



**ISMANTO HADI SANTOSO**

# STATISTIK 1



**PENERBIT  
UWKS PRESS**

# STATISTIK I

Ismanto Hadi Santoso

UWKS PRESS



**PENERBIT  
UWKS PRESS**

## STATISTIK I

ISBN .....  
17,6 × 25 cm  
165 hlm  
Cetakan ke-1, Agustus 2023

**Penulis:**

Ismanto Hadi Santoso

**Editor:**

Reza Syehma Bahtiar

**Penerbit:**

UWKS PRESS

Anggota IKAPI No.206/Anggota Luar Biasa/JTI/2018

Anggota APPTI No.002.071.1.12019

Jl. Dukuh Kupang XXV/54 Surabaya Jawa Timur 60225

Telp. (031) 5677577

Hp. 085745182452 / 081703875858

Email : [uwkspress@gmail.com](mailto:uwkspress@gmail.com) / [uwkspress@uwks.ac.id](mailto:uwkspress@uwks.ac.id)

**Dilarang mengutip sebagian atau seluruh isi buku ini  
dengan cara apapun, termasuk dengan penggunaan  
mesin fotokopi, tanpa izin sah dari penerbit**

*Buat anak-anakku, menantu, dan cucu-  
cucu ku semua*

**Radytia Buyung Permata – Faricha Rahmati**

**Radytia Bayu Sithala – Ana Tri Rahayu**

**Ristiya Gakih Paramita – Galih Satrio Utomo**

**Kanza, Audry. Abdulloh**

**Nadhira**

**Rinjani, Kerinci**

**Ini yang papa persembahkan, buat kenang-kenangan**

UWKS

## SINOPSIS

Buku *Statistik I* merupakan pengantar bagi mahasiswa untuk memahami dasar-dasar statistika, terutama sebagai buku pegangan bagi mahasiswa. Buku ini juga dapat dijadikan sebagai buku pegangan bagi mahasiswa dalam semua bidang studi baik itu sosiologi, psikologi, administrasi bisnis, ataupun bidang ilmu pengetahuan alam. Berbekal ilmu pengetahuan aljabar sekolah menengah atas sudah cukup menjadi bekal untuk dapat memahami konsep-konsep dasar statistika deskriptif dan inferensial yang diuraikan. Meskipun pengalaman menunjukkan bahwa mahasiswa akan lebih mendapatkan manfaat dari mata kuliah ini bila sebelumnya telah mempelajari pengantar mengenai fungsi atau matematika terhingga (*finite mathematics*).

Dalam penjelasan materi dilengkapi dengan contoh soal dan soal latihan yang disesuaikan dengan dinamika yang terjadi di lapangan, dan ditempatkan langsung setelah pasal yang membahasnya dalam setiap bab. Konsep-konsep statistik tambahan dan berbagai prosedur yang didefinisikan di banyak soal latihan dapat diberikan sesuai dengan kehendak dan kebijaksanaan dosen pengajar.

Seluruh materi terdiri dari 9 (sembilan) bab. Berbagai konsep statistika deskriptif disatukan dalam lima bab pertama. Bab 1 membahas Pemahaman Statistik, Bab 2 membahas

Data Statistik, Bab 3 membahas Penyajian Data, Bab 4 membahas Ukuran Pemusatan Data dan Letak Data, dan Bab 5 membahas Ukuran Penyebaran. Untuk melengkapi pemahamannya, uraian tentang persentil, kuartil; dan desil telah diperluas sehingga mencakup baik data yang dikelompokkan ataupun yang belum dikelompokkan, uraian mengenai koefisien kemajuluran kurva juga diberikan, dan bahkan diberikan tambahan untuk menggambarkan kegunaan bilangan acak dalam mengambil suatu contoh acak. Adapun konsep terkait dengan angka indeks dibahas pada Bab 6. Statistika masa kini, yang semakin menekankan aspek inferensianya, didasarkan pada teori peluang. Oleh karena itu pengenalan konsep dasar teori peramalan dengan pemanfaatan data kurun waktu dibahas dalam tiga bab berikutnya. Bab 7 membahas Analisis Serial Waktu (*time series*), Bab 8 membahas Analisis Trend, dan Bab 9 membahas Regresi Dan Korelasi.

Buku ini cukup bahan untuk menyampaikan materi dalam satu semester, dengan 12 sampai 14 pertemuan dengan 3 (tiga) jam perminggu, serta memungkinkan untuk menyampaikan secara terstruktur dan sistematis. Beberapa contoh soal yang ada dapat dikembangkan dengan menambah kasus-kasus terkini untuk menambah kedalaman pemahaman mahasiswa, sekaligus sebagai soal-soal latihan

## KATA PENGANTAR

Buku *Statistik I* merupakan pengantar bagi mahasiswa untuk memahami dasar-dasar statistika. Buku ini dapat dijadikan sebagai buku pegangan bagi mahasiswa dalam semua bidang studi baik itu sosiologi, psikologi, administrasi bisnis, ataupun bidang ilmu pengetahuan alam. Berbekal ilmu pengetahuan aljabar sekolah menengah atas sudah cukup menjadi bekal untuk dapat memahami konsep-konsep dasar statistika deskriptif dan inferensial yang diuraikan. Meskipun pengalaman menunjukkan bahwa mahasiswa akan lebih mendapatkan manfaat dari mata kuliah ini bila sebelumnya telah mempelajari pengantar mengenai fungsi atau matematika terdahulu (*finite mathematics*).

Dalam penjelasan materi dilengkapi dengan contoh soal dan soal latihan yang disesuaikan dengan dinamika yang terjadi di lapangan, dan ditempatkan langsung setelah pasal yang memahaminya dalam setiap bab. Konsep-konsep statistika tambahan dan berbagai prosedur yang didefinisikan di banyak soal latihan dapat diberikan sesuai dengan kehendak dan kebijaksanaan dosen pengajar.

Seluruh materi terdiri dari 9 (sembilan) bab. Berbagai konsep statistika deskriptif disatukan dalam lima bab pertama. Bab 1 membahas Pemahaman Statistik, Bab 2 membahas Data Statistik, Bab 3 membahas Penyajian Data, Bab 4 membahas Ukuran

Pemusatan Data dan Letak Data, dan Bab 5 membahas Ukuran Penyebaran. Untuk melengkapi pemahamannya, uraian tentang persentil, kuartil; dan desil telah diperluas sehingga mencakup baik data yang dikelompokkan ataupun yang belum dikelompokkan, uraian mengenai koefisien kemajuluran kurva juga diberikan, dan bahkan diberikan tambahan untuk menggambarkan kegunaan bilangan acak dalam mengambil suatu contoh acak. Adapun konsep terkait dengan angka indeks dibahas pada Bab 6. Statistika masa kini, yang semakin menekankan aspek inferensianya, didasarkan pada teori peluang. Oleh karena itu pengenalan konsep dasar teori peramalan dengan pemanfaatan data kurun waktu dibahas dalam tiga bab berikutnya. Bab 7 membahas Analisis Serial Waktu (*time series*), Bab 8 membahas Analisis Trend, dan Bab 9 membahas Regresi Dan Korelasi.

Buku ini cukup bahan untuk menyampaikan materi dalam satu semester, dengan 12 sampai 14 pertemuan dengan 3 (tiga) jam perminggu, serta memungkinkan untuk menyampaikan secara terstruktur dan sistematis. Beberapa contoh soal yang ada dapat dikembangkan dengan menambah kasus-kasus terkini untuk menambah kedalaman pemahaman mahasiswa, sekaligus sebagai soal-soal latihan.



# DAFTAR ISI

## BAB I PENDAHULUAN

1.1 Pengertian statistik dan statistika .....	3
1.2 Statistika Deskriptip Dan Statistika Infernsial .....	7
1.3 Landasan Kerja Statistika .....	10
1.4 Karakteristik Pokok Statistika .....	11
1.5 Manfaat Dan Kegunaan Statistik .....	12
1.6 Jenis-jenis Analisis Statistik .....	12

## BAB II DATA STATISTIK:

2.1 Jenis-Jenis Data .....	13
2.2 Skala Pengukuran Data .....	15
2.3 Sumber Data Statistik .....	19

## BAB III PENYAJIAN DATA

3.1 Pengertian dan Kegunaan .....	23
3.2 Tabel Distribusi Frekuensi .....	23
<b>a.</b> Tabel Distribusi Frekuensi .....	23
<b>b.</b> Frekuensi Relatif dan Kumulatif .....	26
3.3 Penyajian Data dalam Bentuk Diagram .....	29
3.4 Histogram, Poligon Frekuensi, dan Ogive .....	35

## BAB IV UKURAN PEMUSATAN DAN LETAK DATA

4.1 Pengertian Nilai Sentral .....	39
<b>a.</b> Mean ( <i>arithmetic mean</i> ) .....	40

b. Median .....	45
c. Mode/Modus .....	48
d. Rata-rata Ukur (Geometric Mean) .....	52
e. Rata-rata Harmonik (H) .....	54

## BAB V UKURAN PENYEBARAN

5.1 Jangkauan (Range) .....	59
5.2 Simpangan kuartil ( <i>Quartile Deviation</i> ) .....	62
5.3 Simpangan Rata-rata ( <i>Mean Deviation</i> ) .....	64
5.4 Ragam dan Standart Deviasi .....	67
5.5 Ukuran sebaran relatif ( <i>Measures of Relative Dispersion</i> ).....	72
5.6 Skewness and Kurtosis .....	74

## BAB VI ANGKA INDEKS

6.1 Angka Perbandingan, Angka Relatif, Angka Indeks.....	79
6.2 Kegunaan Angka Indeks .....	80
6.3 Angka Indeks Sederhana .....	81
6.4 Angka Indeks Gabungan .....	84

## BAB VII ANALISA SERIAL WAKTU (*TIME SERIES*)

7.1 Pengertian .....	88
7.2 Jenis-Kanis Metode Peramalam .....	89
7.3 Time Series atau Derat Waktu .....	90
7.4 Causal Methods atau Sebab Akibat .....	90

## BAB VIII ANALISIS TREND

8.1 Variasi musiman .....	91
---------------------------	----

8.2	Matode Trend	.....91
8.3	Methode Trend Semi Average	.....92
8.4	Methode Trend Moment	.....92
8.5	Methode Trend Kuadrat Terkecil	.....93
8.6	Methode Trend Musiman	.....96
8.7	Mencari Indeks Musiman	.....97
8.8	Metode Dekomposisi	.....99

## BAB IX REGRESI DAN KARELASI

9.1	Pendahuluan	.....101
9.2	Regresi Linier Sederhana	.....107
9.3	Korelasi Linier Sederhana	.....110
PUSTAKA		.....162

# BAB I

## PENDAHULUAN

Statistika adalah ilmu yang mempelajari bagaimana merencanakan, mengumpulkan, menganalisis, menginterpretasi, dan mempresentasikan data. Singkatnya, statistika adalah ilmu yang berkenaan dengan data. Istilah 'statistika' (bahasa Inggris: statistiks) berbeda dengan 'statistik' (statistik). Statistika merupakan ilmu yang berkenaan dengan data, sedang statistik adalah data, informasi, atau hasil penerapan algoritma statistika pada suatu data. Dari kumpulan data, statistika dapat digunakan untuk menyimpulkan atau mendeskripsikan data; ini dinamakan statistika deskriptif. Sebagian besar konsep dasar statistika mengasumsikan teori probabilitas. Beberapa istilah statistika antara lain: populasi, sampel, unit sampel, dan probabilitas. Statistika merupakan pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisannya dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisaan yang dilakukan. Statistika dikelompokkan dalam dua kelompok yaitu *statistika deskriptif* dan *statistika inferensia*. *Statistika deskriptif* adalah metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Sedangkan pengertian *statistika inferensia* adalah metode yang berhubungan dengan analisis sebagian data untuk kemudian sampai pada peramalan

atau penarikan kesimpulan tentang seluruh gugus data induknya

Penggunaan istilah statistika berakar dari istilah-istilah dalam bahasa latin modern *statistikum collegium* ("dewan negara") dan bahasa Italia *statista* ("negarawan" atau "politikus").

**Gottfried Achenwall (1749)** menggunakan Statistik dalam bahasa Jerman untuk pertama kalinya sebagai nama bagi kegiatan analisis data kenegaraan, dengan mengartikannya sebagai "ilmu tentang negara (state)". Pada awal abad ke-19 telah terjadi pergeseran arti menjadi "ilmu mengenai pengumpulan dan klasifikasi data". Sir John Sinclair memperkenalkan nama (*Statistiks*) dan pengertian ini ke dalam bahasa Inggris. Jadi, statistika secara prinsip mula-mula hanya mengurus data yang dipakai lembaga-lembaga administratif dan pemerintahan. Pengumpulan data terus berlanjut, khususnya melalui sensus yang dilakukan secara teratur untuk memberi informasi kependudukan yang berubah setiap saat.

Pada abad ke-19 dan awal abad ke-20 statistika mulai banyak menggunakan bidang-bidang dalam matematika, terutama peluang. Cabang statistika yang pada saat ini sangat luas digunakan untuk mendukung metode ilmiah, statistika inferensi, dikembangkan pada paruh kedua abad ke-19 dan awal abad ke-20 oleh **Ronald Fisher** (peletak dasar statistika inferensi), **Karl Pearson** (metode regresi linear), dan **William Sealey Gosset** (meneliti problem sampel berukuran kecil). Penggunaan statistika pada masa sekarang dapat dikatakan telah

menyentuh semua bidang ilmu pengetahuan, mulai dari astronomi hingga linguistika. Bidang-bidang ekonomi, biologi dan cabang-cabang terapannya, serta psikologi banyak dipengaruhi oleh statistika dalam metodologinya. Akibatnya lahirlah ilmu-ilmu gabungan seperti ekonometrika, biometrika (atau biostatistika), dan psikometrika.

Meskipun ada pihak yang menganggap statistika sebagai cabang dari matematika, tetapi sebagian pihak lainnya menganggap statistika sebagai bidang yang banyak terkait dengan matematika melihat dari sejarah dan aplikasinya. Di Indonesia, kajian statistika sebagian besar masuk dalam fakultas matematika dan ilmu pengetahuan alam, baik di dalam departemen tersendiri maupun tergabung dengan matematika.

Statistika banyak diterapkan dalam berbagai disiplin ilmu, baik ilmu-ilmu alam (misalnya astronomi dan biologi maupun ilmu-ilmu sosial (termasuk sosiologi dan psikologi), maupun di bidang bisnis, ekonomi, dan industri. Statistika juga digunakan dalam pemerintahan untuk berbagai macam tujuan; sensus penduduk merupakan salah satu prosedur yang paling dikenal. Aplikasi statistika lainnya yang sekarang populer adalah prosedur jajak pendapat atau polling (misalnya dilakukan sebelum pemilihan umum), serta hitung cepat (perhitungan cepat hasil pemilu) atau quick count. Di bidang komputasi, statistika dapat pula diterapkan dalam pengenalan pola maupun kecerdasan buatan.

## 1.2 Pengertian statistik dan statistika

Kata statistik dan statistika mungkin tidak asing lagi ditelinga kita, sebab dahulu kita pernah mempelajarinya pada sekolah menengah atas. Statistik dan statistika tidaklah jauh dari kehidupan kita sehari-hari, peranannya sangat membantu kita untuk memecahkan suatu masalah. Masyarakat tanpa menyadari sudah menggunakan statistik dan statistika didalam keseharian, contohnya ketika mereka mencatat pengeluaran dan pemasukan tiap bulannya. Walaupun masyarakat tidak jauh dari statistik dan statistika, namun kebanyakan masyarakat menganggap bahwa statistik dan statistika adalah sama. Ada beberapa hal yang akan dijelaskan lebih terperinci untuk membedakan statistik dan statistika. Berikut adalah penjelasan lebih terperinci mengenai pengertian statistik dan statistika.

### a. Pengertian statistik

Kata statistik bukan merupakan kata dari bahasa Indonesia asli, secara etimologis kata "statistik" berasal dari kata *status* (bahasa latin) yang mempunyai persamaan arti dengan kata *state* (bahasa Inggris) atau kata *staat* (bahasa Belanda), dan yang dalam bahasa Indonesia diterjemahkan menjadi *negara*. Pada mulanya, kata "statistik" diartikan sebagai "kumpulan bahan keterangan (data), baik yang berwujud angka (data kuantitatif) maupun yang tidak berwujud angka (data kualitatif), yang mempunyai arti

penting dan kegunaan yang besar bagi suatu *negara*. Namun, pada perkembangan selanjutnya, arti kata statistik hanya dibatasi pada "kumpulan bahan keterangan yang berwujud angka (data kuantitatif)" saja; bahan keterangan yang tidak berwujud angka (data kualitatif) tidak lagi disebut *statistik*.

Seiring berjalannya waktu kata statistik tidak lagi dibatasi untuk kepentingan-kepentingan Negara saja tapi sudah digunakan dalam keseharian untuk mempermudah masyarakat untuk menganalisis sesuatu yang berkaitan dengan data-data. Sehingga setelah masyarakat memahami statistik dan mulai mempergunakannya dalam kehidupan sehari munculah berbagai macam nama statistik. Statistik yang menjelaskan sesuatu hal biasanya diberi nama statistik mengenai hal yang bersangkutan didalamnya, contohnya kumpulan data yang membahas tentang tingkat produksi suatu perusahaan dinamakan statistik produksi. Banyak persoalan baik itu seperti penelitian ataupun pengamatan yang dinyatakan dalam bentuk bilangan atau angka-angka. Kumpulan angka-angka disusun atau diatur dan disajikan dalam tabel (terkadang dilengkapi dengan gambar baik berupa iagram maupun grafik, hal ini dilakukan bertujuan untuk mempermudah menjelaskan isi dari data) seperti berikut mungkin bisa membantu anda memahami statistik lebih lanjut.



Contoh daftar harga alat tulis menulis disuatu toko

NO	NAMA BARANG	HARGA
1	Pensil	Rp. 2000,00
2	Penghapus	Rp. 1000,00
3	Bulpoin	Rp. 2500,00
4	Stipo (pemutih)	Rp. 4000,00
5	Buku tulis	Rp. 3000,00
6	Penggaris	Rp. 1500,00
7	Buku gambar A4	Rp. 5000,00
8	Spidol	Rp. 4000,00

Contoh tersebut merupakan contoh statistika kuantitatif.

Data kuantitatif, merupakan data yang berbentuk bilangan serta harganya berubah-ubah (bersifat variabel). Dari segi nilainya, data kuantitatif dibagi menjadi dua golongan yaitu data dengan variabel diskrit (data diskrit), dan data dengan variabel kontinu (data kontinu). Selain data kuantitatif dalam statistika dikenal juga istilah data kualitatif, data kualitatif merupakan data yang dikategorikan menurut lukisan kualitas objek yang diteliti atau dipelajari.

Contohnya:

*“seorang mahasiswa datang keruangan kajor untuk mencari nilai kepribadian”*. Nilai kepribadian dikatakan data kualitatif dikarenakan hanya dinyatakan dengan kata yang

didasari oleh data kuantitatif. Maka dapat disimpulkan bahwa statistik merupakan kumpulan data baik berupa bilangan maupun bukan bilangan yang disusun dalam table ataupun diagram yang melukiskan atau menggambarkan suatu persoalan.

Kata statistik bisa juga digunakan untuk menyatakan ukuran sebagai wakil dari kumpulan data mengenai sesuatu hal. Ukuran ini didapat berdasarkan perhitungan menggunakan kumpulan sebagian data yang diambil dari keseluruhan tentang persoalan tersebut (misal : persen dan rata-rata).

#### **b. Pengertian statistika**

Dari data hasil penelitian sering kali diminta suatu uraian, penjelasan atau kesimpulan tentang persoalan yang diteliti. Sebelum kesimpulan dibuat, keterangan data yang telah terkumpul itu terlebih dahulu dipelajari, dianalisis atau diolah dan berdasarkan pengolahan ini baru dibuat kesimpulan. Dari pernyataan diatas tersirat bahwa statistika merupakan pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisisannya dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisisan yang dilakukan.

Maka dari definisi di atas dapat kita simpulkan bahwa ruang lingkup statistika lebih luas daripada

statistik serta statistika mencakup statistik, atau dapat kita analogikan ibarat computer, suatu keutuhan computer merupakan statistika sedangkan alat-alat penyusun dari computer ( LCD, mouse, CPU, keyboard, dll) merupakan statistika.

**Sudjana** (2002) mendefinisikan statistika sebagai pengetahuan yang berhubungan dengan cara-cara pengumpulan fakta, pengolahan serta pembuatan keputusan yang cukup beralasan berdasarkan fakta dan analisa yang dilakukan. Sementara statistik dipakai untuk menyatakan kumpulan fakta, umumnya berbentuk angka yang disusun dalam tabel atau diagram yang melukiskan atau menggambarkan suatu persoalan. Lebih katakan statistika adalah ilmu terdiri dari teori dan metode yang merupakan cabang dari matematika terapan dan membicarakan tentang : bagaimana mengumpulkan data, bagaimana meringkas data, mengolah dan menyajikan data, bagaimana menarik kesimpulan dari hasil analisis, bagaimana menentukan keputusan dalam batas-batas resiko tertentu berdasarkan strategi yang ada.

**Singgih Santoso** (2012) menyatakan, pada prinsipnya statistik diartikan sebagai kegiatan untuk

mengumpulkan data, meringkas/menyajikan data, menganalisa data dengan metode tertentu, dan menginterpretasikan hasil analisis tersebut.

Dalam kaitannya untuk menyelesaikan masalah, pendekatan statistik terbagi dua yaitu pendekatan statistik dalam arti sempit dan luas. Dalam arti sempit (statistik deskriptif), statistika yang hanya mendeskripsikan tentang data yang dijadikan dalam bentuk tabel, diagram, pengukuran rata-rata, simpangan baku, dan seterusnya tanpa perlu menggunakan signifikansi atau tidak bermaksud membuat generalisasi.

Sementara dalam arti luas (statistik inferensi/induktif) adalah alat pengumpul data, pengolah data, menarik kesimpulan, membuat tindakan berdasarkan analisis data yang dikumpulkan dan hasilnya dimanfaatkan / digeneralisasi untuk populasi.

Bidang keilmuan statistika adalah sekumpulan metode untuk memperoleh dan menganalisa data dalam pengambilan suatu kesimpulan. Meski merupakan cabang ilmu matematika, statistika memiliki perbedaan mendasar pada logikanya. Jika matematika menggunakan logika deduktif, sementara statistik menggunakan logika induktif.

Logika statistika, dengan demikian sering disebut dengan logika induktif yang tidak memberikan kepastian namun member tingkat peluang bahwa untuk

premis-premis tertentu dapat ditarik kesimpulan, dan kesimpulannya mungkin benar mungkin juga tidak.

Langkah yang ditempuh dalam logika statistika adalah:

1. Observasi dan eksperimen
2. Munculnya hipotesis ilmiah
3. Verifikasi dan pengukuhan dan berakhir pada
4. Sebuah teori dan hukum ilmiah.

## **1.2 Statistika Deskriptif Dan Statistika Inferensial**

### **a. Statistika deskriptif**

*Statistika deskriptif* merupakan metode-metode yang berkait dengan pengumpulan dan penyajian sekumpulan data, sehingga dapat memberikan informasi yang berguna. Perlu kiranya dimengerti bahwa statistika deskriptif memberikan informasi hanya mengenai data yang dipunyai dan sama sekali tidak menarik kesimpulan yang lebih banyak dan lebih jauh dari data yang ada. Kegiatan memeriksa sifat-sifat penting dari data yang ada itu disebut analisis data secara pemerian (deskripsi). Karenanya bagian statistika demikian dinamakan Statistika Deskriptif atau Statistika Perian. Penyusunan tabel, diagram, modus, kuartil, simpangan baku termasuk dalam kategori statistika deskriptif. Kegiatan itu dilakukan melalui:

- 1) Pendekatan aritmetika yaitu pendekatan melalui pemeriksaan rangkuman nilai atau ukuran-ukuran

penting dari data. Yang dimaksud rangkuman nilai di sini ialah penyederhanaan kumpulan nilai data yang diamati ke dalam bentuk nilai-nilai tertentu. Setiap rangkuman nilai ini disebut statistik. Jadi, statistik menerangkan sifat kumpulan data dalam bentuk nilai yang mudah dipahami, sedangkan statistika adalah suatu ilmu tentang sekumpulan konsep serta metode yang dapat digunakan untuk mengumpulkan, menyajikan dan menganalisis data serta menarik kesimpulan berdasar hasil analisis data tersebut.

