

PENGANTAR

STATISTIK

PENDIDIKAN

Sudarman



# STATISTIK PENDIDIKAN



**Sudarman**

# STATISTIK PENDIDIKAN

## Sudarman

Editor : Susilo  
Triana Fitriastuti  
Kiswanto

Tata Letak : Agus Fitrianto

ISBN : 978-602-6834-02-7

© 2015. Mulawarman University Press

Cetakan Pertama : November 2015

Hak Cipta Dilindungi Undang-Undang.  
Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apapun tanpa izin tertulis dari penerbit

Isi di luar tanggung jawab percetakan.

Sudarman. 2015. Statistik Pendidikan. Mulawarman  
University Press. Samarinda



**Mulawarman**  
**University PRESS**

Penerbit

Mulawarman University PRESS

Gedung LPPM Universitas Mulawarman

Jl. Krayan, Kampus Gunung Kelua

Samarinda – Kalimantan Timur – INDONESIA 75123

Telp/Fax (0541) 747432; Email : [mup@lppm.unmul.ac.id](mailto:mup@lppm.unmul.ac.id)

# PRAKATA

Buku ini pada awalnya merupakan hasil pengembangan bahan ajar sebagai pendukung instrumen riset disertasi saya. Bahan ajar itu menjadi referensi perkuliahan mahasiswa Program Studi Pendidikan Ekonomi FKIP Universitas Mulawarman dan telah mengalami berbagai revisi dari diskusi dan masukan baik oleh mahasiswa maupun rekan sejawat guna mencapai tingkat kedalaman dan keluasan yang memadai sebagai sumber belajar mata kuliah pengantar statistik pendidikan.

Buku teks ini terdiri dari tujuh bab, menyajikan statistik pendidikan secara menyeluruh berdasarkan kompetensi dasar yang harus ditempuh dan dicapai kompetensinya selama satu semester. Adapun butir-butir inti dari setiap bab adalah:

**Bab 1. Konsep Statistika** yang mengupas konsep statistika hingga sumber data statistik, sehingga pemahaman konsep statistik dan statistika itu menjadikan tidak akan terjadi lagi penyalahgunaan dari dua kata tersebut. Selain itu, bab ini juga membahas kegunaan dan peranan statistik dalam kehidupan sehari-hari, dan pembahasan tentang jenis-jenis statistik, jenis data serta berbagai sumber untuk mendapatkan data statistik.

**Bab 2. Penyajian Data** yang mengantarkan untuk menguasai teknik penyajian data berbentuk tabel dan grafik. Bagian ini juga menampilkan penyajian data dalam tabel (struktur dan macam-macam tabel) dan berbentuk gambar (diagram batang, histogram, diagram lingkaran, lambang, garis dan peta).

**Bab 3. Distribusi Frekuensi** yang menawarkan bahasan tentang penyajian data dengan tabel distribusi frekuensi, mengajak untuk memahami konsep distribusi frekuensi melalui pengertian serta

cara penyusunan tabel distribusi frekuensi baik data tunggal maupun data kelompok. Selanjutnya bab ini juga menguraikan bahasan tentang penyajian grafik distribusi frekuensi baik bentuk histogram, poligon, dan ogive. Akhir bab ini menguraikan tentang distribusi frekuensi relatif dan kumulatif.

**Bab 4. Ukuran Pemusatan Data** yang membahas konsep dan rumus ukuran pemusatan, definisi dan perlunya pemahaman ukuran pemusatan, dilanjutkan pemaparan tentang rata-rata hitung, rata-rata data tidak dikelompokkan maupun berkelompok serta rata-rata harmohik. Kupasan lain adalah median dan hubungan ketiganya, kuartil, desil dan persentil.

**Bab 5. Ukuran Penyebaran** yang mengantar pemahaman tentang variabilitas. Bab ini mengulas tentang definisi variabilitas dan ukuran variabilitas (range, deviasi rata-rata, deviasi standar, dan nilai standar dan nilai yang dibakukan).

**Bab 6. Momen, Skewness dan Kurtosis** yang membahas tentang momen, kemiringan, dan kurtosis yang membantu penggunaan statistik deskriptif dalam pendidikan maupun ekonomi, sekaligus memperkuat pengetahuan tentang konsep statistik deskriptif.

**Bab 7. Regresi dan Korelasi Sederhana** yang membahas tentang pemahaman konsep regresi linier, korelasi, penggunaan rumus dari berbagai teknik korelasi, baik skala data nominal, ordinal maupun rasio serta keterkaitan antara regresi dan korelasi.

Masing-masing bab disajikan ringkasan dan konsep kunci, yang didukung contoh dan latihan soal untuk mempermudah pemahaman terhadap masing-masing konsep dan prosedur setiap bahasan. Akhirnya, semoga buku teks ini bisa bermanfaat dalam pengembangan pembelajaran statistik pendidikan.

Samarinda, November 2015

**Sudarman**

# DAFTAR ISI

PRAKATA .....	iii
DAFTAR ISI .....	v
DAFTAR TABEL .....	vii
DAFTAR GAMBAR .....	xi
<b>BAB 1. KONSEP STATISTIKA</b>	
A. Pengantar .....	1
B. Pengertian Statistik dan Statistika .....	3
C. Peranan dan Kegunaan Statistik .....	5
D. Jenis-Jenis Statistik .....	10
E. Data .....	12
F. Sumber Data Statistik.....	17
G. Ringkasan .....	17
H. Soal dan Latihan .....	18
<b>BAB 2. PENYAJIAN DATA</b>	
A. Pengantar .....	19
B. Penyajian Data dalam Bentuk Tabel .....	21
C. Penyajian Data dalam Bentuk Grafik .....	27
D. Ringkasan .....	36
E. Soal dan Latihan .....	37
<b>BAB 3. DISTRIBUSI FREKUENSI</b>	
A. Pengantar .....	38
B. Distribusi Frekuensi Data Kualitatif .....	40
C. Persentase Data Kualitatif .....	41
D. Distribusi Frekuensi Data Kuantitatif.....	42
E. Penyajian Grafik Distribusi Frekuensi.....	49
F. Distribusi Frekuensi Relatif .....	52
G. Distribusi Frekuensi Kumulatif .....	52
H. Distribusi Frekuensi Kumulatif Relatif .....	55
I. Ringkasan .....	56
J. Soal dan Latihan .....	57
<b>BAB 4. UKURAN PEMUSATAN DAN UKURAN LETAK</b>	
A. Pengantar .....	58
B. Pengertian .....	59
C. Rata-Rata Hitung (Arithmetic Mean) .....	60

D. Rata-Rata Ukur (Geometric Mean) .....	68
E. Rata-Rata Harmonik .....	74
F. Median .....	76
G. Modus .....	81
H. Hubungan Rata-Rata, Median dan Modus .....	83
I. Ukuran Letak (Kuartil, Desil dan Persentil) .....	86
J. Ringkasan .....	94
K. Soal dan Latihan .....	95
<b>BAB 5. UKURAN PENYEBARAN</b>	
A. Pengantar .....	97
B. Jangkauan, Inter-Kuartil dan Deviasi Kuartil .....	99
C. Deviasi Rata-Rata (Rata-Rata Simpangan) .....	103
D. Pengukuran Varians dan Simpangan Baku .....	108
E. Nilai yang Dibakukan (Standardized Value) .....	116
F. Dispersi Relatif, Koefisien Variasi .....	118
G. Ringkasan .....	119
H. Soal dan Latihan .....	120
<b>BAB 6. MOMEN, SKEWNESS DAN KURTOSIS</b>	
A. Pengantar .....	122
B. Momen.....	123
C. Ukuran Kemencengan (Skewness) .....	126
D. Keruncingan (Kurtosis) .....	134
E. Ringkasan .....	137
F. Soal dan Latihan .....	138
<b>BAB 7. REGRESI LINIER DAN KORELASI SEDERHANA</b>	
A. Pengantar .....	139
B. Analisis Regresi Linier.....	141
C. Analisis Korelasi Sederhana .....	147
D. Teknik-Teknik Korelasi .....	151
E. Korelasi Data Kualitatif .....	156
F. Ringkasan .....	159
G. Soal dan Latihan .....	160
<b>DAFTAR PUSTAKA .....</b>	<b>162</b>

# DAFTAR TABEL

No.	Tubuh Utama	Halaman
2.1.	IPM di Beberapa Negara .....	21
2.2.	Skema Catatan Pendahuluan .....	22
2.3.	Jumlah Mahasiswa Program Studi Magister Pendidikan Ekonomi .....	25
2.4.	Penjualan PT UM Unggul Menurut Jenis Barang dan Daerah Penjualan Tahun 2012 .....	26
2.5.	Produksi Barang PT Nayamul Menurut Kualitasnya ....	26
2.6.	Distribusi Frekuensi Skor Kinerja Guru SMP Negeri ....	27
2.7.	Penjualan PT Harapan Insan 2006-2011 .....	30
2.8.	Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 .....	31
2.9.	Hubungan Kecerdasan dengan Hasil Produksi .....	35
3.1.	Data 50 Orang Pembeli Komputer dari Beberapa Jenis Produsen di Harco Computer Malang .....	40
3.2.	Distribusi Frekuensi Pembelian 5 Merk Komputer .....	41
3.3.	Distribusi Frekuensi Relatif dan Persentase Pembelian Komputer .....	42
3.4.	Distribusi Frekuensi Data Tunggal .....	44
3.5.	Hasil UTS Statistik Mahasiswa FKIP Unmul 2014 .....	46
3.6.	Distribusi Frekuensi Data Nilai Mata Kuliah Metodologi Penelitian dari 80 Orang Mahasiswa FKIP Unmul .....	49
3.7.	Distribusi Frekuensi Relatif Nilai Metodologi Penelitian	52
3.8.	Distribusi Frekuensi Kumulatif “Kurang Dari” .....	53



3.9. Distribusi Frekuensi Kumulatif “Atau Lebih”.....	53
3.10. Distribusi Frekuensi Kumulatif Relatif “Kurang Dari” .....	56
3.11. Distribusi Frekuensi Kumulatif Relatif “Atau Lebih” .....	56
4.1. Nilai Ekspor Minyak dan Gas Pada Periode Januari Hingga Juni 2012 .....	62
4.2. Transkrip Akademik Mahasiswa .....	62
4.3. Rata-Rata Berat Beras PT Makmur .....	63
4.4. Rataan Berat Beras PT Makmur (Metode Simpangan)...	64
4.5. Data Penimbangan Beras PT Makmur Menggunakan Metode Pengkodean .....	66
4.6. Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya (Tipe “Kurang Dari”) .....	66
4.7. Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya (Tipe Frekuensi Biasa) .....	67
4.8. Perhitungan Rata-Rata Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya (Metoda “Defisional”) .....	67
4.9. Perkembangan Harga Saham PT Inti.....	71
4.10. Rasio Perkembangan Harga Saham PT Inti .....	72
4.11. Perhitungan Rata-Rata Ukur Beras PT Makmur .....	73
4.12. Omset Penjualan Supermarket Daremantep Jaya .....	77
4.13. Omset 6 Supermarket Daremantep Jaya .....	81
4.14. Omset Penjualan PT Abadi Jaya .....	82
4.15. Berat Badan Mahasiswa .....	85
4.16. Nilai Frekuensi Berat Badan Mahasiswa .....	85
4.17. Distribusi Harga Tembakau .....	88

4.18. Titik-Titik Letak Desil.....	90
4.19. Modal 40 Perusahaan .....	91
4.20. Letak Beberapa Titik Persentil.....	92
4.21. Total Penjualan 50 Tenaga Penjual di Perusahaan Baldroc Selama Semester I Tahun 2012 .....	93
5.1. Volume Ekspor Komoditi Pertanian Menurut Wilayah ..	99
5.2. Nilai Range Volume Ekspor Komoditi Pertanian .....	99
5.3. Data Penjualan 6 Salesman CV Dharma Jaya .....	100
5.4. Perhitungan MD Nilai Penjualan di Surabaya .....	104
5.5. Perhitungan MD Nilai Penjualan di Malang .....	104
5.6. Data Harga Saham di BEJ .....	105
5.7. Deviasi Harga Saham di BEJ .....	106
5.8. Perhitungan Deviasi Rata-Rata Nilai Upah Karyawan...	107
5.9. Variasi Penjualan 6 Salesman CV Dharma Jaya .....	109
5.10. Perhitungan Nilai Varians .....	111
5.11. Perhitungan Simpangan Modal 40 Perusahaan .....	113
5.12. Penolong Perhitungan Simpangan .....	113
5.13. Penolong Perhitungan Simpangan Baku .....	114
6.1. Penolong Perhitungan Momen Berat 5 Ikan .....	125
6.2. Penolong Perhitungan Momen dari Daftar Distribusi ...	125
6.3. Nilai Ujian Statistik 40 Mahasiswa FKIP Unmul .....	128
6.4. Penolong Perhitungan Nilai Koefisien Kemencengan ....	128
6.5. Nilai Ujian Matematika Ekonomi 111 Mahasiswa .....	130
6.6. Penolong Perhitungan Kemencengan Data Berkelompok	133
6.7. Penolong Perhitungan Tingkat Keruncingan Kurva .....	136

6.8.	Tinggi 100 Mahasiswa Universitas XYZ .....	137
7.1.	Hasil Produksi dan Skor Tes Kecerdasan 8 Karyawan Pabrik Mainan Anak-anak “Tackey” .....	143
7.2.	Penolong Perhitungan Regresi.....	145
7.3.	Hubungan Biaya Periklanan dan Tingkat Penjualan ....	146
7.4.	Regresi Sederhana Hubungan Biaya Periklanan dan Tingkat Penjualan .....	146
7.5.	Penolong Perhitungan Korelasi .....	152
7.6.	Penolong Perhitungan Korelasi Metode Skor Deviasi ....	152
7.7.	Rank Hipotesis Nilai Buah Versi Nabil dan Agus .....	154
7.8.	Rangking Hasil Penjualan Berdasarkan Nilai Ujian .....	155
7.9.	Hubungan Tingkat Pendidikan dan Konsumsi Susu ...	158

# DAFTAR GAMBAR

No.	Tubuh Utama	Halaman
1.1.	Histogram Hasil Belajar Siswa Kelas Kontrol (Model Pembelajaran Konvensional) .....	11
1.2.	Tingkatan Data .....	16
2.1.	Jumlah Mahasiswa Baru Pada Program Magister Pendidikan Ekonomi di Universitas Negeri Malang Periode 2007 -2010 .....	28
2.2.	Grafik Perkembangan Penjualan .....	30
2.3.	Grafik Perkembangan Ekonomi Dunia .....	30
2.4.	Diagram (Lingkaran) Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 .....	32
2.5.	Diagram (Kue yang Diledakkan) Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 .....	33
2.6.	Realisasi Investasi PMDN Triwulan I Tahun 2011 Berdasarkan Lokasi Proyek .....	34
2.7.	Diagram Pencar Hubungan Skor Tes Kecerdasan dan Hasil Produksi .....	36
3.1.	Histogram Hasil Belajar Metodologi .....	50
3.2.	Histogram Hasil Belajar Metodologi dengan Titik Tengah .....	50
3.3.	Poligon Hasil Belajar Metodologi .....	51
3.4.	Poligon Frekuensi Kumulatif “Kurang Dari” dan “Atau Lebih” .....	54
3.5.	Kurva Ogive Perolehan Belajar Metodologi Penelitian ...	55
4.1.	Kurva Simetris .....	83

4.2.	Kurva Condong Kiri .....	84
4.3.	Kurva Condong Kanan .....	84
5.1.	Distribusi Normal .....	115
6.1.	Bentuk Kecondongan Kurva .....	127
6.2.	Kurva Menceng Kiri .....	129
6.3.	Kurva Leptokurtis, Mesokurtis, dan Platikurtis .....	135
7.1.	Diagram Pencar Hubungan Hasil Produksi dan Skor Tes Kecerdasan .....	143
7.2.	Diagram Pencar dan Titik Koordinatnya dalam Keempat Kuadran .....	149
7.3.	Diagram Pencar dan Titik-Titik Koordinat X dan Y Mempunyai Hubungan Positif .....	149
7.4.	Diagram Pencar dan Titik-Titik Koordinat X dan Y Mempunyai Hubungan Negatif .....	150

## Bab 1.

# KONSEP STATISTIKA

### KOMPETENSI DASAR

Mendesripsikan tentang konsep statistik, kegunaan statistik dan sumber data statistik.

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- a) Menjelaskan tentang pengertian statistik dan statistika
- b) Menjelaskan tentang kegunaan statistik
- c) Menjelaskan tentang jenis-jenis statistik
- d) Menguraikan peranan statistik dalam kehidupan sehari-hari
- e) Menjelaskan tentang jenis-jenis data
- f) Menjelaskan tentang berbagai sumber data statistik

### DESKRIPSI SINGKAT

Bab ini menelusuri tentang konsep statistika hingga sumber data statistik. Mempelajari bab ini diharapkan bisa memahami konsep statistik dan statistika, sehingga tidak terjadi penyalahgunaan dua kata tersebut. Hal ini perlu dipelajari karena masih terdapat simpang siur penggunaan pengertian kedua istilah itu. Selain itu, bab ini juga mengupas kegunaan dan peranan statistik dalam kehidupan, serta pembahasan jenis-jenis statistik, jenis data serta berbagai sumber untuk mendapatkan data statistik.

### KATA KUNCI

Statistik	Nominal	Wawancara
Statistika	Ordinal	Observasi
Deskriptif	Interval	Dokumentasi
Interferensial	Rasio	Participation
Kualitatif	Kuantitatif	Non Participation

## A. PENGANTAR

Kita tentu pernah mendengar sebelumnya bahwa “*Statistik sulit,*” atau dalam pikiran kita sering terlintas pertanyaan “*Apa gunanya ilmu ini untuk saya?*”, “*Apa yang harus saya lakukan selanjutnya?*”, dan berbagai keluhan lainnya. Saat mempelajari pengantar statistik, mahasiswa juga sering mengeluh “*Aku tidak*

*memahaminya!*”. Berbagai komentar tersebut menggambarkan bahwa mata kuliah statistik sebagai mata kuliah sadistik. Bahkan bagi kebanyakan orang, statistik dianggap sebagai ilmu yang ruwet, penuh rumus-rumus yang rumit dan diperlukan ketelitian serta ketepatan dalam menghitungnya.

Namun seperti kita ketahui bersama, hampir setiap bidang baik pendidikan, perekonomian, perindustrian, perdagangan, perusahaan, bahkan setiap pimpinan perusahaan ataupun lembaga pernah menghadapi persoalan yang berhubungan dengan angka. Berdasarkan kumpulan angka-angka tersebut, ia berusaha membuat kesimpulan yang bisa mendeskripsikan atau memberi penjelasan mengenai persoalan itu. Salah satu usaha untuk mempermudah penyajian data tersebut adalah dengan menyusun angka ke dalam sebuah daftar atau tabel. Pandangan awam mengatakan proses penyajian itu adalah statistik. Hal ini memperlihatkan adanya keterkaitan antara kata statistik dengan kumpulan angka-angka yang menggambarkan sebuah data yang dapat menjadi sebuah informasi.

Sebelum mempelajari lebih jauh tentang statistik, ada baiknya mengenal sejarah statistik. Kesadaran manusia akan pentingnya kemampuan menghitung menjadi sebuah ide yang baik (seperti “Berapa banyak barang yang kamu perlukan untuk ditukar dengan salah satu dari barang lain”), dan kemampuan mengumpulkan informasi jadi keahlian yang sangat berguna.

Peristiwa serius terjadi pada abad ke-17 ketika kumpulan data pertama berkaitan dengan populasi dikumpulkan. Sejak saat itu, ilmuwan (kebanyakan ahli matematika, tapi kemudian ada juga ahli fisika dan biologi) merasa perlu untuk mengembangkan perangkat yang lebih spesifik untuk menjawab pertanyaan yang spesifik juga. Sebagai contoh, *Francis Galton* sepupu *Charles Darwin* (1822-1911), pertama kali menggunakan korelasi dalam ilmu penelitian ilmu hayat. Pada era tersebut penggunaan metode

statistik dalam penelitian biologi maupun sosial dianggap sebagai sesuatu yang tidak lazim. Bahkan akhir abad ke-19, kritikan-kritikan keras kerap kali dilontarkan terhadap *Karl Pearson* yang memelopori penggunaan statistik dalam pelbagai penelitian biologi maupun pemecahan persoalan bersifat sosio ekonomis.

Selama 100 tahun terakhir, tiada seorang pun peneliti yang menyangkal betapa pentingnya ilmu statistik bagi penelitian ilmiah. Tanpa metode statistik, peneliti seakan-akan seperti orang buta meraba-raba dalam kegelapan untuk menangkap seekor kucing hitam. Statistik modern seperti kita kenal merupakan produk abad ke-20, berawal ketika R.Fisher memperkenalkan analisis varians ke dalam literatur statistik. Statistik merupakan bidang pengetahuan yang mengalami pertumbuhan pesat dimana metodenya berkembang sejajar dengan penemuan penting lainnya.

Seiring kemajuan pesat bidang komputer, muncul berbagai program yang dibuat khusus untuk membantu pengolahan data statistik. Pengolahan data statistik menjadi jauh lebih mudah tanpa mengurangi ketepatan hasil outputnya. Namun satu hal yang terpenting adalah proses untuk mendapatkan data, proses tersebut seyogyanya juga dimulai dengan proses pengumpulan data yang terpercaya dan dapat diandalkan, karena dengan demikian setelah data dianalisa akan mendapatkan hasil yang bisa dipercaya dan dapat diandalkan.

Berkenaan dengan hal tersebut, bahan ajar ini mengajak untuk mendalami tentang konsep statistik dan statistika serta penerapannya pada konteks statistik deskriptif, selain penerapan alat bantu statistik berupa program SPSS.

## **B. PENGERTIAN STATISTIK DAN STATISTIKA**

Statistik (*statistic*) berasal dari kata *state* (Negara). Mengapa disebut negara? Karena pada mulanya, kata statistik diartikan sebagai keterangan-keterangan yang dibutuhkan oleh negara dan



berguna bagi negara. Kepentingan negara itu meliputi berbagai bidang kehidupan dan penghidupan, sehingga lahir istilah statistik, yang penggunaannya bisa disesuaikan dengan lingkup datanya. Contohnya, dalam kehidupan sehari-hari kita kerap kali membaca berapa rata-rata *annual income* penduduk Indonesia? Data terakhir menunjukkan bahwa pada tahun 2010, rata-rata pendapatan setiap penduduk Indonesia (per kapita/per kepala) sudah berada pada angka USD 3000 per tahun. Padahal ada kalanya data yang dikumpulkan di lapangan tidak disajikan dalam bentuk rata-rata seperti tadi, melainkan disajikan dalam bentuk tabel atau diagram dengan uraian yang lebih rinci dan di bagian atas atau bawah tabel.

Berdasar uraian tersebut ditarik sebuah pengertian bahwa Statistik adalah kumpulan data, bilangan maupun non bilangan yang disusun dalam tabel dan atau diagram yang melukiskan suatu persoalan.

Sekarang mari kita tinjau apa yang dimaksudkan dengan statistika. Sebelum mengemukakan, menjelaskan, menyimpulkan atau memberikan pernyataan tentang sesuatu persoalan, terlebih dahulu kita perlu melakukan penelitian mengenai persoalan yang akan disimpulkan, agar pernyataan yang dibuat cukup beralasan dan dapat dipertanggungjawabkan. Sebelum penelitian dilakukan perlu diketahui batas-batasnya, hingga ruang lingkup yang akan disimpulkan, kemudian dilanjutkan dengan pengumpulan bahan-bahan untuk dianalisis, berdasarkan analisa tersebut baru kesimpulan tentang persoalan tersebut dibuat. Penjelasan diatas merupakan sebuah langkah-langkah dasar dalam statistika.

Berdasarkan uraian tersebut didefinisikan bahwa Statistika merupakan pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau analisisnya serta penarikan kesimpulan berdasarkan kumpulan data dan penganalisaan yang dilakukan. Makna lain dari Statistika yaitu suatu ilmu yang

mempelajari tentang cara-cara mengumpulkan, menganalisis dan menginterpretasikan data. Dengan kata lain, statistika menjadi semacam alat dalam melakukan riset empiris. Para ilmuwan menggambarkan persepsinya tentang suatu fenomena dalam menganalisis data. Deskripsi yang sudah stabil tentang suatu fenomena seringkali mampu menjelaskan suatu teori (walaupun demikian, orang dapat saja berargumentasi bahwa ilmu biasanya menggambarkan *bagaimana* sesuatu terjadi, bukannya mengapa). Penemuan teori baru merupakan proses kreatif yang didapat dengan cara mereka ulang informasi pada teori yang telah ada atau mengeskrak informasi yang diperoleh dari dunia nyata. Pendekatan awal yang umumnya digunakan untuk menjelaskan.

### **C. PERANAN DAN KEGUNAAN STATISTIK**

Perkembangan statistik sebagai metode ilmiah memberikan pengaruh yang besar terhadap setiap aspek kehidupan manusia modern. Metode statistik dibutuhkan sebagai peralatan analisa dan interpretasi data kuantitatif.

#### **1. Peranan di Bidang Penelitian**

##### **a. Bagi Calon Peneliti dan Para Peneliti**

Dalam kehidupan dan penghidupan sehari-hari di tengah ledakan data, kita tidak dapat melepaskan diri dari data, baik data itu bersifat kuantitatif maupun kualitatif. Kedua sifat data tersebut dapat dianalisis secara kuantitatif maupun kualitatif atau gabungan dari keduanya. Dalam menghadapi data yang berserakan itu, aliran kuantitatif yang berakar dari paham positivism memandang bahwa data dan kebenaran itu sudah ada di sekitar kita. Kita ditantang untuk mengumpulkannya melalui teknik pengumpulan data baik melalui pengamatan, wawancara, angket maupun dokumentasi secara objektif. Setelah data itu terkumpul, maka dilanjutkan dengan mengolah data tersebut

dalam bentuk penyajian data. Bentuk mana yang dipilih, hal ini tergantung kebutuhannya masing-masing. Dalam hal ini statistik deskriptif diperlukan karena peneliti akan dapat mendeskripsikan data yang dikumpulkan. Perkembangan selanjutnya, peneliti ingin membedakan data berdasarkan rata-rata kelompoknya atau ingin menghubungkan data yang satu dengan yang lainnya atau ingin meramalkan pengaruh data yang satu dengan yang lainnya sehingga akhirnya peneliti dapat menarik suatu kesimpulan dari data yang telah dianalisisnya. Dalam hal ini teknik statistik inferensial sangatlah diperlukan. Jadi, statistik berperan sebagai alat untuk deskripsi, komparasi, korelasi dan regresi.

b. Bagi Pembaca

Sebagai ilmuwan produktif tentunya kita selalu disibukkan oleh kegiatan membaca khususnya membaca laporan-laporan penelitian, laporan-laporan keadaan kantor/perusahaan, nota keuangan, laju inflasi, GNP, dan lain sebagainya. Masalahnya ialah, “bagaimana pembaca bisa memahami informasi tersebut dengan benar kalau tidak mengerti statistik?” Akibatnya ialah komunikasi antara penulis dengan pembaca tidak efektif. Lebih berbahaya lagi jika pembaca yang buta statistik tadi berani menerapkannya untuk mengambil keputusan.

c. Bagi Pembimbing Penelitian

Peneliti maupun pembimbing yang bijaksana mempunyai pandangan yang luas dalam mencari kebenaran. Peneliti dan pembimbing janganlah terlalu picik menganggap bahwa hanya metode itulah satu-satunya alat yang dipakai mencari kebenaran, karena tidak semua metode kualitatif dapat menyelesaikan permasalahan. Demikian pula, tidak semua metode kuantitatif dapat menyelesaikan semua permasalahan.

Peneliti maupun pembimbing yang terlalu membela bahwa metode kualitatif yang paling benar atau sebaliknya sambil

menjelekkkan metode lain, secara tidak langsung menunjukkan kedangkalan atau ketidaktahuannya terhadap metode lainnya. Sebab belum tentu kita lebih baik dari orang yang dijelekkkan. Apakah kita sudah menguasai metode kualitatif secara penuh, sehingga berani menjelekkkan metode kuantitatif atau sebaliknya.

Seringkali timbul cemooh dari peneliti kuantitatif yang mengatakan bahwa peneliti kualitatif tidak berani menggunakan kuantitatif, karena statistiknya lemah atau tidak paham statistik. Sebaliknya, peneliti kualitatif mencemoohkan peneliti kuantitatif dengan mengatakan bahwa peneliti kuantitatif itu sangat dangkal dan hanya bekerja dengan angka-angka tanpa menyelami makna kualitatif yang ada di balik angka, dan penelitian kuantitatif hanya menguji hipotesis saja sehingga tidak menghasilkan teori-teori baru bagi perkembangan ilmu.

Cemoohan itu tidak perlu membuat peneliti, pembimbing atau penguji terbawa arus pembelaan ekstrem yang hanya membenarkan salah satu metode saja. Sebagai peneliti dan pembimbing yang kritis, kita harus mampu menempatkan kedua metode penelitian pada fungsinya masing-masing. Jika mungkin kedua metode itu dapat saling mengisi. Metode mana yang dipakai dalam penelitian? Jawabnya ialah tergantung dari masalah apa yang akan diteliti. Sebagai contoh, jika masalah yang ingin diteliti adalah sejauhmana peredaran keuangan, maka mungkin metode kuantitatiflah yang paling cocok dipakai. Namun jika kita ingin meneliti masalah proses dan sistem nilai budaya masyarakat, maka metode kualitatiflah yang paling cocok dipakai. Ada kalanya kedua metode itu digunakan secara bersamaan, misalnya untuk bisa mengerti data statistik secara mendalam dibutuhkan metode kualitatif terlebih dahulu untuk memberikan kedalaman aspek terhadap butir tes dalam menyusun angket.

Penelitian kualitatif sebaiknya diikuti penelitian kuantitatif, sehingga dapat memberikan kenyataan yang lebih akurat dan

berguna dalam kegiatan prediksi dan control. Sebagai contoh, kita telah meneliti secara kualitatif tentang adanya pengaruh informasi langsung para petugas dan informasi tidak langsung melalui media massa terhadap modernisasi masyarakat. Jika kita dihadapkan pada pilihan, “*Mana yang harus kita dahulukan untuk mempercepat proses modernisasi itu?*” , maka kita perlu mengadakan penelitian kuantitatif dengan variabel yang tepat.

d. Bagi Penguji Skripsi, Thesis atau Disertasi

Penguji yang menguji skripsi/tesis/disertasi mahasiswa yang menggunakan metode kuantitatif selayaknya memahami statistik sehingga dapat meningkatkan kualitas lulusannya sebagai penguji. Jangan sampai penguji buta statistik tetapi nekat menguji mahasiswa dengan mengajukan sanggahan bahwa korelasinya 0,90 artinya sangat kecil dan mohon dibetulkan. Sementara mahasiswa lain yang turut mendengarkan dapat menilai betapa bodohnya penguji itu, atau karena lemah statistiknya sehingga tidak berani menguji analisis statistiknya.

## **2. Peranan di Bidang Ekonomi dan Manajemen**

Pada bidang ekonomi, metode statistik merupakan alat yang penting dalam proses pengambilan keputusan bagi masing-masing bagian atau bidang diantaranya adalah:

a. Pimpinan (Manajer) dan Administrator, statistik sebagai alat:

- Pengumpulan data baik secara sensus maupun sampling.
- Pengolahan atau analisis data.
- Penyajian data dalam bentuk laporan manajemen.
- Pengambilan keputusan atau perencanaan, dan
- Evaluasi atau pengawasan antara data di lapangan dengan penyimpangan di lapangan.
- Melakukan pemecahan masalah manajerial .

b. Bagian Produksi

Penggunaan statistik dalam proses produksi berkaitan dengan penetapan standar dan pengawasan kualitas, pengawasan terhadap efisiensi kerja dan tes terhadap produk baru

c. Bidang Akuntansi

Sebagian besar penggunaan statistik di akuntansi bertalian dengan penilaian aktiva perusahaan, penyesuaian perubahan harga dan hubungan biaya dan volume produksi.

d. Bidang Pemasaran

Penggunaan statistik dalam bidang pemasaran berhubungan dengan analisa penjualan, pasar dan pemasaran. Analisa statistik pada ketiga hal tersebut berkenaan langsung dengan penyelidikan tentang referensi konsumen, penaksiran potensi pasaran bagi produk baru, penetapan harga dan penelitian mengenai potensi pasar di daerah baru.

### **3. Peranan Bagi Ilmu Pengetahuan**

Statistika sebagai disiplin ilmu berguna untuk kemajuan ilmu dan teknologi. Karena itu, kita dituntut untuk memahami statistii lebih mendalam. Jika tidak, kita semakin ketinggalan dari perkembangan ilmu dan teknologi dengan negara lainnya. Statistika dapat dijadikan sebagai alat:

- *Deskripsi* yaitumenggambarkan data pengukuran dampak dan proses pembangunan melalui indikator ekonomi, indeks harga, inflasi, GNP, laporan nota keuangan dan sebagainya.
- *Komparasi* yaitu membandingkan data pada dua kelompok atau beberapa kelompok.
- *Korelasi* yaitu mencari besarnya hubungan data dalam suatu penelitian.
- *Regresi* yaitu meramalkan pengaruh data yang satu terhadap data lainnya, atau estimasi terhadap kecenderungan peristiwa yang akan terjadi di masa depan.

- *Komunikasi* yaitu merupakan alat penghubung antar pihak berupa laporan data statistik atau analisis statistik sehingga kita maupun pihak lainnya dapat memanfaatkannya dalam membuat suatu keputusan.

## **D. JENIS-JENIS STATISTIKA**

Berdasarkan definisinya, Statistika meliputi pengumpulan data, pengorganisasian data, penyajian data, analisa data dan interpretasi dari hasil analisis tersebut. Berdasarkan pada definisi tersebut, statistika dibagi menjadi 2 (dua) jenis yaitu statistika deskriptif dan statistika induktif.

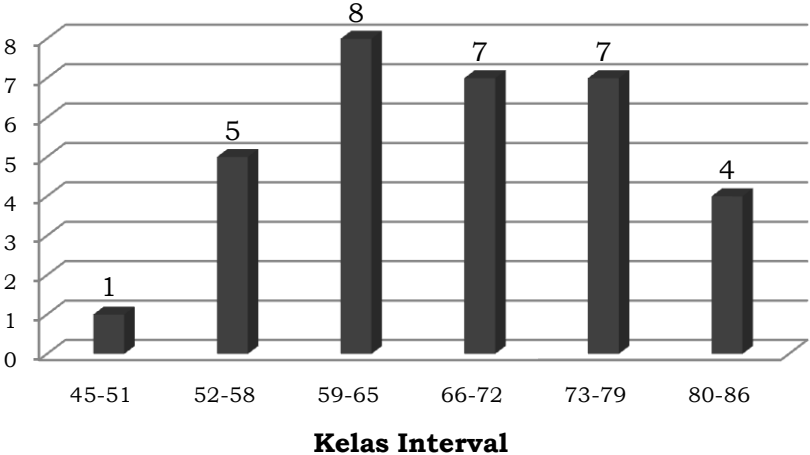
### **1. Statistika Deskriptif**

*Statistika deskriptif* adalah teknik yang digunakan untuk mensarikan data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap orang. Hal tersebut melibatkan proses kuantifikasi dari penemuan suatu fenomena. Berbagai statistik sederhana, seperti rata-rata, dihitung dan ditampilkan dalam bentuk tabel dan grafik. Statistika deskriptif dapat memberikan pengetahuan yang signifikan pada kejadian fenomena yang belum dikenal dan mendeteksi keterkaitan yang ada di dalamnya. Tetapi dapatkah statistika deskriptif memberikan hasil yang diterima secara ilmiah? Statistik merupakan suatu alat pengukuran yang berhubungan dengan keragaman pada karakteristik objek-objek yang berbeda .

*Statistik deskriptif* ialah susunan angka yang memberikan gambaran tentang data dalam bentuk-bentuk tabel, diagram, histogram, polygon frekuensi, ozaiv (ogive), ukuran penempatan (median, kuartil, desil dan persentil), ukuran gejala pusat (rata-rata hitung, rata-rata ukur, rata-rata harmonik, dan modus), simpangan baku, angka baku, kurva normal, korelasi, dan regresi linier. Berikut beberapa contoh dari statistika deskriptif.

Statistik deskriptif berbentuk uraian, misalnya diuraikan sebagai berikut: data hasil penelitian terhadap hasil belajar yang dilakukan kepada 32 orang responden. Berdasarkan hasil tes didapatkan rentang nilai minimum sebesar 45 dan maksimum sebesar 82 dengan rata-rata 74.

Statistik deskriptif disajikan dalam bentuk histogram/ diagram seperti gambar berikut:



Gambar 1.1. Histogram Hasil Belajar Siswa Kelas Kontrol (Model Pembelajaran Konvensional)

**2. Statistika Inferensial**

Statistika dalam arti luas adalah salah satu alat untuk mengumpulkan dan mengolah data, menarik kesimpulan serta membuat keputusan berdasarkan analisis data. Statistik dalam arti luas meliputi penyajian data, yang berarti meliputi statistik dalam arti sempit tadi. Statistik dalam arti luas ini disebut juga dengan istilah statistika inferensial atau statistik induktif.

Selain istilah-istilah di atas, ada istilah statistika matematis dan statistika praktis. Statistika matematis ialah ilmu yang mempelajari asal-usul atau penurunan sifat-sifat, dalil-dalil, rumus-rumus, serta dapat diwujudkan ke dalam model-model



lain bersifat teoritis, sedangkan statistika praktis ialah penerapan statistika matematis ke dalam berbagai bidang ilmu lainnya sehingga lahirlah istilah statistika kedokteran, statistika sosial, dan sebagainya. Bagi mereka yang ingin mendalami statistika praktis secara mendalam sebaiknya memperkuat dasar-dasar statistika matematis terlebih dahulu. Statistik inferensial dibagi menjadi kelompok statistika parametrik dan nonparametrik.

Statistik Parametrik digunakan apabila datanya memenuhi persyaratan: (1) interval, (2) normal, (3) homogen, (4) dipilih acak, dan (5) linier. Contoh-contoh analisis statistik parametrik adalah: pengujian hipotesis, regresi (untuk menyimpulkan), korelasi (untuk menyimpulkan), uji t, dan anova.

Statistik nonparametrik dipakai apabila data kurang dari 30, atau tidak normal (tidak linier), contohnya: tes binomial, tes chi-kuadrat, Kruskal-Wallis, Friedman, tes Kolmogorov-Smirnov, tes run, tes McNemar, tes tanda, tes Wilcoxon, tes Walsh, tes Fisher, tes median, tes U Mann-Whitney, tes run Wald-Wolfowitz, tes reaksi ekstrem Moses, tes Q Cochran, koefisien kontingensi, koefisien rank dari Spearman Brown, koefisien rank dari Kendall, dan uji normalitas dari Lillieford.

## **E. DATA**

Menurut Webster's New World Dictionary, data berarti sesuatu yang diketahui atau dianggap. Dengan demikian, data memberikan gambaran tentang suatu keadaan atau persoalan. Data tentang sesuatu pada umumnya dikaitkan dengan tempat dan waktu. Misalnya harga beras bermutu sedang di pasar besar Malang, pada 30 September 2012 adalah Rp.8000,- per kg. Penyebutan tempat dan waktu sangat penting, sebab harga beras akan berubah sesuai waktu dan tempatnya. Untuk memperoleh gambaran keadaan sosial dan ekonomi, pemerintah harus mengumpulkan data kegiatan ekonomi (produksi, perdagangan,

konsumsi, pendapatan, harga, dan lain-lain) dan kegiatan sosial (pendidikan, kesehatan, kebudayaan). Badan Pusat Statistik (BPS) mengeluarkan publikasi indikator sosial dan indikator ekonomi yang dapat memberikan gambaran tentang keadaan sosial dan ekonomi masyarakat.

Data berbentuk jamak, sedang datum berbentuk tunggal. Jadi data sama dengan datum-datum. Data ialah suatu bahan mentah yang jika diolah dengan baik melalui berbagai analisis dapat melahirkan berbagai informasi.

**Contoh 1.1.** Nilai ulangan susulan statistik dari 6 mahasiswa pendidikan ekonomi adalah 70, 50, 65, 80, 90, 67.

Dalam statistik dikenal istilah-istilah jenis data, tingkatan data, sumber data, penyajian data, dan analisis data. Data dianalisis sesuai dengan jenis dan tingkatannya, karena itu masing-masing tingkatan data mempunyai analisis sendiri khususnya dalam analisis regresi.

Data yang baik tentunya harus yang mutakhir, cocok (*relevant*) dengan masalah penelitian dari sumber yang dapat dipertanggungjawabkan, lengkap, akurat, objektif, dan konsisten. Pengumpulan data sedapat mungkin diperoleh dari tangan pertama. Data yang baik sangat diperlukan dalam penelitian, sebab bagaimanapun canggihnya suatu analisis data jika tidak ditunjang oleh data yang baik, maka hasilnya kurang dapat dipertanggungjawabkan.

## **1. Jenis Data**

Data statistik adalah keterangan atau ilustrasi mengenai sesuatu hal yang bisa berbentuk kategori (misalnya rusak, baik, cerah, berhasil) atau bilangan. Selanjutnya data yang berupa kategori disebut sebagai data kualitatif dan data bilangan disebut data kuantitatif.

Data kualitatif adalah data non-angka (numerik) seperti jenis kelamin, warna kesayangan, dan asal suku. Data kualitatif digunakan apabila kita tertarik melihat proporsi atau bagian yang termasuk dalam kategori. Contohnya berapa persen jenis kelamin pria dibandingkan wanita, warna apa yang disukai oleh sebagian besar penduduk, dan berapa persen suku tertentu dibandingkan dengan suku lain.

Data kuantitatif adalah data angka atau numerik seperti jumlah mobil, jumlah TV yang dijual disuatu toko, berat badan, jarak Solo-Jakarta, dan sebagainya. Semua ukuran itu berupa angka. Berdasarkan cara perolehan, data kuantitatif dibedakan menjadi data dikotomi (data diskrit) dan data kontinum.

Data dikotomi disebut: data deskrit, data kategori atau data nominal. Data-data tersebut yang diperoleh dari hasil menghitung atau membilang dan menghasilkan data kuantitatif yang nilainya khusus dan biasanya berupa bilangan bulat. Data dikotomi adalah data yang paling sederhana yang disusun menurut jenisnya atau kategorinya. Bila kita telah memberi nama kepada sesuatu berarti kita telah menentukan jenis atau kategorinya menurut pengukuran kita. Dalam dikotomi, data dikelompokkan menurut kategorinya dan diberi angka. Angka-angka tersebut hanyalah label belaka, bukan menunjukkan tingkatan (*ranking*). Dasar dalam menyusun kategori data tidak boleh tumpang tindih (*mutually exclusive*). Kalau melakukan kategori secara alamiah, maka disebut data dikotomi sebenarnya (*true dichotomi*) dan jika kategorinya dibuat-buat sendiri (direkayasa), maka disebut data dikotomi dibuat-buat (*artificial dichotomi*).

