# BAB 3:
# NILAI RINGKASAN DATA

Penyajian data dalam bentuk tabel dan grafik memberikan kemudahan bagi kita untuk menggambarkan data dan membuat kesimpulan terhadap sifat data. Namun tabel dan grafik belum bisa secara langsung memberikan informasi ringkas dan baik akan sifat pokok dari data tersebut. Ukuran data, yang meliputi di antaranya nilai rata-rata, median, modus, standar dan deviasi, akan memberikan kepada kita satu informasi penting dan sangat bermakna dengan cepat ringkas.

Nilai ringkasan data adalah sebuah bilangan yang menggambarkan profil (sifat-sifat) data. Terdapat tiga karakteristik atau sifat penting dari data numerik, yaitu:

1. Ukuran pemusatan (tendensi sentral)
2. Penyebaran (variasi), dan
3. Bentuk distribusi (shape of distribustion)

## UKURAN TENDENSI SENTRAL

## a. RATA-RATA HITUNG (*MEAN ARITMETIK*)

Ukuran yang sering disebut dengan istilah "rata-rata" ini, dicari dengan perhitungan (jumlah nilai data) dibagi oleh (banyaknya observasi). Mengingat gugus data yang diamati bisa diperoleh dari populasi atau dari sampel, maka dibedakan antara rata-rata populasi dengan rata-rata sampel. Rata-rata populasi dilambangkan dengan μ (miyu), sedangkan rata-rata sampel dilambangkan dengan x (x bar).

Rumus rata-rata hitung adalah sebagai berikut:

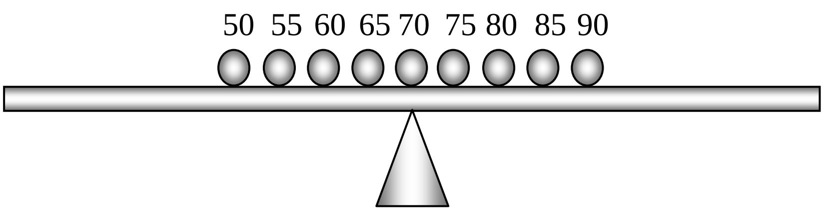
$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Berikut adalah nilai ujian statistika dari sepuluh mahasiswa UNY:

| No. | Nama | Nilai |
|---|---|---|
| 1 | Anto | 50 |
| 2 | Bayu | 55 |
| 3 | Cica | 60 |
| 4 | Deny | 65 |
| 5 | Elan | 70 |
| 6 | Fahri | 75 |
| 7 | Gina | 80 |
| 8 | Hana | 85 |
| 9 | Indy | 90 |
| ***Rata-rata Hitung*** | | ***70*** |

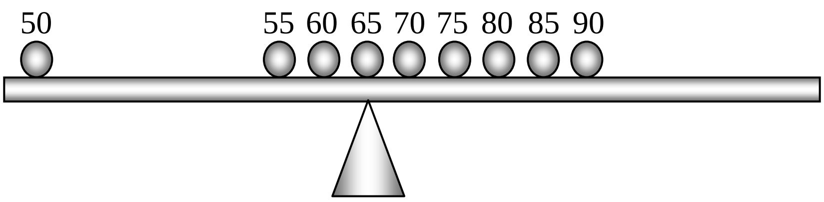
Rata-rata hitung pada baris paling bawah pada tabel di samping (= 70), di dapat dengan cara menjumlahkan nilai kesepuluh mahasiswa (50 + 55 + 60 + ... + 90) kemudian hasilnya dibagi dengan 9 (yaitu jumlah observasi).

Pemahaman makna rata-rata hitung (selanjutnya kita sebut dengan : rata-rata), adalah sebagai berikut:

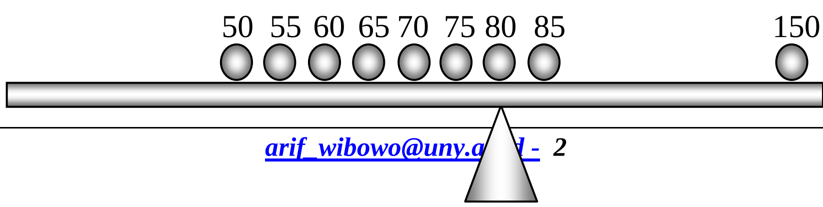



Nilai rata-rata adalah sebuah nilai kesetimbangan yang berfungsi sebagai penyeimbang sehingga observasi-observasi yang nilainya lebih kecil dari mean seimbang dengan observasi-observasi yang lebih lebih besar dari mean.

