RICE

CCC

Fondren Library

Data @ Rice Workshop Series

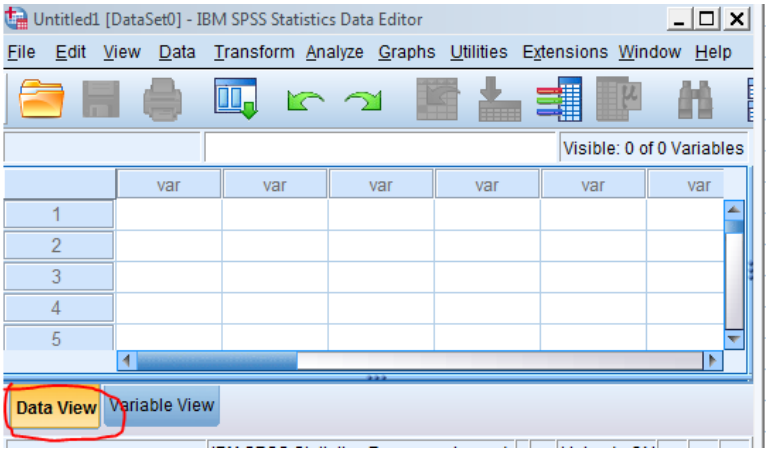
Introduction to SPSS

Overview:

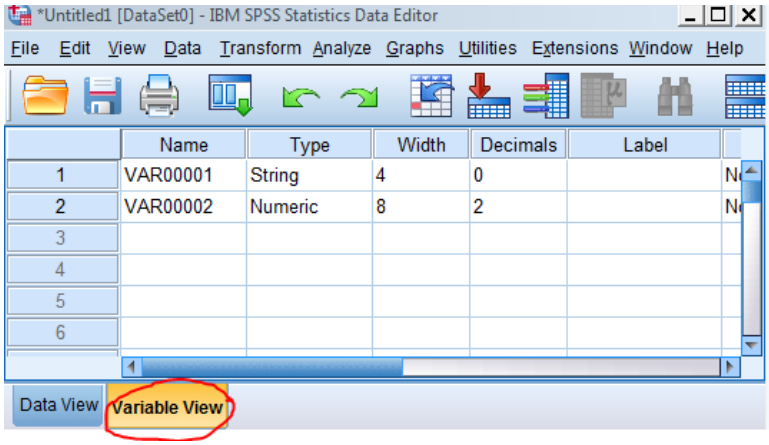
The software name originally stood for **Statistical Package for the Social Sciences** (**SPSS**)

