

مفاهيم أساسية حول علم الإحصاء

أولا - الإحصاء الوصفي Statistique descriptive: وهو ذلك الفرع من الإحصاء الذي يهتم بتعميم

النتائج بعد جمع جزء منها، وترتيبها وتلخيصها وعرضها في صورة جداول ثم تمثيلها بيانيا وتحليلها وتفسيرها لاتخاذ القرارات المناسبة.

أما الإحصاء الاستدلالي فهو يختص بإجراء التنبؤات والتقديرات والاستنتاجات.

ثانيا - مصادر جمع البيانات الإحصائية: يوجد مصدران لجمع البيانات هما:

1- مصدر تاريخي: وهو ما يؤخذ من السجلات المحفوظة.

مثل: - سجلات المواليد والوفيات.

- البيانات الواردة من رسائل الماستر والمجستير والدكتوراه.

- البيانات التي يتم نشرها من خلال المنظمات العالمية.

2- مصدر ميداني: ويتم فيه جمع البيانات مباشرة عن طريق إتصال الباحث بالوحدة محل الدراسة، ويتم جمع البيانات ميدانيا عن طريق الوسائل التالية:

أ- المقابلات الشخصية: يستطيع الباحث بهذه الطريقة تحقيق أعلى درجات الدقة في جمع البيانات. إلا أن هذه الطريقة

وبالرغم مما تمتاز به من دقة المعلومات قد تكون مكلفة. وخاصة في حالة العينات كبيرة الحجم.

ب- الاستمارة الاحصائية (الاستبيان): ويراعى في هذه الاستمارة الشروط التالية:

- أن تكون الأسئلة واضحة وسهلة.

- أن تكون الاستمارة قصيرة.

- يجب التأكيد على سرية البيانات للشخص محل الدراسة حتى لا تكون إجابته بعيدة عن الواقع.

- يجب أن تحقق الاستمارة الأهداف محل الدراسة.

ثالثا - التعريف بالمصطلحات والمفاهيم الإحصائية العامة:

1 - المجتمع Population: وهو المجموعة الكاملة من الافراد أو العناصر محل الدراسة والتي لها خصائص مشتركة.

وكل شخص من المجتمع يدعى وحدة إحصائية (Unité Statistique) وهو العنصر الأساسي عند القيام بتجربة ما

مثال: لو كانت دراستنا حول: أطول طلبة المركز الجامعي ميلة.

في هذه الحالة المجتمع الإحصائي هو جميع طلبة المركز.

- قد يكون المجتمع الإحصائي محدودا حيث يمكن حصر عدد أفرادهِ. مثل: عدد الطلاب في الكلية.

أو غير محدود مثل: عدد النجوم في السماء، عدد حبات القمح المحصودة في مزرعة ما.

2 - العينة Echantillon: وهي جزء من المجتمع التي يتم اختيارها وتجري عليها القياسات. وبعد استخلاص النتائج نعمم

على المجتمع الذي سحبت منه.

مثال: في المثال السابق إذا أخذنا طلبة معهد العلوم الاقتصادية. فيمكن اعتبار هذه المجموعة عينة من المجتمع الذي هو محل

الدراسة.

ملاحظة: لماذا يتم اختيار العينة؟؟

يكون أخذ العينة أفضل من دراسة المجتمع كله للأسباب التالية: أقل وقت، أقل تكلفة، الحسابات تكون أدق، إمكانية الاتصال

بكل مفردات العينة (عمال مثلا).

3 - الوحدة الإحصائية Unité Statistique: وهي الكائن الواحد أو الوحدة الأساسية التي يتكون منها المجتمع الإحصائي.

مثال: في المثال السابق: أطوال طلبة المركز الجامعي ميلة.

الوحدة الإحصائية هي الطالب الواحد

4 - الصفة أو الخاصية Caractère Statistique: وهي الخاصية التي تميز الفرد في المجتمع

(أو القاسم المشترك لكل وحدات المجتمع المدروس).

مثال: ماهي الخاصية المدروسة في المثال السابق؟

الخاصية المدروسة هي الطول.

رابعا - أنواع البيانات (المتغيرات) الإحصائية

تنقسم المتغيرات الإحصائية إلى نوعين:

1 - متغيرات وصفية (نوعية) Variable qualitative: وهي البيانات التي تكون لهُل صفات معينة، لا يمكن قياسها ولا

يعبر عنها بصورة عددية. وتنقسم المتغيرات النوعية إلى قسمين هما:

أ - متغيرات وصفية إسمية qualitative nominale: وهي بيانات غير رقمية لا تتأثر بأي ترتيب منطقي.

