الاحصاء في مجال التربية الرياضية "statistics"

الاحصاء في مجال التربية الرياضية
“statistics”

تعريف علم الإحصاء

”علم الإحصاء هو العلم الذي يعني بـ: جمع البيانات, تنظيمها, تلخيصها, تحليلها, تفسيرها, وعرضها بطرق مختلفة سواء جدوليا أو بيانيا للوصول لنتائج سليمة وقرارات مقبولة في ظل ظروف غير مؤكدة“

Data collection, Organization, Analysis, Interpretation, and Presentation-uncertainty

وظائف علم الإحصاء

من التعريف: أهم وظائف علم الإحصاء هى:

1. وصف البيانات Data Description
2. الاستدلال الإحصائي Statistical Inference
3. التنبؤ Forecasting

أولا: وصف البيانات

- عند جمع البيانات تكون "بيانات خام": Raw Data
- وصف البيانات يشمل:
  - تلخيص, تبويب وعرض في صورة جداول"Tables" أو رسوم بيانية"Graphs"
  - حساب بعض المؤشرات الإحصائية (متوسطات, تشتت,.......)

ثانيا: الإستدلال الإحصائي

- اختيار عينة (sample) بطريقة علمية من مجتمع(population) الدراسة
- جمع البيانات المطلوبة علي هذه العينة
- التقدير(Estimation):حساب مؤشرات إحصائية -إحصاء- (statistics) علي العينة
- الاستدلال (inference) علي معلمات (parameters) المجتمع
- يتم الإستدلال الإحصائي عبر موضوعين/مرحلتين:

خطوات الاستدلال الإحصائي

1. التقدير Estimation كما سبق
2. اختبارات الفروض (Tests of Hypotheses) : استخدام بيانات العينة للوصول إلى قرار علمي سليم بخصوص الفروض المحددة حول معالم المجتمع.

- نوعين من التقدير:
  - التقدير بنقطة (Point Estimate).
  - التقدير بفترة (Interval Estimate).

ثالثاً: التنبؤ

- تستخدم نتائج الاستدلال الإحصائي، الدالة على الماضي في معرفة ما يمكن أن يحدث في الحاضر والمستقبل.
- توجد أساليب إحصائية عديدة للتنبؤ، مثل استخدام معادلات رياضية يتم تقدير معاملاتها من بيانات العينة، ثم بعد ذلك استخدام المعادلات المقدرة في التنبؤ بما يمكن أن يحدث في المستقبل .

أنواع البيانات وطرق قياسها

- نوع البيانات، وطريقة قياسها من أهم الأشياء التي تحدد التحليل الإحصائي المستخدم
- يمكن تقسيم البيانات إلى مجموعتين هما:

1. البيانات الوصفية Qualitative Data
2. البيانات الكمية Quantitative Data

(1) البيانات الوصفية

- قد تكون غير رقمية، أو رقمية مرتبة في شكل مستويات أو في شكل فئات رقمية.

- تقاس البيانات الوصفية بمعيارين هما:

a) بيانات وصفية مقاسة بمعيار اسمي Nominal Scale:

وهي بيانات غير رقمية, من مجموعات متنافية، لا يمكن المفاضلة بينها أو ترتيبها.

b) بيانات وصفية مقاسة بمعيار ترتيبي Ordinal Scales:

وتتكون من مستويات، أو فئات يمكن ترتيبها تصاعديا أو تنازليا

أمثلة للبيانات الوصفية ذات المعيار الاسمي

1. النوع: متغير وصفي تقاس بياناته بمعيار اسمي " ذكر – أنثى " .
2. الحالة الاجتماعية: متغير وصفي تقاس بياناته بمعيار اسمي " متزوج ـ أعزب ـ أرمل ـ مطلق ".
3. أصناف التمور: متغير وصفي يقاس بياناته بمعيار اسمي " برحي ـ خلاص ـ سكري ـ ....".
4. الجنسية: متغير وصفي يقاس بياناته بمعيار اسمي " مصري ـ سعودي ...."

أمثلة للبيانات الوصفية ذات المعيار الترتيبي

1. تقدير الطالب: متغير وصفي تقاس بياناته بمعيار ترتيبي "-D
D+-C-C+-B-B+-A-A+

2. المستوى التعليمي: متغير وصفي تقاس بياناته بمعيار ترتيبي
"أمي – يقرأ ويكتب ـ ابتدائية ـ متوسطة ـ ثانوية ـ جامعية ـ
أعلى من جامعية "

3. تركيز خلات الصوديوم المستخدم في حفظ لحوم الدجاج من البكتريا: متغير وصفي ترتيبي يقاس بياناته بمعيار ترتيبي
“15% - 10% - 5% - 0%"

(2) البيانات الكمية

- هي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية للظاهرة، وتنقسم إلى قسمين هما:

a) بيانات فترة Interval Data

b) بيانات نسبية Ratio Data

بيانات الفترة Interval Data

- بيانات رقمية ، تقاس بمقدار بعدها عن الصفر، ومن أمثلة ذلك :

  - درجة الحرارة: متغير كمي تقاس بياناته بمعيار بعدي، حيث أن درجة الحرارة "0o" ليس معناه انعدام الظاهرة، ولكنه يدل على وجود الظاهرة.

  - درجة الطالب في الاختبار: متغير كمي تقاس بياناته بمعيار بعدي، حيث حصول الطالب على الدرجة "0" لا يعني انعدم مستوى الطالب.

بيانات نسبية Ratio Data

- متغيرات كمية، تدل القيمة "0" على عدم وجود الظاهرة ومن الأمثلة على ذلك:
  - إنتاجية الفدان بالطن/هكتار.
  - مقدار المسافة التي قفزها اللاعب في مسابقة الوثب الطويل.

طرق جمع البيانات

- طريقة جمع البيانات من أهم مراحل البحث الإحصائي
- جمع البيانات بأسلوب علمي صحيح ← نتائج دقيقة
- طرق جمع البيانات تتلخص في:

1. مصادر البيانات.
2. أسلوب جمع البيانات
3. أنواع العينات
4. وسائل جمع البيانات.

(1) مصادر البيانات

هناك مصدرين للبيانات:

1. المصادر الأولية: يجمع الباحث البيانات من المصدر مباشرة.

- المحاسن: الدقة (نسبيا)-الثقة بالمصدر
- المساوئ: تكلفة عالية ← وقت, جهد, مال, الذاكرة

2. المصادر الثانوية: تجمع البيانات بصورة غير مباشرة.

- المحاسن: توفير الوقت والجهد والمال
- المساوئ: درجة ثقة أقل (نسبيا)

(2) أسلوب جمع البيانات

- بالنظر إلي: هدف البحث و حجم المجتمع هناك أسلوبين لجمع البيانات:

✓ أسلوب الحصر الشامل لمجتمع الدراسة: يتم جمع بيانات عن كل مفردة من مفردات المجتمع بلا استثناء

✓ أسلوب المعاينة: يتم اختيار عينة بطريقة علمية سليمة، ودراستها ثم تعميم نتائج العينة على المجتمع

(أ) أسلوب الحصر الشامل لمجتمع الدراسة

- تجمع المعلومات عن كل مفردة في المجتمع
- مميزاته:
  - **الشمول وعدم التحيز**
  - **دقة النتائج**
- عيوبه:
  - **الوقت والمجهود،**
  - **التكلفة العالية.**

(ب) أسلوب المعاينة

- نختار عينة بطريقة علمية وتطبق عليها الدراسة
- مميزاته:
  - تقليل الوقت والجهد.
  - تقليل التكلفة.
  - الحصول على بيانات أكثر تفصيلا، وخاصة إذا جمعت البيانات من خلال استمارة استبيان
- عيوبه:
  - النتائج قد تكون أقل دقة، وخاصة (إذا)
  - العينة المختارة لا تمثل المجتمع تمثيلا جيدا
  - عدم وجود تنوع كاف (تمثيل العينة).

طرق عرض البيانات

- بعد تحديد العينة وجمع البيانات ياتي تبويب البيانات وعرضها بصورة يمكن الاستفادة منها في وصف الظاهرة محل الدراسة، من حيث تمركز/توزيع البيانات، ودرجة تجانسها. وهناك طريقتين لعرض البيانات هما:
  - عرض البيانات جدوليا.
  - عرض البيانات بيانيا.

تابع طرق عرض البيانات

- الفكرة هنا عرض البيانات آخذين في الإعتبار التشابه/الإختلاف بين المفردات
- عرض البيانات في شكل جدول وصفي مبسط
- يمكننا إجراء حسابات بسيطة لتلخيص خصائص العينة/المتغير
- دعنا ننظر إلي مثالين:
  - مثال (1): **جدول تكراري بسيط** لمتغير وصفي
  - مثال (2): **جدول تكراري بسيط** لمتغير كمي

مثال (1): جدول تكراري بسيط لمتغير وصفي

بيانات عينة من 40 فصيله دم :

A	O+	A	B	A	B	A	O
B	AB	B	A	O+	B	O	B
A	O+	B	B	A	O	B	O+
O+	AB	O+	AB	O	B	A	B
A	O+	B	O	AB	O+	B	A

مناقشة المثال
جدول أولي (تفريـغ البيانات)

عدد الاشخاص (التكرارات)	العلامات الإحصائية	فصيله الدم
5	𝍸	O
10	𝍸 𝍸	A
13	𝍸 𝍸 ///	B
8	𝍸 ///	O+
4	////	AB
40		Sum

الجدول التكراري

التوزيع التكراري النسبي*	عدد الاشخاص (التكرارات)(f)	فصيله الدم
$\left(\frac{5}{40}\right) = 0.125$	5	O
$\left(\frac{10}{40}\right) = 0.25$	10	A
$\left(\frac{13}{40}\right) = 0.325$	13	B
$\left(\frac{8}{40}\right) = 0.20$	8	O+
$\left(\frac{4}{40}\right) = 0.10$	4	AB
1.00	40	Sum

مثال (2): جدول تكراري بسيط لمتغير كمي

- نود الآن تكوين جدول تكراري/التوزيع التكراري النسبي ومن ثم إبداء ملاحظات في حالة: متغير كمي
- لدينا درجات 70 طالبا في مقرر "الاحصاء ":
  - كون التوزيع التكراري والتوزيع التكراري النسبي لدرجات الطلاب.
  - ما نسبة الطلاب الحاصلين على درجة ما بين 70 إلى أقل من 80؟
  - ما هي نسبة الطلاب الحاصلين على درجة أقل من 70 درجة؟
  - ما هي نسبة الطلاب الحاصلين على درجة 80 أو أكثر ؟

مثال لمتغير كمي

- درجات " الاحصاء " –افتراضي:

56	65	70	65	55	60	66	70	75	56
60	70	61	67	61	71	67	62	71	66
68	72	57	68	72	69	57	71	69	75
72	62	67	73	58	63	66	73	63	65
58	73	74	76	74	80	81	60	74	58
76	82	77	83	77	80	91	78	94	72
79	64	57	79	55	87	64	88	78	62

تابع مثال لمتغير كمي

- خطوات تكوين الجدول التكراري:
  - حدد أعلي وأدني قيمة (غي المثال هنا درجة الطالب)
  - احسب المدي(R) ("**R**ange"): (الفرق بين أعلي وأدني قيمة)
  - حدد **عدد الفئات** / **طول الفئة**:

