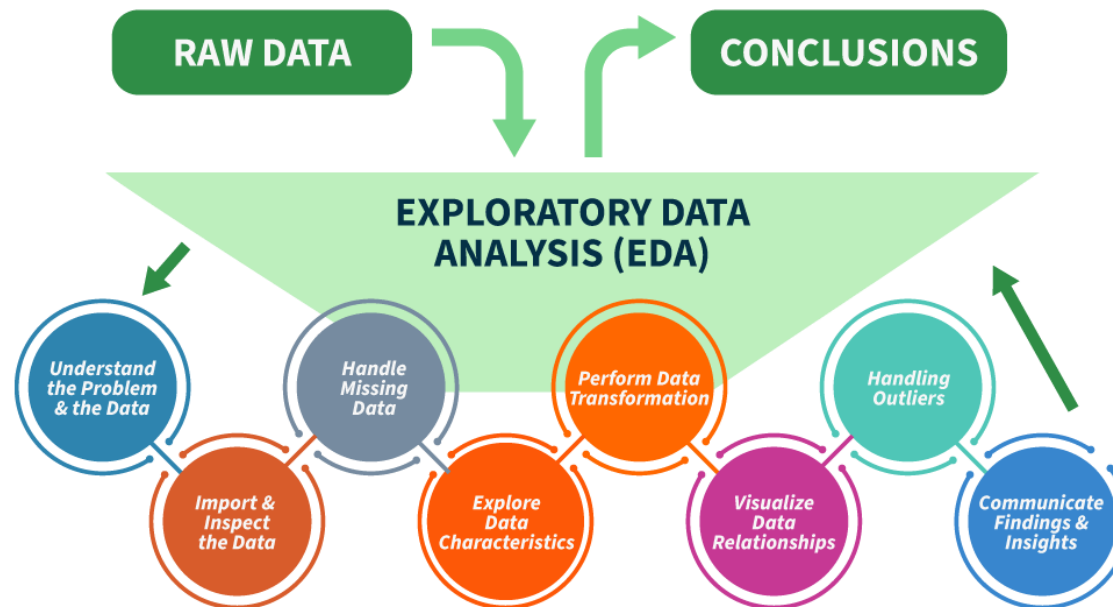


# INTRODUCCIÓN A LOS DATOS



Profesor: Pedro Saa ([pnsaa@uc.cl](mailto:pnsaa@uc.cl))

Año: 1-2025

# OBJETIVOS DE APRENDIZAJE

---

- ▶ **O1:** Comprender los principios de tidy data
- ▶ **O2.** Identificar y explicar los 4 principios del diseño experimental: controlar, replicar, bloquear y aleatoria
- ▶ **O3:** Clasificar un estudio como observacional o experimental, y que tipo de conclusiones se pueden obtener (generalizables o causales)
- ▶ **O4:** Distinguir entre muestreo aleatorio, simple y estratificado
- ▶ **O5.** Explicar la diferencia entre muestreo aleatorio y asignación aleatoria y cuál la ventaja que tiene cada una.

# Principios del diseño experimental

---

# ¿Cuáles son los principios del diseño experimental?

## Aleatorizar

Reducir el sesgo de estudio o incorporar la variabilidad de la población en mi unidad experimental

## Replicar

Existe una variación natural en distintas unidades experimentales, esta se reduce replicando el estudio o experimento

## Fijar/Controlar

Fijar una variable de estudio que puede introducir variación a mi unidad experimental

