Surviving SPSS

http://dataservices.gmu.edu/workshops/spss

http://dataservices.gmu.edu/software/spss

Debby Kermer
George Mason University Libraries
Data Services Research Consultant
Mason Data Services

http://dataservices.gmu.edu

Wendy Mann
Joy Suh
Debby Kermer

datahelp@gmu.edu

Geospatial Analysis and Mapping
Finding & Archiving Data, Data Management Plans
Qualitative & Quantitative Methods, Analysis, & Software
Why SPSS?

• Easy to get started
• Optimized for the behavioral and social sciences
• Easy to maintain weighting for sampled surveys
• Best for working with missing values
<table>
<thead>
<tr>
<th>Statistical Software</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SPSS</strong></td>
<td><strong>STATA</strong></td>
</tr>
<tr>
<td>Easy to start, limited capability</td>
<td>Easy syntax, highly extensible</td>
</tr>
<tr>
<td><strong>Best for</strong> those with infrequent and/or minimal needs</td>
<td><strong>Best for</strong> academics doing cutting-edge research</td>
</tr>
<tr>
<td><strong>SAS</strong></td>
<td><strong>R</strong></td>
</tr>
<tr>
<td>Hard to learn, highly capable</td>
<td>Hard to learn, highly extensible</td>
</tr>
<tr>
<td><strong>Best for</strong> managing huge and/or complex datasets</td>
<td><strong>Best for</strong> those who program and know what they are doing</td>
</tr>
</tbody>
</table>
http://dataservices.gmu.edu/software/spss

Data Services

Software

ArcGIS
NVivo
Qualtrics
R
SAS
SPSS
Stata
Tableau

SPSS

SPSS is a software package used for statistical analysis. It is a powerful tool for data analysis, allowing users to perform complex statistical analyses with ease. SPSS is widely used in research and industry for its user-friendly interface and comprehensive statistical functions.

Handouts from SPSS

Basic Skills

- BruneiASK YouTube Channel
  - from Brunei University ASK
  - A few good tips in SPSS

- SPSS Statistics
  - from Laerd
  - A collection of resources covering SPSS basics

- SPSS Tutorial
  - from Ken Poon
  - Both new and experienced users can benefit from these tips.
  - Shows how to use SPSS for various tasks.

Basic Skills

Textbook-like Resources

- Laerd Statistics (<$5 a month)
  - This is not a free or library-provided site. But, it is the most comprehensive plain-language source. Free access to much of the content is available through searching Google.

- Covers basic (e.g., measures of central tendency) and advanced SPSS skills (e.g., linear regression) with extensive explanations and step-by-step instructions and help on writing up the results.

- Statistics Hell
  - by Andy Field, author of Discovering Statistics Using SPSS
  - Videos and handouts (that look and read like textbook chapters) give detailed assistance with basic and advanced SPSS skills.

Data Analysis

- Resources for advanced users
  - Guides and tutorials for more complex analyses

- FAQs
  - Answers to frequently asked questions about SPSS usage and troubleshooting

News

Get Help

Learn More
IBM SPSS Statistics Windows and Files

Data → .sav

The **Data Window** has two views, each with important information about your data.

- **Data View**
- **Variable View**

Use the tabs at the bottom to **switch views**. Or, double-click the variable #/name in the blue row/column header to keep your place.

**Columns** in both Views can be dragged wider or narrower, just like in Excel.

It is important to **Label** the numeric values of categorical variables. In Data View, you can see either by clicking →

Visit **Edit | Options** to specify if labels or names/values are displayed in Variable Lists and Output, and to set the default decimals.

Output → .spv

The **Output Window** stores a log of what you did and tables and graphs generated.

The log is the **syntax** SPSS actually uses to create the results. Double-click on the log to select and copy the text to use later.

Use the **Outline** to skip to different sections. Clean up by selecting items and pressing **Delete** on the keyboard.

Double-click any item to edit or select text. Choose a Pivot Table **TableLook** in Options to alter the font and borders for all tables.

To save the output as a Word or Excel file, choose **File | Export**. If **Objects to Export** is “All”, extra diagnostic info is included.

Syntax → .sps

The **Syntax Window** stores SPSS code so you can examine, fix, or redo prior steps, or allow others to replicate your analysis.

Instead of clicking Ok, choose to **Paste** the Syntax into the file for you to Run yourself.

Syntax is just text, so you can **edit or copy** it to change variables or options. The **BLUE** is for commands, which end with a period.

To **Run** the code, highlight the command(s) you want to do, then click the green arrow.

