DIKTAT KULIAH

Statistika Dasar

DOSEN

ARI MULYOTO, S.Pd. M.Si.

Bab 1

Pendahuluan

1.1. Pengertian Statistik

Statistik adalah kumpulan keterangan berbentuk angka-angka yang disusun, diatur, dan disajikan dalam bentuk daftar, tabel, atau disertai dengan gambar-gambar yang dinamakan grafik atau diagram.

Contoh 1.1. : Statistik penduduk, statistik pertanian, dll.

1.2. Pengertian Statistika

Statistika adalah pengetahuan yang berhubungan dengan pengumpulan dan pengolahan data, penarikkan kesimpulan dan pengambilan keputusan berdasarkan analisis data yang dikumpulkan tersebut.

Statistika modern mempunyai dua aspek utama, yaitu :

a. *Metode statistika*

   Adalah cara mengumpulkan, mengolah dan menafsirkan data yang berbentuk angka.

b. *Teori statistika*

   Adalah cabang ilmu matematika yang diterapkan (*Applied Mathematics*) dengan menggunakan aksioma-aksioma untuk menyelidiki persoalan-persoalan mengenai perencanaan eksperimen.

Dilihat dari segi pengerjaannya, statistika dibagi dalam dua bagian, yaitu :

a. *Statistika deskriptif*

   Membahas masalah penyusunan data dalam daftar atau tabel, pembuatan dan penyajian data dalam bentuk diagram, dan menarik kesimpulan.

b. *Statistika induktif* (*statistik inferensial*)

   Menarik kesimpulan secara umum dari data yang sudah berbentul tabel, membuat ramalan (*forecasting*), dan menjelaskan sebab akibatnya.

1.3. Kegunaan Statistika

Statistika sebagai disiplin ilmu berguna untuk kemajuan ilmu dan teknologi. Di negara maju, statistika telah digunakan untuk menyelesaikan berbagai masalah, diantaranya dalam bidang ekonomi, pertambangan, dan lain sebagainya. Statistika dapat digunakan sebagai sebagai alat :

a. ***Deskriptif***, yaitu menggambarkan atau menerangkan tentang sesuatu keadaan berdasarkan data yang diperoleh, seperti mengukur dampak dan proses pembangunan melalui indikator-indikator ekonomi, indeks harga konsumen, tingkat inflasi, GNP, laporan nota keuangan negara, menentukan pusat (berkumpulnya) barang tambang (mineral), dll.

b. ***Komparasi***, yaitu membandingkan dua atau lebih kelompok data.

c. ***Korelasi***, yaitu mencari hubungan dua atau lebih kelompok data.

d. ***Regresi***, yaitu meramalkan pengaruh data yang satu terhadap data yang lainnya (untuk estimasi terhadap kecenderungan peristiwa yang akan terjadi di masa depan).

e. ***Komunikasi***, yaitu alat penghubung antar pihak, berupa laporan data statistik atau analisis statistik sehingga pihak lain dapat memanfaatkannya dalam membuat suatu keputusan.

1.4. Populasi

Populasi (*population*) atau *universum* (*universe*) adalah himpunan objek penelitian.

Contoh 1.3. :

1. Misalkan akan diteliti usia rata-rata lampu pijar yang diproduksi oleh sebuah perusahaan lampu pijar. Lampu pijar yang harus diukur usianya dan dihitung rata-ratanya seharusnya meliputi seluruh lampu pijar yang dapat dihasilkan oleh perusahaan tersebut. Usia lampu pijar yang dapat diproduksi oleh perusahaan tersebut membentuk data kuantitatif yang dinamakan populasi.
2. Misalkan akan diteliti minat mahasiswa semester 3 Program Studi Informatika UNPAM terhadap musik tradisional Jawa (Karawitan Jawa). Populasinya adalah semua mahasiswa semester 3 Program Studi Informatika UNPAM.

1.5. Sampel

Sampel adalah unsur dari populasi. *Sampel representatif* adalah sampel yang mewakili karakteristik populasi secara tepat.

Contoh 1.4. :

1. populasi = Semua mahasiswa semester 3 Program Studi Informatika UNPAM.

   sampel = Nina, Dewi, Yuno, Sari, Adi (mereka semua mahasiswa semester 3 Program Studi Informatika UNPAM)

2. Populasi = Semua SMK di Tangerang Selatan.

   Sampel = SMK Negeri 1, SMK Negeri 2, SMK Sasmita Jaya, SMK Borobudur.

1.6. Variabel

Adalah gejala yang bervariasi yang menjadi objek penelitian (perhatian suatu penelitian).

Contoh 1.5. :

Penelitian mengenai hubungan usia dengan angka bunuh diri pria di Indonesia, dari tahun 1990 sampai dengan tahun 2000 dan dari usia 13 tahun sampai 70 tahun. Dalam hal ini usia (misal = x) adalah variabel bebas dan angka bunuh diri (misal = y) adalah variabel terikat.

Latihan

1. Apa perbedaan statistik dengan statistika ?
2. Apa yang dimaksud dengan metode statistika ?
3. Apa yang maksud dengan statistika inferensi ?
4. Apa yang dimaksud dengan statistika deskriptif ?
5. Apa yang dimaksud dengan populasi ?
6. Apa yang maksud dengan variabel ?
7. Apa yang dimaksud dengan sampel ?
8. Sebutkan contoh kegunaan statistika dalam bidang perindustrian, marketing dan pariwisata ?

Bab 2

Data Statistika

2.1. Arti dan Kegunaan Data

Data adalah segala fakta dan angka yang dapat memberikan suatu informasi. Data berarti sesuatu yang *diketahui* atau *dianggap*. *Diketahui* berarti sesuatu yang sudah terjadi atau ada. Karena itu sering ditemui istilah data statistik atau data historis. *Dianggap* berarti merupakan sesuatu pendapat atau asumsi.

Pada dasarnya, kegunaan data adalah untuk membuat keputusan (*decision makers*).

2.2. Syarat Data yang Baik

Data yang tidak baik akan menghasilkan keputusan yang tidak baik. Hal ini bisa mengakibatkan perencanaan yang salah, kontrol tidak efektif, dan evaluasi tidak mengenai sasaran yang benar dan objektif.

Syarat-syarat data yang baik, diantaranya adalah :

a. ***Obyektif***, artinya harus menggambarkan kenyataan yang sebenarnya.

b. ***Tepat waktu*** (*Up to date*), selalu berkaitan dengan keadaan yang berlaku saat ini, khususnya jika digunakan sebagai alat kontrol dan evaluasi.

c. ***Relevan***, yang berarti berkaitan erat dengan masalah-masalah yang diteliti.

d. ***Representatif***, yang berarti dapat mewakili / mengambarkan keadaan secara menyeluruh, bukan secara sebagian-sebagian.

e. Mempunyai kesalahan baku (*standard error*) dan variansi yang kecil.

2.3. Pembagian Data

2.3.1. Menurut Sifat Data

Ditinjau dari sifatnya, data dibagi menjadi dua bagian :

a. *Data kualitatif*

Adalah data yang tidak dinyatakan dalam bentuk angka. Misalnya jenis kelamin, jenis warna, dll.

b. *Data kuantitatif*

Adalah data yang dinyatakan dalam bentuk angka. Misalnya usia, tinggi badan, dll.

2.3.2. Menurut Cara Memperolehnya

Dari cara memperolehnya, data dibedakan menjadi :

a. *Data Primer*

Adalah data yang dikumpulkan dan diterbitkan sendiri oleh orang atau badan yang meneliti tersebut.

Contoh : Badan Pusat Statistik mengumpulkan data tentang ekspor dan impor, jumlah penduduk, kemudian menerbitkan data tersebut.

b. Data sekunder

Adalah data yang dikumpulkan oleh orang atau sebuah badan, sedangkan orang atau badan tersebut tidak langsung mengumpulkan sendiri, tetapi memperoleh dari pihak lain yang telah mengumpulkan terlebih dulu dan menerbitkannya.

2.3.3. Menurut Waktu Pengumpulannya

Ditinjau dari waktu mengumpulkan, data dibagi menjadi :

a. *Cross section*

Adalah data yang dikumpulkan pada saat yang bersamaan (dalam waktu yang sama / bulan yang sama).

b. *Time series*

Adalah data yang dikumpulkan pada beberapa periode yang berbeda.

2.3.4. Menurut Sumbernya

Ditinjau dari waktu pengumpulanya, data dibagi menjadi :

a. *Data intern*

Adalah data yang diperoleh langsung dari suatu instansi atau organisasi.

b. *Data ekstern*

Adalah data yang diperoleh dari luar instansi atau organisasi itu dan sifatnya umum.

2.3.5. Data-data yang Lain

a. *Data diskrit*

Adalah data yang hanya mempunyai jumlah data yang sangat terbatas. Misalnya data tentang jumlah mahasiswa di sebuah sekolah.

b. *Data kontinu*

   Adalah data yang secara teoritis mempunyai nilai pengamatan yang tidak terbatas (terus-menerus). Misalnya waktu, dll.

c. *Data statis*

   Adalah data yang mempunyai nilai tetap dan terbatas dalam setiap periode tertentu. Misalnya jumlah jam dalam satu hari, jumlah bulan dalam satu tahun, dll.

d. *Data dinamis*

   Adalah data yang mempunyai nilai turun-naik, mengikuti situasi tertentu. Misalnya volume penjualan suatu produk industri, dll.

2.4. Pengumpulan Data

Ada beberapa cara untuk mengumpulkan data :

2.4.1. Mengadakan wawancara

Wawancara dilakukan secara langsung kepada objek yang dituju dengan cara tanya-jawab.

2.4.2. Menggunakan Kuesioner

Lembar pertanyaan berisi pertanyaan yang harus dijawab oleh para responden. Bentuk pertanyaan jelas, singkat, dan mencakup semua masalah yang ingin diselidiki.

2.4.3. Mengadakan Pengamatan (Observasi)

Data langsung diambil dari hasil pengamatan. Sangat tergantung pada kemampuan pengamat untuk menarik kesimpulan.

2.4.4. Mengadakan Koleksi

Koleksi dilakukan dengan mencari data lewat selebaran, koran, majalah, brosur, dll. Kebenaran data kurang bisa dipertanggungjawabkan.

2.5. Penyusunan Data

Data yang telah terkumpul selanjutnya diatur, disusun, dan diklasifikasikan. Penyusunan data dibagi menjadi dua bagian, yaitu data tunggal dan data kelompok.

2.5.1. Data Tunggal

Adalah data yang disusun sendiri menurut besarnya.

Contoh : Data nilai Matematika dari 10 mahasiswa : 5, 6, 5, 8, 9, 7, 7, 6, 8, 7.

2.5.2. Data Kelompok

Adalah data disusun secara berkelompok.

Contoh : Data nilai Kalkulus dari 40 mahasiswa yang dikelompokkan menjadi kelompok-kelompok nilai 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, dan 90-99.

2.6. Tabel Distribusi Frekuensi

Sesuai dengan bentuk data yang akan disajikan, tabel distribusi frekuensi dikelompokkan menjadi dua bagian.

2.6.1. Distribusi frekuensi bilangan

Dalam tabel distribusi frekuensi, yang disajikan adalah bilangan (kuantitatif).

Contoh :

Nilai	Jumlah Mahasiswa
40-48	6.00
49-57	3.00
58-66	14.00
67-75	9.00
76-84	2.00
85-93	6.00
Jumlah	40.00

2.6.2. Distribusi frekuensi kategori

Tabel distribusi frekuensi kategori menyajikan penggolongan data berdasarkan kategori-kategori tertentu.

Contoh :

Tabel jumlah SMK di Pulau Jawa di beberapa propinsi pada tahu ajaran 1984 / 1985.

Propinsi	Jumlah Sekolah
DKI Jakarta	90
Jawa Barat	85
Jawa Tengah	137
DI Yogyakarta	41
Jawa Timur	148
Jumlah	40

2. 7. Contoh Tabel

2.7.1. Tabel Baris Kolom

Luas daerah Kalimantan dalam km persegi 1962

Propinsi	Luas
Kalbar	146.76
Kalsel	37.66
Kalteng	152.6
Kaltim	202.44
Jumlah	539.46

2.7.2. Tabel kontingensi

Untuk data yang terdiri dari dua klasifikasi atau dua variabel.

