Introduction to SPSS
"a powerful statistical-analysis and data-management system"

SPSS is a statistical analysis program that allows:

- Data management in a tabular format
- The ability to create graphs and tables
- A broad range of statistical analyses

- SPSS is more *specialized* than Excel, therefore we need to provide data in a more precise way to take advantage of SPSS's functionality.

Goals for this section of the course include:

- Becoming familiar with Statistical Packages and what they can do
- Creating new Datasets
- Importing & exporting Datasets
- Manipulating data in a Dataset
- Basic analysis of data (mainly descriptive statistics)
- Producing professional quality output using Word, Equation Editor, PowerPoint and SPSS results.
- An overview of SPSS's advanced features

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

Examining the Help utility within SPSS

Open up SPSS 17.0 for Windows.
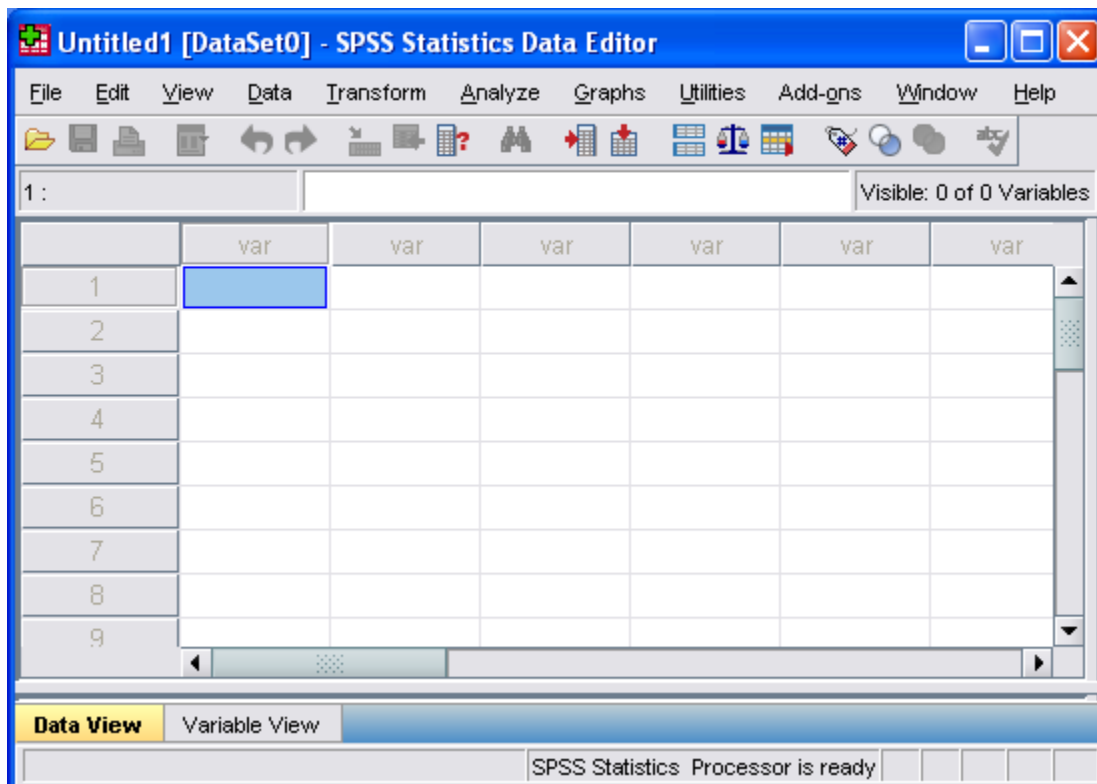
SPSS has a very nice help utility and a tutorial as part of the application. Let's briefly examine each of these features before diving into SPSS. Getting familiar with the help and tutorial features is important as we will only be able to cover a small portion of the features that exist in SPSS.

