



M6 – Teknik Pengambilan Sampel ***Probability Sampling dan Non- Probability Sampling***

-Tim Dosen Metode Statistika-





Outline

- Teknik Pengambilan Sampel
- Probability Sampling
- Non-Probability Sampling



Pendahuluan Teknik Pengambilan Sampel / Teknik Sampling



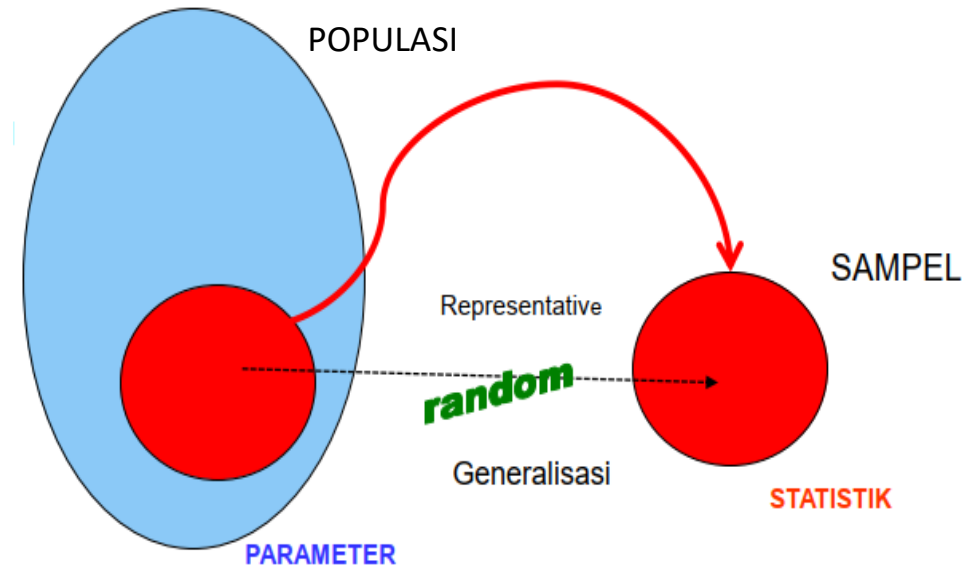
Teknik Pengambilan Sampel / Teknik Sampling



Sampling merupakan proses dalam menyeleksi observasi (sampel) untuk menyediakan deskripsi yang cukup dan menguatkan inferensi dari suatu populasi.

Unsur Sampling:

- Populasi
- Sampel
- Random
- Representatif
- Generalisasi
- Parameter
- Statistik





Konsep Dasar Sampling

Populasi

Kumpulan secara keseluruhan dari observasi yang akan diteliti.

Sampel

Suatu subgroup/bagian dari elemen-elemen populasi yang dipilih dan akan diteliti.

Populasi Target (*Target population*) dapat didefinisikan dalam bentuk:

- ***Elemen*** : responden/sumber informasi
- **Unit Sampling (*Sampled population*)** : unit dasar berisi elemen-elemen populasi yang akan diambil sampelnya
- **Kerangka *Sampling* (*Sampling frame*)** : sebuah daftar atau himpunan dari unit sampel dalam populasi yang digunakan untuk menyeleksi sampel tersebut.



Proses Sampling

1. Definisikan populasi, termasuk didalamnya menentukan populasi target
2. Tentukan Sampling Frame
3. Pilih Teknik Sampling yang akan digunakan (probability/non-probability sampling)
4. Tentukan ukuran sampel
5. Tentukan proses samplingnya

Inti dari Teknik Sampling

POPULASI (=N)



Akan dipelajari berbagai teknik sampling yang sesuai dengan kondisi populasi

Apa saja?

Dan

Akan didiskusikan : faktor apa saja yang menentukan besarnya "n"

Inti dari teknik sampling adalah **menentukan ukuran sampel (n).**

Faktor-faktor yang mempengaruhi n diantaranya:

- Besarnya populasi
- Keragaman populasi → heterogenitas populasi
- Tingkat ketelitian / tingkat kepercayaan
- Rencana analisis

Diskusi



Contoh 1

Suatu survey dilakukan dengan tujuan mengetahui rata-rata pengeluaran mahasiswa ITS .

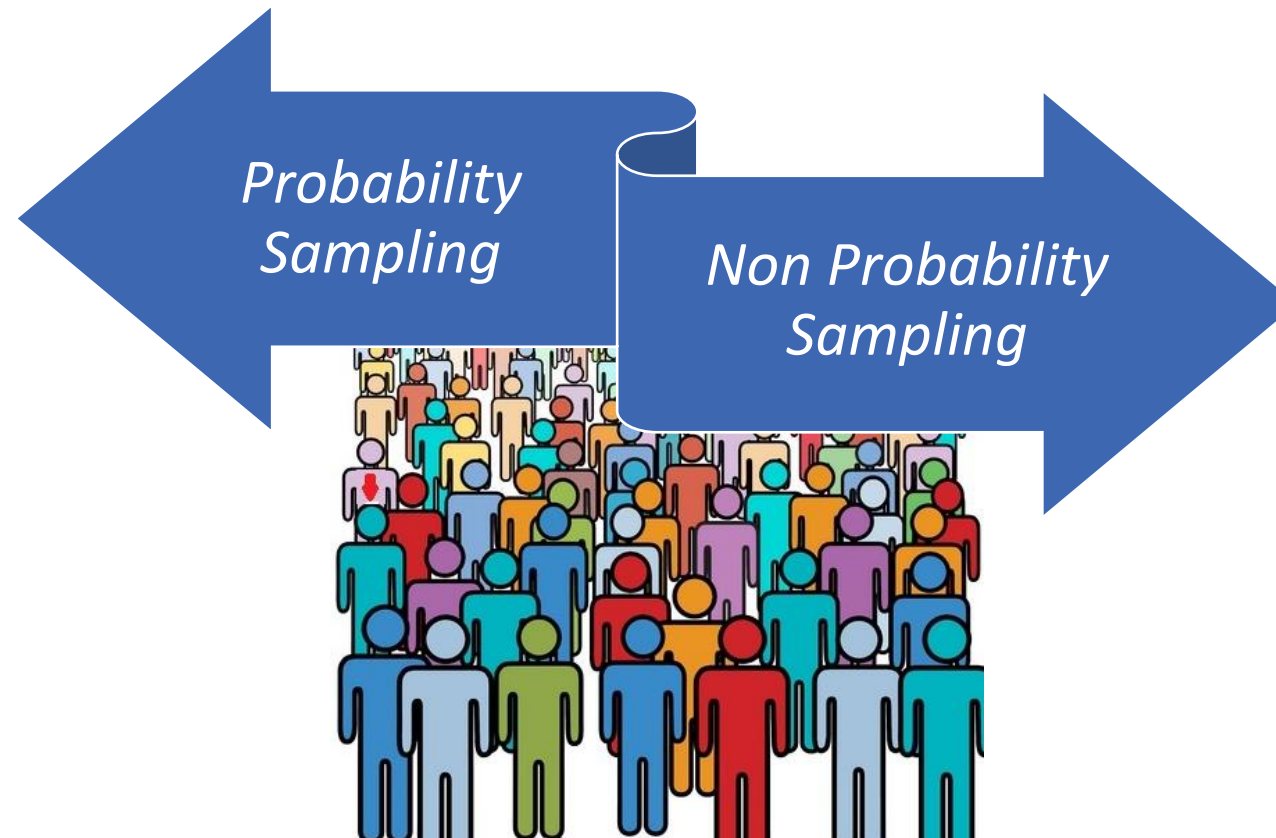
Contoh 2

Suatu survei bertujuan untuk memperkirakan pengeluaran rumah tangga nelayan di Desa “X”

Definisikan: populasi dan populasi target, kerangka sampel, unit sampling, unit observasi, karakteristik yang akan diukur, parameter yang diestimasi

Jenis – Jenis Teknik Sampling

- Simple Random Sampling (SRS)
- Systematic Sampling (SS)
- Stratified Sampling
- Cluster Sampling



