Intro to SPSS

Fall 2011

Intro to SPSS

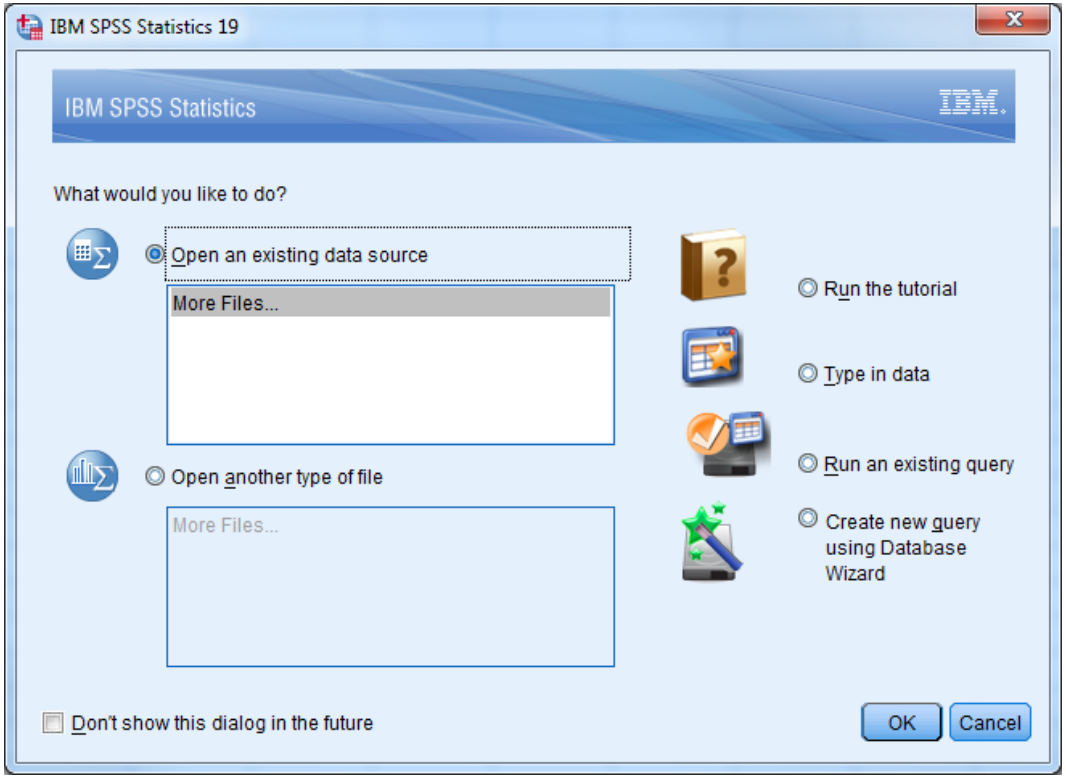
- SPSS is a statistical analysis program that allows:
  - Data management in a spreadsheet-like format
  - The ability to create graphs and tables
  - A broad range of statistical analyses – very detailed analyses
  - You will need to know or remember some basic statistics that I will go over for this section of the course
- SPSS is more *specialized* than Excel, therefore we need to provide data in a more precise way to take advantage of SPSS's functionality.