Interface

Data View



Variable View

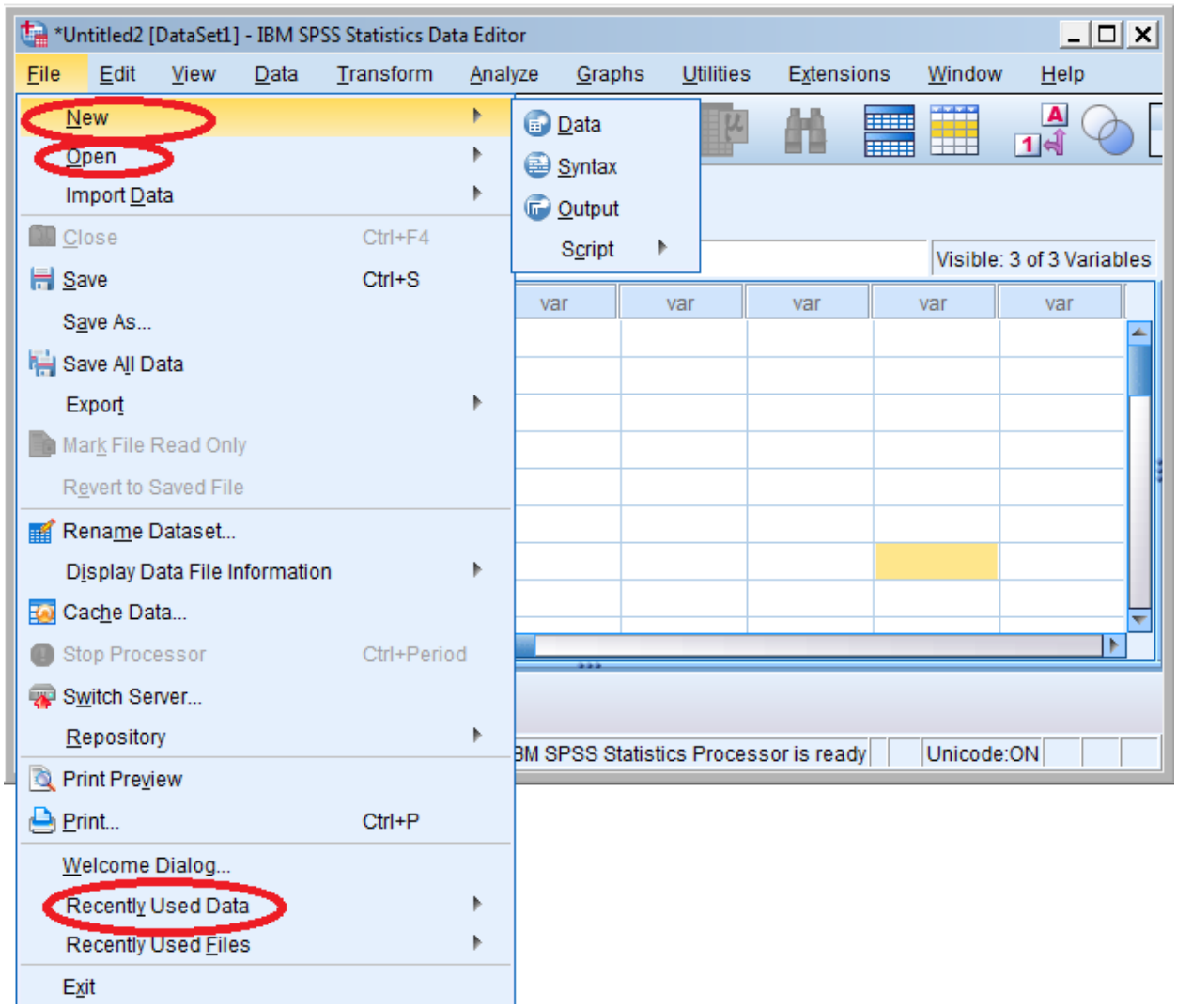


Data entry and change variable properties(name, type, lable, values, measures)

