KUANTITATIF DAN KUALITATIF LINGKUNGAN

(STATISTIKA LINGKUNGAN)

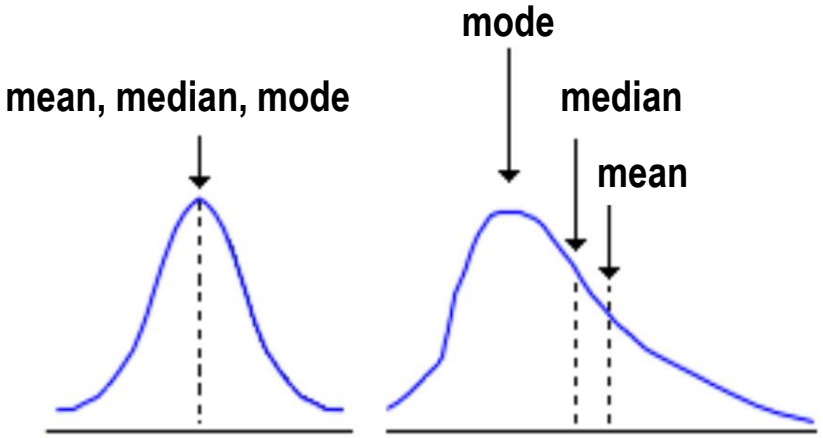
Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN
PASCASARJANA UNIVERSITAS JAMBI

KUANTITATIF DAN KUALITATIF LINGKUNGAN

(STATISTIKA LINGKUNGAN)

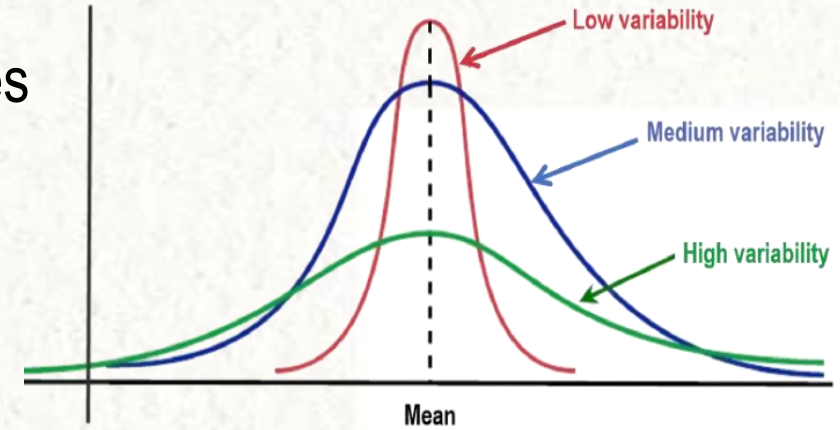
Measures of Central Tencdency



STATISTIKA DESKRIPTIF

- Central Tendency Measures
- Variation or Variability measures

Measures of Variability



Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

/Zed_A

OVERVIEW

GARIS-GARIS BESAR PERKULIAHAN

BAGIAN I. PENGERTIAN DAN KONSEP

1. Pengertian Statistika
2. Populasi dan Sampel
3. Statistika Deskriptif dan Statistika Infrensial
4. Data dan Variabel
5. Kegunaan Statistika

BAGIAN II.STATISTIKA DESKRIPTIF

- Pengukuran Nilai Sentral
- Pengukuran Dispersi
- Distribusi Frekuensi

GARIS-GARIS BESAR PERKULIAHAN (Lanjutan…)

BAGIAN III. STATISTIKA INFERENSIAL

1. Hipotesis
2. Teknik Sampling
3. Chi – Kuadrat
4. Analisis t–test
5. Analisis F–test
   (Analisis Varian Untuk K–Kategori Perlakuan)
6. Korelasi Dan Regresi

