

Incontri di promozione della cultura statistica
Iniziativa Utenti
Aprile, 2024

Rappresentazioni grafiche

Monica Bailot

DCCI - CIC - Servizio di coordinamento della promozione della cultura statistica e della diffusione dei dati a livello territoriale

Indice

- **Funzione**
- **Elementi essenziali**
- **Tipologia più adatta**
- **Grafici per esempi**

Le rappresentazioni grafiche

Le rappresentazioni grafiche fanno parte della statistica descrittiva e consentono di **trasformare le informazioni contenute nelle tabelle di dati**, mettendone in luce le caratteristiche salienti, **in modo intuitivo**.



Comunicare con un grafico

Il grafico è una **rappresentazione schematica di un fenomeno mediante un disegno convenzionale**

Permette di:

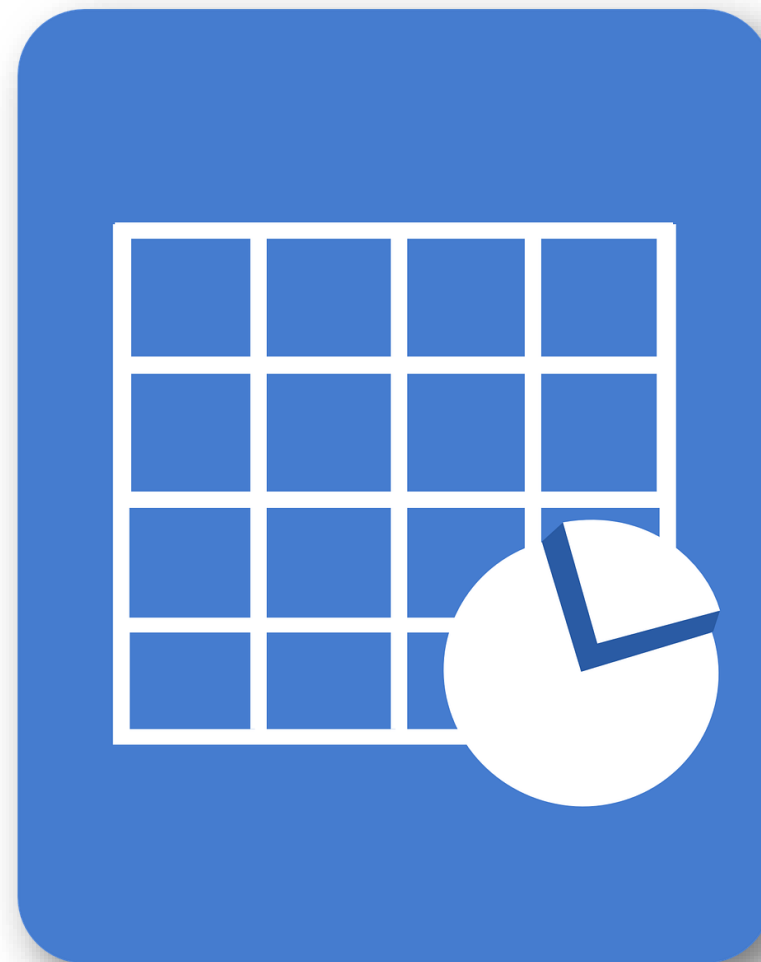
- fare ulteriore **sintesi dei dati**
- trarre **conclusioni più rapidamente**
- scoprire **informazioni nascoste** o poco evidenti
- rendere i **dati accessibili**



Comunicare con un grafico

Sebbene la rappresentazione grafica renda il dato più intuitivo ha lo svantaggio di mettere in evidenza solamente un aspetto del fenomeno di interesse.

Per questo motivo, tabelle e grafici sono spesso utilizzati in modo associato.



