

**Percorso: la rappresentazione dei dati**

# **LA RAPPRESENTAZIONE GRAFICA DEI DATI**

# INDICE

1. Introduzione
2. Rappresentazioni grafiche di caratteri qualitativi
3. Rappresentazioni grafiche di caratteri quantitativi
4. Rappresentazioni grafiche di distribuzioni statistiche doppie
5. La scelta della rappresentazione grafica

# 1. Introduzione

- a) Obiettivo
- b) Potenzialità della rappresentazione grafica
- c) Diverse possibilità di rappresentazione a seconda del livello di misurazione dei caratteri
- d) Le componenti di un grafico
- e) La rilevanza dei dati

## a) Obiettivo

Le **rappresentazioni grafiche** hanno l'obiettivo di illustrare, mediante:

- figure,
- linee o segmenti,
- superfici o aree,
- solidi,
- simboli convenzionali
- ecc.



una distribuzione di frequenze o di intensità in funzione delle modalità di uno o più caratteri.

## b) Potenzialità della rappresentazione grafica

- **Visualizzazione immediata dell'andamento del fenomeno e della struttura della distribuzione**, che consente una efficace e globale descrizione dei dati;
- **Sintesi** e quindi possibilità, in poco spazio, di confrontare più distribuzioni (curve, spezzate ecc.);
- **Potenzialità investigative:**
  - mette in risalto i casi "anomali" (particolari *picchi* grafici), imputabili a errori nei dati o a effettivi casi anomali, che vanno ulteriormente approfonditi,
  - evidenzia le correlazioni tra caratteri aventi tra loro un legame logico,
  - permette di individuare andamenti di fondo (*trend*) interpolabili con funzioni matematiche (ad es. curva normale, retta ecc.).

## c) Diverse possibilità di rappresentazione a seconda del livello di misurazione dei caratteri

Per ogni distribuzione statistica semplice o doppia o multipla esiste il tipo di rappresentazione grafica adatta e una stessa distribuzione può essere rappresentata con più tipologie di grafico.

In generale esistono dei vincoli tra tipo di rappresentazione grafica e livello di misurazione dei caratteri da rappresentare che vanno rispettati affinché questa sia corretta, ossia fornisca un'immagine visiva quanto più possibile fedele del fenomeno e della sua distribuzione statistica.

Affinché una rappresentazione grafica sia utile ed efficace dovrebbe contenere con immediatezza e chiarezza tutte le informazioni necessarie alla comprensione dei dati in essa rappresentati.

## d) Le componenti di un grafico

- **I dati:** sono rappresentati in barre, linee, aree o punti.
- **Le componenti di supporto:** consentono la comprensione dei dati:
  - Il titolo del grafico
  - I titoli degli assi
  - Le etichette degli assi
  - L'unità di misura dei dati
  - La griglia
  - La legenda
  - Le etichette dei dati
  - Le note
  - La fonte dei dati.
- **Gli elementi decorativi:** non sono legati ai dati.

## Le componenti di supporto 1/2

- **Il titolo del grafico** deve essere breve e conciso.  
Ne esistono di due tipi:
  - **Il titolo informativo** contiene le informazioni necessarie per comprendere i dati. Risponde alle tre domande: "Cosa?", "Dove?" e "Quando?".  
*p.e. Tasso di disoccupazione in Italia, anni 2009-2012*
  - **Il titolo descrittivo** sintetizza in poche parole il trend o il pattern rappresentato nel grafico.  
*p.e. L'aumento della disoccupazione in Italia dal 2009 al 2012.*
- **I titoli degli assi** identificano le mutabili o le variabili rappresentate dagli assi. Se si possono evincere dal titolo del grafico non è necessario ripeterli.
- **Le etichette degli assi** identificano le modalità o i valori rappresentati nel grafico.

## Le componenti di supporto 2/2

- **L'unità di misura dei dati** (p.e. "in migliaia" , "%" etc.).  
Se l'unità di misura è ovvia, non è necessario specificarla (p.e. "anni" per le serie storiche).
- **La griglia** può essere aggiunta per agevolare la lettura e il confronto dei dati.
- **La legenda** identifica simboli, tratteggi o colori usati per rappresentare i dati.
- **Le etichette dei dati** visualizzate sopra o vicino alle barre, alle aree o alle linee facilitano la lettura del grafico.
- **Le note** possono essere aggiunte per fornire definizioni o informazioni sulla metodologia.
- **La fonte** di provenienza dei dati.

## e) La rilevanza dei dati 1/2

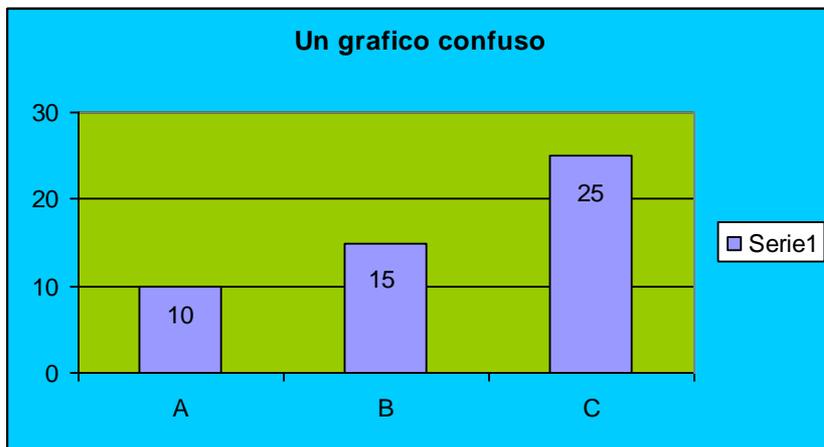
Per massimizzare l'efficacia di un grafico l'attenzione deve essere concentrata sui dati.

Pertanto le **componenti di supporto**:

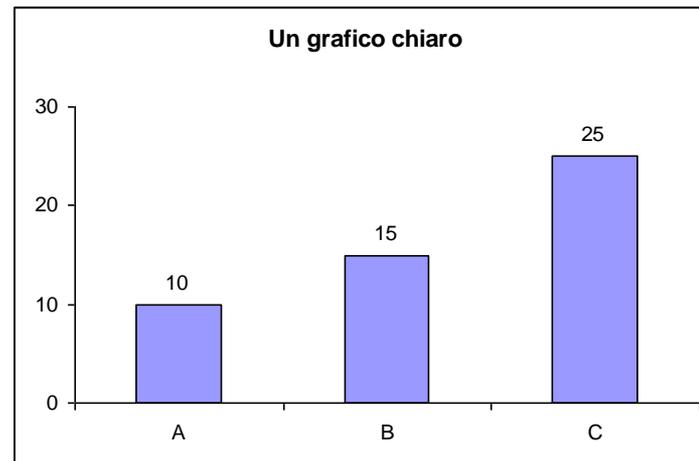
- **devono essere presenti solo se necessarie**: titoli degli assi, legende e etichette in alcuni casi possono essere essenziali per la comprensione del grafico, ma in altri possono essere del tutto inutili.
- **devono essere lievi**: è preferibile usare linee più leggere per gli assi e per la griglia e linee più marcate per i dati.

**Gli effetti decorativi** non devono allontanare l'attenzione del lettore dai dati.

## e) La rilevanza dei dati 2/2



Nel grafico a sinistra tutte le componenti hanno il massimo impatto. Il risultato è un grafico confuso, difficile da leggere anche se sono presenti solo 3 valori.



Il grafico a destra è più facile da leggere. Il ricorso a poche componenti di supporto permette di concentrare l'attenzione sui dati.

## 2. Rappresentazioni grafiche di caratteri qualitativi

- a) Grafici a barre
- b) Diagrammi circolari
- c) Diagrammi in coordinate polari
- d) Cartogrammi, mappe tematiche

## a) Grafici a barre

(1/3)

I grafici a barre (o ortogrammi) sono impiegati per rappresentare graficamente **serie sconnesse** o **rettilinee** e possono essere di due tipi:

-**A colonne**, se sono costituiti da una successione di colonne, segmenti verticali o rettangoli (a base uguale o arbitraria) equidistanti, in numero pari alle modalità del carattere, e hanno *altezza* uguale o proporzionale alla frequenza (assoluta o relativa) o all'intensità della modalità da rappresentare;

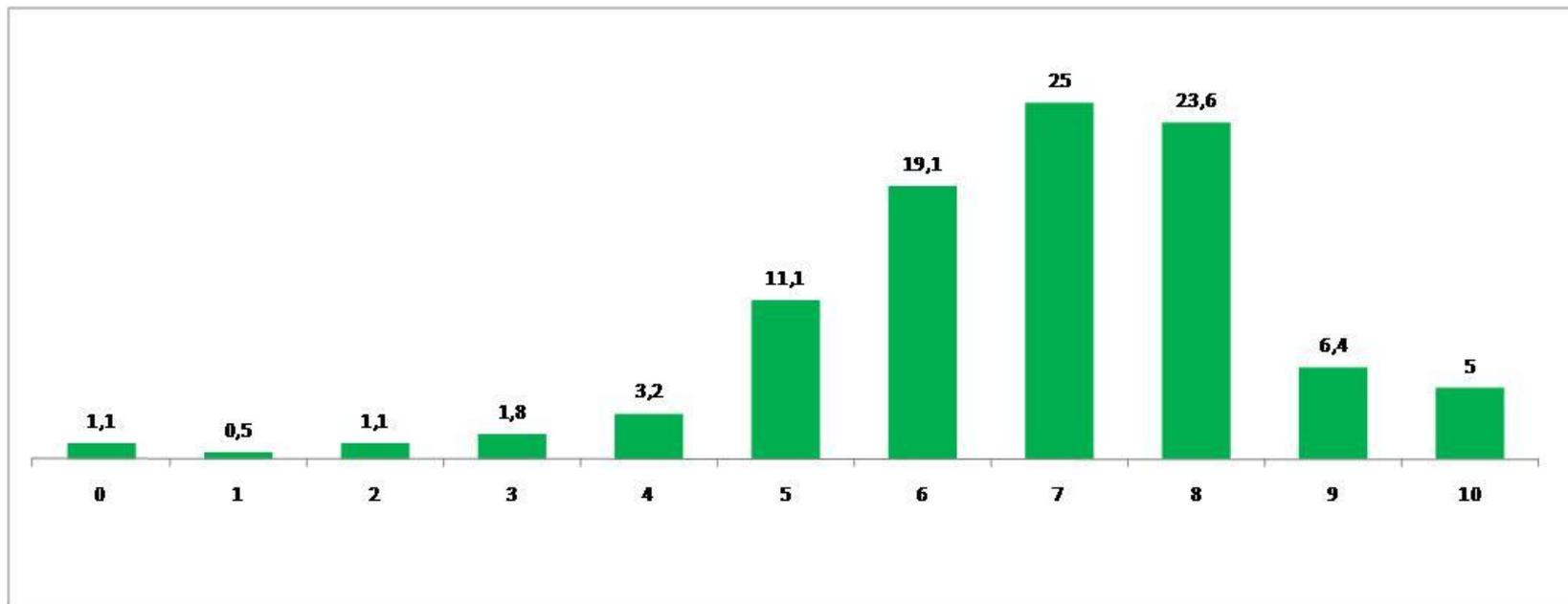
-**A nastri**, se sono costituiti da tanti nastri (segmenti orizzontali, rettangoli) sovrapposti ed equidistanti, in numero pari alle modalità del carattere, e hanno *lunghezza* uguale o proporzionale alla frequenza (assoluta o relativa) o all'intensità della modalità da rappresentare.

I grafici a barre sono pertanto caratterizzati dall'averne un solo asse (verticale nel caso a colonne, orizzontale nel caso a nastri) in scala graduata, secondo l'unità di misura che si è scelta per rappresentare le frequenze o intensità. Sull'altro asse, invece, figureranno le modalità (qualitative), per convenzione equidistanti.

