

TEKNIK SAMPLING DAN DATA WRANGLING

Table of Content

- Sampling & Sensus
- Why Sampling?
- +- Sensus
- Unsur” Sampling
- Populasi & Sampel
- Random
- Parameter & Statistik

Pengertian Teknik Sampling

Metode/Cara pengumpulan Informasi (Data) terhadap sebagian anggota populasi

Tujuan

- Menarik kesimpulan atau perkiraan tentang sifat-sifat (karakteristik) suatu populasi dengan hanya memakai data dari sampel.
- Untuk membuat Sampling lebih efisien dengan biaya rendah dan lebih teliti

Metode Pengumpulan Data

1. Sensus = Mengumpulkan data dari **seluruh** elemen dalam **populasi**
2. Sampling (Survei) = Mengumpulkan data dari **sebagian** elemen **populasi**

Macam-Macam Sampling

1. Probability Sampling : setiap elemen dari populasi mempunyai **probabilitas** kesempatan **yang sama** untuk dipilih menjadi sampel.

- menggunakan kaidah peluang (**probability**) dalam pemilihan **sampel**
- hasil survey digunakan untuk melakukan pendugaan (**estimasi**) terhadap **karakteristik populasi**

2. Non-Probability Sampling : setiap elemen yang **dipilih** hanya berdasarkan **pertimbangan** dari **peneliti**.

- **tidak** menggunakan kaidah peluang (**probability**) dalam pemilihan **sampel**
- hasil survey **tidak** digunakan untuk melakukan pendugaan (**estimasi**) terhadap **karakteristik populasi**

Keuntungan dan Kelemahan Sensus

Keuntungan (+)

- 1.Dapat menyajikan data wilayah kecil
- 2.Dapat dijadikan kerangka sampel (frame)

Kelemahan (-)

- 1.Cakupan variabel terbatas
- 2.Waktu lama
- 3.Biaya besar
- 4.Ketelitian kurang

Kenapa menggunakan sampling?

1.Sumber daya terbatas

2.Waktu yang tersedia terbatas

3.Pengamatan kadang bersifat merusak

4.Mustahil mengamati seluruh anggota populasi

Populasi dan Sampel

Populasi

Merupakan kesuluran objek yang karakteristiknya sama (hendak diuji)

Sampel

Merupakan bagian populasi yang karakteristiknya hendak diuji

No	Rincian	Populasi	Sampel
1	Nilai karakteristik unit ke- i	Y_i	y_i
2	Rata-rata nilai karakteristik	\bar{Y}	$\bar{y} = \hat{Y}$
3	Total nilai karakteristik	Y	\hat{Y}
4	Banyaknya unit sampling	N	n
5	Varians	S^2	s^2
6	Proporsi	P	$p = \hat{P}$
7	Rasio	R	$r = \hat{R}$

