

ELEMENTI DI STATISTICA

SFP- MATERA

La statistica è lo studio dei fenomeni collettivi ossia di quei fenomeni che riguardano una pluralità di oggetti.

Le fasi di un'indagine statistica sono:

- Rilevamento dei dati,
- Elaborazione e interpretazione dei dati.

Un'indagine statistica consiste nella raccolta e nell'analisi dei dati concernenti uno o più fenomeni che riguardano un insieme di soggetti: ognuno di questi soggetti costituisce un'**unità statistica** e il loro insieme si chiama **popolazione** o **universo**.

In un'indagine statistica si considerano, per ogni unità statistica, uno o più aspetti detti **caratteri**. Ciascun carattere si può presentare in diverse **modalità**.

Si distinguono:

caratteri quantitativi, le cui modalità sono numeri o grandezze;

caratteri qualitativi, le cui modalità sono attributi diversi da numeri o grandezze.

Raccogliere i dati di un'intera popolazione può essere difficile o addirittura impossibile. In questi casi l'indagine viene svolta su un sottoinsieme dell'universo statistico, detto **campione**. Le conclusioni dell'indagine possono però non essere vere per l'intera popolazione.

La **statistica inferenziale** studia le modalità con cui è possibile estendere all'intera popolazione le conclusioni di un'indagine svolta su di un campione e permette di valutare il grado di attendibilità di tali conclusioni.

Indagine statistica sul numero di reti segnate in ogni partita del campionato di calcio di serie A dello scorso anno.

Unità statistica: ogni singolo incontro di calcio

Popolazione: l'insieme degli incontri di calcio (380 in tutto)

Carattere: numero di reti segnate
Carattere quantitativo con

modalità: 0, 1, 2, 3, 4, 5, 6, 7, 8.

I dati si riassumono in una tabella di questo tipo: nella prima colonna sono riportate le modalità del carattere, nella seconda è riportato il numero di unità statistiche (ossia di incontri di calcio) che presentano quella modalità.

reti segnate	n.incontri
0	17
1	85
2	104
3	77
4	48
5	28
6	16
7	4
8	1
	380

Chiamiamo **frequenza assoluta** di una modalità il numero di unità statistiche della popolazione per cui il carattere si presenta con quella modalità.

La **frequenza relativa** di una modalità è il rapporto fra la sua frequenza assoluta e il numero di unità statistiche della popolazione . Le frequenze relative spesso si indicano come percentuali.

Esempio: la modalità zero ha frequenza relativa uguale a

$$\frac{17}{380} \cong 0,045 = 4,5\%$$

reti	0	1	2	3	4	5	6	7	8	
FA	17	85	104	77	48	28	16	4	1	380
FR%	4,5%	22,4%	27,4%	20,3%	12,6%	7,4%	4,2%	1,1%	0,3%	100%

Analizziamo i dati relativi al numero dei residenti delle 107 provincie italiane (facendo riferimento ai dati raccolti nel censimento dell'anno X)

Unità statistica: ogni *provincia*

Popolazione: insieme delle provincie

Carattere: numero dei residenti in ciascuna provincia

In questo caso, dove il carattere assume 107 modalità diverse, si usa raggruppare le modalità del carattere in classi di frequenza, ognuna delle quali è un intervallo di valori della modalità, riportando nella tabella le frequenze di ogni classe.

Convenzionalmente l'estremo inferiore viene considerato escluso dalla classe, mentre l'estremo superiore è incluso.

Si chiama **ampiezza** di una classe la differenza tra l'estremo superiore e l'estremo inferiore, mentre il **valore centrale** di una classe è la media aritmetica tra l'estremo superiore e l'estremo inferiore.

