UKURAN PEMUSATAN DATA

4

OBJEKTIF :

1. Mahasiswa dapat mengetahui penyelesaian rata-rata
2. Mahasiswa dapat mengetahui penyelesaian median
3. Mahasiswa dapat mengetahui penyelesaian modus
4. Mahasiswa dapat mengetahui penyelesaian kuartil, desil, persentil, skewness dan kurtosis
5. Mahasiswa dapat mengetahui penyelesaian rata-rata tertimbang
6. Mahasiswa dapat mengetahui penyelesaian rata-rata geometrik

---

Bab 4 Ukuran Pemusatan Data

Ada tiga ukuran pusat data yang banyak digunakan, yaitu: rata-rata hitung (selanjutnya disebut rata-rata), median, dan modus. Sebagai tambahan, akan dijelaskan pula mengenai kuartil, desil, persentil, rata-rata tertimbang, dan rata-rata geometrik.

4.1 Rata-Rata Hitung

Rata-rata hitung, atau lebih dikenal dengan rata-rata, merupakan ukuran pusat data yang paling sering digunakan, karena mudah dimengerti oleh siapa saja dan penghitungannya pun mudah. Rata-rata yang dihitung dari data sampel atau sebagai statistik sampel disimbulkan dengan X (baca: X -bar) dan jika dihitung dari data populasi atau sebagai parameter populasi disimbulkan dengan huruf Yunani f.l, (baca: myu x).

Rata-rata dari Data yang Belum Dikelompokkan

Rata-rata dihitung dengan menjumlahkan seluruh angka data yang selanjutnya dibagi dengan banyaknya (jumlah) data. Jumlah data, untuk data sampel disebut sebagai ukuran sampel yang disimbolkan dengan n dan untuk data populasi disebut sebagai ukuran populasi yang disimbulkan dengan N.

Jika X_1, X_2, X_3, ... , Xn adalah angka-angka data yang banyaknya (jumlahnya) n, maka rata-ratanya dihitung:

Rata-rata hitung adalah jumlah seluruh angka data dibagi dengan banyaknya (jumlah) data.

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Atau

$$\bar{X} = \frac{\sum_{i=1}^{n} Xi}{n}$$

$\overline{X}$: Rata-rata sampel

X_i : Data ke-i variabel acak X; i = 1,2,.....,n

n : Ukuran sampel (banyaknya anggota sampel)

Untuk Populasi :

$$\mu = \frac{\sum_{i=1}^{n} Xi}{N}$$

μ : Rata-rata populasi (dibaca:myu)

X_i : Data ke-i variabel acak X;i = 1,2, ...,N

N : Ukuran populasi (banyaknya anggota populasi)

Contoh :

Misalkan data tinggi badan 10 orang mahasiswa (dalam cm): 162, 161,157, 154, 164, 170, 162, 165, 162, 161.

N = 10 dan $\sum_{i=1} Xi = 162 + 161 + \cdots \ldots\ldots + 161 = 1618$

sehingga: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1618}{10} = 161.8$

Rata-rata dari Data yang Telah Dikelompokkan

Menghitung rata-rata memang lebih menguntungkan jika dihitung dari data yang belum dikelompokkan, karena hasil hitungannya lebih mencerminkan fakta yang sebenarnya. Apakah rata-rata dari data yang telah dikelompokkan tidak mencerminkan data yang sebenarnya? Dalam kehidupan sehari-hari, data yang dibutuhkan seringkali sudah disajikan dalam bentuk distribusi frekuensi, seperti yang banyak disajikan dalam berbagai terbitan maupun laporan-laporan. Sehingga, perhitungan rata-rata dari data yang telah dikelompokkan harus dilakukan walaupun hasilnya tidak mencerminkan fakta yang sebenarnya. Namun, paling tidak mendekati fakta yang sebenarnya.

Pada sub bab 4.1 telah dijelaskan bahwa rata-rata dihitung dengan melibatkan seluruh data observasi, baik dari sampel maupun dari populasi. Untuk data observasi yang telah disajikan dalam bentuk distribusi frekuensi atau yang telah dikelompokkan, sifat keaslian data observasi telah hilang. Dengan demikian, untuk keperluan penghitungan rata-rata diperlukan angka-angka data yang dapat digunakan untuk mengestimasi atau menaksir data observasi yang asli. Dalam hal ini, titik-titik tengah dapat dijadikan sebagai penaksir data asli yang tersebar di masing-masing kelasnya. Ada dua cara yang dapat digunakan untuk menghitung rata-rata data yang telah dikelompokkan, yaitu **metode defisional** dan **metode pengkodean.**

a. **Metode defisional**

Untuk menghitung rata-rata, titik-titik tengah masing-masing kelas, sebagai penaksir data asli, dikali dengan frekuensi masing-masing kelas. Hasil perkalian pada masing-masing kelas tersebut selanjutnya dijumlah dan kemudian hasil penjumlahan tersebut dibagi dengan jumlah data atau jumlah frekuensi seluruh kelas. Metode defisional dapat dirumuskan sebagai berikut:

$$\overline{X} = \frac{Xifi}{n}$$

$\overline{X}$: Rata-rata sampel

X_i : Titik tengah kelas ke-i

f_i : Frekuensi kelas ke-i

n : Ukuran sampel (jumlah frekuensi data sampel)

$$\mu_X = \frac{Xifi}{N}$$

μ_X :Rata-rata populasi

X_i : Titik tengah kelas ke-i

f_i : Frekuensi kelas ke-i

N : Ukuran populasi (jumlah frekuensi data populasi)

Contoh

Selama tahun 1993, PT Asuransi Jiwa Jagat Raya telah berhasil menarik nasabah baru sebanyak 60 orang yang usianya dapat didistribusikan sebagai berikut:

Tabel 4.1 Distribusi Usia 60 Nasabah Baru PT Asuransi Jagat Raya

Usia	Frekuensi
25 - 29	8
30 – 34	14
35 – 39	10
40 – 44	18
45 – 49	7
50 - 54	3
Jumlah	60

Rata-rata usia para nasabah baru tersebut dapat dihitung sebagai berikut:

Tabel 4.2

Perhitungan Rata-rata dengan Menggunakan Metode "Defisional"

Titik Tengah X_i	Frekuensi F_i	$X_i f_i$
27	8	216
32	14	448
37	10	270
42	18	756
47	7	329
52	3	156
Jumlah	60	2.275

$\overline{X}$ = 2275 / 60

$\overline{X}$ = 37,92 atau 37 tahun 11 bulan

b. **Metode Pengkodean**

Seringkali data yang akan dihitung rata-ratanya berbentuk angka-angka yang besar seperti nilai penjualan, pembelian, piutang, dan lain sebagainya. Jika angka-angka yang dihitung dalam satuan yang besar, maka penghitungan rata-rata dengan penggunaan metode pengkodean akan sedikit lebih menyulitkan.

Pada bab 3 telah dijelaskan bahwa interval kelas sebuah distribusi frekuensi, secara umum senantiasa sama. Hanya dalam keadaan tertentu, interval kelas dimungkinkan tidak sama. Interval kelas yang sama ini, salah satunya dapat dilihat beda antar titik tengah senantiasa sama. Angka-angka berikut menunjukkan titik tengah yang dikutip dari tabel 4.2.

Titik Tengah :	27		32		37		42		47		52
Interval Kelas :		5		5		5		5		5	

Dengan interval kelas yang sama ini, sebenarnya, angka-angka titik tengah dapat diubah menjadi suatu skala dengan interval yang sama. Skala titik tengah ini lebih sering disebut sebagai kode titik tengah.