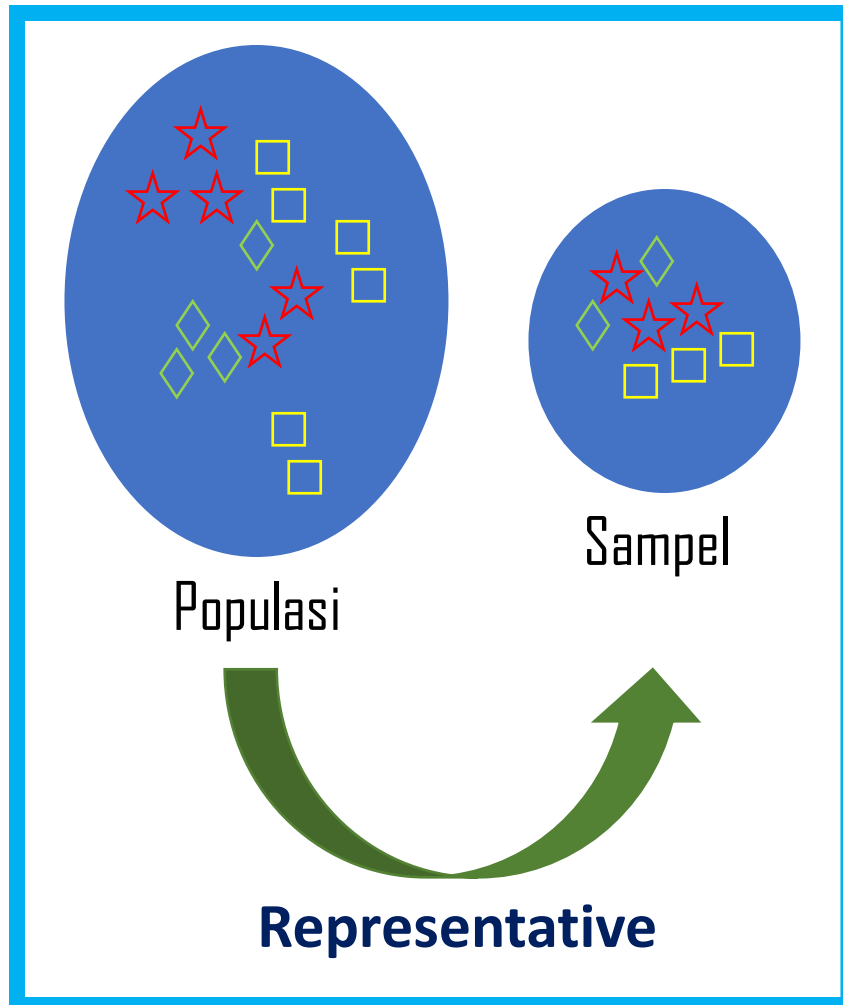
- Convenience / accidental sampling
- Purposive / judgmental sampling
- Quota sampling
- Snowball sampling
- Wildlife sampling



Probability Sampling



Probability Sampling



Sampel yang representatif didapatkan dengan pengambilan sampel yang dilakukan secara **random** atau setiap anggota populasi memiliki kesempatan atau peluang terpilih yang sama

Probability sample lebih akurat daripada non-probability sample

Simple Random Sampling (SRS)



Pengambilan sampel berukuran n dari populasi N dan setiap anggota populasi memiliki kesempatan yang sama untuk terpilih sebagai sampel (*sampling tanpa pengembalian*)

Metode SRS baik digunakan, jika :

- ☐ Karakteristik anggota populasi identik (populasi homogen)
- ☐ *Sampling frame* diketahui
- ☐ Mudah dikerjakan
- ☐ Sampel tersebar, biaya besar

Cara menentukan sampel yang dipilih dapat menggunakan tabel acak, melempar dadu/ undian, paket program statistik, dll.

Simple Random Sampling (SRS) atau Sampling Acak Sederhana (SAS) merupakan dasar dari metode sampling yang lain.

Contoh SRS (1)



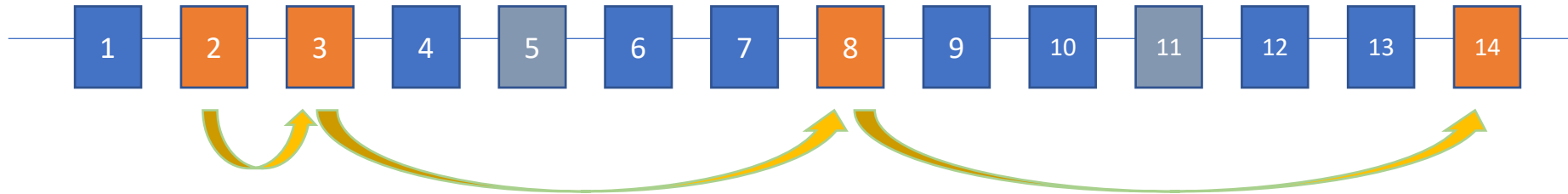
Penelitian dilakukan pada mahasiswa S1 Teknologi Sains Data angkatan 2021, tujuan penelitian untuk mengetahui proporsi mahasiswa yang membawa sepeda motor saat kuliah. Peneliti memiliki *sampling frame* dari mahasiswa S1 Teknologi Sains Data angkatan 2021, dimana setiap mahasiswa memiliki kesempatan yang sama untuk terpilih menjadi sampel. Kemudian, responden terpilih akan disurvei terkait apakah mereka menggunakan sepeda motor saat ke kampus? Dengan demikian, dapat diperkirakan kapasitas tempat parkir prodi Teknologi Sains Data, sehingga dapat menampung semua sepeda motor yang digunakan mahasiswa prodi Teknologi Sains Data.

Contoh SRS (2)



Perumahan mewah, umumnya dihuni penduduk berpenghasilan menengah ke atas. Suatu penelitian ingin mengetahui banyaknya rumah tangga yang menggunakan layanan televisi satelit berlangganan. Dari semua rumah tangga dalam perumahan tersebut, diambil sampel secara SRS, sehingga mengetahui banyaknya penduduk yang menggunakan televisi satelit berlangganan.

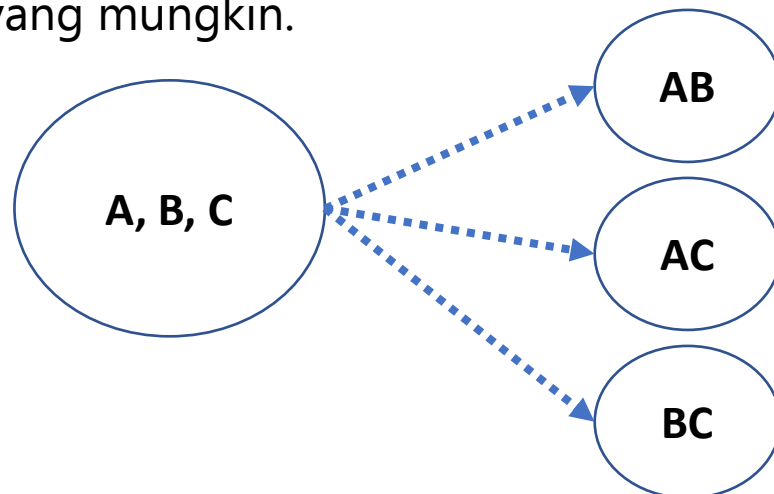
Skema SRS



Periode antara anggota populasi yang terpilih menjadi sampel, bersifat **random** atau **tidak memiliki pola tertentu**.

Contoh 1

Populasi berukuran $N = 14$, akan diambil sampel berukuran $n = 4$. Untuk menghitung kemungkinan sampel terambil yaitu $C(N, n)$. Sehingga untuk contoh tersebut, akan ada $C(14, 4) = 1001$ kombinasi sampel yang mungkin.



Contoh 2

Dari contoh di samping, terdapat $N = 3$ dan akan diambil $n = 2$. Kombinasi sampel yang akan terpilih sebanyak $C(3, 2) = 3$, yaitu AB, AC, BC masing-masing memiliki peluang yang sama untuk terpilih sebagai sampel dengan peluang sebesar $1/3$.

SRS



Pada intinya, sampling digunakan untuk mengestimasi populasi.

Dalam SRS, nilai-nilai yang perlu diestimasi:

- Rata-rata (μ)
- Total (τ)
- Proporsi (p)

Rata-Rata (μ)

Estimator of the population mean μ :

$$\hat{\mu} = \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

Estimated variance of \bar{y} :

$$\hat{V}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$$

Bound on the error of estimation:

$$2\sqrt{\hat{V}(\bar{y})} = 2\sqrt{\left(1 - \frac{n}{N}\right) \frac{s^2}{n}}$$

Sample size required to estimate μ with a bound on the error of estimation B :

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}$$

where

$$D = \frac{B^2}{4}$$

The average amount of money μ for a hospital's accounts receivable must be estimated. Although no prior data are available to estimate the population variance, it is known that most accounts lie within a \$100 range. There are $N = 1000$ open accounts. Find the sample size needed to estimate μ with a bound on the error of estimation $B = \$3$.

