

# BIOSTATISTIK

---

(Saptawati Bardosono)

# Definisi Statistik

---

Statistik merupakan kegiatan untuk:

1. Mengumpulkan data
2. Menyajikan data
3. Menganalisis data dengan cara tertentu
4. Menginterpretasi hasil analisis tersebut

Biostatistik, adalah statistik mengenai kehidupan

# Kegunaan Statistik

---

Sangat banyak, antara lain:

1. Ingin memperkirakan berapa rata-rata asupan kalori keluarga yang tinggal di daerah rawan pangan
2. Apakah ada hubungan antara banyaknya penyuluhan yang diberikan dengan peningkatan status gizi
3. Jika seorang anak semakin banyak makan makanan bergizi tinggi, apakah kecerdasan anak tersebut akan bertambah

# Pembagian Statistik (A)

---

1. Statistik Deskriptif: Bagaimana data dikumpulkan dan disajikan
2. Statistik Inferensi: Bagaimana menganalisis data dan menginterpretasikan hasil analisis tersebut

# Pembagian Statistik (B)

---

1. Statistik Parametrik, berhubungan dengan inferensi statistik yang membahas parameter2 populasi seperti rata2, proporsi, dll, dengan ciri parametrik adalah jenis data interval atau rasio serta distribusi data (populasi) adalah normal atau mendekati normal
2. Statistik Non-Parametrik, inferensi statistik tidak membahas parameter2 populasi dengan ciri adalah data nominal atau ordinal serta distribusi data (populasi) tidak diketahui atau bisa disebut tidak normal

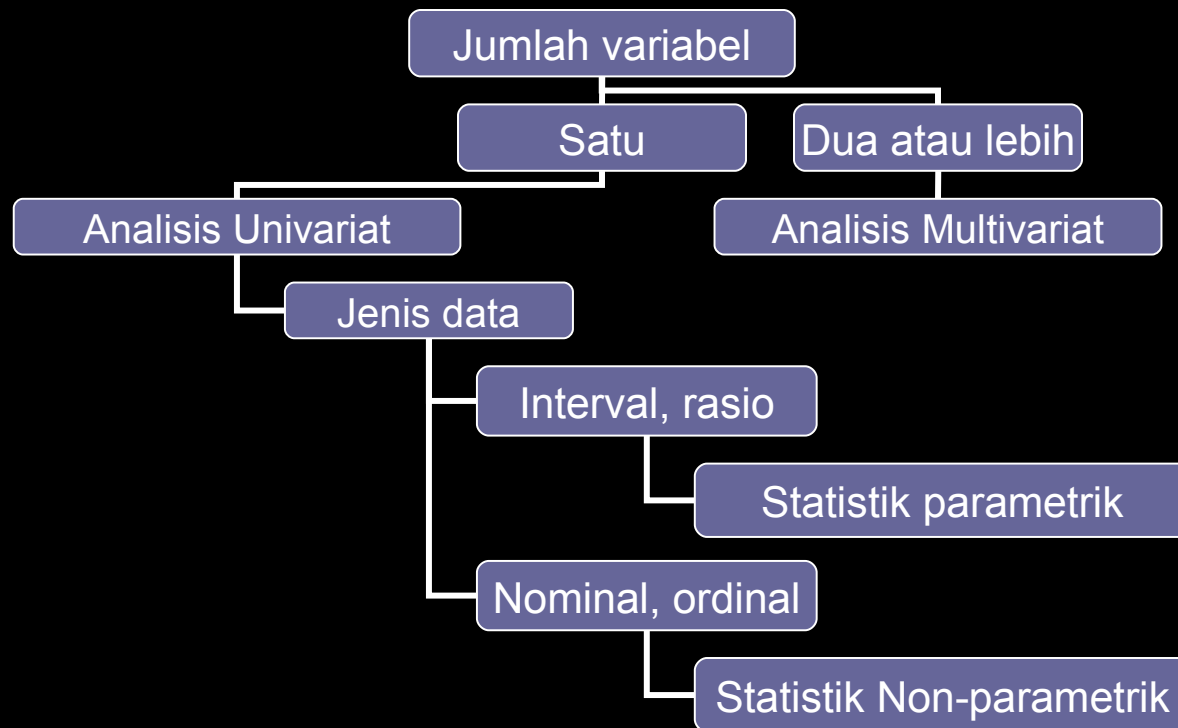
# Pembagian Statistik (C)

---

1. Analisis univariat, di mana hanya ada satu pengukuran (variabel) untuk sejumlah (n) sampel atau pengukuran beberapa variable tetapi masing2 dianalisis tersendiri (contoh: uji t, uji F, ANOVA dll)
2. Analisis multivariat, di mana ada dua atau lebih pengukuran (variabel) untuk sejumlah (n) sampel dan analisis antar variabel dilakukan bersamaan (regresi, korelasi, dll)

# Pembagian Statistik

---



# Data

---

- Data merupakan unsur terpenting dalam penelitian karena dalam data tersimpul bermacam informasi yang diperlukan
- Pembagian jenis data:
  - A. Berdasarkan proses dan hasil pengamatan:
    - 1) Sifat variabel
    - 2) Banyaknya variabel
    - 3) Bentuk susunan

# Data

---

B. Berdasarkan skala pengukuran:

- 1) Skala nominal
- 2) Skala ordinal
- 3) Skala interval
- 4) Skala rasio

# Jenis Data sesuai sifat variabelnya

---

1. Data Kualitatif, adalah data yang dinyatakan dalam bentuk bukan angka, misalnya jenis pekerjaan, status pernikahan, gender, tingkat kepuasan, dll (namun harus dikuantifikasikan agar dapat dianalisis lebih lanjut)
2. Data Kuantitatif, adalah data yang dinyatakan dalam bentuk angka, misalnya usia, tinggi, jumlah asupan, dll.
3. Data semi-kuantitatif, adalah data yang berasal dari data kualitatif kemudian diubah menjadi data kuantitatif dalam bentuk urutan, misalnya tingkat kepuasan akan pelayanan kesehatan, dll

# Jenis Data sesuai banyaknya variabel

---

1. Data univariat, yaitu data untuk 1 variabel
2. Data multivariat, yaitu data untuk variabel lebih dari satu (majemuk), misalnya efek obat hipertensi yang diukur dengan variabel tekanan sistol, diastol, dan denyut nadi sekaligus