Apabila nilai 50 dari data tersebut kita ganti dengan 5, maka mean akan berubah menjadi 65. Gambaran kesetimbangan data menjadi sebagai berikut:




Apabila dari data awal, nilai 90 kita ubah menjadi katakanlah 180, maka nilai mean berubah menjadi 80. Gambaran kesetimbangan juga berubah menjadi sebagai berikut:



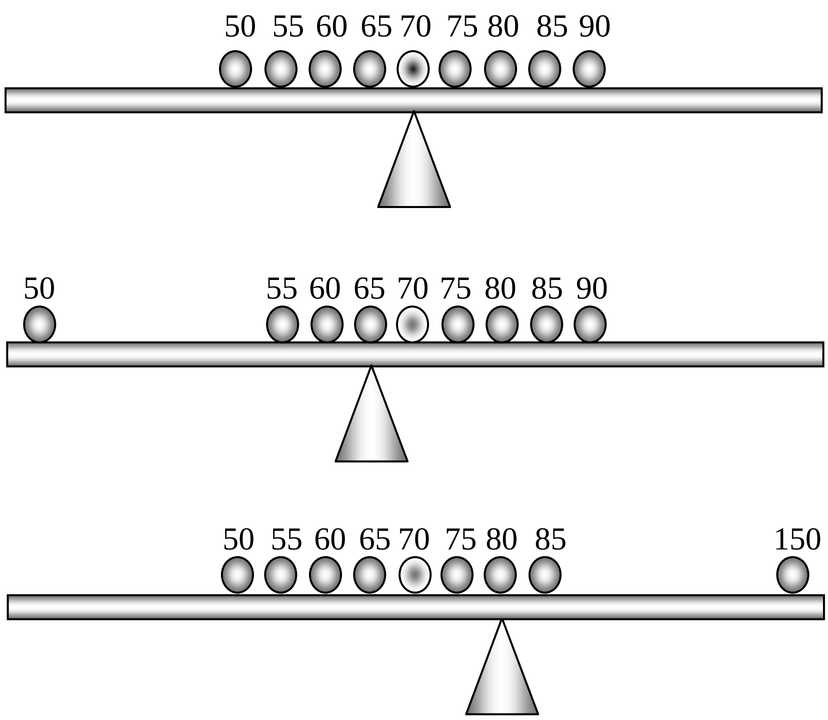
Dari ketiga gambaran di atas, kita lihat bahwa nilai rata-rata sangat tergantung pada besaran tiap-tiap data, termasuk apabila dalam data terdapat nilai ekstreem, yaitu nilai yang sangat kecil atau sangat besar dan jauh berbeda dari kelompok data.

## b. MEDIAN

Adalah nilai tengah dari nilai-nilai pengamatan setelah disusun secara teratur menurut besarnya data. Nilai ini dipengaruhi oleh letak data dalam urutannya, sehingga nilai ini sering disebut dengan "rata-rata posisi". Karena nilai median berada di tengah-tengah dari suatu gugus data (yang disusun berurutan), maka akan terdapat 50% dari jumlah data yang letaknya di bawah median, dan 50% dari jumlah yang lain ada di atas median.

Untuk data yang tidak dikelompokkan, median adalah nilai yang terletak pada posisi: **(N + 1 ) / 2;** dimana N menunjukkan jumlah observasi keseluruhan.

Dari contoh gambaran data di atas, nilai median adalah sebagai berikut




Nilai median dari ketiga contoh tersebut sama yaitu 70. Hal ini karena median semata dipengaruhi oleh posisi data. Median membagi data sama banyaknya dari data yang telah diurutkan mulai dari yang terkecil sampai yang terbesar. Apabila jumlah data genap, maka median akan jatuh di antara dua data, maka nilai median merupakan rata-rata dari dua data tersebut.

