

ESPA4123
Edisi 2

MODUL 01

Pengantar Ilmu Statistika

Rini Yayuk Priyati, S.E., M.Ec., Ph.D.

Daftar Isi

Modul 01	1.1
Pengantar Ilmu Statistika	
Kegiatan Belajar 1	1.4
Apakah Ilmu Statistika Itu?	
Latihan	1.17
Rangkuman	1.19
Tes Formatif 1	1.20
Kegiatan Belajar 2	1.22
Pengumpulan Data dan Teknik <i>Sampling</i>	
Latihan	1.28
Rangkuman	1.30
Tes Formatif 2	1.32
Kunci Jawaban Tes Formatif	1.34
Daftar Pustaka	1.35



Pendahuluan

Ketika kita mendengar istilah statistika, sering kali yang pertama terlintas adalah angka-angka dan perhitungan-perhitungan yang rumit. Tahukah Anda bahwa dalam kehidupan sehari-hari, disadari ataupun tanpa disadari, kita sering menggunakan istilah-istilah statistika, misalnya, kira-kira, kurang lebih, rata-rata, perkiraan, kemungkinan, tren, survei, *pooling*, dan lain-lain.

Ketika kita membaca surat kabar atau artikel, baik yang ilmiah maupun tidak terlalu ilmiah, sering kali penggunaan angka-angka statistik tidak terhindarkan. Tanpa disadari, hal ini telah menjadikan kita lebih familier dengan ilmu statistika itu sendiri. Sebagai mahasiswa ilmu ekonomi, Anda tentu tidak hanya diharapkan mampu membaca angka-angka atau laporan-laporan statistik dengan benar. Anda pada akhirnya diharapkan mampu mengolah dan menyajikan laporan statistik secara benar dan akurat.

Dalam kehidupan sehari-hari, terdapat banyak sekali informasi, sebagian benar, dan dapat dipertanggungjawabkan kebenarannya dan sebagian lainnya tidak. Mengambil kesimpulan atau bahkan keputusan hanya berdasarkan kira-kira saja tentu bukan suatu langkah yang tepat secara ilmiah. Ilmu statistika menyediakan banyak metode yang dapat membantu kita dalam menentukan pilihan-pilihan kesimpulan atau keputusan yang dapat dipertanggungjawabkan secara ilmiah.

Modul ini merupakan modul pembukaan dari BMP *Statistika Ekonomi* (ESPA4123). Modul ini akan menyajikan pengetahuan-pengetahuan mendasar tentang ilmu statistika. Anda disarankan untuk lebih dahulu familier dengan konsep dasar dan istilah-istilah statistika yang akan dibahas dalam modul ini sebelum beranjak lebih jauh ke modul-modul selanjutnya. Modul ini berisi penjelasan-penjelasan berupa konsep-konsep dan uraian-uraian, sedangkan perhitungan-perhitungan yang berkaitan dengan ilmu statistika akan Anda temukan pada Modul 3 dan seterusnya. Setelah mempelajari modul ini, mahasiswa diharapkan mampu

1. menjelaskan konsep dasar ilmu statistika,
2. menghitung dengan menggunakan alat bantu matematika sederhana,
3. menjelaskan metode pengumpulan data,
4. menjelaskan teori *sampling*.

Apakah Ilmu Statistika Itu?

A. STATISTIKA DAN STATISTIK

Sering kali kita dihadapkan dalam pengambilan keputusan, dari yang paling sederhana ke yang lebih rumit. Pengambilan keputusan hendaknya didasarkan atas informasi-informasi yang kita terima dan disesuaikan dengan norma-norma yang berlaku. Ketika kita menghadapi keputusan-keputusan yang lebih rumit, informasi dan norma saja mungkin kurang memadai. Diperlukan alat bantu lain. Di sinilah ilmu statistika dapat dipergunakan untuk menentukan keputusan mana yang paling baik di antara banyak pilihan keputusan berdasarkan data-data statistik.

Dalam lingkup ekonomi, keputusan-keputusan yang diambil oleh para pelaku ekonomi bahkan pemerintah sangat dipengaruhi oleh keadaan ekonomi yang disajikan ke dalam angka-angka statistik. Untuk itulah, di negara mana pun di seluruh dunia terdapat badan tertentu yang mengurus data-data statistik. Di Indonesia, pasti Anda sudah sangat terbiasa mendengar Badan Pusat Statistik (BPS). BPS adalah lembaga resmi pemerintah nonkementerian yang bertanggungjawab langsung kepada presiden yang tugas pokoknya menyediakan data bagi pemerintah dan masyarakat umum. Data dapat dikumpulkan sendiri oleh BPS melalui sensus dan survei (data primer) atau dapat pula berupa data-data sekunder yang didapatkan melalui instansi-instansi lain yang kemudian dipublikasikan oleh BPS. Hampir semua lembaga resmi pemerintah bahkan perusahaan swasta memiliki bagian tersendiri yang mengurus masalah data. Apabila Anda masuk ke situs internet lembaga-lembaga pemerintah ataupun perusahaan-perusahaan besar, Anda akan menemukan jendela khusus yang berisi data-data yang terkait dengan institusi yang bersangkutan.

Contoh 1.1

Cobalah Anda memasuki situs internet kementerian-kementerian yang ada di Indonesia. Periksalah apakah mereka memiliki jendela basis data.

Jika Anda membaca dengan baik penjelasan di atas, Anda akan menemukan dua terminologi yang hampir sama, tetapi memiliki arti yang berbeda, yaitu statistika dan statistik.

Pengertian statistika menurut *Kamus Besar Bahasa Indonesia* (KBBI) sebagai berikut:

1. ilmu tentang cara mengumpulkan, menabulasi, menggolong-golongkan, menganalisis, dan mencari keterangan yang berarti dari data yang berupa angka;
2. pengetahuan yang berhubungan dengan pengumpulan data, penyelidikan, dan kesimpulannya berdasarkan bukti yang berupa catatan bilangan (angka-angka) (KBBI, 2017).

Sementara itu, pengertian statistik sebagai berikut:

1. catatan angka-angka (bilangan); perangkaan;
2. data yang berupa angka yang dikumpulkan, ditabulasi, digolong-golongkan sehingga dapat memberi informasi yang berarti mengenai suatu masalah atau gejala (KBBI, 2017).

Jadi, dapat kita simpulkan bahwa statistik adalah hasil atau informasi berupa angka yang kita peroleh dari metode statistika yang kita lakukan dimulai dari mengumpulkan data, pemrosesan data, dan diakhiri dengan penginterpretasian hasil dan penarikan kesimpulan.

Berikut ini akan disajikan konsep-konsep dasar yang harus Anda pelajari sebelum mempelajari mata kuliah ini lebih jauh lagi. Pemahaman terhadap konsep-konsep tersebut akan membantu Anda menguasai mata kuliah ini lebih jauh.

B. STATISTIKA DESKRIPTIF DAN STATISTIKA INFERENSIAL

Ilmu statistika dapat dibagi ke dalam dua kelompok, yaitu statistika deskriptif dan statistika inferensial. Pada awal-awal perkembangan ilmu statistika, statistika hanya digunakan untuk menggambarkan suatu fenomena dan menyimpulkannya ke dalam angka-angka. Inilah yang sekarang kita kenal sebagai statistika deskriptif.

Statistika deskriptif inilah yang paling umum kita jumpai dan digunakan; tidak hanya untuk kepentingan ilmiah, tetapi juga pada tulisan-tulisan jurnalistik yang biasanya terdapat di koran dan majalah. Penggambaran paling umum dari statistika deskriptif ini dapat Anda lihat dalam bentuk tabel, gambar, dan grafik yang sering digunakan sebagai pendukung suatu pendapat atau opini yang diutarakan.

Contoh 1.2

Bacalah koran atau majalah yang ada di sekitar Anda. Bisakah Anda temukan suatu tulisan yang menggunakan data dalam bentuk tabel, gambar, atau grafik?

Jadi, statistika deskriptif adalah bagian dari ilmu statistika yang berisi metode-metode penyampaian data dan kemudian menyimpulkannya sehingga dapat lebih mudah dimengerti. Statistik deskriptif tidak bisa digunakan untuk memperkirakan atau menyimpulkan hal-hal yang mungkin terjadi di luar jangkauan data yang kita analisis.