Jumlah mahasiswa STIE X menurut jurusan dan jenis kelamin tahun 1994

Jenis Kelamin	MKP	MP	AKT	Jumlah
Laki-laki	500	400	300	1200
Perempuan	400	300	200	900
Jumlah	900	700	500	2100

2.7.3. Tabel Distribusi Frekuensi

Untuk data kuantitatif yang terdiri dari beberapa kelompok.

Umur mahasiswa Universitas Y akhir tahun 1970.

Umur (Th)	Jumlah Mahasiswa
17 – 20	1.172
21 – 24	2.758
25 – 28	2.976
29 – 32	997
33 – 36	205
Jumlah	8.108

Latihan

1. Apa yang dimaksud dengan data statistika ?
2. Sebutkan kegunaan data !
3. Sebutkan syarat-syarat data yang baik !
4. Uraikan dengan jelas, kelebihan dan kekurangan cara kuesioner untuk memperoleh data !
5. Apa yang dimaksud dengan data statis dan data dinamis ?
6. Jelaskan perbedaan data tunggal dan data kelompok !

UNIT 3

Distribusi Frekuensi

Secara garis besar, data dapat disajikan dalam dua bentuk, yaitu bentuk daftar atau table dan bentuk grafik atau diagram. Untuk membuat tabel distribusi frekuensi, diperlukan beberapa langkah berikut.

3.1. Menentukan Daerah Jangkauan (*R*)

Daerah jangkauan atau rentang (range) adalah selisih antara data terbesar dan data terkecil. Jangkauan didefinisikan sebagai :

$$R = x_t - x_r$$

dengan :

R = Jangkauan

x_t = data terbesar

x_r = data terkecil.

Contoh 3.1. :

1. Tentukan jangkauan dari data berikut : 13, 19, 25, 30, 37, 45, 50, 65.

   Jawab :

   Data terbesar = x_t = 65

   Data terkecil = x_r = 13

   Daerah jangkauan = 65–13 = 52.

2. Tentukan daerah jangkauan dari nilai tugas kalkulus 40 mahasiswa semester 2 sebagaberikut :

40	45	65	67	85	90	43	57
63	65	70	75	80	63	60	70
43	45	60	60	60	65	75	86
92	90	50	60	50	65	60	65
70	45	70	75	80	85	65	70

Jawab :

Data terbesar = x_t = 92

Data terkecil = x_r = 40

Daerah jangkauan = 92–40 = 52.

3.2. Menetapkan Banyaknya Kelas (*K*)

Untuk menyusun data dalam tabel distribusi frekuensi tidak ditentukan dengan pasti berapa kelas yang harus dibuat. Prinsip-prinsip yang harus diperhatikan dalam menentukan jumlah kelas interval adalah :

1. Jumlah kelas jangan terlalu sedikit atau terlalu besar, supaya diperoleh gambaran yang jelas tentang kelompok.
2. Berdasarkan kebiasaan yang ada, jumlah kelas berkisar antara 5 sampai dengan 15.
3. Cara lain untuk menentukan jumlah kelas adalah menggunakan rumus Sturges, yaitu :

   $$K = 1 + 3{,}3 \log N$$

   dengan :

   K = jumlah kelas (*interval class*)

   N = banyaknya data (*raw data*).

Untuk N yang terlalu kecil atau terlalu besar, penggunaan rumus Sturges tidak tepat, sehingga tidak mutlak dilakukan. Berikut adalah contoh penggunaan rumus Sturges :

Misalkan banyaknya data = N = 40.

Maka $K = 1 + 3{,}3 \log N$ = 1+ (3,3)(log 40) = 1 + (3,3 x 1,6021) = 6,29 ≈ 6.

3.3. Menentukan Panjang Interval Kelas (*I*)

Panjang interval kelas untuk setiap kelas pada distribusi frekuensi diusahakan sama. Berdasarkan rumus K dari Sturges, panjang c didefinisikan sebagai :

$$c = \frac{R}{K}$$

dengan :

R = jangkauan

K = banyaknya kelas

c = interval kelas.

Dari contoh no. 1 dan 2 di atas :

Daerah jangkauan = R = 52, banyaknya kelas = K = 6.

Interval kelas = $c = \frac{R}{K} = \frac{52}{6} = 8{,}67 \approx 9$.

3.4. Menentukan Batas Kelas

Batas kelas adalah dua buah nilai yang membatasi suatu kelas dengan kelas yang lain. Setiap nilai yang terletak disebelah kiri kelas disebut batas bawah kelas (*lower class limits*); sedangkan nilai yang terletak di sebelah kanan kelas disebut batas atas kelas (*upper class limits*). Nilai batas kelas dibagi menjadi dua bagian, yaitu sebagai berikut.

3.4.1. Batas Semu

Adalah nilai yang ditentukan berdasarkan data yang ada untuk memasukkan frekuensi masing-masing data sesuai dengan interval kelas. Batas semu terdiri dari batas semu atas dan batas semu bawah.

contoh 3.4.1. :

Diberikan tabel sebagai berikut :

Nilai	Jumlah Siswa (Frekuensi)
40 - 48	6
49 - 57	3
58 - 66	14
67 - 75	9
76 - 84	2
85 - 93	6
Jumlah	40

Batas semu bawah adalah nilai-nilai : 40, 49, 58, 67, 76, dan 85.

Batas semu atas adalah nilai-nilai : 48, 57, 66, 75, 84, dan 93.

40 – 48 kelas pertama

49 – 57 kelas kedua

58 – 66 kelas ketiga

67 – 75 kelas keempat

76 – 84 kelas kelima

85 – 93 kelas keenam.

3.4.2. Batas Nyata

Batas nyata disebut juga tepi kelas (*class boundary*), terdiri dari :

a. *Batas Nyata Bawah*

   Adalah batas semu bawah dikurangi 0,5 untuk data dengan ketelitian sampai satu satuan. Sebagai contoh, data dengan ketelitian sampai satu desimal, batas nyata bawah sama dengan batas semu bawah dikurangi 0,05 dan seterusnya.

b. *Batas Nyata Bawah*

   Adalah batas semu atas ditambah 0,5 untuk data dengan ketelitian sampai satu satuan. Sebagai contoh, data dengan ketelitian sampai satu desimal, batas nyata atas sama dengan batas semu atas ditambah 0,05 dan seterusnya.

Contoh 3.4.2. :

Tabel contoh di atas dapat disajikan dalam tabel berikut :

Nilai	Jumlah Siswa (Frekuensi)
39,5 - 48,5	6
48,5 - 57,5	3
57,5 - 66,5	14
66,5 - 75,5	9
75,5 - 84,5	2
84,5 - 93,5	6
Jumlah	40

Setengah dari jumlah kedua nilai batas kelas disebut *nilai tengah* (*mid-point*) atau *class mark* suatu kelas, dan biasanya dilambangkan dengan *X*. Sebagai contoh, dari tabel di atas diperoleh :

$$X_1 = \frac{39,5 + 48,5}{2} = 44 \text{ atau } X_1 = \frac{40 + 48}{2} = 44$$

$$X_2 = \frac{48,5 + 57,5}{2} = 53 \text{ atau } X_2 = \frac{49 + 57}{2} = 53$$

dan seterusnya.

3.5. Menentukan Batas Bawah Terendah dari Kelas Pertama

Batas terendah dari kelas pertama dipilih sedemikian sehingga dalam penyebaran frekuensi tidak terdapat nilai yang tidak masuk dalam kelompok data. Nilai terbesar dari batas bawah kelas pertama adalah nilai terkecil dari seluruh data; sedangkan nilai terkecil dari batas atas kelas terakhir adalah nilai terbesar dari seluruh data.

3.6. Menghitung Frekuensi

Nilai frekuensi untuk masing-masing interval kelas ditentukan dengan menggunakan *sistem turus*.

Contoh 3.6. :

Perhatikan kembali tabel contoh di atas :

Nilai	Turus	Jumlah Siswa (Frekuensi)
39,5 - 48,5	𝍸 I	6
48,5 - 57,5	III	3
57,5 - 66,5	𝍸 𝍸 IIII	14
66,5 - 75,5	𝍸 IIII	9
75,5 - 84,5	II	2
84,5 - 93,5	𝍸 I	6
Jumlah		40

3.7. Membuat Tabel Distribusi Frekuensi

Cara yang paling mudah untuk memasukkan data ke dalam tabel distribusi frekuensi adalah cara *turus* atau *tally* (*milidi*).

Perhatikan contoh berikut :

Diberikan data pemasukkan sayuran dari daerah Produsen ke Pasar Induk Kramat Jati dari 22 Agustus 1999 sampai dengan 29 Nopember 1999 (dalam puluhan ton) :

126	129	123	121	129	120	122	123	128	123
124	129	123	123	126	129	124	125	117	121
123	125	128	125	124	116	114	123	123	129
123	123	116	114	123	123	128	123	118	117
111	107	100	110	119	115	113	108	105	97
107	114	113	102	107	114	104	103	133	123
149	139	140	130	146	120	110	111	130	150
133	132	129	125	129	124	122	128	119	132
130	129	113	119	126	124	130	139	131	130
120	116	119	121	141	138	133	122	124	118

Dari data tersebut kita akan membuat tabel distribusi frekuensi :

1. Susun data dan tentukan daerah jangkauannya.

   Data diurutkan dari urutan terkecil sampai yang terbesar (*array*) :

   100 102 103 104 105 107 107 107 108

   110 111 111 113 113 113 114 114 114

   115 116 116 116 117 117 118 118 119

   119 119 120 120 120 121 121 121 122

   122 123 123 123 123 123 123 123 123

   123 123 123 123 123 124 124 124 124

   128 128 129 129 129 129 129 129 129

   130 130 130 130 130 131 132 132 133

   133 138 139 139 140 141 146 149 150

   Data terbesar = x_t = 150

   Data terkecil = x_r = 97

   Daerah jangkauan = 150–97 = 53.

2. Tentukan banyaknya kelas.

   Banyaknya kelas = $K = 1 + 3{,}3 \log N$

   $= 1 + (3{,}3)(\log 100) = 1 + (3{,}3 \times 2) = 7{,}6 \approx 8.$

3. Tentukan panjang interval kelas.

   Panjang interval = $c = \dfrac{R}{K} = \dfrac{53}{8} = 6{,}6 \approx 7.$

4. Tentukan ujung bawah kelas pertama.

   Ujung bawah kelas pertama dipilih 97. Jadi, kelas pertama adalah 97– 103, kelas kedua adalah 104–10, kelas ketiga : 111–117, dan seterusnya.

5. Tentukan frekuensinya.

   Frekuensi dihitung dengan sistem turus.

   Dari kelima langkah diatas, diperoleh tabel :

Nilai	**Jumlah Siswa (Frekuensi)**
97 – 103	4
104 – 110	8
111 – 117	15
118 – 124	35
125 – 131	25
132 – 138	6
139 – 145	4
146 – 152	3
Jumlah	100

3.8. Frekuensi Kumulatif, Frekuensi Relatif, dan Frekuensi Kumulatif Relatif

3.8.1. Frekuensi Kumulatif

Frekuensi kumulatif dapat dibentuk dari tabel distribusi frekuensi biasa, dengan cara menjumlahkan frekuensi-frekuensinya. Frekuensi kumulatif terdiri dari dua jenis, yaitu frekuensi kumulatif *kurang dari* dan frekuensi kumulatif *lebih dari*.

Contoh 3.8.1.:

a. Tabel disribusi frekuensi kumulatif kurang dari.

Batas Nyata Kurang dari	f_K **Kurang dari** $[f_K <]$
96,5	0
103,5	4
110,5	12
117,5	27
124,5	62
131,5	87
138,5	93
145,5	97
152,5	100

b. Tabel distribusi frekuensi kumulatif lebih dari.