(حسب :رأي الباحث/حجم البيانات/أهداف البحث)

  - كون الجدول من: الفئة/التكرار/التكرار النسبي

تابع مثال لمتغير كمي

- المدي(**R**-Range):

=أعلي درجة - أدني درجة =94-55 = 39

- افرض أن عدد الفئات (**C**) المناسب هو 8
- طول الفئة : $R/C=39/8=4.875\approx 5$
- الآن نحدد كل فئة من الفئات الـ 8

تحديد الفئات-تابع

- كل فئة عبارة قيمتين المسافة بينهما هي طول الفئة (5): القيمة الأولي تسمي (الحد الأدني)؛ والثانية (الحد الأعلى).
- كل فئة تبدأ عند انتهاء الفئة التي قبلها بحيث:
- كل مفردة من المفردات تقع داخل فئة واحدة فقط؛ مثلا:
- الفئة الأولي :
  - الحد الأدني = أقل درجة (55=)
  - الحد الأعلي= الحد الأدني + طول الفئة=55+5=60
  - اذا الفئة الأولي تكون " من 55 إلى أقل من 60 "
  - لاحظ إن الدرجة (60) لا تقع داخل الفئة الأولي

تحديد الفئات-تابع

- الفئة الثانية:
- الحد الأدنى= 60
- الحد الأعلى=60+5=65
  - اذا الفئة الأولي تكون " من 60 إلى أقل من 65 "
- وهكذا (أنظر الجدول لكل الفئات)

الجدول التكراري-مثال كمي

الفئات	عدد الطلاب (التكرارات) *(f)*	التكرار النسبي
55 - 60	10	0.143
60 - 65	12	0.171
65 - 70	13	0.186
70 - 75	16	0.229
75 - 80	10	0.143
80 - 85	4	0.057
85 - 90	3	0.043
90 - 95	2	0.028
المجموع	**70**	1.00

العرض البياني للبيانات الكمية

- هو أحد طرق وصف البيانات
- يوضح شكل وكيفية توزيع البيانات ومدى تمركزها
- تستخدم فيه بعض الأشكال البيانية المختلفة مثل المدرج – المضلع – المنحني – الدائرة ... وهكذا

(1) المدرج التكراري Histogram

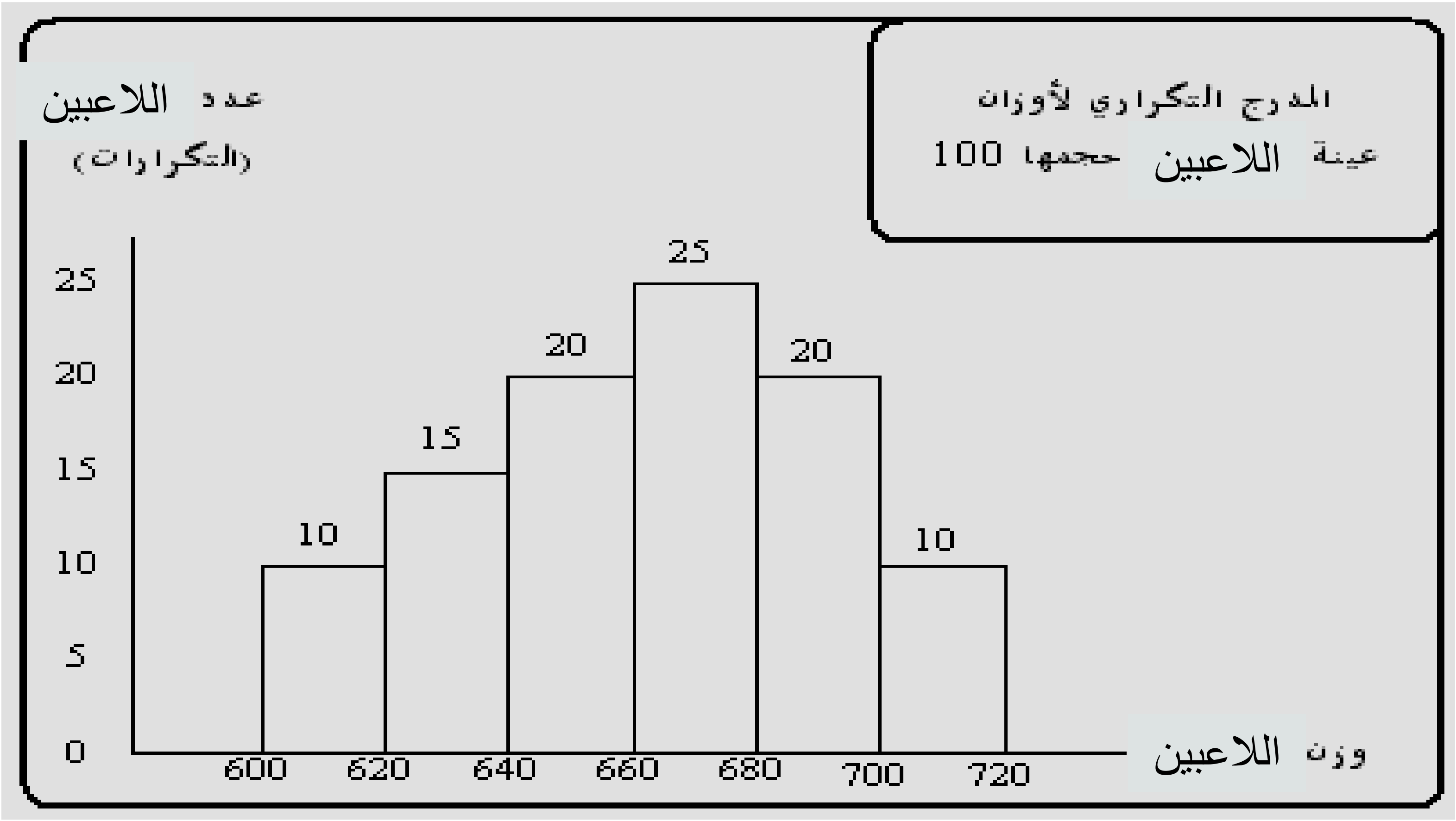
- تمثيل بياني للجدول التكراري البسيط الخاص بالبيانات الكمية
- عبارة عن أعمدة بيانية متلاصقة
- التكرارات على المحور الرأسي، بينما قيم المتغير (حدود الفئات) على المحور الأفقي
- تمثل كل فئة بعمود، ارتفاعه هو تكرار الفئة، وطول قاعدته هو طول الفئة.
- مثال:

المدرج التكراري-مثال

- لدينا التوزيع التكراري التالي لإعداد لاعبي حمل الأثقال حسب الفئات التالية التي يستطيع كل لاعب رفعها :
  - ارسم المدرج التكراري.
  - ارسم المدرج التكراري النسبي، ما ملاحظاتك و تفسيرك للرسم البياني

الوزن المرفوع	-600	-620	-640	-660	-680	700-720	Sum
عدد اللاعبين	10	15	20	25	20	10	100

