الفصل الثاني: البيانات وطرق تلخيصها

لقد تناولنا سابقا في الفصل التمهيدي التعريف بمقياس تحليل المعطيات بالإضافة الى التذكير ببعض اساسيات الجبر الخطي. حيث توصلنا الى ان هذا المقياس يبنى اساسا على مفهوم البيانات. هذه الاساسية تعتبر المادة الأولية لتلك الطرق . غير أنها تتواجد في الواقع بصور هائلة يصعب معالجتها يدويا. لذلك جاء هذا الفصل كخطوة لتبسيط بعض المفاهيم الاساسية المتعلق بهذه البيانات وكذا كيفية التعامل المبدئي معها, من حيث تبويبها ووضعها في صورة تجعلها جاهزة للحساب والتحليل.

II -1 المفاهيم الأساسية « Concepts fondamentaux »

فيما يلي سوف نتطرق الي بعض المفاهيم الاساسية المتعلقة بالبيانات كتعريف البيانات، قواعد البيانات ، مصادرها ،أنواعها.

- **تعريف البيانات**

البيانات هو ما هو معروف او مقبول علي هذا النحو. حيث يمكن للمرء ان يبني منطقا والذي يعمل كنقطة انطلاق للبحث (خاصة في حالة الجمع)، كما تعتبر العنصر الاساسي الذي يبني عليه العمل.

هي المسند او ما هو معروف وهو بمثابة نقطة بداية للتغيير في تجديد هدف حل مشكلة فيما يتعلق بهذه البيانات ، حيث يمكن ان يكون وصفا اوليا للواقع ، نتيجة مقارنة بين حدثين او بعبارة اخري هي عبارة عن ملاحظة او قياس.

تعد البيانات عنصرا هاما، لم يتم تفسيره بعد، يتم وضعه في السياق. وهنا يكمن الفرق بين المعلومات والبيانات. فالمعلومات بحكم تعريفها هي بيانات مفسرة وبعبارة اخري ، فان وضع البيانات في سياق معين يخلق قيمة مضافة لتشكيل المعلومات.

$$\text{Données} + \text{Valeur ajouté} \Leftrightarrow \text{information}$$

فالبيانات الخام تكون خالية من أي منطق ، افتراض، ملاحظة ، احتمالية ، مقارنة كما تعتبره انها غير قابلة للجدل لذلك يمكن استخدامها كأساس للبحث لأي فحص. لذلك يجب ان تخضع البيانات التي قد تختلف اختلافا كبيرا في طبيعتها اعتمادا علي مصدرها للتحويل المسبق قبل المعالجة.

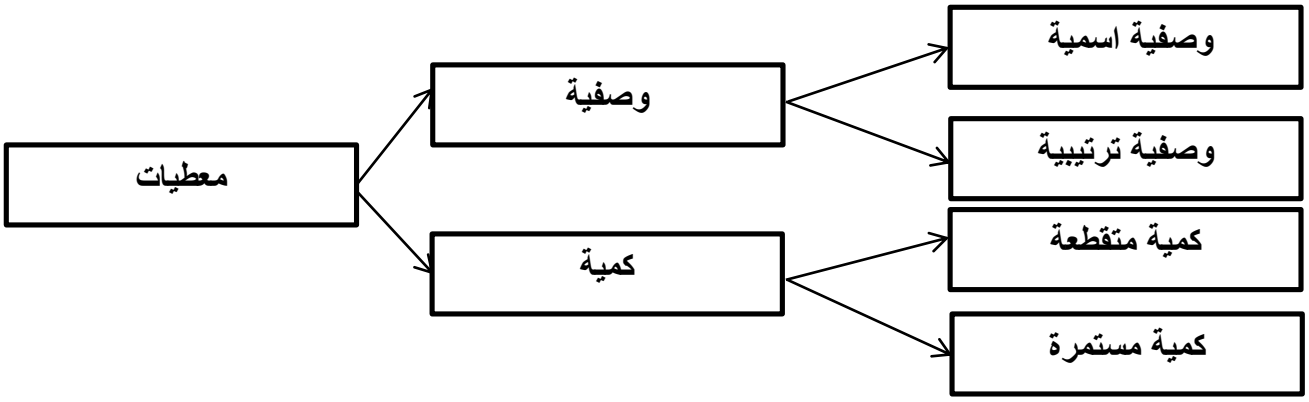
يمكن تعريفها ايضا بأنها مجموعة من المشاهدات أو الملاحظات التي تؤخذ أثناء دراسة معينة، وقد تكون بيانات رقمية (كمية) مثل أطوال وأوزان مجموعة من الطلاب أو دخول مجموعة من الأسر أو بيانات غير رقمية (وصفية) مثل لون البشرة والجنس ... إلخ.

