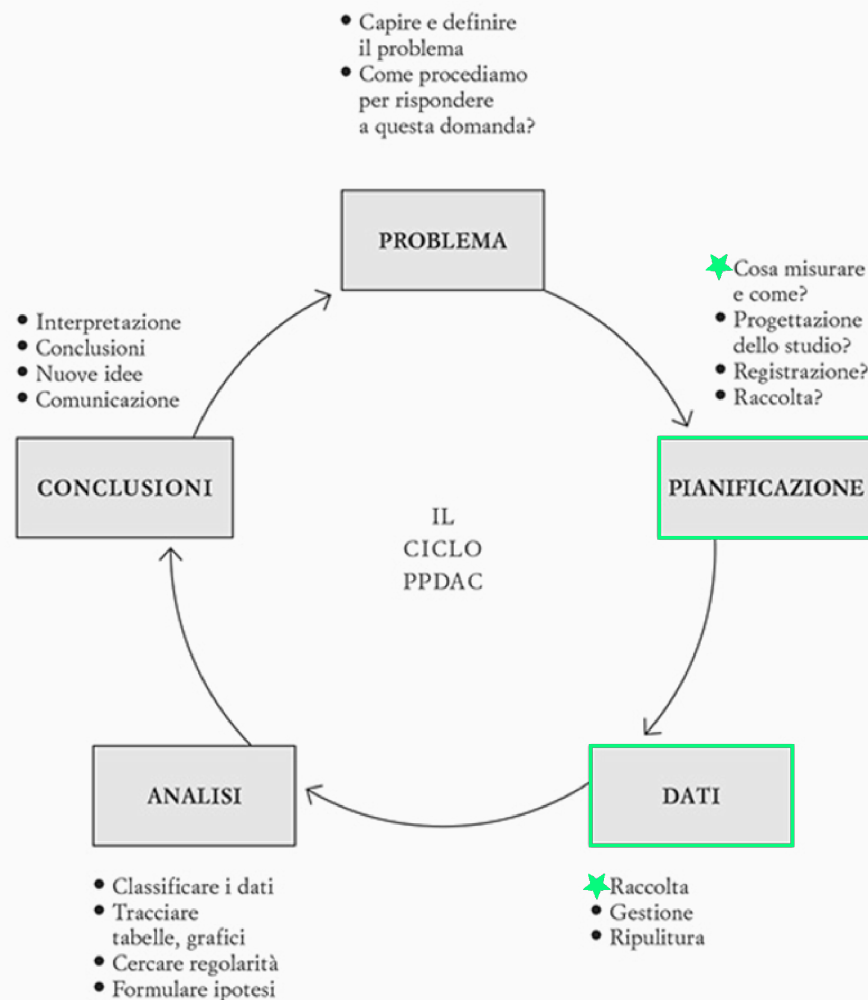


Le variabili

Obiettivi di apprendimento

- Definire cos'è una variabile
- Definire quali sono le cause della variabilità di un'osservazione
- Conoscere le differenze tra i diversi tipi di variabile

Le fasi della ricerca



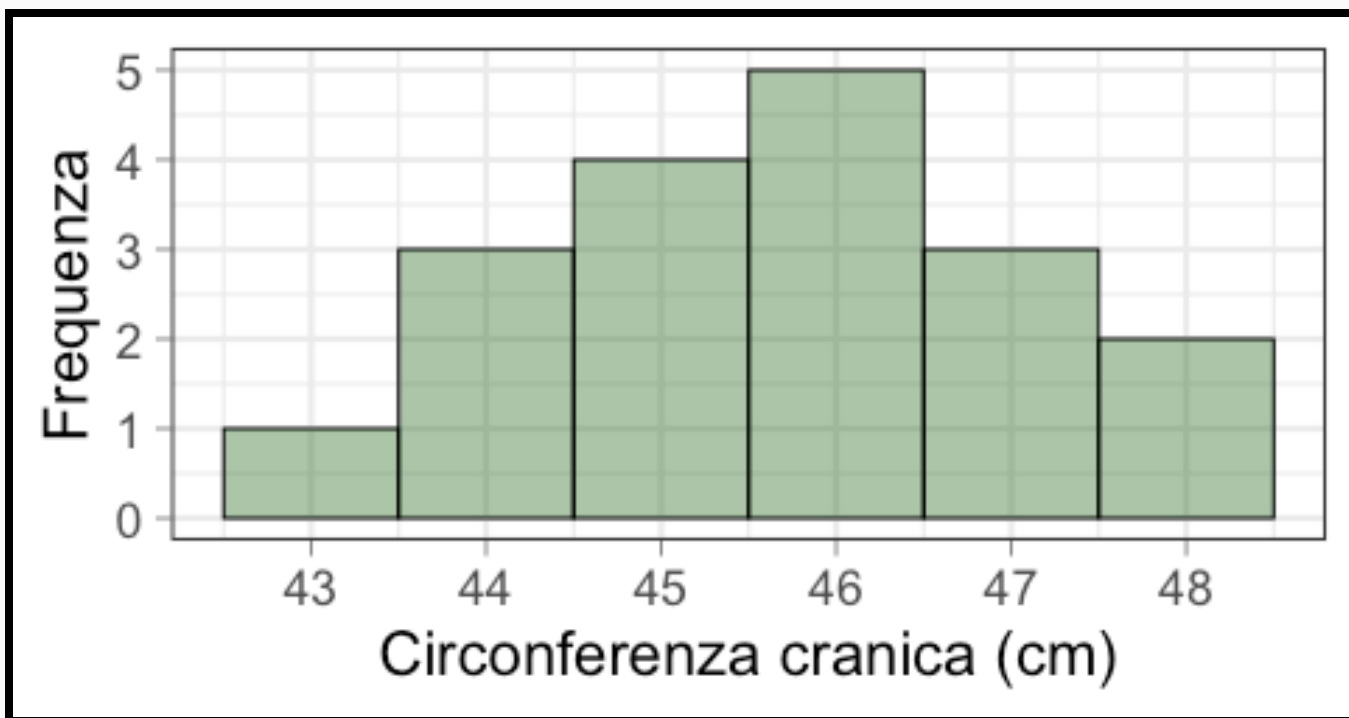
variabile \longleftrightarrow **informazione**

Le variabili...

- sono imperfette: non sempre misurano quello che ci interessa
 - conoscenza di una materia \neq voto all'esame
 - benessere di un Paese \neq prodotto interno lordo
- variano: da una persona all'altra, da un momento all'altro, ...
e lo fanno in modo apparentemente casuale

Da cosa dipende questa variabilità?