For instance, go to Help->Topics->Data Files

-
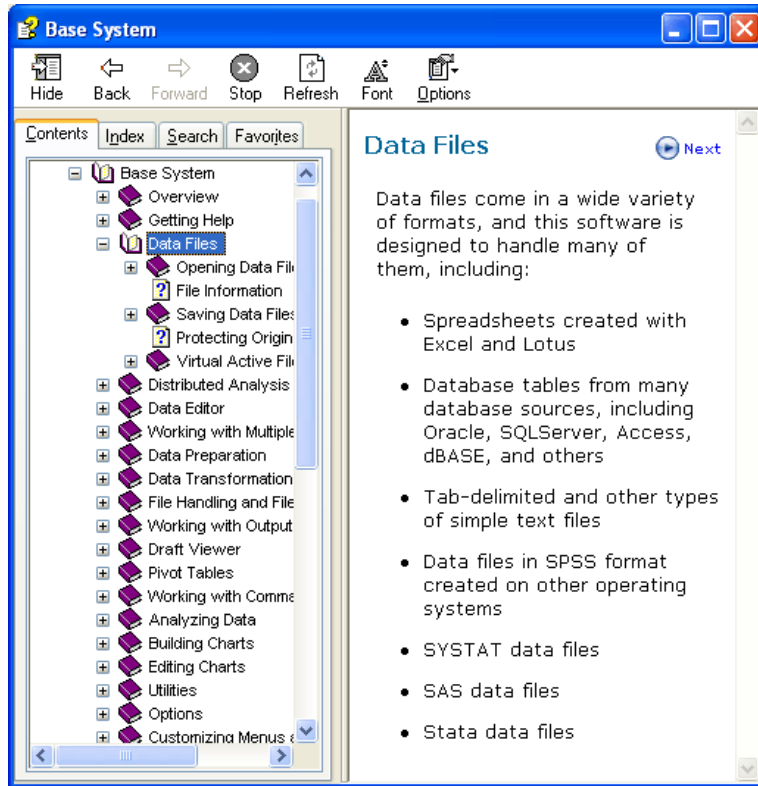
Or go to Help->Tutorial

Creating a Simple Dataset

You notice that SPSS looks somewhat like Excel but there are several important differences:

(1) Column data pertains to a particular variable

Q1: List several examples of what a variable might be.

_____

_____

(2) Row data is considered a case or an observation.

Q1: List several examples of an observation.

_____

_____

 (3) A cell contains a value for a particular variable that is part of a particular observation.

SPSS has two views which you can toggle between by selecting on the correct tab in the lower left-hand corner:

Data View – displays the actual values

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

Q3: List at least three attributes of a variable from the Variable View.

_____

SPSS Worksheet #1

Consider the following dataset

Brand	Name	ServingPerPkg	OzPerPkg	Calories	TotalFatInGrams	SatFatInGrams
M&M/Mars	Snickers Peanut Butter	1.0	2.00	310	20	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

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 (measures):

- Categorical:

  _____
  _____
  _____

  o Nominal:

    _____
    _____
    _____

  o Ordinal

    _____
    _____
    _____

- Scale:

  _____
  _____
  _____

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: _____

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)

Statistics is sometimes broken up into two different areas:

1. **Descriptive Statistics** - a situation is described by the statistics by the collection, summarization, organization and presentation of data.
2. **Inferential Statistics** - where inferences are made from samples of the population (e.g. smokers smoking a pack of cigarettes per day have a higher cholesterol). In this area we get into **Hypothesis testing**.

In the **Descriptive** Statistics world, we are concerned about each of the following.

Give a general description and example of each of the following terms:

Mean _____

Median _____

Mode_____

Problem: Create the dataset from Worksheet #1 using SPSS 17.0 for Windows. Call the data set **candy.sav**.

Step #1: Create the variables using the Variable View. Make sure that each variable has the correct **Type** and **Measure**.

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

Step #3: In the Values column, create the Value Labels for Brand where 1 = "M&M/Mars", 2 = "Hershey", and 3 = "Charms".

Step #4: Change to Data View and enter the data from Worksheet #1. You will need to go back to Variable View and edit some of the settings. Do so as necessary.

Identify each of the following for Total Fat giving your answer to 1 decimal place:

Mean: _____

Standard Deviation: _____

Minimum: _____

Maximum: _____

SPSS Worksheet #2

Problem: 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
10	25
20	35
60	40
40	45
50	35
30	30

You are to report on the following using two different methods.

Mean: _____

Median: _____

Mode: _____

Standard Deviation: _____

Minimum: _____

Maximum: _____

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.

Q1: What are the type and measure values for BrandA _____ and BrandB _____?

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.

Q2: What are the type and measure values for Brand _____ and Fading _____?

Q3: What do the descriptive statistics tell us about the paint with regard to fading?