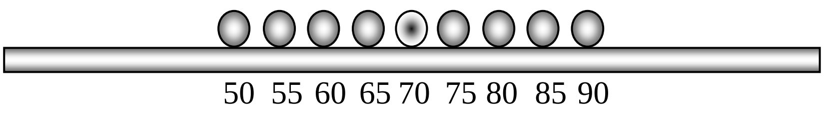
Dari 10 elemen data, maka median jatuh pada data yang terletak pada posisi: **(N + 1 ) / 2** yaitu pada posisi 11/2 = 5.5   Karena median ada pada data ke 5.5, berarti besarnya median merupakan rata-rata dari data ke-5 dan ke-6.  Perhitungan nilai median dengan menjumlahkan data ke-5 dan ke-6 kemudian hasilnya dibagi dua.

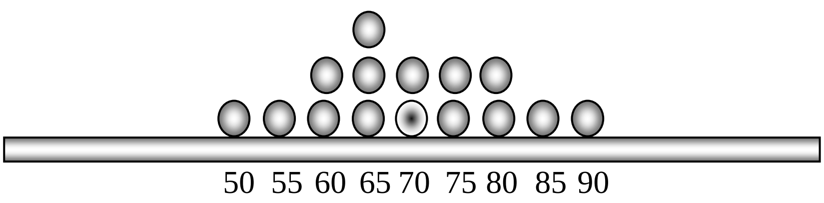
# C. MODUS

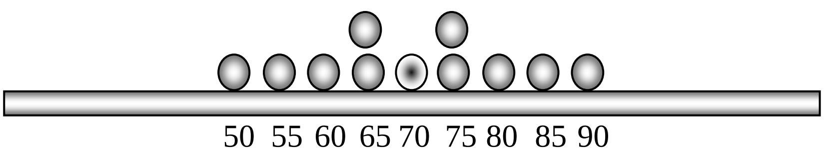
Adalah nilai yang mempunyai frekuensi terbanyak dalam kumpulan data.  Ukuran ini biasanya digunakan untuk mengetahui tingkat seringnya terjadi suatu peristiwa.   (ukuran ini (sebenarnya) cocok digunakan untuk data berskala nominal.

Pada data yang tidak dikelompokkan, modus diperoleh dengan menghitung frekuensi dari masing-masing niai pengamatan, dan kemudian dicari nilai pengamatan yang mempunyai frekuensi observasi paling banyak (nilai data yang paling sering muncul).

Penggambaran contoh data di atas akan kita ubah menjadi sebagai berikut:




Gambar 1, tidak terdapat modus karena semua data mempunyai frekuensi muncul yang sama.

Gambar 2, modus sebesar 65 karena nilai ini yang paling sering muncul.

Gambar 3, modus jatuh pada nilai 65 dan 75. Modus bisa muncul di lebih dari satu tempat, dan kita sebut bimodal.

# MEAN, MEDIAN, DAN MODUS UNTUK DATA BERKELOMPOK

## MEAN UNTUK DATA BERKELOMPOK

Seringkali disamping mencari nilai rata-rata hitung dari gugus data yang masih mengandung nilai-nilai observasi yang lengkap, kita kadang harus menentukan rata-rata hitung dari data yang sudah dikelompokkan dalam distribusi frekuensi. Pada data yang sudah dikelompokkan, sifat asli data sudah tidak nampak, dan sekarang yang nampak adalah sifat kelompoknya. Dengan demikian, kita tidak lagi mengetahui berapa besar nilai-nilai pengamatan sebenarnya sehingga kita tidak dapat mengetahui secara pasti jumlah total nilai-nilai pengamatan yang ada dalam suatu kelas.

Untuk mengatasinya, diberlakukan ***anggapan*** bahwa *besar harga tengah masing-masing interval diperkirakan sama dengan titik tengah interval yang bersangkutan*. Anggapan ini didasarkan atas adanya kemungkinan bahwa nilai-nilai pengamatan yang terletak dalam suatu interval, sebagian lebih kecil dan sebagian lagi lebih besar daripada nilai tengahnya.

Pencarian rata-rata hitung untuk data yang dikelompokkan adalah:

$$mean = \overline{x} = \frac{\sum_{i=1}^{k}(F_i \cdot M_i)}{n}$$

dimana:

$F_i.M_i$ = *nilai tengah dikalikan frekuensi observasi kelas ke-i*
$n$ = *jumlah observasi*
$k$ = *banyaknya kelas.*

Dari rumus di atas, berarti bahwa nilai rata-rata dicari dengan:

1. Mengalikan nilai tengah kelas dengan frekuensi observasi untuk semua kelas
2. Menjuamlahkan hasil pada langkah pertama
3. membagi dengan jumlah observasi (n)

