

الإحصاء- المحاضرة الأولى

Statistics for robot engineering-Lecture 1

Dr Fadi KHALIL

Doctor lecturer in statistics and programing



مقدمة:

أصبح من المعروف في العالم اليوم أنّه لتعلم أو دراسة شيء ما لا بدّ من جمع بيانات تتعلق به. والإحصاء ببساطة هو فن التعلم من هذه البيانات. يهتم بالدرجة الأولى بجمع البيانات، وتوصيفها، ومن ثمّ تحليلها تمهيداً لاستخلاص النتائج منها. وفي الواقع ليس من البديهي تنفيذ هذه المهام خاصة في ظل عدم التأكد والتقلبات الحادة التي أصبحت تتسم بها البيانات هذه الأيام.

من هنا فإنّ علم الهندسة يتناول الإحصاء من ناحية أهميته في قياس دقة المخرجات ودراسة العوامل التي تؤثر على في هذه الدقة (variability). فالمهندس معني أكثر من غيره بفهم مجموعة العوامل العشوائية التي تؤثر على خواص وأبعاد ودقة المنتج النهائي الذي يعمل عليه.

على سبيل المثال من المهام البديهية لمهندس الروبوت هو تصميم ربوت ذكي يحاكي السلوك البشري لتنفيذ مهمة معينة. إذا نقطة البداية ترتكز عند توصيف المهمة، تحديد مقاييس التنفيذ، اقتراح النموذج والتصميم، وتنفيذ التجربة تمهيداً لاختبار جودة الأداء. ولكن هناك احتمال كبير أنّ هذه الخطوات قد لا تتحقق بشكل فعال من المرة الأولى، بل تتطلب أحياناً الإعادة وبالتالي جمع معلومات، تمهيداً لتحليلها وتفسيرها، لفهم خصائص الروبوت من ناحية حساب متوسط سرعة الحركة مثلاً، أو لانحراف ع المقياس المحدد للتنفيذ.

مما سبق يمكن فهم دور الإحصاء كمساعد في اقتراح تصميم منتج جديد، تحسين تصميم حالي، أو تطوير عملية الإنتاج.

يهدف مقرر الإحصاء إلى إطلاع طالب الهندسة على المفاهيم الإحصائية التي تساعد في توظيف الإحصاء في علم الهندسة. هذه المفاهيم تتعلق بأنواع البيانات وطرق تمثيلها بيانياً describing data، ومؤشرات الإحصاء الوصفي describine statistics، بالإضافة إلى أهم أدوات التحليل الإحصائي. يتضمن المقرر أيضاً عرض مبسّط لمفهوم المتغير العشوائي والتوزيع الإحتمائي probability distribution مع شرح لأهم التوزيعات الإحتمائية المنقطعة والمستمرة والعديد من المواضيع ذات الصلة. علاوة على ذلك، يتناول المقرر مفهوم بالغ الأهمية في علوم الهندسة وخاصة التطبيقية منها وهو مراقبة الجودة quality control حيث يعرض كيفية الإستفادة من الأدوات والتقنيات الإحصائية في خدمة هذا الغرض وهو ما يعرف بالعديد من الأمثلة التي ترتبط ما يميز هذا المقرر أنّه لا يكتفي بعرض للمفهوم الإحصائي فقط بل يقوم بعرض العديد من الأمثلة التي ترتبط بالقضايا التي تصادف المهندس في عمله وأبحاثه مما يساعد على استيعاب المفاهيم الإحصائية وتطبيقها في الوقع.

يعمل المقرر في نهايته على تطبيق المواضيع التي تمّ التطرق إليها على برنامج التحليل الإحصائي SPSS وهذا انطلاقاً من عضوية العلاقة بين علم الإحصاء من جهة والبرمجة والحاسوب من جهة أخرى.