- 2) Pendekatan geometrik, yaitu melalui penyajian data dalam bentuk gambar berupa grafik atau diagram.
- 3) Kedua pendekatan mengakibatkan perbedaan dalam penyajian datanya. Penyajian data pertama menekankan angka-angka dan yang kedua menekankan pada gambar.

Berikut merupakan penjelasan mengenai pengertian statistik menurut para ahli.

**Sudjana (2012)** menjelaskan : Fase statistika dimana hanya berusaha melukiskan atau mengalalisa kelompok yang diberikan tanpa membuat atau menarik kesimpulan tentang populasi atau kelompok yang lebih besar dinamakan statistika deskriptif

**Iqbal Hasan (2001)** menjelaskan : Statistik deskriptif atau statistik deduktif adalah bagian dari statistik mempelajari cara pengumpulan data dan penyajian data sehingga mudah dipahami. Statistik deskriptif hanya berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai suatu data atau keadaan atau fenomena. Dengan kata lain, statistik deskriptif berfungsi menerangkan keadaan, gejala, atau persoalan. Penarikan kesimpulan pada statistik deskriptif (jika ada) hanya ditujukan pada kumpulan data yang ada. Didasarkan pada ruang lingkup bahasannya statistik deskriptif mencakup *Distribusi frekuensi* beserta bagian-bagiannya seperti :

- Grafik distribusi (histogram, poligon frekuensi, dan ogif);
- Ukuran nilai pusat (rata-rata, median, modus, kuartil dan sebagainya);
- Ukuran dispersi (jangkauan, simpangan rata-rata, variasi, simpangan baku, dan sebagainya);
- Kemencengan dan keruncingan kurva
- Angka indeks
- Times series/deret waktu atau berkala
- Korelasi dan regresi sederhana

**Bambang Suryoatmono (2004)** menyatakan Statistika

Deskriptif adalah statistika yang menggunakan data pada suatu kelompok untuk menjelaskan atau menarik kesimpulan mengenai kelompok itu saja

- Ukuran Lokasi: **mode, mean, median**, dll
- Ukuran Variabilitas: varians, deviasi standar, range, dll
- Ukuran Bentuk: skewness, kurtosis, plot boks

**Pangestu Subagyo (2003)** menyatakan : Yang dimaksud

sebagai statistika deskriptif adalah bagian statistika mengenai pengumpulan data, penyajian, penentuan nilai-nilai statistika, pembuatan diagram atau gambar mengenai sesuatu hal, disini data yang disajikan dalam bentuk yang lebih mudah dipahami atau dibaca.

### **b. Statistika inferensi**

*Statistik inferensial* merupakan kebalikan dari statistika deskriptif, statistika inferensial merupakan statistik yang berkenaan dengan cara penarikan kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan karakteristik atau ciri dari suatu populasi. Dengan demikian dalam statistik inferensial dilakukan suatu generalisasi (perampatan atau memperumum) dan hal yang bersifat khusus (kecil) ke hal yang lebih luas (umum). Oleh karena itu, statistik



inferensial disebut juga statistik induktif atau statistik penarikan kesimpulan. Pada statistik inferensial biasanya dilakukan pengujian hipotesis dan pendugaan mengenai karakteristik (ciri) dari suatu populasi, seperti mean dan Uji t (Sugiyono, 2006). Statistika inferensi, yang berupa kajian tentang penarikan kesimpulan mengenai keseluruhan objek yang menjadi perhatian namun hanya atas dasar data sebagian objek inilah yang disebut Statistika Inferensial atau Statistika Induktif. Dengan demikian, Statistika Inferensial menyimpulkan makna statistik yang telah dihitung, dianalisis atau disajikan grafik atau diagramnya tersebut. Penarikan kesimpulan tentang keseluruhan populasi didasarkan atas pengamatan terhadap salah satu bagian populasi disebut induksi atau generalisasi. Proses induksi atau generalisasi dalam statistika induktif dapat ditemui dalam berbagai kegiatan ilmiah dan juga dalam kehidupan sehari-hari. Misalnya seorang anak kecil sering melihat balok-balok kayu dapat terbakar, maka ia akan menarik kesimpulan bahwa semua balok kayu dapat terbakar.

### **1.3 Landasan Kerja Statistika**

Menurut Sutrisno Hadi (dalam Riduwan dan Sunarto, 2007) ada tiga jenis landasan kerja statistik meliputi

- 1) **Variasi**; didasarkan atas kenyataan bahwa seorang peneliti atau penyelidik selalu menghadapi persoalan

dan gejala yang bermacam-macam.

- 2) **Reduksi**, memberikan kesempatan pada penyelidik untuk menyelidiki hanya sebagian dari seluruh gejala atau kejadian yang hendak diselidiki (Penelitian Sampling)
- 3) **Generalisasi**; sekalipun penelitian dilakukan terhadap sebagian dari seluruh peristiwa/kejadian yang hendak diteliti, namun kesimpulan dari penelitian akan diperuntukan bagi keseluruhan kejadian/peristiwa atau gejala yang hendak diambil.

#### 1.4 Karakteristik Pokok Statistika

Menurut Riduwan, dkk (2007), bahwa karakteristik atau ciri pokok statistik sebagai berikut :

- 1) Statistik bekerja dengan angka. Angka-angka ini dalam statistik mempunyai dua pengertian, yaitu :
  - a) Angka statistik sebagai jumlah atau frekuensi dan angka statistik sebagai nilai atau harga. Pengertian ini mengandung arti bahwa data statistik adalah data kuantitatif. Contoh : Jumlah pegawai Pemda Provinsi Papua, Jumlah dosen Universitas Cenderawasih, jumlah anggota MPR / DPR dan lain sebagainya .
  - b) Angka statistik sebagai nilai mempunyai arti data kualitatif yang diwujudkan dalam angka. Contoh : Nilai kepribadian, nilai

kecerdasan mahasiswa, metode mengajar dosen, kualitas sekolah, mutu pemberdayaan guru dan lain sebagainya.

- 2) Statistik bersifat objektif. Statistik bekerja dengan angka sehingga mempunyai sifat objektif, artinya angka statistik dapat digunakan sebagai alat pencari fakta, pengungkap kenyataan yang ada dan memberikan keterangan yang benar, kemudian menentukan kebijakan sesuai fakta dan temuannya diungkapkan apa adanya.
- 3) Statistik bersifat universal (umum). Statistik tidak hanya digunakan dalam salah satu disiplin ilmu saja, tetapi dapat digunakan secara umum dalam berbagai bentuk disiplin ilmu pengetahuan dengan penuh keyakinan

### **1.5 Manfaat Dan Kegunaan Statistik**

Statistik dapat digunakan sebagai alat (Riduwan dan Sunarto, 2007) :

1. **Komunikasi.** Adalah sebagai penghubungan beberapa pihak yang menghasilkan data statistik atau berupa analisis statistik sehingga beberapa pihak tersebut akan dapat mengambil keputusan melalui informasi tersebut.
2. **Deskripsi.** Merupakan penyajian data dan mengilustrasikan data, misalnya mengukur tingkat

kelulusan siswa, laporan keuangan, tingkat inflasi, jumlah penduduk, dan seterusnya

3. **Regresi.** Adalah meramalkan pengaruh data yang satu dengan data yang lainnya dan untuk menghadapi gejala-gejala yang akan datang
4. **Korelasi.** Untuk mencari kuatnya atau besarnya hubungan data dalam suatu penelitian
5. **Komparasi.** Yaitu membandingkan data dua kelompok atau lebih.

#### 1.6 Jenis-jenis analisis statistik

Statistika dibedakan berdasarkan jenisnya menjadi dua yaitu:

Statistika deskriptif adalah statistika yang berkaitan dengan metode atau cara mendeskripsikan, menggambarkan, menjabarkan atau menguraikan data. Statistika deskripsi mengacu pada bagaimana menata, menyajikan dan menganalisis data, yang dapat dilakukan misalnya dengan menentukan nilai rata-rata hitung, median, modus, standar deviasi atau menggunakan cara lain yaitu dengan membuat tabel distribusi frekuensi dan diagram atau grafik.

Statistika inferensia adalah statistika yang berkaitan dengan cara penarikan kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan karakteristik dari suatu populasi. Dengan demikian dalam statistika inferensia data yang diperoleh

dilakukan generalisasi dari hal yang bersifat kecil (khusus) menjadi hal yang bersifat luas (umum).

UWKSPRESS

## **BAB II**

### **DATA STATISTIK**

Data adalah sesuatu yang belum mempunyai arti bagi penerimanya dan masih memerlukan adanya suatu pengolahan. Data bisa berwujud suatu keadaan, gambar, suara, huruf, angka, matematika, bahasa ataupun simbol-simbol lainnya yang bisa kita gunakan sebagai bahan untuk melihat lingkungan, obyek, kejadian ataupun suatu konsep.

#### **2.1 Jenis-Jenis Data**

##### **1) Jenis Data Menurut Cara Memperolehnya :**

###### **a. Data Primer**

Data primer adalah secara langsung diambil dari objek / obyek penelitian oleh peneliti perorangan maupun organisasi. Contoh : Mewawancarai langsung penonton bioskop 21 untuk meneliti preferensi konsumen bioskop.

###### **b. Data Sekunder**

Data sekunder adalah data yang didapat tidak secara langsung dari objek penelitian. Peneliti mendapatkan data yang sudah jadi yang dikumpulkan oleh pihak lain dengan berbagai cara atau metode baik secara komersial maupun non komersial. Contohnya adalah pada peneliti yang menggunakan data statistik hasil riset dari surat kabar atau majalah.

## **2)Macam-Macam Data Berdasarkan Sumber Data :**

### **a. Data Internal**

Data internal adalah data yang menggambarkan situasi dan kondisi pada suatu organisasi secara internal. Misal : data keuangan, data pegawai, data produksi, dsb.

### **b. Data Eksternal**

Data eksternal adalah data yang menggambarkan situasi serta kondisi yang ada di luar organisasi. Contohnya adalah data jumlah penggunaan suatu produk pada konsumen, tingkat preferensi pelanggan, persebaran penduduk, dan lain sebagainya.

## **3)Jenis-jenis Data Menurut Waktu Pengumpulannya :**

### **a. Data Cross Section**

Data cross-section adalah data yang menunjukkan titik waktu tertentu. Contohnya laporan keuangan per 31 desember 2006, data pelanggan PT. Angin Ribut bulan mei 2004, dan lain sebagainya.

### **b. Data Time Series / Berkala**

Data berkala adalah data yang datanya menggambarkan sesuatu dari waktu ke waktu atau periode secara historis. Contoh data time series adalah data perkembangan nilai tukar dollar

amerika terhadap euro eropa dari tahun 2004 sampai 2006, jumlah pengikut jamaah nurdin m. top dan doktor azahari dari bulan ke bulan, dll.

**c. Data dengan Variabel bebas dan variabel terikat :**

Variabel bebas adalah data unit atau ukuran yang diubah dalam suatu pengamatan. Dalam hubungan sebab-akibat, variabel terikat berperan sebagai sebab sementara variabel bebas adalah akibat.

Data dengan variabel terikat adalah data unit atau ukuran yang berubah sesuai dengan berubahnya variabel lain. Variabel terikat menjadi hal yang diperhatikan dalam suatu pengamatan.

**d. Data Berkala**

Data berkala adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan suatu perkembangan atau kecenderungan keadaan/peristiwa/ kegiatan. Biasanya jarak dari waktu ke waktu sama. Data berkala disebut juga time series data. Dengan analisis data berkala kita dapat mengetahui perkembangan satu atau beberapa keadaan serta hubungan atau pengaruhnya terhadap keadaan lain.

**2.2 Skala pengukuran data**

Dalam statistik parametris menggunakan analisis data yang berupa,



## 1) Data Nominal

Skala ini menempatkan angka sebagai atribut objek. Tidak memiliki efek evaluatif karena hanya menempatkan angka ke dalam kategori tanpa struktur, tidak memiliki peringkat dan tidak ada jarak.

Contoh Data Variabel :

- a. Ya = 1 dan Tidak = 0
- b. Pria = 1 dan Wanita = 0
- c. Hitam = 1, Abu-abu = 2, Putih = 2

Analisis Statistik :

Angka tidak bermakna matematika. Analisis statistik yang dapat digunakan berada dalam kelompok non-parametrik yaitu frekuensi dan tabulasi silang dengan Chi-square.

Data ini juga sering disebut data diskrit, kategorik, atau dikhotomi. Disebut diskrit karena ini data ini memiliki sifat terpisah antara satu sama lainnya, baik pemisahan itu terdiri dari dua bagian atau lebih; dan di dalam pemisahan itu tidak terdapat hubungan sama sekali. Masing-masing kategori memiliki sifat tersendiri yang tidak ada hubungannya dengan kategori lainnya. Sebagai misal data hasil penelitian dikategorikan kedalam kelompok “ya” dan “tidak” saja.

Contohnya :

- a. laki-laki/wanita (laki-laki adalah ya laki-laki; dan wanita adalah “tidak laki-laki”), kawin /tidak kawin; janda/duda, dan lainnya.
- b. Jenis pekerjaan dapat digolongkan secara terpisah menjadi pegawai negeri, pedagang, dokter, petani, buruh dsb.
- c. Nomor punggung pemain sepak bola, nomor rumah, nomor plat mobil dan lainnya. Nomor-nomor tersebut semata-mata hanya menunjukkan simbol, tanda, atau atribut saja.
- d. Suku, golongan darah, jenis penyakit, bentuk atau konstitusi tubuh

## 2) **Data Ordinal**

Skala ordinal memiliki peringkat, tapi tidak ada jarak posisional objektif antar angka karena angka yang tercipta bersifat relatif subjektif. Skala ini menjadi dasar dalam Skala Likert.

Contoh Data Variabel :

- a. Sangat Tidak Setuju = 1      Tidak Setuju = 2  
    Tidak Tahu = 3
- b. Setuju = 4      Sangat Setuju = 5
- c. Pendek = 1      Sedang = 2  
    Tinggi = 3

d. Tidak enak = 1

Ragu-ragu = 2

Enak = 3

Analisis Statistik :

Angka 1 lebih rendah dari angka 2 dalam peringkat, tapi tidak bisa dilakukan operasi matematika. Data ordinal menggunakan statistik non-parametrik mencakup frekuensi, median dan modus, Spearman rank-order correlation dan analisis varian.

Sehingga data ordinal merupakan data yang menunjuk pada tingkatan atau penjenjangan pada sesuatu keadaan. Berbeda dengan data nominal yang menunjukkan adanya perbedaan secara kategorik, data ordinal juga memiliki sifat adanya perbedaan di antara obyek yang dijenjangkan. Namun dalam perbedaan tersebut terdapat suatu kedudukan yang dinyatakan sebagai suatu urutan bahwa yang satu lebih besar atau lebih tinggi daripada yang lainnya. Kriteria urutan dari yang paling tinggi ke yang yang paling rendah dinyatakan dalam bentuk posisi relatif atau kedudukan suatu kelompok.

Contoh lain data ini misalnya:

a. prestasi belajar siswa diklasifikasikan menjadi kelompok “baik”, “cukup”, dan “kurang”, atau ukuran tinggi seseorang dengan “tinggi”, “sedang”, dan “pendek”

b. Hasil ujian mahasiswa peserta kuliah Statistik Pendidikan Budiman memperoleh skor 90, Rahmat 85, Musyafak 75, dan Mahsunah 65. Berdasarkan skor-skor tersebut dibuatlah suatu jenjang (rangking), sehingga terjadilah urutan jenjang ke 1 (90), ke 2 (85), ke 3 (75), dan ke 4 (65). Data ordinal memiliki harga mutlak (dapat diperbandingkan) dan selisih perbedaan antara urutan-urutan yang berdekatan bisa tidak sama.

### 3) Data Interval

Skala interval adalah skala ordinal yang memiliki poin jarak objektif dalam keteraturan kategori peringkat, tapi jarak yang tercipta sama antar masing-masing angka.

Contoh Data Variabel :

Data Umur

- a. Umur 20-30 tahun = 1
- b. Umur 31-40 tahun = 2
- c. Umur 41-50 tahun = 3

Data Suhu

- a. Suhu 0-50 Celsius = 1
- b. Suhu 51-100 Celsius = 2
- c. Suhu 101-150 Celsius = 3

Analisis Statistik :

Angka 3 berarti lebih tua atau lebih panas dari angka 2 setara dengan angka 2 terhadap angka 1, bisa

operasi penjumlahan dan pengurangan. Statistik parametrik yaitu deviasi mean dan standar, korelasi  $r$ , regresi, analisis varian dan analisis faktor ditambah berbagai multivariat.

Data interval tergolong data kontinum yang mempunyai tingkatan yang lebih tinggi lagi dibandingkan dengan data ordinal karena mempunyai tingkatan yang lebih banyak lagi. Data interval menunjukkan adanya jarak antara data yang satu dengan yang lainnya.

Contoh lain data interval misalnya hasil ujian, hasil pengukuran berat badan, hasil pengukuran tinggi badan, dan lainnya. Satu hal yang perlu diperhatikan bahwa data interval tidak dikenal adanya nilai 0 (nol) mutlak. Dalam hasil pengukuran (tes) misalnya mahasiswa mendapat nilai 0. Angka nol ini tidak dapat diartikan bahwa mahasiswa tersebut benar-benar tidak bisa apa-apa. Meskipun ia memperoleh nilai nol ia memiliki suatu pengetahuan atau kemampuan dalam matakuliah yang bersangkutan. Nilai nol yang diberikan oleh dosen sebetulnya hanya merupakan atribut belaka hanya saja pada saat ujian, pertanyaan yang diujikan tidak pas seperti yang dipersiapkannya. Atau jawaban yang diberikan tidak sesuai dengan yang dikehendaki soal.

#### 4) Data Rasio

Skala rasio adalah skala interval yang memiliki nol mutlak.

Contoh Data Variabel :

0 tahun, 1 tahun, 2 tahun, 3 tahun, ..... dst.

a. ....  $-3C$ ,  $-2C$ ,  $-1C$ ,  $0C$ ,  $1C$ ,  $2C$ ,  $3C$ , ..... dst.

b. ....  $0,71m$  .....  $5,38m$  .....  $12,42m$  ..... dst.

Analisis Statistik :

Berlaku semua operasi matematika. Analisis statistik sama dengan skala interval. Data rasio merupakan data yang tergolong ke dalam data kontinum juga tetapi yang mempunyai ciri atau sifat tertentu. Data ini memiliki sifat interval atau jarak yang sama seperti halnya dalam skala interval. Namun demikian, skala rasio masih memiliki ciri lain. Pertama harga rasio memiliki harga nol mutlak, artinya titik nol benar-benar menunjukkan tidak adanya suatu ciri atau sifat. Misalnya titik nol pada skala sentimeter menunjukkan tidak adanya panjang atau tinggi sesuatu. Kedua angka skala rasio memiliki kualitas bilangan riil yang berlaku perhitungan matematis.

Contohnya : berat badan Rudi 70 kg, sedangkan Saifullah 35 kg. Keadaan ini dapat dirasiokan bahwa berat badan Rudi dua kali berat badan Saifullah. Atau berat badan Saifullah separuh dari berat badan

Rudi. Berbeda dengan data interval misalnya Rudi ujian dapat 70 sementara Saifullah memperoleh 30. Hal ini tidak dapat diartikan bahwa kepandaian Rudi dua kali lipat kepandaian Saifullah.

Data rasio dalam ilmu-ilmu sosial jarang dipergunakan, bahkan hampir tidak pernah dipergunakan. Lapangan penggunaan data berskala rasio ini lebih banyak berada dalam bidang ilmu-ilmu eksakta terutama fisika.

Secara umum perbedaan kriteria antara skala data nominal, ordinal, interval dan rasio dapat digambarkan sebagai berikut:

Kemampuan	Nominal	Ordinal	Interval	Rasio
Frekuensi	√	√	√	√
Modus & median		√	√	√
Mengurutkan nilai		√	√	√
Membedakan kuantitas antar nilai			√	√
Penambahan & Pengurangan			√	√
Perkalian & pembagian				√
Nol mutlak				√

## 2.3 Sumber data statistik

### Metode Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian. Tujuan yang diungkapkan dalam bentuk hipotesis merupakan jawaban sementara terhadap pertanyaan penelitian. metode pengumpulan

data bisa dilakukan dengan cara:

### 1) Wawancara

**Menurut Prabowo (1996)** wawancara adalah metode pengambilan data dengan cara menanyakan sesuatu kepada seseorang responden, caranya adalah dengan bercakap-cakap secara tatap muka. Pada penelitian ini wawancara akan dilakukan dengan menggunakan pedoman wawancara.

Menurut Patton dalam proses wawancara dengan menggunakan pedoman umum wawancara ini, interview dilengkapi pedoman wawancara yang sangat umum, serta mencantumkan isu-isu yang harus diliput tanpa menentukan urutan pertanyaan, bahkan mungkin tidak terbentuk pertanyaan yang eksplisit.

Pedoman wawancara digunakan untuk mengingatkan interviewer mengenai aspek-aspek apa yang harus dibahas, juga menjadi daftar pengecek (check list) apakah aspek-aspek relevan tersebut telah dibahas atau ditanyakan. Dengan pedoman demikian interviwer harus memikirkan bagaimana pertanyaan tersebut akan dijabarkan secara kongkrit dalam kalimat Tanya, sekaligus menyesuaikan pertanyaan dengan konteks actual saat wawancara berlangsung (Patton dalam poerwandari, 1998).



## **2) Observasi**

Disamping wawancara, penelitian ini juga melakukan metode observasi. Menurut Nawawi & Martini (1991) observasi adalah pengamatan dan pencatatan secara sistematik terhadap unsur-unsur yang tampak dalam suatu gejala atau gejala-gejala dalam objek penelitian.

Dalam penelitian ini observasi dibutuhkan untuk dapat memahami proses terjadinya wawancara dan hasil wawancara dapat dipahami dalam konteksnya. Observasi yang akan dilakukan adalah observasi terhadap subjek, perilaku subjek selama wawancara, interaksi subjek dengan peneliti dan hal-hal yang dianggap relevan sehingga dapat memberikan data tambahan terhadap hasil wawancara.

Menurut Patton (dalam Poerwandari 1998) tujuan observasi adalah mendeskripsikan setting yang dipelajari, aktivitas-aktivitas yang berlangsung, orang-orang yang terlibat dalam aktivitas, dan makna kejadian di lihat dari perpektif mereka yang terlihat dalam kejadian yang diamati tersebut.

### **Macam-Macam Observasi**

#### **a) Observasi Partisipatif**

Peneliti mengamati apa yang dikerjakan

orang, mendengarkan apa yang diucapkan dan berpartisipasi dalam aktivitas yang diteliti

**b)Observasi Terus Terang atau Tersamar**

Peneliti berterus terang kepada narasumber bahwa ia sedang melakukan penelitian.

**c)Observasi tak Berstruktur**

Dilakukan dengan tidak Berstruktur karena fokus penelitian belum jelas

- .Angket atau kuesioner (questionnaire)  
Angket atau kuesioner merupakan suatu teknik pengumpulan data secara tidak langsung (peneliti tidak langsung bertanya jawab dengan responden). Instrumen atau alat pengumpulan datanya juga disebut angket berisi sejumlah pertanyaan-pertanyaan yang harus dijawab atau direspon oleh responden. Responden mempunyai kebiasaan untuk memberikan jawaban atau respon sesuai dengan persepsinya.

Kuesioner merupakan metode penelitian yang harus dijawab responden untuk menyatakan pandangannya terhadap suatu persoalan. Sebaiknya pertanyaan dibuat dengan bahasa sederhana yang mudah

dimengerti dan kalimat-kalimat pendek dengan maksud yang jelas. Penggunaan kuesioner sebagai metode pengumpulan data terdapat beberapa keuntungan, diantaranya adalah pertanyaan yang akan diajukan pada responden dapat distandarkan, responden dapat menjawab kuesioner pada waktu luangnya, pertanyaan yang diajukan dapat dipikirkan terlebih dahulu sehingga jawabannya dapat dipercaya dibandingkan dengan jawaban secara lisan, serta pertanyaan yang diajukan akan lebih tepat dan seragam.

