اسم المقرر

مبادئ الإحصاء

د. سعيد سيف الدين

جامعة الملك فيصل

عمادة التعلم الإلكتروني والتعليم عن بعد

الحمد لله رب العالمين ، والصلاة والسلام على خاتم الأنبياء والمرسلين سيدنا ونبينا محمد بن عبد الله وعلى آله وصحبه أجمعين

المحاضرة الثامنة

[تابع] الباب الثالث
مقاييس النزعة المركزية

عناصر المحاضرة

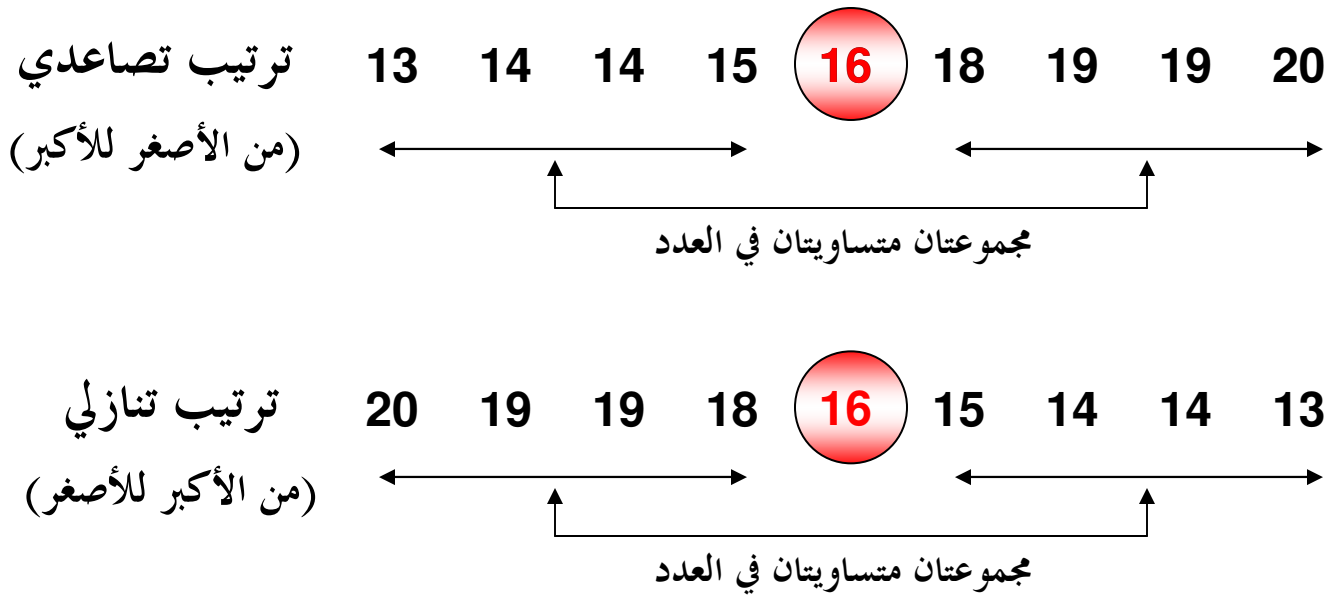
- الوسيط

تعريف الوسيط :

(ببساطة) يُعرف الوسيط [وسنرمز له بالرمز M] لمجموعة من القيم (المرتبة تصاعدياً أو تنازلياً حسب قيمها) على أنه القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد ، أو بتعبير آخر هي القيمة التي في المنتصف

فمثلاً لمجموعة القيم : 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها ٩ قيم] ، إذا قمنا بترتيبها تصاعدياً أو تنازلياً