يمكن تلخيص مخطط المقرر بما يأتي:



- 1. علم الإحصاء، أنواع البيانات وطرق جمعها
- statistical inference مفهوم الاستدلال الإحصائي.
- 3. طرق الوصف الإحصائي Describing data method
- 4. مقاييس النزعة المركزية والتشتت Central tendency and dispersion measures
 - 5. مفهوم التوزيع الاحتمالي Probability distribution
 - a. التوزيعات الإحتمالية المنقطعة Discrete probability distribution
 - b. التوزيعات الاحتمالية المستمرة Continuous probability distribution
 - 6. اختبار الفرضيات Hypothesis Testing
 - 7. قياس العلاقة والتوافق بين المتغيرات Correlation and correspondence
 - 8. مفهوم الانحدار التقليدي Traditional regession
- 9. مفهوم الانحدار المضبوط كأداة لتعلم الآلة Machine learning: regularized regression
 - Ridge regression .a
 - Lasso regression .b
 - Elastic net .c
 - 10. مفهوم مقايس لورينز ومعامل جيني The Lorenz curve and Gini index
 - 11. مراقبة الجودة وخرائط مراقبة الجودة Quality control and control charts
 - 12. تطبيق على برنامج الحسابات والرسوم البيانية R



في المحاضرة الأولى سيتم التطرق للمواضيع الآتية:

- 1. علم الإحصاء، أنواع البيانات وطرق جمعها
- 2. مفهوم الاستدلال الإحصائي statistical inference
 - 3. أنواع المتغيرات، وطرق تمثيلها بيانياً
 - 4. مقاييس النزعة المركزية
 - a. الوسط الحسابي Arithmetic mean



1- علم الإحصاء:

الإحصاء هو مجموعة من الطرق التي تساهم في تقديم ملخص عن ظاهرة مدروسة. وهو يعبر عن إجمالي البيانات الرقمية وغير الرقمية المتعلقة بهذه الظاهرة. يعبر علم الإحصاء أيضاً عن مجموعة الطرق العلمية التي تستخدم لـ

- [1] لجمع ووصف البيانات.
- [2] ترتيب وتصنيف البيانات.
 - [3] عرض وتمثيل البيانات
- [4] وتحليل البيانات باستخدام الادوات الإحصائية.

من الضروري إذا في البداية التعرف على أنواع البيانات ووسائل جميع البيانات، قبل التعرف على الأساليب الإحصائية المستخدمة لتحليل هذه البيانات.

2- أنواع البيانات Data measures:

عموماً يوجد أربعة أنواع أو مستويات للقياس مرتبة تصاعدياً من البسيط إلى الأكثر وضوحاً وهى القياس: الإسمى، والرتبي، الفترى، والنسبي ويمكن المقارنة بين الأنواع الأربعة على النحو الآتي:

أمثلة	الخصائص القياسية	العمليات الرياضية	المستوى
المهنة الجنس الحالة الاجتماعية نوع الحاسوب	-عدد لا يدل على كم أو مقدار (أعداد منفصلة) - الأرقام تحل محل الأسماء، - لا يمكن ترتيب الأرقام ترتيب تصاعدي او تنازلي الأرقام تمثل فئات – وضع الأشخاص في فئات - لا تمثل الأرقام كميات من الخصائص التي يملكها الفرد تميز الأرقام بين المجموعات لا يمكن إجراء العمليات الحسابية على الأرقام	العد	الاسمى Nominal
المستوى التعليمي، درجة الإصابة بمرض ما، درجة أداء محرك السيارة	- كم لا يشار إليه بعدد (قيم منفصلة) - الأرقام مرتبة ترتيباً تنازلياً أو تصاعدياً يمتم بترتيب الأفراد في الخاصية، ولكن لا يعبر الرقم عن كمية امتلاك الفرد من الخاصية، بل تدل على أن الرتبة تملك من الخاصية مقدار أكبر أو أصغر من رتبة أخرى. مثلاً إذا	الترتيب العد	الرتب <i>ي</i> Ordinal