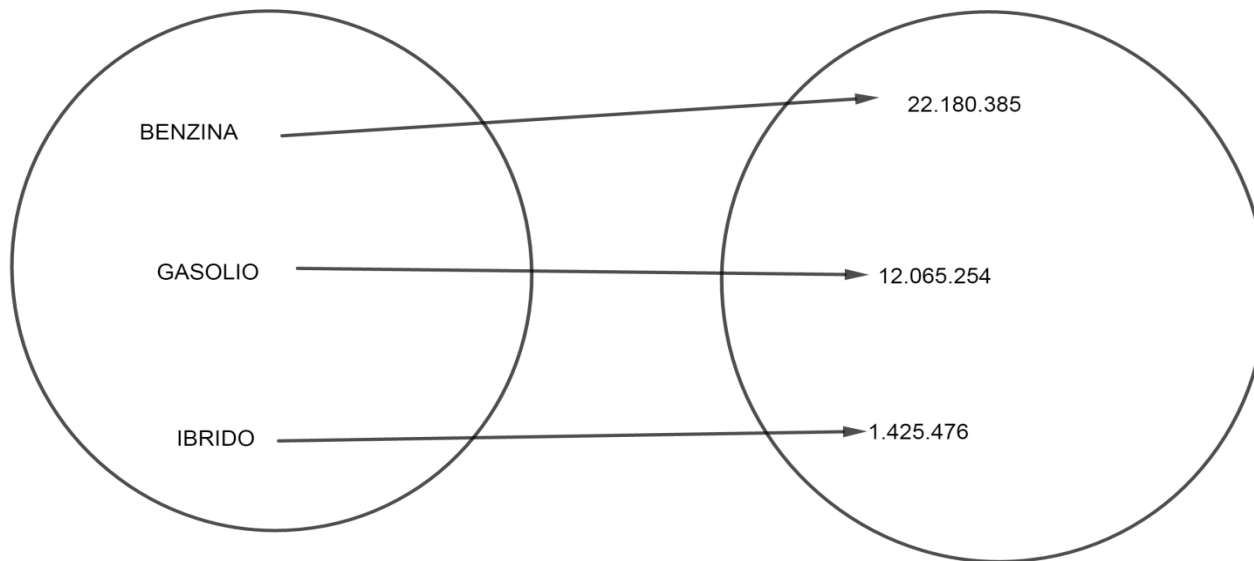
n. residenti	n. provincie
0-100.000	2
100.000-200.000	17
200.000-300.000	24
300.000-400.000	20
400.000-500.00	12
500.000-750.000	11
750.00-1.000.000	12
1.000.000-2.000.000	5
2.000.000-3.000.000	1
3.000.000-4.000.000	3
	107

Talvolta in un'indagine statistica è necessario considerare contemporaneamente due caratteri distinti. Per rappresentare contemporaneamente le frequenze dei due caratteri si usa una tabella a doppia entrata.

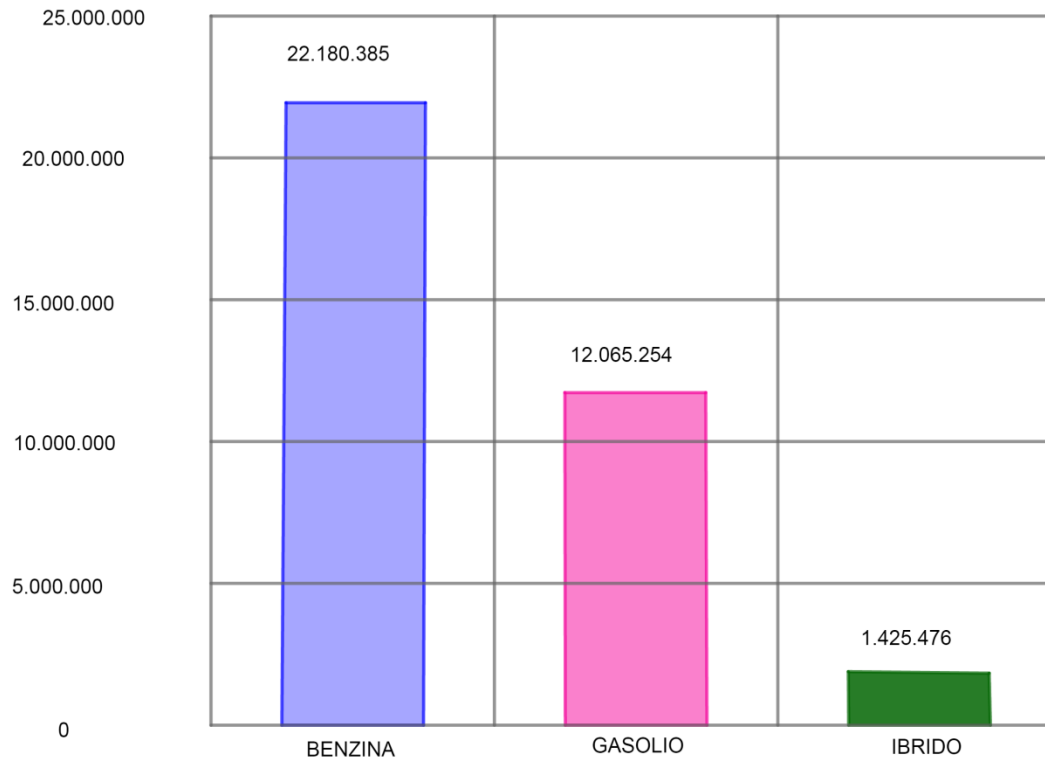
In una tale tabella la prima riga contiene le modalità di uno dei caratteri e la prima colonna le modalità dell'altro carattere.

