



# Introduction to SPSS

**Dr. Md. Atiqul Islam**

M.Sc. (SUST, BD), M.Sc. (UHasselt, BE), PhD (RuG, NL)

Professor

Department of Statistics

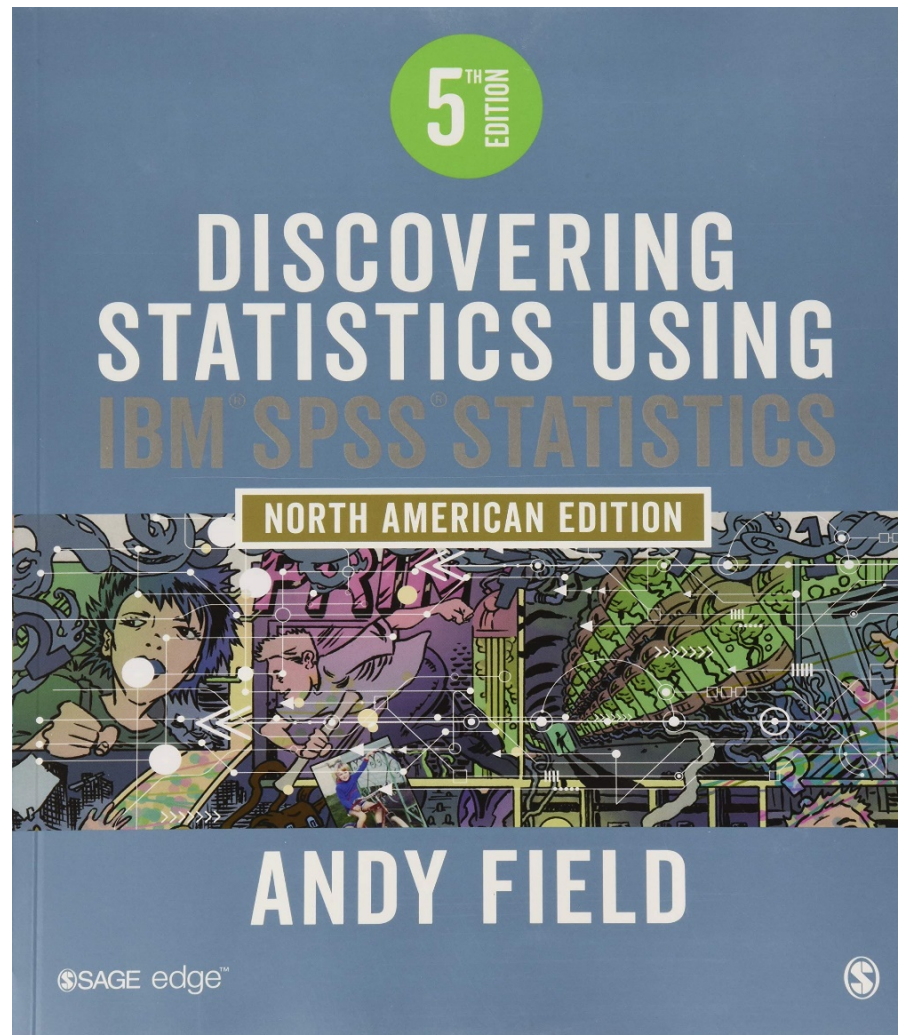
Jagannath University, Dhaka-1100

E-mail: [atique@stat.jnu.ac.bd](mailto:atique@stat.jnu.ac.bd)

*August 24, 2025*

- ▷ Background of SPSS
- ▷ Data Entry in SPSS
- ▷ Types of SPSS Files
- ▷ Import External Data into SPSS
- ▷ Menu Bar
  - ▷ “DATA” Menu Bar
  - ▷ “TRANSFORM” Menu Bar
  - ▷ “Analyze” Menu Bar
- ▷ Clean and edit data
- ▷ Descriptive statistics

- ▷ Field, A. (2018), *Discovering Statistics Using IBM SPSS Statistics*, 5<sup>th</sup> Edition, Sage Publications, Inc.



# Background of SPSS



- ▷ **SPSS** stands for “**S**tatistical **P**ackage for the **S**ocial **S**ciences”.
- ▷ It has also been known as “**PASW**” which stands for **P**redictive **A**alytics **S**oft**W**are.
- ▷ Powerful, user-friendly software package for the manipulation and statistical analysis of data.
- ▷ First version released in 1968. It was acquired by **I**nternational **B**usiness **M**achines (**IBM**) in 2009.
- ▷ It is called now **IBM SPSS**.
- ▷ Latest version of SPSS is **IBM SPSS 31.0.0** (on June 10, 2025).
- ▷ Alternative to SPSS
  - Tableau
  - JASP

# Background of SPSS



- ▷ Most widely used programs for statistical analysis in social and medical sciences.
- ▷ Among several forms of SPSS, **SPSS BASE** is the core program.
- ▷ In addition to the statistical analysis, data management and data documentation are the features of the base software.



## ▷ Descriptive Analysis

- Frequency Distribution, Descriptive Statistics, Cross-tabulation.

## ▷ Graphical Analysis (e.g. Bar diagram, Histogram and many more).

## ▷ Inferential Analysis

- Means, t-test, ANOVA, Correlation, Nonparametric tests etc.

## ▷ Prediction for numerical outcomes

- Linear regression, Logistic regression, GLMs etc.

## ▷ Prediction for identifying groups

- Factor analysis, Cluster analysis, Discriminant analysis etc.

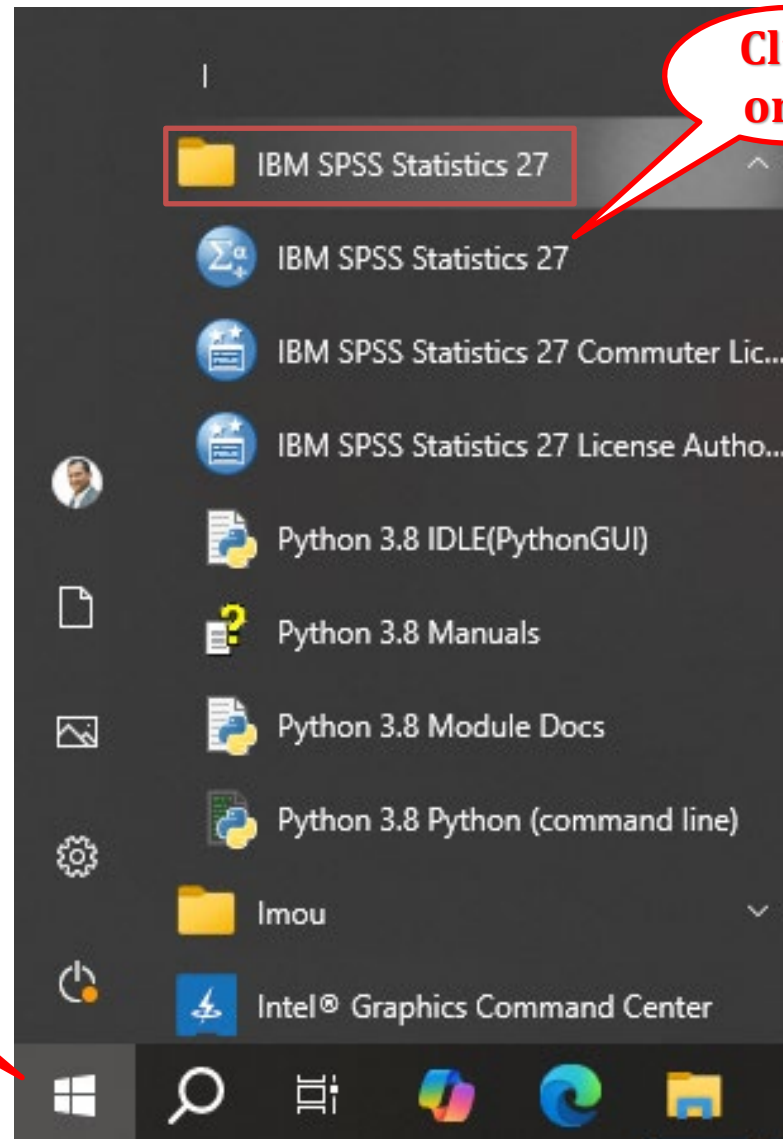
# Data Entry in SPSS



▷ At first open an empty SPSS file from the program.

Click on  → IBM SPSS Statistics → IBM SPSS Statistics 27

# Windows 10 and IBM SPSS 27 Version





# Open a New Dataset Window



Welcome to IBM SPSS Statistics

## IBM SPSS Statistics

**New Files:**

- New Dataset
- New Database Query...