Batas Nyata Lebih dari	f_K **Lebih dari** $[f_K >]$
96,5	100
103,5	96
110,5	88
117,5	73
124,5	38
131,5	13
138,5	7
145,5	3
152,5	0

3.8.2. Frekuensi Relatif

Frekuensi relatif (f_R) suatu interval kelas didefinisikan sebagai :

$$f_R = \frac{f_i}{\sum_{i=1}^{n} f_i}$$

dengan :

f_R = frekuensi ke-*n*.

$\sum_{i=1}^{n} f_i$ = total frekuensi.

Contoh tabel frekuensi relatif dalam desimal dan persen :

Pemasukkan (Ton)	f_i	f_R Desimal	f_R (%)
97 – 103	4	0,04	4
104 – 110	8	0,08	8
111 – 117	15	0,15	15
118 – 124	35	0,35	35
125 – 131	25	0,25	25
132 – 138	6	0,06	6
139 – 145	4	0,04	4
146 – 152	3	0,03	3
Jumlah	100	1,00	100

3.8.3. Frekuensi Kumulatif Relatif

Frekuensi kumulatif relatif (f_{KR}) atau frekuensi kumulatif persentase didefinisikan sebagai :

$$f_{KR} = \frac{f_{Ki}}{\sum_{i=1}^{n} f_i}$$

dengan :

f_{KR}= frekuensi kumulatif relatif

$\sum_{i=1}^{n} f_i$ = total frekuensi.

Contoh tabel frekuensi kumulatif relatif :

Batas Nyata kurang dari	$f_K <$	$f_{KR} <$	$f_K >$	$f_{KR} >$
96,5	0	0	100	100
103,5	4	4	96	96
110,5	12	12	88	88
117,5	27	27	73	73
124,5	62	62	38	38
131,5	87	87	13	13
138,5	93	93	7	7
145,5	97	97	3	3
152,5	100	100	0	0

3.9. Menyajikan Data dalam Grafik atau Diagram

Data dapat disajikan dalam bentuk grafik atau diagram. Jenis diagram antara lain :

1. Diagram batang (*bar chart* / *bar graph*).
2. Diagram garis (*line chart* / *line graph*).
3. Diagram lambang / piktogram.
4. Diagram lingkaran (*pie chart* / *circular graph*).
5. Diagram Peta / kartogram.
6. Diagram pencar / diagram titik.
7. Diagram histogram.
8. Diagram poligon.
9. Kurva ogive positif dan kurva ogive negatif.

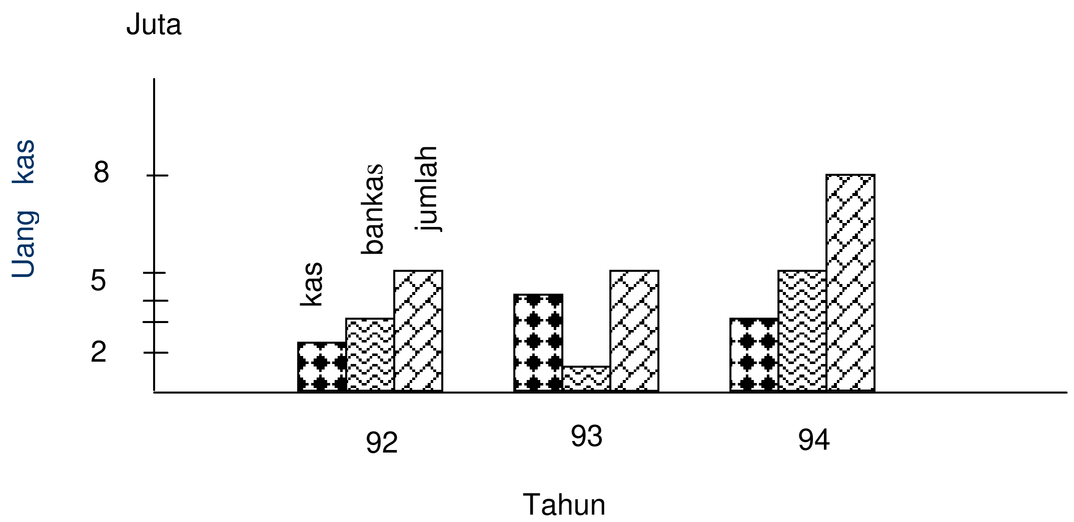
Diagram nomor 1 – 6 paling cocok digunakan untuk data tunggal; sedangkan nomor 8 dan 9 paling sesuai digunakan untuk data kelompok.

3.10. Contoh-contoh Diagram

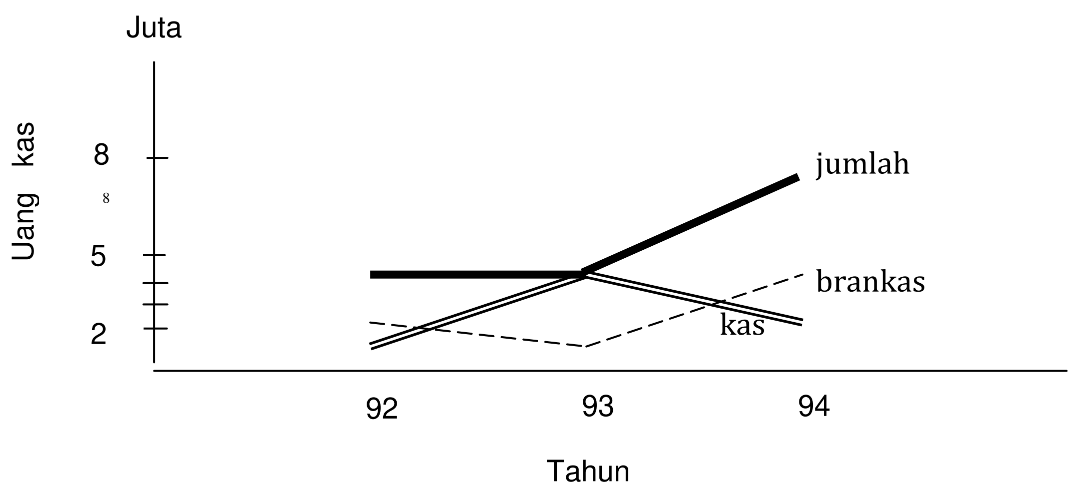
Diagram Batang

Diagram Batang Tunggal

Tahun	Uang Kas (dalam Jutaan)	Uang Di Bank	Jumlah
92	2	3	5
93	4	1	5
94	3	5	8



3.10.2. Diagram Garis

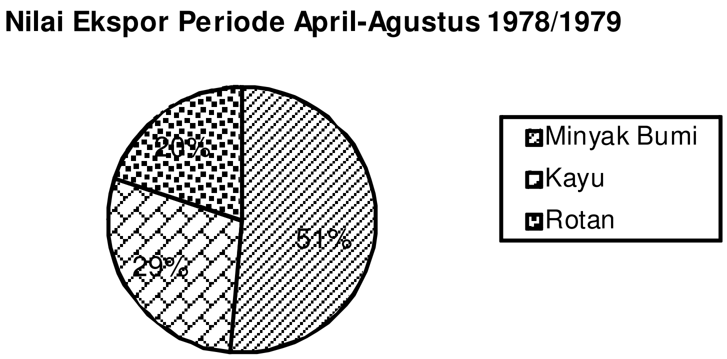


3.10.3. Diagram lingkaran

Untuk menunjukan peranan (%) masing-masing komponen terhadap keseluruhannya.

Contoh.

Jenis Komoditi	Nilai Ekspor
Minyak Bumi	450.00
Kayu	250.00
Rotan	175.00



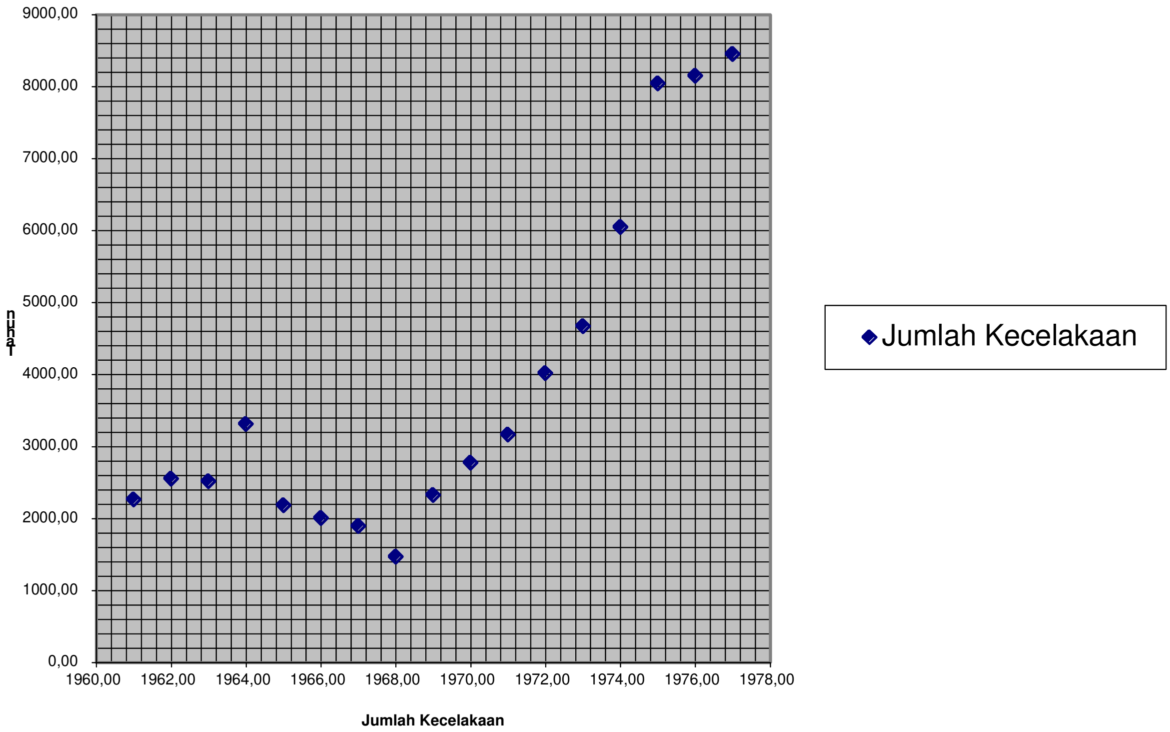
3.10.4. Diagram pencar

Untuk menggambarkan kempulan data yang terdiri atas dua variabel.

Contoh

Tahun	Jumlah Kecelakaan
1961.00	2263.00
1962.00	2550.00
1963.00	2515.00
1964.00	3310.00
1965.00	2183.00
1966.00	2006.00
1967.00	1893.00
1968.00	1467.00
1969.00	2325.00
1970.00	2773.00
1971.00	3164.00
1972.00	4017.00
1973.00	4669.00
1974.00	6047.00
1975.00	8043.00
1976.00	8148.00
1977.00	8452.00

Jumlah Kecelakaan Lalu-lintas di DKI Jakarta Tahun 1961-1977



3.10.5. Diagram Lambang

Sebuah perusahaan mempunyai 500 karyawan. Karyawan wanita berjumlah 300 orang, sedangkan karyawan pria 200 orang.