Contoh dari data dikotomi sebenarnya antara lain adalah: jenis kelamin, umpamanya ada tiga yaitu laki-laki diberi angka 1, banci diberi angka 2 dan perempuan diberi angka 3. Angka 3 pada wanita bukan berarti kekuatan pada wanita sama dengan tiga kali laki-laki. Tetapi seperti yang disebutkan tadi bahwa

angka-angka tersebut hanyalah label belaka. Banyak contoh-contoh data dikotomi sebenarnya seperti macam warna kulit, suku bangsa, bahasa daerah dan sebagainya.

Data dikotomi dibuat apabila data itu belum mempunyai kategori mutlak atau alamiah seperti di atas tadi, oleh sebab itu data tersebut masih bisa diubah-ubah jika memang dikehendaki. Sebagai contoh: tidak lulus diberi angka 1 dan lulus diberi angka 2. Tetapi jika yang tidak lulus ingin diubah menjadi lulus maka bisa saja diadakan ujian ulangan. Data dikotomi ini mempunyai sifat-sifat: eksklusif, tidak mempunyai urutan (*ranking*), tidak mempunyai ukuran baru, dan tidak mempunyai nol mutlak.

Data Kontinyang diperoleh dari hasil pengukuran termasuk dalam data kontinu. Data kontinu merupakan data kuantitatif yang nilainya menempati semua interval pengukuran dan merupakan hasil pengukuran serta bisa berupa bilangan pecahan dan bulat. Contohnya, berat badan bisa 60,1 kg. tinggi badan, luas rumah, panjang jalan dan lain-lain, yang adalah hasil pengukuran digolongkan sebagai data kontinu. Data kontinum terdiri atas tiga macam data yaitu: data ordinal, data interval dan data rasio. Ketiga macam data tersebut diuraikan sebagai berikut:

#### a. Data Ordinal

Data ordinal ialah data yang sudah diurutkan dari jenjang yang paling rendah sampai ke yang paling tinggi, atau sebaliknya tergantung peringkat selera pengukuran yang subjektif terhadap objek tertentu. Data ordinal disebut juga sebagai data berurutan, data berjenjang, data berpangkat, data tata jenjang, data ranks, dan data petala, data bertangga atau data bertingkat.

Pemberian jenjang itu umumnya dilakukan dengan urutkan data mulai terendah sampai tertinggi atau sebaliknya. Kemudian beri angka 1 dan seterusnya mulai tertinggi hingga terendah. Data ordinal bersifat eksklusif, tidak mempunyai ukuran baru, mempunyai urutan, dan tidak mempunyai nilai nol mutlak.

b. Data Interval

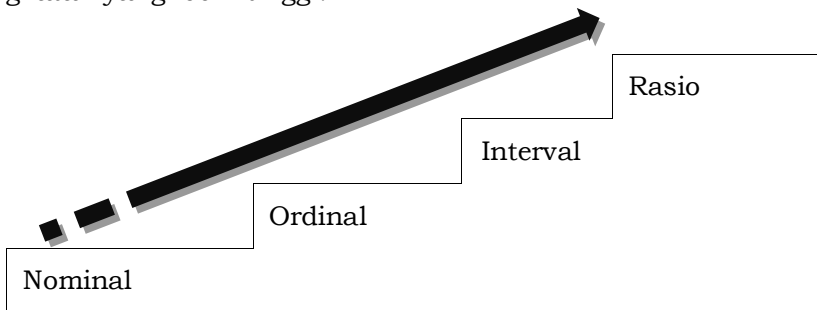
Data interval mempunyai sifat nominal dari data ordinal dan mempunyai nol mutlak, sehingga mempunyai skala interval yang sama jaraknya. Pengukuran data interval tidak memberikan jumlah absolut dari objek yang diukur, contohnya persepsi, tanggapan, dan sebagainya. Dalam penelitian sosial, data interval banyak digunakan. Data interval bersifat eksklusif, mempunyai urutan, ukuran baru, tetapi tidak mempunyai nilai nol mutlak.

c. Data Rasio

Data rasio mengandung sifat-sifat interval dan mempunyai nilai nol mutlak. Contoh data rasio adalah: berat badan, tinggi, panjang atau jarak. Data rasio ini sering dipakai dalam penelitian keilmualan atau enjinereng. Karena data rasio, ordinal dan interval merupakan hasil pengukuran, maka pada ketiga data tersebut ditemui adanya bilangan pecahan. Data rasio bersifat eksklusif, mempunyai urutan, ukuran baru dan nol mutlak.

**2. Tingkatan Data**

Tingkatan data kalau diurutkan dari yang tertinggi ke yang terendah yaitu: 1) rasio, 2) interval, 3) ordinal, dan 4) nominal. Dalam analisis statistik, data yang tinggi dapat diturunkan ke tingkatan yang lebih rendah jika dianggap diperlukan. Namun sebaliknya, data tingkatan rendah tidak dapat dinaikkan kepada tingkatan yang lebih tinggi.



Gambar 1.2. Tingkatan Data

## **F. SUMBER DATA STATISTIK**

Data dapat langsung dikumpulkan peneliti melalui pihak yang disebut sumber primer. Data dikumpulkan peneliti melalui pihak kedua atau tangan kedua disebut pihak sekunder, yaitu data yang diperoleh melalui wawancara kepada pihak lain tentang objek dan subjek yang diteliti, dan mempelajari dokumentasi-dokumentasi tentang objek dan subjek yang diteliti. Dari kedua sumber data tersebut, tentu saja sumber data primer lebih dapat dipertanggungjawabkan dibandingkan data yang didapat melalui sumber sekunder. Pengumpulan data bisa dilakukan melalui teknik wawancara (*interview*), pengamatan (*observation*), angket (*questionnary*) dan dokumentasi (*documentation*). Wawancara bisa dilakukan tidak sistematis maupun sistematis. Pengamatan bisa dilakukan secara tidak langsung (*nonparticipation*) atau langsung (*participation*). Angket dapat dibuat tertutup atau terbuka. Peneliti dapat menggunakan salah satu atau gabungan dari teknik-teknik pengumpulan data tersebut. Masing-masing teknik mempunyai keuntungan dan kerugian masing-masing.

## **G. RINGKASAN**

- 1) Statistik adalah kumpulan data, bilangan maupun non bilangan yang disusun dalam tabel dan atau diagram yang melukiskan suatu persoalan.
- 2) Statistika adalah pengetahuan yang berhubungan dengan cara pengumpulan data, pengolahan atau analisa dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisaan yang dilakukan
- 3) Bagi Calon Peneliti dan Para Peneliti statistik , statistik sangat diperlukan karena peneliti akan dapat mendeskripsikan data yang dikumpulkan, membedakan data berdasarkan rata-rata kelompok atau menghubungkan data satu dengan lainnya dan meramalkan pengaruh data yang satu dengan yang lainnya.

- 4) *Statistika Deskriptif* adalah metode statistika yang digunakan untuk menggambarkan atau mendeskripsikan data yang telah dikumpulkan menjadi sebuah informasi.
- 5) *Statistika Inferensial (Induktif)* adalah metode yang digunakan untuk mengetahui sebuah populasi berdasarkan suatu sampel atau contoh dengan menganalisis dan menginterpretasikan data menjadi sebuah kesimpulan.
- 6) Data ialah suatu bahan mentah yang jika diolah dengan baik melalui berbagai analisis dapat melahirkan berbagai informasi.

## **H. SOAL DAN LATIHAN**

- 1) Jelaskan pentingnya statistika dalam kehidupan sehari-hari, dan siapa saja yang sering menggunakan statistika?
- 2) Jelaskan perbedaan statistik dan statistika !
- 3) Jelaskan perbedaan antara statistika deskriptif dan statistika inferensial ! Berikan contoh dari kasus sehari-hari yang Anda temui!
- 4) Jelaskan jenis data menurut sumbernya, cara memperoleh dan waktu pengumpulannya !
- 5) Jelaskan apa yang dimaksud skala Nominal (N), Ordinal (O), Interval (I) dan Rasio (R) !

## Bab 2.

# PENYAJIAN DATA

### **KOMPETENSI DASAR**

Mendesripsikan cara menyajikan data dalam bentuk tabel dan diagram.

### **TUJUAN PEMBELAJARAN**

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- a) Menjelaskan tentang struktur tabel statistik.
- b) Menjelaskan tentang jenis-jenis tabel yang digunakan untuk menyajikan data.
- c) Menjelaskan tentang jenis-jenis diagram yang digunakan untuk menyajikan data.
- d) Menjelaskan tentang cara penyajian data dalam bentuk tabel maupun diagram.

### **DESKRIPSI SINGKAT**

Setelah memahami konsep statistik dan peranannya dalam kehidupan sehari-hari yang disajikan pada bab sebelumnya, maka bab ini mengantarkan kita untuk menguasai bagaimana membaca dan menyajikan data dalam bentuk tabel dan diagram atau grafik. Sebagai pendukung pemahaman terhadap penyajian data tersebut, pada bagian ini akan dibahas mengenai penyajian data dalam bentuk tabel termasuk struktur tabel dan macam-macam tabel, penyajian data dalam bentuk gambar yang terdiri atas diagram batang, histogram, diagram lingkaran, diagram lambang, diagram garis dan diagram peta.

### **KATA KUNCI**

Tabel Referensi (Reference Table)

Tabel Ikhtisar (Summary Table)

Tabel Umum (General Table)

Diagram Batang

Histogram

Diagram Lingkaran

Diagram Lambang

Diagram Garis

Diagram Peta

## **A. PENGANTAR**

Data statistik tidak hanya cukup dikumpulkan dan diolah, tetapi juga perlu disajikan dalam bentuk yang mudah dibaca dan dimengerti, namun seringkali kita menghadapi masalah dalam



menyajikan sejumlah data statistik yang besar dalam bentuk ringkas dan kompak. Sejumlah informasi dapat diperoleh kembali dengan mudah dan sederhana jika data tersebut diringkas dalam bentuk tabel, grafik dan juga diagram.

Tabel statistik umumnya dapat dibedakan ke dalam *tabel referensi* (*reference table*) dan *tabel ikhtisar* (*summary table*). Tabel referensi memiliki fungsi sebagai “gudang keterangan” karena tabel itu memberi keterangan-keterangan yang terperinci dan disusun khusus guna kepentingan referensi. Tabel-tabel yang terdapat dalam laporan sensus umumnya merupakan tabel yang memberi keterangan-keterangan secara umum bagi kepentingan referensi. Fungsi tabel bersifat umum karena angka-angkanya bisa dipergunakan dalam bermacam-macam cara. Acapkali, tabel sedemikian itu dinamakan *tabel umum* (*general table*).

Namun adakalanya pimpinan perusahaan yang sangat sibuk dan kurang waktu guna membaca tabel statistik umumnya lebih suka mempelajari peristiwa-peristiwa perusahaannya dari grafik statistik. Sebenarnya, penyajian data statistik secara grafis memiliki bermacam-macam kegunaan. Grafik/diagram seringkali dipergunakan dalam iklan dengan maksud agar konsumen memperoleh kesan yang mendalam terhadap ciri-ciri produk yang diiklankan. Kegiatan-kegiatan produksi lebih mudah dilihat dan dinilai secara visual bila kegiatan-kegiatan tersebut dinyatakan dalam angka dan digambarkan secara grafis.

Peta pengawasan kualitas merupakan alat penting dalam melaksanakan pengawasan produk maupun pengawasan proses produksi. Grafik penjualan perusahaan memberi gambaran yang sederhana dan menarik mengenai perkembangan hasil penjualan yang telah dicapai oleh perusahaan yang bersangkutan. Pada hakekatnya, grafik dan tabel seyogyanya dipergunakan secara bersama-sama. Grafik statistik memang lebih mudah menarik perhatian pembacanya daripada tabel statistik. selain daripada

itu, grafik sedemikian itu dapat melukiskan suatu peristiwa secara lebih mengesankan dan tidak menjemukan. Meskipun demikian, penyajian secara grafis sebetulnya hanya bersifat aproksimatif.

Angka-angka yang eksak dan terperinci tentang suatu peristiwa hanya mungkin diperoleh dari tabel statistik. Karena itu, analisa dan interpretasi data kuantitatif umumnya dilakukan dari angka-angka yang terdapat dalam tabel statistik. Bagi seorang guru dengan melihat kecenderungan peningkatan prestasi belajar siswanya dari grafik perkembangan prestasi siswa, dari grafik tersebut bisa menganalisa kecenderungan mata pelajaran apa yang mengalami peningkatan pesat atau sebaliknya. Namun disamping melihat dari tampilan grafik perlu lebih dalam untuk melihat gambaran dalam bentuk tabel, sehingga bisa mengamati langsung data eksak yang diperoleh oleh siswa.

## **B. PENYAJIAN DATA DALAM BENTUK TABEL**

Data yang telah terkumpul dari sebuah penelitian tentu akan dilakukan sebuah langkah lanjutan yakni telaah atas data tersebut, adapun untuk keperluan analisa maka data itu perlu disusun secara teratur. Biasanya data itu disusun dalam tabel.

Tabel 2.1. IPM di Beberapa Negara

Negara	HDI	Indeks Harapan Hidup	Indeks Pendidikan	Indeks PDB	Indeks Kemiskinan	Peringkat
Singapura	0.922	0.907	0.908	0.950	5.2	25
Malaysia	0.811	0.811	0.839	0.783	8.3	63
Thailand	0.781	0.743	0.855	0.745	10.0	78
China	0.777	0.792	0.837	0.703	11.7	81
Filipina	0.771	0.767	0.888	0.657	15.3	90
Vietnam	0.733	0.812	0.815	0.572	15.2	105
Indonesia	0.728	0.745	0.830	0.609	18.2	107
India	0.619	0.645	0.620	0.591	31.3	128
Laos	0.601	0.637	0.663	0.503	34.5	130
Kamboja	0.598	0.550	0.691	0.552	38.6	131
Myanmar	0.583	0.596	0.764	0.389	21.5	132

Sumber:UNDP Human Development Report 2007-2008

Para pembaca akan lebih mudah memahami dan menelaah penyajian data berbentuk daftar (Tabel 2.1). Dari daftar seperti ini pula umumnya kita telah dapat menjawab beberapa pertanyaan yang dikehendaki atau membuat uraian yang diperlukan. Jika diperhatikan di kiri bawah dari daftar itu terdapat perkataan, “sumber:...”, yang hendaknya disertakan setiap kali membuat daftar. Catatan ini menuliskan dari mana data itu diperoleh.

### 1. Struktur Tabel Statistik

Sebuah tabel yang formal umumnya terdiri dari beberapa bagian seperti yang terlihat pada Tabel 2.2. Tabel statistik yang baik dan efisien harus bersifat sederhana dan jelas. *Nama* (titel), *nama kolom* dan *nama kompartimen* harus diusahakan agar jelas dan singkat.

Tabel 2.2. Skema Catatan Pendahuluan

Nama Kompartimen	Nama Kolom			
	Nama	Kolom	Nama	Kolom
Kompartimen				

Pos-pos keterangan dan angka-angka dalam kolom harus disusun secara cermat serta dan mengena. Usaha sedemikian itu akan berhasil bila beberapa syarat-syarat umum dipenuhi. Saya akan memberikan uraian sevara sigkat mengenai syarat-syarat sedemikian itu dan bagaimana cara pemenuhannya.

#### a. Nama (titel) dan identifikasi

Tiap tabel yang baik harus memiliki nama (titel) dan nama kolom, yang harus ditaruh di atas tabel. Nama harus jelas dan singkat. Bila keduanya tidak dapat dipenuhi, kesederhanaan harus lebih diutamakan sedangkan keterangan-keterangan lebih lanjut dapat diberikan dalam catatan di bawah tabel.

Umumnya, susunan redaksi nama harus dapat memberi keterangan tentang ciri-ciri data yang terdapat dalam tabel. Nama harus memberi keterangan tentang ciri-ciri data yang terpenting, baru kemudian boleh melengkapi keterangan tersebut dengan sub keterangan. Sub-keterangan merupakan keterangan yang kurang penting jika dibandingkan keterangan yang terdapat dalam nama. Jika redaksi nama disusun dalam beberapa baris, maka susunannya harus merupakan bentuk piramid yang terbalik atau dalam baris-baris yang rapi dan teratur.

Akhirnya, bila dalam sebuah laporan terdapat beberapa tabel, maka tabel harus diberi nomor urut secara sistematis sehingga dapat diidentifikasi atas dasar nomor urutnya dan bukan atas dasar namanya.

b. *Prefatory Notedan FootNote*

Catatan pendahuluan (*prefatory note*) dan catatan yang terdapat di bawah tabel (*foot note*) sebetulnya merupakan bagian yang integral dari sebuah tabel. Catatan pendahuluan biasanya ditaruh langsung di bawah nama tabel dalam bentuk yang agak kurang menonjol dibandingkan dengan namanya sendiri.

Penjelasan tentang angka-angka dalam kolom maupun pos-pos keterangan umumnya diberikan dalam catatan yang terdapat di bawah tabel itu sendiri. Umumnya, tanda-tanda catatan yang dipergunakan bagi penjelasan angka-angka ialah: \*, †, +, dan sebagainya. Sebaliknya, tanda catatan yang dipergunakan untuk memberi penjelasan ada pos-pos keterangan dalam kompartimen tabel maupun nama kolom umumnya ialah angka-angka romawi.

c. Sumber Data

Sumber data umumnya diletakkan di bawah tabel sesudah catatan. Sumber data harus diusahakan selengkap mungkin. Keterangan-keterangan tentang penulis, nama buku, jilid dan halaman buku, penerbit dan lain-lain harus lengkap dan tidak

meragukan. Bila data diambil dari data sekunder, sumber primer serta sumber sekundernya harus disebutkan secara lengkap.

d. Persentasi

Bila angka-angka persentasi dipergunakan di dalam tabel, maka pos-pos keterangan dalam kompartimen tabel harus memberi perincian yang jelas. Penyusun tabel yang kompeten selalu berusaha untuk menghindari istilah “persentasi” yang meragukan itu dengan jalan menggunakan istilah “persentasi dari jumlah”, “persentasi dari pertambahan atau penurunan”, dan sebagainya.

e. Jumlah

Bila jumlah angka dalam sebuah tabel merupakan sesuatu yang penting dan patut ditonjolkan, maka jumlah tersebut harus diletakkan pada sisi atas dalam kompartimen tabel atau sisi kiri dalam nama kolom. Ada kalanya, jumlah itu dapat juga dicetak atau ditulis dengan huruf tebal atau huruf besar. Bila jumlah itu tidak ingin ditonjolkan, maka dapat diletakkan di bagian bawah kompartimen atau pada sisi kanan nama kolom.

f. Unit

Unit pengukuran angka-angka yang terdapat dalam kolombiasanya sudah jelas dan tidak meragukan. Bila hal itu tidak mungkin, maka ciri-ciri unit pengukurannya harus dijelaskan dalam nama kompartimen atau nama kolom.

## **2. Bentuk Tabel**

Format sebuah tabel yang baik seharusnya jangan terlalu panjang atau terlalu pendek. Sudah tentu, bentuk tabel yang aktual sebenarnya harus disesuaikan dengan besarnya ruang laporan di mana tabel-tabel tersebut disajikan. Secara mendatar, bentuk tabel akan ditentukan oleh beberapa faktor sebagai berikut:

- a) Lebarnya kompartimen tabel dan hal tersebut ditentukan oleh pos-pos keterangan yang terpanjang. Sudah tentu, pos keterangan yang panjang dapat diatur dalam beberapa baris.
- b) Lebarnya tiap kolom dan hal tersebut ditentukan oleh jumlah angka yang terbesar.
- c) Cara mengatur spasi kata-kata.
- d) Cara mengatur tepi.

Sedangkan secara vertikal, bentuk tabel akan ditentukan oleh beberapa faktor sebagai berikut:

- a) Ruang yang dibutuhkan bagi nama, catatan pendahuluan, catatan yang terdapat di bawah tabel dan sumber data.
- b) Jumlah baris yang terdapat dalam tubuh tabel.
- c) Cara mengatur spasi kata-kata.
- d) Cara mengatur tepi.

### 3. Macam-Macam Tabel

Untuk menyusun sekumpulan data yang besar dan kecilnya belum tersusun ke dalam bentuk yang teratur, sebaiknya data itu disajikan dalam sebuah tabel. Tabel yang biasa digunakan terdiri atas tabel klasifikasi, kontingensi, dan distribusi frekuensi.

#### a. Tabel Klasifikasi

Tabel klasifikasi merupakan tabel yang menunjukkan atau memuat pengelompokan data. Tabel klasifikasi dapat berupa tabel klasifikasi tunggal atau ganda.

Tabel 2.3. Jumlah Mahasiswa Program Studi Magister Pendidikan Ekonomi

No	Tahun	Jumlah Mahasiswa Baru
1	2006	0
2	2007	13
3	2008	11
4	2009	29
5	2010	35
Total		88

Sumber: Buku Kinerja PPS UM

Daftar diatas tergolong daftar dengan klasifikasi tunggal, karena hanya satu klasifikasi berdasarkan tahun (2006-2010) dan jumlah mahasiswa. Dari data tersebut diketahui peningkatan dan penurunan jumlah mahasiswa setiap tahunnya.

Tabel 2.4. Penjualan PT.UM Unggul Menurut Jenis Barang dan Daerah Penjualan Tahun 2012 (Dalam Ribuan)/Dos

Jenis Barang	Daerah Penjualan				Total
	I	II	III	IV	
Air Mineral	20	30	50	60	160
Buah Apel	15	25	40	50	130
Cuka Ape	10	20	25	30	85
Total	45	75	115	140	375

Sumber: Image Magazine

Tabel 2.4 adalah tabel dua arah yang menunjukkan hubungan timbal balik antara dua hal : jenis barang dan daerah penjualan. Berdasarkan tabel itu juga akan diperoleh gambaran tentang perbandingan hasil penjualan antara daerah yang satu dengan yang lainnya. Hal itu akan memudahkan kita untuk melakukan analisa guna mengetahui jenis barang apa yang paling laku dan didaerah mana, yang selanjutnya dapat digunakan sebagai dasar penentuan alokasi barang-barang diberbagai daerah.

b. Tabel Kontingensi

Tabel kontingensi atau badan tabel (daftar) terdiri atas beberapa sel sesuai dengan perinciannya. Biasanya merupakan daftar dengan klasifikasi dua. Jika pada judul baris ada  $b$  baris dan pada judul kolom ada  $k$  kolom, maka diperoleh daftar kontingensi berukuran  $b \times k$ . Berikut ini ditampilkan sebuah contoh tabel kontingensi berukuran  $3 \times 3$ .

Tabel 2.5. Produksi Barang PT Nayamul Menurut Kualitasnya

	Kualitas I	Kualitas II	Kualitas III	Jumlah
Mesin A	5.478	1.423	1.002	7.903
Mesin B	8.203	3.042	906	12.151
Mesin C	6.217	2.361	1.568	10.146
Jumlah	19.898	6.826	3.476	30.200

c. Tabel Distribusi Frekuensi

Tabel Distribusi Frekuensi akan diuraikan tersendiri dalam bab selanjutnya. Namun sekedar mengenal tabel jenis ini, berikut diberikan sebuah contoh sederhana.

Tabel 2.6. Distribusi Frekuensi Skor Kinerja Guru SMP Negeri

No	Kelas Interval	Frekuensi
1	124 – 132	1
2	133 – 141	1
3	142 – 150	3
4	151 – 169	41
5	170 – 178	11
6	179 – 187	4
7	188 – 196	2
Jumlah		63

Sumber: Tesis

### C. PENYAJIAN DATA DALAM BENTUK GRAFIK

Cara lain untuk menyajikan data agar dapat memberikan gambaran yang lebih jelas, ialah melukiskannya dalam bentuk diagram. Diagram dalam fungsinya dapatlah disamakan dengan sebuah potret yang dapat memberikan gambaran serta uraian-uraian daripada tempat obyek dari gambar itu diambil.

Secara garis besar diagram terbagi menjadi dua macam, yakni diagram dan kartogram. Diagram yang sering digunakan meliputi diagram batang, diagram garis, diagram lambang dan diagram lingkaran. Penggunaan diagram dikarenakan penyajian data berupa gambar lebih menjelaskan laporan secara visual, untuk maksud tersebut pertama-tama diuraikan diagram batang.

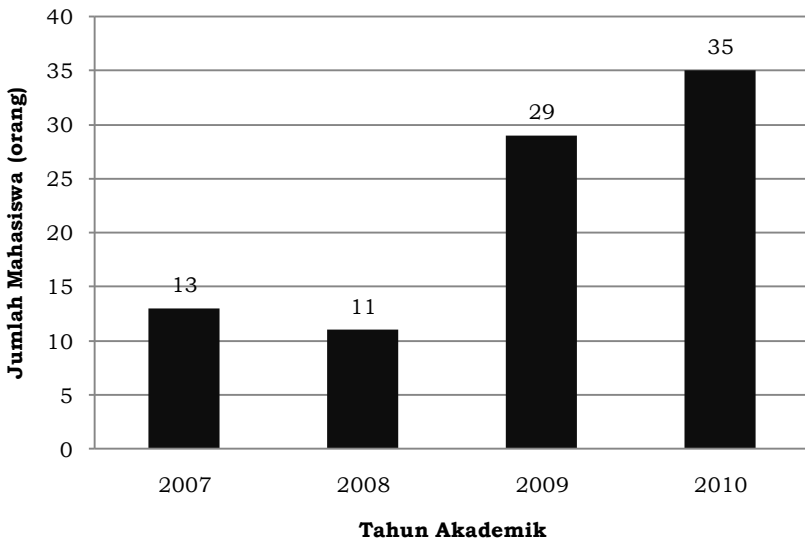
#### 1. Diagram Batang

Diagram Batang sangat cocok untuk menyajikan data yang berbentuk kategori atau atribut. Cara menggambarkan diagram batang sangat sederhana, yakni diagram batang memerlukan sumbu tegak dan sumbu datar yang berpotongan tegak lurus. Sumbu tegak dan sumbu datar terbagi menjadi beberapa skala



bagian yang sama. setelah membuat kedua buah sumbu, maka titik-titik pembagian sumbu datar dibuat garis tegak lurus, lalu diperlebar atau diberi warna, tingginya sama dengan frekuensi dalam tiap pembagian atribut. Pada bagian bawah sumbu datar dituliskan atribut atau waktu dan pada sumbu tegak dituliskan nilai data. Sewaktu membuat diagram perlu dipikirkan diagram mana yang sesuai dengan sekumpulan data, jika kumpulan data itu mengenai hasil observasi kualitatif maka diagram batanglah yang lebih baik.

Marilah kita lakukan pembuatan diagram batang dengan informasi berikut. Perkembangan jumlah mahasiswa Program Studi Magister Pendidikan di Universitas Negeri Malang mengalami peningkatan. Hal ini ditandai dari jumlah mahasiswa pada tahun 2007 sebanyak 13 orang, tahun 2008 sebanyak 11 orang, tahun 2009 sebanyak 29 orang dan tahun 2010 sebanyak 35 orang. Data itu kemudian ditampilkan dalam diagram batang seperti pada Gambar 2.1.



Gambar 2.1. Jumlah Mahasiswa Baru Pada Program Magister Pendidikan Ekonomi di Universitas Negeri Malang Periode 2007-2010

## 2. Diagram Garis

Diagram garis biasanya digunakan untuk menggambarkan suatu data berkelanjutan dalam kurun waktu tertentu, misalnya data tentang produksi dari tahun ke tahun, nilai ekspor suatu jenis barang dari tahun ke tahun dan sebagainya. Diagram garis juga digunakan untuk mengetahui perubahan yang seolah-olah keberlanjutan selama jangka waktu tertentu. Diagram ini juga digunakan untuk mengetahui bagaimana sifat perubahan data dari waktu ke waktu. Apakah perubahan itu suatu peningkatan yang sangat tinggi atau menggambarkan penurunan yang cukup menyolok.

Langkah-langkah dalam membuat diagram garis dijelaskan sebagai berikut :

- a) Buatlah dua sumbu, yaitu sumbu datar dan sumbu tegak. Pada sumbu datar biasanya menunjukkan waktu, sedangkan sumbu tegak menunjukkan frekuensi. Pembagian skala pada setiap sumbu tidak selalu mengambil skala yang sama.
- b) Sesuaikan data pada masing-masing sumbu, artinya data tahun pada sumbu datar ditarik lurus ke samping kanan, sehingga memotong pada satu titik.
- c) Jika semua data sudah disesuaikan masing-masing sumbu, maka akan terdapat sekumpulan titik-titik.
- d) Hubungkan titik-titik itu sehingga berbentuk garis.
- e) Di bagian tengah bawah diagram diberi nomor agar mudah dalam pencarian diagram. Biasanya nomor itu meliputi bab berapa materi itu sedang dibahas dan nomor diagram, lalu di bawahnya disertai penjelasan datanya.

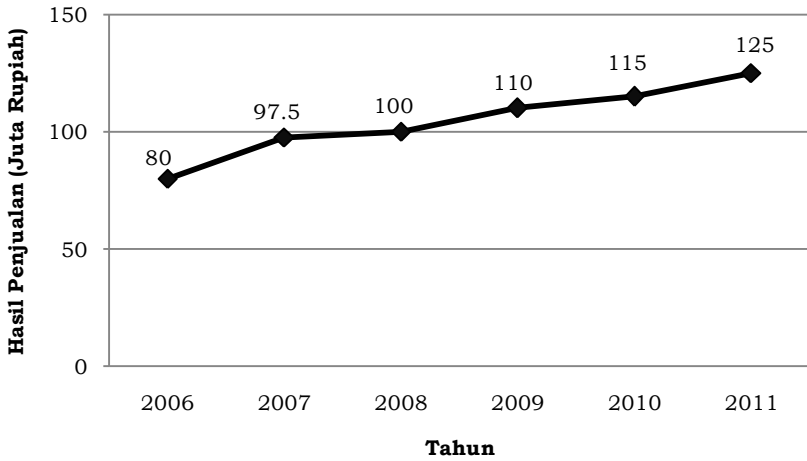
Berikut contoh diagram garis tunggal berupa garis yang menggambarkan perkembangan (trend) dari suatu karakteristik.

**Contoh 2.1.** Diketahui sebuah PT “Harapan Insan” menghasilkan omzet penjualan dari tahun 2006 hingga 2011 seperti pada tabel berikut:

Tabel.2.7. Penjualan PT Harapan Insan 2006-2011 (juta rupiah)

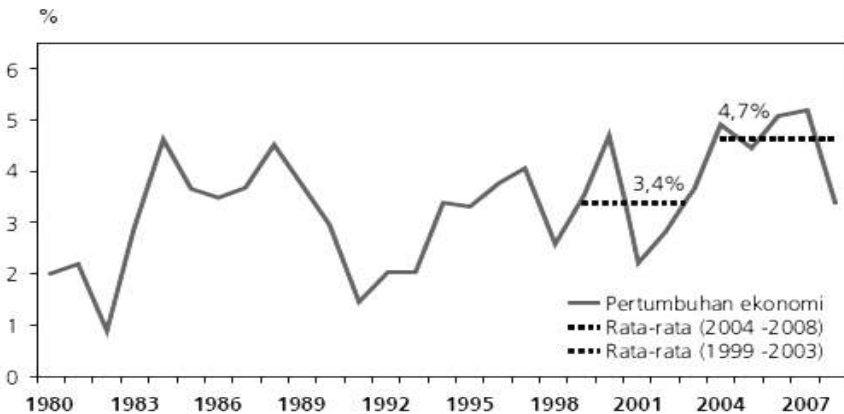
Tahun	2006	2007	2008	2009	2010	2011
Hasil Penjualan	80	97,5	100	110	115	125

Dari data-data pada tabel diatas, dapat dibuat diagram garis sebagai berikut :



Gambar 2.2. Grafik Perkembangan Penjualan

Contoh lain yakni diagram garis perkembangan ekonomi dunia pada tahun 2008, yang dikutip dari Outlook Ekonomi Indonesia 2009-2014



Sumber: WEO, Oktober 2008

Gambar 2.3. Grafik Perkembangan Ekonomi Dunia

### 3. Diagram Lingkaran

Diagram lingkaran biasanya digunakan untuk menyatakan perbandingan jika data itu terdiri atas beberapa kelompok. Terdapat banyak model diagram lingkaran, salah satunya adalah diagram kue yang merupakan diagram lingkaran berbentuk tiga dimensi dan setiap juring menunjukkan persentase kelompok. Diagram kue adalah tipe yang cantik untuk membandingkan bagian-bagian secara keseluruhan. Modifikasi bentuk adalah diagram kue diledakkan. Perubahan gambar tidak mempengaruhi data. Kita bisa menarik perhatian pada potongan kue tertentu yang dipandang signifikan dengan memindahkan kue keluar.

Barangkali penyajian data yang paling rumit adalah dalam bentuk diagram lingkaran. Hal ini disebabkan karena kita mesti menentukan porsi frekuensi dalam sebuah lingkaran, yang mesti menggunakan busur derajat, dengan cara sebagai berikut:

- a) Ubahlah nilai data absolut ke dalam bentuk persentase untuk masing-masing kategori.
- b) Ubahlah nilai data dalam bentuk persentase ke dalam satuan derajat untuk masing-masing kategori.
- c) Gambar lingkaran utuh, dan lalu tarik garis sembarang dari titik tengah lingkaran menuju ke garis lingkaran mana saja.
- d) Masukkan kategori-kategori lainnya ke dalam lingkaran yang sesuai dengan arah jarum jam
- e) Setiap kategori di dalam lingkaran diberi warna yang berbeda.
- f) Setiap kategori di dalam lingkaran hendaknya diberi identitas.

**Contoh 2.2.** Perhatikan tabel jumlah kendaraan bermotor di Indonesia pada tahun 1999 di bawah ini, selanjutnya buatlah diagram lingkaran berdasarkan tabel tersebut.

Tabel 2.8. Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 (dalam ribu)

Jenis Kendaraan	MPU	BIS	Mobil Gerobak (MG)	Sepeda Motor (SM)	Jumlah
(1)	(2)	(3)	(4)	(5)	(6)
Jumlah	925	191	788	4551	6455

Langkah 1. Mengubah data ke dalam bentuk presentase untuk masing-masing jenis kendaraan.

$$\text{MPU} = \frac{925}{6455} \times 100\% = 14\%$$

$$\text{BIS} = \frac{191}{6455} \times 100\% = 3\%$$

$$\text{MG} = \frac{788}{6455} \times 100\% = 12\%$$

$$\text{SM} = \frac{4551}{6455} \times 100\% = 71\%$$

Langkah 2. Mengubah nilai presentase untuk ke dalam satuan derajat masing-masing jenis kendaraan.

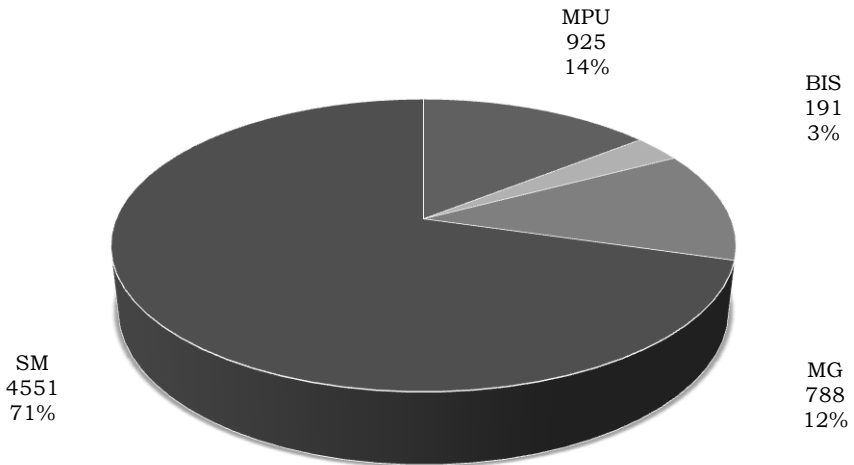
$$\text{MPU} = \frac{925}{6455} \times 360^\circ = 51,59^\circ$$

$$\text{BIS} = \frac{191}{6455} \times 360^\circ = 10,65^\circ$$

$$\text{MG} = \frac{788}{6455} \times 360^\circ = 43,95^\circ$$

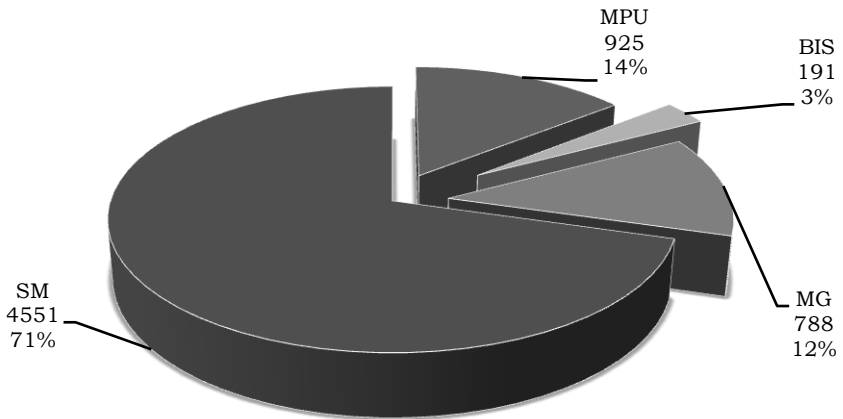
$$\text{SM} = \frac{4551}{6455} \times 360^\circ = 253,81^\circ$$

Langkah 3. Membuat diagram lingkaran seperti dibawah ini :



Gambar 2.4. Diagram (Lingkaran) Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 (dalam Ribuan)

Langkah 4. Membuat diagram kue yang diledakkan sebagaimana disajikan pada gambar dibawah ini :



Gambar 2.5. Diagram (Kue yang Diledakkan) Jumlah Kendaraan Bermotor di Indonesia Menurut Jenis Tahun 1999 (dalam Ribuan)

#### 4. Diagram Lambang

Diagram lambang sangat cocok untuk menyajikan data kasar sesuatu hal dan sebagai alat visual bagi orang awam. Setiap satuan yang disajikan lambang disesuaikan dengan jenis datanya, misalnya untuk data jumlah manusia dibuatkan gambar orang. Satu gambar orang menyatakan sekian jiwa tergantung kebutuhannya. Kelemahannya ialah jika data yang dilaporkan tidak penuh (bulat) maka lambangnya pun menjadi tidak utuh.

Langkah-langkah dalam membuat diagram lambang dapat dijelaskan secara ringkas sebagai berikut :

- a) Kita buat tiga buah kolom, dengan ketentuan sebagai berikut:
  - Kolom pertama berisi nama-nama kategori,
  - kolom kedua berisi lambang yang digunakan,
  - kolom ketiga berisi frekuensinya.
- b) Di bawah diagram diberi catatan berupa lambang (gambarkan) yang mewakili sejumlah obyek tertentu. Bilangan yang dipakai untuk lambang ini jangan terlalu besar dan sebaliknya.

- c) Tulis nama kategori pertama dan gambarkan lambangnya pada kolom lambang serta tuliskan banyak datanya pada kolom frekuensi.
- d) Banyaknya lambang yang digambarkan tidak sama dengan banyaknya yang ada tetapi kalau dikalikan dengan bilangan yang mewakili lambang tersebut sama dengan frekuensinya. Dengan demikian, mungkin ada lambang yang digambarkan secara tidak utuh.

Contoh diagram lambang digambarkan secara sederhana seperti pada gambar di bawah ini :

Lokasi Proyek	Lambang	Nilai Realisasi
Jawa Timur		3 Trilyun
Jawa Barat		2,5 Trilyun
DKI Jakarta		3 Trilyun
Sulawesi Tengah		2 Trilyun
Sulawesi Selatan		1,5 Trilyun

Keterangan :  = 500 Milyar

Gambar 2.6. Realisasi Investasi PMDN Triwulan I Tahun 2011 Berdasarkan Lokasi Proyek

## 5. Diagram Pencar

Diagram pencar sangat cocok untuk menyajikan data yang terdiri atas dua variabel. Diagram dibuat dalam sistem koordinat. Diagram pencar ini berfungsi pula untuk menentukan apakah suatu data itu linier. Diagram ini digunakan untuk mendukung hubungan logis antara dua variabel. Grafik ini disebut Diagram pencar (atau ada juga yang menyebutnya Diagram Tebaran), yang menunjukkan titik-titik tertentu. Setiap titik memperlihatkan suatu hasil yang kita nilai sebagai variabel tak bebas (*dependent*) maupun bebas (*independent*). Diagram Pencar memiliki manfaat untuk: (a) menunjukkan apakah terdapat hubungan bermanfaat antara dua variabel, dan (b) menetapkan tipe persamaan yang menunjukkan gubungan antara kedua variabel tersebut.

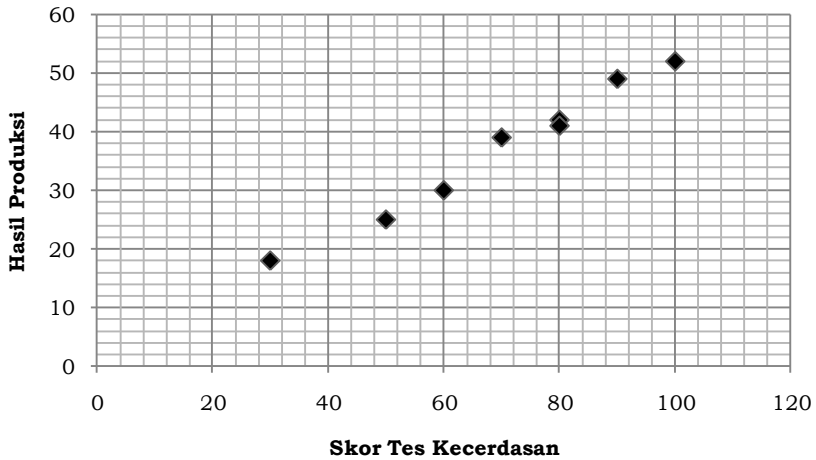
Kita dapat menjelaskan tujuan dan manfaat diagram pencar dengan menggunakan data pada tabel. Sebagai contoh, terdapat data produksi karyawan dalam satuan lusin (sebagai variabel tak bebas) dan hasil *aptitude test* atau tes kecerdasan 8 orang karyawan (sebagai variabel bebas). Jika hasil tes kecerdasan menunjukkan asumsi yang kita perkirakan, rasanya masuk akal bila kita asumsikan bahwa karyawan yang mempunyai nilai (skor) tinggi akan memberi hasil produksi yang lebih tinggi juga.

**Contoh 2.3.** Buatlah Diagram pencar dari data hubungan skor kecerdasan karyawan terhadap hasil produksi yang telah disajikan pada Tabel 2.9.

