الفصل الثالث

عرض وتلخيص البيانات
Data Presentation and Summariza tion

بعد ان تجمع البيانات الاحصائية وفق الاساليب التي ذكرناها في الفصل الثاني تبدأ مرحلة عرض وتلخيص هذه البيانات مستندة على طبيعة البيانات والهدف الاساسي من جمعها وهناك ثلاث طرق أساسية لعرض وتلخيص البيانات وهي (1) طريقة الجداول و (2) طريقة العرض البياني و (3) طريقة حساب المقاييس الاحصائية المختلفة ويركز هذا الفصل على استخدامات هذه الطرق الثلاثة ومزاياها .

3 - 1 العرض الجدولي Tabular presentation

وفي هذه الطريقة تعرض البيانات الاحصائية بجداول ذات تقسيم واحد أو تقسيمات متقاطعة تعطي خلاصة للبيانات وفق تلك التقسيمات وبشكل يسمح للقارىء ادراك ماتتضمنه البيانات من معاني وابعاد واستنتاجات بسهولة ويسر . واليك عدة أمثلة للعرض الجدولي للبيانات .

مثال (3 - 1) : يبين الجدول رقم 3 - 1 عرضا ملخصاً لدرجة أصابة 100 شجرة بآفة زراعية معينة .

جدول (3 - 1) اصابة الاشجار بآفة زراعية معينة

درجة الاصابة	عدد الاشجار (التكرار)
خفيفة	70
متوسطة	20
شديدة	10
	100

ان القارىء يستطيع ان يدرك بسرعة من ان درجة الاصابةهي خفيفة في اكثر الاشجار وذلك لمجرد القاء نظرة على ملخص البيانات المعروضة بالجدول 3 – 1 اما مصطلح « التكرار » Fregueney فانه يعني عدد المرات التي تكررت فيها الصفة أو المجموعة قيد الدرس وعليه فان الجدول المذكور علاه يعتبرتوزيع تكراري للاشجار المائة حسب درجة الاصابة .

مثال (3 – 2) يبين الجدول رقم (3 – 2) ملخصاً للفئات العمرية للحائزين الزراعيين في منطقة ادارية معينة :

جدول (3 – 2) : الفئات العمرية للحائزين الزراعيين لمنطقة ادارية

فئة العمر (بالسنوات	عدد الحائزين الزراعيين (التكرار)	النسبة المئوية ٪
30 – 20	100	10
40 – 30	150	15
60 – 40	700	70
60 فما اكثر	50	5
المجمـــوع	1000	100

ويبدو واضحا ان اكثر الحائزين الزراعيين هم من فئة الاعمار المتقدمة . وهذا قد يعني اعراض الشباب عن اختيار الزراعة مهنة ومصدرا للعيش . الامر الذي يترتب عليه وضع خطط فعالة لتشجيع الشباب على العمل في هذا الميدان الحيوي وتجعل العديد منهم يختاره دون سواه .

مثال 3 – 3 : يبين الجدول رقم (3 – 3) نتائج تعفير البذور باجد السموم لمعرفة تأثيره على الانبات مقارنة مع ترك هذه البذور دون تعفير .

جدول (3 – 3) : تأثير تعفير البذور على الانبات في تجربة معينة

نوع البذور	الانبات		المجموع
	انبتت	لم تنبت	
معفرة	180	20	200
غير معفرة	30	70	100

ان الخلاصة التي يعرضها الجدول رقم (3 – 3) لا تترك مجالا للشك حول فاعلية وأهمية التعفير بالسم المذكور لضمان نسبة عالية للانبات .

ومن الأمثلة الثلاثة أعلاه يتضح ان الهدف من العرض الجدولي هو اعطاء الفكرة التي تتضمنها البيانات باسلوب سريع وبسيط ومباشر وخال من العمليات الحسابية المعقدة نوعا ما .