المدرج التكراري لعينة اللاعبين

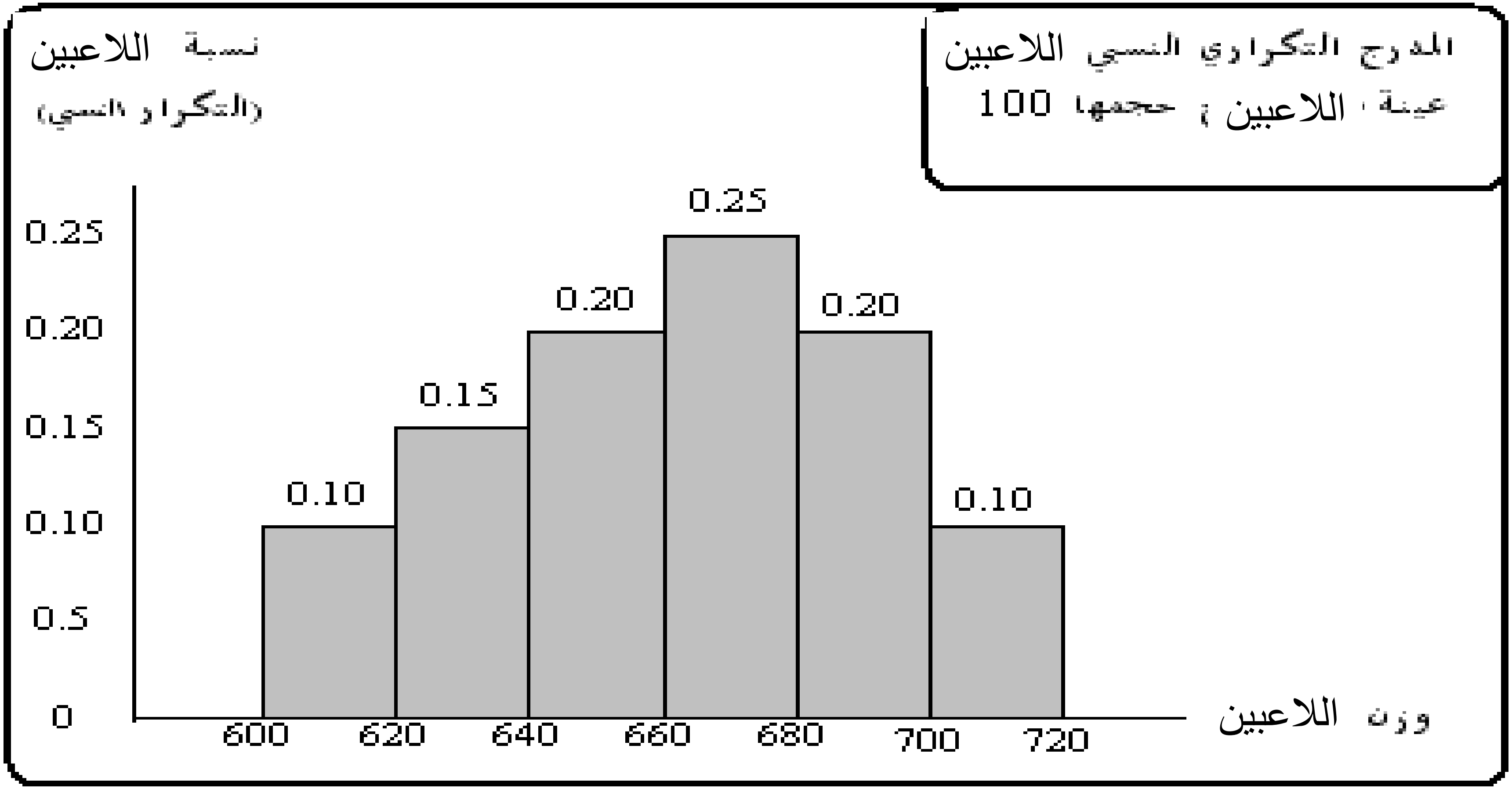


رسم المدرج التكراري النسبي

- **لرسم المدرج التكراري النسبي, أولا نحسب التكرار النسبي:**

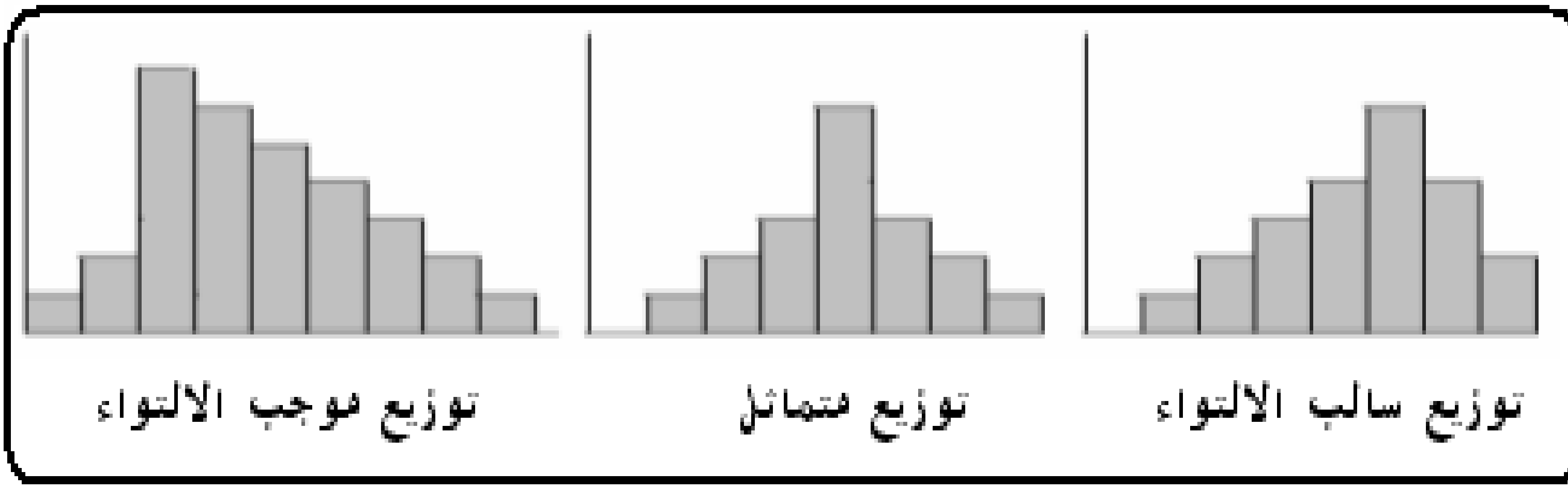
الوزن	-600	-620	-640	-660	-680	700-720	sum
عدد اللاعبين	10	15	20	25	20	10	100
التكرار النسبي	0.10	0.15	0.20	0.25	0.20	0.10	1.0

المدرج التكراري النسبي لعينة اللاعبين



ملاحظات على شكل المدرج التكراري

- أشكال توزيع البيانات في المدرج:



أغلب اللاعبين ترفع أثقال كبيرة ولا توجد مشكلة بالنادي

اعداد اللاعبين متساوية لرفع الأثقال المختلفة

أغلب اللاعبين ترفع أثقال خفيفة ولدينا مشكلة

(2) المضلع التكراري
(RF-Polygon)

- تمثيل بياني أيضا للجدول التكراري البسيط
- تمثل التكرارات على المحور الرأسي ومراكز الفئات على المحور الأفقي
- توصل الإحداثيات بخطوط مستقيمة، وبعد ذلك يتم توصيل طرفي المضلع بالمحور الأفقي.
- يحسب مركز الفئة كما يلي:

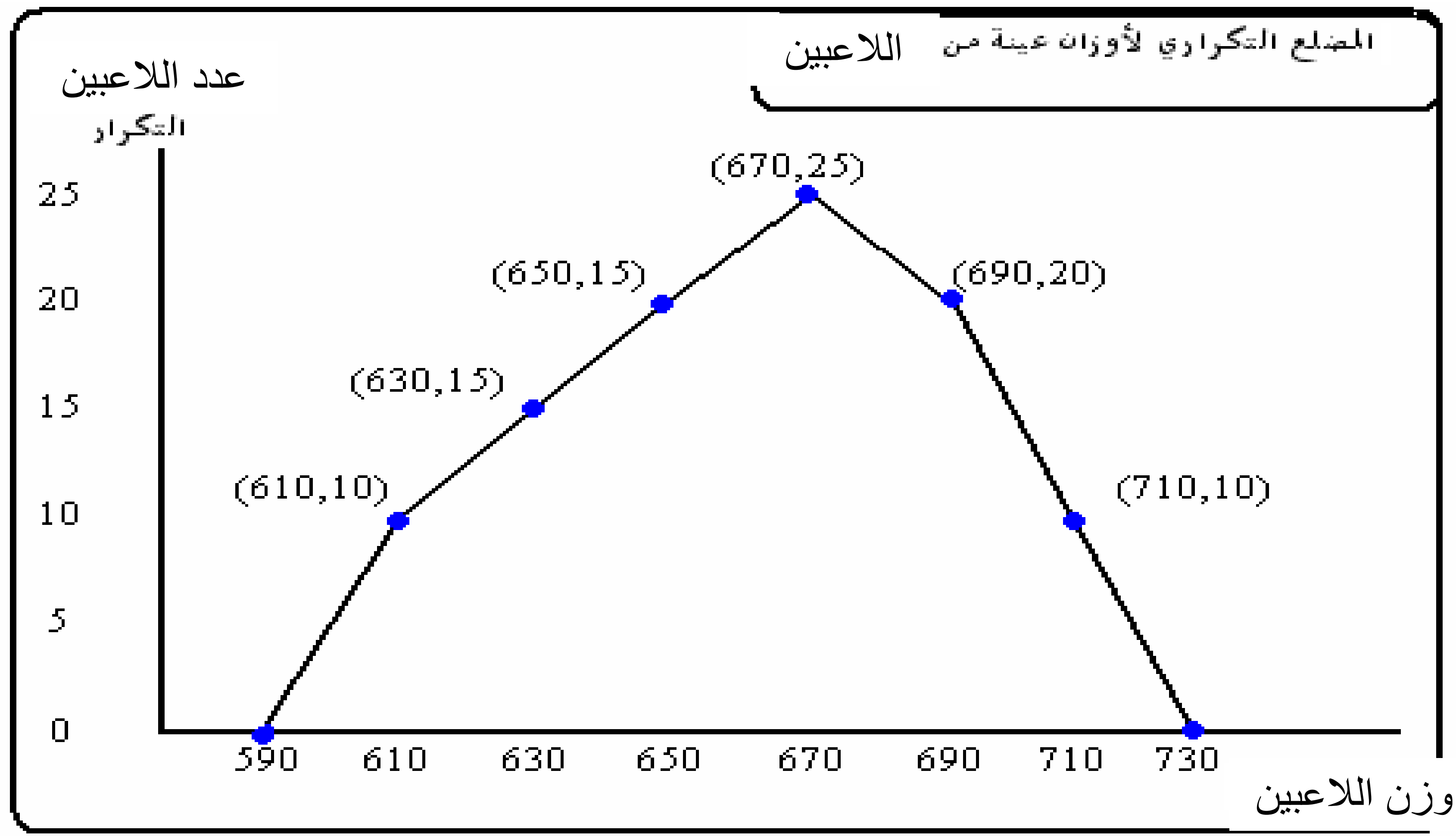
$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{٢} \qquad (٢-٣)$$

$$Midpoint = \frac{Lower + Upper}{2}$$

تكوين/رسم المضلع التكراري

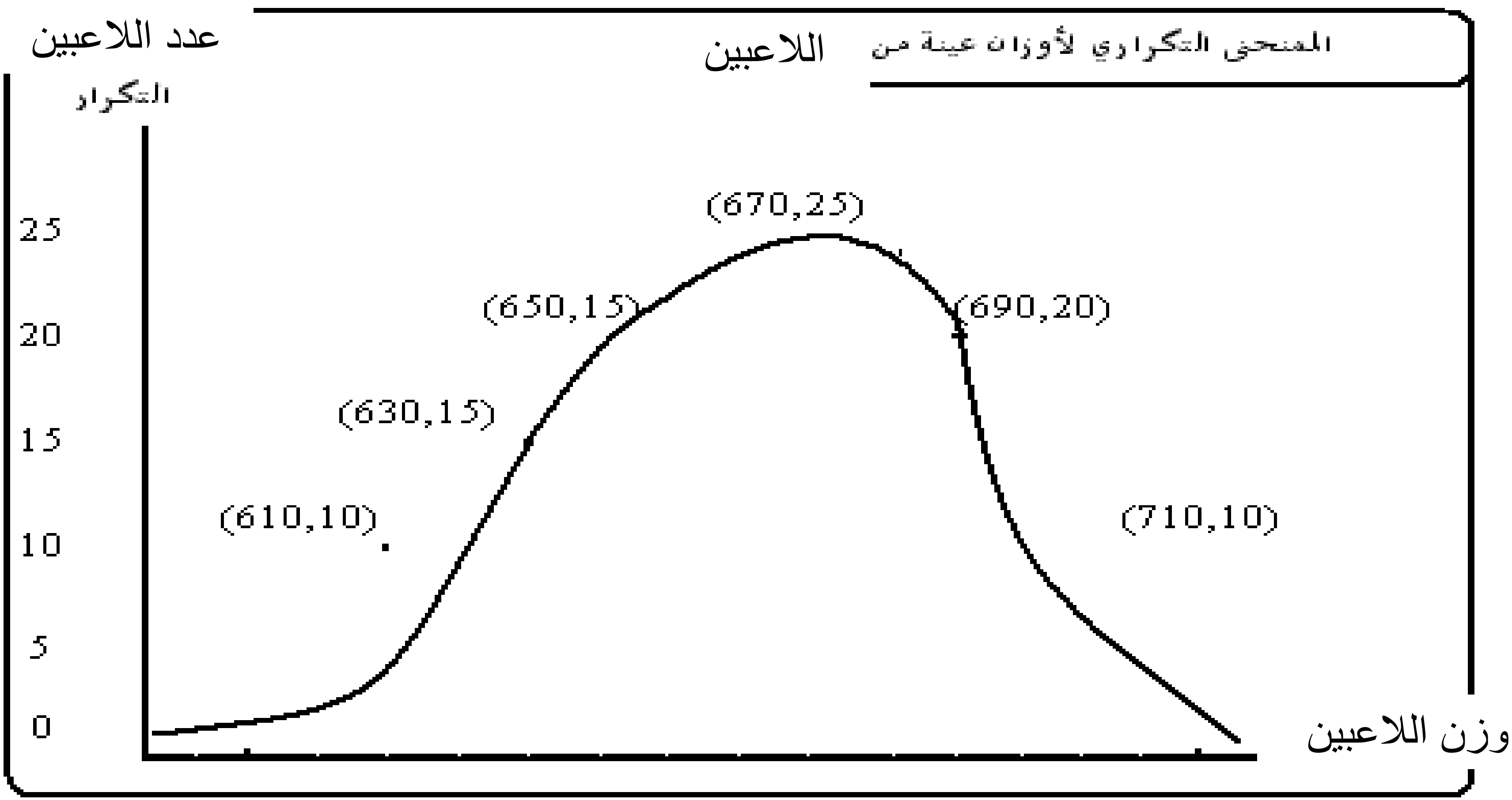
مركز الفئة (x)	(عدد اللاعبين)التكرار (f)	الوزن
(600+620)/2= 610	10	600-
(620+640)/2=630	15	620-
650	20	640-
670	25	660-
690	20	680-
(700+720)/2=710	10	700-720
	100	Sum