## a) Ortogramma a colonne

(2/3)

**Persone di 14 anni e più per livello di soddisfazione per la vita nel complesso (a) in Italia  
Anno 2013** (per 100 persone di 14 anni e più)



Fonte: Istat, Indagine Multiscopo "Aspetti della vita quotidiana"

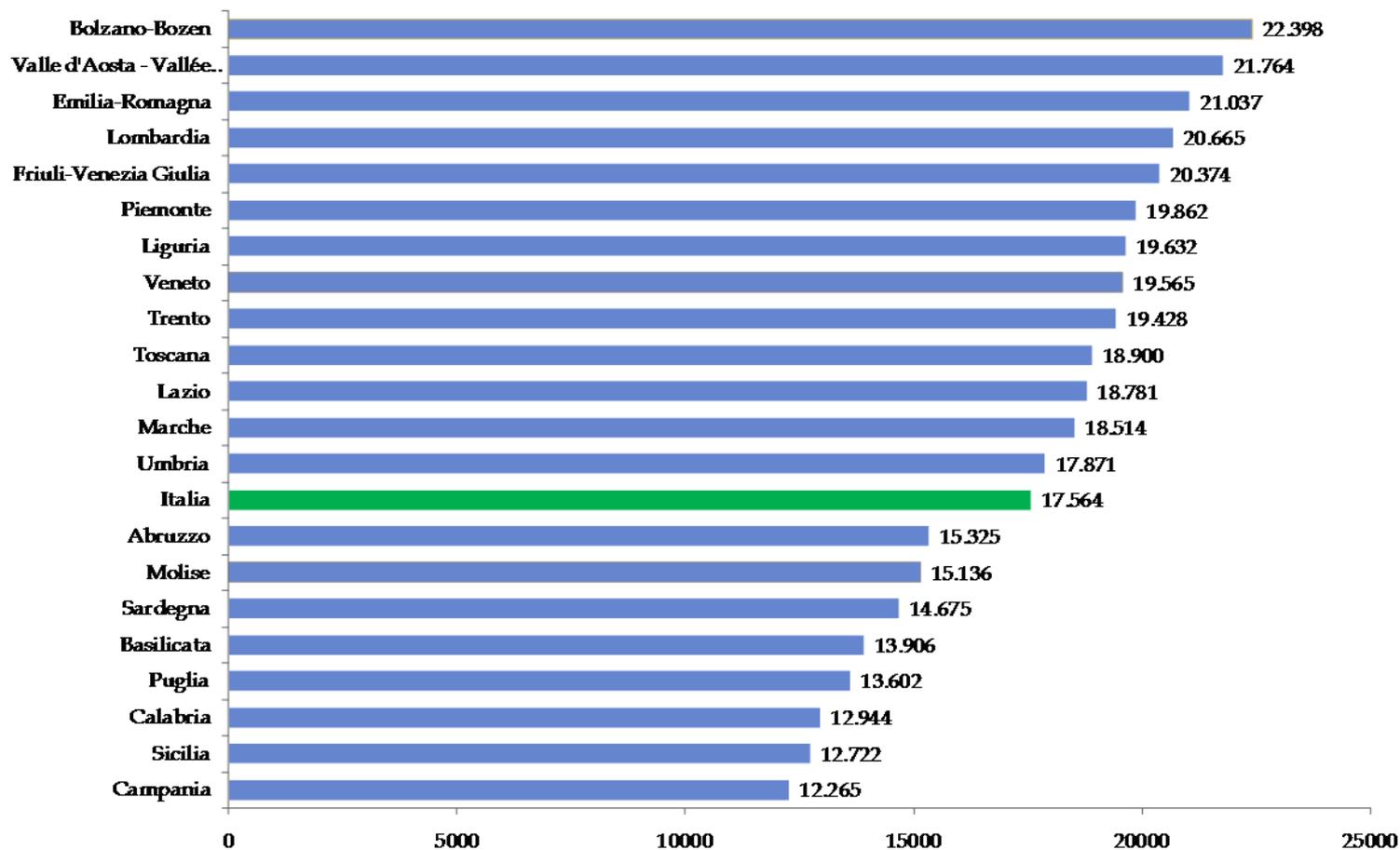
(a) Espresso con un punteggio da 0 (per niente soddisfatto) a 10 (molto soddisfatto)

Se la rappresentazione grafica riguarda una serie sconnessa, l'ordine in cui saranno poste le modalità è arbitrario; se si tratta invece di una serie rettilinea (come in questo caso), le modalità saranno poste nell'ordine naturale che esse presentano nella serie.

## a) Ortogramma a nastri

(3/3)

Reddito disponibile delle famiglie per abitante per regione – Anno 2012 (*valori in euro*)



## b) Diagrammi circolari

(1/2)

I diagrammi circolari (o aereogrammi) per la loro forma circolare, sono comunemente noti come "diagrammi a torta".

Sono particolarmente adatti alle serie **sconnesse o rettilinee**.

Sono efficaci per mettere in evidenza l'importanza relativa delle singole modalità rispetto al totale.

Ne esistono di vari tipi e principalmente:

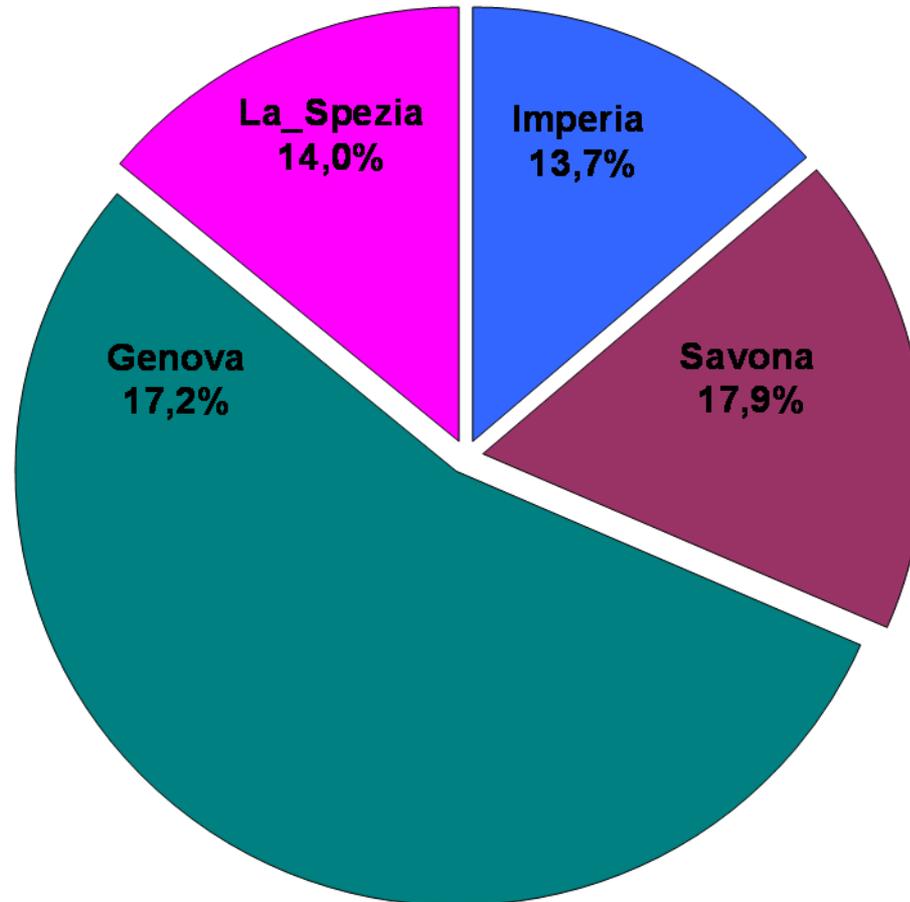
-**A spicchi o settori variabili**, con angoli al centro corrispondenti alle frequenze assolute ( o relative) delle singole modalità e raggio fisso;

-**A spicchi o settori fissi**, con angoli al centro uguali e raggio variabile corrispondente alle frequenze assolute (o relative) delle singole modalità.

## b) Diagrammi circolari

(2/2)

Composizione percentuale della popolazione residente in Liguria per provincia  
Censimento 2011



Fonte: Istat

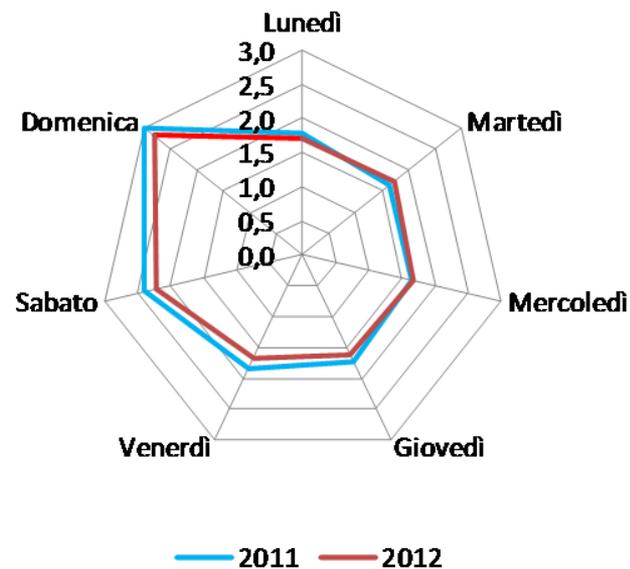
## c) Diagrammi in coordinate polari

Si utilizzano per rappresentare serie cicliche.

Le **serie cicliche** sono distribuzioni di mutabili che dispongono di un ordinamento naturale ma non di una prima e un'ultima modalità.

**Esempio.** I nati, i matrimoni, le vendite di determinati prodotti, fenomeni astronomici o atmosferici e così via secondo i mesi dell'anno, i giorni della settimana ecc.

Indice di mortalità in incidenti stradali per giorno della settimana  
Anni 2011 e 2012



Fonte: Istat

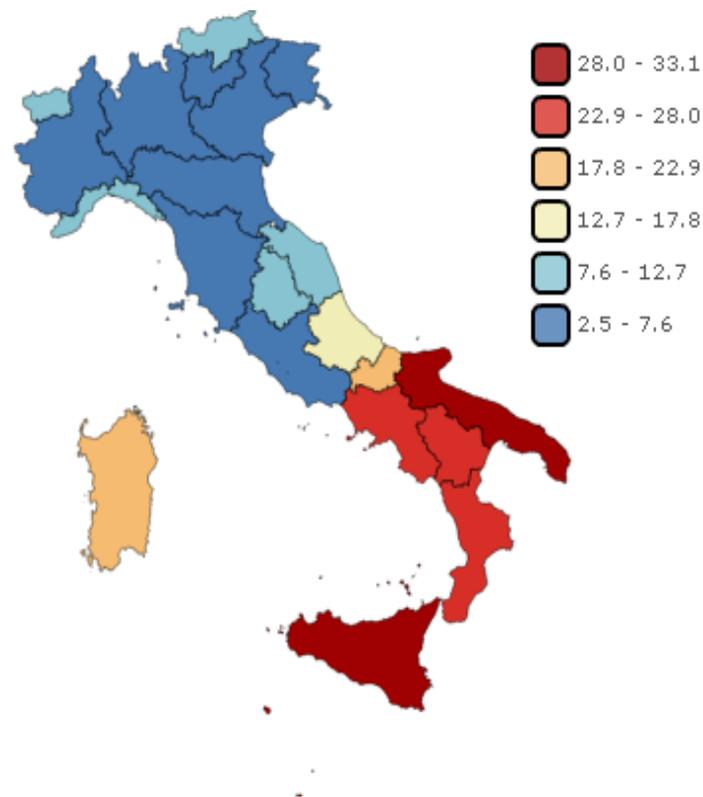
## d) Cartogrammi, mappe tematiche

(1/2)

I cartogrammi sono grafici utili per rappresentare **serie territoriali**.

Per costruire un cartogramma occorre disporre di una carta geografica o topografica in cui siano chiaramente delimitate le diverse zone, regioni, circoscrizioni (geografiche, politiche, amministrative) rispetto alle quali viene analizzata l'intensità o la frequenza di uno o più caratteri (es. nati, morti, reddito pro capite, secondo le Regioni, Province, Comuni).

