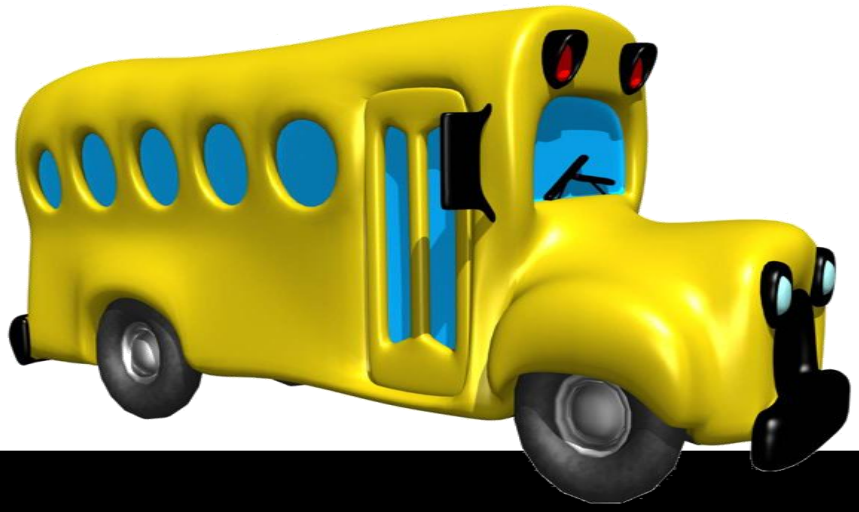
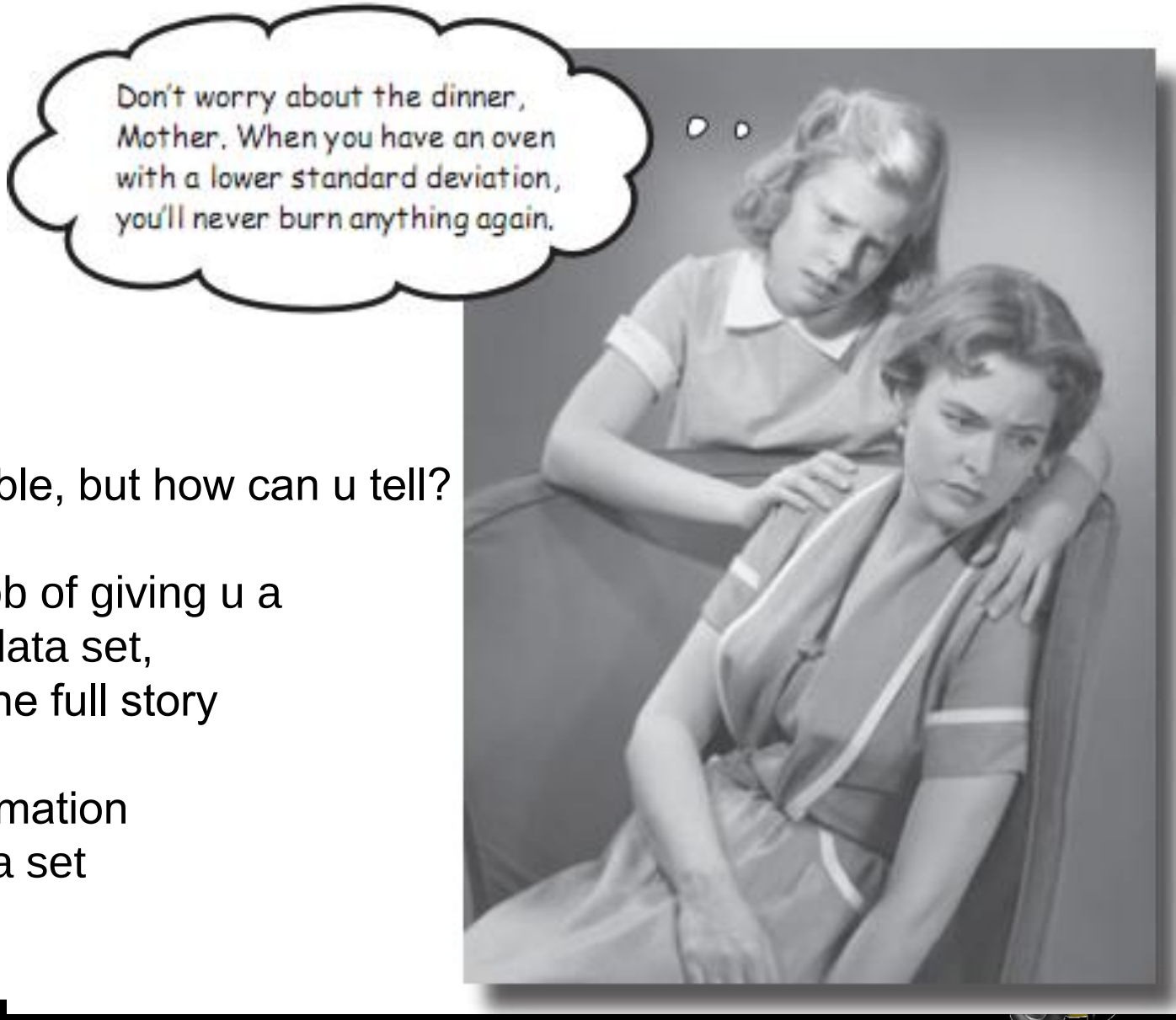
Bab 2
Bagian 2

