



DASAR-DASAR STATISTIK PENELITIAN

Oleh :

Nuryadi, S.Pd.Si., M.Pd

Tutut Dewi Astuti, SE., M.Si, Ak., CA., CTA

Endang Sri Utami, SE., M.Si., Ak., CA

M. Budiantara, SE.,M.Si.,Ak, CA

SIBUKU MEDIA

Dasar-Dasar Statistik Penelitian

Oleh :

Nuryadi, S.Pd.Si., M.Pd

Tutut Dewi Astuti, SE., M.Si, Ak., CA., CTA

Endang Sri Utami, SE., M.Si., Ak., CA

M. Budiantara, SE.,M.Si.,Ak, CA

Diterbitkan oleh:

SIBUKU MEDIA

Ngringinan, Palbapang, Bantul, Bantul, Yogyakarta, 55713

Email: penerbitsibuku@gmail.com

www.sibuku.com

Dicetak oleh:

Gramasurya

Jl. Pendidikan No. 88 Yogyakarta 55182

Telp. /Fax. 0274 413364

Emial: gramasurya@gmail.com

ISBN: 978-602-6558-04-6

Cetakan Ke-1: Januari 2017

Kata Pengantar

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan hidayah-Nya atas kelancaran sehingga penulis dapat menyelesaikan buku yang berjudul “ Dasar-dasar Statistik Penelitian” semoga dengan dibuatkan buku ini pembaca dapat memahami tentang analisis-analisis dalam Statistika yang lebih berorientasi pada Penelitian.

Berbagai sumber referensi dasar dan esensial yang relevan dari artikel ilmiah, buku statistika, dari website memang sengaja dipilih dan digunakan untuk memperkuat pembahasan dan membangun kerangka penyajian yang komprehensif agar mudah dipahami dan dapat memenuhi harapan pembaca. Sasaran dari buku ini pada khususnya adalah mahasiswa yang merasa sulit apabila dihadapkan dengan penelitian kuantitatif yang berhubungan dengan statistik inferensial. Dalam buku ini disajikan juga statistik deskriptif, uji asumsi, dan uji hipotesis (uji t, uji F, dan uji X^2) beserta langkah-langkah dan interpretasi dari hasil output atau hasil perhitungannya.

Akhirnya, diharapkan semoga buku dasar-dasar statistik penelitian ini bermanfaat bagi pembaca khususnya mahasiswa dan para peneliti yang menggunakan pendekatan kuantitatif. Penulis juga mengharapkan kritik dan saran yang bersifat membangun demi pembuatan buku selanjutnya yang masih berhubungan dengan statistika. Oleh karena itu, penulis berharap agar buku ini dapat dijadikan sebagai bahan pembelajaran dan berguna bagi pembaca.

Yogyakarta, Januari 2017

Penulis

Daftar Isi

Kata Pengantar.....	iii
Daftar Isi	v
BAB 1. STATISTIKA DAN STATISTIK.....	1
A. Sejarah Statistika.....	1
B. Pengertian Statistika dan Statistik.....	1
C. Data Statistik.....	2
D. Populasi dan Sampel.....	8
BAB 2. PENYAJIAN DATA	11
A. Tabel.....	11
B. Diagram/Grafik.....	15
BAB 3. DISTRIBUSI FREKUENSI	27
A. Tabel Distribusi Frekuensi	27
B. Statistik Tabulasi Silang.....	34
BAB 4. PEMUSATAN DAN PENYEBARAN DATA.....	43
A. Mengukur Pemusatan Data.....	43
B. Ukuran Letak.....	50
C. Pengukuran Penyebaran (Dispersi)	56
BAB 5. KEMENCENGAN DAN KURTOSIS.....	69
A. Pengukuran Kemencengan.....	69
B. Pengukuran Kurtosis	71
BAB 6. INFERENSI STATISTIK	73
A. Estimasi Parameter	73
B. Uji Hipotesis	74
C. Pengujian Hipotesis.....	75
D. Langkah-langkah Uji Hipotesis.....	75

BAB 7. UJI NORMALITAS DATA DAN HOMOGENITAS DATA.....	79
A. Uji Normalitas.....	79
B. Uji Homogenitas	89
BAB 8. UJI T-TEST (PENGANTAR STATISTIK LANJUT)	95
A. Uji T-Test Satu Sampel (<i>One Sample t-test</i>).....	95
B. <i>Paired Sample t-Test</i>	101
C. <i>Independent Sample t-test</i>	107
BAB 9. UJI CHI SQUARE (Uji Data Kategorik)	117
A. Pendahuluan.....	117
B. Prosedur Sampel Tunggal dengan Chi Kuadrat.....	118
C. Prosedur untuk Sampel Indep-enden	121
D. Prosedur untuk Sampel Dependen.....	122
BAB 10. STATISTIK ANALISIS VARIANS	125
A. Karakter.....	125
B. Uji Hipotesis dalam Analisis Variansi	125
C. Contoh Perhitungan	126
BAB 11. ANALISIS REGRESI DAN KORELASI	133
A. Pendahuluan.....	133
B. Konstanta dan Koefisien Regresi.....	134
C. Koefisien Korelasi	136
D. Penaksiran Nilai Variabel Dependen	139
DAFTAR PUSTAKA	147
LAMPIRAN-LAMPIRAN	149
PROFIL PENULIS	169

Statistika dan Statistik

A. SEJARAH STATISTIKA

Ilmu statistika mempunyai sejarah yang sangat panjang seiring peradaban manusia. Pada zaman sebelum Masehi, bangsa-bangsa di Mesopotamia (Babilonia), Mesir, dan Cina telah mengumpulkan data statistic untuk memperoleh informasi tentang berapa besar pajak yang harus dibayar oleh setiap penduduk, beberapa banyak hasil pertanian yang mampu diproduksi, dan lain sebagainya. Pada abad pertengahan, lembaga gereja menggunakan statistika untuk mencatat jumlah kelahiran, kematian, dan pernikahan.

Statistika pertama kali di temukan oleh Aristoteles dalam bukunya yang berjudul “politea”, dalam buku tersebut ia menjelaskan data tentang keadaan 158 negara yang di sebut sebagai statistika. Pada abad ke-17 di Inggris, statistika di sebut sebagai political arithmetic. Pada abad ke-18, istilah statistika dipopulerkan oleh Sir John Sinclair dalam bukunya berjudul “statistical account of Scotland (1791-1799)”, setelah terlebih dahulu dikemukakan oleh seorang ahli hitung asal Jerman yang bernama Gottfried Achenwell (1719-1772).

B. PENGERTIAN STATISTIKA DAN STATISTIK

Pada umumnya orang tidak membedakan antara statistika dan statistic. Kata statistic berasal dari kata Latin yaitu status yang berarti “Negara” (dalam bahasa inggris adalah state). Pada awalnya kata statistic diartikan sebagai keterangan-

keterangan yang dibutuhkan oleh Negara dan berguna bagi Negara (Anto Dajan, Pengantar Metode Statistik). Misalkan keterangan mengenai jumlah keluarga penduduk suatu Negara, keterangan mengenai usia penduduk, pekerjaan penduduk suatu Negara dan sebagainya.

Agar pengertian statistic sebagai kumpulan angka-angka, tidak mengaburkan perbedaan antara kumpulan angka-angka dengan metode sehingga kumpulan angka tersebut “berbicara”. Dalam arti kumpulan angka tersebut disajikan dalam bentuk table/diagram, selanjutnya dianalisa dan ditarik kesimpulan. Ini semua ternyata merupakan pengetahuan tersendiri yang disebut statistika. Jadi Statistika adalah ilmu pengetahuan, murni dan terapan, mengenai penciptaan, pengembangan, dan penerapan teknik-teknik sedemikian rupa sehingga ketidakpastian inferensia induktif dapat dievaluasi. Statistik adalah kumpulan fakta yang berbentuk angka-angka yang disusun dalam bentuk daftar atau tabel yang menggambarkan suatu persoalan. Perbedaan dari statistic dan parameter adalah statistic merupakan sembarangan nilai yang menjelaskan nilai dari sampel. Sedangkan parameter merupakan sembarangan nilai yang menjelaskan nilai dari populasi.

Statistika dalam pengertian sebagai ilmu dibedakan menjadi dua yaitu :

1. **Statistic deskriptif** mempunyai tujuan untuk mendeskripsikan atau memberi gambaran objek yang diteliti: sebagaimana adanya tanpa menarik kesimpulan atau generalisasi. Dalam statistika deskriptif ini dikemukakan cara-cara penyajian data dalam bentuk table maupun diagram, penentuan rata-rata (mean), modus, median, rentang serta simpangan baku.
2. **Statistic inferensial (induktif)** mempunyai tujuan untuk penarikan kesimpulan. Sebelum penarikan kesimpulan dilakukan suatu dugaan yang dapat diperoleh dari statistic deskriptif.

C. DATA STATISTIK

Setiap kegiatan yang berkaitan dengan statistic, selalu berhubungan dengan data. Menurut kamus Besar Bahasa Indonesia pengertian data adalah keterangan yang benar dan nyata. Data adalah bentuk jamak dari datum. Datum adalah keterangan atau informasi yang diperoleh dari suatu pengamatan sedangkan data adalah segala keterangan atau informasi yang dapat memberikan gambaran tentang

suatu keadaan. Dari contoh-contoh yang telah diberikan sebelumnya, dapat diperoleh bahwa tujuan pengumpulan data adalah :

- Untuk memperoleh gambaran suatu keadaan.
- Untuk dasar pengambilan keputusan.

Syarat data yang baik agar memperoleh kesimpulan tepat dan benar maka data yang dikumpulkan dalam pengamatan harus nyata dan benar, diantaranya:

- Data harus obyektif (sesuai keadaan sebenarnya)
- Data harus mewakili (representative)
- Data harus update
- Data harus relevan dengan masalah yang akan dipecahkan.

1. Macam-macam Data

Data adalah kumpulan keterangan atau informasi yang di peroleh dari suatu pengamatan. Data dibagi menjadi beberapa macam, yaitu :

a) Klasifikasi Data Berdasarkan Jenis Datanya

❖ Data Kuantitatif

Data kuantitatif adalah data yang dipaparkan dalam bentuk angka-angka. Misalnya adalah jumlah pembeli saat hari raya idul adha, tinggi badan siswa kelas 3 ips 2, nilai matematika (...6,7,8,9,10,...) dan lain-lain.

❖ Data Kualitatif

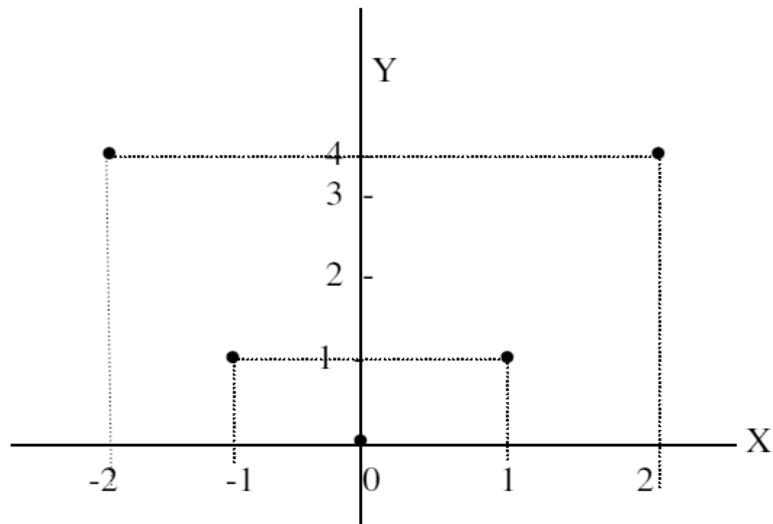
Data kualitatif adalah data yang disajikan dalam bentuk kata-kata yang mengandung makna. Contohnya seperti persepsi konsumen terhadap botol air minum dalam kemasan, anggapan para ahli terhadap psikopat, warna (merah, hijau, biru, kuning, hitam, dll) dan lain-lain.

b) Pembagian Jenis Data Berdasarkan Sifat Data

❖ Data Diskrit (cacahan)

Data diskrit adalah data yang nilainya adalah bilangan asli. Contohnya adalah berat badan ibu-ibu pkk sumber ayu, nilai rupiah dari waktu ke waktu, jumlah peserta yang hadir dalam seminar nasional pendidikan matematika. Jumlah siswa yang lulus try out akbar UAN 2011, jumlah buku yang terdapat pada perpustakaan kampus, dan lain-sebagainya.

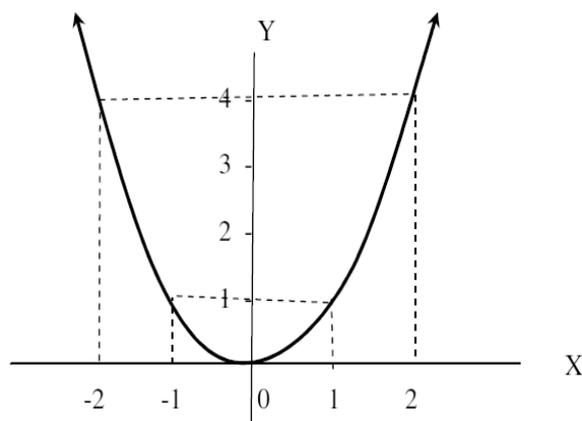
Gambar 1



❖ Data Kontinu (ukuran)

Data kontinu adalah data yang nilainya ada pada suatu interval tertentu atau berada pada nilai yang satu kenilai yang lainnya. Contohnya penggunaan kata sekitar, kurang lebih, kira-kira, dan sebagainya. Dinas pertanian daerah mengimpor bahan baku pabrik pupuk kurang lebih 850 ton.

Gambar 2



c) Jenis-jenis Data Menurut Waktu Pengumpulannya

❖ Data Cross Section

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

❖ **Data Time Series / Berkala**

Data berkala adalah data yang datanya menggambarkan sesuatu dari waktu ke waktu atau periode secara historis. Contoh data time series adalah data perkembangan nilai tukar dollar amerika terhadap euro eropa dari tahun 2004 sampai 2006, dll.

d) Macam-Macam Data Berdasarkan Sumber Data

❖ **Data Internal**

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

❖ **Data Eksternal**

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

e) Jenis Data Menurut Cara Memperolehnya

❖ **Data Primer**

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

❖ **Data Sekunder**

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

2. Skala Pengukuran Pada Data

a) SKALA NOMINAL (KLASIFIKASI)

Skala nominal merupakan skala pengukuran yang paling rendah tingkatannya di antara ke empat skala pengukuran yang lain. Seperti namanya, skala ini membedakan satu obyek dengan obyek lainnya

berdasarkan lambang yang diberikan. Ciri data yang dihasilkan adalah posisi data setara (pegawai negeri tidak lebih tinggi dari wiraswasta meskipun angka tandanya berbeda).

Contoh : Data mengenai barang-barang yang dihasilkan oleh sebuah mesin dapat digolongkan dalam kategori cacat atau tidak cacat. Barang yang cacat bisa diberi angka 0 dan yang tidak cacat diberi angka 1. Data 1 tidaklah berarti mempunyai arti lebih besar dari 0. Data satu hanyalah menyatakan lambang untuk barang yang tidak cacat.

Bilangan dalam Skala Nominal berfungsi hanya sebagai lambang untuk membedakan, terhadap bilangan-bilangan tersebut tidak berlaku hukum aritmetika, tidak boleh menjumlahkan, mengurangi, mengalikan, maupun membagi. \neq dan $=$ adalah hubungan sama dengan dan tidak sama dengan. Statistik yang sesuai dengan data berskala Nominal adalah Statistik Nonparametrik. Contoh perhitungan statistik yang cocok adalah Modus, Frekuensi dan Koefisien Kontingensi.

b) SKALA ORDINAL (RANGKING)

Skala pengukuran berikutnya adalah skala pengukuran ordinal. Skala pengukuran ordinal mempunyai tingkat yang lebih tinggi dari skala pengukuran nominal. Dalam skala ini, terdapat sifat skala nominal, yaitu membedakan data dalam berbagai kelompok menurut lambang, ditambah dengan sifat lain yaitu, bahwa satu kelompok yang terbentuk mempunyai pengertian lebih (lebih tinggi, lebih besar,...) dari kelompok lainnya. Oleh karena itu, dengan skala ordinal data atau obyek memungkinkan untuk diurutkan atau dirangking. Ciri data yang dihasilkan nominal adalah posisi data tidak setara (contoh pangkat seorang TNI diatas, Mayor lebih tinggi dari Kapten, dan Kapten lebih tinggi dari Letnan) dan tidak dapat dilakukan operasi matematika (misalkan pada tingkat kepuasan konsumen : $2 + 3 = 5$, yang berarti tidak puas + cukup puas = sangat puas).