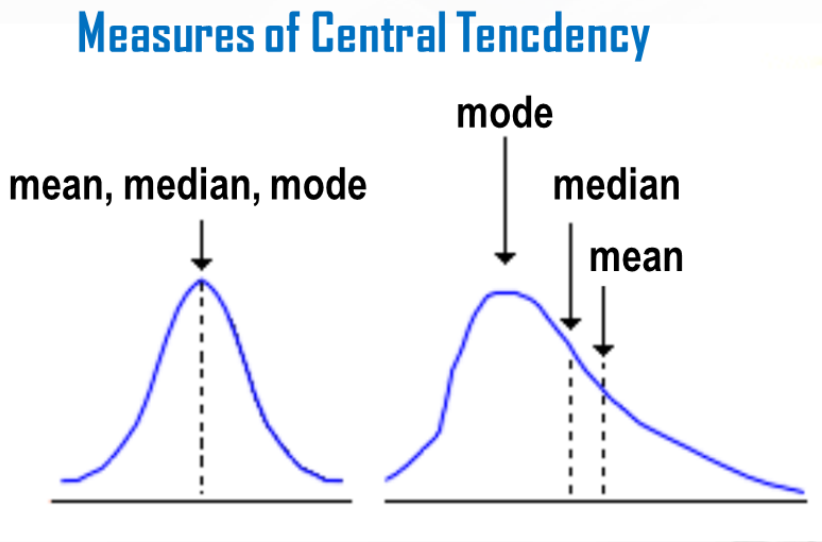
BAGIAN IV. STATISTIKA NON PARAMETRIK

1. Kasus satu sampel
2. Kasus dua sampel dependent
3. Kasus dua sampel independent

STATISTIKA DESKRIPTIF

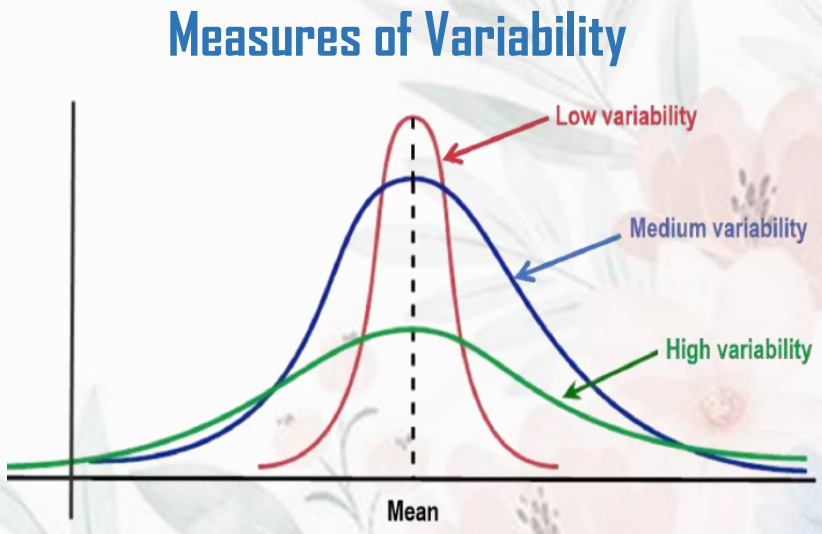
Central Tendency Measures

- Mean (Nilai Tengah)
- Median
- Mode (Modus)



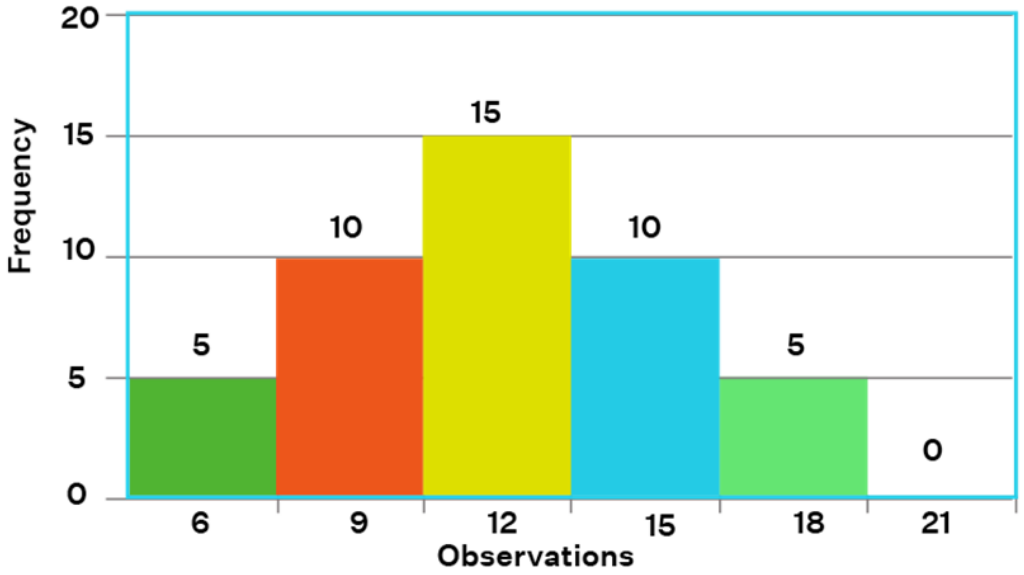
Variance or Variability Measures

- Variance (Ragam)
- Standard deviation (Simpangan baku)
- Koefisien variasi (Coefficient of variation)
- Interquartil Range (Kisaran:



Zulkifli Alamsyah

Measures of Central Tendency



Mean = 12
Median = 12
Mode = 12

Mean (Rata-rata)

Function: Provides the average value of all data. Suitable for use when the data distribution is symmetrical and there are no extreme values (outliers).

$$\text{Mean } X = \frac{X_1 + X_2 + \ldots + X_n}{n} \quad \text{atau} \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Jika data dikelompokkan kedalam kelas:

No.	Selang Kelas Umur (thn)	Frekuensi
1	21 – 25	10
2	26 – 30	15
3	31 – 35	13
4	36 - 40	10
	Jumlah	**48**

$$\text{Mean : } \bar{X} = \frac{\sum_{j=1}^{g} m_j f_j}{n}$$

Dimana:

n = jumlah pengamatan
g = jumlah kelas
m_j = nilai tengah kelas ke-j
f_j = frekuensi kelas ke-j

Zulkifli Alamsyah

Selang Kelas Umur (thn)	Nilai Tengah Kelas (m_j)	Frekuensi (fj)	mj . fj
21 – 25	23	10	230
26 – 30	28	15	420
31 – 35	33	13	429
36 - 40	38	10	380
Jumlah		**48**	1.459

n = 48

g = 4

Σmj fj = 1.459

$$\bar{X} = \frac{1.459}{48} = 30{,}40$$

Kesimpulan:

Rata-rata umur responden yang tersebar dalam 4 kelas umur adalah 30,08 tahun

Zulkifli Alamsyah

Median

- Median Adalah nilai tengah dari sekumpulan data bila diurutkan dari terkecil sampai terbesar.
- Tidak dipengaruhi oleh nilai-nilai ekstrim

Contoh:

No. Obs.	Nilai	No. Obs.	Nilai
1	23	7	23
2	28	8	24
3	26	9	21
4	22	10	24
5	26	11	27
6	25	12	26

- Nilai observasi diurutkan dari terkecil ke terbesar
- Membagi 2 data (50% datas dan 50% dibawah median)

$$\text{Urutan Median} = \frac{n+1}{2}$$

Urutan median antara 5 dan 6 → Nilai Median = (24+25)/2 = 24.5

No. Obs.	Nilai
9	21
4	22
1	23
7	23
8	24
10	24
6	25
3	26
5	26
12	26
11	27
2	28

6 Obs.

6 Obs.

Zulkifli Alamsyah

Median untuk Data Berkelompok:

Menentukan suatu nilai yang membagi dua seluruh data dalam tabel distribusi frekuensi

Contoh: Data 2;

Selang Kelas	Freku-ensi
5 – 8	3
9 – 12	6
13 – 16	5
17 - 20	4
21 - 24	2
Jumlah	**20**

Formula yang digunakan untuk menghitung media (Md) adalah:

$$Md = L + \frac{n/2 - CF}{f} \times i$$

Dimana:

L = nilai bawah kelas yang mengandung nilai median
n = Jumlah data
CF = Frekuensi kumulatif sebelum kelas ybs
f = Frekuensi dalam kelas ybs
i = selang kelas

$$Md = 13 + \frac{20/2 - 9}{5} \times 4 = 13 + {}^{4}/_{5} = 13{,}80$$

Zulkifli Alamsyah

Modus

- Modus adalah angka yang paling sering muncul pada nilai pengamatan.
- Merupakan ukuran pemusatan yang dapat digunakan untuk data kualitatif

Berdasarkan data pada contoh:

Nilai pengamatan 9 paling sering muncul (4 kali).

Oleh karena itu modus data adalah 9.

Contoh:

Nilai	Frekuensi
5	1
6	2
9	4
11	2
13	3
15	2
18	3
20	1
24	2
Jumlah	**20**

Contoh MODUS data berkelompok:

Data1;

Selang Kelas	Freku-ensi
5 – 8	3
9 – 12	6
13 – 16	5
17 - 20	4
21 - 24	2
Jumlah	**20**

Data2;

Selang Kelas	Freku-ensi
2 – 5	2
6 – 9	4
10 – 13	5
14 - 17	5
18 - 21	3
Jumlah	**20**

Untuk Data1;
Modus data adalah kelas 9-12 (Unimodal)

Untuk Data2;
Terdapat dua modus, yaitu kelas 10-13 dan 14-17 (Bimodal)

Distribusi Frekuensi

- Mengorganisasikan data secara sistematik di dalam berbagai macam klasifikasi tanpa mengurangi informasi yang ada dari data tersebut.
- Jika data yang tersedia banyak, maka bisa dibagi ke dalam beberapa kelas.
- Penentuan jumlah kelas, sangat tergantung kepada kisaran (range) data dan interval (lebar) kelas.
- Kaedah **Sturgess** dalam pembuatan tabel distribusi frekuensi:

1. Tentukan Jumlah kelas menggunakan formula sebagai berikut:

$$\Sigma K = 1 + 3{,}3 \text{ Log } N$$

$$2^k > N$$

Dimana: K = kelas
N = jumlah observasi/data

2. Tentukan selang kelas (class interval, CI) dengan formula:

$$Ci = R / \Sigma K$$

Dimana R adalah kisaran (*range*) data:

Range (R) = Nilai terbesar – Nilai terkecil

Contoh:

Data berikut adalah hasil pengamatan terhadap produksi per hektar kedelai dari 30 orang petani sampel dalam satuan kw/ha.

No. Obs.	Kw/ha	No. Obs.	Kw/ha	No. Obs.	Kw/ha	No. Obs.	Kw/ha	No. Obs.	Kw/ha
1	15.0	7	23.7	13	19.7	19	15.4	25	18.3
2	23.0	8	14.2	14	20.8	20	13.5	26	20.7
3	17.4	9	33.8	15	12.9	21	20.3	27	13.7
4	21.4	10	18.3	16	29.8	22	17.1	28	18.9
5	10.3	11	26.1	17	15.7	23	14.0	29	17.8
6	18.6	12	23.2	18	12.9	24	27.1	30	16.6

Jumlah Kelas (K)

= 1 + 3,3 Log (30)

=1 + 4,9 = 6 (dibulatkan)

Kisaran data (Range, R)

= 33,8 – 10,3 = 23,5

Selang kelas (class interval, CI)

= R/K = 23,5/6 =3,92

Kelas pertama:

Nilai terkecil (sampai) Nilai terkecil + CI – 0.01

10.30 sampai (10.30 + 3.92 – 0.01)

→ **10.30 – 14.21**

Kelas selanjutnya → penambahan dgn 3.92

Tabel Distribusi Frekuensi produksi per hektar kedelai dari 30 orang petani sampel

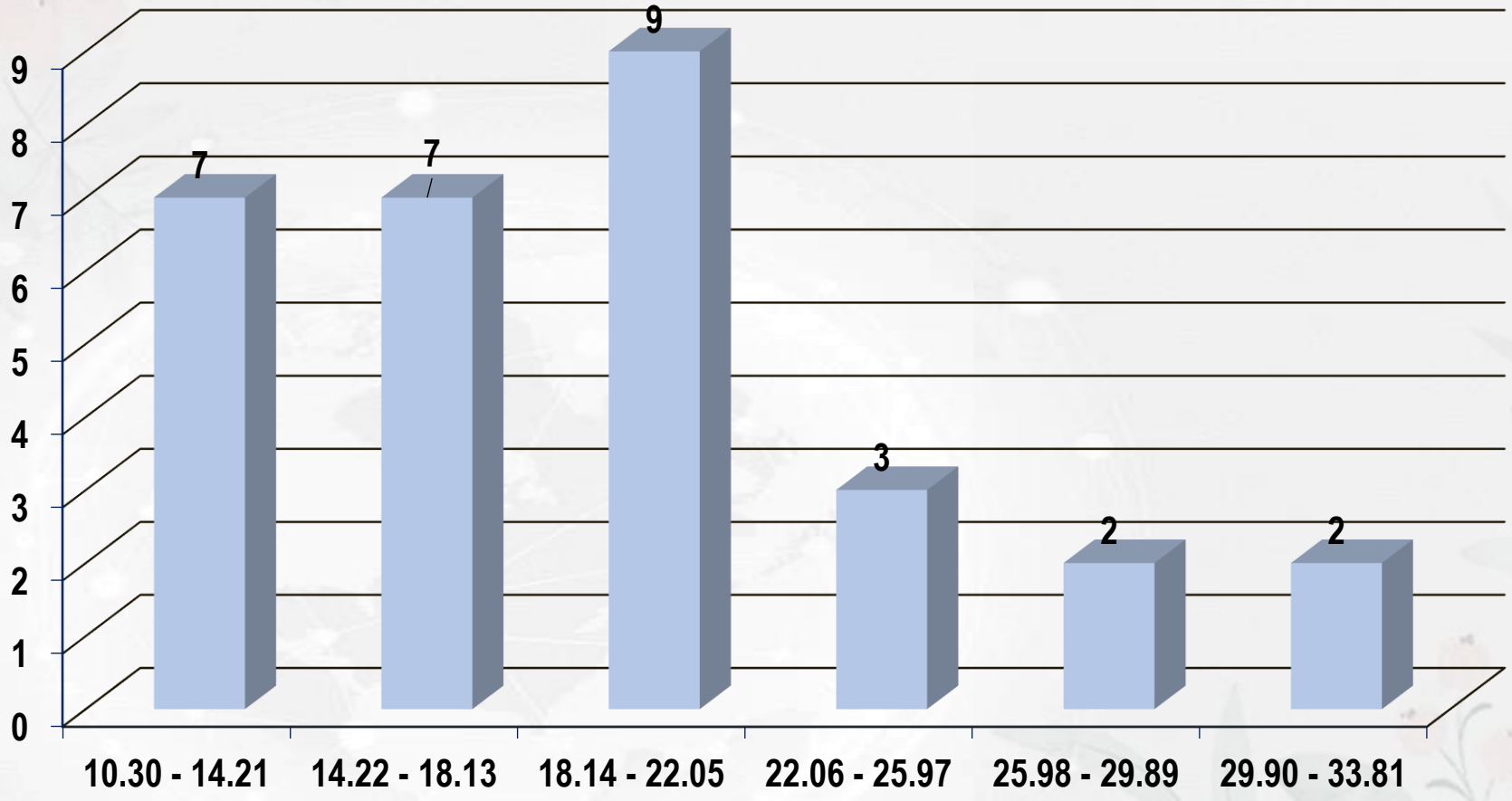
Kelas			Frekuensi
10.30	-	14.21	7
14.22	-	18.13	7
18.14	-	22.05	9
22.06	-	25.97	3
25.98	-	29.89	2
29.90	-	33.81	2
Jumlah			30