Selain menggunakan penggambaran dalam bentuk tabel, grafik, dan gambar, beberapa perhitungan yang paling umum digunakan dalam statistika deskriptif adalah nilai tendensi sentral dan nilai simpangan. Kedua nilai ini akan kita pelajari pada bagian lain dari BMP ini.

Adakalanya menggambarkan data secara keseluruhan dengan menggunakan statistika deskriptif sangat sulit atau bahkan tidak mungkin dilakukan, khususnya untuk data-data yang lingkungannya sangat besar, misalnya satu negara. Contohnya adalah apabila kita ingin mengetahui tinggi badan siswa kelas IX di Indonesia, akan sangat sulit dan memakan waktu apabila kita harus mendata tinggi badan semua siswa kelas IX yang ada di Indonesia. Maka dari itu, yang kita lakukan adalah menggunakan sampel beberapa ratus siswa kelas IX. Statistika inferensial memungkinkan kita untuk menarik kesimpulan dari keseluruhan populasi hanya dengan menggunakan sampel yang kita tentukan. Teknik pengambilan sampel (*sampling*) sangat umum digunakan dalam statistika dan akan dibahas pada bagian lain dari modul ini. Statistika inferensial dapat kita andalkan dalam pengambilan keputusan walaupun kita tetap harus memperhitungkan kesalahan (*error*) yang muncul dari teknik ini. Error ini biasanya dinyatakan dalam persentase. Jadi, statistika inferensial adalah suatu metode statistik yang digunakan untuk menarik kesimpulan atau bahkan memperkirakan keseluruhan populasi hanya dengan menggunakan informasi yang diambil dari sampel.

C. DATA

Setelah mempelajari penjelasan di bagian A di atas, Anda sekalian tentu sudah mengerti perbedaan antara statistika dan statistik. Dalam kehidupan sehari-hari, statistik lebih kita kenal dengan istilah data walaupun sebenarnya ada perbedaan antara data dan statistik. Perbedaan paling mendasar adalah data tidak harus berupa angka, sedangkan statistik, sesuai dengan pengertian di bagian terdahulu, adalah selalu berbentuk angka.

Contoh 1.3

Dapatkah Anda menyebutkan contoh data yang tidak berbentuk angka?

Jawab

Coba perhatikan KTP Anda. Pada setiap barisnya terdapat data-data kependudukan Anda yang tidak semua berisi angka. Anda akan menemukan nama, jenis kelamin, alamat, agama, status perkawinan, pekerjaan, dan kewarganegaraan yang semuanya berupa data, tetapi tidak berbentuk angka.

Akan tetapi, untuk memudahkan, setelah bagian ini kita akan menggunakan istilah data untuk menggantikan istilah statistik.

Data, sesuai dengan cara diperolehnya, dapat dikelompokkan menjadi dua, yaitu data primer dan data sekunder. Apabila Anda telah membaca dengan baik bagian A di atas, Anda akan mengetahui perbedaan keduanya.

Contoh 1.4

Bacalah kembali penjelasan bagian A di atas. Sebutkan perbedaan antara data primer dan data sekunder.

Benar! Data primer adalah data yang dikumpulkan secara langsung oleh si pelapor data dari sumber aslinya. Data yang dikumpulkan dilaporkan sesuai dengan aslinya dan tidak mengalami proses perhitungan. Banyak cara yang bisa dilakukan untuk mengumpulkan data primer, di antaranya melalui observasi, eksperimen, survei, dan sensus.

Data sekunder adalah data yang dikumpulkan dari pihak lain yang telah mengalami pemrosesan dan perhitungan. Ada banyak sumber yang bisa kita rujuk untuk menemukan data sekunder, contohnya adalah statistik yang dipublikasikan oleh BPS, laporan keuangan perusahaan-perusahaan, data-data yang disajikan di koran atau majalah, data-data yang terdapat pada laman internet suatu organisasi, dan sebagainya.

Sesuai dengan bentuknya, data dapat kita bedakan menjadi data kualitatif dan data kuantitatif. Cara paling mudah dan umum yang dilakukan untuk membedakan data kuantitatif dan data kualitatif adalah melihat bentuk datanya. Data kuantitatif adalah data berbentuk angka atau yang dapat diangkakan contoh data kuantitatif adalah jarak, umur, harga, tinggi badan, dan sebagainya.

Sementara itu, data kualitatif adalah data yang berbentuk uraian atau yang berbentuk selain angka, contohnya jenis kelamin, pekerjaan, alamat domisili, warna kulit, etnis, dan sebagainya. Definisi di atas berlaku umum walaupun pada kasus-kasus tertentu tidak harus demikian. Misalnya, dalam beberapa kasus, kita menemukan data yang berbentuk angka, tetapi sebenarnya adalah data kualitatif, misalnya kode pos.

Kode pos berisi serangkaian angka (di Indonesia terdiri atas empat angka), tetapi sama sekali tidak merujuk ke kuantitas, tetapi merujuk ke lokasi. Contoh lain dari bentuk data kualitatif yang berbentuk angka adalah kode barang yang tertera di kemasan barang (biasanya dalam bentuk kode batang/ *barcode*).

Contoh 1.5

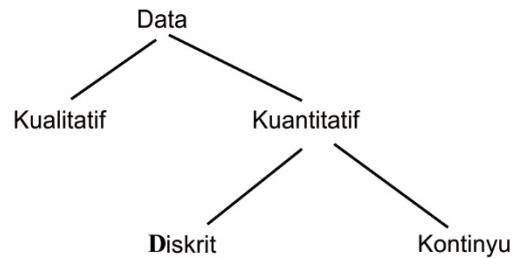
Bisakah Anda menjelaskan alasan mengapa kode batang (*barcode*) merupakan data kualitatif?

Dalam praktiknya, untuk menyederhanakan penyimpanan data, data kualitatif biasanya disimpan dalam bentuk kode/angka yang menunjukkan kategori. Misalnya, untuk kategori jenis kelamin, biasanya dinyatakan dengan angka 0 untuk laki-laki dan 1 untuk perempuan. Dalam hal ini, angka 0 dan 1 tidak berarti nilai, hanya menunjukkan kategori. Dalam penggunaan data kualitatif, hal yang biasanya kita amati adalah berapa jumlah atau proporsi untuk setiap kategorinya, misalnya berapa orang responden yang berjenis kelamin laki-laki (kode 0) dan berapa orang yang berjenis kelamin perempuan (kode 1) atau berapa persen responden yang berjenis kelamin laki-laki dan berapa persen yang berjenis kelamin perempuan.

Pada data kuantitatif, hal yang ingin kita amati adalah nilai dari variabel tersebut, bagaimana nilai tendensi pusatnya, bagaimana sebarannya, dan sebagainya. Data kuantitatif sangat umum digunakan dalam penelitian-penelitian bidang eksakta. Untuk bidang noneksakta, data kuantitatif biasanya lazim digunakan dalam penelitian-penelitian bidang ekonomi.

BMP ini secara keseluruhan akan membahas penggunaan statistika dalam bidang ekonomi. Dalam kehidupan sehari-hari, kita sering sekali membaca atau menyaksikan bagaimana statistika digunakan dalam bidang ekonomi misalnya dalam penggambaran indikator-indikator ekonomi, seperti tingkat pertumbuhan ekonomi, tingkat inflasi, fluktuasi nilai tukar, indeks harga saham, dan sebagainya. Contoh-contoh tersebut merupakan sebagian kecil dari banyak contoh penggunaan statistika dalam bidang ekonomi. Sementara itu, data kualitatif sangat umum digunakan dalam penelitian-penelitian yang berhubungan dengan ilmu sosial/humaniora, misalnya bidang sosiologi, antropologi, hukum, ilmu sejarah, ilmu bahasa, dan sebagainya, walaupun dalam beberapa kasus data kualitatif juga digunakan dalam bidang-bidang eksakta.

Data kuantitatif dapat dibagi menjadi dua, yaitu data diskrit (*discrete*) dan data kontinu (*continuous*). Data diskrit adalah data yang berbentuk angka bulat (bukan pecahan) yang merupakan hasil hitung, sedangkan data kontinu adalah data yang memiliki nilai terus-menerus (dapat berbentuk pecahan) yang merupakan hasil ukur. Apabila kita gambarkan dalam sebuah bagan, urut-urutannya sebagai berikut.



Gambar 1.1
Bagan Data

Untuk memperjelas, perhatikan contoh di bawah ini.