**Restore Points:**

Restore Points Recent Files Sample Files

**What's New:**

```
MATRIX.  
COMPUTE X={0,1,2,3,4,5}.  
COMPUTE PDF_X=PDF.BINOM(X,5,.5).  
COMPUTE CDF_X=CDF.BINOM(X,5,.5).  
PRINT {X;PDF_X;CDF_X}.  
END MATRIX.
```

**Matrix**

Run MATRIX procedure:

{X;PDF_X;CDF_X}					
.000000000	1.000000000	2.000000000	3.000000000	4.000000000	5.000000000
.031250000	.156250000	.312500000	.312500000	.156250000	.031250000
.031250000	.187500000	.500000000	.812500000	.968750000	1.000000000

----- END MATRIX -----

Random variable generation and distribution function evaluation functionality within the MATRIX procedure now matches that available in COMPUTE.

**Help & Support** **Tutorials** **Community**

Visit [SPSS Statistics page](#) to explore available packages and offers.

[Give Feedback](#)  
[Report Issue](#)

☐ Don't show this dialog in the future

[Open](#)

[Close](#)



# Open a New Dataset Window

Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

1: Visible: 0 of 0 Variables

	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1																				
2																				
3																				
4																				
5																				
6																				
7																				
8																				
9																				
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				
19																				
20																				
21																				
22																				
23																				
24																				
25																				
26																				
27																				
28																				
29																				
30																				
31																				
32																				
33																				
34																				
35																				
36																				
37																				

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:OFF

## ▷ Data view

- The place to enter data
- Columns: variables
- Rows: records

## ▷ Variable view

- The place to enter variables
- List of all variables
- Characteristics of all variables

Untitled1 [DataSet0] - SPSS Statistics Data Editor

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

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										

Variable Name

Type of Variable

Details of Variable

Values for the categorical variable

Scale of Variable

Data View

Variable View

SPSS Statistics Processor is ready

12

# Blank Sheet



Untitled2 [DataSet3] - IBM SPSS Statistics Data Editor

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

Visible: 0 of 0 Variables

Columns: variables

Rows: cases

Under Data View

Data View Variable View

IBM SPSS Statistics Processor is ready

13

# Defining the Variables



- ▷ Enter the variable names in the “**Name**” columns in the order you want them.
- ▷ Traditionally variable names had to be 8 characters or less, start with a letter, and contain no embedded blanks.
- ▷ Now, they can be longer than 8 characters, but we will try to keep with names of 8 or fewer characters.
- ▷ Names can contain some special characters, but not all such characters. We only use underscore (\_), but hyphens (-), ampersands (&), and spaces (#) cannot be used.

# Creation of Data File in SPSS



- ▷ Suppose we have the following characteristics in a questionnaire:
- Identification Number
  - Age
  - Sex (Female, Male)
  - Religion (Muslim, Hindu, Buddhist, Christian)
  - Year of education (Illiterate, Primary, Secondary, Higher)
  - Place of residence (Urban, Rural)
  - Type of marriage preferred (Arranged, Love)
  - Height in cm
  - Weight in kg
  - Opinion on abortion (7 different questions)

# Defining the Variables



▷ So, we can choose the variable names as follows:

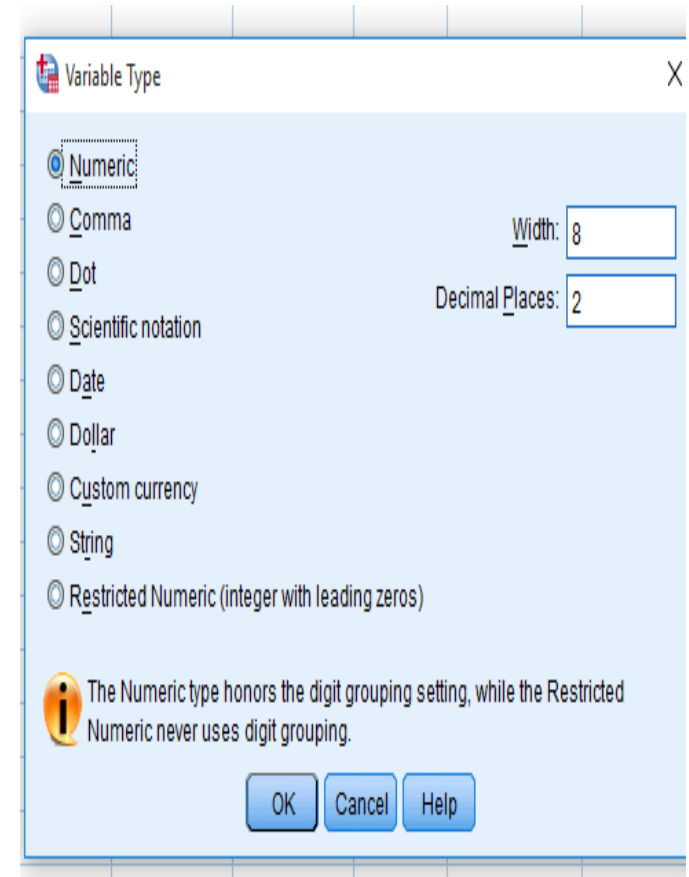
Actual Variable	Variable in SPSS	To Do
Identification Number	ID	ID
Age	AGE	Q1
Sex	GENDER	Q2
Religion	RELIGION	Q3
Year of education	EDUCATION	Q4
Place of residence	RESIDENCE	Q5



# Types of Variables



- ▷ The two basic types we use - **numeric** and **string**.
- ▷ Numeric variables only have numbers assigned.
- ▷ String variables may contain letters or numbers.
- ▷ Even if a string variable happens to contain only numbers, numeric operations on that variable will not be allowed.



# Types of Variables



Variable Type

☒ Numeric

☐ Comma

☐ Dot

☐ Scientific notation

☐ Date

☐ Dollar

☐ Custom currency

☐ String

☐ Restricted Numeric (integer with leading zeros)

Width: 8

Decimal Places: 2

The Numeric type honors the digit grouping setting, while the Restricted Numeric never uses digit grouping.

OK Cancel Help

## ▪ COMMA

Numeric variables that include commas that delimit every three places

*Example:* Thirty-thousand and one half: 30,000.50

## ▪ COMMA

*Example:* Thirty-thousand and one half: 30.000,50

## ▪ SCIENTIFIC NOTATION

Numeric variables whose values are displayed with an E and power-of-ten exponent, with or without a sign, or only with a sign (no E or D).

*Example:* 1.23E2, 1.23D2, 1.23E+2, 1.23+2.

## ▪ DATE

*Example:* Dates: 01/31/2013, 31.01.2013

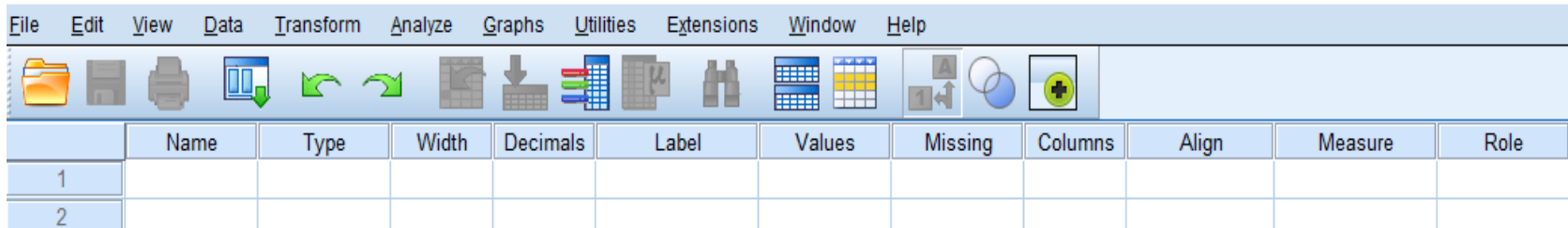
## ▪ DOLLAR

*Example:* Thirty-three thousand dollars and thirty-three cents: \$33,000.33

## ▪ RESTRICTED NUMBER

*Example:* 00000123456 (width 11)

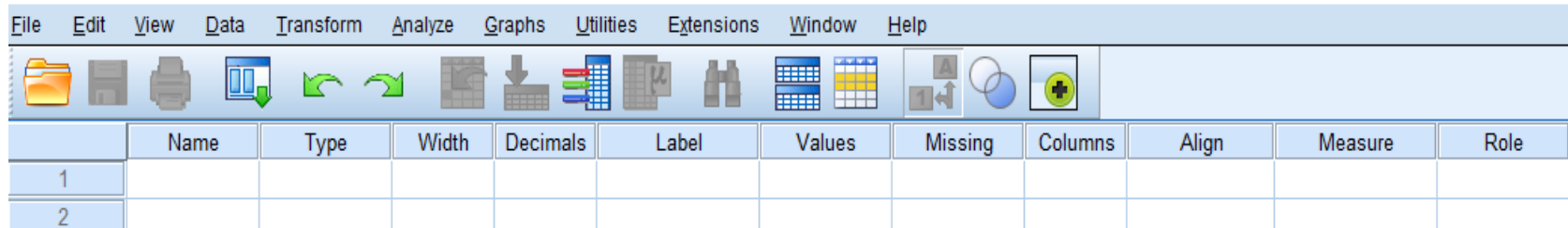
# Variables Attributes



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1											
2											

- ▷ **Width** - Default width of numeric variable entries is eight. It can increase or decrease.
- ▷ **Decimal** - For continuous data decimal is necessary. For discrete data decimal is not so important.
- ▷ **Label** - Enter the variable labels in the “Label” column. Here the variable should be detailed since it not only recognizes the variable but also appears in the result.
- ▷ **Values** - Enter the value labels in the “Values” column. For categorical variables, an integer code should be assigned to each category and variable should be numeric.