In ogni cella è riportata la frequenza con cui i due caratteri si presentano con le modalità corrispondenti alla prima riga e alla prima colonna. Tali frequenze sono dette **frequenze congiunte**.

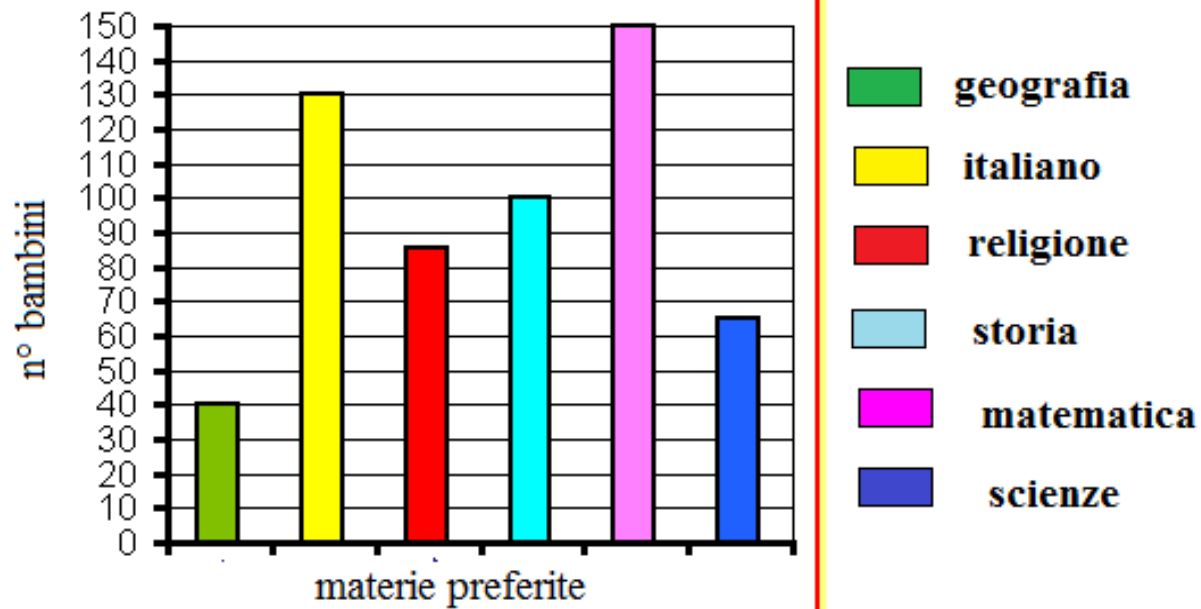
	IN TRASFERTA						
IN CASA	0	1	2	3	4	5	
0	17	26	22	9	1	0	75
1	50	50	16	8	6	1	140
2	32	31	22	5	0	0	90
3	21	14	10	6	0	0	51
4	3	6	5	2	0	0	16
5	1	4	2	0	0	0	7
6	0	0	1	0	0	0	1
	133	131	78	30	7	1	380