Tabel 2.9. Hubungan Kecerdasan dengan Hasil Produksi

Karyawan	Skor Tes Kecerdasan	Hasil Produksi
A	60	30
B	90	49
C	30	18
D	80	42
E	70	39
F	50	25
G	80	41
H	100	52





Gambar 2.7. Diagram Pencar Hubungan Skor Tes Kecerdasan dan Hasil Produksi

Data dari seorang karyawan akan menunjukkan satu titik tertentu pada diagram pencar, seperti yang terlihat pada Gambar 2.7. (Titik-titik yang memperlihatkan karyawan C dan F diberi tanda untuk memperlihatkan bagaimana kedua pengamatan terhadap karyawan tersebut digunakan untuk membuat grafik). Seperti yang kita lihat pada gambar, kedelapan titik membentuk gerakan sedemikian rupa sehingga seakan membentuk suatu garis lurus, dan hubungan yang sangat erat terlihat dengan fakta bahwa semua titik sangat dekat dengan garis lurus yang ditetapkan. Kita juga dapat melihat bahwa ada hubungan yang positif (langsung) antara kedua variabel. Sebagai contoh, bila skor tes kecerdasan meningkat, maka hasil produksi juga meningkat. Tentu ada kemungkinan variabel tertentu terdapat hubungan negatif (berlawanan), yaitu bila X meningkat, Y menurun.

#### D. RINGKASAN

- 1) Tabel statistik umumnya dibedakan ke dalam *tabel referensi* (*reference table*) dan *tabel ikhtisar* (*summary table*). Tabel referensi memiliki fungsi sebagai “gudang keterangan” karena

tabel itu memberi keterangan-keterangan yang terperinci dan disusun khusus guna kepentingan referensi.

- 2) Sebuah tabel formal umumnya terdiri dari beberapa bagian yakni *Nama* (titel), *nama kolom* dan *nama kompartimen*, bagian-bagian tersebut harus jelas dan singkat.
- 3) Bentuk tabel ditentukan oleh beberapa faktor sebagai berikut:
  - a) Lebar kompartimen tabel dan hal tersebut ditentukan pos-pos keterangan terpanjang. Pos keterangan panjang bisa diatur dalam beberapa baris.
  - b) Lebarnya tiap kolom ditentukan jumlah angka terbesar.
  - c) Cara mengatur spasi kata-kata.
  - d) Cara mengatur tepi.
- 4) Tabel yang biasa digunakan terdiri atastabel klasifikasi, tabel kontingensi dan tabel distribusi frekuensi.
- 5) Diagram terbagi atas diagram dan kartogram. Diagram yang sering dikenal adalah diagram batang, diagram garis, diagram lambang dan diagram lingkaran.

## **E. SOAL DAN LATIHAN**

- 1) Jelaskan perbedaan tabel referensi dengan tabel ikhtisar!
- 2) Menurut pendapat Anda, mana penyajian data yang lebih baik dalam bentuk tabel atau diagram? Berikan alasan Anda.
- 3) Berikan contoh penyajian tabel 1 arah, 2 arah dan 3 arah.
- 4) Jelaskan cara pembuatan diagram lingkaran dan penggunaan dalam menyajikan data!
- 5) Buatlah diagram garis dan diagram batang dari data berikut:

Tahun	Kredit Langsung
2003	1.722
2004	2.548
2005	3.742
2006	4.365
2007	6.938

## Bab 3.

# DISTRIBUSI FREKUENSI

### KOMPETENSI DASAR

Mahasiswa dapat menyusun penyajian data dalam bentuk distribusi frekuensi dan bentuk grafik atau diagram.

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- Menjelaskan tentang arti dan manfaat distribusi frekuensi.
- Menyusun tabel distribusi frekuensi relatif.
- Menyusun tabel distribusi kumulatif.
- Menyusun distribusi frekuensi data tunggal.
- Menyajikan data bentuk distribusi frekuensi data kelompok.
- Menyajikan data bentuk grafik (histogram, poligon dan ogive).

### DESKRIPSI SINGKAT

Bab ini membahas tentang penyajian data dengan tabel distribusi frekuensi, dimulai dari pemahaman konsep distribusi frekuensi yang meliputi pengertian dan cara penyusunan tabel distribusi frekuensi baik data tunggal maupun data kelompok. Selanjutnya diuraikan bahasan tentang penyajian grafik distribusi frekuensi baik dalam bentuk histogram, poligon, dan ogive. Akhir bab ini menguraikan tentang distribusi frekuensi relatif dan kumulatif

### KATA KUNCI

Distribusi Frekuensi	Panjang Kelas
Frekuensi Persentase	Batas Kelas
Frekuensi Relatif	Interval Kelas
Distribusi Frekuensi Relatif	Histogram
Distribusi Frekuensi	Poligon
Persentase	Nilai Kelas Interval
Rumus Struges	Kurva Ogive
Range	

## A. PENGANTAR

Setiap kali kegiatan pengumpulan data statistik, umumnya kegiatan itu menghasilkan kumpulan data angka yang keadaanya tidak teratur, berserak dan masih merupakan bahan keterangan bersifat mentah, karena hanya berupa kumpulan angka dengan

kondisi yang belum dapat memberikan informasi secara ringkas dan jelas mengenai ciri atau sifat yang dimiliki angka tersebut. Peranan dan kegunaan statistika sebagai ilmu pengetahuan adalah menyajikan atau mendeskripsikan data angka yang telah berhasil dihimpun secara teratur, ringkas, mudah dimengerti, hingga secara jelas dapat memberikan gambaran yang tepat mengenai ciri atau sifat yang terkandung di dalam angka tersebut.

Data pertama yang diperoleh pada suatu observasi disebut data mentah (*raw data*). Data ini belum tersusun secara numerik. Sebagai contoh data tinggi badan siswa yang penyajiannya masih dalam bentuk presensi kehadiran yang biasanya hanya diurutkan berdasarkan alphabet nama siswa. Seringkali data mentah disajikan berdasarkan urutan naik (*ascending*) atau urutan turun (*descending*). Bentuk penyajian seperti ini disebut array. Selisih antara nilai data terbesar dan terkecil disebut rentang (*range*).

Data yang diperoleh dalam proses pengumpulan data pada sebuah observasi tersebut dapat berupa data kualitatif maupun kuantitatif, Untuk dapat memahami data dengan mudah, data kualitatif maupun data kuantitatif harus disajikan dalam bentuk ringkas dan jelas. Salah satu cara untuk meringkas data adalah berbentuk tabel kontigensi yang telah dibahas pada bab sebelumnya. Selain itu, terdapat cara lain untuk menyajikan data bentuk tabel selain daftar baris – kolom dan tabel kontingensi yaitu tabel distribusi frekuensi, yang sekilas telah disinggung pada bab sebelumnya. Distribusi frekuensi bisa dimaknai sebagai pengelompokan data ke dalam beberapa kelompok (kelas) dan kemudian dihitung jumlah data yang masuk ke dalam tiap kelas.

Berdasarkan dari uraian singkat di atas, dapat disimpulkan definisi distribusi frekuensi adalah *pengelompokan data ke dalam beberapa kategori yang menunjukkan banyaknya data dalam setiap kategori, dan setiap data tidak dapat dimasukkan ke dalam atau lebih kategori.*

Dalam bekerja dengan jumlah data yang besar, biasanya lebih menguntungkan jika data itu disajikan dalam kelas-kelas atau kategori bersamaan dengan frekuensi yang bersesuaian. Frekuensi yang dimaksud adalah banyaknya kejadian yang ada pada kelas-kelas tertentu. Suatu tabel yang menyajikan kelas-kelas data beserta frekuensinya disebut distribusi frekuensi atau tabel frekuensi.

Tabel distribusi frekuensi merupakan susunan data dalam suatu tabel yang telah diklasifikasikan menurut kelas-kelas atau kategori tertentu. Menurut pembagian kelasnya, yaitu distribusi frekuensi kualitatif (kategori) dan distribusi frekuensi kuantitatif (bilangan). Pada distribusi frekuensi kualitatif, pembagian kelas didasarkan pada kategori tertentu dan skala nominal. Sedangkan kategori kelas dalam distribusi frekuensi kuantitatif terdapat dua macam yakni data tunggal dan kategori data berkelompok.

## **B. DISTRIBUSI FREKUENSI DATA KUALITATIF**

Data yang disajikan pada Tabel 3.1. adalah data kualitatif 50 orang pembeli komputer dari lima jenis perusahaan komputer. Dari data tersebut, kita kesulitan untuk mengetahui dengan cepat, jenis komputer mana yang paling banyak diminati pembeli. Untuk menjawab pertanyaan itu, maka datanya perlu disajikan dalam bentuk distribusi frekuensi (Tabel 3.2.).

Tabel 3.1. Data 50 Orang Pembeli Komputer dari Beberapa Jenis Produsen di Harco Computer Malang

Acer	Fujitsu	Fujitsu	Fujitsu	Compaq
Axio	Acer	Fujitsu	Compaq	Axio
Sony	Fujitsu	Acer	Sony	Fujitsu
Fujitsu	Sony	Axio	Acer	Fujitsu
Sony	Fujitsu	Axio	Acer	Sony
Compaq	Compaq	Acer	Axio	Compaq
Compaq	Acer	Axio	Compaq	Acer
Acer	Axio	Compaq	Acer	Axio
Axio	Compaq	Acer	Axio	Compaq
Compaq	Acer	Acer	Axio	Axio

Tabel 3.2. Distribusi Frekuensi Pembelian 5 Merk Komputer

Perusahaan	Frekuensi
Acer	13
Axio	12
Sony	5
Fujitsu	9
Compaq	11
Jumlah	50

Tabel di atas telah menyajikan hasil distribusi frekuensi yang dilakukan terhadap data awal. Dari tampilan data tersebut dapat diketahui dengan cepat bahwa Accer merupakan jenis komputer yang paling banyak diminati dan Sony adalah jenis komputer yang peminatnya paling sedikit.

### C. PERSENTASE DATA KUALITATIF

Distribusi frekuensi menunjukkan jumlah atau banyaknya item dalam setiap kategori atau kelas. Meskipun demikian, kita sering tertarik untuk mengetahui proporsi atau persentase item dalam setiap kelas. *Frekuensi relatif* dari suatu kelas adalah proporsi item dalam setiap kelas terhadap jumlah keseluruhan item dalam data tersebut. Jika sekelompok data memiliki  $n$  observasi, maka frekuensi relatif dari setiap kategori atau kelas akan diberikan sebagai berikut :

$$F_r = \frac{\text{Frekuensi Kelas}}{n} \quad (3-1)$$

Sedangkan frekuensi persentase dari suatu kelas adalah frekuensi relatif kelas tersebut dikalikan dengan 100. Frekuensi relatif merupakan proporsi, persentase = proporsi dikalikan 100. Distribusi frekuensi relatif adalah ringkasan ringkasan berbentuk tabel dari sekelompok data yang menunjukkan frekuensi relatif dari setiap kelas. Distribusi frekuensi persentase merupakan ringkasan bentuk tabel dari sekelompok data yang

menunjukkan frekuensi persentase bagi setiap kelas. Dengan menggunakan rumus frekuensi relatif di atas, kita akan mendapatkan data tentang pembelian komputer. Dari Tabel 3.2 dapat kita hitung frekuensi relatif untuk Acer, yaitu  $\frac{13}{50} = 0,26$ , frekuensi relatif untuk Axio, yaitu  $\frac{12}{50} = 0,24$ , dan seterusnya. Sedangkan untuk mendapatkan frekuensi persentase, frekuensi relatif tersebut dikalikan dengan 100.

Tabel 3.3. Distribusi Frekuensi Relatif dan Persentase Pembelian Komputer

Perusahaan	Frekuensi Relatif	Frekuensi Persentase(%)
Acer	0,26	26
Axio	0,24	24
Sony	0,10	10
Fujitsu	0,18	18
Compaq	0,22	22
Total	1,00	100

## D. DISTRIBUSI FREKUENSI DATA KUANTITATIF

Definisi tentang distribusi frekuensi berlaku baik untuk data kualitatif maupun kuantitatif. Meskipun demikian, kita harus lebih hati-hati dalam menentukan kelas yang digunakan pada distribusi frekuensi. Ada tiga hal yang perlu diperhatikan dalam menentukan kelas bagi distribusi frekuensi untuk data kuantitatif yaitu jumlah kelas, lebar kelas dan batas kelas. Kategori kelas dalam distribusi frekuensi kuantitatif, terdapat dua macam yaitu kategori data tunggal dan kategori data kelompok (bergolong).

### 1. Distribusi Frekuensi Data Tunggal

Distribusi frekuensi tunggal merupakan jenis distribusi frekuensi yang mengelompokkan data mentah berdasarkan kategori tunggal, bukan kelompok (*ungrouped data*). Biasanya jenis ini digunakan untuk jenis data yang jarak interval nilai tertinggi dan terendah maksimal 10.

**Contoh 3.1.** Penelitian mengenai Pendapatan Asli Daerah (PAD) 10 kabupaten/kota di Indonesia secara acak menunjukkan data mentah sebagai berikut (dalam milyar rupiah): 20, 21, 25, 25, 21, 21, 26, 21, 20, 21. Guna membuat data mentah tersebut bermakna, maka perlu dilakukan pengolahan data. Cara yang paling sederhana untuk mengolah data tersebut adalah dengan *Data Array* yaitu mengurutkan data dari terendah ke tertinggi atau sebaliknya. Untuk mengarray data mentah di atas dilakukan sebagai berikut: Dari PAD terendah: 20, 20, 21, 21, 21, 21, 21, 25, 25, 26. Dengan data array akan mempermudah untuk melihat distribusi frekuensinya. Namun apabila data mentah jumlahnya ratusan, penggunaan data array tidak lagi efisien, sehingga digunakan distribusi frekuensi menggunakan tabulasi.

**Contoh 3.2.** Distribusi frekuensi menggunakan tabulasi untuk distribusi frekuensi tunggal adalah sebagai berikut nilai ujian tengah semester mata kuliah statistik pendidikan dari 40 mahasiswa non reguler Pendidikan Ekonomi UNMUL tahun 2011 adalah sebagai berikut :

80, 50, 70, 40, 40, 50, 70, 70, 60, 40, 70, 60, 60, 50, 40, 80, 80, 70, 60, 50, 50, 60, 70, 80, 40, 50, 70, 60, 70, 60, 70, 70, 60, 60, 80, 60, 60, 40, 40, 50.

Langkah-langkah membuat distribusi frekuensi adalah:

- Buatlah tabel dengan tiga kolom, terdiri atas kolom 1: kuantitas komplain, kolom 2: Jari-jari, dan kolom 3: frekuensi.
- Tentukan nilai terendah dan tertingginya, kemudian masukkan sebagai kriteria dalam kelas.
- Lakukan tabulasi berdasarkan kolom atau baris
- Isi frekuensi yang ada berdasarkan hasil jari-jari

Adapun hasil data-data di atas dapat disajikan dalam secara lebih ringkas dalam bentuk tabel distribusi frekuensi data tunggal sebagai berikut :



Tabel 3.4. Distribusi Frekuensi Data Tunggal

Nilai	Turus	Frekuensi
40		7
50		7
60		11
70		10
80		5
Jumlah		$\sum f = 40$

## 2. Distribusi Frekuensi Kelompok

Distribusi frekuensi kelompok merupakan jenis distribusi frekuensi yang mengelompokkan data mentah berdasarkan kategori kelompok (*grouped data*). Sebuah distribusi frekuensi memiliki bagian-bagian pokok sebagai berikut:

### a) Kelas

Kelas merupakan kelompok nilai data (variabel). Penentuan jumlah kelas diserahkan sepenuhnya kepada penyusun distribusi frekuensi. Berapa sebaiknya jumlah kelas yang digunakan untuk distribusi frekuensi? Jawabannya adalah tidak ada pedoman baku yang dapat dijadikan sebagai cara menentukan jumlah kelas. Namun jangan terlalu sedikit dan jangan terlalu banyak. Untuk distribusi frekuensi yang memiliki kelas terlalu sedikit, tujuan pengelompokan data tidak akan tercapai. Sedangkan untuk distribusi frekuensi yang memiliki kelas terlalu banyak, dimungkinkan adanya kelas-kelas yang tidak memiliki data. Sebagai gambaran, jumlah kelas yang dibutuhkan biasanya berkisar dari 5 hingga 15 kelas.

Untuk memudahkannya dapat digunakan perumusan dari Sturges. Walaupun demikian, hasil perhitungan dengan cara perumusan Sturges tersebut sebaiknya tidak harus digunakan secara kaku.

$k = 1 + 3,322 \log n$	(3-2)
------------------------	-------

b) Batas Kelas

Batas kelas adalah nilai-nilai kelas yang membatasi kelas satu dengan kelas yang lain. Ada dua batas kelas, yaitu batas kelas bawah yang terdapat pada sisi kiri setiap kelas, serta batas kelas atas yang terdapat pada sisi kanan setiap kelas. Rumus penentuan batas kelas yakni:

$Bb + i - 1 \rightarrow Bb$ (Batas bawah), $i$ (interval kelas)	(3-3)
---	-------

c) Tepi Kelas

Tepi kelas atau batas riil kelas adalah batas kelas yang tidak memiliki lubang untuk angka tertentu antara kelas yang satu dengan yang lain, yang terdiri atas tepi bawah kelas serta tepi atas kelas. Penentuan tepi kelas adalah dengan mencari titik tengah antara batas atas kelas dengan batas bawah kelas di atasnya.

d) Titik Tengah Kelas

Titik tengah kelas adalah angka atau nilai data yang terletak tepat di tengah suatu kelas. Titik tengah adalah representasi kelas yang bersangkutan.  $TTK = \frac{1}{2}$  (Batas atas + Batas bawah). Titik tengah setiap kelas dapat dijadikan sebagai penaksir data asli yang sudah hilang sebagai akibat proses pengelompokkan. Titik tengah ini sebenarnya merupakan rata-rata hitung suatu kelas yang dihitung dengan membagi hasil jumlah batas kelas bawah dan batas kelas atas dengan angka 2.

e) Interval Kelas

Interval kelas adalah selang yang memisahkan kelas satu dengan kelas yang lain secara berurutan. Interval kelas tersebut ditentukan dari beda antara kelas bawah suatu kelas dengan batas kelas bawah kelas sebelumnya atau sesudahnya. Ada juga yang menyebut interval kelas dengan lebar kelas, yaitu jarak

antara tepi batas kelas bawah dengan tepi batas kelas atas suatu kelas. Namun pada situasi tertentu dimungkinkan interval atau lebar kelas yang tidak sama. Dalam menentukan interval kelas, perlu diketahui terlebih dahulu jangkauan atau beda antara angka data terbesar dan angka data terkecil. Selanjutnya dapat digunakan perumusan sederhana seperti berikut ini:

$$\text{Interval Kelas} = \frac{\text{Jangkauan}}{\text{Jumlah Kelas}} \quad (3-4)$$

f) Frekuensi Kelas

Frekuensi kelas adalah banyaknya data yang termasuk di dalam kelas tertentu.

**Contoh 3.3.** Contoh berikut ini penulis jadikan sebagai penjelasan dari beberapa keterangan tentang bagian-bagian diatas.

Tabel 3.5. Hasil UTS Statistik Mahasiswa FKIP UNMUL 2014

Nilai	Jumlah Mahasiswa
45 – 49	3
50 – 54	5
55 – 59	6
60 – 64	8
65 – 69	12
70 – 74	15
75 – 79	10
80 – 84	7
85 – 89	4
Jumlah	70

{ Kelas Interval } } Frekuensi

Dari tabel distribusi tersebut di atas dapat dijelaskan secara ringkas sebagai berikut :

- Banyaknya kelas adalah 9.
- Batas kelas bawah adalah 45; 50; 55; 60; 65; 70; 75; 80; 85.
- Batas kelas atas adalah 49; 54; 59; 64; 69; 74; 79; 84; 89.
- Tepi bawah 44.5; 49.5; 54.5; 59.5; 64.5; 69.5; 74.5; 79.5; 84.5.
- Tepi atas 49.5; 54.5; 59.5; 64.5; 69.5; 74.5; 79.5; 84.5; 89.5.
- Titik tengah 47; 52; 57; 62; 67; 72; 77; 82; 87.

### 3. Menyusun Tabel Distribusi Frekuensi Kelompok

Jika penyebaran angka/skor/nilai yang akan disajikan dalam bentuk tabel distribusi frekuensi mempunyai data yang besar, maka perlu dibuat tabel seperti disajikan diatas. Proses pembentukan distribusi frekuensi untuk data kelompok tentu melalui tahapan-tahapan, untuk memperjelas tahapan tersebut mari kita lihat contoh-contoh berikut:

**Contoh 3.4.** Diketahui hasil perolehan belajar mata kuliah Metodologi Penelitian 80 orang mahasiswa FKIP UNMUL Program Studi Pendidikan Ekonomi adalah sebagai berikut :

79	49	48	74	81	98	87	80	63	60
80	84	90	91	93	82	70	78	83	82
70	71	92	38	56	81	74	73	60	67
68	72	85	51	65	93	83	86	89	63
90	35	83	73	74	35	86	88	76	63
92	93	76	71	90	72	67	75	88	70
80	91	61	72	97	91	88	81	66	88
70	74	99	95	80	59	71	77	79	75

Menurut Anto Dajan (2008:83), agar data berupa deretan angka yang menunjukkan nilai perolehan belajar dapat disajikan dalam bentuk tabel distribusi frekuensi yang baik, maka perlu ditempuh 3 tahapan yakni penentuan jumlah kelas guna memasukkan angka-angka, memasukkan angka kedalam kelas yang sesuai serta menghitung frekuensinya dan membuat tabel distribusi frekuensi.

#### Langkah 1. Menghitung Besarnya Jangkauan Data

Besarnya jangkauan data dapat dihitung dengan mencari data tertinggi (*highest score*) dan data terendah (*lowest score*). Berdasarkan data diatas, maka diperoleh data tertinggi = 99 dan data terendah=35.

## Langkah 2. Menghitung Jangkauan (Range)

Jangkauan (*range*) dapat dihitung dengan menggunakan cara data tertinggi dikurangi data terendah.

$$R = \text{Data Tertinggi} - \text{Data Terendah}$$

Berdasarkan data diatas maka dapat ditentukan besar jangkauan data adalah  $99-35 = 64$

## Langkah 3. Menghitung Banyak Kelas

Banyaknya kelas dapat dihitung dengan menggunakan cara perumusan Sturges. Banyaknya kelas sebaiknya antara 7 dan 15, atau paling banyak 20, meskipun tidak ada aturan umum yang menentukan jumlah kelas. H.A. Sturges (1926) menulis artikel berjudul: "*The Choice of a Class Interval*" dalam *Journal of the American Statistical Association*, yang mengemukakan suatu rumus untuk menentukan banyaknya kelas sebagai berikut :

$$k = 1 + 3,322 \log n$$

dimana:  $k$  = banyak kelas  
 $n$  = jumlah data

Banyak kelas dari data-data yang disajikan di atas adalah:

$$\begin{aligned} k &= 1 + 3,322 \log 80 \\ &= 1 + 3,322 (1,9031) \\ &= 7.322 \approx 7 \text{ kelas (hasil dibulatkan ke bawah)} \end{aligned}$$

## Langkah 4. Menghitung Panjang Kelas (Interval)

Panjang kelas interval dapat dihitung dengan menggunakan rumus sebagai berikut :

$$I = R : k$$

dimana:  $k$  = banyak kelas  
 $R$  = jangkauan (*range*)  
 $I$  = panjang kelas (*interval*)

Berdasarkan data di atas diketahui panjang kelas (interval) adalah  $64 : 7 = 9,14$  dibulatkan menjadi 10.

### Langkah 5. Menentukan Batas Atas dan Batas Bawah

Batas bawah dan batas atas setiap kelas interval biasanya diambil dari data terkecil sebagai batas bawah. Selanjutnya kelas interval pertama dihitung dengan menjumlahkan batas bawah kelas dengan panjang kelas ( $i$ ) dikurang dengan 1 atau sesuai dengan rumus sebagai berikut :

$Bb + i - 1$
--------------

Nilai terkecil dari data di atas adalah 35, maka batas bawah adalah 35 dan batas atas adalah  $(35+10 - 1) = 44$ , sehingga kelas interval yang terbentuk adalah 35-44, 45-54, 55-64, 65-74, 75-84, 85-94, 95-104.

### Langkah 6. Mempersiapkan Tabel Distribusi Frekuensi

Tabel distribusi frekuensi terdiri atas tiga kolom yang berisi informasi tentang kelas interval, tally atau turus dan frekuensi.

Tabel 3.6. Distribusi Frekuensi Data Nilai Mata Kuliah Metodologi Penelitian dari 80 orang Mahasiswa FKIP UNMUL

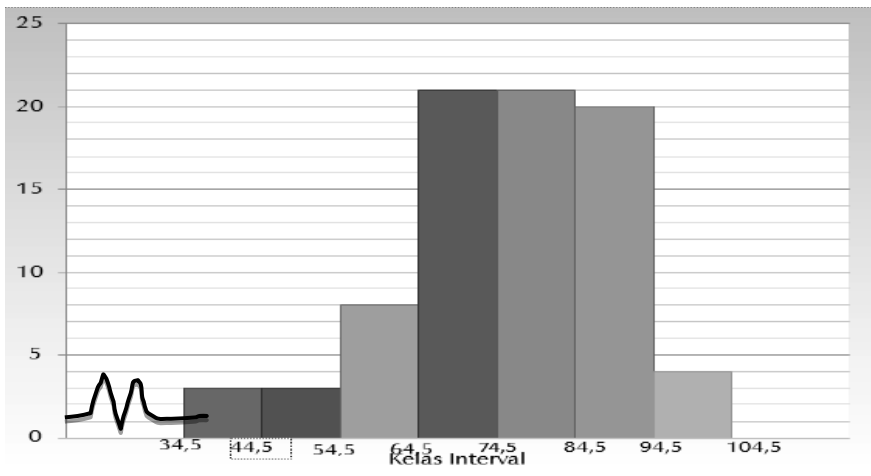
Kelas Interval	Turus	Frekuensi
35 – 44		3
45 – 54		3
55 – 64		8
65 – 74		21
75 – 84		21
85 – 94		20
95 - 104		4

## **E. PENYAJIAN GRAFIK DISTRIBUSI FREKUENSI**

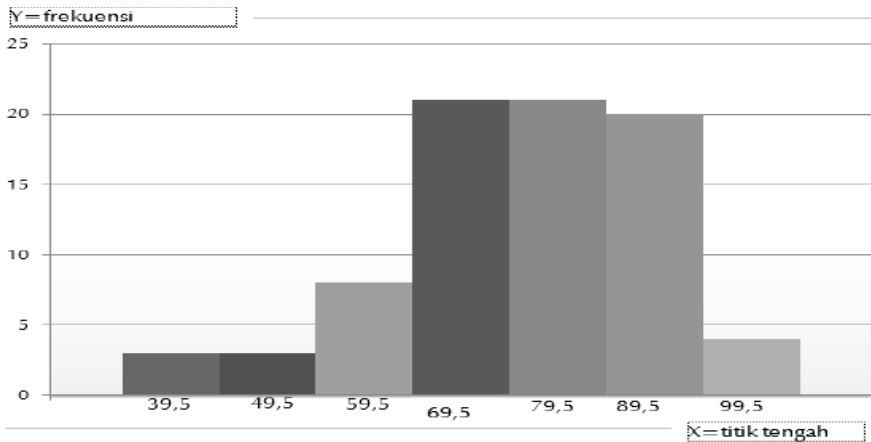
Penyajian data frekuensi dengan grafik frekuensi yang sederhana umumnya lebih menarik perhatian dan mengesankan. Dalam metode statistik, grafik frekuensi yang sering digunakan dalam analisa statistik antara lain (1) Histogram, (2) Poligon Frekuensi (*frequency polygon*), dan (3) Kurva Frekuensi yang Diratakan (*smoothed frequency curve*).

### 1. Histogram Frekuensi

Histogram acapkali dianggap sebagai grafik frekuensi yang bertangga. Salah satu fungsi histogram adalah menggambarkan beda antar kelas dalam distribusi. Dalam hal ini, histogram merupakan serangkaian empat persegi panjang yang memiliki alas sepanjang interval antara kedua tepi kelas dan memiliki luas sebanding dengan frekuensi pada kelas tersebut. Sumbu datar melukiskan data kelas-kelas interval. Jika ujung bawah kelas interval pertama cukup besar harganya, maka pada sumbu datar(horisontal), antara titik asal dan harga terjadi pemutusan.



Gambar 3.1. Histogram Hasil Belajar Metodologi

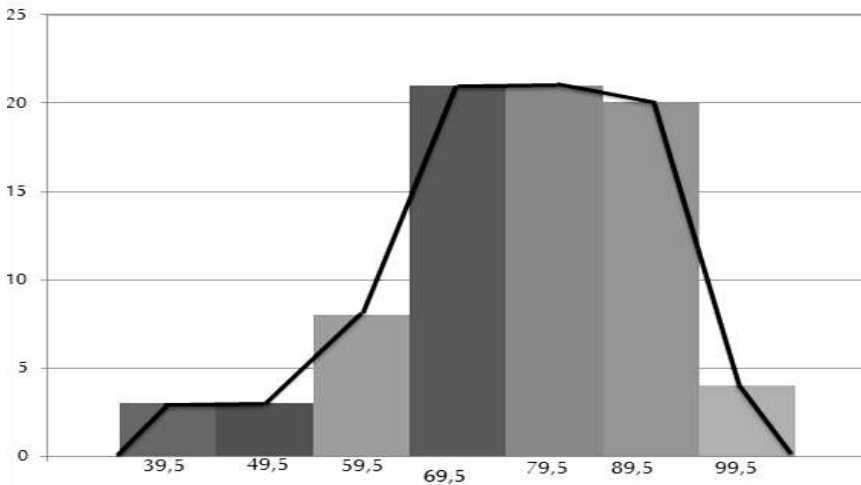


Gambar 3.2. Histogram Hasil Belajar Metodologi dengan Titik Tengah

## 2. Poligon Frekuensi

Jenis grafik lain yang banyak digunakan dalam menyajikan distribusi frekuensi adalah poligon. Dasar pembuatan poligon frekuensi sama halnya dengan pembuatan histogram. Sesuai dengan namanya yang berarti banyak sudut, poligon memang berbentuk garis patah-patah yang menghubungkan antara titik-titik tengah pada setiap puncak batang histogram sehingga tampak seperti benda bersudut banyak.

Bentuk poligon frekuensi dapat digabung dengan histogram seperti disajikan pada Gambar 3.3.



Gambar 3.3 Poligon Hasil Belajar Metodologi

Poligon frekuensi harus ditutup kedua ujungnya dengan menarik garis dari kedua ujuang batang histogram (kiri dan kanan) ke arah sumbu sumbu-x dengan skala yang sama seperti skala kelas interval lainnya. Poligon frekuensi bisa dibuat secara langsung tanpa harus menggambar histogram terlebih dulu. Caranya adalah dengan membuat tanda di atas titik tengah setiap kelas interval dengan jarak yang sesuai dengan frekuensinya kemudian titik-titik ini dihubungkan dengan garis dan ditutup pada kedua ujungnya seperti yang dijelaskan sebelumnya.



## F. DISTRIBUSI FREKUENSI RELATIF

Ada kalanya, analisa data statistik berhubungan erat dengan soal-soal yang berkaitan dengan perbandingan secara persentasi. Dalam hal tersebut, frekuensi distribusi perlu dinyatakan dalam bentuk persentasi atau proporsi. Distribusi yang berfrekuensi seperti itu dinamakan distribusi frekuensi relatif atau  $f$  (rel) atau distribusi presentasi atau  $f$  (%). Adapun nilai frekuensi relatifnya diperoleh dengan membagi nilai frekuensi masing-masing kelas interval dengan jumlah data dikalikan dengan 100% atau dengan menggunakan rumus sebagai berikut :

$$F \text{ (\%)} \text{ baris pertama} = \frac{f \text{ (abs) pertama}}{n} \times 100\% \quad (3-5)$$

Tabel 3.7 di bawah ini menyajikan kembali data dari Tabel 3.6 tentang hasil perolehan belajar metodologi penelitian yang telah ditambah dengan satu frekuensi relatif  $f$ (%).

Tabel 3.7. Distribusi Frekuensi Relatif Nilai Metodologi Penelitian

Kelas Interval	Frekuensi	$f$ (%)
35 – 44	3	3.75
45 – 54	3	3.75
55 – 64	8	10.00
65 – 74	21	26.26
75 – 84	21	26.25
85 – 94	20	25.00
95 - 104	4	5.00
Jumlah	80	100.00

## G. DISTRIBUSI FREKUENSI KUMULATIF

Dalam beberapa jenis analisa statistik, distribusi kumulatif umumnya lebih banyak digunakan daripada distribusi frekuensi biasa. Distribusi frekuensi kumulatif banyak sekali kegunaannya untuk menganalisa tentang upah buruh, perpajakan, penjualan dan sebagainya. Tabel di bawah ini kembali menggunakan data pada tabel-tabel sebelumnya, namun kini disajikan dalam bentuk bentuk distribusi kumulatif.

Tabel 3.8. Distribusi Frekuensi Kumulatif “Kurang Dari”

Kelas Interval	Frekuensi
Kurang dari 35	0
Kurang dari 45	3
Kurang dari 55	6
Kurang dari 65	14
Kurang dari 75	35
Kurang dari 85	56
Kurang dari 95	76
Kurang dari 105	80

Penentuan frekuensi bagi tiap-tiap kelas kedua distribusi di atas dilakukan dengan jalan: a) mengakumulasikan berturut-turut frekuensi kelas-kelas sebelumnya yang terdapat dalam distribusi biasa, b) memasukkan angka ke dalam kelas yang bersangkutan dengan menggunakan data asal serta kemudian menghitung frekuensinya.

Frekuensi kelas kelima diperoleh dengan menjumlahkan frekuensi kelas pertama sampai dengan frekuensi kelas keempat dari distribusi dalam Tabel 3.7. Distribusi kumulatif seperti di atas disebut sebagai Distribusi Kumulatif “Kurang Dari” (*less than distribution*).

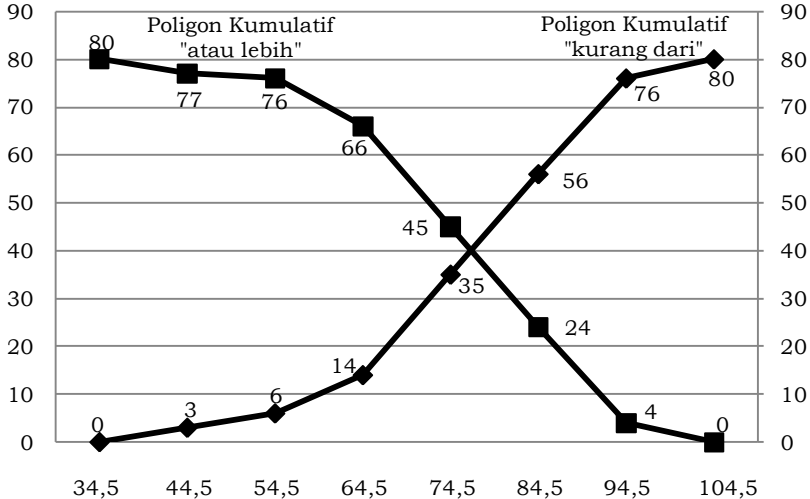
Sedangkan Distribusi Kumulatif “Atau Lebih” (*or more distribution*) dapat dibentuk dengan cara yang sama. Distribusi sedemikian itu dapat dilihat dalam Tabel 3.9.

Tabel 3.9. Distribusi Frekuensi Kumulatif “Atau Lebih”

Kelas Interval	Frekuensi
35 atau lebih	80
45 atau lebih	77
55 atau lebih	76
65 atau lebih	66
75 atau lebih	45
85 atau lebih	24
95 atau lebih	4
105 atau lebih	0

Pada hakekatnya, penggunaan tabel distribusi di atas dapat menghilangkan keraguan dalam memasukkan angka-angka ke dalam kelas-kelas tertentu. Penyajian secara grafis dari distribusi

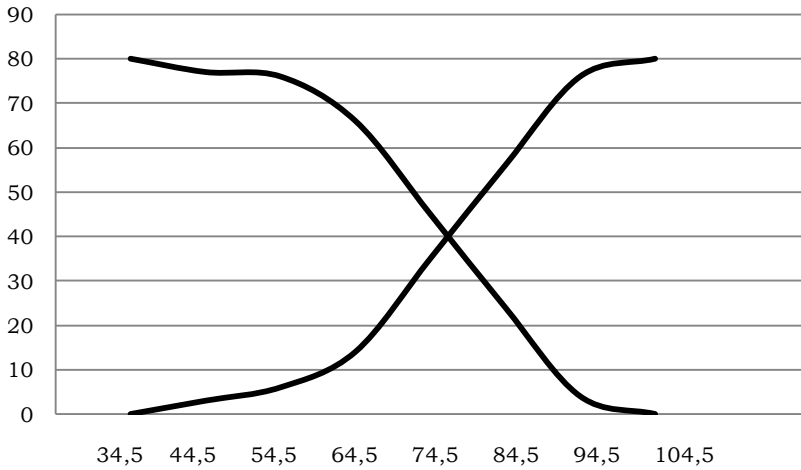
kumulatif “kurang dari” dan “atau lebih” dapat dilakukan dengan menggambarkan poligon frekuensinya. Penggambaran poligon frekuensi dapat dilihat pada gambar di bawah ini:



Gambar 3.4. Poligon Frekuensi Kumulatif "kurang dari" dan "atau lebih"

Poligon distribusi frekuensi kumulatif di atas dinamakan *ogive*. Penggambaran poligon dilakukan dengan menghubungkan semua titik-titik ordinat dari tepi kelas. Penggambaran poligon sedemikian itu hendaknya dimulai dari titik nol yang terdapat pada tepi kelas bawah dari interval kelas pertama. Di sini letak perbedaan dalam cara penggambaran poligon distribusi frekuensi dan poligon distribusi frekuensi kumulatif. Penggambaran poligon distribusi dilakukan dengan menghubungkan semua titik tengah interval kelas, sedangkan cara penggambaran poligon distribusi kumulatif dilakukan dengan jalan menghubungkan titik-titik ordinat tepi kelas.

Bila poligon frekuensi kumulatif yang disajikan di atas diratakan (*smoothed*), maka kita bisa memperoleh *kurva frekuensi kumulatif*. Kurva frekuensi kumulatif itu juga dinamakan *kurva ogive* dan dapat dilihat pada gambar di bawah ini :



Gambar 3.5. Kurva Ogive Perolehan Belajar Metodologi Penelitian

Kurva ogive seperti digambarkan di atas memiliki beberapa kegunaan yang patut diperhatikan. Kurva ogive pada umumnya digunakan untuk menggambarkan kumulasi frekuensi. Keraguan dalam memasukkan angka-angka ke dalam kelas tertentu yang mungkin timbul karena persoalan interval kelas atau batas kelas tidak perlu ada. Penghitungan statistik tentang median, kuartil dan desil lebih mudah dilakukan dengan bantuan kurva ogive di atas. Kurva ogive dapat digunakan bagi tujuan interpolasi.

**H. DISTRIBUSI FREKUENSI KUMULATIF RELATIF**

Seringkali orang tertarik untuk mengetahui dengan cepat banyaknya data yang memiliki nilai di atas atau di bawah nilai tertentu. Keperluan tersebut dapat terpenuhi dengan menyusun daftar atau tabel distribusi frekuensi kumulatif relatif. Daftar distribusi frekuensi kumulatif relatif ialah apabila nilai  $f_{kum}$  dalam frekuensi kumulatif di ubah dalam persen, contoh yakni pada Tabel 3.10 dan 3.11 yang menampilkan kembali data pada Tabel 3.8 dan 3.9 yang telah disajikan dalam bentuk distribusi kumulatif relatif.

Tabel 3.10. Distribusi Frekuensi Kumulatif Relatif “Atau Lebih”

Kelas Interval	Frekuensi	f (%)
35 atau lebih	80	100.00
45 atau lebih	77	96.25
55 atau lebih	76	95.00
65 atau lebih	66	82.50
75 atau lebih	45	56.25
85 atau lebih	24	30.00
95 atau lebih	4	5.00
105 atau lebih	0	0.00

Tabel 3.11. Distribusi Frekuensi Kumulatif Relatif “Kurang Dari”

Kelas Interval	Frekuensi	f (%)
Kurang dari 35	0	0.00
Kurang dari 45	3	3.75
Kurang dari 55	6	7.50
Kurang dari 65	14	17.50
Kurang dari 75	35	43.75
Kurang dari 85	56	70.00
Kurang dari 95	76	95.00
Kurang dari 105	80	100.00

## I. RINGKASAN

- 1) Distribusi frekuensi adalah pengelompokan data ke dalam beberapa kategori yang menunjukkan banyaknya data dalam setiap kategori, dan setiap data tidak dapat dimasukkan ke dalam atau lebih kategori.
- 2) Tabel distribusi frekuensi adalah susunan data dalam suatu tabel yang diklasifikasikan menurut kelas/kategori tertentu.
- 3) Tabel distribusi frekuensi menurut pembagian kelasnya terdiri atas distribusi frekuensi kualitatif dan distribusi frekuensi kuantitatif.
- 4) Frekuensi relatif dari suatu kelas adalah proporsi item dalam setiap kelas terhadap jumlah keseluruhan item dalam data.
- 5) Frekuensi persentase dari suatu kelas adalah frekuensi relatif kelas dikalikan 100. Frekuensi relatif merupakan proporsi, persentase = proporsi dikalikan 100.

- 6) Distribusi frekuensi tunggal adalah jenis distribusi frekuensi yang mengelompokkan data berdasarkan kategori tunggal, bukan kelompok (*ungrouped data*).
- 7) Langkah-langkah dalam membuat distribusi frekuensi adalah:
  - a) Menentukan jumlah kelas dengan rumus:  $k = 1 + 3,322 \log n$  dimana  $k$ =jumlah kelas dan  $n$ =jumlah data. Jumlah kelas minimal mengikuti aturan  $2^k > n$ .
  - b) Menentukan interval kelas yaitu (nilai tertinggi - nilai terendah) / jumlah kelas.
  - c) Melakukan penturusan yaitu memasukkan data ke dalam interval kelas yang ada.
  - d) Ketentuan dalam menyusun distribusi frekuensi adalah:
  - e) Tidak ada kelas tumpang tindih. Setiap data hanya masuk ke dalam satu kelas. Setiap interval kelas harus mempunyai ukuran sama. Jumlah kelas minimal 5 dan tidak lebih 15.
- 8) Penyajian data bisa dilakukan dengan membuat grafik seperti histogram, poligon dan ogive.
- 9) Histogram menghubungkan interval kelas dengan frekuensi.