ولا بد لنا في هذه المرحلة ان نجلب انتباه القارىء الى بعض الاعتبارات الواجب ملاحظتها عند عرض البيانات جدوليا على شكل فئات وتكراراتها . وأهم هذه الاعتبارات ما يلي :

أ – طول الفئة :

ويقصد بطول الفئة المسافة (أو الفترة) بين القيمة الصغرى للفئة والقيمة الكبرى لها معبرا عنها بوحدات القياس الخاصة بالبيانات المراد ترتيبها بفئات . وعليه فان طول الفئة (10 – 20) سنة هو 10 سنوات . ولا يشترط ان تكون اطوال جميع الفئات متساوية خاصة اذا كانت هناك أسباب ومبررات ذات علاقة بالهدف من جمع البيانات تملي على عارض البيانات ترتيبها حسب فئات معينة وذات اطوال مختلفة كما هي الحال عند الرغبة في توزيع سكان المزارع لفئات عمرية تجاري الفئات التي تستخدم لوصف القوى العاملة والتي تأخذ الشكل المبين في الجدول رقم (3 – 3 أ) .

جدول (3 - 1) الفئات العمرية لسكان المزارع حسب مواصفات القرى العاملة

الفئة العمرية بالسنوات	طول الفئة (سنة)
أقل من 15	15
15 الى أقل من 65	50
65 فما فوق	غير محدد

ان الفئة العمرية (15 الى أقل من 65) والتي عادة ما تكون (15 - 65) تضم السكان النشاط الاقتصادي نسبيا تضم الفئات الأخرى السكان غير النشطين اقتصاديا . ويطلق على الفئة ذات الطول غير المحدد بالفئة المفتوحة أو المفتوحة الطرف (كما هي الحال بالنسبة للفئة 65 فأكثر) . ويلاحظ بان اطوال الفئات غير متساوية نتيجة ما تمليه مواصفات القرى العاملة

ب - عدد الفئات :

ويقصد بها عدد الفئات المراد توزيع البيانات عليها . وبهذا فان عدد الفئات في الجدول (3 - 2) يعادل (5) . وعند تحديد عدد الفئات على عدد البيانات الاحصائية المتوفرة قد يزداد ويقل تبعا لها . الا ان عدد الفئات يجب ان يبقى محدودا وكبير بما فيه الكفاية لتلخيصها وتحليلا وعمليا للبيانات .

جـ - توزيع البيانات الى فئات متساوية الطول :

اذا أراد الباحث توزيع بياناته الاحصائية الى فئات متساوية الطول لعدم وجود مبرر لتوزيعها الى فئات غير متساوية الطول . فلا بد له ان يحدد عدد الفئات وطول الفئة الواحدة . ويحسب عدد الفئات (m) حسب المعادلة (3 - 1) :

$$m = 1 + 3.3 \log n \quad \ldots (3 - 1)$$

حيث أن :

log يرمز الى اللوغاريتمات الاعتيادية (للاساس 10)

n يمثل عدد القيم الاحصائية (أي عدد البيانات)

ولا بد من تقريب عدد الفئات الى العدد الصحيح الاعلى اللاحق في حالة الحصول على عد[illegible]عشري كما هومتوقع في أكثر الحالات بسبب طبيعة المعادلة . ويبين الجدول رقم (3 - 4) أمثلة على عدد الفئات متساوية الطول لاعداد معينة من البيانات :

جدول (3 - 1) عدد الفئات متساوية الطول لعدد مختلف من البيانات

عدد القيم (n)	عدد الفئات المتساوية الطول (m)
10	5
100	8
1000	11
10000	15

ويبدو واضحا من الجدول (3 - 4) فعالية المعادلة الخاصة بتحديد عدد الفئات متساوية الاطوال كاسلوب لتلخيص البيانات خاصة عندما يكون عدد البيانات كبيراً .

