جامعة العربي بن مهيدي – أم البواقي-
كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير
قسم العلوم الاقتصادية
الاستاذ : بوزيان محمد

السنة الجامعية: 2025 - 2026
المستوى: ثالثة ليسانس
التخصص: اقتصاد كمي
المدة: ساعة ونصف

الامتحان السداسي الثاني في مقياس تحليل البيانات2

التمرين الاول :

يمثل الجدول التالي نوع الدراسات التي يواصلها الطلبة (جامعة، أقسام تحضيرية، مسارات أخرى) تبعًا للمسار الدراسي في الثانوية (حسب تصنيف 1970):

- A : أدبي ← Littéraire
- B : اقتصاد ← Economie
- C : علوم (رياضيات، فيزياء) ← Science (Math, Physique)
- D : علوم طبيعية (بيولوجيا، كيمياء)← Science (Biologie, Chimie)
- E: تكنولوجي (نظري) ← Technologique (Théorique)
- Techn : Technologique (Appliquée)

	Univ	Classes prépa	Autres
A	13	2	5
BDD'	20	2	8
CE	10	5	5
Techn.	7	1	22

المطلوب:

1. حساب جدول التكرارات النظرية في حالة الاستقلالية .
2. حساب مسافة كاي-تربيع .(χ^2)؟ مع تفسير النتيجة
3. إجراء اختبار الاستقلالية باستعمال .χ^2
4. حساب مصفوفة بروفيلات الصفوف X ومصفوفة بروفيلات الأعمدة Y مع التعليق .
5. في إطار التحليل العاملي للتوافق(AFC) ، لدينا القيم الذاتية التالية :

$$\lambda 2=0.04936 \quad \lambda 1=0.1998,$$

- ماهو العدد المتوقع للمحاور العملية (برر الاجابة) ؟ قم بتفسير النتائج.

6. لتكن احداثيات كل من ملامح السطر والعمود على المحورين العامليين كما يلي :

	Axis1	Axis2
A	-0.2972483	-0.11354938
BDD'	-0.2539284	-0.21594323
CE	-0.3435636	0.40399761
Tech	0.6811363	0.02231108

	Comp1	Comp2
Univ	-0.3405158	-0.14392945
Prepa	-0.4785344	0.62263268
Autres	0.5452783	0.02425365

- قم بتمثيل كل منهما على نفس الخريطة العاملية ؟ - قم بتفسير النتائج ؟

- **التمرين الثاني :**

أُجري تحليل تصنيف هرمي تصاعدي (Classification Ascendante Hiérarchique) على مجموعة من المؤسسات الاقتصادية اعتمادًا على عدة مؤشرات تتعلق بالأداء والإنتاجية.

وقد تم الحصول على شجرة التصنيف (Dendrogramme) التالية:

- في المرحلة الأولى تم دمج:
  - Aمع B عند المسافة 2
  - Dمع E عند المسافة 3
- في المرحلة الثانية:
  - تم دمج العنصر C مع المجموعة (A,B) عند المسافة 5
- في المرحلة الثالثة:
  - تم دمج العنصر F مع المجموعة (D,E) عند المسافة 6
- في المرحلة الأخيرة:
  - تم دمج المجموعتين الكبيرتين ((A,B,C)) و ((D,E,F)) عند المسافة 11

1- أعد رسم شجرة التصنيف (Dendrogramme) انطلاقًا من المعطيات السابقة.
2- حدد عدد الأصناف الممكن الحصول عليها إذا تم قطع الشجرة عند:

1. المسافة 3
2. المسافة 5
3. المسافة 8

مع تحديد عناصر كل صنف في كل حالة.