	كان لدينا المقياس الآتي		
	1- حزين، 2- غير حزين 3- لا بأس 4- سعيد، 5- سعيد جداً		
	هنا نعرف أنّ الرمز 2 يدل على مقدار ما من السعادة أكبر من		
	الرمز 1، ولكن لا نعرف إذا كان الفرق بين 3 و 2 هو نفسه		
	الفرق بين 4 و 5 مثلاً. من هنا نرى أنّ المقياس الرتبي لا يعبر		
	بالضرورة عن فروق متساوية حسابياً في الرتب		
	لا يمكن تنفيذ العمليات الحسابية عليه مباشرة.		
	- عدد يدل على كم أو مقدار (قيم متصلة).		
	- وضع الأشخاص في مقياس متصل يتكون من مسافات		_
	متساویة وله صفر اعتباری ولیس مطلق.		
عند تقسيم درجات	- الفرق بين الرتب يعبر عن الفرق بين الدرجات. مثلاً الفرق	.,	
الطلاب إلى مستوبات وكل	بين 60 و65 درجة حرارة هو قابل للقياس وهو 5 درجات وهو	الجمع	
مستوى يأخذ قيمة	نفسه الفرق بين 40 و 45 درجة.	الضرب	الفترى Interval
معينة، ودرجات الحرارة،	- ليس له مقياس نسبي بمعنى إنّ درجة الحرارة 10 مئوية لا	* **	intervar
قوة المحرك بالحصان	تعني أنّها ضعف درجة الحرارة 20 مئوية، أي إذا النسبة	الطرح	
	10/20 ليس لها معنى مع هذه النوع من البيانات لأنّه مثلاً		
	بتحويل هذه الدرجات إلى مقياس فهرنهايت تصبح 50 و 68		
	على الترتيب وبنسبتهما يظهر رقم أخر50/68.		
	- عدد يدل على كم أو مقدار (قيم متصلة)		
	- وضع الأشخاص في مقياس متصل يتكون من وحدات		
الطول	متساوية وله صفر مطلق		
الوزن	- يمكن استخدام النسب لمقارنة الأرقام.	جميع	
الوري	يمكن إجراء العمليات الحسابية عليه.	. يى العمليات	النسبي Ratio
درجة الأمبير	- له مقياس نسبي، بمعني إذا نسبنا دخل موظف والذي يبلغ	الرياضية	
الفولت	- له مقياس نسبي، بمعنى إدا نسبنا دخل موطف والذي يبلغ 20 ألف على دخل موظف أخر والذي يبلغ 10 آلاف أي		
العولت	20 الله على دخل موطف الحر والذي يبلغ 10 الدف اي 20 الله على الحراد على الله الموطف الأول		
	يملك دخل أكبر من الثاني بمقدار الضعف.		
	يست و حل مير بن مدي بمستدر مسيد		



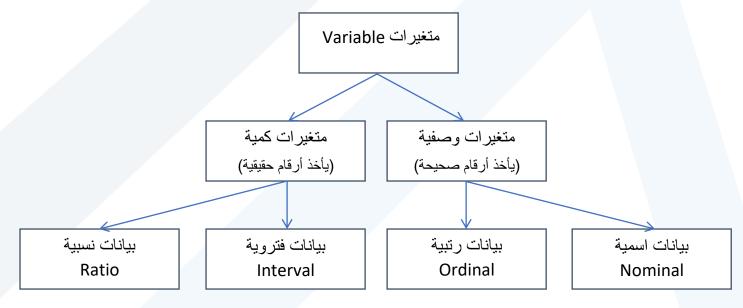
نود الإشارة إلى أن معرفتنا بمستويات قياس المتغيرات التى نستخدمها تسهل علينا اختيار الأداة الإحصائية
 المناسبة لتحليل البيانات المدروسة.

3- أنواع المتغيرات:

يمكن تمييز نوعين رئيسين من المتغيرات:

- [1] متغيرات مستمرة تأخذ قيم مستمرة ضمن مجال معين (مقدار ثخانة سلك ملم،).
- [2] ومتغيرات منقطعة تأخذ قيم صحيحة فقط او متقطعة (مثل نجاح تجربة أم لا).

