

BAB 1

STATISTIKA

STANDAR KOMPETENSI:

Menggunakan aturan statistika, kaidah pencacahan, dan sifat-sifat peluang dalam pemecahan masalah.

Kompetensi Dasar

1. Membaca data dalam bentuk tabel dan diagram batang, garis, lingkaran, dan *ogive*
2. Menyajikan data dalam bentuk tabel dan diagram batang, garis, lingkaran, dan *ogive* serta penafsirannya
3. Menghitung ukuran pemusatan, ukuran letak, dan ukuran penyebaran data, serta penafsirannya

Untuk menyajikan hasil dari suatu pemilihan kepala daerah media masa memilih menggunakan tabel atau diagram, karena tabel dan diagram dianggap paling efektif dan efisien untuk menyampaikan informasi tersebut. Anda tidak mungkin dapat memahami informasi tersebut tanpa memiliki pengetahuan tentang cara-cara menyajikan informasi. Informasi tentang *persentase* kenaikan *rata-rata* harga sembako setelah kenaikan BBM tidak dapat dipahami apabila Anda tidak mengetahui maksud dari istilah presentase dan rata-rata. Membaca dan membuat tabel atau diagram, istilah persentase dan rata-rata merupakan bagian dari pembahasan statistika. Dengan demikian pengetahuan statistika sangat bermanfaat dalam menyajikan dan menafsirkan informasi.

Soal Apersepsi

1. Nyatakanlah $\frac{3}{8}$ dalam persen!
2. Berapa derajat bila memutar jarum jam satu putaran penuh?
3. Suatu lingkaran dibagi menjadi 6 juring yang sama luasnya, berapakah sudut pusat tiap-tiap juring?
4. Apakah 35 termasuk bilangan antara 30 dan 35?
5. Apakah 71 termasuk bilangan mulai 71 sampai dengan 75?

Kontekstual Problem

Setelah pemerintah Indonesia menaikkan harga BBM pada bulan Mei 2008 terjadi kenaikan harga beberapa bahan makanan pokok seperti beras, telur ayam, minyak goreng, dan sebagainya. Harga beras dengan kualitas yang sama pada bulan Mei 2008 tercatat Rp. 5.500,00 sedangkan pada bulan Juli 2008 tercatat Rp. 6.000,00; telur ayam sebelumnya Rp. 9.000,00 per kilogram sekarang menjadi Rp. 14.000,00 per kilogram; daging ayam sebelumnya Rp. 18.000,00 per kilogram sekarang Rp. 22.000,00 per kilogram; minyak goreng sebelumnya Rp. 7.000,00 per liter sekarang harganya menjadi Rp. 13.000,00 per liter; dan daging sapi semula Rp. 48.000,00 per kilogram sekarang menjadi Rp. 50.000,00 per kilogram. Sajikanlah data di atas agar mudah dipahami oleh para pembaca, dan tentukan persentase rata-rata kenaikan harga dari 5 jenis bahan pokok tersebut.

A. Pendahuluan

Statistika adalah suatu ilmu tentang cara-cara mengumpulkan dan menganalisa data, baik data yang bersifat numerik atau data non-numerik. Data yang dikumpulkan mungkin tentang karakteristik individu atau obyek. Karakteristik individu atau obyek itu disebut *variabel*. Berat badan, tinggi badan atau agama yang dianut dari tiap orang merupakan contoh variabel.

Menurut statistika variabel terbagi ke dalam variabel *numerik* dan *non-numerik*. Jenis kelamin, agama yang dianut, dan status pekerjaan merupakan variabel non-numerik, biasa disebut variabel *kategori* atau kelas. Tinggi badan, berat badan, dan besarnya penghasilan merupakan contoh variabel numerik.

Variabel numerik terbagi ke dalam variabel *kontinu*, variabel *diskrit*, dan variabel *ordinal*. Tinggi badan dan berat badan adalah contoh dari variabel kontinu, banyaknya saudara kandung dan banyaknya kendaraan yang dimiliki tiap orang merupakan variabel diskrit, sedangkan tanggal kelahiran merupakan contoh variabel ordinal.

Data biasanya dikumpulkan untuk mencari jawaban dari suatu persoalan tertentu melalui penelitian atau menguji suatu hipotesa. Contoh: (1) Bagaimana tingkat polusi udara di kota Bandung menurut standar Uni Eropa? (2) Pada kelompok umur masyarakat manakah yang paling banyak menggunakan internet? (3) Benarkah siswa di kabupaten A lebih banyak meneruskan pendidikan ke perguruan tinggi dibandingkan dengan siswa dari kabupaten B?

Latihan 1

1. Manakah yang termasuk variabel kategori, variabel diskrit atau variabel kontinu dari data berikut?
 - (a). Banyaknya kesalahan mengetik kata per halaman dari suatu makalah
 - (b). Jenis pekerjaan yang tiap kepala keluarga
 - (c). Banyaknya suara dalam memilih gubernur
 - (d). Banyaknya gempa yang terjadi setiap tahun di Indonesia dari tahun 2000 hingga 2008
 - (e). Berat badan dari sekelompok siswa
 - (f). Banyaknya sepeda motor yang diproduksi dalam setiap bulan
 - (g). Jarak yang ditempuh tiap orang dalam 5 menit
 - (h). Nilai tes mata pelajaran matematika

B. Menyusun dan Menyajikan data

1. Tabel Distribusi Frekuensi

Sebelum dianalisa, data yang telah dikumpulkan biasanya disusun dalam sebuah tabel distribusi frekuensi.

