

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi
Republik Indonesia, 2021

Matematika untuk SMA/SMK Kelas X

Penulis: Dicky Susanto, dkk

ISBN: 978-602-244-526-5

Bab

7

Statistika

Pengalaman Belajar

Setelah mempelajari bab ini, kalian diharapkan dapat:

1. Membedakan berbagai macam jenis data serta membuat grafik yang sesuai dan merepresentasikan data tersebut, serta melakukan analisis data untuk pengambilan kesimpulan.
2. Menggambar dan menginterpretasikan histogram, diagram garis batang, *line plot*;
3. Menentukan ukuran pemusatan dari kumpulan data: *mean*, median, dan modus, pada data tunggal dan data kelompok.
4. Menentukan ukuran penempatan dari kumpulan data: kuartil dan persentil pada data tunggal dan data kelompok.
5. Mengetahui ukuran penyebaran dari kumpulan data: jangkauan inter kuartil, varian, dan simpangan baku pada data tunggal dan data kelompok.
6. Membandingkan 2 kelompok data menggunakan ukuran pemusatan dan penyebaran.

Data Menolong Kita

Dalam kehidupan sehari-hari, kita selalu berhubungan dengan data. Pernahkah kalian melihat kumpulan data, atau pernahkah kalian mengumpulkan data? Di tengah pandemi Covid-19 ini, kita bisa melihat data berapa banyak orang yang terkena infeksi karena virus korona. Dari data harian yang kita kumpulkan, kita bisa melakukan analisis sederhana, seperti daerah mana yang tingkat penularannya sudah semakin turun, mana daerah yang justru tingkat penularannya malah semakin naik. Selain itu, kita juga bisa melihat berapa rata-rata tingkat kesembuhan dari pasien Covid-19 setiap harinya. Untuk dapat menarik kesimpulan dari hal-hal di atas, kita memerlukan data lalu mengolahnya sehingga kita dapat memahami situasi yang sesungguhnya berdasarkan fakta yang aktual, bukan berdasarkan perasaan atau berita hoaks.

Sebagai contoh, ketika vaksin Covid-19 berhasil ditemukan di beberapa negara, muncul berita hoaks yang menyebutkan bahwa vaksin tersebut bermasalah. Dari data yang berhasil dikumpulkan dan diolah, ternyata dari setiap 40.000 orang yang divaksin, rata-rata akan ada 5 orang atau sekitar 0,01% yang mengalami masalah. Hal ini disebabkan berbagai hal seperti usia lanjut, penyakit bawaan, dan sebagainya. Sementara itu, diketahui bahwa rata-rata tingkat kematian dari pasien Covid-19 yang tidak sempat divaksin di Indonesia mencapai 3%, atau jika ada 6.000 pasien baru, maka diperkirakan rata-rata 200 pasien akan berakhir dengan kematian. Dari fakta ini kita bisa melihat bahwa persentase pasien Covid-19 yang tidak divaksin 300 kali lebih berisiko daripada persentase orang yang bermasalah karena vaksin. Wow! 300 kali lebih berisiko tanpa vaksin. Jumlah yang tidak sedikit.

Jadi, dengan melihat data tersebut kita dapat memahami situasi dengan lebih baik sehingga kita mampu mengambil keputusan dengan lebih tepat.

Statistik adalah ilmu yang akan membantu kalian menguasai berbagai hal yang terkait dengan data, mulai dari pengumpulan data, mengolahnya, menganalisis sampai akhirnya mengambil keputusan berdasarkan data.

Pada waktu SMP, kalian telah belajar bagaimana untuk melihat ukuran pemusatan dari sekumpulan data, kalian mencari rata-rata atau *mean*, modus, dan median. Saat ini, kalian akan mempelajari jenis data, ukuran penyebaran serta ukuran lokasi dari sekumpulan data supaya kalian dapat menarik kesimpulan dengan lebih baik.

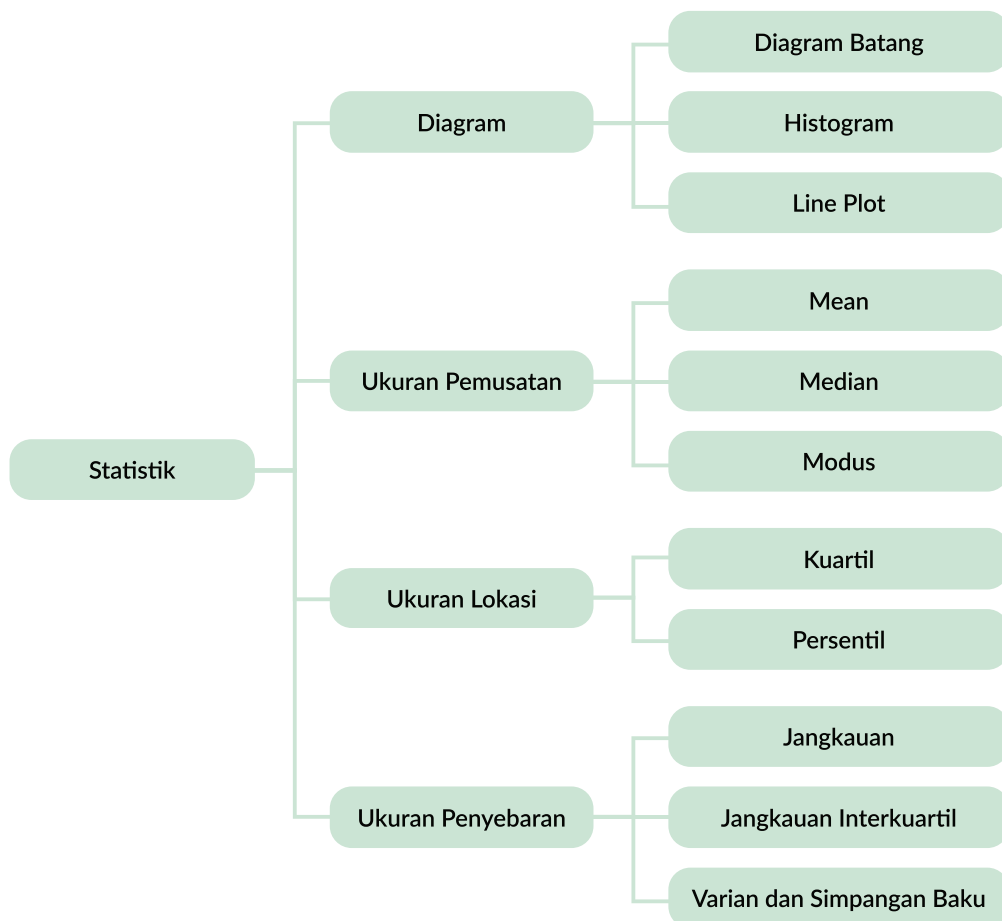
Kata Kunci

Frekuensi, frekuensi relatif, histogram, diagram batang, *line plot*, jangkauan, *mean*, modus, median, simpangan baku, varian, pencilan.

Pertanyaan Pemantik

1. Bagaimana pengolahan data dapat membantu kita dalam pengambilan keputusan?
2. Bagaimana kita menentukan ukuran pemusatan yang paling sesuai dengan konteks masalah yang dihadapi?
3. Bagaimana ukuran pemusatan dan ukuran penyebaran dapat membantu kita untuk membedakan 2 kelompok data?

Peta Konsep





Ayo Mengingat Kembali

1. Diagram lingkaran sederhana atau diagram batang sederhana dapat digunakan untuk menampilkan informasi yang tersedia, baik dalam bentuk data tunggal maupun dalam bentuk tabel frekuensi.
2. Dari data tunggal sederhana atau data dalam tabel distribusi frekuensi, bisa diperoleh ukuran pemusatan: mean, modus, dan median yang dapat memberikan gambaran tentang kumpulan data tersebut.

Eksplorasi 7.1 Penggunaan Diagram Batang untuk Menganalisis Data



Ayo Bereksplorasi

Pandemi Covid-19 melanda seluruh dunia. Setiap harinya jumlah pasien yang terinfeksi virus Covid-19 terus bertambah. Pada tabel berikut, kalian dapat melihat rata-rata pertambahan pasien baru positif Covid-19 setiap minggunya di Provinsi DKI Jakarta.

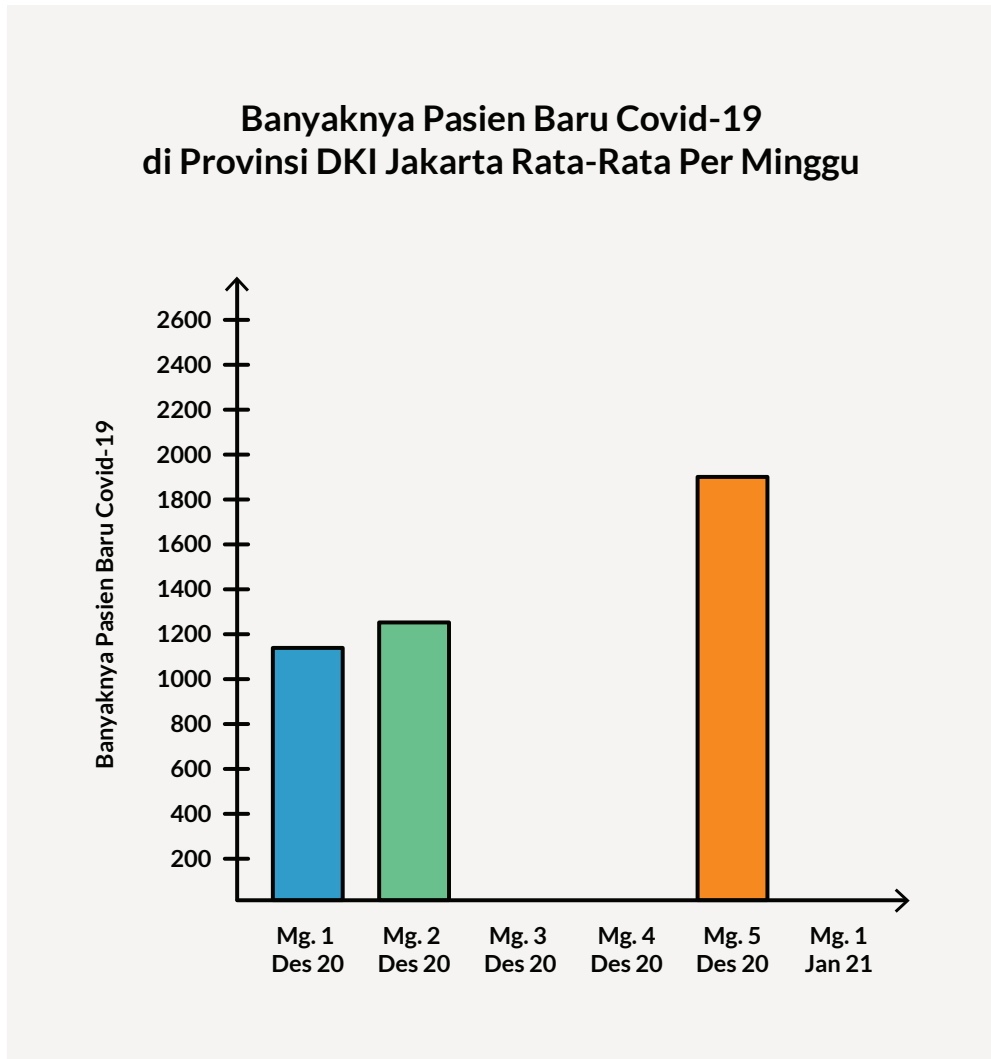
Tabel 7.1 Frekuensi Banyaknya Pasien Baru Covid-19

Tanggal	Banyaknya Pasien Baru Positif Covid-19*
Minggu ke-1 Desember 2020	1.170
Minggu ke-2 Desember 2020	1.220
Minggu ke-3 Desember 2020	1.520
Minggu ke-4 Desember 2020	1.830
Minggu ke-5 Desember 2020	1.900
Minggu ke-1 Januari 2021	2.120

*jumlah dibulatkan ke puluhan terdekat. Sumber: corona.jakarta.go.id

1. Dari Tabel 7.1 di atas, pada minggu ke berapakah yang mengalami rata-rata kenaikan jumlah pasien positif Covid-19 yang paling besar?