Random

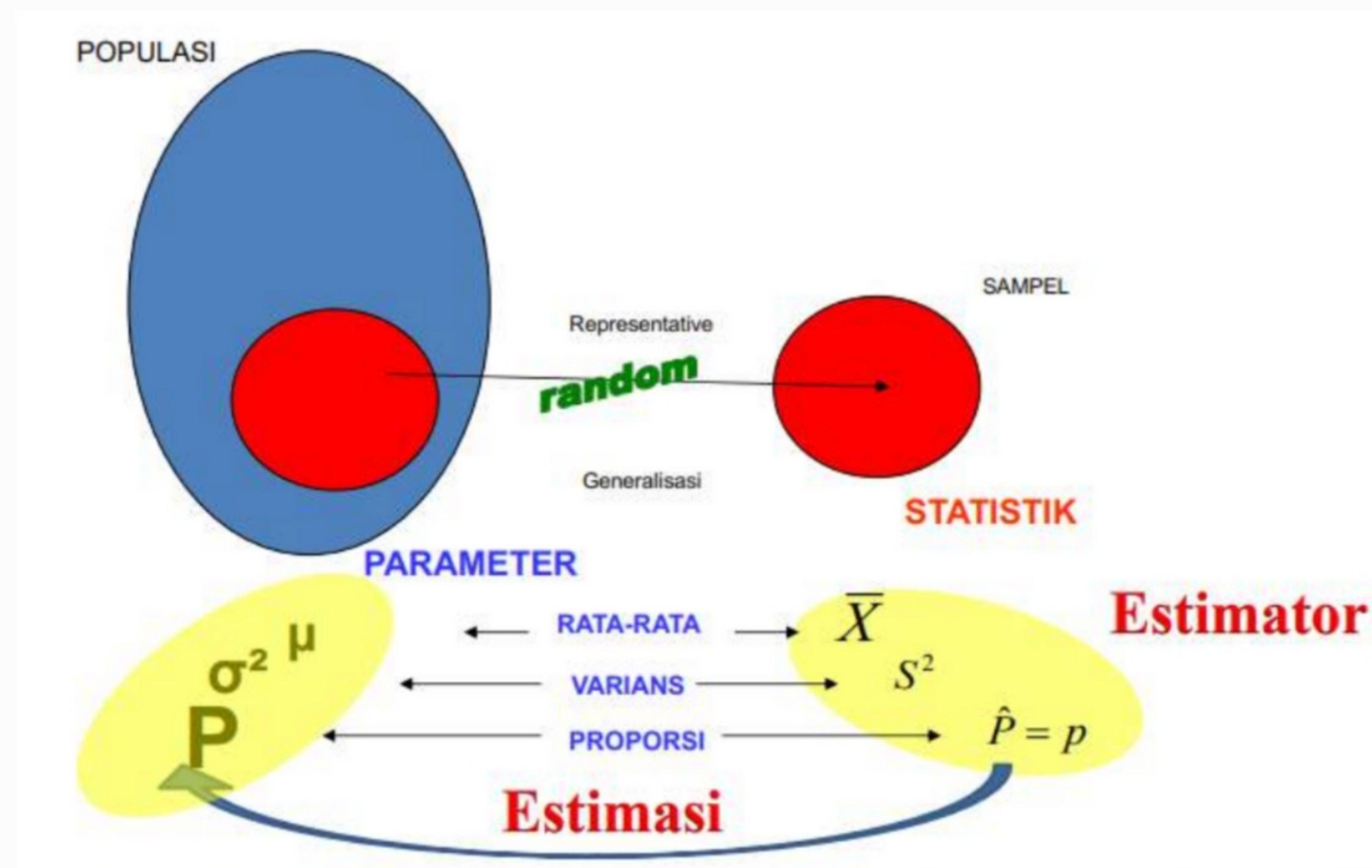
Definisi

Merupakan suatu proses dimana setiap elemen dalam populasi **memiliki peluang yang sama** untuk dipilih menjadi **bagian** dari **sampel**