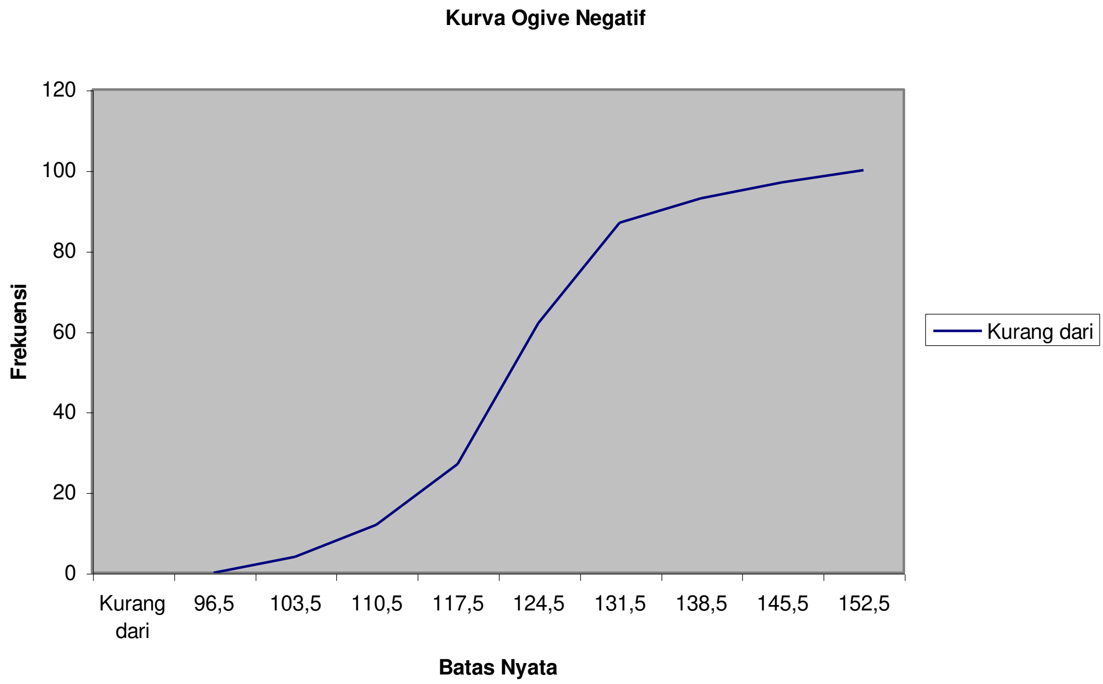
Karyawan	Jumlah
Pria	♣♣
Wanita	♣♣♣
Jumlah	♣♣♣♣♣

3.10.6. Kurva Ogive

a. Kurva Ogive negatif.

Tabel disribusi frekuensi kumulatif kurang dari.

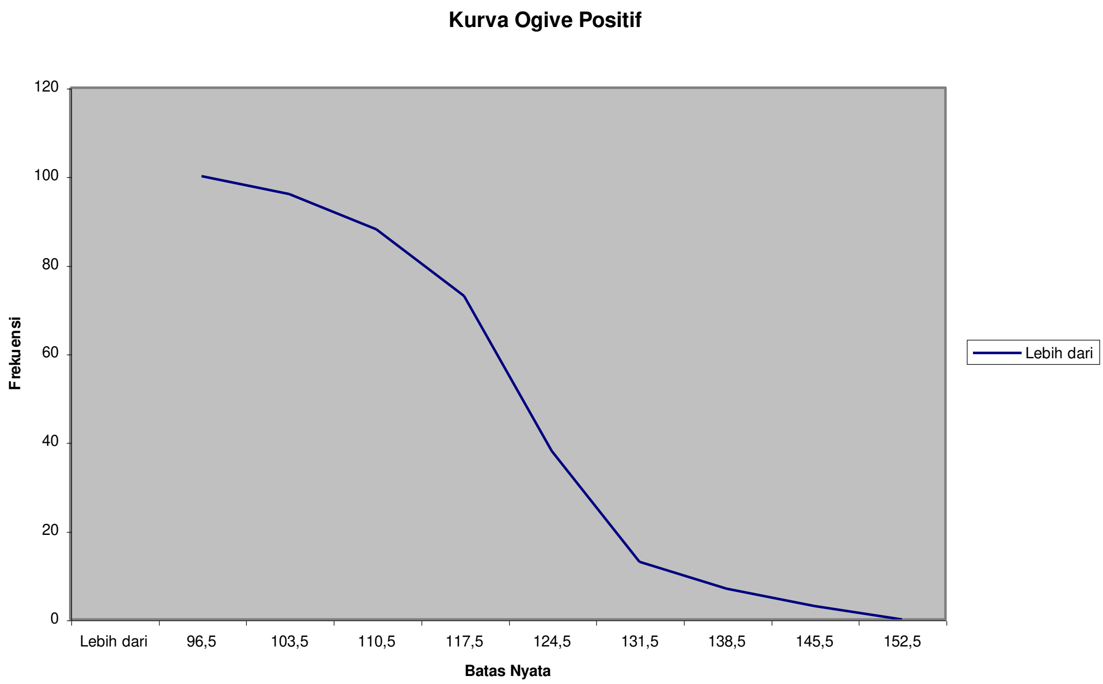
Batas Nyata Kurang dari	f_K Kurang dari $[f_K <]$
96,5	0
103,5	4
110,5	12
117,5	27
124,5	62
131,5	87
138,5	93
145,5	97
152,5	100



b. Kurva ogive positif.

Tabel distribusi frekuensi kumulatif lebih dari.

Batas Nyata Lebih dari	f_K Lebih dari $[f_K >]$
96,5	100
103,5	96
110,5	88
117,5	73
124,5	38
131,5	13
138,5	7
145,5	3
152,5	0



Latihan

1. Sebutkan macam-macam bentuk penyajian data !
2. Dari data berikut, carilah jangkauannya : 45, 40, 35, 30, 31, 29, 20, 19, 30, 29.
3. Diberikan data sebagai berikut :

   108 90 95 95 85 85 86 92

   101 80 93 93 97 90 91 87

   91 79 102 92 97 98 90 91

   94 98 99 82 103 92 98 75

   81 88 89 90 99 91 87 104

   Tentukan :

   a. Jangkauannya.
   b. Banyaknya kelas.
   c. Panjang interval.
   d. distribusi frekuensinya.
   e. Gambar histogram dan tabel frekuensinya.
4. Diberikan tabel jumlah siswa SMEA A yang lulus ujian dari tahun 1992 sampai dengan tahun 1996 :

Tahun	Jumlah Siswa
1992	70
1993	80
1994	100
1995	110
1996	120

Dari data tersebut, buatlah :

a. Diagram garis.
b. Diagram Lingkaran.

UNIT 4

Ukuran Pemusatan Data

Ukuran pemusatan data (*tendensi sentral*) adalah ukuran untuk memberikan gambaran tentang sampel dari suatu populasi. Tendensi sentral berupa rata-rata (mean), modus, dan median.

4.1. Menentukan Mean, Median dan Modus Data Tunggal

4.1.1. Mean (Rata-rata)

Misalkan $x_1, x_2, \ldots, x_n$ adalah nilai-nilai observasi. Mean dari data observasi didefinisikan sebagai :

$$\overline{X} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Atau dapat juga ditulis sebagai : $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Contoh 4.1.1. :

1. Nilai ulangan Biologi dari 7 siswa adalah 7, 8, 9, 6, 7, 7, 8. Tentukan rata-rata hitungnya !

   Jawab :

   $$\overline{X} = \frac{7+8+9+6+7+7+8}{7} = \frac{52}{7} = 7{,}43.$$

2. Diberikan tabel nilai ulangan Matematika :

Nilai Matematika = x	Banyaknya siswa
5	4
6	12
7	14
8	10

   Carilah rata-rata hitungnya !

Jawab :

$$\overline{X} = \frac{(5\text{x}4)+(6\text{x}12)+(7\text{x}14)+(8\text{x}10)}{4+12+14+10} = \frac{20+72+98+80}{40} = \frac{270}{40} = 6{,}75.$$

3. Carilah rata-rata hitung dari data berikut :

Kelas Interval	f_i	x_i	$f_i x_i$
31 – 40	8	35,5	284
41 – 50	5	45,5	227,5
51 – 60	16	55,5	888
61 – 70	21	65,5	1375,5
71 – 80	28	75,5	2114
81 – 90	20	85,5	1710
91 –100	2	95,5	191
Jumlah	100		6790

Rumus: $M = \overline{X} = \frac{6790}{100} = 67{,}90$.

4.1.2. Median

Median adalah nilai sentral dari sebuah rangkaian data yang telah diurutkan dari kecil ke besar. Nilai ini merupakan nilai sentral berhubungan dengan posisi sentral yang dimilikinya dalam sebuah distribusi. Median juga disebut sebagai rata-rata posisi (*positional average*). Secara teoritis, median membagi seluruh jumlah observasi atau pengukuran ke dalam dua bagian yang sama.

Misalkan $x_1, x_2, \ldots, x_n$ adalah nilai-nilai observasi. Bila nilai-nilai tersebut diurutkan dari nilai terkecil hingga nilai terbesar sedemiian rupa sehingga :

$x_1 \le x_2 \le \ldots \le x_n$

maka median adalah nilai x_k jika *n* ganjil dan $n = 2k - 1$. Jika *n* genap dan $n = 2k$, bila $x_k = x_{k+1}$, maka median= $x_k = x_{k+1}$ dan jika $x_k < x_{k+1}$ maka median $= \frac{(x_k + x_{k+1})}{2}$.

Contoh 4.1.2. :

1. Tentukan median dari data : 3, 4, 6, 3, 2, 1, 8, 9 !

   Jawab :

   Data diurutkan : 1, 2, 3, 3, 4, 6, 8, 9.

   Jumlah data ada 8 buah, berarti genap, sehingga $k = \frac{n}{2} = \frac{8}{2} = 4$.

   Karena $x_4 < x_5$, maka median= $\frac{(x_k + x_{k+1})}{2}$. Jadi, median= $\frac{3+4}{2} = 3{,}5$.

2. Tentukan median dari data : 2, 4, 6, 1, 7, 3, 9, 5, 6.

   Jawab :

   Data diurutkan : 1, 2, 3, 4, 5, 6, 6, 7, 9.

   Jumlah data ada 9 buah, berarti ganjil, sehingga $n = 2k - 1 \Leftrightarrow$ $k = \frac{n+1}{2} = \frac{9+1}{2} = \frac{10}{2} = 5$.

   Jadi, median = 5.

4.1.3. Modus

Modus (*mode*) adalah nilai observasi yang memiliki frekuensi tertinggi.

Contoh 4.1.3. :

1. Tentukan modus dari data berikut : 4, 2, 6, 7, 4, 4, 5.

   Jawab :

   Modus = 4.

2. Tentukan modus dari data berikut :

Nilai Matematika	Banyaknya Siswa
5	3
6	6
7	10
8	5
Jumlah	24

   Jawab : Modus = 7.

4.2. Menentukan Mean, Median dan Modus Data Berkelompok

4.2.1. Mean

Serupa dengan data tunggal, rata-rata hitung data berkelompok dapat ditentukan dengan tiga cara berikut ini.

4.2.1.1. Titik Tengah Kelas

Jika x_i merupakan titik tengah kelas ke-*i* dan f_i merupakan frekuensi kelas ke-*i*, maka maka mean didefinisikan sebagai :

$$\overline{X} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}.$$

Contoh 4.2.1.1. :

Carilah rata-rata hitung dari data berikut :

Kelas Interval	f_i	x_i	$f_i x_i$
31 – 40	8	35,5	284
41 – 50	5	45,5	227,5
51 – 60	16	55,5	888
61 – 70	21	65,5	1375,5
71 – 80	28	75,5	2114
81 – 90	20	85,5	1710
91 –100	2	95,5	191
Jumlah	100		6790

Jawab :

$$\overline{X} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{6790}{100} = 67{,}90.$$

4.2.1.2. Simpangan Rata-rata

Untuk mencari simpangan rata-rata, kita harus menentukan nilai rata-rata duga (rataan sementara) atau disebut *assumed mean* (AM), biasanya titik tengah kelas yang frekuensinya terbesar. Selanjutnya, kita menentukan simpangan : $d_i = x_i - AM$. Nilai mean ditentukan oleh rumus :

$$\overline{X} = AM + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}.$$

Contoh 4.2.1.2. :

Carilah rata-rata hitung dari data berikut :

Interval Kelas	f_i	x_i	Simpangan $d_i = x_i - AM$	$f_i d_i$
97 – 103	4	100	– 21	– 84
104 – 110	8	107	– 14	– 112
111 – 117	15	114	– 7	– 105
118 – 124	**35**	**121**	**0**	**0**
125 – 131	25	128	7	175
132 – 138	6	135	14	84
139 – 145	4	142	21	84
146 – 152	3	149	28	84
Jumlah	100			126

Jawab :

Rata-rata sementara (AM) = 121.