3.92 3.92 3.92 3.92

Frekuensi adalah jumlah data yang termasuk dlm kelas yang bersangkutan.

Zulkifli Alamsyah

Penyajian Distribusi Frekuensi dalam bentuk Histogram



Zulkifli Alamsyah

DISPERSION MEASURES

Variance

$$\sigma^2 = \frac{\sum_1^N (x_i - \bar{X})^2}{N}$$

X ⟶ Variable for which variance is to be calculated

x_i ⟶ Observation number i of the variable X. i can take values between 1 and n

N ⟶ Number of observations

$\bar{X}$ ⟶ It is the mean of the variable X

Standard deviation

$$\sigma = \sqrt{\frac{\sum_1^N (x_i - \bar{X})^2}{N}}$$

Statistical range

$$R = Max_x - Min_x$$

R ⟶ It is the range

Max_x ⟶ It is the maximum value of the sample or population

Min_x ⟶ It is the minimum value of the sample or population

X ⟶ Variable for which statistical range is to be calculated

Coefficient of variation

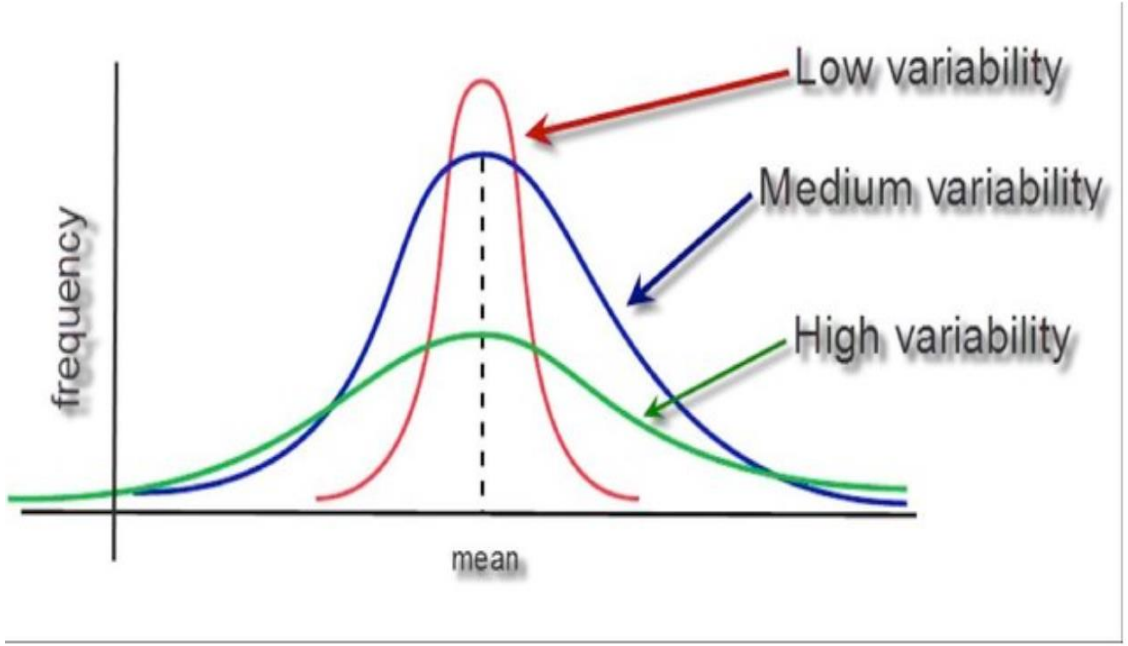
$$CV = \frac{\sigma_x}{|\bar{X}|}$$

X ⟶ Variable for which coefficient of variation is to be calculated

σ_x ⟶ Standard deviation of variable X

$|\bar{X}|$ ⟶ It is the mean of the variable X in absolute value with

Variability of the Data



Variability (also called *dispersion* or *spread*) describes **how far apart the data values are** in a dataset

It tells us **how much the data varies** from the center (mean or median).

- If the values are close together → **low variability**
