Discover. Gather. Create. Share.

UBC Library Research Commons

# 2016 SPSS Workshop II

Facilitators: Wendy and Sarah

# SPSS services

1. **Consultations one day per week**
2. **Workshops (http://elred.library.ubc.ca/libs)**

- Workshop I

  SPSS Overview(Interface, Basic Concepts)

  Data Entry

  One variable analysis

- Workshop II

  Data Management

  Inferential Statistics

# Today's Agenda

1. Brief review to last week's contents
2. Data Management
3. Inferential statistics

## Overview to SPSS workshop 1

**Data: three levels of measurement**

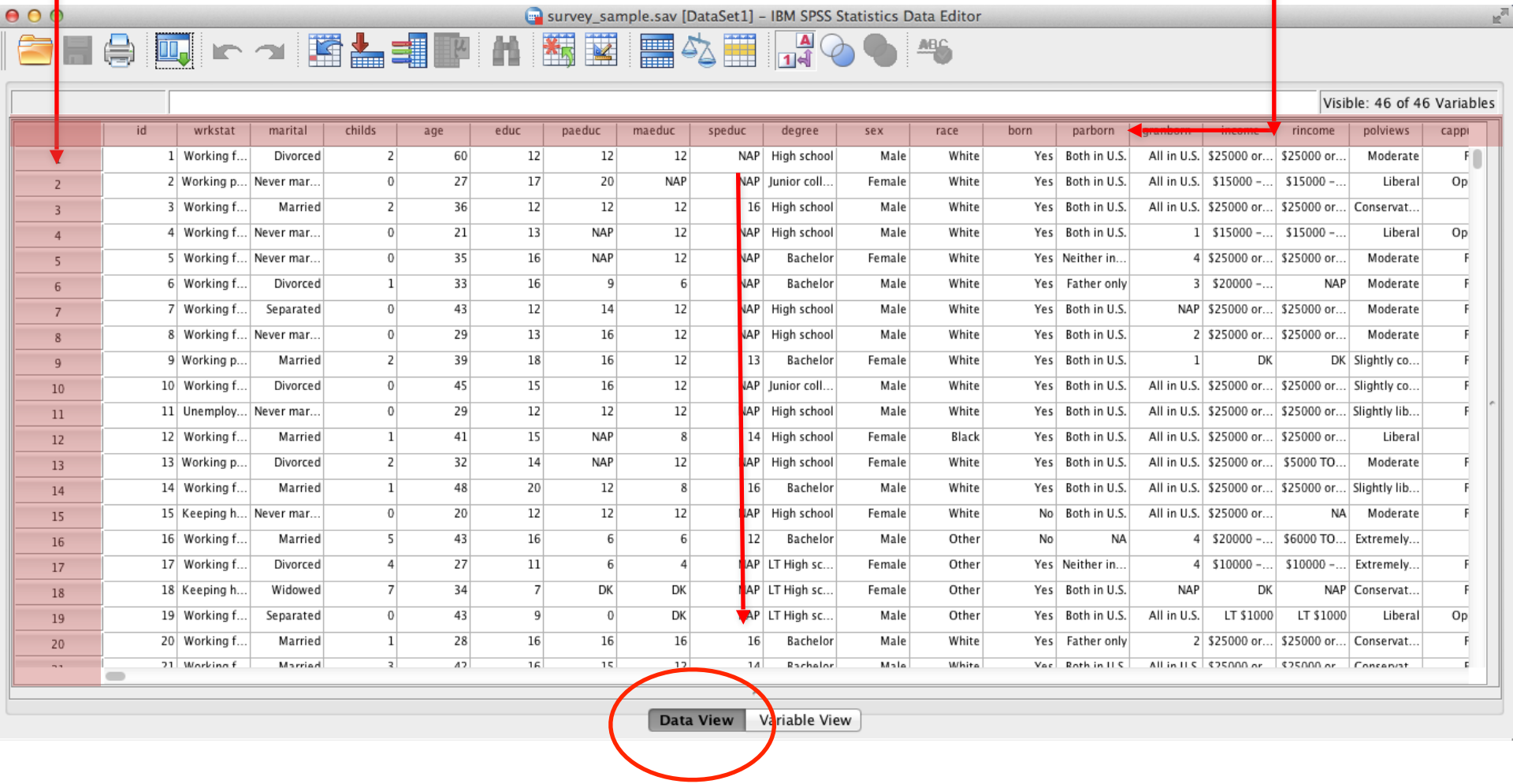
Many SPSS procedures depend on whether the data is qualitative or quantitative, as well as the measurement.

| Basic types of data | Measurement | Example |
|---|---|---|
| Qualitative (Categorical) | Nominal | Nationality, Sex |
| | Ordinal | Rating of a restaurant |
| | | Age Group |
| Quantitative | Interval/Ratio | Age |
| | Scale | |

Rows represent cases

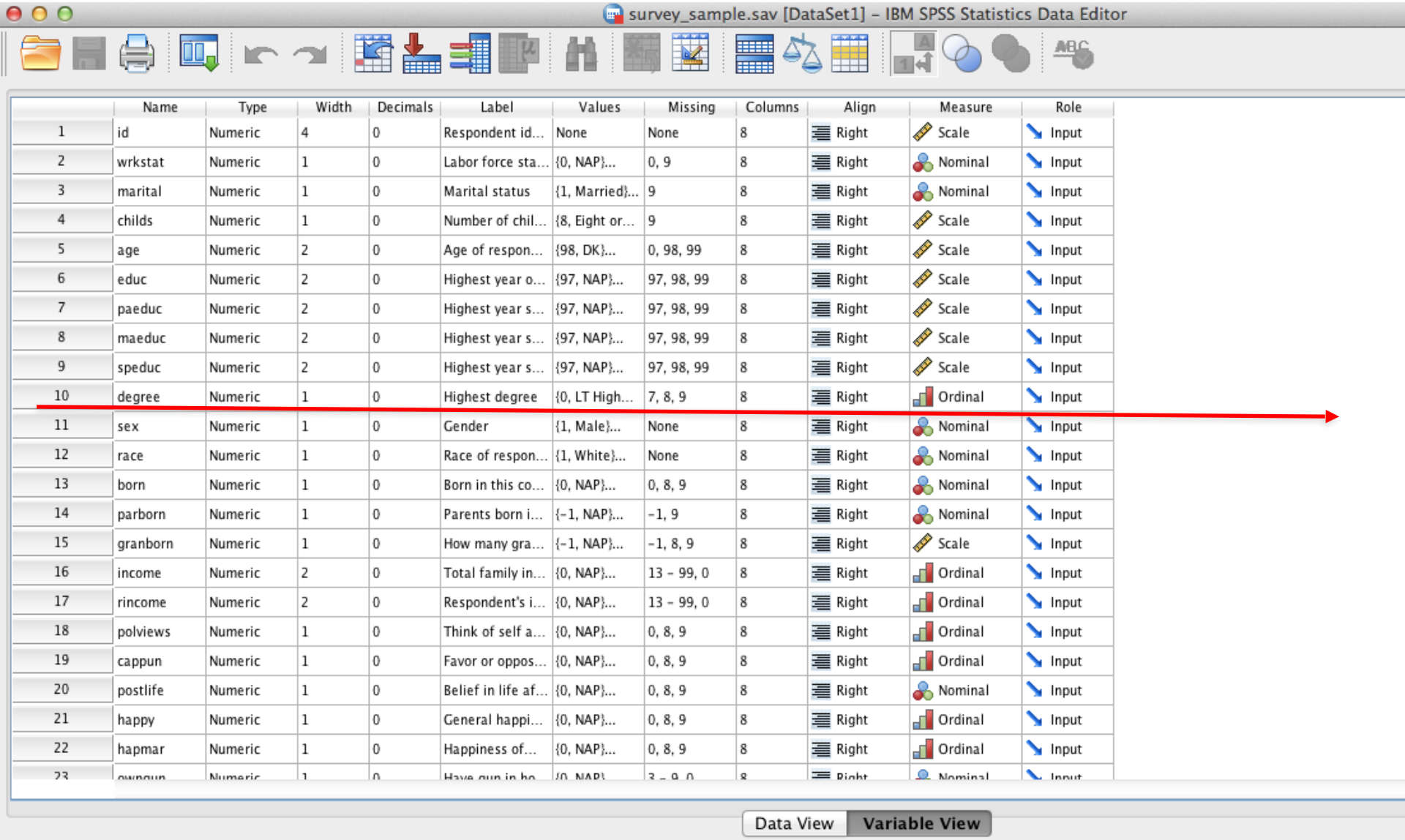
# Data Editor- Data View

Variables




| | id | wrkstat | marital | childs | age | educ | paeduc | maeduc | speduc | degree | sex | race | born | parborn | granborn | income | rincome | polviews | cappi |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | Working f... | Divorced | 2 | 60 | 12 | 12 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Moderate | F |
| 2 | 2 | Working p... | Never mar... | 0 | 27 | 17 | 20 | NAP | NAP | Junior coll... | Female | White | Yes | Both in U.S. | All in U.S. | $15000 -... | $15000 -... | Liberal | Op |
| 3 | 3 | Working f... | Married | 2 | 36 | 12 | 12 | 12 | 16 | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Conservat... | |
| 4 | 4 | Working f... | Never mar... | 0 | 21 | 13 | NAP | 12 | NAP | High school | Male | White | Yes | Both in U.S. | 1 | $15000 -... | $15000 -... | Liberal | Op |
| 5 | 5 | Working f... | Never mar... | 0 | 35 | 16 | NAP | 12 | NAP | Bachelor | Female | White | Yes | Neither in... | 4 | $25000 or... | $25000 or... | Moderate | F |
| 6 | 6 | Working f... | Divorced | 1 | 33 | 16 | 9 | 6 | NAP | Bachelor | Male | White | Yes | Father only | 3 | $20000 -... | NAP | Moderate | F |
| 7 | 7 | Working f... | Separated | 0 | 43 | 12 | 14 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | NAP | $25000 or... | $25000 or... | Moderate | F |
| 8 | 8 | Working f... | Never mar... | 0 | 29 | 13 | 16 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | 2 | $25000 or... | $25000 or... | Moderate | F |
| 9 | 9 | Working p... | Married | 2 | 39 | 18 | 16 | 12 | 13 | Bachelor | Female | White | Yes | Both in U.S. | 1 | DK | DK | Slightly co... | F |
| 10 | 10 | Working f... | Divorced | 0 | 45 | 15 | 16 | 12 | NAP | Junior coll... | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly co... | F |
| 11 | 11 | Unemploy... | Never mar... | 0 | 29 | 12 | 12 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly lib... | F |
| 12 | 12 | Working f... | Married | 1 | 41 | 15 | NAP | 8 | 14 | High school | Female | Black | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Liberal | |
| 13 | 13 | Working p... | Divorced | 2 | 32 | 14 | NAP | 12 | NAP | High school | Female | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $5000 TO... | Moderate | F |
| 14 | 14 | Working f... | Married | 1 | 48 | 20 | 12 | 8 | 16 | Bachelor | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly lib... | F |
| 15 | 15 | Keeping h... | Never mar... | 0 | 20 | 12 | 12 | 12 | NAP | High school | Female | White | No | Both in U.S. | All in U.S. | $25000 or... | NA | Moderate | F |
| 16 | 16 | Working f... | Married | 5 | 43 | 16 | 6 | 6 | 12 | Bachelor | Male | Other | No | NA | 4 | $20000 -... | $6000 TO... | Extremely... | |
| 17 | 17 | Working f... | Divorced | 4 | 27 | 11 | 6 | 4 | NAP | LT High sc... | Female | Other | Yes | Neither in... | 4 | $10000 -... | $10000 -... | Extremely... | F |
| 18 | 18 | Keeping h... | Widowed | 7 | 34 | 7 | DK | DK | NAP | LT High sc... | Female | Other | Yes | Both in U.S. | NAP | DK | NAP | Conservat... | F |
| 19 | 19 | Working f... | Separated | 0 | 43 | 9 | 0 | DK | NAP | LT High sc... | Male | Other | Yes | Both in U.S. | All in U.S. | LT $1000 | LT $1000 | Liberal | Op |
| 20 | 20 | Working f... | Married | 1 | 28 | 16 | 16 | 16 | 16 | Bachelor | Male | White | Yes | Father only | 2 | $25000 or... | $25000 or... | Conservat... | F |
| 21 | 21 | Working f... | Married | 3 | 42 | 16 | 15 | 12 | 14 | Bachelor | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Conservat... | F |

