

STATISTIKA DAN PROBABILITAS (CIV -110)

PERTEMUAN 3 Pengolahan Data Statistik Deskriptif



OUTLINE

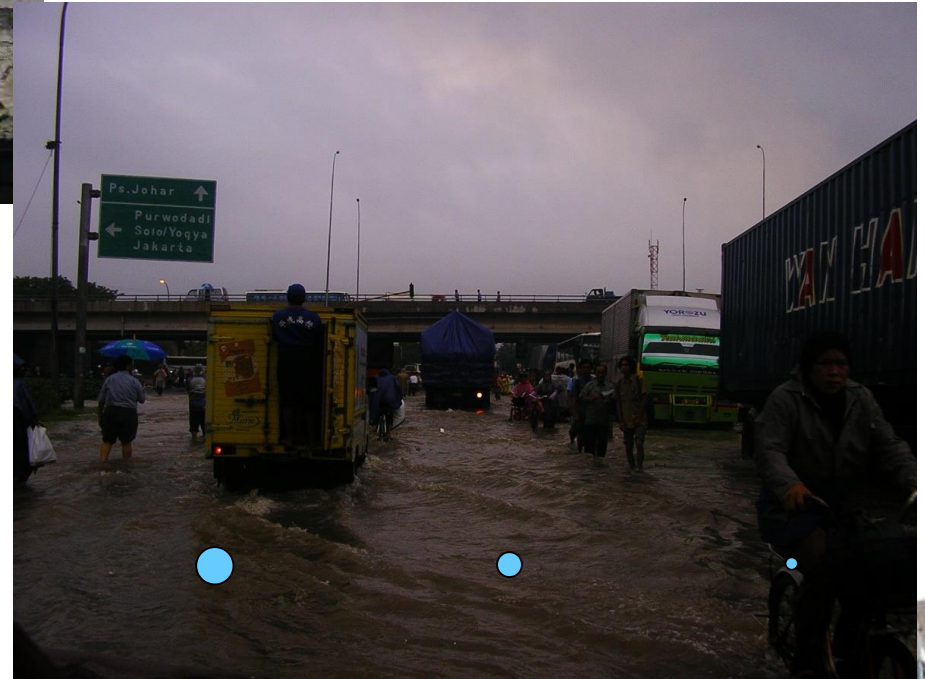
- Penggolongan data statistik
- Definisi statistik deskriptif
- Penyajian data
- Distribusi frekuensi
- Central tendency
- Dispersion
- Pengolahan dengan SPSS ver.23



