

MODUL STATISTIK FASILITAS PELAYANAN KESEHATAN



Disusun Oleh:

**TIM DOSEN PRODI SARJANA TERAPAN MANAJEMEN INFORMASI KESEHATAN
UNIVERSITAS INDONESIA MAJU**

**PROGRAM STUDI SARJANA TERAPAN MANAJEMEN INFORMASI KESEHATAN
FAKULTAS VOKASI**

UNIVERSITAS INDONESIA MAJU

JAKARTA

2022



Modul Statistik Fasilitas Pelayanan Kesehatan

Nama Mahasiswa :

NPM :

Program Studi Sarjana Terapan Manajemen Informasi Kesehatan

Fakultas Vokasi

Universitas Indonesia Maju

2022

KATA PENGANTAR

Buku petunjuk praktikum ini disusun untuk memenuhi kebutuhan mahasiswa sebagai panduan dalam melaksanakan praktikum Statistik Fasilitas Pelayanan Kesehatan, untuk mahasiswa program studi D4 Manajemen Informasi Kesehatan (MIK) UIMA. Dengan adanya buku petunjuk praktikum ini diharapkan akan membantu dan mempermudah mahasiswa dalam memahami dan melaksanakan praktikum Statistik Fasilitas Pelayanan Kesehatan sehingga akan memperoleh hasil yang baik.

Materi yang dipraktikkan merupakan materi yang selaras dengan materi kuliah Statistik Fasilitas Pelayanan Kesehatan. Untuk itu dasar teori yang didapatkan saat kuliah juga akan sangat membantu mahasiswa dalam melaksanakan praktikum ini.

Buku petunjuk ini masih dalam proses penyempurnaan. Insha Allah perbaikan akan terus dilakukan demi kesempurnaan buku petunjuk praktikum ini dan disesuaikan dengan perkembangan ilmu pengetahuan. Semoga buku petunjuk ini dapat dipergunakan sebagaimana mestinya.

Jakarta, September 2022

Penyusun

DAFTAR ISI

HALAMAN SAMPUL.....	i
KATA PENGANTAR	iii
DAFTAR ISI.....	iv
BAB I STATISTIKA DESKRIPTIF	1
BAB II NILAI TENGAH ATAU NILAI PUSAT	7
BAB III UKURAN POSISI.....	17
BAB IV UKURAN VARIASI (<i>DISPERSI</i>)	23
DAFTAR PUSTAKA.....	35

BAB I

STATISTIKA DESKRIPTIF

Statistika deskriptif adalah statistika yang berfungsi untuk mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui data sampel atau populasi sebagaimana adanya, tanpa melakukan analisis dan membuat kesimpulan yang berlaku untuk umum.

Statistika deskriptif merupakan bidang ilmu pengetahuan statistika yang mempelajari tata cara penyusunan dan penyajian suatu data yang dikumpulkan dalam satu penelitian. Proses mengklasifikasi statistika deskriptif dan statistika inferensial dilakukan berdasarkan aktivitas yang dilakukan (Martono, 2010).

Statistika deskriptif adalah metode statistika yang digunakan untuk menggambarkan atau mendeskripsikan data yang telah dikumpulkan menjadi sebuah informasi (Purwanto, 2012). Statistika deskriptif memberikan informasi mengenai data yang dipunyai dan sama sekali tidak menarik kesimpulan apapun tentang gugus induknya yang lebih besar.

A. STATISTIKA DESKRIPTIF VERSUS STATISTIKA INFERENSIAL

Ilmu statistika tersegmentasi menjadi dua kategori deskriptif dan inferensial. Statistika deskriptif digunakan untuk menggambarkan distribusi variabel yang sedang diamati atau dipelajari. Statistika inferensial digunakan untuk menguji hipotesis atau membuat keputusan. Hipotesis memiliki probabilitas atau resiko melakukan kesalahan berdasarkan data yang terkumpul. Probabilitas ini disebut sebagai kesalahan Tipe I atau p-value. Statistika inferensial membantu analisis untuk meramalkan dan meringkas data apakah mereka melihat hasil yang signifikan atau hanya mengamati kejadian yang terjadi karena kebetulan.

Statistika deskriptif yang tepat ditentukan oleh jenis data yang dianalisis. Jika ingin menggambarkan seberapa sering suatu peristiwa terjadi, maka tarif dan proporsi sering digunakan. Misalnya, tingkat mortalitas (*mortality rate*) digunakan untuk mengukur berapa banyak subjek meninggal dibandingkan dengan jumlah keseluruhan subjek. Proporsi adalah statistika yang sesuai untuk digunakan saat menjelaskan kasus MS-DRG tanpa komplikasi atau tanpa penyerta (*complication or co-morbidities (CC)*), dengan komplikasi atau penyerta (CC), dan dengan komplikasi atau penyerta mayor (MCC). Jika ingin menganalisis sebuah variabel yaitu interval atau rasio, maka statistika deskriptif yang sesuai kemungkinan adalah mean, median, atau mode. Misalnya, lama rawat untuk layanan rawat inap dijelaskan dengan menggunakan mean (rata-rata) atau median. Tabel 1.1 mencantumkan jenis data dan statistik deskriptif yang sesuai.

Tabel 1.1

Jenis Data dan Statistik Deskriptif (Sumber: Edgerton, 2016, 512)

Jenis Data	Contoh	Statistik Deskriptif yang sesuai
Nominal-Data Kategori	Jenis kelamin, kode HCPCS atau kode ICD-10, Bagian atau Unit	Frekuensi, proporsi, modus
Ordinal-Data Kategori	Skor Kepuasan Pasien, Skor Keparahannya, Level Pusat Trauma, Ukuran Survey dengan Skala Likert	Frekuensi, proporsi, modus, range
Interval-Data Numerik	Suhu, Level pH	Mean, median, standar deviasi (SD), range
Rasio-Data Numerik	Tarif RS, <i>Length of Stay</i> (LOS), Umur	Mean, median, standar deviasi (SD), <i>range</i> , <i>coeficient of varians</i>

Metode statistik inferensial yang sesuai ditentukan oleh hipotesis yang akan diuji atau pertanyaan yang harus dijawab. Seorang analis mungkin diminta untuk membandingkan lama rawat (*length of stay*) atau tingkat kematian (*mortality rate*) dengan standar nasional atau untuk menentukan apakah tingkat perubahan MS-DRG berbeda dengan nilai yang biasanya diamati pada fasilitas tersebut. Pertanyaan ini bisa dijawab dengan teknik inferensial. Metode statistik yang tepat tergantung pada pertanyaan penelitian dan jenis data yang tersedia.

Pertanyaan penelitian dalam pelayanan kesehatan dasar mungkin sering dianalisis dengan menggunakan *confidence intervals* atau tes hipotesis. *Confidence intervals* (interval kepercayaan) adalah kisaran nilai, sehingga probabilitas kisaran yang mencakup nilai sebenarnya dari parameter adalah probabilitas atau kepercayaan yang ditetapkan. Sebuah uji hipotesis memungkinkan analis untuk menentukan kemungkinan bahwa sebuah hipotesis benar mengingat data dalam sampel dengan tingkat yang dapat diterima yang dapat ditentukan dalam membuat tingkat kesalahan (*level of error*).

Misalnya, tujuan penelitian adalah untuk menentukan waktu tunggu di unit gawat darurat rumah sakit (*hospital's emergency department* atau DE). Sampel acak pasien dipilih dari populasi pasien yang mengunjungi unit gawat darurat selama bulan sebelumnya. Waktu tunggu rata-rata diketahui 53,5 menit, adalah statistik yang dapat menggambarkan waktu tunggu, namun statistik itu sendiri tidak memberikan banyak informasi tentang ketepatan perkiraan atau seberapa pasti waktu tunggu pengunjung unit gawat darurat tersebut sebenarnya adalah kisaran 53,5 menit.

Interval kepercayaan dapat digunakan untuk memberikan informasi tambahan. Interval kepercayaan 95 persen untuk waktu tunggu unit gawat darurat dihitung dan ternyata berada pada kisaran (50,1 menit sampai dengan 56,9 menit). Nilai 95 persen adalah

tingkat kepercayaan atau probabilitas bahwa interval keyakinan mengandung rata-rata populasi. Dalam kasus ini, dengan tingkat kepercayaan 95 persen rata-rata waktu tunggu unit gawat darurat antara 50,1 menit sampai dengan 56,9 menit. Ingatlah bahwa angka-angka ini didasarkan pada sampel dan kita tidak mengetahui keseluruhan nilai populasi. Sampel digunakan untuk membuat kesimpulan atau kesimpulan tentang populasi. Rentang interval kepercayaan adalah ukuran presisi perkiraan. Semakin kecil rentang interval maka akan lebih tepat.