Rata-rata hitung : $\overline{X} = AM + \dfrac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} = 121 + \dfrac{126}{100} = 121 + 1{,}26 = 122{,}26.$

4.2.2. Median

Median (Me) didefinisikan dengan : $Me = L + \left(\dfrac{\frac{1}{2}n - \left(\sum\limits_{i=1}^{n} f_i \right)_{Me}}{f_{Me}} \right)$,

dengan : L = Tepi bawah kelas yang memuat median

n = Jumlah seluruh frekuensi

$\left(\sum f\right)_{Me}$ = Jumlah frekuensi sebelum median

f_{Me} = Frekuensi kelas yang memuat median.

Contoh 4.2.2. :

Tentukan median data berikut :

Pemasukkan Sayur-mayur (puluhan ton)	Banyaknya Hari (frekuensi = f_i)	Frekuensi Kumulatif (f_k) Kurang dari	
97 – 103	4	96.5	0
104 – 110	8	103.5	4
111 – 117	15	110.5	12
118 – 124	35	117.5	27
125 – 131	25	124.5	62
132 – 138	6	131.5	87
139 – 145	4	138.5	93
146 – 152	3	145.5	97
Jumlah	100		

Setengah jumlah frekuensi = $\frac{1}{2}n = \frac{1}{2}(100) = 50$

Kelas median = 118-124 , sehingga L = 118 – 0,5 = 117,5

$\left(\sum f\right)_{Me} = 27$

f_{Me} = 35

c = panjang interval = 124,5 – 117,5 = 7

Median = $Me = L + \left(\frac{\frac{1}{2}n - \left(\sum_{i=1}^{n} f_i \right)_{Me}}{f_{Me}} \right) = 117{,}5 + \left(\frac{50-27}{35} \right) 7$

=117,5 + 4,6 = 122,1.

4.2.3. Modus

Modus dari data yang dikelompokkan, didefinisikan sebagai :

Modus (Mo) = $L + \left(\frac{d_1}{d_1 + d_2} \right) c$,

dengan :

L = Tepi bawah kelas modul

d_1 = Selisih frekuensi kelas modus dengan kelas sebelumnya

d_2 = Selisih frekuensi kelas modus dengan kelas sesudahnya

c = panjang interval kelas

Contoh 4.2.3. :

Tentukan modus data berikut :

Pemasukkan Sayur-mayur (puluhan ton)	Banyaknya Hari (frekuensi = f_i)
97 – 103	4
104 – 110	8
111 – 117	15
118 – 124	35
125 – 131	25
132 – 138	6
139 – 145	4
146 – 152	3
Jumlah	100

Jawab :

Kelas modus = 118 - 124

L = 118 – 0,5

d_1 = 35 – 15 = 20

d_2 = 35 – 25 = 10

c = 124,5 – 117,5 = 7

Modus (Mo) $= L + \left(\frac{d_1}{d_1 + d_2}\right)c = 117{,}5 + \left(\frac{20}{20+10}\right)7 = 117{,}5 + \frac{140}{30}$

= 117,5 + 4,67

= 122,17.

4.3. Rata-rata Ukur (Geometri Mean)

4.3.1. Rata-rata Ukur dari Data Tunggal

Rata-rata ukur atau rata-rata geometri dilambangkan dengan *RU* atau *RG*. Untuk data tunggal, didefinisikan sebagai :

$$RU = RG = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n}$$

Contoh 4.3.1. :

1. Tentukan rata-rata ukur dari data tunggal berikut : 2, 4, 2, 1.

   Jawab :

   $RU = RG = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n} = \sqrt[4]{1 \times 2 \times 2 \times 4} = \sqrt[4]{16} = 2.$

2. Diketahui data 2,5, 9, 19, 38, 75

   Berapakah rata-rata ukurnya?

   Jawab :

   Cara 1. $RG = \sqrt[6]{2.5.9.19.38.75} = 13{,}02$

   Cara 2. $\log RG = \log \sqrt[6]{2.5.9.19.38.75}$

$$= \log (2.5.9.19.38.75)^{1/6}$$

$$= \frac{1}{6} \log (2.5.9.19.38.75)$$

$$= \frac{1}{6} (\log 2 + \log 5 + \log 9 + \log 19 + \log 38 + \log 38 + \log 75)$$

$$= \frac{1}{6} (0{,}3010 + 0{,}6990 + 0{,}9542 + 1{,}2788 + 1{,}5798 + 1{,}8751)$$

$$= \frac{1}{6} (6{,}6879) = 1{,}11465$$

RG = 13,02.

3. Tentukan rata-rata ukur dari data berikut :

Tahun	Prod. (x_i)	Log x_i
88	20	1.3010
89	25	1.3979
90	30	1.4771
91	40	1.6020
n = 4		5.7780

Jawab : $\log RG = \frac{\sum \log xi}{n} = \frac{5{,}7780}{4} = 1{,}4445$

RG = 27,83

4. *4.3.2. Rata-Rata Ukur untuk Data Bergolong*

Untuk data yang berbentuk distribusi frekuensi, rata-rata ukur didefinisikan sebagai :

$$\log RU = \frac{\sum_{i=1}^{K} f_i \log x_i}{N}, \text{ dengan } N = \sum_{i=1}^{K} f_i .$$

Contoh 4.3.2. :

Tentukan rata-rata ukur untuk data berikut :

Nilai Ujian	f_i	x_i	$\log x_i$	$f_i \log x_i$
31 – 40	1	35.5	1.5502	1.5502
41 – 50	2	45.5	1.6580	3.3160
51 – 60	5	55.5	1.7443	8.7215
61 – 70	15	65.5	1.8162	27.2430
71 – 80	25	75.5	1.8779	46.9475
81 – 90	20	85.5	1.9320	38.6400
91 – 100	12	95.5	1.9800	23.7600
Jumlah	80			150.1782

Jawab :

$\sum (f_i \log x_i) = 150{,}1782$; $\sum f_i = 80$

$$\log RG = \frac{150{,}1782}{80} = 1{,}8772$$

RG= 75,37

Jadi, nilai ujian mahasiswa mempunyai rata-rata ukur 75,37.

4.4. Mengukur Tingkat Pertumbuhan Penduduk

Rumus : $P(t) = P_0 \left(1 + \frac{\overline{X}}{100}\right)^t$

dengan : $P(t)$ = Keadaan Akhir

P_0 = Keadaan Awal

$\overline{X}$ = Rata-rata pertumbuhan setiap satuan waktu

t = Satuan waktu yang digunakan

contoh 4.4. :

Penduduk Indonesiaa pada akhir tahun 1946 adalah 60 juta, sedangkan akhir tahun 1956 adalah 78 juta. Berapa laju pertumbuhan penduduk tersebut.

Jawab :

$P_0 = 60$; $P(t) = 78$; $t = 10$

$$P(t) = P_0 \left(1 + \frac{\overline{X}}{100}\right)^t$$

$$\Leftrightarrow 78 = 60\left(1 + \frac{\overline{X}}{100}\right)^{10}$$

$$\Leftrightarrow \log\ 78 = \log 60 + 10 \log \left(1 + \frac{\overline{X}}{100}\right)$$

$$\Leftrightarrow 1{,}8921 = 1{,}7782 + (10).\log \left(1 + \frac{\overline{X}}{100}\right)$$

$$\Leftrightarrow 0{,}1139 = (10).\log \left(1 + \frac{\overline{X}}{100}\right)$$

$$\Leftrightarrow 0{,}01139 = \log \left(1 + \frac{\overline{X}}{100}\right)$$

$$\Leftrightarrow 1{,}0266 = 1 + \frac{\overline{X}}{100}$$

$\Leftrightarrow$ 0,0266 = $\overline{X}/100$

$\Leftrightarrow$ $\overline{X}$ = 2,66.

Jadi, laju rata-rata pertumbuhan penduduk = 2,66%.

Latihan

1. Tentukan mean, median dan modus dari data berikut :

Nilai Ulangan	Frekuensi
9 – 21	4
22 – 34	3
35 – 47	6
48 – 60	12
61 – 73	11
74 – 86	9
87 – 99	5

2. Tentukan mean soal no.1 dengan :
   a. Cara simpangan rata-rata.
   b. Cara coding.

3. Tentukan rata-rata harmonis, rata-rata kuadrat, dan rata-rata geometri dari data berikut : 29, 27, 20, 22, 30, 21, 25.

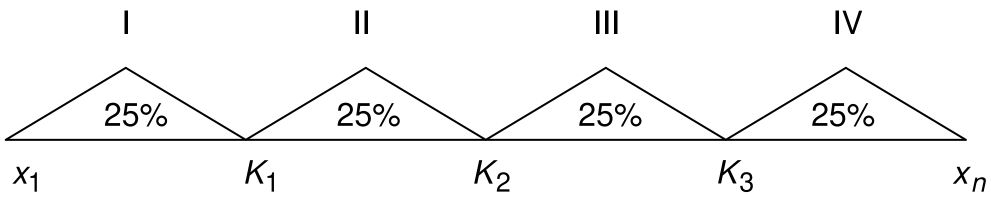
Unit 5

Ukuran Letak Data

Ukuran berdasarkan pada nilai-nilai batas jika data tersebut dibagi empat sama banyak, sepuluh sama banyak, atau dibagi seratus sama banyak. Ukuran letak data yang akan kita bahas adalah kuartil, desil, dan persentil.

5.1. Kuartil

Kuartil adalah nilai batas jika sekumpulan data dibagi menjadi empat bagian yang sama, setelah data diurutkan dari kecil ke besar.



Keterangan :

x_1 = data terkecil (nilai minimum)

K_1 = kuartil pertama (kuartil bawah)

K_2 = kuartil kedua (kuartil tengah = median)

K_3 = kuartil ketiga (kuartil atas)

x_n = data terbesar (nilai maksimum)

Untuk menentukan nilai kuartil, data diurutkan mulai dari yang terkecil ke nilai yang terbesar. Setelah itu menentukan letak kuartil dana akhirnya menentukan nilai kuartil.

Letak kuartil didefinisikan sebagai :

Letak K_i = data ke-$\frac{i(n+1)}{4}$,

dengan :

n = banyak data

i = 1, 2, 3.

Contoh 5.1. :

1. Tentukan kuartil bawah, kuartil tengah, dan kuartil atas dari data : 7, 5, 8, 4, 6, dan 10.

   Jawab :

   Data diurutkan : 4, 5, 6, 7, 8, 10.

   Letak K_1 = data ke-$\frac{1(6+1)}{4} = \frac{7}{4} = 1{,}75$. Jadi, $K_1 = 4 + 0{,}75(5-4) = 4 + 0{,}75 = 4{,}75.$

   Letak K_2 = data ke-$\frac{2(6+1)}{4} = \frac{14}{4} = 3{,}5$. Jadi, $K_2 = 6 + 0{,}5(7-6) = 6 + 0{,}5 = 6{,}5.$

   Letak K_3 = data ke-$\frac{3(6+1)}{4} = \frac{21}{4} = 5{,}25$. Jadi,

   $K_3 = 8 + 0{,}25(10-8) = 8 + 0{,}5 = 8{,}5.$

2. Tentukan K_1, K_2, dan K_3 dari data berikut :

Nilai Ulangan	Banyaknya Siswa
5	1
6	4
7	8
8	2
Jumlah	15

   Jawab :

   Letak K_1 = data ke-$\frac{1(15+1)}{4} = \frac{16}{4} = 4$. Jadi, K_1 = 6.

   Letak K_2 = data ke-$\frac{2(15+1)}{4} = \frac{32}{4} = 8$. Jadi, K_2 = 7.

   Letak K_3 = data ke-$\frac{3(15+1)}{4} = \frac{48}{4} = 12$. Jadi, K_3 = 7.