يكون الوسيط هو العدد الخامس [رتبة الوسيط أي ترتيبه بين القيم] وقيمته 16

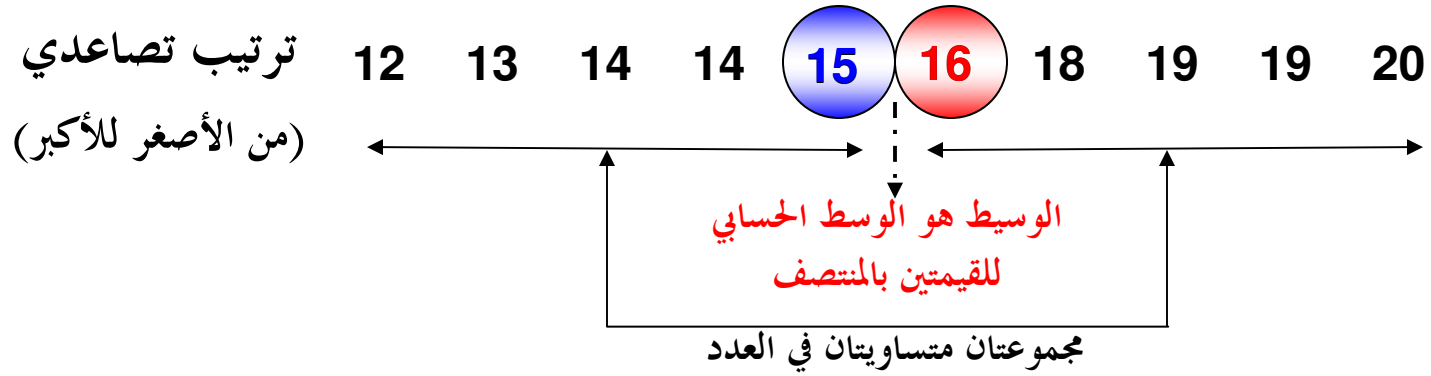


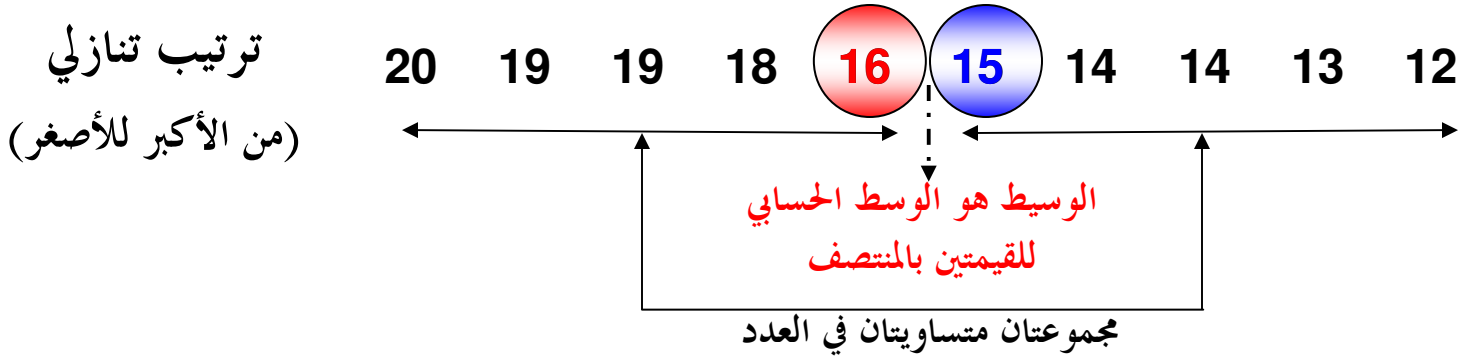
لاحظ هنا أن عدد القيم n [هنا = ٩] فردي وبالتالي هناك قيمة واحدة في منتصف المجموعة

هام جداً

فرق بين رتبة الوسيط وقيمته

أما لمجموعة القيم : 12 , 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها ١٠ قيم (أي زوجي) حيث أضفنا القيمة 12 للمجموعة السابقة] ، إذا قمنا بترتيبها تصاعدياً أو تنازلياً





في هذه الحالة توجد قيمتان بالمنتصف وهما القيمة الخامسة والقيمة السادسة [وهما العددان 16 , 15] ، عندئذ يكون الوسيط هو الوسط الحسابي لهاتين القيمتين ، أي :

$$\frac{15 + 16}{2} = 15.5$$

إذن من السابق يمكن استنتاج **طريقة حساب الوسيط لمجموعة من القيم** كالآتي :

- قم أولاً بترتيب البيانات **تصاعدياً** أو **تنازلياً** .
- حدد ما إذا كانت هناك قيمة واحدة بالمنتصف أم قيمتين ، وهذا يتوقف على قيمة n حيث n عدد القيم

فإذا كانت n فردية

كانت هناك قيمة **واحدة** في المنتصف رتبتها

$$\frac{n+1}{2}$$

وتكون هذه القيمة هي **الوسيط**

وإذا كانت n زوجية

كانت هناك **قيمتان** في المنتصف رتبتهما

$$\frac{n}{2} \quad , \quad \frac{n}{2}+1$$

ويكون الوسط الحسابي لهاتين القيمتين هو **الوسيط**

فمثلاً

$n = 5$

9	2	7	12	10
2	7	9	10	12

ترتيب تصاعدي

هناك قيمة **واحدة** في المنتصف رتبتها :

$$\frac{n+1}{2}=\frac{5+1}{2}=3$$

أي القيمة **الثالثة** . وتكون تلك القيمة هي الوسيط . أي أن :

