المحاضرة الثالثة

المقاييس الإحصائية للبيانات المبوبة

أولا: مقاييس النزعة المركزية (المتوسط - الوسيط - المنوال)

(١) مقاييس النزعة المركزية

المتوسط هو قيمة نموذجية يمكن أن تمثل مجموعة من البيانات بحيث تعطي دلالة معينة لتلك البيانات ، بمعنى أنه عندما ينظر الباحث (أو القارئ لتلك البيانات) ويريد أن يبحث عن شيء يربط هذه البيانات فإن تلك المتوسطات يمكن أن تعطيه بعضاً مما يريده .

وحيث أن مثل هذه القيم (المتوسطات) تميل إلى الوقوع في المركز داخل مجموعة البيانات (عند ترتيبها حسب قيمها) ، فإن هذه المتوسطات تُسمى أيضاً **بمقاييس النزعة المركزية** .

وهناك صور عديدة من هذه المقاييس وإن كان الأكثر شيوعاً :

- الوسط الحسابي (أو باختصار الوسط) .
- الوسيط
- المنوال (أو الشائع)

وغيرها ، وكل منها له مميزاته وعيوبه وهذا يعتمد على البيانات والهدف من استخدامه .

وإلى جانب كونه ممثلاً لمجموعة البيانات يجب أيضاً أن تتوافر في المتوسط عدة شروط ، منها :

- أن يمكن تحديد قيمته بالضبط وتكون عملية حسابه سهلة إلى حد كبير .
- أن يأخذ في الاعتبار جميع البيانات .

ومن الجدير بالذكر أن بعض هذه المقاييس يمكن تحديدها حسابياً بسهولة ، وبعضها يمكن تحديدها بيانياً بسهولة ، والبعض يمكن تحديده حسابياً وبيانياً بسهولة ، لكننا في هذا المقرر سنكتفي بالطريقة الأبسط (للطالب) عند تحديد هذه المقاييس ، وهذه الطريقة الأبسط ستختلف من مقياس لآخر .

(٢) أهمية حساب مقاييس النزعة المركزية

عند معرفتنا بتلك المتوسطات (مقاييس النزعة المركزية) يصبح أمامنا فرصة كبيرة لأن :

- ننظر لمتوسط مجموعة من البيانات لنعرف الكثير عن خصائص تلك المجموعة .
- نعقد مقارنة بين عدة مجموعات من البيانات في وقت واحد وذلك من خلال مقارنة متوسطات تلك المجموعات بعضها ببعض .

الوسط الحسابي

(١) تعريف الوسط الحسابي

يُعرف الوسط الحسابي [وسنرمز له بالرمز $\bar{x}$] لمجموعة من البيانات $x_1, x_2, \ldots, x_n$ [قيم المتغير x وعددها n] كالآتي :

الوسط الحسابي = مجموع قيم البيانات / عددها

أي

$$\bar{x} = \frac{\sum x}{n}$$

س ١ : درجات خمسة طلاب في مقرر ما [الدرجة العظمى 20] هي : **10 , 12 , 7 , 2 , 9** . أوجد الوسط الحسابي لدرجاتهم .

ج ١ :

$$\bar{x} = \frac{\sum x}{n} = \frac{9+2+7+12+10}{5} = \frac{40}{5} = 8$$

من هذا المثال **البسيط** يمكن ملاحظة **الخصائص العامة** التالية للوسط الحسابي :

- يمكن تحديد قيمة الوسط الحسابي بالضبط، كما أن طريقة تحديده سهلة .
- يأخذ في الاعتبار جميع البيانات .
- لا يتأثر بترتيب البيانات .
- لا يُشترط أن يكون الوسط الحسابي عدداً صحيحاً ولا يُشترط أن يكون إحدى قيم البيانات ولكنه قيمة تقع بين أقل قيمة في البيانات وأكبر قيمة فيها .
- يتأثر بالقيم المتطرفة في البيانات [كما يتضح من السؤالين التاليين] .

س ٢ : احسب الوسط الحسابي للقيم :

10 , 15 , 12 , 13 , 9

ج ٢ :

$$\frac{10+15+12+13+9}{5}=\frac{59}{5}=11\overline{8}$$

س ٣ : احسب الوسط الحسابي للقيم :

10 , 15 , 12 , 13 , 900

ج ٣ :

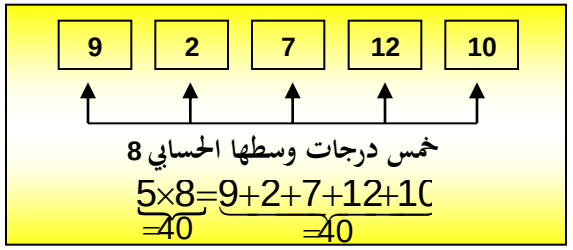
$$\frac{10+15+12+13+900}{5}=\frac{950}{5}=19($$

- حاصل ضرب قيمة الوسط الحسابي في عدد البيانات = مجموع قيم البيانات

وهذا واضح من تعريف الوسط الحسابي :

