

measures of spread

→ How Far Are Points From one another

بعض نقطة تختلف اعشار العدد الراوحة المرتبطة بالماحة بين النقاط ويعقدها المجموع بين اعشار المجموعة

① Range

الفارق

② interquartile range IQR

الفارق الريحي النصف

③ Standard Deviation

اعشار المعياري

④ Variance

التفاوتس

من العددهم اعشار البيانات، هنالك اثنان من المجموعات
المرجعية اعشار البيانات المكتبة باسم المجموعات

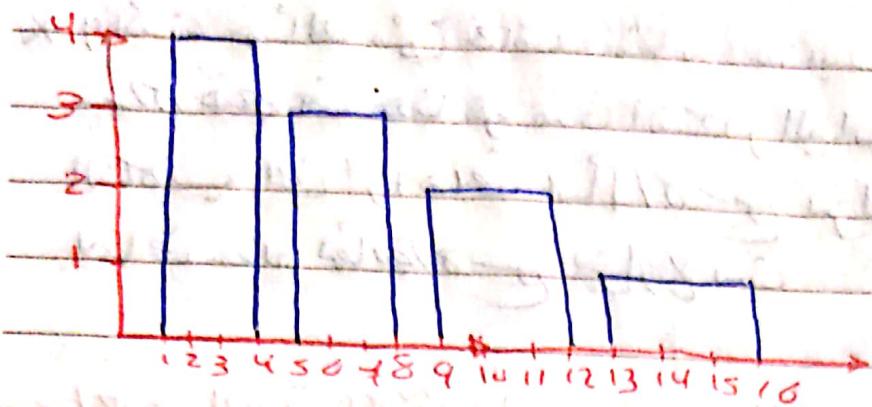
* Histogram : the most common visual for quantitative data

* تطبيقات المجموعات من البيانات

ex: 1, 2, 2, 4 5, 7, 8 9, 12, 15
 ④ 1-4 5-8 9-12 13-16

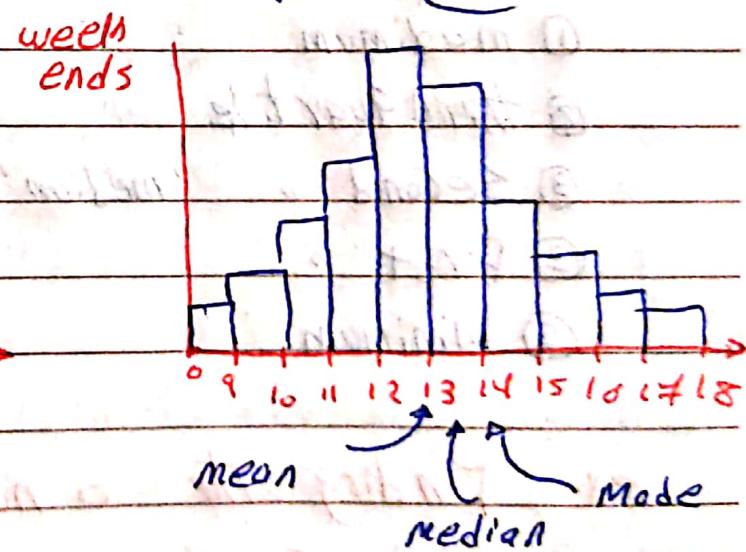
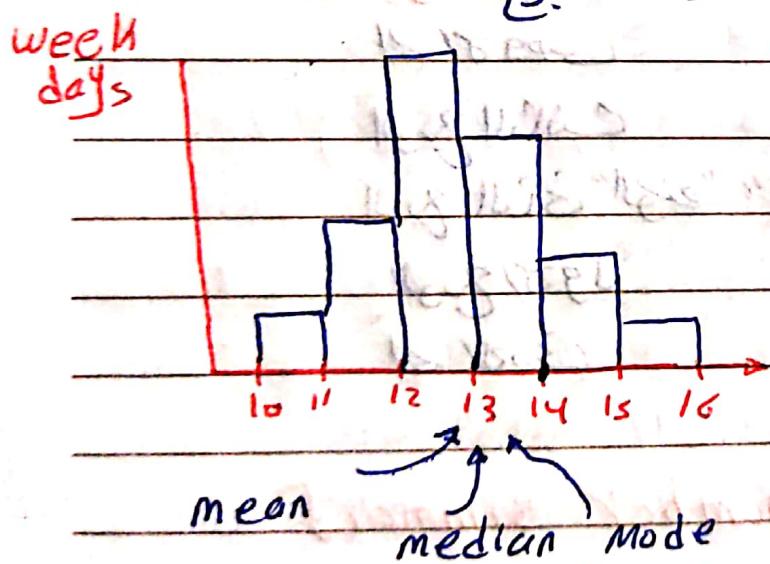
① تم تجزيئ البيانات، بحيث ياباعث كل منش المجموعات المعياري اعشار
التفاوتس على اجزاء اخرست، مما يمثل فئات احادي levels كل فئة يحتوى
على عدد ارقام