We need an estimate of σ^2 , the population variance. Because the range is often approximately equal to $4SD$ (4σ), one fourth of the range will provide an approximate value of σ . Hence,

$$\sigma \approx \frac{\text{Range}}{4} = \frac{100}{4} = 25$$

and

$$\sigma^2 \approx (25)^2 = 625$$

Using Eq. (4.11), we obtain

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}$$

where

$$D = \frac{B^2}{4} = \frac{(3)^2}{4} = 2.25$$

So

$$n = \frac{1000(625)}{999(2.25) + 625} = 217.56$$

That is, we need approximately 218 observations to estimate μ , the mean accounts receivable, with a bound on the error of estimation of \$3.00. Note that the tool will show a sample size of 213 for a 95% confidence interval with margin of error of \$3. ■

Total (τ)



Estimator of the population total τ :

$$\hat{\tau} = N\bar{y} = \frac{N \sum_{i=1}^n y_i}{n}$$

Estimated variance of τ :

$$\hat{V}(\hat{\tau}) = \hat{V}(N\bar{y}) = N^2 \left(1 - \frac{n}{N}\right) \left(\frac{s^2}{n}\right)$$

Bound on the error of estimation:

$$2\sqrt{\hat{V}(N\bar{y})} = 2\sqrt{N^2 \left(1 - \frac{n}{N}\right) \left(\frac{s^2}{n}\right)}$$

Sample size required to estimate τ with a bound on error B :

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}$$

where

$$D = \frac{B^2}{4N^2}$$

An investigator is interested in estimating the total weight gain in 4 weeks for $N = 1000$ chicks fed on a new ration. Obviously, to weigh each bird would be time-consuming and tedious. Therefore, determine the number of chicks to be sampled in this study in order to estimate τ with a bound on the error of estimation equal to 1000 grams. Many similar studies on chick nutrition have been run in the past. Using data from these studies, the investigator found that σ^2 , the population variance, was approximately equal to 36.00 (grams)². Determine the required sample size.

We can obtain an approximate sample size using Eq. (4.13) with σ^2 equal to 36.00 and

$$D = \frac{B^2}{4N^2} = \frac{(1000)^2}{4(1000)^2} = 0.25$$

That is,

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2} = \frac{1000(36.00)}{999(0.25) + 36.00} = 125.98$$

The investigator therefore needs to weigh $n = 126$ chicks to estimate τ , the total weight gain for $N = 1000$ chickens in 0 to 4 weeks, with a bound on the error of estimation equal to 1000 grams. Note that the tool will show a sample size of 123 for a 95% confidence interval with margin of error of 1000 grams. ■

Proporsi (p)



Estimator of the population proportion p :

$$\hat{p} = \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

Estimated variance of \hat{p} :

$$\hat{V}(\hat{p}) = \left(1 - \frac{n}{N}\right) \frac{\hat{p}\hat{q}}{n-1}$$

where

$$\hat{q} = 1 - \hat{p}$$

Bound on the error of estimation:

$$2\sqrt{\hat{V}(\hat{p})} = 2\sqrt{\left(1 - \frac{n}{N}\right) \frac{\hat{p}\hat{q}}{n-1}}$$

Sample size required to estimate p with a bound on the error of estimation B :

$$n = \frac{Npq}{(N-1)D + pq}$$

where

$$q = 1 - p \quad \text{and} \quad D = \frac{B^2}{4}$$

Student government leaders at a college want to conduct a survey to determine the proportion of students who favor a proposed honor code. Because interviewing $N = 2000$ students in a reasonable length of time is almost impossible, determine the sample size (number of students to be interviewed) needed to estimate p with a bound on the error of estimation of magnitude $B = 0.05$. Assume that no prior information is available to estimate p .

Answer:

We can approximate the required sample sizes when no prior information is available by setting $p = 0.5$. We have,

$$D = \frac{B^2}{4} = \frac{0.05^2}{4} = 0.000625$$

Hence,

$$n = \frac{Npq}{(N-1)D + pq} = \frac{(2000)(0.5)(0.5)}{(1999)(0.000625) + (0.5)(0.5)} = 333.56$$

That is, 334 students must be interviewed to estimate the proportion of students who favor the proposed honor code with a bound on the error of estimation of $B = 0.05$

Kelebihan dan Kekurangan SRS



Kelebihan:

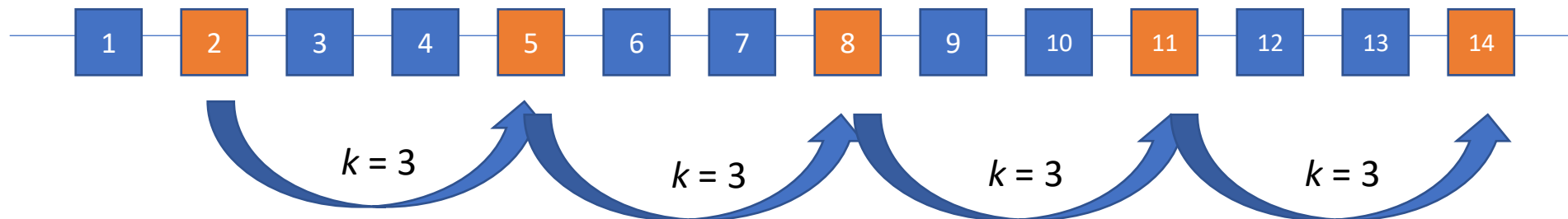
- Sederhana
- Kesalahan sampling mudah dihitung

Kekurangan:

- Perlu sampling frame / daftar unit yang lengkap
- Sample yang diambil tidak selalu memberikan representasi populasi yang baik
- Penggunaan kurang tepat apabila saat populasi sampel berjumlah banyak.

Systematic Sampling (SS)

Pengambilan sampel yang dilakukan **secara sistematis**.
Memulai dari **titik yang dipilih secara random** pada *sampling frame* sebagai sampel pertama, kemudian sampel berikutnya terpilih pada interval yang sama dengan **kelipatan k** .



Lebih efisien daripada SRS
Populasi homogen dan membutuhkan *sampling frame*

Contoh SS (1)



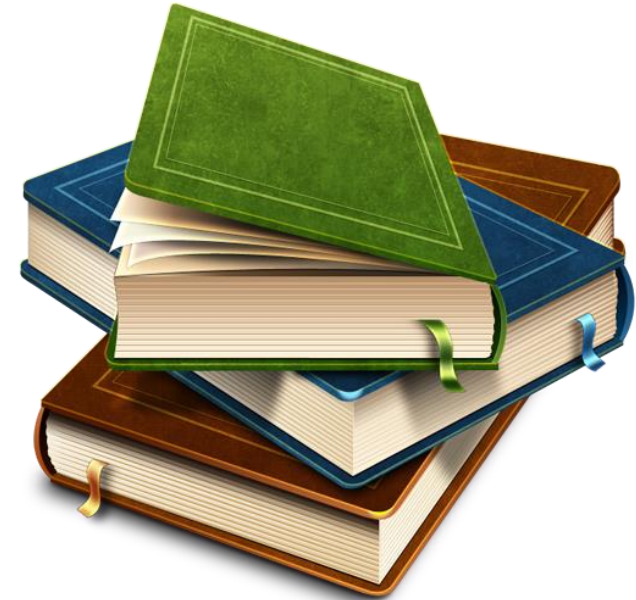
Perusahaan Roti ingin mengurangi kerusakan kemasan roti sehingga dapat mengurangi produk cacat. Bagian pengendalian kualitas (*Quality Control*) melakukan inspeksi kemasan roti. Dalam mendapatkan sampel kemasan roti, karyawan menginspeksi kemasan roti secara sistematis dalam sehari untuk setiap adonan yang diproduksi. Misalkan: petugas inspeksi mengamati kemasan roti tiap adonan dengan kelipatan kemasan roti yang diamati sebesar 5. Selanjutnya, bagian QC dapat memperoleh ukuran proporsi kemasan roti yang cacat.



Contoh SS (2)



Survei penduduk di Perumahan Mulyorejo, dilakukan untuk mengetahui rata-rata pengeluaran rumah tangga pada pendidikan formal anak. Sampling dilakukan dengan memberi nomer urut setiap rumah sebagai *sampling frame*, kemudian diambil sampel secara sistematis.





Rata-rata (μ) & Total (τ)

Estimator of the population mean μ :

$$\hat{\mu} = \bar{y}_{\text{sy}} = \frac{\sum_{i=1}^n y_i}{n}$$

where the subscript sy signifies that systematic sampling was used.

Estimated variance of \bar{y}_{sy} :

$$\hat{V}(\bar{y}_{\text{sy}}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$$

assuming a randomly ordered population.

Estimator of the population total τ :

$$\hat{\tau} = N\bar{y}_{\text{sy}}$$

Estimated variance of τ :

$$\hat{V}(N\bar{y}_{\text{sy}}) = N^2\hat{V}(\bar{y}_{\text{sy}}) = N^2\left(1 - \frac{n}{N}\right)\left(\frac{s^2}{n}\right)$$

assuming a randomly ordered population.

Sample size required to estimate μ with a bound B on the error of estimation:

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}$$

where

$$D = \frac{B^2}{4}$$



Proporsi (p)

Estimator of the population proportion p :

$$\hat{p}_{sy} = \bar{y}_{sy} = \frac{\sum_{i=1}^n y_i}{n}$$

Estimated variance of \hat{p}_{sy} :

$$\hat{V}(\hat{p}_{sy}) = \left(1 - \frac{n}{N}\right) \frac{\hat{p}_{sy} \hat{q}_{sy}}{n - 1}$$

where $\hat{q}_{sy} = 1 - \hat{p}_{sy}$, assuming a randomly ordered population.