- If the values are spread out → **high variability**

Variability helps us understand how consistent, stable, or predictable the data is.

Zulkifli Alamsyah

Kisaran (Range)

Kisaran adalah selisih antara nilai pengamatan terkecil dan terbesar.

Contoh: Berdasarkan DATA1, kisaran = 28 – 21 = 7

Ragam (Variance)

Variance menggambarkan bagaimana data terdistribusi dan fluktuasinya terhadap nilai tengah.

Ragam dari sampel (S^2) dihitung dengan formula sbb:

$$S^2 = \frac{\Sigma(X_i - \bar{X})^2}{n - 1} \quad \text{atau} \quad S^2 = \frac{n\,\Sigma X_i^2 - (\Sigma X_i)^2}{n(n-1)}$$

$$= 48.92 / (12-1) = 4.45$$

Data1;

No. Obs.	Nilai	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
1	23	-1.58	2.51
2	28	3.42	11.67
3	26	1.42	2.01
4	22	-2.58	6.67
5	26	1.42	2.01
6	25	0.42	0.17
7	23	-1.58	2.51
8	24	-0.58	0.34
9	21	-3.58	12.84
10	24	-0.58	0.34
11	27	2.42	5.84
12	26	1.42	2.01
Jumlah (Σ)	295		48.92
Rata2 ($\bar{X}$)	24.58		

Zulkifli Alamsyah

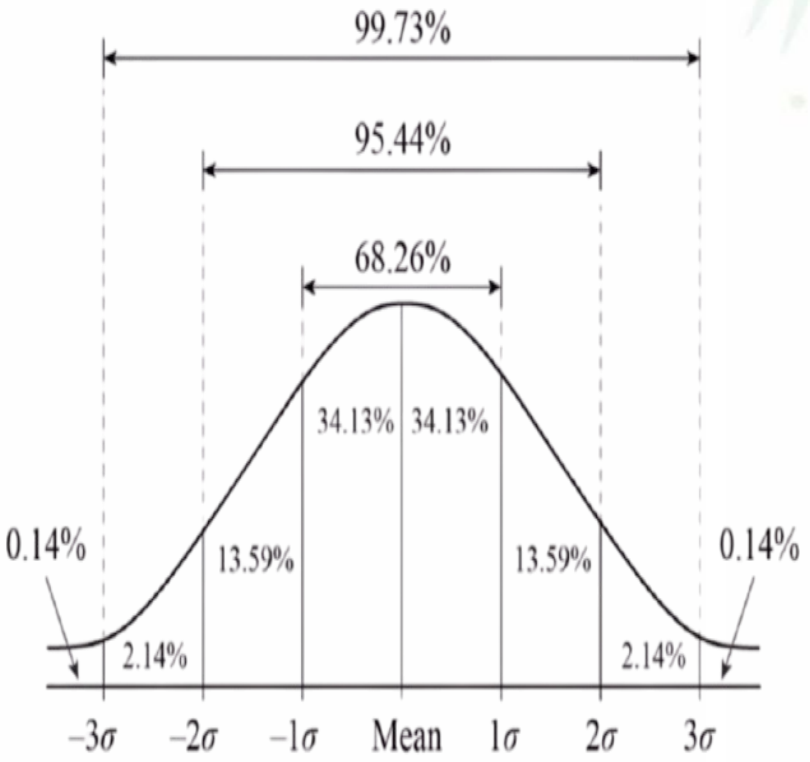
Standard Deviation (Simpangan baku)

- Akar dari nilai varians.
- Merupakan ukuran kesalahan (error) dari setiap n pengukuran (data
- Paling umum digunakan untuk mengukur dispersi karena dinyatakan dalam satuan yang sama dengan data asli.

Berdasarkan **Data1**, $S^2 = 4.45$

Oleh karena itu, $S = \sqrt{4.45} = 2{,}11$.