3- ما هي أول مجموعة تم تشكيلها؟ وما هو تفسير ذلك إحصائيًا؟
4- ما هي آخر عملية دمج تمت في الشجرة؟ وماذا تستنتج حول العلاقة بين المجموعتين الناتجتين؟

5- إذا أراد الباحث الاحتفاظ بتصنيف متجانس وفي نفس الوقت قليل الأصناف، فما هو عدد الأصناف الأنسب؟ علّل إجابتك.
6- حدد عند أي مستوى من المسافة يظهر أكبر فقدان للتجانس بين الأصناف، وما دلالة ذلك في اختيار عدد المجموعات؟
7- إذا تم اعتماد قاعدة القطع عند أكبر قفزة في المسافات، فما هو عدد الأصناف النهائي المقترح؟ علّل إجابتك إحصائيًا.
8- إذا اعتبرنا أن العناصر تمثل زبائن لمؤسسة تجارية، فكيف يمكن استغلال نتائج التصنيف في وضع استراتيجية تسويقية؟
9- ما الفرق بين:

- الدمج عند مسافة صغيرة
- والدمج عند مسافة كبيرة

في إطار تفسير شجرة التصنيف؟

10- اشرح لماذا يعتبر التصنيف الهرمي التصاعدي من الأساليب الاستكشافية غير المراقبة (Méthodes non supervisées).
11- اقترح كيف يمكن التأكد من جودة التصنيف المتحصل عليه باستعمال أدوات إحصائية أو بيانية أخرى.

جامعة العربي بن مهيدي – أم البواقي-
كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير
قسم العلوم الاقتصادية
الاستاذ : بوزيان محمد

السنة الجامعية: 2025 - 2026
المستوى: ثالثة ليسانس
التخصص: اقتصاد كمي
المدة: ساعة ونصف

الاجابة النموذجية

التمرين الاول :..... 10 نقاط

1- جدول التكرارات النظرية يعطى كما يلي : ..... 01 نقاط

Filière	Université	Prépa	Autres	Total
A	10	2	8	20
BDD	15	3	12	30
CE	10	2	8	20
Techn.	15	3	12	30
Total	50	10	40	100

مثلا : $e_{41}=(30\times50) / 100 = 15$

-2

من خلال عبارة كاي تربيع : $\chi2=\sum (oij-eij)^2 / eij$ 01نقاط

Filière	Université	Prépa	Autres
A	0.900	0.000	1.125
BDD	1.667	0.333	1.333
CE	0.000	4.500	1.125
Techn.	4.267	1.333	8.333

مثلا : $(5-2)^2/2 = 4.500$

يبين جدول مساهمات الكاي-تربيع أن أكبر المساهمات في قيمة chi^2 جاءت من الخانات:

- **Techn. – Autres** بقيمة 8.333
- **C – Prépa** بقيمة 4.500
- **Techn. – Univ** بقيمة 4.2674

وهذا يدل على أن أكبر الانحرافات عن حالة الاستقلال تتمثل في:

- توجه طلبة المسار التكنولوجي بشكل كبير نحو فئة"Autres" ،
- توجه الطلبة العلميين نحو الأقسام التحضيرية،
- ضعف توجه طلبة المسار التكنولوجي نحو الجامعة .

أما الخانات ذات المساهمات الضعيفة أو المنعدمة، فتشير إلى تقارب بين القيم المشاهدة والنظرية، وبالتالي تأثير ضعيف في تفسير العلاقة بين المتغيرين.

3- .... 01,5 نقاط

$$\chi2=0.900+0+1.125+1.667+0.333+1.333+0+4.500+1.125+4.267+1.333+8.333$$

$$\chi2\approx24.916$$

بلغت قيمة إحصائية كاي-تربيع: $\chi2\approx24.92$ وهي قيمة مرتفعة مقارنة بالقيمة الحرجة الجدولية عند درجات الحرية: $ddl=(4-1)(3-1)=66$

حيث أن: chi^2_{0.05,6}=12.59

وبما أن: 24.92 > 12.592

فإننا نرفض فرضية الاستقلال بين المتغيرين.

وبالتالي نستنتج وجود علاقة إحصائية معنوية بين نوع المسار الدراسي ونوع الدراسة المتابعة، أي أن توجيه الطلبة نحو الجامعة أو الأقسام التحضيرية أو المسارات الأخرى يتأثر بشكل واضح بنوع التخصص الدراسي الأصلي.