2. Berdasarkan Tabel 7.1 di atas, lengkapi diagram batang di bawah ini.



Ayo Berdiskusi

Setiap siswa memberikan estimasinya disertai alasannya.

3. Jika pola penambahan rata-rata mingguan jumlah pasien positif Covid-19 ini terus bertambah, berikan estimasimu untuk jumlah pasien positif Covid-19 pada minggu ke-2 Januari 2021. Jelaskan alasanmu!



Ayo Bernalar

Pemilihan grafik yang tepat akan memberikan gambaran yang lebih tepat.

4. Saat menentukan kenaikan jumlah pasien Covid-19 yang paling besar, manakah yang lebih mudah digunakan, tabel atau diagram batang? Jelaskan!

A. Histogram

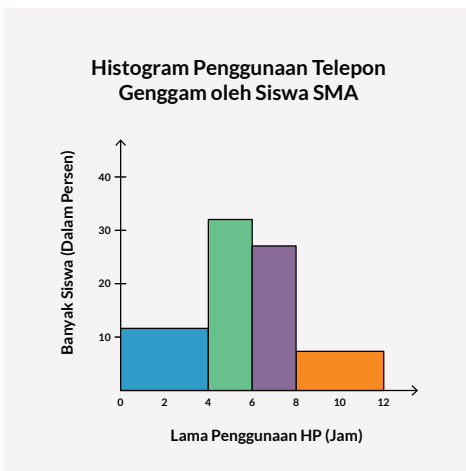
Ada berbagai tipe diagram. Diagram mana yang paling baik untuk digunakan sangat tergantung pada data apa yang kalian miliki dan informasi apa yang ingin kalian sampaikan.

Salah satu diagram yang dapat kalian gunakan adalah histogram. Histogram hampir serupa dengan diagram batang, namun histogram berbeda dengan diagram batang. Gambar 7.1 dan 7.2 menunjukkan contoh histogram dan diagram batang.

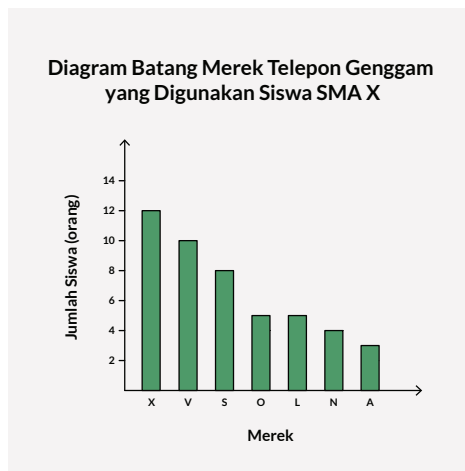


Ayo Berdiskusi

Dari Gambar 7.1 dan Gambar 7.2, carilah perbedaan dari histogram dan diagram batang.



Gambar 7.1 Histogram Penggunaan HP oleh Siswa SMA



Gambar 7.2 Diagram Batang Merek HP yang digunakan Siswa SMA

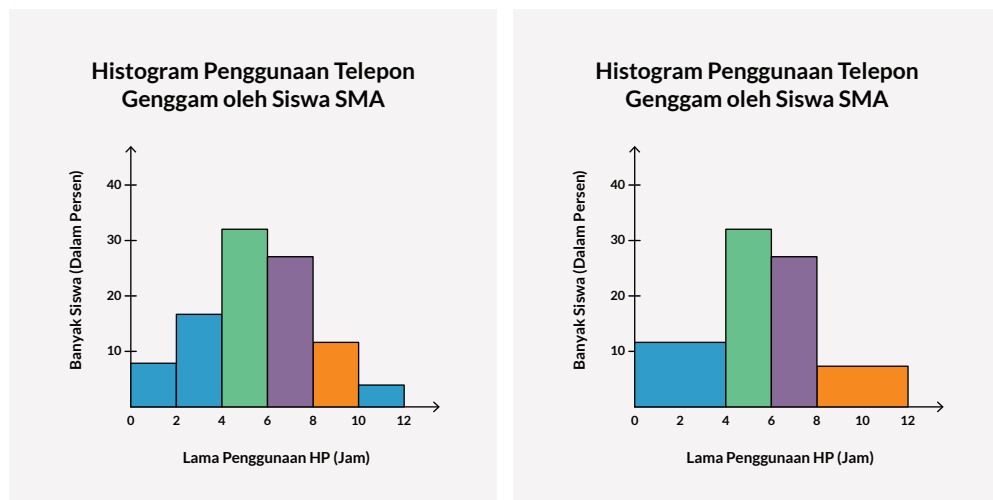
Histogram biasanya digunakan untuk menunjukkan distribusi dari suatu kelompok data, sedangkan diagram batang digunakan untuk membandingkan data. Histogram menampilkan data yang sifatnya kuantitatif dengan rentang data yang dikelompokkan ke dalam interval, sedangkan diagram batang menampilkan data yang sifatnya kategori.

Perbedaan lainnya, pada histogram, gambar batang menempel satu sama lain, sedangkan pada diagram batang, ada spasi antarbatang. Perbedaan terakhir, diagram batang biasanya memiliki batang dengan lebar yang sama, sedangkan lebar batang dalam histogram tidak perlu sama selama luas totalnya seratus persen jika digunakan persen atau luas total sama dengan jumlah data. Oleh karena itu, frekuensi data dalam diagram batang dilihat dari panjang batang, sedangkan frekuensi dalam histogram diberikan berdasarkan area pada masing-masing batang.



Ayo Berdiskusi

Perhatikan Gambar 7.3. Kedua histogram menampilkan data yang sama. Cobalah mencari bagaimana kedua histogram ini menjelaskan data yang sama walaupun terlihat berbeda.



Gambar 7.3 Tampilan Data yang Sama Menggunakan Dua Histogram yang Berbeda

Kalian bisa menggunakan pendekatan luas persegi panjang dalam menggambar histogram.

Pada histogram sebelah kiri:

- Frekuensi Kelas 0-2 adalah 8, luas persegi panjangnya adalah $2 \times 8 = 16$
- Frekuensi Kelas 2-4 adalah 16, luas persegi panjangnya adalah $2 \times 16 = 32$
- Luas gabungan kedua kelas tersebut adalah $16 + 32 = 48$

Pada histogram sebelah kanan:

- Frekuensi Kelas 0-4 adalah 12, luas persegi panjangnya adalah $4 \times 12 = 48$

Jadi, kelas 0-2 dan 2-4 pada histogram kiri memiliki luas yang sama dengan kelas 2-4 pada histogram kanan, sehingga dapat dikatakan bahwa histogram kiri dan histogram kanan menjelaskan data yang sama.

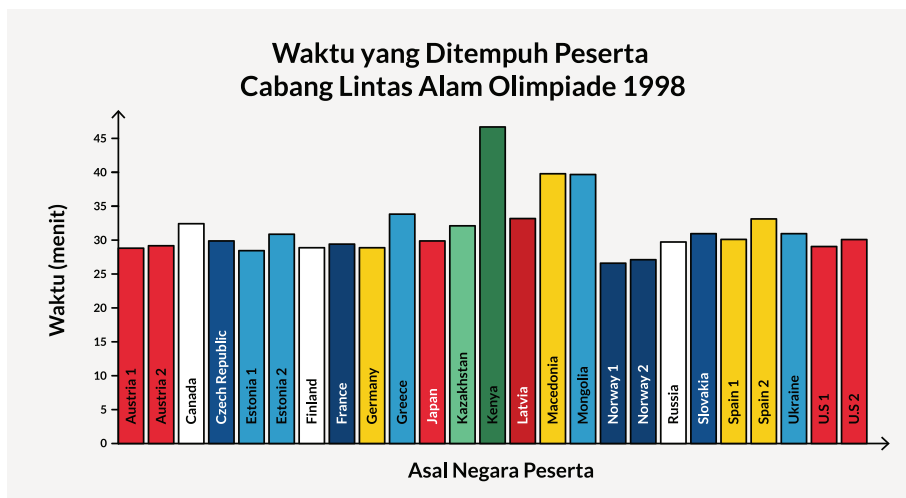


Ayo Mencoba

Coba kalian buktikan mengapa kelas 8-10 dan kelas 10-12 pada histogram kiri dapat digabung menjadi kelas 8-12 pada histogram kanan. Jelaskan jawabanmu!

Latihan 7.1

1. Perhatikan diagram batang berikut. Diagram berikut menunjukkan waktu yang ditempuh oleh para atlet di Olimpiade 1998 cabang Lintas Alam 10 km.



Gambar 7.4 Diagram Batang Waktu yang Ditempuh Peserta Lintas Alam Olimpiade 1998

Sumber: <https://www.olympic.org/nagano-1998/cross-country-skiing>

- a. Dari Gambar 7.4, ada berapa atlet yang berpartisipasi dalam cabang lintas alam? Ada berapa negara yang berpartisipasi dalam cabang ini?



Ayo Berpikir Kritis

Untuk dapat memahami diagram dengan lebih baik, maka kalian perlu memahami situasi dan konteks dari diagram.

- b. Peserta dari negara manakah yang mendapatkan medali emas? Berapakah catatan waktunya?
- c. Berapakah atlet yang menyelesaikan lomba ini dengan interval catatan waktu antara 31 menit dan 32 menit 59 detik?



Ayo Berpikir Kreatif

Ayo berpikir kreatif dalam menyesuaikan tampilan diagram batang untuk menjawab permasalahan.

- d. Gambar 7.4 disusun berdasarkan abjad dari nama depan asal negara atlet. Pikirkanlah cara lain untuk menyusun diagram batang ini. Pertanyaan seperti apakah yang mudah untuk dijawab dari susunan diagram batang yang baru tersebut?

Saat menjawab soal bagian c) di atas, mungkin kalian memerlukan waktu yang lebih lama karena harus melihat catatan waktu dari tiap atlet peserta. Meskipun kalian mudah membaca catatan waktu dari tiap atlet, tidak terlalu mudah untuk menemukan banyaknya atlet yang menyelesaikan lomba dengan interval catatan waktu tertentu.

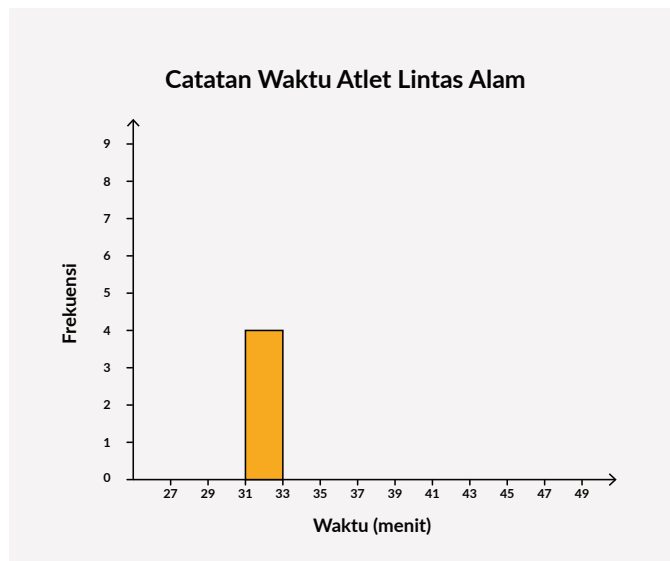
Sekarang, kalian akan menggunakan histogram untuk menampilkan data catatan waktu dari para atlet. Dalam histogram, data dibagi ke dalam interval yang sama, dengan 1 gambar batang untuk setiap interval. Tinggi dari setiap batang menunjukkan banyaknya data yang masuk dalam interval tersebut.