Melanjutkan contoh waktu tunggu di unit gawat darurat ini, misalkan tujuannya adalah untuk menguji apakah rata-rata waktu tunggu memenuhi standar yang ditetapkan rumah sakit 60 menit. Dalam kasus ini, uji hipotesis adalah metode statistik yang benar. Uji hipotesis mensyaratkan definisi hipotesis nol untuk diuji versus hipotesis alternatif. Hipotesis nol biasanya adalah status quo atau netral. Hipotesis alternatif kadang disebut hipotesis penelitian dan merupakan kesimpulan yang biasanya memerlukan beberapa tindakan yang harus dilakukan. Hipotesis nol adalah bahwa waktu tunggu di unit gawat darurat kurang dari atau sama dengan 60 menit; Hipotesis alternatifnya adalah bahwa waktu tunggu di unit gawat darurat lebih besar dari 60 menit. Dalam pengujian hipotesis, tingkat kesalahan Tipe I yang dapat diterima harus dipilih sebelum menghitung hasilnya. Kesalahan Tipe I adalah probabilitas untuk salah menolak hipotesis nol mengingat nilai yang ada dalam sampel. Dalam contoh ini, kesalahan Tipe I membuat kesimpulan bahwa waktu tunggu lebih lama dari 60 menit bila benar-benar kurang dari atau sama dengan 60 menit. Analis juga harus menyadari kesalahan Tipe II dalam pengujian hipotesis. Kesalahan tipe II terjadi ketika hipotesis nol tidak ditolak padahal sebenarnya salah. Kesalahan Tipe II dapat dikendalikan dengan menyesuaikan ukuran sampel yang digunakan untuk penelitian ini; Sampel yang lebih besar akan mengurangi kemungkinan melakukan kesalahan Tipe II.

Implikasi praktis membuat kesalahan Tipe I dalam situasi ini adalah biaya yang mungkin timbul untuk mempelajari akar penyebab dari menunggu lama dan melakukan koreksi operasional. Untuk contoh ini, kesalahan Tipe I ditetapkan menjadi 5 persen atau 0,05. Probabilitas membuat kesalahan Tipe I dapat dihitung dengan menggunakan data sampel dan uji statistika yang sesuai. Probabilitas membuat kesalahan Tipe I berdasarkan kumpulan data tertentu disebut nilai-p (p -value). Jika nilai p atau probabilitas membuat kesalahan Tipe I kurang dari tingkat kesalahan Tipe I yang ditetapkan sebelum pengujian, maka hipotesis nol dapat ditolak. Setelah dihitung uji statistika dan nilai p adalah 0,03, hipotesis nol ditolak dan kesimpulannya dibuat bahwa waktu tunggu di unit gawat darurat lebih lama dari standar.

Langkah pertama dalam penelitian apapun adalah mengidentifikasi pertanyaan penelitian. Banyak proyek analisis data dimulai tanpa gagasan yang jelas tentang pertanyaan yang tepat untuk dijawab. Analis menjadi sangat prihatin dengan merangkum data dan menghasilkan laporan bahwa mereka lupa usaha mereka mungkin terbuang jika pertanyaan tersebut tidak didefinisikan dengan baik sejak awal. Contoh beberapa pertanyaan penelitian yang mungkin menarik berikut ini:

1. Berapa lama waktu tunggu di unit gawat darurat?

2. Apakah waktu tunggu di unit gawat darurat memenuhi standar yang telah ditetapkan?
3. Berapa persen permintaan periksa laboratorium yang tidak ditandatangani oleh dokter?
4. Berapa tingkat akurasi kode pada kode diagnosis sekunder?

Begitu pertanyaan penelitian didefinisikan, unit analisis juga harus ditentukan. Misalnya, untuk menentukan persentase pesanan lab yang tidak ditandatangani oleh dokter, unit analisis adalah keseluruhan rangkaian pesanan lab untuk periode waktu yang diminati. Kumpulan pesanan lab ini berfungsi sebagai populasi yang diminati. Hal ini hampir mustahil untuk mengumpulkan seluruh populasi pesanan laboratorium untuk secara pasti menjawab pertanyaan yang diajukan. Data yang dibutuhkan sering dipilih melalui sampling dan kemudian dianalisis untuk membuat kesimpulan atau kesimpulan tentang persentase pesanan lab yang tidak ditandatangani dalam populasi.

Dalam prakteknya, terkadang sulit untuk menentukan unit analisis. Misalnya, dalam menentukan tingkat akurasi pengkodean untuk kode diagnosis sekunder, apakah unit analisis masing-masing kode atau klaim dari kode itu? Ada perbedaan penting antara dua unit analisis. Jika unit analisisnya adalah kode, maka tingkat yang dihasilkan akan menjadi proporsi kode yang benar dan bukan proporsi klaim yang benar. Jika pertanyaan penelitiannya adalah untuk memperkirakan jumlah klaim yang dikodekan dengan benar, maka klaim tersebut harus menjadi unit analisis.

B. REVIEW PENGHITUNGAN MATEMATIKA DASAR

Sebelum mempelajari untuk menggunakan rumus statistik atau memahami analisis statistik, pengetahuan dan review penghitungan matematika dasar adalah penting. Penghitungan matematika sebagai dasar pada rumus statistik.

1. *Ratio*

Rasio adalah perbandingan dua angka dari jenis data yang sama. Rasio biasanya ditulis sebagai "a per b". Misalnya, analisis pasien dewasa terhadap anak-anak di klinik setempat menemukan bahwa rasio tersebut adalah lima pasien dewasa untuk setiap dua pasien anak-anak. Rasio itu akan diucapkan sebagai "lima per dua" dan akan dituliskan sebagai 5: 2.

Untuk populasi atau kumpulan data dalam jumlah kecil, rasio tersebut dapat dilihat tanpa perhitungan matematis, namun untuk jumlah besar terdapat persamaan pembagian sederhana, yang ditunjukkan dalam rumus berikut. Untuk memastikan rasio dinyatakan dalam jumlah serendah mungkin. Sebagai contoh, sebuah klinik memiliki 5.000 pasien dewasa dan 2.000 pasien anak-anak. Salah satu cara untuk menunjukkan rasionya adalah 5.000: 2.000, namun itu tidak praktis. Dalam kasus ini, kedua bilangan tersebut habis dibagi 1.000 seperti ditunjukkan pada rumus di bawah ini:

Rate

$$\text{Rasio} = \frac{X}{Y} = \frac{\text{Pasien Dewasa}}{\text{Pasien Anak-anak}} = \frac{5000/1000}{2000/1000} = \frac{5}{2} \text{ jadi rasio adalah } 5 : 2$$

Rate adalah jenis rasio di mana kedua angka mewakili hal yang berbeda. Misalnya, angka dapat membandingkan jumlah pasien dan jumlah hari dalam jangka waktu tertentu, atau jumlah pasien dan jumlah pasien yang memiliki penyakit tertentu. *Rate* biasanya dinyatakan sebagai persentase. Untuk alasan ini, sebagian besar *Rate* harus mencakup tanda persen (%) setelah nilainya.

Untuk menghitung *Rate*, ada rumus yang penting untuk diingat. Setelah rumus ini diketahui, semua jenis *Rate* dapat dengan mudah dihitung. Rumus *Rate* adalah:

$$\text{Rate} = \frac{A}{A+B} \times 100$$

Keterangan: A = Jumlah kejadian dalam periode waktu tertentu

B = Jumlah kejadian yang dapat terjadi dalam periode waktu tertentu

Mengalikan *Rate* dengan 100 menempatkan titik desimal di tempat yang tepat untuk satu persentase. Artinya perbedaan antara 0,305 persen dan 30,5 persen. Ini adalah perbedaan yang penting sehingga merupakan langkah yang tidak boleh dilewati.

Misalnya, Ka. Bagian Rekam Medis dan informasi kesehatan (RMIK) atau MIK ingin mengetahui tingkat produktivitas *clinical coder* bulan Juli 2017. Sebanyak 239 rekam medis pada bulan Juli 2017 dan kenyataannya *clinical coder* menyelesaikan koding hanya 192 rekam medis. Untuk menghitung tingkat produktivitas, Ka. Bagian RMIK atau MIK membagi 192 (berapa kali sesuatu terjadi) oleh 239 (berapa kali sesuatu bisa terjadi) dan mengalikan dengan 100, dengan hasilnya adalah 80,33 persen:

$$\text{Rate} = \frac{A}{A+B} \times 100 = \frac{192}{239} \times 100 = 80,33$$

Ketika melaporkan tingkat perawatan kesehatan, penting untuk mengumpulkan sampai dengan dua desimal. Dalam melaporkan statistik kesehatan, sering kali lebih bermakna dengan menunjukkan lebih detail daripada kurang detail. Dalam pelaporan kualitas atau penyakit, 5.09, 5.37, dan 5.49 dapat menandakan perbedaan penting meskipun semuanya bisa dibulatkan ke 5. Misalnya, jika ini adalah hasil laboratorium, perbedaan kecil bisa berarti perbedaan antara tingkat normal dan abnormal.