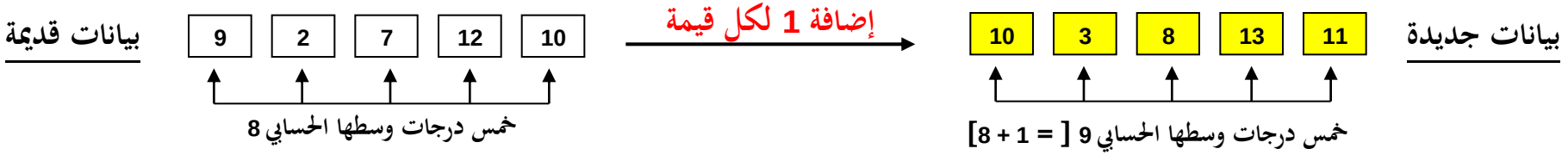
$$\bar{x} = \frac{\sum x}{n} \quad \xrightarrow[\text{أن}]{\text{تعني}} \quad n \times \bar{x} = \sum x$$

فمثلاً



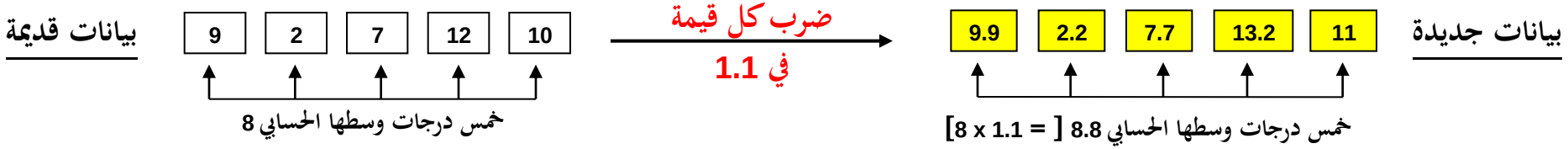
- إذا أضفنا عدد ثابت c لكل قيمة من قيم البيانات ، فإن :

الوسط الحسابي الجديد = الوسط الحسابي القديم + العدد الثابت c



- إذا ضربنا كل قيمة من قيم البيانات في عدد ثابت c ، فإن :

الوسط الحسابي الجديد = الوسط الحسابي القديم x العدد الثابت c



اعتبر نفسك مدرساً للطلاب الخمسة المذكورين في س ١ [كانت درجاتهم (من 20) كالتالي : 12 , 7 , 2 , 9 , 10 ,] وأردت أن تحسن من الوسط الحسابي لدرجاتهم ، أيهما أفضل : أن نزيد درجة كل طالب 5 درجات أم نزيد درجة كل طالب 50% من قيمتها ؟ علل إجابتك .

(٢) حساب الوسط الحسابي لبيانات كمية متقطعة ذات تكرارت

س : أوجد الوسط الحسابي للأرقام :

5 , 5 , 5 , 5 , 5 , 5 , 3 , 3 , 6 , 6 , 4 , 4 , 4 , 4 , 4 , 2 , 2 , 8 , 8 , 8

ج : بتطبيق مباشر للتعريف :

$$\bar{x}=\frac{(5+5+5+5+5+5)+(3+3)+(6+6)+(4+4+4+4+4)+(2+2)+(8+8+8)}{20}=\frac{96}{20}=4.8$$

لاحظ أن الرقم 5 متكرر 6 مرات ، الرقم 3 مرتان ، والرقم 6 مرتان ، والرقم 4 متكرر 5 مرات ، والرقم 2 مرتان ، والرقم 8 ثلاث مرات ، وبالتالي يمكن عمل العملية الحسابية السابقة كالآتي :

$$\bar{x}=\frac{(6 \times 5)+(2 \times 3)+(2 \times 6)+(5 \times 4)+(2 \times 2)+(3 \times 8)}{6+2+2+5+2+3}$$

$$=\frac{30+6+12+20+4+24}{20}=\frac{96}{20}=\underline{\underline{4.8}}$$

وهذا يمكن إنجازه بيسر من خلال الجدول التكراري للبيانات كالتالي :

$$\frac{30+6+12+20+4+24}{20}=\frac{96}{20}$$

نعمل هذا العمود : حاصل ضرب كل قيمة في تكرارها

x

المتغير x	التكرار f	fx
5	6	30
3	2	6
6	2	12
4	5	20
2	2	4
8	3	24
	20	96
	$\sum f=20$	$\sum fx=96$

هذا الجدول التكراري المعطى لنا [مُعطى أو نعمله]

$\therefore$

$$x=\frac{\sum fx}{\sum f}=\frac{96}{20}=4.8$$

أي أنه في حالة البيانات الكمية المتقطعة ذات التكرارات يمكن حساب الوسط الحسابي من العلاقة :

$$x=\frac{\sum fx}{\sum f}$$

حيث $\sum f$ هو مجموع التكرارات

$\sum fx$ هو مجموع حاصل ضرب كل قيمة في تكرارها

س: من مائة رقم يتكرر الرقم 4 عشرون مرة، والرقم 5 أربعون مرة، والرقم 6 ثلاثون مرة، والباقي كانوا الرقم 7 احسب الوسط الحسابي للمائة رقم .