#### **d) Macam-Macam Kuisisioner**

- **Kuesioner tertutup**

Setiap pertanyaan telah disertai sejumlah pilihan jawaban. Responden hanya memilih jawaban yang paling sesuai.

- **Kuesioner terbuka**

Dimana tidak terdapat pilihan jawaban sehingga responden haru memformulasikan jawabannya sendiri.

- **Kuesioner kombinasi terbuka dan tertutup**

Dimana pertanyaan tertutup kemudian

disusul dengan pertanyaan terbuka.

- **Kuesioner semi terbuka**

Pertanyaan yang jawabannya telah tersusun rapi, tetapi masih ada kemungkinan tambahan jawaban.

UWKSPRESS

## **BAB III**

### **PENYAJIAN DATA**

#### **3.1 Pengertian dan kegunaan**

Tujuan penyajian data adalah untuk dapat menyajikan data dengan tepat dalam arti betul dan benar Untuk dapat mempermudah / mempercepat membaca dan menarik kesimpulan agar lebih menarik dan menyenangkan

Data dapat disajikan dalam bentuk :

1. Naskah yaitu cara penyajian data yang ditulis secara narasi
2. Tabel/daftar distribusi frekuensi
3. Diagram / Gambar (dalam fungsinya:sebagai sebuah potret yang dapat memberikan gambaran serta uraian dari tempat atau objek dari mana gambar itu diambil), yang terdiri dari :
  - ✓ Diagram batang / diagram balok / bar chart
  - ✓ Diagram garis / line chart
  - ✓ Diagram pastel /diagram lingkaran /circle chart /pie chart
  - ✓ Diagram lambang / diagram simbol / pictograph
  - ✓ Diagram peta / cartogram / statistikal map

## 3.2 Tabel Distribusi Frekuensi

### 1) Tabel Distribusi Frekuensi

Data yang berukuran besar ( $n > 30$ ) lebih tepat disajikan dalam tabel distribusi frekuensi, yaitu cara penyajian data yang datanya disusun dalam kelas-kelas tertentu.

Langkah-langkah penyusunan tabel distribusi frekuensi adalah sebagai berikut.

- Langkah ke-2 menentukan banyak interval ( $K$ ) dengan rumus "Sturges" yaitu:  $K = 1 + 3,3 \log n$  dengan  $n$  adalah banyak data. Banyak kelas harus merupakan bilangan bulat positif hasil pembulatan.
- Langkah ke-3 menentukan panjang interval kelas ( $I$ ) dengan menggunakan rumus:

$$I = J/K$$

- Langkah ke-4 menentukan batas-batas kelas. Data terkecil harus merupakan batas bawah interval kelas pertama atau data terbesar adalah batas atas interval kelas terakhir.
- Langkah ke-5 memasukkan data ke dalam kelas-kelas yang sesuai dan menentukan nilai frekuensi setiap kelas dengan sistem turus.
- Menuliskan turus-turus dalam bilangan yang bersesuaian dengan banyak turus.

Ingatlah :

Menentukan banyak kelas interval dengan aturan Sturges dimaksudkan agar interval tidak terlalu besar

sebab hasilnya akan menyimpang dari keadaan sesungguhnya. Sebaiknya, jika interval terlalu kecil, hasilnya tidak menggambarkan keadaan yang diharapkan.

Contoh Soal 1 :

Seorang peneliti mengadakan penelitian tentang berat badan dari 35 orang.

Data hasil penelitian itu (dalam kg) diberikan berikut ini:

48 32 46 27 43 46 25 41 40 58 16 36

21 42 47 55 60 58 46 44 63 66 28 56

50 21 56 55 25 74 43 37 51 53 39

Sajikan data tersebut ke dalam tabel distribusi frekuensi.

Jawaban :

1. Jangkauan ( $J$ ) =  $X_m - X_n = 74 - 16 = 58$ .
2. Banyak kelas ( $K$ ) =  $1 + 3,3 \log n = 1 + 3,3 \log 35 = 6,095$ . Banyak kelas dibulatkan menjadi "6".
3. Panjang interval kelas ( $I$ ) adalah  $I = J/K = 58/6 = 9,67$ . Panjang interval kelas dibulatkan menjadi "10".

Dengan panjang interval kelas = 10 dan banyak kelas = 6, diperoleh tabel distribusi frekuensi seperti pada Tabel 4 atau Tabel 5

Cara I:

Batas bawah kelas pertama diambil datum terkecil. Amati Tabel 4. Dari tabel tersebut tampak bahwa frekuensi paling banyak dalam

interval 46–55. Artinya, berat badan kebanyakan berkisar antara 46 kg dan 55 kg.

Tabel 4. Tabel distribusi frekuensi

<b>Interval Kelas</b>	<b>Turus</b>	<b>Frekuensi</b>
16–25	E	5
26–35	C	3
36–45	ED	9
46–55	EE	10
56–65	EA	6
66–75	B	2
	Jumlah	35

Cara II:

Batas atas kelas terakhir diambil datum terbesar.

Amati Tabel 5.

Tabel 5. Tabel distribusi frekuensi

<b>Interval Kelas</b>	<b>Turus</b>	<b>Frekuensi</b>
15–24	C	3
25–34	E	5
35–44	ED	9
45–54	EC	8
55–64	EC	8
65–74	B	2
	Jumlah	35



Dari tabel tampak frekuensi paling sedikit dalam interval 65–74. Artinya, berat badan antara 65 kg dan 74 kg ada 2 orang. Perhatikan interval kelas yang pertama, yaitu 15–24. 15 disebut batas bawah dan 24 disebut batas atas. Ukuran 15–24 adalah hasil pembulatan, ukuran yang sebenarnya terletak pada 14,5–24,5. 14,5 disebut tepi bawah kelas (batas bawah nyata) dan 24,5 disebut tepi atas kelas (batas atas nyata) pada interval kelas 15–24.

Dalam menentukan tepi bawah kelas dan tepi atas kelas pada setiap interval kelas, harus diketahui satuan yang dipakai. Dengan demikian, untuk tepi bawah kelas adalah batas bawah kelas dikurangi  $\frac{1}{2}$  satuan ukuran. Jadi, tepi kelas dari interval kelas 15–24 menjadi 14,5–24,5.

## 2) Frekuensi Relatif dan Kumulatif

Frekuensi yang dimiliki setiap kelas pada tabel distribusi frekuensi bersifat mutlak. Adapun frekuensi relatif dari suatu data adalah dengan membandingkan frekuensi pada interval kelas itu dengan banyak data dinyatakan dalam persen. Contoh: interval frekuensi kelas adalah 20. Total data seluruh interval kelas = 80 maka frekuensi relatif kelas ini adalah  $\frac{20}{80} = \frac{1}{4}$ , sedangkan frekuensi relatifnya adalah  $\frac{1}{4} \times 100\% = 25\%$ .

Dari uraian tersebut, dapatkah Anda menyatakan rumus frekuensi relatif? Cobalah nyatakan rumus frekuensi relatif dengan kata-kata Anda sendiri.

Frekuensi relatif dirumuskan sebagai berikut.

**Frekuensi relatif kelas ke-k = frekuensi kelas ke-k / banyak data**

Frekuensi kumulatif kelas ke-k adalah jumlah frekuensi pada kelas yang dimaksud dengan frekuensi kelas-kelas sebelumnya.

Ada dua macam frekuensi kumulatif, yaitu

- 1) frekuensi kumulatif "kurang dari" ("kurang dari" diambil terhadap tepi atas kelas);
- 2) frekuensi kumulatif "lebih dari" ("lebih dari" diambil terhadap tepi bawah kelas).

**Tepi atas = batas atas +  $\frac{1}{2}$  satuan pengukuran**

**Tepi bawah = batas bawah -  $\frac{1}{2}$  satuan pengukuran**

Contoh Soal 5 :

Dari Tabel 4. untuk interval kelas 46 – 55 (kelas 4), hitunglah

- a. frekuensi relatif;
- b. frekuensi kumulatif "kurang dari";
- c. frekuensi kumulatif "lebih dari".

Penyelesaian :

- a. Frekuensi relatif kelas ke-4 = (frekuensi kelas ke-4 / banyak datum) x 100 % =  $10/35 \times 100\% = 28,57\%$
- b. Frekuensi kumulatif "kurang dari" untuk interval kelas 46 – 55

$$= 5 + 3 + 9 + 10 = 27 \text{ (kurang dari tepi atas kelas } 55,5)$$

- c. Frekuensi kumulatif "lebih dari" untuk interval kelas 46 – 55
- $$= 10 + 6 + 2 = 18 \text{ (lebih dari tepi bawah kelas } 45,5).$$

Misalkan, hasil ulangan Bahasa Indonesia 37 siswa kelas XI SMA 3 disajikan dalam tabel di bawah. Penyajian data pada Tabel 1.1 dinamakan penyajian data sederhana. Dari tabel 1.1, Anda dapat menentukan banyak siswa yang mendapat nilai 9, yaitu sebanyak 7 orang. Berapa orang siswa yang mendapat nilai 5? Nilai berapakah yang paling banyak diperoleh siswa?

Jika data hasil ulangan bahasa Indonesia itu disajikan dengan cara mengelompokkan data nilai siswa, diperoleh tabel frekuensi berkelompok seperti pada Tabel 2. Tabel 2. dinamakan Tabel Distribusi Frekuensi.

Tabel 1. Penyajian data sederhana

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

Tabel 2. Tabel Distribusi Frekuensi

Interval Kelas	Turus	Frekuensi
1–2	EB	7
3–4	C	3
5–6	EC	8
7–8	EE	10
9–10	EC	8
	Jumlah	37

### 3.3 Penyajian Data dalam Bentuk Diagram

Kerap kali data yang disajikan dalam bentuk tabel sulit untuk dipahami. Lain halnya jika data tersebut disajikan dalam bentuk diagram maka Anda akan dapat lebih cepat memahami data itu. Diagram adalah gambar yang menyajikan data secara visual yang biasanya berasal

dari tabel yang telah dibuat. Meskipun demikian, diagram masih memiliki kelemahan, yaitu pada umumnya diagram tidak dapat memberikan gambaran yang lebih detail.

#### a. Diagram Batang

Diagram batang biasanya digunakan untuk menggambarkan data diskrit (data cacahan). Diagram batang adalah bentuk penyajian data statistik dalam bentuk batang yang dicatat dalam interval tertentu pada bidang cartesius.

Ada dua jenis diagram batang, yaitu

1. diagram batang vertikal, dan
2. diagram batang horizontal.

#### Contoh Soal 1 :

Selama 1 tahun, toko "Anggo" mencatat keuntungan setiap bulan sebagai berikut.

Tabel 3. Keuntungan Toko "Anggo" per Bulan (dalam jutaan rupiah)

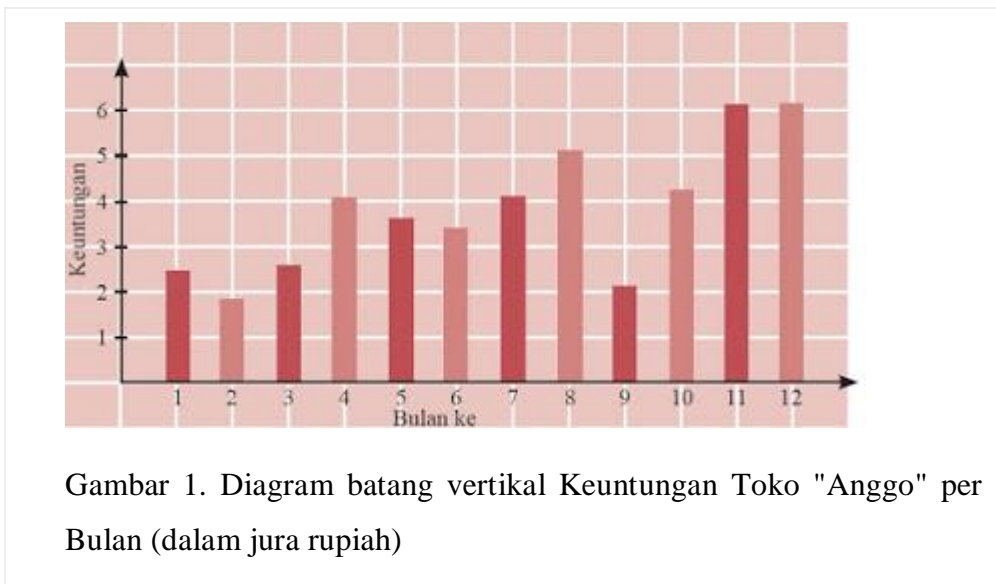
Keuntungan	2,5	1,8	2,6	4,2	3,5	3,3	4,0	5,0	2,0	4,2	6,2	6,2
Bulan ke	1	2	3	4	5	6	7	8	9	10	11	12

- a. Buatlah diagram batang vertikal dari data tersebut.
- b. Berapakah keuntungan terbesar yang diperoleh Toko "Anggo" selama 1 tahun?

c. Kapan Toko "Anggo" memperoleh keuntungan yang sama selama dua bulan berturut-turut?

**Penyelesaian :**

a. **Diagram batang** vertikal dari data tersebut, tampak pada gambar berikut.



Gambar 1. Diagram batang vertikal Keuntungan Toko "Anggo" per Bulan (dalam jura rupiah)

Dari diagram tersebut tampak bahwa keuntungan terbesar yang diperoleh Toko "Anggo" selama 1 tahun adalah sebesar Rp 6.200.000,00.

Toko "Anggo" memperoleh keuntungan yang sama selama dua bulan berturut-turut pada bulan ke-11 dan ke-12.

**b. Diagram Garis**

Pernahkah Anda melihat grafik nilai tukar dolar terhadap rupiah atau pergerakan saham di TV? Grafik

yang seperti itu disebut diagram garis. Diagram garis biasanya digunakan untuk menggambarkan data tentang keadaan yang berkesinambungan (sekumpulan data kontinu). Misalnya, jumlah penduduk setiap tahun, perkembangan berat badan bayi setiap bulan, dan suhu badan pasien setiap jam.

Seperti halnya diagram batang, diagram garis pun memerlukan sistem sumbu datar (horizontal) dan sumbu tegak (vertikal) yang saling berpotongan tegak lurus. Sumbu mendatar biasanya menyatakan jenis data, misalnya waktu dan berat. Adapun sumbu tegaknya menyatakan frekuensi data. Langkah-langkah yang dilakukan untuk membuat diagram garis adalah sebagai berikut.

1. Buatlah suatu koordinat (berbentuk bilangan) dengan sumbu mendatar menunjukkan waktu dan sumbu tegak menunjukkan data pengamatan.
2. Gambarlah titik koordinat yang menunjukkan data pengamatan pada waktu  $t$ .
3. Secara berurutan sesuai dengan waktu, hubungkan titik-titik koordinat tersebut dengan garis lurus.

**Contoh Soal 2 :**

Berikut ini adalah tabel berat badan seorang bayi yang dipantau sejak lahir sampai berusia 9 bulan.

Berat Badan (kg)	3,5	4	5,2	6,4	6,8	7,5	7,5	8	8,8	8,6
Usia (bulan)	0	1	2	3	4	5	6	7	8	9

- Buatlah diagram garisnya.
- Pada usia berapa bulan berat badannya menurun?
- Pada usia berapa bulan berat badannya tetap?

**Pembahasan :**

- Langkah ke-1

Buatlah sumbu mendatar yang menunjukkan usia anak (dalam bulan) dan sumbu tegak yang menunjukkan berat badan anak (dalam kg).

- Langkah ke-2

Gambarlah titik koordinat yang menunjukkan data pengamatan pada waktu  $t$  bulan.

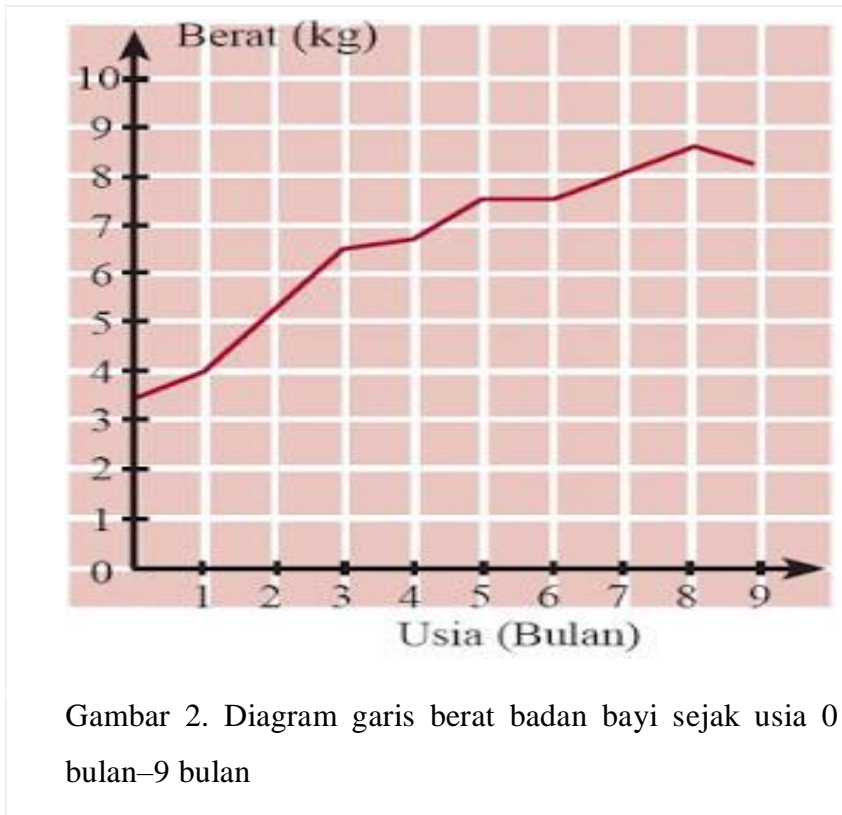
- Langkah ke-3

Secara berurutan sesuai dengan waktu, hubungkan titik-titik koordinat tersebut dengan garis lurus.

- Dari diagram tersebut dapat dilihat bahwa berat badan bayi menurun pada usai 8 sampai 9 bulan.
- Berat badan bayi tetap pada usia 5 sampai 6 bulan. Darimana Anda memperoleh hasil ini? Jelaskan.



Dari ketiga langkah tersebut, diperoleh diagram garis dari data tersebut tampak pada Gambar 2.



### Observasi: Interpolasi dan Ekstrapolasi Data

Anda dapat melakukan observasi terhadap kecenderungan data yang disajikan pada suatu diagram garis. Dari observasi ini, Anda dapat membuat perkiraan-perkiraan dengan cara interpolasi dan ekstrapolasi. Hal ini ditempuh dengan mengganti garis patah pada diagram garis menjadi garis lurus.

Interpolasi data adalah menaksir data atau memperkirakan data di antara dua keadaan (misalnya waktu) yang berurutan. Misalkan, dari gambar grafik Contoh soal 2. dapat diperkirakan berat badan bayi pada usia 5,5 bulan. Coba Anda amati grafik tersebut, kemudian tentukan berat badan bayi pada usia 5,5 bulan.

Ekstrapolasi data adalah menaksir atau memperkirakan data untuk keadaan (waktu) mendatang. Cara yang dapat dilakukan untuk ekstrapolasi adalah dengan memperpanjang ruas garis terujung ke arah kanan. Misalkan, dari gambar grafik soal 2. dapat diperkirakan berat badan bayi pada usia 10 bulan. Jika garis lurus sudah ditentukan, Anda dapat menentukan interpolasi data. Untuk ekstrapolasi data, Anda harus berhati-hati. Menurut diagram garis, berapa kira-kira berat badan bayi pada usia 10 bulan? Berikan alasan Anda.

### **c. Diagram Lingkaran**

Untuk mengetahui perbandingan suatu data terhadap keseluruhan, suatu data lebih tepat disajikan dalam bentuk diagram lingkaran. Diagram lingkaran adalah bentuk penyajian data statistika dalam bentuk lingkaran yang dibagi menjadi beberapa juring lingkaran.

Langkah-langkah untuk membuat diagram lingkaran adalah sebagai berikut.

1. Buatlah sebuah lingkaran pada kertas.
2. Bagilah lingkaran tersebut menjadi beberapa juring lingkaran untuk menggambarkan kategori yang datanya
3. telah diubah ke dalam derajat.

Agar lebih jelasnya, pelajarilah contoh berikut.

**Contoh Soal 3 :**

Tabel berikut menunjukkan banyaknya siswa di suatu kabupaten menurut tingkat sekolah pada tahun 2017.

Tingkat Pendidikan	Banyaknya Siswa
SD	175
SMP	600
SMA	225

- a. Buatlah diagram lingkaran untuk data tersebut.
- b. Berapa persen siswa yang menyelesaikan sekolah sampai pada tingkat SMP?
- c. Berapa persen siswa yang menyelesaikan sekolah sampai pada tingkat SMA?

**Pembahasan :**

- a. Jumlah seluruh siswa adalah 1.000 orang. Seluruh siswa diklasifikasikan menjadi 5 katagori: SD = 175 orang, SMP = 600 orang, dan SMA = 225 orang.

- Siswa SD =  $(175/1.000) \times 100\% = 17,5\%$

Besar sudut sektor lingkaran =  $17,5\% \times 360^\circ = 63^\circ$

- Siswa SMP =  $(600/1.000) \times 100\% = 60\%$

Besar sudut sektor lingkaran =  $60\% \times 360^\circ = 216^\circ$

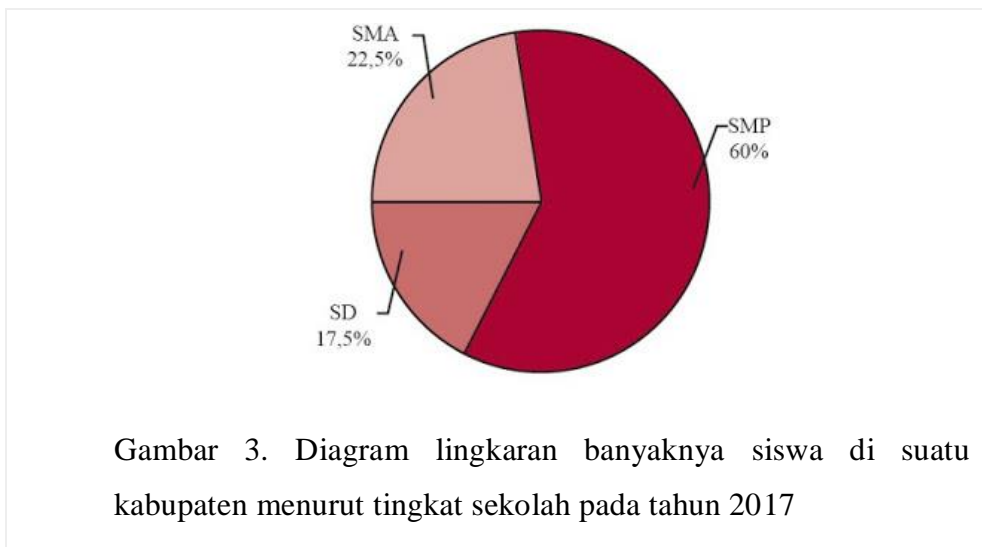
• Siswa SMA =  $(225/1.000) 100\% = 22,5\%$

Besar sudut sektor lingkaran =  $22,5\% \times 360^\circ = 81^\circ$

b. Persentase siswa yang menyelesaikan sekolah sampai pada tingkat SMP adalah 60%.

c. Persentase siswa yang menyelesaikan sekolah sampai pada tingkat SMA adalah 22,5%.

Diagram lingkaran ditunjukkan pada Gambar 3.



Gambar 3. Diagram lingkaran banyaknya siswa di suatu kabupaten menurut tingkat sekolah pada tahun 2017

### 3.4 Histogram, Poligon Frekuensi, dan Ogive

Histogram merupakan diagram frekuensi bertangga yang bentuknya seperti diagram batang. Batang yang berdekatan harus berimpit. Untuk pembuatan histogram, pada setiap interval kelas diperlukan tepi-tepi kelas. Tepi-tepi kelas ini digunakan untuk menentukan titik tengah kelas yang dapat ditulis sebagai berikut.

Titik tengah kelas =  $\frac{1}{2}$  (tepi atas kelas + tepi bawah kelas)

Poligon frekuensi dapat dibuat dengan menghubungkan titik-titik tengah setiap puncak persegi panjang dari histogram secara berurutan. Agar poligon "tertutup" maka sebelum kelas paling bawah dan setelah kelas paling atas, masing-masing ditambah satu kelas.

