#### **TUJUAN PEMBELAJARAN**

Tujuan pembelajaran pada bab ini adalah:

- 1) Pembaca bisa menyebutkan tujuan menentukan penyebaran data;
- 2) Pembaca bisa menentukan nilai penyebaran / dispersi data meliputi nilai jangkauan, (range), simpangan rata-rata, simpangan baku (standar deviasi), ragam, koefisien keragaman (KK), ukuran dispersi pada nilai kuartil dan pencilan (outlier) baik untuk data tunggal dan data kelompok;
- 3) Pembaca bisa memilih data yang stabil dari sejumlah kelompok data.

#### **DESKRIPSI MATERI**

Materi yang dibahas pada bab ini adalah tentang tujuan menentukan nilai penyebaran data /dispersi, keragaman dan kestabilan data, penyebaran data yang terdiri dari jangkauan (range), simpangan rata-rata, simpangan baku (standar deviasi), ragam, koefisien keragaman (KK), ukuran dispersi pada nilai kuartil dan pencilan (outlier).

Ukuran pemusatan dan ukuran letak data belum cukup untuk memberikan gambaran yang jelas tentang suatu data, untuk itu kita perlu untuk melihat tentang penyebaran data. Penyebaran data atau sering disebut dispersi merupakan perserakan nilai-nilai observasi terhadap nilai ratarata atau dapat dikatakan bahwa penyebaran data merupakan suatu ukuran yang menunjukkan seberapa jauh data observasi menyebar dari nilai rata-ratanya. Adapun kegunaan pengukuran penyebaran data antara lain:

- Untuk melihat tingkat variasi data.
   Jika variasi data tinggi maka dikhawatirkan bahwa data sampel tersebut bukan berasal dari populasi yang sama, sehingga tidak bisa digunakan untuk mengukur suatu fenomena tertentu.
- Menentukan data tersebut representative atau tidak.
   Data yang representative diperoleh dari kelompok data homogen, untuk mengetahui representative atau tidaknya data kita perlu untuk mengukur penyebaran data.

Semakin besar variasi nilai data observasi maka semakin jauh titik-titik data tersebut dari nilai rata-rata dan data dengan variasi data yang tinggi ini dikatakan data yang tidak representatif. Data yang representative diperoleh dari populasi yang homogen / kelompok data homogen yaitu populasi yang memiliki sifat yang sama atau mirip sehingga memiliki sebaran data yang kecil dan data relatif lebih stabil. Jika data diperoleh dari populasi yang heterogen / kelompok data

heterogen yaitu populasi yang memiliki sifat yang tidak sama atau tidak mirip sehingga akan memiliki sebaran data yang besar dan data relatif tidak stabil.

Berikut diberikan contoh ilustrasi tentang kelompok data homogen dan kelompok data heterogen agar pembaca mendapat gambaran yang lebih jelas. Misalkan, diberikan data pendapatan 5 orang karyawan PT. ABC sebagai berikut :

Tabel 4.1. Contoh Data dari Kelompok Homogen

No.	Nama Karyawan	Pendapatan / Bulan
1	Radhitya	1.500.000
2	Husin Rozak	1.400.000
3	Maharani	1.600.000
4	Asmaa	1.300.000
5	Kallysta	1.400.000
	Jumlah	7.200.000

Pendapatan rata-rata dari data tabel 4.1. diatas adalah:

$$\bar{x} = \frac{7.200.000}{5} = 1.440.000$$

Jarak data pendapatan masing-masing karyawan dengan nilai rata-ratanya adalah :

Untuk Radhitya = |1.500.000 - 1.440.000| = 60.000

Untuk Husin Rozak = |1.400.000 - 1.440.000| = 40.000

Untuk Maharani = |1.600.000 - 1.440.000| = 160.000

Untuk Asmaa = |1.300.000 - 1.440.000| = 140.000

Untuk Kallysta = |1.400.000 - 1.440.000| = 40.000

Misalkan lagi, diberikan data pendapatan 5 orang karyawan PT. DEF sebagai berikut :

Tabel 4.2. Contoh Data dari Kelompok Heterogen

No.	Nama Karyawan	Pendapatan / Bulan
1	Abdul	3.500.000
2	Rozak	1.200.000
3	Taufik	600.000
4	Dita	300.000
5	Rani	1.600.000
	Jumlah	7.200.000

Pendapatan rata-rata data pada tabel 4.2. diatas adalah:

$$\bar{x} = \frac{7.200.000}{5} = 1.440.000$$

# STATISTIK \$ PROBABILITAS PERTEMUAAN 7. DISPERSI / PENYEBARAN DATA

Jarak data pendapatan masing-masing karyawan dengan nilai rata-ratanya adalah :

Untuk Abdul = |3.500.000 - 1.440.000| = 2.060.000

Untuk Rozak = |1.200.000 - 1.440.000| = 240.000

Untuk Taufik = |600.000 - 1.440.000| = 840.000

Untuk Dita = |300.000 - 1.440.000| = 1.140.000

Untuk Rani = |1.600.000 - 1.440.000| = 160.000

Pendapatan rata-rata pada kedua tabel diatas adalah sama, namun data pendapatan pada PT. ABC relative lebih stabil daripada data pendapatan pada PT DEF, hal ini dilihat dari tingginya tingkat variasi data pada tabel 4.2 lebih tinggi daripada tingkat variasi data pada tabel 4.1. Dengan tingginya variasi data pendapatan tabel 4.2. dikhawatirkan data ini berasal dari populasi yang berbeda.

Beberapa jenis ukuran penyebaran data yang akan dibahas pada bab ini adalah:

- 1. Jangkauan (Range)
- 2. Simpangan Rata-Rata (Deviasi Rata-Rata)
- 3. Simpangan Baku (Deviasi Standar)
- 4. Ragam (Variansi)
- 5. Koefisien Keragaman (KK),
- 6. Ukuran Disperse pada Nilai Kuartil
- 7. Pencilan (outlier)

# 4.1. JANGKAUAN (RANGE)

Penentuan jangkauan merupakan pengukuran penyebaran yang paling sederhana. Jangkauan ini dirumuskan sebagai beda antara pengukuran nilai terbesar dan nilai terkecil yang terdapat dalam sebuah distribusi. Semakin kecil nilai jangkauan maka semakin homogen data dalam suatu himpunan data.

$$R = X_t - X_r \tag{4.1}$$

Keterangan:

R: Range

X<sub>t</sub>: Nilai tertinggiX<sub>r</sub>: Nilai terendah

#### JANGKAUAN UNTUK DATA TUNGGAL

Untuk jangkauan data tunggal kita hanya mengurangkan nilai terbesar dengan nilai terkecil.

