



لجان الرّفعات

STATISTICS



MORPHINE ACADEMY

MORPHINE
ACADEMY

Pharmaceutical Statistics

Lecture 6 Part 1

Descriptive statistics

Measures of Dispersion

Prepared and Presented by

Dr. Muna Oqal

Measures of Dispersion for Grouped Frequency Table

مقاييس التشتت / spread / scatter

المدى

A. Range:

بدرج

The range of grouped frequency table can be calculated as follow:

الحد الأعلى آخر فترة بالجدول

الحد الأدنى في الأول فترة بالجدول

$$R = (\text{Upper limit of last class} - \text{Lower limit for the first Class}) + 1$$

Measures of Dispersion for Grouped Frequency Table

• Example

The following frequency table represents the time in minutes of 50 pharmacists to their work selected from the records of a given pharmaceutical industries company in Jordan:

Time to travel to work	Frequency
الفتره الأولى 1-10	8
11-20	14
21-30	12
31-40	9
الفتره الأخيرة 41-50	7

Range

$$= (\text{الحد الأعلى} - \text{الحد الأدنى}) + 1$$

$$= (50 - 1) + 1$$

$$= 50$$

A. Calculate the range for the time travelled to the work for pharmacists?

Solution:

$$R = (\text{Upper limit of last class} - \text{Lower limit for the first Class}) + 1$$

$$R = (50 - 1) + 1 = 50 \text{ minutes}$$

Measures of Dispersion for Grouped Frequency Table

B. Calculate the value of the sample standard deviation (S) for the time travelled to the work for the pharmacists?

B. Variance and Standard deviation

1. Find the midpoint for each class interval
2. For each class interval multiply the frequency with each midpoint (fx). $\Rightarrow$ ضرب كل midpoint بالتردد
3. Find the sum $\sum (fx)$. $\Rightarrow$ جمع كل $(midpoint * f)$
4. Find the square value for $\sum (fx)$ in step 3. $\Rightarrow$ باي جعنا $(\sum (midpoint * f))^2$ نربعه
5. Then divide the sum in step 4 by the sum of frequencies ($n = \sum f_i$) as follow $\frac{(\sum fx)^2}{n}$ $\Rightarrow$ $\frac{(\sum (midpoint * f))^2}{n}$ $\Rightarrow$ مجموع التكرارات f $\Rightarrow$ كل midpoint $\times$ باي جعنا f برعنا
6. Find the square value of the midpoint for each interval (x^2) $\Rightarrow$ كل midpoint $\times$ باي جعنا f برعنا
7. For each class multiply the frequency (f) with each squared midpoint ($X^2 = fx^2$) $\Rightarrow$ ضرب $(midpoint)^2$ بالترددات f
8. Find the sum $\sum fx^2 \rightarrow \sum (midpoint)^2 * f$ $\Rightarrow$ جمع باي ضربناه $(midpoint)^2 * f$
9. Calculate the sample variance (S^2) as follow:

نطبق على القانون:

$$S^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1}$$

Sample Variance $\leftarrow S^2$

$\sum fx^2$ $\leftarrow$ جمع التكرارات f $\times$ باي جعنا $(midpoint)^2$

$\frac{(\sum fx)^2}{n}$ $\leftarrow$ جمع التكرارات f $\times$ باي جعنا $(midpoint)^2$

Measures of Dispersion for Grouped Frequency Table

- Solution:**

Class	Midpoint (x)	f	x^2	<u>fx</u>	fx^2
1 to 10	5.5	8	30.25	44	242
11 to 20	15.5	14	240.25	217	3363.5
21 to 30	25.5	12	650.25	306	7803
31 to 40	35.5	9	1260.25	319.5	11342.25
41 to 50	45.5	7	2070.25	318.5	14491.75
Total		50		1205	37242.5

$\boxed{fx}$

- To get the sample mean as follow $= \bar{X} = \frac{\sum fx}{n} = \frac{1205}{50} = 24.1$
- The sum of fx is 1205
- The square value of $\sum (\text{fx})$ is 1452025
- The sum of fx^2 is 37242.5.

- To get sample variance as follow $S^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1}$

- $S^2 = \frac{37242.5 - \frac{1452025}{50}}{50-1} = 167.38$, then **S.D** $= \sqrt{167.38} = 12.93$

Ex

Find the Variance and Standard Deviation.

Class	F
1 to 10	8
11 to 20	14
21 to 30	12
31 to 40	9
41 to 50	7
n = 50	

1) نقطہ midPoint کا فرق

(midPoint)

اگر اداً + اگر اداً
2

Class	F	midPoint
1 to 10	8	$\frac{10+1}{2} = 5.5$
11 to 20	14	$\frac{20+11}{2} = 15.5$
21 to 30	12	$\frac{30+21}{2} = 25.5$
31 to 40	9	$\frac{40+31}{2} = 35.5$
41 to 50	7	$\frac{50+41}{2} = 45.5$
Total n = 50		

2) مربع midPoint کا مجموعہ

Class	F	midPoint	(midPoint)
1 to 10	8	$\frac{10+1}{2} = 5.5$	$(5.5)^2 = 30.25$
11 to 20	14	$\frac{20+11}{2} = 15.5$	$(15.5)^2 = 240.25$
21 to 30	12	$\frac{30+21}{2} = 25.5$	$(25.5)^2 = 650.25$
31 to 40	9	$\frac{40+31}{2} = 35.5$	$(35.5)^2 = 1260.25$
41 to 50	7	$\frac{50+41}{2} = 45.5$	$(45.5)^2 = 2070.25$
n = 50			

3 جزو کے midpoint اور $\sum f$ کا فرق

Class	f	midPoint	(midPoint) ²	midpoint * f
1 to 10	8	$\frac{10+1}{2} = 5.5$	$(5.5)^2$	$5.5(8) = 44$
11 to 20	14	$\frac{20+11}{2} = 15.5$	$(15.5)^2$	$15.5(14) = 217$
21 to 30	12	$\frac{30+21}{2} = 25.5$	$(25.5)^2$	$(25.5)(12) = 306$
31 to 40	9	$\frac{40+31}{2} = 35.5$	$(35.5)^2$	$(35.5)(9) = 319.5$
41 to 50	7	$\frac{50+41}{2} = 45.5$	$(45.5)^2$	$(45.5)(7) = 318.5$
n=50				total = 1205

4 جزو کے (midpoint)² اور $\sum f$

Class	f	midPoint	(midPoint) ²	midpoint * f	(midPoint) ² * f
1 to 10	8	$\frac{10+1}{2} = 5.5$	$(5.5)^2 = 30.25$	$5.5(8) = 44$	$(30.25)(8) = 242$
11 to 20	14	$\frac{20+11}{2} = 15.5$	$(15.5)^2 = 240.25$	$15.5(14) = 217$	3363.5
21 to 30	12	$\frac{30+21}{2} = 25.5$	$(25.5)^2 = 650.25$	$(25.5)(12) = 306$	7803
31 to 40	9	$\frac{40+31}{2} = 35.5$	$(35.5)^2 = 1260.25$	$(35.5)(9) = 319.5$	11342.75
41 to 50	7	$\frac{50+41}{2} = 45.5$	$(45.5)^2 = 2070.25$	$(45.5)(7) = 318.5$	14491.75
n=50				total = 1205	total = 37242.5

5] طبق على الفلوت :

$$S^2 = \frac{\sum (\text{midpoint})^2 \times f - \frac{(\sum \text{midpoint} \times f)^2}{n}}{n - 1}$$

$$= \frac{37242.5 - \frac{(1205)^2}{50}}{50 - 1}$$

$$S^2 = 167.38$$

لو طبق ال Standard deviation

$$S.D = \sqrt{S^2}$$

$$= \sqrt{167.38}$$

$$= 12.93$$

Pharmaceutical Statistics

Lecture 6 Part 2

Descriptive statistics

Measures of Position

Prepared and Presented by

Dr. Muna Oqal

الموقع وقياسات Measures of Position

برنا نرد موقع
قيمة معينة من مجموعة
بيانات

Definitions

- Measures of position are used to describe the relative location of an observation. ^{لوصف}

طريقة لتقسيم البيانات إلى أجزاء متساوية
لعدد قريباً تصادفياً

- Quantiles system: are the positions of values in the data set. Data set are divided to ^{موقع} quartiles, ^{القسم} percentiles, percent (25% each), etc., after ordering them from smallest to the largest. ^{من طرق تقسيم الـ DATA}

شرط أساسي

Measures of Position

- The median and quartiles are specific examples of quantiles.