وهناك معادلة أخرى لتحديد عدد الفئات والتي تعطي نفس نتائج المعادلة (3 - 1) اعلاه . أما هذه المعادلة فهي :

$$m = 2.5 \sqrt[4]{n}$$

أما طول الفئة (L) فانه يحسب على النحو التالي :

$$\text{طول الفئة} = \frac{\text{اعلى قيمة - أقل قيمة}}{\text{عدد الفئات}} \qquad L = \frac{R}{m}$$

وبعد تحديد عدد الفئات وطول الفئة يبدأ اعداد جدول تكراري ابتداء من اصغر قيمة في البيانات وانتهاء باكبرها مع مراعاة تساوي اطوال جميع الفئات كما هو موضح في المثال رقم (3 - 4) .

مثال (3 - 4) : البيانات التالية تمثل أوزان اللحم الصافي لـ 15 رأسا من الحملان :

12, 8, 8, 7, 5, 6, 8·5, 10, 8, 9, 11, 19, 7·5, 20, 14 (كغم)

فلو اريد عرض البيانات اعلاه بجدول تكراري ذي فئات متساوية الطول فان عدد الفئات وطول الفئة يحددان على النحو التالي :

(أ) عدد الفئـات =

$$m = 1 + 3{\cdot}3 \log(n)$$
$$= 1 + 3{\cdot}3 \log(15)$$
$$= 1 + (3{\cdot}3)(1{\cdot}176)$$
$$= 4{\cdot}88$$

وعليه ، فان عدد الفئات = 5 (بعد التقريب الى العدد الصحيح الاعلى)

(ب) طول الفئـة = $\frac{\text{أعلى قيمة} - \text{أقل قيمة}}{\text{عـدد الفئـات}}$ = $L = \frac{R}{m}$

$$= \frac{5 - 20}{5}$$

= 3 كغم

وعليه ، فان الجدول التكراري المطلوب مبين في الجدول رقم (3 - 5) :

جدول (3 5) : توزيع اوزان اللحم الصافي المعطاه في المثال (3-4) الى فئات متساوية الأطـوال

الفئـة (الوزن كغم)	التكـرار (عدد الحملان)
8 – 5	4
11 8	6
14 - 11	2
17 – 14	1
20 — 17	2
المجموع	15

ويلاحظ ان كل فئة تشمل على القيم التي تساوي بالضبط حدها الادنى وتقل عن حدها الاعلى . أما الاستثناء لهذه القاعدة فهو شمول أعلى قيمة ضمن الفئة الاخيرة لأن اعطائها فئة خاصة بها سيجعل عدد الفئات ستة بدلاً من خمسة وهذا مخالف لعدد الفئات الذي تم تحديده .

3 - 2 العـرض البيانـي (Graphic presentation)

تعرض البيانات احيانا باشكال مختلفة كالدوائر المجزأة والاعمدة والخطوط المتكسرة وغيرها بحيث يتمكن القارىء من معرفة الافكار والاتجاهات التي تتضمنها البيانات وذلك لمجرد القاء نظرة سريعة على الشكل البياني الذي يمثلها . أما مسألة اختيار هذا أو ذاك الشكل البياني فانها تعتمد على الهدف من العرض والامكانيات الفنية المتوفرة . وعادة ما يجد القارىء اشكالاً بيانية مختلفة في النشرات الخاصة بهذه أو تلك المؤسسة أو في المجلات والكتب العلمية . وسوف نركز على ثلاثة أنواع من العرض البياني بسبب شيوع استخدامها في البحوث العلمية والنشرات واللافتات الدعائية . وهذه الانواع هي :

أ - طريقـة الاعمـدة (Bars Diagram)
ب - طريقـة المـدرج التكـراري (Histogram)
ج - طريقـة المضلـع التكـراري (Polygon)

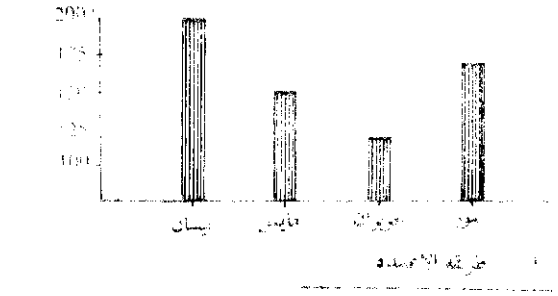
ولتوضيح اوجه الاختلاف بين الطرق الثلاثة للعرض البياني المذكورة أعلاه فاننا سوف نطبقها على المثالين (3 - 5 , 3 - 6)

