

**منهاج مقرر**

**إحصاء الأعمال والاقتصاد**

**إعداد الدكتور**

**عدنان عمورة**

**2019-2018**

## الفصل الأول

### وسائل تلخيص المعطيات و عرضها

## Methods of Summarizing and Presenting Data

### مقدمة:

بعد تحديد مجتمع الدراسة و العينة وقياس قيم المتغير على أفراد العينة نحصل على البيانات (المعطيات) بصورتها الأولية و تسمى البيانات الخام. هذه البيانات تحتاج إلى تنظيم و تلخيص بطرق مناسبة ليتمكن الباحث من الاستفادة منها واستخلاص النتائج المفيدة المتعلقة بالمجتمع.

سندرس أهم الطرق و الوسائل لتنظيم و عرض البيانات وهي:

- (1) الجداول التكرارية.
- (2) المدرجات و المضلعات و المنحنيات التكرارية .
- (3) مخطط الساق و الورقة.
- (4) المخطط الصندوقي.
- (5) مخطط الدائرة (الفطيرة).
- (6) أشكال الانتشار و التبعثر.

عند تحديد الطرق المناسبة لعرض البيانات لا بد من النظر للبيانات على أنها من نوعين بيانات كمية و بيانات اسمية (وصفية) ،وكلا النوعين يمكن تنظيمها في جداول تكرارية (جداول التوزيع التكرارية).

## 1-1 الجداول التكرارية (Frequency Tables) :

### 1-1-1 الجداول التكرارية للبيانات الوصفية (الاسمية):

يحتوي الجدول التكراري للبيانات الوصفية عمودين، يتضمن الأول قائمة الفئات، وهي مجموعة كل الحالات (الصفات) التي تكون البيانات، ويتضمن العمود الثاني (عمود التكرارات) عدد عناصر العينة لكل حالة (تكرارات تلك الحالة في العينة).

الفئات (أسباب الوفاة)	Frequency	Percent
قلبية	18	0.18
أمراض الدم	5	0.05
أمراض عصبية	4	0.04
التسمم	2	0.02
الجهاز التنفسي	5	0.05
حوادث سير	15	0.15
سرطان	17	0.17
مضاعفات الحمل	8	0.08
هضمية	6	0.06
أسباب أخرى متنوعة	20	0.2
Total	100	100.0

الجدول (1-1) التوزيع التكراري  
لمئة متوفى تم  
تصنيفهم حسب سبب الوفاة

### مثال (1-1):

يبين الجدول (1-1) توزيع أسباب الوفاة لعينة من 100 متوفى تم تسجيلهم خلال فترة زمنية محددة في مشفى المجتهد، إذ تتضمن العمود الثاني (عمود التكرارات) عدد الأشخاص من العينة لكل سبب من الأسباب الواردة في العمود

الأول (عمود الفئات)، أما العمود الثالث عمود التكرارات النسبية فيحوي التكرار النسبي لكل فئة أي نسبة تكرارات تلك الفئة إلى حجم العينة .

### 2-1-1 الجداول التكرارية للبيانات الكمية:

تختلف طريقة إنشاء الجداول التكرارية للبيانات الكمية عنها للبيانات الوصفية؛ لأننا لو اعتبرنا كل قيمة من قيم المتغير الكمي فئة لحصلنا على جدول يحوي عدداً كبيراً من الفئات، الشيء الذي يفقد عملنا هدفه، وهو تبسيط وتلخيص المعطيات، لذلك نقسم بياناتنا إلى فئات كما يأتي:

- نجزئ مجال الأعداد الذي يحوي بيانات المتغير إلى فترات منفصلة ومتلاصقة، غالباً تكون متساوية الطول بحيث تغطي جميع قيم المتغير. نعتبر الفئة الأولى هي ذلك الجزء من العينة الذي تقع بياناته في الفترة الأولى، أما الفئة الثانية فهي جزء العينة الذي تقع بياناته في الفترة الثانية وهكذا.. إلى آخر فئة.
- أما عدد الفئات فيتراوح غالباً بين 5 و 15 فئة، وذلك حسب الهدف من البحث وحسب طبيعة البيانات وحجمها، ويمكن استخدام قاعدة ستارج ( Sturges Rule) لتحديد عدد الفئات  $k$  لبيانات عددها  $n$  .

$$k = 1 + 3.322 \ln(n)$$

مع تقريب الناتج لأقرب عدد صحيح، و الجدول الآتي يعطي عداد الفترات  $k$  لأحجام عينات مختلفة  $n$  :

n	25	45	100	200	350	700
k	6	7	8	9	10	11

• تحديد الفئات: نعرف مدى البيانات بأنه الفرق الموجب بين أكبر و أصغر قيمة

$$R = X_{max} - X_{min}$$

ثم نحدد طول كل فترة بأنه نسبة المدى إلى عدد

الفترات، و نرمز له بـ  $w = \frac{R}{k}$  ، ونقرب الناتج إلى أقرب وحدة دقة و هي

الخانة العشرية المستخدمة في تقريب البيانات فتكون الفترة الأولى

$$[L_1 = X_{min}, U_1 = L_1 + w [$$

وهي مجموعة القيم التي لا تقل عن

$$L_1 = X_{min} \text{ ، وتنقص عن } U_1 = L_1 + w .$$

أما الفترة الثانية فهي مجموعة القيم التي لا تقل عن  $L_2 = U_1$  وتنقص عن

$$U_2 = L_2 + w : [L_2 = U_1, U_2 = L_2 + w [$$

وهكذا فإن الفترة الأخيرة هي مجموعة القيم التي لا تقل عن  $L_k = U_{k-1}$

$$\text{حتى أكبر قيمة، وهي } X_{max} [L_k = U_{k-1}, U_k = X_{max} [$$

مثلاً لو كانت البيانات مؤلفة من 45 قيمة أصغرها  $X_{min} = 2.31$  و أكبرها

$$X_{max} = 3.35 \text{ حسب قاعدة ستارج نعتبر } k=7 \text{ ونحسب:}$$

$$\frac{R}{7} = \frac{X_{max} - X_{min}}{7} = \frac{1.04}{7} \cong 0.1458$$

نقرب الناتج إلى أقرب وحدة دقة

(0.01) إلى الأعلى لنجد 0.15 وتكون الفترات هي

الحد الأدنى	الحد الأعلى
2.31	2.46
2.46	2.61
2.61	2.76
2.76	2.91
2.91	3.06
3.06	3.21
3.21	3.36

• نعرف الآن الجدول التكراري للبيانات بأنه جدول من عمودين يتضمن العمود الأول (عمود الفترات) وهي التي تصنف وفقها العينة إلى فئات ويتضمن العمود الثاني (عمود التكرارات). تكرار كل فئة ونرمز بـ  $f_i$  هو عدد جزء العينة الذي قيمه تنتمي للفترة المقابلة ، ويكون مجموع التكرارات مساوياً لحجم العينة  $n = \sum_{i=1}^n f_i$ .

