

# Introduction to SPSS and data management

Αθανάσιος Κυργίδης  
Παναγιώτης Μαυρίδης  
Κωνσταντίνος Τζιόμαλος  
Κωνσταντίνος Τουλής

# Lecture Outline

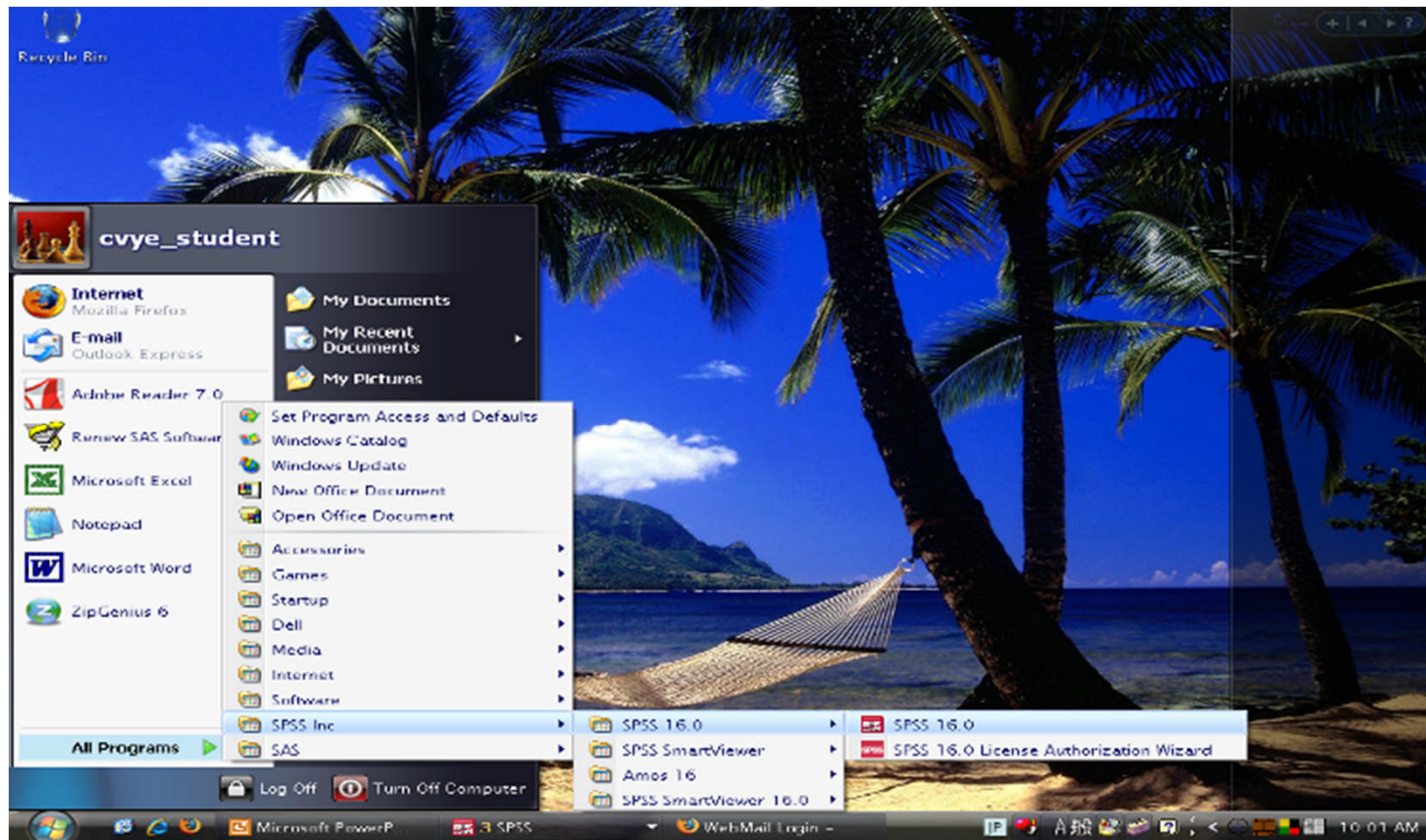


- [Introduction](#)
- [Basics of managing data files](#)
- [Normality Explorations](#)
- Basic Analysis
  - [Descriptive Statistics](#)
  - [Analytical Statistics](#)
- [Exporting Output](#)
- [Hands On](#) (Λυσάρι Ασκήσεων)

# The basics of managing data files

# Opening SPSS

- Start → All Programs → SPSS Inc → SPSS 19.0 → SPSS 19.0



# PASW Statistics

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

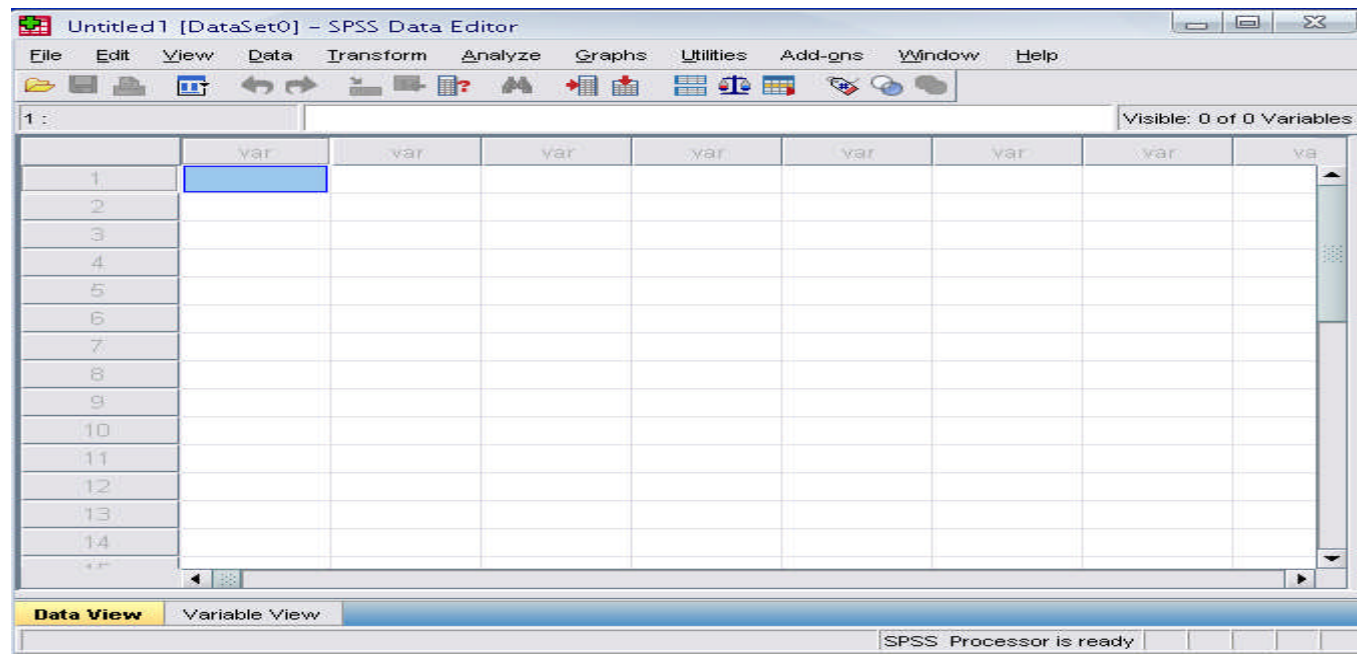
☐ Don't show this dialog in the future

OK

Cancel

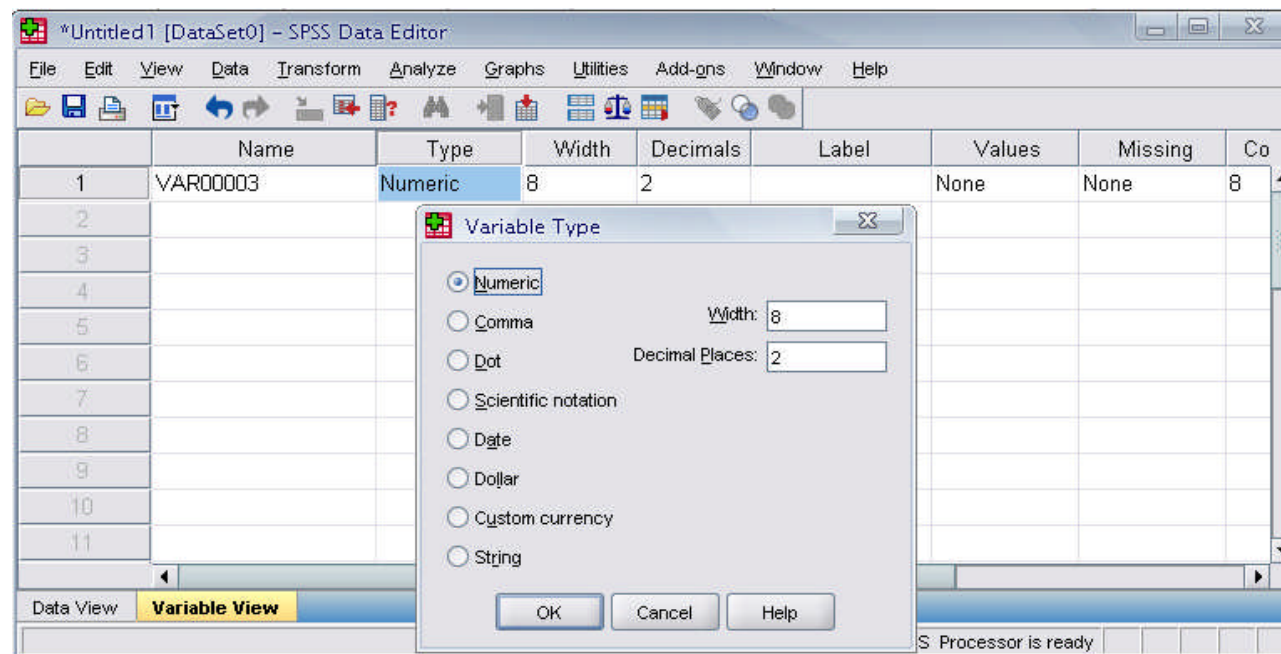
# Opening SPSS

- The default window will have the data editor
- There are two sheets in the window:
  1. Data view
  2. Variable view



