

BAB I

PENDAHULUAN

1.1. Arti Kata Statistik

Istilah statistik mempunyai pengertian yang berbeda-beda bagi orang yang berbeda. Bagi seorang manajer tim sepak bola, statistik bisa dipahami sebagai frekuensi kemenangan, draw, atau kekalahan bagi tim yang dipimpinya. Sedangkan, bagi seorang manajer perusahaan, statistik bisa dipahami sebagai angka penjualan dari tahun ke tahun. Bagi seorang peneliti, statistik sering kali dianggap sebagai alat analisis data. Sedangkan, bagi mahasiswa statistik adalah mata kuliah yang harus dipelajari sebagai dasar untuk membantu pengolahan data pada saat mengambil tugas akhir.

1.1.1. Pengertian statistik

Secara etimologis kata “statistik” berasal dari kata *status* (bahasa Latin) yang mempunyai persamaan arti dengan kata *state* (bahasa Inggris) atau kata *staat* (bahasa Belanda), dan yang dalam bahasa Indonesia diterjemahkan menjadi *negara*. Pada mulanya, kata “statistik” diartikan sebagai ‘kumpulan bahan keterangan (data), baik yang berwujud angka (data kuantitatif) maupun yang tidak berwujud angka (data kualitatif), yang mempunyai arti penting dan kegunaan yang besar bagi suatu *negara*’. Namun, pada perkembangan selanjutnya, arti kata statistik hanya dibatasi pada ‘kumpulan bahan keterangan yang berwujud angka (data kuantitatif)’

saja; bahan keterangan yang tidak berwujud angka (data kualitatif) tidak lagi disebut *statistik*.

Dalam kamus bahasa Inggris akan kita jumpai kata *statistics* dan kata *statistic*. Kedua kata itu mempunyai arti yang berbeda. Kata *statistics* artinya ‘ilmu statistik’, sedang kata *statistic* diartikan sebagai ‘ukuran yang diperoleh atau berasal dari sampel’, yaitu sebagai lawan dari kata “parameter” yang berarti ‘ukuran yang diperoleh atau berasal dari populasi’.

Menurut Kamus Besar Bahasa Indonesia statistik adalah catatan angka-angka (bilangan); perangkaan; data yang berupa angka yang dikumpulkan, ditabulasi, digolong-golongkan sehingga dapat memberi informasi yang berarti mengenai suatu masalah atau gejala.

Banyak persoalan dinyatakan dan dicatat dalam bentuk bilangan atau angka-angka. Kumpulan angka-angka itu sering disusun atau disajikan dalam bentuk daftar atau tabel. Sering daftar atau tabel tersebut disertai dengan gambar-gambar, dan disebut dengan **statistik**. Jadi, kata statistik telah dipakai untuk menyatakan kumpulan data, bilangan maupun nonbilangan yang disusun dalam tabel dan atau diagram, yang menggambarkan suatu persoalan. Statistik yang menjelaskan sesuatu hal biasanya diberi nama statistik mengenai hal yang bersangkutan, misalnya statistik penduduk, statistik kelahiran, dan lain sebagainya.

Kata statistik juga digunakan untuk menyatakan **ukuran** sebagai wakil dari sekumpulan data mengenai sesuatu hal, misalnya 60% siswa nilai matematikanya kurang dari 6,5 maka nilai 60% ini dinamakan statistik. Misalnya lagi jika rata-rata nilai keseluruhan dalam rapor seorang siswa adalah 7,5 maka rata-rata 7,5 tersebut dinamakan statistik.

1.1.2. Pengertian statistika :

Menurut Kamus Besar Bahasa Indonesia, statistika adalah ilmu tentang cara mengumpulkan, menabulasi, menggolong-golongkan, menganalisis, dan mencari keterangan yang berarti dari data yang berupa angka; pengetahuan yang berhubungan dengan pengumpulan data, penyelidikan dan kesimpulannya berdasarkan bukti, berupa catatan bilangan (angka-angka).

Dari hasil pengamatan atau penelitian, dalam laporannya sering diperlukan suatu uraian, penjelasan, atau kesimpulan tentang persoalan yang diamati atau diteliti. Sebelum membuat kesimpulan, keterangan atau data yang terkumpul terlebih dahulu dipelajari, diolah, atau dianalisis, dan berdasarkan pengolahan data inilah baru dibuat kesimpulan. Mulai dari pengumpulan data, pengolahan data, dan pengambilan kesimpulan haruslah mengikuti cara-cara yang benar dan dapat dipertanggungjawabkan. Ini semua merupakan pengetahuan tersendiri yang dinamakan dengan statistika. Jadi, **statistika** adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisisannya, dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisisan yang dilakukan.