الوسيط = 9

تذكر :

الوسط الحسابي لهذه القيم هو :

$$\frac{9+2+7+12+10}{5}=8$$

فمثلاً

$n = 6$

9	2	7	12	10	500
2	7	9	10	12	500

ترتيب تصاعدي

هناك قيمتان في المنتصف رتبتهما :

$$\frac{n}{2}=\frac{6}{2}=3 \quad , \quad 3+1=4$$

أي القيمتان **الثالثة** و**الرابعة** ، وتكون قيمة **الوسيط** هي الوسط الحسابي لهاتين القيمتين ، أي :

$$\frac{9+10}{2}=9.5$$

الوسط الحسابي لهذه القيم هو

$$\frac{9+2+7+12+10+500}{6}=90$$

وواضح تأثره كثيراً بالقيمة المتطرفة 500

هل لاحظت أن الوسيط لم يتأثر بالقيمة المتطرفة 500

مثال : مجموعة الأرقام 9 7 6 6 **5** 4 3 3 2 وسيطها هو **5** [عدد القيم n = 9 (فردي)]

تذكر : الوسط الحسابي لهذه القيم هو : $\bar{x} = \frac{2+3+3+4+5+6+6+7+9}{9} = 5$

مثال آخر : مجموعة الأرقام 18 15 12 **11** **9** 7 5 5 وسيطها هو $\frac{9+11}{2} = 10$ [n = 8 (زوجي)]

تذكر : الوسط الحسابي لهذه القيم هو : $\bar{x} = \frac{5+5+7+9+11+12+15+18}{8} = 10.25$

في السؤال [سلي نفسك – المحاضرة السابعة – شريحة ١٤ – س١] : كانت درجات طالب في ٦ اختبارات هي :

84 , 91 , 72 , 68 , 87 , 78

وطلبنا حساب الوسط الحسابي للدرجات ، أضف لهذا حساب وسيط هذه الدرجات ، وحدد أيهما تفضل (كمتوسط) ولماذا ؟

الوسط الحسابي لدرجات الطالب هو : $\bar{x} = \frac{\sum x}{n} = \frac{84+91+72+68+87+78}{6} = \frac{480}{6} = 80$

ولتحديد الوسيط لابد أولاً من ترتيب القيم (تصاعدياً مثلاً) :

68 , 72 , 78 , 84 , 87 , 91

وحيث أن عدد القيم زوجي ، إذن هناك قيمتان في المنتصف [هما 84 , 78] وسطهما الحسابي 81 $= \frac{78+84}{2}$ = الوسيط

لاحظ في السؤال السابق أن كلاً من المتوسطين : الوسط الحسابي و الوسيط من السهل حسابهما ومن الممكن أن يمثل كل منهما مقياساً للنزعة المركزية للبيانات ، لكن الأفضل (نسبياً هنا) أن نستخدم الوسط الحسابي كمقياس للنزعة المركزية للبيانات حيث أنه يأخذ في الاعتبار جميع قيم البيانات ، بينما يهتم الوسيط بقيم البيانات في المنتصف (وذلك بعد ترتيبها) .

مثال : الأجر (بالريال) في الساعة لخمسة عاملين في مكتب هو : 37 , 92 , 32 , 39 , 25 . احسب الوسط الحسابي للأجور ووسيط هذه الأجور . أيهما تفضل كمقياس لمتوسط أجر الساعة ؟ ولماذا ؟

الوسط الحسابي للأجور هو : $\bar{x} = \frac{\sum x}{n} = \frac{25+39+32+92+37}{5} = \frac{225}{5} = 45$

أما لتحديد الوسيط ، فلابد أولاً من ترتيب القيم (تصاعدياً مثلاً) : 25 , 32 , 37 , 39 , 92

وحيث أن عدد القيم فردي ، إذن هناك قيمة واحدة في المنتصف [هي 37] وهي الوسيط

لاحظ في السؤال السابق أن الوسط الحسابي (بالرغم من عدم احتياجه لترتيب القيم وفي نفس الوقت يأخذ في الاعتبار جميع قيم البيانات) إلا أنه تأثر جداً بالقيمة المتطرفة 92 ، في حين لم يتأثر بها الوسيط لأنه يعتمد على البيانات في المنتصف . لذا يُفضل هنا استخدام الوسيط كمقياس للنزعة المركزية حيث يعطي دلالة أفضل لمتوسط الأجور من الوسط الحسابي .