Useful options under the Menus

File

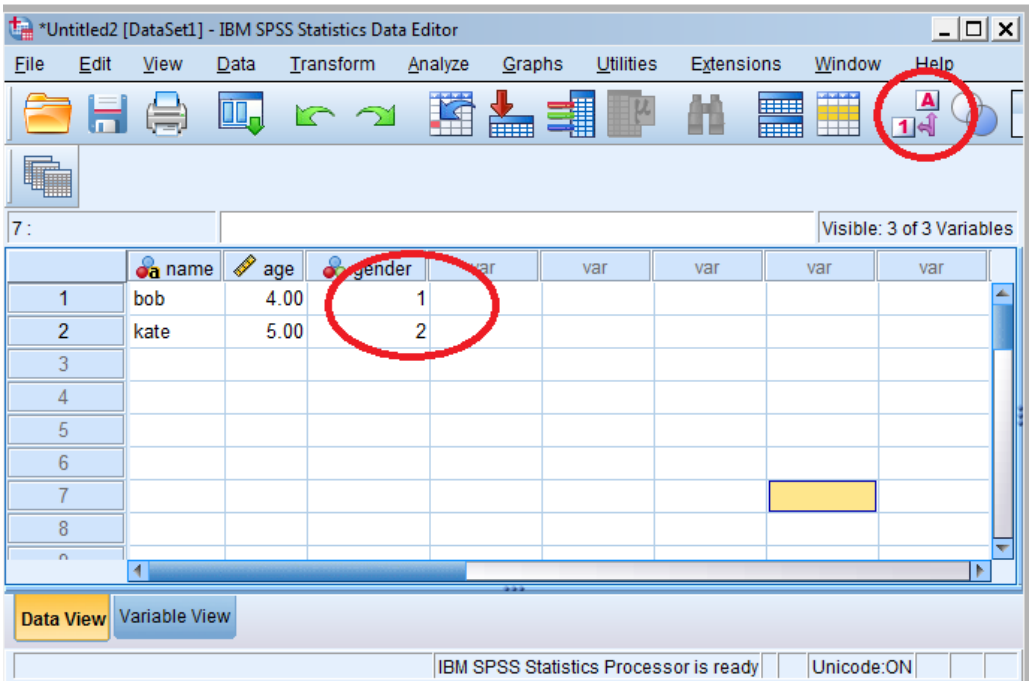
- New
- Open
- Recently used

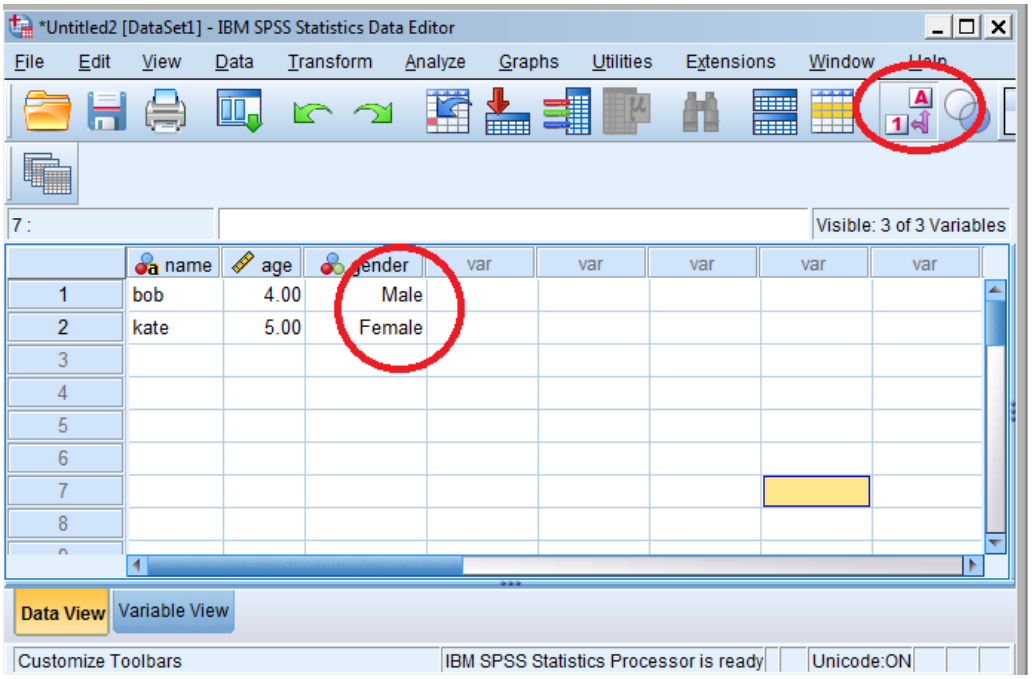


View

