

BAHAN KULIAH
STATISTIKA DASAR



DOSEN PENGAMPU:

VERNITA SARI, S.Pd.,M.Pd.

(digunakan dalam lingkungan sendiri)

PROGRAM STUDI PENDIDIKAN MATEMATIKA
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS MUHAMMADIYAH PAREPARE

2025

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalamualaikum Warahmatullahi Wabarakatuh

Puji syukur kehadiran Allah SWT. yang telah memberikan rahmat serta karunia-Nya sehingga penulis berhasil menyelesaikan bahan kuliah ini dengan lancar. Bahab kuliah ini disusun guna membantu mahasiswa memenuhi kebutuhan dalam memperoleh bahan kuliah khususnya Statistika Dasar.

Bahan kuliah ini merupakan kumpulan materi yang diambil dari beberapa referensi. Penyusun menyadari bahwa baik isi maupun cara penyusunan bahan kuliah ini masih jauh dari sempurna, juga kemungkinan kesalahan ketik ketika tidak dapat dihindarkan. Karena itu, segala saran, dan kritik yang membangun sangat penyusun harapkan. Demikian, semoga bahan kuliah ini dapat berguna dan dimanfaatkan dengan sebaik-baiknya.

Wassalamualaikum Warahmatullahi Wabarakatuh

Parepare, 5 Januari 2025

Penyusun

DAFTAR ISI

KATA PENGANTAR.....	ii
DAFTAR ISI.....	iii
BAB I POPULASI, SAMPEL, DATA DAN PEUBAH	
A. Pengertian Populasi, Sampel, Data dan Peubah.....	1
B. Jenis-jenis Populasi, Sampel, Data dan Peubah	2
C. Teknik Pengambilan Sampel.....	6
D. Teknik Pengumpulan Data.....	10
BAB II STATISTIK DESKRIPTIF DAN STATISTIK INFERENSIAL	
A. Pengertian Statistik Deskriptif dan Statistik Inferensial.....	16
B. Perbedaan Statistik Deskriptif dan Statistik Inferensial.....	17
C. Jenis-jenis Statistik Dekriptif dann Statistik Inferensial.....	18
D. Implikasi Hasil Data.....	21
E. Contoh Statistik Deskriptif dan Statistik Inferensial.....	22
BAB III PENYIAPAN DAN PENYAJIAN DATA	
A. Penyajian dan Penyiapan Data.....	24
B. Jenis-jenis Data.....	24
C. Pembulatan Angka.....	28
BAB IV PEMERIKSAAN DAN PENYAJIAN DATA	
A. Pemeriksaan Data.....	30
B. Penyajian Data.....	31
BAB V DISTRIBUSI FREKUENSI	
A. Pengertian Distribusi Frekuensi.....	40
B. Jenis-jenis Distribusi Frekuensi.....	40
C. Langkah-langkah Distribusi Frekuensi.....	41
BAB VI UKURAN PEMUSATAN DATA	
A. Pengertian Mean.....	45
B. Pengertian Median.....	46
C. Pengertian Modus.....	47
D. Contoh Soal Data Tunggal.....	47
E. Contoh Soal Data Berkelompok.....	48

BAB VII UKURAN LOKASI

A. Pengertian Ukuran Lokasi.....	50
B. Ukuran Letak Data.....	50
C. Macam-macam Ukuran Lokasi.....	57
D. Contoh Soal.....	59

BAB VIII UKURAN PENYEBARAN DATA

A. Pengertian Ukuran Penyebaran Data.....	64
B. Macam-macam Ukuran Penyebaran Data Tunggal.....	64
C. Contoh Soal Penyebaran Data Tunggal.....	66
D. Jenis-jenis ukuran Penyebaran Data Berkelompok.....	69

BAB IX UKURAN KEMIRINGAN DAN KERUNGCIINGAN DATA

A. Ukuran Kemiringan.....	76
B. Ukuran Kerungcingan.....	79

BAB X DISTRIBUSI NOMAL

A. Pengertian Distribusi Normal.....	81
B. Ciri-ciri Distribusi Normal.....	81
C. Parameter dari Distribusi Normal.....	82
D. Tabel Z dan Rumus Distribusi Normal.....	83
E. Contoh Soal Distribusi Normal.....	84

BAB XI UJI HIPOTESIS

A. Pengertian Hipotesis.....	86
B. Kegunaan Hipotesis.....	86
C. Kesalahan Pengujian Hipotesis.....	88
D. Langkah-langkah Pengujian Hipotesis.....	89
E. Menentukan Nilai Kritis.....	91
F. Uji Hipotesis Tentang Rerata.....	93

BAB I

POPULASI, SAMPEL, DATA DAN PEUBAH

A. Pengertian Populasi, Sampel, Data, dan Peubah

1. Populasi

Populasi berasal dari kata bahasa inggris population, yang berarti jumlah penduduk. Populasi adalah generalisasi wilayah yang terdiri atas obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulan (Sugiyono, 2015:80). Populasi juga dapat diartikan sebagai jumlah keseluruhan dari satuan-satuan atau individu-individu yang karakteristiknya hendak diteliti.

Jadi, populasi bukan hanya orang, tetapi juga obyek dan benda-benda alam yang lain. Populasi juga bukan hanya sekedar jumlah yang ada, obyek atau subyek yang dipelajari, tetapi meliputi seluruh karakteristik atau sifat yang dimiliki oleh subjek dan obyek yang diteliti itu.

2. Sampel

Sampel menurut Notoatmodjo (2010) sampel adalah objek yang diteliti dan dianggap mewakili seluruh populasi.

Jadi dapat disimpulkan bahwa sampel itu adalah bagian dari populasi yang memiliki sifat-sifat yang sama dari objek, yang merupakan sumber data. Yang mana secara sederhana sampel dapat dikatakan sebagai bagian dari populasi yang terpilih dan mewakili populasi tersebut.

3. Data

Data berasal dari kata “Datum” yang berarti fakta atau bagian dari fakta yang mengandung arti yang dihubungkan dengan kenyataan yang dapat digambarkan dengan simbol, angka, huruf dan sebagainya.

Data menurut Drs. Jhon J. Longkutoy mengatakan bahwa “Data adalah suatu istilah majemuk dari fakta yang mengandung arti yang dihubungkan dengan kenyataan, simbol, gambar, angka, huruf yang menunjukkan suatu ide, objek, kondisi atau situasi dan lainnya”.

Data disini didapatkan melalui sebuah proses pencarian dan juga pengamatan yang tepat berdasarkan sumber-sumber tertentu.

4. Peubah

Variabel atau peubah adalah sesuatu yang nilainya dapat berubah-ubah dalam suatu hal tertentu. Contoh: IPK, berat badan.

Pengertian variabel menurut Kerlinger adalah konsep yang memiliki macam-macam nilai, dan variabel adalah konsep yang sudah diubah.

Pengertian variabel menurut Moh. Nazir adalah konsep yang mempunyai bermacam-macam nilai.

Jadi, dapat disimpulkan bahwa variabel atau peubah ialah suatu besaran yang dapat diubah atau berubah sehingga bisa mempengaruhi peristiwa atau hasil penelitian.

B. Jenis-jenis Populasi, Sampel, Data, dan Peubah

1. Populasi

Menurut Hendryadi (2019:162-163) Terdapat dua jenis Populasi yaitu Populasi Terbatas dan Populasi Tak Terbatas.

- a. Populasi Terbatas (Finite Population) adalah populasi yang dapat dihitung jumlahnya. Namun, terkadang populasi terbatas sangat besar, sehingga dapat diperlakukan sebagai populasi tak terbatas untuk kesimpulan statistik (generalisasi). Contoh populasi jumlah penduduk Kota Bandung sebesar berapa jiwa.
- b. Populasi Tak Terbatas adalah populasi yang tidak memungkinkan peneliti menghitung jumlah populasi secara keseluruhan. Contoh: meneliti tentang beberapa liter pasang surut air pada bulan purnama.

2. Sampel

Berikut beberapa jenis-jenis sampel yang dapat digunakan oleh peneliti, yaitu:

- a. Sampel acak/random sampling adalah menarik sampel dari populasi secara acak.
- b. Sampel bertingkat adalah sampel yang diambil dengan cara membagi populasi atas tingkatannya. Contoh: berdasarkan umur, pendidikan, kelas ekonomi dll.

- c. Sampel sistematis adalah sampel yang diambil dengan cara membuat daftar anggota populasi secara berurutan. Kemudian ditentukan 1 nomor sebagai patokan.
- d. Sampel cluster adalah sampel yang pengambilannya dilakukan dengan cara membagi populasi dalam kelompok dengan anggota yang tidak sama.
- e. Sampel kuota adalah sampel yang penarikannya dilakukan dengan cara menentukan strata berdasarkan sifat yang paling pengaruh dan yang paling dominan terhadap variabel yang diteliti.
- f. Sampel bertujuan/purposive sampling adalah pengambilan sampel dengan tujuan tertentu. Contoh: untuk menghemat biaya, keterbatasan waktu dan tenaga dll.

3. Data

Semua data yang ada pada hakikatnya merupakan cerminan suatu variabel yang diukur menurut klasifikasinya. Dengan demikian, data dapat diklasifikasikan berdasarkan berikut :

- a. Data Menurut Jenisnya
 - Data Kualitatif yaitu seperti jenis kelamin, pendidikan, warna, suku dan lain sebagainya.
 - Data Kuantitatif yaitu data yang berbentuk angka seperti jumlah penjualan, jarak tinggi badan dan lain sebagainya.
- b. Data Menurut Sifatnya
 - Data Diskrit yaitu data dalam bentuk bilangan bulat. Nilai data yang diperoleh berdasarkan jumlah yang dihitung banyaknya.
 - Data Kontinu yaitu data yang mempunyai nilai terletak dalam sebuah interval, misalnya pendapatan penduduk di kota Jakarta rata-rata empat juta per bulan.
- c. Data Menurut Sumbernya
 - Data Internal yaitu data yang menggambarkan keadaan/aktivitas didalam suatu organisasi. Misalnya data keuangan dan data penjualan.
 - Data Eksternal yaitu data yang menggambarkan keadaan/aktivitas diluar suatu organisasi. Misalnya kepuasan konsumen.
- d. Data Menurut Cara Memperolehnya

- Data Primer merupakan sumber data yang dikumpulkan dan diolah sendiri oleh suatu organisasi atau perorangan langsung dari objeknya.
- Data Sekunder merupakan data penelitian yang diperoleh secara tidak langsung atau data yang diperoleh dalam bentuk yang sudah jadi, sudah dikumpulkan dan diolah oleh pihak lain.

4. Peubah

Berikut jenis-jenis peubah atau variabel berdasarkan sifatnya, yaitu:

- a. Peubah/Variabel Bebas atau Peubah Tetap adalah sejumlah gejala atau faktor atau unsur yang menentukan atau mempengaruhi ada atau munculnya gejala atau respons penelitian. Variabel ini dapat diubah atau dibuat secara bebas oleh peneliti. Peubah bebas juga dapat diartikan sebagai peubah yang mempengaruhi peubah lain.
- b. Peubah Tak-bebas atau Peubah Terikat adalah respons suatu penelitian atau percobaan, yaitu sejumlah gejala atau respons yang muncul karena adanya peubah bebas. Misalnya perbedaan berat badan ayam Broiler akibat diberikan jenis pakan yang berbeda. Jadi : Peubah bebasnya Jenis Pakan dan Peubah terikatnya adalah berat badan.
- c. Peubah Kontrol (Controle Variable) adalah sejumlah gejala atau faktor atau unsur yang dengan sengaja dikendalikan, atau disamakan agar tidak mengganggu atau mempengaruhi peubah bebas atau peubah terikat. Misalnya pada percobaan ayam Broiler dengan jenis pakan yang berbeda, maka jenis kelaminnya dan kandangnya harus sama. Jadi jenis kelamin dan kandang ayam disebut peubah Kontrol. Peubah ini selama penelitian dipertahankan tetap atau tidak berubah.
- d. Peubah Sampingan (Intervining Variable) adalah sejumlah gejala yang tidak dapat dikontrol, akan tetapi dapat diperhitungkan pengaruhnya terhadap peubah terikat atau respons hasil penelitian. Oleh karena peubah ini berpengaruh terhadap peubah bebas, maka akan menyebabkan peubah terikat yang muncul tidak murni akibat peubah bebas, sehingga perlu diketahui seberapa besar pengaruh peubah ini. Salah satu cara untuk memperhitungkan pengaruhnya adalah dengan melakukan pembelokan atau pengelompokan.

Misalnya : bila kita ingin meneliti semua jenis kelamin ayam broiler kita harus mengelompokkan jantan dan betina, jadi Jenis kelamin bukan lagi merupakan peubah Kontrol melainkan sudah dijadikan peubah sampingan.

- e. Peubah Galat atau Peubah Ektrane (Extraneous Variable) adalah : sejumlah gejala yang tidak dapat dikontrol dan tidak dapat pula diperhitungkan pengaruhnya ataupun dielemisir pengaruhnya terhadap peubah bebas dan atau peubah terikat, peubah ini mungkin bersumber dari kondisi sampel dan mungkin pula berada diluar sampel. Peubah ini akan muncul pada saat penelitian berlangsung, peubah ini akan mempengaruhi ketelitian penelitian. Adanya peubah ini dapat dilihat pada besarnya kuadrat tengah galat, makin besar kuadrat tengah galat berarti peubah ini makin besar pengaruhnya.

Berdasarkan bentuknya variabel dibagi menjadi dua yaitu:

- a. Variabel kualitatif merupakan variabel kategori. Yang termasuk variabel kualitatif adalah variabel nominal dan variabel ordinal.
- b. Variabel kuantitatif itu bentuknya numerik (bilangan) yang dibagi menjadi variabel diskrit dan variabel kontinu. Variabel diskrit merupakan variabel yang besarannya tidak dapat menempati semua nilai, nilai variabel diskrit selalu berupa bilangan bulat dan umumnya diperoleh dari hasil pencacahan. Variabel kontinu merupakan variabel yang besarannya dapat menempati semua nilai yang ada di antara dua titik dan umumnya diperoleh dari hasil pengukuran sehingga pada variabel ini dapat dijumpai nilai-nilai pecahan ataupun nilai-nilai bulat.

Contoh:

- “usia” adalah gejala kualitatif, akan tetapi gejala yang bersifat kualitatif itu dilambangkan dengan angka; misalnya: 17 tahun, 25 tahun dan sebagainya.
- “nilai ujian” pada dasarnya adalah gejala kualitatif yang dilambangkan dengan angka, seperti; 5, 7, 8, 10, 60, 70 dan sebagainya.

C. Teknik Pengambilan Sampel

Menurut Supardi (1993), teknik pengambilan dan penarikan sampel atau teknik sampling merupakan suatu cara atau teknik yang digunakan dalam menentukan sampel penelitian.

Pengambilan sampel bertujuan untuk membantu peneliti dalam mengatasi keterbatasan-keterbatasan yang dapat peneliti jumpai di lapangan seperti:

- Apabila populasi terlalu banyak atau jangkauan terlalu luas sehingga tidak memungkinkan bagi kita untuk melakukan pengambilan data pada seluruh populasi.
- Terkendala dalam hal keterbatasan tenaga, waktu, dan biaya.
- Adanya asumsi awal bahwa keseluruhan dalam populasi bersifat seragam sehingga bisa diwakili oleh beberapa sampel yang akan kita ambil.