Contoh : Sistem kepangkatan dalam dunia militer adalah satu contoh dari data berskala ordinal Pangkat dapat diurutkan atau dirangking dari Prajurit sampai Sersan berdasarkan jasa, dan lamanya pengabdian.

c) SKALA INTERVAL

Skala pengukuran Interval adalah skala yang mempunyai semua sifat yang dipunyai oleh skala pengukuran nominal, dan ordinal ditambah dengan satu sifat tambahan. Dalam skala interval, selain data dapat dibedakan antara yang satu dengan yang lainnya dan dapat dirangking, perbedaan (jarak/interval) antara data yang satu dengan data yang lainnya dapat diukur.

Contoh : Data tentang suhu empat buah benda A, B, C, dan D yaitu masing-masing 20, 30, 60, dan 70 derajat Celcius, maka data tersebut adalah data dengan skala pengukuran interval karena selain dapat dirangking, peneliti juga akan tahu secara pasti perbedaan antara satu data dengan data lainnya. Perbedaan data suhu benda pertama dengan benda kedua misalnya, dapat dihitung sebesar 10 derajat, dst.

Bilangan pada skala interval fungsinya ada tiga yaitu :

- 1) Sebagai lambang untuk membedakan
- 2) Untuk mengurutkan peringkat, misal, makin besar bilangannya, peringkat makin tinggi ($>$ atau $<$).
- 3) Bisa memperlihatkan jarak/perbedaan antara data obyek yang satu dengan data obyek yang lainnya. Titik nol bukan merupakan titik mutlak, tetapi titik yang ditentukan berdasarkan perjanjian.

Statistik yang sesuai dengan data berskala Interval adalah Statistik Nonparametrik dan Statistik Parametrik. Contoh perhitungan statistik yang cocok adalah Rata-rata, Simpangan Baku, dan Korelasi Pearson.

d) SKALA RASIO

Skala rasio merupakan skala yang paling tinggi peringkatnya. Semua sifat yang ada dalam skala terdahulu dipunyai oleh skala rasio. Sebagai tambahan, dalam skala ini, rasio (perbandingan) antar satu data dengan data yang lainnya mempunyai makna.

Contoh : Data mengenai berat adalah data yang berskala rasio. Dengan skala ini kita dapat mengatakan bahwa data berat badan 80 kg adalah 10 kg lebih berat dari yang 70 kg, tetapi juga dapat mengatakan bahwa data 80 kg

adalah 2x lebih berat dari data 40 kg. Berbeda dengan interval, skala rasio mempunyai titik nol yang mutlak.

Bilangan pada skala Rasio fungsinya ada tiga yaitu :

- 1) Sebagai lambang untuk membedakan
- 2) Untuk mengurutkan peringkat, misal, makin besar bilangannya, peringkat makin tinggi (> atau <),
- 3) Bisa memperlihatkan jarak/perbedaan antara data obyek yang satu dengan data obyek yang lainnya.
- 4) Rasio (perbandingan) antar satu data dengan data yang lainnya dapat diketahui dan mempunyai arti. Titik nol merupakan titik mutlak.

Statistik yang sesuai dengan data berskala Rasio adalah Statistik Nonparametrik dan Statistik Parametrik. Contoh perhitungan statistik yang cocok adalah Rata-rata kur, Koefisien Variasi dan statistik-statistik lain yang menuntut diketahuinya titik nol mutlak.

D. POPULASI DAN SAMPEL

Populasi adalah seluruh objek yang menjadi sasaran penelitian atau pengamatan dan memiliki sifat-sifat yang sama. Sampel adalah bagian dari populasi yang diambil untuk dijadikan objek pengamatan langsung dan dijadikan dasar dalam pengambilan kesimpulan. Dengan kata lain, populasi adalah himpunan keseluruhan objek yang diteliti, sedangkan sampel adalah bagian yang di ambil dari populasi.

Contoh-contoh populasi dan sampel :

Untuk mengetahui prestasi matematika SMP kelas IX di provinsi DKI Jakarta, dicatat prestasi dari beberapa sekolah di masing-masing kotamadya (Jakarta Pusat, Jakarta Selatan, Jakarta Barat, dan Jakarta Timur).

- Populasi : seluruh siswa SMP kelas IX di provinsi DKI Jakarta.
- Sampel : siswa SMP kelas IX dari beberapa sekolah di masing-masing kotamadya.

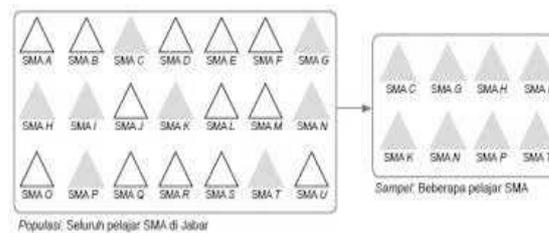
Penelitian ada dua macam yaitu sensus dan sampling. Sensus adalah penelitian yang melibatkan keseluruhan anggota populasi. Sampling adalah penelitian yang hanya melibatkan sebagian anggota populasi.

1. Teknik Penarikan Sampel

Teknik penarikan sampel merupakan salah satu proses yang penting dalam melakukan sebuah penelitian. Karena kesalahan dalam penarikan sample dapat mengakibatkan ketidaksesuaian hasil data penelitian dengan kenyataan. Ada 4 teknik penarikan sampel yang sering digunakan oleh para peneliti :

a) Sampel acak sederhana (Random)

Gambar 3



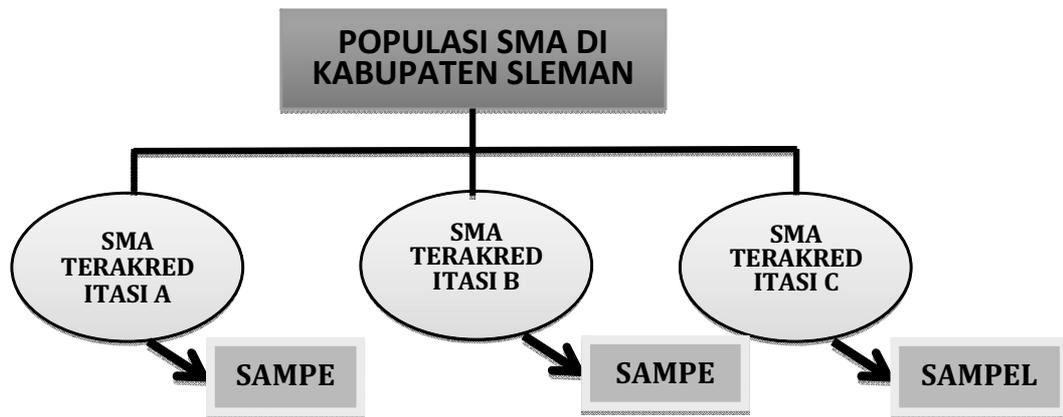
Untuk menghilangkan kemungkinan bias, kita perlu mengambil sampel random sederhana atau sampel acak. Pengambilan sampel dari semua anggota populasi dilakukan secara acak tanpa memperhatikan strata yang ada dalam anggota poipulasi. Hal ini dapat dilakukan apabila anggota populasi dianggap homogen. Prosedurnya :

- 1) Susun "sampling frame"
- 2) Tetapkan jumlah sampel yang akan diambil
- 3) Tentukan alat pemilihan sampel
- 4) Pilih sampel sampai dengan jumlah terpenuhi

b) Sampel stratifikasi

Teknik ini digunakan apabila populasi mempunyai anggota/karakteristik yang tidak homogen dan berstrata secara proportional. Sebagai contoh suatu organisasi mempunyai personil yang terdiri dari latar belakang pendidikan yang berbeda yaitu: SMP, SMA, S1, dan S2 dengan jumlah setiap kelas pendidikan juga berbeda. Jumlah anggota populasi untuk setiap strata pendidikan tidak sama atau bervariasi. Jumlah sampel yang harus diambil harus meliputi strata pendidikan yang ada yang diambil secara proporsional.

Gambar 4



c) Sampel sistematis

Teknik sampling ini merupakan teknik penarikan sampel dengan cara penentuan sampel berdasarkan urutan dari anggota populasi yang telah diberi nomor urut. atau teknik penarikan sampel yang mengambil setiap unsure ke- k dalam populasi, untuk dijadikan contoh dengan titik awal ditentukan secara acak diantara k unsur yang pertama. Sebagai contoh jumlah anggota populasi sebanyak 200 orang. Anggota populasi diberi nomor urut dari no 1 sampai nomor 200. Selanjutnya pengambilan sampel dilakukan dengan memilih nomor urut ganjil, atau genap saja, atau kelipatan dari bilangan tertentu, seperti bilangan 5 dan lainnya.

d) Sampel kelompok (cluster)

Teknik sampling daerah (cluster sampling) digunakan untuk menentukan sampel bila obyek yang akan diteliti atau sumber data sangat luas, misalnya penduduk suatu negara, propinsi atau kabupaten. Untuk menentukan penduduk mana yang akan dijadikan sumber data, maka pengambilan sampelnya berdasarkan daerah dari populasi yang telah ditetapkan.

Sebagai contoh Indonesia terdiri dari 33 propinsi, sampel yang akan diambil sebanyak 5 propinsi, maka pengambilan 5 propinsi dari 30 propinsi dilakukan secara random. Suatu hal yang perlu diingat adalah bahwa karena propinsi yang ada di Indonesia juga berstrata, maka pengambilan sampel untuk 5 propinsi juga dilakuykan dengan menggunakan teknik stratified random sampling. Teknik cluster sampling dilakukan dalam dua tahap yaitu: (1) menentukan sampel daerah, dan (2) menentukan orang-orang yang ada pada daerah dengan cara sampling juga.

Penyajian Data

A. TABEL

Penyajian tabel digunakan sebagai pilihan yang sering dipakai oleh peneliti atau penyaji informasi. Pengolahan data untuk keperluan analisis awal atau analisis lanjutan akan lebih baik apabila di sajikan terlebih dahulu dalam tabel yang baik.

1. TABEL REFERENSI DAN TABEL IKHTIAR

Tabel referensi memiliki fungsi sebagai “gudang keterangan” karena tabel tersebut menyajikan keterangan yang rinci dan disusun secara khusus untuk kepentingan referensi. **Misal**, tabel-tabel yang terdapat dalam laporan sensus umumnya merupakan tabel yang memberikan keterangan secara umum bagi kepentingan referensi. Seringkali tabel semacam ini disebut tabel umum (*general table*),

Tabel ikhtisar atau juga dinamakan tabel naskah (*text table*), umumnya berbentuk singkat, sederhana dan mudah dimengerti. Fungsi tabel ikhtisar adalah memberi gambaran yang sistematis tentang peristiwa-peristiwa yang merupakan hasil penelitian atau observasi. Tabel ikhtisar dapat berasal dari tabel referensi atau dari beberapa tabel ikhtisar yang lainnya. **Tabel ikhtisar banyak digunakan** dalam penulisan laporan perusahaan maupun tulisan ilmiah. Salah satu jenis tabel ikhtisar adalah tabel yang isinya menggambarkan perbandingan. Angka yang dibandingkan tentu saja diletakkan dalam kolom yang berdampingan. Jika angka-angka absolut yang diperbandingkan terlalu

besar, maka dapat disajikan dalam bentuk rasio atau persentase untuk lebih memudahkan. Stressing atau penekanan hal-hal yang dianggap penting dapat dilakukan dengan cara meletakkan angka-angka dalam kolom yang berada di sisi kiri dan yang tidak diberi penekanan diletakkan dalam kolom yang berada di sisi kanannya.

2. CARA PENYUSUNAN POS-POS KETERANGAN DALAM KOMPARTIMEN TABEL

a. *Penyusunan secara alfabetis*

Tabel ini menyajikan data berdasarkan kolom nama kompartemen yang disusun secara alfabetis dimulai dari alfabet paling awal yang ada dalam kolom tersebut (ascending).

b. *Penyusunan secara geografis*

Tabel ini menyajikan data berdasarkan kolom nama kompartemen yang disusun secara geografis dimulai dari lokasi paling barat, misalnya di Indonesia adalah provinsi Banda Aceh.

c. *Penyusunan menurut besaran angka-angka*

Tabel ini menyajikan data berdasarkan kolom yang diberikan penekanan dan disusun menurut besarnya angka-angka, baik dari kecil ke besar (ascending) maupun dari besar ke kecil (descending).

d. *Penyusunan secara historis*

Data yang disajikan dalam tabel diklasifikasikan secara kronologis atau historis, biasanya dimulai dari waktu yang paling dahulu atau paling lama.

e. *Penyusunan atas dasar kelas-kelas yang lazim*

Penyajian data dalam tabel dimana penyusunan pos-pos keterangan dalam kompartemen tabel dilakukan berdasarkan kelas-kelas yang umum digunakan dalam dunia statistik. Misalnya Impor, seringkali dibagi ke dalam tiga kategori ekonomi, yaitu: a. barang konsumsi, b. bahan mentah serta bahan pelengkap, dan c. barang modal.

f. *Penyusunan secara progresif*

Pada tabel ini, penyusunan pos-pos keterangan dalam kompartemen tabel harus dilakukan sedemikian rupa agar angka akhir dari tiap pos harus merupakan hasil perkembangan angka-angka yang telah ada sebelumnya.

Cara penyusunan yang digunakan dalam menyusun pos-pos keterangan dalam kompartemen tabel harus diusahakan agar tabel referensi disusun untuk tujuan referensi, sedangkan tabel ikhtisar disusun untuk tujuan perbandingan serta penekanan pada pospos yang dianggap penting oleh penyusun.

3. STRUKTUR TABEL STATISTIK

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

a. *Nama/titel dan identifikasi*

Tabel yang baik harus memiliki nama (titel) dan nama tersebut harus diletakkan di atas tabel. Nama tabel harus jelas dan singkat, jika tidak maka yang utama adalah kesederhanaan sedangkan catatan-catatan tambahan dapat diberikan dalam catatan di bawah tabel. Umumnya susunan redaksi nama harus menggambarkan tentang ciri-ciri data yang terdapat dalam tabel.

b. *Catatan pendahuluan (prefatory note) dan catatan di bawah tabel (foot-note)*

Catatan pendahuluan dan catatan yang terdapat di bawah tabel sebetulnya merupakan bagian yang integral dari sebuah tabel. Catatan pendahuluan biasanya disimpan langsung dibawah nama tabel dalam bentuk yang kurang menonjol dibandingkan dengan nama tabel.

c. *Sumber data*

Sumber data, umumnya ditempatkan langsung di bawah tabel setelah catatan. Sumber data harus diusahakan selengkap mungkin, berisi keterangan penulis, nama buku, jilid dan halaman buku, penerbit, dan lain-lain yang tidak meragukan. Jika data diambil dari data sekunder, sumber primer serta sumber sekundernya harus disebutkan secara lengkap.

d. *Presentase*

Bila angka presentase digunakan dalam tabel, maka pos-pos keterangan dalam kompartemen tabel harus rinci dan jelas. Istilah

'presentase' yang meragukan dapat dihindari, misalnya dengan menggunakan istilah 'presentase dari jumlah', 'presentase dari penambahan atau penurunan', dsb.

e. Jumlah

Jika jumlah angka merupakan hal yang penting, maka jumlah tersebut harus diletakkan pada sisi atas dalam kompartemen tabel atau sisi kiri dalam nama kolom. Cara lain adalah dengan menuliskannya dalam huruf tebal. Jika dianggap tidak penting maka dapat diletakkan di bawah kompartemen atau pada sisi kanan nama kolom.

f. Unit

Unit pengukuran angka-angka yang terdapat dalam kolom tabel harus jelas dan tidak meragukan. Jika tidak, maka ciri-ciri unit pengukurannya harus dijelaskan dalam nama kompartemen atau nama kolom.

g. Bentuk tabel

Tabel sebaiknya jangan terlalu panjang atau terlalu pendek, tetapi disesuaikan dengan ruang laporan dimana tabel diletakkan.