- Value lables

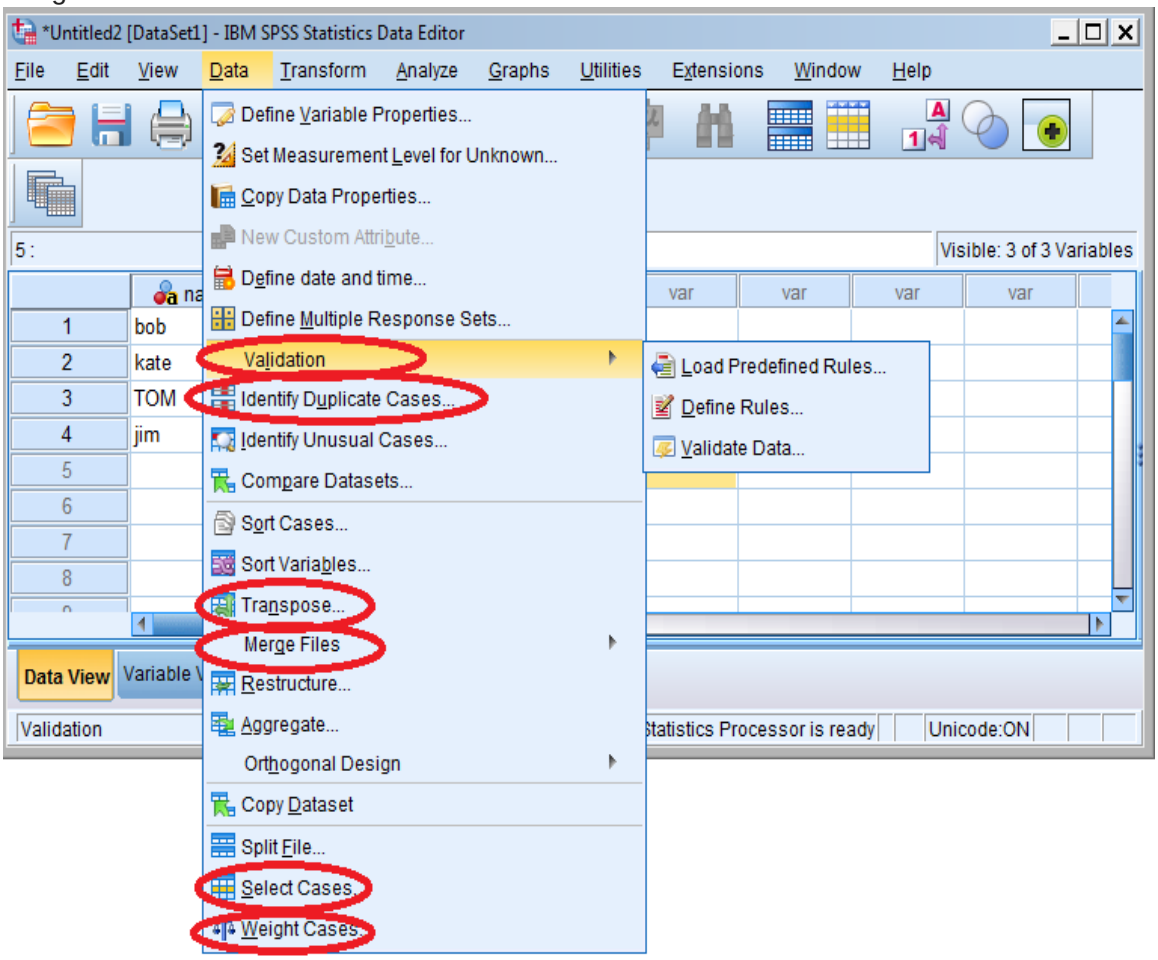
When it's off, it shows the numaric values. When it's on, it shows the text values input in the Variable View