ج : بتكوين الجدول التكراري للأرقام المذكورة ، ثم بضرب كل قيمة في تكرارها والتجميع [عمود fx] يكون الوسط الحسابي للأرقام المذكورة هو :

الجدول التكراري		
المتغير x	التكرار f	fx
4	20	80
5	40	200
6	30	180
7	10	70
	100	530
	$\sum f=100$	$\sum fx=530$

$$\bar{x}=\frac{\sum fx}{\sum f}=\frac{530}{100}=5.3$$

(٣) حساب الوسط الحسابي لبيانات كمية متصلة

عندما نتعامل مع بيانات متصلة تُعطى فيها قيم المتغير على صورة فترات، فيمكن اعتبار أن جميع القيم داخل الفترة مطابقة لمركز الفئة ، وبالتالي يمكن استخدام الصيغة السابقة لحساب الوسط الحسابي :

$$\bar{x} = \frac{\sum f x_0}{\sum f}$$

حيث $\sum f$ هو مجموع التكرارات ، $\sum f x_0$ هو مجموع حاصل ضرب مركز كل فئة في تكرار الفئة

ففي المثال التالي والذي يوضح اطوال سيقان الزهار بالسنتيمتر، يكون الوسط الحسابي لأطوال سيقان الأزهار هو:

الجدول التكراري

الفئة	المتغير x (الطول)	التكرار f	مركز الفئة x_0	fx_0
الأولى	$0 \le x < 20$	4	10	40
الثانية	$20 \le x < 30$	16	25	400
الثالثة	$30 \le x < 35$	12	32.5	390
الرابعة	$35 \le x < 40$	10	37.5	375
الخامسة	$40 \le x < 50$	6	45	270
السادسة	$50 \le x < 60$	2	55	110
		$\sum f = 50$		$\sum f x_0 = 1585$

$$x = \frac{\sum f x_0}{\sum f} = \frac{1585}{50} = 31.7$$

مزايا وعيوب الوسط الحسابي

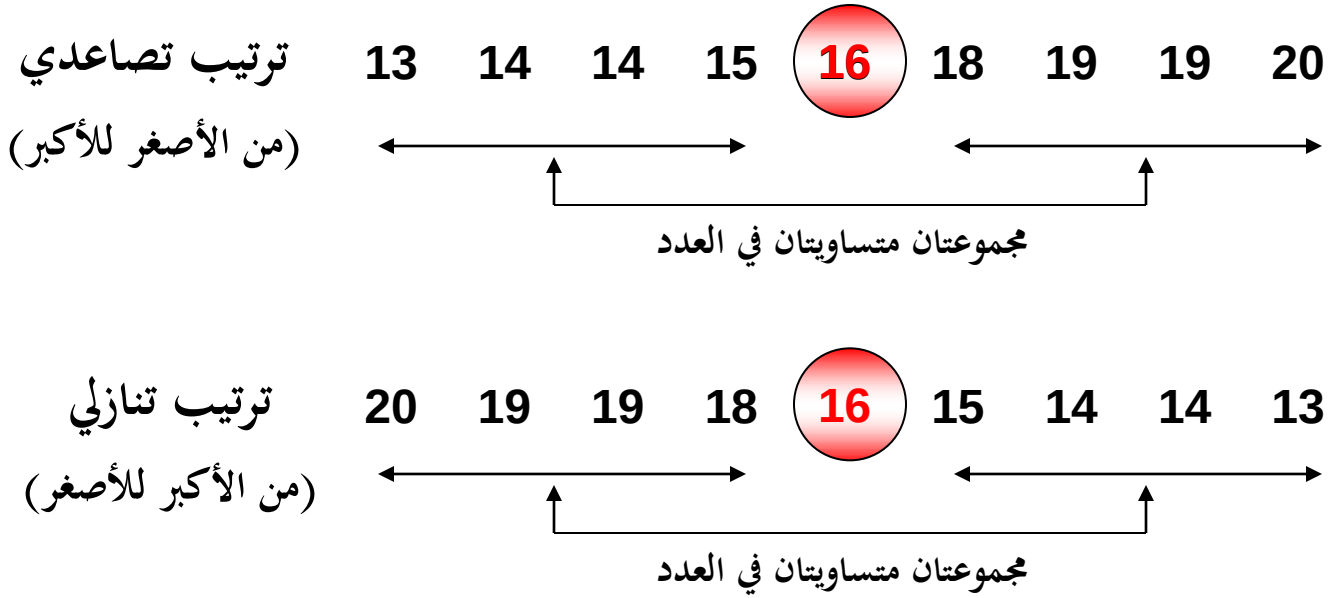
من كل ما سبق يمكن استعراض مزايا وعيوب الوسط الحسابي كالتالي :

- يمكن تحديد قيمة الوسط الحسابي بالضبط ، كما أن طريقة تحديده سهلة [ميزة].
- يأخذ في الاعتبار جميع البيانات [ميزة] .
- لا يتأثر بترتيب البيانات [ميزة] .
- يتأثر بالقيم المتطرفة في البيانات [عيب] .
- لا يمكن حسابه بالرسم ، أي بيانياً [عيب] .

