

Statistik Bisnis

Week 2

Numerical Descriptive Measures

Agenda

Time	Activity
First Session	
90 minutes	Central Tendency
Second Session	
60 minutes	Variation and Shape
30 minutes	Exploring Numerical Data

Objectives

In this chapter, you learn:

- To describe the properties of central tendency, variation,
- and shape in numerical data
- To construct and interpret a boxplot
- To compute descriptive summary measures for a population

Numerical Descriptive Measures



Central
Tendency



Variation and
Shape



Exploring
Numerical
Data



Numerical
Descriptive
Measures for
a Population

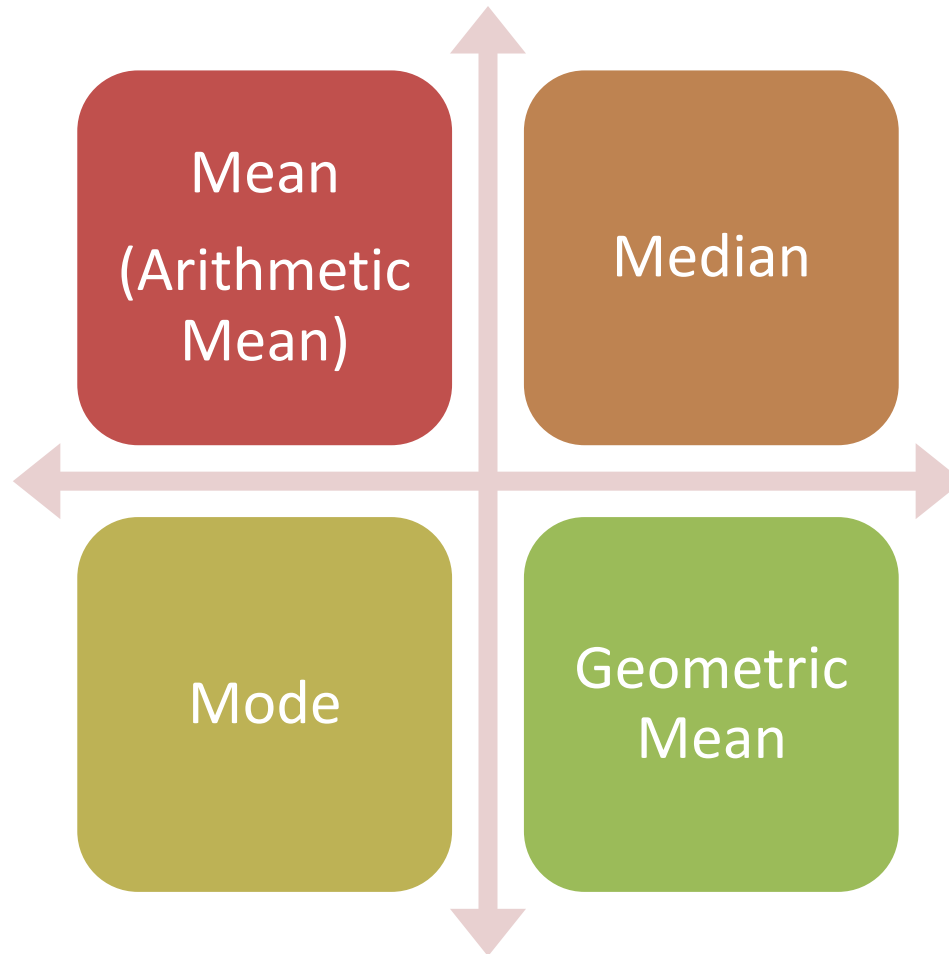


The
Covariance
and The
Coefficient of
Correlation



CENTRAL TENDENCY

Central Tendency



Mean (Rata-rata Hitung)

Perhatikan data berikut:

160 157 162 170 168 174 156 173 157

Berapakah rata-rata tinggi badan?

Mean

Pronounced
x-bar

The i^{th} value

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Sample size

Observed values

Mean

Sekarang, perhatikan data pengeluaran bulanan mahasiswa statistik bisnis berikut:

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Berapakah rata-ratanya?

Mean

In this case we can only **ESTIMATE** the **MEAN**...

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Keyword: "MIDPOINTS"

Estimated Mean

Midpoint	Frequency	Mid * f
250000	2	500000
750000	7	5250000
1250000	13	16250000
1750000	5	8750000
Total	27	30750000

$$\textit{Estimated Mean} = \frac{30750000}{27} = 1138888.89$$

Mean

Berikut adalah data nilai “Mahasiswa A”:

Mata Kuliah	SKS	Nilai
Matematika Bisnis	3	60
Bahasa Inggris	2	80
Perilaku Organisasi	3	100
Statistik	4	90
Manajemen Operasi	3	70

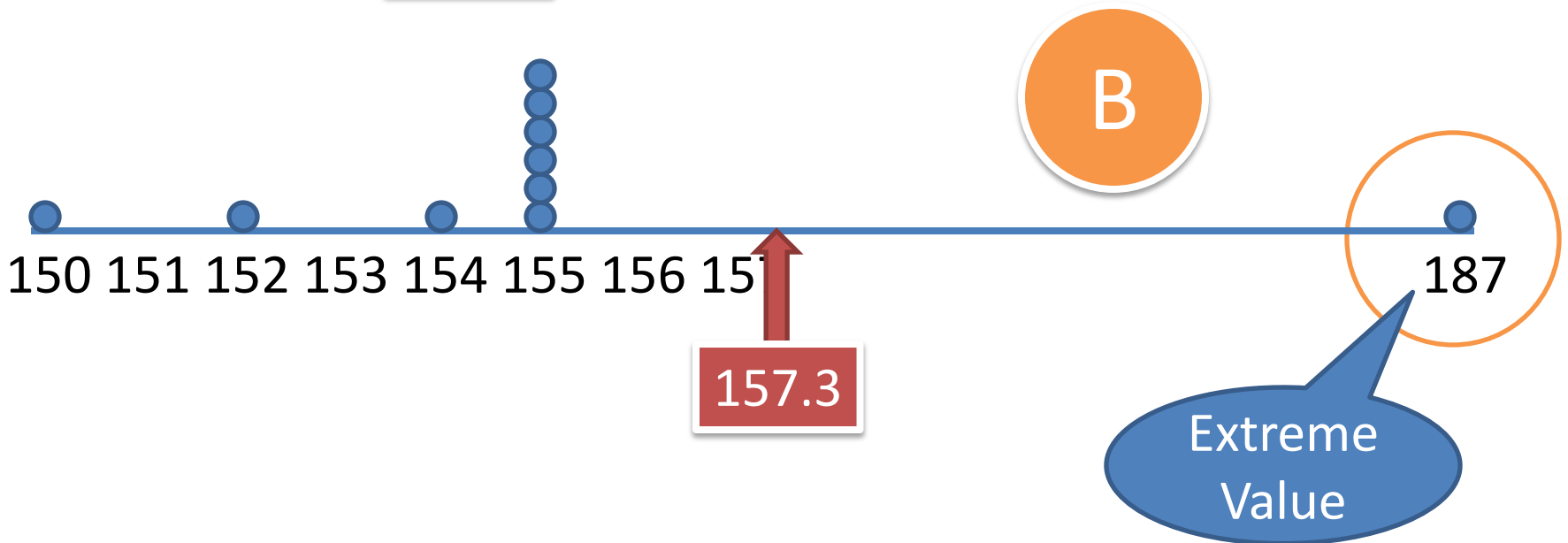
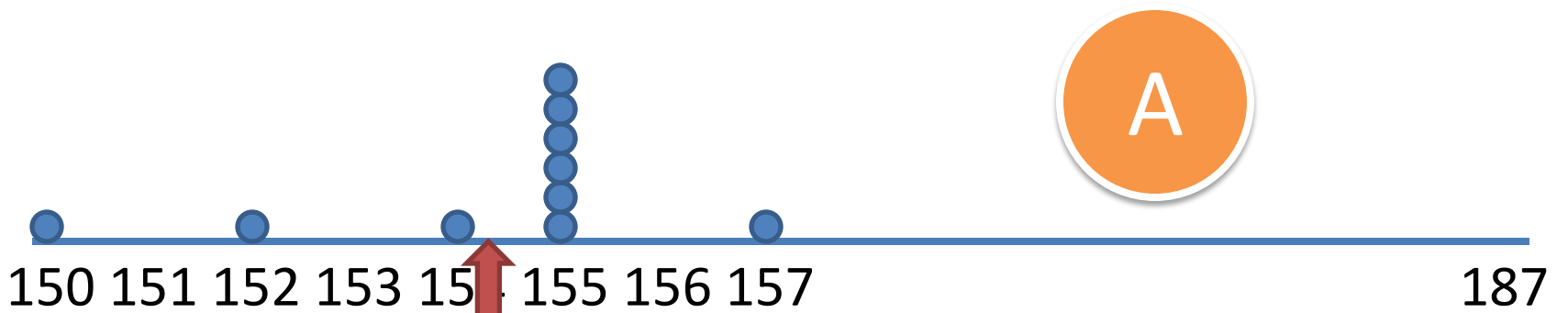
Berapakah nilai rata-rata “Mahasiswa A”?

Mean

Perhatikan dua data berikut ini:

A	150 152 154 155 155 155 155 155 155 157	Rata-rata?
B	150 152 154 155 155 155 155 155 155 187	Rata-rata?