Ukuran Sebaran





Not everything's reliable, but how can u tell?

Average do a great job of giving u a typical value in your data set,
but they don't tell u the full story

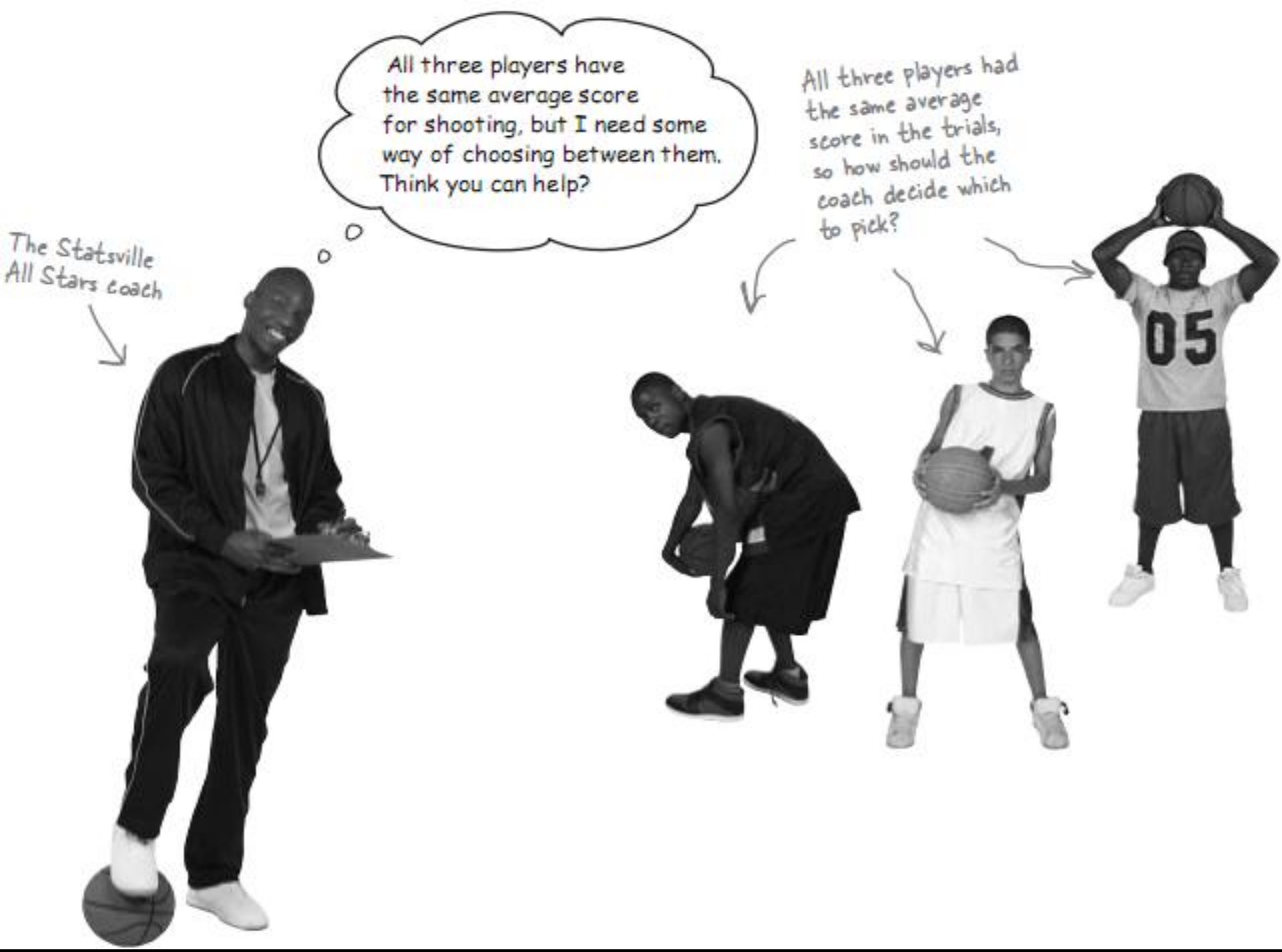
→are'nt enough information in summarizing a data set