② كل فئه لها وظيفة ارجاع كل فرط في المجموعات المعياري



$1 \rightarrow 4$ 4element
 $5 \rightarrow 6$ 3 ,
 $9 \rightarrow 12$ 2 ,
 $13 \rightarrow 16$ 1 "

* لدينا هنا مرحباً بـ تذاكر بـ المقارنة بين عدد الألتراب إلى مـ الـ تـ حـ اـ فـ
أيام / ٢٠ سـ يـ وـ عـ دـ رـ هـ مـ فـ عـ طـ حـ ةـ دـ خـ اـ رـ يـ ١٤ـ سـ يـ عـ جـ مـ الـ عـ اـ مـ



$$\text{mean} = \text{median} = \text{mode} = 13$$

◎

* وكانت هناك سُرّ تخلف ميشاً عن هذلية التوراتيَّة
ـ وله ذريعة هنا هو مني انتشار اليهودية في كل مكان وله ذريعة
الذئب إلى إراها في أيام ١٨ يوم يزاوج بيت ١٥ - ١٦ للثانية
ليهودي عائلة لها أيام ١٨ يوم يزاوج بيت ١٨ - ٥ للثانية

لحدى المفاهيم المهمة في العلوم الاجتماعية هي انتشار بياضاتنا

* most common ways to measure the spread

* Five number summary

→ Gives values for calculating the range and interquartile range

و نعطيها فيما يلي ملخص المذاق والمعدى الرابع

① madimum

امانه

② Third Quartile

الربيع الثالث

③ Second "median" الوسائط الربع الثاني

الربيع الأول

(5) minimum

۱۳۰۸/۲

ex: Finding the 5 number summary

~~25% of 60~~ 5, 8, 3, 2, 1, 3, 10

our data = φ_1 \rightarrow its roof = φ_3 المكتب

1, 2, 3, 3, 5, 8, 10

mark

median

• m, n

"second quartile"

Subject: _____

Date: _____

* median " Q_2 " : A value such that 2/4, or 50% of our data fall below the middle value

