

Data analysis using SPSS

Final Project
Prepared by Walid Soliman

Lecture 9





مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY
BY EVERYONE'S SMART UNIVERSITY

Validity and Reliability

Lecture 8

Final Project
Prepared by Walid Soliman

www.midocean.ae

Chapter 06

Contents

1. Introduction
2. First: **Validity**
3. The concept of **Validity**
4. Types of **Validity**
 - Face **Validity**
 - Content **Validity**
 - Criterion **Validity**
 - Predictive **Validity**
 - Factor **Validity**



جامعة الجميع الذكية
EVERYONE'S SMART UNIVERSITY

Chapter 06

Contents

1. Second: **Reliability**
2. The concept of **Reliability**
3. Methods for measuring **Reliability**
 - The first method: test and re-test
 - The second method: half-partition (Split-half method)
 - The third method: Cronbach's stability coefficient alpha
4. Factors affecting the **Reliability** of the questionnaire
5. The relationship between test **Validity** and **Reliability**



جامعة الجميع الذكية
EVERYONE'S SMART UNIVERSITY

1. **Data Types:** (Quantitative & Qualitative)
2. **The first stage:** Presentation and analysis of data using the SPSS program
3. **The second stage:** Describing the data using statistical measures.
4. **The third stage:** Data analysis using SPSS software

1. Learn about qualitative and quantitative data types.
2. Learn how to display and analyze data using SPSS, including displaying data using different graphic formats
3. Analyze the data by comparing the means of two different groups.
4. Analyze the data by one-way analysis of variance
5. Analyze the data by measuring the correlation coefficient between the variables.
6. Analyze the data through regression analysis between variables.

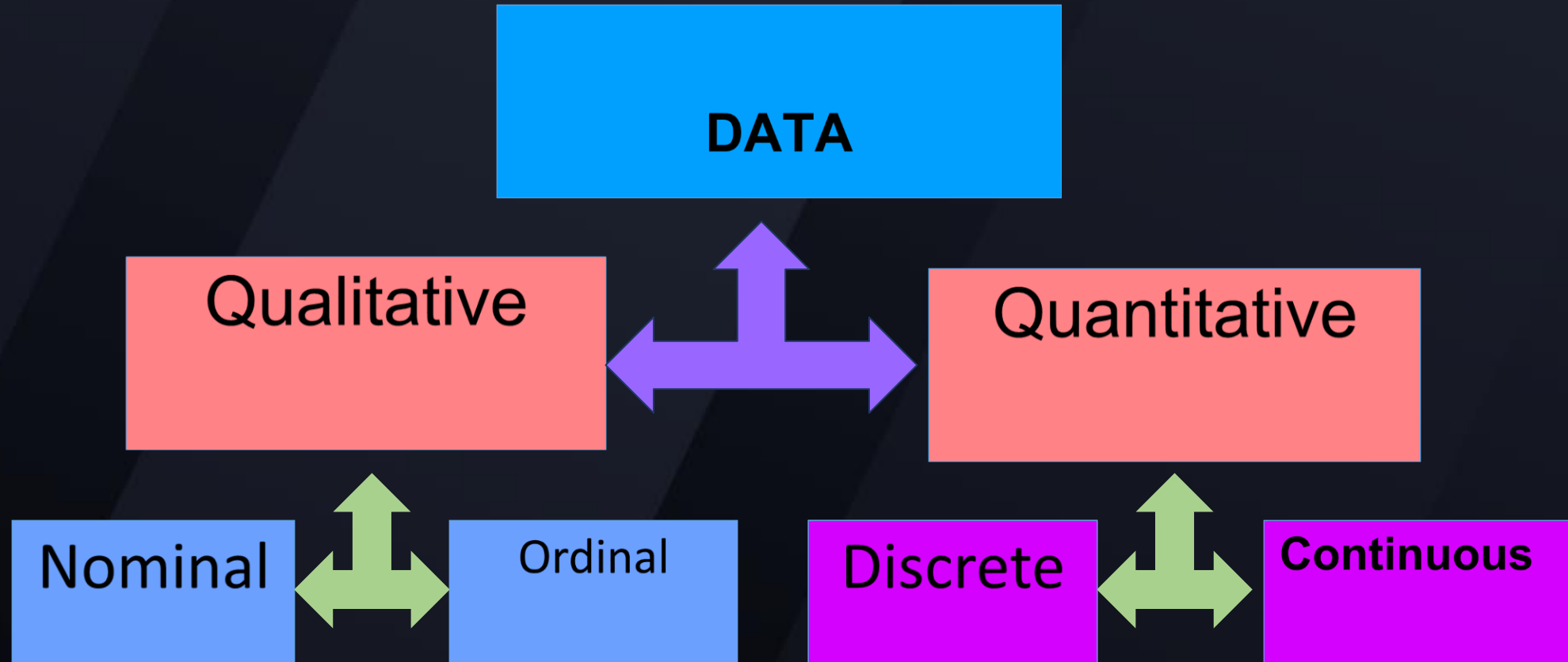
By the end of this chapter you will possess the following Skills:

- The ability to distinguish between different types of data and the appropriate statistical measures for each.
- Data presentation skills using the SPSS program using appropriate graphic figures.
- Data analysis skills on the SPSS program, which includes a comparison between averages, one-way analysis of variance, measurement, and interpretation of the correlation coefficient between variables, regression analysis, and interpretation of results.