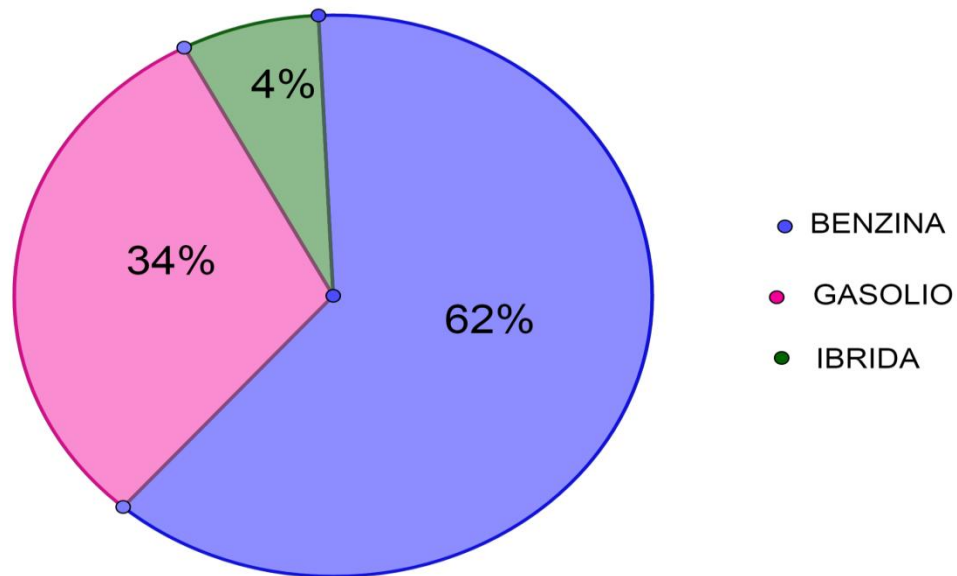
Gli **istogrammi** sono rappresentazioni grafiche che utilizzano rettangoli aventi le basi uguali e le altezze proporzionali alle frequenze delle diverse modalità.



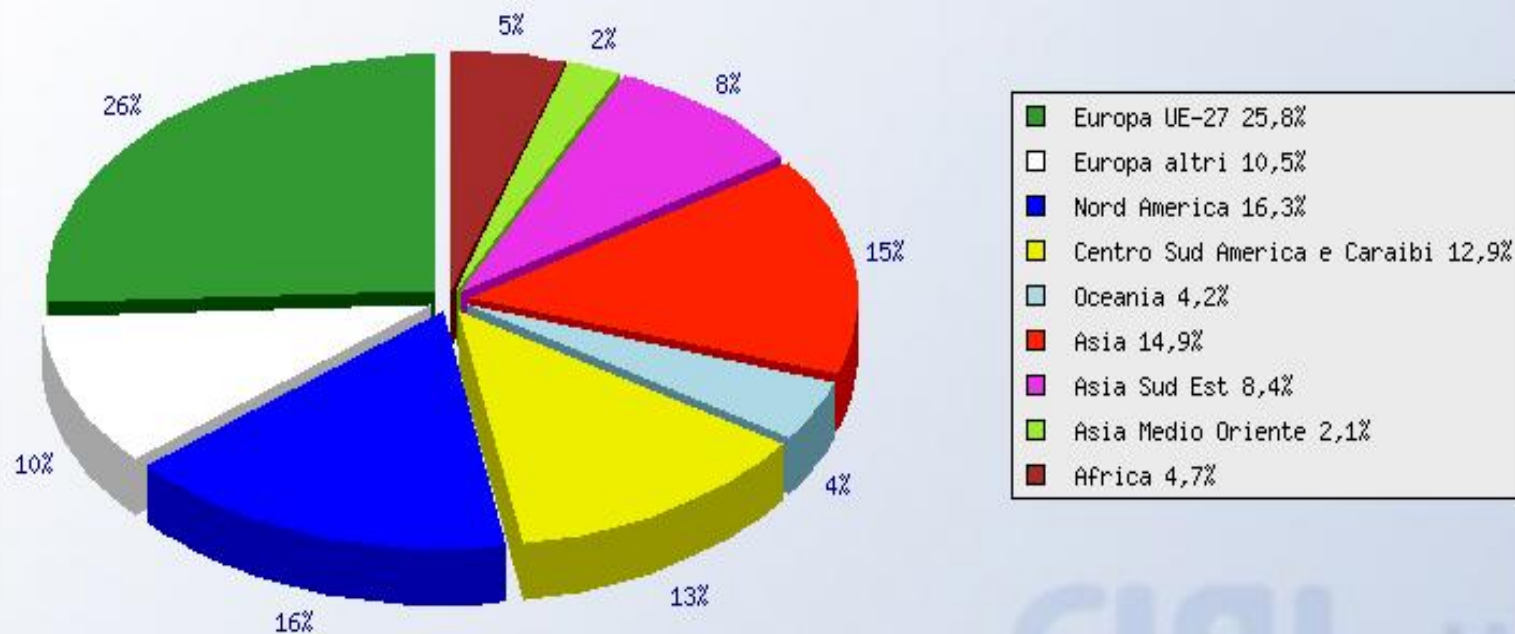
materie preferite bambini di 4° elementare