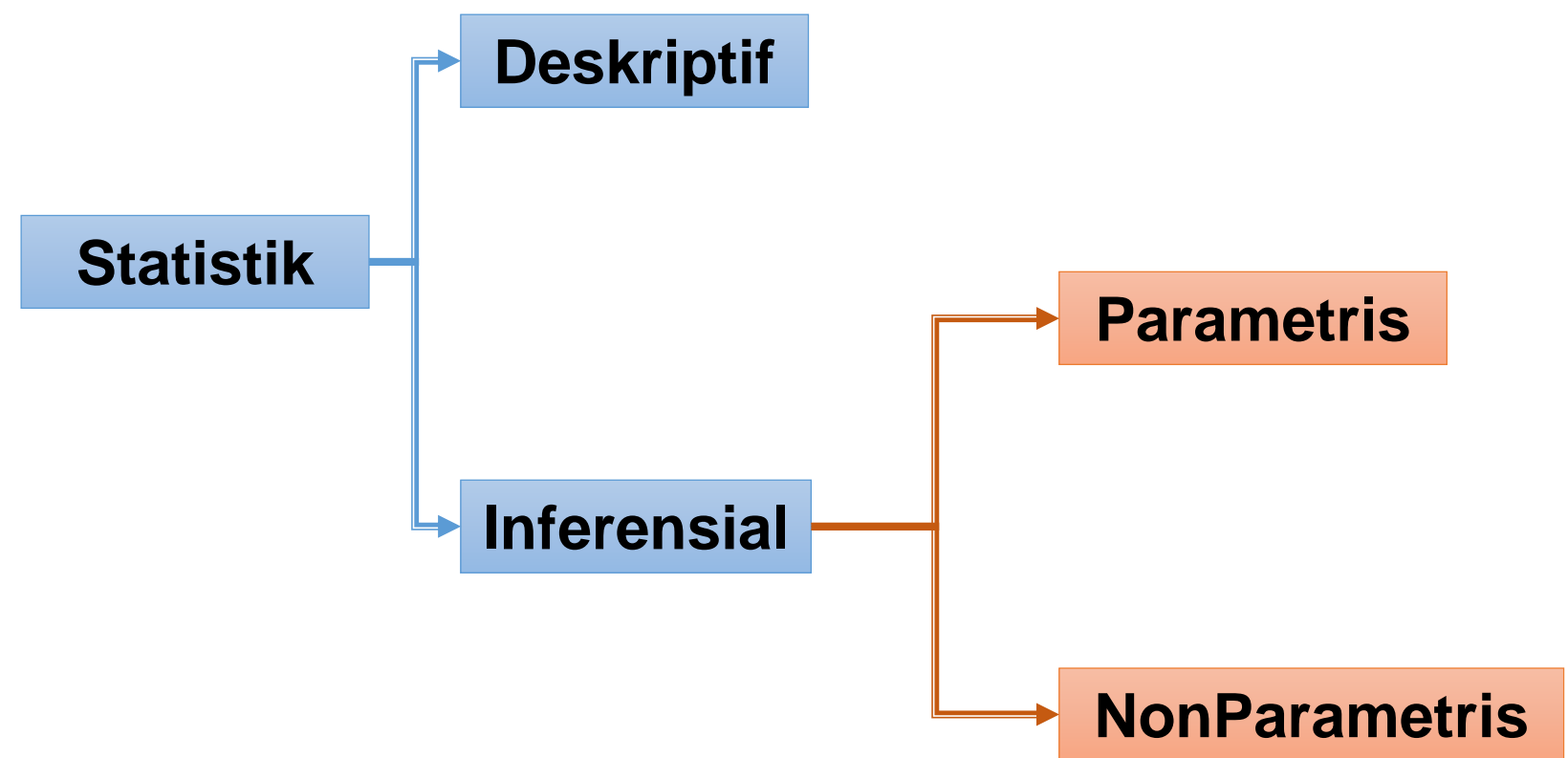


Diukur
(Kuantitatif)

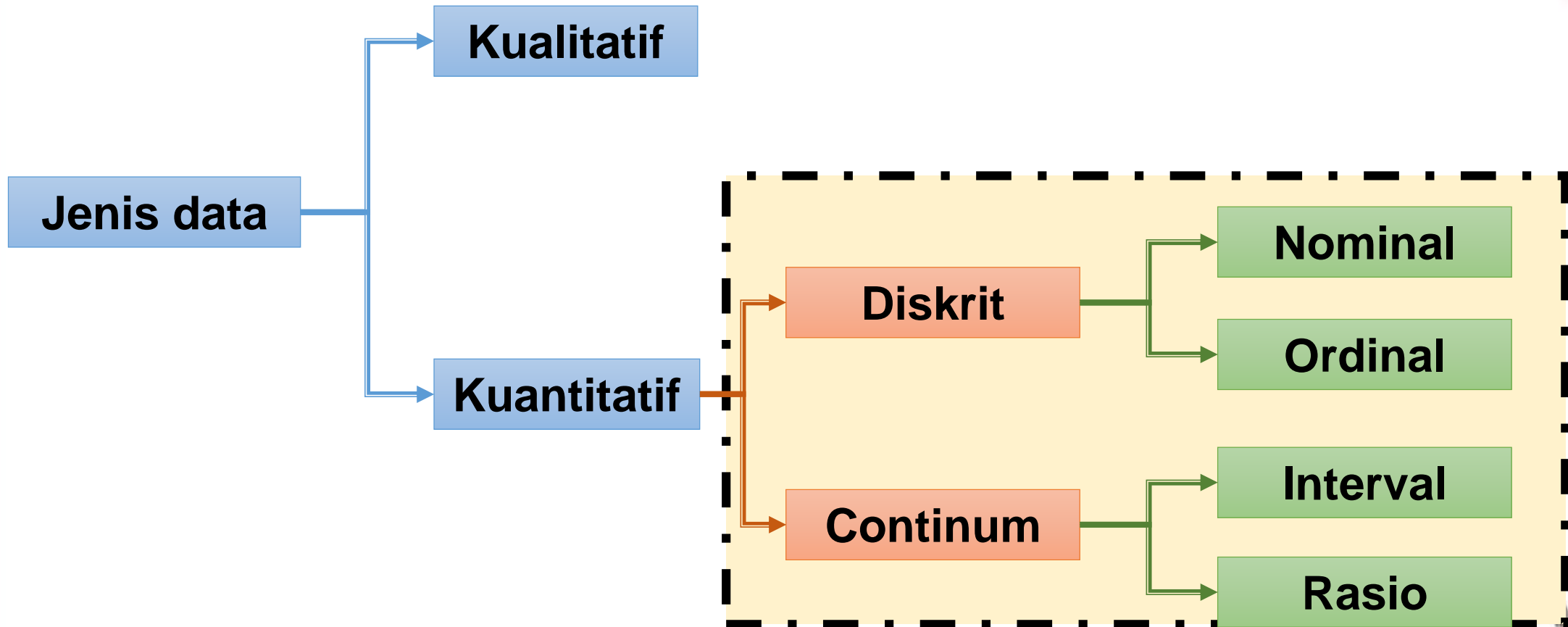
Didiskripsikan
(kualitatif)



Recall....



