

PERTEMUAN 2: UKURAN STATISTIK

Contents

1. Data Tunggal dan Data Kelompok.....	2
2. Ukuran Statistik	2
2.1. Mean.....	3
2.1.1. Data Tunggal:	3
2.1.2. Data Kelompok:.....	3
2.2. Median.....	4
2.2.1. Data Tunggal:	4
2.2.2. Data Kelompok:.....	4
2.3. Modus.....	6
2.3.1. Data Tunggal:	6
2.3.2. Data Kelompok:.....	6
2.4. Kuartil	7
2.4.1. Data Tunggal:	8
2.4.2. Data Kelompok:.....	9
2.5. Desil.....	11
2.5.1. Data Tunggal:	11
2.5.2. Data Kelompok:.....	12
2.6. Persentil	14
2.6.1. Data Tunggal:	14
2.6.2. Data Kelompok:.....	15
3. Tambahan	17

1. Data Tunggal dan Data Kelompok

Data tunggal adalah data yang disajikan dengan cara yang sederhana. Data-data tersebut dapat dikatakan belum tersusun atau dikelompokkan ke dalam kelas-kelas interval.

Contoh:

74, 76, 77, 78, 80, 81, 82, 84, 85, 85, 85, 87, 88, 89, 90, 90, 91, 93, 96, 98.

Data kelompok adalah data yang sudah dikelompokkan ke dalam kelas-kelas interval.

Contoh:

	Kelas	Frekuensi	Titik Tengah
1	75 – 78	4	76
2	79 – 83	3	81
3	84 – 88	6	86
4	89 – 93	5	91
5	94 – 98	2	96

Pada R:

```
> # Mendeklarasikan data tunggal
> ungrouped_data = c(74, 76, 77, 78, 80, 81, 82, 84, 85, 85, 85, 87, 88, 89, 90, 90, 91, 93, 96, 98)
> ungrouped_data
[1] 74 76 77 78 80 81 82 84 85 85 85 87 88 89 90 90 91 93 96 98
>
> # Mendeklarasikan data kelompok
> Kelas = c("74-78", "79-83", "84-88", "89-93", "94-98")
> Frekuensi = c(4, 3, 6, 5, 2)
> Titik_tengah = c(76, 81, 86, 91, 96)
> grouped_data <- data.frame(Kelas, Frekuensi, Titik_tengah)
> print(grouped_data)
  Kelas Frekuensi Titik_tengah
1 74-78         4           76
2 79-83         3           81
3 84-88         6           86
4 89-93         5           91
5 94-98         2           96
>
> # Menghitung banyak data tunggal
> n_ungrouped = length(ungrouped_data)
> n_ungrouped
[1] 20
>
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)
> n_grouped
[1] 20
```

2. Ukuran Statistik

Dalam ilmu statistika terdapat dua ukuran, yaitu ukuran pemusatan dan ukuran penyebaran data.

Ukuran pemusatan data adalah suatu nilai tunggal yang mewakili keseluruhan data. Tujuannya untuk memberikan gambaran umum tentang data. Jenis-jenis ukuran pemusatan data yakni: mean, median, dan modus.

Sedangkan **ukuran penyebaran data** adalah suatu ukuran untuk mengetahui seberapa jauh data menyebar dari nilai rata-ratanya. Tujuannya untuk melihat seberapa beragam

data tersebut. Jenis-jenis ukuran penyebaran data yakni: kuartil, desil, persentil, variansi, dan standar deviasi. Jika ukuran penyebaran seperti variansi dan standar deviasi lebih fokus pada “jarak” antara data, maka kuartil, desil, dan persentil lebih kepada “posisi” data.

Pada pertemuan 2, kita akan membahas mean, median, modus, kuartil, desil, dan persentil.

2.1. Mean

Mean (\bar{x}) adalah nilai rata-rata dari keseluruhan data. Pada data tunggal, mean diperoleh dari membagi total nilai data ($\sum x_i$) dengan banyaknya data (n). Sedangkan data kelompok diperoleh dari hasil bagi antara perkalian frekuensi dan nilai tengah setiap kelas ($\sum f_i \cdot x_i$) dengan jumlah frekuensi ($\sum f_i$).

2.1.1. Data Tunggal:

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

Di mana: $\sum x_i$: total nilai data

n : banyaknya data

Pada contoh:

$$\bar{x} = \frac{74 + 76 + 77 + 78 + 80 + 81 + 82 + 84 + 85 + 85 + 85 + 87 + 88 + 89 + 90 + 90 + 91 + 93 + 96 + 98}{20}$$

$$\bar{x} = \frac{1709}{20} = 85,45$$

Pada R:

```
> # Mean data tunggal
> mean_ungrouped = mean(ungrouped_data)
> mean_ungrouped
[1] 85.45
```

2.1.2. Data Kelompok:

Rumus:

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

Di mana: $\sum f_i$: Frekuensi

x_i : Titik tengah.