Tabel untuk Variabel Kategori

Contoh 1.1

Di tempat parkir suatu sekolah terdapat 50 buah mobil. Adapun warna-warna mobil itu terdiri dari warna merah, hitam, perak, putih, biru, dan warna lainnya. Data tersebut disajikan dalam Tabel 1.1 berikut ini.

Tabel 1.1
Distribusi Warna 50 Mobil
di Suatu Tempat Parkir

Warna Mobil	Banyaknya Mobil (frekuensi)
Hitam	11
Biru	6
Merah	6
Perak	8
Putih	15
Lain-lain	4

Tabel untuk Variabel Diskrit

Contoh 1.2

Banyak saudara kandung dari 40 siswa adalah sebagai berikut.

Tabel 1.2
Distribusi Saudara Kandung 50 siswa

Banyaknya Saudara Kandung	Jumlah Siswa (frekuensi)
0	5
1	21
2	9
3	3
4	2

Tabel untuk Variabel Kontinu

Contoh 1.3

Tinggi 50 siswa pada saat memasuki SMA diukur pada satuan cm adalah sebagai berikut.

Tabel 1.3
Distribusi Frekuensi Tinggi 50 siswa

Tinggi	Banyaknya siswa (frekuensi)
150 – 154	3
155 – 159	9
160 – 164	21

165 – 169	13
170 – 174	4

Pada tabel distribusi untuk variabel kontinu beberapa istilah yang perlu diketahui.

- (1) *Banyak kelas*, tabel di atas terdiri dari 5 kelas, yaitu kelas 150 – 154, 155 – 159, 160 – 164, 165 – 169, dan 170 – 174.
- (2) *Batas bawah* dan *batas atas* tiap kelas, perhatikan kelas 150 – 154 dan 155 – 159. Batas bawah kelas 155 – 159 adalah $(154 + 155)/2 = 154,5$, sekaligus menjadi batas atas kelas 150 – 154.
- (3) *Panjang kelas* atau *interval* adalah selisih batas atas dan batas bawah kelas
- (4) *Titik tengah kelas* adalah jumlah batas atas dan batas bawah dibagi dua.

Untuk memahami istilah-istilah tersebut perhatikan Tabel 1. 4 berikut ini

Tabel 1.4

Tinggi	Batas Bawah	Batas Atas	Titik Tengah	Panjang Kelas (interval)	Frekuensi
150 – 154	149,5	154,5	152	5	3
155 – 159	154,5	159,5	157	5	9
160 – 164	159,5	164,5	162	5	21
165 – 169	164,5	169,5	167	5	13
170 – 174	169,5	174,5	172	5	4

Cara Menyusun Distribusi Frekuensi variabel kontinu

- (1) Tentukan banyak kelas. Bila banyak datanya antara 40 hingga 100, maka banyak kelas adalah 5 hingga 7. Untuk data yang lebih besar dari 100 biasa digunakan aturan Sturges, yaitu banyak kelas $k = 1 + \log n$, dengan n adalah banyak data. atau 6 kelas.
- (2). Menentukan panjang kelas, diusahakan setiap kelas memiliki panjang kelas yang sama. Panjang kelas diperoleh dari selisih data terbesar dan data terkecil dibagi banyak kelas.
- (3) Menyusun interval tiap-tiap kelas sehingga semua data tercakup pada distribusi frekuensi tersebut.

Contoh 1.4

Buatlah Distribusi frekuensi dari nilai tes matematika dari 50 siswa sebagai berikut.

50	52	65	67	73	92	80	67	63	70
55	57	60	73	77	83	85	75	80	50
95	65	67	70	73	77	65	83	87	67
63	73	70	75	67	77	75	90	65	67
73	60	57	80	63	65	57	70	70	60

1. Banyak data $n = 50$, kita tentukan banyak kelas 6

2. Data terbesar adalah 95 dan data terkecil 50, selisihnya 45. Panjang kelas atau interval adalah $45/6 = 7,5$ di bulatkan ke atas menjadi 8
3. Setelah kita menetapkan buat tabel untuk mengetahui frekuensi lakukan penurunan, sehingga diperoleh Tabel 1.5 berikut.

Tabel 1.5
Distribusi Frekuensi Nilai Tes Matematika 50 siswa

Nilai Tes matematika	Tally (Turus)	Frekuensi
50 – 57	///// -/	6
58 – 65	///// -///// - //	12
66 – 73	///// -///// - ///// /	16
74 – 81	///// - /////	9
82 – 89	////	4
90 – 97	///	3

Tabel distribusifrekuensi kumulatif

Tabel distribusi frekuensi variabel kontinu dapat dibuat tabel distribusi kumulatif. Tabel distribusi kumulatif terbagi dua yaitu “*lebih dari*” atau “*lebih dari*”.

Contoh 1. 5

Buatlah tabel distribusi frekuensi kumulatif dari Tabel 1.3 di atas

Tabel 1.3
Distribusi Frekuensi Tinggi 50 siswa

Tinggi	Banyaknya siswa (frekuensi)
150 – 154	3
155 – 159	9
160 – 164	21
165 – 169	13
170 – 174	4

Untuk membuat tabel distribusi frekuensi kumulatif “*lebih dari*” , perlu dicari batas bawah dari setiap kelas untuk mengisi kolom setiap kelas. Untuk memahaminya perhatikan Tabel 1.6 berikut.

Tabel 1.6
**Distribusi Frekuensi Kumulatif “lebih dari”
Tinggi 50 siswa**

Tinggi	Banyaknya siswa (frekuensi)
Lebih dari 149,5	50
Lebih dari 154,5	47
Lebih dari 159,5	38
Lebih dari 164,5	17

Lebih dari 169,5	4
------------------	---

Untuk membuat tabel distribusi frekuensi kumulatif “kurang dari” , perlu dicari batas atas dari setiap kelas untuk mengisi kolom setiap kelas. Untuk memahaminya perhatikan Tabel 1.7 berikut.

