" button *as desired*; choose an appropriate chart for the variable type
- Click **Paste**





When you ***run*** your descriptive analyses, you will be able to look at the statistics table for the measures of central tendency and dispersion as in the table on the left. The full frequency table, on the right, has the absolute "Frequency" (the count) and the relative frequency ("Valid Percent").

Statistics

How interested are you in politics?

N	Valid	136
	Missing	4
Mean		4.65
Median		5.00
Std. Deviation		1.736
Range		6
Minimum		1
Maximum		7

How interested are you in politics?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not at all interested	9	6.4	6.6	6.6
	Fairly uninterested	16	11.4	11.8	18.4
	Somewhat uninterested	7	5.0	5.1	23.5
	Neutral	11	7.9	8.1	31.6
	Somewhat interested	48	34.3	35.3	66.9
	Fairly interested	28	20.0	20.6	87.5
	Very interested	17	12.1	12.5	100.0
	Total	136	97.1	100.0	
Missing	System	4	2.9		
Total		140	100.0		

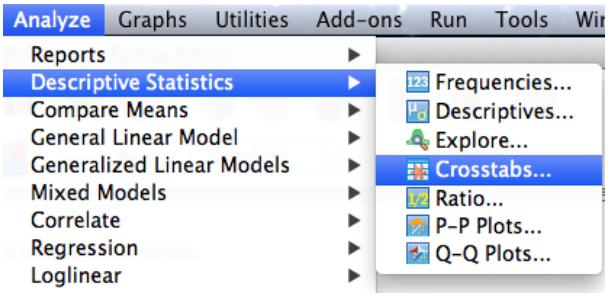
Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

Chi-Square

Compare groups (i.e., two nominal variables). Example: party affiliation & gender

Create Crosstabs

Analyze > Descriptive statistics > Crosstabs > Select variables



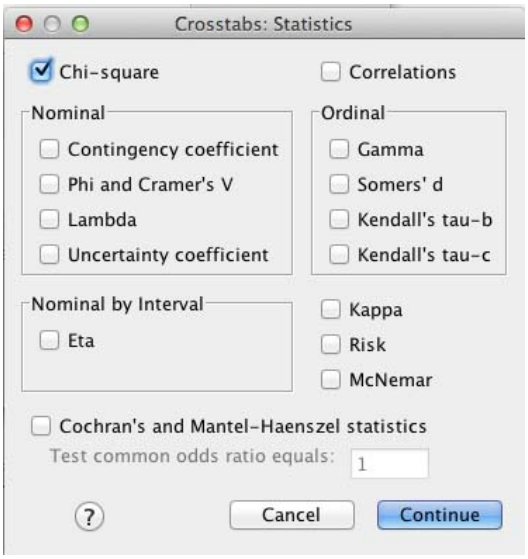
Rows should be the categories you want to compare (such as gender or party affiliation)

Columns should be the values of the variables you're comparing these groups of respondents across (such as opinion of the president)

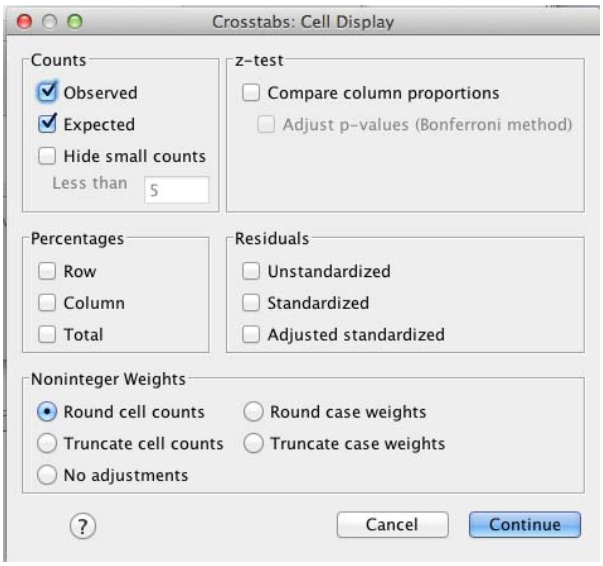
Click the **Statistics** button and check the **Chi-square** box

Click the **Cells** button. In the **Counts** section**,** make sure both the **observed** and **expected** boxes are selected**.** To better interpret relationships, find the P**ercentages** section and select **Row**. This will show the relative frequency within each category you want to compare.