Langkah pertama dalam memberi kode titik tengah adalah menetapkan kelas yang nantinya diberi kode atau skala nol. Dalam menentukan kelas yang berkode nol ini sebenarnya tidak ada pedoman yang baku. Akan tetapi, sebaiknya kelas yang akan diberi kode nol adalah kelas yang berfrekuensi tertinggi. Langkah berikutnya adalah menetapkan kode-kode untuk kelas-kelas yang lain dengan mengurutkan mulai dari kelas berkode nol dengan interval yang sama. Interval kelas ini umumnya adalah satu. Dari tabel 4.2 di atas, kelas yang akan diberi kode nol adalah kelas ke-4. Dengan demikian, titik tengah, frekuensi dan kodenya adalah sebagai berikut:

Titik tengah:	27	32	37	42	47	52
Frekuensi:	8	14	10	18	7	3
Kode:	-3	-2	-1	0	1	2

Dalam literatur-literatur statistika, kode tersebut sering disimbulkan dengan huruf U. Selanjutnya, menghitung rata-rata. dengan menggunakan metode pengkodean dapat dirumuskan sebagai berikut:

$$\overline{X} = X_a + i\,\frac{\text{Uifi}}{n}$$

$\overline{X}$: Rata-rata sampel (μ_x jika populasi)

X_a : Titik tengah pada kelas yang berkode nol

i : Interval kelas

U_i : Kode titik tengah pada kelas ke-i

f_i : Frekuensi kelas ke-i

n : Ukuran sampel (N jika populasi)

Dari contoh berikut dapat dibandingkan tingkat kesulitan dalam menghitung rata-rata dengan menggunakan kedua metode di atas.

Contoh :

Lihat contoh di atas nilai kontrak asuransi keenampuluh nasabah baru tersebut dapat didistribusikan sebagai berikut:

Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya

Nilai Kontrak	Frekuensi
< 10.000.000,00	0
< 20.000.000,00	5
< 30.000.000,00	17
< 40.000.000,00	31
< 50.000.000,00	46
< 60.000.000,00	54
< 70.000.000,00	60

Untuk bisa menghitung rata-rata nilai kontrak dengan menggunakan metode defisional, bentuk penyajian di atas bentuk distribusi frekuensi kumulatif tipe "kurang dari" harus diubah menjadi bentuk distribusi frekuensi yang biasa. Hasilnya adalah:

Distribusi Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya

Nilai Kontrak	Frekuensi
Rp10.000.000,00 - < 20.000.000,00	0
Rp 20.000.000,00 - < 30.000.000,00	5
Rp 30.000.000,00 - <40.000.000,00	17
Rp 40.000.000,00 - < 50.000.000,00	31
Rp 50.000.000,00 - < 60.000.000,00	46
Rp 60.000.000,00 - < 70.000.000,00	54
	60

Proses pengerjaan berikutnya adalah:

Perhitungan Rata-rata Nilai Kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya dengan Metode "Defisional"

X_i	f_i	$X_i f_i$
Rp 15.000.000,00	5	Rp 75.000.000,00
Rp 25.000.000,00	12	Rp 300.000.000,00
Rp 35.000.000,00	14	Rp 490.000.000,00
Rp 45.000.000,00	15	Rp 675.000.000,00
Rp 55.000.000,00	8	Rp 440.000.000,00
Rp 65.000.000,00	6	Rp 390.000.000,00
Jumlah	60	Rp 2.370.000,00

$$\bar{X} = \frac{2.370.000.000,00}{60}$$

$$= \text{Rp } 39.500.000,00$$

Dengan menggunakan metode "pengkodean" penghitungannya disajikan pada tabel di bawah.

Perhitungan Rata – Rata Nilai kontrak Asuransi 60 Nasabah Baru PT Asuransi Jagat Raya dengan Metode "Pengkodean"

U_i	F_i	U_iF_i
-3	5	-15
-2	12	-24
-1	14	-14
0	15	0
1	8	8
2	6	12
	60	-33

$\overline{X}$ = 45.000.000 + 10.000.000 x $\frac{-33}{60}$

$\overline{X}$ = Rp 39.500.000,00

Dapat dibandingkan bahwa perhitungan rata-rata dengan metode defisional ternyata memerlukan waktu lebih banyak, khususnya dalam proses perkalian, daripada dengan menggunakan metode pengkodean. Akan tetapi metode pengkodean hanya dapat digunakan untuk distribusi frekuensi dengan interval kelas yang sama. Sedangkan untuk distribusi frekuensi dengan kelas yang tidak sama, metode defesional lah yang dapat digunakan.

Bagaimana proses perhitungan rata-rata untuk distribusi frekuensi dengan interval kelas yang tidak sama? Proses penghitungan rata-rata untuk distribusi frekuensi dengan interval kelas yang tidak sama tidak memiliki perbedaan dengan penghitungan rata-rata dari distrbusi frekuensi yang memiliki interval kelas yang sama.

Keunggulan rata-rata:

1. Lebih dikenal, sehingga penggunaannya pun lebih mudah.
2. Dapat digunakan pada data kuantitatif dan hanya memiliki satu rata-rata.
3. Karena kumpulan data hanya memiliki satu rata-rata, maka ukuran pusat data ini dapat digunakan dengan baik dalam prosedur statistika, seperti perbandingan dua atau lebih kumpulan data.

Kelemahan rata-rata:

1. Sangat peka terhadap data ekstrim.
2. Tidak dapat digunakan untuk menentukan ukuran pusat data kualitatif.
3. Untuk data berkelompok, hasil perhitungan tidak mencerminkan rata-rata sesungguhnya.
4. Untuk data berkelompok dengan kelas terbuka, rata-ratanya tidak dapat dihitung.

4.2 Median

Berbeda dengan rata-rata, penghitungan median tidak dilaksanakan dengan melibatkan seluruh angka data, namun lebih menekankan pada posisi atau letak data.

Median adalah ukuran pusat data yang nilainya terletak di tengah-tengah rangkaian data yang terurut.

Terletak di tengah-tengah artinya bahwa letak median tersebut membagi deretan data menjadi dua bagian yang sama. Jika $X_{(1)}$, $X_{(2)}$, $X_{(3)}$,$X_{(4)}$, $X_{(5)}$, $X_{(6)}$, $X_{(7)}$ adalah variabel-variabel acak sekelompok data kuantitatif yang terurut, maka mediannya adalah $X_{(4)}$ ini membagi ketujuh data tersebut menjadi dua bagian yang sama; jumlah angka data sebelum median sama dengan jumlah angka data sesudah median.

Median untuk Data Tak-berkelompok

Median adalah ukuran pusat data yang nilainya terletak di tengah-tengah kumpulan data yang terurut. Untuk menentukan median, mula-mula data harus disusun berupa *Array*, yaitu data yang diurutkan dari yang terkecil sampai dengan yang terbesar:

$$X_{(1)}, X_{(2)}, X_{(3)}, \ldots . X_{(n)}$$

$X_{(1)}$adalah data terkecil (= X_{min}), sedangkan $X_{n)}$ adalah data terbesar (= X_{max}).

Posisi median adalah :

$$\boxed{P_{med} = \frac{n+1}{2}}$$

Sehingga median adalah :

$$\text{Med} = X_{\left(\frac{n+1}{2}\right)} \qquad (\,n \text{ ganjil})$$

$$\text{Med} = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} \qquad (\,n \text{ genap})$$

Contoh :

Lihat kembali data tinggi badan 10 orang mahasiswa pada contoh:

162, 161, 157, 154, 164, 170, 162, 165, 162, 161. Data diurutkan dalam bentuk *Array* (dibaca: er-rei) sebagai berikut:

$X_{(1)}$= 154	$X_{(6)}$= 162
$X_{(2)}$= 157	$X_{(7)}$= 162
$X_{(3)}$= 161	$X_{(8)}$= 164
$X_{(4)}$= 161	$X_{(9)}$= 165
$X_{(5)}$= 162	$X_{(10)}$= 170

n = 10, sehingga posisi median adalah P_{med}= (10 +1)/ 2 = 5.5.

