المحور الثاني: تنظيم وعرض البيانات الاحصائية

إنّ من أهم الخدمات التي يقدمها الإحصاء للبحوث المختلفة؛ كيفية تنظيم واختصار البيانات بشكل يسمح للعقل أن يفهمها ويستوعبها، ومن اهم الوسائل التي يستخدمها الاحصائيون لهذا الغرض هو التوزيع التكراري الى جانب التمثيل البياني.

أولا: جداول التوزيع التكراري

أ-تعريف التوزيع التكراري: هو عبارة عن توزيع البيانات المأخوذة من ظاهرة معينة على أصناف أو فئات محددة؛ بحيث تقع كل مفردة في فئة واحدة أو صنف واحد.

ب- طرق بناء جدول التوزيع التكراري

ب-1 / بالنسبة للتوزيع التكراري البسيط:

مثال 1: الدرجات التالية تمثل أطوال 15تلميذ مقدرا بالبوصة

75,69,70,69,70,68,79,65,66,75,66,70,75,69,70

المطلوب: قم بإعداد جدول توزيع تكراري لهذه البيانات.

الحل: $N = 15$ (عدد المفردات)

79 أكبر قيمة

65 أصغر قيمة

القيم	65	66	68	69	70	75	79	المجموع
التكرار	1	2	1	3	4	3	1	15

مثال 2: لنفترض أننا قمنا بجمع بيانات حول لون العين لمجموعة من 30 شخصًا. البيانات التي تم جمعها هي كالتالي:

أزرق، بني، أخضر، بني، أزرق، عسلي، بني، أزرق، أخضر، بني، بني، أزرق، عسلي، بني، أزرق، أخضر، بني، أزرق، بني، أخضر، أزرق، بني، عسلي، أزرق، بني، أخضر، بني، أزرق، بني، أخضر.

المطلوب: قم بإعداد جدول توزيع تكراري لهذه البيانات.

الحل:

لون العين	أزرق	بني	أخضر	عسلي	المجموع
التكرار	9	12	6	3	30

ب-2 / بالنسبة للتوزيع التكراري للبيانات المجمعة:

مثال: لنفترض أن لدينا نتائج اختبار القفز الطويل (بالسنتيمتر) لـ 30 طالبًا من طلاب التربية البدنية. هذه هي البيانات الخام:

185, 200, 215, 190, 220, 205, 195, 230, 210, 180,
225, 200, 218, 192, 208, 222, 198, 235, 212, 188,
202, 217, 193, 207, 228, 211, 197, 232, 209, 183

المطلوب: قم بإعداد جدول توزيع تكراري لهذه البيانات.

الحل:

الخطوة الأولى: تحديد المدى(Range)

- المدى العام هو الفرق بين أعلى قيمة وأقل قيمة في البيانات.
- أعلى قيمة = 235 سم
- أقل قيمة = 180 سم
- المدى العام = 235 - 180 = 55 سم

الخطوة الثانية: تحديد طول الفئة C

$$C = \frac{\mathrm{R}}{\mathrm{K}} = \frac{\mathrm{X_{Max}} - \mathrm{X_{Min}}}{1 + 3.322 \log \mathrm{N}}$$

- C يساوي مدى أو طول الفئة.
- R المدى العام، علما أن: $\mathrm{R} = \mathrm{X_{Max}} - \mathrm{X_{Min}}$
- K عدد الفئات، علما أن: $K = 1 + 3.322 \log \mathrm{N}$
- $\mathrm{X_{Max}}$ أكبر قيمة في البيانات.
- $\mathrm{X_{Min}}$ أصغر قيمة في البيانات.
- N عدد البيانات محل الدراسة.

كما أنّ N تمثل مجموع التكرارات

اذن: $c = \frac{R}{K} = \frac{X_{max} - X_{min}}{1+3.322 log N} = \frac{235-180}{1+3.322 log 30}$

$$c \cong \frac{55}{6} \cong 9{,}17$$

لسهولة التعامل، يمكننا تقريب طول الفئة إلى 10 سم، أو نكتفي بـ 9,17سم

الخطوة الثالثة: تحديد حدود الفئات(Class Limits) وتفريغ البيانات

تحديد حدود الفئات كالاتي:

- نبدأ بالحد الأدنى للفئة الأولى (عادةً ما يكون أقل من أو يساوي أصغر قيمة في البيانات). لنبدأ بـ 180 سم.
- الفئة الأولى: 180 - (180 + 10) = 190 - 180 سم
- الفئة الثانية: 190 - (190 + 10) = 200 - 190 سم
- الفئة الثالثة: 200 - (200 + 10) = 210 - 200 سم
- الفئة الرابعة: 210 - (210 + 10) = 220 - 210 سم
- الفئة الخامسة: 220 - (220 + 10) = 230 - 220 سم
- الفئة السادسة: 230 - (230 + 10) = 240 - 230 سم

ملاحظة: لا توجد قاعدة ثابتة لتحديد عدد الفئات، ولكن عادةً ما يتراوح بين 5 و 15 فئة

تحديد طول الفئات وذلك بقسمة المدى العام على عدد الفئات المفترض. و يرى البعض أنه يجب أن لا يتعدى عدد الفئات 15 فئة، لان الغرض من الجداول التكرارية هو اختصار البيانات، و إذا زاد عدد الفئات عن هذا الحد تفقد الجداول التكرارية قيمتها.

تفريغ البيانات ويكون بالقيم او التشطيب؛ كما يلي:

الفئات	التفريغ بالقيم	التكرار
190 - 180	185, 180, 188,183	4
200 - 190	190, 195, 192, 198, 188, 193,197	6
210 - 200	200, 205, 200, 208, 202,207,209	7
220 - 210	215, 210,218,212,217,211	6
230 - 220	220, 225,222,228	4
240 - 230	230, 235, 232	3
المجموع		30

الفئات	التفريغ بالحزم او التشطيب	التكرار
190 - 180	𝍸	4
200 - 190	𝍸 𝍸	6
210 - 200	𝍸 𝍸	7
220 - 210	𝍸 𝍸	6
230 - 220	𝍸	4
240 - 230	𝍸	3
المجموع		30