## **J. SOAL DAN LATIHAN**

- 1) Jelaskan tentang pengertian distribusi frekuensi ?
- 2) Apakah perbedaan frekuensi relatif dan frekuensi persentase ?
- 3) Jelaskan ketentuan dalam menyusun distribusi frekuensi
- 4) Buatlah tabel distribusi frekuensi, tabel distribusi frekuensi kumulatif dan gambarkan grafik poligon frekuensi dari data nilai pelajaran IPS di Samarinda sebagai berikut :

79	49	48	74	81	98	87	80	63	60
80	84	90	91	93	82	70	78	83	82
70	71	92	38	56	81	74	73	60	67
68	72	85	51	65	93	83	86	89	63
90	35	83	73	74	35	86	88	76	63
92	93	76	71	90	72	67	75	88	70
80	91	61	72	97	91	88	81	66	88
70	74	99	95	80	59	71	77	79	75

## Bab 4.

# UKURAN PEMUSATAN DAN UKURAN LETAK

### KOMPETENSI DASAR

- Mendeskripsikan konsep penerapan prosedur statistik dalam menghitung ukuran pemusatan data (mean, median, modus).
- Mendeskripsikan konsep dan penerapan prosedur statistik dalam menghitung ukuran letak data (kuartil, desil, persentil).

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- Mendeskripsikan ukuran pemusatan dan ukuran letak.
- Mengidentifikasi jenis ukuran pemusatan dan ukuran letak.
- Menghitung ukuran pemusatan data (mean, median, modus).
- Menghitung ukuran letak meliputi kuartil, desil dan persentil.
- Menghitung modus data tidak dikelompokkan dan kelompok.
- Menjelaskan hubungan antara rata-rata, median dan modus.
- Menghitung nilai kuartil, desil dan persentil.

### DESKRIPSI SINGKAT

Bab ini membahas tentang konsep dan rumus yang berkaitan dengan ukuran pemusatan. Pada bagian pengantar dijelaskan tentang definisi dan perlunya memahami ukuran pemusatan, dilanjutkan dengan pemaparan tentang rata-rata hitung, rata-rata baik data yang tidak dikelompokkan maupun data yang berkelompok serta rata-rata harmonik. Kupasannya adalah tentang median dan hubungan ketiganya, serta diakhiri dengan bahasan mengenai kuartil, desil dan persentil.

### KATA KUNCI

Ukuran Pemusatan	Modus
Rata-Rata Hitung	Kuartil
Rata-Rata Ukur	Desil
Rata-Rata Median	Persentil

## A. PENGANTAR

Pada bab sebelumnya telah dijelaskan mengenai distribusi frekuensi, teknik penyajian data baik dalam tabel maupun diagram. Pada bab ini, akan dijelaskan tentang ukuran deskripsi data yaitu ukuran pemusatan dan ukuran letak data, baik data

mentah (data yang belum dikelompokkan) maupun data yang telah berbentuk distribusi frekuensi (data telah dikelompokkan).

Penyusunan data ke dalam bentuk distribusi frekuensi merupakan cara yang paling sederhana guna mengatur data secara sistematis dan mudah dimengerti. Penyajian data ke dalam bentuk grafik juga bertujuan memberi gambaran yang jelas tentang suatu peristiwa secara kuantitatif secara visual. Data tentang perolehan hasil belajar metodologi yang dibahas pada bab sebelumnya seharusnya dapat disederhanakan dan dinyatakan dalam suatu nilai atau karakteristik yang tunggal. Nilai tunggal tersebut merupakan nilai yang cukup representatif bagi seluruh nilai-nilai yang terdapat dalam data bersangkutan.

## **B. PENGERTIAN**

Keperluan menganalisis data lebih lanjut menjadi suatu hal yang sangat diperlukan dalam kegiatan/aktivitas keseharian. Di samping digunakan tabel dan grafik, diperlukan juga ukuran-ukuran yang dapat mewakili data yang akan dianalisis dan dapat digunakan untuk membandingkan keadaan dalam berbagai kelompok data. Untuk keperluan tersebut, statistik menyediakan suatu nilai berupa nilai tunggal yang mewakili keseluruhan nilai yang terdapat dalam data.

Nilai tunggal yang dianggap dapat mewakili keseluruhan nilai dalam data dianggap sebagai *mean* karena nilai itu dihitung berdasarkan keseluruhan nilai yang terdapat dalam data. Nilai mean itulah yang disebut dengan ukuran nilai pusat atau ukuran tendensi pusat. Dengan kata lain, salah satu tugas dari statistik adalah mencari angka di sekitar nilai-nilai dalam suatu distribusi memusat. Angka yang menjadi pusat sesuatu distribusi disebut tendensi pusat atau lazim disebut dengan tendensi sentral.

Para ahli statistik menganggap data tunggal tersebut adalah rata-rata (*averages*) karena merupakan nilai cukup representatif



bagi penggambaran nilai-nilai yang terdapat dalam data dan rata-rata ini dianggap sebagai *nilai sentral* dan dapat digunakan sebagai pengukuran lokasi distribusi frekuensi.

Rata-rata (*average*) adalah nilai yang mewakili himpunan atau sekelompok data (*a set of data*). Nilai rata-rata umumnya cenderung terletak di tengah suatu kelompok data yang disusun menurut besar/kecilnya nilai. Dengan perkataan lain, nilai rata-rata mempunyai kecenderungan memusat, sehingga juga disebut ukuran kecenderungan memusat (*measures of central tendency*). Beberapa jenis rata-rata yang sering dipergunakan ialah rata-rata hitung (*arithmetic mean* atau disingkat *mean*), rata-rata ukur (*geometric mean*), dan rata-rata harmonis (*harmonic mean*). Selain mempunyai keunggulan, setiap rata-rata tersebut juga memiliki kelemahan, dan ketepatan penggunaannya sangat tergantung pada sifat data dan tujuannya (misalnya untuk analisis).

Dalam bahan ajar ini, yang dimaksud dengan rata-rata ialah rata-rata hitung, kecuali apabila ada penjelasan lain. Dalam kehidupan sehari-hari, istilah rata-rata lebih banyak dikenal. Contohnya, rata-rata gaji/upah karyawan perusahaan swasta per bulan, rata-rata produksi beras per tahun, dan sebagainya. Rata-rata hitung, yang selanjutnya disebut rata-rata, sering digunakan sebagai dasar perbandingan dua kelompok nilai atau lebih.

### **C. RATA-RATA HITUNG (ARITHMETIC MEAN)**

Rata-rata hitung (atau sering disebut dengan rata-rata) merupakan suatu bilangan tunggal yang dipergunakan untuk mewakili nilai sentral dari sebuah distribusi. Dalam pemakaian sehari-hari orang awam lebih mempergunakan istilah rata-rata dari istilah rata-rata hitung. Bagi sekelompok data, rata-rata adalah nilai rata-rata dari data itu. Secara teknis dapat dikatakan bahwa rata-rata dari sekelompok variabel adalah jumlah nilai pengamatan dibagi dengan banyaknya pengamatan. Rata-rata

aritmatika atau rata-rata (mean) dari  $n$  buah data  $X_1, X_2, X_3, \dots, X_n$  dari data sampel dinyatakan dengan  $\bar{X}$  dibaca "X bar" sedangkan rata-rata yang diambil dari data populasi dinyatakan dengan  $\mu_x$  (baca : Myu X).

### 1. Rata-Rata Data yang Belum Dikelompokkan

Rata-rata dapat diperoleh dari menjumlahkan seluruh angka data yang selanjutnya dibagi dengan jumlah data. Jumlah data untuk data sampel disebut sebagai ukuran sampel yang disimbolkan dengan  $n$  dan data untuk populasi disebut sebagai ukuran populasi yang disimbolkan  $N$ , Jika  $X_1, X_2, X_3, \dots, X_n$  adalah angka-angka data yang banyaknya (jumlahnya) adalah  $n$ , maka rata-rata dihitung dengan menggunakan rumus sebagai berikut :

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \text{ atau } \bar{X} = \frac{\sum X_n}{n},$$

dimana  $\bar{X}$  = rata-rata,  $X_n$  = data ke- $n$  (4-1)

**Contoh 4.1.** Sebuah contoh secara acak berat 5 ekor ikan patin hasil pemancingan yaitu 700, 680, 750, 840 dan 810 gram. Hitunglah rata-rata berat ikan. Nilai rata-rata berat ikan tersebut adalah:

$$\bar{X}_B = \frac{700 + 680 + 750 + 840 + 810}{5} = 756$$

Jika data tersebut diambil dari populasi maka rumusnya adalah sebagai berikut:

$$\mu = \frac{\sum X_i}{N}, \text{ dimana } \mu = \text{rata-rata, } X_i = \text{data ke-}i, N = \text{data}$$

**Contoh 4.2.** Berikut ini adalah nilai ekspor minyak dan gas pada periode Januari hingga Juni 2012 (dalam juta US\$) yang dikutip dari Indikator Ekonomi BPS dengan data sebagaimana tercantum dalam Tabel 4.1. Berapakah rata nilai ekspor setiap bulannya ?

Tabel 4.1. Nilai Ekspor Minyak dan Gas Pada Periode Januari hingga Juni 2012

Bulan	Nilai Ekspor Minyak dan Gas (Juta US\$)
Januari	3142,6
Pebruari	3555,5
Maret	3,486,1
April	3560,7
Mei	3724,9
Juni	2789,1

Sumber : Indikator Ekonomi BPS

Maka penyelesaian soal di atas dapat dilakukan sebagai berikut :

$$\mu = \frac{3142,6+3555,5+3486,1+3560,7+3724,9+2789,1}{6} = 2795,467$$

Apabila nilai  $X_1, X_2, \dots, X_n$  masing-masing memiliki frekuensi  $W_1, W_2, \dots, W_n$ , maka mean dihitung sebagai berikut:

$$\bar{X} = \sum \frac{W \cdot X}{W} = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n}{w_1 + w_2 + \dots + w_n}, W = \text{timbangan}$$

**Contoh 4.3.** Berikut ini adalah transkrip akademik seorang mahasiswa Pendidikan Ekonomi Universitas Mulawarman.

Tabel 4.2. Transkrip Akademik Mahasiswa

Mata Kuliah	Nilai Mutu	Angka Mutu( $x_i$ )	SKS ( $W_i$ )	$W_i x_i$
Ekonomi Mikro	B	3	2	6
Teori Ekonomi	A	4	4	16
Ekonometrika	C	2	3	6
Pengantar Manajemen	A	4	3	12
$\Sigma$		14	12	40

## 2. Rata-Rata Data yang Dikelompokkan

Menghitung rata-rata memang lebih menguntungkan jika dihitung dari data yang belum dikelompokkan, karena hasil hitungannya lebih mencerminkan fakta yang sebenarnya. Apakah rata-rata data yang telah dikelompokkan merupakan data yang sebenarnya? Dalam kehidupan sehari-hari, data yang dibutuhkan

sering sudah disajikan dalam bentuk distribusi frekuensi, seperti yang telah disajikan dalam berbagai terbitan maupun laporan.

Pada data observasi yang disajikan dalam bentuk distribusi frekuensi (telah dikelompokkan), sifat keaslian data telah hilang, sehingga penghitungan rata-rata memerlukan angka-angka yang dapat mengestimasi data asli. Titik tengah bisa dijadikan sebagai penaksir data asli yang tersebar pada masing-masing kelas dalam distribusi frekuensi. Ada dua cara yang dapat digunakan untuk menghitung rata-rata yang telah dikelompokkan yaitu metode defisional dan metode pengkodean.

a. Metode Defisional

Untuk menghitung rata-rata, titik-titik tengah masing-masing kelas, sebagai penaksir data asli, dikali dengan frekuensi masing-masing kelas. Hasil perkalian pada masing-masing kelas tersebut selanjutnya dijumlah dan kemudian hasil penjumlahan tersebut dibagi dengan jumlah data atau jumlah frekuensi seluruh kelas. Metode defisional dapat dirumuskan sebagai berikut:

$$\bar{X} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i}$$

**Contoh 4.4.** Berdasarkan data penimbangan berat 65 karung beras PT Makmur (dalam Kg), yang sudah diolah seperti pada Tabel 4.3. Berapakah rata-rata berat beras setiap karungnya?

Tabel 4.3. Rata-Rata Berat Beras PT Makmur

Beras (Kg)	Titik Tengah (Xi)	Jumlah Karung (fi)	fi.xi
45 – 50	47,5	5	237,5
51 – 56	53,5	7	374,5
57 – 62	59,5	10	595,0
63 – 68	65,5	20	1310,0
69 – 74	71,5	12	858,0
75 – 80	77,5	8	620,0
81 – 86	83,5	3	250,5
$\Sigma$		65	4245,5

Penyelesaian soal di atas dapat dilakukan sebagai berikut :

$$\bar{X} = \frac{\sum_1^n f_i x_i}{\sum_1^n f_i} = \frac{4245.5}{65} = 65.32 \text{kg}$$

b. Metode Simpangan

Metode lain yang dapat digunakan untuk menentukan nilai rata-rata hitung adalah menggunakan metode simpangan. Metode ini dilakukan dengan cara menentukan nilai M atau mean hitung sementara yang ditentukan dari nilai titik tengah kelas yang mengandung modus, lalu dijumlahkan dengan hasil pembagian, jumlah perkalian frekuensi dengan selisih titik tengah dengan rata-rata sementara (X-M) terhadap jumlah frekuensi totalnya sehingga dapat dirumuskan sebagai berikut:

$$\bar{X} = M + \frac{\sum fd}{\sum f}, \text{ dimana } M = \text{titik tengah kelas Modus, } d = X-M$$

**Contoh 4.5.** Berdasarkan data pada contoh 4.4 diatas, tentukan nilai rata-rata hitung dengan menggunakan metode simpangan.

Tabel 4.4 Rataan Berat Beras PT Makmur (Metode Simpangan)

Berat Beras (Kg)	(Xi)	f	d (X-M)	f.d
45 – 50	47,5	5	-18	-90
51 – 56	53,5	7	-12	-84
57 – 62	59,5	10	-6	-60
63 – 68	65,5	20	0	0
69 – 74	71,5	12	6	72
75 – 80	77,5	8	12	96
81 – 86	83,5	3	18	54
Σ		65	0	-12

$$\bar{X} = 65.5 + \frac{-12}{65} = 65.32$$

c. Metode Pengkodean

Seringkali data yang akan dihitung rata-ratanya berbentuk angka-angka besar seperti nilai penjualan, pembelian, piutang,

dan sebagainya. Jika angka-angka yang dihitung dalam satuan yang besar, maka penghitungan rata-rata dengan penggunaan metode defisional akan sedikit lebih menyulitkan.

Pada pertemuan sebelumnya telah dijelaskan bahwa interval kelas sebuah distribusi frekuensi, secara umum senantiasa sama. Hanya pada keadaan tertentu, interval kelas dimungkinkan tidak sama. Interval kelas yang sama, salah satunya dapat dilihat beda antar titik tengah yang senantiasa sama. Angka-angka berikut ini menunjukkan titik tengah dari data yang dikutip pada Tabel 4.3.

Titik Tengah	:	47,5		53,5		59,5		65,5		71,5
Interval	:		66		66		66		66	

Dengan interval kelas yang sama ini, sebenarnya, angka-angka titik tengah dapat diubah menjadi suatu skala dengan interval yang sama. Skala titik tengah ini lebih sering disebut sebagai kode titik tengah.

Langkah pertama dalam memberi kode titik tengah adalah menetapkan kelas yang nantinya diberi kode atau skala nol. Dalam menentukan kelas yang berkode nol ini sebenarnya tidak ada pedoman yang baku, akan tetapi sebaiknya kelas yang akan diberi kode nol adalah kelas yang berfrekuensi tinggi. Langkah berikutnya adalah menetapkan kode untuk kelas lainnya dengan mengurutkan mulai dari kelas berkode nol dengan interval yang sama. Interval kelas ini umumnya adalah satu, dari Tabel 4.3.diatas kelas yang akan diberi kode nol adalah kelas ke-4. Sehingga, Metode Pengkodean dirumuskan sebagai berikut:

$$\bar{X} = X_0 + i \left( \frac{\sum c_i f_i}{\sum f_i} \right),$$

dimana

$\bar{X}$  = rata-rata;

$X_0$  = titik tengah pada kelas berkode nol

$i$  = interval kelas

$c_i$  = kode titik tengah pada kelas ke-i

$\sum f_i$  = jumlah frekuensi

**Contoh 4.6.** Berdasarkan data penimbangan berat 65 karung beras PT Makmur (dalam Kg), yang sudah diolah seperti pada Tabel 4.3. Berapakah rata-rata berat beras setiap karungnya? (Gunakan Metode Pengkodean)

Tabel 4.5. Data Penimbangan Beras PT Makmur Menggunakan Metode Pengkodean

Titik Tengah( $X_i$ )	Jumlah Karung( $f_i$ )	$c$	$c f_i$
47,5	5	-3	-15
53,5	7	-2	-14
59,5	10	-1	-10
65,5	20	0	0
71,5	12	1	12
77,5	8	2	16
83,5	3	3	9
	$\Sigma(f_i) = 65$		$\Sigma(c f_i) = -2$

Diketahui:

$$X_0 = 65,5, \quad i = 6, \quad \Sigma c f_i = -2, \quad \Sigma f_i = 65$$

$$\begin{aligned} \bar{X} &= X_0 + i \left( \frac{\Sigma c f_i}{\Sigma f_i} \right) \\ &= 65,5 + 6 \left( \frac{-2}{65} \right) \\ &= 65,5 + 6 (-0,0307) \\ &= 65,5 - 0,185 \\ &= 65,315 \approx 65,32 \end{aligned}$$

**Contoh 4.7.** Nilai kontrak asuransi Jagat Raya pada 60 nasabah baru didistribusikan sebagai berikut:

Tabel 4.6. Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya (Tipe “Kurang Dari”)

Nilai Kontrak	Frekuensi
<Rp.10.000.000,00	0
<Rp.20.000.000,00	5
<Rp.30.000.000,00	17
<Rp.40.000.000,00	31
<Rp.50.000.000,00	46
<Rp.60.000.000,00	54
<Rp.70.000.000,00	60

Untuk menghitung rata-rata nilai kontrak menggunakan metoda defisional, bentuk penyajian di atas – bentuk distribusi frekuensi kumulatif tipe “kurang dari” – harus diganti menjadi bentuk distribusi frekuensi yang biasa, sehingga hasilnya menjadi sebagai berikut :

Tabel 4.7. Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Jagat Raya (Tipe Frekuensi Biasa)

Nilai Kontrak	Frekuensi
<Rp.10.000.000,00 - <Rp.20.000.000,00	5
<Rp.20.000.000,00 - <Rp.30.000.000,00	12
<Rp.30.000.000,00 - <Rp.40.000.000,00	14
<Rp.40.000.000,00 - <Rp.50.000.000,00	15
<Rp.50.000.000,00 - <Rp.60.000.000,00	8
<Rp.60.000.000,00 - <Rp.70.000.000,00	6
Jumlah	60

Setelah itu, proses pengerjaan berikutnya adalah :

Tabel 4.8. Penghitungan Rata-rata Nilai Kontrak Asuransi 60 Nasabah Baru PT Jagat Raya (Metoda “Defisional”)

$X_i$	$f_i$	$X_i \cdot f_i$
Rp. 15.000.000,00	5	Rp. 75.000.000,00
Rp. 25.000.000,00	12	Rp. 300.000.000,00
Rp. 35.000.000,00	14	Rp. 490.000.000,00
Rp. 45.000.000,00	15	Rp. 675.000.000,00
Rp. 55.000.000,00	8	Rp. 440.000.000,00
Rp. 65.000.000,00	6	Rp. 390.000.000,00
Jumlah	60	Rp. 2.370.000.000,00

$$\begin{aligned} \bar{X} &= \frac{2.370.000.000,00}{60} \\ &= \text{Rp.}39.500.000,00 \end{aligned}$$

### 3. Memahami Sifat Rata-Rata Hitung

- Setiap kelompok baik dalam bentuk skala interval maupun rasio mempunyai rata-rata hitung.
- Semua nilai data harus dimasukkan ke dalam perhitungan rata-rata hitung.



- c) Satu kelompok baik kelas maupun satu kesatuan dalam populasi dan sampel hanya mempunyai satu rata-rata hitung.
- d) Rata-rata hitung untuk membandingkan karakteristik dua atau lebih populasi atau sampel.
- e) Rata-rata hitung nilainya sangat dipengaruhi nilai ekstrem yaitu nilai yang sangat besar dan nilai yang sangat kecil.
- f) Bagi data dan sekelompok data yang sifatnya terbuka (lebih dari atau kurang dari) tidak mempunyai rata-rata hitung.
- g) Jumlah simpangan, selisih antara tiap data dengan rata-rata hitungnya adalah 0 atau ditulis dalam bentuk  $\Sigma (x_i - \bar{x}) = 0$
- h) Jumlah kuadrat dari simpangan-simpangan selalu lebih kecil atau sama dengan jumlah kuadrat antara bilangan-bilangan tersebut dikurangi oleh suatu bilangan sebaran. Secara matematis ditulis dengan notasi  $\Sigma (x_i - \bar{x})^2 \leq \Sigma (x_i - a)^2$ .
- i) Jika  $n_1$  data mempunyai rata-rata  $\bar{x}_1$ , jika  $n_2$  data mempunyai rata-rata  $\bar{x}_2$ , Jika  $n_3$  data mempunyai rata-rata  $\bar{x}_3$ , jika  $n_k$  data mempunyai rata-rata  $\bar{x}_k$ , maka rata-rata gabungan data tersebut adalah:

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3 + \dots + n_k\bar{x}_k}{n_1 + n_2 + n_3 + \dots + n_k}$$

**D. RATA-RATA UKUR (GEOMETRIC MEAN)**

Pada bidang bisnis dan ekonomi sering diperlukan data untuk mengetahui rata-rata persentase perubahan sepanjang waktu. Nilai rata-rata memerlukan data berkala (*time series*) untuk mengetahui kecenderungan, misalkan indeks ekonomi, tingkat pendapatan, tingkat produksi, rata-rata penjualan tiap tahun, dan lain-lain. Berapa besar rata-rata persentase tingkat perubahan per tahun? Nilai konstanta yang dapat menjelaskan tingkat perubahan per tahunnya. Nilai konstanta ini dapat dicari menggunakan rata-rata geometrik. Rata-rata ukur ini dipakai untuk menggambarkan keseluruhan data, khususnya bila data

tersebut mempunyai ciri tertentu, yaitu banyaknya nilai data yang satu sama lain saling berkelipatan sehingga perbandingan tiap dua data berurutan tetap atau hampir tetap. Rata-rata ukur dari  $k$  buah nilai adalah akar pangkat  $k$  dari hasil perkalian  $k$  peubah nilai tersebut.

Ada 2 (dua) kegunaan dari rata-rata ukur, yaitu:

- a) menentukan rata-rata pertambahan persentase penjualan barang dagangan/jasa, atau pertambahan persentase produksi dan lain-lain dari satu waktu ke waktu berikutnya.
- b) menentukan rata-rata persentase, indeks dan nisbah/relatif.

**1. Rata-Rata Ukur Data Tidak Berkelompok**

Jika ada data yang nilainya terdiri dari  $X_1, X_2, X_3, \dots, X_n$ , maka besar nilai rata-rata ukur untuk nilai tersebut merupakan akar pangkat  $n$  dari hasil masing-masing nilai kelompok tersebut. Untuk mencari rata-rata ukur data yang tidak berkelompok, terdapat 2 (dua) bentuk yang membedakannya, tergantung dari kebutuhan, yaitu:

- a. Mencari tingkat perubahan suatu data/nilai, yang biasanya dinyatakan dalam periode

$G = [\text{antilog } \frac{1}{n} (\log X_n - \log X_0)] - 1$
---

Dimana :

- G = Rata-rata ukur (Geometric mean)
- $n$  = Jumlah satuan waktu (periode)
- $X_n$  = Besarnya data/nilai pada akhir perkembangan
- $X_0$  = Besarnya data/nilai pada waktu permulaan

- b. Mencari jumlah rata-rata dari nilai yang di observasi

Rata-rata geometrik bagi nilai  $n$  bilangan positif  $X_1, X_2, \dots, X_n$ , adalah akar pangkat  $n$  dari hasil kali semua bilangan itu, yang dapat dirumuskan sebagai berikut :

$$G = \sqrt[n]{X_1 \cdot X_2 \dots X_n}$$

G merupakan rata-rata geometrik dari  $X_1, X_2, \dots, X_n$ . Nilai  $X_1, X_2, \dots, X_n$  menunjukkan rata-rata relatif. Selanjutnya gunakan logaritma pada masing-masing ruas dengan menggunakan rumus sebagai berikut :

$$\log G = \frac{1}{n} [\sum_{i=1}^n \log X_i]$$

Rata-rata geometrik G diselesaikan dengan mengambil anti-logaritmanya, dengan rumus sebagai berikut :

$$G = \text{antilog} \frac{1}{n} [\sum_{i=1}^n \log X_i]$$

Perhatikan bahwa logaritma rata-rata geometrik untuk  $n$  buah bilangan positif, sama dengan rata-rata hitung logaritma masing-masing bilangan.

## 2. Hubungan Rata-Rata Ukur dengan Bunga Majemuk

Rata-rata geometrik dapat digunakan untuk menentukan rasio laju kenaikan produksi, perubahan tingkat suku bunga, pertumbuhan penduduk, perkembangan bakteri dalam bejana, dan lain-lain. Hubungan tingkat bunga majemuk dengan rata-rata geometrik dapat dijelaskan berikut ini : Jika  $P_0$  merupakan jumlah uang mula-mula dan  $P_n$  adalah jumlah uang setelah tahun ke- $n$ , maka uang tersebut dibunga-majemukan dengan tingkat suku bunga  $r\%$  per tahun selama  $n$  tahun, sehingga :

$$P_n = P_0 (1 + r)^n$$

Namun bila tingkat suku bunga berubah-ubah dari waktu ke waktu yaitu  $r_1, r_2, \dots, r_n$  maka masing-masing nilai dibagi dengan  $P_0$ , dengan rumus sebagai berikut :

$$P_0 (1 + r)^n = P_0 (1 + r_1) \cdot (1 + r_2) \dots (1 + r_n)$$

Sehingga perhitungan tersebut bisa disederhanakan dengan menggunakan rumus sebagai berikut:

$$(1 + r) = \sqrt[n]{(1+r_1) \cdot (1+r_2) \cdot \dots \cdot (1+r_n)}$$

**Contoh 4.8.** Seorang pengusaha mempunyai uang Rp 1.000.000, yang ditabung dengan bunga majemuk 3% pertahun. Berapakah uang tersebut setelah 5 tahun?

Jawab:

$$\begin{aligned} P_0 &= 1.000.000 & r &= 3\% = 0,03 & n &= 5 \\ P_n &= 1.000.000 (1+ 0,03)^5 \\ &= 1.000.000 (1,03)^5 \\ &= 1.000.000 (1,159274) \\ &= \text{Rp } 1.159.274 \end{aligned}$$

**Contoh 4.9.** Pada tanggal 1 Januari, seseorang menabung di bank sebesar Rp. 10 jt dengan tingkat suku bunga 2% per bulan. Bila selama tahun itu tabungan tidak diambil, hitunglah jumlah rata-rata uang yang ada di bank selama 5 bulan.

$$\begin{aligned} G &= P_0 \sqrt[5]{(1+0,02) \cdot (1+0,02) \cdot (1+0,02) \cdot (1+0,02) \cdot (1+0,02)} \\ &= 10.000.000 \times 1,002 \\ &= 10.200.000 \end{aligned}$$

Jadi, rata-rata tabungan selama 5 bulan adalah Rp.10.200.000,-

**Contoh 4.10.** Hitunglah pertumbuhan harga saham PT Inti jika perkembangan harga per lembar saham selama minggu terakhir bulan Juni 2012 di Bursa Saham Suralaya sebagai berikut:

Tabel 4.9. Perkembangan Harga Saham PT Inti

Hari	Harga
Senin	Rp. 9.900,-
Selasa	Rp. 10.100,-
Rabu	Rp. 10.200,-
Kamis	Rp. 10.550,-
Jum'at	Rp. 10.800,-
Sabtu	Rp. 11.200,-

Langkah untuk menghitung rasio pertumbuhan saham itu ditampilkan pada tabel di bawah ini :

Tabel 4.10. Rasio Perkembangan Harga Saham PT Inti

Hari	Harga	Rasio
Senin	Rp. 9.900,-	-
Selasa	Rp. 10.100,-	$10.100/9.900 = 1,0201$
Rabu	Rp. 10.200,-	$10.200/10.100 = 1,0099$
Kamis	Rp. 10.550,-	$10.550/10.200 = 1,0343$
Jum'at	Rp. 10.800,-	$10.800/10.550 = 1,0237$
Sabtu	Rp. 11.200,-	$11.200/10.800 = 1,0370$

Rasio pertumbuhannya adalah rasio faktor pertumbuhan dikurang satu. Berdasarkan data pada tabel di atas, misalkan rasio pertumbuhan pada hari Selasa adalah 0,0201 (1,0201-1). Dengan menggunakan perumusan rata-rata hitung, maka rata-rata pertumbuhan harga saham adalah :

$$\begin{aligned}\bar{X} &= \frac{1,0202+1,0099+1,0343+1,0237+1,0370}{5} \\ &= 1,02502\end{aligned}$$

Jika dibuktikan dengan menggunakan rata-rata geometrik, penyelesaian soalnya sebagai berikut :

$$\begin{aligned}G &= \sqrt[5]{1,0202 \times 1,0099 \times 1,0343 \times 1,0237 \times 1,0370} \\ &= 1,025 \text{ (hasil pembulatan)}\end{aligned}$$

**Contoh 4.11.** Pendapatan nasional suatu negara pada tahun 2000 adalah US\$ 400 milyar dan pada tahun 2004 menjadi US\$ 500 milyar. Coba hitung dan analisis berapa rata-rata tingkat kenaikan/ pertumbuhannya?

$$n = 4 \rightarrow \text{tahun 2000-2004}$$

$$x_0 = 400 \rightarrow \text{pendapatan nasional tahun 2000}$$

$$x_n = 500 \rightarrow \text{pendapatan nasional 2004}$$

$$\begin{aligned}G &= [\text{antilog}_{\frac{1}{n}}(\log x_n - \log x_0)] - 1 \\ &= [\text{antilog}_{\frac{1}{4}}(\log 500 - \log 400)] - 1 \\ &= [\text{antilog}_{\frac{1}{4}}(2,69897 - 2,60206)] - 1\end{aligned}$$

$$\begin{aligned}
&= [\text{antilog}_{\frac{1}{4}}(0,09691)] - 1 \\
&= (\text{antilog } 0,02423) - 1 \\
&= 1,05737 - 1 \\
&= 0,05737
\end{aligned}$$

Jadi, rata-rata tingkat pertumbuhan pendapatan nasional selama 4 tahun adalah 0,05737 atau 5,737% per tahun.

### 3. Hubungan Rata-Rata Ukur dengan Bunga Majemuk

Seringkali kita menemukan atau menjumpai data atau nilai yang sudah dikelompokkan. Untuk mencari rata-rata ukur untuk data yang berkelompok adalah sebagai berikut:

$$G = \text{antilog} \left\{ \frac{\sum f_i \log x_i}{\sum f_i} \right\}$$

Dimana :

G = Rata-rata ukur / geometric mean

$f_i$  = Frekuensi tiap kelas

$x_i$  = Nilai tengah tiap kelas

**Contoh 4.12.** Berdasarkan data pada Contoh Soal 4.4 (tentang rata-rata hitung) yakni data penimbangan berat 65 karung beras milik PT Makmur (dalam Kg), sudah diolah dalam tabel tersebut. Berapakah rata-rata ukur berat beras setiap karungnya?

Tabel 4.11. Perhitungan Rata-Rata Ukur Beras PT Makmur

Berat Beras (Kg)	Titik Tengah ( $X_i$ )	Banyaknya Karung ( $f_i$ )	Log $x_i$	f. Log $x_i$
45 – 50	47,5	5	1.676694	8,38347
51 – 56	53,5	7	1.728835	12,101845
57 – 62	59,5	10	1.774517	17,74517
63 – 68	65,5	20	1.816241	36,32482
69 – 74	71,5	12	1.854306	22,251672
75 – 80	77,5	8	1.889302	15,114416
81 – 86	83,5	3	1.921687	5,765061
$\Sigma$		65		117,686454

Sumber : Ilustrasi

$$\begin{aligned}
G &= \text{antilog} \left\{ \frac{\sum f \log x}{\sum f} \right\} \\
&= \text{antilog} (117,686454 / 65) \\
&= \text{antilog} 1,810564 \\
&= 64,65
\end{aligned}$$

Jadi nilai rata-rata ukur setiap kantong berasnya adalah 64,65

### E. RATA-RATA HARMONIK

Nilai rata-rata harmonik H paling sering digunakan untuk merata-ratakan kecepatan jarak tempuh, menentukan harga rata-rata komoditi tertentu, menghitung investasi sejumlah uang tertentu setiap periode dan lain-lain. Rata-rata harmonik bagi nilai  $n$  bilangan positif  $X_1, X_2, \dots, X_n$ , adalah  $n$  dibagi dengan jumlah kebalikan bilangan itu, seperti rumus berikut ini:

$$H = \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}}$$

dimana :

H = rata-rata harmonik

$X_n$  = data ke- $n$

$n$  = jumlah data (sampel)

Pada data yang berbentuk kelompok maka nilai harmonik merupakan hasil bagi jumlah frekuensi dengan jumlah hasil bagi frekuensi dengan nilai  $x_i$ , sehingga dirumuskan sebagai berikut :

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

**Contoh 4.13.** Hitunglah rata-rata beras Rojolele per kg dengan menggunakan rata-rata harmonik. Pada minggu pertama terjual seharga Rp.10.000/kg, minggu kedua terjual seharga Rp.9.000/kg, minggu ketiga terjual seharga Rp.8.000/kg dan minggu keempat terjual seharga Rp. 9.600/kg.

$$\begin{aligned}
 H &= \frac{4}{\left[ \frac{1}{10.000} + \frac{1}{9.000} + \frac{1}{8.000} + \frac{1}{9.800} \right]} \\
 &= \frac{4}{\frac{864 + 960 + 1.080 + 900}{8.640.000}} \\
 &= 9.085
 \end{aligned}$$

Jadi harga rata-rata beras Rojolele adalah Rp. 9.085 per Kg.

**Contoh 4.14.** Nilai 10 mahasiswa yang mengikuti perkuliahan Statistika Jurusan Pendidikan Ekonomi Universitas Mulawarman adalah sebagai berikut: 56, 76, 34, 59, 62, 56, 68, 60, 73, dan 81. Hitunglah rata-rata harmonik dari 10 mahasiswa tersebut.

Rata-rata harmonik diperoleh dengan menggunakan rumus sebagai berikut

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

Langkah berikutnya yakni memasukkan data nilai dari kesepuluh mahasiswa dengan  $n=10$ , dan masing-masing  $x_i$  sebagai berikut:

$$H = \frac{10}{\left( \frac{1}{56} \right) + \left( \frac{1}{76} \right) + \left( \frac{1}{34} \right) + \left( \frac{1}{59} \right) + \left( \frac{1}{62} \right) + \left( \frac{1}{56} \right) + \left( \frac{1}{68} \right) + \left( \frac{1}{60} \right) + \left( \frac{1}{73} \right) + \left( \frac{1}{81} \right)}$$

Masing-masing perhitungan dicari hasilnya agar diperoleh nilai rata-rata harmonis sebagai berikut:

$$\begin{aligned}
 H &= \frac{10}{0,018+0,013+0,029+0,017+0,016+0,018+0,015+0,017+0,014+0,012} \\
 &= \frac{10}{0,169} \\
 &= 59,17
 \end{aligned}$$

Maka nilai rata-rata harmonis dari nilai 10 mahasiswa tersebut adalah 59,17



## F. MEDIAN

Median merupakan salah satu ukuran pemusatan atau sebuah nilai yang berada ditengah-tengah data, setelah data tersebut diurutkan. Mengapa perlu median setelah mempelajari rata-rata hitung? Berdasarkan sifatnya, rata-rata hitung sangat dipengaruhi oleh data ekstrem baik terbesar maupun terkecil, sehingga nilai rata-rata tidak mencerminkan kondisi sebenarnya.

Median merupakan nilai sentral dari distribusi frekuensi, yang terhubung dengan posisi sentral yang dimilikinya dalam sebuah distribusi. Median juga disebut sebagai rata-rata posisi (*positional average*). Secara teoritis, median membagi jumlah observasi atau pengukuran ke dalam 2 bagian yang sama. Median juga terbagi menjadi dua kategori yakni median dari data yang belum dikelompokkan dan telah berbentuk kelompok.

### 1. Median dari Data yang Belum Dikelompokkan

Median untuk data tidak berkelompok adalah nilai yang letaknya ditengah data yang telah diurutkan, namun belum dikelompokkan dalam kelas/kategori tertentu atau belum dalam bentuk distribusi frekuensi. Cara mencari letak dan nilai median untuk data tidak berkelompok terdapat dua kemungkinan. Kemungkinan pertama adalah data itu ganjil dan kemungkinan kedua adalah data itu genap. Bila nilai-nilai observasi  $X_i$  sejumlah  $n$  disusun dari nilai terkecil hingga nilai terbesar sedemikian rupa sehingga  $X_1 \leq X_2 \leq \dots \leq X_n$ , maka median nilai-nilai itu adalah  $X_k$ .

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

$$\text{Nilai Me} = X_{k+1} \text{ dengan } k = \frac{1}{2} (n-1)$$

**Contoh 4.14.** Diketahui data nilai ujian mahasiswa yang telah diuraikan dari nilai terkecil sampai nilai terbesar adalah : 63, 65, 67, 70, 73, 73, 77, 78, 81, 81, 81, 82, 84 dan 88. Berapakah nilai median ?

Langkah untuk menyelesaikan soal di atas adalah mengurutkan data dari terkecil sampai terbesar, yakni :63, 65, 67, 70, 73, 73, 77, 78, 81, 81, 81, 82, 84 , 88,85. Posisi Median =  $\frac{1}{2} (15+1) =8$  dengan  $k = \frac{1}{2} (15-1) = 7$ , jadi nilai median berada pada urutan ke  $(7+1)$  atau 8, yakni 78.

Sementara jika jumlah data pada data mentah atau data yang belum dikelompokkan tersebut adalah genap maka rumus yang digunakan adalah sebagai berikut:

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

$$\text{Nilai Me} = \frac{1}{2} (X_k + X_{k+1}) \text{ dengan } k = \frac{1}{2} n$$

**Contoh 4.15.** Omset penjualan dari 20 cabang Supermarket “Daremantep Jaya” pada Agustus 2012 diperoleh data 6 cabang supermarket yang memiliki omset penjualan tertinggi seperti ditampilkan pada Tabel 4.12. Berapakah nilai media dari omset tersebut?

Tabel 4.12. Omset Penjualan Supermarket Daremantep Jaya

Supermarket	Omset
Daremantep Jaya 7	Rp. 170.000.000
Daremantep Jaya 15	Rp. 179.500.000
Daremantep Jaya 11	Rp. 192.500.000
Daremantep Jaya 20	Rp. 195.500.000
Daremantep Jaya 8	Rp. 215.000.000
Daremantep Jaya 10	Rp. 225.750.000

Letak mediannya adalah  $\frac{1}{2}(6+1)=3,5$  sementara letak titik ini berada diantara data ke-3 dan ke-4 antara Rp.192.500.000 dan Rp.195.500.000. Nilai median bisa diketahui dengan cara mencari rata-rata hitung dari kedua angka itu :

$$\frac{1}{2} (192.500.000+195.500.000)= 194.000.000$$

Sebenarnya seluruh angka pada interval Rp.192.500.000 dan Rp.195.500.000 dapat menjadi median. Namun karena sulit menentukan angka pasti untuk menjadi median, maka median ditentukan dengan mencari rata-rata kedua angka diatas.

## 2. Median dari Data yang Dikelompokkan

Langkah pertama dalam menetapkan median dari data yang telah dikelompokkan adalah menentukan letak sebuah titik yang nilainya akan menjadi median. Titik ini membagi deretan angka yang terurut menjadi dua bagian sama banyak. Jika pada data yang belum diurutkan menggunakan rumus  $(n+1)/2$ , maka data yang dikelompokkan menggunakan rumus  $\frac{1}{2}n$ . Setelah diketahui posisi titik tersebut, langkah berikutnya adalah menentukan kelas yang didalamnya terdapat titik tersebut. Untuk data yang berkelompok, nilai median dapat dicari dengan interpolasi yang rumusnya adalah sebagai berikut:

$$Me = Bm + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\}$$

$Bm$  = Nilai batas bawah dari kelas yang mengandung atau memuat nilai median

$n$  = Banyaknya observasi (jumlah semua frekuensi)

$(\sum f_{km})$  = Jumlah frekuensi dari semua kelas di bawah kelas yang mengandung median (kelas yang mengandung median tidak termasuk)

$f_m$  = Frekuensi dari kelas yang mengandung median

$i$  = Besarnya kelas interval, jarak antara kelas yang satu dengan lainnya atau besarnya kelas interval yang mengandung median

**Contoh 4.16.** Misalkan  $X$  adalah nilai upah tahunan karyawan sebuah perusahaan asing. Berapakah besarnya nilai median upah karyawan, jika terdapat 40 orang karyawan yang sedang diselidiki dan hasil pertanyaan yang diajukan tentang besarnya upah tahunan (dalam ribu dolar) adalah sebagai berikut:

146	147	147	148	149	150	150	152	153	154
156	157	158	161	163	164	165	168	173	176
119	125	126	128	132	135	135	135	136	138
138	140	140	142	142	144	144	145	145	146

Penyelesaian soal di atas dapat dilakukan dengan cara mengurutkan nilai dari terkecil ( $X_1$ ) sampai ke nilai terbesar ( $X_{40}$ ), yakni sebagai berikut :

$$\begin{array}{cccccc}
 X_1 = 119 & X_2 = 125 & X_3 = 126 & X_4 = 128 & X_5 = 132 & \\
 X_6 = 135 & X_7 = 135 & X_8 = 135 & X_9 = 136 & X_{10} = 138 & \\
 X_{11} = 138 & X_{12} = 140 & X_{13} = 140 & X_{14} = 142 & X_{15} = 142 & \\
 X_{16} = 144 & X_{17} = 144 & X_{18} = 145 & X_{19} = 145 & X_{20} = 146 & \\
 X_{21} = 146 & X_{22} = 147 & X_{23} = 147 & X_{24} = 148 & X_{25} = 149 & \\
 X_{26} = 150 & X_{27} = 150 & X_{28} = 152 & X_{29} = 153 & X_{30} = 154 & \\
 X_{31} = 156 & X_{32} = 157 & X_{33} = 158 & X_{34} = 161 & X_{35} = 163 & \\
 X_{36} = 164 & X_{37} = 165 & X_{38} = 168 & X_{39} = 173 & X_{40} = 176 & 
 \end{array}$$

Menentukan nilai  $k$  dengan cara  $40 = 2k \rightarrow k = (40/2) = 20$ , sehingga mediannya adalah sebagai berikut :

$$\begin{aligned}
 Me &= \frac{1}{2} (X_n + X_{n+1}) \\
 &= \frac{1}{2} (X_{20} + X_{20+1}) \\
 &= \frac{1}{2} (146 + 146) \\
 &= 146
 \end{aligned}$$

Upah dari 40 orang karyawan itu disajikan dalam tabel frekuensi sebagai berikut:

Upah (Ribuan \$)	$f$
(1)	(2)
118 - 126	3
127 - 135	5
136 - 144	9
145 - 153	12
154 - 162	5
163 - 171	4
172 - 180	2
Jumlah	40

50% observasi (jumlah frekuensi nilai observasinya sama atau lebih kecil dari median). Kelas mengandung median terletak pada nilai frekuensi  $f_m = 12$ .