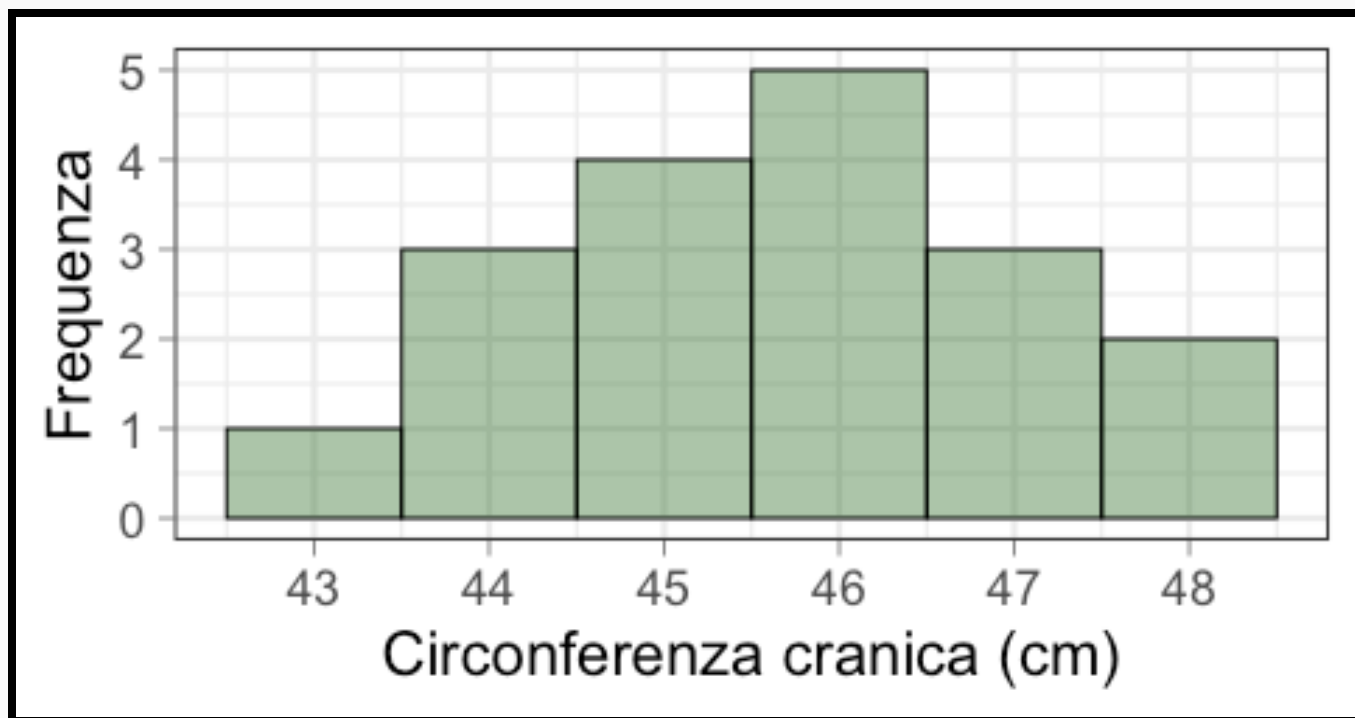
- 🎯 In uno studio sull'età neonatale, 18 pediatri misurano la circonferenza cranica di un bambolotto che simula un bambino di un anno



Perché i dati variano?

Da cosa dipende questa variabilità?

- 🎯 In uno studio sull'età neonatale, 18 pediatri misurano la circonferenza cranica di un bambolotto che simula un bambino di un anno



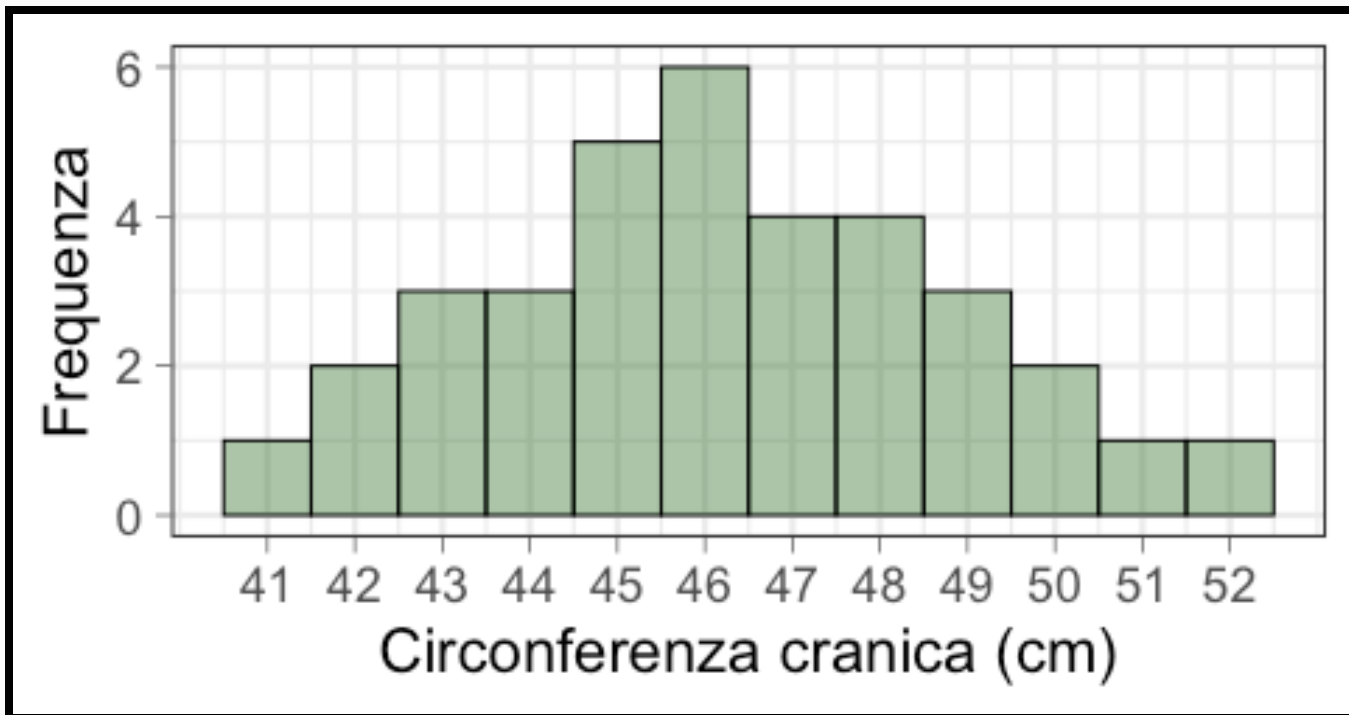
Perché i dati variano?

Errori di misurazione

Da cosa dipende questa variabilità?



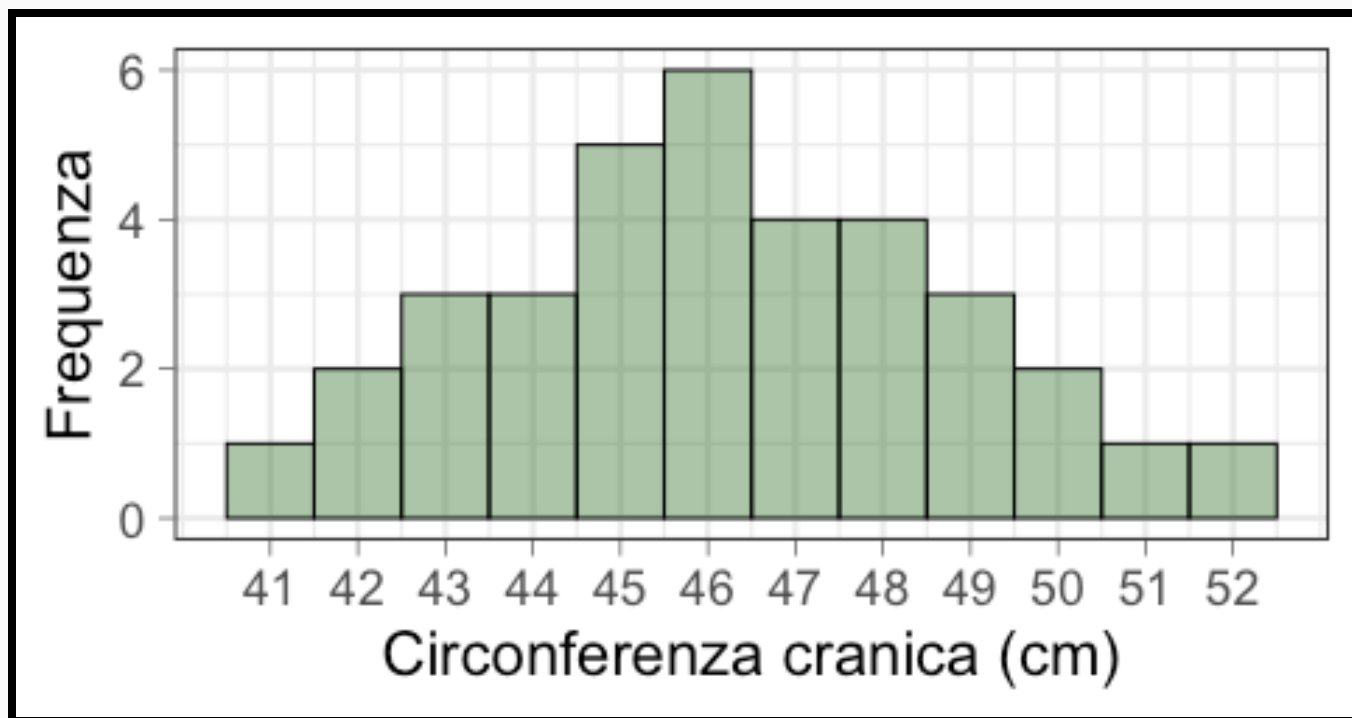
Nello stesso studio, un pediatra misura la circonferenza cranica di 35 bambini di un anno