1) Tabel mendatar

Bentuk tabel ini ditentukan oleh beberapa faktor sebagai berikut:

- a) Lebar kompartemen tabel, yang ditentukan oleh pos-pos keterangan yang terpanjang.
- b) Lebar tiap kolom, yang ditentukan oleh jumlah angka yang terbesar.
- c) Cara mengatur spasi kata-kata.
- d) Cara mengatur tepi.

2) Tabel vertikal

Bentuk tabel ini 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 dan mengatur tepi

B. DIAGRAM/GRAFIK

1. FUNGSI GRAFIK STATISTIK

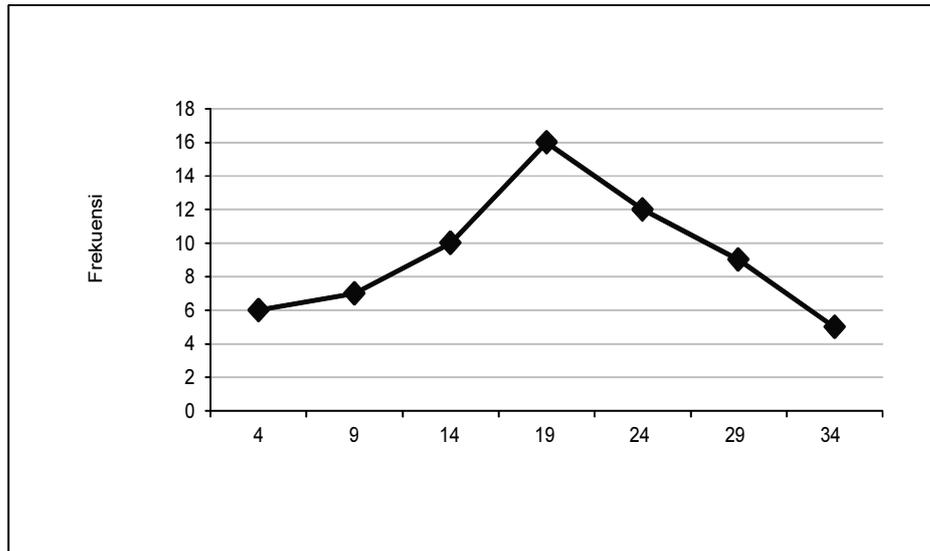
Data statistik dapat disajikan dalam bentuk tabel atau grafik. Penyajian data dalam bentuk grafik umumnya lebih menarik perhatian dan mengesankan. Penyajian data statistik secara grafis mempunyai berbagai fungsi. Grafik atau diagram seringkali digunakan dalam iklan dengan maksud agar konsumen memperoleh kesan yang mendalam terhadap ciri-ciri produk yang diiklankan. Kegiatan produksi lebih mudah dilihat dan dipelajari secara visual bila dinyatakan dalam angka-angka dan digambarkan secara grafis. Peta pengawasan kualitas merupakan alat yang penting dalam melakukan pengawasan produk maupun pengawasan proses produksi. Grafik penjualan suatu perusahaan memberi gambaran yang sederhana dan menarik mengenai perkembangan hasil penjualan yang telah dicapai oleh perusahaan yang bersangkutan. Pada hakekatnya grafik dan tabel seyogyanya digunakan secara bersama-sama. Grafik statistik lebih mudah dan menarik dibanding tabel statistik. Selain itu, grafik dapat melukiskan suatu peristiwa secara lebih mengesankan dan tidak membosankan. Namun demikian, penyajian secara grafis hanyalah bersifat aprosimatif. Angka-angka yang pasti dan rinci tentang suatu peristiwa dimuat dalam tabel. Oleh karena itu, analisis dan interpretasi data umumnya dilakukan terhadap data yang terdapat dalam tabel statistik.

2. JENIS GRAFIK STATISTIK

a) *Diagram garis*

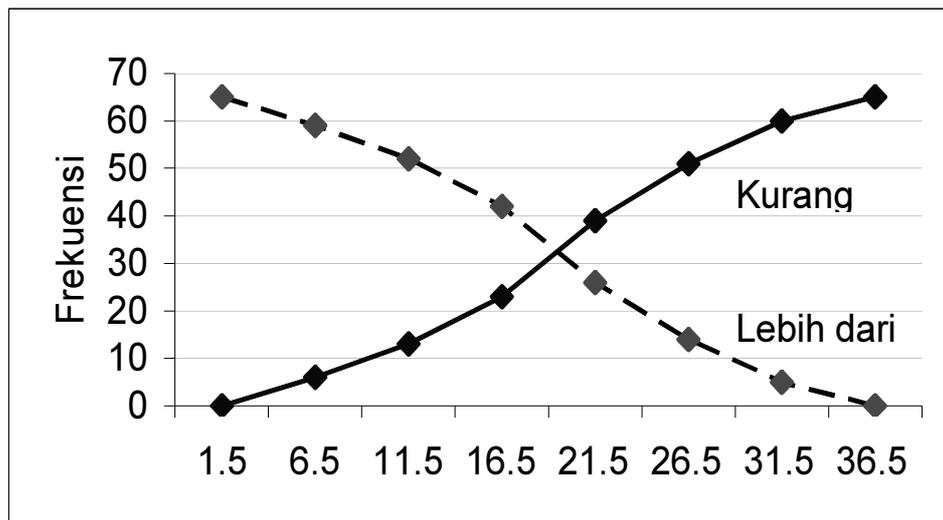
Diagram garis sering disebut juga peta garis (line chart) atau kurva (curve), merupakan bentuk penyajian yang paling banyak dipakai dalam berbagai laporan perusahaan maupun penelitian ilmiah. Data statistik dapat diklasifikasikan atas ciri-ciri kronologis, geografis, kuantitatif maupun kualitatif. Salah satu bentuk data yang dapat diklasifikasi secara kronologis adalah data deret berkala (time series). Sebagian besar distribusi data dapat diklasifikasi secara kuantitatif dalam bentuk distribusi frekuensi. Hasil kedua cara klasifikasi tersebut dapat digambarkan secara visual dalam bentuk kurva. Sedangkan data yang diklasifikasikan berdasarkan geografis maupun kualitatif, jarang digambarkan dalam

bentuk kurva. Data demikian dapat digambarkan dengan peta balok (bar chart) atau bentuk peta lainnya.



Gambar 5: poligon frekuensi

Diagram Ogif dibuat dengan menghubungkan antara batas kelas interval dengan frekuensi kumulatif (jumlah frekuensi; lebih dari atau kurang dari batas kelas interval).



Gambar 6

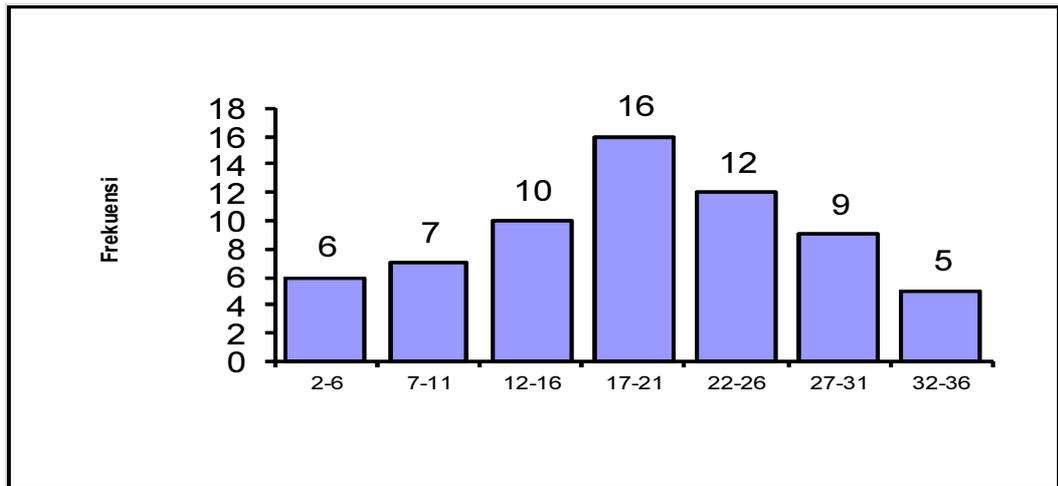
Kurva deret berkala : Metode untuk menggambarkan deret berkala secara visual tergantung pada jenis data yang akan disajikan. Data tersebut dapat

dibedakan ke dalam *data periode* (period data) dan *data titik* (point data). Data periode umumnya menggunakan periode waktu sebagai dasar pengukuran. Misalnya, data jumlah penjualan per bulan, rata-rata penjualan bulanan per tahun dan harga rata-rata selama tahun tertentu. Data titik menggunakan titik periode tertentu sebagai dasar pengukuran. Misalnya, nilai persediaan bahan baku pada suatu titik waktu tertentu dan harga-harga barang pada suatu titik waktu tertentu. Jika data kronologis dilukiskan dengan menggambarkan kurva, maka tahun, bulan, minggu, hari atau unit-unit kronologis lainnya harus dinyatakan pada sumbu mendatar, sedangkan variabel yang bergerak mengikuti waktu harus diletakkan pada sumbu vertikal. Jika *data periode* digambarkan dalam kurva, maka periode waktunya dapat diletakkan di bawah garis vertikal atau diletakkan di bawah spasi antara dua periode. Cara ini dipandang lebih baik karena memiliki kesan visual bahwa waktu atau periode sebenarnya memiliki durasi. Sedangkan jika *data titik* digambarkan dalam kurva, maka spasi periode harus dinyatakan pada sumbu mendatar dan observasinya harus diletakkan dalam spasi pada titik dimana peristiwanya terjadi.

Kurva distribusi frekuensi : Penggambaran grafik sebuah distribusi frekuensi umumnya dilakukan berdasarkan data kuantitatif yang terdapat dalam tabel distribusi frekuensi. Data yang terdapat dalam tabel distribusi frekuensi tersebut digambarkan dalam bentuk diagram kolom yang dinamakan histogram frekuensi. Diagram kolom atau histogram frekuensi ini harus dibedakan dengan diagram balok yang lebih umum dalam penggambaran peristiwa secara visual. Kurva distribusi frekuensinya dapat diperoleh dengan cara menghubungkan *titik tengah* (mid point) tiap-tiap kolom atau balok.

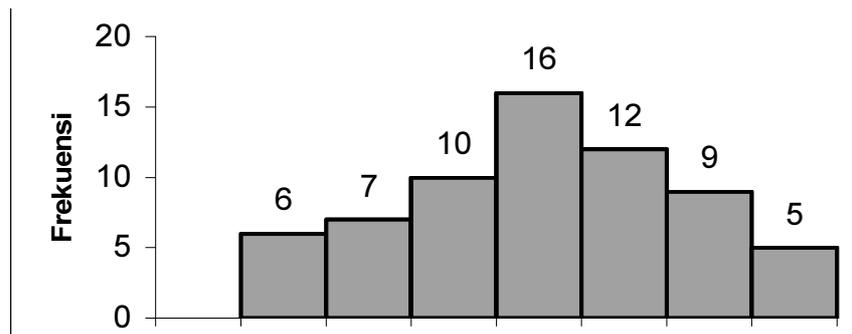
b) *Peta balok/diagram batang* (bar chart)

Diagram ini digunakan untuk memahami persoalan secara visual. Dalam diagram batang, lebar kelas diambil dari selang kelas distribusi frekuensi, sedangkan frekuensi masing-masing kelas ditunjukkan oleh tinggi batang.



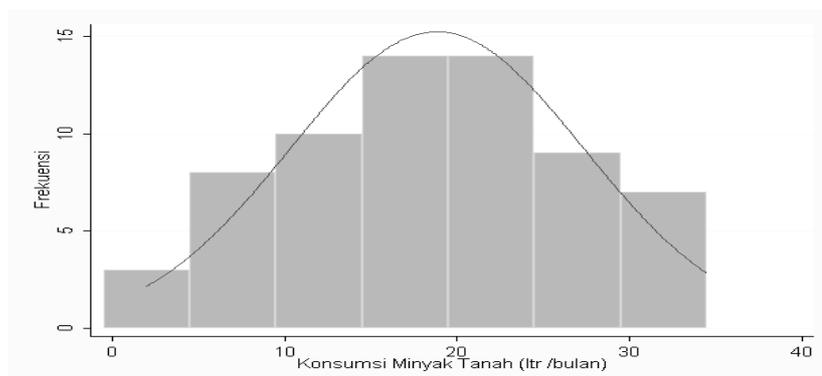
Gambar 7 :Diagram Batang

Diagram Histogram



Gambar 8. Distribusi frekuensi konsumsi minyak tanah oleh rumah tangga di Desa Sinduadi, 2007

Atau



Gambar 9. Distribusi frekuensi konsumsi minyak tanah oleh rumah tangga di Desa Sinduadi, 2007

Diagram histogram berbeda dengan diagram batang dalam hal lebar, yaitu batang digunakan batas kelas dan bukan limit kelasnya. Ini untuk menghilangkan jeda antar batang sehingga antar batang memberikan kesan "padat".

c) Diagram lingkaran (pie diagram)

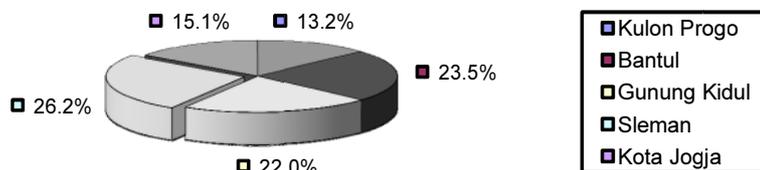
Diagram lingkaran biasanya digunakan untuk menyatakan perbandingan jika data terdiri atas beberapa kelompok atau kategori. Misal persentase penduduk di Wilayah DI Yogyakarta.