أمثلة: - الجنسية (جزائري، تونسي، ...) وصفي إسمي

- الجنس (ذكر، أنثى) وصفي إسمي ذو فرعين

- اللون (أبيض، أسود،...)

- الحالة العائلية (متزوج، أعزب، أرمل، مطلق)

ب - متغيرات وصفية ترتيبية qualitative ordinale : وهي بيانات غير رقمية، تتكون من مستويات تتبع ترتيبيا منطقيا معد مسبقا أو متفق عليه.

أمثلة: - المستوى التعليمي (ابتدائي، متوسط، ثانوي، جامعي)

- مقياس اللباس (s.m.l.xl) وصفي ترتبي

2 - متغيرات كمية (عددية) Variable quantitatives: وهي بيانات رقمية، يمكن التعبير عنها في صورة عددية. مثل:

الطول، الوزن، عدد الأفراد في الأسرة... الخ.

وتنقسم بدورها إلى قسمين هما:

أ - متغيرات كمية مستمرة (متصلة) quantitative Continue: وهي بيانات يمكن أن تأخذ أي قيمة عددية في مدى معين (تأخذ قيما غير محدودة وغير منتهية).

أمثلة عن المتغيرات الكمية المستمرة: الطول، الوزن، السرعة، كميات الأمطار... الخ.

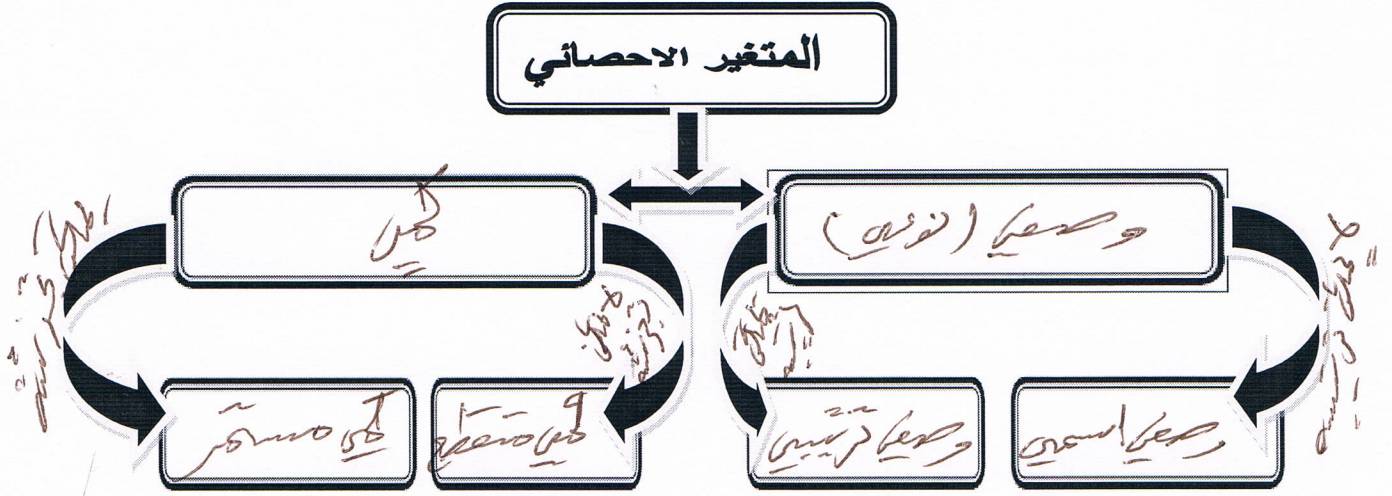
فإذا كان x هو متغير الطول فإن x يمكن أن يأخذ القيم 15 متر، 11.3 متر، 14.75 متر، أي أن المتغير x يمكن أن يأخذ أي قيمة في فترة زمنية معينة.

ب - متغيرات كمية متقطعة (منفصلة) quantitative discret: وهي البيانات التي تأخذ أعداد صحيحة، لا يمكن تجزئتها، وهي محدودة يمكن عدّها وحصرها بالعد.

أمثلة عن المتغيرات الكمية المتقطعة: عدد الأفراد في الأسرة، عدد الغرف في المسكن، عدد الأهداف المسجلة، عدد السيارات المباعة... الخ.