Perché i dati variano?

Da cosa dipende questa variabilità?

- 🎯 Nello stesso studio, un pediatra misura la circonferenza cranica di 35 bambini di un anno

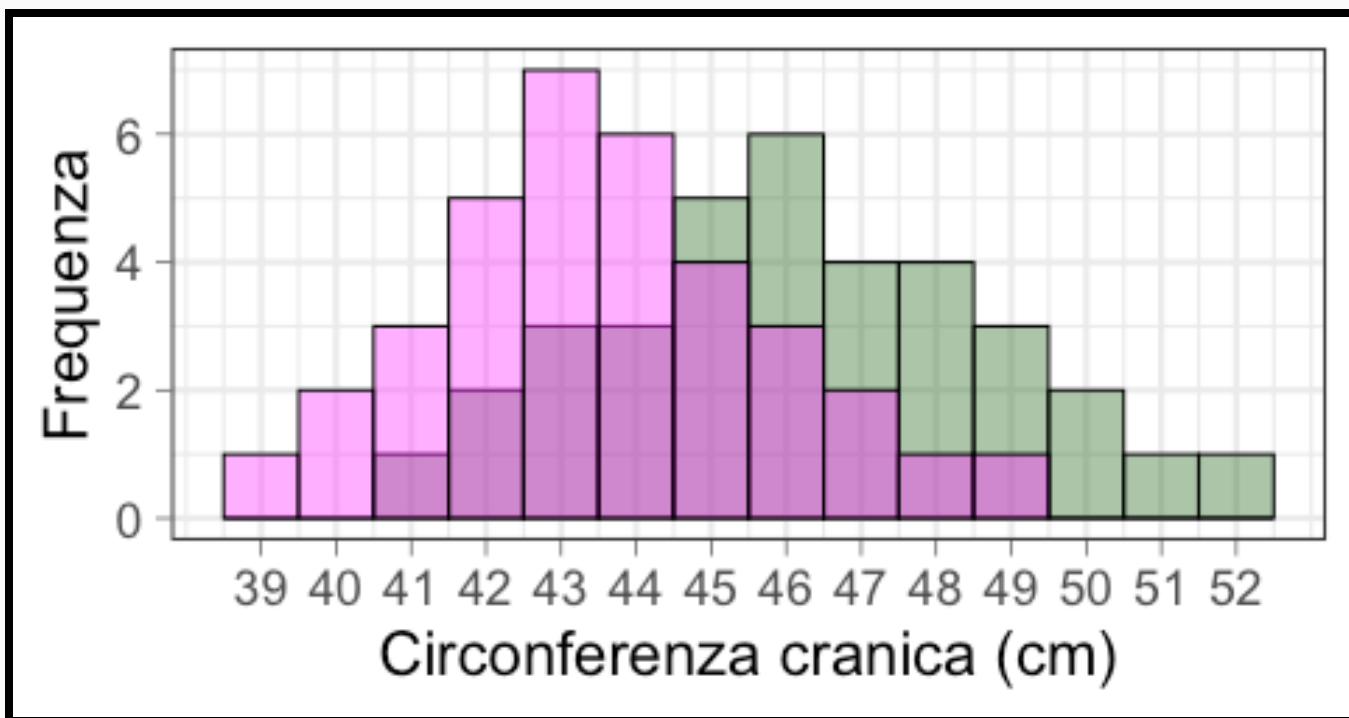


Perché i dati variano?

Differenze individuali

Perché c'è variabilità?

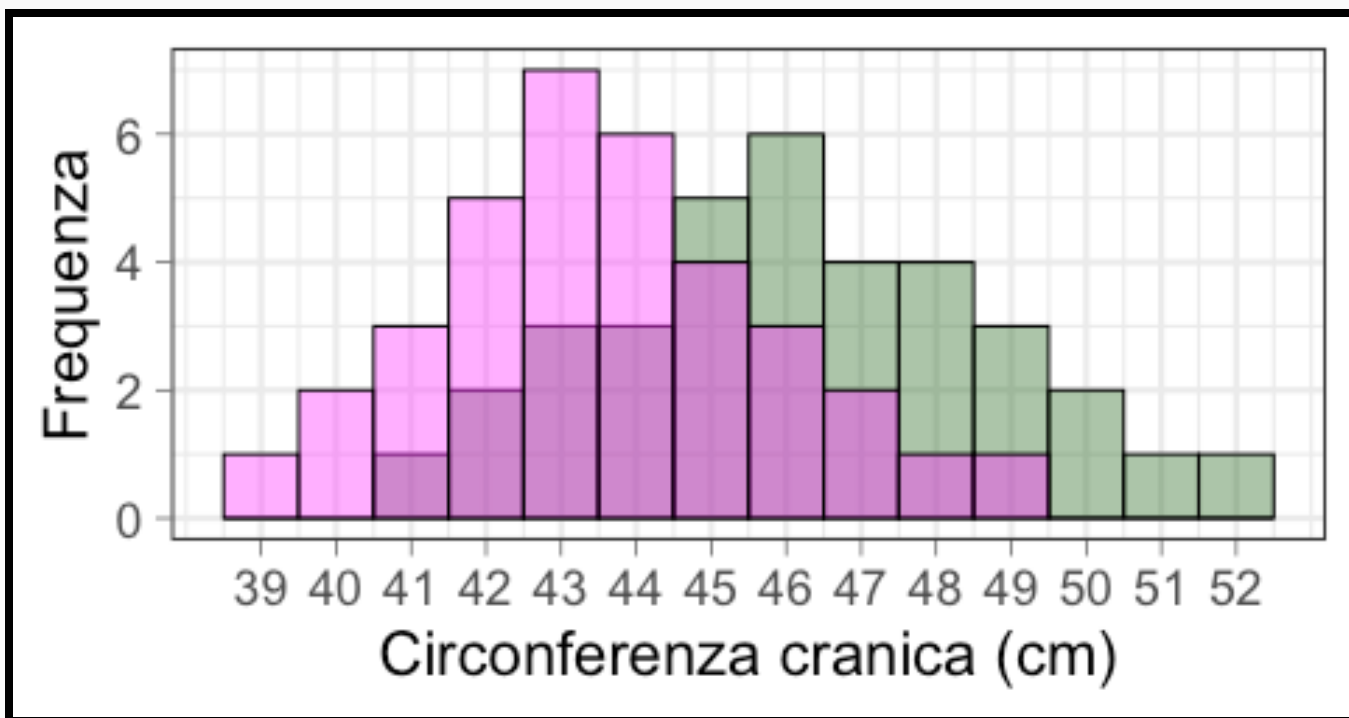
- 🎯 Nello stesso studio, lo stesso pediatra misura altri 35 bambini di un anno che sono però esposti al fumo passivo



Perché i dati variano?