Ukuran sebaran

- Variabilitas → ukuran kuantitatif yang menunjukkan seberapa jauh data dalam suatu distribusi menyebar

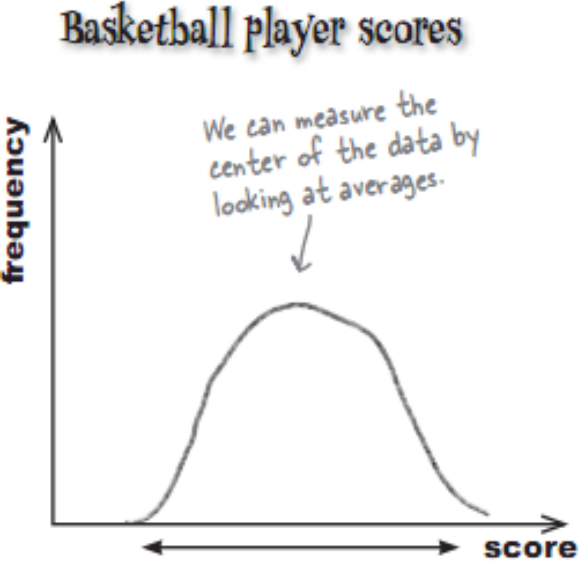
- Jika X sangat dekat dengan rata-rata → rata-rata menjadi ukuran yang baik

- Seberapa "jauh" suatu data dari rata-rata? (Misalkan rata-rata adalah "**best guess**")


All three players have the same average score for shooting, but I need some way of choosing between them. Think you can help?
The Statsville All Stars coach
All three players had the same average score in the trials, so how should the coach decide which to pick?
05

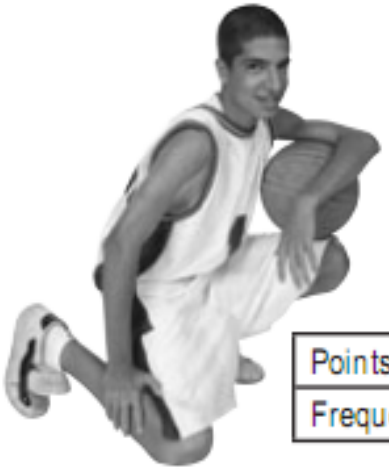
Misal:

data berisi tentang skor poin per-permainan



Points scored per game	7	8	9	10	11	12	13
Frequency	1	1	2	2	2	1	1

Here, frequency tells us the number of games where the player got each score. This player scored 9 points in 2 games, and 12 points in 1 game.