Average (Rata-rata)

Perhitungan penting berikutnya untuk dipahami adalah rata-rata. Ini juga disebut sebagai rata-rata aritmatika, atau mean aritmetik. Mean aritmatika adalah tipe rata-rata yang paling umum. Untuk menghitung mean aritmetika, jumlah semua bilangan dalam kelompok data dibagi dengan jumlah item dalam kelompok data tersebut. Misalnya, mungkin ada kebutuhan untuk mengetahui usia rata-rata pasien yang menghadiri seminar diabetes di klinik. Sepuluh pasien hadir:

- **Langkah 1: Tambahkan semua usia peserta. $49 + 62 + 68 + 54 + 42 + 31 + 62 + 48 + 38 + 36 = 490$**
- **Langkah 2: Bagilah total dengan jumlah item dalam data. $490/10 = 49$**
Usia rata-rata peserta seminar diabetes adalah 49 tahun.

BAB II

NILAI TENGAH ATAU NILAI PUSAT

A. PENGERTIAN NILAI TENGAH

Nilai tengah (*Central Tendency*) adalah suatu nilai yang dapat bersifat representatif dari sekumpulan data. Pengukuran nilai tengah ini digunakan untuk gejala pusat dari suatu distribusi frekuensi yang dapat mewakili keseluruhan data atau populasi.

Pengukuran nilai tengah adalah ukuran yang digunakan untuk menunjukkan nilai tengah (gejala pusat) dari suatu distribusi frekuensi yang dapat keseluruhan data atau populasi.

Syarat pengukuran nilai tengah dikatakan baik, sebagai berikut:

1. Bersifat representatif terhadap sekumpulan data atau populasi.
2. Mempunyai formulasi dan prosedur atau langkah-langkah yang jelas sehingga mudah dipahami.

B. JENIS NILAI TENGAH DATA

1. Mean

Mean adalah statistik yang paling umum digunakan. Mean diperoleh dengan menambahkan semua nilai dan membagi dengan jumlah pengamatan dalam sampel. Mean dari segi aritmetika adalah jumlah nilai-nilai dibagi dengan jumlah individu (Hadi, 2015: 45). Istilah rata-rata hitung atau *arithmetic mean* dapat juga disederhanakan menjadi rata-rata atau mean saja.

Menurut Setiawan (2013: 107), mean memiliki kelebihan dan kekurangan. Kelebihan mean, antara lain: 1) mudah diingat, dimengerti, dipahami dan dihitung; 2) tingkat perubahan data tidak terlalu mempengaruhi prosedur perhitungan; 3) berdasarkan populasi atau sampel yang ada. Sedangkan kekurangan mean antara lain: 1) dipengaruhi nilai ekstrim besar ataupun nilai ekstrim kecil; 2) kelas terbuka sulit ditentukan mean-nya.

Mean biasanya dipilih seseorang sebagai pengukuran *central tendency*, terutama bilamana distribusi mendekati normal.

2. Median

Median adalah nilai tengah dalam sampel. Untuk menemukan median, data diurutkan dari nilai terkecil sampai terbesar dan nilai tengahnya dipilih sebagai median. Jika sampel memiliki jumlah observasi genap, maka kedua nilai tengahnya dirata-ratakan untuk menentukan mediannya.

Median (Me) adalah nilai suatu data yang tepat berada di tengah-tengah nilai data yang lain. Berarti 50% data lebih besar dari ($50% > Me$) dan 50% nilai data lebih kecil dari median ($50% < Me$).

Menurut Setiawan (2013: 122), median memiliki kelebihan dan kekurangan. Kelebihan median yaitu tidak bergantung banyak sedikitnya data, dan nilai-nilai ekstrim tidak berpengaruh. Sedangkan kekurangan median yaitu tidak dapat digunakan untuk menghitung banyak data yang genap secara pasti.

3. Modus

Modusnya adalah nilai dengan frekuensi kejadian tertinggi. Mungkin tidak ada mode unik untuk beberapa variabel kontinu seperti biaya karena setiap nilai mungkin berbeda dalam sampel.

Modus (M_o) adalah data yang mempunyai frekuensi kemunculan terbanyak dibandingkan frekuensi kemunculan data yang lain atau disebut juga data yang banyak muncul.

Menurut Setiawan (2013: 116), modus memiliki kelebihan dan kekurangan. Kelebihan modus yaitu dapat digunakan untuk memprediksi tingkat kejadian atau peristiwa tanpa menghitung terlebih dulu grafiknya mudah dibaca. Sedangkan kekurangan modus yaitu jumlah data/peristiwa tau kejadian harus relatif banyak jika sedikit penyimpangannya relatif besar dan tidak semua peristiwa atau kejadian mempunyai modus, sehingga menimbulkan kesulitan dalam menganalisis.

C. CARA MENGHITUNG NILAI TENGAH

1. Mean

a. Perhitungan Mean Data Tunggal

Data sampel nilainya diperhitungkan secara individual baik data yang kurang maupun yang lebih dari 30 data, misal ada sejumlah data yaitu $X_1, X_2, X_3, X_4, \dots, X_n$ maka rata-rata dapat dihitung dengan rumus :

$$\text{Mean} = \frac{X_1 + X_2 + X_3 + X_4 + \dots + X_n}{n} \text{ atau } \text{Mean} = \frac{\sum X_i}{n}$$

Keterangan:

X_1 = Nilai data kesatu

X_2 = Nilai data kedua

X_3 = Nilai data ketiga

X_4 = Nilai data keempat

X_n = Nilai data ke-n

n = Banyak sampel (data)

X_i = nilai data kesatu

i = mewakili bilangan 1,2,3,4,..... n

Contoh:

- 1) Data lama rawat pasien RS "X" sebagai berikut: 2 hari, 5 hari, 9 hari, 1 hari, 6 hari
Maka perhitungan mean sebagai berikut :

$$\text{Mean} = \frac{X_1 + X_2 + X_3 + X_4 + \dots + X_n}{n}$$

$$\text{Mean} = \frac{2+5+9+1+6}{5}$$

Mean = 4,6 jadi rata-rata lama rawat pasien (LOS) adalah 4,6 hari.

- 2) Data usia pasien (tahun) seperti dalam master Tabel 4.1 pada bab penyajian data berikut ini:

Tabel 2.1.

Daftar Pasien RS "X" tanggal 17 Agustus 2017

NO.RM	Jenis Pasien (Baru/Lama)	Usia (th)	Jenis Kelamin	Gol Darah	TB (cm)	BB (kg)	TD (mmHg)
21-01-01	Baru	17	L	B	149	66	110/90
21-01-02	Baru	23	P	A	167	56	100/80
21-01-03	Lama	21	P	AB	156	65	110/80
21-01-04	Baru	10	P	O	138	37	100/80
21-01-05	Lama	12	P	O	140	31	100/70
21-01-06	Baru	45	L	O	141	67	130/90
21-01-07	Lama	55	L	B	169	72	110/80
21-01-08	Baru	78	P	O	150	79	100/70
21-01-09	Baru	51	L	O	172	56	110/90
21-01-10	Baru	63	L	O	166	70	120/80
21-01-11	Baru	71	P	B	157	54	100/70
21-01-12	Baru	12	P	A	135	40	100/80

Maka perhitungan mean sebagai berikut :

$$\text{Mean} = \frac{X_1 + X_2 + X_3 + X_4 + \dots + X_n}{n}$$

$$\text{Mean} = \frac{17+23+21+10+12+45+55+78+51+63+71+12}{12}$$

Mean = 458/12 = 38,1 jadi rata-rata usia pasien adalah 38,1 tahun.

b. Perhitungan Mean Data Berkelompok

Data yang diukur relatif banyak perlu dikelompokkan ke dalam tabel distribusi frekuensi. Perhitungan mean, data yang dikelompokkan memerlukan frekuensi maka rumusnya :

$$\text{Mean} = \frac{F_1M_1 + F_2M_2 + \dots + F_nM_n}{N}$$

Dimana :

F_1 = frekuensi kelas interval kesatu

F_2 = frekuensi kelas interval kedua

F_n = frekuensi kelas interval ke-n

M_1 = titik tengah interval kesatu

M_2 = titik tengah interval kedua

M_n = titik tengah interval ke-n

N = banyak data

Rumus diatas dapat disederhanakan menjadi :

$$\text{Mean} = \frac{\sum FiMi}{N}$$

Dimana :

F_i = frekuensi kelas interval ke- i

M_i = titik tengah interval ke- i

i = mewakili bilangan 1,2,3 n

Contoh:

Data berkelompok dapat dilihat pada Tabel 4.6 pada bab sebelumnya seperti di bawah ini:

Tabel 2.2.