2. Dari data catatan waktu para atlet cabang Lintas Alam pada Gambar 7.4,
 - a. Lengkapilah kolom Frekuensi pada Tabel 7.2.

Tabel 7.2 Tabel Distribusi Frekuensi Catatan Waktu Atlet

Catatan Waktu Atlet (menit:detik)	Frekuensi
27:00–28:59	
29:00–30:59	
31:00–32:59	
33:00–36:59	
37:00–38:59	
39:00–40:59	
41:00–42:59	
43:00–44:59	
45:00–46:59	
47:00–48:59	

- b. Buatlah histogram yang menunjukkan banyaknya atlet yang menyelesaikan lomba Lintas Alam dalam tiap interval catatan waktu. Satu batang untuk interval waktu 31:00–32:59 telah digambar pada histogram di bawah ini.



- c. Interval waktu manakah yang memiliki jumlah atlet paling banyak?
- d. Bentuk dari susunan batang-batang pada histogram menunjukkan distribusi dari data-data yang ada. Distribusi data menunjukkan bagaimana data tersebar, seperti di mana kebanyakan data berada, di mana tidak ditemui data apa pun, dan di mana data sangat sedikit. Apa yang dapat kamu simpulkan dari distribusi data catatan waktu para atlet di atas?

B. Frekuensi Relatif

Frekuensi pada histogram tidak harus selalu menunjukkan banyaknya data yang ada dalam setiap interval. Histogram juga dapat menggunakan persentase sebagai frekuensi relatif dari setiap kelas intervalnya.

Eksplorasi 7.2 Frekuensi Relatif dalam Histogram

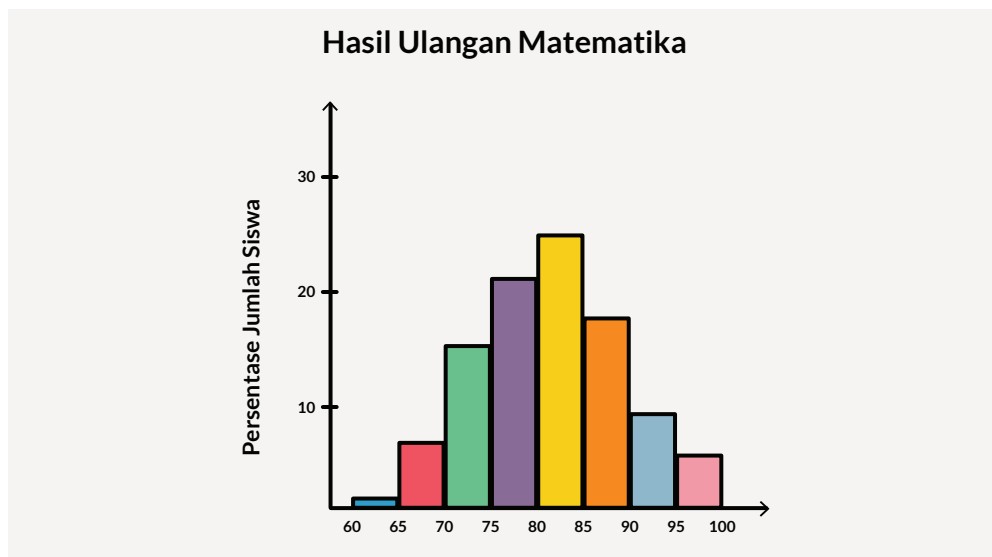


Ayo Bereksplorasi



Ayo Berdiskusi

Ayo berdiskusi dengan menjawab pertanyaan berikut.



Gambar 7.5 Histogram Hasil Nilai Ulangan Matematika

Dari histogram pada Gambar 7.5, ditunjukkan bahwa ada 16% siswa yang mendapatkan nilai matematika antara 70 sampai 75.

- Apakah ini berarti ada 16 siswa yang berada di kelas tersebut? Jelaskan!
- Interval kelas manakah yang memiliki persentase terbesar? Berapa persen kelas dengan interval tersebut?

Misalkan ada 200 siswa yang mengikuti ulangan matematika tersebut. Berapakah banyaknya siswa yang mendapatkan nilai 85 ke atas tapi di bawah 90?

Apabila kalian menambahkan seluruh persen pada setiap interval, berapakah seharusnya jumlah total persen yang kalian peroleh? Jelaskan!

Histogram dengan frekuensi relatif sangat efektif jika digunakan untuk membandingkan dua kelompok data dengan jumlah data yang berbeda, misalnya, jika kalian ingin membandingkan data harian berapa persen penduduk di Jakarta dengan penduduk di Bali yang telah sembuh dari Covid-19. Karena jumlah total penduduk yang terinfeksi Covid-19 di Jakarta berbeda dengan Bali, maka penggunaan persentase sebagai frekuensi relatif memberikan gambaran yang lebih baik.

Latihan 7.2



Ayo Bekerja Sama

Ayo bekerja sama dalam melengkapi tabel di bawah ini agar waktu yang diperlukan menjadi lebih sedikit.

1. Kalian pernah belajar mengenai perkalian dari 0×0 sampai 12×12 . Lengkapilah tabel perkalian di bawah ini.

$\times$	0	1	2	3	4	5	6	7	8	9	10	11	12
0													
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													

Kalian bisa membuat kelompok hasil perkalian di atas dengan mengelompokkan ke dalam kelas-kelas dengan panjang kelas 10. Sebagai contoh, kelas pertama adalah kelas 0-9, kelas kedua: 10-19, kelas ketiga: 20-29, dan seterusnya sampai kelas: 140-149.

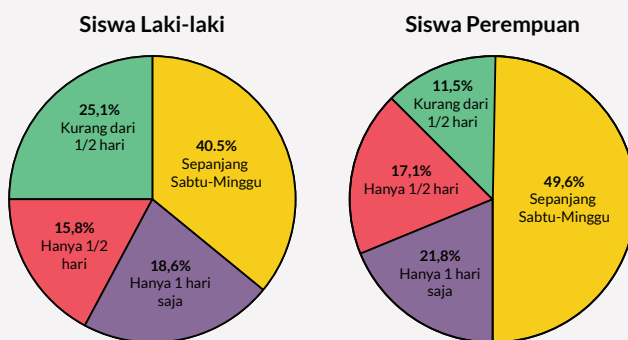
- a. Buatlah tabel frekuensi dengan panjang kelas 10.
 - b. Lalu gambarlah histogramnya.
 - c. Menurut kalian, apakah hasil kali tersebut akan terdistribusi merata ke setiap kelas yang panjang kelasnya 10? Atau apakah ada kelas tertentu yang memiliki hasil kali lebih banyak dari kelas lainnya?
 - d. Sekarang buatlah histogram lainnya dengan panjang kelas 20, dimulai dari 0-19, 20-39, 40-59, dan seterusnya.
 - e. Jelaskanlah persamaan dan perbedaan dari kedua histogram yang kalian hasilkan.
2. Dari sebuah survei terhadap siswa SMP mengenai berapa banyak waktu yang mereka habiskan bersama orang tua mereka di akhir pekan, diperoleh hasil survei sebagai berikut.

Tabel 7.3 Waktu yang Dhabiskan Siswa di Akhir Pekan

Waktu yang Dhabiskan Bersama Keluarga di Akhir Pekan	Siswa Laki-Laki (persen)	Siswa Perempuan (persen)
Sepanjang Sabtu dan Minggu	40,5	49,6
Hanya di salah satu hari saja	18,6	21,8
Hanya $\frac{1}{2}$ hari saja	15,8	17,1
Kurang dari $\frac{1}{2}$ hari	25,1	11,5

Jika kalian ingin membandingkan hasil survei siswa laki-laki dengan hasil survei siswa perempuan, kalian bisa menampilkannya dalam 2 buah diagram lingkaran.

Waktu yang Dhabiskan Bersama Keluarga di Akhir Pekan



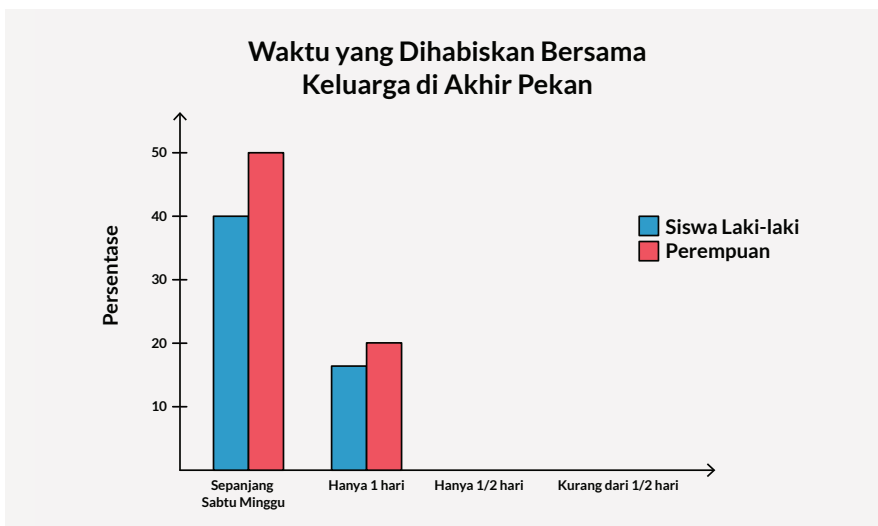
Gambar 7.6 Perbandingan Diagram Lingkaran Siswa Laki-Laki dan Perempuan



Ayo Berpikir Kritis

- Mengapa data yang ditampilkan dalam bentuk persentase?
- Kalian juga bisa menampilkan data-data di atas dalam grafik batang ganda, di mana dalam setiap kategori memiliki 2 batang, yang satu menunjukkan persentase banyaknya siswa laki-laki di kategori tersebut dan yang lainnya menunjukkan persentase siswa perempuan.

Lengkapilah diagram batang berikut.



Gambar 7.7 Diagram Batang Ganda Waktu Akhir Pekan Siswa



Ayo Bernalar

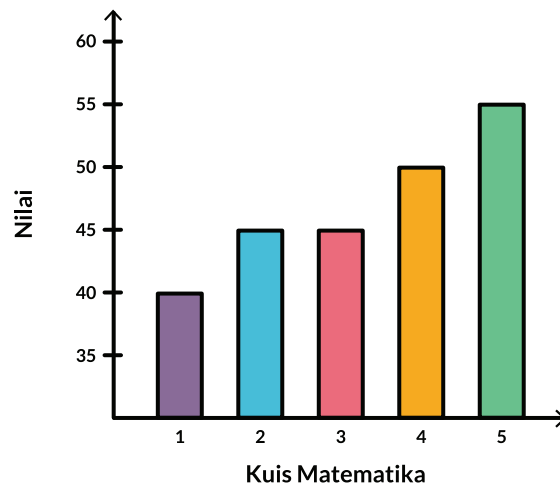
Dengan memilih representasi grafik yang tepat, akan memudahkan kita dalam membandingkan 2 kelompok data.

- c. Menurut kalian, diagram manakah yang lebih mudah digunakan untuk membandingkan 2 kelompok data? Berikan alasan dari pilihanmu.
3. Dani sering bermain *games online* sehingga nilai kuis matematikanya jelek. Orang tua Dani melarang Dani untuk bermain *games online* sampai hasil nilai kuis matematika Dani berubah secara signifikan. Guru matematika Dani setiap minggu memberikan kuis matematika dengan nilai tertinggi 100. Dani membuat grafik batang untuk menunjukkan kepada orang tuanya bahwa nilai kuisnya sudah membaik dalam 5 minggu terakhir.



