

# Statistika Deskriptif\_1

Pertemuan 2 - Mata Kuliah Statistika

@harsawara

54	75	46	75	83	61	74	86	49	50	49	56	51	90	64
70	63	46	51	48	48	66	62	51	86	73	87	74	86	51
75	88	63	79	88	71	63	78	50	77	71	70	59	64	81
46	53	72	85	75	92	74	89	53	79	90	84	70	91	49
51	49	86	61	57	58	65	53	61	60	53	48	57	88	76
45	68	63	62	75	75	87	91	57	58	46	81	65	60	88
75	91	71	86	59	49	51	81	67	85	66	66	60	91	76
72	51	63	78	89	75	86	79	80	47	65	77	47	74	68
90	45	76	89	87	49	92	73	89	84	59	51	88	48	50
87	58	79	88	83	65	51	84	82	84	50	68	55	73	52

- informasi apa yang bisa kita peroleh dari data-data tersebut?
- kesimpulan seperti apa yang bisa dibuat?

# Beberapa Kajian dalam Statistika Deskriptif

## Tabel Distribusi

Frekuensi, Frekuensi Relatif, dll

1

## Ukuran Penyebaran

Range, Variansi, dll

4

## Grafik dan Diagram

Comparison, Relationships, dll

2

## Ukuran Lokasi

Quartil, Desil & Persentil

5

## Ukuran Pemusatan

Mean, Median dan Modus

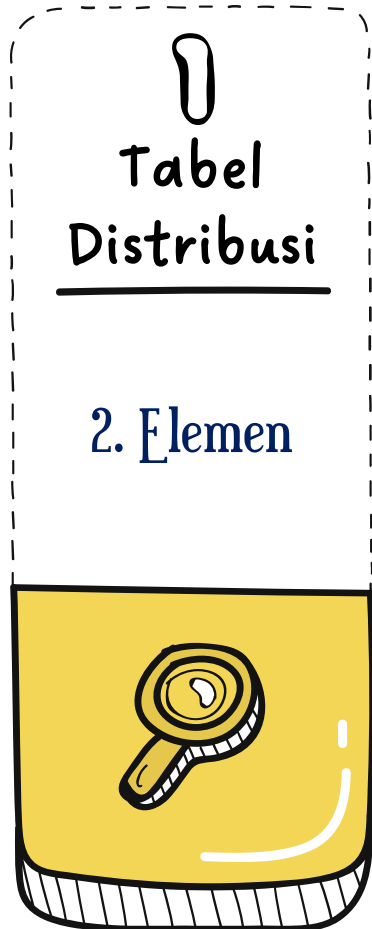
3



digunakan untuk menyajikan ringkasan  
data kategorik ataupun data numerik yang  
dikategorikan

○ untuk data kualitatif

○ untuk data kuantatif



Tabel Preferensi Jenis Pekerjaan dari 100 Alumni

Jenis Pekerjaan	Frekuensi
Perusahaan Swasta	44
Entrepreneur	23
PNS	16
Studi S2	17
Jumlah	100

Judul Tabel  
Caption Kolom  
Kategori

kolom frekuensi  
variabel  
frekuensi

Tabel Preferensi Penghasilan dari 100 Alumni

Penghasilan Mingguan (dalam Dollar)	Jumlah Karyawan (f)
301 – 400	9
401 – 500	16
501 – 600	33
601 – 700	22
701 – 800	14
801 – 900	6

kolom frekuensi  
variabel  
frekuensi kelas ke-1  
kelas ke-1  
batas atas kelas ke-4  
batas bawah kelas ke-4

76	81	81	84	79	79	77	84	82	82
85	79	79	72	75	81	82	80	85	77
74	80	83	83	73	78	82	84	75	83

- Aturan Sturges : jumlah kelas =  $1 + 3,332 \log n$
- Aturan Lind :  $n \sim 2^k$
- Aturan Kauro Ishikawa

Jumlah Data	Jumlah Kelas
< 50	5 - 7
50 - 100	6 - 10
100 - 250	7 - 12
> 250	10 - 25



Tinggi (inch)	f
72 - 74	3
75 - 77	5
78 - 80	7
81 - 83	10
84 - 86	5
	$\Sigma f = 30$

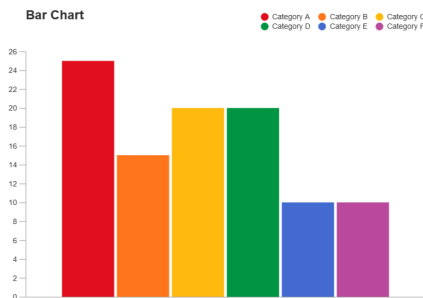
Jumlah Kelas : 5

$$\text{Lebar Kelas} : \frac{X_{\text{maks}} - X_{\text{min}}}{\text{Jumlah Kelas}} = 2,6 \sim 3$$