#### **KASUS 4.1.**

Terdapat data sebagai berikut : 30 35 37 12 32 56 45 56 34, Tentukan jangkauan (range) dari data-data diatas !

# Penyelesaian:

$$x_{maks} = 56$$
 dan  $x_{min} = 12$ , maka nilai jangkauan adalah :

$$R = 56 - 12 = 44$$

#### JANGKAUAN UNTUK DATA BERKELOMPOK

Untuk jangkauan data kelompok, nilai tertinggi diambil dari nilai tengah kelas tertinggi dan nilai terendah diambil dari nilai tengah kelas terendah. Untuk menentukan jangkauan pada data berkelompok kita menggunakan tabel distribusi frekuensi absolut.

# **KASUS 4.2.**

Tentukan jangkauan (range) untuk tabel distribusi frekuensi yang telah dibuat (tabel 2.21)!

# Penyelesaian:

Tabel 4.3. Menentukan Jangkauan (Range) pada Data Berkelompok

Interval Kelas	Frekuensi (f <sub>i</sub> )	Titik Tengah (x <sub>i</sub> )
8-13	5	10.5
14-19	9	
20-15	18	
26-31	21	
32-37	15	
38-43	5	
44-49	7	46,5

$$x_{min} = titik tengah kelas terendah = \frac{8+13}{2} = 10.5$$

$$x_{max} = titik \ tengah \ kelas \ tertinggi = \frac{44 + 49}{2} = 46.5$$

Sehingga nilai jangkauan (range) adalah :

$$R = 46.5 - 10.5 = 36.5$$

# 4.2. SIMPANGAN RATA-RATA (DEVIASI RATA-RATA)

Simpangan rata-rata merupakan ukuran yang menyatakan seberapa besar jarak rata-rata tiap nilai data terhadap nilai rata-ratanya yang artinya kita melihat sebaran titik-titik data

terhadap nilai rata-rata. Semakin kecil nilai simpangan rata-rata untuk suatu kelompok data maka semakin homogen data tersebut yang bermakna bahwa data akan semakin representatif.

Untuk mencari jarak antara suatu titik data  $(x_i)$  dengan nilai rata-rata  $(\bar{x})$  kita dapat mencari nilai mutlak dari nilai titik data  $(x_i)$  dikurangi nilai rata-rata  $(\bar{x})$ . Mengapa kita menggunakan nilai mutlak ? jawabannya karena ada dua kemungkinan keadaan nilai titik data dan nilai rata-rata yaitu nilai titik data  $(x_i)$  < nilai rata-rata  $(\bar{x})$  atau nilai titik data  $(x_i)$  > nilai rata-rata  $(\bar{x})$ , dengan menggunakan harga mutlak maka akan nilai jarak yang dihasilkan selalu akan positif. Pada saat nilai titik data  $(x_i)$  = nilai rata-rata  $(\bar{x})$  maka jarak kedua data ini sama dengan nol.

#### SIMPANGAN RATA-RATA UNTUK DATA TUNGGAL

Persamaan yang digunakan untuk menentukan simpangan rata-rata untuk data tunggal adalah:

$$SR = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} \tag{4.2}$$

Keterangan:

SR = Simpangan rata-rata

n = Banyaknya data

 $x_i = data \text{ ke } i \{i = 1, 2, ..., n\}$ 

 $\bar{x}$  = rata-rata hitung

 $\Sigma$  = notasi sigma yang artinya penjumlahan

# KASUS 4.3.

Tentukan simpangan rata-rata untuk data 30 35 37 12 32 56 45!

# Penyelesaian:

Diketahui n = 7, langkah pertama yang kita lakukan adalah menentukan rata-rata,

$$\bar{x} = \frac{30 + 35 + 37 + 12 + 32 + 56 + 45}{7} = \frac{247}{7} = 35.29$$

Lalu kita menentukan simpangan rata-rata,

$$SR = \frac{|30 - 35.29| + |35 - 35.29| + \dots + |45 - 35.29|}{7} = 9.18$$

jarak data pertama  $(x_1)$  terhadap nilai rata-rata adalah |30-35,29|=5,29, jarak data kedua  $(x_2)$  terhadap nilai rata-rata adalah |35-35,29|=0,29, dan seterusnya. Dan rata-rata jarak penyebaran masing-masing titik data terhadap nilai rata-rata hitung adalah 9.18.

#### SIMPANGAN RATA-RATA UNTUK DATA BERKELOMPOK

Untuk menghitung nilai simpangan rata-rata untuk data berkelompok kita menggunakan tabel distribusi frekuensi absolut. Persamaan untuk menghitung simpangan rata-rata data berkelompok adalah:

$$SR = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{\sum_{i=1}^{n} f_i}$$
 (4.3.)

Keterangan:

SR = Simpangan Rata-Rata

n = banyak kelas

x<sub>i</sub> = nilai tengah kelas ke i

 $\bar{x} = \text{rata-rata hitung}$ 

f<sub>i</sub> = frekuensi kelas ke i

# KASUS 4.4.

Tentukan simpangan rata-rata untuk tabel distribusi frekuensi absolut (Tabel 2.21)!

# Penyelesaian:

Kita akan menggunakan persamaan 4.3. dan untuk menggunakan persamaan 4.3. ini harus mencari nilai rata-rata. Nilai rata-rata hitung untuk tabel 2.21. telah kita cari pada **KASUS** 3.2 (ada tiga cara mencari rata-rata hitung yaitu dengan menggunakan titik tengah, menggunakan rata-rata sementara dan menggunakan coding). Nilai rata-rata yang telah kita peroleh adalah  $\bar{x} = 28,125$ .