Ayo Berpikir Kritis

- a. Panjang batang nilai kuis 5 Dani tiga kali lebih tinggi dari panjang batang nilai kuis 1-nya. Apakah nilai kuis 5-nya tiga kali dari nilai kuis 1-nya?
- b. Orang tua Dani mengatakan bahwa grafik batang yang dibuat Dani menyesatkan karena dari grafik ini terlihat ada perbaikan signifikan dari nilai kuis Dani dibandingkan dengan kenyataannya. Hal manakah pada grafik ini yang menyebabkan grafik ini memberikan kesimpulan yang salah?



Gambar 7.8 Diagram Batang Buatan Dani



Penguatan Karakter

- c. Buatlah diagram batang yang baru yang dapat memberikan hasil yang lebih akurat dan menggambarkan performa Dani yang sebenarnya di kuis matematika mingguan.



Ayo Berefleksi

Dalam subbab ini, kalian sudah belajar mengenai diagram batang dan histogram. Selain itu, kalian juga telah mengenal frekuensi dan frekuensi relatif dalam diagram batang dan histogram.

- Apa saja perbedaan diagram batang dengan histogram?
- Kapan kita sebaiknya menggunakan frekuensi relatif daripada frekuensi?

C. Ukuran Pemusatan

1. Modus dan Median

Modus dan median adalah dua ukuran pemusatan untuk melihat kecenderungan kumpulan data.

Median adalah nilai data yang berada tepat di tengah ketika seluruh data diurutkan dari yang terkecil sampai yang terbesar. Untuk mencari letak median, bagilah banyaknya data dengan 2.

- Jika hasilnya adalah bilangan bulat, m , maka median terletak di tengah-tengah antara urutan ke- m dan ke- $(m + 1)$.
- Jika hasil baginya bukan merupakan bilangan bulat, bulatkanlah hasilnya ke atas, maka median terletak di urutan sesuai hasil pembulatan.

Modus dari sebuah kumpulan data adalah data yang paling sering muncul atau memiliki frekuensi paling besar. Kedua ukuran pemusatan ini memiliki keuntungan, yaitu tidak terpengaruh jika kumpulan data memiliki data pencilan atau data yang berbeda dari kumpulan datanya.

Selain modus dan median, kalian bisa melihat rentang dari kumpulan data melalui *range* atau jangkauan. **Jangkauan** adalah selisih antara data terkecil dengan data terbesar.

Eksplorasi 7.3 *Line Plot*



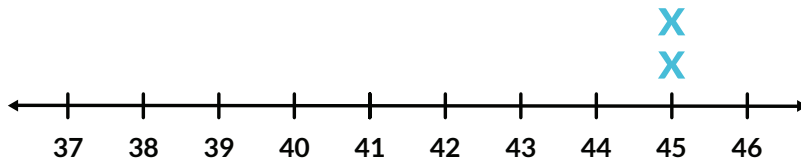
Ayo Bereksplorasi

Basket merupakan olahraga yang digandrungi banyak siswa SMA/MA, khususnya pria. Untuk dapat bermain basket, kalian perlu menggunakan sepatu olahraga. Berikut adalah data penjualan sepatu olahraga di toko A yang terdiri dari beberapa merek dan ukuran pada akhir pekan pertama bulan Januari.

Tabel 7.4 Data Penjualan Sepatu di Toko A

No	Merek	Ukuran	No	Merek	Ukuran
1	A	43	14	A	44
2	B	44	15	C	40
3	C	38	16	D	41
4	A	43	17	B	42
5	C	44	18	D	43
6	D	42	19	E	42
7	A	42	20	A	40
8	A	39	21	A	45
9	B	43	22	C	41
10	E	43	23	A	41
11	C	44	24	A	42
12	E	45	25	C	43
13	B	44			

- Buatlah diagram *line plot* untuk menunjukkan ukuran sepatu yang terjual pada akhir pekan pertama bulan Januari. Diagram line plot adalah sebuah garis bilangan dengan banyaknya tanda X yang menunjukkan banyaknya data yang muncul dengan nilai tertentu. Sebagai contoh, 45 muncul tiga kali. Jadi, kalian tuliskan tanda X di atas angka 45.



Gambar 7.9 *Line Plot* Ukuran Sepatu

- b. Jelaskan bentuk dari line plot yang kamu hasilkan. Bagaimana bentuk line plot ini bisa menjelaskan distribusi data ukuran sepatu di atas?

Ketika kalian mendeskripsikan sebuah kumpulan data, biasanya kalian juga perlu menentukan data terkecil dan data terbesar dari kumpulan data tersebut.

- c. Tentukanlah data terkecil dan data terbesar dari kumpulan data ukuran sepatu yang terjual.
- d. Bagaimana kalian dapat menemukan data terkecil dan data terbesar dengan melihat diagram *line plot*?
- e. Tentukanlah jangkauan dari data ukuran sepatu pada Tabel 7.4.
- f. Tentukanlah modus dari data ukuran sepatu pada Tabel 7.4.
- g. Bagaimana kalian bisa menentukan modus sekumpulan data dari diagram *line plot*?
- h. Urutkanlah data ukuran sepatu di atas dari yang terkecil sampai yang terbesar, lalu tentukanlah mediannya.
- i. Bagaimana kalian dapat menentukan median dari sekumpulan data dengan melihat diagram *line plot*?



Ayo Berdiskusi

Jika terjadi penambahan data baru, bagaimana modus, median, dan jangkauan akan terpengaruh?

- j. Ternyata ada data penjualan di toko sepatu A yang tertinggal. Data-data tersebut adalah 41, 43, 44, 44, dan 46. Berapakah nilai dari jangkauannya sekarang? Berapakah modusnya sekarang?



Ayo Berdiskusi

Bagaimana perbedaan mencari median pada kelompok dengan banyaknya data ganjil dan genap?

- k. Ketika banyaknya data adalah bilangan genap, maka tidak ada data yang diambil sebagai median tunggal. Dalam hal ini, median diambil dari nilai tengah di antara dua nilai data yang berada di tengah. Carilah median dari kumpulan data yang baru.

2. Mean (Rerata atau Rata-Rata)

Rerata atau *mean* adalah ukuran pemusatan lain selain median dan modus. *Mean* dari sebuah kumpulan data adalah bilangan yang diperoleh dengan mendistribusikan secara merata ke seluruh anggota dari kumpulan data. Kalian bisa menghitung *mean* dengan cara menambahkan seluruh nilai data dan membagi dengan total banyaknya data.

Atau jika ditulis dalam bentuk formula: $\bar{x} = \frac{\sum x}{n}$ di mana:

$\bar{x}$ adalah *mean*, dibaca x bar. $\sum x$ menyatakan jumlah total data dan n menunjukkan banyaknya data.

Eksplorasi 7.4 Mean



Ayo Bereksplorasi

OSIS Sekolah A yang beranggotakan 10 orang akan melakukan aksi sosial untuk membantu para korban bencana alam. Mereka sepakat untuk mengumpulkan pakaian bekas layak pakai untuk membantu para korban bencana alam. Adapun jumlah baju yang dikumpulkan setiap pengurus OSIS adalah sebagai berikut.

3 5 7 10 5 3 4 6 9 8

- a. Tentukanlah nilai *mean*, median, dan modus dari jumlah baju yang dikumpulkan oleh para pengurus tersebut.



Ayo Berdiskusi

Bagaimana penambahan data berpengaruh terhadap ukuran pemusatan?

- b. Keesokan harinya, ada dua siswa yang bukan pengurus OSIS, namun mereka terinspirasi dengan aksi sosial yang dilakukan oleh para pengurus OSIS. Mereka langsung ikut menyumbangkan baju layak pakai sebanyak 20 dan 22 buah. Tentukan *mean*, median, dan modus dari kumpulan data yang baru.

3. Penggunaan Ukuran Pemusatan

Setelah kalian mempelajari cara menentukan *mean*, median, dan modus, maka hal yang juga penting adalah mengetahui karakteristik dari setiap ukuran pemusatan ini, agar kita dapat memilih ukuran pemusatan mana yang paling tepat sesuai dengan konteks permasalahan.

Eksplorasi 7.5



Ayo Bereksplorasi

Masih dari kisah para pengurus OSIS Sekolah A sebelumnya. Bagaimana hasil pengamatan kalian setelah membandingkan *mean*, median, dan modus data sumbangan 10 pengurus OSIS dengan *mean*, modus dan median data sumbangan ke-12 siswa?

Di antara *mean*, median, dan modus, manakah nilai yang tetap? Manakah nilai yang berubah? Jelaskan!



Ayo Berdiskusi

Cobalah berpikir ekstrem dengan mengganti 1 data dengan nilai yang sangat berbeda, lalu amati perubahannya.

Bagaimana jika seandainya siswa ke-12 bukan menyumbang 22 buah, namun menyumbang 100 pakaian. Menurut kalian, tanpa menghitung dulu *mean*, median, dan modus yang baru, manakah di antara *mean*, median, dan modus yang nilainya berubah? Manakah yang nilainya tetap? Jelaskan!

Sekarang cobalah kalian menghitung *mean*, median, dan modus yang baru. Apakah analisis kalian di atas sudah benar?

Modus	Digunakan ketika jenis data adalah data kualitatif, atau jenis data kuantitatif yang memiliki 1 modus atau 2 modus (bimodal).
Median	Digunakan untuk jenis data kuantitatif. Biasanya median digunakan ketika ada data yang memiliki nilai yang ekstrem (pencilan), sehingga data ekstrem tersebut tidak memiliki dampak yang besar seperti pada <i>mean</i> .
Mean	Digunakan untuk jenis data kuantitatif dan menggunakan seluruh data. Namun, mean terpengaruh oleh data dengan nilai yang ekstrem.

a. Mean/Rata-Rata Data Kelompok

Data penjualan sepatu di toko A pada Tabel 7.4 merupakan kumpulan data tunggal. Kalian dapat mengelompokkan data-data ini menjadi data kelompok dengan panjang kelas sama dengan 2 sehingga menjadi tabel frekuensi data kelompok sebagai berikut.

Tabel 7.5 Distribusi Frekuensi Data Kelompok Penjualan Sepatu di Toko A

Ukuran	37-39	40-42	43-45	46-48
Frekuensi	2	11	16	1

Cara menghitung rata-rata dari data kelompok di atas adalah menggunakan nilai tengah dari tiap kelompok. Data tunggal dalam kelompok diasumsikan tersebar secara merata, sehingga nilai tengah dari setiap kelompok dapat diasumsikan mewakili kelompok tersebut.

Nilai tengah kelompok 37-39 adalah 38, Nilai tengah kelompok 40-42 adalah 41, Nilai tengah kelompok 43-45 adalah 44, dan Nilai tengah kelompok 46-48 adalah 47.

Rata-rata dari kelompok di atas:

$$\begin{aligned}
 \bar{x} &= \frac{\sum x}{n} = \frac{(2 \times 38) + (11 \times 41) + (16 \times 44) + (1 \times 47)}{2 + 11 + 16 + 1} \\
 &= \frac{1278}{30} = 42,6
 \end{aligned}$$



Ayo Berdiskusi

Bagaimana rata-rata data tunggal dibandingkan dengan rata-rata data kelompok? Apakah masih bisa merepresentasikan kelompok data?

Bandingkanlah hasil rata-rata data kelompok ini dengan hasil rata-rata data tunggal dari penjualan sepatu di toko A. Apakah menurut kalian, kedua hasil rata-rata masih cukup dekat?