Karena n genap, maka median adalah:

$$\text{Med} = \frac{X_{(5)}+X_{(6)}}{2} = \frac{162+162}{2} = 162$$

Seandainya n = 9 dan $X_{(10)}$= 170 tidak ada, maka posisi median adalah P_{med}= (9 +1) / 2 = 5 dan median adalah *Med* = $X_{(5)}$= 162.

Tabel 4.3

Omzet Penjualan 7 Supermarket "Mataram Raya"

selama Bulan Desember 1993

Supermarket	Omset
"Mataram Raya 1"	Rp 65.000.000
"Mataram Raya 2"	Rp 80.000.000
"Mataram Raya 3"	Rp 85.000.000
"Mataram Raya 4"	Rp 90.000.000
"Mataram Raya 5"	Rp 95.000.000
"Mataram Raya 6"	Rp 115.000.000
"Mataram Raya 7"	Rp 170.000.000

⟶ **Median**

Untuk data ganjil, letak median dapat ditentukan dengan mudah. Berbeda dengan jumlah data genap, maka penentuan letak median tidak dapat ditetapkan begitu saja. Jika jumlah datanya 10, maka letak mediannya adalah data ke 5,5 yang dihitung dengan (10 + 1)/2.

Median untuk Data Berkelompok

Langkah pertama dalam menetapkan median dari data yang telah dikelompokkan adalah menentukan letak sebuah titik yang nilainya akan menjadi median. Titik ini, seperti pada uraian sebelumnya, membagi deretan angka data yang terurut menjadi dua bagian yang sama banyak. Jika pada data yang belum diurutkan digunakan perumusan (n + 1)/2, maka untuk data yang telah dikelompokkan, banyak penulis menggunakan perumusan yang lebih sederbana yaitu n/2. Akan tetapi, dengan menggunakan perumusan sebelumnya pun bukanlah suatu kesalahan.

Setelah diketahui posisi titik tersebut, langkah berikutnya adalah menentukan kelas yang didalamnya terdapat titik tersebut.

$$\text{Med} = B_{med} + \left[\frac{(n/2) - fk_{med}}{f_{med}}\right] i$$

Med : Median

B_{med} : Tepi batas kelas bawah pada kelas median (*Lower Class Boundary*)

i : Interval kelas

n : Ukuran sampel

fk_{med} : Frekuensi kumulatif sebelum kelas median

f_{med} : Frekuensi pada kelas median

Contoh :

Lihat kembali data berat badan 64 mahasiswa Psikologi Gunadarma pada tabel 4.4.

Tabel 4.4

Distribusi Frekuensi Berat Badan 64 Mahasiswa Psikologi Gunadarma 2003

Berat badan (kg)	Frekuensi	Frekuensi Kumulatif
36-44	20	20
45-53	19	39
54-62	17	56
63-71	5	61
72-80	1	62
81-89	1	63
90-98	1	64
Jumlah	64	

Titik posisi median = 32. Kelas posisi median yaitu kelas ke-2.

B_{med} = 44.5

I = 9

fk_{med} = 20

f_{med} = 19

$$\text{Med} = \text{B}_{\text{med}} + \left[\frac{\left(n/2\right) - fk_{med}}{f_{med}}\right] i$$

$$= 44.5 + \left[\frac{32 - 20}{19}\right] 9 = 50{,}18$$

Keunggulan median:

1. Tidak dipengaruhi oleh data ekstrim
2. Mudah dimengerti dan mudah dihitung, baik dari data tidak berkelompok maupun data berkelompok. Juga dapat dihitung untuk data berkelompok dengan kelas terbuka.
3. Dapat digunakan untuk data kuantitatif maupun data kualitatif.

Kelemahan median:

1. Hanya dapat ditentukan dari data yang telah diurutkan sehingga membutuhkan waktu yang tidak sedikit.
2. Dihitung bukan berdasarkan nilai data, tetapi berdasarkan jumlah data, sehingga sulit dijadikan sebagai ukuran pusat data untuk menggambarkan kumpulan datanya.

4.3 MODUS

Berdasarkan jumlah data maka median sulit dijadikan sebagai ukuran pusat data yang dapat menggambarkan rangkaian datanya. Misalnya, dua rangkaian· data berikut memiliki median yang sama akan tetapi median pada rangkaian kedua jelas tidak representatif jika dijadikan sebagai ukuran pusat data.

6, 9, 12, 14, 17,21

6,9, 12,100,140,500

Modus, sebagai ukuran pusat data, berbeda dengan rata-rata hitung dalam penentuannya. Modus lebih mirip median dalam penentuannya yang tidak melalui proses aritmatik seperti halnya penentuan rata-rata. Modus adalah suatu nilai yang terdapat dalam serangkaian data yang memiliki frekuensi tertinggi.

Modus untuk Data Tak-berkelompok

Modus adalah nilai yang memiliki frekuensi tertinggi.

Contoh :

Lihat kembali data tinggi badan 10 orang mahasiswa pada contoh 3.1: 162, 161, 157, 154, 164, 170, 162, 165, 170, 161. Modus akan lebih mudah ditentukan jika data tersusun dalam distribusi frekuensi seperti di bawah ini.

Tabel 4.5

Distribusi frekuensi tinggi badan 10 mahasiswa

Tinggi Badan (cm)	Frekuensi
154	1
157	1
161	2
162	3
164	1
165	1
170	1

Tampak bahwa modus data adalah *Mo* = 162

Modus untuk Data Berkelompok

$$\boxed{\mathrm{Mo} = \mathrm{B_{mo}} + \left[\frac{d_i}{d_1 + d_2}\right]\mathrm{i}}$$

Mo : Modus

B_{mo} : Tepi batas kelas bawah pada kelas modus

I : Intervasl kelas

d_1 : Frekuensi kelas modus dikurangi frekuensi kelas sebelum kelas modus

d_2 : Frekuensi kelas modus dikurangi frekuensi kelas sesudah kelas modus

Contoh :

Lihat kembali data berat badan 64 mahasiswa Psikologi Gunadarma dan distribusi frekuensi beserta frekuensi kumulatifnya pada tabel 4.6. Kelas posisi modus yaitu kelas pertama.

Tabel 4.6

Distribusi Frekuensi Berat Badan 64 Mahasiswa Psikologi Gunadarma 2003

Berat badan (kg)	Frekuensi	Frekuensi Kumulatif
36-44	20	20
45-53	19	39

54-62	17	56
63-71	5	61
72-80	1	62
81-89	1	63
90-98	1	64
Jumlah	64	

B_{mo} = 35.5

I = 9

d_1 = 20 - 0 = 20

d_2 = 20 - 19 = 1

$$Mo = B_{mo} + \left[\frac{d_i}{d_1 + d_2}\right]i$$

$$= 35{,}5 + \left[\frac{20}{20+1}\right]9 \qquad = 44{,}07$$

Keunggulan modus:

1. Dapat digunakan untuk data kualitatif maupun kuantitatif.
2. Tidak dipengaruhi oleh data ekstrim.
3. Dapat dihitung untuk data berkelompok dengan kelas terbuka.

Kelemahan modus:

1. Dalam kasus-kasus tertentu, kumpulan data tidak memiliki modus.
2. Jika modus justru lebih daripada satu, tidak dapat digunakan sebagai ukuran pusat data.

