UKURAN GEJALA PUSAT, LETAK, PENCARAN, KEMIRINGAN

Pengertian

- Ukuran gejala pusat merupakan suatu ukuran atau nilai yang letaknya cenderung terletak dipusat data
- The purpose of these measures is to summarize in a single value the typical size, middle property, or central location of a set of values. The most familiar measure of central tendency is, of course, the arithmetic mean, which is simply the sum of the values of a group of items divided by the number of such items (Smidth dan Sanders)

- Nilai rata-rata dapat diartikan sebagai nilai tipikal atau representatif atau perwakilan dari suatu set data. Beberapa contoh dari ukuran gejala pusat atau rata-rata adalah rata-rata aritmatik (arithmetic mean), median, modus. Di antara berbagai ukuran gejala pusat tersebut memiliki kelebihan dan kekurangan, bergantung pada data dan tujuan yang dimaksud (Spiegel dan Stephens)
- Data often have a tendency to congregate about some central value, and this central value may then be used as a summary measure to describe the general data pattern." (Smidth dan Sanders)

Nilai	Nilai	Nilai	Nilai	Nilai
1	5	9	12	16
2	6	10	13	17
3	7	11	14	18
4	8	11	15	

Jumlah Keseluruhan Nilai (Sum)

Andaikan terdapat n buah nilai, yakni $X1$,$X2$, $X3$, … , Xn. Jumlah dari keseluruhan nilai tersebut dihitung dengan rumus sebagai berikut.

$$jumlah\ keseluruhan\ nilai = \sum_{i=1}^{n} X_i$$

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + X_3 + \cdots + X_n$$

Jumlah keseluruhan nilai untuk data pada Tabel diatas adalah 1 + 2 + 3 + ⋯ + 18 = 172

Rata-Rata Aritmatik atau Rata-Rata Hitung

- Rata-rata aritmatik atau sering disebut juga dengan nama rata-rata hitung, merupakan jumlah seluruh nilai dari data, dibagi dengan banyaknya data. Rumusnya adalah

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

RATA – RATA ARITMATIK PADA TABEL DIATAS

$$\bar{X} = \frac{1 + 2 + 3 + \cdots + 18}{19}$$

$$\bar{X} = 9{,}578947$$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 11, 12, 13, 14, 15, 16, 17, 18

Perhatikan bahwa rata-rata hitung 9,57 cenderung terletak di pusat data.

Modus (Mode)
Modus merupakan nilai data dengan frekuensi atau jumlah kemunculan paling banyak. Berdasarkan data pada Tabel diatas, nilai dengan frekuensi kemunculan paling banyak adalah nilai 11, yakni muncul sebanyak dua kali.

MEDIAN
Median juga disebut juga dengan nilai tengah (middle value) atau rata-rata aritmatik dari dua nilai tengah. Nilai dari median membagi data menjadi dua bagian yang sama˜. Sebelum menghitung nilai median, terlebih dahulu data diurutkan dari yang terkecil hingga terbesar. Rumus menghitung median untuk data dengan jumlah genap.

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

RUMUS UNTUK MENGHITUNG MEDIAN UNTUK DATA JUMLAH GANJIL

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

1,2,3,4,5,6,7,8,9,10,11,11,12,13,14,15,16,17,18.

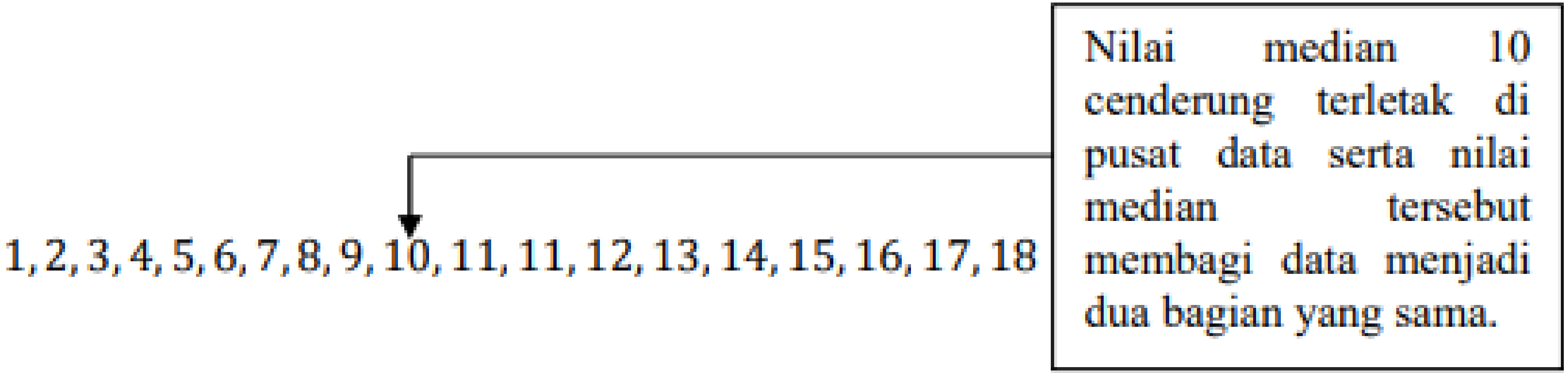
Diketahui banyaknya nilai n = 19, sehingga banyaknya data adalah ganjil.