1.2. Sejarah Statistika

Istilah statistika sudah sangat tua. Statistika bermula sebagai suatu cara berhitung untuk membantu pemerintah yang ingin mengetahui kekayaan dan banyaknya warganya dalam usaha menarik pajak atau berperang. William si penakluk memerintahkan diadakannya survei di seluruh Inggris untuk tujuan pajak dan tugas kemiliteran. Hasil survei ini dikumpulkan dalam sebuah kumpulan yang disebut *Domesday Book*.

Beberapa abad setelah *Domesday Book*, ditemukan suatu penerapan peluang empirik dalam asuransi perkapalan, yang tampaknya sudah tersedia bagi kapal-kapal bangsa Flem pada abad ke-14. Perjudian, dalam bentuk permainan, telah mengantarkan kita ke teori peluang. Teori ini pertama kali dikembangkan oleh Pascal dan Fermat sekitar abad ke-17, karena mereka tertarik pada pengalaman-pengalaman judi Chevalier de Mere.

Kurva normal telah terbukti sangat penting dalam pengembangan statistika. Persamaan kurva ini pertama kali diumumkan pada tahun 1733 oleh de Moivre. De Moivre sama sekali tidak tahu bagaimana menerapkan penemuannya tersebut pada data hasil percobaan, dan karyanya ini tetap tidak diketahui sampai Karl Pearson menemukannya di suatu perpustakaan pada 1924. Walaupun demikian, hasil yang sama dikembangkan kemudian oleh dua astronom matematik, Laplace (1749–1855) dan Gauss (1777–1855), secara terpisah.

Pada abad ke-19, Charles Lyell telah mengajukan suatu argumentasi yang pada dasarnya bersifat statistik terhadap suatu masalah geologi. Dalam periode 1830–1833, diterbitkan 3 jilid *Principles of Geology* karya Lyell, yang mengurutkan batu-batuan zaman Tertier, serta sekaligus memberi nama pada masing-masing batuan. Bersama dengan M. Deshayes, seorang ahli biologi dari Prancis, mereka mengidentifikasi dan mendaftarkan spesies-spesies fosil yang terdapat dalam satu atau lebih strata, dan meramalkan proporsi jenis-jenis yang masih hidup di bagian-bagian laut tertentu. Berdasarkan proporsi-proporsi tersebut, mereka memberi nama Pleistosen, Pliosen, Miosen, dan Eosen. Argumentasi Lyell sesungguhnya bersifat statistika. Sayangnya, setelah ditetapkan dan diterimanya nama-nama tersebut, metodenya segera dilupakan orang. Hal ini terjadi baik di bidang ilmu-ilmu biologi maupun fisika.

Pada abad ke-19 pula, perlunya landasan yang lebih kokoh bagi statistika menjadi semakin jelas. Karl Pearson, seorang ahli fisika matematik, menerapkan matematika pada biologi. Pearson melewati hampir setengah abad dalam penelitian statistika yang serius. Di samping itu, ia juga mendirikan jurnal *Biometrika* dan sebuah aliran statistika. Dengan demikian, kajian statistika memperoleh dorongan besar.

Sementara Pearson hanya memperhatikan contoh besar (*large samples*), teori sampel besar yang dikembangkan ternyata tidak memuaskan peneliti yang selalu berhubungan dengan sampel kecil (*small samples*). Di antara mereka adalah W.S. Gosset, 1876–1937, murid Karl Pearson. Namun, kemampuan matematika Gosset belum memadai untuk mendapatkan sebaran-sebaran pasti dari simpangan baku sampel, rasio antara rata-rata sampel dengan simpangan baku sampel, dan koefisien korelasi; statistik-statistik yang paling banyak diperhatikannya. Akibatnya, ia terpaksa mendasarkan pada kartu; mengocok, mengambil, dan kemudian membuat sebaran frekuensi empiriknya. Makalah yang membuat hasil penelitiannya ini muncul dalam *Biometrika* pada 1908, dan ia menggunakan nama *student*. Sekarang ini sebaran *t-Student* merupakan alat dasar bagi statistikawan dan peneliti; dan *student*-kan merupakan istilah yang lazim dalam statistika. Kini penggunaan sebaran *t-Student* begitu meluas, dan menarik untuk diperhatikan bahwa seorang astronom Jerman, Helmert, telah mendapatkannya secara matematika jauh sebelumnya, yaitu pada 1875.