* Q_1, Q_3 : the medians of the data on either

side of Q_2

نقدر نصف الوسطيات بـ Q_1, Q_2, Q_3

25% \downarrow
النسبة
المنسبة
 $\frac{1}{4}$ \rightarrow Q_1
نحو ثلثها

min

Q_1

Q_3

Q_2

ed: 1, 2, 3, 3, 5, 8, 10, max

min Q_1 Q_2 "median" Q_3 max
1 2 3 8 10

ex: 5, 8, 3, 2, 1, 3, 10, 10, 5

$\frac{2+3}{2} = 2.5$ 1, 2, 3, 3, 5, 8, 10, 10, 5

min Q_1 Q_2 Q_3 max
1 2.5 $\frac{5+3}{2} = 4$ $\frac{8+10}{2} = 9$ 10, 5

وـ range في المجموع 5 أعداد

Range = maximum - minimum

$$\text{ed}_1 = 10 - 1 = 9 \quad \text{ed}_2 = 10, 5 - 1 = 10$$

interquartile range \rightarrow

$$IQR = Q_3 - Q_1$$

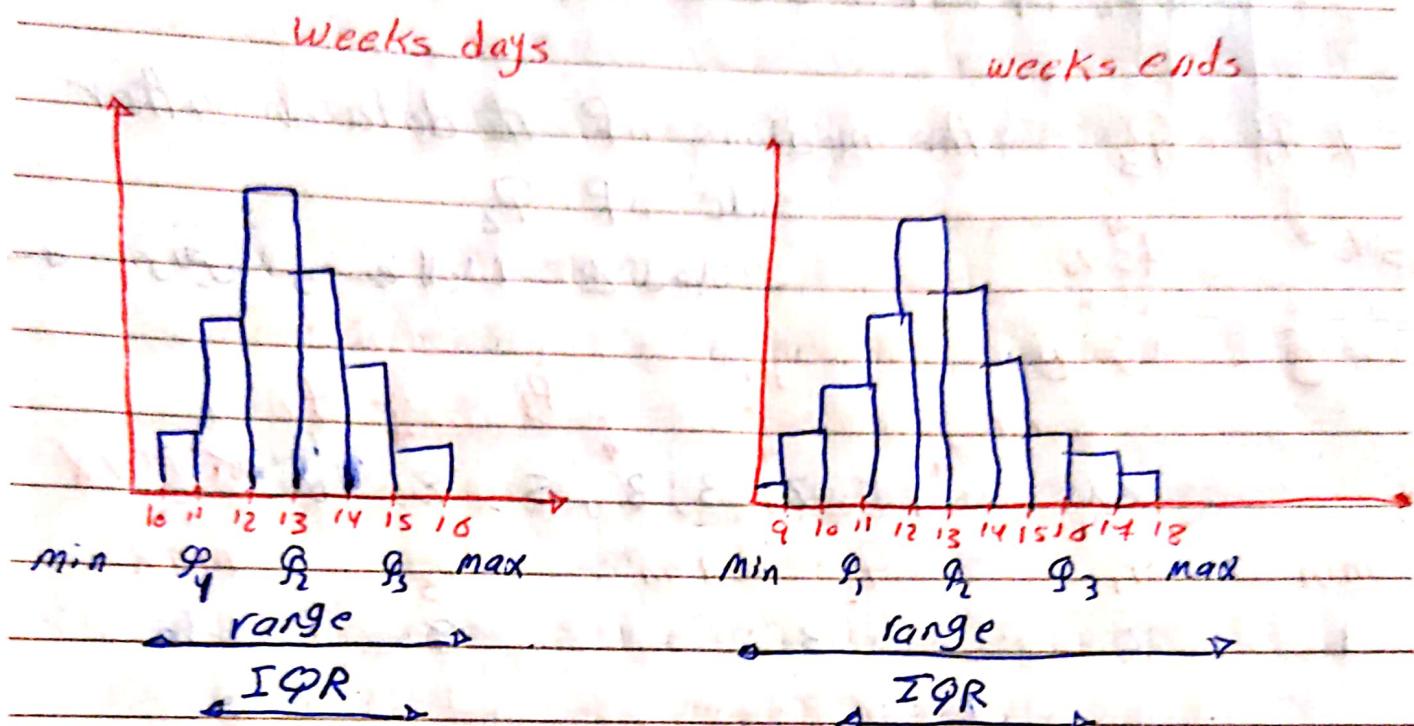
$$\text{ed}_1 = 8 - 2 = 6$$

$$\text{ed}_2 = 9 - 2, 5 = 6, 5$$

Subject:-

Date:

بالنطاق الموزعات التي وجدت لها معنٰى ايجاباً



* Box Plot : useful for quickly comparing the spread of two data set

مقياس مساحة المقارنة بين استمار ملحوظ من البيانات

ياللهم اسألك أسماء الحسنات وارجوك ملائكة السبع

Subject : _____

Date: _____

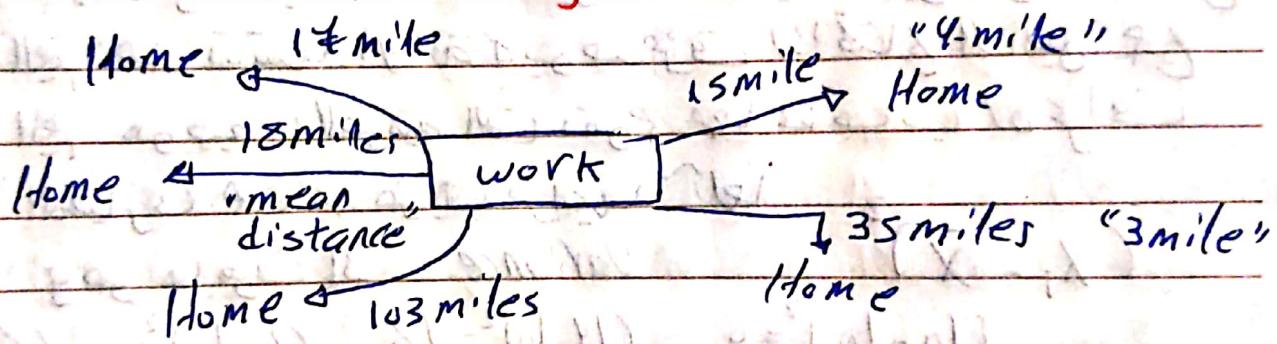
المفهوم الآخر هو مفهوم المعرفة المكتسبة من الممارسة المعاصرة
الممارسة، ذات فتحة واحدة تحيط بغيرها، أو غرفة الممارسة أو
ال ENCOUNTERS، هنا سرير ينبع من غرف الممارسة، ولكن مختلف
المعنى الذي يعطي الممارسة هي الممارسة

* Standard Deviation

→ on average, how much each point varies from the mean of the point

ـ هو يزيد كل نقطة عن متوسط المقادير "mean"

ed: How far are employee located from their work?