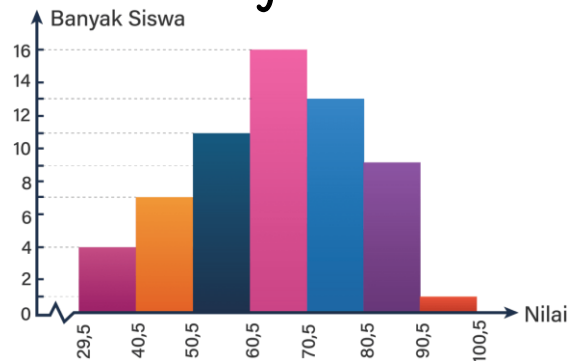
Batas Bawah Kelas-1:

ini bisa menggunakan nilai terendah dalam data atau lebih kecil

### BAR Chart



### Histogram

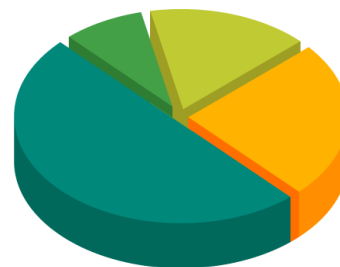


# Grafik & Diagram

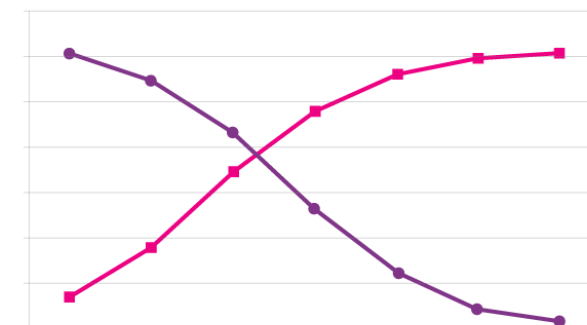
Tinggi (inch)	f	Tepi Kelas	Frekuensi Relatif	Frekuensi Relatif Persen	Frekuensi Kumulatif <	Frekuensi Kumulatif >
72 – 74	3	71,5 – 74,5	$3/30 = 0,100$	$3/30 \times 100\% = 10\%$	3	30
75 – 77	5	74,5– 77,5	$5/30 = 0,167$	$5/30 \times 100\% = 16,7\%$	8	27
78 – 80	7	77,5 – 80,5	$7/30 = 0,233$	$7/30 \times 100\% = 23,3\%$	15	22
81 – 83	10	80,5 – 83,5	$10/30 = 0,333$	$10/30 \times 100\% = 33,3\%$	25	15
84 – 86	5	83,5 – 86,5	$5/30 = 0,167$	$5/30 \times 100\% = 16,7\%$	30	5
	$\Sigma f = 30$		$\Sigma = 1,000$	$\Sigma = 100\%$		