Pada contoh:

$$\bar{x} = \frac{(76 \cdot 4 + 81 \cdot 3 + 86 \cdot 6 + 91 \cdot 5 + 96 \cdot 2)}{(4 + 3 + 6 + 5 + 2)}$$

$$\bar{x} = \frac{1710}{20} = 85,5$$

Pada R:

```
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)
> n_grouped
[1] 20

> # Mean data kelompok
> mean_grouped = sum(grouped_data$Frekuensi * grouped_data$Titik_tengah) / n_grouped
> mean_grouped
[1] 85.5
```

2.2. Median

Median adalah nilai tengah dari keseluruhan data yang telah diurutkan.

2.2.1. Data Tunggal:

$$Me = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{jika banyak data ganjil} \\ \frac{1}{2} \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right), & \text{jika banyak data genap} \end{cases}$$

Pada contoh:

Karena banyak data adalah 20, maka genap.

$$Me = \frac{1}{2} \left(x_{\left(\frac{20}{2}\right)} + x_{\left(\frac{20}{2}+1\right)} \right)$$

$$Me = \frac{1}{2} (x_{10} + x_{11})$$

74, 76, 77, 78, 80, 81, 82, 84, 85, 85, 87, 88, 89, 90, 90, 91, 93, 96, 98

$$Me = \frac{1}{2} (85 + 85) = 85$$

Pada R:

```
> # Median data tunggal
> median_ungrouped = median(ungrouped_data)
> median_ungrouped
[1] 85
```

2.2.2. Data Kelompok:

$$Median = TB + \left(\frac{\frac{n}{2} - fkum}{fm} \right) \cdot i$$

Di mana: TB : Tepi bawah kelas median

n : banyaknya data

$fkum$: frekuensi kumulatif sebelum kelas median

fm : frekuensi kelas median

i : panjang kelas interval

Identifikasi kelas median

Kelas median: kelas di mana $\frac{n}{2}$ (setengah total frekuensi) berada.

Pada contoh:

	Kelas	Frekuensi	Titik Tengah
1	75 – 78	4	76
2	79 – 83	3	81
3	84 – 88	6	86
4	89 – 93	5	91
5	94 – 98	2	96

Langkah-langkah

1. Tentukan kelas median.

$$\frac{n}{2} = \frac{20}{2} = 10$$

Kelas median ditunjukkan oleh data ke-10, di mana itu terletak di kelas ke-3 dengan kelas interval 84-88.

2. Hitung fm , $fkum$, i , dan TB .

$$fm = 6$$

$$fkum = 4 + 3 = 7$$

Kelas interval 84-88 terdiri dari 84, 85, 86, 87, 88.

$$i = 5$$

TB = Batas Bawah Kelas – 0,5

$$(84 - 88) \rightarrow TB = 84 - 0,5 = 83,5$$

3. Hitung median.

$$Median = 83,5 + \left(\frac{10 - 7}{6}\right) \cdot 5$$

$$Median = 83,5 + \left(\frac{3}{6}\right) \cdot 5$$

$$Median = 83,5 + \left(\frac{15}{6}\right) = 86$$

Pada R:

```
> # Menghitung banyak data kelompok  
> n_grouped = sum(grouped_data$Frekuensi)
```

```

> # Median data kelompok
> n_grouped
[1] 20
> cumulative_frequency = cumsum(grouped_data$Frekuensi)
> kelas_median <- which(cumulative_frequency >= n_grouped / 2)[1]
> TB_median <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_median])) - 0.5 # TB = Batas bawah - 0.5
> fkum <- if (kelas_median > 1) cumulative_frequency[kelas_median - 1] else 0 # Jika kelas_median 1, maka fkum 0
> f_median <- grouped_data$Frekuensi[kelas_median]
> i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_median])) + 0.5) - TB_median # i = TA - TB
>
> median_grouped <- TB_median + (((n_grouped / 2) - fkum) / f_median) * i
> median_grouped
[1] 86

```

2.3.Modus

Modus adalah nilai yang paling sering muncul. Pada data kelompok, modus menunjukkan nilai yang paling sering muncul dalam suatu kelompok data yang telah dikelompokkan ke dalam interval-interval kelas.

2.3.1. Data Tunggal:

Pada contoh:

74, 76, 77, 78, 80, 81, 82, 84, **85, 85, 85**, 87, 88, 89, 90, 90, 91, 93, 96, 98.

Modusnya adalah 85 dengan frekuensi 3.

Pada R:

```

> # Modus data tunggal
> modus_ungrouped <- as.numeric(names(sort(table(ungrouped_data), decreasing = TRUE)[1]))
> modus_ungrouped
[1] 85

```

2.3.2. Data Kelompok:

$$Modus = TB + \left(\frac{d1}{d1 + d2} \right) \cdot i$$

Di mana: *TB*: Tepi bawah kelas modus

$$d1 = fm - f_{\text{kelas sebelumnya}}$$

$$d2 = fm - f_{\text{kelas setelahnya}}$$

fm: frekuensi kelas modus

i: panjang kelas interval

Identifikasi kelas modus

Kelas modus: kelas dengan frekuensi terbesar

Pada contoh:

	Kelas	Frekuensi	Titik Tengah
1	75 – 78	4	76
2	79 – 83	3	81
3	84 – 88	6	86

4	89 – 93	5	91
5	94 – 98	2	96

Langkah-langkah

1. Tentukan kelas modus.

Kelas modus berada pada kelas interval 84-88 dengan frekuensi (fm) 6.