"SPSS is the most used statistical program by researchers to analyze information in the field of human, administrative and economic sciences, and helps the researcher to obtain accurate results while saving time and effort."

(Bhandari, P , 2022)

Data Types



First: Qualitative Data

a) Nominal data: male-female, marital status, nationality

No calculations can be performed on it

b) Ordinal data: educational level, degrees of agreement or disagreement.

Second: Quantitative Data

a) Discrete data (discontinuous): number of family members, or workers.

b) Continuous data: sale price, age, income.

(Bhandari, P , 2022)

The first stage: Data presentation.

a) Circle: The proportions of distribution or structural composition of a variable

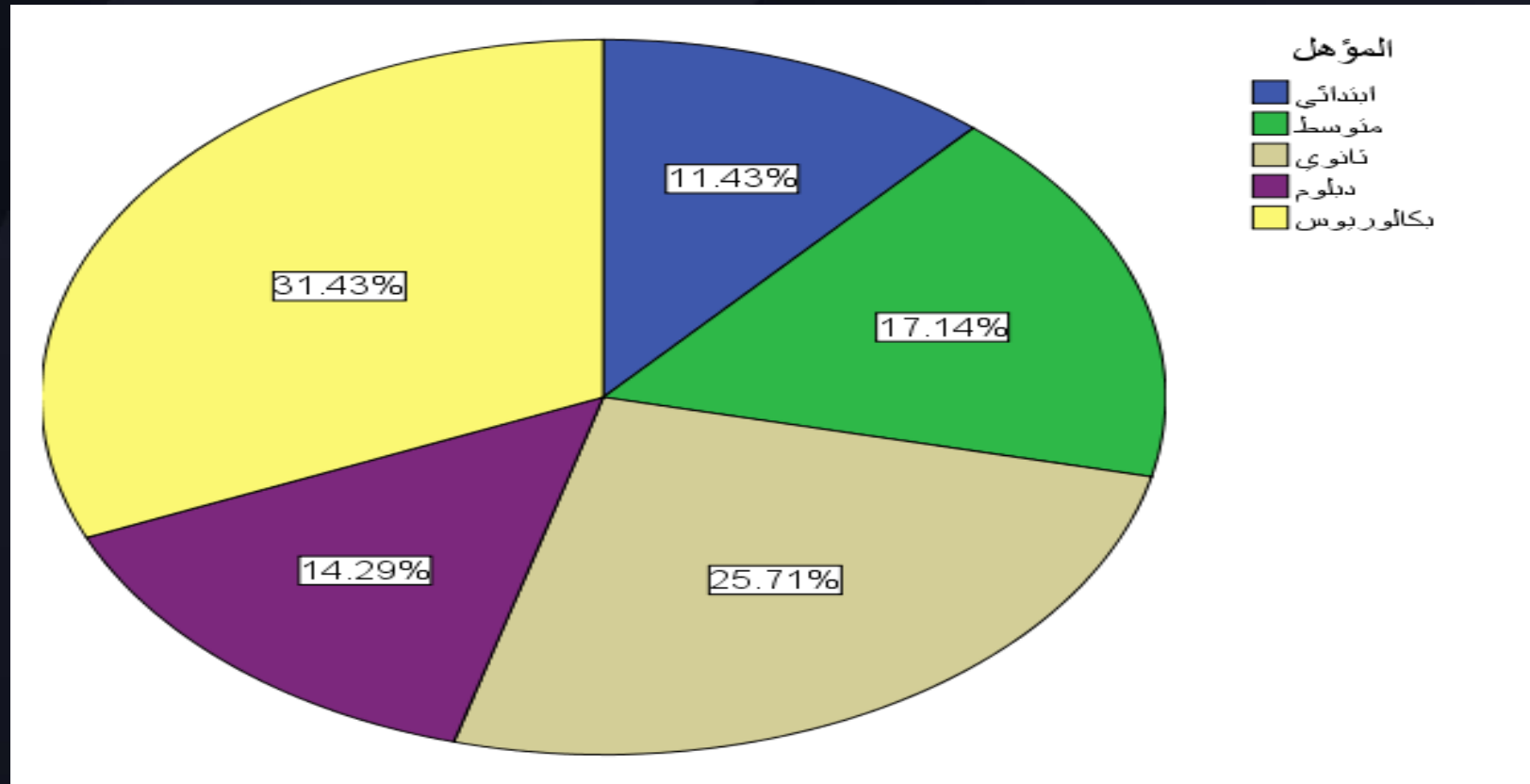
The left screenshot shows the SPSS Data Editor window with the dataset 'staffsurvey5ED.sav' open. The 'Graphs' menu is open, and the 'Legacy Dialogs' > 'Pie' path is highlighted. The data table shows columns for 'id', 'city', 'age', and 'service'.

	id	city	age	service
1	454	8	4	
2	431	8	4	
3	432	8	2	5. E-00
4	433	8	4	5. E-00
5	434	8		
6	435	8	2	
7	436	8	5	
8	437	8	4	
9	442	8	5	
10	445	8	4	1
11	446	8	5	
12	447	8	4	

The right screenshot shows the 'Define Pie: Summaries for Groups of Cases' dialog box. The 'Slices Represent' section has 'N of cases' selected, and the 'Define Slices by' section has '[E] الميزان' selected. Both are circled in red.