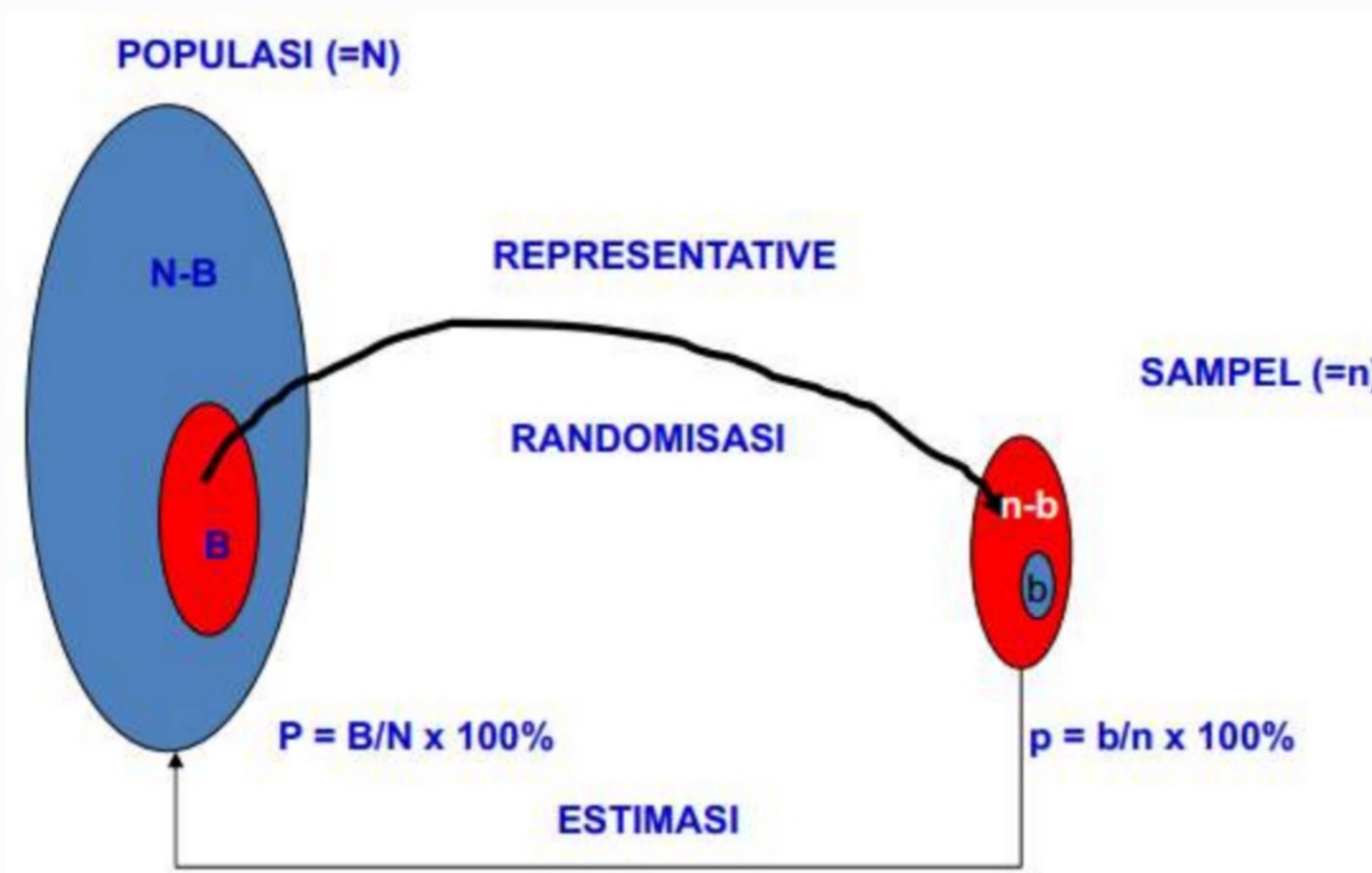
Contoh dari suatu proses yg Random

Pelemparan dadu -> Ketika sebuah dadu dilempar, hasilnya bisa berupa angka 1, 2, 3, 4, 5, atau 6, dan setiap angka **memiliki peluang yang sama** untuk **muncul**, yaitu $1/6$. Hasil dari pelemparan ini **tidak dapat diprediksi** dengan pasti sebelum dadu dilempar, sehingga merupakan contoh dari proses random

Parameter & Statistik



Contoh Proses Sampling



Konsep Dasar Teori Sampling

Populasi

Suatu agregat dari semua elemen, membagi beberapa himpunan dengan karakteristik yang sama yang disesuaikan dengan tujuan umum dari masalah dalam suatu riset.

Macam-macam Populasi Berdasarkan Jumlah

- Finite Population (dipelajari pada matkul ini) = unitnya terbatas
- Infinte Population = unitnya tidak terbatas

Konsep Dasar Teori Sampling

1. Populasi Target

sub populasi dari elemen yang ada pada populasi yang berbagai indikatornya akan dicari

2. Unit Sampling

Suatu subgroup dari elemen-elemen populasi yang dipilih untuk berpartisipasi dalam suatu riset

3. Karakteristik

Suatu subgroup dari elemen-elemen populasi yang dipilih untuk berpartisipasi dalam suatu riset.

4. Nilai Karakteristik

yang dihitung (diestimasi) dapat berupa rata-rata, total, rasio, proporsi, persentase, dan sebagainya

Konsep Dasar Teori Sampling

5. Elemen

unit yang digunakan untuk mendapatkan informasi, misalnya individu, rumah tangga, perusahaan, dsb

6. Unit Observasi

unit dimana informasinya diperoleh baik secara langsung maupun melalui responden tertentu

7. Unit sampling

unit yang dijadikan dasar penarikan sampel baik berupa elemen maupun kumpulan elemen (klaster)

8. Unit analisis

unit yang digunakan pada tahap tabulasi data, bisa berupa elemen atau kumpulan elemen

9. Kerangka Sampel

daftar semua unit yang akan dijadikan sampling unit

Contoh Kasus

Suatu survei bertujuan untuk memperkirakan total biaya produksi dari rumah tangga yang mempunyai UMKM di Desa Waru Sidoarjo pada Agustus 2024

Definisikan apa yang menjadi **populasi, populasi target, kerangka sampel, unit sampling, unit observasi, karakteristik yang diteliti, nilai karakteristik yang diestimasi, dan unit analisisnya**

Populasi	Semua rumah tangga di Desa Waru Sidoarjo pada Agustus 2024
populasi target	Semua rumah tangga yang memiliki UMKM di Desa Waru Sidoarjo pada Agustus 2024
kerangka sampel	Daftar rumah tangga yang memiliki UMKM di Desa Waru Sidoarjo pada Agustus 2024
unit sampling	Rumah tangga yang mempunyai UMKM
unit observasi	Rumah tangga yang mempunyai UMKM (Kepala Keluarga)
karakteristik yang diteliti	Biaya Produksi
nilai karakteristik yang diestimasi	Total
unit analisisnya	Rumah tangga yang mempunyai UMKM

WHY SAMPLING?