Contoh :

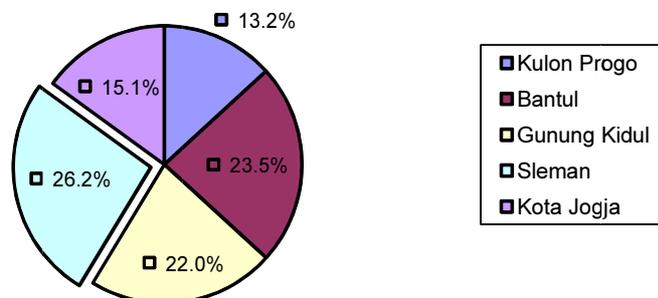
Tabel 1. Jumlah Penduduk DI Yogyakarta, 2006

Kode	Kab /Kota	Penduduk	Persentase
01	Kulon Progo	457.778	13,2%
02	Bantul	813.052	23,5%
03	Gunung Kidul	760.128	22,0%
04	Sleman	907.694	26,2%
71	Kota Jogja	520.780	15,1%
	DIY	3.459.432	100%

Sumber: Dinas Kependudukan DIY, 2007



Distribusi Penduduk di Provinsi DI Yogyakarta, 2006



Gambar Diagram Lingkaran

Contoh 1

Tabel Sederhana

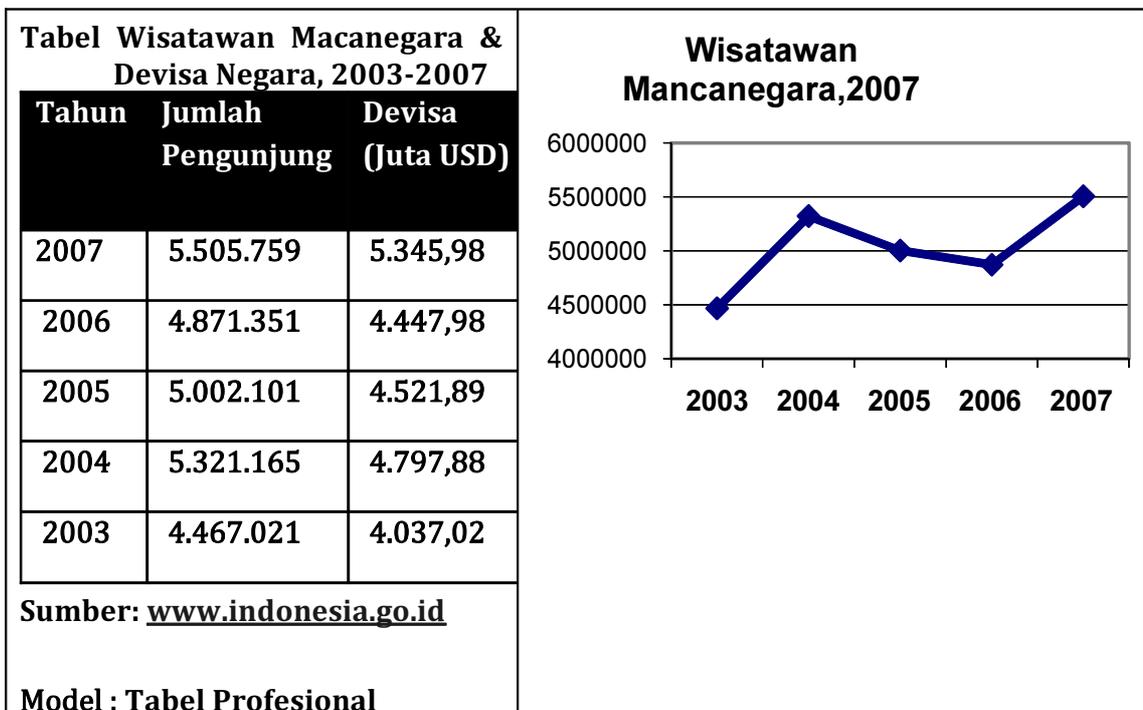
Tahun	Jumlah Pengunjung	Devisa (Juta USD)
2007	5.505.759	5.345,98
2006	4.871.351	4.447,98
2005	5.002.101	4.521,89
2004	5.321.165	4.797,88
2003	4.467.021	4.037,02

Sumber: www.indonesia.go.id

Keterangan Model: Simple1

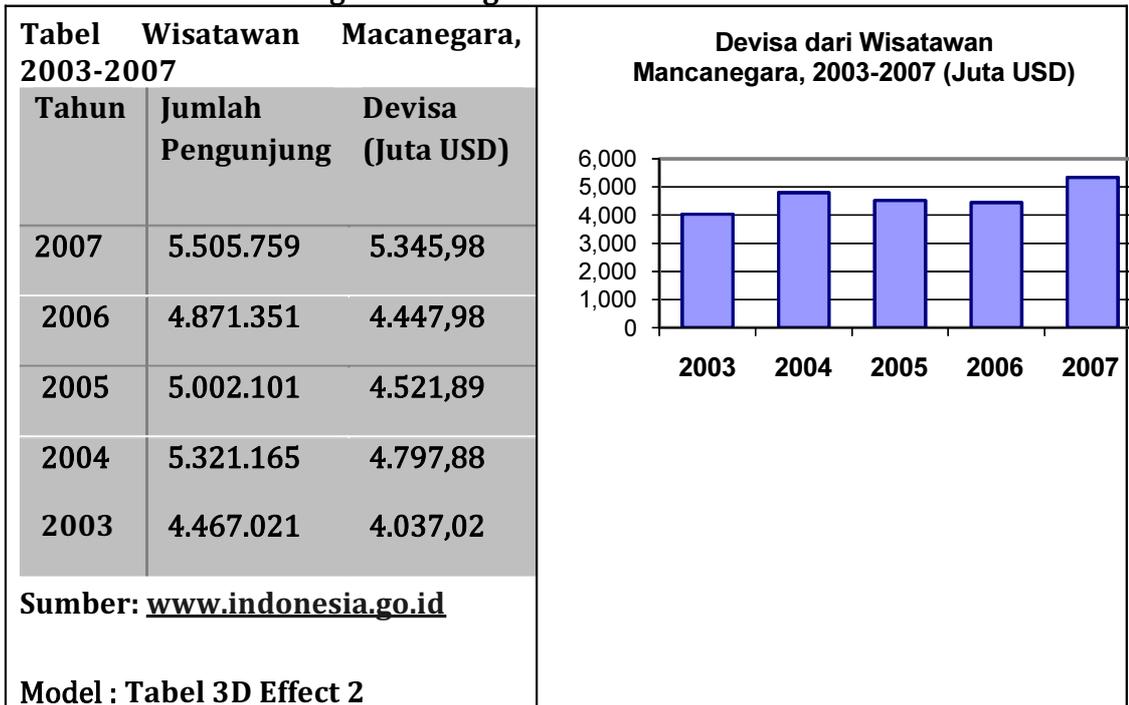
Contoh 2

Tabel Profesional dan Diagram Garis



Contoh 3

Tabel Effect 3D dan Diagram Batang



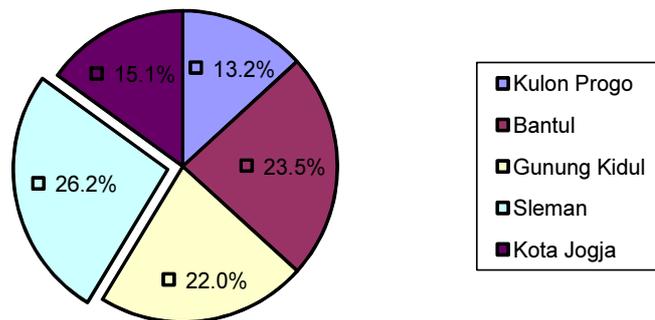
Contoh 4

Tabel Model Classic dan Grafik Lingkaran
Tabel Jumlah Penduduk di Yogyakarta, 2006

No	Kab/Kota	Jumlah				Total
		Laki-Laki	%	Perempuan	%	
1	Kulon Progo	223,613	48.8%	234,165	51.2%	457,778
2	Bantul	398,975	49.1%	414,077	50.9%	813,052
3	Gunung Kidul	371,385	48.9%	388,843	51.2%	760,128
4	Sleman	449,673	49.5%	458,021	50.5%	907,694
71	Kota Jogja	267,496	51.4%	253,284	48.6%	520,780
	DIY	1,711,142	49.5%	1,748,390	50.5%	3,459,432

Sumber:Kependudukan DIY, 2007

Diagram Lingkaran



3. BEBERAPA PERATURAN UMUM TENTANG MENGGANBAR GRAFIK

a) *Pemilihan jenis grafik*

Jenis grafik statistik yang akan disajikan oleh pembuat laporan harus dipilih agar dapat menyajikan gambaran mengenai suatu data secara efektif bagi pembaca. Jika dilihat dari fungsinya, setiap jenis grafik statistik memiliki kelebihan-kelebihan khusus. Namun demikian, grafik yang baik harus bersifat sederhana dan jelas. Grafik yang rumit biasanya disajikan untuk orang yang sangat mengerti permasalahan atau yang sangat mahir dalam ilmu statistik. Pemilihan jenis grafik yang akan disajikan oleh pembuat laporan tidak dapat semata-mata diserahkan pada kebijakan penggambar grafik, kecuali bila pembuat laporan yakin bahwa penggambar memiliki pengetahuan yang baik tentang data yang disajikan, tujuan penyajian, dan kemampuan pembaca dalam menarik kesimpulan dari grafik.

b) *Nama (titel), skala sumbu, sumber dan catatan*

Kegunaan serta pengaturan nama, sumber dan catatan dalam sebuah tabel berlaku juga untuk grafik statistik. Nama grafik dapat diletakkan di atas atau di bawah gambar grafik. Meski demikian banyak statistisi berpendapat bahwa peletakan nama di atas grafik akan lebih efektif jika dibandingkan dengan di bawah grafik. Skala *horizontal* dan *vertical* dalam peta garis, diagram kolom, dan peta balok sebenarnya memiliki kesamaan dalam arti dengan nama kolom dan kompartemen dalam tabel statistik.

c) *Skala dan garis kisi-kisi*

Jarak yang sama pada skala grafik sebenarnya menyatakan jarak nilai yang sama pula. Nilai skala bertujuan memberi gambaran yang aproksimatif tentang jumlah kuantitatif, sedangkan jumlah yang eksak dan rinci secara seksama harus dibaca dari tabel statistiknya.

Garis kisi-kisi harus digambarkan secara lebih tipis dari pada garis skalanya. Peta garis umumnya memiliki garis kisi-kisi baik yang bersifat mendatar maupun vertikal. Peta kolom hanya membutuhkan garis kisi-kisi yang mendatar. Peta balok mendatar membutuhkan garis kisi-kisi vertikal. Pada beberapa penyajian grafik, garis kisi-kisi demikian dapat juga tidak digambarkan sama sekali atau hanya digambarkan secara sebagian saja.

d) *Pemberian tekanan pada penggambaran grafik*

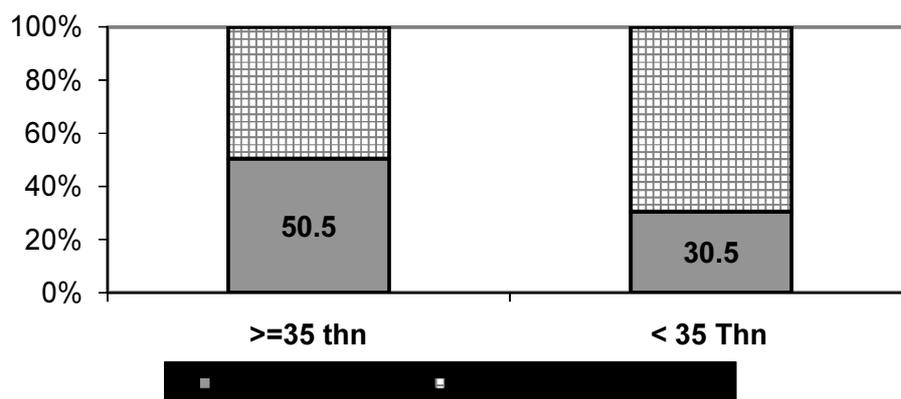
Penekanan tentang suatu peristiwa yang tertentu dalam penyajian grafik dapat dilakukan dengan cara memberi warna yang berbeda, tanda silang, atau garis yang berbeda. Garis dalam peta yang sama juga dapat dibedakan dengan menggunakan warna yang berbeda, garis terputus-putus, garis padat (solid line) atau garis tebal. Garis padat lebih memberi tekanan dari pada garis terputus-putus, sedangkan garis tebal lebih menarik perhatian dari pada garis yang tipis.

LATIHAN

1. Tahun 2012 sebuah perusahaan otomotif sedang melakukan penelitian tentang orang yang memakai kendaraan bermotor berdasarkan merk (Honda, Suzuki, Kawasaki, Yamaha, Mocin). Dari 100 subyek di dapat data sebagai berikut :

Kendaraan	jumlah
Honda	39
Suzuki	20
Kawasaki	19
Yamaha	18
Mocin	4
Total	100

- a. Gambarlah grafik masing-masing merk untuk 5 tahun kedepan, jika setiap tahunnya bertambah 5% .
 - b. Tentukan besarnya prosentase penggunaan merk sepeda motor dalam bentuk pie chart.
2. Jelaskan Perbedaan statistik deskriptif dan statistik inferensial!
 3. Jelaskan pengertian Sampel dan Populasi!
 4. Jelaskan Pengertian Data, Variabel dan Parameter!
 5. Penyajian Data dalam format tabel & grafik memiliki beberapa tujuan. Jelaskan tujuan ke 2 penyajian data tersebut!
 6. Narasikan tabel IPM tingkat Provinsi Seluruh Indonesia! Informasikan kepada pembaca agar dapat mengerti informasi berdasar tabel tersebut!
 7. Berdasarkan Jurnal Pemanfaatan Buku Pelajaran oleh Siswa dan Guru SD tahun 2008; diberikan hasil grafik berikut ini:



Gambar 1. Karakteristik Responden menurut Golongan Umur di Wilayah Kerja UPTD Dinas Pendidikan Kedungadem, Kabupaten Bojonegoro tahun 2007.

Berilah Narasi untuk menjelaskan kepada pembaca mengenai grafik tersebut!

Rank	Provinsi	Usia harapan hidup (tahun)	% melek huruf, dewasa	Rata-rata lama pendidikan (tahun)	Pengeluaran per kapita (000 Rp)	IPM
1	Jakarta*	71	98	9.7	593	72.5
2	Yogyakarta	71	85	7.9	598	68.7
3	Kalimantan Timur	69	94	7.8	578	67.8
4	Riau	68	96	7.3	580	67.3
5	Maluku	67	96	7.6	577	67.2
10	Bali	70	83	6.8	588	65.7
12	Aceh	68	93	7.2	563	65.3
13	Bengkulu	65	93	7.0	577	64.8
14	Jawa Tengah	68	85	6.0	584	64.6
15	Jawa Barat	64	92	6.8	584	64.6
22	Jawa Timur	66	81	5.9	579	61.8
23	Kalimantan Barat	64	83	5.6	571	60.6
24	Nusa Tenggara Timur	64	81	5.7	577	60.4
25	Papua (Irian Jaya)	65	71	5.6	580	58.8
26	Nusa Tenggara Barat	58	73	5.2	566	54.2

Distribusi Frekuensi

A. TABEL DISTRIBUSI FREKUENSI

Statistik Distribusi Frekuensi merupakan rumus statistik deskriptif yang dapat digunakan untuk mengetahui distribusi frekuensi gejala dalam satu variabel. Untuk mempermudah memahami karakteristik suatu data observasi, data tersebut dikelompokkan ke dalam beberapa kelompok (kelas) yang mana masing-masing kelas menampung sebagian data observasi. Penyajian data observasi yang sudah dikelompokkan ini disusun ke dalam suatu tabel yang disebut tabel distribusi frekuensi. Tabel distribusi frekuensi dibuat bertujuan agar data observasi tersebut lebih mudah dipahami. Dalam tabel distribusi frekuensi terdapat beberapa kelas yang masing-masing kelas menampung sejumlah data observasi.

Menurut Algifari (1994:8) langkah-langkah yang dilakukan untuk membuat tabel distribusi frekuensi adalah sebagai berikut :

1. Urutkan data dari nilai data tertinggi ke nilai data terendah.
2. Tentukan jumlah kelas yang akan digunakan pada tabel distribusi.

Ada cara untuk menentukan jumlah kelas seperti dikemukakan oleh Sturges (Algifari, 1994:8), yaitu dengan menggunakan formula :

$$K = 1 + 3,33 \log N$$

yang menyatakan bahwa :

K = jumlah kelas.

N = banyaknya data observasi.

3. Menentukan interval kelas.

Besarnya interval kelas yang digunakan pada tabel distribusi frekuensi juga bebas ditentukan oleh pembuatnya. Akan tetapi perlu diingat bahwa besarnya interval kelas untuk semua kelas adalah sama.

Ada formula yang dapat digunakan untuk menentukan besarnya interval kelas, yaitu :

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

yang menyatakan bahwa :

C_i = interval kelas

R = selisih nilai data tertinggi dengan nilai data terendah (Range)

K = jumlah kelas.

4. Menyusun data ke dalam tabel distribusi frekuensi.

Sebelum menyusun data ke dalam tabel distribusi frekuensi terlebih dahulu ditentukan nilai terendah pada kelas yang pertama. Misalnya menentukan nilai terendah dari kelas yang pertama terlalu kecil, dengan jumlah kelas dan interval kelas yang sudah ditentukan, sehingga kelas yang pertama tersebut tidak menampung data observasi (frekuensi kelasnya nol).

CONTOH KASUS 1.

Berikut ini adalah data mengenai nilai 30 orang peserta ujian Statistik di UMB Yogyakarta :

60	55	61	72	59	49
57	65	78	66	41	52
42	47	50	65	74	68
88	68	90	63	79	56
87	65	85	95	81	69

Buatlah tabel distribusi frekuensi mengenai nilai 30 peserta ujian Statistik tersebut.