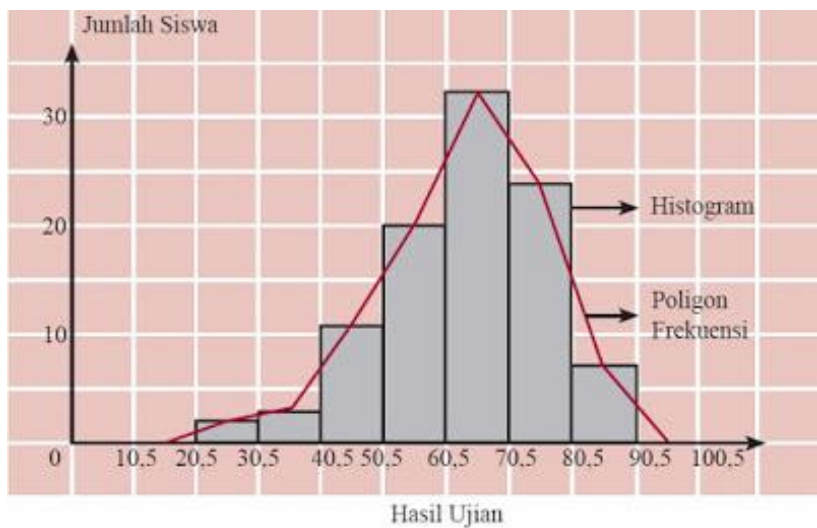
**Contoh Soal 6 :**

Tabel distribusi frekuensi hasil ujian matematika Kelas XI SMA Cendekia di Kalimantan Barat diberikan pada Tabel 6. Buatlah histogram dan poligon frekuensinya.

Tabel 6. Tabel distribusi frekuensi hasil ujian matematika Kelas XI SMA Cendekia di Kalimantan Barat

Kelas Interval	Frekuensi
21–30	2
31–40	3
41–50	11
51–60	20
61–70	33
71–80	24
81–90	7
	100

Jawaban :



Gambar 4. Histogram hasil ujian matematika Kelas XI SMA Cendekia di Kalimantan Barat.

Dari histogram tersebut tampak bahwa kebanyakan siswa memperoleh nilai antara 60,5 dan 70,5. Coba Anda ceritakan hal lain dari histogram tersebut.

#### d. Ogive (Ogif)

Grafik yang menunjukkan frekuensi kumulatif kurang dari atau frekuensi kumulatif lebih dari dinamakan poligon kumulatif.

Untuk populasi yang besar, poligon mempunyai banyak ruas garis patah yang menyerupai kurva sehingga poligon frekuensi kumulatif dibuat mulus, yang hasilnya disebut ogif.

Ada dua macam ogif, yaitu sebagai berikut.

- a. Ogif dari frekuensi kumulatif kurang dari disebut ogif positif.
- b. Ogif dari frekuensi kumulatif lebih dari disebut ogif negatif.

Contoh Soal 7 :

Tabel 7. dan 8. berturut-turut adalah tabel distribusi frekuensi kumulatif "kurang dari" dan "lebih dari" tentang nilai ulangan Biologi Kelas XI SMA 3.

Tabel 7. Tabel distribusi frekuensi kumulatif "kurang dari" tentang nilai ulangan Biologi Kelas XI SMA 3.

Nilai	Frekuensi
< 20,5	0
< 30,5	2
< 40,5	5
< 50,5	16
< 60,5	36
< 70,5	69
< 80,5	93
< 90,5	100

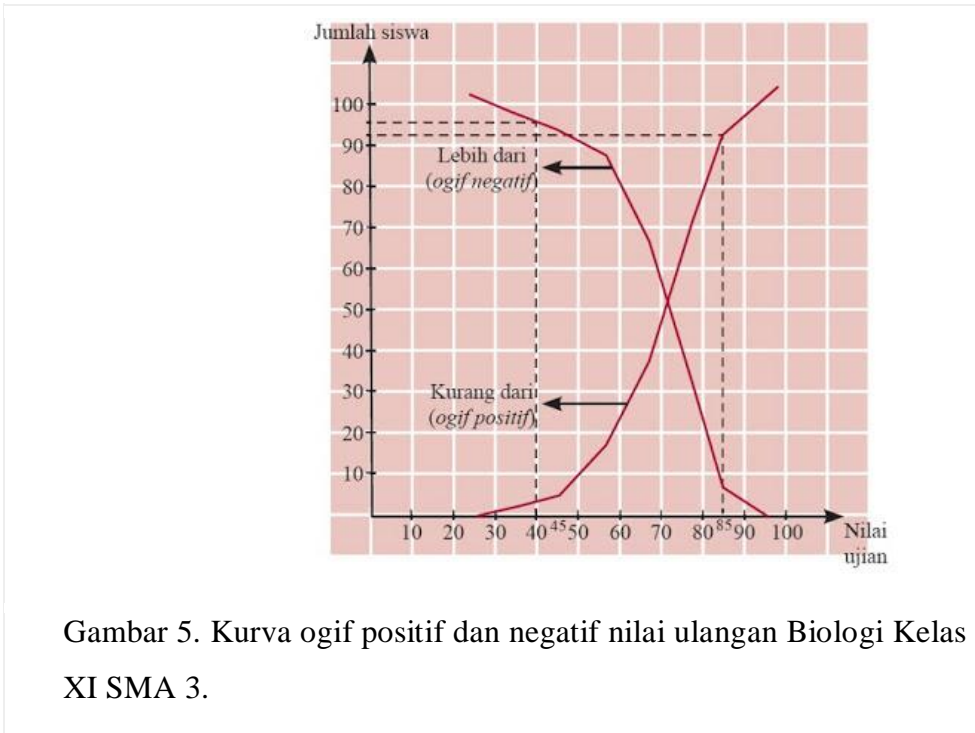
Tabel 8. Tabel distribusi frekuensi kumulatif "lebih dari" tentang nilai ulangan Biologi Kelas XI SMA 3.

<b>Nilai</b>	<b>Frekuensi</b>
> 20,5	100
> 30,5	98
> 40,5	95
> 50,5	84
> 60,5	64
> 70,5	31
> 80,5	7
> 90,5	0

- a. Buatlah ogif positif dan ogif negatif dari tabel tersebut.
- b. Berapakah jumlah siswa yang mempunyai nilai Biologi kurang dari 85?
- c. Berapakah jumlah siswa yang mempunyai berat badan lebih dari 40?



Pembahasan :



Gambar 5. Kurva ogif positif dan negatif nilai ulangan Biologi Kelas XI SMA 3.

- Ogif positif dan ogif negatif dari tabel tersebut tampak pada gambar 5.
- Dari kurva ogif positif, tampak siswa yang mempunyai nilai kurang dari 85 adalah sebanyak 93 orang.
- Dari kurva ogif negatif, tampak siswa yang mempunyai nilai lebih dari 40 adalah sebanyak 96 orang.

## BAB IV

# UKURAN PEMUSATAN DAN LETAK DATA

## 4.1 Pengertian Nilai Sentral

Nilai Sentral suatu rangkaian data merupakan nilai dalam rangkaian data yang dapat mewakili rangkaian data tersebut. Salah satu aspek yang paling penting untuk menggambarkan distribusi data adalah nilai pusat data pengamatan (*Central Tendency*). Setiap pengukuran aritmatika yang ditujukan untuk menggambarkan suatu nilai yang mewakili nilai pusat atau nilai sentral dari suatu gugus data (himpunan pengamatan) dikenal sebagai **ukuran pemusatan data (*tendensi sentral*)**. Terdapat tiga ukuran pemusatan data yang sering digunakan, yaitu:

1. Mean (Rata-rata hitung/rata-rata aritmetika)
2. Median
3. Mode/Modus

yang dalam pengembangannya dalam pemusatan data selain mengukur rata-rata aritmetika, juga dilakukan pengukuran:

4. Rata-rata Ukur (Geometric Mean)
5. Rata-rata Harmoni (Harmonic Mean)
6. Rata-rata Kuadrat (Quadratic Mean)

Pada bab ini akan di bahas mengenai pengertian beberapa **ukuran pemusatan data (Mean, Median, Mode/Modus)** yang dilengkapi dengan contoh perhitungan, baik untuk data tunggal ataupun data yang sudah dikelompokkan dalam tabel distribusi frekuensi. Selain

ukuran statistik di atas, akan dibahas juga mengenai beberapa ukuran statistik lainnya, seperti **Rata-rata Ukur** (*Geometric Mean*), **Rata-rata Harmonik** (H) serta beberapa karakteristik penting yang perlu dipahami untuk ukuran tendensi sentral yang baik serta bagaimana memilih atau menggunakan nilai tendensi sentral yang tepat.

### 1) Mean (*arithmetic mean*)

Rata-rata hitung atau arithmetic mean atau sering disebut dengan istilah mean saja merupakan metode yang paling banyak digunakan untuk menggambarkan ukuran tendensi sentral. Mean dihitung dengan menjumlahkan semua nilai data pengamatan kemudian dibagi dengan banyaknya data. Definisi tersebut dapat dinyatakan dengan persamaan berikut:

#### a. Data tidak berkelompok

Sampel:  $n$

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

Populasi:  $N$

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

Keterangan:

$\Sigma$  = lambang penjumlahan semua gugus data pengamatan

$n$  = banyaknya sampel data

$N$  = banyaknya data populasi

$\bar{x}$  = nilai rata-rata sampel

$\mu$  = nilai rata-rata populasi

Mean dilambangkan dengan  $\bar{x}$  (dibaca "x-bar") jika kumpulan data ini merupakan contoh (sampel) dari populasi, sedangkan jika semua data berasal dari populasi, mean dilambangkan dengan  $\mu$  (huruf kecil Yunani mu).

Sampel statistik biasanya dilambangkan dengan huruf "Inggris" atau ditulis dengan huruf "kecil", sementara parameter-parameter populasi biasanya dilambangkan dengan huruf "Yunani", misalnya  $\mu$ , atau huruf "besar".

### 1. Rata-rata hitung (Mean)

**Untuk data tunggal (data tidak dikelompokkan / data mentah)**

#### **Contoh 1:**

Hitunglah nilai rata-rata dari nilai ujian matematika kelas 3 SMU berikut ini: 2; 4; 5; 6; 6; 7; 7; 7; 8; 9

**Jawab:**

$$2 + 4 + 5 + 6 + 6 + 7 + 7 + 7 + 8 + 9 / 10 = 61/10 = 6,1$$

**Untuk data berkelompok (data dalam tabel frekuensi):**

Nilai rata-rata dari data yang sudah dikelompokkan bisa dihitung dengan menggunakan formula berikut:

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum f_i x_i}{\sum f_i}$$

Keterangan:

$\Sigma$  = lambang penjumlahan semua gugus data pengamatan

$f_i$  = frekuensi data ke-i

$n$  = banyaknya sampel data = nilai rata-rata sampel

**Contoh 2:**

Berapa rata-rata hitung pada tabel frekuensi berikut:

$x_i$	70	69	45	80	56
$f_i$	5	6	3	1	1

Catatan:

Tabel frekuensi pada tabel di atas merupakan tabel frekuensi untuk data tunggal, bukan tabel frekuensi dari data yang sudah dikelompokkan berdasarkan selang/kelas tertentu.

**Jawab:**

$x_i$	$f_i$	$f_i x_i$
70	5	350
69	6	414
45	3	135
80	1	80
56	1	56
Jumlah	16	1035

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$\bar{x} = \frac{1035}{16} = 64.6$$

## 2. Mean dari data distribusi Frekuensi atau dari gabungan:

Distribusi Frekuensi: Rata-rata hitung dari data yang sudah disusun dalam bentuk tabel distribusi frekuensi dapat ditentukan dengan menggunakan formula yang sama dengan formula untuk menghitung nilai rata-rata dari data yang sudah dikelompokkan, yaitu:

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Keterangan:

$\Sigma$  = lambang penjumlahan semua gugus data pengamatan

$f_i$  = frekuensi data ke- $i$

$\bar{x}$  = nilai rata-rata sampel

**Contoh 3:**

Tabel berikut ini adalah nilai ujian statistik 80 mahasiswa yang sudah disusun dalam tabel frekuensi. Berbeda dengan contoh 2, pada contoh ke-3 ini, tabel distribusi frekuensi dibuat dari data yang sudah dikelompokkan berdasarkan selang/kelas tertentu (banyak kelas = 7 dan panjang kelas = 10).

Kelas ke-	Nilai Ujian	$f_i$
1	31 - 40	2
2	41 - 50	3
3	51 - 60	5
4	61 - 70	13

5	71 - 80	24
6	81 - 90	21
7	91 - 100	12
	Jumlah	80

Jawab:

Buat daftar tabel berikut, tentukan nilai wakilnya ( $x_i$ ) dan hitung  $f_i x_i$ .

Kelas ke-	Nilai Ujian	$f_i$	$x_i$	$f_i x_i$
1	31 - 40	2	35.5	71.0
2	41 - 50	3	45.5	136.5
3	51 - 60	5	55.5	277.5
4	61 - 70	13	65.5	851.5
5	71 - 80	24	75.5	1812.0
6	81 - 90	21	85.5	1795.5
7	91 - 100	12	95.5	1146.0



Kelas ke-	Nilai Ujian	f <sub>i</sub>	x <sub>i</sub>	f <sub>i</sub> x <sub>i</sub>
	Jumlah	80		6090.0

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$\bar{x} = \frac{6090}{80} = 76.1$$

Catatan:

Pendekatan perhitungan nilai rata-rata hitung dengan menggunakan distribusi frekuensi kurang akurat dibandingkan dengan cara perhitungan rata-rata hitung dengan menggunakan data aktualnya. Pendekatan ini seharusnya hanya digunakan apabila tidak memungkinkan untuk menghitung nilai rata-rata hitung dari sumber data aslinya.

Rata-rata Gabungan atau rata-rata terboboti (Weighted Mean)

Rata-rata gabungan (disebut juga grand mean, pooled mean, atau rata-rata umum) adalah cara yang tepat untuk menggabungkan rata-rata hitung dari beberapa sampel.

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Contoh 4:

Tiga sub sampel masing-masing berukuran 10, 6, 8 dan rata-ratanya 145, 118, dan 162. Berapa rata-ratanya?

## 2) Median

Median dari  $n$  pengukuran atau pengamatan  $x_1, x_2, \dots, x_n$  adalah nilai pengamatan yang terletak di tengah gugus data setelah data tersebut diurutkan. Apabila banyaknya pengamatan ( $n$ ) ganjil, median terletak tepat ditengah gugus data, sedangkan bila ngenap, median diperoleh dengan cara interpolasi yaitu rata-rata dari dua data yang berada di tengah gugus data. Dengan demikian, median membagi himpunan pengamatan menjadi dua bagian yang sama besar, 50% dari pengamatan terletak di bawah median dan 50% lagi terletak di atas median. Median sering dilambangkan dengan  $\tilde{x}$  (dibaca "x-tilde") apabila sumber datanya berasal dari sampel  $\tilde{\mu}$  (dibaca " $\mu$ -tilde") untuk median populasi. Median tidak dipengaruhi oleh nilai-nilai aktual dari pengamatan melainkan pada posisi mereka. Prosedur untuk menentukan nilai median, pertama urutkan data terlebih dahulu, kemudian ikuti salah satu prosedur berikut ini:

- **Banyak data ganjil** → mediannya adalah nilai yang berada tepat di tengah gugus data

- **Banyak data genap** → mediannya adalah rata-rata dari dua nilai data yang berada di tengah gugus data

**a. Median data tunggal:**

Untuk menentukan median dari data tunggal, terlebih dulu kita harus mengetahui letak/posisi median tersebut. Posisi median dapat ditentukan dengan menggunakan formula beriku

$$Posisi\ Median = \frac{(n + 1)}{2}$$

dimana:

n = banyaknya data pengamatan.

➤ **Median apabila n ganjil:**

**Contoh 5:**

Hitunglah median dari nilai ujian matematika kelas 3 SMU berikut ini:

8; 4; 5; 6; 7; 6; 7; 7; 2; 9; 10

**Jawab:**

data: 8; 4; 5; 6; 7; 6; 7; 7; 2; 9; 10

setelah diurutkan: 2; 4; 5; 6; 6; 7; 7; 7; 8; 9;

10

banyaknya data (n) = 11

posisi Me =  $\frac{1}{2}(11+1) = 6$

jadi Median = 7 (data yang terletak pada urutan ke-6)

➤ **Median apabila n genap:**

**Contoh 6:**

Hitunglah median dari nilai ujian matematika kelas 3 SMU berikut ini:  
8; 4; 5; 6; 7; 6; 7; 7; 2; 9

**Jawab:**

data: 8; 4; 5; 6; 7; 6; 7; 7; 2; 9

setelah diurutkan: 2; 4; 5; 6; 6; 7; 7; 7; 8; 9

banyaknya data (n) = 10

posisi Me =  $\frac{1}{2}(10+1) = 5.5$

Data tengahnya: 6 dan 7

jadi Median =  $\frac{1}{2} (6+7) = 6.5$  (rata-rata dari 2 data yang terletak pada urutan ke-5 dan ke-6)

Nilai Ujian	2	4	5	6	6	7	7	7	8	9	10
Urutan data ke-	1	2	3	4	5	6	7	8	9	10	11

**b. Median dalam distribusi frekuensi:**

Formula untuk menentukan median dari tabel distribusi frekuensi adalah sebagai berikut:

b = batas bawah kelas median dari kelas selang yang mengandung unsur atau memuat nilai median

p = panjang kelas median

n = ukuran sampel/banyak data

f = frekuensi kelas median

F = Jumlah semua frekuensi dengan tanda kelas lebih kecil dari kelas median ( $\sum f_i$ )

**Contoh 7:**

Tentukan nilai median dari tabel distribusi frekuensi pada Contoh 3 di atas!

**Jawab:**

Kelas ke-	Nilai Ujian	$f_i$	$f_{kum}$	
1	31 - 40	2	2	
2	41 - 50	3	5	
3	51 - 60	5	10	
4	61 - 70	13	23	
5	71 - 80	24	47	←letak kelas median
6	81 - 90	21	68	
7	91 - 100	12	80	

Kelas ke-	Nilai Ujian	$f_i$	$f_{kum}$	
8	Jumlah	80		

Letak kelas median: Setengah dari seluruh data = 40, terletak pada kelas ke-5 (nilai ujian 71-80)

$$b = 70.5, p = 10$$

$$n = 80, f = 24$$

$$f = 24 \text{ (frekuensi kelas median)}$$

$$F = 2 + 3 + 5 + 13 = 23$$

### 3) Mode/Modus

Mode adalah data yang paling sering muncul/terjadi. Untuk menentukan modus, pertama susun data dalam urutan meningkat atau sebaliknya, kemudian hitung frekuensinya. Nilai yang frekuensinya paling besar (sering muncul) adalah modus. Modus digunakan baik untuk tipe data numerik atau pun data kategoris. Modus tidak dipengaruhi oleh nilai ekstrem. Beberapa kemungkinan tentang modus suatu gugus data:

Apabila pada sekumpulan data terdapat dua mode, maka gugus data tersebut dikatakan bimodal.

Apabila pada sekumpulan data terdapat lebih dari dua mode, maka gugus data tersebut dikatakan multimodal.

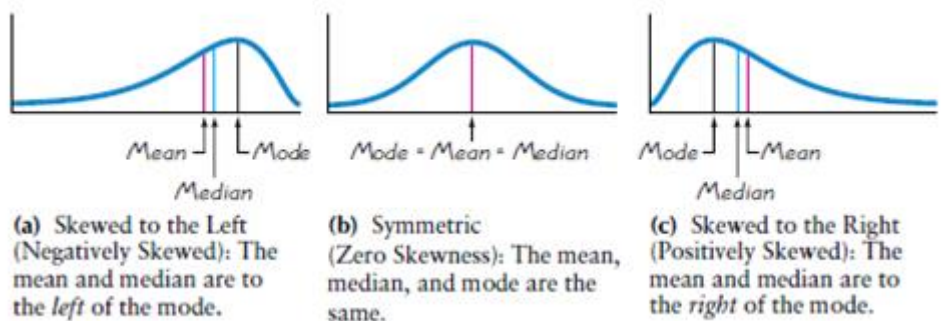
Apabila pada sekumpulan data tidak terdapat mode, maka gugus data tersebut dikatakan tidak mempunyai modus.

Meskipun suatu gugus data mungkin saja tidak memiliki modus, namun pada suatu distribusi data kontinyu, modus dapat ditentukan secara analitis.

Untuk gugus data yang distribusinya simetris, nilai mean, median dan modus semuanya sama.

Untuk distribusi miring ke kiri (negatively skewed):  
 $\text{mean} < \text{median} < \text{modus}$

untuk distribusi miring ke kanan (positively skewed):  
terjadi hal yang sebaliknya, yaitu  $\text{mean} > \text{median} > \text{modus}$ .



Hubungan antara ketiga ukuran tendensi sentral untuk data yang tidak berdistribusi normal, namun hampir simetris dapat didekati dengan menggunakan rumus empiris berikut:

$$\text{Mean} - \text{Mode} = 3 (\text{Mean} - \text{Median})$$

### a. Modus Data Tunggal:

#### Contoh 8:

Berapa modus dari nilai ujian matematika kelas 3 SMU berikut ini:

2, 4, 5, 6, 6, 7, 7, 7, 8, 9

2, 4, 6, 6, 6, 7, 7, 7, 8, 9

2, 4, 6, 6, 6, 7, 8, 8, 8, 9

2, 4, 5, 5, 6, 7, 7, 8, 8, 9

1, 2, 3, 4, 5, 6, 7, 8, 9, 10

#### Jawab:

2, 4, 5, 6, 6, 7, 7, 7, 8, 9 → Nilai yang sering muncul adalah angka 7 (frekuensi terbanyak = 3), sehingga Modus (M) = 7

2, 4, 6, 6, 6, 7, 7, 7, 8, 9 → Nilai yang sering muncul adalah angka 6 dan 7 (masing-masing muncul 3 kali), sehingga Modusnya ada dua, yaitu 6 dan 7. Gugus data tersebut dikatakan bimodal karena mempunyai dua modus. Karena ke-2 mode tersebut nilainya berurutan, mode sering dihitung dengan menghitung nilai rata-rata keduanya,  $\frac{1}{2}(6+7) = 6.5$ .

2, 4, 6, 6, 6, 7, 8, 8, 8, 9 → Nilai yang sering muncul adalah angka 6 dan 8 (masing-masing muncul 3 kali), sehingga Modusnya ada dua, yaitu 6 dan 8. Gugus data tersebut dikatakan bimodal



karena mempunyai dua modus. Nilai mode tunggal tidak dapat dihitung karena ke-2 mode tersebut tidak berurutan.

2, 4, 5, 5, 6, 7, 7, 8, 8, 9 → Nilai yang sering muncul adalah angka 5, 6 dan 7 (masing-masing muncul 2 kali), sehingga Modusnya ada tiga, yaitu 5, 6 dan 7. Gugus data tersebut dikatakan multimodal karena modusnya lebih dari dua.

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 → Pada gugus data tersebut, semua frekuensi data sama, masing-masing muncul satu kali, sehingga gugus data tersebut dikatakan tidak mempunyai modusnya

#### **b. Mode dalam Distribusi Frekuensi:**

$$Mo = b + p \left( \frac{b_1}{b_1 + b_2} \right)$$

dimana:

Mo = modal = kelas yang memuat modus

b = batas bawah kelas modal

p = panjang kelas modal

b<sub>mo</sub> = frekuensi dari kelas yang memuat modus (yang nilainya tertinggi)

b<sub>1</sub> = b<sub>mo</sub> - b<sub>mo-1</sub> = frekuensi kelas modal - frekuensi kelas sebelumnya

b<sub>2</sub> = b<sub>mo</sub> - b<sub>mo+1</sub> = frekuensi kelas modal - frekuensi kelas sesudahnya

**Contoh 9:**

Tentukan nilai median dari tabel distribusi frekuensi pada Contoh 3 di atas!