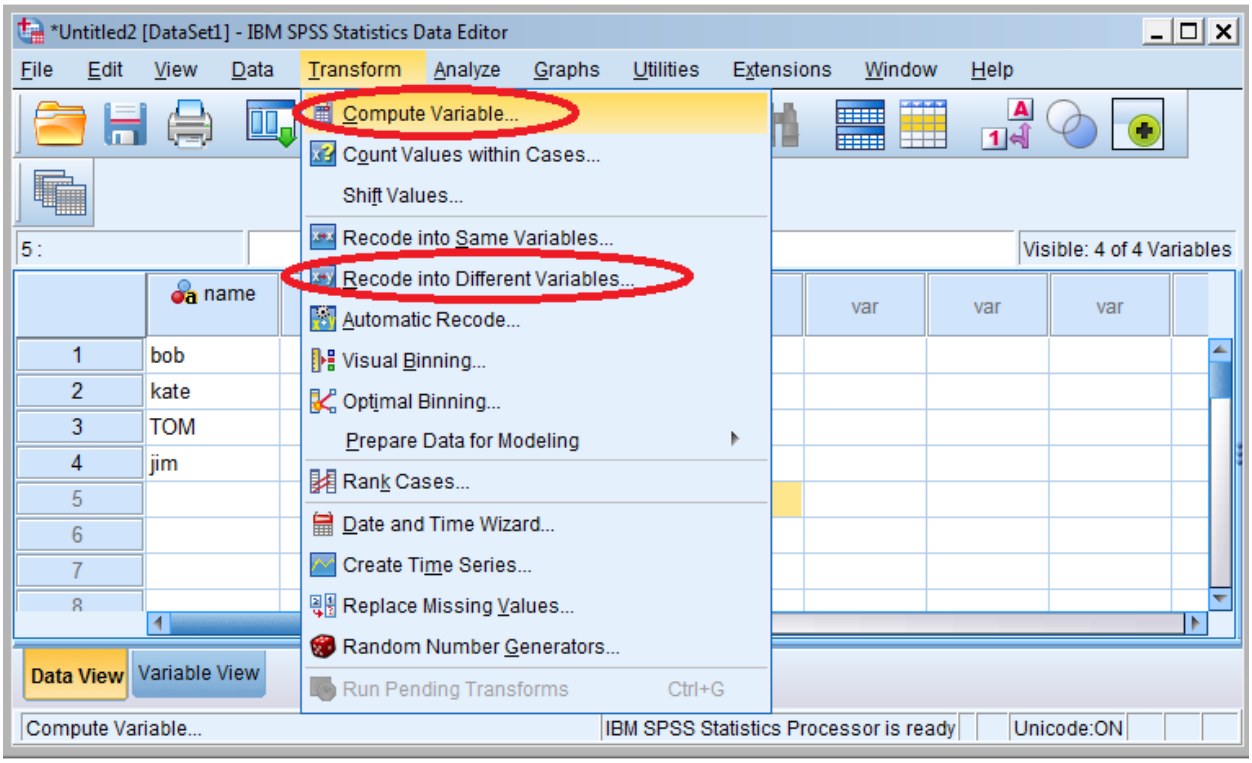
Data

- Identify duplicated case
- Transpose
- Merge files
- Select cases
- Weight cases