Da cosa dipende questa variabilità?

- 🎯 Nello stesso studio, lo stesso pediatra misura altri 35 bambini di un anno che sono però esposti al fumo passivo

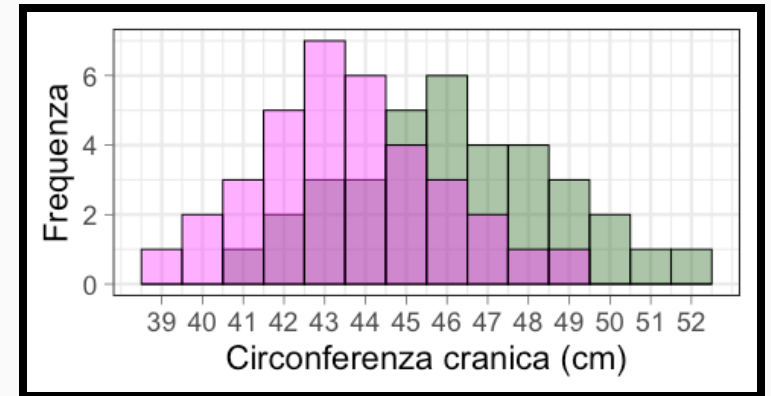
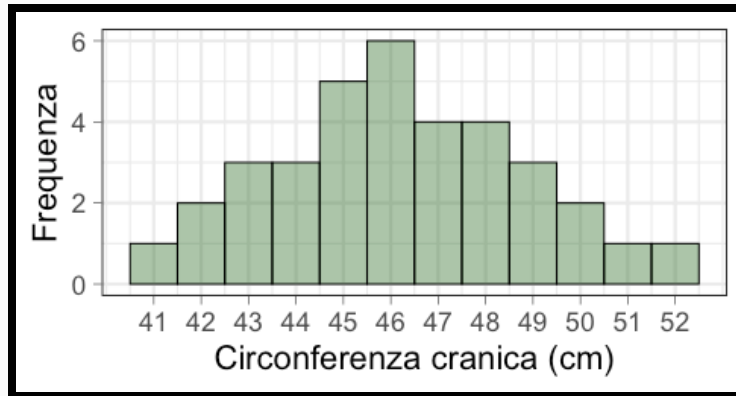
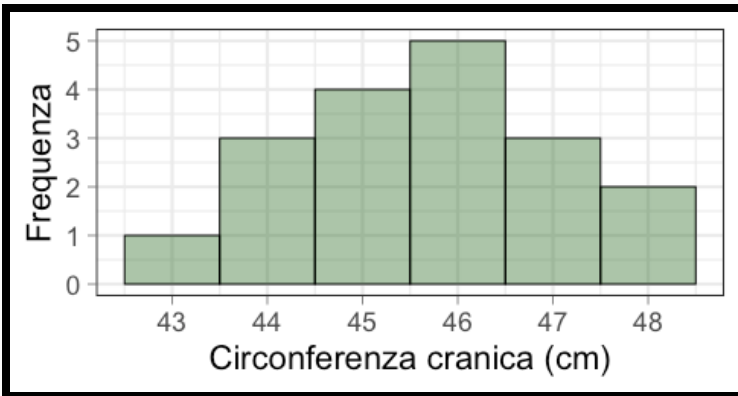


Perché i dati variano?

Diversa condizione
sperimentale?

Perché dobbiamo conoscere la differenza?

- 🎯 Vedremo come determinare se la variabilità dipende dalle condizioni sperimentali o se dipende da differenze individuali e/o da errori di misurazione

