

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



المحور الأول: مفاهيم أساسية حول البيانات وطرق تحليلها



الهدف من المحور:



- ✓ التعريف بتحليل البيانات متعددة المتغيرات
- ✓ التعرف على البيانات، الافراد، المتغيرات وأنواعها
- ✓ التعرف على مستويات القياس
- ✓ التعرف على الجداول وأنواعها
- ✓ مراجعة لأهم المفاهيم (المقاييس) الاحصائية

1- تعريف تحليل البيانات (متعددة المتغيرات)

تمهيد:

« تحليل البيانات يتألف من مجموعة من الأساليب التي يمكن استخدامها عند إجراء عدة قياسات على كل مفردة (أو كائن) في عينة واحدة أو أكثر..... »

« تحليل البيانات عبارة عن مجموعة من التقنيات لاكتشاف البنية المعقدة لجدول أرقام متعدد المتغيرات وتحويلها إلى بنية أبسط تلخصها بشكل أفضل يمكن تمثيل هذه البنية في أغلب الأحيان بيانياً »

1- تعريف تحليل البيانات (متعددة المتغيرات)

يوجد مجموعتان رئيسيتان من طرق (أساليب) تحليل البيانات:

أولاً: الأساليب العاملة (طرق تحليل العوامل):

✓ تعتمد على البحث عن العوامل (أو المحاور) الرئيسية لمجموعة البيانات

✓ تهدف إلى تقليص العدد الكبير من المتغيرات إلى عدد أقل من العوامل

ثانياً: الأساليب التصنيفية (طرق أو تقنيات التصنيف):

✓ تقنيات تُستخدم لتصنيف البيانات إلى فئات أو مجموعات بناءً على مجموعة

من الخصائص أو المتغيرات المتعددة

✓ تهدف إلى تقسيم البيانات إلى مجموعات متجانسة، مما يسهل فهم البنية

الكامنة للبيانات واتخاذ القرارات بناءً عليها.

2- مفاهيم حول البيانات، الأفراد والمتغيرات

1-2 البيانات: " Data "

- ✓ هي حقائق أولية عشوائية غير منظمة وهي غير مفيدة
- ✓ يمكن أن تكون على هيئة أرقام، حروف، كلمات، أو حتى صور
- ✓ يتم جمعها من خلال تجارب، أو ملاحظات، أو استبيانات أو مقابلات.... الخ

2-2 الأفراد: = المشاهدات = الوحدات الإحصائية = الوحدات التجريبية

- ✓ يعتبر الكيان الأساسي الذي تتمحور حوله الدراسة (أو البحث)
- ✓ يقوم الباحث بإجراء عدد معين من القياسات عليه
- ✓ يمكن أن يكون جميع الأفراد عبارة عن عينة أخذت من مجتمع ما أو جميع أفراد المجتمع

2- مفاهيم حول البيانات، الأفراد والمتغيرات

2-3 المتغير: " Variable "

- ✓ هو الصفة (أو الخاصية) التي يتمتع بها الأفراد، قابل للملاحظة والقياس
- ✓ يتغير من وحدة إحصائية إلى أخرى
- ✓ أو يتغير في نفس الوحدة الإحصائية إذا تم قياسه أو ملاحظته في أوقات زمنية مختلفة
- ✓ يمكن أن تكون الخصائص (المتغير) نوعية مثل: لون الشعر، الجنس، الديانة.... الخ
- ✓ يمكن أن تكون الخصائص كمية مثل: الطول، الوزن، درجة الحرارة.... الخ

3- مستويات قياس المتغيرات

تنقسم عادة مستويات القياس إلى أربعة مستويات:

1-3 المستوى الاسمي (Nominal) :

- ✓ الأرقام لا تدل على معنى كمي، وإنما تستخدم كبديل عن الأسماء
- ✓ لا يمكن استخدام العمليات الحسابية في هذا المستوى
- ✓ أفضل مقياس احصائي هو المنوال

3- مستويات قياس المتغيرات

2-3 المستوى الترتيبي (Ordinal) :

- ✓ الأرقام في هذا المستوى من القياس لها معنى يفيد الترتيب
- ✓ المسافات بين الرتب (الأرقام) لا تكون متساوية
- ✓ أفضل مقياس احصائي هو الوسيط
- ✓ لا نستطيع القيام بالعمليات الحسابية إلا الترتيب كأكبر من أو أصغر من

3- مستويات قياس المتغيرات

3-3 المستوى الفئوي (Interval) :

- ✓ تمتلك الأرقام في المستوى الفئوي معنى كمي
- ✓ يمكن ترتيبها، وتكون المسافة (أو الفروق) بين الأرقام متساوية.
- ✓ ما يميز هذا المستوى أن الصفر فيه ليس صفراً حقيقياً (ليس مطلقاً)
- ✓ أفضل مقياس احصائي المتوسط الحسابي. لا يمكن استخدام النسب.