Gli elementi essenziali

Le rappresentazioni grafiche sono sempre corredate dai **metadati**

Il titolo descrive:

- il dato statistico rappresentato e il suo significato
- l'unità di misura in cui è espresso
- il periodo di riferimento

La fonte riporta:

Il soggetto che ha prodotto il dato, la pubblicazione/link dove è possibile consultarlo

La legenda e le etichette spiegano:

il significato di colori e simboli presenti nell'immagine



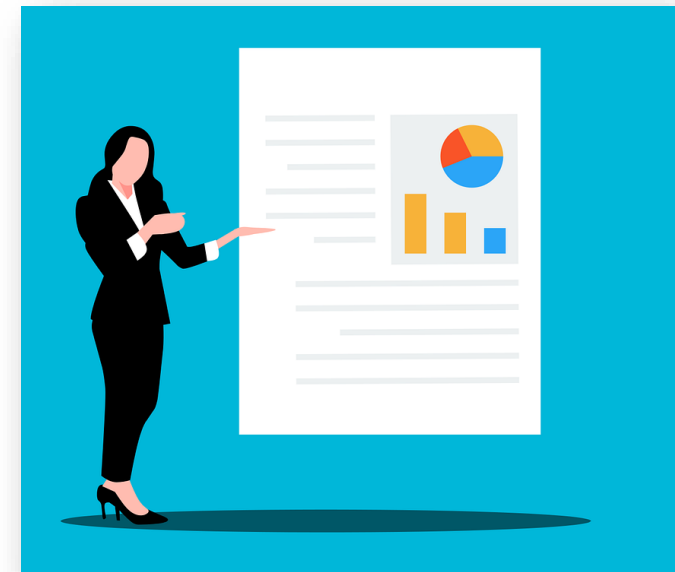
La tipologia più adatta

La tipologia di grafico più adatta al dato osservato **dipende** dal **tipo** di:

- **variabile** (qualitativa o quantitativa)
- **distribuzione** (univariata o bivariata)

E da quale **informazione** si desidera mettere **in evidenza**

Nella presentazione dei dati, per aiutare la lettura, è preferibile **definire un formato standard** (colori, carattere, dimensioni ecc.). Ad es. utilizzare sempre lo stesso colore per una determinata modalità.



La tipologia più adatta

Per scegliere come **presentare i dati**, ad esempio con una tabella, un grafico o una mappa, occorre **effettuare delle valutazioni**.

Quali dati si desidera visualizzare?

Chi sono i destinatari?

Quale tipo di presentazione sarà più utile?

Spesso nelle pubblicazioni si sceglie di **combinare tabelle, grafici e mappe** posizionando eventualmente i dati in tabella accanto all'immagine.



Alcune tipologie di rappresentazioni grafiche

- Areogramma (grafico a torta) o ad Anello
- Grafico a barre o nastri
- Diagramma a barre affiancate o cumulate
- Diagramma a bastoncino/canne d'organo
- Istogramma
- Diagramma a linee
- Cartogramma/Mappa
- *Dispersione/Scatter*
- *Box-plot*