Gli **areogrammi** sono rappresentazioni grafiche che utilizzano aree di estensioni proporzionali alle frequenze delle diverse modalità



2008 - Produzione Latte Bovino nel Mondo (Fonte FAO)



LAUREATI IN CERCA DI OCCUPAZIONE

Numero di laureati in cerca di occupazione

Valori in migliaia.
Gennaio 2011
FONTE ISTAT
Dati del 2009




Copyright cartogrammi.it [2010]

I **cartogrammi** sono carte geografiche suddivise in zone. Ogni zona viene riempita con un colore corrispondente a una classe di frequenze di un certo carattere quantitativo.

ABRUZZO	7	MOLISE	2
BASILICATA	3	PIEMONTE	11
CALABRIA	13	PUGLIA	18
CAMPANIA	27	SARDEGNA	7
EMILIA ROMAGNA	13	SICILIA	19
FRIULI-VENEZIA GIULIA	4	TOSCANA	13
LAZIO	34	TRENTINO-ALTO ADIGE	2
LIGURIA	5	UMBRIA	4
LOMBARDIA	28	VALLE D'AOSTA	0,1
MARCHE	6	VENETO	14

Gli **ideogrammi** sono rappresentazioni grafiche che utilizzano immagini che richiamano l'oggetto dell'indagine.

