الإحصاء والاحتمالات – المحاضرة الأولى

Statistics and probabilities-Lecture 1

Dr. Fadi Khalil, Dr.Ali Ahmed

Dr Soummaya Abdul-Hak

Doctor lecturer in statistics and programing

2025

١.١ مقدمة:

لتعلم أو دراسة شيءٍ ما، لا بدّ من جمع بيانات تتعلق به. وأصبح من المعروف في العالم اليوم أن الإحصاء ببساطة هو فنّ التعلم من البيانات. يهتم بالدرجة الأولى بجمع البيانات، وتوصيفها، ومن ثم تحليلها لاستخلاص النتائج منها.

في الواقع، ليست هذه المهام بديهية التنفيذ، خاصة في ظل عدم التأكد والتقلبات الحادة التي تتسم بها البيانات هذه الأيام.في علم الهندسة، يتناول الإحصاء من ناحية أهميته في قياس دقة المخرجات ودراسة العوامل التي تؤثر في هذه الدقة (variability) فالمهندس معنيّ أكثر من غيره بفهم مجموعة العوامل العشوائية التي تؤثر على خواص وأبعاد ودقة المنتج النهائي الذي يعمل عليه.من مهام المهندس بشكل عام توصيف المشكلة، تحديد العوامل المؤثرة بها، اقتراح نموذج التصميم، وتنفيذ التجربة تمهيدًا لاعتماد النموذج وتطويره لاحقًا.كل هذه الخطوات تتطلب جمع معلومات، تحليلها وتفسيرها لفهم كيفية عمل النموذج وحل المشكلة قيد الدراسة.مما سبق يمكن فهم دور الإحصاء كمساعد في اقتراح تصميم منتج جديد، تحسين تصميم حالي، أو تطوير عملية الإنتاج.

يهدف مقرر الإحصاء والاحتمالات إلى إطلاع طالب الهندسة على المفاهيم الإحصائية التي تساعد في توظيف الإحصاء في المجال الهندسي.

هذه المفاهيم تتعلق بـ:

- أنواع البيانات وطرق تمثيلها بيانياً(Describing Data)
- مؤشرات الإحصاء الوصفي(Descriptive Statistics)
- عرض مبسّط لمفهوم المتغير العشوائي والتوزيع الاحتمالي(Probability Distribution)
- أهم التوزيعات الاحتمالية المنقطعة والمستمرة
- مفهوم مراقبة الجودة (Quality Control) وكيفية الاستفادة من الأدوات والتقنيات الإحصائية في هذا المجال (Statistical Quality Control)

ما يميّز هذا المقرر أنه لا يكتفي بعرض المفاهيم النظرية، بل يقدم أمثلة واقعية ترتبط بالقضايا التي يواجهها المهندس في عمله وأبحاثه، مما يساعد على فهم المفاهيم الإحصائية وتطبيقها في الواقع. في نهاية المقرر يتم تطبيق المواضيع على برنامج التحليل الإحصائيSPSS ، انطلاقًا من العلاقة بين الإحصاء والبرمجة والحاسوب.

٢.١ مفردات المقرر:

١. مفاهيم عامة في الإحصاء.

٢. أنواع الإحصاء (إحصاء وصفي – إحصاء استدلالي).

٣. الإحصاء الوصفي:

a. العرض الجدولي.

b. العرض البياني.

c. المقاييس الإحصائية(النزعة المركزية والتشتت).

٤. مفهوم التوزيع الاحتمالي(**Probability Distribution**)

- التوزيعات الاحتمالية المنقطعة(**Discrete**) (برنوللي–الثنائي–بواسون–الهندسي)
- التوزيعات الاحتمالية المستمرة(**Continuous**) (الأسي–الطبيعي–ستودنت–كاي مربع – فيشر)

٥. الإحصاء الاستدلالي:

- التقدير بنقطة.
- مجالات الثقة.
- اختبار الفرضيات.
- الارتباط والانحدار.

٦. تطبيق عملي على الحزمة الإحصائية(**SPSS**)

مفاهيم عامة في الإحصاء

علم الإحصاء: الإحصاء هو مجموعة من الطرق التي تساهم في تقديم ملخّص عن ظاهرة مدروسة.

ويعبّر عن مجموعة من الطرق العلمية التي تُستخدم لـ:

١. جمع ووصف البيانات.

٢. ترتيب وتصنيف البيانات.