2. Hitung TB, d1, d2, i.

TB = Batas Bawah Kelas – 0,5

$$(84 - 88) \rightarrow TB = 84 - 0,5 = 83,5$$

Kelas interval 84-88 terdiri dari 84, 85, 86, 87, 88.

$$d1 = 6 - 3 = 3$$

$$d2 = 6 - 5 = 1$$

$$i = 5$$

3. Hitung modus.

$$Modus = 83,5 + \left(\frac{3}{3+1} \right) \cdot 5$$

$$Modus = 83,5 + \left(\frac{3}{4} \right) \cdot 5$$

$$Modus = 83,5 + 3,75 = 87,25$$

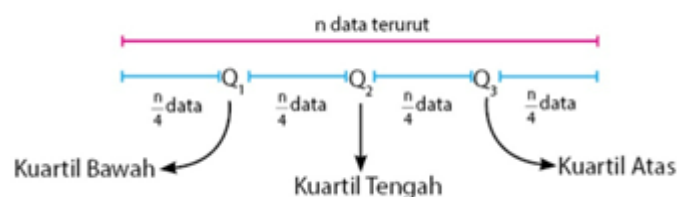
Pada R:

```
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)

> # Modus data kelompok
> n_grouped
[1] 20
> kelas_modus <- which.max(grouped_data$Frekuensi)
> TB_modus <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_modus])) - 0.5 # TB = Batas bawah - 0.5
> f_modus <- grouped_data$Frekuensi[kelas_modus]
> d1 <- if (kelas_modus > 1) f_modus - grouped_data$Frekuensi[kelas_modus - 1] else f_modus
> d2 <- if (kelas_modus < n_grouped) f_modus - grouped_data$Frekuensi[kelas_modus + 1] else f_modus
> i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_modus])) + 0.5) - TB_modus # i = TA - TB
>
> modus_grouped <- TB_modus + (d1 / (d1 + d2)) * i
> modus_grouped
[1] 87.25
```

2.4.Kuartil

Kuartil adalah suatu nilai yang bisa membagi kumpulan data menjadi empat bagian sama besar. Syarat untuk mendapatkan kuartil adalah data harus diurutkan dari data terkecil terlebih dahulu.



8

quantile() untuk menghitung kuartil, yang memiliki 9 metode berbeda (disebut "type 1" hingga "type 9"). Metode default di R adalah Type 7.

Jika ingin hasil R sesuai perhitungan manual, gunakan metode yang konsisten dengan interpolasi manual, misalnya type = 2 atau type = 6 dalam fungsi quantile().

```
> quartiles_ungrouped <- quantile(ungrouped_data, probs = c(0.25, 0.5, 0.75), type = 6)
> quartiles_ungrouped
 25%  50%  75%
80.25 85.00 90.00
```

2.4.2. Data Kelompok:

Rumus Kuartil Q_k (ke- k , $k = 1,2,3$)

$$Q_k = TB + \left(\frac{\frac{k}{4} \cdot n - fkum}{f} \right) \cdot i$$

Di mana: TB : Tepi bawah kelas kuartil

$fkum$: frekuensi kumulatif sebelum kelas kuartil

f : frekuensi kelas kuartil

i : panjang kelas interval

Identifikasi kelas kuartil

Kelas kuartil: kelas di mana $\frac{n}{4}, \frac{n}{2}, \frac{3n}{4}$ berada.

Pada contoh:

	Kelas	Frekuensi	Titik Tengah	Frekuensi Kumulatif
1	75 – 78	4	76	4
2	79 – 83	3	81	7
3	84 – 88	6	86	13
4	89 – 93	5	91	18
5	94 – 98	2	96	20

Langkah-langkah:

1. Tentukan kelas kuartil.

$$Q_1 \rightarrow \frac{n}{4} = \frac{20}{4} = 5$$

$$Q_2 \rightarrow \frac{n}{2} = \frac{20}{2} = 10$$

$$Q_3 \rightarrow \frac{3n}{4} = \frac{60}{4} = 15$$

Kelas kuartil Q_1 berada pada kelas 2, Q_2 berada pada kelas 3, Q_3 berada pada kelas 4.