مثال 3 - 5: يمثل الجدول 3 - 6 سعار الكيلوغرام الواحد من سلعة غذائية خلال أربعة أشهر متتاليـة :

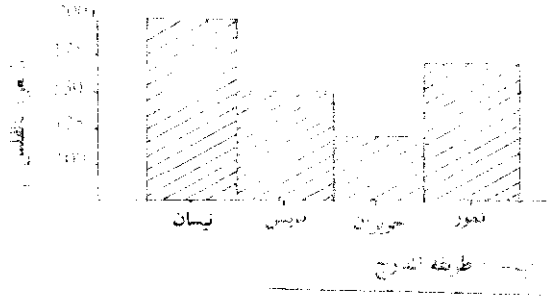
جدول (3 - 6) : اسعار سلعة غذائية حسب الشهر

الشهـر	السعر (بالفلس / كغم)
نيسان	200
مايس	150
حزيران	120
تمـوز	170

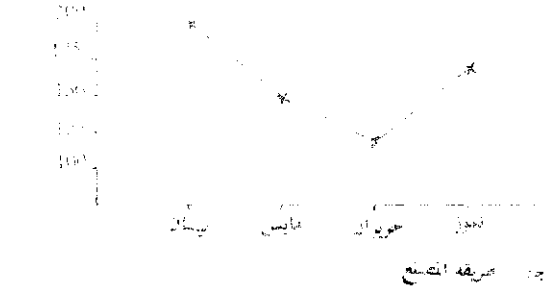
ويبين الشكل (3 - 1) طريقة عرض البيانات المشار إليها أعلاه حسب الأشكال البيانية الثلاثة - الأعمدة والمضلع والمدرج والتي تهدف جميعا إلى إعطاء الفكرة التي تتضمنها البيانات بأقل جهد وأسرع وقت ممكنين حيث تشير العروض البيانية الثلاثة إلى أن السعر بدأ عاليا ثم استمر في الانخفاض حتى شهر حزيران ثم عاد وارتفع ارتفاعا ملحوظا



أ - طريقة الأعمدة



ب - طريقة المدرج



جـ - طريقة المضلع

الشكل 3 - 1) عرض بيانات الجدول (3 - 5) حسب **طريقة الأعمدة** (أ) وطريقة المدرج (ب) والمضلع التكراري (جـ)

خلال شهر تموز . اما مسألة اختيار احدى الطريقتين للعرض البياني فانها تعتمد على الهدف من العرض . اذ قد يكون المدرج أكثر فاعلية في الاعلانات الدعائية كبيرة الحجم وخاصة في المحلات العامة والمطاعم حيث يتم التركيز على الاحجام النسبية لأجزاء المدرج مع الاستعانة بالقليل من الملاحظات الكتابية بالحروف البارزة . وقد يميل كاتب البحث العلمي الى اختيار طريقتي الاعمدة او المضلع تلافيهما في هذا المجال في أكثر الاحيان .

