UNIVERSITY OF BATNA1 –HADJ LAKHDAR –
FACULTY OF ECONOMICS, BUSINESS AND MANAGEMENT SCIENCES
DEPARTMENT OF MANAGEMENT SPECILITY: FINANCIAL MANAGEMENT (BMD3)
COURSE: DATA ANALYSIS (تحليل البيانات)

TEACHER: DR. LEILA BOUSSAAD

MODEL ANSWER

PART: 01 [08 POINTS]

Q1- Why do we **standardize the data** before applying PCA? **(--/02)**

س1. لماذا نقوم بتعيير البيانات قبل تطبيق التحليل بالمركبات الرئيسية؟

Answer- We standardize the data to give all variables the same importance, especially when they have different units or scales. This avoids one variable dominating the results just because it has larger values.

Q2- What is the role of the chi-square statistic in **Correspondence Analysis** (CA) **(--/02)**

س2. ما هو دوز اختبار كاي-تربيع في التحليل التوافقي (CA)؟

Answer- The chi-square statistic measures the strength of the association between rows and columns in the contingency table. It tells us how much the observed values differ from what we would expect if there were no relationship.

Q3- What is the purpose of a **dendrogram**? **(--/02)** **س3. ما هو دور المخطط dendrogram؟**

Answer- A dendrogram is used to show how clusters are formed step by step in hierarchical clustering. It helps to decide the number of clusters by cutting the tree at a chosen level.

Q4-Explain the goal of '**Optimal Scaling**' and its role in preparing data for analysis **(--/02)**

س4. اشرح بإيجاز هدف "Optimal Scaling" ودوره في إعداد البيانات للتحليل

Answer- 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 output displayed by SPSS:

لتكن النتيجة التي يظهرها برنامج SPSS:

Table 1

Age	Sport: football	Swimming	Walking	Cycling	Active Margin
< 15	25	10	10	45	90
15 - 30	8	55	22	57	142
30 - 60	6	24	40	3	73
Active Margin	39	89	72	105	305

Table 2

Age	Sport: football	Swimming	Walking	Cycling	Active Margin
< 15	,278	,111	,111	,500	1,000
15 - 30	,056	,387	,155	,401	1,000
30 - 60	,082	,329	,548	,041	1,000
Mass	,128	,292	,236	,344	

(*)

Table 3 :

Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation 2
1	,500	,250			,729	,729	,047	,111
2	,305	,093			,271	1,000	,052	
Total		,343	104,711	,000[a]	1,000	1,000		

a. 6 degrees of freedom

Q1- Which analysis method is used to generate this output?

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

Answer- Factor Correspondence Analysis

Q2- What is the objective of this method and the data requirements?

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

Answer-

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**.**

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

Answer- Analysis variables: **Age, Sport**. Sample size= **305.**

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

Answer- 39 cases with the modality '**Football'** for the variable **Sport** represent **12.8%** of the total sample of **305**.

Q5- Write the SPSS command that allows obtaining the previous output

س5. أكتب تعليمة SPSS التي تسمح بالحصول على النتيجة السابقة

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

Q6- Give the **titles** of Tables 1, and 2 along with their **contents**. **س6. أعط عناوين الجداول 1 و2 مع محتوياتها.**

Answer- 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).**

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

Answer-

1. The analysis resulted in two dimensions, both of which are significant and together explain the entirety of the total inertia (100%), which suggests a high quality and well-structured solution.
2. The Chi-square test for independence (Chi-square = 104.711, df = 6, sig. = 0) indicates a highly significant association between the rows and columns.

GOOD LUCK