Mean



It is DANGEROUS
to ONLY use
MEAN in
describing a data



Median

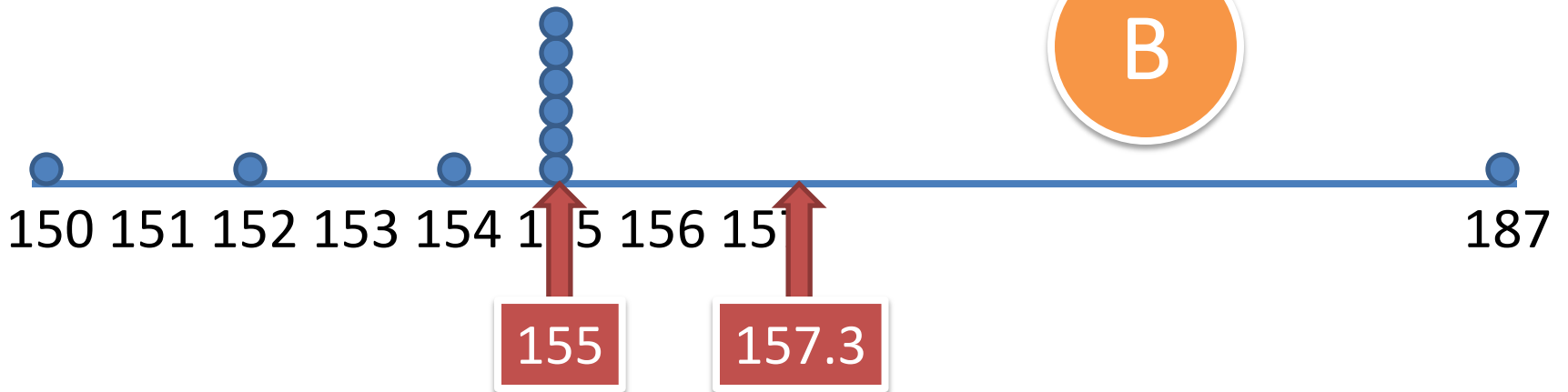
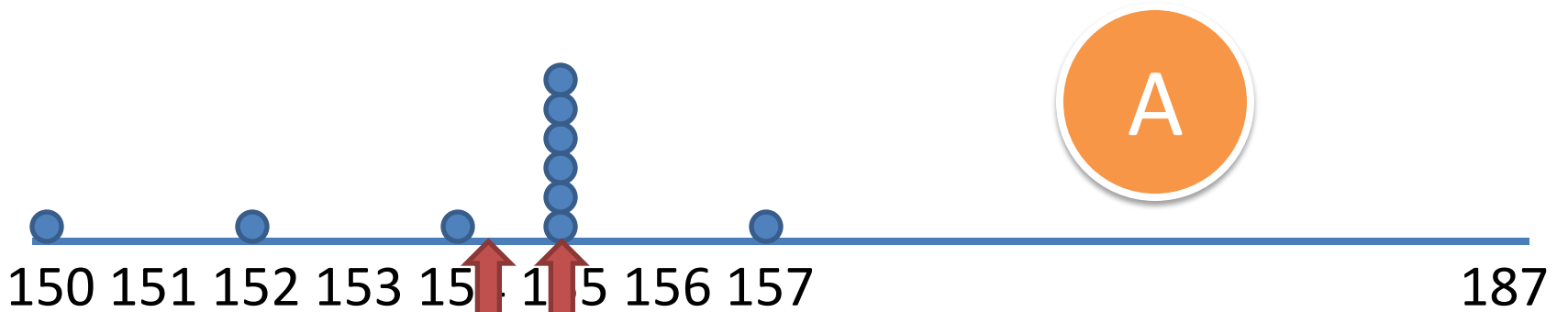
Median position = $\frac{n+1}{2}$ position in the ordered data

Median

Perhatikan dua data berikut ini:

A	150 152 154 155 155 155 155 155 155 157	Median?
B	150 152 154 155 155 155 155 155 155 187	Median?

Median



Median

Berapakah median dari data tinggi badan berikut?

160 157 162 170 168 174 156 173 157

Hitung pula median dari data berikut?

160 157 162 170 168 174 156 173 157 150

Median

Perhatikan lagi data pengeluaran bulanan mahasiswa statistik bisnis berikut:

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Berapakah mediannya?

Median

The MEDIAN group of monthly spending is Rp. 1.000.000 but less than Rp. 1.500.000

Or
ESTIMATE
the
MEDIAN!!



Estimated Median

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Estimated Median = Rp. 1.173.076,92

Estimated Median

$$\text{Estimated Median} = L + \frac{(n/2) - cf_b}{f_m} \times w$$

where:

- **L** is the lower class boundary of the group containing the median
- **n** is the total number of data
- **cf_b** is the cumulative frequency of the groups before the median group
- **f_m** is the frequency of the median group
- **w** is the group width

Mode (Modus)

Berapakah modus data berikut:

160 157 162 170 168 174 156 173 157

Berapa pula modus data berikut:

160 157 162 170 168 174 156 173 150

Mode

Perhatikan lagi data pengeluaran bulanan mahasiswa statistik bisnis berikut:

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Berapakah modusnya?

Mode

The MODAL group of monthly spending is Rp. 1.000.000 but less than Rp. 1.500.000

But the actual **Mode** may not even be in that group!





Mode

Without the raw data we don't really know...

However, we can
ESTIMATE the **MODE**



Estimated Mode

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Estimated Mode

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 - kurang dari Rp. 1.000.000	7
Rp. 1.000.000 - kurang dari Rp. 1.500.000	13
Rp. 1.500.000 - kurang dari Rp. 2.000.000	5

Estimated Mode = Rp. 1.214.285,72

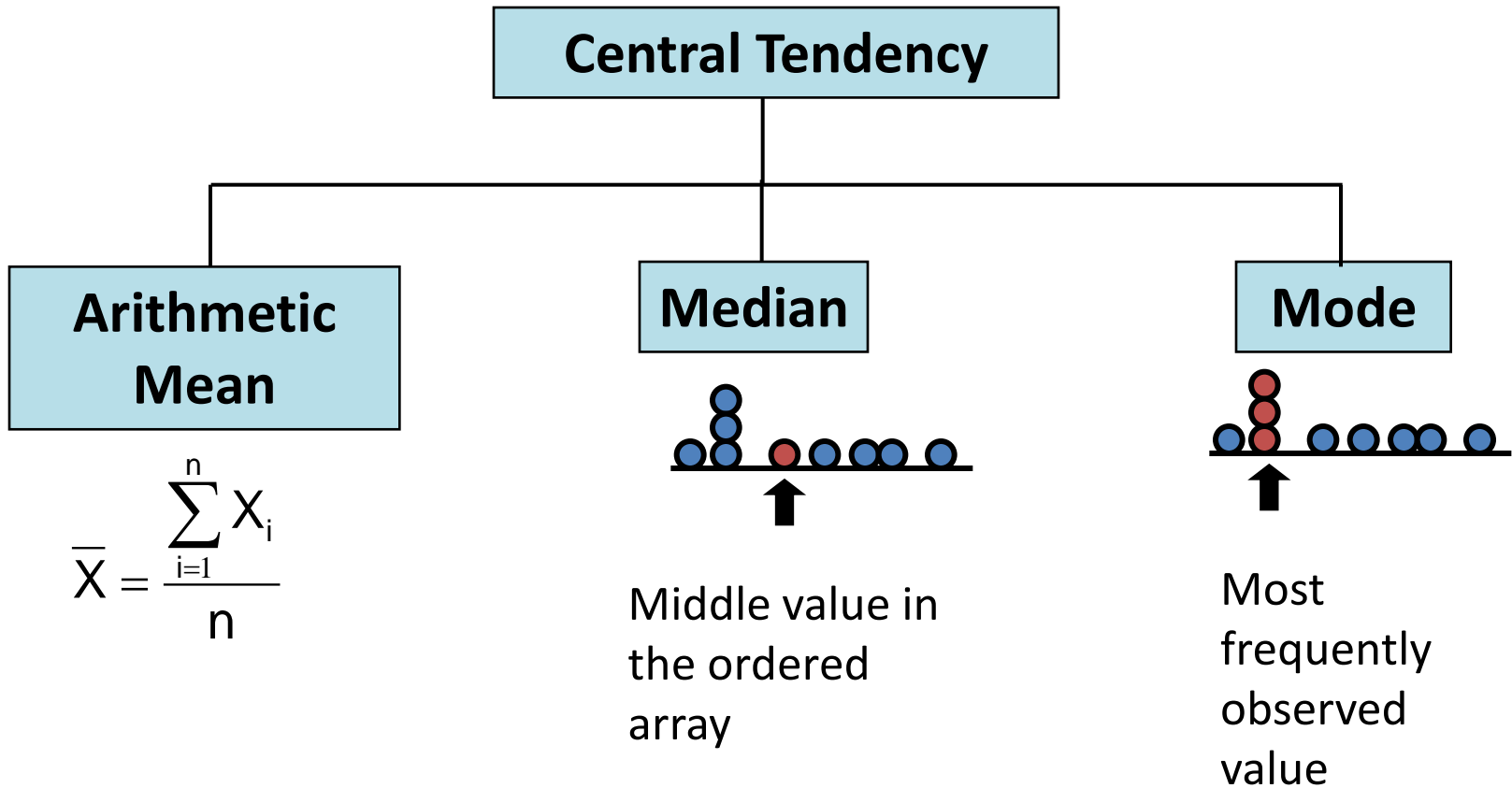
Estimated Mode

$$\text{Estimated Mode} = L + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} \times w$$

where:

- L is the lower class boundary of the modal group
- f_{m-1} is the frequency of the group before the modal group
- f_m is the frequency of the modal group
- f_{m+1} is the frequency of the group after the modal group
- w is the group width

Central Tendency



EXERCISE

3.10

Data berikut adalah data pengeluaran yang dilakukan oleh sampel sembilan orang konsumen untuk makan siang di sebuah restoran cepat saji (dalam \$):

4,20 5,03 5,86 6,45 7,38 7,54 8,46 8,47 9,87

Tentukan rata-rata dan median!

3.12

Berikut adalah data konsumsi bensin per kilometer dari mobil-mobil SUV tahun 2010:

24	23	22	21	22	22	18	18	26
26	26	19	19	19	21	21	21	21
21	18	29	21	22	22	16	16	

Tentukan median and modus!