titik tengah dari limit atas suatu kelas dan limit bawah dari kelas berikutnya

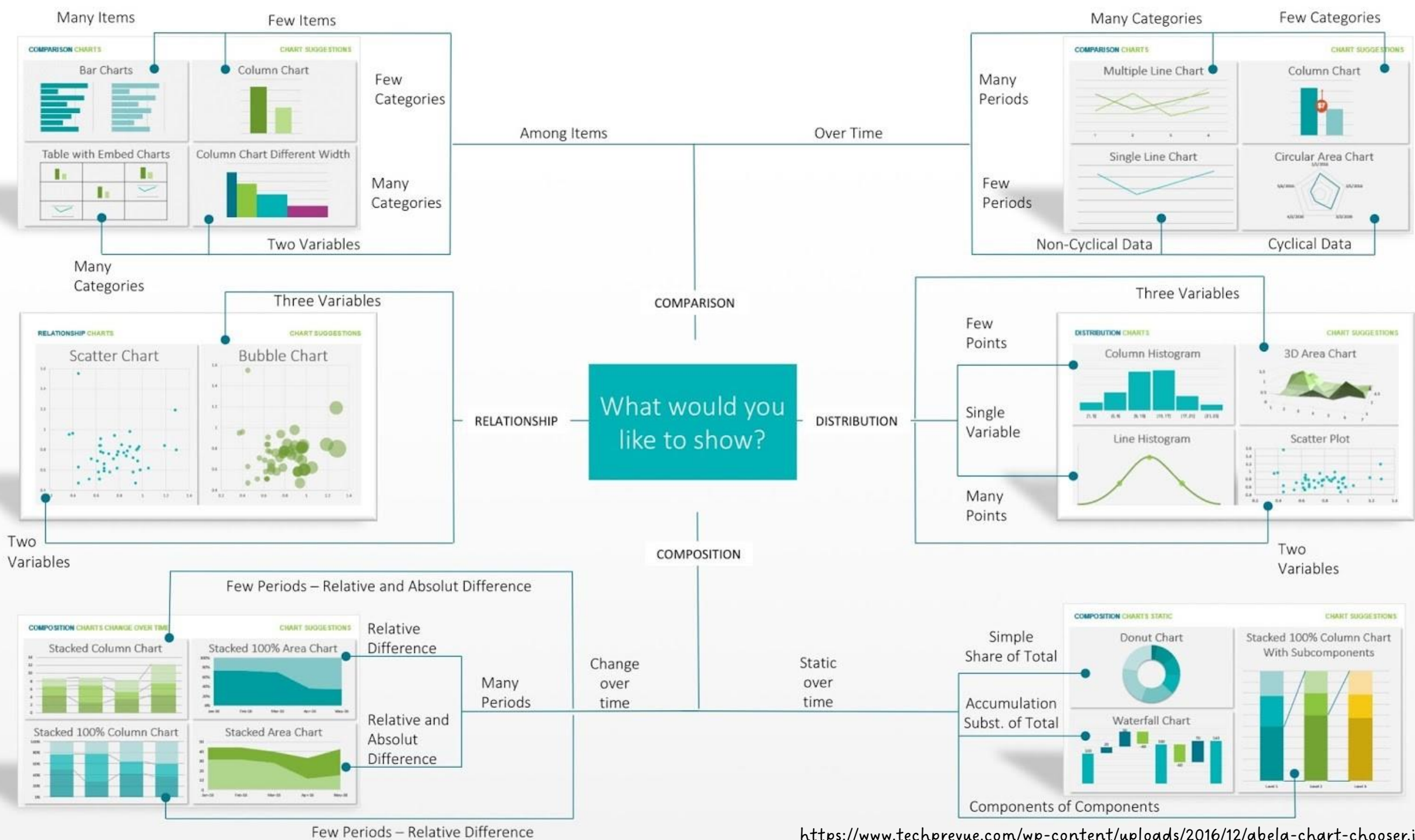
lebar kelas = tepi atas - tepi bawah



### PIE Chart



















### Ogive





# Beberapa bentuk grafik lainnya

- diagram gambar/lambang

Varities of Apples in a food store	
Red Delicious	  
Golden Delicious	  
Red Rome	   
McIntosh	 
Jonathan	   

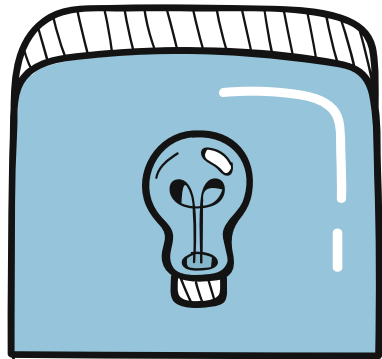
 = 10 apples     = 5 apples

- diagram dahan-daun

Tes Kedua		Tes Pertama
6	4	4 7
4 3 2 2	5	6 9
9 7 6 6 3 3 3 1	6	1 3 4 5 6 8
9 9 6 6 5 5 2 1 0	7	0 1 1 1 1 2 3 3 3 4 5 5 6 6 6 9
9 5 4 4 2 2 1 0	8	3 4 4 5 5 9
9 9 9 9 8 2	9	1 4 5 7

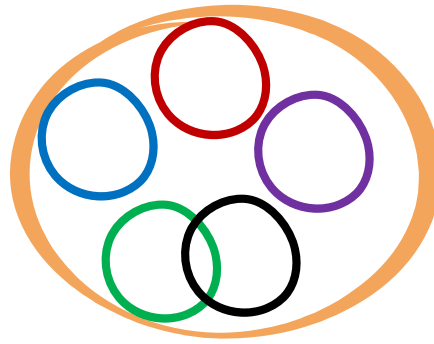
nilai **pencilan** atau **outlier**, yaitu nilai yang sangat kecil atau sangat besar atau jauh dari nilai observasi lainnya

titik tempat di mana nilai-nilai suatu gugus data cenderung mengelompok, menunjukkan titik tengah suatu histogram atau kurva distribusi frekuensi



Ukuran  
Pemusatan

Rerata  
Hitung  
(Mean)



sampel

Nilai  $\bar{x}$  (mean sampel) memungkinkan bervariasi, karena diambil dari sampel yang berbeda, dan bergantung pada nilai observasi tiap sampel

populasi

$$\bar{x} = \frac{\sum x}{n} \quad \text{atau} \quad \mu = \frac{\sum x}{N}$$

Kelemahan:

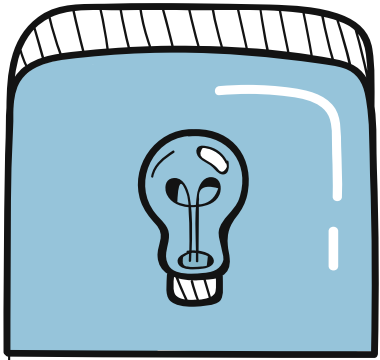
rentan terhadap keberadaan nilai pencilan

data  
pencilan

Negara Bagian	Washington	Oregon	Alaska	Hawaii	California
Populasi (ribuan)	5.136	2.977	587	1.160	30.867

Mean tanpa California :  $(5.136 + 2.977 + 587 + 1.160) / 4 = 2.465$

Mean dengan California :  $(5.136 + 2.977 + 587 + 1.160 + 30.867) / 5 = 8.145,4$



Ukuran  
Pemusatan

Rerata  
Hitung  
(Mean)

Data tunggal (ungrouped data)

$$\bar{x} = \frac{\sum x}{n} \quad \text{atau} \quad \mu = \frac{\sum x}{N}$$

Data berkelompok (grouped data)

$$\bar{x} = \frac{\sum m.f}{\sum f} \quad \text{atau} \quad \mu = \frac{\sum m.f}{\sum f}$$