Dalam melakukan pengambilan sampel terdapat sistematika atau tahapan yang harus kita ikuti dengan cermat. Berikut merupakan tahapan umum dalam teknik pengambilan sampel:

- Mendefinisikan populasi yang akan diamati.
- Menentukan kerangka sampel dan kumpulan semua peristiwa yang dapat terjadi.
- Menentukan teknik atau metode sampling yang tepat.
- Melakukan pengambilan sampel (pengumpulan data).
- Melakukan pemeriksaan ulang pada proses sampling.

Terdapat beragam teknik pengambilan sampel. Macam teknik pengambilan sampel ini kita gunakan tergantung dari jenis penelitian yang kita pilih. Meski begitu, secara garis besar metode pengambilan sampel terbagi menjadi dua yaitu: probability sampling (random sampel) yaitu teknik pengambilan sampel secara acak serta non-probability sampling (non-random sampel) teknik pengambilan tidak acak.

1. Probability Sampling

Probability sampling merupakan jenis dalam teknik pengambilan sampel yang melakukan pengambilan sampelnya dengan random atau acak.

Teknik jenis ini sesuai digunakan untuk populasi yang besaran anggotanya dapat kita tentukan terlebih dahulu. Metode ini menggunakan analisis statistik untuk membantu penentuan sampel terpilihnya. Terdapat beberapa model atau jenis lain dari teknik random, yaitu:

a. Pengambilan Sampel Acak Sederhana (Simple Random Sampling)

Jenis ini melakukan pengambilan sampel secara acak melalui cara yang sederhana seperti pengundian atau menggunakan pendekatan bilangan acak.

Kelebihan penggunaan metode ini yaitu dapat mengurangi bias atau kecenderungan berpihak pada anggota populasi tertentu dan dapat mengetahui adanya kesalahan baku (standard error) dalam penelitian. Sementara itu, kelemahan dalam penggunaan metode ini yaitu rendahnya jaminan mengenai sampel yang terpilih dapat bersifat representatif atau dapat mewakili populasi yang dituju.

Contoh Pengambilan Sampel Metode Acak Sederhana: Dibutuhkan 15 sampel dari populasi penelitian dengan jumlah 90 orang. Peneliti terlebih dahulu membuat undian untuk mendapatkan sampel pertama dari 90 populasi tersebut. Setelah sampel pertama didapatkan, nama yang terpilih sebagai sampel tersebut dikembalikan lagi agar populasi tetap utuh, berjumlah 90 orang. Mengembalikan sampel terpilih memungkinkan responden berikutnya akan tetap sama dengan responden yang sudah dipilih pertama. Hal ini dilakukan terus menerus hingga jumlah 15 sampel terpenuhi.

a. Pengambilan Sampel Acak Sistematis (Systematic Random Sampling)

Pengambilan sampel pada teknik ini menetapkan sampel awal secara acak kemudian sampel selanjutnya dipilih secara sistematis berdasarkan pola tertentu. Pola umum dari teknik ini adalah mengambil bilangan kelipatan dari jumlah anggota populasi dengan jumlah sampel yang akan diambil.

Kelebihan dari penggunaan metode ini adalah cara ini lebih cepat, lebih mudah dan lebih mudah pelaksanaannya dibanding cara lainnya. Cara ini juga memudahkan peneliti karena memungkinkan kita untuk mengambil sampel di lapangan tanpa harus menggunakan kerangka sampel. Kekurangan Metode ini adalah kita tidak dapat memprediksi variasi dari populasi jika urutan yang dilakukan tidak sepenuhnya acak. Selain itu, jika populasi memiliki pengulangan karakteristik yang relatif tetap maka sampel akan cenderung sama atau bersifat seragam.

Contoh Pengambilan Sampel Metode Acak Sistematis: Misalnya, diambil sampel dari populasi dengan jumlah 40 orang yang akan masuk ke sebuah ruangan. Setiap orang yang masuk ke urutan dari kelipatan 4 akan diambil sebagai sampel, artinya orang ke-4, 8, 12, 16 dan seterusnya akan dijadikan sampel penelitian hingga 40 populasi.

a. Pengambilan Sampel Acak Berstrata (Stratified Random Sampling)

Teknik pengambilan sampel ini melakukan penentuan sampel penelitian dengan menetapkan pengelompokan anggota populasi dalam kelompok-kelompok tingkatan tertentu seperti tingkat tinggi, sedang, dan rendah.

Contoh Pengambilan Sampel Metode Acak Berstrata: Misalnya penelitian masyarakat terhadap partisipasi pemilihan umum yang dikelompokkan berdasarkan usia pemilih. Tingkatan dari kelompok tersebut akan ditentukan dari usia yang paling rendah hingga ke yang paling tinggi atau sebaliknya

a. Pengambilan Sampel Acak Berdasar Area atau Wilayah (Cluster Random Sampling)

Teknik pengambilan sampel ini menentukan sampel berdasar kelompok wilayah dari anggota populasi penelitian. Pada teknik ini subyek penelitian akan dikelompokkan menurut area atau tempat domisili anggota populasi. Tujuannya antara lain untuk meneliti tentang suatu hal pada bagian-bagian yang berbeda di dalam suatu wilayah tertentu.

Contoh Pengambilan Sampel Metode Acak Berdasarkan Area/Wilayah: Misalnya peneliti ingin mengetahui tingkat partisipasi masyarakat kota Yogyakarta terhadap program pemerintah daerah. Peneliti akan menentukan

sampel dari wilayah-wilayah yang tersebar di kota Yogyakarta. Baik pada tingkat kecamatan, desa, hingga dusun.

2. Non-Probability Sampling

Teknik pengambilan sampel non-probability merupakan cara pengambilan sampel dengan tidak memberi peluang atau kesempatan sama bagi setiap unsur atau anggota populasi yang dipilih menjadi sampel.

Menurut Supardi (1993), teknik sampling non probability akan sesuai apabila dipilih untuk populasi yang sifatnya infinit atau besaran anggota populasinya belum atau tidak dapat ditentukan terlebih dahulu sebelumnya. Berikut macam-macam dari teknik pengambilan sampel menggunakan Non-Probability Sampling:

a. Purposive Sampling

Teknik purposive sampling adalah teknik penentuan sampel yang didasarkan pada pertimbangan peneliti mengenai sampel-sampel mana yang paling sesuai, bermanfaat dan dianggap dapat mewakili suatu populasi (representatif).

Kelebihan dari metode ini di antaranya tujuan dari penelitian dapat dengan mudah terpenuhi, sampel dapat bersifat lebih relevan dengan desain penelitian, cara ini cenderung lebih murah dan mudah untuk dilaksanakan. Sementara itu kekurangannya sama dengan teknik pengambilan sampel secara acak yaitu tidak adanya jaminan bahwa sampel dapat mewakili populasi yang ditentukan.

Teknik pengambilan sampel ini cenderung lebih tinggi kualitas sampelnya. Karena peneliti telah membuat kisi atau batas berdasarkan kriteria tertentu yang akan dijadikan sampel penelitian. Misal seperti didasarkan pada ciri demografi, gender, jenis pekerjaan, umur dan lain sebagainya. Teknik ini termasuk teknik pengambilan sampel yang cukup sering digunakan dalam penelitian.

a. Snowball Sampling

Biasa dikenal juga dengan teknik pengambilan sampel bola salju. Teknik ini menentukan sampel berdasarkan wawancara dengan sampel sebelumnya atau dengan cara korespondensi.

Melakukan pengambilan sampel dengan teknik ini artinya kita bisa meminta informasi dari sampel pertama untuk mendapatkan sampel berikutnya, demikian secara terus menerus hingga akhirnya seluruh kebutuhan sampel penelitian dapat terpenuhi.

Teknik pengambilan sampel dengan metode bola salju ini sangat cocok untuk penelitian mengenai hal-hal yang sifatnya cukup sensitif dan membutuhkan privasi tingkat tinggi dari respondennya. Misal penelitian tentang penyintas kekerasan seksual, penderita HIV, kelompok waria serta kelompok-kelompok khusus lainnya.

a. Accidental Sampling

Sesuai dengan namanya, teknik pengambilan sampel jenis ini menentukan sampel secara tidak sengaja (accidental). Peneliti akan mengambil sampel pada orang yang kebetulan ditemuinya pada saat itu.

Misalnya penelitian dilakukan pada populasi pelanggan toko A, peneliti cukup menunggu di depan toko A lalu menetapkan sampel kepada siapapun orang yang melakukan transaksi jual-beli di toko A tanpa melihat umur, gender, profesi, dan lain sebagainya.

b. Quota Sampling

Teknik pengambilan sampel ini dilakukan dengan menentukan kuota atau jumlah dari sampel penelitian terlebih dahulu. Prinsip penentuannya sama dengan accidental sampling. Tetapi peneliti menetapkan terlebih dahulu jumlah sampel yang akan diperlukan.

Kelebihan menggunakan teknik ini dalam pengambilan sampel yaitu bersifat praktis karena sampel penelitian sudah diketahui sebelumnya. Sementara kekurangannya yaitu bias penelitian yang cenderung cukup tinggi dapat terjadi.

Misal peneliti menetapkan penelitian dilakukan setiap hari selama satu minggu dengan menetapkan jumlah sampel penelitian sebanyak 100 orang. Apabila peneliti pada hari itu telah memenuhi kuota dengan memperoleh 100 orang maka selesai tugas peneliti untuk mencari sampel penelitian.

D. Teknik Pengumpulan Data

Teknik pengumpulan data merupakan teknik atau metode yang digunakan untuk mengumpulkan data yang akan diteliti. Artinya, teknik ini memerlukan langkah yang strategis dan juga sistematis untuk mendapatkan data yang valid dan juga sesuai dengan kenyataannya. Secara umum, teknik pengumpulan data digunakan peneliti untuk dapat mengumpulkan data atau informasi berdasarkan fakta pendukung yang ada di lapangan demi keperluan penelitian dan teknik yang dilakukan sangat ditentukan oleh metodologi penelitian yang dipilih oleh peneliti itu sendiri. Sebelum memulai melakukan teknik pengumpulan data, ada beberapa hal yang harus diperhatikan yaitu:

1. Data biasanya sudah ditentukan oleh beberapa variabel penelitian. Ketika semua data terkumpul, langkah berikutnya adalah mengolah data, sehingga data yang dikumpulkan memiliki arti karena diolah dengan sistematis.
2. Data yang sudah diolah tersebut dipakai dan dipilih berdasarkan data yang berhubungan atau relevan dengan konsep, kejadian, atau objek penelitian. Datanya bisa berbentuk huruf, angka, simbol, gambar, dan lainnya.
3. Setelah itu, pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian yang diungkap dalam bentuk hipotesis yaitu jawaban sementara terhadap pertanyaan penelitian.
4. Data yang sudah dikumpulkan ditentukan oleh berbagai variabel yang ada di dalam hipotesis dan dikumpulkan dalam bentuk sampel yang sudah ditentukan sebelumnya dan sampelnya digunakan untuk menganalisis sasaran penelitian.

Dalam teknik pengumpulan data, tentu saja ada proses yang harus dilakukan. Prosesnya harus terlaksana secara sistematis dan terarah agar data yang dikumpulkan bisa dibuktikan kebenarannya. Karena pada dasarnya, proses pengumpulan data dalam teknik mengumpulkan data ini nanti harus bisa membuktikan hipotesis dari data yang hasilnya sudah dikumpulkan oleh peneliti. Berikut ini, ada 8 tahap atau proses yang harus dilakukan sebagai tahapan pengumpulan data:

- a. Tinjau literatur dan konsultasi dengan ahli.

Proses atau tahap pertama yang harus dilakukan untuk mengumpulkan data yakni mengumpulkan berbagai informasi yang berhubungan dengan masalah penelitian. Informasi ini diperoleh melalui tinjauan literatur dan konsultasi dengan para ahli sehingga peneliti benar-benar mengerti isu, konsep, dan variabel yang ada di dalam penelitian.

- b. Mempelajari dan melakukan pendekatan terhadap kelompok masyarakat di mana data akan dikumpulkan.

Tahap kedua atau proses yang dilakukan setelah tinjauan literatur adalah peneliti harus mempelajari dan melakukan pendekatan terhadap kelompok masyarakat yang kemudian penelitiannya bisa diterima dan juga berkaitan dengan tokoh-tokoh yang bersangkutan.

- c. Membina dan memanfaatkan hubungan yang baik dengan responden dan lingkungannya.

Tahap selanjutnya adalah membina hubungan baik dengan responden dan lingkungannya. Ini termasuk pada mempelajari bagaimana kebiasaan yang dilakukan responden dan cara berpikir mereka, melakukan sesuatu, bahasa yang digunakan, dan lain sebagainya untuk mendukung berlangsungnya penelitian.

- d. Uji coba atau pilot study.

Selanjutnya, tahapan yang harus dilakukan adalah melakukan uji coba instrumen penelitian pada kelompok masyarakat yang merupakan bagian dari populasi, bukan sampel. Maksudnya untuk mengetahui apakah instrumen yang digunakan cukup dipahami, bisa digunakan, komunikatif atau tidak, dan lain sebagainya.

- e. Merumuskan dan menyusun pertanyaan.

Setelah itu, instrumen yang sudah didapatkan disusun dalam bentuk pertanyaan yang relevan dengan tujuan penelitian. Pertanyaan yang dirumuskan harus mengandung makna yang signifikan dan substantif.

- f. Mencatat dan memberi kode (recording and coding).

Setelah instrumen penelitian disiapkan, dilakukan pencatatan terhadap data yang dibutuhkan dari setiap responden. Berbagai informasi yang diperoleh ini perlu dicatat guna memudahkan proses analisis.

g. Cross checking, validitas, dan reliabilitas.

Setelah itu, dilakukan metode cross checking terhadap data yang didapatkan untuk menguji lagi kebenarannya dan memeriksa sehingga tidak ada keraguan terhadap validitas dan reliabilitasnya.

h. Pengorganisasian dan kode ulang data yang telah terkumpul supaya dapat dianalisis.

Terakhir, setelah data terkumpul, penulis harus melakukan koordinasi terhadap berbagai data yang sudah dikumpulkan, dan Anda bisa mulai menganalisis data tersebut sehingga tidak ada data yang kurang valid.