• يمكن أن يضاف إلى جدول التكرارات عمود ثالث ندعوه عمود التكرارات النسبية، والتكرار النسبي للفئة  $i$  ليس إلا نسبة تكراراتها إلى  $n$  أي  $\frac{f_i}{n}$  ، ونلاحظ أن مجموع التكرارات النسبية يساوي الواحد  $\sum_{i=1}^n \frac{f_i}{n} = 1$ .

• وقد نضيف إلى عمود التكرارات النسبية أو بدلاً عنه عمود تكرارات النسب المئوية، و التكرار المئوي للفترة  $i$  هو ناتج ضرب تكرارها النسبي بـ 100 أي يساوي  $100 \times \frac{f_i}{n}$  ، وهو ليس إلا النسبة المئوية لتكرار الفئة  $i$ ، فإذا كانت  $n=150$  و  $f_i = 30$  فيكون التكرار النسبي 0.2 والتكرار المئوي 20% ونلاحظ أن مجموع التكرارات المئوية يساوي 100.

### مثال (1-2):

البيانات التالية تمثل مستوى الهيموغلوبين لعينة مكونة من خمسين شخصاً

17	-	17.7	-	15.9	-	16.2	-	16.2	-	17.1	-	15.7	-	17.3	-	14.6	-	15.8
16.4	-	13.7	-	16.2	-	16.4	-	16.1	-	14	-	16.2	-	16.4	-	14.9	-	17.8
18.3	-	15.9	-	15.3	-	13.9	-	16.8	-	15.9	-	16.3	-	17.4	-	15	-	15.3
15.1	-	17.4	-	16.5	-	14.4	-	16.3	-	16.3	-	15.9	-	16.7	-	16.1	-	15.5
15.1	-	15.8	-	13.5	-	17	-	15.8	-	17.5	-	17.3	-	14.2	-	16.1	-	15.7

أوجد الجدول التكراري لتلك البيانات المكون من 6 فئات و يحوي عمودي التكرارات النسبية والمئوية.

**الحل:**

نوع البيانات كمية حجمها  $n=50$  نعين طول الفئة

$$\frac{X_{\max}-X_{\min}}{7} = \frac{18.3-13.5}{6} = 0.8$$

طول كل فترة يساوي 0.8 نجد حدود الفترات

13.5 - 14.3 - 15.1 - 15.9 - 16.7 - 17.5 - 18.3

نرتب الفئات في العمود الأول ، ونسجل في العمود الثاني تكرارات كل فئة (عدد قيم العينة التي تنتمي لكل فترة)

مستوى الهيموغلوبين الفترات	التكرارات $f_i$	التكرارات النسبية $\frac{f_i}{50}$	التكرارات المئوية $(\frac{f_i}{50}) \times 100\%$
13.5-14.3	5	0.1	10%
14.3-15.1	4	0.08	8%
15.1-15.9	10	0.2	20%
15.9-16.7	18	0.36	36%
16.7-17.5	9	0.18	18%
17.5-18.3	4	0.08	8%
	50	1	100

الجدول (2-1) التكرارات مع التكرارات النسبية والمئوية لمستوى الهيموغلوبين

### 3-1-1 الجداول التكرارية التراكمية (Cumulative frequency tables):

نعرف التكرار المتجمع الصاعد لفئة بأنه عدد البيانات المتضمنة في تلك الفئة وفي الفئات التي تسبقها. أما التكرار المتجمع الهابط لفئة فيساوي عدد البيانات

الواقعة في هذه الفئة وفي الفئات التي تليها. وجدول التكرارات المتجمعة جدول، من ثلاثة أعمدة، عموده الأول يحوي الفترات (الفئات) ، أما الثاني والثالث فيحوي التكرارات المتجمعة الصاعدة والتكرارات المتجمعة الهابطة على الترتيب للفئات المقابلة.

ولا بد هنا من الإشارة لأهمية جدول التكرارات الصاعدة في تقدير عدد البيانات التي تقل عن قيمة معينة وعدد البيانات الواقعة بين قيمتين مفروضتين ، وكذلك جدول التكرارات المتجمعة الهابطة فيساعد في تقدير عدد البيانات التي تزيد على أو تساوي قيمة معينة.

### مثال (3-1):

نبين في الجدول (3-1) التكرارات المتجمعة الصاعدة والهابطة للبيانات التي تمثل مستوى الهيموغلوبين.

مستوى الهيموغلوبين	التكرارات	التكرارات الصاعدة	التفسير	التكرارات الهابطة	التفسير
13.5-14.3	5	5		50	
14.3-15.1	4	9	5 قيم أقل من 14.3	45	45 قيمة لا تقل عن 14.3
15.1-15.9	10	19	9 بيانات أقل من 15.1	41	41 قيمة لا تقل عن 15.1
15.9-16.7	18	37	19 قيمة أقل 15.9	31	31 قيمة لا تقل عن 15.9
16.7-17.5	9	46	27 قيمة أقل من 16.7	13	13 قيمة لا تقل عن 16.7
17.5-18.3	4	50	46 قيمة أقل من 17.5	4	أربع قيم لا تقل عن 17.5

الجدول (3-1) تفسير التكرارات الصاعدة والهابطة.