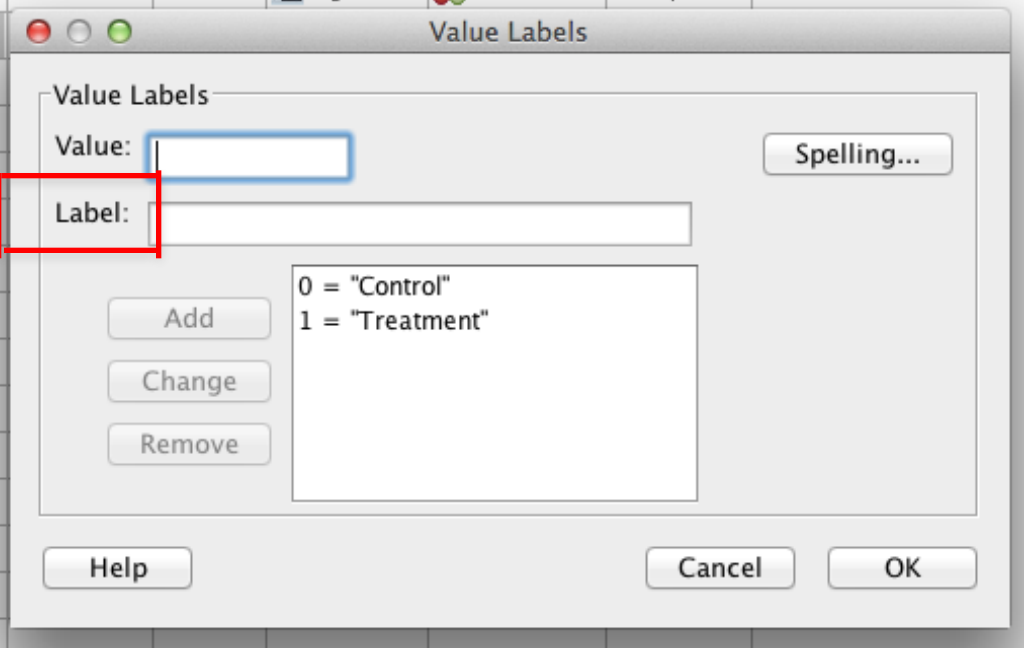
# Data Editor- Variable View




| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure | Role |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | id | Numeric | 4 | 0 | Respondent id... | None | None | 8 | Right | Scale | Input |
| 2 | wrkstat | Numeric | 1 | 0 | Labor force sta... | {0, NAP}... | 0, 9 | 8 | Right | Nominal | Input |
| 3 | marital | Numeric | 1 | 0 | Marital status | {1, Married}... | 9 | 8 | Right | Nominal | Input |
| 4 | childs | Numeric | 1 | 0 | Number of chil... | {8, Eight or... | 9 | 8 | Right | Scale | Input |
| 5 | age | Numeric | 2 | 0 | Age of respon... | {98, DK}... | 0, 98, 99 | 8 | Right | Scale | Input |
| 6 | educ | Numeric | 2 | 0 | Highest year o... | {97, NAP}... | 97, 98, 99 | 8 | Right | Scale | Input |
| 7 | paeduc | Numeric | 2 | 0 | Highest year s... | {97, NAP}... | 97, 98, 99 | 8 | Right | Scale | Input |
| 8 | maeduc | Numeric | 2 | 0 | Highest year s... | {97, NAP}... | 97, 98, 99 | 8 | Right | Scale | Input |
| 9 | speduc | Numeric | 2 | 0 | Highest year s... | {97, NAP}... | 97, 98, 99 | 8 | Right | Scale | Input |
| 10 | degree | Numeric | 1 | 0 | Highest degree | {0, LT High... | 7, 8, 9 | 8 | Right | Ordinal | Input |
| 11 | sex | Numeric | 1 | 0 | Gender | {1, Male}... | None | 8 | Right | Nominal | Input |
| 12 | race | Numeric | 1 | 0 | Race of respon... | {1, White}... | None | 8 | Right | Nominal | Input |
| 13 | born | Numeric | 1 | 0 | Born in this co... | {0, NAP}... | 0, 8, 9 | 8 | Right | Nominal | Input |
| 14 | parborn | Numeric | 1 | 0 | Parents born i... | {-1, NAP}... | -1, 9 | 8 | Right | Nominal | Input |
| 15 | granborn | Numeric | 1 | 0 | How many gra... | {-1, NAP}... | -1, 8, 9 | 8 | Right | Scale | Input |
| 16 | income | Numeric | 2 | 0 | Total family in... | {0, NAP}... | 13 - 99, 0 | 8 | Right | Ordinal | Input |
| 17 | rincome | Numeric | 2 | 0 | Respondent's i... | {0, NAP}... | 13 - 99, 0 | 8 | Right | Ordinal | Input |
| 18 | polviews | Numeric | 1 | 0 | Think of self a... | {0, NAP}... | 0, 8, 9 | 8 | Right | Ordinal | Input |
| 19 | cappun | Numeric | 1 | 0 | Favor or oppos... | {0, NAP}... | 0, 8, 9 | 8 | Right | Ordinal | Input |
| 20 | postlife | Numeric | 1 | 0 | Belief in life af... | {0, NAP}... | 0, 8, 9 | 8 | Right | Nominal | Input |
| 21 | happy | Numeric | 1 | 0 | General happi... | {0, NAP}... | 0, 8, 9 | 8 | Right | Ordinal | Input |
| 22 | hapmar | Numeric | 1 | 0 | Happiness of... | {0, NAP}... | 0, 8, 9 | 8 | Right | Ordinal | Input |
| 23 | owngun | Numeric | 1 | 0 | Have gun in ho... | {0, NAP}... | 3 - 9, 0 | 8 | Right | Nominal | Input |

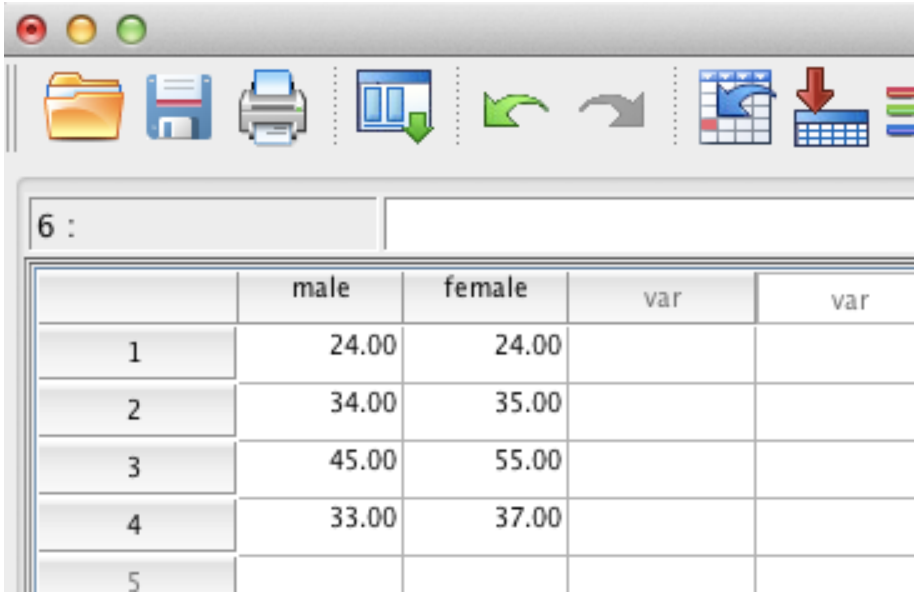
# Variable Name, Variable Label, and Variable Values (codes and label of the value)

| Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure | Role |
|---|---|---|---|---|---|---|---|---|---|---|
| id | Numeric | 2 | 0 | Pt. ID | None | None | 8 | Right | Nominal | Input |
| group | Numeric | 2 | 0 | Treatment group | {0, Control}... | | | | | |
| gender | Numeric | 2 | 0 | Female pts. | None | | | | | |
| age | Numeric | 2 | 0 | Pt. age | None | | | | | |
| los | Numeric | 2 | 0 | Hospital LOS | None | | | | | |
| diabetic | Numeric | 2 | 0 | Diabetes mellitus | {0, No}... | | | | | |
| hypertns | Numeric | 2 | 0 | Hypertensive | {0, No}... | | | | | |
| afib | Numeric | 2 | 0 | Atrial fibrillation | {0, No}... | | | | | |
| priorstr | Numeric | 2 | 0 | Prior stroke | {0, No}... | | | | | |
| smoker | Numeric | 2 | 0 | Current smoker | {0, No}... | | | | | |
| psd | Numeric | 2 | 0 | Post-stroke depres... | {0, No}... | | | | | |
| travel | Numeric | 2 | 0 | Travel ADL | {0, Same as... | | | | | |
| cooking | Numeric | 2 | 0 | Cooking ADL | {0, Plans an... | | | | | |
| housekpg | Numeric | 2 | 0 | Housekeeping ADL | {0, As befor... | | | | | |



# Things to keep in mind in data entry

1. For nominal data, you may need to use this a grouping variable when entering the data in SPSS

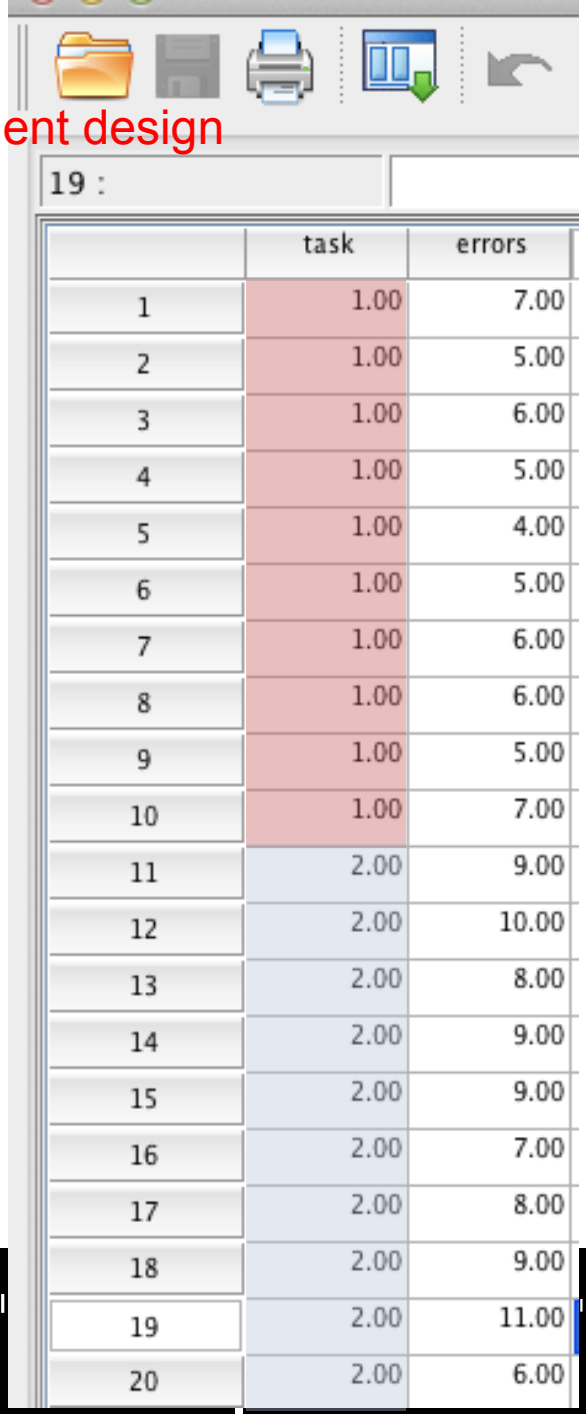


6 :

| | male | female | var | var |
|---|---|---|---|---|
| 1 | 24.00 | 24.00 | | |
| 2 | 34.00 | 35.00 | | |
| 3 | 45.00 | 55.00 | | |
| 4 | 33.00 | 37.00 | | |
| 5 | | | | |

Is this data entry correct?

Independent design



19 :

| | task | errors |
|---|---|---|
| 1 | 1.00 | 7.00 |
| 2 | 1.00 | 5.00 |
| 3 | 1.00 | 6.00 |
| 4 | 1.00 | 5.00 |
| 5 | 1.00 | 4.00 |
| 6 | 1.00 | 5.00 |
| 7 | 1.00 | 6.00 |
| 8 | 1.00 | 6.00 |
| 9 | 1.00 | 5.00 |
| 10 | 1.00 | 7.00 |
| 11 | 2.00 | 9.00 |
| 12 | 2.00 | 10.00 |
| 13 | 2.00 | 8.00 |
| 14 | 2.00 | 9.00 |
| 15 | 2.00 | 9.00 |
| 16 | 2.00 | 7.00 |
| 17 | 2.00 | 8.00 |
| 18 | 2.00 | 9.00 |
| 19 | 2.00 | 11.00 |
| 20 | 2.00 | 6.00 |

13 :

| | daytime | nighttime |
|---|---|---|
| 1 | 7.00 | 8.00 |
| 2 | 9.00 | 10.00 |
| 3 | 6.00 | 8.00 |
| 4 | 5.00 | 7.00 |
| 5 | 8.00 | 8.00 |
| 6 | 6.00 | 10.00 |
| 7 | 9.00 | 11.00 |
| 8 | 6.00 | 7.00 |
| 9 | 5.00 | 8.00 |
| 10 | 7.00 | 7.00 |

Repeated measures/ related design

# Data Management

**Why do you need to management data?**

- To analyze variables that are not in the original data file but are derived from the original variables 1
- To display the data differently
- To code variables differently
- To perform analyses on only a certain portion of the data set

# The Sort Cases command

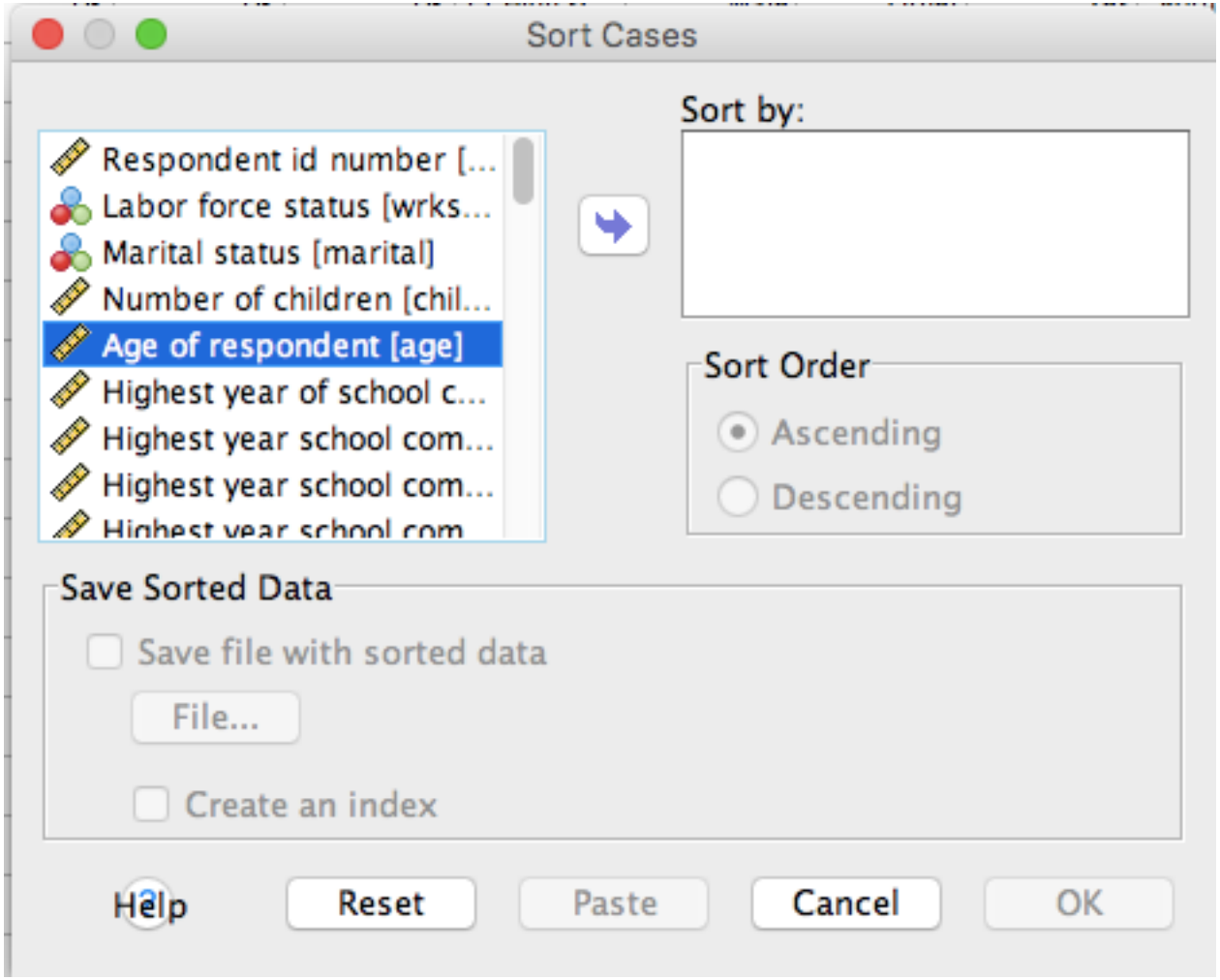
**What**: rearranging your data

**When:**

◇you may want to examine more closely some specific cases which may be easier to find if they are sorted by age, by grades, etc.

◇Other procedures such as Merge require the data to be sorted first.

# Menu: Data→ Sort Cases



## Demonstration

Example:
You may want to list the respondents by age.