## MEDIAN UNTUK DATA BERKELOMPOK

Untuk mencari median dari data yang telah dikelompokkan, dua langkah yang harus dilakukan: pertama kali ditentukan pada kelas mana letak median berada, dengan rumus: **posisi median = pada data ke- (n+1)/2**. Kedua, menghitung besar nilai median dengan rumus:

$$\text{Median} = Tb + \frac{(N/2) - cfb}{fm}.i$$

Dimana:

Tb = Tepi bawah dari kelas yang mengandung median

N = Banyaknya observasi

cfb = frekuensi kumulatif dari kelas dibawahnya.

fm = frekuensi dari kelas yang mengandung median.

i = Interval kelas

# MODUS UNTUK DATA BERKELOMPOK

Langkah mencari modus dari data yang telah dikelompokkan (hampir) sama seperti kita mencari median, yaitu: pertama, tentukan pada kelas mana letak modus berada, dengan anggapan bahwa modus berada pada kelas dengan frekuensi observasi paling banyak. Kedua, menghitung besar modus dengan rumus:

$$\text{Modus} = Tb + \frac{d1}{d1 + d2}.i$$

dimana:

Tb = Tepi bawah dari kelas yang mengandung modus

d1 = Selisih frekuensi observasi antara kelas yang mengandung modus dengan kelas sebelumnya.

d2 = Selisih frekuensi observasi antara kelas yang mengandung modus dengan kelas sesudahnya.

i = interval kelas

## CONTOH MENGHITUNG MEAN, MEDIAN, DAN MODUS DARI DATA BERKELOMPOK:

Dari contoh yang ada pada bab 2, perhitungan ketiga ukuran pemusatan adalah sebagai berikut.

Tabel distribusi frekuensi selengkapnya dengan tambahan satu kolom, yaitu kolom $F_i.M_i$, adalah sebagai berikut:

| Batas Kelas | *Tepi Kelas* | *M* | F | *cf* | *Fi.Mi.* |
|---|---|---|---|---|---|
| 15 - 20 | *14,5 - 20,5* | *17,5* | 6 | *6* | 105,0 |
| 21 - 26 | *20,5 - 26,5* | *23,5* | 14 | *20* | 329,0 |
| 27 - 32 | *26,5 - 32,5* | *29,5* | 24 | *44* | 708,0 |
| 33 - 38 | *32,5 - 38,5* | *35,5* | 25 | *69* | 887,5 |
| 39 - 44 | *38,5 - 44,5* | *41,5* | 10 | *79* | 415,0 |
| 45 - 50 | *44,5 - 50,5* | *47,5* | 7 | *86* | 332,5 |
| 51 - 56 | *50,5 - 56,5* | *53,5* | 4 | *90* | 214,0 |
| ***Jumlah*** | | | **90** | | **2991,0** |

### 1. MEAN

$$\bar{x} = \frac{\sum_{i=1}^{k}(F_i \cdot M_i)}{n}$$

Dari tabel di atas, diketahui bahwa:

**Σ(*Fi.Mi*) = 2991.0**

**n = 90**

Jadi:

$$\bar{x} = \frac{2991}{90} = 33.2$$

### 2. MEDIAN

**Langkah pertama adalah mencari posisi median:**

Median berada pada data ke (n + 1) /2 = (90 + 1) / 2 = 45.5

Dari kolom cf:

Pada kelas-1 baru terdapat 6 elemen data
Sampai pada kelas ke-2 ada sejumlah 20 elemen data.
Sampai pada kelas ke-3 ada sejumlah 44 elemen data.
Sampai pada kelas ke-4 ada sejumlah 69 elemen data.

Berarti data ke-45.5 ada pada kelas keempat (pedapatan antara 33 – 38)

**Langkah kedua adalah menghitung besarnya median:**

Median = $Tb + \frac{(N/2) - cfb}{fm}.i$

Dari tabel di atas, diketahui:

| | |
|---|---|
| Tb | = Tepi bawah kelas ke-4 = 32.5 |
| N/2 | = 90/2 = 45 |
| cfb | = cf dari kelas ke-3 = 44 |
| fm | = frekuensi kelas ke-4 = 25 |
| I | = 6 |

maka, median = $32.5 + \frac{(90/2) - 44}{25}.6$

= 32.5 + 0.36

= 32.86

## 3. MODUS

**Langkah pertama adalah mencari posisi modus:**

Digunakan anggapan bahwa modus berada pada kelas dengan frekuensi paling banyak, yaitu pada kelas ke-4 (kelas pendapatan 33 – 38).