ولكن بالإضافة لهذين النوعين الرئيسيين، تلعب أنواع البيانات التي ذكرناها سابقاً دوراً في تحديد انواع أخرى. ويمكن تلخيص أنواع المتغيرات وفقاً للمخطط التدفقي الآتي:



4- مصادر جمع البيانات (Collecting data Sources):

تنقسم مصادر جمع البيانات إلى نوعين: مصادر تاريخية ومصادر ميدانية:

- [1] المصادر التاريخية: ويقصد بها السجلات المحفوظة والبيانات التي يتم نشرها من قبل الشركات والمؤسسات أو المنظمات الدولة.
 - [2] المصادر الميدانية: يتم وفقاً لها جمع البيانات بطريقة مباشرة من قبل الباحث وذلك عن طريق:
- a. المقابلة الشخصية (Personal interview): حيث يقوم الباحث بمقابلة أفراد المجتمع او العينة المراد دراستها، ويتم توجيه الأسئلة لكل فرد وتسجيل إجابته. ومن عيوب هذه الطريقة أنّها تستغرق جهداً وتكاليف مادية عالية.



- لاستمارة الإحصائية (الاستبيان Questionary): حيث يقوم الباحث بتصميم استمارة تشتمل أسئلة تتعلق بالبحث، ويقوم بتوجيه الاستبيان إما عن طريق الحضور الشخصي، البريد، الانترنت، عبر الهاتف، أو مواقع التواصل الاجتماعي.
- c. التجربة (Experiment): حيث يقوم الباحث بإجراء تجربة ويسجل نتائجها، وذلك للحصول على معلومات مفيدة. فمثلاً لقياس تأثير بعض العوامل في نظام ميكانيكي معين يتم بتصميم التجربة لقياس أداء هذا النظام ومراقبة المؤشرات المتعلقة بكوناته.

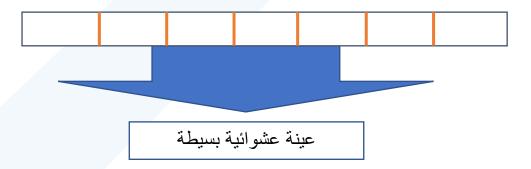
5- أساليب جمع البيانات (Collecting data methods):

يتم جمع البيانات بأحد الأسلوبين التاليين:

[1] الحصر الشامل (Complete Census): حيث يتم جمع البيانات من جميع أفراد مجتمع الدراسة، ويستخدم في المجتمعات الصغيرة (مدرسة، شركة، ..) أو في حالات تنفيذ البحث على المستوى الوطني مثل التعداد السكاني. وتكمن قوة هذا الأسلوب في إعطاء الباحث صورة حقيقية وكاملة عن مجتمع الدراسة، ومن عيوبه تكاليفه الباهظة وطول المدة الزمنية اللازمة لإجرائه.

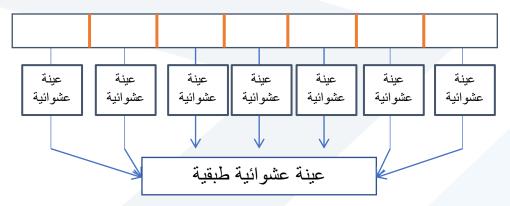
العينات (Sampling): يستخدم أسلوب العينات عند دراسة مجتمعات كبيرة جداً، ويمكن تعريف العينة على العينات (Sampling): يستخدم أسلوب العينات عند دراسة ويمثل جميع خصائص المجتمع بصدق. ويوجد عدة طرق لسحب العينة:

a. العينة العشوائية البسيطة (Simple Random Sample): هنا يتم سحب العينة من المجتمع بشكل مباشر وتستخدم في الحالات التي يكون فيها المجتمع متجانس.