* How does the "distance to work" vary from employee to employee?

مثلاً في المثلث $\triangle ABC$ إذا كان $a = 5$, $b = 7$, $c = 8$ ، فإن مجموع المثلث هو $5 + 7 + 8 = 20$ ، ونسبة المثلث هي $\frac{20}{3} \approx 6.67$.

Subject: _____

Date: _____

ex1 Data set 10, 14, 10, 6

$$\bar{x} = \frac{1}{4} \sum_{i=1}^4 x_i = \frac{10+14+10+6}{4} = 10$$

حساب المعدل ①

$$x_i - \bar{x}$$

الخطوة ② حساب مجموع الفariance

$$10-10=0, 14-10=4, 10-10=0, 6-10=-4$$

$$\bar{x} = \frac{0+4+0-4}{4} = 0$$

الخطوة ③ حساب مجموع الفariance

الخطوة 4: حساب مجموع الفariance حيث يزيد الفariance كلما زاد الفariance، حيث تقصى اواقة الوجود مرفق لـ انتشار الاعداد جميعاً، فيما هو عادي / الطريقة التي تقوم بها الممارسة باردة عدداً

$(x_i - \bar{x})^2$ Variance وهو ما يعرف بالانحراف المعياري

$$(10-10)^2=0, (14-10)^2=16, (10-10)^2=0, (6-10)^2=16$$

$$\text{Variance} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$= \frac{1}{4} \times (0 + 16 + 0 + 16) = \frac{32}{4} = 8$$

$$* s = \sqrt{8} = \sqrt{2,83}$$

* **Variance:** Average square difference of each observation from the mean

Variance الـ الربيعي \Rightarrow standard deviation

$$\text{standard deviation} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\text{variance}}$$

↓ square root of the variances

* **Standard deviation** الـ ميادن

① used to compare spread of different groups

• يوضح المقدار الذي يختلف فيه مجموعات مختلفة في استدراجهن
في المعايير المترادفة

② High standard deviation is associated with higher risk.

• حيث تختلف البيانات - تكون واردة في ارتفاع اعلى اوجهها
الجباري يدخل ارتفاع معدل المفترض (متوسط) في المقارنة بين احتمال ورقة
وال KING، وباختصار احتمال العالية الذي يتمتع بارتفاع اعلى اوجه المعايير
بعض الوقت يغير المفترض مطورة من حركة ورقة احتمال العالية المفترض
و باختصار احتمال اعلى اوجه المعايير

③ Fair Comparisons require the same unit

• لكي يتمكن اعملاً مترادفون من جعل البيانات ذات نفس الوحدات

④ Variance has square unit of their original dataset

• غالباً ما يدرس وحدات مترادفة مختلفة فنحو ليس لها نفس الوحدات لذلك

* **standard deviation:** and the data is measured

☞ ~~عمر جميع اصحاب الربيعي~~

measures of shape

الإيجازات

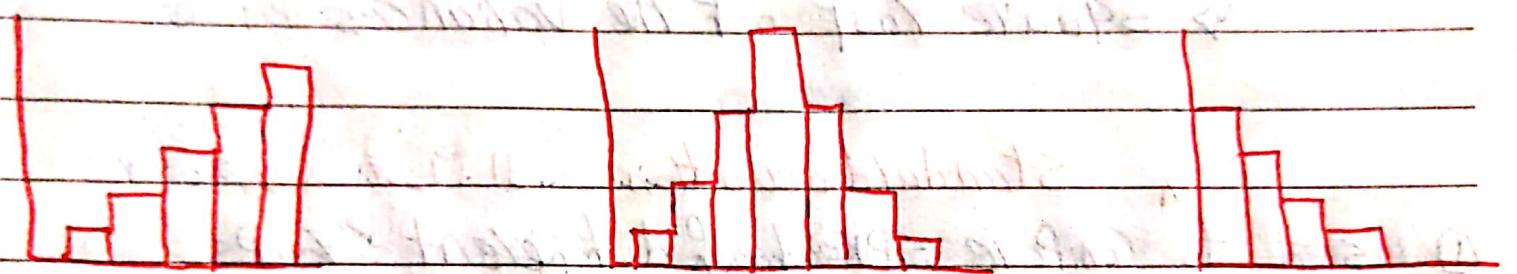
البيانات في توزيعها تختلف في شكلها، وذلك يعتمد على طبيعة البيانات.