4- .... 01,5 نقاط
تمثل هذه القيم مقدار المعلومات أو التباين (Inertie) الذي يفسره كل محور عاملي في تحليل التوافق (AFC). حيث نلاحظ أن:

lambda_1 > \lambda_2 ($\lambda1>\lambda2$)

وبفارق كبير، مما يعني أن المحور الأول هو المحور الأكثر أهمية في تفسير العلاقات الموجودة بين المتغيرات، بينما يضيف المحور الثاني معلومات مكملة لكنها أقل تأثيرًا.

كما يمكن حساب نسب التفسير: (80% F1) و (20% F2) أي أن: المحور الأول يفسر حوالي 80% من القصور الذاتي الكلي، والمحور الثاني يفسر حوالي 20%. وبالتالي فإن معظم البنية الإحصائية للبيانات يمكن تفسيرها انطلاقًا من المحور الأول فقط، مما يدل على أن الخريطة العاملية تعطي تمثيلًا جيدًا وواضحًا للعلاقات بين الفئات.

5- بروفايلات الاسطر والاعمدة .....2.5 نقاط

Filière	Université	Prépa	Autres	Total

A	(13/20=0.650)	(2/20=0.100)	(5/20=0.250)	1
B	(20/30=0.667)	(2/30=0.067)	(8/30=0.266)	1
C	(10/20=0.500)	(5/20=0.250)	(5/20=0.250)	1
Techn.	(7/30=0.233)	(1/30=0.033)	(22/30=0.733)	1

تُظهر بروفيلات السطر أن المسارين A و B يمتلكان سلوكًا متقاربًا، حيث يتجه أغلب الطلبة نحو الجامعة. أما المسار C فيتميز بنسبة معتبرة من الطلبة المتوجهين نحو الأقسام التحضيرية، مما يعكس الطابع العلمي والانتقائي لهذا المسار. في المقابل، يُظهر المسار التكنولوجي سلوكًا مختلفًا بوضوح، إذ تتركز أغلبية طلبته ضمن فئة «أخرى»، وهو ما يدل على توجههم نحو تكوينات تطبيقية أو مهنية بديلة عن المسار الجامعي التقليدي.

جدول بروفايلات العمو د :

Filière	**Université**	**Prépa**	**Autres**
A	(13/50=0.260)	(2/10=0.200)	(5/40=0.125)
B	(20/50=0.400)	(2/10=0.200)	(8/40=0.200)
C	(10/50=0.200)	(5/10=0.500)	(5/40=0.125)
Techn.	(7/50=0.140)	(1/10=0.100)	(22/40=0.550)

تُبين بروفيلات العمود أن الجامعة تستقبل أساسًا الطلبة القادمين من المسارين A و .B أما الأقسام التحضيرية فتضم نسبة كبيرة من طلبة المسارC ، مما يؤكد الارتباط القوي بين التخصصات العلمية والمسارات الانتقائية. في حين تهيمن الفئة التكنولوجية على فئة «أخرى»، وهو ما يعكس خصوصية هذا المسار وارتباطه بالتكوينات غير الجامعية التقليدية.

6- 2.5نقاط

تُظهر الخريطة العاملية وجود علاقة واضحة بين نوع المسار الدراسي ونوع الدراسة المتابعة. حيث ترتبط الفئة التكنولوجية بقوة بفئة «أخرى»، مما يدل على توجهها نحو تكوينات تطبيقية أو غير جامعية، بينما ترتبط الفئة العلمية C بالأقسام التحضيرية، وهو ما يعكس توجهًا نحو المسارات الانتقائية. في المقابل، تتمركز الفئتان A و Bبالقرب من الجامعة، مما يشير إلى ميلهما نحو الدراسة الجامعية التقليدية. وبالتالي فإن المحور الأول يميز أساسًا بين المسارات الأكاديمية التقليدية والمسارات التكنولوجية التطبيقية، في حين يعكس المحور الثاني الفرق بين التعليم الانتقائي والتعليم الجامعي العادي.