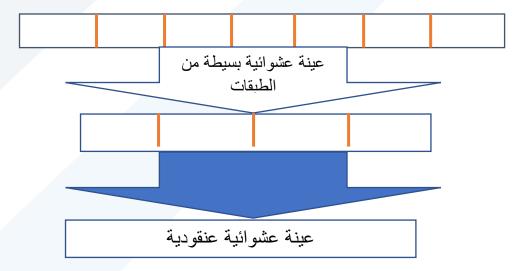


d. العينة العشوائية الطبقية (Stratified Random Sample): تستخدم هذه الطريقة عندما يكون المجتمع غير متجانس، حيث يتم تقسيم المجتمع إلى طبقات ومن ثم عينة عشوائية بسيطة من كل طبقة. مثل تقسيم طلاب الجامعة إلى اختصاصات وسحب عينة عشوائية بسيطة من كل اختصاص، والاختصاص هنا يمثل الطبقة.





العينة العشوائية العنقودية(Cluster Random Sample): تستخدم هذه الطريقة عندما يكون المجتمع غير متجانس ويمكن تقسيمة إلى عدد كبير من الطبقات، هنا يتم سحب عينة عشوائية بسيطة من الطبقات، ومن ثم سحب عينة عشوائية بسيطة من كل طبقة. مثلاً، لسحب عينة لدراسة سرعة الاستجابة في مخدمات الإنترنت لشركة غوغل العملاقة ، حيث يتم تقسيم فروع غوغل الكثيرة والمنتشرة في العالم إلى طبقات ومن ثم يتم سحب عينة عشوائية بسيطة من هذه الفروع (الطبقات) تمهيدا لسحب عينة من المخدمات من كل فرع.



d. العينة العمدية او القصدية (Purposive sample): حيث يتم اختيار عينة تناسب هدف البحث، بحيث توفر في كل عنصر من العينة شروط محددة يرى الباحث أنّها تساعده في الحصول إلى نتائج أفضل في دراسته (مثال اختيار الطلاب الذكور أو الإناث لإجراء دراسة عليهم أو اختيار مجموعة من المصابين بمرض معين لتطبيق الدراسة عليهم).

من أسلوب جمع البيانات عن طريق سحب العينة ينشأ مفهوم مهم جداً وهو ما يسمى بالاستدلال (Statistical Inference).



6- مفهوم الاستدلال الإحصائي:

الاستدلال الإحصائي هو أسلوب أو منهج عمل يدل على دراسة خصائص مجتمع ما عن طريق عينة مأخوذة من هذا المجتمع. أو بعبارة أخرى هي تعميم الخصائص المستمدة من عينة ما على المجتمع الذي أخذت منه هذه العينة. وبالتالي قد يكون هذا المفهوم مرادف لمصطلح التعميم الإحصائي.

بالتأكيد قد يكون الاستدلال الإحصائي مترافقاً مع أخطاء تسمى أخطاء القياس والتقدير. ويقاس هذا الخطأ عن طريق أدوات أو صيغ رياضية تسمى إحصائيات.

وبمقدار ما تكون طريقة سحب العينة مثالية وحقيقية بمقدار ما يكون خطأ التقدير صغيراً، والعينة المثالية هي العينة التي تمثل المجتمع الأم أفضل تمثيل. على سبيل المثال، إذا كان لدينا عينة تتألف من مفردات مأخوذة من مدرسة واحدة لطلاب الصف الخامس، فمن المحتمل أن تكون هذه العينة غير قابلة للتعميم على مجتمع الصف الخامس في المدينة بأكملها وبالتالي خطأ التقدير قد يكون كبيراً.

يقول Grosso modo من أجل ضمان أن تكون العينة ممثلة للمجتمع بشكل جيّد، من الضروري أن تسحب مفردات العينة بشكل كافي. حيث أنّه كلّما اقترب عدد مفردات العينة من عدد مفردات المجتمع الأم، يكون خطأ تقدير معالم المجتمع صغيراً.

