Handbook Table of Contents

Compute a Variable…………………………………….2

Recode a Variable……………………………………….4

Select Cases………………………………………………..6

Split File…………………………………………………….9

Histogram with Frequency Table………………..11

Descriptive Statistics……………………………….. 14

Mean, Median, Mode, Standard Deviation, Range, Skew

Internal Reliability Analysis…………………………17

Dependent Samples t-test…………………………….20

Independent Samples t-test …………………………23

One-Way ANOVA with post hoc test………………26

Factorial (2-way) ANOVA……………………………..30

Repeated Measures ANOVA…………………………..35

Correlation…………………………………………………..39

Linear Regression………………………………………..42

Scatterplot with Regression Line………………….47

Goodness of Fit Chi-Square………………………….51

Contingency Table Chi-Square……………………54

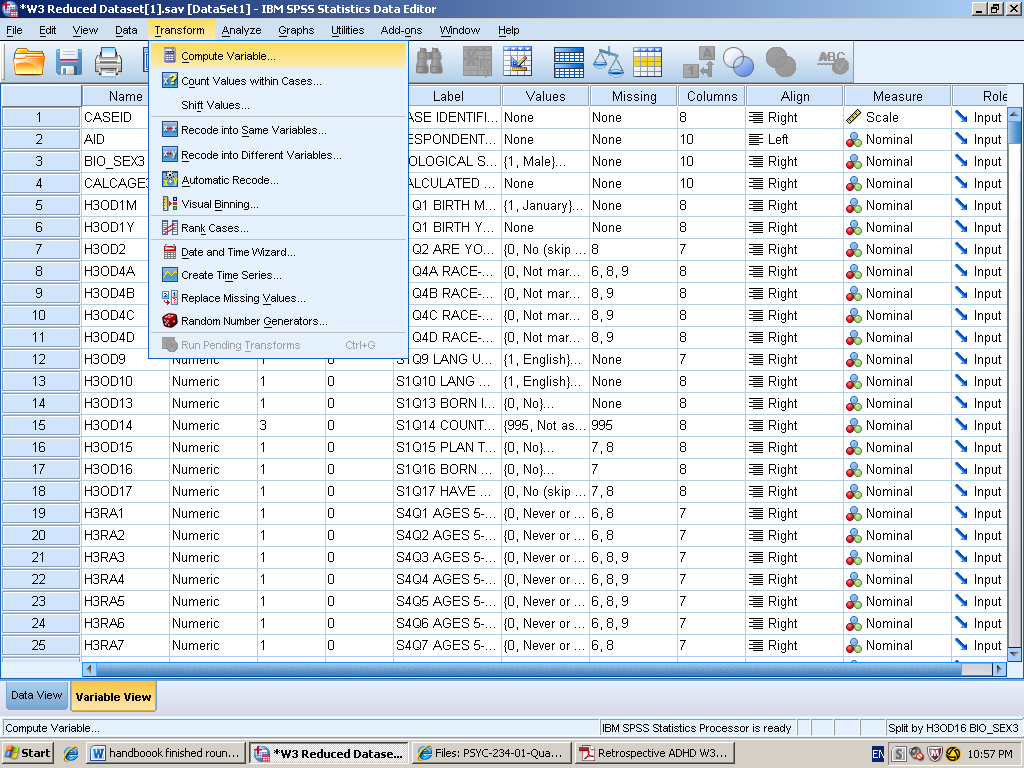
Compute a Variable

**Description of Statistic**

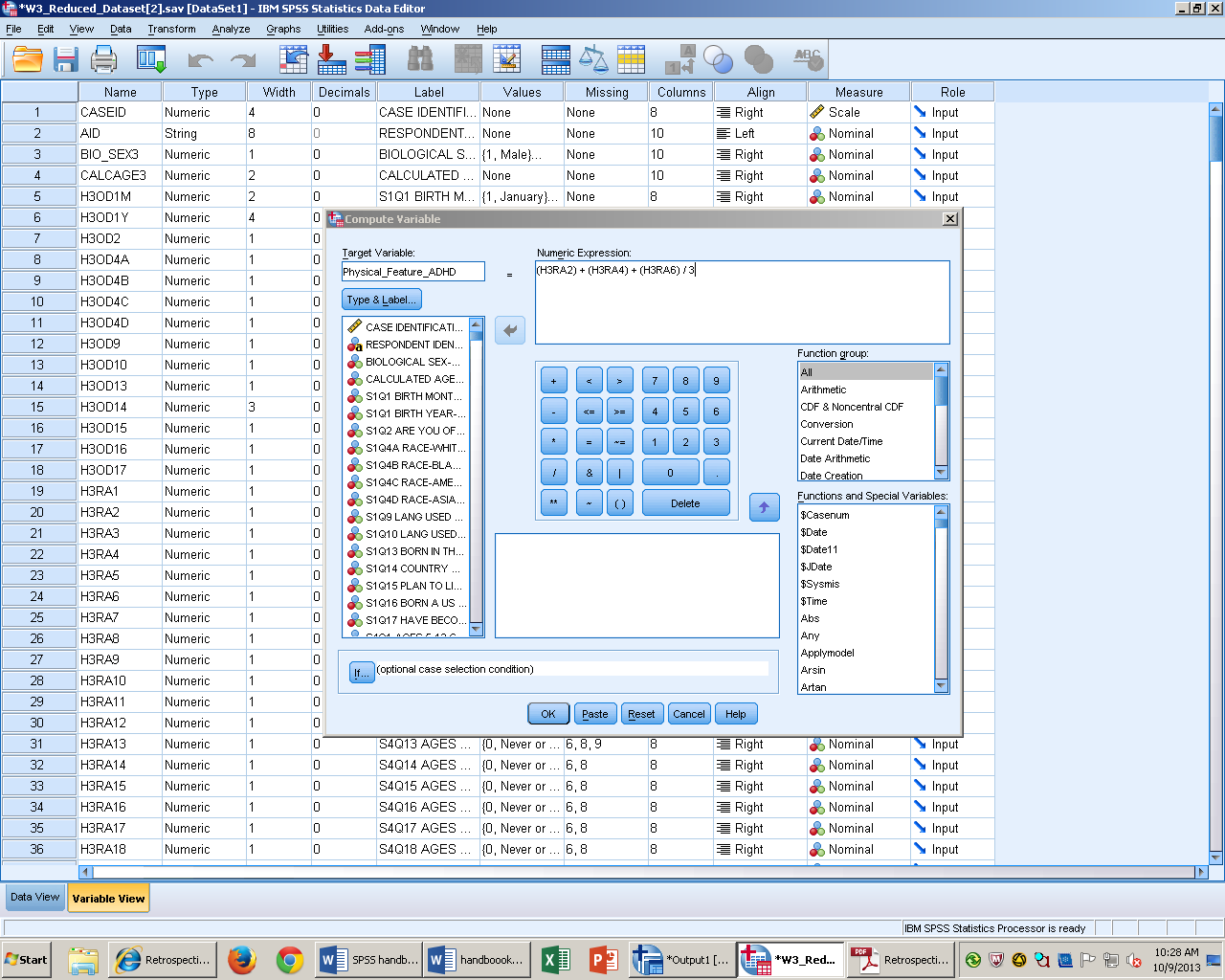
Computing a variable is used when you want to combine existing variables in order to make a new variable. When computing variables they need to be of the same scale of measurement to be able to interrupt the results. An example of computing a variable is a midterm grade. A midterm grade is combined of quizzes, tests, and paper variables that create one single variable, a midterm grade.

**SPSS Steps**

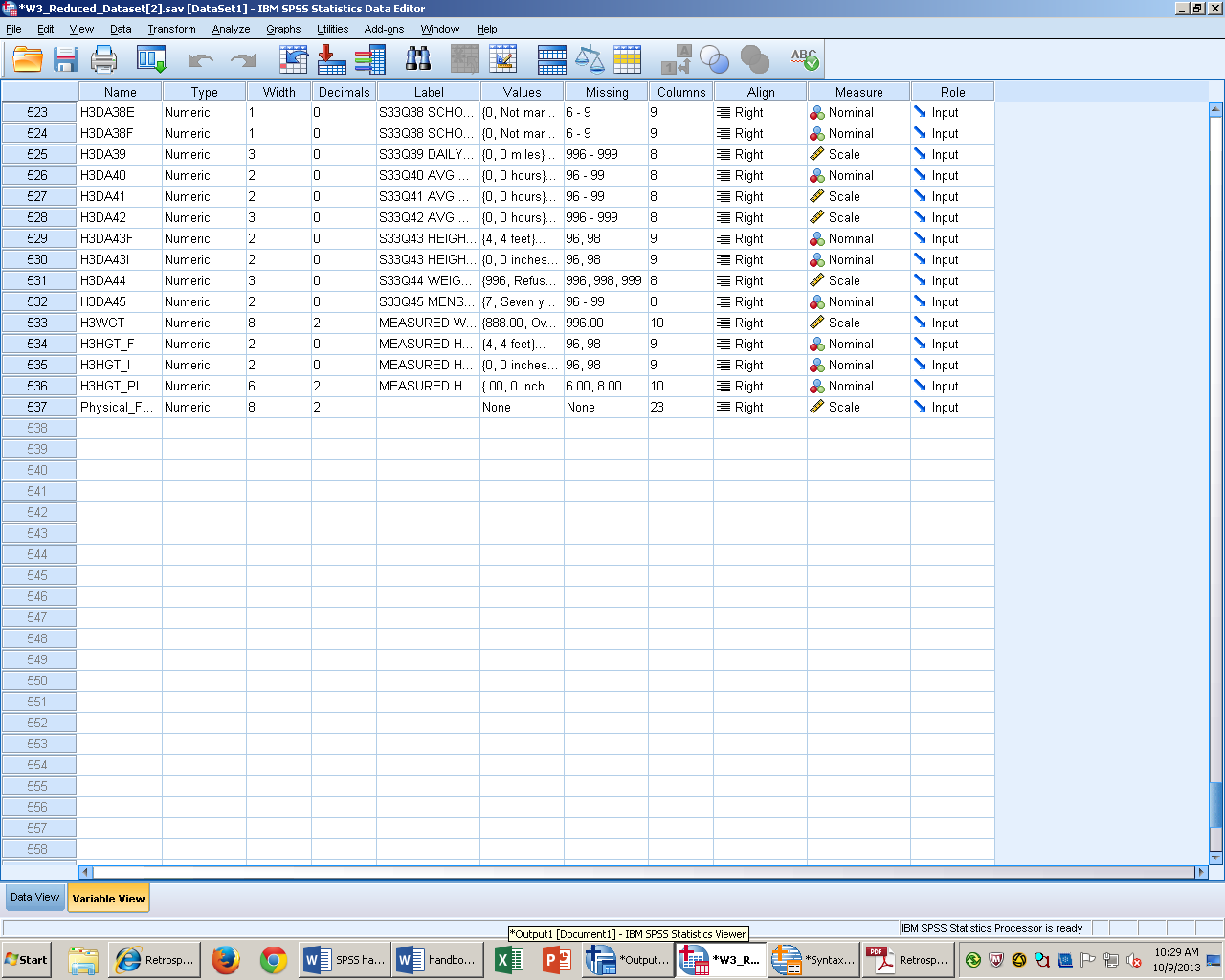
1. *Transform* *Compute Variable*



1. Select the variables of interest from the left hand column to the “Numeric Expressions” box.
2. Give the target variable a name in the “Target Variable” box. Again, SPSS does not like spaces. Then add together the variables you went to combine by putting parentheses around the variables and plus signs in-between. To get the mean divide by the number of variables you have selected.



1. Click paste to put it into syntax, and then run the analysis.
2. Scroll to the bottom of the data set the computed variable will be the last variable in the data set.



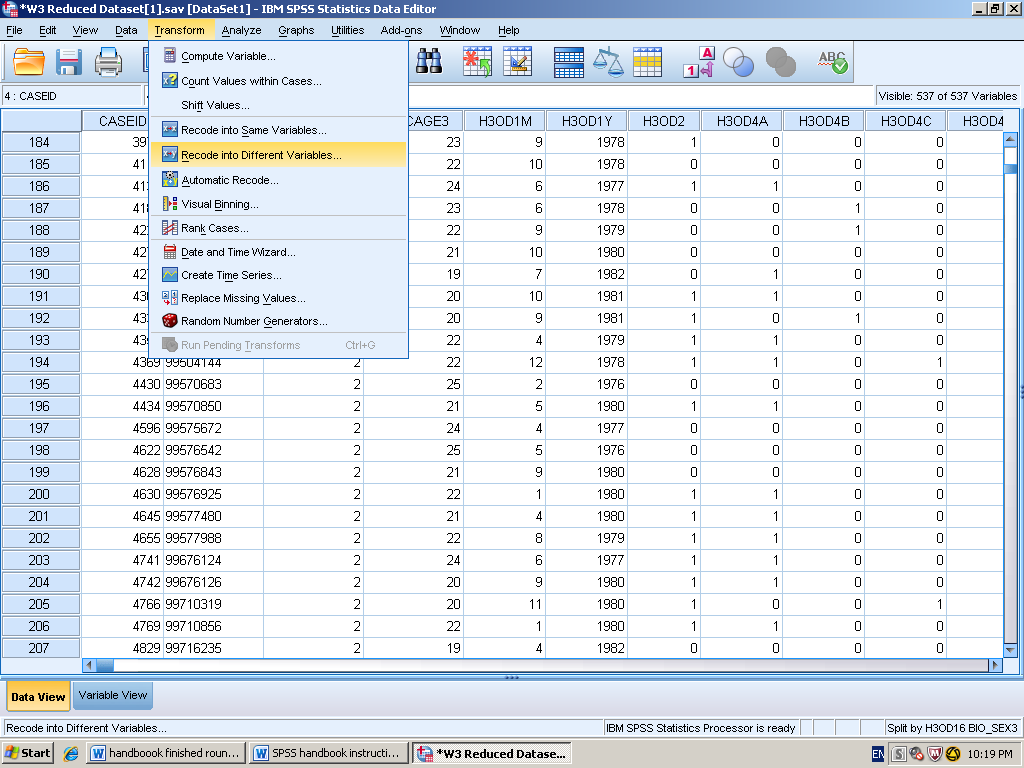
Recode a Variable

**Description of Statistic**

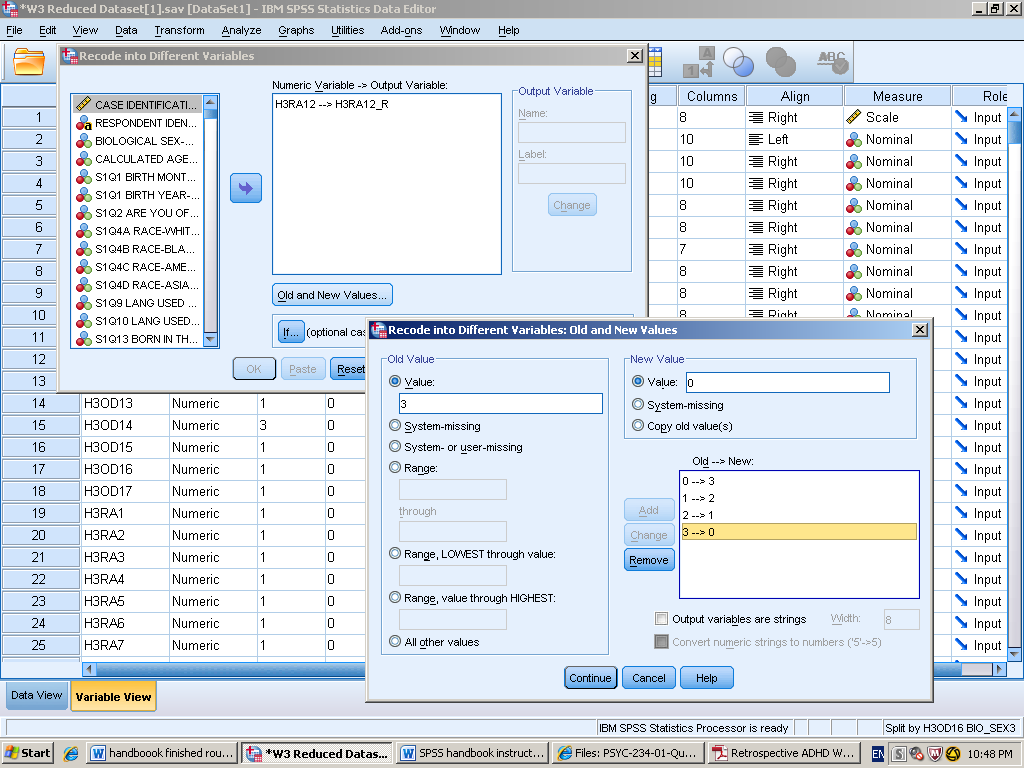
Recoding a variable is used when you want to create a *different meaning* for a variable. Recoding a variable can be done in either one of two ways. First, reverse recode where you give a reversed scale of measurement to the variable. Second, recoding condenses lots of variables from a ratio scale variable into a nominal scale. The disadvantage to this function is the variable you reversed scale doesn’t exactly measure the variable you are reversing.

**SPSS Steps**

1. *Transform*  *Recode into Different Variables*



1. Select variable of interest from the left hand side into the “Numeric Variable – Output Variable” box. Then rename the variable under the “Output variable” box. Remember, SPSS does not like spaces.
2. Click change and then click “Old and New variables.” This is when you flip the order around to compare. Then click continue.



1. Click paste to put into syntax.

**Syntax**

RECODE H3RA12 (0=3) (1=2) (2=1) (3=0) INTO H3RA12\_R.

EXECUTE.

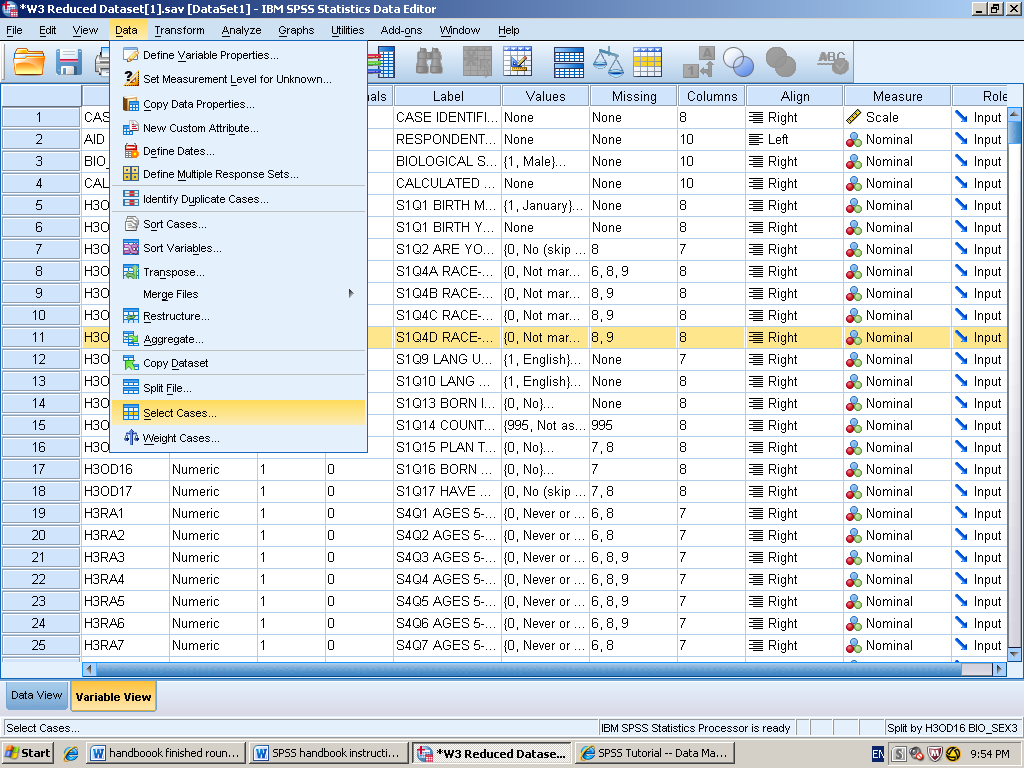
Select Cases

**Description of statistics**

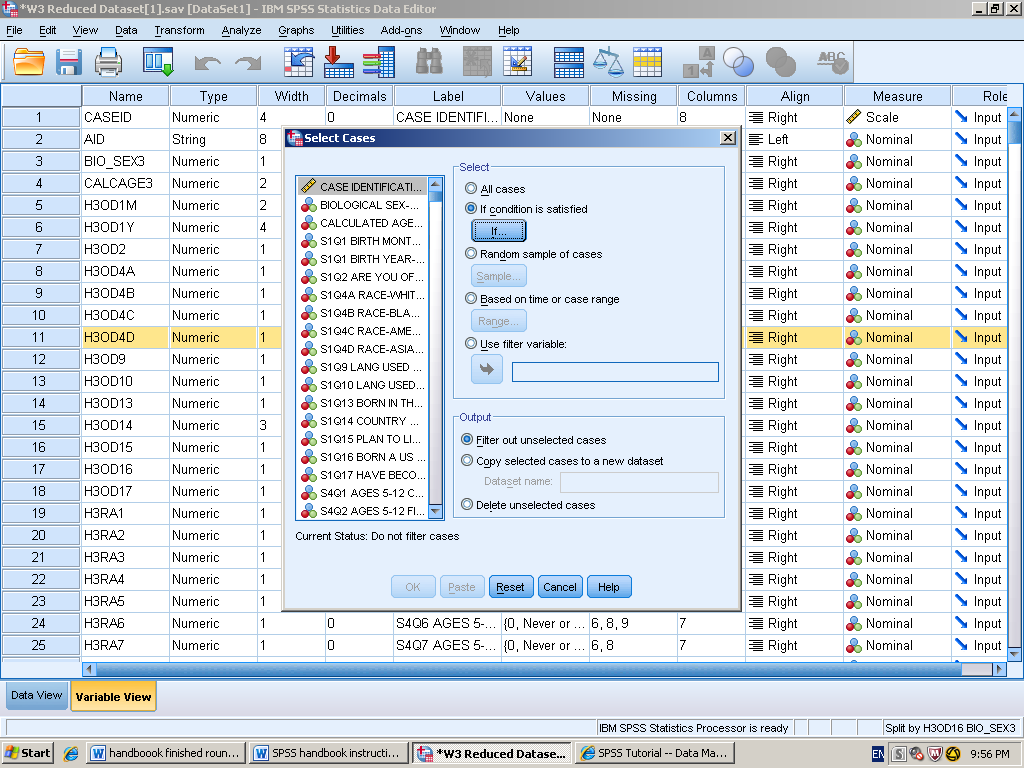
Select Cases allows us to select part of the data set for further analysis while excluding the remaining cases. Select cases focuses on one group or subgroup of a case, but does not do anything in isolation.