Incidenza della povertà relativa per regione  
Anno 2012 (valori percentuali)



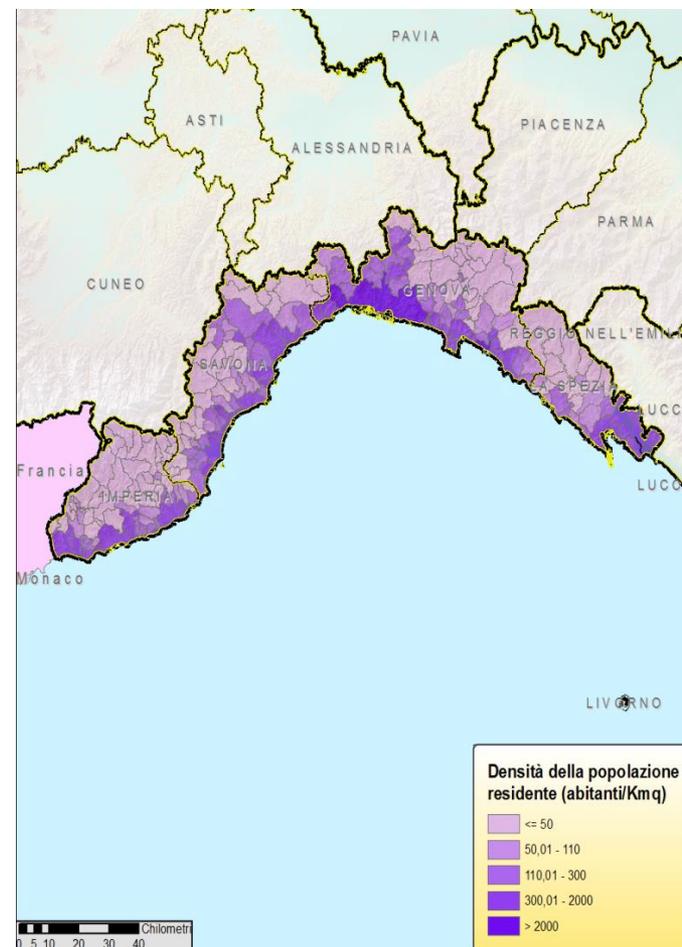
Fonte: Istat

## d) Cartogrammi, mappe tematiche

(2/2)

Le mappe tematiche consentono, anche se molto empiricamente e in prima approssimazione, di tener conto dell'autocorrelazione spaziale, intesa come possibilità che un fenomeno che interessa un certo territorio sia influenzato nelle sue manifestazioni dalla contiguità spaziale esistente tra i luoghi in cui il fenomeno è osservato.

Per esempio nella mappa della Liguria qui a destra, si nota come la densità della popolazione residente sia maggiore nei comuni costieri e in quelli circostanti i comuni capoluogo rispetto a quelli dell'entroterra.



Fonte: Istat

### 3. Rappresentazioni grafiche di caratteri quantitativi

- a) Istogrammi
- b) Diagrammi cartesiani a segmenti
- c) Poligono e curve di frequenza per variabili continue
- d) Rappresentazioni di tipo informatico:
  - a ramo e foglia (stem and leaf plot)
  - a scatola (boxplot)
  - le facce di Chernoff
  - ideogrammi o pictogrammi

## a) Istogrammi

(1/3)

Gli istogrammi si impiegano per rappresentare graficamente distribuzioni di frequenza di **caratteri quantitativi** le cui modalità sono costituite da **classi di valori**.

L'ipotesi di base su cui si fonda la rappresentazione per istogrammi è che **all'interno di ciascuna classe le unità siano equidistribuite**.

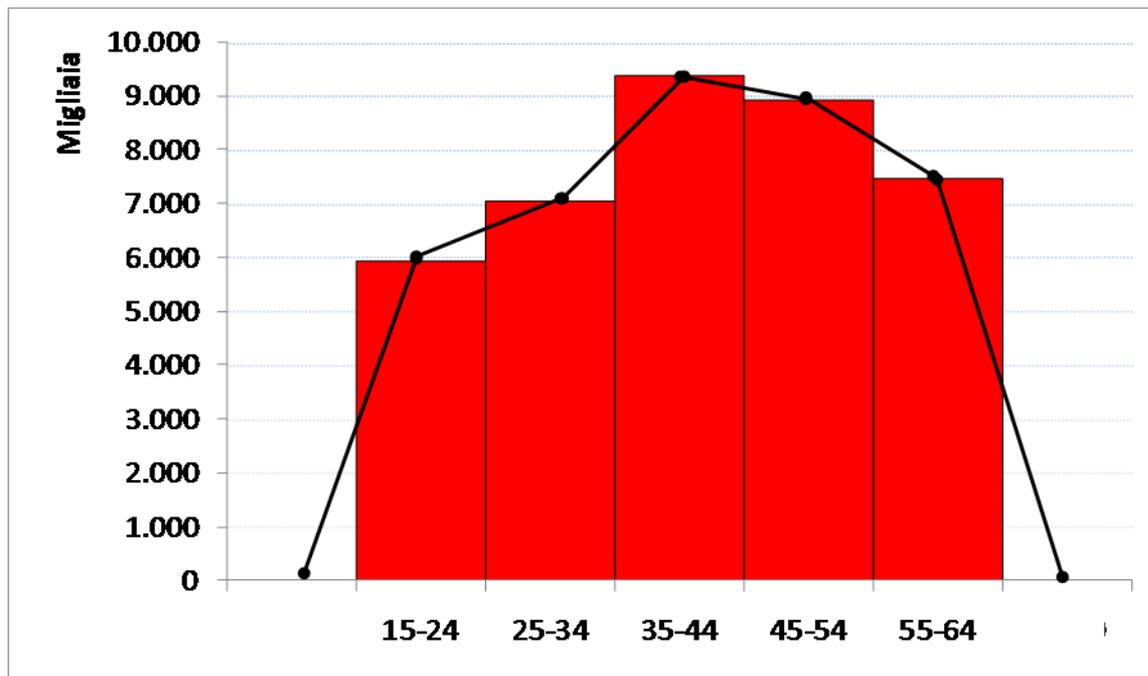
Per rendere più evidente l'andamento e la forma della distribuzione, si possono congiungere con segmenti di retta i punti centrali dei lati superiori dei rettangoli ottenendo una linea spezzata detta **poligono di frequenza**, la quale deve essere chiusa toccando l'asse delle ascisse all'esterno delle classi estreme, in modo che l'area all'interno del poligono di frequenza equivalga a quella dell'istogramma.

## a) Istogrammi

(2/3)

-Le **classi di valori** hanno **uguale ampiezza**. In questo caso avremo tanti rettangoli contigui, ciascuno avente base uguale all'ampiezza della classe e altezza uguale o proporzionale alla frequenza (assoluta o relativa) assunta nell'insieme delle unità della classe;

**Istogramma e poligono di frequenza della popolazione residente in Italia in età lavorativa per classi decennali di età – Censimento 2011 (popolazione in migliaia)**

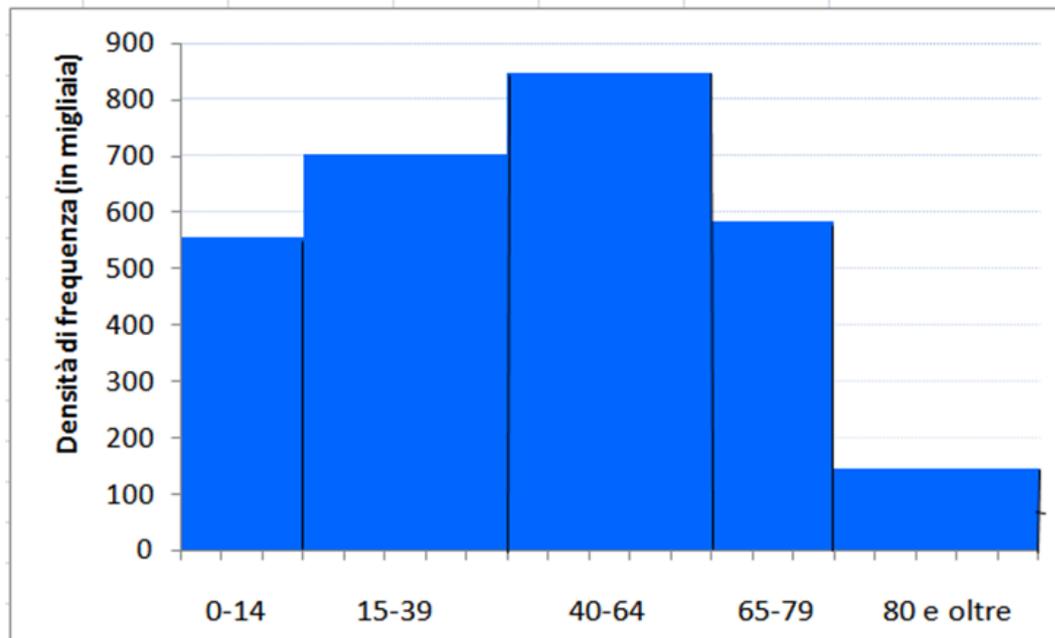


## a) Istogrammi

(3/3)

- Le **classi di valori** hanno **diversa ampiezza**. In quest'altro caso avremo una serie di rettangoli aventi basi diverse uguali all'ampiezza delle classi e altezze da calcolarsi, in modo che le frequenze siano proporzionali alle aree dei rispettivi rettangoli.

**Istogramma della popolazione residente in Italia per classi di età – Censimento 2011 (popolazione in migliaia)**



In ordinata abbiamo le **densità di frequenza** date dal rapporto tra la frequenza (assoluta o relativa) di ciascuna classe e la relativa ampiezza.

Fonte: Istat

## b) Diagrammi cartesiani a segmenti (1/2)

Sono impiegati per rappresentare graficamente caratteri quantitativi discreti, non divisi in classi, e possono configurarsi a segmenti verticali.

**Esempio.** Numero dei componenti per famiglia, numero delle stanze delle abitazioni, numero di unità locali delle aziende e così via.

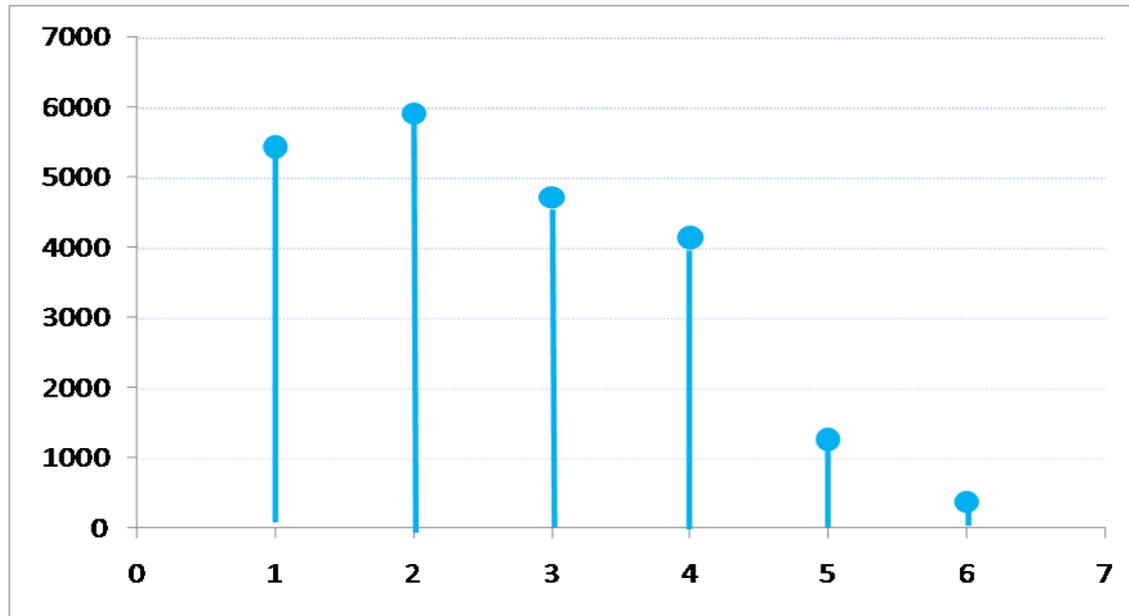
Essi si costruiscono come gli usuali diagrammi cartesiani aventi due assi perpendicolari: l'asse delle ascisse ( $x$ ) e l'asse delle ordinate ( $y$ ), aventi origine comune in zero. Ogni coppia ordinata di valori  $(x_i, y_i)$  determinerà un punto nel piano e l'insieme di tutte le coppie ( $x_i =$  modalità quantitativa  $i$ -esima,  $y_i =$  frequenza della modalità  $i$ -esima) determinerà l'insieme dei punti nel piano che costituiscono la rappresentazione grafica della distribuzione considerata.