7- تبويب، عرض وتمثيل البيانات:

بعد جمع البيانات مهما كان مصدرها (تاريخية أو ميدانية) أو أسلوب جمعها (حصر شامل، أو عن طريق عينة)، تكون البيانات خاماً ليست مرتبة أو منظمة. وبالتالي يصعب دراستها أو معرفة خصائصها وتحليلها. وبالتالي لا بدّ من تبويها وترتيها، ومن ثم عرضها وتمثيلها بيانياً لكي يسهل دراستها واستخلاص بعض النتائج منها.

7-1- الجداول التكرارية (Frequency table):

بالإضافة لأنواع المتغيرات وبياناتها التي رأيناها سابقاً، يمكن ذكر تصنيف آخر للبيانات يتعلق بالبيانات الوصفية (اسمية، رتبية) و الكمية (فئوية نسبية). حيث تختلف طريقة عرض وتمثيل كل نوع من هذه البيانات كما سنرى:

7-1-1- الجداول التكرارية والتجميعية للبيانات الوصفية:

يتضمن الجدول التكراري للبيانات الوصفية بشكل رئيسي عمودين الاول يتضمن يحتوي الوصفية والثاني يحتوي على التكرار المقابل لكل صفة. ويمكن أن يتضمن أعمدة أخرى مثل النسب التكراراية أو التكرار التجميعي الصاعد. ويطلق على الجدول التكراري اسم التوزيع التكراري لأنّه يوزع لكل صفة التكرار المقابل لها.

مثال 1:

ليكن لدينا الجدول الآتي الذي يبين مستوبات 60 معاجل حاسوبي في أحد المختبرات:



المستوى	التكرار Frequency	النسبة Percent	النسب التجميعية
ممتاز	14	23,3	23,3
جيد جداً	14	23,3	46,7
جيد	12	20,0	66,7
مقبول	12	20,0	86,7
ضعیف	8	13,3	100,0
Total	60	100,0	

7-1-2 الجداول التكرارية والتجميعية للبيانات الكمية Freauency table :

تختلف طريقة إنشاء الجداول التكرارية للبيانات الكمية عن البيانات الوصفية، لأننا لو اعتبرنا أنّ كل قيمة من قيم المتغير الكمي هي فئة بحد ذاتها لنتج لدينا جدول توزيع تكرار يحوي عدد كبير من الفئات. لذلك يعتبر من الضروري تجميع كل مجموع من القيم ضمن فئة، بحيث يتم تقسيم جميع القيم إلى فئات متساوية الطول وإعطاء كل فئة عدد التكرارات للقيم المتضمنة في هذه الفئة. وبتم إنشاء الفئات باتباع الخطوات الآتية:

- [1] ترتيب البيانات ترتيباً تصاعدياً من الأصغر إلى الأكبر.
- [2] حساب عدد الفئات بما بتناسب مع البيانات، بحيث يتراوح بين 5 و 15 فئة، أو يمكن استخدام قاعدة ستارجس (Sturges rules) لتحديد عدد الفئات (k) لبيانات عددها (n):

$$k = 1 + 3.322log(n)$$

- [3] تحديد الفئات بحساب طول الفئة، وذلك عن طريق حساب المدى وهو الفرق بين أكبر قيمة (Xmax) وأصغر قيمة $R = X_{max} X_{min}$ (Xmin) قيمة (Xmin)، وقسمته على عدد الفئات (W = R/k) التي تم حسابها في الخطوة السابقة. وبالتالي يكون طول الفئة (W = R/k) مساوياً لـ W = R/k
- و X_{min} و X_{min} و القيم بين $L_1=X_{min}$ ، ويتم وضع فيها كل القيم بين X_{min} و X_{min} ... [4] يتم إنشاء حدود أول فئة الثانية، الثانية، X_{min} ... X_{min} ...

مثال 2:

ليكن لدينا البيانات الآتية عن 30 طالب المتعلقة بعلامات أحد المقررات في كلية الهندسة:

45 40 57 43 56 63 39 68 44 49 61 47 58 34 45 59 48 54 75 54 48 37 62 40 50 54 58 48 49 46

المطلوب إنشاء جدول التوزيع التكراري لتمثيل علامات الطلاب في 7 فئات.