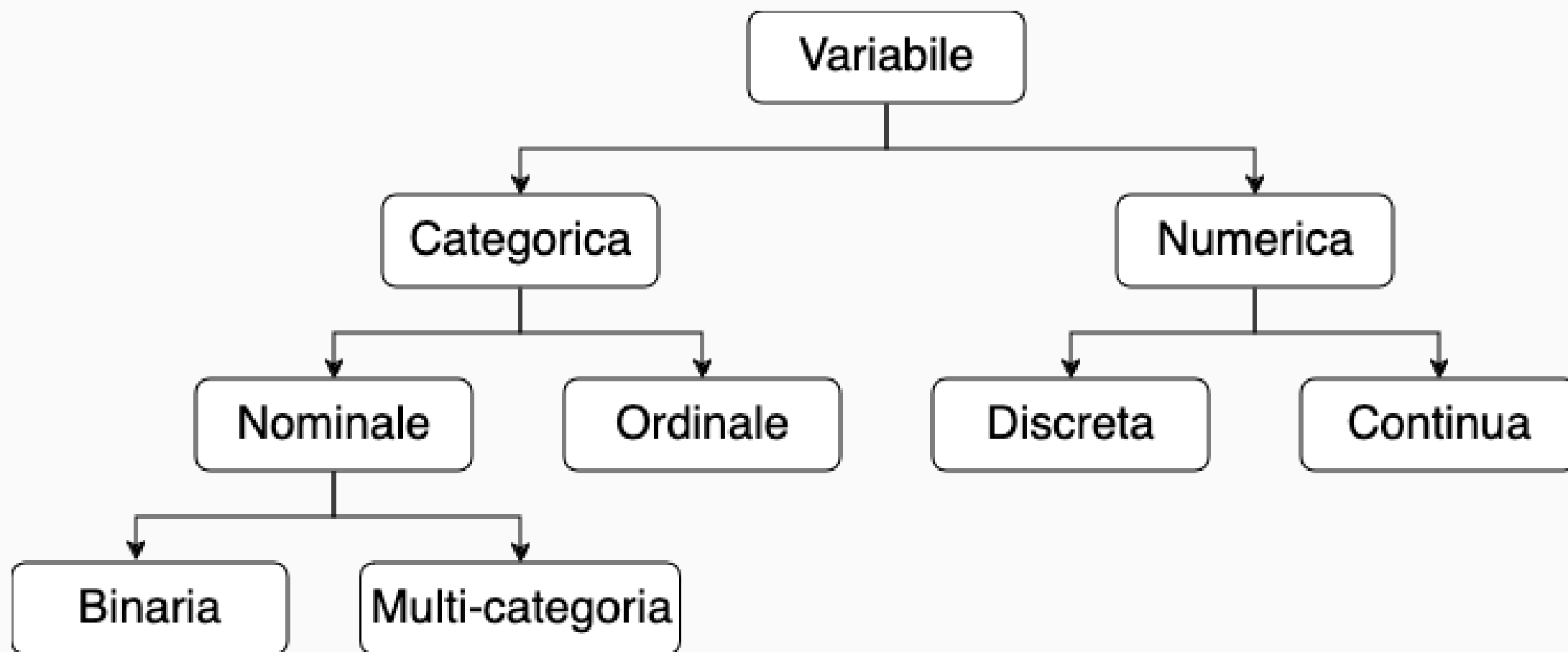


Variabili di esito e di esposizione*

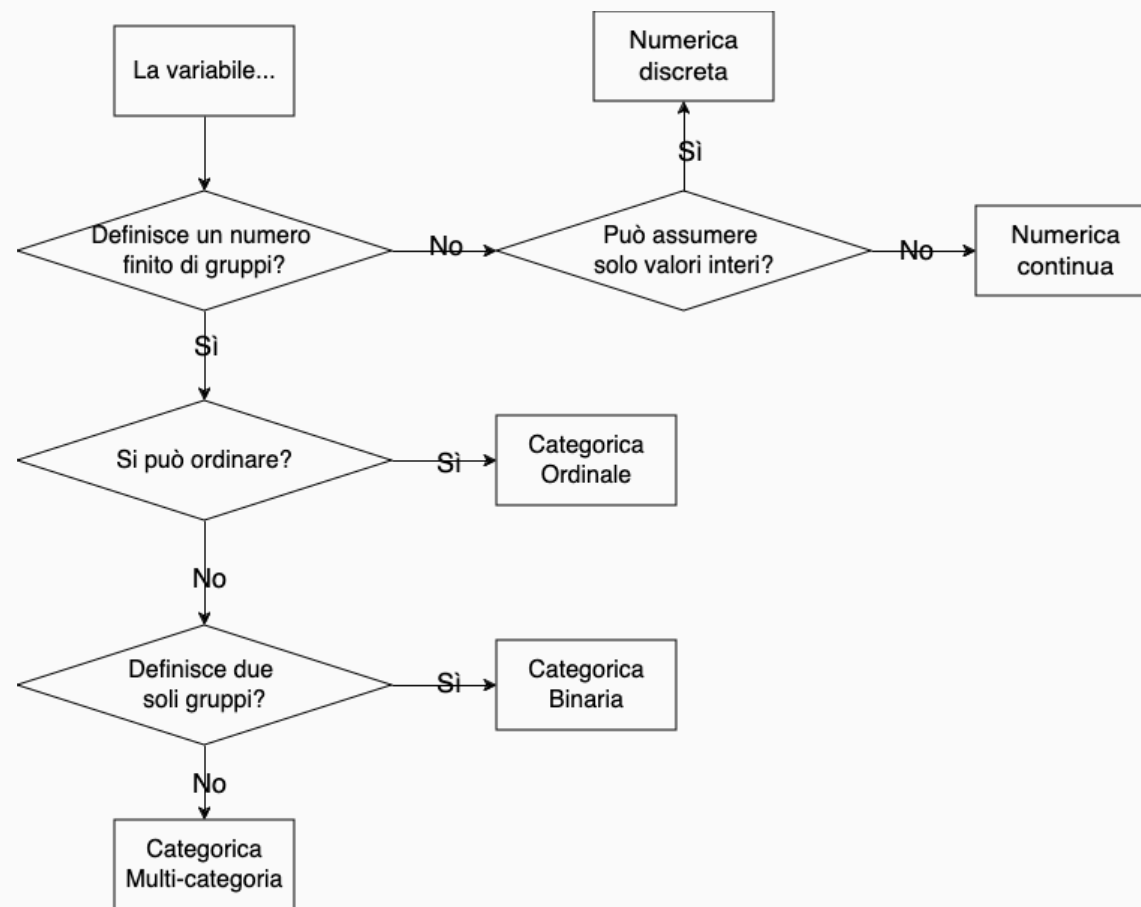
- la variabile di cui studiamo la variazione \longrightarrow variabile di esito
 - Esempio: la circonferenza cranica
- le cause della variabilità \longrightarrow variabili di esposizione
 - Esempio: il fumo passivo

* o dipendenti e indipendenti, o di risposta ed esplicative o (in uno studio caso-controllo), stato e fattore di rischio. In uno studio clinico randomizzato, l'esposizione è il gruppo di trattamento.

I tipi di variabile

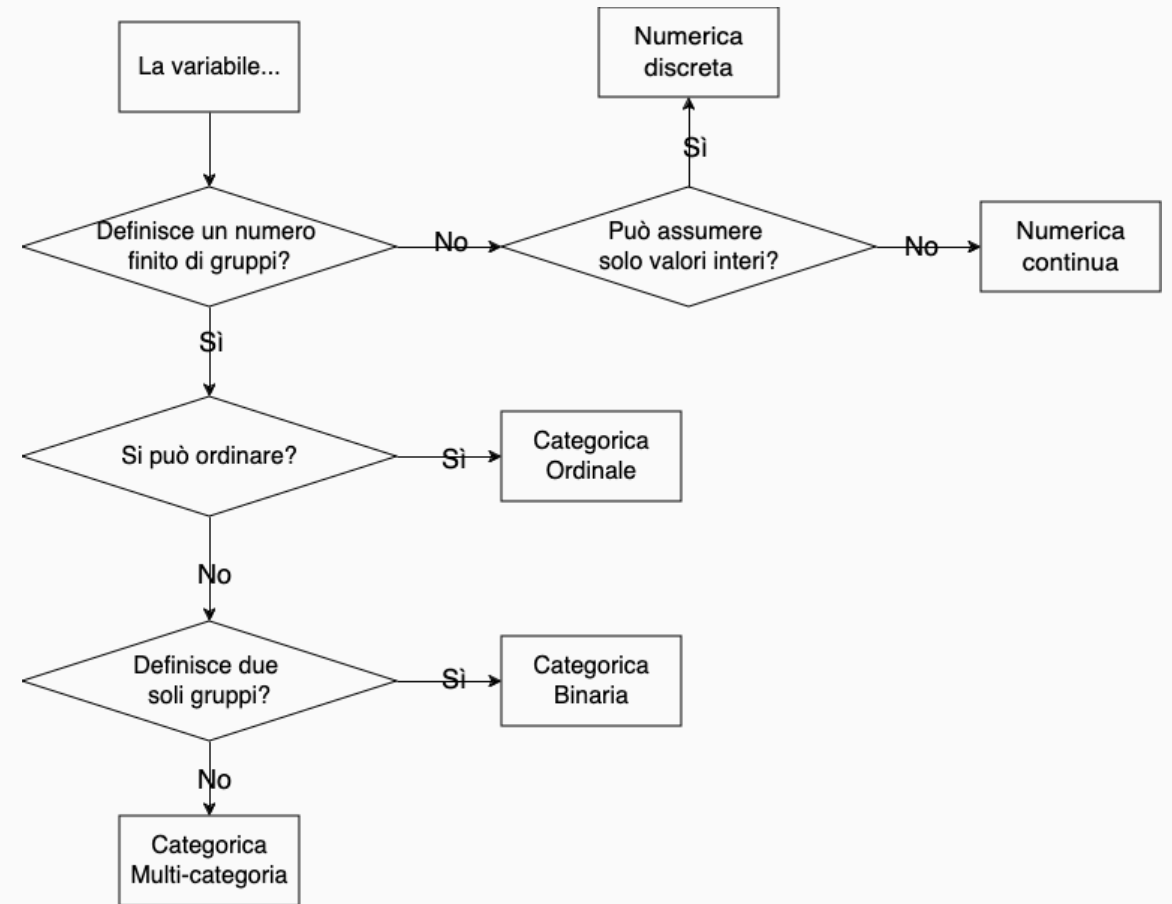


Di che tipo di variabile si tratta?



