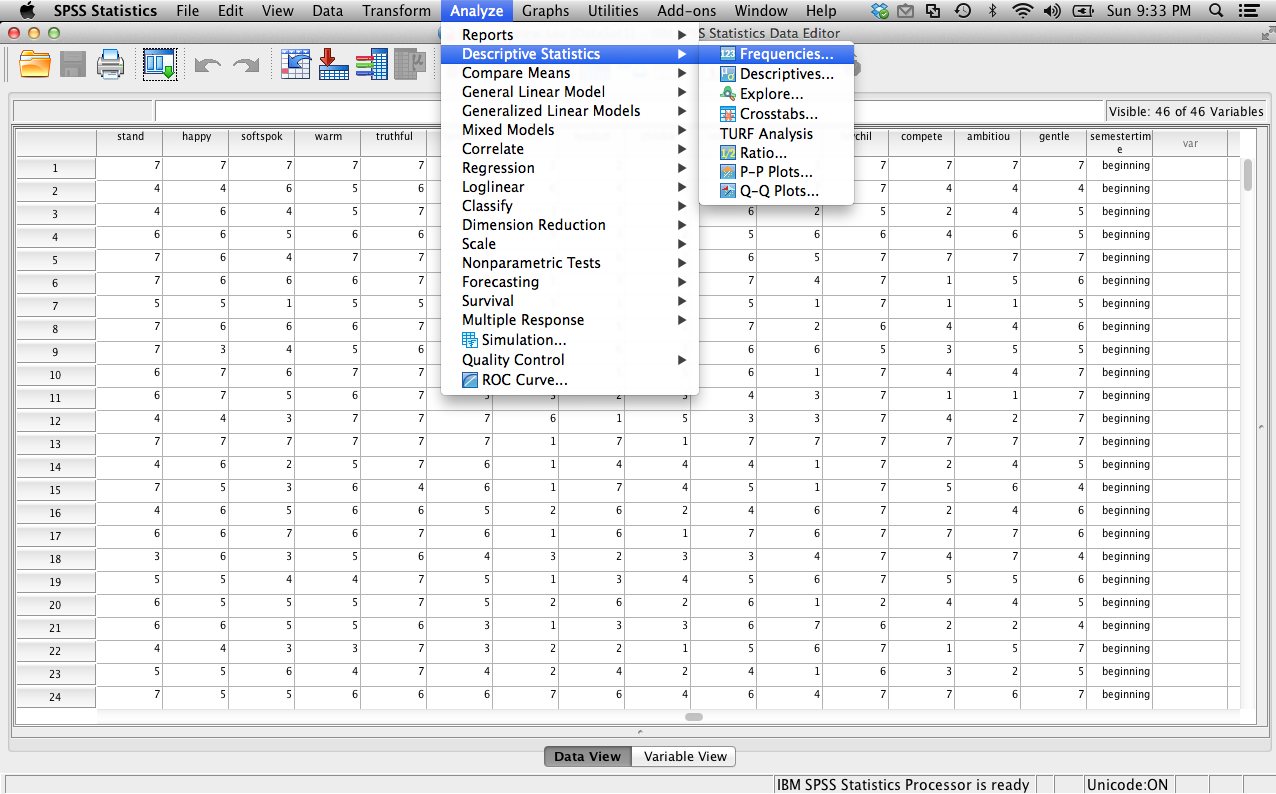
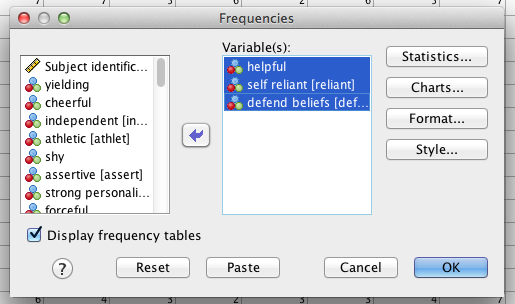
Introductory Statistics

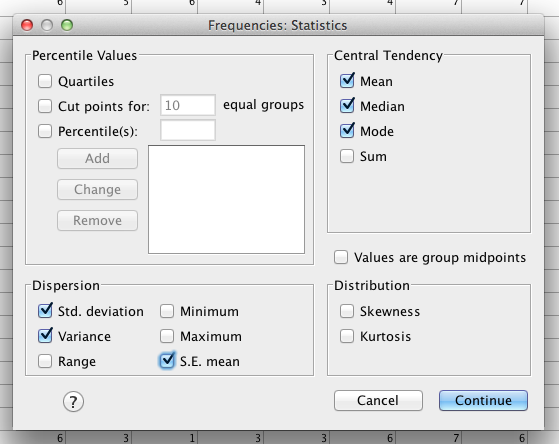
1. Types of Statistics:
   1. Descriptive Statistics: create a picture of the data – describe
      1. Mean – average of all scores
      2. Mode – score that appears the most often
      3. Median – score that appears in the middle when arranged in order
      4. Variance – average distance of scores from the mean
      5. Standard deviation – standardized variance (or standard average distance from the mean)
      6. Standard error – standardized standard deviation (the estimate of SD for the population = SD / square root (N)).
      7. Distribution drawings!
   2. Inferential statistics
      1. Infer information about the data.
      2. Tells you if your data is different from some known sample OR some other data set.
   3. Parametric versus non-parametric
      1. Parametric – used on interval and ratio data, numbers that are continuous in nature
         1. Requires more assumptions
         2. Most of the stats we will use (ANOVA, regression)
      2. Non-parametric – used on all data types (especially nominal, categorical)
         1. Does not require same assumptions
         2. Chi-square is most common.
2. Descriptives How-To:
   1. Analyze > Frequencies
   2. Analyze > Descriptives



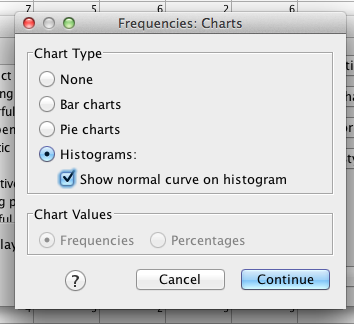
* 1. Frequencies:
     1. Move over the variables you want to get descriptive information from.
     2. Hit the Statistics button on the right side.



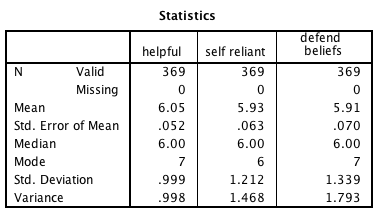
* + 1. Select the options you want (mean, mode, standard deviation, etc.)



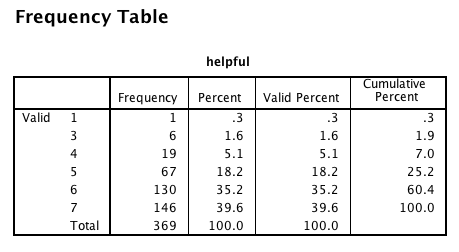
* + 1. Hit continue. You will be back at the options for frequencies.
    2. Hit the charts button.
       1. You will be able to select Histograms (you will use this information a lot). I usually also select view normal curve – it helps you see how skewed the data can be.