نلاحظ أنّ هذه البيانات من النوع الكمي حيث إذا اعتبرنا أنّ لكل قيمة فئة ينتج لدينا الكثير من الفئات، وبالتالي لا بدّ من تقسيم هذه القيم إلى فئات كما يأتي:



[1] طالما أنّ عدد الفئات مُعطى وهو 7 فئات فلا داعي لاستخدام قاعدة ستاردج (sturges).

على عدد $R=X_{max}-X_{min}=75-34=41$ على عدد $R=X_{max}-X_{min}=75$ على عدد الفئات k=7

قوم بإنشاء الفئة الأولى:
$$W = \frac{R}{k} = \frac{41}{7} = 5.7$$
 $[L_1 = X_{min}, U_1 = L_1 + w[= [34,34+6[= [34,40[:34,40[:34]]]]]$ $[L_2 = U_1, U_2 = L_2 + w[= [40,40+6[= [40,46[:40,46[:34]]]]]]$ الفئة الثانية : $[L_3 = U_2, U_3 = L_3 + w[= [46,46+6[= [46,52[:34]]]]]]$ الفئة الرابعة: $[L_4 = U_3, U_4 = L_4 + w[= [52,52+6[= [52,58[:34]]]]]]$ الفئة الخامسة: $[L_5 = U_4, U_5 = L_5 + w[= [58,58+6[= [58,64[:34]]]]]]$ الفئة السادسة: $[L_6 = U_5, U_6 = L_6 + w[= [64,64+6[= [64,70[]]]]]]$ الفئة السادسة: $[L_7 = U_6, U_7 = X_{max}]$

[4] يتم إنشاء جدول التوزيع التكراري كما يلي:

النسبة التجميعية	f_i/n النسبة	f_i التكرار	الفئات
%10	$\frac{3}{30} * 100 = 10\%$	3	[34,40[
%30	$\frac{6}{30} * 100 = 20\%$ $\frac{8}{30} * 100 = 27\%$	6	[40,46[
%57	30	8	[46,52[
%77	$\frac{6}{30} * 100 = 20\%$	6	[52,58[
%94	$\frac{5}{30} * 100 = 17\%$	5	[58,64[
%97	$\frac{1}{30}$ * 100 = 3%	1	[64,70[
%100	$\frac{1}{30} * 100 = 3\%$	1	[70,75]
	1	30	المجموع



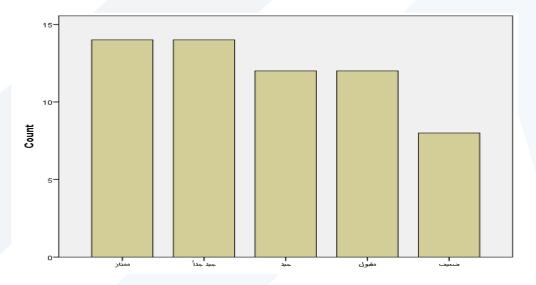
7-2- العرض البياني للجداول التكرارية Graphic representation:

7-2-1- المدرج التكراري (Frequency Histogram):

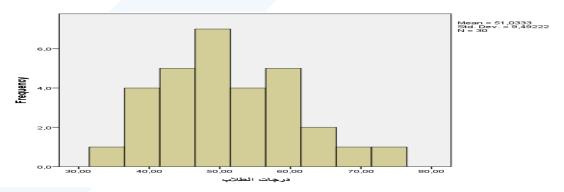
المدرج التكراري يتكون من مجموعة من الأعمدة المتجاورة التي تتوضع على المحور الأفقي (التي تمثل حالات الصفة المدروسة عندما تكن البيانات وصفية وفترات الفئات عندما تكون البيانات كمية) بحيث يساوي ارتفاع كل عمود مقدار التكرار المقابل للفئة.