Points scored per game	7	9	10	11	13
Frequency	1	2	4	2	1

Points scored per game	3	6	7	10	11	13	30
Frequency	2	1	2	3	1	1	1

1. Rentang (Range)

Range is a way of measuring how spread out a set a value are.

- Untuk data tak berkelompok

  R = Nilai Maks-Nilai Min

- Untuk data berkelompok

  R = Nilai tengah kelas terakhir – Nilai tengah kelas pertama

  R = Batas atas kelas terakhir – Batas bawah kelas pertama

contoh

Kelas		Titik Tengah M	Frekuensi
BB	BA		
30	39	34.5	2
40	49	44.5	3
50	59		11
60	69		20
70	79		32
80	89		25
90	99		7

Rentang = 60
Rentang = 69

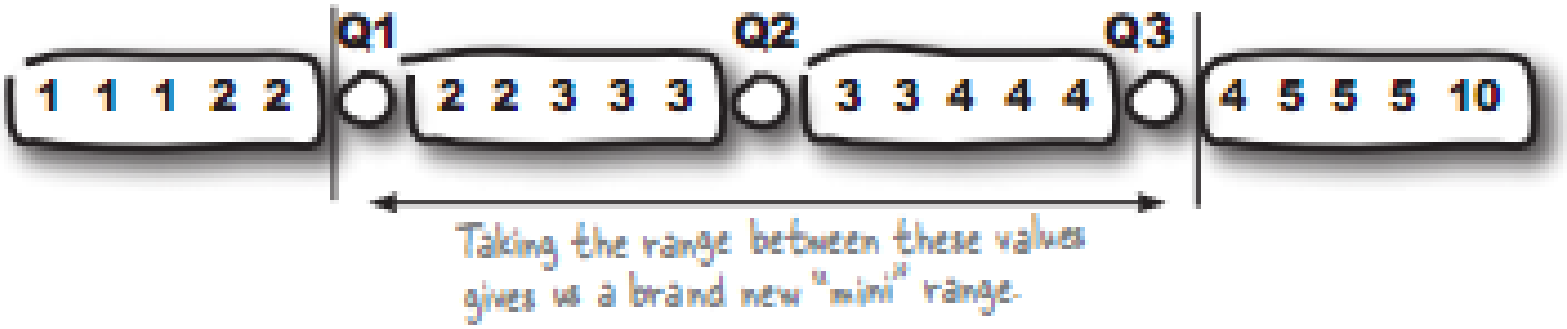
Range… cont

•Range → cara mudah mencari ukuran penyebaran suatu data namun “sering bukan yang terbaik”

- Jika data berisi outlier, range bisa jadi *misleading* sebagai ukuran penyebaran karena sensitif terhadap outlier