Distribusi Usia Pasien RS "X" 17 Agustus 2017

Usia Pasien (Tahun)	Frekuensi
10 – 24	6
25 – 39	0
40 – 54	2
55 – 69	2
70 – 84	2
Total	12

Jawab :

Langkah pertama menentukan titik tengah masing-masing kelas interval. Rumus titik tengah kelas interval (M) :

$$M_1 = \frac{10+24}{2} = 17$$

$$M_2 = \frac{25+39}{2} = 32$$

$$M_3 = \frac{40+54}{2} = 47$$

$$M_4 = \frac{55+69}{2} = 62$$

$$M_5 = \frac{70+84}{2} = 77$$

Usia Pasien (Tahun)	Frekuensi	Titik Tengah (M)	<i>FiMi</i>
10 – 24	6	17	102
25 – 39	0	32	0
40 – 54	2	47	94
55 – 69	2	62	124
70 – 84	2	77	144
Total	12		464

Maka : Mean = $\frac{\sum FiMi}{N}$

$$\text{Mean} = \frac{464}{12}$$

Mean = 38,6 jadi rata-rata usia pasien adalah 38,6 tahun

2. Median

a. Perhitungan Median Data Tunggal

Median merupakan suatu nilai yang berada di tengah-tengah data, setelah data tersebut diurutkan. (Purwanto S, 2012)

Dalam menghitung median, jika jumlah data ganjil maka median adalah nilai tengah dari urutan data namun jika jumlah data genap maka untuk menentukan mediannya diambil 2 data tengah dijumlah, kemudian dibagi 2.

Penghitungan ukuran nilai median data tunggal dengan langkah-langkah berikut ini :

- 1) Menyusun data secara urut (*array*) yaitu data diurutkan dari data terkecil sampai dengan data terbesar.
- 2) Menentukan letak median dengan formulasi $\frac{n+1}{2}$
- 3) Menentukan nilai median secara langsung data yang paling tengah

Contoh sama dengan mean :

- 1) Data lama rawat pasien RS "X" sebagai berikut: 2 hari, 5 hari, 9 hari, 1 hari, 6 hari

Langkah-langkah, yaitu :

- a) Mengurutkan data terkecil ke data terbesar: 1, 2, 5, 6, 9
- b) Menentukan letak median, dimana banyak data atau $n = 5$, yaitu :

$$Lm = \frac{5+1}{2} = 3$$

- c) Nilai median terletak diantara data ke-3
- d) Berarti nilai median adalah 5. Jadi median lama rawat pasien adalah 5 hari.

Median tidak dipengaruhi oleh nilai ekstrim besar maupun ekstrim kecil (*outlier*).

Sebagai contoh data lama rawat di atas, jika pasien dengan lama rawat (LOS) 9 hari digantikan dengan lama rawat (LOS) 20 hari, nilai rata-rata LOS akan berubah menjadi 6,8 hari sementara median LOS tetap 5 hari. Pasien dengan LOS 20 hari dianggap

sebagai outlier (ekstrim besar), sehingga menyebabkan nilai rata-rata LOS menjadi lebih tinggi, tetapi median tidak berubah dengan adanya nilai ekstrim besar tersebut.

- 2) Data usia pasien (tahun) seperti dalam master tabel 4.1 pada bab penyajian data berikut ini:
 - a) Mengurutkan data terkecil ke data terbesar: 10, 12, 12, 17, 21, 23, 45, 51, 55, 63, 71, 78.
 - b) Menentukan letak median, dimana banyak data atau $n = 5$, yaitu :

$$Lm = \frac{12+1}{2} = 6.5$$
 - c) Nilai median terletak diantara data ke-6 dan data ke 7
 - d) Berarti nilai median antara 23 dan 45 yaitu $\frac{23+45}{2} = 34$ Jadi median usia pasien adalah 34 tahun.

c. *Perhitungan Mean Data Berkelompok*

Nilai median untuk data yang dikelompokkan perlu tabel distribusi frekuensi karena data relatif cukup banyak. Rumusnya sebagai berikut:

$$Me = TKB + \frac{N_2 - FKKDB}{FKKDA - FKKDB} Ci$$

Keterangan:

Me = Median

TKB = Tepi Kelas Bawah

N = Banyak data

FKKDA = Frekuensi Kumulatif Kurang Dari Atas

FKKDB = Frekuensi Kumulatif Kurang Dari Bawah

Ci = *Class Interval* = panjang kelas = jarak kelas interval

Contoh:

Tekanan Darah Minimum	80 – 89	90 – 99	90 – 109	110 - 119	120-129	130 -139	Total n
Jumlah Pasien	5	15	13	25	10	7	75

Tentukan berapa tekanan darah minimum paling tengah dari pasien yang periksa ke Puskesmas ?

Jawab :

- 1) Menentukan letak median yaitu $LM = \frac{(n+1)}{2} = \frac{(75+1)}{2} = 38$

- 2) Membuat tabel distribusi frekuensi

Tekanan Darah Minimum	Jumlah Pasien	Tepi Kelas	FKKD
80 – 89	5	79,5	0
90 – 99	15	89,5	5
100 – 109	13	99,5	20
110 – 119	25	109,5	33
120 – 129	10	119,5	58
130 – 139	7	129,5	68
		139,5	75

Nilai median berada di tekanan darah minimum antara 110 s/d 119 sebanyak 25 pasien. Sehingga dapat ditentukan:

$$\text{TKB} = 109,5 \quad \text{N} = 75 \quad \text{FKKDB} = 33 \quad \text{FKKDA} = 58 \quad \text{Ci} = 10$$

3) Menghitung nilai median

$$\text{Me} = \text{TKB} + \frac{\text{N}_2 - \text{FKKDB}}{\text{FKKDA} - \text{FKKDB}} \text{Ci}$$

$$\text{Me} = 109,5 + \frac{75 - 33}{58 - 33} 10 = 109,5 + 1,8 = 111,3$$

3. Modus

a. Perhitungan Modus Data Tunggal

Nilai modus untuk data tunggal ditentukan dengan memilih data yang mempunyai frekuensi kemunculan terbanyak diantara data lainnya dan tidak menggunakan rumus statistik.

Contoh:

- 1) Data lama rawat pasien RS "X" sebagai berikut: 2 hari, 5 hari, 9 hari, 1 hari, 6 hari. Berarti tidak ada nilai modus pada data lama rawat pasien di RS "X".
- 2) Apabila Data lama rawat pasien RS "X" dirubah menjadi: 2 hari, 5 hari, 9 hari, 1 hari, 6 hari, 5 hari. Berarti nilai modus pada data lama rawat pasien di RS "X" adalah 5 hari, karena lama rawat 5 hari paling banyak muncul atau 2 kali muncul sementara yang lainnya masing-masing hanya 1 kali. Nilai modus lama rawat 5 hari disebut unimodal.
- 3) Apabila Data lama rawat pasien RS "X" dirubah menjadi: 2 hari, 5 hari, 9 hari, 1 hari, 6 hari, 5 hari, 6 hari.

Berarti nilai modus pada data lama rawat pasien di RS "X" adalah 5 dan 6 hari, karena lama rawat 5 hari dan 6 hari paling banyak muncul atau masing-masing 2 kali muncul sementara yang lainnya masing-masing hanya 1 kali. Nilai modus lama rawat 5 hari dan 6 hari disebut bimodal.

b. Penghitungan Modus pada Data Berkelompok

Nilai modus untuk data berkelompok dihitung dengan menggunakan tabel distribusi frekuensi. Rumus modus yaitu :

$$Mo = TKB + \frac{d1}{d1+d2} Ci \text{ atau } Mo = TKA - \frac{d1}{d1+d2} Ci$$

Keterangan:

Mo = Modus

TKB = Tepi Kelas Bawah

TKA = Tepi Kelas Atas

d1 = Selisih frekuensi modus berada dengan modus sebelumnya.

d2 = Selisih frekuensi modus berada dengan modus sesudahnya.