نعود للمثالين السابقين:

تمثيل بيانات المثال 1:



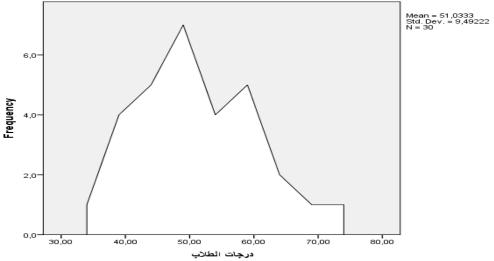
مثال 2:



7-2-2 المضلع التكراري (Frequency Polygon):

هو عبارة عن خط منكسر يصل بقطع مستقيمة النقاط التي مساقطها مراكز الفئات، ويبدأ الخط من الصفر ثم ينتهي بالصفر، للتوضيح أكثر لنرى الشكل البياني الآتي والذي يمثل المضلع التكراري لدرجات الطلاب:





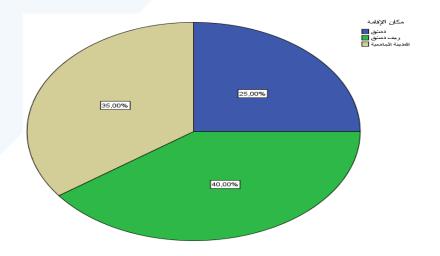
7-2-3 مخطط الدائرة (Pie Charts):

تعتبر هذه الوسيلة من وسائل العرض البياني التي تتيح للمتهم التعرف بنظرة سريعة على الظاهرة المدروسة. حيث يتم تجزئة الدائرة (360) إلى مجموعة من القطاعات التي تتناسب قياسات زواياها مع تكرارات الفئات.

مثال 3:

ليكن لدينا البيانات الآتية عن مكان إقامة 100 طالب إحدى الجامعات:

مكان الإقامة	التكرار
المدينة الجامعية	35
دمشق	25
ریف دمشق	40
المجموع	100





حيث يتم حساب زاوية القطاع بضرب 360 بالنسبة المقابلة لكل فئة:

الزاوية	التكرار	مكان الإقامة
360*(35/100)=126	35	المدينة الجامعية
360*(25/100)=90	25	دم <i>ش</i> ق
360*(40/100)=144	40	ریف دم <i>ش</i> ق
360	100	المجموع

4-2-7- شكل الانتشار (Scatter diagram):

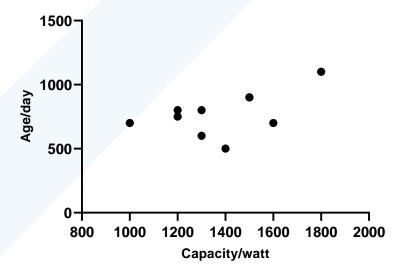
يمثل هذا النوع من الرسوم البيانية توزع قيم متغير (محور العامودي) بالنسبة لمتغير آخر (محور أفقي)، حيث يناسب البيانات الكمية. كذلك يمكن الاستفادة منه في معرفة مدى وجود علاقة بين المتغيرين.

مثال 4:

ليكن لدينا البيانات الآتية عن الإستطاعة والعمر (بالأيام) لـ 10 أجهزة:

1200	1300	1400	1600	1500	1800	1200	1300	1500	1000	watt/ capacity
800	600	500	700	900	1100	750	800	900	700	day/Age

يمكن استخدام شكل الانتشار لتمثيل العلاقة:



يساعد هذا الأسلوب من الرسم البياني على الكشف عن وجود علاقة بين المتغيرين المدورسين، وفيما إذا كان انتشار قيم متغير وفقاً لمتغير آخر تشكل نموذج (pattern) واضح (خطي أو لا خطي). إضافة لتحديد شكل هذه العلاقة فيما إذا كانت طردية أو عكسية.



يمكن تلخيص أنواع وسائل العرض البياني السابقة وفقاً لنوع البيانات كما يأتي:

نوع البيانات Data type	طريقة العرض البياني
وصفية Qualitative	المدرج التكراري، المضلع التكراري، الدائرة
کمیة Quantitative	المدرج التكراري (فئات)، المضلع التكراري (فئات)، شكل الانتشار