**Program Files-IBM-SPSS-Statistics-23-Samples-English-Survey_sample.sav**

# The Select Cases command

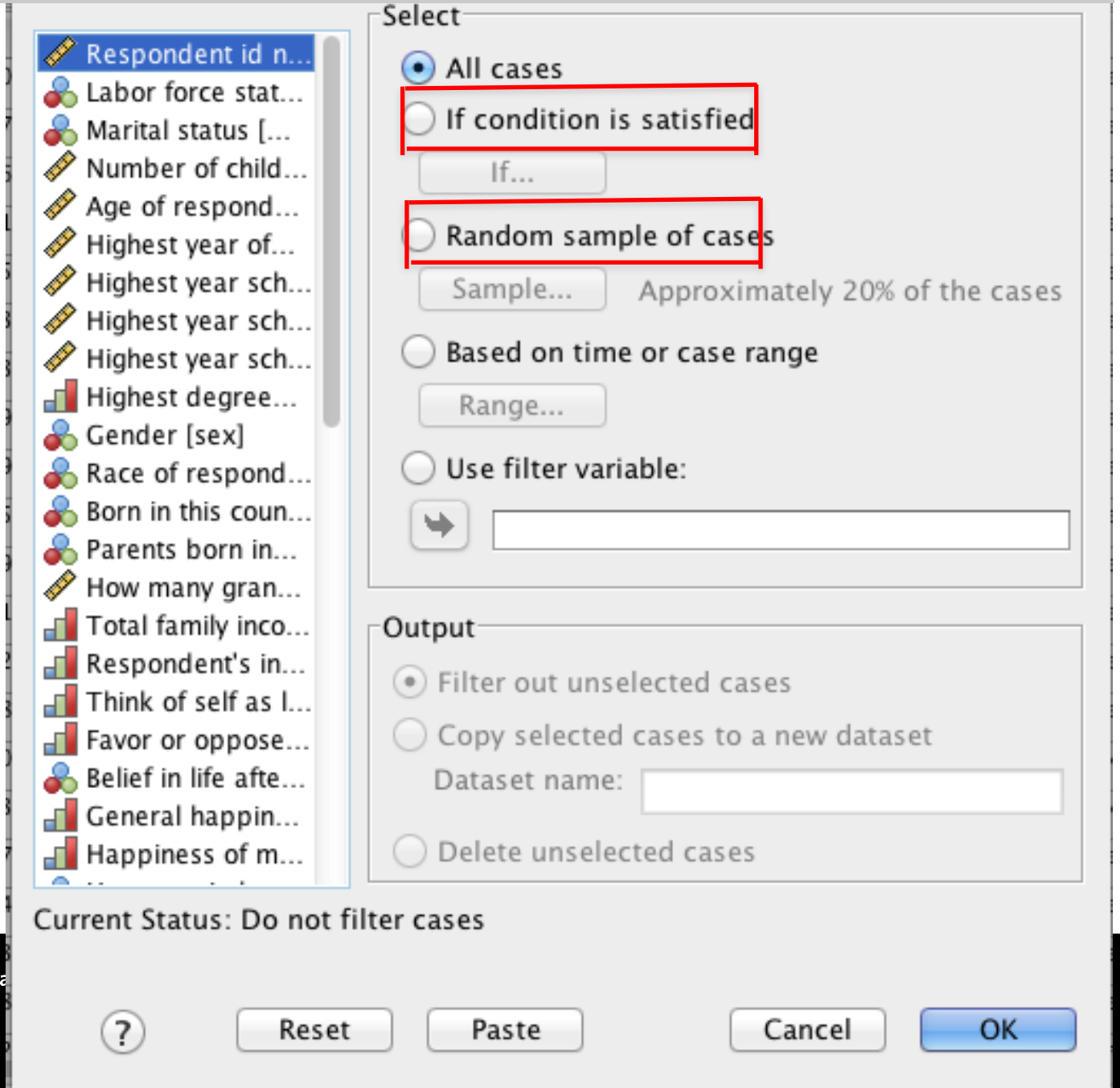
**What:** selecting a portion of the data for analysis

**When:** to do specific analysis in a data file on some of the cases.

◇The data file is too big—to randomly select a small number of cases

◇To do an analysis only on a specified group of cases

# Menu: Data→ Select Cases



## The conditions can be combined.

Condition 1 And Condition 2

Condition 1 OR Condition 2

Condition 1 AND Condition 2 AND (NOT Condition 3)

# Sort vs Select.

| | id | wrkstat | marital | childs | age | educ | paeduc | maeduc | speduc | degree | sex | race | born | parborn | granborn | income | rincome | polviews | capp |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1958 | School | Never mar... | 0 | 18 | 12 | 16 | 13 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | LT $1000 | NAP | Liberal | Op |
| 2 | 1452 | Working p... | Never mar... | 0 | 18 | 11 | 12 | 12 | NAP | LT High sc... | Female | White | Yes | Both in U.S. | 2 | $25000 or... | LT $1000 | Moderate | F |
| 3 | 2180 | Working p... | Never mar... | 0 | 18 | 11 | 16 | 12 | NAP | LT High sc... | Male | White | Yes | Both in U.S. | NAP | $25000 or... | $15000 -... | Moderate | F |
| 4 | 2448 | Working f... | Never mar... | 0 | 18 | 10 | 12 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $6000 TO... | Slightly lib... | F |
| 5 | 1041 | School | Never mar... | 0 | 18 | 12 | NAP | 18 | NAP | High school | Female | White | Yes | Both in U.S. | All in U.S. | DK | NAP | Conservat... | Op |
| 6 | 1048 | School | Divorced | 1 | 18 | 11 | 8 | 8 | NAP | LT High sc... | Male | Other | Yes | Both in U.S. | All in U.S. | DK | NAP | DK | F |
| 7 | 891 | Working p... | Never mar... | 2 | 19 | 10 | 12 | 12 | NAP | LT High sc... | Female | Black | Yes | Both in U.S. | All in U.S. | LT $1000 | NAP | Moderate | F |
| 8 | 921 | Keeping h... | Never mar... | 0 | 19 | 10 | NAP | 7 | NAP | LT High sc... | Female | Other | Yes | Neither in... | 4 | LT $1000 | NAP | Slightly co... | Op |
| 9 | 1013 | School | Never mar... | 0 | 19 | 11 | NAP | 8 | NAP | LT High sc... | Female | Other | Yes | Both in U.S. | NAP | $1000 TO... | NAP | Slightly co... | F |
| 10 | 1818 | Keeping h... | Never mar... | 0 | 19 | 12 | NAP | 12 | NAP | LT High sc... | Female | Black | Yes | Mother only | DK | $1000 TO... | NAP | Liberal | Op |
| 11 | 1943 | Working p... | Never mar... | 0 | 19 | 12 | NAP | 12 | NAP | High school | Male | Other | No | Both in U.S. | NAP | $3000 TO... | NAP | Slightly lib... | Op |

| | id | wrkstat | marital | childs | age | educ | paeduc | maeduc | speduc | degree | sex | race | born | parborn | granborn | income | rincome | polviews |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | Working f... | Divorced | 2 | 60 | 12 | 12 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Moderate |
| 2 | 2 | Working p... | Never mar... | 0 | 27 | 17 | 20 | NAP | NAP | Junior coll... | Female | White | Yes | Both in U.S. | All in U.S. | $15000 -... | $15000 -... | Liberal |
| 3 | 3 | Working f... | Married | 2 | 36 | 12 | 12 | 12 | 16 | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Conservat... |
| 4 | 4 | Working f... | Never mar... | 0 | 21 | 13 | NAP | 12 | NAP | High school | Male | White | Yes | Both in U.S. | 1 | $15000 -... | $15000 -... | Liberal |
| 5 | 5 | Working f... | Never mar... | 0 | 35 | 16 | NAP | 12 | NAP | Bachelor | Female | White | Yes | Neither in... | 4 | $25000 or... | $25000 or... | Moderate |
| 6 | 6 | Working f... | Divorced | 1 | 33 | 16 | 9 | 6 | NAP | Bachelor | Male | White | Yes | Father only | 3 | $20000 -... | NAP | Moderate |
| 7 | 7 | Working f... | Separated | 0 | 43 | 12 | 14 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | NAP | $25000 or... | $25000 or... | Moderate |
| 8 | 8 | Working f... | Never mar... | 0 | 29 | 13 | 16 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | 2 | $25000 or... | $25000 or... | Moderate |
| 9 | 9 | Working p... | Married | 2 | 39 | 18 | 16 | 12 | 13 | Bachelor | Female | White | Yes | Both in U.S. | 1 | DK | DK | Slightly co... |
| 10 | 10 | Working f... | Divorced | 0 | 45 | 15 | 16 | 12 | NAP | Junior coll... | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly co... |
| 11 | 11 | Unemploy... | Never mar... | 0 | 29 | 12 | 12 | 12 | NAP | High school | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly lib... |
| 12 | 12 | Working f... | Married | 1 | 41 | 15 | NAP | 8 | 14 | High school | Female | Black | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Liberal |
| 13 | 13 | Working p... | Divorced | 2 | 32 | 14 | NAP | 12 | NAP | High school | Female | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $5000 TO... | Moderate |
| 14 | 14 | Working f... | Married | 1 | 48 | 20 | 12 | 8 | 16 | Bachelor | Male | White | Yes | Both in U.S. | All in U.S. | $25000 or... | $25000 or... | Slightly lib... |
| 15 | 15 | Keeping h... | Never mar... | 0 | 20 | 12 | 12 | 12 | NAP | High school | Female | White | No | Both in U.S. | All in U.S. | $25000 or... | NA | Moderate |

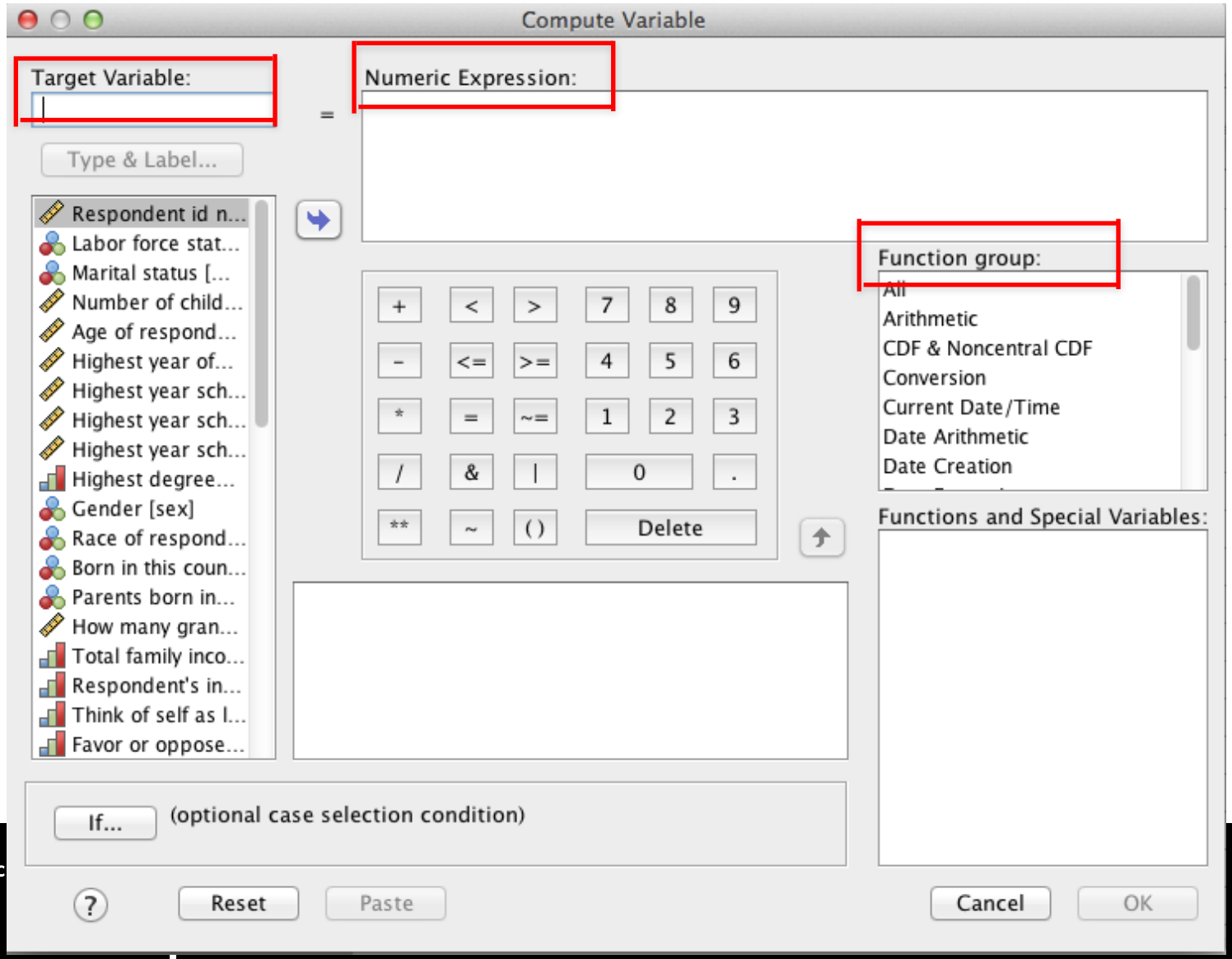
## The Compute Command

**What**: Arithmetical operations that usually result in a new variable being produced.

**When:** you need to create a new variable by making calculations on the existing variables

# The Compute Command

## Menu: Transform- Compute Variable…



# The Compute Command

**Demonstration**

Using the dataset *driving task.sav*

Scenario:

- Existing variables:

  Daytime (errors in daytime task)

  Nighttime (errors in nighttime task)

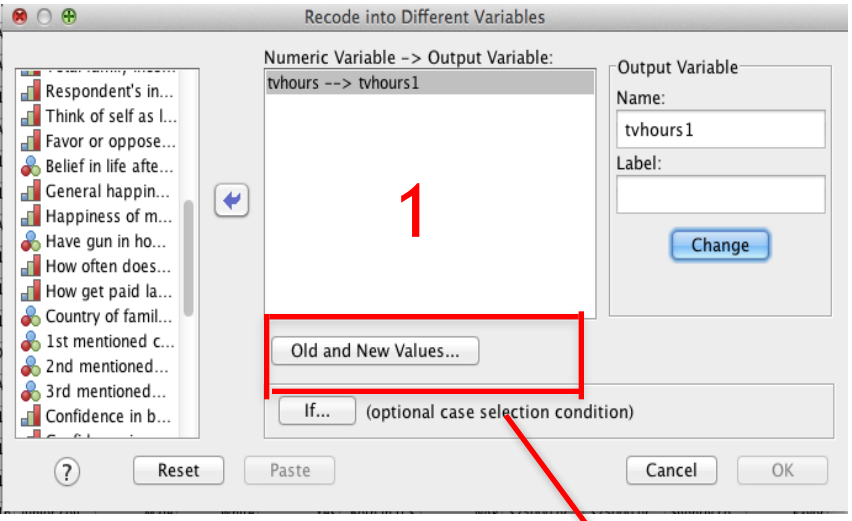
- You want to get the average errors for daytime and nighttime task

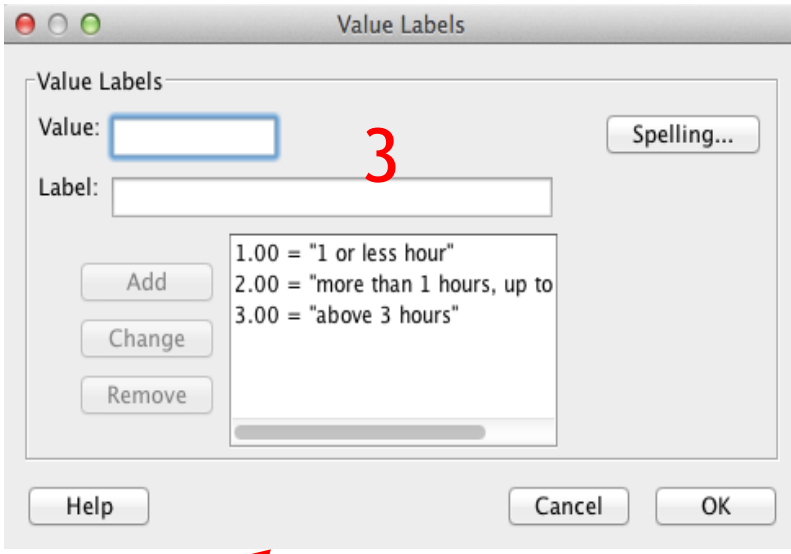
# The Recode command

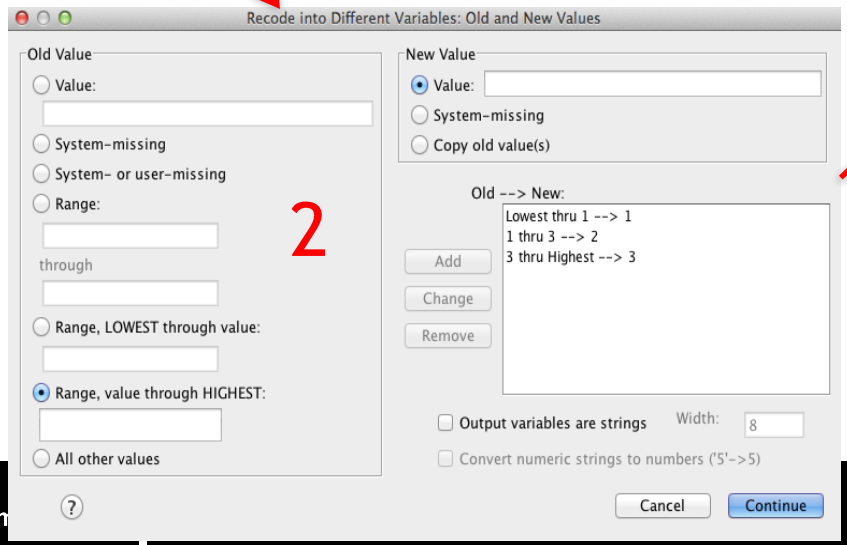
**What**: recode variables to replace a value or a range of values by a single new one

**When:** to reduce a large number of values into a small number of categories.

# Menu: Transform→ Recode into Different Variables…



## Demonstration

Example: for survey_sample.sav you may want to group the values of the hours per day watching TV in your sample into three groups.