GEOMETRIC MEAN

Compounding Data



Interest
Rate

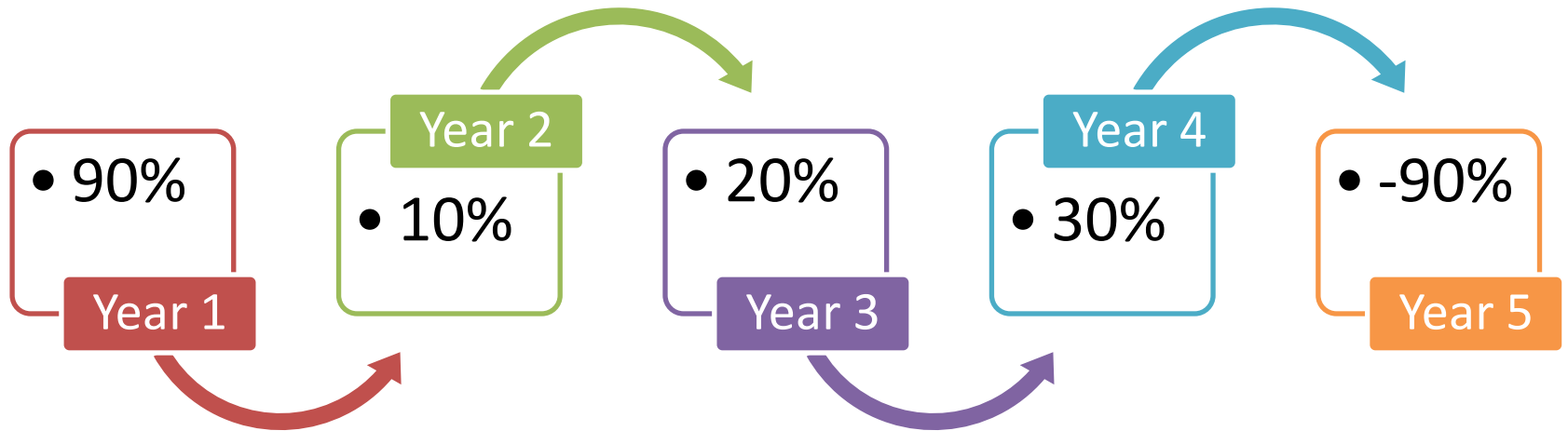
Growth
Rate

Return
Rate

Compounding Data

Misalkan anda telah menginvestasikan uang anda pada bursa saham selama lima tahun. Jika tingkat pengembalian tiap tahunnya adalah 90%, 10%, 20%, 30% dan -90%, berapakah rata-rata tingkat pengembalian per tahun pada periode ini?

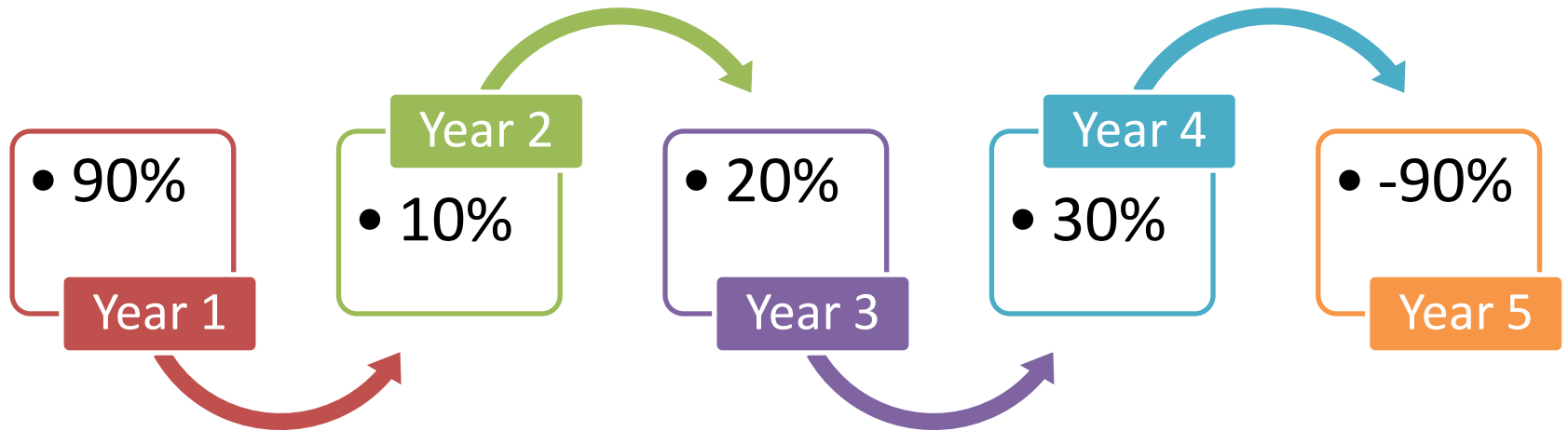
Compounding Data



If we use arithmetic mean in this case

The average return during this period = **12%**

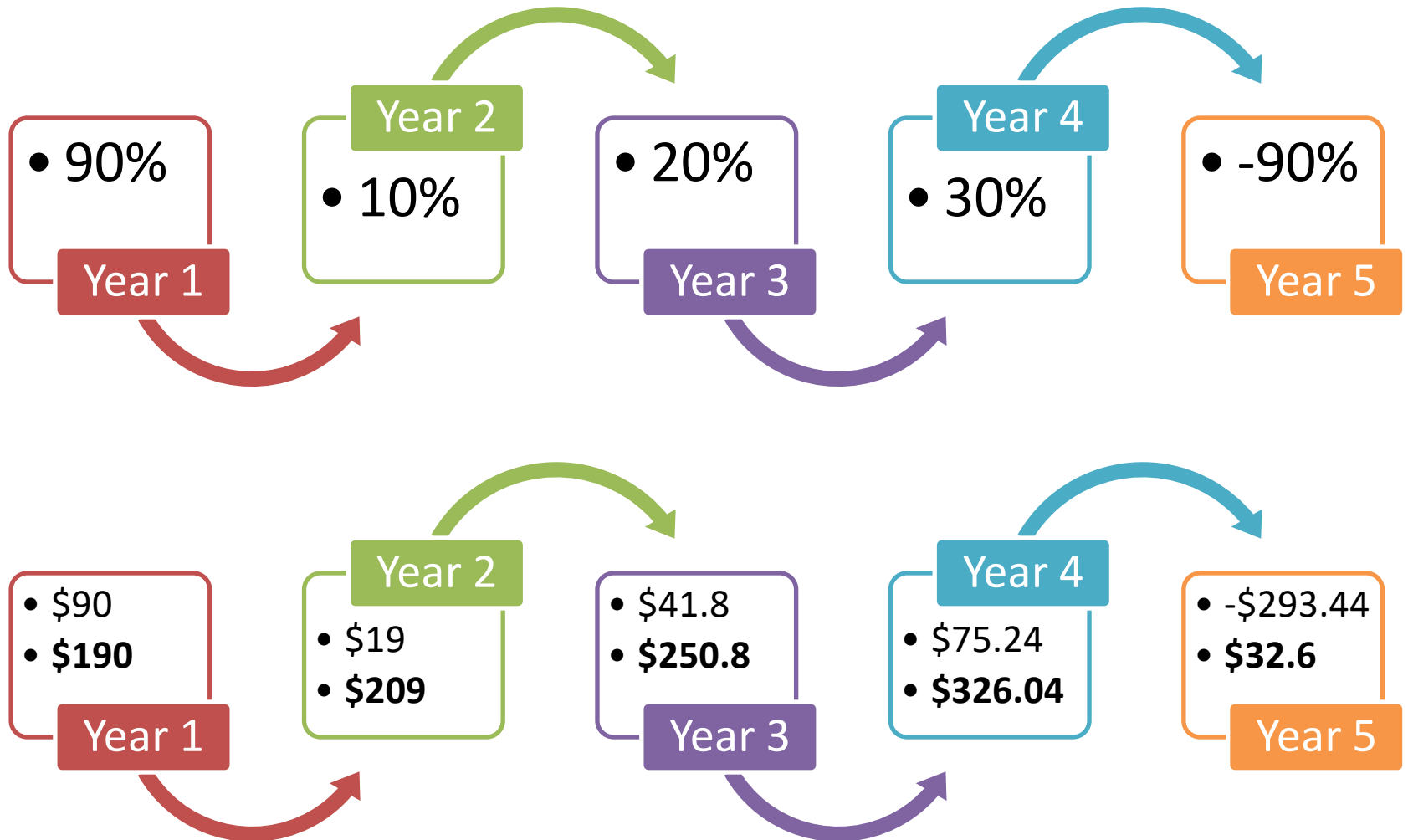
Compounding Data



Let say that you invest \$100 in year 0

How much your stocks worth in year 5?

Compounding Data



Geometric Mean

$$GM = \sqrt[5]{1.9 \times 1.1 \times 1.2 \times 1.3 \times 0.1} - 1$$

$$GM = -20.08\%$$

Well, that's pretty bad...

This is called **geometric mean rate of return**



Measure of Central Tendency For The Rate Of Change Of A Variable Over Time:

The Geometric Mean & The Geometric Rate of Return

- Geometric mean
 - Used to measure the rate of change of a variable over time

$$\bar{X}_G = (X_1 \times X_2 \times \Lambda \times X_n)^{1/n}$$

- Geometric mean rate of return
 - Measures the status of an investment over time

$$\bar{R}_G = [(1 + R_1) \times (1 + R_2) \times \Lambda \times (1 + R_n)]^{1/n} - 1$$

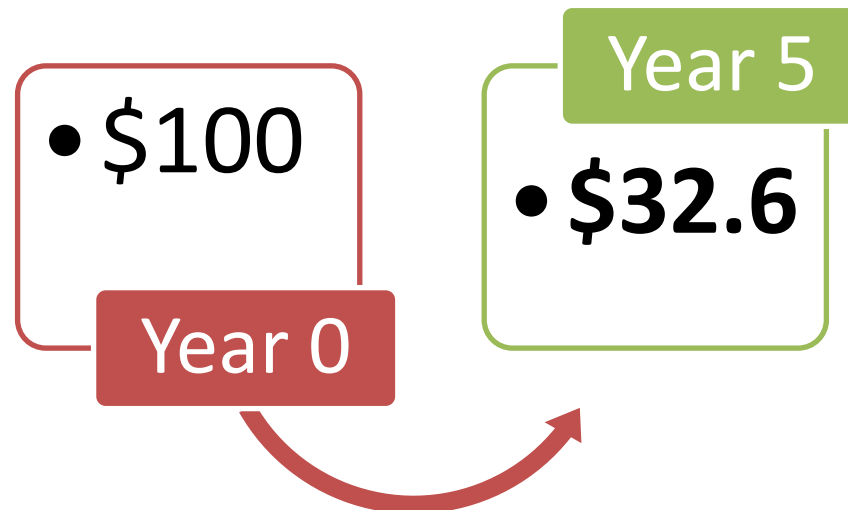
- Where R_i is the rate of return in time period i

Geometric Mean

$$GM = \sqrt[n]{\frac{\text{End of Period Value}}{\text{Beginning of period Value}}} - 1$$

Geometric Mean

Mari perhatikan permasalahan tadi. Diketahui bahwa kita berinvestasi saham senilai \$100 pada tahun ke-0. Namun, pada akhir tahun ke-5 nilai saham tersebut menjadi \$32.6. Hitunglah rata-rata tingkat pengembalian tahunan!



Geometric Mean

$$GM = \sqrt[5]{\frac{32.6}{100}} - 1$$

$$GM = -20.08\%$$

- This value consistent with what we found earlier



Example

Data penduduk Jawa Barat:

- Tahun 2000: 35.729.537 jiwa
- Tahun 2010: 43.053.732 jiwa

Tingkat pertumbuhan penduduk per tahun?