Penggolongan data statistik



Data Nominal

- ❑ Data yang berfungsi hanya sebagai pengganti nama atau sebutan gejala.
- ❑ Angka klasifikasi
- ❑ Contoh: jenis kelamin, jenis pekerjaan, tingkat pendidikan, asal daerah.
- ❑ Teknik statistik yang digunakan antara lain: *Uji Chi Kuadrat, Mc Nemartest, Uji Peluang Fisher*



Data Ordinal

- ❑ Data yang selain berfungsi sebagai pengganti nama atau sebutan suatu gejala juga menunjukkan bahwa masing-masing gejala mempunyai perbedaan intensitas
- ❑ Berdasarkan rangking atau tingkatan
- ❑ Contoh: kelas, juara, semester, peringkat
- ❑ Teknik statistik yang digunakan antara lain: *Uji kolmogorov smirnov, sign test, Mann Whitney, Korelasi Rank Spearman*



Data Interval

- ❑ Data yang mempunyai ciri-ciri skala ordinal, namun memiliki jarak yang sama dari ciri atau sifat yang diukur.
- ❑ Angka-angka interval data dapat dijumlahkan, dibagi, dan dikalikan.
- ❑ Contoh: nilai, skor IQ, temperatur
- ❑ Teknik statistik yang dapat digunakan antara lain: *Uji t, Anova, Pearson Product moment*



Data Rasio

- ❑ Data yang mempunyai ciri-ciri skala interval, namun mempunyai bilangan nol yang sebenarnya
- ❑ *Contoh : Berat, volume, jumlah orang*
- ❑ *Teknik statistic yang umum digunakan : Uji t, Anova, Pearson Product moment*



Resume

DATA KUALITATIF/ DATA NON METRIK

Mempunyai SIFAT TIDAK DAPAT DILAKUKAN OPERASI MATEMATIKA seperti:

- PENAMBAHAN/PENGURANGAN,
- PERKALIAN/PEMBAGIAN.

diukur pada skala:

- NOMINAL DAN ORDINAL

DATA KUANTITATIF/ DATA METRIK

Dapat disebut sebagai data berupa ANGKA DALAM ARTI SEBENARNYA. jadi, berbagai OPERASI MATEMATIKA DAPAT DILAKUKAN.

Diukur pada skala

- INTERVAL DAN RASIO.



Resume....

STATISTIK DESKRIPTIF

Berusaha menjelaskan/menggambarkan berbagai karakteristik **data**, seperti berapa nilai rata-rata (**mean**), seberapa jauh data-data bervariasi (**standard deviation**).

STATISTIK INDUKTIF /INFERENSIAL

Berusaha membuat berbagai inferensi terhadap sekumpulan data yang berasal dari suatu **sampel**. Tindakan inferensi tersebut seperti melakukan **perkiraan, peramalan, pengambilan keputusan**.



Penyajian Data

Tabel

- Tabel arah tunggal : *one way table*
- Tabel arah majemuk : *two way table, three way table*

Grafik

- Grafik batang (*bar chart*)
- Grafik garis (*line chart*)
- Grafik lingkaran (*pie chart*)
- Diagram pencar (*scatter diagram*)
- Kartogram (*cartogram*)
- Piktogram (*pictogram*)



Penyajian Data : Tabel distribusi frekuensi

- Tentukan data terkecil dan data terbesar
- Tentukan rentang data, $R = D_{maks} - D_{min}$
- Tentukan banyak kelas dengan kaidah empiris *Sturges* :

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

- Tentukan panjang kelas interval (l) dengan rumus : $l = \frac{r}{k}$
- Tentukan kelas-kelasnya sedemikian sehingga mencakup semua nilai data
- Tentukan frekuensi tiap kelas



Penyajian Data : Tabel distribusi frekuensi

Berikut disajikan data mentah yang diperoleh dari suatu survei penelitian berbentuk data interval :

66	76	79	86	89
97	125	110	105	111
111	102	101	121	117
122	120	101	115	100
115	95	122	139	144
137	128	138	136	136
92	97	106	81	112
107	106	133	133	103

Tentukan distribusi frekuensinya



PENYELESAIAN SOAL :

Langkah-langkah penyajian data dalam bentuk tabel distribusi frekuensi adalah :

- a) Banyak Data $n = 40$
- b) Data Minimum $D_{\min} = 66$
- c) Data Maximum $D_{\max} = 144$
- d) Rentang data $r = D_{\max} - D_{\min} = 144 - 66 = 78$
- e) Banyak kelas $k = 1 + 3,3 \log(n) = 1 + 3,3 \log(40) = 6,29$

dibulatkan menjadi **7**. (*Catatan*, banyak kelas selalu berupa bilangan bulat dan pembulatan selalu ke atas agar dapat menjangkau seluruh data.)