Masalah … dengan Range …

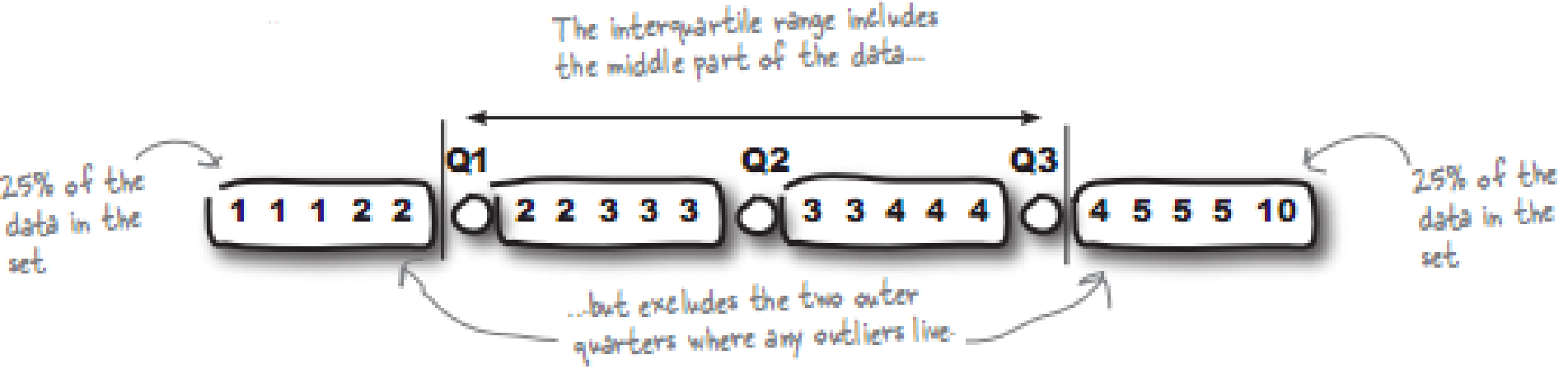
- Mendeskripsikan lebar data
- Sensitif terhadap outlier → alternatif : *mini range ; quartile*



2. IQR

Alternatif lain …→IQR

- Lot less sensitive then range



IQR = Q3-Q1

Pagar Dalam & Pagar Luar

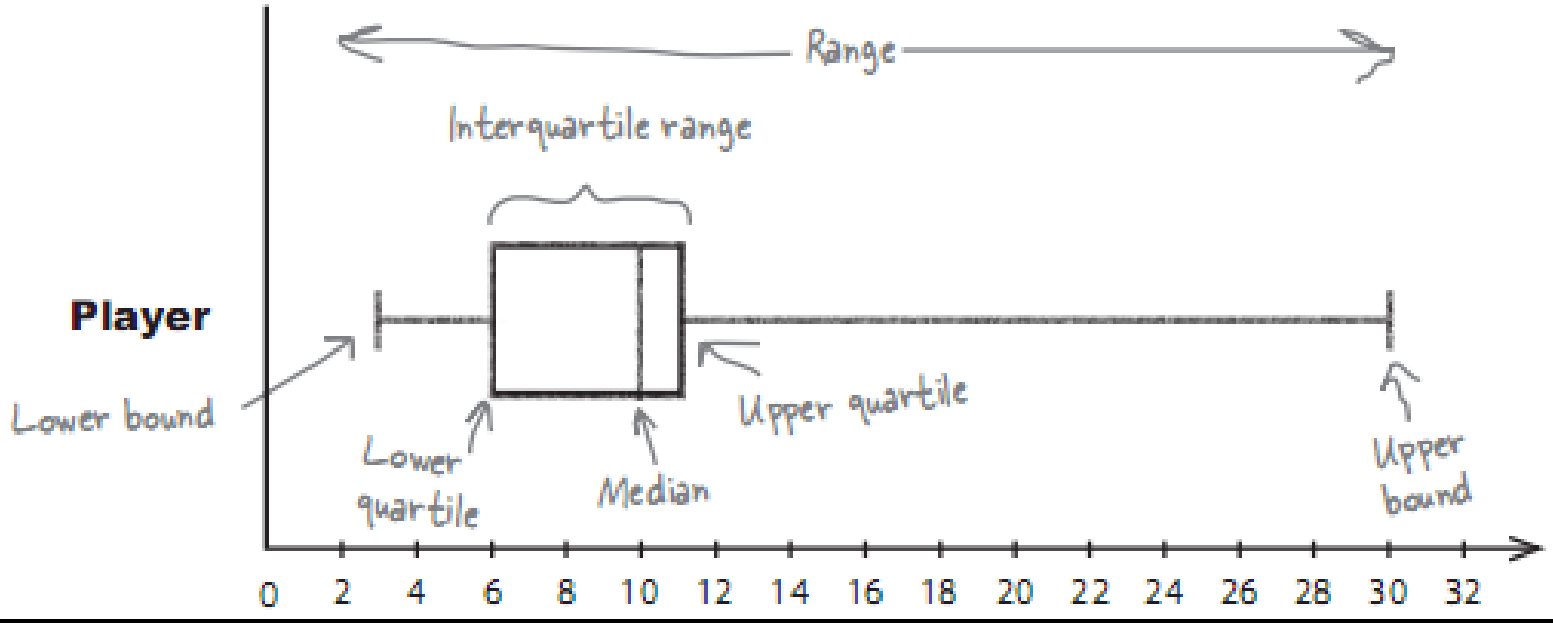
$$IF = Q_1 - 1.5(IQR) \quad \& \quad Q_3 + 1.5(IQR)$$