٣. عرض وتمثيل البيانات.

٤. تحليل البيانات باستخدام الأدوات الإحصائية.

قبل دراسة الأساليب الإحصائية، من الضروري التعرف على أنواع البيانات ووسائل جمعها المستخدمة لتحليلها.

أنواع البيانات(Data Measures)

يوجد أربعة أنواع رئيسية لمستويات القياس، مرتبة تصاعديًا من الأبسط إلى الأكثر دقة: الاسمي – (Nominal) الرتبي – (Ordinal) الفتري – (Interval) النسبي(Ratio) .

المستوى	الخصائص	أمثلة
Nominal	الأرقام مجرد رموز لا تمثل كميات. لا يمكن إجراء عمليات حسابية عليها.	المهنة – الجنس – الحالة الاجتماعية
Ordinal	الأرقام مرتبة تصاعدياً أو تنازلياً، لكن لا تعبر عن فروق متساوية.	المستوى التعليمي – درجة الإصابة بمرض – أداء محرك
Interval	قيم متصلة بفروق متساوية، لكن الصفر اعتباري. لا يمكن إجراء نسب.	درجة الحرارة – درجات الطلاب
Ratio	قيم متصلة بفروق متساوية وصفر مطلق، يمكن إجراء جميع العمليات الحسابية.	الطول – الوزن – الدخل – الفولت

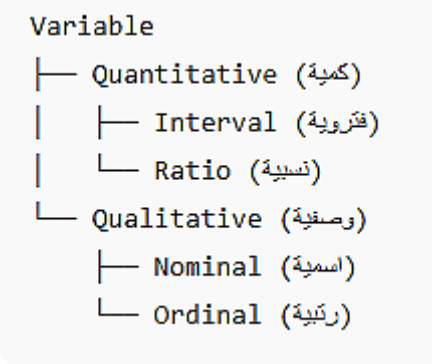
أنواع المتغيرات

يمكن تمييز نوعين رئيسيين من المتغيرات:

١. متغيرات مستمرة :تأخذ قيماً مستمرة ضمن مجال معين (مثل الطول، السمك، الزمن).

٢. متغيرات منقطعة :تأخذ قيماً صحيحة فقط (مثل عدد الطلاب الناجحين).

وتلعب أنواع البيانات دورًا في تحديد نوع المتغير وفق المخطط الآتي:



مصادر جمع البيانات(Collecting Data Sources)

تنقسم إلى نوعين:

١. **مصادر تاريخية** :مثل السجلات والبيانات المنشورة من الشركات أو الجهات الرسمية.

٢. **مصادر ميدانية** :وهي التي يجمعها الباحث مباشرة، وتشمل:

- **المقابلة الشخصية(Personal Interview)**

حيث يلتقي الباحث بالأفراد ويسجل إجاباتهم، لكنها مكلفة وتستهلك وقتًا طويلًا.

- **الاستبيان(Questionnaire)**

باستخدام استمارة ورقية أو إلكترونية عبر الإنترنت أو الهاتف.

- **التجربة(Experiment)**

مثل اختبار تأثير نوع السماد على المحصول الزراعي.

أساليب جمع البيانات(Collecting Data Methods)

١. **الحصر الشامل(Complete Census)**

يتم جمع البيانات من كل أفراد المجتمع، كما في التعداد السكاني.

يعطي صورة دقيقة لكنه مكلف وطويل المدة.

٢. العينات(Sampling)

تُستخدم عند دراسة مجتمعات كبيرة جدًا، وتُختار **عينة** تمثل خصائص المجتمع بدقة.

أنواع العينات:

- **العينة العشوائية البسيطة(Simple Random Sample)**

  سحب مباشر من المجتمع المتجانس.

- **العينة الطبقية(Stratified Sample)**

  تقسيم المجتمع إلى طبقات وسحب عينة من كل طبقة (مثل اختصاصات الجامعة).

- **العينة العنقودية(Cluster Sample)**

  تستخدم للمجتمعات غير المتجانسة عبر تقسيمها إلى طبقات فرعية وسحب عينات متداخلة.

- **العينة القصدية(Purposive Sample)**

  يختار الباحث عناصر محددة تحقق شروطًا معينة تخدم هدف الدراسة.

الإحصاء الوصفي:

هو فرع من فروع الإحصاء يركز على تلخيص وتقديم البيانات بطريقة مفهومة وبشكل يسهل اتخاذ القرارات.