**SPSS Steps**

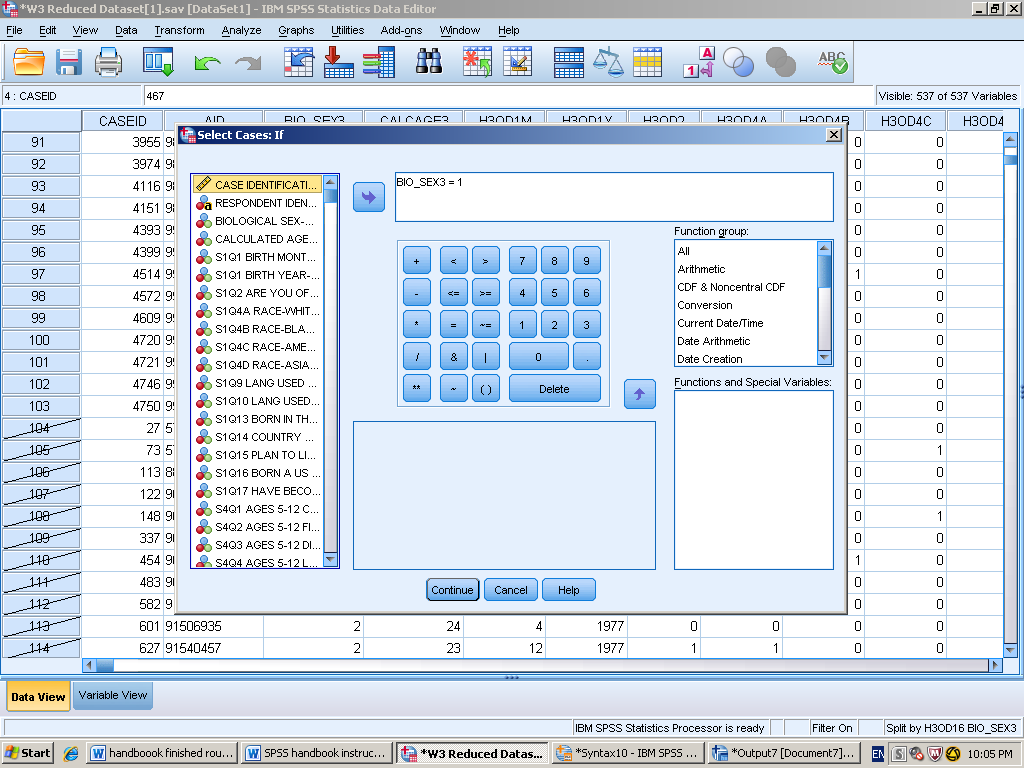
1. *Data* *Select Cases*



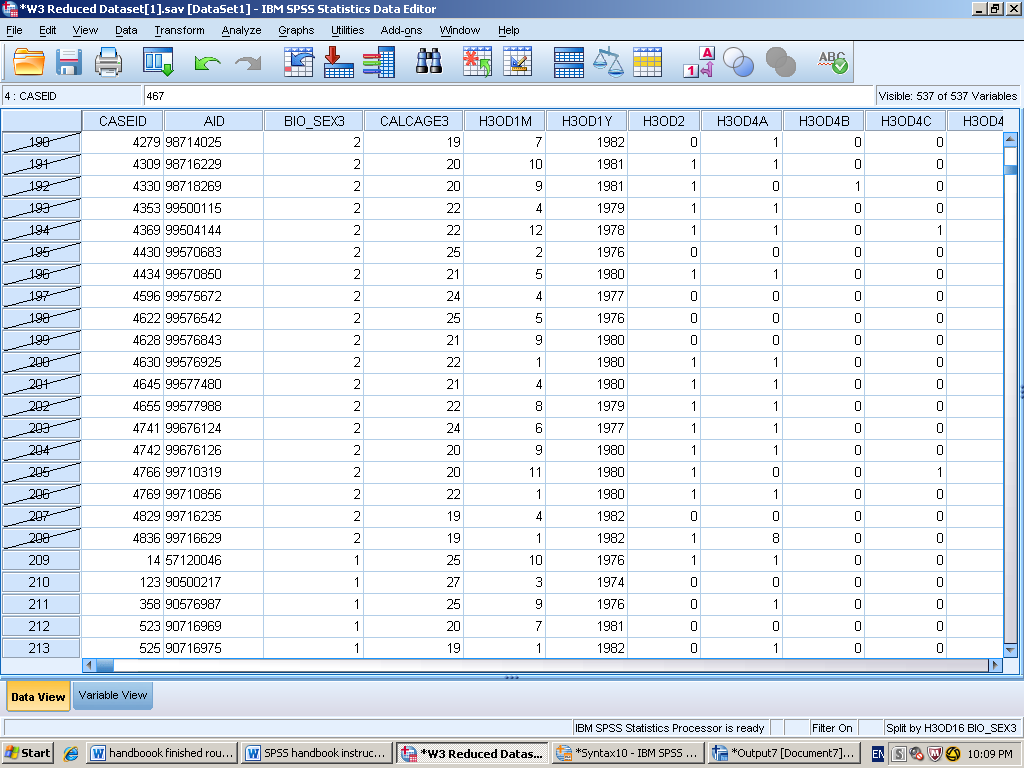
1. In the “Select” box click on “If condition is satisfied.”



1. Select a variable from the left hand side that you would want to create a subgroup of. Enter = 1 to “BIO\_SEX3” in the long box on the left side to single out only the men in the case. Then press continue



1. Click paste to put into the syntax, and then look at data. SPSS crosses out the cases not in use for the analysis you are about to run.



**Syntax**

USE ALL.

COMPUTE filter\_$=(BIO\_SEX3 = 1 ).

VARIABLE LABELS filter\_$ 'BIO\_SEX3 = 1 (FILTER)'.

VALUE LABELS filter\_$ 0 'Not Selected' 1 'Selected'.

FORMATS filter\_$ (f1.0).

FILTER BY filter\_$.

EXECUTE.

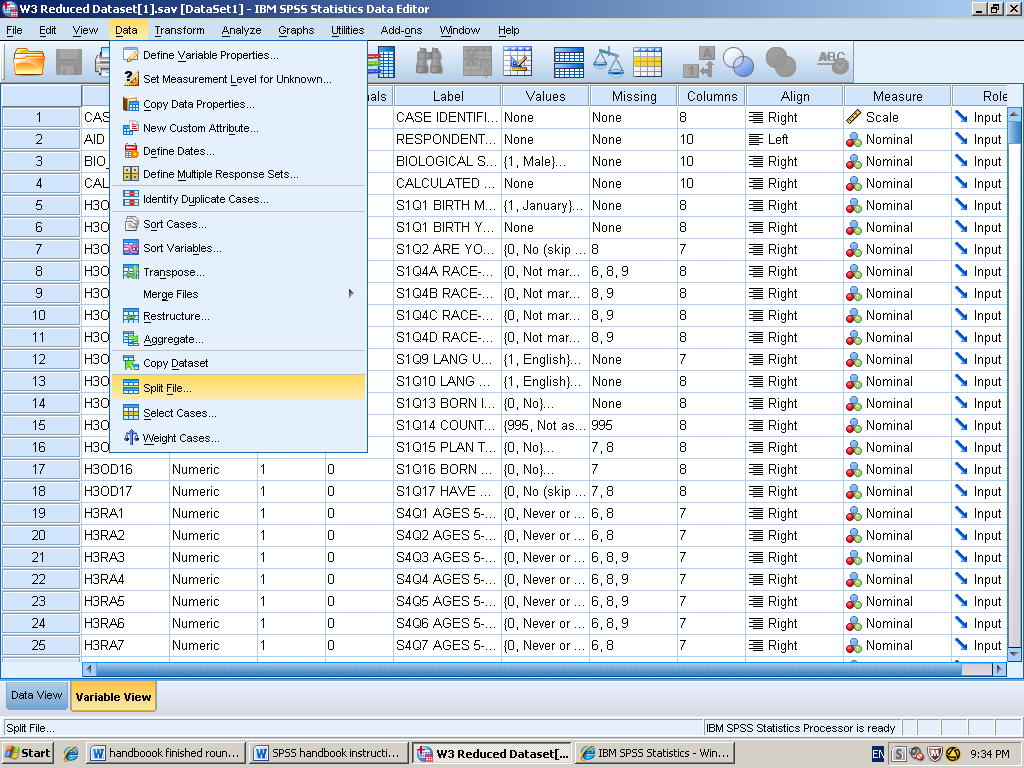
Split File

**Description of statistics**

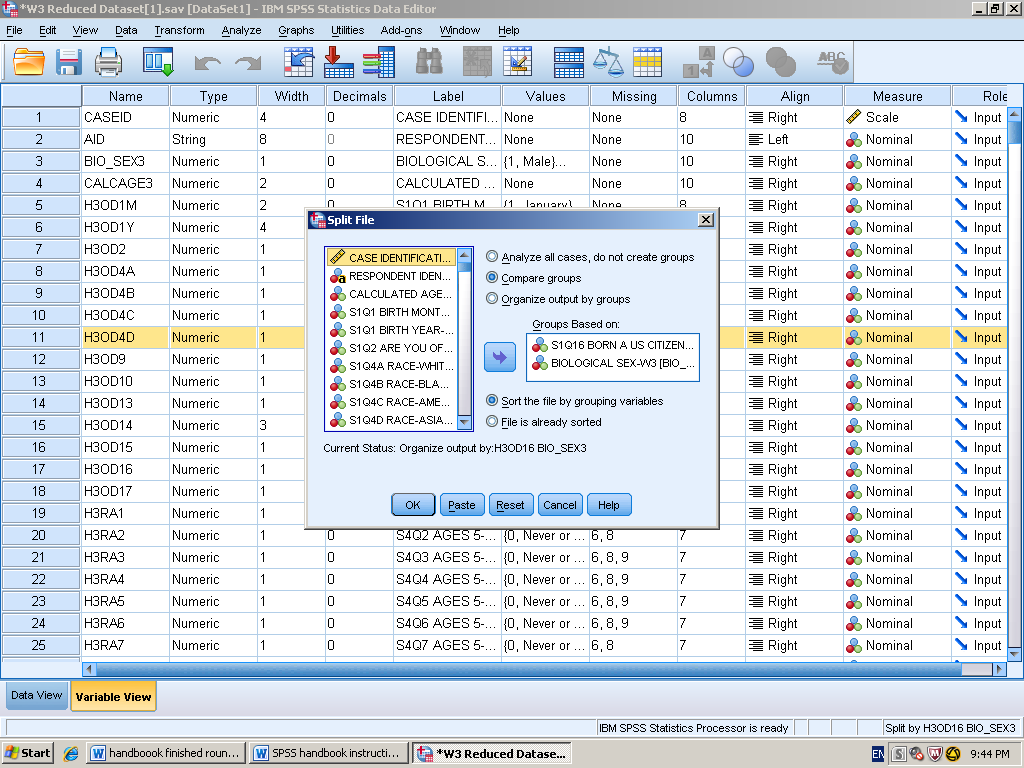
Split File splits the data into separate groups for analysis based on the characteristics for one or more variables. Multiple groups of variables are categorized within each variable for analysis. For example, if you select gender as the first grouping variable and then born in the United States as the second group variable. The data will be categorized by birth place within each gender classification. Split file focuses on organization of groups but does not do an analysis in isolation.

**SPSS Steps**

1. *Data* *Split File*



1. Select the variables of interest and click the arrow to move it from left hand list of variables to the right hand “Groups based on box.” You can select up to 8 variables to be grouped on.
2. Select “compare groups” on the top of the right hand side. Then click “sort the file by grouping variables” on the bottom of the right hand side.



1. Click paste to put into the syntax, again split file does not run an analysis independently you have to tell SPSS what to do next.

**Syntax**

SORT CASES BY H3OD16 BIO\_SEX3.

SPLIT FILE LAYERED BY H3OD16 BIO\_SEX3.

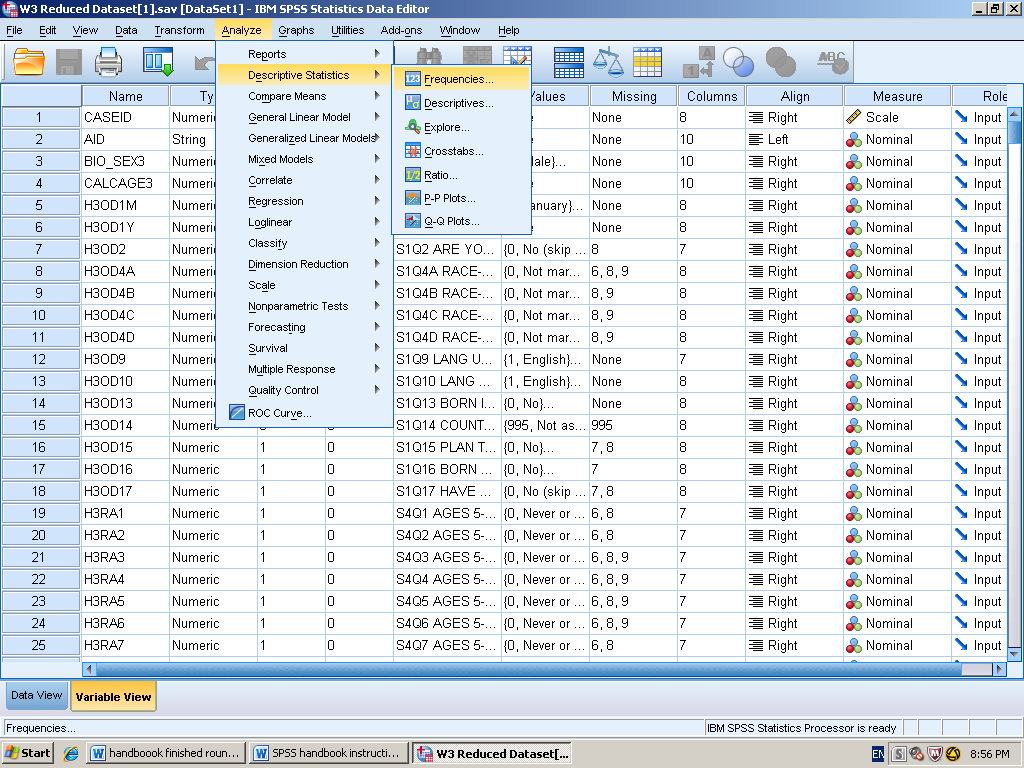
Histogram with Frequency Table

**Description of Statistics:**

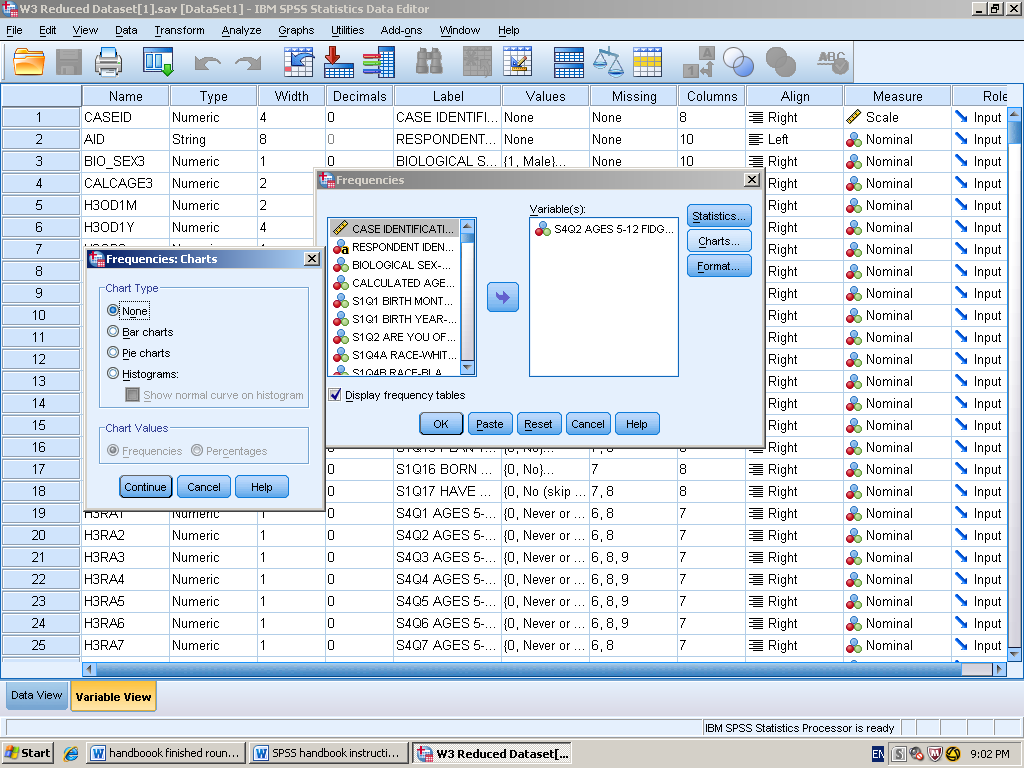
Histogram with frequency table is used to help understand the descriptive statistics data in a visual sense. Histograms are a type of bar graph. Each bar represents frequency for a particular value. A grouped frequency table displays scores organized into class intervals. To most accurately and clearly share information these tables should have class intervals all the same width. They should have continuous class intervals. Also, should use around 5-15 intervals, too many intervals can become confusing. Histograms can also help catch user errors, such as entering in a wrong number. The mode in a histogram is the highest point in the distribution. The histogram can also tell you if the data is unimodal, bimodal, and multi-modal.

**SPSS Steps**

1. *Analyze* *Descriptive Statistics* *Frequencies*



1. Select the variable of interest and click the arrow to move it from the left hand list of variables to the right hand “Variable(s)” box.
2. Click on “charts” on the right hand side. Then in the “chart box” click on Histograms and then click continue.



1. Click paste to put into syntax, and then run the analysis.

**Syntax**

FREQUENCIES VARIABLES=H3RA2

/STATISTICS=STDDEV RANGE MEAN MEDIAN MODE SKEWNESS SESKEW

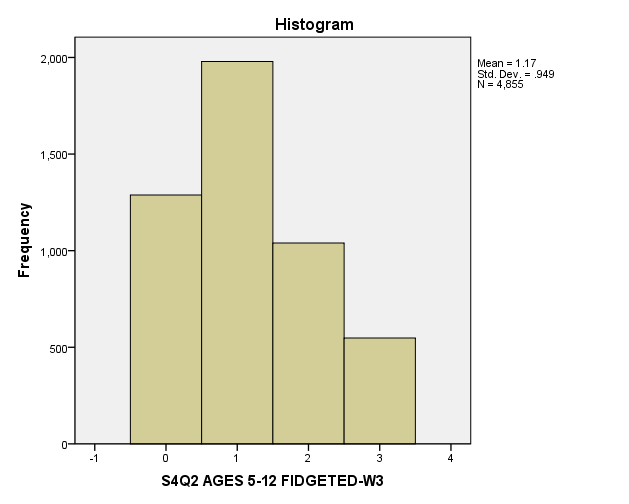
/HISTOGRAM

/ORDER=ANALYSIS.

| **S4Q2 AGES 5-12 FIDGETED-W3** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | Never or rarely | 1288 | 26.4 | 26.5 | 26.5 |
| Sometimes | 1979 | 40.5 | 40.8 | 67.3 |
| Often | 1040 | 21.3 | 21.4 | 88.7 |
| Very often | 548 | 11.2 | 11.3 | 100.0 |
| Total | 4855 | 99.4 | 100.0 |  |
| Missing | Refused | 1 | .0 |  |  |
| Don't know | 25 | .5 |  |  |
| System | 1 | .0 |  |  |
| Total | 27 | .6 |  |  |
| Total | | 4882 | 100.0 |  |  |

**Output Interpretation**

| **Statistics** | | |
| --- | --- | --- |
| S4Q2 AGES 5-12 FIDGETED-W3 | | |
| N | Valid | 4855 |
| Missing | 27 |
| Mean | | 1.17 |
| Median | | 1.00 |
| Mode | | 1 |
| Std. Deviation | | .949 |
| Skewness | | .439 |
| Std. Error of Skewness | | .035 |
| Range | | 3 |



The data states that the mode the variable that appears the most is 1. The mode is also the highest point on the histogram. The histogram is positively skewed, the data is the most clumped on the right side of the chart. The histogram is unimodal, meaning there is only one peak in the data. Mode, the variable that appears the most is 1, sometimes was the most common respond when participants were ask to respond if they fidgeted with their hands, feet, or squirmed in their seat.

