UNIVERSITY OF BATNA 1 –HADJ LAKHDAR –
FACULTY OF ECONOMICS, BUSINESS AND MANAGEMENT SCIENCES
DEPARTMENT OF COMMERCIAL SCIENCES   **SPECIALITY: Marketing (BMD3)**
COURSE: MARKETING DATA ANALYSIS **(تحليل البيانات التسويقية)**

TEACHER: DR. LEILA BOUSSAAD

TYPICAL CORRECTION (S5)

PART: 01 [08 POINTS]

Q1- What are the main differences between Factor Analysis and Cluster Analysis? **(--/02)**

س1- ما هو الفرق بين التحليل العاملي والتحليل العنقودي؟

R-

Factor Analysis: Grouping is based on correlation. We form a group of variables based on several people's responses to those variables.

Cluster Analysis: Grouping is based on the distance (proximity). We group cases based on their responses to several variables.

Q2- Define the terms '**loadings**' and '**communalities**' in Principal Component Analysis, and provide the formulas used to calculate them. **(--/03)**

س2. عرف المصطلحين '**loadings**' و '**communalities**' في تحليل المركبات الرئيسية، و اعط الصيغ (العلاقات) المستخدمة لحسابهما.

R- Loadings represent the correlation or contribution of each variable to a specific factor in factor analysis.

$$L_{ij} = \text{correlation}(X_i, F_j)$$ $\boxed{L_{ij}: \text{Loading of variable } X_i \text{ on factor } F_j.}$

Communalities measure the proportion of each variable's variance that is explained by the extracted factors.

$$h_i^2 = \sum_{j=1}^{m} L_{ij}^2$$

Q3- Compare briefly the three clustering methods available in SPSS **(--/03)**

س3.   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.

PART: 02 [12 POINTS]   Consider the output displayed by S   لتكن النتيجة التى يظهرها برنامج SPSS:

Table 1

EC	HC Black	Brown	Red	Blonc	Active Margin
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	10	93
Green	5	29	14	16	64
Active Margin	108	286	71	127	592

Table 2

EC	HC Black	Brown	Red	Blonc	Active Margin
Brown	,309	,541	,118	,032	1,000
Blue	,093	,391	,079	,437	1,000
Hazel	,161	,581	,151	,108	1,000
Green	,078	,453	,219	,250	1,000
Mass	,182	,483	,120	,215	

(*)

Summary

Dimension	Singular Value	Inertia	Table 3 : Chi Square	Sig.	Proportion of Inertia Accounted for	Cumulative	Confidence Singular Value Standard Deviation	Correlation 2
1	,457	,209			,894	,894	,032	,134
2	,149	,022			,095	,989	,042	
3	,051	,003			,011	1,000		
Total		,234	138,290	,000ᵃ	1,000	1,000		

a. 9 degrees of freedom

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

س1. ما هي طريقة التحليل المستخدمة للحصول على هذا الناتج؟

R- Factor Correspondence Analysis

Q2- What is the objective of this method and the data requirements? **(--/02)**

س2. ما هو هدف هذه الطريقة وما هو نوع البيانات اللازمة؟

R- Objective: Analyze and visualize relationships between rows and columns in a contingency table.
 Data Requirements: Categorical data summarized in a contingency table.

Q3- Based on Table 1, provide the analysis variables and the total sample size. **(--/02)**

س3. من الجدول 1، أعط متغيرات التحليل وحجم العينة الإجمالي.

R- Analysis variables: **HC, EC**.   Sample size= **592**

Q4- What does the value in the (*) frame represent? **(--/01)**   س4. ماذا تمثل القيمة الموجودة في الإطار (*)؟

R- 108 cases with the modality 'Black' for the variable HC represent 18.2% of the total sample of 592.

Q5- Write the SPSS command that allows obtaining the previous output **(--/01.5)**

س5. أكتب تعليمة SPSS التي تسمح بالحصول على النتيجة السابقة

R- Analyze -> Dimension Reduction -> Correspondence Analysis…

Q6- Give the **titles** of Tables 1, and 2 along with their **contents**. **(--/01.5)**

س6. أعط عناوين الجداول 1 و2 مع محتوياتها.

R- Table1: Correspondence table. It shows the count for each category within each variable.

Table2: Row profiles. It shows the proportion of each column category within each row.

Q7- Provide the conclusions that can be drawn from the table 3 (**Summary**). **(--/03)**

س7. ماهي الاستنتاجات التي يمكن استخلاصها من الجدول 3   (**Summary**) ؟

R-

1. Three dimensions were generated, but only two are meaningful, accounting for 98.9 % of the total inertia.
2. The sum of the first two partial inertias accounts for 98.9 % of the total inertia, indicating that the analysis is of good quality.
3. The Chi-square test for independence (Chi-square = 138.29, df = 9, sig. = 0) indicates a highly significant association between the rows and columns.