Latihan 7.3

1. Jika data penjualan sepatu di toko A pada Tabel 7.4 kita ubah menjadi tabel Frekuensi data tunggal sebagai berikut:

Tabel 7.6 Tabel Frekuensi Data Tunggal Penjualan Sepatu di Toko A

Ukuran	38	39	40	41	42	43	44	45	46
Frekuensi	1	1	2	4	5	7	7	2	1

- a. Tentukanlah modus, median, dan *mean* dari kumpulan data di atas.



Problem Solving

Mengambil keputusan berdasarkan data.

- b. Untuk menentukan rencana pemesanan sepatu bulan depan, jelaskan mengapa pemilik toko sebaiknya menggunakan modus.
2. Data berikut menunjukkan jumlah kue yang dijual melalui situs online setiap harinya:

0 3 2 7 4 2 3 0 4 0 6 5 5 2 4 0

- a. Tentukanlah modus dan median dari data di atas.
- b. Menurutmu, ukuran pemusatan manakah yang lebih untuk data di atas, modus atau median? Jelaskan!



Ayo Berpikir Kreatif

3. Buatlah kumpulan data dengan banyaknya data, ada sebanyak 13 buah dan memenuhi kondisi:
 - Data terkecil = 3
 - Data terbesar = 13
 - Modus = 4, dan
 - Median = 8
4. Dari 2 kelas siswa SD di sekolah “Pancasila” diperoleh data tinggi siswa (dalam cm) sebagai berikut:
 - Kelas A: 117, 117, 119, 122, 127, 127, 114, 137, 99, 107, 114, 127, 122, 114, 120, 125, 119
 - Kelas B: 130, 147, 137, 142, 140, 135, 135, 142, 142, 137, 135, 132, 135, 120, 119, 125, 142
 - a. Untuk masing-masing kelas, buatlah grafik *line plot*.
 - b. Tentukanlah *range*, modus, dan median dari setiap kelas.
 - c. Kedua kelas berasal dari tingkat yang berbeda. Kelas manakah menurut kalian yang memiliki tingkat yang lebih tinggi?
 - d. Berapa persen siswa dari kelas B yang memiliki tinggi sama atau lebih tinggi dari median tinggi badan siswa kelas A?
5. Pernahkah kalian mendengar bahwa Indonesia adalah salah satu paru-paru dunia? Hutan tropis di Indonesia memiliki peranan yang sangat penting untuk memberikan sumbangan terhadap lingkungan dunia. Pohon Borneo adalah salah satu jenis pohon yang banyak ditemukan di hutan Kalimantan.

Tabel 7.7 Tabel Frekuensi Data Kelompok Diameter Pohon Borneo di Daerah A

Diameter Pohon Borneo (cm)	19-21	22-24	25-27	28-30
Frekuensi	4	17	25	14

- a. Tentukanlah kelas modus.
- b. Prediksi nilai *mean* dari data kelompok di atas.
- c. Tentukan kelas median.



Ayo Berefleksi

Dalam subbab ini, kalian sudah belajar mengenai ukuran pemusatan: mean, median, dan modus. Kalian juga telah menentukan manakah ukuran pemusatan yang sesuai.

- Ukuran pemusatan manakah yang terpengaruh dengan pencilan? Manakah yang tidak terpengaruh pencilan?
- Saat data tunggal dikelompokkan, lalu kalian menghitung mean data tunggal dan mean data kelompok, bagaimana hasil dari kedua mean tersebut? Apakah berbeda jauh atau berbeda sedikit?

b. Median dan Kelas Modus Data Kelompok

Eksplorasi 7.6 Membandingkan Modus dan Median Data Tunggal dengan Data Kelompok



Ayo Bereksplorasi

Kita masih akan menggunakan data penjualan sepatu di toko A pada Tabel 7.5 yang merupakan Tabel Distribusi Data Kelompok.

Sekarang, mari kita bandingkan modus. Pada data tunggal, kelompok data ini memiliki dua modus atau disebut bimodal, yaitu 43 dan 44 karena kedua data tersebut memiliki frekuensi yang paling tinggi yaitu 7.

Pada data kelompok, kita dapat melihat bahwa kelas modus adalah kelas 43-45 yaitu dengan frekuensi 16. Jadi, walaupun data tunggal diubah ke dalam data kelompok, ternyata kelas modus tetap dapat memberikan gambaran estimasi di mana data modus berada.

Bagaimana dengan median?

Untuk data tunggal, karena jumlah data ada sebanyak 30 data, maka karena 30 dibagi 2 adalah 15, sehingga median terletak di antara data ke-15 dan data ke-16. Data yang terletak di urutan ke-15 adalah 43 dan data di urutan ke-16 adalah 43. Maka median dari kelompok data tunggal adalah $\frac{43 + 43}{2} = 43$.

Untuk mencari median dari data kelompok, kita akan menggunakan interpolasi. Bagaimana interpolasi bekerja? Pertama, tentukan dahulu kelas median. Karena

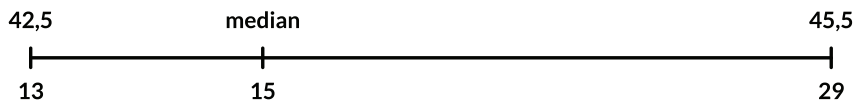
jumlah data sebanyak 30, maka data median berada di urutan ke $\frac{1}{2} \times 30 = 15$. Data ke-15 berada di kelas 43-45.

Tepi bawah kelas 43-45 adalah 42,5 dan tepi atasnya adalah 45,5. Setelah itu kalian perlu menentukan banyaknya data yang nilainya di bawah 42,5 dan 45,5.

Banyaknya data yang nilainya di bawah 42,5 yaitu banyaknya data di kelas 37-39 dan kelas 40-42 yaitu ada sebanyak $2 + 11 = 13$.

Banyaknya data yang nilainya di bawah 45,5 yaitu banyaknya data di kelas 37-39, kelas 40-42 dan kelas 43-45 yaitu ada sebanyak $2 + 11 + 16 = 29$.

Semua data yang diperoleh, diletakkan dalam garis bilangan berikut:



Bilangan di atas garis merupakan tepi bawah dan tepi atas dari kelas median. Bilangan di bawah garis merupakan banyaknya data yang terletak di bawah 42,5, di bawah urutan median, dan di bawah 45,5.

Lalu, kalian tinggal membandingkan selisih dari bilangan-bilangan yang ada pada garis bilangan tersebut:

$$\frac{\text{Median} - 42,5}{15 - 13} = \frac{45,5 - 42,5}{29 - 13}$$

$$\frac{\text{Median} - 42,5}{2} = \frac{3}{16}$$

$$\text{Median} - 42,5 = \frac{6}{16} \approx 0,375$$

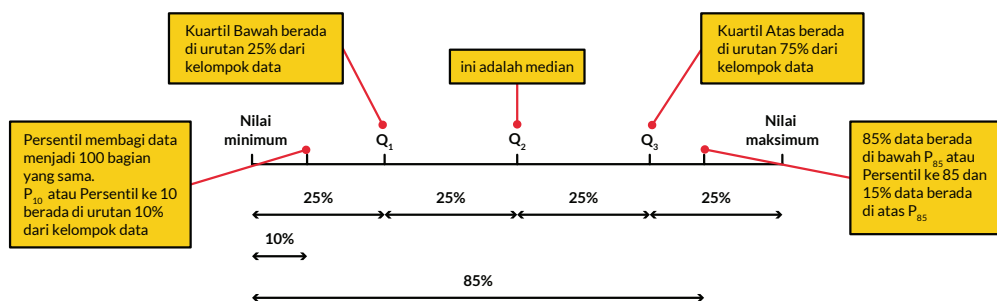
$$\text{Median} = 42,5 + 0,375 = 42,875$$

Ternyata median dari data berkelompok, yaitu 42,875 tidak jauh berbeda dengan median dari data tunggal, yaitu 43. Jadi, walaupun data dikelompokkan, median data kelompok dapat tetap mewakili median dari data tunggal.

D. Ukuran Penempatan (*Measure of Location*)

1. Kuartil Data Tunggal

Sebelumnya kalian telah mempelajari mengenai median. Median membagi kumpulan data yang telah diurutkan menjadi 2 sama besar (50%). Kalian bisa menentukan ukuran penempatan lainnya seperti kuartil dan persentil.



Gambar 7.10 Letak Kuartil dan Persentil dalam Kelompok Data

Serupa dengan mencari letak median, maka untuk mencari letak kuartil bawah atau Q_1 , bagilah banyaknya data dengan 4.

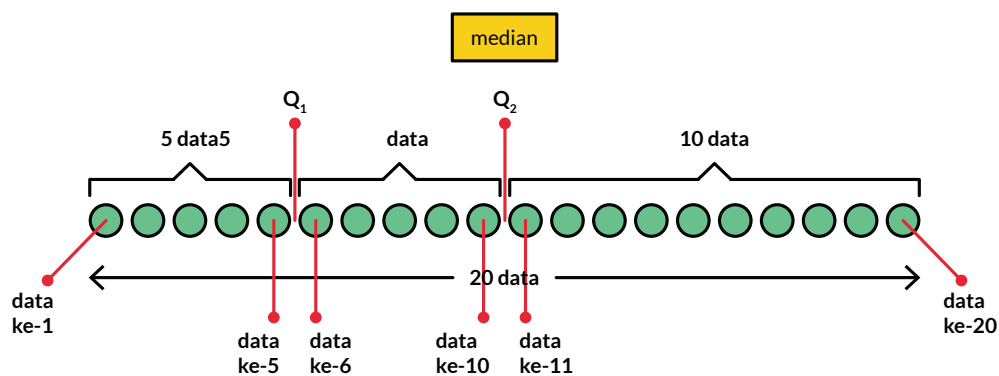
- Jika hasilnya adalah bilangan bulat, m , maka Q_1 terletak di tengah-tengah antara urutan ke- m dan ke- $(m + 1)$.
- Tetapi jika hasil baginya bukan merupakan bilangan bulat, bulatkanlah hasilnya ke atas, maka Q_2 terletak di urutan sesuai hasil pembulatan.

Misalkan jika banyaknya data ada 20 buah, di manakah letak median? Di manakah letak Q_1 ?

Untuk median, 20 dibagi 2 = 10, maka median terletak di antara data urutan ke-10 dan ke-11.

Untuk Q_1 , 20 dibagi 4 = 5, maka Q_1 terletak di antara data urutan ke-5 dan ke-6.

Agar lebih jelas, kalian dapat melihat ilustrasi berikut.



Gambar 7.11 Letak Q_1 dan Q_2 dalam Kelompok Data $n = 20$

Dari ilustrasi di atas, kalian bisa melihat bahwa Median = Q_2 , yaitu membagi kumpulan data menjadi 2 sama besar yaitu, 10 data di sebelah kiri dan 10 data di sebelah kanan.

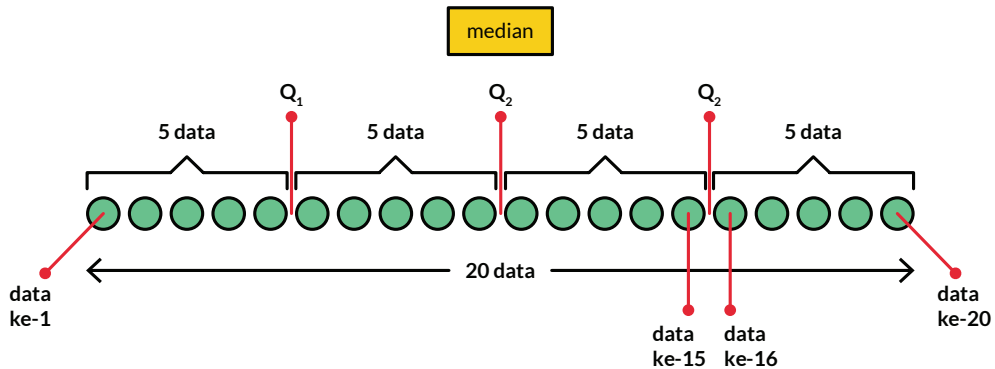
Sedangkan Q_1 membagi dua ke-10 data yang berada di sebelah kiri menjadi masing-masing sebanyak 5 data.