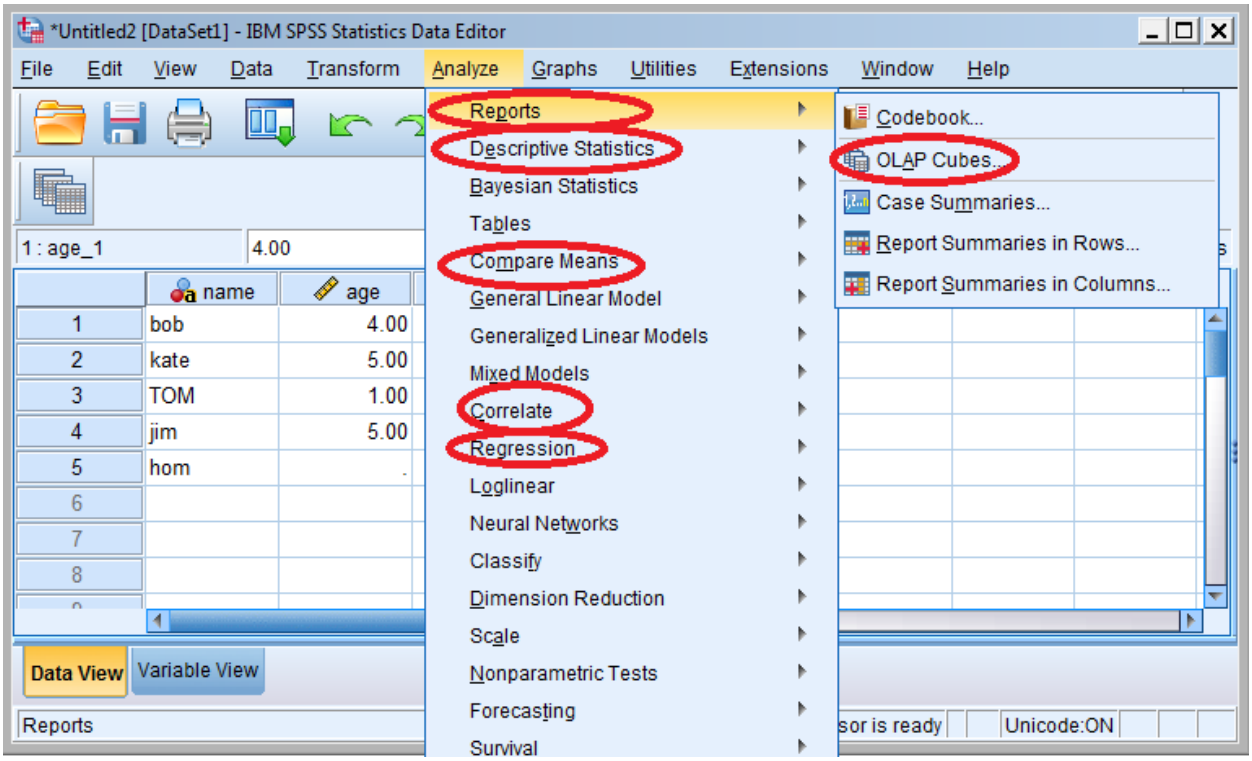
Transformation

- Compute variable
- Recode into different variable



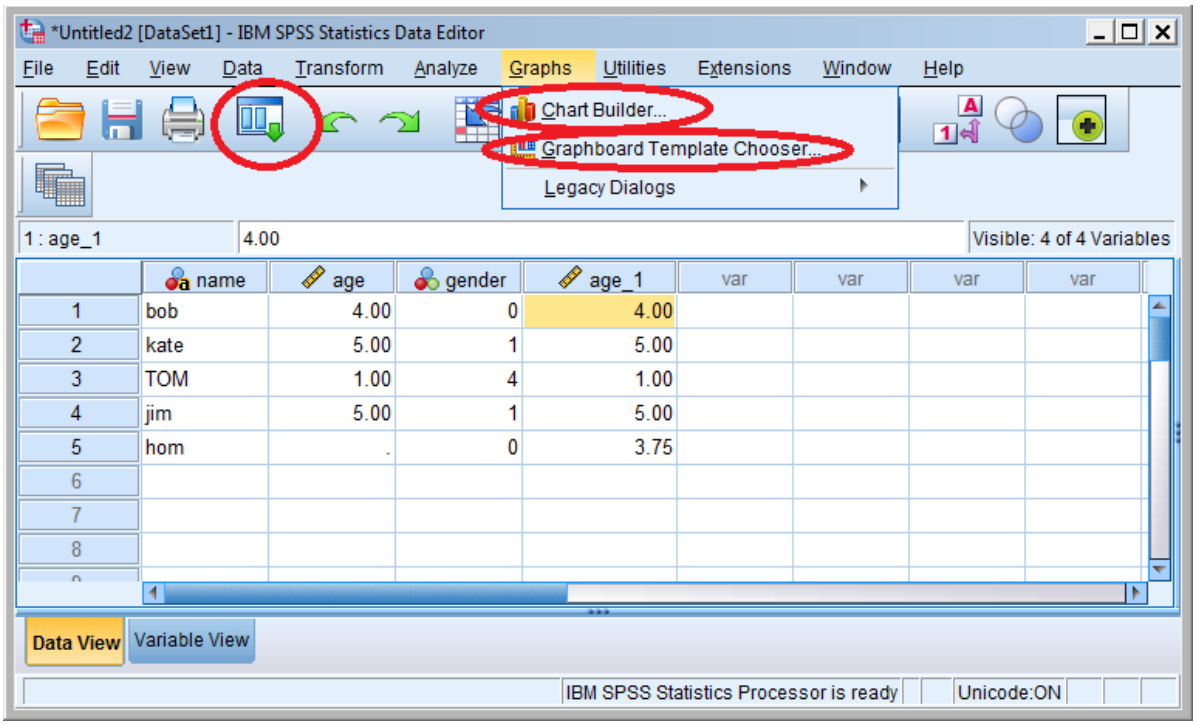
Analyze

- Report-OLAP Cubes
- Descriptive Statistics
- Compare Means
- Correlate
- Regression