Jawab:

Kelas ke-	Nilai Ujian	$f_i$	
1	31 - 40	2	
2	41 - 50	3	
3	51 - 60	5	
4	61 - 70	13	
			$\rightarrow b_1 = (24 - 13) = 11$
5	71 - 80	24	$\leftarrow$ kelas modal (frekuensinya paling besar)
			$\rightarrow b_2 = (24 - 21) = 3$
6	81 - 90	21	
7	91 - 100	12	
8	Jumlah	80	

Kelas modul = kelas ke-5

$$b = 71 - 0.5 = 70.5$$

$$b_1 = 24 - 13 = 11$$

$$b_2 = 24 - 21 = 3$$

$$p = 10$$

Selain tiga ukuran tendensi sentral di atas (mean, median, dan mode), terdapat ukuran tendensi sentral lainnya, yaitu rata-rata ukur (Geometric Mean) dan rata-rata harmonis (Harmonic Mean)

#### 4) Rata-rata Ukur (Geometric Mean)

Untuk gugus data positif  $x_1, x_2, \dots, x_n$ , rata-rata geometrik adalah akar ke- $n$  dari hasil perkalian unsur-unsur datanya. Secara matematis dapat dinyatakan dengan formula berikut:

Dimana:

$U$  = rata-rata ukur (rata-rata geometrik)

$n$  = banyaknya sampel

$\Pi$  = Huruf kapital  $\pi$  ( $\pi$ ) yang menyatakan jumlah dari hasil kali unsur-unsur data.

Rata-rata geometrik sering digunakan dalam bisnis dan ekonomi untuk menghitung rata-rata tingkat perubahan, rata-rata tingkat pertumbuhan, atau rasio rata-rata untuk data berurutan tetap atau hampir tetap atau untuk rata-rata kenaikan dalam bentuk persentase.

**a. Rata-rata ukur untuk data tunggal**

**Contoh 10:**

Berapakah rata-rata ukur dari data 2, 4, 8?

**Jawab:**

$$U = \sqrt[3]{(2)(4)(8)} = \sqrt[3]{64} = 4$$

atau

$$\text{Log}(U) = \frac{\sum \log(x_i)}{n}$$

$$U = 10^{0.6021} = 4$$

**b. Distribusi Frekuensi:**

$$\text{Log}(U) = \frac{\sum(f_i \cdot \log(x_i))}{\sum f_i}$$

xi = tanda kelas (nilai tengah)

fi = frekuensi yang sesuai dengan xi

**Contoh 11:**

Tentukan rata-rata ukur dari tabel distribusi frekuensi pada Contoh 3 di atas!

**Jawab**

Kelas ke-	Nilai Ujian	fi	xi	log xi	fi.log xi
-----------	-------------	----	----	--------	-----------

1	31 - 40	2	35.5	1.5502	3.1005
2	41 - 50	3	45.5	1.6580	4.9740
3	51 - 60	5	55.5	1.7443	8.7215
4	61 - 70	13	65.5	1.8162	23.6111
5	71 - 80	24	75.5	1.8779	45.0707
6	81 - 90	21	85.5	1.9320	40.5713
7	91 - 100	12	95.5	1.9800	23.7600
8	Jumlah	80			149.8091

### 5) Rata-rata Harmonik (H)

Rata-rata harmonik dari suatu kumpulan data  $x_1, x_2, \dots, x_n$  adalah kebalikan dari nilai rata-rata hitung (aritmetik mean). Secara matematis dapat dinyatakan dengan formula berikut:

$$H = \frac{n}{\sum \left( \frac{1}{x_i} \right)}$$

Secara umum, rata-rata harmonic jarang digunakan. Rata-rata ini hanya digunakan untuk data yang bersifat khusus. Misalnya, rata-rata harmonik sering digunakan sebagai ukuran tendensi sentral untuk kumpulan data yang menunjukkan adanya laju perubahan, seperti kecepatan.

**a. Rata-rata harmonic untuk data tunggal**

**Contoh 12:**

Si A bepergian pulang pergi. Waktu pergi ia mengendarai kendaraan dengan kecepatan 10 km/jam, sedangkan waktu kembalinya 20 km/jam. Berapakah rata-rata kecepatan pulang pergi?

**Jawab:**

Apabila kita menghitungnya dengan menggunakan rumus jarak dan kecepatan, tentu hasilnya 13.5 km/jam! Apabila kita gunakan perhitungan rata-rata hitung, hasilnya tidak tepat!

$$\bar{x} = \frac{(10 + 20)}{2} = 15 \text{ km/jam}$$

Pada kasus ini, lebih tepat menggunakan rata-rata harmonik:

$$\bar{x} = \frac{2}{\frac{1}{10} + \frac{1}{20}} = \frac{40}{3} = 13.5 \text{ km/jam}$$

**b. Rata-rata Harmonik untuk Distribusi Frekuensi:**

$$H = \frac{\sum f_i}{\sum \left( \frac{f_i}{x_i} \right)}$$

**Contoh 13:**

Berapa rata-rata Harmonik dari tabel distribusi frekuensi pada Contoh 3 di atas!

**Jawab**

Kelas ke-	Nilai Ujian	fi	xi	fi/xi
1	31 - 40	2	35.5	0.0563
2	41 - 50	3	45.5	0.0659
3	51 - 60	5	55.5	0.0901
4	61 - 70	13	65.5	0.1985
5	71 - 80	24	75.5	0.3179
6	81 - 90	21	85.5	0.2456

Kelas ke-	Nilai Ujian	fi	xi	fi/xi
7	91 - 100	12	95.5	0.1257
8	Jumlah	80		1.1000

$$H = \frac{\sum f_i}{\sum \left( \frac{f_i}{x_i} \right)} = \frac{80}{1.10000} = 72.7283$$

Perbandingan Ketiga Rata-rata (Mean):

$$\bar{x} = 76.10; ; U = 74.58; ; H = 72.73$$

$$H \leq U \leq \bar{x} = 76.10$$

Karakteristik penting untuk ukuran tendensi sentral yang baik

Ukuran nilai pusat/tendensi sentral (average) merupakan nilai perwakilan dari suatu distribusi data, sehingga harus memiliki sifat-sifat berikut:

- Harus mempertimbangkan semua gugus data
- Tidak boleh terpengaruh oleh nilai-nilai ekstrim.
- Harus stabil dari sampel ke sampel.



- Harus mampu digunakan untuk analisis statistik lebih lanjut.

Dari beberapa ukuran nilai pusat, Mean hampir memenuhi semua persyaratan tersebut, kecuali syarat pada point kedua, rata-rata dipengaruhi oleh nilai ekstrem. Sebagai contoh, jika item adalah 2; 4; 5; 6; 6; 6; 7; 7; 8; 9 maka mean, median dan modus semua bernilai sama, yaitu 6. Jika nilai terakhir adalah 90 bukan 9, rata-rata akan menjadi 14.10, sedangkan median dan modus tidak berubah. Meskipun dalam hal ini median dan modus lebih baik, namun tidak memenuhi persyaratan lainnya. Oleh karena itu Mean merupakan ukuran nilai pusat yang terbaik dan sering digunakan dalam analisis statistik.

Kapan kita menggunakan nilai tendensi sentral yang berbeda?

Nilai ukuran pusat yang tepat untuk digunakan tergantung pada sifat data, sifat distribusi frekuensi dan tujuan. Jika data bersifat kualitatif, hanya modus yang dapat digunakan. Sebagai contoh, apabila kita tertarik untuk mengetahui jenis tanah yang khas di suatu lokasi, atau pola tanam di suatu daerah, kita hanya dapat menggunakan modus. Di sisi lain, jika data bersifat kuantitatif, kita dapat menggunakan salah satu dari ukuran nilai pusat tersebut, mean atau median atau modus. Meskipun pada jenis data kuantitatif kita dapat

menggunakan ketiga ukuran tendensi sentral, namun kita harus mempertimbangkan sifat distribusi frekuensi dari gugus data tersebut.

Bila distribusi frekuensi data tidak normal (tidak simetris), median atau modus merupakan ukuran pusat yang tepat.

- Apabila terdapat nilai-nilai ekstrim, baik kecil atau besar, lebih tepat menggunakan median atau modus.
- Apabila distribusi data normal (simetris), semua ukuran nilai pusat, baik mean, median, atau modus dapat digunakan. Namun, mean lebih sering digunakan dibanding yang lainnya karena lebih memenuhi persyaratan untuk ukuran pusat yang baik.
- Ketika kita berhadapan dengan laju, kecepatan dan harga lebih tepat menggunakan rata-rata harmonik.
- Jika kita tertarik pada perubahan relatif, seperti dalam kasus pertumbuhan bakteri, pembelahan sel dan sebagainya, rata-rata geometrik adalah rata-rata yang paling tepat.

UWKSPRESS

## BAB V

### UKURAN PENYEBARAN (MEASURES OF DISPERSION)

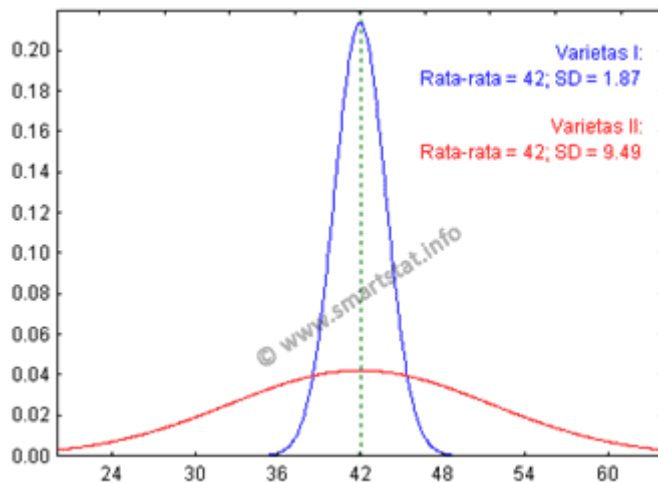
Ukuran tendensi sentral (*mean, median, mode*) merupakan nilai perwakilan dari suatu distribusi frekuensi, tetapi ukuran tersebut tidak memberikan gambaran informasi yang lengkap mengenai bagaimana penyebaran data pengamatan terhadap nilai sentralnya.

Sebagai contoh, kita mempunyai distribusi hasil panen dua varietas padi (kg per plot), masing-masing terdiri dari 5 plot. Andaikan distribusi datanya sebagai berikut:

<b>Varietas</b>	<b>I</b>	<b>:</b>	<b>45</b>	<b>42</b>	<b>42</b>	<b>41</b>
			<b>40</b>			
<b>Varietas</b>	<b>II</b>	<b>:</b>	<b>54</b>	<b>48</b>	<b>42</b>	<b>36</b>
			<b>30</b>			
<b>Varietas</b>	<b>III</b>	<b>:</b>	<b>45</b>	<b>40</b>	<b>44</b>	<b>41</b>
			<b>40</b>			

Kita dapat melihat bahwa nilai mean varietas I dan II bernilai sama, 42 kg, namun apabila kita perhatikan, keragaman kedua varietas tersebut berbeda. Varietas I mungkin lebih dipilih karena lebih konsisten. Hal ini terlihat dari data hasil pada varietas I lebih seragam dibandingkan dengan Varietas II. Pada Varietas I, hasilnya tidak terlalu jauh dari nilai pusatnya, 42 kg, sedangkan pada Varietas II, sebaran datanya sangat beragam

(perhatikan gambar berikut).



Pada contoh tersebut, jelas bahwa ukuran tendensi sentral saja tidak cukup untuk menggambarkan distribusi frekuensi. Selain itu kita harus memiliki ukuran persebaran data pengamatan. Ukuran penyebaran atau ukuran keragaman pengamatan dari nilai rata-ratanya disebut simpangan (*deviation/dispersi*). Terdapat beberapa ukuran untuk menentukan dispersi data pengamatan, seperti jangkauan/rentang (*range*), simpangan kuartil (*quartile deviation*), simpangan rata-rata (*mean deviation*), dan simpangan baku (*standard deviation*).

Materi ukuran penyebaran (measures of dispersion) ini meliputi:

- 1) Range (Rentang)
- 2) Simpangan Kuartil (Quartile Deviation)
- 3) Simpangan Rata-rata (Mean Deviation)

- 4) Ragam (Varians) dan Standar deviasi
- 5) Koefisien Keragaman (Coefficient of Variation)
- 6) Skewness dan Kurtosis

### 5.1 Jangkauan (Range)

Ukuran penyebaran yang paling sederhana adalah **range** (*Jangkauan/Rentang*, terkadang di beberapa literatur diterjemahkan dengan istilah **wilayah**). Range dari suatu kelompok data pengamatan adalah selisih antara nilai minimum dan maksimum.

$$\text{Range} = \text{nilai maksimum} - \text{nilai minimum}$$

Misalnya, range untuk Varietas I pada tabel di atas adalah  $45 - 40 = 5$  (45 adalah nilai maksimum dan 40 adalah nilai minimum). Seringkali kita mengatakan range dengan pernyataan seperti “hasil berkisar antara 40 – 45 kg per petak”. Kisarannya lebih sempit dibandingkan dengan pernyataan “hasil berkisar antara 40 – 60 kg per petak”. Pernyataan pertama menggambarkan bahwa variasi hasil padi tidak terlalu beragam, sedangkan pada pernyataan kedua, terjadi hal sebaliknya.

**Range** hanya memperhitungkan dua nilai, yaitu nilai maksimum dan nilai minimum dan tidak memperhitungkan semua nilai, sehingga sangat **tidak stabil** atau **tidak dapat diandalkan** sebagai indikator dari ukuran penyebaran. Hal ini terjadi karena range sangat dipengaruhi oleh nilai-nilai **ekstrim**. Pada contoh di atas, jika hasil tertinggi varietas I adalah 60 kg/petak, bukan 45 kg/petak, maka

range-nya =  $60-40= 20$  kg/petak.

Jelas, interpretasi kita akan berubah. Kita lebih sepakat mengatakan bahwa variasi hasil sangat beragam. Benarkah demikian? Apabila kita perhatikan kembali, nilai hasil padi lainnya hampir seragam, berkisar antara 40-44 kg/petak. Namun dengan adanya pencilan hasil, 60 kg/petak, interpretasinya jadi lain, kita cenderung mengatakan bahwa hasil beragam, padahal keragaman tersebut sebenarnya tidak mewakili semua nilai dalam sampel/populasinya.

Hasil sebesar 60 kg/petak merupakan contoh dari nilai ekstrem dan tidak biasa. Nilai tersebut merupakan pencilan (*outlier*) dan sebaiknya di periksa kembali kebenaran datanya atau dihilangkan dari data pengamatan, karena akan menghasilkan kesimpulan yang tidak tepat.

**Contoh 2:**

Contoh kasus lain yang bisa menimbulkan salah interpretasi mengenai ukuran penyebaran data dengan menggunakan Range adalah sebagai berikut:

Berikut ini adalah nilai Quiz ke-1 dan ke-2 Matakuliah Statistik.

Tentukan Range untuk masing-masing Quiz. Apa kesimpulan Anda?

Quiz ke-1:	1	20	20	20	20	20	20	20	20	20	20
Quiz ke-2:	2	3	4	5	6	14	15	16	17	18	19

**Jawab:**

Quiz 1: range = 20-1 = 19

Quiz 2: range = 19-2 = 17

**Kesimpulan:**

Quiz ke-1 lebih bervariasi di banding Quiz 2 karena nilai range Quiz 1 > Quiz 2. Bandingkan dengan kesimpulan yang diperoleh dengan menggunakan simpangan kuartil dan Standar deviasi.

Kelemahan lain dari Range adalah tidak menggambarkan sebaran data terhadap nilai pusatnya. Perhatikan contoh dan gambar berikut.

**Contoh 3:**

Tentukan Mean dan Range dari kedua Varietas berikut. Kesimpulan apa yang bisa Anda tarik berdasarkan nilai mean (rata-rata) dan range-nya?

Varietas I	45	42	42	41	40
Varietas III	45	40	44	41	40

**Jawab:**

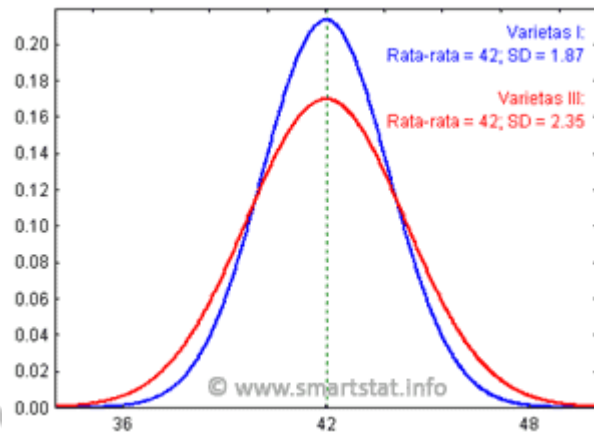
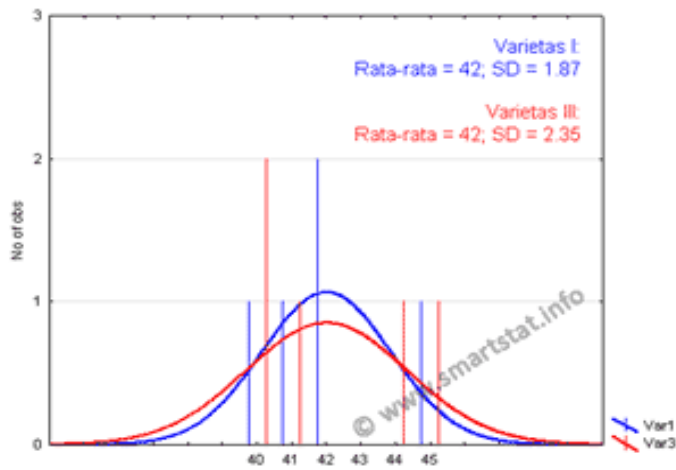
Varietas I: Mean = 42; range = 5

Varietas II: Mean = 42; range = 5

**Kesimpulan:**

Kedua Varietas, I dan III mempunyai nilai mean dan range yang sama, yaitu mean = 42 dan range = 5.





Apabila kita hanya menggunakan ukuran range sebagai ukuran penyebaran, pasti kita mengatakan bahwa keragaman hasil kedua varietas sama. Namun apabila kita perhatikan bagaimana sebaran data kedua varietas terhadap nilai pusatnya, mungkin kita lebih memilih Varietas I, karena pada Varietas I sebaran datanya tidak jauh dari nilai pusatnya.

Untuk menghindari kelemahan range seperti di atas, ukuran

dispersi lain seperti **simpangan kuartil** lebih disukai.

## 5.2 Simpangan kuartil (*Quartile Deviation*)

Simpangan kuartil dihitung dengan cara menghapus nilai-nilai yang terletak di bawah kuartil pertama dan nilai-nilai di atas kuartil ketiga, sehingga nilai-nilai ekstrem, baik yang berada di bawah ataupun di atas distribusi data, dihilangkan.

Simpangan kuartil didapatkan dengan cara menghitung nilai rata-rata dari kedua kuartil tersebut,  $Q_1$  dan  $Q_3$ .

$$\frac{(Q_3 - Q_2) + (Q_2 - Q_1)}{2} = \frac{(Q_3 - Q_1)}{2}$$

Simpangan kuartil lebih stabil dibandingkan dengan Range karena tidak dipengaruhi oleh nilai ekstrem. Nilai-nilai ekstrim sudah dihapus. Meskipun demikian, sama seperti Range, simpangan kuartil juga tetap tidak memperhatikan dan memperhitungkan penyimpangan semua gugus datanya. Simpangan kuartil hanya memperhitungkan nilai pada kuartil pertama dan kuartil ketiga saja.

### Contoh 4

Tentukan nilai simpangan kuartil pada Contoh 2.

#### Jawab:

Untuk menentukan nilai kuartil, terlebih dahulu sampel data harus diurutkan. Kebetulan pada contoh ini, data sudah terurut.

Selanjutnya tentukan **letak** dari kuartil tersebut dan terakhir **tentukan nilai** kuartilnya.

	1	2	3	4	5	6	7	8	9	10	11
Quiz 1:	1	20	20	20	20	20	20	20	20	20	20
Quiz 2:	2	3	4	5	6	14	15	16	17	18	19

= 11

$$\text{Letak } Q_i = \frac{i}{4}(n + 1)$$

### Quiz 1:

Letak  $Q_1 = \frac{1}{4}(11+1) = 3$  sehingga nilai  $Q_1$  adalah data yang terletak pada urutan ke-3, yaitu 20

Letak  $Q_3 = \frac{3}{4}(11+1) = 9$  sehingga nilai  $Q_3$  adalah data yang terletak pada urutan ke-9, yaitu 20

$$\text{simpangan kuartil} = \frac{(Q_3 - Q_1)}{2} = \frac{(20 - 20)}{2} = 0$$

### Quiz 2:

Letak  $Q_1 = \frac{1}{4}(11+1) = 3$  sehingga nilai  $Q_1$  adalah data yang terletak pada urutan ke-3, yaitu 5

Letak  $Q_3 = \frac{3}{4}(11+1) = 9$  sehingga nilai  $Q_3$  adalah data yang terletak pada urutan ke-9, yaitu 17

$$\text{simpangan kuartil} = \frac{(Q_3 - Q_1)}{2} = \frac{(17 - 5)}{2} = 6$$

### Kesimpulan:

Berdasarkan simpangan kuartil, Quiz ke-2 lebih bervariasi dibandingkan dengan Quiz ke-1. (kesimpulannya berbeda dengan kesimpulan berdasarkan range)

### 5.3 Simpangan Rata-rata (*Mean Deviation*)

Simpangan rata-rata merupakan penyimpangan nilai-nilai individu dari nilai rata-ratanya. Rata-rata bisa berupa mean atau median. Untuk data mentah simpangan rata-rata dari median cukup kecil sehingga simpangan ini dianggap paling sesuai untuk data mentah. Namun pada umumnya, simpangan rata-rata yang dihitung dari mean yang sering digunakan untuk nilai simpangan rata-rata. Simpangan rata-rata dihitung dengan formula berikut:

$$\text{Simpangan rata-rata} = \frac{\sum(x_i - \bar{x})}{n}$$

Formula tersebut tentu memenuhi dua kriteria sebelumnya, dihitung dari semua data dan menunjukkan dispersi rata-rata dari mean, tetapi tidak memenuhi kriteria ketiga. Bagaimanapun dispersi dari data, semua perhitungan dengan rumus ini akan selalu menghasilkan nilai nol. Hal ini karena pembilang dari rumus di atas  $\sum(x_i - \bar{x})$  menunjukkan bahwa hasil penjumlahannya akan selalu sama dengan nol.

Terdapat dua cara untuk mengantisipasi masalah ini, keduanya akan menghilangkan tanda-tanda negatif dari perhitungan.

**Cara pertama** adalah dengan menggunakan formula berikut:

**Sampel:**

$$\text{Simpangan rata - rata} = \frac{\sum |x_i - \bar{x}|}{n}$$

**Populasi:**

$$\text{Simpangan rata - rata} = \frac{\sum |x_i - \mu|}{N}$$

Untuk data yang sudah disusun dalam bentuk tabel frekuensi:

- **Data Tunggal** (tidak di grupkan berdasarkan selang kelas):
- **Data kelompok** (sudah digrupkan berdasarkan selang tertentu):

Simpangan rata-rata yang dihitung dari distribusi frekuensi data yang dikelompokkan menggunakan nilai data perkiraan, bukan data aslinya. Data perwakilan tersebut disimbolkan dengan  $m$ . Untuk membuat perhitungan dari data yang sudah dikelompokkan kita harus menganggap, bahwa semua nilai dalam sebuah kelas, sama dengan nilai wakilnya (tanda kelasnya,  $m_i$ ). Selanjutnya, nilai perkiraan simpangan rata-rata dapat dihitung dengan menggunakan rumus:

Pada formula di atas, pembilangnya akan selalu bernilai positif, karena yang diambil adalah nilai mutlaknya, perhatikan tanda modulus  $||$  yang berarti baik hasilnya negatif ataupun positif akan selalu diperlakukan sebagai

data positif.

**Cara kedua** adalah dengan menggunakan jumlah kuadrat dari semua nilai simpangan datanya. Cara ini dikenal dengan istilah Ragam (varians) dan **standar deviasi**.

**Contoh 5**

Tentukan nilai simpangan rata-rata pada Contoh 2.