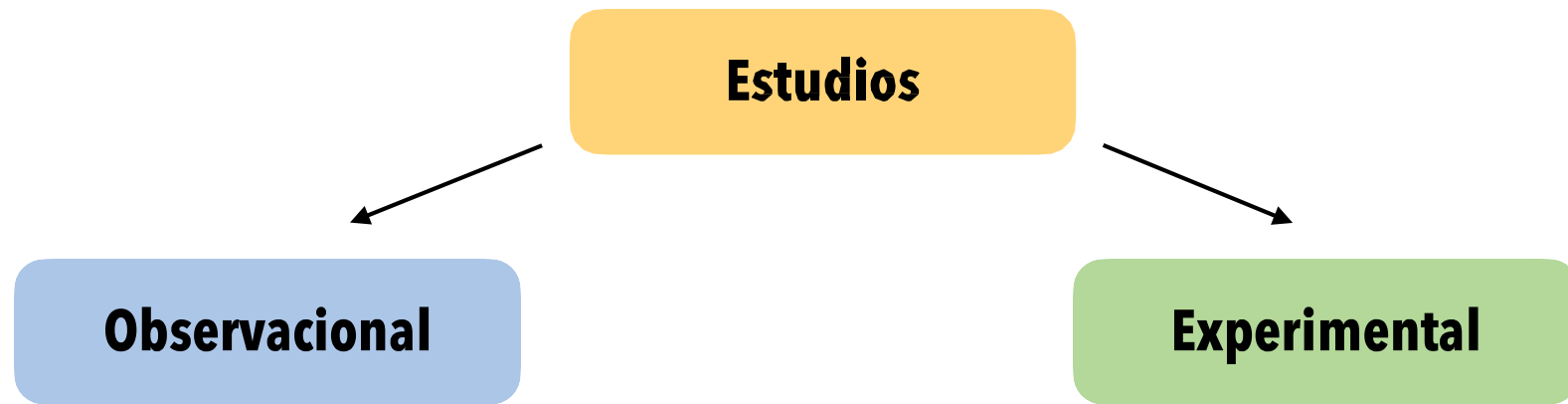
## Bloquear

Bloquear por variables que sospechas y **no tienes control** sobre ellas que pueden afectar tu estudio

# Tipos de estudio

---

Existen dos macro clasificaciones de tipos de estudios, estos son **observacionales** y **experimentales**, la diferencia entre ellos recae en como se reúnen los datos...



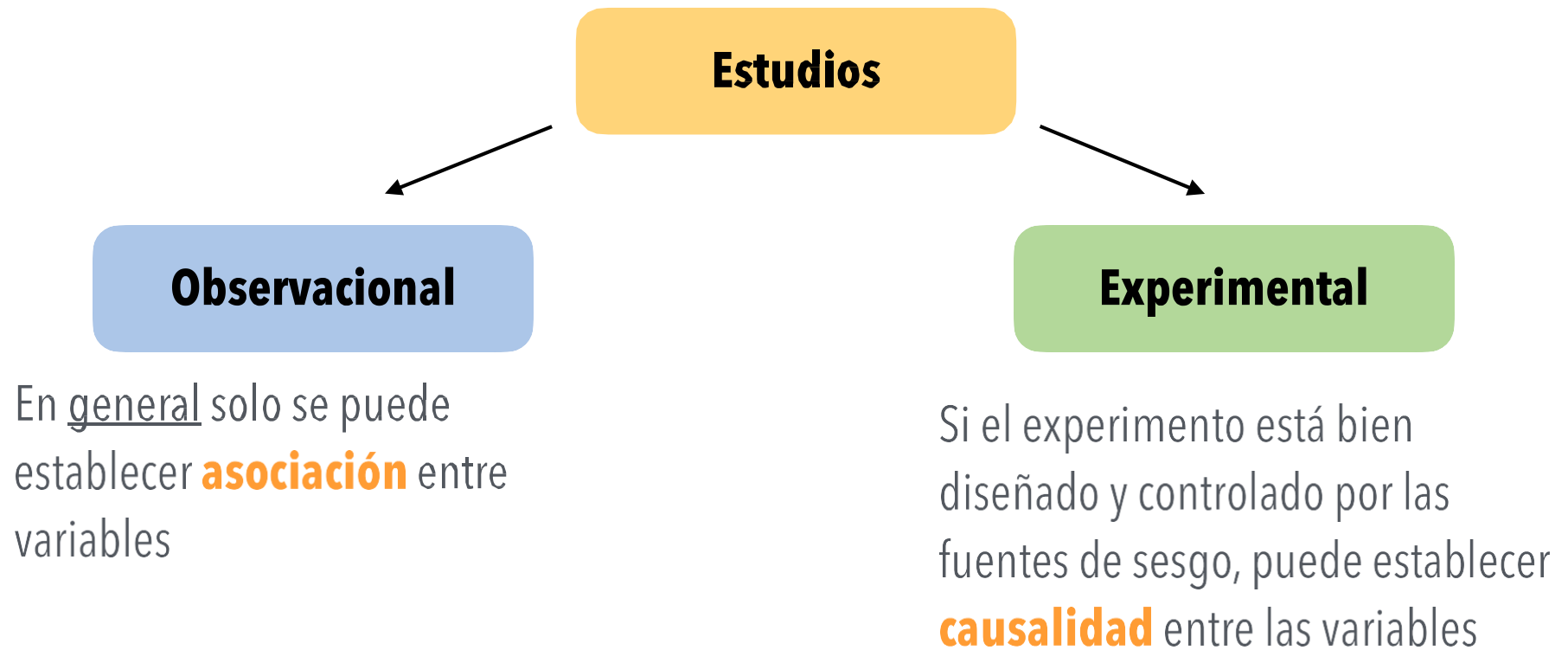
Reúnen datos de tal forma que no interviene en como el dato se origina