Tabel 1.7
Distribusi Frekuensi Kumulatif “Kurang dari”
Tinggi 50 siswa

Tinggi	Banyaknya siswa (frekuensi)
Kurang dari 154,5	3
Kurang dari 159,5	12
Kurang dari 164,5	33
Kurang dari 169,5	46
Kurang dari 174,5	50

Latihan 2

1. Buatlah distribusi frekuensi dari hasil nilai tes matematika dari 40 siswa berikut ini.

80	62	53	76	59	78	84	66	71	50
79	69	64	56	65	58	78	75	60	71
68	56	79	73	48	83	64	58	75	91
80	59	55	73	81	62	64	69	51	45

2. Buatlah distribusi kumulatif “lebih dari” dan “kurang dari” dari distribusi berikut.

Tabel 1.8
Distribusi Frekuensi Berat 50 siswa

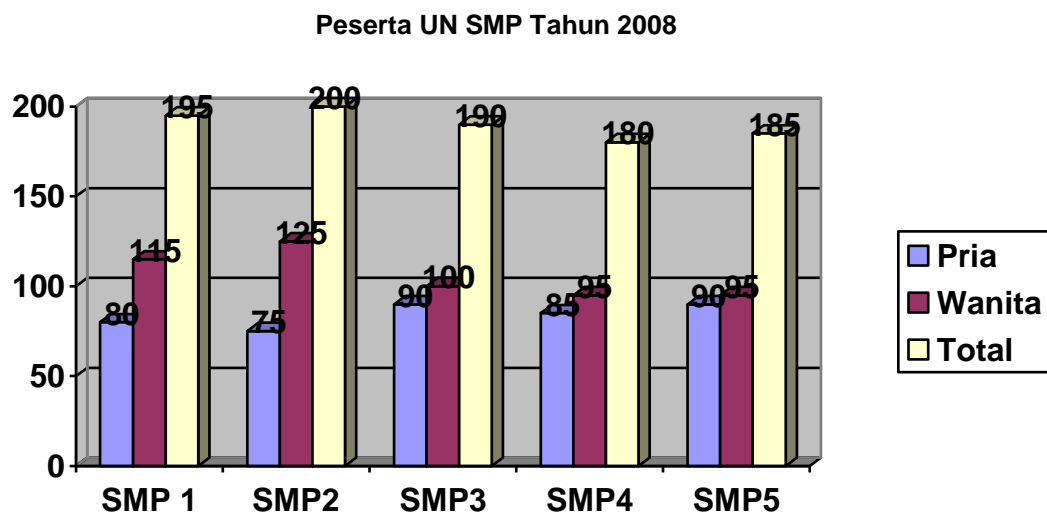
Berat Badan (Kg)	Banyaknya siswa (frekuensi)
40 – 44	5
45 – 49	12
50 – 54	23
55 – 59	15
60 – 64	6
65 – 69	4

2. Jenis-jenis Diagram

Diagram digunakan untuk menampilkan menjadi efektif dan menarik. Ada berbagai jenis diagram untuk menampilkan suatu data, antara lain diagram batang, diagram garis, diagram lingkaran, dan lain sebagainya. Jenis diagram yang digunakan untuk menyajikan data itu disesuaikan dengan keperluan dan jenis variabel dari data tersebut. *Diagram Batang* batang biasa digunakan untuk menampilkan data dengan variabel tidak kontinu, yaitu data dengan variabel kategori atau diskrit. Diagram batang untuk variabel kontinu disebut *histogram*. Diagram garis biasa digunakan perkembangan data dari waktu ke waktu, sedangkan untuk menonjolkan proporsi dari data dengan variabel kategori biasa digunakan diagram lingkaran. Diagram garis untuk distribusi frekuensi kumulatif disebut *ogive*.

Contoh 1.5

Diagram batang banyaknya peserta UN tahun 2008 dari lima SMP sebagai berikut.

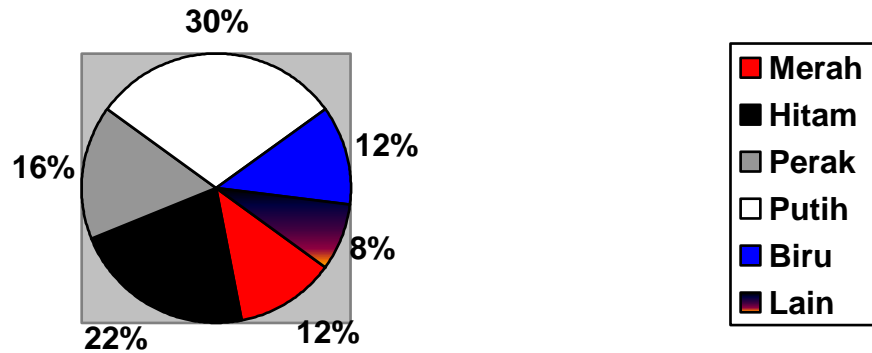


Gambar 1.2

Contoh 1.6.

Di tempat parkir suatu sekolah terdapat 50 buah mobil. Adapun warna-warna mobil itu terdiri dari warna merah, hitam, perak, putih, biru, dan warna lainnya. Data tersebut disajikan dalam diagram lingkaran sebagai berikut.