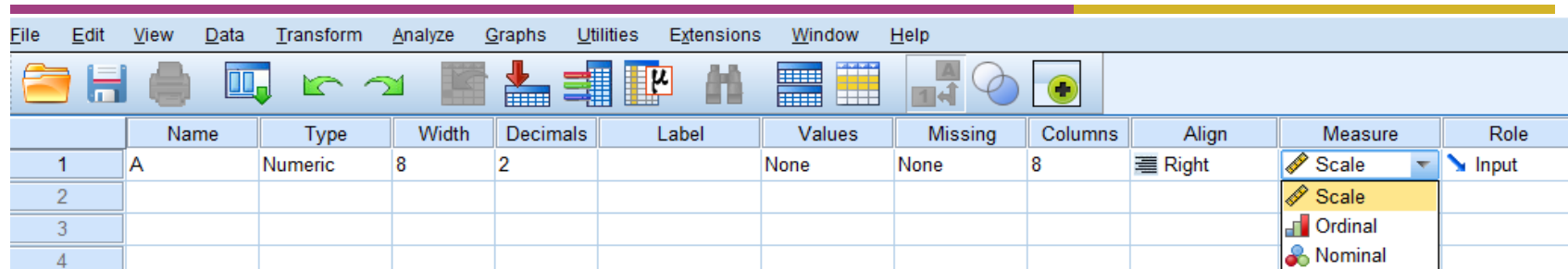
# Variables Attributes (cont'd)



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1											
2											

- ▶ **Missing** - Enter the missing values in the “Missing” column. SPSS recognizes the **period (.)** symbol as missing value. Other values like 9/99/999 are used by putting the **Missing Values dialogue box**.
- ▶ **Columns** - Width of the variable column in the **Data View**. The default cell width for numerical variables is eight. It can increase or decrease.
- ▶ **Align** - alignment of variable entries. The SPSS default is to align numerical variables to the right-hand side of a cell and string variables to the left.

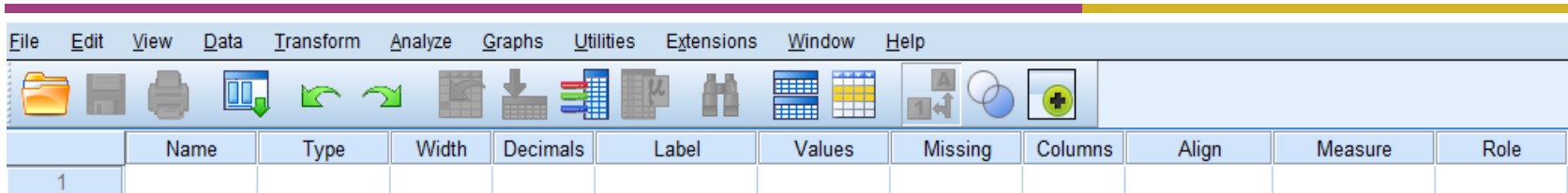
# Variables Attributes (cont'd)



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	A	Numeric	8	2		None	None	8	Right	Scale	Input
2										Scale	
3										Ordinal	
4										Nominal	

- ▷ **Measure** - measurement scale of the variable.
  - ▷ For “**numeric**” type variables the default measurement scale is a ratio or interval scale (referred to by SPSS as “**Scale**”).
  - ▷ For “**string**” type variables the default is a **nominal scale**.
  - ▷ The option “**Nominal**” is for categorical variables without ordered categories but is not used by default.
  - ▷ The third option “**Ordinal**” is for categorical variables with ordered categories but is not used by default.
- ▷ A summary of variable characteristics can be obtained from the **Utilities** drop-down menu.

# Variables Attributes (cont'd)



## ▷ Role – variable's role

- ▷ **Input:** The variable will be used as a predictor (independent variable). This is the default assignment for variables.
  - ▷ **Target:** The variable will be used as an outcome (dependent variable).
  - ▷ **Both:** The variable will be used as both a predictor and an outcome (independent and dependent variable).
  - ▷ **None:** The variable has no role assignment.
  - ▷ **Partition:** The variable will partition the data into separate samples.
  - ▷ **Split:** Used with the IBM® SPSS® Modeler (not IBM® SPSS® Statistics).
- ▷ **Lets go to SPSS window to create the data file.**

# Enter variables



Based on your code book!

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Variable View' tab is active, displaying a list of variables. The 'Age' variable is selected in the 'Values' column. A blue cloud with the text 'Based on your code book!' has arrows pointing to the 'Value Labels' dialog boxes. The top 'Value Labels' dialog is empty, while the bottom one shows '1 = Male' and '2 = Female' entered.

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
Q01	Numeric	8	0	Age	None	None	8	Right	Ordinal	Input

Value Labels

Value:

Label:

Add

Change

Spelling...

Value Labels

Value: 1

Label: Male

1 = "Male"

2 = "Female"

Add

Change

Remove

Spelling...

OK Cancel Help

# The Final SPSS DATA File (Variable View)



\*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

— □ ×

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



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Id	Numeric	8	2	Identification number	None	None	8	Right	Scale	Input
2	Q1	Numeric	8	2	age	None	None	8	Right	Scale	Input
3	Q2	Numeric	8	2	gender	{1.00, male}...	None	8	Right	Scale	Input
4	Q3	Numeric	8	2	religion	{1.00, musli...	None	8	Right	Scale	Input
5	Q4	Numeric	8	2	year of education	None	None	8	Right	Scale	Input
6	Q5	Numeric	8	2	residence	{1.00, rural}...	None	8	Right	Scale	Input
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											

Data View Variable View

IBM SPSS Statistics Processor is ready



# Data Entry in Data View Spreadsheet



- ▷ When variables are ready in **Variable View**, data values can be entered in the **Data View** spreadsheet. SPSS assumes that rows represent cases and columns represent variables.
- ▷ The appearance of the **Data View** spreadsheet is controlled by the **View** drop-down menu.
- ▷ When labels have been assigned to the category codes of a categorical variable, these can be displayed by checking **Value Labels**.

# Enter Cases (Data View)



\*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

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

Visible: 6 of 6 Variables

	Id	Q1	Q2	Q3	Q4	Q5	var
1	1.00	45.00	male	muslim	16.00	rural	
2	2.00	35.00	female	hindu	12.00	rural	
3	3.00	33.00	female	others	14.00	urban	
4	4.00	23.00	male	hindu	12.00	rural	
5	5.00	18.00	male	muslim	15.00	urban	
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

Data View Variable View

IBM SPSS Statistics Processor is ready

# Practice Problem



- ▷ Open the **Exercise.sav** data file.
- ▷ Practice defining the correct attributes to each variable by following the code book below.

Name	Label	Value Label	Missing Values	Measure
IDnum				Scale
sex	Respondent's Sex	1 = Male 2 = Female		Nominal
race	Race of Respondent	1 = White 2 = Black 3 = Other		Nominal
region	Region of the United States	1 = North East 2 = South East 3 = West		Nominal
happy	General Happiness	0 = NAP 1 = Very Happy 2 = Pretty Happy 3 = Not too Happy 8 = DK 9 = NA	0, 8, 9	Ordinal

# Practice Problem



## ▷ Code book:

life	Is Life Exciting or Dull	0 = NAP	0, 8, 9	Ordinal
		1 = Exciting		
		2 = Routine		
		3 = Dull		
		8 = DK		
		9 = NA		
sibs	Number of Brothers and Sisters	98 = DK	98, 99	Scale
		99 = NA		
childs	Number of Children	8 = Eight or More	9	Scale
		9 = NA		
age	Age of Respondent	98 = DK	0, 98, 99	Scale
		99 = NA		
educ	Highest Year of School Completed	97 = NAP	97, 98, 99	Scale
		98 = DK		
		99 = NA		

# Practice Problem



## ▷ Code book:

paeduc	Highest Year School, Father	97 = NAP	97, 98, 99	Scale
		98 = DK		
		99 = NA		
maeduc	Highest Year School, Mother	97 = NAP	97, 98, 99	Scale
		98 = DK		
		99 = NA		
seeduc	Highest Year School, Spouse	97 = NAP	97, 98, 99	Scale
		98 = DK		
		99 = NA		
prestg80	Occupational Prestige Score	0 = DK,NA,NAP	0	Scale
occcat80	Occupational Category	1 = Managerial and Professional		Nominal
		2 = Technical and Sales		
		3 = Service		
		4 = Farming, Forest, and Fishing		
		5 = Production and Craft		
		6 = General Labor		

# Types of SPSS file



## ▷ Three parts of SPSS

- ▷ Data Editor Enter and View Data Values, files end in .sav or .por
- ▷ Output Viewer View Statistical Results, files end in .spv
- ▷ Syntax Window Write & Run Syntax Command, files end in .sps
- ▷ For example -

