

## STATISTIK dan PROBABILITAS

### 1. Pendahuluan

#### ▪ Pengertian

Setiap hari kita berhadapan dengan data dan kadang harus membuat kesimpulan/keputusan berdasarkan data tersebut. Statistik adalah disiplin ilmu yang menyediakan metode untuk membantu kita memahami data. Metode statistik yang digunakan secara cerdas, menawarkan seperangkat alat untuk mendapatkan wawasan tentang dunia di sekitar kita. Saat ini penggunaan analisis statistik telah meluas ke berbagai bidang seperti bisnis, kedokteran, pertanian, ilmu sosial, alam ilmu, dan rekayasa. Statistik mengajarkan kita bagaimana membuat keputusan cerdas terhadap variabilitas dalam sampel dan populasi. Populasi tanpa variabilitas sangat langka, variabilitas hampir universal. Variabilitas perlu dipahami untuk dapat mengumpulkan, mendeskripsikan, menganalisis, dan menarik kesimpulan dari data dalam cara yang masuk akal.

Statistik adalah ilmu yang menyediakan metode untuk membantu kita memahami data.

Ada beberapa pendapat mengenai pengertian statistik :

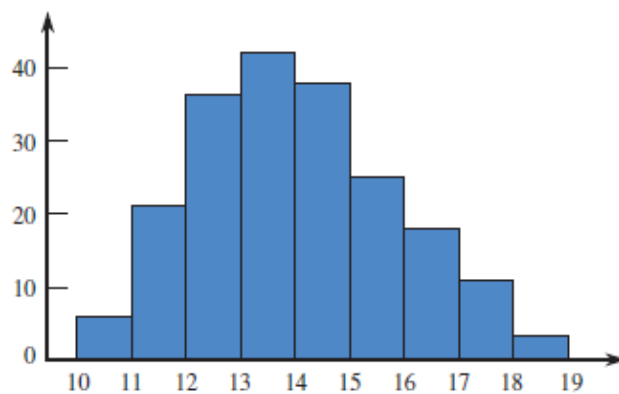
1. Statistik adalah sekumpulan cara maupun aturan-aturan yang berkaitan dengan pengumpulan, pengolahan (Analisis), penarikan kesimpulan, atas data-data yang berbentuk angka dengan menggunakan suatu asumsi-asumsi tertentu. (*Prof. Dr. H. Agus Irianto*)
2. Statistik adalah metode yang memberikan cara-cara guna menilai ketidak-tentuan dari penarikan kesimpulan yang bersifat induktif. (*Stoel dan Torrie*)

Kedua pendapat diatas menyatakan statistik adalah metode untuk mendapatkan informasi yang ada pada serangkaian data.

#### Definisi Statistika

Statistika adalah cabang ilmu yang mempelajari tentang bagaimana mengumpulkan, menganalisis dan menginterpretasikan data. Atau dengan kata lain, statistika menjadi semacam alat dalam melakukan suatu riset empiris.

Penggunaan metode statistik yang benar dapat memberi wawasan yang lebih memadai tentang dunia di sekitar kita. Contoh, sebagai bagian dari upaya menjaga kualitas air. Dilakukan pemantauan air di lima tempat melalui alat kontrol kontaminasi air. Konsentrasi kontaminan dalam per juta (ppm) diukur dari masing-masing lokasi pemantauan. Kemudian rata-rata dari lima pengukuran tersebut dihitung. Histogram berikut merangkum nilai-nilai kontaminasi rata selama 200 hari.



Andaikan telah terjadi tumpahan bahan kimia pada salah pabrik yang berjarak 1 km dari sumur. Tidak diketahui apakah tumpahan akan mencemari tanah pada daerah tumpahan dan, jika demikian, apakah tumpahan ini juga akan mempengaruhi kualitas air sumur yang jauh dari sumber tumpahan. Satu bulan setelah tumpahan, lima spesimen air dikumpulkan dari sumur, dan kontaminasi rata-rata adalah 15,5 ppm. Jika sebelumnya tumpahan, konsentrasi kontaminan rata bervariasi dari hari ke hari dengan nilai rata-rata 15,5 maka nilai rata-rata 15,5 ppm setelah tumpahan ini mengindikasikan bahwa kualitas air sumur tidak dipengaruhi oleh tumpahan.

Dalam menganalisis data, para ilmuwan menggambarkan persepsinya tentang suatu fenomena. Pendekatan awal yang umumnya digunakan untuk menjelaskan suatu fenomena adalah *statistika deskriptif*.