Statistics…



Cells…



Time saving tip: You can request multiple cross tabs at once and SPSS will generate them all in separate tables

Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

Interpret Chi-Square output:

1. Examine the chi-square test to see if it is significant

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.855[a]	2	.054
Likelihood Ratio	5.886	2	.053
Linear-by-Linear Association	2.080	1	.149
N of Valid Cases	136		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.90.

2. *If* it is significant, look for the patterns to interpret

   a) Compare the *expected* count to *actual* count in each box.
   The more difference there is, the greater the relationship.

partyid * What is your gender? Crosstabulation

			What is your gender?		
			Male	Female	Total
partyid	Republican	Count	19	21	40
		Expected Count	17.9	22.1	40.0
	Independent	Count	19	12	31
		Expected Count	13.9	17.1	31.0
	Democrat	Count	23	42	65
		Expected Count	29.2	35.8	65.0
Total		Count	61	75	136
		Expected Count	61.0	75.0	136.0

   b) Compare the row percentages ("% within") in each column.
   Percentages in the same column would be equal if there were no relationship.

has_pet Has a Pet? * is_happy Happy? Crosstabulation

			is_happy Happy?		
			0 Not Happy	1 Happy	Total
has_pet Has a Pet?	0 No Pet	Count	404	2224	2628
		% within has_pet Has a Pet?	15.4%	84.6%	100.0%
	1 Has a Pet	Count	496	2981	3477
		% within has_pet Has a Pet?	14.3%	85.7%	100.0%
Total		Count	900	5205	6105
		% within has_pet Has a Pet?	14.7%	85.3%	100.0%

Independent Sample T-Test

Compare means of an interval dependent variable by a nominal independent variable with 2 levels/values. Example: comparing means in some opinion by party affiliation.

NOTE: If you want to compare means on two **variables** because you have one group, use a **Paired Sample T-test**. Enter each DV as the "paired comparison".

Independent samples is for when *different* people/participants are reporting on the *same* DV. In other words, men vs. women (not the same people).

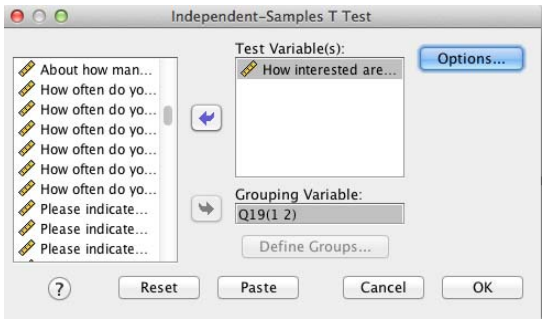
Analyze > Compare Means > Independent Sample T-test

The **Test Variable** is your DV. If you have multiple DVs with the same IV, you can run multiple t-tests at once, though beware of additive error!

The **Grouping Variable** is your IV, which groups you are interested in comparing.

After putting your IV in the box, click on **Define Groups…** to specify which groups to compare

- You can only include **2 groups** (if 3+ groups, you'd need to use ANOVA).
- Make sure to enter the values *as coded in your dataset* (e.g., usually 0/1 or 1/2).



Interpret T-Test output

1. Check whether **Levene's Test** was significant. If it ***is*** significant, skip to the *second* (bottom) row for the t-test because equal variances cannot be assumed. Otherwise, use the first (top) row.

2. Check whether the t-test was significant overall (where the red box is)

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
How interested are you in politics?	Equal variances assumed	1.146	.286	2.963	133	.004	.863	.291	.287	1.440
	Equal variances not assumed		1	2.961	126.215	2 .004	.863	.292	.286	1.440

3. If significant, compare the means for the two groups and interpret

Group Statistics

	What is your gender?	N	Mean	Std. Deviation	Std. Error Mean
How interested are you in politics?	Male	60	3 5.12	1.688	.218
	Female	75	4.25	1.677	.194

Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

ANOVA

Compare groups and variables (i.e., compare the means of an interval dependent variable by a nominal independent variable). Use ANOVA when you have ***any*** of the following:

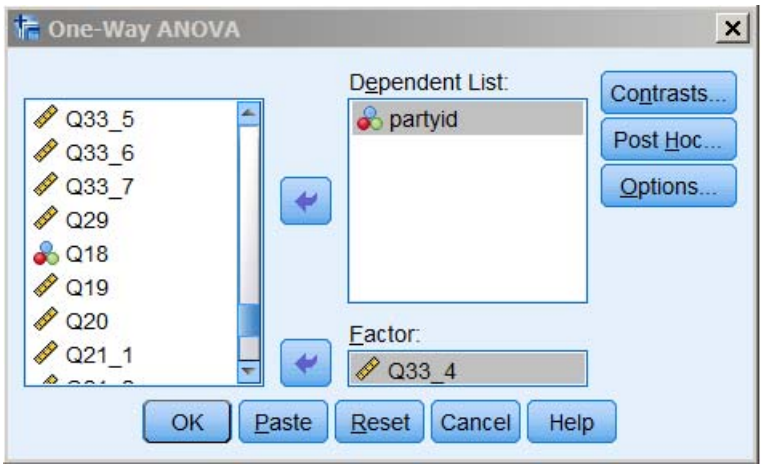
- 3+ categories/values in 1 IV (**One-Way ANOVA**)
- 2+ IVs and their interaction (**Two-Way ANOVA, Three-Way etc.,** or **Factorial ANOVA**)
- Control variables or Covariates (**ANCOVA**)

Go to Analyze > General Linear Model > Univariate

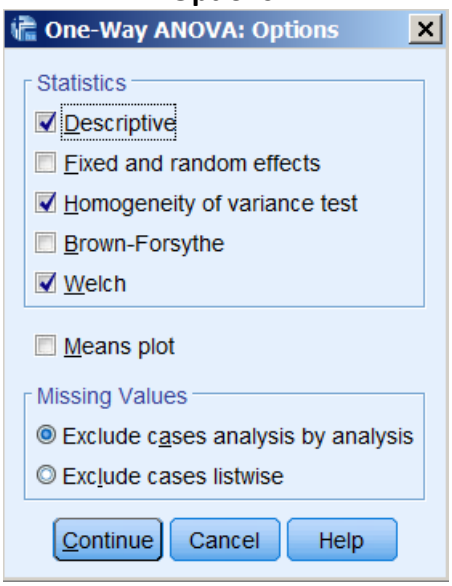
Analyze > Compare Means > One-Way ANOVA

- Place your numeric DV into the "**Dependent List**" box
- Place your nominal IV into the "**Factor**" box
- Click the "**Options…**" button to get **statistics** if required
- Click the **"Post Hoc…"** button for **post-hoc comparisons** if any variable has 3+ categories

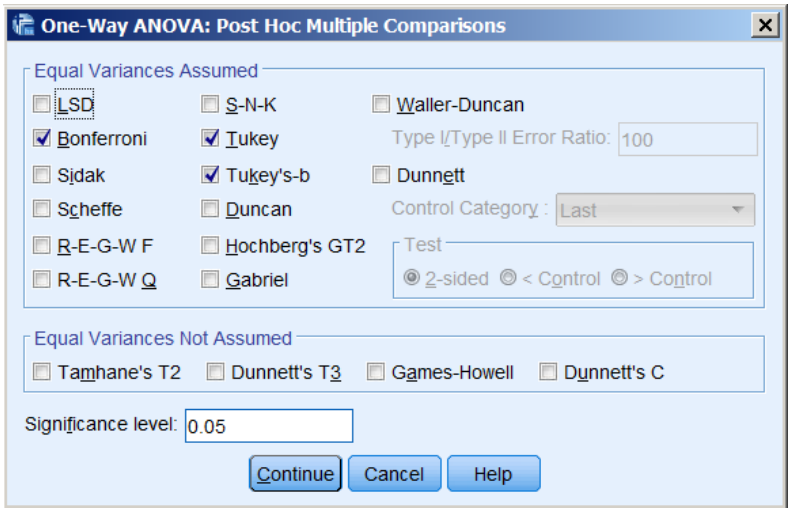
Main Window



Options…



Post Hoc…



Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

Interpret ANOVA Output

1. Examine whether the "omnibus test" (Corrected Model) is significant: is the model useful *overall*?

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.888	6	.315	.412	.870
Within Groups	98.517	129	.764		
Total	100.404	135			

2. If the effect of a variable with 3+ values is significant, examine the **post-hoc comparisons** to see which pairs of values are significantly different.

Pairwise Comparisons

Dependent Variable:Please indicate your level of agreement with each of the following statements:-There are too many political rants on Facebook.