**Jawab:**

Quiz I: rata-rata = 18.27

Quiz 2: rata-rata = 10.82

No	Quiz 1 ( $x_i$ )	$x_i - \bar{x}$	$ x_i - \bar{x} $		Quiz 2 ( $x_i$ )	$x_i - \bar{x}$	$ x_i - \bar{x} $
1	1	-17.27	17.27		2	-8.82	8.82
2	20	1.73	1.73		3	-7.82	7.82
3	20	1.73	1.73		4	-6.82	6.82
4	20	1.73	1.73		5	-5.82	5.82
5	20	1.73	1.73		6	-4.82	4.82
6	20	1.73	1.73		14	3.18	3.18
7	20	1.73	1.73		15	4.18	4.18
8	20	1.73	1.73		16	5.18	5.18
9	20	1.73	1.73		17	6.18	6.18
10	20	1.73	1.73		18	7.18	7.18
11	20	1.73	1.73		19	8.18	8.18

		<b>Jumlah</b>	<b>34.55</b>			<b>Jumlah</b>	<b>68.18</b>
--	--	---------------	--------------	--	--	---------------	--------------

**Quiz 1:**

$$\text{Simpangan rata - rata} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{34.55}{11} = 3.141$$

**Quiz 2:**

$$\text{Simpangan rata - rata} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{68.18}{11} = 6.198$$

**Kesimpulan:**

Berdasarkan simpangan rata-rata, Quiz ke-2 lebih bervariasi dibandingkan dengan Quiz ke-1. (kesimpulannya berbeda dengan kesimpulan berdasarkan range)

**Catatan:**

Untuk menentukan simpangan rata-rata dari tabel frekuensi, caranya mirip dengan **contoh 7** dan **8**.

**Contoh Tambahan:**

Dengan cara yang sama seperti di atas, nilai simpangan rata-rata untuk ketiga varietas:

$$\text{Varietas I} = 1.2$$

$$\text{Varietas II} = 7.2$$

$$\text{Varietas III} = 2$$

#### **5.4 Ragam dan Standar deviasi**

Ukuran penyebaran dengan menggunakan perhitungan simpangan rata-rata diperoleh dengan mengabaikan tanda-tanda penyimpangan.

Secara matematis hal tersebut tidak benar. Cara kedua,

yaitu dengan mengkuadratkan nilai simpangan sehingga nilai negatif berubah menjadi positif. Cara ini lebih tepat. Rata-rata dari jumlah nilai simpangan dikenal dengan **ragam**(*varians*). Setelah nilai ragam diperoleh, selanjutnya nilai ragam tersebut diakarkan untuk mendapatkan kembali satuan asal dari variabel tersebut (*bukan kg<sup>2</sup>/petak<sup>2</sup>, tapi kg/petak :-*) . Cara pengukuran keragaman seperti ini dikenal dengan **Standar deviasi**.

Secara matematis, standar deviasi dapat dihitung dengan menggunakan formula:

**Standar deviasi populasi** disimbolkan dengan  $\sigma$  (baca ‘sigma’) dan **standar deviasi sampel** disimbolkan dengan  $s$ . Standar deviasi sampel yang baik seharusnya merupakan ukuran yang **tidak bias** terhadap standar deviasi populasinya, karena kita menggunakan ukuran standar deviasi sampel untuk memperkirakan nilai standar deviasi populasi. Untuk itu, nilai  $n$  pada formula di atas diganti dengan  $n - 1$  sehingga formula untuk standar deviasi sampel adalah sebagai berikut:

**Data pada tabel distribusi frekuensi:**

**Data Tunggal:**

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

**Data kelompok** (sudah digrupkan berdasarkan selang tertentu):



Sama seperti pada perhitungan simpangan rata-rata. Standar deviasi dan ragam yang dihitung dari distribusi frekuensi data yang sudah dikelompokkan menggunakan nilai data perkiraan, bukan data aslinya. **Data perwakilan tersebut disimbolkan dengan  $m$ .** Untuk membuat perhitungan dari data yang sudah dikelompokkan kita harus menganggap, bahwa semua nilai dalam sebuah kelas, sama dengan nilai wakilnya (tanda kelasnya,  $m_i$ ). Selanjutnya, nilai perkiraan standar deviasi dapat dihitung dengan menggunakan rumus:

$$\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}$$

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

Nilai kuadrat dari standar deviasi dikenal dengan ragam (*variance*).

Pada teknik **analisis varian**, dikenal dengan **Jumlah Kuadrat** (*Sum of Square*), dan **ragam** (*varian*) dikenal dengan istilah **Kuadrat Tengah/Rata-rata Jumlah Kuadrat** (*Mean Square*).

Standar deviasi merupakan ukuran penyebaran yang paling banyak digunakan. Semua gugus data dipertimbangkan sehingga lebih stabil dibandingkan dengan ukuran lainnya. Namun, apabila dalam gugus data tersebut terdapat nilai ekstrem, standar deviasi menjadi tidak sensitif lagi, sama

halnya seperti mean.

Standar Deviasi memiliki beberapa karakteristik khusus lainnya. SD tidak berubah apabila setiap unsur pada gugus datanya di tambahkan atau dikurangkan dengan nilai konstan tertentu. SD berubah apabila setiap unsur pada gugus datanya dikali/dibagi dengan nilai konstan tertentu. Bila dikalikan dengan nilai konstan, standar deviasi yang dihasilkan akan setara dengan hasil kali dari nilai standar deviasi aktual dengan konstan.

**Contoh 6 :**

Apabila data nilai Quiz pada contoh 2 diambil dari sampel, tentukan nilai ragam dan standar deviasinya.

**Jawab:**

Untuk mencari nilai standar deviasi sampel, kita bisa menggunakan salah satu formula berikut:

Formula pertama adalah formula secara definitif. Formula yang direkomendasikan untuk perhitungan secara manual adalah formula yang ke-2. Cara perhitungan dengan formula yang ke-2 bisa di lihat pada contoh 7 dan 8. Pada contoh ini, sebagai latihan, kita gunakan formula yang pertama. Untuk perhitungan dengan formula pertama, kita memerlukan nilai rata-ratanya, sehingga terlebih dahulu kita harus menghitung nilai rata-ratanya.

Quiz I: rata-rata = 18.27, Quiz 2: rata-rata = 10.82

No	Quiz 1( $x_i$ )	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	Quiz 2( $x_i$ )	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
1	1	-17.27	298.35	2	-8.82	77.76
2	20	1.73	2.98	3	-7.82	61.12
3	20	1.73	2.98	4	-6.82	46.49
4	20	1.73	2.98	5	-5.82	33.85
5	20	1.73	2.98	6	-4.82	23.21
6	20	1.73	2.98	14	3.18	10.12
7	20	1.73	2.98	15	4.18	17.49
8	20	1.73	2.98	16	5.18	26.85
9	20	1.73	2.98	17	6.18	38.21
10	20	1.73	2.98	18	7.18	51.58
11	20	1.73	2.98	19	8.18	66.94
		<b>Jumlah</b>	<b>3.281.818</b>			<b>4.536.364</b>

**Quiz 1:**

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} \quad s = \sqrt{\frac{328.18}{11 - 1}} = 5.73$$

$$ragam = s^2 = 5.73^2 = 32.82$$

**Quiz 2:**

$$s = \sqrt{\frac{453.64}{11 - 1}} = 6.74$$

$$ragam = s^2 = 6.74^2 = 45.36$$

**Kesimpulan:**

Berdasarkan nilai ragam dan standar deviasi, Quiz ke-2 lebih bervariasi dibandingkan dengan Quiz ke-1. (kesimpulannya berbeda dengan kesimpulan berdasarkan range)

### Contoh 7

Hitung nilai standar deviasi dan ragam dari tabel frekuensi data tunggal berikut:

No	$x_i$	$f_i$
1	70	5
2	69	6
3	45	3
4	80	1
5	56	1
<b>Jumlah</b>	<b>320</b>	<b>16</b>

### Jawab:

Untuk kemudahan dalam perhitungan secara manual, kita gunakan formula standar deviasi berikut:

Selanjutnya kita buat tabel seperti pada tabel berikut:

No	$x_i$	$f_i$	$f_i \cdot x_i$	$f_i \cdot x_i^2$
1	70	5	350	24500
2	69	6	414	28566
3	45	3	135	6075
4	80	1	80	6400
5	56	1	56	3136
<b>Jumlah</b>	<b>320</b>	<b>16</b>	<b>1035</b>	<b>68677</b>

Dari tabel tersebut didapat:

$$n = 16$$

$$\text{mean} = 1035/16 = 64.69$$

**Standar deviasi:**

$$s = \sqrt{\frac{68677 - \frac{(1035)^2}{16}}{16 - 1}} = 10.72$$

$$\text{atau } s = \sqrt{\frac{16(68677) - (1035)^2}{16(16 - 1)}} = 10.72$$

$$\text{Ragam } s^2 = (10.72)^2 = 115.03$$

### Contoh 8

Hitung nilai standar deviasi dan ragam dari tabel frekuensi yang sudah dikelompokkan:

Tabel berikut ini adalah nilai ujian statistik 80 mahasiswa yang sudah disusun dalam tabel frekuensi. Berbeda dengan contoh di atas, pada contoh ini, tabel distribusi frekuensi dibuat dari data yang sudah dikelompokkan berdasarkan selang/kelas tertentu (banyak kelas = 7 dan panjang kelas = 10).

Kelas ke-	Nilai Ujian	$f_i$
1	31 – 40	2
2	41 – 50	3
3	51 – 60	5
4	61 – 70	13
5	71 – 80	24
6	81 – 90	21
7	91 – 100	12
	<b>Jumlah</b>	<b>80</b>

**Jawab:**

Untuk kemudahan dalam perhitungan secara manual, kita gunakan formula standar deviasi berikut:

Selanjutnya kita buat daftar tabel berikut, tentukan nilai tengah kelas/pewakilnya ( $m_i$ ) dan lengkapi kolom berikutnya.

Kelas ke-	Nilai Ujian	$f_i$	$m_i$	$f_i \cdot m_i$	$f_i \cdot m_i^2$
1	31 – 40	2	35.5	71.0	2520.5

Kelas ke-	Nilai Ujian	$f_i$	$m_i$	$f_i \cdot m_i$	$f_i \cdot m_i^2$
2	41 – 50	3	45.5	136.5	6210.8
3	51 – 60	5	55.5	277.5	15401.3
4	61 – 70	13	65.5	851.5	55773.3
5	71 – 80	24	75.5	1812.0	136806.0
6	81 – 90	21	85.5	1795.5	153515.3
7	91 – 100	12	95.5	1146.0	109443.0
	<b>Jumlah</b>	<b>80</b>	<b>458.5</b>	<b>6090.0</b>	<b>479670.0</b>

Dari tabel tersebut didapat:

$$n = 80$$

$$\text{mean} = 6090/80 = 76.13$$

Standar deviasi dan ragam:

$$\text{Ragam } s^2 = (77.92)^2 = 6070.81$$

Dengan cara yang sama seperti di atas, nilai standar deviasi untuk ketiga varietas:

$$\text{Varietas I} = 1.87$$

$$\text{Varietas II} = 9.49$$

$$\text{Varietas III} = 2.35$$

### 5.5 Ukuran sebaran relatif (*Measures of Relative Dispersion*)

Perhatikan contoh kasus Varietas I vs Varietas II di atas.

Kedua varietas tersebut mempunyai nilai rata-rata yang sama. Untuk dua distribusi data dengan nilai rata-rata yang

sama atau hampir sama, kita dapat secara langsung membandingkan keragaman kedua distribusi tersebut dengan melihat nilai standar deviasinya. Kita sepakat untuk mengatakan bahwa Varietas II lebih beragam dibandingkan dengan Varietas I. Namun apabila rata-rata dari kedua distribusi data tersebut jauh berbeda, kita tidak dapat membandingkan keragamannya dengan menggunakan nilai standar deviasinya secara langsung. Pada kasus tersebut, untuk membandingkan tingkat keragaman dari kedua distribusi datanya, kita harus menggunakan ukuran penyebaran relatif.

Terdapat beberapa ukuran penyebaran relatif untuk Range, Simpangan Kuartil, Simpangan Rata-rata, dan Standar deviasi. **Koefisien Keragaman** (*coefficient of variation*) yang paling penting dan sering digunakan adalah **ukuran penyebaran relatif dari Standar Deviasi**.

Koefisien Keragaman Standar deviasi dihitung dengan formula berikut:

$$KK = C.V = \frac{s}{\bar{x}} \times 100\%$$

Koefisien Keragaman merupakan ukuran yang bebas satuan dan selalu dinyatakan dalam bentuk persentase. Nilai KK yang kecil menunjukkan bahwa data tidak terlalu beragam dan di katakan lebih konsisten. KK tidak dapat diandalkan apabila nilai rata-rata hampir sama dengan 0 (nol). KK juga tidak stabil apabila skala pengukuran data



yang digunakan bukan skala rasio.

**Contoh 9:**

Perhatikan gugus data untuk Kelompok A dan  
Kelompok B

A	2	4	5	6	6	7	7	7	8	9
B	3	6	7	9	9	10	10	10	11	12

Kelompok A: Rata-rata = 6.1;  $s = 2.0$

Kelompok B: Rata-rata = 8.7;  $s = 2.7$

$$KK(A) = \frac{s}{\bar{x}} \times 100\% = \frac{2.0}{6.1} \times 100\% = 33.2\%$$

$$KK(B) = \frac{s}{\bar{x}} \times 100\% = \frac{2.7}{8.7} \times 100\% = 30.7\%$$

Nilai Koefisien keragaman kelompok B lebih kecil dibandingkan dengan kelompok A. Namun apabila kita perhatikan nilai standar deviasinya, justru hal sebaliknya yang terjadi, SD A lebih kecil dibandingkan dengan SD B. Dengan demikian, **untuk melihat keragaman relatif suatu gugus data, kita jangan hanya menyandarkan pada nilai standar deviasinya.**

### 5.6 Skewness and Kurtosis

Rata-rata dan ukuran penyebaran dapat menggambarkan distribusi data tetapi tidak cukup untuk menggambarkan **sifat distribusi**. Untuk dapat

menggambarkan karakteristik dari suatu distribusi data, kita menggunakan konsep-konsep lain yang dikenal sebagai kemiringan (*skewness*) dan keruncingan (*kurtosis*).

### **Ukuran Kemencengan Kurva (*Skewness*)**

Kurva yang tidak simetris dapat menceng ke kiri atau ke kanan. Di dalam kurva yang simetris, letak rata-rata, median, dan modus sama. Perhatikan tiga bentuk kurva berikut:

Untuk menghitung tingkat kemencengan suatu kurva digunakan rumus berikut:

Suatu kurva disebut menceng kanan jika tingkat kemencengannya positif. Kurva disebut simetris jika tingkat kemencengannya nol. Sedangkan, jika tingkat kemencengannya negatif maka kurva disebut menceng kiri.

### **Ukuran Keruncingan Kurva (*Kurtosis*)**

Dilihat dari tingkat keruncingannya, bentuk kurva dibagi menjadi tiga, yaitu **leptokurtis**, **platykurtis**, dan **mesokurtis**. Bentuk kurva dari masing-masing jenis tersebut adalah sebagai berikut:

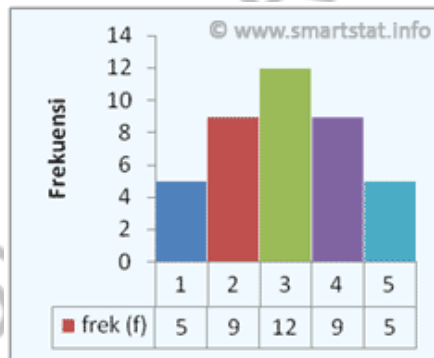
Untuk menghitung tingkat keruncingan suatu kurva digunakan moment koefisien keruncingan (*moment coefficient of kurtosis*) yang rumusnya sebagai berikut:

Suatu kurva disebut meruncing (*leptokurtis*) jika tingkat keruncingannya lebih dari 3. Kurva disebut normal (*mesokurtis*) jika tingkat keruncingannya sama dengan 3. Sedangkan, jika tingkat keruncingannya kurang dari 3 maka kurva disebut mendatar (*platykurtis*).

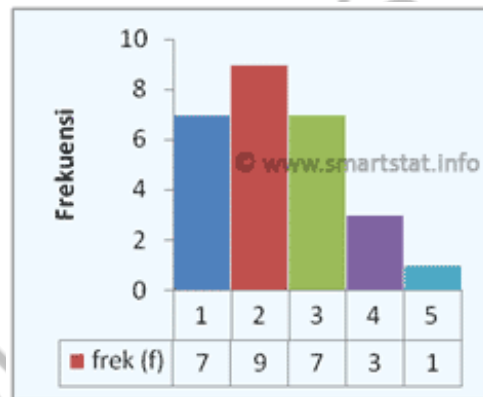
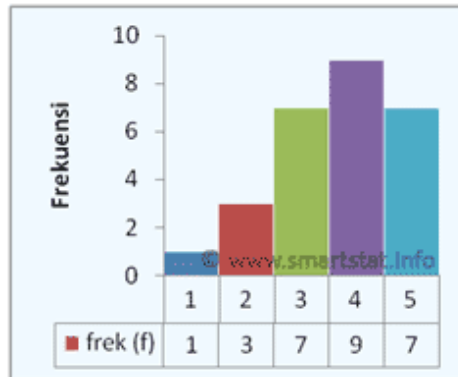
### a. Skewness

Kemiringan (*skewness*) berarti ketidaksimetrisan. Sebuah distribusi dikatakan simetris apabila nilai-nilainya tersebar merata disekitar nilai rata-ratanya. Sebagai contoh, distribusi data berikut simetris terhadap nilai rata-ratanya.

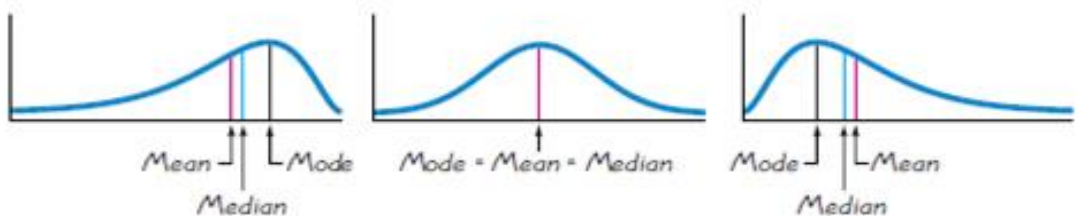
x	1	2	3	4	5
frek (f)	5	9	12	9	5



Pada contoh gambar berikut, distribusi data tidak simetris. Gambar pertama miring (menjulang) ke arah kiri dan gambar ke-2 miring ke arah kanan.



Pada distribusi data yang simetris, mean, median dan modus bernilai sama.



**(a) Skewed to the Left (Negatively Skewed):** The mean and median are to the *left* of the mode.

**(b) Symmetric (Zero Skewness):** The mean, median, and mode are the same.

**(c) Skewed to the Right (Positively Skewed):** The mean and median are to the *right* of the mode.

Beberapa langkah-langkah perhitungan digunakan untuk menyatakan arah dan tingkat kemiringan dari sebaran data. Langkah-langkah tersebut diperkenalkan oleh Pearson.

Koefisien kemiringan(*Coefficient of Skewness*):

$$S_k = \frac{3(\text{mean} - \text{median})}{\text{standar deviasi}}$$

**Interpretasi:** Untuk distribusi data yang simetris,  $S_k = 0$ . Apabila distribusi data **menjuler ke kiri** (*negatively skewed*),  $S_k$  bernilai **negatif**, dan apabila **menjuler ke kanan** (*positively skewed*),  $S_k$  bernilai **positif**. Kisaran untuk  $S_k$  antara -3 dan 3.

Ukuran kemiringan yang lain adalah koefisien

$\beta_1$  (baca 'beta-satu'):

$$\text{populasi : } \beta_1 = \frac{\mu_3^2}{\mu_2^3}; \quad \text{sampel : } b_1 = \frac{m_3^2}{m_2^3}$$

dimana:

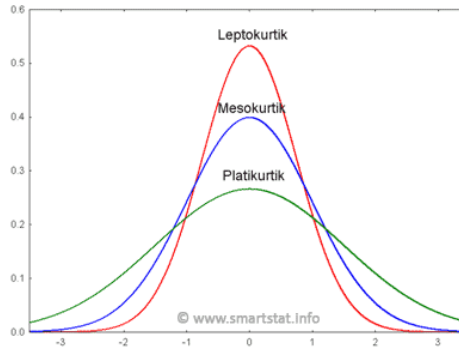
$$m_3 = \frac{\sum(x_i - \bar{x})^3}{n - 1}; \quad \text{dan } m_2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

**Interpretasi:**

Distribusi dikatakan simetris apabila nilai  $b_1 = 0$ . Skewness positif atau negatif tergantung pada nilai  $b_1$  apakah bernilai positif atau negatif.

## b. Kurtosis

Kurtosis merupakan ukuran untuk mengukur keruncingan distribusi data.



Distribusi pada gambar di atas semuanya simetris terhadap nilai rata-ratanya. Namun bentuk ketiganya tidak sama. Kurva berwarna biru dikenal sebagai **mesokurtik** (kurva normal), kurva berwarna merah dikenal sebagai **leptokurtik** (kurva runcing) dan kurva berwarna hijau dikenal sebagai **platikurtik** (kurva datar).

Kurtosis dihitung dengan menggunakan koefisien Pearson,  $\beta_2$  (baca 'beta – dua').

$$\text{populasi : } \beta_2 = \frac{\mu_4}{\mu_2^2}; \quad \text{sampel : } b_2 = \frac{m_4}{m_2^2}$$

dimana:

$$m_4 = \frac{\sum(x_i - x)^4}{n - 1}; \quad \text{dan } m_2 = \frac{\sum(x_i - x)^2}{n - 1}$$

### Interpretasi:

Distribusi dikatakan:

Mesokurtik (Normal) jika  $b_2 = 3$

Leptokurtik jika  $b_2 > 3$

Platikurtik jika  $b_2 < 3$

## BAB VI

### ANGKA INDEKS

#### 6.1 Angka Perbandingan, Angka Relatif, Angka Indeks

*Angka Perbandingan* merupakan suatu bilangan yang diperoleh dengan membandingkan bilangan dari keadaan yang satu dengan bilangan dari keadaan yang lain.

*Angka Relatif* adalah angka perbandingan yang dinyatakan dalam bentuk persen.

*Angka Indeks* adalah angka relatif tetapi simbol persen (%) nya tidak dicantumkan.

#### Contoh 1

Diketahui jumlah produksi pada perusahaan B selama kurun waktu 3 tahun sebagai berikut :

Tahun	Jumlah Produksi (Unit)
2000	100

2001	75
2002	250

a. Bandingkan jumlah produksi tahun 2001 dengan tahun 2000 !

- **Angka Perbandingan** = 
$$\frac{\text{jumlah produksi 2001}}{\text{jumlah produksi 2000}} = \frac{75}{100}$$
- **Angka Relatif** = 
$$\frac{75}{100} \times 100\% = 75\%$$
- **Angka Indeks** = 
$$\frac{75}{100} \times 100 = 75$$

Pada tahun 2001 jumlah produksi mengalami penurunan sebesar 25% dari tahun 2000.

b. Bandingkan jumlah produksi tahun 2002 dengan tahun 2000 !

- **Angka Perbandingan** = 
$$\frac{\text{jumlah produksi 2002}}{\text{jumlah produksi 2000}} = \frac{250}{100}$$
- **Angka Relatif** = 
$$\frac{250}{100} \times 100\% = 250\%$$
- **Angka Indeks** = 
$$\frac{250}{100} \times 100 = 250$$



Jumlah produksi pada tahun 2002 mengalami kenaikan sebesar 150% dari tahun 2000.

## 6.2 Kegunaan Angka Indeks

1. Dengan angka indeks kita bisa mengetahui naik/turun atau maju/mundur suatu usaha/kegiatan.
2. Angka indeks kurang dari 100 ( $I < 100$ ) menunjukkan kemunduran/penurunan suatu usaha/kegiatan. Penurunannya sebesar  $(100 - I)$ .

### Contoh 2

Angka indeks = 75 → terjadi penurunan sebesar  $100 - 75 = 25$ .

3. Angka indeks lebih dari 100 ( $I > 100$ ) menunjukkan kemajuan/kenaikan suatu usaha/kegiatan. Kenaikannya sebesar  $(I - 100)$ .

### Contoh 3

Angka indeks = 250 → terjadi kenaikan sebesar  $250 - 100 = 150$ .