#### ▪ Statistika Deskriptif

*Statistika deskriptif* adalah tehnik yang digunakan untuk mensarikan data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap

orang. Proses kuantifikasi ini kemudian dituangkan dalam berbagai bentuk statistik sederhana, seperti rata-rata hitung dan ditampilkan dalam bentuk tabel dan grafik.

- **Statistika Induktif**

Pada kasus dimana tidak mungkin seluruh populasi diobservasi-karena berbagai alasan ekonomis ataupun praktis. Mengambil kesimpulan tentang suatu populasi berdasarkan data dari sampel yang terbatas merupakan tujuan dari suatu proses pengambilan keputusan *inferensial* atau *statistik induktif*

- Istilah dasar

## 2. Data dan Menampilkan Data

- Arti data

Data ialah bahan mentah yang perlu diolah sehingga hasilkan informasi atau keterangan, baik kualitatif maupun kuantitatif yang menunjukkan fakta. Sedangkan perolehan data seyogyanya relevan artinya data yang ada hubungannya langsung dengan masalah penelitian, *mutakhir* artinya data yang diperoleh masih hangat dibicarakan, dan diusahakan oleh orang pertama (data primer). Data yang sudah memenuhi syarat perlu diolah. Pengolahan data merupakan kegiatan terpenting dalam proses dan kegiatan penelitian. Keliruan memilih analisis dan perhitungan akan berakibat fatal pada kesimpulan, generalisasi maupun interpretasi. Hal ini perlu dikaji secara mendalam hal-hal yang menyangkut pengolahan data, supaya bisa memilih dan menentukan secara tepat dalam pengolahan data.

- Jenis data

Data menurut jenisnya ada dua yaitu data kualitatif dan data kuantitatif:

- **Data Kualitatif** yaitu data yang berhubungan dengan kategorisasi, karakteristik berwujud pertanyaan atau berupa kata-kata.

Contohnya: Wanita itu cantik, pria itu tampan, baik, buruk, senang, sedih, harga minyak turun harga dollar naik, rumah itu besar sekali. pohon itu rindang, laut Banda dalam sekali dsb. Data ini biasanya didapat dari wawancara dan bersifat subjektif sebab data tersebut ditafsirkan lain

oleh orang yang berbeda. Data kualitatif dapat diangkakan dalam bentuk ordinal atau ranking.

- **Data Kuantitatif** yaitu data yang berwujud angka-angka.

Contohnya: harga solar Rp 4.300,-/liter, yang diterima menjadi PNS 125 orang, penghasilan PT XYZ sebesar 2 milyar/tahun, Pembayaran SPP TK Harapan Rp 50.000/bulan, dan sebagainya. Data ini diperoleh dari pengukuran langsung maupun dari angka-angka yang diperoleh dengan, mengubah data kualitatif menjadi data kuantitatif. Data kuantitatif bersifat objektif dan bisa ditafsirkan sama oleh semua orang.

- **Skala pengukuran**

- **Skala Nominal** yaitu skala yang paling sederhana disusun menurut jenis (kategori) atau fungsi bilangan hanya sebagai simbol untuk membedakan sebuah karakteristik lainnya. Contoh :

- Jenis Kulit: Hitam (1), Kuning (2), Putih (3). Angka 1, 2, 3, hanya sebagai label saja.
- Suku Jawa (1), Madura (2), Bugis (3), Sunda (4), Batak (5), Minang (6)

- **Skala Ordinal** : skala yang didasarkan pada ranking, diurutkan dari jenjang lebih tinggi sampai jenjang terendah atau sebaliknya. Contoh :

Nilai	100	90	80	70
Rangking	1	2	3	4

- **Skala Interval** : skala yang menunjukkan jarak antara satu data dengan data yang lain dan mempunyai bobot yang sama.
- **Skala ratio** skala pengukuran yang mempunyai nilai nol (0) mutlak dan mempunyai jarak yang sama.

- **Populasi dan sampel**

Populasi merupakan objek atau subjek yang berada pada suatu wilayah dan memenuhi syarat-syarat tertentu berkaitan dengan masalah penelitian.

Ada dua jenis populasi, yaitu:

- **Populasi Terbatas:**

Adalah populasi yang mempunyai sumber data yang jelas batasnya secara kuantitatif sehingga dapat dihitung jumlahnya.