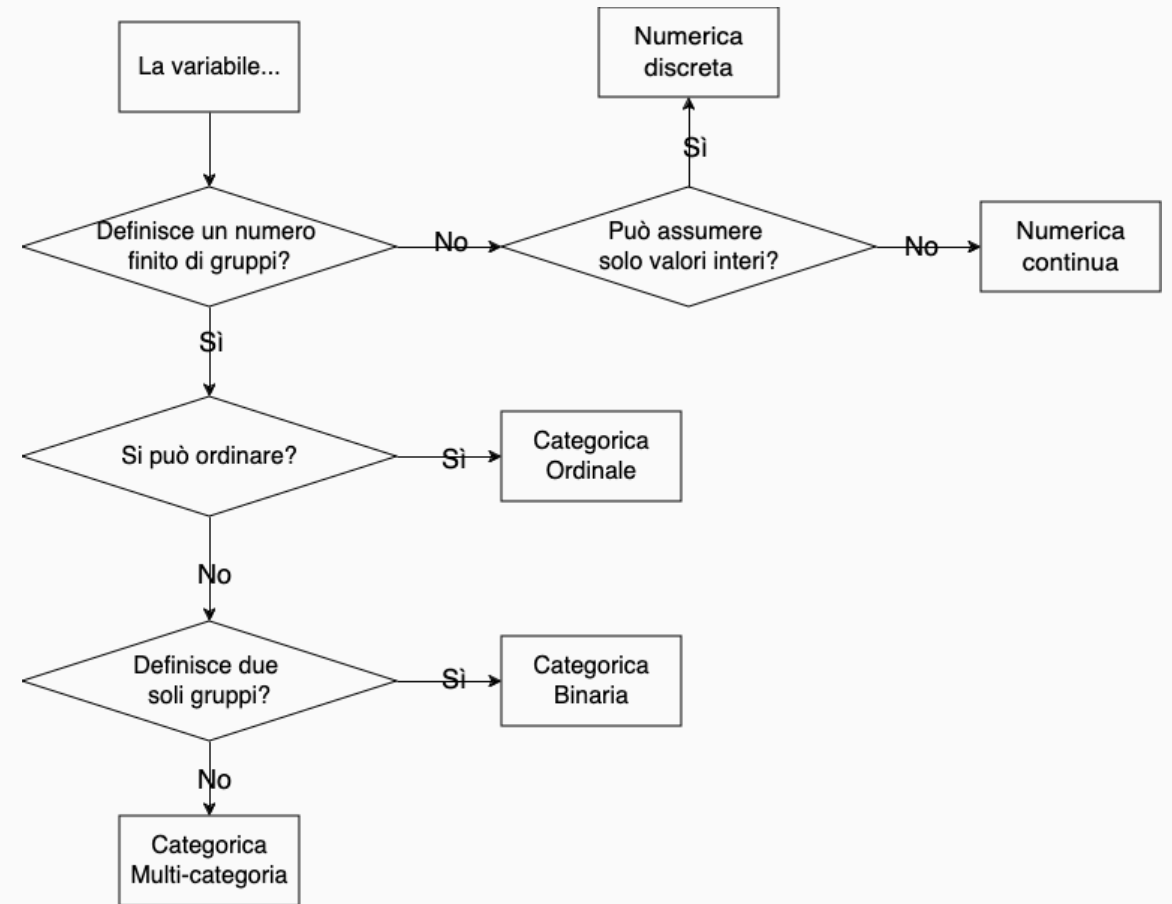
Di che tipo di variabile si tratta?

? La taglia di una maglietta



Di che tipo di variabile si tratta?

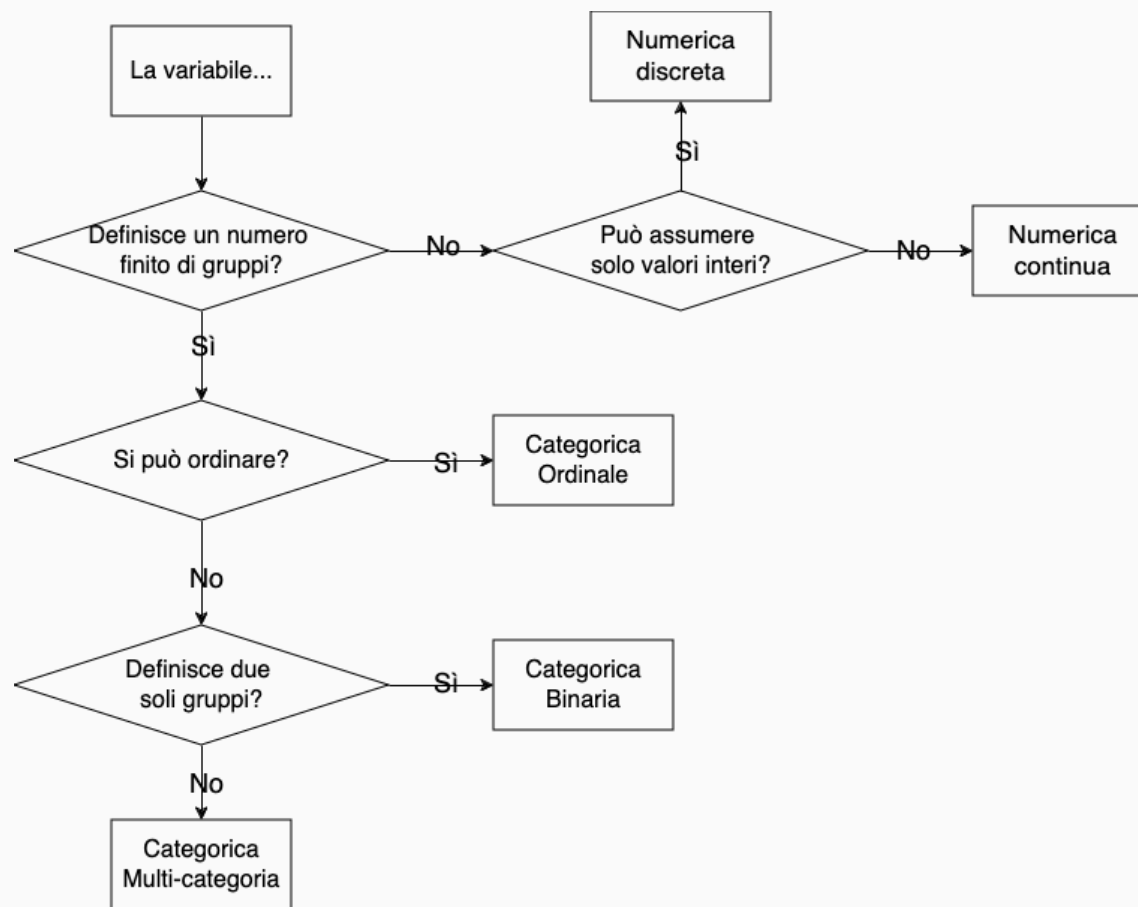
? Il numero di studenti in un corso



Di che tipo di variabile si tratta?