76	81	81	84	79	79	77	84	82	82
85	79	79	72	75	81	82	80	85	77
74	80	83	83	73	78	82	84	75	83

$$\bar{x} = \frac{\sum x}{n} = \frac{2319}{30} = 77.3$$

Tinggi (inch)	f	Nilai tengah (m)	m.f
72 - 74	3	73	219
75 - 77	5	76	380
78 - 80	7	79	553
81 - 83	10	82	820
84 - 86	5	85	425
	$\Sigma f = 30$		$\Sigma m.f = 2397$

$$\bar{x} = \frac{\sum m.f}{\sum f} = \frac{2397}{30} = 79.9$$

# Statistika Deskriptif

#part2

nilai yang terletak pada **tengah** suatu **data** di mana data tersebut **telah diurutkan**



$$Me = X_{\frac{n+1}{2}}$$

Data  
berjumlah  
ganjil

$$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$$

Data  
berjumlah  
genap

Data tunggal (ungrouped data)

76	81	81	84	79	79	77	84	82	82
----	----	----	----	----	----	----	----	----	----

76	77	79	79	81	81	82	82	84	84
----	----	----	----	----	----	----	----	----	----

$X_{\frac{n}{2}}$   $X_{\frac{n}{2}+1}$

Data berkelompok (grouped data)

$$Me = B_m + i \left( \frac{\frac{n}{2} - f_{km}}{f_m} \right)$$

di mana :

$B_m$  = tepi bawah kelas median     $f_{km}$  = frekuensi kumulatif sebelum median

$i$  = interval kelas

$f_m$  = frekuensi pada kelas median

$n$  = ukuran sampel data

Tinggi (inch)	f
72 – 74	3
75 – 77	5
78 – 80	7
81 – 83	10
84 – 86	5
	$\Sigma f = 30$

Karena data berjumlah genap, maka data median terletak diantara data ke-15 dan data ke-16

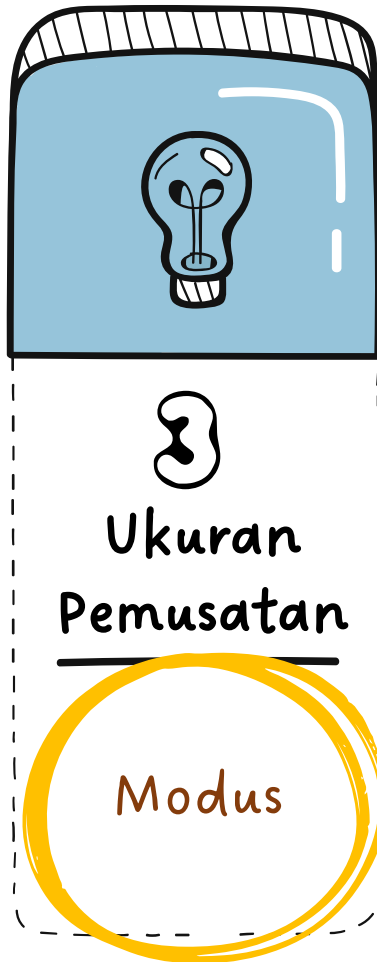
Dalam table, data ke-15 terletak pada kelas ke-3, sementara data ke-16 terletak pada kelas ke-4

Maka perhitungan median-nya adalah

$$Me = B_m + i \left( \frac{\frac{n}{2} - f_{km}}{f_m} \right) = 80.5 + 3 \left( \frac{\frac{30}{2} - 15}{10} \right) = 80.5$$

Data yang hanya memiliki 1 modus disebut **unimodal**; 2 modus dengan frekuensi sama disebut **bimodal** dan lebih dari 2 modus disebut **multimodal**

nilai yang memiliki **frekuensi tertinggi** dalam suatu gugus data



Data tunggal (ungrouped data)

Data berkelompok (grouped data)

$$Mo = B_m + i \left( \frac{d_1}{d_1 + d_2} \right)$$

di mana :

$B_m$  = tepi bawah kelas median

$i$  = interval kelas

$d_1$  = frekuensi kelas modus - frekuensi sebelum kelas modus

$d_2$  = frekuensi kelas modus - frekuensi setelah kelas modus

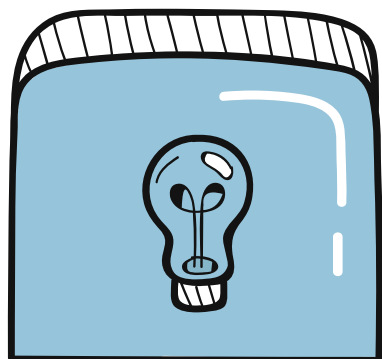
Tinggi (inch)	f
72 – 74	3
75 – 77	5
78 – 80	7
81 – 83	10
84 – 86	5
	$\Sigma f = 30$

Data tersaji hanya memiliki 1 kelas modus, maka data modus terletak pada kelas ke-4

Maka perhitungan modus-nya adalah

$$Mo = B_m + i \left( \frac{d_1}{d_1 + d_2} \right) = 80.5 + 3 \left( \frac{3}{3 + 5} \right) = 80.5 + 1,125 = 81.625$$