نظام تقسيم البيانات لتسهيل دراستها

المعلومات تنقسم

- Quantile systems that cut data into more than four ranges are really only useful where there are quite large numbers of observations. Such as quintiles, deciles and centiles (percentiles).

مفيد

لأن

لأنه
يكون
عند
عدد
كبير
من
الملاحظات

والتالي بقسمها إلى
أجزاء 4 أجزاء

- There are four quintiles, which divide data into five ranges, nine deciles for ten ranges and 99 centiles that produce 100 ranges.

2
قسم
نصفه
خلال
السلالات

- The ninth decile is thus equivalent to the 90th centile and both indicate a point that ranks 10% from the top of a set of values.

Measures of Position



Quantile systems

⇒ * نظام الهدف منه تقسيم البيانات إلى أجزاء متساوية بعد ترتيبها من الأصغر إلى الأكبر.

Quantile systems divide ranked data sets into groups with equal numbers of observations in each group. Specifically:

- 3 *Quartiles* divide data into four equal-sized groups.
- 4 *Quintiles* divide it five ways.
- 9 *Deciles* divide it ten ways.
- 99 *Centiles* divide it 100 ways.

Measures of Position

A. Percentiles

النسب

B. Quartiles

الأرباع

C. Five Number Summary

ملخص الأرقام الخمسة

D. Boxplot

!! لقد الآن متى طلعك منك إلا انك تعرفت شو الهون منك
Quantile system
والأشياء الثانيه حلا بتوضع ان شاء الله.

Measures of Position

- **Percentiles** or (centiles): are measures which divide the data set into 100 equal parts ($P_1, P_2, \dots, P_{100}$)
- Percentiles: are measures indicating the value below a given percentage of observations in group of observation falls.
- For example, the 40th percentile (P_{40}) is the value (or score) below which 40% of the observations may be found.

Ex

$$P_{40} = 15 \Rightarrow \begin{cases} \rightarrow 40\% \text{ جابوا تى ال 15} \\ \rightarrow 60\% \text{ جابوا فوق ال 15} \end{cases}$$

كيف نحس الـ Percentile

① نرتب الـ DATA تصاعدياً .

② نستخدم القانون

$$L = \frac{K \cdot n}{100}$$

القيمة
رقم النسبة المئوية
موقع القيمة داخل البيانات
Ex: P_{30}
 $K=30$

③ نزيد 0.5 عشوائياً

نطلع القيمة بدقة

Example

$n=24$

① ترتيبه جازم ② القانون $L = \frac{K}{100} \cdot n$
 $n=24$
 $K=75$
 $L = \frac{75}{100} \cdot (24) = 18$
 $18 + 0.5 = 18.5$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
58, 59, 68, 69, 74, 75, 78, 78, 78, 79, 79, 81, 83, 84, 85,
16 17 18 18.5 19 20 21 22 23 24
85, 86, 86, 88, 88, 89, 90, 90, 92

Solution:

Use

$$L = \frac{K}{100} \cdot n$$

Then $L = 75 / 100 \cdot 24 = 18$ (position) + 0.5 = 18.5
which corresponds to $(86+88)/2 = 87$

Measures of Position

Finding the Score Given a Percentile

$$L = \frac{K}{100} \cdot n$$

1] ترتیب البيانات من الأصغر إلى الأكبر

2] نستخدم القانون
↓
P50
↓
K=50
عدد القيم
↓
L = $\frac{K \cdot n}{100}$
موقع القيمة داخل البيانات

n total number of values in the data set

k percentile being used

L locator that gives the *position* of a value

P_k k th percentile

Then, you add 0.5 to find the exact position

3] نزيد 0.5 للقيمة
بلي طلعن طبع ليس؟؟
معدل على موقع دقيق

Measures of Position

P₇₅

Example: we want to know the 75th percentile of the ordered grades from smallest to largest,

n=24

k=75

① ترتيبه جازه ② التاذف

نزيد 0.5

18 + 0.5 = 18.5

$$L = \frac{k}{100} \cdot n = \frac{75}{100} \cdot 24 = 18$$

58, 59, 68, 69, 74, 75, 78, 78, 78, 79, 79, 81, 83, 84, 85, 85, 86, 86, 88, 88, 89, 90, 90, 92

Solution:

$$\frac{86+88}{2} = 87$$

Use

$$L = \frac{K}{100} \cdot n$$

Then $L = 75 / 100 * 24 = 18$ (position) + 0.5 = 18.5
which corresponds to $(86+88)/2 = 87$

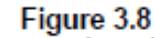
*هناك خلعنا اول طريقة لتقسيم البيانات وهي النسبة المئوية التي تقسم البيانات إلى ١٠ أجزاء متساوية كل جزء ١٠٪

③

الخيار ٤

الموقع

ربا عیاف



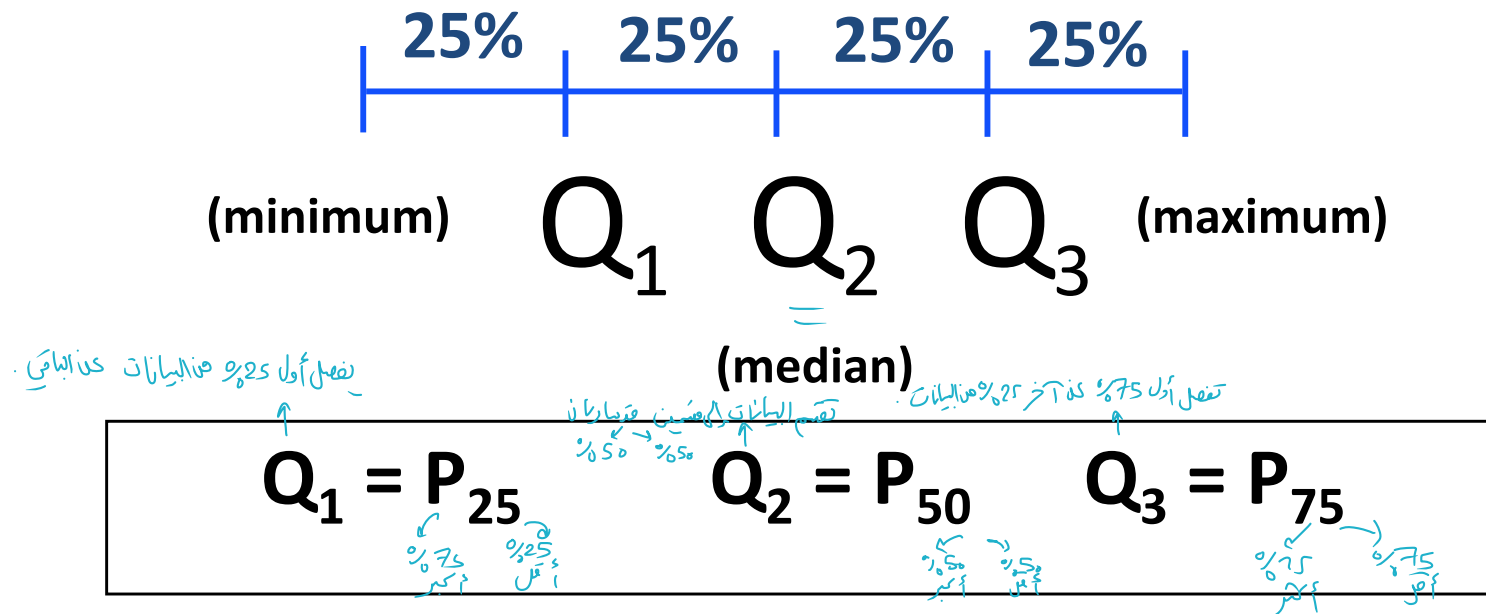
The middle half of the observations in a frequency distribution lie within the interquartile range

Quartiles $\Rightarrow$ الأرباع

Q_1, Q_2, Q_3

تقسم البيانات إلى أربع أجزاء متساوية

divides ranked scores into four equal parts



Q=quartile, P= Percentile

Measures of Position

• Definition

(a) First (lower) Quartile (Q1) = Median = P_{25}

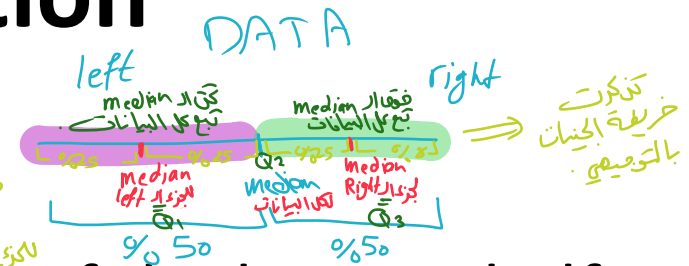
The first (lower) quartile (Q1) is the median of the bottom half of the ordered observation for a data set (to the left of the median), or it is the median of the data set (lies at or below the median (MD)).