والآن ماذا عن الوسيط لبيانات كمية متصلة

أعتقد أننا نستطيع تحديده بسهولة [حيث نوهنا لذلك في الباب الثاني] فهو :

- قيمة المتغير المناظرة لنقطة تقاطع المضلعين : المتجمع الصاعد والمتجمع الهابط للبيانات .

أو

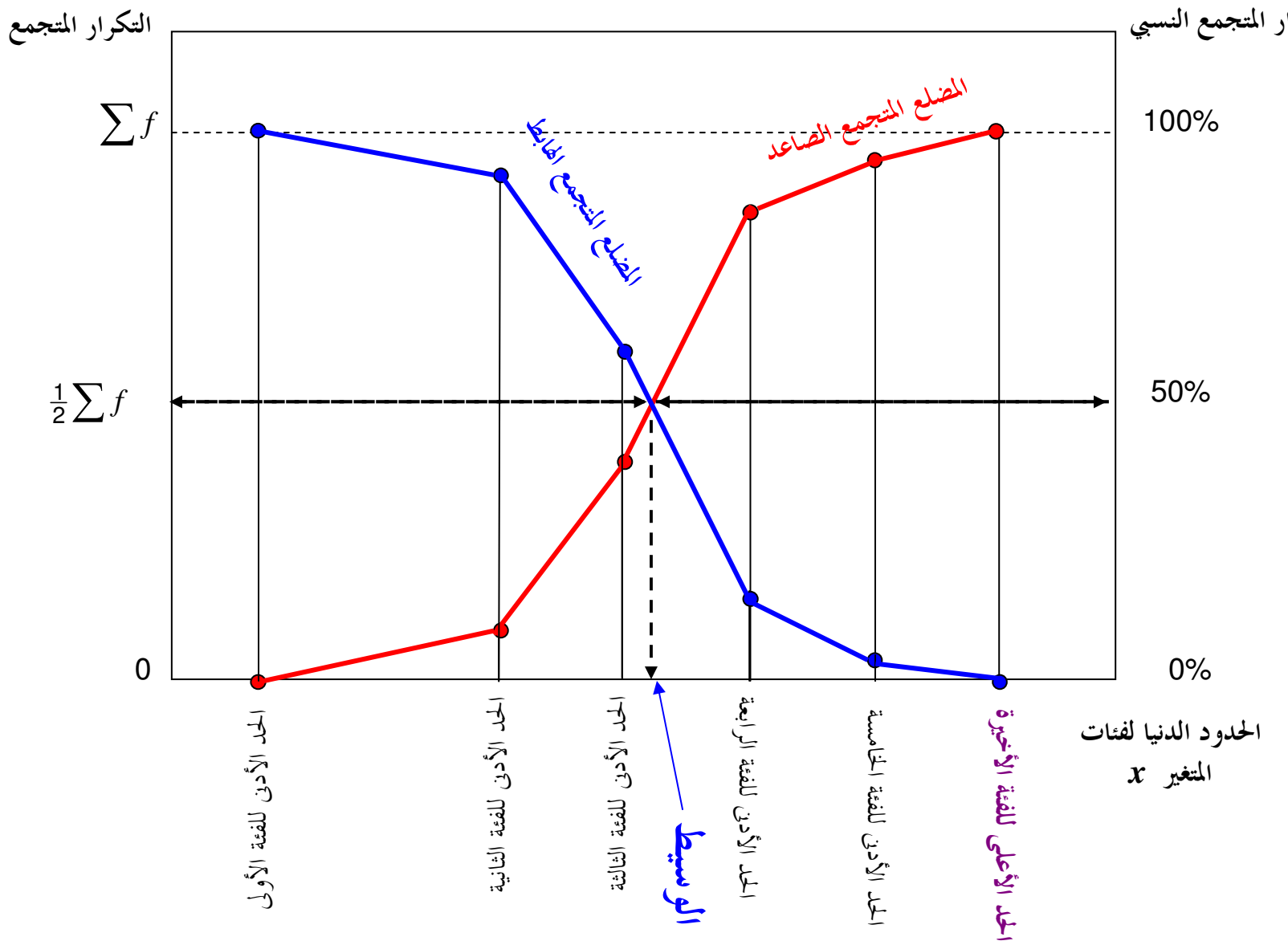
- القيمة التي يناظرها تكرار متجمع = نصف مجموع التكرارات

أو

- القيمة التي يناظرها تكرار نسبي متجمع = 50%

طريقة تحديد الوسيط من :

* من المضلع المتجمع الصاعد فقط
* من المضلع المتجمع الهابط فقط .
* من المضلعين معاً



مثال : في دراسة جغرافية لعدد من مساحات مجموعة من الأراضي لمنطقة سكنية بالرياض تبين أن التوزيع التكراري لها كما هو مبين .