2. Hitung Q_1 .

$$\frac{n}{4} = 5$$

$$(79 - 83) \rightarrow TB = 79 - 0,5 = 78,5$$

$$fkum = 4$$

$$f = 3$$

$$i = 5$$

$$Q_1 = 78,5 + \left(\frac{5-4}{3}\right) \cdot 5$$

$$Q_1 = 78,5 + \frac{5}{3} \approx 80,166$$

3. Hitung Q2.

$$\frac{n}{4} = 10$$

$$(84 - 88) \rightarrow TB = 84 - 0,5 = 83,5$$

$$fkum = 7$$

$$f = 6$$

$$i = 5$$

$$Q_2 = 83,5 + \left(\frac{10-7}{6}\right) \cdot 5$$

$$Q_2 = 83,5 + \frac{5}{2} = 86$$

4. Hitung Q3.

$$\frac{3n}{4} = 15$$

$$(89 - 93) \rightarrow TB = 89 - 0,5 = 88,5$$

$$fkum = 13$$

$$f = 5$$

$$i = 5$$

$$Q_3 = 88,5 + \left(\frac{15-13}{5}\right) \cdot 5$$

$$Q_3 = 88,5 + 2 = 90,5$$

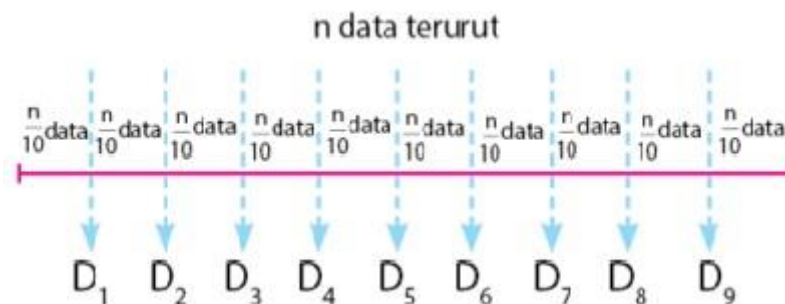
Pada R:

```
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)

> # Kuartil data kelompok
> quartiles_grouped <- numeric(3) # Placeholder untuk Q1, Q2, Q3
> for (k in 1:3) {
+   posisi_kuartil <- k / 4 * n_grouped
+   kelas_kuartil <- which(cumsum(grouped_data$Frekuensi) >= posisi_kuartil)[1]
+   TB_kuartil <- as.numeric(sub("-", ".", grouped_data$Kelas[kelas_kuartil])) - 0.5 # TB = Batas bawah - 0.5
+   fkum_kuartil <- if (kelas_kuartil > 1) cumsum(grouped_data$Frekuensi)[kelas_kuartil - 1] else 0
+   f_kuartil <- grouped_data$Frekuensi[kelas_kuartil]
+   i <- (as.numeric(sub("-", ".", grouped_data$Kelas[kelas_kuartil])) + 0.5) - TB_kuartil # i = TA - TB
+   quartiles_grouped[k] <- TB_kuartil + ((posisi_kuartil - fkum_kuartil) / f_kuartil) * i
+ }
> quartiles_grouped
[1] 80.16667 86.00000 90.50000
```

2.5.Desil

Desil adalah suatu nilai yang bisa membagi kumpulan data menjadi 10 bagian sama besar. Syarat untuk mendapatkan desil seperti kuartil, yaitu data harus diurutkan terlebih dahulu dari data terkecil.



2.5.1. Data Tunggal:

Rumus Desil D_k ($k=1, 2, 3, \dots, 9$)

$$D_k = x_{\left(\frac{k}{10} \cdot (n+1)\right)}$$

Pada contoh:

$$D_1 = x_{\left(\frac{1}{10} \cdot (20+1)\right)} = x_{\frac{21}{10}} = x_{2,1}$$

$$D_6 = x_{\left(\frac{6}{10} \cdot (20+1)\right)} = x_{\frac{126}{10}} = x_{12,6}$$

$$D_2 = x_{\left(\frac{2}{10} \cdot (20+1)\right)} = x_{\frac{42}{10}} = x_{4,2}$$

$$D_7 = x_{\left(\frac{7}{10} \cdot (20+1)\right)} = x_{\frac{147}{10}} = x_{14,7}$$

$$D_3 = x_{\left(\frac{3}{10} \cdot (20+1)\right)} = x_{\frac{63}{10}} = x_{6,3}$$

$$D_8 = x_{\left(\frac{8}{10} \cdot (20+1)\right)} = x_{\frac{168}{10}} = x_{16,8}$$

$$D_4 = x_{\left(\frac{4}{10} \cdot (20+1)\right)} = x_{\frac{84}{10}} = x_{8,4}$$

$$D_9 = x_{\left(\frac{9}{10} \cdot (20+1)\right)} = x_{\frac{189}{10}} = x_{18,9}$$

$$D_5 = x_{\left(\frac{5}{10} \cdot (20+1)\right)} = x_{\frac{105}{10}} = x_{10,5}$$

74, 76	77, 78	80, 81	82, 84	85, 85	85, 87	88, 89	90, 90	91, 93	96, 98
D1	D2	D3	D4	D5	D6	D7	D8	D9	

Jika posisi D_k bukan bilangan bulat, lakukan interpolasi linier.