(b) Second Quartile (Q2) = Median = P_{50}

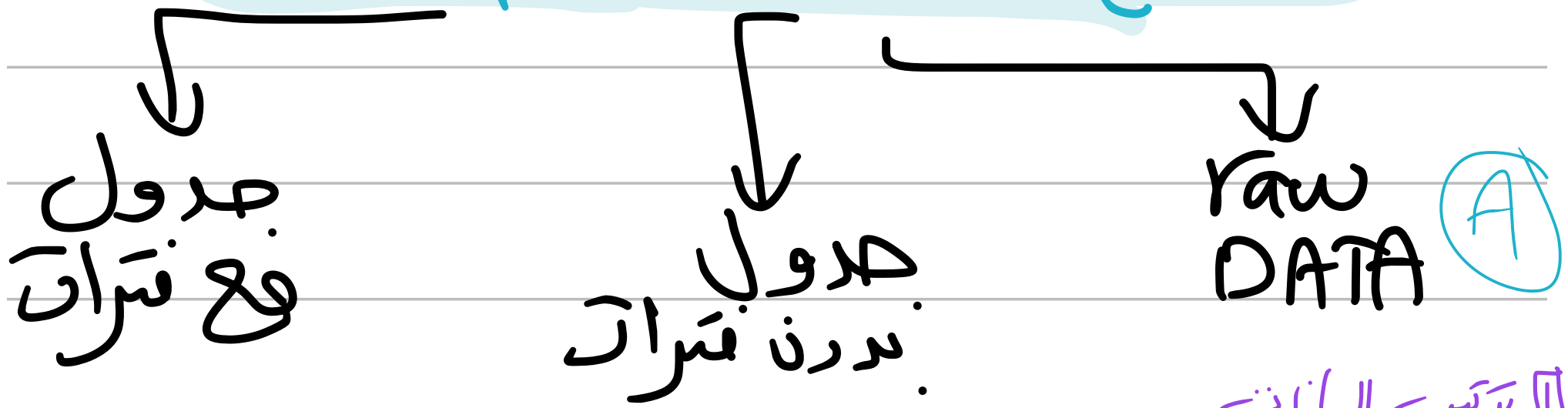
The second quartile (Q2) is the median of the data set, that is, $Q2 = MD$. It separates the lowest 50% of the data from highest 50%.

(c) Third (Upper) Quartile (Q3) = Median = P_{75}

The third (upper) quartile (Q3) is the median of the top half of the ordered observations for a data set (to the right of the median), or it is the median of the data set (lies at or above the median (MD)).



کل ماخص ال Quartiles من حسابات



1 ترتيب البيانات
(الأصغر ← الأكبر)

2 نفس ال Median

تبع كل البيانات يلي هو Q_2

كيف أحسب ال median

* ترتيب (أنا أصغر مرتبة)
* n (عدد القيم) ← زوج (even) زوجي
* هيل بنكون الـ median موقع الـ median
بعد بنشوف قيمته.

3

جزء يسار	جزء يمين
Median تبع الجزء اليسار Q_1	Median تبع كل البيانات Q_2

Median
تبع الجزء
اليمين
 Q_3

Example
Find the Quartiles

5, 10, 10, 10, 12, 15, 20, 20, 25, 30, 30, 40, 40, 60

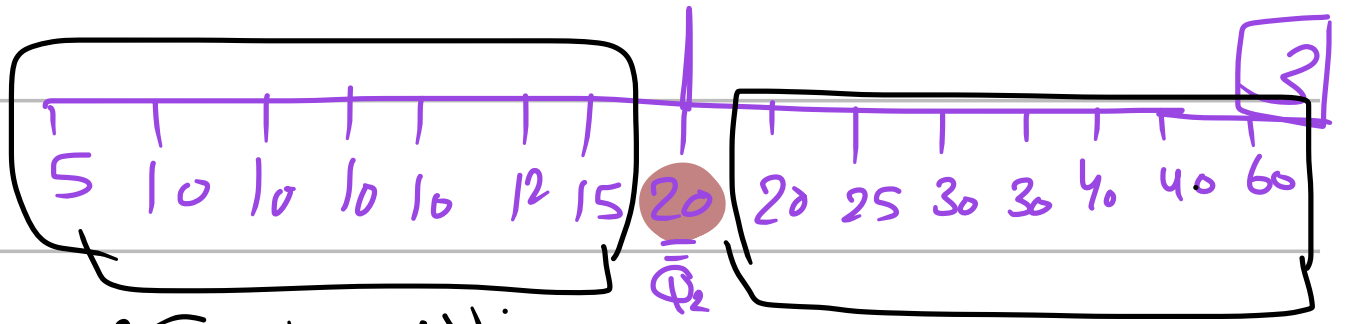
1 مرتبة ✓

2 نجد ال Median تبع كل البيانات

موقع ال median $n=15$ (odd)
 $\frac{15+1}{2} = 8$
 ← موقعه 8

5, 10, 10, 10, 12, 15, 20, 20, 25, 30, 30, 40, 40, 60
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

$Q_2 = P_{50} = \text{Median} = 20$
 تبع كل البيانات



بدنا الـ median تبع
هاي البيانات
وينطلع Q_1

$Q_2 = P_{50}$
= median
تبع البيانات

هنا بدنا الـ median
هاي البيانات
وينطلع Q_3

$n = 7 \leftarrow$ median
تبع الجزء
الأكبر
 $\frac{7+1}{2} = 4$ له
لعمق 4
 Q_1

$n = 7 \leftarrow$ median
تبع الجزء
الأصغر
 $\frac{7+1}{2} = 4$ له
لعمق 4
 Q_3

5, 10, 10, 10, 10, 12, 15
1 2 3 4 5 6 7

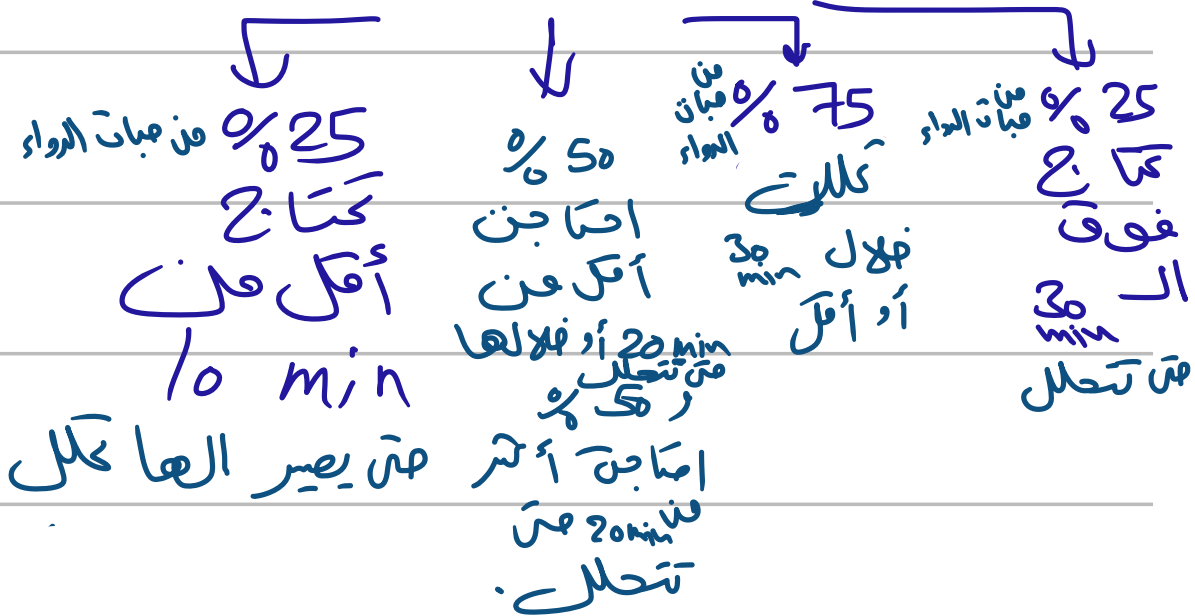
20, 25, 30, 30, 40, 40, 60
1 2 3 4 5 6 7

$P_{25} = Q_1 = \text{median} = 10$

$P_{75} = Q_3 = \text{median} = 30$

بالسؤال أخطانا
حتى يصير لها
15 tablets
وبدنا نشوف كم بها
مبة دواء

15 tablets



B

Find the quartile Q_1 , Q_2 and Q_3 for number of accidents per week?

number of Accidents (x)	2	7	12	17	22
Number of weeks (f)	5	19	13	8	5

1) اُلجوزو ڏيو ۽ (c.f) cumulative frequency

$$C.P = C.P_{\text{المف}} + P_{\text{المالي}}$$

Number of a.c(x)	Frequency (f)	C.f
2	5	$C.f + f = 0 + 5 = 5$
7	19	$C.f + f = 5 + 19 = 24$
12	13	37
17	8	45
22	5	50
		$n = 50$

2

$$\frac{n}{4}$$

لج بعين
موقع

$$\frac{n}{y} = \frac{50}{4} = 12.5$$

$$\frac{n}{2}$$

2) عرض

موقع اول $\bar{\text{median}}_{\Phi_2}$

$$\frac{n}{2} = \frac{50}{2} = 25$$

$$\frac{3n}{4}$$

2) $\frac{1}{3}$

$$\frac{3^4}{4} = \frac{3(50)}{4} = 37.5$$

3) لنذكر على $\frac{n}{2}$ في جدول C.f صادية له أو باعثة

$$\frac{n}{4} = 12.5$$

نذكر على 12.5 أو 1 برفضا باعثة في خانة C.f تكون
رشفة X تكون هي 11.