المطلوب حساب الوسط الحسابي والوسيط لمساحة الأراضي .

المساحة (بالفدان)	عدد قطع الأراضي
1 -	14
3 -	29
5 -	18
7 - 10	9

المتغير x هنا هو مساحة الأرض (بالفدان) ، في حين يمثل عدد قطع الأراضي التكرار f .

أولاً : الوسط الحسابي : نستكمل الجدول التكراري كما هو مبين :

الجدول التكراري

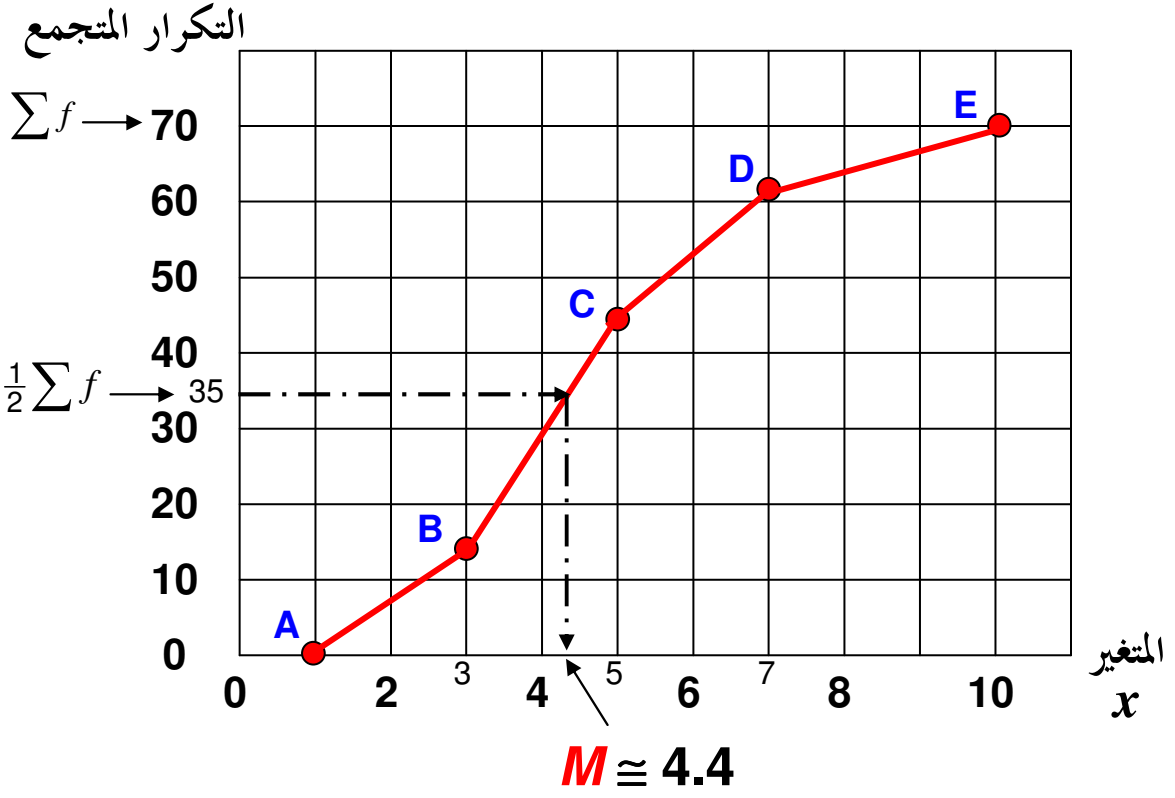
الفئة	المتغير (المساحة) x	التكرار f	المركز x_0	$f\,x_0$
الأولى	$1 \le x < 3$	14	2	28
الثانية	$3 \le x < 5$	29	4	116
الثالثة	$5 \le x < 7$	18	6	108
الرابعة	$7 \le x < 10$	9	8.5	76.5
		$\sum f = 70$		$\sum f\,x_0 = 328.5$

$$\therefore \quad \bar{x} = \frac{\sum f\,x_0}{\sum f} = \frac{328.5}{70} = 4.692857143 \cong \underline{4.7}$$