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 أي رقم فردي] ، إذا قمنا بترتيبها تصاعدياً أو تنازلياً



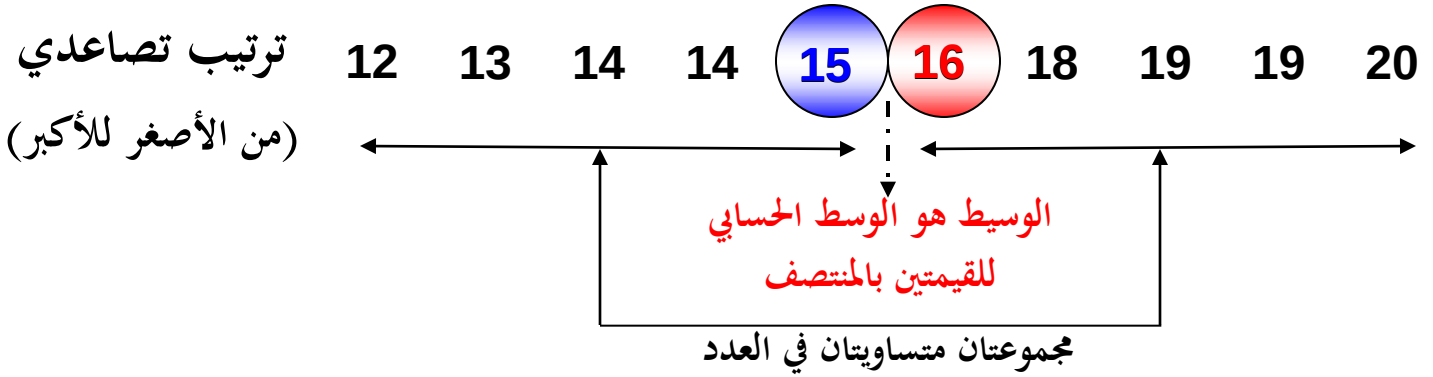
يكون الوسيط هو العدد الخامس [رتبة الوسيط أي ترتيبه بين القيم] وقيمته 16

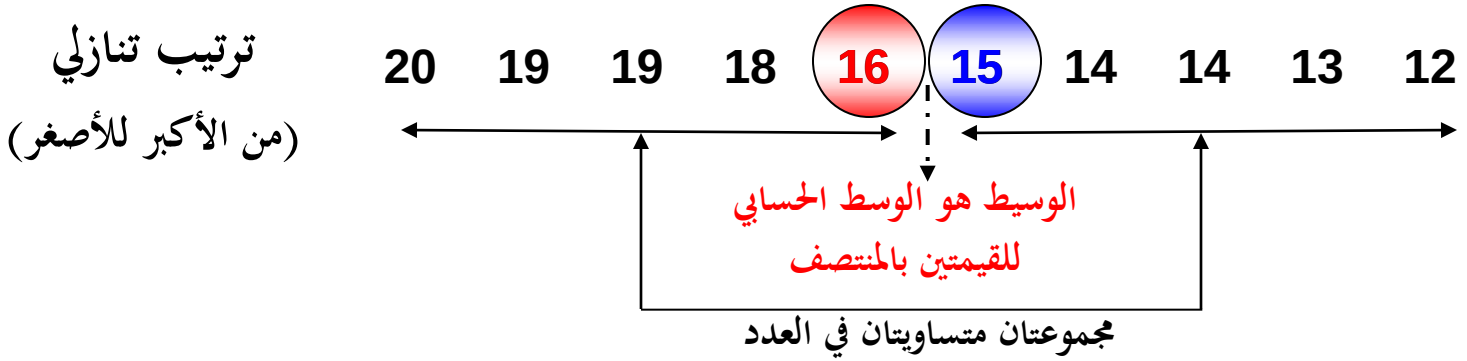
هام جداً

فرق بين رتبة الوسيط وقيمته

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

أما لمجموعة القيم : 12 , 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها ١٠ قيم (أي رقم 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

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

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 :

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

أي القيمتان الثالثة والرابعة ، 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

لاحظ من الأمثلة السابقة 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 آخر : 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 هذا السؤال 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 .

الوسيط لبيانات كمية متصلة

يمكن حساب الوسيط للبيانات الكمية المتصلة من خلال الرسم

وكذلك من خلال المعادلات الاحصائية بسهولة

الوسيط من خلال الرسم البياني يتم كالتالي:

- قيمة المتغير المناظر لنقطة تقاطع المنحنيين: المتجمع الصاعد والمتجمع الهابط للبيانات .

