

## TOPIK 5

# PENGUMPULAN DAN PENGORGANISASIAN DATA

### 5.1 Metode Pengumpulan Data

### 5.2 Klasifikasi Data



#### Reference :

Heumann, C., & Schomaker, M. (2016). *Introduction to statistics and data analysis*. Springer International Publishing Switzerland.

Ott, Lyman. (2001). *An introduction to statistical methods and data analysis*. 5th ed. Duxbury Thomson Learning.

## 5.1 Metode Pengumpulan Data

Metode pengumpulan data adalah teknik atau cara yang sistematis untuk mengumpulkan data. Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian.

Tahap Pengumpulan Data :

1. Menentukan tujuan dari eksperimen, survey, atau observasi
2. Mengidentifikasi variabel yang ingin diketahui
3. Menentukan rancangan survey atau eksperimen yang sesuai
4. Mengumpulkan data

Data adalah sesuatu yang belum memiliki arti bagi penerimanya dan masih membutuhkan adanya suatu pengolahan. Data bisa memiliki berbagai wujud, mulai dari gambar, suara, huruf, angka, bahasa, simbol, bahkan keadaan. Semua hal tersebut dapat disebut sebagai data asalkan dapat kita gunakan sebagai bahan untuk melihat lingkungan, obyek, kejadian, ataupun suatu konsep.

### 5.1.1 SURVEY

*“survey research studies large and small population (or universe) by selecting and studying samples chosen from the population to discover the relative incident, distribution, and interrelations of sociological and psychological variables”.*

- Biasanya pengumpulan data dilakukan dengan mengajukan pertanyaan atau menyediakan kuesioner untuk partisipan.
- Idealnya survey dilakukan terhadap partisipan sampel terpilih, yang merepresentasikan populasi yang diamati.
- Teknik yang dilakukan untuk menentukan sampel dari sebuah populasi disebut sebagai teknik pengambilan sampel.

#### Teknik Pengambilan Sample

1. Menentukan populasi
2. Mencari data akurat untuk unit populasi
3. Memilih sample yang representatif
4. Menentukan jumlah sample

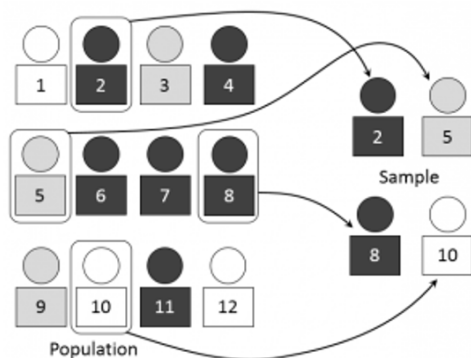
Terdapat berbagai teknik pengambilan sampel yang digunakan. Teknik sampling berdasarkan adanya randomisasi, yakni pengambilan subyek secara acak dari kumpulannya, dapat dikelompokkan menjadi 2 yaitu sampling nonprobabilitas dan sampling probabilitas. Teknik-teknik sampling tersebut dapat dilihat pada skema berikut.



## 1. Probability Sampling

Probability sampling adalah salah satu teknik pengambilan sampel yang memberikan peluang yang sama bagi setiap unsur (anggota) populasi untuk dipilih menjadi anggota sampel. Dengan probability sampling, maka pengambilan sampel secara acak atau random dari populasi yang ada.

### a. Simple Random Sampling



- Semua unit memiliki peluang yang sama untuk dipilih
- Biasanya digunakan ketika karakteristik setiap unit observasi dalam populasi relatif mirip
- Tahapannya adalah dengan membuat daftar semua anggota populasi (sampling frame), kemudian setiap anggota ditandai dengan nomor tertentu (misalnya ada jumlah populasi 12, kemudian diberi nomor dari 1 hingga 12).
- Dari populasi tersebut, peneliti memilih sampel sebanyak 4 secara acak dengan dua cara yaitu tabel bilangan acak dan software generator bilangan acak. Misal diperoleh secara acak adalah no 2,5,8 dan 10 ◀

## b. Systematic Sampling

*Systematic random sampling* adalah suatu metode statistik dimana sampel dari populasi dipilih secara acak namun dengan interval yang telah ditentukan. Interval ini disebut sebagai *sampling interval*, yang didapat dari pembagian populasi menjadi sampel yang diperlukan.