ثانياً : الوسيط : نكون الجدول التكراري المتجمع الصاعد [أو الهابط]

الجدول التكراري		
الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

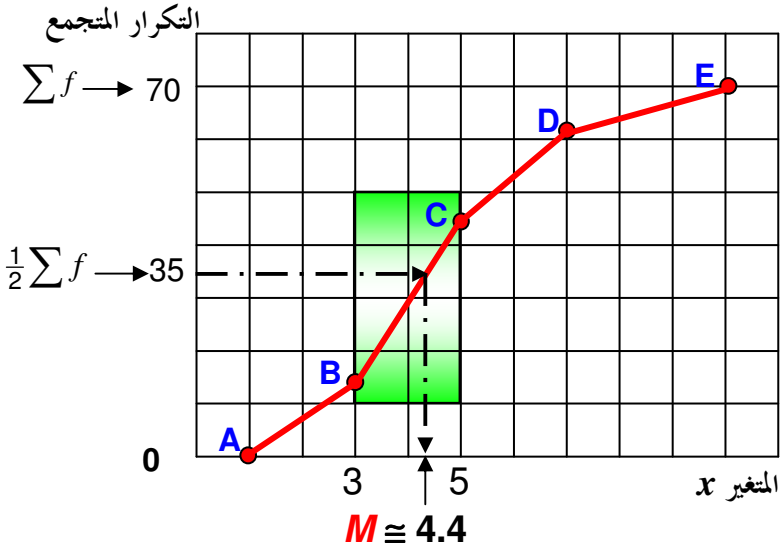
الجدول التكراري المتجمع الصاعد		
المتغير x	التكرار المتجمع	النقطة على المضلع
< 1	0	A (1 , 0)
< 3	14	B (3 , 14)
< 5	43	C (5 , 43)
< 7	61	D (7 , 61)
< 10	$\sum f = 70$	E (10 , 70)



كده انتهى السؤال ، أي أن الوسط الحسابي للمساحة $\cong$ 4.7

والوسيط $\cong$ 4.4

لكن هناك ملحوظة هامة جداً مش عارف أنت لاحظتها أم لا



الوسيط يقع بين النقطتين B(3 , 14) , C(5 , 43) [أي داخل الفئة $3 \le x < 5$] هذه الفئة تُسمى بـــ **الفئة الوسيطية**

أي أن الفئة الوسيطية هي تلك الفئة التي يقع داخلها الوسيط

وهنا يتبادر إلى الذهن سؤالان هامان :

السؤال الأول : هل من الممكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً أم لازم نعمل الجدول التكراري المتجمع الصاعد ونرسم المضلع التكراري المتجمع الصاعد ؟ .

السؤال الثاني : هل من الممكن [بعد تحديد الفئة الوسيطية] تحديد الوسيط من الجدول التكراري مباشرة دون الحاجة للجدول التكراري المتجمع الصاعد أو المضلع التكراري المتجمع الصاعد ؟ .

والإجابة على السؤالين : نعم يمكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً ، ثم بعد ذلك يمكن أيضاً من هذا الجدول التكراري تحديد قيمة الوسيط دون أن نحتاج لعمل جدول تكراري متجمع صاعد ورسم المضلع التكراري المتجمع الصاعد ، **وتعالي نشوف إزاي**

بالنسبة للسؤال الأول [تحديد الفئة الوسيطية]

(١) احسب أولاً نصف مجموع التكرارات .

(٢) ابدأ بالرقم صفر في ذهنك وزود تكرارات الفئات على التوالي وكل مرة قارن بنصف مجموع التكرارات السابق . أول ما يزيد الناتج عن نصف المجموع السابق أو يساويه يبقى آخر فئة زودنا تكرارها تكون هي الفئة الوسيطية .

يالله نشوف

الجدول التكراري		
الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

- احسب $\frac{1}{2}\sum f$: $\frac{1}{2}\sum f = \frac{70}{2} = 35$ ← ماشي يا عم .. طلع
- نبدأ بالصفر [في ذهننا]
- نزود على الصفر السابق تكرار الفئة الأولى [14] ينتج 14

14 أقل من 35 ، يبقى الفئة الأولى ليست الفئة الوسيطية

- نزود على الـ 14 الأخيرة تكرار الفئة الثانية [29] ينتج 43

43 أكبر من 35 ، يبقى الفئة الثانية هي الفئة الوسيطية

وبالنسبة للسؤال الثاني [تحديد الوسيط (بعد ما حددنا الفئة الوسيطية)]

(١) حدد الحد الأدنى للفئة الوسيطية وأيضاً طولها .