فمثلا إذا كان x متغير يمثل عدد أفراد الأسرة، فإنه يمكن أن يأخذ القيم 1، 2، 3، 4، 5 ... ولا يمكن أن يأخذ x القيم 1.5، 3.25، 5.17.

مخطط يبين أنواع البيانات أو المتغيرات



تمرين: حدد كلا من المجتمع الاحصائي، الوحدة الاحصائية، الخاصية المدروسة ونوعيتها بدقة لكل من:

- 1- الأجور السنوية لأساتذة الجامعة.
- 2- عدد الأولاد في كل عائلة لأحد الأحياء السكنية.
- 3- كمية الأمطار المتساقطة خلال فصول السنة في ولاية ميلة.
- 4- جنسية مجموعة من السياح.
- 5- وزن اللاعبين.
- 6- توزيع عينة من 30 فرد حسب المستوى التعليمي.

المحور الأول

تنظيم وعرض البيانات

تمهيد:

يعتبر تنظيم وعرض البيانات الاحصائية أول مرحلة للتحليل الاحصائي، وتتقيد طريقة تنظيم وعرض البيانات على نوع هذه البيانات سواء كانت وصفية (نوعية) أو كمية. وعليه يمكن تنظيم وعرض البيانات إما عن طريق تصميم التوزيعات أو الجداول التكرارية أو باستعمال الرسوم البيانية. والهدف من كل ذلك تقديم البيانات بطريقة مبسطة ومختصرة.

أولاً- التوزيعات والجداول التكرارية:

يعتبر تبويب (تفرغ) وعرض البيانات في جداول من الخطوات الأساسية للحصول على المعلومات. الهدف من استخدام الجداول هو تلخيص البيانات وتسهيل فهمها ودراساتها، لذا يجب أن يكون الجدول المسمى بجدول التوزيعات التكرارية بسيط وغير غامض بقدر الإمكان.

- حسب نوعية المتغير (وصفي أو كمي) يمكن تصميم الجداول التالية:

1- الجداول التكرارية للبيانات الوصفية (النوعية): وهو عبارة عن جدول يحتوي في صورته البسيطة على العناصر التالية:

أ- نوع المتغير: يتم رصد كل صفات المتغير في العمود الأول.

ب- التكرار المطلق: وهو يمثل عدد المرات التي يتكرر فيها نفس المتغير ونرمز له بالرمز n_i ويكون في العمود الثاني.

- مع مراعاة مجموع الأعداد الموجودة في عمود التكرارية يساوي عدد المفردات أو مجموع التكرارات

$$\sum n_i = n_1 + n_2 + n_3 + \dots + n_K$$

المتغير	التكرار المطلق n_i
الصفة 1	n_1
الصفة 2	n_2
.	.
.	.
.	.
الصفة 3	n_K
المجموع Σ	Σn_i

ج- التوزيع التكراري النسبي والمئوي:

- التكرار النسبي: ونرمز له بالرمز f_i

وهو حاصل قسمة تكرار كل صفة أو فئة على مجموع التكرارات أي $f_i = \frac{n_i}{\sum n_i}$

ملاحظة: دائماً مجموع التكرار النسبي يساوي 1 $\sum f_i = 1$

- التكرار النسبي المئوي: ونرمز له بالرمز $f_{i\%}$

$$f_{i\%} = f_i \times 100$$

أو

$$f_{i\%} = \frac{n_i}{\sum n_i} \times 100$$

– التوزيع التكراري التجميعي الصاعد: ونرمز له بالرمز $N_i \nearrow$ وبحسب كالاتي:

$$N_1 \nearrow = n_1$$

$$N_2 \nearrow = n_1 + n_2 \quad \gg \quad N_2 \nearrow = N_1 \nearrow + n_2$$

$$N_k \nearrow = \sum n_i$$

ملاحظة: التكرار التجميعي الصاعد الأول دائما يساوي التكرار المطلق الأول $N_1 \nearrow = n_1$ والتكرار التجميعي الصاعد الأخير دائما يساوي مجموع التكرارات $N_k \nearrow = \sum n_i$

– التوزيع التكراري التجميعي النسبي الصاعد: ونرمز له بالرمز $F_i \nearrow$

$$F_i \nearrow = \frac{n_i}{\sum n_i} \quad \text{أو} \quad \text{بحسب بنفس الطريقة}$$