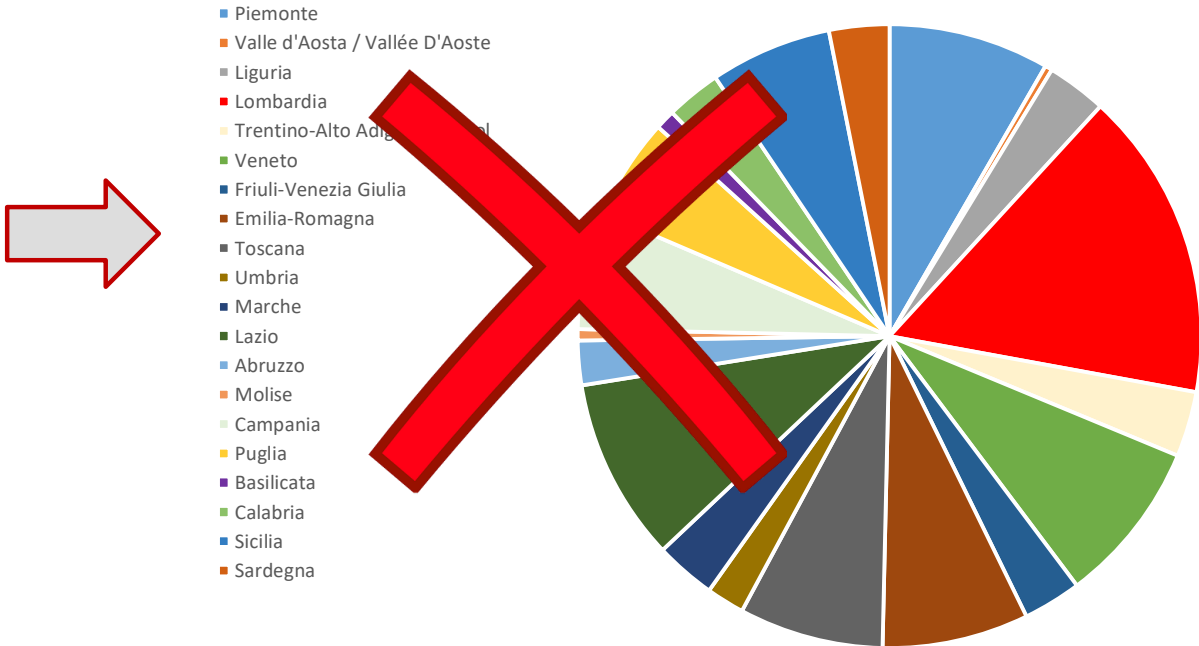


Areogramma e grafico ad anello

Istituzioni non profit e dipendenti per regione. Anno 2021 (valori assoluti e composizioni %) (dati definitivi)

Territorio	Valori assoluti		Composizioni %	
	Istituzioni non profit	Dipendenti	Istituzioni non profit	Dipendenti
Piemonte	30.049	73.666	8,3%	8,2%
Valle d'Aosta / Vallée D'Aoste	1.362	2.030	0,4%	0,2%
Liguria	11.199	22.762	3,1%	2,5%
Lombardia	57.927	197.563	16,1%	22,1%
Trentino-Alto Adige / Südtirol	12.201	24.933	3,4%	2,8%
- Bolzano	5.730	10.714	1,6%	1,2%
- Trento	6.471	14.219	1,8%	1,6%
Veneto	30.597	80.941	8,5%	9,1%
Friuli-Venezia Giulia	10.902	20.807	3,0%	2,3%
Emilia-Romagna	27.375	84.867	7,6%	9,5%
Toscana	27.005	55.173	7,5%	6,2%
Umbria	7.129	12.276	2,0%	1,4%
Marche	11.357	19.119	3,1%	2,1%
Lazio	34.266	114.546	9,5%	12,8%
Abruzzo	8.298	11.952	2,3%	1,3%
Molise	2.109	3.604	0,6%	0,4%
Campania	21.911	37.916	6,1%	4,2%
Puglia	18.966	42.266	5,3%	4,7%
Basilicata	3.687	6.422	1,0%	0,7%
Calabria	10.266	11.618	2,8%	1,3%
Sicilia	22.747	47.070	6,3%	5,3%
Sardegna	11.272	24.210	3,1%	2,7%
Italia	360.625	893.741	100%	100%