المضلع التكراري لعينة اللاعبين



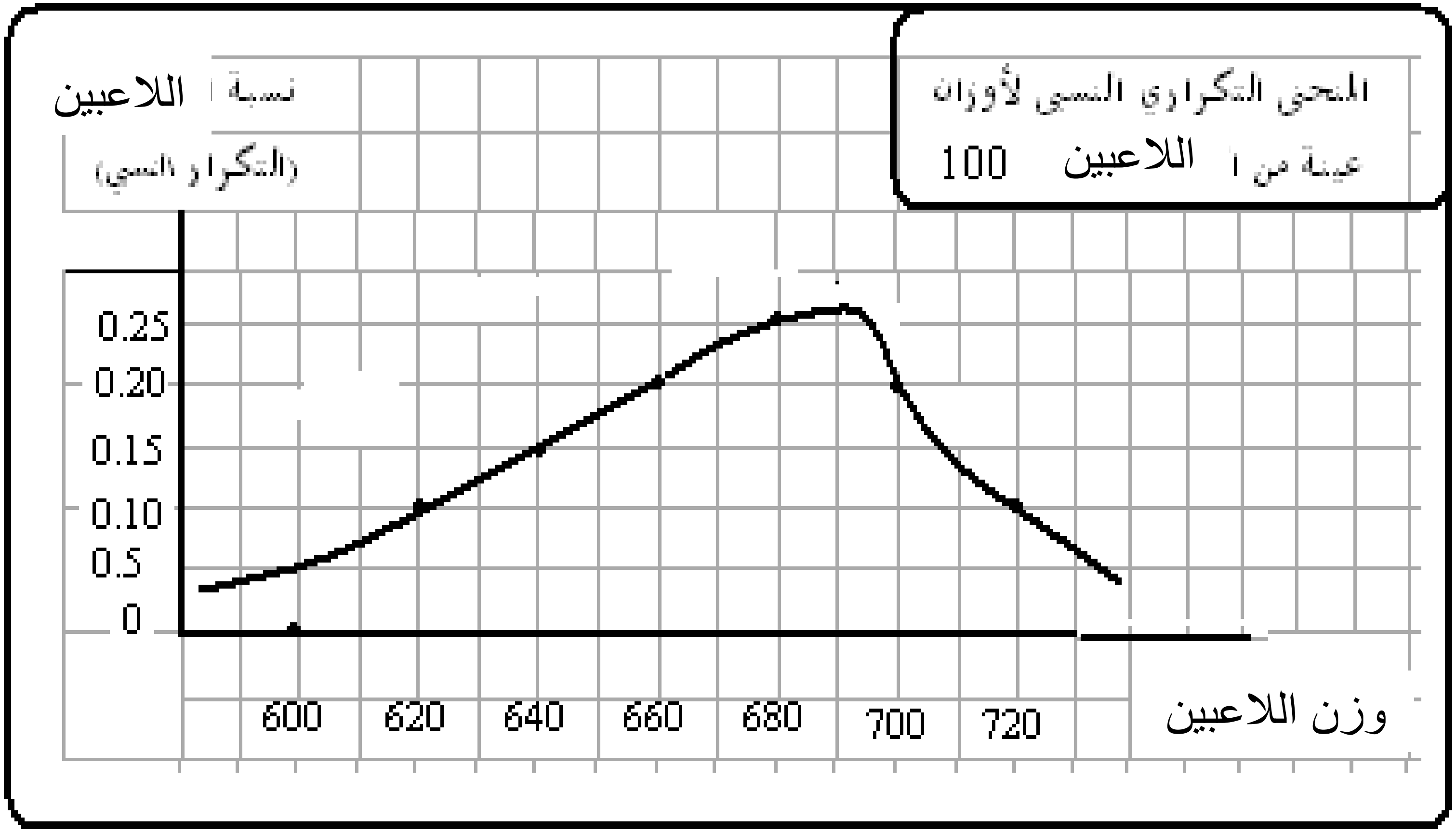
(3) المنحني التكراري

- مثل المضلع التكراري مع استبدال الخط الواصل بين النقاط بمنحني



المنحني التكراري النسبي

- يمكن رسم المنحني التكراري النسبي-شبيه بالمضلع التكراري النسبي مع استبدال اخط بمنحني



التوزيع التكراري المتجمع الصاعد

- الجدول التكراري أدناه يبين توزيع 40 لاعب في نادي حسب أعمارهم , المطلوب:
  - كون جدول التوزيع التكراري المتجمع الصاعد.
  - كون جدول التوزيع التكراري المتجمع الصاعد النسبي.
  - ارسم المنحنى التكراري المتجمع الصاعد النسبي

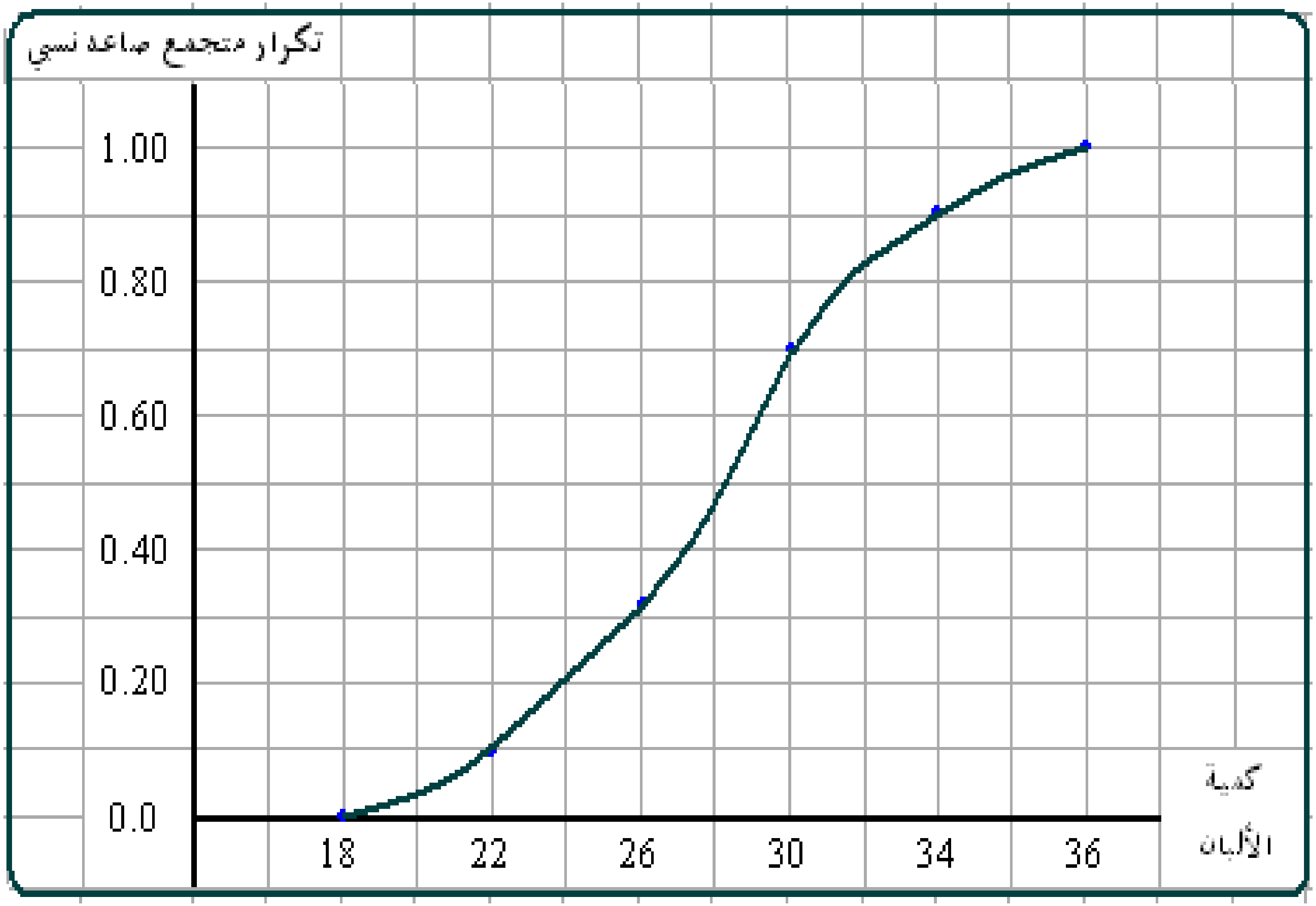
العمر	18-	22-	26-	30-	34-38	Sum
عدد اللاعبين	4	9	15	8	4	40

التوزيع التكراري المتجمع الصاعد والنسبي

- يمكن تكوين هذين التكرارين كالآتي:

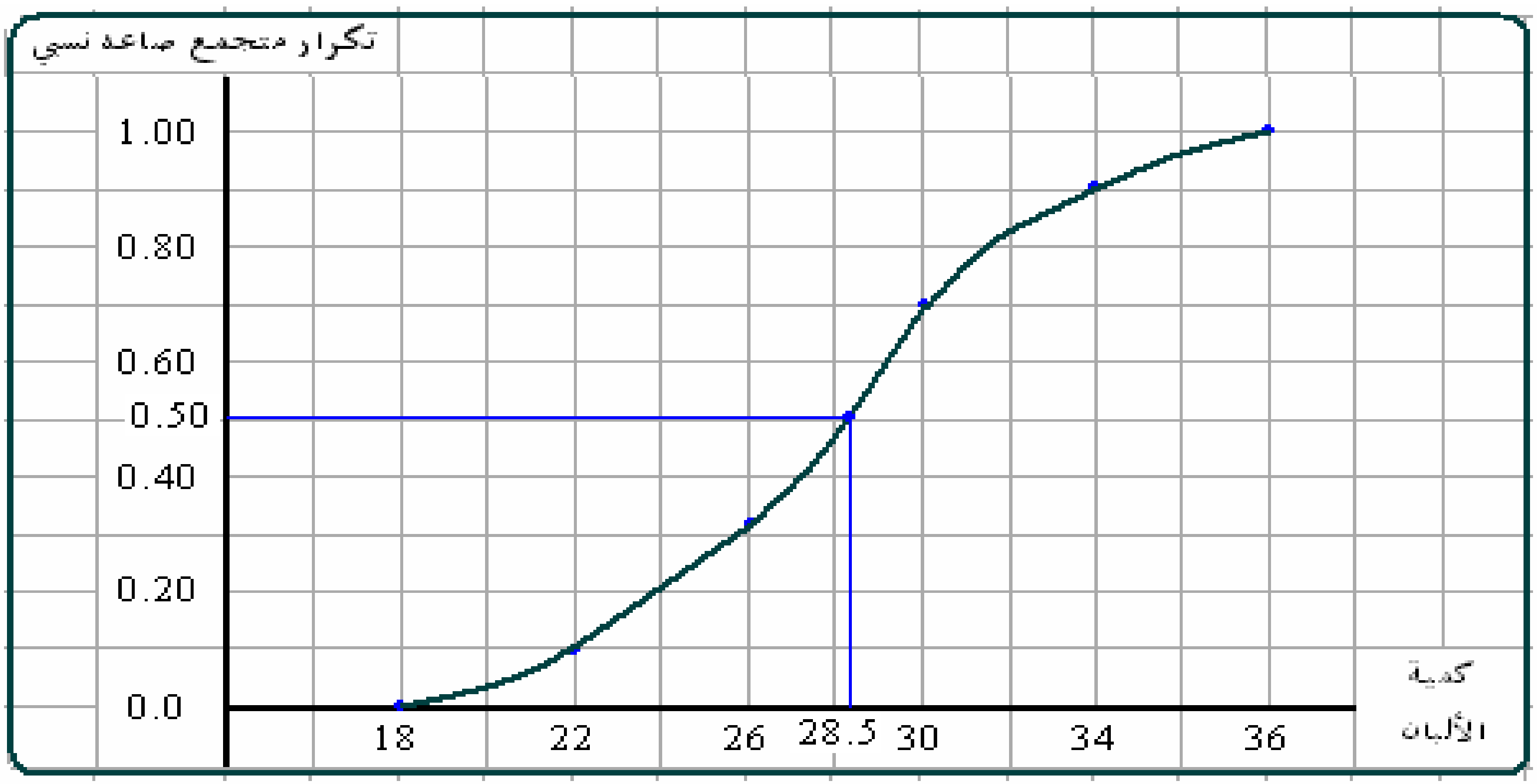
تكرار متجمع صاعد نسبي	تكرار متجمع صاعد	أقل من
0.00	0	أقل من 18
0.10	4	أقل من 22
0.325	13	أقل من 26
0.70	28	أقل من 30
0.90	36	أقل من 34
1.00	40	أقل من 38