– التوزيع التكراري التجميعي النسبي المتوي الصاعد: ونرمز له بالرمز $F_{i\%} \nearrow$

$$F_{i\%} \nearrow = F_i \nearrow \times 100$$

– التوزيع التكراري التجميعي النازل: ونرمز له بالرمز $N_i \nwarrow$ وبحسب كالاتي:

$$N_1 \nwarrow = \sum n_i$$

$$N_2 \nwarrow = \sum n_i - n_1 \quad \gg \quad N_2 \nwarrow = N_1 \nwarrow - n_1$$

.

$$N_k \nwarrow = n_k$$

ملاحظة: التكرار التجميعي النازل الأول دائما يساوي مجموع التكرارات $N_1 \nwarrow = \sum n_i$ والتكرار التجميعي النازل الأخير $N_k \nwarrow = n_k$

– التوزيع التكراري التجميعي النسبي النازل: ونرمز له بالرمز $F_i \nwarrow$

– التوزيع التكراري التجميعي النسبي المتوي النازل: ونرمز له بالرمز $F_{i\%} \nwarrow$

$$F_{i\%} \nwarrow = F_i \nwarrow \times 100 \quad \text{وبحسب كالاتي:}$$

تمرين: سئل 20 شاب عن رأيهم في أداء مقدم البرنامج التلفزيوني الجديد وكان على كل شاب أن يختار واحد من خمسة إجابات وهي – ضعيف، – تحت لمتوسط، – متوسط، – فوق المتوسط، – ممتاز. فكانت الإجابات كالاتي:

ممتاز، تحت المتوسط، متوسط، فوق المتوسط، ممتاز، ضعيف، فوق المتوسط، ممتاز، تحت المتوسط، فوق المتوسط، ممتاز، فوق المتوسط، تحت المتوسط، فوق المتوسط، متوسط، متوسط، تحت المتوسط، ضعيف، ممتاز، فوق المتوسط. المطلوب: – أعرض البيانات في شكل جدول تكراري.

- كون التكرار النسبي الصاعد والنازل.
- كون التكرار التجميعي الصاعد والنازل.
- كون التكرار التجميعي النسبي الصاعد والنازل.
- حساب النسبة المئوية للشباب الذي كان رأيهم في أن الأداء كان أعلى من درجة متوسط (أي فوق المتوسط + ممتاز).

الحل:

المتغير	التكرار المطلق n_i	التكرار النسبي f_i	التكرار النسبي المئوي $f_{i\%}$	التكرار التجميعي الصاعد N_i	التكرار التجميعي النازل N_i	التكرار التجميعي النسبي الصاعد F_i
ضعيف	2	0,10	10	2	20	0,10
تحت المتوسط	4	0,20	20	6	18	0,30
متوسط	3	0,15	15	9	14	0,45
فوق المتوسط	6	0,30	30	15	11	0,75
ممتاز	5	0,25	25	20	5	1
المجموع	20	1	100	—	—	—

$$f_i = \frac{n_i}{\sum n_i}$$

$$f_1 = \frac{2}{20} = 0,10$$

$$f_{i\%} = f_i \times 100$$

$$f_{i\%} = 0,10 \times 100 = 10$$

- النسبة المئوية للشباب الذي كان رأيهم في أن الأداء كان أعلى من درجة متوسط هي 55% = 25+30

2- الجداول التكرارية للبيانات الكمية:

أ- الجدول التكراري للبيانات الكمية المتقطعة (المنفصلة): يكون الشكل العام للجدول التكراري في حالة البيانات الكمية المتقطعة كالتالي:

- في العمود الأول: نضع قيم المتغير X بصورة فردية ومرتبطة ترتيبا تصاعديا.
- العمود الثاني: يمثل عدد المرات التي تتكرر فيها نفس القيمة.

قيم المتغير	التكرار المطلق n_i
X_1	n_1
X_2	n_2
.	.
.	.
.	.
X_K	n_K
المجموع Σ	Σn_i

مثال: ليكن لديك السلسلة التالية والتي تمثل عدد الأطفال في كل أسرة في حي يتكون من 20 أسرة.

4	4	3	4	3	4	3	6	4	5
6	1	1	4	2	3	2	4	2	8

المطلوب: عرض البيانات في شكل جدول تكراري.

الحل: 1- الترتيب: 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 6, 6, 8

2- الجدول التكراري:

قيم المتغير x	ni
1	2
2	3
3	4
4	7
5	1
6	2
8	1
Σ	20

من الملاحظ أنه بمجرد أن توضع البيانات في جدول تكراري يصبح من السهل ملاحظة الوتيرة التي تظهر بها قيم المتغير (عدد الأطفال). فمثلاً: من السهل تحديد عدد الأسر التي بها عدد الأطفال يساوي 8 وهي أسرة واحدة.