Setelah memahami pengertian dan juga proses pengumpulan data, selanjutnya akan dijelaskan mengenai berbagai teknik pengumpulan data. Setidaknya ada empat teknik. Berikut ini merupakan teknik pengumpulan data menurut Sugiyono (2017).

a. Wawancara

Wawancara adalah teknik pengumpulan data yang dilakukan melalui tatap muka dan tanya jawab langsung antara peneliti dan narasumber. Seiring perkembangan teknologi, metode wawancara dapat pula dilakukan melalui media-media tertentu, misalnya telepon, email, atau video call melalui Zoom atau skype. Wawancara terbagi atas dua kategori, yakni wawancara terstruktur dan tidak terstruktur.

b. Wawancara terstruktur

Dalam wawancara terstruktur, peneliti telah mengetahui dengan pasti informasi apa yang hendak digali dari narasumber. Pada kondisi ini, peneliti biasanya sudah membuat daftar pertanyaan secara sistematis. Peneliti juga bisa menggunakan berbagai instrumen penelitian seperti alat bantu recorder, kamera untuk foto, serta instrumen-instrumen lain.

c. Wawancara tidak terstruktur

Wawancara tidak terstruktur adalah wawancara bebas. Peneliti tidak menggunakan pedoman wawancara yang berisi pertanyaan-pertanyaan spesifik, namun hanya memuat poin-poin penting dari masalah yang ingin digali dari responden.

d. Observasi

Observasi adalah metode pengumpulan data yang kompleks karena melibatkan berbagai faktor dalam pelaksanaannya. Metode pengumpulan data observasi tidak hanya mengukur sikap dari responden, namun juga dapat digunakan untuk merekam berbagai fenomena yang terjadi. Teknik pengumpulan data observasi cocok digunakan untuk penelitian yang bertujuan untuk mempelajari perilaku manusia, proses kerja, dan gejala-gejala alam. Metode ini juga tepat dilakukan pada responden yang kuantitasnya tidak terlalu besar. Metode pengumpulan data observasi terbagi menjadi dua kategori, yakni:

e. Participant observation

Dalam participant observation, peneliti terlibat secara langsung dalam kegiatan sehari-hari orang atau situasi yang diamati sebagai sumber data. Contoh observasi partisipasi adalah ketika meneliti adat/tradisi pada kelompok masyarakat tertentu. Dalam hal ini, peneliti tidak hanya menonton tapi juga menjadi bagian dari kelompok tersebut.

f. Non participant observation

Berlawanan dengan participant observation, non-participant observation merupakan observasi yang peneliti tidak ikut secara langsung dalam kegiatan atau proses yang sedang diamati. Metode ini biasa dipakai untuk meneliti hal-hal yang berhubungan dengan sikap dan perilaku negatif. Contohnya penelitian tentang perilaku membuang sampah sembarangan siswa di sekolah tertentu.

a) Angket (kuesioner)

Kuesioner merupakan metode pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawab. Kuesioner merupakan metode pengumpulan data yang lebih efisien bila peneliti telah mengetahui dengan pasti variabel yang akan diukur dan tahu apa yang diharapkan dari responden. Selain itu, kuesioner juga cocok digunakan bila jumlah responden cukup besar dan tersebar di wilayah yang luas.

Berdasarkan bentuk pertanyaannya, kuesioner dapat dikategorikan dalam dua jenis, yakni kuesioner terbuka dan kuesioner tertutup. Kuesioner terbuka adalah kuesioner yang memberikan kebebasan kepada objek penelitian untuk

menjawab. Sementara itu, kuesioner tertutup adalah kuesioner yang telah menyediakan pilihan jawaban untuk dipilih oleh objek penelitian. Seiring dengan perkembangan, beberapa penelitian saat ini juga menerapkan metode kuesioner yang memiliki bentuk semi terbuka. Dalam bentuk ini, pilihan jawaban telah diberikan oleh peneliti, namun objek penelitian tetap diberi kesempatan untuk menjawab sesuai dengan kemauan mereka.

Contohnya seorang petugas kecamatan memberikan formulir yg memuat data keluarga.

b) Studi Dokumen

Studi dokumen adalah metode pengumpulan data yang tidak ditujukan langsung kepada subjek penelitian. Studi dokumen adalah jenis pengumpulan data yang meneliti berbagai macam dokumen yang berguna untuk bahan analisis. Dokumen yang dapat digunakan dalam pengumpulan data dibedakan menjadi dua, yakni:

a. Dokumen primer

Dokumen primer adalah dokumen yang ditulis oleh orang yang langsung mengalami suatu peristiwa, misalnya: autobiografi, yaitu riwayat hidup seseorang yang ditulis oleh dirinya sendiri.

b. Dokumen sekunder

Dokumen sekunder adalah dokumen yang ditulis berdasarkan oleh laporan/ cerita orang lain, misalnya: biografi adalah karya sastra yang berisikan riwayat hidup seorang tokoh yang dianggap penting oleh masyarakat tertentu.

BAB II

STATISTIK DESKRIPTIF DAN STATISTIK INFERENSIAL

A. Pengertian Statistik Deskriptif dan Statistik Inferensial

Statistik deskriptif adalah jenis statistik yang digunakan untuk menggambarkan atau merangkum data. Statistik deskriptif dapat memberikan informasi tentang pemusatan data, penyebaran data, dan bentuk distribusi data. Statistik deskriptif memberikan informasi dari suatu sampel atau populasi dalam bentuk yang ringkas. Tujuan penggunaan statistik deskriptif adalah untuk mengetahui bagaimana kondisi data yang dimiliki.

Ukuran pemusatan adalah metode paling lazim yang digunakan dalam analisis statistik deskriptif. Metode ini berfokus untuk menggambarkan kondisi data di titik pusat. Secara umum, dapat dilihat bagaimana kondisi data dengan melihat dimana letak pusat data tersebut.

Statistik inferensial adalah jenis statistik yang digunakan untuk menguji hipotesis dan membuat kesimpulan berdasarkan data sampel yang diambil dari populasi. Statistik inferensial menggunakan data sampel untuk mendapatkan gambaran informasi dari suatu kelompok data populasi. Dari statistik inferensial memungkinkan untuk melakukan prediksi, mencari hubungan atau ketergantungan suatu himpunan data yang lebih besar/banyak berdasarkan data sampel yang dimiliki.

Statistik inferensial menggunakan metode estimasi parameter dan uji hipotesis. Estimasi parameter adalah suatu metode yang memakai perhitungan dari hasil sampel meliputi standar deviasi, modus, median, dan mean dalam analisa populasi. Uji hipotesis adalah menggunakan cara pelaksanaan kegiatan perbandingan statistik berdasarkan rata-rata dari sampel sebanyak dua sampel sekaligus.

Statistik deskriptif maupun inferensial digunakan pada waktu seseorang hendak menggambarkan tentang sifat dari kumpulan data. Untuk statistik deskriptif menyajikan data dalam wujud tabel atau diagram. Sedangkan waktu yang tepat untuk menggunakan jenis inferensial adalah ketika akan melakukan perbandingan perbedaan tindakan di sejumlah kelompok.

B. Perbedaan Statistik Deskriptif dan Statistik Inferensial

Perbedaan utama antara statistik deskriptif dan statistik inferensial adalah bahwa statistik deskriptif hanya digunakan untuk merangkum data, sedangkan statistik inferensial digunakan untuk menguji hipotesis dan membuat kesimpulan berdasarkan sampel yang diambil dari populasi.

Statistik Deskriptif	Statistik Inferensial
Merupakan metode untuk menggambarkan, meringkas, dan menganalisis data secara numerik atau grafis	Merupakan metode yang digunakan untuk membuat kesimpulan umum tentang populasi berdasarkan sampel yang diambil dari populasi
Tidak digunakan untuk membuat generalisasi atau inferensi tentang populasi secara keseluruhan	Digunakan untuk membuat generalisasi atau inferensi tentang populasi secara keseluruhan
Contoh statistik deskriptif adalah mean, median, modus, deviasi standar, dan korelasi	Contoh statistik inferensial adalah uji hipotesis, interval kepercayaan, analisis regresi, dan analisis variansi
Statistik deskriptif dapat digunakan untuk menggambarkan data dalam bentuk tabel, grafik, dan diagram	Statistik inferensial memerlukan penghitungan probabilitas untuk menentukan signifikansi atau kepercayaan pada kesimpulan yang dibuat
Tujuan dari statistik deskriptif adalah untuk memberikan gambaran yang jelas tentang data yang dianalisis	Tujuan dari statistik inferensial adalah untuk membuat kesimpulan tentang populasi berdasarkan sampel yang diambil
Tidak ada pengujian hipotesis yang dilakukan pada statistik deskriptif	Pengujian hipotesis merupakan salah satu teknik penting dalam statistik inferensial

Digunakan pada situasi di mana peneliti hanya ingin mengetahui karakteristik dari sampel yang dianalisis	Digunakan pada situasi di mana peneliti ingin membuat inferensi tentang populasi berdasarkan sampel yang diambil
--	--

C. Jenis-jenis Statistik Deskriptif dan Statistik Inferensial

Pada umumnya ada tiga model penyajian data dalam statistika deskriptif, meliputi:

1. Distribusi frekuensi

Distribusi adalah cara untuk menunjukkan frekuensi atau banyaknya objek di setiap kelas. Tujuan distribusi adalah untuk mendapatkan informasi lebih dalam tentang kumpulan data yang tidak bisa diperoleh secara cepat hanya melihat data asli.

Distribusi ditampilkan dengan grafik atau tabel. Peneliti dapat meringkas frekuensi dari suatu variabel dalam bentuk angka atau persentase. Sebagai contoh, berikut adalah jumlah motor terjual pada bulan Juni 2022.

Merek Motor	Jumlah Terjual
Honda	35
Yamaha	26
Suzuki	9
Kawasaki	16

Dari tabel di atas dapat diketahui beberapa hal:

- Merek motor yang terjual paling banyak adalah Honda.
- Merek motor yang terjual paling sedikit adalah Suzuki.
- Yamaha menduduki urutan kedua merek motor yang terjual paling banyak dibulan Juni.

2. Tendensi Sentral

Tendensi sentral berfokus pada nilai rata-rata dari kumpulan data. Peneliti dapat menggunakan tendensi sentral untuk merangkum dan mendeskripsikan kelompok variabel. Jenis statistika deskriptif ini merujuk pada kecenderungan letak pusat. Tendensi sentral menggunakan tiga metode untuk mencari hasil, yaitu mean, median, dan modus.

a. Mean (rata-rata)

Mean adalah nilai rata-rata dari suatu kumpulan data. Untuk mencari mean, kita menjumlahkan semua nilai kemudian dibagi dengan jumlah data.

b. Median

Median adalah nilai tengah dari data yang diurutkan dari yang terkecil hingga terbesar. Jika jumlah data genap, maka median diambil dari rata-rata dua nilai tengah.

c. Modus

Modus adalah nilai yang paling sering muncul dalam kumpulan data. Jika tidak ada nilai yang muncul lebih dari satu kali, maka tidak ada modus.

3. Variabilitas

Variabilitas berfungsi menganalisis bagaimana persebaran distribusi dalam suatu kumpulan data. Ada tiga aspek dalam variabilitas yaitu:

a. Range

Range adalah selisih antara nilai tertinggi dan nilai terendah dalam suatu kumpulan data. Range memberikan gambaran tentang jarak antara nilai terbesar dengan nilai terkecil dalam kumpulan data. Rumus Range = nilai tertinggi – nilai terendah.

b. Standard Deviation

Nilai yang berfungsi menentukan persebaran data dalam suatu sampel. Standard deviation menunjukkan seberapa dekat data tersebut dengan mean.

c. Variance

Variance adalah ukuran seberapa jauh data tersebar dari mean. Semakin besar variance, semakin jauh data tersebar dari mean.

Adapun jenis-jenis statistik inferensial sebagai berikut.

1. Statistik parametrik

Statistik parametrik adalah jenis statistik inferensial yang membahas parameter-parameter populasi yang membahas parameter-parameter populasi, jenis datanya interval atau rasio, dan distribusi data normal.

Contoh uji statistik parametrik: analisis korelasi. Analisis korelasi pada uji statistik parametrik digunakan untuk menguji hubungan antar variabel. Hubungan yang dimaksud disini adalah keeratan hubungan antara dua variabel atau lebih tanpa memperhatikan ada tidaknya hubungan kausal di antara variabel-variabel yang dianalisis.

Analisis korelasi yang digunakan untuk menguji data yang bersifat parametrik sering juga disebut dengan analisis Pearson. Metode analisis korelasi Pearson merupakan metode pengujian yang pada dasarnya dilakukan terhadap suatu variabel pengukuran yang menyajikan data-data yang bersifat kuantitatif. Metode Pearson ini dapat ditemukan pada aplikasi Minitab, baik Minitab 14 maupun Minitab 15.

2. Statistik Nonparametrik

Statistik nonparametrik adalah jenis inferensial yang tidak membahas parameter-parameter populasi, jenis datanya nominal atau ordinal, dan distribusi data tidak diketahui.

Contoh uji statistik non parametrik: chi square test. Uji chi square pada uji statistik nonparametrik adalah salah satu cara yang digunakan untuk menyampaikan atau menunjukkan keberadaan hubungan antara variabel yang diteliti. Misalkan sebagai peneliti, kita hendak melakukan uji terhadap perilaku mahasiswa.

Karakter yang akan diuji adalah perilaku mahasiswa yang dikategorikan menjadi dua kategori. Kategori pertama yaitu mahasiswa yang mendukung program kampus dan kedua adalah yang acuh terhadap program kampus. Kondisi tersebut memungkinkan kita untuk melakukan uji hipotesis mengenai perbedaan perilaku mahasiswa tersebut dilihat dari frekuensinya.

Ada beberapa jenis metode statistik inferensial, di antaranya adalah uji t, ANOVA, chi-square, dan regresi.

a. Uji t

Uji t adalah metode statistik yang digunakan untuk membandingkan rata-rata dua kelompok yang independen.

b. Analisis Varian (ANOVA)

Analisis varian (ANOVA) adalah metode statistik yang digunakan untuk membandingkan rata-rata tiga kelompok atau lebih.

c. Chi-square

Uji chi-square adalah metode statistik yang digunakan untuk menentukan apakah ada hubungan antara dua variabel kategori.

d. Regresi

Analisis regresi adalah metode statistik yang digunakan untuk menentukan hubungan antara dua variabel atau lebih.

D. Implikasi dari Hasil Data yang Didapat Menggunakan Statistik Deskriptif dan Statistik Inferensial.

Hasil analisis data yang didapat menggunakan statistik deskriptif dan statistik inferensial dapat memiliki implikasi yang besar terhadap pengambilan keputusan. Hasil analisis data dapat membantu kita dalam mengambil keputusan yang lebih tepat dan akurat. Berikut adalah beberapa implikasi yang lebih spesifik dari hasil analisis data yang dapat menggunakan statistik deskriptif dan statistik inferensial:

a. Mengidentifikasi masalah dan solusi yang mungkin terjadi

Misalnya, jika analisis data menunjukkan bahwa penjualan produk menurun secara signifikan pada bulan tertentu, maka dapat dianggap sebagai masalah yang perlu diatasi. Dari analisis ini, kita dapat mengambil tindakan yang tepat untuk mengatasi masalah tersebut, seperti meningkatkan pemasaran atau memperbaiki kualitas produk.

b. Meningkatkan efisiensi dan efektivitas keputusan

Contohnya, jika data menunjukkan bahwa karyawan yang memiliki lebih banyak pengalaman cenderung menghasilkan lebih banyak produk atau layanan, maka dapat dipertimbangkan untuk merekrut karyawan yang memiliki pengalaman yang lebih tinggi.

c. Mengukur kinerja dan kemajuan

Statistik deskriptif dan inferensial dapat digunakan untuk mengukur kinerja dan kemajuan dalam bisnis atau organisasi. Misalnya, jika kita menggunakan analisis regresi untuk mengukur keterkaitan antara pengeluaran pemasaran dengan penjualan produk, maka kita dapat mengevaluasi kinerja pemasaran kita dan membuat perubahan yang diperlukan untuk meningkatkan penjualan.

d. Mengetahui preferensi dan perilaku pelanggan

Misalnya, dengan menggunakan analisis cluster, kita dapat mengidentifikasi kelompok pelanggan yang memiliki preferensi yang sama dan memberikan pelayanan atau produk yang sesuai dengan preferensi mereka.

e. Menemukan tren dan pola dalam data

Statistik deskriptif dan inferensial dapat membantu kita menemukan tren dan pola dalam data. Misalnya, jika kita menggunakan analisis regresi untuk mengukur keterkaitan antara pengeluaran pemasaran dengan penjualan produk, kita dapat mengetahui apakah ada tren peningkatan atau penurunan penjualan terkait dengan pengeluaran pemasaran.

f. Mengoptimalkan penggunaan sumber daya

Hasil analisis data dapat membantu kita mengoptimalkan penggunaan sumber daya. Misalnya, jika analisis data menunjukkan bahwa sebagian besar pelanggan kita berasal dari wilayah tertentu, maka dapat dipertimbangkan untuk memfokuskan pemasaran di wilayah tersebut untuk mengoptimalkan penggunaan sumber daya.