Contoh 1.6



Perhatikan gambar kucing dan bebek di samping. Data apa yang Bisakah Anda temukan sesuai dengan data kualitatif, diskrit dan kontinu?

Jawab

Data kualitatif:

- Nama hewan di kiri: kucing
- Nama hewan di kanan: bebek
- Cara berkembang biak hewan di sebelah kiri: melahirkan
- Cara berkembang biak hewan di sebelah kanan: bertelur
- Dan sebagainya (bisakah Anda menyebutkan kriteria kualitatif lain dari kucing dan bebek?)

Data diskrit:

- Jumlah kucing: 1
- Jumlah bebek: 2
- Jumlah kaki setiap kucing: 4
- Jumlah kaki setiap bebek: 2
- Dan seterusnya.

Data kontinu:

- Berat kucing: 3,5 kg (contoh)
- Berat 1 ekor bebek: 1,5 kg (contoh)
- Dan seterusnya.

Jadi, data diskrit adalah data yang bisa dihitung dan data kontinu adalah data yang bisa diukur.

D. KOEFISIEN DAN VARIABEL

Koefisien dan variabel terutama sangat terkait dengan statistika inferensial. Statistika inferensial biasanya dinyatakan dalam persamaan matematis. Dalam kaitan itu, sangat penting untuk membedakan istilah yang digunakan dalam persamaan matematis. Perhatikan persamaan matematis berikut ini.

$$y = 3x_1 - 2x_2 + 5$$

Ada empat anggota dari persamaan di atas, yaitu y , $3x_1$, $-2x_2$, dan 5 . Variabel dinyatakan dalam huruf y , x_1 , dan x_2 . Walaupun dinyatakan dalam huruf, variabel sebenarnya berisi serangkaian angka yang bisa ditemukan pada data yang jumlahnya bervariasi sehingga diwakilkan dengan menggunakan huruf.

Koefisien adalah nilai yang berada di depan variabel. Pada contoh di atas, angka 3 menunjukkan koefisien untuk x_1 dan -2 menunjukkan koefisien untuk x_2 . Sementara itu, untuk variabel yang tidak memiliki koefisien di depannya (y) nilai koefisiennya adalah 1 .

Selain variabel dan koefisien, persamaan di atas juga mengandung konstan. Konstan adalah suatu nilai yang tidak memiliki variabel. Disebut konstan karena nilainya tidak bisa berubah. Pada contoh di atas, yang dimaksud dengan konstan adalah 5 . Konstan biasanya dinyatakan dengan lambang c .

Contoh 1.7

Tunjukkan variabel, koefisien, dan konstan dari persamaan di bawah ini.

$$2y = 3x^2 + 6xy$$

Jawab

Variabel : y , x , dan xy

Koefisien : 2 untuk y , 3 untuk x , dan 6 untuk xy

Konstan : $c = 0$

E. VARIABEL BEBAS DAN VARIABEL TERIKAT

Variabel dapat dibedakan menjadi dua kelompok, yaitu variabel bebas (variabel independen) dan variabel terikat (variabel dependen). Variabel bebas adalah suatu variabel yang perubahan nilainya bisa memengaruhi nilai variabel terikat. Variabel terikat bisa kita definisikan sebagai variabel yang nilainya dipengaruhi oleh perubahan nilai variabel bebas. Pengaruh terhadap variabel terikat inilah yang kemudian kita observasi.

Contoh 1.8

Dilakukan penelitian tentang pengaruh lama waktu penggunaan telepon genggam terhadap prestasi belajar siswa SMP. Bisakah Anda menentukan mana variabel bebas dan variabel terikatnya?

Jawab

Variabel bebas: rata-rata lama waktu penggunaan telepon genggam oleh siswa SMP per hari.

Variabel terikat: prestasi belajar siswa SMP (misalnya menggunakan indikator nilai rapor.)

Lazimnya, dalam suatu persamaan matematis, variabel terikat diletakkan di sebelah kiri tanda sama dengan dan variabel bebasnya di sebelah kanan tanda sama dengan. Apabila digambarkan dalam sumbu x dan y , variabel terikat biasanya diletakkan pada sumbu vertikal dan variabel bebas diletakkan pada sumbu horizontal.

F. SKALA PENGUKURAN**Contoh 1.9**

Tunjukkan variabel bebas dan variabel terikat dari persamaan di bawah ini.

$$y = 3x_1 - 2x_2 + 5$$

Jawab

Variabel bebas : x_1 dan x_2

Variabel terikat : y

Ilmu statistika sangat bergantung dengan data dan perhitungan. Untuk itulah diperlukan skala pengukuran. Skala pengukuran digunakan untuk pengategorian atau penguantifikasian suatu variabel. Penggunaan skala pengukuran yang tepat akan membantu kita menghasilkan pengukuran yang akurat terhadap suatu permasalahan tertentu.

Terdapat empat skala pengukuran yang digunakan dalam statistika dimulai dari yang terlemah ke yang terkuat, yaitu nominal, ordinal, interval, dan rasio. Masing-masing skala pengukuran tersebut akan dijelaskan di bawah ini.

1. Skala Nominal

Skala nominal adalah pengelompokan objek-objek penelitian ke dalam kategori-kategori tertentu yang telah ditetapkan. Data kualitatif biasanya masuk dalam kategori skala nominal ini. Contoh termudah adalah memberikan skala pada jenis kelamin,

misalnya 0 untuk laki-laki dan 1 untuk wanita. Walaupun dilambangkan dalam bentuk angka, besar kecilnya angka tidak menunjukkan nilai sama sekali, misalnya hanya karena perempuan dilambangkan dengan angka 1 (lebih besar dari angka laki-laki yang memiliki angka 0) tidak berarti menunjukkan bahwa perempuan memiliki nilai lebih dari laki-laki. Angka di sini hanya digunakan untuk penggolongan atau pengategorian saja. Contoh lain dengan kategori lebih banyak dapat Anda lihat di bawah ini.

Contoh 1.10

Dilakukan survey kesehatan terhadap pengungsi banjir di “Daerah X”. Pertanyaan yang ditanyakan adalah:

Keluhan kesehatan apa yang **paling** Anda rasakan selama berada dalam pengungsian?

1. Batuk-batuk
2. Sesak nafas
3. Gatal-gatal
4. Diare
5. Demam
6. Sulit tidur
7. Depresi
8. lainnya...

Pada contoh di atas, para pengungsi diminta untuk mengisi jenis keluhan kesehatan yang paling mengganggu mereka. Penomoran 1 sampai 8 sama sekali tidak merujuk pada nilai atau sama sekali tidak menunjukkan bahwa penyakit yang satu lebih mengkhawatirkan/parah dari penyakit yang lain. Penomoran hanya digunakan untuk mengategorikan jenis penyakit. Skala nominal adalah skala pengukuran yang paling lemah.

2. Skala Ordinal

Apabila skala nominal hanya digunakan sebagai pengategorian, skala yang sedikit lebih tinggi adalah skala ordinal. Dengan menggunakan skala ordinal, kita bisa melihat *rangking* dari hal yang kita amati. Akan lebih mudah dipahami jika Anda memperhatikan contoh di bawah ini.

Contoh 1.11

Dilakukan survei terhadap pengguna jasa pesawat terbang. Para responden diminta untuk memberikan *rangking* terhadap maskapai-maskapai di bawah ini sesuai dengan aspek pelayanan di kabin dengan angka 1 sampai 5, yaitu 1 menyatakan maskapai yang paling tinggi tingkat pelayanannya dan 5 adalah maskapai yang paling rendah tingkat pelayanannya. Kuesioner yang diisi oleh salah satu responden menunjukkan hasil sebagai berikut.

<i>Rangking</i>	Nama Maskapai
4	Aman Air
2	Sukses Air
1	Cendrawasih Air
3	Harimau Air
5	Nusantara Air

Pada contoh di atas, responden bisa menyatakan pilihannya dalam *rangking*, yaitu maskapai yang paling baik pelayanannya adalah Cendrawasih Air dan yang paling jelek pelayanan kabinnnya adalah Nusantara Air. Walaupun demikian, dengan menggunakan skala ordinal, kita tidak bisa menguantifikasi perbedaan tingkat kenyamanan pelayanan satu maskapai dengan maskapai lainnya. Yang bisa kita lakukan hanya menyatakan bahwa orang akan merasa lebih puas jika memilih Sukses Air dibandingkan dia memilih terbang dengan Aman Air (Sukses Air memiliki *rangking* 2 dan Aman Air memiliki *rangking* 4). Namun, kita tidak bisa menyatakan bahwa orang yang memilih Sukses Air akan merasa dua kali lipat lebih puas daripada jika dia memilih Aman Air.