Data file → **.sav** (*data.sav*)

Output file → **.spo** (**data.spo**)

Syntax file → **.sps** (**data.sps**)

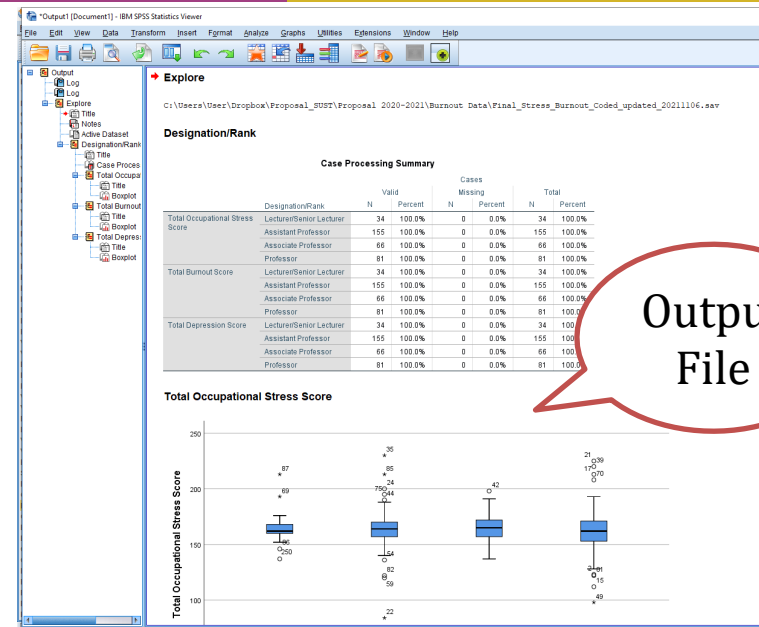
# Parts of SPSS



Final\_Stress\_Burnout\_Coded\_updated\_20211106.sav [1] - IBM SPSS Statistics Data Editor

ID	Gender	University	Place	Birth	Age	Religion	Education	Designation	Experience
1	Male	SUST	Rural	40	Muslim	PhD	Associate Profes...		
2	Male	SUST	Urban	43	Muslim	PhD	Professor		
3	Male	SUST	Rural	42	Muslim	Master	Associate Profes...		
4	Male	SUST	Rural	35	Hindu	Master	Associate Profes...		
5	Female	SUST	Urban	27	Muslim	Master	Associate Profes...		
6	Male	SUST	Urban	38	Hindu	Master	Associate Profes...		
7	Male	SUST	Rural	35	Muslim	Master	Associate Profes...		
8	Male	SUST	Urban	51	Hindu	PhD	Professor		
9	Male	SUST	Rural	32	Muslim	M.Phil	Assistant Profes...		
10	Male	SUST	Rural	28	Hindu	Master	Associate Profes...		
11	Male	SUST	Urban	43	Muslim	PhD	Professor		
12	Male	SUST	Urban	45	Muslim	PhD	Professor		
13	Male	SUST	Urban	34	Hindu	Bachelor	Assistant Profes...		
14	Female	SUST	Urban	36	Muslim	Master	Associate Profes...		
15	Male	SUST	Rural	42	Muslim	PhD	Professor		
16	Male	SUST	Rural	29	Hindu	Bachelor	Assistant Profes...		
17	Male	SUST	Urban	54	Muslim	PhD	Professor		
18	Male	SUST	Urban	39	Muslim	M.Phil	Associate Profes...		
19	Male	SUST	Rural	42	Muslim	Master	Associate Profes...		
20	Male	SUST	Rural	43	Hindu	Master	Associate Profes...		
21	Male	SUST	Urban	42	Muslim	PhD	Professor		
22	Female	SUST	Rural	39	Muslim	PhD	Associate Profes...		
23	Male	SUST	Rural	44	Muslim	PhD	Professor		
24	Male	SUST	Urban	42	Muslim	Master	Associate Profes...		
25	Male	SUST	Rural	41	Muslim	PhD	Professor		
26	Male	SUST	Rural	34	Hindu	PhD	Associate Profes...		
27	Male	SUST	Rural	39	Muslim	PhD	Associate Profes...		
28	Male	SUST	Urban	43	Muslim	PhD	Professor		
29	Female	SUST	Rural	39	Muslim	PhD	Associate Profes...		
30	Male	SUST	Urban	42	Muslim	PhD	Professor		
31	Male	SUST	Rural	40	Muslim	M.Phil	Associate Profes...		
32	Female	SUST	Rural	40	Muslim	PhD	Professor		
33	Male	SUST	Urban	42	Muslim	PhD	Professor		
34	Male	SUST	Urban	32	Muslim	Bachelor	Assistant Profes...		
35	Male	SUST	Urban	38	Muslim	PhD	Associate Profes...		
36	Male	SUST	Rural	36	Muslim	Master	Associate Profes...		

Data Editor



Output File

Burnout\_20210625.sps - IBM SPSS Statistics Syntax Editor

```

Encoding: windows-1252
1. GET DATA
2.   /FILE=C:\Users\User\Dropbox\Proposal_SUST\Proposal 2020-2021\Burnout Data\
3.   /SHEET='stress burnout'
4.   /CELLRANGE=FULL
5.   /READNAMES=ON
6.   /DATATYPEMIN PERCENTAGE=95.0
7.   /HIDDEN IGNORE=YES.
8. EXECUTE.
9.
10. GET DATA
11.   /FILE=C:\Users\User\Dropbox\Proposal_SUST\Proposal 2020-2021\Burnout Data\Stress_Burnout_PHQ_20210625.xlsx
12.   /SHEET='simburn'
13.   /CELLRANGE=FULL
14.   /READNAMES=ON
15.   /DATATYPEMIN PERCENTAGE=95.0
16.   /HIDDEN IGNORE=YES.
17. EXECUTE.
18.
19. SAVE OUTFILE = "Stress_Burnout_PHQ_20210625.sav".
20. GET FILE = "Stress_Burnout_PHQ_20210625.sav".
21.
22. GET DATA
23.   /FILE=C:\Users\User\Dropbox\Proposal_SUST\Proposal 2020-2021\Burnout Data\simburn_updated.xlsx
24.   /SHEET='simburn'
25.   /CELLRANGE=FULL
26.   /READNAMES=ON
27.   /DATATYPEMIN PERCENTAGE=95.0
28.   /HIDDEN IGNORE=YES.
29. EXECUTE.
30.
31. SAVE OUTFILE = "simburn_updated.sav".
32. GET FILE = "simburn_updated.sav".
33.
34. GET FILE = "Stress_Burnout_PHQ_20210625.sav".
35.
36. ADD FILES /FILE=*
37.   /FILE=C:\Users\User\Dropbox\Proposal_SUST\Proposal 2020-2021\Burnout Data\simburn_updated.sav.
38. EXECUTE.
39.
40. SAVE OUTFILE = "Final_Stress_Burnout_updated.sav".
41. EXECUTE.
42.
43. GET FILE = "Final_Stress_Burnout_updated.sav".
44.
45. VALUE LABELS
46.   /GENDER
47.   1 'Male'
48.   0 'Female'
49. EXECUTE.
50.
51. COMPUTE

```

Syntax File

# Import Data File

---

- ▷ Import data from an SPSS Existing File (\*.sav)
- ▷ Import data from a spreadsheet (\*.xls / \*.xlsx)
- ▷ Import data from an ASCII text file (\*.txt, \*.dat)
- ▷ Import data from a database file (dBase, \*.dbf)
  
- ▷ An useful guideline is available here

<http://libguides.library.kent.edu/SPSS/ImportData>



# Import data from SPSS/SAS/STATA



## ▷ To open an SPSS data file

### ▷ Manually :

File → Open → Data →  
Browse file from source

### ▷ Syntax:

GET FILE = 'Drive:\source----\file name'.