- **Populasi Tak Terbatas (Tak Terhingga)**

Populasi tak terbatas yaitu sumber datanya tidak dapat ditentukan batas-batasannya sehingga relatif tidak dapat dinyatakan dalam bentuk jumlah.

Berdasarkan sifatnya, populasi dapat digolongkan menjadi:

- *Populasi Homogen* adalah sumber data yang unsurnya memiliki sifat yang sama sehingga tidak perlu mempersoalkan jumlahnya secara kuantitatif.
- *Populasi Heterogen* adalah sumber data yang unsurnya memiliki sifat atau keadaan yang berbeda (bervariasi) sehingga perlu ditetapkan batas-batasnya, baik secara kualitatif maupun kuantitatif.

Dalam melaksanakan penelitian, walaupun tersedia populasi yang terbatas dan homogen, adakalanya peneliti tidak melakukan pengumpulan data secara populasi, tetapi mengambil sebagian dari populasi yang dianggap mewakili populasi (*representatif*). Bagian dari populasi yang mempunyai ciri-ciri atau keadaan tertentu yang akan diteliti disebut sampel. Pertimbangan logis penggunaan sampel adalah kepraktisan, keterbatasan biaya, waktu, tenaga dan kadang ada percobaan yang bersifat merusak (*destruktif*).

- **Pengolahan data**

Pengumpulan data dan analisis memungkinkan peneliti untuk menjawab pertanyaan berbagai fenomena. Proses analisis data dapat dilihat sebagai urutan langkah-langkah yang terarah dari rencana pengumpulan data hingga didapat kesimpulan berdasarkan informasi yang dihasilkan data. Proses ini dapat dibagi dalam enam langkah berikut:

1. Memahami sifat dari masalah.

Efektifitas analisis data memerlukan pemahaman tentang masalah penelitian, tujuan penelitian, apa pertanyaan yang diharapkan untuk terjawab dan memiliki arah yang jelas sebelum pengumpulan data untuk memastikan bahwa data yang dikumpulkan dapat menjawab pertanyaan yang diharapkan dapat terjawab.

2. Memutuskan apa yang harus diukur dan bagaimana mengukurnya.

Pada tahapan ini adalah memutuskan informasi apa yang dibutuhkan untuk menjawab pertanyaan yang menarik untuk dijawab. Misalnya, dalam sebuah studi tentang hubungan antara berat pemain sepak bola dan posisi bermain,

kita perlu untuk mengumpulkan data tentang pemain badan pemain dan posisi yang sering mereka tempati di lapangan. Pada kasus lain pilihan informasi tidak dapat diukur langsung. Misalnya, dalam sebuah studi tentang hubungan antara gaya belajar yang disukai gaya dan inteligensi. Pada kasus ini ada masalah dengan bagaimana mengukur gaya belajar yang baik dan apa ukuran kecerdasan akan yang akan digunakan. Disini penting untuk secara hati-hati dalam menentukan variabel apa yang harus dipelajari dan metode apa tepat untuk menentukan nilainya.

### 3. Pengumpulan data.

Langkah pengumpulan data sangat penting. Pada tahap ini harus diputuskan apakah sumber data yang ada saat ini telah memadai atau perlu pengumpulan data baru. Karena jenis analisis yang sesuai dan kesimpulan yang diambil tergantung pada bagaimana data dikumpulkan.

### 4. Data summarization dan analisis awal.

Setelah data dikumpulkan, Langkah selanjutnya biasanya melibatkan analisis awal yang meliputi meringkas Data grafis dan numerik. Analisis awal memberikan wawasan penting karakteristik data dan dapat memberikan panduan dalam memilih yang tepat metode untuk analisa lebih lanjut.

### 5. Analisis data.

Langkah analisis data dibutuhkan untuk memilih dan menerapkan metode statistik.

### 6. Interpretasi hasil.

Interpretasi hasil adalah langkah final untuk menjawab beberapa pertanyaan harus ditangani misal

1. Apa yang dapat kita pelajari dari data?
2. Kesimpulan apa yang dapat yang ditarik dari analisis?
3. Bagaimana hasil penelitian dapat jadi pemandu penelitian berikutnya?

Interpretasi hasil sering menghasilkan pertanyaan penelitian baru, yang pada gilirannya mengarah kembali ke langkah pertama.