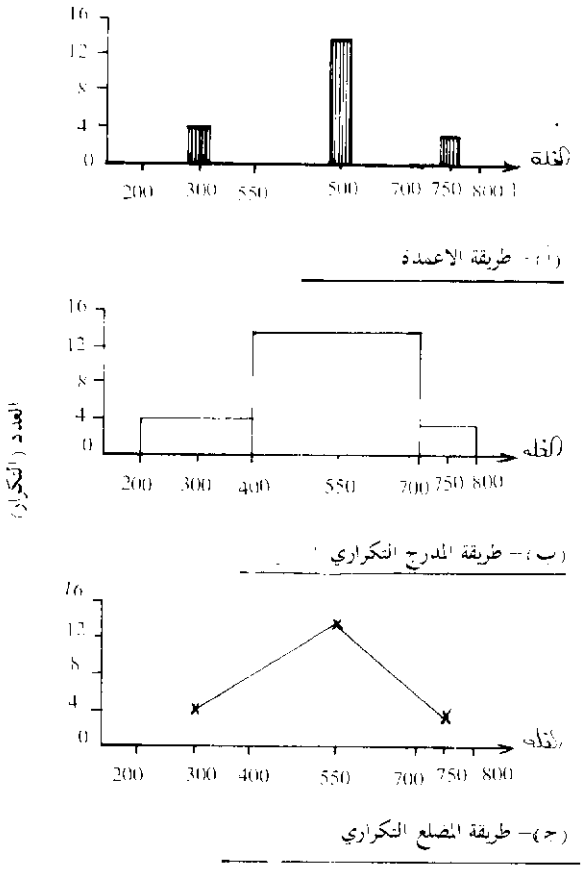
مثال (3 - 6) مثل الجدول التكراري (3 - 7) لعينة 20 سلالة من سلالات محصول الحنطة

جدول (3 - 7) عينة 20 سلالة من الحنطة

العلة للدونم (كغم)	عدد السلالات (التكرار)
400 - 200	4
600 - 400	13
800 - 600	3

وبين الشكل (3 - 2) كيفية عرض هذه البيانات بالاشكال البيانية الثلاثة .

ونجدر الاشارة الى أن العرض البياني بطريقة المدرج التكراري أكثر ملائمة للبيانات ذات الفئات وذلك لأنها تبين حدود كل فئة وحجمها النسبي وبذا تعطي الفكرة التي تتضمنها البيانات بشكل متكامل ومتفق مع خصوصية البيانات المجدولة وفق فئات معينة . أما اسلوبا العرض الاخران (الاعمدة والمضلع التكراري) فانهما مفيدان أيضا في حالة الرغبة في استخدامهما في مثل هذه الحالات مع مراعاة تمثيل كل فئة بمركزها والذي يحسب على النحو التالي : -



الشكل (3 - 2) عرض بيانات المثال (3 – 6) حسب طريقة الاعمدة (أ) ـ المدرج التكراري (ب) والمضلع التكراري (ج

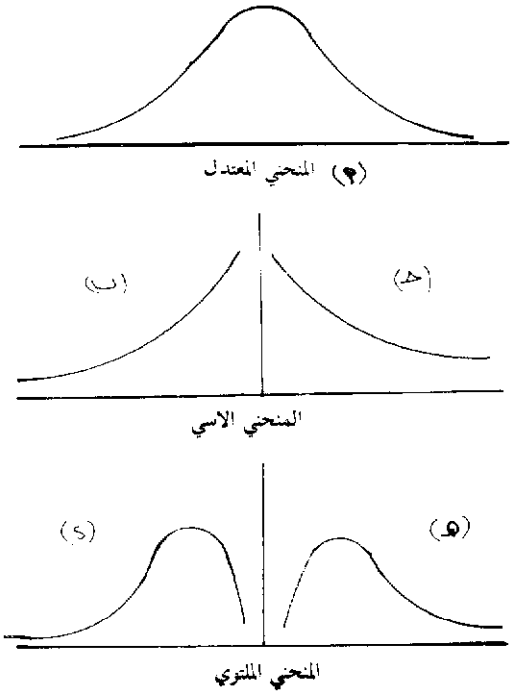
$$\text{مركز الفئة} = \frac{\text{قيمة الحد الادنى + قيمة الحد الاعلى}}{2}$$

ولابد من اشارة الى ان ازدياد عدد القيم في المحور الافقي (محور س) ، للمضلع التكراري وتقاربها يمهد الخط المتكسر الذي يمثل المضلع ويخفف من انكساراته فاذا كانت الصفة (المتغير) المعروضة بياننا وفق هذه الطريقة هي صفة مستمرة (متغير مستمر) كالطول والوزن ، حيث يمكن ان تتصل قيمها بعضها ببعض على المحور الافقي . فان المضلع