رسم المنحنى التكراري المتجمع الصاعد النسبي



قراءة المنحنى التكراري المتجمع الصاعد النسبي

- تحديد نسبة المفردات الأقل من قيمة محددة
- تحديد نسبة المفردات الواقعة بين قيم محددة
- تحديد قيم معينة—مثل الوسيط

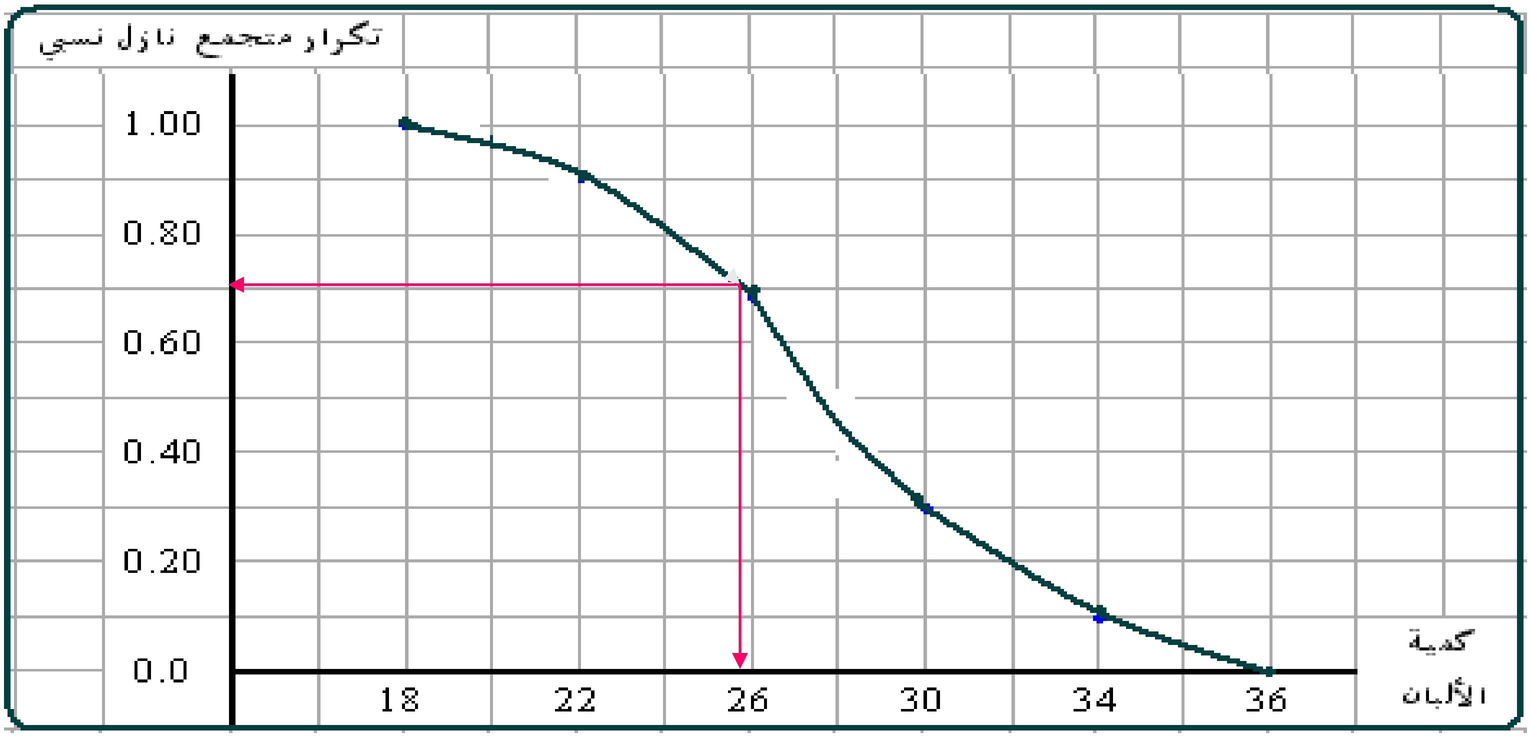


التوزيع التكراري المتجمع الهابط

- يمكن تكوين التوزيع المتجمع الهابط والنسبي للمثال السابق كما يلي:

تكرار متجمع صاعد نسبي	تكرار متجمع صاعد	أكثر من أو يساوي
1.00	40	18 أو أكثر
0.90	36	22 أو أكثر
0.70	28	26 أو أكثر
0.325	13	30 أو أكثر
0.10	4	34 أو أكثر
0.00	0	38 أو أكثر

رسم المنحنى التكراري المتجمع النسبي الهابط



العرض البياني للبيانات الوصفية:
(الدائرة البيانية Pie Chart-)

- لعرض بيانات المتغير الوصفي في شكل دائرة، توزع الـ 360o حسب التكرار النسبي لمجموعات المتغير
- يمكن تحديد مقدار الزاوية الخاصة بأية مجموعة بتطبيق المعادلة التالية:

مقدار الزاوية = 360o× التكرار النسبي للمجموعة

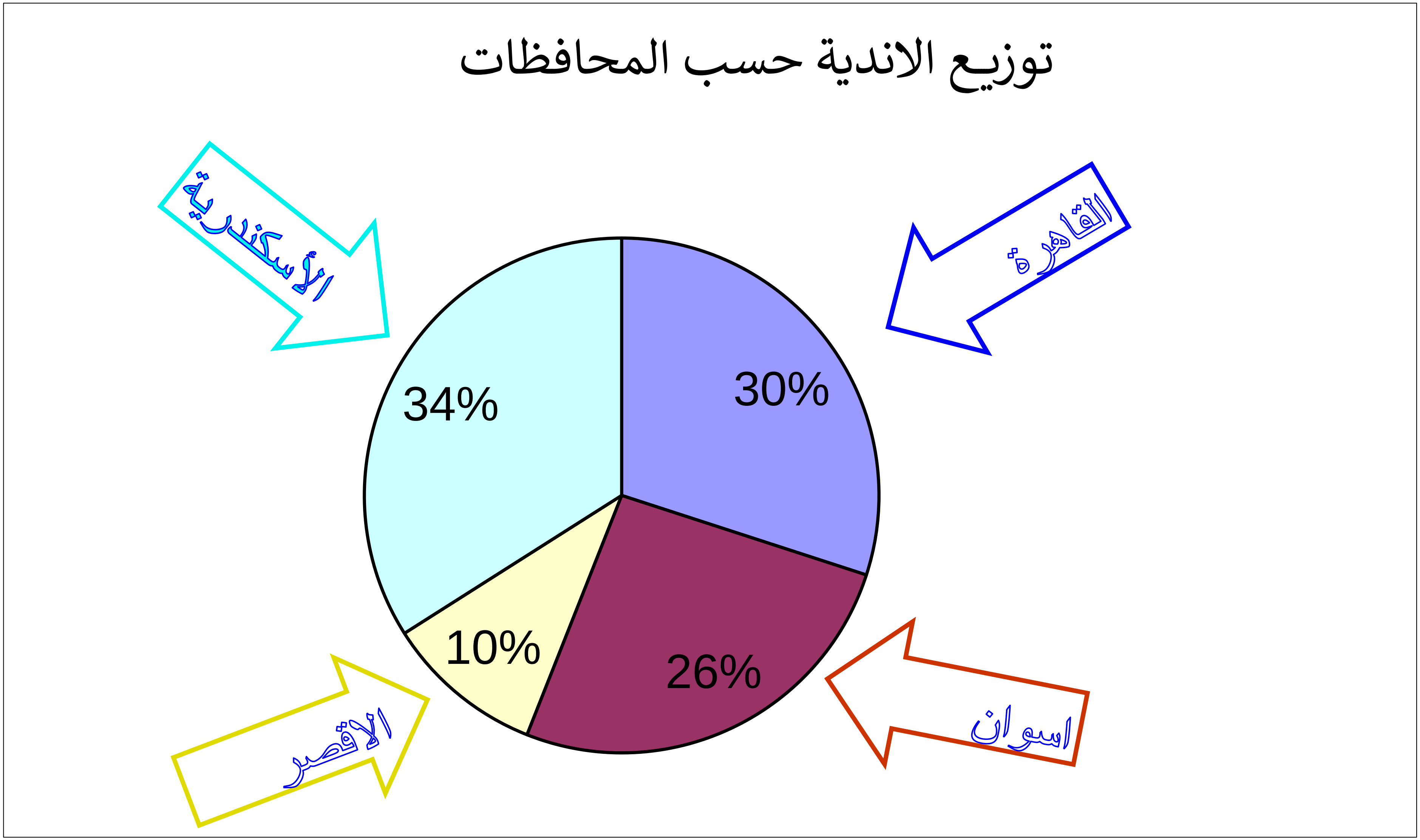
- العديد من البرامج الإحصائية (Excel مثلا) تقوم بهذه العملية بيسر

مثال
(الدائرة البيانية -Pie Chart)

- الجدول التكراري التالي يبين توزيع عينة حجمها 500 نادي حسب المحافظات.