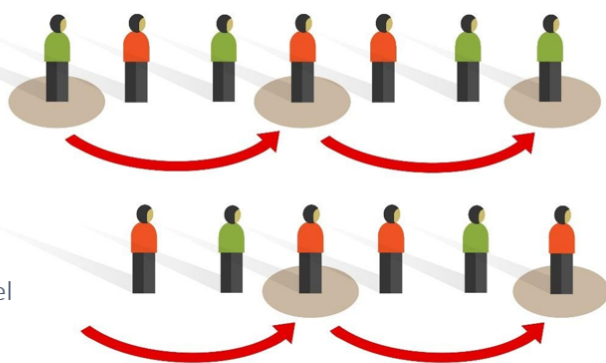
### Tahapan Systematic Random Sampling

Berikan nomor pada setiap penyusun populasi. Misalnya ada 50 orang dalam populasi. Beri label angka 1 sampai 50 pada anggota populasi;

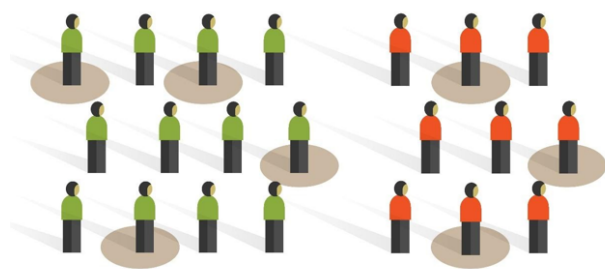
Tentukan ukuran sampel yang Anda perlukan. Misalnya Anda memerlukan 10 orang; Bagi populasi berdasarkan ukuran sampel Anda.

Pada contoh *systematic random sampling* ini berarti 50 dibagi ukuran sampel 10.  $50/10 = 5$ . Maka angka 5 ini menjadi *sampling interval* pada penelitian Anda. Di bawah ini dapat Anda lihat sampel yang dipilih mengikuti *sampling interval* 5, maka pilihlah anggota setiap kelipatan 5.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36  
37 38 39 40 41 42 43 44 45 46 47 48 49 50



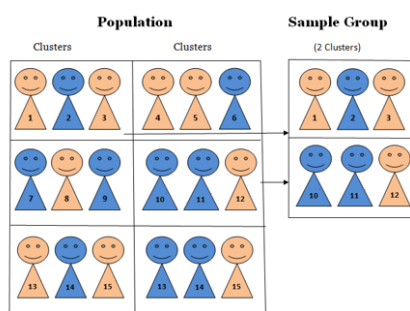
## c. Stratified Random Sampling



- Digunakan ketika ada perbedaan karakteristik yang cukup signifikan dari setiap unit observasi
- Setiap unit dikelompokkan terlebih dahulu berdasarkan karakteristiknya, kemudian random sampling dilakukan untuk setiap kelompok karakteristik tersebut.
- Satu kelompok mewakili tingkatan tertentu dalam populasi



## d. Cluster Random Sampling



- *Cluster random sampling* adalah suatu jenis teknik *sampling* dimana seorang peneliti membagi populasi menjadi beberapa kelompok yang terpisah yang disebut sebagai *cluster*.
- Dari beberapa *cluster* ini diambil beberapa sampel yang dipilih secara *random* atau acak. Analisis penelitian dari teknik *cluster random sampling* ini diambil dari data sampel *cluster-cluster* tersebut.
- *Cluster random sampling* adalah teknik *sampling* yang diterapkan ketika pada populasi didapati kelompok-kelompok yang nampak seragam namun secara internal tetap berlainan.