3- مستويات قياس المتغيرات

3-3 المستوى النسبي (Ratio) :

- ✓ يمكن ترتيب القيم (رتبياً). والفروق (المسافة) بين القيم متساوية.
- ✓ يحتوي هذا المستوى من القياس على نقطة صفر مطلقة
- ✓ يسمح باستخدام العمليات الرياضية كافة
- ✓ إمكانية استخدام النسب

4- الجداول وأنواعها:

1-4 تعريف الجداول:

«الجدول هو طريقة لعرض البيانات بشكل منظم باستخدام صفوف وأعمدة لتسهيل فهم البيانات ومقارنتها.»

«الجدول الإحصائي هو أداة تستخدم لتلخيص البيانات الرقمية أو النصية وتصنيفها لعرضها بشكل واضح. كما يساعدنا في تحليل الأنماط والعلاقات بين المتغيرات.»

«الجدول هو تمثيل مرئي للبيانات باستخدام صفوف وأعمدة، حيث يتم تنظيمها بشكل منطقي لتسهيل التحليل والمقارنة. يمكن أن يحتوي الجدول على أرقام أو رموز»

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جداول القياسات (الجداول الكمية):

- ✓ هي جداول شائعة الاستعمال
- ✓ أعمدها عبارة عن متغيرات (خصائص) ذات قيم عددية متصلة
- ✓ وصفوفها عبارة عن أفراد أو ملاحظات

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جدول القياسات (الجدول الكمي):

		الخصائص (المتغيرات)					
		العمر X^1	الدخل الخاضع للضريبة x^2	الراتب الإجمالي X^i	الأقدمية X^p
الأفراد	1	X^1_1	x^2		X^i		X^{p_1}
	2	X^1_2	x^2		X^i_2		X^{p_2}
	i	X^1_i	x^2_i		X^i_i		X^{p_i}
	n	X^1_n	x^2_n		X^i_n		X^{p_n}

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جداول الترميز

1. جدول الترميز المتصل (مكثف)

- ✓ عبارة عن جداول من المتغيرات النوعية (الاسمية أو الترتيبية)
- ✓ صفوفها عبارة عن أفراد أو ملاحظات
- ✓ أعمدها عبارة عن متغيرات (خصائص) نوعية لديها فئات (مستويات)
- ✓ القيم العددية عبارة عن رموز لفئات المتغيرات النوعية

4- الجداول وأنواعها:

2-4 أنواع الجداول:

لنأخذ مثال حول أسئلة الاستبيان التالي:

- الجنس؟
1. ذكر
2. أنثى
- المستوى الوظيفي؟
1. عامل/موظف
2. مشرف
3. مدير تنفيذي
- الحالة الاجتماعية؟
1. أعزب
2. متزوج
3. أرمل
4. مطلق

➤ جدول الترميز المتصل (المكثف)

		الخصائص (المتغيرات)		
		الجنس	المستوى الهرمي	الحالة الاجتماعية
الأفراد	1	2	1	2
	2	2	1	2
	3	1	1	3
	4	1	3	1
	5	1	1	4
	6	2	2	2

4- الجداول وأنواعها:

2-4 أنواع الجداول:

2. جدول الترميز المنفصل بالكامل

- ✓ نفس خصائص النوع السابق من الجداول
- ✓ تُفصل كل فئة من فئات المتغيرات (الخصائص)، أي يصبح لها عمود خاص بها.
- ✓ القيم العددية عبارة عن رقمين ثنائيين 1 يدل على وجود الخاصية و 0 عدم وجودها
- ✓ يسمى أيضا الجداول الثنائية.

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جدول الترميز المنفصل بالكامل

		الخصائص (المتغيرات)								
		الجنس		المستوى الهرمي			الحالة الاجتماعية			
		ذكر	أنثى	عامل، موظف	مشرف	مدير تنفيذي	عازب	متزوج	أرمل	مطلق
الأفراد	1	1	0	0	1	0	0	1	0	0
	2	0	1	1	0	0	0	1	0	0
	...									
	i	1	0	0	0	1	1	0	0	0
	...									
	n	0	1	1	0	0	0	0	1	0

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جداول التوفيقات Tableaux de contingence

- ✓ هي جداول إحصائية تُستخدم لعرض وتلخيص العلاقة بين متغيرين نوعيين
- ✓ الأعمدة عبارة عن فئات المتغير النوعي الأول.
- ✓ الصفوف عبارة عن فئات المتغير النوعي الثاني
- ✓ يحتوي جدول على تكرارات (تعدادات) التي تربط بين فئات المتغيرين.