### Example:

#### ▷ SPSS:

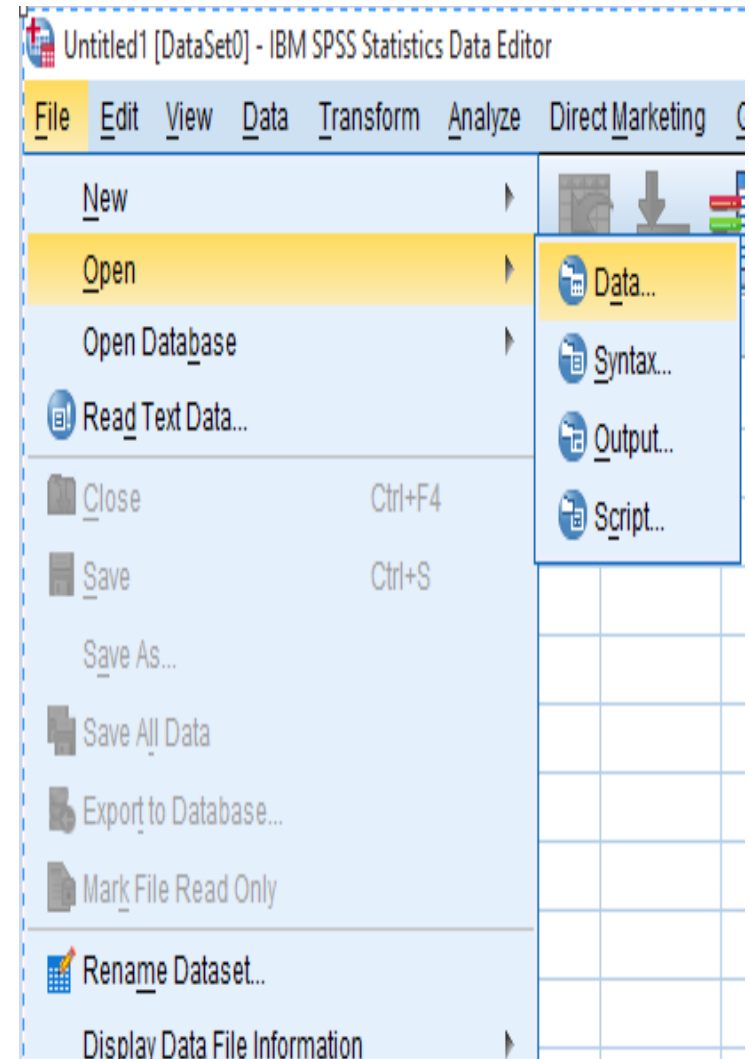
**GET FILE**='C:\SPSS  
Training\DayI\Obesity.sav'.

#### ▷ SAS:

**GET SAS FILE**="C:\SPSS Training\Day I\  
Obesity.sas7bdat".

#### ▷ STATA:

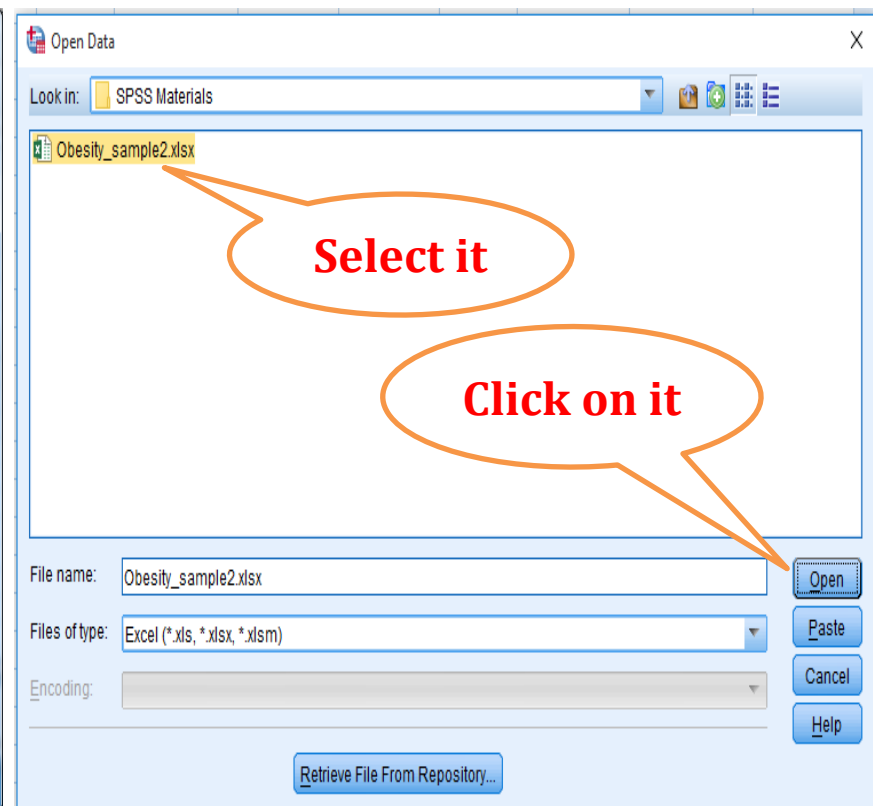
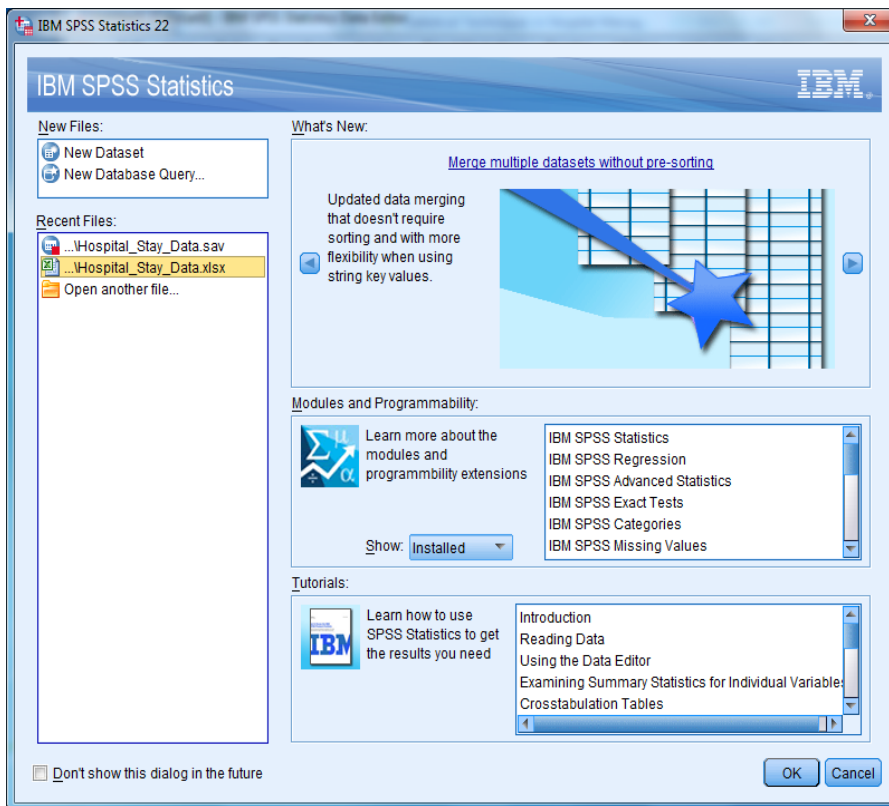
**GET STATA FILE**="C:\SPSS Training\Day I\  
Obesity.dta".



# Import data from Excel



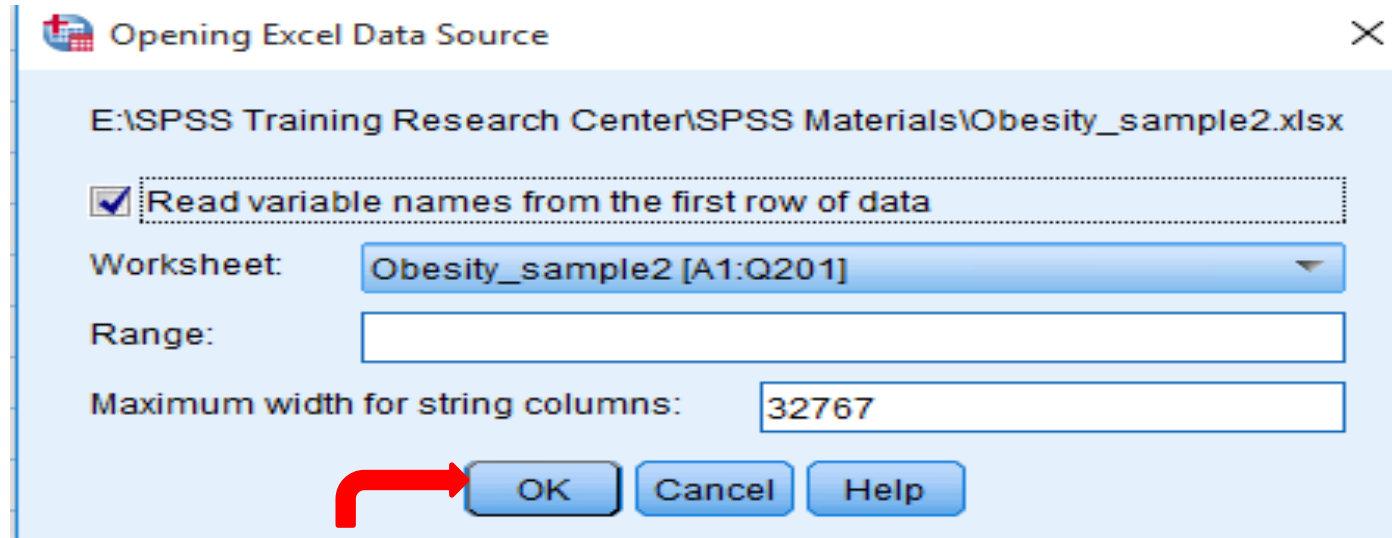
- ▷ Select File → Open → Data
- ▷ Choose **Excel** as file type
- ▷ Select the file you want to import
- ▷ Then click Open



