




Metode Statistika (STK211)

Pertemuan III
Statistika Dasar (Basic Statistics)

Pertanyaan

Jika  punya data mengenai daya hidup dari baterai HP merk "XXX" 

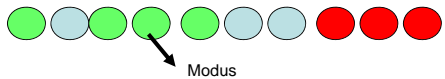
- Dimana "lokasi" atau "pusat" dari data? → ukuran pemusatan
- Seberapa besar variasi dari data → ukuran penyebaran

Ukuran Pemusatan

- Modus (Mode): Nilai pengamatan yang paling sering muncul
- Median: Pengamatan yang ditengah-tengah dari data terurut
- Quartil: Nilai-nilai yang membagi data terurut menjadi 4 bagian yang sama
- Mean: merupakan pusat massa (centroid) sehingga simpangan kiri dan simpangan kanan sama besar

Modus (Mode)

- Merupakan nilai pengamatan yang paling sering muncul
- Dalam satu gugus data dapat mengandung lebih dari satu modus
- Dapat digunakan untuk semua jenis data, tapi paling banyak digunakan untuk data kategorik atau data diskret dengan hanya sedikit nilai yang mungkin muncul



Median

- Pengamatan yang ditengah-tengah dari data terurut
- Nama lain dari percentil ke-50
- Nama lain dari kuartil 2 (Q2)
- Digunakan untuk menggambarkan lokasi dari data numerik
- Kekar terhadap adanya pencilan

Cara menghitung median contoh

Urutkan data dari terkecil sampai terbesar

Jika jumlah data **ganjil**, nilai median merupakan nilai di tengah

Data I: 2 8 3 4 1

Data terurut: 1 2 3 4 8

↑
Median

Cara menghitung median contoh

Urutkan data dari terkecil sampai terbesar

Jika jumlah data **genap**, nilai median merupakan rata-rata dari dua nilai di tengah

Data II: **2 8 3 4 1 8**

Data terurut: **1 2 3 4 8 8**

$$\text{Median} = (3+4)/2 = 3.5$$

Perhatikan data I dan data III

Data I terurut: **1 2 3 4 8**

Median

Data III terurut: **1 2 3 4 100**

Median

Secara umum langkah teknis untuk menghitung median contoh

- Urutkan data dari kecil ke besar
- Cari posisi median ($n_{med} = (n+1)/2$)
- Nilai median
 - Jika n_{med} bulat, maka $\text{Median} = X_{(n+1)/2}$
 - Jika n_{med} pecahan, maka $\text{Median} = (X_{(n)/2} + X_{(n)/2+1})/2$ (rata-rata dua pengamatan yang berada sebelum dan setelah posisi median)

Kuartil

- Nilai-nilai yang membagi data terurut menjadi 4 bagian yang sama
- Q0 (dibaca kuartil 0) merupakan nilai minimum dari data
- Q1 (dibaca kuartil 1) merupakan nilai yang membagi data 25% data di kiri dan 75% data di kanan
- Q2 (dibaca kuartil 2) merupakan median, membagi data menjadi 50%
- Q3 (dibaca kuartil 3) merupakan nilai yang membagi data 75% data di kiri dan 25% data di sebelah kanan
- Q4 (dibaca kuartil 4) merupakan nilai maksimum dari data
- Nilai Q1, Q2, dan Q3 kekar terhadap pencilan

Langkah Teknis memperoleh Kuartil (Quartile)

Metode Belah dua

- Urutkan data dari kecil ke besar
- Cari posisi kuartil
 - $n_{Q2} = (n+1)/2$
 - $n_{Q1} = (n_{Q2} + 1)/2 = n_{Q3}$, n_{Q2} posisi kuartil dua terpangkas (pecahan dibuang)
- Nilai kuartil 2 ditentukan sama seperti mencari nilai median. Kuartil 1 dan 3 prinsipnya sama seperti median tapi kuartil 1 dihitung dari kiri, sedangkan kuartil 3 dihitung dari kanan.