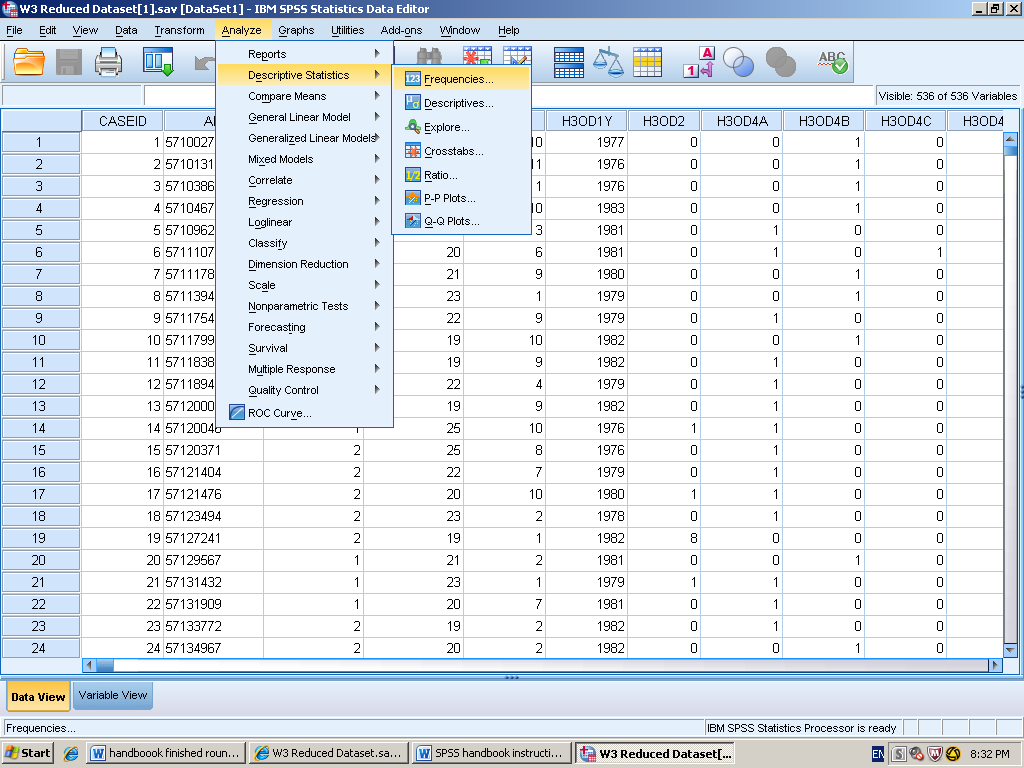
Descriptive Statistics

**Description of Statistics:**

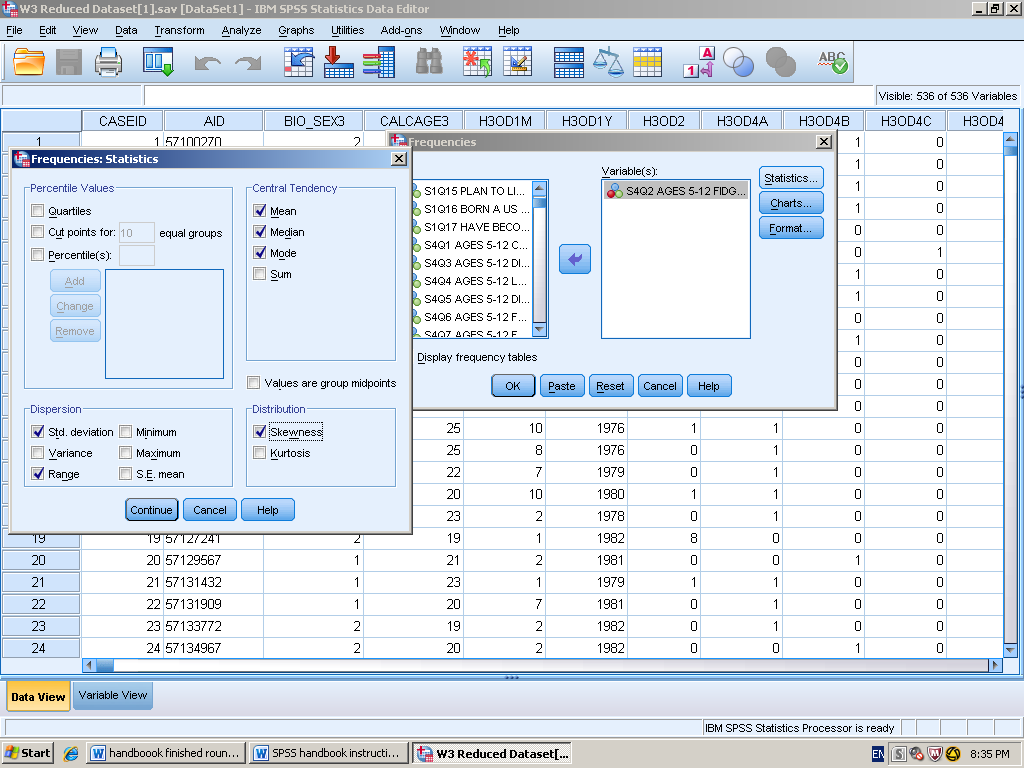
Descriptive statistics quantitatively describe the basic features of data. They provide summaries about the data and measurements. A measure of central tendency (mode, median, and mean) gives us a descriptive value. Mode is the most common score in a distribution. Mode is also the least affected by outliers. Median is the middle score in a distribution. Mean is the balance point of the distribution of scores. It is also the most common measure of central tendency used in psychology. Measures of variability (range and variance) tell how spread out the scores are in a distribution. Range is the difference between the highest and the lowest scores. Variance is the average of each score’s squared difference from the mean. Standard deviation is the most common way of describing variability. It is the average amount the scores differ from the mean. Skew defines the data is clustered at one end. Positively skewed means the data is clustered at the left. Negatively skewed means the data is clustered at the right.

**SPSS Steps**

1. *Analyze Descriptive Statistics Frequencies*



1. Select the variable of interest and click the arrow to move it from the left hand list of variables to the “Variables” box. Then click “Statistics.”
2. In the “Frequencies: Statistics” box under central tendency click mean, median, and mode. In the “Dispersion” box click standard deviation and range. Then in the “Distribution” box click Skewness. Than press continue.



1. Click paste to put into syntax, and then run the analysis.

**Syntax**

DATASET ACTIVATE DataSet1.

FREQUENCIES VARIABLES=H3RA2

/STATISTICS=STDDEV RANGE MEAN MEDIAN MODE SKEWNESS SESKEW /ORDER=ANALYSIS.

**Output Interpretation**

| **Statistics** | | |
| --- | --- | --- |
| S4Q2 AGES 5-12 FIDGETED-W3 | | |
| N | Valid | 4855 |
| Missing | 27 |
| Mean | | 1.17 |
| Median | | 1.00 |
| Mode | | 1 |
| Std. Deviation | | .949 |
| Skewness | | .439 |
| Std. Error of Skewness | | .035 |
| Range | | 3 |

| **S4Q2 AGES 5-12 FIDGETED-W3** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | Never or rarely | 1288 | 26.4 | 26.5 | 26.5 |
| Sometimes | 1979 | 40.5 | 40.8 | 67.3 |
| Often | 1040 | 21.3 | 21.4 | 88.7 |
| Very often | 548 | 11.2 | 11.3 | 100.0 |
| Total | 4855 | 99.4 | 100.0 |  |
| Missing | Refused | 1 | .0 |  |  |
| Don't know | 25 | .5 |  |  |
| System | 1 | .0 |  |  |
| Total | 27 | .6 |  |  |
| Total | | 4882 | 100.0 |  |  |

The data states that the mode, the variable that appears the most is 1, sometimes was the most common respond when participants were ask to respond if they fidgeted with their hands, feet, or squirmed in their seat. The *M* the balance point of the distribution of scores is 1.17. The median, the middle score in the distribution is also 1. The *SD* is .949, stating how far the variables are spread out from the mean. The range is 3, meaning that every option was selected from all the cases. Participants fidgeted with their hands, feet,or squirmed in their seat

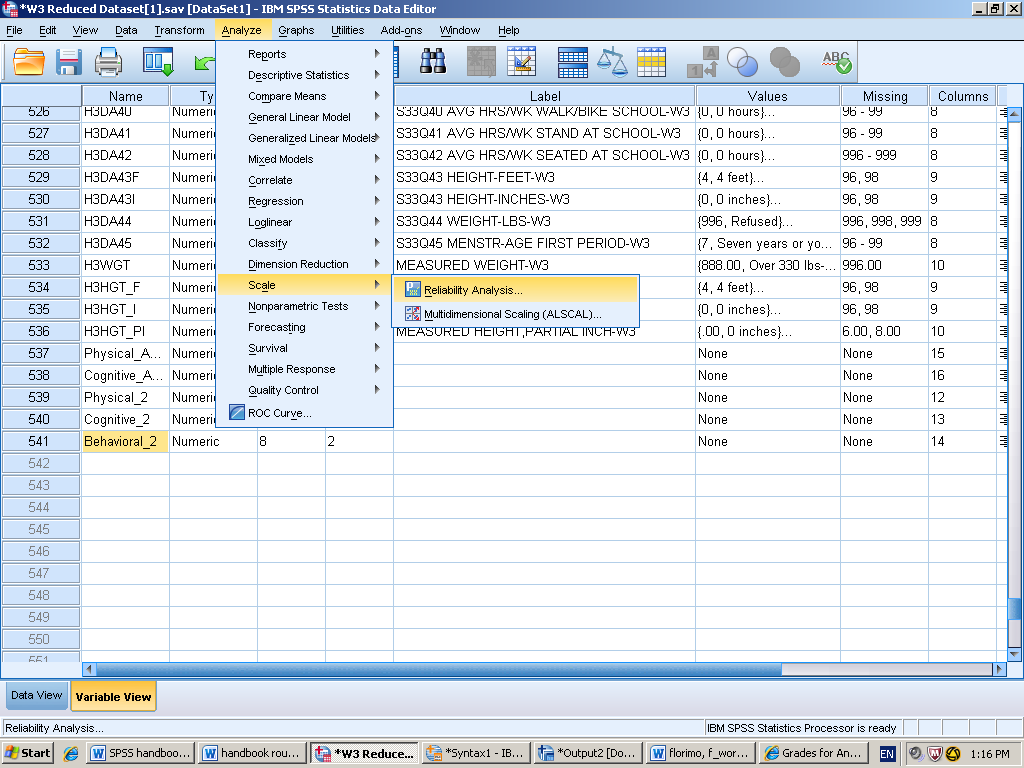
Internal Reliability Analysis

**Description Statistics**

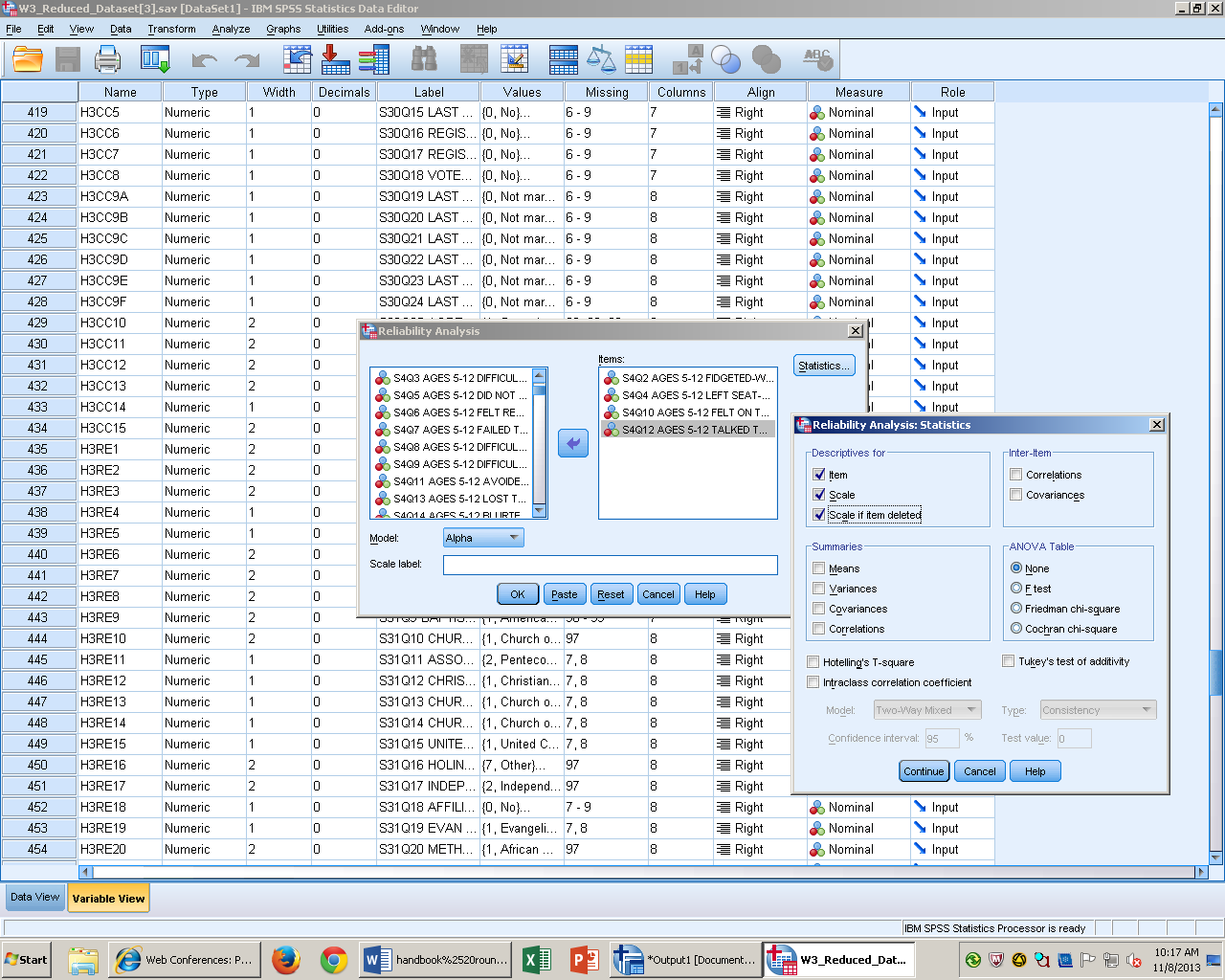
Reliability is the degree of consistency or stability of measure. There are four types of reliability; inter-rater, test-retest, split-half, and internal consistency reliability. Inter-rater reliability is when the main measure is an observation or coding of written or spoken material. Test-retest is when the same group of people is given the same measure but at two different points in time. Split-half reliability is where a test is split in two and the scores of half of the test are compared to each other. Internal consistency reliability is used when you divide measures into halves, of all possible halves. Then, find correlation of all possible halves and then find the mean. Cronbach Alpha is the most common measure of internal consistency. When making meaningful and reliable new variables follow the guidelines common idea, same scale of measure, and that the numbers mean the same thing.

**SPSS Steps**

1. *Analyze Scale Reliability Analysis*



1. Select the variables of interest and click the arrow to move it from the left hand list of variables to the right hand “items” box. You can select more than one variable to test at a time.
2. Set the model to alpha. Next click on “statistics” and under descriptive click item, scale, scale if item deleted. Then press continue.



1. Click paste to put into syntax, then run the analysis.

**Syntax**

RELIABILITY

/VARIABLES=H3RA2 H3RA4 H3RA10 H3RA12

/SCALE('ALL VARIABLES') ALL

/MODEL=ALPHA

/STATISTICS=DESCRIPTIVE SCALE

/SUMMARY=TOTAL.

**Output interpretation**

| **Case Processing Summary** | | | |
| --- | --- | --- | --- |
|  | | N | % |
| Cases | Valid | 4821 | 98.8 |
| Excludeda | 61 | 1.2 |
| Total | 4882 | 100.0 |
| 1. Listwise deletion based on all variables in the procedure. | | | |
|  | | | |

| **Reliability Statistics** | |
| --- | --- |
| Cronbach's Alpha | N of Items |
| .611 | 4 |

| **Item-Total Statistics** | | | | |
| --- | --- | --- | --- | --- |
|  | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted |
| S4Q2 AGES 5-12 FIDGETED-W3 | 2.48 | 3.936 | .381 | .549 |
| S4Q4 AGES 5-12 LEFT SEAT-W3 | 3.23 | 4.492 | .399 | .545 |
| S4Q10 AGES 5-12 FELT ON THE GO-W3 | 2.72 | 3.948 | .401 | .533 |
| S4Q12 AGES 5-12 TALKED TOO MUCH-W3 | 2.54 | 3.558 | .403 | .536 |

After running the analysis, the cronbach alpha is .61, this is not a good rating for reliability. A cronbach alpha is 0- 1, but anything above .8 is respectable. Under cronbach alpha “if item deleted” tells us if we delete that question, what our new alpha will be. If I took out any of the questions my cronbah alpha would still be considered poor. In order to take out a question you would want smaller numbers under “corrected item-total correlation” and larger numbers under “Cronbach’s Alpha if item deleted” column.”

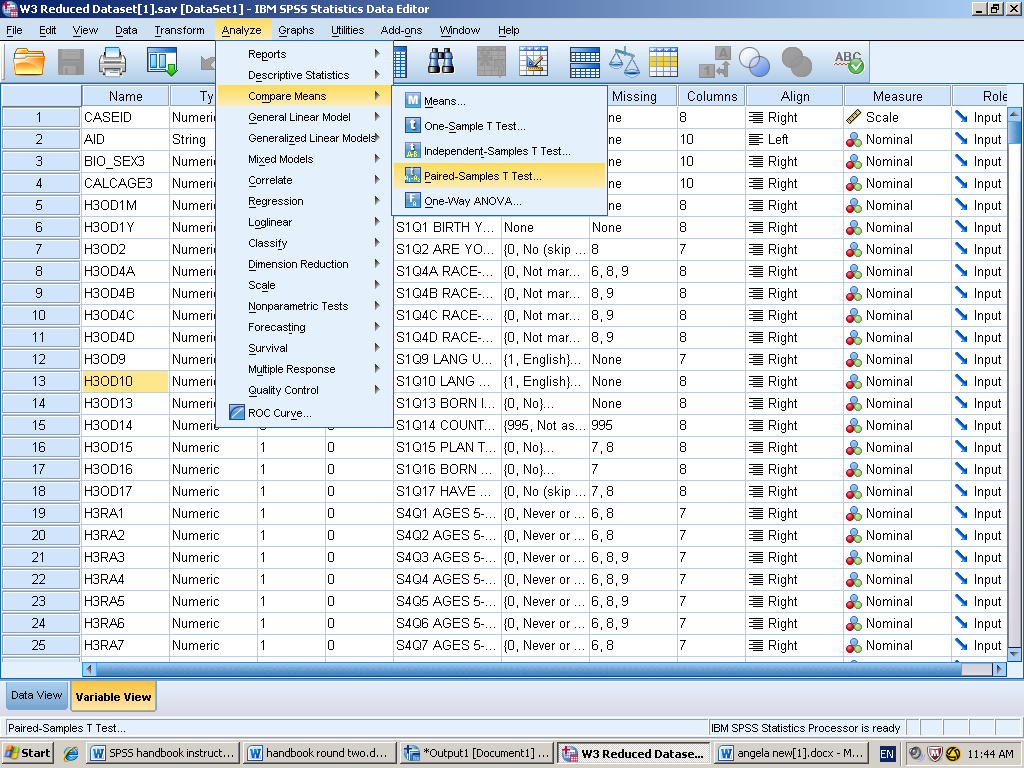
Dependent Samples *t*-test

**Description of Statistic**

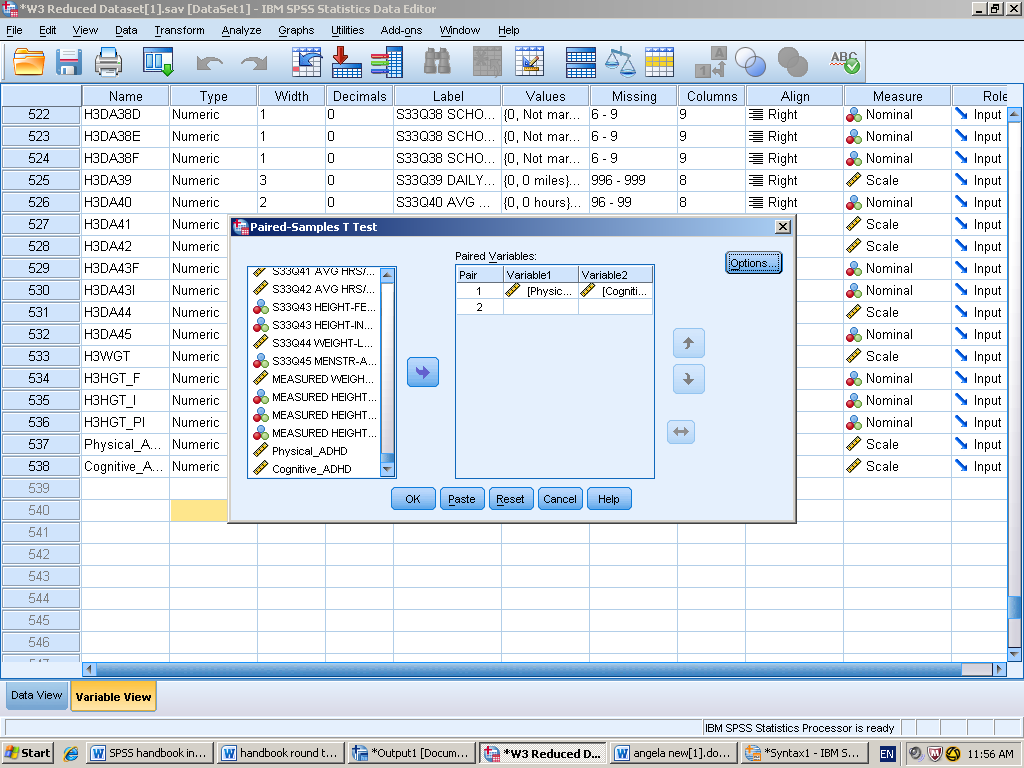
The dependent samples t-test is used when the population mean and the population standard deviation are unknown. This test requires a repeated measure or within subjects design. A within subjects design is when one population has 2 or more scores. The dependent samples t-test differs from the single sample t-test because we use the *difference* of scores. Also, assume the population mean (of the difference of scores) is zero.