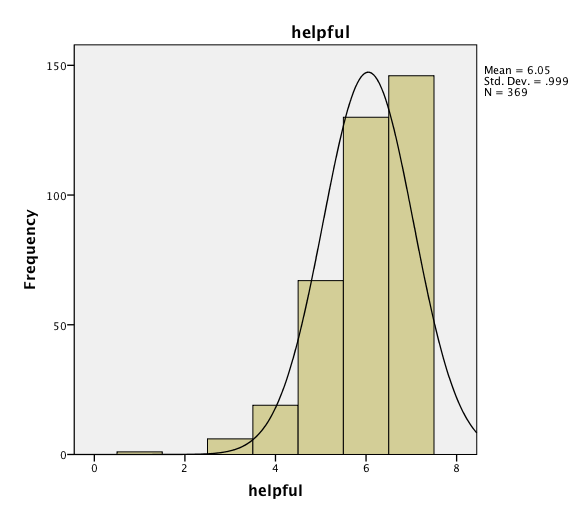
* 1. Frequency Output example:



This box gives you information on the number of subjects (N) in the study (valid just means how many lines there were in the dataset it could use). Then the missing line tells you how many missing data points you have from that specific variable (variables are the columns). Lastly, you will get the descriptive statistics you asked for (Mean, SE, Median, Mode, SD, and SD2 in this example).

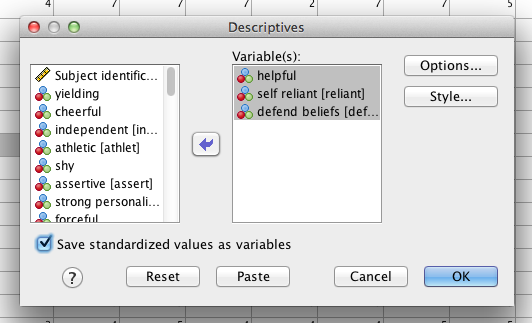


This box is a frequency table. Frequency tables have VALUES in the first column (what the scores could be), frequencies of those values in the second column, percentages in the third column, valid percent (important if there is missing data), and cumulative percent (adds up as you go down). These boxes are useful to look for incorrectly coded data, to see where most of the scores are, etc. They are not very helpful if you have data with true decimals (too big!).

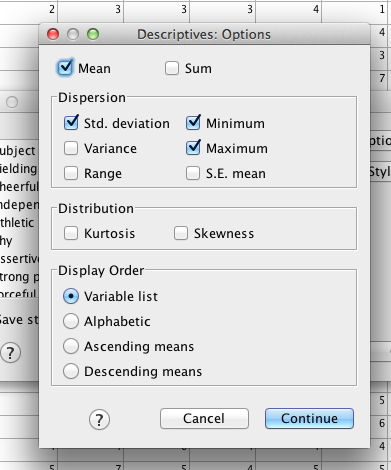


Histograms are frequency tables in bar chart form (and are infinitely more useful). You can now see where most of the data falls (4-6 in this example), skew, and kurtosis.

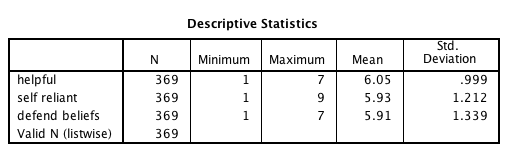
* 1. Histograms terms:
     1. Distribution descriptors: Unimodal, bimodal, multimodal, rectangular – describes the number of “humps” the distribution has.
     2. Normal distributions – centered at the mean, unimodal, symmetric
     3. Skew – how “not normal” a distribution is (how much it leans).
        1. Negative skew: scores are on the top (right side) of the distribution, tail is on the left side, sometimes called a ceiling effect.
        2. Positive skew: scores are on the bottom (left side) of the distribution, tail is on the right side, sometimes called a floor effect.
     4. Kurtosis – how different from normal in shape a distribution is (skinnier, flatter).
  2. Descriptives
     1. If you do analyze > descriptive statistics > descriptives you have less options for output, but it is quicker because it’s sort of pre-set to the normal things people want.
     2. Again, move over the variables you want to get descriptives from.



* + 1. Hit options to get different types of descriptives.
    2. Pick your favorite, hit continue and ok.



* 1. Descriptives output (very basics):
     1. Does not do median, mode
     2. Will give you standardized scores (z-scores)

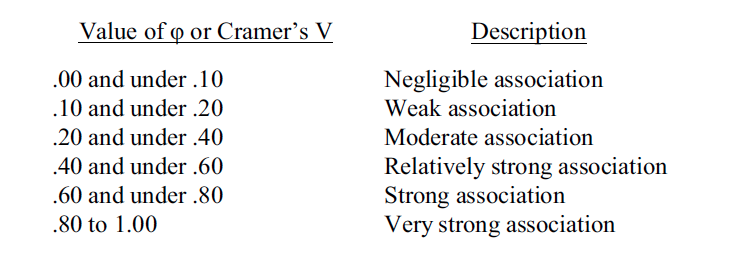


In the descriptives output, you get variables as rows and statistics as columns (hello backwards from before!). You will need N (number of valid rows/participants), and then the rest of the numbers you asked for (here is what comes without selecting any other options). Missing data is harder to see in this option (which is why I vote for the other!).

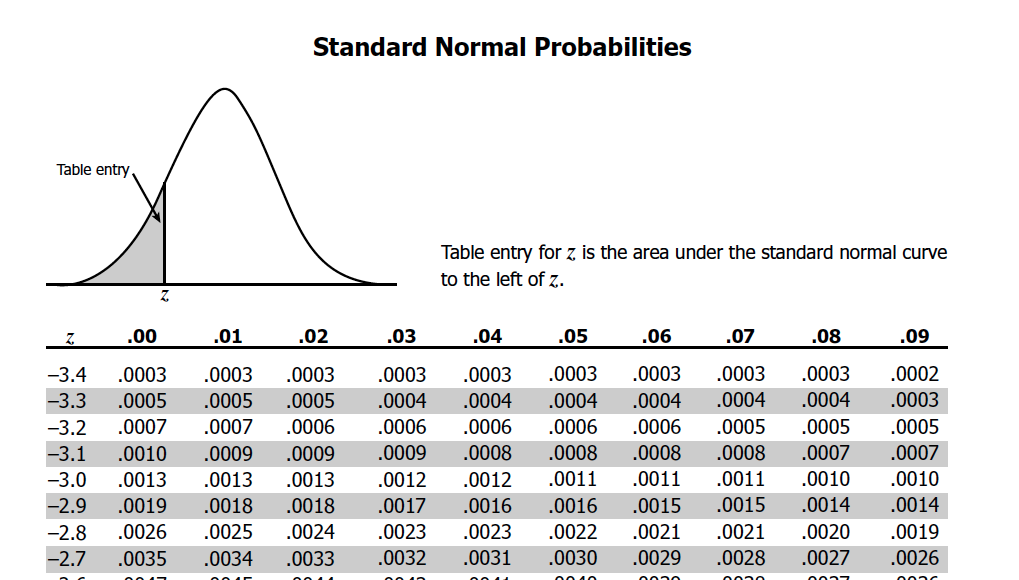
Go back to the data, and if you selected “save standardized values as variables”, you will see new columns at the END (far right) of your data set. These are the standardized scores of your data (z-scores described below), which are useful for talking about univariate outliers.