التمرين الثاني :..... 10 نقاط

1- الترتيب الزمني لعمليات الدمج هو كما يلي: .....01,5 نقاط

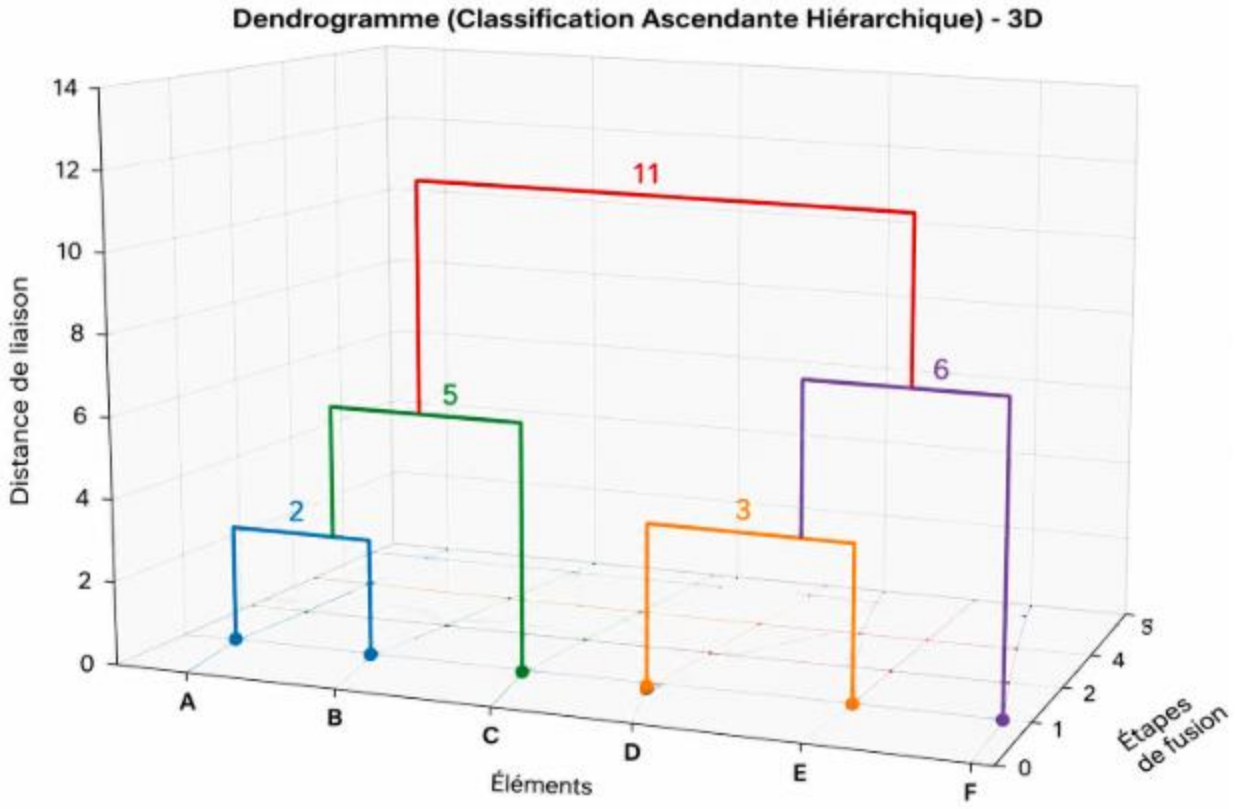
دمج A مع B عند المسافة 2 .

دمج D مع E عند المسافة 3 .

دمج C مع المجموعة (A, B) عند المسافة 5 .

دمج F مع المجموعة (D, E) عند المسافة 6 .

دمج المجموعتين (A, B, C) و (D, E, F)عند المسافة 11.



2- تحديد عدد الأصناف عند مستويات قطع مختلفة : .....01,5 نقاط

أ) القطع عند المسافة 3 : والذي يمكننا من الحصول على عدد الأصناف = 4 وهي

{F} {C}. {D, E} ,{A, B}

وهنا نأخذ فقط الدمجات التي تمت عند مسافة ≤ 3.