Istituzioni non profit per regione



Troppe modalità, il grafico non è leggibile

Fonte: Censimento permanente delle istituzioni non profit

Areogramma e grafico ad anello

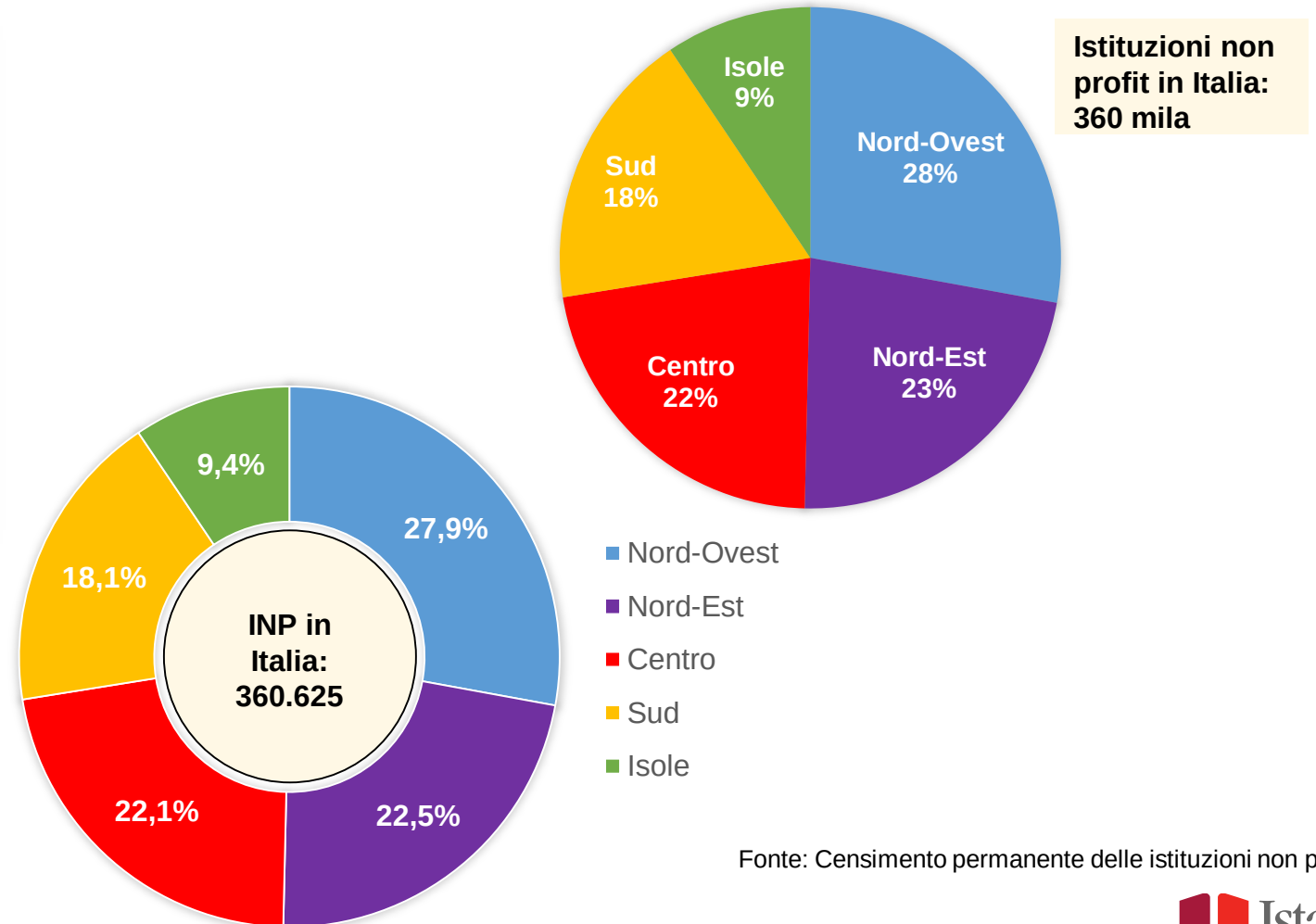
Istituzioni non profit e dipendenti per ripartizione geografica. Anno 2021 (valori assoluti) (dati definitivi)

Ripartizioni	Istituzioni non profit	Dipendenti
Nord-Ovest	100.537	296.021
Nord-Est	81.075	211.548
Centro	79.757	201.114
Sud	65.237	113.778
Isole	34.019	71.280
Italia	360.625	893.741

Fonte: Censimento permanente delle istituzioni non profit

- Si utilizzano solitamente per
- variabili qualitative con poche modalità
 - composizioni %

Istituzioni non profit per ripartizione geografica. Anno 2021 (composizione %)