Contoh :

Hitung rata-rata, median, dan modus distribusi frekuensi berikut:

Tabel 4.7

Distribusi frekuensi IPK 32 orang mahasiswa

IPK	Frekuensi
$X < 1.5$	1
$1.5 \leq X < 2.0$	4
$2.0 \leq X < 2.5$	5
$2.5 \leq X < 3.0$	7
$3.0 \leq X < 3.5$	11
$X \geq 3.5$	4
Jumlah	32

Tabel 4.8

Perhitungan rata-rata, median, dan modus IPK 32 orang mahasiswa

X_j	f_j	$X_j f_j$	Frekuensi kumulatif
1.52	1	1.25	1
1.75	4	7.00	5
2.25	5	11.25	10
2.75	7	19.25	17
3.25	11	35.75	28
3.75	4	15.00	32
Jumlah	32	89.50	

a. Rerata:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{89.50}{32} = 2.80$$

b. Median:
Kelas posisi median adalah kelas keempat.

$$Med = B_{med} + \left[\frac{(n/2) - fk_{med}}{f_{med}}\right] i$$

$$= 2.50 + \left[\frac{16-10}{7}\right] 0.5 = 2.93$$

c. Modus:
Kelas posisi modus adalah kelas kelima.

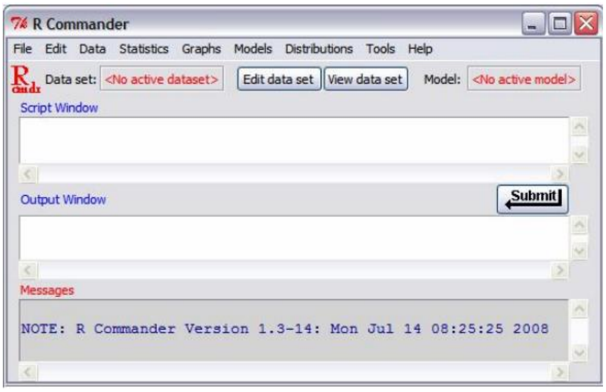
$$Mo = B_{mo} + \left[\frac{d_i}{d_1 + d_2}\right] i$$

$$= 3.00 + \left[\frac{4}{4+7}\right] 0.5 = 3.18$$

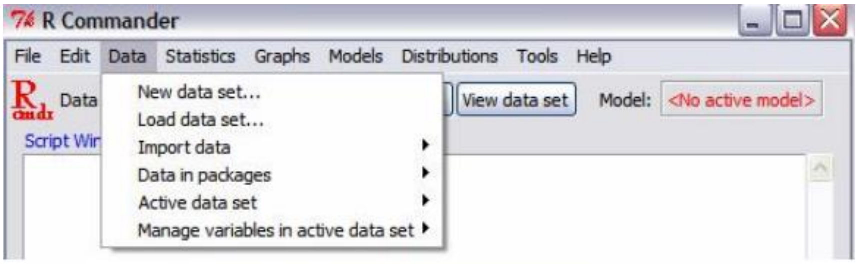
Contoh :

Diketahui data umur pegawai PT DOFI yaitu 19 40 38 31 42 20 27 22 37 42 untuk mencari nilai-nilai ukuran statistik data (rata-rata, median, kuartil ke-1 dan kuartil ke-3) tersebut dengan menggunakan program R, ikutilah langkah-langkah berikut :

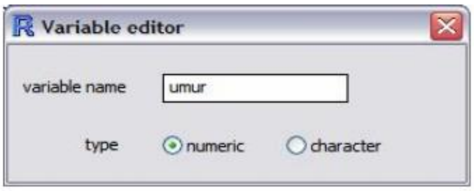
1. Tekan icon R Commander pada desktop, kemudian akan muncul tampilan seperti gambar di bawah ini.

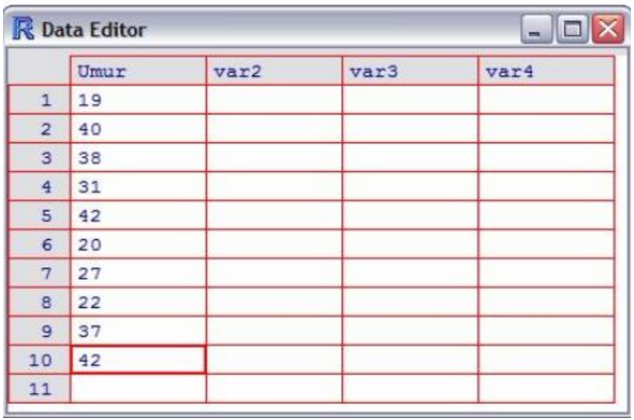


2. Pilih menu Data, New Data Set. Masukkan nama dari data set adalah umur, lalu tekan tombol OK

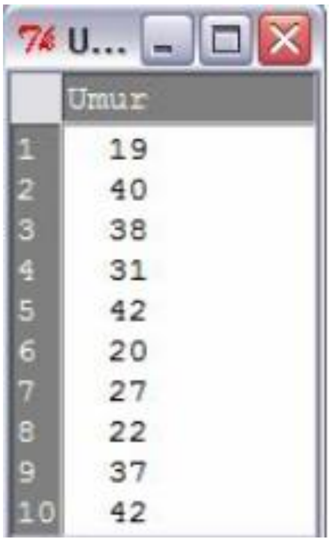


3. Masukkan data umur pegawai PT. DOFI. Jika data editor tidak aktif maka dapat diaktifkan dengan menekan RGui di taskbar windows pada bagian bawah layar monitor. Jika sudah selesai dalam pengisian data tekan tombol Close. Untuk mengubah nama dan tipe variabel, dapat dilakukan dengan cara Double Click pada variabel yang ingin di setting.

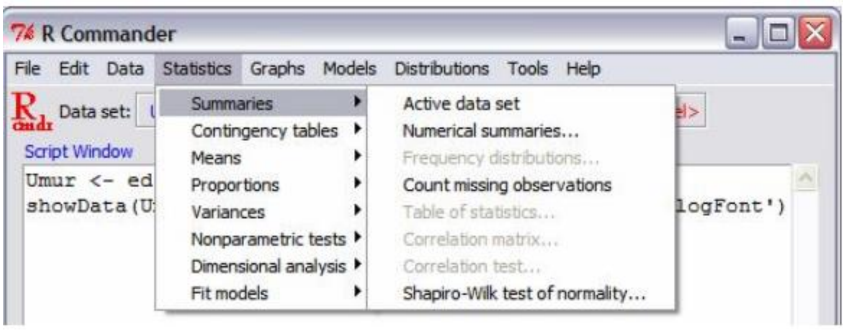




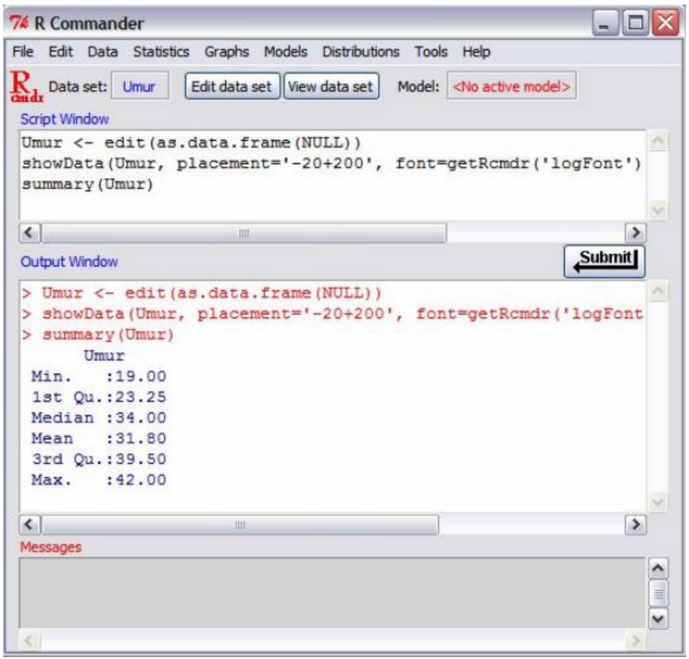
4. Untuk mengecek kebenaran data yang sudah dimasukkan, tekan tombol View Data Set maka akan muncul tampilan seperti gambar di bawah ini. Jika ada data yang salah, tekan tombol Edit Data Set, lalu perbaiki data yang salah.