1. Hypothesis Testing (inferential statistics)
   1. Basic gist: You are pitting two rival answers against each other. Think of it like your favorite sport. You have a Favorite team (Research hypothesis) versus your enemy team (Null hypothesis). You want them to be different. You want your team to win (reject the null!).
   2. Terms:
      1. Null hypothesis: You expect to find: 1) no difference between group scores (t-tests, ANOVA families), 2) no relationship between variables (regression families. Chi-square).
      2. Research/alternative hypothesis: You expect to find: 1) differences in group scores, 2) relationships between variables.
      3. Hypothesis testing = rock’em sock’em robots of statistics.
   3. Why is this called *Null Hypothesis Significance Testing (NHST)?*
      1. How do you determine if you were correct (aka things were significant?)
         1. *p*-values (most common)
         2. Cut off scores (tables online)
         3. Fit indices (EFA)
      2. Rejecting the null / statistically significant – when your team wins! You find that the probability of the null hypothesis is very low, so you reject the idea that everything is equal (or that your team would never win).
      3. Retaining the null / not statistically significant – the probability of the null hypothesis is not low enough, it could be that the groups are equal, or that your team might not win.
   4. Terrible terminology:
      1. P-values – the probability of getting that results (t-value, f-value, chi-square, etc.) if the NULL were true
         1. You want your team to win! So you want the null to be false. Therefore, you want the probability of being wrong to be very low.
      2. Alpha – the probability of a Type 1 error (please note that alpha / = pvalue … you SET alpha as a criterion for a low type 1 error, which generally is *p*<.05, or *p*<.01, but it’s not the same thing as the *p*-actual found in your experiment).
         1. Type 1 error – rejecting the null hypothesis when it is FALSE.
            1. Memory mnemonic: First mistake = worst mistake. Saying something happened when it did not.
      3. Beta – the probability of a Type 2 error, the opposite of power.
         1. Type 2 error – failing reject the null hypothesis when you should reject the null (aka your research hypothesis is supported but you missed it…bummer).
      4. Power – the probability of rejecting the null when you should reject the null (aka your research hypothesis is supported and you showed that … yeah!).
         1. G\*Power is fantastical!
         2. Power is normally used for sample size calculation, to determine how many participants you need to find statistical significance, given a set effect size and analysis type.
   5. Other related issues:
      1. One tail test – you only want one direction (higher or lower) – note that a lot of multivariate statistics will not let you make this assumption.
      2. Two tail test – you are not sure if it is higher or lower, so you hedge your bets and look for both of them.
      3. Assumptions: Things that must be true for your test to return an answer that is reasonably correct
         1. GIGO – if you feed garbage data into SPSS, then you will get a garbage answer out.
         2. Therefore, when the assumptions are not met, you do not know what the answer you got actually *means.*
   6. Cut off Scores:
      1. Usually you learn about cut off scores and the score has to be greater than the cut off score to be significant
      2. You are finding the point in which the probability of that score is less than 5% or 1%.
      3. Now we are going to use the SIG column or the precise p-value (it’s much easier!).
         1. SPSS will give you the p-value. You want your p-values to be less than .05 or .01.
      4. Eliminates the need for cut off scores (*sort of*)
         1. Not for z-tests
         2. Not for post hoc tests
         3. Not for one tailed tests
         4. Always uses a two-tailed test (if applicable).
2. Effect size
   1. Effect size is a measure of “how big” an effect was in your experiment. For example, you might reject the null hypothesis (yay the experiment worked!), but then determine that the group differences or relationships were small.
   2. Effect size is considered “a measure of strength of a phenomenon” for a technical definition.
   3. Common effect sizes:
      1. Those based on mean differences (Used for: any time you have two means: *t*-tests, ANOVA post hoc tests)
         1. Cohen’s *d* - Cohen’s *d* is one of the most well-known calculations for effect size. The general formula gives the standardized distance between the two population means, or how much the two populations do not overlap. The formula for *d* is very adaptive and can be used for many different between samples and within samples tests.
         2. Hedge’s *g –* As Cohen’s *d* tends to be positively biased, Hedge’s *g* corrects for this to create an unbiased calculation.
         3. Glass’s delta – Glass’ delta is a form of effect size found by dividing the difference of the two groups by the standard deviation of only the control group.
         4. Sizes:
            1. Small .2
            2. Medium .5
            3. Large .8
            4. Can get very big or be negative.
   4. Those based on variance overlap (Used for: ANOVA overalls, regression)
      1. η 2 (eta squared) and R2 – These statistics are based on the amount of variance that you have accounted for by your manipulation (groups) or independent variable (like predictors in regression) out of the total variance.
      2. ω2 – omega squared is an estimate of the population effect size for eta and r squared (so it’s usually smaller than the other two), and accounts for sample size and the amount of error variance in your study in a different way that eta and r squared.
      3. Eta squared is the most common for ANOVA, R2 is more common for regression.
      4. There are also partial versions of all three of these statistics for when you have more than one IV … that means that you can calculate the effect size of each piece separately, rather than the experiment as a whole (useful to know which variable was the “best”).
      5. Sizes (the rules for this are not as set in stone as the *d*):
         1. Small .01
         2. Medium .09
         3. Large .25
         4. Since this statistic is the proportion of variance over a total, it ranges from 0 to 1 and cannot be negative.
   5. For categorical variables:
      1. Odds-ratios – gives you the odds of one group membership over another.
      2. φ and Cramer’s V – chi-square statistic (for independence tests only, see below) that is loosely based on Cohen’s *d*.
      3. Sizes:



1. Univariate versus Multivariate
   1. Univariate – 1+ IVs to 1 DV
      1. Most common used
      2. Limit you to only one DV
      3. Types
         1. *Z*
         2. *t-tests*
         3. *ANOVA/ANCOVA*
         4. *Correlation*
         5. *Regression*
         6. *Chi-Square (non-parametric)*
   2. Multivariate – 1+ IVs to 1+ DVs
      1. Traditionally multivariate means “more than one DV”, but the term is loosely applied to “all things not vanilla ANOVA/regression”.
      2. Types
         1. *MANOVA/MANCOVA*
         2. *Profile Analysis (repeated measures)*
         3. *Discriminant Analysis (discriminate?)*
         4. *Log Regression*
         5. *Factor Analysis*
         6. *Canonical Correlations*
         7. *Frequency Analysis*
         8. *And many more…*
2. Z Walk Through
   1. Z-Score
      1. When: one person, population mean, and population standard deviation are known (previous research … like standardized tests)
      2. Assumptions: Normal Distribution
      3. Formula: (X – μ) / σ
         1. Person’s score – average score divided by the standard deviation
      4. Example: A personnel psychologist has to decide which of three employees to place in a particular job that requires a high level of coordination. All three employees have taken tests of coordination, but each took a different test. Employee A scored 15 on a test with a mean of 10 and a standard deviation of 2; Employee B scored 350 on a test with a mean of 300 and a standard deviation of 40; and Employee C scored 108 on a test with a mean of 100 and a standard deviation of 16. (On all three tests, higher scores mean greater coordination.)
      5. Who’s the best?
         1. First calculate Z.
         2. Then look to see if that Z score:
            1. Is less than a cut off score.
            2. P value is less than your alpha (remember most of the time we use *p*<.05).
      6. Z-tables! Look online at the table.