Fonte: Censimento permanente delle istituzioni non profit

Areogramma e grafico ad anello: breve descrizione

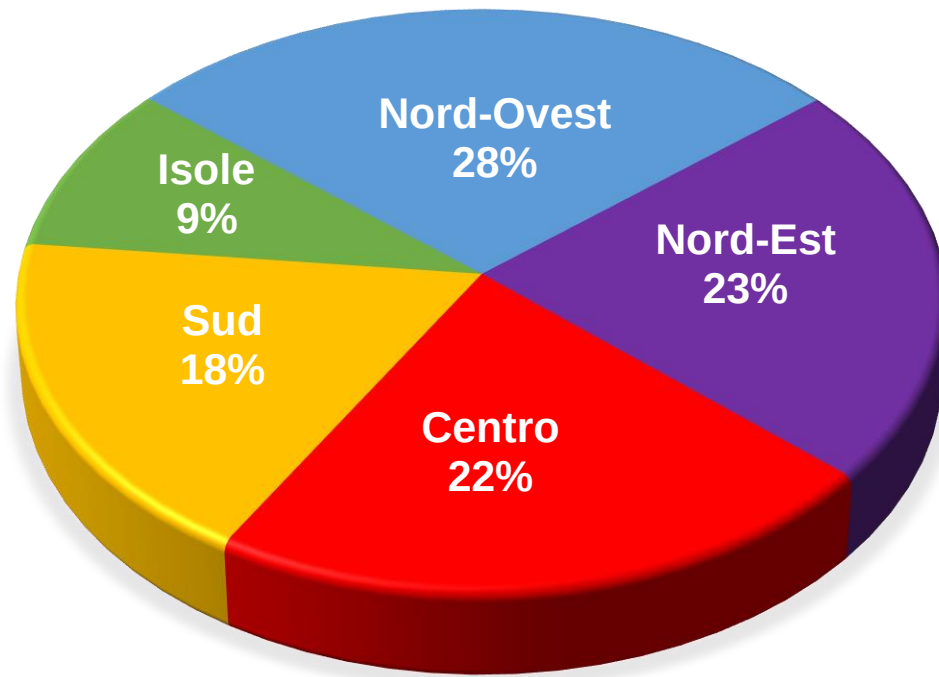
Quando utilizzare il grafico a torta/anello?

Per rappresentare la distribuzione di un determinato fenomeno tra 100 unità.

È un grafico con il quale si descrivono i dati su di un'area circolare

- Le modalità sono le parti di un cerchio (anello) o «fette» della torta
- I settori circolari hanno una dimensione proporzionale alla frequenza
- La somma dei valori delle «fette» è uguale al totale delle unità osservate/stimate
- L'immagine è sempre accompagnata da un'etichetta che specifica il totale delle unità osservate/stimate

La torta 3D



Le rappresentazioni grafiche tridimensionali sono esteticamente accattivanti... ma...

Aggiungono una dimensione che nei dati non esiste e falsano la lettura dell'informazione statistica.