$Median$ = (Xn+1)/2

$Median$ = (X19+1)/2

$Median$ = X10

X10 berarti nilai median terletak pada data dengan urutan ke−10, yakni 10.



Perhatikan bahwa nilai median membagi menjadi dua bagian yang sama. Bagian pertama adalah {1,2,3,4,5,6,7,8,9}, dan bagian kedua adalah {11,11,12,13,14,15,16,17,18}. Perhatikan bahwa masing-masing bagian terdiri dari 9 nilai.

Berdasarkan uraian tersebut, keuntungan menggunakan median sebagai ukuran gejala pusat adalah median tidak terpengaruh oleh outlier (data pencilan). Oleh karena itu, median lebih disukai dibandingkan rata-rata atau mean (rata-rata aritmatik) sebagai ukuran gejala pusat, untuk data yang mengandung outlier.

Ukuran Letak (Measure of Position)

Kuartil ($\boldsymbol{K}$)

- Ukuran kuartil terdiri dari tiga buah nilai yang membagi data menjadi empat bagian yang sama.

1,2,3,4,5,6,7,8,9,10,11.

Nilai kuartil dikelompokkan atas tiga, yakni kuartil pertama ($K1$), kuartil kedua ($K2$), dan kuartil ketiga ($K3$). Angka 3, 6, dan 9 masing-masing merupakan $K1$, $K2$, dan $K3$.

RUMUS MENGHITUNG KUARTIL

$$K_i = \frac{i(n+1)}{4} \; ; i = 1,2,3$$

Perhatikan bahwa K_i merupakan nilai dari kuartil ke$-i$ dengan $i = 1, 2,$ dan 3. Berikut **contohnya**

1,2,3,4,5,6,7,8,9,10,11,11,12,13,14,15,16,17,18

Diketahui banyaknya nilai data $n = 19$. Berikut akan dihitung nilai dari K_1, K_2, dan K_3.

$$K_1 = \frac{1(19+1)}{4}$$

$$K_1 = 5$$

$K_1 = 5$ berarti nilai K_1 terletak pada data dengan urutan ke-5, yakni 5.

$$K_2 = \frac{2(19+1)}{4}$$

$$K_2 = 10$$

$K_2 = 10$ berarti nilai K_2 terletak pada data dengan urutan ke-10, yakni 10.

$$K_3 = \frac{3(19+1)}{4}$$

$$K_3 = 15$$

$K_3 = 15$ berarti nilai K_3 terletak pada data dengan urutan ke-15, yakni 14. Ketiga nilai kuartil tersebut membagi data menjadi empat bagian yang sama. Bagian pertama adalah $\{1,2,3,4\}$, bagian kedua adalah $\{6,7,8,9\}$, bagian ketiga adalah $\{11,11,12,13\}$, dan bagian keempat adalah $\{15,16,17,18\}$. Perhatikan bahwa banyaknya nilai untuk masing-masing bagian adalah 4.

Desil

Ukuran desil terdiri dari sembilan nilai yang membagi data menjadi sepuluh bagian yang sama.

1,2,3,4,5,6,7,8,9,10,11,11,12,13,14,15,16,17,18.