Warna Mobil di Tempat Parkir

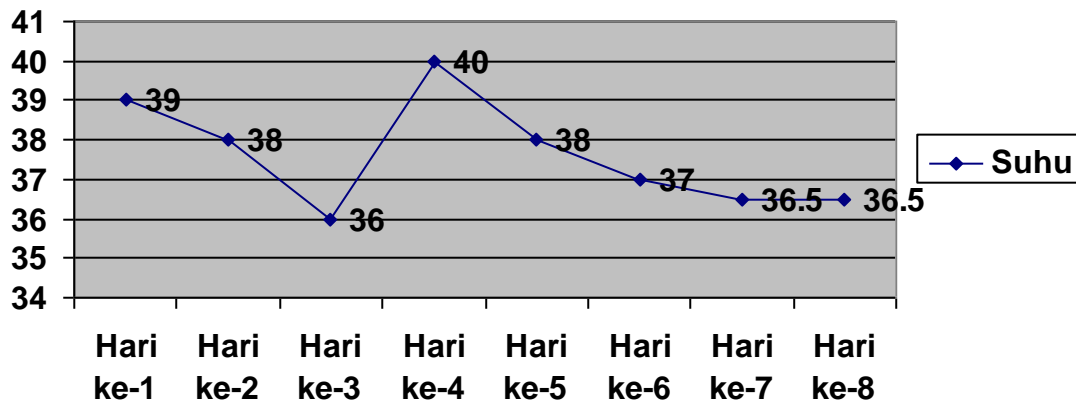


Gambar 1.3

Contoh 1.7

Suhu seorang pasien sejak hari pertama hingga hari ke delapan disajikan dalam diagram garis berikut.

Suhu Seorang Pasien Demam Berdarah

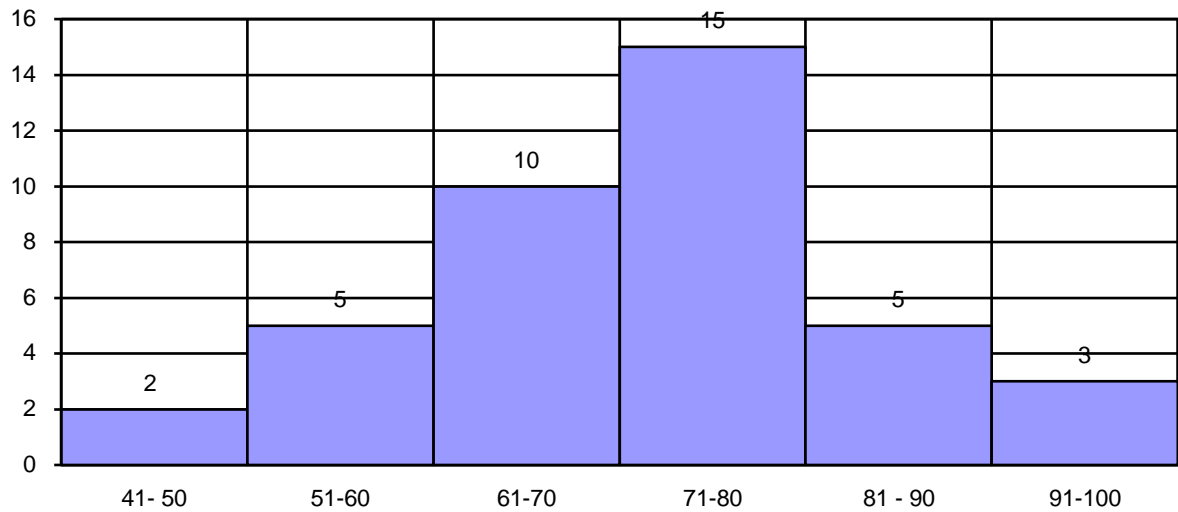


Gambar 1.4

Contoh 1.8

Nilai tes matematika dari 40 siswa disajikan dalam histogram berikut.

Nilai Tes Matematika



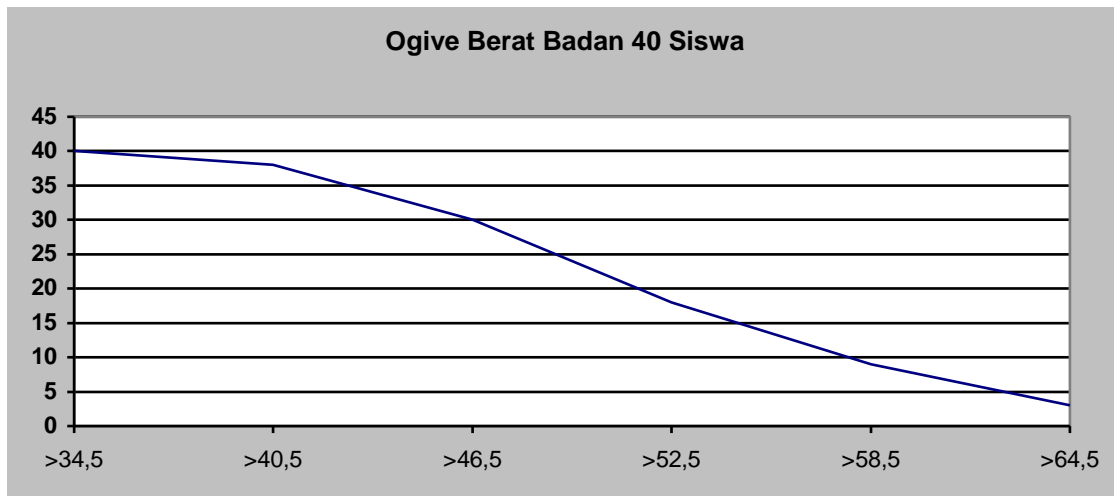
Gambar 1.5

Contoh 1. 9.