* Comment between an asterisk and period, or add explanation anywhere /* no period */

****** Be creative for headings ******
Download the Dataset

Go to:  http://dataservices.gmu.edu/workshops/spss/

Find SPSS I then click Datafile

Complete the form and read the user agreement

Click Submit and Download Data

Save the file where you can find it

Open the file (by default “PST 11-05c.sav”)

   Double-Click or   File | Open... (if SPSS is open)
Data View

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3.02</td>
<td>101705</td>
<td>Spanish</td>
<td>727</td>
<td>Florida</td>
<td>SOUTH</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.79</td>
<td>100605</td>
<td>English</td>
<td>386</td>
<td>Florida</td>
<td>SOUTH</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2.56</td>
<td>101705</td>
<td>Spanish</td>
<td>770</td>
<td>Georgia</td>
<td>SOUTH</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2.35</td>
<td>101705</td>
<td>Spanish</td>
<td>718</td>
<td>New York</td>
<td>NORTHEA...</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2.69</td>
<td>100505</td>
<td>English</td>
<td>215</td>
<td>Pennsylvania</td>
<td>NORTHEA...</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>4</td>
<td>2.13</td>
<td>101705</td>
<td>Spanish</td>
<td>786</td>
<td>Florida</td>
<td>SOUTH</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>5</td>
<td>1.29</td>
<td>100505</td>
<td>English</td>
<td>336</td>
<td>North Carolina</td>
<td>SOUTH</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>5</td>
<td>3.15</td>
<td>101705</td>
<td>Spanish</td>
<td>716</td>
<td>New York</td>
<td>NORTHEA...</td>
<td></td>
</tr>
</tbody>
</table>
Look at the Data Window

Confirm you are in Data View

Where are Variables?

Where are Cases?
Navigating Data View

Variables are in columns
Find variable q2
Make q2 wider

Cases are in rows
Determine the # of cases (scroll down)
Values & Labels

Press or use View | Value Labels

**Value** = What is stored and analyzed; the "actual" value

**Label** = What is displayed; the characteristic or group

Value $\rightarrow$ Label

1 $\rightarrow$ “English”

2 $\rightarrow$ “Spanish”
Switching Views

#1: Click on the Tabs

or

#2: Double-Click on the Blue Header area

- Variable View
- Data View

From Data View, double-click q2 to switch views
### Variable View

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Missing</th>
<th>Columns</th>
<th>Align</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>praid</td>
<td>Numeric</td>
<td>4</td>
<td>0</td>
<td>Unique ID</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>respid</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>Record</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>weight</td>
<td>Numeric</td>
<td>4</td>
<td>2</td>
<td>Interview date</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>int_date</td>
<td>Numeric</td>
<td>6</td>
<td>0</td>
<td>Language of Int...</td>
<td>(1, English)...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>lang</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Area code</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>state</td>
<td>Numeric</td>
<td>2</td>
<td>0</td>
<td>State from FIPS</td>
<td>(1, Alabama...)</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>cregion</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Census region</td>
<td>(1, NORTH...)</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>msa</td>
<td>Numeric</td>
<td>5</td>
<td>0</td>
<td>MSA code</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>fips</td>
<td>Numeric</td>
<td>5</td>
<td>0</td>
<td>FIPS code</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>usr</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Community typ...</td>
<td>(R, Rural...)</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
</tr>
<tr>
<td>usr1</td>
<td>Numeric</td>
<td>6</td>
<td>0</td>
<td>Community typ...</td>
<td>{1, Urban}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
</tr>
<tr>
<td>tz</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Time zone</td>
<td>(C, Central)</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
</tr>
</tbody>
</table>
Variables

**Name** – "Nickname"

**Label** – Display Label

**Role** – Theoretical Role (only for new / fancy analyses)

Change *sex* (15) to *gender*

Change the Label for *sex*

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Missing</th>
<th>Columns</th>
<th>Align</th>
<th>Measure</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>usr</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Community type - alpha</td>
<td>{R, Rural}...</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
<td>Input</td>
</tr>
<tr>
<td>usr1</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>Community type - numeric</td>
<td>{1, Urban}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>tz</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Time zone</td>
<td>{C, Central}...</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
<td>Input</td>
</tr>
<tr>
<td>form</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Sample split</td>
<td>{1, Form 1}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
<tr>
<td>sex</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Respondent's sex</td>
<td>{1, Male}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
<tr>
<td>e2</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Q3: Generally, how would you say</td>
<td>{1, Very happy}</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
</tbody>
</table>
## Variable Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage &amp; Analysis</td>
<td>Level of Measurement</td>
</tr>
<tr>
<td>Numeric</td>
<td>Scale = Interval or Ratio</td>
</tr>
<tr>
<td></td>
<td>Prevents incorrect analysis</td>
</tr>
<tr>
<td>String</td>
<td>String type must be Nominal</td>
</tr>
<tr>
<td></td>
<td>For essay text; only keeps text up to</td>
</tr>
<tr>
<td></td>
<td>the Width</td>
</tr>
</tbody>
</table>

### Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Missing</th>
<th>Columns</th>
<th>Align</th>
<th>Measure</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>usr</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Community type - alpha</td>
<td>{R, Rural}...</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
<td>Input</td>
</tr>
<tr>
<td>usr1</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>Community type - numeric</td>
<td>{1, Urban}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>tz</td>
<td>String</td>
<td>1</td>
<td>0</td>
<td>Time zone</td>
<td>{C, Central}...</td>
<td>None</td>
<td>8</td>
<td>Left</td>
<td>Nominal</td>
<td>Input</td>
</tr>
<tr>
<td>form</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Sample split</td>
<td>{1, Form 1}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
<tr>
<td>sex</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Respondent’s sex</td>
<td>{1, Male}...</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
<tr>
<td>q2</td>
<td>Numeric</td>
<td>1</td>
<td>0</td>
<td>Q2 Generally, how would you say?</td>
<td>{1, Very happy}</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Ordinal</td>
<td>Input</td>
</tr>
</tbody>
</table>
Values

For state (7):

• Change District of Columbia to DC
• Label 3 “Not used”
• Remove Label for 3

NOTE: Use this method ONLY when entering your own data. It is not replicable.
Q.2

Q.2  Generally, how would you say things are these days in your life -- would you say that you are very happy, pretty happy, or not too happy?

<table>
<thead>
<tr>
<th>Value</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very happy</td>
</tr>
<tr>
<td>2</td>
<td>Pretty happy</td>
</tr>
<tr>
<td>3</td>
<td>Not too happy</td>
</tr>
<tr>
<td>9</td>
<td>Don’t know/Refused</td>
</tr>
</tbody>
</table>

Name: q2
Type: Numeric
Width: 1
Measure: Ordinal
Values:

1  = “Very happy”
2  = “Pretty happy”
3  = “Not too happy”
9  = “Don’t know/Refused”
Missing Values

**System Missing**
- *(a period)*

**User Missing**
- 7 = “N/A”
- 8 = “Don’t Know”
- 9 = “Refused”
User Missing Values

Add 9 as a User Missing value for q2

Add 7 & 9 as User Missing values for q3a

Copy User Missing values from q3a to all q3-series

Right Click and choose Copy

Select All the q3a cells

Right Click and choose Paste
Data View Display

**Decimals** – # of decimals shown (in output, too)

**Column** – Width of the column

**Align** – Justification of values in column

only affects the **look** in Data View
Analyze Menu

- Frequencies...
- Crosstabs...
- Descriptive Statistics
- Explore...
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
Descriptive Statistics

How happy are people with their life?
Finding Variables

Find \( q_2 \)

Type name VERY FAST

Right Click to re-sort
Choosing Variables

Move **q2** to the box then **click OK**

- **Click the arrow**
- **Double-Click the name**
Output Window / Viewer

Syntax

Click and press the Delete Key

Outline
Output Tables

<table>
<thead>
<tr>
<th>Valid</th>
<th>Variable Name</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>1 Very happy</td>
<td>2108</td>
<td>34.1</td>
<td>34.5</td>
<td>34.5</td>
</tr>
<tr>
<td></td>
<td>2 Pretty happy</td>
<td>3097</td>
<td>50.1</td>
<td>50.7</td>
<td>85.2</td>
</tr>
<tr>
<td></td>
<td>3 Not too happy</td>
<td>901</td>
<td>14.6</td>
<td>14.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6105</td>
<td>98.7</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Missing</th>
<th>Value Label</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Don’t know/Refused</td>
<td></td>
<td>82</td>
<td>1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6188</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SPSS Options

Edit | Options
Options: General

- Display names
- No scientific notation for small numbers in tables
- Raise viewer window
- Open syntax window at startup
Options: Output

Outline Labeling
- Variables in item labels shown as:
  - Names and Labels
  - Values and Labels

One Click Descriptives
- Suppress tables with many categories
- Maximum number of categories: 50
- Include a chart in the output

Output Display
- Display method:
  - Model Viewer
  - Pivot tables and charts
- Output display method applies to:
  - Generalized Linear Mixed Models
  - Nonparametric Tests

Pivot Table Labeling
- Variable in labels shown as:
  - Names and Labels
  - Values and Labels

Screen Reader Accessibility
- Read full row and column labels for each cell in pivot tables
- Read only row or column labels that change when you move to a different cell
### Goal

Is there a relationship between **happiness** and having **pets**?