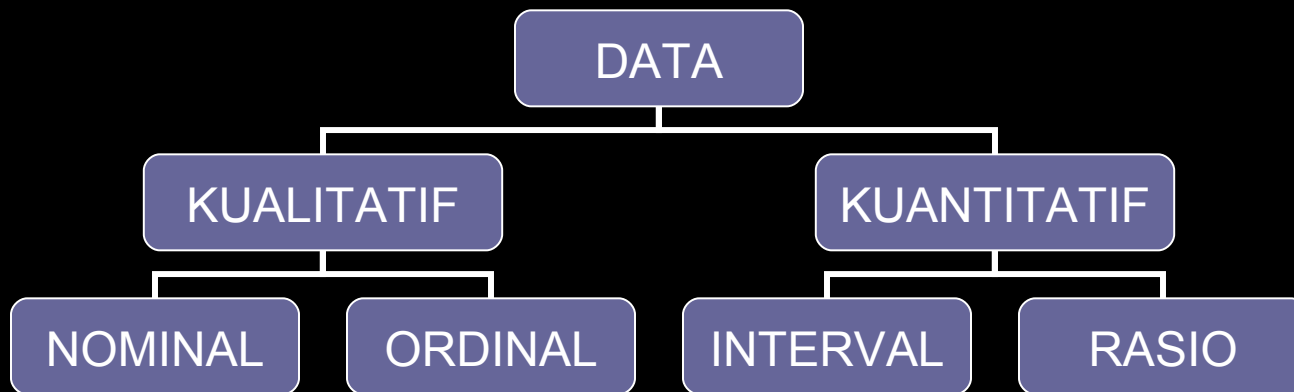
# Jenis Data sesuai bentuk susunannya

---

1. Bentuk susunan perbandingan bebas maupun berpasangan dengan perbandingan tunggal ataupun majemuk
2. Bentuk korelasi dan regresi sederhana maupun majemuk dengan bentuk korelasi sesuai sifat variabelnya (korelasi linier, non-linier, kategorial, biserial, rank, dll).
3. Bentuk sekuensial dengan pengumpulan data dalam waktu lama untuk dianalisis secara bertahap sampai terpenuhi besar sampel yang diperlukan

# Jenis Data sesuai skala pengukuran

---



# Data Nominal

---

- Skala ini mempunyai derajat pengukuran terendah
- Adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi, dengan ciri2:
  1. Posisi data setara
  2. Tidak bisa dilakukan operasi matematika
- Contoh: kebangsaan, kesukuan, jenis kelamin, stadium penyakit dll
- Analisis statistik: non-parametrik untuk komparasi dan korelasi

# Data Ordinal

---

Adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi tetapi di antara data tersebut terdapat hubungan, dengan ciri2:

- 1) Posisi data tidak setara
- 2) Tidak bisa dilakukan operasi matematika
- 3) Adanya sistim jenjang urutan, misalnya kelompok usia (0-1 tahun, 1-3 tahun dan 3-5 tahun)
- 4) Mempunyai ciri2 data nominal

Analisis statistik: non-parametrik

# Data Interval

---

- Adalah data yang diperoleh dengan cara pengukuran dengan jarak dua titik pada skala diketahui dan ciri2nya:
  1. Tidak ada kategorisasi
  2. Bisa dilakukan operasi matematika
  3. Ada jarak antar 2 kelompok satuan
  4. Mempunyai ciri2 seperti skala nominal dan ordinal
- Contoh: Suhu tubuh bisa diukur dalam skala selsius atau fahrenheit dengan masing2 mempunyai skala sendiri2
- Analisis statistik: parametrik dan non-parametrik kecuali koefisien variasi

# Data Rasio

---

- Adalah data yang diperoleh dengan cara pengukuran dan ciri2nya:
  1. Mempunyai ciri2 seperti skala nominal, ordinal dan interval
  2. Mempunyai titik 0 yang definitif
  3. Sembarang 2 titik pada skala tidak bergantung pada unit pengukurannya
- Contoh: Mengukur BB dalam gram dan pounds hasilnya identik
- Analisis statistik: parametrik dan non-parametrik

# Benar atau Salah?

---

- Kanker A secara statistik 2 kali lebih ganas dari kanker B

# Sumber Data

---

1. Data primer, adalah data yang diperoleh dari pengukuran yang dilakukan oleh peneliti sendiri melalui kegiatan penelitian
2. Data sekunder, adalah data yang diperoleh dari hasil pengukuran pihak lain, misalnya laporan, rekam medis, dll

# Pengumpulan Data

---

1. Observasi
2. Wawancara
3. Pemeriksaan klinis
4. Pemeriksaan antropometris
5. Pemeriksaan laboratorium
6. DII

# Pengumpulan Data

---

Kriteria penting dalam pengumpulan data:

- 1) Semua teknik dapat dipergunakan dalam semua jenis penelitian
- 2) Teknik pengumpulan data harus dapat dilakukan dengan mudah, cepat dan tepat sehingga reliable dan valid
- 3) Disebut reliable bila dilakukan oleh siapapun dan sesering apapun tetap memberi hasil yang sama
- 4) Disebut valid bila dengan teknik tersebut akan didapatkan data yang diinginkan

# Teknik observasi

---

## Jenis observasi:

- 1) Observasi terlihat, bila peneliti aktif berpartisipasi pada aktivitas yang sedang diteliti
- 2) Observasi sistematis, bila mempunyai kerangka/struktur yang jelas berisi semua faktor yang diperlukan dan materi observasi mempunyai lingkup lebih sempit dan terbatas serta terarah
- 3) Observasi eksperimental, bila ingin mengetahui adanya perubahan akibat eksperimental yang dilakukan

# Teknik observasi

---

Kendala dalam teknik observasi:

- 1) Yang akan diobservasi adalah perilaku individu sendiri dan perilaku individu dalam hubungannya dengan orang lain
- 2) Peneliti harus mencatat saja tanpa memberikan intepretasi langsung untuk menghindari terjadinya penyimpangan fakta
- 3) Catatan observasi dibuat selengkap mungkin
- 4) Untuk mencapai reliabilitas, pengamatan dilakukan oleh beberapa orang dan hasilnya dibandingkan dan didiskusikan.