Tabel 4.4. Menentukan Simpangan Rata-Rata untuk Data Berkelompok

Interval Kelas	$\mathbf{f_i}$	Xi	$ x_i - \overline{\overline{x}} $	$f_i x_i-\overline{\overline{x}} $
8-13	5	10.5	17.625	88.125
14-19	9	16.5	11.625	104.625
20-15	18	22.5	5.625	101.25
26-31	21	28.5	0.375	7.875
32-37	15	34.5	6.375	95.625
38-43	5	40.5	12.375	61.875
44-49	7	56.5	18.375	128.625
Jumlah	80			588

Sehingga diperoleh rata-rata hitung,

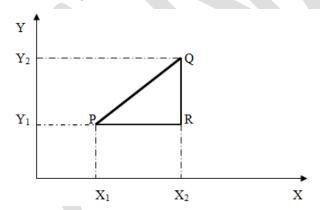
$$SR = \frac{588}{80} = 7,35$$

Pada data berkelompok ini kita menghitung jarak titik tengah masing-masing kelas interval  $(x_i)$  dengan nilai rata-rata. Artinya jarak rata-rata setiap data dengan rata-rata hitung adalah 7,35

# 4.3. SIMPANGAN BAKU (STANDAR DEVIASI)

Istilah simpangan baku pertama kali diperkenalkan oleh Karl Pearson pada tahun 1894. Simpangan baku merupakan salah satu teknik dalam statistika untuk melihat penyebaran data / dispersi data, dimana nilai ini memberikan kita gambaran tentang bagaimana titik-titik data tersebar terhadap nilai rata-rata. Simpangan baku dapat juga dikatakan sebagai rata-rata jarak penyimpangan titik-titik data terhadap nilai rata-rata, nilai ini memperlihatkan keberagaman data dalam suatu kumpulan data. Semakin dekat nilai simpangan baku dengan nilai 0 maka semakin homogen data nilai dalam himpunan tersebut, pada saat nilai simpangan baku sama dengan 0 maka nilai semua data dalam himpunan tersebut adalah sama.

Pada mata kuliah kalkulus, kita menghitung jarak antara dua buah titik pada bidang sistem koordinat. Misalkan, terdapat dua buah titik  $P(x_1, y_1)$  dan  $Q(x_2, y_2)$ , jarak antara titik  $P(x_1, y_2)$ 



Gambar 4.1. Jarak Antara Dua Titik

Untuk mencari jarak antara P dan Q kita bisa menggunakan teorema Pythagoras.

$$\begin{split} |PQ|^2 &= |PR|^2 + |QR|^2 \\ |PQ|^2 &= |X_2 - X_1|^2 + |Y_2 - Y_1|^2 \\ |PQ| &= \sqrt{|X_2 - X_1|^2 + |Y_2 - Y_1|^2} \end{split}$$

Misalkan terdapat dua titik P(6,7) dan Q(10,10), tentukan jarak antara kedua titik tersebut!

Nilai 
$$x_1 = 6$$
,  $y_1 = 7$ ,  $x_2 = 10$ ,  $y_2 = 10$   

$$D(x,y) = \sqrt{(10-6)^2 + (10-7)^2}$$

$$= \sqrt{4^2 + 3^2}$$

$$= \sqrt{16 + 9}$$
$$= \sqrt{25} = 5$$

Dengan kerangka berfikir yang sama, kita dapat mencari jarak antara masing-masing titik data dengan nilai rata-ratanya dengan mencari nilai akar dari penjumlahan hasil kuadrat nilai titik data dikurangi nilai rata-rata. Rumus pada simpangan rata-rata ini didasarkan pada teorema phytagoras (mencari jarak terpendek dari dua buah titik), dimana kita mencari jarak terpendek antara nilai  $x_i$  dengan nilai rata-rata.

Untuk apa kita mencari nilai simpangan baku ? tujuan mencari nilai simpangan baku ini adalah untuk mengetahui apakah sampel yang diambil homogen atau tidak, apakah dapat mewakili populasi, apakah termasuk sampel yang representatif. Misalkan kita mengumpulkan data tentang gaji SATPAM di Kota Palembang, tentu data gaji yang kita kumpulkan akan homogen (gajinya hampir sama), jika ada gaji yang terlalu jauh berbeda dikawatirkan data atau sampel tersebut bukan berasal dari populasi yang sama sehingga tidak bisa mewakili populasi tersebut.

Persamaan untuk menentukan simpangan baku dibedakan menjadi 2, yaitu standar baku untuk populasi ( $n \ge 30$ ) dan standar baku untuk sampel (untuk n < 30). Standar baku untuk populasi dilambangkan dengan  $\sigma$  dan standar baku untuk sampel dilambangkan dengan  $\sigma$ . Semakin kecil nilai simpangan baku semakin kecil jarak titik-titik data dengan nilai rata-ratanya yang maknanya semakin stabil sekelompok data dan semakin representative data tersebut.

#### SIMPANGAN BAKU UNTUK DATA TUNGGAL

Persamaan untuk menghitung simpangan baku sampel adalah:

$$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \tag{4.4}$$

Persamaan untuk menghitung simpangan baku populasi adalah:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}} \tag{4.5}$$

Keterangan:

n = banyak data

 $\sigma$  = Simpangan baku untuk populasi

S = Simpangan baku untuk sampel

 $x_i = Data ke i$ 

 $\bar{x}$  = Rata-rata hitung

#### KASUS 4.5.

Tentukan simpangan baku (standar deviasi) untuk data berikut :

# Penyelesaian:

a) Jumlah data (n) = 5, karena n < 30 maka kita menggunakan persamaan 4.4. Perhatikan persamaan 4.4., pada persamaan ini kita memerlukan informasi tentang rata-rata hitung, sehingga kita harus mencari nilai rata-rata hitung terlebih dahulu dan diperoleh:

$$\bar{x} = \frac{32 + 29 + 25 + 22 + 33}{5} = \frac{141}{5} = 28.2$$

Tabel 4.5. Menentukan Simpangan Baku untuk 4.5. a

Xi	$x_i - \overline{\overline{x}}$	$(x-\overline{\overline{x}})^2$
32	3.8	14.44
29	0.8	0.64
25	-3.2	10.24
22	-6.2	38.44
33	4.8	23.04
	<b>Tumlah</b>	86.8