Sample size required to estimate p with a bound B on the error of estimation:

$$n = \frac{Npq}{(N - 1)D + pq}$$

where

$$q = 1 - p \quad \text{and} \quad D = \frac{B^2}{4}$$

Kelebihan dan Kekurangan SS



Kelebihan:

- Sederhana
- Kesalahan sampling mudah dihitung

Kekurangan:


- Perlu daftar unit secara lengkap
- Hanya dapat digunakan jika populasi merupakan populasi periodik (artinya kejadian yang sama terjadi secara periodik)



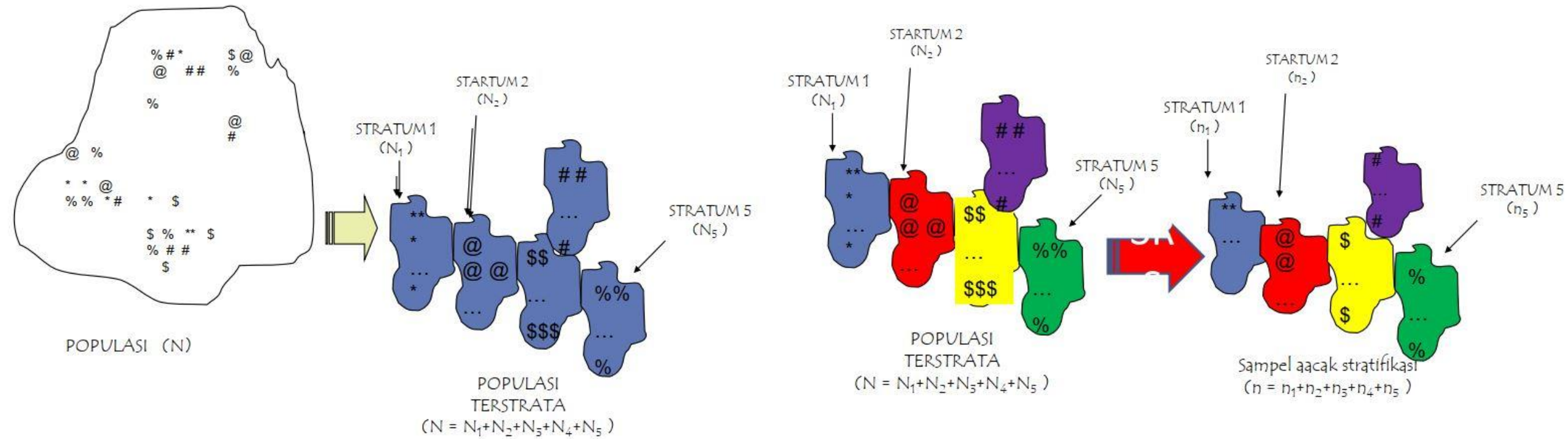
Stratified Sampling (StS)

 **Membagi populasi** berdasarkan karakteristik tertentu menjadi beberapa subkelompok yang homogen (**strata**).

 Elemen **dalam masing-masing strata** bersifat **homogen** tetapi **antar strata** bersifat **heterogen**.

 **Pengambilan sampel** dilakukan dengan cara mengambil sampel dari masing-masing strata sesuai dengan **proporsi pada masing-masing strata**.

Ilustrasi Stratified Sampling



Dasar Pembentukan Stratified Sampling



- Unit sampling itu sendiri
- Variabel wilayah administrasi: kabupaten, kecamatan, desa, wilayah
- Variabel letak geografis: pantai, bukan pantai
- Variabel lain: kepadatan penduduk, lapangan usaha
- Untuk perusahaan: besar, sedang, kecil
- Untuk sekolah: negeri, swasta
- Dll.

Alasan Memilih Stratified Sampling

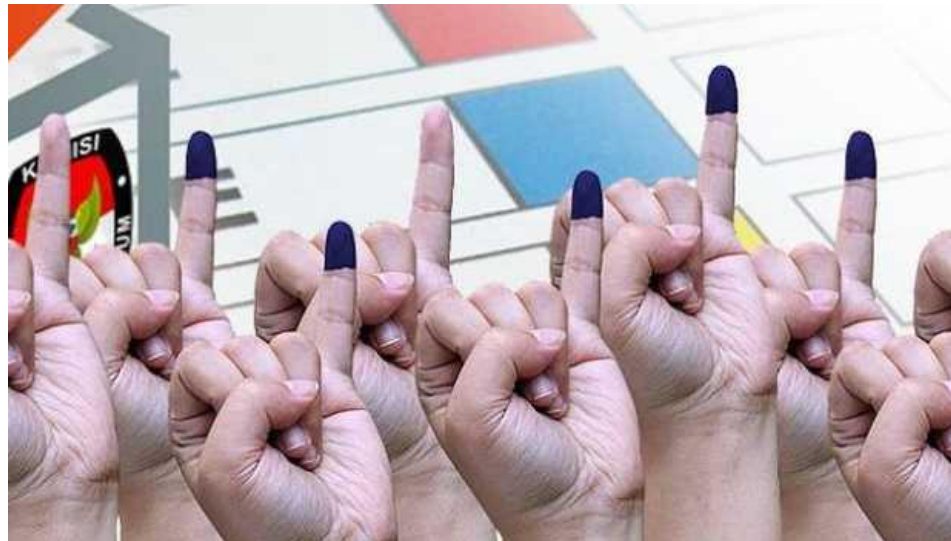


- Stratifikasi DAPAT menghasilkan suatu batas kekeliruan taksiran yang kecil dibandingkan dengan yang dihasilkan oleh SRS. Bila menginginkan precision/ketelitian untuk sub populasi tertentu
- Biaya per pengamatan dalam survei dapat ditekan dengan mengelompokkan elemen-elemen populasi kedalam strata
- Taksiran dari parameter dari populasi dimungkinkan untuk tiap kelompok yang dinyatakan dalam strata

Contoh Stratified Sampling



Misal peneliti ingin mengadakan penelitian rata-rata pemilih tetap yang golput pada suatu pemilu. Pemilih tersebut dibagi menjadi beberapa strata atau tingkatan berdasarkan usia pemilih, yaitu 17-26 tahun, 27-36 tahun, 37-46 tahun, dan >46 tahun. Kemudian, dari masing-masing strata tersebut ditentukan jumlah sampel yang digunakan untuk penelitian.





Rata-rata (μ) & Total (τ)

Estimator of the population mean μ :

$$\bar{y}_{st} = \frac{1}{N} [N_1 \bar{y}_1 + N_2 \bar{y}_2 + \cdots + N_L \bar{y}_L] = \frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i$$

Estimated variance of \bar{y}_{st} :

$$\begin{aligned} \hat{V}(\bar{y}_{st}) &= \frac{1}{N^2} [N_1^2 \hat{V}(\bar{y}_1) + N_2^2 \hat{V}(\bar{y}_2) + \cdots + N_L^2 \hat{V}(\bar{y}_L)] \\ &= \frac{1}{N^2} \left[N_1^2 \left(1 - \frac{n_1}{N_1} \right) \left(\frac{s_1^2}{n_1} \right) + \cdots + N_L^2 \left(1 - \frac{n_L}{N_L} \right) \left(\frac{s_L^2}{n_L} \right) \right] \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i} \right) \left(\frac{s_i^2}{n_i} \right) \end{aligned}$$

Estimator of the population total τ :

$$N \bar{y}_{st} = N_1 \bar{y}_1 + N_2 \bar{y}_2 + \cdots + N_L \bar{y}_L = \sum_{i=1}^L N_i \bar{y}_i$$

Estimated variance of $N \bar{y}_{st}$:

$$\hat{V}(N \bar{y}_{st}) = N^2 \hat{V}(\bar{y}_{st}) = \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i} \right) \left(\frac{s_i^2}{n_i} \right)$$



Approximate sample size required to estimate μ or τ with a bound B on the error of estimation:

$$n = \frac{\sum_{i=1}^L N_i^2 \sigma_i^2 / a_i}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2}$$

where a_i is the fraction of observations allocated to stratum i , σ_i^2 is the population variance for stratum i , and

$$D = \frac{B^2}{4} \quad \text{when estimating } \mu$$

$$D = \frac{B^2}{4N^2} \quad \text{when estimating } \tau$$

Alokasi sampel untuk
estimasi rata-rata



Approximate allocation that minimizes cost for a fixed value of $V(\bar{y}_{st})$ or minimizes $V(\bar{y}_{st})$ for a fixed cost:

$$\begin{aligned} n_i &= n \left(\frac{N_i \sigma_i / \sqrt{c_i}}{N_1 \sigma_1 / \sqrt{c_1} + N_2 \sigma_2 / \sqrt{c_2} + \cdots + N_L \sigma_L / \sqrt{c_L}} \right) \\ &= n \left(\frac{N_i \sigma_i / \sqrt{c_i}}{\sum_{k=1}^L N_k \sigma_k / \sqrt{c_k}} \right) \end{aligned}$$

where N_i denotes the size of the i th stratum, σ_i^2 denotes the population variance for the i th stratum, and c_i denotes the cost of obtaining a single observation from the i th stratum. Note that n_i is directly proportional to N_i and σ_i and inversely proportional to $\sqrt{c_i}$.