- f) Interval kelas (i) dengan aturan $i = \frac{r}{k} = \frac{78}{7} = 11,2$ dibulatkan ke atas menjadi **12**.
- g) Tentukan Tepi Kelas (*Class Boundaries*) dengan rumus $BBK-0,5$ dan $BAK+0,5$



PENYELESAIAN SOAL :

Kelas	Interval		Tepi Kelas	Tally	Frekuensi Absolut	F _{Relative} (f _r) %
	BBK	BAK				
1	66	- 77	65,5 – 77,5	II	2	5
2	78	- 89	77,5 – 89,5	IIII	4	10
3	90	- 101	89,5 – 101,5	IIII II	7	17,5
4	102	- 113	101,5 – 113,5	IIII IIII	10	25
5	114	- 125	113,5 – 125,5	IIII III	8	20
6	126	- 137	125,5 – 137,5	IIII I	6	15
7	138	- 149	137,5 – 149,5	III	3	7,5
	Jumlah			-	40	100

Keterangan: BBK = Batas Bawah Kelas

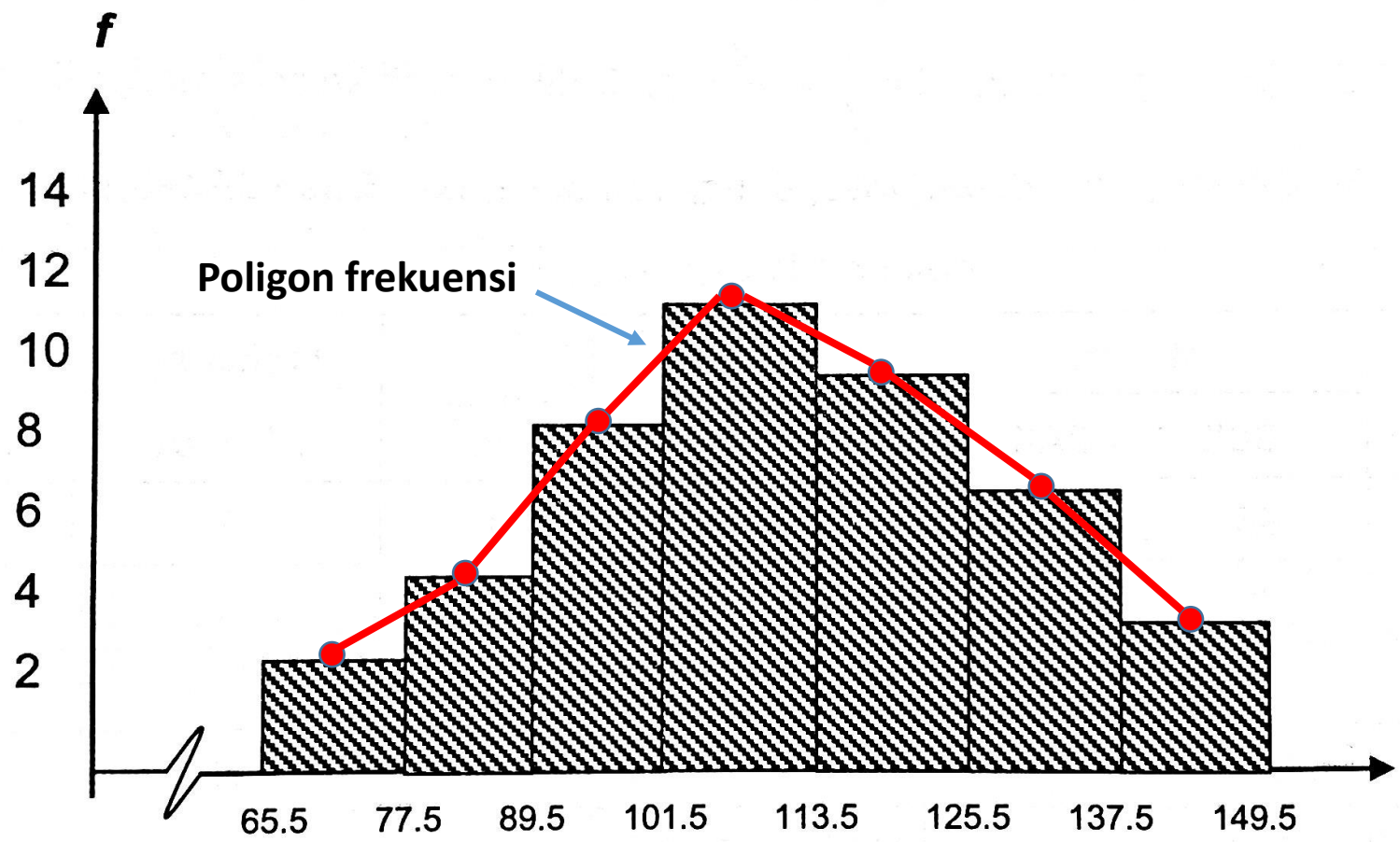
BAK = Batas Atas Kelas



PENYELESAIAN SOAL :