E. Contoh Statistik Deskriptif dan Statistik Inferensial

Contoh statistika deskriptif yaitu tabel, diagram, grafik, dan besaran-besaran lain di majalah dan koran-koran. Contoh statistik deskriptif adalah ketika mengumpulkan data tentang nilai ujian semua siswa kelas 11 di sekolah selama tiga tahun. Statistika deskriptif dapat digunakan untuk mendapatkan gambaran singkat tentang nilai pada tahun-tahun tersebut. Kemudian dapat membandingkan skor ujian rata-rata dengan skor rata-rata sekolah lain.

Contoh statistik inferensial adalah memilih secara acak sampel siswa kelas 11 di suatu kota dan mengumpulkan data tentang nilai ujian mereka serta

karakteristik lainnya. Dalam hal ini, dapat menggunakan statistik inferensial untuk membuat dan menguji hipotesis tentang populasi siswa kelas 11 di kota tersebut berdasarkan data sampel.

BAB III

PENYIAPAN DAN PENYAJIAN DATA

A. Penyajian dan Penyiapan Data

Menurut Encyclopedia Britannica (2020), data adalah fakta atau angka-angka yang dikumpulkan, dianalisis dan diringkas untuk disajikan. Contoh sederhananya seperti informasi tentang harga sepatu yang berbeda-beda di banyak toko tadi.

Penyajian data adalah upaya untuk menampilkan atau memaparkan data yang didapatkan secara visual. Tujuan penyajian data untuk mempermudah proses pemahaman dan analisis data yang jumlahnya banyak merupakan alasan mengapa data perlu disajikan.

B. Jenis-jenis Data

Jenis-jenis data dapat dibedakan menurut beberapa kategori, misalnya: menurut sifatnya, cara memperolehnya, berdasarkan sumber datanya, waktu pengumpulannya, dan skala pengukurannya. Berikut akan dijelaskan jenis-jenis data berdasarkan keempat kategorisasi tersebut.

1. Berdasarkan Sifat

Berdasarkan sifatnya, data dikelompokkan ke dalam dua kelompok, yaitu data kualitatif dan data kuantitatif.

a. Data kualitatif

Data yang dinyatakan dalam bentuk non-angka atau non-numerik atau biasa juga disebut atribut. Dalam istilah komputer disebut data bertipe string. Misalnya, persepsi mahasiswa UT terhadap layanan tutorial di UPBJJ-UT. Data yang diberikan berupa persepsi atau pendapat, sehingga jawaban yang mungkin adalah sangat baik, baik, cukup, kurang baik, sangat kurang baik.

b. Data kuantitatif

Data yang dinyatakan dalam bentuk angka (data numerik). Dalam komputer dikenal sebagai data numerik. Misalnya, data jumlah pembelian modul UT, jumlah mahasiswa Jurusan Statistika Masa Registrasi 2010.2, dan

sebagainya. Data yang dapat dinyatakan dalam angka (dapat dikuantifikasi) dinamakan data kuantitatif. Data kuantitatif dikelompokkan menjadi dua, yaitu :

1) Data diskret

Data kuantitatif diskret adalah data hasil pencacahan dan berupa bilangan bulat (dalam komputer dikenal sebagai integer). Misalnya: jumlah mahasiswa statistika di UPBJJ-UT Bandung, jumlah mata kuliah yang ada di Jurusan Statistika FMIPA-UT, dan jumlah dosen Jurusan Statistika FMIPA-UT.

2) Data kontinu

Data kuantitatif kontinu adalah data hasil proses pengukuran dan dapat berupa bilangan pecahan (bilangan real). Misalnya: Rata-rata berat badan bayi yang baru lahir adalah 2,95 kg, tinggi badan Budi adalah 150,5 cm, dan IQ Budi adalah 125.

2. Berdasarkan Cara Memperolehnya

Berdasarkan cara memperolehnya, data dikelompokkan ke dalam dua kelompok, yaitu data primer dan data sekunder.

a. Data primer

Data primer adalah data yang diperoleh atau dikumpulkan langsung oleh peneliti dari sumber datanya. Beberapa teknik pengumpulan data primer antara lain: observasi, wawancara, diskusi terfokus (focus group discussion, FGD), dan penyebaran kuesioner.

b. Data sekunder

Data sekunder adalah data yang diperoleh secara tidak langsung dari objek penelitian. Data sekunder dapat diperoleh dari berbagai sumber seperti: Biro Pusat Statistik (BPS), buku, laporan, jurnal, dan sumber data lainnya.

3. Berdasarkan Sumber Datanya

Berdasarkan sumbernya, data dikelompokkan ke dalam dua kelompok, yaitu data internal dan data eksternal.

a. Data internal

Data internal adalah data yang menggambarkan situasi dan kondisi pada suatu organisasi atau instansi secara internal. Misalnya: data keuangan, data pegawai, data produksi, dan sebagainya.

b. Data eksternal

Data eksternal adalah data yang menggambarkan situasi dan kondisi yang ada di luar organisasi. Misalnya: data jumlah penggunaan modul oleh mahasiswa di luar UT, tingkat kepuasan mahasiswa UT terhadap layanan akademik dan administratif UT, dan data sebaran mahasiswa UT.

4. Berdasarkan Waktu Pengumpulan Data

Berdasarkan waktu pengumpulan datanya, data dikelompokkan ke dalam dua kelompok, yaitu data cross section, dan data time series (berkala).

a. Data cross section

Data cross section adalah data yang menunjukkan titik waktu tertentu. Misalnya: laporan nilai tutorial online mahasiswa Jurusan Statistika FMIPA-UT Masa Registrasi 2010.2, data jumlah mahasiswa yang meregistrasi mata kuliah Pengumpulan dan Penyajian Data Masa Registrasi 2010.1.

b. Data time series

Data time series (berkala) adalah data yang datanya menggambarkan sesuatu dari waktu ke waktu atau periode secara historis. Contoh data time series adalah data perkembangan jumlah mahasiswa Jurusan Statistika FMIPA-UT dari 2005 sampai dengan 2010, jumlah mahasiswa Jurusan Statistika yang mengikuti ujian TAP dari tahun 2005 sampai dengan 2010.

5. Berdasarkan Skala

Berdasarkan skala pengukurannya, data dibedakan menjadi 4 kelompok, yaitu: data nominal, ordinal, interval dan rasio. Rincian mengenai ciri-ciri dan contoh masing-masing data dibawah ini:

a. Nominal

Digunakan untuk mengklarifikasikan informasi/data. Adapun ciri-ciri nominal :

- Dibedakan dalam kategori tanpa memperhatikan urutan.
- Satu pengukuran hanya menghasilkan satu-satunya kategori.
- Setiap kategori dianggap sama (tanpa tingkatan).
- Data paling “rendah” dalam level pengukuran data.
- Tidak bisa dioperasikan secara matematis.

Contoh :

- Jenis kelamin.
 - Jenis sabun.
 - Suku.
 - Data alamat.
 - Tanggal/tempat lahir.
 - Agama.
- b. Ordinal

Digunakan untuk mengklarifikasikan serta memiliki tingkatan. Adapun ciri-ciri ordinal :

- Dibedakan dalam kategori berdasarkan urutan.
- Memiliki tingkatan data.
- Lebih “tinggi” dibanding data nominal dalam level pengukuran data.
- Tidak bisa dioperasikan secara matematis.

c. Interval

Adapun ciri-ciri interval :

- Urutan bertingkat dan dikuantifikasi (diberi nilai).
- Memiliki interval tertentu.
- Lebih “tinggi” dibanding data ordinal dalam level pengukuran.
- Dapat dianalisis dengan uji statistik parametrik.

d. Rasio

Adapun ciri-ciri rasio :

- Data bersifat angka dalam arti yang sesungguhnya.
- Memiliki angka nol absolut.
- Memiliki kedudukan paling “tinggi” dalam level pengukuran data.
- Dapat dioperasikan secara matematis.

C. Pembulatan Angka

Dalam kehidupan sehari-hari kita sering kali dihadapkan dengan bilangan-bilangan pecahan yang memiliki angka desimal dengan banyak angka di belakang koma. Dengan alasan kepraktisan dalam penyajian data, kita diharuskan untuk membulatkan bilangan-bilangan pecahan tersebut.

Dalam Statistika terdapat aturan-aturan pembulatan bilangan yang harus dijadikan acuan dalam mengolah data satistika. Terdapat tiga aturan pembulatan dalam statistika, yaitu:

Aturan 1

Jika angka terkiri dari angka yang harus dihilangkan kurang dari 5, maka angka terkanan dari angka yang mendahuluinya tetap (tidak berubah).

Contoh :

50,16482 ton akan dibulatkan hingga dua angka di belakang koma menjadi 50,16 ton angka yang harus dihilangkan adalah 482 dengan angka terkiri 4 (kurang dari 5). maka angka terkanan dari angka yang mendahuluinya (yaitu 6) tetap (tidak berubah).

Aturan 2

Jika angka terkiri dari angka yang harus dihilangkan lebih dari 5 atau angka 5 diikuti oleh angka-angka bukan nol semua, maka angka terkanan dari angka yang mendahuluinya bertambah dengan satu.

Contoh:

50,14652 akan dibulatkan hingga dua angka di belakang koma menjadi 50,15.

50,14501 akan dibulatkan hingga dua angka di belakang koma menjadi 50,15.

Aturan 3

Jika angka terkiri dari angka yang harus dihilangkan sama dengan 5 atau angka 5 diikuti oleh angka-angka nol semua, maka angka terkanan dari angka

yang mendahuluinya tetap jika angka tsb genap, dan bertambah satu jika angka tsb ganjil.

Contoh:

50,14500 akan dibulatkan hingga dua angka di belakang koma menjadi 50,14.

50,13500 akan dibulatkan hingga dua angka di belakang koma menjadi 50,14.

BAB IV

PEMERIKSAAN DAN PENYAJIAN DATA

A. Pemeriksaan Data

Pemeriksaan adalah pengecekan atau penelitian kembali data yang telah dikumpulkan untuk mengetahui dan menilai kesesuaian dan relevansi data yang dikumpulkan untuk bisa diproses lebih lanjut. Pemeriksaan data atau editing dilakukan terhadap jawaban yang telah ada dalam kuesioner dengan memperhatikan hal-hal meliputi: kelengkapan pengisian jawaban, kejelasan tulisan, kejelasan makna jawaban, serta kesesuaian antara jawaban.

Proses editing merupakan proses dimana peneliti melakukan klarifikasi, konsistensi dan kelengkapan data yang sudah terkumpul.

1. Proses klarifikasi menyangkut memberikan penjelasan mengenai apakah data yang sudah terkumpul akan menciptakan masalah konseptual atau teknis pada saat peneliti melakukan analisa data. Dengan adanya klarifikasi ini diharapkan masalah teknis atau konseptual tersebut tidak mengganggu proses analisis sehingga dapat menimbulkan bias penafsiran hasil analisis. Hal ini berkaitan dengan apakah data yang sudah terkumpul secara logis dapat digunakan sebagai penafsiran terhadap hasil analisa.
2. Konsistensi mencakup keajegan jenis data berkaitan dengan skala pengukuran yang akan digunakan.
3. Kelengkapan mengacu pada terkumpulnya data secara lengkap sehingga dapat digunakan untuk menjawab masalah yang sudah dirumuskan dalam penelitian tersebut.

Pemeriksaan data adalah mengoreksi apakah data yang terkumpul sudah cukup lengkap, sudah benar dan sudah sesuai atau relevan dengan masalah. Misalkan, seorang guru mencatat hasil ulangan matematika seluruh siswanya. Sebelum mencari nilai rata-ratanya, ia perlu memeriksa untuk memastikan data yang diperolehnya tidak salah catat. Ia juga perlu memeriksa apakah ada nilai-nilai yang harus dibulatkan atau tidak. Kesalahan catatan dan pembulatan data ini akan menyebabkan nilai rata-rata ulangan Matematika pada kelas tersebut tidak sesuai dengan data yang sebenarnya. Tahap pemeriksaan data, Tahap ini memiliki

tujuan untuk mengetahui kelayakan data guna melanjutkan analisis data pada tahap selanjut.

B. Penyajian Data

Penyajian data merupakan salah satu kegiatan dalam pembuatan laporan hasil penelitian yang telah dilakukan agar data yang telah dikumpulkan dapat dipahami dan dianalisis sesuai dengan tujuan yang diinginkan. Setelah memperoleh data Penyajian data adalah bentuk pengemasan suatu data secara visual sedemikian sehingga data lebih mudah dipahami. Tanpa ada penyajian yang tepat, seorang peneliti akan kesulitan untuk menganalisis hasil akhir penelitian. Data yang disajikan harus sederhanaan jelas agar mudah dibaca. Penyajian data juga dimaksudkan agar para pengamat dapat dengan mudah memahami apa yang kita ajukan untuk selanjutnya dilakukan penilaian atau perbandingan dan lain lain.

Prinsip dasar penyajian data adalah komunikatif dan lengkap, dalam arti data yang disajikan dapat menarik perhatian orang lain untuk membaca dan memahami isinya. Penyajian data yang komunikatif dapat dilakukan dengan: penyajian data dapat dibuat berwarna, dan bila data yang disajikan dalam bentuk banyak maka perlu bervariasi penyajiannya (tidak hanya dengan tabel saja). Beberapa cara penyajian data menurut sugiyono yang dikemukakan disini adalah: dengan tabel, grafik atau diagram dan piktogram.

1. Penyajian Data dalam bentuk tabel

Tabel adalah penyajian data dalam bentuk kolom-kolom untuk mempermudah pengelompokan data data dalam tabel dapat berupa angka dan frekuensi untuk menyajikan data dalam bentuk tabel ada beberapa langkah yang harus dilakukan yaitu:

- 1.) Mengurutkan data yang tersedia dari bilangan terkecil hingga terbesar.
- 2.) Membuat tabel distribusi frekuensi.
- 3.) Jika data yang tersedia berupa makanan warna kesukaan tokoh idola dan lain sebagainya maka data yang tersedia tidak perlu diurutkan. Secara umum, tabel dibedakan menjadi tiga, yaitu sebagai berikut :

a. Tabel Baris Kolom

Tabel baris kolom merupakan jenis tabel yang paling sederhana. Pada tabel ini, hanya memuat keterangan dengan satu variabel frekuensi.

Contoh

Tabel pertumbuhan kecambah hari ke-1 sampai ke-4.

Hari Ke-	Pertumbuhan (cm)
1	2 mm
2	4 mm
3	2 mm
4	3 mm

b. Tabel Kontingensi

Tabel kontingensi adalah jenis tabel yang memuat lebih dari satu variabel frekuensi. Perhatikan contoh berikut.

Jenis Buah	Kondisi	
	Fresh (Kg)	Busuk (Kg)
Strawberry	5.000	20
Kiwi	2.500	10
Apel Fuji	3.000	25
Anggur Hijau	1.500	5

c. Tabel Distribusi Frekuensi

Tabel distribusi frekuensi ini biasa digunakan untuk data yang memiliki sebaran dan frekuensi cukup banyak. Jika data-data tersebut dijadikan data tunggal, tentu tidak akan efektif. Itulah mengapa perlu tabel distribusi frekuensi.