## **2. Nonprobability sampling**

Nonprobability sampling adalah salah satu teknik pengambilan sampel yang tidak memberi peluang/kesempatan yang sama bagi setiap unsur atau anggota populasi untuk dipilih menjadi sampel. Jenis teknik sampling ini antara lain:

- a. **Convenience sampling** adalah pengambilan sampel didasarkan pada ketersediaan elemen dan kemudahan untuk mendapatkannya. Sampel diambil/terpilih karena sampel tersebut ada pada tempat dan waktu yang tepat. Cara ini nyaris tidak dapat diandalkan, tapi paling murah dan cepat dilakukan karena peneliti memiliki kebebasan untuk memilih siapa saja yang mereka temui. Penarikan sampel ini bermanfaat penggunaannya pada tahap awal penelitian eksploratif yang ditujukan untuk mencari petunjuk awal tentang suatu kondisi yang menarik perhatian. Hasil yang diperoleh dengan cara ini seringkali dapat menyediakan bukti-bukti yang cukup melimpah sehingga terkadang pengambilan sampel yang lebih canggih tidak diperlukan lagi.

**Contoh:** Penelitian yang dilakukan oleh seorang dosen mengenai keinginan mahasiswa seluruh perguruan tinggi negeri dan swasta terhadap pemerintah RI berkaitan dengan langkah-langkah yang harus ditempuh dalam menghadapi pandemi yang telah berjalan satu tahun. Disini dapat lebih tepat (dari segi waktu dan kemudahan mendapatkan data) menggunakan teknik *convenience sampling* dengan mengambil sampel mahasiswa dari setiap kelas pada saat dosen tersebut mengajar.

- b. **Sampling Kuota atau Quota Sampling** adalah teknik untuk menentukan sampel dari populasi yang mempunyai ciri-ciri tertentu sampai jumlah (kuota) yang diinginkan. Teknik ini jumlah populasi tidak diperhitungkan akan tetapi diklasifikasikan dalam beberapa kelompok. Sampel diambil dengan memberikan jatah atau quorum tertentu terhadap kelompok. Pengumpulan data dilakukan langsung pada unit sampling. Setelah jatah terpenuhi, maka pengumpulan data dihentikan. Teknik ini biasanya digunakan dan didesain untuk penelitian yang menginginkan sedikit sampel dimana setiap kasus dipelajari secara mendalam. Dan bahayanya, jika sampel terlalu sedikit, maka tidak akan dapat mewakili populasi.

**d. Judgment Sampling Purposive** pada dasarnya merupakan suatu bentuk Convenience sampling bila ditinjau dari cara pengambilan unit-unit sampelnya. Sampel diambil berdasarkan pada kriteria-kriteria yang telah dirumuskan terlebih dahulu oleh peneliti. Dalam perumusan kriterianya, subjektifitas dan pengalaman dari peneliti sangat berperan. Pemilihan sekelompok subjek dalam purposive sampling, didasarkan atas ciri-ciri tertentu yang dipandang mempunyai sangkut paut yang erat dengan ciri-ciri populasi yang sudah diketahui sebelumnya. Maka dengan kata lain, unit sampel yang dihubungi disesuaikan dengan kriteria-kriteria tertentu yang diterapkan berdasarkan tujuan penelitian atau permasalahan penelitian.

**e. Snowball Sampling** adalah teknik pengambilan sampel yang awal mula jumlahnya kecil, kemudian sampel ini disuruh memilih teman-temannya untuk dijadikan sampel. Dan begitu seterusnya, sehingga jumlah sampel makin lama makin banyak. Ibaratkan sebuah bola salju yang menggelinding, makin lama semakin besar. Pada penelitian kualitatif banyak menggunakan sampel purposive dan snowball.

**Contoh:** Akan diteliti pendapat para ahli penyakit dalam senior Indonesia terhadap pengobatan penyakit dalam menggunakan tenaga dalam, maka pengambilan sampel dapat dilakukan dengan Snowball sampling. Pertimbangan tersebut dikaitkan dengan kenyataan bahwa populasi yang berupa ahli penyakit dalam senior di Indonesia sangat spesifik, jumlahnya sedikit dengan lokasi yang tersebar dan karena profesi yang sama maka kemungkinan besar mereka saling mengenal satu dengan yang lain. Pertama, dicari seorang ahli penyakit dalam senior. Selanjutnya dari satu orang ini diminta menunjukkan beberapa ahli lain yang dapat diwawancarai, demikian seterusnya sehingga diperoleh sejumlah responden yang diperlukan.