Per rendere maggiormente visibili tali punti, si tracciano dei segmenti verticali congiungenti l'ascissa ( $x_i$ ) con il punto del piano corrispondente all'ordinata ( $y_i$ ).

## b) Diagrammi cartesiani a segmenti

(2/2)

Distribuzione delle famiglie residenti in Italia per numero di componenti – Censimento 2001 (valori in migliaia)



Fonte: Istat

E' da notare che in questo caso è scorretto costruire il poligono o spezzata di frequenza congiungendo tra loro i punti poiché il carattere considerato è discreto e quindi non possiede i valori intermedi a quelli indicati dalle modalità quantitative.

Una spezzata di frequenza che unisse tra loro le modalità, infatti, attribuirebbe anche valori intermedi alle modalità stesse.

## c) Poligono e curva di di frequenza per variabili continue (1/4)

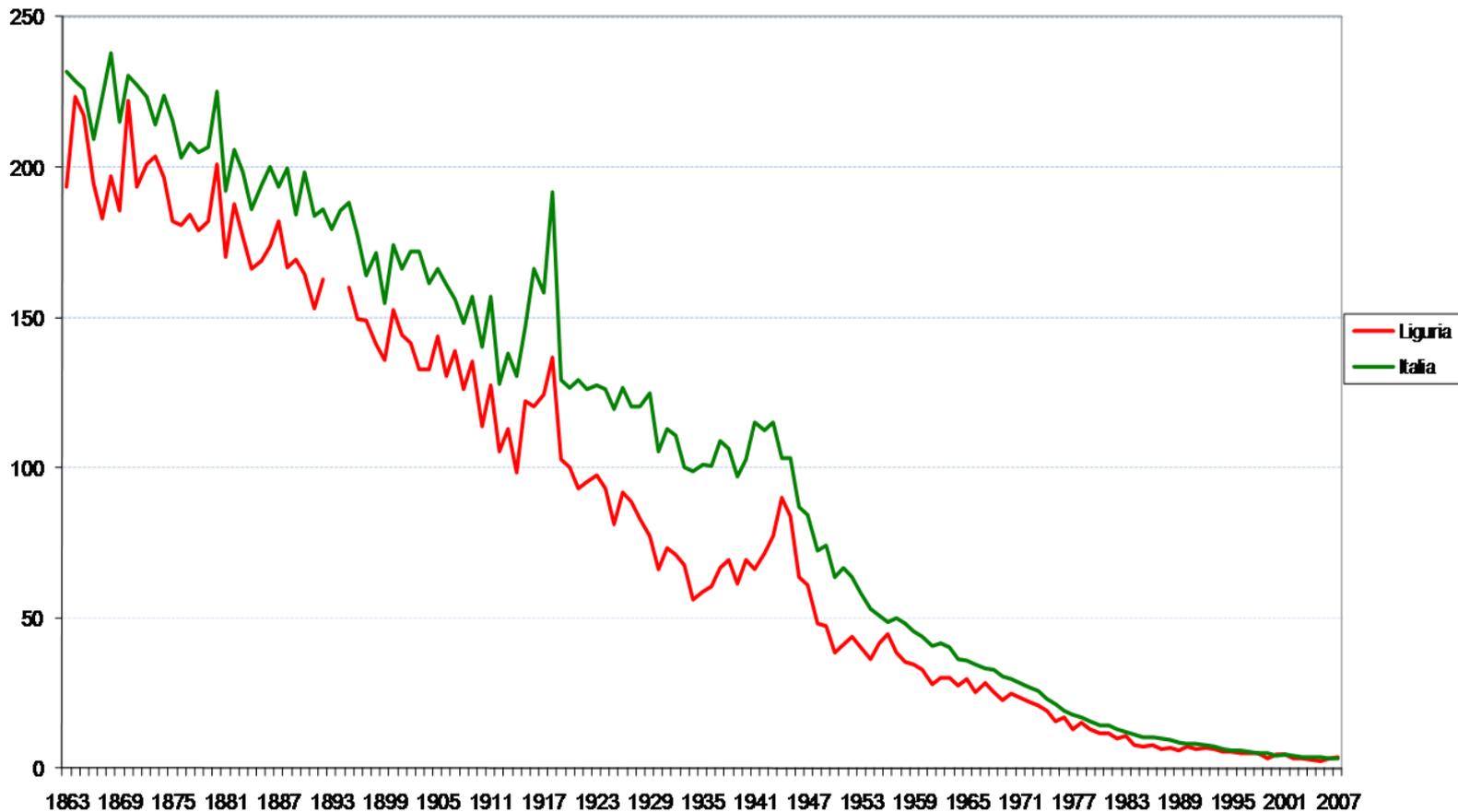
I diagrammi cartesiani ortogonali sono impiegati anche per rappresentare graficamente i caratteri quantitativi continui (come ad esempio età, prezzi) o, nel caso delle serie storiche, per quei caratteri che si suppone si modifichino con continuità nel tempo.

La costruzione di tali diagrammi è del tutto analoga a quella vista per i diagrammi ad aste o segmenti salvo che, in questo caso, una volta ottenuti nel piano cartesiano i punti  $(x_i, y_i)$  rappresentanti la distribuzione considerata, essi vanno uniti mediante una spezzata detta **poligono di frequenza**.

All'aumentare del numero di unità rilevate il poligono di frequenza si approssima sempre più a una linea continua detta **curva di frequenza**,

## c) Poligono e curva di frequenza per variabili continue (2/4)

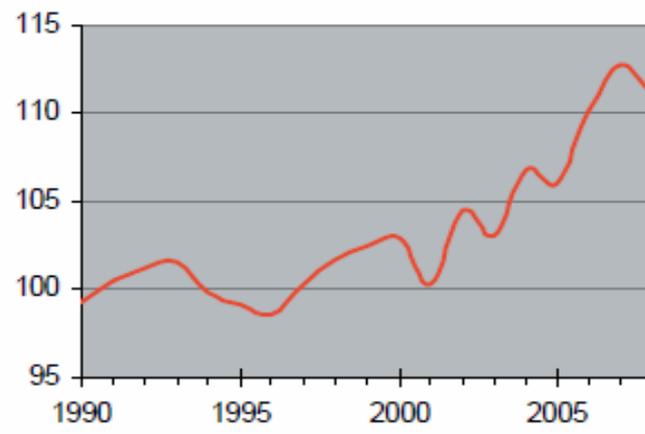
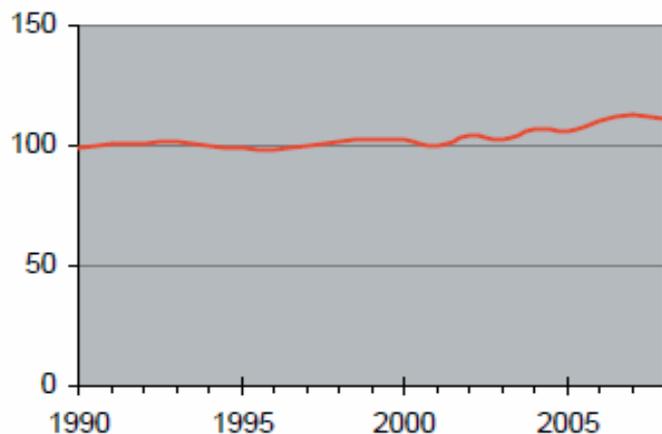
Evoluzione della mortalità infantile in Liguria e in Italia - Anni 1863 - 2007  
(morti 0-1 anno per 1.000 nati vivi)



Fonte: Istat

## c) Poligono e curva di frequenza per variabili continue (3/4)

**Problemi di scala:** questi grafici rappresentano la stessa distribuzione ma l'andamento della serie viene percepito in maniera diversa.



Fonte: UNECE, Making Data Meaningful Part 2: A guide to presenting statistics, United Nations, Geneva 2009.

Il grafico a sinistra mostra un trend stabile per circa 10 anni, seguito da una moderata crescita.

Riducendo la scala sull'asse delle ordinate (grafico a destra), risultano più evidenti le variazioni intervenute nei primi dieci anni e sembra più marcata la crescita dell'ultimo periodo.

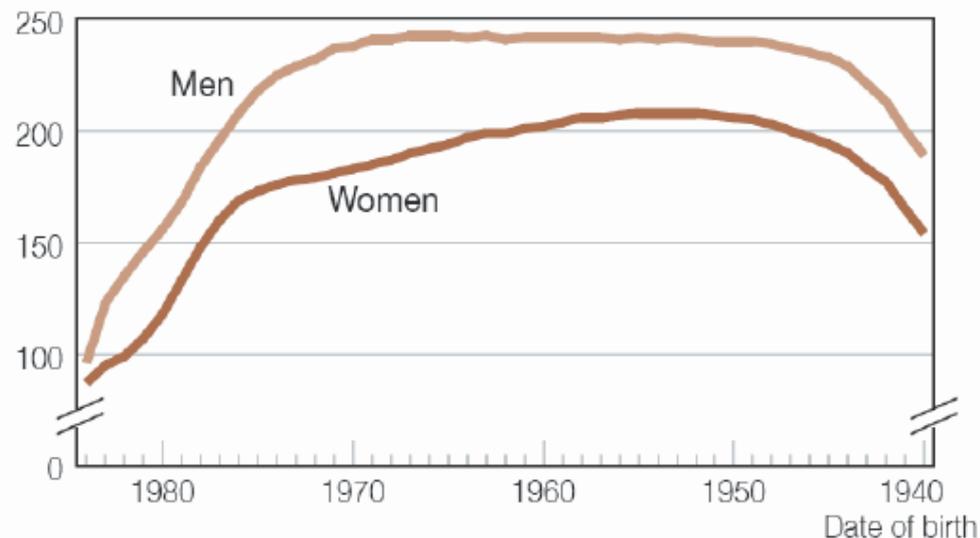
## c) Poligono e curva di frequenza per variabili continue

(4/4)

**Problemi di scala:** quando l'asse delle ordinate non inizia dallo zero è opportuno evidenziarlo, in modo da richiamare l'attenzione del lettore su quest'aspetto.

**GOOD EXAMPLE** of a chart with a y-axis not starting at zero

**Pensionable income for persons aged 20-64 in Sweden, 2004**  
Average income in thousands Swedish Krona



Source: Statistics Sweden (2006), *Women and Men in Sweden: Facts and figures 2006*<sup>8</sup>.

## d) Rappresentazioni di tipo informatico (1/8)

Ne possiamo distinguere quattro tipologie principali:

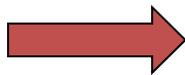
- **a ramo a foglia** (stem and leaf).

Si considerino i dati relativi ai voti di maturità di 10 ragazzi:

65 74 79 83 83 86 88 92 97 99

Il primo passaggio è quello di dividere le decine dalle unità come:

6 5
7 4
7 9
8 3
8 3
8 6
8 8
9 2
9 7
9 9



Il secondo passaggio è quello di raggruppare i rami posizionando le relative foglie una accanto all'altra:

6 5
7 49
8 3368
9 279

Nella rappresentazione a ramo e foglia si considerano le due cifre iniziali del quoziente d'intelligenza come "rami" e la terza cifra come "foglia" ottenendo:

## d) Rappresentazioni di tipo informatico (2/8)

Questa tipologia di rappresentazione grafica pone immediatamente in evidenza sia quali sono i dati aventi come prima cifra 6, 7, 8, 9 sia il loro numero.