**SPSS Steps**

1. *Analyze Compare means Paired Samples t-test*



1. Select the variable of interest and click the arrow to move it from the left hand list of variables to the right hand “Paired Variables” box. Then select the next variable of interest and it will be placed under the “variable 2” column. No options need to be selected.



1. Click paste to put into syntax, then run the analysis.

**Syntax**

T-TEST PAIRS=Physical\_ADHD WITH Cognitive\_ADHD (PAIRED)

/CRITERIA=CI(.9500)

/MISSING=ANALYSIS.

**Output interpretation**

| **Paired Samples Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | Physical\_ADHD | .8280 | 4787 | .57130 | .00826 |
| Cognitive\_ADHD | .7241 | 4787 | .54938 | .00794 |

| **Paired Samples Correlations** | | | | |
| --- | --- | --- | --- | --- |
|  | | N | Correlation | Sig. |
| Pair 1 | Physical\_ADHD & Cognitive\_ADHD | 4787 | .709 | .000 |

| **Paired Samples Test** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
| Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Pair 1 | Physical\_ADHD - Cognitive\_ADHD | .10389 | .42763 | .00618 | .09178 | .11601 | 16.809 | 4786 | .000 |

To test the prediction that individuals report more physical than cognitive symptoms regarding ADHD, I conducted a dependent- samples t-test on self- reports of allergy symptoms. The test was significant, *t*(4786) = 16.81 p <.001. In support of my hypothesis, physical symptoms (*M*= .8280, *SD*= .57130) reported more ADHD symptoms than cognitive symptoms (*M*= .7241. *SD*= .54931).

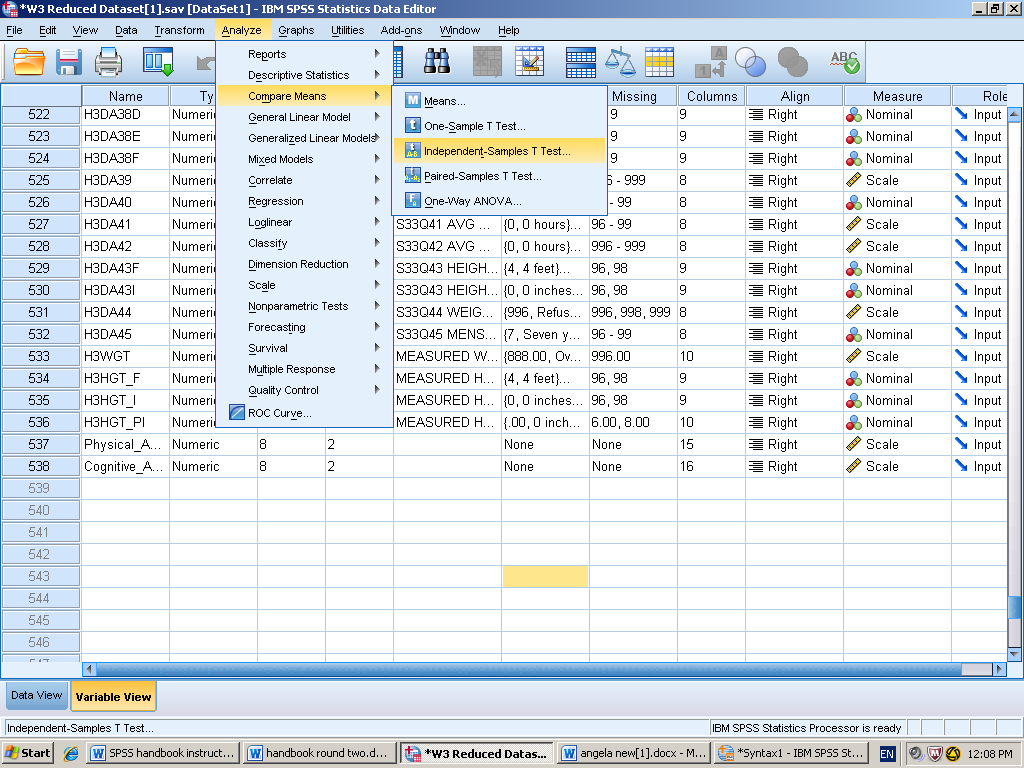
Independent Samples *t*-test

**Description of Statistics**

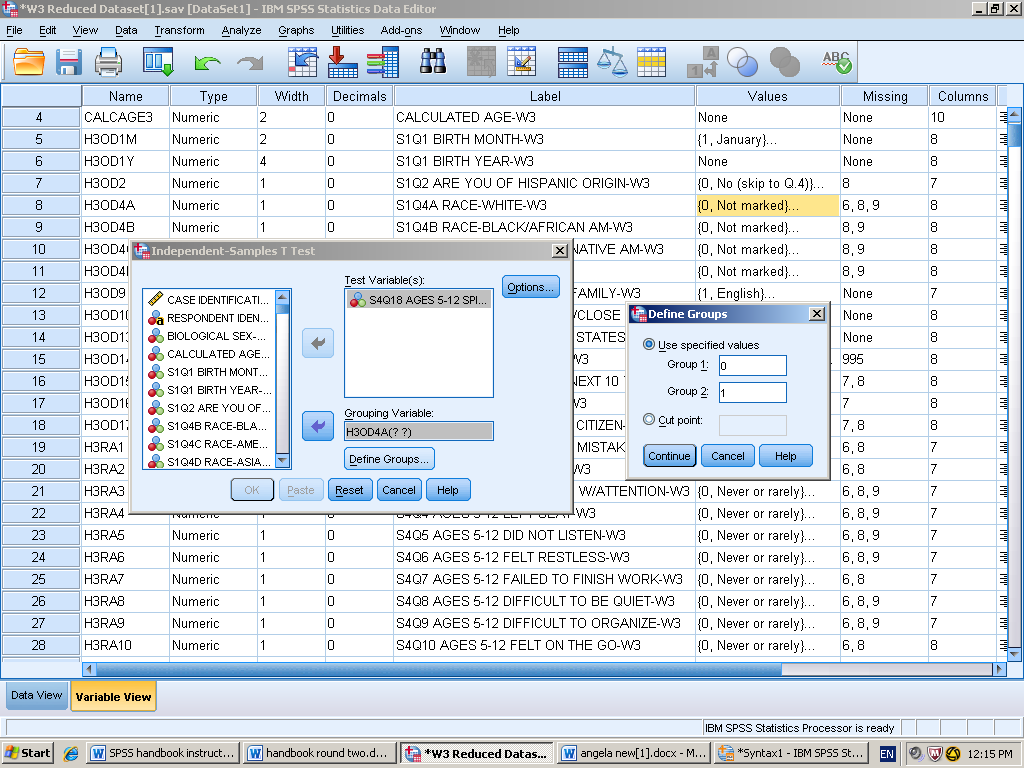
Independent Samples t-test is used when our sample mean information is *independent.* This tell is applied to situations in which we have a between subjects research design. The people we are comparing still come from 2 samples but the people are completely independent from each other. This test is used when we do not know population mean or population standard deviation information. The main difference between the independent and dependent is not the information you know but where your mean information comes from.

**SPSS Steps**

1. *Analyze Compare Means Independent Samples t-test*



1. Select the independent variable and click the arrow to move it from the left hand list of variables to the right hand “test variables” box. You can select more than one variable to test.
2. Select the dependent variable and click the arrow to move it from the left hand list of variables to the right hand “grouping variables” box. Next select define groups and press continue.



1. Click paste to put it in syntax, then run the analysis.

**Syntax**

T-TEST GROUPS=H3OD4A(0 1)

/MISSING=ANALYSIS

/VARIABLES=H3RA18

/CRITERIA=CI(.95).

**Output Inteperation**

| **Group Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | S1Q4A RACE-WHITE-W3 | N | Mean | Std. Deviation | Std. Error Mean |
| S4Q18 AGES 5-12 SPITEFUL/VINDICTIVE-W3 | Not marked | 1460 | .45 | .674 | .018 |
| Marked | 3354 | .44 | .700 | .012 |

| **Independent Samples Test** | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| S4Q18 AGES 5-12 SPITEFUL/VINDICTIVE-W3 | Equal variances assumed | .858 | .354 | .107 | 4812 | .915 | .002 | .022 | -.040 | .045 |
| Equal variances not assumed |  |  | .109 | 2874.139 | .913 | .002 | .021 | -.040 | .044 |

To test the prediction that the means of non-whites compared to the means of whites will report more vindictive or spiteful behavior I conducted an independent sample t-test on self-reports of ADHD behavior. The test was not significant t(4812)= 1.37, p=.982 in rejection of my hypothesis, individuals marked as “non-white” (*M*=.45 SD=.674) compared to individuals who marked “white” (*M*=.44, *SD*=,700.) on spiteful and vindictive.

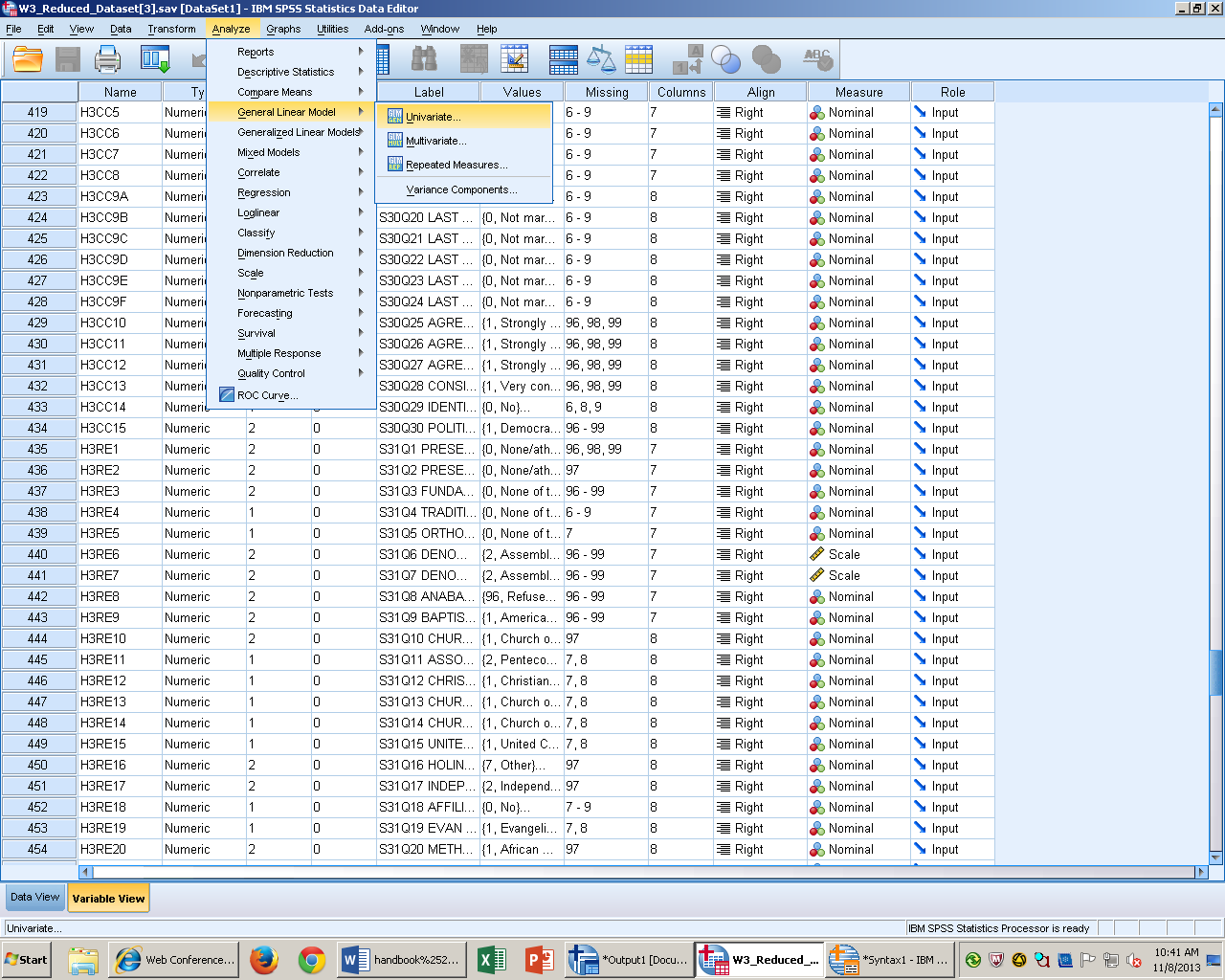
One-Way ANOVA with post hoc test

**Description of Statistic**

The one-way ANOVA is used for comparing more than 2 groups of scores, when the groups of people are entirely separate, when the dependent variable is continuous (interval or ratio scale), and when the independent variable is nominal (categorical). The null hypothesis always states that means of the populations are equal to each other. When you reject the null hypothesis this means at least one of the means are different from the others. We prefer one-way ANOVA instead of multiple t-tests because we want to avoid increasing our type one error. A post hoc test will compare each group to every other group. It is also how we follow up after a significant result. There are many post hoc test available. They vary in terms of how conservative they are or not. A tukey test is a very common and balances between the liberal and conservative tests.

**SPSS Steps**

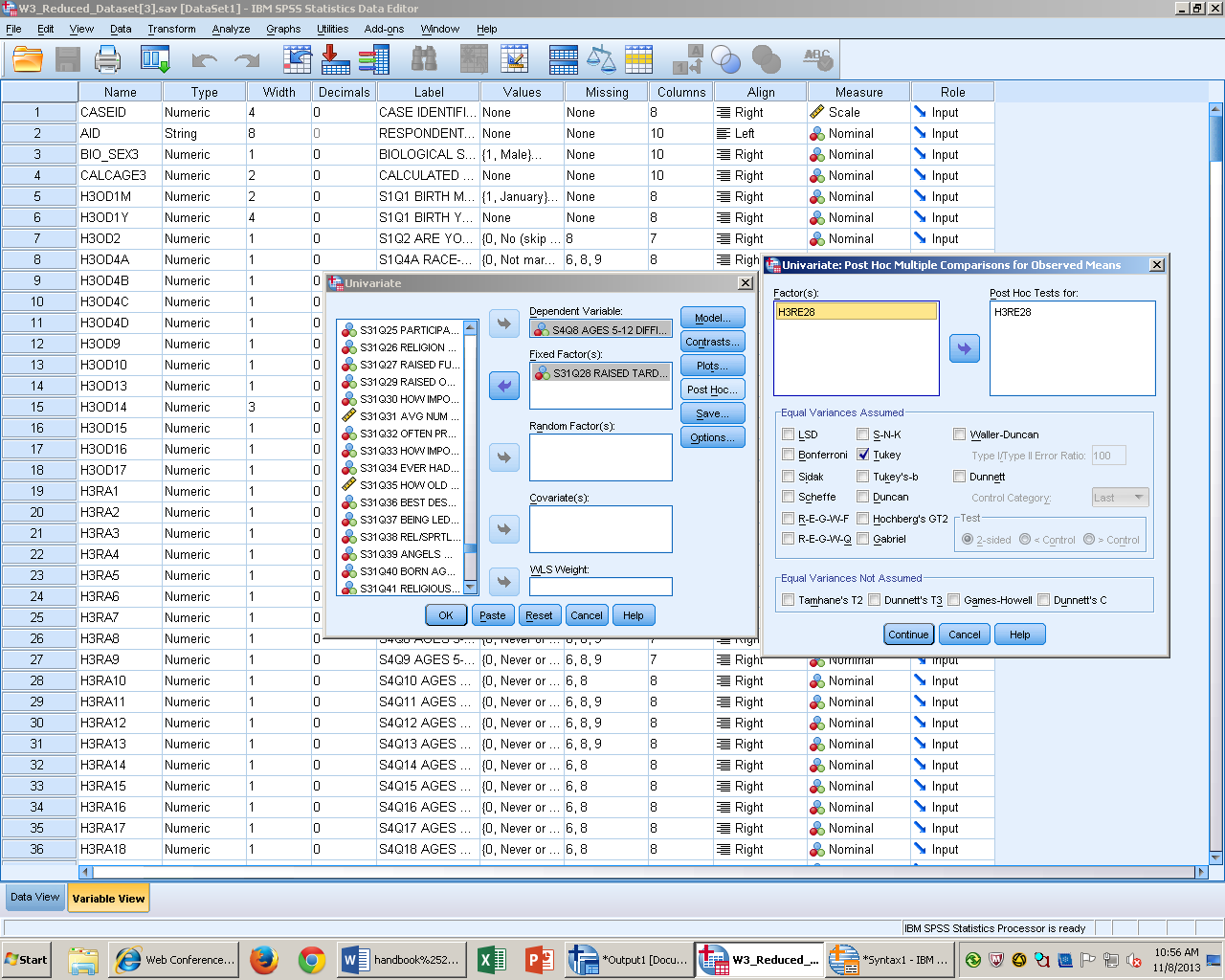
1. Analyze Compare Means Univariate



2. Add the dependent variable into the “Dependent Variable” box. Add the independent variable into “factor box.”

3. Select the blue POST HOC button. Click the box next to “Tukey.”

4. Click paste to put it in syntax, then run the analysis.



**Syntax**

ONEWAY H3OD9 BY H3RA14

/MISSING ANALYSIS

/POSTHOC=TUKEY ALPHA(0.05).

**Output Interpretation**

| **ANOVA** | | | | | |
| --- | --- | --- | --- | --- | --- |
| S4Q14 AGES 5-12 BLURTED OUT ANSWERS-W3 | | | | | |
|  | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 7.979 | 5 | 1.596 | 2.147 | .057 |
| Within Groups | 3610.645 | 4859 | .743 |  |  |
| Total | 3618.624 | 4864 |  |  |  |

| **Multiple Comparisons** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| S4Q14 AGES 5-12 BLURTED OUT ANSWERS-W3  Tukey HSD | | | | | | |
| (I) S1Q9 LANG USED MOST W/FAMILY-W3 | (J) S1Q9 LANG USED MOST W/FAMILY-W3 | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
| Lower Bound | Upper Bound |
| English | Spanish | .099 | .067 | .675 | -.09 | .29 |
| Another European language | -.464 | .352 | .775 | -1.47 | .54 |
| An Asian language | .316 | .140 | .214 | -.08 | .72 |
| Half English and half another language | .076 | .096 | .969 | -.20 | .35 |
| Other | .286 | .249 | .862 | -.42 | 1.00 |
| Spanish | English | -.099 | .067 | .675 | -.29 | .09 |
| Another European language | -.563 | .358 | .616 | -1.58 | .46 |
| An Asian language | .217 | .154 | .722 | -.22 | .66 |
| Half English and half another language | -.023 | .115 | 1.000 | -.35 | .31 |
| Other | .187 | .257 | .979 | -.55 | .92 |
| Another European language | English | .464 | .352 | .775 | -.54 | 1.47 |
| Spanish | .563 | .358 | .616 | -.46 | 1.58 |
| An Asian language | .781 | .379 | .308 | -.30 | 1.86 |
| Half English and half another language | .541 | .365 | .675 | -.50 | 1.58 |
| Other | .750 | .431 | .505 | -.48 | 1.98 |
| An Asian language | English | -.316 | .140 | .214 | -.72 | .08 |
| Spanish | -.217 | .154 | .722 | -.66 | .22 |
| Another European language | -.781 | .379 | .308 | -1.86 | .30 |
| Half English and half another language | -.240 | .169 | .715 | -.72 | .24 |
| Other | -.031 | .285 | 1.000 | -.84 | .78 |
| Half English and half another language | English | -.076 | .096 | .969 | -.35 | .20 |
| Spanish | .023 | .115 | 1.000 | -.31 | .35 |
| Another European language | -.541 | .365 | .675 | -1.58 | .50 |
| An Asian language | .240 | .169 | .715 | -.24 | .72 |
| Other | .209 | .266 | .970 | -.55 | .97 |
| Other | English | -.286 | .249 | .862 | -1.00 | .42 |
| Spanish | -.187 | .257 | .979 | -.92 | .55 |
| Another European language | -.750 | .431 | .505 | -1.98 | .48 |
| An Asian language | .031 | .285 | 1.000 | -.78 | .84 |
| Half English and half another language | -.209 | .266 | .970 | -.97 | .55 |

| **S4Q14 AGES 5-12 BLURTED OUT ANSWERS-W3** | | | |
| --- | --- | --- | --- |
| Tukey HSDa,b | | | |
| S1Q9 LANG USED MOST W/FAMILY-W3 | N | Subset for alpha = 0.05 | |
| 1 | 2 |
| An Asian language | 38 | .55 |  |
| Other | 12 | .58 | .58 |
| Spanish | 174 | .77 | .77 |
| Half English and half another language | 82 | .79 | .79 |
| English | 4553 | .87 | .87 |
| Another European language | 6 |  | 1.33 |
| Sig. |  | .851 | .061 |

I compared six languages on their means of blurting out answers, using a one-way ANOVA with post hoc, test was not significant *F*(4,5489) = 2.417, *p=.*057.I conducted a post hoc test and found none of the languages spoken were significantly different from each other.