Proporsi (p)



Estimator of the population proportion p :

$$\hat{p}_{st} = \frac{1}{N^2} (N_1 \hat{p}_1 + N_2 \hat{p}_2 + \cdots + N_L \hat{p}_L) = \frac{1}{N} \sum_{i=1}^L N_i \hat{p}_i$$

Estimated variance of \hat{p}_{st} :

$$\begin{aligned} \hat{V}(\hat{p}_{st}) &= \frac{1}{N^2} [N_1^2 \hat{V}(\hat{p}_1) + N_2^2 \hat{V}(\hat{p}_2) + \cdots + N_L^2 \hat{V}(\hat{p}_L)] \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \hat{V}(\hat{p}_i) \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i}\right) \left(\frac{\hat{p}_i \hat{q}_i}{n_i - 1}\right) \end{aligned}$$

Approximate sample size required to estimate p with a bound B on the error of estimation:

$$n = \frac{\sum_{i=1}^L N_i^2 p_i q_i / a_i}{N^2 D + \sum_{i=1}^L N_i p_i q_i}$$

where a_i is the fraction of observations allocated to stratum i , p_i is the population proportion for stratum i , and $D = B^2/4$.

Approximate allocation that minimizes cost for a fixed value of $V(\hat{p}_{st})$ or minimizes $V(\hat{p}_{st})$ for a fixed cost:

$$n_1 = n \left(\frac{N_i \sqrt{p_i q_i / c_i}}{N_1 \sqrt{p_1 q_1 / c_1} + N_2 \sqrt{p_2 q_2 / c_2} + \cdots + N_L \sqrt{p_L q_L / c_L}} \right)$$
$$= n \left(\frac{N_i \sqrt{p_i q_i / c_i}}{\sum_{k=1}^L N_k \sqrt{p_k q_k / c_k}} \right)$$

where N_i denotes the size of the i th stratum, p_i denotes the population proportion for the i th stratum, and c_i denotes the cost of obtaining a single observation from the i th stratum.

Alokasi sampel untuk
estimasi proporsi





Cluster Sampling

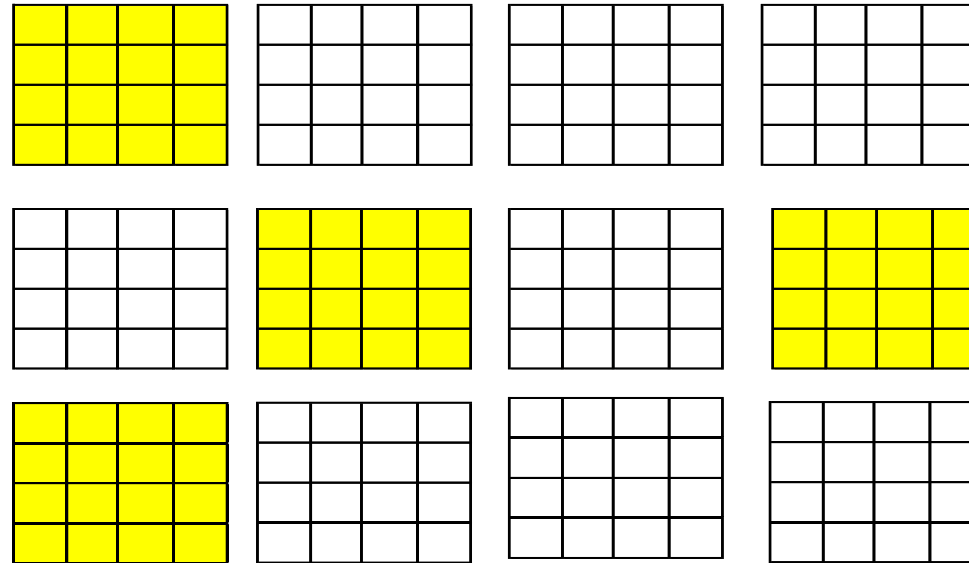
- **Unit observasi** dikumpulkan dalam unit sampel yang lebih besar, disebut **klaster**.
- Digunakan bila populasi dapat dibagi dalam kelompok-kelompok dan karakteristik yang dipelajari ada dalam setiap kelompok.
- Tiap unit (individu) yang ada dalam kelompok yang terpilih akan diambil sebagai sampel.

Peneliti kekurangan *sampling frame* yang baik, sehingga tidak memerlukan *sampling frame* untuk metode *cluster sampling*.

Dapat dilakukan
ketika

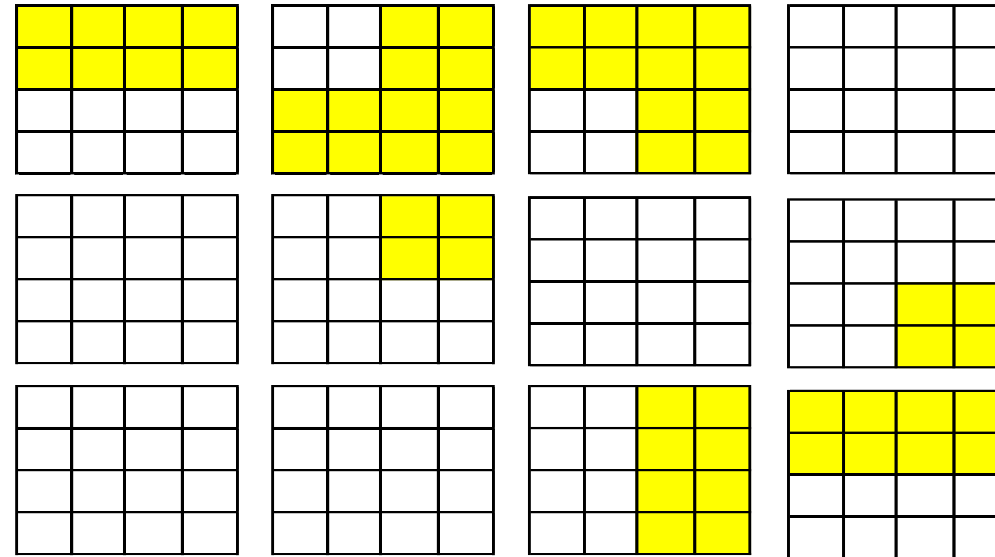
Biaya yang digunakan untuk menjangkau elemen sangat tinggi.

Contoh One Stage Cluster Sampling



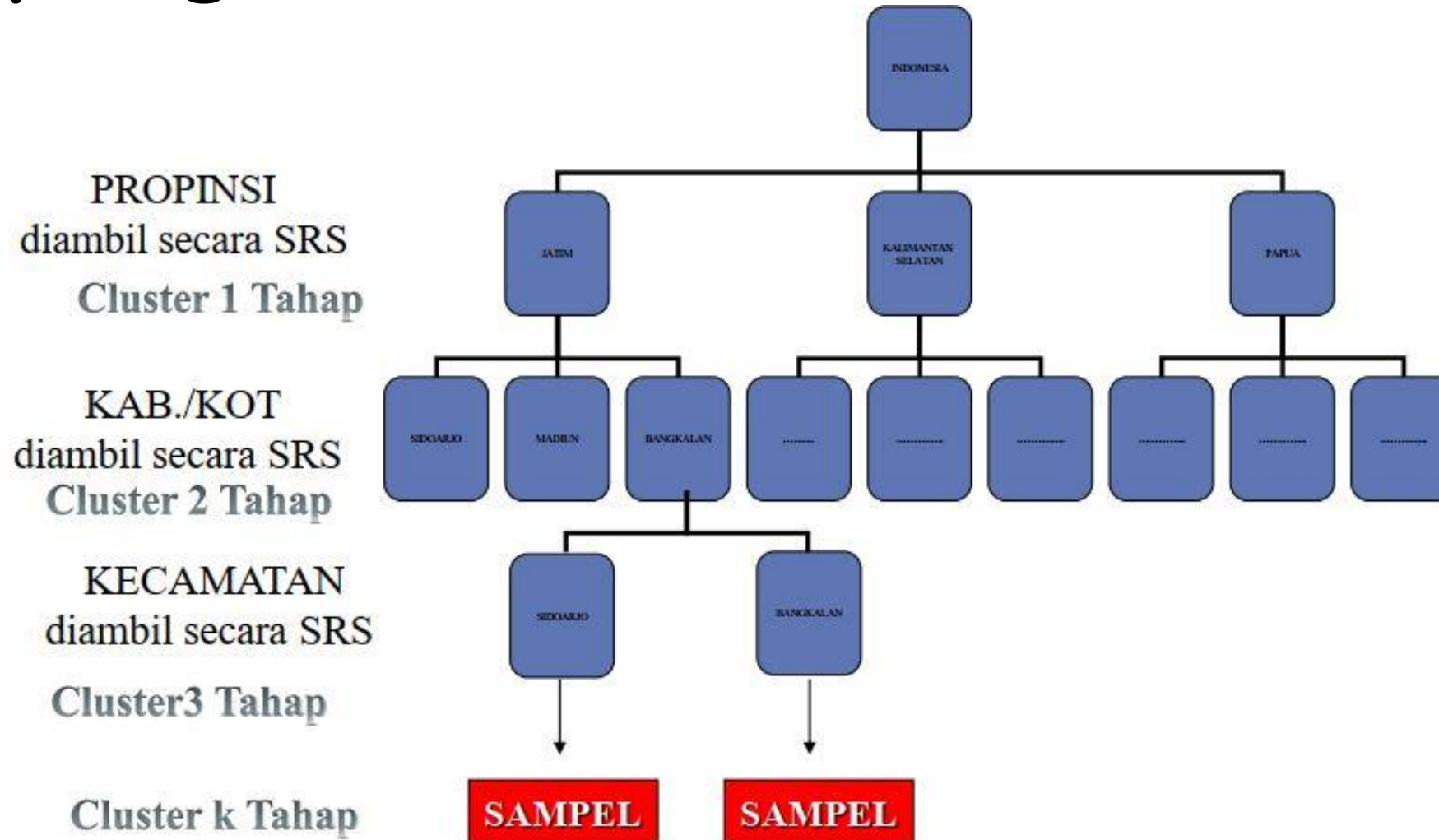
Peneliti akan melakukan penelitian mengenai rata-rata jumlah ternak sapi yang dimiliki peternak sapi di suatu daerah tertentu. Di daerah tersebut terdapat 192 peternak sapi. Kemudian peternak dibagi menjadi klaster yang setiap klaster beranggotakan 16 orang, sehingga klaster yang terbentuk sebanyak 12 klaster. Dari semua klaster dipilih secara acak sebanyak 4 klaster sehingga jumlah sampel yang digunakan adalah 64 orang peternak.