EXERCISE

3.22

Pada tahun 2006-2009, harga logam mulia cepat berubah. Tabel berikut menunjukkan total tingkat pengembalian (dalam persentase) untuk platina, emas, dan perak dari tahun 2006 hingga tahun 2009:

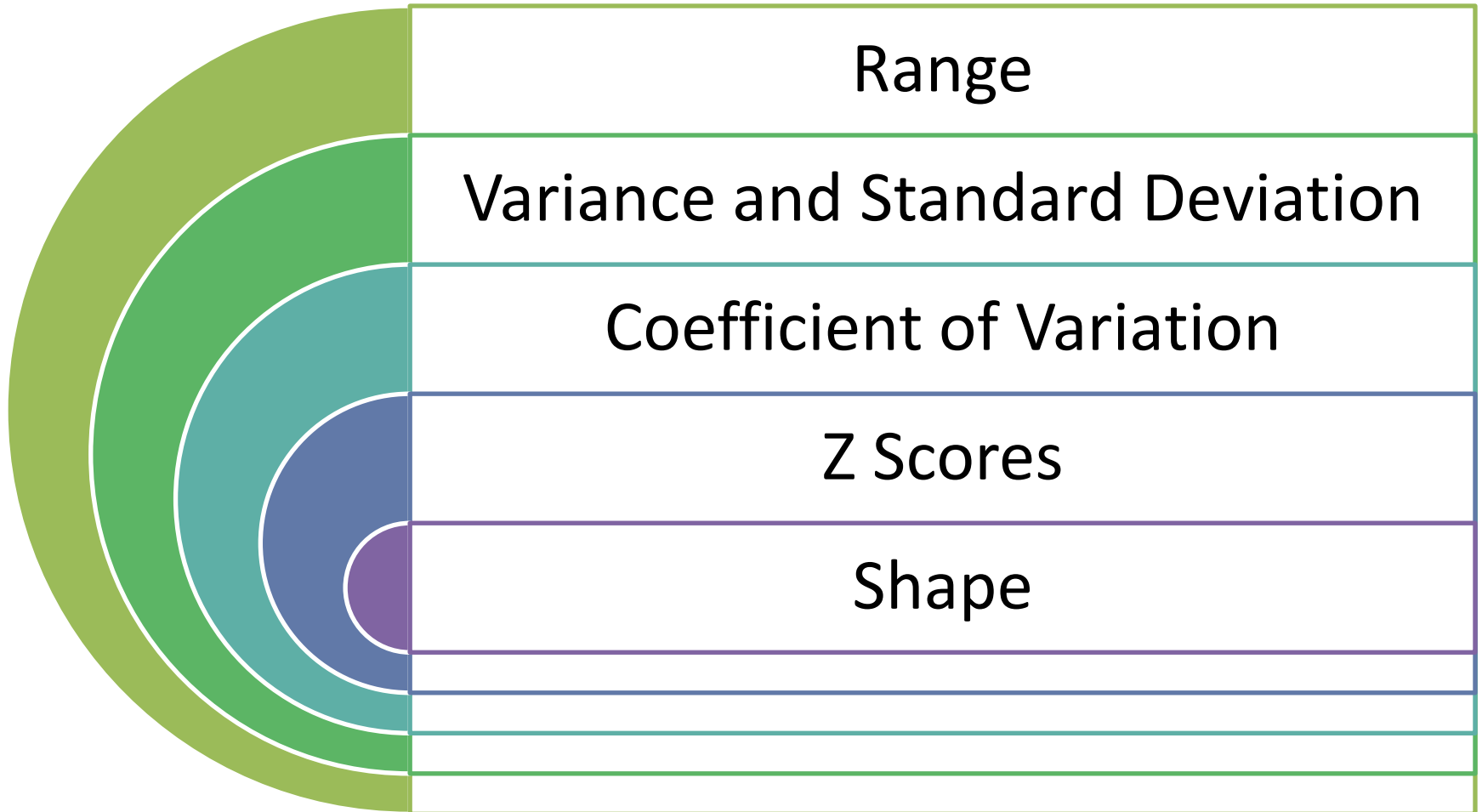
Tahun	Platina	Emas	Perak
2009	62.7	25.0	56.8
2008	-41.3	4.3	-26.9
2007	36.9	31.9	14.4
2006	15.9	23.2	46.1

3.22

- a. Hitung rata-rata tingkat pengembalian per tahun untuk platina, emas, dan perak dari tahun 2006 hingga tahun 2009.
- b. Apakah kesimpulan yang bisa kita tarik mengenai rata-rata tingkat pengembalian per tahun dari tiga logam mulia tersebut?

VARIATION AND SHAPE

Variation and Shape



Review on Central Tendency

Perhatikan data berikut:

160 157 162 170 168 174 156 173 157 150

Hitunglah rata-rata, median, dan modus?

RANGE

Range (Rentang)

Perhatikan data berikut:

160 157 162 170 168 174 156 173 157 150

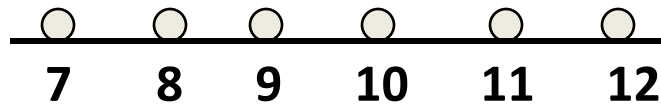
Berapakah rentangnya?

Range

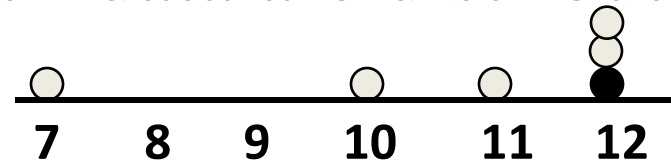
$$\textit{Range} = X_{\max} - X_{\min}$$

Measures of Variation: Why The Range Can Be Misleading

- Ignores the way in which data are distributed



$$\text{Range} = 12 - 7 = 5$$



$$\text{Range} = 12 - 7 = 5$$

- Sensitive to outliers

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,5

$$\text{Range} = 5 - 1 = 4$$

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,120

$$\text{Range} = 120 - 1 = 119$$

VARIANCE AND STANDARD DEVIATION

Variance and Standard Deviation



Deviation

Perhatikan kembali data berikut:

160 157 162 170 168 174 156 173 157 150


Berapakah rata-ratanya?

Mean = 162.7

Deviation

Data	Deviation
160	-2.7
157	-5.7
162	-0.7
170	7.3
168	5.3
174	11.3
156	-6.7
173	10.3
157	-5.7
150	-12.7

$$\textit{Deviation} = X_i - \bar{X}$$


$$= 156 - 162.7$$

Variance and Standard Deviation

Data	Deviation	(Dev)^2
160	-2.7	7.29
157	-5.7	32.49
162	-0.7	0.49
170	7.3	53.29
168	5.3	28.09
174	11.3	127.69
156	-6.7	44.89
173	10.3	106.09
157	-5.7	32.49
150	-12.7	161.29

Sum of Squares
= 594.1

Variance and Standard Deviation

Sample size (n) = 10

$$\text{Variance} = \frac{594.1}{10 - 1} = 66.01$$

$$\text{SD} = \sqrt{66.01} = 8.125$$

Variance and Standard Deviation

- Sample

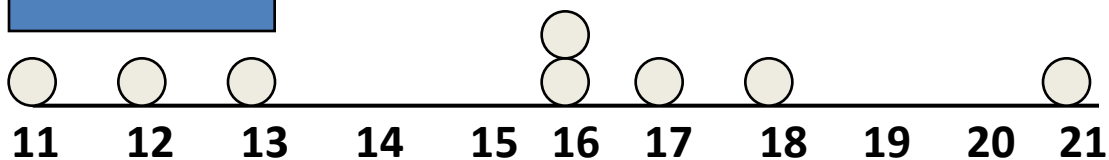
$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

- Population

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{N}$$

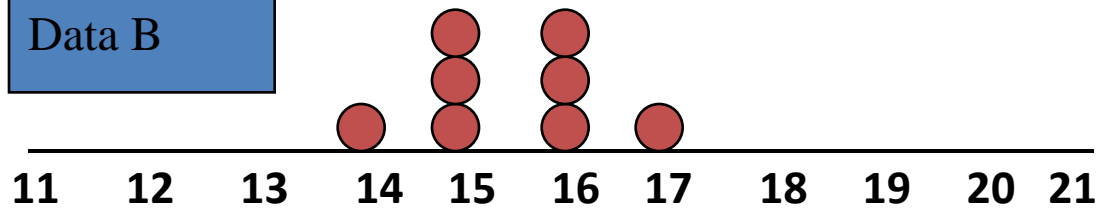
Measures of Variation: Comparing Standard Deviations

Data A



Mean = 15.5
 $S = 3.338$

Data B



Mean = 15.5
 $S = 0.926$

Data C



Mean = 15.5
 $S = 4.570$

Standard Deviation

Sekarang, perhatikan data pengeluaran bulanan mahasiswa statistik bisnis berikut:

Pengeluaran Bulanan	Frekuensi
Kurang dari Rp. 500.000	2
Rp. 500.000 – kurang dari Rp. 1.000.000	7
Rp. 1.000.000 – kurang dari Rp. 1.500.000	13
Rp. 1.500.000 – kurang dari Rp. 2.000.000	5

Berapakah SIMPANGAN BAKU?

Standard Deviation

Sekarang, perhatikan data pengeluaran bulanan mahasiswa statistik

Pengeluaran	Frekuensi
Kurang dari Rp. 500.000	
Rp. 500.000 – kurang dari Rp. 1.000.000	
Rp. 1.000.000 – kurang dari Rp. 1.500.000	13
Rp. 1.500.000 – kurang dari Rp. 2.000.000	5

E.S.T.I.M.A.T.I.O.N

Berapakah SIMPANGAN BAKU?

Estimated Standard Deviation

Midpoint	Frequency	Dev ²	(Dev ²)*f
250000	2	790123456790.12	1580246913580.25
750000	7	151234567901.24	1058641975308.64
1250000	13	12345679012.35	160493827160.49
1750000	5	373456790123.46	1867283950617.28
Total	27		46666666666666.67

$$\text{Variance} = \frac{46666666666666.67}{27-1} = 179487179487.18$$

$$SD = \sqrt{179487179487.18} = 423659.27$$

THE COEFFICIENT OF VARIATION

The Coefficient of Variation



Height

Weight

The Coefficient of Variation

Perhatikan kembali data berikut:

160 157 162 170 168 174 156 173 157 150

Berapakah rata-rata dan simpangan baku?