This is test is problematic because my independent variables all have very different sample sizes which could lead to my results being inaccurate.

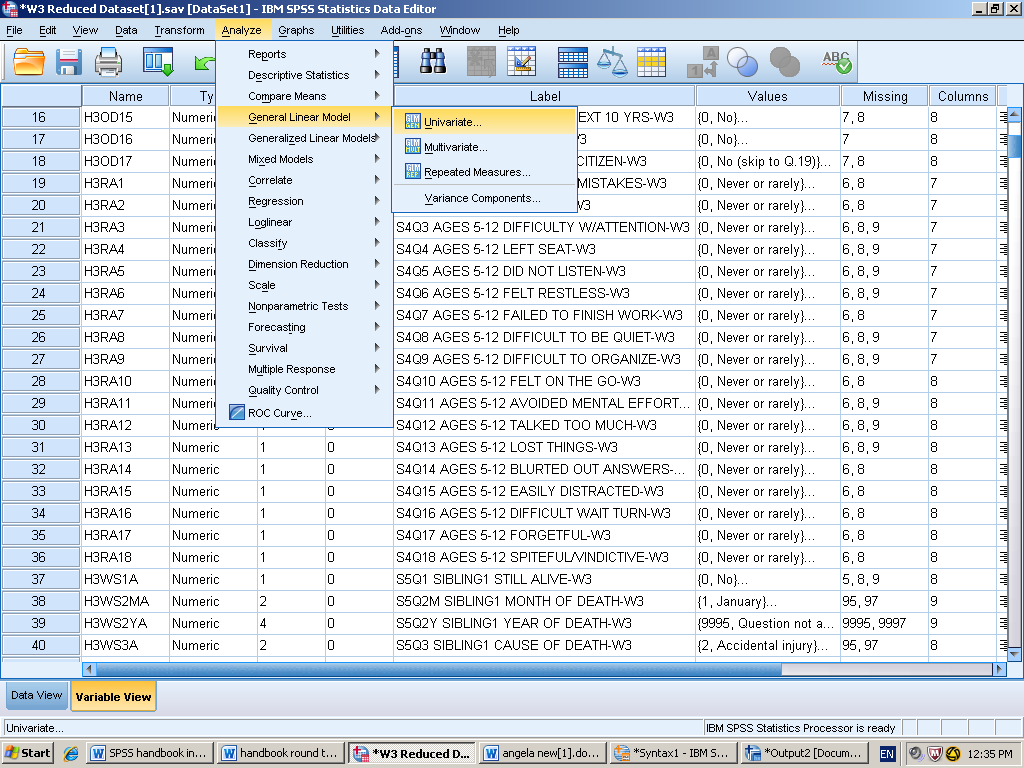
Factorial (2-Way) ANOVA

**Description of Statistics**

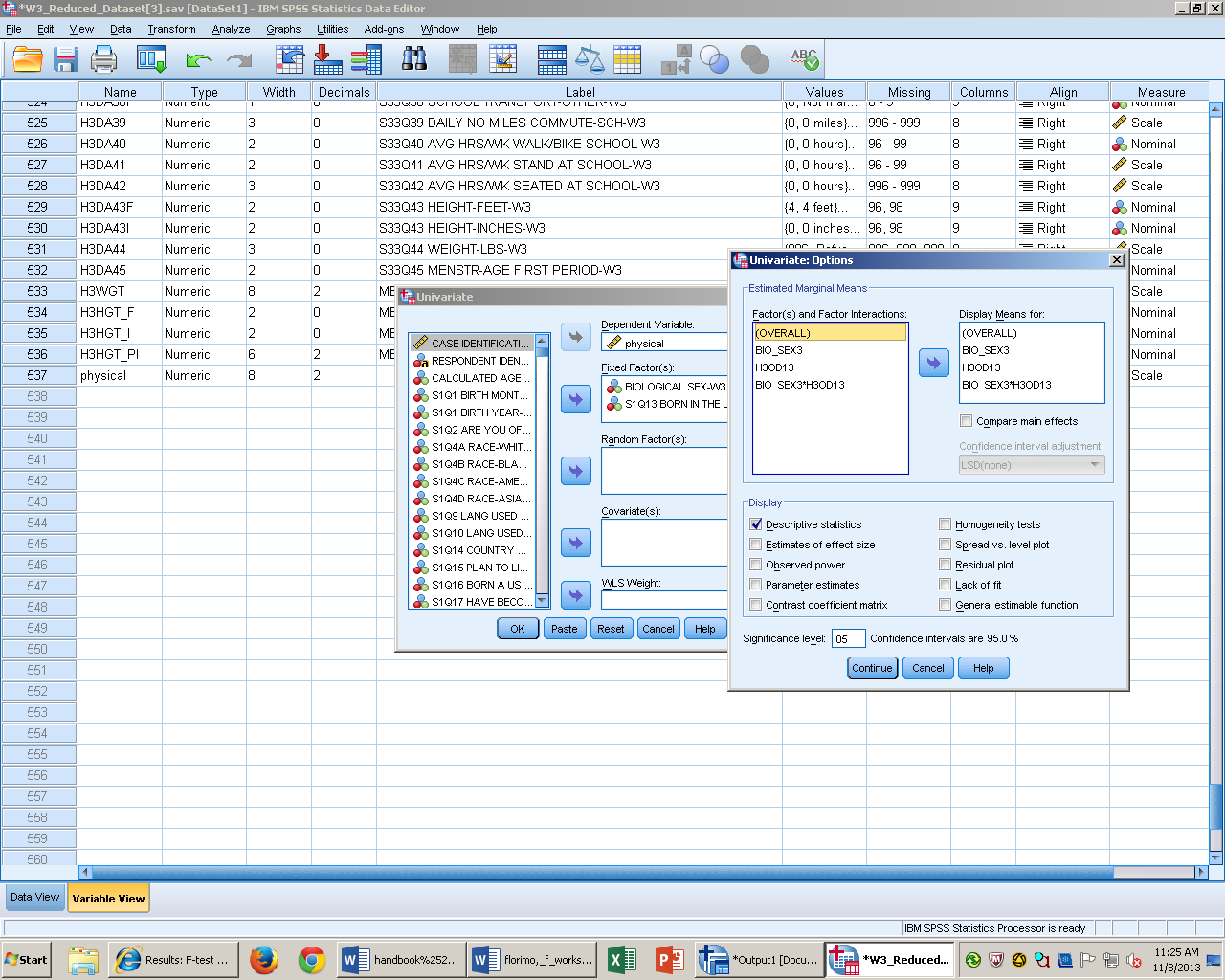
A factorial (2-way) ANOVA is used when we 2 or more independent variables that each have 2 or more levels combine and influence the dependent variable. We can describe our factorial research design using numbers. A shorthand or nomenclature design 3 X 2 is explained by the number of numbers is the number of independent variables (2 independent variables) and the value of the number defines the number of levels (independent variable number 1 has three levels) can represent this is any order. A main effect is an effect of an independent variable on a dependent variable. Each independent variable has one main effect. An interaction is whenever the effect of one independent variable *depends* on the level of another.

**SPSS Steps**

1. *Analyze General Linear Model Univariate*



1. Enter the dependent variable into the “Dependent variable” box. Select the independent variables and click the arrow to move them into the “fixed factor” box.
2. Select the blue option button the left side and select “descriptive statistics.” Then select continue. You would select a post hoc test if had more than 2 levels.



1. Click paste to put into syntax, and then run the analysis.

**Syntax**

UNIANOVA physical BY BIO\_SEX3 H3OD13

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/POSTHOC=BIO\_SEX3 H3OD13(TUKEY)

/EMMEANS=TABLES(OVERALL)

/EMMEANS=TABLES(BIO\_SEX3)

/EMMEANS=TABLES(H3OD13)

/EMMEANS=TABLES(BIO\_SEX3\*H3OD13)

/PRINT=DESCRIPTIVE

/CRITERIA=ALPHA(.05)

/DESIGN=BIO\_SEX3 H3OD13 BIO\_SEX3\*H3OD13.

**Output Interpretation**

| **Between-Subjects Factors** | | | |
| --- | --- | --- | --- |
|  | | Value Label | N |
| BIOLOGICAL SEX-W3 | 1 | Male | 2224 |
| 2 | Female | 2597 |
| S1Q13 BORN IN THE UNITED STATES-W3 | 0 | No | 259 |
| 1 | Yes (skip to Q.20) | 4562 |

| **Descriptive Statistics** | | | | |
| --- | --- | --- | --- | --- |
| Dependent Variable:physical | | | | |
| BIOLOGICAL SEX-W3 | S1Q13 BORN IN THE UNITED STATES-W3 | Mean | Std. Deviation | N |
| Male | No | .8900 | .59178 | 125 |
| Yes (skip to Q.20) | .9846 | .64050 | 2099 |
| Total | .9793 | .63811 | 2224 |
| Female | No | .7481 | .54969 | 134 |
| Yes (skip to Q.20) | .8634 | .61251 | 2463 |
| Total | .8574 | .60986 | 2597 |
| Total | No | .8166 | .57369 | 259 |
| Yes (skip to Q.20) | .9192 | .62839 | 4562 |
| Total | .9137 | .62594 | 4821 |

| **Tests of Between-Subjects Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| Dependent Variable:physical | | | | | |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
| Corrected Model | 20.542a | 3 | 6.847 | 17.658 | .000 |
| Intercept | 743.539 | 1 | 743.539 | 1917.405 | .000 |
| BIO\_SEX3 | 4.236 | 1 | 4.236 | 10.923 | .001 |
| H3OD13 | 2.695 | 1 | 2.695 | 6.950 | .008 |
| BIO\_SEX3 \* H3OD13 | .026 | 1 | .026 | .067 | .796 |
| Error | 1867.956 | 4817 | .388 |  |  |
| Total | 5912.938 | 4821 |  |  |  |
| Corrected Total | 1888.498 | 4820 |  |  |  |
| a. R Squared = .011 (Adjusted R Squared = .010) | | | | | |

| **1. Grand Mean** | | | |
| --- | --- | --- | --- |
| Dependent Variable:physical | | | |
| Mean | Std. Error | 95% Confidence Interval | |
| Lower Bound | Upper Bound |
| .872 | .020 | .833 | .911 |

| **2. BIOLOGICAL SEX-W3** | | | | |
| --- | --- | --- | --- | --- |
| Dependent Variable:physical | | | | |
| BIOLOGICAL SEX-W3 | Mean | Std. Error | 95% Confidence Interval | |
| Lower Bound | Upper Bound |
| Male | .937 | .029 | .881 | .994 |
| Female | .806 | .028 | .752 | .860 |

| **3. S1Q13 BORN IN THE UNITED STATES-W3** | | | | |
| --- | --- | --- | --- | --- |
| Dependent Variable:physical | | | | |
| S1Q13 BORN IN THE UNITED STATES-W3 | Mean | Std. Error | 95% Confidence Interval | |
| Lower Bound | Upper Bound |
| No | .819 | .039 | .743 | .895 |
| Yes (skip to Q.20) | .924 | .009 | .906 | .942 |

| **4. BIOLOGICAL SEX-W3 \* S1Q13 BORN IN THE UNITED STATES-W3** | | | | | |
| --- | --- | --- | --- | --- | --- |
| Dependent Variable:physical | | | | | |
| BIOLOGICAL SEX-W3 | S1Q13 BORN IN THE UNITED STATES-W3 | Mean | Std. Error | 95% Confidence Interval | |
| Lower Bound | Upper Bound |
| Male | No | .890 | .056 | .781 | .999 |
| Yes (skip to Q.20) | .985 | .014 | .958 | 1.011 |
| Female | No | .748 | .054 | .643 | .854 |
| Yes (skip to Q.20) | .863 | .013 | .839 | .888 |

To test the prediction that males experience more physical symptoms of ADHD than females, but only if the individual was female, I conducted a 2(gender: male vs. female) x 2(birth place: born in U.S vs. not born in the U.S.) factorial ANOVA on self-reports of physical ADHD symptoms. The main effect of birth place was significant, *F* (1, 4817) = 6.950, *p*= .008. Individuals born in the United States (*M* =.92, *SD*= .009) reported more physical symptoms than individuals not born in the United States( *M*= .82, *SD* = .04). The main effect of gender was also significant, *p* = .001. Males (*M*= .99 *SD* = .014 ) reported more physical symptoms than females (*M* = .86 *SD* =.013) The test of the interaction was not significant, *F* (1, 4817) = .067, *p*= .796.

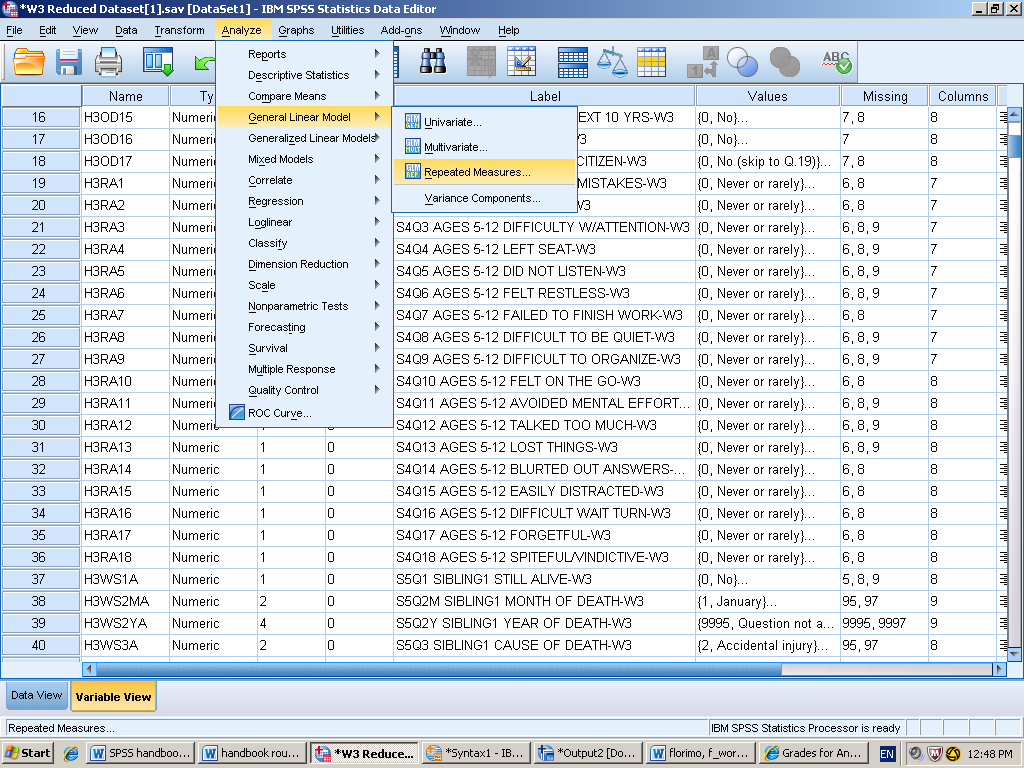
Repeated Measures ANOVA

**Description of Statistic**

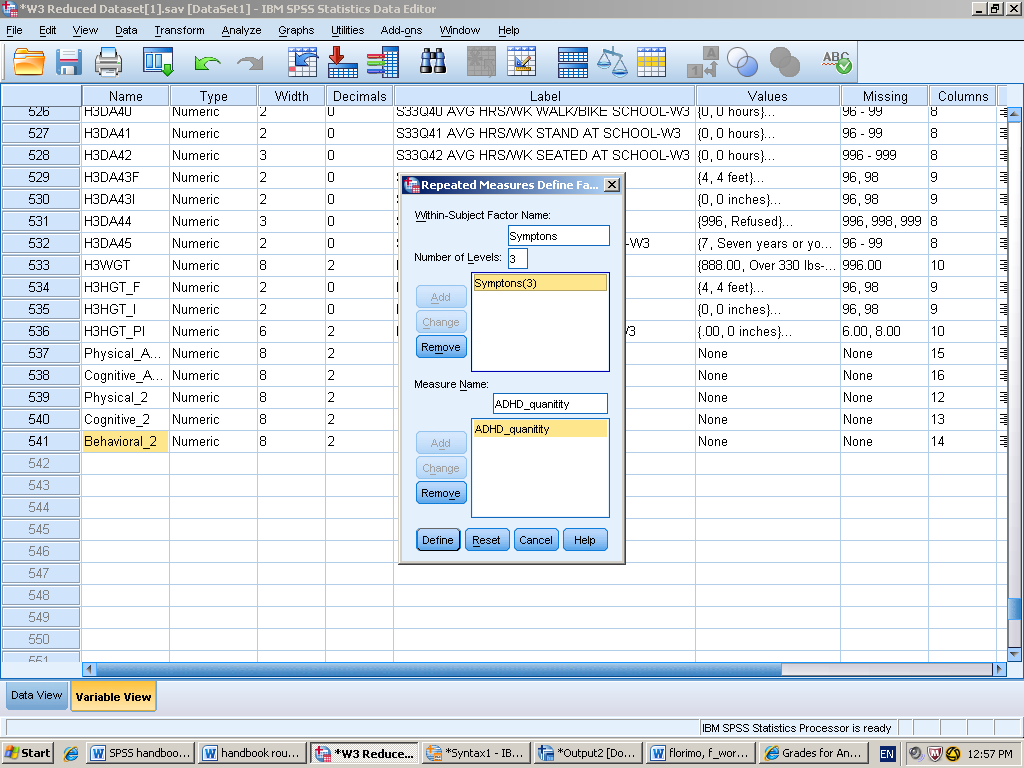
The repeated measure ANOVA is used when we have the same participants in all (2+) conditions of the independent variable. The repeated measures ANOVA allows us to add in another independent variable compared to the dependent t-test. Another way to use the repeated measures ANOVA design is to use *time* as the independent variable. This test can be used for completely different activities but with everyone doing the actives or for the same activity but over time in both cases the means are linked to each other. A significant F tells us that we probably have samples coming from the same population with different means. However we are not sure where the significant difference is all it tells us is we have at least one mean different than the other means.

**SPSS Steps**

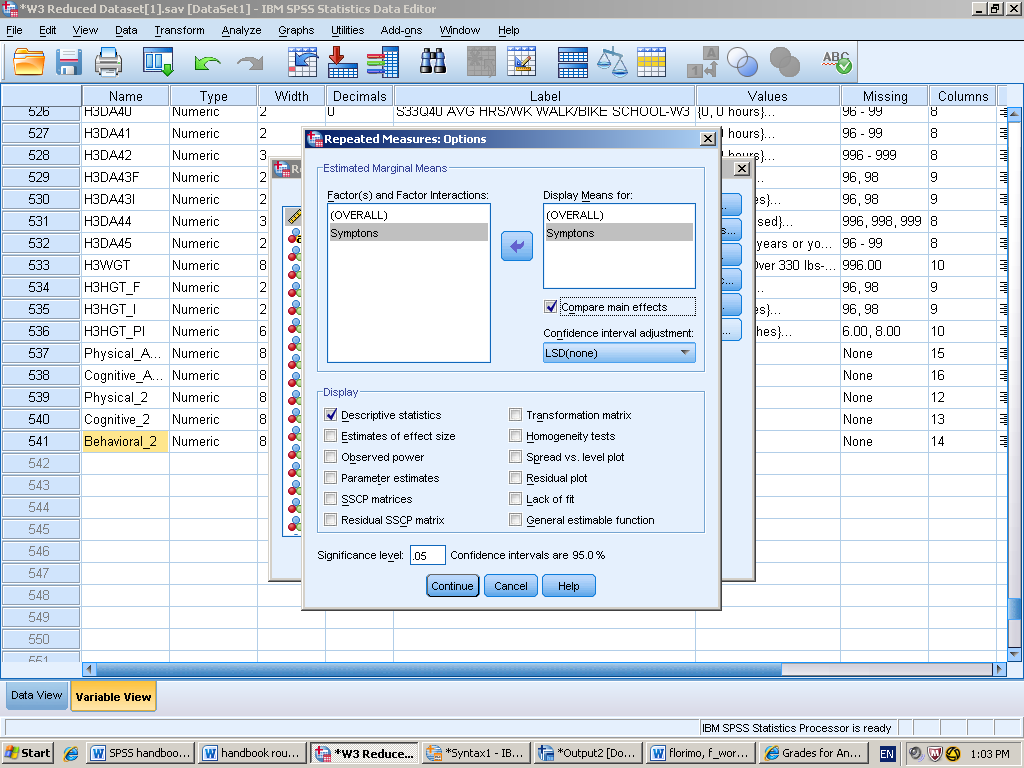
1. *Analyze General Linear Model Repeated Measure*



1. Identify the within-subjects factor and define it. This is the independent Variable. Also, define the amount of levels. A level is the condition of the IV. Also define the independent variable in the “Measure Name” box. Then select define.



1. Select the variables you wish to enter as your levels for the within subjects design box.
2. Under “options” you select compare main effects, description statistics, and activity means.