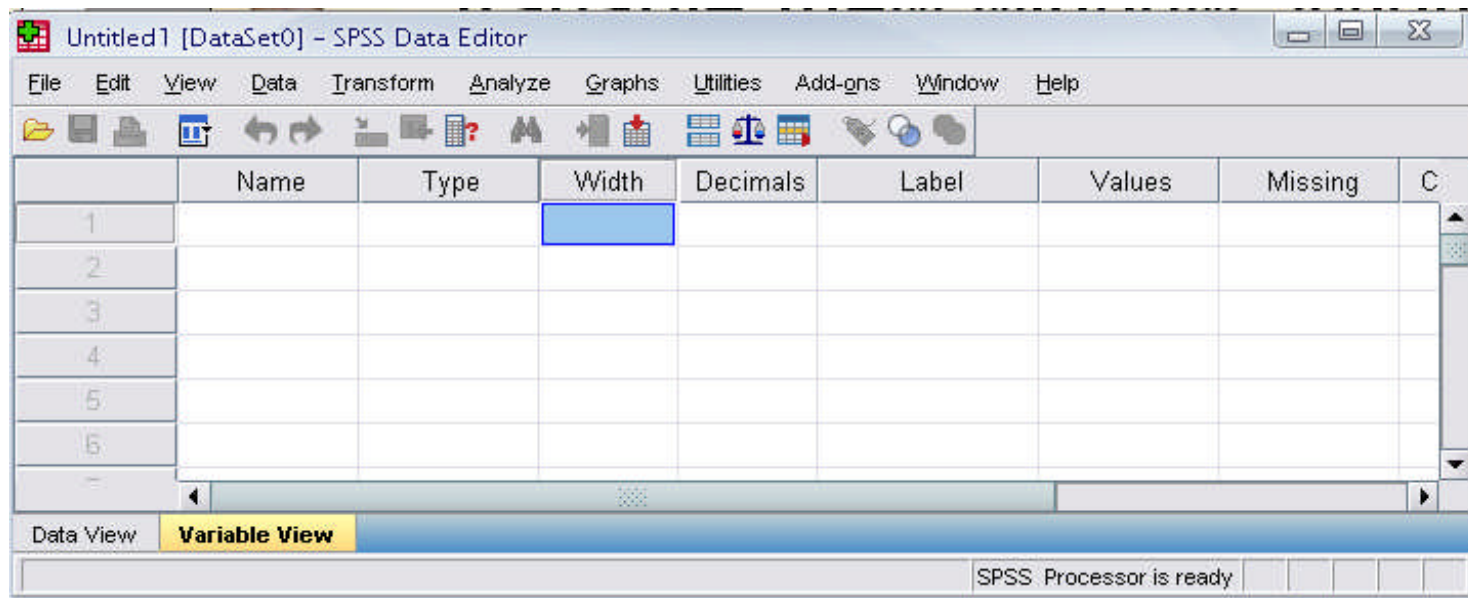
# Variable View window: Type

- Type
  - Click on the ‘type’ box. The two basic types of variables that you will use are numeric and string. This column enables you to specify the type of variable.



# Variable View window: Width

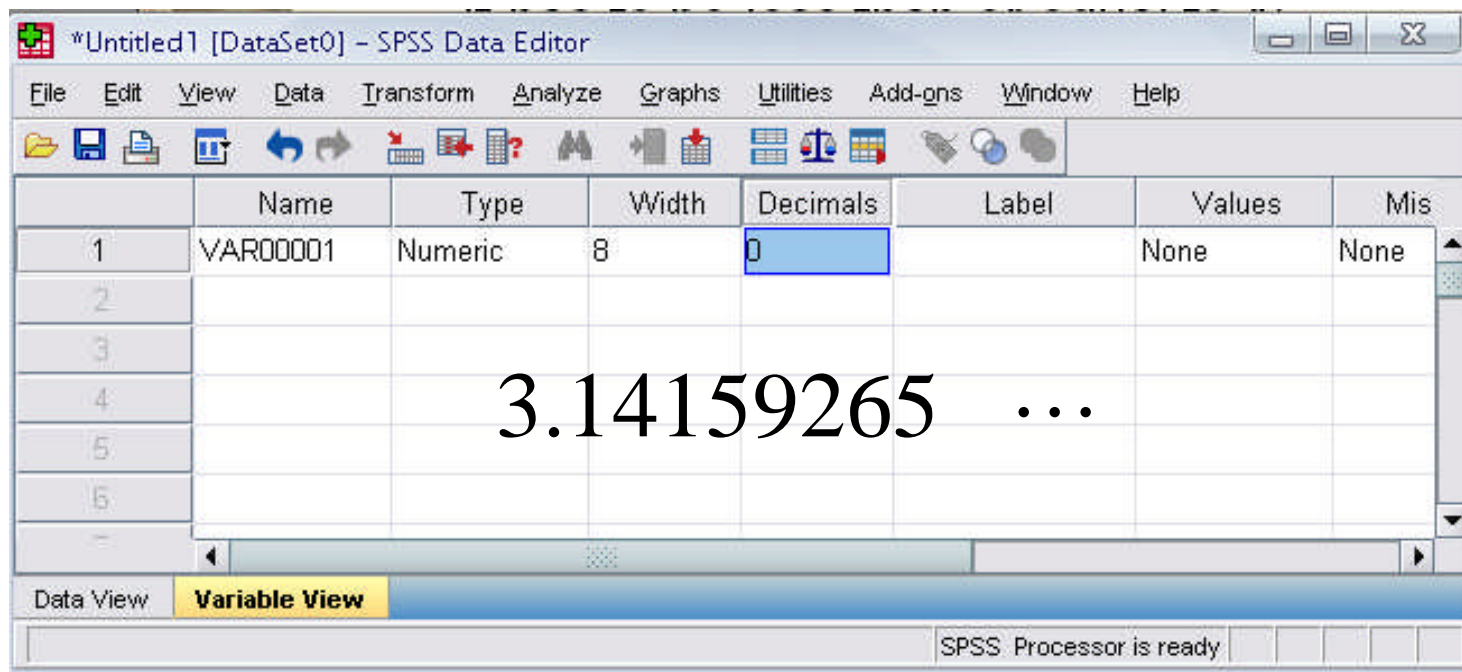
- Width
  - Width allows you to determine the number of characters SPSS will allow to be entered for the variable





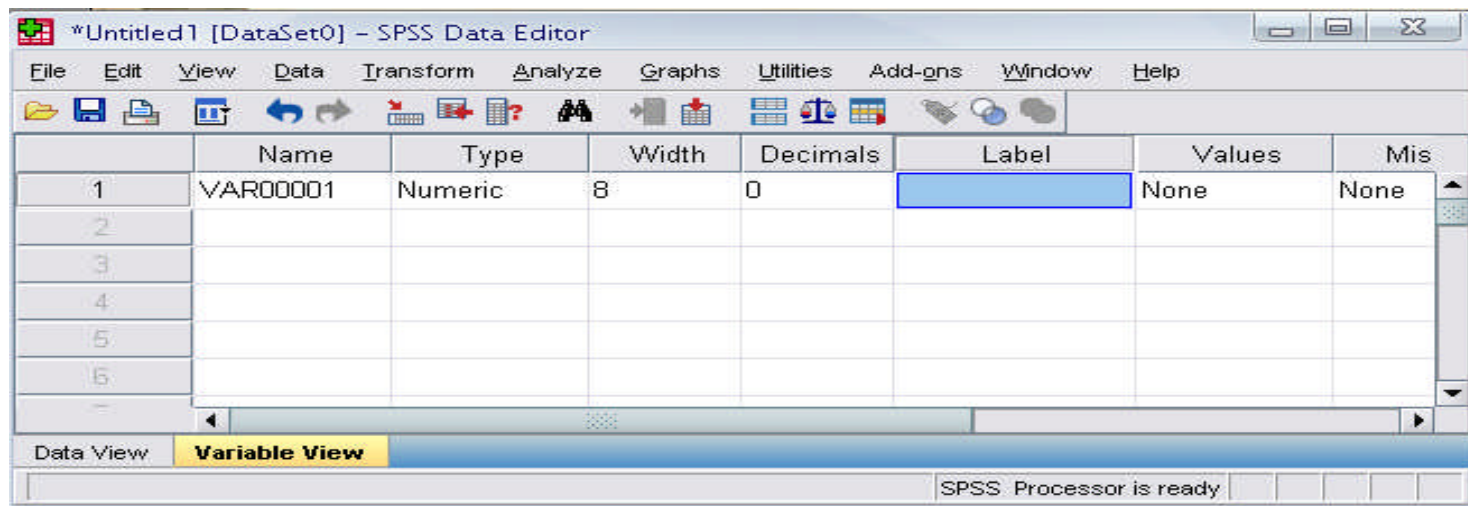
# Variable View window: Decimals

- Decimals
  - Number of decimals
  - It has to be less than or equal to 16



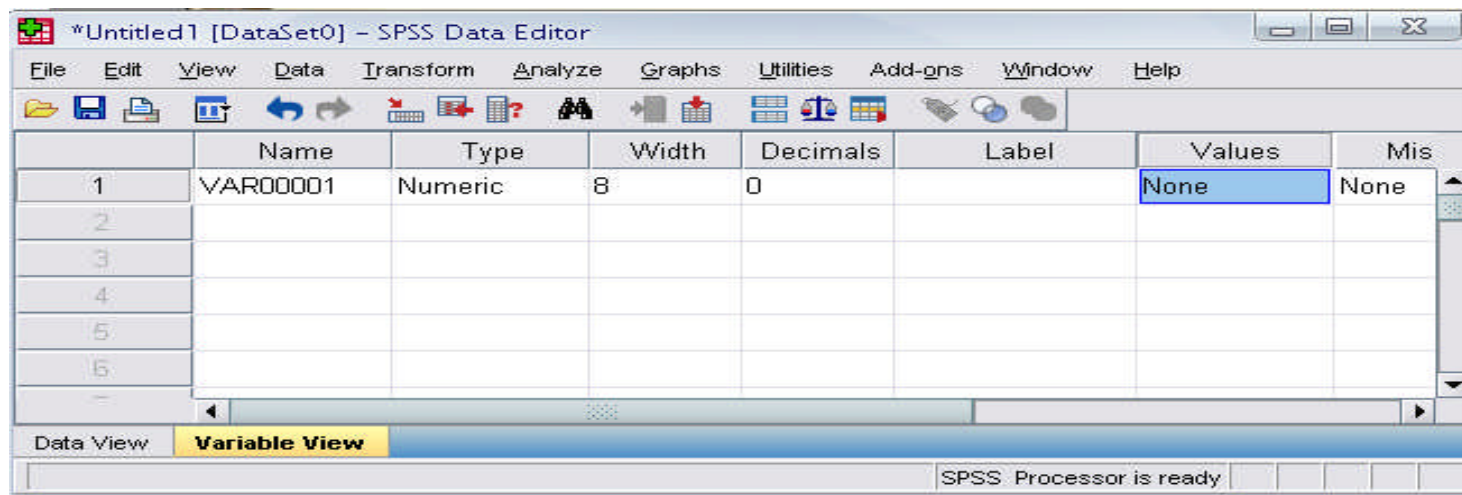
# Variable View window: Label

- Label
  - You can specify the details of the variable
  - You can write characters with spaces up to 256 characters