3. Skala Interval

Dalam skala pengukuran, skala interval memiliki tingkatan yang lebih tinggi daripada skala nominal dan ordinal. Selain bisa digunakan untuk menyatakan *rangking*, skala interval juga bisa digunakan untuk menyatakan jarak absolut dari dua pengukuran. Skala interval ini sudah menggunakan data kuantitatif sebagai pengukurnya. Namun, sebagaimana skala ordinal, besarnya jarak antara satu pengukuran dan pengukuran yang lain tidak menunjukkan arti apa-apa. Perhatikan contoh berikut.

Contoh 1.12

Untuk mengetahui kemampuan berbahasa Inggris para calon mahasiswa baru, Universitas X melakukan tes bahasa Inggris berupa tes TOEFL. Dari puluhan calon mahasiswa baru, berikut ini adalah hasil yang diperoleh oleh Amir, Budi, Candra, dan Dewi.

Amir : 500
Budi : 600
Candra : 500
Dewi : 400

Dari contoh di atas, kita mendapatkan informasi bahwa skor yang diperoleh Budi adalah 100 lebih tinggi daripada skor yang diperoleh Amir. Kita juga bisa menyatakan bahwa skor yang diperoleh Amir adalah 100 poin lebih tinggi dari skor yang diperoleh Dewi. Maka dari itu, jarak absolut Budi-Amir dan dari Amir-Dewi adalah sama-sama bernilai 100.

Selain itu, kita juga bisa menyimpulkan bahwa Budi adalah calon mahasiswa dengan kemampuan berbahasa Inggris terbaik karena memiliki nilai skor 600. Sebaliknya, Dewi memiliki skor terendah (400) sehingga dapat kita simpulkan bahwa Dewi memiliki kemampuan berbahasa Inggris terjelek dari empat orang calon mahasiswa di atas. Besarnya skor Budi adalah 50% lebih tinggi dari skor yang didapat oleh Dewi. Namun, hal ini tidak menunjukkan bahwa Budi 50% lebih pintar dari Dewi dalam hal kemampuan bahasa Inggris.

4. Skala Rasio

Dari keempat skala pengukuran, skala rasio inilah yang paling tinggi tingkatannya. Dengan skala rasio, kita tidak hanya bisa menyatakan *ranking*, perbedaan absolut, tetapi juga rasio dari data. Skala rasio selalu dinyatakan dalam data kuantitatif dan tidak bisa bernilai kurang dari nol. Perhatikan contoh di bawah ini.

Contoh 1.13

Amir dan Budi sama-sama bekerja di PT X. Amir memperoleh gaji bulanan sebesar Rp5 juta dan Budi Rp10 juta.

Kesimpulan yang bisa kita ambil dari contoh di atas adalah Budi memiliki gaji bulanan yang lebih tinggi dari gaji Amir dengan selisih absolut Rp 5 juta yang juga berarti bahwa Budi memiliki gaji dua kali lebih banyak dari Amir.

Banyak contoh yang bisa kita gunakan untuk menyatakan skala rasio, antara lain panjang jalan, jumlah orang, umur, berat badan, harga barang, dan sebagainya.

Perhatikan contoh di bawah ini.

Contoh 1.14

Perhatikan kuesioner kepuasan pelanggan Restoran X di bawah ini. Masuk dalam kategori manakah skala pengukuran ini?

Kuesioner Kepuasan Pelanggan Restoran X

Jawablah pertanyaan di bawah ini dengan memberikan tanda centang \surd pada aspek kepuasan menurut pengalaman Anda.

Nilai kepuasan:

- 1 : Sangat tidak puas 3 : Ragu-ragu 5 : Sangat puas
2 : Tidak puas 4 : Puas

No	Aspek kepuasan	1	2	3	4	5
1	Pilihan menu					
2	Rasa masakan					
3	Penyajian/tampilan makanan					
4	Kebersihan					
5	Pelayanan					
6	Harga					

Skala pengukuran dalam contoh kuesioner kepuasan pelanggan Restoran X di atas lebih kita kenal dengan skala Likert (*Likert scale*). Skala pengukuran di atas biasanya sering kita jumpai dalam survei-survei berbentuk kuesioner. Lalu, masuk ke dalam kategori yang manakah skala pengukuran Likert ini? Belum ada pendapat yang seragam antara para ahli statistika mengenai skala Likert ini, ada yang memasukkannya dalam kategori skala ordinal dan ada pula yang mengategorikannya sebagai skala interval. Namun, mayoritas berpendapat bahwa skala Likert termasuk skala interval karena persepsi orang terhadap jarak antarkategori kepuasan adalah sama, contohnya persepsi antara jarak sangat tidak puas dan tidak puas adalah sama dengan persepsi antara jarak puas dan sangat puas.

G. KEMAMPUAN MATEMATIKA DASAR YANG DIBUTUHKAN

Pada dasarnya, statistika adalah cabang dari ilmu matematika yang khusus mempelajari data dan pengolahan data. Sebelum mempelajari mata kuliah ini lebih jauh, ada baiknya kita mengulang kembali beberapa keahlian dasar matematika. Karena mata kuliah ini merupakan statistika dasar, Anda tidak perlu khawatir apabila Anda tidak terlalu menguasai perhitungan matematis yang rumit. Sebagian besar penjelasan yang

digunakan pada mata kuliah ini hanya membutuhkan kemampuan matematis sederhana. Kemampuan dasar matematika, seperti penjumlahan, pengurangan, perkalian, pembagian, perpangkatan, dan pengakaran, sudah cukup untuk menguasai mata kuliah ini. Untuk perhitungan yang lebih rumit, Anda dipersilakan menggunakan alat bantu hitung seperti kalkulator dan perangkat lunak komputer. Namun demikian, keahlian dasar matematis tetap diperlukan untuk membantu Anda memahami bagaimana perhitungan-perhitungan tersebut berasal. Berikut ini akan disajikan beberapa contoh keahlian dasar matematis yang mungkin Anda perlukan dalam mata kuliah ini.

Contoh 1.15

Temukan nilai x dari persamaan di bawah ini.

$$2x - 5 = 3$$

Jawab

$$2x - 5 = 3$$

$$2x = 3 + 5 = 8$$

$$x = 4$$

Contoh 1.16

Temukan nilai x dari persamaan di bawah ini.

$$\frac{x}{2} = 15$$

Jawab

$$\frac{x}{2} = 15$$

$$x = 15 \cdot (2) = 30$$

Contoh 1.17

Temukan nilai x dari persamaan di bawah ini.

$$x = \frac{(4-2)^3 + (5+7)^2}{(5-3)^2}$$

Jawab:

$$x = \frac{(4-2)^3 + (5+7)^2}{(5-3)^2}$$

$$x = \frac{2^3 + 12^2}{2^2} = \frac{8 + 144}{4} = 38$$



Latihan

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

Cocokkan jawaban di bawah ini dengan cara menarik garis pada jawaban yang sesuai.