# Open Excel files in SPSS

## ▷ Manually

File → Open → Data → Browse Excel file → OK



## ▷ Syntax

```
GET DATA /TYPE=XLSX
```

```
  /FILE='C:\SPSS Training\DayI\Obesity_sample2.xlsx'
```

```
  /SHEET=name Obesity_sample2'
```

```
  /CELLRANGE=full
```

```
  /READNAMES=on
```

```
  /ASSUMEDSTRWIDTH=32767.
```

```
EXECUTE.
```

# Open Excel files in SPSS (Cont'd)



\*Untitled3 [DataSet3] - IBM SPSS Statistics Data Editor

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

Visible: 17 of 17 Variables

	ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIST	ARM	F.AGE	M.AGE	F.OCC	M.OCC	F.EDU	M.EDU	INCOME	EXF
1	1	3	9.0	0	135.0	25.10	53.0	19.0	43	41	1	0	3	4	65000	
2	2	1	5.0	1	106.0	15.50	50.0	16.0	41	26	2	0	2	1	30000	
3	3	2	8.0	0	137.0	30.60	67.0	20.0	40	30	3	0	4	3	40000	
4	4	4	10.0	1	121.0	22.00	61.0	17.0	40	36	2	0	4	2	50000	
5	5	1	5.0	1	102.0	14.20	47.0	15.0	43	41	1	0	3	4	65000	
6	6	3	9.0	1	145.0	41.30	79.0	23.0	42	36	1	1	3	3	100000	
7	7	2	9.0	1	121.0	24.00	57.0	19.0	40	30	2	0	3	4	80000	
8	8	3	8.0	0	141.0	36.30	67.0	21.0	45	30	1	0	4	1	120000	
9	9	1	5.0	1	106.0	15.30	55.0	16.0	45	35	1	0	3	4	40000	
10	10	3	10.0	1	145.0	38.40	76.0	23.0	45	36	1	0	1	1	50000	
11	11	2	9.0	1	135.0	26.20	65.0	17.0	45	36	1	0	1	1	30000	
12	12	5	11.0	0	156.0	39.70	66.0	24.0	45	36	1	0	1	1	30000	
13	13	3	8.0	0	120.0	24.40	57.0	19.0	45	36	1	0	2	2	50000	
14	14	3	8.0	1	131.0	26.40	58.0	17.0	45	36	1	0	3	3	40000	
15	15	1	6.0	0	102.0	12.70	49.0	13.5	35	25	4	0	2	0	70000	
16	16	2	7.0	0	117.0	17.00	51.4	15.0	35	25	4	0	2	0	70000	
17	17	2	8.0	1	126.5	31.20	68.0	21.5	49	35	2	0	4	3	150000	
18	18	1	7.0	1	125.0	38.00	71.5	23.0	52	36	2	0	2	1	15000	
19	19	5	10.0	0	148.0	41.40	79.0	23.0	45	35	2	0	3	1	50000	
20	20	5	11.0	0	150.0	50.70	98.0	27.5	60	36	3	0	3	2	20000	
21	21	5	11.0	1	136.0	22.00	42.0	15.0	47	37	10	0	3	3	30000	
22	22	1	6.0	1	117.0	19.40	58.0	18.5	36	34	8	8	4	2	40000	
23	23	3	8.0	1	122.0	20.00	58.0	17.0	41	33	2	0	3	3	40000	

Save this file as SPSS data

Data View Variable View

Ref: Sample data from childhood obesity project (2015)

# Import data from CSV file

▷ To open an CSV Data File (.txt, .tab, .csv, .dat)

▷ Manually

File → Open → Data → Browse file → OK.

Please follow the steps.....

## Step 1

Graphs Utilities Add-ons Window Help

Text Import Wizard - Step 1 of 6

Welcome to the text import wizard!  
This wizard will help you read data from your text file and specify information about the variables.

Does your text file match a predefined format?

☐ Yes ☒ No [Browse...](#)

Text file: E:\SPSS Training Research Center\SPSS Materials\Obesity\_sample1.csv

	var1	var2	var3	var4
1				
2				
3				
4				

0 10 20 30 40 50 60

	ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIST	ARM	F. AGE	M. AGE	F. OCC	M. OCC
1	1	3	9	0	135	25.1	53	19	43	41	1	0
2	1	3	9	0	135	25.1	53	19	43	41	1	0
3	2	1	5	1	106	15.5	50	16	41	26	2	0
4	3	2	8	0	137	30.6	67	20	40	30	3	0
5	4	4	10	1	121	22.6	61	17	40	36	2	0
6	5	1	5	1	102	14.2	47	15	43	41	1	0
7	6	3	9	1	145	41.3	79	23	42	36	1	1
8	7	2	9	1	121	24	57	19	40	29	2	8

< Back Next > Finish Cancel Help

# Arrange variables (Step 2)



Text Import Wizard - Step 2 of 6

How are your variables arranged?

☒ Delimited - Variables are delimited by a specific character (i.e., comma, tab).  
☐ Fixed width - Variables are aligned in fixed width columns.

Are variable names included at the top of your file?

☒ Yes  
☐ No

Text file: E:\SPSS Training Research Center\SPSS Materials\Obesity\_sample1.csv

	0	10	20	30	40	50	60											
1	1	3	9	0	135	25.1	53	19	43	41	1	0	3	4	65000	50000	5	
2	2	1	5	1	106	15.5	50	16	41	26	2	0	2	1	30000	25000	4	
3	3	2	8	0	137	30.6	67	20	40	30	3	0	4	3	40000	22000	5	
4	4	4	10	1	121	22	61	17	40	36	2	0	4	2	50000	45000	4	
5	5	1	5	1	102	14.2	47	15	43	41	1	0	3	4	65000	50000	5	
6	6	2	0	1	145	41	2	70	22	42	26	1	1	2	3	100000	50000	5

< Back Next > Finish Cancel Help

# Checking Line number (Step 3)



Text Import Wizard - Delimited Step 3 of 6

The first case of data begins on which line number?

How are your cases represented?

☒ Each line represents a case

☐ A specific number of variables represents a case:

How many cases do you want to import?

☒ All of the cases

☐ The first  cases.

☐ A random percentage of the cases (approximate):  %

Data preview

	0	10	20	30	40	50	60
1	1	3	9	0	135	25.1	53
2	2	1	5	1	106	15.5	50
3	3	2	8	0	137	30.6	67
4	4	4	10	1	121	22	61

< Back Next > Finish Cancel Help

# Choosing delimiters (Step 4)



Text Import Wizard - Delimited Step 4 of 6

Which delimiters appear between variables?

☐ Tab ☒ Space

☒ Comma ☐ Semicolon

☐ Other:

What is the text qualifier?

☒ None

☐ Single quote

☐ Double quote

☐ Other:

Data preview

ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIS
1	3	9	0	135	25.1	53
2	1	5	1	106	15.5	50
3	2	8	0	137	30.6	67
4	4	10	1	121	22	61
5	1	5	1	102	14.2	47
6	3	9	1	145	41.3	79
7	2	9	1	121	24	57
8	3	8	0	141	36.3	67
9	1	5	1	106	15.3	55
10	3	10	1	145	38.4	76

< Back Next > Finish Cancel Help



# Specification for Variables (Step 5)



Text Import Wizard - Step 5 of 6

Specifications for variable(s) selected in the data preview

**i** Data format is determined from the values present in the first 200 records. If a column contains multiple data types in the first 200 records, the variable type is set to string. The length (number of characters) for string variables is determined by the longest value present in the first 200 records. If subsequent records have longer values, they will be truncated.

Variable name:  Original Name: ID

Data format:

Data preview

ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIS
1	3	9	0	135	25.1	53
2	1	5	1	106	15.5	50
3	2	8	0	137	30.6	67
4	4	10	1	121	22	61

< Back Next > Finish Cancel Help

# Step 6



Graphs Utilities Add-ons Window Help

Text Import Wizard - Step 6 of 6

You have successfully defined the format of your text file.

Would you like to save this file format for future use?

☐ Yes ☐ No [Save As...](#)

Would you like to paste the syntax?

☐ Yes ☒ No ☒ [Cache data locally](#)

Press the Finish button to complete the text import wizard.