Adapun contoh tabel distribusi frekuensi. Tabel distribusi frekuensi tinggi badan siswa SD Taruna Jaya.

Tinggi Badan (cm)	Frekuensi
133-137	20
138-142	25
132-147	10
148-152	9
153-157	11

Pada tabel di atas terdapat interval atau rentang untuk tinggi badan siswa SD Taruna Jaya. Banyaknya interval atau rentang di setiap baris harus sama, ya. Misalnya, interval $133 - 137 = 5$, $138 - 142 = 5$, dan seterusnya sampai mencapai data terbesarnya. Untuk menentukan banyaknya interval (lebar kelas), kamu bisa menggunakan aturan Sturges seperti yang telah dijelaskan pada artikel statistika deskriptif.

3. Penyajian Data dalam Bentuk Diagram atau Grafik

Diagram adalah bentuk penyajian data berupa gambar dua dimensi yang merepresentasikan keseluruhan data. Secara umum, diagram dibagi menjadi tiga yaitu diagram batang, diagram garis, dan diagram lingkaran. Grafik adalah metode penyajian data dengan menggunakan gambar-gambar yang umumnya digunakan untuk melihat perubahan yang terjadi dalam sebuah variabel atau untuk membandingkan beberapa variabel yang memiliki karakteristik yang sama.

a. Diagram Batang

Diagram batang adalah diagram yang berbentuk persegi panjang. Penggambaran batangnya bisa mengarah vertikal maupun horizontal. Diagram ini biasa digunakan untuk menyajikan data seperti hasil penjualan, jumlah kunjungan perpustakaan, hasil panen tahunan, dan sebagainya.

Langkah-langkah dalam membuat diagram batang dari suatu data yaitu:

- 1.) Membuat tabel distribusi frekuensi Buat bidang koordinat kartesius yang terdiri dari garis mendatar (horizontal) dan garis tegak (vertikal).
- 2.) Tuliskan nama data pada garis mendatar atau tegak dan frekuensi dari masing-masing nama serta label keterangan pada kedua sumbu tersebut.
- 3.) Buat persegi panjang yang sesuai antara nama masing-masing data dengan frekuensinya.
- 4.) Beri nama atau judul pada diagram.

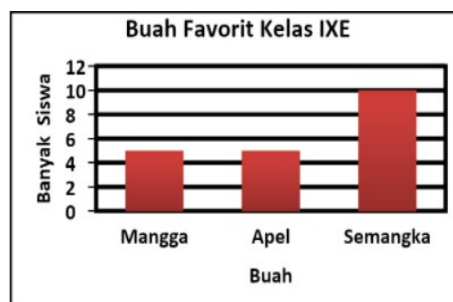
Contoh diagram batang sebagai berikut:

Berikut adalah data buah Favorit siswa kelas IX. Mangga, Apel, Semangka, Semangka, Mangga, Apel, Semangka, Semangka, Semangka, Mangga Semangka, Mangga, Apel, Semangka, Apel, Semangka, Semangka, Mangga, Apel, Semangka Dari informasi di atas, sajikan data dengan menggunakan diagram batang.

Penyelesaian : Tabel distribusi frekuensi.

Buah	Frekuensi (f)
Mangga	5
Apel	5
Semangka	10
Jumlah	20

Maka diagram batangnya adalah sebagai berikut :



b. Diagram Garis

Diagram garis adalah diagram yang digambarkan sebagai garis lurus yang menghubungkan antara data satu dan lainnya. Sebelum membuat garis lurus, kamu harus menentukan koordinat atau titik data. Lalu, hubungkan titik-titik data tersebut hingga diperoleh suatu grafik.

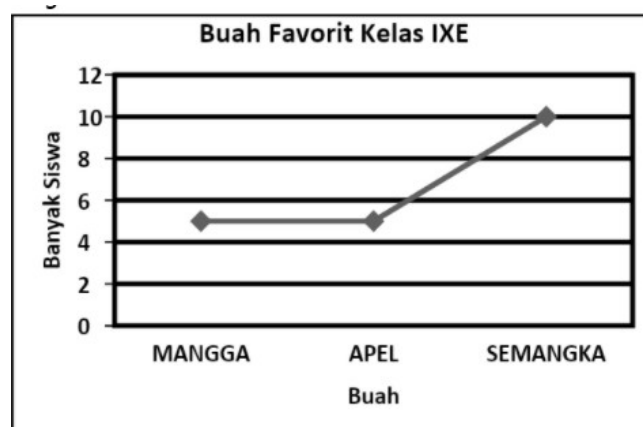
Contoh diagram batang sebagai berikut:

Berikut adalah data buah Favorit siswa kelas IX. Mangga, Apel, Semangka, Semangka, Mangga, Apel, Semangka, Semangka, Semangka, Mangga Semangka, Mangga, Apel, Semangka, Apel, Semangka, Semangka, Mangga, Apel, Semangka. Dari informasi di atas, sajikan data dengan menggunakan diagram batang.

Penyelesaian :Tabel distribusi frekuensi.

Buah	Frekuensi (f)
Mangga	5
Apel	5
Semangka	10
Jumlah	20

Maka diagram barisnya sebagai berikut :



c. Diagram Lingkaran

Gambar pada diagram lingkaran akan terbagi atas beberapa juring-juring dengan sudut pusat juring menggambarkan jumlah data yang diwakili. Untuk menggambarkan diagram lingkaran dalam kategori yang datanya telah diubah ke dalam persentase (%), rumus yang digunakan adalah sebagai berikut ini:

$$\% \text{ juring} = \frac{\text{Frekuensi}}{\text{Jumlah Data}} \times 100 \%$$

Contoh diagram batang sebagai berikut:

Berikut adalah data buah Favorit siswa kelas IX. Mangga, Apel, Semangka, Semangka, Mangga, Apel, Semangka, Semangka, Semangka, Mangga Semangka, Mangga, Apel, Semangka, Apel, Semangka, Semangka, Mangga, Apel, Semangka Dari informasi di atas, sajikan data dengan menggunakan diagram batang.

Penyelesaian : Tabel distribusi frekuensi

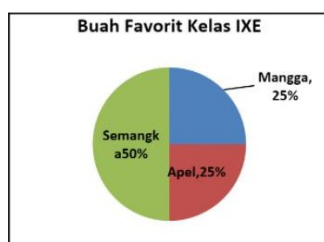
Buah	Frekuensi (f)
Mangga	5
Apel	5
Semangka	10
Jumlah	20

Sehingga , Mangga = $\frac{5}{20} \times 100\% = 25\%$

Apel = $\frac{5}{20} \times 100\% = 25\%$

Semangka = $\frac{10}{20} \times 100\% = 50\%$

Maka gambar diagram lingkaran untuk % adalah sebagai berikut. :



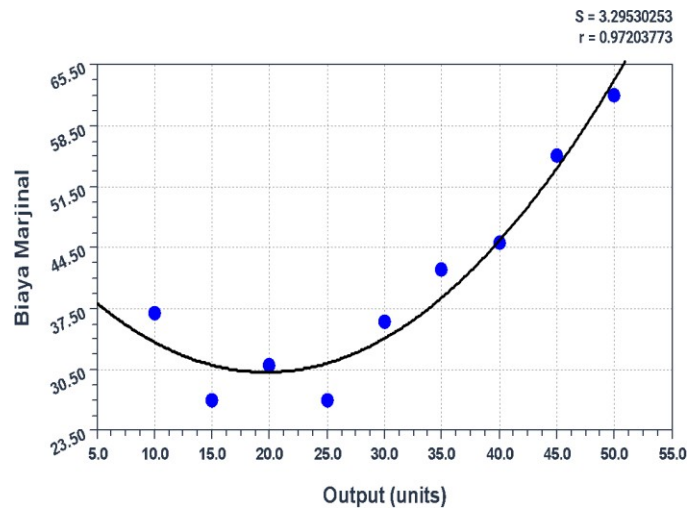
d. Grafik Kurva

Kurva sebenarnya sama saja dengan diagram garis kecuali tujuannya adalah untuk melukiskan suatu hubungan yang kontinu. Umumnya kurva digunakan untuk data hasil eksperimen dalam upaya untuk menggambarkan hubungan antara dua deret data. Kurva dua dimensi juga terdiri dari sumbu tegak dan sumbu mendatar. Setelah data diplot, kemudian dibuat kurva yang mendekati titik-titik pada grafik. Biasanya untuk menggambarkan kurva ini digunakan pendekatan melalui persamaan matematis. Lihat contoh berikut ini. Biaya Marginal dan Output dari suatu proses produksi.

Tabel Biaya Marginal dan output dari suatu proses produksi yaitu:

Biaya Marginal	Output Unit
37	10
27	15
31	20
27	25
36	30
42	35
45	40
55	45
62	50

Setelah data diplot, kemudian dicari persamaan kurva yang paling mendekati titik-titik kedua deret data di atas. Dengan menggunakan perangkat lunak komputer yang tersedia di pasaran, kurva ini bisa dengan mudah diperoleh seperti yang disajikan dalam Gambar berikut



Kurva Biaya Marjinal dan Output suatu proses produksi Persamaan : $y = 44,3476 - 1,4381x + 0,0366x^2$. Garis kurva dalam Gambar diatas diperoleh dengan memasukkan nilai-nilai x (output) ke dalam persamaan polinomial derajat dua seperti yang ditunjukkan pada keterangan gambar di atas. Semakin cocok persamaan yang dibangun maka semakin banyak titik-titik yang mendekati garis kurva.

4. Penyajian data dalam bentuk piktogram

Diagram gambar atau piktogram adalah diagram dimana datanya disajikan dalam bentuk gambar atau lukisan untuk mewakili benda yang menampilkan banyak benda sesungguhnya. Untuk penyajian data yang lain tentu saja kita bisa menggunakan gambar-gambar yang lebih menarik seperti gambar mobil, pohon, uang, dan sebagainya. Gambar yang digunakan disesuaikan dengan objek yang dideskripsikan yang digunakan untuk mewakili sejumlah objek.

Cara penyajian data menggunakan piktogram memiliki kelebihan tersendiri misalnya ada gambar orang yang hanya tampak setengahnya saja, itu digunakan untuk mewakili jumlah siswa yang hanya 10 orang. Namun tentu akan sulit untuk menggambarkan jumlah siswa yang hanya sepertiga atau seperempat dari jumlah yang bisa diwakilkan dengan gambar tersebut. Oleh karenanya, penyajian data dengan menggunakan piktogram agak jarang digunakan.

Perhatikan contoh berikut. Jajanan paling sering dibeli siswa.

Bentuk Tabel :

No	Jajanan	Frekuensi
1	Es Sirop	42
2	Batagor	38
3	Kue Pancong	24
4	Cilok	44
5	Somay	16

Sekarang jika satu gambar mewakili 2 buah, coba kamu sajikan data tersebut dalam bentuk piktogram!

Berdasarkan tabel diatas maka bentuk piktogram dari data diatas adalah sebagai berikut:

Jajanan Paling Sering Dibeli Siswa

Es Sirop	
Batagor	
Kue Pancong	
Cilok	
Siomay	

BAB V

DISTRIBUSI FREKUENSI

A. Pengertian Distribusi Frekuensi

Jika diartikan secara bahasa, distribusi frekuensi merupakan proses penyaluran frekuensi. Dalam ilmu statistika, frekuensi adalah berapa kali suatu variabel dilambangkan menggunakan angka atau bilangan. Kondisi yang dilakukan secara berulang kali dalam deret angka tersebut, frekuensi juga diartikan sebagai kekerapan, keseringan hingga jarang-jarang.

Distribusi frekuensi adalah kondisi yang menggambarkan adanya frekuensi mulai dari gejala yang muncul berupa angka, kemudian tersalur, terbagi hingga terpancar. Penyajian angka dalam penunjukkan frekuensi ini bisa berupa table, grafik dan gambar. Hingga kemudian muncul istilah yang paling sering disebut, yakni tabel distribusi frekuensi.

B. Jenis jenis Distribusi Frekuensi

Distribusi frekuensi memiliki tiga jenis yang berbeda berdasarkan kriterianya, yaitu:

1. Distribusi Frekuensi Biasa

Distribusi ini berisikan jumlah frekuensi dari setiap kelompok data, Dalam hal ini dibedakan kembali menjadi dua, yakni sebagai berikut:

- a. Distribusi frekuensi kategori adalah dengan mengelompokkan datanya disusun berbentuk kata-kata (kualitatif).
- b. Distribusi frekuensi numeric adalah distribusi penyatuan kelas-kelasnya (disusun secara interval) didasarkan pada angka-angka.

1. Distribusi Frekuensi Relatif

Distribusi frekuensi relative berisikan nilai-nilai hasil bagi antara frekuensi kelas dan jumlah pengamatan, serta menyatakan proporsi data yang berada pada suatu kelas interval. Distribusi jenis ini dalam suatu kelas didapatkan dengan cara membagi frekuensi dengan total data yang ada dari pengamatan atau observasi.

2. Distribusi Frekuensi Kumulatif

Distribusi frekuensi kumulatif berisikan frekuensi yang telah dijumlahkan. Terdapat dua macam distribusi frekuensi kumulatif yaitu distribusi frekuensi kumulatif kurang dari ($<$) dan distribusi frekuensi lebih dari ($>$).

Dalam melakukan studi statistic, peneliti harus mengumpulkan data untuk variabel tertentu yang diteliti. Untuk menggambarkan situasi dan membuat kesimpulan tentang suatu peristiwa, peneliti harus mengatur data dengan cara yang benar. Metode peorganisasian data yang paling mudah adalah dengan membuat distribusi frekuensi. Data yang didapat dalam jumlah banyak akan lebih muda dibaca saat data tersebut sudah dikelompokkan. Contoh studi kasus: sebuah peneliti ingin meneliti tinggi badan dari 30 siswa kelas VII-A SMA N 4 BARRU. Langkah pertama, peneliti harus mendapatkan data tinggi badan siswa tersebut. Berikut ini merupakan data tinggi badan dari 30 siswa kelas VII-A SMA N 4 BARRU.

175	177	170	179	183	180	184	185	179	181
180	172	181	178	180	175	179	185	186	176
172	180	177	175	175	185	179	182	187	179

C. Langkah-langkah Distribusi Frekuensi

1. Mengurutkan Data

Data tinggi badan dari 30 siswa kelas VII-A SMA N 4 BARRU diurutkan dari data terkecil hingga terbesar. Berikut merupakan hasil pengurutan data.