**Langkah kedua mencari besarnya modus.**

Modus = $Tb + \frac{d1}{d1 + d2}.i$

Dari tabel kita, diketahui bahw:

Tb = 32.5

d1 = 25 – 24 = 1

d2 = 25 – 10 = 15

I = 6,

maka:

Modus = $32.5 + \frac{1}{1 + 15}.6$

= 32.5 + 0.375

= 32.875

# UKURAN PENYEBARAN (*VARIATION*)

## d. RANGE (=rentang = jangkauan)

Adalah ukuran variasi yang dihitung dari selisih antara nilai yang terbesar dengan nilai terkecil. Range sangat mudah dihitung tetapi memang sangat jarang digunakan sebagai ukuran penyimpangan. Biasanya range digunakan dalam pengendalian mutu, atau dalam melihat fluktuasi harga, dan ramalan cuaca.

Range = Xh – Xl,

dimana :

Xh = data tertinggi
Xl = data terendah

Misalnya dari contoh gugus data di depan, kita ketahui bahwa data tertinggi adalah 53 dan data terkecil adalah 15, berarti range dari gugus data kita adalah:

Range = 53 – 15

= 38

## e. VARIASI dan STANDAR DEVIASI

Standar deviasi ini merupakan ukuran variasi yang paling banyak digunakan, karena nilainya paling memenuhi kriteria statistika.

Standar deviasi adalah akar kuadrat dari variasi. Variasi dicari dengan menghitung selisih dari setiap elemen data dengan rata-rata.

Variasi dibedakan antara Variasi populasi ($\sigma^2$) dengan variasi sampel ($S^2$), demikian juga kita mengenal standar deviasi populasi ($\sigma$) dan standar deviasi sampel (S).

Rumus Variasi untuk sampel dan populasi adalah sebagai berikut:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}$$

Sedangkan standar deviasi populasi dan sampel adalah:

$\sigma = \sqrt{\sigma^2}$ $\quad\quad\quad\quad$ $s = \sqrt{s^2}$

Contoh perhitungan variasi dan standar deviasi dari data yang belum dikelompokkan adalah sebagai berikut:

Misalnya data usia 5 mahasiswa manajemen UNY adalah**: 20 : 19 : 21 : 22 : 18**

Rata-rata usia kelima mahasiswa = (20+19+21+22+18)/5 = **20**

Untuk memudahkan perhitungan, kita susun data ke dalam kolom-kolom sebagai berikut:

| X | x - rata-rata | (x-rata-rata)$^2$ |
|---|---|---|
| 20 | 0 | 0 |
| 19 | -1 | 1 |
| 21 | 1 | 1 |
| 22 | 2 | 4 |
| 18 | -2 | 4 |
| **Jumlah (sigma)** | | **10** |

**rata-rata =20*

Varians = $\sigma^2 = \dfrac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$

$= \dfrac{10}{5}$

$= 2$

Standar Deviasi = $\sigma = \sqrt{\sigma^2}$

$= \sqrt{2}$

$= 1.41$

Perhitungan standar deviasi sedikit berbeda apabila data sudah dikelompokkan ke dalam tabel distribusi frekuensi:

Dari contoh pendapatan tahunan penduduk DIY, perhitungan standar deviasi adalah sebagai berikut:

| Batas Kelas | M | F | (Mi - x) | $(Mi - x)^2$ | $Fi(Mi - x)^2$ |
|---|---|---|---|---|---|
| 15 - 20 | *17,5* | 6 | *-15,73* | 247,54 | 1485,226667 |
| 21 - 26 | *23,5* | 14 | *-9,73* | 94,74 | 1326,328889 |
| 27 - 32 | *29,5* | 24 | *-3,73* | 13,94 | 334,5066667 |
| 33 - 38 | *35,5* | 25 | *2,27* | 5,14 | 128,4444444 |
| 39 - 44 | *41,5* | 10 | *8,27* | 68,34 | 683,3777778 |
| 45 - 50 | *47,5* | 7 | *14,27* | 203,54 | 1424,764444 |
| 51 - 56 | *53,5* | 4 | *20,27* | 410,74 | 1642,951111 |
| ***Jumlah*** | | **90** | **15,866667** | **1043,96** | **7025,60** |