Ayo Mencoba

Nah, bisakah kalian mencari di mana letak Q_3 atau kuartil atas agar dia membagi ke-10 data di sebelah kanan Q_2 sama banyak?

Benar sekali, Q_3 terletak di antara data ke-15 dan ke-16.



Gambar 7.12 Letak Kuartil dalam Kelompok Data $n = 20$

Dari ilustrasi di atas, kalian dapat melihat bahwa Q_1 , Q_2 dan Q_3 membagi kumpulan data menjadi 4 bagian yang sama besar, yaitu masing-masing terdiri dari 5 data.

Atau dapat dikatakan bahwa di antara Q_1 dan Q_2 terdapat 25% data. Demikian juga di antara Q_2 dan Q_3 terdapat 25% data.

Bisakah kalian menentukan rumus untuk mencari Q_3 ? Coba pikirkan dulu sejenak.

Serupa dengan mencari letak median, maka untuk mencari letak kuartil bawah atau Q_1 , bagilah banyaknya data dengan 4.

- Jika hasilnya adalah bilangan bulat, m , maka Q_1 terletak di tengah-tengah antara urutan ke- m dan ke- $(m + 1)$.
- Tapi jika hasil baginya bukan merupakan bilangan bulat, bulatkanlah hasilnya ke atas, maka Q_2 terletak di urutan sesuai hasil pembulatan.

Misalkan jika banyaknya data ada sebanyak 20 buah, di manakah letak median? Di manakah letak Q_1 ?

Benar, serupa dengan mencari letak median dan Q_1 , maka untuk mencari letak kuartil bawah atau Q_3 adalah dengan mengalikan banyaknya data dengan $\frac{3}{4}$.

- Jika hasilnya adalah bilangan bulat m , maka Q_3 terletak di tengah-tengah antara urutan ke- m dan ke- $(m + 1)$.
- Tapi jika hasil baginya bukan merupakan bilangan bulat, bulatkanlah hasilnya ke atas, maka Q_3 terletak di urutan sesuai hasil pembulatan.

Jika banyaknya data ada 20 buah, maka $\frac{3}{4} \times 20 = 15$. Karena 15 merupakan bilangan bulat, maka letak Q_3 ada di antara data ke-15 dan ke-16.

Bandingkan hasilnya dengan ilustrasi di atas. Apakah sama atau berbeda?



Ayo Mencoba

Carilah Q_1 , Q_2 dan Q_3 dari data penjualan sepatu pada Tabel 7.6.

2. Kuartil Data Kelompok

Sama seperti menentukan median (Q_2) dalam data kelompok, menentukan Q_1 dan Q_3 juga menggunakan cara yang sama, yaitu dengan cara interpolasi.

Dalam data kelompok, letak Q_1 , Q_2 dan Q_3 adalah sebagai berikut:

Kelompok data ditampilkan dalam tabel frekuensi kumulatif, lalu letak kuartil adalah sebagai berikut:

- Q_1 = data ke $\frac{1}{4}$ dari total data
- Q_2 = data ke $\frac{1}{2}$ dari total data
- Q_3 = data ke $\frac{3}{4}$ dari total data

Mari kita gunakan contoh penjualan sepatu di toko A pada Tabel 7.5.

Karena total data ada sebanyak 30 buah, maka letak Q_1 ada di data ke $\frac{1}{4} \times 30 = 7,5$

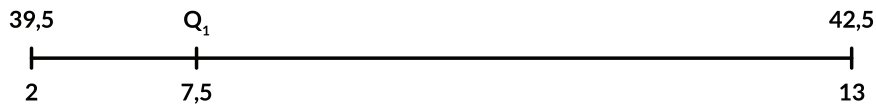
Pada tabel, data ke 7,5 terletak pada kelas 40-42. Masihkah kalian ingat metode interpolasi?

Tepi bawah kelas 40-42 adalah 39,5 dan tepi atas kelas 40-42 adalah 42,5.

Banyaknya data yang berada sebelum 39,5 ada sebanyak 2 buah.

Banyaknya data yang berada sebelum 42,5 ada sebanyak 13 buah.

Tempatkan angka-angka tersebut dalam garis bilangan sebagai berikut.



$$\frac{Q_1 - 39,5}{7,5 - 2} = \frac{42,5 - 39,5}{13 - 2}$$

$$\frac{Q_1 - 39,5}{5,5} = \frac{3}{11}$$

$$Q_1 - 39,5 = \frac{3}{11} \times 5,5$$

$$Q_1 - 39,5 = 1,5$$

$$Q_1 = 41$$

$Q_1 = 41$, artinya 25% dari sepatu yang terjual memiliki ukuran lebih kecil sama dengan 41, atau ukuran 38, 39, 40, dan 41.

Sebanyak 75% sepatu yang terjual merupakan sepatu dengan ukuran di atas 41.



Ayo Mencoba

Carilah Q_3 dari data berkelompok penjualan sepatu di toko A pada Tabel 7.5. Bandingkanlah hasil Q_1 dan Q_3 dari data berkelompok dengan Q_1 dan Q_3 dari data tunggal. Jelaskan!

Jadi, saat menghitung kuartil, pastikan terlebih dahulu apakah data yang kalian akan hitung adalah data tunggal atau data kelompok agar metode yang dipilih lebih tepat.

3. Persentil Data Kelompok

Sebelumnya kalian telah mempelajari bahwa kuartil membagi data menjadi 4 bagian sama besar. Ukuran penempatan yang lain adalah persentil. Hanya saja persentil membagi data menjadi 100 bagian sama besar. Persentil ke-10 ditulis dengan simbol P_{10} artinya sebelum P_{10} terdapat 10% data dan sesudah P_{10} terdapat 90% data.

Cara menentukan persentil dalam data kelompok, sama dengan cara menentukan kuartil dalam data kelompok, yaitu dengan cara interpolasi.

Kelompok data ditampilkan dalam tabel frekuensi kumulatif, lalu letak persentil adalah sebagai berikut:

- P_{10} = data ke $\frac{10}{100}$ dari total data
- P_{85} = data ke $\frac{85}{100}$ dari total data

Eksplorasi 7.7



Ayo Bereksplorasi

Mari kita lihat data berikut. Data berikut menampilkan lamanya waktu yang diperlukan ketika seseorang mengurus KTP di kelurahan M selama 1 minggu.

Waktu yang diperlukan, t (menit)	20–29	30–39	40–49	50–59	60–69
Frekuensi	6	10	18	13	2

- a. Hitunglah P_{65} .



Ayo Berpikir Kritis

Ketika kita melihat pengumuman atau klaim dari seseorang atau siapa pun, ada baiknya kita mempertanyakan dasar dari klaim atau pengumuman tersebut, tidak begitu saja menerimanya tanpa data pendukung. Kehati-hatian ini menjadi bekal yang sangat penting saat menghadapi berbagai masalah.

- b. Di papan pengumuman kantor kelurahan tertulis poster sebagai berikut:

Untuk pengurusan KTP
Hanya 10% dari warga yang perlu menunggu lebih dari 56 menit

Dengan menghitung persentil yang sesuai, berikan komentarmu tentang benar atau tidaknya isi dari poster tersebut.

Solusi

- a. Karena data di atas merupakan data kelompok, maka kita akan menggunakan interpolasi untuk menemukan persentil ke-65.

Pertama kita tentukan dulu letak P_{65} . Total frekuensi ada sebanyak 49 buah.

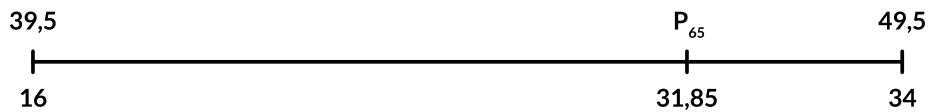
- P_{65} terletak pada data ke $\frac{65}{100} \times 49 = 31,85$
- Jadi P_{65} ada di kelas dengan interval 40–49.

Tepi bawah kelas 40–49 adalah 39,5 dan tepi atas kelas 40–49 adalah 49,5.

Banyaknya data sebelum 39,5 ada sebanyak 16 data.

Banyaknya data sebelum 49,5 ada sebanyak 34 data.

Tempatkan angka-angka tersebut dalam garis bilangan sebagai berikut:



Lalu kita gunakan interpolasi:

$$\frac{P_{65} - 39,5}{31,85 - 16} = \frac{49,5 - 39,5}{34 - 16}$$

$$\frac{P_{65} - 39,5}{15,85} = \frac{10}{18}$$

$$P_{65} - 39,5 = \frac{5}{9} \times 15,85$$

$$P_{65} - 39,5 = 8,81$$

$$P_{65} = 48,31$$

$P_{65} = 48,31$ artinya 65% warga menunggu kurang dari 48,31 menit atau 35% warga menunggu lebih dari 46,31 menit.

- b. Untuk menentukan apakah isi poster tersebut benar atau tidak, maka kalian perlu mencari Persentil ke-90.

Pertama kita tentukan dulu letak P_{90} .

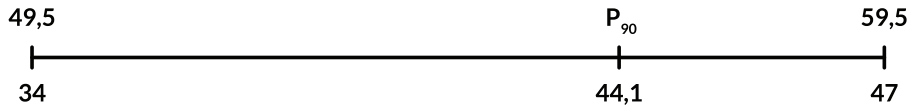
- P_{90} terletak pada data ke $\frac{90}{100} \times 49 = 44,1$
- Data ke-44,1 terletak pada kelas 50–59
- Jadi P_{90} ada di kelas 50–59.

Kelas 50–59 memiliki tepi bawah = 49,5 dan tepi atas = 59,5.

Banyaknya data sebelum 49,5 ada sebanyak 34 data.

Banyaknya data sebelum 59,5 ada sebanyak 47 data.

Tempatkan angka-angka tersebut dalam garis bilangan sebagai berikut:



Lalu, kita kembali menggunakan interpolasi:

$$\frac{P_{90} - 49,5}{44,1 - 34} = \frac{59,5 - 49,5}{47 - 34}$$

$$\frac{P_{90} - 49,5}{10,1} = \frac{10}{13}$$

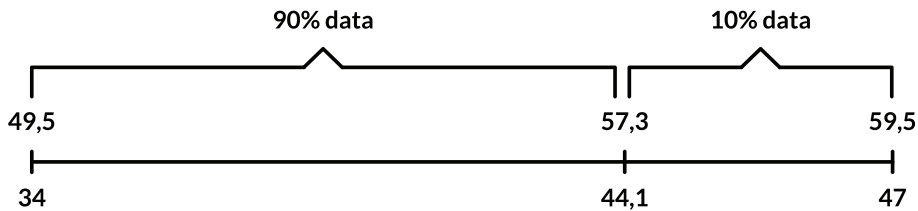
$$P_{90} - 49,5 = \frac{10}{13} \times 10,1$$

$$P_{90} - 49,5 = 7,8$$

$$P_{90} = 49,5 + 7,8$$

$$P_{90} = 57,3$$

Interpretasi dari hasil $P_{90} = 57,3$ artinya 90% warga menunggu pengurusan KTP sampai 57,3 menit dan ada 10% warga yang menunggu lebih dari 57,3 menit. Jadi isi poster yang menyebutkan bahwa hanya 10% warga yang menunggu lebih dari 56 menit tidak tepat. Karena pasti lebih dari 10% warga yang menunggu lebih dari 56 menit.



Alternatif lain, kalian bisa mencari banyaknya warga yang menunggu lebih dari 56 menit.