The first stage: Data presentation. (Sample of 35 employees)

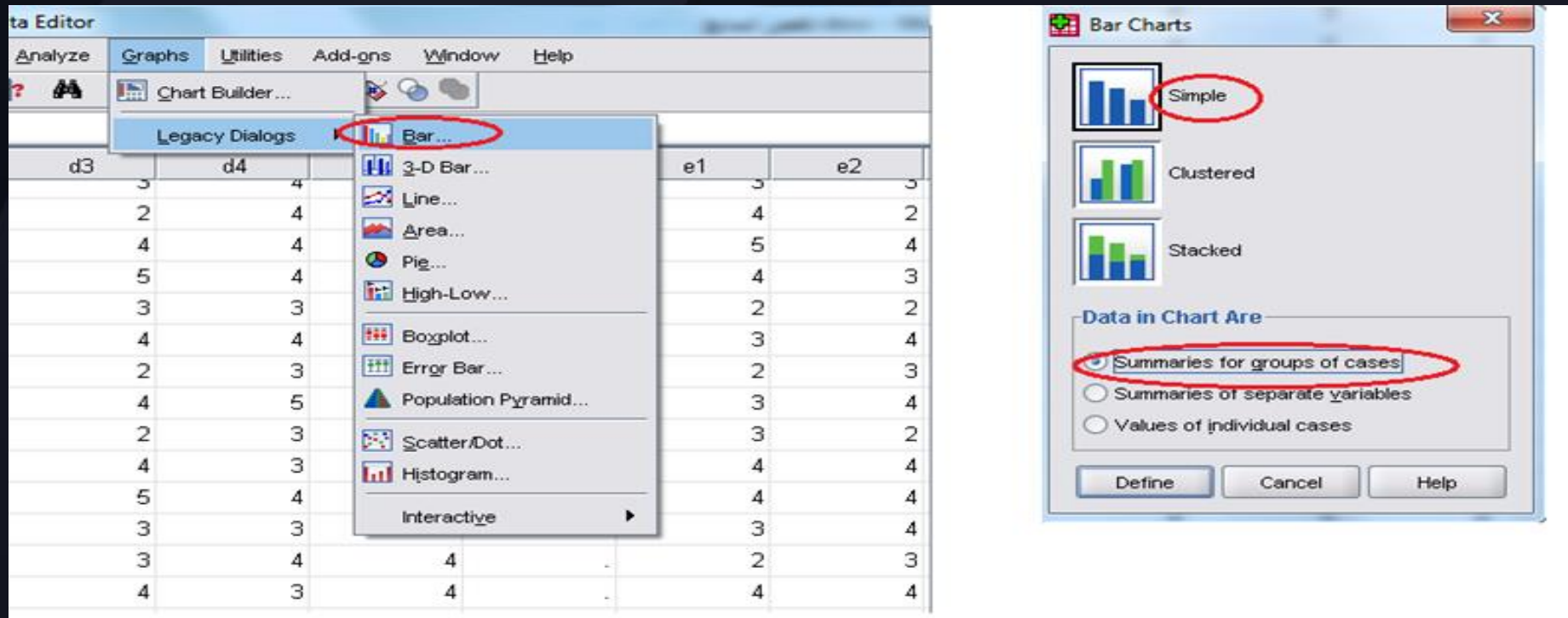
a) Circle: The proportions of distribution or structural composition of a variable



View and analyze data using SPSS software

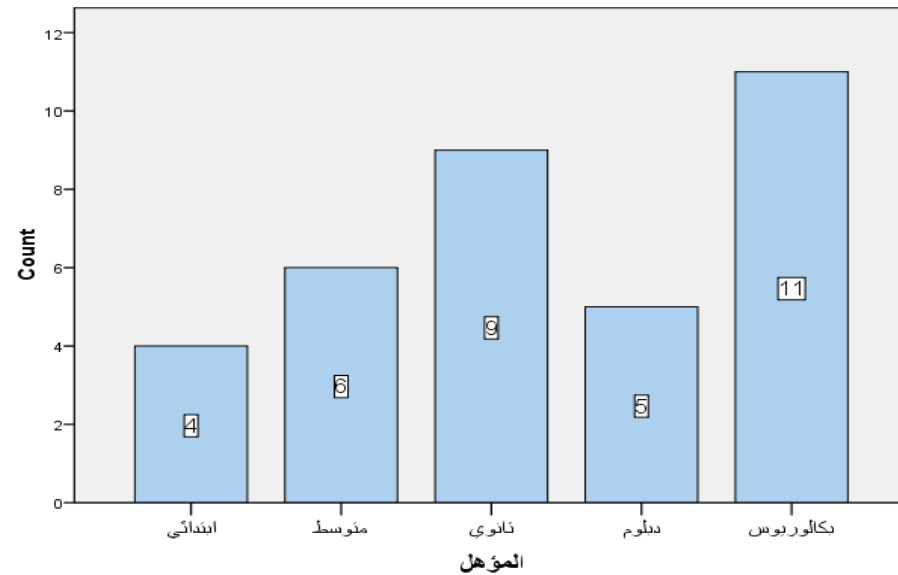
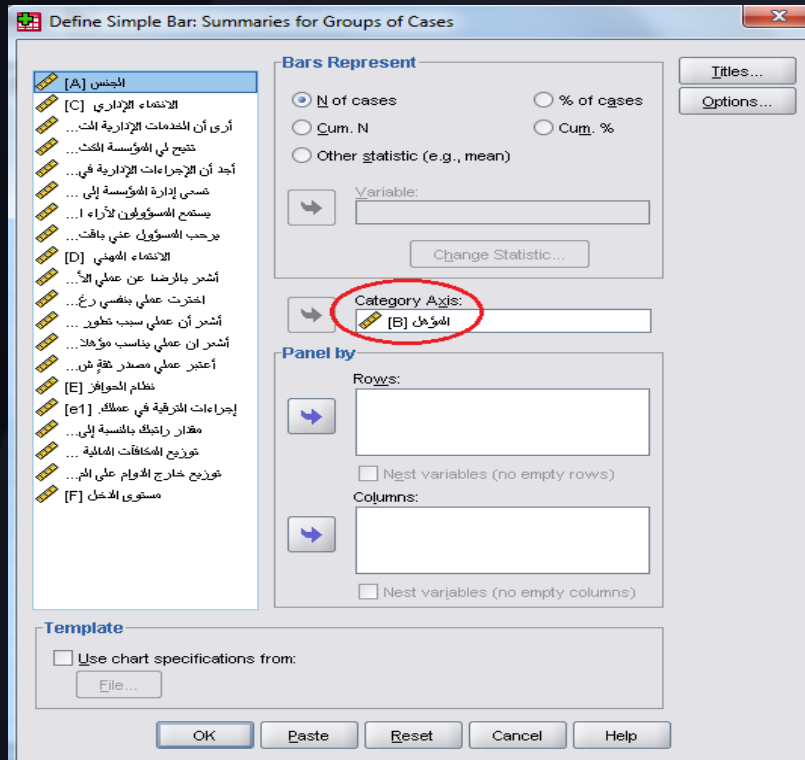
The first stage: Data presentation (Sample of 35 employees).

