TEACHER: DR. LEILA BOUSSAAD

TYPICAL CORRECTION (S5)

PART: 01 [08 POINTS]

Q1- Compare briefly the three clustering methods available in SPSS **(--/03)**

س1.   قارن بإيجاز طرق التحليل العنقودي المتوفرة في برنامج   SPSS

R-

Two-Step CA: Automatically determines the optimal number of clusters, supports both continuous and categorical data, and handles large datasets.

K-Means CA: It requires the number of clusters to be specified, ideal for continuous variables.

Hierarchical CA: Creates a dendrogram for visualizing nested clusters but it is not ideal for large datasets.

Q2- Define '**inertia**' in Factorial Correspondence Analysis (FCA) and provide the formula that relates **total inertia** to the **chi-square statistic**. **(--/03)**

س2. عرف المصطلح 'inertia' في التحليل التوافقي، و أعط الصيغة (العلاقة) التي تربط   total inertia   بـ   chi-square

R- Inertia: Quantifies row/column contribution to data information.

Total inertia = Chi-square statistic / Total sample size

Q3- Explain the goal of '**Optimal Scaling**' and its role in preparing data for analysis **(--/02)**

س3.   اشرح بإيجاز هدف "Optimal Scaling" ودوره في إعداد البيانات للتحليل

R- Optimal Scaling aims to convert categorical data into numerical values while preserving their relationships for better analysis.

It prepares data by making it suitable for statistical methods such as regression and principal component analysis.

PART: 02 [12 POINTS]   Consider the outputs displayed by SPSS:

لتكن النتيجة التي يظهرها برنامج SPSS:

Table 1

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,598
Bartlett's Test of Sphericity	Approx. Chi-Square	10,197
	df	3
	Sig.	,017

Table 2   Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,233	74,432	74,432	2,233	74,432	74,432
2	,581	19,379	93,811			
3	,186	6,189	100,000			

Extraction Method: Principal Component Analysis.

Q1- Which analysis method is used to generate this output? **(--/01)**

س1. ما هي طريقة التحليل المستخدمة للحصول على هذه النتيجة؟

R- Factor analysis (Principal Component Analysis)

Q2- Write the SPSS command that allows obtaining this output (--/**01.5**)

س2. أكتب تعليمة SPSS التي تسمح بالحصول على هذه النتيجة.

R- Analyze -> Dimension Reduction -> Factor…

Q3- Define **'KMO and Bartlett's Test of Sphericity'** and interpret the results presented in **Table 1 (--/04)**

س3. عرف **'KMO and Bartlett's Test of Sphericity'** وفسر النتائج المقدمة في الجدول 1.

R- KMO and Bartlett's Test of Sphericity assess data suitability for factor analysis. The KMO test checks sampling adequacy, while Bartlett's test evaluates whether correlations exist between variables.

KMO = 0.598, which is above 0.5, it indicates an acceptable level of sampling adequacy for factor analysis. Combined with the significant Bartlett's test (p = 0.017), the data is suitable for factor analysis.

Q4- Based on the results; can you determine the number of analysis variables? If yes, how many? (--/**01.5**)

س4. بناءً على النتائج، هل يمكنك تحديد عدد متغيرات التحليل؟ إذا كانت الاجابة نعم، فكم عددها؟

R- Yes, the number is three.

Q5- Based on the results in Table 2, how many components are retained according to the eigenvalue criterion? Justify your answer. **(--/02)**

س5. من الجدول 2، ما عدد المركبات التي تم الاحتفاظ بها وفقًا لمعيار القيمة الذاتية؟ برر إجابتك

R- Only 1 component should be retained, as there is just one eigenvalue greater than 1 (2.233).

Q6- Give another method that can be used to determine how many factors should be retained. **(--/01)**

س6. أعط طريقة أخرى يمكن استخدامها لتحديد عدد المركبات (العوامل) التي يجب الاحتفاظ بها

R- Another method to determine how many factors should be retained is the **Scree plot.**

Q7- Based on Table 2, what is the percentage of total variance explained by the selected components? (--/01)

س7. من الجدول 2، ما هي نسبة التباين الكلي التي تفسرها المركبات المختارة؟

R- The percentage of total variance explained by the selected components is 74.432%.