Lecture Outline

• Merging Data
• Data Types
• Variable Creation
  • Transforming dates
  • Calculate a new variable from one or more existing variables
  • Create a categorical variable from a numeric variable
  • Create an indicator (dummy) variable
Merging Data

• Datasets being merged must share at least one variable to match on and must be sorted by this variable prior to merging
  • Use menus to select Data > Sort Cases
  • Choose the matching variables in the ‘Sort Cases’ dialog

• Types of merges in SPSS
  • Use menus to select Data > Merge Files
    • Add Variables – same cases but different variables
    • Add Vases – different cases but same variables
Data Types

Determine how you summarize data, display data in a table or graph, and analyze data using statistical methods.

Qualitative / Categorical
- Ordinal – the order among the categories is important
- Nominal – no inherent ordering
  - Dichotomous or binary – two categories or levels

Quantitative / Numeric
- Discrete – integers and counts
- Continuous – not restricted to integers
Data Types

Determine how you summarize data, display data in a table or graph, and analyze data using statistical methods

Qualitative / Categorical
- Ordinal – the order among the categories is important
- Nominal – no inherent ordering
  - Dichotomous or binary – two categories or levels

Quantitative / Numeric
- Discrete – integers and counts
- Continuous – not restricted to integers
Transforming Dates

• Date data may be qualitative or quantitative depending on how it is manipulated

• In SPSS, date variables are stored as a number corresponding to the number of days since October 15, 1582

• Use menus to select Transform > Compute Variable
  • Date Arithmetic – compute time between dates as a new variable
  • Date Extraction – extract parts of a date (i.e. day, month, year)
  • Date Creation – create a new date variable
Demo

• SPSS sample data file dietstudy.sav
• Each case represents a separate subject
• Pre-, interim-, and post-diet weights and triglyceride levels
Calculate a New Variable

• Menus to select Transform > Compute Variable
  • Name and label new variable
  • Combine existing variables with arithmetic operations/functions

• New variable occupies the last row of variable view and last column of data view
Create a Categorical Variable from a Numeric Variable

- **Menus to select Transform > Recode into Different Variables**
  - Name and label new variable
  - Specify range parameters in the ‘Old and New Values’ dialog
  - Edit formatting of new variable in the variable view

- **Menus to select Transform > Visual Binning**
  - Name and label new variable
  - ‘Make Cutpoints’
    - Equal Width Intervals
    - Equal Percentiles Based on Scanned Cases
    - Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases
  - Specify whether cutpoints should be included or excluded
Create an Indicator or Dummy Variable

• Menus to select Transform > Recode into Different Variables
  • Name and label new variable
  • Specify value parameters in the ‘Old and New Values’ dialog
• Repeat for additional indicator variables
Next Lecture

• Lecture 3: Summarizing Data
  • Descriptive statistics
  • Graphical displays
  • By a grouping variable

• Questions? Email: laura.berbert@childrens.harvard.edu