Perhatikan ilustrasi data I

- Posisi $Q2 = n_{Q2} = (5+1)/2 = 3$
- Posisi $Q1 = \text{Posisi } Q3 = (3+1)/2 = 2$

Data terurut: **1 2 3 4 8**


Median
Q1 Q3

Perhatikan ilustrasi data II

- Posisi $Q_2 = n_{Q_2} = (6+1) / 2 = 3.5$
- Posisi $Q_1 = \text{Posisi } Q_3 = (3+1)/2 = 2$

Data terurut: 1 2 3 4 8 8


↑
↑
↑
Q1
Median
Q3



Langkah Teknis memperoleh Kuartil (Quartile)

Metode Interpolasi

- Urutkan data dari kecil ke besar
- Cari posisi kuartil
 - $n_{q_1} = (1/4)(n+1)$
 - $n_{q_2} = (2/4)(n+1)$
 - $n_{q_3} = (3/4)(n+1)$
- Nilai kuartil dihitung sebagai berikut:
 - $X_{q_i} = X_{a,i} + h_i (X_{b,i} - X_{a,i})$
 - $X_{a,i}$ = pengamatan sebelum posisi kuartil ke-i, $X_{b,i}$ = pengamatan setelah posisi kuartil ke-i dan h_i adalah nilai pecahan dari posisi kuartil




Perhatikan ilustrasi data I

- Posisi $Q_2 = n_{Q_2} = (5+1) / 2 = 3$
- Posisi $Q_1 = 1/4(5+1) = 1.5$
- Posisi $Q_3 = 3/4(5+1) = 4.5$

Data terurut: 1 2 3 4 8

↑
↑
↑
Q1 = 1 + 0.5(2-1) = 1.5
Median
Q3 = 4 + 0.5(8-4) = 6




Perhatikan ilustrasi data II

- Posisi $Q_2 = n_{Q_2} = (6+1) / 2 = 3.5$
- Posisi $Q_1 = 1/4(6+1) = 1.75$
- Posisi $Q_3 = 3/4(6+1) = 5.25$

Data terurut: 1 2 3 4 8 8

↑
↑
↑
Q1 = 1 + 0.75(2-1) = 1.75
Median
Q3 = 8 + 0.25(8-8) = 8




Statistik 5 serangkai

Q2	
Q1	Q3
Q0	Q4


Berdasarkan metode Interpolasi

Data I	Data II												
<table style="width: 100%; text-align: center;"> <tr><td colspan="2">3</td></tr> <tr><td>1.5</td><td>6</td></tr> <tr><td>1</td><td>8</td></tr> </table>	3		1.5	6	1	8	<table style="width: 100%; text-align: center;"> <tr><td colspan="2">3.5</td></tr> <tr><td>1.75</td><td>6</td></tr> <tr><td>1</td><td>8</td></tr> </table>	3.5		1.75	6	1	8
3													
1.5	6												
1	8												
3.5													
1.75	6												
1	8												



Mean (rata-rata)

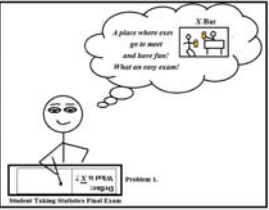

- Merupakan pusat massa (centroid)
- Jika menggambarkan populasi di tuliskan sebagai μ , huruf Yunani "mu"
- Jika menggambarkan contoh dituliskan sebagai \bar{x} , disebut "xbar"
- Digunakan untuk tipe data numerik
- Tidak bisa digunakan untuk tipe data kategorik dan diskret
- Sangat resisten terhadap pencilan



Langkah Teknis memperoleh mean

- Rata-rata (Mean)
 - Populasi: $\mu = \frac{\sum_{i=1}^N x_i}{N}$
 - Sampel: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$


Data I
(merupakan data contoh):
2 8 3 4 1
 $\bar{x} = \frac{2+8+3+4+1}{5} = 3.6$
Jangan dibulatkan!!!!

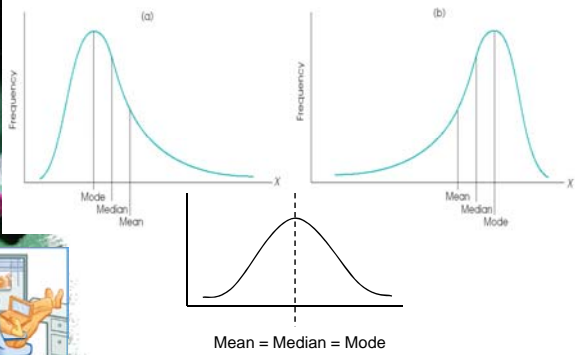
Perhatikan data I dan data III

Data I terurut: 1 2 3 4 8
 $\bar{x} = \frac{1+2+3+4+8}{5} = 3.6$
Median

Data III terurut: 1 2 3 4 100
 $\bar{x} = \frac{1+2+3+4+100}{5} = 22$
Median



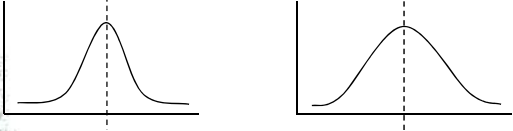
Kaitan antar bentuk sebaran dengan ukuran pemusatan