### 5.1.2 Eksperimen

Sejalan dengan metode penelitian survey, metode penelitian eksperimen merupakan salah satu bentuk metode penelitian dengan jenis kuantitatif pula. Namun, jika metode survey dapat dilakukan tanpa perlakuan atau percobaan, maka metode eksperimen ini merupakan metode yang wajib menggunakan perlakuan atau percobaan. Metode penelitian eksperimen digunakan untuk mengetahui pengaruh variabel bebas (perlakuan) terhadap variabel terikat dalam kondisi yang dapat dikendalikan. Pengendalian kondisi dimaksudkan agar tidak ada variabel lain selain variabel bebas yang dapat mempengaruhi variabel terikat. Untuk dapat mengendalikan kondisi pada penelitian eksperimen digunakan kelompok kontrol.

Penelitian eksperimen biasa digunakan di laboratorium. Dengan demikian, untuk bidang farmasi teknologi dan bidang kandungan bahan alam, pelaksanaan penelitian yang menggunakan metode ini dilaksanakan di laboratorium, kecuali untuk bidang farmasi rumah sakit dan bidang farmasi komunitas.

#### Jenis Eksperimen :

- Eksperimen Alami (Quasy-Eksperiment) : Eksperimen alami melibatkan prediksi atau membentuk hipotesis dan kemudian mengumpulkan data dengan mengamati suatu sistem. Variabelnya tidak dikontrol dalam percobaan alami.
- Eksperimen Terkendali : Eksperimen laboratorium adalah eksperimen terkontrol, meskipun dapat juga dilakukan di luar pengaturan laboratorium . Dalam eksperimen terkontrol, kita membandingkan grup eksperimen dengan sebuah kontrol kelompok. Idealnya, kedua kelompok ini identik kecuali satu variabel yaitu variabel bebas.
- Eksperimen Lapangan: Eksperimen lapangan dapat berupa eksperimen alami atau eksperimen terkontrol. Itu terjadi di pengaturan dunia nyata, bukan dalam kondisi lab. Misalnya, percobaan yang melibatkan hewan di habitat aslinya akan menjadi percobaan lapangan.

### **5.1.3 Observasi**

*Observation is the process of watching someone or something. Often, observation is an informal action, but it can also be formal and involve data collection. An observation can also be the collected information itself. Explore different examples of observations from scientific to social.*

Data observasi adalah data yang dikumpulkan secara berkala untuk memantau perilaku alami dari objek observasi. Tidak ada campur tangan dari peneliti dalam proses observasi.

Observasi dapat dilakukan tanpa perlu dilakukan perancangan survey atau pun pelaksanaan eksperimen.

Contoh Observasi :

- Seorang dokter mengawasi pasien untuk mengetahui reaksinya setelah diberi suntikan obat tertentu
- Seorang tukang roti mengawasi kuenya di oven untuk memastikan kuenya mengembang dan matang dengan benar
- Seorang astronom melihat langit malam dan merekam data mengenai pergerakan benda yang dilihatnya
- Seorang ahli zoologi mengamati singa di sarang setelah diperlihatkan mangsa untuk menentukan kecepatan respons hewan
- Pengendali penerbangan udara mengawasi monitor untuk mengawasi pergerakan dan lokasi pesawat

## 5.2 Klasifikasi Data

### 5.2.1 Berdasarkan Sumbernya

#### Data primer

Data primer, adalah data yang dikoleksi langsung oleh peneliti dan berada pada karakter originalnya. Data primer dapat dikumpulkan melalui eksperimen (penelitian eksperimental) dan survey .

#### Data Sekunder

Data sekunder, adalah data yang dikoleksi oleh orang lain dan telah melalui proses statistik. Ketika peneliti menggunakan data sekunder, maka dia harus melihat ke dalam berbagai sumber dari mana ia bisa mendapatkan mereka. Dalam hal ini ia tentu tidak dihadapkan dengan masalah yang biasanya berhubungan dengan pengumpulan data asli.