E' preferibile non utilizzarle

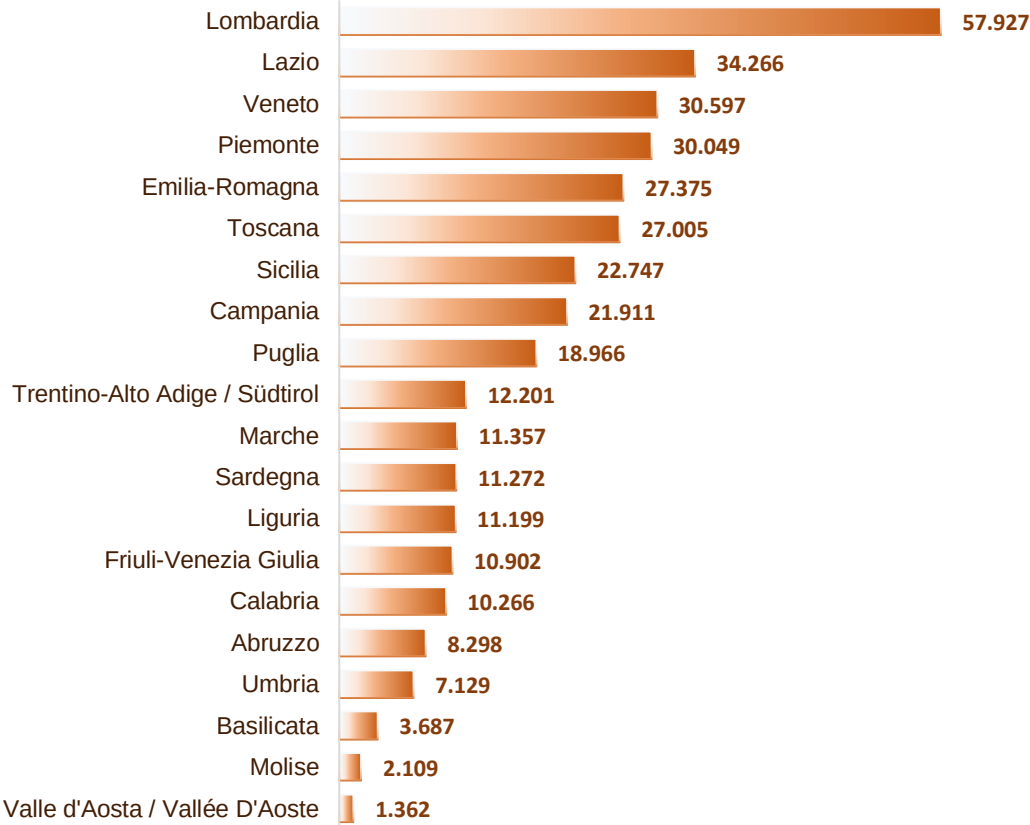
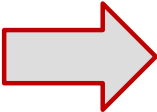
Diagramma a barre e/o nastri

Istituzioni non profit e dipendenti per regione. Anno 2021 (valori assoluti e composizioni %) (dati definitivi)

Territorio	Valori assoluti		Composizioni %	
	Istituzioni non profit	Dipendenti	Istituzioni non profit	Dipendenti
Piemonte	30.049	73.666	8,3%	8,2%
Valle d'Aosta / Vallée D'Aoste	1.362	2.030	0,4%	0,2%
Liguria	11.199	22.762	3,1%	2,5%
Lombardia	57.927	197.563	16,1%	22,1%
Trentino-Alto Adige / Südtirol	12.201	24.933	3,4%	2,8%
- Bolzano	5.730	10.714	1,6%	1,2%
- Trento	6.471	14.219	1,8%	1,6%
Veneto	30.597	80.941	8,5%	9,1%
Friuli-Venezia Giulia	10.902	20.807	3,0%	2,3%
Emilia-Romagna	27.375	84.867	7,6%	9,5%
Toscana	27.005	55.173	7,5%	6,2%
Umbria	7.129	12.276	2,0%	1,4%
Marche	11.357	19.119	3,1%	2,1%
Lazio	34.266	114.546	9,5%	12,8%
Abruzzo	8.298	11.952	2,3%	1,3%
Molise	2.109	3.604	0,6%	0,4%
Campania	21.911	37.916	6,1%	4,2%
Puglia	18.966	42.266	5,3%	4,7%
Basilicata	3.687	6.422	1,0%	0,7%
Calabria	10.266	11.618	2,8%	1,3%
Sicilia	22.747	47.070	6,3%	5,3%
Sardegna	11.272	24.210	3,1%	2,7%
Italia	360.625	893.741	100%	100%

Fonte: Censimento permanente delle istituzioni non profit

Istituzioni non profit per regione. Anno 2021 (valori assoluti)



Fonte: Censimento permanente delle istituzioni non profit

Diagramma a barre e/o nastri: breve descrizione

E' utile per rappresentare distribuzioni semplici di variabili **qualitative** e **quantitative discrete**.