Mean = 162.7 and SD = 8.125

The Coefficient of Variation

Mahasiswa pada data tinggi badan sebelumnya, memiliki berat badan sebagai berikut:

50 55 57 52 55

69 60 65 71 70

Berapakah rata-rata dan simpangan baku?

Mean = 60.4 and SD = 7.8

The Coefficient of Variation

	Height	Weight
Mean	162.7	60.4
SD	8.125	7.8

Variabel manakah yang datanya lebih bervariasi?

Coefficient of Variation:

$$CV_{\text{Height}} = 4.99\%$$

$$CV_{\text{Weight}} = 12.92\%$$

The Coefficient of Variation

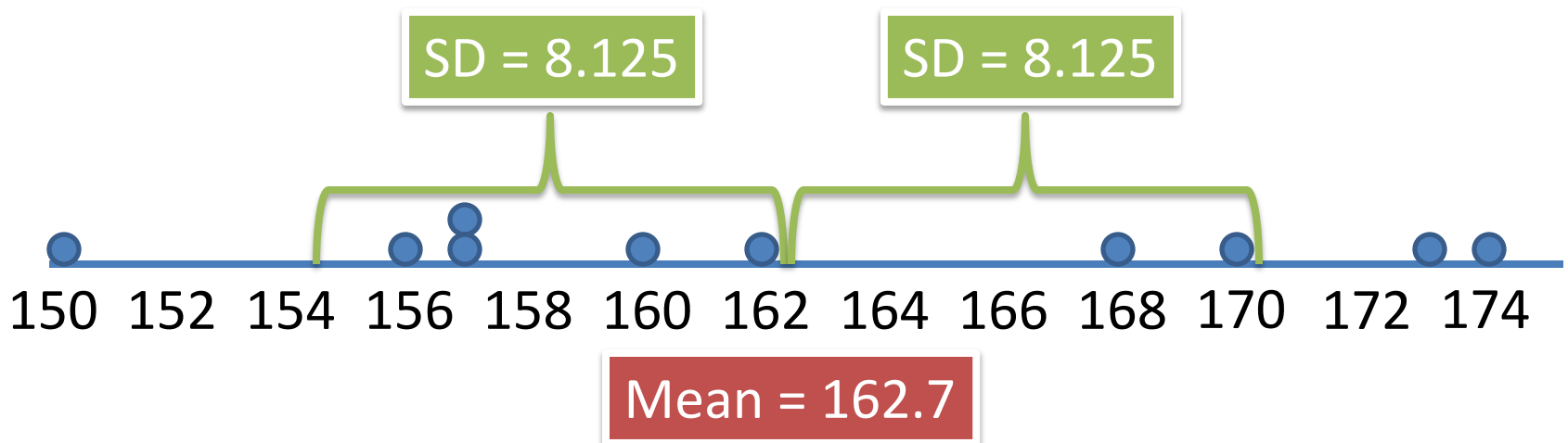
$$CV = \left(\frac{SD}{\bar{X}} \right) \cdot 100\%$$

LOCATING EXTREME OUTLIERS: Z SCORE

Locating Extreme Outliers: Z Score

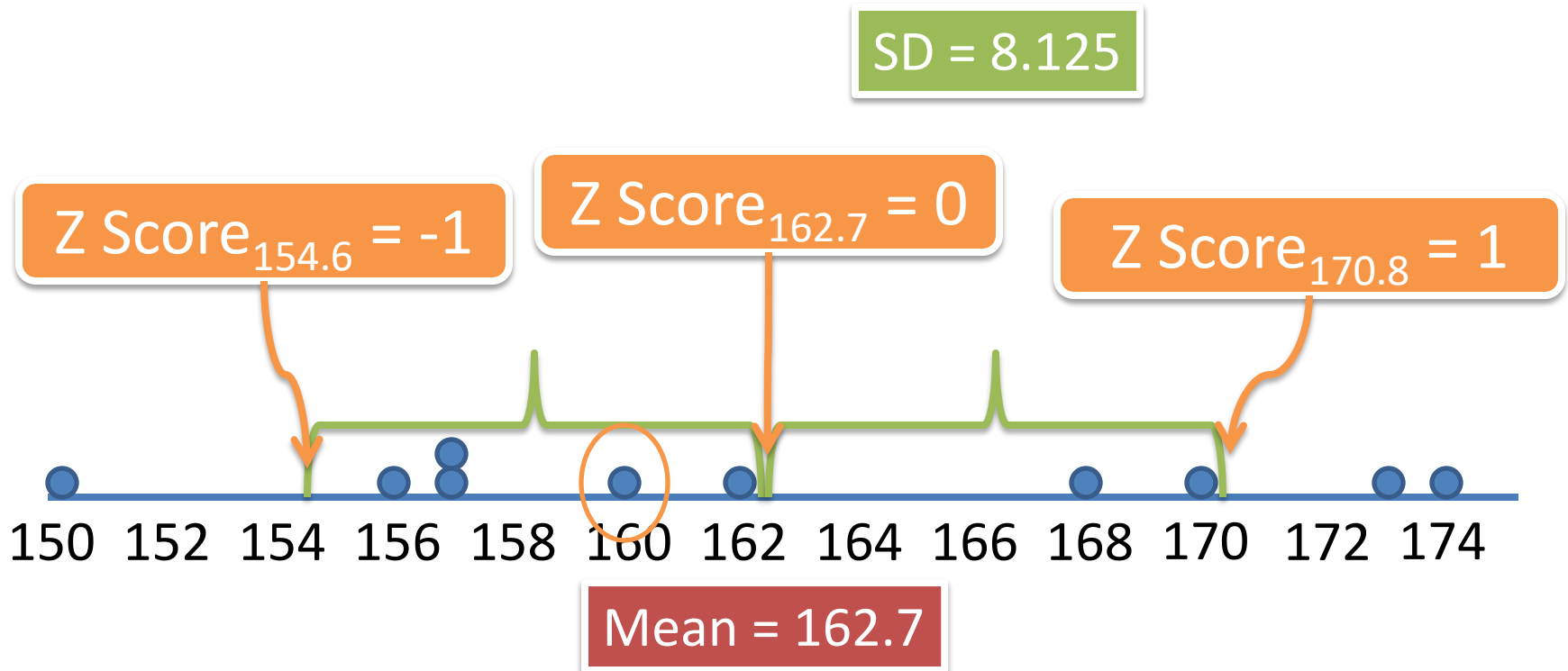
Perhatikan lagi data tinggi badan berikut:

160 157 162 170 168 174 156 173 157 150



Locating Extreme Outliers: Z Score

Maka, Z Score untuk 160 adalah?



Locating Extreme Outliers: Z Scores

Perhatikan kembali data berikut:

160 157 162 170 168 174 156 173 157 150

Berapakah Z Scores dari 160, 174, 168 dan 150?

$Z_{160} = -0.33$, $Z_{174} = 1.39$, $Z_{168} = 0.65$, and

$Z_{150} = -0.56$

Locating Extreme Outliers: Z Score

$$Z_X = \frac{X - \bar{X}}{SD}$$

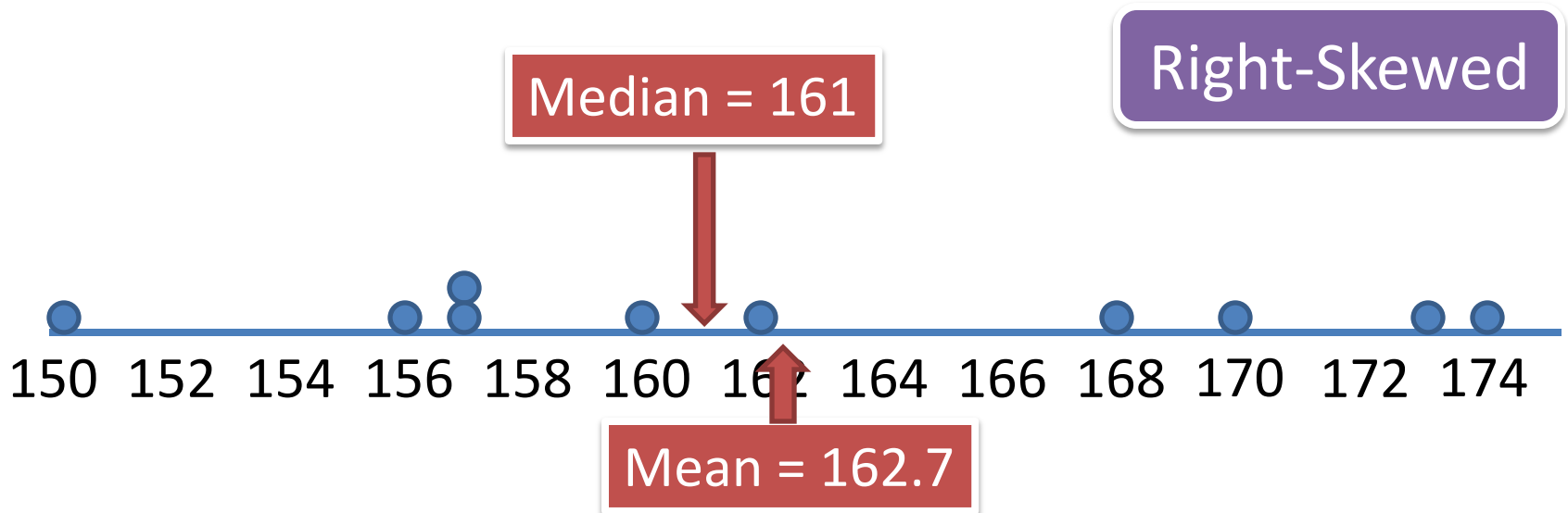
- A data value is considered an extreme outlier if its Z-score is less than -3.0 or greater than +3.0.
- The larger the absolute value of the Z-score, the farther the data value is from the mean.