Se define un estudio **retrospectivo** cuando ocupa data reunida en el pasado

Se define un estudio **prospectivo** la data se reúne durante el estudio

Investigadores asignan de forma aleatorias a sujetos de estudio a distintos tipos de tratamiento

Existen dos macro clasificaciones de tipos de estudios, estos son **observacionales** y **experimentales**, la diferencia entre ellos recae en como se reúnen los datos...



**Nunca olvidar...**  
**Correlación no implica causalidad**



# Tipos de muestreo

---

# ¿Por qué se toman muestras de la población y no se ocupa toda la población en el estudio?

- Muy Costoso
- Impráctico
- Lento de realizar
- ...

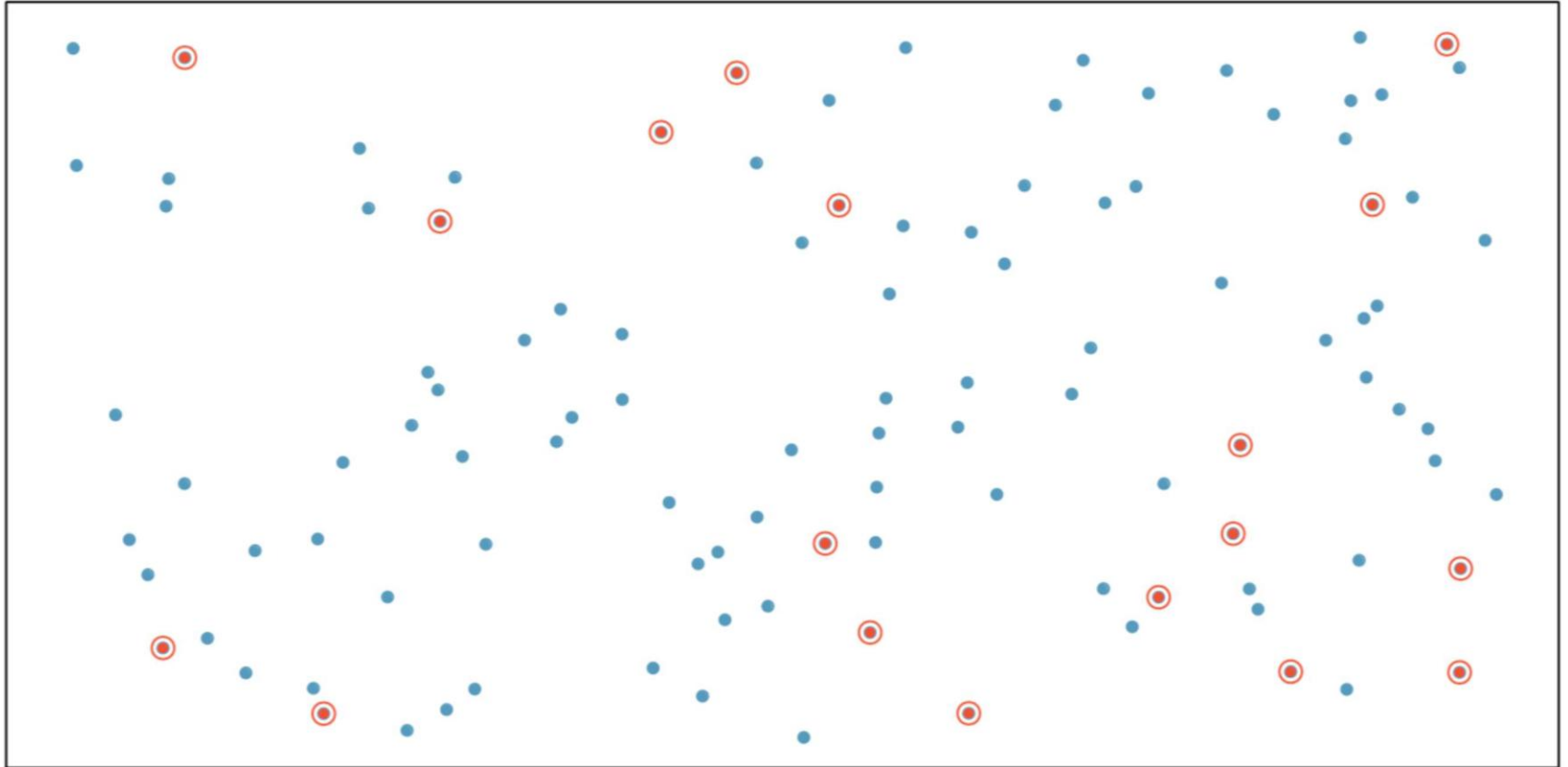
Para conocer el estado de la sopa, no necesitas tomarte toda la sopa.