Perhatikan bahwa nilai-nilai yang dilingkar merupakan nilai-nilai desil. Nilai-nilai tersebut membagi data menjadi 10 bagian yang sama. Masing-masing bagian terdiri dari 1 nilai. Terdapat sembilan nilai desil, yakni desil pertama (D_1), desil kedua (D_2), dan sampai dengan desil kesembilan (D_9). Berikut rumus untuk menghitung nilai desil.

$$D_i = \frac{i(n+1)}{10} \; ; i = 1,2,3,\ldots,9$$

$$D_1 = \frac{1(19+1)}{10} = 2$$

Nilai desil ke−1 terletak pada data dengan urutan ke-2, yakni 2.

$$D_8 = \frac{8(19+1)}{10} = 16$$

Nilai desil ke−8 terletak pada data dengan urutan ke-16, yakni 15.

$$D_9 = \frac{9(19+1)}{10} = 18$$

Nilai desil ke−9 terletak pada data dengan urutan ke-18, yakni 17. Sembilan nilai desil tersebut membagi data menjadi sepuluh bagian yang sama dengan banyaknya nilai untuk masing-masing bagian adalah 1.

Ukuran Pencaran atau Dispersi atau Sebaran

Data 1	70	70	70	70	70	$\bar{X} = 70$
Data 2	50	60	70	80	90	$\bar{X} = 70$
Data 3	20	60	70	100	100	$\bar{X} = 70$
Data 4	20	20	10	100	200	$\bar{X} = 70$

Nilai rata-rata untuk data manakah yang dapat mewakili data dengan baik?

Berdasarkan pengamatan, nilai rata-rata dari data 1 dapat mewakili data 1 dengan baik (secara sempurna), nilai rata-rata dari data 2 cukup baik dalam mewakili data 2, namun nilai rata-rata dari data 3 dan data 4 kurang baik dalam mewakili data 3 dan data 4

- Ukuran pencaran atau dispersi merupakan suatu nilai yang mengukur tingkat pencaran atau sebaran nilai-nilai data terhadap nilai rata-ratanya. Nilai pencaran yang tinggi menunjukkan nilai-nilai data cenderung terletak cukup jauh terhadap nilai rata-rata dari data tersebut. Dengan kata lain, data semakin bervariasi atau heterogen
- Ukuran pencaran diantaranya adalah range, variance, dan standar deviasi

Perhatikan data pada table berikut :

Nilai	Nilai	Nilai	Nilai	Nilai
10	20	30	40	50
10	30	30	40	50
10	30	30	40	50
20	30	30	50	

Nilai Maksimum

Nilai maksimum merupakan nilai yang paling tinggi dari suatu data. Berdasarkan data pada Tabel diatas, nilai maksimum adalah nilai 50.

Nilai Minimum

Nilai minimum merupakan nilai yang paling rendah dari suatu data. Berdasarkan data pada Tabel diatas, nilai minimum adalah nilai 10.

Range

Range merupakan selisih antara nilai maksimum dengan nilai minimum. Diketahui nilai maksimum adalah 50 dan nilai minimum adalah 10, sehingga nilai range adalah 50 − 10 =40. Ukuran range sama seperti rata-rata aritmatik, yakni memiliki kelemahan ketika dalam suatu data mengandung outlier

Sebagai contoh misalkan diberikan data dengan nilai 1, 2, 3, 4, 5, 100. Nilai range berdasarkan data tersebut adalah 100 − 1 = 99. Seandainya data dengan nilai 100 tidak diikutsertakan dalam penghitungan nilai range, maka diperoleh nilai range 5 − 1 = 4. Perhatikan bahwa nilai range menurun, dari 100 menjadi 4. Nilai data 100 merupakan outlier
(data pencilan).

Variance

Variance (dalam hal ini *variance* untuk sampel) dilambangkan dengan s^2. Berikut rumus untuk menghitung nilai *variance*.

$$s^2 = \frac{|X - \bar{X}|^2}{n - 1}.$$

Nilai *variance* sampel (s^2) berdasarkan data pada Tabel 4.3 adalah

$$s^2 = \frac{3 \times |10 - 31{,}6|^2 + 2 \times |20 - 31{,}6|^2 + \cdots + 4 \times |50 - 31{,}6|^2}{19 - 1}$$

$$s^2 = 180{,}7018$$

Standar Deviasi

Standar deviasi merupakan akar kuadrat positif variance. Nilai dari standar deviasi dapat diinterpretasi sebagai nilai yang menunjukkan seberapa dekat nilai-nilai data menyebar atau berkumpul di sekitar rata-ratanya. Standar deviasi merupakan salah satu dari ukuran pencaran yang paling sering digunakan. Diketahui nilai variance adalah 180,7018, sehingga nilai standar deviasi adalah $\sqrt{180,7018} = 13,4425$.