<table>
<thead>
<tr>
<th>Happiness</th>
<th>q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Happy = 1 &amp; 2</td>
</tr>
<tr>
<td>0</td>
<td>Not Happy = 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Have pets</th>
<th>q75</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Has Pets = 1</td>
</tr>
<tr>
<td>0</td>
<td>No Pets = 2+</td>
</tr>
</tbody>
</table>

Use Edit | Go to Variable...
Transform Menu

Working with Variables
**Goal:** Make an indicator (0/1) for people with pets

Transform | Compute Variable...

New Variable Name: `has_pet`

What to indicate: `q75 = 1`
Goal: Recode q2 into is_happy: Happy vs Unhappy

Transform | Recode into Different Variables...

1. q2 → is_happy
2. Name: is_happy
3. Label: Happy?
4. Old and New Values...
5. Old Value: 1
6. New Value: 1
7. Press Add
Labeling with Syntax

Data | Define Variable Properties...

Select the variables to scan. They should be categorical (nominal or ordinal) for best results. You can change the measurement level setting in the next panel.

Variables:
- givediff
- agegen3pt
- givedac
- agegen5pt
- havgkids2
- gimme5
- givboth3
- m1stat2
- has_pet
- is_happy

Limit number of cases scanned to: 200

Continue | Cancel | Help
Variable Properties

- **Current Variable**: `has_pet` with the label set to `Has pet?`.
- **Measurement Level**: Nominal.
- **Role**: Input.
- **Unlabeled values**: 0.
- **Value Label Grid**:
  - Value 1: 0 No Pet
  - Value 2: 1 Has a Pet

- **Cases scanned**: 3014
- **Value list limit**: 200
Add Labels

Variable Name: has_pet
Variable Label: Has a Pet?
Value Labels: 0 = No Pet 1 = Has a Pet

is_happy
Variable Label: Happy?
Value Labels: 0 = Not Happy 1 = Happy
Variable Name: has_pet
Variable Label: Has a Pet?
Value Labels: 0 = No Pet, 1 = Has a Pet

Variable Name: is_happy
Variable Label: Happy?
Value Labels: 0 = Not Happy, 1 = Happy
Crosstabulation Analysis
Crosstabs

Analyze | Descriptive Statistics | Crosstabs...
Among those **with no pet**, **84.6%** report being generally happy.

Among those who **have a pet**, **85.7%** report being generally happy.
Chi-Square

Analyze | Descriptive Statistics | Crosstabs...
Data Menu

Working with Cases
Weighting

# of people a case represents, accounts for the over- or under-representation of a group

In “Readme PST 11-05.txt”

The WEIGHT variable is the weighting variable

Data | Weight Cases...

Look for

or Select:

Weight cases by: weight
Selecting [a Subset of] Cases

**Goal:** Do analysis on only those who have a pet.

**Data | Select Cases...**

- **Filtering**
  - ignored cases
  - 0 or missing value
  - included cases
  - every other value

[Diagram showing select cases options]

Options:
- Use filter variable:
  - has_pet
- Filter out unselected cases
- Copy selected cases to a new dataset
- Delete unselected cases
# Tables

**Crosstabs: q76dog & is_happy**

<table>
<thead>
<tr>
<th></th>
<th>is_happy Happy?</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 Not Happy</td>
<td>1 Happy</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>q76dog Owns Dog - From Q76</td>
<td>156</td>
<td>909</td>
<td>1065</td>
<td></td>
</tr>
<tr>
<td>% within q76dog Owns Dog - From Q76</td>
<td>14.6%</td>
<td><strong>85.4%</strong></td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td>1 Yes</td>
<td>341</td>
<td>2072</td>
<td>2413</td>
<td></td>
</tr>
<tr>
<td>% within q76dog Owns Dog - From Q76</td>
<td>14.1%</td>
<td><strong>85.9%</strong></td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>497</td>
<td>2981</td>
<td>3478</td>
<td></td>
</tr>
<tr>
<td>% within q76dog Owns Dog - From Q76</td>
<td>14.3%</td>
<td><strong>85.7%</strong></td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>
Splitting a File for Comparison

**Goal:** Compare males and females with pets

**Data | Split File...**

- Turn off Grouping
- Add group as a factor for each analysis
- Do analysis for one group, then for the other

Choose **Compare groups**
Move **gender** to box
Guys with dogs are more likely to be happy.

Girls with dogs are less likely to be happy.