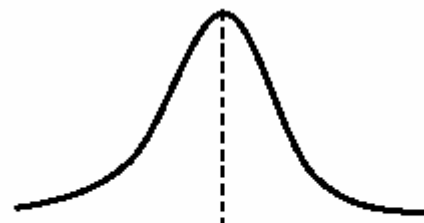
## Ukuran Pemusatan

Hubungan  
antara Mean,  
Median dan  
Modus



mode median mean

Nilai mean akan sangat dipengaruhi oleh pencilan di ekor sebelah kanan (**pencilan mayor**)



mean  
median  
mode

Untuk suatu histogram yang **simetris**, dan kurva frekuensi dengan sebuah puncak, **nilai mean, media dan modus adalah sama**



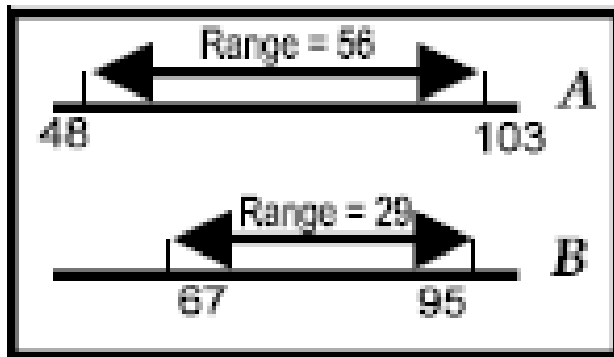
mean median mode

Nilai mean akan sangat dipengaruhi oleh pencilan di ekor sebelah kiri (**pencilan minor**)

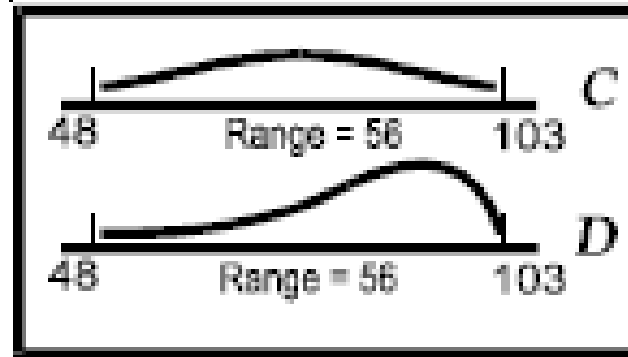
# 0. Range

$$\text{Range} = X_{\text{maks}} - X_{\text{min}}$$

- hanya ditentukan oleh 2 data (nilai yang lain dalam data diabaikan)
- Sepertihalnya mean, nilai range **dipengaruhi** oleh adanya **'outlier/pencilan'**



skor **Kelas B** memiliki **variabilitas yang lebih rendah** (lebih berkerumun rapat) daripada skor Kelas A



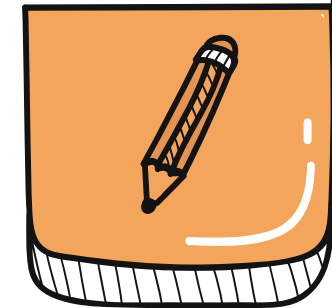
Kelas C dan D memiliki rentang yang sama, **tetapi memiliki penyebaran skor yang berbeda**



Ukuran  
Penyebaran

Kegunaan

digunakan  
untuk  
mengukur  
ketersebaran  
suatu data



**mengukur penyimpangan**  
setiap skor **dari rata-rata**  
- dan kemudian  
menghitung rata-rata  
semua penyimpangan

# 2. Standard Deviation

digunakan untuk menunjukkan seberapa dekat nilai-nilai suatu data dengan nilai rerata

$$\begin{aligned}
 s_x &= \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \\
 &= \sqrt{\frac{576,7}{29}} \\
 &= \sqrt{19,89} \\
 s_x &= 4,46
 \end{aligned}$$

76	81	81	84	79	79	77	84	82	82
85	79	79	72	75	81	82	80	85	77
74	80	83	83	73	78	82	84	75	83

$$\bar{x} = \frac{\sum x}{n} = \frac{2319}{30} = 77,3$$

$(76 - 77,3)^2$	$(85 - 77,3)^2$	$(74 - 77,3)^2$
$(81 - 77,3)^2$	$(79 - 77,3)^2$	$(80 - 77,3)^2$
$(81 - 77,3)^2$	$(79 - 77,3)^2$	$(83 - 77,3)^2$
$(84 - 77,3)^2$	$(72 - 77,3)^2$	$(83 - 77,3)^2$
$(79 - 77,3)^2$	$(75 - 77,3)^2$	$(73 - 77,3)^2$
$(79 - 77,3)^2$	$(81 - 77,3)^2$	$(78 - 77,3)^2$
$(77 - 77,3)^2$	$(82 - 77,3)^2$	$(82 - 77,3)^2$
$(84 - 77,3)^2$	$(80 - 77,3)^2$	$(84 - 77,3)^2$
$(82 - 77,3)^2$	$(85 - 77,3)^2$	$(75 - 77,3)^2$
$(82 - 77,3)^2$	$(77 - 77,3)^2$	$(83 - 77,3)^2$