ب- الجدول التكراري للبيانات الكمية المستمرة (المتصلة): حين تشتمل البيانات على عدد كبير من القيم يفضل تجميعها في فئات حتى يسهل عرضها بصورة واضحة. ويراعى هنا ألا يكون عدد الفئات كبير فتضيع الفائدة من عملية التجميع. وألا يكون عدد الفئات صغير فتضيع معالم التوزيع ويفقد الكثير من تفاصيله.

ولتصميم جدول تكراري يحتوي على عدد مناسب من الفئات يمكن اتباع الخطوات التالية:

1- حساب المدى: ونرمز له بالرمز E

المدى هو الفرق بين أكبر قيمة وأصغر قيمة في التوزيع.

$$E = X_{\text{Max}} - X_{\text{Min}}$$

2- تحديد عدد الفئات: ونرمز لها بالرمز NC

- غالباً ما يتراوح عدد الفئات بين 5 إلى 20 فئة، ويعتمد هذا على عدد المفردات في العينة أو المجتمع.

- ما يلاحظ أنه لا يوجد قاعدة نظرية محددة لتحديد عدد الفئات، إلا أنه يستحسن الاستعانة ب:

الطريقة 1: قاعدة sturges

$$NC = 1 + 3.33 \log (N)$$

حيث N تمثل عدد مفردات العينة أو المجتمع.

الطريقة 2: قاعدة yule

$$NC = 2.5 \times \sqrt[4]{N}$$

3- حساب طول الفئة: ونرمز له بالرمز L

ويحسب كالتالي:

$$L = \frac{E}{NC} \quad \text{المدى} \quad \text{عدد الفئات}$$

مع مراعاة تقريب الناتج بالزيادة إلى أقرب عدد صحيح مناسب.

4- تحديد الفئة الأولى: نختار أصغر قراءة في البيانات وهي لتكون بداية الفئة الأولى، ويضاف إليها طول الفئة فتحصل على نهاية الفئة الأولى وبداية الفئة الثانية وهكذا...

- مراكز الفئات (منتصف الفئات): ونرمز له بالرمز X_i ويحسب كالآتي:

$$X_i = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2}$$

ملاحظة: عند تفرغ البيانات فإنه يجب أن تنتمي كل مفردة إلى فئة واحدة فقط.

تمرين: سحبت عينة من 30 مزرعة للتعرف على مردوديتها من القمح (بالطن) خلال موسم فكانت النتائج كالتالي:

29	14	20	20	17	25	20	14	12	16	17	16	12	15	20
12	20	15	14	25	20	17	15	20	14	15	12	16	14	20

المطلوب:- حدد المجتمع الاحصائي، الوحدة الاحصائية، الخاصية المدروسة ونوعيتها.

- أعرض البيانات في جدول تكراري؟

- حدد كلا من f_i , $f_i\%$, N_i

- ماهي نسبة المزارع التي تفوق مردوديتها من القمح أو يساوي 15 (طن) ويقل تماما عن 24 (طن)؟

الحل:

المجتمع الاحصائي	الوحدة الاحصائية	الخاصية المدروسة	نوعيتها
المزارع	المزرعة	مردودية القمح	كمي مستمر

عرض البيانات في جدول تكراري:

- حساب المدى: $E = X_{\text{Max}} - X_{\text{Min}}$

$$E = 29 - 12 = 17$$

- تحديد عدد الفئات: NC

نستخدم قاعدة Sturges

$$NC = 1 + 3.33 \log(N)$$

$$NC = 1 + 3.33 \log(30) = 5.87 \approx 6 \text{ فئات}$$

- حساب طول الفئة: L

$$L = \frac{E}{Nc} = \frac{17}{6} = 2,83 \approx 3$$

أي أن الفئة الأولى هي:

[12 , 15 [

C	n_i	f_i	$F_i\%$	N_i ↗
[12 , 15 [	9			
[15 , 18 [	10			
[18 , 21 [	8			
[21 , 24 [	0			
[24 , 27 [	2			
[27 , 30 [	1			
Σ	30			

ثانياً- التمثيل البياني لتوزيعات التكرارية:

الأشكال الآتية تمثل أهم طرق تمثيل البيانات:

1- شكل الأعمدة:

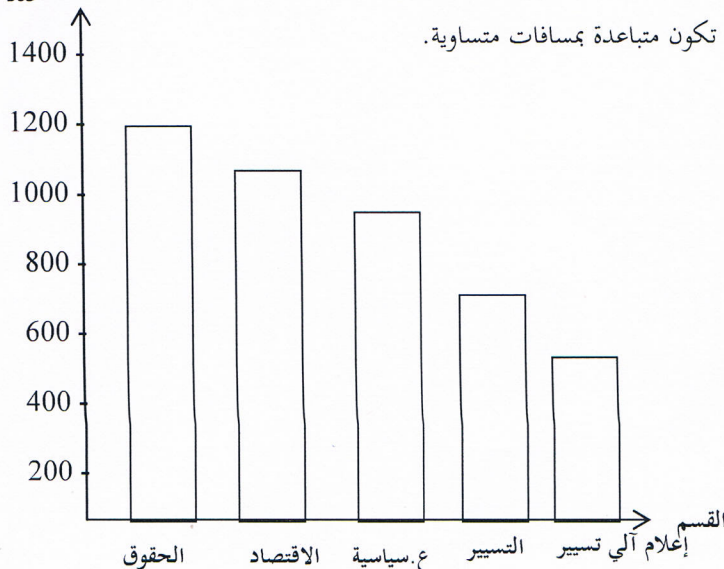
أ- الأعمدة المستطيلة والأعمدة البسيطة: وهو الرسم البياني الملائم لتوزيع متغير وصفي (نوعي) أو متغير كمي متقطع (منفصل).

مثال 01: إليك الجدول التكراري التالي مثل بيانيا هذا المتغير بواسطة أعمدة مستطيلة.

القسم	الحقوق	الاقتصاد	التسيير	علوم سياسية	إعلام آلي تسيير	المجموع
عدد الطلبة	1200	1000	800	600	400	4000

- الأعمدة المستطيلة: وهي عبارة عن مجموعة من الأعمدة ذات القواعد المتساوية إلا أن ارتفاعها يتناسب مع تكرار كل خاصية، كما أن هذه الأعمدة تكون متباعدة بمسافات متساوية.

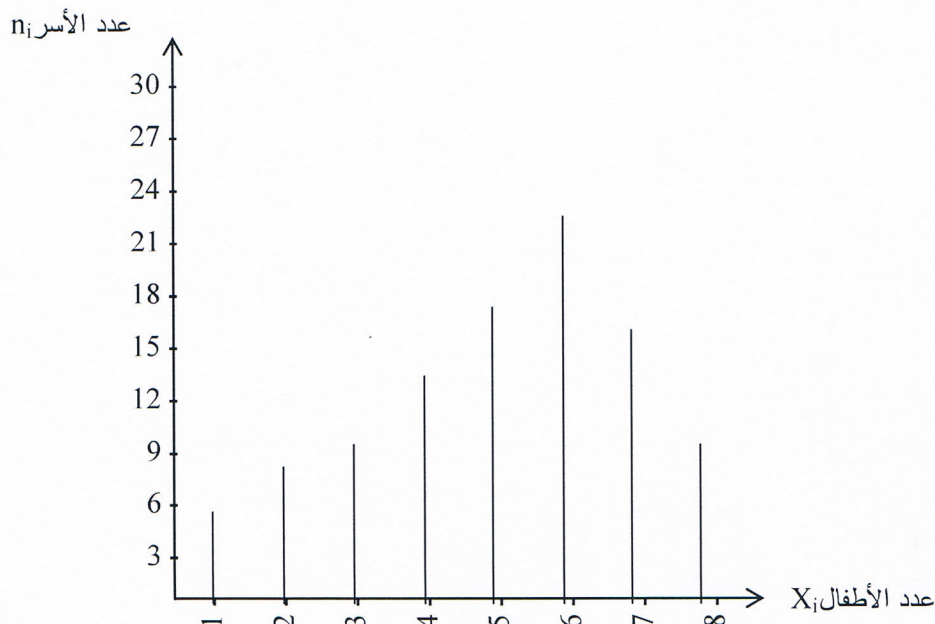
الحل:



مثال 02: يبين الجدول التالي عدد الأطفال في كل أسرة، في حي يتكون من 110 أسرة، المطلوب عرض هذه البيانات بواسطة أعمدة بسيطة.