## ▪ Menampilkan/ Menyajikan data

Data populasi atau sampel yang sudah terkumpul dengan baik, apabila digunakan untuk keperluan informasi, laporan atau analisis lanjutan hendaknya diatur,

disusun, dan disajikan dalam bentuk yang, jelas, rapih, serta komunikatif dengan cara menampilkan atau menyajikan data yang lebih menarik publik. Hal yang harus diperhatikan dalam penyajian data ini adalah siapa yang memerlukannya dan untuk apa. Dalam sistem manajemen kita mengenal ada tiga tingkatan manajerial dalam suatu organisasi yakni manajemen tingkat bawah, tingkat menengah dan tingkat atas/puncak. Masing-masing tingkatan ini memerlukan data sesuai dengan fungsinya masing-masing. Pada tingkat bawah data atau informasi yang dimiliki masih utuh dan lengkap seperti apa yang diperoleh di lapangan. Pada tingkatan selanjutnya data sudah disaring sedemikian rupa sehingga lebih ringkas dan padat. Dari tingkatan ini biasanya data diolah lebih lanjut untuk menghasilkan alternatif-alternatif dalam proses pengambilan keputusan oleh tingkat manajemen puncak. Adalah sesuatu yang tidak efisien bagi seorang pimpinan puncak untuk membaca seluruh data yang ada, sementara dia harus membuat keputusan yang cepat dan akurat.

Secara umum ada beberapa cara penyajian data statistik yang sering digunakan yaitu *tabel*, *grafk*, *diagram* baik untuk data ungroup dan group.

### **Penyajian data ungroup**

#### 1. Tabel data

Tabel statistik memuat semua data yang telah diringkas baik secara kuantitatif maupun kualitatif. Tabel statistik dapat dibuat dalam berbagai bentuk mulai dari tabel klasifikasi satu arah, dua arah, tiga arah dan lebih dari tiga arah. Yang mana yang akan dipilih tergantung dari ragam data yang diperoleh. Bentuk tabel biasanya akan terdiri dari komponen-komponen : judul tabel, judul kolom, isi tabel dan sumber. Berikut ini diberikan beberapa contoh tabel klasifikasi satu, dua dan tiga arah.

- Tabel satu arah

Adalah tabel yang hanya terdiri atas satu kategori atau karakteristik data. Contoh tabel satu arah sebagai berikut.

**TABEL 1 . EXPORT IKAN**

JUDUL KOLOM

ISI TABEL

SUMBER

Tahun	Vokume (Ribu ton)
(1)	(2)
1991	217,2
1996	240
1997	333,6
1998	301,4
1999	318,2
2000	336,2
2001	21,4

JUDUL

SUMBER : *Biro Pusat Statistik*

▪ Tabel dua arah

Adalah tabel yang terdiri atas dua kategori atau karakteristik data. Contoh tabel dua arah sebagai berikut

No	Jurusan	Asal Daerah Sekolah		Jumlah
		SMA	SMK	
(1)	(2)	(3)	(4)	(5)
1	Sistem Informasi	200	100	300
2	Teknik Informatika	240	121	361
3	Desain Grafis	155	102	257
Total		495	323	918

Sumber : Data buatan

▪ Tabel tiga arah

Adalah tabel yang terdiri atas tiga kategori atau karakteristik data. Contoh tabel dua arah sebagai berikut

No	Jurusan	Asal Daerah Sekolah	Jenis
----	---------	---------------------	-------



		Transportasi			
		SMA	SMK	Umum	Motor
(1)	(2)	(3)	(4)	(5)	(6)
1	Sistem Informasi	200	100	100	200
2	Teknik Informatika	240	121	300	61
3	Desain Grafis	155	102	100	107
Total		495	323	500	368