# Teknik wawancara

---

Wawancara adalah suatu cara yang digunakan untuk mendapatkan data secara lisan dari responden melalui pertemuan/percakapan

Jenis wawancara:

- 1) Wawancara tidak terpimpin
- 2) Wawancara tepimpin
- 3) Wawancara bebas terpimpin

# Teknik wawancara

---

## Wawancara tidak terpimpin:

- 1) Tidak ada topik yang menjadi fokus wawancara
- 2) Pertanyaan2 tidak sistematis
- 3) Tidak menggunakan pedoman
- 4) Data akan sulit diolah/dianalisis

# Teknik wawancara

---

Wawancara terpimpin:

- 1) Dilakukan berdasarkan pedoman yang berupa kuesioner yang telah dipersiapkan secara baik
- 2) Pengumpulan data dan pengolahannya dapat berjalan dengan teliti
- 3) Hasilnya dapat disajikan secara kualitatif maupun kuantitatif
- 4) Dapat dilakukan oleh beberapa orang

# Teknik wawancara

---

Wawancara bebas terpimpin:

- 1) Merupakan cara kombinasi
- 2) Mempunyai ciri keluwesan dan arah yang jelas
- 3) Untuk menggali gejala2 lain yang bersifat pribadi

# Pengolahan Data

---

1. Editing
2. Koding
3. Data dasar
4. Analisis statistik
5. Penyajian data
6. Interpretasi data

# Pengolahan data

---

Editing, yaitu memeriksa kuesioner apakah mudah dibaca, semua pertanyaan diisi, keserasian jawaban ataupun adanya kesalahan2 lain

Koding, yaitu memberikan angka2/kode2 tertentu yang telah disepakati terhadap data tertentu

# Nilai Tendensi Tengah

---

Menggambarkan data dalam berbagai ukuran pusatnya seperti mean, median, modus, persentil dll

# Mean

---

- ❖ Merupakan rata2 secara aritmatika dari sejumlah nilai, dengan rumus

$$\hat{Y} = \Sigma(Y_i / n)$$

- ❖ Contoh: Untuk menentukan lama tinggal di RS setelah operasi, dikumpulkan data dari 5 pasien dengan lama tinggal sbb 1,3,2,4,5 hari. Rata2 lama tinggalnya adalah

$$\hat{Y} = (1+3+2+4+5 / 5) = 3 \text{ hari}$$

# Mean

---

## ❖ Ciri2 mean:

- 1) Bila semua nilai dikurangi dengan mean maka jumlahnya adalah 0
- 2) Bila jumlah sampelnya kecil maka nilai mean menjadi sangat sensitif terhadap nilai ekstrim, karena itu nilai mean akan menjerumuskan bila digunakan pada data yang tidak normal/tak simteris (skew). Contoh, bila lama tinggal pasien ke 5 adalah 50 hari, maka mean menjadi 12 hari dan bila lama tinggalnya 500 hari, maka mean menjadi 102 hari
- 3) Nilai mean tidak harus sama dengan salah satu nilai sampel

# Median

---

- ❖ Bila nilai sampel diurutkan dari besarnya, maka median adalah nilai yang ada di tengah (bagi jumlah sampel ganjil) atau rata2 dari 2 nilai di tengah (bagi jumlah sampel genap)
- ❖ Contoh:
  - 1) Lama tinggal di RS dari 9 pasien pasca-operasi diobservasi. Nilainya diurutkan dari besarnya sbb 1,1,3,4,8,9,12,13,dan 15, maka median adalah 8 hari
  - 2) Lama tinggal di RS dari 6 pasien lainnya diobservasi dan diurutkan sbb 1,6,7,10,11 dan 15, maka median =  $(7+10)/2 = 8,5$  hari

# Median

---

## ❖ Ciri2 median:

- 1) Bermanfaat untuk data yang tidak normal/skew karena tidak sensitif terhadap nilai ekstrim. Contoh, bila lama tinggal tertinggi 15 hari diganti dengan 150 hari, nilai median akan tetap
- 2) Pada data dengan distribusi normal/simetris, nilai mean dan median adalah sama
- 3) Nilai median tidak harus sama dengan salah satu nilai sampel

# Modus

---

- ❖ Adalah nilai yang tersering ditemukan dalam sekelompok data sampel
- ❖ Contoh:
  - 1) Dari data sampel sbb 1,4,3,1,2,5, maka nilai 1 tersering ditemukan, jadi nilai modus = 1.
  - 2) Dari data sampel sbb 2,4,2,3,1,5,1, maka nilai 1 dan 2 sama2 tersering ditemukan, jadi nilai modulusnya ada 2 (1 dan 2)

# Nilai Variabilitas

---

1. Sebaran (range) adalah perbedaan antara nilai terbesar dengan nilai terkecil dari suatu kelompok data. Karena mudah diukur, maka nilai ini digunakan untuk mengukur variabilitas secara sederhana dan cepat
2. Varians dan simpang baku (SD), merupakan ukuran<sup>2</sup> variabilitas yang lebih mewakili dibandingkan range karena menggunakan semua data dari kelompok data dan masing<sup>2</sup> penyimpangan dari mean

# Varians

---

Merupakan kuadrat rata2 penyimpangan data dari nilai mean (karena bila tidak dikuadratkan, jumlahnya = 0)

$$S^2 = \Sigma(Y_i - \hat{y}) / (n-1)$$

Varians:  $6/4 = 1.5 \text{ hari}^2$

Pasien	Lama tinggal ( $Y_i$ )	Deviasi ( $Y_i - \hat{y}$ )	( $Y_i - \hat{y}$ ) <sup>2</sup>	( $Y_i$ ) <sup>2</sup>
1	3	(3-3) = 0	0	9
2	5	(5-3) = 2	4	25
3	2	(2-3) = -1	1	4
4	3	(3-3) = 0	0	9
5	2	(2-3) = -1	1	3
Total	15	0	6	51