Puedes hacer un **análisis exploratorio**, tomar un **muestreo representativo**, y hacer una **inferencia** de como está

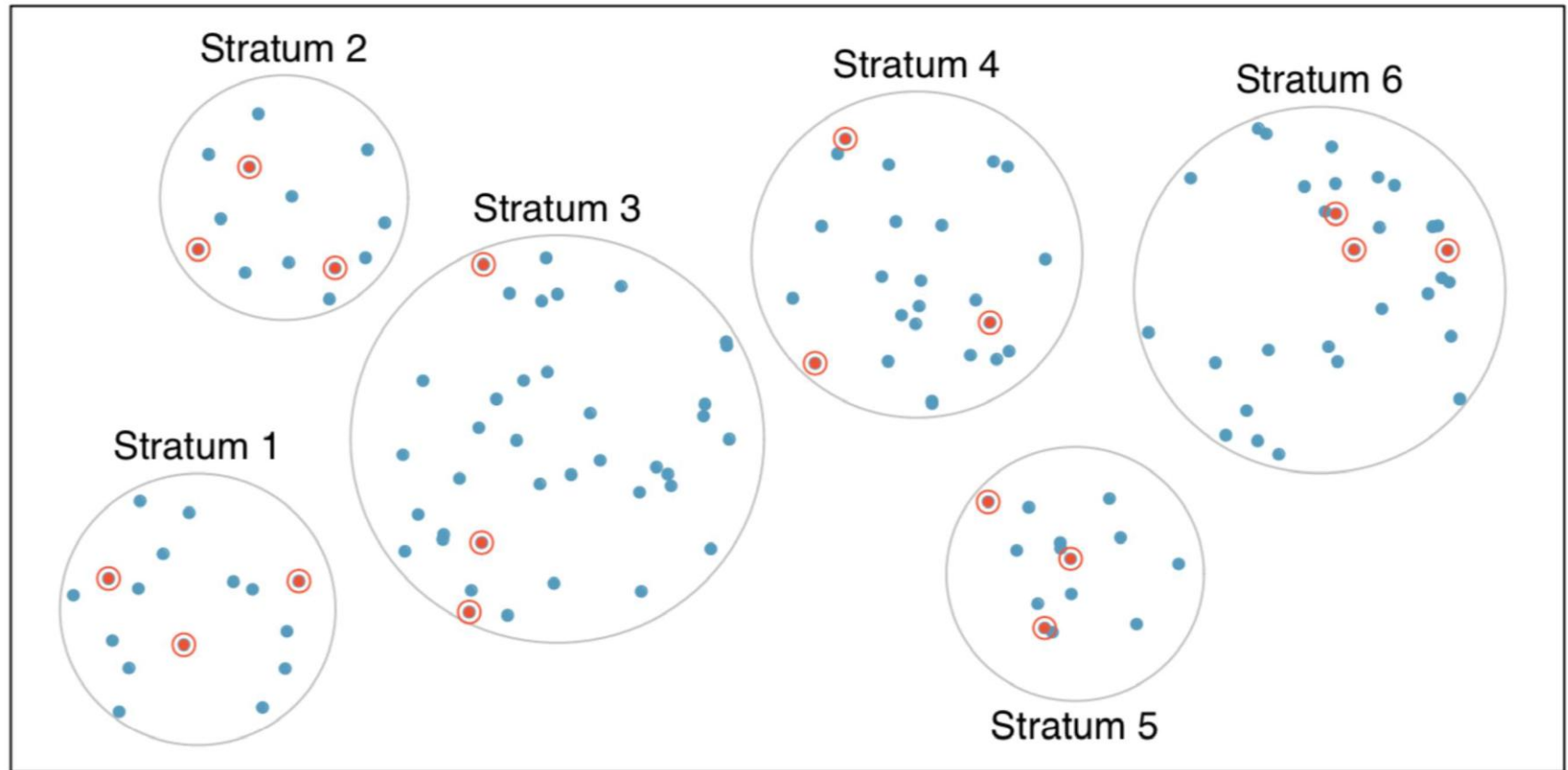


La importancia de que el **muestreo representativo sea aleatorio** es que elimina las fuentes naturales de **sesgo** al incorporar la **variabilidad** que tiene la población

**Tipos de muestreo:** El **muestreo aleatorio simple** es un muestreo en que cada caso de la población tiene la **misma** probabilidad de ser escogida



**Tipos de muestreo:** En un **muestreo estratificado** la población se divide en estratos homogéneos y se toma una muestra al azar dentro de cada estrato para una mejor representación de la población



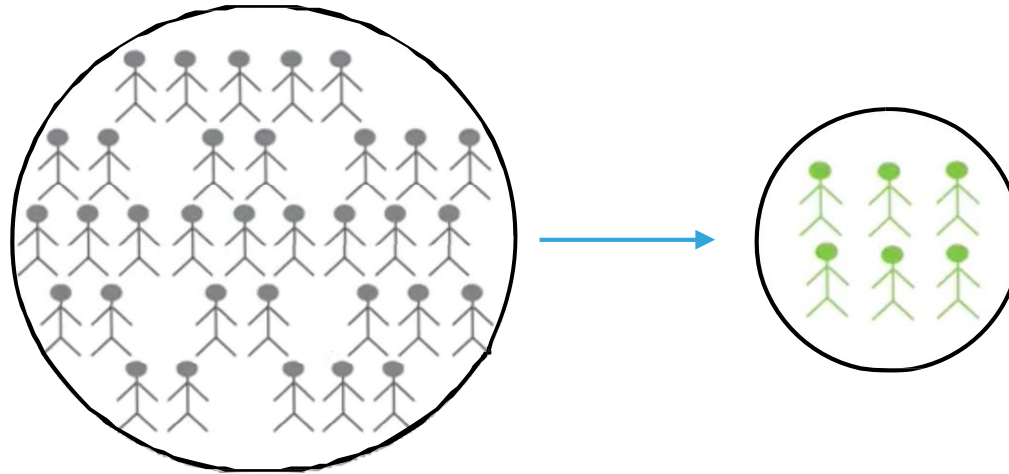