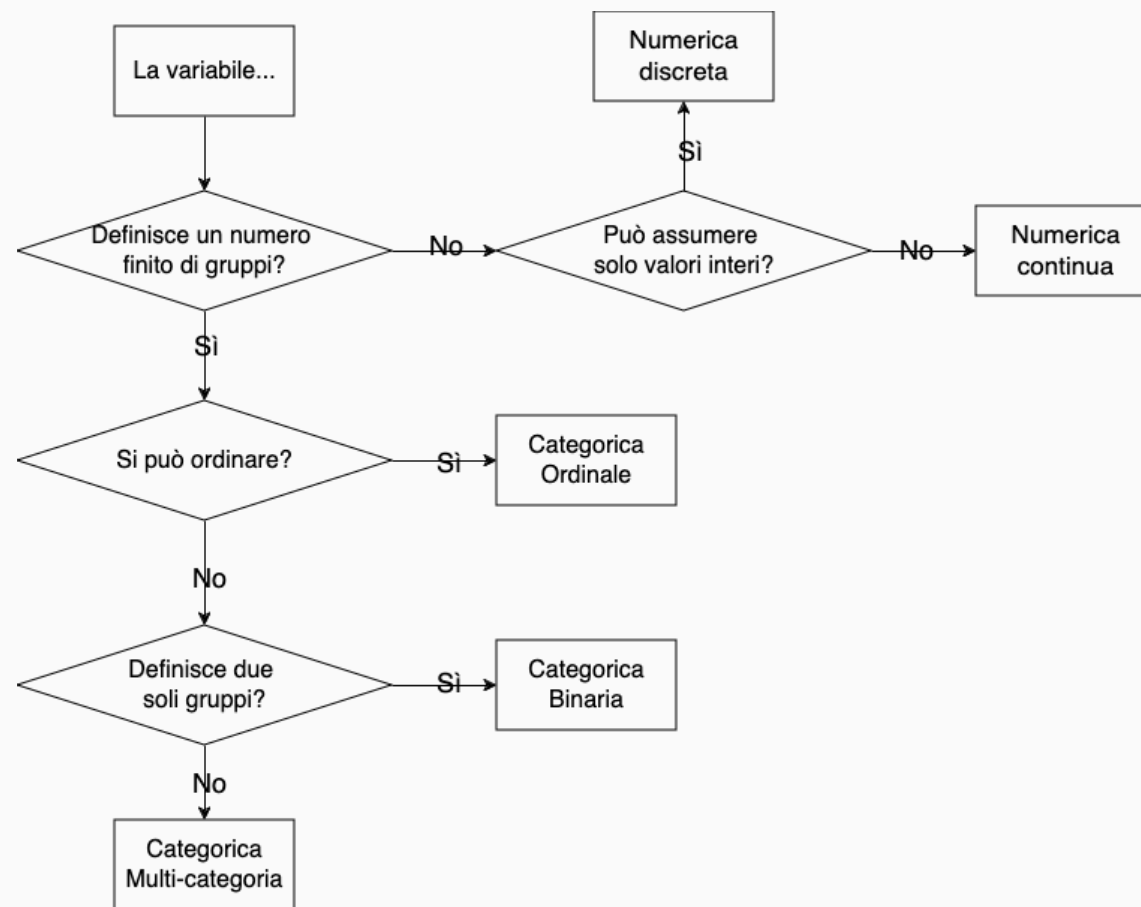


La nazionalità



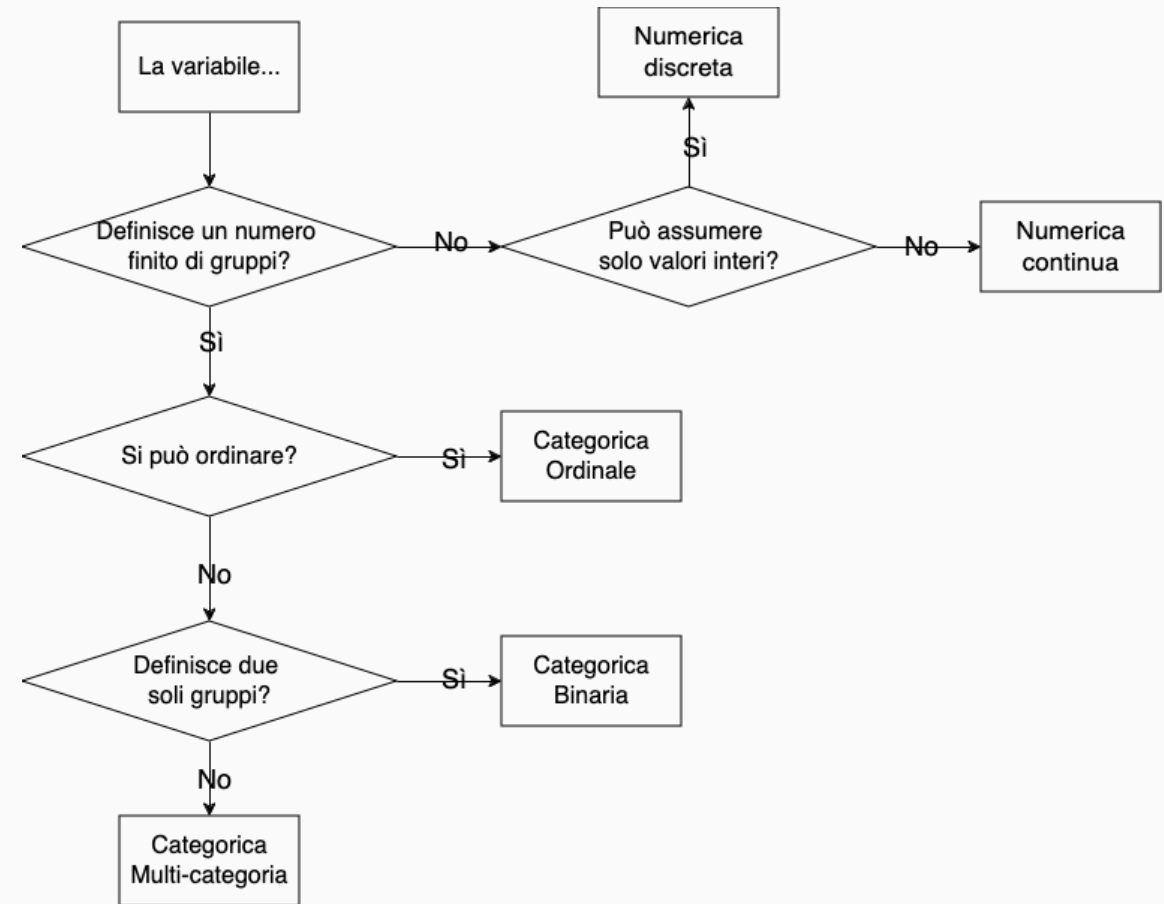
Di che tipo di variabile si tratta?

? La glicemia a digiuno



Di che tipo di variabile si tratta?

? Superare l'esame per la patente



Esercizio #1

? Di che tipo sono le variabili in questa tabella?

01:00

Visconti A., et al., *Total serum N-glycans associate with response to immune checkpoint inhibition therapy and survival in patients with advanced melanoma*, BMC Cancer, 2023 doi:10.1186/s12885-023-10511-3

Table 1 Patient characteristics.

All cohorts
N (pre-treatment)
N (follow-up)
Sex
Male
Female
Age (years)
BMI (kg/m²)
BRAF mutant
LDH (\leqULN)
Metastatic stage
Stage III unresectable
M1a
M1b
M1c
M1d
ECOG performance status
0
1
2
3
ICI therapy
Ipilimumab
Pembrolizumab
Nivolumab
Ipilimumab + Nivolumab

Esercizio #1 -- Soluzione

? Di che tipo sono le variabili in questa tabella?