Untuk menentukan kuartil-kuartil dari data yang berbentuk distribusi frekuensi, pada prinsipnya sama dengan menentukan mediannya. Pertama, tabel dilengkapi dengan frekuensi kumulatif kurang dari (grafik ogive positif), kemudian kuartil-kuartil ditentukan dengan rumus :

$$K_i = L_{K_i} + \left(\frac{\frac{i}{4}N - \left(\sum_{i=1}^{k} f_i \right)_{K_i}}{f_{K_i}} \right) c$$

dengan :

L_{K_i} = tepi bawah kelas yang memuat K_i.

N = jumlah seluruh frekuensi.

$\left(\sum_{i=1}^{k} f_i \right)_{K_i}$ = jumlah frekuensi sebelum K_i.

c = panjang interval kelas.

i = 1, 2, 3.

3. Tentukan K_1, K_2, dan K_3 dari data berikut :

Berat (ton)	f_i
97-103	4
104-110	8
111-117	15
118-124	35
125-131	25
132-138	6
139-145	4
146-152	3
Jumlah	100

Jawab :

Interval Kelas	f_i	Frekuensi Kumulatif (f_k) Kurang dari	
97-103	4	96.5	0
104-110	8	103.5	4
111-117	15	110.5	12
118-124	35	117.5	27
125-131	25	124.5	62
132-138	6	131.5	87
139-145	4	138.5	93
146-152	3	145.5	97
Jumlah	100		

Kuartil bawah = K_1

$\frac{1}{4}N = \frac{1}{4}(100)$=25, maka kelas K_1 adalah 111-117.

L_{K_i} = 110,5.

$\left(\sum_{i=1}^{k} f_i\right)_{K_i}$ = 15.

c = 117,5 – 110,5 = 7.

Jadi, $K_1 = 110{,}5 + \left(\frac{25-12}{15}\right)7$ = 110,5 + 6,07 = 116,57.

Kuartil tengah = K_2 = median.

$\frac{2}{4}N = \frac{1}{2}(100)$=50, maka kelas K_1 adalah 118-124.

Jadi, $K_2 = 117{,}5 + \left(\frac{50-27}{35}\right)7$ = 117,5 + 4,6 = 122,1.

Kuartil atas = K_3.

$\frac{3}{4}N = \frac{3}{4}(100)$=75, maka kelas K_1 adalah 125-131.

Jadi, $K_3 = 124{,}5 + \left(\frac{75-62}{25}\right)7$ = 124,5 + 3,64 = 128,14.

5.2. Desil

Desil adalah nilai batas jika sekumpulan data dibagi menjadi sepuluh bagian yang sama, setelah data diurutkan dari kecil ke besar.

$x_1 \quad D_1 \quad D_2 \quad D_3 \quad D_4 \quad D_5 \quad D_6 \quad D_7 \quad D_8 \quad D_9 \quad x_n$

D_1 = desil pertama

D_2 = desil kedua

$\vdots$

D_9 = desil ke-9.

Untuk menentukan letak desil data tunggal, pada prinsipnya sama dengan menentukan titik kuartil. Letak desil data tunggal ditentukan dengan rumus :

Letak D_i = data ke- $\dfrac{i(n+1)}{10}$

dengan :

n = banyak data

i = 1, 2, …, 9.

Untuk menentukan desil pada data berbentuk distribusi frekuensi, pada prinipnya sama dengan kuartil data berkelompok. Desil data berkelompok didefinisikan sebagai :

$$D_i = L_{D_i} + \left(\frac{\frac{i}{10}N - \left(\sum_{i=1}^{k} f_i \right)_{D_i}}{f_{D_i}} \right) c$$

dengan :

L_{D_i} = tepi bawah kelas yang memuat D_i.

N = jumlah seluruh frekuensi.

$\left(\sum_{i=1}^{k} f_i \right)_{D_i}$ = jumlah frekuensi sebelum D_i.

f_{D_i} = frekuensi kelas yang memuat D_i.

C = panjang interval kelas.

I = 1, 2, …, 9.

Contoh 5.2. :

Perhatikan tabel pada contoh sebelumnya. Tentukan D_3, D_5, dan D_8.

Jawab :

Desil ketiga = D_3.

$\dfrac{3}{10}N = \dfrac{3}{10}(100) = 30$.

Jadi, $D_3 = 117{,}5 + \left(\dfrac{30-27}{35} \right) 7$ = 117,5 + 0,6 = 118,1.

Desil ketiga = D_5.

$$\frac{5}{10}N = \frac{5}{10}(100) = 50.$$

Jadi, $D_5 = 117{,}5 + \left(\frac{50-27}{35}\right)7$ = 117,5 + 4,6 = 122,1.

Desil ketiga = D_8.

$$\frac{8}{10}N = \frac{8}{10}(100) = 80.$$

Jadi, $D_8 = 124{,}5 + \left(\frac{80-62}{25}\right)7$ = 124,5 + 5,04 = 129,54.

5.3. Persentil

Persentil adalah nilai batas jika sekumpulan data dibagi menjadi seratus bagian yang sama, setelah data diurutkan dari kecil ke besar : $P_1 \le P_2 \le \ldots \le P_{99}$,

dengan :

P_1 = persentil pertama

P_2 = persentil kedua

$\vdots$

P_{99} = persentil ke-99.

Untuk data tunggal, letak persentil didefinisikan sebagai :

Letak P_i = data ke-$\frac{i(n+1)}{100}$

dengan :

n = banyak data

i = 1, 2, …, 99.

Pada data berkelompok, persentil didefinisikan sebagai :

$$P_i = L_{P_i} + \left(\frac{\frac{i}{100}N - \left(\sum_{i=1}^{k} f_i\right)_{P_i}}{f_{P_i}}\right)c$$

dengan :

L_{P_i} = tepi bawah kelas yang memuat P_i.

N = jumlah seluruh frekuensi.

$\left(\sum_{i=1}^{k} f_i\right)_{P_i}$ = jumlah frekuensi sebelum P_i.

f_{P_i} = frekuensi kelas yang memuat P_i.

c = panjang interval kelas.

I = 1, 2, …, 99.

Contoh 5.3. :

Tentukan P_{15}, P_{50}, dan P_{95} dari data berikut ini :

Nilai Ulangan	Banyaknya Siswa
5	1
6	4
7	8
8	2
Jumlah	15

Jawab :

Letak P_{15} = data ke-$\frac{15(15+1)}{100} = \frac{240}{100} = 2{,}4$.

$P_{15} = 6 + 0{,}4(6-6) = 6 + 0 = 6.$

Letak P_{50} = data ke-$\frac{50(15+1)}{100} = \frac{800}{100} = 8$.

$P_{50} = 7$.

Letak P_{95} = data ke-$\frac{95(15+1)}{100} = \frac{1.520}{100} = 15{,}2.$

P_{95} = tidak ada.

Latihan

1. Tentukan kuartil bawah, kuartil tengah, dan kuartil atas dari data berikut : 60, 70, 90, 80, 50, 100.

2. Tentukan kuartil bawah, kuartil tengah, dan kuartil atas dari data berikut :

Nilai	1	2	3	4	5	6	7	8	9
Frekuensi	1	2	5	7	10	8	4	2	1

3. Diberikan data sebagai berikut :

Nilai	f_i
40 – 49	12
50 – 59	25
60 – 69	19
70 – 79	3
80 – 89	10
90 – 99	4
100 – 109	2

Tentukan :

a. Desil ke-8.

b. Persentil ke-65.

UNIT 6

Ukuran Penyebaran Data

6.1. Pengertian dan Kegunaan

Penyebaran atau dispersi adalah perserakan dari nilai observasi terhadap nilai rata-ratanya. Rata-rata dari serangkaian nilai observasi tidak dapat diinterpretasikan secara terpisah dari hasil dispersi nilai-nilai tersebut sekitar rata-ratanya. Makin besar variasi nilai x_i, makin kurang representatif rata-rata distribusinya.

Contoh 6.1. :

Diberikan tabel hasil test mahasiswa A dan B :

Mahasiswa	Hasil Tes					
A	60	65	50	60	65	60
B	30	90	50	70	60	60

Mahasiswa A : $\overline{X}_A = 60$, variasi nilai dari 50 sampai 65.

Mahasiswa B : $\overline{X}_B = 60$, variasi nilai dari 30 sampai 90.

Bisa kita lihat $\overline{X}_A = \overline{X}_B$.

Meskipun rata-rata hasil tes mereka sama, tetapi dispersi hasil tes mahasiswa B lebih besar dari pada mahasiswa A. Nilai A lebih konsisten (stabil) dari pada nilai B. Sedang nilai B kadang baik, kadang jelek. Hal ini berarti prestasi nilai A lebih baik (stabil) dari pada B.

Berdasarkan besar kecilnya penyebaran, kelompok data dibagi menjadi dua, yaitu :

a. Kelompok data homogen

   Penyebaran relatif kecil; jika seluruh data sama, maka disebut kelompok data homogen 100%.

b. Kelompok data heterogen

   Penyebarannya relatif besar.

Kegunaan ukuran penyebaran antara lain sebagai berikut :

a. Ukuran penyebaran dapat digunakan untuk menentukan apakah nilai rata-ratanya benar-benar representatif atau tidak. Apabila suatu kelompok data mempunyai penyebaran yang tidak sama terhadap nilai rata-ratanya, maka dikatakan bahwa nilai rata-rata tersebut tidak representatif. Perhatikan contoh berikut :

Karyawan	Upah (Rp)
A	40000
B	50000
C	55000
D	65000
E	390000
Jumlah	600000

Rata-rata upah karyawan = Rp $\frac{600.000}{5}$ = Rp 120.000,00

Jelas nilai rata-rata ini tidak representatif, karena ada 4 karyawan yang upahnya dibawah rata-rata. Hal ini diakibatkan oleh sebaran data yang sangat heterogen.

b. Ukuran penyebaran dapat digunakan untuk mengadakan perbandingan terhadap variabilitas data.

c. Ukuran penyebaran dapat membantu penggunaan ukuran statistika, misalnya dalam pengujian hipotesis, apakah dua sampel berasal dari populasi yang sama atau tidak.

6.2. Pengukuran Jangkauan (Range)

Penentuan jangkauan atau rentang sebuah distribusi merupakan pengukuran dispersi yang paling sederhana. Jangkauan sebuah distribusi frekuensi dirumuskan sebagai beda antara pengukuran nilai terbesar dan nilai terkecil yang terdapat dalam sebuah distribusi.

Rumus : $R = X_t - X_r$, dengan :

R = Range

X_t = Nilai tertinggi

X_r = Nilai terendah

Contoh 6.2. :

1. Pandang tabel nilai ujian mahasiswa FE UI :

Tabel nilai mahasiswa FE UI

53,53	63,14	49,03	55,15	67,79
63,49	58,63	50,84	51,77	41,22
73,55	50,74	56,00	46,98	46,33
62,66	66,60	59,16	50,37	44,82
52,49	53,35	61,61	55,54	50,94

Jangkauan distribusi dari nilai mahasiswa FE UI adalah = Nilai tertinggi – nilai terendah = 73,33 – 41,22 = 32,33.

2. Diberikan tabel distribusi frekuensi dari nilai 111 mahasiswa FE UI.

Nilai Ujian	Jumlah Mahasiswa
20,00-27,49	3
27,50-34,99	5
35,00-42,49	7
42,50-49,99	23
50,00-57,49	40
57,50-64,99	20
65,00-72,49	10
72,50-79,99	3

Bila nilai-nilai observasi telah dikelompokkan ke dalam distribusi frekuensi, maka jangkauan distribusi dirumuskan sebagai beda antara pengukuran nilai titik tengah kelas pertama dan nilai titik tengah kelas terakhir.

Jangkauan distribusi nilai mahasiswa FE UI adalah :

Nilai titik tengah kelas pertama = $\frac{(27{,}49 - 20{,}00)}{2}$ = 24,995

Nilai titik tengah kelas terakhir = $\frac{(79{,}99 - 72{,}50)}{2}$ = 74,995

Jangkauan distribusi = nilai titik tengah kelas pertama – nilai titik tengah kelas terakhir = 74,995 – 24,995 = 50,00.