Sumber Data buatan

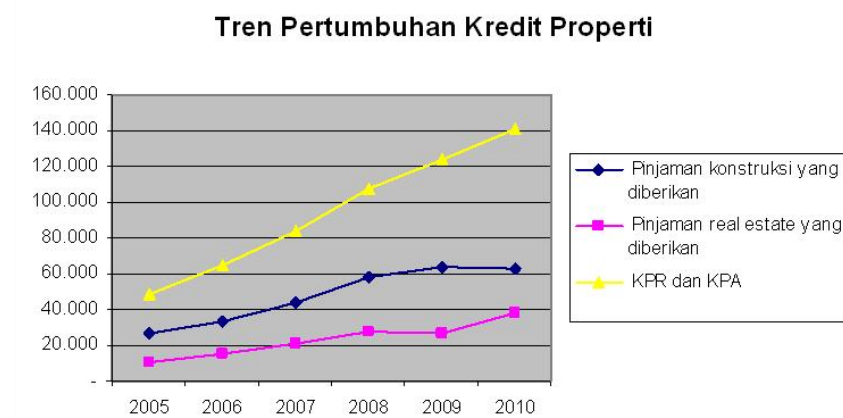
Pada prinsipnya untuk membuat tabel klasifikasi, tentukan terlebih dahulu apa yang menjadi penekanan utamanya, kemudian penekanan keduanya, ketiga dan selanjutnya, Data dengan penekanan utama ditempatkan pada kolom dan penekanan kedua pada baris. Untuk klasifikasi lebih tinggi lagi, proses ini diulang terus hingga selesai

## 2. Grafik

Data yang tersaji dalam bentuk angka-angka kerap kali tidak begitu disukai. Metode grafik atau dikenal pula dengan istilah diagram adalah cara lain yang dapat untuk menampilkan informasi yang ringkas namun memberikan arti yang tinggi. Sebuah grafik hendaknya dibuat sesederhana mungkin dan menekankan pada aspek-aspek penting dari data yang disajikan. Kecenderungan untuk memasukkan terlalu banyak informasi dalam sebuah grafik haruslah dihindarkan karena hal ini tidak akan memberikan informasi tambahan malah mungkin dapat membingungkan pembacanya. Berikut adalah beberapa jenis grafik yang umum digunakan untuk menampilkan data.

### ▪ Diagram garis

Diagram garis umumnya digunakan untuk melihat perkembangan suatu gejala atau fakta yang terkait dengan dimensi waktu. Dengan diagram ini maka konsumen akan lebih mudah melihat bagaimana berfluktuasinya nilai data yang sedang diamati. Untuk menggambarkannya sumbu tegak (sumbu-y) digunakan untuk menyatakan data yang diamati, sedangkan sumbu mendatar (sumbu-x) untuk menyatakan waktu.

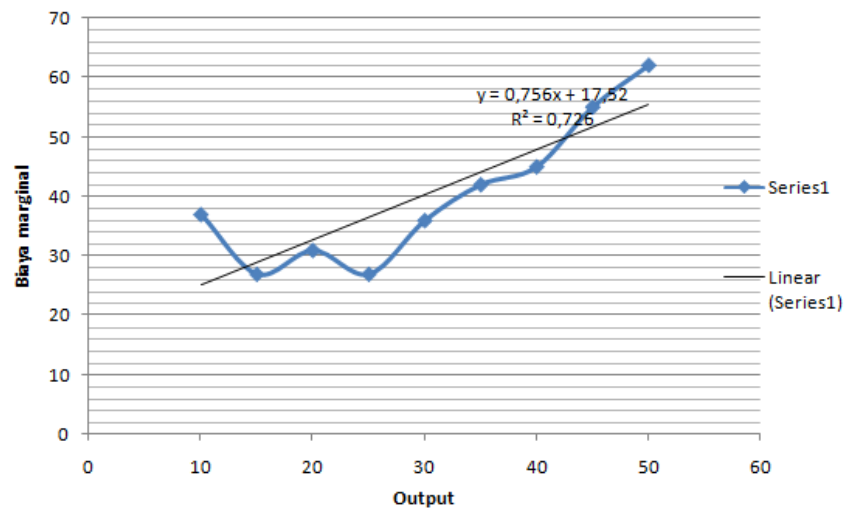


#### ▪ Kurva

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

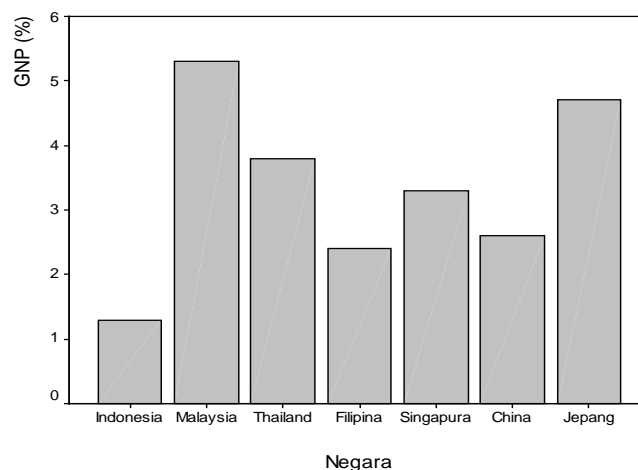
Biaya Marginal	Output (unit)
37	10
27	15
31	20
27	25
36	30
42	35
45	40
55	45
62	50