5. Jika data sudah benar, pilih menu Statistic, Summaries, Active Data Set



6. Akan muncul tampilan :



Maka kita bisa mengetahui bahwa dari data umur pegawai PT. DOFI, memiliki nilai :

Minimum	: 19.00
Kuartil 1	**: 23.25**
Median	**: 34.00**
Mean (rata-rata)	**: 31.80**
Kuartil 3	**: 39.50**
Maximum	: 42.00

4.4 Kuartil, Desil, Persentil, Skewness, Kurtosis

Jika tiga ukuran di atas merupakan ukuran lokasi yang cenderung bertindak sebagai ukuran pusat data, maka ketiga ukuran ini hanya merupakan ukuran lokasi. Kendati bukan sebagai ukuran pusat data, ukuran ini banyak bermanfaat bagi para pengambil keputusan. Pada akhir sub bab akan disajikan satu contoh penggunaan ukuran ini. Tiga ukuran tersebut adalah kuartil, desil, dan persentil. Untuk ketiga-tiganya, pembahasan akan ditekankan untuk data yang telah dikelompokkan saja. Dalam perhitungan nanti, ketiga ukuran ini tidak berbeda dengan perhitungan median.

Kuartil untuk Data Tak-berkelompok

Kuartil membagi sederetan data terurut menjadi empat bagian yang sama. Terdapat tiga kuartil, yaitu kuartil pertama (Q_1), kuartil kedua (Q_2), dan kuartil ketiga (Q_3).

Posisi kuartil (n < 30):

$$\text{Posisi } Q_1 = \frac{n+2}{4}$$

$$\text{Posisi } Q_2 = \frac{2n+2}{4} = \frac{n+1}{2} \quad = \text{posisi median}$$

$$\text{Posisi } Q_3 = \frac{3n+2}{4}$$

Kuartil untuk Data Berkelompok

Nilai kuartil (n > 30):

$$Q_1 = B_q + i\left[\frac{n/4 - f_{kq}}{f_q}\right]$$

$$Q_3 = B_q + i\left[\frac{3n/4 - f_{kq}}{f_q}\right]$$

Keterangan :

Q_1 : Kuartil pertama·

Q_3 : Kuartil ketiga

B_q : Tepi batas kelas bawah pada kelas kuartil.

I : Interval kelas

n : Ukuran sampel

f_{kq} : Frekuensi kumulatif sebelum kelas kuartil

f_q : Frekuensi pada kelas kuartil

Contoh

Lihat tabel di bawah tentukanlah kuartil pertama dan kuartil ketiga!

Kelas	Frekuensi
20 - < 30	7
30 - < 40	8
40 - < 50	10
50 - < 60	15
60 - < 70	25
70 - < 80	10
80 - < 90	5
Jumlah	80

Kuartil pertama:

Titik kuartil pertama: 80/4 = 20

B_q : 40

i : 10

f_{kq} : 15

f_q : 10

$$Q = 40 + 10\left[\frac{20 - 15}{10}\right]$$

$Q_1 = 45$

Kuartil ketiga:

Titik kuartil ketiga : 3n / 4 = 240/4 = 60

B_q : 60

I : 10

f_{kq} : 40

f_q : 25

$$Q = 60 + 10\left[\frac{60 - 40}{25}\right]$$

$Q_1 = 68$

Desil dan Persentil

Jika pada kuartil deretan data terurut dibagi menjadi 4, maka pada desil, deretan data terurut dibagi menjadi 10 bagian yang sama. Perumusan yang digunakan pun tidak jauh berbeda. Yang berbeda hanya bagian rumus yang menentukan titik -titik desil. Berikut tabel yang memuat bagian rumus yang menentukan sembilan titik desil:

Desil ke-1:	n/10	
Desil ke-2:	2n/10	
Desil ke-3:	3n/10	
Desil ke-4:	4n/10	
Desil ke-5:	5n/10	$\longrightarrow$ Median
Desil ke-6:	6n/10	
Desil ke-7:	7n/10	
Desil ke-8:	8n/10	
Desil ke-9:	9n/10	

Adapun bagian-bagian lainnya menyesuaikan letak titik desil yang bersangkutan.

Contoh

Lihat tabel pada contoh kuartil. Tentukanlah desil ke-7!

Letak titik desil ke-7	= (80x7) / 10 = 56
Bd (tepi batas bawah kelas desil)	= 60
fkd (frekuensi kumulatif sebelum kelas desil)	= 40
fd (frekuensi pada kelas desil)	= 25

Desil ke-7 (d_7) = $60 + 10\left[\frac{56-40}{25}\right]$

Desil ke-7 = 66,4

Demikian pula dalam menentukan persentil. Bagian rumus yang berubah hanyalah bagian yang menentukan letak titik persentil, dan bagian-bagian yang lainnya menyesuaikan persentil yang dimaksud.

Posisi beberapa titik persentil:

Persentil ke-1	$n/100$
Persentil ke-12	$12n/100$
Persentil ke-27	$27n/100$
Persentil ke-87	$87n/100$
Persentil ke-99	$99n/100$

Contoh

Lihat tabel di bawah ini. Tentukan persentil ke-67!

Kelas	Frekuensi
20 - < 30	7
30 - < 40	8
40 - < 50	10
50 - < 60	15
60 - < 70	25
70 - < 80	10
80 - < 90	5
Jumlah	80

Letak titik persentil ke-67 = (80x67) / 100 = 56

B_p (tepi batas bawah kelas persentil) = 60

f_{kp} (frekuensi kumulatif sebelum kelas persentil) = 40

f_p(frekuensi pada kelas persentil) = 25

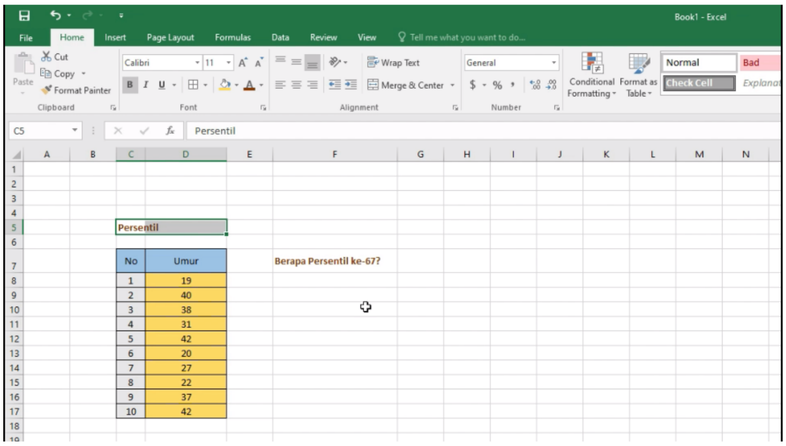
Persentil ke-67 (p) = 60 +10 $\left[\frac{53,6-40}{25}\right]$

Persentil ke-67 = 65,44

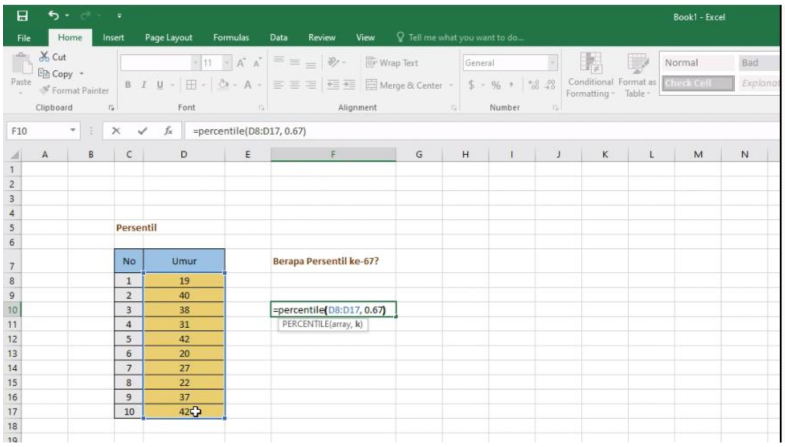
Persentil untuk data tidak berkelompok

Diketahui data umur pegawai PT DOFI yaitu 19 40 38 31 42 20 27 22 37 42 untuk mencari persentil ke-67 dengan menggunakan Microsoft Excel, ikutilah langkah-langkah berikut :