Mean = Median = Mode

Ukuran Penyebaran

- Menggambarkan suatu UKURAN KUANTITATIF tingkat penyebaran atau pengelompokan dari data
- Keragaman biasanya didefinisikan dalam bentuk jarak :
 - Seberapa jauh jarak antar titik-titik tersebut satu sama lain
 - Seberapa jauh jarak antara titik-titik tersebut terhadap rataannya
 - Bagaimana tingkat keterwakilan nilai tersebut terhadap kondisi data keseluruhan




Wilayah (Range)

- Merupakan selisih dari nilai terbesar – nilai terkecil
 $R = X_{max} - X_{min}$
- Hanya memperhitungkan nilai terkecil dan terbesar, sedangkan sebaran nilai antara dua nilai tersebut tidak diperhitungkan
- Resisten terhadap nilai yang ekstrim

Data I terurut: 1 2 3 4 8
R = 8-1 = 7

Data III terurut: 1 2 3 4 100
R = 100-1 = 99



Jangkauan antar Kuartil (Interquartile Range)

- Merupakan selisih antara kuartil 3 dengan kuartil 1
 $IQR = Q_3 - Q_1$
- Memperhitungkan sebaran antara nilai minimum dan nilai maksimum
- Kekar terhadap adanya nilai-nilai yang ekstrim (pencilan)

Statistik 5 serangkai dari data I (metode belah dua)


3	
2	4
1	8

$IQR = 4-2 = 2$

Statistik 5 serangkai dari data III (metode belah dua)

3	
2	4
1	100

$IQR = 4-2 = 2$



Deviasi

- Ukuran penyebaran yang lebih kompleks adalah bagaimana data tersebut mengelompok di sekitar rataannya
- Deviasi merupakan selisih dari data terhadap rataannya.
- Ukuran keragaman dari deviasi adalah rataaan deviasi = $\sum (x - \mu) / n$
- $\sum (x - \mu) / n \approx 0$

Data 1

Data	Deviasi
1	-2.6
2	-1.6
3	-0.6
4	0.4
8	4.4
Rataan	3.6

Ragam

Data 1

Data	(X-μ)	(X-μ) ²
1	-2.6	6.76
2	-1.6	2.56
3	-0.6	0.36
4	0.4	0.16
8	4.4	19.36
Rataan	3.6	5.84

- Untuk menghilangkan +/- maka deviasi dikuadratkan terlebih dahulu sebelum dirata-ratakan.
- Ukuran semacam ini disebut ragam = $\sum (x - \mu)^2 / n$
- $\sum (x - \mu)^2$ merupakan jumlah kuadrat dari deviasi disekitar rataannya

- Ragam (Variance)

➤ Populasi $\sigma^2 = \frac{\sum (x - \mu)^2}{N}$

➤ Contoh $s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$ Derajat bebas = db

Untuk menghitung ragam contoh maka perlu dihitung rataaan contoh, maka data terakhir tergantung dari data-data sebelumnya. Hanya 1 yang tidak bebas, sedangkan n-1 data lainnya bebas variasinya

Data 1 $\sigma^2 = \frac{\sum (x - \mu)^2}{N} = \frac{292}{5} = 58.4$ $s^2 = \frac{\sum (x - \bar{x})^2}{n-1} = \frac{292}{4} = 73$

Perhatikan permainan berikut

Banu mengajak Anda main tebak-tebakan. Banu mempunyai tiga kaleng. Salah satu dari kaleng tersebut berisi bola. Yang manakah yang berisi bola?