Untuk menjawab kasus 1, langkah-langkah yang harus dilakukan :

1. Urutkan data observasi dari nilai terendah ke nilai tertinggi

41	60	72
42	61	74
47	63	78
49	65	79
50	65	81
52	65	85
55	66	87
56	68	88
57	68	90
59	69	95

2. Menentukan jumlah kelas pada tabel distribusi frekuensi.

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

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

$$K = 1 + 3,3(1,48)$$

$$K = 1 + 4,8745$$

$$K = 5,8745 \approx 6$$

3. Menentukan interval kelas (C_i) dengan formula

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

$$R = 95 - 41 = 54$$

$$K = 5,8745$$

Sehingga

$$C_i = \frac{54}{5,8745} = 9,19$$

$$C_i = 9,19 \approx 10$$

4. Menyusun data observasi pada tabel distribusi frekuensi.

Untuk kasus 1, misalnya kita tentukan nilai terendah pada kelas pertama adalah 40, sehingga dengan interval kelas = 10, maka masing-masing kelas sebagai berikut :

Tabel 1

NILAI
40 - 49
50 - 59
60 - 69
70 - 79
80 - 89
90 - 99

Apabila data observasi merupakan bilangan pecahan (satu atau lebih angka dibelakang koma), maka kelas pada tabel distribusi frekuensi dapat dibuat seperti yang terlihat pada tabel 2 berikut ini :

Tabel 2

NILAI
40 < 49
50 < 59
60 < 69
70 < 79
80 < 89
90 < 99

Setelah selesai membuat kelas-kelas yang diperlukan dalam tabel distribusi frekuensi, maka masukkan semua data observasi ke dalam tabel tersebut, sehingga diperoleh tabel sebagai berikut :

NILAI	SCORE	BANYAKNYA DATA (FREKUENSI)
40 - 49	IIII	4
50 - 59	IIII I	6
60 - 69	IIII IIII	10
70 - 79	IIII	4
80 - 89	IIII	4
90 - 99	II	2
Σ		30

Setelah selesai menyusun data observasi ke dalam tabel, maka diperoleh suatu tabel distribusi frekuensi. Dari tabel distribusi frekuensi tersebut diperoleh beberapa informasi sebagai berikut :

1. Nilai terendah adalah antara 40 sampai dengan 49. Banyaknya peserta yang memperoleh nilai terendah tersebut sebanyak 4 orang.
2. Nilai tertinggi adalah antara 90 sampai dengan 99. Banyaknya peserta yang memperoleh nilai tertinggi tersebut sebanyak 2 orang.
3. Sebagian besar peserta memperoleh nilai antara 60 sampai dengan 69 sebanyak 10 orang.

CONTOH KASUS 2

Seorang kepala madrasah ingin mengetahui distribusi frekuensi siswa berdasarkan jenis kelamin, latar belakang kesartrian, dan kerajinan membayar SPP dengan data sebagai berikut.

NO	N A M A	X1	X2	X3
1	Abimanyu	1	1	1
2	Baladewa	1	2	1
3	Banowati Duryudana	2	3	3
4	Drupadi Puntadewa	2	3	2
5	D u r n a	1	2	2
6	Dursasana	1	3	2
7	Duryudana	1	2	2
8	Harjuna	1	1	2
9	Kr e s n a	1	1	2
10	Kunti Talibrata	2	1	2
11	Larasati Harjuna	2	1	1
12	Mustakaweni	2	3	3
13	N a k u l a	1	1	1
14	Puntadewa	1	1	1
15	S a d e w a	1	1	1
16	Sengkuni	1	3	3
17	Srikandi Harjuna	2	1	4
18	Surtikanti Karna	2	3	3
19	Utari Abimanyu	2	1	4
20	Werkudara	1	2	2

KETERANGAN

X1 = Jenis kelamin (1=Pria; 2=Wanita)

X2 = Kesantrian (1=Santri Total; 2=Santri Kalong; 3= Bukan Santri)

X3 = Kerajinan Membayar SPP (1=Sangat Rajin; 2=Rajin; 3=Malas;4=Sangat Malas)

Perhitungan:

Dari perhitungan data jenis kelamin (X1) diketahui distribusi frekuensinya sbb:

1. Siswa pria sebanyak 12 anak atau 60 persen.
2. Siswa wanita sebanyak 8 anak atau 40 persen.

Dari perhitungan data latar belakang kesantrian (X2) diketahui distribusi frekuensinya sbb:

1. Siswa yang berlatar belakang santri total sebanyak 10 anak atau 50 persen.
2. Siswa yang berlatar belakang santri kalong sebanyak 4 anak atau 20 persen.
3. Siswa yang berlatar belakang bukan santri sebanyak 6 anak atau 30 persen.

Dari perhitungan data kerajinan membayar SPP (X3) diketahui distribusi frekuensinya sbb:

1. Siswa yang sangat rajin membayar SPP sebanyak 6 anak atau 30 persen.
2. Siswa yang rajin membayar SPP sebanyak 8 anak atau 40 persen.
3. Siswa yang malas membayar SPP sebanyak 4 anak atau 20 persen.
4. Siswa yang sangat malas membayar SPP sebanyak 2 anak atau 10 persen.

Kesimpulan:

Siswa pria lebih banyak daripada siswa wanita.

1. Kebanyakan siswa berlatarbelakang santri, baik santri total maupun santrikalong; dalam hal ini jumlah siswa yang berlatar belakang santri total lebih dua kali lipat daripada santri kalong.
2. Kebanyakan siswa rajin dan sangat rajin membayar SPP; meski ada pula yang sangat malas membayar SPP.

CONTOH KASUS 3 :

Manajer Bengkel Hudson Auto berkeinginan melihat gambaran yang lebih jelas tentang distribusi biaya perbaikan mesin mobil. Untuk itu diambil 50 pelanggan sebagai sampel, kemudian di catat data tentang biaya perbaikan mesin mobilnya (\$). Berikut hasilnya :

91	78	93	57	75	52	99	80	97	62
71	69	72	89	66	75	79	75	72	76
104	74	62	68	97	105	77	65	80	109
85	97	88	68	83	68	71	69	67	74
62	82	98	101	79	105	79	69	62	73

Penyelesaian :

Banyaknya kelas (k) = 6

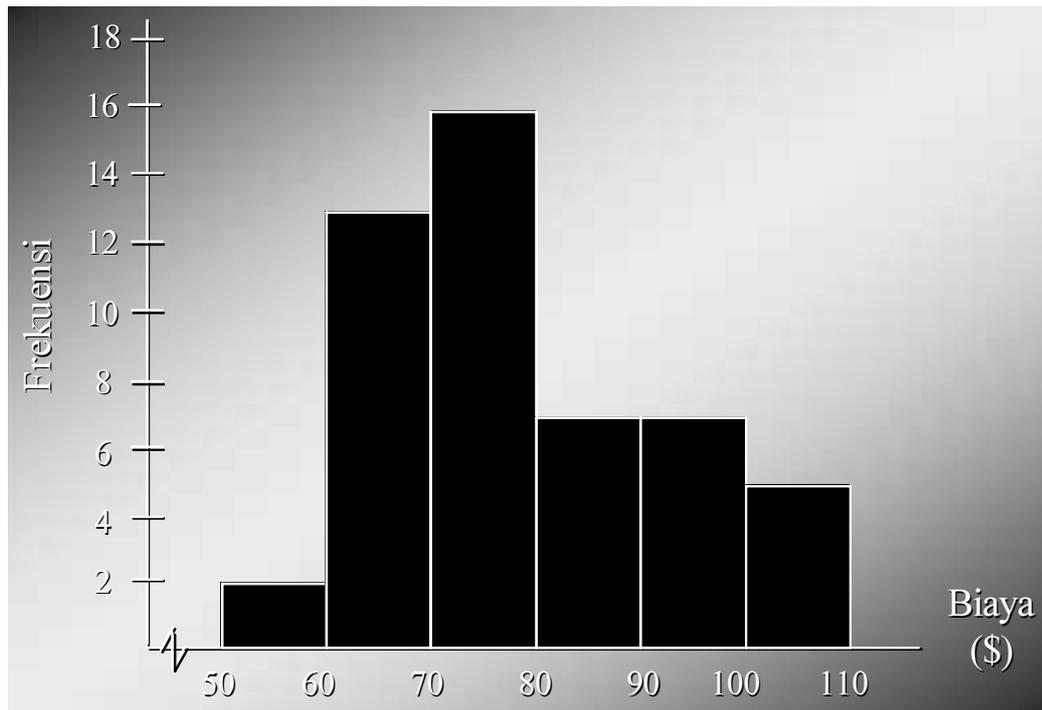
Panjang kelas (d) = $(109 - 52) / 6$
= 9,5 (dibulatkan menjadi 10)

Biaya (\$)	Frekuensi	Frekuensi relative	Frekuensi kumulatif	Frek. Relatif Kumulatif
50 – 59	2	0,04	2	0,04
60 – 69	13	0,26	15	0,30
70 – 79	16	0,32	31	0,62
80 – 89	7	0,14	38	0,76
90 – 99	7	0,14	45	0,90
100 – 109	5	0,10	50	1,00
Total	50	1,00		

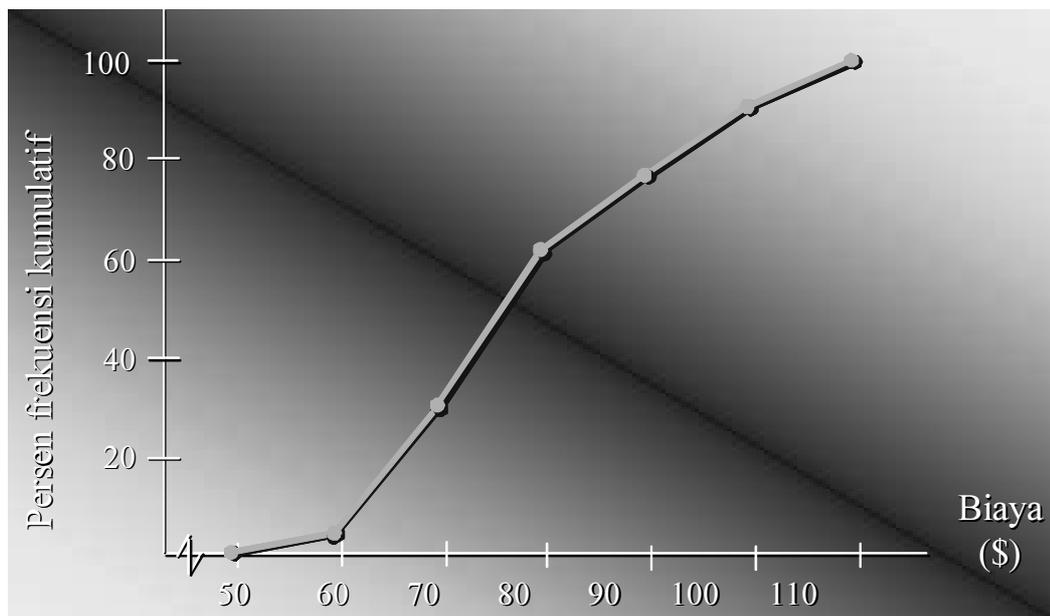
Analisis tabel distribusi frekuensi :

1. Hanya 4% pelanggan bengkel dengan biaya perbaikan mesin \$50-59.
2. 30% biaya perbaikan mesin berada di bawah \$70.
3. Persentase terbesar biaya perbaikan mesin berkisar pada \$70-79.
4. 10% biaya perbaikan mesin adalah \$100 atau lebih

Contoh : Bengkel Hudson



Ogive



B. STATISTIK TABULASI SILANG

1. KARAKTER

Statistik Tabulasi Silang merupakan rumus statistik deskriptif korelatif yang dapat digunakan untuk mengetahui distribusi frekuensi gejala dalam suatu variabel apabila variabel tersebut dihubungkan dengan variabel yang lain.

2. SPESIFIKASI

Statistik Tabulasi Silang efektif dijalankan untuk data yang tidak terlalu bervariasi.

3. CONTOH KASUS

Seorang kepala madrasah ingin mengetahui distribusi frekuensi siswa berdasarkan jenis kelamin, latar belakang kesantrian, dan kerajinan membayar SPP kalau ketiga variabel tersebut saling dihubungkan.

4. KETERANGAN

Statistik Tabulasi Silang hanya dapat dijalankan untuk dua atau lebih variabel.

CONTOH PERHITUNGAN

Permasalahan:

Seorang kepala madrasah ingin mengetahui distribusi frekuensi siswa berdasarkan jenis kelamin, latar belakang kesantrian, dan kerajinan membayar SPP kalau ketiga variabel tersebut saling dihubungkan.

NO	N A M A	X1	X2	X3
1	Abimanyu	1	1	1
2	Baladewa	1	2	1
3	Banowati Duryudana	2	3	3
4	Drupadi Puntadewa	2	3	2
5	Du r n a	1	2	2
6	Dursasana	1	3	2
7	Duryudana	1	2	2
8	Harjuna	1	1	2
9	K r e s n a	1	1	2
10	Kunti Talibrata	2	1	2
11	Larasati Harjuna	2	1	1
12	Mustakaweni	2	3	3
13	N a k u l a	1	1	1
14	Puntadewa	1	1	1
15	S a d e w a	1	1	1
16	Sengkuni	1	3	3
17	Srikandi Harjuna	2	1	4
18	Surtikanti Karna	2	3	3
19	Utari Abimanyu	2	1	4
20	Werkudara	1	2	2

KETERANGAN

X1 = Jenis kelamin (1=Pria; 2=Wanita)

X2 = Kesantrian (1=Santri Total; 2=Santri Kalong; 3= Bukan Santri)

X3 = Kerajinan Membayar SPP (1=Sangat Rajin; 2=Rajin; 3=Malas;4=Sangat Malas)

Perhitungan:

Hubungan antara jenis kelamin (X1) dengan latar belakang kesantrian siswa (X2) dapat dijelaskan sebagai berikut.

Tabel 1:

**HUBUNGAN ANTARA JENIS KELAMIN DENGAN
LATAR BELAKANG KESANTRIAN SISWA**

X1 \ X2	<u>1</u>	<u>2</u>	Σ
<u>1</u>	6	4	10
<u>2</u>	4	0	4
<u>3</u>	2	4	6
Σ	12	8	20

Penafsiran:

Dari perhitungan dalam Tabel 1 tersebut di atas dapat ditafsirkan hal-hal sebagai berikut.

1. Tidak ada seorang pun siswa wanita yang berlatar belakang sebagai santri kalong.
2. Separo dari keseluruhan siswa mempunyai latar belakang sebagai santri total.
3. Hanya ada 6 siswa atau 30 persen yang latar belakangnya bukan sebagai santri.

Hubungan antara jenis kelamin (X1) dengan kerajinan membayar SPP siswa (X3) dapat dijelaskan sebagai berikut.

Tabel 2:
HUBUNGAN ANTARA JENIS KELAMIN DENGAN
KERAJINAN MEMBAYAR SPP SISWA

X1 \ X3	<u>1</u>	<u>2</u>	Σ
<u>1</u>	5	1	6
<u>2</u>	6	2	8
<u>3</u>	1	3	4
<u>4</u>	0	2	2
Σ	12	8	20

Penafsiran:

Dari perhitungan dalam Tabel 2 tersebut di atas dapat ditafsirkan hal-hal sebagai berikut.

1. Para siswa pada umumnya rajin dan sangat rajin membayar SPP, meskipun ada pula yang sangat malas.
2. Siswa pria pada umumnya lebih rajin membayar SPP daripada siswa wanita.
3. Terdapat 2 siswa wanita atau 10 persen yang sangat malas membayar SPP.
4. Hanya ada 1 siswa pria atau 5 persen yang malas membayar SPP; dan tidak seorang pun yang sangat malas.

Hubungan antara latar belakang kesiantrian (X2) dengan kerajinan membayar SPP siswa (X3) dapat dijelaskan sebagai berikut.