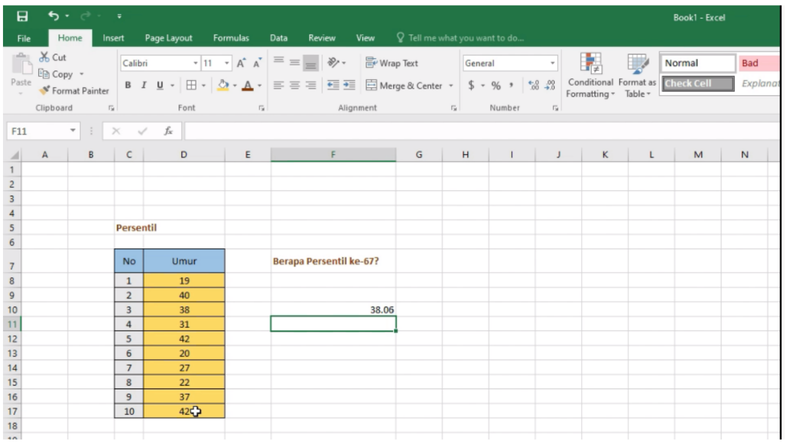
1. Buka Microsoft Excel lalu masukan data yang ingin dicari persentilnya.



2. Untuk mencari persentil pada data tersebut ketikkan rumus =Percentile(D8:D17,0.67) atau tuliskan rumus buka kurung lalu blok data dari cell D8 sampai D17 dan untuk persentil karena membagi menjadi 100 bagian 67 jadi 0,67 tutup kurung lalu tekan enter.



3. Maka akan muncul hasil persentil ke-67 yaitu sebesar 38,06



Ukuran Kecondongan (Skewness)

Kita telah membicarakan ukuran pemusatan dengan menggunakan rata-rata hitung ($\overline{x}$), median (*Md*) dan modus (*Mo*). Hubungan $\bar{x}$, *Md, Mo* apabila $\overline{x}$ = *Md* =*Mo* maka kurva akan simetris. Apabila *x ≠ Md ≠ Mo* maka akan terjadi kecondongan kurva *(skewness)*.

Kecondongan dapat berbentuk kecondongan positif dan negatif. Kecondongan positif terjadi bila ada pengamatan yang nilainya sangat besar sehingga rata-rata hitung menjadi besar dan μ = $\overline{x}$ <*Md* <*Mo*. Ukuran kecondongan dirumuskan sebagai berikut :

$$\text{sk} = \frac{\mu - Mo}{\sigma} \qquad \text{atau} \qquad \text{sk} = \frac{3\,(\mu - Md)}{\sigma}$$

Dimana :

sk = Koefisien kecondongan

μ = Rata-rata hitung

Mo = Modus

σ = Standar deviasi

Md = Median

Contoh :

Tabel di bawah yaitu data kepadatan jumlah penduduk Kabupaten Bengkulu Selatan pada tahun 2003.

Kecamatan	Kepadatan Penduduk
Manna	129
Kota Manna	342
Kedurang	53
Seginim	171
Pino	62
Pino Raya	68

Hitunglah koefisien kecondongan dari kepadatan jumlah penduduk, apabila koefisien negatif condong ke kiri berarti penduduk mengarah ke perkotaan dan sebaliknya.

Jawab:

Rumus = Sk = [3 (μ - Md)]/ σ

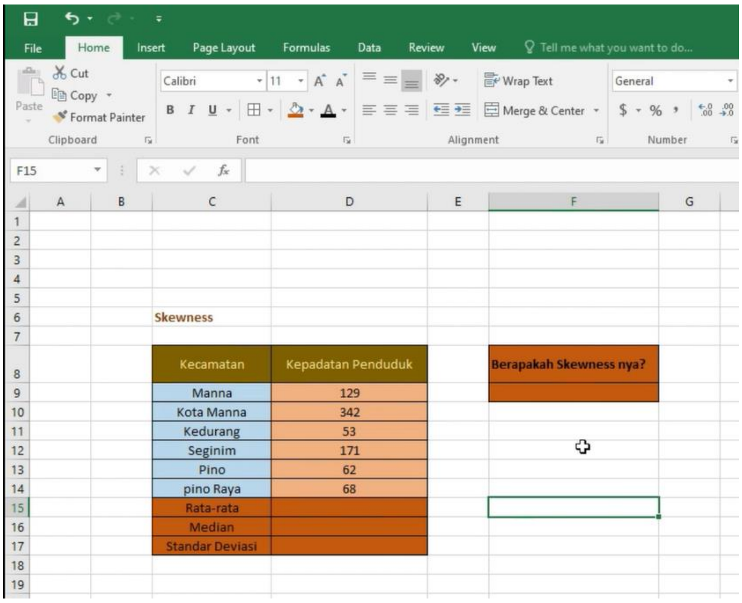
Dengan menggunakan bantuan komputer diperoleh bahwa: μ = 137,5 ; Md = 98,5 dan σ = 110

Sk = [3 ((μ - Md)]/σ = [3 (137,5 - 98,5)]/ 110 = 1,06

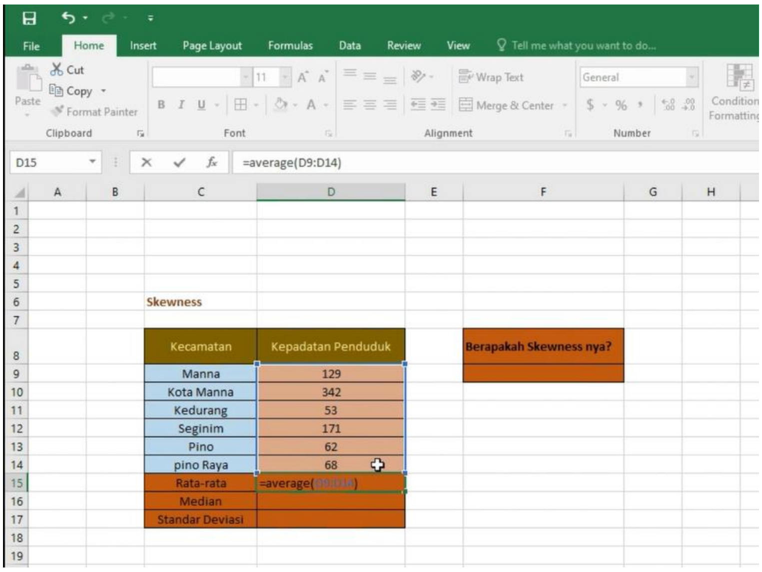
Nilai SK positif dan relatif besar dibandingkan dengan nol, ini menunjukkan bahwa kepadatan penduduk relatif condong ke kanan, atau banyak penduduk yang tinggal di pedesaan dibandingkan dengan perkotaan.