Data distribusi frekuensi kumulatif “lebih dari” pada Tabel 1.9 dapat ditampilkan dalam dalam Gambar 1.6 yang disebut ogive..

Tabel 1.9
Frekuensi kumulatif “lebih dari”
Berat badan 40 siswa

Berat Badan (Kg)	Frekuensi Kumulatif
Lebih dari 34,5	40
Lebih dari 40,5	38
Lebih dari 46,5	30
Lebih dari 52,5	18
Lebih dari 58,5	9
Lebih dari 64,5	3



Gambar 1.6.

Latihan 3

1. Data siswa suatu sekolah dari tahun 2001 hingga tahun 2007

Tahun	2001		2002		2003		2004		2005		2006		2007	
Jenis Kelamin	L	P	L	P	L	P	L	P	L	P	L	P	L	P
Banyak siswa	50	70	110	130	180	220	190	210	230	250	250	270	240	280
120		240		400		400		480		520		520		

Buatlah diagram batang dari data tersebut!

. Warna mobil siswa yang diparkir di halaman sekolah adalah sebagai berikut.

Warna	Merah	Hitam	Silver	Putih	Biru	Lainnya
Banyak mobil	35	4	24	20	28	9

Buatlah diagram lingkaran dari data tersebut

3. Pasokan dan permintaan minyak tanah di suatu kelurahan dalam satuan liter

	Januari	Februari	Maret	April	Mei	Juni
Pasokan	12.000	10.000	10.000	8.000	6.000	6.000
Permintaan	10.000	11.000	8.000	10.000	8.000	7.000

Buatlah diagram garis dari data tersebut.

4. Tinggi badan siswa dari 50 siswa adalah sebagai berikut

Tinggi Badan	Frekuensi
140 – 145	2
146 – 151	7
152 – 157	13
158 – 163	15
164 – 169	10
170 - 175	3
	50

- Buatlah histogram dari data tersebut.
- Buatlah frekuensi kumulatif “kurang dari”, dan ogive data tersebut.

C. Ukuran Statistik

1. Ukuran Pemusatan Data

Ukuran pemusatan adalah terbagi tiga macam yaitu, mean (rata-rata/rataan), modus dan median.

Jika suatu himpunan n data bernilai $x_1, x_2, x_3, \dots, x_n$, maka nilai mean dari data tersebut

adalah $\frac{\sum x_i}{n}$

Modus adalah data yang paling banyak muncul, sedangkan median adalah data yang terletak di tengah setelah data diurutkan dari yang terkecil hingga terbesar.

Contoh 1. 10

Sebanyak 15 siswa diminta untuk menyelesaikan sebuah soal yang sederhana. Waktu yang digunakan tiap siswa dalam satuan detik adalah sebagai berikut.

21, 29, 25, 18, 25, 27, 23, 35, 42, 31, 34, 47, 25, 24, 19.

Tentukan mean, modus, dan median dari data tersebut.

Jawab:

$$\text{Mean} = \frac{\sum x_i}{n} = \frac{21 + 29 + 25 + 18 + 25 + 27 + 23 + 35 + 42 + 31 + 34 + 47 + 25 + 24 + 19}{15} = \frac{425}{15} = 28,1$$

Bila data di atas diurutkan dari data terkecil hingga terbesar adalah sebagai berikut.

18, 19, 21, 23, 24, 25, 25, 25, 27, 29, 31, 34, 35, 42, 47. Data yang terletak di tengah dari 15 data adalah data ke- 8 yaitu 25. Jadi median data tersebut adalah 25

Pada data di atas, data yang paling banyak muncul adalah 25, maka modus data tersebut adalah 25.

Untuk data yang disusun dalam bentuk distribusi frekuensi, untuk memperoleh ukuran pemusatan biasa digunakan rumus-rumus sebagai berikut.

$$\text{Mean} = \frac{\sum f_i x_i}{n},$$

x_i : titik tengah interval kelas ke-i

f_i : frekuensi data kelas ke-i

n : banyak data

$$\text{Median} = L + C \left(\frac{\frac{n}{2} - f_c}{f_d} \right)$$

L : batas bawah kelas yang memuat median

C : panjang interval

f_d : fekuensi kelas yang memuat median

f_c : frekuensi kumulatif sebelum kelas median

n : banyak data

$$\text{Modus} = L + C \left(\frac{s_1}{s_1 + s_2} \right)$$

L : batas bawah kelas yang memuat modus

C : panjang interval

s_1 : selisih frekuensi kelas modus dan kelas sebelumnya

s_2 : selisih frekuensi kelas modus dan kelas sesudahnya

Contoh 1.11

Tentukan mean, median dan modus dari dari distribusi frekuensi berat badan 40 siswa seperti berikut.

Tabel 1.10

Berat Badan (Kg)	Frekuensi
35 – 40	2
41 – 46	8
47 – 52	12
53 – 58	9
59 – 64	6
65 – 70	3

Jawab:

Untuk mencari nilai mean diperlukan titik tengah masing-masing kelas, dan hasil kali titik tengah dengan frekuensinya seperti terlihat dalam tabel 1. 11 berikut.