Proses Desain Sampling

- 1. Definisikan Populasinya**
- 2. Tentukan Sampling Frame**
- 3. Pilih Teknik**
- 4. Tentukan Ukuran Sampel**
- 5. Tentukan Proses Sampling**

Contoh Sampling

1. Probability Sampling

- Simple Random Sampling
- Systematic Sampling
- Stratified Sampling
- Cluster Sampling

2. Non-Probability Sampling

- Convenience Sampling
- Judgmental Sampling
- Quota Sampling
- Snowball Sampling

Probability Sampling

1. Simple Random Sampling

setiap elemen dari populasi telah diketahui dan mempunyai probabilitas yang sama untuk terpilih.

2. Stratified Sampling

suatu sampel yang diambil dari beberapa level/strata yang terdapat dalam populasi

3. Cluster Sampling

Sampel yang dipilih dari populasi yang dibagi menjadi kelompok (area sampling) dan setiap unitnya dipilih secara acak.

4. Systematic Sampling

Sampel yang dipilih dengan cara memilih secara random titik awal dan memilih setiap elemen ke-i dari sampling frame

Non-Probability Sampling

1. Convenience Sampling

Ambil sampel dari orang yang paling mudah dijangkau (contoh: survei di jalan).

2. Judgmental Sampling

Peneliti pilih sampel karena dianggap paling sesuai tujuan riset.

3. Quota Sampling

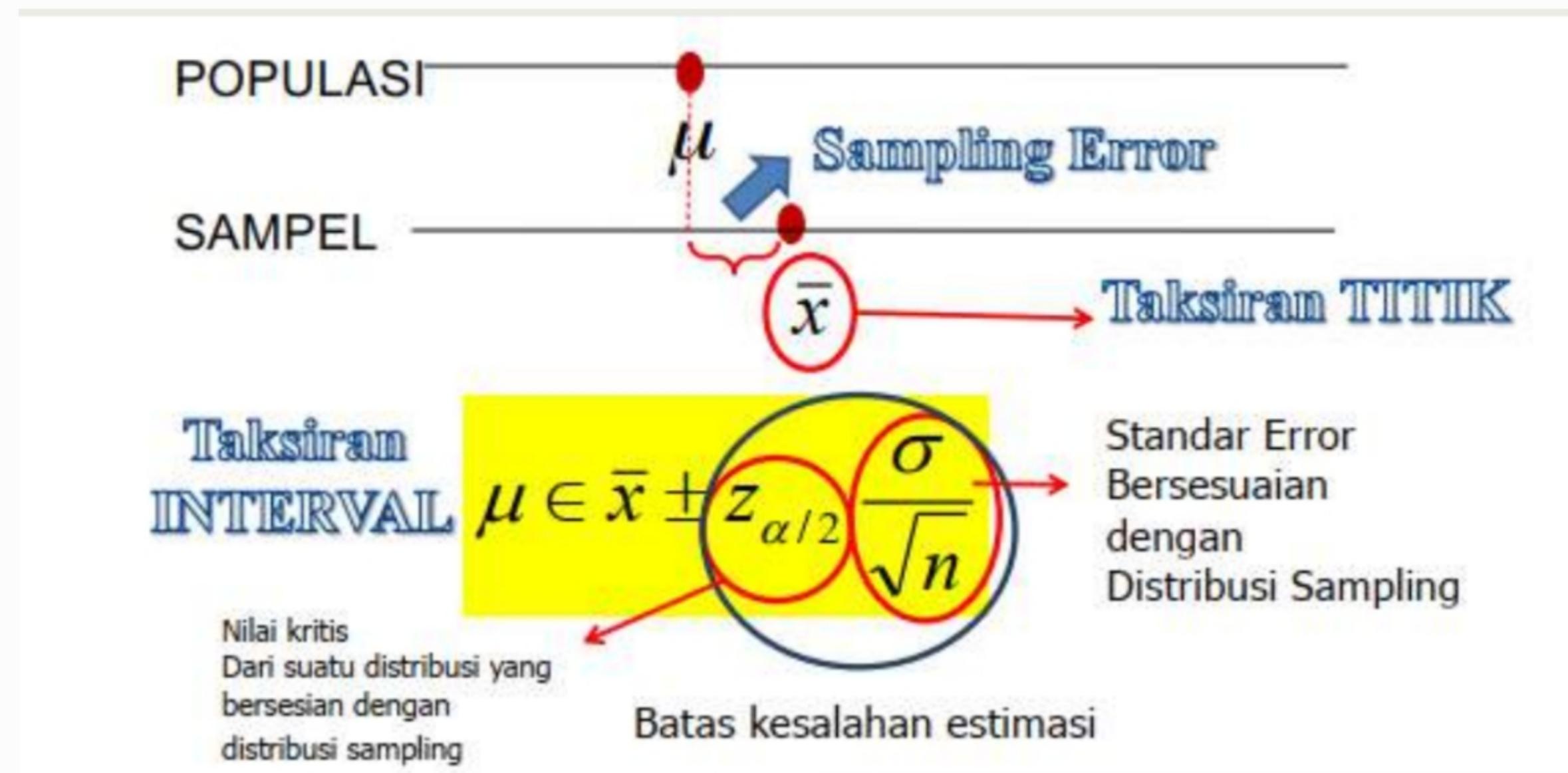
Tentukan kategori populasi, lalu ambil sampel sesuai kuota tiap kategori.