# Simpang Baku (SD)

---

Karena nilai varians menggunakan kuadrat dari deviasi, maka ukuran varians juga dikuadratkan (contoh: 1,5 hari<sup>2</sup>). Untuk menghindarkan hal tersebut maka nilai positif dari varians diakarkan dan disebut sebagai SD

$$s = \sqrt{s^2} = \sqrt{\sum(Y_i - \hat{y}) / (n-1)}$$

Jadi SD dari contoh =  $\sqrt{(6/4)} = 1,22$

# Simpang Baku (SD)

---

Untuk memudahkan perhitungan, berikut adalah rumus untuk menghitung SD

$$s = \sqrt{(\sum Y_i^2 - \sum Y_i^2)/(n-1)}$$

$$s = \sqrt{(51 - 15^2/(4))} = \sqrt{1,5} = 1,22$$

# Simpang Baku (SD)

---

Untuk memudahkan perhitungan, berikut adalah rumus untuk menghitung SD

$$s = \sqrt{(\sum Y_i^2 - \sum Y_i^2)/(n-1)}$$

$$s = \sqrt{(51 - 15^2/(4))} = \sqrt{1,5} = 1,22$$

# Data distribusi normal (Gaussian)

---

Kurva data dengan distribusi normal berbentuk lonceng dan simetris antara nilai mean nya, artinya:

- a) 50% data berada di atas nilai mean dan 50% di bawah nilai mean
- b) Hampir 68% data berada dalam 1 SD dari mean
- c) Hampir 95% data berada dalam 1,96 SD dari mean (atau 2 SD)
- d) Hampir 99% data berada dalam 2,58 SD dari mean (atau 3 SD)

# Nilai kemiringan dan kerampingan

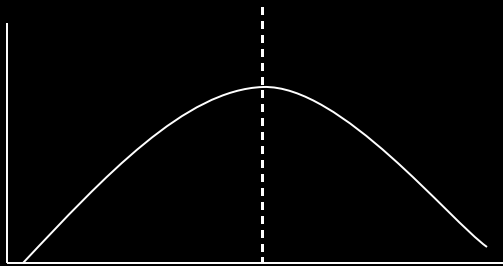
---

Nilai kemiringan (skewness) dan nilai kerampingan (kurtosis) digunakan untuk menentukan distribusi normal/simetris dari data bergantung dari bentuk kurva distribusi data

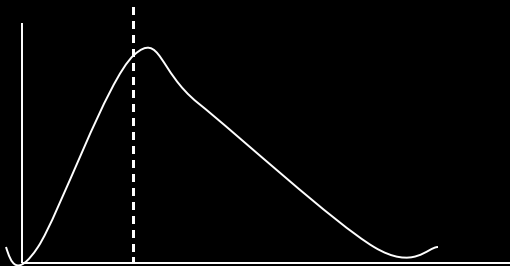
# Nilai kemiringan dan kerampingan

---

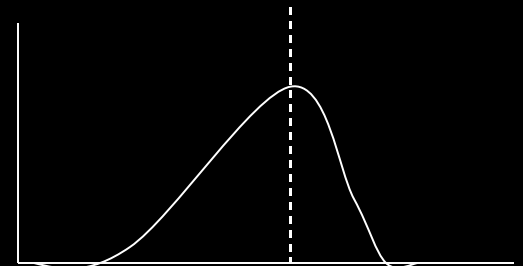
Distribusi normal/  
simetris



Miring (skew) ke  
kiri



Miring (skew) ke  
kanan



# Nilai kemiringan dan kerampingan

---

Contoh:

Bila diketahui skewness 0,052 dan standard error skewness 0,277, maka rasio skewness =  $0,052/0,277 = 0,0675$

Dengan kurtosis -1,181 dan standard error kurtosis 0,548, maka rasio kurtosis =  $-1,181/0,548 = -2,155 (\cong -2)$

Sehingga rasio skewness dan kurtosis keduanya berada di antara interval angka -2 dan +2 atau distribusi data normal atau simetris

# Distribusi normal?

---

Untuk mengetahui apakah sebuah distribusi normal atau mendekati normal:

1. Dengan uji Kolmogorov-Smirnov (KS), uji Shapiro-Wilk (SW), dll
2. Membuat grafik dengan prosedur tertentu

# Distribusi normal?

---

1. KS hitung  $<$  KS tabel
2. Signifikansi KS  $>0,05$
3. Signifikansi SW  $>0,05$
4. Histogram dalam area kurva normal

# Penyajian Data

---

Cara penyajian data:

- 1) Bentuk tulisan
- 2) Bentuk tabel
- 3) Bentuk grafik

# Bentuk tulisan

---

Tujuannya untuk memberikan suatu keterangan keseluruhan prosedur, hasil2 dan kesimpulan2 yang dibuat dengan menggunakan tulisan

Biasanya bentuk tulisan tidak dapat mencakup banyak gambaran statistik

# Bentuk tabel

---

Tujuannya untuk menyajikan suatu kumpulan data2 numerik dalam suatu bentuk tabel, di mana data2 disusun dalam baris dan kolom sedemikian rupa sehingga dapat memberikan perbandingan2

# Bentuk tabel

---

Paling sedikit ada 4 bagian dalam tabel, yaitu:

- 1) Judul
- 2) Stub
- 3) Box head
- 4) Body

# Judul tabel

---

- 1) Biasanya ditempatkan di atas tabel
- 2) Harus jelas, singkat, lengkap dan mengenai sasaran
- 3) Dapat menjawab pertanyaan: what, when dan where

# Stub

---

- 1) Kolom paling kiri, termasuk kepala kolom
- 2) Memberikan suatu keterangan secara terperinci tentang gambaran pada tiap baris dan badan tabel

# Box head

---

- 1) Bagian atas tabel termasuk kepala kolom
- 2) Memberi keterangan/penjelasan secara terperinci tentang gambaran tiap kolom badan tabel

Body, terdiri atas kolom2 dab hanya berisi angka2

# Bagian lain dari tabel

---

- 1) Total
- 2) Nomor tabel
- 3) Catatan kaki
- 4) Sumber data

# Jenis tabel

---

- 1) Tabel induk
- 2) Text table
- 3) Tabel distribusi:
  - a) tabel distribusi frekuensi relatif
  - b) tabel distribusi frekuensi kumulatif

# Tabel induk: Jumlah kelahiran menurut ras ibu, tempat melahirkan dan yang menolong kelahiran di kota X, 1999

Suku Ibu	Jumlah kelahiran				Jumlah
	Ditolong dokter		Oleh bidan	Oleh lain2	
	Di RS	nonRS			
Pribumi	606	14	10	4	634
WNI	334	94	30	16	474
Jumlah	940	108	40	20	1108