***Mean = 33,2***

$$\sigma^2 = \frac{\sum Fi(Mi - \overline{X})^2}{N}$$

$$= \frac{7025.6}{90}$$

$$= 78.06$$

Standar Deviasi = σ = $\sqrt{\sigma^2}$ = $\sqrt{78.06}$

= 8.84

## *ARTI ANGKA STANDAR DEVIASI:*

Angka standar deviasi (dan varians) tidak akan pernah bernilai negatif. Hal ini karena angka ini didapatkan dari hasil pengkuadratan beda tiap observasi. Namun demikian, secara teoritis angka ini masih ada kemungkinan bernilai nol, yaitu apabila tidak terdapat variasi sama sekali dalam kelompok data. Namun demikian, kasus seperti ini (tidak ada variasi) hampir tidak pernah ada.

Berbagai fenomena random yang kita selidiki selalu mempunyai mempunyai variasi dalam nilai-nilainya. Karena sifat variasi ini, maka apabila kita ingin mempelajari sifat-sifat suatu data secara utuh, maka kita tidak boleh hanya mempelajari ukuran tendensi sentralnya saja, tetapi harus dipelajari juga ukuran variasi untuk memberikan gambaran bagaimana kelompok data tersebut menyebar.

Perhatikan tiga gambaran kelompok data berikut:

**Data nilai Statistika Manajemen 2008**

| No. | Kelas A | Kelas B | Kelas C |
|---|---|---|---|
| 1 | 10 | 6 | 2 |
| 2 | 10 | 7 | 4 |
| 3 | 10 | 8 | 6 |
| 4 | 10 | 9 | 8 |

| | | | |
|---|---|---|---|
| 5 | 10 | 10 | 10 |
| 6 | 10 | 11 | 12 |
| 7 | 10 | 12 | 14 |
| 8 | 10 | 13 | 16 |
| 9 | 10 | 14 | 18 |
| rata rata | **10** | **10** | **10** |

Rata-rata nilai statistika kelas A sama dengan rata-rata nilai kelas B, sama dengan rata-rata nilai kelas C yaitu 10. Namun demikian persebaran data (variasinya) sangat jauh berbeda.

Untuk mengetahui seberapa besar tingkat variasi data kita gunakan standar deviasi.

| ***Kelas A*** | **$(Xi\text{-}X)^2$** |
|---|---|
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *10* | 0 |
| *Jumlah* | **0** |

| ***Kelas B*** | **$(Xi\text{-}X)^2$** |
|---|---|
| *6* | 16 |
| *7* | 9 |
| *8* | 4 |
| *9* | 1 |
| *10* | 0 |
| *11* | 1 |
| *12* | 4 |
| *13* | 9 |
| *14* | 16 |
| *Jumlah* | **60** |

| ***Kelas C*** | **$(Xi\text{-}X)^2$** |
|---|---|
| *2* | 64 |
| *4* | 36 |
| *6* | 16 |
| *8* | 4 |
| *10* | 0 |
| *12* | 4 |
| *14* | 16 |
| *16* | 36 |
| *18* | 64 |
| *Jumlah* | **240** |

$\sigma_a = \sqrt{(0/9)}$
$= \mathbf{0}$

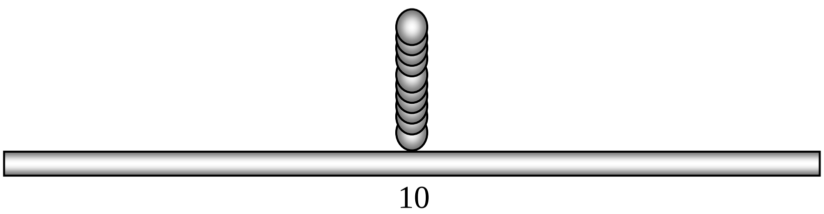
$\sigma_b = \sqrt{(60/9)}$
$= \mathbf{2.58}$

$\sigma_c = \sqrt{(240/9)}$
$= \mathbf{5.16}$

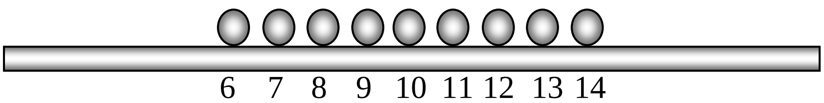
Standar deviasi kelas A = 0; berarti sama sekali tidak ada variasi data
Standar deviasi kelas B = 2.58; terdapat variasi tetapi tidak terlalu besar
Standar deviasi kelas C = 5.16; variasinya lebih besar daripada dua kelas sebelumnya.