- **انواع البيانات « Typologie de données »**

باعتبار البيانات هي عبارة عن وصف او مقدار والتي من خلالها يمكن تصنيف افراد في مجتمع معين. فانه يمكن التمييز بين نوعين من البيانات: كمية ووصفية (نوعية).فتحديد نوع البيانات له دور كبير في تحديد الاسلوب الاحصائي المتبع لذلك اولي علم الاحصاء اهمية بالغة لذلك. فهناك ما هو اجتماعي كدرجة التمدرس، درجة الفقر، درجة التشبع, وماهو اقتصادي: كمية الاستهلاك، الدخل، التضخم ,وماهو رياضي: معامل الرشاقة، الكتلة العضلية، قوة السحب والدفع ...الخ.

ويمكن تلخيص مختلف انواع البيانات في المخطط التالي:

المخطط 1.II: أنواع البيانات



- **البيانات الوصفية:** هي بيانات ويمكن قياسها أي ان مستويات التصنيف(Modalités) فهي عبارة عن اسماء غير عددية وغير رقمية وهي نوعان :
  - **وصفية اسمية**: هي بيانات غير رقمية تتكون من مجموعات متنافية ، كل مجموعة لها خصائص تميزها عن المجموعة الأخرى كما ان هذه المجموعات لا يمكن المفاضلة بينها أي ان معيار القياس غير قرشي مثال :
    - الحالة العائلية : اعزب- متزوج – مطلق
    - نوع الجنس: ذكر- انثي.

معيار القياس هو الذي يسمح قبلا بالمعالجة الممكنة لهذه البيانات ، وفي هذه الحالة فانه لا يمكن اجراء عمليات رياضية علي مجموع البيانات ما عدا المساواة والتي تكون بين مختلف مستويات التصنيف.

مثل:1=1 أي ان اي طالب فهو ينتمي الي نفس الصنف مع طالب اخر.

  - **وصفية ترتيبية:** وتتكون من فئات او مستويات تصنيف يمكن ترتيبها تصاعديا او تنازليا أي ان معيار القياس فيها هو الترتيب.

مثال:

- المستوي الدراسي: أمي، ابتدائي، ثانوي، جامعي.

حيث هنا يمكن اعطاء ارقام مرتبة من 1 الي غاية اعلي مستوي او العكس .

في هذه الحالة يمكن استخدام المساواة وكذا عمليات الترتيب اكبر واصغر فقط ام باقي العمليات الرياضية فلا يمكن اجراءها.

- **البيانات الكمية:** وهي البيانات التي يمكن قياسها وبالتالي تكون قيمها مقدرة عدديا فهي اذن رقمية .وتنقسم بدورها الي قسمين:

- **كمية منقطعة**: وهي البيانات التي تأخذها قيما عددية مستقلة عن بعضها البعض بحيث تكون صحيحة وغير قابلة للتجزئة.

مثال:

- عدد افراد العائلة : 1-2-3.

هنا جميع العمليات الرياضية تكون لها معني في هذه الحالة. بالإضافة الي ان 0 في هذه الحالة فله معني الغياب او الانعدام للظاهرة.

- **كمية مستمرة :** هي بيانات كمية تأخذ فيما في مجال معين كما يكن ان تأخذ جميع القيم داخل ذلك المجال. ومعيار القياس فيما هو معيار تعدي أي بالمقارنة الي نده لذلك فان 0 في هذه الحالة لا يعني انعدام الظاهرة. كما ان جميع العمليات الرياضية لها معين.

مثال :

- درجة الحرارة : 200^{o} -،50- ، 0 ، 70 .