Beberapa poin penting yang dapat dijelaskan dari tabel di atas adalah sebagai berikut :

1. Upah dianggap sebagai bilangan yang didistribusikan secara kontinu. Median merupakan upah yang mempunyai ciri/sifat

sedemikian rupa sehingga  $\frac{1}{2}$  atau 50% dari observasi (jumlah frekuensi) yaitu  $(40/2)=20$  observasi, terletak di bawah median dan setengah lainnya di atas median tersebut.

2. Jumlah tiga frekuensi pertama,  $f_1 + f_2 + f_3 = 3 + 5 + 9 = 17$ , observasi belum mencapai 20, atau belum ada setengahnya. Untuk mencapai 20 observasi, diperlukan tiga observasi dari kelas keempat yang frekuensinya  $f_4 = 12$ . Jadi, median terletak dalam kelas keempat.
3. Karena kelas interval keempat, yaitu 145-153, sama dengan (setelah memperhitungkan bahwa upah merupakan data yang kontinu) 144.5-153.5, sehingga median akan terletak di posisi  $(3/12)$  dalam jarak kelas interval 144.5-153.5.
4. Hasil perhitungan sebagai berikut:

- $B_m = 144.5$  (nilai batas kelas bawah dari kelas yang memuat median, setelah diadakan koreksi kontinuitas),  $\frac{1}{2}n = (40/2) = 20$  dan  $(\sum f_i) = f_1 + f_2 + f_3 = 17$ , maka  $f_m = 12$  dan besarnya kelas interval  $(i) = (153.5 - 144.5) = 9$ , jarak antara nilai batas bawah dan nilai batas atas kelas interval yang memuat median atau jarak suatu kelas dengan kelas berikutnya, baik diukur dengan nilai atas bawah atau batas atas.  $145 - 136 = 9$ , atau  $153 - 144 = 9$  (sama saja).
- Cara ini disebut interpolasi, sehingga hasilnya tidak tepat seperti halnya dihitung langsung. Walaupun demikian, hasilnya akan mendekati nilai sebenarnya.
- Di dalam prakteknya, kita tidak mengetahui data aslinya: data itu sudah disajikan dalam tabel frekuensi, dibuat kelas/kelompok/kategori. Jadi, rumus interpolasi ini sangat penting untuk menghitung median. Dengan menggunakan rumus di atas, maka akan dihasilkan angka berikut:

$$Me = B_m + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\} = 144.5 + 9 \left\{ \frac{\frac{40}{2} - 17}{12} \right\} = 146.75$$

## G. MODUS

Modus sebagai ukuran pusat data, berbeda dengan rata-rata hitung penentuannya. Modus mirip dengan median dalam penentuannya tidak melalui proses aritmatik seperti penentuan rata-rata. Modus adalah nilai yang paling sering muncul dalam data. Modus sering disingkat Mo. Sejumlah data bisa tidak punya modus, satu modus (*Unimodal*), dua modus (*Bimodal*), atau lebih dari dua modus (*Multimodal*). Cara mencari modus dibedakan data tunggal dan data berkelompok.

### A. Modus Data Tidak Berkelompok

Modus dari data yang tidak berkelompok adalah nilai yang mempunyai frekuensi tertinggi, atau nilai yang paling banyak terjadi di dalam suatu distribusi nilai.

**Contoh 4.17.** Omset penjualan yang diperoleh 6 supermarket “Daremantep Jaya”, pada bulan Januari 2013 adalah seperti pada Tabel 4.13., tentukan nilai modulusnya!

Tabel 4.13. Omset 6 Supermarket Daremantep Jaya

Supermarket	Omset
Daremantep Jaya 1	Rp. 90.000.000
Daremantep Jaya 2	Rp. 95.000.000
Daremantep Jaya 3	Rp. 100.000.000
Daremantep Jaya 4	Rp. 100.000.000
Daremantep Jaya 5	Rp. 105.000.000
Daremantep Jaya 6	Rp. 110.750.000

Modus dari data diatas dapat ditentukan dengan mudah yakni dengan mudah yaitu Rp.100.000.000 dikarenakan nilai omset tersebut memiliki frekuensi lebih tinggi dibandingkan dengan nilai omset yang lainnya.

### B. Modus Data Berkelompok

Modus atau mode yang telah dikelompokkan diperkirakan berada pada kelas yang memiliki frekuensi tinggi. Kendatipun hal

tersebut hanya bersifat estimatif dan berbeda sifat estimatifnya dengan rata-rata dan median. Apabila data sudah dikelompokkan dalam tabel frekuensi, maka dalam mencari modulusnya harus dipergunakan rumus berikut ini :

$$Mod = B_m + i \left\{ \frac{(d_1)}{(d_1) + (d_2)} \right\}$$

dimana :

$B_m$  = nilai batas bawah, kelas yang memuat modulus

$f_{m0}$  = frekuensi kelas yang memuat modulus

$(d_1)$  =  $f_{m0} - f_{(m0 - 1)}$  {selisih frekuensi kelas yang memuat modulus dengan frekuensi kelas sebelumnya (bawahnya)}

$(d_2)$  =  $f_{m0} - f_{(m0 + 1)}$  {selisih frekuensi kelas yang memuat modulus dengan frekuensi kelas sesudahnya (atasnya)}

$i$  = besarnya jarak antara nilai batas atas dan nilai batas bawah dari kelas yang memuat modulus.

**Contoh 4.18.** PT Abadi Jaya melakukan pengelompokan cabang perusahaan berdasarkan omset penjualan sebagaimana ditampilkan pada Tabel 4.14. Tentukan nilai modulus dari omset perusahaan-perusahaan tersebut !

Tabel 4.14. Omset Penjualan PT Abadi Jaya

Interval Omset Penjualan (Rp juta)	Jumlah Perusahaan
200 – 219	7
220 – 239	9
240 – 259	11
260 – 279	18
280 – 299	12
300 – 319	5

Kelas yang berisi modulus,  $f_{m0} = 18$



Diketahui :

$$B_m = \frac{1}{2} (259 + 260) = 259.5;$$

$$i \text{ (interval)} = 220 - 200 = 20$$

$$f_{(m0-1)} = 11; f_{(m0+1)} = 12; (d_1) = 18 - 11 = 7; (d_2) = 18 - 12 = 6$$

$$\begin{aligned}
 \text{Mod} &= B_m + i \left\{ \frac{(d_1)}{(d_1) + (d_2)} \right\} \\
 &= 259.5 + 20 \left( \frac{7}{7+6} \right) \\
 &= 270.27
 \end{aligned}$$

Jadi, nilai modus dari omset perusahaan-perusahaantersebut adalah Rp.270,27 juta atau Rp.270.270.000,-

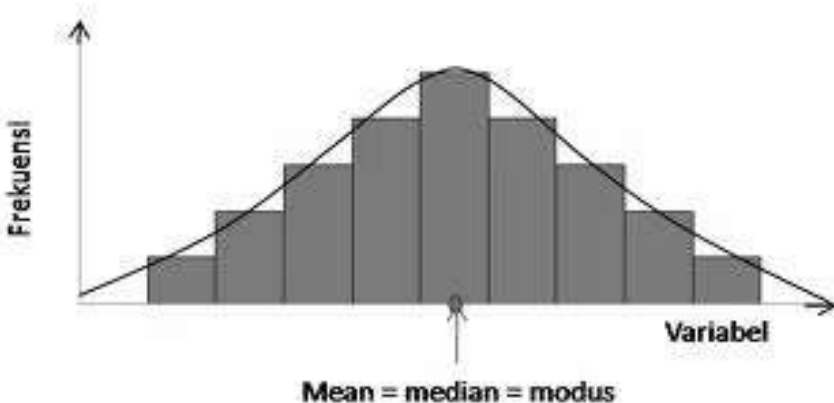
## H. HUBUNGAN RATA-RATA, MEDIAN DAN MODUS

Rata-rata hitung, median dan modus adalah tiga ukuran pemusatan yang digunakan untuk mengetahui bentuk kurva poligon dari data baik berupa sampel maupun populasi. Bentuk kurva poligon bisa berupa kurva normal atau simetris, condong ke kiri (*skewed negatif*) atau condong ke kanan (*skewed positif*). Hubungan antara rata-rata hitung, median dan modus dijelaskan sebagai berikut:

### a) Kurva Simetris

Kurva simetris adalah kurva dimana sisi kanan dan kiri sama, sehingga kalau dilipat dari titik tengahnya maka terdapat dua bagian yang sama. Untuk kurva simetris, nilai rata-rata hitung, median dan modus mempunyai nilai sama.

$$\bar{X} = \text{Med} = \text{Mod}$$



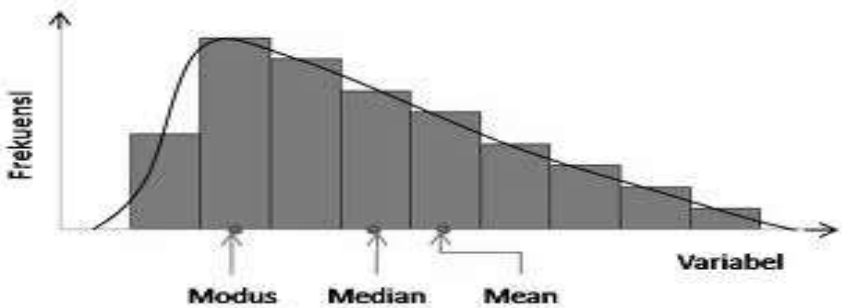
Gambar 4.1. Kurva Simetris



b) Kurva Condong Kiri

Kurva condong kiri atau condong positif disebabkan nilai rata-rata hitung lebih besar dibandingkan dengan median dan modus. Hal tersebut terjadi karena adanya nilai ekstrem tinggi yang mempengaruhi nilai rata-rata hitung, sedangkan median dan modus tidak terpengaruhi. Pada kejadian seperti ini data pada umumnya bernilai rendah, tetapi ada beberapa yang ekstrem bernilai sangat tinggi.

$$\bar{X} > Me, Mod$$

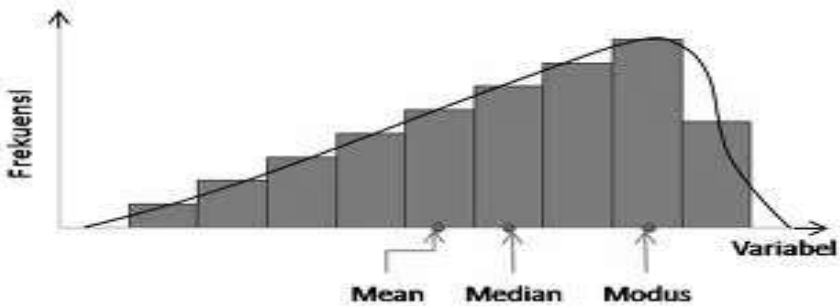


Gambar 4.2. Kurva Condong Kiri

c) Kurva Condong Kanan

Kurva condong ke kanan atau condong negatif disebabkan nilai rata-rata hitung lebih kecil daripada nilai median dan modus. Peristiwa ini karena adanya nilai ekstrem rendah yang mempengaruhi nilai rata-rata hitung.

$$\bar{X} < Me, Mod$$



Gambar 4.3. Kurva Condong Kanan

Hubungan antara nilai rata-rata, median dan modus akan bersifat empiris apabila ketiganya mempunyai nilai hampir sama. Sebagaimana yang dikemukakan oleh Karl Pearson dalam Dajan (2008) yang menyatakan bahwa bila distribusi dari variabel yang kontinu memiliki modus tunggal serta menceng secara moderat, maka mediannya akan terletak kira-kira  $2/3$  dari seluruh yang dihitung dari modus ke arah rata-rata hitungannya.

Sehingga dapat dirumuskan sebagai berikut:

$$\text{Mod} = \bar{X} - 3 (\bar{X} - \text{Me})$$

**Contoh 4.19.** Berdasarkan tabel distribusi frekuensi berat badan 100 orang mahasiswa Fakultas Ekonomi UM tahun 2012, tentukan rata-rata, median dan modusnya.

Tabel 4.15. Berat Badan Mahasiswa

Berat Badan	Frekuensi (f)
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8

Penyelesaian soal di atas dapat diawali dengan mencari nilai rata-rata seperti ditampilkan pada tabel di bawah ini:

Tabel 4.16. Nilai Frekuensi Berat Badan Mahasiswa

Berat Badan	$X_i$	Frekuensi (f)	$X_i \cdot f_i$
60-62	61	5	305
63-65	64	18	1.152
66-68	67	42	2.814
69-71	70	27	1.890
72-74	73	8	584
Jumlah		$\sum f = 100$	$\sum X_i \cdot f_i = 6.745$

$$\bar{X} = \frac{\sum_{i=1}^n f_i X_i}{\sum_{i=1}^n f_i} = \frac{6.745}{100} = 67.45$$

$$\text{Median (Me)} = Bm + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\} = 65.5 + 3 \left\{ \frac{50 - 23}{42} \right\} = 65.64$$

$$\begin{aligned}
\text{Modus (Mod)} &= \bar{X} - 3(\bar{X} - Me) \\
&= 67,45 - 3(67,45 - 65,64) \\
&= 62,02
\end{aligned}$$

Sehingga disimpulkan bahwa nilai , “Mean>Median, Modus”

## I. UKURAN LETAK (KUARTIL, DESIL DAN PERSENTIL)

Jika sekelompok data dibagi menjadi dua bagian yang sama, maka nilai yang berada di tengah (50%) disebut dengan median. Konsep median dapat diperluas yaitu kelompok data yang telah diurutkan (membesar atau mengecil) dibagi menjadi empat bagian sama banyak. Ketiga ukuran di atas (mean, median dan modus) merupakan ukuran lokasi yang cenderung bertindak sebagai ukuran pusat data, maka ketiga ukuran ini hanya merupakan ukuran lokasi. Kendati bukan sebagai ukuran pusat data, ukuran ini banyak bermanfaat bagi pengambil keputusan.

Pada akhir sub-bab disajikan contoh penggunaan ukuran ini. Tiga ukuran tersebut adalah kuartil, desil dan persentil.

### 1. Kuartil Data Tidak Berkelompok

Apabila kita bicara tentang median, maka nilai ini seolah-olah membagi kelompok data menjadi 2 bagian yang sama. Artinya, 50% dari kelompok data ini (seluruh nilai observasi) mempunyai nilai sama atau lebih kecil dari median, sedangkan 50% lainnya mempunyai nilai yang sama atau lebih besar dari median tersebut. Ingat bahwa nilai median merupakan salah satu dari nilai observasi/pengamatan.

$\frac{50\%}{(X \leq \text{med})} \text{med} \frac{50\%}{(X \geq \text{med})}$
--

Untuk kelompok data di mana  $n \geq 4$ , kita tentukan tiga nilai, misalkan  $Q_1, Q_2, Q_3$ , yang membagi kelompok data tersebut menjadi 4 bagian yang sama, yaitu setiap bagian memuat data

yang sama atau jumlah observasinya sama. Nilai-nilai tersebut dinamakan kuartil pertama, kuartil kedua, dan kuartil ketiga. Pembagian itu adalah sedemikian rupa sehingga nilai 25% data/observasi sama atau lebih kecil dari  $Q_1$ , 50% data/observasi sama atau lebih kecil dari  $Q_2$ , 75% data/observasi sama atau lebih kecil dari  $Q_3$ . Jika suatu kelompok data atau nilai sudah diurutkan dari yang terkecil ( $X_n$ ), maka untuk menghitung  $Q_1$ ,  $Q_2$ , dan  $Q_3$ , harus dipergunakan rumus.

**Contoh 4.20.** Berikut ini adalah data upah bulanan dari 13 karyawan (dalam ribu rupiah), yaitu 40, 30, 50, 65, 45, 55, 70, 60, 80, 35, 85, 95, 100, ( $n = 13$ ). Cari nilai  $Q_1$ ,  $Q_2$ , dan  $Q_3$ .

Untuk menyelesaikan soal di atas, pertama-tama data diurutkan dahulu:  $X_1 = 30$ ,  $X_2 = 35$ ,  $X_3 = 40$ ,  $X_4 = 45$ ,  $X_5 = 50$ ,  $X_6 = 55$ ,  $X_7 = 60$ ,  $X_8 = 65$ ,  $X_9 = 70$ ,  $X_{10} = 80$ ,  $X_{11} = 85$ ,  $X_{12} = 95$ ,  $X_{13} = 100$ .

$$Q_1 \text{ adalah nilai ke } \frac{i(n+1)}{4} = \frac{1(13+1)}{4} = 3\frac{1}{2}$$

nilai ke- $3\frac{1}{2}$  adalah rata-rata dari  $X_3$  dan  $X_4$

$$\begin{aligned} Q_1 &= \frac{1}{2} (X_3 + X_4) \\ &= \frac{1}{2} (40 + 45) \\ &= 42.5 \end{aligned}$$

$$\begin{aligned} Q_2 &= \text{nilai ke } - \frac{2(13+1)}{4} \\ &= \text{nilai ke } -7, \text{ jadi nilai } Q_2 = 60 \end{aligned}$$

$$\begin{aligned} Q_3 &= \frac{1}{2} (X_{10} + X_n) \\ &= \frac{1}{2} (80 + 85) \\ &= 82.5 \end{aligned}$$

(Nilai kuartil tidak perlu sesuai nilai data asli).

## 2. Kuartil Data Berkelompok

Jika dalam menentukan titik letak median sederetan data terurut dibagi menjadi dua, maka kuartil membagi sederetan data

terurut menjadi empat bagian yang sama. Dengan demikian, nantinya akan terdapat tiga kuartil yaitu kuartil pertama ( $Q_1$ ), kuartil kedua atau median, dan kuartil ketiga ( $Q_3$ ). Titik lokasi ketiga kuartil (untuk data yang telah dikelompokkan) tersebut secara sederhana dapat dirumuskan sebagai berikut:

$$Q_1 = n/4, Q_2 = 2n/4 = n/2 = \text{medi}, Q_3 = 3n/4, Q_k = kn/4$$

Selanjutnya, dengan memperhatikan rumus di atas, kuartil pertama dan kuartil ketiga (kuartil kedua sama dengan median) dapat dirumuskan sebagai berikut:

$$Q_1 = B_q + i \left[ \frac{n/4 - f_{kq}}{f_q} \right], \quad Q_3 = B_q + i \left[ \frac{3n/4 - f_{kq}}{f_q} \right],$$

dimana:

- $Q_1$  = Kuartil pertama
- $Q_3$  = Kuartil ketiga
- $B_q$  = Tepi batas kelas bawah pada kelas kuartil
- $i$  = Interval kelas
- $n$  = Ukuran sampel
- $f_{kq}$  = Frekuensi kumulatif sebelum kelas kuartil
- $f_q$  = Frekuensi pada kelas kuartil

**Contoh 4.21.** Berdasarkan data pengamatan dari 100 petani tembakau terhadap harga tembakau per kg, didapatkan seperti tabel berikut, hitunglah nilai  $Q_1$ , dan  $Q_3$ .

Tabel 4.17. Distribusi Harga Tembakau (ribu/kg)

Nilai Kelas	$f$
72,2 – 72,4	2
72,5 – 72,7	5
72,8 – 73,0	10
73,1 – 73,3	13
73,4 – 73,6	27
73,7 – 73,9	23
74,0 – 74,2	16
74,3 – 74,5	4
Jumlah	$\sum f_i = 100$

- Untuk menghitung  $Q_1$ :  $f_1 + f_2 + f_3 = 17$  belum mencapai 25% (25). Agar bisa mencapai jumlah frekuensi 25, maka frekuensi kelas yang ke-4 harus ikut dijumlahkan, sehingga diketahui kelas ke-4 memuat  $Q_1$ .
- dari data,  $(\sum f_i) = 17$ ;  $n = 100$ ;  $f_q = 13$ . Nilai batas bawah dan batas atas dari kelas yang memuat  $Q_1$ , masing-masing adalah  $\frac{1}{2}(73,0 + 73,10) = 73,05$  dan  $\frac{1}{2}(73,3 + 73,4)$ . Jadi,  $i = 73,35 - 73,05 = 0,30$ .
- Setelah didapatkan nilai setiap komponen, maka dimasukkan kedalam rumus, sehingga hasilnya sebagai berikut:

$$Q_1 = B_q + i \cdot \left[ \frac{n/4 - f_{kq}}{f_q} \right], = 73,05 + 0,30 \left\{ \frac{100 - 17}{13} \right\} = 73,23$$

- Untuk menghitung  $Q_3$ :  $f_1 + f_2 + f_3 + f_4 + f_5 = 2 + 5 + 10 + 13 + 27 = 57$ , belum mencapai angka 75% (=75), masih kurang  $(75 - 57) = 18$ . Kekurangan ini ditambah dengan frekuensi kelas ke-6, sehingga kelas ke-6 memuat  $Q_3$ . Dari data,  $(\sum f_i) = 57$ ;  $n = 100$ , dan  $f_q = 23$ .
- Nilai batas bawah dan batas atas dari kelas yang memuat  $Q_3$ , masing-masing adalah  $\frac{1}{2}(73,6 + 73,7) = 73,65$  dan  $\frac{1}{2}(73,9 + 74,0) = 73,95$ ; jadi  $B_q = 73,65$ ;  $i = 73,95 - 73,65 = 0,30$ .

$$Q_3 = B_q + i \cdot \left[ \frac{3n/4 - f_{kq}}{f_q} \right], = 73,65 + 0,30 \left\{ \frac{300 - 57}{23} \right\} = 73,89$$

### 3. Desil Data Tidak Berkelompok

Jika sekelompok data dibagi menjadi sepuluh bagian yang sama, artinya setiap bagian mempunyai jumlah observasi yang sama, sedemikian rupa sehingga nilai 10% observasi sama atau lebih kecil dari  $D_1$ , nilai 20% observasi sama atau lebih kecil dari  $D_2$ , dan seterusnya. Nilai tersebut disebut desil atau disingkat  $D$  yaitu  $D_1, D_2, D_3, \dots, D_9$ . Nilai desil ke- $k$  yaitu  $D_k$  dapat ditentukan dengan rumus sebagai berikut :

$$D_k = \text{nilai yang ke } - \left[ \frac{k(n+1)}{10} \right], \text{ dimana } k = 1, 2, \dots, 9$$

**Contoh 4.22.** Tentukan Desil  $D_3$  dan  $D_7$  dari data upah harian 13 karyawan (ribu rupiah) berikut ini :40, 30, 50, 65, 45, 55, 70, 60, 80, 35, 85, 95, 100.

Soal di atas bisa diselesaikan dengan mengurutkan data menjadi: 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 95, 100.

$$D_k = \text{nilai yang ke } - \frac{k(n+1)}{10}, k=1, 2, \dots, 9$$

$$D_3 = \text{nilai yang ke } - \frac{3(13+1)}{10}, \text{ nilai ke } 4,2$$

$$D_3 = \text{Nilai yang ke } - 4 + (0,2) \times (\text{nilai ke-5} - \text{nilai ke-4})$$

$$D_3 = 45 + (0,2) \times (50 - 45) = 46$$

$$D_7 = \text{nilai yang ke } - \frac{7(13+1)}{10}, \text{ nilai ke } 9,8$$

$$D_7 = \text{Nilai yang ke } - 9 + (0,8) \times (\text{nilai ke-10} - \text{nilai ke-9})$$

$$D_7 = 70 + (0,8) \times (80 - 70) = 78$$

#### 4. Desil Data Berkelompok

Jika pada kuartil deretan data terurut dibagi menjadi 4, maka pada desil, deretan data terurut dibagi menjadi 10 bagian. Perumusan yang digunakan tidak jauh berbeda, kecuali hanya bagian rumus yang menentukan titik-titik desil. Berikut tabel yang memuat bagian rumus menentukan sembilan titik.

Tabel 4.18. Titik-titik Letak Desil

Desil ke-1	:	$n/10$
Desil ke-2	:	$2n/10$
Desil ke-3	:	$3n/10$
Desil ke-4	:	$4n/10$
Desil ke-5	:	$5n/10$
Desil ke-6	:	$6n/10$
Desil ke-7	:	$7n/10$
Desil ke-8	:	$8n/10$
Desil ke-9	:	$9n/10$

← **Median**

Adapun bagian lain menyesuaikan titik desil bersangkutan, yang dapat dirumuskan sebagai berikut:

$$D_k = B_{dk} + i \left[ \frac{\frac{k}{10}n - F_{kd}}{f_{dk}} \right]$$

dimana:

$D_k$  = Desil ke- $k$

$B_{dk}$  = Batas bawah kelas yang mengandung  $D_k$

$i$  = Interval kelas

$F_{kd}$  = Jumlah frekuensi sebelum kelas yang mengandung  $D_k$

$f_{dk}$  = Frekuensi kelas yang mengandung  $D_k$

**Contoh 4.23.** Tentukan Desil ke-3 dan ke-7 dari modal (dalam juta rupiah) dari 40 perusahaan yang disajikan tabel distribusi frekuensi berikut ini :

Tabel 4.19. Modal 40 Perusahaan

Modal	Nilai Tengah (X)	Frekuensi ( f )	f . X
112 - 120	116	4	464
121 - 129	125	5	652
130 - 138	134	8	1072
139 - 147	143	12	1716
148 - 156	152	5	760
157 - 165	161	4	644
166 - 174	170	2	340
		$\Sigma f = 40$	$\Sigma f . X = 5621$

Soal di atas termasuk dalam Desil Data Berkelompok, sehingga menggunakan rumus dibawah ini :

$$D_k = B_{dk} + i \left[ \frac{\frac{k}{10}n - F_{kd}}{f_{dk}} \right]$$

$D_3$  membagi data 30% ke bawah dan 70% ke atas sehingga berada pada kelas 130-138, sedangkan  $D_7$  membagi data 70% ke bawah dan 30% ke atas sehingga berada pada pada kelas 139-147.

$$D_3 = 130 + 9 \left[ \frac{\frac{3}{10}40 - 9}{8} \right] = 130 + 9 \left[ \frac{12 - 9}{8} \right] = 133.375$$

$$D_7 = 139 + 9 \left[ \frac{\frac{7}{10}40 - 17}{12} \right] = 139 + 9 \left[ \frac{28 - 17}{12} \right] = 147.25$$



## 5. Persentil Data Tidak Berkelompok

Jika sekelompok data dibagi menjadi 100 (seratus) bagian yang sama banyaknya, maka terdapat 99 pembagi yang masing masing disebut *persentil (P)* yaitu  $P_1, P_2, P_3, \dots, P_{99}$ . Nilai *Persentil* ke- $i$ , yaitu  $P_i$  dihitung dengan rumus :

$$P_k = \text{nilai yang ke-} \frac{k(n-1)}{100}, \text{ dimana } k = 1, 2, 3, \dots, 99.$$

**Contoh 4.24.** Tentukan letak  $P_{20}$  serta nilainya dari data berikut ini: 35, 40, 70, 80, 91, 50, 61, 25, 95.

Soal di atas dapat diselesaikan dengan mengurutkan data dari terkecil hingga terbesar : 15, 35, 40, 50, 61, 70, 80, 91, 95.

Letak persentil ke 20 ( $P_{20}$ ) adalah  $\frac{20(9+1)}{100} = 2$

Jadi persentil ke 20 terletak pada data ke-2, yaitu 35.

## 6. Persentil Data Berkelompok

Rumus yang berubah hanya bagian yang menentukan letak titik persentil dan bagian lainnya yang menyesuaikan persentil yang dimaksud. Berikut tabel yang memuat bagian rumus yang menentukan titik persentil.

Tabel 4.20. Letak Beberapa Titik Persentil

Persentil ke-1	$n/100$
Persentil ke-12	$12n/100$
Persentil ke-27	$27n/100$
Persentil ke-87	$87n/100$
Persentil ke-99	$99n/100$

$$P_k = B_{pk} + i \left[ \frac{\frac{k}{100}n - F_{kp}}{f_{pk}} \right]$$

**Contoh 4.25.** Perusahaan Baldroc bergerak dalam penjualan bahan bangunan mempekerjakan 50 tenaga penjual (salesman)

yang beroperasi dari rumah ke rumah. Selama semester pertama tahun 2012, total penjualan masing-masing tenaga penjualan dapat disajikan sebagai berikut:

Tabel 4.21. Total Penjualan 50 Tenaga Penjual di Perusahaan Baldroc selama Semester I Tahun 2012

Nilai Penjualan (juta rupiah)	Tenaga Penjual
Rp. 100,- < Rp. 150,-	4
Rp. 150,- < Rp. 200,-	9
Rp. 200,- < Rp. 250,-	11
Rp. 250,- < Rp. 300,-	15
Rp. 300,- < Rp. 350,-	7
Rp. 350,- < Rp. 400,-	4
Jumlah	50

Pimpinan Perusahaan Baldroc menetapkan bahwa tenaga penjual yang bisa mencapai penjualan Rp. 275.000.000,- atau lebih akan menerima bonus sebesar 10% dari nilai penjualan. Perkiraan jumlah tenaga penjual yang bisa menerima bonus tersebut!

Penyelesaian soal di atas dapat diselesaikan dengan terlebih dahulu menentukan nilai penjualan sebesar Rp. 275.000.000,- sebagai ukuran lokasi. Dalam kasus di atas, ukuran lokasi yang dapat digunakan adalah persentil.

$$P_k = B_{pk} + i \left[ \frac{\frac{k}{100}n - F_{kp}}{f_{pk}} \right]$$

$$275 = 250 + 50 \left[ \frac{50n - 24}{15} \right]$$

$$275 = 250 + 166.666,67n - 80$$

$$n = 0,63$$

keterangan :

275 = Persentil ke- ...

250 = Tepi batas kelas bawah pada kelas persentil

50 = Interval kelas

50 = Ukuran sampel

n = Rasio persentil ke- ...

24 = Frekuensi kumulatif sebelum kelas persentil

15 = Frekuensi pada kelas persentil

n sebesar 0,63 ini merupakan rasio jumlah tenaga penjual yang nilai penjualannya kurang dari Rp. 275.000,00 atau yang tidak mendapatkan bonus. Sehingga tenaga penjual yang mendapatkan bonus sebesar 0,37 atau sebanyak 18,5 atau 19 orang.

**Contoh 4.26.** Tentukan persentil ke 75 dari modal (dalam juta rupiah) dari 40 perusahaan yang disajikan dalam tabel distribusi frekuensi sebagaimana disajikan pada Tabel 4.19.

Letak  $P_{75} = \frac{75 \times 40}{100} = 30$ , sehingga termasuk kelas ke-5 (148-156) yang didapatkan dari jumlah frekuensi tiap kelas sampai mencapai 30 yaitu ( $4 + 5 + 8 + 12 + 5 = 34$ )

Untuk  $P_{75}$ , maka  $B_{pk} = 147,5$ ;  $F_{kp} = 4 + 5 + 8 + 12 = 29$ ;  $i = 9$  dan  $f_{pk} = 5$ , sehingga diperoleh nilai sebagai berikut:

$$P_k = 147,5 + 9 \left[ \frac{\frac{75}{100} \cdot 40 - 29}{5} \right] = 149,3$$

## J. RINGKASAN

- 1) Rata-rata (*average*) adalah nilai yang mewakili sekelompok data (*a set of data*). Nilai rata-rata umumnya terletak di tengah data-data yang disusun menurut besar atau kecilnya nilai.
- 2) Nilai rata-rata mempunyai kecenderungan memusat, sehingga disebut ukuran kecenderungan memusat (*measures of central tendency*).
- 3) Jenis rata-rata yang dipergunakan dalam ukuran pemusatan (*measures of central tendency*) ialah rata-rata hitung (*arithmetic mean* atau disingkat *mean*), rata-rata ukur (*geometric mean*), dan rata-rata harmonis (*harmonic mean*).
- 4) Rata-rata hitung (rata-rata) merupakan suatu bilangan tunggal yang dipergunakan untuk mewakili nilai sentral dari sebuah distribusi. Rumus rata-rata hitung adalah sebagai berikut :

Data Tidak Berkelompok		Data Berkelompok	
Sampel	Populasi	Sampel	Populasi
$\bar{x} = \frac{\sum X_n}{n}$	$\mu = \frac{\sum x_i}{N}$	$\bar{X} = \frac{\sum f_i x_i}{n}$	$\mu = \frac{\sum f_i x_i}{N}$

5) Median didefinisikan sebagai titik tengah dari semua nilai data yang telah diurutkan dari nilai terkecil ke nilai terbesar, atau sebaliknya dari nilai terbesar ke nilai terkecil. Rumus median adalah sebagai berikut :

Data Tidak Berkelompok		Data Berkelompok	
Sampel	Populasi	Sampel	Populasi
Letak Median $\frac{1}{2} (n+1)$	Letak Median $\frac{1}{2} (n+1)$	Letak Median $\frac{1}{2}n$	Letak Median = $\frac{1}{2}N$
Median = $\frac{1}{2} (X_k+X_{k+1})$	Median = $\frac{1}{2} (X_k+X_{k+1})$	Median = $Bm + i \left\{ \frac{\frac{N}{2} - (\sum f_{km})}{f_m} \right\}$	Median = $Bm + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\}$

- 6) Modus dari suatu kelompok adalah nilai kelompok itu yang mempunyai frekuensi tertinggi, atau nilai paling banyak terjadi di suatu kelompok nilai. Untuk selanjutnya kita singkat Mod.
- 7) Konsep median dapat diperluas yaitu kelompok data yang telah diurutkan (membesar/mengecil) dibagi menjadi empat bagian sama banyak, nilai itu dinamakan kuartil pertama, kuartil kedua, dan kuartil ketiga.
- 8) Jika pada kuartil deretan data terurut dibagi menjadi 4, maka pada desil, deretan data terurut dibagi menjadi 10 bagian yang sama dan jika sekelompok data dibagi menjadi 100 (seratus) bagian yang sama banyaknya, maka akan terdapat 99 pembagi yang masing masing disebut *persentil (P)* yaitu  $P_1, P_2, P_3, \dots, P_{99}$ .

## K. SOAL DAN LATIHAN

- 1) Apa yang dimaksud dengan rata-rata hitung, rata-rata ukur (*geometric mean*), dan rata-rata harmonis (*harmonic mean*).
- 2) Jelaskan hubungan antara mean, median dan modus dalam distribusi frekuensi.

- 3) PT Sasa Jaya mempunyai waralaba Pisang Gapit di 10 kota di Pulau Jawa. Hitunglah nilai rata-rata hitung, median dan modus dari data pendapatan bersih dari setiap cabang pada tahun 2010 disajikan sebagai berikut:

Cabang	Pendapatan (Rp juta)
Jakarta	80
Serang	10
Tangerang	50
Malang	40
Semarang	40
Jogyakarta	50
Surabaya	90
Bandung	40
Jember	20
Solo	50

- 4) Tentukan Kuartil ( $Q_i$ ), Desil ( $D_i$ ), dan Persentil ( $D_i$ ) dari data upah bulanan 13 karyawan (dalam ribuan rupiah) berikut ini 40, 30, 50, 65, 45, 55, 70, 60, 80, 35, 85, 95, 100.
- 5) Tentukan Kuartil ( $Q_i$ ), Desil ( $D_i$ ) dan Persentil ( $D_i$ ) dari modal 40 perusahaan (dalam juta rupiah) yang disajikan dalam tabel distribusi frekuensi berikut ini :

Modal	Nilai Tengah (X)	Frekuensi (f)	f . X
112 - 120	116	4	464
121 - 129	125	5	652
130 - 138	134	8	1072
139 - 147	143	12	1716
148 - 156	152	5	760
157 - 165	161	4	644
166 - 174	170	2	340
		$\sum f = 40$	$\sum f . X = 5621$

## Bab 5.

# UKURAN PENYEBARAN

### KOMPETENSI DASAR

Mahasiswa dapat menjelaskan konsep dan menghitung ukuran penyebaran data meliputi jangkauan, inter-kuartil deviasi rata-rata, deviasi standar dan koefisien variasi.

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- Menjelaskan tentang pengertian ukuran penyebaran.
- Menjelaskan jangkauan, inter-kuartil dan deviasi kuartil.
- Menghitung nilai jangkauan, inter-kuartil dan deviasi kuartil.
- Menjelaskan deviasi rata-rata, varians dan deviasi standar.
- Menghitung deviasi rata-rata, varians dan deviasi standar.
- Menjelaskan tentang koefisien variasi.
- Menghitung nilai koefisien variasi.

### DESKRIPSI SINGKAT

Bab ini mengantarkan untuk memahami mengenai variabilitas, menguraikan tentang pengertian variabilitas, macam ukuran variabilitas (range, deviasi rata-rata, deviasi standar dan nilai yang dibakukan)

### KATA KUNCI

Ukuran Penyebaran  
Inter-Kuartil  
Deviasi Rata-Rata  
Deviasi Standar

Range (Jangkauan)  
Deviasi Kuartil  
Varians  
Nilai yang Dibakukan

## A. PENGANTAR

Pada bab sebelumnya telah dibahas karakteristik ukuran pusat data (ukuran gejala pusat). Bahasan itu telah menjelaskan sifat-sifat sekumpulan data. Ada ukuran lain yakni variabilitas data (dispersi/penyebaran data). Variabilitas data yang akan dipelajari antara lain rentang, rata-rata simpangan, simpangan baku, rentang kuartil, koefisien variasi, varians dan ukuran kecondongan dan keruncingan.

Pada bab ini akan dibahas ukuran-ukuran tersebut diatas, namun untuk lebih memperjelas terhadap apa yang dimaksud dengan ukuran penyebaran perlu dipahami definisi dari ukuran penyebaran tersebut.

Variabilitas sangat penting bagi penggambaran serangkaian data, terutama jika seseorang ingin membandingkan dua atau lebih rangkaian data. Dalam usaha membandingkan rangkaian data, penggunaan ukuran pusat saja tidak akan menghasilkan yang baik, bahkan memberikan hasil yang menyesatkan. Karena mengetahui nilai tengah saja belum cukup, tanpa diketahui seberapa besar data tersebut menyebar disekitar nilai tengahnya. Dengan memahami unsur penyebaran data diharapkan kita tidak menarik kesimpulan yang salah.

Berikut contoh yang dapat memperjelas pentingnya ukuran penyebaran. Misalnya ketika Anda ingin belajar berenang, Anda menanyakan berapa kedalaman rata-rata kolam renang. Misalkan kedalaman kolam tersebut adalah 1,5 meter, jika tinggi Anda adalah 1,65 m, mungkin Anda berpikir tidak akan tenggelam. Namun yang perlu diingat rata-rata tersebut termasuk mulai dari kolam anak-anak yang kedalamannya hanya 30 cm sampai dengan orang dewasa yang kedalamannya hingga 3 meter. Apabila tidak berhati-hati dan belajar berenang, maka Anda bisa tenggelam.

Contoh lain produksi roti yang dihasilkan oleh perusahaan rata-rata tiap hari rusak 10%. Pada suatu hari dihasilkan roti sebanyak 1000 bungkus. Akankah kita memperoleh 100 roti yang rusak ? jika diantara 1000 bungkus roti itu terdapat 250 yang rusak, bagaimana kesimpulan kita mengenai proses dan keadaan mesin. Dapatkah kita mengambil kesimpulan bahwa pegawainya ceroboh atau mesin yang rusak?

Contoh di atas membuat kita lebih memahami pentingnya ukuran penyebaran. Apabila kumpulan data tidak mempunyai

nilai ekstrem rendah dan tinggi, maka ukuran pemusatan relatif baik untuk digunakan. Namun jika datanya mengandung nilai ekstrem yang tinggi, maka perlu untuk mengetahui sebarannya.

## B. JANGKAUAN, INTER-KUARTIL DAN DEVIASI KUARTIL

### 1. Jangkauan atau Range

Jangkauan atau kisaran nilai (range) merupakan ukuran yang paling sederhana dari ukuran penyebaran. Jangkauan merupakan perbedaan antara nilai terbesar dan terkecil dalam suatu kelompok data baik data populasi maupun data sampel. Semakin kecil ukuran jangkauan menunjukkan karakter yang lebih baik, karena berarti mendekati nilai pusat dan kompak.

**Jangkauan (range) = Nilai Terbesar – Nilai Terkecil**

**Contoh 5.1.** Berikut adalah volume ekspor komoditi pertanian tahun 2003-2007 dari Afrika, Amerika, Asia Pasifik, Eropa dan Timur Tengah. Hitunglah jangkauan (range) nya! Bagaimana komentar Anda?

Tabel 5.1. Volume Ekspor Komoditi Pertanian Menurut Wilayah

No	Wilayah	Volume (000 ton)				
		2003	2004	2005	2006	2007
1	Afrika	647,0	640,3	909,3	1.115,5	1.004,1
2	Amerika	1.157,9	1.553,4	1.831,0	2.357,0	2.575,6
3	Asia Pasifik	3.492,7	4.800,2	5.266,2	7.431,8	10.485,0
4	Eropa	1.279,4	1.704,7	2.202,9	2.677,6	3.622,7
5	Timur Tengah	90,9	96,2	241,9	243,8	389

Sumber : BPS, diolah Direktorat Pemasaran Internasional

Penyelesaian Soal:

Tabel 5.2. Nilai Range Volume Ekspor Komoditi Pertanian

Nilai	Afrika	Amerika	Asia Pasifik	Eropa	Timur Tengah
Tertinggi	1.115,5	2.575,6	10.485,0	3.622,7	389
Terendah	640,5	1.157,9	3.492,7	1.279,4	90,9
Range	475,00	1.417,70	6.992,30	2.343,30	298,10



Nilai range terkecil adalah Timur Tengah diikuti oleh Afrika. Besarnya range menunjukkan selisih nilai terbesar dengan nilai terkecil sehingga menunjukkan juga fluktuasi. Sehingga dapat disimpulkan Asia Pasifik mempunyai nilai range volume terbesar, maka fluktuasi hasil panen di wilayah tersebut juga mempunyai perubahan yang besar.

**Contoh 5.2.** Berikut data penjualan per hari dari salesman CV Dharma Jaya di Malang dan Surabaya. Tentukan jangkauan (range) dari nilai penjualan dua kota tersebut.

Tabel 5.3. Data Penjualan 6 Salesman CV Dharma Jaya

Tenaga Penjual	Surabaya	Malang
David	900.000,00	1.600.000,00
Eliza	1.100.000,00	1.400.000,00
Farah	2.200.000,00	1.500.000,00
Galih	1.400.000,00	1.500.000,00
Handoyo	1.600.000,00	1.700.000,00
Indah	1.800.000,00	1.300.000,00

Penyelesaian Soal:

Jangkauan (range) penjualan CV Dharma Jaya di Surabaya dan Malang adalah:

Surabaya :  $\text{Rp } 2.200.000,00 - \text{Rp } 900.000,00 = \text{Rp } 1.300.000,00$

Malang :  $\text{Rp } 1.700.000,00 - \text{Rp } 1.300.000,00 = \text{Rp } 400.000,00$

Berdasarkan data range diketahui bahwa penjualan di Surabaya memiliki variabilitas lebih tinggi dibanding di Malang.