Sehingga diperoleh rata-rata jarak penyimpangan titik-titik data terhadap nilai rata-rata atau simpangan baku adalah sebesar :

$$S = \sqrt{\frac{86.8}{4}} = 4.66$$

b) Jumlah data (n) = 30, karena n ≥ 30 maka kita menggunakan persamaan 4.5. Perhatikan persamaan 4.5., pada persamaan ini kita memerlukan nilai μ (lambang rata-rata untuk populasi),

$$\mu = \frac{32 + 33 + 34 + \dots + 29}{30} = \frac{836}{30} = 27,8667$$

Tabel 4.6. Menentukan Simpangan Baku untuk 4.5. b

Xi	x <sub>i</sub> - μ	$(\mathbf{x_i} - \boldsymbol{\mu})^2$
32	4,133	17,084
33	5,133	26,351
34	6,133	37,618
29	1,133	1,284
33	5,133	26,351
31	3,133	9,818
33	5,133	26,351
19	-8,867	78,618
21	-6,867	47,151

20	-7,867	61,884
23	-4,867	23,684
34	6,133	37,618
23	-4,867	23,684
33	5,133	26,351
33	5,133	26,351
19	-8,867	78,618
23	-4,867	23,684
31	3,133	9,818
29	1,133	1,284
28	0,133	0,018
23	-4,867	23,684
24	-3,867	14,951
31	3,133	9,818
32	4,133	17,084
34	6,133	37,618
19	-8,867	78,618
25	-2,867	8,218
26	-1,867	3,484
32	4,133	17,084
29	1,133	1,284
Jumlah	836	795,467

Sehingga diperoleh simpangan baku sebesar,

$$\sigma = \sqrt{\frac{795,467}{30}}$$

$$\sigma = \sqrt{27,430}$$

$$\sigma = 5,24$$

# SIMPANGAN BAKU UNTUK DATA BERKELOMPOK

Simpangan baku untuk data berkelompok dapat dicari dengan menggunakan persamaan berikut ini:

$$S = \sqrt{\frac{\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}{n-1}}$$
 (4.6)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i(x_i - \mu)^2}{n}} \tag{4.7}$$

# KASUS 4.6.

Tentukan simpangan baku untuk tabel distribusi frekuensi absolut (Tabel 2.21)!

# Penyelesaian:

Catatan:  $\mu = 28.125$ 

Tabel 4.7. Menentukan Simpangan Baku untuk Data Berkelompok

Kelas Interval	$(\mathbf{f_i})$	Xi	$x_i - \mu$	$(x-\mu)^2$	$f_i(x-\mu)^2$
8-13	5	10.5	-17.63	310.64	1553.20
14-19	9	16.5	-11.63	135.14	1216.27
20-15	18	22.5	-5.63	31.64	569.53
26-31	21	28.5	0.38	0.14	2.95
32-37	15	34.5	6.38	40.64	609.61
38-43	5	40.5	12.38	153.14	765.70
44-49	7	56.5	28.38	805.14	5635.98
Jumlah	80				10353.25

Sehingga diperoleh nilai simpangan baku,

$$\sigma = \sqrt{\frac{10.353,25}{80}} = 11.38$$

# 4.4. RAGAM (VARIANSI)

Variansi adalah rata-rata dari jumlah kuadrat simpangan baku tiap data. Ragam data dirumuskan sebagai :

$$Ragam = S^2 (4.8)$$

Artinya ragam diperoleh dari nilai simpangan baku dikuadratkan.

Berdasarkan kasus 4.5. (a) kita memperoleh nilai Ragam =  $4,66^2$  = 21,7156 dan dari kasus 4.6. kita memperoleh nilai Ragam =  $11.38^2$  = 129,42. Semakin kecil nilai variansi maka tingkat variasi data juga akan semakin kecil sehingga data lebih representative.

#### 4.5. KOEFISIEN KERAGAMAN (KK)

Koefisien Keragaman (KK) digunakan untuk melihat kestabilan sekelompok data. Semakin kecil nilai Koefisien Keragaman (KK) maka semakin stabil keadaan suatu data. Implementasi dari Koefisien Keragaman (KK) adalah kita bisa membandingkan beberapa kelompok data dan menemukan kelompok data mana yang lebih stabil dari kelompok data yang lain. Persamaan untuk mencari nilai Koefisien Keragaman (KK) adalah simpangan baku (S) dibagi dengan rata-rata hitung.

$$KK = \frac{S}{\bar{x}} \tag{4.9}$$

Keterangan:

KK = Koefisien Keragaman

S = Simpangan Baku

 $\bar{x}$  = Nilai Rata-Rata

# KOEFISIEN KERAGAMAN UNTUK DATA TUNGGAL

# **KASUS 4.7.**

Tentukan nilai Koefisien Keragaman (KK) untuk data tunggal berikut ini :

# Penyelesaian:

Kita akan menggunakan persamaan 4.9., Koefisien keragaman adalah simpangan baku (S) dibagi dengan nilai rat-rata ( $\bar{x}$ ).

$$\bar{x} = \frac{111+116+111+112+104}{5} = \frac{554}{5} = 110.8$$

Tabel 4.8. Menentukan Simpangan Baku untuk Kasus Koefisien Keragaman

Xi	$x_i - \overline{\overline{x}}$	$(x-\overline{\overline{x}})^2$
111	0.2	0.04
116	5.2	27.04
111	0.2	0.04
112	1.2	1.44
104	-6.8	46.24
J	lumlah	74.8

Diperoleh nilai simpangan baku (S) dan koefisien keragaman (KK):

$$S = \sqrt{74.8} = 4.32$$

$$KK = \frac{4.32}{110.8} = 0.039$$

# KOEFISIEN KERAGAMAN UNTUK DATA BERKELOMPOK

# **KASUS 4.8.**

Tentukan koefisien keragaman untuk tabel distribusi frekuensi absolut (Tabel 2.21)!

# Penyelesaian:

Nilai koefisien keragaman diperoleh dari membagi nilai simpangan baku (S) dengan nilai rata-rata  $(\bar{x})$ .

- Nilai rata-rata telah dicari pada **kasus 3.2** diperoleh nilai  $\bar{x} = 28,125$
- Nilai simpangan baku telah dicari Pada **kasus 4.6** diperoleh S = 11,38.