أو

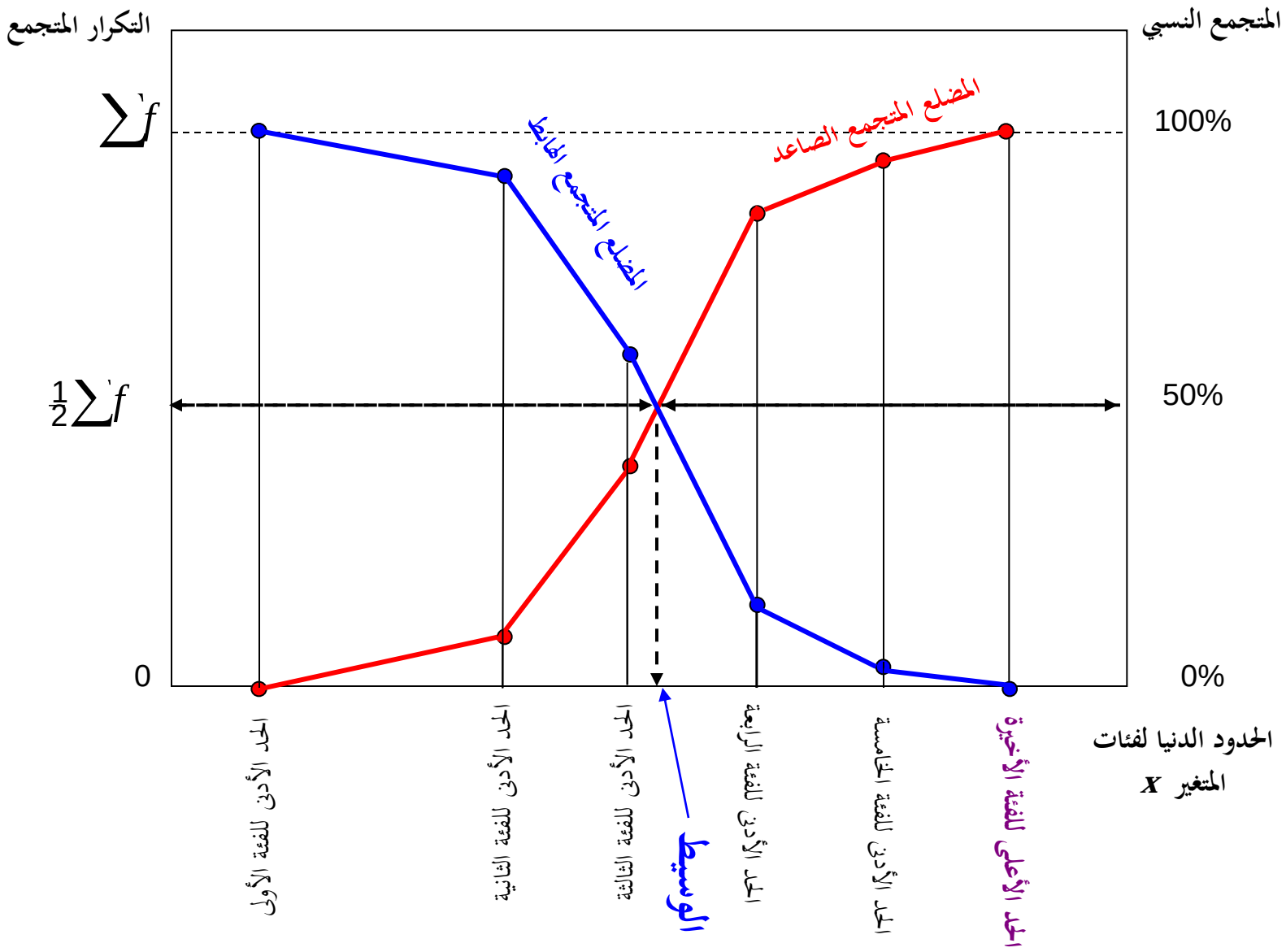
- القيمة التي يناظرها تكرار متجمع = نصف مجموع التكرارات

أو

- القيمة التي يناظرها تكرار نسبي متجمع = 50%

طريقة تحديد الوسيط من :

* من المنحنى المتجمع الصاعد فقط
* من المنحنى المتجمع الهابط فقط .
* من المنحنيين معاً



الوسيط من خلال المعادلات الاحصائية

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

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 .

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

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

الجدول التكراري

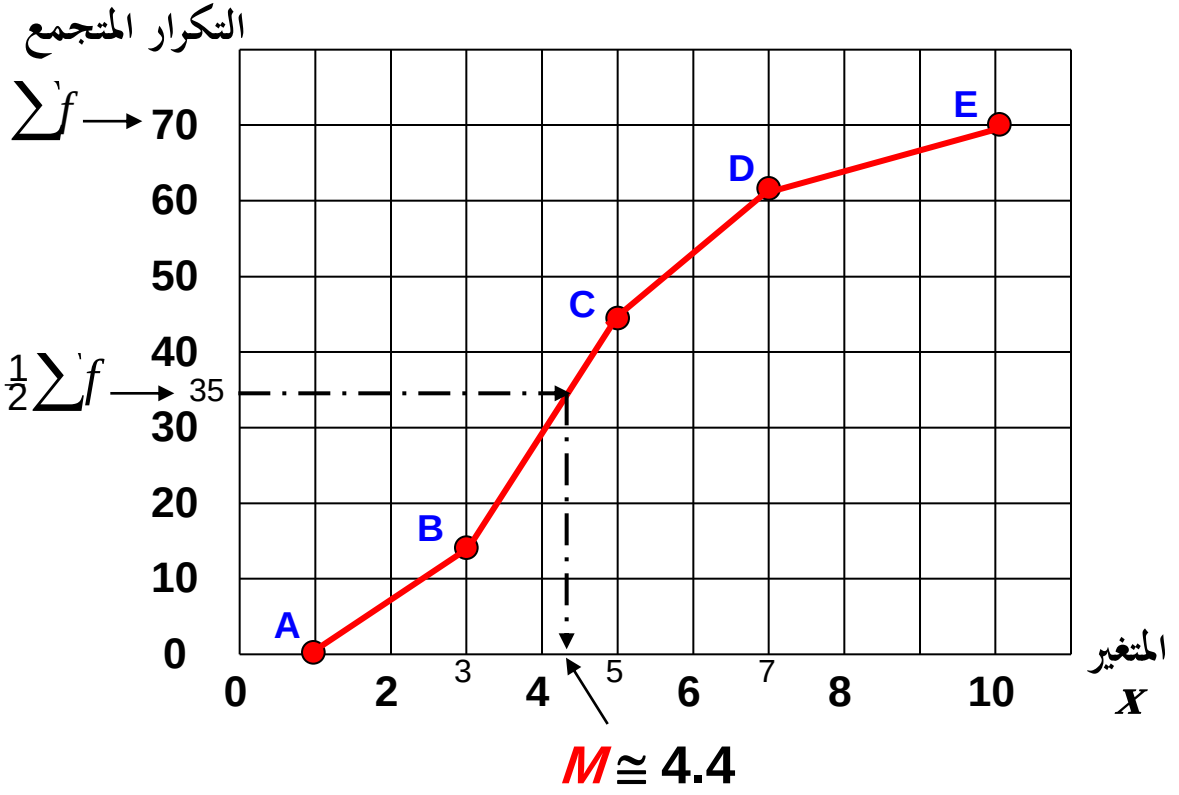
الفئة	المتغير (المساحة) x	التكرار f	المركز x_0	$f x_0$
الأولى	$1 \le x < 3$	14	2	28
الثانية	$3 \le x < 5$	29	4	116
الثالثة	$5 \le x < 7$	18	6	108
الرابعة	$7 \le x < 10$	9	8.5	76.5
		$\sum f = 70$		$\sum f x_0 = 328.5$