1. Click paste to put it in syntax, then run the analysis.

**Syntax**

GLM Physical\_2 Cognitive\_2 Behavioral\_2

/WSFACTOR=Symptons 3 Polynomial

/MEASURE=ADHD\_quanitity

/METHOD=SSTYPE(3)

/EMMEANS=TABLES(OVERALL)

/EMMEANS=TABLES(Symptons) COMPARE ADJ(LSD)

/PRINT=DESCRIPTIVE

/CRITERIA=ALPHA(.05)

/WSDESIGN=Symptons.

**Output interpretation**

| **Within-Subjects Factors** | |
| --- | --- |
| Measure:ADHD\_quanitity | |
| Symptons | Dependent Variable |
| 1 | Physical\_2 |
| 2 | Cognitive\_2 |
| 3 | Behavioral\_2 |
|  |  |

| **Tests of Within-Subjects Effects** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Measure:ADHD\_quanitity | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
| Symptons | Sphericity Assumed | 267.990 | 2 | 133.995 | 1074.273 | .000 |
| Greenhouse-Geisser | 267.990 | 1.798 | 149.072 | 1074.273 | .000 |
| Huynh-Feldt | 267.990 | 1.798 | 149.019 | 1074.273 | .000 |
| Lower-bound | 267.990 | 1.000 | 267.990 | 1074.273 | .000 |
| Error(Symptons) | Sphericity Assumed | 1191.677 | 9554 | .125 |  |  |
| Greenhouse-Geisser | 1191.677 | 8587.726 | .139 |  |  |
| Huynh-Feldt | 1191.677 | 8590.757 | .139 |  |  |
| Lower-bound | 1191.677 | 4777.000 | .249 |  |  |

| **Pairwise Comparisons** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Measure:ADHD\_quanitity | | | | | | |
| (I) Symptons | (J) Symptons | Mean Difference (I-J) | Std. Error | Sig.a | 95% Confidence Interval for Differencea | |
| Lower Bound | Upper Bound |
| 1 | 2 | .335\* | .008 | .000 | .319 | .351 |
| 3 | .179\* | .007 | .000 | .165 | .193 |
| 2 | 1 | -.335\* | .008 | .000 | -.351 | -.319 |
| 3 | -.156\* | .006 | .000 | -.168 | -.144 |
| 3 | 1 | -.179\* | .007 | .000 | -.193 | -.165 |
| 2 | .156\* | .006 | .000 | .144 | .168 |
| Based on estimated marginal means | | | | | | |
| \*. The mean difference is significant at the .05 level.  a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments). | | | | | | |

We conducted a repeated measure ANOVA to compare three means on retrospective ADHD. The ANOVA was significant *F* (2,9554) = 1074.27, *p* = <.005. I conducted a post hoc test *p*= <.05 and found cognitive symptoms (*M*= .58 *SD*= .57), physical symptoms, (*M*= .91, *SD*=.63) and behavioral symptoms (*M*= .74 *SD*=.57 ) were all significantly different from each other on retrospective ADHD.

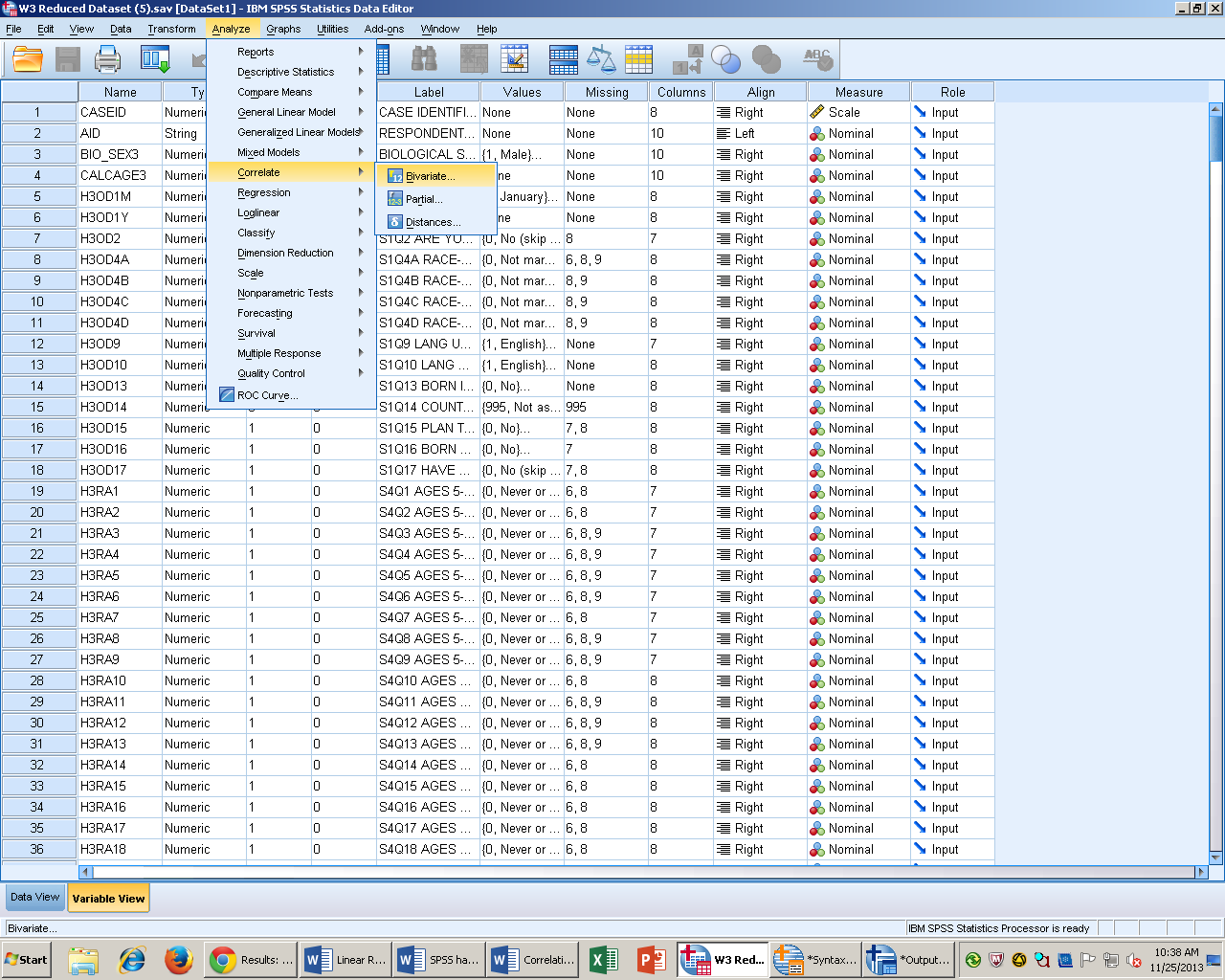
Correlation

**Descriptive Statistics**

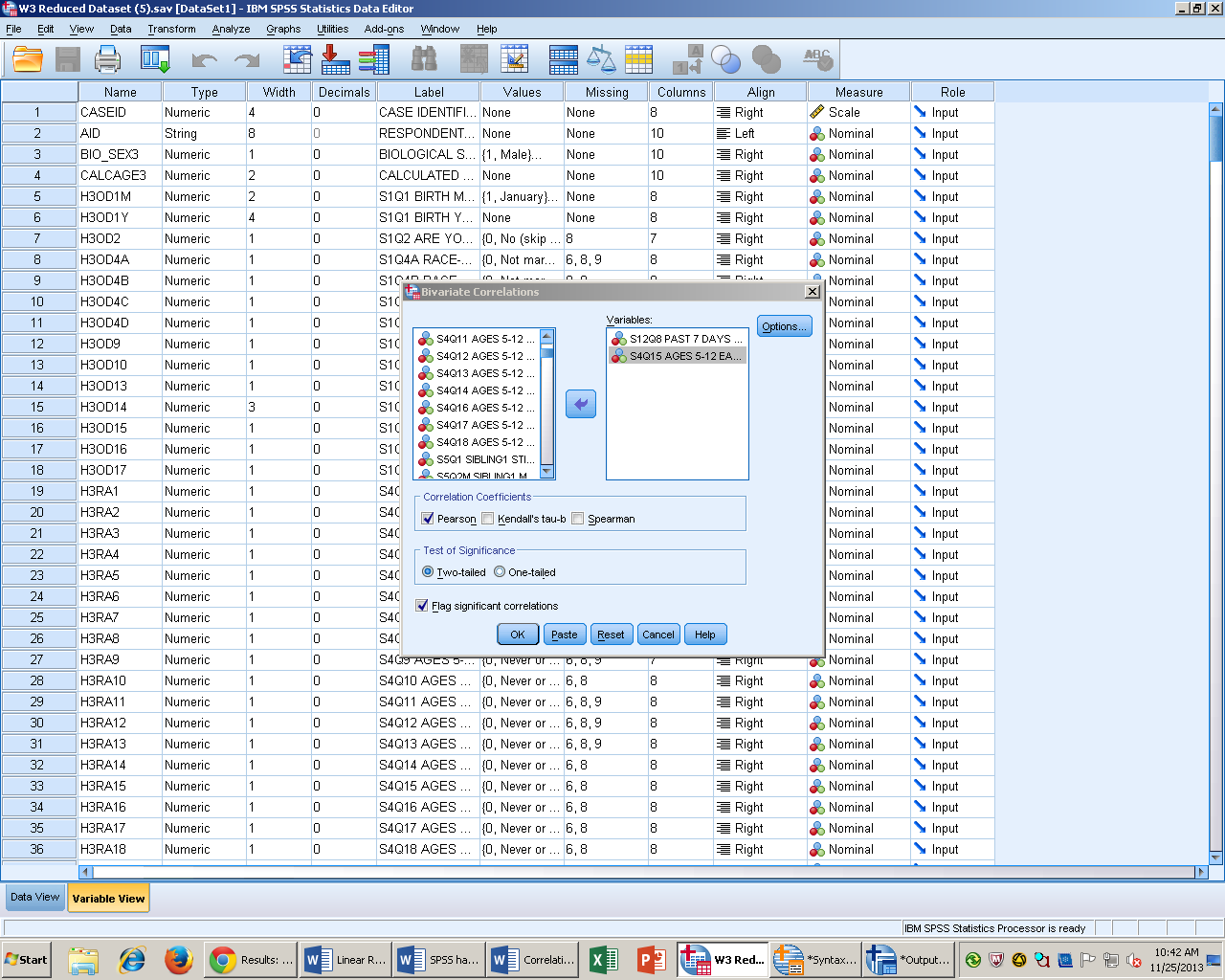
A correlation is the association between two variables scores. The requirements needed for a correlation are two and only two variables and equal interval variables on the ratio or interval scale. A scatterplot represents every data point and can show us a relationship or an association between two variables. A correlation coefficient is a numerical representation of the direction and strength of the relationship. Before examining the correlation coefficient, you should review the scatterplot because if you do not have a linear relationship SPSS will calculate a correlation coefficient and the results will be misleading. There are two factors that are also included in the deviation score is the number of people and variation of the scores. The statistic can range from 0 to 1. The sign, tells you the direction of the linear correlation either positive or negative. The number tells you the strength of the liner correlation, ranges from 0 to 1.00. The correlation troublemakers are outliers, full range, and curvilinear.

**SPSS Steps**

1. Analyze Correlation Bivariate



1. Select the variable of interest and click the arrow to move it from the left hand list of variables to the right hand “variables” box. It doesn’t matter the order in which you select your independent or dependent variable. No options need to be selected.



1. Click paste to put into syntax then run the analysis.

**Syntax**

CORRELATIONS

/VARIABLES=H3SP8 H3RA15

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

**Output Interpretation**

| **Correlations** | | | |
| --- | --- | --- | --- |
|  | | S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | S4Q15 AGES 5-12 EASILY DISTRACTED-W3 |
| S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | Pearson Correlation | 1 | .224\*\* |
| Sig. (2-tailed) |  | .000 |
| N | 4876 | 4861 |
| S4Q15 AGES 5-12 EASILY DISTRACTED-W3 | Pearson Correlation | .224\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| N | 4861 | 4864 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

To measure the correlation of having trouble keeping your mind on what you were doing in this past seven days is related to when you were between the ages of 5 and 12 if you were easily distracted. I reject the null hypothesis. My results are significant *r* (4861) = .224, *p* <.001. Having trouble keeping your mind on what you were doing in the past seven days is related to when you were between the ages of 5 and 12 you were easily distracted.

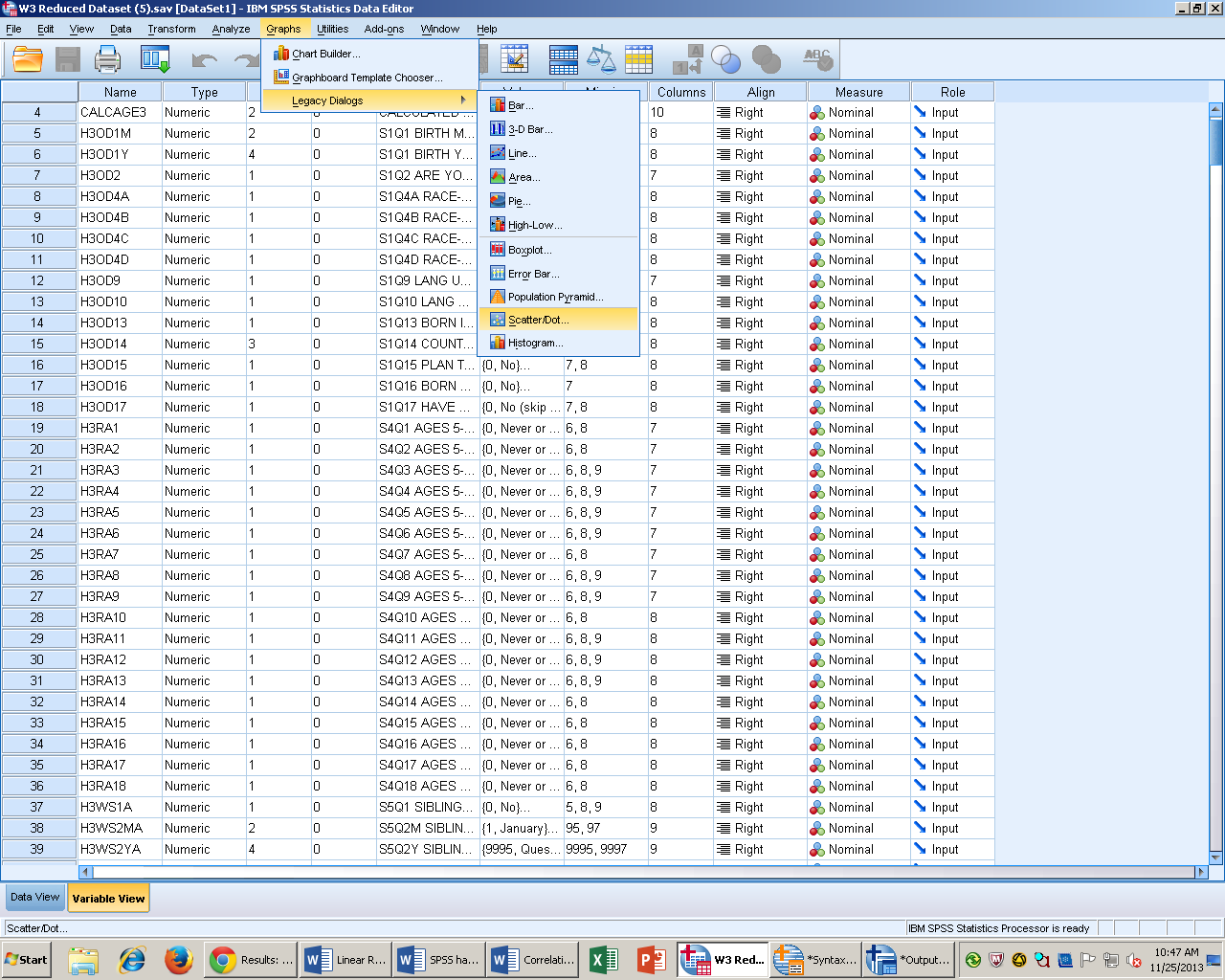
Scatterplot with Regression Line

**Descriptive Statistics**

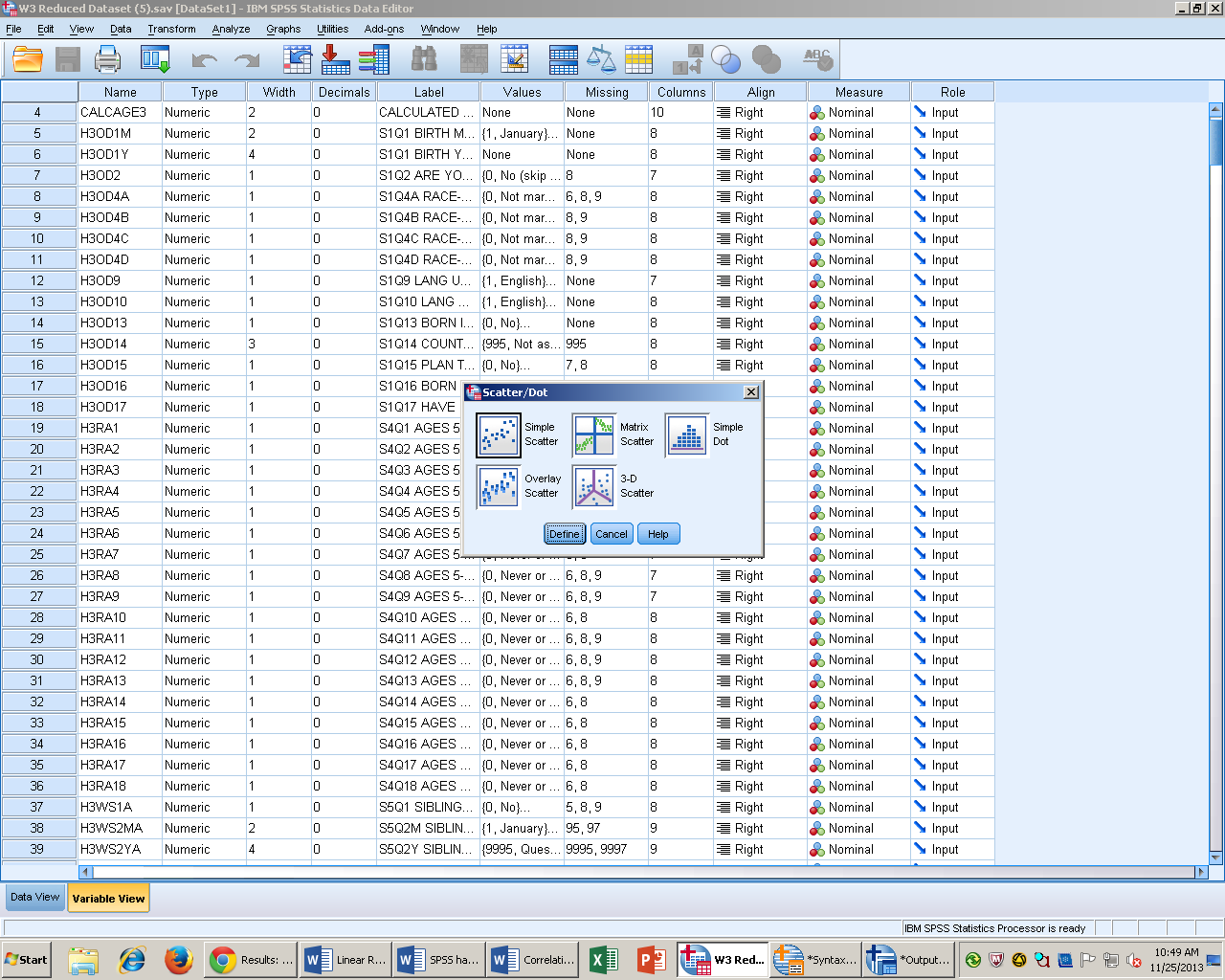
A scatterplot represents every data point and can show a relationship and/or an association between variables. Scatterplots can tell us about the pattern and strength of a correlation. The patterns of correlation can be defined as linear, curvilinear, or no pattern. A linear correlation follows close to a straight line. It doesn’t matter which variable goes on which axis because we do not have true independent or dependent variables. Positive and negative can also be analyzed on a scatterplot. Positive correlation as variables increase one variable, increases on the next. Negative correlation as variables increase on one variable they decrease on the next. The strength of the correlation can be defined as perfect, strong, or weak.