تبويب، عرض وتمثيل البيانات (Data Classification, Display, and Representation)

١.الجداول التكرارية(Frequency Tables)

يمكن تصنيف البيانات – بالإضافة إلى أنواع المتغيرات التي رأيناها سابقًا – إلى بيانات وصفية (Qualitative) وبيانات كمية (Quantitative) وتختلف طريقة عرض كل نوع منها كما يلي:

١.١الجداول التكرارية والتجميعية للبيانات الوصفية

يتضمن الجدول التكراري للبيانات الوصفية عمودين رئيسيين:

١. العمود الأول: يحتوي على **القيم أو الصفات.**

٢. العمود الثاني: يحتوي على **التكرار المقابل لكل صفة.**

وقد يتضمن الجدول أيضًا أعمدة أخرى مثل **النسبة التكرارية** أو **التكرار التجميعي الصاعد.**

ويُطلق على الجدول التكراري اسم **التوزيع التكراري(Frequency Distribution)** لأنّه يوضح توزيع التكرارات على الصفات.

مثال(1)

جدول يبين مستويات ٦٠ طالبًا في أحد المقررات:

المستوى	التكرار (Frequency)	النسبة (%)	النسبة التجميعية
ممتاز	14	23.3%	23.3%
جيد جدًا	14	23.3%	46.7%
جيد	12	20.0%	66.7%
مقبول	12	20.0%	86.7%
ضعيف	8	13.3%	100.0%
المجموع	**60**	**100%**	—

٢.١الجداول التكرارية والتجميعية للبيانات الكمية(Quantitative Data)

تختلف طريقة إنشاء الجداول التكرارية للبيانات الكمية عنها في الوصفية.

فلو اعتبرنا أن كل قيمة فئة مستقلة، سينتج لدينا عدد كبير جدًا من الفئات، لذا يتم تجميع القيم في فئات متساوية الطول.

خطوات إنشاء الجداول التكرارية:

١. **ترتيب البيانات تصاعديًا.**

٢. **تحديد عدد الفئات(k)** ، ويمكن استخدام قاعدة **Sturges Rule**:

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

حيث nهو عدد القيم.

٣. **حساب المدى:(Range)**

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

٤. **حساب طول الفئة:(w)**

$$w = \frac{R}{k}$$

٥. **تحديد حدود الفئات(Class Boundaries)** بدءًا من أصغر قيمة.

مثال:(2)

لنفترض أن لدينا **30طالبًا** وعلاماتهم كما يلي:

46, 49, 48, 58, 54, 51, 41, 62, 37, 48, 54, 75, 54, 48, 59, 45, 34, 58, 47, 61, 49, 44, 68, 39, 63, 56, 43, 57, 41, 45.

المطلوب: إنشاء جدول توزيع تكراري في **7فئات**.

نرتب البيانات تصاعدياً:

٦١,٦٢,٦٣,٦٨,٧٥ ,٣٤,٣٧,٣٩,٤١,٤١,٤٣,٤٤,٤٥,٤٥,٤٦,٤٧,٤٨,٤٨,٤٨,٤٩,٤٩,٥١,٥٤,٥٤,٥٤,٥٦,٥٧,٥٨,٥٨,٥٩

نحسب المدى:

$$R = 75 - 34 = 41$$

طول الفئة:

$$w = \frac{41}{7} = 5.7 \approx 6$$

نحدد الفئات كالتالي:

رقم الفئة	حدود الفئة (Class Interval)
1	]40 – 34]
2	]46 – 40]
3	]52 – 46]
4	]58 – 52]
5	]64 – 58]
6	]70 – 64]
7	[75 – 70]

الجدول التكراري الناتج

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

توضيح حساب حدود الفئات:

نقوم بإنشاء الفئة الأولى

$$[L_1 = X_{min}, U_1 = L_1 + w[= [34{,}34 + 6[= [34{,}40[$$

الفئة الثانية :

$$[L_2 = U_1, U_2 = L_2 + w[= [40{,}40 + 6[= [40{,}46[$$

الفئة الثالثة:

$$[L_3 = U_2, U_3 = L_3 + w[= [46{,}46 + 6[= [46{,}52[$$

الفئة الرابعة:

$$[L_4 = U_3, U_4 = L_4 + w[= [52{,}52 + 6[= [52{,}58[$$

الفئة الخامسة:

$$[L_5 = U_4, U_5 = L_5 + w[= [58{,}58 + 6[= [58{,}64[$$

الفئة السادسة:

$$[L_6 = U_5, U_6 = L_6 + w[= [64{,}64 + 6[= [64{,}70[$$

الفئة السابعة:

$$[L_7 = U_6, U_7 = X_{max}[= [70{,}75]$$

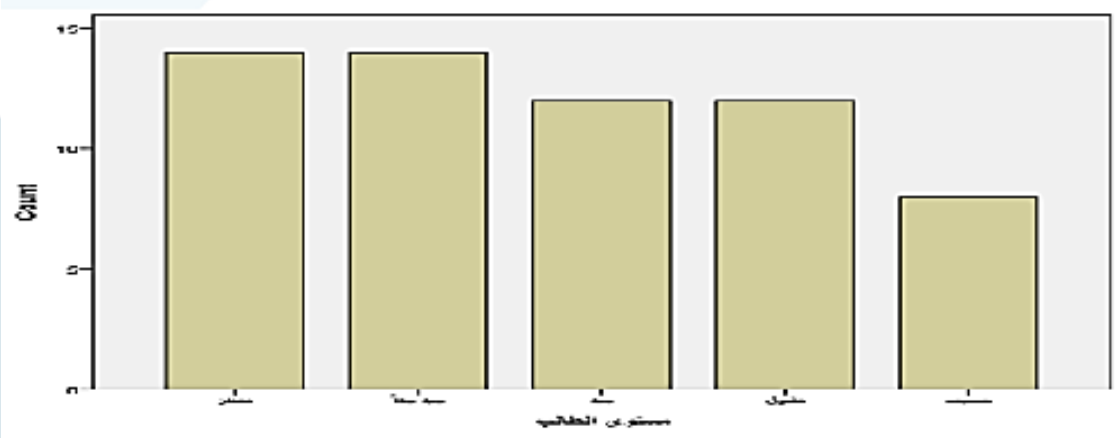
العرض البياني للجداول التكرارية(Graphic Representation)

١ .المدرج التكراري(Histogram)

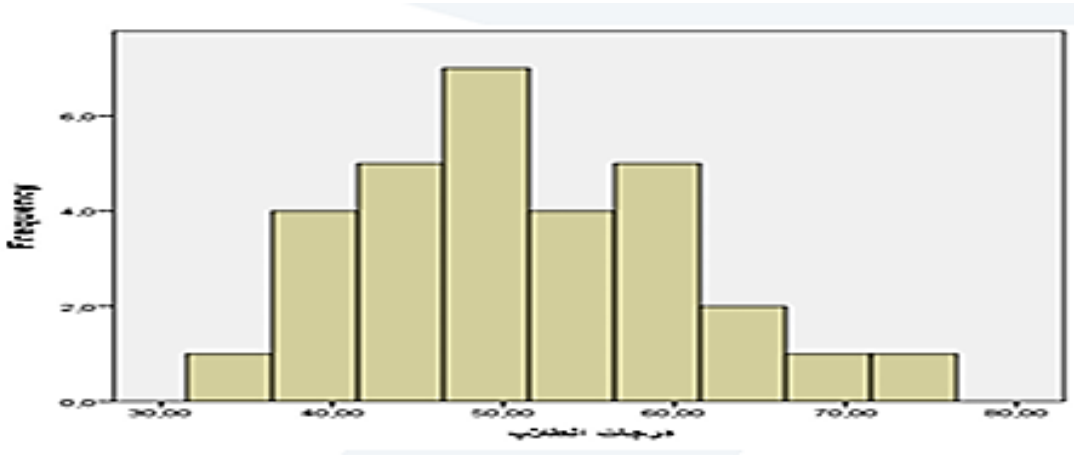
المدرج التكراري يتكون من مجموعة أعمدة متجاورة، تمثل كل فئة بعمود، ويكون ارتفاع العمود مساويًا لتكرار الفئة.يُستخدم عندما تكون البيانات كمية أو وصفية فئوية.