$$\therefore \quad x = \frac{\sum f x_0}{\sum f} = \frac{328.5}{70} = 4.692857142 \approx 4.7$$

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
		$\sum f = 70$

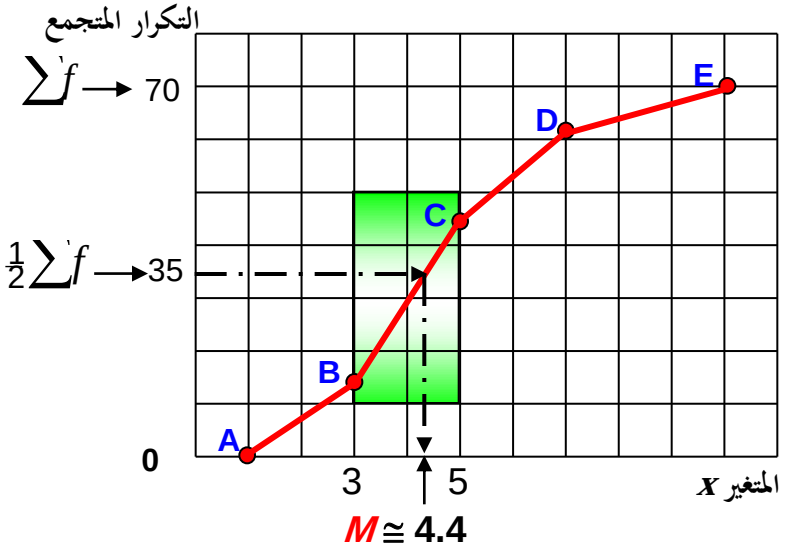
الجدول التكراري المتجمع الصاعد		
المتغير x	التكرار المتجمع	النقطة على المضلع
< 1	0	A (1 , 0)
< 3	14	B (3 , 14)
< 5	43	C (5 , 43)
< 7	61	D (7 , 61)
< 10	$\sum f = 70$	E (10 , 70)

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



هكذا انتهت الإجابة على السؤال، أي أن الوسط الحسابي للمساحة $\cong$ 4.7 والوسيط $\cong$ 4.4

لكن هناك ملاحظة هامة جداً يجي ملاحظتها



الوسيط يقع بين النقطتين C(5 , 43) , B(3 , 14) [أي داخل الفئة $3 \le x < 5$] هذه الفئة تُسمى بـ **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 ، **وذلك كالتالي:**

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 تكون آخر فئة زودنا تكرارها تكون هي الفئة الوسيطية .

ويتم ذلك كالتالي

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 $\frac{1}{2}\sum f$: $\frac{1}{2}\sum f = \frac{70}{2} = 35$ ⟵
- 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 بـ "التكرار المتجمع السابق" = مجموع تكرار الفئات السابقة للفئة الوسيطية

(٣) احسب الوسيط من العلاقة :

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

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 :

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

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

مثال: طُلِب 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	$6 \leq x < 12$	25
الثالثة	$12 \leq x < 15$	35
الرابعة	$15 \leq x < 18$	18
		$\sum f = 98$

المجموعة (٢)

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

المجموعة (٣)

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

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

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

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

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

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

- 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
		$\sum f = 98$

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

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
		$\sum f = 98$

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

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
		$\sum f = 98$

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

المنوال

تعريف المنوال [الشائع] :

يُعرف المنوال لمجموعة من القيم على أنه القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعاً [لذا يُسمى في بعض الأحيان بال "الشائع"] . وأحياناً يُرمز للمنوال بالرمز $\hat{X}$

فمثلاً :

مجموعة القيم 18 16 12 11 10 10 9 9 9 7 5 2 2 لها منوال 9

ومجموعة القيم 16 15 12 10 8 5 3 9 ليس لها منوال [أو عديمة المنوال]

ومجموعة القيم 9 7 7 7 5 5 4 4 4 3 2 لها منوالان 7 , 4

أي أن مجموعة القيم قد تكون وحيدة المنوال [لها منوال واحد] ، وقد تكون عديدة المنوال [منوالان أو أكثر] وقد تكون عديمة المنوال [لا يوجد لها منوال]

أما مجموعة القيم 7 7 6 6 5 5 4 4 فقد تتسرع وتقول أنها رباعية المنوال ومناولها : 4 , 5 , 6 , 7

لكن [حيث أن جميع القيم لها نفس التكرار] هذه المجموعة الأخيرة عديمة المنوال

والمنوال [مقارنةً بالوسط الحسابي والوسيط] به العديد من العيوب منها :

- أنه لا يأخذ في الاعتبار جميع البيانات ولكنه يهتم فقط بالقيم الأكثر تكراراً .
- أنه قد لا يتواجد أو قد يكون هناك أكثر من منوال للبيانات .