(٢) احسب ما يُسمى بـــ "التكرار المتجمع السابق" = مجموع تكرار الفئات السابقة للفئة الوسيطية

(٣) احسب الوسيط من العلاقة :

$$\text{الوسيط } M = \text{الحد الأدنى للفئة الوسيطية} + \left[\frac{\text{(نصف مجموع التكرارات} - \text{التكرار المتجمع السابق)}}{\text{تكرار الفئة الوسيطية}} \times \text{طول الفئة الوسيطية} \right]$$

يالله نشوف

الجدول التكراري

الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 \le x < 3$	14
الثانية ← الفئة الوسيطية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

- الفئة الوسيطية هي الفئة الثانية :

حدها الأدنى 3 وطولها 2 [= 3 – 5] وتكرارها 29

- التكرار المتجمع السابق :

يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطية [أي تكرار الفئة الأولى فقط] = 14

• بالتعويض في القانون السابق : [ونعمل الحسابات واحدة واحدة الله يسترها معاكم]

$$M = 3 + \left[\frac{35 - 14}{29} \times 2\right] = 3 + \left[\frac{21}{29} \times 2\right] = 3 + 1.44827 = 4.44827 \cong 4.4$$

تُسمى الطريقة الحسابية السابقة (لحساب الوسيط) بـــ "طريقة الاستكمال"

مثال جميل : طُلب من ٣ مشرفين بإحدى المدارس تقسيم طلبة المدرسة إلى ٣ مجموعات متساوية على أن يقوم كل مشرف بتقديم بيان عن فئات العمر المختلفة لطلبة مجموعته وعدد الطلبة في كل فئة من فئات العمر ، فكانت الجداول التكرارية المبينة :

المجموعة (١)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$15 \le x < 18$	18
		$\sum f = 98$

المجموعة (٢)		
الفئة	العمر x	العدد f
الأولى	$6 \le x < 12$	20
الثانية	$12 \le x < 15$	25
الثالثة	$15 \le x < 18$	35
الرابعة	$x \ge 18$	18
		$\sum f = 98$

المجموعة (٣)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$x \ge 15$	18
		$\sum f = 98$

هل يمكن من خلال البيانات السابقة حساب الوسط الحسابي لعمر الطلبة في كل مجموعة ؟ علل إجابتك . وإذا كانت الإجابة بـــ "لا" احسب مقياساً مناسباً يُعطي دلالة لمتوسط العمر في كل مجموعة .

المجموعة (١)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$15 \le x < 18$	18
		$\sum f = 98$

المجموعة (٢)		
الفئة	العمر x	العدد f
الأولى	$6 \le x < 12$	20
الثانية	$12 \le x < 15$	25
الثالثة	$15 \le x < 18$	35
الرابعة	$x \ge 18$	18
		$\sum f = 98$

المجموعة (٣)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$x \ge 15$	18
		$\sum f = 98$

الإجابة هي " لا " للمجموعات الثلاث [أي لا يمكن حساب الوسط الحسابي للعمر] ، وها هي الأسباب :

- في المجموعة الأولى : الحد الأدنى للفئة الأولى غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من أسفل]
- في المجموعة الثانية : الحد الأعلى للفئة الأخيرة غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من أعلى]
- في المجموعة الثالثة : الحد الأدنى للفئة الأولى غير معروف وأيضاً الحد الأعلى للفئة الأخيرة غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من الطرفين]

مثل هذه الجداول تُسمى جداول تكرارية مفتوحة :

- من أسفل [إذا كان الحد الأدنى للفئة الأولى غير معروف]
- من أعلى [إذا كان الحد الأعلى للفئة الأخيرة غير معروف]
- من الطرفين [إذا كان الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة غير معروفين]

وحيث أن الوسيط لأي مجموعة من البيانات يعتمد في حسابه على البيانات الموجودة في المنتصف ، إذن يمكن استخدام الوسيط كمتوسط للدلالة على متوسط العمر في كل مجموعة :

بالنسبة للمجموعة الأولى من الطلبة :

المجموعة (١)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$15 \le x < 18$	18
		$\sum f = 98$

تحديد الفئة الوسيطية :