In tal modo è possibile visualizzare più modalità quantitative che differiscono tra di loro, ad esempio, sulle unità.

Le foglie possono anche disporsi in ordine crescente ottenendo:

<b>Ramo</b>	<b>Foglia</b>		<b>Totale</b>
6	5		1
7	49		2
8	3368		4
9	279		3
			<b>10</b>

Giusto per avere un'idea, da questa tavola si legge che un solo ragazzo ha un voto di maturità pari a 65, tre superiori a 90.

## d) Rappresentazioni di tipo informatico (3/8)

- **a scatola** (boxplot), la quale consente di disporre delle principali caratteristiche di una distribuzione statistica semplice.

Nell'esempio di rappresentazione a boxplot riportata di seguito:

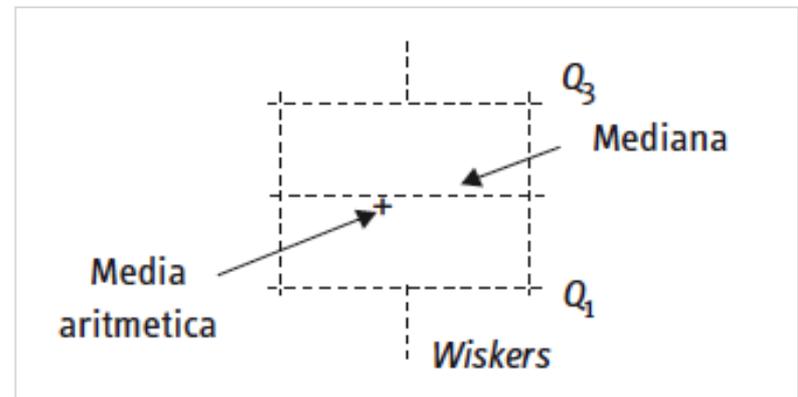


figura un rettangolo in cui:

- I lati inferiore e superiore indicano rispettivamente il 1° quartile ( $Q_1$ ) e il 3° quartile ( $Q_3$ ) della distribuzione considerata
- La linea tratteggiata all'interno del rettangolo indica la posizione della mediana ( $Q_2=2^\circ$  quartile)
- Il simbolo "+" rappresenta la media aritmetica
- La linea verticale tratteggiata (whiskers) indica l'estensione della distribuzione prima di  $Q_1$  e dopo  $Q_3$ .

## d) Rappresentazioni di tipo informatico (4/8)

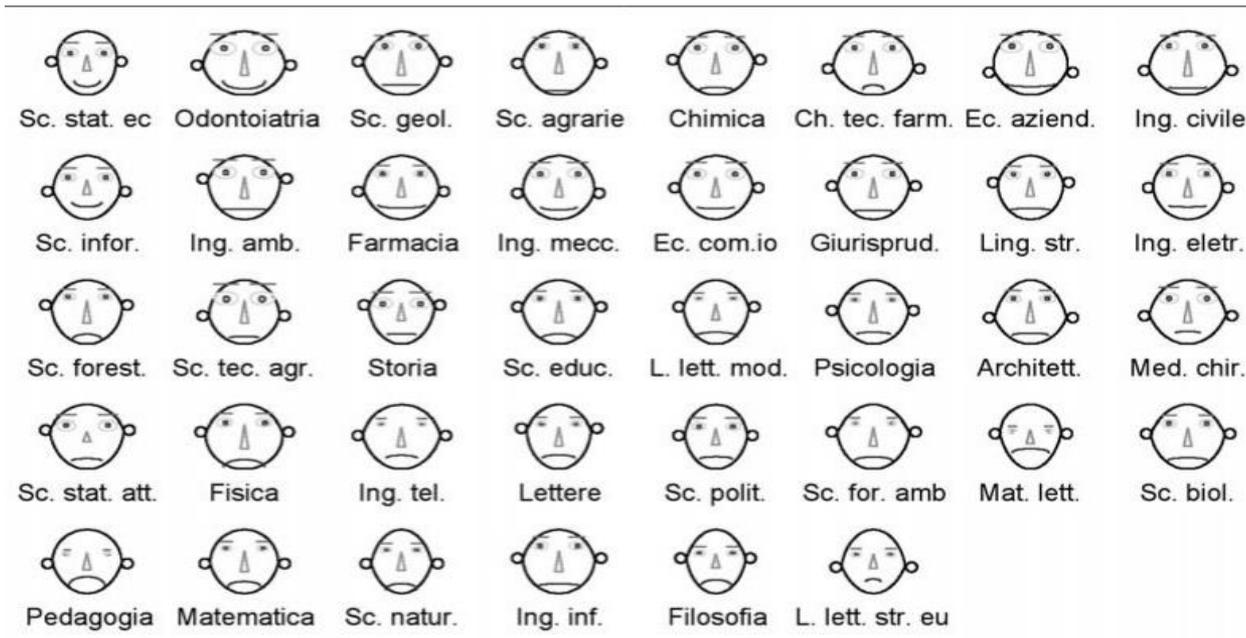
- le **facce di Chernoff**, una rappresentazione grafica particolare, introdotta nel 1973 ed eseguibile soltanto tramite computer e software apposito, in cui si utilizzano i tratti di un volto (occhi, naso, bocca) per rappresentare le diverse componenti di un'informazione.

Tali rappresentazioni sono adatte sia per distribuzioni statistiche semplici che multiple.

Un limite di queste rappresentazioni, originali ed efficaci, è rappresentato dal fatto che non sempre è immediatamente chiaro quali elementi dei grafici rappresentano i dati. La forma e l'espressione del viso, infatti, possono risultare dalla combinazione dei valori corrispondenti a più variabili e l'informazione che se ne ricava sarà più di tipo qualitativo (somiglianze, dissomiglianze) che quantitativo.

## d) Rappresentazioni di tipo informatico (5/8)

Livelli di soddisfazione di alcuni aspetti del lavoro dei laureati occupati dell'Ateneo fiorentino nell'anno 2000



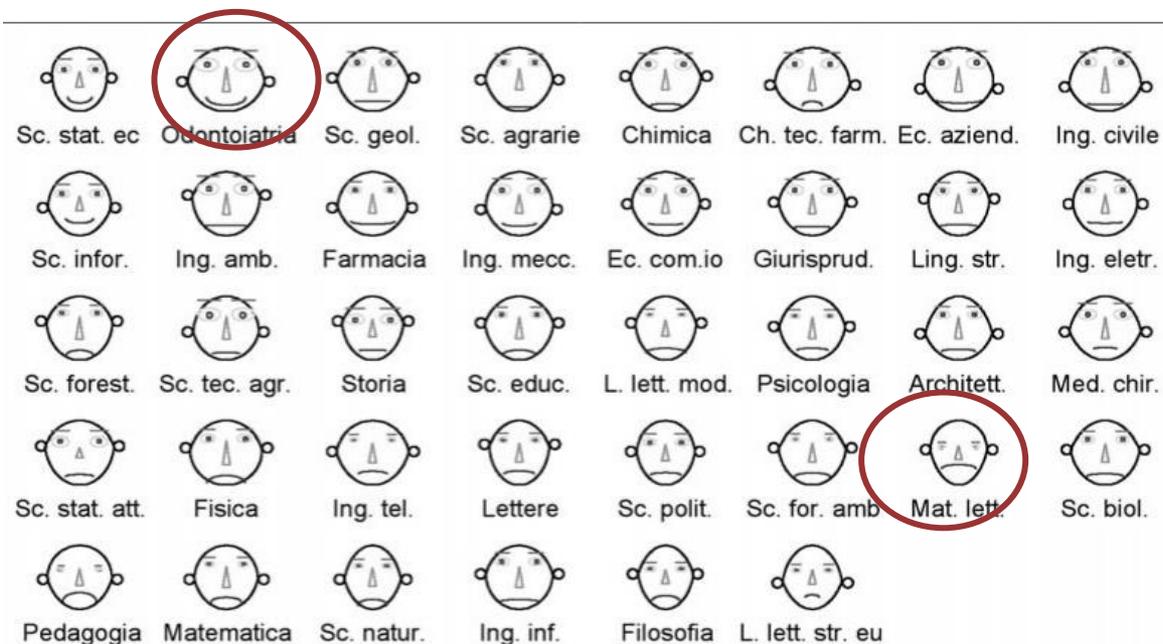
Fonte: Chiandotto B., Sulla misura della qualità della formazione universitaria, in Studi e Note di Economia 3/2004

### Legenda:

- larghezza della faccia: coerenza studi fatti
- livello delle orecchie: tempo libero
- ecc della fronte: stabilità/sicurezza
- ecc p.inf: acquisizioni professionali
- lunghezza naso: rispondenza interessi culturali
- centratura bocca: indipendenza/autonomia
- curvatura bocca: possibilità di guadagnare
- larghezza bocca: locazione sede di lavoro
- altezza occhi: rapporti con i colleghi
- larghezza occhi: possibilità di carriera

## d) Rappresentazioni di tipo informatico (6/8)

Livelli di soddisfazione di alcuni aspetti del lavoro dei laureati occupati dell'Ateneo fiorentino nell'anno 2000



LEGENDA: faccia/lar = Coerenza studi fatti, orec/liv = Tempo libero, fronte/ecc = Stabilità/sicurezza, p.inf/ecc = Acquis. professional, naso/lun = Rispon. interessi cu, bocca/cent = Indipendenza/autonom, bocca/curv = Possibilità di guad, bocca/lar = Locazione sede lavor, occhi/al = Rapporti colleghi, occhi/lar = Possibilità di carriera

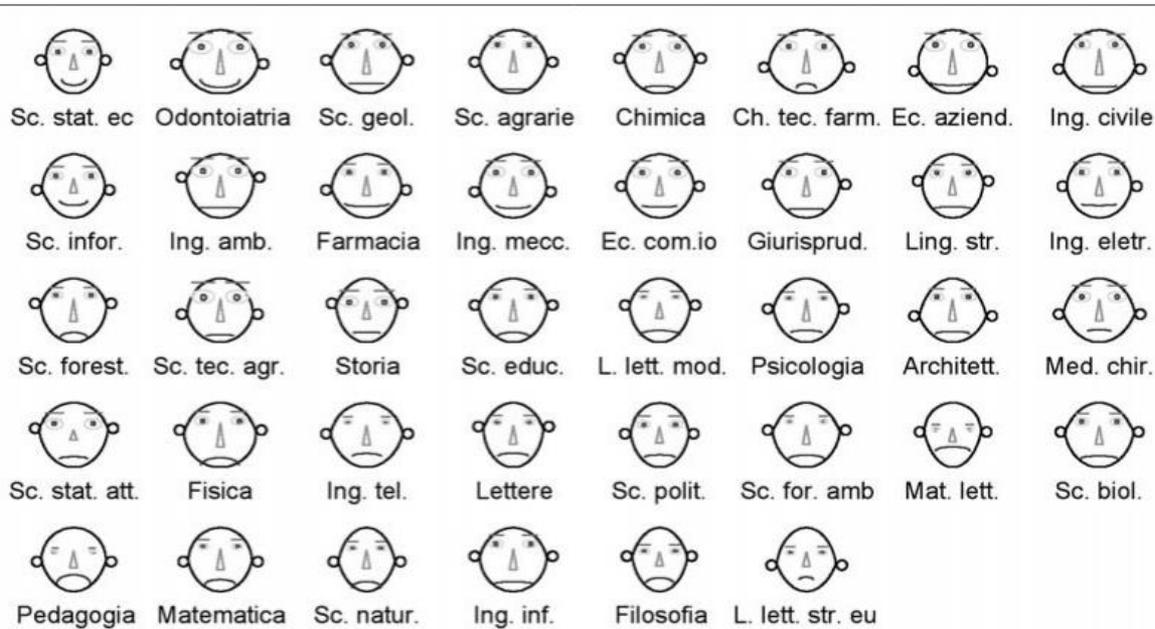
La larghezza della faccia, p.e. rappresenta la coerenza degli studi fatti con l'attività lavorativa svolta.