إلا أنه أيضاً يتميز ببعض المزايا منها :

- أنه أسرع في تحديده من الوسط والوسيط
- من الممكن تحديده للتوزيعات التكرارية للبيانات المنفصلة سواء كانت تلك البيانات كمية متقطعة أو نوعية [والبيانات الأخيرة (النوعية) لا يمكن حساب الوسط الحسابي لها أو الوسيط]

بيانات منفصلة

درجات طلاب في مقرر الإحصاء	
عدد الطلاب	درجة الطالب
28	12
24	14
39	16
9	18

بيانات كمية متقطعة

لها منوال وحيد وهو "الدرجة 16"

درجات طلاب في مقرر الإنجليزي	
عدد الطلاب	درجة الطالب
23	12
30	14
30	16
17	18

بيانات كمية متقطعة

لها منوالان وهما "14 , 16"

درجات طلاب في مقرر الفقه	
عدد الطلاب	درجة الطالب
25	12
25	14
25	16
25	18

بيانات كمية متقطعة

ليس لها منوال

سيارات في أحد المواقف	
عدد السيارات	لون السيارة
10	أحمر R
23	أزرق B
12	أبيض W
5	أصفر Y

بيانات نوعية

لها منوال وهو "اللون الأزرق"

وماذا عن التوزيعات التكرارية للبيانات الكمية المتصلة

الموضوع في غاية البساطة :

- حدد **الفئة المنوالية** [وهي الفئة التي يناظرها أكبر كثافة تكرار]
- حدد **المنوال** [وهو (تقريباً) مركز الفئة المنوالية]

فمثلاً في التوزيع التكراري التالي:

	الجدول التكراري				
	المتغير x	التكرار f	طول الفئة c	مركز الفئة x_0	كثافة التكرار
الفئة لأولى	$0 \le x < 20$	4	20	10	0.2
الفئة الثانية	$20 \le x < 30$	16	10	25	1.6
الفئة الثالثة	$30 \le x < 35$	12	5	32.5	2.4
الفئة الرابعة	$35 \le x < 40$	10	5	37.5	2
الفئة الخامسة	$40 \le x < 50$	6	10	45	0.6
الفئة السادسة	$50 \le x < 60$	2	10	55	0.2
		$\sum f = 50$			

أكبر كثافة تكرار

إذن الفئة المنوالية هي **الفئة الثالثة**

والمنوال = **32.5**

وبالرغم من هذه البساطة إلا أن هناك **تنبيه** و **تحذير**

بالنسبة للتنبيه : فالطريقة السابقة [اعتبار أن مركز الفئة المنوالية هو المنوال] هي طريقة تقريبية ، لكن هناك طرق أخرى حسابية وبيانية تعطي قيمة المنوال أكثر دقة ، لكننا لن نتطرق لهذه الطرق في هذا المكان وذلك للتبسيط

وبالنسبة للتحذير : فالفئة المنوالية هي الفئة التي يناظرها أكبر كثافة تكرار وليس أكبر تكرار ، ويتفق اللفظان "أكبر كثافة تكرار" و "أكبر تكرار" فقط إذا كانت أطوال الفئات واحدة .

فمثلاً في المثال السابق ، إذا قلنا أن الفئة المنوالية هي الفئة المناظرة لأكبر تكرار فستكون تلك الفئة هي الفئة الثانية [وهذا خطأ جسيم]

في حين أن الفئة المنوالية هي الفئة الثالثة [كما سبق وبينا] .

الجدول التكراري

	المتغير x	التكرار f	طول الفئة c	مركز الفئة x_0	كثافة التكرار
الفئة لأولى	$0 \le x < 20$	4	20	10	0.2
الفئة الثانية ← الفئة المنوالية [خطأ جسيم]	$20 \le x < 30$	16	10	25	1.6
الفئة الثالثة ← الفئة المنوالية [صح]	$30 \le x < 35$	12	5	32.5	2.4
الفئة الرابعة	$35 \le x < 40$	10	5	37.5	2
الفئة الخامسة	$40 \le x < 50$	6	10	45	0.6
الفئة السادسة	$50 \le x < 60$	2	10	55	0.2
		$\sum f = 50$			

مقارنة بين المتوسطات الثلاثة : الوسط ، الوسيط ، المنوال

الوسط الحسابي

مزاياه :

- سهولة حسابه
- يأخذ في الاعتبار جميع البيانات
- لا يحتاج إلى ترتيب معين للبيانات

عيوبه :

- يتأثر بشدة بالقيم المتطرفة
- لا يمكن إيجاده بالرسم [بيانياً]
- لا يمكن حسابه في حالات التوزيعات التكرارية المفتوحة

الأكثر استخداماً

الوسيط

مزاياه :