Tabel 3:
HUBUNGAN ANTARA LATAR BELAKANG KESANTRIAN
DENGAN KERAJINAN MEMBAYAR SPP SISWA

X2 \ X3	<u>1</u>	<u>2</u>	<u>3</u>	Σ
<u>1</u>	5	1	0	6
<u>2</u>	3	3	2	8
<u>3</u>	0	0	4	4
<u>4</u>	2	0	0	2
Σ	10	4	6	20

Penafsiran:

Dari perhitungan dalam Tabel 3 tersebut di atas dapat ditafsirkan hal-hal sebagai berikut.

1. Para siswa pada umumnya rajin dan sangat rajin membayar SPP, meskipun ada pula yang sangat malas.
2. Siswa yang rajin dan sangat rajin membayar SPP umumnya berlatar belakang sebagai santri; baik santri total maupun santri kalong.
3. Tidak satu pun siswa yang berlatar belakang bukan santri yang sangat rajin atau sangat malas membayar SPP.

Selanjutnya hubungan antara jenis kelamin (X1), latar belakang kesantrian (X2), dengan kerajinan membayar SPP siswa (X3) dapat dijelaskan sebagai berikut.

Tabel 4:
HUBUNGAN ANTARA JENIS KELAMIN, KESANTRIAN,
DENGAN KERAJINAN MEMBAYAR SPP SISWA

X3 =>		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	Σ
X1	X2					
<u>1</u>	<u>1</u>	4	2	0	0	6
	<u>2</u>	1	3	0	0	4
	<u>3</u>	0	1	1	0	2
<u>2</u>	<u>1</u>	1	1	0	2	4
	<u>2</u>	0	0	0	0	0
	<u>3</u>	0	1	3	0	4
Σ		6	8	4	2	20

Penafsiran:

Dari perhitungan dalam Tabel 4 tersebut di atas dapat ditafsirkan hal-hal sebagai berikut.

1. Separo atau 50 persen dari siswa tersebut berlatar belakang sebagai santri total; di sisi lain tidak ada seorang siswa wanita pun yang berlatar belakang sebagai santri kalong.
2. Kebanyakan siswa, tepatnya 14 anak atau 70 persen, ternyata rajin dan sangat rajin membayar SPP.
3. Siswa yang berlatar belakang santri total dan santri kalong pada umumnya rajin dan sangat rajin membayar SPP, meskipun adasiswa wanita berlatar belakang santri total yang sangat malas membayar SPP.

Kesimpulan:

1. Jumlah siswa pria lebih banyak daripada siswa wanita.
2. Sebagian besar siswa memiliki latar belakang kesantrian, baik santri total maupun santri kalong; meskipun tidak ada seorang siswa wanita pun yang berlatar belakang santri kalong.

3. Kebanyakan siswa rajin dan sangat rajin membayar SPP meskipun ada juga siswa yang sangat malas.
4. Latar belakang kesantrian berhubungan positif dengan kerajinan pembayaran SPP siswa; maksudnya siswa yang memiliki latar belakang kesantrian umumnya rajin atau sangat rajin dalam hal pembayaran SPP.

Latihan :

1. Data hasil ujian akhir mata kuliah statistika dari 60 orang mahasiswa :

23	60	79	32	57	74	52	70	82	36
80	77	81	95	41	65	92	85	55	76
52	10	64	75	78	25	80	98	81	67
41	71	83	54	64	72	88	62	74	43
60	78	89	76	84	48	84	90	15	79
34	67	17	82	69	74	63	80	85	61

Lakukan analisis dari distribusi frekuensi dan gambarlah diagramnya?

2. The Roth Young Personnel Service reported that annual salaries for department store assistant managers range from \$28,000 to \$57,000 (National Business Employment Weekly, October 16–22, 1994). Assume the following data are a sample of the annual salaries for 40 department store assistant managers (data are in thousands of dollars).

48	35	57	48	52	56	51	44
40	40	50	31	52	37	51	41
47	45	46	42	53	43	44	39
50	50	44	49	45	45	50	42
52	55	46	54	45	41	45	47

- a. What are the lowest and highest salaries reported?
- b. Use a class width of \$5000 and prepare tabular summaries of the annual salary data. Compare the result with the Sturges Method.
- c. What proportion of the annual salaries are \$35,000 or less?
- d. What percentage of the annual salaries are more than \$50,000?

3. Seorang guru ingin mengetahui kemampuan peserta didik kelas X SMA Mercuru Buana. Untuk itu, dia melakukan ujian tes prestasi terhadap 30 peserta didik dan didapat data hasil tes sebagai berikut :

Table 1. hasil prestasi belajar

70	80	65	90	55	85
75	85	70	78	65	55
90	45	70	73	70	65
66	65	55	68	70	76
54	78	60	66	80	75

Maka tentukan :

- Rata-rata nilai ujian tes prestasi?
- Lakukanlah analisis distribusi frekuensi dengan parameter jumlah nilai terendah (40 – 60), nilai sedang(61 – 79), dan tinggi (80 – 100)
- Bagaimana sebaran kemampuan peserta didik tersebut?
- Buatlah data kelompok dari table 1 diatas!

Pemusatan dan Penyebaran Data

A. Mengukur Pemusatan Data

Rumus yang digunakan untuk mengukur pemusatan data selalu dibedakan untuk data yang tidak dikelompokkan dan data yang dikelompokkan.

1. Rerata (mean)

Rerata merupakan konsep secara awam mengenai rata-rata. Merupakan titik berat dari seperangkat data atau observasi sensitif terhadap nilai ekstrim. Digunakan terutama bila teknik statistik lain, seperti pengujian hipotesis akan dilakukan pada data.

a. Untuk data yang tidak dikelompokkan (data tunggal)

$$\bar{x} = \frac{\sum x}{n}$$

$$\mu = \frac{\sum x}{N}$$

dimana :

\bar{x} = rerata

Σ = huruf besar Yunani sigma, yang berarti jumlahkan

x = nilai suatu hasil pengamatan atau observasi

Σx = jumlahkan semua observasi

n = jumlah semua observasi

b. Untuk data yang dikelompokkan

$$\bar{x} = \frac{\sum f \cdot x_i}{\sum f}$$

dimana :

\bar{x} = titik tengah (*mid point*) kelas interval ke I

x_i = titik tengah interval kelas

f = frekwensi observasi pada kelas interval ke i

$\sum fx$ = jumlahkan frekwensi tiap kelas interval

Contoh :

Data tinggi badan mahasiswa FKIP UMB- Yogyakarta diambil 50 mahasiswa secara random :

Tabel 1. Hasil Pengukuran tinggi badan

Interval Kelas	f_i
164,5 - 167,5	6
167,5 - 170,5	7
170,5 - 173,5	8
173,5 - 176,5	11
176,5 - 179,5	7
179,5 - 182,5	6
182,5 - 185,5	5
Jumlah	50

Jawab :

Interval Kelas	F	X_i	$f \cdot x_i$
164,5 - 167,5	6	166	996
167,5 - 170,5	7	169	1183
170,5 - 173,5	8	172	1376
173,5 - 176,5	11	175	1925
176,5 - 179,5	7	178	1246
179,5 - 182,5	6	181	1086
182,5 - 185,5	5	184	920
Jumlah	50		8732

Maka

$$\bar{x} = \frac{\sum f \cdot x_i}{\sum f} = \frac{8732}{50} = 174,64$$

2. Median

Median merupakan nilai tengah dari sekelompok data yang nilai tiap observasi telah disusun dari yang terkecil ke terbesar. Tidak sensitif terhadap nilai ekstrim. Median digunakan untuk mengukur pemusatan kalau distribusi mencong (*skewed*) secara jelas. Dapat dihitung pada distribusi yang tidak komplit sekalipun, misalnya distribusi yang berakhir terbuka (contoh 150-169 ; 170-189; 190-209; 210+).

a. Untuk data yang tidak dikelompokkan

- 1) Bila jumlah observasi ($=n$) ganjil, maka median adalah nilai observasi ke $:\frac{n+1}{2}$ dari urutan nilai observasi kecil ke besar.

Contoh : 5, 4, 5, 6, 7, 1, 5, 3, 4, 6, 9. Tentukan median

Urutkan data : 1, 3, 4, 4, 5, 5, 5, 6, 6, 7, 9

$$\text{Median } (M_e) = \frac{n+1}{2} = \frac{11+1}{2} = 6$$

- 2) Bila banyaknya observasi ($=n$) genap, maka median adalah nilai di antara observasi ke $:\frac{n}{2}$ dan $\frac{n}{2}+1$, diambil rata-rata.

Contoh :

1, 3, 4, 4, 5, 5, 5, 6, 6, 7

$$n = 10 \rightarrow \frac{n}{2} = 5 \text{ dan } 5 + 1 = 6$$

$$Me = \frac{5 + 6}{2} = 5,5$$

b. Untuk data yang dikelompokkan

$$Me = Lm + \frac{w \left(\frac{n}{2} - cf \right)}{f_m}$$

dimana :

Me = median

lm = batas bawah dari kelas interval dimana median berada (kelas median)

n = banyaknya observasi

cf = frekwensi kumulatif dari kelas interval sebelum kelas median

w = lebar kelas interval dimana median berada

contoh :

Tentukan median dari data kelompok dibawah ini

Jawab :

Interval Kelas	f_i
164,5 - 167,5	6
168,5 - 171,5	7
172,5 - 175,5	8
176,5 - 179,5	11
180,5 - 183,5	7
184,5 - 187,5	6
188,5 - 191,5	5
Jumlah	50

→ cf

→ fm

Jawab : (sebagai latihan mahasiswa)

Menentukan kelas median $= \frac{n}{2} = \frac{50}{2} = 25$

$$Me = Lm + \frac{w \left(\frac{n}{2} - cf \right)}{f_m} = 176 + \frac{7(25 - 21)}{11}$$
$$= 176 + \frac{28}{11} = 178,54$$

3. Modus (Mode)

Modus merupakan nilai yang paling sering muncul (frekuensi terbesar) dari seperangkat data atau observasi. Mencerminkan yang paling tipikal atau kasus yang paling umum. Kalau kita ingin segera mengetahui nilai pemusatan, maka kita menghitung modus. Seperangkat data dapat saja tidak memiliki modus, tetapi sebaliknya dapat pula memiliki beberapa modus. Kalau satu modus saja disebut unimodal, dua modus disebut bimodal dan kalau tanpa modus disebut nonmodal.

a. Untuk data yang tidak dikelompokkan

Modus (*crude mode*) = nilai yang paling sering muncul

Contoh : 1, 3, 4, 4, 5, 5, 5, 6, 6, 7

$M_0 = 5$

b. Untuk data yang dikelompokkan

Modus = titik tengah dari kelas interval yang memiliki frekwensi terbesar.

$$M_0 = Bb + w \left(\frac{d_1}{d_1 + d_2} \right)$$

Interval Kelas	f_i
164,5 - 167,5	6
168,5 - 171,5	7
172,5 - 175,5	8
176,5 - 179,5	11
180,5 - 183,5	7
184,5 - 187,5	6
188,5 - 191,5	5
Jumlah	50

$$\begin{aligned}
 M_0 &= 176 + 7 \frac{11 - 8}{(11 - 8) + (11 - 7)} \\
 &= 176 + 7 \left(\frac{3}{7} \right) \\
 &= 176 + 3 \\
 &= 179
 \end{aligned}$$

CONTOH :

1. Untuk data yang tidak dikelompokkan

Berikut ini data mengenai lama perawatan sepuluh penderita yang dirawat di bangsal perawatan Psikiatri dari suatu rumah sakit :

Pasien ke	Lama perawatan (hari)	Pasien ke	Lama perawatan (hari)
1	29	6	14
2	14	7	28
3	11	8	14
4	24	9	18
5	14	10	22

Hitung : rerata, median, modul lama perawatan dari pasien-pasien ini !

1. Rata-rata

$$\bar{x} = \frac{\sum x}{n} = \frac{11 + 14 + 14 + \dots + 24 + 28 + 29}{10}$$

$$\bar{x} = \frac{188}{10} = 18.8 \text{ hari}$$

2. Median

Urutan nilai observasi adalah sebagai berikut :

11; 14; 14; 14; 14; 18; 22; 24; 28; 29

Karena banyaknya observasi genap, maka median merupakan rata-rata nilai

dari observasi ke $\frac{n}{2} = \frac{10}{2} = 5$ dan $\frac{n}{2} + 1 = 6$

Jadi :

$$\text{Median} = \frac{14+18}{2} = 16 \text{ hari}$$

3. Modus

Oleh karena 14 hari adalah nilai yang paling sering muncul, maka modus adalah 14 hari

2. Untuk data yang dikelompokkan

Dari sejumlah penderita *typhus abdominalis* yang dirawat di bangsal penyakit menular suatu Rumah Sakit, diperoleh data sebagai berikut :

Masa inkubasi (hari) dari 170 penderita *typhus abdominalis*

Masa inkubasi (hari)	Jumlah penderita
2	25
6	80
10	30
14	15
18	12
22	6
24*	2
	total = 170

* tidak ada pasien dengan masa inkubasi 30 hari atau lebih.

Hitung : rerata, median dan modus.

Masa inkubasi (hari)	Banyaknya pasien (f)	Titik tengah (x)	fx	fx^2	Frekuensi kumulatif (cf)
2- 5	25	4	100	400	25
6-9	80	8	640	5120	105
10- 13	30	12	360	4320	135
14- 17	15	16	240	3840	150
18- 21	12	20	240	4800	162
22-25	6	24	144	3456	168
26- 29	2	28	56	1568	170
	Total = 170		$fx =$ 1780	2350 4	

1. Rerata

$$\bar{x} = \frac{fx}{n} = \frac{1780}{170} = 10,47 \text{ hari}$$

2. Median

$$Md = L_m + \frac{\frac{n}{2} - cf}{fm} \cdot w$$

$\frac{n}{2} = \frac{170}{2} = 85$, kelas interval dimana median berada (kelas median) adalah: 6, maka

$$lm = 6$$

$$cf \text{ kelas interval sebelumnya} = 25$$

$$fm = 80$$

$$w = 10 - 6 = 4$$

$$Md = 6 + \frac{\frac{170}{2} - 25}{80}$$

$$Md = 6 + \frac{60}{80} \cdot 4$$

$$Md = 6 + 3 = 9$$

3. Modus

$Mo = 8$, oleh karena frekuensi tertinggi dimiliki kelas interval 6 - dan titik tengah kelas interval ini adalah : 8.

Latihan :

Berdasarkan data pada contoh kasus 1. Tentukan nilai mean, median, modus

NILAI	SCORE	BANYAKNYA DATA (FREKUENSI)
40 - 49	III	4
50 - 59	IIII I	6
60 - 69	IIII IIII	10
70 - 79	III	4
80 - 89	III	4
90 - 99	II	2
Σ		30

B. Ukuran Letak

Agar kita dapat mengetahui lebih jauh mengenai karakteristik data observasi dengan beberapa ukuran sentral, kita sebaiknya mengetahui beberapa ukuran lain, yaitu ukuran letak. Ada tiga macam ukuran letak yang akan di bahas pada bagian ini, yaitu Kuartil, Desil, dan Persentil.

1. Kuartil

Kuartil adalah ukuran letak yang membagi data observasi menjadi empat bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 25% data observasi. Pada satu set data observasi mempunyai tiga buah kuartil, yaitu K_1, K_2, K_3 .

Untuk menentukan nilai kuartil data observasi yang tidak berkelompok (*ungrouped data*) melalui langkah-langkah sebagai berikut ini :

- 1) Urutkan data observasi dari kecil ke besar
- 2) Tentukan letak kuartilnya

Menentukan letak K_1, K_2, K_3 dapat digunakan formulasi sebagai berikut :

$$\text{Letak } K_1 = \frac{N+1}{4}$$

$$\text{Letak } K_2 = \frac{2(N+1)}{4}$$

$$\text{Letak } K_3 = \frac{3(N+1)}{4}$$

- 3) Tentukan nilai kuartilnya.