$$OF = Q_1 - 3(IQR) \quad \& \quad Q_3 + 3(IQR)$$

BOX PLOT

Here's a reminder of the data.

3 3 6 7 7 10 10 10 11 13 30



Work out the mean, lower bound, upper bound, and range for the following sets of data, and sketch the charts. Are values dispersed in the same way? Does the range help us describe these differences?

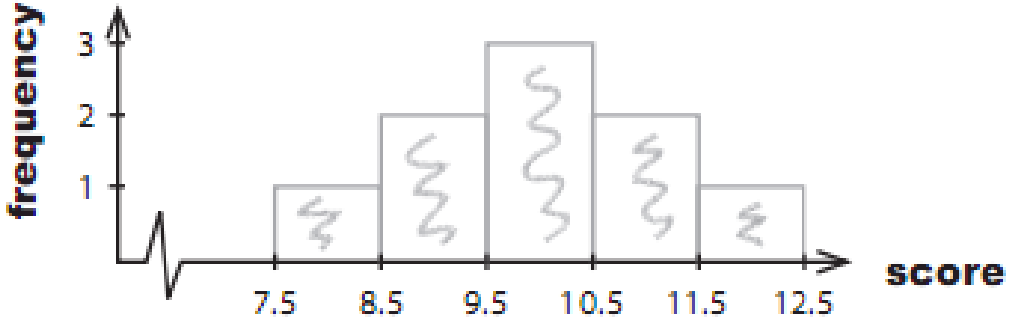
Score	8	9	10	11	12
Frequency	1	2	3	2	1

Score	8	9	10	11	12
Frequency	1	0	8	0	1

Hitung rata-rata, range dari data pada Tabel di atas.
Apakah nilai-nilai data mempunyai simpangan yang sama?
Apakah Range dapat membantu kita dalam menggambarkan penyebaran ini?

Work out the mean, lower bound, upper bound, and range for the following sets of basketball scores, and sketch the charts. Are values dispersed in the same way? Does the range help us describe these differences?

Score	8	9	10	11	12
Frequency	1	2	3	2	1



$\mu = 10$

Lower bound = 8

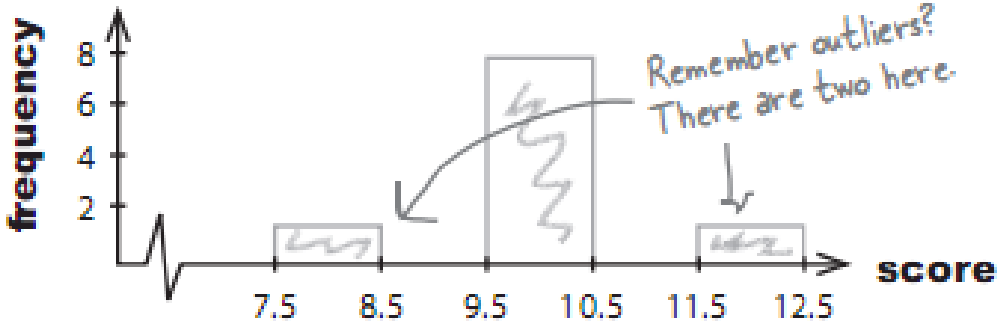
Upper bound = 12

Range = 12 − 8

= 4

Look, these results are the same even though the data's different.