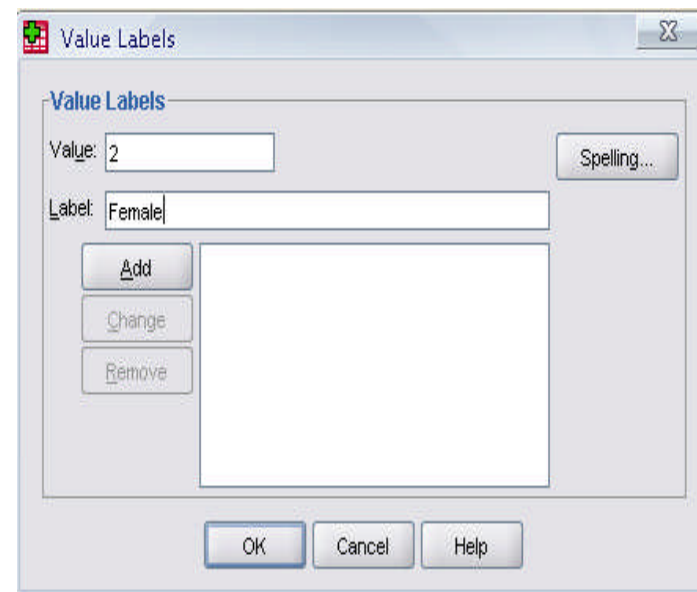
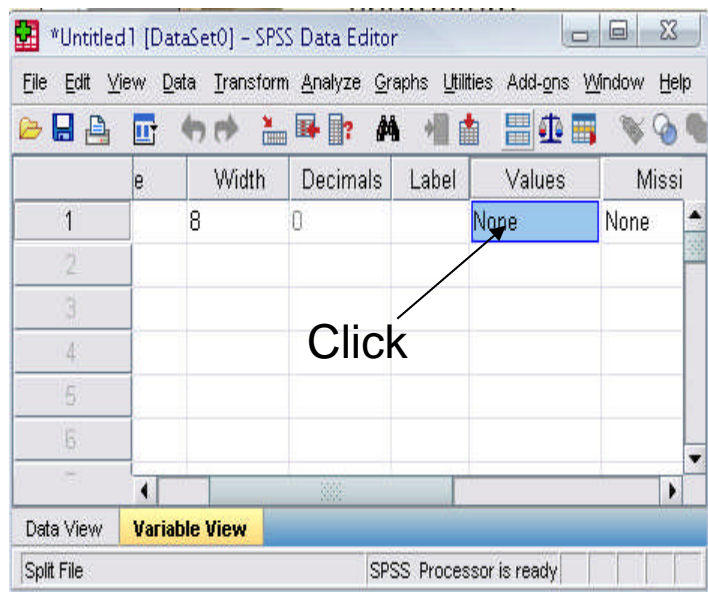
# Variable View window: Values

- Values
  - This is used and to suggest which numbers represent which categories when the variable represents a category



# Defining the value labels

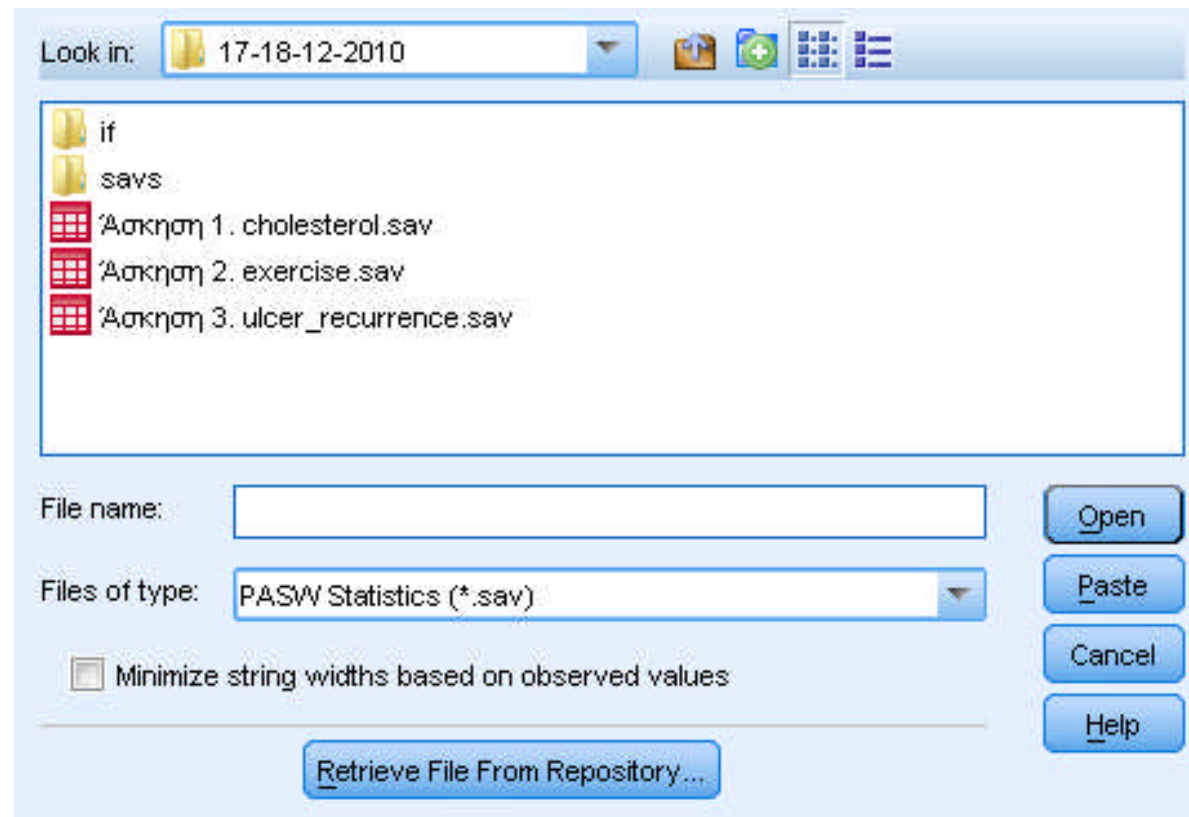
- Click the cell in the values column as shown below
- For the value, and the label, you can put up to 60 characters.
- After defining the values click add and then click OK.



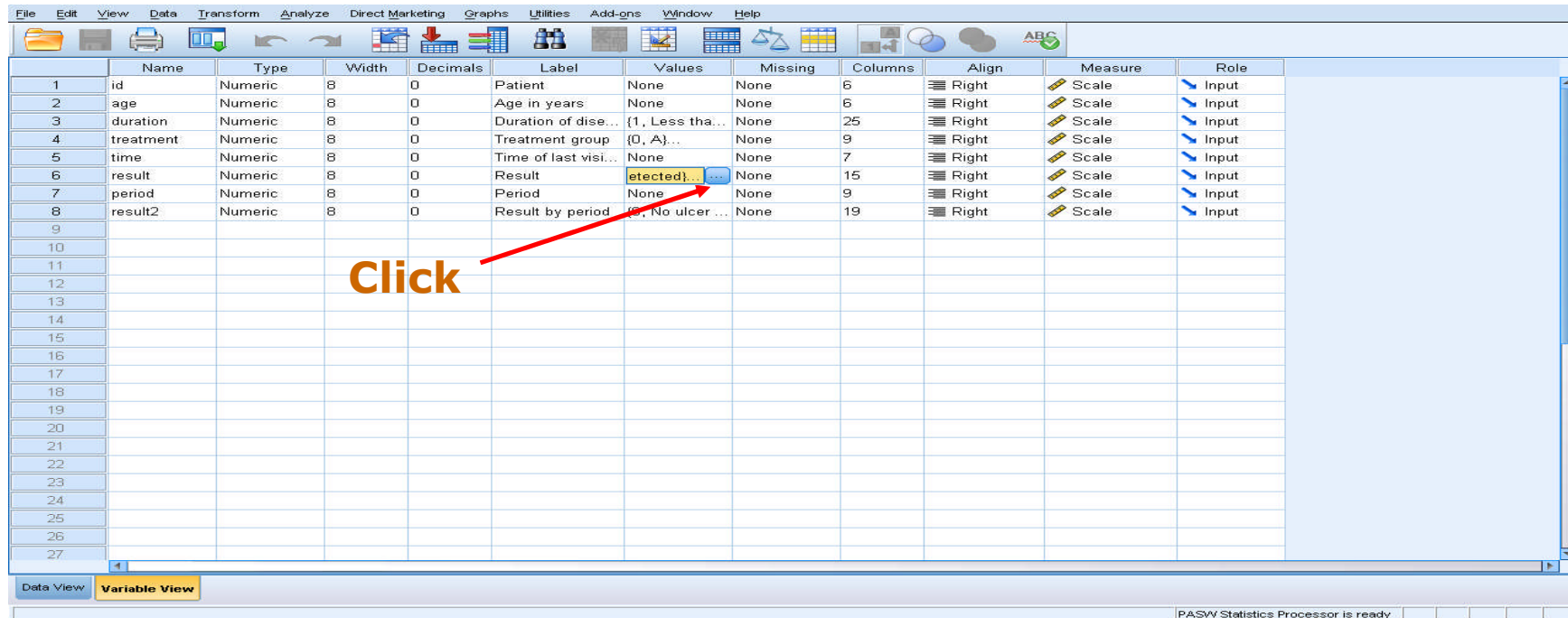
# How to make a new data set?

- We will create now a data set on our own.
- Types of variables
  - String
  - Numeric
- Are you ready for this creation procedure?