Histogram

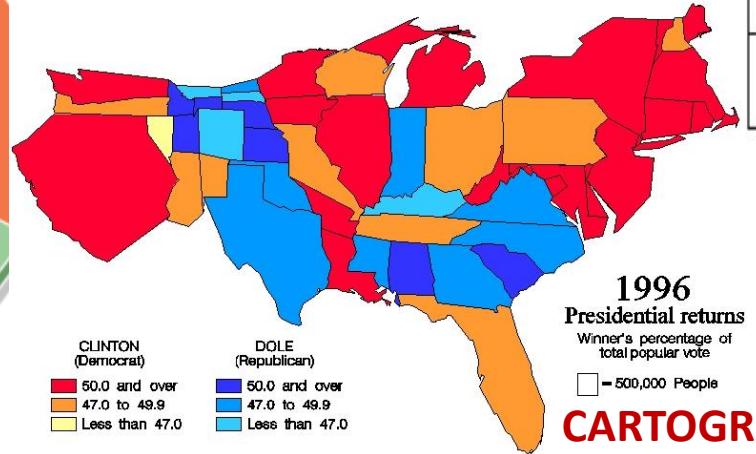
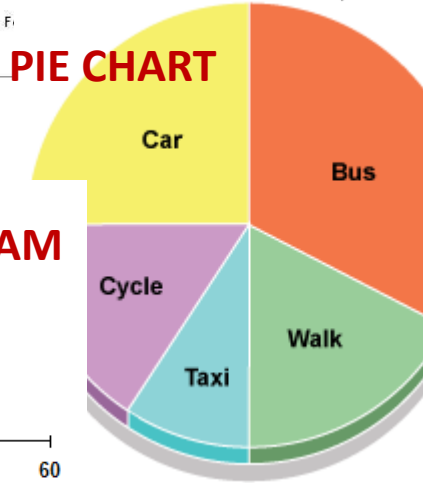
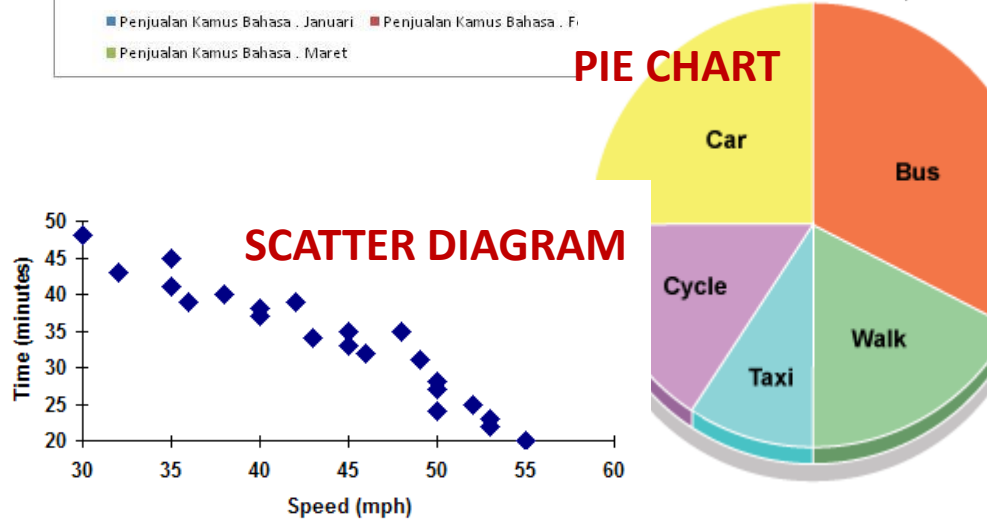
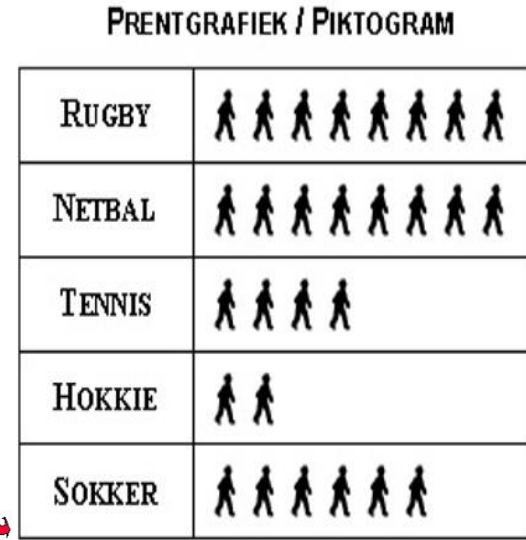
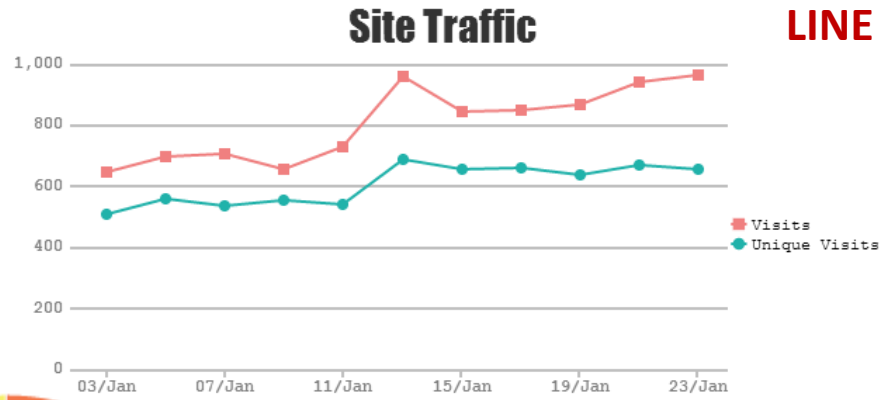
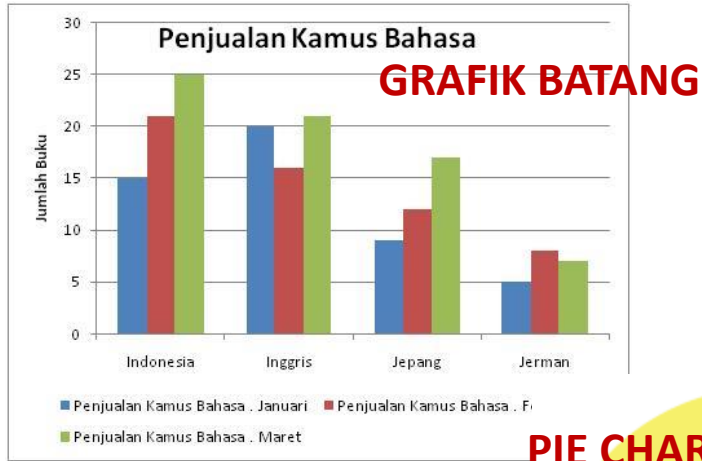
- Histogram merupakan salah satu bentuk penyajian data yang menggambarkan hasil penelitian yang menjelaskan pasang-surut keadaan data statistic dengan garis
- Gunakan tepi kelas untuk titik absis sumbu x dan frekuensi absolut untuk titik absis sumbu y
- Gambarkan balok dengan tinggi sesuai dengan frekuensi absolut dan lebar sesuai dengan tepi masing-masing kelas



Gambar 4.1 Histogram Distribusi Frekuensi



Contoh grafik



PIKTOGRAM = 20



ANALISIS STATISTIK DESKRIPTIF

1. UKURAN PEMUSATAN

- Mean
- Median
- Modus
- Kuartil

2. UKURAN SEBARAN

- Varian
- Standar Deviasi
- Range
- rentangan semi kuartil
- Rentangan persentil

3. UKURAN BENTUK

- Skewness
- Kurtosis



UKURAN PEMUSATAN (Central Tendency)

Rata-rata Aritmatik (Mean)

- ❑ Rata-rata aritmatik untuk sampel disimbolkan dengan \bar{X} atau \bar{Y}
- ❑ Rata-rata aritmatik untuk populasi disimbolkan dengan μ

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$



Median

- Merupakan letak data yang membagi 2 bagian yang sama atau disebut juga **nilai tengah**

$$\text{Letak posisi } Me = \frac{n + 1}{2}$$

n = banyaknya data

Modus

Nilai yang muncul dengan frekuensi terbanyak, tetapi **tidak semua** himpunan bilangan memiliki modus.