Sehingga diperoleh nilai Koefisien Keragaman sebesar,

$$KK = \frac{11,38}{28,125} = 0,40$$

# 4.6 UKURAN DISPERSI PADA NILAI KUARTIL

Kita telah membahas tentang kuartil pada subbab 3.2.2, Ada tiga buah nilai kuartil yaitu kuartil pertama  $(Q_1)$ , kuartil kedua  $(Q_2)$  dan kuartil ketiga  $(Q_3)$ . Ketiga nilai kuartil ini membagi

data yang telah diurutkan ke dalam 4 bagian yang sama besar (lihat gambar 3.1). Ada empat bagian yang terbentuk yaitu bagian 1 dari nilai data terkecil sampai ke nilai  $Q_1$ , bagian kedua dari nilai  $Q_1$  sampai ke  $Q_2$ , bagian ketiga dari nilai  $Q_2$  sampai ke  $Q_3$  dan bagian keempat dari  $Q_3$  sampai ke  $Q_4$ . Pada ukuran penyebaran pada nilai kuartil ini kita akan mencari selisih antara nilai  $Q_1$  dan  $Q_3$ , maknanya disini adalah kita akan melihat tingkat konsentrasi data distribusi tengah (pada bagian kedua dan bagian ketiga) seluas 50% dari seluruh distribusi. Semakin kecil nilai ukuran penyebaran pada nilai kuartil maka semakin tinggi tingkat konsentrasi titiktitik data pada distribusi tengah seluas 50% dari seluruh distribusi.

Beberapa ukuran penyebaran pada nilai kuartil, yaitu

# • Hamparan (H) atau Jangkauan antar kuartil (JK),

$$JK = H = Q_3 - Q_1 \tag{4.10}$$

Prinsip pemikiran pada jangkauan antar kuartil ini hampir sama dengan mencari nilai jangkauan atau rentang pada data tunggal (subbab 3.11.), dimana kita mencari rentang antar nilai tertinggi dan nilai terendah dengan cara mengurangkan nilai tertinggi dengan nilai terendah. Demikian juga pada jangkauan antar kuartil / hamparan ini, kita juga melihat hamparan data dengan mencari rentang nilai kuartil tertinggi  $(Q_3)$  dengan nilai kuartil terendah  $(Q_1)$  dengan cara mengurangkan nilai kuartil tertinggi  $(Q_3)$  dengan nilai kuartil terendah  $(Q_1)$  seperti yang terlihat pada persamaan 4.10.

# • Jangkauan semi antar kuartil atau simpangan kuartil (SK),

$$SK = \frac{Q_3 - Q_1}{2} \tag{4.11}$$

Nilai simpangan kuartil ini merupakan nilai yang menunjukkan setengah kali hamparan yang diperoleh dengan mengurangkan nilai kuartil tertinggi  $(Q_3)$  dengan nilai kuartil terendah  $(Q_1)$  lalu dibagi 2.

# • Rataan Kuartil (RK)

Nilai rataan kuartil diperoleh dengan mencari rata-rata antara nilai kuartil tertinggi  $(Q_3)$  dengan nilai kuartil terendah  $(Q_1)$ .

$$RK = \frac{Q_1 + Q_2}{2} \tag{4.12}$$

# • Rataan Tiga Kuartil (RTK)

Nilai ini diperoleh dengan mencari rata-rata antara tiga nilai kuartil, Q<sub>1</sub>, Q<sub>2</sub> dan Q<sub>3</sub>.

$$RTK = \frac{Q_1 + 2Q_2 + Q_3}{4} \tag{4.13}$$

# DISPERSI PADA KUARTIL UNTUK DATA TUNGGAL KASUS 4.9.

# STATISTIK \$ PROBABILITAS PERTEMUAAN 7. DISPERSI / PENYEBARAN DATA

Data sebagai berikut:

Tentukan:

- a) Hamparan (H) atau Jangkauan antar kuartil (JK),
- b) Jangkauan semi antar kuartil atau simpangan kuartil (SK),
- c) Rataan Kuartil (RK)
- d) Rataan Tiga Kuartil (RTK)

# Penyelesaian:

a) Pertama kita urutkan dari data terkecil ke data terbesar,

n = 7 (ganjil) dan 7+1 = 8 (habis dibagi 4) sehingga kita menggunakan persamaan 3.10.

$$Q_1 = X_{\frac{7+1}{4}} = X_{\frac{8}{4}} = X_2 = 23$$

$$Q_2 = X_{\frac{2(7+1)}{4}} = X_{\frac{16}{4}} = X_4 = 25$$

$$Q_3 = X_{\frac{3(7+1)}{4}} = X_{\frac{24}{4}} = X_6 = 29$$

Sehingga,

• 
$$H = IK = 29 - 23 = 6$$

• 
$$SK = \frac{6}{2} = 3$$

• RK = 
$$\frac{23+29}{2}$$
 = 26

• 
$$RTK = \frac{23 + (2 \times 25) + 29}{4} = 25.5$$

b) Pertama kita urutkan dari data terkecil ke data terbesar,

N=9 (ganjil) dan 9+1=10 (tidak habis dibagi 4) sehingga kita menggunakan persamaan 3.11.

$$Q_{1} = \frac{X_{\frac{n-1}{4}} + X_{\frac{n+3}{4}}}{2} = \frac{X_{\frac{9-1}{4}} + X_{\frac{9+3}{4}}}{2} = \frac{X_{2,5} + X_{3}}{2}$$

$$Q_{1} = \frac{(X_{2} + 0,5(X_{3} - X_{2})) + X_{3}}{2} = \frac{(20 + 0,5(21 - 20)) + 21}{2} = \frac{20,5 + 21}{2} = 20,75$$

$$Q_{2} = X_{\frac{2(n+1)}{4}} = X_{\frac{2(9+1)}{4}} = X_{\frac{20}{4}} = X_{5} = 22$$

$$Q_3 = \frac{X_{3n+1} + X_{3n+5}}{\frac{4}{2}} = \frac{X_{3*9+1} + X_{3*9+5}}{\frac{4}{2}} = \frac{X_7 + X_8}{2} = \frac{27 + 28}{2} = \frac{55}{2} = 27,5$$