# Open “Άσκηση 3. ulcer\_recurrence”



# Example 1



The screenshot shows the SPSS Variable View window. A red arrow points to the 'Values' column for the 'result' variable, which contains the text 'etected)...'. The word 'Click' is written in orange text next to the arrow. The table below represents the data shown in the Variable View window.

|    | Name      | Type    | Width | Decimals | Label                | Values           | Missing | Columns | Align | Measure | Role  |
|----|-----------|---------|-------|----------|----------------------|------------------|---------|---------|-------|---------|-------|
| 1  | id        | Numeric | 8     | 0        | Patient              | None             | None    | 6       | Right | Scale   | Input |
| 2  | age       | Numeric | 8     | 0        | Age in years         | None             | None    | 6       | Right | Scale   | Input |
| 3  | duration  | Numeric | 8     | 0        | Duration of dise...  | {1, Less tha...  | None    | 25      | Right | Scale   | Input |
| 4  | treatment | Numeric | 8     | 0        | Treatment group      | {0, A}...        | None    | 9       | Right | Scale   | Input |
| 5  | time      | Numeric | 8     | 0        | Time of last visi... | None             | None    | 7       | Right | Scale   | Input |
| 6  | result    | Numeric | 8     | 0        | Result               | etected)...      | None    | 15      | Right | Scale   | Input |
| 7  | period    | Numeric | 8     | 0        | Period               | None             | None    | 9       | Right | Scale   | Input |
| 8  | result2   | Numeric | 8     | 0        | Result by period     | {0, No ulcer ... | None    | 19      | Right | Scale   | Input |
| 9  |           |         |       |          |                      |                  |         |         |       |         |       |
| 10 |           |         |       |          |                      |                  |         |         |       |         |       |
| 11 |           |         |       |          |                      |                  |         |         |       |         |       |
| 12 |           |         |       |          |                      |                  |         |         |       |         |       |
| 13 |           |         |       |          |                      |                  |         |         |       |         |       |
| 14 |           |         |       |          |                      |                  |         |         |       |         |       |
| 15 |           |         |       |          |                      |                  |         |         |       |         |       |
| 16 |           |         |       |          |                      |                  |         |         |       |         |       |
| 17 |           |         |       |          |                      |                  |         |         |       |         |       |
| 18 |           |         |       |          |                      |                  |         |         |       |         |       |
| 19 |           |         |       |          |                      |                  |         |         |       |         |       |
| 20 |           |         |       |          |                      |                  |         |         |       |         |       |
| 21 |           |         |       |          |                      |                  |         |         |       |         |       |
| 22 |           |         |       |          |                      |                  |         |         |       |         |       |
| 23 |           |         |       |          |                      |                  |         |         |       |         |       |
| 24 |           |         |       |          |                      |                  |         |         |       |         |       |
| 25 |           |         |       |          |                      |                  |         |         |       |         |       |
| 26 |           |         |       |          |                      |                  |         |         |       |         |       |
| 27 |           |         |       |          |                      |                  |         |         |       |         |       |

Value Labels

Value:

Label:

0 = "No ulcer detected"  
1 = "Ulcer detected"

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Click

|    | Name      | Type    | Width | Decimals | Label                | Values           | Missing | Columns | Align | Measure | Role  |
|----|-----------|---------|-------|----------|----------------------|------------------|---------|---------|-------|---------|-------|
| 1  | id        | Numeric | 8     | 0        | Patient              | None             | None    | 6       | Right | Scale   | Input |
| 2  | age       | Numeric | 8     | 0        | Age in years         | None             | None    | 6       | Right | Scale   | Input |
| 3  | duration  | Numeric | 8     | 0        | Duration of dise...  | {1, Less tha...  | None    | 25      | Right | Scale   | Input |
| 4  | treatment | Numeric | 8     | 0        | Treatment group      | {0, A}...        | None    | 9       | Right | Scale   | Input |
| 5  | time      | Numeric | 8     | 0        | Time of last visi... | None             | None    | 7       | Right | Scale   | Input |
| 6  | result    | Numeric | 8     | 0        | Result               | etected}...      | None    | 15      | Right | Scale   | Input |
| 7  | period    | Numeric | 8     | 0        | Period               | None             | None    | 9       | Right | Scale   | Input |
| 8  | result2   | Numeric | 8     | 0        | Result by period     | {0, No ulcer ... | None    | 19      | Right | Scale   | Input |
| 9  |           |         |       |          |                      |                  |         |         |       |         |       |
| 10 |           |         |       |          |                      |                  |         |         |       |         |       |
| 11 |           |         |       |          |                      |                  |         |         |       |         |       |
| 12 |           |         |       |          |                      |                  |         |         |       |         |       |
| 13 |           |         |       |          |                      |                  |         |         |       |         |       |
| 14 |           |         |       |          |                      |                  |         |         |       |         |       |
| 15 |           |         |       |          |                      |                  |         |         |       |         |       |
| 16 |           |         |       |          |                      |                  |         |         |       |         |       |
| 17 |           |         |       |          |                      |                  |         |         |       |         |       |
| 18 |           |         |       |          |                      |                  |         |         |       |         |       |
| 19 |           |         |       |          |                      |                  |         |         |       |         |       |
| 20 |           |         |       |          |                      |                  |         |         |       |         |       |
| 21 |           |         |       |          |                      |                  |         |         |       |         |       |
| 22 |           |         |       |          |                      |                  |         |         |       |         |       |
| 23 |           |         |       |          |                      |                  |         |         |       |         |       |
| 24 |           |         |       |          |                      |                  |         |         |       |         |       |
| 25 |           |         |       |          |                      |                  |         |         |       |         |       |
| 26 |           |         |       |          |                      |                  |         |         |       |         |       |
| 27 |           |         |       |          |                      |                  |         |         |       |         |       |

Data View Variable View

PASW Statistics Processor is ready

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1: id 5 Visible: 9 of 9 Variables

|    | id | age | duration          | treatment | time | result            | period | result2           | var | var | var |
|----|----|-----|-------------------|-----------|------|-------------------|--------|-------------------|-----|-----|-----|
| 1  | 5  | 56  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 2  | 8  | 41  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 3  | 14 | 27  | Less than 5 years | A         | 6    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 4  | 16 | 54  | Less than 5 years | A         | 6    | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 5  | 21 | 25  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 6  | 22 | 58  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 7  | 24 | 41  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 8  | 26 | 56  | Less than 5 years | A         | 3    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 9  | 29 | 72  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 10 | 33 | 76  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 11 | 37 | 44  | Less than 5 years | A         | 12   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 12 | 40 | 23  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 13 | 42 | 61  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 14 | 2  | 73  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 15 | 3  | 54  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 16 | 7  | 71  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 17 | 9  | 23  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 18 | 10 | 37  | Less than 5 years | B         | 5    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 19 | 11 | 38  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 20 | 15 | 47  | Less than 5 years | B         | 6    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 21 | 17 | 38  | Less than 5 years | B         | 10   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 22 | 20 | 75  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 23 | 23 | 63  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 24 | 25 | 47  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 25 | 30 | 59  | Less than 5 years | B         | 12   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 26 | 31 | 63  | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |

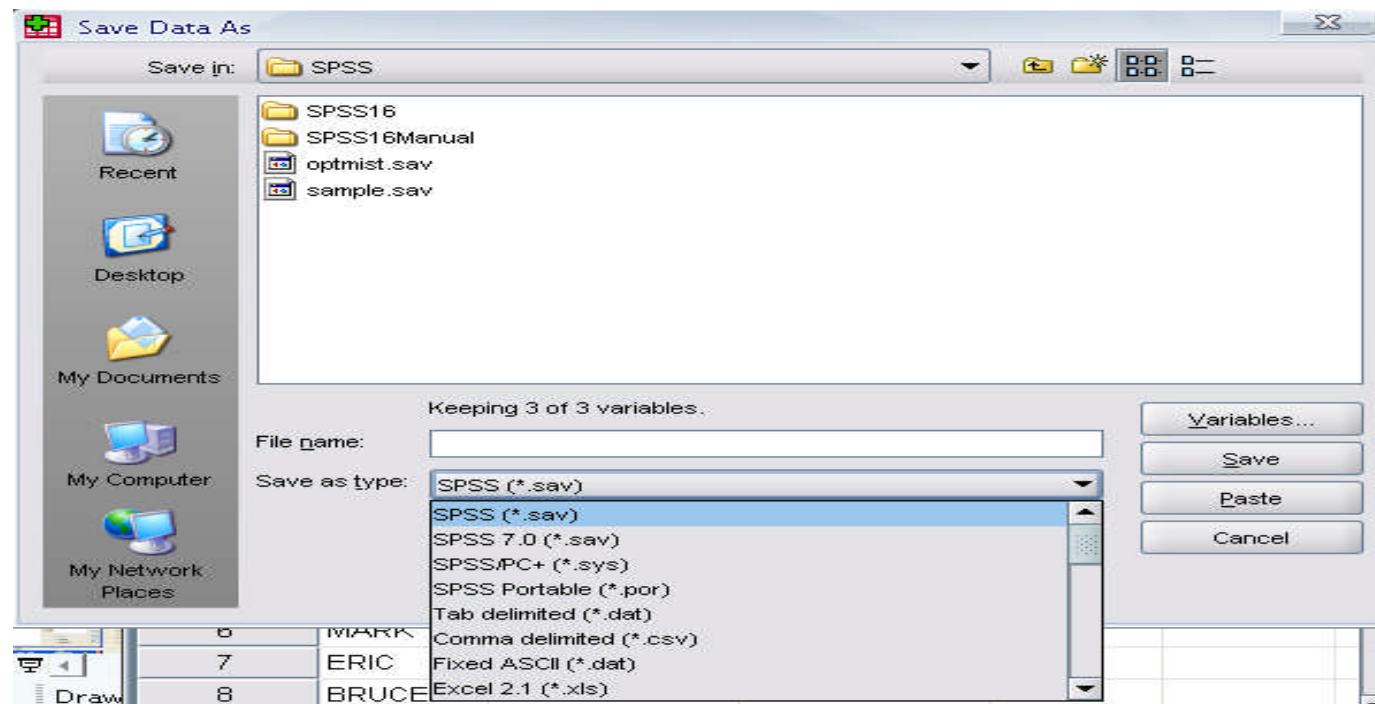
Data View Variable View

PASW Statistics Processor is ready



# Saving the data

- To save the data file you created simply click 'file' and click 'save as.' You can save the file in different forms by clicking "Save as type."



# Sorting the data

- Click 'Data' and then click Sort Cases

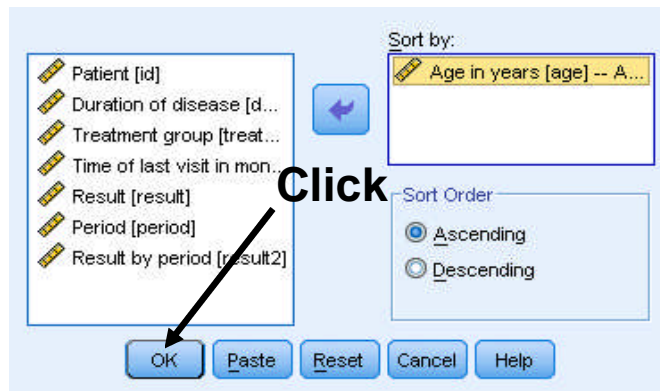
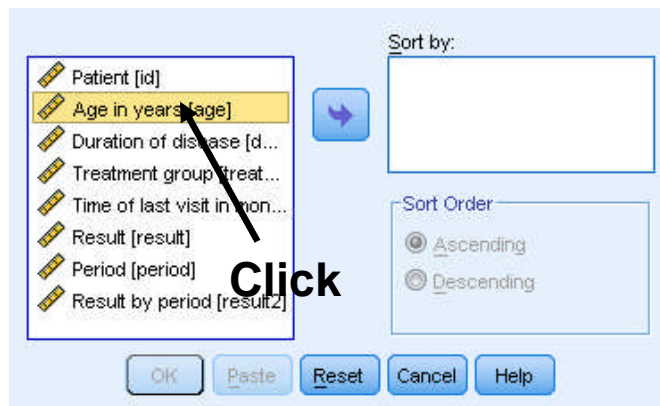
The screenshot shows the SPSS Data View window. The 'Data' menu is open, and 'Sort Cases...' is highlighted. The data table has 26 rows and 10 columns. The first column is 'id', and the next three are 'on', 'treatment', and 'time'. The remaining columns are 'result', 'period', 'result2', and three empty columns labeled 'var'. The data is as follows:

|    | id | on                | treatment | time | result            | period | result2           | var | var | var |
|----|----|-------------------|-----------|------|-------------------|--------|-------------------|-----|-----|-----|
| 1  | 1  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 2  | 2  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 3  | 3  | Less than 5 years | A         | 6    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 4  | 4  | Less than 5 years | A         | 6    | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 5  | 5  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 6  | 6  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 7  | 7  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 8  | 8  | Less than 5 years | A         | 3    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 9  | 9  | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 10 | 10 | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 11 | 11 | Less than 5 years | A         | 12   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 12 | 12 | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 13 | 13 | Less than 5 years | A         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 14 | 14 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 15 | 15 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 16 | 16 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 17 | 17 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 18 | 18 | Less than 5 years | B         | 5    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 19 | 19 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 20 | 20 | Less than 5 years | B         | 6    | Ulcer detected    | 1      | Ulcer detected    |     |     |     |
| 21 | 21 | Less than 5 years | B         | 10   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 22 | 22 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 23 | 23 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 24 | 24 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |
| 25 | 25 | Less than 5 years | B         | 12   | Ulcer detected    | 1      | No ulcer detected |     |     |     |
| 26 | 26 | Less than 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |     |     |     |

The bottom status bar shows 'Sort Cases...' and 'PASW Statistics Processor is ready'.