المحافظة	القاهرة	اسوان	الأقصر	الأسكندرية	sum
عدد الأندية	150	130	50	170	500
التكرار النسبي	0.3	0.26	0.10	0.34	1.00

تابع مثال الدائرة البيانية



مقاييس النزعة المركزية
Measures of Central Tendency

- تسمى أيضا بمقاييس الموضع أو المتوسطات
- هى القيم التى تتركز القيم حولها وهي:
- الوسط الحسابي
- الوسيط Median
- المنوال Mode

- مقاييس النزعة المركزية و شكل توزيع البيانات
- الرباعيات Quartiles

الوسط الحسابي ($\bar{X}$)
Arithmetic Mean

- من أهم مقاييس النزعة المركزية
- وأكثرها استخداما في النواحي التطبيقية
- ويمكن حسابه للبيانات المبوبة وغير المبوبة
- خصائص الوسط الحسابي
- مزايا وعيوب الوسط الحسابي

الوسط الحسابي للبيانات غير المبوبة

- بشكل عام الوسط الحسابي هو مجموع القيم مقسوما على عددها
- لدينا n من القيم ، ويرمز لها بالرمز : $x_1, x_2, \ldots, x_n$
- الوسط الحسابي لهذه القيم ، $\bar{X}$ يحسب بالمعادلة التالية:

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (٣-١)$$

- تمرين: فيما يلي درجات 8 طلاب في مقرر 122 إحصاء تطبيقي .
- 40 36 40 35 37 42 32 34

ماهو الوسط الحسابي لدرجة الطالب في الامتحان؟

الوسط الحسابي للبيانات المبوبة

- يمكننا حساب الوسط الحسابي من الجدول التكراري
- إذا كان لدينا (k) فئة ومراكز هذه الفئات هي:

$$x_1, x_2, \dots, x_k$$

تكرارات هذه الفئات هي $f_1, f_2, \dots, f_k$

- فإن الوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} \qquad (٣-٢)$$

- مثال: احسب الوسط الحسابي لتوزيع 40 تلميذ حسب أوزانهم بالجدول

فئات الوزن	32-34	34-36	36-38	38-40	40-42	42-44
عدد التلاميذ	4	7	13	10	5	1

حل المثال

فئات الوزن (C)	التكرارات f	مراكز الفئات x	$x\ f$
32-34	4	$(32+34)\div 2=33$	$4\times 33=132$
34-36	7	35	$7\times 35=245$
36-38	13	37	$13\times 37=481$
38-40	10	39	$10\times 39=390$
40-42	5	41	$5\times 41=205$
42-44	1	43	$1\times 43=43$
المجموع	40		1496

$$\overline{x}=\frac{\sum_{i=1}^{6} x_i\ f_i}{\sum_{i=1}^{6} f_i}=\frac{1496}{40}=37.4\ k.g$$

خصائص الوسط الحسابي

1. الوسط الحسابي للمقدار الثابت يساوى الثابت نفسه ، أي أنه إذا كانت قيم (x) هي : (a, a, a,...,a) ، فإن الوسط الحسابي هو:

$$\boxed{\bar{x} = \frac{a + a + \ldots + a}{n} = \frac{na}{n} = a} \qquad (٣-٣)$$

2. مجموع "انحرافات" القيم عن وسطها الحسابي يساوى صفرا ، ويعبر عن هذه الخاصية بالمعادلة:

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

3. عند اضافة مقدار ثابت a لكل قيمة للقيم فإن الوسط الحسابي الجديد $\bar{y}$ يساوى الوسط الحسابي القديم $\bar{x}$ مضافا إليه هذا المقدار الثابت:

$$\boxed{\bar{y} = \bar{x} + a} \qquad (٣-٥)$$

تابع خصائص الوسط الحسابي

4. إذا ضرب مقدار ثابت (a) في كل قيمة من القيم ، فإن الوسط الحسابي الجديد ($\bar{y}$) يساوى الوسط الحسابي القديم ($\bar{x}$) مضروبا في هذا المقدار الثابت:

$$\boxed{\bar{y} = a\ \bar{x}} \qquad (٣-٦)$$

5. مجموع مربعات انحرافات القيم عن وسطها الحسابي أقل ما يمكن:

$$\boxed{\sum (x-\bar{x})^2 \langle \sum (x-a)^2 \quad if \quad a \neq \bar{x}} \qquad (٣-٧)$$

الوسط الحسابي المرجح weighted arithmetic mean

- بعض الأحيان يكون لكل قيمة من قيم المتغير (x_i)أهمية نسبية تسمى وزن (weight "w_i")
- أخذ هذه الأوزان بالاعتبار يجعل الوسط الحسابي دقيقا ومعبرا ويحسب من:

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i}$$

- **تمرين**: الجدول التالي درجات الطلاب لمقرر الإحصاء وعدد ساعات الاستذكار (الأوزان) بالأسبوع . أوجد الوسط الحسابي المرجح:

مسلسل	1	2	3	4	5	sum
X (الدرجة)	23	40	36	28	46	173
W (عدد ساعات الاستذكار)	1	3	3	2	4	13

مزايا وعيوب الوسط الحسابي

- يتميز الوسط الحسابي بالمزايا التالية :
  - أنه سهل الحساب .
  - يأخذ في الاعتبار كل القيم .
  - أنه أكثر المقاييس استخداما وفهما .
- ومن عيوبه .
  - أنه يتأثر بالقيم الشاذة والمتطرفة .
  - يصعب حسابه في حالة البيانات الوصفية .
  - قد يكون خادعا مالم يصاحبه مقياس للتشتت .

الوسيط
Median

- **هو أحد مقاييس النزعة المركزية**
- **يأخذ في الاعتبار رتب القيم**
- **ويعرف الوسيط بأنه القيمة** التي تقع وسط القيم بعد ترتيبها إما تصاعديا أو تنازليا : إذا نصف القيم تكون أقل من الوسيط ونصفها الآخر أكبر منه.

الوسيط للبيانات غير المبوبة

- رتب القيم تصاعديا .
- إذا كان عدد القيم (n) فردي فإن:

$$\text{الوسيط} = \text{القيمة رقم} \left(\frac{n+1}{2}\right)$$

- إذا كان عدد القيم زوجي:

$$\text{الوسيط} = \frac{\text{القيمة رقم} \left(\frac{n}{2}\right) + \text{القيمة رقم} \left(\frac{n}{2}+1\right)}{2}$$

- **مثال**: تم تطبيق اسلوبين من التدريب الاسلوب(a) تم تطبيقه على 7 لاعبين والاسلوب (b) تم تطبيقه على 10 لاعبين ، وبعد انتهاء الموسم الرياضي ، تم تسجيل قوة تحمل جميع اللاعبين في المجموعتين ، وكانت:

النوع (a)	1.2	2.75	3.25	2	3	2.3	1.5			
النوع (b)	4.5	1.8	3.5	3.75	2	2.5	1.5	4	2.5	3

- احسب الوسيط لكل اسلوب تدريبي.

1. الوسيط للاسلوب الأول (a): بعد ترتيب القيم

قيمة الوسيط

الإنتاج	1.2	1.5	2	2.3	2.75	3	3.25
الرتبة	1	2	2	4	5	6	7

رتبة الوسيط

- عدد القيم فردى n=7 اذا الرتبه $\left((n+1)/2=(7+1)/2=4\right)$
- ويكون الوسيط هو القيمة رقم 4 ، أي أن وسيط الاسلوب a: (2.3)

2. حساب وسيط الاسلوب (b):

$$\text{قيمة الوسيط} = \frac{2.5 + 3}{2}$$

2.75

الإنتاج	1.5	1.8	2	2.5	2.5	3	3.5	3.75	4	4.5
الرتبة	1	2	3	4	5	6	7	8	9	10

5.5

رتبة الوسيط

- عدد القيم زوجي $n=10$ الوسيط هو متوسط القيمتين $(n/2 +1)$و $(n/2)$:

الوسيط للبيانات المبوبة

- من جدول توزيع تكراري ، يتم إتباع الخطوات التالية:
- تكوين الجدول التكراري المتجمع الصاعد
- تحديد رتبة الوسيط: $\left(\frac{\sum f}{2}\right) = \left(\frac{n}{2}\right)$
- تحديد فئة الوسيط كما في الشكل التالي :

تكرار متجمع صاعد سابق f_1	الحد الأدنى لفئة الوسيط (A)
رتبة الوسيط $(n/2)$ ⟵	الوسيط Med
تكرار متجمع صاعد لاحق f_2	الحد الأعلى لفئة الوسيط

تابع

- يحسب الوسيط ، بتطبيق المعادلة:

$$\boxed{Med = A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L} \qquad (١١-٣)$$

- حيث أن :

L هي طول فئة الوسيط، وتحسب بالمعادلة التالية:

طول الفئة = الحد الأعلى – الحد الأدنى

L = Upper - Lower

مثال

- فيما يلي توزيع 50 رياضي، حسب الاحتياجات اليومية من الغذاء بالكيلوجرام:

فئات الاحتياجات اليومية	1.5 -	4.5 -	7.5 -	10.5 -	13.5 – 16.5
عدد الرياضيين f	4	12	19	10	5

- والمطلوب : حساب الوسيط : أ - حسابيا ب- بيانيا

حل المثال

- <u>حساب الوسيط حسابيا</u> :
- رتبة الوسيط :

$$\frac{n}{2}=\frac{\sum f}{2}=\frac{50}{2}=25$$

- الجدول التكراري المتجمع الصاعد :

أقل من	1.5	4.5	7.5 A	**الوسيط**	10.5	13.5	16.5
تكرار متجمع صاعد	0	4	16 f_1	**25**	35 f_2	45	50

تابع

- من معادلة الوسيط نجد أن:

$$A = 7.5 \ , \ \ f_1 = 16 \ , \ \ f_2 = 35 \ , \ L = 10.5 - 7.5 = 3$$

- إذا الوسيط قيمته هي :