**SPSS Steps**

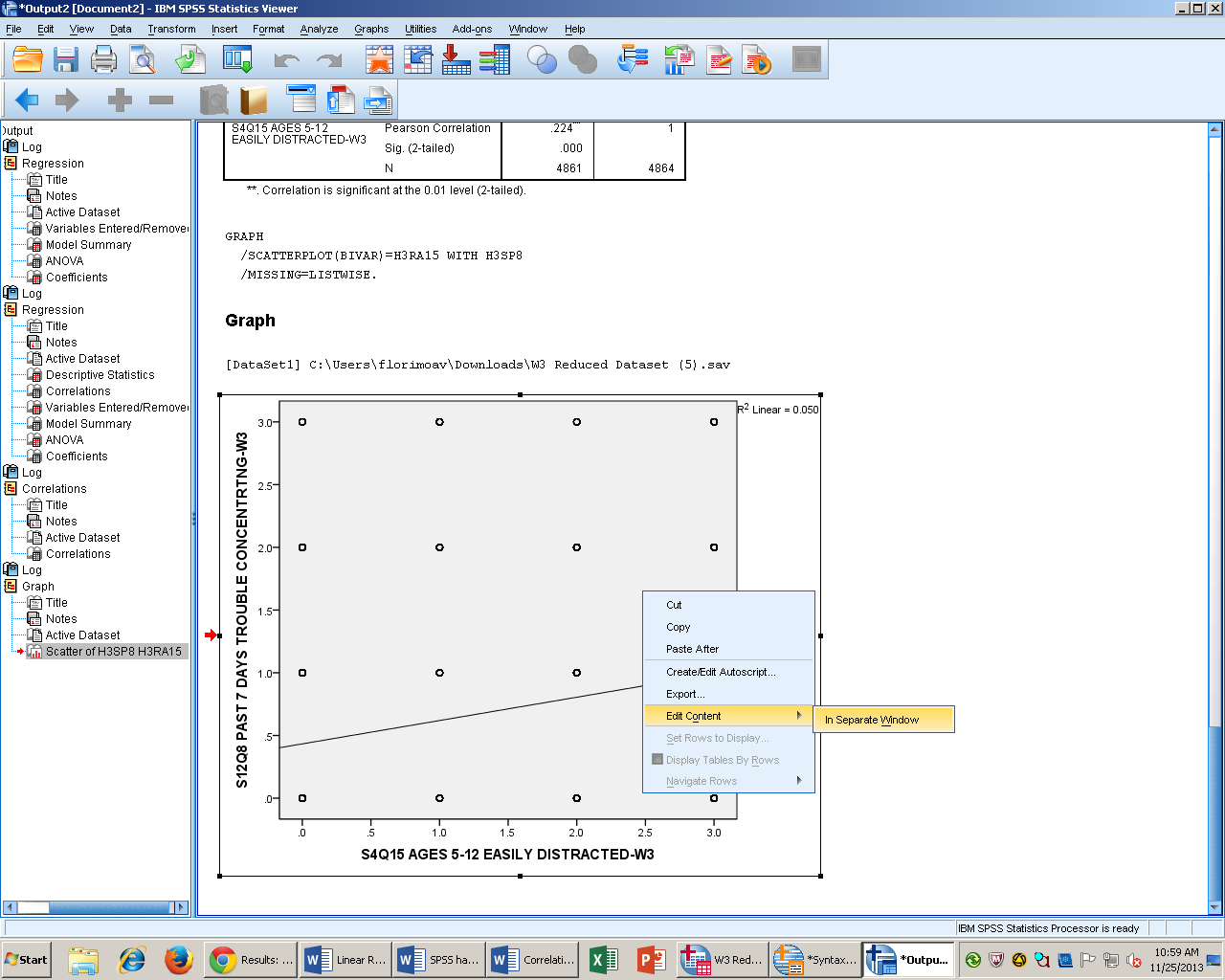
1. Graphs Legacy Dialogues Scatter/Dot



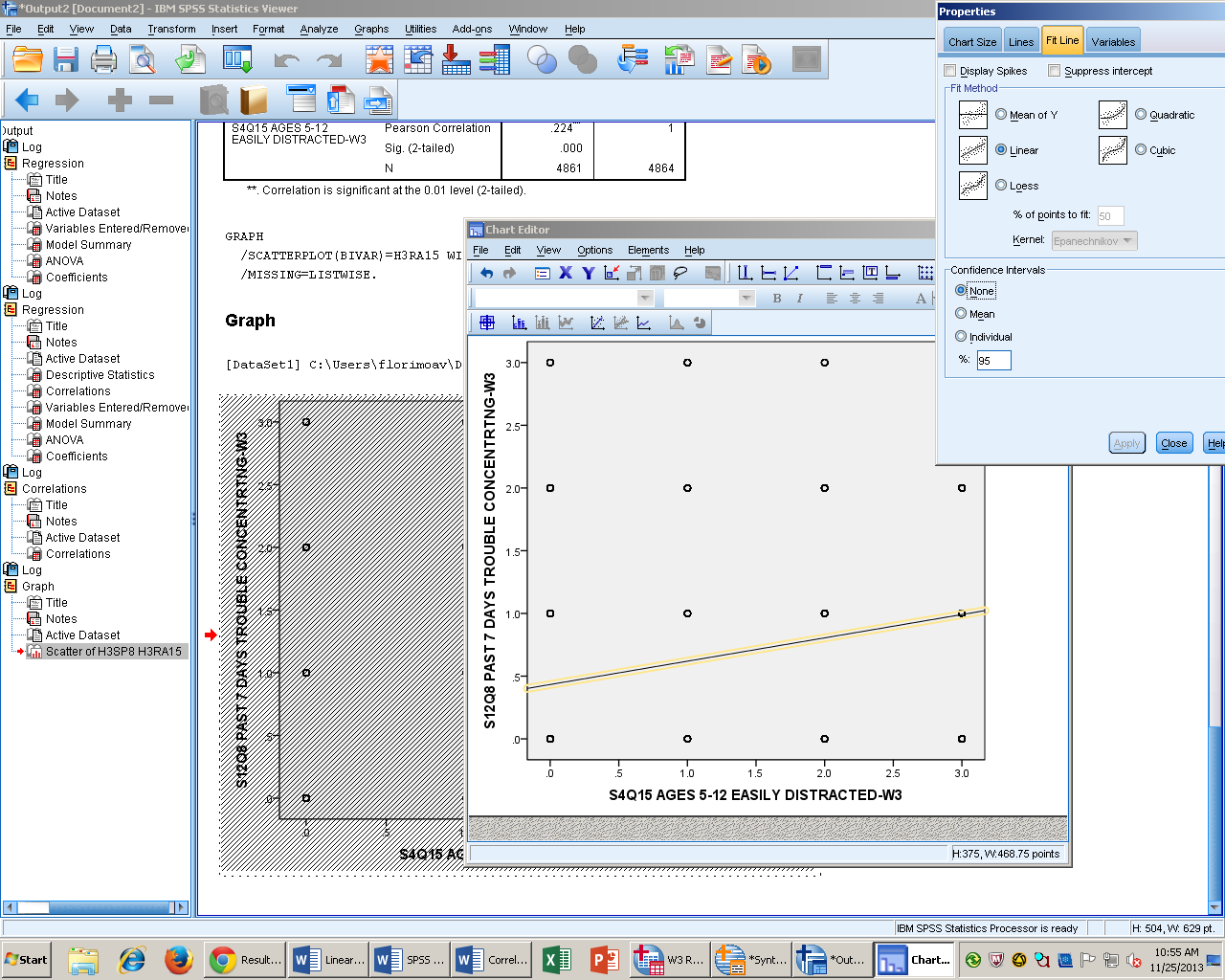
1. Select Simple Scatter and then click define.



1. Select the dependent variable of interest and the independent variable of interest and move them into the right column.
2. The put data into syntax, and run the analysis.
3. Right click on the scatterplot and then click edit content in separate window.



1. Then click “Add Fit line.” Then click close.



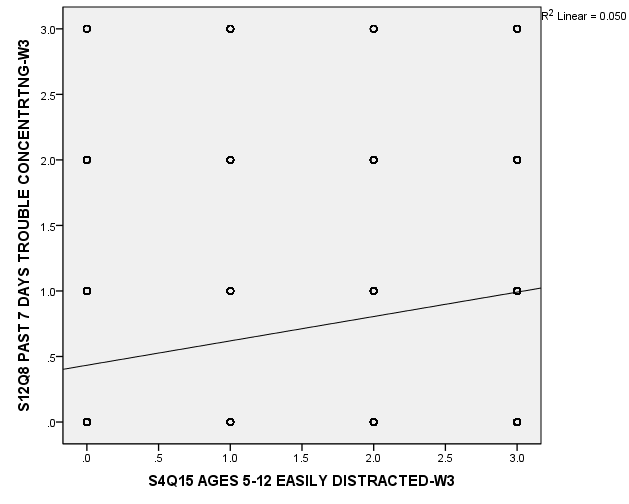
**Syntax**

GRAPH

/SCATTERPLOT(BIVAR)=H3RA15 WITH H3SP8

/MISSING=LISTWISE.

**Output Interpretation**



My variables appear to have a relationship of weak strength and positive correlation. I did not analyze any outliers or full range troublemakers.

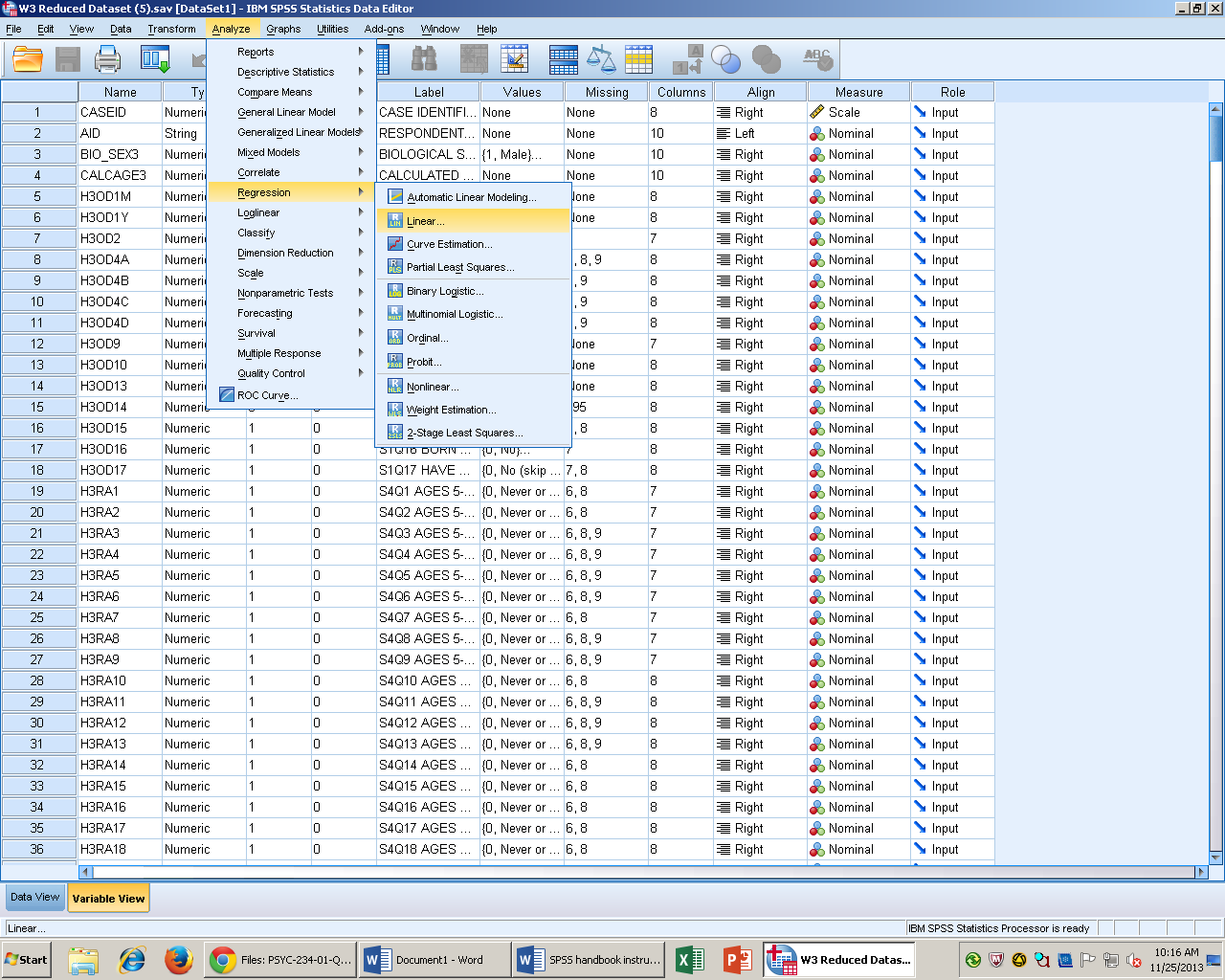
Linear Regression

**Description of Statistics**

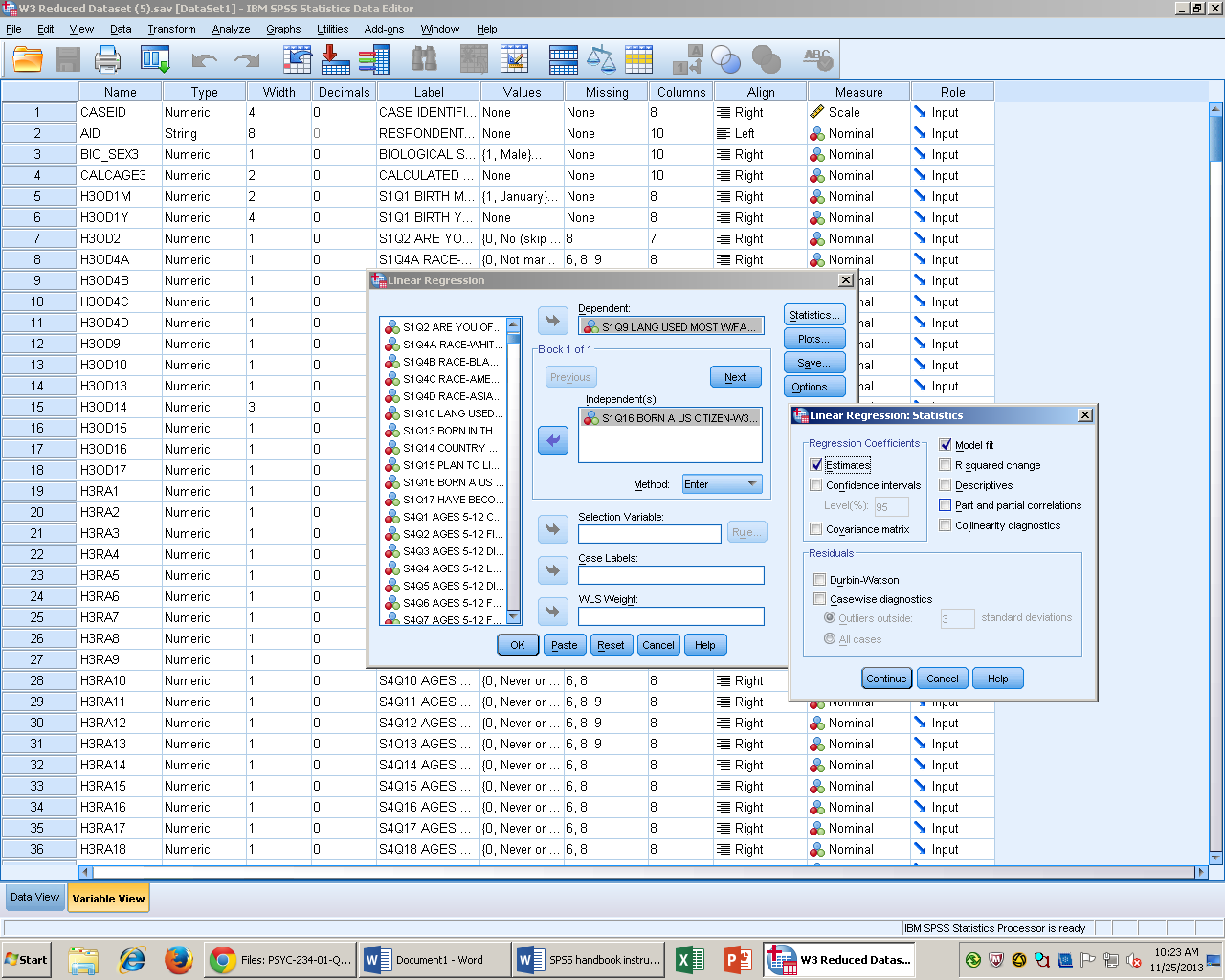
Correlation is a building block for regression. Regression is making predications about one variable, based on information from a second variable. *We cannot conduct a regression analysis if our variables are not correlated.* In regression, the independent variable is called the predictor variable whereas the dependent variable, Y is called the criterion variable. In order to predict a persons score on Y; you must start with a baseline number (a). Then add it to the result of the multiplying a special predictor variable (b) by the persons score on (x). Y = a + b (x). The slope of regression line is the amount the line moves up for every unity of X it moves across. The intercept is the predicated score on Y when X is zero. Least Squares Criterion goal is to minimize the amount of error.

**SPSS Steps**

1. Analyze Regression Linear Regression



1. Select the dependent variable, or the criterion variable and click the arrow to move it to the “dependent test box.” Select the independent variable, or the predictor variable and move it into the “independent test box.”
2. Then select the blue statistics box and select descriptive statistics. Model Fit and Estimates are clicked by default.



1. Click paste to put it in syntax, then run the analysis.

**Syntax**

DATASET ACTIVATE DataSet1.

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT H3SP8

/METHOD=ENTER H3RA5.

**Output Interpretation**

| **Descriptive Statistics** | | | |
| --- | --- | --- | --- |
|  | Mean | Std. Deviation | N |
| S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | .62 | .747 | 4861 |
| S4Q5 AGES 5-12 DID NOT LISTEN-W3 | .54 | .726 | 4861 |

| **Correlations** | | | |
| --- | --- | --- | --- |
|  | | S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | S4Q5 AGES 5-12 DID NOT LISTEN-W3 |
| Pearson Correlation | S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | 1.000 | .151 |
| S4Q5 AGES 5-12 DID NOT LISTEN-W3 | .151 | 1.000 |
| Sig. (1-tailed) | S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | . | .000 |
| S4Q5 AGES 5-12 DID NOT LISTEN-W3 | .000 | . |
| N | S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | 4861 | 4861 |
| S4Q5 AGES 5-12 DID NOT LISTEN-W3 | 4861 | 4861 |

| **Model Summary** | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .151a | .023 | .023 | .739 |
| a. Predictors: (Constant), S4Q5 AGES 5-12 DID NOT LISTEN-W3 | | | | |
|  | | | | |

| **Coefficientsa** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | .536 | .013 |  | 40.652 | .000 |
| S4Q5 AGES 5-12 DID NOT LISTEN-W3 | .155 | .015 | .151 | 10.625 | .000 |
| a. Dependent Variable: S12Q8 PAST 7 DAYS TROUBLE CONCENTRTNG-W3 | | | | | | |

A regression test was conducted to measure the relationship between concentration as adults and the inability to listen between the ages of 5 and 12. The test was significant *R*(4861) =.151, *p >.*00.R2 accounts for 2.3% of the variability in our sample. The regression equation would be Ŷ= .536 + .16(2). My score would be .883 on concentration as adults and inability to listen as children. Concentration as adults and inability to listen as children is dependent on each other.

|  |
| --- |

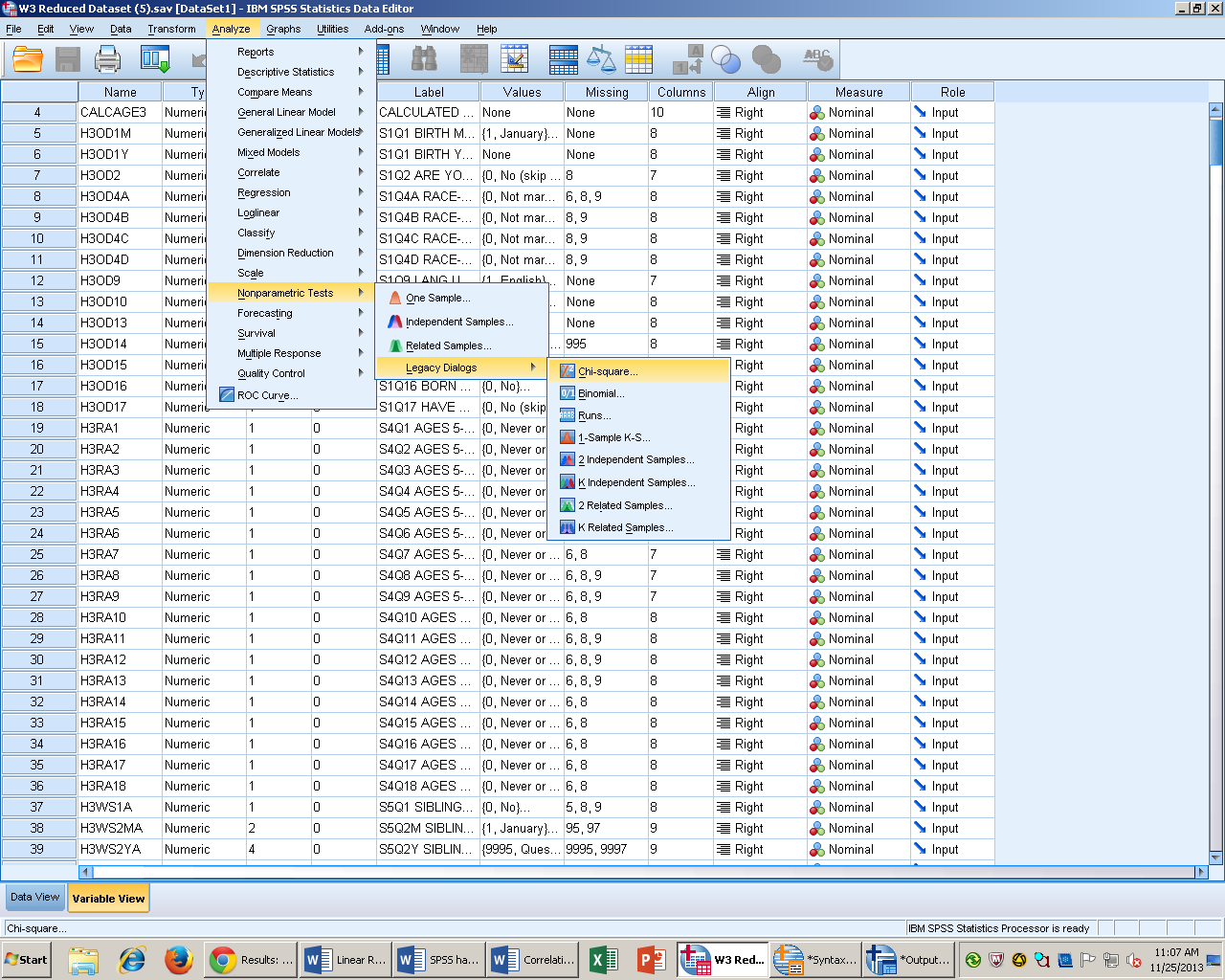
Goodness of Fit Chi Square

**Description of Statistic**

For a chi square goodness of fit test you need one variable that must be on the nominal scale of measurement. When using nominal variables we cannot use means instead we focus on frequency data. For the chi square goodness of fit test you compare *actual or observed* frequencies to *expected* frequencies. The chi square distribution cannot be less than zero or negative. The shape is right skew. The degree of freedom for this test is number of categories -1. The chi square goodness of fit is not used often in psychology.