# Sorting the data (cont'd)

- Double Click 'Name of the students.' Then click ok.

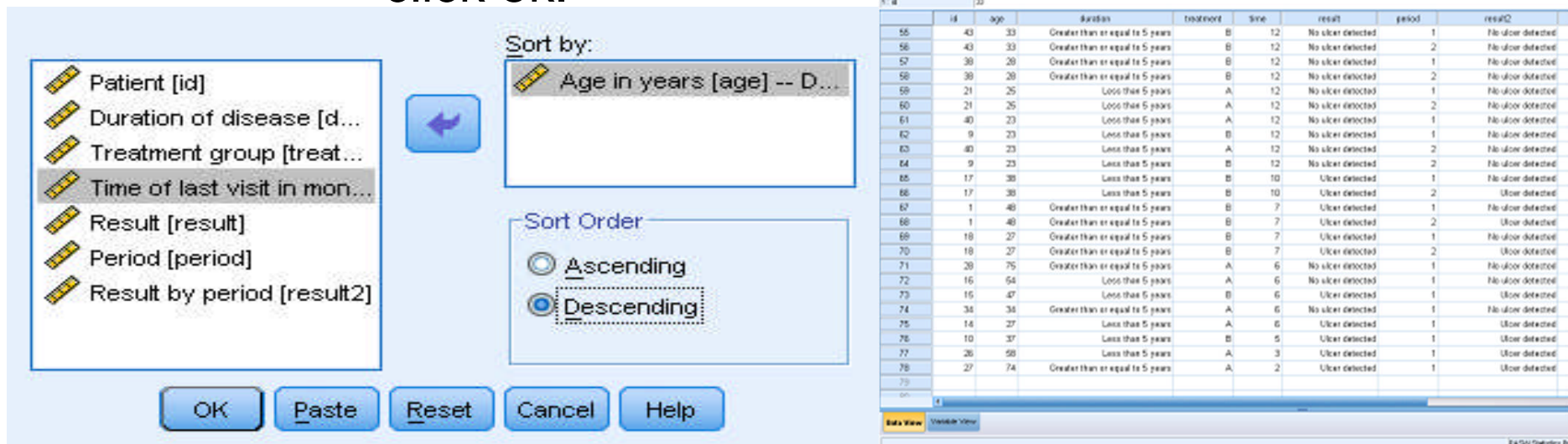


| File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help |    |     |                                  |           |      |              |
|---|----|-----|----------------------------------|-----------|------|--------------|
| 1: id 40  |    |     |                                  |           |      |              |
|   | id | age | duration                         | treatment | time | result       |
| 1   | 40 | 23  | Less than 5 years                | A         | 12   | No ulcer det |
| 2   | 9  | 23  | Less than 5 years                | B         | 12   | No ulcer det |
| 3   | 40 | 23  | Less than 5 years                | A         | 12   | No ulcer det |
| 4   | 9  | 23  | Less than 5 years                | B         | 12   | No ulcer det |
| 5   | 21 | 25  | Less than 5 years                | A         | 12   | No ulcer det |
| 6   | 21 | 25  | Less than 5 years                | A         | 12   | No ulcer det |
| 7   | 14 | 27  | Less than 5 years                | A         | 6    | Ulcer det    |
| 8   | 18 | 27  | Greater than or equal to 5 years | B         | 7    | Ulcer det    |
| 9   | 18 | 27  | Greater than or equal to 5 years | B         | 7    | Ulcer det    |
| 10  | 38 | 28  | Greater than or equal to 5 years | B         | 12   | No ulcer det |
| 11  | 38 | 28  | Greater than or equal to 5 years | B         | 12   | No ulcer det |
| 12  | 43 | 33  | Greater than or equal to 5 years | B         | 12   | No ulcer det |
| 13  | 43 | 33  | Greater than or equal to 5 years | B         | 12   | No ulcer det |
| 14  | 34 | 34  | Greater than or equal to 5 years | A         | 6    | No ulcer det |
| 15  | 35 | 36  | Less than 5 years                | B         | 12   | No ulcer det |
| 16  | 35 | 36  | Less than 5 years                | B         | 12   | No ulcer det |
| 17  | 10 | 37  | Less than 5 years                | B         | 5    | Ulcer det    |
| 18  | 11 | 38  | Less than 5 years                | B         | 12   | No ulcer det |
| 19  | 17 | 38  | Less than 5 years                | B         | 10   | Ulcer det    |
| 20  | 11 | 38  | Less than 5 years                | B         | 12   | No ulcer det |
| 21  | 17 | 38  | Less than 5 years                | B         | 10   | Ulcer det    |
| 22  | 13 | 38  | Greater than or equal to 5 years | A         | 12   | No ulcer det |
| 23  | 13 | 38  | Greater than or equal to 5 years | A         | 12   | No ulcer det |
| 24  | 8  | 41  | Less than 5 years                | A         | 12   | No ulcer det |
| 25  | 24 | 41  | Less than 5 years                | A         | 12   | No ulcer det |
| 26  | 8  | 41  | Less than 5 years                | A         | 12   | No ulcer det |

Data View Variable View

## Example 2

- How would you sort the data by the time of their last visit of patients in descending order?
- Answer
  - Click data, sort cases, double click 'time of last visit in months,' click 'descending,' and finally click ok.



The image shows the SPSS Sort Cases dialog box and the Data View window. In the dialog box, 'Time of last visit in months' is selected in the 'Sort by' list, and 'Descending' is selected in the 'Sort Order' section. The Data View window shows a table with columns: id, age, duration, treatment, time, result, period, and result2.

**Sort Cases Dialog Box:**

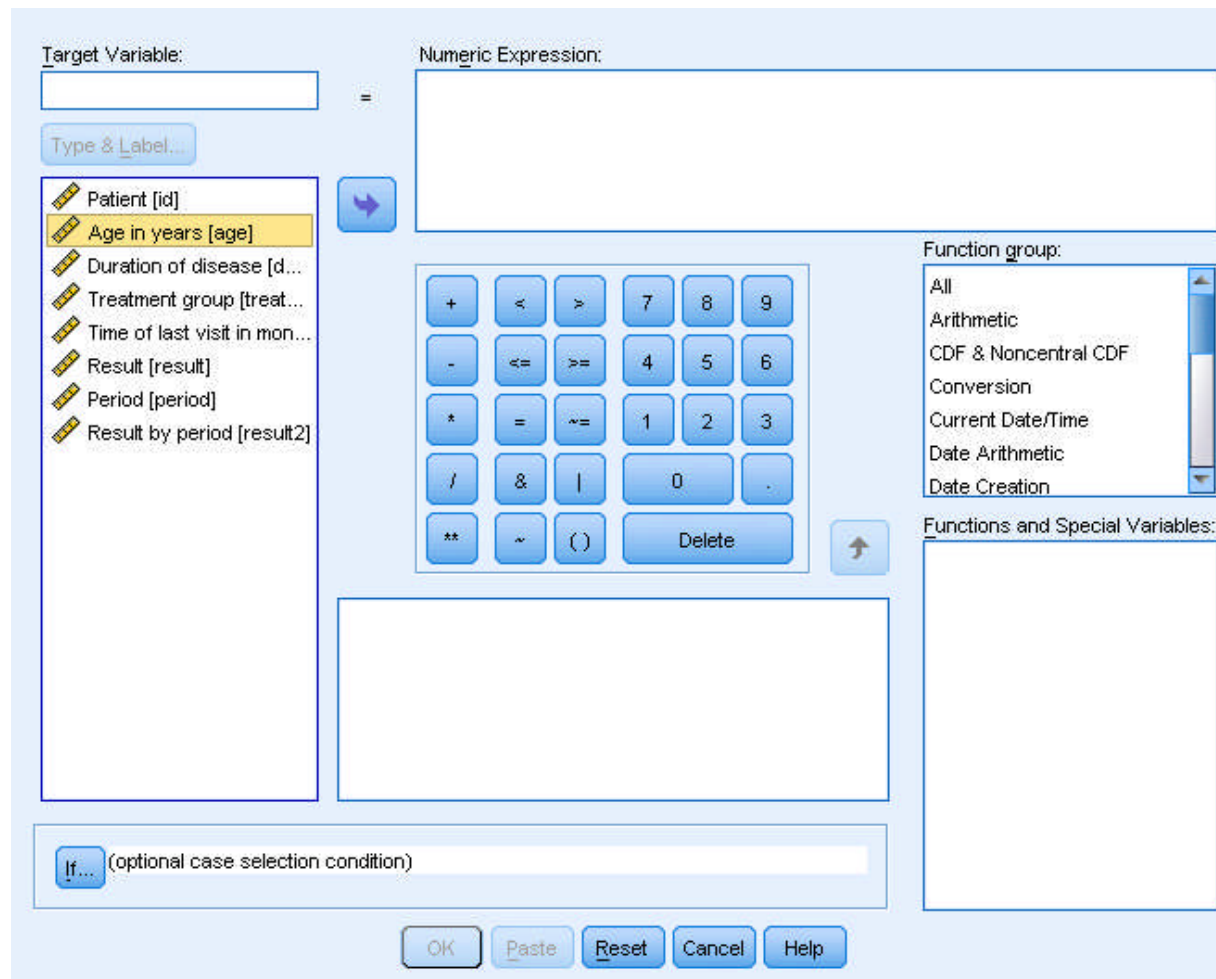
- Sort by:** Time of last visit in months
- Sort Order:** Descending