1)

Jarak Jakarta ke
Yogyakarta

Warna pakaian

Jumlah pakaian
dalam lemari

Kualitatif

Diskrit

Kontinu

2)

Pengaruh pemberian
ASI terhadap bayi
prematurn

Karakteristik
vegetasi tepi pantai

Statistika deskriptif

Statistika inferensial

3)

Jarak Jakarta ke Yogyakarta

Jawaban benar dan salah

Urutan preferensi konsumen pada merek deterjen

Perbedaan nilai IQ pada individu A dan B

Skala nominal

Skala ordinal

Skala interval

Skala rasio

Petunjuk Jawaban Latihan

1)

Jarak Jakarta ke Yogyakarta

Warna pakaian

Jumlah pakaian dalam lemari

Kualitatif

Diskrit

Kontinu

2)

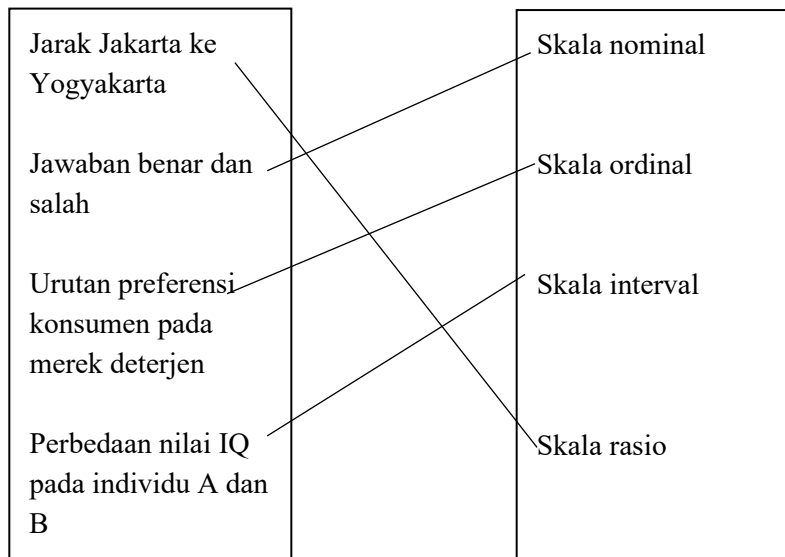
Pengaruh pemberian ASI terhadap bayi prematur

Karakteristik vegetasi tepi pantai

Statistika deskriptif

Statistika inferensial

3)



Rangkuman

1. Statistika adalah metode yang digunakan dalam mengumpulkan data, pemrosesan data, dan diakhiri dengan penginterpretasian hasil dan penarikan kesimpulan.
2. Statistik adalah hasil atau informasi berupa angka yang kita peroleh dari metode statistika yang kita lakukan dimulai dari mengumpulkan data, pemrosesan data, dan diakhiri dengan penginterpretasian hasil dan penarikan kesimpulan.
3. Statistika deskriptif adalah bagian dari ilmu statistika yang berisi metode-metode penyampaian data dan kemudian menyimpulkannya sehingga dapat lebih mudah dimengerti.
4. Statistika inferensial adalah metode penarikan kesimpulan dari keseluruhan populasi hanya dengan menggunakan sampel yang kita tentukan.
5. Data sesuai dengan cara diperolehnya dapat dikelompokkan menjadi dua, yaitu data primer dan data sekunder. Data primer adalah data yang dikumpulkan secara langsung oleh si pelapor data dari sumber aslinya. Data sekunder adalah data yang dikumpulkan dari pihak lain yang telah mengalami pemrosesan dan perhitungan.
6. Sesuai dengan bentuknya, data dapat kita bedakan menjadi data kualitatif dan data kuantitatif. Data kuantitatif adalah data berbentuk angka atau yang dapat diangkakan, contohnya jarak, umur, harga, tinggi badan, dan sebagainya. Sementara itu, data kualitatif adalah data yang berbentuk uraian atau yang berbentuk selain angka, contohnya jenis kelamin, pekerjaan, alamat domisili, warna kulit, etnis, dan sebagainya.
7. Data kuantitatif dapat dibagi menjadi dua, yaitu data diskrit (*discrete*) dan data kontinu (*continuous*). Data diskrit adalah data yang berbentuk angka bulat (bukan pecahan), sedangkan data kontinu adalah data yang memiliki nilai terus-menerus (dapat berbentuk pecahan).
8. Variabel bebas (variabel independen) adalah variabel yang perubahan nilainya bisa memengaruhi nilai variabel terikat (variabel dependen). Variabel terikat adalah variabel yang nilainya dipengaruhi oleh perubahan nilai variabel bebas.

9. Skala pengukuran digunakan untuk pengategorian atau penguantifikasian suatu variabel. Penggunaan skala pengukuran yang tepat akan membantu kita menghasilkan pengukuran yang akurat terhadap suatu permasalahan tertentu.
10. Terdapat empat skala pengukuran yang digunakan dalam statistika, dimulai dari yang terlemah ke yang terkuat, yaitu nominal, ordinal, interval, dan rasio.
11. Skala nominal adalah pengelompokan objek-objek penelitian ke dalam kategori-kategori tertentu yang telah ditetapkan.
12. Skala ordinal adalah skala yang menunjukkan *ranking* dari hal yang kita amati.
13. Skala interval adalah skala yang menyatakan jarak absolut dari dua pengukuran.
14. Skala rasio adalah skala yang bisa menunjukkan nilai dan rasio dari data.



Tes Formatif 1

Pilihlah satu jawaban yang paling tepat!

- 1) Kumpulan metode pengumpulan, pengolahan, dan penginterpretasian data disebut
 - A. statistik
 - B. data
 - C. statistika
 - D. observasi
- 2) Yang merupakan data diskrit adalah
 - A. jumlah orang
 - B. berat barang
 - C. alamat rumah
 - D. kode barang

Persamaan berikut ini untuk soal nomor 3 dan 4.

$$A = \alpha_0 + \alpha_1 B_1 + \alpha_2 B_2 + e$$

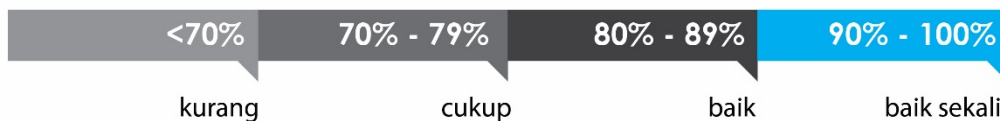
- 3) Berikut ini yang merupakan koefisien adalah
 - A. α_0
 - B. α_1
 - C. A
 - D. B

- 4) Berikut ini yang merupakan variabel terikatnya adalah
- A. A
 - B. B_1
 - C. B_2
 - D. e
- 5) Tentukan x jika $\frac{x^2}{4} = 9$
- A. 36
 - B. 9
 - C. 6
 - D. 4

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 1 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 1.

$$\text{Tingkat Penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100$$

Arti tingkat penguasaan



Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan Kegiatan Belajar 2. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 1, terutama bagian yang belum dikuasai.

Pengumpulan Data dan Teknik *Sampling*

A. METODE PENGUMPULAN DATA

Bagi peneliti, data bisa dikumpulkan dengan dua cara sehingga kita mengenal dua macam data, yaitu data primer dan data sekunder. Data sekunder adalah data yang dikumpulkan dari suatu sumber data yang telah diolah dan dipublikasikan oleh pihak lain. Pada suatu titik tertentu, semua data tentu berasal dari sumber pertama (primer), yaitu data dikumpulkan dari sumber pertama dan kemudian diolah oleh pengumpul data.

Sesuai dengan lingkungannya, data bisa dikumpulkan dengan dua macam cara, yaitu sensus dan *sampling*. Keduanya akan dijelaskan di bawah ini.

1. **Sensus** adalah metode pengumpulan data yang melibatkan seluruh populasi yang akan diteliti. Karena memakan biaya dan sumber daya yang besar, sensus biasanya dilakukan oleh negara. Contoh sensus yang dilakukan secara reguler oleh negara (dalam hal ini BPS) adalah sensus penduduk, sensus ekonomi, sensus pertanian, dan sebagainya.
2. **Sampling** adalah metode pengumpulan data menggunakan sebagian kecil dari keseluruhan populasi yang diamati. Pada umumnya, peneliti tidak memiliki sumber daya yang cukup untuk melakukan pengambilan data secara sensus sehingga teknik *sampling* ini sangat lazim dilakukan. Pembahasan tentang *sampling* ini akan disajikan secara lebih lengkap di bagian selanjutnya.

Berikut ini adalah beberapa metode pengumpulan data primer yang sering dilakukan.