$(x_i - \bar{x})$

1,69	59,29	10,89
13,69	2,89	7,29
13,69	2,89	32,49
44,89	28,09	32,49
2,89	5,29	18,49
2,89	13,69	0,49
0,09	22,09	22,09
44,89	7,29	44,89
22,09	59,29	5,29
22,09	0,09	32,49

$(x_i - \bar{x})^2$

## 3. Variansi

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



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



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

Data berkelompok (grouped data)

Tinggi (inch)	f	Nilai tengah (m)	$f_i(m_i - \bar{x})^2$
72 - 74	3	73	142,83
75 - 77	5	76	76,05
78 - 80	7	79	5,67
81 - 83	10	82	44,1
84 - 86	5	85	130,05
	$\Sigma f = 30$		$\Sigma_{i=1}^k = 398,7$

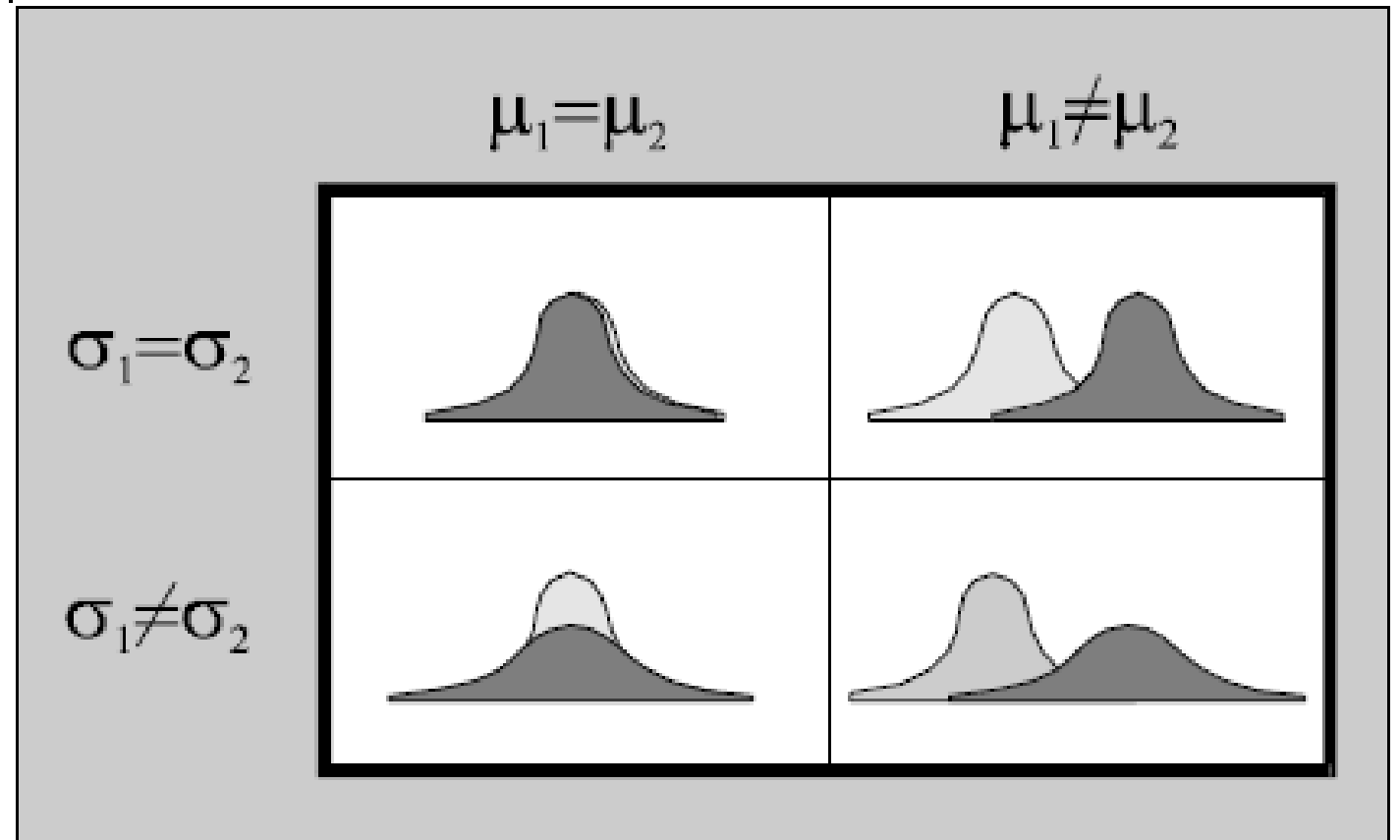
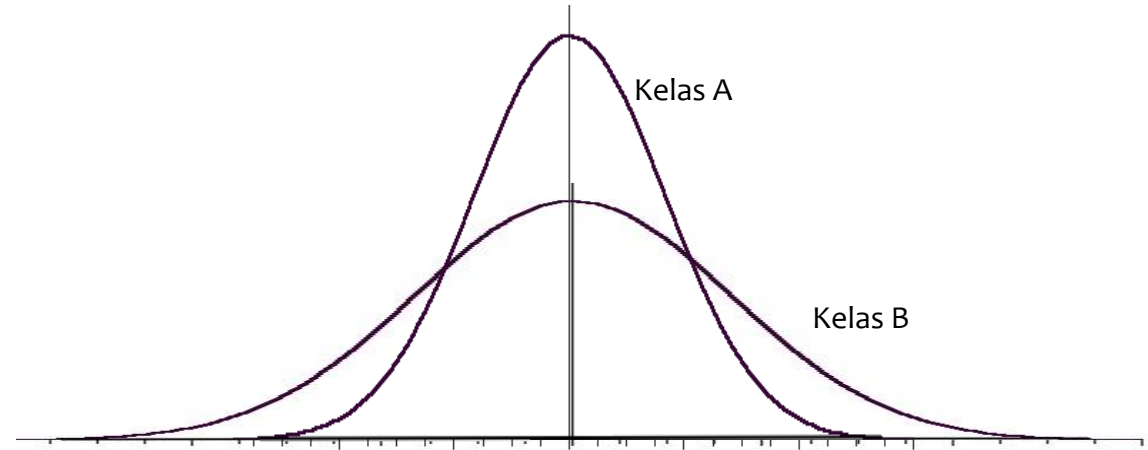
$$s_x = \sqrt{\frac{\sum_{i=1}^k f_i(m_i - \bar{x})^2}{\sum_{i=1}^k f_i}}$$