- **مصادر البيانات « Source des données »**

من الوظائف الاساسية التي تختص بها علم الاحصاء هو جمع البيانات .ويكون بأسلوب علمي صحيح وذلك من اجل الوصول الي نتائج دقيقة يعتمد عليها في التحليل بحيث تكون من احدي المصادر التالية:

- **المصادر الاولية:** وهي المصادر التي يتحصل عليها الباحث بشكل مباشر بحيث يقوم بجمع البيانات من المصدر مباشرة وذلك عن طريق اجراء المقابلات ، الاستبيانات وغيرها. ويتميز هذا النوع من المصادر بالدقة والثقة ، واهم ما يعاب عليها انها تحتاج الي وقت ومجهود كبيرين هذا من ناحية ، كما انها مكلفة جدا من الناحية المادية من ناحية اخري.
- **المصادر الثانوية:** وهي المصادر التي يتحصل فيها الباحث علي البيانات بشكل غير مباشر، بمعني آخر يتدخل عنصر الوسيط سواء كان ذلك بشريا او ماديا مثل الهيئات الدولية والرسمية او اشخاص معينين.

من مزايا هذا النوع ، توفير الوقت والجهد والمال الا ان درجة ثقة الباحث فيها ليست بنفس الدرجة في حالته المصادر الاولية وهذا لوجود الكثير من الاختلافات التناقضات بين مصادر بيانات هيئة معينة بالمقارنة بهيئة اخري وهذا الاعتبارات عدة.

- **قواعد البيانات « Base de données »**

في الحقيقة تتوفر البيانات في شكل متعدد وبصورة كبيرة لذلك تكون في شكل قواعد تستخدم لتخزين البيانات وتنظيمها وتحليلها. فقواعد البيانات التي يرمز لها بالرمز (Base de données=BDD) هي مجموعة من البيانات المنظمة التي تسمح بالبحث بسهولة وادارتها وتحديثها. داخل هذه القواعد يتم تنظيم البيانات في صفوف واعمدة وجداول ، يتم فهرستها بحيث يمكنك بسهولة العثور علي البيانات التي تبحث عنها

باستخدام برامج كمبيوتر خاصة في كل مرة يتم فيها اضافة معلومات جديدة ، كما يتم تحديث البيانات وربما حذفها.

هذه الجداول التي تحوي مجموع البيانات يطلق عليها عادة اسم المصفوفات. كما انها توجد في اشكال مختلفة. كما يمكن تعريفها علي انها مجموعة من البيانات المرتبطة وذات الصلة مرتبطة بطريقة معينة بحيث يمكن البحث فيها وتحديثها بسهولة ويتم فيها في شيء التكرار.

2-II تشكيل جداول البيانات « Construction des tableaux de données »

وتنطلق هذه العملية من عملية ادخال البيانات المحصلة في جداول خاصة بهدف ابرازها وتوضيحها في أضيف نطاق ممكن لتكوين فكرة عنها وهذا بعد معالجتها. ويختلف اسلوب تبويب البيانات تبعا لطبيعتها .

فبالنظر الى انواع البيانات فانه يمكن تبويبها ايضا الي كمية ووصفية. ويختلف شكل الجداول طبقا لنوع البيانات، وحسب عدد المتغيرات. يمكن التمييز بين ثلاث انواع من الجداول في علم الاحصاء والتي يتم تشكيلها, حيث نجد : جداول البيانات , جداول التوزيع للمتغيرات و كذا جداول الاقتران. سوف نتطرق الى كل منها وكذا الى مختلف المميزات الخاصة بكل منها وكذا الوظائف الخاصة المنوطة بها (من الجانب الاحصائي طبعا). حيث نهتم في ذلك بعدد الهوامش المشكلة لها وكذا طبيعة كل هامش.

وفيما يلي عرض لبعض انواع الجداول الممكن تشكيلها :

- **جداول المعطيات « Tableau de donnée »**

تكون جداول المعطيات عادة في حالة الإجرائيات الاحصائية العلمية , حيث تمثل اول الاشياء التي يتم اعداها. وهي الجداول التي تسمح بتسهيل ووضع في الاعتبار معالجة البيانات. هذه الجداول يتم عرضها عموما كجداول كبيرة الحجم لا نها تحتوي الكبير من الاسطر بقدر عدد المواضيع. كل خانة منها تعطي لنا معلومة عن هذه المواضيع حيث ان :

هوامش او حدود الجدول هي أثنان. تتموضع على اليسار وكذا اعلى الجدول. الاولى غالبا تكون غير مسماة, حيث تمثل قائمة المواضيع . أما الثانية وتمثل قائمة المتغيرات. القائمة الاولى وتحوي العناصر 1, 2, 3 الخ . أو a, b, c الخ . بحيث ان كل عنصر يعرف لنا موضوع معين. القائمة الثانية وتحوي مجموع العناصر التي تتمثل في المتغيرات. في بعض الاحيان ما يكون ترتيب هذه المتغيرات عشوائيا اما في الحالات الاخرى فيكون وفقا لتنظيم معين. في جميع الاحالات فان هذه الهوامش يمكن قراءتها قبل كل شيء على انها قوائم إعدادية وليس كتغيرات للمتغير .