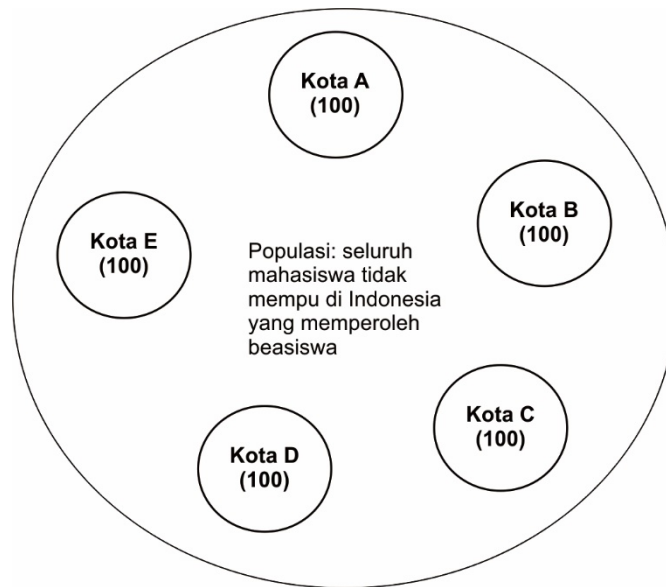
1. **Observasi** adalah cara pengumpulan data dengan pengamatan, yaitu pengumpul data atau peneliti tidak bisa memengaruhi nilai atau respons yang diberikan oleh responden. Pada umumnya, penelitian-penelitian bidang ekonomi menggunakan cara observasi ini walaupun saat ini sudah berkembang penelitian ekonomi eksperimental.
2. **Percobaan (eksperimen)** adalah cara mengumpulkan data ketika para responden diberikan beberapa perlakuan yang berbeda dan kemudian diamati pengaruh perbedaan perlakuan itu terhadap hasil yang diperoleh. Teknik eksperimental ini sangat lazim digunakan untuk penelitian-penelitian bidang eksakta, seperti penelitian di bidang kedokteran, biologi, fisika, dan sebagainya.

3. **Wawancara** adalah tanya jawab langsung yang dilakukan oleh pengumpul data dan sumber data (narasumber). Wawancara bisa dilakukan langsung (berhadapan) ataupun tidak langsung, misalnya melalui percakapan telepon atau *video conference*. Pemilihan antara wawancara langsung dan tidak langsung tentunya sangat terkait dengan jarak antara pewawancara dan narasumber serta ketersediaan biaya. Wawancara secara langsung memiliki kelebihan, yaitu informasi dapat diperoleh secara lebih lengkap karena dengan wawancara langsung si pewawancara bisa menangkap tidak hanya bahasa lisan, tetapi juga bahasa tubuh dari narasumber. Selain itu, informasi yang tidak jelas juga bisa langsung bisa diklarifikasi sehingga tidak menimbulkan kesalahan interpretasi. Kelemahan wawancara langsung adalah biayanya mahal dan memakan waktu yang lama. Selain itu, kelemahan lainnya adalah kemungkinan akan terjadi bias jawaban dari narasumber karena pertanyaan yang diarahkan atau narasumber yang merasa terintimidasi apabila jawaban yang diberikan tidak sesuai dengan pendapat yang berlaku umum.
4. **Kuesioner** adalah sekumpulan pertanyaan penelitian yang dibuat secara tertulis dan membutuhkan respons secara tertulis pula. Ada dua macam pertanyaan yang diajukan dalam kuesioner, yaitu berupa pertanyaan tertutup dan pertanyaan terbuka. Pertanyaan tertutup adalah pertanyaan yang sudah memiliki alternatif jawaban sehingga para responden hanya diminta untuk memilih jawaban yang paling mewakili mereka. Pertanyaan terbuka adalah pertanyaan yang tidak memiliki pilihan jawaban sehingga para responden dibebaskan menjawab apa saja sesuai dengan jawaban yang akan mereka utarakan. Kelebihan dari teknik ini adalah bisa mengurangi kemungkinan kesalahan yang dilakukan oleh pewawancara dan yang diwawancarai. Kelemahannya sebagai berikut. Pertama, tidak adanya kontrol dari peneliti apakah kuesioner dijawab oleh responden yang bersangkutan (jika kuesioner dikirim melalui pos). Kedua, tingkat respons balik yang rendah terutama terjadi jika kuesioner dikirim melalui pos. Terakhir adalah peneliti tidak bisa mengklarifikasi jawaban responden yang kurang jelas.

B. POPULASI DAN SAMPEL

Sesuai dengan lingkup objek penelitiannya, kita mengenal dua macam istilah, yaitu populasi dan sampel. Misalnya, seorang peneliti akan melakukan penelitian dengan topik “Pengaruh Pemberian Beasiswa terhadap Prestasi Belajar Mahasiswa Tidak Mampu di Indonesia”. Yang menjadi populasi dari penelitian ini adalah seluruh mahasiswa tidak mampu di Indonesia yang menerima beasiswa. Namun, karena keterbatasan sumber daya, penelitian hanya akan dilakukan di lima kota besar dengan jumlah sampel di masing-masing kota dibatasi hanya 100 mahasiswa saja atau secara keseluruhan menjadi 500 mahasiswa saja. Dari penjelasan tersebut, dapat kita simpulkan bahwa populasi adalah sekelompok unit atau objek penelitian yang memiliki

karakteristik tertentu yang akan kita pelajari. Sementara itu, sampel adalah bagian dari populasi yang akan kita pelajari yang dianggap bisa mewakili keseluruhan populasi. Teknik pengambilan data sampel ini disebut dengan istilah *sampling*. Untuk memperjelas, perhatikan gambar di bawah ini.



Gambar 1.2
Populasi dan Sampel

Pengambilan data populasi ini sangat mungkin dilakukan untuk penelitian dengan lingkup kecil atau penelitian yang melibatkan objek penelitian yang jumlahnya kecil. Namun, untuk penelitian dengan lingkup yang sangat besar, penggunaan teknik *sampling* sangat umum dilakukan. Walaupun harus selalu diingat bahwa dengan melakukan *sampling*, kita juga harus memperhatikan kemungkinan kesalahan (*sampling error*) yang harus tetap diperhitungkan dalam pengambilan keputusan. Beberapa alasan mengapa teknik *sampling* sangat umum dilakukan sebagai berikut.

1. **Murah**
Pengambilan data dengan menggunakan metode sampel tentu akan jauh menurunkan biaya penelitian.
2. **Cepat**
Di samping biaya yang lebih murah, teknik ini juga akan mempercepat waktu pengambilan data.
3. **Akurat**
Pada beberapa kasus dengan menggunakan sampel, data yang kita peroleh akan menjadi lebih akurat. Untuk penelitian-penelitian yang membutuhkan partisipasi berulang dari responden, kemungkinan kesalahan akan lebih besar pada jumlah observasi yang besar.

4. Efek samping

Penelitian dengan kemungkinan efek samping yang buruk tidak mungkin dilakukan dengan menggunakan keseluruhan populasi, contohnya adalah penelitian daya tahan mobil dari tabrakan dan penelitian tentang pengaruh obat baru terhadap manusia.

5. Jumlah populasi yang tidak jelas

Adakalanya jumlah responden dapat terus bertambah atau berkurang sebelum masa penelitian berakhir. Kita ambil contoh penelitian mengenai beasiswa terhadap mahasiswa di atas. Pada awal penelitian dirancang, mungkin jumlah populasi mahasiswa adalah sebanyak n orang. Pada perjalanannya, jumlah populasi mahasiswa dapat bertambah atau berkurang karena faktor-faktor yang tidak bisa dikontrol oleh peneliti.

Walaupun lebih sulit dilakukan, beberapa penelitian dengan skala besar tetap harus menggunakan data populasi. Salah satu contoh pengambilan data populasi dengan skala besar adalah sensus yang dilakukan oleh pemerintah (sensus penduduk, sensus ekonomi, sensus pertanian, dan sebagainya).

C. MACAM-MACAM SAMPEL

Beberapa macam sampel yang digunakan dalam statistika sebagai berikut.

1. *Convenience sample* adalah teknik pengambilan data sampel yang paling mudah dijangkau oleh peneliti. Teknik ini kemungkinan besar akan menghasilkan kesimpulan penelitian yang tidak akurat.

Contoh 1.18

Menghentikan pejalan kaki di suatu jalan dan menanyakan pertanyaan penelitian.

2. *Stratified sample*, yaitu teknik sampel ketika sampel dikelompok-kelompokkan ke dalam dua atau lebih strata.

Contoh 1.19

Pengambilan sampel pegawai dari suatu kementerian yang digolongkan berdasar tingkat pendidikannya, yaitu ada pegawai dengan tingkat pendidikan SLTA, diploma, S-1, S-2, dan S-3.

3. *Systematic sample* adalah teknik pengambilan sampel ketika peneliti menentukan sampel berdasarkan urutan yang sama jaraknya.

Contoh 1.20

Terdapat 1000 orang populasi dari penelitian yang akan dilakukan. Peneliti menentukan bahwa yang menjadi sampelnya adalah orang ke-10 secara berurutan. Apabila setiap individu diurutkan secara berurutan, yang menjadi sampel penelitian adalah orang ke-10, ke-20, ke-30, ke-40, dan seterusnya sampai ke-1000.