عدد الأسر	عدد الأطفال في كل أسرة
6	1
9	2
10	3
14	4
18	5
25	6
17	7
11	8
110	المجموع

الحل:



نلاحظ بسهولة أن العمود الذي يقابل القيمة 6 هو أطول الأعمدة وتكراره يساوي 25 ويعني ذلك أن أغلب العائلات لها 6 أطفال.

ب- العمود المجزأ: (عادة ما يستخدم في توزيع متغير وصفي ترتيبي)، وهو عبارة عن مستطيل مقسم إلى عدة أجزاء كل جزء يقابل تكرار معين للخاصية المدروسة.

- لرسم العمود المجزأ نستعمل النسب المئوية المقابلة لكل تكرار حيث طول المستطيل يساوي 100%.

2- شكل الدائرة: وهو الرسم البياني الملائم لتوزيع متغير وصفي (نوعي) أو متغير كمي متقطع (منفصل).

وهو عبارة عن قرص كامل يتم تقسيمه إلى أجزاء. ونتحصل على مقدار الزاوية المطلوب لكل جزء وذلك بضرب التكرار

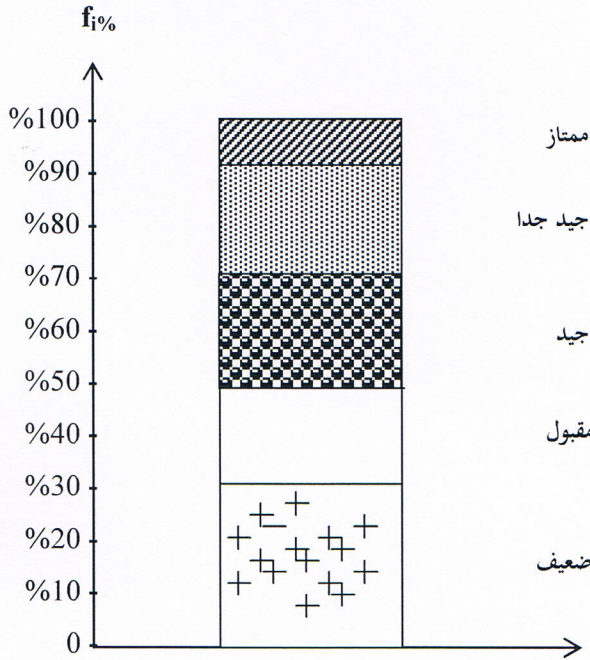
$$\alpha_i^\circ = f_i \times 360$$

النسبي في 360.

مثال 03: أعرّض البيانات التالية الخاصة بتقديرات 20 طالب في مادة الاحصاء بواسطة العمود المجزأ.

التقديرات X_i	ضعيف	مقبول	جيد	جيد جدا	ممتاز	المجموع
عدد الطلبة	6	4	4	5	1	20

الحل:



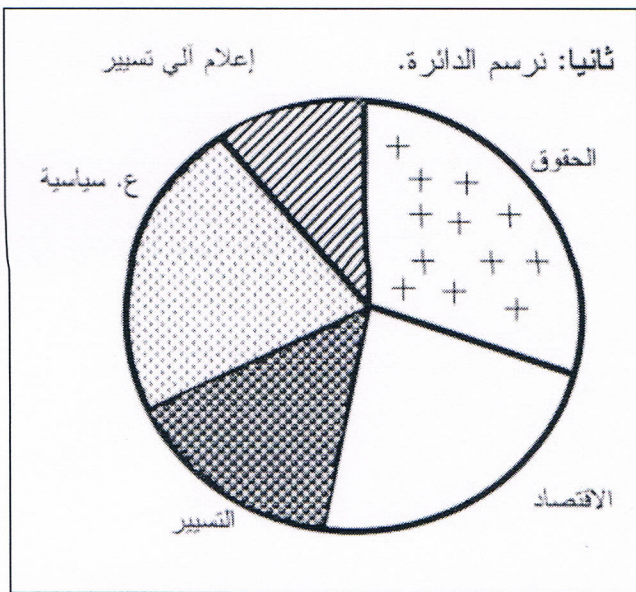
$f_i\%$	n_i	X_i
30	6	ضعيف
20	4	مقبول
20	4	جيد
25	5	جيد جدا
5	1	ممتاز
%100	20	المجموع

مثال 04: مثل بيانات معطيات المتغير في المثال 01 بواسطة دائرة.