Contoh Two Stage Cluster Sampling



Peneliti akan melakukan penelitian mengenai karakteristik orang tua mahasiswa Prodi S1 Teknologi Sains Data , yaitu proporsi banyaknya tanggungan orang tua (anak). Orang tua mahasiswa tersebut kemudian dibagi menjadi 12 klaster menurut asal daerah/ domisili. Kemudian dipilih secara acak didapatkan hanya 7 klaster yang digunakan dalam pengambilan sampel. Setelah itu, pada masing-masing klaster dipilih kembali secara acak yang menjadi sampel sesuai dengan keinginan dan kebutuhan peneliti. Sehingga jumlah sampel yang digunakan sebanyak 56 orang tua mahasiswa Prodi Teknologi Sains Data Universitas Airlangga.

Contoh K -Stage Cluster Sampling





Rata-rata (μ) & Total (τ)

Ratio estimator of the population mean μ :

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n m_i}$$

Estimated variance of \bar{y} :

$$\hat{V}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s_r^2}{n\bar{M}^2}$$

where

$$s_r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}m_i)^2}{n-1}$$

Here \bar{M} can be estimated by \bar{m} if M is unknown.

Estimator of the population total τ :

$$M\bar{y} = M \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n m_i}$$

Estimated variance of $M\bar{y}$:

$$\hat{V}(M\bar{y}) = M^2 \hat{V}(\bar{y}) = N^2 \left(1 - \frac{n}{N}\right) \frac{s_r^2}{n}$$

Estimator of the population total τ , which does not depend on M :

$$N\bar{y}_t = \frac{N}{n} \sum_{i=1}^n y_i$$

Estimated variance of $N\bar{y}_t$:

$$\hat{V}(N\bar{y}_t) = N^2 \hat{V}(\bar{y}_t) = N^2 \left(1 - \frac{n}{N}\right) \frac{s_t^2}{n}$$

where

$$s_t^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}_t)^2}{n-1}$$

Approximate sample size required to estimate μ , with a bound B on the error of estimation:

$$n = \frac{N\sigma_r^2}{ND + \sigma_r^2}$$

where σ_r^2 is estimated by s_r^2 and $D = (B^2\overline{M}^2)/4$.

Approximate sample size required to estimate τ , using $M\bar{y}$, with a bound B on the error of estimation:

$$n = \frac{N\sigma_r^2}{ND + \sigma_r^2}$$

where σ_r^2 is estimated by s_r^2 and $D = B^2/4N^2$.

Approximate sample size required to estimate τ , using $N\bar{y}_t$ with a bound B on the error of estimation:

$$n = \frac{N\sigma_t^2}{ND + \sigma_t^2}$$

where σ_t^2 is estimated by s_t^2 and $D = B^2/4N^2$.

Proporsi (p)



Estimator of the population proportion p :

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n m_i}$$

Estimated variance of \hat{p} :

$$\hat{V}(\hat{p}) = \left(1 - \frac{n}{N}\right) \frac{s_p^2}{nM^2}$$

where

$$s_p^2 = \frac{\sum_{i=1}^n (a_i - \hat{p}m_i)^2}{n - 1}$$



Non-Probability Sampling



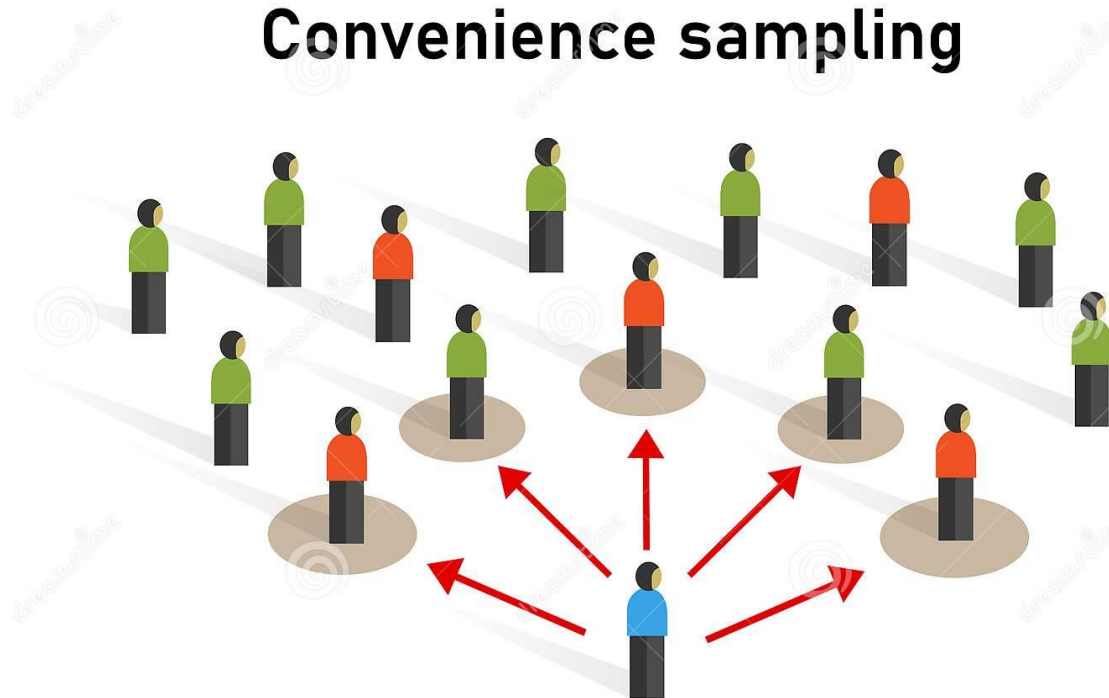


Non-Probability Sampling

- Dalam sampel non-probabilitas, **individu dipilih berdasarkan kriteria non-acak dan tidak setiap individu memiliki peluang untuk disertakan.**
- Jenis sampel ini **lebih mudah dan lebih murah untuk diakses**, tetapi **memiliki resiko bias pengambilan sampel yang lebih tinggi** daripada sampel probabilitas. Artinya kesimpulan yang dibuat tentang populasi lebih lemah dibandingkan sampel probabilitas dan kesimpulan ini mungkin lebih terbatas. Jika peneliti menggunakan sampel non-probabilitas, sebisa mungkin sampel yang terambil mewakili/merepresentasikan populasi.
- Teknik pengambilan sampel non-probabilitas **sering digunakan dalam penelitian eksplorasi dan kualitatif**. Dalam jenis penelitian ini, tujuannya bukan untuk menguji hipotesis tentang populasi dalam scope luas, tetapi untuk mengembangkan pemahaman awal tentang populasi scope kecil atau populasi yang sedang diteliti saja.