**Ejemplo:** Estudiar el nivel de contaminación de empresas mineras. Puede que la distribución de emisiones de las empresas sean muy dispares, por ejemplo 60% tienen emisiones medias, 30% emisiones muy bajas y 10% con emisiones muy altas.

# Muestreo aleatorio vs Asignación aleatoria

---

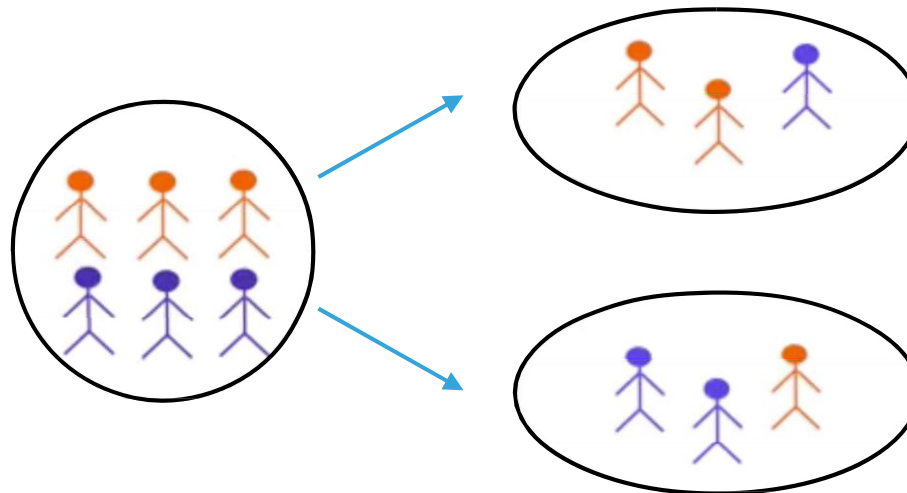
El **muestreo aleatorio** y la **asignación aleatoria** cumplen objetivos distintos, el primero permite **generalizar** los experimentos o estudios y el segundo permite determinar **causalidad** de un diseño experimental