I laureati in odontoiatria registrano una maggiore coerenza rispetto ai laureati in materie letterarie.

Fonte: Chiandotto B., Sulla misura della qualità della formazione universitaria, in Studi e Note di Economia 3/2004.

## d) Rappresentazioni di tipo informatico (7/8)

Livelli di soddisfazione di alcuni aspetti del lavoro dei laureati occupati dell'Ateneo fiorentino nell'anno 2000



LEGENDA: faccia/lar = Coerenza studi fatti, orec/liv = Tempo libero, fronte/ecc = Stabilità/sicurezza, p.inf/ecc = Acquis. professional, naso/lun = Rispon. interessi cu, bocca/cent = Indipendenza/autonom, bocca/curv = Possibilità di guada, bocca/lar = Locazione sede lavor, occhi/al = Rapporti colleghi, occhi/lar = Possibilità di carriera

La curvatura della bocca rappresenta la possibilità di guadagnare.

Quali sono le facoltà che danno opportunità maggiori di guadagno?

Osservate i sorrisi!

Fonte: Chiandotto B., Sulla misura della qualità della formazione universitaria, in Studi e Note di Economia 3/2004.

## d) Rappresentazioni di tipo informatico (7/8)

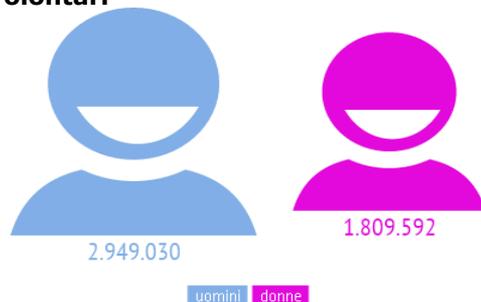
- **ideogrammi** o **pittogrammi**, rappresentazioni grafiche aventi carattere molto divulgativo perché si avvalgono di figure, simboli, generalmente tutti simili tra loro, aventi un'immediata attinenza con il carattere considerato (figure umane, oggetti ecc.) e di grandezza o numero variabile per indicare l'entità della frequenza o dell'intensità del carattere rappresentato.

Qualsiasi carattere statistico (qualitativo o quantitativo) può essere rappresentato con questa tipologia di grafici.

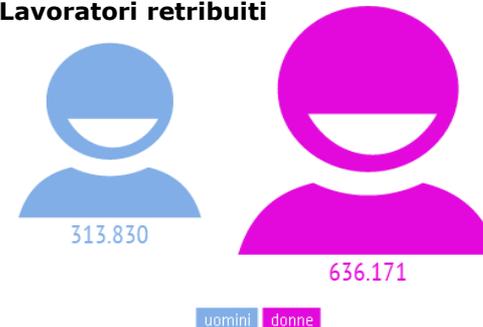
E' da notare, tuttavia, che tali rappresentazioni possono prestarsi ad equivoci perché non sempre è agevole leggerli ed essi hanno una scarsa utilità scientifica.

### Volontari e lavoratori retribuiti nelle istituzioni non profit per genere – Italia Censimento 2011

#### Volontari



#### Lavoratori retribuiti



## 4. Rappresentazioni grafiche di distribuzioni statistiche doppie

- a) Nuvola dei punti
- b) Stereogramma
- c) Diagrammi a barre e cartodiagrammi
- d) Piramide delle età

## a) Nuvola dei punti o scatter plot

(1/2)

È una rappresentazione grafica molto utilizzata nel caso di una distribuzione statistica doppia in cui entrambi i caratteri sono quantitativi perché consente di visualizzare la correlazione esistente fra le variabili.

Lo scatter plot è un diagramma cartesiano in cui sull'asse delle ascisse ( $x$ ) e su quello delle ordinate ( $y$ ) sono riportati i valori assunti dalle due variabili ( $X, Y$ ) e ciascun punto  $P$  del piano rappresenta l'unità statistica avente come valore delle due variabili le coordinate  $(x_i, y_i)$ .

La nuvola dei punti consente di vedere la dispersione tra le unità statistiche ossia la loro vicinanza o distanza, indicanti rispettivamente la loro somiglianza o dissomiglianza rispetto a due caratteri considerati contemporaneamente.

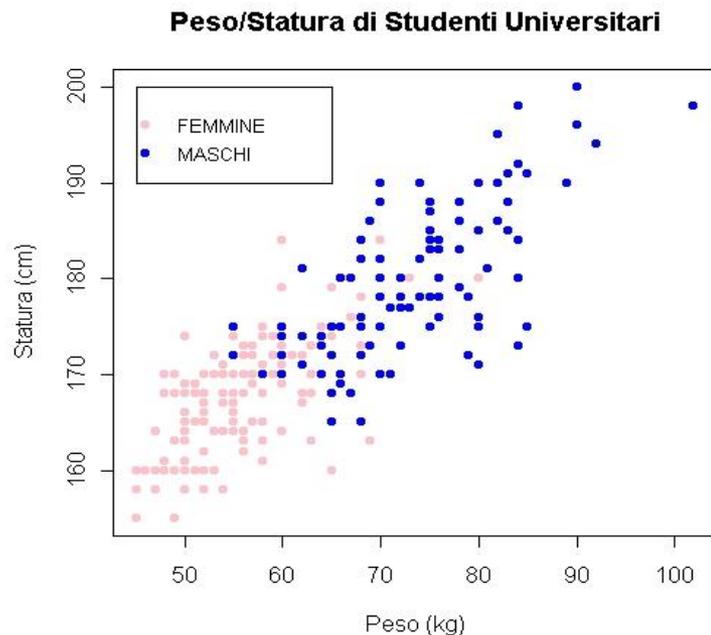
Importante inoltre è la forma assunta dalla nuvola dei punti, che può fornire indicazioni sul tipo di relazione esistente tra due variabili.

## a) Nuvola dei punti o scatter plot

(2/2)

Ad esempio:

- se i punti tendono a disporsi lungo una retta, le due variabili hanno una relazione di tipo lineare ( $Y = B_0 + B_1X$ )
- se la nuvola di punti assume una forma sferica, tra le due variabili non vi è un legame di tipo lineare.



In questo grafico si vede che:

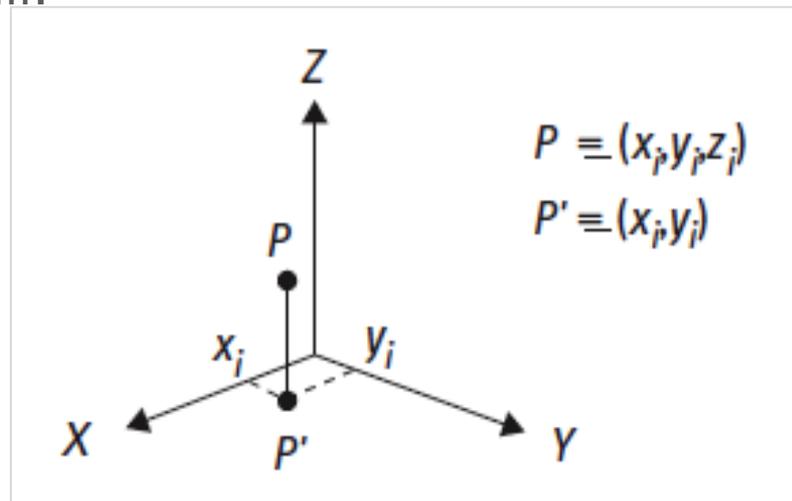
- al crescere del peso aumenta la statura;
- i valori di peso e statura delle femmine sono, in media, inferiori ai valori dei maschi.

## b) Stereogramma

(1/4)

Anche lo stereogramma è una rappresentazione grafica particolarmente indicata per **tabelle a doppia entrata di correlazione**.

Esso è costituito da una diagramma cartesiano ortogonale in uno spazio a tre dimensioni, in  $R^3$ , in cui sui tre assi (x, y e z) sono riportati rispettivamente: i valori della variabile X, i valori della variabile Y e le frequenze di associazione  $Z = f(x,y)$  delle due variabili.



## b) Stereogramma

(2/4)

Nella costruzione dello stereogramma, a seconda che le variabili siano entrambe discrete, una discreta e una continua oppure entrambe continue, occorre distinguere tra:

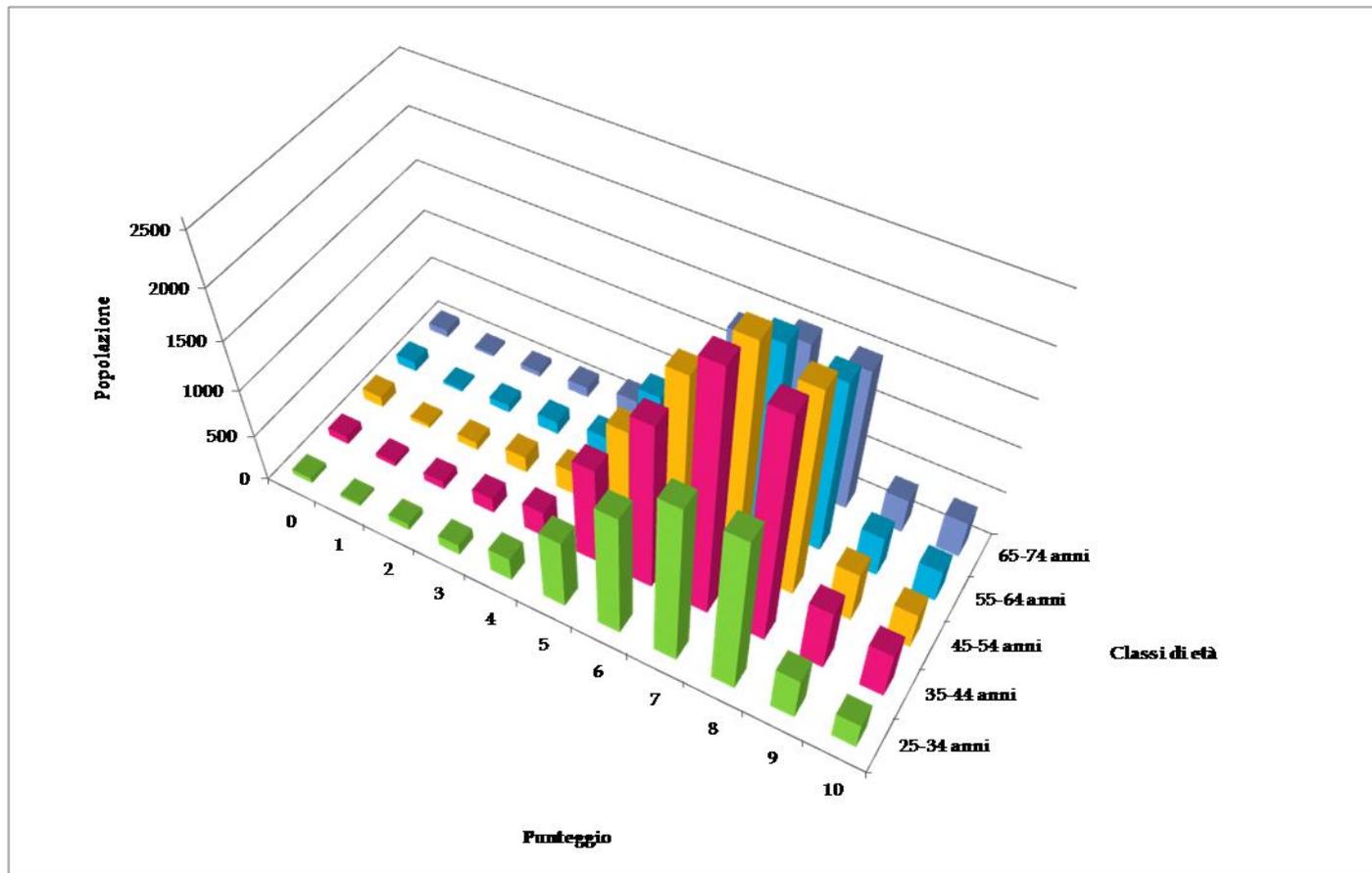
- **Stereogramma a pali**, impiegato nel caso di due variabili discrete oppure di variabili divise in classi di valori considerando il valore centrale di ciascuna classe;
- **Stereogramma a curve**, utilizzato nel caso in cui entrambe le variabili sono continue oppure una è discreta e l'altra continua.