SHAPE

Shape

Perhatikan kembali data berikut:

160 157 162 170 168 174 156 173 157 150



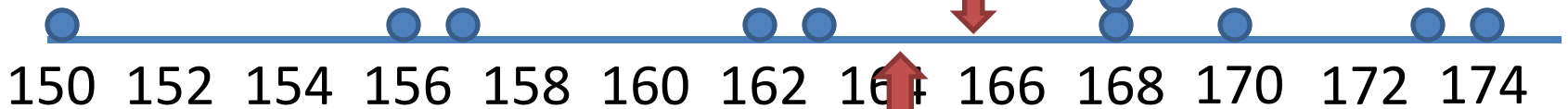
Shape

Bagaimana jika datanya seperti berikut:

163 168 162 170 168 174 156 173 157 150

Left-Skewed

Median = 165.5



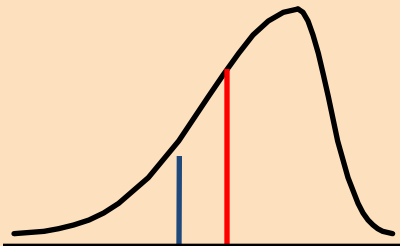
Mean = 164.1

Shape

Describes how data are distributed

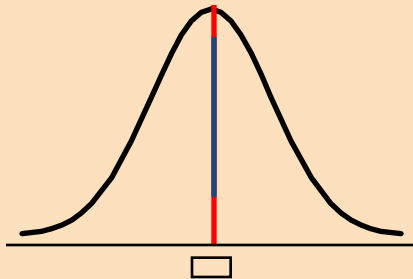
Left-Skewed

Mean < Median



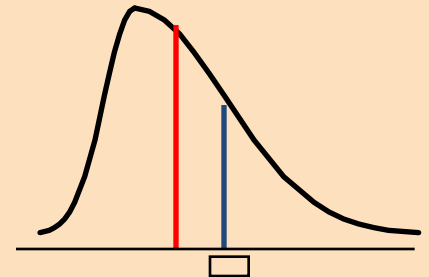
Symmetric

Mean = Median



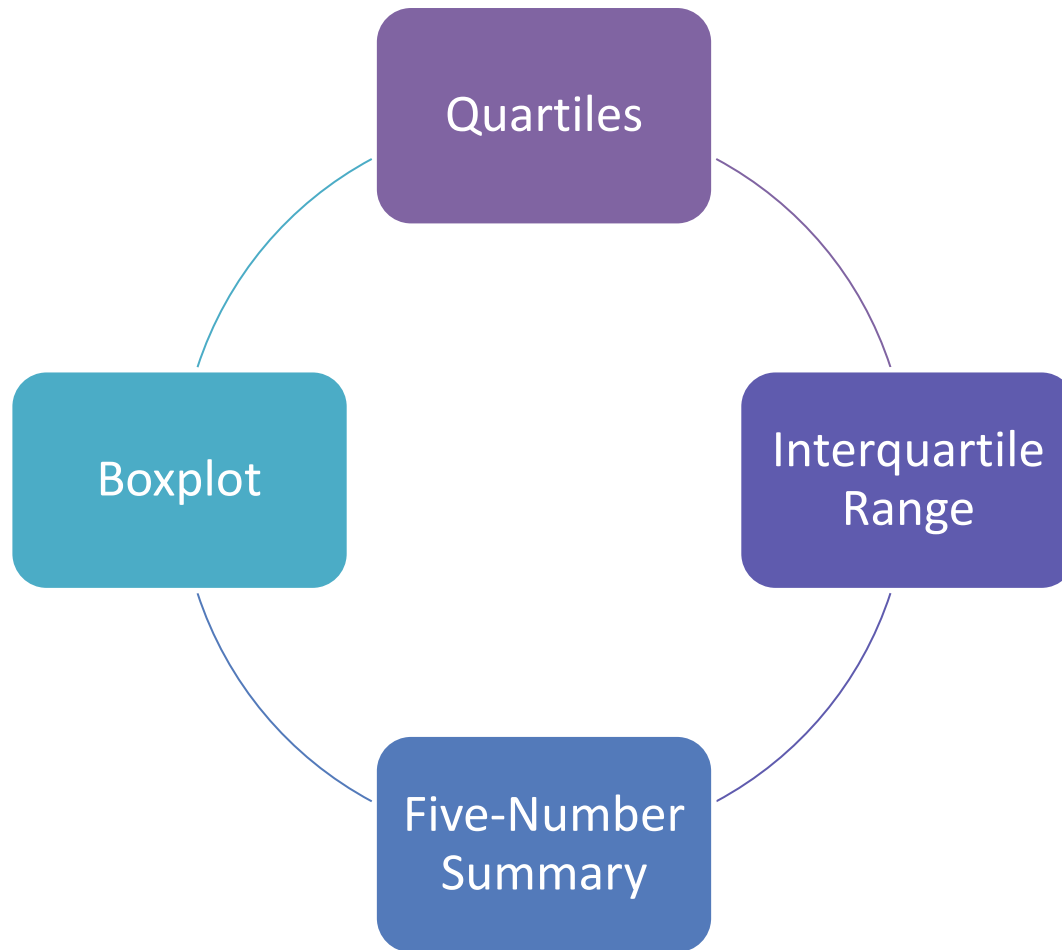
Right-Skewed

Median < Mean



EXPLORING NUMERICAL DATA

Exploring Numerical Data



QUARTILES

Quartiles

1st Quartile

Q_1

2nd Quartile

Q_2

Median

3rd Quartile

Q_3

Quartiles

Perhatikan data tinggi badan berikut:

160 157 162 170 168 174 156

Berapakah Q_1 , Q_2 dan Q_3 ?

$$Q_1 = 157$$

$$Q_2 = 162 \text{ (Median)}$$

$$Q_3 = 170$$

Quartiles

Perhatikan juga data berikut:

160 157 162 170 168 174 156 173 150

Berapa Q_1 , Q_2 dan Q_3 ?

$$Q_1 = 156.5$$

$$Q_2 = 162 \text{ (Median)}$$

$$Q_3 = 171.5$$

Quartiles

Dan perhatikan pula data berikut:

160 157 162 170 168 174 156 173 157 150

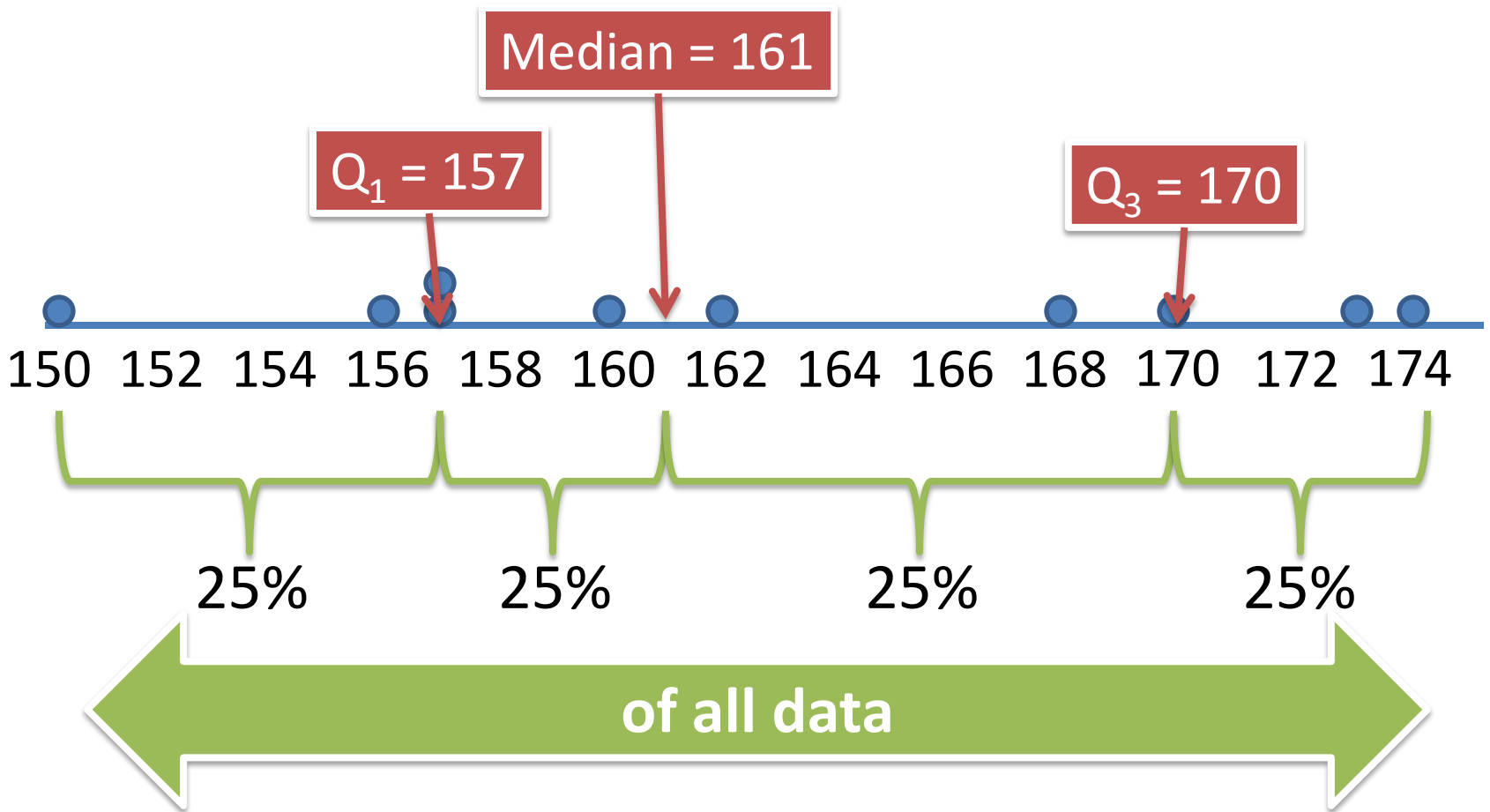
Berapakah Q_1 , Q_2 dan Q_3 ?