SPSS

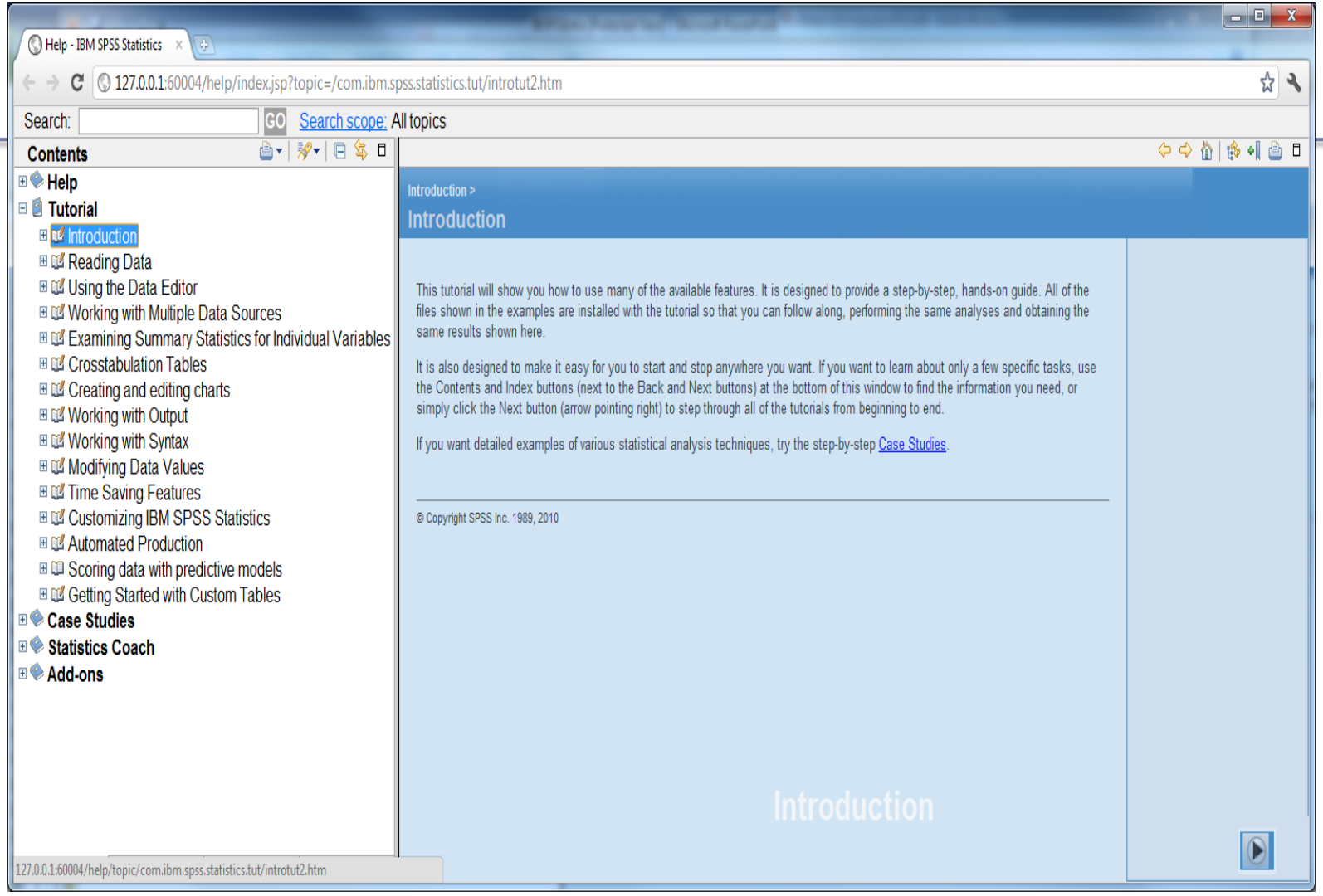
- Goals for this section of the course include:
  - Becoming familiar with Statistical Packages and what they can do
  - Creating new Datasets
  - Importing & exporting Datasets
  - Manipulating data in a Dataset
  - Basic analysis of data (mainly descriptive statistics)
  - An overview of SPSS's advanced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