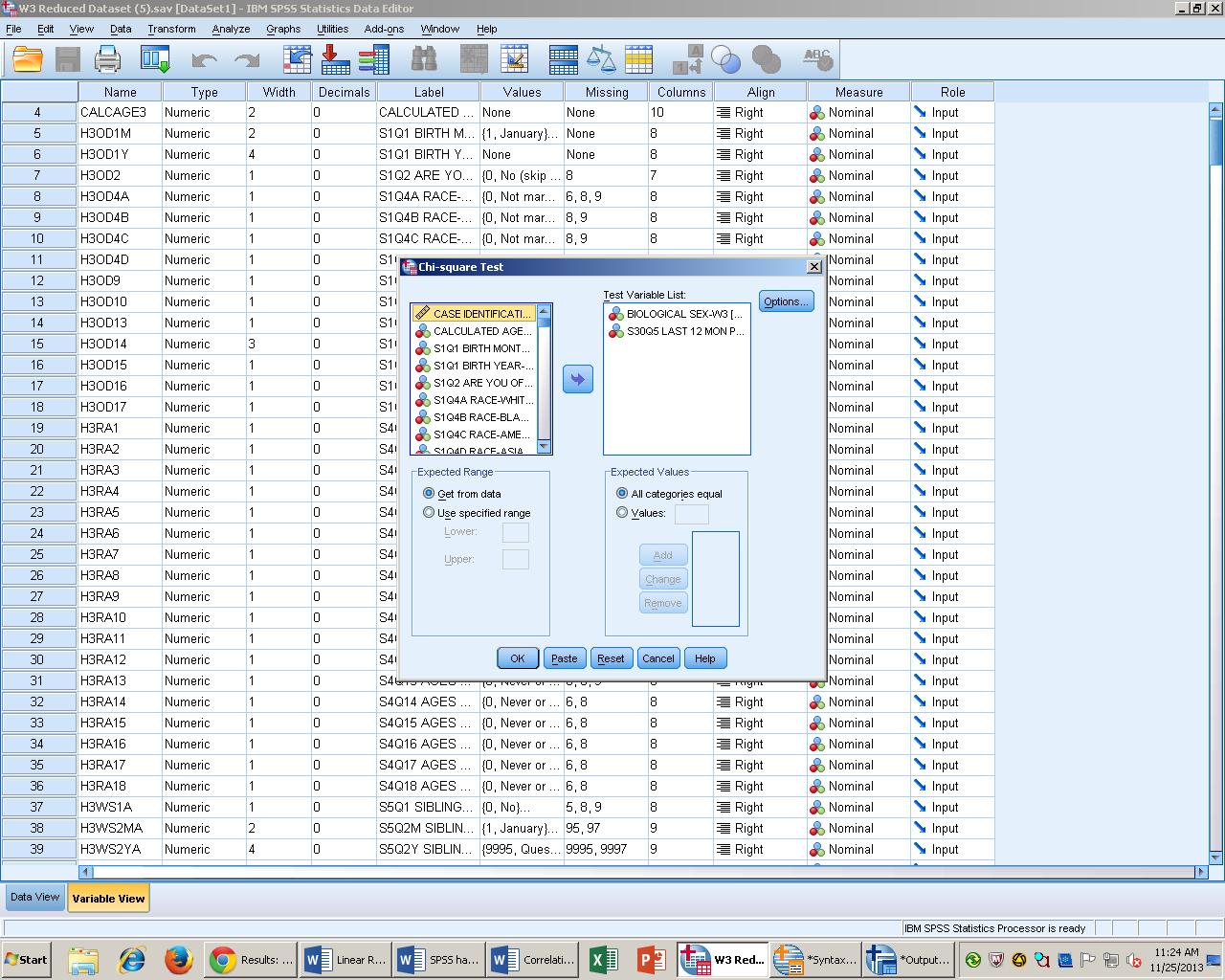
**SPSS Steps**

1. Analyze Nonparametric Tests Legacy Dialogues Chi Square



2. Select the variable(s) of interest and move it into the “test variable list.”

3. Under expected values, “all categories equal” is the default option.



4. Select paste to put it into syntax and run the results.

**Syntax**

**NPAR TESTS**

**/CHISQUARE=BIO\_SEX3 H3CC3**

**/EXPECTED=EQUAL**

**/STATISTICS DESCRIPTIVES**

**/MISSING ANALYSIS.**

**Ouput Interperation**

| **S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3** | | | |
| --- | --- | --- | --- |
|  | Observed N | Expected N | Residual |
| No (skip to Q.5) | 3454 | 2430.5 | 1023.5 |
| Yes | 1407 | 2430.5 | -1023.5 |
| Total | 4861 |  |  |

| **BIOLOGICAL SEX-W3** | | | |
| --- | --- | --- | --- |
|  | Observed N | Expected N | Residual |
| Male | 2253 | 2441.0 | -188.0 |
| Female | 2629 | 2441.0 | 188.0 |
| Total | 4882 |  |  |

| **Test Statistics** | | |
| --- | --- | --- |
|  | BIOLOGICAL SEX-W3 | S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 |
| Chi-Square | 28.959a | 862.006b |
| df | 1 | 1 |
| Asymp. Sig. | .000 | .000 |
|  | | |

A chi square goodness of fit test was conducted to compare actual frequencies to expected frequencies. The test of gender was significant χ2 (1, *n* = 4882) = 28.95, *p* < .001. The biological sex, female expected count (2441) was greater than the observed count (2629). The biological sex, male expected count (2441) was less than the observed count (2253). The test of volunteer work was also significant χ2 (1, n = 862.00) = 862.00, *p* >.001. The “yes” expected (2430.05) is greater than the observed (1407). The “no” expected (2430.05) is greater than the observed (3454).

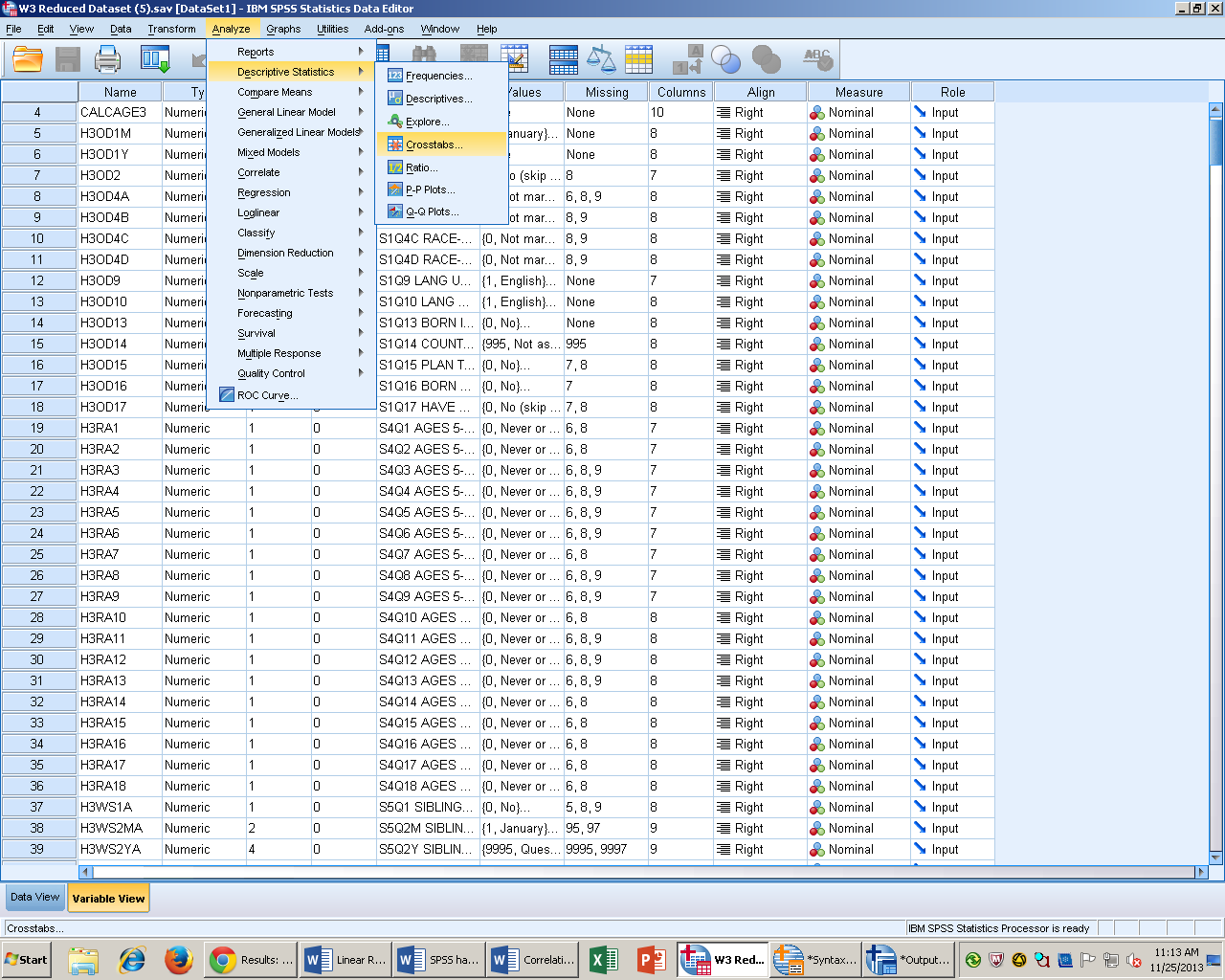
Chi Square Test of Independence

**Descriptive Statistics**

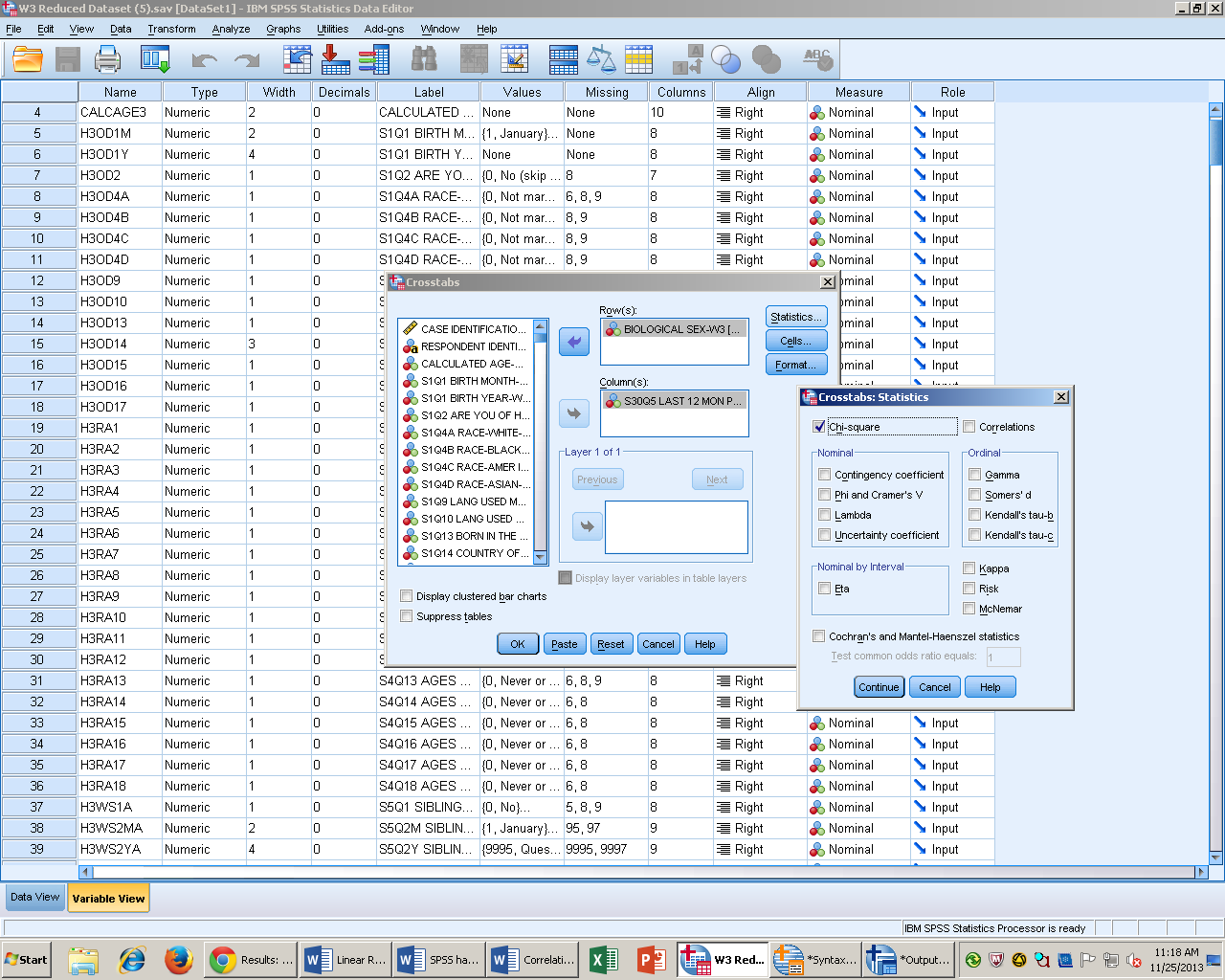
The chi square test of the independence requires that you have two variables that are on the nominal scale of measurement. Rather than focusing on simple frequencies, we are going to need to look at *proportions* of people in different combinations. Assumption of the chi square test is that each score must not have any special relationship to any other scores. Therefore you cannot use chi square for repeated measures design.

SPSS Steps

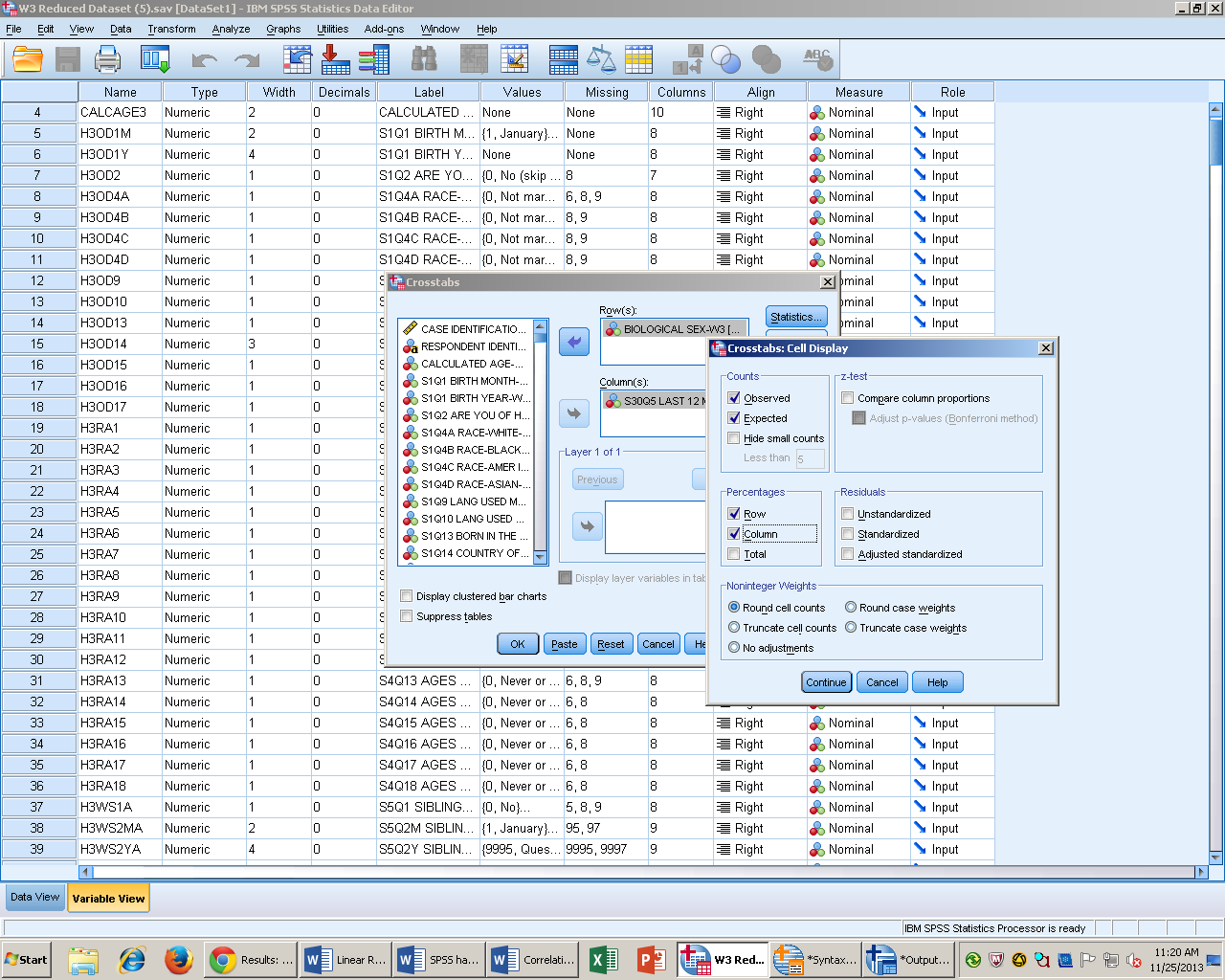
1. Analyze Descriptive Statistics Crosstabs



1. Select the row variable of interest and click the arrow to move it from the left hand list of variables to the right hand “row variables” box. Next select the column variable of interest and into the “column variables” box.
2. Select the blue box labeled “Statistics” and click on chi square. Then press continue.



1. Select the blue box labeled “Cells.” Under counts click on expected, observed should be selected by default. Under percentages select row and column. Then press continue.



1. Click paste to put it into syntax, and then run the Analysis.

**Syntax**

CROSSTABS

/TABLES=BIO\_SEX3 BY H3CC3

/FORMAT=AVALUE TABLES

/STATISTICS=CHISQ

/CELLS=COUNT EXPECTED ROW COLUMN

/COUNT ROUND CELL.

**Output Interpretation**

| **Case Processing Summary** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Cases | | | | | |
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| BIOLOGICAL SEX-W3 \* S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 | 4861 | 99.6% | 21 | .4% | 4882 | 100.0% |

| **BIOLOGICAL SEX-W3 \* S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 Crosstabulation** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | | S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 | | Total |
| No (skip to Q.5) | Yes |
| BIOLOGICAL SEX-W3 | Male | Count | 1637 | 604 | 2241 |
| Expected Count | 1592.4 | 648.6 | 2241.0 |
| % within BIOLOGICAL SEX-W3 | 73.0% | 27.0% | 100.0% |
| % within S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 | 47.4% | 42.9% | 46.1% |
| Female | Count | 1817 | 803 | 2620 |
| Expected Count | 1861.6 | 758.4 | 2620.0 |
| % within BIOLOGICAL SEX-W3 | 69.4% | 30.6% | 100.0% |
| % within S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 | 52.6% | 57.1% | 53.9% |
| Total | | Count | 3454 | 1407 | 4861 |
| Expected Count | 3454.0 | 1407.0 | 4861.0 |
| % within BIOLOGICAL SEX-W3 | 71.1% | 28.9% | 100.0% |
| % within S30Q5 LAST 12 MON PERFORM VOLTRY WORK-W3 | 100.0% | 100.0% | 100.0% |

| **Chi-Square Tests** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | Value | df | Asymp. Sig. (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
| Pearson Chi-Square | 8.025a | 1 | .005 |  |  |
| Continuity Correctionb | 7.847 | 1 | .005 |  |  |
| Likelihood Ratio | 8.045 | 1 | .005 |  |  |
| Fisher's Exact Test |  |  |  | .005 | .003 |
| Linear-by-Linear Association | 8.024 | 1 | .005 |  |  |
| N of Valid Cases | 4861 |  |  |  |  |
| a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 648.65.  b. Computed only for a 2x2 table | | | | | |

To test the prediction that the individual’s biological sex is dependent of performing any unpaid community service in the last twelve months. I used a chi square goodness of fit test. I reject the null hypothesis. My results are significant *χ2* (1, n = 4861) = 8.0215, *p*=.005. The observed count (803) of female and community service is greater than the expected count (758.4) The observed count (1637) of males who do not participate in community service is less than the expected count (1592.4).