Distribusi Normal



Sekitar 68% seluruh data berada dalam rentang ±1 simpangan baku dari mean.

Mean $- \mathbf{1}\boldsymbol{\sigma}$ *sampai* **Mean** $+ \mathbf{1}\boldsymbol{\sigma}$

Zulkifli Alamsyah

Koefisien Variasi (coefficient of variation)

- Koefisien variasi menggambarkan sebaran data dari nilai tengah berdasarkan pengukuran relatif (dalam persentase).
- Koefisien variasi didefinisikan sebagai simpangan baku terhadap nilai tengah yang dinyatakan dalam bentuk persentase

Koefisien variasi dihitung dengan formulas sbb:

$$CV = \frac{S}{\bar{X}} \times 100\%$$

Dari DATA1,

$S = 2.11$

$\bar{X} = 24.58$

$CV = (2.11 / 24.58) \times 100\%$

$= \mathbf{8.58\%}$.

- Variasi atau penyebaran data hanya sekitar 8.58% dari nilai rata-ratanya.
- Data sangat homogen, atau stabil dan konsisten, tidak bervariasi besar.

Zulkifli Alamsyah

Data 2 Kelompok Sampel

Interpretasi nilai Koefisien Variasi

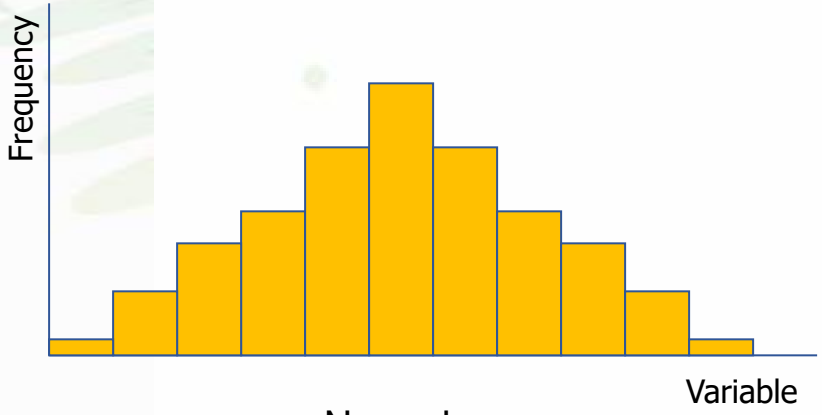
Nilai CV	Interpretasi
< 10%	Variasi sangat rendah (data sangat homogen)
10–20%	Variasi rendah–sedang
20–30%	Variasi sedang
> 30%	Variasi tinggi (data heterogen)

Perhatikan data berikut. Mana yang lebih baik?

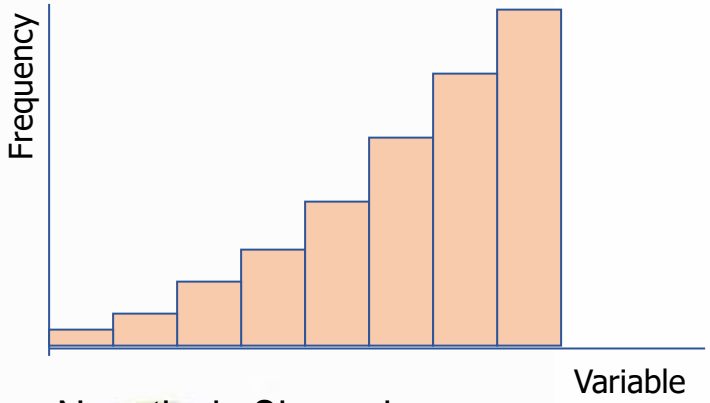
No.	Group A	Group B
1	65	42
2	66	54
3	67	58
4	68	62
5	71	67
6	73	77
7	74	77
8	77	85
9	77	93
10	77	100
Mean	71.5	71.5
Median	72	72
Modus	77	77

Zulkifli Alamsyah

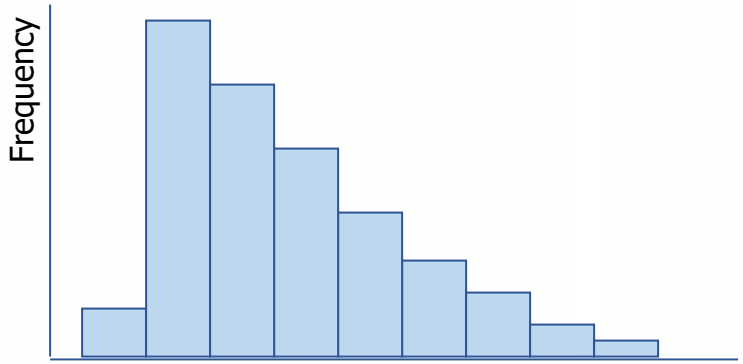
Bentuk-bentuk Sebaran data:



Normal



Negatively Skewed



Positively Skewed

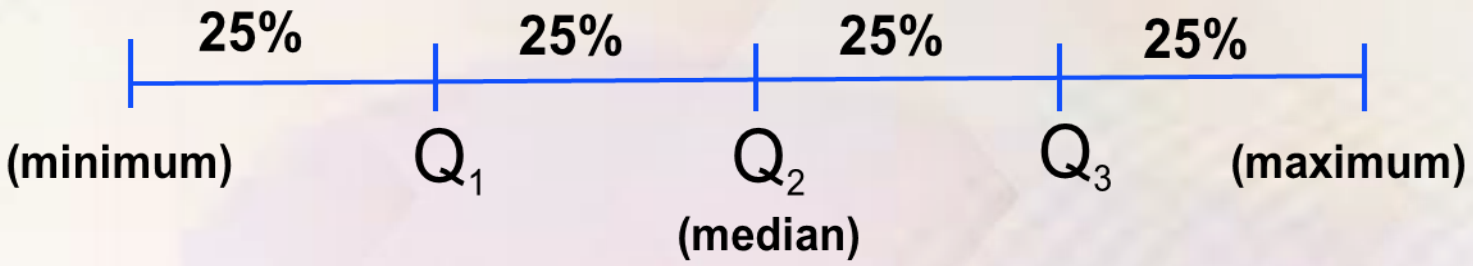
Zulkifli Alamsyah

Quantiles (kuantil)

- Kuantil biasa digunakan untuk memperoleh gambaran lebih rinci dari sekumpulan data.
- Bila median membagi dua data sekumpulan data, kuantil membagi kumpulan data tersebut menjadi 4 (quartil), 10 (desil) dan 100 (persentil).

Kuartil (quartiles)

- Q1 = nilai yang berada pada kuarter pertama yang membagi 25% data yang lebih kecil dan 75% data yang lebih besar.
- Q2 = median
- Q3 = nilai yang berada pada kuarter ketiga yang membagi 75% data yang lebih kecil dan 25% data yang lebih besar.



Interquartile Range / IQR

- **Interkuartil (IQR)** adalah ukuran penyebaran (dispersion) dalam statistika deskriptif yang menunjukkan *rentang nilai tengah* dari suatu data, yaitu selisih antara kuartil ketiga (Q3) dan kuartil pertama (Q1).
- Secara matematis: **IQR** $= \boldsymbol{Q3 - Q1}$
- Menggambarkan penyebaran 50% data di bagian Tengah → seberapa jauh variasi data yang berada di tengah distribusi—yaitu dari posisi 25% hingga 75%.
- Tidak dipengaruhi oleh outlier Tidak seperti rentang (range), IQR tidak sensitif terhadap nilai ekstrim (sangat kecil atau sangat besar) Karena itu, IQR lebih stabil untuk data yang mempunyai pencilan.
- Digunakan untuk mendeteksi outlier

Nilai dianggap outlier jika: **Nilai** $< \boldsymbol{Q1 - 1.5 \times IQR}$

Atau **Nilai** $> \boldsymbol{Q3 + 1.5 \times IQR}$

Zulkifli Alamsyah

Contoh Sederhana

Data:

5, 7, 8, 9, 10, 12, 13, 15, 18, 20

- Q1 = median dari 5 data pertama → median dari (5, 7, 8, 9, 10) = **8**
- Q3 = median dari 5 data terakhir → median dari (12, 13, 15, 18, 20) = **15**

$$\text{IQR} = 15 - 8 = 7$$

Interpretasi:
50% data berada dalam rentang antara 8 hingga 15, atau hanya 7 satuan.

IQR hanya fokus pada data tengah (25%–75%), sehingga relatif stabil.

Perbedaan Karakteristik Pengukuran Dispersi Data

Ukuran	Apa yang diukur?	Rumus / Cara	Sensitif terhadap outlier?	Kapan digunakan?
Range	Seluruh rentang data (min sampai max)	Max − Min	**Sangat sensitif**	Bila ingin gambaran kasar penyebaran
IQR	Penyebaran 50% data tengah	Q3 − Q1	**Tidak sensitif**	Bila data punya outlier atau distribusi tidak normal
Standar Deviasi	Rata-rata deviasi data dari mean	√ Variance	**Sensitif**	Analisis statistik lanjutan, distribusi normal

Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN

PASCASARJANA UNIVERSITAS JAMBI