Visconti A., et al., *Total serum N-glycans associate with response to immune checkpoint inhibition therapy and survival in patients with advanced melanoma*, BMC Cancer, 2023 doi:10.1186/s12885-023-10511-3

Table 1 Patient characteristics.

All cohorts	
N (pre-treatment)	
N (follow-up)	
Sex	Binario
Male	
Female	
Age (years)	Numerico
BMI (kg/m ²)	Continuo
BRAF mutant	Binario
LDH (\leq ULN)	Binario
Metastatic stage	Ordinale
Stage III unresectable	
M1a	
M1b	
M1c	
M1d	
ECOG performance status	Ordinale
0	
1	
2	
3	
ICI therapy	Multi-categoria
Ipilimumab	
Pembrolizumab	
Nivolumab	
Ipilimumab + Nivolumab	

Esercizio #2

? Avere una diagnosi di autismo è una variabile categorica dicotomica?

a) Vero b) Falso

? Il numero di iscritti a un evento è una variabile quantitativa discreta?

a) Vero b) Falso

? Il livello di colesterolo è una variabile qualitativa ordinale?

a) Vero b) Falso

Esercizio #2 -- Soluzione

? Avere una diagnosi di autismo è una variabile categorica dicotomica?

a) Vero ☒ b) Falso

? Il numero di iscritti a un evento è una variabile quantitativa discreta?

a) Vero b) Falso

? Il livello di colesterolo è una variabile qualitativa ordinale?

a) Vero b) Falso

Esercizio #2 -- Soluzione

? Avere una diagnosi di autismo è una variabile categorica dicotomica?

a) Vero ☒ b) Falso

? Il numero di iscritti a un evento è una variabile quantitativa discreta?

a) Vero ☒ b) Falso

? Il livello di colesterolo è una variabile qualitativa ordinale?

a) Vero b) Falso

Esercizio #2 -- Soluzione

? Avere una diagnosi di autismo è una variabile categorica dicotomica?

a) Vero ☒ b) Falso

? Il numero di iscritti a un evento è una variabile quantitativa discreta?

a) Vero ☒ b) Falso

? Il livello di colesterolo è una variabile qualitativa ordinale?

a) Vero b) Falso ☒

Esercizio #3

? Un ricercatore si chiede se il numero di figli dipenda dal livello di educazione. Per rispondere a questa domanda divide le persone in 4 categorie (licenza media e/o elementare, diploma, laurea, dottorato) e ne trascrive il numero di figli.

Che tipo di variabili raccoglie?

Educazione

- a) multi-categoria
- b) ordinale
- c) discreta
- d) continua

Numero di figli


- a) multi-categoria
- b) ordinale
- c) discreta
- d) continua

Esercizio #3 -- Soluzione

? Un ricercatore si chiede se il numero di figli dipenda dal livello di educazione. Per rispondere a questa domanda divide le persone in 4 categorie (licenza media e/o elementare, diploma, laurea, dottorato) e ne trascrive il numero di figli.

Che tipo di variabili raccoglie?

Educazione

- a) multi-categoria
- b) ordinale 
- c) discreta
- d) continua

Numero di figli

- a) multi-categoria
- b) ordinale
- c) discreta
- d) continua

Esercizio #3 -- Soluzione

- ? Un ricercatore si chiede se il numero di figli dipenda dal livello di educazione. Per rispondere a questa domanda divide le persone in 4 categorie (licenza media e/o elementare, diploma, laurea, dottorato) e ne trascrive il numero di figli.

Che tipo di variabili raccoglie?

Educazione

- a) multi-categoria
- b) ordinale ☒
- c) discreta
- d) continua

Numero di figli

- a) multi-categoria
- b) ordinale
- c) discreta ☒
- d) continua

Esercizio #4

? Un ricercatore disegna uno studio in cui pazienti con alti livelli di colesterolo sono assegnati in modo casuale a prendere o un nuovo farmaco o un placebo. Sapendo che c'è una differenza nel modo in cui uomini e donne rispondono al trattamento, ne registra il sesso. I livelli di colesterolo e vengono misurati prima e dopo 3 mesi di trattamento

Che tipo di variabili raccoglie?

- Trattamento :
- Sesso :
- Livelli di colesterolo :

Esercizio #4 -- Soluzione

? Un ricercatore disegna uno studio in cui pazienti con alti livelli di colesterolo sono assegnati in modo casuale a prendere o un nuovo farmaco o un placebo. Sapendo che c'è una differenza nel modo in cui uomini e donne rispondono al trattamento, ne registra il sesso. I livelli di colesterolo e vengono misurati prima e dopo 3 mesi di trattamento

Che tipo di variabili raccoglie?

- Trattamento : binaria
- Sesso : binaria
- Livelli di colesterolo : continua

Una sola variabile, tanti tipi



Esempio: età

Una sola variabile, tanti tipi



Esempio: età

- Data di nascita + data visita

Una sola variabile, tanti tipi



Esempio: età

- Data di nascita + data visita
- Quanti anni hai?

Una sola variabile, tanti tipi



Esempio: età

- Data di nascita + data visita
- Quanti anni hai?
- A quale di queste fasce d'età appartieni?

Discretizzazione

- Da una serie di valori continui a intervalli (o classi)
- La scelta del numero di classi e degli estremi è arbitraria
- Le classi devono essere mutualmente esclusive
- Riportate anche le classi nulle (frequenza zero)


$$\text{IMC (BMI)} = \frac{\text{peso}_{\text{kg}}}{\text{altezza}_m^2}$$

IMC (BMI)	Classificazione
< 18,5	Sottopeso
18,5 – 24, 9	Normopeso
25,0 – 29,9	Sovrappeso
30,0 – 34,9	Obesità 1° grado
35,0 – 39,9	Obesità 2° grado
> 39,9	Obesità 3° grado

Discretizzazione

Table 1 Patient characteristics.	
	All cohorts
N (pre-treatment)	88
LDH (\leq ULN)	58 (65.9%)


Visconti A., et al., Total serum N-glycans associate with response to immune checkpoint inhibition therapy and survival in patients with advanced melanoma, BMC Cancer, 2023 doi:10.1186/s12885-023-10511-3

Esercizio #5

- ?
- Un ricercatore ha registrato l'età del campione come una variabile discreta. Può trasformarla in una variabile categorica ordinale?
- a) Sì, sempre
 - b) Dipende dal contesto
 - c) No, mai

Esercizio #5 -- Soluzione


? Un ricercatore ha registrato l'età del campione come una variabile discreta. Può trasformarla in una variabile categorica ordinale?

- a) Sì, sempre 
- b) Dipende dal contesto
- c) No, mai

Esercizio #6

- ?
- Un ricercatore ha registrato l'età del campione come una variabile categorica ordinale. Può riottenere l'esatta informazione numerica?
- a) Sì, sempre
 - b) Dipende dal contesto
 - c) No, mai

Esercizio #6 -- Soluzione

- ? Un ricercatore ha registrato l'età del campione come una variabile categorica ordinale. Può riottenere l'esatta informazione numerica?
- a) Sì, sempre
 - b) Dipende dal contesto
 - c) No, mai 

Cosa abbiamo imparato?

- Le variabili descrivono delle caratteristiche di interesse
- Le differenze che esistono tra le osservazioni (variabilità) dipendono sia da fattori noti o conoscibili (come le differenze individuali e le condizioni sperimentali), sia da effetti casuali (rumore statistico, come gli errori di misurazione)
- Le variabili possono essere di diverso tipo, a seconda della caratteristica che descrivono
- Le variabili numeriche possono essere discretizzate (quindi trasformate in variabili categoriche)