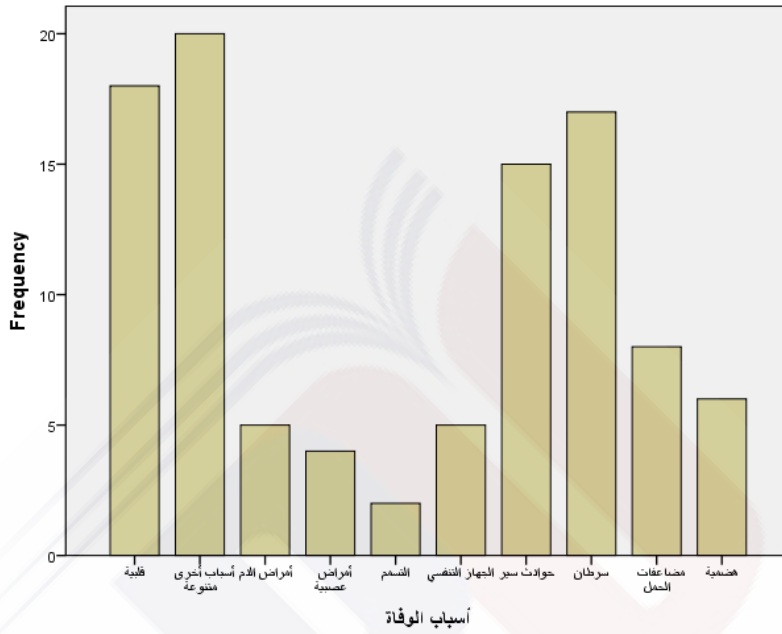


## 1-2 العرض البياني للجداول التكرارية:

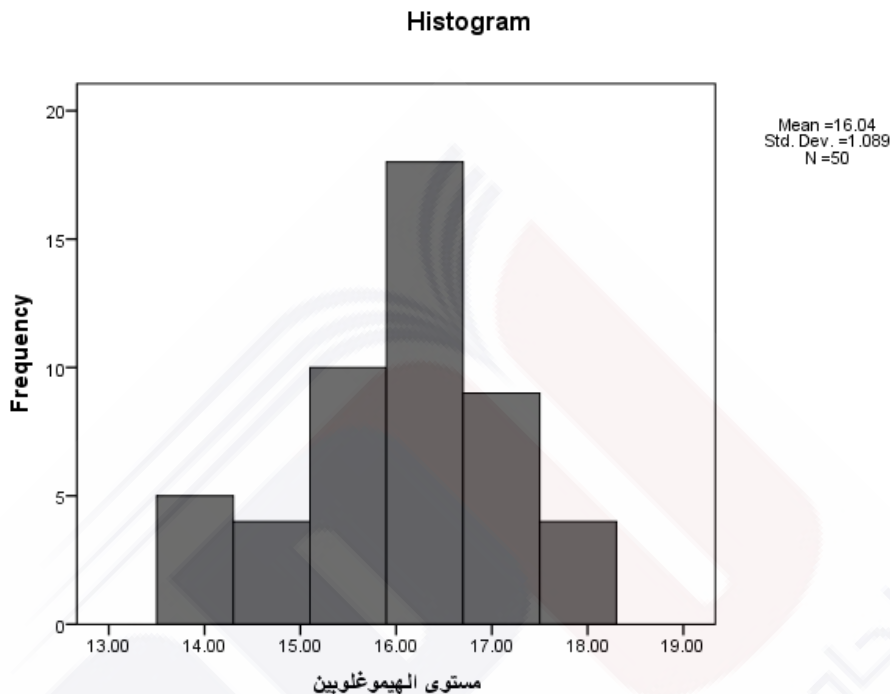
إن تلخيص وتنظيم البيانات خطوة أساسية لا بد منها ، لكنها غير كافية لعرض البيانات لذلك سنتعرف على أسلوب آخر لتقديم البيانات ، وتمثيلها بيانياً ليتاح للمهتم التعرف عليها دون عناء وتعب.

### 1-2-1 المدرج التكراري (Frequency Histogram) :

المدرج التكراري للبيانات يتكون من مجموعة من الأعمدة المتجاورة المقامة على المحور الأفقي Ox (الذي يمثل عليه حالات الصفة المدروسة عندما تكون البيانات وصفية وفترات الفئات عندما تكون البيانات كمية) يساوي ارتفاع كل عمود تكرارات الفئة المقام عليها. انظر الشكل (1-1) الذي يمثل المدرج التكراري للأسباب المختلفة للوفاة لعينة من المرضى المتوفين خلال فترة محددة في مشفى المواساة والشكل (2-1) يمثل تكرارات المستويات المختلفة للهيموغلوبين لعينة من خمسين شخصاً ، فيمكن بنظرة سريعة لكل من المدرجين إدراك كيفية توزع العينة. أي الحالات لها أعلى تكرار ، وأيها لها أدنى تكرار. ويستفاد من أسلوب العرض هذا في مقارنة عينتين أو أكثر . كما سنرى في المثال الآتي:



الشكل (1-1) المدرج التكراري لأسباب الوفاة



الشكل ( 2-1 ) المدرج التكراري لمستوى الهيموغلوبين

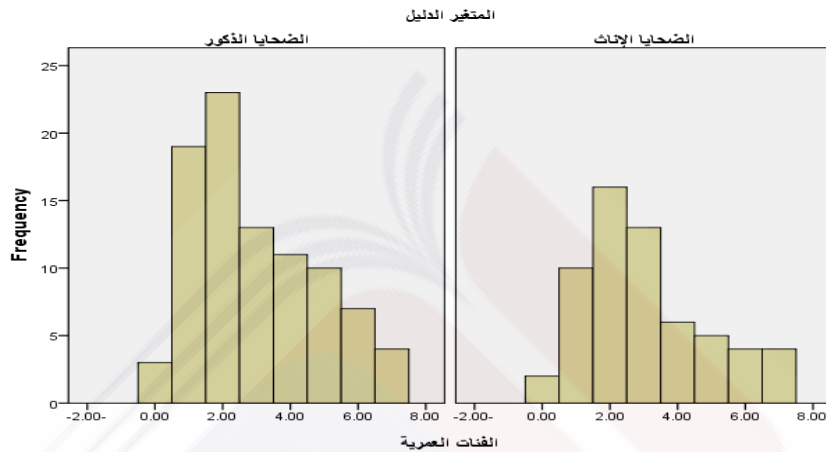
**مثال (1-4):**

الجدول (3-1) يلخص التوزيع و التوزيع النسبي المئوي، لأعمار عينة من الأشخاص الذين توفوا في حوادث مرورية تتكون من 150 شخصاً ( 90 من الذكور و 60 من الإناث). إذ جُرِّت العينة إلى فئات عمرية الأولى قبل سن العاشرة والثانية مجموعة الأعمار ما بين 10 و 20 سنة والثالثة فئة العشرينات وهكذا.... الفئة الثامنة من كانت أعمارهم في السبعينات أو أكثر.

يظهر الشكل (3-1) التوزيع التكراري لأعمار الضحايا الذكور - أ - والضحايا الإناث - ب - . أما الشكل (4-1) فيظهر مقارنة بين أعداد الضحايا الذكور والإناث في كل فئة عمرية باستخدام مخطط الأعمدة .