# Tabel distribusi frekuensi relatif: BB bayi yang mengunjungi Puskesmas X dalam bulan Juni 2004

BB (kg)	Jumlah bayi (%)
3,0 -	8
4,0-	28
5,0-	40
6,0-	20
7,0-8,0	4
Jumlah	100%

# Tabel distribusi frekuensi kumulatif:

BB bayi yang mengunjungi Puskesmas X dalam bulan Juni 2004

BB (kg)	Jumlah bayi
Kurang dari 3,0	0
Kurang dari 4,0	2
Kurang dari 5,0	9
Kurang dari 6,0	19
Kurang dari 7,0	24
Kurang dari 8,0	25

# Tabel distribusi frekuensi kumulatif:

BB bayi yang mengunjungi Puskesmas X dalam bulan Juni 2004

BB (kg)	Jumlah bayi
Lebih dari 3,0	25
Lebih dari 4,0	23
Lebih dari 5,0	16
Lebih dari 6,0	6
Lebih dari 7,0	1
Lebih dari 8,0	0

# Bentuk grafik

---

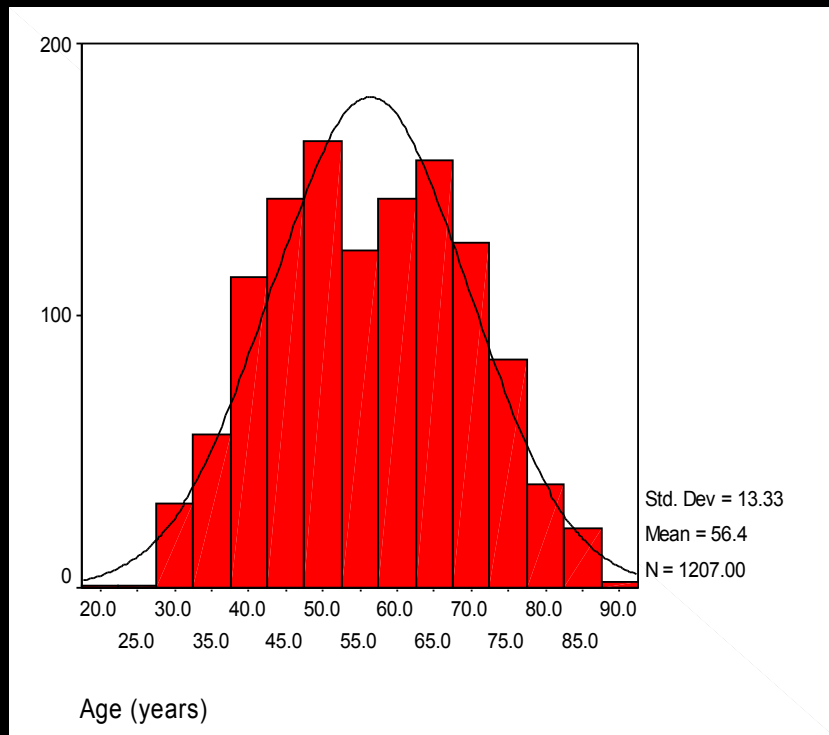
A. Untuk meramalkan sifat2 suatu kumpulan data:

- 1) Histogram
- 2) Frekuensi poligon
- 3) Ogive

B. Untuk membandingkan sifat2 yang ada dalam tabel:

- 1) Diagram garis
- 2) Diagram batang
- 3) Diagram pie
- 4) Diagram sebaran
- 5) Piktogram

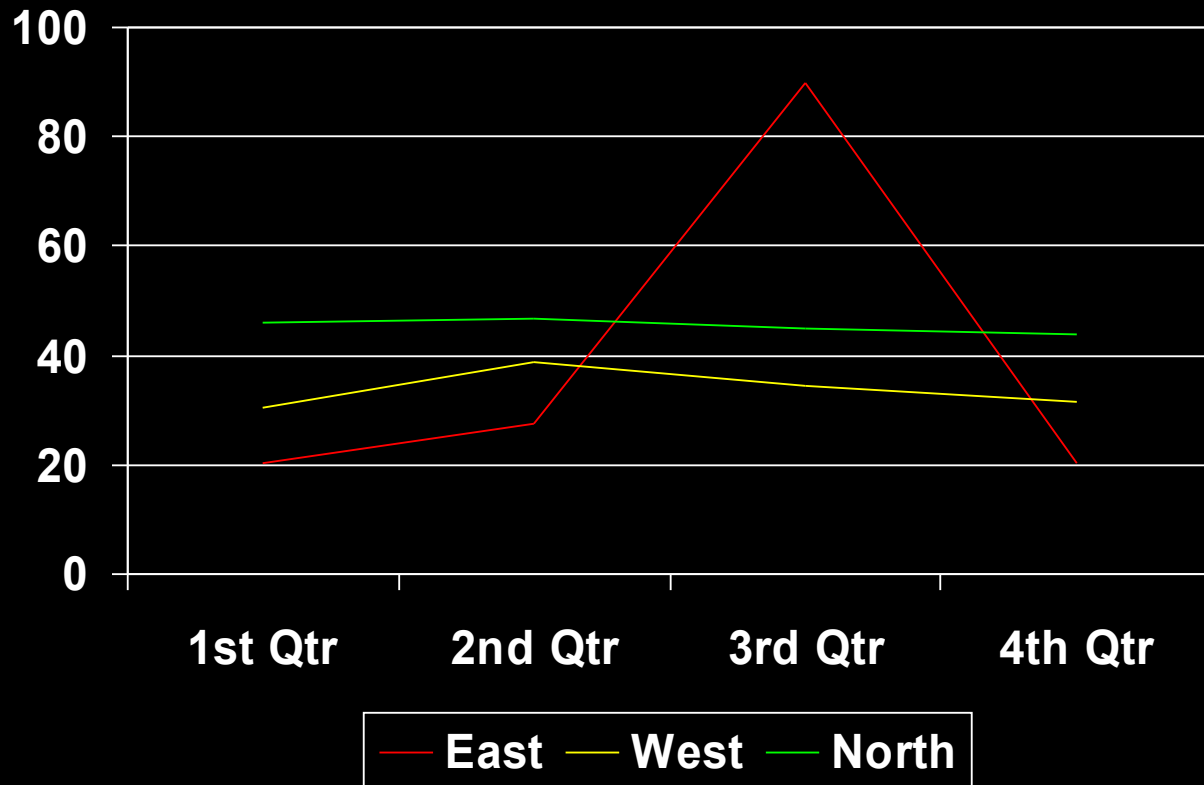
# Histogram



- Untuk menyajikan suatu distribusi frekuensi data2 kontinyu
- Harus ada batas kelas untuk tiap2 kelas interval
- Kelas interval di sumbu horizontal dan nilai2 frekuensi di sumbu vertikal
- Tiap kelas interval dalam bentuk bar vertikal