Pengerjaan dengan Excel

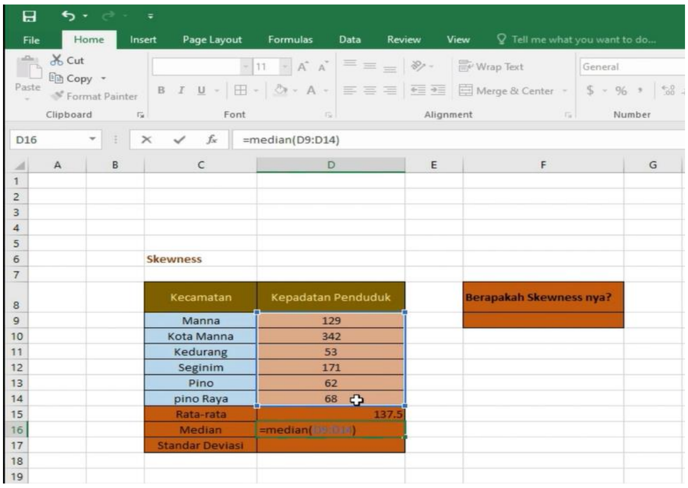
1. Buka Microsoft Excel lalu masukan data yang ingin dicari skewnessnya.



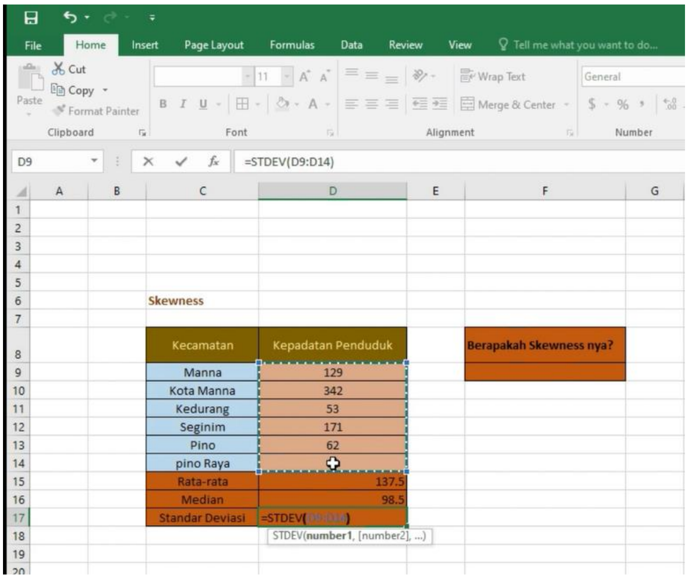
2. Untuk mencari rata-rata pada data tersebut ketikkan rumus =average(D9:D14) di cell D15 lalu tekan enter.



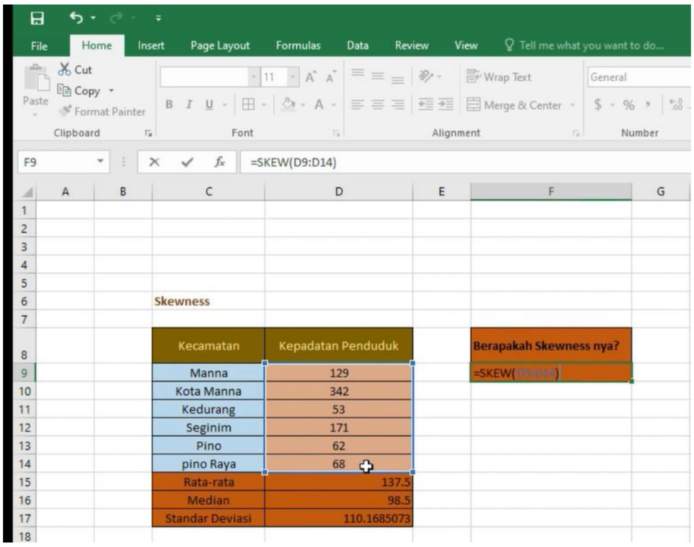
3. Untuk mencari median pada data tersebut masukkan rumus =median(D9:D14) di D16 lalu tekan enter.



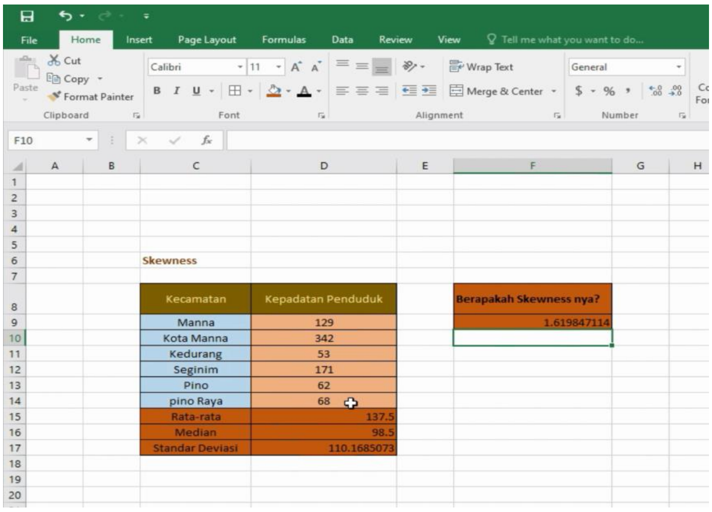
4. Untuk mencari standar deviasinya yaitu masukkan rumus=STDEV(D9:D14) pada cell D17 lalu tekan enter.



5. Ada cara yang lebih cepat yaitu dengan langsung memasukkan rumus skewnessnya =SKEW(D9:D14) dengan memblok data pada cell D9 sampai D14 lalu enter.



6. Maka akan muncul nilai skewnessnya yaitu sebesar 1,6



Ukuran Keruncingan *(Kurtosis)*

Ukuran keruncingan suatu kurva dapat dilakukan dengan membandingkan dengan kurva yang simetris (kurva normal). Tiga bentuk kurva keruncingan yaitu : kurva yang distribusi puncaknya berbentuk tidak mendatar dan tidak meruncing disebut *mesokurtic* (distribusi normal). Kurva yang puncaknya berbentuk mendatar yang disebut *platykurtic* dan kurva yang distribusi puncaknya meruncing disebut *leptokurtic.*

Koefisien keruncingan dapat diperoleh dengan menggunakan rumus:

$$\alpha^4 = \frac{{}^{1}\!/_{n} \sum f\,.(X-\mu)^4}{\sigma^4}$$

Dimana:

α^4 = Koefisien Kurtosis

n = Jumlah Data

f = Jumlah Frekuensi Kelas

X = Nilai Tengah Kelas

μ = Nilai Rata - rata Hitung Data

σ = Standar Deviasi

Contoh

Tabel di bawah yaitu pertumbuhan Ekonomi Beberapa Negara Asia Tahun 2002

Negara	2002
Cina	7,4
Philipina	4,0
Hongkong	1,4
Indonesia	3,2
Kamboja	5,0
Korea Selatan	6,0
Malaysia	4,5
Singapura	3,9
Thailand	3,8
Vietnam	5,7

Hitunglah koefisien keruncingannya !

Penyelesaian :

X	$(X-\mu)$	$(X-\mu)^2$	$(X-\mu)^4$
7,4	2,9	8,4	70,7
4,0	-0,5	0,3	0,1
1,4	-3,1	9,6	92,4
3,2	-1,3	1,7	2,9
5,0	0,5	0,3	0,1
6,0	1,5	2,3	5,1
4,5	0,0	0,0	0,0