Number of acc(x)	Frequency (f)	C.f
2	5	C.f + f = 0 + 5 = 5
7	19	C.f + f = 5 + 19 = 24
12	13	37
17	8	45
22	5	50
n = 50		

$Q_1 = 7$ ←

→ أكبر من 12.5 باعثة

4) نذكر على $\frac{n}{2}$ في جدول C.f صادية له أو باعثة
باعثة وشفة X تكون الـ Q_2

$$\frac{n}{2} = 25$$

Number of acc(x)	Frequency (f)	C.f
2	5	C.f + f = 0 + 5 = 5
7	19	C.f + f = 5 + 19 = 24
12	13	37
17	8	45
22	5	50
n = 50		

$Q_2 = 12$ ←

→ أكبر من 25 باعثة

5) نذكر على $37.5 = \frac{3n}{4}$ في جدول C.f نصفها
أو برفضا باعثة

Number of accidents (x)	Frequency (f)	C.f
2	5	$\begin{matrix} \text{C.f} + f \\ \text{سابقہ} + \text{حالیہ} \\ = 0 + 5 = 5 \end{matrix}$
7	19	$\begin{matrix} \text{C.f} + f \\ \text{سابقہ} + \text{حالیہ} \\ 5 + 19 = 24 \end{matrix}$
12	13	37
17	8	45
22	5	50
$n = 50$		

③ = 17 ←

→ اسی سے 37.5 دے رہی ہیں

• **Conclusion:**

The results mean that:

- ☐ 25% of weeks have less than 7 accidents and 75% of weeks have more than 7 accidents.
- ☐ 50% of weeks have less than 12 accidents and 50% of weeks have more than 12 accidents.
- ☐ 75% of weeks have less than 17 accidents and 25% of weeks have more than 17 accidents.

③ جدول مع تكرارات :

① جدار $C.F$:

② نجد الفترة التي يوجد داخلها Q_1

نستخدم هذا القانون How??

$$Q_1 = L_1 + \left(\frac{\left(\frac{n}{4}\right) - CF_1}{f_1} \right) * i$$

↓
The lower boundary of the Q_1 class interval.
↓
النكسر الخاص بالفترة.

↓
The class interval width.
↓
الفاصل بين الفترات = $(x_j - x_{j-1})$

↓
الفترة CF_1

↓
التردد f_1

↓
الحد الأدنى للفترة = $\frac{n}{4} - \frac{1}{2}$

② نجد الفترة التي يوجد داخلها Q_2 :

$$Q_2 = L_2 + \left(\frac{\left(\frac{n}{2}\right) - CF_2}{f_2} \right) * i$$

④ نجد الفترة التي يوجد داخلها Q_3 :

$$Q_3 = L_3 + \left(\frac{\left(\frac{3n}{4}\right) - CF_3}{f_3} \right) * i$$

Example: - Find the quartiles:

Class (Sales Volume in j)	Number of Days (f)
53-56	3
57-60	5
61-64	9
65-68	7
69-72	6

① جداول C.F

Class (Sales Volume in j)	Number of Days (f)	C.F
53-56	3	0+3=3
57-60	5	3+5=8
61-64	9	8+9=17
65-68	7	17+7=24
69-72	6	24+6=30

$n=30$

← تقع Q_1 في هذه الفترة

← تقع Q_2 في هذه الفترة

← تقع Q_3 في هذه الفترة

② $15 = \frac{30}{2} \leq \frac{n}{2}$ نجد

نبحث عن 15 أو أكبر منها مباشرة عن موقع Q_2

③ $7.5 = \frac{30}{4} \leq \frac{n}{4}$ نجد

نبحث عن 7.5 أو أكبر منها مباشرة عن موقع Q_1

④ $22.5 = \frac{3(30)}{4} \leq \frac{3n}{4}$ نجد

نبحث عن 22.5 أو أكبر منها مباشرة عن موقع Q_3

$Q_3 = L_3 + \left(\frac{\frac{3n}{4} - c.f_3}{f_3} \right) i$ (65-68) ① $L_3 = 65 - \frac{1}{2} = 64.5$ ② $\frac{3n}{4} = 22.5$ ③ $c.f_3 = 17$ $\rightarrow$ بنوفا $c.f$ للسطر الثالث بالجدول ④ $f_3 = 7$ $\rightarrow$ بنوفا f للسطر الحالي ⑤ $i = 4$ $Q_3 = 64.5 + \left(\frac{22.5 - 17}{7} \right) (4)$ $Q_3 = 67.64$	$Q_2 = L_2 + \left(\frac{\frac{n}{2} - c.f_2}{f_2} \right) i$ (61-64) ① $L_2 = 61 - \frac{1}{2} = 60.5$ ② $\frac{n}{2} = \frac{30}{2} = 15$ ③ $c.f_2 = 8$ $\rightarrow$ بنوفا $c.f$ للسطر الثاني بالجدول ④ $f_2 = 9$ $\rightarrow$ بنوفا f للسطر الحالي ⑤ $i = (64 - 61) + 1 = 4$ $Q_2 = 60.5 + \left(\frac{15 - 8}{9} \right) (4)$ $Q_2 = 63.61$	$Q_1 = L_1 + \left(\frac{\frac{n}{4} - c.f_1}{f_1} \right) i$ (57-60) ① $L_1 = 57 - \frac{1}{2} = 56.5$ ② $\frac{n}{4} = \frac{30}{4} = 7.5$ ③ $c.f_1 = 3$ $\rightarrow$ بنوفا $c.f$ للسطر الأول بالجدول ④ $f_1 = 5$ $\rightarrow$ بنوفا f للسطر الحالي ⑤ $i = (60 - 57) + 1 = 4$ $Q_1 = 56.5 + \left(\frac{7.5 - 3}{5} \right) (4)$ $Q_1 = 60.1$
--	---	---

Conclusion

For this pharmacy, the results mean that:

- 25% of days have sales volume less than 61.1 JD and 75% of days have sales volume more than 61.1 JD.
- 50% of days have sales volume less than 63.61 and 50% of days have sales volume more than 63.61 JD.
- 75% of days have sales volume less than 67.64 JD and 25% of days have sales volume more than 67.64 JD.

Measures of Position

موجود بالمختص

* حساب الـ Quartiles لما يكون عندي قيم مفردة:

A. The Quartiles for Raw Data (ungrouped data)

To find the quartiles for a set of raw data, do the following:

1. Arrange the data set from the smallest to highest (ordered array).
ترتيب البيانات من الأصغر للأكبر.
2. Calculate the median ($MD=Q2$) for the data set.
حساب الـ median
3. For the half of the data set to the left of the MD calculate their median to get the first quartile ($Q1$).
خذ النصف الأيسر من البيانات وحسب الـ median
4. For the half of the data set to the right of the MD, calculate their median to get the third quartile ($Q3$).
خذ النصف الأيمن من البيانات وحسب الـ median

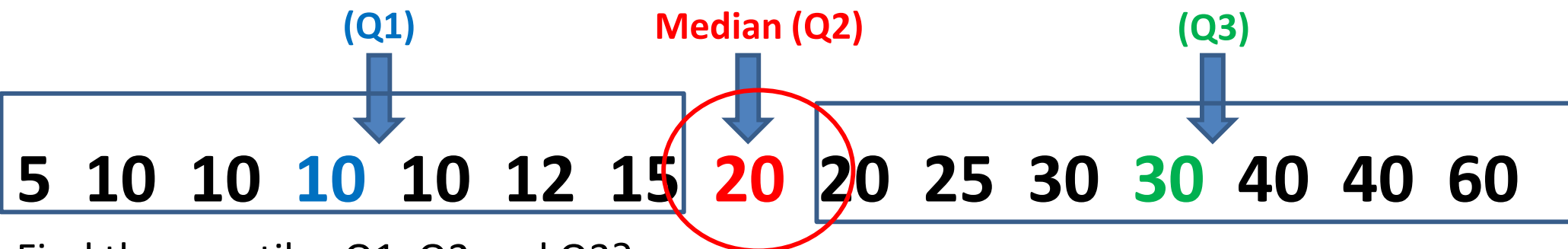
Measures of Position

تم طرحه بالخطأ فوق

← عليه تفكك حبة الدواء إلى أجزاء (مفكر عند ملاصقتها لسائل مثل اللعاب)

• Example (n is odd)

The times in minutes needed for a random of 15 tablets to disintegrate are as follow:



Find the quartiles Q1, Q2 and Q3?