$$Q_1 = 157$$

$$Q_2 = 161 \text{ (Median)}$$

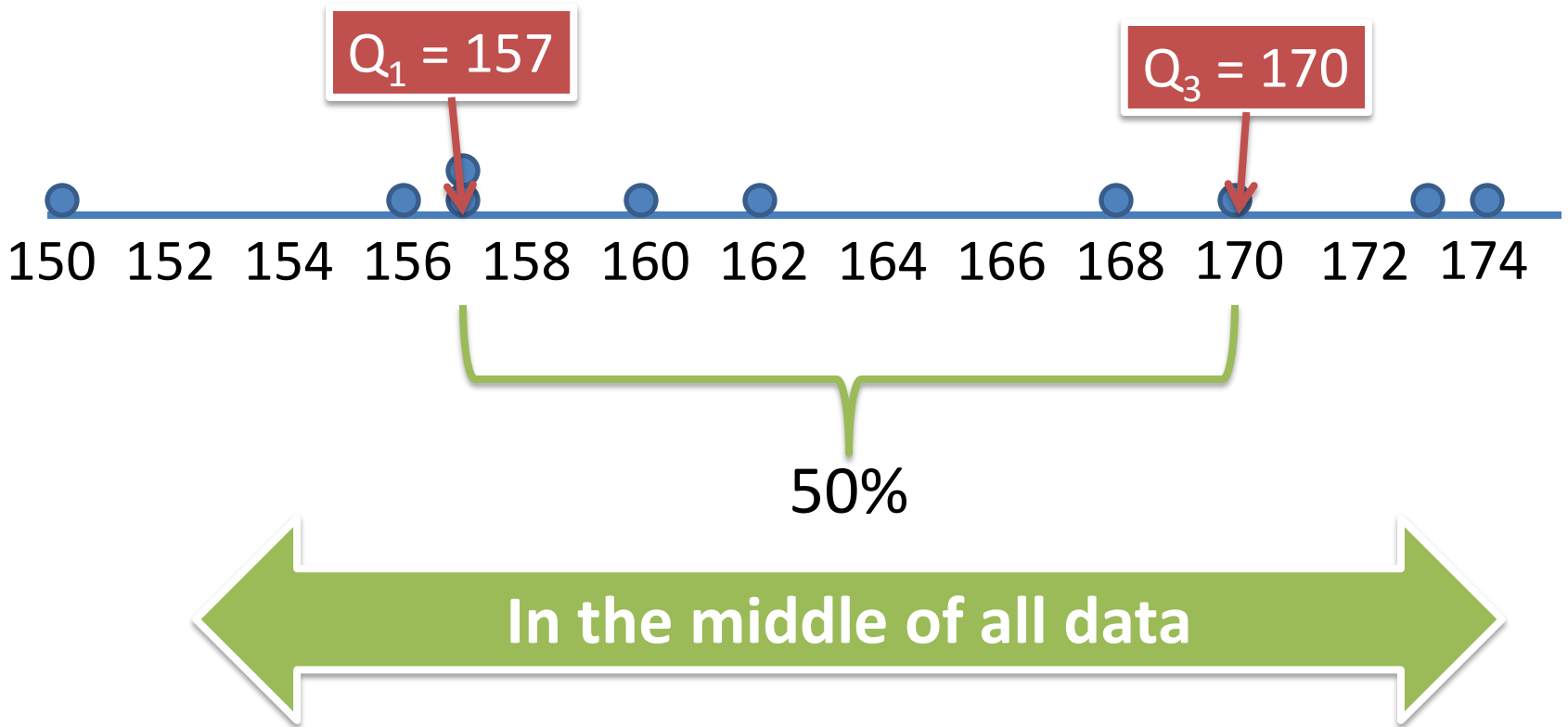
$$Q_3 = 170$$

Quartiles

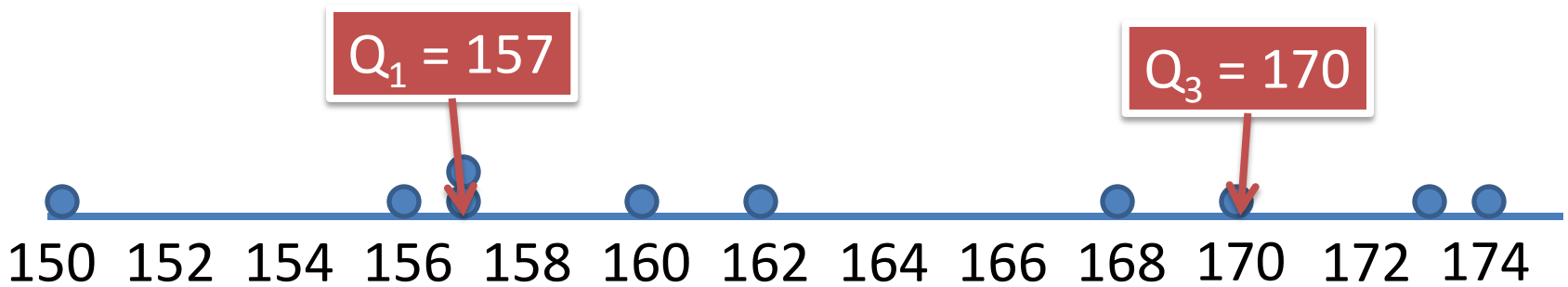


INTERQUARTILE RANGE

Interquartile Range



Interquartile Range



Berapakah rentang Interquartile?

$$\text{Interquartile Range} = 170 - 157 = 13$$

Interquartile Range

$$\textit{Interquartile Range} = Q_3 - Q_1$$

FIVE-NUMBER SUMMARY

Five-Number Summary

X_{\min} Q_1 *Median* Q_3 X_{\max}

Five-Number Summary

Perhatikan data tinggi badan berikut:

160 157 162 170 168 174 156 173 157 150

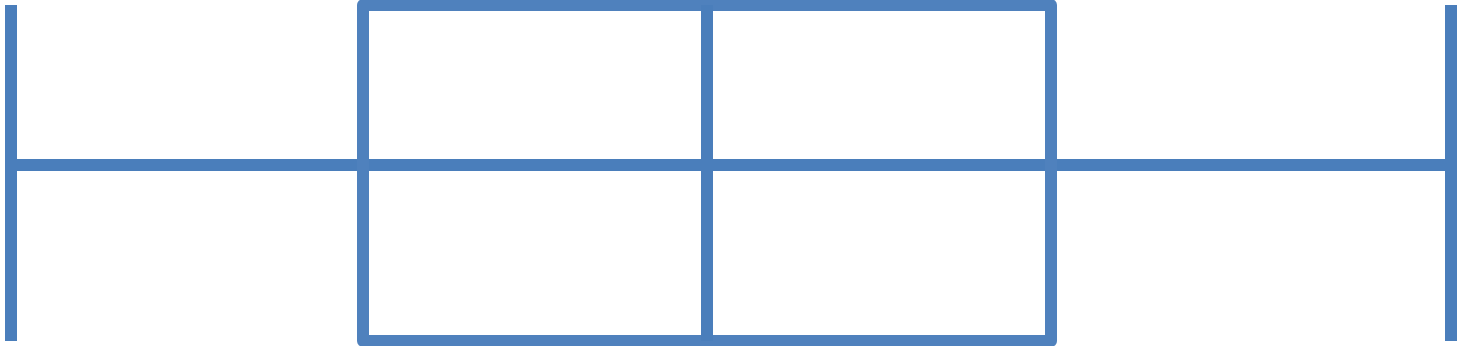
Apakah *Five-Number Summary*?

150 157 161 170 174

BOXPLOT

Boxplot

X_{\min} Q_1 Median Q_3 X_{\max}



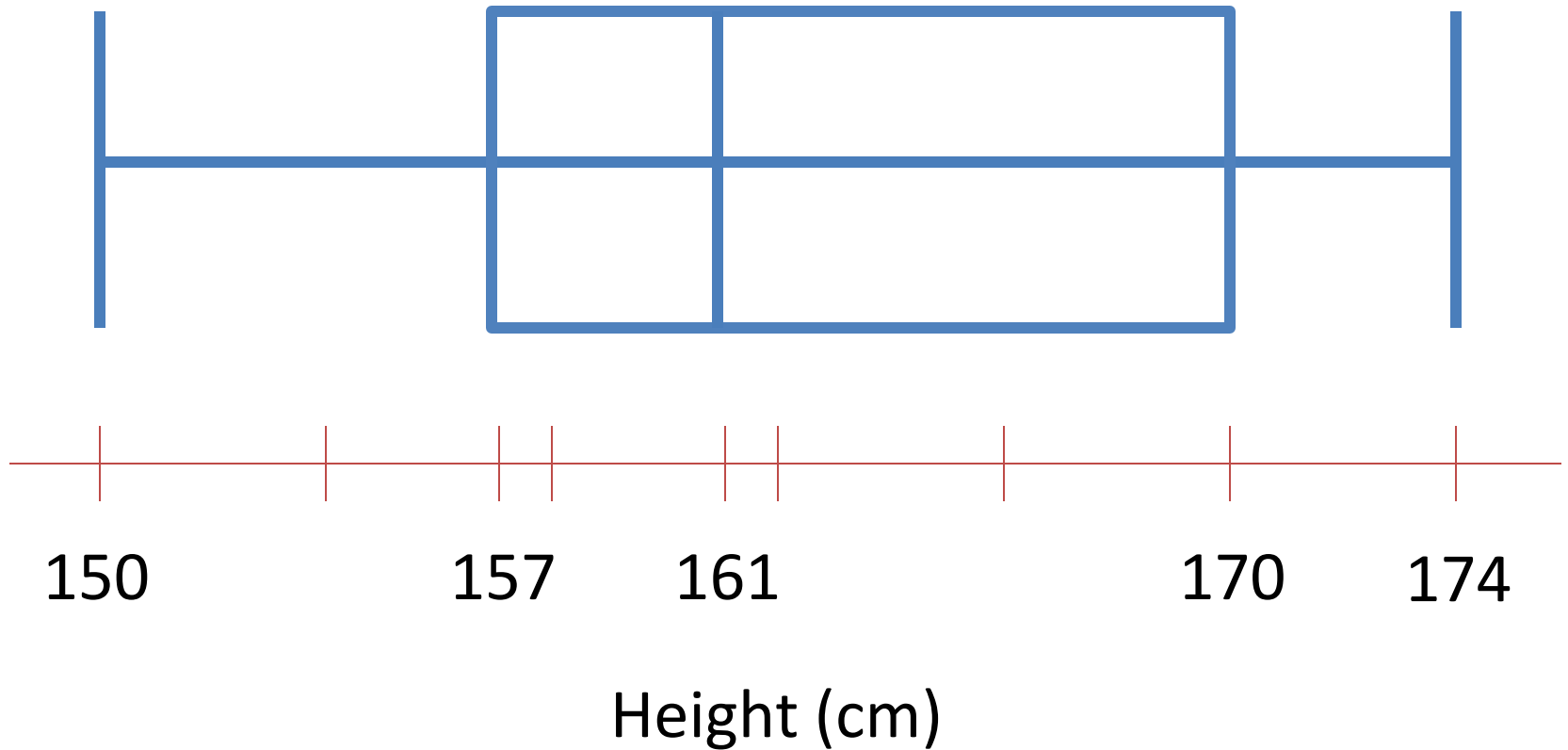
Boxplot

Perhatikan data berikut:

160 157 162 170 168 174 156 173 157 150

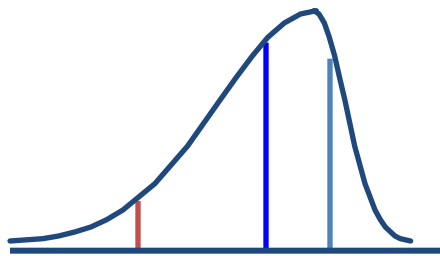
Buatlah *Boxplot*-nya?