Age	Male		Female	
	Frequency	Percent	Frequency	Percent
أقل من عشر سنوات	3	3.3	2	3.3
بين العشر والعشرين سنة	19	21.1	10	16.7
فئة العشرينات	23	25.6	16	26.7
فئة الثلاثينات	13	14.5	13	21.7
فئة الأربعينات	11	12.2	6	10
فئة الخمسينات	10	11.1	5	8.3
فئة الستينات	7	7.8	4	6.65
السبعينات وما يزيد	4	4.4	4	6.65
Total	90	100	60	100

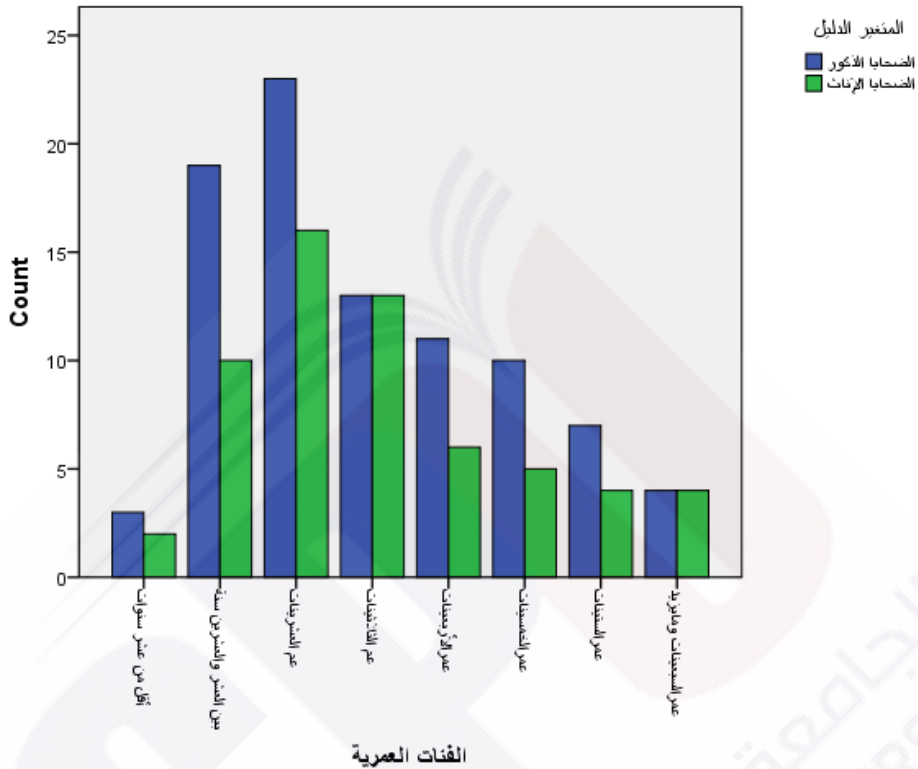
الجدول (4-1) الجدول التكراري والتكراري النسبي المئوي لعينة ضحايا الحوادث المرورية.



أعمار الذكور - ب -

أعمار الإناث - أ -

الشكل (3-1) التوزيع التكراري لأعمار ضحايا حوادث السير للذكور والإناث.



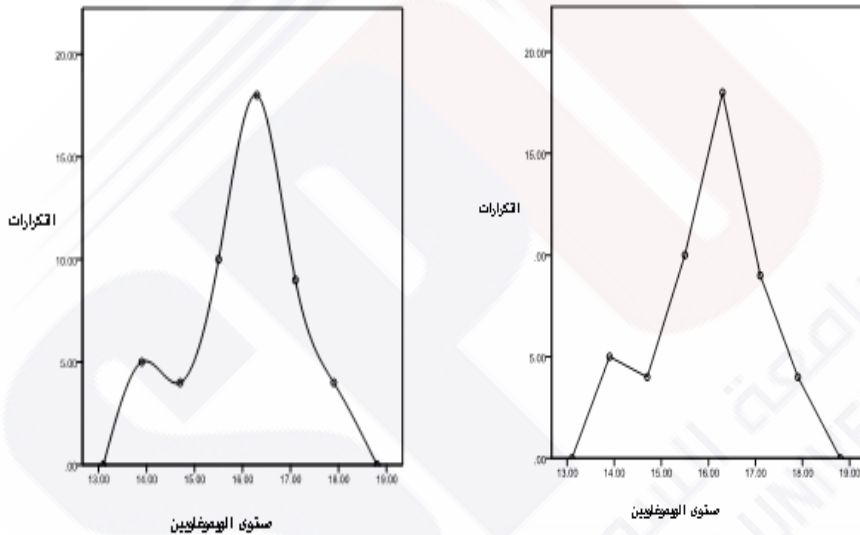
الشكل (4-1) مخطط الأعمدة المركب لمقارنة أعداد العينتين في كل فئة عمرية

## 2-2-1 المضلع التكراري والمضلع التكراري النسبي ( Frequency Polygon):

هناك تمثيل بياني هندسي آخر بديل عن المدرج التكراري ندعوه المضلع التكراري، وهو عبارة عن خط منكسر يصل بقطع مستقيمة النقاط التي مساقطها مراكز الفئات وترتيبها تكرارات الفئات المحددة فوقها النقطة وبدايته على المحور الأفقي فوق فترة وهمية تكرارها صفر ونهايته كذلك على فترة وهمية تلي جميع الفترات تكرارها صفر، أما المضلع التكراري النسبي فهو خط منكسر يصل النقاط التي لها

نفس المساقط المذكورة ، لكن تراتيبها هي التكرارات النسبية للفئات فهو لا يختلف من حيث الشكل عن المضلع التكراري العادي.

يمثل الشكل (1-5) المضلع التكراري والمنحني التكراري لمستوى الهيموغلوبين عند عينة من الأشخاص .انظر المثال (1-2) .ونحصل على المنحني التكراري إذا وصلنا النقاط السابقة بخط منحنٍ .