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جداول التوفيقات Tableaux de contingence

		الحالة الاجتماعية			الإجمالي
		أعزب	متزوج	مطلق	
الجنس	ذكر	30	50	10	90
	أنثى	40	60	20	120
الإجمالي		70	110	30	210

4- الجداول وأنواعها:

2-4 أنواع الجداول:

➤ جداول التوفيقات Tableaux de contingence

		لون الشعر				الإجمالي
		بني	كستنائي	أحمر	أشقر	
لون العيون	عسلي	68	119	26	7	220
	بندقي	15	54	14	10	93
	أخضر	5	29	14	16	64
	أزرق	20	84	17	94	215
الإجمالي		108	286	71	127	592

5- مراجعة مفاهيم احصائية:

1-5 المتوسط الحسابي:

- ✓ من أهم المقاييس النزعة المركزية
- ✓ المتوسط الحسابي هو القيمة التي تمثل "مركز" مجموعة من الأرقام.
- ✓ يُحسب عن طريق جمع جميع القيم ثم قسمة الناتج على عددها وفق القانون:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n xi = \frac{\sum_{i=1}^n xi}{n}$$

5- مراجعة مفاهيم احصائية:

1-5 المتوسط الحسابي:

✓ مثال: يحتوي الجدول التالي على نقاط طلبة في مادة الرياضيات. احسب المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^n x}{n} = \frac{52}{5} = 10,4$$

الطلبة	نقطة الرياضيات
1	12
2	11
3	9
4	8
5	12
المجموع	52

5- مراجعة مفاهيم احصائية:

2-5 التباين والانحراف المعياري:

✓ من أهم مقاييس التشتت

✓ يقيس التباين والانحراف المعياري مدى تشتت القيم حول المتوسط الحسابي

✓ التباين يعبر عن تشتت البيانات بوحدة مربعة (مثل متر²)

✓ الانحراف المعياري: يعبر عن تشتت البيانات بنفس وحدة القيم الأصلية (مثل متر)

$$V = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$$

$$\delta = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}} = \sqrt{V}$$

5- مراجعة مفاهيم احصائية:

أجور الشركة A	أجور الشركة B
3100	1800
2500	2000
2800	1900
3200	4500
4000	6000
2500	5000
3000	1600
2700	2400
300	2500
2900	2000

2-5 التباين والانحراف المعياري:

مثال: لديك جدول لأجور موظفين وعمال في شركتين A و B. المطلوب:

1- احسب المتوسط الحسابي لكلا الشركتين. ماذا تستنتج؟

2- احسب التباين والانحراف المعياري

أجور الشركتين بمتوسط حسابي متساوي 2970

$$V_A = 168100, \delta = 410$$

$$V_B = 2246100, \delta = 1498.70$$

5- مراجعة مفاهيم احصائية:

2-5 التباين المشترك:

- ✓ التباين المشترك هو مقياس إحصائي يوضح مدى تغير متغيرين معًا
- ✓ إذا كان التباين المشترك موجبًا، فإن المتغيرين يميلان إلى التحرك في نفس الاتجاه
- ✓ إذا كان التباين المشترك سالبًا، فإن المتغيرين يميلان إلى التحرك في اتجاهين متعاكسين
- ✓ إذا كان التباين المشترك صفرًا، فلا يوجد علاقة خطية بين المتغيرين

$$COV(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{n}$$

5- مراجعة مفاهيم احصائية:

2-5 معامل الارتباط:

- ✓ الارتباط هو مقياس إحصائي يوضح قوة واتجاه العلاقة الخطية بين متغيرين
- ✓ إذا كان التباين المشترك موجبًا، فإن المتغيرين يميلان إلى التحرك في نفس الاتجاه
- ✓ إذا كان التباين المشترك سالبًا، فإن المتغيرين يميلان إلى التحرك في اتجاهين متعاكسين
- ✓ إذا كان التباين المشترك صفرًا، فلا يوجد علاقة خطية بين المتغيرين

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{Y})^2}}$$

$$r(x; y) = \frac{COV(x, y)}{\delta_x \delta_y}$$

5- مراجعة مفاهيم احصائية:

2-5 معامل الارتباط:

✓ تمرينين تطبيقيين:

الطلاب	(X) ساعات الدراسة	(Y) الدرجة
1	2	50
2	4	60
3	6	70
4	8	80
5	10	90

اليوم	(X) درجة الحرارة	(Y) مبيعات الآيس كريم
1	25	100
2	28	120
3	30	150
4	22	90
5	35	200