1. Data is already in an order from smallest to highest
2. The **MD** is $X_{((n+1)/2)} = X_{(8)} = \mathbf{20}$ minutes.
3. For the half to the left of the median, $n=8$, then $X_{((n+1)/2)} = X_{(4.5)} = \mathbf{Q1} = (10+10)/2 = \mathbf{10}$ minutes.
4. The second quartiles is $\mathbf{Q2=MD=20}$ minutes.
5. For the half to the right with the MD, $n=8$, then then $X_{((n+1)/2)} = X_{(4.5)} =$ the third quartile which is $\mathbf{Q3} = (30+30)/2 = \mathbf{30}$ minutes.
6. The quartiles are: **Q1=10 minutes**, **Q2=20 minutes**, **Q3=30 minutes**.

When n is odd, you can include the median in both quartiles (**Q₁** and **Q₃**)

Measures of Position

موجود بالملف فوق

Since $Q_1 = P_{25}$, $Q_2 = P_{50}$, $Q_3 = P_{75}$, so we can find Q_1 , Q_2 , and Q_3 as following:

Q1=25%=25/100*15=3.75 + 0.5= 4.25th (**position**),

which corresponds to **10** ((10+10)/2).

Q2=50/100*15=7.5 +0.5= 8th (**position**),

which corresponds to **20**.

Q3=75/100*15=11.25+0.5=11.75th (**position**),

which corresponds to **30** ((30+30)/2).

5	10	10	10	10	12	15	20	20	25	30	30	40	40	60
---	----	----	-----------	----	----	----	-----------	----	----	----	-----------	----	----	----

Measures of Position

وجود بالمقد
فوق

Conclusion

- This means that 25% of the tablets need less than 10 minutes to disintegrate.
- 50% of the tablets need 20 minutes to disintegrate.
- Before 30 minutes 75% of all tables were disintegrated.
- 25% only of these tablets need more than 30 minutes to disintegrate.

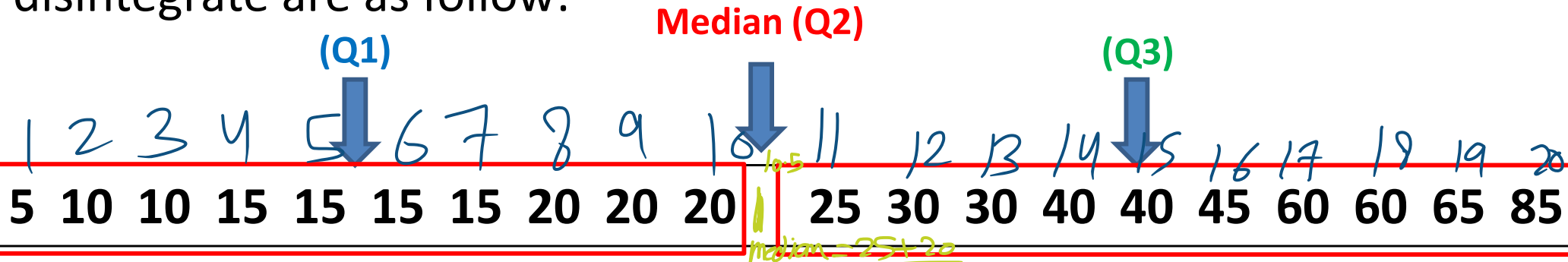
Disintegration time (min)	Frequency
5	1
10	4
12	1
15	1
20	1
25	2
30	2
40	2
60	1
Total	15

Measures of Position

نفسی از Ex یایی حلیناه
بس الفرقا انه هون عدد القيم
زر جي $n=20$ هون $\frac{n+1}{2}$ 2 نون
موقع ال Median 2 يكون ساره من
قيمتين جمعهم وتنقسم الى 2

• Example (n is even)

The time in minutes needed a random sample of 20 capsule to disintegrate are as follow:



Find the quartiles Q1, Q2, and Q3?

$$= 22.5$$

$$n = 20$$

$$\frac{20+1}{2} = \frac{21}{2} = 10.5$$

موقع ال Median

1. Data is already in an order from smallest to highest
2. The **MD** is $X_{((n+1)/2)} = X_{(10.5)} = (20+25)/2 = \mathbf{22.5}$ minutes.
3. For the half to the left of the median, $n=10$, then the first quartiles is $\mathbf{Q1} = (15+15)/2 = \mathbf{15}$ minutes
4. The second quartiles is $Q2 = MD = 22.5$ minutes.
5. For the half to the right with the MD, $n=10$, then the third quartile is $\mathbf{Q3} = (40+45)/2 = \mathbf{42.5}$ minutes.
6. The quartiles are: **Q1=15 minutes**, **Q2=22.5 minutes**, **Q3=42.5 minutes**.

$$Q2 = P_{50} = \text{median} = 22.5$$

القيمة الوسطى

Measures of Position

Since $Q_1 = P_{25}$, $Q_2 = P_{50}$, $Q_3 = P_{75}$, so we can find Q_1 , Q_2 , and Q_3 as following:

Q1 = $25\% = 25/100 * 20 = 5 + 0.5 = 5.5^{\text{th}}$ (**position**),

which corresponds to **$((15+15)/2) = 15$** .

Q2 = $50/100 * 20 = 10 + 0.5 = 10.5^{\text{th}}$ (**position**),

which corresponds to **$((20+25)/2) = 22.5$** .

Q3 = $75/100 * 20 = 15 + 0.5 = 15.5^{\text{th}}$ (**position**),

which corresponds to **$((40+45)/2) = 42.5$** .

$$Q_2 = 22.5$$

$$Q_1 = 15.5$$

$$Q_3 = 42.5$$

(Q1)

Median (Q2)

(Q3)

5 10 10 15 15 15 15 20 20 20 25 30 30 40 40 45 60 60 65 85

Measures of Position

Conclusion

- This means that 25% of the capsules need less than 15 minutes to disintegrate.
- 50% of the capsules need 22.5 minutes to disintegrate.
- Before 42.5 minutes 75% of all capsules were disintegrated.
- 25% only of these capsules need more than 42.5 minutes to disintegrate.

Disintegration Time (minutes)	Frequency
5	1
10	2
15	4
20	3
25	1
30	2
40	2
45	1
60	2
65	1
85	1
Total	20

Measures of Position

B. The quartiles for simple (ungrouped) frequency table

جدول
بغیر
گروہ بندی
فrequencies

The quartiles for simple frequency table can be calculated as follows:

1. Find the cumulative frequency (cf).
2. Find the values of $(n/4)$, $(n/2)$, $(3n/4)$, where n is the total number of observation (the sum of frequencies).
3. The first quartile (Q_1) is the first value having cumulative frequency (cf) greater than or equal to $(n/4)$.
4. The second quartile (Q_2) is the first value having cumulative frequency (cf) greater than or equal to $(n/2)$, that is the median (MD).
5. The third quartile (Q_3), is the first value having cumulative frequency (cf) greater than or equal to $(3n/4)$.

Measures of Position

- **Example**

The Gulf Pharmaceutical Industries Company in UAE planning to improve safety plan in its factory. For this, accident data for the last 50 weeks was compiled. These data were listed into simple frequency table as shown below:

Number of Accidents (x)	2	7	12	17	22
Number of weeks (f) <i>frequency</i>	5	19	13	8	5

Find the quartiles Q1, Q2 and Q3 for number of accidents per week?