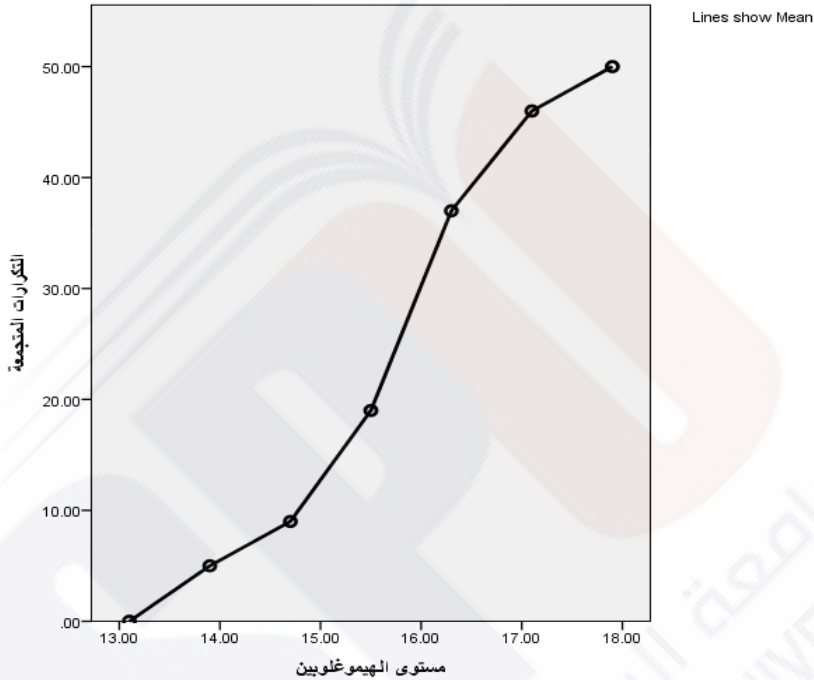
الشكل ( 1-5 ) المضلع التكراري والمنحني التكراري لمستوى الهيموغلوبين

### 3-2-1 المضلع التكراري الصاعد و المنحني التكراري الصاعد

: (Cumulative frequency polygon curve)

ندعو الخط المنكسر الواصل بين مجموعة النقاط التي مساقطها على المحور الأفقي مراكز الفئات وعلى المحور الشاقولي التكرارات المتجمعة الصاعدة لتلك

الفئات بالمضلع التكراري المتجمع الصاعد والمنحني الذي يصل بين تلك النقاط بالمنحني التكراري المتجمع الصاعد انظر الشكل (1-6).



الشكل ( 1-6 ) المضلع التكراري المتجمع ( التراكمي ) الصاعد لمستوى الهيموغلوبين

### 3-1 مخطط الساق والورقة:

هناك أسلوب آخر لتنظيم البيانات أعده (Tukey) يشبه أسلوب الجداول التكرارية والأعمدة، وهو مخطط الساق والورقة، فقد استبدل الأعمدة بالأعداد نفسها، فالساق هو القسم الصحيح من العدد والورقة القسم العشري.

**مثال (1-5):** يبين الجدول التالي قيم المتغير الكمي الدال على حجم الزفير القسري بالثانية لخمسين طالباً من طلاب كلية الطب:



2.85	2.98	3.04	3.10	3.10	3.19	3.30	3.39	3.42	3.48
3.50	3.54	3.52	3.54	3.57	3.60	3.69	3.75	3.78	3.83
3.90	3.96	4.05	4.08	4.10	4.14	4.14	4.16	4.20	4.20
4.30	4.30	4.32	4.44	4.47	4.47	4.50	4.56	4.56	4.68
4.70	4.78	4.80	4.80	4.90	5	5.1	5.1	5.2	5.3

والترتيب الآتي هو مخطط الساق و الورقة لتلك البيانات

2	8	9
3	0	1 1 1 3 3 4 4 5 5 5 5 5 6 6 7 7 8 9 9
4	0	0 1 1 1 1 2 2 3 3 3 4 4 4 5 5 5 6 7 7 8 8 9
5	0	1 1 2 3

ففي الساق الأولى 2 مثلاً ورقتان 8 و9 تمثلان العددين 2.8 و 2.9 للدلالة على القياسين 2.85 و 2.98 من العينة و مخطط الساق و الورقة شكل بسيط يوضح لنا كيفية توزع القياسات.

#### 1-4 المخطط الصندوقي:

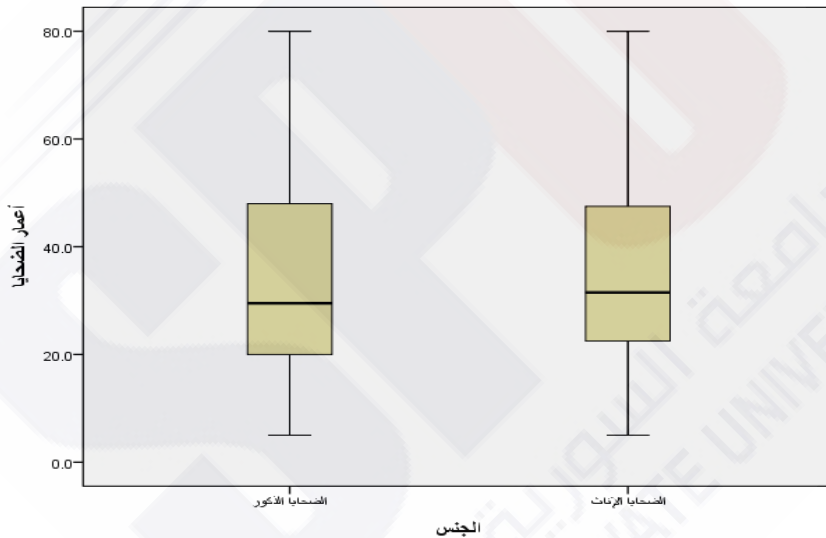
نرغب في كثير من التطبيقات بعرض توزيع البيانات بشكل مبسط لمقارنة عينتين أو أكثر، حينئذٍ قد يكون من المفيد استخدام الربيعيات . و كما نعلم لكل مجموعة من القياسات الترتيبية (على الأقل) ثلاثة ربيعيات ، الربيعي الأول و الثاني و الثالث ، ويعرف كل ربيعي بالقياس الذي تسبقه على الترتيب 25% و 50% و 75% من القياسات بعد ترتيبها تصاعدياً إذ تستخدم هذه المقاييس لوصف توزيع وتشتت البيانات .

فالمخطط الصندوقي عبارة عن صندوق ذي قرنين ممتدين بشكل موازٍ للمحور الشاقولي الذي تتوزع عليه القياسات ، و تشير حافتا الصندوق السفلية والعليا للربيعيين الأول و الثالث على الترتيب ، وهذا يعني أن حافتي الصندوق تتضمنان

نصف القياسات متوسطة القيم ، ويفصل بينها خط أفقي يمثل الوسيط (الربيعي الثاني). و تقع ربع القياسات ذات القيم الأصغر تحت الصندوق و ربع القياسات ذات القيم الأعلى فوقه .