Data preview

ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIS
1	3	9	0	135	25.1	53
2	1	5	1	106	15.5	50
3	2	8	0	87	30.6	67
4	4	10	1		22	61
5	1	5	1		14.2	47
6	3	9	1		41.3	79
7	2	9	1		24	57
8	3	8	0		36.3	67
9	1	5	1		15.3	55
10	3	10	1		38.4	76

**Click on it**

< Back Next > Finish Cancel Help

# Final Data file



Untitled3 [DataSet3] - IBM SPSS Statistics Data Editor

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

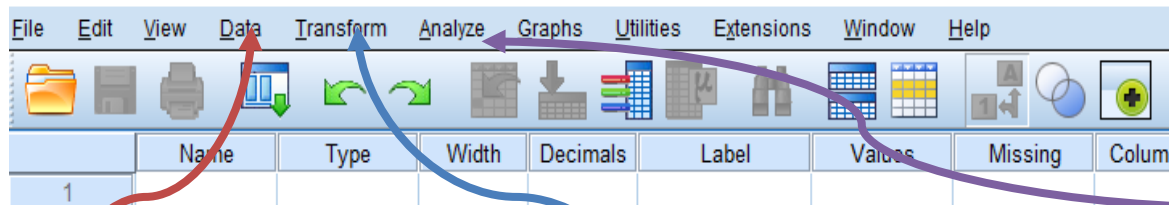
Visible: 17 of 17 Variables

	ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIST	ARM	F.AGE	M.AGE	F.OCC	M.OCC	F.EDU	M.EDU	INCOME	EXP
1	1	3	9.0	0	135.0	25.10	53.0	19.0	43	41	1	0	3	4	65000	
2	2	1	5.0	1	106.0	15.50	50.0	16.0	41	26	2	0	2	1	30000	
3	3	2	8.0	0	137.0	30.60	67.0	20.0	40	30	3	0	4	3	40000	
4	4	4	10.0	1	121.0	22.00	61.0	17.0	40	36	2	0	4	2	50000	
5	5	1	5.0	1	102.0	14.20	47.0	15.0	43	41	1	0	3	4	65000	
6	6	3	9.0	1	145.0	41.30	79.0	23.0	42	36	1	0	3	4	100000	
7	7	2	9.0	1	121.0	24.00	57.0	19.0	40				2	4	80000	
8	8	3	8.0	0	141.0	36.30	67.0	21.0	45							
9	9	1	5.0	1	106.0	15.30	55.0	16.0					1		40000	
10	10	3	10.0	1	145.0	38.40	76.0	23.0	4						50000	
11	11	2	9.0	1	135.0	26.20	65.0	17.0					1		30000	
12	12	5	11.0	0	156.0	39.70	66.0	24.0							30000	
13	13	3	8.0	0	120.0	24.40	57.0	19.0					2	2	50000	
14	14	3	8.0	1	131.0	26.40	58.0	17.0					4	3	40000	
15	15	1	6.0	0	102.0	12.70	49.0	13.5	35		4	0	2	0	70000	
16	16	2	7.0	0	117.0	17.00	51.4	15.0	35	25	4	0	2	0	70000	
17	17	2	8.0	1	126.5	31.20	68.0	21.5	49	35	2	0	4	3	150000	
18	18	1	7.0	1	125.0	38.00	71.5	23.0	52	36	2	0	2	1	15000	
19	19	5	10.0	0	148.0	41.40	79.0	23.0	45	35	2	0	3	1	50000	
20	20	5	11.0	0	150.0	50.70	98.0	27.5	60	36	3	0	3	2	20000	
21	21	5	11.0	1	136.0	22.00	42.0	15.0	47	37	10	0	3	3	30000	
22	22	1	6.0	1	117.0	19.40	58.0	18.5	36	34	8	8	4	2	40000	
23	23	3	8.0	1	122.0	20.00	58.0	17.0	41	33	2	0	3	3	40000	

Save this file as SPSS data

Data View Variable View

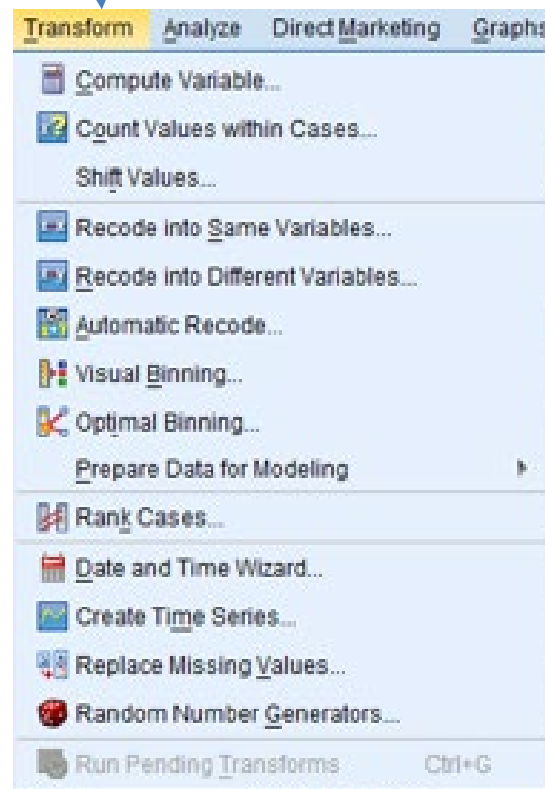
# Menu Bar



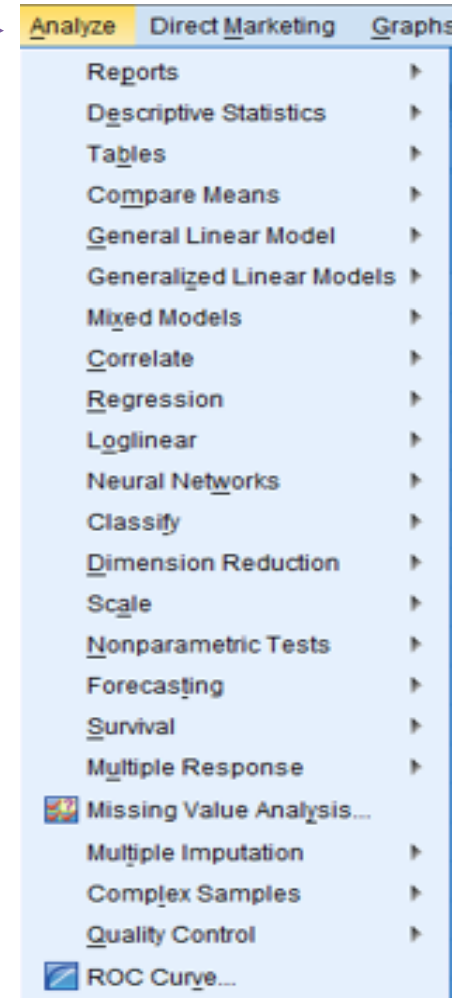
**“Data” Menu Bar**



**“Transform” Menu Bar**



**“Analyze” Menu Bar**



# Clean data after import data file



- ▷ Run **cases summaries** for all variables.
- ▷ Run **frequency** for qualitative variables and **Descriptive** for quantitative variables.
- ▷ Check outputs to see if you have variables with wrong values.
- ▷ Check missing values and physical surveys if you use paper surveys, and make sure they are real missing.
- ▷ Sometimes, you need to recode string variables into numeric variables.

# Cases summaries



IBM SPSS Statistics Data Editor window showing the 'Reports' menu and the 'Summarize Cases' dialog box.

The 'Reports' menu is open, showing options like 'Codebook...', 'OLAP Cubes...', 'Case Summaries...', 'Report Summaries in Rows...', and 'Report Summaries in Columns...'. The 'Case Summaries...' option is highlighted.

The 'Summarize Cases' dialog box is open, showing the 'Variables' list (ID, CLASS, AGE, GENDER, HEIGHT, WEIGHT) and the 'Grouping Variable(s)' field. The 'Display cases' checkbox is checked, and the 'Limit cases to first' value is set to 100. The 'OK' button is highlighted with a red circle and the text 'Click on it'.

The background data table shows variables: EIGHT, WAIST, ARM, F.AGE, M.AGE, F.OCC.

	EIGHT	WAIST	ARM	F.AGE	M.AGE	F.OCC
1	25.10	53.0	19.0	43	41	
2	15.50	50.0	16.0	41	26	

Buttons at the bottom: Data View, Variable View, OK, Paste, Reset, Cancel, Help.

# Cases summaries



**Case Processing Summary<sup>a</sup>**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
ID	100	100.0%	0	0.0%	100	100.0%
CLASS	100	100.0%	0	0.0%	100	100.0%
AGE	100	100.0%	0	0.0%	100	100.0%
GENDER	100	100.0%	0	0.0%	100	100.0%
HEIGHT	100	100.0%	0	0.0%	100	100.0%
WEIGHT	100	100.0%	0	0.0%	100	100.0%
WAIST	100	100.0%	0	0.0%	100	100.0%
ARM	100	100.0%	0	0.0%	100	100.0%
F.AGE	100	100.0%	0	0.0%	100	100.0%
M.AGE	100	100.0%	0	0.0%	100	100.0%
F.OCC	100	100.0%	0	0.0%	100	100.0%
M.OCC	100	100.0%	0	0.0%	100	100.0%
F.EDU	100	100.0%	0	0.0%	100	100.0%
M.EDU	100	100.0%	0	0.0%	100	100.0%
INCOME	100	100.0%	0	0.0%	100	100.0%
EXPENSE	100	100.0%	0	0.0%	100	100.0%
FAM.SIZE	100	100.0%	0	0.0%	100	100.0%

a. Limited to first 100 cases.