Boxplot for the Height of Business Statistic's Student 2014



Distribution Shape and The Boxplot

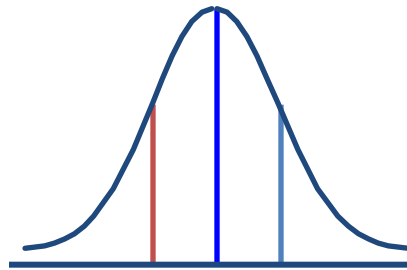
Left-Skewed



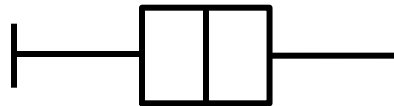
Q_1 Q_2 Q_3



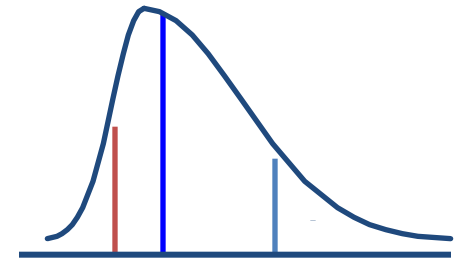
Symmetric



Q_1 Q_2 Q_3



Right-Skewed

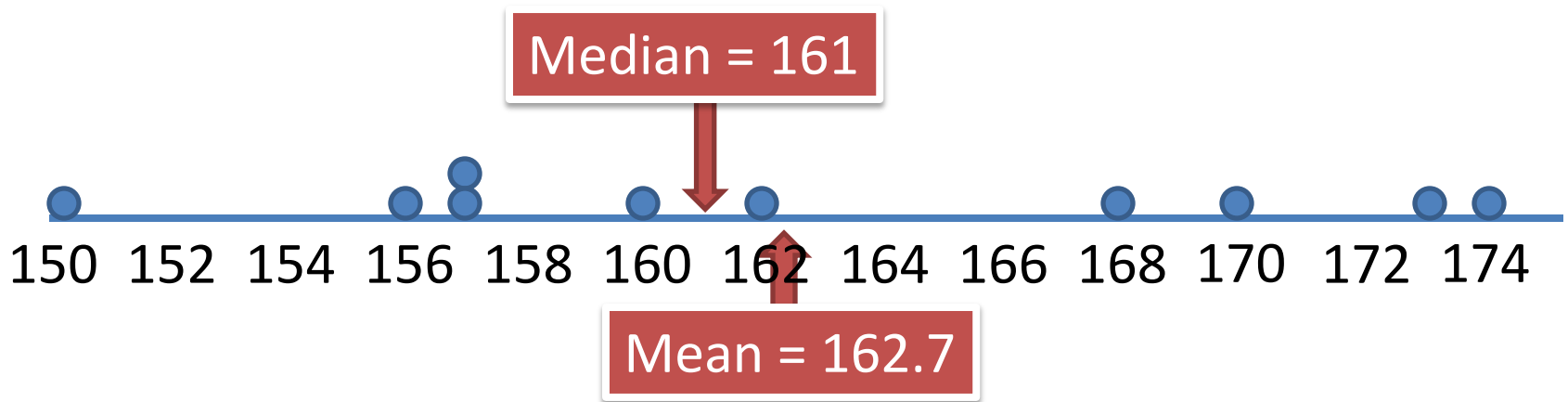


Q_1 Q_2 Q_3



MEASURE OF SKEWNESS

Karl Pearson's Measure of Skewness

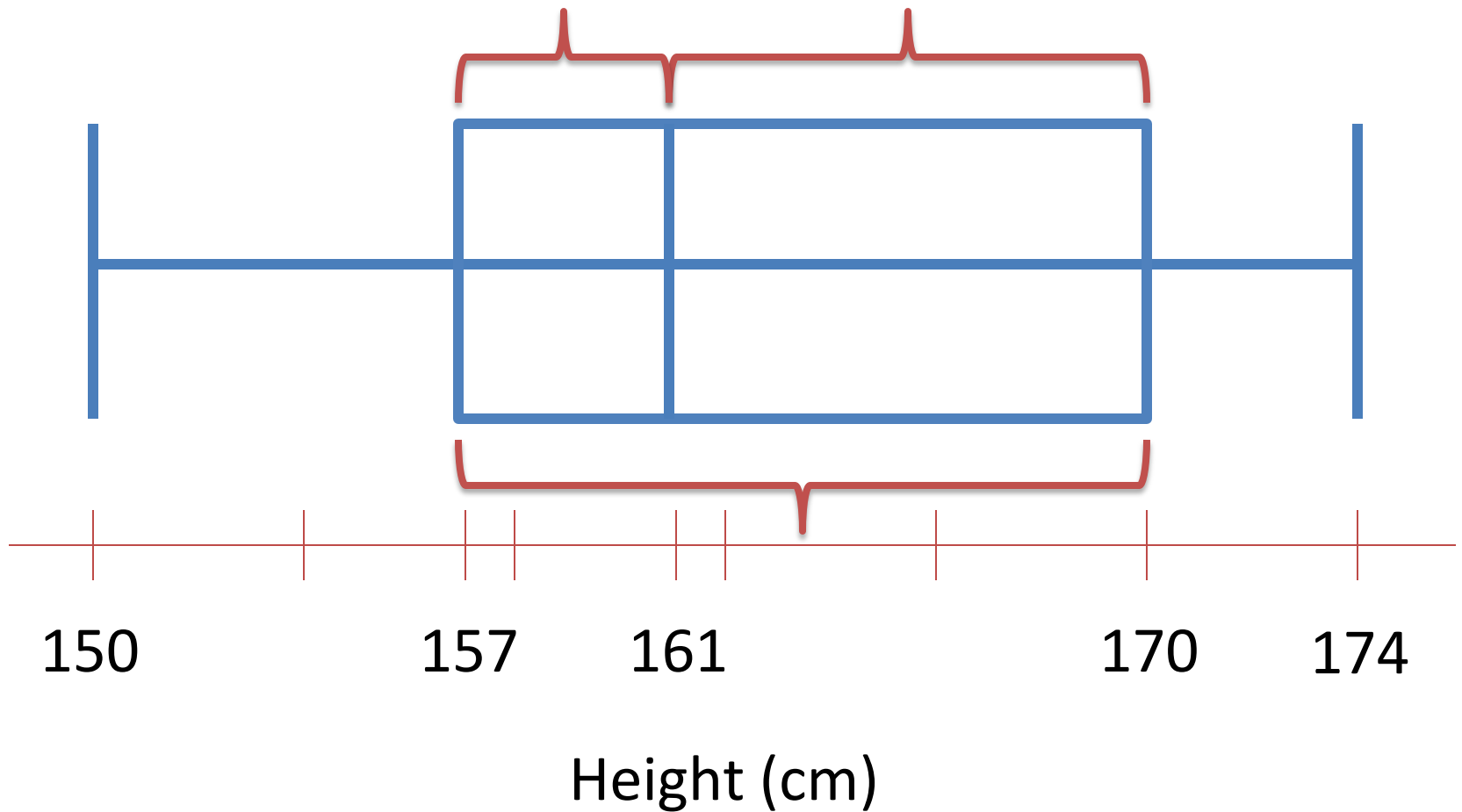


$$S_k = \frac{3(162.7 - 161)}{8.125} = 0.63$$

Karl Pearson's Measure of Skewness

$$S_k = \frac{3(\bar{X} - \textit{Median})}{S}$$

Bowley's Formula for Measuring Skewness



Bowley's Formula for Measuring Skewness

$$S_k = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

EXERCISE

3.10

Data berikut adalah data pengeluaran yang dilakukan oleh sampel sembilan orang konsumen untuk makan siang di sebuah restoran cepat saji (dalam \$):

4,20 5,03 5,86 6,45 7,38 7,54 8,46 8,47 9,87

- Hitung variansi, simpangan baku (standard deviation), rentang, dan koefisien variasi.
- Apakah datanya menceng? Jika ya bagaimana?
- Berdasarkan hasil perhitungan diatas, apa kesimpulan yang bisa anda ambil mengenai jumlah pengeluaran konsumen untuk makan siang tersebut?

HOMEWORK

3.62

Salah satu produk perusahaan asuransi adalah asuransi jiwa. Proses approval meliputi review formulir aplikasi hingga diterbitkan dan dikirimnya polis ke pelanggan. Kemampuan perusahaan untuk mengirim polis asuransi secepat mungkin merupakan hal yang penting bagi pelayanan. Dalam satu bulan kebelakang dipilih sampel acak 14 polis yang disetujui. Berikut adalah waktu proses approval ke 14 polis tersebut.

3.62

73 19 16 64 28 28 31 90 60 56 31 56 22 18

- a. Hitung rata-rata, median, kuartil pertama dan kuartil ketiga.
- b. Hitung rentang, rentang antarkuartil, variansi, deviasi standar/simpangan baku.
- c. Apakah datanya menceng? Jika iya, bagaimana?
- d. Apa yang akan anda katakan kepada konsumen yang ingin membeli produk asuransi ini jika mereka bertanya tentang waktu yang dibutuhkan untuk proses approval?

3.-

Berikut data biaya listrik pada bulan juli 2010 dari sampel acak 20 apartemen dengan satu kamar tidur di kota besar:

Biaya Listrik	Frekuensi
80 – kurang dari 100	3
100 – kurang dari 120	3
120 – kurang dari 140	2
140 – kurang dari 160	4
160 – kurang dari 180	4
180 – kurang dari 200	2
200 – kurang dari 220	2

3.-

- a. Hitung rata-rata, median, modus.
- b. Hitung simpangan baku.

3.22

Tabel berikut merupakan data kalori dan total lemak (dalam gram per sajian) dari sampel 12 veggie burger

Kalori	Total lemak
110	3.5
110	4.5
90	3.0
90	2.5
120	6.0
130	6.0
120	3.0
100	3.5
140	5.0
70	0.5
100	1.5
120	1.5

3.22

- a. Hitung rata-rata, median, modus, kuartil pertama dan kuartil ketiga.
- b. Hitung rentang, rentang antarkuartil, variansi, deviasi standar/simpangan baku, deviasi rata-rata.
- c. Apakah datanya menceng? Jika iya, bagaimana?
- d. Apa kesimpulan yang anda dapatkan mengenai kalori dan total lemak tersebut?



THANK YOU