- احسب $\frac{1}{2}\sum f$: $\frac{1}{2}\sum f = \frac{98}{2} = 49$
- نبدأ بالصفر [في ذهننا]
- نزود على الصفر تكرار الفئة الأولى [20] ينتج 20 [أقل من 49]
- نزود على الـ 20 الأخيرة تكرار الفئة الثانية [25] ينتج 45 [أيضاً أقل من 49]
- نزود على الـ 45 الأخيرة تكرار الفئة الثالثة [35] ينتج 80 [أكبر من 49]

إذن الفئة الثالثة هي الفئة الوسيطية

تحديد الوسيط :

- الحد الأدنى للفئة الوسيطية = 12
- طول الفئة الوسيطية = 3 [= 12 – 15]
- تكرار الفئة الوسيطية = 35
- التكرار المتجمع السابق = مجموع تكرارات الفئات السابقة للفئة الوسيطية

[أي الفئتين الأولى والثانية] = 20 + 25 = 45 . إذن

المجموعة (١)

الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة ← الفئة الوسيطية	$12 \le x < 15$	35
الرابعة	$15 \le x < 18$	18
		$\sum f = 98$

$$M = 12 + \left[\frac{(49-45)}{35} \times 3\right] = 12 + \left[\frac{4}{35} \times 3\right] = 12 + 0.342857 = 12.342857 \cong \underline{\underline{12.3}}$$

وبنفس الطريقة يمكن التعامل مع المجموعتين (٢) ، (٣) ، وعليك التأكد من صحة الحل

المجموعة (٢)

الفئة	العمر x	العدد f
الأولى	$6 \le x < 12$	20
الثانية	$12 \le x < 15$	25
الثالثة ← الفئة الوسيطية	$15 \le x < 18$	35
الرابعة	$x \ge 18$	18
		$\sum f = 98$

$$M = 15 + \left[\frac{(49-45)}{35} \times 3\right] \cong \underline{\underline{15.3}}$$

المجموعة (٣)

الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة ← الفئة الوسيطية	$12 \le x < 15$	35
الرابعة	$x \ge 15$	18
		$\sum f = 98$

$$M = 12 + \left[\frac{(49-45)}{35} \times 3\right] \cong \underline{\underline{12.3}}$$

سلي نفسك بهذا السؤال : سبق وحسبنا الوسط الحسابي لأطوال سيقان الأزهار [مثال (٢-٤)/محاضرة ٧/ص١٣] وكان 31.7 تقريباً وحسبنا الوسيط (بيانياً) للأطوال وكان 32 تقريباً . أيضاً سبق وحسبنا الوسط الحسابي للأجر السنوي لمجموعة من العمال [مثال (٢-٦)/محاضرة ٧/ص١٣] وكان 83.75 تقريباً وحسبنا الوسيط (بيانياً) للأجور وكان 80 تقريباً . والآن مطلوب من سعادتك حساب الوسيط للمثالين بطريقة الاستكمال السابقة ومقارنة الحلول ببعضها .

لمساعدتك على الحل وتنظيم تفكيرك قم باستكمال البيانات الناقصة في المثالين

مثال (٢-٤)

الجدول التكراري		
الفئة	المتغير x	التكرار f
الأولى	$0 \le x < 20$	4
الثانية	$20 \le x < 30$	16
الثالثة	$30 \le x < 35$	12
الرابعة	$35 \le x < 40$	10
الخامسة	$40 \le x < 50$	6
السادسة	$50 \le x < 60$	2
		$\sum f = 50$

الفئة الوسيطية هي : .....

حدها الأدني ..... وطولها هو ..... وتكرارها .....

التكرار المتجمع السابق = ...

إذن الوسيط M [وتحسبه] يطلع 32.1 تقريباً

مثال (٢-٦)

الجدول التكراري		
الفئة	المتغير (الأجر) x	التكرار f
الأولى	$50 \le x < 60$	6
الثانية	$60 \le x < 70$	9
الثالثة	$70 \le x < 80$	15
الرابعة	$80 \le x < 90$	12
الخامسة	$90 \le x < 100$	9
السادسة	$100 \le x < 120$	6
السابعة	$120 \le x < 180$	3
		$\sum f = 60$

الفئة الوسيطية هي : .....

حدها الأدني ..... وطولها هو ..... وتكرارها .....

التكرار المتجمع السابق = ...

إذن الوسيط M [وتحسبه] يطلع 80 بالضبط

تمت بحمد الله