ومع ذلك فان جدول المعطيات الحقيقي يمكن ان يعرف عن طريق حجمه. ففي الحقيقة فان هذا الجدول هو الاول من حيث الاعداد في الدراسات الاحصائية لمجتمع ما . أما الجداول الاخرى فهي معدة من خلال هذا الجدول ونتيجة لذلك يكون حجمها أقل . حيث يمنح لنا الرؤية الفورية لمجموع البيانات.

- **جدول II-1** : مثال عن جدول المعطيات

	Annonce première marge	*Deuxième marge : liste des variables informées*						
	N° de sujet	sexe	âge	question 1			question2	
				réponse1	réponse 2	réponse3	réponse 1	réponse2
Première marge : liste des sujets	1	F	22	x			x	
	2	F	21	x				x
	3	M	20	x				x
	4	M	22	x			x	
	5	F	23	x			x	
	6	M	21		x			x

هذا الجدول يساعد على:

- الاستغلال السريع لتوزيع المتغيرات. ينظر إليه وفقا للعمود او وفقا لمجموعة من الاعمدة.
- المقارنة بين توزيع متغيرتين من خلال التجاور المرئي لعمودين,
- تسليط الضوء على مواضيع معينة وبالتالي القراءة عن طريق الاسطر ، انظر تجاور السطور.

تنشأ المهام التي تطورت لهذا النوع من الجداول من الوظائف المدرجة. ربما تكون المهمة الأولى المرتبطة بهذا النوع من الجداول هي اعداده. يتطلب تحديد المتغيرات وتسمياتها ، وتحديد مجتمع الدراسة ، وهو أمر بعيد عن الوضوح.

تتعلق المهام الأخرى بالاستخدام المباشر لجدول البيانات. من بينها ، دعنا نقتبس استخراج عمود وتكوين جدول إحصائي يتوافق مع وصف توزيع متغير ، واستخراج عمود واحد أو أكثر ومجاورهم ، واستخراج سطر واحد أو أكثر ومجاورتهم.

- **جدول التوزيعات الاحصائية « Tableau statistique de distribution »**
- **جدول التوزيع لمتغيرة :**

هذا النوع من الجداول هو الاكثر شيوعا. يتم تشكيله او اعداده بعد استخراج احد الاعمدة او أكثر لجدول المعطيات. ويتم الحصول عليه بواسطة تجميع الخلايا المتماثلة للأعمدة المستخرجة و عن طريق تعدادها. يقوم بشرح توزيع متغيرة بمعنى أخر نقل المراسلات بين مختلف مستويات تصنيف المتغيرة وكذا عدد المواضيع التي يمثلها مستوى التصنيف.

هذا الجدول عادة ما يتكون من اثنين أو ثلاث أسطر. تظهر القائمة حتما وهي بمثابة الهامش الايسر: انها القائمة التي يتم تحديد ترتيبها بواسطة الانضباط . أولا يكون اسم المتغير, ثم يليه المصطلح " مشاهدات" أو " عدد الافراد...», ثم في النهاية المصطلح " تكرارات ". عادة ما يكون الهامش الثاني على يمسن الجدول : ويحوي العناصر المتمثلة في المجاميع وكذا عدد المفردات ثم في الاخير النسب المئوية.

الهامش الثالث يمكن تحديده مكانيا , حيث يتمثل في السطر العلوي للجدول والذي يحوي مختلف الصفات (او مستويات التصنيف). الخانة العلوية على اليسار تلعب دور المشر او المعلن عن هذه القائمة.

اذن يمثل هذا الجدول جدول التوفيقات يتكون من مجاورة او تركيب مجموعة من القوائم وفقا لترتيب محدد. الجزء الداخلي للجدول هو رقمي أي مجموعة من الارقام الصحيحة والعشرية.