Score	8	9	10	11	12
Frequency	1	0	8	0	1



$\mu = 10$

Lower bound = 8

Upper bound = 12

Range = 12 − 8

= 4

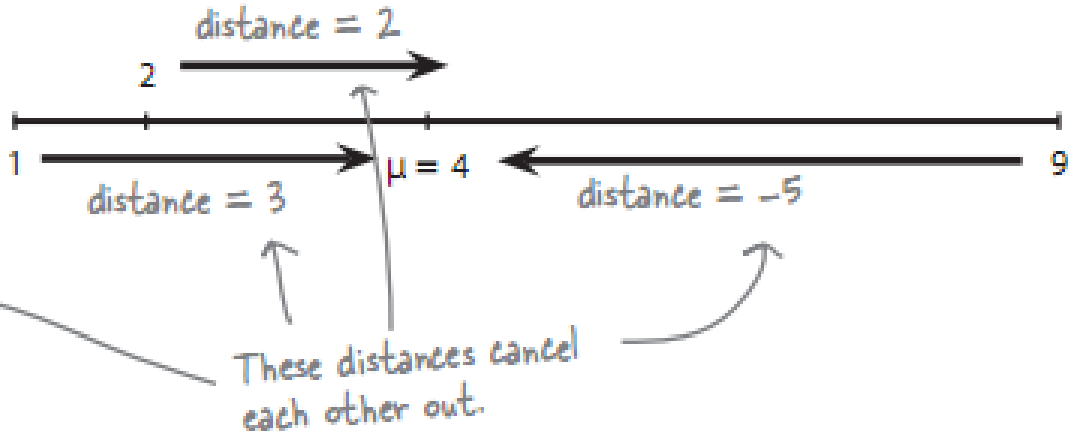
3. Variansi

Misal kita punya data 1, 2 dan 9 dengan rata-rata =4. Berapakah *average* jarak data tsb dari rata-ratanya ?

$$\text{Average distance} = \frac{(1 \text{ to } \mu) + (2 \text{ to } \mu) + (9 \text{ to } \mu)}{3}$$