**Data View Window:**

|    | id | age | duration                         | treatment | time | result            | period | result2           |
|----|----|-----|----------------------------------|-----------|------|-------------------|--------|-------------------|
| 55 | 43 | 33  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |
| 56 | 43 | 33  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 2      | No ulcer detected |
| 57 | 38 | 28  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected |
| 58 | 38 | 28  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 2      | No ulcer detected |
| 59 | 21 | 25  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected |
| 60 | 21 | 25  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected |
| 61 | 40 | 23  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected |
| 62 | 9  | 23  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected |
| 63 | 40 | 23  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected |
| 64 | 9  | 23  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected |
| 65 | 17 | 38  | Less than 5 years                | B         | 10   | Ulcer detected    | 1      | No ulcer detected |
| 66 | 17 | 38  | Less than 5 years                | B         | 10   | Ulcer detected    | 2      | Ulcer detected    |
| 67 | 1  | 48  | Greater than or equal to 5 years | B         | 7    | Ulcer detected    | 1      | No ulcer detected |
| 68 | 1  | 48  | Greater than or equal to 5 years | B         | 7    | Ulcer detected    | 2      | Ulcer detected    |
| 69 | 18 | 27  | Greater than or equal to 5 years | B         | 7    | Ulcer detected    | 1      | No ulcer detected |
| 70 | 18 | 27  | Greater than or equal to 5 years | B         | 7    | Ulcer detected    | 2      | Ulcer detected    |
| 71 | 28 | 75  | Greater than or equal to 5 years | A         | 6    | No ulcer detected | 1      | No ulcer detected |
| 72 | 16 | 64  | Less than 5 years                | A         | 6    | No ulcer detected | 1      | No ulcer detected |
| 73 | 16 | 47  | Less than 5 years                | B         | 6    | Ulcer detected    | 1      | Ulcer detected    |
| 74 | 34 | 34  | Greater than or equal to 5 years | A         | 6    | No ulcer detected | 1      | No ulcer detected |
| 75 | 14 | 27  | Less than 5 years                | A         | 6    | Ulcer detected    | 1      | Ulcer detected    |
| 76 | 10 | 37  | Less than 5 years                | B         | 5    | Ulcer detected    | 1      | Ulcer detected    |
| 77 | 26 | 58  | Less than 5 years                | A         | 3    | Ulcer detected    | 1      | Ulcer detected    |
| 78 | 27 | 74  | Greater than or equal to 5 years | A         | 2    | Ulcer detected    | 1      | Ulcer detected    |

# Transforming data

- Click 'Transform' and then click 'Compute Variable...'



The image shows the 'Compute Variable' dialog box in SPSS. It is used to create a new variable based on an expression of existing variables and functions.

**Target Variable:** A text box for naming the new variable. Below it is a button labeled 'Type & Label...'. To the right of this box is an equals sign (=).

**Numeric Expression:** A large text box for entering the expression. To its left is a blue arrow button pointing right.

**Variable List:** A list of variables on the left side of the dialog. The variables are: Patient [id], Age in years [age], Duration of disease [d...], Treatment group [treat...], Time of last visit in mon..., Result [result], Period [period], and Result by period [result2]. The 'Age in years [age]' variable is highlighted with a yellow background.

**Operator Pad:** A grid of buttons for mathematical and logical operators. The buttons are: +, <, >, 7, 8, 9, -, <=, >=, 4, 5, 6, \*, =, ~=, 1, 2, 3, /, &, |, 0, ., \*\*, ~, (, ), and a 'Delete' button.

**Function group:** A list box showing categories of functions. The visible options are: All, Arithmetic, CDF & Noncentral CDF, Conversion, Current Date/Time, Date Arithmetic, and Date Creation. A scroll bar is on the right.

**Functions and Special Variables:** A large empty text box for listing specific functions and variables.

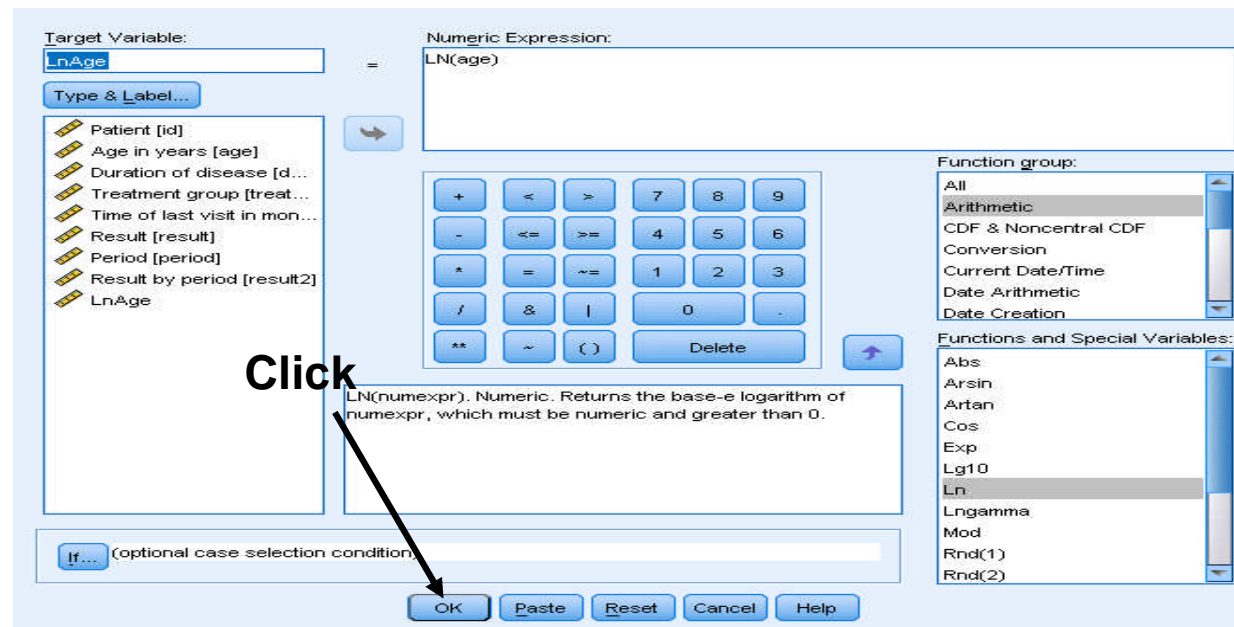
**If... (optional case selection condition):** A text box at the bottom left for specifying conditions under which the variable should be computed.

**Buttons:** At the bottom of the dialog are five buttons: OK, Paste, Reset, Cancel, and Help.



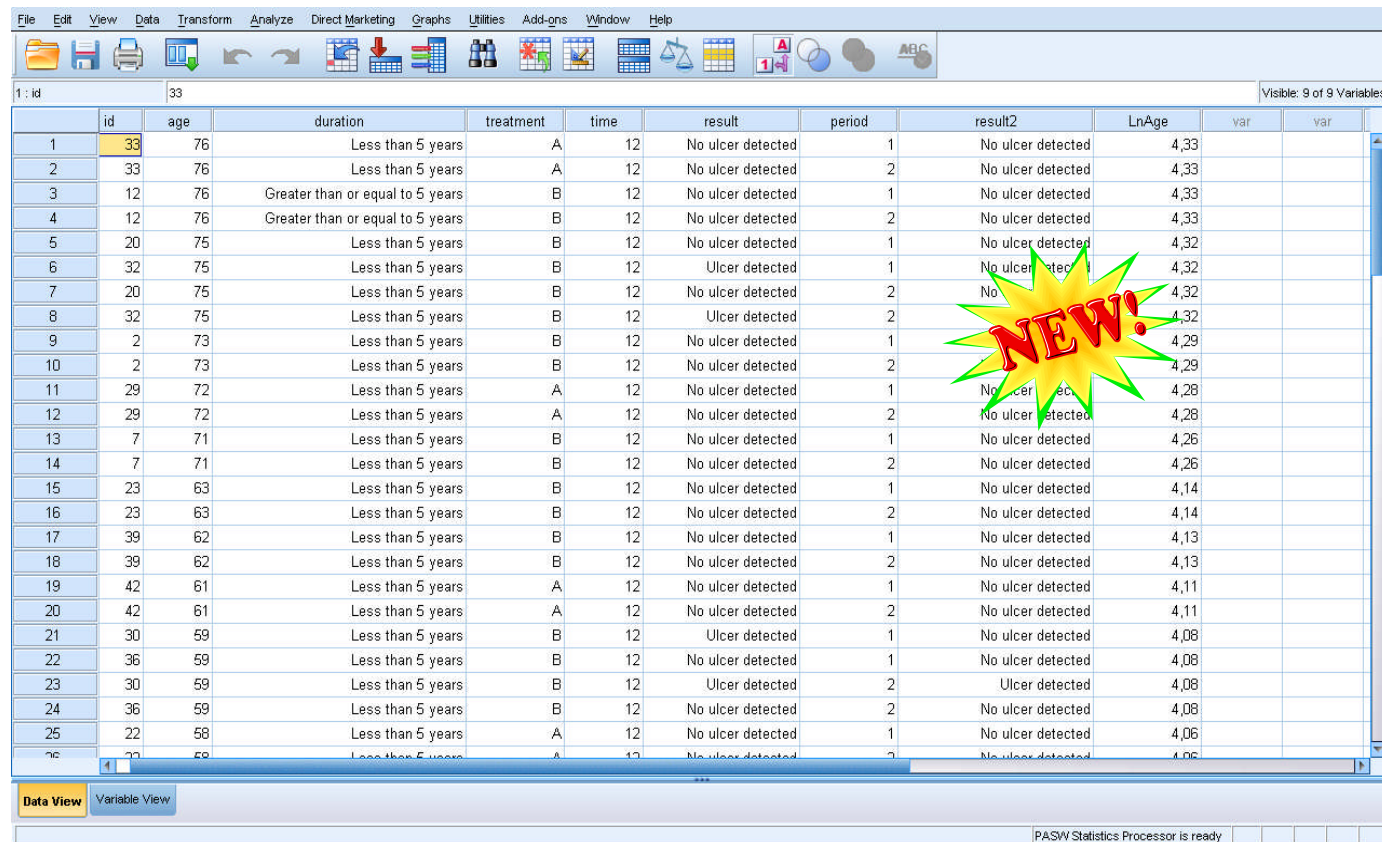
# Transforming data (cont'd)

- Example: Adding a new variable named 'LnAge' which is the natural log of height
  - Type in LnAge in the 'Target Variable' box. Insert 'ln(Age in years [age])' in the 'Numeric Expression' box. Click OK