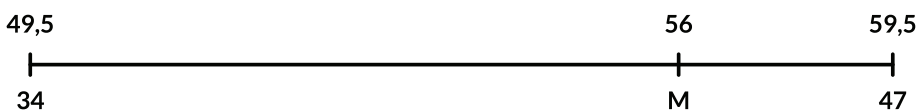
Data 56 menit berada di kelas 50–59.

Tepi bawah dan tepi atas kelas 50–59 adalah 49,5 dan 59,5.

Banyaknya data sebelum 49,5 ada sebanyak 34 data.

Banyaknya data sebelum 59,5 ada sebanyak 47 data.

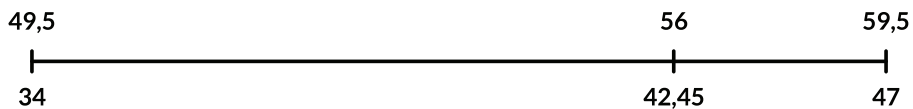
Tempatkan angka-angka tersebut dalam garis bilangan sebagai berikut:



Sama seperti sebelumnya, kita kembali menggunakan metode interpolasi:

$$\begin{aligned}\frac{56 - 49,5}{M - 34} &= \frac{59,5 - 49,5}{47 - 34} \\ \frac{6,5}{M - 34} &= \frac{10}{13} \\ (13)(6,5) &= 10(M - 34) \\ 84,5 &= 10(M - 34) \\ M - 34 &= 8,45 \\ M &= 34 + 8,45 \\ M &= 42,45\end{aligned}$$

Setelah mendapatkan hasil di atas kita coba interpretasikan hasilnya.



Total data = 49.

42,45 warga menunggu pengurusan KTP kurang dari 56 menit.

Banyaknya warga menunggu pengurusan KTP lebih dari 56 menit =

$$49 - 42,45 = 6,55$$

Persentase warga yang menunggu lebih dari 56 menit =

$$\frac{6,55}{49} \times 100\% = 13,4\%$$

Jadi, klaim kantor kelurahan bahwa hanya 10% warga yang menunggu lebih dari 56 menit tidak benar.

Latihan 7.4

1. Secara geografis Indonesia dilalui oleh garis khatulistiwa sehingga hanya terdapat 2 musim, yaitu musim panas dan musim hujan. Oleh karena itu, sangat penting bagi kita untuk memahami faktor penting apa saja yang ada pada kedua musim tersebut agar kita dapat mempersiapkan diri dengan lebih baik dalam usaha hidup berdamai dengan alam. Salah satu faktor yang penting yang menentukan musim adalah faktor curah hujan.

Curah hujan adalah jumlah air hujan yang turun pada suatu daerah dalam kurun waktu tertentu. Dengan kata lain, curah hujan adalah volume air hujan yang terkumpul dalam bidang datar dalam periode tertentu.

Biasanya curah hujan dinyatakan dalam satuan milimeter. Data curah hujan yang ditampilkan adalah ketinggian air hujan yang terkumpul di tempat datar seluas 1 meter persegi. Jadi, jika curah hujan sebesar 1 mm artinya volume air hujan yang terkumpul pada tempat datar seluas 1 meter persegi ada sebanyak 1 liter. Pada umumnya curah hujan dikategorikan menjadi 3 kategori, yaitu rendah (0-100 mm), menengah (100-300 mm) dan tinggi (300-500 mm).

Perhatikan data curah hujan di Kota Samarinda sepanjang tahun 2017 berikut (dalam mm, dibulatkan ke satuan terdekat):

161 139 88 343 309 421 161 250 100 152 219 223

Sumber: <https://samarindakota.bps.go.id>

- a. Tentukanlah median dari data tersebut.
 - b. Tentukanlah Q_1 dan Q_3 dari data tersebut. Apakah kalian perlu melakukan interpolasi?
2. Menjelang Hari Raya Kurban, biasanya para peternak sapi mempersiapkan sapi-sapi yang akan dijual. Berikut data berat 31 ekor sapi yang akan dijual oleh peternak.

Berat sapi (kg)	300-349	350-399	400-449	450-499	500-549
Frekuensi	3	6	10	7	5

- a. Tentukanlah estimasi dari median berat sapi di atas.
 - b. Carilah Q_1 . Apakah kalian perlu melakukan interpolasi?
 - c. Carilah Q_3 .
 - d. Interpretasikanlah hasil Q_3 yang kamu dapatkan di bagian c.
 - e. Carilah P_{10} , lalu interpretasikan hasilnya.
3. Indonesia adalah negara yang kaya dan terkenal dengan faunanya yang beraneka ragam. Bahkan, banyak hewan yang hanya terdapat di Indonesia karena keunikan kondisi alamnya. Karena itulah kita harus melestarikan dan memperhatikan hewan langka yang masih tersisa agar kelak generasi selanjutnya tetap dapat menyaksikan kelangsungan hidup hewan langka ini. Salah satu contoh hewan langka adalah burung elang jawa (*Nisaetus bartelsi*). Jumlahnya saat ini diperkirakan hanya tinggal sekitar 300-500 ekor saja.

Tabel di bawah ini menunjukkan panjang bentang sayap elang jawa dalam meter yang berhasil dikumpulkan oleh para peneliti lingkungan.

Panjang bentang sayap (cm)	166–170	171–175	176–180	181–185	Lebih dari 186
Frekuensi	4	20	37	28	11

- Tentukanlah Q_1 dan Q_3 .
- Tentukanlah persentil ke-80 dan interpretasikanlah hasilnya.
- Jelaskan mengapa tidak mungkin bisa menemukan persentil ke-90.



Ayo Berefleksi

Dalam subbab ini, kalian sudah belajar mengenai ukuran lokasi: kuartil dan persentil.

- Kuartil berapakah yang sama dengan median?
- Ada berapa persen datakah yang di atas Q_3 ?
- Ada berapa persen datakah yang di atas Q_1 ?
- Ada berapa persen datakah yang di bawah P_{15} ?
- Kuartil berapakah yang nilainya sama dengan P_{25} ? P_{75} ?

E. Ukuran Penyebaran

1. Jangkauan Inter Kuartil

Ukuran penyebaran dari sekumpulan data mengukur seberapa jauh data-data tersebut tersebar. Dua kelompok data yang memiliki *mean* yang sama, bisa memiliki uluran penyebaran yang sangat berbeda.

Eksplorasi 7.8 Membandingkan Ukuran Penyebaran dari Dua Kelompok Data Tunggal



Ayo Bereksplorasi

Kelompok pertama yang terdiri dari 12 orang memiliki umur: 13, 14, 15, 15, 16, 16, 17, 17, 17, 17, 17, 18

Kelompok kedua yang juga terdiri dari 12 orang memiliki umur: 1, 3, 4, 5, 7, 8, 12, 27, 28, 29, 32, 36

Hitunglah *mean*, Q_1 , dan Q_3 dari kedua kelompok di atas.

Rata-rata umur dari kelompok pertama maupun kelompok kedua adalah 16 tahun.

Walaupun kedua kelompok memiliki *mean* yang sama, namun jika kalian memperhatikan setiap data dari kedua kelompok, manakah yang menurut kalian lebih mewakili kelompok umur siswa? Manakah yang lebih mewakili umur orang dewasa dan anak kecil? Jelaskan alasanmu.

Salah satu ukuran penyebaran yang telah kalian pelajari sebelumnya adalah jangkauan (*range*).

$$\text{Range kelompok pertama} = 18 - 13 = 5$$

$$\text{Range kelompok kedua} = 36 - 1 = 35$$

Range kelompok kedua lebih besar dari *range* kelompok pertama, berarti data pada kelompok kedua jauh lebih tersebar dibanding kelompok pertama.

Ukuran penyebaran lain yang dapat digunakan adalah jangkauan interkuartil. Jangkauan interkuartil diperoleh dengan cara mencari selisih antara kuartil atas (Q_3) dan kuartil bawah (Q_1).

Menghitung Q_1 dan Q_3 kelompok pertama, tidak perlu menggunakan metode interpolasi karena data merupakan data tunggal.

$$\text{Karena } \frac{1}{4} \text{ data} = \frac{1}{4} \times 12 = 3, \text{ maka } Q_1 \text{ terletak di antara data ke-3 dan ke-4}$$

$$\text{Sedangkan } \frac{3}{4} \times 12 = 9, \text{ maka } Q_3 \text{ terletak di antara data ke-9 dan ke-10}$$

$$\text{Kelompok pertama: } Q_1 = 15 \text{ dan } Q_3 = 17$$

$$\text{Kelompok kedua: } Q_1 = 4,5 \text{ dan } Q_3 = 28,5$$

Jangkauan interkuartil kelompok pertama = $17 - 15 = 2$, sedangkan jangkauan interkuartil kelompok kedua = $28,5 - 4,5 = 24$.

Jika hasil di atas kita tampilkan dalam tabel:

Tabel 7.8 Perbandingan Mean, Range dan Jangkauan Interkuartil Antara Kelompok Pertama dan Kedua

Kelompok	Mean	Range	Jangkauan Interkuartil
Pertama	16	5	2
Kedua	16	35	24

Maka kita dapat menyimpulkan bahwa walaupun kedua kelompok memiliki rata-rata umur yang sama yaitu 16, kalian akan dapat menemukan teman-teman yang seumuran dengan kalian pada kelompok pertama daripada kelompok kedua. Hal ini dikarenakan data-data yang tersebar pada kelompok pertama memiliki ukuran penyebaran (*range* dan jangkauan interkuartil) yang lebih kecil dibanding kelompok kedua. Jadi, data pada kelompok pertama banyak yang besarnya di sekitar *mean*.

2. Varian dan Simpangan Baku Data Tunggal

Ukuran penyebaran lainnya yang biasa digunakan untuk mengetahui sebaran data adalah varian.

Semakin kecil varian, maka data-data dalam kelompok tersebut semakin seragam mendekati *mean* kelompok. Demikian juga sebaliknya.

Varian diperoleh dengan cara mengurangi setiap data dengan *mean*, atau dengan rumus berikut:

Varian = $\frac{\sum(x - \bar{x})^2}{n}$, di mana $\bar{x}$ adalah *mean*.

Varian sering diberikan ditulis dalam simbol σ^2 .

Sedangkan simpangan baku adalah akar dari varian. Simbol untuk simpangan baku adalah σ .

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

Eksplorasi 7.9 Membandingkan Varian dari Dua Kelompok Data Tunggal



Ayo Bereksplorasi

Kembali ke soal kelompok umur:

Kelompok pertama yang terdiri dari 12 orang memiliki umur: 13, 14, 15, 15, 16, 16, 17, 17, 17, 17, 17, 18.

Kelompok kedua yang juga terdiri dari 12 orang memiliki umur: 1, 3, 4, 5, 7, 8, 12, 27, 28, 29, 32, 36.

Hitunglah varian dan simpangan baku dari kedua kelompok umur ini.

Rata-rata dari kelompok pertama maupun kedua = 16. Selanjutnya, mari kita hitung varian masing-masing kelompok.

$$\begin{aligned}\sigma^2 \text{ Kelompok 1} &= \frac{(13-16)^2 + (14-16)^2 + 2(15-16)^2 + 2(16-16)^2 + 5(17-16)^2 + (18-16)^2}{12} \\ &= \frac{9 + 4 + 2 + 0 + 5 + 4}{12} = \frac{24}{12} = 2\end{aligned}$$

σ^2 Kelompok 2

$$\begin{aligned}&= \frac{(1-16)^2 + (3-16)^2 + (4-16)^2 + (5-16)^2 + (7-16)^2 + (8-16)^2 + (12-16)^2 + (27-16)^2 + (28-16)^2 + (29-16)^2 + (32-16)^2 + (36-16)^2}{12} \\ &= \frac{225 + 169 + 144 + 121 + 81 + 64 + 16 + 121 + 144 + 169 + 256 + 400}{12} = \frac{1910}{12} = 159,2\end{aligned}$$

Simpangan baku (σ) kelompok 1 = $\sigma^1 = \sqrt{2} = 1,41$

Simpangan baku (σ) kelompok 2 = $\sigma^2 = \sqrt{159,2} = 12,62$

Kita dapat melihat bahwa nilai σ^2 yang besar menunjukkan bahwa data-data umur pada kelompok 2 memiliki sebaran yang jauh dari *mean* kelompok 2. Sedangkan kelompok 1 memiliki data-data yang relatif seragam dan mendekati *mean* kelompok 1.