Ci = *Class interval* atau panjang kelas interval

Contoh:

Selama tahun 2011 total premi pemegang polis asuransi kesehatan dari 60 pemegang polis kesehatan baru dapat didistribusikan sebagai berikut :

Premi Kesehatan (x Rp 1000)	Pemegang Polis Kesehatan
Rp 10.000 – Rp. 19.999	6
Rp 20.000 – Rp. 29.999	9
Rp 30.000 – Rp. 39.999	10
Rp 40.000 – Rp. 49.999	20
Rp 50.000 – Rp. 59.999	10
Rp 60.000 – Rp. 69.999	5

Tentukan modus premi asuransi kesehatan yang dibayarkan para pemegang polis kesehatan baru tersebut!

Premi Kesehatan (x Rp 1000)	Pemegang Polis Kesehatan	Tepi Kelas
Rp 10.000 – Rp. 19.999	6	9.999,5
Rp 20.000 – Rp. 29.999	9	19.999,5
Rp 30.000 – Rp. 39.999	10	29.999,5
Rp 40.000 – Rp. 49.999	20	39.999,5
Rp 50.000 – Rp. 59.999	10	49.999,5
Rp 60.000 – Rp. 69.999	5	59.999,5
		69.999,5
Jumlah	60	-

Diketahui :

Kelas interval keempat frekuensi terbanyak yaitu 20 memiliki TKB = 39.999,5 dan TKA = 49.999,5. Sedangkan $d1 = 20 - 10 = 10$ dan $d2 = 20 - 10 + 10$, $Ci = 10.000$. nilai modus dapat dicari dengan dua rumus, sebagai berikut :

$$1) Mo = TKB + \frac{d1}{d1+d2} Ci$$

$$Mo = 39.999,5 + \frac{10}{10+10} 10.000$$

$$Mo = 39.999,5 + 5.000 = 44.999,5$$

$$2) Mo = TKA - \frac{d1}{d1+d2} Ci$$

$$Mo = 49.999,5 - \frac{10}{10+10} 10.000$$

$$Mo = 49.999,5 - 5.000 = 44.999,5 \text{ dibulatkan } 45.000, \times 1000.$$

Jadi modus premi asuransi kesehatan baru sebesar Rp. 44.999.500- (Empat puluh empat juta sembilan ratus sembilan puluh sembilan ribu lima ratus rupiah atau dibulatkan menjadi Rp. 45.000 (empat puluh lima juta ribu rupiah).

Data Kategori

Pada data kategori nilai tengah dapat diwakili oleh nilai modus atau proporsi (persentase) dari data tersebut. Contoh data kategori adalah jenis kelamin dan golongan darah pasien rawat jalan RS "X" seperti dalam master tabel berikut ini.

Tabel 2.3.

Daftar Pasien RS "X" tanggal 17 Agustus 2017

NO.RM	Jenis Pasien (Baru/Lama)	Usia (th)	Jenis Kelamin	Gol Darah	TB (cm)	BB (kg)	TD (mmHg)
21-01-01	Baru	17	L	B	149	66	110/90
21-01-02	Baru	23	P	A	167	56	100/80
21-01-03	Lama	21	P	AB	156	65	110/80
21-01-04	Baru	10	P	O	138	37	100/80
21-01-05	Lama	12	P	O	140	31	100/70
21-01-06	Baru	45	L	O	141	67	130/90
21-01-07	Lama	55	L	B	169	72	110/80
21-01-08	Baru	78	P	O	150	79	100/70
21-01-09	Baru	51	L	O	172	56	110/90
21-01-10	Baru	63	L	O	166	70	120/80
21-01-11	Baru	71	P	B	157	54	100/70
21-01-12	Baru	12	P	A	135	40	100/80

Cara menghitung nilai tengah adalah bisa dilihat dalam Tabel 5.3 yaitu dengan menghitung nilai proporsi atau persentase dari data tersebut.

Tabel 2.4.

Distribusi Jenis Kelamin Pasien RS "X" 17 Agustus 2017

Jenis Kelamin	Frekuensi	Persentase (%)
Laki-laki	5	41.7
Perempuan	7	58.3
Total	12	100

Tabel 2.5.

Distribusi Golongan Darah Pasien RS "X" 17 Agustus 2017

Golongan Darah	Frekuensi	Persentase (%)
A	2	16.7
B	3	25.0
AB	1	8.3
O	6	50.0
Total	12	100

Dapat dijelaskan bahwa nilai pusat pada distribusi data jenis kelamin pasien adalah perempuan sebesar 58.3% dan golongan darah adalah golongan darah O sebesar 50.0%.

BAB III UKURAN POSISI

A. KWARTIL

Kwartil merupakan nilai yang membagi frekuensi distribusi data menjadi empat kelompok yang sama besar. Dengan kata lain kwartil merupakan nilai yang membagi tiap-tiap 25% frekuensi dalam distribusi.

Kwartil ada 3, yaitu kwartil pertama (K1) yaitu suatu nilai yang membatasi 25% frekuensi bagian bawah dengan 75% frekuensi bagian atas.

Kwartil kedua (K2) yaitu suatu nilai yang membatasi 50% frekuensi bagian bawah dengan 50% frekuensi bagian atas. K2 akan memiliki nilai yang sama dengan median.

Kwartil ketiga (K3) yaitu suatu nilai yang membatasi 75% frekuensi bagian bawah dengan 25% frekuensi bagian atas.

$$\text{Rumus menghitung kwartil: } K_n = Bb + \left[\frac{n/4N - fkb}{fd} \right] i$$

Keterangan:

K_n = Kwartil ke n

Bb = Batas bawah interval kelas yang mengandung kwartil ke n

fkb = frekuensi kumulatif interval kelas di bawah interval kelas yang mengandung kwartil ke n

fd = frekuensi interval kelas yang mengandung kwartil ke n

i = lebar interval kelas

N = jumlah frekuensi dalam distribusi

Contoh :

Carilah nilai yang membatasi 25 orang yang mempunyai nilai ujian statistik paling tinggi pada tabel berikut ini:

Tabel3.1
Distribusi Frekuensi nilai ujian statistik mahasiswa

No.	Interval kelas	f	Fk
1	90 – 99	2	100
2	80 – 89	6	98
3	70 – 79	10	92
4	60 – 69	14	82
5	50 – 59	20	68
6	40 – 49	18	48
7	30 – 39	16	30

No.	Interval kelas	f	Fk
8	20 – 29	10	14
9	10 – 19	2	4
10	0 – 9	2	2
	Σ	100	

Mencari interval kelas yang mengandung $K_3 \rightarrow \frac{3}{4} \times 100 = 75 \rightarrow$ berada pada frekuensi 82 dengan kelas data antara **60 – 69**.

Dari tabel diketahui :

$$Bb = 59,5$$

$$fkb = 68$$

$$fd = 14$$

$$i = 10$$

$$N = 100$$

Masukkan ke dalam rumus kuartil:

$$\begin{aligned}
 K_3 &= 59,5 + \left[\frac{\frac{3}{4} \times 100 - 68}{fd} \right] i \\
 &= 59,5 + \left[\frac{75 - 68}{14} \right] 10 \\
 &= 59,5 + 5 = 64,5
 \end{aligned}$$

Jadi nilai yang membagi 25 orang dengan nilai ujian tertinggi dan 75 orang dengan nilai ujian yang rendah adalah 64,5.

B. DESIL

Desil merupakan nilai yang membagi frekuensi distribusi data menjadi sepuluh kelompok yang sama besar. Dengan kata lain desil merupakan nilai yang membagi tiap-tiap 10% frekuensi dalam distribusi. Desil ada 9, yaitu D1, D2, D3, D4, D5, D6, D7, D8, dan D9.

Desil pertama (D1) adalah suatu nilai yang membatasi 10% frekuensi bagian bawah dengan 90% frekuensi bagian atas.

Desil keenam (D6) yaitu suatu nilai yang membatasi 60% frekuensi bagian bawah dengan 40% frekuensi bagian atas.

Desil kedelapan (D8) yaitu suatu nilai yang membatasi 80% frekuensi bagian bawah dengan 20% frekuensi bagian atas.

