# Intro to SPSS

Fall 2015

Intro to SPSS

- SPSS is a statistical analysis program that allows:
  - Data management
  - Graphs and tables
  - Statistical analyses
  - You will need: some basic statistics
    - We will discuss these
- SPSS is more *specialized* than Excel
- Provide data in a more precise way

SPSS

- Goals for this section of the course include:
  - Becoming familiar with Statistical Packages
  - Creating new Datasets
  - Importing & exporting Datasets
  - Manipulating data in a Dataset
  - Basic analysis of data (mainly descriptive statistics)
  - An overview of SPSS's advanced features
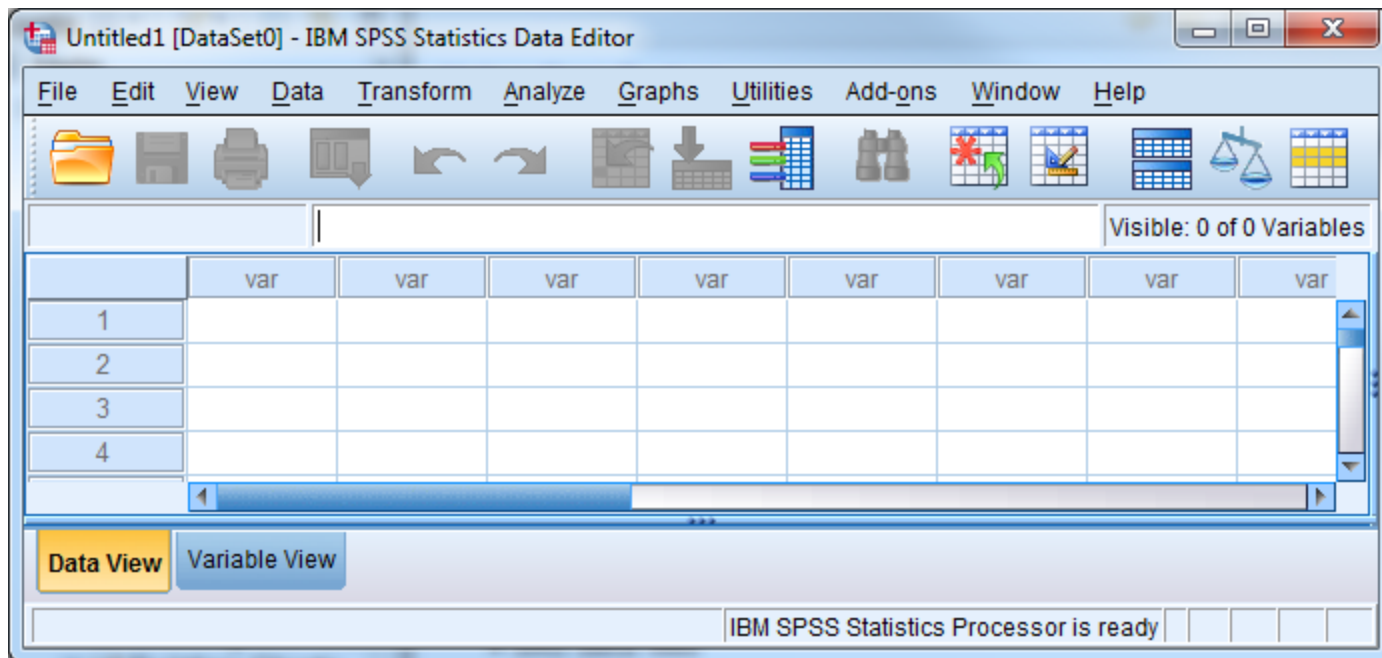 - Examining the Help utility within SPSS

Note: This is not a statistics course such as Math 207. We will only concentrate on basic statistical concepts.