4. Angka indeks dapat memperlihatkan perbandingan antara kegiatan yang sama dalam dua waktu yang berbeda.

### Contoh 4

Membandingkan jumlah produksi tahun 2001 dengan 2000.

5. Dalam membuat angka indeks diperlukan dua macam waktu, yakni :

a. Waktu / tahun dasar.

Adalah waktu/tahun dimana suatu kegiatan digunakan sebagai dasar perbandingan.

b. Waktu/tahun yang bersangkutan.

Adalah waktu/tahun dimana kegiatannya akan dibandingkan terhadap kegiatan pada waktu dasar.

### 6.3 Angka Indeks Sederhana

#### a. Indeks Harga

Digunakan untuk mengukur perubahan harga barang.

$$I_p = \frac{P_t}{P_0} \times 100 \quad (6.1)$$

Keterangan :

$I_p$  = indeks harga tahun yang bersangkutan

$P_0$  = harga pada tahun dasar

$P_t$  = harga pada tahun yang bersangkutan

Catatan : indeks harga pada tahun dasar selalu bernilai 100 ( $I_{P_0} = 100$ )

#### Contoh 5

Hitunglah indeks harga barang X untuk tahun-tahun berikut, dengan menggunakan :

- 1). Tahun dasar 2005
- 2). Tahun dasar 2005 – 2007

1) Lengkapi tabel di bawah ini !

Tabel 1. Harga Dan Indeks Harga Barang X  
Dengan Tahun Dasar 2005

Tahun	Harga (Rp)	Indeks Harga (2005=100)
2005*	4000	
2006	4200	
2007	4800	
2008	5500	

**Catatan :**

- ✓ 2005=100 artinya tahun dasar yang digunakan adalah tahun 2005 sehingga indeks harga pada tahun tersebut adalah 100.
- ✓  $P_0 = 4000$ , maka untuk mengisi indeks harga gunakan :

$$I_p = \frac{P_t}{4000} \times 100$$

2). Lengkapi tabel di bawah ini !

Tabel 2. Harga Dan Indeks Harga Barang X  
Dengan Tahun Dasar 2005 – 2007

Tahun	Harga (Rp)	Indeks Harga (2005-2007=100)
2005	4000	

Catatan :

- ✓ 2005-2007=100 artinya tahun dasar yang digunakan adalah tahun 2005-2007.
- ✓ Karena tahun dasarnya lebih dari satu waktu (3 tahun), maka  $P_0$ -nya merupakan "rata-rata harga tahun dasar"

✓  $P_0 = \frac{4000 + 4200 + 4800}{3}$  , maka untuk mengisi indeks harga

$P_0 = 4333,33$

gunakan :

$$I_p = \frac{P_t}{4333,33} \times 100$$

- ✓ Jika indeks harga dari 3 tahun dasar tersebut dirata-ratakan maka hasilnya akan sama dengan 100.
- ✓  $\bar{X}_{I_p} =$

### b. Indeks Kuantitas/Jumlah

Digunakan untuk mengukur perubahan kuantitas/jumlah barang.

$$I_Q = \frac{Q_t}{Q_0} \times 100 \quad (6.2)$$

Dengan :  $I_Q$  = indeks kuantitas/jumlah tahun yang bersangkutan

$Q_0$  = kuantitas pada tahun dasar

$Q_t$  = kuantitas pada tahun yang bersangkutan

Catatan : indeks kuantitas pada tahun dasar selalu bernilai 100 ( $I_{Q0} = 100$ )

#### Contoh 6

Hitunglah indeks jumlah barang X untuk tahun-tahun berikut, dengan menggunakan :

- 1). Tahun dasar 2000
- 2). Tahun dasar 2000 – 2001

Tabel 3. Jumlah produksi Dan Indeks Jumlah Produksi

Barang X Dengan Tahun Dasar 2000

Tahun	Jumlah Produksi (unit)	Indeks Jumlah Produksi (2000=100)
2000	150	
2001	120	

2002	250	
2003	300	

**Catatan :**

2000=100 artinya tahun dasar yang digunakan adalah tahun 2000 sehingga indeks jumlah produksi pada tahun tersebut adalah 100.

$Q_0 = 150$ , maka untuk mengisi indeks jumlah produksi gunakan :

$$I_Q = \frac{Q_t}{150} \times 100$$

Tabel 4. Jumlah produksi Dan Indeks Jumlah Produksi Barang X Dengan Tahun Dasar 2000-2001

Tahun	Jumlah Produksi (unit)	Indeks Jumlah Produksi (2000-2001=100)
2000	150	
2001	120	
2002	250	
2003	300	

**Catatan :**

✓ 2000-2001=100 artinya tahun dasar yang digunakan adalah tahun 2000-2001.

✓ Karena tahun dasarnya lebih dari satu waktu (2 tahun), maka  $Q_0$ -nya merupakan "rata-rata jumlah produksi tahun dasar" 125

✓  $Q_0 = \frac{150 + 120}{2}$  ,

$Q_0 = 135$

maka untuk mengisi indeks jumlah produksi gunakan

## 6.4 Angka Indeks Gabungan

### 1. Angka Indeks Gabungan Tidak Diboboti

a. Metode Agregatif Sederhana

### 2. Angka Indeks Gabungan Diboboti

a. Metode Laspeyres

b. Metode Paasche

c. Metode Drobish

d. Metode Fisher

## Penyusunan Angka Indeks Gabungan

### 1. Angka Indeks Gabungan Tidak Diboboti

No.	Metode	Indeks Harga	Indeks Jumlah
1.	Metode Agregatif Sederhana	$I_{AP} = \frac{\sum P_t}{\sum P_0} \times 100$	$I_{AQ} = \frac{\sum Q_t}{\sum Q_0} \times 100$

### 2. Angka Indeks Gabungan Diboboti

N	Metode	Harga	Indeks Jumlah
o.			

1.	<b>Metode Laspeyres</b>	$L_P = \frac{\sum (P_t \times Q_o)}{\sum (P_o \times Q_o)} \times 100$	$I_Q = \frac{\sum (Q_t \times P_o)}{\sum (Q_o \times P_o)} \times 100$
2.	<b>Metode Paasche</b>	$P_P = \frac{\sum (P_t \times Q_t)}{\sum (P_o \times Q_t)} \times 100$	$I_Q = \frac{\sum (Q_t \times P_t)}{\sum (Q_o \times P_t)} \times 100$
3.	<b>Metode Drobish</b>	$D_P = \frac{L_P + P_P}{2}$	$D_Q = \frac{I_Q + P_Q}{2}$
4.	<b>Metode Fisher</b>	$F_P = \sqrt{L_P \times P_P}$	$F_Q = \sqrt{I_Q \times P_Q}$

**Keterangan :**

$P_o$  = harga pada tahun dasar

$P_t$  = harga pada tahun yang bersangkutan

$Q_o$  = jumlah pada tahun dasar

$Q_t$  = jumlah pada tahun yang bersangkutan

$n$  = banyak jenis barang

**Contoh 7**

Berikut adalah data mengenai harga (\$) dan jumlah (unit) empat jenis barang yang dijual selama tahun 2000-2001 :

Jenis barang	Harga		Jumlah	
	2000	2001	2000	2001
A	50	75	20	30



B	80	120	50	40
---	----	-----	----	----

Tentukan indeks harga gabungan dan indeks jumlah gabungan keempat jenis barang tersebut untuk tahun 2001 dengan menggunakan seluruh metode ! (tahun dasar 2000)

**Jawab :**

Jenis barang	Harga		Jumlah		P <sub>t</sub> · Q <sub>0</sub>	P <sub>0</sub> · Q <sub>0</sub>	Q <sub>t</sub> · P <sub>0</sub>	P <sub>t</sub> · Q <sub>t</sub>
	2000	2001	2000	2001				
A	50	75	20	30				
B	80	120	50	40				
TOTAL								

#### a. Metode Agregatif Sederhana

- **Harga**

$$I_{AP} = \frac{\sum P_t}{\sum P_0} \times 100 =$$

Pada tahun 2001, harga kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

- **Jumlah**

$$I_{AQ} = \frac{\sum Q_t}{\sum Q_0} \times 100 =$$

Pada tahun 2001, jumlah produksi kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

### b. Metode Laspeyres

- **Harga**

$$L_P = \frac{\sum (P_t \times Q_o)}{\sum (P_0 \times Q_o)} \times 100 =$$

Pada tahun 2001, harga kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

- **Jumlah**

$$L_Q = \frac{\sum (Q_t \times P_o)}{\sum (Q_0 \times P_o)} \times 100 =$$

Pada tahun 2001, jumlah produksi kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

### c. Metode Paasche

- **Harga**

$$P_P = \frac{\sum (P_t \times Q_t)}{\sum (P_0 \times Q_t)} \times 100 =$$

Pada tahun 2001, harga kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

- **Jumlah**

$$P_Q = \frac{\sum(Q_t \times P_t)}{\sum(Q_0 \times P_t)} \times 100 =$$

Pada tahun 2001, jumlah produksi kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

#### d. Metode Drobish

- **Harga**

$$D_P = \frac{L_P + P_P}{2} =$$

Pada tahun 2001, harga kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

- **Jumlah**

$$D_Q = \frac{L_Q + P_Q}{2} =$$

Pada tahun 2001, jumlah produksi kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

**e. Metode Fisher**

• **Harga**

$$F_P = \sqrt{L_P \times P_P} =$$

Pada tahun 2001, harga kedua jenis barang mengalami sebesarjika dibandingkan dengan tahun 2000.

• **Jumlah**

$$F_Q = \sqrt{L_Q \times P_Q} =$$

Pada tahun 2001, jumlah produksi kedua jenis barang mengalami sebesar jika dibandingkan dengan tahun 2000.

## BAB VII

### ANALISA SERIAL WAKTU (*TIME SERIES*)

#### 7.1 Pengertian

Data time series adalah nilai-nilai suatu variabel yang berurutan menurut waktu (misal: hari, minggu, bulan, tahun). Ada 4 faktor yang mempengaruhi data *time series*. Dalam data ekonomi biasanya kita mendapatkan adanya fluktuasi/ variasi dari waktu ke waktu atau disebut dengan variasi *time series*. Variasi ini biasanya disebabkan oleh adanya faktor Trend (*trend factor*), Fluktuasi siklis (*cyclical fluctuation*), Variasi musiman (*seasonal variation*), dan pengaruh random (*irregular/random influences*).

**Trend** adalah keadaan data yang menaik atau menurun dari waktu ke waktu. Contoh yang menunjukkan trend menaik yaitu pendapatan per kapita, jumlah penduduk. **Variasi musiman** adalah fluktuasi yang muncul secara reguler setiap tahun yang biasanya disebabkan oleh iklim, kebiasaan (mempunyai pola tetap dari waktu ke waktu). Contoh yang

menunjukkan variasi musiman seperti penjualan pakaian akan meningkat pada saat hari raya, penjualan buku dan tas sekolah akan meningkat pada saat awal sekolah. **Variasi siklis** muncul ketika data dipengaruhi oleh fluktuasi ekonomi jangka panjang, variasi siklis ini bisa terulang setelah jangka waktu tertentu. Variasi siklis biasanya akan kembali normal setiap 10 atau 20 tahun sekali, bisa juga tidak terulang dalam jangka waktu yang sama. ini yang membedakan antara variasi siklis dengan musiman. Gerakan siklis tiap komoditas mempunyai jarak waktu muncul dan sebab yang berbeda-beda, yang sampai saat ini belum dapat dimengerti. Contoh yang menunjukkan variasi siklis seperti industri konstruksi bangunan mempunyai gerakan siklis antara 15-20 tahun sedangkan industri mobil dan pakaian gerakan siklisnya lebih pendek lagi. **Variasi random** adalah suatu variasi atau gerakan yang tidak teratur (*irregular*). Variasi ini pada kenyataannya sulit diprediksi. Contoh variasi ini dalam data *time series* karena adanya perang, bencana alam dan sebab-sebab unik lainnya yang sulit diduga. Total variasi dalam data *time series* adalah merupakan hasil dari keempat faktor tersebut yang mempengaruhi secara bersama-sama. Dalam tulisan ini hanya akan dianalisa dua variasi pertama, sedangkan dua variasi terakhir tidak dianalisa karena memang pola variasi tersebut tidak tersistem dengan baik selain membutuhkan waktu yang sangat lama untuk mendapatkan data yang panjang. Pengalaman dan *feeling so*

*good* dari pengambil keputusan dapat membantu *adjustment* pada hasil ramalan.

Model *Time Series* adalah suatu peramalan nilai-nilai masa depan yang didasarkan pada nilai-nilai masa lampau suatu variabel dan atau kesalahan masa lampau. Model *time series* biasanya lebih sering digunakan untuk suatu peramalan/prediksi. Dalam tehnik peramalan dengan *time series* ada 2 kategori utama yang perlu dilakukan pengujian, yaitu pemulusan (*smoothing*) dan dekomposisi (*decomposition*). Metode pemulusan mendasarkan ramalannya dengan prinsip rata-rata dari kesalahan masa lalu (*Averaging smoothing past errors*) dengan menambahkan nilai ramalan sebelumnya dengan persentase kesalahan (*percentage of the errors*) antara nilai sebenarnya (*actual value*) dengan nilai ramalannya (*forecasting value*). Metoda dekomposisi mendasarkan prediksinya dengan membagi data *time series* menjadi beberapa komponen dari Trend, Siklis, Musiman dan pengaruh Random; kemudian mengkombinasikan prediksi dari komponen-komponen tersebut (kecuali pengaruh random yang sulit diprediksi). Pendekatan lain untuk peramalan adalah metoda causal atau yang lebih dikenal dengan sebutan regresi. Tehnik pemulusan dan regresi akan dibahas pada sesi tulisan yang lain.

## 7.2 Jenis- Jenis Metode Peramalan

Untuk melakukan peramalan diperlukan metode tertentu dan metode mana yang digunakan tergantung dari data dan informasi yang akan diramal serta tujuan yang hendak dicapai. Dalam prakteknya terdapat berbagai metode peramalan antara lain :

### **1. Time Series atau Deret Waktu**

Analisis time series merupakan hubungan antara variabel yang dicari (dependent) dengan variabel yang mempengaruhi-nya (independent variable), yang dikaitkan dengan waktu seperti mingguan, bulan, triwulan, catur wulan, semester atau tahun.

Dalam analisis time series yang menjadi variabel yang dicari adalah waktu.

Metode peramalan ini terdiri dari :

- a. Metode Smoting, merupakan jenis peramalan jangka pendek seperti perencanaan persediaan, perencanaan keuangan. Tujuan penggunaan metode ini adalah untuk mengurangi ketidakteraturan data masa lampau seperti musiman.
- b. Metode Box Jenkins, merupakan deret waktu dengan menggunakan model matematis dan digunakan untuk peramalan jangka pendek.
- c. Metode proyeksi trend dengan regresi, merupakan metode yang dignakan baik untuk jangka pendek



maupun jangka panjang. Metode ini merupakan garis trend untuk persamaan matematis.

## **2. Causal Methods atau sebab akibat**

Merupakan metode peramalan yang didasarkan kepada hubungan antara variabel yang diperkirakan dengan variabel lain yang mempengaruhinya tetapi buakn waktu.

Dalam prakteknya jenis metode peramalan ini terdiri dari :

- a. Metode regresi dan kolerasi, merupakan metode yang digunakan baik untuk jangka panjang maupun jangka pendek dan didasarkan kepada persamaan dengan teknik least squares yang dianalisis secara statis.
- b. Model Input Output, merupakan metode yang digunakan untuk peramalan jangka panjang yang biasa digunakan untuk menyusun trend ekonomi jangka panjang.
- c. Model ekonometri, merupakan peramalan yang digunakan untuk jangka panjang dan jangka pendek.

## **BAB VIII**

### **ANALISIS TREND**

#### **8.1 Pengertian**

**Analisis trends** merupakan suatu metode analisis statistika yang ditujukan untuk melakukan suatu estimasi atau peramalan pada masa yang akan datang. Untuk melakukan peramalan dengan baik maka dibutuhkan berbagai macam informasi (data) yang cukup banyak dan diamati dalam periode waktu yang relatif cukup panjang, sehingga hasil analisis tersebut dapat mengetahui sampai berapa besar fluktuasi yang terjadi dan faktor-faktor apa saja yang memengaruhi terhadap perubahan tersebut.

Secara teoristis, dalam analisis runtun waktu (time series) hal yang paling menentukan adalah kualitas dan keakuratan dari data-data yang diperoleh, serta waktu atau periode dari data-data tersebut dikumpulkan. Jika data yang dikumpulkan tersebut semakin banyak maka semakin baik pula estimasi atau peramalan yang diperoleh. Sebaliknya, jika data yang dikumpulkan semakin sedikit maka hasil estimasi atau peramalannya akan semakin jelek.

Metode yang dapat digunakan untuk analisis time series ini adalah

- Metode Garis Linier Secara Bebas (Free Hand Method),
- Metode Setengah Rata-Rata (Semi Average Method),

- Metode Rata-Rata Bergerak (Moving Average Method) dan
- Metode Kuadrat Terkecil (Least Square Method).

## 8.2 Metode Trend Bebas

Cenderung digunakan sebagai analisis pendahuluan yang akan memberikan gambaran awal suatu permasalahan yang akan dihadapi, dengan mencoba melihat pola data pengamatan melalui tebaran titik dari pasangan data penjualan pada setiap waktunya, lalu diperkirakan trendnya. Dengan metode ini hanya bisa mendapat gambaran trendnya saja tanpa dapat memperkirakan jumlahnya.

## 8.3 Metode Trend Semi Average

Metode ini dapat digunakan bila datanya genap sehingga dapat dibagi jadi 2 kelompok.

Langkah-langkahnya :

- a. Data dibagi jadi dua kelompok
- b. Hitung rata-rata masing-masing kelompok
- c. Rata-rata kelompok satu disebut a (konstanta)
- d. Lalu hitung b, dengan cara menghitung selisih rata-rata kelompok 1 dengan rata-rata kelompok 2 dibagi jumlah data yang ada dalam suatu kelompok
- e. Buat persamaan  $Y = a + bX$

$Y$  = forecast tahun yang diinginkan,

$a$  = bilangan konstanta,

$b$  = selisih kelompok 1 dan 2 dibagi jumlah data dalam suatu kelompok

f. Menentukan nilai  $X = 0$  pada tahun pertengahan kelompok 1. Jika data genap data selanjutnya dimulai dengan 1 dengan jarak dua. Jika data ganjil, setelah nol adalah 2 dengan jarak dua.

#### 8.4 Metode Trend Moment

Dalam penerapannya data tidak harus genap. Metode ini memberikan nilai  $X$  mulai dari nol untuk tahun awal dan selanjutnya berurutan.

Rumusnya adalah :

$$\Sigma Y = a.n + b.\Sigma X$$

$$\Sigma XY = a. \Sigma X + b.\Sigma X^2$$

#### 8.5 Metode Trend Kuadrat Terkecil (Least Square)

Metode Least Square : Metode yang digunakan untuk analisis time series adalah Metode Garis Linier Secara Bebas (Free Hand Method), Metode Setengah Rata-Rata (Semi Average Method), Metode Rata-Rata Bergerak (Moving Average Method) dan Metode Kuadrat Terkecil (Least Square Method). Dalam hal ini akan lebih dikhususkan untuk membahas analisis time series dengan

metode kuadrat terkecil yang dibagi dalam dua kasus, yaitu kasus data genap dan kasus data ganjil. Secara umum persamaan garis linier dari analisis time series adalah :  $Y = a + b X$ . Keterangan : Y adalah variabel yang dicari trendnya dan X adalah variabel waktu (tahun).

Sedangkan untuk mencari nilai konstanta (a) dan parameter (b) adalah :

$$a = \frac{\sum Y}{N} \text{ dan}$$

$$b = \frac{\sum XY}{\sum X^2}$$

Langkah-langkahnya :

a. Tentukan nilai X apakah data ganjil atau genap. Kalau ganjil, nilai  $X = 0$  pada pertengahan data dan tahun sebelumnya dikurang 1 dan tahun sesudah ditambah satu dengan jarak 1 (-2,-1,0,1,2). Jika data genap maka nilai nol terletak di antara 2 tahun pertengahan dengan jaraknya dua (-5,-3,-1,1,3,5)

b. Rumusnya :

$$a = \frac{\sum Y}{n}$$

$$b = \frac{\sum XY}{\sum X^2}$$

c. Buat persamaan  $Y = a + bX$

Contoh Kasus Data Ganjil :

Tabel : Volume Penjualan Barang "X" (dalam 000 unit)  
Tahun 1995 sampai dengan 2003

Tahun	Penjualan (Y)	X	XY	X <sup>2</sup>
1995	200	- 4	- 800	16
1996	245	- 3	- 735	9
1997	240	- 2	- 480	4
1998	275	- 1	- 275	1
1999	285	0	0	0
2000	300	1	300	1
2001	290	2	580	4
2002	315	3	945	9
2003	310	4	1.240	16
Jumlah	2.460		775	60

Untuk mencari nilai a dan b adalah sebagai berikut :

$$a = 2.460 / 9 = 273,33 \text{ dan}$$

$$b = 775 / 60 = 12,92$$

Persamaan garis liniernya adalah :  $Y = 273,33 + 12,92 X$ .  
 Dengan menggunakan persamaan tersebut, dapat diramalkan penjualan pada tahun 2010 adalah :  $Y = 273,33 + 12,92$   
 (untuk tahun 2010 nilai X adalah 11), sehingga :  $Y = 273,33$   
 $+ 142,12 = 415,45$  artinya penjualan barang “X” pada tahun  
 2010 diperkirakan sebesar 415.450 unit

Contoh Kasus Data Genap:

Tabel : Volume Penjualan Barang “X” (dalam 000 unit)  
 Tahun 1995 sampai dengan 2002

Tahun	Penjualan (Y)	X	XY	X <sup>2</sup>
1995	200	- 7	- 1.400	49
1996	245	- 5	- 1.225	25
1997	240	- 3	- 720	9
1998	275	- 1	- 275	1
1999	285	1	285	1
2000	300	3	900	9
2001	290	5	1.450	25
2002	315	7	2.205	49
Jumlah	2.150		1.220	168

Untuk mencari nilai a dan b adalah sebagai berikut :

$$a = 2.150 / 8 = 268,75 \text{ dan}$$

$$b = 1.220 / 168 = 7,26$$

Persamaan garis liniernya adalah :  $Y = 268,75 + 7,26$

X. Berdasarkan persamaan tersebut untuk meramalkan penjualan pada tahun 2008 adalah :  $Y = 268,75 + 7,26$  (untuk tahun 2008 nilai X adalah 19), sehingga :  $Y = 268,75 + 137,94 = 406,69$  artinya penjualan barang "X" pada tahun 2008 diperkirakan sebesar 406,69 atau 406.690 unit.

Selain dengan menggunakan metode tersebut di atas, juga dapat dipakai dengan metode sebagai berikut :

Tabel : Volume Penjualan Barang "X" (dalam 000 unit)  
Tahun 1995 sampai dengan 2002

Tahun	Penjualan (Y)	X	XY	X <sup>2</sup>
1995	200	- 3	- 700	12,25
1996	245	- 2 ½	- 612,5	6,25
1997	240	- 1 ½	- 360	2,25
1998	275	- ½	- 137,5	0,25
1999	285	½	142,5	0,25
2000	300	1 ½	450	2,25
2001	290	2 ½	725	6,25
2002	315	3 ½	1102,5	12,25
Jumlah	2.150		610,0	42,00



Untuk mencari nilai a dan b adalah sebagai berikut :

$$a = 2.150 / 8 = 268,75 \text{ dan}$$

$$b = 610 / 42 = 14,52$$

Persamaan garis liniernya adalah :  $Y = 268,75 + 14,52 X$ . Berdasarkan persamaan tersebut untuk meramalkan penjualan pada tahun 2008 adalah :  $Y = 268,75 + 14,52$  (untuk tahun 2008 nilai X adalah  $9\frac{1}{2}$ ), sehingga :  $Y = 268,75 + 137,94 = 406,69$  artinya penjualan barang “X” pada tahun 2008 diperkirakan sebesar 406.690 unit

## 8.6 Variasi musiman

Salah satu komponen yang mempengaruhi data *time series* adalah komponen musiman. Gerakan musiman (*seasonal movement*) merupakan gerakan yang teratur artinya naik turunnya terjadi pada waktu-waktu yang sama. Disebut gerakan musiman oleh karena terjadinya bertepatan dengan pergantian musim didalam satu tahun atau dalam waktu yang singkat. misal:

- Harga beras akan turun pada saat musim panen padi.
- Penjualan buku akan meningkat pada awal sekolah.
- Jumlah pengunjung ke gedung bioskop akan naik pada malam minggu.