ب) القطع عند المسافة 5 : والذي يمكن من الحصول على عدد الأصناف = 3

{A, B, C}و{D, E} و {F}

حيث نأخذ الدمجات التي تمت عند مسافة ≤ 5.

ج) القطع عند المسافة 8 : و نأخذ الدمجات التي تمت عند مسافة ≤ 8.

(A, B, C)

(D, E, F) حيث نتحصل على عدد الأصناف = 2 وهي {A, B, C}

{D, E, F}

2- أول مجموعة تم تشكيلها: .....01,5 نقاط

بعد أول عملية دمج هي: A مع B عند المسافة 2 .

التفسير الإحصائي

المسافة 2 هي أصغر مسافة في الشجرة، مما يعني أن A و Bهما الأكثر تشابهًا من بين جميع العناصر، أي أن خصائصهما متقاربة جدًا.

4- آخر عملية دمج : .....01 نقاط

آخر عملية دمج تمت عند المسافة 11 بين: و (A, B, C) (D, E, F)

الاستنتاج

المسافة 11 كبيرة نسبيًا، ما يدل على أن المجموعتين تختلفان بشكل واضح، وأنهما تمثلان فئتين متميزتين جدًا.

5- العدد الأنسب للأصناف : 01نقاط

الاختيار الأنسب هو = **2 وهما** {A, B, C} {D, E, F}

التعليل

- يعطي تصنيفًا بسيطًا وسهل التفسير .
- يحافظ على تجانس جيد داخل كل صنف .
- يوجد فرق واضح بين المجموعتين .

6-أكبر فقدان للتجانس : .....01 نقاط

الفرق	الانتقال
1	$2 \rightarrow 3$
2	$3 \rightarrow 5$
1	$5 \rightarrow 6$
5	$6 \rightarrow 11$

أكبر قفزة هي:

- $6 \rightarrow 11$ فرق = 5

الدلالة

دمج المجموعتين عند هذا المستوى يؤدي إلى فقدان كبير للتجانس، لذا يُفضّل التوقف قبل هذا الدمج.

7- قاعدة أكبر قفزة.....0,5 نقاط

بما أن أكبر قفزة حدثت بين 6 و11، فإن القطع يتم قبل المسافة 11. لذلك يعطي التصنيف النهائي المقترح

عدد الأصناف = 2

8- الاستغلال التسويقي للنتائج .....01 نقاط

إذا كانت العناصر تمثل زبائنًا: فان الشريحة الأولى{A, B, C} : زبائن لهم خصائص وسلوكيات متقاربة.

الشريحة الثانية{D, E, F} : زبائن يختلفون عن الشريحة الأولى.

التطبيقات التسويقية وتتمثل في :

- تصميم عروض خاصة لكل شريحة .
- تخصيص الحملات الإعلانية .
- تحسين برامج الولاء .
- رفع فعالية الاستهداف وتقليل التكاليف .

9- الفرق بين الدمج عند مسافة صغيرة وكبيرة.....0,5 نقاط

الدمج عند مسافة صغيرة يدل على تشابه قوي جدًا. والعناصر متجانسة بدرجة عالية . اما الدمج عند مسافة كبيرة فيدل على اختلاف واضح .وقد يؤدي إلى انخفاض التجانس الداخلي .

10- لماذا يعد أسلوبًا غير مراقب؟.....01.نقاط

التصنيف الهرمي التصاعدي لا يعتمد على متغير تابع أو فئات معروفة مسبقًا، بل يكتشف البنية الداخلية للبيانات اعتمادًا فقط على التشابه أو التباعد بين العناصر.

11- طرق التأكد من جودة التصنيف.....01 نقاط

- التمثيل البياني بالمركبات الأساسية(ACP) : يسمح برؤية مدى انفصال المجموعات بصريًا.
- مقارنة عدة طرق ربط مثل : Single Linkage و Complete Linkage و Ward
- مؤشرات التجانس الداخلي والتباعد الخارجي :للتحقق من صلابة التصنيف.