# Transforming data (cont'd)

- A new variable 'Lnheight' is added to the table



Visible: 9 of 9 Variables

|    | id | age | duration                         | treatment | time | result            | period | result2           | LnAge | var | var |
|----|----|-----|----------------------------------|-----------|------|-------------------|--------|-------------------|-------|-----|-----|
| 1  | 33 | 76  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,33  |     |     |
| 2  | 33 | 76  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,33  |     |     |
| 3  | 12 | 76  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,33  |     |     |
| 4  | 12 | 76  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,33  |     |     |
| 5  | 20 | 75  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,32  |     |     |
| 6  | 32 | 75  | Less than 5 years                | B         | 12   | Ulcer detected    | 1      | No ulcer detected | 4,32  |     |     |
| 7  | 20 | 75  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,32  |     |     |
| 8  | 32 | 75  | Less than 5 years                | B         | 12   | Ulcer detected    | 2      | No ulcer detected | 4,32  |     |     |
| 9  | 2  | 73  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,29  |     |     |
| 10 | 2  | 73  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,29  |     |     |
| 11 | 29 | 72  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,28  |     |     |
| 12 | 29 | 72  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,28  |     |     |
| 13 | 7  | 71  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,26  |     |     |
| 14 | 7  | 71  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,26  |     |     |
| 15 | 23 | 63  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,14  |     |     |
| 16 | 23 | 63  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,14  |     |     |
| 17 | 39 | 62  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,13  |     |     |
| 18 | 39 | 62  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,13  |     |     |
| 19 | 42 | 61  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,11  |     |     |
| 20 | 42 | 61  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,11  |     |     |
| 21 | 30 | 59  | Less than 5 years                | B         | 12   | Ulcer detected    | 1      | No ulcer detected | 4,08  |     |     |
| 22 | 36 | 59  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,08  |     |     |
| 23 | 30 | 59  | Less than 5 years                | B         | 12   | Ulcer detected    | 2      | Ulcer detected    | 4,08  |     |     |
| 24 | 36 | 59  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,08  |     |     |
| 25 | 22 | 58  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,06  |     |     |
| 26 | 22 | 58  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,06  |     |     |

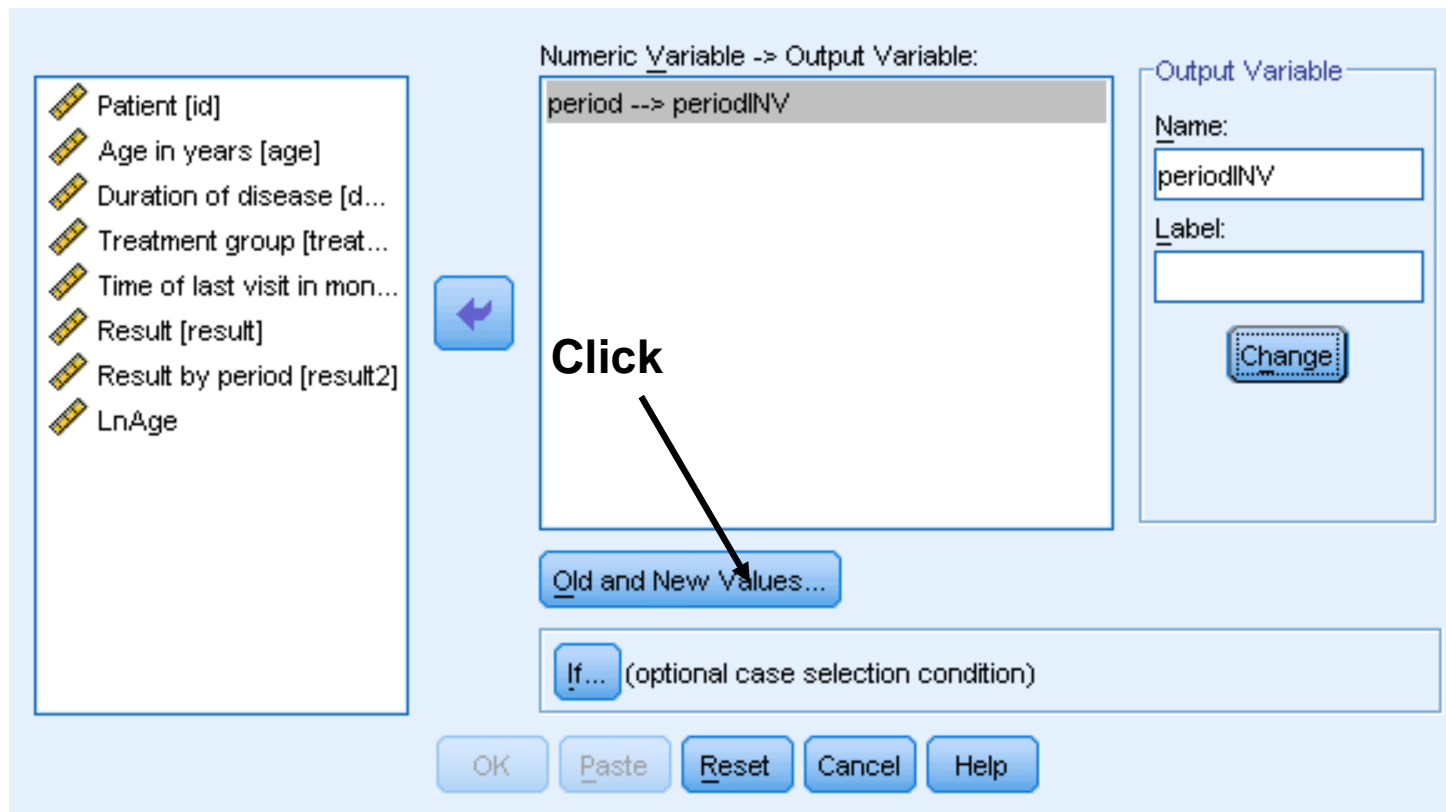
\*\*\*

Data View Variable View

PASW Statistics Processor is ready

# Reverse Coding

- Transform > Recode > Into Different Variables





# Reverse Coding

Old Value

☒ Value:

☐ System-missing

☐ System- or user-missing

☐ Range:

through

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

New Value

☒ Value:

☐ System-missing

☐ Copy old value(s)

Old --> New:

1 --> 2  
2 --> 1

Add

Change

Remove

☐ Output variables are strings      Width: 

8

☐ Convert numeric strings to numbers ('5' -> 5)

Continue

Cancel

Help

# Example 3

- Create a new variable named “sqrtAge” which is the square root of Age
- Answer

Target Variable:  = Numeric Expression:

Type & Label...

Function group:

- All
- Arithmetic
- CDF & Noncentral CDF
- Conversion
- Current Date/Time
- Date Arithmetic
- Date Creation

Functions and Special Variables:

- Cos
- Exp
- Lg10
- Ln
- Lngamma
- Mod
- Rnd(1)
- Rnd(2)
- Rnd(3)
- Sin
- Sqrt

Calculator:

NUMEXPR:

OK Paste Reset Cancel Help

4: treatment

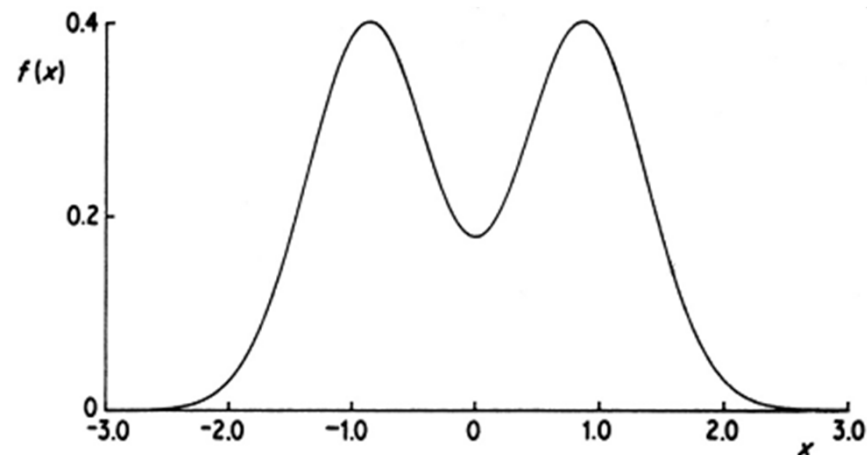
|    | duration                         | treatment | time | result            | period | result2           | LnAge | SQRTAge | var | var | var |
|----|----------------------------------|-----------|------|-------------------|--------|-------------------|-------|---------|-----|-----|-----|
| 1  | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,33  | 8,72    |     |     |     |
| 2  | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,33  | 8,72    |     |     |     |
| 3  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,33  | 8,72    |     |     |     |
| 4  | Greater than or equal to 5 years | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,33  | 8,72    |     |     |     |
| 5  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,32  | 8,66    |     |     |     |
| 6  | Less than 5 years                | B         | 12   | Ulcer detected    | 1      | No ulcer detected | 4,32  | 8,66    |     |     |     |
| 7  | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,32  | 8,66    |     |     |     |
| 8  | Less than 5 years                | B         | 12   | Ulcer detected    | 2      | Ulcer detected    | 4,32  | 8,66    |     |     |     |
| 9  | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,29  | 8,54    |     |     |     |
| 10 | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,29  | 8,54    |     |     |     |
| 11 | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,28  | 8,49    |     |     |     |
| 12 | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,28  | 8,49    |     |     |     |
| 13 | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,26  | 8,43    |     |     |     |
| 14 | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,26  | 8,43    |     |     |     |
| 15 | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,14  | 7,94    |     |     |     |
| 16 | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,14  | 7,94    |     |     |     |
| 17 | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,13  | 7,87    |     |     |     |
| 18 | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,13  | 7,87    |     |     |     |
| 19 | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,11  | 7,81    |     |     |     |
| 20 | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,11  | 7,81    |     |     |     |
| 21 | Less than 5 years                | B         | 12   | Ulcer detected    | 1      | No ulcer detected | 4,08  | 7,68    |     |     |     |
| 22 | Less than 5 years                | B         | 12   | No ulcer detected | 1      | No ulcer detected | 4,08  | 7,68    |     |     |     |
| 23 | Less than 5 years                | B         | 12   | Ulcer detected    | 2      | Ulcer detected    | 4,08  | 7,68    |     |     |     |
| 24 | Less than 5 years                | B         | 12   | No ulcer detected | 2      | No ulcer detected | 4,08  | 7,68    |     |     |     |
| 25 | Less than 5 years                | A         | 12   | No ulcer detected | 1      | No ulcer detected | 4,06  | 7,62    |     |     |     |
| 26 | Less than 5 years                | A         | 12   | No ulcer detected | 2      | No ulcer detected | 4,06  | 7,62    |     |     |     |

Data View Variable View

PASW Statistics Processor is ready

# Applications

- If the distribution turns out to be bimodal, a median or mean will be of little use.
- More often bimodal distributions give us a clear signal that we have lumped together two (or more) data sets which should have been separate.
  - Example: If women are paid less than men, income data for a given level of education turns to be bimodal.