b) Bar graphs: discrete ordinal and quantitative data.



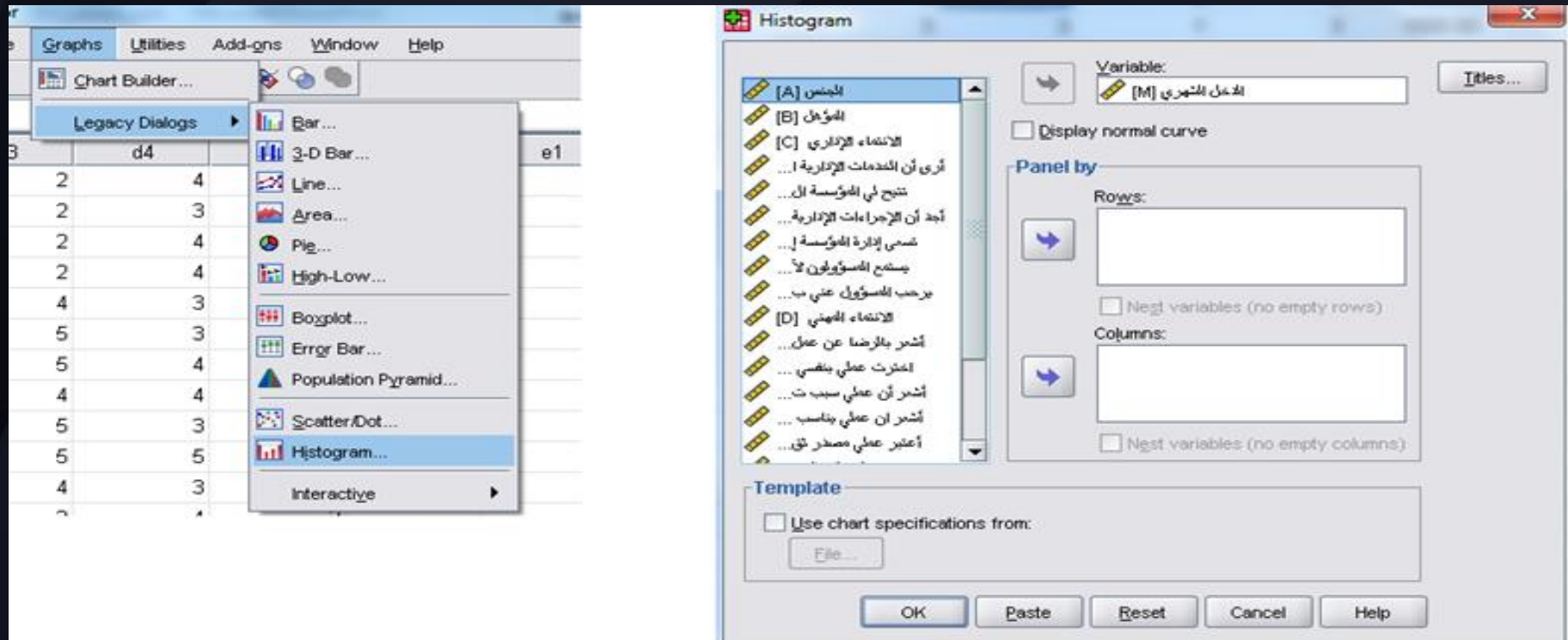
The first stage: Data presentation.

b) Bar graphs: discrete ordinal and quantitative data.



The first stage: Data presentation.