بالعودة للمثالين السابقين:

المثال(١):



المثال(٢):

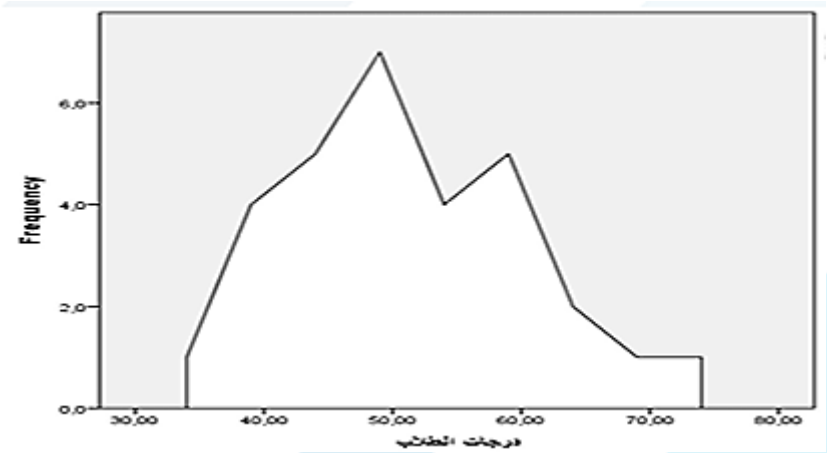


٢.المضلع التكراري(Frequency Polygon)

هو خط منكسر يصل بين نقاط تمثل مراكز الفئات على المحور الأفقي، والتكرارات على المحور العمودي.

يبدأ الخط من الصفر وينتهي بالصفر أيضًا.

بالعودة للمثال(٢):



٣.مخطط الدائرة(Pie Chart)

يُستخدم لإظهار النسب المئوية لكل فئة بطريقة مرئية سريعة.تُجزّأ الدائرة (٣٦٠°) إلى قطاعات تتناسب زواياها مع التكرارات.

مساحة القطاع الدائري = $\frac{\text{التكرار}}{\text{التكراراتمجموع}} \times 360$ = **التكرار النسبي** × ٣٦٠ .

مثال:(3)

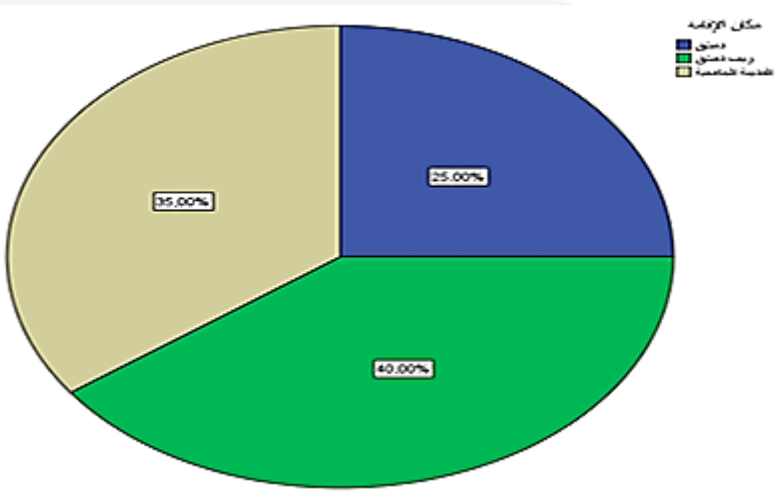
الجدول الآتي يبين مكان إقامة ١٠٠ طالب في إحدى الجامعات:

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

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

مكان الإقامة	التكرار	الزاوية
المدينة الجامعية	35	360*(35/100)=126
دمشق	25	360*(25/100)=90
ريف دمشق	40	360*(40/100)=144
المجموع	100	360

ويكون مخطط الدائرة:



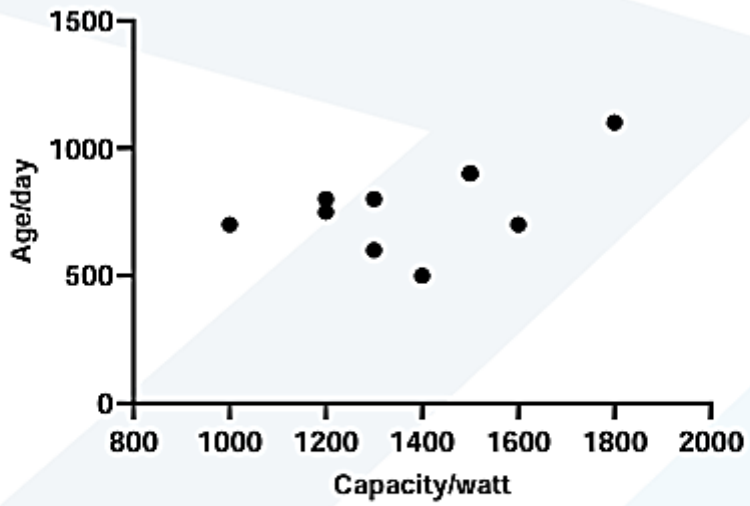
٤ .شكل الانتشار(Scatter Diagram)