جدول II-2 : مثال توضيحي عن جدول احصائي لتوزيع متغير

		Troisième marge : liste des modalités de la variable					
Première marge	nombre de pages annonce →	[0,100]	]100,200]	]200,300]	]300,500]	total	*Deuxième marge*
	nombre de livres	6	10	12	7	35	
	fréquences	17 %	29 %	34 %	20 %	100 %	

كما يمكن ان يظهر على الشكل التالي :

جدول I-3 : مثال معدل للجدول السابق

		Troisième marge : liste des modalités de la variable					
Première marge	x_i	[0,100]	]100,200]	]200,300]	]300,500]	total	*Deuxième marge*
	n_i	6	10	12	7	35	
	f_i	17 %	29 %	34 %	20 %	100 %	

الوظيفة الاساسية من ملا هذا الجدول هو تسهيل او تبسيط وصف المتغير المقصود دراسته. أما بالنسبة للمهام المنسوبة اليه هي مقارنة الاعمدة من حيث تشابه او اختلاف الافراد أو من حيث تشابه واختلاف الاصناف.

- **جدول التوزيعات الاحصائية**

هي الجداول التي يتم تشكيلها عن طريق التصاق الجداول الاحصائية لمتغير واحد. عدد اسطر الجدول الابتدائي يكون مضاعف. هذا الجدول يسمح بمقارنة التوزيعات لنفس المتغيرة من أجل مجتمعات مختلفة.

هذه الجداول, مثل الجداول الاحصائية البسيطة , يحدها هوامش على اليسار والتي تتمثل في قائمة عددية. وحده العنصر الاول الذي يظهر في القائمة يكون محددا: وهو اسم المتغير الذي سكون محل الدراسة. ثم يأتي بالترتيب الذي هو مؤلف الجدول في عرض غير شامل وهي عدد المواضيع او الاشياء لمختلف المجتمعات.

كما في السابق يمكننا تعريف هامش علوي, والذي يتمثل في قائمة مستويات التصنيف للمتغير الدراسة. هذا الهامش, يكون في النهاية معلنا في الخانة العلوية يسارا. أما بالنسبة لمحتوى الجدول او داخل الجدول فيتكون من الارقام الصحيحة.

جدول II-4: مثال عن هذا النوع من الجداول

		Troisième marge : liste des modalités de la variable					
Première marge	nombre de pages annonce →	[0,100]	]100,200]	]200,300]	]300,500]	total	*Deuxième marge*
	nombre de manuels de mathématiques	6	10	12	7	35	
	nombre de manuels de français	10	15	16	8	49	
	nombre de manuels d'histoire	14	25	4	18	61	

- **جدول الاقتران "Tableau de contingence"**

تتشكل هذه الجداول من تقاطع متغيرتين (في الغالب وصفيين). حيث أن اعداده يتم بالعودة الى الجدول الاولي للبيانات وكذلك عد الاشياء التي تمثل في ان واحد القيمتين المقابلة للمتغيرتين . كما ان هذه الجداول يحدها اربعة حدود. الحدين الأساسين على اليسار والاعلى. واللذان يمثلان مستويات التصنيف الخاصة بكل متغيرة. فاذا كان لدينا المتغيرتين Yو X فان مستويات تصنيف كل منهما هي علي الترتيب $(X_1, X_2, \dots, X_r)$ و$(Y_1, Y_2, \dots, Y_s)$. يقوم الجدول المشكل من اعطاء عدد المفردات التي تحمل في ان واحد مستوي التصنيف للمتغير x_i ومستوي التصنيف y_j للمتغير Y . أما الحدين المتبقين فاحدهما على اليمين والاخر في الاسفل فيمثلان قيم المجاميع الهامشية على السطر وكذا العمود. جدول الاقتران يكون من الشكل :

جدول II-5 : مثال توضيحي عن جدول الاقتران

	Annonces	*Deuxième marge : variations de nombre de pages*				*Annonce*	
Première marge variations discipline	nombre de pages/ discipline → ↓	[0,100]	]100,200]	]200,300]	]300,500]	sous totaux ↓	Troisième marge
	mathématiques	6	10	12	7	35	
	français	10	15	16	8	49	
	histoire	14	25	4	18	61	
Annonce	sous totaux →	30	50	32	33	145	
		Quatrième marge					

تتمثل وظيفة هذا الجدول في السماح بدراسة العلاقة بين هذين المتغيرين وتحديد نوع الرابط الذي يدمع بينهما.