Jangkauan (range) merupakan alat pengukuran variabilitas yang sederhana sehingga alat ukur ini memiliki kelemahan yakni tidak melibatkan seluruh data. Beberapa catatan pengukuran dan penggunaan jangkauan diantaranya:

a) Pengukuran jangkauan dalam pengawasan kualitas

Hasil pengukuran jangkauan sudah menggambarkan dispersi (variasi) nilai observasi dengan cara paling sederhana. Bila kita

ingin memperoleh hasil pengukuran dispersi secara kasar dan cepat, pengukuran jangkauan dapat digunakan. Pengukuran range banyak digunakan untuk pengawasan kualitas karena kesederhanaan pengukurannya.

b) Evaluasi hasil Pengukuran Jangkauan

Jangkauan bukan pengukuran dispersi yang memuaskan karena pengukurannya tergantung kepada dua nilai, yaitu nilai tertinggi dan nilai terendah, yang merupakan nilai-nilai ekstrim dalam distribusi. Sehingga, range akan mempunyai fluktuasi sangat besar, tergantung kepada nilai-nilai ekstrim. Kelemahan lain dari *range* sebagai alat pengukuran variabilitas adalah karena *range* tidak memenuhi definisi menjadi alat semacam itu. Seperti diketahui sebelumnya bahwa variabilitas menunjukkan penyebaran nilai-nilai sekitar tendensi sentral, dan dalam *range* tidak jelas petunjuk dimana letak tendensi sentralnya. *Range* tidak menunjukkan bentuk distribusi.

## 2. Inter-Kuartil

Ukuran Inter-kuartil dihitung dengan menentukan beda antara kuartil ketiga dan kuartil pertama yang dirumuskan sebagai berikut :

$$\mathbf{IQR = Q_3 - Q_1, Q_1 = \text{kuartil pertama, } Q_3 = \text{kuartil ketiga}}$$

Seperti halnya dengan cara perhitungan jangkauan, maka perhitungan inter-kuartil tidak melibatkan seluruh data.

**Contoh 5.3.** Berdasarkan data pada Contoh 5.2, tentukan nilai inter-kuartilnya !

### **Surabaya**

$$\text{Letak } Q_1 = \frac{n+1}{4} = \frac{6+1}{4} = \frac{7}{4} = 1,75$$

$$\begin{aligned} \text{Nilai } Q_1 &= 900.000,00 + 0,75(1.100.000,00 - 900.000,00) \\ &= 1.050.000,00 \end{aligned}$$

$$\text{Letak } Q_3 = \frac{3(n+1)}{4} = \frac{21}{4} = 5,25$$

$$\begin{aligned}\text{Nilai } Q_3 &= 1.800.000,00 + 0.25(2.200.000,00 - 1.800.000,00) \\ &= 1.900.000,00\end{aligned}$$

$$\text{Inter-kuartil} = 1.900.000,00 - 1.050.000,00 = 850.000,00$$

### **Malang**

$$\text{Letak } Q_1 = \frac{n+1}{4} = \frac{6+1}{4} = \frac{7}{4} = 1,75$$

$$\begin{aligned}\text{Nilai } Q_1 &= 1.300.000,00 + 0.75(1.400.000,00 - 1.300.000,00) \\ &= 1.375.000,00\end{aligned}$$

$$\text{Letak } Q_3 = \frac{3(n+1)}{4} = \frac{21}{4} = 5,25$$

$$\begin{aligned}\text{Nilai } Q_3 &= 1.600.000,00 + 0.25(1.700.000,00 - 1.600.000,00) \\ &= 1.625.000,00\end{aligned}$$

$$\text{Inter-kuartil} = 1.625.000,00 - 1.375.000,00 = 250.000,00$$

Berdasarkan nilai inter-kuartil, penjualan kota Surabaya memiliki variabilitas yang lebih tinggi dibanding di Malang.

### **3. Inter-Kuartil**

Deviasi kuartil mengukur variabilitas dengan menentukan rata-rata hitung inter-kuartilnya, dengan rumus sebagai berikut:

$$\text{Deviasi Kuartil (QD)} = (Q_3 - Q_1) / 2 ;$$

$$Q_1 = \text{kuartil pertama, } Q_3 = \text{kuartil ketiga}$$

**Contoh 5.4.** Dari data Contoh 5.2 tentukan Deviasi Kuartil!

#### **Surabaya**

$$\text{IQR} = 1.900.000,00 - 1.050.000,00 = 850.000,00$$

$$\text{QD} = \frac{1.900.000,00 - 1.050.000,00}{2} = \frac{850.000,00}{2} = 425.000,00$$

#### **Malang**

$$\text{IQR} = 1.625.000,00 - 1.375.000,00 = 250.000,00$$

$$\text{QD} = \frac{1.625.000,00 - 1.375.000,00}{2} = \frac{250.000,00}{2} = 125.000,00$$

Berdasarkan deviasi kuartil, nilai penjualan di Surabaya menunjukkan variabilitas yang lebih tinggi.

### C. DEVIASI RATA-RATA (RATA-RATA SIMPANGAN)

Mean Deviasi (*Average Deviation*) atau Deviasi Mean dari deviasi nilai-nilai Mean dalam suatu distribusi, diambil nilainya yang absolut. Dalam hal ini, deviasi absolut adalah nilai-nilai positif. Secara aritmetik, mean deviasi dapat didefinisikan sebagai mean harga mutlak dari deviasi nilai-nilai individual. Selanjutnya untuk menyelesaikan pekerjaan mencari mean deviasi, haruslah ditemukan mean terlebih dahulu. Kemudian ditentukan besarnya penyimpangan tiap-tiap dari nilai mean. Dalam statistika, deviasi diberi simbol huruf-huruf kecil seperti  $x$ ,  $y$ ,  $z$ ,  $d$ , dan sebagainya.

#### 1. Deviasi Rata-Rata Data Belum Dikelompokkan

Berbeda dengan tiga cara sebelumnya, deviasi rata-rata melibatkan seluruh data dalam penghitungan. Variabilitas diukur dengan membandingkan data observasi secara individual dengan pusat data (biasanya rata-rata). Perhitungan dilakukan dengan mencari rata-rata beda absolut antara data observasi dengan pusat datanya. Apabila tersedia data  $X_1, X_2, \dots, X_i, \dots, X_n$ , dan rata-rata  $\bar{X} = \frac{1}{n} \sum X_i$ , maka simpangan rata-rata hitung adalah:

$$(X_1 - \bar{X}), (X_2 - \bar{X}), \dots, (X_i - \bar{X}), \dots, (X_n - \bar{X}).$$

Deviasi rata-rata (*MD*) atau rata-rata simpangan adalah rata-rata hitung dari nilai absolut simpangan dengan rumus :

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

**MD = Deviasi Rata-Rata;  $X_i$  = data ke-i;  $\bar{X}$  = Rata-Rata**

Perhatikan bahwa hasil pengurangan data observasi dengan rata-rata berada pada dua tanda garis tegak, yang menunjukkan bahwa hasil pengurangan tersebut berbentuk absolut /mutlak (senantiasa positif).

**Contoh 5.5.** Berdasarkan data pada Contoh 5.2, hitunglah nilai deviasi rata-rata nilai penjualan pada dua kota tersebut ?

**Rata-rata penjualan di Surabaya**

$$\begin{aligned}\bar{X} &= \frac{900.000 + 1.100.000 + 2.200.000 + 1.400.000 + 1.600.000 + 1.800.000}{6} \\ &= 1.500.000\end{aligned}$$

**Rata-rata penjualan di Malang**

$$\begin{aligned}\bar{X} &= \frac{1.600.000 + 1.400.000 + 1.500.000 + 1.500.000 + 1.700.000 + 1.300.000}{6} \\ &= 1.500.000\end{aligned}$$

Perhitungan Deviasi Rata-rata (MD) adalah sebagai berikut

Tabel 5.4. Perhitungan MD Nilai Penjualan di Surabaya

$X_i$	$\bar{X}$	$ X_i - \bar{X} $
900.000,00	1.500.000,00	600.000,00
1.100.000,00	1.500.000,00	400.000,00
2.200.000,00	1.500.000,00	700.000,00
1.400.000,00	1.500.000,00	100.000,00
1.600.000,00	1.500.000,00	100.000,00
1.800.000,00	1.500.000,00	300.000,00
Jumlah		2.200.000,00

$$\text{Deviasi rata-rata (MD)} = \frac{2.200.000}{6} = 366.667$$

Tabel 5.5 Perhitungan MD Nilai Penjualan di Malang

$X_i$	$\bar{X}$	$ X_i - \bar{X} $
1.600.000,00	1.500.000,00	100.000,00
1.400.000,00	1.500.000,00	100.000,00
1.500.000,00	1.500.000,00	0
1.500.000,00	1.500.000,00	0
1.700.000,00	1.500.000,00	200.000,00
1.300.000,00	1.500.000,00	200.000,00
Jumlah		600.000,00

$$\text{Deviasi rata-rata (MD)} = \frac{600.000}{6} = 100.000$$

Jika deviasi rata-rata dihitung berdasarkan simpangan terhadap median, maka diartikan  $(X_1 - \text{med})$ ,  $(X_2 - \text{med})$ , ...  $(X_n - \text{med})$ , sehingga simpangan terhadap median dirumuskan:

$$\mathbf{MD = \frac{1}{n} \Sigma | X_i - \text{med} |}$$

MD = Mediasi Rata-Rata;  $X_i$  = Data ke-i;  
Med = Nilai Median; n = Ukuran Data

**Contoh 5.6.** Cari rata-rata simpangan, baik terhadap rata-rata hitung maupun median dari data  $X_1 = 30, X_2 = 40, X_3 = 50, X_4 = 60, X_5 = 70$

$$\bar{X} = \frac{1}{5}(50 + 40 + 30 + 60 + 70) = 50, \text{ med} = 50$$

$$\begin{aligned} \text{MD} &= \frac{1}{5} \sum |X_i - \bar{X}| = \frac{1}{5} |0| + |-10| + |-20| + |10| + |20| = 12 \\ &= \frac{1}{5} \sum |X_i - \text{med}| = 12 \end{aligned}$$

## 2. Deviasi Rata-Rata Data Telah Dikelompokkan

Seperti halnya ketika menentukan ukuran pusat data telah dikelompokkan, diperlukan penaksir data asli dari kelas-kelas data yang terdapat dalam sebuah distribusi frekuensi, yaitu titik-titik tengah masing-masing kelas. Bila nilai-nilai observasi sudah dikelompokkan ke dalam bentuk distribusi frekuensi, maka deviasi rata-ratanya dirumuskan sebagai berikut:

**Untuk n (sampel)**  
**Deviasi rata-rata (MD) =  $\frac{\sum |x_i - \bar{X}| \cdot f_i}{n}$**

dimana :

$X_i$  : Titik tengah kelas ke- $I$        $\bar{X}$  : Rata-rata sampel

$f_i$  : Frekuensi kelas ke- $I$        $n$  : Ukuran sampel

**Contoh 5.7.** Berikut data yang sudah dikelompokkan dari harga saham pilihan pada bulan Juni 2007 di BEJ. Hitunglah *Deviasi rata-rata* dari data tersebut!

Tabel 5.6. Data Harga Saham di BEJ

Kelas ke-	Interval	Jumlah Frekuensi (F)
1	160 – 303	2
2	304 – 447	5
3	448 – 591	9
4	592 – 735	3
5	736 – 878	1

Penyelesaian soal di atas dapat dilakukan dengan langkah-langkah sebagai berikut :

- Untuk memperoleh nilai tengah data ( $X$ ), nilai tengah kelas dikalikan dengan frekuensi masing-masing kelas ( $fX$ ). Jumlah perkalian antara frekuensi dengan nilai tengah kelas dibagi dengan jumlah data mendapatkan nilai rata-rata hitung data berkelompok ( $\sum fX/n$ ).
- Langkah kedua menghitung deviasi setiap kelas dengan cara mengurangkan nilai tengah kelas dengan rata-rata hitungnya ( $|X_i - \bar{X}|$ ).
- Langkah ketiga mengalikan frekuensi dengan deviasi setiap kelas  $f(|X_i - \bar{X}|$ ).
- Menjumlahkan hasil perkalian frekuensi dengan deviasi setiap kelas kemudian membaginya dengan jumlah data ( $n$ ).

Langkah-langkah di atas disajikan ringkas dalam bentuk tabel sebagai berikut :

Tabel 5.7. Deviasi Harga Saham di BEJ

Interval	Titik tengah ( $X_i$ )	$f$	$f.X_i$	$X_i - \bar{X}$	$f X_i - \bar{X} $
160 – 303	231,5	2	463,	-259,2	518,4
304 – 447	375,5	5	1.877,5	-115,2	576,0
448 – 591	519,5	9	4.677,5	28,8	259,2
592 – 735	663,5	3	1.990,0	172,8	518,4
736 – 878	807,0	1	807,0	316,3	316,3
		20	9.813,5		2.118,3

$$\sum f.X_i = 9.813,5 \text{ dan } \sum f |X - \bar{X}| = 2.118,3$$

$$\bar{X} = (\sum f.X_i) / n = 9.813,5 / 20 = 490,7$$

$$MD = (\sum f |X - \bar{X}|) / n = 2.118,3 / 20 = 109,4$$

**Contoh 5.8.** Upah 50 orang karyawan PT ABADI (Ribu Rp/hari), datanya telah diolah dalam Tabel 5.7. Hitunglah *Deviasi rata-rata* dari data tersebut!

Tabel 5.8. Perhitungan Deviasi Rata-Rata Nilai Upah Karyawan

Upah	$X_i$	$f_i$	$fixX_i$	$ X_i - \bar{X} $	$ X_i - \bar{X}  \cdot f_i$
130-139	134,5	4	538	30,6	122,4
140-149	144,5	6	867	20,6	123,6
150-159	154,5	8	1236	10,6	84,8
160-169	164,5	12	1974	0,6	7,2
170-179	174,5	9	1570,5	9,4	84,6
180-189	184,5	7	1291,5	19,4	135,8
190-199	194,5	4	778	29,4	117,6
		50	8255	120,6	676

$$\bar{X} = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i} = \frac{8255}{50} = 165,1$$

$$MD = \frac{\sum |X_i - \bar{X}| \cdot f_i}{n} = \frac{676}{50} = 13,2$$

Hasil penghitungan variabilitas menggunakan deviasi rata-rata ini tentu saja lebih baik daripada menggunakan jangkauan (*range*), inter-kuartil, dan deviasi-kuartil, karena penghitungan deviasi melibatkan seluruh data observasi. Akan tetapi deviasi rata-rata masih memiliki kelemahan, untuk memperoleh nilai-nilai beda data observasi dengan rata-rata yang positif, metode ini menganggap sama antara nilai-nilai negatif dan positif. Secara matematik, kedua sifat bilangan tersebut baik positif maupun negatif harus dibedakan dengan tegas.

Keunggulan mean deviasi terhadap pengukuran variabilitas dengan *range* tersebut adalah dipenuhinya definisi tentang variabilitas oleh mean deviasi itu, yaitu penyebaran nilai-nilai yang ditinjau dari tendensi sentral. Akan tetapi mean deviasi mempunyai satu kelemahan pokok, karena cara perhitungannya mengabaikan tanda-tanda plus dan minus. Oleh karena itu mean deviasi tidak dapat dikenai perhitungan-perhitungan matematik yang tetap mempertahankan nilai-nilai plus dan minus. Untuk mengatasi kelemahan itu, maka timbullah cara pengukuran variabilitas lain, yaitu "standard deviasi".



## D. PENGUKURAN VARIANS DAN SIMPANGAN BAKU

### 1. Varians Data Belum Dikelompokkan

Pengertian varians mirip dengan deviasi rata-rata. Hanya saja, untuk memperoleh hasil perhitungan dalam bilangan positif tidak diwujudkan dalam bilangan absolut, namun dikuadratkan, atau dengan kata lain varians adalah alat ukur variabilitas serangkaian data yang dihitung dengan mencari rata-rata selisih/ beda kuadrat antara data observasi dengan pusat datanya. Varians adalah rata-rata hitung dari kuadrat simpangan setia pengamatan terhadap rata-rata hitungnya, di mana  $(X_i - \mu)$  adalah simpangan (deviasi) dari observasi terhadap rata-rata sampel. Apabila kita mempunyai suatu populasi dengan jumlah elemen sebanyak  $N$  dan sampel dengan  $n$  elemen, dan selanjutnya nilai karakteristik tertentu kita kumpulkan (umur, hasil penjualan, harga barang, produksi barang, nilai ujian), maka kita akan memperoleh sekumpulan nilai observasi sebagai berikut:

**Populasi:  $X_1, X_2, \dots, X_N$**

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

**= rata-rata sebenarnya dari  $X$  (rata-rata populasi)**

Seperti pada rata-rata, dalam varians pun ada yang disebut sebagai varians populasi dan varians sampel. Simbol dari varians populasi adalah  $\sigma^2$  (dibaca sigma kuadrat) merupakan varians sebenarnya dari  $X$ , dengan rumus sebagai berikut :

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

$(X_i - \mu)$  = deviasi dari observasi terhadap rata-rata sebenarnya

Rumus varians populasi itu dapat disederhanakan menjadi seperti berikut :

$$\sigma^2 = \frac{N \sum X_i^2 - (\sum X_i)^2}{N^2}$$

Sedangkan varians sampel ( $S^2$ ) menurut Karl Pearson dirumuskan sebagai berikut:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Bagi distribusi sampel dengan  $n < 100$ , Fisher, Wilks dan beberapa statistisi memberikan perumusan varians sebagai berikut:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Rumus varians sampel itu dapat disederhanakan menjadi sebagai berikut :

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

**Contoh 5.9.** Berdasarkan Tabel 5.3. yang disajikan pada Contoh 5.2. terdapat data penjualan oleh 6 Salesman CV Dharma Jaya. Tentukan variasi nilai penjualan di Surabaya dan Malang.

Tabel 5.9. Variasi Penjualan 6 Salesman CV Dharma Jaya

$X_i$	$\bar{X}$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
900.000	1.500.000	- 600.000,00	360.000.000.000
1.100.000	1.500.000	400.000,00	160.000.000.000
2.200.000	1.500.000	700.000,00	490.000.000.000
1.400.000	1.500.000	-100.000,00	10.000.000.000
1.600.000	1.500.000	100.000,00	10.000.000.000
1.800.000	1.500.000	300.000,00	90.000.000.000
			1.120.000.000.000

Dengan menggunakan rumus :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

maka varians data sampel di atas adalah :

$$\begin{aligned}
 S^2 &= \frac{1}{6-1} (1.120.000.000.000) \\
 &= 224.000.000.000
 \end{aligned}$$

## 2. Varians Data Telah Dikelompokkan

Bila varians dan deviasi standar dihitung dari distribusi frekuensi, maka titik tengah tiap-tiap kelas umumnya dianggap sebagai nilai tunggal yang cukup representatif bagi semua nilai-nilai observasi  $X_i$  yang dikelompokkan ke dalam kelas-kelas yang bersangkutan. Simbol varians populasi adalah  $\sigma^2$  (dibaca sigma kuadrat) yang merupakan varians sebenarnya dari  $X$ . Rumus varians dari distribusi frekuensi dapat disajikan sebagai berikut :

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

Seringkali angka yang dihadapi tergolong besar sehingga dapat menyulitkan dalam proses penghitungan. Untuk rumus itu bisa disederhanakan menggunakan pengkodean sebagai berikut :

$$\sigma^2 = i^2 \left[ \frac{\sum u_i^2 \cdot f_i - \frac{(\sum u_i \cdot f_i)^2}{N}}{N} \right]$$

dimana:

- |            |                           |       |                        |
|------------|---------------------------|-------|------------------------|
| $\sigma^2$ | = Varians Populasi,       | $i$   | = Interval kelas       |
| $u_i$      | = Kode u pada kelas ke-I, | $f_i$ | = Frekuensi kelas ke-i |
| $N$        | = Ukuran Populasi         |       |                        |

Sedangkan varians sampel ( $S^2$ ) dapat menggunakan rumus sebagai berikut:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \cdot f_i$$

atau dirumuskan dengan menggunakan cara pengkodean sebagai berikut:

$$S^2 = i^2 \left[ \frac{\sum c_i^2 \cdot f_i - \frac{(\sum c_i \cdot f_i)^2}{N}}{n-1} \right]$$

**Contoh 5.10.** Hitunglah varians dari data berkelompok seperti yang tertuang dalam Contoh 5.7.

Tabel 5.10. Perhitungan Nilai Varians

Titik tengah (X)	fi	f.X	Xi - $\bar{X}$	(Xi - $\bar{X}$ ) <sup>2</sup>	fi. (Xi - $\bar{X}$ ) <sup>2</sup>
231,5	2	463,0	-259,2	67.185	134.369
375,5	5	1.877,5	-115,2	13.271	66.355
519,5	9	4.677,5	28,8	829	7.465
663,5	3	1.990,0	172,8	29.860	89.580
807,0	1	807,0	316,3	100.046	100.046
	20	9.813,5			397.815

$$\bar{X} = \frac{9.813,5}{20} = 490,7$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{397.815}{20-1} = \frac{397.815}{19} = 20.938$$

### 3. Deviasi Standar Data Belum Dikelompokkan

Pada praktiknya, ukuran variabilitas yang sering digunakan adalah simpangan baku atau deviasi standar. Simpangan baku merupakan salah satu ukuran dispersi yang diperoleh dari akar kuadrat positif varians. Di antara ukuran dispersi atau variasi, simpangan baku adalah yang paling banyak dipergunakan, sebab mempunyai sifat-sifat matematis (*mathematical property*) yang sangat penting dan berguna sekali untuk pembahasan teori dan analisis. Rumus dan simbol dari deviasi standar populasi adalah:

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

$\sigma = \text{deviasi standar atau simpangan baku dari } X$

Pada prinsipnya simpangan baku adalah akar dari varians. Pada prakteknya, pengumpulan data didasarkan pada sampel tidak menghasilkan varians (simpangan baku) sebenarnya, tetapi hanya suatu perkiraan saja dengan rumus sebagai berikut:

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

$S = \text{simpangan baku sampel}$

Bisa ditunjukkan secara statistik matematis bahwa jika pembagiannya (penyebutnya)  $n - 1$ ,  $(S^2) = \sigma^2$ , artinya  $S^2$  “unbiased estimator” dari  $\sigma^2$ , sehingga dalam prakteknya, rumus-rumus di atas banyak digunakan.

**Contoh 5.11.** Berdasarkan Contoh 5.9., tentukan simpangan baku nilai penjualan CV. Dharma Jaya di Surabaya.

Nilai Varians Contoh 5.8 adalah:  $S^2 = \frac{1}{6-1} (1.120.000.000.000) = 224.000.000.000$ , dari data itu maka simpangan baku atau deviasi standarnya adalah:  $S = \sqrt{224.000.000.000} = 473.286,3826$

#### 4. Deviasi Standar Data Telah Dikelompokkan

Untuk data yang berkelompok dan sudah disajikan dalam tabel frekuensi, rumus simpangan baku populasi adalah sebagai berikut:

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \mu)^2}{N}}$$

$X_i = \text{nilai tengah dari kelas ke-}i, i = 1, 2, \dots, k.$

atau

$$\sigma = \sqrt{\frac{1}{N} \left\{ \sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{N} \right\}}$$

pada distribusi frekuensi dengan kelas interval yang tidak sama dapat menggunakan rumus dengan metode pengkodean sebagai berikut :

$$\sigma = i \sqrt{\frac{\sum f_i c_i^2}{N} - \left( \frac{\sum f_i c_i}{N} \right)^2}$$

*untuk kelas interval yang sama*

untuk data sampel diperoleh simpangan baku sampel dengan rumus sebagai berikut :

$$S = i \sqrt{\frac{\sum f_i c_i^2}{n-1} - \left(\frac{\sum f_i c_i}{n-1}\right)^2}$$

untuk kelas interval yang sama

$$S = i \sqrt{\frac{1}{n-1} \left\{ \sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{n-1} \right\}}$$

untuk kelas interval yang tidak sama

**Contoh 5.12.** Hitunglah simpangan baku dari modal 40 perusahaan (dalam juta rupiah) dalam data dalam distribusi frekuensi yang telah dikelompokkan dalam tabel sebagai berikut:

Tabel 5.11. Perhitungan Simpangan Modal 40 Perusahaan

Modal (M)	Nilai Tengah	<i>f</i>
118 – 126	122	3
127 – 135	131	5
136 – 144	140	9
145 – 153	149	12
154 – 162	158	5
163 – 171	167	4
172 – 180	176	2
Jumlah		40

Untuk data berkelompok, harus diperhatikan jangkauan kelas satu dengan kelas berikutnya harus sama, atau selisih nilai tengah satu dengan nilai tengah lainnya harus sama, yaitu sebesar (131- 122) = (140 – 131) = ... = 9, jadi *i* = 9. Titik asal asumsi = *X<sub>i</sub>* = 149, yaitu kelas 145 – 153. Dengan demikian, nilai simpangan (deviasi) dari setiap nilai tengah terhadap titik asal asumsi tersebut dapat diperoleh sebagai berikut:

Tabel 5.12. Penolong Perhitungan Simpangan

Kelas	<i>f</i>	<i>c<sub>i</sub></i>	<i>c<sub>i</sub><sup>2</sup></i>	<i>fci</i>	<i>fci<sup>2</sup></i>
118 – 126	3	-3	9	-9	27
127 – 135	5	-2	4	-10	20
136 – 144	9	-1	1	-9	9
145 – 153	12	0	0	0	0
154 – 162	5	1	1	5	5
163 – 171	4	2	4	8	16
172 – 180	2	3	9	6	18
Jumlah	40	0	28	$\sum f_i c_i = -9$	$\sum f_i c_i^2 = 95$

$$\sigma = i \sqrt{\frac{\sum f_i c_i^2}{N} - \left(\frac{\sum f_i c_i}{N}\right)^2}, \sigma = 9 \sqrt{\frac{95}{40} - \left(\frac{-9}{40}\right)^2} = 13,72$$

Perhitungan ini merupakan pendekatan (aproksimasi) dan nilainya tidak sama dengan hasil perhitungan langsung data asli yang masih merupakan data mentah). Titik asumsi ditentukan secara sembarang, tetapi lebih baik kalau dipilih kelas dengan nilai frekuensi terbesar.

**Contoh 5.13.** Berdasarkan data yang sudah dikelompokkan dari contoh di atas, hitunglah simpangan baku.

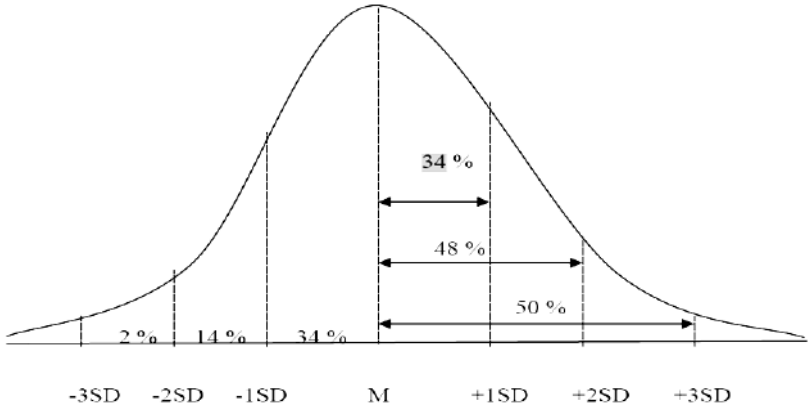
Tabel 5.13. Penolong Perhitungan Simpangan Baku

$X_i$	$X_i^2$	$f$	$fX_i$	$fX_i^2$
122	14.884	3	366	44.652
131	17.161	5	655	85.805
140	19.600	9	1.26	176.400
149	22.201	12	1.788	266.400
158	24.964	5	790	124.820
167	27.889	4	668	111.556
176	30.976	2	352	61.952
Jumlah		$\sum f_i = 40$	$\sum f_i X_i = 5.879$	$\sum f_i X_i^2 = 871.597$

$$\begin{aligned} \sigma &= \sqrt{\frac{1}{N} \left\{ \sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{N} \right\}} \\ &= \sqrt{\frac{1}{40} \left\{ 871.597 - \frac{(5.879)^2}{40} \right\}} \\ &= 13,72 \end{aligned}$$

Pada umumnya, setiap hasil pengumpulan data (melalui pengukuran/pengamatan langsung seperti observasi, wawancara, dan sebagainya) menghasilkan suatu kelompok data, misalnya  $X_1, X_2, \dots, X_i, \dots, X_N$ , di mana masing-masing nilai akan berbeda satu sama lain. Seringkali kita ingin mengetahui berapa selisih simpangan (deviasi) dari masing-masing nilai tersebut (misalnya  $X_i$ , observasi atau nilai ke- $i$ ) terhadap rata-rata hitungnya (yaitu terhadap  $\mu = \frac{1}{N} \sum X_i$ ).

Simpangan (deviasi) dari  $X_i$  terhadap  $\mu = (X_i - \mu)$ , diukur dengan simpangan baku  $\sigma$ . Jadi simpangan baku merupakan satuan ukuran (*unit of measurement*) dari simpangan atau deviasi. Seperti halnya kg, ton untuk ukuran berat: cm, m, km untuk mengukur panjang, maka  $\sigma$  = simpangan baku digunakan untuk mengukur simpangan atau deviasi masing-masing nilai individu dari suatu kelompok data terhadap rata-rata hitungnya. Standard Deviasi (SD) membagi range menjadi beberapa bagian yang sama lebarnya, pembagian mana dimulai pertama-tama dari mean distribusi, membentang ke atas dan ke bawah dengan tanda-tanda plus dan minus. Di bawah ini diberikan gambaran arti standard deviasi yang diterapkan pada salah satu bentuk distribusi, yaitu distribusi normal (*normal distribution*).



Gambar 5.1 Distribusi Normal

Apabila distribusi berbentuk normal, maka banyaknya individu yang mendapatkan nilai dari M sampai + 1 SD kira-kira ada 34 %; dari M sampai +2 SD ada 48%; dan dari M sampai +3 SD ada 50%. Demikian pula antara M sampai -1 SD = 34%; antara M sampai -2SD = 48%; dan antara M sampai -3SD = 50%. Persentase tersebut adalah persentase pembulatan. Presisinya adalah sebagai berikut:



Dari M sampai 1 SD = 34,12%

Dari M sampai 2 SD = 47,72%

Dari M sampai 3 SD = 49,87%

Dari -1 SD sampai +1 SD = 68%

Dari -2 SD sampai +2 SD = 96%

Apabila kita hitung sebelah-menyebelah mean hasilnya adalah sebagai berikut:

Dari -1 SD sampai +1 SD = 68%

Dari -2 SD sampai +2 SD = 96%

Dari -3 SD sampai +3 SD = 100%

### **E. NILAI YANG DIBAKUKAN (STANDARDIZED VALUE)**

Standar deviasi sering digunakan untuk menentukan skor baku. Skor baku adalah suatu ukuran relatif untuk menyatakan penyimpangan data dari nilai rata-rata yang diukur berdasarkan nilai standar deviasinya. Skor baku digunakan untuk menghitung luas kurva normal baku dan membandingkan data pengamatan dari dua atau lebih populasi atau sampel yang berbeda dalam menentukan tingkatan atau rangking relatifnya.

Variabel X mempunyai rata-rata  $\mu$  dengan simpangan baku  $\sigma$ . Jadi,  $\frac{X_i}{\sigma}$  merupakan nilai baku dari  $X_i$ , sehingga nilai skor baku (*standardized*) dirumuskan sebagai berikut:

$$Z_i = \frac{X_i - \mu}{\sigma} \text{ (populasi) atau } Z_i = \frac{X_i - \bar{X}}{s} \text{ (sampel)}$$

**Contoh 5.14.** Dari data  $X_1 = 2, X_2 = 8, X_3 = 10, X_4 = 4, X_5 = 1, N = 5, Z_i = \frac{X_i - \mu}{\sigma}$  dan  $i = 1, 2, \dots, 5$ . Hitunglah nilai-nilai  $\mu, \sigma, \frac{X}{\sigma}, Z_i, \mu_z$  dan  $\sigma_z$ .

Penyelesaian soal di atas dapat dilakukan dengan langkah-langkah sebagai berikut :

$$\mu = \frac{1}{N} \sum X_i = \frac{1}{N} (X_1 + X_2 + \dots + X_5) = \frac{1}{5} (2 + 8 + \dots + 1) = \frac{25}{5} = 5$$

$$\sigma = \sqrt{\frac{1}{N} \sum (X_i - \mu)^2} = \sqrt{\frac{1}{5} (2 - 5)^2 + (8 - 5)^2 + \dots + (1 - 5)^2} = 3,5$$

$$\frac{X_1}{\sigma} = \frac{2}{3,5} = 0,57;$$

$$\frac{X_2}{\sigma} = \frac{8}{3,5} = 2,29;$$

$$\frac{X_3}{\sigma} = \frac{10}{3,5} = 2,86;$$

$$\frac{X_4}{\sigma} = \frac{4}{3,5} = 1,14;$$

$$\frac{X_5}{\sigma} = \frac{1}{3,5} = 0,29;$$

$$Z_1 = \frac{2 - 5}{3,5} = -0,86;$$

$$Z_2 = \frac{8 - 5}{3,5} = 0,86;$$

$$Z_3 = \frac{10 - 5}{3,5} = 1,43;$$

$$Z_4 = \frac{4 - 5}{3,5} = -0,29;$$

$$Z_5 = \frac{1 - 5}{3,5} = -1,14;$$

$$\mu_z = \frac{1}{N} \sum Z_i = \frac{1}{5} (Z_1 + Z_2 + Z_3 + Z_4 + Z_5) = \frac{1}{5} (0) = 0$$

Rata-rata simpangan yang dibakukan = 0, sehingga:

$$\begin{aligned} \Sigma &= \sqrt{\frac{1}{N} \left\{ \sum Z_i^2 - \frac{(\sum Z_i)^2}{N} \right\}} \\ &= \sqrt{\frac{1}{5} \left\{ (-0,86)^2 + (0,86)^2 + \dots + (-1,14)^2 - \frac{(0)^2}{5} \right\}} \\ &= 0,9907 = 1 \end{aligned}$$

## F. DISPERSI RELATIF, KOEFISIEN VARIASI

Simpangan baku mempunyai satuan yang sama dengan satuan data aslinya. Hal ini merupakan suatu kelemahan jika kita ingin membandingkan dua kelompok data, misalnya modal dari perusahaan besar di Amerika dengan yang ada di Indonesia, harga 10 mobil (juta rupiah) dengan harga 10 ekor ayam (ribu rupiah) dan berat 10 ekor gajah dengan berat 10 ekor semut. Walaupun nilai simpangan baku untuk berat gajah atau harga mobil lebih besar, nilai ini belum tentu lebih heterogen atau lebih bervariasi daripada berat semut dan harga ayam.

Variasi aktual (dispersi aktual) yang ditentukan dari deviasi standar atau berbagai ukuran dispersi yang lain disebut sebagai dispersi mutlak (absolut). Namun demikian suatu variasi sebesar 10 cm dalam pengukuran jarak 10 km tentu berbeda dengan variasi yang sama namun pada jarak 1 km. Ukuran terhadap efek ini dinyatakan melalui dispersi relatif, yang didefinisikan oleh

$$\text{Dispersi relative} = \frac{\text{dispersi mutlak}}{\text{rata-rata}}$$

Untuk perbandingan dua kelompok nilai dipergunakan Koefisien Variasi (KV), yang bebas dari satuan data asli. Koefisien variasi merupakan rasio antara simpangan baku dengan rata-ratanya yang rumuskan sebagai berikut:

$$\text{KV} = \frac{\sigma}{\mu} \times 100\%, \text{ untuk populasi}$$

Jika ada dua kelompok data dengan  $KV_1$  dan  $KV_2$ , di mana  $KV_1 > KV_2$ , maka kelompok pertama lebih bervariasi atau lebih heterogen daripada kelompok kedua.

**Contoh 5.13.** Harga 5 motor bekas masing-masing adalah Rp. 4.000.000, Rp. 4.500.000, Rp. 5.000.000, Rp. 4.750.000, serta Rp. 4.250.000 dan harga 5 ayam masing-masing Rp. 600, Rp.800,

Rp.900, Rp.550, dan Rp.1000. hitunglah simpangan baku harga mobil ( $\sigma_m$ ) dan harga ayam ( $\sigma_a$ ). Mana yang lebih bervariasi (heterogen), harga motor atau harga ayam?

$$\begin{aligned}\mu_m &= \frac{1}{5}(4.000.000 + 4.500.000 + \dots + 4.250.000) \\ &= Rp. 4.500.000\end{aligned}$$

$$\sigma_m = \sqrt{\frac{1}{5}\sum(X_i - \mu_m)^2} = Rp. 353.550$$

$$\begin{aligned}\mu_a &= \frac{1}{5}(600 + 800 + \dots + 1.000) \\ &= Rp. 770\end{aligned}$$

$$\begin{aligned}\sigma_a &= \sqrt{\frac{1}{5}\sum(X_i - \mu_a)^2} \\ &= Rp. 172,05\end{aligned}$$

$$\begin{aligned}KV_m &= \frac{\sigma_m}{\mu_m} \times 100\% \\ &= \frac{353.550}{4.500.000} \times 100\% = 7,86\%\end{aligned}$$

$$KV_a = \frac{\sigma_a}{\mu_a} \times 100\% = \frac{172,05}{770} \times 100\% = 22,34\%$$

Karena  $KV_a > KV_m$ , maka harga ayam lebih bervariasi (heterogen) dibandingkan harga mobil.

## G. RINGKASAN

- 1) Ukuran penyebaran adalah suatu ukuran baik parameter atau statistik untuk mengetahui seberapa besar penyimpangan data dengan nilai rata-rata hitungnya.
- 2) Ukuran penyebaran, didefinisikan sebagai derajat penyebaran nilai-nilai variabel suatu tendensi sentral dalam distribusi. Terdapat berbagai macam ukuran variabilitas, yakni: a) range, range antar kuartil, deviasi kuartil; b) mean deviasi, dan c) standard deviasi.

- 3) Jangkauan merupakan perbedaan antara nilai terbesar dan terkecil dalam kelompok data baik populasi maupun sampel. Semakin kecil ukuran jangkauan menunjukkan karakter yang lebih baik, karena berarti mendekati nilai pusat dan kompak.
- 4) Range, merupakan jarak antara nilai tertinggi dengan nilai terendah. Kelemahan range adalah: a) sangat tergantung dua nilai ekstrim, b) tidak dapat menunjukkan bentuk distribusi, dan c) tidak memenuhi definisi variabilitas.
- 5) Kelemahan pokok mean deviasi adalah dalam perhitungannya mengabaikan tanda-tanda plus dan minus. Oleh karena itu, mean deviasi tidak dapat dikenai perhitungan matematik yang tetap mempertahankan nilai-nilai plus dan minus.
- 6) Variasi adalah alat ukur variabilitas serangkaian data yang dihitung dengan mencari rata-rata selisih/beda kuadrat antara data observasi dengan pusat datanya.
- 7) Ukuran penyebaran yang sering digunakan adalah simpangan baku atau deviasi standar. Simpangan baku merupakan salah satu ukuran dispersi yang diperoleh dari akar kuadrat positif varians.
- 8) Secara matematik, standard deviasi didefinisikan sebagai akar dari jumlah deviasi kuadrat dibagi banyaknya individu dalam distribusi. Langkah pertama dalam mencari standard deviasi adalah dengan menghitung besaran (nilai) mean.

## **H. SOAL DAN LATIHAN**

- 1) Jelaskan apa yang dimaksud variabilitas? Berikan contoh dari macam-macam ukuran variabilitas?
- 2) Apa yang dimaksudkan dengan range?
- 3) Sebutkan kelemahan range dalam mendiskripsi variabilitas?
- 4) Apa yang dimaksudkan dengan Range Inter-Kuartil?
- 5) Jelaskan apa yang dimaksudkan dengan deviasi mean?

- 6) Carilah jangkauan himpunan bilangan (a) 12,6,7,3,15,10,18,5 dan (b) 9,3,8,8,9,8,9,18. Bagaimana interpretasi terhadap hasil jangkauan kedua data itu serta apakah jangkauan merupakan ukuran yang baik untuk ukuran penyebaran.
- 7) Hitunglah jangkauan (range) dari berat badan 100 mahasiswa Pendidikan Ekonomi berikut ini .

Berat Badan (kg)	Banyaknya Mahasiswa
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8

## Bab 6.

# MOMEN, SKEWNESS DAN KURTOSIS

### KOMPETENSI DASAR

Mahasiswa dapat mendeskripsikan dan melakukan perhitungan tentang momen, skewness dan kurtosis.

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- Mendefinisikan tentang konsep momen.
- Menghitung nilai momen pada distribusi data.
- Menjelaskan tentang konsep kemiringan.
- Memformulasikan koefisien kemiringan Pearson.
- Menggambarkan bentuk kemiringan.
- Menjelaskan tentang konsep kurtosis.
- Menghitung nilai koefisien kurtosis pada sebuah data.
- Menggambarkan bentuk kurva kurtosis

### DESKRIPSI SINGKAT

Bab ini secara garis besar membahas tentang momen, kemiringan dan kurtosis, yang akan membantu dalam penggunaan statistik deskriptif dalam pengkajian permasalahan pendidikan maupun ekonomi. Selain itu juga dapat memperluas dan memperkuat pengetahuan tentang konsep-konsep statistik deskriptif.

### KATA KUNCI

Momen	Skewness	Kurtosis
Leptokurtik	Platikurtik	Mesokurtik
Koefisien Kemencengan		

## A. PENGANTAR

Sebuah kurva yang simetris, rata-rata hitung, median dan modusnya berada pada satu titik. Akan tetapi kurva yang tidak simetris bentuknya dapat menjulur ke kiri atau kanan. Ukuran kemenjuluran/kemencengan (skewness) adalah derajat asimetris distribusi frekuensi. Disamping tingkat kemencengan, satu hal yang perlu juga kita pelajari adalah tingkat keruncingan sebuah kurva, jika kurva dalam distribusi normal maka mempunyai

distribusi yang tidak mendatar dan tidak meruncing, bagaimana jika kurva mempunyai tingkat keruncingan yang tinggi atau bahkan lebih tampak mendatar?. Bab ini akan membahas lebih dalam untuk memberikan pemahaman terhadap Skewness dan Kurtosis.

Rata-rata dan varians sebenarnya merupakan hal istimewa dari kelompok ukuran lain yang disebut *momen*. Dari momen ini pula beberapa ukuran lain dapat diturunkan. Bentuk-bentuk sederhana dari momen dan ukuran-ukuran yang didapat akan diuraikan di dalam bab ini.