Setelah diurutkan dari nilai paling kecil ke paling besar (n= 40)

66	97	106	115	133	
76	97	106	117	133	
79	100	107	120	136	
81	101	110	121	136	----- Data ke-20
86	101	111	122	137	----- Data ke-21
89	102	111	122	138	
92	103	112	125	139	
95	105	115	128	144	

Nilai median adalah data ke-20,5 sehingga median = 110.5



Setelah diurutkan dari nilai paling kecil ke paling besar (n= 40)

66	97	106	115	133
76	97	106	117	133
79	100	107	120	136
81	101	110	121	136
86	101	111	122	137
89	102	111	122	138
92	103	112	125	139
95	105	115	128	144

Nilai modus bisa lebih dari satu bahkan tidak memiliki modus.

Untuk data di atas modus terkecil = 97 dan modus terbesar adalah 136



Kuantil

- ❑ Nilai yang membagi suatu jajaran data menjadi bagian yang sama
- ❑ Terdiri dari **kuartil, desil dan persentil**

1. Kuartil : ukuran letak yang membagi distribusi menjadi 4 bagian yang sama

$$Q_i = \text{nilai ke} - i \frac{n + 1}{4}, i = 1, 2, 3$$

2. desil : ukuran letak yang membagi distribusi menjadi 10 bagian yang sama

$$D_i = \text{nilai kecil} \frac{n + 1}{10}, i = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

3. persentil : ukuran letak yang membagi distribusi menjadi 100 bagian yang sama

$$P_i = \text{nilai ke} \frac{i(n + 1)}{100}, i = 1, 2, \dots, 99$$



Setelah diurutkan dari nilai paling kecil ke paling besar (n= 40)

	66	97	106	115	133	
Data ke-10	76	97	106	117	133	
Data ke-11	79	100	107	120	136	
	81	101	110	121	136	Data ke-20
	86	101	111	122	137	Data ke-21
	89	102	111	122	138	Data ke-30
	92	103	112	125	139	Data ke-31
	95	105	115	128	144	

Nilai $Q_1 = 27,5$
 Nilai $Q_2 = 110,5$
 Nilai $Q_3 = 124,25$

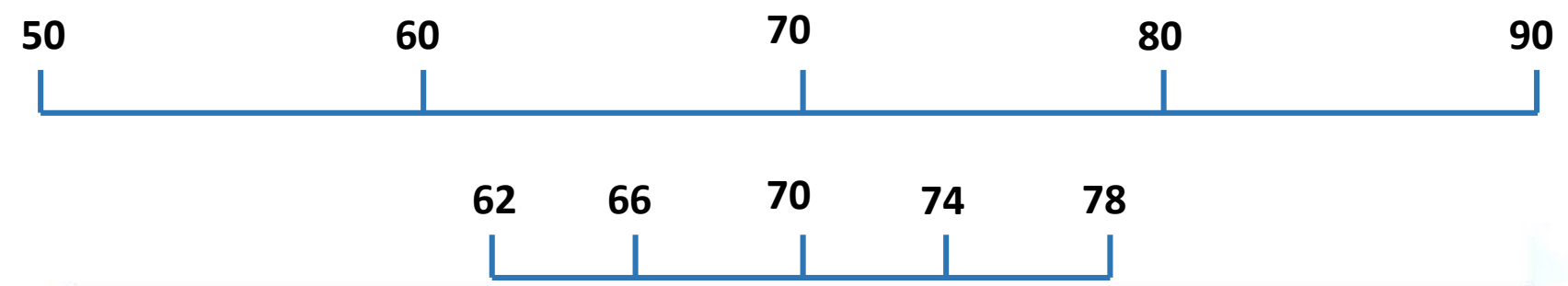


UKURAN PENYEBARAN (*Dispersion*)

- ❑ Tingkatan dimana distribusi data memiliki kecenderungan untuk menyebar di sekitar nilai reratanya.
- ❑ Ukuran dispersi yang kecil menunjukkan nilai data saling berdekatan (perbedaan kecil) dan sebaliknya.