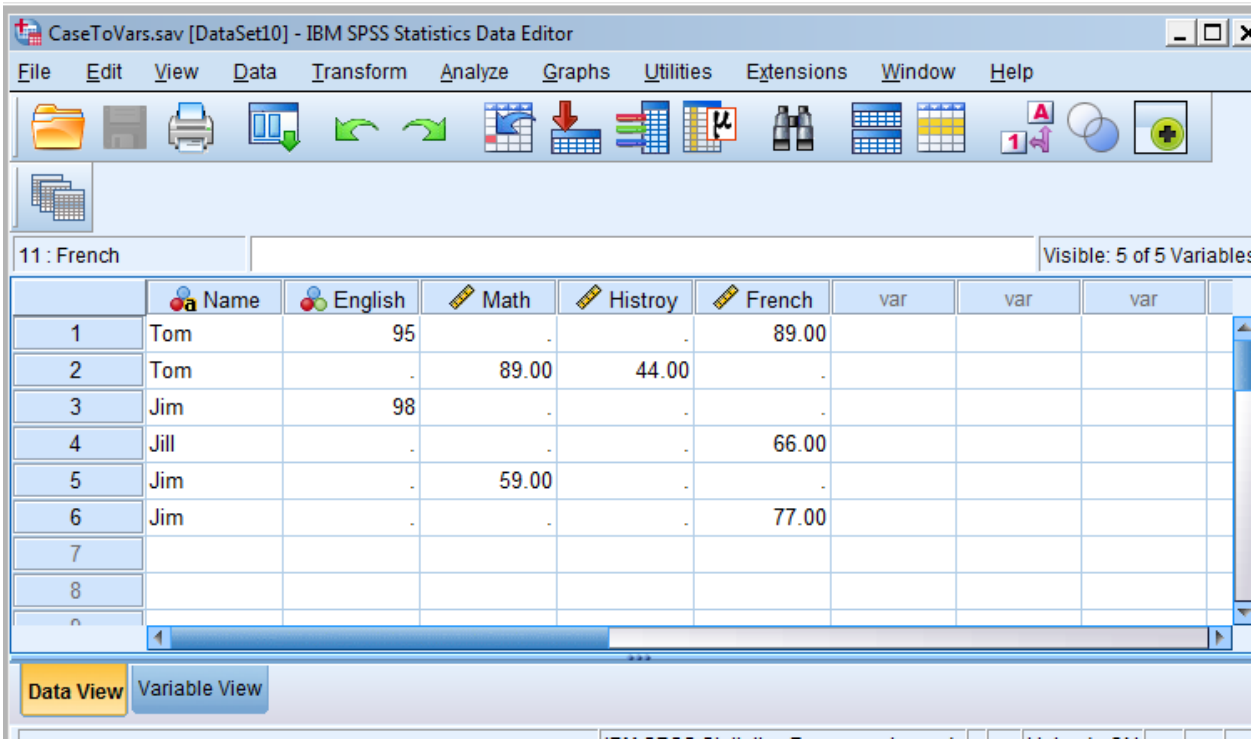
Graphs

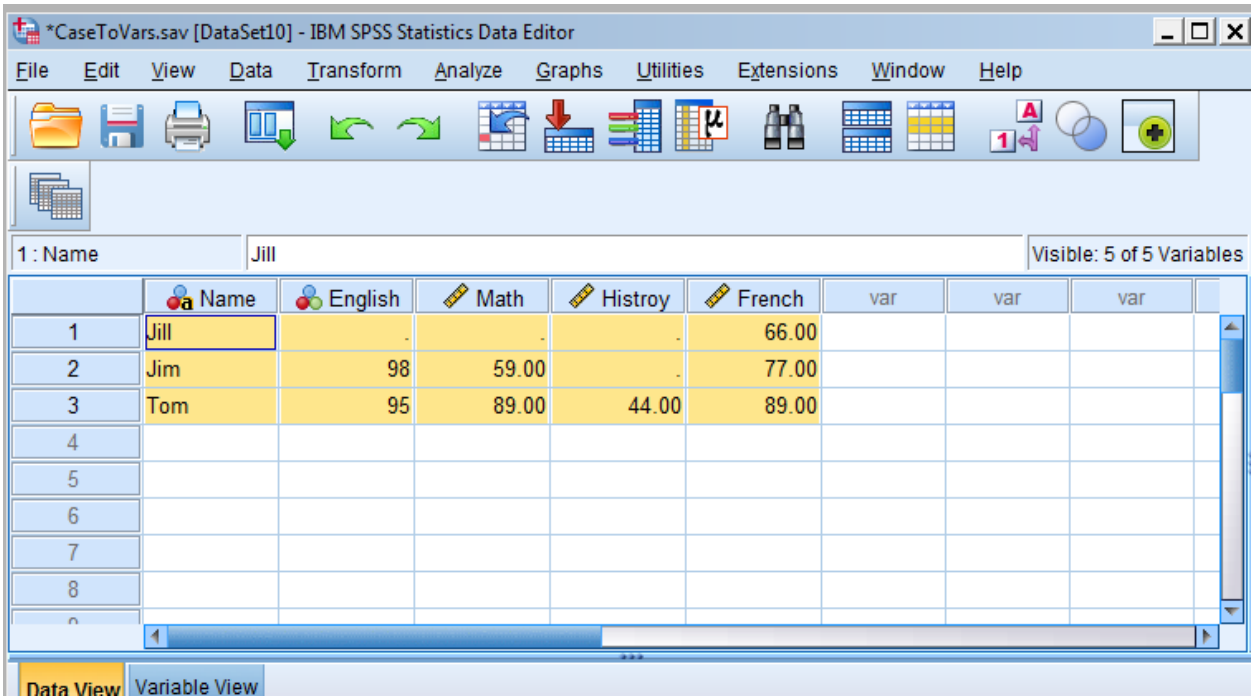
- Chart Builder
- Graphboard Template Chooser



Recently used and Customize

Scrips





Terms

Data type

- String
- Numeric

Level of measures

- Norminal (No ranking, "Race", "Gender")
- Ordinal (Scale, "Strongly Disagree","Strongly Agree")
- Scale ("Age", "Height", "Income")

Cross Tablulation

Cross-tabluations are frequency distributitons for two variables together. It gives you a basic picture of how two variables inter-relate.