$$D_k = x_y + 0, * \cdot (x_{y+1} - x_y)$$

$$D_1 = 76 + 0,1 \cdot (77 - 76) = 76 + 0,1 \cdot (1) = 76,1$$

$$D_2 = 78 + 0,2 \cdot (80 - 78) = 78 + 0,2 \cdot (2) = 78,4$$

$$D_3 = 81 + 0,3 \cdot (82 - 81) = 81 + 0,3 \cdot (1) = 81,3$$

$$D_4 = 84 + 0,4 \cdot (85 - 84) = 84 + 0,4 \cdot (1) = 84,4$$

$$D_5 = 85 + 0,5 \cdot (85 - 85) = 85$$

$$D_6 = 87 + 0,6 \cdot (88 - 87) = 87 + 0,6 \cdot (1) = 87,6$$

$$D_7 = 89 + 0,7 \cdot (90 - 89) = 89 + 0,7 \cdot (1) = 89,7$$

$$D_8 = 90 + 0,8 \cdot (91 - 90) = 90 + 0,8 \cdot (1) = 90,8$$

$$D_9 = 93 + 0,9 \cdot (96 - 93) = 93 + 0,9 \cdot (3) = 95,7$$

Pada R:

```
> # Desil data tunggal
> deciles_ungrouped <- quantile(ungrouped_data, probs = seq(0.1, 0.9, 0.1), type = 6)
> deciles_ungrouped
10% 20% 30% 40% 50% 60% 70% 80% 90%
76.1 78.4 81.3 84.4 85.0 87.6 89.7 90.8 95.7
```

2.5.2. Data Kelompok:

Rumus Desil D_k ($k=1,2,3, \dots, 9$)

$$D_k = TB + \left(\frac{\frac{k}{10} \cdot n - f_{kum}}{f} \right) \cdot i$$

Di mana: TB : Tepi bawah kelas desil

f_{kum} : frekuensi kumulatif sebelum kelas desil

f : frekuensi kelas desil

i : panjang kelas interval

Identifikasi kelas desil

Kelas desil: kelas di mana $\frac{k}{10} \cdot n$ berada.

Pada contoh:

	Kelas	Frekuensi	Titik Tengah	Frekuensi Kumulatif
1	75 – 78	4	76	4
2	79 – 83	3	81	7
3	84 – 88	6	86	13
4	89 – 93	5	91	18
5	94 – 98	2	96	20

Tentukan D_3 , D_5 dan D_8 ?

Langkah-langkah:

1. Tentukan kelas desil.

$$D_3 \rightarrow \frac{3}{10} \cdot 20 = 6$$

$$D_5 \rightarrow \frac{5}{10} \cdot 20 = 10$$

$$D_8 \rightarrow \frac{8}{10} \cdot 20 = 16$$

Kelas desil D_3 berada pada kelas 2, D_5 berada pada kelas 3, D_8 berada pada kelas 4.

2. Hitung D_3 .

$$\frac{3}{10} \cdot 20 = 6$$

$$(79 - 83) \rightarrow TB = 79 - 0,5 = 78,5$$

$$fkum = 4$$

$$f = 3$$

$$i = 5$$

$$D_3 = 78,5 + \left(\frac{6 - 4}{3}\right) \cdot 5$$

$$D_3 = 78,5 + \frac{10}{3} \approx 81,83$$

3. Hitung D5.

$$\frac{5}{10} \cdot 20 = 10$$

$$(84 - 88) \rightarrow TB = 84 - 0,5 = 83,5$$

$$fkum = 7$$

$$f = 6$$

$$i = 5$$

$$D_5 = 83,5 + \left(\frac{10 - 7}{6}\right) \cdot 5$$

$$D_5 = 83,5 + \frac{5}{2} = 86$$

4. Hitung D8.

$$\frac{8}{10} \cdot 20 = 16$$

$$(89 - 93) \rightarrow TB = 89 - 0,5 = 88,5$$

$$fkum = 13$$

$$f = 5$$

$$i = 5$$

$$D_8 = 88,5 + \left(\frac{16 - 13}{5}\right) \cdot 5$$

$$D_8 = 88,5 + 3 = 91,5$$

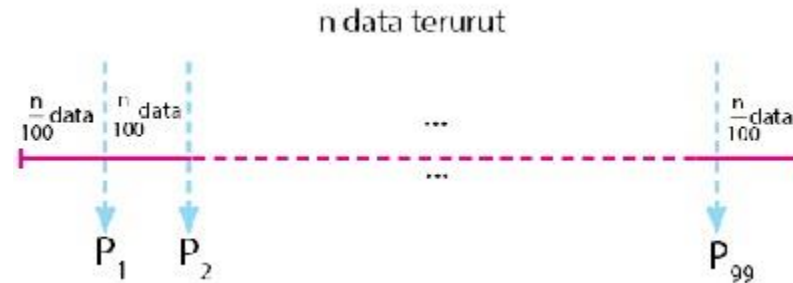
Pada R:

```
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)

> # Desil data kelompok
> deciles_grouped <- numeric(9) # Placeholder untuk D1 hingga D9
> for (k in 1:9) {
+   posisi_desil <- k / 10 * n_grouped
+   kelas_desil <- which(cumsum(grouped_data$Frekuensi) >= posisi_desil)[1]
+   TB_desil <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_desil])) - 0.5 # TB = Batas bawah - 0.5
+   fkum_desil <- if (kelas_desil > 1) cumsum(grouped_data$Frekuensi)[kelas_desil - 1] else 0
+   f_desil <- grouped_data$Frekuensi[kelas_desil]
+   i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_desil])) + 0.5) - TB_desil # i = TA - TB
+   deciles_grouped[k] <- TB_desil + ((posisi_desil - fkum_desil) / f_desil) * i
+ }
> deciles_grouped
[1] 76.00000 78.50000 81.83333 84.33333 86.00000 87.66667 89.50000 91.50000 93.50000
```

2.6.Persentil

Persentil adalah suatu nilai yang bisa membagi kumpulan data menjadi 100 bagian sama besar. Syarat untuk mendapatkan persentil seperti kuartil dan desil, yaitu data harus diurutkan terlebih dahulu dari data terkecil.