## Muestreo aleatorio



**Generalizable**

## Asignación Aleatoria



**Causalidad**

El **muestreo aleatorio** y la **asignación aleatoria** cumplen objetivos distintos, el primero permite **generalizar** los experimentos o estudios y el segundo permite determinar **causalidad** de un diseño experimental

	Random assignment	No random assignment	
Random sampling	causal and generalizable	not causal, but generalizable	Generalizability
No random sampling	causal, but not generalizable	neither causal nor generalizable	No generalizability
	Causation	Association	

*ideal experiment* (arrow pointing to Random assignment / Random sampling)

*most observational studies* (arrow pointing to No random assignment / Generalizability)

*most experiments* (arrow pointing to No random sampling / Causation)

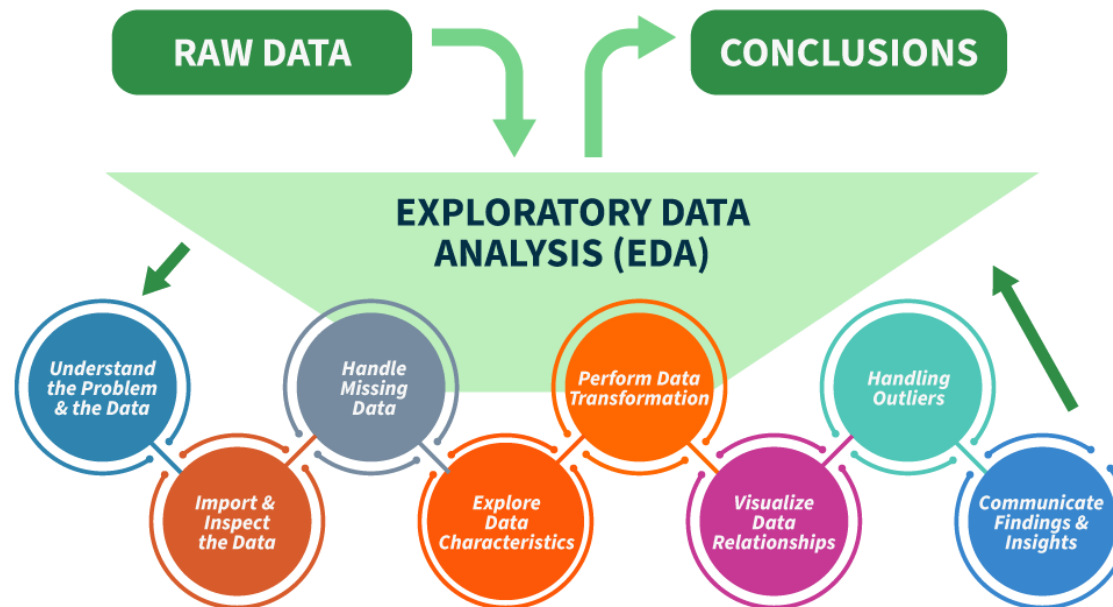
*bad observational studies* (arrow pointing to No random assignment / No generalizability)



# Resumen

- Idealmente nos interesa inferir respecto de una población **determinada**, sin embargo, raramente es posible “muestrearla” completamente.
- Diferentes tipos de muestreo de la población afectarán las conclusiones que podremos extraer.
- Los muestreos aleatorios permiten **generalizar**, mientras que las asignaciones aleatorias permiten determinar **causalidad**.

# INTRODUCCIÓN A LOS DATOS



Profesor: Pedro Saa ([pnsaa@uc.cl](mailto:pnsaa@uc.cl))

Año: 1-2025