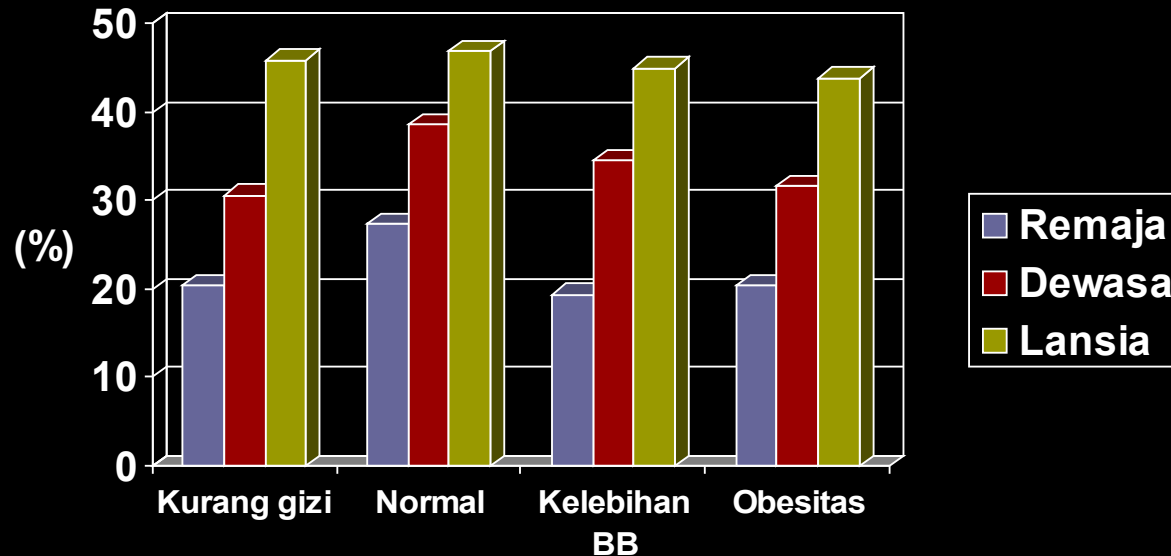
# Diagram garis: Untuk membandingkan perubahan2 2 variabel atau lebih

---



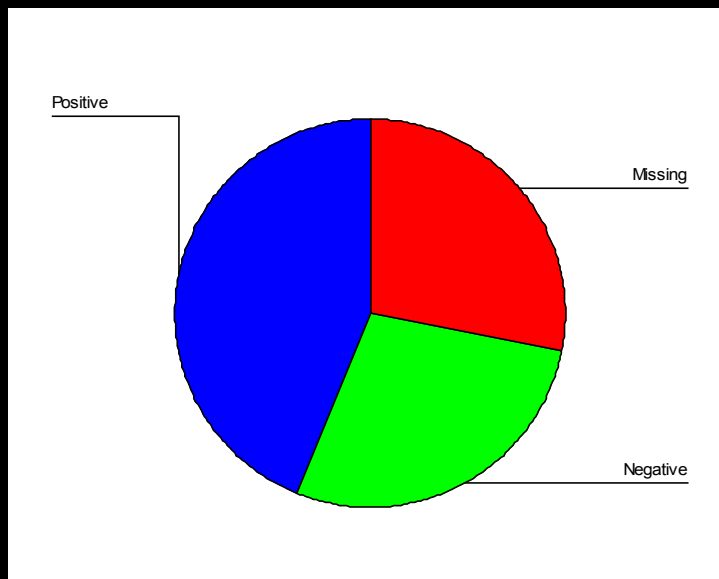
# Diagram batang: Untuk menyajikan suatu data dikret

## Hubungan antara status gizi dengan kelompok usia



# Diagram pie:

---

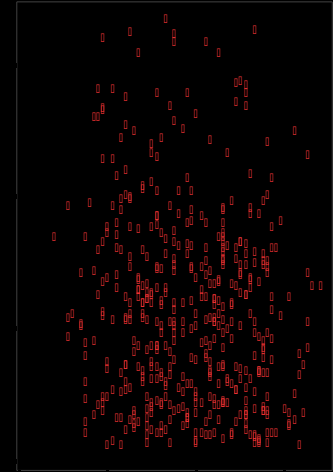
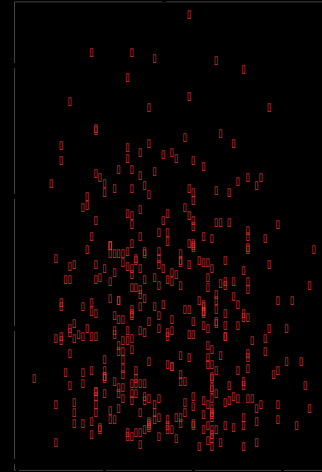


Untuk penyajian data kualitatif sebagai bagian2 komponen perbandingan dari keseluruhan

# Diagram sebaran titik

---

Untuk menyajikan sepasang pengamatan dari 2 variabel untuk memperlihatkan adanya atau tidak adanya hubungan antara 2 variabel tersebut



# Piktogram

---

Menyajikan data berupa gambar, di mana setiap gambar mewakili sebuah jumlah tertentu

Tahun

1980 \* \* \* \*

1990 \* \* \* \* \* \* \*

\* = 1000 penderita DHF

# Penyajian Data

---

Untuk data nominal, data disajikan dalam sebaran frekuensi

Contoh: Dari 100 klien yang datang ke klinik gizi dibagi menjadi 2 kelompok berdasarkan IMT yang disajikan sesuai tabel berikut

## Tabel: Sebaran frekuensi status gizi pengunjung klinik gizi

Status gizi	Frekuensi	Frekuensi relatif (%)
Kurang	20	20
Normal	35	35
Kelebihan BB	20	20
Obesitas	25	25
Total	100	100