4. *Cluster sample* adalah pengelompokan sampel berdasarkan letak geografisnya. Kelompok geografis tersebut lalu dibagi lagi ke dalam area-area lebih kecil. Dalam setiap area tersebut, kita mengambil beberapa sampel secara acak atau yang disebut klaster (*cluster*). Seluruh populasi yang ada di klaster tersebut lalu kita observasi.

Contoh 1.21

Dilakukan survei tentang pengeluaran rumah tangga di suatu kabupaten. Dari tujuh kecamatan yang ada di kabupaten tersebut, dipilih tiga kabupaten saja. Untuk setiap kecamatan yang terpilih, dipilih dua keluarahan. Jadi, jumlah seluruh keluarahan yang terpilih adalah enam keluarahan. Seluruh keluarga yang ada di enam kelurahan tersebut merupakan sampel dari survei ini.

5. *Representative sample* adalah teknik pengambilan sampel berdasarkan keterwakilan (representasi) dari setiap golongan individu yang kita teliti.

Contoh 1.22

Dilakukan penelitian tentang besarnya pengeluaran untuk pangan dari rumah tangga. Maka kita akan mengambil beberapa perwakilan sampel dari setiap golongan pendapatan: golongan dengan pendapatan tinggi, golongan pendapatan menengah, dan golongan pendapatan rendah.

Yang harus kita perhatikan dari teknik ini adalah proporsi setiap sampel terhadap populasi golongannya. Apabila sampel yang kita ambil tidak proporsional, kesimpulan yang kita ambil akan menjadi tidak valid.

6. *Random sample* atau sampel acak adalah metode pengambilan sampel yang dianggap paling akurat karena setiap anggota populasi memiliki kesempatan yang sama untuk terpilih sebagai sampel.

Contoh 1.23

Ketika berlangsungnya pemilihan presiden (pilpres), banyak sekali lembaga survei yang melakukan survei mengenai pemenang pilpres dengan hasil yang berbeda-beda. Hal ini disebabkan oleh data yang dikumpulkan dari beberapa lembaga survei sangat terkonsentrasi pada kelompok tertentu sehingga hasil yang diperoleh tidak mencerminkan hasil akhir pilpres.

Pengacakan dengan jumlah populasi yang lebih besar dapat dilakukan dengan bantuan komputer, contohnya pengundian hadiah dari bank dan nomor lotere.

D. BIAS SAMPEL

Walaupun tanpa melibatkan seluruh populasi, penggunaan sampel seharusnya tetap dapat menggambarkan keseluruhan populasi yang diwakilinya. Namun demikian, tidak ada jaminan dengan menggunakan teknik sampel ini jawaban penelitian yang kita dapatkan akan benar-benar mewakili jawaban seluruh populasi. Salah satu penyebabnya adalah adanya bias sampel. Bias sampel dapat disebabkan oleh pemilihan sampel yang terkonsentrasi pada sekelompok responden dengan karakteristik yang sama sehingga jawaban penelitian yang kita peroleh akan mencerminkan jawaban penelitian kelompok tersebut bukan keseluruhan populasi.

Beberapa cara yang bisa dilakukan untuk menghindari bias sampel ini adalah memperbesar jumlah sampel, pemilihan teknik sampel yang benar dan sesuai dengan tujuan penelitian, serta menghindari subjektivitas dalam penelitian.

Contoh 1.24

Contoh paling sederhana dari *random sampling* ini adalah pengundian *doorprize*, yaitu setiap peserta suatu acara berhak atas satu nomor undian yang kemudian diacak.

E. KESALAHAN SAMPEL (*SAMPLING ERROR*)

Walaupun pemilihan dan teknik sampel sudah dilakukan dengan cara yang benar, hasil suatu penelitian tetap akan menghasilkan suatu kesalahan (*error*) karena tidak mungkin sampel yang kita gunakan akan secara sempurna mewakili keseluruhan populasi. Kesalahan sampel (*sampling error*) adalah selisih antara estimasi yang

dihasilkan oleh sampel dengan estimasi populasi. *Sampling error* ini biasanya diakomodasi dalam model regresi dan diletakkan di bagian paling kanan dari persamaan regresi tersebut. Perhatikan persamaan regresi di bawah ini.

$$Y = c + \alpha_1 x_1 + \alpha_2 x_2 + e$$

Sampling error dinotasikan oleh e . Secara matematis, *sampling error* bisa dinyatakan dalam persamaan berikut.

$$\text{error}(e) = |\hat{\theta} - \theta|$$

Perlu diketahui, $\text{error}(e)$ adalah *sampling error*, $\hat{\theta}$ adalah estimasi yang dihasilkan oleh sampel, dan θ parameter dari populasi. Nilai dari *sampling error* ini dapat kita perkecil dengan teknik pemilihan sampel yang tepat, tidak bias, dan dengan menggunakan sebanyak mungkin sampel. Beberapa hal berikut ini juga bisa berkontribusi terhadap bias.

1. Pertanyaan yang bersifat sensitif sehingga memungkinkan responden untuk menjawab secara tidak jujur.
2. Kurangnya respons dari kelompok tertentu. Misalnya, apabila survei dilakukan pada jam kerja, kemungkinan untuk memperoleh responden laki-laki dewasa akan kecil.
3. Tidak bersedia menjawab pertanyaan atas kemauan sendiri. Hal ini bisa terjadi apabila pertanyaan yang diajukan diramalkan akan menimbulkan kontroversi sehingga sebagian responden tidak bersedia menjawab pertanyaan penelitian.
4. Bias yang dihasilkan oleh pewawancara bisa disebabkan oleh pertanyaan yang terlalu diarahkan atau karakteristik dari pewawancara. Misalnya, seorang responden pria akan merasa tidak nyaman diwawancarai oleh pewawancara wanita apabila pertanyaan yang diajukan adalah seputar masalah reproduksi sehingga kemungkinan jawaban yang diutarakan juga akan tidak tepat.
5. Sebab-sebab teknis lainnya, seperti kesalahan perekaman, kesalahan pencatatan, dan sebagainya.



Latihan

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

- 1) Apakah penelitian di bawah ini lebih tepat dilakukan dengan teknik sensus atau *sampling*? Jelaskan alasannya!
 - a) Lama belajar siswa kelas 4 di SD X.
 - b) Persepsi ibu-ibu rumah tangga terhadap tiga merek sabun pencuci piring.

- 2) Teknik apa yang paling cocok digunakan untuk mengumpulkan data berikut ini? Cocokkan jawaban Anda!

Pengaruh imunisasi X terhadap daya tahan tubuh anak usia di bawah 5 tahun

Survei kepuasan mahasiswa terhadap tutorial *online*

Pendapat ahli olahraga terhadap jalannya pertandingan sepak bola antara negara A dan negara B

Wawancara

Eksperimen

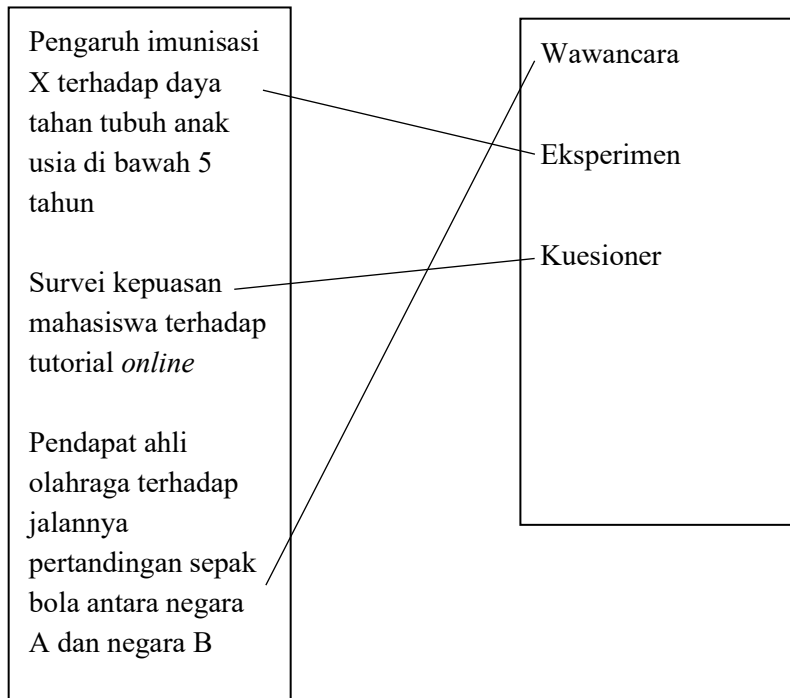
Kuesioner

- 3) Apakah penyebab bias dari hal-hal berikut ini?
- Guru melakukan wawancara terhadap siswa-siswanya mengenai kinerja guru yang bersangkutan.
 - Wawancara mengenai pendapat ibu-ibu rumah tangga terhadap harga kebutuhan pokok di pasar tradisional.
 - Penelitian mengenai kadar gula darah penduduk kota Z yang menggunakan sampel mayoritas pria usia di atas 60 tahun.