### مثال(1-6):

بهدف مقارنة تشتت أعمار ضحايا حوادث الطرق للذكور و الإناث رسمنا المخطط الصندوقي لعينتي الذكور و الإناث انظر الشكل (1-7)



الشكل (1-7) مخطط صندوقي لتوزع وتشتت أعمار ضحايا حوادث المرور

### 5-1 مخطط الفطيرة (الدائرة) (Pie Charts):

تبقى كل الوسائل التي ذكرناها إجراءات أولية لا بد منها لإلقاء نظرة سريعة على البيانات ، لكنها غير كافية ، ولا تغني الدارس عن استخدام تقنيات متطورة لتحليل المعطيات ، فعلى أي حال تدعونا الحاجة أحياناً لإيجاد سبل وأساليب معينة

لعرض البيانات على شكل تتيح للمهتم التعرف وبمنظرة سريعة على الظاهرة المدروسة.

من هذه الأساليب رسم مخطط الفطيرة وتجزئة الدائرة ( 360 درجة ) إلى مجموعة من القطاعات الزاوية تتناسب وقياساتها مع تكرارات الفئات (المجموعات الجزئية التي تكون المعطيات ).

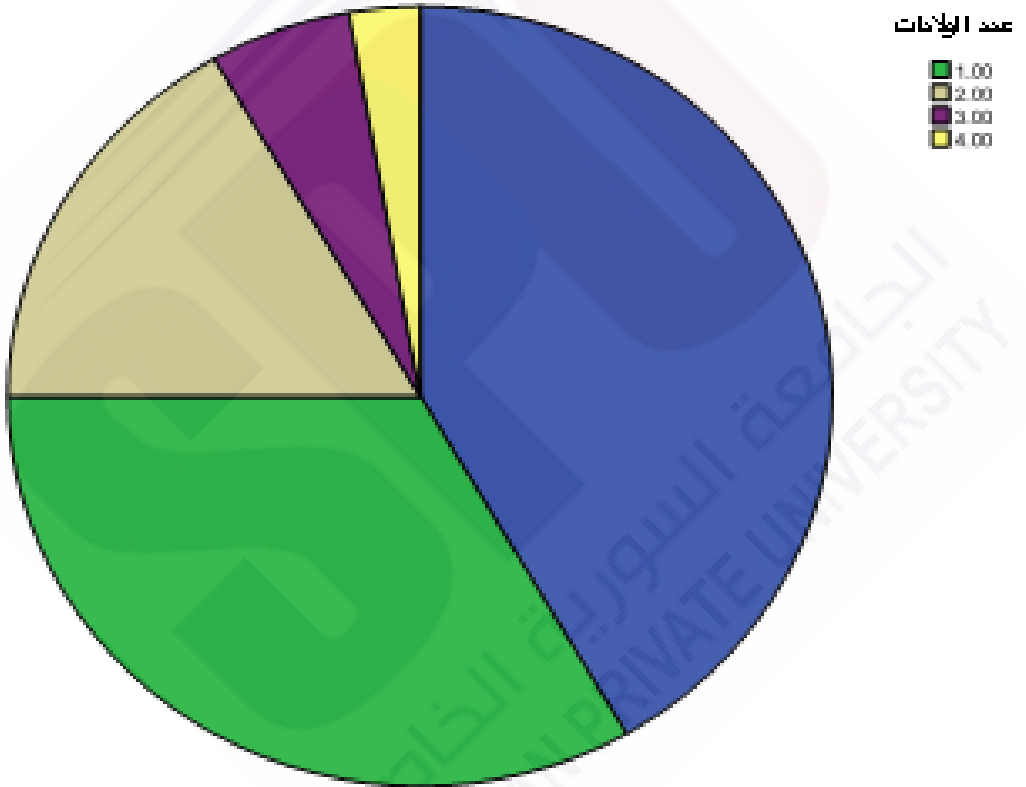
#### مثال (1-7) :

الجدول ( 1-5 ) هو الجدول التكراري لبيانات كمية تدل على عدد الولادات السابقة لعينة من 108 امرأة حامل دخلن مشفى التوليد

قياسات القطاعات الزاوية	التكرارات	عدد الولادات
$\frac{45}{108} \times 360 = 150$	45	0
$\frac{36}{108} \times 360 = 120$	36	1
$\frac{18}{108} \times 360 = 60$	18	2
$\frac{6}{108} \times 360 = 20$	6	3
$\frac{3}{108} \times 360 = 10$	3	4
360	108	

الجدول ( 1-5 ) قياسات القطاعات الزاوية لفئات عدد الولادات

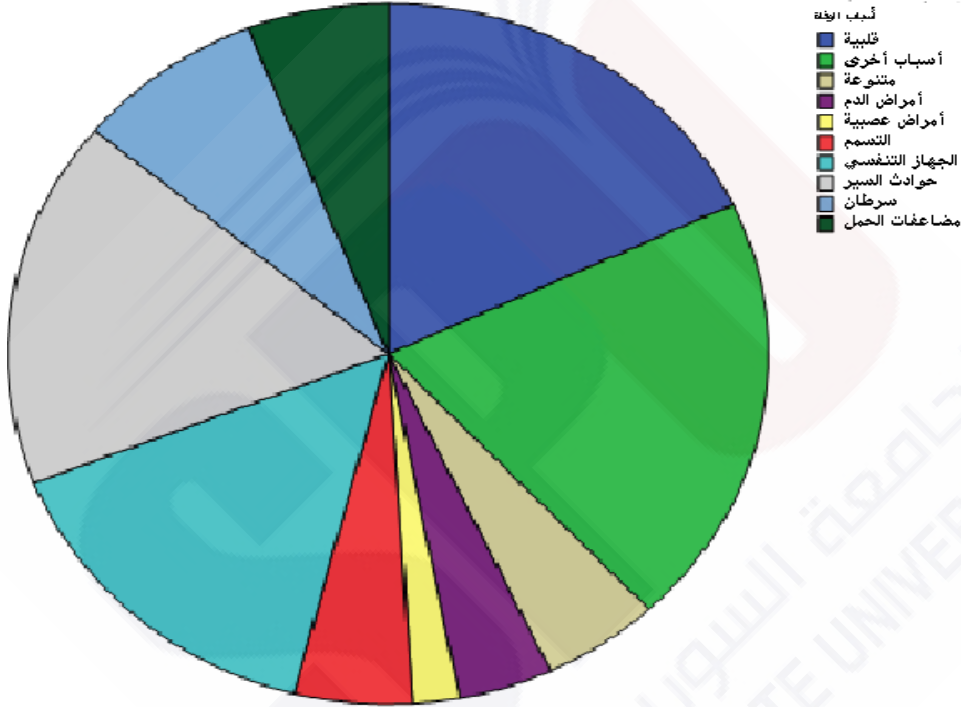
يتضمن العمود الأخير قياسات زوايا الفئات الخمس ، فنجد قياس زاوية الفئة i  $\frac{f_i}{n} \times 360$  درجة. أي قياس زاوية الفئة الثالثة مثلاً يساوي  $60 = \frac{18}{108} \times 360$  .  
والشكل ( 8-1 ) الآتي مخطط الفطيرة لهذه البيانات ، فتدل المساحة الأكبر على نسبة النساء اللاتي ينتظرن المولود الأول والمساحة الأصغر لنسبة النساء اللاتي لهن أربعة أولاد ...