Dengan hati-hati, sebelum menggunakan data sekunder, data sekunder haruslah memiliki karakteristik sebagai berikut:

- Keandalan data
- Kesesuaian data
- Kecukupan data



### 5.2.2 Berdasarkan Implisit/eksplisit Data

Data diklasifikasikan dalam 2 jenis, yaitu terstruktur dan tidak terstruktur. Data terstruktur direpresentasikan dalam skema yang jelas sehingga mudah untuk dianalisa maupun diintegrasikan dengan data terstruktur lainnya. Sedangkan data tidak terstruktur direpresentasikan dalam berbagai bentuk sehingga sangat sulit untuk dianalisa maupun diintegrasikan dengan sumber data lain

#### 1. Data Berstruktur

- ✓ Pada data terstruktur, nilai yang ingin diketahui dari suatu unit observasi direpresentasikan dalam suatu variable tunggal yang jelas batasan nilainya.
- ✓ Data terstruktur biasanya berbentuk tabel relasi atau data statistik

Jenis data structured (data tradisional) dapat diproses, disimpan, dan diambil dalam format tetap. Jenis data ini disimpan dalam bentuk tabel, baris dan kolom yang normalnya disimpan dalam excel atau spreadsheet, dimana informasi pada data sangat terorganisir dan dapat dengan mudah diakses dari database dengan algoritma mesin pencari sederhana.

Contoh data terstruktur adalah, data sensor, data penjualan pada suatu perusahaan, data karyawan dalam database perusahaan dengan detail yang terstruktur seperti detail data diri karyawan, posisi pekerjaan, gaji, dan lainnya ditampilkan secara terorganisir.

Budi	Budiman	13011910002	Surakarta	Jawa Tengah
------	---------	-------------	-----------	-------------

#### 2. Data Semi-Terstruktur

- ✓ Data semi terstruktur memiliki struktur yang relatif fleksibel.
- ✓ Data semi terstruktur sering juga disebut “*schemaless*” atau “*self-describing*”
- ✓ Ciri:
  1. Data dapat berisi variabel yang belum diketahui pada saat perancangan
  2. Untuk satu jenis data, bisa jadi direpresentasikan dalam lebih dari satu cara, misalnya tanggal bisa direpresentasikan dalam satu kesatuan (1 Desember 2020) atau dipisahkan (Tanggal 1 Bulan Desember Tahun 2020)
  3. Bisa jadi ada banyak variabel yang didefinisikan di awal tetapi tidak ada nilainya.

Jenis data semi-structured merupakan jenis data yang dimasukkan ke dalam sebuah tabel, tetapi skemanya tidak sama dengan tabel biasa yang hanya terdiri dari baris dan kolom. Data semi-terstruktur mengandung format data terstruktur dan tidak terstruktur. Walaupun belum diklasifikasi oleh repository tertentu (database), namun

```
<Data Mahasiswa>
<Nama> Budi Budiman </Nama>
<NIM> 13011910002 </NIM>
<Kota> Surakarta </Kota>
<Provinsi> Jawa Tengah </Provinsi>
</Data Mahasiswa>
```

mengandung informasi yang penting. Contohnya adalah data dalam bentuk file csv, file xml, dan file json.

### 3. Data Tidak Berstruktur

- ✓ Data tidak terstruktur merupakan data yang berbentuk rangkaian symbol, umumnya dituliskan dalam bahasa alami
- ✓ Contoh data tidak terstruktur biasanya didapatkan dari isian kuesioner yang berbentuk teks bebas, atau isi dari suatu email.

Jenis data unstructured adalah data dengan bentuk yang tidak dikenal, harus disimpan dengan format khusus karena tidak memiliki struktur yang spesifik seperti jenis data structured. Raw data dari jenis data ini hanya dapat menghasilkan nilai setelah diproses dan dianalisa. Menyimpan data jenis ini pun memiliki kerumitan seperti memerlukan penggunaan sistem penyimpanan yang memadai, seperti database NoSQL (MongoDB dan CouchDB).

Contoh jenis data tidak terstruktur seperti data teks, berformat foto atau gambar, video, atau suara. Selain itu, bisa juga dalam bentuk keluhan pelanggan, kontrak, ataupun email internal. Contoh dari data jenis ini dapat ditemukan dalam social

Budi Budiman adalah mahasiswa dengan NIM 13011910002. Ia berasal dari Kota Surakarta atau biasa dikenal dengan Kota Solo, Provinsi Jawa Tengah.

media, seperti komentar, likes, followers, dan data click pada setiap aktivitas di akun media sosialmu.