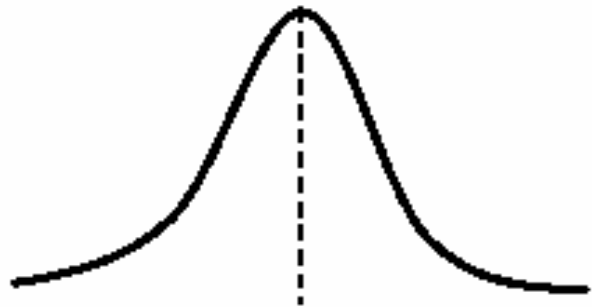
$$= \sqrt{\frac{398,7}{5}}$$

$$s_x = 8,93$$

$$\bar{x} = \frac{\sum m \cdot f}{\sum f} = \frac{2397}{30} = 79.9$$

Apa sebenarnya yang dapat dimanfaatkan dari simpangan baku dan variansi





mean  
median  
mode

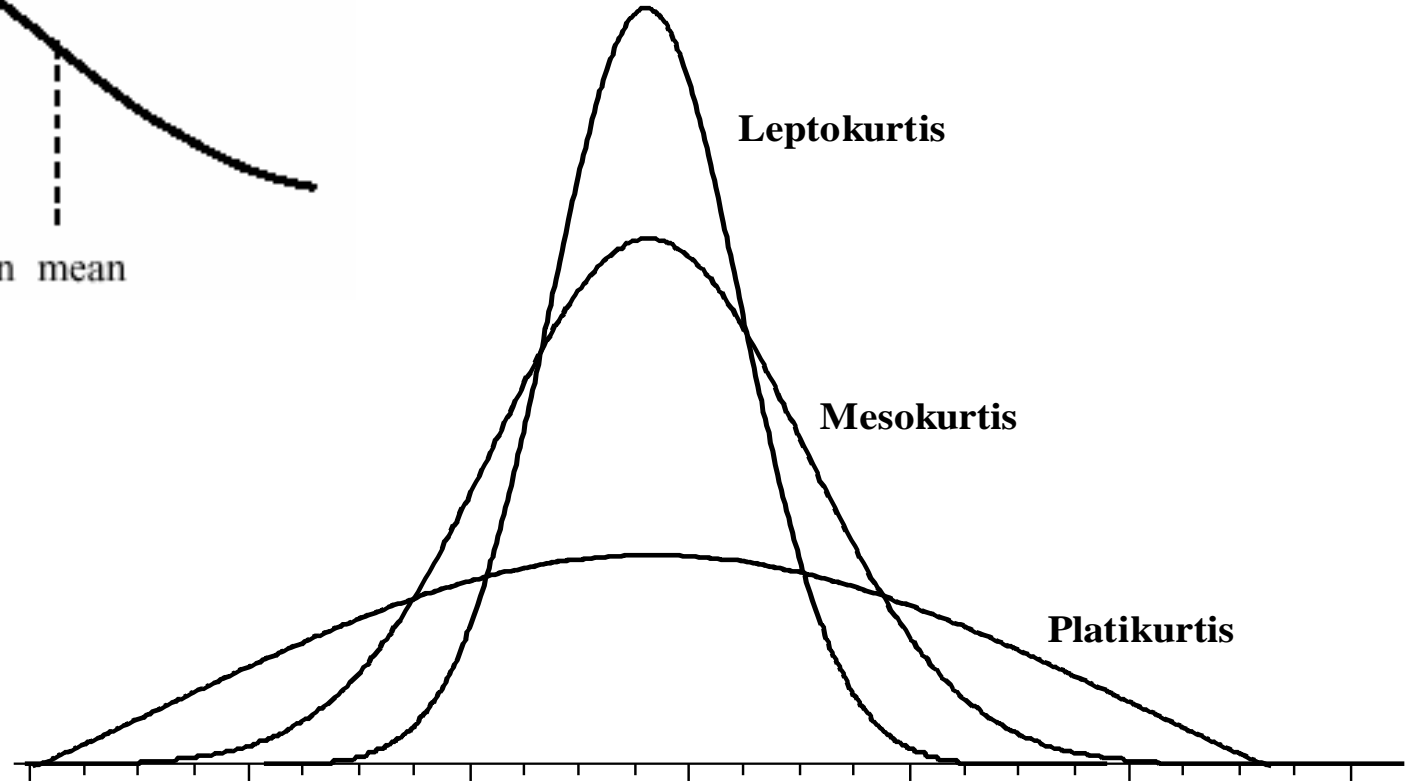


mode median mean



mean median mode

Kemiringan  
(Skewness)



Kelancipan  
(Kurtosis)

## Kuartil

$$Q_j = B_q + i \left( \frac{k \frac{n}{4} - f_{k_q}}{f_q} \right)$$

## Desil

$$D_j = B_d + i \left( \frac{k \frac{n}{10} - f_{k_d}}{f_d} \right)$$

## Persentil

$$P_j = B_p + i \left( \frac{k \frac{n}{100} - f_{k_p}}{f_p} \right)$$

di mana :

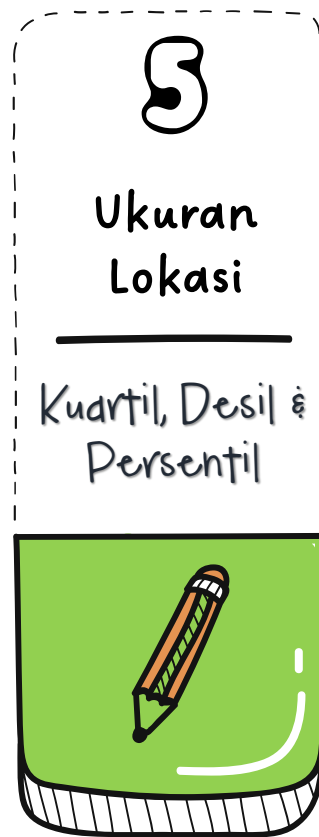
$B_m$  = tepi bawah kelas Kuartil/Desil/Persentil yang diinginkan

$i$  = interval kelas

$n$  = ukuran sampel data

$f_{km}$  = frekuensi kumulatif sebelum kelas Kuartil/Desil/Persentil yang diinginkan

$f_m$  = frekuensi pada kelas Kuartil/Desil/Persentil yang diinginkan