الشكل ( 8-1 ) مخطط الدائرة لعدد الولادات

### مثال (1-8):

تفيد هذه الطريقة في عرض البيانات جميع أنواع البيانات ولاسيما الاسمية الشكل (1-9) يظهر توزيع أسباب الوفاة لعينة من المتوفين الواردة في المثال (1-1).



الشكل (1-9) مخطط الدائرة لتوزيع بيانات أسباب الوفاة الاسمية

### 6-1 مخطط الانتشار (Scatter diagram):

لكل نوع من البيانات الأسلوب الأفضل الذي يناسبه ، فعند دراسة العلاقة بين متغيرين نرى من الأجدى أن نجعل المستوي ساحة لتمثيل العينة حيث تمثل قياسات المتغير الأول على المحور الأفقي والمتغير الثاني على المحور الشاقولي،

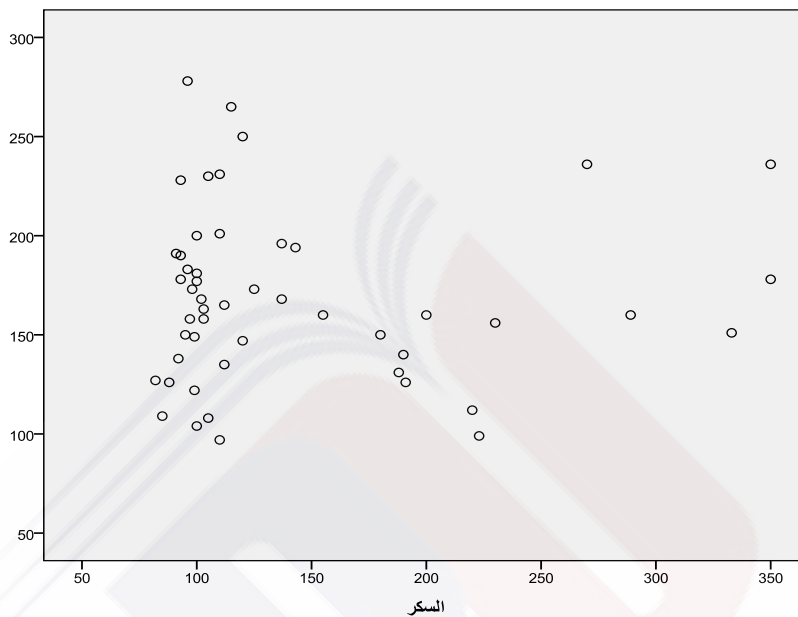
ومن ثمَّ تمثّل العينة بنقاط منتشرة في المستوى مبعثرة أحياناً ومتجمعة أحياناً أخرى، يدلنا شكل توزيعها على طبيعة العلاقة بين المتغيرين.

### مثال (9-1):

سكر الدم	الكوليسترول	سكر الدم	الكوليسترول	سكر الدم	الكوليسترول	سكر الدم	الكوليسترول	سكر الدم	الكوليسترول
230	156	270	236	99	122	100	181	143	194
91	191	97	158	333	151	137	196	289	160
95	150	105	108	96	183	350	236	100	104
180	150	190	140	120	147	110	201	223	99
188	131	103	163	112	135	100	200	137	168
200	160	92	138	155	160	120	250	88	126
103	158	98	173	93	178	110	97	82	127
115	265	99	149	125	173	105	230	93	228
350	178	102	168	112	165	85	109	191	126
100	177	220	112	110	231	93	190	96	278

الجدول (4-1) نسبة السكر والكوليسترول في الدم لخمسين مريضاً يعانون نشباتٍ دماغيةً مختلفةً .

تمثّل البيانات في الجدول الآتي قياسات معدل السكر في الدم والكوليسترول لخمسين مريضاً راجعوا مشفى الأسد الجامعي يعانون احتشاءاتٍ دماغيةً مختلفةً. لأخذ تصور أولي إذا كانت هناك علاقة ظاهرة بين السكر والكوليسترول لمجتمع الاحتشاءات ، نرسم شكل الانتشار ، ندخل قياسات أحد المتغيرين وليكن الأول (السكر) ونمثله على المحور الأفقي وقياسات المتغير الثاني ، ونمثله على المحور العمودي ، ونظهر شكل الانتشار للقياسات .



الشكل (10-1) الانتشار لمعدل الكوليسترول مع السكر في الدم لخمسين مريضاً

## تمارين ومسائل

1. البيانات الآتية تمثل ألوان عينة لنوع من الزهور:

حمراء - بيضاء - صفراء - زرقاء - زرقاء - بيضاء - بيضاء - خضراء -  
خضراء - حمراء - بيضاء - صفراء - زرقاء - زرقاء - بيضاء - زرقاء - صفراء -  
زرقاء - بيضاء - حمراء - خضراء - صفراء - حمراء - خضراء - حمراء -  
حمراء - حمراء - زرقاء - صفراء - زرقاء - صفراء - صفراء - خضراء - حمراء -  
صفراء - حمراء - حمراء - خضراء - خضراء - حمراء - صفراء - حمراء - حمراء -  
خضراء - حمراء .

أ- ما نوع هذه البيانات؟

ب- أوجد التوزيع التكراري والتكراري النسبي والتكراري المئوي.

ت- مثل التوزيع التكراري بيانياً.

ث- ارسم مخطط الفطيرة لهذه البيانات.

2. تمثل البيانات الآتية أطوال عينة من طلاب كلية الطب:

170 - 167 - 167 - 169 - 166 - 166 - 167 - 167 - 169 - 166  
169 - 168 - 169 - 170 - 167 - 168 - 168 - 170 - 166 - 168  
167 - 170 - 169 - 169 - 167 - 166 - 168 - 170 - 168 - 166  
169 - 170 - 170 - 168 - 168 - 170 - 168 - 167 - 167 - 169  
166 - 169 - 166 - 169 - 167 - 167 - 167 - 170 - 170 - 170

و المطلوب:



أ- ما نوع البيانات ؟

ب- أوجد التوزيع التكراري النسبي.

ت- مثل التوزيع التكراري بيانياً.

ث- مثل البيانات باستخدام مخطط الدائرة (الفطيرة).