Rumus menghitung desil:

$$D_n = Bb + \left[\frac{n/10N - fkb}{fd} \right] i$$

Keterangan:

D_n = Desil ke n

Bb = Batas bawah interval kelas yang mengandung desil ke n

fk_b = frekuensi kumulatif interval kelas di bawah interval kelas yang mengandung desil ke n

fd = frekuensi interval kelas yang mengandung desil ke n **i** = lebar interval kelas

N = jumlah frekuensi dalam distribusi

Contoh:

Carilah nilai yang membatasi 40 orang yang mempunyai nilai ujian statistik paling rendah pada tabel berikut ini:

Tabel 3.2
Distribusi Frekuensi nilai ujian statistik mahasiswa

No.	Interval kelas	F	Fk
1	90 – 99	2	100
2	80 – 89	6	98
3	70 – 79	10	92
4	60 – 69	14	82
5	50 – 59	20	68
6	40 – 49	18	48
7	30 – 39	16	30
8	20 – 29	10	14
9	10 – 19	2	4
10	0 – 9	2	2
	Σ	100	

Mencari interval kelas yang mengandung $D_4 \rightarrow 4/10 \times 100 = 40 \rightarrow 40 - 49$ Dari tabel diketahui :

Bb = 39,5

fk_b = 30

fd = 18

i = 10

N = 100

Masukkan ke dalam rumus desil :

$$\begin{aligned} D_3 &= 39,5 + \left[\frac{4/10 \times 100 - 30}{40 - 30} \right] 10 \\ &= 39,5 + \left[\frac{10 - 30}{10} \right] 10 \\ &= 39,5 + \left[\frac{18}{10} \right] 10 \\ &= 39,5 + 5,555 = 45,05 \end{aligned}$$

Jadi nilai yang membagi 40 orang dengan nilai ujian tertinggi dan 60 orang dengan nilai ujian yang rendah adalah 45,05.

C. PERSENTIL

Persentil merupakan nilai yang membagi frekuensi distribusi data menjadi seratus kelompok yang sama besar. Dengan kata lain persentil merupakan nilai yang membagi tiap-tiap 1% frekuensi dalam distribusi.

Persentil ada 99, yaitu P1 - P99.

Persentil pertama (P1) adalah suatu nilai yang membatasi 1% frekuensi bagian bawah dengan 99% frekuensi bagian atas.

Persentil ketigapuluhtujuh (P37) yaitu suatu nilai yang membatasi 37% frekuensi bagian bawah dengan 63% frekuensi bagian atas.

Persentil kedelapanpuluh enam (P86) yaitu suatu nilai yang membatasi 86% frekuensi bagian bawah dengan 14% frekuensi bagian atas.

Rumus menghitung persentil:

$$P_n = Bb + \left[\frac{n / 100N - fkb}{fd} \right] i$$

Keterangan:

P_n = Persentil ke n

Bb = Batas bawah interval kelas yang mengandung persentil ke n

fkb = frekuensi kumulatif interval kelas di bawah interval kelas yang mengandung persentil ke n

fd = frekuensi interval kelas yang mengandung persentil ke

i = lebar interval kelas

N = jumlah frekuensi dalam distribusi

Contoh:

Carilah nilai yang membatasi 67 orang yang mempunyai nilai ujian statistik paling rendah pada tabel berikut ini:

Tabel 3.3

Distribusi Frekuensi nilai ujian statistik mahasiswa

No.	Interval kelas	F	Fk
1	90 – 99	2	100
2	80 – 89	6	98
3	70 – 79	10	92
4	60 – 69	14	82
5	50 – 59	20	68
6	40 – 49	18	48
7	30 – 39	16	30
8	20 – 29	10	14
9	10 – 19	2	4
10	0 – 9	2	2
	Σ	100	

Mencari interval kelas yang mengandung P₆₇ → $67/100 \times 100 = 67 \rightarrow 50 - 59$

Dari tabel diketahui :

$$Bb = 49,5$$

$$fkb = 48$$

$$fd = 20$$

$$i = 10$$

$$N = 100$$

Masukkan ke dalam rumus persentil:

$$\begin{aligned} P_{67} &= 49,5 + \left[\frac{67/100 \times 100 - 48}{20} \right] 10 \\ &= 49,5 + \left[\frac{67 - 48}{20} \right] 10 \\ &= 49,5 + 9,5 = 59 \end{aligned}$$

Jadi nilai yang membagi 67 orang dengan nilai ujian tertinggi dan 33 orang dengan nilai ujian yang rendah adalah 59.

JENJANG PERSENTIL (*Percentile rank*)

Jenjang persentil (JP) adalah bilangan yang menunjukkan frekuensi dalam persen (%) yang terdapat pada suatu bilangan tertentu dan dibawahnya.

Rumus jenjang persentil:

$$JP = \left\{ \left(\frac{X - Bb}{i} \right) fd + fkb \right\} \frac{100}{N}$$

Keterangan:

X = Suatu nilai yang diketahui

Bb = Batas bawah interval kelas yang mengandung X

fkb = frekuensi kumulatif interval kelas di bawah interval kelas yang mengandung X

fd = frekuensi interval kelas yang mengandung X

i = lebar interval kelas

N = jumlah frekuensi dalam distribusi

Tabel 3.4
Distribusi Frekuensi Nilai Ujian Statistik Mahasiswa

No.	Interval kelas	F	Fk
1	90 – 99	2	100
2	80 – 89	6	98
3	70 – 79	10	92
4	60 – 69	14	82
5	50 – 59	20	68
6	40 – 49	18	48
7	30 – 39	16	30
8	20 – 29	10	14
9	10 – 19	2	4
10	0 – 9	2	2
	Σ	100	

Contoh:

Dari tabel di atas, berapa orangkah yang mendapat nilai ujian 73 ke bawah?

Jawab:

Mencari interval kelas yang mengandung nilai 73 → **70 - 79**

Dari table diketahui:

$$X = 73$$

$$Bb = 69,5$$

$$fkb = 82$$

$$fd = 10$$

$$i = 10$$

$$N = 100$$

Masukkan ke dalam rumus,

$$JP = \left\{ \left(\frac{73 - 69,5}{10} \right) 10 + 82 \right\} \frac{100}{100}$$

$$= \left\{ \left(\frac{3,5 - 869,5}{10} \right) 82 \right\} 1 = 85,5 \%$$

Dari perhitungan diperoleh JP adalah 85,5%, dengan demikian jumlah orang yang mendapat nilai 73 ke bawah adalah $\frac{85,5}{100} \times 100 = 85,5 \sim \mathbf{86 \text{ Orang}}$

BAB IV

UKURAN VARIASI (DISPERSI)

A. PENGERTIAN UKURAN DISPERSI

Menurut Hasan (2011: 101) ukuran dispersi atau ukuran variasi atau ukuran penyimpangan adalah ukuran yang menyatakan seberapa jauh penyimpangan nilai-nilai data dari nilai-nilai pusatnya atau ukuran yang menyatakan seberapa banyak nilai-nilai data yang berbeda dengan nilai-nilai pusatnya.

Ukuran dispersi pada dasarnya adalah pelengkap dari ukuran nilai pusat dalam menggambarkan sekumpulan data. Jadi, dengan adanya ukuran dispersi maka penggambaran sekumpulan data akan menjadi lebih jelas dan tepat.

Macam-macam ukuran dispersi adalah jangkauan, rerata deviasi, variansi, dan deviasi baku.

B. JENIS UKURAN VARIASI (DISPERSI)

1. Jangkauan (range , R)

Menurut Hasan (2011: 101), jangkauan atau ukuran jarak adalah selisih nilai terbesar data dengan nilai terkecil data. Menurut Riduwan dan Akdon (2013 : 39) range (rentangan) ialah data tertinggi dikurangi data terendah. Sedangkan menurut Siregar (2010: 40), rentang atau daerah jangkauan adalah selisih antara nilai terbesar dengan nilai terkecil dari serangkaian data. Menurut Usman dan Akbar (2008 : 95), rentang ialah ukuran variasi yang paling sederhana yang dihitung dari datum terbesar dikurang datum data terkecil.

Jadi jangkauan adalah selisih antara nilai tertinggi dengan nilai terendah dari serangkaian data. Berikut adalah rumus jangkauan (range) untuk data tunggal dan data kelompok menurut Hasan (2011: 101) adalah sebagai berikut :

a. Data tunggal

Bila ada sekumpulan data tunggal $x_1, x_2, x_3, \dots, x_n$ maka jangkauannya adalah
Jangkauan = $x_n - x_1$

Contoh soal :

Tentukan jangkauan data : 1, 4, 7, 8, 9, 11 Penyelesaian :

$$X_6 = 11 \text{ dan } X_1 = 1$$

$$\text{Jangkauan} = X_6 - X_1 = 11 - 1 = 10$$

b. Data kelompok

Untuk data berkelompok, jangkauan dapat ditentukan dengan dua cara yaitu menggunakan titik atau nilai tengah dan menggunakan tepi kelas.