The coding of your new categories will be as follows:

| Old value | New value | New label |
|---|---|---|
| 0,1,2 | 1 | 2 hour or less |
| 3 and above | 2 | 3 hours and more |

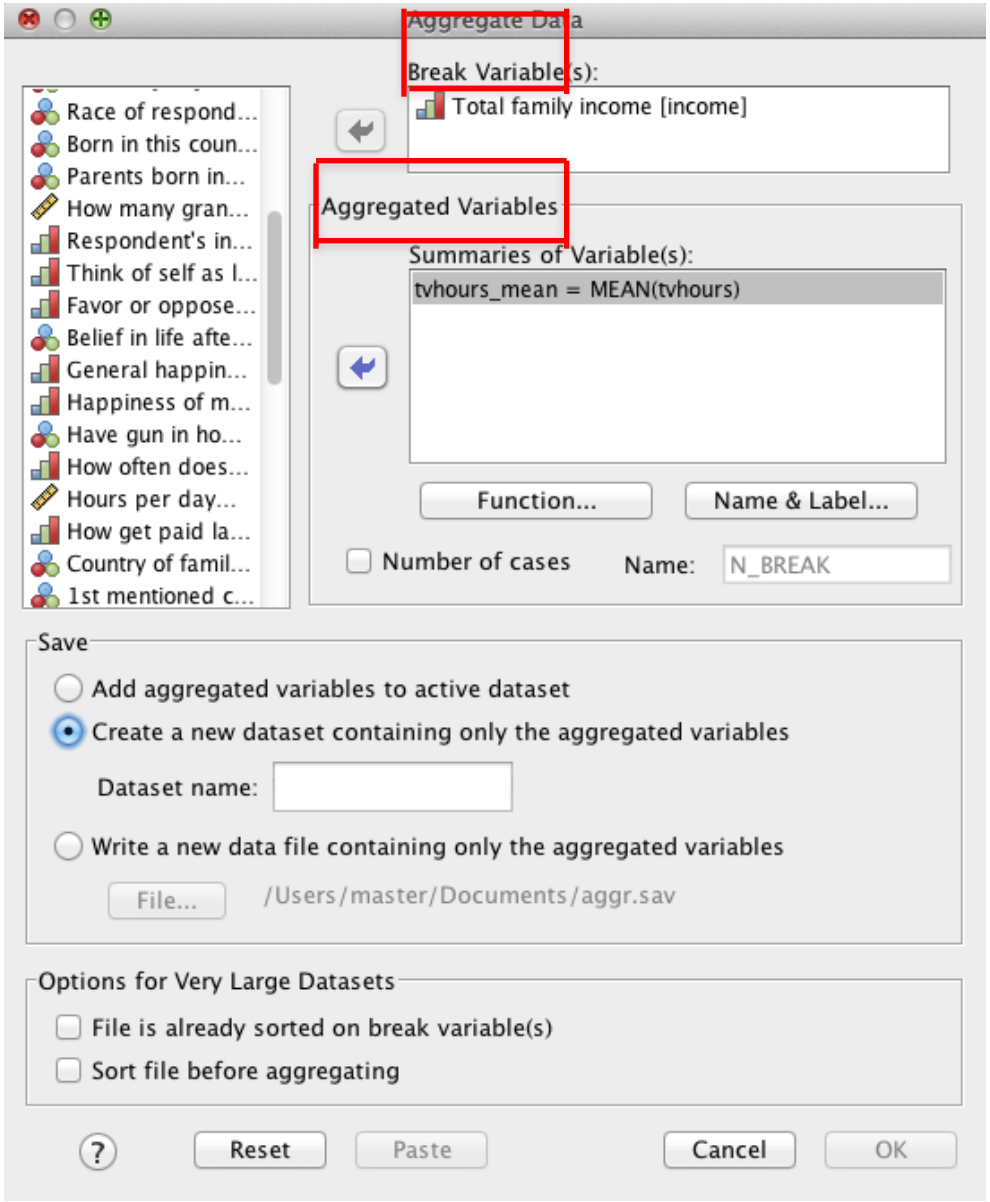
# The Aggregate Command

**What:** collapse a variable by its categories and aggregate these categories by calculating a statistics like mean.

**When:** for compiling statistics for subgroups of your sample, and writing the results into a form that allows further analysis and manipulation.

**Example:** if you want to know the average hours of watching TV of the families with different levels of family income.

## Menu: Data→ Aggregate …



| | income | tvhours_mean |
|---|---|---|
| 1 | LT $1000 | 4.71 |
| 2 | $1000 TO... | 3.07 |
| 3 | $3000 TO... | 2.75 |
| 4 | $4000 TO... | 4.10 |
| 5 | $5000 TO... | 3.59 |
| 6 | $6000 TO... | 5.21 |
| 7 | $7000 TO... | 3.76 |
| 8 | $8000 TO... | 3.48 |
| 9 | $10000 -... | 3.61 |
| 10 | $15000 -... | 3.01 |
| 11 | $20000 -... | 3.01 |
| 12 | $25000 or... | 2.47 |
| 13 | REFUSED | 2.95 |
| 14 | DK | 3.54 |
| 15 | NA | 4.19 |

| | | | Hours per daywatching TV into three groups | | | |
|---|---|---|---|---|---|---|
| | | | 1 or less hour | more than 1 hours, up to 3 hours | above 3 hours | Total |
| Respondent's income | LT $1000 | Count | 11 | 15 | 16 | 42 |
| | | % within Respondent's income | 26.2% | 35.7% | 38.1% | 100.0% |
| | $1000 TO 2999 | Count | 17 | 14 | 11 | 42 |
| | | % within Respondent's income | 40.5% | 33.3% | 26.2% | 100.0% |
| | $3000 TO 3999 | Count | 14 | 12 | 14 | 40 |
| | | % within Respondent's income | 35.0% | 30.0% | 35.0% | 100.0% |
| | $4000 TO 4999 | Count | 11 | 9 | 7 | 27 |
| | | % within Respondent's income | 40.7% | 33.3% | 25.9% | 100.0% |
| | $5000 TO 5999 | Count | 15 | 9 | 8 | 32 |
| | | % within Respondent's income | 46.9% | 28.1% | 25.0% | 100.0% |
| | $6000 TO 6999 | Count | 10 | 18 | 15 | 43 |
| | | % within Respondent's income | 23.3% | 41.9% | 34.9% | 100.0% |
| | $7000 TO 7999 | Count | 13 | 14 | 14 | 41 |
| | | % within Respondent's income | 31.7% | 34.1% | 34.1% | 100.0% |
| | $8000 TO 9999 | Count | 22 | 23 | 18 | 63 |

**Respondent's income * Hours per day watching TV Crosstabulation**

| | | | Hours per day watching TV | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Respondent's income | LT $1000 | Count | 1 | 3 | 7 | 8 | 5 | 6 | 1 | 1 | 1 | 1 | 0 |
| | | % within Respondent's income | 2.9% | 8.8% | 20.6% | 23.5% | 14.7% | 17.6% | 2.9% | 2.9% | 2.9% | 2.9% | 0.0% |
| | $1000 TO 2999 | Count | 1 | 5 | 7 | 7 | 10 | 1 | 0 | 0 | 0 | 0 | 0 |
| | | % within Respondent's income | 3.2% | 16.1% | 22.6% | 22.6% | 32.3% | 3.2% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| | $3000 TO 3999 | Count | 1 | 4 | 8 | 4 | 6 | 5 | 2 | 0 | 0 | 0 | 1 |
| | | % within Respondent's income | 3.2% | 12.9% | 25.8% | 12.9% | 19.4% | 16.1% | 6.5% | 0.0% | 0.0% | 0.0% | 3.2% |
| | $4000 TO 4999 | Count | 2 | 4 | 5 | 4 | 2 | 1 | 3 | 1 | 0 | 0 | 0 |
| | | % within Respondent's income | 9.1% | 18.2% | 22.7% | 18.2% | 9.1% | 4.5% | 13.6% | 4.5% | 0.0% | 0.0% | 0.0% |
| | $5000 TO 5999 | Count | 1 | 6 | 4 | 5 | 2 | 3 | 1 | 0 | 0 | 0 | 0 |
| | | % within Respondent's income | 4.5% | 27.3% | 18.2% | 22.7% | 9.1% | 13.6% | 4.5% | 0.0% | 0.0% | 0.0% | 0.0% |
| | $6000 TO 6999 | Count | 0 | 5 | 9 | 9 | 7 | 3 | 3 | 1 | 0 | 0 | 1 |
| | | % within Respondent's income | 0.0% | 13.2% | 23.7% | 23.7% | 18.4% | 7.9% | 7.9% | 2.6% | 0.0% | 0.0% | 2.6% |
| | $7000 TO 7999 | Count | 1 | 5 | 8 | 6 | 5 | 5 | 1 | 1 | 1 | 0 | 0 |
| | | % within Respondent's income | 3.0% | 15.2% | 24.2% | 18.2% | 15.2% | 15.2% | 3.0% | 3.0% | 3.0% | 0.0% | 0.0% |
| | $8000 TO 9999 | Count | 1 | 13 | 12 | 11 | 6 | 5 | 4 | 0 | 1 | 0 | 1 |
| | | % within Respondent's | | | | | | | | | | | |

# Exercise 1

**Program Files-IBM-SPSS-Statistics-23-Samples-English-Survey_sample.sav**

Please create a new variable by recoding the variable ***Think of self as a liberal or conservative (polviews)***.

The old variable has 7 categories (1=Extremely liberal, 2=liberal, 3=slightly liberal, 4=moderate, 5=slightly conservative, 6=conservative, 7=extremely conservative) plus three missing values (0,8,9). You want to recode the 7 categories into three: 1=liberal, 2=moderate, 3=conservative.

# Workshop II

**Bivariate analysis**

- differences between scores for two variables
- relationships between two variables

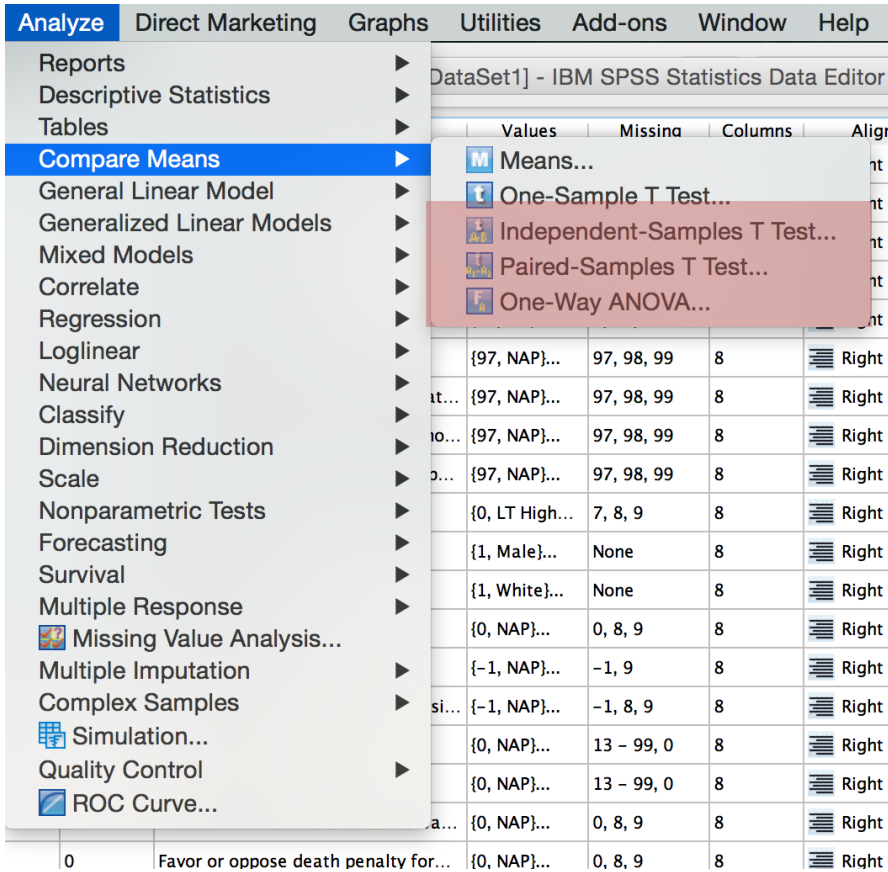
# Bivariate analysis:

## Scale variables

↓

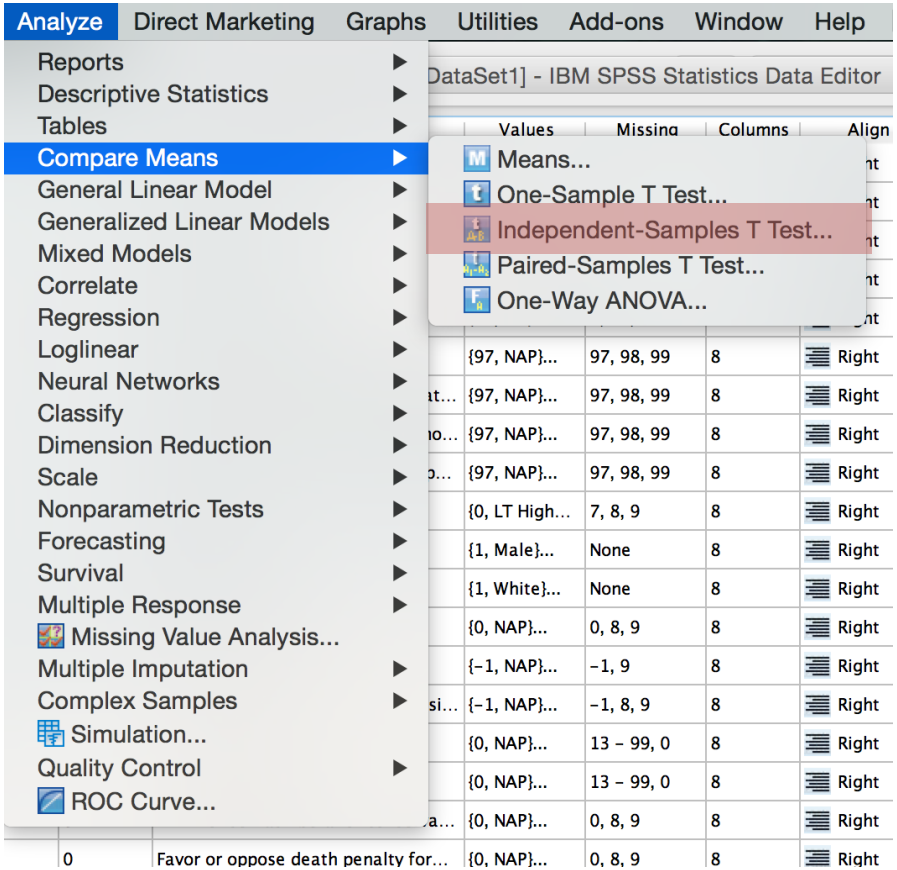
Compare means

- Independent T Test
- Paired-sample T-Test
- One Way ANOVA

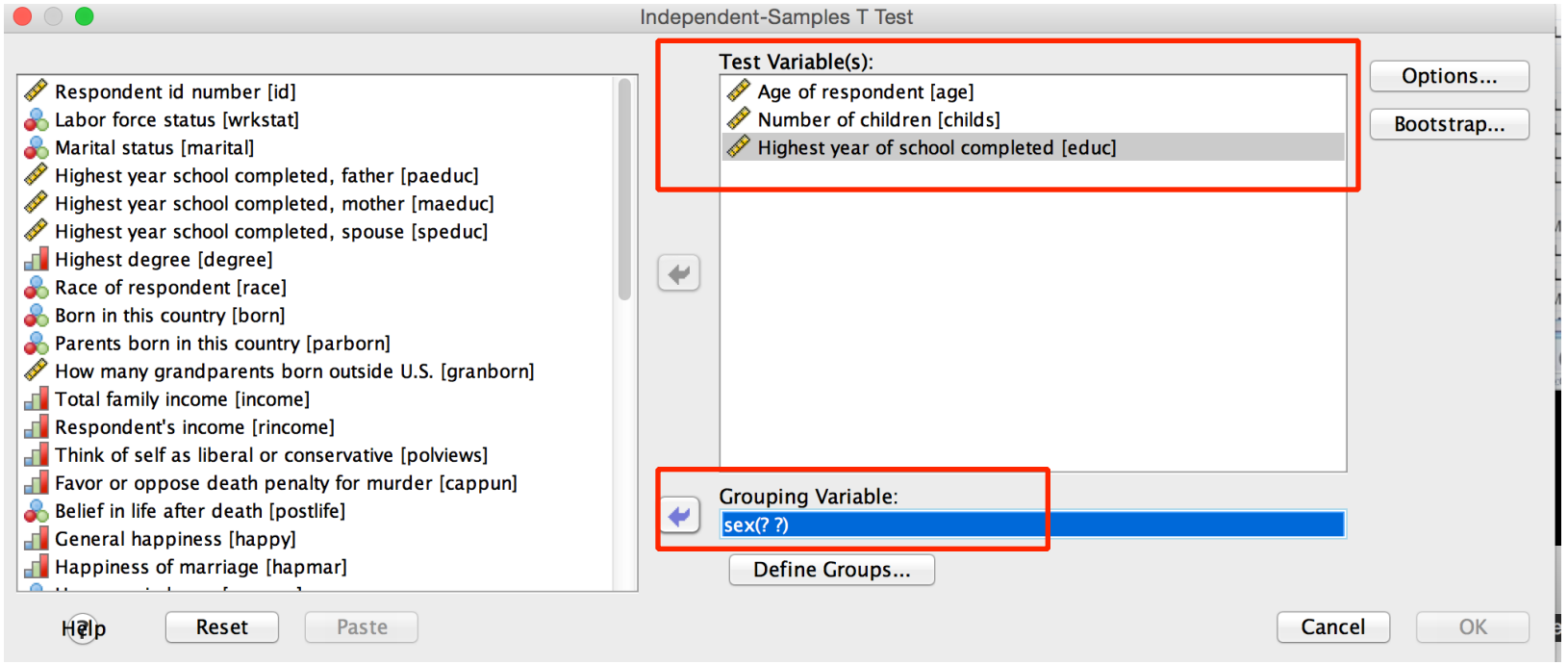


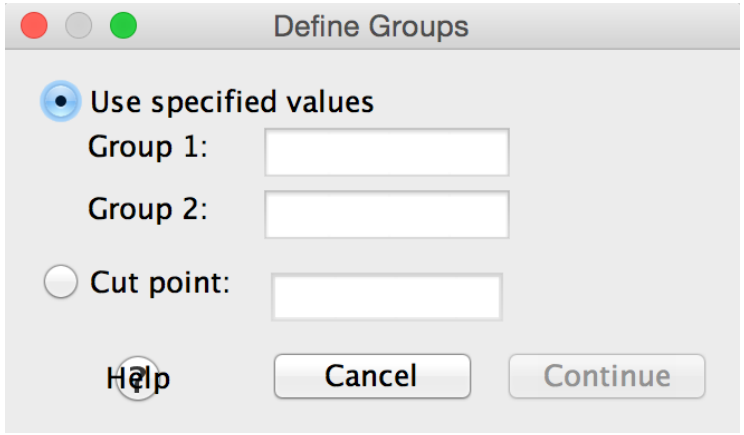
# Independent T Test

TASK: compare men and women's age, # of child, highest years of school completed




Is used when you want to compare the means of a normally distributed interval dependent variable for two independent groups (eg. males, females)


Independent-Samples T Test
Respondent id number [id]
Labor force status [wrkstat]
Marital status [marital]
Highest year school completed, father [paeduc]
Highest year school completed, mother [maeduc]
Highest year school completed, spouse [speduc]
Highest degree [degree]
Race of respondent [race]
Born in this country [born]
Parents born in this country [parborn]
How many grandparents born outside U.S. [granborn]
Total family income [income]
Respondent's income [rincome]
Think of self as liberal or conservative [polviews]
Favor or oppose death penalty for murder [cappun]
Belief in life after death [postlife]
General happiness [happy]
Happiness of marriage [hapmar]
Test Variable(s):
Age of respondent [age]
Number of children [childs]
Highest year of school completed [educ]
Options...
Bootstrap...
Grouping Variable:
sex(? ?)
Define Groups...
Help
Reset
Paste
Cancel
OK


Define Groups
Use specified values
Group 1:
Group 2:
Cut point:
Help
Cancel
Continue

```
T-TEST GROUPS=sex(1 2)
  /MISSING=ANALYSIS
  /VARIABLES=age childs educ
  /CRITERIA=CI(.95).
```

## T-Test

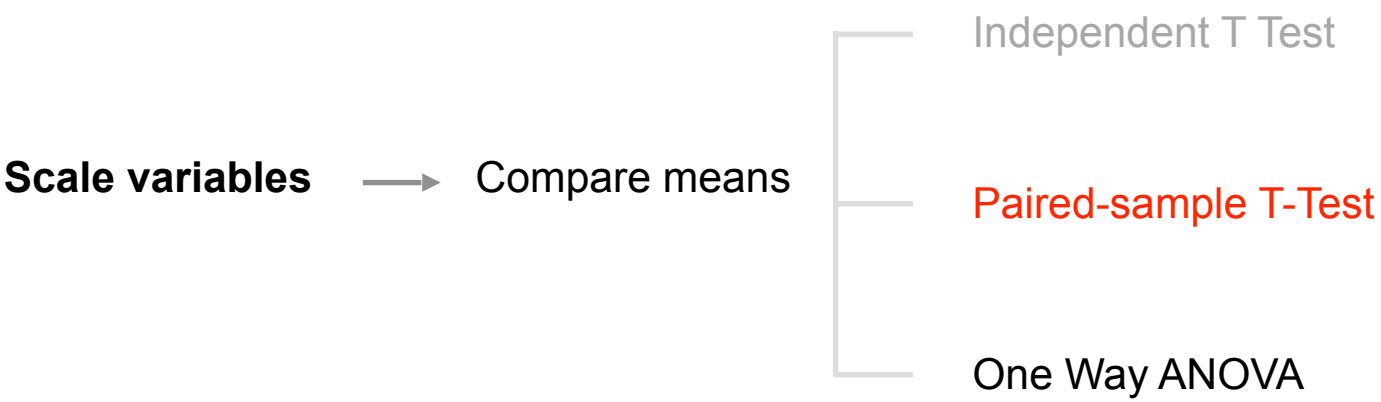
**Group Statistics**

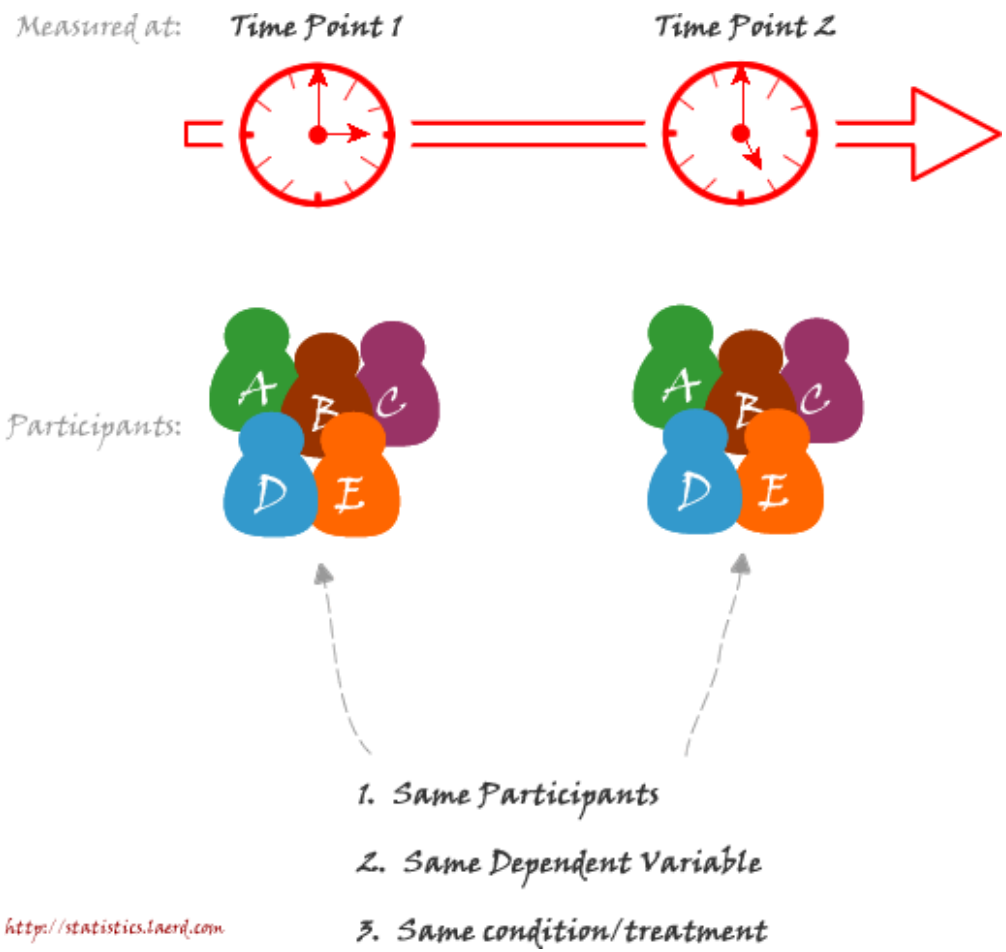
| | Gender | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Age of respondent | Male | 1228 | 44.57 | 16.333 | .466 |
| | Female | 1600 | 46.32 | 17.634 | .441 |
| Number of children | Male | 1229 | 1.65 | 1.667 | .048 |
| | Female | 1596 | 1.96 | 1.695 | .042 |
| Highest year of school completed | Male | 1225 | 13.36 | 2.966 | .085 |
| | Female | 1595 | 13.17 | 2.896 | .073 |

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| Age of respondent | Equal variances assumed | 13.917 | .000 | -2.701 | 2826 | .007 | -1.750 | .648 | -3.021 | -.479 |
| | Equal variances not assumed | | | -2.728 | 2728.627 | .006 | -1.750 | .642 | -3.008 | -.492 |
| Number of children | Equal variances assumed | 1.444 | .230 | -4.887 | 2823 | .000 | -.312 | .064 | -.437 | -.187 |
| | Equal variances not assumed | | | -4.898 | 2662.350 | .000 | -.312 | .064 | -.437 | -.187 |
| Highest year of school completed | Equal variances assumed | 2.064 | .151 | 1.738 | 2818 | .082 | .193 | .111 | -.025 | .411 |
| | Equal variances not assumed | | | 1.733 | 2601.434 | .083 | .193 | .112 | -.025 | .412 |

# Bivariate analysis:

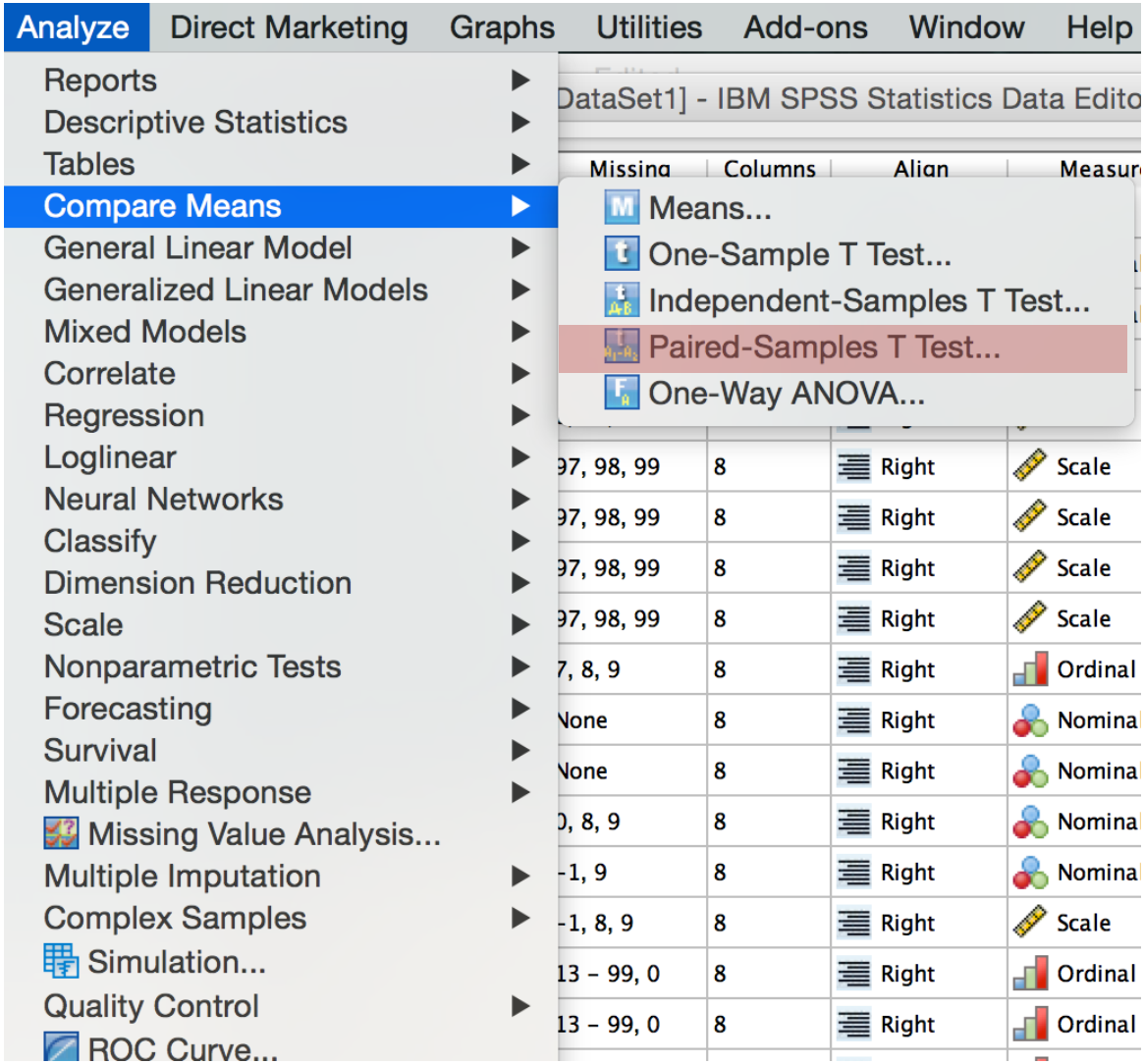





## Paired-sample T-Test

Same group of people.