No	Tinggi Badan	No	Tinggi Badan	No	Tinggi Badan
1	170	11	178	21	181
2	172	12	179	22	181
3	172	13	179	23	182
4	175	14	179	24	183
5	175	15	179	25	184

6	175	16	179	26	185
7	175	17	180	27	185
8	176	18	180	28	185
9	177	19	180	29	189
10	177	20	180	30	187

2. Menentukan range atau jangkauan dari data tersebut

Range atau jangkauan merupakan ukuran penyebaran atau ukuran disperse dari data. Jangkauan adalah selisih nilai terbesar dan terkecil dari data. Jangkauan menunjukkan seberapa terbesarnya nilai-nilai dalam suatu deret. Jika jangkauannya merupakan angka yang terbesar, maka nilai-nilai dalam deret tersebut dekat satu sama lain. Berikut merupakan perhitungan range pada studi kasus ini.

$$\text{Range} = \text{Data Terbesar} - \text{Data Terkecil}$$

$$\text{Range} = 187 - 170 = 17$$

3. Menentukan Jumlah Kelas

Kelas-kelas adalah kelompok nilai data atau variabel dari suatu data acak. Dalam menentukan jumlah kelas menggunakan Aturan Sturgess, yakni aturan dalam statistika yang diturunkan dari distribusi binomial digunakan untuk menentukan banyaknya kelas pada distribusi frekuensi data berkelompok, dengan rumus:

$$k = 1 + 3,3 \log n$$

Dimana:

k = jumlah kelas

n = jumlah data

Berikut merupakan perhitungan jumlah kelas pada studi kasus ini:

$$k = 1 + 3,3 \log 30$$

$$k = 5,874 = 6$$

4. Menentukan panjang interval kelas

Panjang interval kelas atau luas kelas adalah jarak antara tepi atas kelas dan tepi bawah kelas. Berikut merupakan rumus dalam menentukan panjang interval kelas:

$$C = R/k$$

Dimana:

C = lebar kelas

R = range

k = jumlah kelas

Berikut merupakan perhitungan panjang interval kelas pada studi kasus ini:

$$C = 17/5,874$$

$$C = 2,894 = 3$$

Sehingga didapatkan interval kelas,

Tinggi Badan
170-172
173-175
176-178
179-181
182-184
185-187

5. Menentukan tepi bawah dan tepi atas kelas

Dalam menentukan tepi bawah dan tepi atas kelas, dilakukan dengan mengurangi 0,5 pada batas kelas bawah dan menambahkan 0,5 pada batas kelas atas. Prinsip dasarnya adalah batas kelas harus memiliki nilai tempat desimal yang sama dengan data, tetapi tepi bawah dan tepi atas kelas harus memiliki satu tambahan nilai tempat desimal dan berakhir di 5. Contoh pada studi kasus ini, yaitu:

$$\text{Tepi bawah} = \text{kelas bawah} - 0,5 = 170 - 0,5 = 169,5$$

$$\text{Tepi atas} = \text{kelas atas} + 0,5 = 172 + 0,5 = 172,5$$

Tinggi Badan	Tepi Kelas
170-172	169,5-172,5
173-175	172,5-175,5
176-178	175,5-178,5
179-181	178,5-181,5
182-184	181,5-184,5
185-187	184,5-187,5

6. Menentukan frekuensi dari setiap kelas

Frekuensi kelas adalah banyaknya data yang termasuk ke dalam kelas tertentu dari data acak. Berikut merupakan frekuensi dari setiap yang didapatkan:

Tinggi Badan	Tepi Kelas	Frekuensi
170-172	169,5-172,5	3
173-175	172,5-175,5	4
176-178	175,5-178,5	4
179-181	178,5-181,5	11

Dari pengolahan data table distribusi frekuensi 30 siswa kelas VII-A, didapatkan range atau jangkauan sebesar 17, jumlah kelas sebesar 6, dan panjang interval kelas sebesar 3. Pada batas kelas 169,5-172,5 didapatkan frekuensi sebanyak 3, batas 172,5-175,5 sebanyak 4, batas 175,5-178,5 sebanyak 4, batas 178,5-181,5 sebanyak 11, batas 181,5-184,5 sebanyak 3, dan batas 184,5-187,5 sebanyak 5.

BAB VI

UKURAN PEMUSATAN DATA

A. Pengertian Mean

Mean adalah perhitungan jumlah nilai data dibagi dengan banyaknya observasi atau banyaknya individu. Nama lain dari mean adalah rata-rata. Oleh sebab itu, pengertian lain dari mean adalah nilai tengah atau rata-rata dari suatu gugus data pengamatan.

Mean dibedakan menjadi dua bentuk. Jenis mean adalah rata-rata sampel yang dilambangkan dengan $\bar{x}$ dan rata-rata populasi yang biasa dilambangkan dengan μ . Kedua jenis tersebut yang membedakannya adalah rumusnya.

Fungsi mean adalah dapat menghitung rata-rata dari data yang tidak berkelompok maupun data yang berkelompok. Data yang berkelompok adalah data yang disajikan dalam tabel frekuensi dengan ciri-cirinya terbagi menjadi kelompok-kelompok kelas tertentu.

Adapun yang dimaksud dengan data tidak berkelompok adalah data yang tidak disajikan dengan tabel frekuensi serta tidak dibedakan menjadi kelas tertentu. Dengan demikian, mean dari data tidak berkelompok hanya penjumlahan dari tiap data dibagi dengan jumlah pengamatan.

Untuk data tidak berkelompok, rumus mean adalah sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n x_i}{N}$$

Mean dari populasi

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Mean dari sampel

Keterangan:

x_i = data pertama

n = jumlah data

Berdasarkan rumus tersebut, langkah yang dilakukan untuk menghitung data kelompok mean adalah dengan cara mengalikan nilai tengah kelas dengan frekuensi data untuk semua kelas.

Setelah itu, jumlahkan hasil sebelumnya yaitu dari perkalian untuk semua kelas, lalu barulah membaginya dengan jumlah data (n).

B. Pengertian Median

Median adalah nilai tengah dari data yang sudah diurutkan dan disusun secara teratur berdasarkan besar kecilnya data. Oleh sebab itu, perhitungan median dilakukan setelah Anda menyusun data mulai dari yang terkecil hingga terbesar, kemudian barulah ditentukan nilai tengahnya.

Penentuan median ini menunjukkan bahwa ada 50% dari jumlah data atau frekuensi distribusi yang letaknya dibawah median serta 50%-nya yang lain di atas median.

Rumus median untuk data tidak berkelompok terbagi menjadi dua yaitu dalam bentuk jumlah data yang genap dan pada jumlah data yang ganjil.

Sebelumnya, langkah-langkah untuk menentukan median adalah mengurutkan data dari terkecil sampai terbesar. Setelah itu, hitung jumlah data tersebut. Apabila n (jumlah data) ganjil dengan data yang tidak terlalu banyak dapat ditentukan langsung dengan melihat data yang ada di posisi tengah itulah mediannya.

Apabila jumlah data genap, maka sebagai mediannya diambil dari rata-rata dua data yang ada di tengah. Adapun rumus mudah yang dipakai untuk menentukan posisi median untuk data ganjil dan genap tidak berkelompok (tunggal) yaitu:

Rumus:

a. Distribusi Frekwensi Genap

$$Me = \frac{X_{N/2} + X_{(N/2) + 1}}{2}$$

b. Distribusi Frekwensi Ganjil

$$Me = X_{(N+1)/2}$$

Keterangan:

Me = Nilai Median
Xi = Data ke i
N = Banyaknya data

Untuk data berkelompok, penentuan posisi median adalah dengan menggunakan rumus:

Nilai median untuk data berkelompok berdasarkan rumus di atas menunjukkan bahwa langkah pertama yang diambil adalah menentukan kelas mana letak median berdasarkan rumus: $(n+1)/2$, setelah itu menghitung besar nilai median seperti pada rumus di atas.

C. Pengertian Modus

Modus adalah nilai yang sering muncul. Maksudnya adalah dalam kumpulan data, nilai dengan frekuensi terbanyak disebut dengan modus. Ukuran modus digunakan untuk mengetahui tingkat seringnya terjadi dalam suatu peristiwa.

Untuk modus yang tidak berkelompok, cara menemukan modus cukup mudah yaitu dengan menghitung frekuensi dari masing-masing data, kemudian dicari nilai yang mempunyai frekuensi terbanyak atau yang sering muncul.

D. Contoh Soal Data Tunggal

Nilai matematika yang diperoleh siswa kelas VIII C SMP Bumi Pertiwi disajikan dalam data berikut:

8,7,6,8,9,9,8,7,6,9,7,8,8,7,9,8,8,6,8,7

Berdasarkan data di atas, carilah mean, median, dan modus!

- Mean

$$\begin{aligned}\bar{x} &= \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \\ \bar{x} &= \frac{8+7+6+8+9+9+8+7+6+9+7+8+8+7+9+8+8+6+8+7}{20} \\ \bar{x} &= 7,65\end{aligned}$$

- Median

Untuk mendapatkan median, pada contoh soal mean, median, modus data tunggal serta data kelompok di atas, urutkan dari bilangan terkecil hingga terbesar.

6,6,6,7,7,7,7,7,8,8,8,8,8,8,8,8,9,9,9,9. Karena jumlah datanya 20 (genap), maka gunakan rumus:

$$\begin{aligned}Me &= \frac{1}{2} \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right) \\ Me &= \frac{1}{2} \left(x_{\left(\frac{20}{2}\right)} + x_{\left(\frac{20}{2}+1\right)} \right) = \frac{1}{2} (x_{10} + x_{11})\end{aligned}$$

Karena x_{10} adalah 8 dan x_{11} adalah 8, maka:

$$Me = \frac{1}{2}(8 + 8) = 8$$

- Modus

Pada data yang disajikan, nilai 6 muncul 3 kali, nilai 7 muncul 5 kali, nilai 8 muncul 8 kali, dan nilai 9 muncul 4 kali. Jadi, modus pada soal di atas adalah 8.

E. Contoh Soal Data Kelompok

Diketahui data table sebagai berikut:

Nilai	Frekuensi
61-70	8
71-80	17
81-90	20
91-100	15

Hitung mean, median, dan modus!

- Mean

Nilai	Frekuensi	M_i	$M_i f_i$
61-70	8	65,5	524
71-80	17	75,5	1.283,5
81-90	20	85,5	1.710
91-100	15	95,5	1.432,5

Maka, rata-ratanya yaitu:

$$\begin{aligned}\bar{x} &= \frac{4.950}{60} \\ &= 82,5\end{aligned}$$

- Median

Sebelum menghitung dengan rumus, buat tabel frekuensi kumulatif untuk mengetahui kelas median.

Nilai	Frekuensi	f_k
61-70	8	8
71-80	17	25
81-90	20	45
90-100	15	60

Karena data berjumlah 60, maka median berada di antara data ke-30 dan ke-31, yang mana keduanya berada di kelas interval ke-3 (81-90).

Maka interval ke-3 inilah yang disebut kelas median. Sekarang masukkan rumusnya:

$$\begin{aligned}
 Me &= Tb + \left(\frac{\frac{n}{2} - f_k}{f_i} \right) p \\
 Me &= 80,5 + \left(\frac{\frac{60}{2} - 25}{20} \right) 10 \\
 &= 80,5 + \left(\frac{5}{20} \right) 10 = 83
 \end{aligned}$$

- Modus

Modus pada data yang disajikan terletak di kelas interval ke-3 karena frekuensinya paling banyak. Maka, modusnya adalah:

$$\begin{aligned}
 Mo &= T_b + \left(\frac{b_1}{b_1 + b_2} \right) p \\
 &= 80,5 + \left(\frac{20 - 17}{(20 - 17) + (20 - 15)} \right) 10 \\
 &= 80,5 + \left(\frac{3}{8} \right) 10 \\
 &= 84,25
 \end{aligned}$$

BAB VII

UKURAN LOKASI

A. Pengertian Ukuran Lokasi

Ukuran lokasi atau ukuran letak dimaksudkan sebagai besaran atau ukuran untuk mendapatkan gambaran yang lebih jelas berdasarkan letak data dari sekumpulan data yang mempunyai ukuran ini sangat berarti dalam rangka melakukan analisis data.

Menurut Andi (2007:69), ukuran lokasi adalah besaran atau ukuran untuk mendapatkan gambaran yang lebih jelas berdasarkan letak dari sekumpulan data yang mempunyai. Ukuran lokasi dalam statistika digunakan untuk menentukan posisi relatif nilai suatu data terhadap sekumpulan data.

Ukuran letak suatu rangkaian data adalah ukuran yang didasarkan pada letak ukuran tersebut. Ukuran letak merupakan ukuran untuk melihat dimana letak salah satu data dari sekelompok banyak data yang ada. Yang termasuk ukuran letak antara lain adalah kuartil, desil, dan persentil. Dalam penentu ketiganya adalah mengurutkan distribusi data dari yang terkecil sampai terbesar.

B. Ukuran Letak Data (Ukuran Lokasi)

Dalam ukuran lokasi ada tiga ukuran letak data yang harus diketahui, yaitu: kuartil, desil, dan persentil yang diistilahkan dengan fraktil yang semuanya itu merupakan pengembangan dari Median yang merupakan bagian dari pemusatan data.

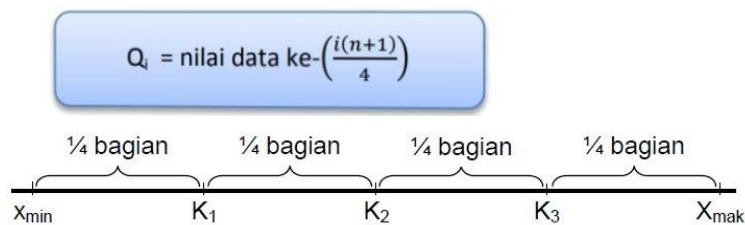
- **Kuartil**

Jika median dapat dikatakan sebagai ukuran perdua, maka kuartil disebut dengan ukuran Perempatan. Kuartil dalam kehidupan sehari-hari disebut dengan quarter dan disimbolkan dengan K atau Q. Kuartil adalah ukuran letak yang membagi distribusi data menjadi empat bagian sama besar, yaitu masing-masing $\frac{1}{4}n$. Jadi akan dijumpai tiga buah kuartil, yaitu kuartil pertama atau kuartil bawah (K1), kuartil kedua atau kuartil tengah (K2) yang merupakan median, dan kuartil ketiga atau kuartil atas (K3). Menentukan kuartil pertama, kedua, dan ketiga sama seperti menentukan median, data harus diurutkan dari nilai terkecil ke nilai terbesar atau sebaliknya dari nilai terbesar ke nilai terkecil.

Kuartil pertama atau bawah (K1) adalah suatu nilai yang membatasi 25% distribusi bagian bawah dan 75% distribusi bagian atas. Kuartil kedua atau tengah (K2) adalah nilai yang membatasi 50% distribusi bagian bawah dan 50% distribusi bagian atas. Dalam hal ini kuartil kedua dapat diidentikkan dengan median (Me). Kuartil ketiga atau atas (K3) adalah nilai yang membatasi 75% distribusi bagian bawah dan 25% distribusi bagian atas. Jika data diurutkan dari nilai terkecil ke nilai terbesar, maka K1 akan terletak di sebelah kiri K2 dan K3 terletak di sebelah kanan K2.

Sedangkan jika data diurutkan dari nilai terbesar ke terkecil, maka K1 akan terletak di sebelah kanan K2 dan K3 akan terletak di sebelah kiri K2. Dengan posisi K2 selalu di tengah-tengah K1 dan K3. Secara visualisasi dapat diilustrasikan sebagai berikut.

Sebelum menentukan nilai kuartil berapa dari segugusan data tunggal maupun data berkelompok, terlebih dahulu yang harus dicari adalah letak kuartil. Adapun rumus yang digunakan untuk menentukan letak kuartil adalah sebagai berikut.



Keterangan :

Q_i = kuartil ke- i

n = banyaknya data

$$Q_i = x_n + d \cdot (x_{n+1} - x_n)$$

Keterangan:

Q_i = kuartil ke-i, dengan $i = 1, 2$, atau 3

x_n = nilai data ke-n

x_{n+1} = nilai data ke-($n + 1$)

d = nilai desimal dari $\left(\frac{i(n+1)}{4}\right)$

Misalnya Q_i (kuartil) terletak di antara nilai data ke-n dan nilai data ke-($n+1$) atau tidak bulat, maka kita bisa menggunakan rumus:

Contoh Soal:

Tentukan Q_1 , Q_2 dan Q_3 dari data : 7,3,8,5,9,4,8,3,10,2,7,6,8,7,2,6,9 !