- 1) Jangkauan adalah selisih titik tengah kelas tertinggi dengan titik tengah kelas terendah.
- 2) Jangkauan adalah selisih tepi atas kelas tertinggi dengan tepi bawah kelas terendah.

Contoh soal :

Tentukan jangkauan dari distribusi frekuensi berikut !

Tabel 4.1
Pengukuran Tinggi Badan 50 Mahasiswa

Tinggi Badan (cm)	Frekuensi
140 – 144	2
145 – 149	4
150 – 154	10
155 – 159	14
160 – 164	12
165 – 169	5
170 – 174	3
Jumlah	80

Penyelesaian :

Titik tengah kelas terendah = 142

Titik tengah kelas tertinggi = 172

Tepi bawah kelas terendah = 139,5

Tepi atas kelas tertinggi = 174,5

1) Jangkauan = 172 – 142 = 30

2) Jangkauan = 174,5 – 139,5 = 35

2. Rerata Deviasi (Simpangan Rata-rata)

Menurut Hasan (2011: 105) deviasi rata-rata adalah nilai rata-rata hitung dari harga mutlak simpangan-simpangannya. Cara mencari deviasi rata-rata, dibedakan antara data tunggal dan data kelompok.

a. Deviasi rata-rata data tunggal

Untuk data tunggal, deviasi rata-ratanya dapat dihitung dengan menggunakan rumus:

$$DR = \frac{1}{n} \sum |X - \bar{X}| = \frac{\sum |Xi - \bar{X}|}{n}$$

Contoh soal :

Tentukan deviasi rata-rata dari 2, 3, 6, 8, 11!

$$\text{Rata-rata hitung} = \bar{X} = \frac{2+3+6+8+11}{5} = 6$$

$$\sum |Xi - \bar{X}| = |2 - 6| + |3 - 6| + |6 - 6| + |8 - 6| + |11 - 6| = 14$$

$$\begin{aligned} DR &= \frac{\sum |Xi - \bar{X}|}{n} \\ &= \frac{14}{5} = 2,8 \end{aligned}$$

b. *Deviasi rata –rata untuk data kelompok*

$$DR = \frac{1}{n} \sum |X - \bar{X}| = \frac{\sum |X_i - \bar{X}|}{n}$$

Contoh soal :

Tentukan deviasi rata-rata dari distribusi frekuensi pada Tabel 1 Pengukuran Tinggi Badan 50 Mahasiswa !

Penyelesaian :

Tinggi Badan (cm)	X	F	xifi	 X - X̄ 	f X - X̄
140-144	142	2	284	15,7	31,4
145-149	147	4	588	10,7	42,8
150-154	152	10	1520	5,7	57
155-159	157	14	2198	0,7	9,8
160-164	162	12	1944	4,3	51,6
165-169	167	5	835	9,3	46,5
170-174	172	3	516	14,3	42,9
Jumlah	-	50	7885	-	282

$$\bar{x} = \frac{\sum (xi fi)}{\sum fi} = \frac{7885}{50} = 157,7$$

$$DR = \frac{\sum f|X - \bar{x}|}{n} = \frac{282}{50} = 5,64$$

3. Variance (Varians)

Menurut Riduwan dan Akdon (2013: 43), variance (varians) adalah kuadrat dari simpangan baku. Fungsinya untuk mengetahui tingkat penyebaran atau variasi data. Sedangkan menurut Hasan (2011: 107), varians adalah nilai tengah kuadrat simpangan dari nilai tengah atau simpangan rata-rata kuadrat. Untuk sampel, variansnya (varians sampel) disimbolkan dengan s^2 . Untuk populasi, variansnya (varians populasi) disimbolkan dengan σ^2 (baca: sigma).

a. *Varians data tunggal*

Untuk seperangkat data $x_1, x_2, x_3, \dots, x_n$ (data tunggal), variansnya dapat ditentukan dengan dua metode, yaitu metode biasa dan metode angka kasar.

1) Metode biasa

a) Untuk sampel besar ($n > 30$) :

$$S^2 = \frac{\sum (X - \bar{X})^2}{n}$$

b) Untuk sampel kecil ($n \leq 30$):

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

2) Metode angka kasar

a) Untuk sampel besar ($n > 30$) :

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

b) Untuk sampel kecil ($n \leq 30$) :

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

Contoh soal:

Tentukan varians dari data 2, 3, 6, 8, 11!

Penyelesaian:

$n = 5$

$$\bar{X} = \frac{2+3+6+8+11}{5} = 6$$

X	$X - \bar{X}$	$(X - \bar{X})^2$	X^2
2	-4	16	4
3	-3	9	9
6	0	0	36
8	2	4	64
11	5	25	121
30		54	234

$$s^2 = \frac{\sum (X - \bar{X})^2}{n-1} \quad \left| \quad s^2 = \frac{\sum X^2}{n-1} - \left(\frac{\sum X}{n-1} \right)^2 \right.$$

$$= \frac{54}{5-1} \quad \left| \quad = \frac{234}{5-1} - \frac{(30)^2}{5(5-1)} \right.$$

$$= 13,5 \quad \left| \quad = 13,5 \right.$$

b. Varians data berkelompok

Untuk data berkelompok (distribusi frekuensi), variansnya dapat ditentukan menggunakan tiga metode, yaitu metode biasa, metode angka kasar, dan metode coding.

1) Metode biasa

a) Untuk sampel besar ($n > 30$)

$$s^2 = \frac{\sum f(X - \bar{X})^2}{n}$$

b) Untuk sampel kecil ($n \leq 30$)

$$S^2 = \frac{\sum (X - \bar{x})^2}{n-1}$$

2) Metode angka kasar

a) Untuk sampel besar ($n > 30$):

$$S^2 = C^2 \cdot \frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n} \right)^2$$

b) Untuk sampel kecil ($n \leq 30$):

$$S^2 = \frac{\sum fX^2}{n-1} - \frac{(\sum fX)^2}{n(n-1)}$$

3) Metode coding

a) Untuk sampel besar ($n > 30$):

$$S^2 = C^2 \cdot \frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n} \right)^2$$

b) Untuk sampel kecil ($n \leq 30$):

$$S^2 = C^2 \cdot \frac{\sum fu^2}{n-1} - \frac{(\sum fu)^2}{n(n-1)}$$

Keterangan :

C = panjang interval kelas

$$u = \frac{d}{C} = \frac{X - M}{C}$$

M = rata – rata hitung sementara

Contoh soal :

Tentukan varian dari distribusi frekuensi berikut!

Tabel 4. 2
Pengukuran Diameter Pipa

Diameter (mm)	Frekuensi
65-67	2
68-70	5
71-73	13
74-76	14
77-79	4
80-82	2
Jumlah	40

Penyelesaian :

1) Dengan metode biasa

Diameter	X	F	(xifi)	X - \bar{x}	(x - \bar{x}) ²	f(x - \bar{x}) ²
65-67	66	2	132	-7,425	55,131	110,262
68-70	69	5	345	-4,425	19,581	97,905
71-73	72	13	936	-1,425	2,031	26,403
74-76	75	14	1050	1,575	2,481	34,734
77-79	78	4	312	4,575	20,931	83,724
80-82	81	2	162	7,575	57,381	114,762
Jumlah	-	40	2937	-	-	467,790

$$\bar{x} = \frac{\sum(Xi - fi)}{\sum fi} = \frac{2.937}{40} = 73,425$$

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n} = \frac{467,790}{40} = 11,694$$

2) Dengan metode angka kasar

Diameter	X	F	X ²	fX	fX ²
65-67	66	2	4.356	132	8.712
68-70	69	5	4.761	345	23.805
71-73	72	13	5.184	936	67.392
74-76	75	14	5.625	1.050	78.750
77-79	78	4	6.084	312	24.336
80-82	81	2	6.561	162	13.122
Jumlah	-	40	-	2.937	216.117

$$s^2 = \frac{\sum fX^2}{n} - \left(\frac{\sum fX}{n} \right)^2 = \frac{216.117}{40} - \left(\frac{2.937}{40} \right)^2 = 5402,925 - 5391,231 = 11,694$$

3) Dengan metode coding

Diameter	X	f	U	u ²	Fu	fu ²
65-67	66	2	-3	9	-6	18
68-70	69	5	-2	4	-10	20
71-73	72	13	-1	1	-13	13
74-76	75	14	0	0	0	0
77-79	78	4	1	1	4	4
80-82	81	2	2	4	4	8
Jumlah	-	40	-	-	-21	63

$$\begin{aligned}
s^2 &= c^2 \left(\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n} \right)^2 \right) \\
&= 3^2 \left(\frac{63}{40} - \left(\frac{-21}{40} \right)^2 \right) \\
&= 9 (1,575 - 0,276) \\
&= 11,694
\end{aligned}$$

4. Simpangan Baku (Standar Deviasi)

Menurut Riduwan dan Akdon (2013 : 40), *standard deviation* (simpangan baku) ialah suatu nilai yang menunjukkan tingkat (derajat) variasi kelompok atau ukuran standar penyimpangan dari reratanya. Sedangkan menurut Hasan (2011 : 112) Simpangan baku adalah akar dari tengah kuadrat simpangan dari nilai tengah atau akar simpangan rata-rata kuadrat. Untuk sampel, simpangan bakunya (simpangan baku sampel) disimbolkan dengan s . Untuk populasi, simpangan bakunya (simpangan baku populasi) disimbolkan σ . Untuk menentukan nilai simpangan baku, caranya ialah dengan menarik akar dari varians.