Measures of Position

- Solution:**

Step-1: Calculate the value of cumulative frequency (cf) as follows:

	Number of Accidents (x)	Frequency (f)	cf
	2	5	5
Q1	7	19	24
Q2	12	13	37
Q3	17	8	45
	22	5	50
	Total	50	

Measures of Position

- **Step-2:** Find the following values:
 - $n/4 = 50/4 = 12.5$
 - $n/2 = 50/2 = 25$
 - $3n/4 = (3)(50)/4 = 37.5$
- **Step-3:** The first quartile (Q1) is the first value having cumulative frequency (cf) greater than or equal to 12.5, then from the table the value of the first quartile (Q1) is **Q1=7** accident per week.
- **Step-4:** the second quartile (median) (Q2=MD) is the first value having cumulative frequency (cf) greater than or equal to 25, then from the table, the value of the second quartile (Q2) is **Q2=12** accident per week.
- **Step-5:** The third quartile (Q3) is the first value having cumulative frequency (fc) greater than or equal to 37.5, then from the table, the value of the first quartile (Q3) is **Q3=17** accident per weeks.

Measures of Position

- **Conclusion:**

The results mean that:

- ☐ 25% of weeks have less than 7 accidents and 75% of weeks have more than 7 accidents.
- ☐ 50% of weeks have less than 12 accidents and 50% of weeks have more than 12 accidents.
- ☐ 75% of weeks have less than 17 accidents and 25% of weeks have more than 17 accidents.

Measures of Position

C. The quartiles for grouped frequency table

The quartiles for grouped frequency table (frequency distribution) can be calculated as follows:

For the first quartile (Q1)

Step-1: Construct the cumulative frequency distribution.

Step-2: Determine the first quartile class interval (**First Quartile Class**), that is, the first class having cumulative frequency (cf) greater than or equal to $(n/4)$.

Step-3: Find the first quartile (Q1) by using the following formula:

$$Q_1 = L_1 + \left(\frac{\left(\frac{n}{4}\right) - cf_1}{f_1} \right) * i$$

Where:

n = the total number of frequencies.

cf_1 = cumulative frequency prior to the first quartile class interval.

i = the class interval width.

L_1 = the lower boundary (limit) of the first quartile class interval.

f_1 = the frequency of the first quartile class interval.

Measures of Position

C. The quartiles for grouped frequency table

For the second quartile (Median) (Q₂=MD)

Step-1: Construct the cumulative frequency distribution.

Step-2: Determine the second quartile class interval (**Second Quartile Class**), that is, the first class having cumulative frequency (cf) greater than or equal to (n/2).

Step-3: Find the second quartile (Q₂) by using the following formula:

$$Q_2 = L_2 + \left(\frac{\left(\frac{n}{2}\right) - cf_2}{f_2} \right) * i$$

Where:

n = the total number of frequencies.

cf_2 = cumulative frequency prior to the second quartile class interval.

i = the class interval width.

L_2 = the lower boundary (limit) of the second quartile class interval.

f_2 = the frequency of the second quartile class interval.

Measures of Position

C. The quartiles for grouped frequency table

For the third quartile (Q3)

Step-1: Construct the cumulative frequency distribution.

Step-2: Determine the third quartile class interval (**Third Quartile Class**), that is, the first class having cumulative frequency (cf) greater than or equal to $(3n/4)$.

Step-3: Find the third quartile (Q2) by using the following formula:

$$Q_3 = L_3 + \left(\frac{\left(\frac{3n}{4} \right) - cf_3}{f_3} \right) * i$$

Where:

n = the total number of frequencies.

cf_3 = cumulative frequency prior to the third quartile class interval.

i = the class interval width.

L_3 = the lower boundary (limit) of the third quartile class interval.

f_3 = the frequency of the third quartile class interval.

Measures of Position

- **Example:**

The following frequency table represents the daily sales volume in JD for a period of 30 days selected from the sales records of given pharmacy in Jordan:

Class (sales volume in JD)	53 - 56	57 - 60	61 - 64	65 - 68	69 - 72
Number of Days (f)	3	5	9	7	6

Find the quartiles Q_1 , Q_2 , and Q_3 for the sales volume in JD?

Measures of Position

Solution

Step-1: calculate the value of cumulative frequency (cf) as follows:

Class Interval (Daily Sales Voume in JD)	f_i	cf_i	
53 - 56	3	3	
57 - 60	5	8	Q1 Class
61 - 64	9	17	Q2 Class
65 - 68	7	24	Q3 Class
69 - 72	6	30	
Total	30		

Step-2: Calculate the interval width (i)

$i = 57 - 53 = 4$ so $i = 4$

Measures of Position

- **Step-3: calculate the value of the Q1 as follows:**

$n/4 = 30/4 = 7.5$, then (57-60) is the first quartile class, then

$$Q_1 = L_1 + \left(\frac{\left(\frac{n}{4}\right) - cf_1}{f_1} \right) * i = 56.5 + \left(\frac{7.5 - 3}{5} \right) * 4 = 60.1 \text{ JD.}$$

- **Step-4: calculate the value of the Q2 as follows:**

$n/2 = 30/2 = 15$, then (61-64) is the second quartile class (median class), then

$$Q_2 = L_2 + \left(\frac{\left(\frac{n}{2}\right) - cf_2}{f_2} \right) * i = 60.5 + \left(\frac{15 - 8}{9} \right) * 4 = 63.61 \text{ JD.}$$

- **Step-5: calculate the value of the Q3 as follows:**

$3n/4 = (3*30)/4 = 22.5$, then (65-68) is the third quartile class, then

$$Q_3 = L_3 + \left(\frac{\left(\frac{3n}{4}\right) - cf_3}{f_3} \right) * i = 64.5 + \left(\frac{22.5 - 17}{7} \right) * 4 = 67.64 \text{ JD.}$$

Measures of Position

Conclusion

For this pharmacy, the results mean that:

- 25% of days have sales volume less than 61.1 JD and 75% of days have sales volume more than 61.1 JD.
- 50% of days have sales volume less than 63.61 and 50% of days have sales volume more than 63.61 JD.
- 75% of days have sales volume less than 67.64 JD and 25% of days have sales volume more than 67.64 JD.

Range

10-90
Percentile
Range

$$= P_{90} - P_{10}$$

midquartile
Range

$$= \frac{Q_3 + Q_1}{2}$$

Semi-interquartile
Range

$$= \frac{Q_3 - Q_1}{2}$$
$$= \frac{IQR}{2}$$

Interquartile
Range
(IQR)

$$= Q_3 - Q_1$$

$$Q_1 = P_{25}$$

$$Q_2 = P_{50}$$

$$Q_3 = P_{75}$$

Measures of Position

Quartiles

تقسیم آردل 25%
منه DATA نه البای.
↑
Q₁ = P₂₅ ← آردل 25%
↑
Q₂ = P₅₀ ← آردل 50%
↑
Q₃ = P₇₅ ← آردل 75%
↑
P = Percentile