Nilai K_1, K_2, K_3 adalah data observasi yang terletak pada letak K_1, K_2, K_3

Contoh kasus :

Berikut ini adalah data mengenai nilai 7 orang peserta ujian Statistik di UMB Yogyakarta :

78 56 66 48 80 70 76

Tentukan K_1, K_2, K_3

Jawab :

Untuk menentukan K_1, K_2, K_3 , maka langkah-langkah yang digunakan adalah sebagai berikut :

- Urutkan nilai tersebut dari kecil ke besar

48 56 66 70 76 78 80

- Tentukan letak K_1, K_2, K_3 dengan formula

$$\text{Letak } K_1 = \frac{7+1}{4} = 2$$

$$\text{Letak } K_2 = \frac{2(7+1)}{4} = 4$$

$$\text{Letak } K_3 = \frac{3(7+1)}{4} = 6$$

Jadi letak K_1 pada urutan data ke 2, letak K_2 pada urutan data ke 4, dan letak K_3 pada urutan data ke 6

- Tentukan nilai K_1, K_2, K_3

No urut	1	2	3	4	5	6	7
nilai	48	56	66	70	76	78	80



Nilai K_2 adalah juga merupakan median dari nilai peserta ujian tersebut. Apabila banyaknya data observasi menunjukkan bilangan genap, maka median terletak diantara dua nomor urut.

Kuartil (K_1, K_2, K_3) data observasi berkelompok dapat ditentukan dengan langkah-langkah sebagai berikut ini :

1. Tentukan kelas K_1, K_2, K_3 dengan formula :

Kelas kuartil 1 (K_1) :

$$K_1 = \frac{N}{4}$$

Kelas kuartil 2 (K_2) :

$$K_2 = \frac{2N}{4}$$

Kelas kuartil 3 (K_3) :

$$K_3 = \frac{3N}{4}$$

2. Tentukan K_1, K_2, K_3 dengan menggunakan formula

$$K_1 = B_{K_1} + \left(\frac{\frac{N}{4} - Cf_1}{f_{K_1}} \right) \cdot C_i$$

Yang menyatakan bahwa :

K_1 = Kuartil 1

B_{K_1} = tepi kelas bawah kelas kuartil 1

N = banyaknya data observasi ($\sum f$)

Cf_1 = frekuensi kumulatif kelas sebelum kelas kuartil 1

f_{K_1} = frekuensi kumulatif kelas kuartil 1

C_i = interval kelas

$$K_2 = B_{K_2} + \left(\frac{\frac{2N}{4} - Cf_2}{f_{K_2}} \right) \cdot C_i$$

Yang menyatakan bahwa :

K_2 = Kuartil 2

B_{K_2} = tepi kelas bawah kelas kuartil 2

N = banyaknya data observasi ($\sum f$)

Cf_2 = frekuensi kumulatif kelas sebelum kelas kuartil 2

f_{K_2} = frekuensi kumulatif kelas kuartil 2

C_i = interval kelas

K_2 nilainya sama dengan nilai median

$$K_3 = B_{K_3} + \left(\frac{\frac{3N}{4} - Cf_3}{f_{K_3}} \right) \cdot C_i$$

Yang menyatakan bahwa :

K_3 = Kuartil 3

B_{K_3} = tepi kelas bawah kelas kuartil 3

N = banyaknya data observasi ($\sum f$)

Cf_3 = frekuensi kumulatif kelas sebelum kelas kuartil 3

f_{K_3} = frekuensi kumulatif kelas kuartil 3

C_i = interval kelas

Contoh kasus :

Tentukan K_1, K_2 dan K_3 nilai 30 peserta ujian statistik seperti yang tampak pada tabel 3.1

NILAI	FREKUENSI	TEPI KELAS	FREKUENSI KUMULATIF
40 – 49	4	39,5	4
50 – 59	6	49,5	10
60 – 69	10	59,5	20
70 – 79	4	69,5	24
80 – 89	4	79,5	28
90 – 99	2	89,5	30
Σ	30		

2. Desil

Desil adalah ukuran letak yang membagi data observasi menjadi sepuluh bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 10% data observasi. Pada satu set data observasi mempunyai sembilan buah desil, yaitu D_1, D_2, \dots, D_9 .

Untuk data tunggal, jika banyak data n dan D_i adalah desil ke- i , maka

$$\text{Letak } D_i = \text{data ke-} \frac{i(n+1)}{10} \text{ dengan } i = 1, 2, 3, 4, \dots, 9$$

Contoh;

Tentukan D_3 , dan D_5 dari ; 6, 4, 6, 4, 7, 5, 6, 5, 8, 7, 7, 7, 8, 6 !

Penyelesaian;

Data diurutkan menjadi ; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8

Data	4	4	5	5	6	6	6	6	7	7	7	7	8	8
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

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

$$\text{Letak } D_3 = \text{data ke-} \frac{3(14+1)}{10}$$

$$= \text{data ke- } 4,5 (X_{4,5})$$

Dengan interpolasi diperoleh :

$$D_3 = X_4 + 0,5(X_5 - X_4)$$

$$= 5 + 0,5(6 - 5)$$

$$= 5,5$$

$$\text{Letak } D_5 = \text{data ke-} \frac{5(14+1)}{10}$$

$$= \text{data ke- } 7,5 (X_{7,5})$$

Dengan interpolasi diperoleh :

$$\begin{aligned} D_5 &= X_7 + 0,5(X_8 - X_7) \\ &= 6 + 0,5(6 - 6) \\ &= 6 \end{aligned}$$

Desil data berkelompok dapat dihitung dengan rumus :

$$D_i = T_b + p \left(\frac{\frac{i}{10}n - F}{f} \right)$$

Dimana $i = 1, 2, 3, 4, \dots, 9$

Dengan $D_i = \text{desil ke-}i$

T_b = tepi bawah interval kelas D_i

P = panjang kelas interval D_i

$n = \sum f$ = banyak data

F = frekuensi kumulatif sebelum kelas D_i

f = frekuensi pada kelas D_i

Contoh.

Hitung nilai D_5 dan D_8 dari data berdistribusi kelompok berikut :

Interval	F	F_k
21-25	3	3
26-30	9	12
31-35	4	16
36-40	10	26
41-45	3	29
46-50	11	40

Penyelesaian ;

Desil ke-5 terletak pada $\left(\frac{5}{10} \cdot 40\right) = 20$ (kelas interval 36-40)

$$\begin{aligned} D_5 &= 35,5 + \frac{5(20 - 16)}{10} \\ &= 37,5 \end{aligned}$$

Desil ke-8 terletak pada $\left(\frac{8}{10} \cdot 40\right) = 32$ (kelas interval 46-50)

$$D_8 = 45,5 + \frac{5(32 - 29)}{11}$$

$$= 46,9$$

3. Persentil

Persentil adalah ukuran letak yang membagi data observasi menjadi seratus bagian yang sama besar. Oleh karena itu masing-masing bagian mengandung 1 % data observasi. Pada satu set data observasi mempunyai 99 persentil, yaitu : P_1, P_2, \dots, P_{99} .

Persentil data tunggal maka :

$$\text{Letak } P_i = \text{data ke-} \left(\frac{i(n+1)}{100}\right), \text{ dengan } i = 1,2,3,\dots,99$$

Contoh;

Tentukan P_{30} , dan P_{75} dari ; 6, 4, 6, 4, 7, 5, 6, 5, 9, 7, 10, 7, 10, 6 !

Penyelesaian;

Data diurutkan menjadi ; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 9, 10, 10

Data	4	4	5	5	6	6	6	6	7	7	7	9	10	10
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

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

$$\text{Letak } P_{30} = \text{data ke-} \frac{30(14+1)}{100}$$

$$= \text{data ke- } 4,5 (X_{4,5})$$

Dengan interpolasi diperoleh :

$$P_{30} = X_4 + 0,5(X_5 - X_4)$$

$$= 5 + 0,25(6 - 5)$$

$$= 5,25$$

$$\text{Letak } P_{75} = \text{data ke-} \frac{75(14+1)}{100}$$

$$= \text{data ke- } 11,25 (X_{11,25})$$

Dengan interpolasi diperoleh :

$$P_{75} = X_{11} + 0,5(X_{12} - X_{11})$$

$$= 7 + 0,25(9 - 7)$$

$$= 7,5$$

Persentil data berkelompok dapat dihitung dengan rumus :

$$P_i = T_b + p \left(\frac{\frac{i}{100}n - F}{f} \right)$$

Dimana $i = 1, 2, 3, 4, \dots, 99$

Dengan P_i = persentil ke- i

T_b = tepi bawah interval kelas P_i

P = panjang kelas interval P_i

$n = \sum f$ = banyak data

F = frekuensi kumulatif sebelum kelas P_i

f = frekuensi pada kelas P_i

Contoh.

Hitung nilai P_{25} dari data berdistribusi kelompok berikut :

Interval	F	F_k
21-25	3	3
26-30	9	12
31-35	4	16
36-40	10	26
41-45	3	29
46-50	11	40

Penyelesaian ;

Persentil ke-25 terletak pada $\left(\frac{25}{100} \cdot 40 \right) = 10$ (kelas interval 26-30)

$$\begin{aligned} P_{25} &= 25,5 + \frac{5(10 - 3)}{9} \\ &= 29,4 \end{aligned}$$

C. Pengukuran Penyebaran (Dispersi)

1. Pengertian Tentang Disperse.

Digunakan untuk menunjukkan keadaan berikut :

a. Gambaran variabilitas data

Yang dimaksud dengan variabilitas data adalah suatu ukuran yang menunjukkan besar kecilnya perbedaan data dari rata-ratanya. Ukuran ini

dapat juga disebutkan sebagai ukuran yang menunjukkan perbedaan antara data satu dengan yang lainnya. Ukuran pemusatan (Mean, Median, dan Modus) ini dapat kita gunakan untuk menggambarkan keadaan sekumpulan data, tetapi gambaran itu masih kurang lengkap apabila tidak disertai dengan ukuran-ukuran penyebaran. Hal ini disebabkan karena dengan ukuran gejala pusat saja mungkin beberapa kumpulan data sebenarnya berbeda dapat disimpulkan sama.

b. Perbedaan nilai satu observasi terhadap nilai observasi lainnya

Rata-rata dari serangkaian nilai-nilai observasi tidak dapat diinterpretasikan secara terpisah dengan dispersi (sebaran) nilai-nilai tersebut terhadap rata-ratanya. Jika terdapat keseragaman/kesamaan nilai-nilai observasi, X_i , maka dispersi nilai-nilai tersebut akan sama dengan nol, dan rata-ratanya akan sama dengan nilai X_i . Semakin besar variasi nilai-nilai X_i , maka rata-rata distribusi semakin kurang representatif.

Contoh:

Tabel 7-1 Rata-rata hitung hasil test mata kuliah statistik deskriptif kelompok A dan B.

kelompok	hasil test					
A	60	65	50	60	65	60
B	65	90	50	70	60	60

Mahasiswa A: $\bar{X} = 360/6 = 60$

Mahasiswa B: $\bar{X} = 360/6 = 60$

Rata-rata hasil test kedua mahasiswa tersebut tidak berbeda, namun dispersi hasil test mahasiswa B (30 sampai dengan 90) jauh lebih besar dari pada varisasi hasil test mahasiswa A (50 sampai dengan 65). Hal ini berarti hasil test mahasiswa A jauh lebih konsisten (stabil) dibanding mahasiswa B. Tingkat dispersi berhubungan erat dengan sifat kesamaan/kesejenan data. Misalnya data tentang besarnya modal pedagang kaki lima khusus makanan, akan kecil variasinya jika dibandingkan dengan data seluruh pedagang kaki lima tanpa melihat jenis dagangannya. Secara umum, suatu rata-rata akan cukup representatif bagi

serangkaian nilai-nilai observasi X_i bila nilai-nilai tersebut diperoleh dari data yang bersifat sejenis bagi tujuan pengamatan tertentu.

2. Pengukuran Jarak (Range)

Pengukuran jarak sebuah distribusi merupakan pengukuran dispersi yang paling sederhana. Jarak sebuah distribusi frekuensi dirumuskan sebagai “selisih atau beda antara pengukuran nilai terbesar dan nilai terkecil yang terdapat dalam sebuah distribusi frekuensi”. Atau secara matematis dapat ditulis sebagai berikut:

$$R = X_u - X_i$$

Keterangan :

R = range data observasi

X_u = nilai tertinggi

X_i = nilai terendah

Beberapa Catatan Tentang Pengukuran dan Penggunaan Jarak

- 1) Hasil pengukuran jarak (range) sebenarnya sudah dapat menggambarkan disperse (variasi) nilai-nilai observasi dengan cara yang paling sederhana. Jika kita ingin memperoleh hasil pengukuran dispersi secara kasar dan cepat, maka ukuran range dapat digunakan.
- 2) Range bukan merupakan pengukuran dispersi distribusi yang memuaskan karena hasil pengukurannya jelas tergantung pada kedua nilai ekstrim tanpa mengikutsertakan pola dispersi nilai-nilai observasi X_i secara keseluruhan.

Contoh kasus :

Berikut ini adalah nilai ulangan harian 10 siswa mata pelajaran statistika di SMA Mercuru Buana Yogyakarta:

56 66 78 94 48 82 50 76 80 70

Range nilai 10 siswa yang ikut ulangan harian statistika tersebut dapat ditentukan dengan menggunakan formula :

$$\begin{aligned} R &= X_u - X_i \\ &= 94 - 48 = 46 \end{aligned}$$

Range data observasi berkelompok (grouped data) adalah data selisih antara tepi kelas atas kelas yang terakhir dengan tepi kelas bawah kelas pertama.

Contoh kasus :

Tabel 2.1 berikut ini data mengenai nilai 30 peserta ujian Matematika di SMA Mercu Buana Yogyakarta

Tabel 2.1

NILAI	FREKUENSI (f)
40 - 49	6
50 - 59	10
60 - 69	4
70 - 79	4
80 - 89	2
90 - 99	4

Range nilai 30 peserta ujian matematika dapat ditentukan dengan menggunakan Rumus :

$$R = B_u - B_i$$

Dengan nilai-nilai

$B_u = 99,5$ (tepi kelas atas kelas yang terakhir)

$B_i = 39,5$ (tepi kelas bawah kelas yang pertama)

Sehingga besarnya Range (R)

$$R = 99,5 - 39,5 = 60$$

3. Pengukuran Deviasi Kuartil.

Nilai-nilai X_i yang ordinatnya membagi seluruh distribusi dalam 4 (empat) bagian yang sama dinamakan nilai-nilai kuartil. Q1 merupakan kuartil pertama, Q2 merupakan kuartil kedua dan sama dengan median ($Q2 = md$), sedangkan Q3 dinamakan kuartil ketiga. Dalam distribusi kuartil, 50% dari semua nilai-nilai observasi seharusnya terletak antara Q1 dan Q3. Jarak antara Q1 dan Q3 dinamakan jarak inter-kuartil (inter-quartilrange). Makin kecil jarak tersebut, maka makin tinggi tingkat konsentrasi distribusi tengah seluas 50% dari seluruh distribusi.

Secara teoritis, pengukuran deviasi kuartil sebuah sampel dapat dirumuskan sebagai:

$$dQ = \frac{Q_3 - Q_1}{2}$$

Selanjutnya dapat dikatakan bahwa deviasi kuartil adalah sebesar +dQ atau - dQ dari mediannya.

Pada dasarnya, pengukuran deviasi kuartil sama seperti pengukuran jarak (range). Pengukurannya didasarkan pada jarak antara Q1 dan Q3. Pengukuran tersebut tidak dipengaruhi oleh dispersi dari seluruh nilai-nilai observasi, deviasi kuartil hanya mengikutsertakan dispersi nilai-nilai observasi X_i yang didistribusikan di tengah-tengah seluruh distribusi seluas 50% saja.

4. Pengukuran Deviasi Rata-rata (Mean Deviation)

a. Deviasi rata-rata dari data yang belum dikelompokkan

Dispersi serangkaian nilai-nilai observasi akan kecil bila nilai-nilai tersebut berkonsentrasi sekitar rata-ratanya. Sebaliknya, dispersinya akan besar bila nilai-nilai observasi tersebar jauh dari rata-ratanya.