Beberapa statistisi cenderung menggunakan beda antara tepi bawah kelas pertama dengan tepi atas kelas terakhir :

Tepi bawah dari kelas pertama = 20,00

Tepi atas kelas terakhir = 79,99

Jangkauan distribusi = 79,99 – 20,00 = 60,00

3. Tentukan jangkauan dari tabel distribusi frekuensi berikut :

Interval Kelas	*Fi*
97 - 103	4
104 - 110	8
111 - 117	15
118 - 124	35
125 - 131	25
132 - 138	6
139 - 145	4
146 - 152	3
	100

Jawab :

Nilai titik tengah kelas pertama = $\frac{(103-97)}{2} = 100$

Nilai titik tengah kelas terakhir = $\frac{(152-146)}{2} = 149$

Jangkauan distribusi = nilai titik tengah kelas pertama – nilai titik tengah kelas terakhir = 149 – 100 = 49.

Atau :

Jangkauan = 152 – 97 = 55.

6.3. Pengukuran Deviasi Kuartil

Median didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi dua bagian yang sama.

Dengan cara yang sama, kuartil didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi empat bagian yang sama. Ketiga nilai tersebut dinamakan nilai-nilai kuartil dan dilambangkan dengan :

Q_1 = kuartil pertama

Q_2 = kuartil kedua

Q_3 = kuartil ketiga.

Pada distribusi kuartil, 50% dari semua nilai observasi seharusnya terletak di antara Q_1 dan Q_3. Jangkauan antara Q_1 dan Q_3 dinamakan jangkauan inter-kuartil (*inter-quartile-range*). Makin kecil jangkauan tersebut, makin tinggi tingkat konsentrasi distribusi tengah seluas 50% dari seluruh distribusi. Rumus jangkauan kuartil adalah :

$H = Q_3 - Q_1$.

Pengukuran dispersi atas dasar jangkauan inter-kuartil dinamakan deviasi *kuartil* atau *simpangan kuartil* (*quartile deviation*) :

$$d_q = \frac{Q_3 - Q_1}{2}.$$

Contoh 6.3. :

1. Pandang tabel tingkat kematian karena bunuh diri laki-laki usia 25-34 tahun (per 100.000 orang) pada tahun 1971 :

Negara	Jumlah	Negara	Jumlah
Kanada	22	Italia	7
Israel	9	Belanda	8
Jepang	22	Polandia	26
Austria	29	Spanyol	4
Perancis	16	Swedia	28
Jerman	28	Swiss	22
Hongaria	48	Inggris	10
Italia	7	Amerika Serikat	20

Data tersebut kita urutkan dari yang terkecil menuju yang terbesar :

<u>Negara</u>	<u>Jumlah</u>	
Hongaria	48	
Austria	29	
Swedia	28	
Jerman	28	Kuartil ketiga = Q_3
Polandia	26	
Swiss	22	
Kanada	22	
Jepang	22	Kuartil kedua = Q_2
Amerika Serikat	20	
Perancis	16	
Inggris	10	
Israel	9	Kuartil pertama = Q_1
Belanda	8	
Italia	7	
Spanyol	4	

Jangkauan kuartil :

$H = Q_3 - Q_1 = 28 - 9 = 19$.

Deviasi kuartil (rentang antar kuartil) :

$$d_q = \frac{Q_3 - Q_1}{2} = \frac{28-9}{2} = \frac{19}{2} = 9{,}5$$

2. Pandang data jumlah penduduk Kanada tahun 1851-1961 (dalam jutaan) :

Tahun Sensus	Penduduk
1851	2.44
1861	3.23
1871	3,69
1881	4,32
1881	4.32
1891	4.83
1901	5,37
1911	7,21
1911	7.21
1921	8.79
1931	10,38
1941	11,51
1941	11.51
1951	14.01
1961	18.24

Kuartil pertama = $Q_1 = \frac{3{,}69 + 4{,}32}{2} = 3{,}46$

Kuartil kedua = $Q_2 = \frac{5{,}37 + 7{,}21}{2} = 6{,}29$

Kuartil ketiga = $Q_3 = \frac{10{,}38 + 11{,}51}{2} = 10{,}95$

Jangkauan kuartil :

$$H = Q_3 - Q_1 = 10{,}95 - 3{,}46 = 7{,}49.$$

Deviasi kuartil :

$$d_q = \frac{Q_3 - Q_1}{2} = \frac{10{,}95 - 3{,}46}{2} = \frac{7{,}49}{2} = 3{,}74.$$

6.4. Rata-rata Simpangan

Rata-rata simpangan adalah suatu simpangan nilai untuk observasi terhadap rata-rata. Rata-rata simpangan sering disebut simpangan rata-rata atau mean deviasi, yang dilambangkan dengan "*SR*". Untuk data tunggal, rata-rata simpangan ditentukan dengan rumus :

$$SR = \frac{\sum_{i=1}^{n} |x_i - \overline{X}|}{n}.$$

Untuk data berkelompok, rata-rata simpangan ditentukan dengan rumus :

$$SR = \frac{\sum_{i=1}^{K} f_i \left| x_i - \overline{X} \right|}{N} .$$

Contoh 6.4. :

1. Tentukan simpangan rata-rata dari 7, 5, 8, 4, 6, dan 10 !

Jawab :

x	$\left\| x_i - \overline{X} \right\|$
7	0,33
5	-1,67
8	1,33
4	-2,67
6	-0,67
10	3,33
40	10
6.6	
7	

$$SR = \frac{\sum_{i=1}^{6} \left| x_i - 6{,}67 \right|}{6} = \frac{10}{6} = 1{,}67 .$$

2. Tentukan simpangan rata-rata dari distribusi frekuensi berikut :

x	f_i
5	1
6	4
7	8
8	2

Jawab :

x	f_i	$f_i x_i$	$\left\| x_i - \overline{X} \right\|$	$f_i \left\| x_i - \overline{X} \right\|$
5	1	5	1.73	1.73
6	4	24	0.73	2.92
7	8	56	0.27	2.16
8	2	16	1.27	2.54
	15	101		9.35
		6.73		

$$SR = \frac{\sum_{i=1}^{15} f_i \left| x_i - 6{,}73 \right|}{15} = \frac{9{,}35}{15} = 0{,}62 .$$

6.5. Pengukuran Variansi dan Deviasi Standar

Rumus variansi dan deviasi standar populasi adalah :

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

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}\ .$$

dengan :

N = Jumlah observasi dalam populasi

μ = Rata-rata populasi.

Untuk populasi yang berjumlah besar, sangat tidak mungkin untuk mendapatkan nilai μ dan σ. Untuk mengestimasi (menaksir) nilai μ dan σ, diambil sampel data. Nilai μ diestimasi oleh $\overline{X}$ dan σ diestimasi oleh S.

6.6. Variansi dan deviasi standar dari data data tunggal

Simpangan baku atau deviasi standar (*Standard Deviation*) merupakan ukuran penyebaran yang paling baik, karena menggambarkan besarnya penyebaran tiap-tiap unit observasi. Karl Pearson menamakannya deviasi standar dan dirumuskan sebagai :

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2}\ .$$

Kuadrat dari deviasi standar dinamakan ***variansi*** :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2\ .$$

Contoh 6.6. :

Pandang tabel jumlah pemakaian tenaga listrik per bulan di DKI Jakarta tahun 1978.

Bulan	Jumlah Pemakaian dalam Juta Kw H = *X*	$x_i - \overline{X}$	$(x_i - \overline{X})^2$
Januari	111	-8,67	75,11

Februari	109	-10,67	113,78
Maret	105	-14,67	215,11
April	118	-1,67	2,78
Mei	117	-2,67	7,11
Juni	125	5,33	28,44
Juli	123	3,33	11,11
Agustus	123	3,33	11,11
September	126	6,33	40,11
Oktober	120	0,33	0,11
Nopember	128	8,33	69,44
Desember	131	11,33	128,44
$\sum$	1436	0	702,67
$\overline{X}$	119,67		

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2 = \frac{1}{12}702{,}67 = 58{,}56$$

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2} = \sqrt{58{,}56} = 7{,}65.$$

6.7. Rumus Fisher dan Wilks

Untuk distribusi sampel dengan n < 100, Fisher, Wilks dan beberapa statistisi memberi perumusan tentang variansi dan deviasi standar sebagai berikut :

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}\ .$$

Deviasi standar sampel di atas sebetulnya digunakan sebagai penaksir tak bias (*unbiased estimate*) bagi deviasi standar populasi σ. Banyak statistisi yang menganjurkan penggunaan pembagi n-1 dalam menghitung deviasi standar sampel guna menaksir deviasi standar populasi. Bila jumlah *n* tidak besar, hasil penggunaan kedua rumus mungkin mempunyai perbedaan yang berarti. Tapi jika jumlah *n* besar sekali, beda kedua rumus di atas tidak berarti.

6.8. Rumus Alternatif bagi Variansi dan Deviasi Standar Sampel

$$S^2 = \frac{\sum_{i=1}^{n} x_i - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} x_i - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n}}\ .$$

Contoh 6.8. :

Pandang kembali tabel jumlah pemakaian tenaga listrik per bulan di DKI Jakarta tahun 1978.

Bulan	Jumlah Pemakaian dalam Juta Kw H = *X*	x_i^2
Januari	111	12321
Februari	109	11881
Maret	105	11025
April	118	13924
Mei	117	13689
Juni	125	15625
Juli	123	15129
Agustus	123	15129
September	126	15876
Oktober	120	14400
Nopember	128	16384
Desember	131	17161
	1436	172544
	119.67	

$$S^2 = \frac{\sum_{i=1}^{n} x_i - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n} = \frac{1436 - \frac{1436^2}{12}}{12} = 58{,}56$$

$$S = \sqrt{S} = \sqrt{58{,}56} = 7{,}65.$$

6.9. Cara Menghitung Variansi dan Deviasi Standar Secara Singkat

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - x_0)^2 - \frac{\left(\sum_{i=1}^{n} (x_i - x_0)\right)^2}{n}}{n}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x_0)^2 - \frac{\left(\sum_{i=1}^{n} (x_i - x_0)\right)^2}{n}}{n}}$$

dengan : x_0 = titik asal deviasi secara arbriter.

Contoh 6.9. :

Pandang kembali tabel jumlah pemakaian tenaga listrik per bulan di DKI Jakarta tahun 1978.

Bulan	Jumlah Pemakaian dalam Juta Kw H = *X*	$x_i - x_0$	$(x_i - x_0)^2$
Januari	111	-12	144
Februari	109	-14	196
Maret	105	-18	324
April	118	-2	4
Mei	117	-6	36
Juni	125	2	4
Juli	123	0	0
Agustus	123	0	0
September	126	3	9
Oktober	120	-3	9
Nopember	128	5	25
Desember	131	8	64
	1436	-37	815
	119.67		

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - x_0)^2 - \frac{\left(\sum_{i=1}^{n}(x_i - x_0)\right)^2}{n}}{n} = \frac{815 - \frac{(-34)^2}{12}}{12} = 58{,}41$$

$$S = \sqrt{58{,}41} = 7{,}64.$$

6.10. Variansi dan Deviasi Standar dari Data yang Telah Dikelompokkan

Bila variansi dan deviasi standar dihitung dari sebuah distribusi frekuensi, maka titik tengah tiap-tiap kelas umumya dianggap sebagai nilai tunggal yang cukup representatif bagi semua nilai-nilai observasi yang dikelompokkan ke dalam kelas-kelas yang bersangkutan. Rumus variansi dan deviasi standar dari distribusi frekuensi sedemikian itu dapat diberikan sebagai :

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}(m_i - \overline{X})^2 f_i$$

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(m_i - \overline{X})^2 f_i}$$

dengan :

m_i = titik tengah tiap-tiap kelas

f_i = jumlah frekuensi kelas.