3. جمع عالم تصنيف نباتات عينة من خمسة أنواع من النباتات في رحلة برية و حين فرزها وجد الآتي:

الصف	الأول	الثاني	الثالث	الرابع	الخامس
التكرارات	20	35	18	25	10

أ- ما نوع هذه البيانات ؟

ب- أوجد التوزيع التكراري و التكراري النسبي.

ت- ارسم المضلع التكراري النسبي.

ث- أوجد مخطط الساق و الورقة (اعتبر العشرات و المئات هي الساق و الأحاد هي الورقة).

4. أجرى طبيب صيدلاني تجارب على عينة من الفئران، و سجل النتائج ، ثم نظمها في الجدول التكراري الآتي:

مجال الوزن (غرام)	51-100	101-150	151-200	201-250
التكرار (عدد الفئران)	20	25	15	10

أ- ما نوع البيانات ؟

ب- أوجد التوزيع التكراري النسبي و المئوي.

ت- أوجد المدرج التكراري المتجمع الصاعد.

5. البيانات الآتية عبارة عن وزن السمكة (بالكيلوغرام):

2 - 2.5 - 2.5 - 4 - 3 - 2.5 - 1.5 - 1.5 - 4 - 2.5 - 1.5 - 3.5 - 2 - 3  
1.5 - 4 - 2.5 - 1.5 - 1.5 - 2.5 - 3 - 2.5 - 4 - 3.5 - 2.5 - 1.5 - 2  
3 - 1 - 2.5 - 2.5 - 2.5 - 2.5 - 3.5 - 1.5 - 2 - 3.5 - 2.5 - 1.5 - 3.5  
1.5 - 1.5 - 2 - 2 - 2.5 - 2 - 3.5 - 3.5 - 2.5 - 2

أ- أوجد مخطط الساق و الورقة.

ب- رتب البيانات تصاعدياً ، واحسب  $Q_1, Q_2, Q_3$  الربيعيات الأول والثاني والثالث.

ج- ارسم المخطط الصندوقي.

6. قيسست كمية البوتاسيوم المأخوذة من 50 محضراً في لتر من الدم ، و قدرت بالوحدات الدولية 1L of mm فكانت

4.8 - 4.6 - 4.3 - 4 - 4.9 - 4.7 - 3.3 - 4.3 - 4.9 - 4.8  
4.9 - 4.8 - 4.9 - 4.9 - 4.7 - 4.8 - 4.7 - 4.1 - 4.3 - 5.5  
4.6 - 4.4 - 4.6 - 4.9 - 4.1 - 4.2 - 4.5 - 4.3 - 4 - 4.7  
4.8 - 4.3 - 4.7 - 4.4 - 5.3 - 4.6 - 4.5 - 5 - 4.9 - 4.4  
4.7 - 4.2 - 5.1 - 4.1 - 4.4 - 4.1 - 4.4 - 4.9 - 4.7 - 4.8

و المطلوب:

أ- نظم هذه البيانات في جدول تكراري، و أوجد التكرارات المتجمعة الصاعدة.

ب- ارسم المنحني التكراري.

ت- احسب الربيعيات، وارسم المخطط الصندوقي.

7. نبين في الجدول الآتي احتمال البقاء على قيد الحياة لمختلف الأعمار لمجتمعين مختلفين:

الفئة العمرية	احتمال البقاء على قيد الحياة للمجتمع الأول	احتمال البقاء على قيد الحياة للمجتمع الثاني
0-10	0.96	0.93
10-20	0.95	0.9
20-30	0.94	0.88
30-40	0.92	0.86
40-50	0.88	0.80
50-60	0.76	0.75
60-70	0.53	0.6
70-80	0.21	0.4
80-90	0.02	0.1
90-100	0.01	0.05

والمطلوب:

ارسم مخطط الأعمدة لهذه البيانات لمقارنة احتمال البقاء على قيد الحياة بين المجتمعين لكل فئة عمرية، ماذا تستنتج؟

8. يبين الجدول الآتي قياسات التريغليسريد مصل الدم ل 60 طفلاً:

0.21 - 0.30 - 0.34 - 0.39 - 0.42 - 0.47 - 0.52 - 0.58 - 0.66 - 0.83 -  
0.26 - 0.32 - 0.35 - 0.40 - 0.44 - 0.48 - 0.54 - 0.50 - 0.79 - 0.88 -  
0.20 - 0.30 - 0.34 - 0.39 - 0.42 - 0.47 - 0.52 - 0.58 - 0.66 - 0.81 -  
0.27 - 0.32 - 0.36 - 0.40 - 0.44 - 0.48 - 0.53 - 0.60 - 0.72 - 0.96 -  
0.28 - 0.33 - 0.37 - 0.40 - 0.45 - 0.49 - 0.56 - 0.62 - 0.78 - 1.03 -  
0.29 - 0.15 - 0.46 - 0.50 - 0.64 - 0.78 - 0.90 - 0.92 - 0.98 - 1.20

أ- أوجد التوزيع التكراري لهذه البيانات مرة باستخدام أربع فئات فقط

(0.20,0.44) , (0.45,0.69) , (0.70,0.94) , (0.95,1.20)

و مرة باستخدام 10 فئات.

(0.20-0.29) , (0.30-0.39) , ..., (1.10-1.20)

ب- ارسم المدرج التكراري في كل مرة. أي الرسمتين أفضل لتمثيل البيانات؟ وهل تقترح عدداً آخر للفئات؟

ج- أوجد تمثيلاً بيانياً آخر تراه مناسباً للبيانات.

1. تمثل البيانات الآتية مستويات سكر الدم لمجموعة من طلاب السنة الأولى كلية الطب (mmol/l of 1L)

3.3 - 4.5 - 3.4 - 4.6 - 3.6 - 3.3 - 3.3 - 4.8 - 3.9 - 4 -  
4.4 - 4 - 3.6 - 4.1 - 4.7 - 2.2 - 3.8 - 3.6 - 4.7 - 5.1 -  
3.7 - 2.9 - 3.6 - 3.7 - 5 - 4.4 - 4.1 - 4.2 - 3.4 - 4.7 -  
4.9 - 4.4 - 3.8 - 4.1 - 3.8 - 4 - 3.4 - 6 - 4.3 - 4.9 -

و المطلوب:

أ- أنشئ مخطط الساق و الورقة لتنظيم البيانات.

ب- أوجد التوزيع التكراري للبيانات باتخاذ طول فئة يساوي 0.5.

ت- ارسم مخطط الدائرة.

ث- ارسم المضلع التكراري.