Test the before and after difference




## Paired-sample T-Test

Analyze

↓

Compare Means

↓

Paired-Samples T Test

A driving simulator was used to investigate the number of errors made during daytime and night-time driving tasks. Ten people were given the same task of driving in daytime and in night-time. The number of driving errors were as follows:

| Participants: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Daytime task: | 7 | 9 | 6 | 5 | 8 | 6 | 9 | 6 | 5 | 7 |
| Night-time: | 8 | 10 | 8 | 7 | 8 | 10 | 11 | 7 | 8 | 7 |

**Let's do this**

Use the given data set and try Paired sample T test

# Bivariate analysis:

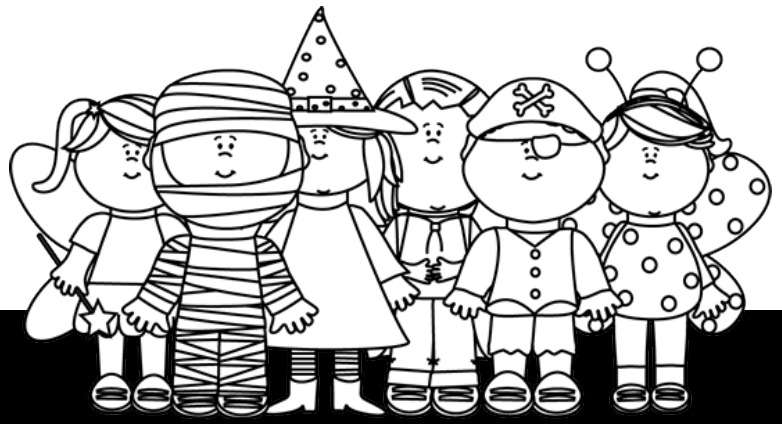
**Scale variables** → Compare means

- Independent T Test
- Paired-sample T-Test
- One Way ANOVA

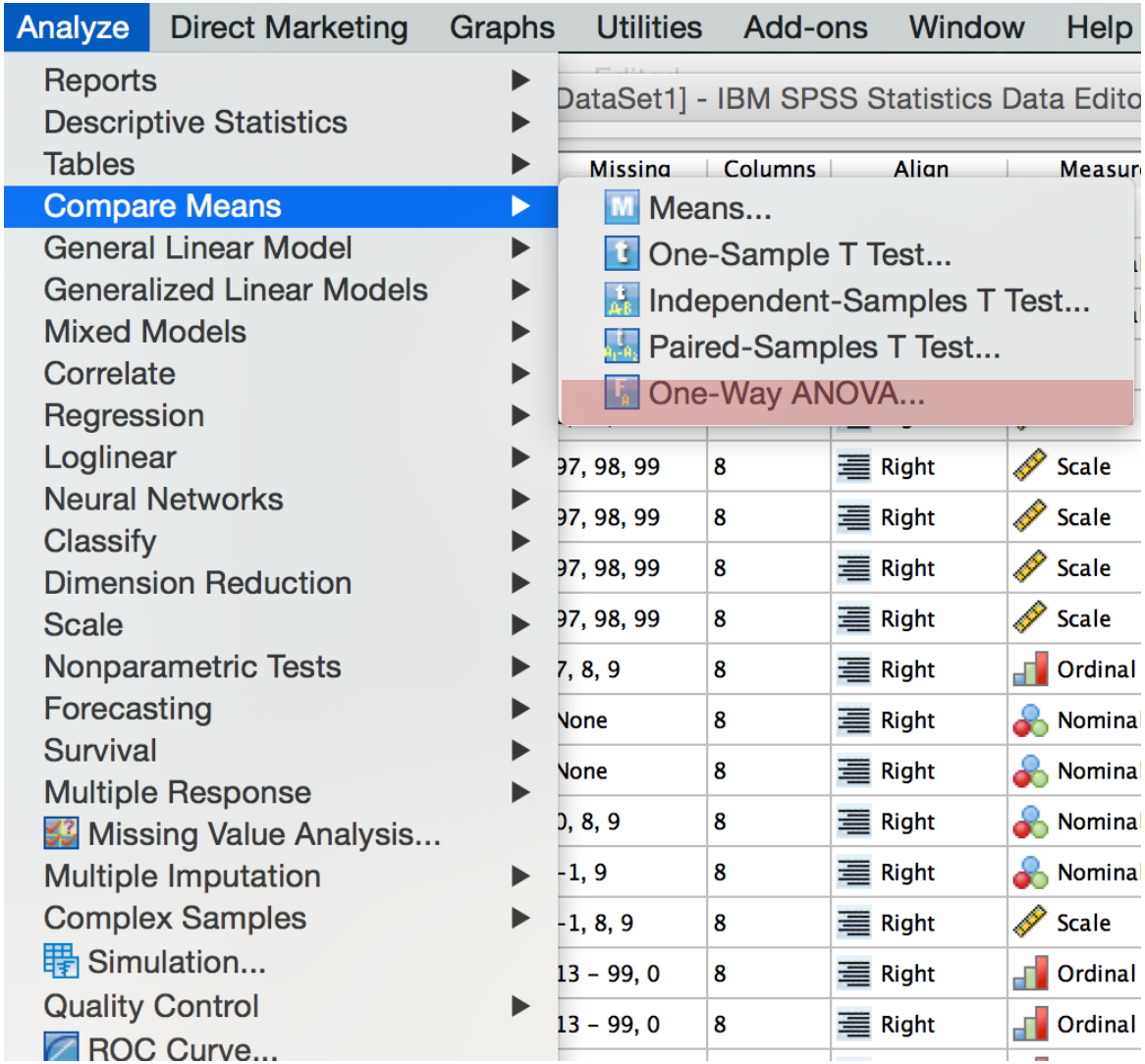
# One-way Anova

- a categorical independent variable (nominal/ordinal) (>2 categories)
- a normally distributed interval dependent variable (scale)
- test for differences in the means of the dependent variable broken down by the levels of the independent variable.

TASK: test the differences of **number of children** broken down by **happiness of marriage**



# One-way Anova



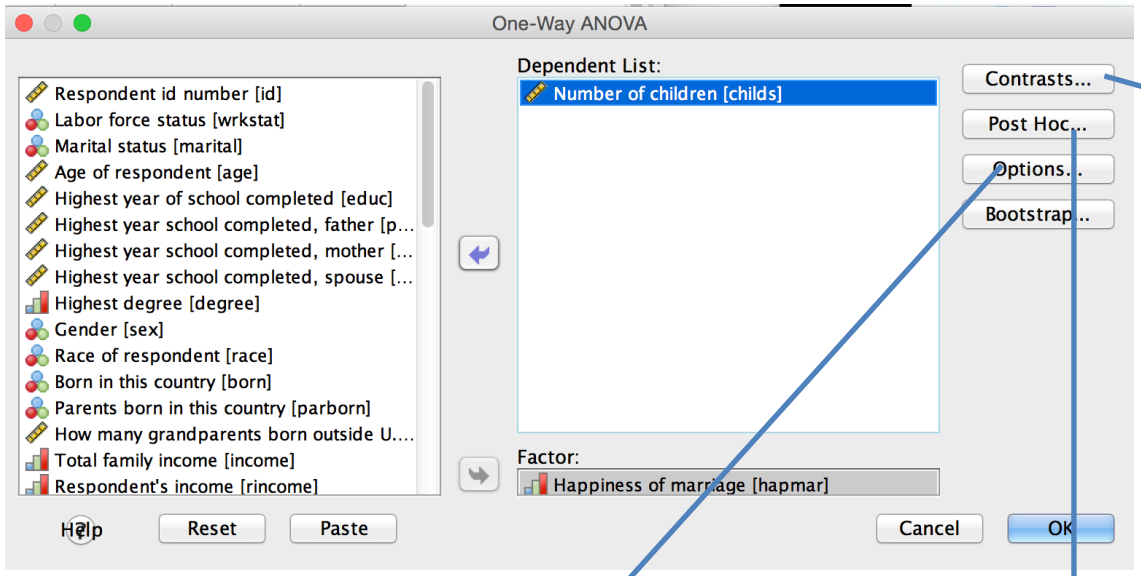

Analyze

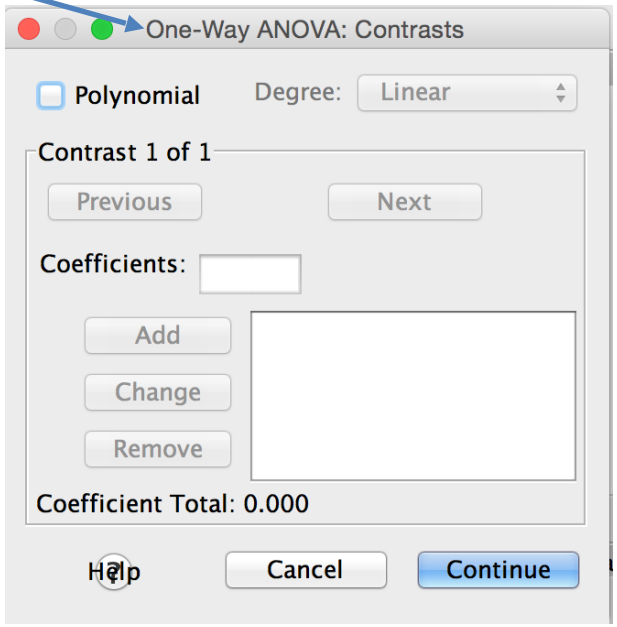
↓

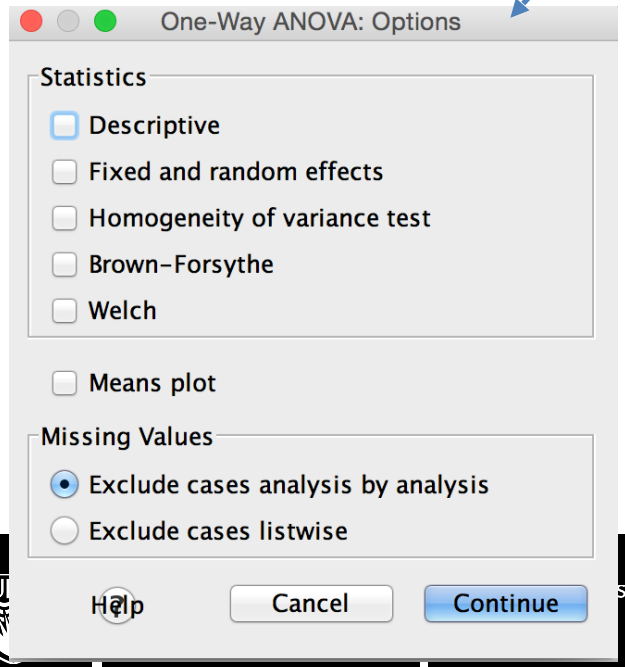
Compare Means

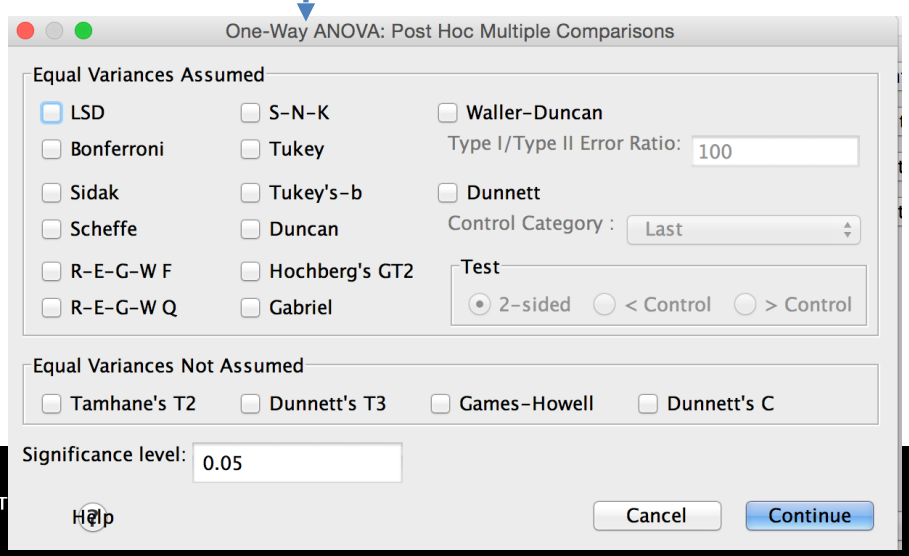
↓

One-Way ANOVA


One-Way ANOVA
Respondent id number [id]
Labor force status [wrkstat]
Marital status [marital]
Age of respondent [age]
Highest year of school completed [educ]
Highest year school completed, father [p...
Highest year school completed, mother [...
Highest year school completed, spouse [...
Highest degree [degree]
Gender [sex]
Race of respondent [race]
Born in this country [born]
Parents born in this country [parborn]
How many grandparents born outside U....
Total family income [income]
Respondent's income [rincome]
Dependent List:
Number of children [childs]
Factor:
Happiness of marriage [hapmar]
Contrasts...
Post Hoc...
Options...
Bootstrap...
Help
Reset
Paste
Cancel
OK


One-Way ANOVA: Contrasts
Polynomial
Degree:
Linear
Contrast 1 of 1
Previous
Next
Coefficients:
Add
Change
Remove
Coefficient Total: 0.000
Help
Cancel
Continue


One-Way ANOVA: Options
Statistics
Descriptive
Fixed and random effects
Homogeneity of variance test
Brown-Forsythe
Welch
Means plot
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Help
Cancel
Continue