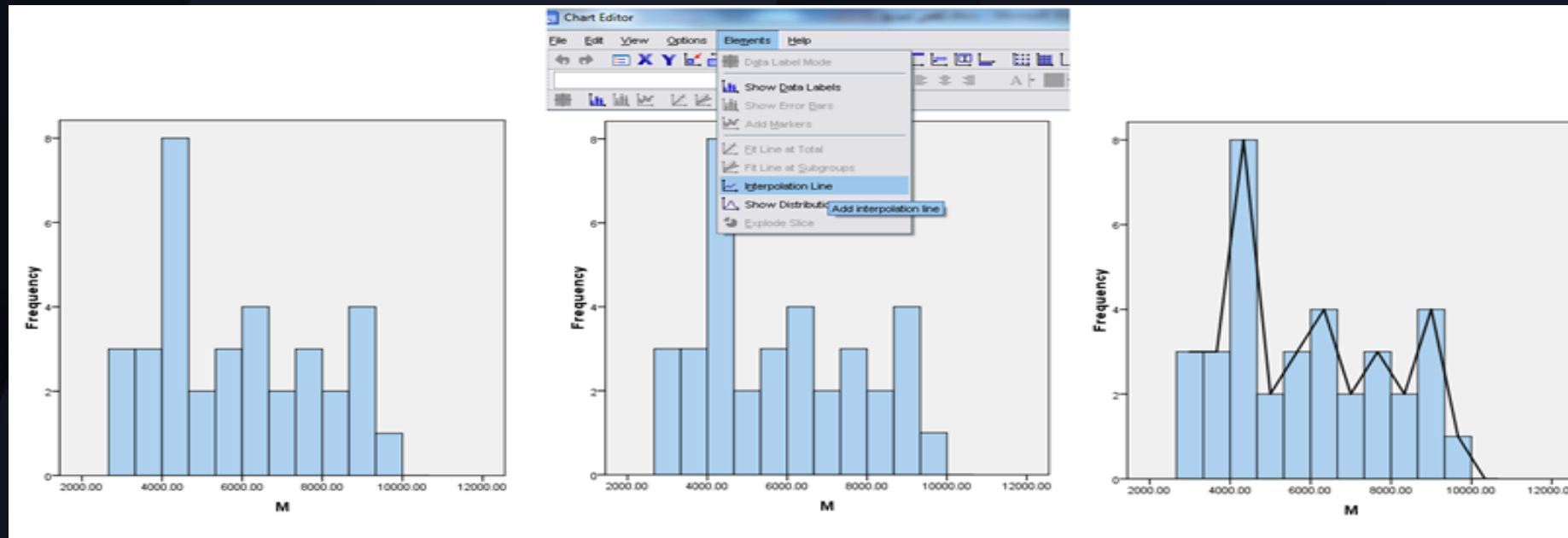
c) Histogram and Interpolation line: connected quantitative data.



The first stage: Data presentation.

c) Histogram and Interpolation line: connected quantitative data.

We choose the Chart editor, and then we choose the Interpolation line from the list of elements.



The second stage: describing the data using statistical measures

A good scale:

- Its exact value should be determined and not left to objectivity.
- To enter all data or observations in his account.
- Easy to understand and calculate.
- His ability to coercive dealing.
- Not affected by outliers (extremes).

The second stage: describing the data using statistical measures

(descriptive statistical measures)

Definition of central tendency

It is the tendency of the data to center around a certain value in the middle of the data.

A. The measure of central tendency:

- Mean
- Median
- Mod

Mean

It is the most common and used measure in all scientific and practical fields, and the process of measuring the arithmetic mean is a very simple process, as in the first step all the values are collected and in the second step the sum of the values is divided by their number to obtain the arithmetic mean for a group of values

Median

It is the value that lies in the middle of the data distribution

Mod

The measure determines the value or variable that occurs frequently in the data distribution. The mode can be calculated for any type of quantitative or descriptive data.

B. Frequency Table:

It is a table that organizes and summarizes statistical data, whether descriptive or quantitative, and distributes them to categories that determine the number of individuals who belong to each category.

C. Measures of Dispersion:

To determine the extent of convergence, homogeneity, or divergence of the given data, we need measures other than measures of central tendency. These measures are called measures of dispersion.

The dispersion of the data (i.e. the divergence around the mean value of the data) can be measured by a number of measures:

- **Variance**
- **Standard Deviation**
- **RANGE**

❖ Variance

To calculate it, the arithmetic mean must be found first, then subtract each value given in the data from the arithmetic mean, square it, then add all the results from the squaring process, then divide the result by the number of values.