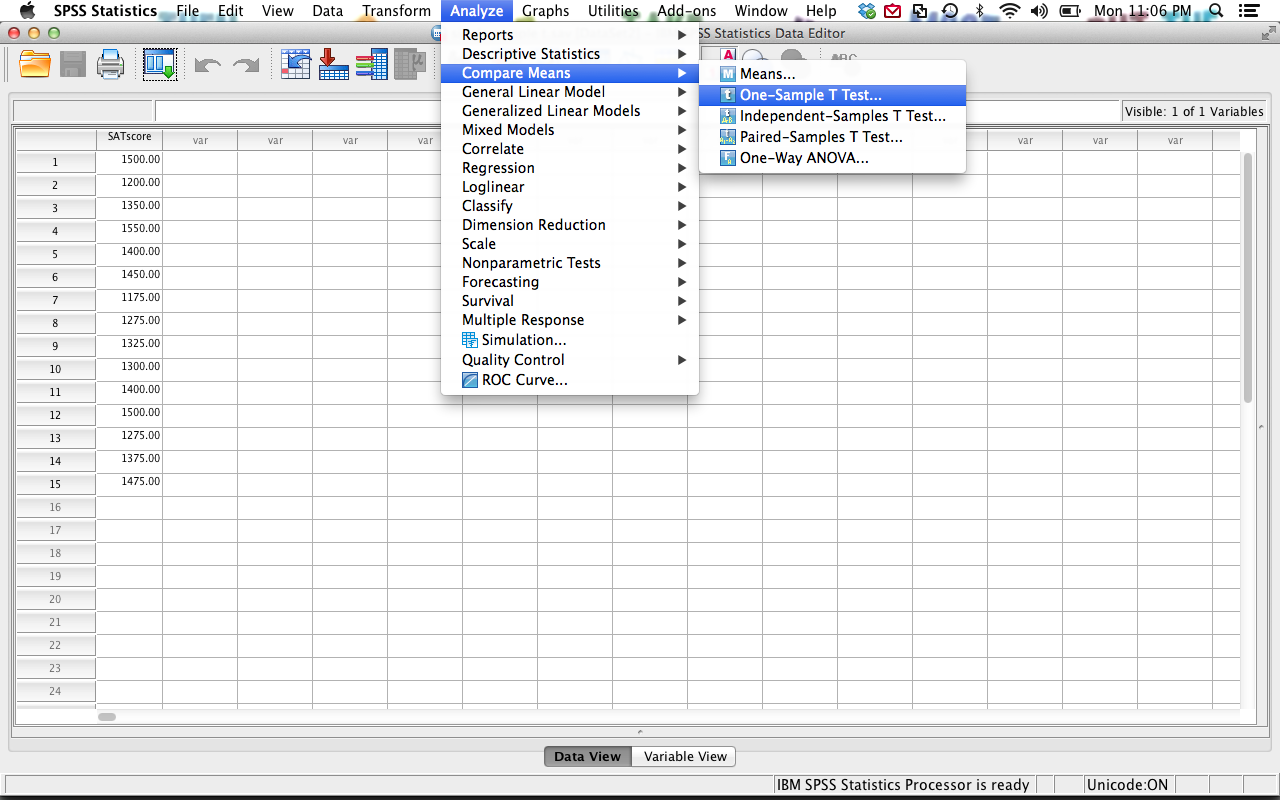
What’s going on? In the first column, you have the found z-score to one decimal place. (Anything above 3.4 is approaching zero … so very very small values). Then across the top, you have the second decimal for the found z value. So for a Z score = -2.75, the entry is 0.0030. That number is the *p* value to the LEFT of the z-score. That’s the probability of getting that score OR LOWER for one tailed test (want two-tails? *p* X 2). Remember that normal distributions probability = 1 for the whole distribution, so to the RIGHT = 1 – *p* (sometimes listed as *q*).

Handy cut off scores for hypothesis testing.

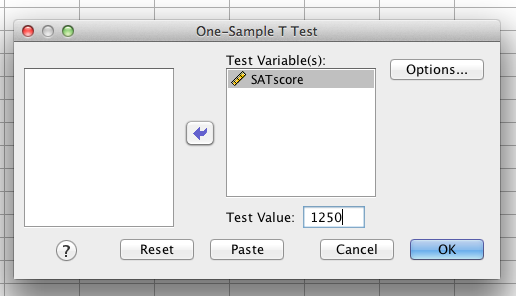
|  |  |  |
| --- | --- | --- |
|  | *p*<.05 | *p*<.01 |
| One tailed | 1.64 for right tail  -1.64 for left tail | 2.33 for right tail  -2.33 for left tail |
| Two tailed | 1.96 or -1.96 | 2.58 or -2.58 |

* 1. Z-Tests:
     1. When: one group of people, population mean, population standard deviation are known
     2. Assumptions: Normal distribution
     3. Formula: (M – μ) / σm
        1. Sample mean – population mean divided by population standard error
           1. Standard error = standard deviation (σ) / √ N
     4. Example: In a study to see if children from lower socio-economic status (SES) neighborhoods have lower than average test-taking skills, a psychologist administered a standard measure of test-taking skills to a set of seven randomly chosen children from a low SES neighborhood and found them to have a score of 38. The average score on this measure for the population in general is 50 with a standard deviation of 10. Using the .05 level of significance, what conclusions should be drawn about whether children from low SES neighborhoods have lower test-taking ability?
        1. First find standard error.
        2. Then find Z.
        3. Then look to see if that Z score:
           1. Is less than a cut off score.
           2. P value is less than your alpha (remember most of the time we use *p*<.05).
     5. Write up: *Z* = 2.56, *p*<.05, *d* = 1.25
     6. Effect size: see MOTE user guide.

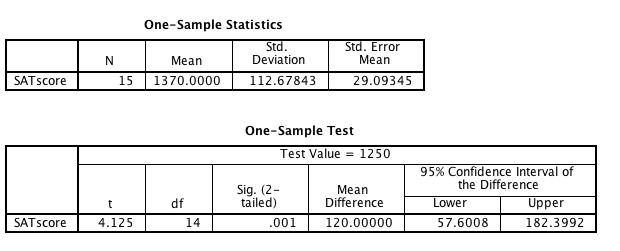
1. T-Tests
   1. Types
      1. Single sample – one group of people, a population mean, NO population standard deviation.
      2. Dependent – one group of people tested twice!
      3. Independent – two groups of people.
   2. Assumptions:
      1. Normal Curves
      2. Homogeneity – equal variances for each group
      3. Linearity – variables are linearly related (bit of an odd one for categorical groups).
   3. Formulas
      1. While we are going to talk a lot about formulas, since SPSS will do these for you, one thing to know is that these formulas are basically the same idea as Z tests.
      2. Z and t tests are all versions of the following:
         1. Some Mean You Care About – Some Other Mean You Care About / Standard Error (estimated from standard deviation).
         2. Think about how that relates to the null hypothesis.
   4. Single sample example
      1. Uses: when you have one group of people to compare to a population mean.
      2. A school has a gifted/honors program that they claim is significantly better than others in the country. The national average for gifted programs is a SAT score of 1250.
      3. Use the file single sample t-test here.
      4. Analyze > Compare means > one-sample t-test



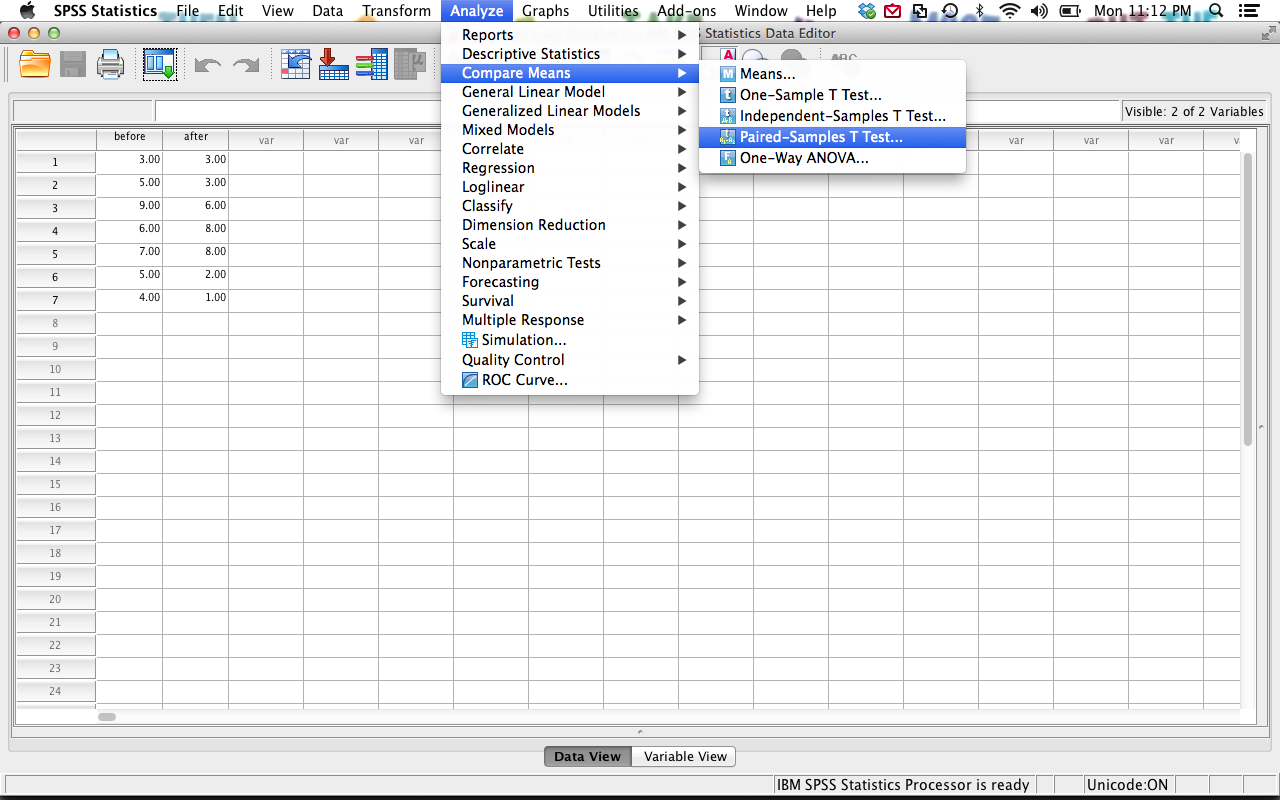
* + 1. Move over the variable you want to test.
    2. Be sure to enter your population mean in the test-value spot and hit ok.