- سهولة حسابه حسابياً أو بيانياً
- لا يتأثر بالقيم المتطرفة
- يمكن حسابه في حالة التوزيعات التكرارية المفتوحة

عيوبه :

- يحتاج إلى ترتيب للبيانات أولاً
- لا يأخذ في الاعتبار جميع البيانات

يفضل استخدامه في الحالات التي لا نستطيع فيها حساب الوسط الحسابي

المنوال

مزاياه :

- سهولة حسابه
- لا يتأثر كثيراً بالقيم المتطرفة
- لا يحتاج لترتيب البيانات

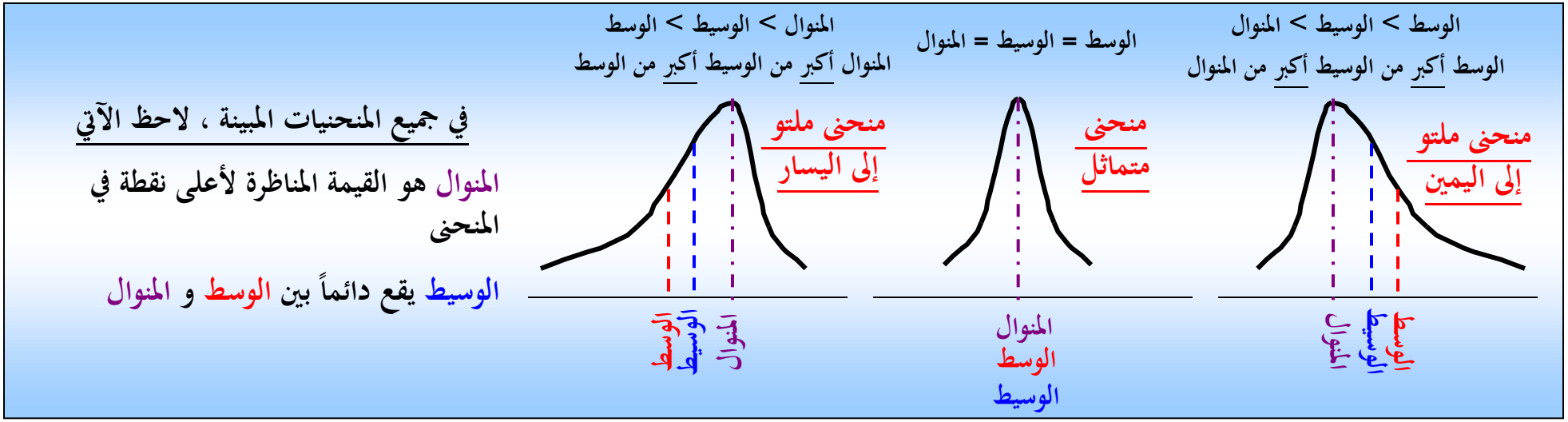
عيوبه :

- قد لا يتواجد وقد يكون له أكثر من قيمة

أقل مقاييس النزعة المركزية استخداماً

يمكن حسابها للبيانات الكمية

في بعض الحالات يمكن تحديده للبيانات النوعية



والمنحنيات التكرارية **وحيدة المنوال والبسيطة الالتواء** تحقق العلاقة الاعتبارية التالية :

$$\text{الوسط} - \text{المنوال} = ٣ \times (\text{الوسط} - \text{الوسيط})$$

وهذه العلاقة يمكن وضعها على أي صورة من الصور التالية

$$\text{الوسط} = \frac{(٣ \times \text{الوسيط}) - \text{المنوال}}{٢}$$

وهذه الصورة مفيدة عندما يكون الوسيط و المنوال معلومان ونريد معرفة الوسط الحسابي

$$\text{المنوال} = (٣ \times \text{الوسيط}) - (٢ \times \text{الوسط})$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و الوسيط معلومان ونريد معرفة المنوال

$$\text{الوسيط} = \frac{(٢ \times \text{الوسط}) + \text{المنوال}}{٣}$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و المنوال معلومان ونريد معرفة الوسيط

- فمثلاً إذا كان المنوال لمجموعة من القيم = 95 ، والوسيط لها = 85 ، فإن :

$$\text{الوسط} = \frac{(٣ \times \text{الوسيط}) - \text{المنوال}}{٢} = \frac{(٣ \times ٨٥) - ٩٥}{٢} = \frac{٢٢٥ - ٩٥}{٢} = \frac{١٦٠}{٢} = ٨٠$$

- وإذا كان الوسط الحسابي لمجموعة من القيم = 80 ، والوسيط لها = 85 ، فإن :

$$\text{المنوال} = (٣ \times \text{الوسيط}) - (٢ \times \text{الوسط}) = (٣ \times ٨٥) - (٢ \times ٨٠) = ٢٥٥ - ١٦٠ = ٩٥$$

- وإذا كان الوسط الحسابي لمجموعة من القيم = 80 ، والمنوال لها = 95 ، فإن :

$$\text{الوسيط} = \frac{(٢ \times \text{الوسط}) + \text{المنوال}}{٣} = \frac{(٢ \times ٨٠) + ٩٥}{٣} = \frac{٩٥ + ١٦٠}{٣} = \frac{٢٥٥}{٣} = ٨٥$$

سؤال: المنحنى التكراري للبيانات المذكورة في أيٍ من الأمثلة السابقة :

◯ ملتو لليمين ◯ ملتو لليسار ◯ متماثل