Convenience/accidental sampling

Convenience sampling



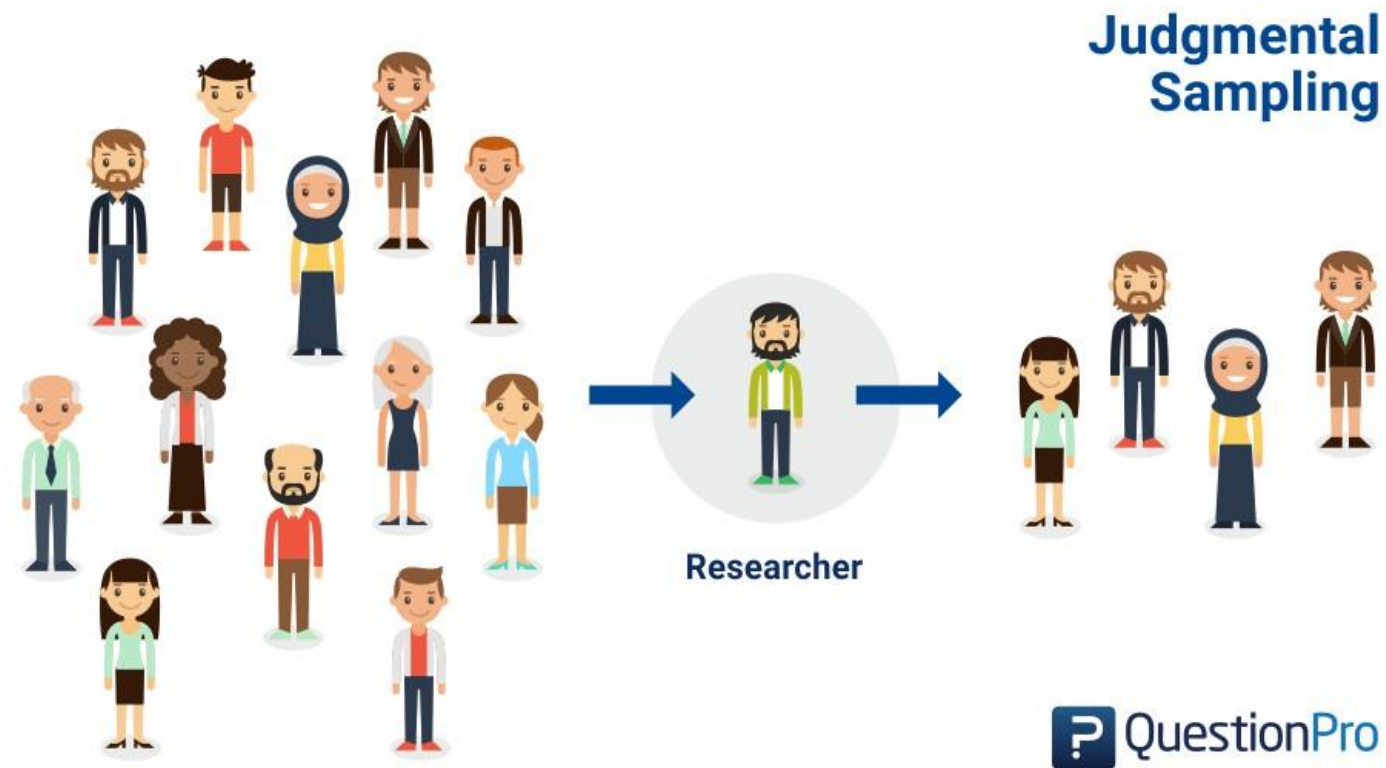
Convenience/accidental sampling



- Sampel diambil berdasarkan pada ketersediaan elemen dan kemudahan untuk mendapatkannya. Sampel terpilih karena berada pada waktu dan tempat yang tepat. Dalam banyak kasus, anggota mudah didekati untuk menjadi bagian dari sampel.
- Kelebihan :
 - Mudah, murah dan cepat (hemat biaya dan waktu)
 - Pengumpulan data dapat dilakukan dengan cepat
 - Tidak mahal untuk membuat sampel
 - Mudah untuk melakukan penelitian
 - Biaya murah
 - Sampel tersedia
 - Lebih sedikit aturan untuk diikuti
- Kekurangan :
 - Hasilnya tidak dapat diandalkan karena tidak ada cara untuk mengetahui apakah sampel sudah mewakili populasi. Sehingga tidak dapat memberikan hasil yang dapat digeneralisasikan.
 - Tidak dapat digunakan bila populasi tidak dapat didefinisikan



Purposive/judgmental sampling



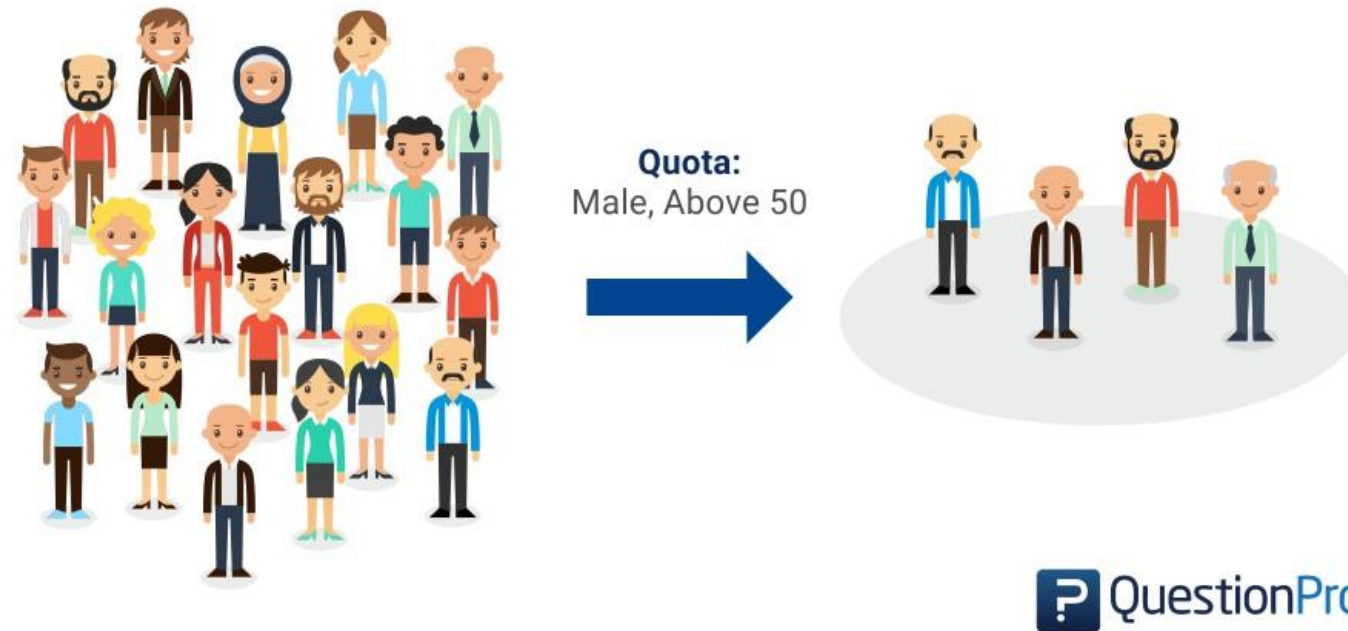
Purposive/judgmental sampling



- Judgment sampling merupakan pengambilan sampel berdasarkan kriteria yang telah ditentukan.
- Pengambilan sampel metode ini dapat digunakan untuk berbagai alasan. Secara umum, tujuan pengambilan sampel adalah untuk sengaja memilih unit (misalnya: individu, peristiwa, objek) yang paling sesuai untuk memungkinkan peneliti menjawab pertanyaan penelitian mereka.
- Hal ini sering dilakukan ketika populasi yang diteliti sangat kecil, atau karakteristik unit yang diinginkan sangat jarang, sehingga pengambilan sampel probabilistik tidak dapat dilakukan.
- Kelebihan :
 - Jika probability sampling tidak dapat digunakan sama sekali
 - Sampel sangat kecil
 - Peneliti memahami obyek penelitian dengan sangat baik
- Kekurangan :
 - Perlu kejelian peneliti dalam mendefinisikan populasi dan membuat pertimbangannya.