2.6.1. Data Tunggal:

Rumus Persentil P_k (ke- k , $k = 1, 2, 3, \dots, 99$)

$$D_k = x_{\left(\frac{k}{100} \cdot (n+1)\right)}$$

Pada contoh: Tentukan P20, P66, dan P69?

$$P_{20} = x_{\left(\frac{20}{100} \cdot (20+1)\right)} = x_{\frac{420}{100}} = x_{4,2}$$

$$P_{20} = x_{\left(\frac{20}{100} \cdot (20+1)\right)} = x_{\frac{420}{100}} = x_{4,2}$$

$$P_{66} = x_{\left(\frac{66}{100}(20+1)\right)} = x_{\frac{1386}{100}} = x_{13,86}$$

$$P_{69} = x_{\left(\frac{69}{100} \cdot (20+1)\right)} = x_{\frac{1449}{100}} = x_{14,49}$$

74, 76, 77, 78 | 80, 81, 82, 84, 85, 85, 85, 87, 88 | 89 | 90, 90, 91, 93, 96, 98
P20 P66 P69

Jika posisi P_k bukan bilangan bulat, lakukan interpolasi linier.

$$P_k = x_y + 0,* \cdot (x_{y+1} - x_y)$$

$$P_{20} = 78 + 0,2 \cdot (80 - 78) = 78 + 0,2 \cdot (2) = 78,4$$

$$P_{66} = 88 + 0,86 \cdot (89 - 88) = 88 + 0,86 \cdot (1) = 88,86$$

$$P_{69} = 89 + 0,49 \cdot (90 - 89) = 89 + 0,49 \cdot (1) = 89,49$$

Pada R:

```
> # Persentil data tunggal
> percentiles_ungrouped <- quantile(ungrouped_data, probs = seq(0.01, 0.99, 0.01), type = 6)
> percentiles_ungrouped
```

1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	11%	12%	13%	14%	15%	16%	17%	18%	19%	20%
74.00	74.00	74.00	74.00	74.10	74.52	74.94	75.36	75.78	76.10	76.31	76.52	76.73	76.94	77.15	77.36	77.57	77.78	77.99	78.40
21%	22%	23%	24%	25%	26%	27%	28%	29%	30%	31%	32%	33%	34%	35%	36%	37%	38%	39%	40%
78.82	79.24	79.66	80.04	80.25	80.46	80.67	80.88	81.09	81.30	81.51	81.72	81.93	82.28	82.70	83.12	83.54	83.96	84.19	84.40
41%	42%	43%	44%	45%	46%	47%	48%	49%	50%	51%	52%	53%	54%	55%	56%	57%	58%	59%	60%
84.61	84.82	85.00	85.00	85.00	85.00	85.00	85.00	85.00	85.00	85.00	85.00	85.26	85.68	86.10	86.52	86.94	87.18	87.39	87.60
61%	62%	63%	64%	65%	66%	67%	68%	69%	70%	71%	72%	73%	74%	75%	76%	77%	78%	79%	80%
87.81	88.02	88.23	88.44	88.65	88.86	89.07	89.28	89.49	89.70	89.91	90.00	90.00	90.00	90.00	90.00	90.17	90.38	90.59	90.80
81%	82%	83%	84%	85%	86%	87%	88%	89%	90%	91%	92%	93%	94%	95%	96%	97%	98%	99%	
91.02	91.44	91.86	92.28	92.70	93.18	93.81	94.44	95.07	95.70	96.22	96.64	97.06	97.48	97.90	98.00	98.00	98.00	98.00	

2.6.2. Data Kelompok:

Rumus Persentil P_k (ke- k , $k = 1, 2, 3, \dots, 99$)

$$P_k = TB + \left(\frac{\frac{k}{100} \cdot n - f_{kum}}{f} \right) \cdot i$$

Di mana: TB : Tepi bawah kelas persentil

f_{kum} : frekuensi kumulatif sebelum kelas persentil

f : frekuensi kelas persentil

i : panjang kelas interval

Identifikasi kelas persentil

Kelas persentil: kelas di mana $\frac{k}{100} \cdot n$ berada.

Pada contoh:

	Kelas	Frekuensi	Titik Tengah	Frekuensi Kumulatif
1	75 – 78	4	76	4
2	79 – 83	3	81	7
3	84 – 88	6	86	13
4	89 – 93	5	91	18
5	94 – 98	2	96	20

Tentukan P_{20} , P_{66} , dan P_{94} ?