Data A terdiri dari nilai :	62	66	70	74	78
Data B terdiri dari nilai :	50	60	70	80	90

Rata-rata kedua kelompok data tersebut =70, namun dengan sebaran berbeda



UKURAN PENYEBARAN (*Dispersion*)

Rentang data (range)

Selisih data terbesar dengan data terkecil

$$R = D_{maks} - D_{min}$$

Rentang semi kuartil (Q)

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

Q_1 Q_3 = kuartil kesatu dan ketiga

Rentang persentil 10 - 90

$$\text{Persentil } 10 - 90 = P_{90} - P_{10}$$

Standar Deviasi

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

varian

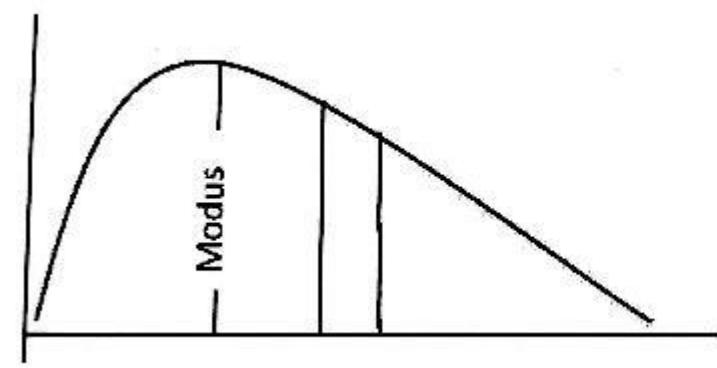
$$s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}$$



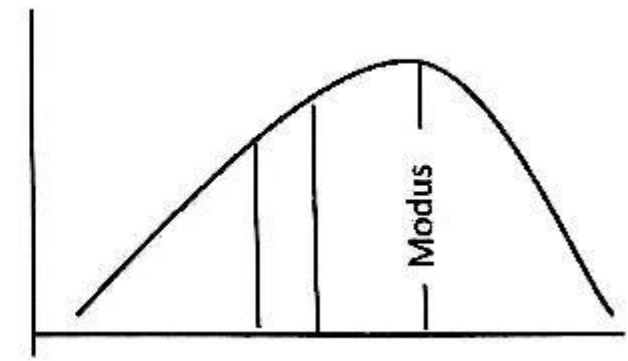
UKURAN BENTUK

Kemiringan (skewness)

Merupakan **derajat ketidaksimetrisan** atau penyimpangan dari kesimetrisan suatu distribusi



Kemiringan positif



Kemiringan negatif

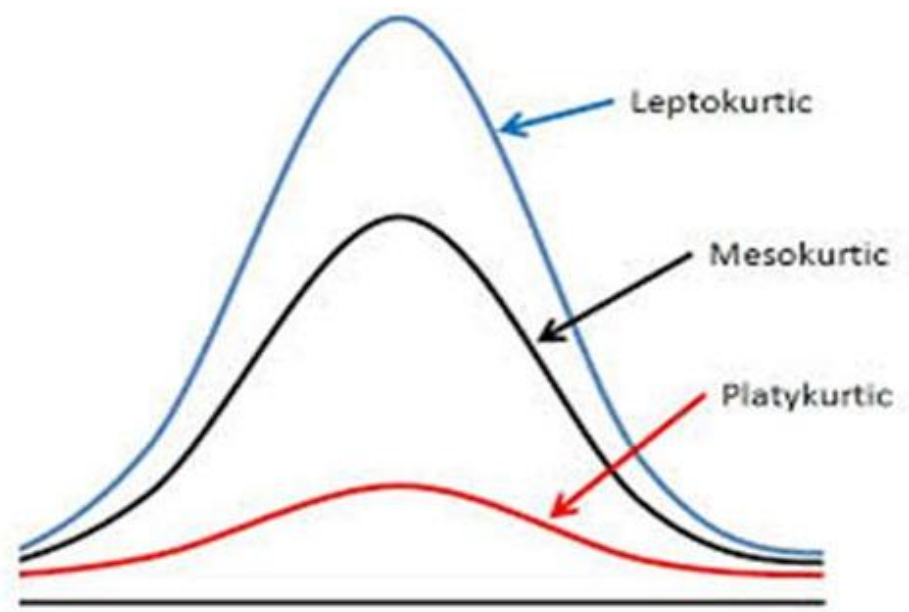
$$s = \frac{n}{(n-1)(n-2)} \sum (x_i - \bar{X})^3$$



UKURAN BENTUK

Keruncingan (kurtosis)

- Sangat erat kaitannya dengan kurva normal, Kurtosis menunjukkan keruncingan kurva
- Bila memiliki puncak lebih runcing dari kurva normal = **leptokurtik**
- Apabila kurva sebaran berbentuk kurva normal = **mesokurtik**
- Bila memiliki puncak lebih rendah dari kurva normal = **platikurtik**



$$Kurtosis = \left\{ \frac{n}{(n-1)(n-2)(n-3)} \sum \frac{(X_i - \bar{X})^4}{s} \right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