Quota sampling

Quota Sampling



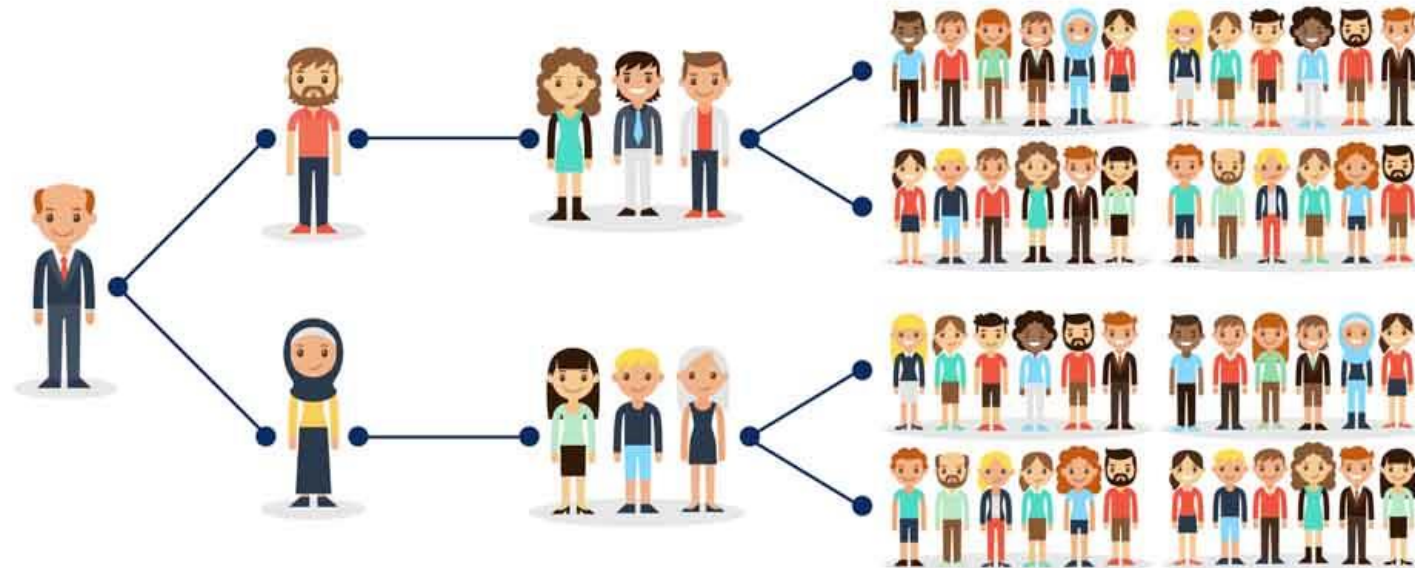
Quota sampling



- Pengambilan sampel kuota didefinisikan sebagai metode pengambilan sampel non-probabilitas di mana peneliti membuat sampel yang melibatkan individu yang mewakili suatu populasi. Peneliti memilih individu-individu ini menurut sifat atau kualitas tertentu. Sampel ini dapat digeneralisasikan untuk seluruh populasi. Bagian terakhir akan ditentukan hanya menurut pengetahuan pewawancara atau peneliti tentang populasi.
- Kelebihan :
 - Menghemat waktu
 - Kenyamanan penelitian
 - Menghemat uang
- Kekurangan :
 - tingginya tingkat kesulitan dalam merumuskan hasil penelitian karena data yang diperoleh sangat beragam
 - tidak adanya prosedur baku bagi pewawancara dan teknik wawancara akan berpengaruh pada terjadinya bias

Snowball sampling

SNOWBALL SAMPLING



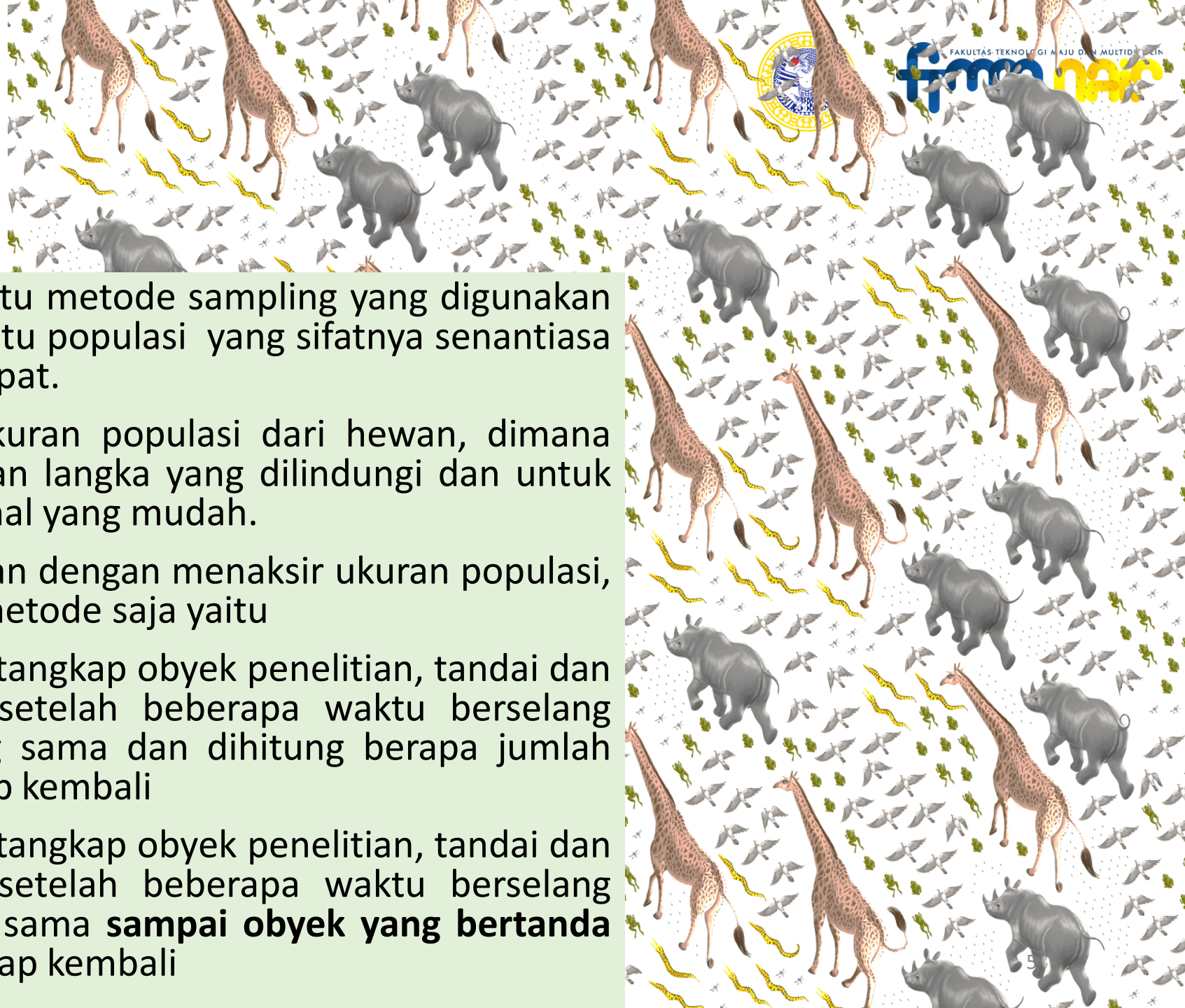


Snowball sampling

- Pengambilan sampel snowball didefinisikan sebagai teknik pengambilan sampel non-probabilitas di mana sampel memiliki sifat yang jarang ditemukan. Dalam teknik pengambilan sampel ini, subyek yang ada memberikan rujukan untuk merekrut sampel yang diperlukan untuk studi penelitian.
- Kelebihan :
 - Peneliti dapat menjangkau populasi yang sulit diambil sampelnya
 - Prosesnya murah, sederhana dan hemat biaya
 - Membutuhkan sedikit perencanaan dan tenaga kerja
- Kekurangan :
 - Peneliti memiliki sedikit kendali atas metode pengambilan sampel.
 - Keterwakilan sampel tidak dijamin.
 - Bias pengambilan sampel

Wildlife sampling

- Wildlife sampling merupakan suatu metode sampling yang digunakan untuk memperkirakan ukuran suatu populasi yang sifatnya senantiasa bergerak / berpindah-pindah tempat.
- Umumnya untuk mengetahui ukuran populasi dari hewan, dimana hewan ini merupakan jenis hewan langka yang dilindungi dan untuk menjumpainya pun bukan suatu hal yang mudah.
- Ada beberapa metode sehubungan dengan menaksir ukuran populasi, namun hanya akan dibahas dua metode saja yaitu
 - Direct Sampling. Prosedurnya : tangkap obyek penelitian, tandai dan lepaskan kembali. Kemudian setelah beberapa waktu berselang ditangkap lagi di wilayah yang sama dan dihitung berapa jumlah obyek yang bertanda tertangkap kembali
 - Invers Sampling. Prosedurnya : tangkap obyek penelitian, tandai dan lepaskan kembali. Kemudian setelah beberapa waktu berselang ditangkap lagi di wilayah yang sama **sampai obyek yang bertanda** dalam jumlah tertentu tertangkap kembali





Bagaimana cara menentukan teknik sampling ??

- Catat tujuan penelitian. Umumnya, yang perlu diperhatikan yaitu kombinasi dari biaya, presisi, atau akurasi.
- Identifikasi teknik pengambilan sampel yang efektif yang berpotensi mencapai tujuan penelitian.
- Ujilah masing-masing metode ini dan periksa apakah mereka membantu dalam mencapai tujuan penelitian.
- Pilih metode yang paling sesuai untuk penelitian.



Probability vs nonprobability sampling

	Probability sampling	Non-probability sampling
Definisi	Probability sampling merupakan teknik pengambilan sampel, dimana subjek dari populasi mendapatkan kesempatan yang sama untuk dipilih sebagai sampel yang representatif	Nonprobability sampling merupakan salah satu metode pengambilan sampel dimana tidak diketahui individu mana dari populasi yang akan dipilih sebagai sampel
Dasar pemilihan sampel	Random (acak)	Arbitrarily (sembarang)
Peluang sampel terpilih	Ditentukan dan diketahui	Tidak ditentukan & tidak diketahui
Penelitian	Menghasilkan kesimpulan / inferensia	Eksplorasi
Hasil sampling	Tak bias	Bias
Metode	Obyektif	Subyektif



TERIMA KASIH

-Tim Dosen Metode Statistika-