الحل:

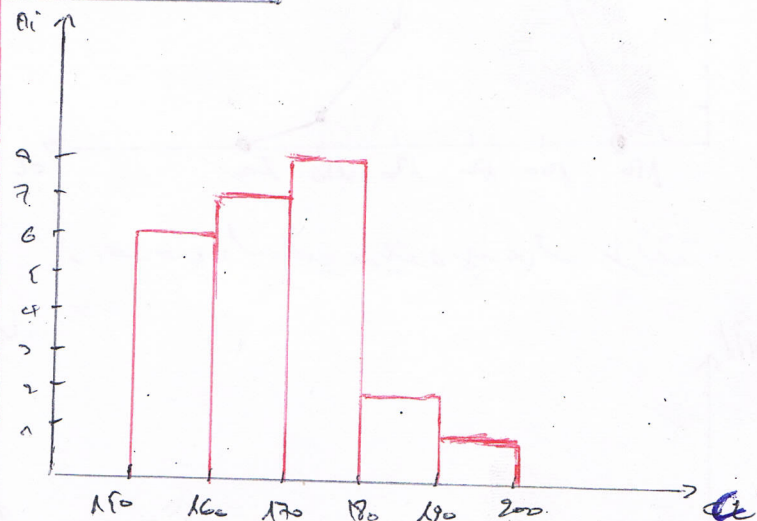
أولاً: نحسب مقدار الزاوية: $\alpha_i^\circ = f_i \times 360$

القسم	عدد الطلبة	f_i	مقدار الزاوية α°
الحقوق	1200	0,3	108°
الاقتصاد	1000	0,25	90°
التسيير	800	0,2	72°
ع. سياسية	600	0,15	54°
إ. آلي تسيير	400	0,1	36°
المجموع	4000	1	360°



[illegible][illegible]

	n_i
$[160, 160[$	6
$[160, 170[$	7
$[170, 180[$	8
$[180, 190[$	2
$[190, 200[$	1



(ایم ایف) کے لئے مسابقتی اسٹریٹیجی (فریڈرک ڈیون)

[illegible]

Frequency	ni
[20. 25[	5
[25. 35[	15
[35. 40[	20
[40. 55[	25
[55. 75[	30
[75. 80[	5
Σ	100

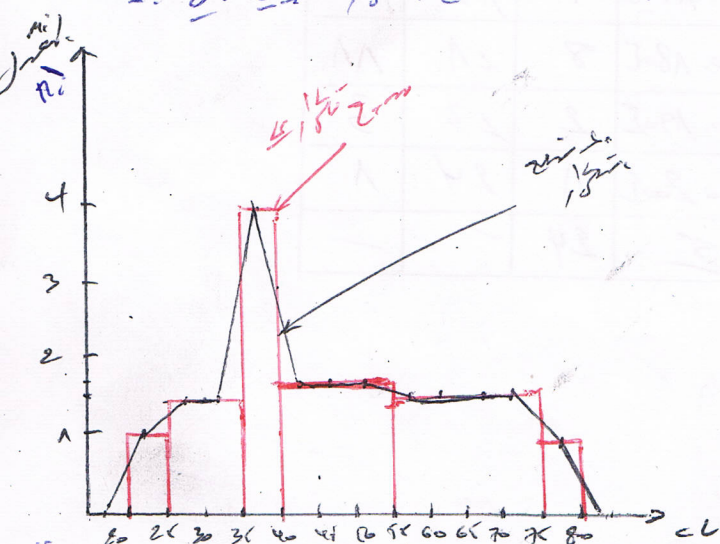
الحل: ما هي مخرجات كل حل؟ اكتب في الجدول
عبر مقارنة البيانات التي تم جمعها من التجارب

از طرفی آنست که منزه است از هر شیئی که در آن باشد از جهت آنکه
منزه است از تمام و بیرون است از هر شیئی که در آن باشد
و در هر شیئی که در آن باشد از جهت آنکه منزه است از تمام

$$\frac{\cancel{\text{Jas. 100}} \times \frac{\cancel{\text{100}}}{\cancel{\text{100}}} = \text{Jas. 100}$$

مسئله: اگر $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$ و $\frac{1}{x} - \frac{1}{y} = \frac{1}{z}$ باشد، $\frac{1}{x}$ و $\frac{1}{y}$ را بیابید.

$\frac{L}{n}$	n	L	$\frac{L}{n}$
12.25	5	5	$\frac{5}{5} = 1$
25.35	15	10	1.5
38.45	20	5	4
51.55	25	15	1.66
64.65	30	20	1.5
77.75	5	5	1
Σ	100		

[illegible]