4. Snowball Sampling

Responden pertama menunjuk responden berikutnya, dan seterusnya.

Probability	Non-Probability
Besar Sampel dengan Rumus	Besar Sampel tanpa Rumus
Menggunakan Konsep Peluang	Tanpa Konsep Peluang
Gambaran Populasi Akurat	Gambaran Populasi Kasar

Konsep dasar Estimasi



Estimator

Syarat Estimator

Penduga ($\hat{\theta}$) dikatakan *unbiased estimator* dari parameter θ jika *expected value*-nya sama dengan θ

$$E(\hat{\theta}) = \theta$$

Jika $E(\hat{\theta}) \neq \theta$, Maka penduga $\hat{\theta}$ bisa dikatakan *biased estimator* dari θ

Kesalahan dalam pengumpulan Data

Sampling Error

kesalahan yang terjadi karena proses penarikan sampel

Non-Sampling Error

1. *Mis-specification of sample subjects* : kesalahan dalam penentuan responden.
2. *Coverage Error* : ketidaklengkapan daftar cakupan.
3. *Non Response Error* : ketidaklengkapan respon.
4. *Measurement Error* : kesalahan dalam pengukuran.

Simple Random Sample

Simple Random Sampling With Replacement (SRSWR)

Setelah satu elemen dipilih, elemen itu dikembalikan lagi ke populasi (satu elemen bisa terpilih lebih dari sekali)

$$\frac{(N + n - 1)!}{n!(N - 1)!}$$

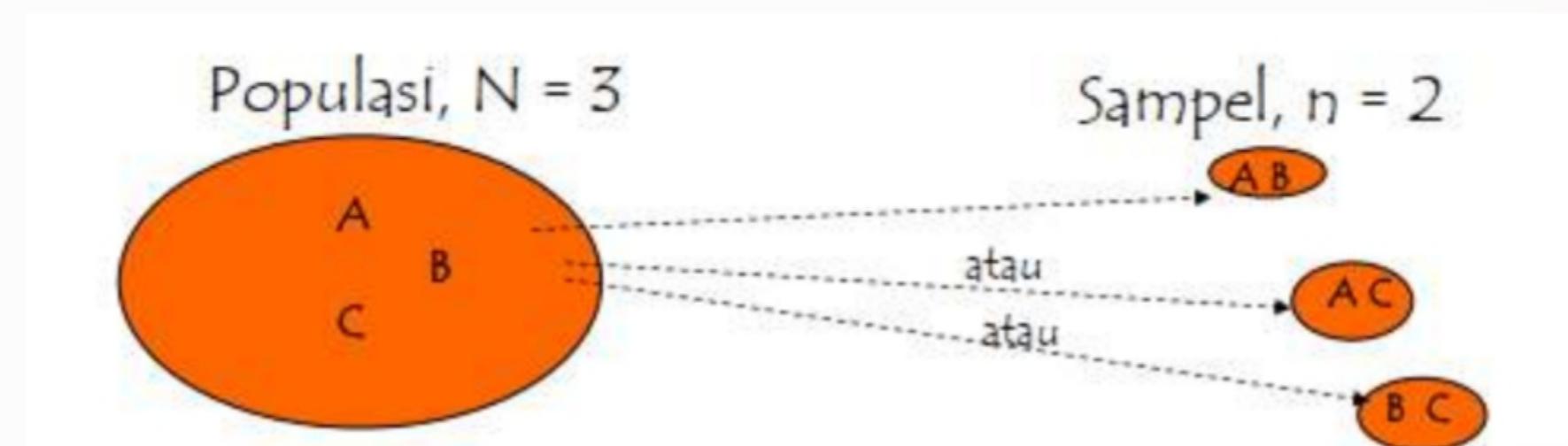
Simple Random Sampling Without Replacement (SRSWOR)