* Shape: How to use Histograms to determine shape
Associated with data

left skewed

symmetric

right skewed



* التوزيع المُنحني الذي ينبع من حجم المفرد الذي يقع في النهاية اليسرى أو اليمين ويعكس طبيعة البيانات أو ينمّي إلى الباردة وهو طول المدى غير ممكناً إلى العصى العلوي والمنسق العلوي

* يُعرف أكثر التوزيعات المنتشرة بـ "التوزيع العادي" ويشتهر بـ "حاجز العجز" "normal distribution" "the bell curve"

يعتمد لشكل التوزيع بجزء الكثرة ما يليق بقيمة المترن والمترن

الدوريجات متساوية طبق المعدل
"normal distribution"

mean = median = mode

متساوية طبق المعدل في التوزيع

النوع الثاني المعدل هو المعدل mode

النوع الثالث المعدل هو المعدل mode

right skewed ③

mode

معدل سائقي تطويق دليل التوزيع mean

mode < median < mean

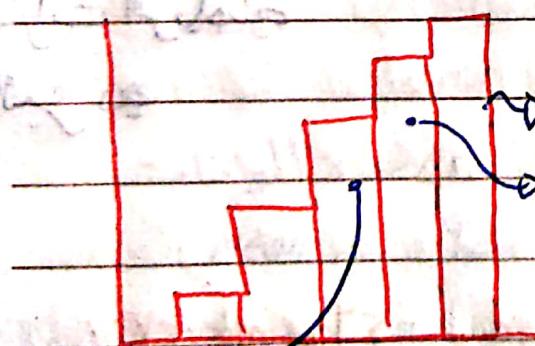
median

mean

mean > median

\rightarrow

left skewed ④



mean

معدل سائقي تطويق mean

median > mean

\rightarrow



Subject:

Date:

* Quick Plot: when you are working with data
you can build a quick plot to see the shape

ed-of PP bell-shape "normal"

student Heights , weights , scores , precipitation
أوزان طلاب ، أوزان ، درجات امتحان ، كميات الأمطار
Distribution \bar{x} , manufacturing
نمط التوزيع (نطاء صناعة المنتج)

ed-of left-skewed

GPA , Age of death , price

ed-of right-skewed

blood , distribution , Athletic Abilities
نوع الدم ، توزيع المروحة ، القدرة الرياضية
of wealth

المرات الرياضية
نوع المروحة
نوع الدم
البراعة
معيار الربح المتفق في
الوقت

* Distributions

right skewed , left-skewed , symmetric

هي الرسم من اهتماماته التي يهتم بها في العالم
المتفق في توزيعه ونوعه ونوعه
+ توزيعه في توزيعه

Aspects of Quantitative Variables

Center
spread

shape
outliers

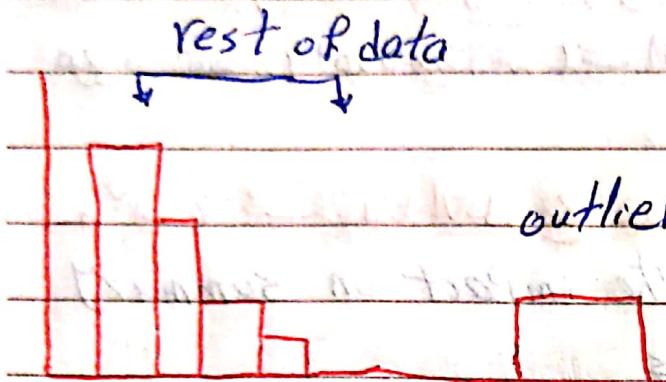
ALAQSA

outliers

میرا بیٹا

* outliers: data points that fall very far from the rest of the values in our data set

هي تفاصيل البيانات تقع ليس احدي عتائق الهم في ملحوظة البيانات
في الواقع السائدة هي هناك عدة طرق مختلفة