Perhatikan table berikut

Data						Rata-Rata	*Range*	*Variance*	Standar Deviasi
Data 1	70	70	70	70	70	70	0	0	0
Data 2	50	60	70	80	90	70	40	250	15,811
Data 3	20	60	70	100	100	70	80	1100	33,166
Data 4	20	20	10	100	200	70	180	6600	81,240

Nilai standar deviasi data 1 bernilai 0, data 2 bernilai 15,811, data 3 bernilai 33,166, dan data 4 bernilai 81,240. Perhatikan bahwa pada data 1, seluruh nilai data sama, yakni seluruhnya 70, sehingga nilai standar deviasinya 0 (begitu juga dengan nilai range dan variance). Dapat dilihat bahwa semakin besar nilai standar deviasi dari suatu data, maka sebaran data cenderung jauh terhadap rata-ratanya (walaupun ada beberapa data yang dekat dengan rata-ratanya)

				Rata-Rata	Range	Variance	Standar Deviasi
Data 5	13	14	15	14	2	1	1
Data 6	12	14	16	14	4	4	2
Data 7	8	14	20	14	12	36	6
Data 8	1	14	27	14	26	169	13

Nilai rata-rata untuk data 5 sampai data 8 adalah 14. Untuk data 5, jarak 13 ke 14 adalah 1, yakni $|14 - 13| = 1$, begitu juga jarak dari 15 ke 14, yakni $|15 - 14| = 1$. Nilai standar deviasinya adalah 1. Untuk data 6, jarak dari 12 ke 14 adalah 2, yakni $|14 - 12| = 2$, begitu juga jarak dari 16 ke 14, yakni $|16 - 14| = 2$. Nilai standar deviasinya adalah 2. Semakin besar nilai standar deviasi dari suatu data, maka sebaran data cenderung jauh terhadap rata-ratanya

Koefisien Variasi (Coefficient of Variation

Siswa	Berat Badan	Uang Jajan
1	54,33	20000
2	58,89	20000
3	64,33	19000
4	54,21	20000
5	53,45	19000
Rata-Rata	57,042	19600
Standar Deviasi	4,604554	547,722558
Koefisien Variasi	0,080722	0,02794503

Data mana yang lebih bervariasi atau heterogen, apakah data berat badan atau data uang jajan?

Perhatikan bahwa satuan data untuk berat badan (puluhan) dan uang jajan (puluhan ribu) berbeda. Berdasarkan Tabel diketahui nilai standar deviasi dari uang jajan, yakni 547,722, lebih besar dari pada nilai standar deviasi dari berat badan, yakni 4,604. Namun belum tentu berarti bahwa data uang jajan lebihbervariasi atau heterogen dibandingkan data berat badan. Hal ini dikarenakan satuan data berbeda.

- Untuk itu dapat digunakan koefisien variasi untuk membandingkan tingkat variasi atau heterogen di antara dua atau lebih kelompok, ketika satuan data berbeda-beda.

Nilai dari koefisien variasi dihitung sebagai berikut.

$Koefisien\, Variasi$ (KV) = standart deviasi / rata rata (s/X)

Berdasarkan Tabel diatas, diketahui koefisien variasi untuk data berat badan adalah 0,080722, sementara koefisien variasi untuk data uang jajan adalah 0,02794503. Sehingga data berat badan lebih bervariasi atau heterogen dibandingkan data uang jajan

Data yang Dibakukan (Standardized Data)

Suatu variabel yang mengukur deviasi dari rata-rata, dalam unit atau satuan standar deviasi, disebut variabel yang dibakukan (standardized variable)

data dalam bentuk standar atau baku sangat berguna untuk tujuan perbandingan distribusi dari beberapa kelompok data. Suatu data dari variabel asli X, dapat ditransformasi dalam bentuk standar dengan rumus sebagai berikut.

$$Z = \frac{X - \bar{X}}{s}$$

Tabel 4.8

Siswa	Berat	Uang Jajan	Z_Baku	Z_Uang Jajan
1	54,33	20000	-0,588982091	0,730296743
2	58,89	20000	0,401341779	0,730296743
3	64,33	19000	1,582780781	-1,095445115
4	54,21	20000	-0,615043245	0,730296743
5	53,45	19000	-0,780097224	-1,095445115
Rata-Rata	57,042	19600	0	0
Standar Deviasi	4,604554	547,722558	1	1
Koefisien Variasi	0,080722	0,02794503		

Berdasarkan Tabel 4.8, nilai standar atau baku untuk uang jajan 20000 adalah 0,730296743. Nilai tersebut diperoleh sebagai berikut.