Jika data *time series* dipengaruhi oleh variasi musiman, maka diperlukan metoda peramalan yang lebih baik yang memperhatikan keterlibatan variasi musiman didalam data.

Untuk keperluan analisa seringkali data *time series* dinyatakan dalam bentuk angka indeks. Apabila kita ingin menunjukkan ada tidaknya gerakan musiman perlu dibuat indeks musiman (*seasonal index*). Indeks musiman adalah suatu angka yang bervariasi terhadap nilai dasar 100. Jika suatu periode musiman mempunyai nilai indeks 100, nilai ini menunjukkan bahwa pada bulan tersebut tidak ada pengaruh musiman. Ada beberapa metode untuk menghitung angka indeks musiman, antara lain adalah metode rata-rata sederhana (*simple average method*).

### **8.7 Mencari indeks musiman dengan metoda rata-rata sederhana**

Indeks musiman dapat digunakan untuk menguraikan perkiraan/ ramalan penjualan tahunan menjadi perkiraan penjualan per bulan pada tahun mendatang. Untuk mencari indeks musiman dengan metode rata-rata sederhana, pertama perlu dicari nilai rata-rata untuk setiap bulannya dengan maksud untuk menghilangkan pengaruh trend. Berapa banyak tahun yang digunakan untuk mendapatkan nilai rata-rata tergantung dari banyak tahun terulangnya gerakan siklis yang maksudnya untuk menghilangkan pengaruh dari gerakan siklis (misal: 5 tahun, 10 tahun atau lebih).

Dari nilai rata-rata tersebut selanjutnya dicari besaran persentasenya terhadap total atau jumlah nilai rata-rata dimana jumlah nilai rata-rata tersebut menjadi nilai 100 dalam besaran persentase. Indeks musiman didapat dengan cara mengalikan besaran persentase masing-masing bulan dengan konstanta 12.

Tabel 2. data bulanan selama 11 tahun hasil penjualan (miliar Rp) PT. "Maju Terus".

<u>Bln</u>	<u>Tahun</u>										
	'00	'01	'02	'03	'04	'05	'06	'07	'08	'09	'10
<u>Jan</u>	3,8	2.8	3.3	3.4	3.0	2.9	2.5	3.0	3.2	3.1	2.6
<u>Peb</u>	4,0	2.9	3.4	3.5	3.1	3.0	2.6	3.1	3.3	3.3	2.7
<u>Mar</u>	3,9	2.9	3.3	3.5	3.2	3.0	2.5	3.0	3.3	3.2	2.7
<u>Apr</u>	3,9	3.0	3.3	3.6	3.2	3.0	2.6	3.2	3.4	3.3	2.7
<u>Mei</u>	4,2	3.0	3.5	3.6	3.3	3.1	2.7	3.2	3.4	3.4	2.8
<u>Jun</u>	4,2	3.1	3.6	3.7	3.2	3.1	2.7	3.3	3.5	3.5	2.9
<u>Jul</u>	4,3	3.1	3.7	3.8	3.3	3.2	2.8	3.2	3.5	3.5	2.9
<u>Agst</u>	4,3	3.1	3.7	3.9	3.3	3.4	2.8	3.3	3.6	3.6	2.8
<u>Sept</u>	4,3	3.2	3.8	3.9	3.3	3.4	2.9	3.4	3.7	3.6	2.9
<u>Okt</u>	4,4	3.2	3.8	4.0	3.4	3.5	2.9	3.4	3.7	3.6	3.0
<u>Nov</u>	4,4	3.2	3.8	3.9	3.4	3.4	2.9	3.4	3.7	3.6	3.0
<u>Des</u>	4,3	3.0	3.6	3.7	3.2	3.1	2.7	3.2	3.4	3.4	2.8
<u>jml</u>	50.0	36.5	43.0	44.5	38.9	38.1	32.6	38.2	41.7	41.1	33.8

Contoh 2.: menggunakan data pada tabel 2.

<u>Bulan</u>	<u>Jumlah penjualan</u>	<u>Rata-Rata Penjualan</u>	<u>Persentase terhadap Total</u>	<u>Indeks Musiman</u>
Jan	33,6	3,05	7,64*	91,68
Feb	34,9	3,17	7,94	95,28
....	....	....	....	....
Nov	38,7	3,52	8,82	105,84
Des	36,4	3,31	8,30	99,6
Jumlah	-	39,9	100	1200

Indeks musiman = persentase thd total x 12

$$* \frac{3.05}{39.9} \times 100\%$$

Perkiraan hasil penjualan bulanan tahun 2011 dapat dinyatakan sebagai berikut:

$$\frac{\text{indeks musiman}}{1200} \times \text{perkiraan tahunan}$$

$$\text{Januari: } \frac{91.68}{1200} \times \text{Rp. } 35,3 \text{ miliar} = \text{Rp. } 2,7 \text{ miliar}$$

$$\text{Februari: } \frac{95.28}{1200} \times \text{Rp. } 35,3 \text{ miliar} = \text{Rp. } 2,8 \text{ miliar}$$

.....

$$\text{Nopember: } \frac{105.84}{1200} \times \text{Rp. } 35,3 \text{ miliar} = \text{Rp. } 3,1 \text{ miliar}$$

$$\text{Desember: } \frac{99.60}{1200} \times \text{Rp. } 35,3 \text{ miliar} = \text{Rp. } 2,9 \text{ miliar}$$

Rp 35,3 miliar merupakan perkiraan hasil penjualan tahun 2011 yang diprediksi dengan menggunakan trend.

## 8.8 Metoda Dekomposisi

Dekomposisi adalah suatu prosedur dalam menganalisa data serial waktu dengan cara mengidentifikasi faktor-faktor komponen yang ada dalam suatu periode data. Setiap komponen diidentifikasi secara terpisah sehingga pola serial waktu dapat digunakan untuk peramalan kegiatan masa depan baik untuk jangka pendek maupun jangka panjang.

Pada dasarnya ada 3 komponen yang membentuk pola suatu data serial waktu. Ketiga komponen tersebut adalah gerakan trend, musiman (*seasonal*) dan siklis (*cyclical*). Dekomposisi mengasumsikan bahwa data dibentuk seperti berikut ini:

$$\begin{aligned}\text{Data} &= \text{Pola} + \text{Error} \\ &= \text{fungsi (trend, musiman, siklis)} + \text{Error}\end{aligned}$$

- **Trend** adalah suatu gerakan yang menunjukkan arah perkembangan (kecenderungan menaik atau menurun).
- **Gerakan musiman** adalah suatu gerakan yang mempunyai pola tetap dari waktu ke waktu.
- **Gerakan siklis** adalah gerakan jangka panjang disekitar garis trend (berlaku untuk data tahunan), gerakan siklis ini akan terulang dalam jangka waktu tertentu atau bisa juga dalam jangka waktu yang tidak sama.
- **Error/irregular** adalah gerakan yang sporadis atau yang tidak tertentu. Gerakan ini ditimbulkan oleh suatu kejadian yang tak terduga seperti perang, gempa bumi dan sebagainya.

Apabila gerakan trend, musiman, siklis dan error masing-masing diberi simbol T, S, C dan I maka data serial waktu Y merupakan hasil kali dari 4 komponen tersebut, yaitu:

$$Y = T \times S \times C \times I$$

## **BAB IX**

### **REGRESI DAN KORELASI**

#### **9.1 Pendahuluan**

Penggunaan metode ini didasarkan kepada variabel yang ada dan yang akan mempengaruhi hasil peramalan.

Hal- hal yang perlu diketahui sebelum melakukan peramalan dengan metode regresi adalah mengetahui terlebih dahulu mengetahui kondisi- kondisi seperti :

- a. Adanya informasi masa lalu
- b. Informasi yang ada dapat dibuatkan dalam bentuk data (dikuantifikasikan)
- c. Diasumsikan bahwa pola data yang ada dari data masa lalu akan berkelanjutan dimasa yang akan datang.

Adapun data- data yang ada dilapangan adalah :

- a. Musiman (Seasonal)
- b. Horizontal (Stationary)
- c. Siklus (Cylikal)
- d. Trend

Dalam menyusun ramalan pada dasarnya ada 2 macam analisis yang dapat digunakan yaitu :

1. Analisi deret waktu(Time series), merupakan analisis antarvariabel yang dicari dengan variabel waktu

2. Analisis Cross Section atau sebab akibat (Causal method), merupakan analisis variabel yang dicari dengan variabel bebas atau yang mempengaruhi.

Ada dua pendekatan untuk melakukan peramalan dengan menggunakan analisis deret waktu dengan metode regresi sederhana yaitu :

1. Analisis deret waktu untuk regresi sederhana linier
2. Analisis deret untuk regresi sederhana yang non linier

Untuk menjelaskan hubungan kedua metode ini kita gunakan notasi matematis seperti:

$$Y = F(x)$$

Dimana :

Y = Dependent variable (variabel yang dicari)

X = Independent variable (variabel yang mempengaruhinya)

Notasi regresi sederhana dengan menggunakan regresi linier (garis lurus) dapat digunakan sebagai berikut :

$$Y = a + b x$$

Dimana a dan b adalah merupakan parameter yang harus dicari. Untuk mencari nilai a dapat digunakan dengan menggunakan rumus :

$$a = \bar{Y} - b \bar{x}$$

kemudian nilai b dapat dicari dengan rumus :



$$b = \frac{\sum XY - \bar{X} \bar{Y}}{\sum X^2 - \bar{X} \sum X}$$

- Persamaan regresi: Persamaan matematik yang memungkinkan peramalan nilai suatu peubah tak-bebas (*dependent variable*) dari nilai peubah bebas (*independent variable*)
- Diagram Pencar = *Scatter Diagram*

Diagram yang menggambarkan persebaran nilai-nilai observasi peubah tak-bebas dan peubah bebas.

Nilai peubah bebas: ditulis pada sumbu X  
(sumbu horizontal)

Nilai peubah takbebas: ditulis pada sumbu Y  
(sumbu vertikal)

Nilai peubah tak-bebas ditentukan oleh nilai peubah bebas

Anda sudah dapat menentukan mana peubah tak-bebas dan peubah bebas?

Contoh 1:

Umur vs Tinggi Tanaman (X : Umur, Y : Tinggi)

Biaya Promosi vs Volume penjualan (X : Biaya Promosi, Y : Vol. penjualan)

### **Analisis Dengan Regresi Linier Cross Section**

Cross section method atau casual method atau sebab akibat merupakan peramalan yang kita lakukan

untuk mengukur peramalan dalam suatu periode dengan faktor yang mempengaruhinya bukan waktu.

Penggunaan rumusan yang kita gunakan untuk cross section sama dengan penerapan untuk metode time series, begitu pula dengan hasil peramalannya.

Jadi penjualan =  $f(x, x, x, \dots)$

X = harga, mutu pendapatan, promosi dll

$$Y = a + b x$$

Dimana x adalah variabel bukan waktu.

### **Analisis Deret Waktu Dengan Regresi**

Ada 2 pendekatan untuk melakukan peramalan dengan menggunakan analisis deret waktu dengan metode regresi sederhana, yaitu :

1. Analisis deret waktu untuk regresi sederhana linier
  2. Analisis deret waktu untuk regresi sederhana yang non linier
- Jenis-jenis Persamaan Regresi :
    - a. Regresi Linier :
      - Regresi Linier Sederhana
      - Regresi Linier Berganda
    - b. Regresi Non-linier
      - Regresi Eksponensial
  - a) Regresi Linier
    - Bentuk Umum Regresi Linier Sederhana

$$Y = a + bX$$

Y : peubah tak-bebas

X : peubah bebas

a : konstanta

b : kemiringan

- Bentuk Umum Regresi Linier Berganda

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Y : peubah tak-bebas                      a :  
konstanta

X<sub>1</sub> : peubah bebas ke-1                      b<sub>1</sub> :  
kemiringan garis ke-1

X<sub>2</sub> : peubah bebas ke-2                      b<sub>2</sub> :  
kemiringan garis ke-2

X<sub>n</sub> : peubah bebas ke-n                      b<sub>n</sub> :  
kemiringan garis ke-n

b) Regresi Non Linier

- Bentuk umum Regresi Eksponensial

$$Y = ab^x$$

$$\log Y = \log a + (\log b) x$$

Dalam analisis deret waktu yang linier adalah analisis pola hubungan yang dicari dengan satu variabel yang mempengaruhinya : waktu. Sedangkan analisis deret waktu yang non linier, merupakan analisis hubungan antara variabel yang dicari dengan hanya satu (1) yang mempengaruhinya, yaitu variabel waktu.

Untuk menjelaskan hubungan kedua metode ini kita gunakan notasi matematis seperti :

$$Y = F(x)$$

Dimana :

Y = Dependent variable (variabel yang dicari)

X = Independent variable (variabel yang mempengaruhinya)

Notasi regresi sederhana dengan menggunakan regresi linier (garis lurus) dapat digunakan sebagai berikut :

$$Y = a + b X \dots\dots\dots$$

Dimana a dan b adalah merupakan parameter (koefisien regresi) yang harus dicari. Untuk mencari nilai a dapat digunakan dengan menggunakan rumus :

$$a = \frac{\sum Y}{n} - b \frac{\sum X}{n}$$

atau :

$$a = \bar{Y} - b \bar{X}$$

kemudian nilai b dapat dicari dengan rumus :

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2}$$

atau

$$b = \frac{\sum XY - \bar{X} \sum Y}{\sum X^2 - \bar{X} \sum X}$$

Langkah selanjutnya adalah melakukan pengujian terhadap hasil yang diperoleh dengan :

*Pertama*, uji Test Koefisien Penentu ( $R^2$ ), pengujian ini untuk mengetahui tepat tidaknya variabel yang mempengaruhi besarnya penjualan yang diramalkan adalah waktu.

*Kedua*, Test Significance (T.Test) atau F test yaitu pengujian untuk mengetahui apakah benar persamaan regresi itu adalah linier.

### ***Pengujian $R^2$ .***

Test rumusan yang digunakan adalah:

$$R^2 = 1 - \frac{\sum e^2}{\sum y^2}$$

Dimana,

$$\sum e^2 = Y^2 - b^2 (\sum x^2)$$

$$\sum y^2 = Y^2 - n (\bar{Y})^2$$

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

### ***Test Significance***

Tujuan test ini menguji dan meneliti apakah regresi yang digunakan dalam menyusun ramalan adalah benar linier, dimana data yang diteliti tepat berada disekitar garis linier.

#### 1. F. Test

Tujuannya adalah untuk mengetahui apakah nilai estimasi dari a dan b dapat bervariasi karena pengaruh sampling/ random.

Persamaan F. Test adalah sebagai berikut :

$$F = \frac{\sum (Y - \hat{Y})^2}{k - 1} \bigg/ \frac{\sum (Y - \bar{Y})^2}{n - k}$$

Dimana :

k = jumlah variabel (dalam regresi sederhana = 2)

n = jumlah tahun

atau kita juga menggunakan rumus sebagai berikut

:

$$F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}}$$

Hasil  $F_{\text{rasio}}$  kemudian kita bandingkan dengan  $F_{\text{tabel}}$  apabila  $F_{\text{rasio}} > F_{\text{tabel}}$ , maka secara statistik koefisien b adalah *significance* berbeda dengan nol (0), sehingga persamaan regresi dapat dilakukan secara benar dengan bentuk persamaan sebagai berikut :

$$Y = a + b x$$

Demikian pula sebaliknya jika  $F_{\text{rasio}} < F_{\text{tabel}}$

## 2. Persamaan T. Test

Test ini dikenal dengan nama student-t didistribusikan untuk menguji a dan b dengan formula :

$$t_{\text{test a}} = \frac{a}{\tau a} \qquad t_{\text{test b}} =$$

Hasilnya jika diperoleh :

$T_{\text{test}} > T_{\text{tabel}}$  ( $T_{\text{distribusi}}$ ), maka tingkat keyakinan tertentu (R) dapat disimpulkan bahwa nilai koefisien regresi a dan b secara statistic berbeda dari (0) dan demikian pula sebaliknya.

## 9.2 Regresi Linier Sederhana

Metode Kuadrat terkecil (*least square method*): metode yang paling populer untuk menetapkan persamaan regresi linier sederhana.

- Bentuk Umum Regresi Linier Sederhana :

$$Y = a + bX$$

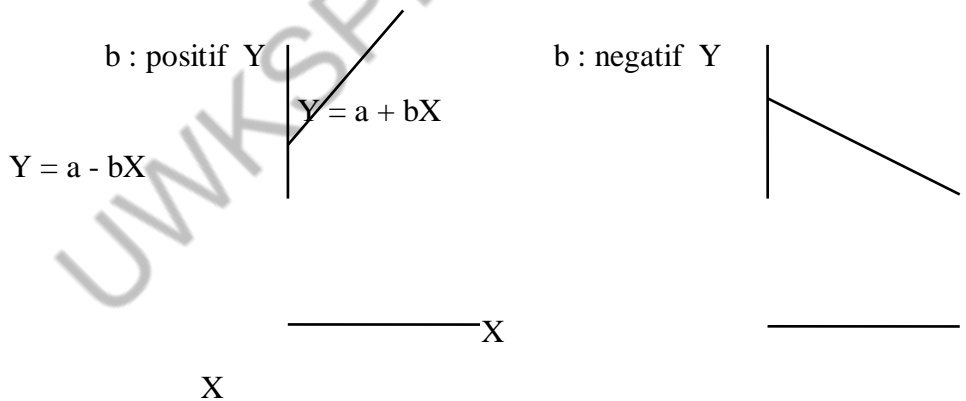
Y : peubah tak-bebas  
peubah bebas

X :

a : konstanta  
kemiringan

b :

Nilai b dapat positif (+) dapat negatif (-)



Penetapan Persamaan Regresi Linier Sederhana



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

$$a = \bar{y} - b\bar{x} \quad \text{sehingga} \quad a = \frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n x_i}{n}$$

$n$  : banyaknya pasangan data

$y_i$  : nilai peubah tak bebas Y ke- $i$

$x_i$  : nilai peubah bebas X ke- $i$

Contoh 2:

Berikut adalah data Biaya Promosi dan Volume Penjualan PT BIMOIL perusahaan Minyak Gosok.

Tahun	X Biaya Promosi (Juta Rupiah)	Y Volume Penjualan (Ratusan Juta Liter)	xy	$x^2$	$y^2$
1992	2	5	10	4	25
1993	4	6	24	16	36
1994	5	8	40	25	64
1995	7	10	70	49	100
1996	8	11	88	64	121

$\Sigma$	$\Sigma x = 26$	$\Sigma y = 40$	$\Sigma xy = 232$	$\Sigma x^2 = 158$	$\Sigma y^2 = 346$
----------	-----------------	-----------------	-------------------	--------------------	--------------------

$$n = 5$$

bentuk umum persamaan regresi linier sederhana :  $Y = a + bX$

$$b = \frac{5 \times 232 - (26)(40)}{5 \times 158 - 26^2} = \frac{1160 - 1040}{790 - 676} = \frac{120}{114} = 1,053$$

$$a = \frac{40}{5} - 1,053 \frac{26}{5} = 8 - 5,470 = 2,530$$

$$Y = a + b X \rightarrow Y = 2,530 + 1,053 X$$

### Peramalan dengan Persamaan Regresi

Contoh 3:

Diketahui hubungan Biaya Promosi (X dalam Juta Rupiah) dan Y (Volume penjualan dalam Ratusan Juta liter) dapat dinyatakan dalam persamaan regresi linier berikut:

$$Y = 2,530 + 1,053 X$$

Perkirakan Volume penjualan jika, dikeluarkan biaya promosi Rp. 10 juta?

Jawab :

$$Y = 2,530 + 1,053 X$$

$$X = 10$$

$$Y = 2,53 + 1,053 (10) = 2,53 + 10,53 = 13,06 \text{ (ratusan juta liter)}$$

$$\text{Volume penjualan} = 13,06 \times 100\,000\,000 \text{ liter}$$

### 9.3 Korelasi Linier Sederhana

Koefisien Korelasi (r): ukuran hubungan linier peubah X dan Y

Nilai r berkisar antara (+1) sampai (-1)

Nilai r yang (+) ditandai oleh nilai b yang (+)

Nilai r yang (-) ditandai oleh nilai b yang (-)

Jika nilai r mendekati +1 atau r mendekati -1 maka X dan Y memiliki korelasi linier yang tinggi.

Jika nilai r = +1 atau r = -1 maka X dan Y memiliki korelasi linier sempurna.

Jika nilai  $r = 0$  maka X dan Y tidak memiliki relasi (hubungan) linier (dalam kasus  $r$  mendekati 0, anda dapat melanjutkan analisis ke regresi eksponensial).

Koefisien Determinasi Sampel =  $R = r^2$

Ukuran proporsi keragaman total nilai peubah Y yang dapat dijelaskan oleh nilai peubah X melalui hubungan linier.

Penetapan & Interpretasi Koefisien Korelasi dan Koefisien Determinasi

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left[ n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right] \left[ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right]}}$$

Contoh 4:

Lihat Contoh 2, setelah mendapatkan persamaan Regresi  $Y = 2.530 + 1.053 X$ , hitung koefisien korelasi ( $r$ ) dan koefisien determinasi ( $R$ ). Gunakan data berikut (lihat Contoh 2)

$$\Sigma x = 26 \quad \Sigma xy = 232 \quad \Sigma y^2 = 346$$

$$\Sigma y = 40 \quad \Sigma x^2 = 158$$

Nilai  $r = 0,9857$  menunjukkan bahwa peubah X (biaya promosi) dan Y (volume penjualan) berkorelasi linier yang positif dan tinggi

$$= 0,97165\dots = 97 \%$$

Nilai  $R = 97\%$  menunjukkan bahwa 97% proporsi keragaman nilai peubah Y (volume penjualan) dapat dijelaskan oleh nilai peubah X (biaya promosi) melalui hubungan linier. Sisanya, yaitu 3 % dijelaskan oleh hal-hal lain.

UWKSPRES

## DAFTAR PUSTAKA

(<http://www.psychologymania.com/2013/01/macam-macam-indeks.html>)

(<http://ekonomi-holic.blogspot.com/2012/06/pengertian-ciri-dan-peranan-indeks.html>)

Anton Dajan ( 1981). *Pengantar Metode Statistik Jilid I.* Jakarta: Lembaga Penelitian, Pendidikan dan Penerangan Ekonomi dan Sosial

Djarwanto (2013). *Statistik Induktif*, Edisi edisi 5. Penerbit BPEP, Yogyakarta.

Hartono (2016), *Analisis Data Statistik dan Penelitian*, Cetakan ke VII. SPSS 16.0, Penerbit Pustaka Pelajar, Yogyakarta.

Husaini Usman (2020), *Pengantar Statitik, Cara Mudah Memahami Statitik*, Edisi ke 3. Penerbit Bumi Aksara, Gramedia, Jakarta.

Indian Agricultural Statistics Research Institute: <http://www.iasri.res.in/>

Lindley, D. Making Decisions. John Wiley. Second Edition 1985. ISBN 0-471-90808-8

Mario Triola. 2004. *Elementary Statistiks*. 9<sup>th</sup> Edition. Pearson Education.

Ronald E.Walpole. *Pengantar Statistika, halaman 22-27"*. 1993. Jakarta : PT Gramedia Pustaka Utama. ISBN 979-403-313-8

Singgih Santoso (2012). *Statistik Parametrik*, Elex Media Komputindo, Jakarta.

----- (2022). *Panduan Lengkap SPSS 26*, Percetakan PT. Gramedia, Jakarkata.

Statistikal dispersion: [http://en.wikipedia.org/wiki/Statistikal\\_dispersion](http://en.wikipedia.org/wiki/Statistikal_dispersion)

Stephen Bernstein and Ruth Bernstein. (1999). *Elements of Statistiks I: Descriptive Statistiks and Probability*. The McGraw-Hill Companies, Inc

Sudjana (1992). *Peranan Statistik dalam Penelitian*. Penerbit Tarsito, Universitas Muhammadiyah. Purworejo.

Sugiono.(2010). *Statistik untuk Penelitian* Jakarta: Alfabeta.

UWKSPRESS