# Cases summaries



Case Summaries<sup>a</sup>

	ID	CLASS	AGE	GENDER	HEIGHT	WEIGHT	WAIST	ARM	F.AGE	M.AGE	F.OCC	M.OCC	F.EDU	M.EDU	INCOME	EXPENSE	FAM.SIZE
1	1	3	9.0	0	135.0	25.10	53.0	19.0	43	41	1	0	3	4	65000	50000	5
2	2	1	5.0	1	106.0	15.50	50.0	16.0	41	26	2	0	2	1	30000	25000	4
3	3	2	8.0	0	137.0	30.60	67.0	20.0	40	30	3	0	4	3	40000	22000	5
4	4	4	10.0	1	121.0	22.00	61.0	17.0	40	36	2	0	4	2	50000	45000	4
5	5	1	5.0	1	102.0	14.20	47.0	15.0	43	41	1	0	3	4	65000	50000	5
6	6	3	9.0	1	145.0	41.30	79.0	23.0	42	36	1	1	3	3	100000	50000	5
7	7	2	9.0	1	121.0	24.00	57.0	19.0	40	29	2	8	3	4	80000	70000	6
8	8	3	8.0	0	141.0	36.30	67.0	21.0	45	38	1	2	3	4	120000	70000	5
9	9	1	5.0	1	106.0	15.30	55.0	16.0	55	40	2	0	1	1	40000	35000	6
10	10	3	10.0	1	145.0	38.40	76.0	23.0	45	38	2	0	1	1	50000	45000	4
11	11	2	9.0	1	135.0	26.20	65.0	17.0	45	30	4	0	1	1	30000	15000	4
12	12	5	11.0	0	156.0	39.70	66.0	24.0	45	30	4	0	1	1	30000	15000	4
13	13	3	8.0	0	120.0	24.40	57.0	19.0	38	30	4	0	2	2	50000	30000	8
14	14	3	8.0	1	131.0	26.40	58.0	17.0	39	26	10	0	4	3	40000	35000	5
15	15	1	6.0	0	102.0	12.70	49.0	13.5	35	25	4	0	2	0	70000	10000	4
16	16	2	7.0	0	117.0	17.00	51.4	15.0	35	25	4	0	2	0	70000	10000	4
17	17	2	8.0	1	126.5	31.20	68.0	21.5	49	35	2	0	4	3	150000	100000	4
18	18	1	7.0	1	125.0	38.00	71.5	23.0	52	36	2	0	2	1	15000	15000	5
19	19	5	10.0	0	148.0	41.40	79.0	23.0	45	35	2	0	3	1	50000	40000	4
20	20	5	11.0	0	150.0	50.70	98.0	27.5	60	36	3	0	3	2	20000	15000	4
21	21	5	11.0	1	136.0	22.00	42.0	15.0	47	37	10	0	3	3	30000	50000	4
22	22	1	6.0	1	117.0	19.40	58.0	18.5	36	34	8	8	4	2	40000	35000	7
23	23	3	8.0	1	122.0	20.00	58.0	17.0	41	33	2	0	3	3	40000	25000	9
24	24	3	10.0	1	141.0	40.50	70.0	26.0	41	33	2	0	2	1	2700000	500000	22
25	25	4	10.0	0	133.0	25.00	62.0	17.5	46	36	2	0	0	0	300000	80000	22
26	26	2	8.0	0	115.5	18.60	54.0	17.0	41	33	2	0	2	1	2700000	500000	22
27	27	3	9.0	0	136.0	36.90	77.0	25.0	41	33	2	0	2	1	2700000	500000	22
28	28	1	8.0	1	125.0	25.80	55.0	21.0	45	38	10	0	4	2	30000	20000	5
29	29	4	8.5	0	128.5	19.50	52.0	14.3	43	33	9	0	4	4	70000	60000	7
30	30	1	5.5	0	121.7	19.20	54.0	15.5	43	33	9	0	4	4	70000	60000	7
31	31	5	11.0	1	137.0	26.20	62.5	16.0	46	40	2	0	3	0	20000	20000	10
32	32	5	12.0	1	143.5	25.00	60.5	16.5	45	30	10	0	4	2	30000	25000	10
33	33	2	8.0	0	125.5	21.00	54.5	18.0	45	30	10	0	4	2	30000	25000	10
34	34	1	7.0	0	118.0	20.00	56.0	15.5	40	32	10	0	4	3	55000	50000	5
35	35	3	9.0	1	138.0	36.00	68.0	23.0	40	30	10	0	3	1	30000	22000	5
36	36	2	8.0	0	121.0	23.40	60.0	18.5	50	30	2	5	2	4	70000	55000	7
37	37	1	6.0	0	112.0	17.00	54.0	16.0	33	28	2	0	1	2	150000	100000	6
38	38	4	10.0	0	125.0	21.00	53.0	16.0	40	35	8	0	2	2	40000	35000	5
39	39	2	8.0	0	121.0	25.80	55.5	20.0	66	44	2	0	2	2	100000	70000	8
40	40	3	11.0	0	144.0	50.00	79.0	26.0	45	33	9	0	3	2	60000	40000	6
41	41	5	12.0	1	154.0	38.30	72.0	20.0	34	30	4	0	0	1	100000	50000	4
42	42	4	9.0	0	137.0	30.60	67.0	20.0	56	45	2	0	2	1	100000	35000	7
100	100	4	9.0	0	142.0	29.00	69.0	16.5	50	40	4	0	2	2	60000	20000	4
Total	N	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	Minimum	1	1	5.0	0	96.0	11.50	42.0	13.5	28	1	0	0	0	11000	9000	3
	Maximum	100	5	14.0	1	156.0	50.70	98.0	27.5	70	50	10	4	4	2700000	500000	22

a. Limited to first 100 cases.



# Descriptive Statistics



- ▷ Find wrong entries.
- ▷ Have basic knowledge about the sample and targeted variables in a study.
- ▷ Summarize data.

# Qualitative/Categorical Variables

Analyze → Descriptive statistics → Frequency

Visible: 17 of 17 Variables

ID	CL	HEIGHT	WEIGHT	WAIST	ARM	F.AGE	M.AGE	F.OCC
1	1	135.0	25.10	53.0	19.0	43	41	
2	2	106.0	15.50	50.0	16.0	41	26	
3	3	137.0	30.60	67.0	20.0	40	30	
4	4	121.0	22.00	61.0	17.0	40	26	
5	5	10						
6	6	14						
7	7	12						
8	8	14						
9	9	10						
10	10	14						
11	11	13						
12	12	15						
13	13	12						
14	14	13						
15	15	10						
16	16	11						
17	17	12						
18	18	12						
19	19	14						
20	20	12						
21	21	11						
22	22	12						
23	23	12						

Variable(s):

- CLASS
- GENDER
- F.OCC
- M.OCC
- F.EDU
- M.EDU
- FAM.SIZE

☒ Display frequency tables

OK Paste Reset Cancel Help

Click on it

# Qualitative/Categorical variables



**GENDER**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid F	92	46.0	46.0	46.0
M	108	54.0	54.0	100.0
Total	200	100.0	100.0	

**FATHER'S OCCUPATION**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	1	.5	.5	.5
DOCTOR	9	4.5	4.5	5.0
BUSINESS	89	44.5	44.5	49.5
PHARMACIST	4	2.0	2.0	51.5
IMIGRANT	31	15.5	15.5	67.0
BANKER	6	3.0	3.0	70.0
LAWER	3	1.5	1.5	71.5
ENGINEER	1	.5	.5	72.0
TEACHER	11	5.5	5.5	77.5
INSURANCE	3	1.5	1.5	79.0
OTHER	42	21.0	21.0	100.0
Total	200	100.0	100.0	

**MOTHER'S OCCUPATION**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid HW	176	88.0	88.0	88.0
DOCTOR	1	.5	.5	88.5
GOVT SERVICE	1	.5	.5	89.0
BANKER	3	1.5	1.5	90.5
TEACHER	12	6.0	6.0	96.5
OTHERS	7	3.5	3.5	100.0
Total	200	100.0	100.0	

# Quantitative variables



Analyze → Descriptive statistics → Descriptives...

Visible: 17 of 17 Variables

Variable(s):

- AGE
- HEIGHT
- WEIGHT

Save standardized values as variables

OK Paste Reset Cancel Help

Click on it

ID	CLASS	ST	ADM	FACE	MAGE	F.OCC	M.OCC	F.EDU	M.EDU	INCOME	EXP
1	1	3									
2	2	1									
3	3	2									
4	4	4									
5	5	1									
6	6	3									
7	7	2									
8	8	3									
9	9	1									
10	10	3									
11	11	2									
12	12	5									
13	13	3									
14	14	3									
15	15	1									
16	16	2									
17	17	2									
18	18	1									
19	19	5									
20	20	5									
21	21	5									
22	22	1									
23	23	3									

Data View Variable View

Descriptives...

IBM SPSS Statistics Processor is ready Unicode:ON

# Quantitative variables



## Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AGE	200	5.00	14.00	8.2850	2.06825
height in cm	200	95.00	156.00	126.2900	13.00618
Weight in kg	200	11.50	55.00	25.4656	9.09687
Valid N (listwise)	200				

*Thank You!*