II-3 : تلخيص البيانات « Synthèse de données »

بعد الحصول علي البيانات سواء من مصادرها الاولية أو الثانوية , فإنها عادة ما تكون بيانات رقمية خام وغير منتظمة, يصعب تحليلها. (تكون مقدمة في شكل سلسلة من القيم في حالة المتغيرات الكمية). لذلك دعت الحاجة الي وضعها في شكل جداول. فتقديمها بهذه الطريقة، يجعلها قابلة للتحويل والمعالجة من أجل الحصول علي معلومات "ذات معني". لذلك كان من الضروري ايجاد طرق لتلخيص هذه البيانات خصوصا وأنها تتواجد في الواقع بكميات هائلة. فالإحصاء الوصفي هو الوسيلة التي تستخدم في تلخيص الكم الهائل من هذه البيانات وهذا عن طريق سواء العرض البياني والذي يكون في شكل رسومات بيانية خاصة أو عن طريق العرض الرياضي حيث يمكن التعبير عن مجموع القيم بالمتوسطات والتي هي عبارة عن القيمة النموذجية الممثلة لمجموع هذه البيانات .وفيها تميل القيمة النموذجية إلي الوقوع في المركز، لذلك تسمى هذه القيم بمقاييس النزعة المركزية. ومن بين المتوسطات الأكثر شيوعاً المتوسط الحسابي والوسيط والمنوال. بينما لا تعتبر مقاييس النزعة المركزية كافية للتعبير عن البيانات و الاختلافات بين المجموعات وخاصة في حال وجود عدد كبير منها. .لذلك نشأت الحاجة إلي إيجاد مقاييس تقيس درجة تجانس أو تقارب او تشتت مفردات البيانات بعضها عن بعض , وتعرف هذه المقاييس بمقاييس التشتت وسالتي سوف نستعرض منها كلا من المدى والتباين وكذا الانحراف المعياري.

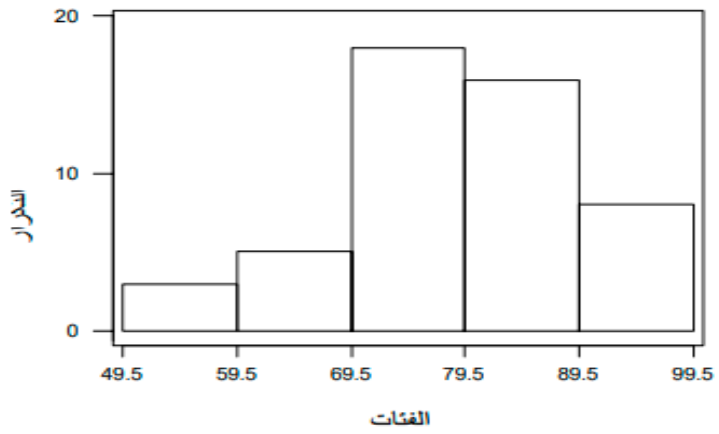
- **طريقة الرسومات البيانية**

يعتبر العرض البياني شكلا من الأشكال الهامة في تلخيص البيانات وفي الحقيقة يعطينا معلومات عن البيانات أسرع من الجداول، أما عن الرسومات البيانية الخاصة بالمتغيرات الكمية والاكثر استعمالا هي:المدرج التكراري، المنحني التكراري والمضلع التكراري.

- **المدرج التكراري**

يرسم المدرج التكراري علي محورين متعامدين أحدهما أفقي يمثل الفئات والثاني رأسي يمثل التكرار. ونرسم مستطيلات متلاصقة علي الفئات قاعدتها طول الفئة محسوبا من الحدود الحقيقية، وارتفاعاتها عبارة عن تكرار هذه الفئات. فمثلا بالنسبة إلي الفئة الأولي يكون المستطيل قاعدته بادئة من الحد الأدنى للفئة الأولي، ومنتهية بالحد الأعلى للفئة الأولي. وارتفاع المستطيل هو تكرار الفئة الأولي. وهكذا لباقي المستطيلات التي تمثل باقي التكرارات والمدرج التكراري للبيانات الموجودة في الجدول.