Interquartile **Range** (or IQR): $Q_3 - Q_1$

Semi-interquartile **Range**: $\frac{Q_3 - Q_1}{2}$

Midquartile **Range**: $\frac{Q_1 + Q_3}{2}$

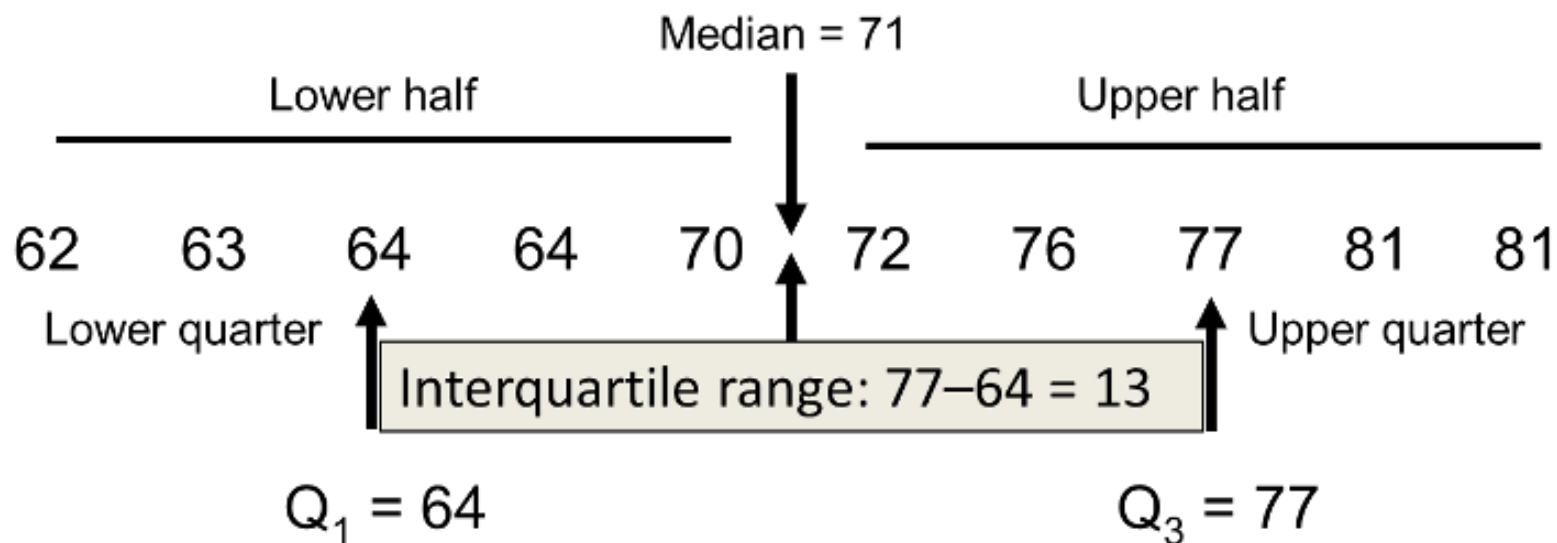
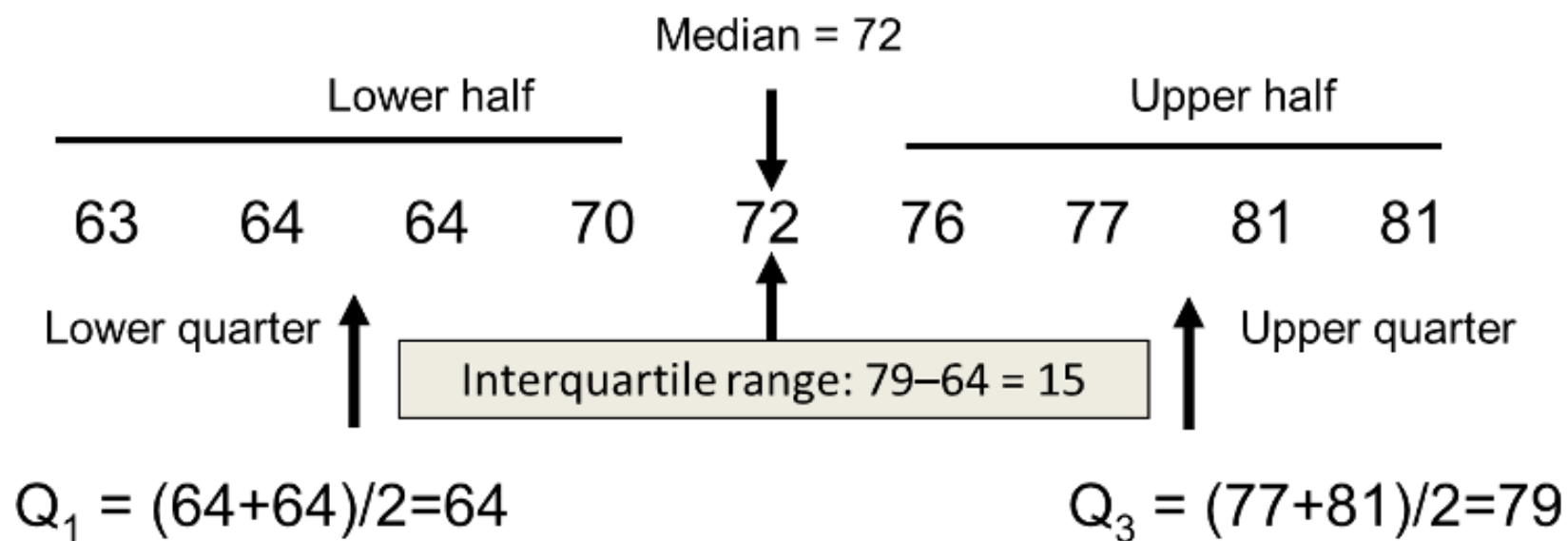
10 - 90 Percentile **Range**: $P_{90} - P_{10}$

Measures of Position

The Interquartile Range (IQR)

- The interquartile range (IQR) is a robust measure of variation that is based on the quartiles. عَوِي
- The IQR is defined as the range of middle 50% of observations in the data set. مقياس مدى انتشار البيانات الوسطى
- It is the difference between the third quartile (Q3) and the first quartile (Q1), and it is found by using following formula: الفارق

$$IQR = Q3 - Q1$$



Ex 1

63, 64, 64, 70, 72, 76, 77, 81, 81

موقع
البيانات
Median
البيانات
$$= \frac{n+1}{2} = \frac{5+1}{2} = 2.5$$

Median
$$= \frac{64+64}{2} = 64$$

$$Q_1 = 64$$

Median
$$= Q_2$$

$$Q_2 = 72$$

موقع
البيانات
Median
البيانات
$$= \frac{n+1}{2} = \frac{4+1}{2} = 2.5$$

$$Q_3 = \text{Median} = \frac{77+81}{2} = 79$$

$$\begin{aligned} IQR &= Q_3 - Q_1 \\ &= 79 - 64 \\ &= 15 \end{aligned}$$

Ex (2)

5-5
62, 63, 64, 64, 70, 72, 76, 77, 81, 81
1 2 3 4 5 6 7 8 9 10

$$\text{موقع median} = Q_2 = \frac{10+1}{2} = 5.5$$

$$\text{median} = \frac{70+72}{2}$$

$$Q_2 = \frac{142}{2} = 71$$

Q₁

$$\text{موقع median} = \frac{5+1}{2} = 3$$

الجزء اليسار

$$Q_1 = 64$$

Q₃

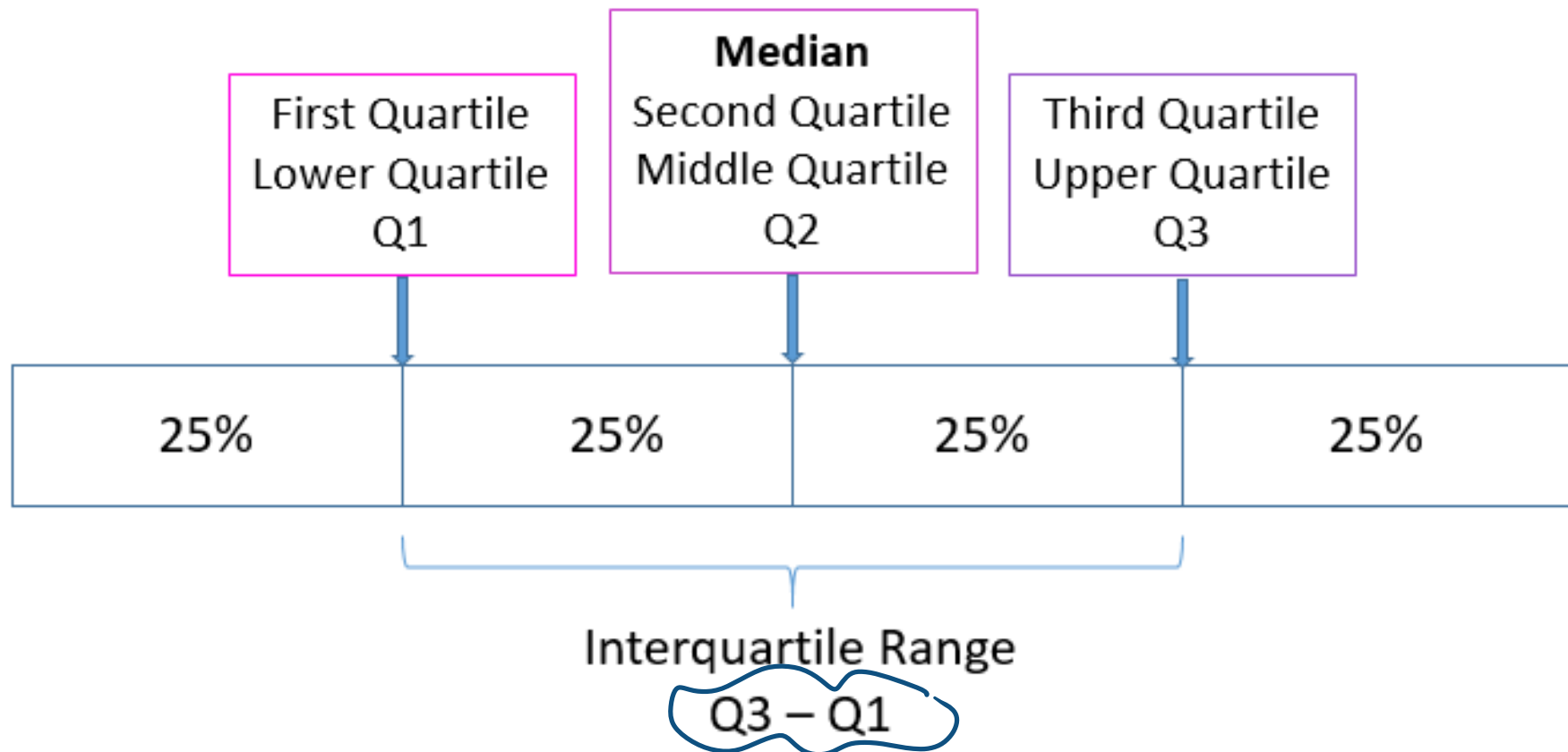
$$\text{موقع median} = \frac{5+1}{2} = 3$$

الجزء اليمين

$$Q_3 = 77$$

$$\begin{aligned} IQR &= 77 - 64 \\ &= 13 \end{aligned}$$

Median and Quartiles



Median and inter-quartile range are robust indicators of central tendency and dispersion

The median (second quartile) and inter-quartile range can be used as an alternative method for describing the central tendency and dispersion of a set of measured data. Both are robust and can be useful where there are occasional extreme values.