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

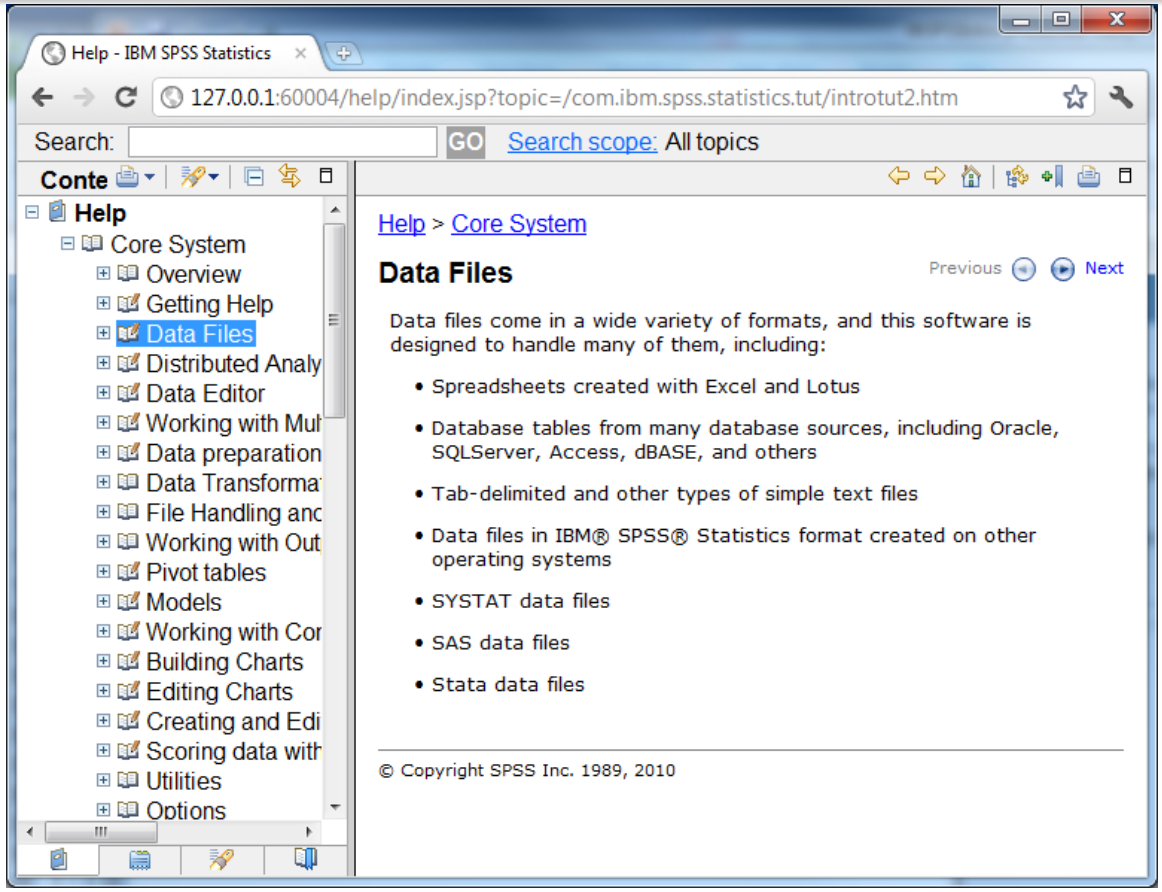
Open PASW


IBM SPSS Statistics 19
IBM SPSS Statistics
IBM
What would you like to do?
Open an existing data source
More Files...
Open another type of file
More Files...
Run the tutorial
Type in data
Run an existing query
Create new query using Database Wizard
Don't show this dialog in the future
OK
Cancel

