Using SPSS to get the Most out of your CORE Alcohol and other Drug Data

Eric S. Davidson, MA, CSADP
Director, Illinois Higher Ed. Center
August 31, 2010
Introductions

• Name
• Institution
• Position/Title
• Why you are here today!!!
Opening Data

• Open -> Data->
Data View vs. Variable View

• Data view = the actual data entered
• Variable view = description/definition of data contained in data view
Variable View

• Essential to have. Without it you can analyze data, but you’ll be looking back at answer codes

• Variable data = place for words

• Data view – you can put words in, however SPSS will not analyze data.
Variable View - Name

- Names of your variables
- Any length, but SPSS likes names less than 10 characters
- Must begin with a letter
- May not end with a period
- Can use @, #, _, or $
- Duplicate names are not allowed
- Names are not case sensitive, Alc, ALC, AlC are all identical
- Certain variable names may not be used – see page 31 bottom
Variable View - Type

- Numeric variables are most common
- Can include comma, dot, scientific notation, date, dollar, currency
Variable View - Width

• Number of characters you expect variable entry to have
Variable View - Decimals

- Number of decimal places for each variable
Variable View - Label

- Ability to give meaning to variable name
Variable View - Values

• Every variable and level of a variable must be given some form of numerical representation/value
• E.g. gender: 1 = female, 2 = male
• Up to 60 characters
Variable View - Missing

- Rarely used
- Used to differentiate those who did not answer vs those who answered differently
Variable View – Column Width

• Column width
• Can be manipulated to allow you to see more or less columns
Variable View - Measure

- Nominal – groups, no order of ranking. Fr/So/Jr/Sr – one is not greater than the other – circles
- Ordinal – many of the CORE Scales, ranking without – Bar chart
- Interval Scale – temperature
- Ratio Scale – age
Variable View - Measure

• Interval Scale – temperature
  – ruler
• Ratio Scale – age
  – ruler
Entering Data

• By Variable- data view
• By Variable – variable view
• By case or subject
Editing Data

• Changing a Cell value – click on cell and enter data
• Inserting a new case – click on case number above where you want the new case to be
• Inserting a new variable – click on variable to the right of the variable where you want to add
• Copying and Pasting Cells
Replacing Missing Data – Categorical Data

• ethnicity as an example, code missing data as an extra level. For example, if you have 5 levels – code the missing data as a 6th level.
Replacing Missing Data – Continuous Data

- SPSS has several different mechanisms to do so
- Transform -> Replace Missing Values
- Series Mean – mean or average value of all cases
- Mean of Nearby points – mean of surrounding cases
- Median of Nearby points – median of surrounding cases
Creating Variables
Computing Variables

- Summing up all quizzes, test grades
- Transform->Compute Variable
- Target Variable Box = name of new variable
- Numeric Expression Box = expression that will define the new variable
- Functions – math functions that you can use
Recoding into Different Variables

• Creates new variables by dividing existing categories
  – Example – recoding %’s into letter grades
Recoding into Different Variables

• Transform -> Recode into Different Variables
• Grades (Q9) -> numeric variable/output variable
• Put in new variable name and label
• Click old and new values
Recoding into Different Variables Cont

- Place one level of “old variable”
  - $A^+ = 13$, now $A^+ = 4$
- Click add
- Continue with $A=12$, now $A=4$, etc.
- Click Continue
- Click Okay
Graphing

- Bar
- Line
- Pie
- Box
- Error Bar
- Histograms
- Scatter Plots
Graphing

- Graphs->chart builder->choose graph type and move to box
- Determine x axis and y axis
- Element properties -> x-axis, mark sort by label under categories
- Element properties -> y-axis indicate scale range if you want it changed
- Hit apply
- Hit okay in chart builder
- Graph will come out in output
- Double Click on chart to edit
Reports - Codebook

• Analyze->reports->codebook
Frequencies

- Analyze->descriptive statistics->frequencies
- Determine descriptive statistics
- Determine charts if wanted or desired
- Hit okay
Descriptive Statistics