(I) partyid	(J) partyid	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Republican	Independent	-.556	.382	.148	-1.311	.199
	Democrat	-.720*	.316	.024	-1.345	-.096
Independent	Republican	.556	.382	.148	-.199	1.311
	Democrat	-.164	.351	.641	-.858	.530
Democrat	Republican	.720*	.316	.024	.096	1.345
	Independent	.164	.351	.641	-.530	.858

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).
*. The mean difference is significant at the

3. Then, look at the **group means** to interpret those differences

Estimates

Dependent Variable:Please indicate your level of agreement with each of the following statements:-There are too many political rants on Facebook.

partyid	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Republican	3.878[a]	.249	3.385	4.370
Independent	4.434[a]	.288	3.864	5.003
Democrat	4.598[a]	.195	4.212	4.983

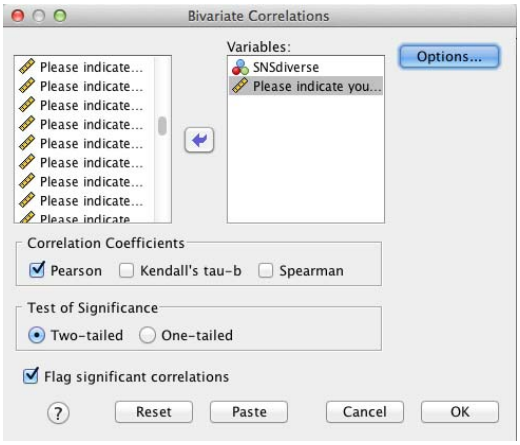
a. Covariates appearing in the model are evaluated at the following values: What is the highest level of education you've completed? = 3.42.

Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017

Pearson's Correlation

Compare the relationship between two interval-level variables.

Analyze > Correlate > Bivariate



Put your variables in the box. Include as many as you are interested in.
Note that it will only present the correlation for each *pair* (no controls or interactions).

Interpret Correlation Output

See if the relationship is significant (in red), then interpret the correlation coefficients (in yellow)

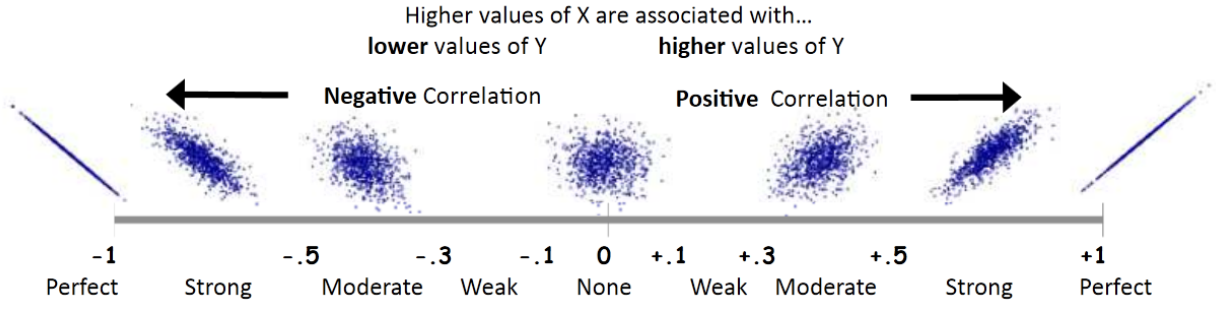
Correlations

		Q33_1	Q33_4
Q33_1 How often in the past three months have you:-Contacted or tried to contact an elected official	Pearson Correlation	1	.613**
	Sig. (2-tailed)		.000
	N	139	138
Q33_4 How often in the past three months have you:-Signed a petition	Pearson Correlation	.613**	1
	Sig. (2-tailed)	.000	
	N	138	138

**. Correlation is significant at the 0.01 level (2-tailed).

Also describe the *Direction* and *Magnitude* of the correlation in words, as described below:

- A **positive correlation** means that if we think the relationship is causal, as the value of variable X increases, we predict that the value of variable Y will also icrease. And, if the value of variable X decreases, we predict that the value variable Y also decreases.
- A **negative correlation** means we predict higher values of X are associated with lower values of Y, and lower values of X are associated with higher values of Y.



Originally by Emily Vraga, George Mason University
Revised by Debby Kermer, Library Data Services © 2017