Gender Gender * Happy happy Crosstabulation

Count

		happy			
		Happy	Sad	Bored	Total
Gender	Male	3	2	9	14
	Female	10	0	1	11
Total		13	2	10	25

Chi square is used to test the relationship between two norminal or ordinal variables. (If **p-value** is less than **0.05**)

Correlation measures the strength and direction (-1,1) of association between two quantitative variables.

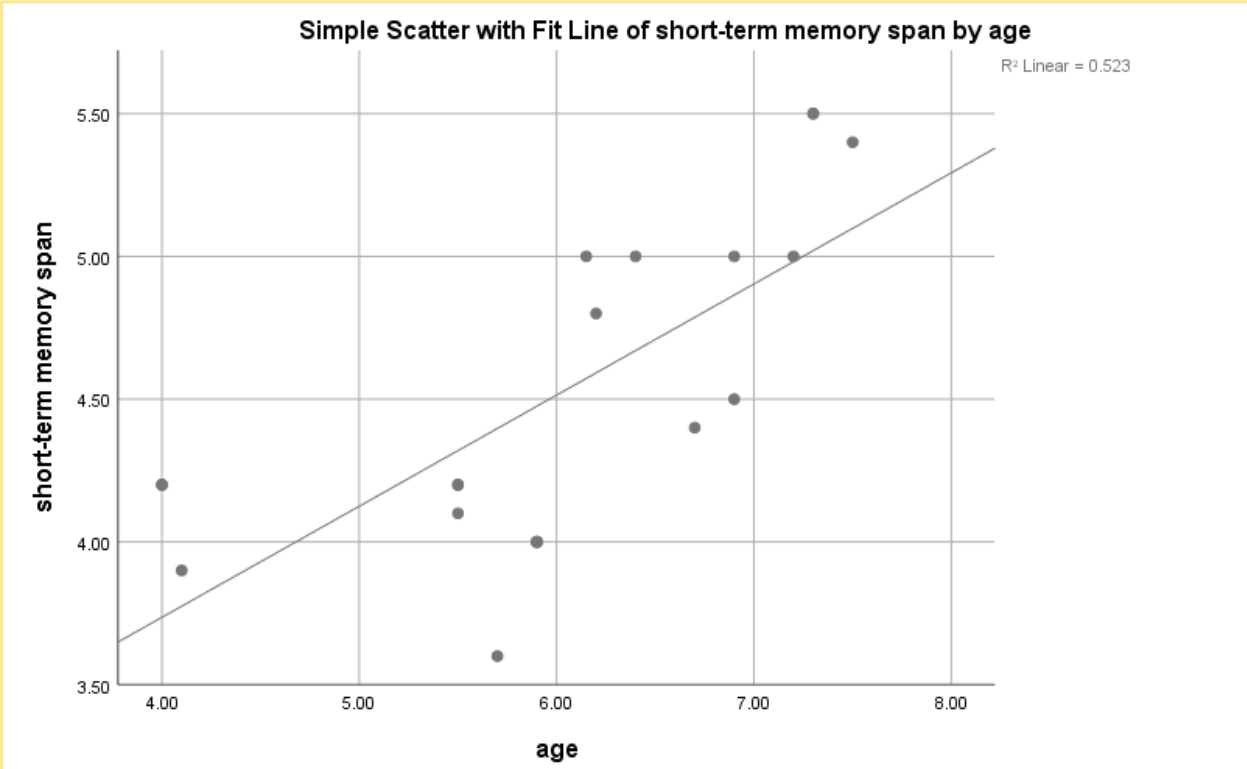
Correlations

		short-term memory span	age
short-term memory span	Pearson Correlation	1	.723**
	Sig. (2-tailed)		.000
	N	20	20
age	Pearson Correlation	.723**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Regression predicts the value one variable(dependent) base on another variable(independent, predictor)

GGraph



Y=2.2+0.39X

Exercise

1. **Find out how many girls vs boys in our dataset**
2. **Make a pie chart for gender distribution**
3. **What's the average age of the children?**
4. **Recode "Age" into "Age1"(below or equal 6 code as 1, above code as 2)**
   Recode "Read Ability" into "ReadAbility1" (below or equal to 6 code as 1, above code as 2)
5. **Crosstab between "Age1" and "ReadAbility1"**
6. **Correlation between "Age" and "Memory Span"**
7. **Linear Regression "Age" and "Memory Span"**
8. **Plot "Age" and "Memory Span" with a trend line**
9. **OLAP Cubes Average Grad in 6 years for Black/White weighted by Cohort counts**