### Biaya Marginal Proses Produksi



#### ■ Grafik/diagram batang

Diagram batang merupakan salah satu bentuk penyajian data secara visual. Diagram batang sangat tepat digunakan jika variabel berbentuk kategori maupun atribut. Demikian pula data tahunan, asalkan jumlah tahunnya tidak terlalu banyak. Diagram batang bisa digambarkan secara vertikal (sejajar dengan sumbu-y) dan horizontal (sejajar dengan sumbu-x).



GNP beberapa negara Asia

#### ■ Grafik Lingkaran

Jika diagram batang dapat digunakan untuk membandingkan sejumlah variabel, selama variabel ini bisa diukur paling sedikit satu aspek, maka diagram lingkaran dapat digunakan bilamana variabel-variabel dijumlahkan ke dalam suatu jumlah tertentu sedemikian rupa sehingga

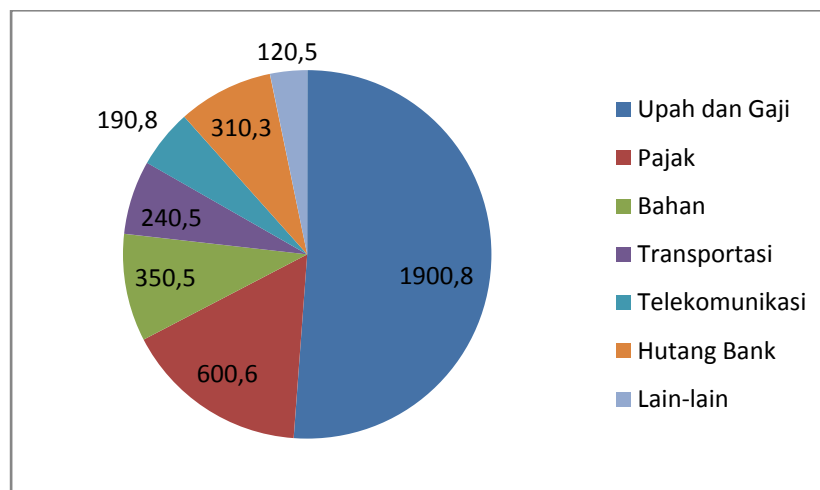
setiap variabel dapat dinyatakan sebagai persentase dari jumlah tersebut. Lingkaran yang memiliki jumlah sudut  $360^0$  dapat dibagi kedalam segmen-segmen untuk menunjukkan nilai penting relatif dari setiap variabel.

Pengeluaran PT. xyz tahun 2010

Jenis Pengeluaran	Jumlah (juta rupiah)
Upah dan Gaji	1900,8
Pajak	600,6
Bahan	350,5
Transportasi	240,5
Telekomunikasi	190,8
Hutang Bank	310,3
Lain-lain	120,5
Jumlah	3.714,00

SUMBER : *Databuatan*

Grafik



### Penyajian data group

Pada data dengan jumlah besar sering digunakan cara pengelompokan (group) data ke dalam interval kelas yang sama panjang. Tabel pengelompokan data ini dikenal dengan nama tabel distribusi. Keuntungan dari penggunaan cara ini adalah dapat terlihat bentuk distribusi/penyebaran datanya. Istilah-istilah dalam tabel distribusi.

- **Interval Kelas**
  - Tiap-tiap kelompok disebut interval kelas atau sering disebut interval atau kelas saja
- **Batas Kelas**
  - Batas bawah
  - Batas atas.
- **Tepi Kelas (Batas Nyata Kelas)**
  - Tepi bawah = batas bawah – 0,5
  - Tepi atas = batas atas + 0,5
- **Lebar kelas**
  - Lebar kelas = tepi atas – tepi bawah
- Tahapan Penyajian data group
  1. Tentukan range data
$$\text{Range data} = \text{Data}_{\max} - \text{Data}_{\min}$$
  2. Tentukan jumlah interval  
Tentukan jumlah kelas interval dengan rumus Sturges
$$K = 1 + 3,3 \log n \quad (n = \text{jumlah data})$$
  3. Tentukan panjang kelas interval p
$$p = \text{Range data} / K$$
  4. Tentukan batas bawah interval