One-Way ANOVA: Post Hoc Multiple Comparisons
Equal Variances Assumed
LSD
Bonferroni
Sidak
Scheffe
R-E-G-W F
R-E-G-W Q
S-N-K
Tukey
Tukey's-b
Duncan
Hochberg's GT2
Gabriel
Waller-Duncan
Type I/Type II Error Ratio:
100
Dunnett
Control Category :
Last
Test
2-sided
< Control
> Control
Equal Variances Not Assumed
Tamhane's T2
Dunnett's T3
Games-Howell
Dunnett's C
Significance level:
0.05
Help
Cancel
Continue

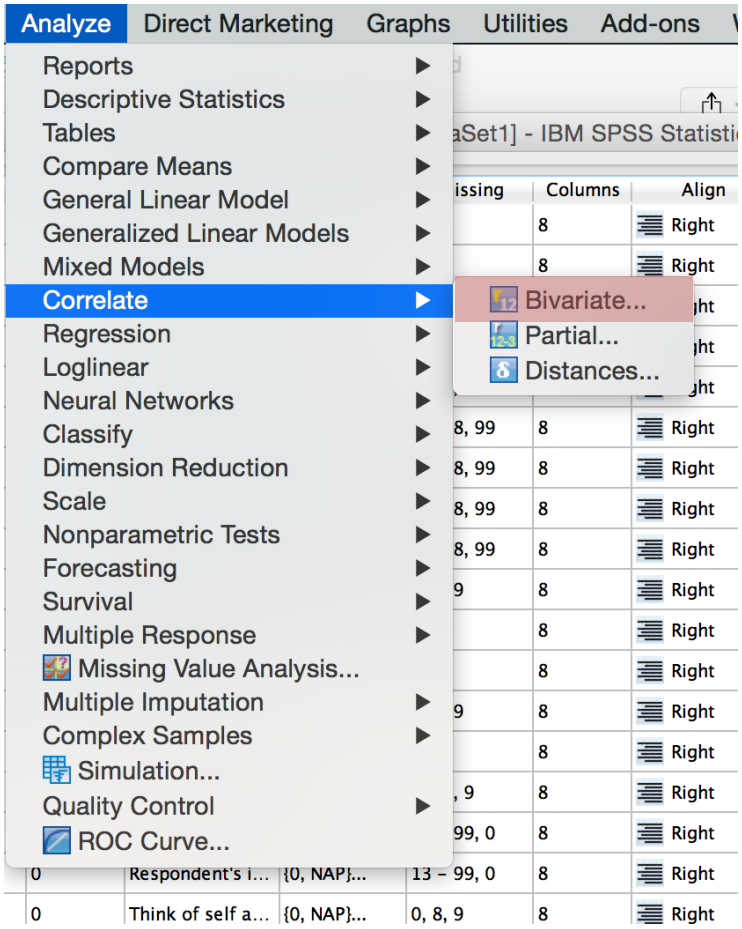
## Bivariate analysis

differences between scores on two variables

relationships between two variables

- Correlation
- Regression
- Crosstabulations

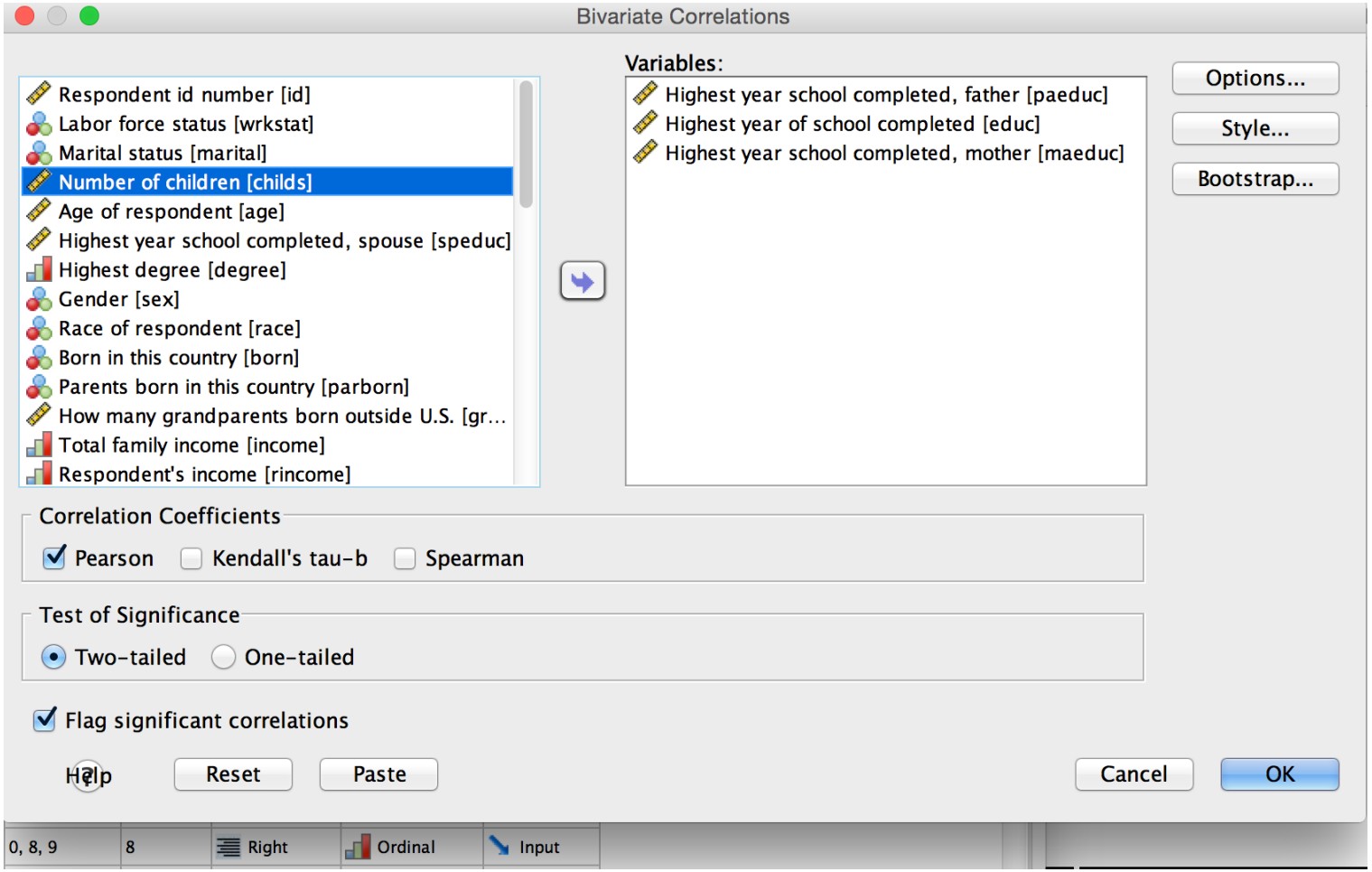
# Correlation



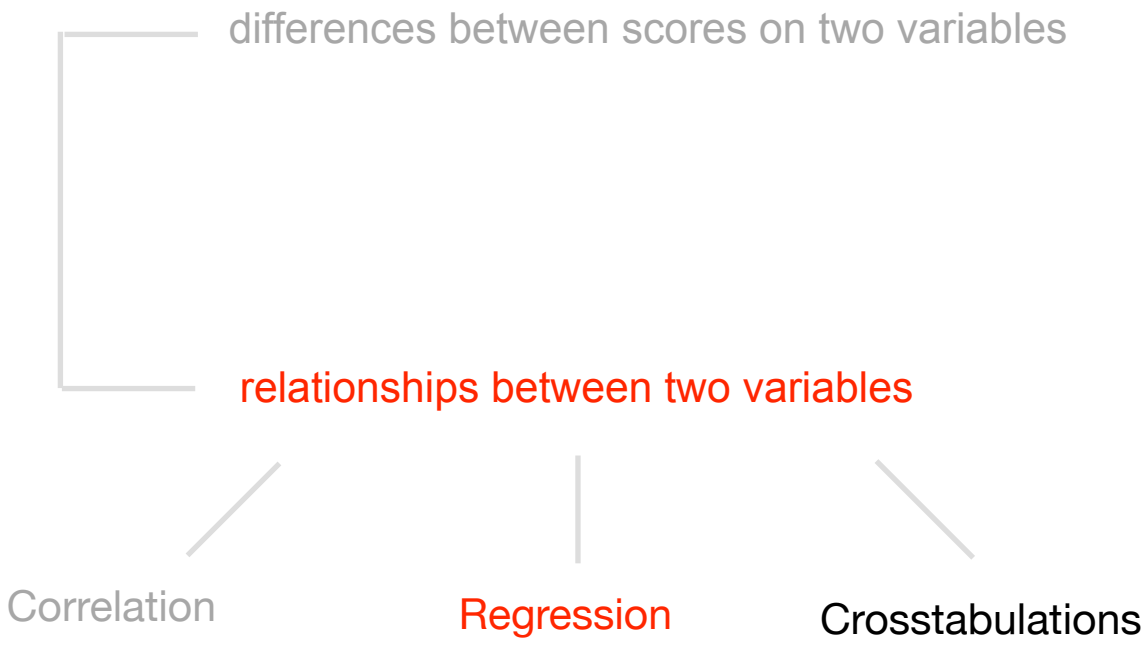

TASK: correlation between

**highest year school completed** and

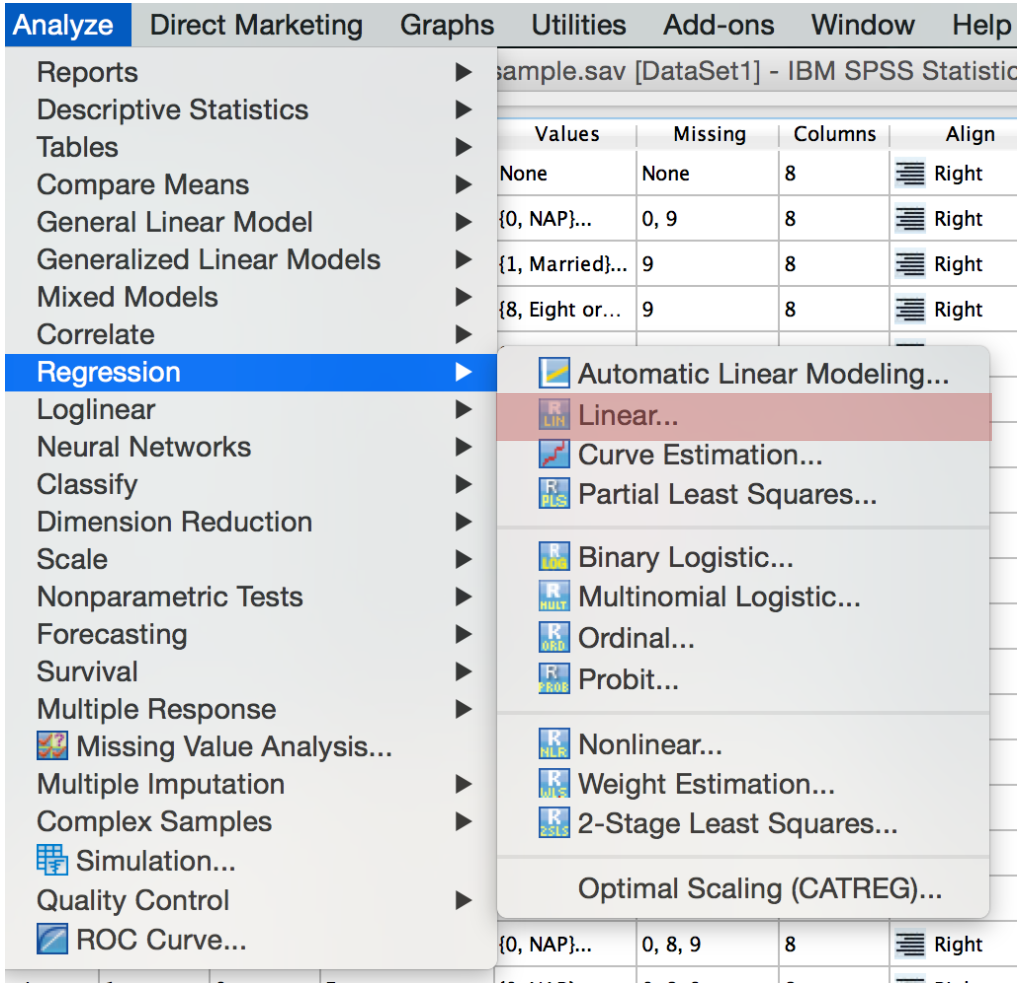
**highest year school completed for father**


Bivariate Correlations
Respondent id number [id]
Labor force status [wrkstat]
Marital status [marital]
Number of children [childs]
Age of respondent [age]
Highest year school completed, spouse [speduc]
Highest degree [degree]
Gender [sex]
Race of respondent [race]
Born in this country [born]
Parents born in this country [parborn]
How many grandparents born outside U.S. [gr...
Total family income [income]
Respondent's income [rincome]
Variables:
Highest year school completed, father [paeduc]
Highest year of school completed [educ]
Highest year school completed, mother [maeduc]
Options...
Style...
Bootstrap...
Correlation Coefficients
Pearson
Kendall's tau-b
Spearman
Test of Significance
Two-tailed
One-tailed
Flag significant correlations
Help
Reset
Paste
Cancel
OK
0, 8, 9
8
Right
Ordinal
Input