Langkah-langkah:

1. Tentukan kelas persentil.

$$P_{20} \rightarrow \frac{20}{100} \cdot 20 = 4$$

$$P_{66} \rightarrow \frac{66}{100} \cdot 20 = 13,2$$

$$P_{94} \rightarrow \frac{94}{100} \cdot 20 = 18,8$$

Kelas persentil P_{20} berada pada kelas 1, P_{66} berada pada kelas 4, P_{94} berada pada kelas 5.

2. Hitung P_{20} .

$$\frac{20}{100} \cdot 20 = 4$$

$$(74 - 78) \rightarrow TB = 74 - 0,5 = 73,5$$

$$fkum = 0$$

$$f = 4$$

$$i = 5$$

$$P_{20} = 73,5 + \left(\frac{4 - 0}{4}\right) \cdot 5$$

$$P_{20} = 73,5 + 5 = 78,5$$

3. Hitung P66.

$$\frac{66}{100} \cdot 20 = 13,2$$

$$(89 - 93) \rightarrow TB = 89 - 0,5 = 88,5$$

$$fkum = 13$$

$$f = 5$$

$$i = 5$$

$$P_{66} = 88,5 + \left(\frac{13,2 - 13}{5}\right) \cdot 5$$

$$P_{66} = 88,5 + 0,2 = 88,7$$

4. Hitung P94.

$$\frac{94}{100} \cdot 20 = 18,8$$

$$(94 - 98) \rightarrow TB = 94 - 0,5 = 93,5$$

$$fkum = 18$$

$$f = 2$$

$$i = 5$$

$$P_{94} = 93,5 + \left(\frac{18,8 - 18}{2}\right) \cdot 5$$

$$P_{94} = 93,5 + \left(\frac{0,8}{2}\right) \cdot 5 = 95,5$$

Pada R:

```
> # Menghitung banyak data kelompok
> n_grouped = sum(grouped_data$Frekuensi)

> # Persentil data kelompok
> percentiles_grouped <- numeric(99) # Placeholder untuk P1 hingga P99
> for (k in 1:99) {
+   posisi_persentil <- k / 100 * n_grouped
+   kelas_persentil <- which(cumsum(grouped_data$Frekuensi) >= posisi_persentil)[1]
+   TB_persentil <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_persentil])) - 0.5 # TB = Batas bawah - 0.5
+   fkum_persentil <- if (kelas_persentil > 1) cumsum(grouped_data$Frekuensi)[kelas_persentil - 1] else 0
+   f_persentil <- grouped_data$Frekuensi[kelas_persentil]
+   i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_persentil])) + 0.5) - TB_persentil # i = TA - TB
+   percentiles_grouped[k] <- TB_persentil + ((posisi_persentil - fkum_persentil) / f_persentil) * i
+ }
> print(percentiles_grouped)
[1] 73.75000 74.00000 74.25000 74.50000 74.75000 75.00000 75.25000 75.50000 75.75000 76.00000 76.25000
[12] 76.50000 76.75000 77.00000 77.25000 77.50000 77.75000 78.00000 78.25000 78.50000 78.83333 79.16667
[23] 79.50000 79.83333 80.16667 80.50000 80.83333 81.16667 81.50000 81.83333 82.16667 82.50000 82.83333
[34] 83.16667 83.50000 83.83333 84.16667 84.50000 84.83333 85.16667 85.50000 85.83333 86.16667 86.50000
[45] 86.83333 87.16667 87.50000 87.83333 88.16667 88.50000 88.83333 89.16667 89.50000 89.83333 90.16667
[56] 90.50000 90.83333 91.16667 91.50000 91.83333 92.16667 92.50000 92.83333 93.16667 93.50000 93.83333
[67] 94.16667 94.50000 94.83333 95.16667 95.50000 95.83333 96.16667 96.50000 96.83333 97.16667 97.50000
[78] 97.83333 98.16667 98.50000 98.83333 99.16667 99.50000 99.83333 100.00000 100.33333 100.66667 101.00000
[89] 101.33333 101.66667 102.00000 102.33333 102.66667 103.00000 103.33333 103.66667 104.00000 104.33333 104.66667
```