# Penyajian Data

---

Untuk data ordinal, data disajikan dalam sebaran frekuensi, median dan modus

Contoh: Dari 20 klien lansia ditanyakan mengenai rasa formula susu baru dengan skala 1) sangat tidak suka 2) tidak suka 3) biasa saja 4) suka 5) suka sekali yang disajikan sesuai tabel berikut

## Tabel: Sebaran frekuensi kualitas rasa formula susu oleh 20 lansia

Skala rasa	Frekuensi	Frekuensi relatif (%)
Sangat tidak suka	3	15
Tidak suka	3	15
Biasa saja	4	20
Suka	7	35
Suka sekali	3	15
Total	20	100

# Penyajian Data

---

Median rasa formula susu adalah 3 atau biasa saja

Modus rasa formula adalah skala rasa 4 atau suka

# Penyajian Data

---

Untuk data interval dan rasio, data disajikan dalam segala bentuk cara statistik, namun dipengaruhi oleh bentuk distribusi data

Untuk distribusi data yang simetris, maka nilai mean dan median adalah sama

Contoh: Lama yang dibutuhkan untuk menurunkan kadar BB 5 klien adalah 3,5,4,4,dan 29 bulan

Mean = 9 bulan dan median = 4 bulan

# Penyajian Data

---

Data dapat disajikan dalam bentuk grafik seperti bar, pie, garis, histogram dsb.

Grafik bar (batang) digunakan untuk menyajikan data kualitatif, misalkan status gizi pada 3 kelompok usia.

Grafik bar juga dapat digunakan untuk menyajikan data 3 dimensi, misalkan status gizi pada 3 kelompok usia berdasarkan gender

# Penyajian Data

---

Histogram adalah sejenis grafik jenis bar yang digunakan untuk menggambarkan suatu sebaran frekuensi, dan juga untuk melihat apakah sebuah data berdistribusi normal/simetris atau tidak

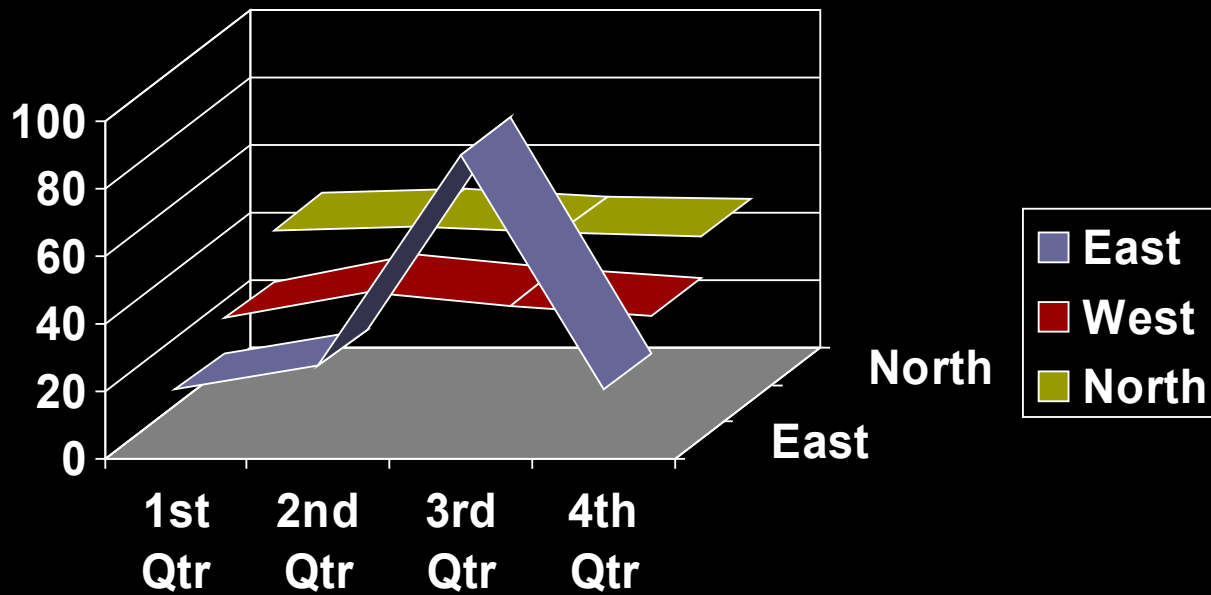
Boxplot digunakan untuk menggambarkan suatu sebaran frekuensi dan melihat apakah ada data yang ekstrim dan outlier

Scatterplot digunakan untuk memperlihatkan pola hubungan antara 2 variabel

# Penyajian Data

---

- Grafik pie, berbentuk lingkaran digunakan untuk menggambarkan data yang bersifat kualitatif (misal, gender, tingkat pendidikan, dll).
- Grafik garis, dot/titik, drop-line ataupun pita digunakan untuk menggambarkan proses/perkembangan suatu variabel berdasarkan waktu atau variabel terkait lainnya



# Teori Kemungkinan (probabilitas)

---

Untuk keperluan komunikasi dan mencegah terjadinya salah interpretasi dari suatu informasi medis di antara para ahli dan antara seorang ahli dengan pasiennya, maka yang terbaik adalah menentukan kemungkinan dengan istilah frekuensi relatif (proporsi).

Kemungkinan (P) bahwa suatu kejadian (E) akan terjadi yaitu  $P(E)$  diestimasi dengan rumus:

$$P(E) = \frac{\text{frekuensi terjadinya E}}{\text{frekuensi akan terjadinya E}}$$

# Teori Kemungkinan (probabilitas)

---

Nilai probabilitas berada antara 0 dan 1:

- a) Nilai 0 artinya kejadian tidak akan terjadi
- b) Nilai 1 artinya kejadian pasti terjadi
- c) Nilai 0,5 artinya kemungkinan kejadian akan terjadi sama dengan kejadian tidak akan terjadi

Jumlah dari probabilitas (frekuensi relatif) dari semua kejadian yang dapat terjadi dalam sampel harus 1 (atau 100%)

# Teori Kemungkinan (probabilitas)

---

Contoh: Dari 1047 sampel usia 40-59 tahun diamati kadar kolesterol serumnya, ingin diketahui kemungkinan bahwa kadar kolesterol serum seseorang yang dipilih secara acak berada pada interval 160 dan 179 mg/dL dan probabilitas lelaki usia 50 tahun dengan kolesterol serum kurang dari 200 mg/dL.