Jika bola tersebut dianggap sebagai rataaan sampel maka ada sebanyak 3-1 = 2 kaleng yang ditebak bebas → db = n-1

Jika kaleng I dan II Anda angkat namun tidak terdapat bola maka sudah pasti kaleng ke-3 yang berisi bola

- Simpangan baku (standard deviation)

➤ Ragam merupakan ukuran jarak kuadrat, sehingga untuk mendapatkan jarak yang sebenarnya adalah dengan mengakarkan ragam → simpangan baku

➤ σ simpangan baku populasi dan s simpangan baku sampel

Latihan :

- a. 3 9 7 4 10 3
- b. 4 9 3 8 6

Tentukan nilai :

Mean, Median, Q1, Q3, Ragam, Simpangan Baku, Range, dan IQR

untuk kedua gugus data di atas



Ilustrasi Data

No	Sex	Tinggi	Berat	Agama
1	1	167	63	Islam
2	1	172	74	Islam
3	0	161	53	Kristen
4	0	157	47	Hindu
5	1	165	58	Islam
6	0	167	60	Islam
7	1	162	52	Budha
8	0	151	45	Katholik
9	0	158	54	Kristen
10	1	162	63	Islam
11	1	176	82	Islam
12	1	167	69	Islam
13	0	163	57	Kristen
14	0	158	60	Islam
15	1	164	58	Katholik
16	0	161	50	Islam
17	1	159	61	Kristen
18	1	163	65	Islam
19	1	165	62	Islam
20	0	169	59	Islam
21	1	173	70	Islam

Data pada ilustrasi data diolah menggunakan MINITAB

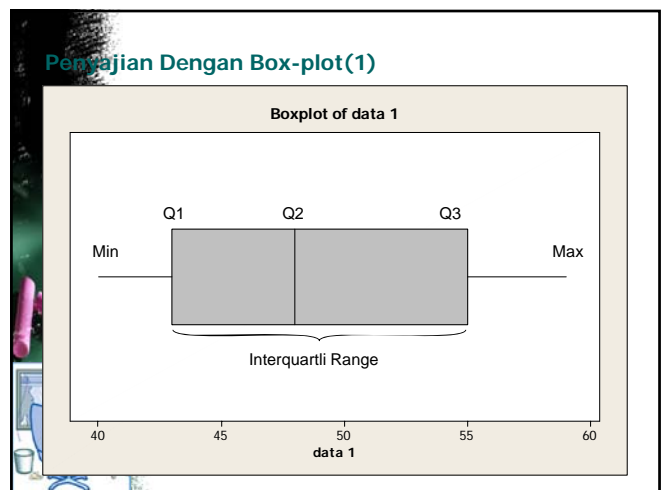
Descriptive Statistics: Tinggi, Berat

Variable	N	Mean	StDev	Variance	Minimum	Q1	Median	Q3	Maximum
Tinggi	21	163.81	5.85	34.26	151.00	160.00	163.00	167.00	176.00
Berat	21	60.10	8.86	78.49	45.00	53.50	60.00	64.00	82.00

Variable	Range	IQR
Tinggi	25.00	7.00
Berat	37.00	10.50



- Informasi yang diperoleh dari diagram kotak garis
- Melihat ukuran penyebaran dan ukuran pemusatan data
 - Melihat adanya data pencilan
 - Sebagai alat pembandingan sebaran dua kelompok data atau lebih



Cara membuat box plot

- Hitung Statistik lima serangkai
- Hitung Pagar Dalam Atas (PAD1) : $Q3 + 1.5(Q3-Q1)$
- Hitung Pagar Dalam Atas (PAD2) : $Q3 + 3(Q3-Q1)$
- Hitung Pagar Dalam Bawah (PBD1): $Q1 - 1.5(Q3-Q1)$
- Hitung Pagar Dalam Bawah (PBD)2: $Q1 - 3(Q3-Q1)$
- Identifikasi data. Jika data < PBD atau data > PAD maka data dikatakan outlier
- Gambar kotak dengan batas Q1 dan Q3
- Jika tidak ada pencilan : Tarik garis dari Q1 sampai data terkecil dan tarik garis dari Q3 sampai data terbesar
- Jika ada pencilan : Tarik garis Q1 dan atau Q3 sampai data sebelum pencilan
- Pencilan digambarkan dengan asterik