Penyelesaian:

Data terurut : 2,2,3,3,4,5,6,6,7,7,7,8,8,8,9,9,10

$$\text{Letak } Q_1 = \frac{1(17 + 1)}{4} = \frac{18}{4} = 4,5$$

$$\begin{aligned} Q_1 &= X_4 + 0,5(X_5 + X_4) \\ &= 3 + 0,5(4 - 3) \\ &= 3,5 \end{aligned}$$

$$n = 17$$

- Desil

Desil merupakan nilai yang membagi data menjadi sepuluh bagian sama besar. Desil sering dilambangkan dengan D.

Sekumpulan data yang sudah diurutkan dari terkecil hingga terbesar dapat dibagi menjadi sepuluh bagian. Ukuran lokasi yang membagi sekumpulan data menjadi sepuluh bagian yang sama dinamakan desil. Masing-masing bagian mengandung 10% data. Dengan demikian, dari suatu kumpulan data dapat diperoleh 9 buah desil, yakni $D_1, D_2, \dots, D_9$.

$$\text{Letak } Q_2 = \frac{2(17+1)}{4} = \frac{36}{4} = 9$$

$$Q_2 = X_9 = 7$$

$$\text{Letak } Q_3 = \frac{3(17+1)}{4} = \frac{54}{4} = 13,5$$

$$Q_3 = X_{13} + 0,5(X_{14} - X_{13})$$

$$= 8 + 0,5(8 - 8)$$

$$= 8$$

Nilai desil ke- i suatu data tunggal dapat ditentukan dengan rumus berikut.

Keterangan:

D_i = desil ke- i , dengan $i = 1, 2, 3, 4, \dots, 9$

$$D_i = \text{nilai data ke-}\left(\frac{i(n+1)}{10}\right)$$

n = banyak data

$$\left(\frac{i(n+1)}{10}\right)$$

Jika nilai bukan bilangan bulat, nilai ditentukan dengan menggunakan interpolasi linear seperti berikut.

$$D_i = x_n + d \cdot (x_{n+1} - x_n)$$

Misalkan D_i terletak di antara nilai data ke- n dan nilai data ke- $(n+1)$, maka:

Contoh Soal:

Keterangan:

D_i = desil ke- i , dengan $i = 1, 2, 3, 4, \dots, 9$

x_n = nilai data ke- n

x_{n+1} = nilai data ke- $(n+1)$

d = nilai desimal dari $\left(\frac{i(n+1)}{10}\right)$

- Persentil

Dari data 14, 12, 8, 6, 15, 10, 2, 9, 4, 3, tentukan desil ke-2.

Pembahasan

Data diurutkan menjadi 2, 3, 4, 6, 8, 9, 10, 12, 14, 15. Dan banyak data $n = 10$.

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

$$D_2 = \text{data ke} - \frac{2(10+1)}{10}$$

$$D_2 = \text{data ke} - \frac{22}{10} = 2,2$$

$$D_2 = x_2 + 0,2(x_3 - x_2)$$

$$D_2 = 3 + 0,2(4 - 3)$$

$$D_2 = 3 + 0,2 \cdot 1 = 3,2$$

Sekumpulan nilai yang sudah diurutkan dari yang terkecil hingga terbesar dapat dibagi menjadi seratus bagian. Ukuran lokasi atau letak yang membagi sekumpulan data menjadi seratus bagian yang sama dinamakan persentil. Masing-masing bagian mengandung 1% data. Pada sekumpulan data mempunyai 99 buah persentil, yaitu $P_1, P_2, \dots, P_{99}$.

Persentil pertama (P_1) adalah sebuah nilai yang menyatakan 1% dari keseluruhan data atau observasi nilainya lebih kecil dari P_1 dan 99% nya lagi nilainya lebih besar dari nilai P_1 .

Persentil kedua (P_2) adalah sebuah nilai yang menyatakan bahwa 2% dari keseluruhan data atau observasi nilainya lebih kecil dari P_2 dan 98% nya lagi nilainya lebih besar dari nilai P_2 .

Persentil kelima puluh (P_{50}) adalah sebuah nilai yang menyatakan bahwa 50% dari keseluruhan data atau observasi nilainya lebih kecil dari P_{50} dan 50% nya lagi nilainya lebih besar dari nilai P_{50} . Dapat dikatakan P_{50} sama dengan Median.

Persentil kedua (P99) adalah sebuah nilai yang menyatakan bahwa 99% dari keseluruhan data atau observasi nilainya lebih kecil dari P99 dan 1% nya lagi nilainya lebih besar dari nilai P99.

Persentil ke- i (P_i) adalah sebuah nilai yang menyatakan bahwa $1i\%$ dari keseluruhan data atau observasi nilainya lebih kecil dari P_i dan $(100\% - 1i\%)$ nya lagi nilainya lebih besar dari nilai P_i .

$$P_i = \text{nilai data ke-} \left(\frac{i(n+1)}{100} \right)$$

Nilai persentil ke- i suatu data tunggal dapat ditentukan dengan rumus berikut.

Keterangan:

P_i = persentil ke- i , dengan $i = 1, 2, 3, 4, \dots, 99$

n = banyak data

$$\left(\frac{i(n+1)}{100} \right)$$

Jika nilai bukan bilangan bulat, nilai P_i ditentukan dengan menggunakan interpolasi linear seperti berikut.

$$P_i = x_n + d \cdot (x_{n+1} - x_n)$$

Keterangan:

P_i = persentil ke- i , dengan $i = 1, 2, 3, 4, \dots, 99$

x_n = nilai data ke- n

x_{n+1} = nilai data ke- $(n + 1)$

d = nilai desimal dari $\left(\frac{i(n+1)}{100} \right)$

Misalkan P_i terletak di antara nilai data ke- n dan nilai data ke- $(n + 1)$, maka:

Contoh Soal:

Dari data 6, 7, 3, 4, 8, 3, 7, 6, 10, 15, tentukan persentil ke – 30.

Pembahasan

Data diurutkan menjadi 3, 3, 4, 6, 6, 7, 7, 8, 10, 15.
Dan banyak data $n = 10$.

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

$$P_{30} = \text{data ke} - \frac{30(10+1)}{100}$$

$$P_{30} = \text{data ke} - \frac{330}{100} = 3,3$$

$$P_{30} = x_3 + 0,3(x_4 - x_3)$$

$$P_{30} = 4 + 0,3(6 - 4)$$

$$P_{30} = 4 + 0,3 \cdot 2 = 4 + 0,6 = 4,6$$

Dari data 6, 7, 3, 4, 8, 3, 7, 6, 10, 15, tentukan persentil ke – 30.

Pembahasan

Data diurutkan menjadi 3, 3, 4, 6, 6, 7, 7, 8, 10, 15.
Dan banyak data $n = 10$.

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

$$P_{30} = \text{data ke} - \frac{30(10+1)}{100}$$

$$P_{30} = \text{data ke} - \frac{330}{100} = 3,3$$

$$P_{30} = x_3 + 0,3(x_4 - x_3)$$

$$P_{30} = 4 + 0,3(6 - 4)$$

$$P_{30} = 4 + 0,3 \cdot 2 = 4 + 0,6 = 4,6$$

C. Macam-macam Ukuran Lokasi

Terdapat pula macam-macam ukuran lokasi diantaranya:

1. Kuartil

Kuartil adalah suatu nilai yang bisa membagi data menjadi empat bagian sama besar. Syarat untuk mendapatkan kuartil ini adalah data harus diurutkan terlebih dahulu.

Data berkelompok adalah kumpulan data yang ditulis dalam bentuk interval. Kuartil data berkelompok adalah suatu nilai yang membagi data-data interval menjadi empat bagian sama besar.

Rumus Kuartil Data Berkelompok

Ada beberapa elemen yang harus di tentukan sebelumnya, seperti letak kuartil yang dicari, frekuensi kumulatif data, tepi bawah kuartil yang dicari, dan interval kelas. Adapun langkah menentukan kuartil data berkelompok adalah sebagai berikut.

a. Menentukan letak kuartil

$$\text{letak } Q_i = \frac{i}{4}n$$

Keterangan :

Q_i = Kuartil ke-i

i = Letak kuartil ke-i

n = Banyak data

b. Menentukan kuartil yang dimaksud dengan rumus berikut.

$$Q_i = Tb_i + p \left(\frac{n_i - f_k}{f} \right)$$

Keterangan :

Q_i = Kuartil ke-i

Tb_i = Tepi bawah kelas kuartil ke-i

p = Interval kelas

f_k = Frekuensi kumulatif sebelum kuartil ke-i

f = Frekuensi kuartil ke-i

n = Banyaknya data

i = Posisi kuartil yang dicari (1-3)

2. Desil

Desil adalah suatu istilah yang menunjukkan pembagian data menjadi sepuluh bagian sama besar. Sebelum menentukan desil, data harus diurutkan terlebih dahulu. Ada sembilan desil yang nantinya bisa membagi data menjadi sepuluh bagian sama besar.

Rumus Desil Data Berkelompok

$$D_i = Tb + \left(\frac{\frac{i}{10} \times n - f_{kk_{D_i}}}{f_{D_i}} \right) \times l$$

Keterangan :

D_i = Nilai desil ke-i

Tb = Tepi bawah kelas desil ke-i

i = 1,2,3,...,9

n = Jumlah frekuensi/banyak data

$f_{kk_{D_i}}$ = Frekuensi kumulatif kurang dari sebelum kelas desil ke-i

f_{D_i} = Frekuensi kelas letak desil ke-i

l = Panjang Kelas

3. Persentil

Persentil adalah nilai atau angka yang mewakili posisi persentase pada suatu kumpulan berbagai jenis data penelitian. Adapun dalam statistik dan statistika istilah persentil kerap kali dipergunakan untuk memahami dan menginterpretasikan data, yang menunjukkan nilai di bawah yang persentase tertentu dari suatu kumpulan data. Oleh karena itulah persentil ke- i dari sekumpulan data adalah nilai di mana i persen data berada di bawahnya.

Rumus Persentil Data Berkelompok

$$P_i = Tb + \left(\frac{\frac{i}{100}n - f_k}{f_i} \right) p$$

Keterangan :

n = Jumlah seluruh frekuensi.

Tb = Tepi bawah kelas persentil.

i = Bilangan bulat yang kurang dari 100 (1, 2, 3, ..., 99).

p = Panjang kelas interval.

f_i = Frekuensi kelas persentil.

f_k = Jumlah frekuensi sebelum kelas persentil.

D. Contoh Soal

1. Desil

Dalam rangka memperingati Hari Pendidikan Nasional, Dinas Pendidikan kota Makassar mengadakan Seminar Pendidikan pada 60 orang dengan rentang usia yang berbeda-beda seperti berikut.

Rentang Usia (th)	Jumlah Peserta
16-20	4
21-25	10
26-30	6

31-35	15
36-40	8
41-45	14
46-50	3

Tentukan kuartil ke-3 dari data di atas!

Jawab

Mula-mula, tentukan dahulu frekuensi kumulatif pada tabel.

Rentang Usia (th)	Jumlah Peserta	Frkuensi Kumulatif (fk)
16-20	4	4
21-25	10	14
26-30	6	20
31-35	15	35
36-40	8	43
41-45	14	57
46-50	3	60

Banyaknya data (n) = 60

Selanjutnya, tentukan letak kuartil ke-3

$$\text{letak } Q_i = \frac{i}{4} n$$

$$\text{letak } Q_3 = \frac{3}{4} (60)$$

$$\text{letak } Q_3 = 45$$

Karena letak kuartilnya 45, maka kuartil tersebut berada di rentang usia 41-45

$$Tb_3 = 41 - 0,5 = 40,5$$

$$p = \text{panjang data} = 5$$

Substitusikan nilai elemen-elemen yang diketahui pada persamaan berikut.

$$Q_i = Tb_i + p \left(\frac{n_i - f_k}{f} \right)$$

$$Q_3 = 40,5 + 5 \left(\frac{\frac{3(60)}{4} - 43}{14} \right)$$

$$= 40,5 + 0,714$$

$$= 41,21$$

Jadi, kuartil ke-3 dari data tersebut adalah 41,21.

2. Desil

Perhatikan tabel data kelompok tersebut!

Berat Badan (kg)	Frekuensi
41-45	3
46-50	6
51-55	16
56-60	8
61-65	7

Berdasarkan data yang disajikan pada tabel diatas, tentukan desil ke-5

Jawab

$$\frac{i \times n}{10} = \frac{5 \times 40}{10} = 20$$

$$Tb = 51 - 0,5 = 50,5$$

$$l = 5$$

$$fkk_{Di} = 6 + 3 = 9$$

$$f_{Di} = 16$$

$$D_5 = Tb + l \left(\frac{\frac{i}{10} \times n - f_{kk_{Di}}}{f_{Di}} \right)$$

$$D_5 = 50,5 + 5 \left(\frac{\frac{5}{10} \times 40 - 9}{16} \right)$$

$$D_5 = 50,5 + 5 \left(\frac{20 - 9}{16} \right)$$

$$D_5 = 50,5 + 5 \times 0,6875$$

$$D_5 = 50,5 + 3,4375$$

$$= 53,9375$$

3. Persentil

Tentukan letak persentil kelompok ke-75 dari kelompok data seperti tabel dibawah ini.

X	F
41-45	3
46-50	6
51-55	16
56-60	8
61-65	7

Jawab

X	F	F Kumulatif
41-45	3	3
46-50	6	9
51-55	16	25

56-60	8	33
61-65	7	40

$$\text{Letak persentil ke-75} = \frac{75}{100}40 = 30$$

Data ke-30 pada tabel dan kelas pada persentil ke-75 = 56-60

$$P_1 = Tb + \left(\frac{\frac{i}{100} n - f_k}{f_i} \right) p$$

$$P_{75} = 50,5 + \left(\frac{\frac{75}{100} 40 - 25}{8} \right) 5$$

$$P_{75} = 50,5 + \left(\frac{30 - 25}{8} \right) 5$$

$$P_{75} = 50,5 + 3,125$$

$$= 53,625.$$

BAB VIII

UKURAN PENYEBARAN DATA

A. Pengertian Ukuran Penyebaran Data

Ukuran penyebaran data adalah suatu ukuran yang menyatakan seberapa besar nilai-nilai data berbeda atau bervariasi dengan nilai ukuran pusatnya atau seberapa besar penyimpangan nilai-nilai data dengan nilai pusatnya.

Ukuran penyebaran data. Ukuran Ini berguna untuk menunjukkan seberapa jauhnya suatu data menyebar dari rata-ratanya. Misalnya kita hendak membandingkan tingkat produktivitas dua perusahaan Tempe. Seumpama kita telah mendapatkan data bahwa kedua perusahaan tersebut memiliki rata-rata produksi 300 Tempe sehari, namun kita tidak dapat disimpulkan bahwa kedua perusahaan tersebut memiliki tingkat produktivitas yang sama karena mungkin saja salah satu perusahaan cenderung lebih homogen, dalam arti jumlah produksi tidak jauh dari kisaran rata-rata sedangkan perusahaan lainnya cenderung heterogen, dalam arti jumlah produk jauh dan sangat beragam dari kisaran rata-rata.

Ukuran penyebaran data adalah suatu ukuran yang menyatakan seberapa besar nilai-nilai data berbeda atau bervariasi dengan nilai ukuran pusatnya atau seberapa besar penyimpangan nilai-nilai data dengan nilai pusatnya. Ukuran penyebaran data adalah ukuran yang menunjukkan seberapa jauh data suatu menyebar dari rata-ratanya.