## b) Stereogramma

(3/4)

### Stereogramma a pali

Popolazione dai 25 ai 74 anni per classe di età decennale e livello di soddisfazione della vita nel complesso (a) – Italia 2013 – (popolazione in migliaia)



a) Espresso con un punteggio da 0 (per niente soddisfatto) a 10 (molto soddisfatto)

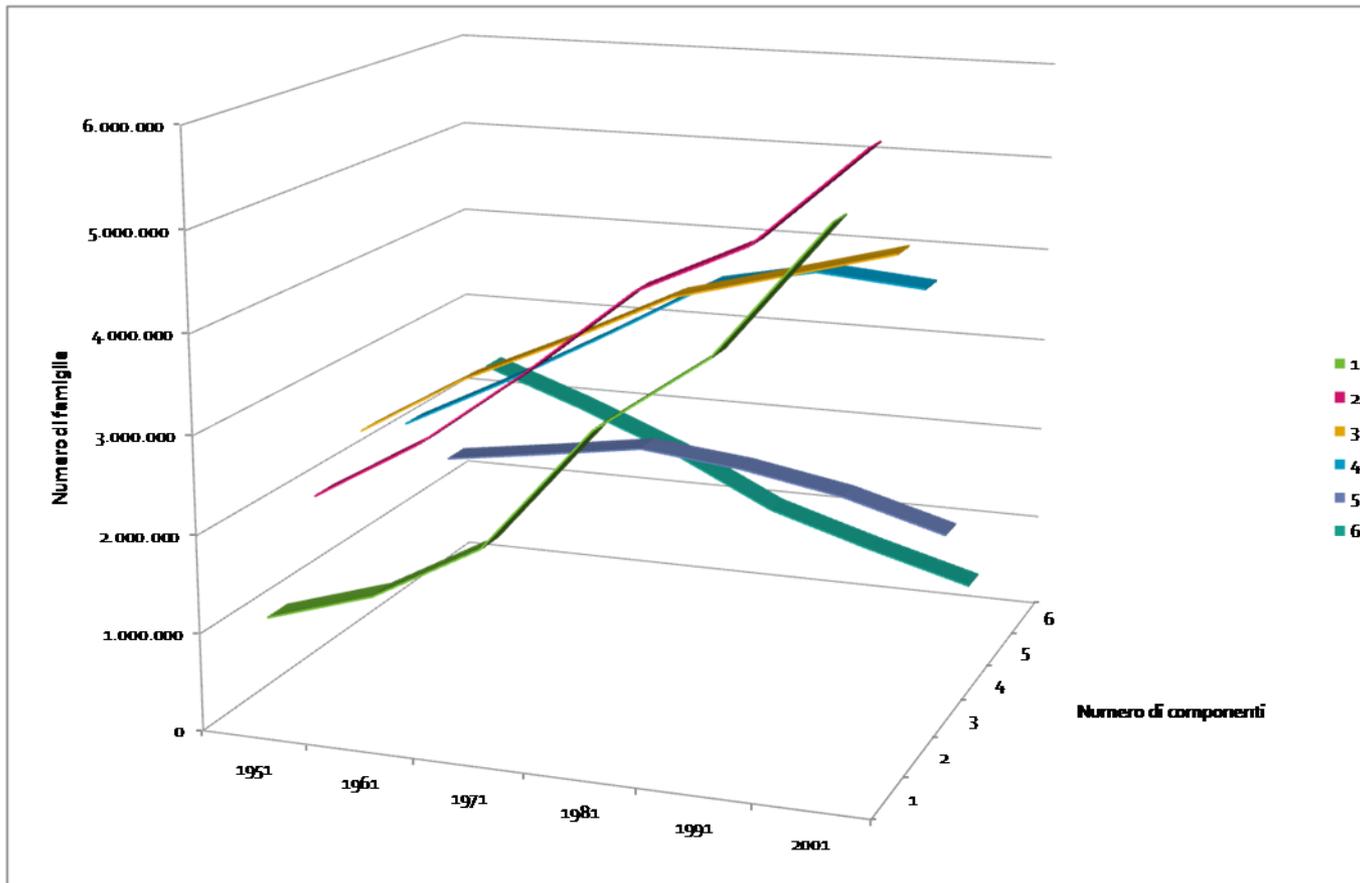
Fonte: Istat

## b) Stereogramma

(4/4)

### Stereogramma a curve

Famiglie residenti per numero di componenti per famiglia ai censimenti 1951- 2001, ai confini dell'epoca.



Fonte: Istat

## c) Diagrammi a barre e cartodiagrammi (1/4)

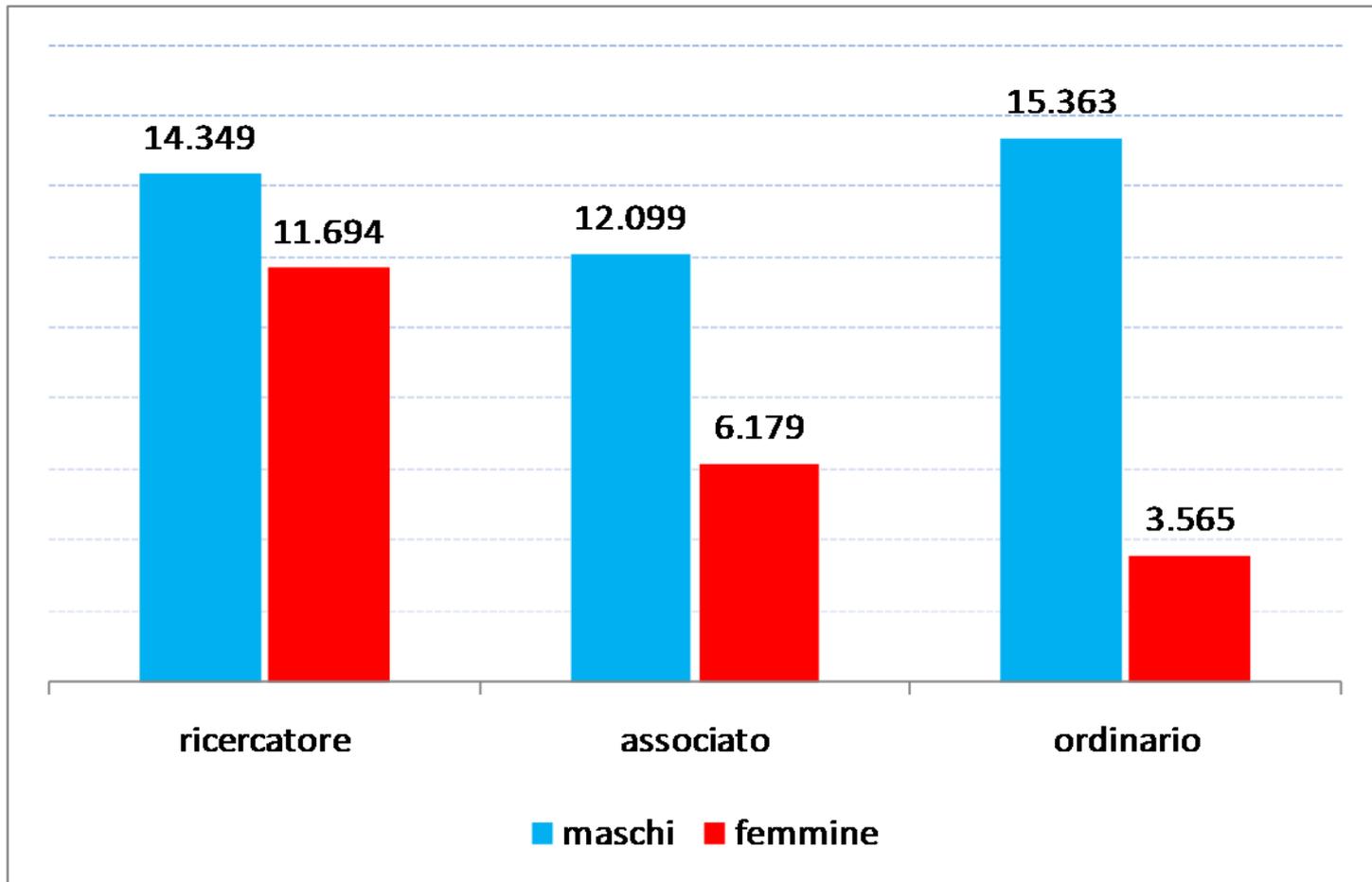
Nel caso di distribuzioni statistiche doppie in cui vi sono due mutabili o una variabile e una mutabile (**tabelle di contingenza**) esistono diverse possibili rappresentazioni grafiche, a seconda della natura dei due caratteri considerati e degli scopi della rappresentazione grafica.

Tra le rappresentazioni più impiegate vi è quella già vista per le mutabili statistiche semplici, il diagramma a barre, anche se in questo caso sono rappresentati i due caratteri e le frequenze di associazione delle modalità dei medesimi: **Diagramma a barre verticali** e **Diagramma a barre per impilamento**.