R.A. Fisher (1890–1962), yang dipengaruhi oleh Karl Pearson dan Student, memberikan sumbangan yang sangat banyak dan penting bagi statistika. Ia dan murid-muridnya memberikan dorongan yang besar bagi penggunaan prosedur-prosedur statistika dalam banyak bidang, terutama dalam bidang-bidang

pertanian, biologi, dan genetika.

J. Neyman (1895) dan E.S.Pearson (1895), mengemukakan teori pengujian hipotesis pada 1936 dan 1938. Teori ini merancang sejumlah besar penelitian dan banyak hasilnya mempunyai kegunaan praktis.

Pada 1902–1950, Abraham Wald menulis dua buku yang sangat bermanfaat hingga saat ini, yakni '*Sequential Analysis*' dan '*Statistical Decision Functions*'. Dalam abad inilah (hingga saat ini) hampir semua metode statistika yang kini digunakan itu dikembangkan.

1.3. Pembagian Statistika

Sebagaimana telah dijelaskan pada subbab sebelumnya bahwa statistika adalah pengetahuan tentang bagaimana memahami data. Untuk memahami data dibutuhkan beberapa aktivitas terkait dengan data tersebut yang dimulai dengan mengumpulkan data, menyajikan data, membuat ringkasan data untuk selanjutnya dilakukan pengolahan data dan penganalisisan data, membuat estimasi/pendugaan mengenai ukuran statistik yang pada akhirnya membuat kesimpulan secara general (populatif) mengenai ukuran statistik yang diperoleh dari data yang diambil secara sampling. Berdasarkan kegiatan-kegiatan tersebut statistika kemudian dibagi menjadi dua bagian, yaitu Statistika Deskriptif dan Statistika Induktif/Inferensial.

1.3.1. Statistika deskriptif

Statistika deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian sekumpulan data sehingga dapat memberikan informasi yang berguna. Perlu kiranya dipahami bahwa statistika deskriptif memberikan

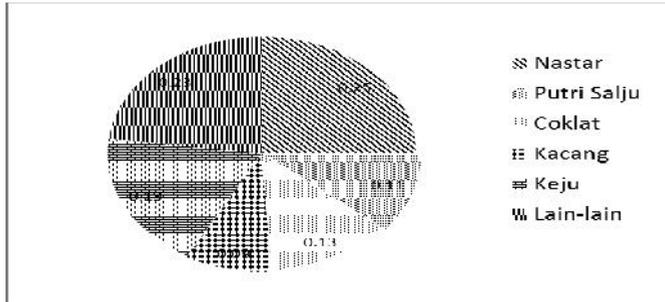
informasi hanya mengenai data yang dipunyai dan sama sekali tidak menarik kesimpulan yang lebih banyak dan lebih jauh dari data yang ada. Kegiatan memeriksa sifat-sifat atau karakteristik penting dari data yang ada itu disebut analisis data secara deskripsi. Karenanya, bagian statistika demikian dinamakan Statistika Deskriptif. Penyusunan tabel, diagram, modus, kuartil, simpangan baku termasuk dalam kategori statistika deskriptif. Kegiatan itu dilakukan melalui:

- a. Pendekatan aritmetika, yaitu pendekatan melalui pemeriksaan rangkuman nilai atau ukuran-ukuran penting dari data. Yang dimaksud rangkuman nilai di sini ialah penyederhanaan kumpulan nilai data yang diamati ke dalam bentuk nilai-nilai tertentu. Setiap rangkuman nilai ini disebut statistik. Jadi, statistik menerangkan sifat kumpulan data dalam bentuk nilai yang mudah dipahami, sedangkan statistika adalah suatu ilmu tentang sekumpulan konsep serta metode yang dapat digunakan untuk mengumpulkan, menyajikan, dan menganalisis data serta menarik kesimpulan berdasar hasil analisis data tersebut.
- b. Pendekatan geometrik, yaitu melalui penyajian data dalam bentuk gambar berupa grafik atau diagram. Kedua pendekatan mengakibatkan perbedaan dalam penyajian datanya. Penyajian data pertama menekankan angka-angka dan yang kedua menekankan pada gambar.