❖ Standard deviation

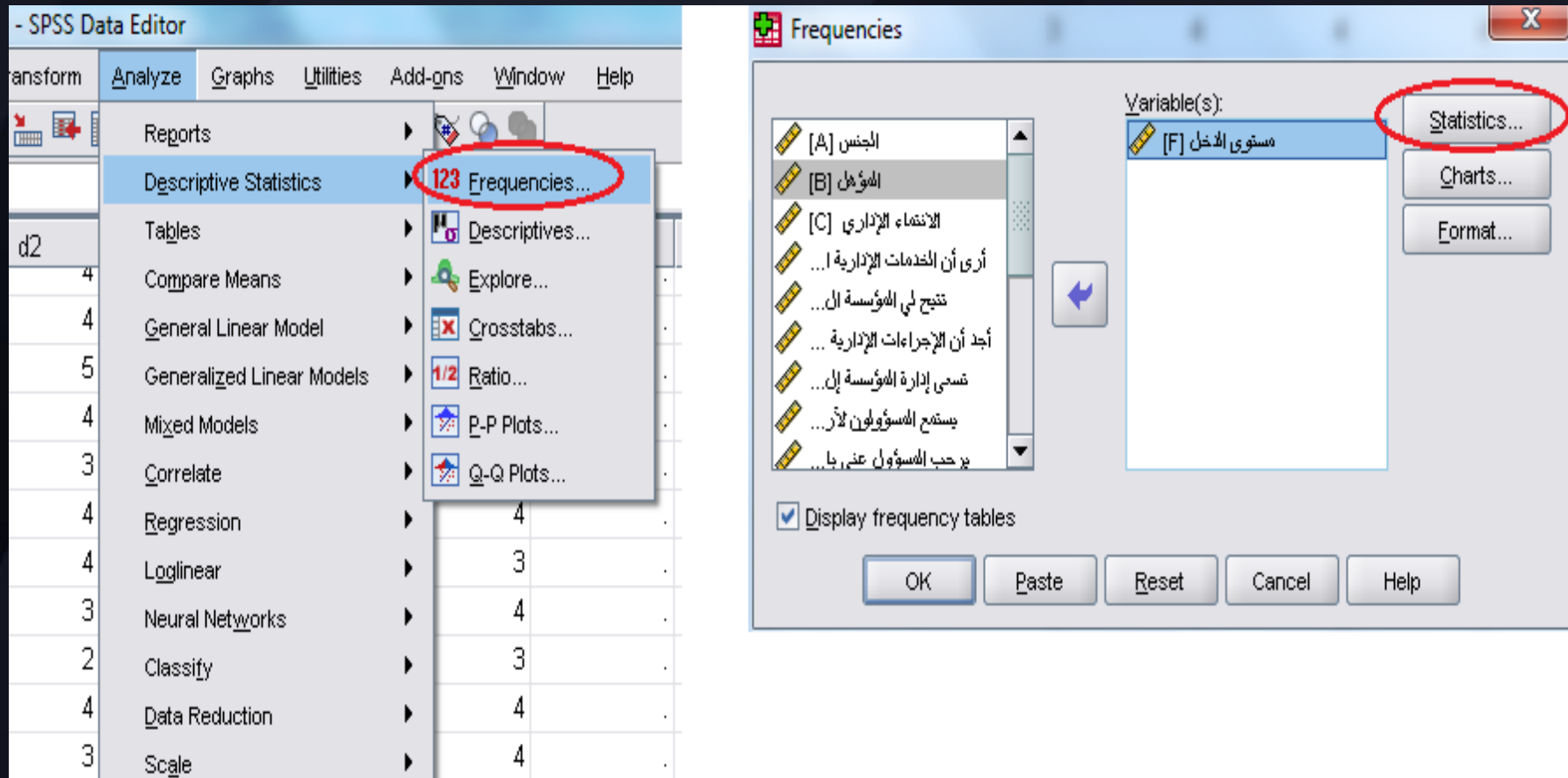
It is a measure of dispersion, which measures how far away or close the data are from their arithmetic mean

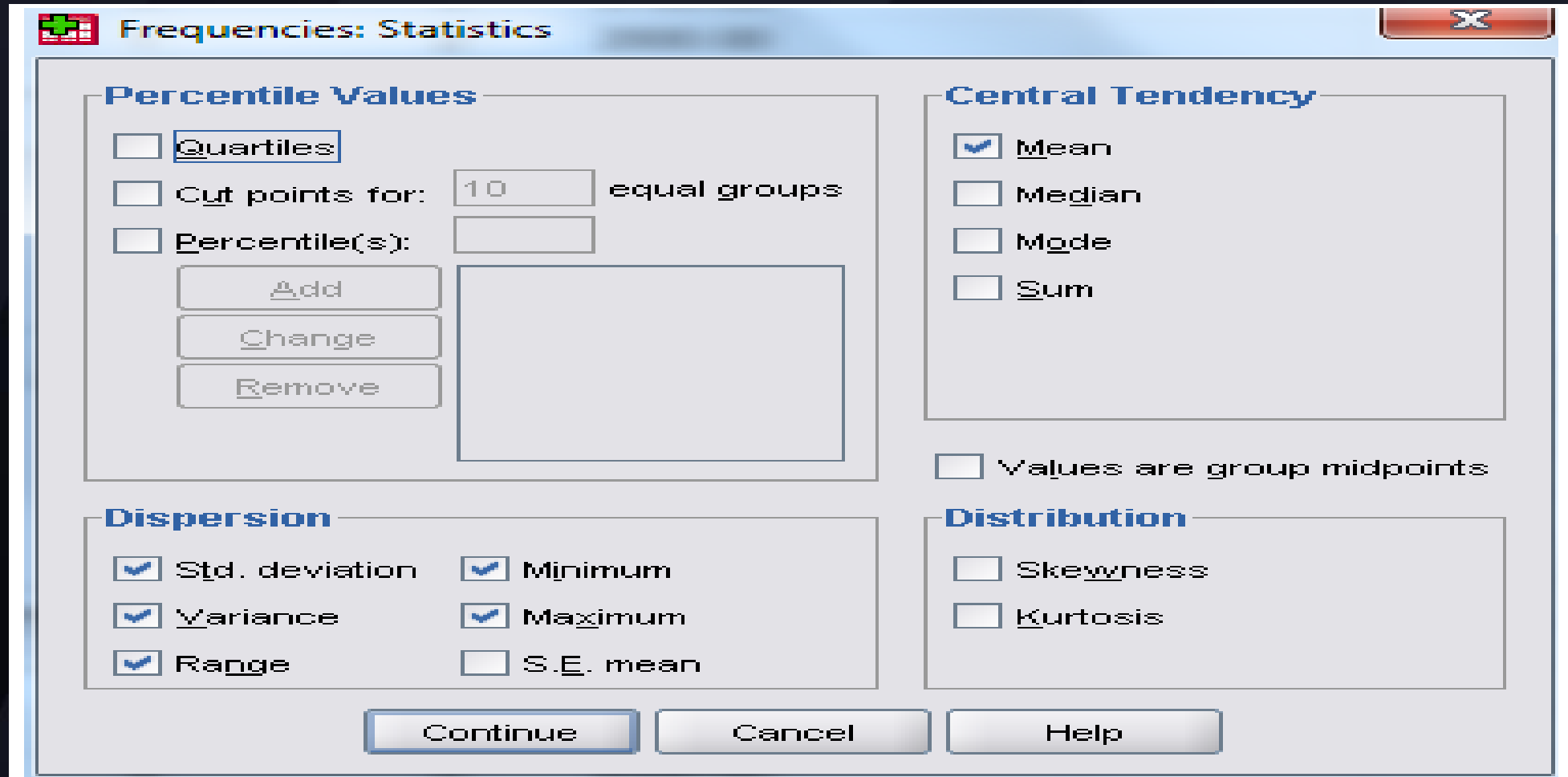
is the positive square root of the variance