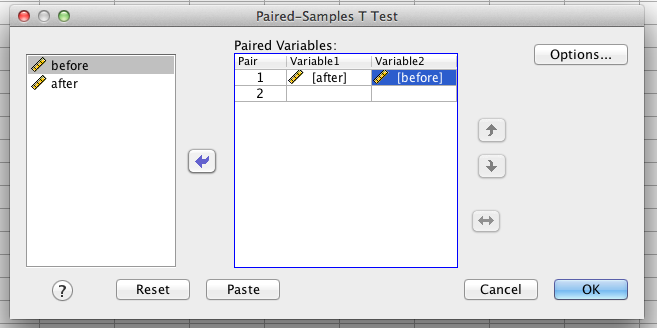
* + 1. Output example:



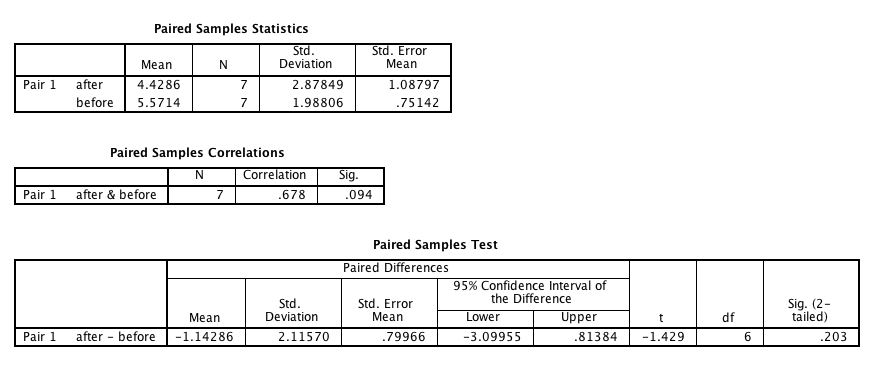
* + - 1. N = number of people
      2. Mean = group mean
      3. SD = standard deviation of group
      4. SE = standard error.
      5. One sample box: t – found t-value (called the *test statistic*).
      6. Df = degrees of freedom (OMG WHAT ARE THESE?!)
      7. Sig = p-value (want this to be less than .05)
      8. Write up example:
         1. *M* = 1370.00, *SD*  = 112.68, *t*(14) = 4.125, *p* = .001, *d* = 1.06
    1. Effect size: see MOTE user guide.
  1. Dependent t-test
     1. Use: when you have one group of people tested twice, before/after scores, etc.
     2. Example: In a study to test the effects of science fiction movies on people's belief in the supernatural, seven people completed a measure of belief in the supernatural before and after watching a popular science fiction movie. Participants' scores are listed below with high scores indicating high levels of belief. Carry out a t test for dependent means to test the experimenter's assumption that the participants would be less likely to believe in the supernatural after watching the movie.
     3. Use dependent t example here.
     4. Analyze > compare means > paired samples t-test



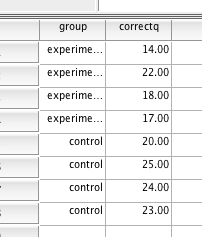
* + 1. Move over the variables you wish to test into variable 1 and variable 2. You can test lots of combinations at once. I suggest doing TIME 2 – TIME 1 if that is how your data is set up.
       1. Why?



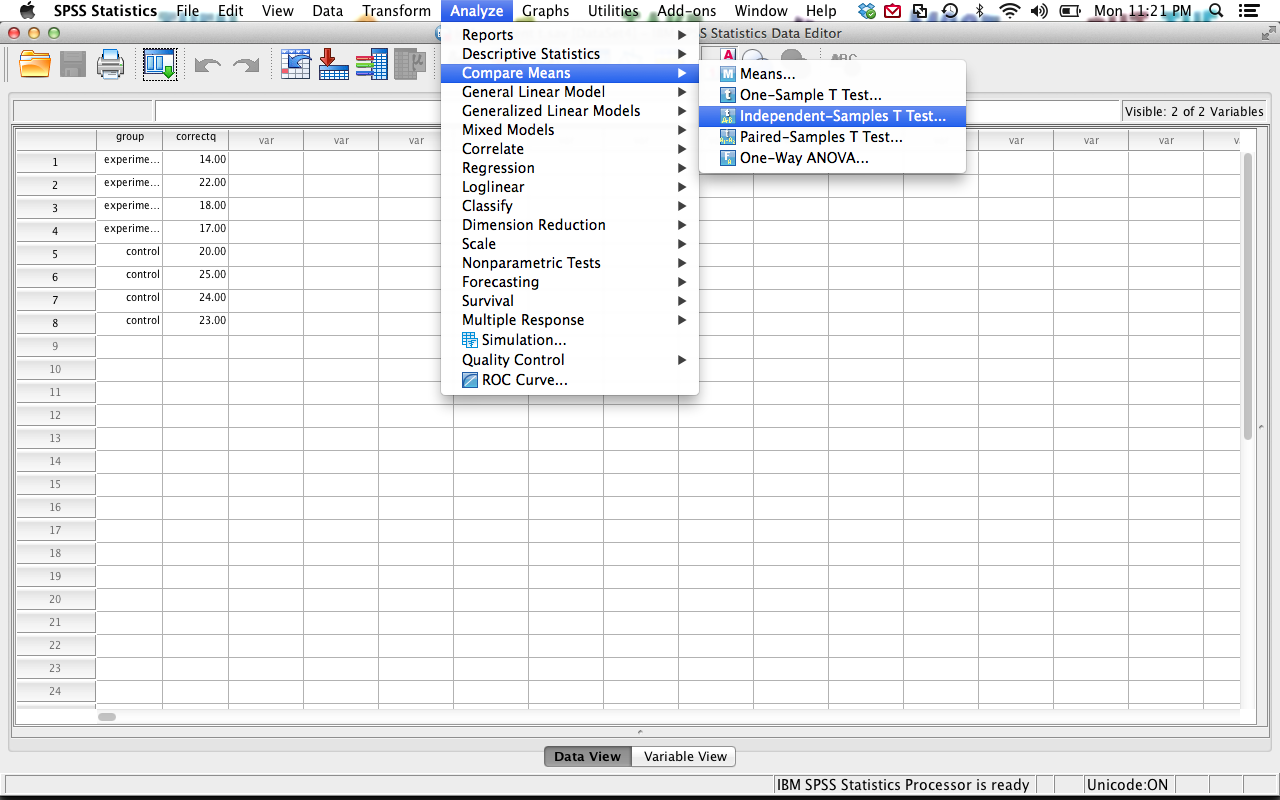
* + 1. Output:



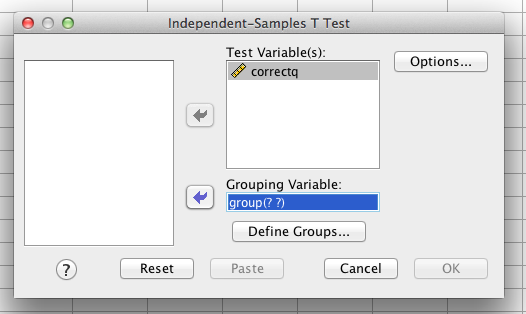
* + - 1. First box tells you the means, standard deviations, and standard errors for each time measurement.
      2. The second box tells you the correlation between the two time measurements.
      3. Unless you have a specific question about both of those things, most people just look at box three.
      4. The mean is the mean difference between time measurements.
      5. Standard deviation and standard error of the difference between time measurements.
      6. Confidence interval of the mean difference between time measurements.
      7. T-values, with degrees of freedom and p-value.
      8. How to write: *Mdiff* = -1.14, *SDdiff* = 2.12, *t*(6) = 1.429, *p* = .203, *d* = .54
    1. Effect size: see MOTE user guide.
  1. Independent t-test
     1. Use: Two groups (only two, no more) of completely separate people.
     2. Example: A forensic psychologist conducted a study to examine whether being hypnotized during recall affects how well a witness can remember facts about an event. Eight participants watched a short film of a mock robbery, after which each participant was questioned about what he or she had seen. The four participants in the experimental group were questioned while they were hypnotized and gave 14, 22, 18, and 17 accurate responses. The four participants in the control group gave 20, 25, 24, and 23 accurate responses. Using the .05 significance level, do hypnotized witnesses perform differently than witnesses who are not hypnotized?
     3. Use the independent t-test example.
        1. Look at how we used value labels and entered the data…remember each person gets their own row, with the IV and the DV labeled as two different columns.



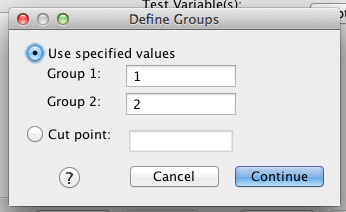
* + 1. Analyze > compare means > independent samples t-test.

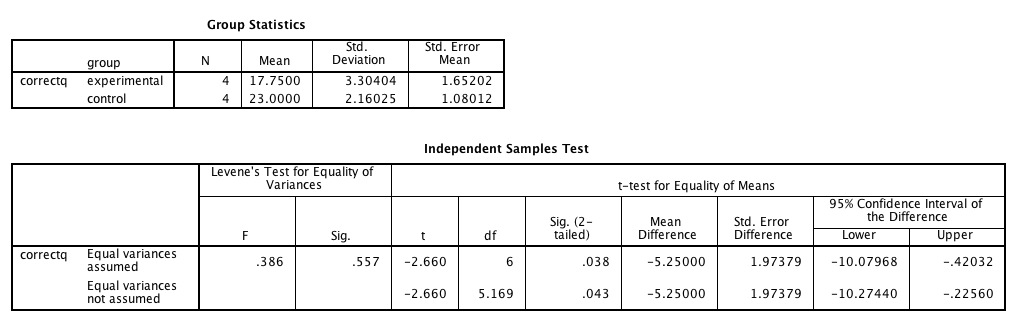


* + 1. Move over the “scores”- the thing you measured into the first box (the DV)



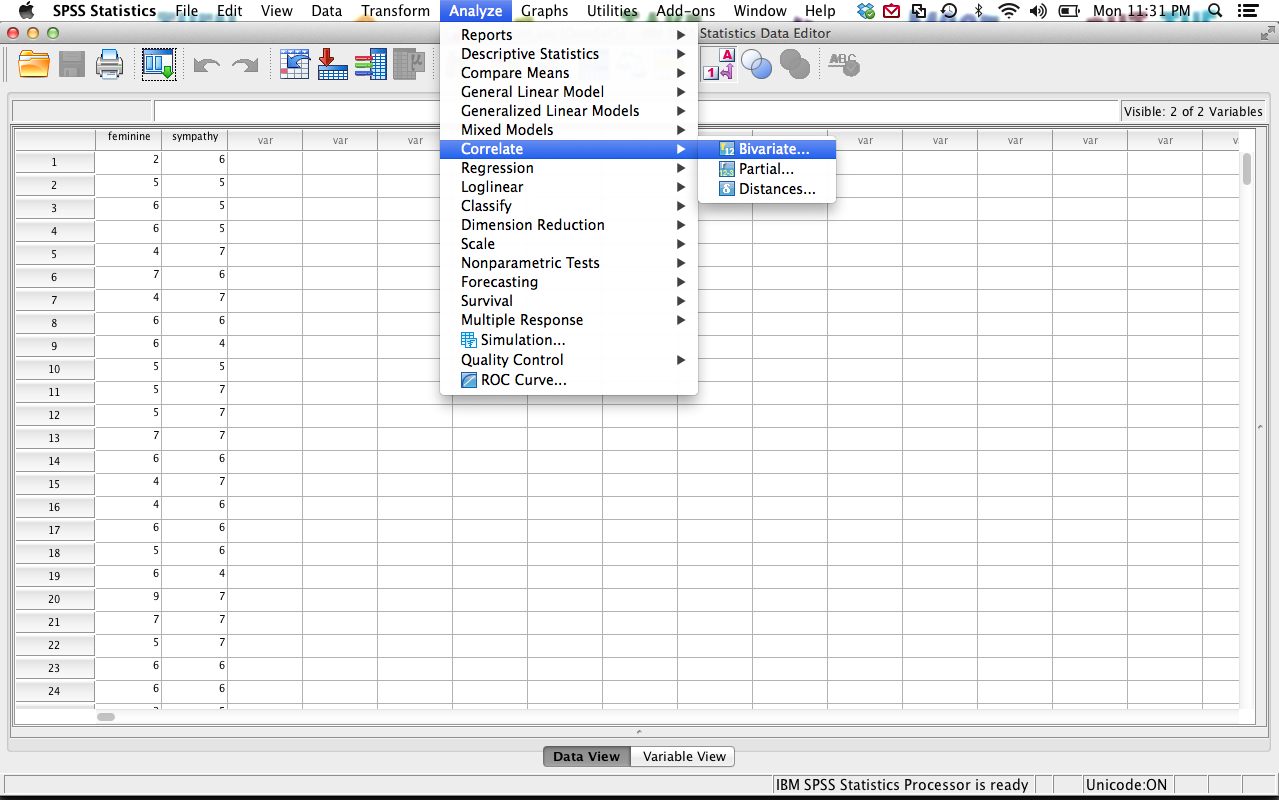
* + 1. Move the group labels (the IV) into the second box (go back and see how to do the value labels). See the question marks? We can’t get any output until we fix them (basically tell SPSS what to do).
       1. Hit define groups.
       2. Type in the numbers you used to label the groups (it’s easiest to go with 1 and 2, but use the numbers you labeled them as).
       3. Hit continue and ok