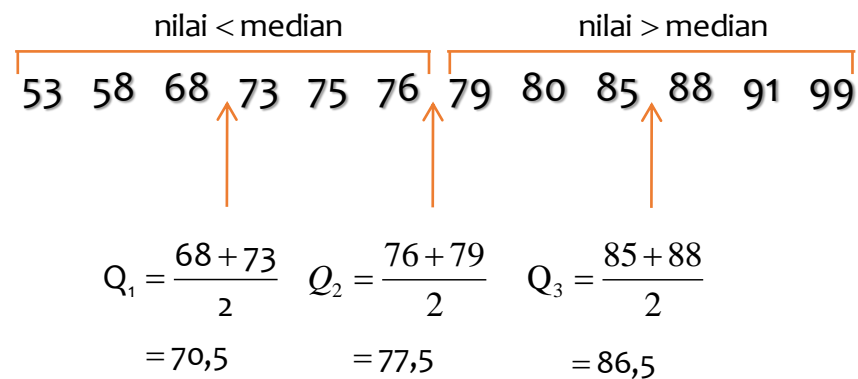


## Contoh

75 80 68 53 99 58 76 73 85 88 91 79

Tentukan nilai ketiga kuartil dan tentukan pula posisi 88 dalam hubungannya dengan kuartil!

Pertama, urutkan data dari yang terkecil menuju ke yang terbesar



$Q_1 = 70,5$ , menyatakan bahwa ±25% nilai pada sampel mendapat nilai < 70,5

nilai < median      nilai > median

53 58 68 73 75 76 79 80 85 88 91 99

$$Q_1 = \frac{68+73}{2} = 70,5$$

$$Q_2 = \frac{76+79}{2} = 77,5$$

$$Q_3 = \frac{85+88}{2} = 86,5$$

$$D_7 = \frac{7 \cdot (n+1)}{10} = \dots \quad 85.3$$

±70% nilai pada sampel  
mendapat nilai < 85,3

$$P_{15} = \frac{15 \cdot (n+1)}{100} = \dots \quad 57.75$$

±15% nilai pada sampel  
mendapat nilai < 57,75

$$Q_1 = \frac{1 \cdot (n+1)}{4}$$

$$= \frac{1 \cdot (12+1)}{4} = 3,25$$

○ diterjemahkan sebagai :

$$X_3 + 0,25(X_4 - X_3)$$

$$\text{sehingga nilai } Q_1 = 68 + 0,25(5) = 69.25$$

Q1 = 69,25; menyatakan bahwa  
±25% nilai pada sampel mendapat  
nilai < 69,25



selesai

---