Gambaran ketiga kelas adalah sebagai berikut:

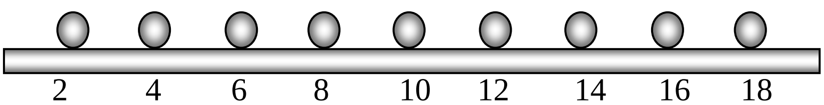
Kelas A:




Kelas B:




Kelas C:



Dari tiga kelompok data dengan rata-rata yang sama, perbedaan standar deviasi memberikan gambaran seperti tersebut di atas.

# e. KOEFISIEN VARIASI

Adalah ukuran variasi relatif yang bertujuan membandingkan variasi dari beberapa gugus data yang mempunyai satuan berbeda. Dengan ukuran Satuan Variasi (KV), maka besaran nilai tidak dipengaruhi oleh satuan pengukuran data aslinya, sehingga parameter yang sama dari beberapa populasi yang menggunakan unit pengukuran berbeda pun dapat diperbandingkan. Koefisien variasi sangat berguna dalam membandingkan dua (atau lebih) kelompok data yang mempunyai besaran berbeda.

Rumus untuk Koefisien Korelasi adalah:

$$KV = \frac{\sigma}{\mu}.100\% \qquad\qquad KV = \frac{s}{\bar{x}}.100\%$$

Koefisien variasi sangat berguna dalam membandingkan variasi dari dua (atau lebih) kelompok data yang mempunyai besaran berbeda.

Misalnya, dari dua kelompok masyarakat industri di kabupaten bantul, (Industri kecil dan industri menengah) ingin diketahui kelompok mana yang mempunyai tingkat penghasilan lebih seragam (diindikasikan dengan standar deviasi yang kecil).

Didapat data rata-rata penghasilan bulanan dari responden di kedua kelompok industri, sebagai berikut:

*Rata-rata penghasilan bulanan dua kelompok industri di Kapubaten Bantul (**dalam juta Rupiah**)*

| Industri kecil | Industri menengah |
|---|---|
| 10 | 80 |
| 15 | 75 |
| 13 | 85 |
| 14 | 78 |
| 18 | 80 |
| 13 | 80 |
| 12 | 75 |
| 16 | 82 |
| 18 | 75 |
| 17 | 78 |

Rata-rata dan standar deviasi kedua kelompok data tersebut adalah:

| | |
|---|---|
| x = 14.6 | x = 78.8 |
| **s = 2.54** | **s = 3.12** |

Dilihat dari standar deviasi, kelompok industri menengah memounyai nilai s yang lebih besar. Dari pemahaman kita tentang tentang standar deviasi kita tahu bahwa standar deviasi yang besar mengindikasikan penyebaran data yang lebih bervariasi dalam kelompok data tersebut. Namun, apabila kita cermati bahwa besaran data dari kedua kelompok data di atas tidaklah sama (kelompok industri kecil dengan besaran “Belasan juta” sedangkan kelompok industri menengah dengan besaran Tujuh puluh hingga delapan puluhan juta rupiah).

Dengan membandingkan variasi dari dua kelompok data dengan besaran yang tidak sama, tidak bisa diambil kesimpulan bahwa kelompok data dengan standar deviasi yang lebih besar berarti variasi data lebih besar.  Untuk kepentingan ini, maka kita harus menggunakan KOEFISIEN VARIASI.

Perhitungan koefisien variasi dari dua kelompok data di atas adalah sebagai berikut:

| ***Industri Kecil*** | ***Industri Menengah*** |
|---|---|
| $x_k$ = 14.6 | $x_m$ = 78.8 |
| **$s_k$ = 2.54** | **$s_m$ = 3.12** |
| **$KV_k = s_k/x_k$<br>= 2.54/14.6<br>= 0.17** | **$KV_m = s_m/x_m$<br>= 3.12/78.8<br>= 0.04** |

Dari koefisien variasi, terlihat bahwa kelompok industri menengah (ternyata) mempunyai variasi yang lebih kecil (dengan KV = 0.04) dibandingkan kelompok industri kecil (KV = 0.17).