Petunjuk Jawaban Latihan

- Populasi karena jumlah anggota populasi yang kecil.
 - Sampling* karena jumlah populasi ibu-ibu dan jumlah pasar tradisional yang besar jumlahnya.

2)



- 3)
- Siswa kemungkinan besar akan berbohong dan memberikan jawaban yang diinginkan guru karena merasa terintimidasi.
 - Jawaban kemungkinan akan cenderung seragam dengan menyatakan harga kebutuhan pokok semua mahal karena kurang objektif.
 - Sampel yang terkonsentrasi di kelompok tertentu sehingga kurang proporsional karena penduduk kota Z tentu terdiri atas laki-laki dan perempuan dengan berbagai golongan tingkat umur.



Rangkuman

- Sesuai dengan lingkungannya, data bisa dikumpulkan dengan dua macam cara, yaitu sensus dan *sampling*.
- Sensus adalah metode pengumpulan data yang melibatkan seluruh populasi yang akan diteliti. Karena memakan biaya dan sumber daya yang besar, sensus biasanya dilakukan oleh negara. Contoh sensus yang dilakukan secara reguler oleh negara (dalam hal ini BPS) adalah sensus penduduk, sensus ekonomi, sensus pertanian, dan sebagainya.
- Sampling* adalah metode pengumpulan data menggunakan sebagian kecil dari keseluruhan populasi yang diamati. Pada umumnya, peneliti tidak memiliki sumber daya yang cukup untuk melakukan pengambilan data secara sensus sehingga teknik *sampling* ini sangat lazim dilakukan.
- Sesuai dengan metode atau tekniknya, data bisa dikumpulkan dengan cara observasi, percobaan, wawancara, dan dengan menggunakan kuesioner.

5. Observasi adalah cara pengumpulan data dengan pengamatan ketika pengumpul data atau peneliti tidak bisa memengaruhi nilai atau respons yang diberikan oleh responden.
6. Percobaan (eksperimen) adalah cara mengumpulkan data ketika para responden diberikan beberapa perlakuan yang berbeda dan kemudian diamati pengaruh perbedaan perlakuan itu terhadap hasil yang diperoleh.
7. Wawancara adalah tanya jawab langsung yang dilakukan oleh pengumpul data dan sumber data (narasumber). Wawancara bisa dilakukan langsung (berhadapan) ataupun tidak langsung misalnya melalui percakapan telpon atau *video conference*. Pemilihan antara wawancara langsung dan tidak langsung tentunya sangat terkait dengan jarak antara pewawancara, narasumber, dan ketersediaan biaya.
8. Kuesioner adalah sekumpulan pertanyaan penelitian yang dibuat secara tertulis dan membutuhkan respons secara tertulis pula.
9. Sesuai dengan lingkup objek penelitiannya kita mengenal dua macam istilah, yaitu populasi dan sampel. Populasi adalah sekelompok unit atau objek penelitian yang memiliki karakteristik tertentu yang akan kita pelajari. Sementara itu, sampel adalah bagian dari populasi yang akan kita pelajari yang dianggap bisa mewakili keseluruhan populasi.
10. Alasan mengapa teknik *sampling* sangat umum dilakukan adalah murah, cepat, akurat, mengurangi efek samping, dan mengurangi kemungkinan perubahan jumlah populasi.
11. Beberapa macam sampel yang digunakan dalam statistika adalah *convenience sample*, *stratified sample*, *systematic sample*, *cluster sample*, *representative sample*, dan *random sample*.
12. *Convenience sample* adalah pengambilan sampel yang paling mudah dijangkau.
13. *Stratified sample*, yaitu teknik sampel ketika sampel dikelompokkan ke dalam dua atau lebih strata.
14. *Systematic sample* adalah teknik pengambilan sampel ketika peneliti menentukan sampel berdasarkan urutan yang sama jaraknya.
15. *Cluster sample* adalah pengelompokan sampel berdasarkan letak geografisnya. Kelompok geografis tersebut lalu dibagi lagi ke dalam area-area lebih kecil. Dalam setiap area tersebut, kita mengambil beberapa sampel secara acak atau yang disebut klaster (*cluster*). Seluruh populasi yang ada di klaster tersebut lalu kita observasi.
16. *Representative sample* adalah teknik pengambilan sampel berdasarkan keterwakilan (representasi) dari setiap golongan individu yang kita teliti.
17. *Random sample* atau sampel acak adalah metode pengambilan sampel yang dianggap paling akurat karena setiap anggota populasi memiliki kesempatan yang sama untuk terpilih sebagai sampel.
18. Bias sampel adalah pemilihan sampel yang terkonsentrasi pada sekelompok responden dengan karakteristik yang sama sehingga jawaban penelitian yang kita peroleh akan mencerminkan jawaban penelitian kelompok tersebut bukan keseluruhan populasi.
19. Kesalahan sampel (*sampling error*) adalah selisih antara estimasi yang dihasilkan oleh sampel dengan estimasi populasi.



Tes Formatif 2

Pilihlah satu jawaban yang paling tepat!

- 1) Metode pengumpulan data dengan menggunakan beberapa perlakuan dikenal dengan istilah
 - A. observasi
 - B. eksperimen
 - C. wawancara
 - D. kuesioner

- 2) Pertanyaan pada soal no. 1 merupakan jenis pertanyaan
 - A. tertutup
 - B. terbuka
 - C. terstruktur
 - D. terintegrasi

- 3) Pengetesan terhadap daya tahan lampu merek X dilakukan dengan menggunakan sampel karena alasan
 - A. lebih mudah
 - B. lebih akurat
 - C. daya rusak
 - D. populasi tidak terbatas

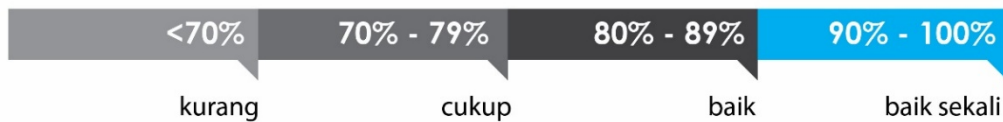
- 4) Teknik sampel yang paling mudah dilakukan adalah
 - A. *stratified sample*
 - B. *systematic sample*
 - C. *cluster sample*
 - D. *convenience sample*

- 5) Pertanyaan mengenai penghasilan biasanya dilakukan dengan menggunakan pilihan *range* penghasilan tertentu. Hal ini disebabkan pertanyaan mengenai penghasilan dinilai
 - A. sensitif
 - B. sulit diukur
 - C. intimidasi
 - D. sangat subjektif

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 2 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 2.

$$\text{Tingkat Penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100$$

Arti tingkat penguasaan



Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan modul selanjutnya. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 2, terutama bagian yang belum dikuasai.

Kunci Jawaban Tes Formatif

Tes Formatif 1

- 1) C
- 2) A
- 3) B
- 4) A
- 5) C

Tes Formatif 2

- 1) B
- 2) A
- 3) C
- 4) D
- 5) A

Daftar Pustaka

- Brase, C.H., & Brase, C.P. (2004). *Understanding basic statistics*. Boston: Houghton Mifflin.
- Carlson, K.A., & Winqvist, J.R. (2014). *An introduction to statistics: An active learning approach*. California: Sage Publications.
- Conover, W.J., & Iman R.L. (1983). *Introduction to modern business statistics*. New Jersey: John Wiley & Sons, Inc.
- Kenkel, J.A. (1995). *Introductory statistics for management and economics*. California: Duxbury Press.
- Statistika. (2017). *Dalam kamus besar bahasa Indonesia*. Dikutip dari <http://kbbi.web.id/statistika>.