 = 100 000 abitanti

Roma 

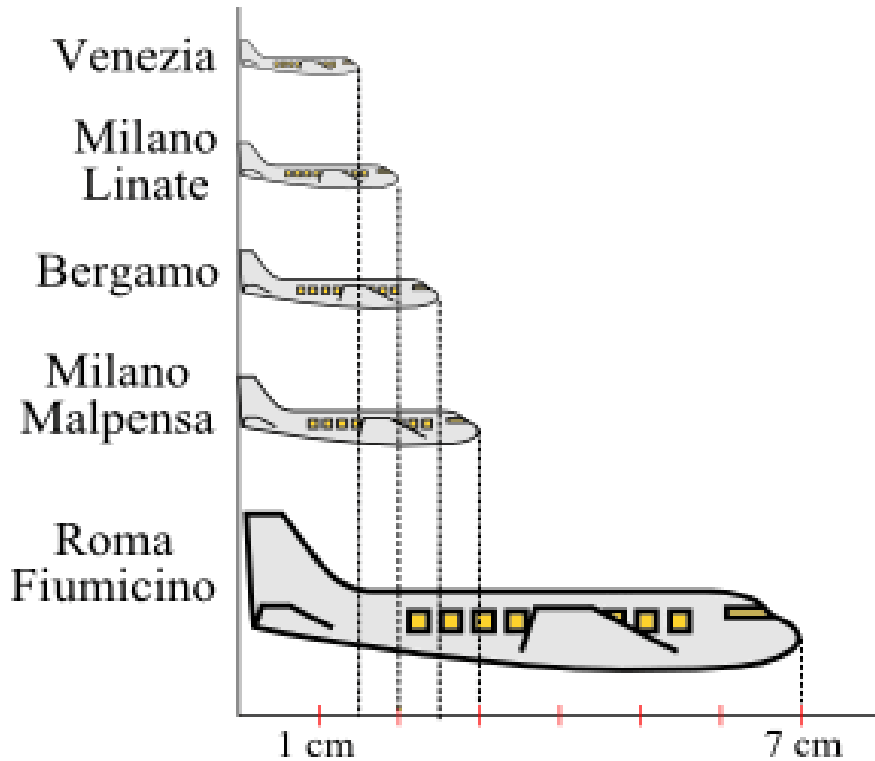
Milano 

Napoli 

Torino 

Palermo 

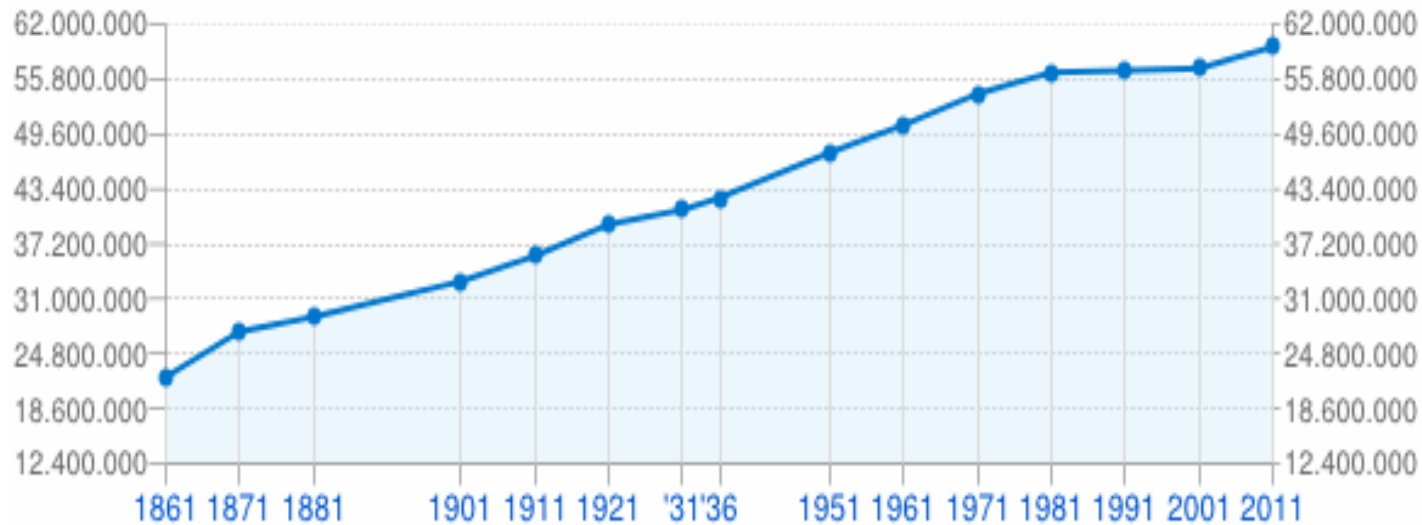
Numero di aerei atterrati in un anno in alcuni aeroporti italiani:



Roma Fiumicino	700.000
Milano Malpensa	300.000
Bergamo	250.000
Milano Linate	200.000
Venezia	150.000

Ideogramma per dimensione: supponiamo che un aereo lungo 1 centimetro equivalga a 100.000 aerei atterrati

I diagrammi cartesiani sono particolarmente usati per rappresentare le serie storiche: si rappresentano sull'asse delle ascisse i tempi, ovvero gli istanti in cui sono state effettuate le rilevazioni, e sull'asse delle ordinate i valori rilevati. I grafici così ottenuti sono costituiti da punti isolati; tali punti possono essere uniti da linee per mettere in rilievo l'andamento del fenomeno.



Popolazione residente ai censimenti

ITALIA - Dati ISTAT - Elaborazione TUTTITALIA.IT

Temperatura media in alcune capitali europee