Deviasi rata-rata dari seluruh nilai-nilai observasi X_i dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum(x_i - \bar{x})}{n}$$

Sedangkan pengukuran deviasi atas dasar nilai-nilai absolut dapat dirumuskan sebagai:

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

Contoh :

Carilah deviasi rata-rata data berikut ini :

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Jawab :

Dimana $i=1,2,3,4,\dots,20$

$$d_{\bar{x}} = \frac{\sum|x_i - \bar{x}|}{n} = \frac{706}{20} = 35,31$$

b. Deviasi rata-rata dari data yang telah dikelompokkan

Apabila nilai-nilai observasi sudah dikelompokkan ke dalam bentuk distribusi frekuensi, maka deviasi rata-ratanya dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum_{i=1}^k f_i |m_i - \bar{x}|}{n}$$

Dimana:

m_i = titik tengah kelas frekuensi

f_i = frekuensi dari kelas distribusi ke-i

k = jumlah kelas distribusi

Dalam beberapa kondisi tertentu, median dapat digunakan sebagai pengukuran rata-rata secara memuaskan. Deviasi rata-rata sebuah distribusi dapat juga diukur dari median distribusi yang bersangkutan seperti dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum |x_i - md|}{n}$$

Atau

$$d_{\bar{x}} = \frac{\sum f_i |m_i - md|}{n}$$

Umumnya deviasi rata-rata merupakan pengukuran dispersi yang lebih baik jika dibandingkan dengan jarak atau deviasi kuartil. Hasil pengukuran deviasi rata-rata mencerminkan dispersi tiap-tiap nilai observasi dari rata-ratanya dan bukan hanya tergantung pada kedua nilai ekstrim.

Contoh :

Dari data tunggal dibawah ini, rubahlah menjadi data kelompok :

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Dan carilah Deviasi rata-ratanya.

Jawab :

Data setelah dikelompokkan

Nilai	F	Mi
40 - 47	2	43,5
48 - 55	5	51,5
56 - 63	2	59,5
64 - 71	5	67,5
72 - 79	2	75,5
80 - 87	4	83,5
Σ	20	

$$\begin{aligned} \text{Median (Md)} &= Mo + P \left(\frac{\frac{n}{2} - fm}{fc} \right) \\ &= 63,5 + 8 \left(\frac{10 - 9}{5} \right) \\ &= 63,5 + 8(0,2) \\ &= 63,5 + 1,6 \\ &= 65,1 \end{aligned}$$

$$\begin{aligned} d_{\bar{x}} &= \frac{\Sigma f_i |x_i - md|}{n} \\ &= \frac{|(43,5 - 65,1) \cdot 2 + (51,5 - 65,1) \cdot 5 + \dots + (83,5 - 65,1) \cdot 4|}{20} \\ &= \frac{228,8}{20} = 11,44 \end{aligned}$$

5. Pengukuran Varians dan Deviasi Standar

Varians digunakan untuk melihat kehomogenan data secara kasar, dimana nilai hasil perhitungan varians sebagai titik pusat dari penyebaran data.

Contoh 1:

Seorang guru matematika melakukan tes prestasi dengan membagi siswa dalam 3 kelompok, yaitu A,B, dan C. Dalam satu kelompok terdapat 5 siswa. Walaupun dibentuk kelompok namun untuk tes dikerjakan secara individu. Didapat hasil sebagai berikut :

KELOMPOK	NILAI					\bar{x}
A	50	50	50	50	50	50
B	60	40	50	55	45	50
C	30	70	90	10	50	50

a. *Varians dan deviasi standar dari data yang belum dikelompokkan*

Karl Pearson merumuskan pengukuran varians sebagai:

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

Standarisasi unit-unit pengukuran di atas dilakukan melalui proses pengakaran, dan dinamakan deviasi standar, sebagai berikut:

$$s = \sqrt{s^2}$$

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

b. *Varians dan deviasi standar dari data yang belum dikelompokkan*

- Rumus Fisher dan Wilks

Varians dari Fisher dan Wilks:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- Deviasi standar dari Fisher dan Wilks:

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

- Varians dan deviasi standar populasi

Varians polupasi:

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

- Deviasi standar populasi:

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

c. **Varians dan deviasi standar dari data yang telah dikelompokkan**

- Varians dari data sampel yang telah dikelompokkan:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \cdot f_i$$

- Deviasi standar dari data sampel yang telah dikelompokkan:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \cdot f_i}$$

dimana:

x_i = titik tengah tiap-tiap kelas

f_i = jumlah frekuensi kelas

d. **Variansi dan deviasi standar dengan cara transformasi**

Seperti halnya dengan mencari nilai mean data kelompok. Kita juga dapat mencari nilai variansi dapat dicari dengan cara transformasi.

$$u_i = x_i - a$$

Dimana :

x_i : titik tengah interval kelas ke-i

a : sembarang harga titik tengah interval kelas (biasanya yang memiliki frekuensi terbanyak)

sehingga rumus VARIANSI (S^2) adalah :

$$s^2 = c^2 u^2$$

c = lebar kelas/panjang kelas

dimana :

$$s_{\bar{u}}^2 = \frac{1}{n-1} \sum_{i=1}^k f_i (u_i - \bar{u})^2$$

Atau dapat juga ditulis :

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n f_i u_i^2 - \frac{1}{n} \left(\sum_{i=1}^k f_i u_i \right)^2 \right]$$

Contoh :

Dari data tinggi badan (cm) 50 mahasiswa Pendidikan Matematika FKIP Universitas Mercu Buana Yogyakarta didapat data :

Tabel 1. Perhitungan variansi data berkelompok

Interval Kelas	x_i	u_i	f_i	u_i^2	$f_i u_i$	$f_i u_i^2$
164,5 - 167,5	166	166-175=-9	6	81	6*-9=-54	6*81 =486
167,5 - 170,5	169	169-175=-6	7	36	7*-6=-42	7*36 = 252
170,5 - 173,5	172	-3	8	9	-24	72
173,5 - 176,5	175	0	11	0	0	0
176,5 - 179,5	178	178-175= 3	7	9	21	63
179,5 - 182,5	181	6	6	36	36	216
182,5 - 185,5	184	9	5	81	45	405
Jumlah			50		-18	1494

Berdasarkan tabel 1 dengan menggunakan rumus transformasi, maka variansinya :

$$\begin{aligned}
 s^2 &= \frac{1}{n-1} \left[\sum_{i=1}^n f_i u_i^2 - \frac{1}{n} \left(\sum_{i=1}^k f_i u_i \right)^2 \right] \\
 &= \frac{1}{50-1} \left(1494 - \frac{1}{50} (-18)^2 \right) = 30,35 \\
 s &= \sqrt{30,4} = 5,50
 \end{aligned}$$

e. *Beberapa catatan tentang varians dan deviasi standar dari data yang telah dikelompokkan*

- Koreksi Sheppard (Sheppard's Correction): Jika distribusi frekuensi simetris atau mendekati simetris, maka hasil rata-rata hitung yang diperoleh dari distribusi frekuensi tersebut kurang lebih sama dengan hasil rata-rata yang diperoleh dari data kasar (yang belum dikelompokkan).
- Distribusi normal sebenarnya merupakan distribusi teoritis (mengikuti "hukum normal") karena pada dasarnya gejala-gejala alami tidak seluruhnya bersifat normal.

Latihan :

Dari data di bawah ini :

NO	NILAI	F
1	5 – 9,99	6
2	10 – 14,99	12
3	15 – 19,99	19
4	20 – 24,99	20
5	25 – 29,99	14
6	30 – 34,99	8
7	35 – 39,99	2
	JUMLAH	80

Maka tentukan :

1. Gambarlah diagram batang, garis
2. Tentukan Mean, median, Modus, Variansi, SD
3. Tentukan Variansi dan SD dengan cara transformasi

Latihan-latihan

1. Populasi beranggotakan orang dengan ukuran masing-masing : 4,5,6,7,10,12,14. Diambil 2 sampel ukuran dengan pengambilan sampel dilakukan tanpa pengembalian. Buatlah distribusi sampling rata-ratanya?
2. Di ketahui sebuah data tentang nilai prestasi matematika siswa kelas X SMA Mercuru Buana.

Nilai (X)	f
4	2
5	3
6	8
7	4
8	5
9	3
10	2
3	3

Dari tabel di atas maka tentukan :

- a. Proporsi, Kumulasi Proporsi Bawah dan Proporsi Kumulasi atas
- b. Simpangan x, Jumlah Kuadrat Simpangan (JK(X)), Variansi, dan Simpangan Baku.

3. Data dikotomi tentang motivasi belajar dan prestasi belajar matematika

SISWA	MOTIVASI		PRESTASI BELAJAR	
	RENDAH (X1)	TINGGI (X2)	RENDAH (Y1)	TINGGI (Y2)
A	0	1	1	0
B	0	1	1	0
C	0	1	0	1
D	1	0	0	1
E	1	0	0	1
F	0	1	1	0
G	0	1	1	0
H	1	0	0	1
I	0	1	0	1
J	0	1	1	0
K	0	1	1	0
L	0	1	1	0
M	0	1	0	1
N	0	1	0	1
O	0	1	0	1
P	0	1	1	0
Q	0	1	0	1
R	0	1	1	0
S	1	0	0	1
T	1	0	0	1
U	1	0	0	1
V	0	1	0	1
W	1	0	0	1
X	1	0	1	0
Y	1	0	1	0
Z	1	0	0	1
AA	1	0	0	1
AB	0	1	1	0
AC	0	1	0	1
AD	0	1	0	1
AE	0	1	1	0
AF	1	0	0	1
AG	1	0	0	1
AH	1	0	0	1
AI	1	0	0	1

Dari data diatas maka tentukan : Rerata, variansi, simpangan baku untuk X1, X2, Y1 dan Y2

Kemencengan dan Kurtosis

A. Pengukuran Kemencengan

Rata-rata hitung serta deviasi standar dua distribusi mungkin sama meskipun bentuk kurva frekuensi kedua distribusi tersebut berbeda karena tingkat kemencengannya berbeda.

1. Koefisien Pearson Tentang Kemencengan

Rata-rata hitung dipengaruhi oleh nilai-nilai ekstrimnya. Modus tidak dipengaruhi oleh nilai-nilai ekstrim, sedangkan median hanya dipengaruhi oleh kedudukannya. Jika sebuah distribusi simetris, maka rata-rata hitung = median = modus. Sebaliknya jika distribusi tidak simetris, maka rata-rata hitung \neq median \neq modus. Pengukuran tingkat kemencengan (skewness) pertama kali dirumuskan oleh Karl Pearson dalam bentuk koefisien Pearson sebagai:

$$sk = \frac{\bar{X} - m_0}{s}$$

Dimana

sk = kemencengan

\bar{X} = rata-rata hitung

m_0 = modus

s = deviasi standard

a. *Modifikasi koefisien $(\bar{X} - mo)/s$*

Perumusan ko-efisien Pearson membutuhkan data statistik rata-rata hitung, modus, dan deviasi standar. Namun banyak para statistisi yang kurang merasa puas dengan penggunaan modus bagi pengukuran kemencengan distribusi, karena pengukuran modus distribusi sampel umumnya bersifat aproksimatif (kira-kira) dan seringkali memiliki selisih yang relatif besar terhadap modus dari data asalnya. Selanjutnya, Pearson merumuskan kembali pengukuran kemencengan menjadi sebagai berikut:

$$sk = \frac{(\bar{X} - [\bar{X} - 3(\bar{X} - md)])}{s}$$

atau:

$$sk = 3 \left(\frac{\bar{X} - md}{s} \right)$$

b. *Interpretasi hasil ko-efisien Pearson*

Berdasarkan pengalaman empiris, Croxton dan Crowden beranggapan bahwa hasil ko-efisien kemencengan distribusi dapat bervariasi antara +3. Meskipun demikian, mereka berpendapat bahwa hasil ko-efisien kemencengan jarang sekali mencapai +1. Hasil demikian kemungkinan diperoleh berdasarkan karya Hostelling dan Solomon dimana mereka membuktikan bahwa $(\bar{X} - md)/s$ seharusnya terletak antara +1

c. *Rumus Bowley Tentang Kemencengan*

Sebuah perumusan tentang pengukuran kemencengan yang lebih sederhana dibandingkan rumus dari Pearson telah dikembangkan oleh A.L. Bowley. Ia mengembangkan ko-efisiennya atas dasar hubungan antara statistik Q_1, Q_2, Q_3 dari sebuah distribusi. Jika sebuah distribusi simetris, maka jarak antara kedua kuartil dari mediannya adalah sama. Sementara, jika sebuah distribusi tidak simetris, maka jarak antara kedua kuartil dari mediannya tidak akan sama.

Ko-efisien Bowley dirumuskan sebagai berikut:

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

2. Pengukuran Kemencengan Relatif

Kemencengan relatif α_3 sangat tergantung pada bentuk kurva frekuensi dan seringkali digunakan sebagai pengukuran kemencengan sekitar rata-rata distribusi teoritis. Perumusan α_3 secara umum untuk data yang belum dikelompokkan dapat ditulis sebagai:

$$\alpha_3 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^3}{\sigma^3}$$

Sedangkan untuk data yang telah dikelompokkan adalah:

$$\alpha_3 = \frac{\frac{1}{n} \sum_{i=1}^n (m_i - \mu)^3 \cdot f_i}{\sigma^3}$$

Jika distribusi simetris sekitar rata-ratanya, maka $\sum_{i=1}^k (m_i - \mu)^3 = 0$, sehingga

$\alpha_3 = 0$. Sebaliknya jika distribusi menceng sekitar rata-ratanya, maka α_3 akan menghasilkan nilai positif atau negatif sesuai dengan arah kemencengan distribusi.

B. Pengukuran Kurtosis

1. Pengertian Tentang Kurtosis.

Pengukuran kurtosis (peruncingan) sebuah distribusi teoritis kadang-kadang disebut juga dengan istilah eksese (excess) dari sebuah distribusi. Sesungguhnya kurtosis dapat dianggap sebagai suatu distorsi dari kurva normal. Kurtosis pada umumnya diukur dengan cara membandingkan bentuk peruncingan kurvanya dengan kurva normal. Jika bagian tengah dari kurva frekuensi memiliki puncak (peak) yang lebih runcing dari pada yang dimiliki kurva normal, maka distribusi tersebut dinamakan distribusi leptokurtik (leptokurtic). Sedangkan jika bagian tengah kurva distribusi frekuensi memiliki puncak yang lebih datar dari pada yang dimiliki oleh kurva normal, maka distribusinya dinamakan distribusi platikurtik (platykurtic). Distribusi normal atau disebut dengan distribusi mesokurtik (mesokurtic) pada dasarnya berada diantara leptokurtik dan platikurtik.

2. Pengukuran Kurtosis

Secara teoritis, pengukuran kurtosis sebuah distribusi dapat dilakukan dengan menggunakan α_4 yang dirumuskan untuk data yang belum dikelompokkan sebagai:

$$\alpha_4 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^4}{\sigma^4}$$

dan bagi data yang sudah dikelompokkan sebagai:

$$\alpha_4 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^4 \cdot f_i}{\sigma^4}$$

Distribusi yang sangat meruncing akan memiliki α_4 yang tinggi, sedangkan distribusi dengan puncak yang datar akan menghasilkan α_4 yang rendah. Saat ini statistisi mengetahui bahwa bentuk keruncingan (kurtosis) distribusi sebenarnya tidak berkaitan dengan nilai α_4 . Dua buah distribusi yang berbeda dapat memiliki α_4 yang sama. Pada hakekatnya sebuah kurtosis distribusi jarang sekali dihitung. Pengukuran kurtosis sendiri sebetulnya penting sekali dalam distribusi student dan distribusi normal. Penerapan kurva frekuensi teoritis dapat dibenarkan jika kurtosis kurva frekuensi tidak berbeda secara mencolok dari kurtosis distribusinya sendiri. Misalnya jika taksiran kurtosis populasi adalah sebesar $-0,104$ maka bagi sebuah kurva normal, nilai kurtosis di atas seharusnya menjadi nol. Bagi distribusi Poisson dengan l yang besar sekali, kurtosis seharusnya mendekati nol sehingga distribusinya dapat diterapkan dengan kurva normal.