# Normality explorations

# Data Analysis

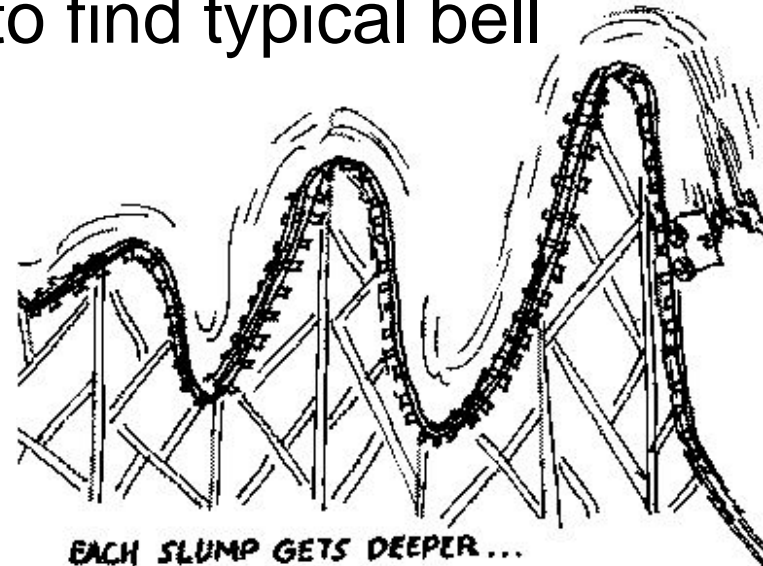
- Data analysis embraces both the problem of finding an appropriate model, on the one hand, and model estimation and testing, on the other.
- In this context normality assumption becomes important.
- In social sciences, it is hard to find typical bell shaped normal distribution.

# Data Analysis

- Data analysis embraces both the problem of finding an appropriate model, on the one hand, and model estimation and testing, on the other.
- In this context normality assumption becomes important.
- In social sciences, it is hard to find typical bell shaped normal distribution.

# Data Analysis

- Data analysis embraces both the problem of finding an appropriate model, on the one hand, and model estimation and testing, on the other.
- In this context normality assumption becomes important.
- In social sciences, it is hard to find typical bell shaped normal distribution.



# Open “Άσκηση 1. cholesterol”

The screenshot displays the PASW Statistics Processor interface. The 'Analyze' menu is open, and the 'Explore...' option is selected. The 'Explore' dialog box is shown with the following settings:

- Dependent List:** Cholesterol
- Factor List:** Group
- Label Cases by:** (empty)
- Display:** ☒ Both, ☐ Statistics, ☐ Plots

The background data table is as follows:

|    | Patient | Group | Day |
|----|---------|-------|-----|
| 6  | 6       | 1     | 2   |
| 7  | 7       | 1     | 2   |
| 8  | 8       | 1     | 2   |
| 9  | 9       | 1     | 2   |
| 10 | 10      | 1     | 2   |
| 11 | 11      | 1     | 2   |
| 12 | 12      | 1     | 2   |
| 13 | 13      | 1     | 2   |
| 14 | 14      | 1     | 2   |
| 15 | 15      | 1     | 2   |
| 16 | 16      | 1     | 2   |
| 17 | 17      | 1     | 2   |
| 18 | 18      | 1     | 2   |
| 19 | 19      | 1     | 2   |
| 20 | 20      | 1     | 2   |
| 21 | 21      | 1     | 2   |
| 22 | 22      | 1     | 2   |
| 23 | 23      | 1     | 2   |
| 24 | 24      | 1     | 2   |
| 25 | 25      | 1     | 2   |
| 26 | 26      | 1     | 2   |
| 27 | 27      | 1     | 2   |
| 28 | 28      | 1     | 2   |
| 29 | 1       | 1     | 4   |
| 30 | 2       | 1     | 4   |

At the bottom of the window, the status bar indicates 'PASW Statistics Processor is ready'.



**Boxplots**

☐ Factor levels together

☐ Dependents together

☒ None

**Descriptive**

☐ Stem-and-leaf

☐ Histogram

☒ Normality plots with tests

**Spread vs Level with Levene Test**

☒ None

☐ Power estimation

☐ Transformed Power: Natural log

☐ Untransformed

Continue Cancel Help

| Tests of Normality |   |                                 |    |                   |              |    |      |
|--------------------|---|---------------------------------|----|-------------------|--------------|----|------|
| Group              |   | Kolmogorov-Smirnov <sup>a</sup> |    |                   | Shapiro-Wilk |    |      |
|                    |   | Statistic                       | df | Sig.              | Statistic    | df | Sig. |
| Cholesterol        | 1 | ,063                            | 75 | ,200 <sup>*</sup> | ,989         | 75 | ,740 |
|                    | 2 | ,126                            | 30 | ,200 <sup>*</sup> | ,939         | 30 | ,088 |

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Dependent List:

Cholesterol

Factor List:

Label Cases by:

Display

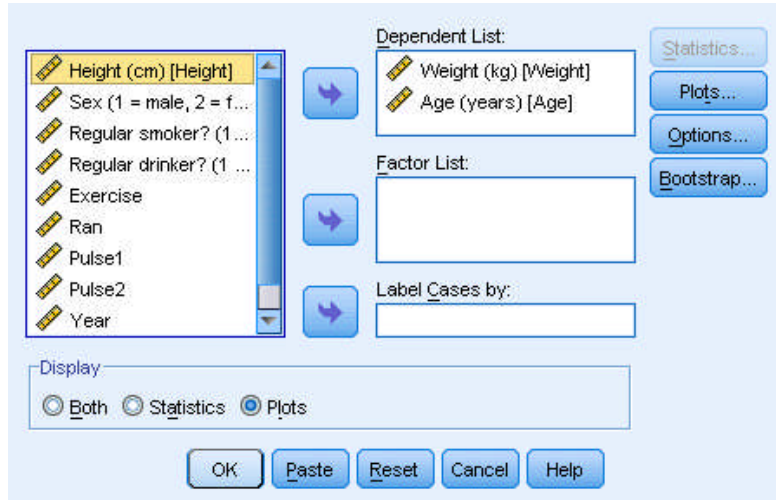
☒ Both ☐ Statistics ☐ Plots

OK Paste Reset Cancel Help

| Tests of Normality |                                 |     |      |              |     |      |
|--------------------|---------------------------------|-----|------|--------------|-----|------|
|                    | Kolmogorov-Smirnov <sup>a</sup> |     |      | Shapiro-Wilk |     |      |
|                    | Statistic                       | df  | Sig. | Statistic    | df  | Sig. |
| Cholesterol        | ,074                            | 105 | ,193 | ,983         | 105 | ,198 |

a. Lilliefors Significance Correction

# Open “Άσκηση 2. exercise”



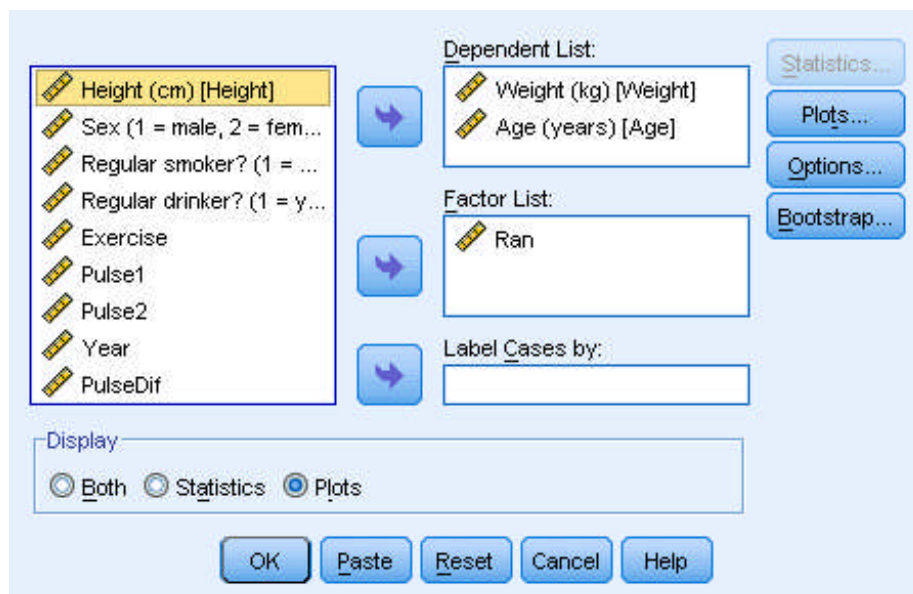
The image shows the SPSS One-Sample Kolmogorov-Smirnov Test dialog box. On the left is a list of variables: Height (cm) [Height], Sex (1 = male, 2 = f...), Regular smoker? (1...), Regular drinker? (1 ...), Exercise, Ran, Pulse1, Pulse2, and Year. The 'Dependent List' contains 'Weight (kg) [VWeight]' and 'Age (years) [Age]'. The 'Factor List' is empty. The 'Label Cases by:' field is also empty. On the right are buttons for 'Statistics...', 'Plots...', 'Options...', and 'Bootstrap...'. At the bottom, there is a 'Display' section with radio buttons for 'Both', 'Statistics', and 'Plots' (which is selected). Below this are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

## Tests of Normality

|             | Kolmogorov-Smirnov <sup>a</sup> |     |      | Shapiro-Wilk |     |      |
|-------------|---------------------------------|-----|------|--------------|-----|------|
|             | Statistic                       | df  | Sig. | Statistic    | df  | Sig. |
| Weight (kg) | ,126                            | 107 | ,000 | ,956         | 107 | ,001 |
| Age (years) | ,269                            | 107 | ,000 | ,563         | 107 | ,000 |

a. Lilliefors Significance Correction

# Open “Άσκηση 2. exercise”



The image shows the SPSS One-Way ANOVA dialog box. On the left is a list of variables: Height (cm) [Height], Sex (1 = male, 2 = fem...), Regular smoker? (1 = ...), Regular drinker? (1 = y..., Exercise, Pulse1, Pulse2, Year, and PulseDif. In the center, there are three boxes with arrows pointing to them: 'Dependent List:' containing 'Weight (kg) [Weight]' and 'Age (years) [Age]', 'Factor List:' containing 'Ran', and 'Label Cases by:' which is empty. On the right are buttons for 'Statistics...', 'Plots...', 'Options...', and 'Bootstrap...'. At the bottom left is a 'Display' section with radio buttons for 'Both', 'Statistics', and 'Plots' (which is selected). At the bottom are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

**Tests of Normality**

| Ran         |   | Kolmogorov-Smirnov <sup>a</sup> |    |      | Shapiro-Wilk |    |      |
|-------------|---|---------------------------------|----|------|--------------|----|------|
|             |   | Statistic                       | df | Sig. | Statistic    | df | Sig. |
| Weight (kg) | 1 | ,115                            | 45 | ,166 | ,967         | 45 | ,230 |
|             | 2 | ,141                            | 62 | ,004 | ,922         | 62 | ,001 |
| Age (years) | 1 | ,325                            | 45 | ,000 | ,638         | 45 | ,000 |
|             | 2 | ,281                            | 62 | ,000 | ,554         | 62 | ,000 |

a. Lilliefors Significance Correction