## B. MOMEN

Pada bagian ini dibahas tentang momen, meliputi konsep: momen ke-r di sekitar bilangan tetap (asli), momen ke-r di sekitar nol, momen ke-r di sekitar rata-rata, momen ke-r dalam bentuk data distribusi frekuensi, momen ke-r dengan cara sandi. Supaya Anda dapat memahami konsep momen tersebut dengan jelas, pelajarilah uraian berikut ini dengan tuntas.

Misalkan diberikan variable x data sebanyak  $n$ :  $X_1, X_2, \dots, X_n$ , Jika  $A$  = sebuah bilangan tetap dan  $r = 0, 1, 2, \dots, n$ , maka momen ke-r sekitar  $A$ , disingkat  $m_r'$  didefinisikan oleh hubungan

$$m_r' = \frac{\sum (x_i - A)^r}{n}$$

Untuk  $A = 0$  didapat *momen ke-r* sekitar nol atau disingkat *momen ke-r* ( $m_r$ ), didefinisikan oleh hubungan: (data tidak berkelompok) :

$$m_r' = \frac{1}{n} \sum_{i=1}^n X_i^r$$

Dari persamaan diatas maka untuk  $r = 1$  didapat rata-rata  $\bar{x}$ , Jika  $A = \bar{x}$  maka kita peroleh momen ke-r sekitar rata-rata, biasa disingkat  $m_r$ .



$$m_r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r$$

Untuk data yang sudah dikelompokkan menjadi  $k$  kelas, dan  $X_i$  merupakan nilai tengah kelas ke- $i$ , maka perumusan momen ke- $r$  ( $m_r$ ) adalah sebagai berikut:

$$m_r = \frac{\sum f_i \cdot X_i^r}{n}$$

Untuk  $r = 1$ , maka  $M_1$  (momen pertama) merupakan rata-rata hitung. Momen tersebut merupakan momen terhadap titik asal, sedangkan momen terhadap rata-rata hitung adalah sebagai berikut:

$$m_r = \frac{1}{n} \sum_{i=1}^k f_i (X_i - \bar{X})^r$$

Untuk  $r = 2$ , maka  $m_2$  (momen kedua) merupakan varians (=kuadrat dari simpangan baku =  $S^2$ ). Momen ketiga dan keempat, yaitu  $m_3$  dan  $m_4$  masing-masing berguna untuk mengukur **kemencengan** (*skewness*) dan **keruncingan** (*kurtosis*) dari suatu distribusi frekuensi.

Jika menggunakan cara pengkodean maka dapat dijabarkan rumus sebagai berikut:

$$m'_r = i^r \left( \frac{\sum f_i \cdot c_i^r}{n} \right)$$

Harga  $m_r$  ditentukan berdasarkan hubungan yang dapat dirumuskan sebagai berikut:

$$\begin{aligned} m_2 &= m'_2 - (m'_1)^2 \\ m_3 &= m'_3 - 3m'_1 m'_2 + 2(m'_1)^3 \\ m_4 &= m'_4 - 4m'_1 m'_3 + 6(m'_1)^2 m'_2 - 3(m'_1)^4 \end{aligned}$$

**Contoh 6.1.** Seorang peternak ikan hias mempunyai 5 ikan arwana yang dipelihara dan diberi makanan spesial pada suatu hari ikan tersebut ditimbang. Setelah ditimbang 5 ikan tersebut mempunyai berat sebagai berikut (dalam kg): 2, 3, 7, 8, 10, Tentukan momen pertama, kedua, ketiga dan keempat di sekitar rata-rata sebuah data tersebut.

Soal di atas dapat diselesaikan dengan langkah berikut :

Langkah pertama menentukan nilai rata-rata dari berat tersebut,

$$\bar{X} = \frac{2 + 3 + 7 + 8 + 10}{5} = \frac{30}{5} = 6$$

Langkah kedua untuk memudahkan membuat tabel penolong.

Tabel 6.1. Penolong Perhitungan Momen Berat 5 Ikan

$X_j$	$(X_j - \bar{X})$	$(X_j - \bar{X})^2$	$(X_j - \bar{X})^3$	$(X_j - \bar{X})^4$
2	-4	16	-64	256
3	-3	9	-27	81
7	1	1	1	1
8	2	4	8	16
10	4	16	64	256
Jumlah	0	46	-18	610

$$m_1 = (0/5) = 0$$

$$m_2 = (46/5) = 9,2$$

$$m_3 = (-18/5) = -3,60$$

$$m_4 = (610/5) = 122$$

**Contoh 6.2.** Berdasarkan data pada contoh-contoh sebelumnya, tentukan empat buah momen sekitar rata-rata untuk data dalam daftar distribusi seperti pada kita lakukan sebagai berikut :

Tabel 6.2. Penolong Perhitungan Momen dari Daftar Distribusi

Kelas	$f_i$	$c_i$	$c_i^2$	$f_i c_i$	$f_i c_i^2$	$f_i c_i^3$	$f_i c_i^4$
118 – 126	3	-3	9	-9	27	-81	243
127 – 135	5	-2	4	-10	20	-40	80
136 – 144	9	-1	1	-9	9	-9	9
145 – 153	12	0	0	0	0	0	0
154 – 162	5	1	1	5	5	5	5
163 – 171	4	2	4	8	16	32	64
172 – 180	2	3	9	6	18	54	162
Jumlah	40	0	28	-9	95	-39	563

Dengan menggunakan rumus di atas dapat ditentukan nilai  $m_r$  sebagai berikut:

$$m'_1 = 9^1 \left( \frac{-9}{40} \right) = 9(-0,225) = -2,025$$

$$m'_2 = 9^2 \left( \frac{95}{40} \right) = 81(2,375) = 192,375$$

$$m'_3 = 9^3 \left( \frac{-39}{40} \right) = 729(-0,975) = 710,775$$

$$m'_4 = 9^4 \left( \frac{563}{40} \right) = 6.561(14,075) = 92.346,075$$

Maka untuk menentukan rata-rata adalah

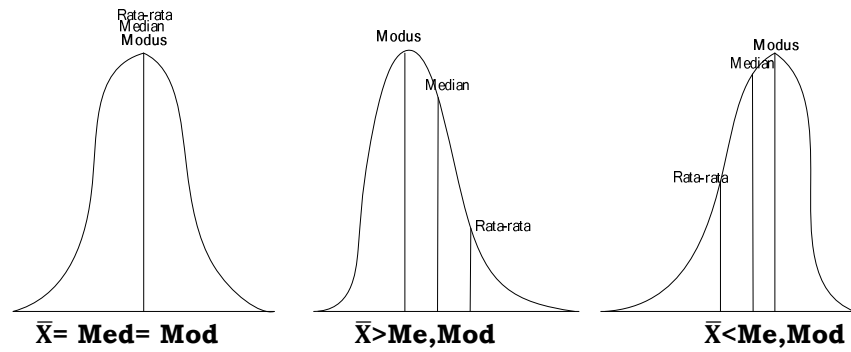
$$m_2 = m'_2 - (m'_1)^2$$

$$m_2 = 192,375 - (-2,025)^2 = 188,274$$

### C. UKURAN KEMENCENGAN (SKEWNESS)

Kemencengan atau kecondongan (skewness) adalah derajat ketidaksimetrisan (penyimpangan) dari kesimetrian dari sebuah distribusi. Sebuah distribusi yang tidak simetris akan memiliki rata-rata, median, dan modus yang tidak sama besarnya ( $\bar{X} \neq Me \neq Mo$ ), sehingga distribusi akan terkonsentrasi pada salah satu sisi dan kurvanya akan menceng. Jika distribusi memiliki ekor yang lebih panjang ke kanan daripada yang ke kiri maka distribusi disebut *menceng ke kanan* atau memiliki *kemencengan positif*. Sebaliknya, jika distribusi memiliki ekor yang lebih panjang ke kiri daripada yang ke kanan maka distribusi disebut *menceng ke kiri* atau memiliki *kemencengan negatif*.

Untuk distribusi miring, mean akan cenderung berada pada posisi sama dengan modus di ekor kurva yang lebih panjang. Kurva yang tidak simetris dapat menceng ke kiri atau ke kanan. Di dalam kurva yang simetris, letak modus, median, dan rata-rata ( $\bar{X}$ ) sama. Perhatikan bentuk-bentuk kecondongan kurva seperti ditampilkan pada gambar di bawah ini :



Gambar 6.1. Bentuk Kecondongan Kurva

Jadi ukuran ketidaksimetrisan bisa diperoleh dari selisih atau perbedaan nilai mean dan modus. Ukuran ini dapat dibuat menjadi ukuran tanpa dimensi atau satuan jika kita membaginya dengan suatu ukuran dispersi misalnya deviasi standar, untuk mengetahui bahwa konsentrasi distribusi menceng ke kanan atau menceng ke kiri.

### 1. Koefisien Kemencengan Pearson

Koefisien Kemencengan Pearson merupakan nilai selisih rata-rata dengan modus dibagi simpangan baku. Koefisien Kemencengan Pearson dirumuskan sebagai berikut:

$$sk = \frac{\bar{X} - \text{mod}}{S}$$

Apabila secara empiris, hubungan antar nilai pusat adalah sebagai berikut :

$$\bar{X} - \text{mod} = 3(\bar{X} - \text{med})$$

Maka rumus kemencengan dapat diubah sebagai berikut:

$$sk = \frac{3(\bar{X} - \text{med})}{S}$$

Jika nilai  $sk$  dihubungkan dengan keadaan kurva maka:

- a)  $sk = 0 \rightarrow$  kurva memiliki bentuk simetris

- b)  $sk > 0 \rightarrow$  nilai-nilai terkonsentrasi pada sisi kanan ( $\bar{X}$  terletak di sebelah kanan modus), sehingga kurva memiliki ekor memanjang ke kanan, kurva menceng ke kanan atau menceng positif.
- c)  $sk < 0 \rightarrow$  nilai-nilai terkonsentrasi pada sisi kiri ( $\bar{X}$  terletak di sebelah kiri modus). Sehingga kurva memiliki ekor memanjang ke kiri, kurva menceng ke kiri atau menceng negatif.

**Contoh 6.3.** Berikut ini adalah data nilai ujian statistik dari 40 mahasiswa program studi Pendidikan Ekonomi FKIP Universitas Mulawarman pada tahun 2010. Tentukan nilai  $sk$  dan uji arah kemencengannya (gunakan kedua rumus tersebut)! Gambarlah kurvanya !

Tabel 6.3. Nilai Ujian Statistik 40 Mahasiswa FKIP Unmul

Nilai Ujian	Frekuensi
31-40	4
41-50	3
51-60	5
61-70	8
71-80	11
81-90	7
91-100	2

Penyelesaian soal di atas dapat dilakukan dengan langkah-langkah sebagai berikut :

Tabel 6.4. Penolong Perhitungan Nilai Koefisien Kemencengan

Nilai	X	f	c	c <sup>2</sup>	fc	fc <sup>2</sup>
31 – 40	35,5	4	-4	16	-16	64
41 – 50	45,5	3	-3	9	-9	27
51 – 60	55,5	5	-2	4	-10	20
61 – 70	65,5	8	-1	1	-8	8
71 – 80	75,5	11	0	0	0	0
81 – 90	85,5	7	1	1	7	7
91 – 100	95,5	2	2	4	4	8
Jumlah		40			-32	134

$$\bar{X} = X_0 + i \left( \frac{\sum c_i f_i}{\sum f_i} \right) = 75,5 + 10 \left( \frac{-32}{40} \right) = 67,5$$

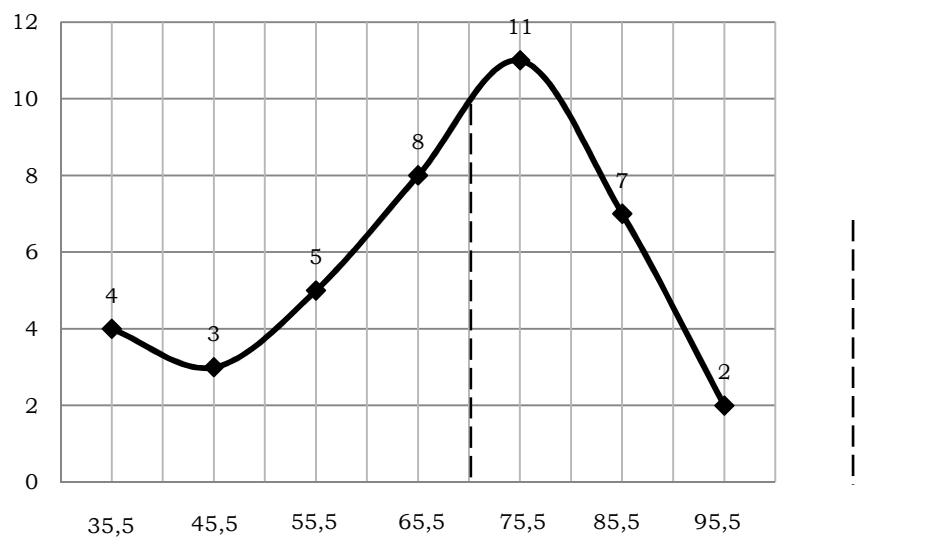
$$S = i \sqrt{\frac{\sum_{i=1}^k f_i c_i^2}{n} - \left[ \frac{\sum_{i=1}^k f_i c_i}{n} \right]^2} = 10 \sqrt{\frac{134}{40} - \left[ \frac{-32}{40} \right]^2} = 16,2$$

$$Me = B_m + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\} = 60,5 + 10 \left\{ \frac{20 - 12}{8} \right\} = 60,5 + 10 = 70,5$$

$$\text{mod} = B_m + i \left\{ \frac{(d_1)}{(d_1) + (d_2)} \right\} = 70,5 + 10 \left\{ \frac{(4)}{(4) + (5)} \right\} = 70,5 + 4,444 = 74,94$$

$$sk = \frac{\bar{X} - \text{mod}}{S} = \frac{67,5 - 74,94}{16,2} = -0,46$$

Oleh karena nilai sk-nya negatif maka kurvanya menceng kiri



Gambar 6.2. Kurva Menceng Kiri

## 2. Koefisien Kemencengan Bowley

Sebuah perumusan tentang pengukuran kemencengan yang lebih sederhana dari perumusan Pearson dikembangkan oleh A.L. Bowley. Bowley mengembangkan koefisien didasarkan hubungan antar statistik, koefisien kemencengan Bowley berdasarkan pada hubungan kuartil-kuartil ( $Q_1$ ,  $Q_2$  dan  $Q_3$ ) dari sebuah distribusi.

sedangkan distribusi yang menceng secara negatif, maka  $(Q_3 - Q_2) < (Q_2 - Q_1)$ . Koefisien kemencengan Bowley dirumuskan secara umum sebagai berikut :

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

Koefisien kemencengan Bowley sering juga disebut Koefisien Kuartil Kemencengan. Apabila nilai  $sk_B$  dihubungkan dengan keadaan kurva, didapatkan :

- 1) Jika  $Q_3 - Q_2 > Q_2 - Q_1$  maka distribusi akan menceng ke kanan atau menceng secara positif.
- 2) Jika  $Q_3 - Q_2 < Q_2 - Q_1$  maka distribusi akan menceng ke kiri atau menceng secara negatif.
- 3)  $sk_B$  positif, berarti distribusi menceng ke kanan.
- 4)  $sk_B$  negatif, berarti distribusi menceng ke kiri.
- 5)  $sk_B = \pm 0,10$  menggambarkan distribusi yang menceng tidak berarti dan  $sk_B > 0,30$  menggambarkan kurva yang menceng berarti.
- 6) Ukuran kemencengan kuartil selain disimbolkan dengan  $sk_B$  juga sering pula disimbolkan  $QCS = \text{Quartile Coefficient of Skewness}$  (Kuartil Koefisien Kemencengan).

**Contoh 6.4.** Tentukan kemencengan kurva distribusi frekuensi Nilai Ujian Matematika Ekonomi 111 mahasiswa angkatan 2010

Tabel 6.5. Nilai Ujian Matematika Ekonomi 111 Mahasiswa

Nilai Ujian	Frekuensi
20,00 – 29,99	4
30,00 – 39,99	9
40,00 – 49,99	25
50,00 – 59,99	40
60,00 – 69,99	28
70,00 – 79,99	5
Jumlah	111

Soal di atas dapat diselesaikan dengan langkah-langkah sebagai berikut :

Kelas  $Q_1$  = kelas ke-3

$$Q_1 = B_q + i \left[ \frac{\frac{n}{4} - f_{kq}}{f_q} \right] = 39,995 + 10 \left[ \frac{27,75 - 13}{25} \right] = 45,895$$

Kelas  $Q_2$  = kelas ke-4

$$Q_2 = B_q + i \left[ \frac{\frac{n}{2} - f_{kq}}{f_q} \right] = 49,995 + 10 \left[ \frac{55,5 - 38}{40} \right] = 54,37$$

Kelas  $Q_3$  = kelas ke-5

$$Q_3 = B_q + i \left[ \frac{\frac{3}{4} - f_{kq}}{f_q} \right] = 59,995 + 10 \left[ \frac{83,25 - 78}{28} \right] = 61,87$$

$$sk_B = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{61,87 - 2(54,37) + 45,895}{61,87 - 45,895} = -0,06$$

Karena  $sk_B$  negatif (-0,06) maka kurva menceng ke kiri dengan kemencengan yang berarti.

### 3. Koefisien Kemencengan Persentil

Koefisien kemencengan persentil didasarkan atas hubungan antar persentil ( $P_{90}$ ,  $P_{50}$  dan  $P_{10}$ ) dari sebuah distribusi. Koefisien kemencengan persentil dirumuskan:

$$Sk_p = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{(P_{50} - P_{10})}$$

dimana :

$sk_p$  = koefisien kemencengan persentil,

P = persentil

Simbol kemencengan persentil selain  $sk_p$  yang sering dijumpai adalah PCS atau *Percentile Coefficient of Skewness* (Kuartil Koefisien Kemencengan).



#### 4. Koefisien Kemencengan Momen

Koefisien Kemencengan Momen berdasarkan perbandingan momen ke-3 dengan pangkat tiga simpangan baku. Koefisien kemencengan momen dilambangkan  $\alpha_3$ . Koefisien kemencengan momen disebut juga kemencengan relatif.

Apabila nilai  $\alpha_3$  dihubungkan dengan keadaan kurva, maka didapatkan:

- Untuk distribusi simetris (normal), nilai  $\alpha_3 = 0$ ,
- Untuk distribusi menceng ke kanan, nilai  $\alpha_3 =$  positif,
- Untuk distribusi menceng ke kiri, nilai  $\alpha_3 =$  negatif,

Menurut Karl Pearson, distribusi yang memiliki nilai  $\alpha_3 > \pm 0,50$  adalah distribusi yang sangat menceng. Menurut Kenney dan Keeping, nilai  $\alpha_3$  bervariasi antara  $\pm 2$  bagi distribusi yang menceng. Untuk mencari besaran nilai momen ( $\alpha_3$ ), dibedakan proses perhitungannya yang terdiri dari perhitungan terhadap data tunggal dan data berkelompok.

##### a. Data Tunggal

Koefisien Kemencengan Momen untuk data tunggal dapat menggunakan rumus sebagai berikut:

$$\alpha_3 = \frac{M_3}{S^3} = \frac{\frac{1}{2} \sum (X - \bar{X})^3}{S^3}$$

$\alpha_3 =$  koefisien kemencengan momen

##### b. Data Berkelompok

Koefisien Kemencengan Momen untuk data berkelompok dapat menggunakan rumus sebagai berikut:

$$\alpha_3 = \frac{M_3}{S^3} = \frac{\frac{1}{2} \sum (X - \bar{X})^3 f}{S^3}$$

Dalam pemakaiannya, rumus kedua lebih praktis dan lebih mudah perhitungannya. Berikut diberikan Contoh 6.5 agar rumus mudah dipahami sehingga tepat untuk sarana latihan.

**Contoh 6.5.** Berdasarkan data berkelompok dari contoh modal 40 populasi perusahaan sebelumnya (Contoh 6.2), hitunglah  $sk$  (tingkat kemencengan) dan  $\alpha_3$  (ukuran tingkat kemencengan).

Soal di atas dapat diselesaikan dengan langkah-langkah sebagai berikut :

Tabel 6.6. Penolong Perhitungan Kemencengan Data Berkelompok

Kelas	$X_i$	$f$	$fX_i$	$c$	$fc$	$fc^2$	$fc^3$	$fc^4$
118 - 126	122	3	366	-3	-9	27	-91	243
127 - 135	131	5	655	-2	-10	20	-40	80
136 - 144	140	9	1.260	-1	-9	9	-9	9
145 - 153	149	12	1.788	0	0	0	0	0
154 - 162	158	5	790	1	5	5	5	5
163 - 171	167	4	668	2	8	16	32	64
172 - 180	176	2	352	3	6	18	54	162
Jumlah		40	5.879		-9	95	-39	563

Untuk menghitung  $\alpha_4$

$$\bar{X} = \frac{\sum f_i x_i}{\sum f_i} = \frac{5.879}{40} = 146,975$$

$$Me = B_m + i \left\{ \frac{\frac{n}{2} - (\sum f_{km})}{f_m} \right\} \text{ dengan, } i = 9; B_m = 144,5$$

$$med = 144,5 + 9 \left( \frac{20-17}{12} \right) = 146,75$$

$$S = i \sqrt{\frac{\sum f_i c_i^2}{n} - \left( \frac{\sum f_i c_i}{n} \right)^2} = 9 \sqrt{\frac{95}{40} - \left( \frac{-9}{40} \right)^2} = 13,72$$

$$sk = \frac{3(\bar{X} - med)}{S} = \frac{3(146,975 - 146,75)}{13,72} = 0,049$$

$$\begin{aligned} \alpha_3 &= \frac{9^3}{(13,72)^3} \left\{ \frac{-39}{40} - 3 \left( \frac{95}{40} \right) \left( \frac{-9}{40} \right) + 2 \left( \frac{-9}{40} \right)^3 \right\} \\ &= 0,282 \text{ (0,605)} \\ &= 0,17 \end{aligned}$$

Semakin besar nilai  $\alpha^3$ , maka kurva suatu distribusi makin menceng atau miring). Batas-batas nilai ukuran kemencengan beserta artinya dijelaskan sebagai berikut :

1.  $0,0 \leq (sk = \alpha^3) < 0,1$  maka bentuk kurva distribusinya bisa dianggap normal.

2.  $0,0 \leq (sk = \alpha^3) < 0,3$  maka bentuk kurva distribusinya menceng. Bila bernilai negatif menceng ke kiri, bila bernilai positif menceng ke kanan.
3.  $(sk = \alpha^3) \geq 0,3$  maka bentuk kurva distribusinya sangat menceng. Bila bernilai negatif sangat menceng ke kiri, bila bernilai positif sangat menceng ke kanan.

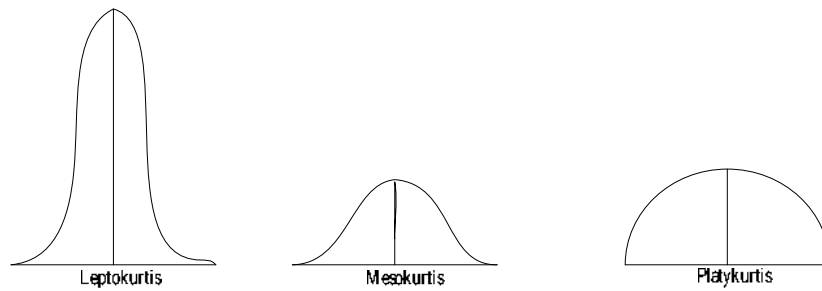
## **D. KERUNCINGAN (KURTOSIS)**

### **1. Pengertian Kurtosis**

Pengukuran kurtosis (peruncingan) sebuah distribusi teoritis adakalanya dinamakan pengukuran eksek (excess) dari sebuah distribusi. Sebenarnya kurtosis bisa dianggap sebagai suatu distorsi dari kurva normal. Keruncingan atau kurtosis adalah tingkat ketinggian puncak atau keruncingan dari sebuah distribusi yang biasanya diambil secara relatif terhadap suatu distribusi normal. Berdasarkan keruncingannya, kurva distribusi dapat dibedakan atas tiga macam, yaitu sebagai berikut:

- a) Leptokurtik, adalah distribusi yang memiliki puncak relatif tinggi.
- b) Platikurtik, adalah distribusi yang memiliki puncak hampir mendatar.
- c) Mesokurtik, merupakan distribusi yang memiliki puncak tidak tinggi dan tidak mendatar. Bila distribusi merupakan distribusi simetris, maka distribusi mesokurtik dianggap normal.

Pada kurva simetris, jika skala tegak lurus kurva normal ditarik secara memanjang dan skala horisontalnya dipersempit maka kurvanya akan menjadi tinggi dan ramping. Sebaliknya, jika skala tegak lurus diperpendek dan skala horisontal diperlebar, maka kurvanya akan menjadi pendek dan melebar.



Gambar 6.3 Kurva Leptokurtis, Mesokurtis, dan Platykurtis.

Keruncingan (koefisien kurtosis) dilambangkan dengan  $\alpha_4$  (alpha 4). Jika hasil perhitungan koefisien keruncingan diperoleh:

- Nilai lebih kecil dari 3, maka distribusinya adalah platikurtik.
- Nilai lebih besar dari 3, maka distribusinya adalah leptokurtik.
- Nilai sama dengan 3, maka distribusinya adalah mesokurtik.

## 2. Pengukuran Kurtosis

Tingkat keruncingan suatu kurva distribusi bisa dihitung dengan menggunakan  $\alpha_4$ , yaitu *moment coefficient of kurtosis* yang untuk data tunggal menggunakan rumus sebagai berikut:

$$\alpha_4 = \frac{M_4}{S^4} = \frac{\frac{1}{n} \sum_{i=1}^n (X - \bar{X})^4}{S^4} = \frac{\frac{1}{n} \sum (X - \bar{X})^4}{S^4}$$

**Contoh 6.6.** Berikut diketahui Kuota Anggaran Beasiswa (dalam milyar rupiah) bagi 5 Kabupaten di wilayah utara Kalimantan Timur yang meliputi Tana Tidung, Bulungan, Nunukan, Malinau, Berau dan Tarakan masing-masing sebagai berikut: 2, 3, 6, 8, 11, tentukan tingkat keruncingan kurva (kurtosis) dari data tersebut !

3.16

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} = \frac{30}{5} = 6$$

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}} = \sqrt{\frac{52}{4}} = 3,66$$

Tabel 6.7. Penolong Perhitungan Tingkat Keruncingan Kurva

X	$X - \bar{X}$	$(X - \bar{X})^2$	$(X - \bar{X})^4$
2	-4	16	256
3	-3	9	81
6	0	0	0
8	2	4	16
11	5	25	625
Jumlah	0	54	978

$$\alpha_4 = \frac{\frac{1}{n} \sum (X - \bar{X})^4}{s^4} = \frac{\frac{1}{5}(978)}{(3,66)^4} = \frac{195,6}{179,4} = 1,09$$

Karena nilainya 1,09 (lebih kecil dari 3) maka distribusinya adalah distribusi platikurtik.

Jika data tersebut berbentuk kelompok, maka rumus yang digunakan adalah sebagai berikut:

$$\alpha_4 = \frac{M_4}{S^4} = \frac{\frac{1}{n} \sum_{i=1}^k f_i (M_i - \bar{X})^4}{S^4}$$

$$\begin{aligned} \alpha_4 &= \frac{6.561}{35.433,68} \{14,075 - (-0,975) - (-0,225) + 6(2,375) - 3\left(\frac{-9}{40}\right)^4\} \\ &= 0,185 \{14,075 - 4(0,129) + 6(0,120) - 3(0,0020)\} \\ &= 0,185 (14,075 - 0,876 + 0,72 - 0,0078) \\ &= 2,57 \end{aligned}$$

Jika  $\alpha_4 > 3$  dihasilkan kurva *leptokurtis* (meruncing)

$\alpha_4 = 3$  dihasilkan kurva *mesokurtis* (normal)

$\alpha_4 < 3$  dihasilkan kurva *platykurtis* (mendatar)

Rumus yang lain menggunakan Koefisien Kurtosis Persentil dilambangkan dengan  $K$  (kappa). Untuk distribusi normal, nilai  $K = 0,263$ . Koefisien kurtosis persentil menggunakan rumus sebagai berikut:

$$K = \frac{\frac{1}{2} (Q_3 - Q_1)}{P_{90} - P_{10}}$$

**Contoh 6.7.** Berikut ini disajikan tabel frekuensi dari tinggi 100 mahasiswa universitas XYZ (data rekayasa).

- Tentukan koefisien kurtosis persentil (K)
- Apakah distribusinya termasuk distribusi normal

Tabel 6.8. Tinggi 100 Mahasiswa Universitas XYZ

Tinggi (inchi)	Frekuensi (f)
60 – 62	5
63 – 65	18
66 – 68	42
69 – 71	27
72 – 74	8

$$\text{Kelas } Q_1 = \text{kelas ke-3, } Q_1 = 65,5 + 3 \left( \frac{\frac{1 \cdot 100}{4} - 23}{42} \right) = 65,64$$

$$\text{Kelas } Q_3 = \text{kelas ke-4, } Q_3 = 68,5 + 3 \left( \frac{\frac{3 \cdot 100}{4} - 65}{27} \right) = 69,61$$

$$\text{Kelas } P_{10} = \text{kelas ke-2, } P_{10} = 62,5 + 3 \left( \frac{\frac{10 \cdot 100}{100} - 5}{18} \right) = 63,33$$

$$\text{Kelas } P_{90} = \text{kelas ke-4, } P_{90} = 68,5 + 3 \left( \frac{\frac{90 \cdot 100}{100} - 65}{27} \right) = 71,28$$

Koefisien kurtosis persentil (K) adalah:

$$K = \frac{\frac{1}{2}(Q_3 - Q_1)}{P_{90} - P_{10}} = \frac{\frac{1}{2}(69,61 - 65,64)}{71,28 - 63,33} = 0,25$$

Karena  $K = 0,25$  ( $K < 0,263$ ) maka bukan distribusi normal.

## E. RINGKASAN

- Momen ke- $r$  sekitar  $A$ , disingkat  $m_r'$  dirumuskan  $m_r' = \frac{\sum(x_i - A)^r}{n}$
- Jika  $A = 0$  didapat momen ke- $r$  sekitar nol atau disingkat *momen ke- $r$  ( $m_r$ )* sebagai berikut :

Data Tidak Berkelompok	Data Berkelompok
$m_r = \frac{1}{n} \sum_{i=1}^n X_i^r$ atau $m_r = \frac{X_i^r}{n}$	$m_r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r$

- 3) Data yang sudah dikelompokkan menjadi  $k$  kelas, dan  $X_i$  merupakan nilai tengah kelas ke- $i$ , maka perumusan momen ke- $r$  ( $m_r$ ) adalah sebagai berikut:

$$m_r = \frac{1}{n} \sum_{i=1}^k f_i X_i^r \text{ atau } m_r = \frac{\sum f_i X_i^r}{n}$$

- 4) Momen rata-rata hitung, dirumuskan  $m_r = \frac{1}{n} \sum_{i=1}^k f_i (X_i - \bar{X})^r$

Harga  $m_r$  berdasarkan hubungan:

$$m_2 = m_2' - (m_1')^2$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2m_2' - 3(m_1')^4$$

- 5) Kemencengan (*skewness*) adalah derajat ketidaksimetrisan atau penyimpangan dari kesimetrian dari sebuah distribusi.

## F. SOAL DAN LATIHAN

- Carilah momen-momen (a) pertama, (b) kedua, (c) ketiga dari himpunan bilangan 2, 3, 7, 8, 10.
- Carilah momen-momen (a) pertama, (b) kedua, (c) ketiga dari rata-rata bilangan pada soal di atas.
- Hitunglah empat buah momen sekitar rata-rata untuk data dalam daftar distribusi frekuensi berikut:

Data	60-62	63-65	66-68	69-71	72-74	Jumlah
$f_i$	5	18	42	27	8	100

- 4) Carilah koefisien (a) pertama dan (b) kedua person kemiringan untuk distribusi upah 65 karyawan PMA seperti tabel berikut:

x\$	255	265	275	285	295	305	315	Jumlah
$f_i$	8	10	16	14	10	5	2	100

- 5) Hitung koefisien momen kurtosis dari berat badan mahasiswa universitas ADAAZA sesuai dengan tabel distribusi berikut:

Data	61	64	67	70	73	Jumlah
$f_i$	5	18	42	27	8	100

## Bab 7.

# REGRESI LINIER DAN KORELASI SEDERHANA

### KOMPETENSI DASAR

Mahasiswa dapat menggunakan konsep regresi linier dan korelasi sederhana dalam statistik.

### TUJUAN PEMBELAJARAN

Setelah mempelajari bagian ini, mahasiswa diharapkan dapat :

- Menjelaskan pengertian regresi dan korelasi.
- Menjelaskan peubah tak bebas (*dependent variables*) dengan peubah bebas (*independent variables*).
- Menjelaskan asumsi yang mendasari analisis regresi linier.
- Menentukan penduga intersep dan slope pada regresi linier sederhana.
- Menjelaskan konsep koefisien determinasi.
- Menghitung nilai korelasi dari dua variabel.
- Menentukan penggunaan rumus korelasi sesuai jenis data.

### DESKRIPSI SINGKAT

Salah satu tujuan analisis data ialah memperkirakan besarnya efek kuantitatif dari perubahan suatu kejadian. Permasalahan ini bisa dibantu pemecahannya melalui penggunaan konsep regresi dan korelasi. Bahasan kedua konsep itu akan dikupas dalam bab ini, diawali dengan pemahaman konsep regresi linier, penggunaan rumus persamaan regresi yang terdiri atas metode penggunaan, pemahaman korelasi dan penggunaan rumus dari berbagai teknik korelasi, baik skala data nominal, ordinal maupun rasio serta keterkaitan antara regresi dan korelasi

### KATA KUNCI

Regresi	Korelasi Rank Order
Metode Kuadrat Terkecil	Contingency Coefficient
Analisis Korelasi	Korelasi Pearson Produk Momen
Koefisien Determinasi	

## A. PENGANTAR

Kehidupan senantiasa berhadapan dengan berbagai gejala yang meliputi berbagai variabel, contohnya berat badan dalam taraf tertentu tergantung pada tinggi badannya, produktivitas



pada keadaan tertentu tergantung pada efisiensi dan efektivitas kerjanya. Keterkaitan antar variabel ini juga menjadi hal yang fundamental dalam berbagai penelitian kuantitatif baik bidang pendidikan, ekonomi, sosial, pertanian dan juga perikanan.

Contoh kongkret di dalam analisa ekonomi dan bisnis, dalam mengolah data sering digunakan analisis regresi dan korelasi. Analisa regresi dan korelasi telah dikembangkan untuk mempelajari pola dan mengukur hubungan statistik antara dua atau lebih variabel. Namun karena bab ini hanya membahas tentang regresi linier dan korelasi sederhana, maka hanya dua variabel yang digunakan. Sedangkan sebaliknya, jika lebih dari dua variabel yang terlibat, maka disebut regresi dan korelasi berganda. Analisa ini memberikan hasil apakah antara variabel-variabel yang sedang diteliti atau sedang dianalisis terdapat hubungan, baik saling berhubungan, saling mempengaruhi dan seberapa besa tingkat hubungannya. Pada dasarnya, analisis ini menganalisis hubungan dua variabel dimana membutuhkan dua kelompok hasil observasi atau pengukuran sebanyak  $n$  (data).

Analisis masalah distribusi pasangan variabel dinamakan dengan analisa bivariat, analisa seperti ini membutuhkan data yang terdiri dari dua kelompok hasil observasi atau pengukuran. Data demikian itu dapat diperoleh dari berbagai kegiatan diatas yang menghasilkan pasangan observasi sebanyak  $n$  yang dinyatakan sebagai  $(X_i, Y_i)$ , dimana  $i=1,2,3\dots n$ . Variabel  $X$  bisa motivasi kerja dan variabel  $Y$  bisa Kinerja atau contoh yang paling dekat adalah dalam dunia pendidikan variabel  $X$  adalah minat belajar dan variabel  $Y$  adalah hasil belajar.

Data hubungan antara variabel  $X$  dan  $Y$  berdasarkan pada dua hal, yaitu:

- 1) Penentuan bentuk persamaan yang sesuai guna meramalkan rata-rata  $Y$  melalui  $X$  atau rata-rata  $X$  melalui  $Y$  dan menduga kesalahan selisih peramalan. Hal ini menitikberatkan pada

observasi variabel tertentu, sedangkan variabel-variabel lain dikonstantir pada berbagai tingkat atau keadaan. Hal ini yang dinamakan regresi.

- 2) Pengukuran derajat keeratan antara variabel X dan Y. Derajat ini tergantung pada pola variasi atau interelasi yang bersifat simultan dari X dan Y. Pengukuran ini disebut korelasi.

Hubungan antara variabel X dan Y mungkin merupakan hubungan dependen sempurna dan mungkin juga merupakan hubungan independen sempurna. Variabel X dan Y dikatakan berasosiasi atau berkorelasi secara statistik jika terdapat batasan antara dependen dan independen sempurna. Metode analisis ini juga digunakan untuk mengestimasi atau menduga besarnya suatu variabel yang lain telah diketahui nilainya. Salah satu contoh adalah untuk menganalisis hubungan antara tingkat pendapatan dan tingkat konsumsi.

Pada dasarnya, masalah hubungan antara variabel X dan Y umumnya berkisar pada dua hal yang terkadang susah ditarik garis pemisahannya. Analisa kedua variabel tersebut dapat berupa analisis regresi dan korelasi mengenai dua variabel yang didasarkan pada garis lurus akan dibicarakan dalam dua bagian, yaitu bagian pertama analisis secara umum dan bagian kedua analisis dengan sampel. Pada bab ini dibicarakan analisis secara umum yang menggambarkan hubungan antara dua variabel tertentu.

## **B. ANALISIS REGRESI LINIER**

Analisis regresi dapat digunakan untuk dua hal pokok yaitu memperoleh suatu persamaan dan garis yang menunjukkan persamaan hubungan antara dua variabel. Persamaan dan garis yang didapat disebut dengan persamaan regresi, yang berbentuk linear maupun non linear. Di samping itu dapat juga digunakan untuk menaksir satu variabel, yang disebut *dependent variabel*,

yang dalam bab ini adalah variabel Y, dengan variabel lain yang disebut *independen variabel* atau variabel X, berdasarkan hubungan yang ditunjukkan oleh regresi.

Analisis regresi adalah alat statistik yang memanfaatkan hubungan antara dua atau lebih peubah kuantitatif sehingga salah satu peubah bisa diramalkan dari peubah-peubah lainnya. Misalnya, jika kita tahu hubungan antara kemampuan awal mahasiswa dan prestasi belajarnya pada tahun pertama di perguruan tinggi, maka kita dapat meramal prestasi belajar mahasiswa melalui analisis regresi jika kemampuan awal mahasiswa telah kita ketahui.

Tujuan utama subbab ini adalah bagaimana menghitung suatu persamaan regresi yang akan menjelaskan hubungan antara dua variabel. Pembahasan kita akan terbatas pada regresi sederhana, yaitu mengenai hubungan antara dua variabel yang biasanya cukup tepat dinyatakan dalam sebuah garis lurus.

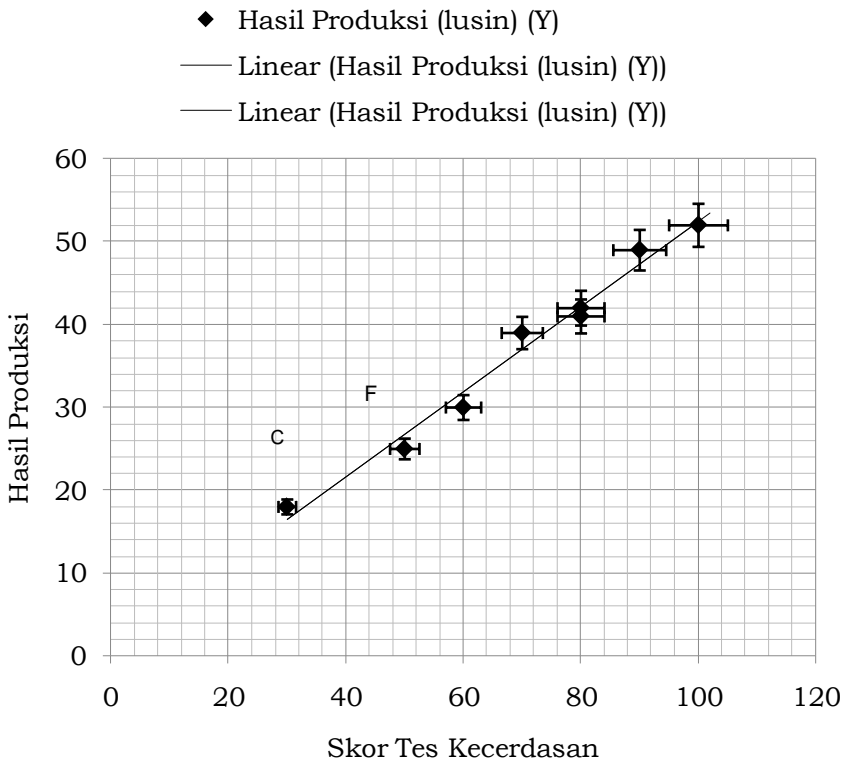
Setelah ditetapkan bahwa terdapat hubungan logis di antara variabel, maka untuk mendukung analisis lebih jauh, barangkali tahap selanjutnya adalah menggunakan grafik. Grafik ini disebut diagram pencar yang menunjukkan titik-titik tertentu. Setiap titik memperlihatkan suatu hasil yang kita nilai sebagai variabel tak bebas (*dependent*) maupun bebas (*independent*). Diagram pencar ini memiliki 2 manfaat, yaitu: (1) membantu menunjukkan apakah terdapat hubungan bermanfaat antara dua variabel, dan (2) membantu menetapkan tipe persamaan yang menunjukkan hubungan antara kedua variabel tersebut.

Kita dapat menjelaskan tujuan dan manfaat diagram pencar menggunakan data pada Tabel 7.1. Data itu menggambarkan hasil produksi karyawan yang dinyatakan dalam satuan lusin (sebagai variabel tak bebas) dan hasil *aptitude test* atau tes kecerdasan karyawan (sebagai variabel bebas), terhadap 8 orang karyawan pabrik mainan anak-anak “Tackey”. Jika hasil tes

kecerdasan menunjukkan asumsi yang kita perkirakan, rasanya masuk akal bila kita asumsikan karyawan yang mempunyai nilai (skor) tinggi memberikan hasil produksi yang lebih tinggi juga.