Contoh 6.9. :

Nilai Ujian	m_i	$(m_i - \overline{X})^2$	f_i	$(m_i - \overline{X})^2 f_i$
0 - 9,99	4.99	2649.16	1	2649.16
10 - 19,99	14.99	1719.76	4	6879.04
20 - 29,99	24.99	990.36	7	6932.53
30 - 39,99	34.99	460.96	31	14289.79
40 - 49,99	44.99	131.56	42	5525.56
50 - 59,99	54.99	2.16	54	116.69
60 - 69,99	64.99	72.76	33	2401.11
70 - 79,99	74.99	343.36	24	8240.66
80 - 89,99	84.99	813.96	22	17907.14
90 - 99,99	94.99	1484.56	8	11876.49
			226	76818.16

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(m_i - \overline{X}\right)^2 f_i = \frac{1}{226} 76818{,}16 = 339{,}90$$

$$S = \sqrt{339{,}90} = 18{,}44.$$

6.10. Cara Menghitung Variansi dan Deviasi Standar Secara Singkat

$$S = i\sqrt{\frac{\sum_{i=1}^{k} u_i^2 f_i}{n} - \left(\frac{\sum_{i=1}^{k} u_i f_i}{n}\right)^2}.$$

Contoh 6.10. :

Nilai Ujian	m_i	f_i	u_i	u_i^2	$u_i f_i$	$u_i^2 f_i$
0 - 9,99	4.99	1	-5	25	-5	25
10 - 19,99	14.99	4	-4	16	-16	64
20 - 29,99	24.99	7	-3	9	-21	63
30 - 39,99	34.99	31	-2	4	-62	124
40 - 49,99	44.99	42	-1	1	-42	42
50 - 59,99	54.99	54	0	0	0	0
60 - 69,99	64.99	33	1	1	33	33
70 - 79,99	74.99	24	2	4	48	96
80 - 89,99	84.99	22	3	9	66	198
90 - 99,99	94.99	8	4	16	32	128
		226			33	773

$$S = i\sqrt{\frac{\sum_{i=1}^{k} u_i^2 f_i}{n} - \left(\frac{\sum_{i=1}^{k} u_i f_i}{n}\right)^2} = \sqrt{\frac{773}{226} - \left(\frac{33}{226}\right)^2} = 18{,}4.$$

6.11. Pengertian Distribusi Normal

Distribusi normal merupakan distribusi teoritis. Pada abad permulaan 19, sebagian besar sarjana beranggapan bahwa distribusi hasil observasi mengikuti *hukum normal* tersebut. Sebetulnya tidak semua distribusi hasil observasi bersifat normal, karena para sarjana mulai menemukan distribusi lain (seperti distribusi Poisson, Fisher, dll.). Meskipun demikian, kenyatan menunjukkan bahwa distribusi-distribusi hasil observasi memilki kurva frekuensi yang bermodus tunggal dengan kedua ujung yang mendatar ke arah kiri dan kanan serta cenderung simetris. Kurva simetris itu dekat sekali persamaannya dengan *kurva normal* yang biasa disebut *kurva Gauss*.

Distribusi normal dengan $\mu = 0$ dan $\sigma = 1$ dinamakan distribusi normal standar atau distribusi normal baku. Nilai distribusi normal baku sudah dibuat tabelnya, sehingga kita dapat menghitung nilai standar dengan mudah, dengan membakukan nilai observasi. Caranya adalah sebagai berikut :

Misalkan $x_1, x_2, \ldots, x_n$ adalah nilai observasi dengan rata-rata $\overline{X}$ dan deviasi standar S. Nilai observasi dapat diubah menjadi nilai standar, dinotasikan dengan Z, dengan menggunakan rumus :

$$Z_i = \frac{x_i - \overline{X}}{S}$$

Nilai standar $Z_1, Z_2, \ldots, Z_n$ mempunyai $\mu = 0$ dan $\sigma = 1$.

Contoh 6.11 :

1. Suatu kelompok data mempunyai rata-rata 25 dan simpangan standar 4. Salah satu datanya bernilai 30. Nyatakan nilai mentah itu ke dalam nilai standar.
   Jawab :

   Diketahui : $x_i = 30$, $\mu = 25$ dan $\sigma = 4$

   $$Z = \frac{x_i - \overline{X}}{S} = \frac{30 - 25}{4} = \frac{5}{4} = 1{,}25.$$

2. Seorang siswa SMK mendapat nilai ujian akhir Matematika 85. Rata-rata ujian Matematika 76 dan simpangan bakunya 9. Untuk bidang studi akuntansi, siswa tersebut mendapat nilai 90 dengan rata-rata ujian akuntansi 80 dan simpangan baku 15. Dalam mata pelajaran manakah ia mendapat kedudukan lebih baik ?

   Jawab :

   Nilai standar untuk matematika : $Z = \frac{85 - 76}{9} = 1$

   Nilai standar untuk akuntansi : $Z = \frac{90 - 80}{15} = 0{,}67$.

Nilai tersebut menggambarkan bahwa siswa tersebut mendapat satu simpangan di atas rata-rata nilai matematika dan mendapat 0,67 simpangan di atas rata-rata nilai akuntansi. Hal itu berarti kedudukan siswa tersebut lebih tinggi dalam mata pelajaran matematika.

6.12. Pengukuran Dispersi Relatif

6.12.1. Pengertian

Pengukuran jangkauan, deviasi kuartil, deviasi rata-rata dan deviasi standar merupakan pengukuran yang absolut. Pengukuran demikian itu sebetulnya hanya dapat digunakan bagi penggambaran dispersi nilai-nilai observasi sebuah distribusi secara definitive. Bila kita ingin melakukan perbandingan tingkat dispersi antara dua atau beberapa distribusi dan bila jumlah nilai-nilai observasi dari dua atau beberapa distribusi di atas tidak sama, maka pengukuran dispersi secara absolut sebagai metode guna membandingkan dispersi akan memperoleh hasil yang menyesatkan.
Contoh 6.12.1. :

Seorang pengusaha bangunan ingin membandingkan variasi gaji buruh ekstranya dengan variasi gaji stafnya. Gaji buruh dibayar secara harian, sedangkan gaji staf dibayar sebulan sekali. Rata-rata gaji buruh Rp 500,00 dengan deviasi standar Rp 150,00 ; sedangkan gaji rata-rata staf Rp 30.000,00 dengan deviasi standar Rp 15.000,00.

Perbandingan langsung dari hasil perhitungan deviasi standar tentu tidak memungkinkan. Gaji staf dibayar per bulan tentu jumlahnya lebih besar dari pada gaji buruh yang dibayar harian, sehingga dispersi gaji staf lebih besar dari dispersi gaji buruh.

6.12.2. Cara Menghitung Ko-efisien Variansi

Dalam membandingkan tingkat variasi dua atau lebih distribusi hendaknya rata-rata distribusi digunakan sebagai dasar pengukuran variasinya secara relatif dan dinamakan **ko-efisien variasi** (*co-efficient of variation*) :

$$V = \frac{S}{\overline{X}}$$

dengan :

S = deviasi standar sampel

$\overline{X}$ = rata-rata hitung sampel

Contoh 6.12.2. :

1. Sepeda motor jenis A dapat dipakai dalam kondisi prima rata-rata selama 40 bulan dengan simpangan baku 8 bulan. Jenis B 36 bulan dengan simpangan standar 6 bulan. Tentukan koefisien variasi dari masing-masing jenis sepeda motor tesebut dan interpretasinya.

   Jawab :

   $V_A = \frac{S}{\overline{X}} = \frac{8}{40}$ x 100 % = 20 %.

   $V_B = \frac{S}{\overline{X}} = \frac{6}{36}$ x 100 % = 16,7 %.

   Nilai tersebut berarti masa pakai sepeda motor B dalam kondisi prima lebih seragam (uniform) bila dibandingkan dengan masa pakai kondisi prima sepeda motor A.

2. Tentukan koefisien variasi dari data tabel berikut :

Interval Kelas	*fi*
97-103	4
104-110	8
11-117	15
118-124	35
125-131	25
132-138	6
139-145	4
146-152	3
	100

   Jawab :

   Dari hasil perhitungan diperoleh :

   $\overline{X} = 122{,}26$ dan $S_S = 10{,}21$.

   Jadi, $V = \frac{10{,}21}{122{,}26}$ x 100% = 8,35%.

6.12.3. Cara Menghitung Ko-efisien Variasi Kuartil

Salah satu rumus yang paling sering digunakan adalah :

$$V_Q = \frac{(Q_3 - Q_1)/2}{md}.$$

Bila nilai *md* tidak diperoleh, maka niali *md* dapat dicari dengan rumus : $(Q_3 + Q_1)/2$, sehingga rumus di atas menjadi :

$$V_Q = \frac{(Q_3 - Q_1)/2}{(Q_3 + Q_1)/2} \quad \Leftrightarrow \quad V_Q = \frac{(Q_3 - Q_1)}{(Q_3 + Q_1)}.$$

Contoh 6.12.3. :

Dari contoh 6.3. soal nomor 2 :

Jika diberikan : Q_1 = 3,46; Q_2 = 6,29; Q_3 = 10,95, maka

$$V_Q = \frac{(Q_3 - Q_1)}{(Q_3 + Q_1)} = \frac{(10{,}95 - 3{,}46)}{(10{,}95 + 3{,}46)} = \frac{7{,}49}{14{,}41} = 0{,}52.$$

Latihan

1. Tentukan jangkauan semi interkuartil dari data : 12, 9, 8, 19, 20, 7, 5, 19, 16, 13, 18, 18.

2. Tentukan simpangan rata-rata, simpangan baku, dan jangkauan 50 – 90 persentil dari data berikut :

Tinggi (cm)	Frekuensi
150 – 154	5
155 – 159	20
160 – 164	42
165 – 169	26
170 – 174	7

3. Tentukan koefisien variasi data soal no. 2.

4. Hasil ulangan mata pelajaran Matematika yang diikuti oleh 20 siswa adalah sebagai berikut :

   62, 95, 54, 38, 77, 68, 61, 70, 92

   45, 65, 78, 81, 66, 50, 67, 75, 90, 83.

   a. Berapakah rata-rata nilai Matematika siswa di atas ?
   b. Berapakah deviasi standarnya ?
   c. Berilah komentar singkat tentang hasil penghitungan anda !

5. Seorang siswa memperoleh nilai 70 untuk ulangan Matematika dan 90 untuk ulangan Akuntansi. Hasil rata-rata ulangan Matematika bagi seluruh kelas adalah 64 dan berdeviasi standar 12. Sedangkan hasil rata-rata ulangan Akuntansi bagi seluruh kelas adalah 72 dan berdeviasi standar 10. Keterangan apa yang mungkin anda peroleh mengenai kemampuan siswa dalam kedua mata pelajaran tersebut !

Daftar Pustaka

Anto Dajan.1987. *Pengantar Metode Statistika Jilid 1 dan 2*. Jakarta : LP3ES.

B.H. Erickson dan T.A. Nosanchuk.1983. *Memahami Data*. Jakarta : LP3ES.

Edward J. Dudewicz dan Satya N. Mishra.1987. *Modern Mathematical Statistics*. New York : John Wiley & Sons.

Hadi Sutrisno.1986. *Statistika*. Jakarta : Universitas Terbuka.

Husaini Usman, dan R. Purnomo Setiady Akbar.1995. *Pengantar Statistika*. Jakarta : Bumi Aksara.

Lee j. Bain dan Max Engelhardt.1992. *Introduction to Probability and Mathematical Statistics*. California : Duxbury Press.

Leonard J. Kazmier.1976. *Bussiness Statistics*. New York : Mc.Graw-Hill.

Marzuki Mustafa.1986. *Pengantar Teori Statistika*. Jakarta : STIE Perbanas.

Mason dan Brambel.1987. *Understanding and Conducting Research*. New York : Mc. Graw-Hill.

Sudjana.1982. *Metode Statistika*. Bandung : Tarsito.

Supranto.1986. *Statistika Teori dan Aplikasi*. Jakarta : Erlangga.

Tim Matematik SMK.2001. *Matematika 3*. Jakarta : PT Galaxy Puspa Mega.