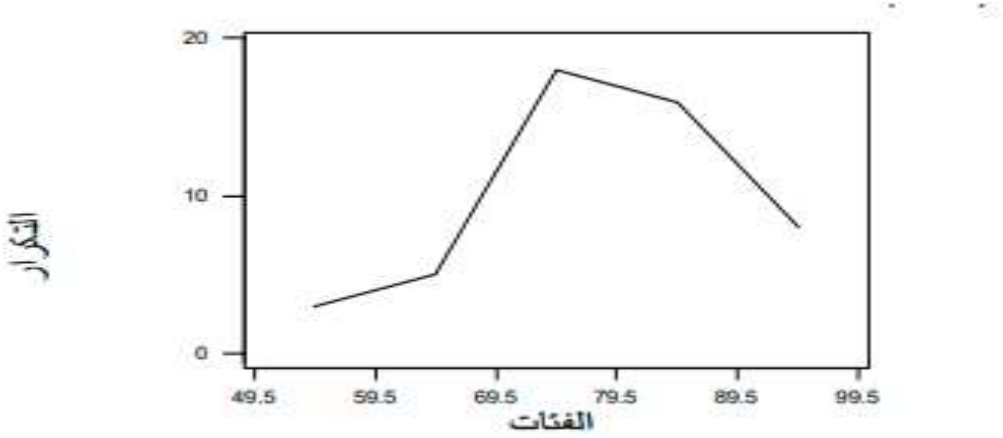
المخطط 2.II: مثال توضيحي عن المدرج التكراري



- **المضلع التكراري**

يرسم المضلع التكراري علي محورين، الأفقي يمثل الفئات والرأسي يمثل التكرار، مثل ما ورد شرحه في طريقة رسم المدرج التكراري، وبدلا من رسم مستطيل ارتفاعه يمثل التكرار نضع نقطة واحدة فقط علي ارتفاع يمثل التكرار لهذه الفئة وذلك عند منتصف الفئة. ويكرر رسم النقاط لباقي التكرار بحيث تكون ارتفاعاتها ممثلة لتكرار تلك الفئات وذلك من منتصفاتها، لأننا نفترض انتظام توزيع التكرارات داخل كل فئة. وبعد ذلك نصل بخط مستقيم كل نقطتين متجاورتين فنحصل علي المضلع التكراري.

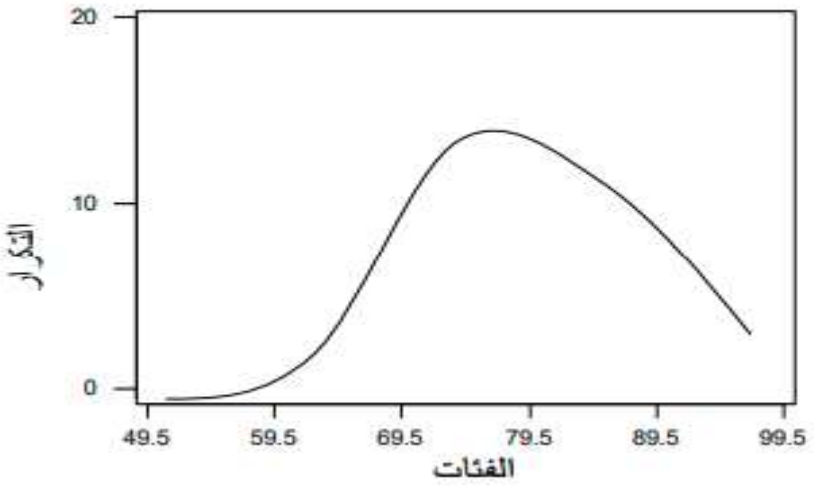
المخطط II.3: مثال توضيحي عن المضلع التكراري



- **المنحنى التكراري**

يرسم المنحنى التكراري على محورين متعامدين الأفقي يمثل الفئات والرأسي يمثل التكرار. ويتم رسم النقاط مثل ما اتبع في المضلع التكراري، ويمهد المنحنى التكراري باليد كي يأخذ شكلاً انسيابيا حتى لو لزم الأمر عدم المرور ببعض النقاط. المنحنى التكراري الممهد للبيانات في الجدول .