Examining Output

Redo Crosstabs: q76dog & is_happy

<table>
<thead>
<tr>
<th>gender</th>
<th>Respondent's gender</th>
<th>is_happy Happy?</th>
<th>0 Not Happy</th>
<th>1 Happy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Male</td>
<td>q76dog Owns Dog - From Q76</td>
<td>0 No</td>
<td>Count</td>
<td>80</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>17.0%</td>
<td>83.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td>1 Yes</td>
<td>Count</td>
<td>148</td>
<td>1015</td>
<td>1163</td>
</tr>
<tr>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>12.7%</td>
<td>87.3%</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Count</td>
<td>228</td>
<td>1405</td>
<td>1633</td>
</tr>
<tr>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>14.0%</td>
<td>86.0%</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td>2 Female</td>
<td>q76dog Owns Dog - From Q76</td>
<td>0 No</td>
<td>Count</td>
<td>76</td>
<td>519</td>
</tr>
<tr>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>12.8%</td>
<td>87.2%</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 Yes</td>
<td>Count</td>
<td>192</td>
<td>1057</td>
<td>1249</td>
</tr>
<tr>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>15.4%</td>
<td>84.6%</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Count</td>
<td>268</td>
<td>1576</td>
<td>1844</td>
</tr>
<tr>
<td></td>
<td>% within q76dog Owns Dog - From Q76</td>
<td>14.5%</td>
<td>85.5%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>
### Status Bar

Shows status of:

- **Filter**
- **Split**
- **Weighting**

**Data View**

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Weight</th>
<th>Label</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>psaid</td>
<td>Numeric</td>
<td>4</td>
<td>Unique ID</td>
<td>None</td>
</tr>
<tr>
<td>respd</td>
<td>Numeric</td>
<td>8</td>
<td>Record</td>
<td>None</td>
</tr>
<tr>
<td>weight</td>
<td>Numeric</td>
<td>4</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>int_date</td>
<td>Numeric</td>
<td>6</td>
<td>Interview date</td>
<td>None</td>
</tr>
<tr>
<td>lang</td>
<td>Numeric</td>
<td>1</td>
<td>Language of Interview</td>
<td>{1, English}...</td>
</tr>
<tr>
<td>area</td>
<td>Numeric</td>
<td>3</td>
<td>Area code</td>
<td>None</td>
</tr>
<tr>
<td>state</td>
<td>Numeric</td>
<td>2</td>
<td>State from FIPS</td>
<td>{1, Alabama}...</td>
</tr>
<tr>
<td>cregon</td>
<td>Numeric</td>
<td>1</td>
<td>Census region</td>
<td>{1, NORTHEAST}...</td>
</tr>
<tr>
<td>msa</td>
<td>Numeric</td>
<td>5</td>
<td>MSA code</td>
<td>None</td>
</tr>
<tr>
<td>fips</td>
<td>Numeric</td>
<td>5</td>
<td>FIPS code</td>
<td>None</td>
</tr>
</tbody>
</table>
| usr      | String  | 1      | Community type-alpha | (R, Rural)...
| usr1     | Numeric | 8      | Community type-numerical | {1, Urban}...
| tz       | String  | 1      | Time zone | {1, Central}...
| form     | Numeric | 1      | Sample split | {1, Form I}...
| gender   | Numeric | 1      | Respondent’s sex | {1, Male}...
| q2       | Numeric | 1      | Q.2 Generally, how would you say things... | {1, Very happy}...
| q3a      | Numeric | 1      | Q.3a Your standard of living -- of what you... | {1, Very Satisfied}...
| q3b      | Numeric | 1      | Q.3b Your household income... | {1, Very Satisfied}...
| q3c      | Numeric | 1      | Q.3c Your family life... | {1, Very Satisfied}...
| q3d      | Numeric | 1      | Q.3d Your job -- the kind of work you do... | {1, Very Satisfied}...
| q3e      | Numeric | 1      | Q.3e Your housing situation... | {1, Very Satisfied}...
| q3f      | Numeric | 1      | Q.3f Your free time -- the time when you... | {1, Very Satisfied}...
| q3h      | Numeric | 1      | Q.3h Your relationship with your spouse... | {1, Very Satisfied}...
Return to Normal

Turn off Grouping
Data | Split File...

- Analyze all cases, do not create groups
- Compare groups
- Organize output by groups

Current Status: Analysis by groups is off.

Turn off Filtering
Data | Select Cases...

- All cases
- If condition is satisfied

Current Status: Do not filter cases
Finishing Up
Save Files

File | Save As...

Data File &
Output File &
Syntax File

To Exit, click “Yes”
Questions

datahelp@gmu.edu

http://dataservices.gmu.edu

http://dataservices.gmu.edu/software/spss/