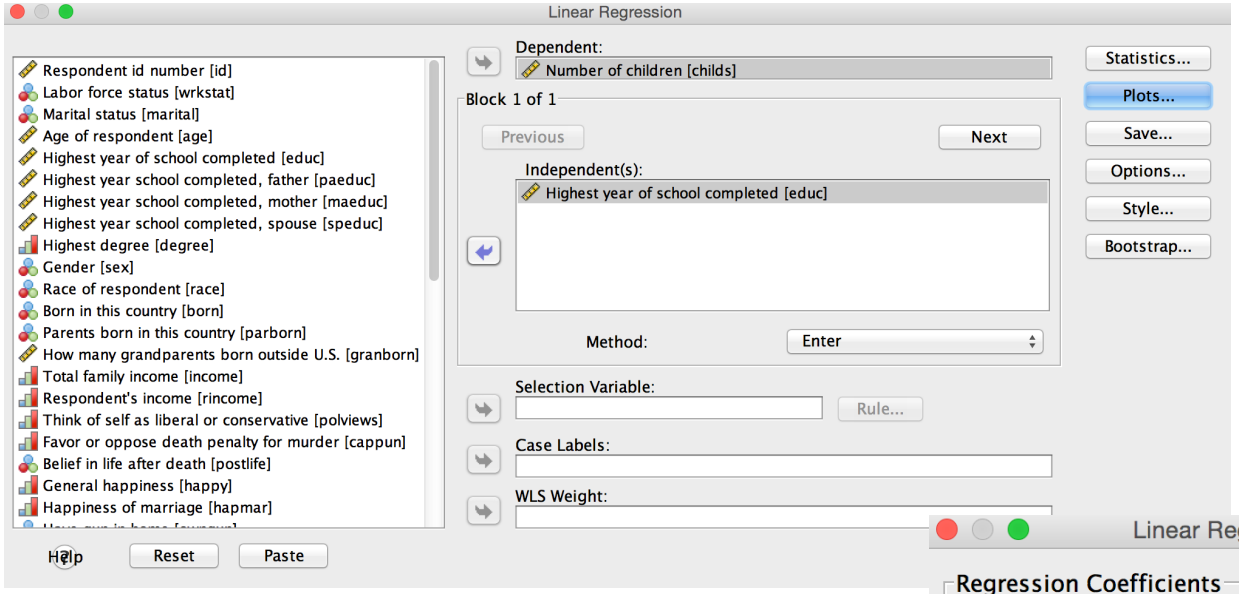
## Bivariate analysis

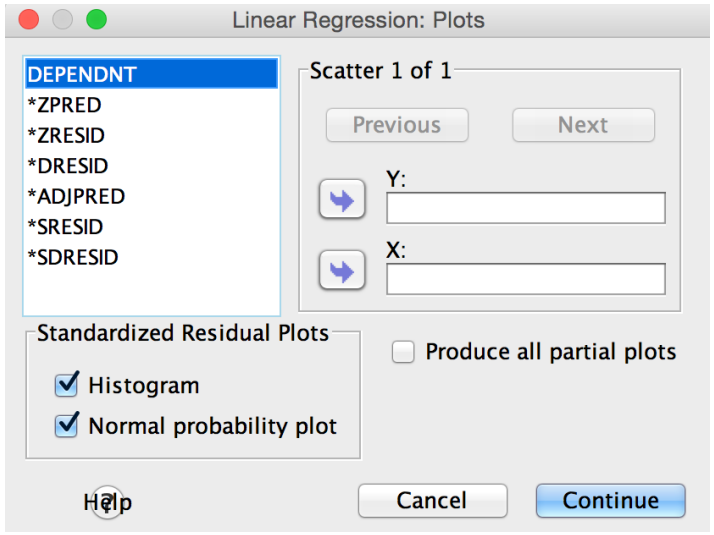


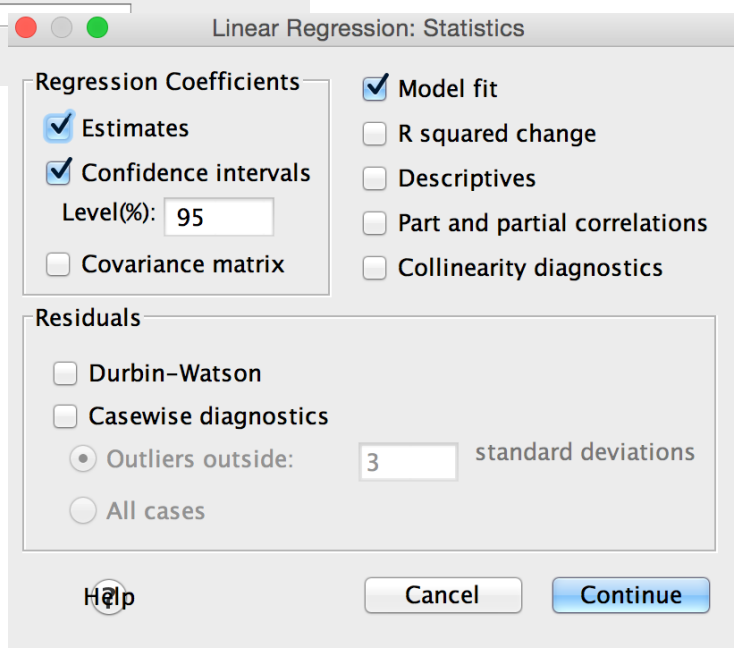
# Regression




TASK: Association between "number of children" and " highest year of school completed"


Linear Regression
Respondent id number [id]
Labor force status [wrkstat]
Marital status [marital]
Age of respondent [age]
Highest year of school completed [educ]
Highest year school completed, father [paeduc]
Highest year school completed, mother [maeduc]
Highest year school completed, spouse [speduc]
Highest degree [degree]
Gender [sex]
Race of respondent [race]
Born in this country [born]
Parents born in this country [parborn]
How many grandparents born outside U.S. [granborn]
Total family income [income]
Respondent's income [rincome]
Think of self as liberal or conservative [polviews]
Favor or oppose death penalty for murder [cappun]
Belief in life after death [postlife]
General happiness [happy]
Happiness of marriage [hapmar]
Dependent:
Number of children [childs]
Block 1 of 1
Previous
Next
Independent(s):
Highest year of school completed [educ]
Method:
Enter
Selection Variable:
Rule...
Case Labels:
WLS Weight:
Statistics...
Plots...
Save...
Options...
Style...
Bootstrap...
Help
Reset
Paste


Linear Regression: Plots
DEPENDNT
*ZPRED
*ZRESID
*DRESID
*ADJPRED
*SRESID
*SDRESID
Scatter 1 of 1
Previous
Next
Y:
X:
Standardized Residual Plots
Histogram
Normal probability plot
Produce all partial plots
Help
Cancel
Continue


Linear Regression: Statistics
Regression Coefficients
Estimates
Confidence intervals
Level(%): 95
Covariance matrix
Model fit
R squared change
Descriptives
Part and partial correlations
Collinearity diagnostics
Residuals
Durbin-Watson
Casewise diagnostics
Outliers outside: 3 standard deviations
All cases
Help
Cancel
Continue

TASK: Gender and Think of self as liberal or conservative

http://www.ats.ucla.edu/stat/spss/faq/xtab_int.htm

Step 4: How interpret the results

SPSS->Help

Step 3: Find the test in SPSS

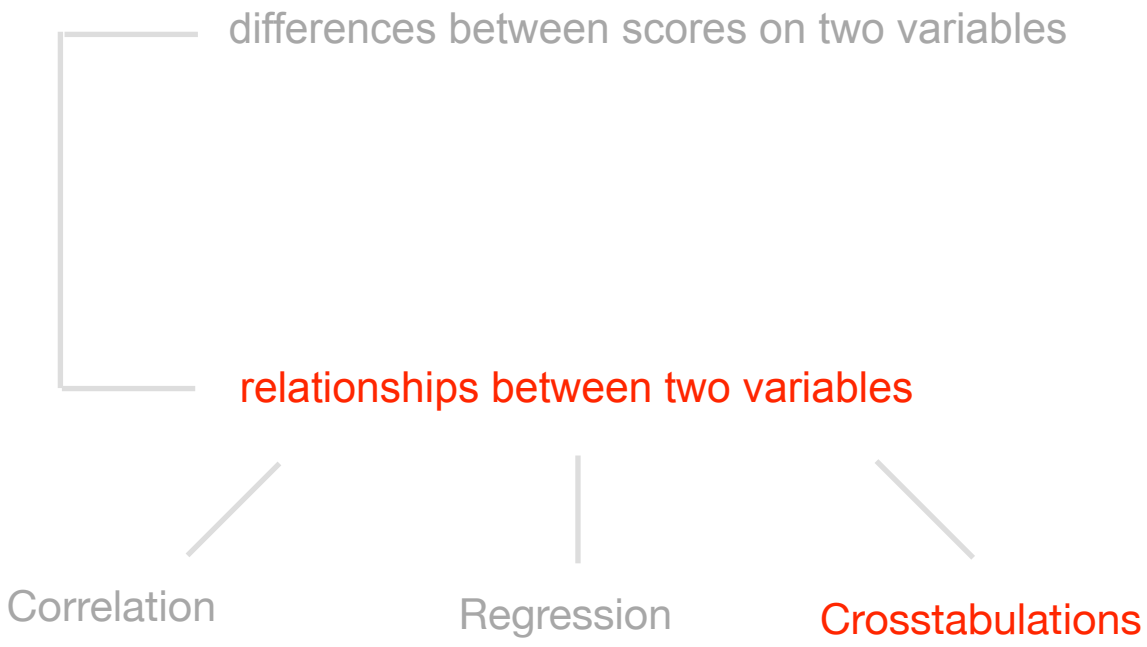
SPSS->Help->Statistics Coach -> Related info.

Step 2: Assumptions, options

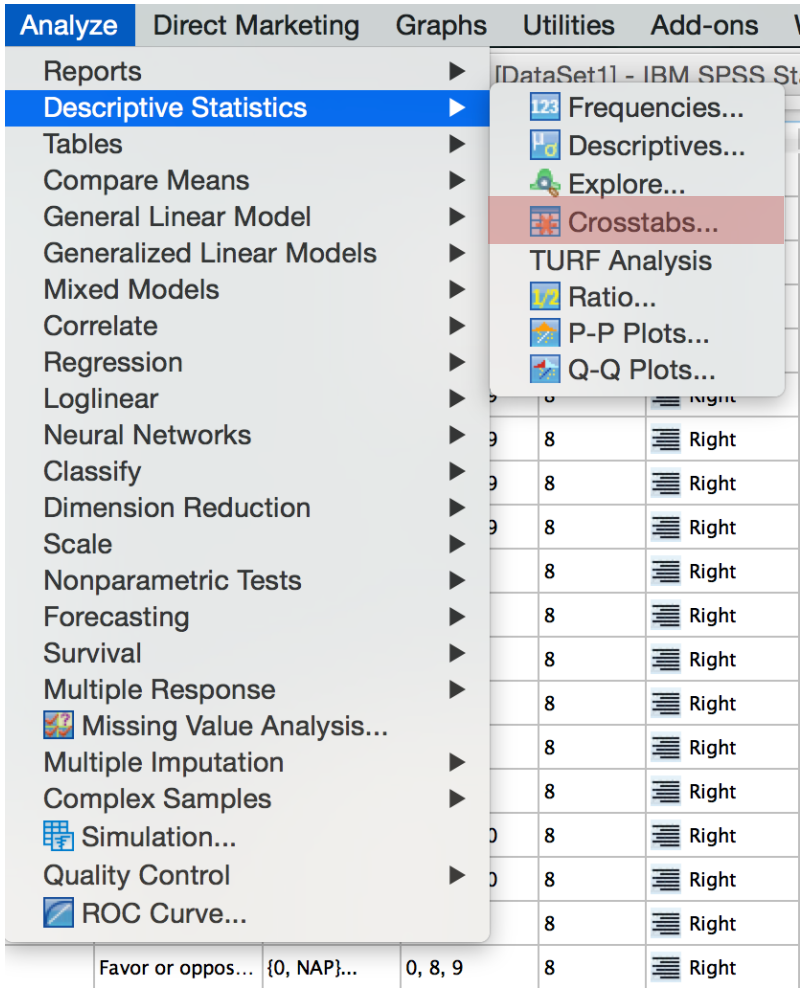
SPSS->Help->Statistics Coach.

Step1: Choose an appropriate test

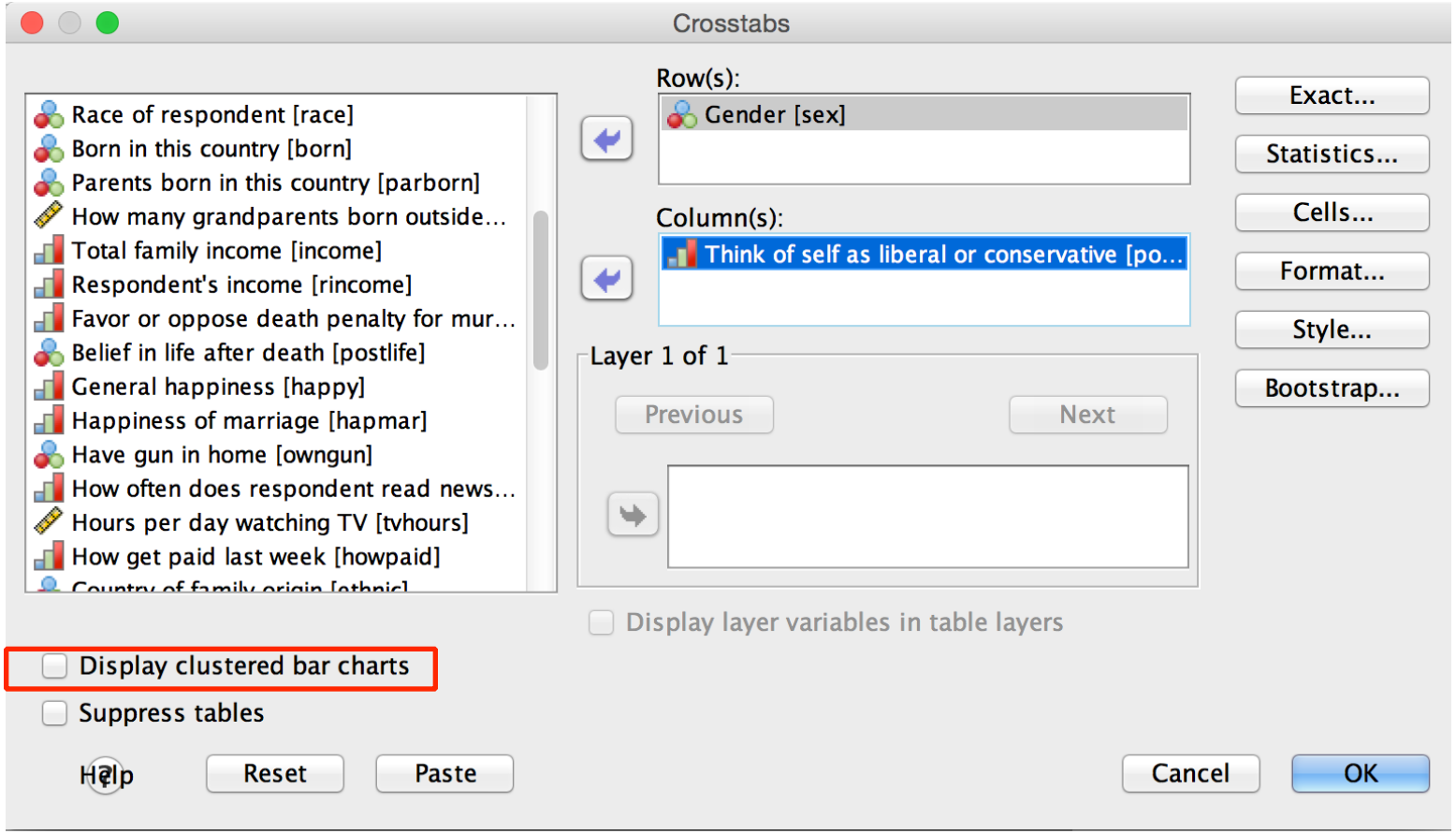
## Bivariate analysis



# Crosstabulations




# TASK: Gender and Think of self as liberal or conservative



TASK: Use “Statistical Resources Tool” to run test for the following scenario

**Say we wish to test whether the mean for age is the same for males and females.**

You want to compare the means of a normally distributed interval dependent variable for two independent groups.

Statistics Coach > Statistics Coach > Compare groups for significant differences > Scale numeric data divided into groups > Two groups or variables > One scale, numeric dependent variable divided into two unrelated groups


Google ubc research commons feedback form

All Images Maps News Videos More Search tools

About 12,600 results (0.56 seconds)

**Research Commons | Koerner Library**
koerner.library.**ubc**.ca/services/**research-commons**/
Apr 3, 2014 - ... of our team? We'd love to get your feedback - just fill out the **Research Commons** Student **Feedback Form**. ... Book a consultation via our Consultation Request Form. ... Koerner Library, Level 2, **research.commons@ubc**.ca ...
You've visited this page 4 times. Last visit: 02/02/16

**Thesis Formatting**
Need help with formatting your thesis? The Research ...

**Graduate Student Workshops**
Graduate Student Workshops. Research in a large research ...

**SPSS Support**
The Research Commons now offers SPSS Software Support.

**Feedback Form**
Research Commons Feedback Form. Last updated on June 6 ...

**Citation Management**
You are able to view upcoming weekly citation management ...

**FIREtalks**
Join us for FIREtalks, an interdisciplinary discussion ...

More results from ubc.ca »

http://koerner.library.ubc.ca/services/research-commons/research-commons-feedback-form/


Thank you!