• Analyze->descriptive statistics->descriptive
Crosstabulations

- Analyze->descriptive statistics->crosstabs
- Place one variable in rows
- Place another variable in columns
- Click on statistics
- Click on Chi Square
Crosstabulations Continued

- Click on cells, click on observed and expected in counts
- Click okay
- Look at output, look at Pearson’s Chi Square
- If Asymp sig is less than 0.05, findings are statistically significant
- If Asymp sig is equal to or greater than 0.05, and equal to or less than 0.10 findings are marginally significant
Means Procedures

• Analyze->CompareMeans->Means

• Dependent Variable – Must be scale (interval or ordinal)

• Independent Variable – grouping variable

• Means Options Box
T-Tests – Independent Samples

• Used to determine a statistically significant difference in means between two groups
• Since we’re looking at means, must have scale level data (interval or ratio)
  – Ave # of drinks per week
• Analyze->CompareMeans->Independent Samples T-test
• Test Variable – Dependent Variable
• Grouping Variable – Independent Variable
• Define Groups – need to know the numbers used for labeling independent variable
T-Test Independent Samples Continued

- Options = confidence level. 0.05 considered most typical cut-off
- Levene’s Test for equality
  - If significance of F is =/< .05, variance differences exist between the two groups – use the unequal line to determine if t-test is significant
  - If significance of F is >.05, variance differences do not exist between the two groups, use the equal line to determine if t-test is significant.
T-Tests – One Sample

• Use to compare mean of a distribution with some standard objective.
• Use to compare your mean average of drinks consumed with state average (5.2)
• Analyze->CompareMeans->One Sample T-test
• Test value = 5.2
Paired Samples T-test

• Within subjects needed – two test scores
• Also looks at correlation. Do those who score high on test one score high on test two
One-Way ANOVA

• Used to determine a statistically significant difference in means between three or more groups
• One dependent variable and one independent variable
• Since we’re looking at means, must have scale level data (interval or ratio)
  – Ave # of drinks per week
One-Way ANOVA Continued

• Post-hoc tests will be needed to determine where differences exist, ANOVA only tells you if differences exist—not where.

• Analyze->CompareMeans->One-Way ANOVA

• Review of AdHoc Post tests
Two-Way ANOVA

• Like 1-Way ANOVA, but looks at the influence of 2 categorical variables on an scale dependent variable.

• Gender and Greek Status by drinks per week

• READ Chapter 13
Two Way ANOVA

- Analyze->General Linear Model ->Univariate
- Move DV to Dependent Variable
- Move IV’s to Fixed Factors
- Determine Post-Hocs – ESD prefers Tukeys B
- Click Options – Descriptive Stats, Estimates of effect size, and observed power
- Click continue and click OK
Correlations

- Chapter 10
- Correlation does not imply causality
Pearson’s Correlation

• Used when dependent variable is scale level (rankings)
  – Interval
  – ratio
• Analyze->Correlate->Bivariate
• Move Variables Over
• Click on Pearson Correlation
Pearson’s Correlations

- Determine if you’re using one or two tailed – Two-Tailed will be selected by default
- Click on one-tail if you have clear knowledge of the direction of your correlation
- Pairwise – include subject/case even if missing a few values
- List-wise - exclude case if missing any data
Correlations - Spearman's

- Used when dependent variable is ordinal level (rankings)
  - First, second, third
  - Core questions when ranges are used (0-2, 3-5, etc.)

- Analyze->Correlate->Bivariate
- Move Variables Over
- Click on Spearman Correlation
Correlations - Spearman's

- Determine if you’re using one or two tailed – Two-Tailed will be selected by default
- Click on one-tail if you have clear knowledge of the direction of your correlation
- Pairwise – include subject/case even if missing a few values
- List-wise- exclude case if missing any data
Presenter Information

Eric S. Davidson, MA, CSADP
Director, Illinois Higher Education Center
Eastern Illinois University

600 Lincoln Avenue
Charleston, IL 61920
217/581-2019
esdavidson@eiu.edu