## c) Diagrammi a barre e cartodiagrammi (2/4)

### Diagramma a barre verticali

Docenti universitari per qualifica e genere – Italia – Anno 2008

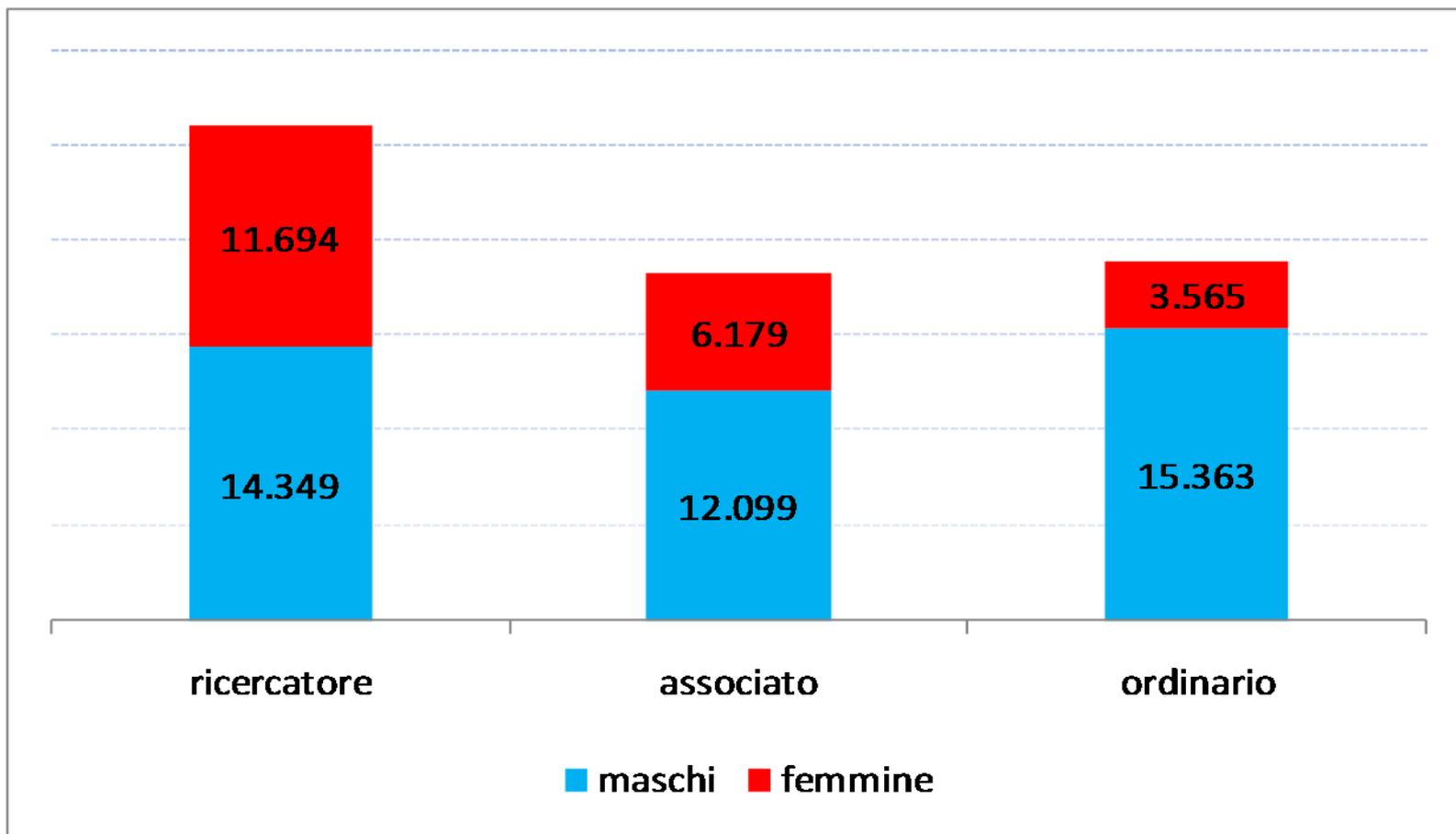


Fonte: Istat

## c) Diagrammi a barre e cartodiagrammi (3/4)

### Diagramma a barre per impilamento

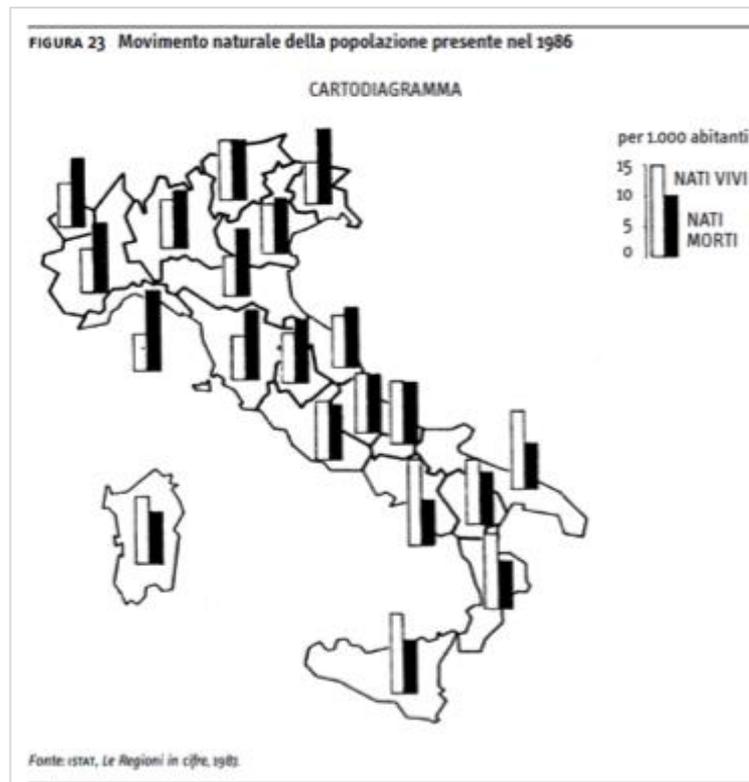
Docenti universitari per qualifica e genere – Italia – Anno 2008



## c) Diagrammi a barre e cartodiagrammi (4/4)

I cartodiagrammi non sono altro che dei cartogrammi in cui, anziché delle serie territoriali semplici, vengono rappresentate delle serie territoriali di due o più caratteri.

Esempio: I nati vivi e i morti per 1.000 abitanti nelle 20 Regioni italiane nel 1986.

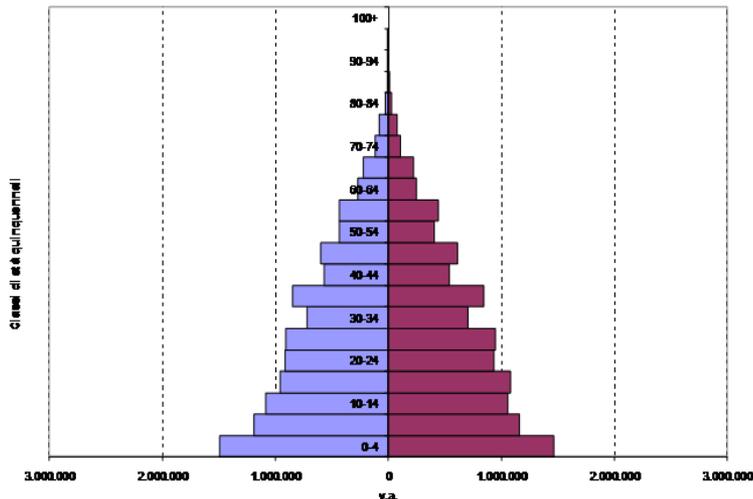


## d) Piramide delle età

(1/2)

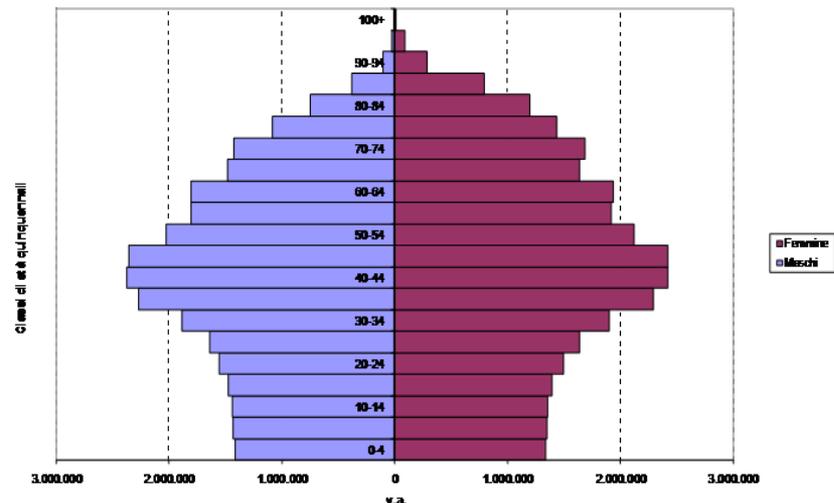
E' una particolare ed efficace rappresentazione grafica della struttura per età e sesso di un dato collettivo o popolazione. E' rappresentata da due istogrammi rovesciati, corrispondenti ai due sessi, sul cui asse verticale comune sono indicate le modalità del carattere età (in classi) e sul cui asse orizzontale sono riportate le frequenze (assolute o relative) sia dei maschi che delle femmine, corrispondenti a ciascuna classe di età considerata.

**Struttura della popolazione per età e genere, Italia - Censimento 1861**



Fonte: Istat da Ministero di agricoltura, industria e commercio

**Struttura della popolazione per età e genere, Italia - Censimento 2011**



Fonte: Istat

## d) Piramide delle età

(2/2)

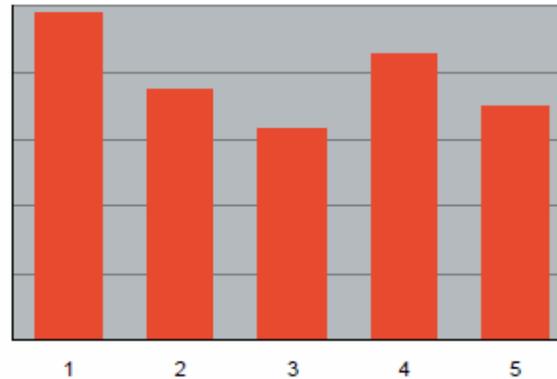
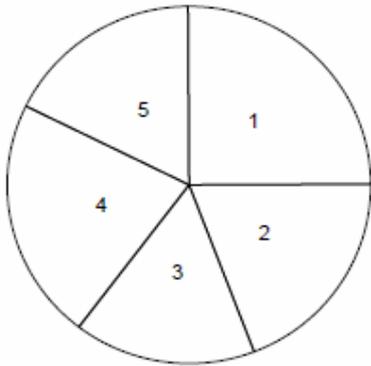
Dalla forma assunta dalla piramide è possibile trarre indicazioni sia sui fattori che caratterizzano la struttura per età e sesso attuale che sull'evoluzione passata nonché previsioni per un arco di tempo non superiore a un secolo.

Tali indicazioni possono essere tratte analizzando, in particolare:

- La **base**, che fornisce indicazioni circa il flusso delle nascite. Se è molto larga, si ha un flusso di nascite in forte aumento; se è sufficientemente larga, le nascite sono costanti o in lieve aumento; se è stretta significa che il flusso delle nascite è in diminuzione.
- L'**inclinazione dei lati**, che fornisce indicazioni circa il livello generale di eliminazione per morte. Se l'obliquità dei lati è forte, si ha un'alta mortalità; se è debole, si ha una bassa mortalità.
- La presenza di **rigonfiamenti o strozzature** per particolari classi d'età, che fornisce indicazione dell'intervento di particolari fattori di perturbazione (guerre, epidemie etc.).

# La scelta della rappresentazione grafica 1/2

Questi 2 grafici rappresentano la stessa distribuzione.  
Qual è più chiaro?



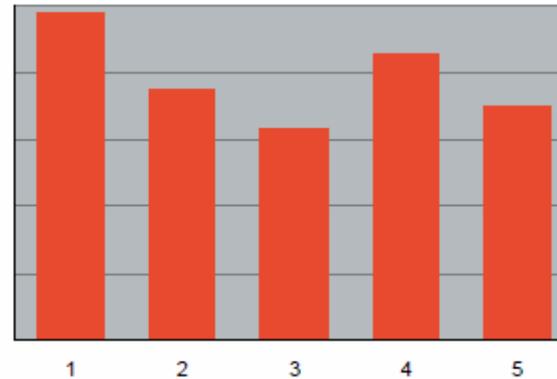
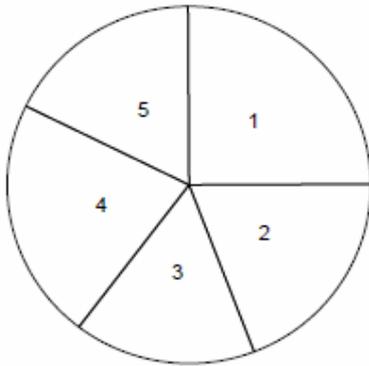
Fonte: UNECE, Making Data Meaningful Part 2: A guide to presenting statistics, United Nations, Geneva 2009.

Quale settore del diagramma circolare è maggiore?

## La scelta della rappresentazione grafica 2/2

Per la maggior parte delle persone è più facile confrontare segmenti piuttosto che angoli.

Nel diagramma circolare i settori numero 1 e 4 sembrano identici, mentre nel diagramma a barre è evidente la differenza



Fonte: UNECE, Making Data Meaningful Part 2: A guide to presenting statistics, United Nations, Geneva 2009.

E' opportuno rappresentare la stessa distribuzione con più grafici per individuare quello che meglio rappresenta il messaggio che si vuole veicolare.

## Bibliografia

- Leti G., *Statistica Descrittiva*, Il Mulino, Bologna, 1983
- UNECE, *Making Data Meaningful Part 2: A guide to presenting statistics*, United nations, Geneva 2009.
- <http://www3.istat.it/servizi/studenti/valoredati>

...e adesso...  
buon lavoro!

Rete per la promozione  
della cultura statistica

Per ulteriori moduli didattici relativi alle scuole secondarie di secondo grado si prega di consultare anche la piattaforma Scuola di statistica – Lab (accessibile dal link <http://scuoladistatistica-lab.istat.it/>)