B. Macam-macam Ukuran Penyebaran Data Tunggal

Terdapat pula macam-macam ukuran penyebaran diantaranya:

1. Jangkauan/rentang (range)

Jangkau dari suatu data didefinisikan sebagai selisih antara data terbesar dengan data terkecil.

$$R = X_{maks} - X_{min}$$

Keterangan :

R = Jangkauan atau range

X_{min} = Nilai atau data terkecil

X_{maks} = Nilai atau data terbesar

2. Hamparan

Hamparan (H) adalah selisih antara kuartil ketiga dengan kuartil pertama.

$$H = Q_3 - Q_1$$

Keterangan:

H = Hamparan

Q_3 = Kuartil ke-3

Q_1 = Kuartil ke-1

3. Simpangan Kuartil (Q_d)

Simpangan kuartil adalah setengah kali panjang hamparan.

$$Q_d = \frac{1}{2}H \text{ atau } Q_d = \frac{1}{2}(Q_3 - Q_1)$$

Keterangan:

Q_d = Simpangan kuartil

H = Hamparan

Q_3 = Kuartil ke-3

Q_1 = Kuartil ke-1

4. Simpangan Rata-rata

Simpangan rata-rata menyatakan jarak rata-rata suatu data terhadap rataannya.

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

Keterangan:

SR = Simpangan rata-rata

n = Banyaknya data

x_i = Nilai data ke-i

$\bar{x}$ = Rataan hitung

5. Ragam/Varians (S^2) dan Simpangan Baku (S)

Ragam menyatakan rata-rata kuadrat jarak suatu data terhadap rataannya.

$$\text{Ragam/Varians: } S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

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

Keterangan:

S^2 = Ragam (varians)

S = Simpangan baku

n = Banyaknya data

x_i = Nilai data ke-i

$\bar{x}$ = Rataan hitung

C. Contoh Soal Ukuran Penyebaran Data Tunggal

1. Diketahui data: 3, 5, 5, 4, 10, 9, 8, 7, 9. Tentukan jangkauan data, hamparan, dan simpangan kuartil dari data tersebut.

Pembahasan:

$$x_{maks} = 10$$

$$x_{min} = 3$$

$$\text{Jangkauan: } J = x_{maks} - x_{min}$$

$$= 10 - 3$$

$$J = 7$$

Data diurutkan menjadi: 3, 4, 5, 5, 7, 8, 9, 9, 10

$$Q_1 = \frac{4+5}{2} = 4,5$$

$$Q_2 = 7$$

$$Q_3 = \frac{9+9}{2} = 9$$

$$\text{Hamparan: } H = Q_3 - Q_1$$

$$= 9 - 4,5$$

$$H = 4,5$$

$$\text{Simpangan Kuartil: } Q_d = \frac{1}{2}H$$

$$= \frac{1}{2}(4,5)$$

$$Q_d = 2,25$$

2. Diketahui data 3, 5, 4, 11, 10, 9, 7. Simpangan rata-rata tersebut adalah.....

Pembahasan:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$= \frac{3+5+4+11+10+9+7}{7}$$

$$= \frac{49}{7}$$

$$\bar{x} = 7$$

Simpangan Rata-rata:

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

$$= \frac{1}{7} (|3 - 7| + |5 - 7| + |4 - 7| + |11 - 7| + |10 - 7| + |9 - 7| + |7 - 7|)$$

$$= \frac{1}{7} (4 + 2 + 3 + 4 + 3 + 2 + 0)$$

$$= \frac{1}{7} 18$$

$$SR = 2,57$$

3. Diketahui data: 4, 5, 3, 7, 9. 8. Ragam dan simpangan baku data tersebut adalah....

Pembahasan:

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n x_i}{n} \\ &= \frac{4+5+3+7+9+8}{6}\end{aligned}$$

$$= \frac{36}{6}$$

$$\bar{x} = 6$$

Ragam/Varians:

$$\begin{aligned}S^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \frac{1}{6} ((4 - 6)^2 + (5 - 6)^2 + (3 - 6)^2 + (7 - 6)^2 + (9 - 6)^2 + (8 - 6)^2) \\ &= \frac{1}{6} (4 + 1 + 9 + 1 + 9 + 4) \\ &= \frac{1}{6} 28\end{aligned}$$

$$S^2 = \frac{14}{3} \approx 4,67$$

Simpangan Baku:

$$\begin{aligned}S &= \sqrt{S^2} \\ &= \sqrt{\frac{14}{3}} \\ &= \frac{\sqrt{14}}{\sqrt{3}} \\ &= \sqrt{14}/\sqrt{3} \times \sqrt{3}/\sqrt{3} \\ &= \sqrt{42}/3 \approx 2,16.\end{aligned}$$

D. Jenis-jenis Ukuran Penyebaran Data Berkelompok

Beberapa jenis ukuran penyebaran data diantaranya sebagai berikut:

1. Jangkauan (Range)

Dengan menggunakan jangkauan maka dapat diketahui gambaran secara kasar tentang variasi suatu distribusi data. Nilai jangkauan ini sangat kasar karena tidak mempertimbangkan nilai-nilai yang lain selain nilai ekstrimnya. (Rohmad dan Supriyanto, 2015:76). Jangkauan data merupakan selisih antara nilai maksimum dan nilai minimum pada suatu data. Jangkauan merupakan selisih data terbesar dan data terkecil. Jangkauan sering dilambangkan dengan R.

Untuk menghitung jangkauan data dalam bentuk data berkelompok ada dua cara yaitu sebagai berikut :

- a. Menghitung nilai Tengah tiap kelas dengan nilai terkecil dan terbesar jangkauan datanya adalah selisih antara nilai tengah pada kelas dengan nilai terbesar dan nilai tengah pada kelas dengan nilai terkecil.

$$R = x_{i(max)} - x_{i(min)}$$

- b. Menghitung batas bawah kelas dengan nilai terendah dan batas atas dengan nilai tertinggi kelas pada data tersebut jangkauan datanya adalah selisih antara batas bawah kelas dengan nilai terendah dan batas atas dengan nilai tertinggi.

$$R = BA_i (max) - BB_i (min)$$

Contoh: Daftar berat badan 100 mahasiswa

Berat Badan (Kg)	Banyak Mahasiswa (f)
60 – 62	5
63 – 65	18
66 – 68	42
69 – 71	27
72 – 74	8

- Hitunglah jangkauan data dari data tersebut!

Cara I :

Kelas dengan nilai tertinggi : 72 – 74

$$x_{i(\max)} = \frac{72 + 74}{2} = 73$$

Kelas dengan nilai terendah : 60 – 62

$$x_{i(\min)} = \frac{60 + 62}{2} = 61$$

Jangkauan datanya:

$$R = x_{i(\max)} - x_{i(\min)} = 73 - 61 = 12$$

Cara II :

Kelas dengan nilai tertinggi : 72 – 74

$$BA_{i(\max)} = 74,5$$

Kelas dengan nilai terendah

$$BB_{i(\min)} = 59,5$$

$$\text{Jangkauan datanya : } R = BA_{i(\max)} - BB_{i(\min)} = 74,5 - 59,5 = 12$$

Di Dalam jangkauan kuartil terdapat dua jenis yaitu :

a). Jangkauan Antar Kuartil/Interkuartil (Jangkauan Kuartil)

Jangkauan antar kuartil sendiri merupakan selisih antara kuartil atas dengan kuartil bawah. Jadi, cara untuk menentukan simpangan kuartil adalah dengan menentukan nilai dari kuartil atas dan juga kuartil bawah. Dari data kita bisa menentukan nilai kuartil baik kuartil ke-1 (Q_1) kuartil ke-2 (Q_2) dan kuartil ke-3 (Q_3), untuk cara menentukan nilai kuartil sudah kita bahas pada pertemuan sebelumnya selanjutnya dari nilai-nilai kuartil tersebut juga berlaku ukuran penyebaran yaitu jangkauan antara kuartil (hamparan) yang kita simbolkan dengan (Q_R).

$$\text{Rumus Jangkaun Antar Kuartil yaitu: } Q_R = Q_3 - Q_1$$

b). Jangkauan Semi Antar Kuartil/Interkuartil (Simpangan Kuartil)

Simpangan kuartil adalah setengah dari jangkauan antara kuartil simpangan kuartil merupakan transformasi dari jangkauan kuartil yang digunakan untuk menjelaskan variabilitas karena semi kuartil lebih terfokus pada pertengahan atau 50% dari distribusi sehingga kondisi kurang dipengaruhi oleh skor yang ekstrem. Jangkauan semi antar kuartil (simpangan kuartil) biasa kita

simbolkan dengan Q_d .

Rumus Simpangan Kuartil yaitu:

$$Q_d = \frac{1}{2}(Q_3 - Q_2)$$

Contoh: tentukan jangkauan antarkuartil dan simpangan kuartil pada data berkelompok berikut:

Nilai	f_i	f_k
31 – 40	1	1
41 – 50	2	3
51 – 60	5	8
61 – 70	15	23
71 – 80	25	48
81 – 90	20	68
91 – 100	12	80

$$\text{Letak } Q_1 = \frac{1(80+1)}{4} = 20,25$$

Kelas kuartil 3 : 61 – 70

$$b = 60,5$$

$$p = 10$$

$$f_k = 8$$

$$f_q = 15$$

$$\text{Letak } Q_3 = \frac{3(80+1)}{4} = 60,75$$

Kelas kuartil 3 : 81 – 90

$$b = 80,5$$

$$p = 10$$

$$f_k = 48$$

$$f_Q = 20$$

Maka,

$$Q_3 = b + p \left(\frac{\frac{3n}{4} - f_k}{f_Q} \right) \rightarrow 80,5 + 10 \left(\frac{60 - 48}{20} \right) = 86,5$$

$$Q_1 = b + P \left(\frac{\frac{1n}{4} - f_k}{f_Q} \right) \rightarrow 60,5 + 10 \left(\frac{20 - 48}{15} \right) = 68,5$$

Sehingga,

$$Q_R = Q_3 - Q_1 \rightarrow 86,5 - 68,5 = 18$$

$$Q_d = \frac{1}{2}(Q_3 - Q_1) \rightarrow \frac{1}{2}(86,5 - 68,5) = 9$$

Jadi, jangkauan kuartilnya adalah 18 dan simpangan kuartilnya adalah 9

2. Simpangan Rata-rata (Deviasi)

Simpangan rata-rata adalah rata-rata yang hitung nilai absolut simpangan untuk menentukan kekurangan dari nilai jangkauan maka bisa dihitung nilai simpangan rata-rata simpangan rata-rata dengan memperhitungkan nilai-nilai lain selain nilai ekstrim pendistribusian data.

Simpangan rata-rata atau deviasi rata-rata adalah ukuran yang menyatakan seberapa besar penyebaran tiap nilai data terhadap nilai-nilai atau rata-ratanya, biasa disimbolkan dengan *SR*. Rumus menghitung simpangan rata-rata pada data berkelompok yaitu:

$$SR = \frac{\sum f|x_i - \bar{x}|}{n}$$

Keterangan :

SR = simpangan rata-ratanya

n = banyak kelas

x_i = nilai data ke - i

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

f = frekuensi data ke-i

Contoh : Tentukan simpangan rata-rata pada data berkelompok berikut ini: Daftar berat badan 100 mahasiswa

Berat badan (kg)	Banyak mahasiswa (f)
60 – 62	5
63 – 65	18
66 – 68	42
69 – 71	27
72 – 74	8

Penyelesaian

- Cari nilai tengah dari data tersebut

Berat badan (kg)	F	x_i	$f_i x_i$	$ x_i - \bar{x} $	$f x_i - \bar{x} $
60 – 62	5	61	305	6,45	32,25
63 – 65	18	64	1152	3,45	62,1
66 – 68	42	67	2814	0,45	18,9
69 – 71	27	70	1890	2,55	68,85
72 – 74	8	73	584	5,55	44,4
Jumlah	100		6745		226,5

- Hitunglah rata-rata data tersebut :

$$\bar{x} = \frac{\sum f_i \cdot x_i}{\sum f_i} = \frac{6745}{100} = 67,45$$

Dari bantuan tabel diatas maka dapat dihitung simpangan rata-ratanya :

$$SR = \frac{\sum f|x_i - \bar{x}|}{n} = \frac{226,5}{100} = 2,265$$

Jadi, simpangan rata-ratanya adalah 2,265.

3. Simpangan Baku (Standar Deviasi)

Simpangan baku ini paling banyak digunakan dalam ukuran penyebaran data. Dengan menggunakan simpangan rata-rata hasil pengamatan penyebaran sudah memperhitungkan seluruh nilai yang ada pada data. Namun, karena dalam perhitungan menggunakan nilai absolut maka tidak diketahui arah penyebarannya.

Sehingga, perlu digunakan simpangan baku karena simpangan baku memuat nilai pangkat 2 dari skor simpangan. simpangan baku merupakan ukuran penyebaran paling teliti. Simpangan baku atau disebut juga deviasi standar merupakan akar dari jumlah kuadrat deviasi dibagi banyaknya data.

Sekumpulan data kuantitatif yang dikelompokkan dapat dinyatakan oleh $x_1, x_2, \dots, x_n$ dan masing-masing data mempunyai frekuensi $f_1, f_2, \dots, f_n$. simpangan baku atau S dari data tersebut diperoleh dengan menggunakan rumus:

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

keterangan:

SB = simpangan baku

x_i = data ke - i

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

f_i = frekuensi data ke - i

n = banyaknya data

Contoh: Tentukan simpangan baku dari data berkelompok

berikut ini!Daftar berat badan 100 mahasiswa

Berat badan (kg)	Banyak mahasiswa (f)
60 – 62	5
63 – 65	18
66 – 68	42
69 – 71	27
72 – 74	8

- Mencari nilai tengah dari data tersebut

Berat badan (kg)	f	x_i	$f_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
60 – 62	5	61	305	-6,45	41,6025	208,01
						25

63 – 65	18	64	1152	-3,45	11,9025	214,24 5
66 – 68	42	67	2814	-0,45	0,2025	8,505
69 – 71	27	70	1890	2,55	6,5025	175,56 75
72 – 74	8	73	584	5,55	30,3025	242,42
Jumlah	100		6745			848,75

- Hitunglah rata-rata dari data tersebut

$$\bar{x} = \frac{\sum f_i \cdot x_i}{\sum f_i} = \frac{6745}{100} = 67,45$$

Dari tabel bantuan diatas maka dapat dihitung simpangan bakunya :

$$SB = \sqrt{\frac{\sum f (x_i - \bar{x})^2}{n - 1}} \rightarrow \sqrt{\frac{848,75}{100 - 1}} = 2,93$$

Jadi, simpangan baku dari data berkelompok diatas adalah 2,93

4. Ragam (Variansi)

Ragam atau variansi adalah rata-rata dari jumlah kuadrat simpangan tiap data. Ragam bisa dirumuskan sebagai:

$$\text{Ragam} = SB^2$$

Artinya ragam diperoleh dari nilai simpangan baku dikuadratkan.

Contoh:

Dari contoh soal yang berkaitan dengan simpangan baku data berkelompok di atas, diperoleh simpangan bakunya adalah 2,93. Sehingga nilai ragamnya (Variansi) adalah: $\text{Ragam} = SB^2 = 2,93^2 = 8,5849$.

Jadi, ragam dari simpangan baku pada data berkelompok diatas adalah 8,5849.