Jadi, $s = \sqrt{\text{variens}}$

Cara mencari simpangan baku, dibedakan antara data tunggal dan berkelompok.

a. Simpangan baku data tunggal

Untuk seperangkat data $x_1, x_2, x_3, \dots, x_n$ (data tunggal) simpangan bakunya dapat ditentukan dengan dua metode, yaitu metode biasa dan metode angka kasar.

1) Metode biasa

a) Untuk sampel besar ($n > 30$) :

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

b) Untuk sampel kecil ($n \leq 30$) :

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

2) Metode angka kasar

a) Untuk sampel besar ($n > 30$) :

$$s = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum X}{n} \right)^2}$$

b) Untuk sampel kecil ($n \leq 30$) :

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

Contoh soal:

- 1) Tentukan simpangan baku (standar deviasi) dari data 2, 3, 6, 8, 11! Penyelesaian:

Dari perhitungan diperoleh varians (s^2) = 13,5

Dengan demikian simpangan bakunya adalah

$$s = \sqrt{\text{variens}} = \sqrt{13,5} = 3,67$$

- 2) Berikut ini adalah sampel nilai mid test statistik 1 dari sekelompok mahasiswa di sebuah universitas.

30, 35, 42, 50, 58, 66, 74, 82, 90, 98

Tentukan simpangan baku dari data di atas!

Penyelesaian :

$n = 10$

X	$x - \bar{x}$	$(x - \bar{x})^2$	x^2
30	-32,5	1.056,25	900
35	-27,5	756,25	1.225
42	-20,5	420,25	1.764
50	-12,5	156,25	2.500
58	-4,5	20,25	3.364
66	3,5	12,25	4.356
74	11,5	132,25	5.476
82	19,5	380,25	6.724
90	27,5	756,25	8.100
98	35,5	1.260,25	9.604
625		4.950,5	44.013

$$\begin{aligned}\bar{x} &= \frac{30 + 35 + 42 + 50 + 58 + 66 + 74 + 82 + 90 + 98}{10} \\ &= \frac{625}{10} \\ &= 62,5\end{aligned}$$

- 1) Dengan metode biasa

$$\begin{aligned}s &= \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{4.950,5}{10 - 1}} \\ &= \sqrt{550,056} \\ &= 23,45\end{aligned}$$

2) Dengan metode angka kasar

$$s = \sqrt{\frac{\sum x^2}{n-1} - \frac{(\sum X)^2}{n(n-1)}}$$
$$s = \sqrt{\frac{44,013}{10-1} - \frac{(625)^2}{10(10-1)}}$$
$$= \sqrt{4.890,33 - 4.340,28}$$
$$= 23,45$$

b. *Simpangan baku data berkelompok*

Untuk data berkelompok (distribusi frekuensi), simpangan bakunya dapat ditentukan dengan tiga metode, yaitu metode biasa, metode angka kasar, dan metode coding.

1) Metode biasa

a) Untuk sampel besar ($n > 30$)

$$s = \sqrt{\frac{\sum f(x - \bar{x})^2}{n}}$$

b) Untuk sampel kecil ($n \leq 30$)

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

2) Metode angka kasar

a) Untuk sampel besar ($n > 30$):

$$s = \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fX}{n}\right)^2}$$

b) Untuk sampel kecil ($n \leq 30$):

$$s = \sqrt{\frac{\sum fx^2}{n-1} - \left(\frac{\sum fX}{n-1}\right)^2}$$

3) Metode coding

a) Untuk sampel besar ($n > 30$):

$$s = C \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum u}{n}\right)^2}$$

b) Untuk sampel kecil ($n \leq 30$):

$$s = C \sqrt{\frac{\sum fu^2}{n-1} - \left(\frac{\sum u}{n-1}\right)^2}$$

Keterangan:

C = panjang interval kelas

$$u = \frac{d - X - M}{c}$$

M = rata – rata hitung sementara

Contoh soal :

- 1) Tentukan simpangan baku dari distribusi frekuensi pada contoh Tabel 2! Penyelesaian: Dari perhitungan didapatkan varians (s^2) = 11,694. Dengan demikian simpangan bakunya adalah $s = \sqrt{\text{varians}} = \sqrt{11,694} = 3,42$
- 2) Tentukan simpangan baku dari distribusi frekuensi berikut (gunakan ketiga rumus)!

Tabel 6. 3
Berat Badan 100 Mahasiswa Universitas “B”

Berat Badan (kg)	Frekuensi (f)
40-44	8
45-49	12
50-54	19
55-59	31
60-64	20
65-69	6
70-74	4
Jumlah	100

Penyelesaian:

- 1) Dengan metode biasa

Berat Badan	X	f	fX	$x - \bar{x}$	$(x - \bar{x})^2$	$f \cdot (x - \bar{x})^2$
40-44	42	8	336	-13,85	191,8225	1.534,58
45-49	47	12	564	-8,85	78,3225	939,87
50-54	52	19	988	-3,85	14,8225	281,63
55-59	57	31	1.767	1,15	1,3225	40,99
60-64	62	20	1.240	6,15	37,8225	756,45
65-69	67	6	402	11,15	124,3225	745,94
70-74	72	4	288	16,15	260,8225	1.043,29

$$\bar{x} = \frac{\sum fX}{\sum f}$$

$$x = \frac{5.585}{100} = 55,85$$

$$s = \sqrt{\frac{\sum f(x-x)^2}{n}}$$

$$= \sqrt{\frac{5.432,75}{100}}$$

$$= 7,31$$

2) Dengan metode angka kasar

Berat Badan	<i>f</i>	<i>X</i>	<i>X</i> ²	<i>fX</i>	<i>fX</i> ²
40-44	8	42	1.764	336	14.112
45-49	12	47	2.209	564	26.508
50-54	19	52	2.704	988	51.376
55-59	31	57	3.249	1.767	100.719
60-64	20	62	3.844	1.240	76.880
65-69	6	67	4.489	402	26.934
70-74	4	72	5.184	288	20.736
Jumlah	100			5.585	317.265

$$s = \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fX}{n}\right)^2}$$

$$= \sqrt{\frac{317.265}{100} - \left(\frac{5.585}{100}\right)^2}$$

$$= 7.31$$

3) Dengan metode coding

Berat Badan	<i>X</i>	<i>f</i>	<i>u</i>	<i>u</i> ²	<i>fu</i>	<i>Fu</i> ²
40-44	42	8	-3	9	-24	72
45-49	47	12	-2	4	-24	48
50-54	52	19	-1	1	-19	19
55-59	57	31	0	0	0	0
60-64	62	20	1	1	20	20
65-69	67	6	2	4	12	24
70-74	72	4	3	9	12	36
Jumlah		100			-23	219

$$c = 5$$

$$\begin{aligned} s &= C \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2} \\ &= \sqrt{\frac{317.265}{100} - \left(\frac{5.585}{100}\right)^2} \\ &= 7,31 \end{aligned}$$

DAFTAR PUSTAKA

Akdon dan Riduwan .2013. *Rumus dan Data dalam Analisis Statistika*. Bandung: Alfabeta.

Hasan, M. Iqbal. 2011. Pokok – Pokok Materi Statistika 1 (Statistik Deskriptif). Jakarta: PT Bumi Aksara

Sugiyono. 2014. *Statistika untuk Penelitian*. Bandung: Alfabeta

Supranto, Johanes. *Statistik : Teori dan Aplikasi*. Edisi 8 Jilid 1. Jakarta. Erlangga. 2016.

Victorianus, Aries Siswanto. *Strategi dan Langkah-langkah Penelitian*. Yogyakarta. Graha Ilmu. 2012