Open SPSS
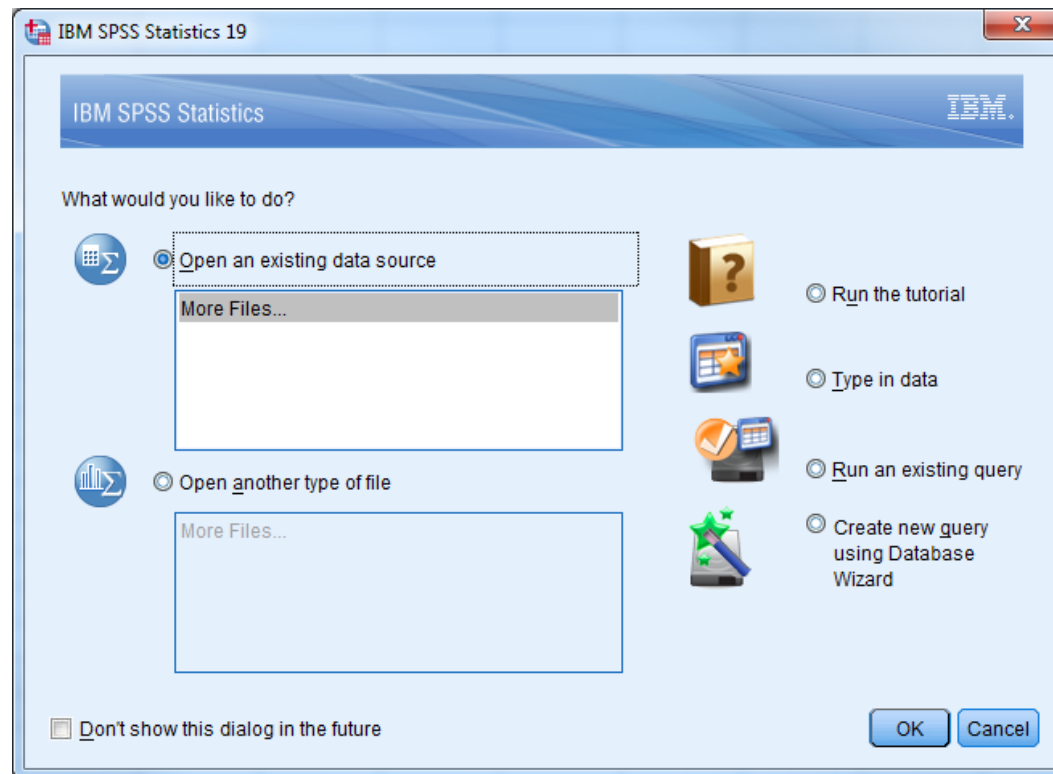
Create a Simple Dataset

- SPSS looks somewhat like Excel BUT there are several important differences
- Select the Data View tab

Excel versus SPSS Differences

- Column data pertains to a particular **variable**

List several examples of what a variable might be

- Row data is considered a case, an observation, or an individual

List several examples of an observations

- A cell contains a value for a particular variable that is part of a part of a particular observation

SPSS Views

- Data View – displays the actual values of the data set

- Variable View – contains the descriptions of each variable's attributes in the data file

List at least three attributes of a variable from the Variable View

Dog Dataset Example

Breed	Age	Weight
Collie	2	23.2
Collie	3	35.7
Setter	5	45.4
Shepard	1	65.9
Setter	2	72.2

Dataset Questions

- Using the SPSS Tutorial, SPSS Help, or Web define each of the following terms and give a real life example of each. SPSS contains the following data types (**variable measurement level**):
    - Categorical/Qualitative Variables
        - Nominal
        - Ordinal
    - Quantitative Variables
        - Scale

Qualitative vs. Quantitative

- Qualitative: classify individuals into categories
- Quantitative: tell how much or how many of something there is

- Which are qualitative and which are quantitative?
  - Person's Age
  - Person's Gender
  - Mileage (in miles per gallon) of a car
  - Color of a car

Qualitative: Ordinal vs. Nominal

- Ordinal variables:
  - One whose categories have a natural ordering
  - Example: grades

- Nominal variables:
  - One whose categories have no natural ordering
  - Example: state of residence

Quantitative

- Discrete variables: Variables whose possible values can be listed

  – Example: number of children

- Continuous variables: Variables that can take any value in an interval

  – Example: height of a person

Both have the measure: scale

Dataset Questions

- Using the SPSS, SPSS Tutorial, and SPSS Help:
  - What are the **types** available in the Variable View?