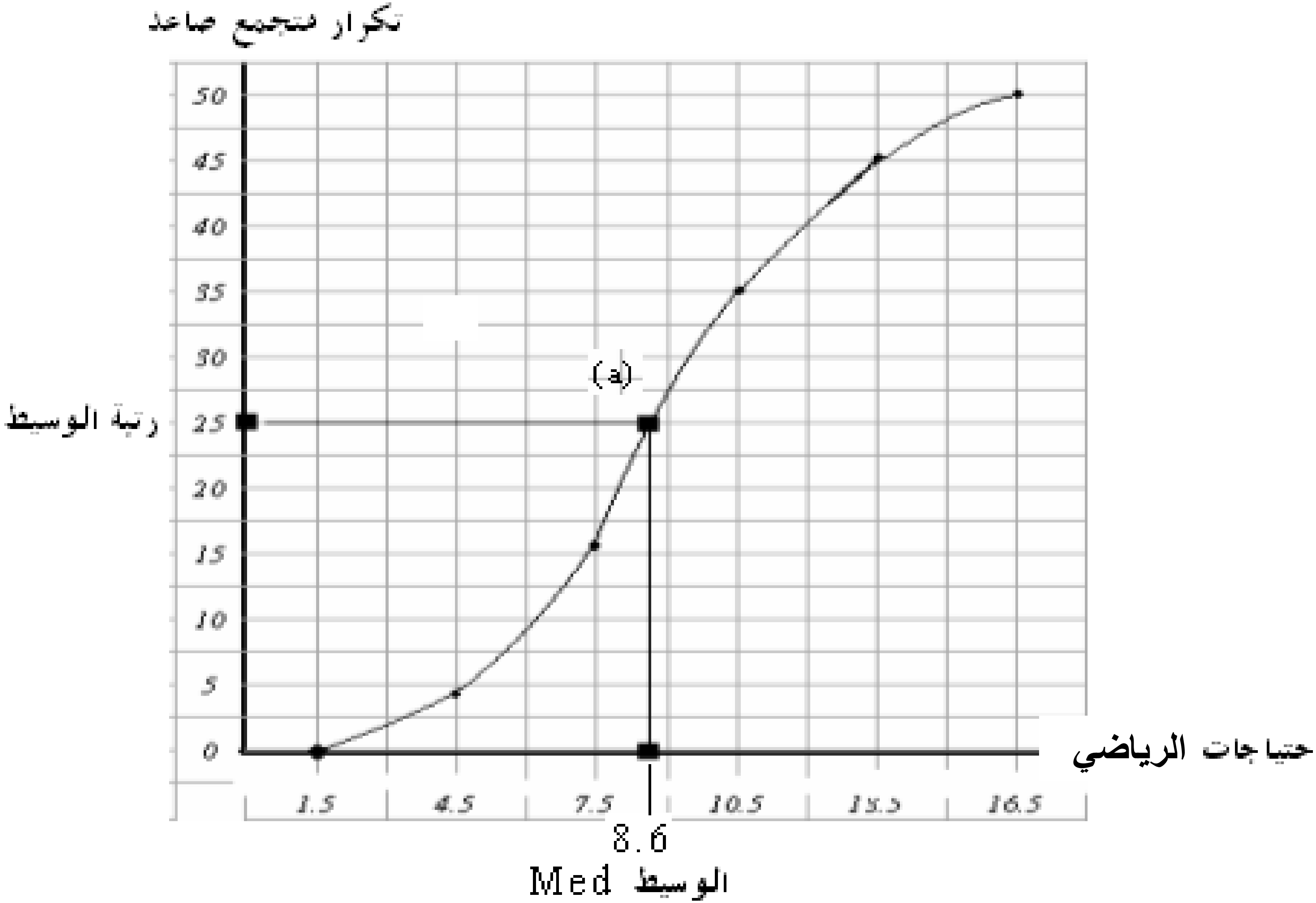
$$Med = A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L = 7.5 + \frac{25-16}{35-16} \times 3$$

$$= 7.5 + \frac{9}{19} \times 3 = 7.5 + \frac{27}{19} = 7.5 + 1.421 = 8.921 \ k.g$$

الوسيط-بيانيا

- تمثيل جدول التوزيع التكراري المتجمع الصاعد بيانيا
- تحديد رتبة الوسيط (25) على المنحنى التكراري المتجمع الصاعد . ثم رسم خط مستقيم أفقي حتى يلقى المنحنى في النقطة (a) .
- إسقاط عمود رأسي من النقطة (a) على المحور الأفقي .
- نقطة تقاطع الخط الرأسي مع المحور الأفقي تعطى قيمة الوسيط .
- الوسيط كما هو مبين في الشكل $Med = 8.6$

الوسيط-بيانيا



مزايا وعيوب الوسيط

- **من مزاياه:**
  - **لا يتأثر بالقيم الشاذة أو المتطرفة .**
  - **كما أنه سهل في الحساب .**
  - **مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيم أخرى:**

$$\sum |x - Med| \leq \sum |x - a| \ , \ \ a \neq Med$$

مزايا وعيوب الوسيط

- ومن عيوبه:
  - أنه لا يأخذ عند حسابه كل القيم في الاعتبار، فهو يعتمد على قيمة أو قيمتين فقط .
  - يصعب حسابه في حالة البيانات الوصفية المقاسة بمعيار اسمي (nominal)

Mode - المنوال

- المنوال هو القيمة الأكثر شيوعا أو تكرارا
- يكثر استخدامه في حالة البيانات الوصفية ، لمعرفة النمط (المستوى) الشائع
- بالنسبه للبيانات المبوبة – القيمة ذات أكثر تكرار

المنوال:مثال (بيانات غير مبوبة)

- فيما يلي أوزان الطلاب لعشرة عينات عشوائية (طلاب) من أقسام كلية التربية الرياضية:
- أحسب المنوال لكل قسم ؟

علوم الصحة	80	77	75	77	77	77	65	70	58	67
العاب جماعية	88	68	60	75	93	65	77	85	95	90
المنازلات	80	65	69	80	65	88	76	65	86	80
الإدارة	85	73	69	85	73	69	69	73	72	85

المنوال للبيانات المبوبة

- **(طريقة الفروق)** : يحسب المنوال من القانون:

$$\boxed{Mod = A + \frac{d_1}{d_1 + d_2} \times L} \qquad (١٣-٣)$$

- حيث:
- **A : الحد الأدنى لفئة المنوال (الفئة ذات أكبر تكرار) .**
  - **D_1 : الفرق الأول = (تكرار فئة المنوال – التكرار سابق)**
  - **D_2 : الفرق الثاني = (تكرار فئة المنوال – التكرار لاحق)**
  - **L : طول فئة المنوال .**

المنوال للبيانات المبوبة-مثال

- فيما يلي توزيع 30 أسرة حسب الإنفاق الاستهلاكي الشهري (ألف جنيه)

فئات الإنفاق	2 -	5 -	8 -	11 -	14 - 17
عدد الأسر (f)	4	7	10	5	4

- والمطلوب حساب المنوال، باستخدام طريقة الفروق

◆ من الشكل التالي نجد كل مكونات القانون:

الفئات	التكرارات
2 -	4
5 -	7
8 -	10
11 -	5
14 - 17	4

فئة المنوال

$A = 8$

$d_1 = 10 - 7 = 3$

أكبر تكرار

$d_2 = 10 - 5 = 5$

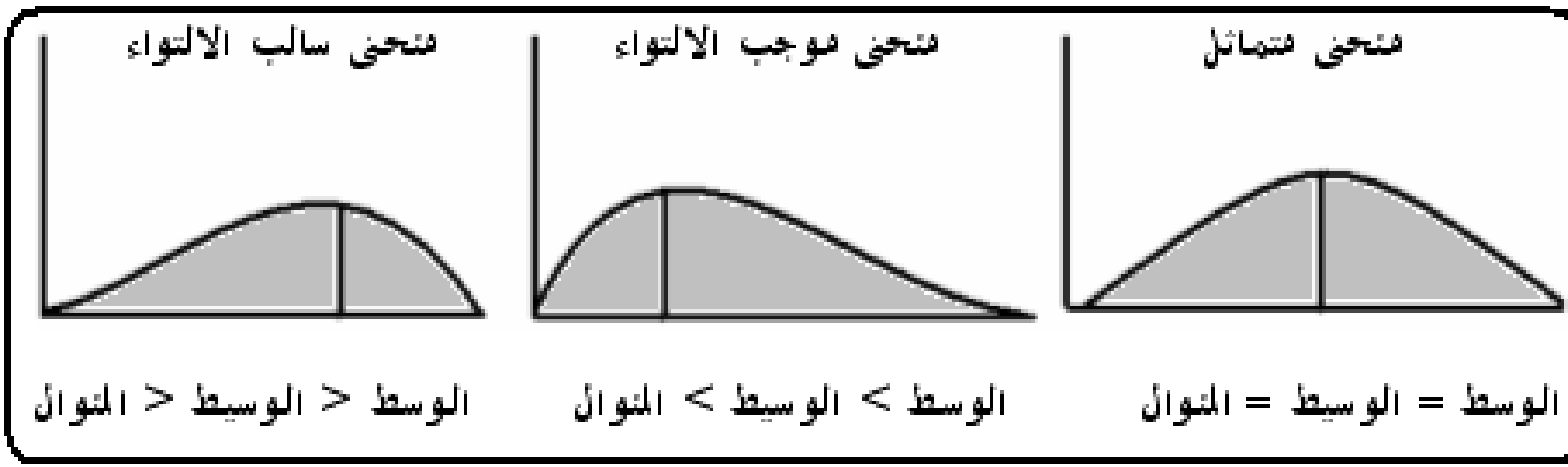
◆ وبتطبيق القانون نحسب المنوال كالتالي:

$$Mode = A + \frac{d_1}{d_1 + d_2} \times L$$

$$= 8 + \frac{3}{3+5} \times 3 = 8 + 1.125 = 9.125$$

استخدام مقاييس النزعة المركزية لتحديد شكل توزيع البيانات

- يمكن استخدام الوسط الحسابي والوسيط والمنوال في وصف المنحنى التكراري، والذي يعبر عن شكل توزيع البيانات:



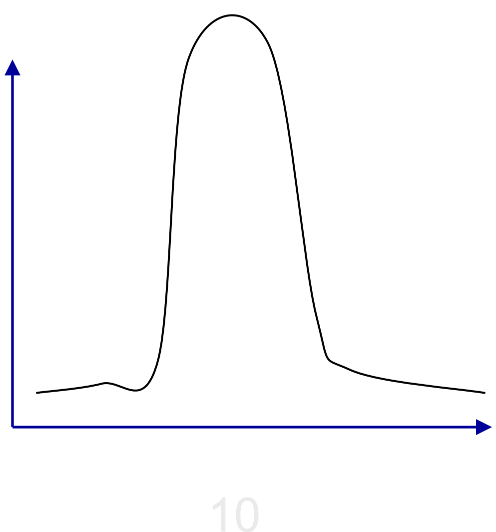
- الوسط يتأثر بالقيم المتطرفة لذا يكون اقرب الي "ذيل" منحني التوزيع—الوسيط أقل تأثرا (حساسية) لتلك القيم.