$$= \frac{3 + 2 + (-5)}{3}$$

$$= 0$$



…Variansi

→ Deviasi dari satu observasi pengukuran thp rata-rata

Populasi

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

Sampel

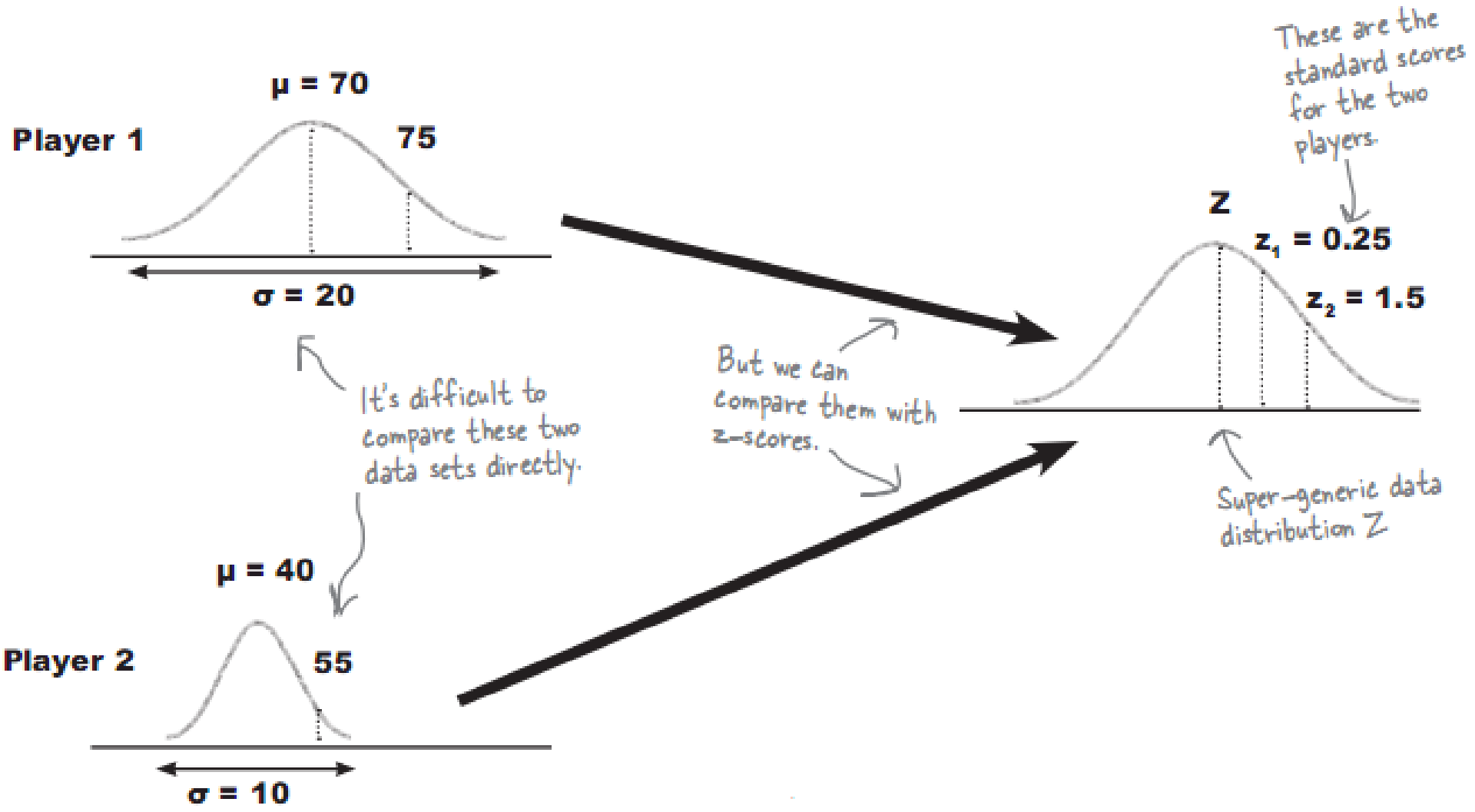
$$s^2 == \frac{\sum_{i=1}^{n}\left(X_i - \bar{x}\right)^2}{n-1} \Rightarrow s = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{x}\right)^2}{n-1}}$$

Nilai Standar (Baku)

Nilai Standar (baku) memberikan cara dalam perbandingan suatu nilai apabila nilai tersebut berasal dari himpunan atau distribusi yang sama.

$$z = \frac{x - \mu}{\sigma}$$

Ilustrasi



Which players better ?

Latihan

Seorang pelatih basket mempunyai record pertandingan 3 pemain. Rata-rata tiap pemain adalah 10. Jika pelatih juga menghitung ukuran simpangan pemain-pemain itu. Menurut Anda pemain mana yang paling reliabel untuk dapat dipilih?

Player 1

Score	7	9	10	11	13
Frequency	1	2	4	2	1

Player 2

Score	7	8	9	10	11	12	13
Frequency	1	1	2	2	2	1	1

Player 3

Score	3	6	7	10	11	13	30
Frequency	2	1	2	3	1	1	1

Do U have same conclusion when U work out with Range ?