التكراري يبدأ بفقدان الانكسارات ويصبح ممهداً أكثر فأكثر مع زيادة عدد القيم. وقد يفقد الانكسارات تماما بعد مرحلة معينة فيصبح منحنى يطلق عليه « المنحنى التكراري (Freguency curve) اما شكل هذا المنحنى فانه يعتمد على طبيعة المتغير الذي يمثله وقيمته المركزية (المتوسط الحسابي مثلا) وتشتت قيم افراده عن تلك القيمة . وهنـــاك انواع متعددة من المنحنيات التكرارية . وسنذكر منها الامثلة التالية لاهميتها وشيوعها في المجال الزراعي(بشقيه النباتي والحيواني)بشكل خاص وبالمجالات البايولوجية بشكل عام :

أ- المنحنى التكراري المعتدل (Normal) أو الطبيعي :

وهو منحنى ذو قمة واحدة وعلى شكل ناقوس متماثل (Symmetrical) الجانبين (انظر الجزء آ من الشكل 3 - 3) . وعادة ماتكون المنحنيات التكرارية التي تصف الظواهر البايولوجية (الكمية منها بصفة خاصة) من نوع المنحنى المعتدل . **ولاتختلف المنحنيات**



(أ) المنحنى المعتدل

المنحنى الاسي

المنحنى الملتوي

الشكل (3 - 3) بعض الامثلة للمنحنيات التكرارية

المعتدلة في شكلها العام ولكنها تباين في حدة قمتها ومدى تباعد اطرافها عن مركز قممها اعتمادا على درجة التفاوت بين قيم افرادها . فكلما كان التفاوت قليلا زادت حدة القمة وتقلص انتشار المنحنى المعتدل والعكس بالعكس .

ب – المنحنى الاسي Exponential curve

وهو منحنى يتميز بتصاعد بطيء ثم يتزايد شيئا فشيئا (انظر الجزء ب من الشكل (3 - 3) او بتناقص سريع يتلوه تباطؤ ملحوظ (انظر الجزء ج من الشكل (3 - 3) وقد يمثل المنحنى الاسي المبين في الجزء (ب) . على سبيل المثال . عدد الخلايا البكتيرية خلال فترة معينة اما المنحنى الاسي المبين في الجزء (ج) فقد يمثل مستوى مضاد حيوي في دم حيوان معين خلال فترة زمنية بعد ادخاله في جسم ذلك الحيوان .

جـ – المنحنيات الملتوية Skewed curves

وتتميز هذه المنحنيات بتركيز قيمها في جانب من المنحنى بحيث يكون الانحراف هو مبين في الجزئين د و هـ من الشكل (3 - 3) .

3 - 3 المقاييس الاحصائية

ان عرض البيانات بطريقتي الجداول او الأشكال البيانية هو اجراء مفيد وفعال في تسهيل ادراك الصورة التي تعكسها البيانات . ورغم ذلك . فان اسلوبي العرض المذكورين لايخلوان من نقاط ضعف من أهمها :

أ – ان بامكان الباحث اعطاء افكار واستنتاجات مختلفة حول نفس مجموعة البيانات وذلك عن طريق اختيار تصنيف او تصنيفات معينة وفئات ذات اطوال خاصة وأشكال بيانية بوحدات قياسية كبيرة او صغيرة بحيث تسخر جميع هذه العوامل لاعطاء الفكرة التي يريد الباحث ان يعكسها للقارىء .

ب – ان هناك طرقاً أخرى لعرض وتلخيص البيانات والتي تأخذ بنظر الاعتبار خصوصية تلك البيانات بما في ذلك تركيزها حول قيمة معينة ومدى تشتتها عن بعضها البعض . وهذه الطرق عادة ماتعطي قيمة واحدة قادرة على اعطاء القارىء تصورا مفيداً لطبيعة توزيع قيم المتغير المدروس . وهذه القيم تمثل مقاييس احصائية