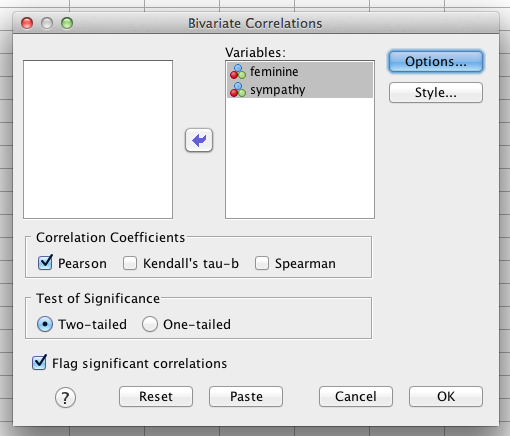


* + 1. Output:
       1. The first box contains the means and standard deviations for each group. You will need this information.
       2. Most people use the equal variances assumed line – which will depend on how many people you have and if you tested for it (see data analysis chapter).
       3. The first two sections are Levine’s test for equal variances. You want this to be NOT significant (i.e. p>.05).
          1. You want equal variances … remember that the null is the equal hypothesis, so you do not want to reject the null and say they are unequal. That’s bad for you. ☹ (data screening is weird, more on this in the data screening chapter).
       4. T-values, degrees of freedom, and p-values next.
       5. Most people don’t talk about mean differences because that’s more common for a dependent t-test
       6. Write up:
       7. Experimental group *M* = 17.75, *SD* = 3.30
       8. Control group *M* = 23.00, *SD* = 2.16
       9. *t*(6) = -2.66, *p* = .04, *d* = 1.09
    2. Effect size: see MOTE user guide.

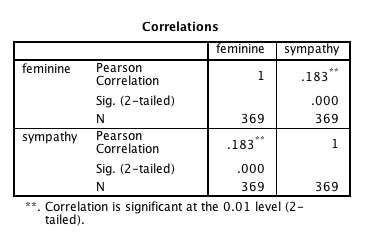
1. Correlation (Regression gets its own chapter)
   1. Uses: when you have two variables, but do not know which one *caused* the other one. You should be using at least mildly continuous variables.
   2. Types
      1. Pearson’s r
      2. Spearman’s rho
   3. Assumptions
      1. Normality
      2. Homogeneity
      3. Homoscedasticity – the spread of the errors for the X variable is the same all the way across the Y variable (equal errors)
      4. Linearity
   4. Example: Scores were measured for femininity and sympathy (see correlation.sav). Is there a correlation between those two variables?
      1. Analyze > Correlate > Bivariate



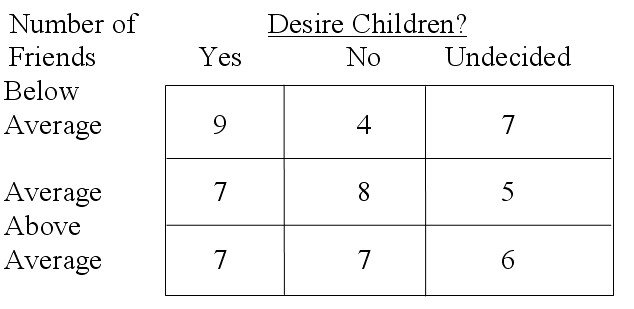
* + 1. Move all the variables you want to correlate to the right side and hit ok.

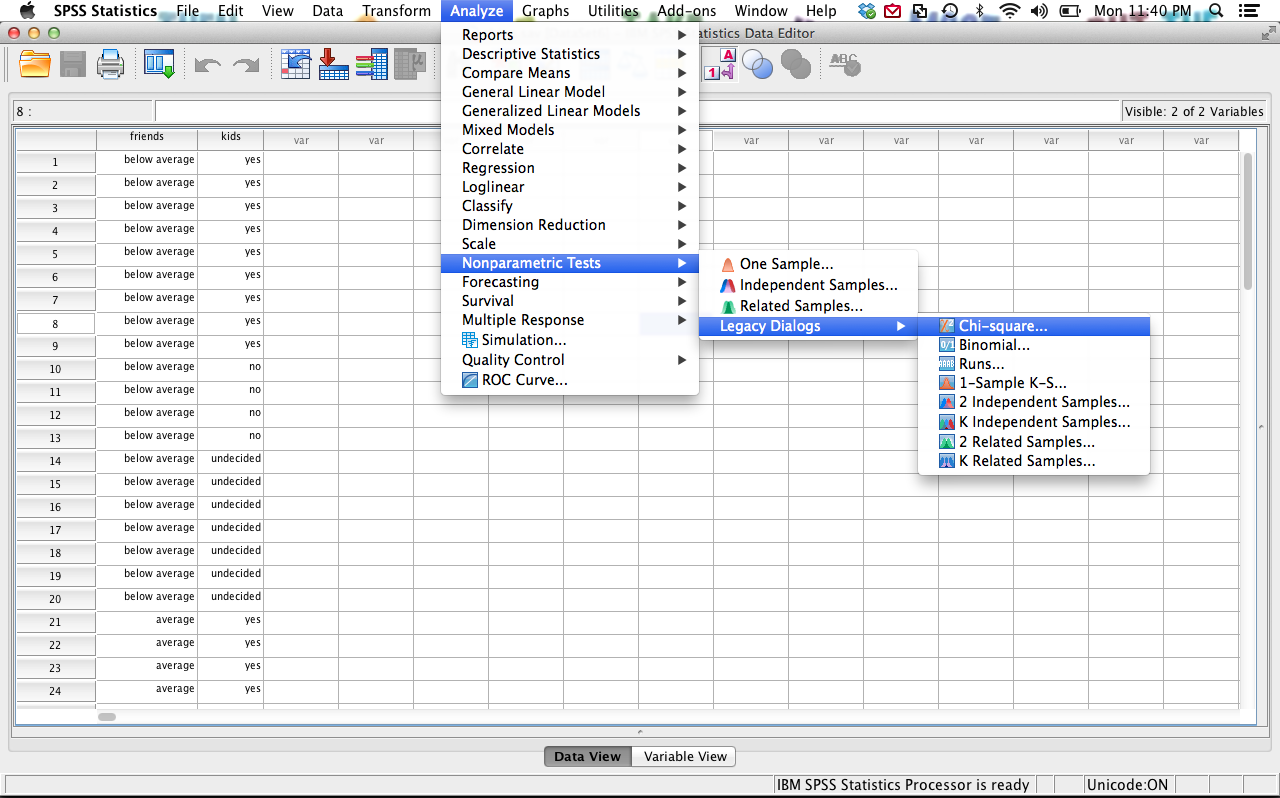


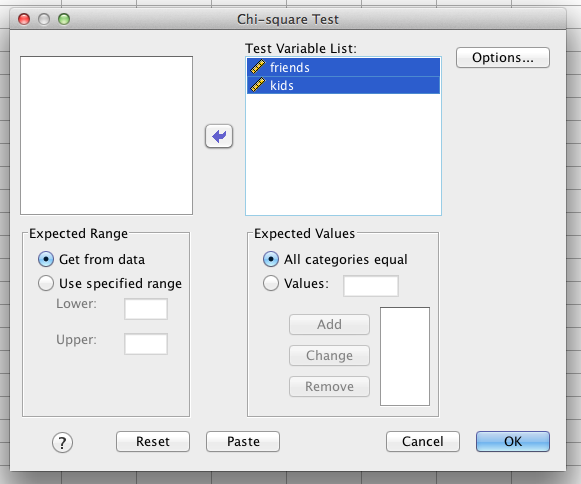
* + 1. Output: you will get a box of each variable paired with all the other variables. Basically you’ll get each one twice.