# g. KUARTIL (Q)

Kuartil adalah nilai-nilai yang membagi data yang telah diurutkan menjadi empat bagian yang sama.  Suatu gugus data akan mempunyai tiga tempat kuartil, yaitu kuartil 1, kuartil 2, dan kuartil 3, dimana kuartil 2 sebenarnya sama dengan median (yaitu data yang letaknya ditengah-tengah gugus data)

Posisi kuartil dari data yang sudah diurutkan bisa dicari dengan rumus:

Q$k$ = nilai yang ke- $\frac{(n+1)}{4}.k$ ; dimana $k$ = 1, 2, 3 (=posisi kuartil ke-i)

Dari posisi kuartil yang sudah diketahui, nilai kuartil adalah nilai data yang berada ada posisi data tersebut.

Untuk data yang sudah dikelompokkan, Nilai kuartil ini bisa dicari dengan rumus:

$$Q_k = Tb + \frac{(kN/4) - cfb}{f_Q}.i$$

Dimana:

Qk = Kuartil ke-*k*

Tb = Batas bawah nyata kelas yang mengandung Qk

cfb = frekuensi kumulatif di bawah kelas yang berisi Qk

fQ = Frekuensi observasi kelas yang mengandung Qk

i = Interval Kelas

k = 1, 2, 3

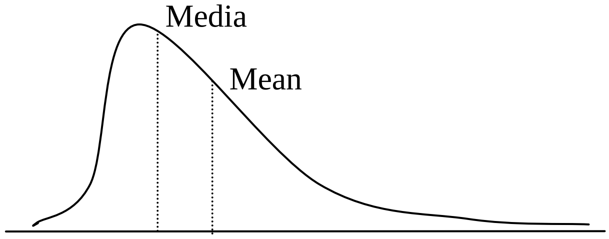
N = banyaknya observasi

# BENTUK DISTRIBUSI DATA

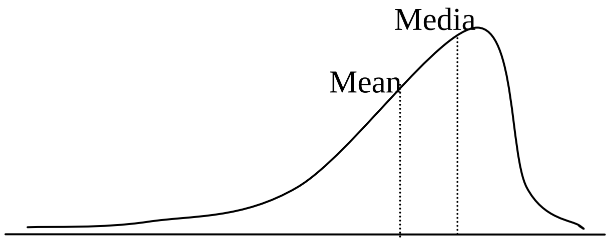
Distribusi data dilihat dari kememcengannya bisa berbentuk simetris, menceng ke kiri, atau menceng ke kanan. Pada kurva yang simetris, nilai median, rata-rata, dan modus sama besar dan berada di tengah-tengah data.

Sebuah distribusi data akan berbentuk simetris apabila banyaknya data dibawah nilai rata-rata secara reatif sama dengan banyaknya data yang ada di atas nilai rata-rata. Pada distribusi data yang simetris, nilai rata-rata sama atau hampir sama dengan median (dan modus). Salah satu dari distribusi data simetris yang banyak digunakan dalam statistika adalah distribusi normal. Suatu distribusi dikatakan berdistribusi normal apabila data berdistribusi simetris dan bersifat unimodal (memiliki satu modus).

Bentuk distribusi data yang tidak memiliki kriteria di atas dikenal dengan distribusi yang tidak simetris. Secara umum, pada distribusi yang tidak simetris dikenal distribusi yang menceng ke kiri atau menceng ke kanan. Data berdistribusi menceng ke kanan apabila ada beberapa data yang terlalu besar (ekstrim). Pada data berdistribusi menceng ke kanan dijumpai **rata-rata > median**. Tampilan distribusi data yang menceng ke kanan adalah sebagai berikut:



Data berdistribusi menceng ke kiri bila ada satu atau beberapa data yang secara ekstrim terlalu kecil sehingga dijumpai kondisi **rata-rata < median**. Tampilan distribusi data yang menceng ke kanan adalah sebagai berikut:




Selain secara visual, kemencengan suatu kurva bisa diketahui melalui koefisien Pearson, dengan formula sebagai berikut:

$$TK = \frac{3(\bar{x} - Med)}{S} :$$

Dimana : x = rata-rata hitung

Med = Median

S = Standar Deviasi

Apabila TK ≈ 0 → distribusi simetris

Apabila TK > 0 → distribusi menceng ke kanan

Apabila TK < 0 → distribusi menceng ke kiri

***Demikian, WaLLahu a'lam***

***al arif 2012***