المخطط II.4: مثال توضيحي عن المدرج التكراري



- **الطريقة الرياضية:**

من المعروف عادة أن الرسوم البيانية تكون غير دقيقة خصوصا في حالة الكم الهائل من البيانات، لذلك يجب أن يكون لدينا مقاييس عددية تصف لنا هذه البيانات. وسوف نتعرض في هذا الباب إلى نوع مهم من المقاييس الإحصائية وهو ما يسمى بمقاييس النزعة المركزية أو مقاييس الموضع أو المتوسطات. وهى مقاييس عددية تعين موقع التوزيع، وهى مهمة في حالة المقارنة بين التوزيعات المختلفة بوجه عام. وتكون فائدتها أكثر فى حالة التوزيعات المتشابهة في طبيعتها وأشكالها ولكنها مختلفة في مواقعها. فمثلا: عند دراسة عينة من البيانات الإحصائية التي تخص بعض الأسر من الريف حسب فئات الأنفاق الاستهلاكي السنوي، وعينة من البيانات الإحصائية التي تخص بعض الأسر من الحضر حسب فئات الانفاق الاستهلاكي أيضا، فإن حساب المتوسط السنوي للإنفاق لكل من الريف والحضر يمكننا من المقارنة بينهما.

- **عن طريق مقاييس النزعة المركزية:**
  - **الوسط الحسابي (المتوسط)**

يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداما في الإحصاء والحياة العملية، إذ يستخدم عادة في الكثير من المقارنات بين الظواهر المختلفة. ولو أسندت قيمة المتوسط لكل مشاهدة فإن مجموع هذه القيم الجديدة يكون مساويا لمجموع المشاهدات الأصلية

- **الوسيط**

عند ترتيب البيانات (المشاهدات) ترتيبا تصاعديا أو تنازليا فإن الوسيط يكون البيان (المشاهدة (التي يقع 50 % من البيانات قبلها في الترتيب و 50 % من البيانات بعدها في الترتيب. فإذا كان عدد البيانات فرديا فإن الوسيط يكون المشاهدة التي تقع في المنتصف، وإذا كان عدد البيانات زوجيا فإن الوسيط هو متوسط المشاهدتين اللتين تقعان في المنتصف.

- **المنوال**

يعرف المنوال علي أنه القيمة الأكثر شيوعا (تكرارا) في مجموعة البيانات. وقد يكون لمجموعة البيانات منوال واحد ولذلك يطلق عليها وحيدة المنوال، أو يكون لها أكثر من منوال وتسمي متعددة المنوال. وقد لا يكون لمجموعة البيانات أي منوال وبذلك تسمي عديمة المنوال.

- **تلخيص البيانات عن طريق مقاييس التشتت**

لقد سبق لنا دراسة طرق عرض البيانات جدوليا وبيانيا والتعرف علي أشكالها وتوزيعاتها المختلفة وكذلك دراسة مقاييس النزعة المركزية (المتوسطات) وذلك لوصف البيانات عدديا لهذه التوزيعات المختلفة. ولكن طرق عرض البيانات وحساب المتوسطات للمجموعات المختلفة من البيانات غير كافٍ للمقارنة بين هذه المجموعات.

- **المدى**

يعرف المدي بأنه الفرق بين أكبر قراءة وأصغر قراءة في مجموعة القراءات أي أن: المدي = R أكبر قيمة – أصغر قيمة وذلك في حالة البيانات المباشرة. أما في حالة البيانات المبوبة فإن المدي يعرف بأكثر من طريقة، نذكر منها الطريقتين الآتيتين:

١-المدى = الفرق بين مركزي الفئة العليا والفئة الدنيا .

٢- المدى = الحد الأعلى للفئة العليا مطروحا منه الحد الأدنى للفئة الدنيا.

❖ **الانحراف المعياري**

من الصعب التعامل رياضيا (تحليليا) مع الانحراف المتوسط، ولذلك دعت الحاجة إلي استخدام مقياس للتشتت بنفس قوة الانحراف المتوسط، ولكي يكون من السهل التعامل مع تحليليها، وبما أن الفكرة هي التخلص من الإشارات للانحرافات فإن تربيع الانحرافات يخلصنا من الإشارة. ولهذا فإن الانحراف المعياري يعرف عن طريق التباين والذي يعرف علي أنه متوسط مربع انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز σ .