3,9	-0,6	0,4	0,1
3,8	-0,7	0,5	0,2
5,7	1,2	1,4	2,1

Σ X = 44,9 ; $\mu = \Sigma X/n = 44{,}9/10 = 4{,}5$

$\Sigma (X - \mu)^2 = 24{,}8$ $\Sigma (X - \mu)^4 = 173{,}6$

Dari data diatas $\Sigma (X - \mu)^4 = 173{,}6$

Standart Deviasi $\sigma = \sqrt{}\ \Sigma (X - \mu)^2 /n = \sqrt{}\ 24{,}8/10 = \sqrt{2{,}48} = 1{,}6$

$$\alpha^4 = \frac{{}^{1}/_{n} \sum f \,.(X-\mu)^4}{\sigma^4} = \alpha^4 = \frac{{}^{1}/_{10}\,.173{,}7}{1{,}6^4} = \frac{17{,}37}{6{,}25} = 2{,}8$$

jadi nilai α^4= 2,8 dan lebih kecil dari 3, maka kurvanya termasuk *platykurtic*

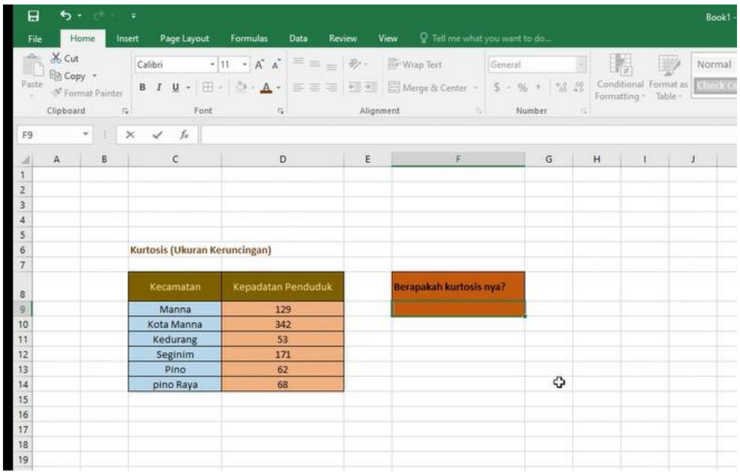
Contoh ke 2:

Kecamatan	Kepadatan Penduduk
Manna	129
Kota Manna	342
Kedurang	53
Seginim	171
Pino	62
Pino Raya	68

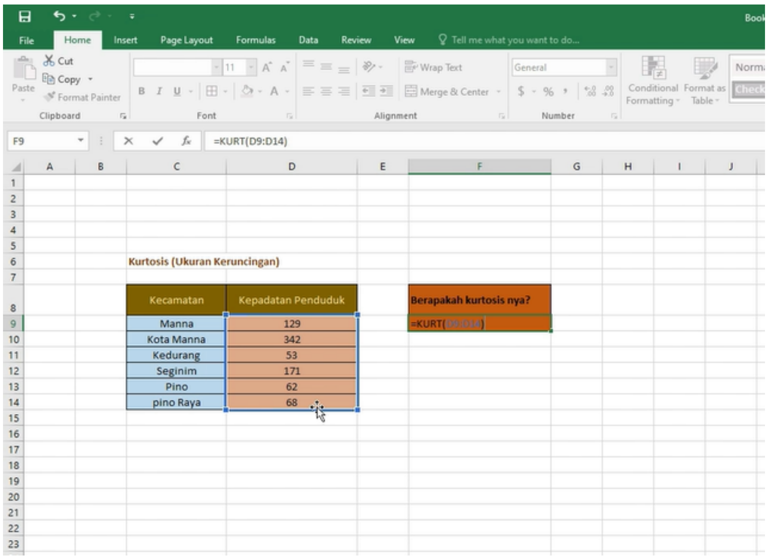
Carilah kurtosisnya!

Penyelesaian dengan microsoft excel

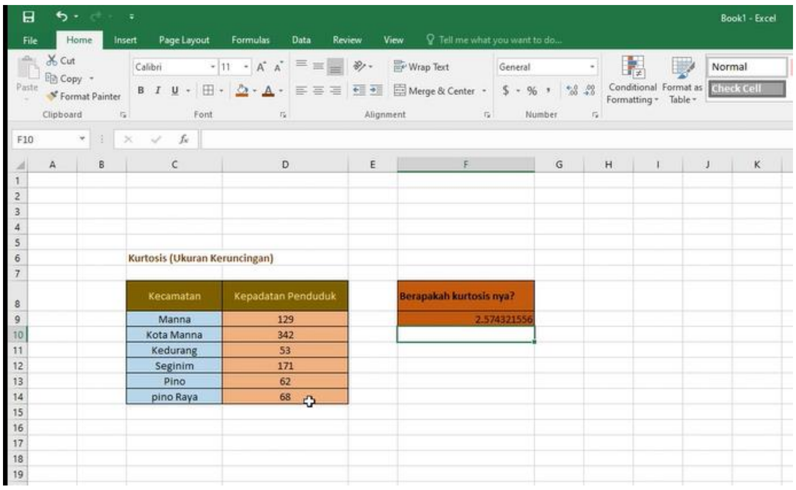
1. Bukalah program Microsoft Excel lalu masukkan data yang ingin dicari nilai kurtosisnya



2. Kemudian masukkan rumus kurtosis =KURT(D9:D14) dengan memblok datanya pada cell D9 sampai D14 lalu tekan enter.



3. Maka akun muncul nilai kurtosisnya yaitu sebesar 2,57



4.5 Rata – Rata Tertimbang

Misalkan W menyatakan rata-rata tertimbang untuk data X_1, X_2, . . . , X_n, masing-masing dengan penimbang (bobot) W_1, W_2, . . . , W_n, maka rata-rata tertimbang adalah:

$$\boxed{W = \frac{\sum_{i=1}^{n} W_i X_i}{\sum_{i=1}^{n} W_i}}$$

Jika penimbang dinyatakan dalam proporsi (atau persentase), maka $\sum w_i = 1$, sehingga:

$$\boxed{W = \sum_{i=1}^{n} W_i X_i}$$

Contoh :

Misalkan mahasiswa Y mendapatkan nilai 90 untuk tugas harian mata kuliah Statistika, 80 untuk Ujian Tengah Semester, dan 60 untuk Ujian Akhir Semester. Jika bobot tugas harian, UTS, dan UAS masing-masing adalah 10%, 60%, dan 30%, maka nilai akhirnya (dihitung sebagai rata-rata tertimbang)adalah:

$$W = \sum_{i=1}^{n} w_i x_i$$

$$= (0.10)(90) + (0.60)(80) + (0.30)(60) = 75$$

4.6 Rata – Rata Geometrik

Tidak jarang, seseorang harus menghitung rata-rata pertumbuhan suatu kualitas atau nilai sesuatu, misalnya rata-rata pertumbuhan nilai penjualan, rata-rata pertumbuhan jumlah penduduk, dan lain sebagainya. Untuk menghitungnya, penggunaan rata-rata hitung tidak dapat digunakan lagi dan tentunya diperlukan cara lain, yaitu rata-rata **geometrik** atau **rata-rata ukur**.

Misalkan G menyatakan rata-rata geometrik untuk data X_1 , X_2 , . . . , Xn , maka:

$$\boxed{G^n = X_1.X_2........X_n = \prod_{i=1}^{n} X_i}$$

$$\text{Log } G^n = \log (X_1, X_2 \text{........,} X_n = \log X_1 + \log X_2 + ... + \log X_n$$

$$n \log G = \sum_{i=1}^{n} logX_i$$

sehingga
$$\boxed{\text{Log } G = \frac{\sum_{i=1}^{n} logX_i}{n}}$$

atau
$$\boxed{G = \sqrt[n]{X_1 X_2 \ldots X_n} = \sqrt[n]{\prod_{i=1}^{n} X_i}}$$

Misalkan jumlah kasus DBD (Demam Berdarah Dengue) di kota B pada tahun 2000, 2001, 2002, dan 2003 masing-masing adalah 124, 130, 143, dan 158. Rata-rata geometriknya adalah:

$$G = \sqrt[n]{X_1 X_2 \ldots X_n}$$

$$= \sqrt[4]{(124)(130)(143)(158)} = 138{,}15$$

DAFTAR PUSTAKA

Harlan, Johan. 2004. Metode Statistika 1. Jakarta: gunadarma.

Kustituanto, Bambang., dan Rudy Badrudin. 1994. *Buku Statistika I (Deskriptif)*. Jakarta: gunadarma.

Modul Ilab Statistika 1