Tabel 1.11

Berat Badan (Kg)	Titik Tengah x_i	Frekuensi f_i	$f_i x_i$
35 – 40	37,5	2	75
41 – 46	43,5	8	348
47 – 52	49,5	12	594
53 – 58	55,5	9	499,5
59 – 64	61,5	6	369
65 – 70	67,5	3	202,5
		40	2088

$$\text{Mean} = \frac{\sum f_i x_i}{n} = \frac{2088}{40} = 52,2$$

Untuk menentukan median perhatikan Tabel 1.12., data yang di tengah adalah data ke 20 dan ke 21 yang terletak pada kelas 47 -52

Tabel 1.12

Berat Badan (Kg)	Frekuensi	
35 – 40	2	
41 – 46	8	
47 – 52	12	Kelas yang memuat data median
53 – 58	9	
59 – 64	6	
65 – 70	3	

Batas bawah kelas median $L = 46,5$, panjang interval $C = 6$, frekuensi kelas yang memuat data median adalah $f_d = 12$, sedangkan frekuensi kumulatif sebelum kelas median adalah $f_c = 10$, dan $n = 40$

$$\text{Median} = L + C \left(\frac{\frac{n}{2} - f_c}{f_d} \right) = 46,5 + 6 \frac{\frac{40}{2} - 10}{12} = 46,5 + \frac{10}{12} = 46,5 + 4,98 = 51,48$$

Untuk menentukan modus perhatikan Tabel 1.13., kelas yang frekuensinya paling banyak adalah disebut kelas modus yaitu kelas 47 -52 dengan frekuensi 12.

Tabel 1.13

Berat Badan (Kg)	Frekuensi	
35 – 40	2	
41 – 46	8	
47 – 52	12	Kelas yang memuat data modus
53 – 58	9	
59 – 64	6	
65 – 70	3	

Batas bawah kelas modus $L = 46,5$, panjang interval $C = 6$, selisih frekuensi kelas modus dengan kelas sebelumnya adalah $s_1 = 4$, sedangkan selisih kelas modus dengan kelas sesudahnya adalah $s_2 = 3$.

$$\text{Modus} = L + C \left(\frac{s_1}{s_1 + s_2} \right) = 46,5 + 6 \frac{4}{4 + 3} = 46,5 + 3,43 = 49,93$$

2. Kuartil dan Desil dari Himpunan Data

Sebelumnya kita telah mengetahui median dari suatu himpunan data, adalah data yang terletak di tengah setelah diurutkan dari data terkecil hingga terbesar. Data ke-seperempat disebut kuartil pertama dilambangkan dengan K_1 , data ke dua-perempat atau setengah disebut kuartil kedua sama dilambangkan sebagai K_2 , sedangkan data ke tiga-perempat disebut kuartil ketiga dilambangkan K_3 . Dalam hal ini dapat disimpulkan bahwa K_2 sama dengan median.

Contoh 1.12

Tentukan kuartil pertama, kedua dan ketiga dari data berikut.

21, 29, 25, 18, 25, 27, 23, 35, 42, 31, 34, 47, 25, 24, 19.

Jawab:

Bila diurutkan dari data terkecil hingga terbesar diperoleh susunan sebagai berikut.

18, 19, 21, 23, 24, 25, 25, 25, 27, 29, 31, 34, 35, 42, 47

K_1	K_2	K_3

Data kuartil pertama adalah 23, kuartil kedua = median adalah 25, dan kuartil ketiga adalah 34.

Selanjutnya dari suatu himpunan data setelah diurutkan dari terkecil hingga terbesar, dapat dicari data yang terletak pada satu-persepuluh, dua-persepuluh dan seterusnya. Data yang terletak pada satu-persepuluh disebut desil pertama, data yang terletak pada dua-persepuluh disebut desil kedua dan seterusnya. Dapat disimpulkan bahwa desil ke-5 sama dengan kuartil kedua sama dengan median. Desil ke i dilambangkan dengan D_i dengan $i = 1, 2, \dots, 9$

Contoh 1.13

Tentukan desil keempat dari data berikut.

21, 29, 25, 18, 25, 27, 23, 35, 42, 31, 34, 47, 25, 24, 19.

Jawab:

Setelah diurutkan dari data terkecil hingga terbesar diperoleh susunan sebagai berikut.

18, 19, 21, 23, 24, 25, 25, 25, 27, 29, 31, 34, 35, 42, 47

Desil keempat adalah data ke empat-persepuluh dari 15 data yaitu data ke $\frac{4}{10}$ dari 15 data yaitu data ke 6. Jadi D_4 adalah 25

Untuk menentukan kuartil dan desil untuk data dari suatu distribusi frekuensi biasa digunakan rumus-rumus sebagai berikut.

$$\text{Rumus kuartil ke } i \text{ adalah } K_i = L + C \left(\frac{\frac{n}{4}i - f_c}{f_d} \right)$$

L: batas bawah kelas yang memuat data kuartil ke $\frac{n}{4}i$

C : panjang interval

f_d : frekuensi kelas yang memuat data kuartil ke- $\frac{n}{4}i$

f_c : frekuensi kumulatif kelas yang terletak sebelum kelas kuartil ke- $\frac{n}{4}i$

n : banyak data

$$\text{Rumus desil ke } i \text{ adalah } D_i = L + C \left(\frac{\frac{n}{10}i - f_c}{f_d} \right)$$

L: batas bawah kelas yang memuat data desil ke $\frac{n}{10}i$

C : panjang interval

f_d : frekuensi kelas yang memuat data desil ke- $\frac{n}{10}i$

f_c : frekuensi kumulatif kelas yang terletak sebelum kelas desil ke- $\frac{n}{10}i$

n : banyak data

Contoh 1.14

Tentukan kuartil ke-3 dan desil desil ke 2 dari data berikut.