Tabel 7.1. Hasil Produksi dan Skor Tes Kecerdasan 8 Karyawan Pabrik Mainan Anak-anak “Tackey”

Karyawan	Hasil Produksi (lusin) (Y)	Skor Tes Kecerdasan (X)
A	30	60
B	49	90
C	18	30
D	42	80
E	39	70
F	25	50
G	41	80
H	52	100



Gambar.7.1. Diagram Pencar Hubungan Hasil Produksi dan Skor Tes Kecerdasan

Data dari seorang karyawan akan menunjukkan satu titik tertentu pada diagram pencar, seperti yang terlihat pada Gambar 7.1. Seperti pada gambar itu, kedelapan titik membentuk suatu garis lurus, dan hubungan yang sangat erat terlihat dengan fakta bahwa semua titik sangat dekat dengan garis lurus. Kita juga bisa melihat bahwa ada hubungan yang positif (atau langsung) antara kedua variabel. Sebagai contoh, bila skor tes kecerdasan meningkat, maka hasil produksi meningkat. Tentu saja, ada kemungkinan pada variabel tertentu terdapat hubungan negatif (atau berlawanan), yaitu bila X meningkat, Y menurun.

### 1. Persamaan Regresi Linear

Garis lurus pada Gambar 7.1 memperlihatkan hubungan antara variabel disebut garis regresi atau garis perkiraan. Seperti telah diketahui, persamaan yang digunakan untuk mendapatkan garis regresi pada data diagram pencar disebut persamaan regresi atau persamaan perkiraan\*\*. Jika kita ingat pembahasan pada bagian akhir tentang perhitungan *trend* linear, mungkin kita sedikit menemui kesukaran dengan persamaan regresi di sini. Mengapa? Karena persamaan umum garis lurus diperlihatkan di sini juga digunakan untuk menempatkan garis regresi pada data yang diperoleh. Oleh karena itu, metode kuadrat terkecil (*method of least squares*) akan kita gunakan untuk menempatkan garis pada data yang diamati, sehingga bentuk dari persamaan regresi adalah sebagai berikut:

$$\hat{Y} = a + bX$$

a = Bilangan Konstan (Nilai  $\hat{Y}$  bila  $X = 0$ )

b = Kemiringan garis regresi (kenaikan atau penurunan  $\hat{Y}$  untuk setiap perubahan satu satuan X) atau koefisien regresi, yang mengukur pengaruh X terhadap Y kalau X naik satu unit.

X = Nilai tertentu dari variabel bebas

$\hat{Y}$  = Nilai yang dihitung pada variabel tidak bebas (baca: *Ye topi*)

Bentuk persamaan regresi itu sering dibaca sebagai regresi X atas Y, artinya regresi X sebagai variabel prediktor sedangkan Y sebagai kriteriumnya. Koefisien regresi linier dinyatakan dengan huruf b yang juga menyatakan perubahan rata-rata variabel Y untuk setiap variabel X sebesar satu bagian. Cara menghitung persamaan regresi terdiri dua cara yakni dengan menggunakan metode skor deviasi (dengan rata-rata hitung) dan metode skor mentah (tanpa rata-rata).

**a. Metode Skor Deviasi**

Perhitungan analisis regresi dan analisis korelasi dapat dipermudah menggunakan rumus berbentuk penyimpangan nilai tengah Variabel X dan Y, yaitu penyimpangan dari  $\bar{X}$  dan  $\bar{Y}$ . Oleh karena itu, dapat digunakan simbol berikut ini:

$$x = (X - \bar{X}), y = (Y - \bar{Y}) \text{ dan } (Y - \bar{Y}) = b (X - \bar{X})$$

Nilai dari a dan b pada persamaan regresi tersebut dihitung dengan rumus berikut ini:

$$b = \frac{\sum x_i y_i}{\sum x_i^2}, a = \bar{Y} - b\bar{X}$$

**Contoh 7.1.** Carilah persamaan regresi berdasarkan data yang terdapat pada Tabel 7.1.

Langkah 1. Buat tabel penolong dan tentukan  $\bar{y}, \bar{x}, \sum xy, \sum xy, \sum x^2$

Tabel 7.2 Penolong Perhitungan Regresi

Karyawan	(Y)	(X)	$\frac{y}{(Y - \bar{Y})}$	$\frac{x}{(X - \bar{X})}$	xy	$x^2$	$y^2$
A	30	60	-7	-10	70	100	49
B	49	90	12	20	240	400	144
C	18	30	-9	-40	760	1600	361
D	42	80	5	10	50	100	25
E	39	70	2	0	0	0	4
F	25	50	-12	-20	240	400	144
G	41	80	4	10	40	100	16
H	52	100	15	30	450	900	225
	296	560	0	0	1850	3600	968

$$\bar{Y} = \frac{\sum Y}{N} = \frac{296}{8} = 37; \bar{X} = \frac{\sum X}{N} = \frac{560}{8} = 70$$

Langkah 2. Masukkan ke dalam rumus

$$b = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{1850}{3600} = 0,5139 \quad a = \bar{Y} - b\bar{X} = 37 - 0,5139(70) = 1,027$$

Langkah 3. Masukkan nilai a dan b ke dalam persamaan regresi

$$Y = 1,027 + 0,5139X$$

### b. Metode Skor Mentah

Kostanta a dan b diberikan oleh persamaan normal berikut:

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

**Contoh 7.2.** Diketahui penelitian terhadap hubungan antara nilai biaya periklanan dengan tingkat penjualan sebuah koperasi adalah sebagai berikut (dalam ribu). Cari persamaan regresinya.

Tabel 7.3. Hubungan Biaya Periklanan dan Tingkat Penjualan

Biaya Periklanan	Tingkat Penjualan
50	40
51	46
52	44
53	55
54	49

Langkah 1 : Menentukan variabel X dan variabel Y.

Dalam soal di atas, biaya periklanan merupakan variabel X dan tingkat penjualan merupakan variabel Y.

Langkah 2 : Membuat Tabel Regresi Sederhana

Tabel 7.4. Regresi Sederhana Hubungan Biaya Periklanan dan Tingkat Penjualan

Biaya Periklanan (X)	Tingkat Penjualan (Y)	(X) <sup>2</sup>	(Y) <sup>2</sup>	(XY)
50	40	2500	1600	2000
51	46	2601	2116	2346
52	44	2704	1936	2288
53	55	2809	3025	2915
54	49	2916	2401	2646
260	234	13530	11078	12195

Langkah 3 : Menentukan koefisien a dan koefisien b

$$b = \frac{n \sum XY - \sum X \cdot \sum Y}{n \sum X^2 - (\sum X)^2} = \frac{5(12195) - (260)(234)}{5(13530) - (260)^2} = 2,7$$

$$a = \frac{\sum Y - b \sum X}{n} = \frac{(234) - 2,7(260)}{5} = -93,6$$

Langkah 4 : Menentukan persamaan regresi linier sederhana

$$Y = a + bX$$

$$Y = -93,6 + 2,7X$$

### **C. ANALISIS KORELASI SEDERHANA**

Francis Galton merupakan orang pertama memperkenalkan alat statistik yang bernama analisis regresi dan analisis korelasi. Ia lahir di Birmingham (Inggris) pada 16 Februari 1822. Galton mempelajari hubungan antara tinggi badan ayah dan anak. Ia mengamati bahwa tinggi anak dari tinggi dan pendek cenderung “mundur (regress)” ke rata-rata grup. Kecenderungan ini dinamakan regresi ke arah “mediokritas (sedang)”. Ia lalu mengembangkan deskripsi matematis bagi kecenderungan regresi ini.

Analisis korelasi merupakan alat untuk mengukur keeratan hubungan antara dua variabel. Perhitungan dari derajat keeratan didasarkan pada persamaan regresi. Akan tetapi analisis korelasi dapat dilakukan tanpa adanya persamaan regresi. Suatu catatan bahwa tingginya tingkat korelasi tidak menunjukkan hubungan sebab akibat antara variabel. Mungkin diperoleh korelasi yang tinggi antara dua variabel tetapi tidak menunjukkan tingkat pengaruh antara dua variabel. Sebagai contoh naiknya hasil belajar dihubungkan dengan tinggi badan mahasiswa. Kita tidak bisa menyimpulkan bahwa tinggi badan seseorang menyebabkan naiknya hasil belajar. Tingkat korelasi tinggi hanya menunjukkan hasil matematis. Kita diharuskan membuat sebuah kesimpulan berdasarkan pandangan yang logis berdasarkan hasil penelitian ilmiah.



## 1. Koefisien Korelasi dan Kegunaannya

Hubungan dua variabel bisa positif dan negatif. Hubungan X dan Y dikatakan positif apabila kenaikan (penurunan) X pada umumnya diikuti kenaikan (penurunan) Y. Sebaliknya dikatakan negatif bila kenaikan (penurunan) X pada umumnya diikuti oleh penurunan (kenaikan) Y.

Contoh hubungan positif:

X = pupuk,

Y = produksi

X = biaya iklan,

Y = hasil penjualan

X = berat badan,

Y = tekanan darah

Contoh hubungan negatif:

X = jumlah akseptor,

Y = jumlah kelahiran

X = harga suatu barang,

Y = permintaan barang

X = pendapatan masyarakat,

Y = kejahatan ekonomi

Bila variabel X dan Y dikorelasikan, titik-titik koordinat dalam diagram pencar bertendensi membentuk lingkaran yang trend agak definitif arahnya. Pada Gambar 7.2, titik asal dari sumbu digambarkan pada titik koordinat  $(\bar{X}, \bar{Y})$  didistribusikan dalam keempat kuadran bidang  $X'Y'$ . Tanda-tanda aljabar dalam keempat kuadran diberikan sebagai berikut:

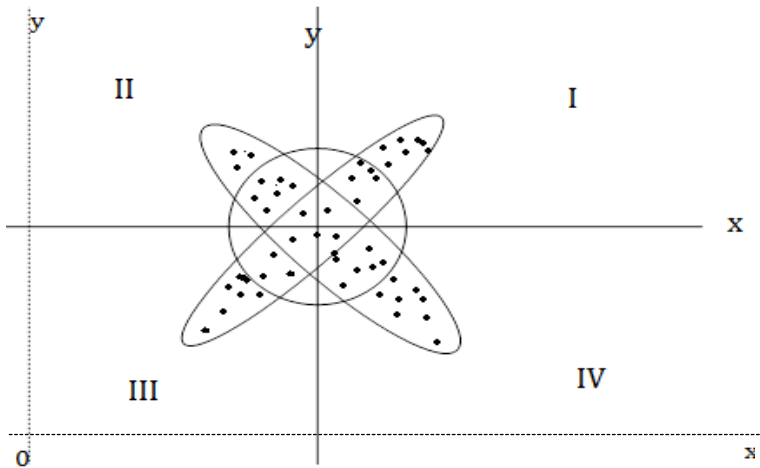
I -----> X' dan Y' positif

II -----> X' negatif dan Y' positif

III -----> X' dan Y' negatif

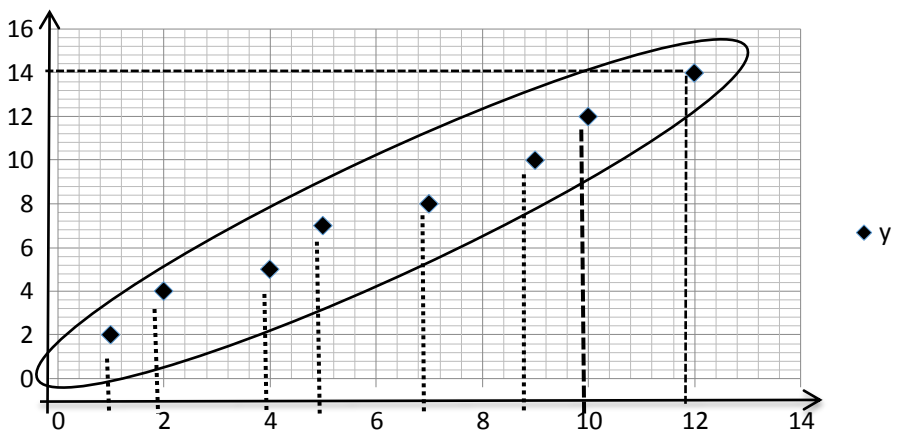
IV -----> X' positif dan Y' negatif

Hasil perkalian  $X'Y'$  akan menjadi bernilai positif bagi semua titik yang terdapat dalam Kuadran I dan III dan akan menjadi bernilai negatif bagi semua titik yang terdapat dalam kuadran II dan IV

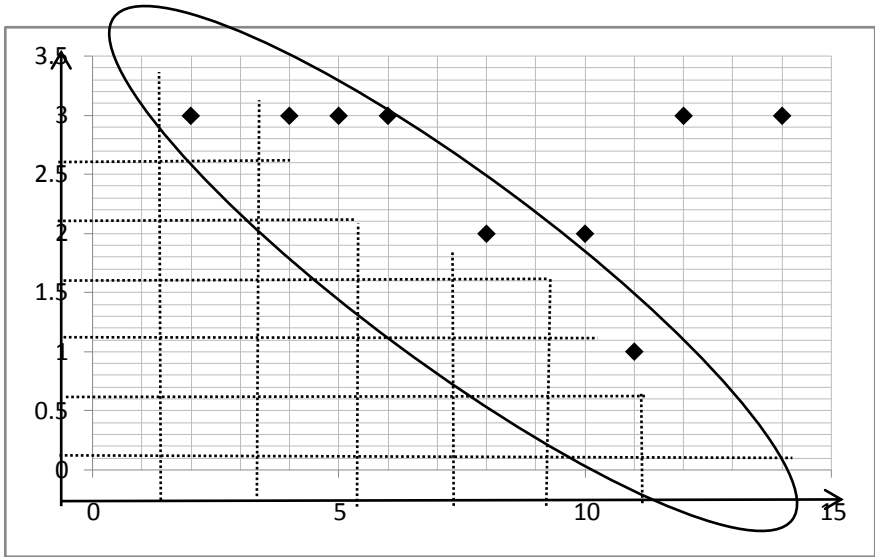


Gambar 7.2. Diagram Pencar dan Titik Koordinatnya dalam Keempat Kuadran.

Hubungan positif dan negatif antar variabel  $x$  dan  $y$  secara terperinci dijelaskan bahwa apabila variabel  $X$  dan  $Y$  terdapat hubungan, maka bentuk diagram pencarnya mulus/teratur, seperti terlihat dalam Gambar 7.3 dan Gambar 7.4, di mana yang pertama menunjukkan gerakan diagram pencar dari kiri bawah ke kanan atas (hubungan positif), sedangkan yang bergerak dari kiri atas ke kanan bawah (hubungan negatif).



Gambar 7.3. Diagram Pencar dan Titik-Titik Koordinat  $X$  dan  $Y$  Mempunyai Hubungan Positif



Gambar 7.4. Diagram Pencar dan Titik-Titik Koordinat X dan Y Mempunyai Hubungan Negatif

Apabila bentuk diagram pencar tidak teratur, artinya kenaikan/penurunan X pada umumnya tidak diikuti naik turunnya Y, maka dikatakan X dan Y tidak berkorelasi. Dengan kata lain, jika naik turunnya variabel X tidak mempengaruhi Y, dikatakan X dan Y bebas (*independen*), atau tidak ada hubungan atau hubungannya begitu lemahnya sehingga bisa diabaikan.

Kuat dan tidaknya hubungan antara X dan Y apabila dapat dinyatakan dengan fungsi linear (paling tidak mendekati), diukur dengan suatu nilai yang disebut Koefisien Korelasi. Nilai koefisien korelasi ini paling sedikit -1 dan paling besar 1. Jadi, jika  $r =$  koefisien korelasi, maka nilai  $r$  dapat dinyatakan sebagai berikut:

$$-1 \leq r \leq 1$$

Jika  $r = 1$ , hubungan X dan Y sempurna dan positif (mendekati 1, yaitu hubungan sangat kuat dan positif).

Jika  $r = -1$ , hubungan X dan Y sempurna dan negatif (mendekati -1, yaitu hubungan sangat kuat dan negatif).

Jika  $r = 0$ , hubungan X dan Y sangat lemah (tidak ada).

Di sini, X dikatakan mempengaruhi Y, jika berubahnya nilai X akan menyebabkan perubahan nilai Y; artinya naik/turunnya X membuat nilai Y naik/turun sehingga nilai Y akan bervariasi, baik terhadap rata-rata Y maupun terhadap garis linear yang mewakili diagram pencar. Namun, naik/turunnya Y sedemikian rupa sehingga nilai Y bervariasi, tidak semata-mata disebabkan X, karena masih ada faktor lain yang menyebabkan. Misalnya Y = hasil penjualan, X = biaya iklan (*advertensi*), maka naik/turunnya Y selain disebabkan variabel X juga faktor (variabel) lain (seperti pendapatan masyarakat, harga, selera, dan lain-lain).

## D. TEKNIK-TEKNIK KORELASI

### 1. Korelasi Pearson (Pearson's product moment coefficient)

Korelasi Pearson Produk Momen merupakan salah satu teknik korelasi yang paling banyak digunakan dalam berbagai penelitian. Besarnya korelasi atau koefisien korelasi dinyatakan dalam lambang  $r$ . Cara menentukan nilai koefisien korelasi ( $r$ ) terdiri dua cara yakni metode skor deviasi (rata-rata hitung) dan metode skor mentah (tanpa rata-rata).

#### a. Metode Skor Mentah (tanpa nilai rata-rata)

Rumus yang digunakan untuk menentukan nilai  $r$  dengan perhitungan tanpa rata-rata adalah sebagai berikut:

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

- Jika  $r = 0$  atau mendekati 0 maka hubungannya lemah.
- Jika  $r = (-1)$  maka hubungan sangat kuat, bersifat tidak searah.
- Jika  $r = (+1)$  maka hubungan sangat kuat, bersifat searah.

**Contoh 7.3.** Diketahui suatu penelitian terhadap hubungan antara biaya periklanan dengan tingkat penjualan dari sebuah koperasi seperti Contoh 7.2, tentukan besar koefisien korelasinya.

Langkah 1 : Menentukan variabel X dan variabel Y.

Biaya periklanan (variabel X) dan tingkat penjualan (variabel Y).

Langkah 2 : Membuat Tabel Penolong Korelasi

Tabel 7.5. Penolong Perhitungan Korelasi

Biaya Periklanan (X)	Tingkat Penjualan (Y)	(X) <sup>2</sup>	(Y) <sup>2</sup>	(XY)
50	40	2500	1600	2000
51	46	2601	2116	2346
52	44	2704	1936	2288
53	55	2809	3025	2915
54	49	2916	2401	2646
260	234	13530	11078	12195

Langkah 3 : Memasukkan data dari tabel ke rumus korelasi

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} = \frac{5(12195) - (260)(234)}{\sqrt{[5(13530 - (260)^2)][5(11078) - (234)^2]}}$$

$$= 0,76$$

**b. Metode Skor Deviasi (dengan nilai rata-rata)**

Cara menghitung *r* adalah sebagai berikut:

$$r_{xy} = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}}$$

jika  $x = X - \bar{X}$  dan  $y = Y - \bar{Y}$

**Contoh 7.4.** Carilah nilai korelasi berdasarkan data yang terdapat pada Tabel 7.1.

Langkah 1. Buat tabel penolong dan tentukan  $\bar{y}$ ,  $\bar{x}$ ,  $\sum xy$ ,  $\sum xy$ ,  $\sum x^2$

Tabel 7.6. Penolong Perhitungan Korelasi Metode Skor Deviasi

Karyawan	(Y)	(X)	$y$ ( $Y - \bar{Y}$ )	$x$ ( $X - \bar{X}$ )	$xy$	$x^2$	$y^2$
A	30	60	-7	-10	70	100	49
B	49	90	12	20	240	400	144
C	18	30	-9	-40	760	1600	361
D	42	80	5	10	50	100	25
E	39	70	2	0	0	0	4
F	25	50	-12	-20	240	400	144
G	41	80	4	10	40	100	16
H	52	100	15	30	450	900	225
	296	560	0	0	1850	3600	968

$$\bar{Y} = \frac{\sum Y}{N} = \frac{296}{8} = 37;$$

$$\bar{X} = \frac{\sum X}{N} = \frac{560}{8} = 70$$

Langkah 2. Masukkan ke dalam rumus

$$r_{xy} = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}} = \frac{1850}{\sqrt{[3600][968]}} = \frac{1850}{\sqrt{3484800}} = \frac{1850}{1866.76} = 0,99$$

Jadi nilai koefisien korelasinya adalah  $r=0,99$

## 2. Koefisien Determinasi

Koefisien determinasi dilambangkan dengan  $r^2$ , merupakan kuadrat dari koefisien korelasi. Koefisien ini dapat digunakan untuk menganalisis apakah variabel yang diduga atau diramal (Y) dipengaruhi oleh variabel (X) atau seberapa variabel independen (bebas mempengaruhi variabel dependen (tak bebas).

## 3. Kesalahan Standar Estimasi

Untuk mengetahui ketepatan persamaan estimasi dapat digunakan dengan mengukur besar kecilnya kesalahan standar estimasi. Semakin kecil nilai kesalahan standar estimasi, maka semakin tinggi ketepatan persamaan estimasi dihasilkan untuk menjelaskan nilai variabel sesungguhnya. Sebaliknya, semakin besar nilai kesalahan standar estimasi maka semakin rendah ketepatan persamaan estimasi dihasilkan untuk menjelaskan variabel dependen sesungguhnya. Kesalahan standar estimasi diberi simbol  $Se$  ditentukan rumus berikut:

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

## 4. Koefisien Korelasi Data Berkelompok

Rumus untuk menghitung koefisien korelasi yang sudah dibahas sebelumnya adalah untuk data yang tidak berkelompok (data yang belum disajikan dalam bentuk tabel frekuensi, dengan kelas atau kategori). Untuk data berkelompok rumusnya adalah :

$$r = \frac{n(\sum uvf) - (\sum uf_u)(\sum vf_v)}{\sqrt{n(\sum u^2 f_u) - (\sum uf_u)^2} \sqrt{n(\sum v^2 f_v) - (\sum vf_v)^2}}$$

### 5. Korelasi Rank (Peringkat)

Korelasi Rank Order (peringkat) dipergunakan mengukur hubungan dua variabel yang keduanya mempunyai tingkat pengukuran ordinal, parameter yang bisa menyatakan hubungan keduanya ini juga kerap kali disebut dengan korelasi Spearman atau Spearman's coefficient of (rank) correlation.

Sebagai **Contoh 7.5.**, jika ada dua orang, Nabil dan Agus, yang sama-sama penggemar buah. Kedua orang itu diminta untuk memberikan nilai terhadap 10 jenis buah. Buah paling digemari diberi nilai 1 dan seterusnya sampai buah yang tidak disenangi diberi nilai 10. Dengan kata lain, Nabil dan Agus diminta untuk memberikan rank (peringkat). Pemberian rank ini bisa dibalik, maksudnya buah yang disenangi diberi nomor 10 dan seterusnya, sementara yang tidak disenangi diberi nilai 1. Hasil pemberian rank adalah sebagai berikut:

Tabel 7.7. Rank Hipotesis Nilai Buah Versi Nabil dan Agus

Nomor Urut	Nama Buah	Rank Nabil	Rank Agus
(1)	(2)	(3)	(4)
1	Salak	9	8
2	Jambu	5	3
3	Markisa	10	9
4	Apel	1	2
5	Manggis	8	7
6	Pepaya	7	10
7	Melon	3	4
8	Mangga	4	6
9	Jeruk	2	1
10	Sawo	6	5

Apabila dibuat koefisien korelasi antara rank dari Agus dan Nabil terhadap 10 jenis buah tersebut, maka akan diperoleh *koefisien korelasi buah*.

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$$

dimana  $d_i$  = selisih dari pasangan rank ke- $i$ ;  $n$  = jumlah pasangan rank (dalam hal ini  $n = 10$ ). Rumus ini disebut Spearman.

Tabel 7.8. Tabel Penolong Perhitungan Korelasi

Rank Nabil	8	3	9	2	7	10	4	6	1	5
Rank Agus	9	5	10	1	8	7	3	4	2	6
Selisih rank ( $d$ )	-1	-2	-1	1	-1	3	1	2	-1	-1
$d^2$	1	4	1	1	1	9	1	4	1	1

$$r_{\text{rank}} = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} = 1 - \frac{6(1+4+1+1+1+9+1+4+1+1)}{10(100-1)} = 1 - 0,1455 = 0,85$$

**Contoh 7.6.** Ada 10 calon salesman yang diuji mengenai teknik penjualan. Setelah mereka selesai diuji kemudian ditugaskan untuk melakukan penjualan. Hasilnya adalah  $X$  = hasil ujian,  $Y$  = hasil penjualan tahun pertama. Nilai  $X$  dan  $Y$  dari 10 salesman termasuk rank-nya dan selisih rank ( $d$ ) adalah sebagai berikut:

Tabel 7.8. Rangkang Hasil Penjualan Berdasarkan Nilai Ujian

Nama	Nilai Ujian	Rank	Penjualan Tahun ke-1 (ribu rupiah)	Rank	Selisih Rank	
	( $X$ )		( $Y$ )		( $d$ )	$d^2$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Amin	48	3	312	2	1	1
Joni	32	6	164	8	-2	4
Tono	40	5	280	4	1	1
Amir	34	7	196	7	0	0
Ahmad	30	8	200	6	2	4
Paulus	50	1,5	288	3	-1,5	2,25
Purwanto	26	9	146	10	-1	1
Bambang	50	1,5	361	1	0,5	0,25
Jatmiko	22	10	149	9	1	1
Aryo	43	4	252	5	-1	1

Karena Paulus dan Bambang mempunyai nilai sama, maka rank mereka harus sama yaitu:  $\frac{1+2}{2} = 1,5$ . Mula-mula Paulus diberi



rank 1, Bambang diberi rank 2 (atau sebaliknya), kemudian dirata-ratakan. Ada kemungkinan tiga orang mempunyai nilai yang sama dan jatuh pada rank 5, 6, dan 7, dimana masing-masing mendapat rank yang sama yaitu:  $\frac{5+6+7}{3} = 6$ . Dalam hal ini, masing-masing diberi rank sesuai dengan urutannya, kemudian dicari rata-rata ranknya. Apabila ada beberapa rank yang sama, maka biasanya koefisien korelasi rank tidak sama dengan koefisien korelasi yang dihitung berdasarkan rumus Pearson (*product moment coefficient of correlation*).

$$r_{\text{rank}} = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} = 1 - \frac{6(1+4+1+0+4+2,25+1+0,25+1+1)}{10(10^2-1)} = 1 - \frac{6(15,5)}{10(99)} = 0,91$$

## E. KORELASI DATA KUALITATIF

Rumus koefisien korelasi yang telah dibahas dalam subbab sebelumnya hanya berlaku untuk data kuantitatif. Banyak sekali hasil pengumpulan data penelitian yang menghasilkan data kualitatif berupa kategori-kategori. Misalnya penduduk memiliki pendapatan tinggi, menengah dan rendah; penduduk pendapatan sedang dikategorikan mempunyai mobil sedan berukuran besar, sedang dan kecil.

Kita sering ingin mengetahui apakah ada hubungan antara selera konsumen dengan letak geografis; kedudukan orang tua kedudukan anak (kalau orang tua pangkatnya tinggi, apakah anaknya juga); pendidikan orang tua dan anak; pendidikan dan pendapatan (makin tinggi pendidikan, makin tinggi pendapatan?), tingkat kepuasan pelanggan dengan umur, pendapatan, masa kerja, pendidikan, kedudukan sosial, dan sebagainya. Untuk data kualitatif yang digunakan dalam mengukur kuatnya hubungan disebut Contingency Coefficient (Koefisien Bersyarat) yang mempunyai pengertian sama seperti koefisien korelasi. Misalnya, hasil penelitian disajikan dalam bentuk *p x q table*.

Telah dijelaskan di atas bahwa koefisien bersyarat ( $C_c$ ), dipergunakan untuk mengukur kuatnya hubungan data kualitatif yang mempunyai arti seperti koefisien korelasi, di mana nilai  $C_c$  sebesar nol, yang berarti tidak ada hubungan. Akan tetapi, batas atas  $C_c$  tidak sebesar satu, tergantung atau sebagai fungsi banyaknya kategori (baris atau kolom). Batas tertinggi nilai  $C_c$  ialah  $\sqrt{(r-1)/r}$ , di mana nilai  $r$  ialah banyaknya baris atau kolom. Kalau banyaknya baris tidak sama dengan banyaknya kolom, pilih nilai yang terkecil.

Adapun menghitung nilai koefisien bersyarat ( $C_c$ ) digunakan rumus:

$$C_c = \sqrt{\frac{x^2}{x^2+n}}$$

Sedangkan menentukan Chi Kuadrat ( $x^2$ ) menggunakan rumus sebagai berikut :

$$x^2 = \sum \frac{(f_o - f_h)^2}{f_h}, \quad f_h = \frac{n_b \times n_k}{n_{tot}}$$

$x^2$  = Chi Kuadrat,

$f_o$  = Frekuensi yang diobservasi

$f_h$  = Frekuensi yang diharapkan

$n_b \times n_k$  = Jumlah Frekuensi Baris x Jumlah Frekuensi Kolom

Jika nilai perbandingan  $C_c$  dengan batas tertinggi  $< 0,50$ , maka hubungannya lemah, jika terletak antara  $0,50$  dan  $0,75$  maka hubungannya sedang, antara  $0,75$  dan  $0,90$  maka sangat kuat, sama dengan  $1$  maka hubungannya sempurna.

**Contoh 7.7.** Apakah ada hubungan antara tingkat pendidikan ibu rumah tangga dengan konsumsi susu dari anggota keluarga mereka, dilakukan penelitian yang hasilnya disajikan dalam tabel dibawah ini. Hitung  $C_c$  untuk mengukur hubungan antara tingkat pendidikan dan konsumsi susu tersebut.

Tabel 7.9. Hubungan Tingkat Pendidikan dan Konsumsi Susu

Pendidikan	Konsumsi		
	Kurang	Cukup	Tinggi
(1)	(2)	(3)	(4)
Tidak Tamat SLA	82	65	12
Tamat SLA	59	112	24
Pernah masuk Perguruan Tinggi	37	94	42

Untuk memudahkan menyelesaikan soal di atas, kita harus menyusun tabel berikut.

I \ II	1	2	3	Jumlah
(1)	(2)	(3)	(4)	(5)
1	82 (53,70)	65 (81,76)	12 (23,53)	$n_1 = 159$
2	59 (65,86)	112 (100,28)	24 (28,86)	$n_2 = 195$
3	37 (58,43)	94 (88,96)	42 (25,61)	$n_3 = 173$
Jumlah	$n_1 = 178$	$n_2 = 271$	$n_3 = 78$	$n = 527$

$$f_{h11} = \frac{n_{1,n1}}{n} = \frac{(159)(178)}{527} = 53,70, \quad f_{h12} = \frac{n_{1,n1}}{n} = \frac{(159)(271)}{527} = 81,76$$

$$f_{h13} = \frac{n_{1,n1}}{n} = \frac{(159)(78)}{527} = 23,53, \quad f_{h21} = \frac{n_{1,n1}}{n} = \frac{(195)(178)}{527} = 65,86$$

$$f_{h22} = \frac{n_{1,n1}}{n} = \frac{(195)(271)}{527} = 100,28, \quad f_{h23} = \frac{n_{1,n1}}{n} = \frac{(195)(78)}{527} = 28,86$$

$$f_{h31} = \frac{n_{1,n1}}{n} = \frac{(173)(178)}{527} = 58,43, \quad f_{h32} = \frac{n_{1,n1}}{n} = \frac{(173)(271)}{527} = 88,96$$

$$f_{h33} = \frac{n_{1,n1}}{n} = \frac{(173)(78)}{527} = 25,61$$

$$\begin{aligned} x^2 &= \sum \frac{(f_o - f_h)^2}{f_h} \\ &= \frac{(28,30)^2}{53,70} + \frac{(-16,76)^2}{81,76} + \frac{(-11,53)^2}{25,53} + \frac{(-6,86)^2}{65,86} + \frac{(11,72)^2}{100,28} + \frac{(-4,86)^2}{28,86} + \frac{(-21,43)^2}{58,43} + \frac{(5,04)^2}{88,96} + \frac{(16,39)^2}{25,61} \\ &= 45,54 \end{aligned}$$

$$C_c = \sqrt{\frac{x^2}{x^2 + n}} = \sqrt{\frac{45,54}{45,54 + 527}} = 0,28$$

Karena jumlah baris dan jumlah kolom sama dengan 3, maka r sama dengan 3. Batas atas  $C_c$  ialah  $\sqrt{(3-1)/3} = 0,82$ . Karena nilai

perbandingan  $C_c$  dengan batas atas lebih kecil dari 0,50, maka hubungan atau korelasi antara tingkat pendidikan ibu rumah tangga dengan tingkat konsumsi susu anggota rumah tangga tidak begitu kuat, bahkan bisa dikatakan lemah.

**F. RINGKASAN**

- 1) Analisis regresi dapat digunakan untuk dua hal yaitu untuk memperoleh persamaan dan menunjukkan garis hubungan antara dua variabel. Persamaan dan garis hubungan yang diperoleh disebut persamaan regresi, berbentuk linear maupun non linear.
- 2) Metode kuadrat terkecil (method of least squares) untuk menempatkan garis pada data yang diamati, sehingga bentuk dari persamaan regresi adalah sebagai berikut:  $\hat{Y} = a + b X$
- 3) Persamaan regresi terdiri dua cara yakni metode skor deviasi (rata-rata hitung) dan metode skor mentah (tanpa rata-rata).

Metode Skor Deviasi	Metode Skor Mentah
$(Y - \bar{Y}) = b(X - \bar{X})$ $b = \frac{\sum x_i y_i}{\sum x_i^2}, a = \bar{Y} - b\bar{X}$	$a = \frac{\sum Y - b \sum X}{n}$ $b = \frac{n \sum XY - \sum X \cdot \sum Y}{n \sum X^2 - (\sum X)^2}$

- 4) Analisis korelasi adalah alat yang dipakai untuk mengukur keeratan hubungan antara dua variabel. Perhitungan dari derajat keeratan didasarkan pada persamaan regresi.
- 5) Korelasi Pearson Produk Momen merupakan teknik korelasi yang paling banyak digunakan dalam penelitian. Besarnya angka korelasi atau koefisien korelasi dinyatakan dalam r.
- 6) Cara menentukan nilai koefisien korelasi (r) terdiri dua cara yakni metode skor deviasi (rata-rata hitung) dan metode skor mentah (tanpa rata-rata).

Metode Skor Deviasi	Metode Skor Mentah
$r_{xy} = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}}$ $x = X - \bar{X}$ dan $y = Y - \bar{Y}$	$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$

- 7) Koefisien determinasi ( $r^2$ ) adalah kuadrat koefisien korelasi, yang digunakan untuk menganalisis variabel yang diduga (Y) dipengaruhi variabel (X) atau seberapa variabel independen (bebas) mempengaruhi variabel dependen (tak bebas).
- 8) Korelasi Rank Order (peringkat) untuk mengukur hubungan dua variabel yang keduanya mempunyai tingkat pengukuran ordinal, parameter yang menyatakan hubungan keduanya ini sering disebut dengan korelasi Spearman atau Spearman's coefficient of (rank) correlation.
- 9) Rumus korelasi rank order adalah  $r = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$
- 10) Data kualitatif yang dipergunakan mengukur hubungan disebut Contingency Coefficient (Koefisien Bersyarat) yang mempunyai pengertian sama dengan koefisien korelasi.

## G. SOAL DAN LATIHAN

- 1) Tabel dibawah menampilkan tinggi badan X dan Y dari sampel terdiri dari 12 pasangan ayah (X) dan anak laki-laki tertua (Y) (dalam cm). Cobalah a) buat diagram pencar, dan b) cari persamaan regresi dari data tersebut.

X	165	163	167	164	168	162	170	166	168	167	169	171
Y	168	166	168	165	169	166	168	165	171	167	168	170

- 2) Berdasarkan tabel di atas, tentukan nilai koefisien korelasi dan koefisien determinasi, berikan penjelasan terhadap hasil nilai koefisien tersebut.
- 3) Jelaskan perbedaan penggunaan korelasi produk momen, korelasi rank order dan chi-kuadrat.
- 4) Tabel dibawah ini menyajikan distribusi frekuensi nilai akhir 100 orang mahasiswa peserta kuliah matematika ekonomi dan statistik. Berdasarkan data itu a) tentukan jumlah mahasiswa yang mendapat nilai 70-79 untuk matematika ekonomi dan 80-89 untuk statistik, b) tentukan korelasi kedua mata kuliah, serta c) tentukan koefisien korelasi kedua nilai mata kuliah itu.

NILAI AKHIR		Matematika Ekonomi					
		40-49	50-59	60-69	70-79	80-89	90-99
Statistik	90-99				2	4	4
	80-89			1	4	6	5
	70-79			5	10	8	1
	60-69	1	4	9	5	2	
	50-59	3	6	6	2		
	40-49	3	5	4			

# DAFTAR PUSTAKA

- Agung, I. Gusti Ngurah. 2000. Metode Penelitian Sosial 1 dan 2. Gramedia. Jakarta.
- Alfifari. 1994. Statistik Ekonomi: Teori, Kasus dan Solusi. STIE YKPN. Yogyakarta.
- Anas Sudijono. 2006. Pengantar Statistik Pendidikan. PT Raja Grafindi Persada.
- Arikunto, S. 2007. Prosedur Penelitian Suatu Pendekatan Praktik. Rineka Cipta. Jakarta.
- Cresswell, J. W. 1998. Research Design: Qualitative and Quantitative Approaches. SAGE Publicational. London.
- Cresswell, J. W. 2008. Educational Research: Planning, Conductiong, and Evaluating Quantitative and Qualitative Research. Pearson Merrill Prentice Hall. New York.
- Dajan, Anto. 2008. Pengantar Metode Statistik Jilid 1. LP3ES. Jakarta.
- Eddy Marlianto. 2004. Metode Statistika. USU Press. Medan.
- Endang Soenarya. 1993. Statistika Dasar Universitas. M2S Bandung.
- Fraenkel, Jack R. dan Norman E. Wallen. 2009. How to Design and Evaluate Research. Sevent Edition. McGraw-Hill Inc. Singapore.
- Furqon. 2010. Statistika Terapan untuk Penelitian. Alfabeta. Bandung.
- Gall, M.D. et al. 2001. Educational Research: an Introduction. Allyn and Bacon. New York.
- Gronlund, N.E. 1976. Measurement and Evaluation in Teaching. McMillan Publishing Co.Inc. New York.

- Guilford, J.P., 1956. *Fundamental Statistics in Psychology and Education*. McGraw Hill Book Co. Inc. New York.
- Husaini Usman. 2006. *Pengantar Statistika*. Bumi Aksara. Jakarta.
- Joseph L. Schafer. 2003. *Analysis of Categorical Data*. Dept of Statistics Pennsylvania State University.
- Kuswadi dan Erna Mutiara. 2004. *Delapan Langkah dan Tujuh Alat Statistik*. Gramedia. Jakarta.
- M. Iqbal Hasan. 1999. *Pokok-pokok Materi Statistik 1 (Statistik Deskriptif)*. Bumi Aksara. Jakarta.
- Maxwell. 1996. *Qualitative and Quantitative Research*. Alyn and Bacon. New York.
- McMillan, J.H dan Schumacher, S. (2001). *Research in Education: A Conceptual Introduction*(5th ed.). Longman. Inc. USA.
- McMillan. 2008. *Educational Research: Fundamental for the Consumer*. Pearson Merrill Prentice Hall. New York.
- Minium, E. and King B. (1993). *Statistical Reasoning in Psychology and Education*. John Wiley & Sons. Inc. Canada.
- Nurgana, E. 1993. *Statistik Untuk Penelitian*. Bandung.
- Permadi. Nurgiantoro, B. 2001. *Statistik Terapan Penelitian Ilmu Sosial*. UGM Press. Yogyakarta.
- Ruseffendi, E. T. 1991. *Penilaian Pendidikan dan Hasil Belajar Siswa Khususnya dalam Pengajaran Matematika*. Tidak diterbitkan. Bandung.
- Ruseffendi, E.T. 1994. *Dasar-dasar Penelitian Pendidikan dan Bidang Non Eksakta Lainnya*. Unnes Press. Semarang.
- Ruseffendi, E.T. 1998. *Statistika Dasar untuk Penelitian Pendidikan*. IKIP Bandung Press. Bandung.
- Speigel, Murray R and Stephen Lary J. 2007. *Schaum's Outlines Teori dan Soal-Soal*. Erlangga. Jakarta.
- Sudjana, N. 2008. *Penilaian Hasil Proses Belajar Mengajar*. PT Remaja Rosdakarya. Bandung.



- Sudjana, N. dan Ibrahim, R. 2001. Penelitian dan Penilaian Pendidikan. Sinar Baru Algesindo. Bandung.
- Sudjana. 1995. Metode Statistika. Tarsito. Bandung.
- Sudjana. 1996. Metode Statistik. Tarsito. Bandung.
- Sugiyono. 1997. Statistika Untuk Penelitian. Alfabeta. Bandung.
- Suhardi dan Purwanto, S. K. 2004. Statistika untuk Ekonomi dan Keuangan Modern. Salemba Empat. Jakarta.
- Suparman I. A. 1995. Statistik Sosial. PT. Raja Grafindo Persada. Jakarta.
- Supranto, J. 2008. Statistik Teori dan Aplikasi Jilid 2. Erlangga. Jakarta.
- Susetyo, B. 2010. Statitik Terapan untuk Mengolah Data Penelitian. Refika Aditama. Bandung.
- Walpole, R.E. 1989. Pengantar Statistika, PT Gramedia. Jakarta
- Wand, E. dan Gerald, W. B. (1957). Essentials of Educational Evaluation. Holt Rinehart and Winston. New York.

# Pengantar Statistik Pendidikan

Buku ini pada awalnya merupakan hasil pengembangan bahan ajar sebagai pendukung instrumen riset disertasi saya. Bahan ajar tersebut menjadi referensi perkuliahan mahasiswa Program Studi Pendidikan Ekonomi di Fakultas Keguruan dan Ilmu Pendidikan (FKIP) Universitas Mulawarman, dan telah mengalami berbagai revisi dari hasil diskusi dan masukan baik oleh mahasiswa maupun rekan sejawat, guna mencapai tingkat kedalaman dan keluasan yang memadai sebagai sumber belajar mata kuliah Pengantar Statistik Pendidikan.

Buku ini terdiri atas 7 (tujuh) bab, yang menyajikan tentang statistik pendidikan secara menyeluruh berdasarkan kompetensi dasar yang harus ditempuh dan dicapai kompetensinya selama satu semester.



**Dr. Sudarman, S.Pd., M.Pd.**  
Dosen Universitas Mulawarman

Bidang khusus yang ditekuni saat ini antara lain :

- Komunikasi dalam Pembelajaran
  - Pengembangan Media Pembelajaran
  - Penelitian Pembelajaran dan Statistik
  - Pengembangan Kurikulum dan Desain Pembelajaran
- 

 **Mulawarman  
University PRESS**

Penerbit  
Mulawarman University PRESS  
Gedung LPPM Universitas Mulawarman  
Kampus Gunung Kelua, Jl. Krayan, Samarinda  
Provinsi Kalimantan Timur, INDONESIA 75123  
Telp/Fax. (0541) 747432, Email : mup@lppm.unmul.ac.id

ISBN 978-602-6834-02-7



9786026834027