يمثل هذا النوع من الرسوم توزيع قيم متغير بالنسبة لمتغير آخر .يُستخدم عادة مع البيانات الكمية للكشف عن وجود علاقة بين متغيرين.

مثال(٤) لدينا بيانات عن الاستطاعة الكهربائية (Capacity/Watt) والعمر بالأيام (Age/Day) لـ ١٠ أجهزة:

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

يكون شكل الانتشار كما يأتي:



يمكن الإستفادة من شكل الانتشار للكشف فيما إذا كان انتشار نقاط احد المتغيرا ت بالنسبة لمتغير آخر يتم بشكل عشوائي أو وفقاً لنموذج ذو معالم، بمعنى آخر يسمح شكل الانتشار بتحديد طبيعة واتجاه العاقة بين المتغيرين المدروسين.

٥.مخطط الساق والورقة(Stem and Leaf Diagram)

وسيلة لعرض البيانات الكمية بطريقة تُظهر شكل التوزيع وعدد التكرارات بسرعة.

تُقسّم القيم إلى قسمين:

- الساق :(Stem) يمثل الأرقام الأولى.
- الورقة :(Leaf) تمثل الأرقام الأخيرة.

مثال(٥) البيانات التالية تمثل قوة الضغط لـ ٨٠سبيكة من الألمنيوم:(lb/ps)

105	221	183	186	121	181	180	143
97	154	153	174	120	168	167	141
245	228	174	199	181	158	176	110
163	131	154	115	160	208	158	133
207	180	190	193	194	133	156	123
134	178	76	167	184	135	229	146
218	157	101	171	165	172	158	169
199	151	142	163	145	171	148	158
160	175	149	87	160	237	150	135
196	201	200	176	150	170	118	149

بالضبع من الصعوبة استخراج خصائص قوة التحمل لسبائك الألمنيوم من هذا الشكل الخام للبيانات، وبالتالي من الضروري عرض هذه البيانات بشكل مفيد أكثر وذلك باستخدام مخطط الساق والورقة:

Stem	Leaf	Frequency
7	6	1
8	7	1
9	7	1
10	5 1	2
11	5 8 0	3
12	1 0 3	3
13	4 1 3 5 3 5	6
14	2 9 5 8 3 1 6 9	8
15	4 7 1 3 4 0 8 8 6 8 0 8	12
16	3 0 7 3 0 5 0 8 7 9	10
17	8 5 4 4 1 6 2 1 0 6	10
18	0 3 6 1 4 1 0	7
19	9 6 0 9 3 4	6
20	7 1 0 8	4
21	8	1
22	1 8 9	3
23	7	1
24	5	1

بمقارنة أرقام الجدول مع أرقام المخطط يلاحظ مثلاً أن الرقمين 105، و 101 ، تم تمثياهما بوضع 10 على الساق و وبضع ورقتين تمثلان العدد1 والعدد 5 .

من المخطط السابق يمكن استخلاص خاصية أنّ غالبية سبائك الألمنيوم المدورسة تتحمل ضغط يتراوح 100 و 200 باوند في البوصة المربعة. بينما القيم المركزية تتراوح بين 142 و169.

العمود الأخير في هذه المخطط يعطي معلومات عن تكرار كل قيمة مثلاً يوجد قيمة واحد للرقم 76 بينما يوجد 3 قيم للرقم 12 وبمعنى آخر يوجد 3 أوراق للعدد 12. كذلك يوجد 8 أوراق Leaf على الساق رقم 14.

ملخص أنواع العرض البياني

نوع البيانات	طرق العرض المناسبة
وصفية (Qualitative)	المدرج التكراري، المضلع التكراري، مخطط الدائرة
كمية (Quantitative)	المدرج التكراري (بفئات)، المضلع التكراري، شكل الانتشار، مخطط الساق والورقة

<< انتهت المحاضرة>>