SPSS

- We can build the Dog data sheet together
  - Variable View

Name	Type	Measure
Breed		
Age		
Weight		

  - Enter Data
    (see next slide)

SPSS

Breed	Age	Weight
Collie	2	23.2
Collie	3	35.7
Setter	5	45.4
Shepard	1	65.9
Setter	2	72.2

SPSS

- Save the data that you just created

- What is the file extension?

- How do you open the file again?

- Notice the output window. What is its purpose?

Candy Dataset Example

Brand	Name	ServingPerPkg	OzPerPkg	Calories	TotalFatInGrams	SatFatInGrams
M&M/Mars	Snickers Peanut Butter	1.0	2.00	310	20.0	7.0
Hershey	Cookies 'n Mint	1.0	1.55	230	12.0	6.0
Hershey	Cadbury Dairy Milk	3.5	5.00	220	12.0	8.0
M&M/Mars	Snickers	3.0	3.70	170	8.0	3.0
Charms	Sugar Daddy	1.0	1.70	200	2.5	2.5

More Dataset Questions

- For the given dataset, what is the <u>type</u> and <u>measure</u> for the data for each of the variables? Why?
  - Brand
  - Name
  - ServingPerPkg
  - OzPerPkg
  - Calories
  - TotalFatInGrams
  - SatFatInGrams

Problem 8.1

Create the dataset Candy8.1 in SPSS from the Candy data given to you on the previous slide

- Create the variables using the Variable View. Make sure that each variable has the correct **Type** and **Measure**.

- Set the decimals column as follows: Brand: 0, Name: 0, ServingPerPkg: 1, OzPerPkg: 2, Calories: 0, TotalFatInGrams: 1, and SatFatInGrams: 1.

Problem 8.1 (continued)

- In the Values column, create the Value Labels for Brand where: value is 1 and Label is "M&M/Mars". Add labels for 2 = "Hershey", and 3 = "Charms".
- Change to Data View and enter the candy data.
    - When you enter the Brand, select View Value Labels from the View menu or toolbar to select M&M, Hershey, or Charms
    - You will need to go back to Variable View and edit some of the settings. Do so as necessary.

Types of Data Analysis

- When doing data analysis, we are interested in two types of summaries:
  - Statistical Summaries (e.g. descriptive, hypothesis testing)

  - Visual Summaries (e.g. tables, graphs)

Areas of Statistics

- **Descriptive Statistics**
  - describe and summarize data

- **Inferential Statistics**
  - Infer from samples
  - e.g. smokers smoking a pack of cigarettes per day have higher cholesterol
  - Hypothesis testing

Descriptive Statistics

- We are concerned, among other things, the following:
  - Mean:

  - Median:

  - Mode:

Mean vs. Median

- Mean is influenced by extreme values unlike the median

- Example: Five families live in an apartment building. Their incomes in dollars are: 25,000, 31,000, 34,000, 44,000, and 56,000. The first family won the $1,000,000 in the lottery. What are the mean and the median?

Problem 8.1 Continued

- Identify each of the following for Total Fat giving your answer to 1 decimal place:
  - Minimum:
  - Maximum:
  - Mean:
  - Standard Deviation:

- How would you find the median and the mode?

Problem 8.2

A paint manufacturer tested two experimental brands of paint over a period of months to determine how long they would last without fading. Here are the results:

Brand A	Brand B	Report on the following
10	25	-Mean
20	35	-Median
60	40	-Mode
40	45	-Std Deviation
50	35	-Minimum
30	30	-Maximum

Solution - Method 1

One way has two variable columns where the first is BrandA and the second is BrandB. Enter the above data and find the asked for information. Save this file as BrandMethod1.sav.

What are the type and measure values for:

BrandA _____ and
BrandB _____

Solution – Method 2

The second way has two columns where the first column is a variable called Brand and the second column is called Fading. Create value labels where 1="BrandA" and 2="BrandB". Enter the information and find the asked for information. Save this file as BrandMethod2.sav.

What are the type and measure values for
      Brand _____ and
      Fading _____

What do the descriptive statistics tell us about the paint with regard to fading?