Sehingga,

• 
$$H = JK = 29 - 22 = 7$$

• 
$$SK = \frac{29-22}{2} = 3.5$$

• RK = 
$$\frac{22+29}{2}$$
 = 25,5

• 
$$RTK = \frac{22 + (2 \cdot 24, 5) + 29}{4} = 25$$

c) Pertama kita urutkan dari data terkecil ke data terbesar,

n = 10 (genap) dan 10+1 = 11 tidak habis dibagi 4 maka kita menggunakan persamaan 3.13.

$$Q_1 = x_{\frac{10+2}{4}} = x_3 = 22$$

$$Q_2 = \frac{x_{10} + x_{10n}}{2} = x_{5,5} = 24 + 0.5(25 - 24) = 24.5$$

$$Q_3 = x_{\frac{3.10+2}{4}} = x_8 = 29$$

Sehingga,

• 
$$H = IK = 34.30 - 21.5 = 12.8$$

• 
$$SK = \frac{34,30-21,5}{2} = 6,4$$

• RK = 
$$\frac{21,5+34,30}{2}$$
 = 27,9

• 
$$RTK = \frac{21,5 + (2*27,79) + 34,30}{4} = 27,845$$

# 4.7. PENCILAN (OUTLIER)

Dalam statistik, pencilan (outlier) adalah nilai data yang menyimpang sangat jauh dari nilai data lainnya dalam sekelompok data yang ada. Adanya data yang menyimpang sangat jauh ini merupakan suatu peringatan bagi peneliti tentang adanya kesalahan atau ketidaknormalan data yang dapat berasal dari kesalahan alat ukur atau instrumen penelitian ataupun kesalahan dalam menentukan sampel (sampel tidak berasal dari populasi yang homogen).

Untuk menentukan apakah sebuah data termasuk pencilan atau bukan, kita harus melihat apakah nilai data tersebut masih termasuk di dalam Pagar Dalam (PD) atau Pagar Luar (PL). Pagar Dalam (PD) dan Pagar Luar (PL) ini memiliki batas atas dan batas bawah.

• Nilai batas bawah Pagar Dalam (PD) adalah nilai data yang berada satu langkah di bawah nilai kuartil pertama (Q<sub>1</sub>) dan nilai batas atas Pagar Dalam (PD) adalah nilai data yang berada satu langkah di atas nilai kuartil ketiga (Q<sub>3</sub>). Data yang tidak termasuk didalam Pagar Dalam (PD) dinamakan pecilan minor (minor outlier).

• Nilai batas bawah Pagar Luar (PL) adalah nilai data yang berada dua langkah di bawah nilai kuartil pertama (Q<sub>1</sub>) dan nilai batas atas Pagar Luar (PL) adalah nilai data yang berada dua langkah di atas nilai kuartil ketiga (Q<sub>3</sub>). Data yang tidak termasuk did ala Pagar Luar (PL) dinamakan pencilan mayor (mayor outlier).

Data yang berada diluar Pagar Dalam (PD) dan Pagar Luar (PL) ini dapat kita buang dan tidak usah dianalisis.

Misalkan, kita memiliki data pemakaian internet (jam) sebagai berikut:

Pada sekumpulan data diatas ada nilai 100 yang menyimpang sangat jauh dari nilai data-data lainnya. Kita perlu mengenali bahwa nilai ini berpotensi menjadi pencilan. Langkah-langkah yang dapat dilakukan untuk mengenali pencilan ini adalah:

1) Susun data dari data terkecil sampai data terbesar,

2) Kita tentukan nilai kuartil pertama (Q<sub>1</sub>), kuartil kedua (Q<sub>2</sub>) dan kuartil ketiga (Q<sub>3</sub>)

Cara menentukan nilai kuartil ini telah kita bahas pada subbab 3.2.2, karena jumlah data (n) sedikit, hanya sebanyak 9 data maka kita bisa menggunakan gambar piramida, dimana nilai  $Q_1$  terletak diantara data kedua dan data ketiga, nilai  $Q_2$  terletak pada data kelima dan nilai  $Q_3$  terletak diantara data ketujuh dan data kedelapan.

39 **40 40** 41 **44** 44 **45 47** 100 
$$Q_1 = (40+40)/2 = 40$$
  $Q_2 = 44$   $Q_3 = (45+47)/2 = 46$ 

3) Menentukan jarak antar kuartil (JK)

Jarak antar kuartil (JK) telah dibahas pada subbab 4.6., nilai ini ditentukan dengan mengurangkan nilai  $Q_3$  dengan  $Q_1$ , seperti pada persamaan 4.10.

$$JK = H = 46-40 = 6$$

4) Menentukan Pagar Dalam (PD)

kita perlu menentukan batas bawah dan batas atas Pagar Dalam (PD). Untuk itu kita mencari nilai Langkah (L) terlebih dahulu dengan cara mengalikan 1,5 dengan jangkauan antar kuartil (JK) atau hamparan, nilai L diperoleh menggunakan persamaan:

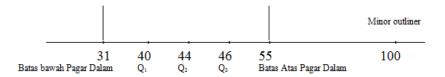
$$L = 1.5(Q_3 - Q_1) \tag{4.14}$$

Sehingga diperoleh,

$$L = 1.5(46 - 40) = 9$$

Batas bawah Pagar Dalam (PD) =  $Q_1 - L = 40 - 9 = 31$ 

Batas atas Pagar Dalam (PD) =  $Q_3 + L = 46+9 = 55$ 



Gambar 4.2. Pagar Dalam (PD) dan Minor Outlier

# 5) Menentukan Pagar Dalam (PD) dan Pagar Luar (PL)

Kita juga perlu menentukan batas bawah dan batas atas Pagar Luar (PDL). Untuk itu kita mencari nilai Langkah (L) terlebih dahulu dengan cara mengalikan 3 dengan jangkauan antar kuartil (JK) atau hamparan, nilai L diperoleh menggunakan persamaan:

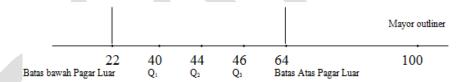
$$L = 3(Q_3 - Q_1) (4.15)$$

Sehingga diperoleh,

$$L = 3(46 - 40) = 18$$

Batas bawah Pagar Luar (PL) =  $Q_1 - L = 40-18 = 22$ 

Batas atas Pagar Luar (PL) =  $Q_3 + L = 46+9 = 64$ 



Gambar 4.3. Pagar Luar (PL) dan Mayor Outlier

Kita lihat nilai rata-rata yang dihasilkan jika nilai data 100 dibuang maka diperoleh nilai rata-rata sebesar 42,5 dan jika nilai data 100 diikut sertakan maka diperoleh nilai rata-rata sebesar 48,9.

$$\bar{x} = \frac{39 + 40 + 40 + 41 + 44 + 44 + 45 + 47 + 100}{9} = 48,9$$

$$\bar{x} = \frac{39 + 40 + 40 + 41 + 44 + 44 + 45 + 47}{8} = 42,5$$

Jelas sekali terlihat bahwa kumpulan data kita terletak dikisaran nilai 39 sampai 47, akan sangat janggal jika nilai rata-ratanya adalah 48,9 diatas nilai 47. Jadi nilai data 100 ini termasuk pencilan mayor, maka data ini harus dihapuskan karena akan dapat mengubah hasil penelitian.

#### KASUS PENYEBARAN DATA DAN PEMBAHASANNYA

# **KASUS 4.10.**

Juni

Juli

Jumlah

PT. Sejahtera memiliki 3 jenis barang yang dijual yaitu menjual perlengkapan bayi, menjual tas, dan menjual sepatu. Keuntungan bersih dari hasil pendapatan dalam 6 bulan terakhir disajikan pada tabel dibawah ini:

Bulan Pendapatan setiap Jenis Usaha (Juta) (2016)Perlengkapan Bayi Tas Sepatu Januari 6 7 10 Februari 7 5 5 Maret 5 6 6 April 4 5 5

6

7

6

5

6

7

Tabel 4.7. Data Awal untuk Kasus 4.10

Perusahaan ini akan mempertahankan 2 usaha dari 3 usaha yang dimilikinya. Tentukan 2 bidang usaha yang akan dilanjutkan !

#### Penyelesaian:

Catatan penting yang harus dipahami adalah usaha yang akan dipertahankan adalah usaha yang pendapatannya paling stabil. Pendapatan yang paling stabil dilihat dari penyebaran data (dispersi) pendapatan pada masing-masing usaha. Pada penjelasan diatas (subbab 4.5), Koefisien Keragaman (KK) digunakan untuk melihat kestabilan sekelompok data untuk itu kita akan menghitung Koefisien Keragaman (KK). Semakin kecil nilai Koefisien Keragaman semakin stabil data yang disajikan. Langkah yang dilakukan adalah :

- 1) Menentukan nilai rata-rata menggunakan persamaan 3.1.
- 2) Menentukan nilai simpangan baku (S) menggunakan persamaan 4.4.
- 3) Dan Terakhir menentukan nilai koefisien keragaman (KK) menggunakan persamaan 4.9.

# • USAHA PERLENGKAPAN BAYI

$$\bar{x} = \frac{6+7+5+4+6+7}{6} = \frac{35}{6} = 5,833$$

Tabel 4.8. Simpangan Baku untuk Usaha Perlengkapan Bayi

Bulan (2016)	Xi	$x_i - \overline{\overline{x}}$	$(x-\overline{\overline{x}})^2$
Januari	6	0,167	0,028
Februari	7	1,167	1,361
Maret	5	-0,833	0,694
April	4	-1,833	3,361
Juni	6	0,167	0,028
Juli	7	1,167	1,361
Jumlah	35		6,833

Diperoleh nilai simpangan baku (S) dan koefisien keragaman (KK) untuk usaha percetakan:

$$S = \sqrt{\frac{6,833}{5}} = 1,169$$

$$KK = \frac{1,169}{5,833} = 0,200$$

#### • USAHA MENJUAL TAS

$$\bar{x} = \frac{7+5+6+5+6+7}{6} = \frac{36}{6} = 6$$

Tabel 4.9. Simpangan Baku Usaha Menjual Tas

Bulan (2016)	Xi	$x_i - \bar{\bar{x}}$	$(x-\bar{\bar{x}})^2$
Januari	7	1	1
Februari	5	-1	1
Maret	6	0	0
April	5	-1	1
Juni	6	0	0
Juli	7	1	1
Jumlah	36		4

Diperoleh nilai simpangan baku (S) dan koefisien keragaman (KK) untuk usaha kuliner :

$$S = \sqrt{\frac{4}{5}} = 0.894$$

$$KK = \frac{0.894}{6} = 0.149$$

# USAHA MENJUAL SEPATU

$$\bar{x} = \frac{10+5+6+5+6+5}{6} = \frac{37}{6} = 6,167$$

Tabel 4.10. Simpangan Baku untuk Usaha Menjual Sepatu

Bulan (2016)	Xi	$x_i - \bar{\bar{x}}$	$(x-\bar{\bar{x}})^2$
Januari	10	3,833	14,694
Februari	5	-1,167	1,361
Maret	6	-0,167	0,028
April	5	-1,167	1,361
Juni	6	-0,167	0,028
Juli	5	-1,167	1,361
Jumlah	37		18,833

Diperoleh nilai simpangan baku (S) dan koefisien keragaman (KK) untuk usaha transportasi :

$$S = \sqrt{\frac{18,833}{5}} = 1,941$$

$$KK = \frac{1,941}{6,167} = 0,315$$

Semakin kecil nilai koefisien keragaman semakin homogen dan semakin stabil suatu kelompok data. Usaha yang paling stabil adalah usaha menjual tas dengan koefisien keragaman = 0,149 kemudian diikuti usaha menjual perlengkapan bayi dengan koefisien keragaman = 0,200 dan yang terakhir adalah usaha menjual sepatu dengan koefisien keragaman = 0,315. Dua usaha yang disarankan untuk dipertahankan adalah usaha menjual tas dan usaha menjual perlengkapan bayi.

# MENGHITUNG UKURAN PENYEBARAN DENGAN MS. EXCEL

# **KASUS 4.11.**

Menentukan jangkauan (range), variansi/ragam, simpangan baku (standar deviasi) dan koefisien keragaman dengan Ms. Excel.