Tabel 1.14

Berat Badan (Kg)	Frekuensi
35 – 40	2
41 – 46	8
47 – 52	12
53 – 58	9
59 – 64	6
65 – 70	3

Jawab:

Data kuartil ketiga dari 40 data adalah data ke $\frac{40}{4} \times 3 = 30$

Data yang ke -30 terletak pada kelas 53 -58 dengan frekuensi $9 = f_d$

Kelas kuartil ke-3 memiliki batas bawah $52,5 = L$ dengan panjang interval $C = 6$,
sedangkan frekuensi kumulatif sebelum kelas kuartil ke-3 adalah $f_c = 22$

$$\text{Jadi } K_3 = 52,5 + 6 \cdot C \left(\frac{\frac{40}{4} 3 - 22}{9} \right) = 52,5 + 6 \cdot \frac{30 - 22}{9} = 52,5 + 49/9 = 52,5 + 6,125 =$$

58,625

Data desil ke-2 dari 40 data adalah data ke $\frac{40}{10} \times 2 = 8$.

Data yang ke -8 terletak pada kelas 41 -46 dengan frekuensi $8 = f_d$

Kelas desil ke-2 memiliki batas bawah $40,5 = L$ dengan panjang interval $C = 6$,
sedangkan frekuensi kumulatif sebelum kelas desil ke-2 adalah $f_c = 2$

$$\text{Jadi } D_2 = 40,5 + 6 \left(\frac{\frac{40}{10} 2 - 2}{8} \right) = 40,5 + 6 \cdot \frac{6}{8} = 40,5 + 4,5 = 45$$

Latihan 4

1. Waktu yang diperlukan (dalam satuan hari) penyembuhan dari 30 pasien yang suatu penyakit adalah sebagai berikut.

5	12	7	24	1	23	15	20	23	15
20	5	13	15	16	9	2	13	34	21
19	12	2	13	12	10	3	4	6	35

- Tentukan mean, median dan modus dari data tersebut.
- Tentukan kuartil ke -1 dan ke-3 data tersebut.
- Tentukan desil ke-3 dan ke – 8 data tersebut.

2. Perhatikan data tentang tinggi badan dari 50 siswa dalam satuan cm

Tinggi Badan	Frekuensi
140 – 145	2
146 – 151	7
152 – 157	13
158 – 163	15
164 – 169	10
170 - 175	3
	50

Carilah mean, median dan modus dari data tersebut.

3. Ukuran Penyebaran Data

Beberapa ukuran penyebaran data yang biasa digunakan antara lain, jangkauan, simpangan kuartil, variansi, dan simpangan baku. Jangkauan adalah selisih data terbesar dengan data terkecil, simpangan kuartil adalah $\frac{1}{2} (K_3 - K_1)$. Simpangan baku adalah akar pangkat dua dari variansi, sedangkan variansi memiliki rumus sebagai berikut.

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

Contoh 1.15

Tentukan jangkauan, simpangan kuartil, variansi dan simpangan baku dari data berikut.
21, 29, 22, 18, 25, 27, 23, 35, 42, 31, 34, 45, 25, 24, 19.

Jawab:

Setelah diurutkan data di atas diperoleh

18, 19, 21, 22, 23, 24, 25, 25, 27, 29, 31, 34, 35, 42, 45.

Jangkauan = $45 - 18 = 27$

$K_1 = 22$ dan $K_3 = 34$, maka simpangan kuartil = $\frac{1}{2} (34 - 22) = 6$

Untuk memperoleh variansi kita perlu mencari dulu mean dari data tersebut.

$$\bar{x} = \frac{18 + 19 + 21 + 22 + 23 + 24 + 25 + 25 + 27 + 29 + 31 + 34 + 35 + 42 + 45}{15}$$

$$\frac{420}{15} = 28$$

$$(x_1 - \bar{x})^2 = (18 - 28)^2 = (-10)^2 = 100$$

$$(x_2 - \bar{x})^2 = (19 - 28)^2 = (-9)^2 = 81$$

$$(x_3 - \bar{x})^2 = (21 - 28)^2 = (-7)^2 = 49$$

$$(x_4 - \bar{x})^2 = (22 - 28)^2 = (-6)^2 = 36$$

$$(x_5 - \bar{x})^2 = (23 - 28)^2 = (-5)^2 = 25$$

$$(x_6 - \bar{x})^2 = (24 - 28)^2 = (-4)^2 = 16$$

$$(x_7 - \bar{x})^2 = (25 - 28)^2 = (-3)^2 = 9$$

$$\begin{aligned}
(x_8 - \bar{x})^2 &= (25 - 28)^2 = (-3)^2 = 9 \\
(x_9 - \bar{x})^2 &= (27 - 28)^2 = (-1)^2 = 1 \\
(x_{10} - \bar{x})^2 &= (29 - 28)^2 = (1)^2 = 1 \\
(x_{11} - \bar{x})^2 &= (31 - 28)^2 = (3)^2 = 9 \\
(x_{12} - \bar{x})^2 &= (34 - 28)^2 = (6)^2 = 36 \\
(x_{13} - \bar{x})^2 &= (35 - 28)^2 = (7)^2 = 49 \\
(x_{14} - \bar{x})^2 &= (42 - 28)^2 = (14)^2 = 196 \\
(x_{15} - \bar{x})^2 &= (45 - 28)^2 = (17)^2 = 289
\end{aligned}$$

Variansi dari data tersebut adalah

$$\begin{aligned}
s^2 &= \frac{\sum_{i=1}^{15} (x_i - \bar{x})^2}{n-1} = \frac{100+81+49+36+25+16+9+9+1+1+9+36+49+196+289}{14} \\
&= \frac{857}{14} = 61,2142
\end{aligned}$$

Simpangan baku dari data tersebut $s = \sqrt{s^2} = \sqrt{61,2142} = 7,82$.