3. Tambahan

```
# Mendeklarasikan data tunggal
ungrouped_data = c(74, 76, 77, 78, 80, 81, 82, 84, 85, 85, 85, 87, 88, 89, 90, 90, 91, 93, 96, 98)
ungrouped_data

# Mendeklarasikan data kelompok
Kelas = c("74-78", "79-83", "84-88", "89-93", "94-98")
Frekuensi = c(4, 3, 6, 5, 2)
Titik_tengah = c(76, 81, 86, 91, 96)
grouped_data <- data.frame(Kelas, Frekuensi, Titik_tengah)
print(grouped_data)

# Menghitung banyak data tunggal
n_ungrouped = length(ungrouped_data)
n_ungrouped

# Menghitung banyak data kelompok
n_grouped = sum(grouped_data$Frekuensi)
n_grouped

# Mean data tunggal
mean_ungrouped = mean(ungrouped_data)
mean_ungrouped

# Mean data kelompok
mean_grouped = sum(grouped_data$Frekuensi * grouped_data$Titik_tengah) / n_grouped
mean_grouped

# Median data tunggal
median_ungrouped = median(ungrouped_data)
median_ungrouped

# Median data kelompok
n_grouped
cumulative_frequency = cumsum(grouped_data$Frekuensi)
kelas_median <- which(cumulative_frequency >= n_grouped / 2)[1]
TB_median <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_median])) - 0.5 # TB = Batas bawah - 0.5
fkum <- if (kelas_median > 1) cumulative_frequency[kelas_median - 1] else 0 # Jika kelas_median 1, maka fkum 0
f_median <- grouped_data$Frekuensi[kelas_median]
i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_median])) + 0.5) - TB_median # i = TA - TB

median_grouped <- TB_median + (((n_grouped / 2) - fkum) / f_median) * i
median_grouped
TB_median
i
fkum
f_median

# Modus data tunggal
modus_ungrouped <- as.numeric(names(sort(table(ungrouped_data), decreasing = TRUE)[1]))
modus_ungrouped

# Modus data kelompok
n_grouped
kelas_modus <- which.max(grouped_data$Frekuensi)
TB_modus <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_modus])) - 0.5 # TB = Batas bawah - 0.5
f_modus <- grouped_data$Frekuensi[kelas_modus]
d1 <- if (kelas_modus > 1) f_modus - grouped_data$Frekuensi[kelas_modus - 1] else f_modus
d2 <- if (kelas_modus < n_grouped) f_modus - grouped_data$Frekuensi[kelas_modus + 1] else f_modus
i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_modus])) + 0.5) - TB_modus # i = TA - TB

modus_grouped <- TB_modus + (d1 / (d1 + d2)) * i
modus_grouped
```

```

# Kuartil data tunggal
quartiles_ungrouped <- quantile(ungrouped_data, probs = c(0.25, 0.5, 0.75), type = 6)
quartiles_ungrouped

# Kuartil data kelompok
quartiles_grouped <- numeric(3) # Placeholder untuk Q1, Q2, Q3
for (k in 1:3) {
  posisi_kuartil <- k / 4 * n_grouped
  kelas_kuartil <- which(cumsum(grouped_data$Frekuensi) >= posisi_kuartil)[1]
  TB_kuartil <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_kuartil])) - 0.5 # TB = Batas bawah - 0.5
  fkum_kuartil <- if (kelas_kuartil > 1) cumsum(grouped_data$Frekuensi)[kelas_kuartil - 1] else 0
  f_kuartil <- grouped_data$Frekuensi[kelas_kuartil]
  i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_kuartil])) + 0.5) - TB_kuartil # i = TA - TB

  quartiles_grouped[k] <- TB_kuartil + ((posisi_kuartil - fkum_kuartil) / f_kuartil) * i
}
quartiles_grouped

# Desil data tunggal
deciles_ungrouped <- quantile(ungrouped_data, probs = seq(0.1, 0.9, 0.1), type = 6)
deciles_ungrouped

# Desil data kelompok
deciles_grouped <- numeric(9) # Placeholder untuk D1 hingga D9
for (k in 1:9) {
  posisi_desil <- k / 10 * n_grouped
  kelas_desil <- which(cumsum(grouped_data$Frekuensi) >= posisi_desil)[1]
  TB_desil <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_desil])) - 0.5 # TB = Batas bawah - 0.5
  fkum_desil <- if (kelas_desil > 1) cumsum(grouped_data$Frekuensi)[kelas_desil - 1] else 0
  f_desil <- grouped_data$Frekuensi[kelas_desil]
  i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_desil])) + 0.5) - TB_desil # i = TA - TB

  deciles_grouped[k] <- TB_desil + ((posisi_desil - fkum_desil) / f_desil) * i
}
deciles_grouped

# Persentil data tunggal
percentiles_ungrouped <- quantile(ungrouped_data, probs = seq(0.01, 0.99, 0.01), type = 6)
percentiles_ungrouped

# Persentil data kelompok
percentiles_grouped <- numeric(99) # Placeholder untuk P1 hingga P99
for (k in 1:99) {
  posisi_persentil <- k / 100 * n_grouped
  kelas_persentil <- which(cumsum(grouped_data$Frekuensi) >= posisi_persentil)[1]
  TB_persentil <- as.numeric(sub("-", "", grouped_data$Kelas[kelas_persentil])) - 0.5 # TB = Batas bawah - 0.5
  fkum_persentil <- if (kelas_persentil > 1) cumsum(grouped_data$Frekuensi)[kelas_persentil - 1] else 0
  f_persentil <- grouped_data$Frekuensi[kelas_persentil]
  i <- (as.numeric(sub("-", "", grouped_data$Kelas[kelas_persentil])) + 0.5) - TB_persentil # i = TA - TB

  percentiles_grouped[k] <- TB_persentil + ((posisi_persentil - fkum_persentil) / f_persentil) * i
}
print(percentiles_grouped)

```