د) ملحوظات المقدمة

outliers أي من النقاط البيانات

۱۴۷

* يساعد على إنشاء مخطط التوقيع للبيانات - ٤) فهم المترافق قرارة ومتغيره
أو متوجه المترافق الذي يمكن أن تؤدي إليه القيم المترافق بحسب طريقة
(متوجه مومزة عامل روابط رواد ادراة عمال

* Annual Earnings in thousands "\$"

45, 68, 92, 53, 105, 56, 24, 15, 155

and \$16 billion

ولقد ذكرنا في مقدمة دراسة تأثير العوامل المدارية على نتائج القياسات

mean = 160 million / year

(هـ) ممثل يذكر سبب ورثة كيوجي رجل أعمال يدعى كيوجي الواتـ

* zero entrepreneurs

earned this or close it

* median = 62,000

٥) الفعل مؤثر في آخر يابع الحقن المدحون انه يكتبهما رائدا اعجمان ملائكة ببيانات

١٠٥٣٦ ملايين اصحاب اعمال في العالم

٤٨٢ مليون ~ standard deviation

most entrepreneurs mark

٩٥٪ من اصحاب الاعمال ارباحاً مرتفعاً جداً لكن هناك ليس الجميع اصحاب ارباحاً مرتفعاً بل هناك جماعة اصحاب ارباحاً متوسطة واجهة ايجابية

؟ كل ذلك يكتب علينا التحليل مع هذه القيم المفترضة

① note they edit and the impact on summary statistics

١٠٥٣٦ ملايين اصحاب اعمال في العالم

standard deviation mean

if typo, remove or fix

(ادخلت عبارة عن اصحاب اعمال ايجابية او اصحاب اعمال سلبية)

سبعين في المائة المقصود اصحاب اعمال ايجابية

ناتج الصيغة

② understanding why they edit, And the impact on questions we are trying to answer

١٠٥٣٦ ملايين اصحاب اعمال في العالم

reporting the 5 number summary is better than

mean and standard deviation when outliers

are present

١٠٥٣٦ ملايين اصحاب اعمال في العالم

٩٥٪ من اصحاب اعمال ايجابية



④ be careful in reporting know how to ask
the right questions

ـ يجتاز اليمين الرقّ الواقع في آخر المقالة كولما يجتاز بالفعل في بياضنا
ـ تكون نسبات القيمة أكثر دلالة من غيرها إذا كانت مسحات المعلمات
ـ تكون النسبة المئوية والذى تشير لها فيها من المهم معرفة قيمة كل
ـ حصة القيمة المئوية بالذى يجتاز كولما يجتاز

ماداً) أنت المنهى يا ساد المستشارين؟ هل يحق الضرادات عنة

- ① Plot your data
 - ② make a decision on our outliers

- ۳) اکست نیوں کا یہ سیاہ ت مورنگہ نیک عرض ہے۔

الصيغة حول البيانات، ما هي؟ أون المخارق والمدو ط وقط

* bell-shaped data

We can learn a lot with the mean and standard deviation in normally distributed

* Shewed data

Five number summary is best for skewed data

وأدلة البيانات، منها ورقة لفوجي ملخص أسلوب القياس (روي) ملحوظات

أكبر مجموعات البيانات = صناعة الترجمة التي تقدمها أجهزة المحتوى والذكاء الاصطناعي

ما هي الأدوات التي تستخدم في تحليل البيانات
Visual بحرياً

* Descriptive statistics

لبيانات تم جمعها لفهمها (بيانات)
البيانات التي تم جمعها

* Descriptive statistics

describing Collected data

بيانات التي تم جمعها

* Inferential statistics

Drawing conclusions about A Population
Based on data collected from a sample of
individuals from that Population

بيانات التي تم جمعها حول مجموعة ملائمة من
البيانات التي تم جمعها من تلك المجموعة المثلية

ex: دراسة نسبت الشرب بين الجنسين
نسبة شرب الماء بين الجنسين

Total population : 100,000 students

sample : 5000

statistics : 73%

Parameter: Proportion of all 100,000
students that drink water

Subject:

Date:

* descriptive statistics → بيانات كمية
ex: 73% of 5000 response say
they drink coffee

* inferential statistics → بيانات احصائية
"what does this mean about all university
student?"

Population → المجموعة الكلية

Sample → مجموعة ممثلة من المجموعة الكلية

Statistics → بيانات تم الحصول عليها من المجموعة

Parameter → متغير ينتمي للمجموعة الكلية