Untuk mencari variansi data yang cukup banyak, misalnya biasa digunakan rumus berikut,

$$s^2 = \frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)}$$

$$\begin{array}{lll}
x_1 = 18 & \Rightarrow & x_1^2 = 324 \\
x_2 = 19 & \Rightarrow & x_2^2 = 361 \\
x_3 = 21 & \Rightarrow & x_3^2 = 441 \\
x_4 = 22 & \Rightarrow & x_4^2 = 484 \\
x_5 = 23 & \Rightarrow & x_5^2 = 529 \\
x_6 = 24 & \Rightarrow & x_6^2 = 576 \\
x_7 = 25 & \Rightarrow & x_7^2 = 625 \\
x_8 = 25 & \Rightarrow & x_8^2 = 625 \\
x_9 = 27 & \Rightarrow & x_9^2 = 729 \\
x_{10} = 29 & \Rightarrow & x_{10}^2 = 841 \\
x_{11} = 31 & \Rightarrow & x_{11}^2 = 961 \\
x_{12} = 34 & \Rightarrow & x_{12}^2 = 1156 \\
x_{13} = 35 & \Rightarrow & x_{13}^2 = 1225 \\
x_{14} = 42 & \Rightarrow & x_{14}^2 = 1764 \\
x_{15} = 45 & \Rightarrow & x_{15}^2 = 2025
\end{array}$$

$$\begin{aligned}
\sum x_i &= 420 & \sum x_i^2 &= 12666 \\
(\sum x_i)^2 &= 420^2 = 176400
\end{aligned}$$

$$s^2 = \frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)} = \frac{15.12666 - 176400}{15.14} = \frac{189990 - 176400}{210} = \frac{13590}{210} = 64,71$$

Simpangan baku $s = \sqrt{64,71} = 8,04$

Kedua rumus itu menghasilkan nilai variansi dan simpangan baku yang tidak jauh berbeda, apalagi bila datanya lebih banyak lagi.

Latihan 5

1. Waktu yang diperlukan (dalam satuan hari) penyembuhan dari 20 pasien yang suatu penyakit adalah sebagai berikut.

5	12	7	24	15	20	15	20	5
13	15	16	9	21	12	12	10	6

- Tentukan simpangan kuartil dari data tersebut.
- Tentukan variansi dan simpangan baku data tersebut.

2. Perhatikan data tentang tinggi badan dari 50 siswa dalam satuan cm

Tinggi Badan	Frekuensi
140 – 145	2
146 – 151	7
152 – 157	13
158 – 163	15
164 – 169	10
170 - 175	3
	50

Carilah variansi dan simpangan baku dari data tersebut.

Perdalam konsepmu!

- A. Jelaskan bedanya variabel diskrit dengan variabel kontinu!
- B. Diagram manakah yang paling tepat untuk menyajikan data perkembangan nilai tukar mata uang rupiah dengan dolar Amerika dari hari ke hari?
- C. Ukuran pemusatan data manakah yang paling tepat untuk digunakan untuk mengetahui trend model pakaian yang digunakan tahun 2007?
- D. Ukuran penyebaran data manakah yang paling tepat untuk digunakan untuk mengetahui simpangan penghasilan kelas menengah di suatu kelurahan.

RANGKUMAN

1. Statistika adalah suatu ilmu tentang cara-cara mengumpulkan dan menganalisa data.
2. Data terbagi ke dalam data dengan variabel numerik dan variabel non-numerik, variabel variabel numerik terbagi sebagai variabel diskrit dan kontinu.
3. Diagram yang sering digunakan untuk menyajikan data adalah diagram batang, diagram garis, dan diagram lingkaran.
4. Diagram batang untuk menyajikan data yang disusun dalam distribusi frekuensi disebut histogram.
5. Diagram garis untuk menyajikan data yang disusun dalam distribusi frekuensi kumulatif disebut ogive.
6. Ukuran pemusatan data terdiri dari mean (rata-rata/rerata/rataan), median dan modus.
7. Mean adalah jumlah nilai data dibagi banyaknya data
8. Modus adalah data yang paling banyak muncul.
9. Median adalah data yang terletak di tengah-tengah setelah data diurutkan dari terkecil hingga terbesar.
10. Kuartil ke 1 (Q_1), kuartil ke 2 (Q_2) dan kuartil ke 3 (Q_3), berturut-turut adalah data yang ke satu -perempat, dua-perempat, dan tiga-perempat dari keseluruhan data setelah diurutkan dari terkecil hingga terbesar.
11. Desil ke 1 (D_1), desil ke 2 (D_2) dan seterusnya hingga desil ke 9 (D_9), berturut-turut adalah data yang ke satu -persepuluh, dua-persepuluh, dan sembilan -persepuluh dari keseluruhan data setelah diurutkan dari terkecil hingga terbesar.
12. Mean untuk data yang disusun dalam distribusi frekuensi

$$\text{Mean} = \frac{\sum f_i x_i}{n},$$

13. Median untuk data yang disusun dalam distribusi frekuensi.

$$\text{Median} = L + C \left(\frac{\frac{n}{2} - f_c}{f_d} \right)$$

14. Modus untuk data yang disusun dalam distribusi frekuensi.

$$\text{Modus} = L + C \left(\frac{s_1}{s_1 + s_2} \right)$$

15. Ukuran penyebaran data yang biasa digunakan adalah range, simpangan kuartil, simpangan baku.

16. Range adalah selisih data terbesar dengan data terkecil.

17. Simpangan kuartil = $\frac{1}{2} (Q_3 - Q_1)$.

18. Simpangan baku untuk data yang relatif sedikit, $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$.

Simpangan baku untuk data yang disusun dalam distribusi frekuensi,

$$s = \sqrt{\frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)}}$$