



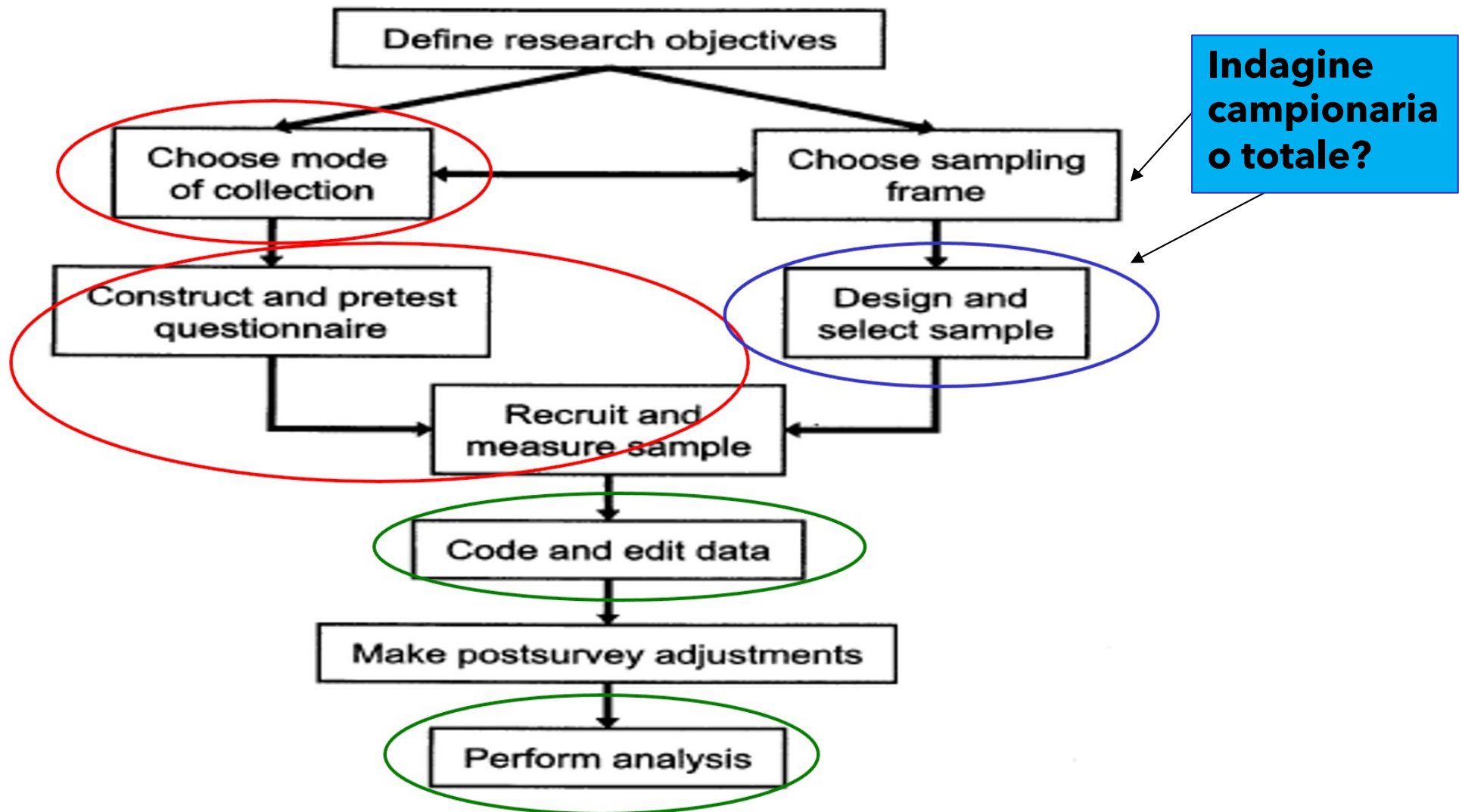
# Tecniche di indagine statistica

---

## Lezione 13



# Ciclo di vita indagine come processo (produttivo) – Fasi indagine

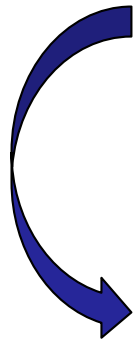


# Indagine campionaria vs indagine totale (esaustiva o censuaria)

(tenuto conto di obiettivi, ampiezza pop.ne, risorse, ...)

indagine campionaria consente:

- costi più contenuti
- maggiore tempestività
- maggiore accuratezza (se **selezione** condotta in modo **opportuno**)



indagine **totale**:

- no incertezza dovuta a campionamento ma non priva di altre *fonti di errore* (**non campionario**), difficili da tenere sotto controllo con N elevato

indagine **campionaria**:

- incertezza dovuta a campionamento (errore campionario) ma che può essere definita (calcolata) e altre *fonti di errore* (**non campionario**) che può essere più agevole tenere sotto controllo poiché  $n \ll N$

# Distorsione da selezione e indagine campionaria

*Distorsione da selezione (**selection bias**) emerge se:*

*La popolazione **obiettivo** (target) è diversa dalla popolazione rappresentata dal **campione** (survey)*

Come può verificarsi?