Setelah satu elemen dipilih, elemen itu tidak dikembalikan (satu elemen hanya bisa terpilih sekali)

$$C_n^N = \frac{N!}{(N - n)! n!}$$

Konsep SRS

Populasi berukuran $N = 3$ diambil sampel berukuran $n = 2$, maka banyak sampel yang mungkin adalah 3 yaitu AB, AC, dan BC yang masing-masing mempunyai peluang/kesempatan yang sama besar untuk terpilih menjadi sampel, dengan peluang sebesar $1/3$.



Keuntungan & Kelemahan **SRS**

Keuntungan:

- cara pengambilan sampel dan teknik estimasi parameteranya sederhana.

Kelemahan:

- hanya cocok untuk populasi yang relatif homogen
- hanya cocok untuk cakupan survei yang tidak terlalu luas, karena membutuhkan kerangka sampel sampai elemen
- biaya tinggi untuk populasi yang besar.

Contoh Soal

https://docs.google.com/spreadsheets/d/1Ea1mnZ_gsIC0tmRxcUV6_ah1JJbSr22sF7tOYx7e9w/edit?gid=0#gid=0

Estimasi Rata-Rata

Nilai yang diestimasi	SRS	
	WR	WOR
Rata-rata	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$	
Varians rata-rata	$v(\bar{y}) = \frac{s^2}{n}$	$v(\bar{y}) = \frac{N - n}{N} \cdot \frac{s^2}{n}$
Standar error	$se(\bar{y}) = \sqrt{v(\bar{y})}$	
Relative standar error (RSE)	$rse(\bar{y}) = \frac{se(\bar{y})}{\bar{y}} \times 100\%$	
(1 - α)% Confidence Interval	$\bar{y} - Z_{\alpha/2} \cdot se(\bar{y}) < \bar{Y} < \bar{y} + Z_{\alpha/2} \cdot se(\bar{y})$	

Estimasi Total

Nilai yang diestimasi	SRS	
	WR	WOR
Total	$\hat{Y} = \frac{N}{n} \sum_{i=1}^n y_i = N\bar{y}$	
Varians total	$v(\hat{Y}) = N^2 \cdot \frac{s^2}{n} = N^2 \cdot v(\bar{y})$	$v(\hat{Y}) = N^2 \cdot \frac{N-n}{N} \cdot \frac{s^2}{n} = N^2 \cdot v(\bar{y})$
Standar error	$se(\hat{Y}) = \sqrt{v(\hat{Y})}$	
Relative standar error (RSE)	$rse(\hat{Y}) = \frac{se(\hat{Y})}{\hat{Y}} \times 100\%$	
(1 - α)% Confidence Interval	$\hat{Y} - Z_{\alpha/2} \cdot se(\hat{Y}) < Y < \hat{Y} + Z_{\alpha/2} \cdot se(\hat{Y})$	

Estimasi Proporsi

Nilai yang diestimasi	SRS	
	WR	WOR
Rata-rata	$p = \frac{1}{n} \sum_{i=1}^n y_i$	
Varians rata-rata	$v(p) = \frac{pq}{n-1}$	$v(p) = \frac{N-n}{N} \cdot \frac{pq}{n-1}$
Standar error	$se(p) = \sqrt{v(p)}$	
Relative standar error (RSE)	$rse(p) = \frac{se(p)}{p} \times 100\%$	
(1 - α)% Confidence Interval	$p - Z_{\alpha/2} \cdot se(p) < P < p + Z_{\alpha/2} \cdot se(p)$	

Thank You