Buat tabel penjualan sebagai berikut:

Tanggal	Sales	Smartphone	Tablet	Laptop
3 januari 2019	Mira Amira	1	4	2
4 januari 2019	Rina Andina	4	3	3
4 januari 2019	Romi Simaroang	4	1	5
4 januari 2019	Rina Andina	21	11	6
15 januari 2019	Rina Andina	12	17	21
16 januari 2019	Ferfy Ramadhan	19	13	12
16 januari 2019	Ferfy Ramadhan	21	9	13
19 januari 2019	Mira Amira	29	5	21
20 januari 2019	Romi Simaroang	1	6	9
20 januari 2019	Ferfy Ramadhan	23	8	7
5 februari 2019	Romi Simaroang	3	2	8
10 februari 2019	Mira Amira	9	4	12
15 februari 2019	Mira Amira	16	6	21
16 februari 2019	Mira Amira	15	15	19
17 febrauri 2019	Rina Andina	11	11	13
19 februari 2019	Romi Simaroang	9	13	6
20 februari 2019	Romi Simaroang	8	9	7
1 maret 2019	Mira Amira	5	5	3
4 maret 2019	Mira Amira	7	3	1
6 maret 2019	Mira Amira	2	7	1
14 maret 2019	Rina Andina	1	6	4
20 maret 2019	Ferfy Ramadhan	6	9	8

20 maret 2019	Romi Simaroang	7	3	9
25 maret 2019	Ferfy Ramadhan	14	2	2

		Smartphone	Tablet	Laptop
Penyebaran	Jangkauan / Range	Jangkauan / Range 28		20
	Variansi untuk Populasi	59,3889	18,4722	39,5260
	Variansi untuk Sampel	61,9710	19,2754	41,2446
	SD untuk Populasi	7,7064	4,2979	6,2870
	SD untuk Sampel	7,8722	4,3904	6,4222
	Koefisien Keragaman	0,7458	0,5997	0,7084

Langkah-langkah yang dilakukan:

1) Tulis **Penyebaran** pada cell C30, tulis Jangkauan / Range pada cell D30, lalu tulis rumus menentukan jangkauan pada cell E30 :

**=MAX(E4:E27)-MIN(E4:E27)** 

Drag dari cell E30 sampai G30;

2) Tulis **Variansi untuk Popuasi** pada cell D31, lalu tulis rumus untuk menentukan variansi untuk populasi pada cell E31 :

=VAR.P(E4:E27)

**Drag** dari cell E31 sampai G31;

3) Tulis **Variansi untuk Sampel** pada cell D32, lalu tulis rumus untuk menentukan variansi untuk sampel pada cell E32 :

=VAR.S(E4:E27)

Drag dari cell E32 sampai G32;

4) Tulis **SD untuk Populasi** pada cell D33, lalu tulis rumus untuk menentukan standar deviasi untuk populasi pada cell E33 :

**=STDEV.P(E4:E27)** 

Drag dari cell E33 sampai G33;

5) Tulis **SD untuk Sampel** pada cell D34, lalu tulis rumus untuk menentukan standar deviasi untuk sampel pada cell E34 :

**=STDEV.S(E4:E27)** 

**Drag** dari cell E34 sampai G34;

6) Tulis **Koefisien Keragaman** pada cell D35, lalu tulis rumus untuk menentukan koefisien keragaman pada cell E35 :

**=E33/AVERAGE(E4:E27)** 

Drag dari cell E35 sampai G35;

#### **KASUS 4.12**

Dari	Ke	Smartphone	Nilai Bukan Nol	Tablet	Nilai Bukan Nol
8:00 AM	9:00 AM	1	1	4	4
9:00 AM	10:00 AM	0		3	3
10:00 AM	11:00 AM	4	4	1	1
11:00 AM	12 noon	43	43	44	44
12 noon	1:00 PM	32	32	52	52
1:00 PM	2:00 PM	57	57	65	65
2:00 PM	3:00 PM	38	38	83	83
3:00 PM	4:00 PM	45	45	87	87
4:00 PM	5:00 PM	76	76	98	98
5:00 PM	6:00 PM	95	95	82	82
6:00 PM	7:00 PM	102	102	92	92
7:00 PM	8:00 PM	98	98	78	78
8:00 PM	9:00 PM	110	110	62	62
9:00 PM	10:00 PM	51	51	54	54
10:00 PM	11:00 PM	12	12	23	23
11:00 PM	12 midnight	2	2	11	11
12 midnight	1:00 AM	0		12	12
1:00 AM	2:00 AM	0		4	4
2:00 AM	3:00 AM	0		1	1
3:00 AM	4:00 AM	0		0	
4:00 AM	5:00 AM	0		0	
5:00 AM	6:00 AM	0		0	
6:00 AM	7:00 AM	1	1	0	
7:00 AM	8:00 AM	0		2	2

	Smartphone	Smartphone	Tablet	Tablet
Jumlah	767		858	
Tidak Ada Traksaksi	16		20	
Rata-rata	48	48	43	45
Standar Deviasi		37		35
Koefisien Keragaman		0,78		1,29

Dapat dilihat bahwa nilai koefisien keragaman penjualan tablet < koefisien keragaman penjualan smartphone

# Langkah yang dilakukan:

1) Pada cell F4, tuliskan rumus untuk menentukan 'jika jumlah transaksi = 0 maka kosongkan cell"

=IF(E4=0;" ";E4)

# STATISTIK \$ PROBABILITAS PERTEMUAAN 7. DISPERSI / PENYEBARAN DATA

Drag dari cell F4 sampai cell F27

2) Pada cell E29, tuliskan rumus untuk menjumlahkan penjualan,

**=SUM(E4:E27)** 

3) Pada cell E30, tuliskan rumus untuk menentukan jumlah jam tanpa transaksi,

**=COUNT(E4:E27) - COUNTIF(E4:E27;0)** 

4) Pada cell E31, tulis rumus untuk menentukan rata-rata transaksi,

=E29/E30

5) Pada cell F31, tuliskan rumus untuk menentukan rata-rata transaksi,

**=AVERAGE(F\$3:F\$26)** 

6) Pada cell F32, tuliskan rumus untuk menentukan standar deviasi,

**=STDEV.P(F\$3:F\$26)** 

7) Pada cell F33, tuliskan rumus untuk menentukan koefisien keragaman,

=F31/F32