# Distorsione da selezione – possibili situazioni

Si può indurre *selection bias* se:

- Si selezionano le unità più facili da raggiungere
- Si selezionano le unità che paiono “rappresentative” a giudizio dell’intervistatore/ricce o del ricercatore/ricce
- Si utilizzano nel campione unità che si selezionano sulla base della volontarietà
- Si specifica male la popolazione obiettivo
- Si usano cattivi frame (es., sottocopertura)
- Si sostituiscono le unità del campione (per assicurarsi la numerosità del campione desiderata)
- Si riscontrano mancate risposte (per irreperibilità o rifiuto)

***l’uso (possibilmente rigoroso) di tecniche di campionamento probabilistico permette di evitare alcune fonti di selection bias***

# Indagine campionaria – disegno di campionamento per selezione unità della popolazione

## disegno di campionamento:

- insieme delle **regole** (e decisioni) seguite nella formazione del campione
- **obiettivo**: selezionare un sottoinsieme  $C$  di unità della pop.ne, esaminare le unità statistiche del campione  $C$  per studiare una (o più) variabili  $Y$  nella popolazione
- **probabilistico** (selezione *casuale* delle unità):
  1. ogni elemento della popolazione ha una probabilità **nota** (non nulla) di essere estratto (che deriva dall'assegnazione di una distribuzione di probabilità su tutti i possibili campioni di ampiezza  $n$ , selezionabili dalla pop.ne)
  2. sono usate tecniche (in modo appropriato) per la selezione casuale delle unità
- **non probabilistico** selezione
  - unità più accessibili, “facili”
  - con scelta soggettiva/ragionata
  - combinando “accessibilità selezione “per quote” (combinazione dei due)

# Campionamento probabilistico (disegno) e errore

**L'uso (possibilmente rigoroso) di tecniche di campionamento probabilistico permette di evitare alcune fonti di selection bias**

In generale, l'incertezza dovuta al campionamento (*sampling error*) è funzione di:

- 1) ciascun elemento della pop.ne ha una **probabilità nota, non nulla** di far parte del campione  
(**decisione su:** selezione **casuale** delle unità e probabilità di essere selezionata/**inclusa** nel campione)
- 2) il campione è selezionato in modo tale da tener conto dell'eventuale esistenza **di sottopopolazioni**  
(**decisione su:** **struttura del campione** collegata a caratteristiche della lista - sampling frame- delle unità che compongono la pop.ne)
- 3) le unità sono selezionate **individualmente** o in **gruppo**?  
(**decisione su:** **struttura del campione** collegata a caratteristiche della lista - sampling frame- delle unità che compongono la pop.ne)
- 4) quanto **grande** è il campione?  
(**decisione su:** precisione delle stime)

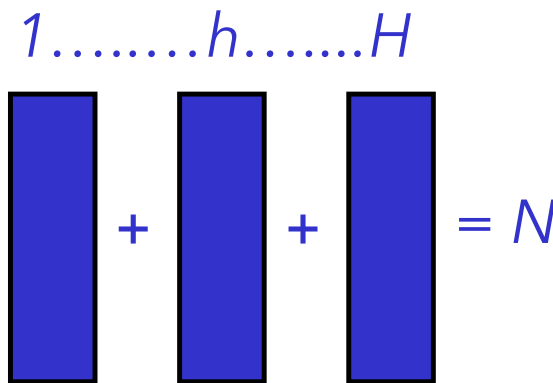
# Struttura del campione

SEMPLICE: lista di etichette che corrispondono uno a uno alle unità della popolazione (campione casuale semplice)

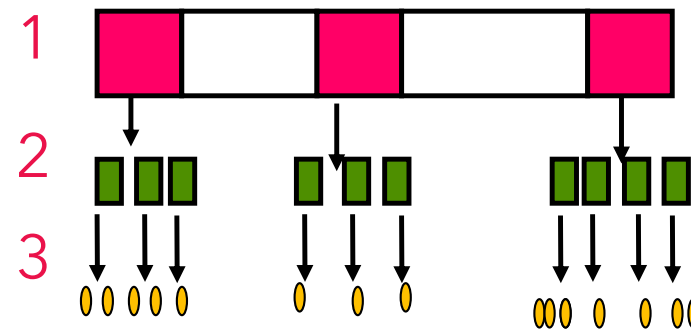
COMPLESSO: varie liste

- distinte per sottopopolazioni (campionamento stratificato)
- gerarchiche (campionamento su più stadi)
- dinamiche (campionamento ripetuto nel tempo; campionamento ruotato)

## STRATI



## STADI



## DINAMICHE