Me
Q1 Q3
Q0 Q4

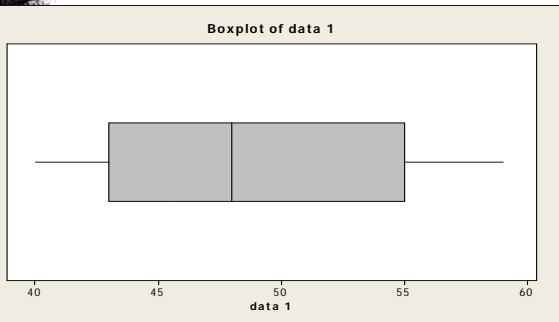
Ilustrasi (1)

- Statistik 5 serangkai dari data sbb:

Me	48
Q1 Q3	43 55
Min Max	40 59

- $PDA = 55 + 1.5 (55 - 43) = 73$
- $PDB = 43 - 1.5 (55 - 43) = 25$
- Tidak ada pencilan

Boxplot of data 1



Sebaran data tidak simetrik, karena nilai median lebih dekat ke Q1 → miring ke kanan
Tidak ada pencilan

Ilustrasi (4)

Stem-and-leaf of data 1 N = 23
Leaf Unit = 1.0

```

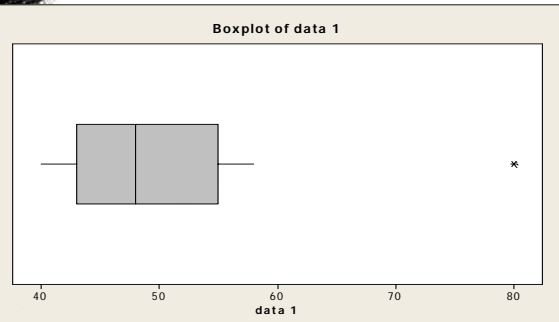
9 4 002233344
(5) 4 68899
9 5 02
7 5 556788
1 6
1 6
1 7
1 7
1 7
1 8 0
    
```

Me	48
Q1 Q3	43 55
Min Max	40 80

$PDA = 55 + 1.5 (55 - 43) = 73$
 $PDB = 43 - 1.5 (55 - 43) = 25$

Pencilan : 80

Boxplot of data 1



Sebaran data tidak simetrik, karena nilai median lebih dekat ke Q1 → miring ke kanan
Terdpat nilai pencilan (80)

Contoh data:

Jawa Barat			Jawa Tengah		
No.	Kota/Kab	Pert. Pend.	No.	Kota/Kab	Pert. Pend.
1	Pandeglang	2.15	1	Cilacap	1.28
2	Lebak	2.48	2	Banyumas	1.78
3	Bogor	4.52	3	Prabalingga	1.42
4	Sukabumi	2.51	4	Banarnegara	1.49
5	Cianjur	2.33	5	Kebumen	1.09
6	Bandung	3.31	6	Purworejo	0.62
7	Garut	2.35	7	Wonosobo	1.64
8	Tasikmalaya	2.15	8	Magelang	1.31
9	Ciamis	1.21	9	Boyolali	1.08
10	Kuningan	1.97	10	Klaten	1.19
11	Cirebon	2.73	11	Sukoharjo	2.10
12	Majalengka	2.01	12	Wonorejo	0.51
13	Sumedang	1.41	13	Karanganyar	2.07
14	Indramayu	2.53	14	Sragen	1.85
15	Subang	1.89	15	Grobogan	1.52
16	Purwakarta	2.32	16	Bera	1.27
17	Karawang	2.31	17	Rembang	2.08
18	Bekasi	3.57	18	Pati	1.62
19	Tangerang	4.04	19	Kudus	2.03
20	Sarangani	2.85	20	Jepara	1.87
21	Kota Bogor	2.60	21	Demak	1.38
22	Kota Sukabumi	1.48	22	Semarang	0.46
23	Kota Bandung	2.20	23	Temanggung	1.83
24	Kota Cirebon	2.51	24	Kendal	0.83
			25	Batang	1.70
			26	Pekalongan	1.90
			27	Pemalang	1.79
			28	Tegal	2.67
			29	Brebes	2.09
			30	Kota Magelang	1.25
			31	Kota Surakarta	1.99
			32	Kota Salatiga	2.30
			33	Kota Semarang	5.21
			34	Kota Pekalongan	1.95
			35	Kota Tegal	2.44

Rata-Rata:	Jabar	2.48
	Jateng	1.68
Minimum:	Jabar	1.00
	Jateng	1.00
Maksimum:	Jabar	23.00
	Jateng	34.00