Cara lain dalam menghitung varian:

$$\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

Mari kita menghitung ulang nilai dari varian kelompok 1 dengan rumus di atas dan membandingkan hasilnya dengan cara sebelumnya:

$$\begin{aligned}\sum x &= 13 + 14 + 15 + 15 + 16 + 17 + 17 + 17 + 17 + 17 + 18 = 192 \\ \sum x^2 &= 13^2 + 14^2 + 15^2 + 15^2 + 16^2 + 16^2 + 17^2 + 17^2 + 17^2 + 17^2 + 18^2 = 3096\end{aligned}$$

Karena jumlah data ada sebanyak 12, maka $n = 12$

$$\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2 = \frac{3096}{12} - \left(\frac{192}{12} \right)^2 = 2$$



Ayo Berdiskusi

Mengapa rumus kedua bisa memberikan hasil yang sama?

Bagaimana hasil varian dengan cara ini dibanding cara sebelumnya? Apakah sama?



Ayo Mencoba

1. Kalian dapat mencoba untuk mencari varian dengan rumus $\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$ untuk kelompok yang kedua.

2. Jika ada kelompok ketiga yang juga beranggotakan 12 orang, namun semuanya berusia 16 tahun, tanpa melakukan perhitungan menggunakan rumus, bisakah kalian menentukan mean, varian, dan simpangan baku dari kelompok ketiga ini?

3. Varian dan Simpangan Baku Data Kelompok

Sama halnya seperti mencari mean dari data kelompok, kita akan selalu mengasumsikan bahwa data-data yang terdapat dalam kelas interval tertentu diasumsikan tersebar merata sehingga kita dapat menggunakan nilai tengah dari setiap kelas interval. Mari kita lihat soal berikut.

Eksplorasi 7.10 Varian dalam Data Kelompok



Ayo Bereksplorasi

Dari suatu penelitian mengenai lamanya baterai HP, diperoleh data sebagai berikut:

Tabel 7.9 Tabel Distribusi Frekuensi Data Kelompok Durasi Baterai HP

Durasi baterai (jam)	6-10	11-15	16-20	21-25	26-30
Frekuensi	2	10	18	45	5

Tentukanlah varian dan simpangan dari durasi baterai tersebut.

Untuk data berkelompok, maka kita perlu menentukan nilai tengah dari masing-masing kelas terlebih dahulu. Lalu mencari nilai Σx dan Σx^2 . Agar lebih mudah, kita tempatkan semua nilai dalam tabel berikut.

Durasi baterai (jam)	Nilai tengah, x_i	Frekuensi, f	$f \cdot x_i$	$f \cdot x_i^2$
6-10	8	2	16	128
11-15	13	10	130	1.690
16-20	18	18	324	5.832
21-25	23	45	1.035	23.805
26-30	28	5	140	3.920
		80	1.645	35.375

Dari tabel di atas kita memperoleh:

$$\Sigma f = 80 \qquad \Sigma(fx) = 1.645 \qquad \Sigma(fx^2) = 35.375$$

$$\text{Maka varian } \sigma^2 = \frac{\Sigma(fx^2)}{\Sigma(f)} - \left(\frac{\Sigma(fx)}{\Sigma(f)} \right)^2 = \frac{35375}{80} - \left(\frac{1645}{80} \right)^2 = 19,37$$

$$\text{Simpangan baku } \sigma = \sqrt{19,37} = 4,4$$

Latihan 7.5

1. Dari suatu survei tentang banyaknya buku yang dibaca oleh siswa SMA dalam 1 bulan, diperoleh hasil yang diambil secara acak. Banyaknya buku yang dibaca 7 orang siswa adalah sebagai berikut:

3 4 6 2 8 8 5

Tentukanlah varian dan simpangan dari data tersebut.

2. Sebelum pandemi Covid-19, sekolah mencatat waktu yang diperlukan oleh siswa untuk makan siang di kantin (dibulatkan ke menit terdekat). Hasilnya adalah sebagai berikut:

Waktu yang diperlukan, t (menit)	35	36	37	38
Frekuensi	3	17	29	34

- a. Tentukanlah rata-rata dari data tersebut.
 - b. Tentukanlah simpangan bakunya.
3. Diketahui sekumpulan data memiliki data-data sebagai berikut:

$$\Sigma x = 24 \qquad \Sigma x^2 = 78 \qquad n = 8$$

Carilah:

- a. *mean*
 - b. Varian, σ^2
 - c. Simpangan baku, σ
4. Dari data kelompok pertama yang terdiri dari 10 bilangan diperoleh sebagai berikut:

$$\Sigma x = 50 \qquad \Sigma x^2 = 310$$

Sedangkan kelompok kedua yang terdiri dari 15 bilangan diperoleh sebagai berikut:

$$\Sigma x = 86 \qquad \Sigma x^2 = 568$$

Tentukanlah *mean* dan simpangan baku dari gabungan kedua kelompok tersebut yang terdiri dari 25 bilangan.

5. Guru berbeda mengajar 2 kelas yang berbeda, kelas A dan kelas B, dengan beda metode mengajar. Siswa dari kedua kelas tersebut mengikuti ujian yang sama pada akhir semester. Berikut hasil ujian dari kedua kelas.

Hasil Ujian	20-29	30-39	40-49	50-59	60-69	70-79	80-89
Frekuensi A	1	3	6	6	11	10	8
Frekuensi B	1	2	4	13	15	6	3

- Hitunglah *mean* dari masing-masing kelompok.
- Dari hasil a, menurut kalian, apakah metode guru yang satu lebih baik dari metode guru lainnya? Jelaskan alasan dari jawabanmu!

Refleksi

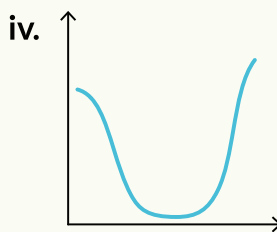
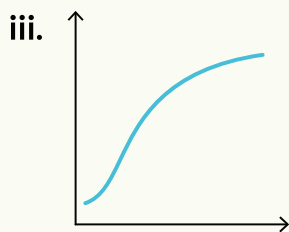
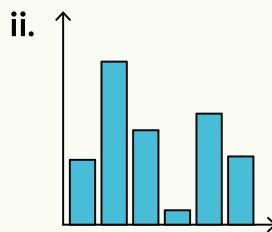
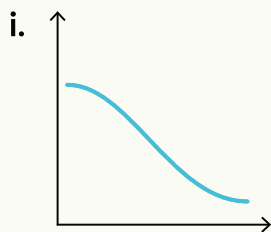
Dalam bab ini, kalian sudah belajar mengenai ukuran pemusatan, ukuran lokasi, dan ukuran penyebaran dari suatu kelompok data dan menggunakan berbagai ukuran tersebut dalam melakukan pengambilan keputusan.

- a. Bagaimana menemukan *mean*, modus, dan median data kelompok?
- b. Bagaimana menemukan ukuran lokasi seperti persentil dan kuartil baik dalam data tunggal maupun dalam data kelompok?
- c. Bagaimana menemukan varian dan simpangan baku baik dalam data tunggal maupun dalam data kelompok?

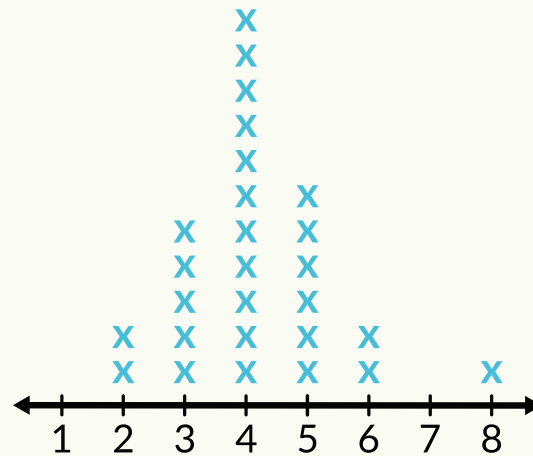
Bagaimana memilih ukuran pemusatan yang tepat dan sesuai dengan konteks masalah?

Uji Kompetensi

1. Di antara keempat grafik di bawah ini, manakah yang merupakan grafik dari:
 - a. Perubahan berat badan seekor kucing dari lahir sampai usia 2 tahun.
 - b. Aktivitas kegiatan anak dari sebelum tidur dan setelah tidur.
 - c. Jumlah penduduk di 6 kota yang berbeda.
 - d. Ketinggian permukaan air laut dari kondisi pasang ke kondisi surut.



2. Saat pelajaran matematika, para siswa di kelas 10 menggambar *line plot* yang menunjukkan banyaknya anggota keluarga dari setiap siswa.



- Berapakah banyaknya orang yang terdapat dalam keluarga para siswa di kelas tersebut?
 - Seorang siswa yang bernama Jono berkata, "*Line plot* ini salah! Anggota keluarga saya berjumlah 8 orang. Saya memiliki jumlah anggota keluarga yang terbanyak, mengapa tanda X di atas angka 8 malah jadi yang paling pendek?" Jawablah pertanyaan Jono.
3. Hasil 4 ulangan matematika Dodi adalah 81, 79, 90, dan 70. Ulangan ke-5 baru akan dibagikan. Guru Dodi menyampaikan ke Dodi bahwa Dodi boleh memilih apakah mau menggunakan median atau *mean* sebagai nilai rapornya, namun Dodi harus menentukan sebelum ia menerima hasil tes matematika yang ke-5.
- Hitunglah *mean* dan median dari keempat hasil ulangan matematika Dodi.
 - Jika Dodi tidak yakin dengan hasil ulangan ke-5 nya, manakah yang sebaiknya ia pilih, *mean* atau median? Jelaskan.
 - Jika Dodi yakin dengan hasil ulangan ke-5 nya, manakah yang sebaiknya ia pilih, *mean* atau median? Jelaskan.
4. Dalam ujian Fisika, rata-rata nilai dari delapan siswa adalah 65. Rata-rata grup kedua yang berjumlah 12 siswa adalah 72. Hitunglah rata-rata gabungan dari kedua kelompok ini yang berjumlah 20 siswa.

5. Selama tahun ajaran yang lalu, diperoleh data banyaknya hari di mana siswa tidak hadir.

Jumlah hari absen	0	1	2	3	4
Frekuensi	12	20	10	7	5

- Hitunglah Q_1 dari data ini, lalu interpretasikan hasilnya
 - Hitunglah jangkauan interkuartil dari data ini.
 - Hitunglah standar deviasi dari data jumlah hari absen tersebut.
6. Dalam suatu lomba lari, diperoleh data catatan waktu sebagai berikut:

Waktu yang ditempuh, t (menit)	20-29	30-39	40-49	50-59	60-69
Frekuensi	5	10	36	20	9

- Hitunglah *mean*.
- Gunakanlah interpolasi untuk menghitung jangkauan interkuartil.
- Jika diketahui $\Sigma fx = 3.740$ dan $\Sigma fx^2 = 183.040$ di mana x adalah nilai tengah dari tiap kelas, maka tentukanlah nilai dari varian dan simpangan baku dari catatan waktu para pelari.