❖ Range

It is the difference between the largest value and the lowest value of the data, and it is symbolized by the symbol (R).

❖ The second stage: describing the data using statistical measures





The image shows the 'Frequencies: Statistics' dialog box in SPSS. The dialog is divided into four main sections: Percentile Values, Central Tendency, Dispersion, and Distribution. The 'Percentile Values' section has three checkboxes: 'Quartiles' (unchecked), 'Cut points for: 10 equal groups' (unchecked), and 'Percentile(s):' (unchecked). Below these are three buttons: 'Add', 'Change', and 'Remove'. The 'Central Tendency' section has four checkboxes: 'Mean' (checked), 'Median' (unchecked), 'Mode' (unchecked), and 'Sum' (unchecked). Below this is a checkbox for 'Values are group midpoints' (unchecked). The 'Dispersion' section has six checkboxes: 'Std. deviation' (checked), 'Variance' (checked), 'Range' (checked), 'Minimum' (checked), 'Maximum' (checked), and 'S.E. mean' (unchecked). The 'Distribution' section has two checkboxes: 'Skewness' (unchecked) and 'Kurtosis' (unchecked). At the bottom are three buttons: 'Continue', 'Cancel', and 'Help'.

Frequencies: Statistics

Percentile Values

- ☐ Quartiles
- ☐ Cut points for: equal groups
- ☐ Percentile(s):
-
-
-

Central Tendency

- ☒ Mean
- ☐ Median
- ☐ Mode
- ☐ Sum
- ☐ Values are group midpoints

Dispersion

- ☒ Std. deviation
- ☒ Variance
- ☒ Range
- ☒ Minimum
- ☒ Maximum
- ☐ S.E. mean

Distribution

- ☐ Skewness
- ☐ Kurtosis

❖ The second stage: describing the data using statistical measures

Monthly income variable	
Statistics	Value
Mean	5897.1
Std. Deviation	2026.6
Variance	4107000
Range	6700
Minimum	2800
Maximum	9500

❖ The Second Stage: describing the data using statistical measures

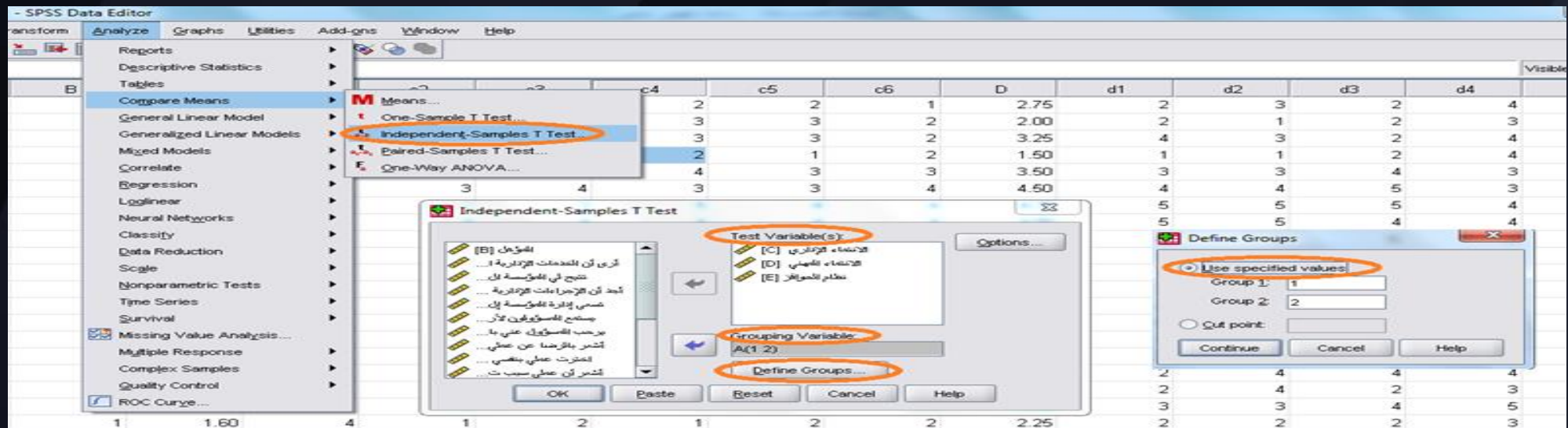
Frequency

The Degree	Frequency	Percent
Primary	4	11.4
Middle School	6	17.1
High School	9	25.7
Diploma	5	14.3
BSC	11	31.4
Total	35	100.0

❖ The third stage:

a) Comparing the mean of two different groups:

Example: In our previous example, we will test the difference or equality between the averages of the three variables (administrative affiliation - professional affiliation - incentive system) according to each gender (male - female).