Measures of Position

- Example: we want to know the IQR, Semi-interquartile Range, and the Midquartile of the ordered grades from smallest to largest, $n=24$

58, 59, 68, 69, 74, 75, 78, 78, 78, 79, 79, 81, 83, 84, 85,
85, 86, 86, 88, 88, 89, 90, 90, 92

Measures of Position

Example: we want to know the Q1, Q2, Q3, IQR, Semi-interquartile Range , and the Midquartile, and the 10-90 percentile range of the ordered grades from smallest to largest, n=24

58, 59, 68, 69, 74, 75, 78, 78, 78, 79, 79, 81, 83, 84, 85, 85, 86, 86, 88, 88, 89, 90, 90, 92

$Q1 = 25\% = 25/100 * 24 = 6 + 0.5 = 6.5$ (position) which corresponds to 76.5.

$Q2 = 50 / 100 * 24 = 12 + 0.5$ (position) which corresponds to 82.

$Q3 = 75 / 100 * 24 = 18 + 0.5$ (position) which corresponds to 87.

$$\text{IQR} = Q3 - Q1 = 87 - 76.5 = 10.5$$

Semi-interquartile Range= $10.5/2=5.25$

Midquartile= $Q1+Q3/2 = 81.75$

10-90 percentile range= $P_{90}-P_{10}$.

$P_{90} = 0.9 * 24 = 21.6 + 0.5 = 22 = 90$, $P_{10} = 0.1 * 24 = 2.4 + 0.5 = 2.9 = 68$

10-90 percentile range= $P_{90}-P_{10}=90-68=22$

Example Find IQR, Semi IQR
and midquartile: $n=24$

58, 59, 68, 69, 74, 75, 78, 78, 78, 79, 79, 81, 83, 84, 85,
85, 86, 86, 88, 88, 89, 90, 90, 92

1. القيم مرتبة صاعداً ✓

2. اخرج Q_3 , Q_1 , Q_2

المتوسط
median
المتوسط

$$= \frac{24+1}{2} = 12.5$$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
58	59	68	69	74	75	78	78	78	79	79	81	83	84	85
85	86	86	88	88	89	90	90	92						

Arrows from 6.5 point to the 6th and 7th positions (75 and 78).
Arrows from 12.5 point to the 12th and 13th positions (81 and 83).

6.5

$$Q_2 = \frac{81+83}{2} = 82$$

3

$$= \frac{12+1}{2} = 6.5$$

$$Q_1 = \frac{75+78}{2} = 76.5$$

موقع
 Median
 من جداول

$$= \frac{12+1}{2} = 6.5$$

$$Q_3 = \frac{86+88}{2} = 87$$

*

$$IQR = Q_3 - Q_1 \quad \boxed{3}$$

$$= 87 - 76.5$$

$$= 10.5$$

*

$$Semi IQR = \frac{IQR}{2}$$

$$= \frac{10.5}{2} = 5.25$$

*

$$Mid quartile = \frac{Q_3 + Q_1}{2} = \frac{87 + 76.5}{2} = 81.75$$

* 10-90 Percentile Range

$$= P_{90} - P_{10}$$

$$= 90 - 68$$

$$= 22$$

توزيع
جدار P

$$L = \frac{K}{100} (n)$$

$$L = \frac{90}{100} (24) = 0.9 (24) = 21.6$$

موقع

$$P_{90} = L + 0.5$$

$$= 21.6 + 0.5 = 22.1 \approx 22$$

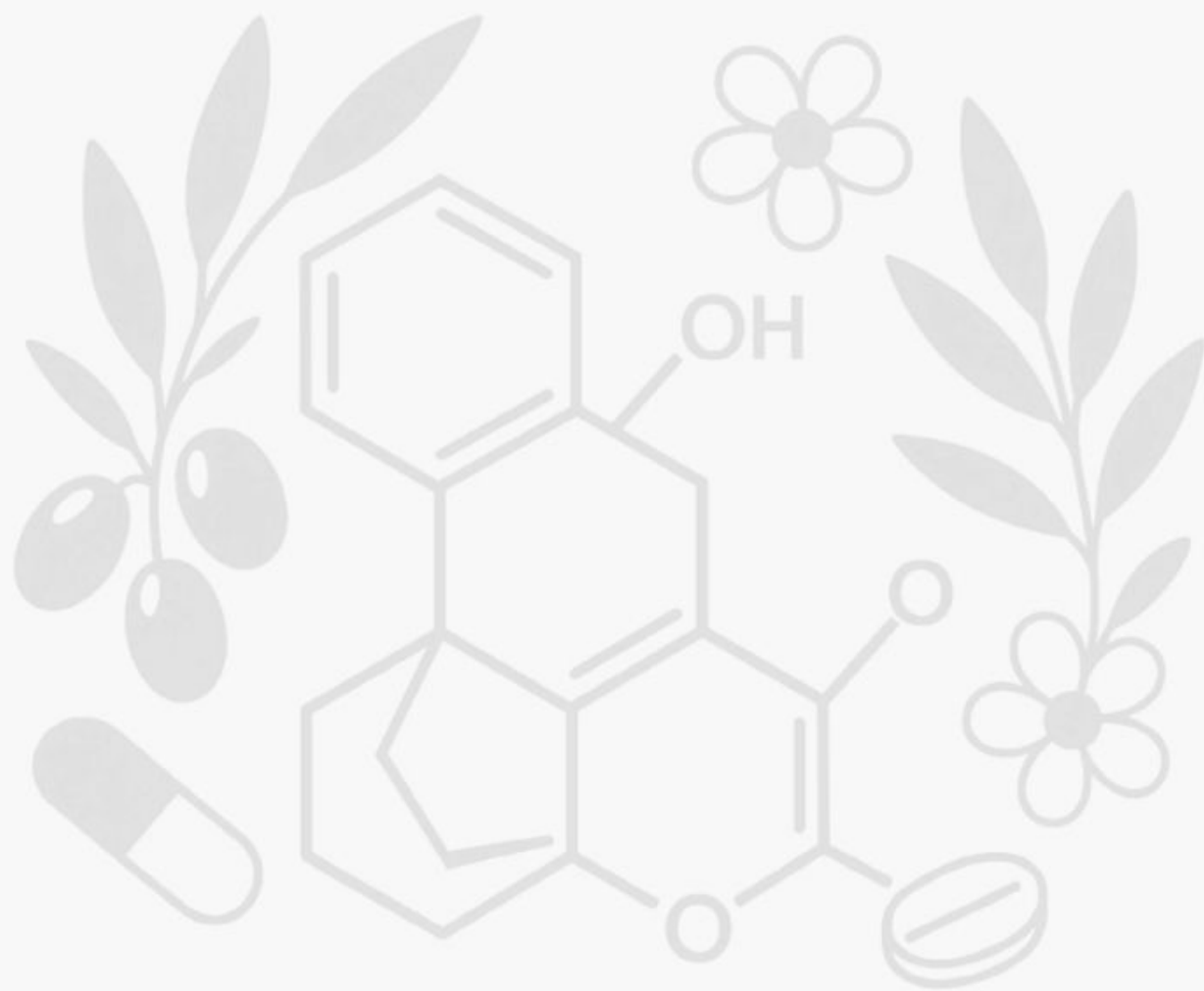
$$\Rightarrow \boxed{P_{90} = 90}$$

كيفية P10

$$P_{10} \Rightarrow L = \frac{K}{100} (n)$$

$$= \frac{10}{100} (24) = 0.1 (24) = 2.4$$

$$2.4 + 0.5 = 2.9 \approx 3$$



MORPHINE
ACADEMY