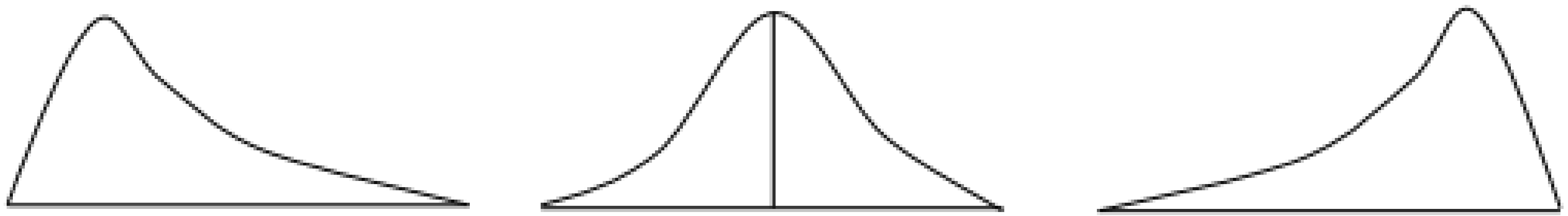
$$Z = \frac{X - \bar{X}}{s} = \frac{20000 - 19600}{547,722558} = 0,730296743$$

Nilai standar atau baku untuk berat badan 54,33 adalah -0,588982091. Nilai tersebut diperoleh sebagai berikut.

$$Z = \frac{X - \bar{X}}{s} = \frac{54,33 - 57,042}{4,604554} = -0,588982091$$

Ukuran Kemiringan (Skewness)

Ukuran kemiringan atau *skewness* merupakan suatu nilai yang mengukur ketidaksimetrisan distribusi data. Suatu data dikatakan berdistribusi simetris sempurna bila nilai rata-rata, median, dan modus dalam data adalah sama.



(1) (2) (3)

Gambar 4.1

Pada Gambar 4.1 (1) kurva cenderung condong ke kanan atau disebut kurva positif, sementara Gambar 4.1 (2) kurva bersifat simetris. Pada Gambar 4.1 (3) kurva cenderung condong ke kiri atau disebut kurva negatif. Berikut rumus untuk menghitung nilai kemiringan suatu data.

$$Kemiringan = \frac{n}{(n-1)(n-2)}\left(\frac{\sum(X-\bar{X})^3}{s^3}\right)$$

Bila nilai kemiringan < 0 atau negatif, maka kurva cenderung condong ke kiri (kurva negatif). Jika nilai kemiringan > 0 atau positif, maka kurva cenderung condong ke kanan (kurva positif). Jika nilai kemiringan mendekati 0 atau 0, maka kurva cenderung simetris.

Tabel 4.10

	X	f	fX	$f\sum(X-\bar{X})^3$
	1	1	1	-17,576
	2	2	4	-8,192
	3	5	15	-1,08
	4	3	12	0,192
	5	2	10	5,488
	6	2	12	27,648
Jumlah		15	54	6,48
Rata-rata ($\bar{X}$)	3,6			
Standar deviasi (s)	1,454058			

Misalkan diberikan data seperti pada Tabel 4.9. Berdasarkan data pada Tabel 4.9, berikut akan dihitung nilai kemiringan. Dari Tabel 4.10, diketahui $\bar{X} = 3{,}6$ dan $s = 1{,}454058$, sehingga nilai kemiringan dapat dihitung sebagai berikut.

$$Kemiringan = \frac{n}{(n-1)(n-2)}\left(\frac{\sum(X-\bar{X})^3}{s^3}\right)$$

$$Kemiringan = \frac{15}{(15-1)(15-2)}\left(\frac{6{,}48}{1{,}454058^3}\right)$$

$$Kemiringan = 0{,}17372$$

Tabel 4.9

Nilai (X)	Nilai (X)	Nilai (X)	Nilai (X)
1	3	4	5
2	3	4	6
2	3	4	6
3	3	5	

Rata-rata > median > modus (miring ke kanan).
Rata-rata < median < modus (miring ke kiri).
Rata-rata = median = modus (simetri)