## Kelompok sebaran frekuensi kadar kolesterol serum pada lelaki sehat usia 40-59 tahun

Kolesterol serum (mg/dL)	Frekuensi	Frekuensi relatif (%)	Frekuensi kumulatif (%)
120-139	10	1,0	1,0
140-159	21	2,0	3,0
160-179	37	3,5	6,5
180-199	97	9,3	15,8
200-219	152	14,5	30,3
220-239	206	19,7	50
240-259	195	18,6	68,6
260-279	131	12,5	81,1
280-299	96	9,2	90,3
300-319	47	4,5	94,8
320-339	30	2,9	97,7
340-359	13	1,2	98,9
360-379	6	0,6	99,5
380-399	4	0,4	99,9
400-419		0	99,9
420-439	1	0,1	100,0
440-459		0	100,0
460-479	1	0,1	100,1
Total	1047	100,1	

---

Dari tabel tersebut didapatkan proporsi (probabilitas) dari sampel dengan interval kadar kolesterol 160-179 mg/dL adalah 37 dari 1047 atau  $37/1047 = 0,35$

Frekuensi kumulatif dari sampel dengan kadar kolesterol serum kurang dari 200 mg/dL adalah 15,8% atau  $(10+21+37+97)$  dari 1047 sampel =  $165/1047 = 0,158$

# Tabel 2X2

---

Tabel 2X2 digunakan untuk menjelaskan hukum probabilitas sbb:

Tabel terdiri dari 2 baris dan 2 kolom sehingga ada 4 sel. Total dari masing2 baris dan kolom disebut sebagai total marginal dan total dari baris dan kolom disebut sebagai total keseluruhan

# Tabel: Hasil tes diagnostik standar dan diagnostik eksperimental

	Penyakit +	Penyakit -	Total
Hasil +	7	4	11
Hasil -	3	86	89
Total	10	90	100

# Tabel 2X2

---

- Hasil disebut + apabila melebihi ambang batas yang ditentukan (cut-off point)
- Hasil disebut – apabila melebihi ambang batas yang ditentukan (cut-off point)
- Hasilnya: Dari 100 orang yang diteliti, 10 dinyatakan menderita penyakit berdasarkan tes diagnostik standar (gold-standard) dan 90 dinyatakan bebas penyakit. Dari 90 orang yang bebas penyakit, 86 mempunyai hasil tes - dan 4 mempunyai hasil tes +. Dari 10 orang yang sakit, 3 hasil tesnya – dan 7 hasil tesnya + berdasarkan tes diagnostik eksperimental.

# Tabel 2X2

---

- Bagaimana probabilitas dari 100 sampel dinyatakan berpenyakit berdasarkan tes diagnostik standar?  $P(\text{penyakit } +) = 10/100 = 0,1$
- Bagaimana probabilitas dari 100 sampel mempunyai hasil tes + dengan tes diagnostik eksperimental?  $P(\text{hasil } +) = 11/100 = 0,11$

# Aturan dalam hukum probabilitas

---

- Probabilitas gabungan dari 2 atau lebih kejadian klinis merupakan probabilitas yang dapat terjadi secara bersamaan, dan dituliskan sebagai  $P(A+B)$

Contoh: Berapa probabilitas sampel yang bebas penyakit mempunyai hasil tes –? Lihat kolom Penyakit – dan Hasil -, ada 86 dari 100 sampel yang secara bersamaan tanpa penyakit dan hasil tes – atau  $P(A+B) = 86/100 = 0,86$

# Aturan dalam hukum probabilitas

---

- Probabilitas terkondisi adalah probabilitas suatu kejadian akan terjadi setelah kejadian lain telah terjadi  $P(A/B)$

Contoh-1: Berapa probabilitas sampel yang kadar kolesterolnya antara 120-139 mg/dL dari mereka yang kadarnya di bawah 240 mg/dL? Mereka yang kadar kolesterolnya di bawah 240 adalah 523 dan yang diantara 120-139 = 10.  $P(B/A) = 10/523=0,19$

# Aturan dalam hukum probabilitas

---

Contoh-2: Berapa probabilitas sampel yang dinyatakan sakit dari mereka yang hasil tesnya +? Mereka yang hasil tesnya + = 11 dari 100 sampel. Dari 11 tsb yang dinyatakan sakit = 7.  $P(\text{penyakit+}/\text{hasil+}) = 7/11 = 0,64$ , artinya dari mereka yang hasil tesnya + ada 64% dinyatakan penyakit +.

# Aturan dalam hukum probabilitas

---

Probabilitas terkondisi versus probabilitas tak terkondisi:

- a) Tak terkondisi artinya diasumsikan hasil tes belum diketahui (pretest)=  
 $P(\text{penyakit+}) = 10/100 = 0,10$
- b) Terkondisi artinya dinyatakan penyakit + setelah diketahui hasil tes + (posttest) =  
 $P(\text{penyakit+} / \text{hasil+}) = 7/11 = 0,64$

# Rumus Probabilitas

---

Probabilitas terkondisi atau

$$P(A/B) = P(A \text{ dan } B) / P(B)$$

Contoh: Berapa probabilitas seseorang terkena penyakit mempunyai hasil tes +?

$$P(\text{penyakit+} / \text{hasil+}) = P(\text{penyakit + \& hasil+}) \\ \text{dibagi } P(\text{hasil+}) = 7/100 : 11/100 = 7/11 = \\ 0,64$$

# Rumus Probabilitas

---

Probabilitas gabungan atau

$$P(A \& B) = P(A / B) P(B) = P(A) P(B)$$

Contoh: Berapa probabilitas seseorang bebas penyakit & mempunyai hasil tes -?

$$P(\text{penyakit-} / \text{hasil-}) = P(\text{penyakit -} / \text{hasil-})$$

$$P(\text{hasil-}) = (86/89) (89/100) = 86/100 = 0,86$$

# Rumus Probabilitas

---

Probabilitas gabungan atau

$$P(A \text{ atau } B) = P(A) + P(B) - P(A \& B)$$

Contoh: Berapa probabilitas seseorang tanpa penyakit atau hasil tesnya -?

$$\begin{aligned} P(\text{penyakit- atau hasil-}) &= P(\text{penyakit-}) + \\ &P(\text{hasil-}) - P(\text{penyakit- dan hasil-}) = \\ &90/100 + 89/100 - 86/100 = 93/100 = 0,93 \end{aligned}$$