❖ The third stage:

a) Comparing the mean of two different groups:

Variable	M&F	N	Mean	Std. Deviation
administrative affiliation	M	23	3.087	.92013
	F	12	3.200	.82682
professional affiliation	M	23	3.511	1.04305
	F	12	3.583	.57735
Incentive System	M	23	3.250	.90139
	F	12	3.500	.48850

❖ The third stage: a) Comparing the mean of two different groups:

Independent Samples Test						
Variable		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig.
administrative affiliation	Equal variances assumed	.142	.709	-.357	33	.724
	Equal variances not assumed			-.369	24.66	.715
professional affiliation	Equal variances assumed	10.628	.003	-.223	33	.825
	Equal variances not assumed			-.264	32.80	.793
Incentive System	Equal variances assumed	8.049	.008	-.891	33	.380
	Equal variances not assumed			-1.064	32.89	.295

❖ The third stage: b) One-way ANOVA:

The screenshot displays the SPSS Data Editor interface. The 'Analyze' menu is open, and 'One-Way ANOVA...' is selected. The 'One-Way ANOVA' dialog box is shown, with 'الانتماء الإداري' (C), 'الانتماء المهني' (D), and 'نظام الحوافز' (E) in the 'Dependent List' and 'المؤهل' (B) in the 'Factor' field. The background data table is as follows:

	c4	c5	c6	D	d1
2	2	2	1	2.75	
3	3	3	2	2.00	
3	3	3	2	3.25	
2		1	2	1.50	
4		3	3	3.50	
3	4	3	4	4.50	

View and analyze data using SPSS software

❖ The third stage: b) One-way ANOVA:

Sum of Squares:

- Between Groups
- Within Groups
- df
- Mean Square
- F ratio

Null hypothesis: The means of the sampled populations are equal.

Alternative hypothesis: There are at least two sampled populations with different means.

b) One-way ANOVA The analysis of variance test indicates that the value of F for the three variables is significant at 1%, which means that there is a significant difference in the averages of these variables according to qualification

ANOVA						
Variable		Sum of Squares	Df	Mean Square	F	Sig.
administrative affiliation	Between Groups	20.858	4	5.215	29.031	.000
	Within Groups	5.389	30	.180		
	Total	26.247	34			
professional affiliation	Between Groups	24.294	4	6.073	54.407	.000
	Within Groups	3.349	30	.112		
	Total	27.643	34			
Incentive System	Between Groups	18.177	4	4.544	48.408	.000
	Within Groups	2.816	30	.094		
	Total	20.993	34			

❖ C) Measuring the correlation between variables: Correlation

- Correlation coefficient between +1 and -1
- Correlation coefficient between two variables between ± 0.01 , ± 0.3 (**weak correlation**)
- Correlation coefficient ranged from ± 0.3 , with ± 0.6
- The correlation coefficient is between ± 0.6 and ± 0.99 (**strong correlation**)
- Correlation coefficient = 1, **perfect correlation**
- Pearson Correlation

D) Correlation

The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and the 'Correlate' option is selected, leading to the 'Bivariate...' dialog box. The dialog box is titled 'Bivariate Correlations' and contains the following elements:

- Variables:** A list of variables is shown, including [A] الجنس, [B] المؤهل, [C] الانتماء الإداري, [D] الانتماء المهني, [E] نظام الحوافز, and [M] الدخل الشهري. The variable [C] is highlighted.
- Correlation Coefficients:** The 'Pearson' checkbox is checked, while 'Kendall's tau-b' and 'Spearman' are unchecked.
- Test of Significance:** The 'Two-tailed' radio button is selected, while 'One-tailed' is unselected.
- Flag significant correlations:** This checkbox is checked.
- Buttons:** 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons are at the bottom.

The background shows a data table with columns labeled c2, c3, c4, c5, c6, D, d1, and d2. The data values are as follows:

c2	c3	c4	c5	c6	D	d1	d2
2	2	2	2	1	2.75	2	
1	1						
3	3						
1	1						
4	4						
3	3						
5	5						
3	3						
3	3						
2	2						
4	4						
4	4						
2	2						
4	4						
3	3						
2	2						

D) Correlation

Variables		administrat ive affiliation	professional affiliation	Incentive System	Monthly Income
administrative affiliation	Pearson Correlation	1			
	Sig. (2-tailed)				
professional affiliation	Pearson Correlation	.841**	1		
	Sig. (2-tailed)	.000			
Incentive System	Pearson Correlation	.840**	.901**	1	
	Sig. (2-tailed)	.000	.000		
Monthly Income	Pearson Correlation	.781**	.846**	.789**	1
	Sig. (2-tailed)	.000	.000	.000	

E) Linear Regression analysis

- The causal relationship between the variables
- The dependent variable and the dependent variable (or variables).
- Simple linear regression: $Y = a + b X + u$
- Multiple linear regression:
 - $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n + u$



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY
BY EVERYONE'S SMART UNIVERSITY



Thank You