Contoh Penggunaan Statistika Deskriptif

Sejak tahun 2000, Ina Cookies telah menjual kue kering dengan berbagai varian rasa. Saat ini ada 8 jenis kue yang dijual, yaitu nastar, putri salju, cokelat, kacang, keju, sagu, lidah kucing, dan semprit. Penjualan yang telah berhasil dilakukan

oleh Ina Cookies mendekati 150 ribu toples selama 2014 dalam berbagai varian rasa. Ringkasan data disajikan dalam gambar 1.1 berikut ini:



Gambar 1.1 Volume Penjualan kue kering Ina Cookies

Berdasarkan grafik tersebut, secara jelas terlihat bahwa penjualan kue nastar (25%) merupakan penjualan tertinggi dari berbagai jenis kue yang dijual oleh Ina Cookies dan diikuti oleh kue keju sebesar 19% dan cokelat sebesar 13%. Ketika gambar 1.1 menjelaskan volume terjual dari berbagai katagori kue, ini merupakan contoh penggunaan statistika deskriptif. Penggunaan statistika deskriptif dapat pula digunakan untuk menjelaskan data yang disajikan dalam tabel. Contoh penyajian data ke dalam tabel disajikan pada Tabel 1.1:

Tabel 1.1 Jumlah Pelajar di Daerah A

Jenis Kelamin	SD	SMP	SMA	Jumlah
Laki-laki	4.758	2.795	1.459	9.012
Perempuan	4.032	2.116	1.256	7.404
Jumlah	8.790	4.911	2.715	16.416

1.3.2. Statistika Inferensial

Statistik inferensial merupakan kebalikan dari statistika deskriptif, statistika inferensial merupakan statistik yang berkenaan dengan cara penarikan kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan karakteristik atau ciri dari suatu populasi. Dengan demikian dalam statistik inferensial dilakukan suatu generalisasi dan hal yang bersifat khusus (kecil) ke hal yang lebih luas (umum). Oleh karena itu, statistik inferensial disebut juga statistik induktif atau statistik penarikan kesimpulan. Pada statistik inferensial biasanya dilakukan pengujian hipotesis dan pendugaan mengenai sifat (ciri) dari suatu populasi, seperti mean dan Uji t (Sugiyono, 2006). Dengan demikian, Statistika Inferensial menyimpulkan makna statistik yang telah dihitung, dianalisis, atau disajikan grafik atau diagramnya tersebut. Penarikan kesimpulan tentang keseluruhan populasi-populasi yang didasarkan atas pengamatan terhadap salah satu bagian populasi disebut induksi atau generalisasi. Proses induksi atau generalisasi dalam statistika induktif dapat ditemui dalam berbagai kegiatan ilmiah dan juga dalam kehidupan sehari-hari. Misalnya, seorang anak kecil sering melihat balok-balok kayu dapat terbakar maka ia akan menarik kesimpulan bahwa semua balok kayu dapat terbakar.

Contoh Penggunaan Statistika Inferensial

Seorang dosen suatu perguruan tinggi ingin mengetahui gambaran bagaimana prestasi mahasiswanya, seluruh mahasiswa yang terdaftar di perguruan tinggi tersebut ada sebanyak 5.000 orang. Untuk mengukur prestasi mahasiswa digunakan Indeks Prestasi (IP). Untuk mendapat gambaran mengenai prestasi mahasiswa tersebut maka dosen tersebut

harus mengumpulkan data IP seluruh maha-siswa kemudian diambil rata-ratanya. Ketika rata-rata IP diperoleh dengan melibatkan semua data IP mahasiswa, maka dosen tersebut sedang menerapkan statistika deskriptif. Namun, ketika dosen tersebut mengumpulkan data IP sebagian mahasiswa saja sebagai sampel sehingga diperoleh nilai dugaan rata-rata IP, dan nilai dugaan tersebut digunakan untuk menyimpulkan bagaimana prestasi seluruh mahasiswa perguruan tinggi tersebut maka dosen tersebut sedang menggunakan Statistika Inferensial.

1.4. Populasi dan Sampel

1.4.1. Populasi

Pengertian populasi pada statistika berbeda dengan pengertian populasi pada disiplin ilmu lainnya. Pada statistika, totalitas semua nilai yang mungkin, hasil menghitung, ataupun mengukur, kuantitatif maupun kualitatif mengenai sifat tertentu dari semua anggota kumpulan yang lengkap dan jelas yang ingin dipelajari sifat sifatnya disebut populasi. Pengertian lain mengenai populasi merupakan keseluruhan (universum) dari objek penelitian yang dapat berupa manusia, hewan, tumbuh-tumbuhan, gejala, nilai, peristiwa, sikap hidup, dan sebagainya yang menjadi pusat perhatian dan menjadi sumber data penelitian. Apabila kita lihat definisi tersebut, pengertian populasi bisa sangat beragam sehingga kita harus mendefinisikan populasi tersebut dengan jelas dan tepat.