Run the Tutorial

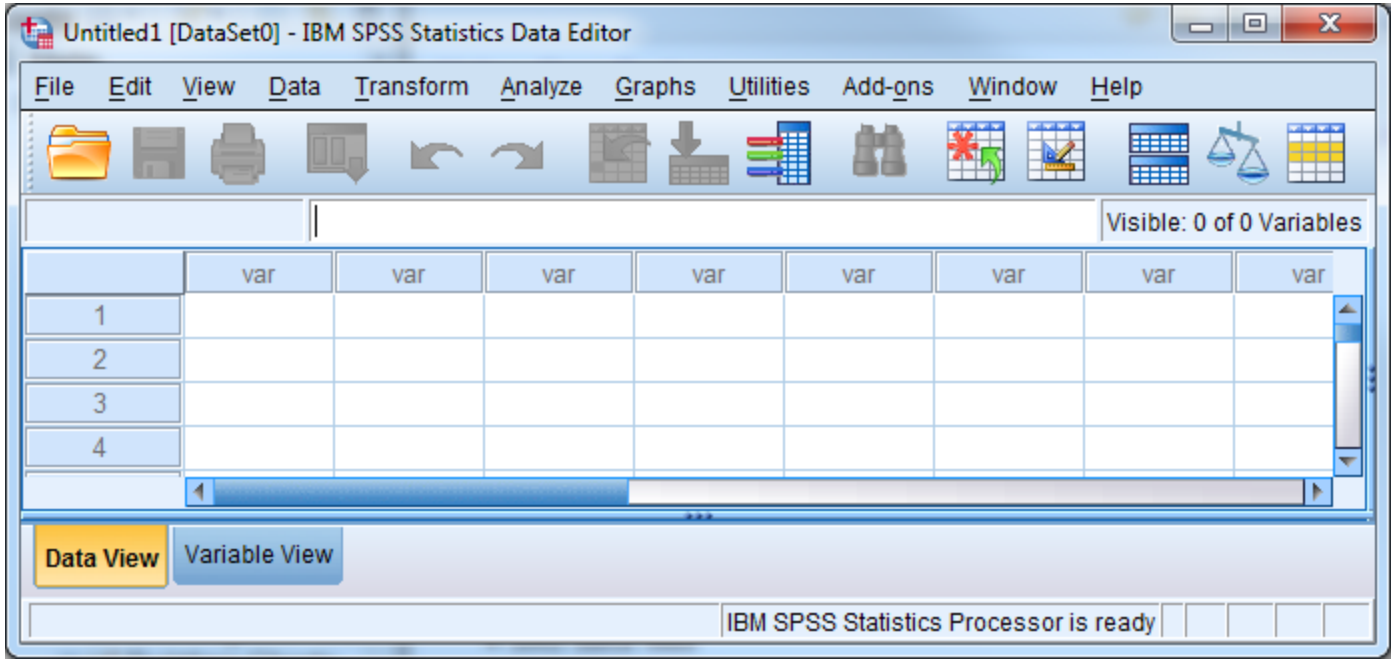


PASW Help


Help - IBM SPSS Statistics
127.0.0.1:60004/help/index.jsp?topic=/com.ibm.spss.statistics.tut/introtut2.htm
Search:
GO
Search scope: All topics
Help
Core System
Overview
Getting Help
Data Files
Distributed Analy
Data Editor
Working with Mul
Data preparation
Data Transforma
File Handling and
Working with Out
Pivot tables
Models
Working with Cor
Building Charts
Editing Charts
Creating and Edi
Scoring data with
Utilities
Options
Help > Core System
Data Files
Previous
Next
Data files come in a wide variety of formats, and this software is designed to handle many of them, including:
• Spreadsheets created with Excel and Lotus
• Database tables from many database sources, including Oracle, SQLServer, Access, dBASE, and others
• Tab-delimited and other types of simple text files
• Data files in IBM® SPSS® Statistics format created on other operating systems
• SYSTAT data files
• SAS data files
• Stata data files
© Copyright SPSS Inc. 1989, 2010

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 or an observation

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 attribute in the data file

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

Candy Dataset Example

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

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 (measures):
  - Categorical
  - Nominal
  - Ordinal
  - Scale

More Dataset Questions

- For the given dataset, what is the type and measure 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)

Areas of Statistics

- **Descriptive Statistics** - a branch of statistics dealing with description and summarization of collections of data
- **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**.

Descriptive Statstics

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

Problem 8.1

Create the dataset Candy8.1 from the Candy dataset using SPSS 19.0

1. Create the variables using the Variable View. Make sure that each variable has the correct **Type** and **Measure**.
2. Set the decimals column as follows: Brand: 0, Name: 0, ServingPerPkg: 1, OzPerPkg: 2, Calories: 0, TotalFatInGrams: 1, and SatFatInGrams: 1.
3. In the Values column, create the Value Labels for Brand where 1 = "M&M/Mars", 2 = "Hershey", and 3 = "Charms".
4. Change to Data View and enter the candy data. You will need to go back to Variable View and edit some of the settings. Do so as necessary.

Problem 8.1 Continued

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

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?