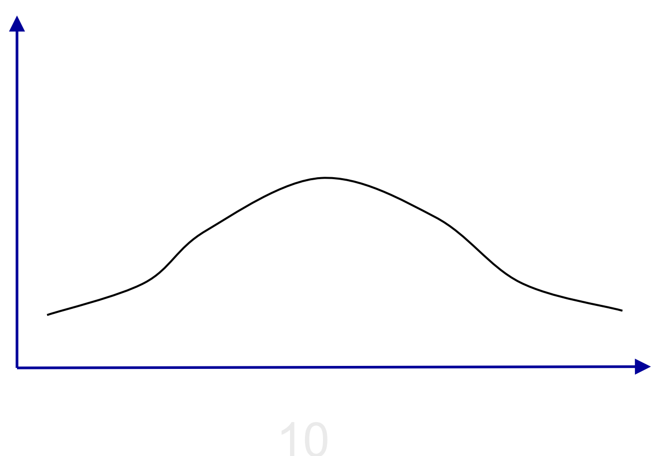
مقاييس التشتت
Measures of Dispersion

- مقدمة
- المدي والانحرافات
- التباين
- الانحراف المعياري

مقدمة

- نحتاج كثيرا الي مقارنة مجموعتين أو اكثر من البيانات
- يمكن استخدام شكل التوزيع التكراري، المنحنى التكراري ، وكذلك بعض مقاييس النزعة المركزية (؟؟؟؟)—كاف؟؟؟لا





المدي-
Range

- هو أبسط مقاييس التشتت ,وأيسرها حسابا (وحدة القياس؟)
- في حالة البيانات غير المبوبة:

المدى في حالة البيانات غير المبوبة = أكبر قراءة − أقل قراءة

$$Rang = Max - Min$$

(٤-١)

- في حالة البيانات المبوبة-عدة طرق منها:

المدى في حالة البيانات المبوبة = مركز الفئة الأخيرة − مركز الفئة الأولى

(٤-٢)

المدي-أمثلة

1. احسب المدى للبيانات التالية:

4.8	6.21	5.4	5.18	5.29	5.18	5.08	4.63	5.03

2. احسب المدى للبيانات التالية:

المساحة	15-20	20-25	25-30	30-35	35-40	40-45
عدد المزارع	3	9	15	18	12	3

مزايا وعيوب المدى

- من مزايا المدى:
  - أنه بسيط وسهل الحساب
  - يكثر استخدامه عند الإعلان عن حالات الطقس، مثل درجات الحرارة، والرطوبة، والضغط الجوي.
- ومن عيوبه:
  - أنه يعتمد على قيمتين فقط ، ولا يأخذ جميع القيم في الحسبان
  - يتأثر بالقيم الشاذة؟؟

التباين
Variance

- أكثر مقاييس التشتت استخداما في النواحي التطبيقية.
- يعبر عن متوسط مربعات انحرافات القيم عن وسطها الحسابي
- يمكن حساب:
  - التباين في المجتمع (σ^2)
  - التباين في العينة (s^2)

التباين في المجتمع (σ^2)

- افرض لدينا كل مفردات المجتمع:

$$(x_1, x_2, x_3, \ldots, x_n)$$

- التباين في المجتمع، يرمز له بالرمز (σ^2) يحسب باستخدام المعادلة:

$$\boxed{\sigma^2 = \frac{\sum (x-u)^2}{N}} \qquad (٤-٦)$$

- حيث (μ) هو الوسط الحسابي للمجتمع ويحسب من: $\mu = \sum x / N$

- مثال:
- التالي عدد سنوات الخبرة للمدربين ، عددهم 15 مدرب. بفرض أن هذه البيانات تم جمعها عن كل مفردات المجتمع أوجد التباين؟

5	13 7	14	12	9 6	8	10	13	14	6	11	12	10

الحل

- أولا نحسب الوسط الحسابي في المجتمع كالآتي:

$$\mu = \frac{1}{N}\sum x$$

$$= \frac{1}{15}(5+13++7+...+12+10) = \frac{1}{15}(150) = 10$$

- نكون الجدول التالي لحساب مربعات الانحرافات:

سنوات الخبرة x	$(x-\mu)$	$(x-\mu)^2$
5	5-10 = -5	25
13	3	9
7	-3	9
14	4	16
12	2	4
9	-1	1
6	-4	16
8	-2	4
10	0	0
13	3	9
14	4	16
6	-4	16
11	1	1
12	2	4
10	0	0
Sum 150	**0**	**130**

تابع

- ثم نحسب تباين سنوات الخبرة للمدربين في المجال الرياضي:

$$\sigma^2 = \frac{\sum(x-u)^2}{N} = \frac{130}{15} = 8.67(year^2)$$

- طريقة أخري لحل المثال:

تابع

- يمكن فك المجموع $\sum(x-\mu)^2$ كالتالي:

$$\sum(x-\mu)^2 = \sum\left(x^2 - 2x\mu + \mu^2\right)$$
$$= \sum x^2 - 2\mu\sum x + \sum\mu^2$$
$$= \sum x^2 - 2N\mu^2 + N\mu^2$$
$$= \sum x^2 - N\mu^2$$

- إذا التباين في المجتمع يمكن صياغته كالتالي:

$$\boxed{\sigma^2 = \frac{1}{N}\sum x^2 - \mu^2} \qquad (٤-٧)$$

تابع

◆ وبالتطبيق على المثال السابق:

سنوات الخبرة x	x^2
5	25
13	169
7	49
14	196
12	144
9	81
6	36
8	64
10	100
13	169
14	196
6	36
11	121
12	144
10	100
150	1630

$$\sum x = 150 \ , \ \sum x^2 = 1630$$

$$\mu = \frac{1}{N}\sum x = \frac{1}{15}(150) = 10$$

إذا التباين هو

$$\sigma^2 = \frac{1}{N}\sum x^2 - \mu^2$$

$$= \frac{1}{15}1630 - 10^2 = 108.67 - 100 = 8.67$$

وهي نفس النتيجة التي تم الحصول عليها باستخدام الصيغة (**6–4**) .

التباين في العينة

(s^2)

- غالبا يكون تباين المجتمع (σ^2) غير معلوم
- وعندئذ يتم سحب عينة من هذا المجتمع، ويحسب التباين من:

$$\boxed{s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}} \qquad (٨-٤)$$

مثال

- عينة حجمها 5 مديريين فنيين، عدد سنوات خبرتهم كالتالي, احسب التباين.

8	13	10	5	9

- نحسب الوسط الحسابي: $\bar{X}$ =9
- لحساب التباين تكون الجدول:

						sum
سنوات الخبرة X	8	13	10	5	9	45
$(x-\bar{x})$	-1	4	1	-4	0	0
$(x-\bar{x})^2$	1	16	1	16	0	34

تابع

- بالتالي:

$$s^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{34}{(5-1)} = \frac{34}{4} = 8.5$$

- كما يمكن حساب تباين العينة من:

$$\boxed{s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)} \qquad (٤-١٠)$$

التمرين (4)

- استخدم القانون أدناه لحساب التباين للمفردات التالية (سنوات الخبرة لعينة من المدربين بالسنوات)
- القانون

$$\boxed{s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right)} \qquad (١٠-٤)$$

- العينة:

23	17	22	20	21	17	22	19	18

الانحراف المعياري
Standard Deviation

- **من مقاييس التشتت**
- **يقاس بوحدات قياس المتغير محل الدراسة**
- **هو الجذر التربيعي الموجب للتباين** أي:

$$\boxed{\text{الانحراف المعيارى} = \sqrt{\text{التباين}}} \qquad (٤-١١)$$

تابع

- **من مثال سابق الانحراف المعياري لسنوات الخبرة لعمال المصنع (المجتمع):**

$$\sigma = \sqrt{\frac{1}{N}\sum x^2 - \mu^2}$$

$$= \sqrt{\frac{1}{15}1630 - 10^2} = \sqrt{8.67} = 2.94(\text{سنة})$$

الانحراف المعياري للبيانات المبوبة

- الانحراف المعياري يحسب بتطبيق المعادلة التالية :

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}}$$

or

$$s = \sqrt{\frac{\sum x^2 f - \frac{\left(\sum xf\right)^2}{n}}{n-1}} \qquad (١٢-٤)$$

- حيث كل مكونات القانون كما بينا سابقا.

الانحراف المعياري للبيانات المبوبة (مثال)

- من مثال سابق نحسب الانحراف المعياري كما يلي:

الإنفاق	عدد الأسر f	مركز الفئة X	xf	$x^2 f$
2-5	1	3.5	3.5	12.25
5-8	8	6.5	52	338
8-11	13	9.5	123.5	1173.25
11-14	10	12.5	125	1562.5
14-17	8	15.5	124	1922
sum	40		428	5008

$$n = \sum f = 40$$

$$\sum xf = 428$$

$$\sum x^2 f = 5008$$

تابع

- وبتطبيق المعادلة ، نجد أن الانحراف المعياري قيمته هي:

$$s = \sqrt{\frac{\sum x^2 f - \frac{\left(\sum xf\right)^2}{n}}{n-1}}$$

$$= \sqrt{\frac{5008 - \frac{(428)^2}{40}}{40-1}} = \sqrt{\frac{5008 - 4579.6}{39}}$$

$$= \sqrt{10.984615} = 3.314(000SR)$$

خصائص الانحراف المعياري

- الانحراف المعياري للمقدار الثابت يساوي صفرا
- إذا أضيف مقدار ثابت إلى كل قيمة لا يتأثر الانحراف الانحراف المعياري بذلك
- إذا ضرب كل قيمة من قيم المفردات في مقدار ثابت فإن الانحراف المعياري للقيم الجديدة ، يساوي الانحراف المعياري للقيم الأصلية مضروبا في الثابت
- إذا كان لدينا التوليفة الخطية : $y = aX + b$

فإن الانحراف المعياري للمتغير (y) يكون: $S_y = aS_x$

مثال

- مجموعتين من اللاعبين، تم استخدام طريقتين تغذية وبعد فترة تم جمع بيانات عن أوزان المجموعتين بالكيلوجرام ، وتم الحصول على المقاييس أدناه:المطلوب مقارنة درجة تشتت المجموعتين .

المجموعة الثانية	المجموعة الأولى	المقاييس
198	173	$\bar{X} =$
25	23	$S =$

تابع

- معامل الاختلاف النسبي للمجموعة الأولى:

$$CV_1 = \frac{S}{\overline{\overline{X}}} \times 100 = \frac{23}{173} \times 100 = 13.3\%$$

- معامل الاختلاف النسبي للمجموعة الثانية:

$$CV_2 = \frac{S}{\overline{\overline{X}}} \times 100 = \frac{25}{195} \times 100 = 12.8\%$$

- درجة تشتت أوزان المجموعة الثانية أقل قليلا من درجة تشتت أوزان المجموعة الأولى

الدرجة المعيارية
Standardized degree(z-score)

- تقيس الدرجة المعيارية لقيمة معينة عدد وحدات الانحراف المعياري التي تزيد أو تقل بها هذه القيمة عن الوسط الحسابي

- الدرجة المعيارية للقيمة (x_i) ويرمز لها بالرمز(z) تحسب باستخدام المعادلة التالية:

$$z = \frac{x_i - \bar{x}}{s}$$

مثال

- في المثال السابق اختير أحد اللاعبين من المجموعة الأولى بعد تطبيق البرنامج ، ووجد أن وزنه 178 كيلوجرام، وبالمثل أحد اللاعبين من المجموعة الثانية، ووجد أن وزنه 180 كيلوجرام ، قارن بين هذين القيمتين من حيث أهمية كل منها في المجموعة التي تنتمي إليها:

المجموعة الثانية	المجموعة الأولى	
198 25	173 23	$\bar{x} =$ $s =$
180	178	القيمة.

تابع الحل

- للمقارنة بين الوحدتين يتم حساب الدرجة المعيارية لوزن كل منها، بتطبيق المعادلة:
  - الدرجة المعيارية لوزن الوحدة المسحوبة من المجموعة الأولى (178 Kg.) هي :

$$z = \frac{x - \bar{x}}{s} = \frac{178 - 173}{23} = 0.22$$

  - الدرجة المعيارية لوزن الوحدة المسحوبة من المجموعة الثانية (180 Kg.) هي

$$z = \frac{x - \bar{x}}{s} = \frac{180 - 198}{25} = -0.75$$

تابع الحل

- من النتائج نجد أن الوزن 178 كيلوجرام يزيد عن الوسط الحسابي بـ 0.22 انحراف معياري ، بينما نجد أن الوزن 180 كيلوجرام يقل عن الوسط الحسابي بـ 0.75 انحراف معياري . ومن ثم في هذه الحالة الوزن الأول أهميته النسبية أعلى من الوزن الثاني.