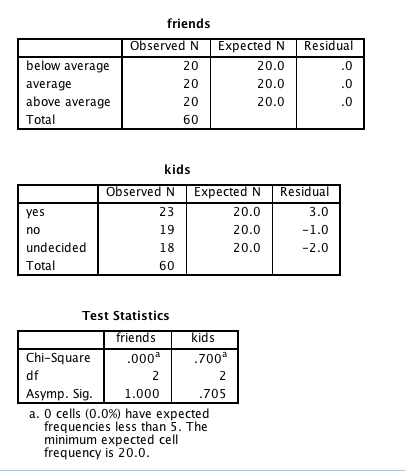
* + 1. Look for the variable combination of X and Y – here it’s .183 – the star means it is significant at p<.05. The second line is the p-value.
    2. Write: *r* = .18, *p* < .001.

1. Chi-Square
   1. Chi-square is a non-parametric test – meaning that you do not need the normal parametric assumptions.
   2. Assumptions:
      1. Each person can only go into one category.
      2. You need enough people in each category (no small frequencies or small expected frequencies).
   3. Uses: when you have nominal (discrete) data and want to understand if the categories are equal in frequency.
   4. Example: The following table shows results of a survey conducted at a particular high school in which students who had a small, average, or large number of friends were asked whether they planned to have children.
   5. 
   6. Goodness of fit (one variable at a time – really useful if you want to know if your categorical variables are split equally … log regression screener).
      1. This analysis tells you if your categories are roughly equal.
      2. Analyze > nonparametric tests > legacy dialogs > chi-square

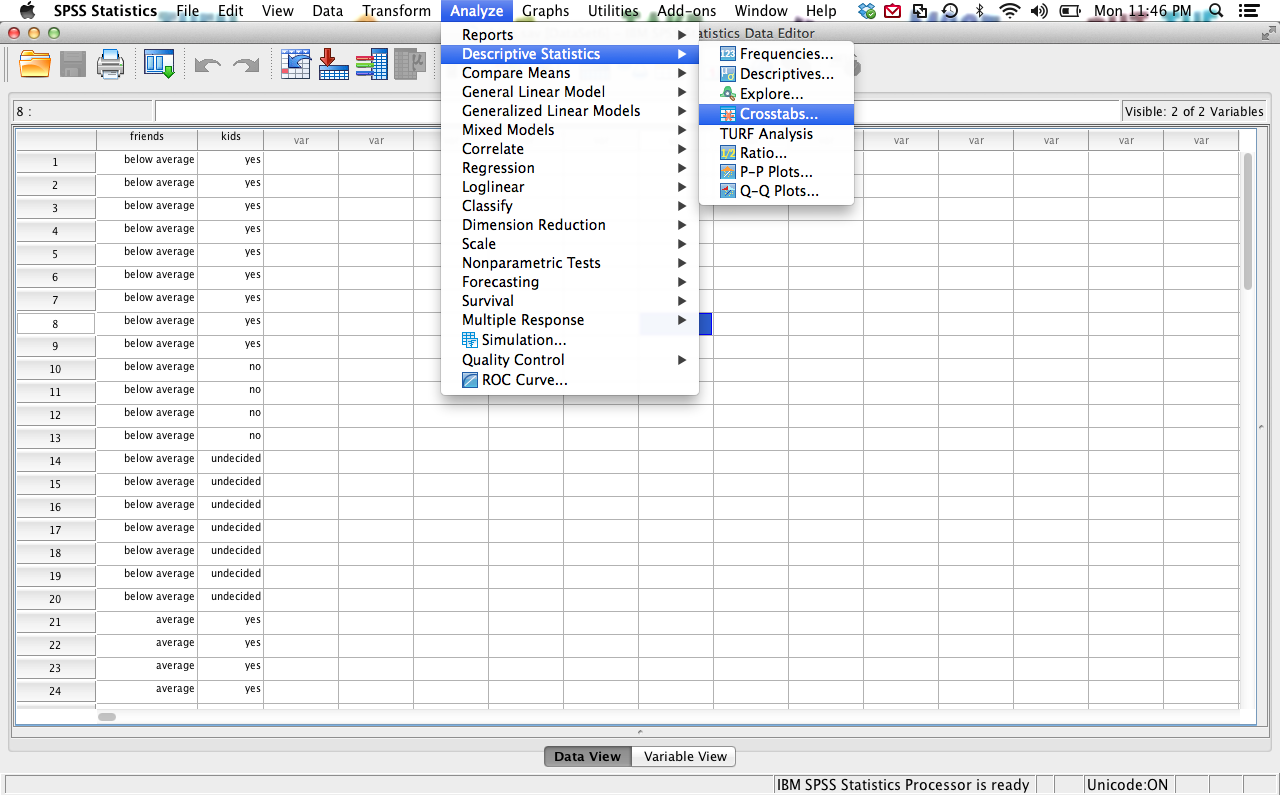


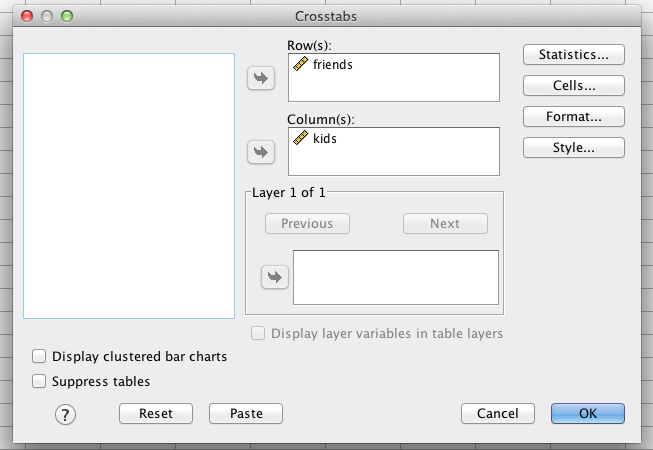


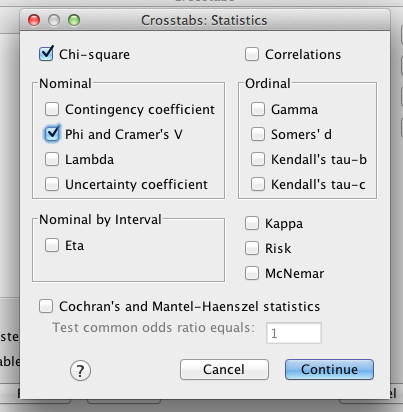
* 1. Output for goodness of fit.
     1. First two boxes tell you the observed number of people in each box and what they expected to find.
     2. The last box gives you the chi-square value, degrees of freedom and the p-value (asymp sig).
     3. Write: *X*2(2) = .70, *p* = .71
     4. (not really a normal effect size for this one).



* 1. Independence Test – examines if the variables are “related” by looking at the pattern of frequencies across rows and columns.
     1. This analysis tells you if rows and columns are what you would expect (aka there’s no relationship between variables, null hypothesis) OR if there are larger/smaller values than what you might expect (aka there is a relationship between variables, research hypothesis).
     2. Analyze > descriptive statistics > cross tabs
        1. Move one variable to the rows and one variable to the columns boxes.
        2. Hit statistics > chi-square, phi & Cramer’s V.







* 1. Output
     1. You really just want to look at the last box. The first boxes tell you the frequencies for each box and the expected values for each box.
     2. The last box tells you the chi-square (first line), degrees of freedom, and p-value.
     3. Next, you get the effect size box – either Phi or Cramer’s V are appropriate, but V is more commen.
     4. Write: *X*2(4) = 2.05, *p* = .73, V = .13.
  2. Effect size: Not quite in MOTE yet ☹
     1. 
     2. *X*2 is the chi-square value
     3. N = total number of participants
     4. k = the smaller number of rows or columns
        1. If you have 2 rows and 3 columns, that would be 2-1
     5. The difference between phi and v is that phi does not include (k-1) because it was meant to be only for 2X2 (2 rows, 2 columns), so the k-1 was the same either way and equal to 1. Most people use V because phi = V when you have an equal number of rows and columns and V is more appropriate for unequal rows and columns.