I dati sono descritti con dei **rettangoli verticali o orizzontali** (nastri)

- Le barre hanno tutte la stessa base e sono equidistanti tra loro, mai contigue
- L'altezza (lunghezza) delle barre è proporzionale alla frequenza assoluta/percentuale osservata
- Un asse riporta le modalità, l'altro la metrica (valore del dato)
- Se le descrizioni delle modalità sono lunghe, il diagramma a nastri è preferibile
- Spesso le modalità sono presentate in graduatoria secondo le frequenze assolute/percentuali

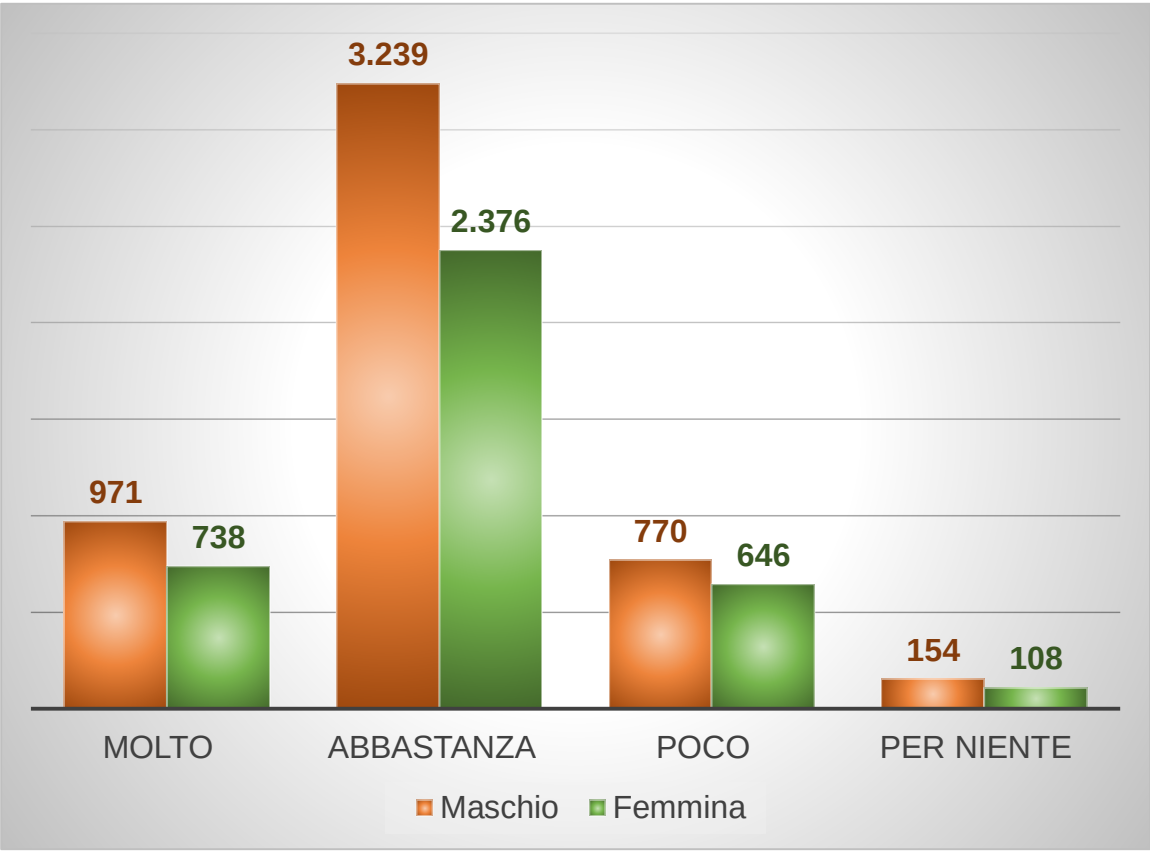
Diagramma a barre affiancate o cumulate per tabelle di contingenza

Persone tra 25 e 44 anni secondo la soddisfazione per il lavoro e genere. Italia. Anno 2022 (valori in migliaia)

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	971	738	1.709
abbastanza	3.239	2.376	5.615
poco	770	646	1.416
per niente	154	108	262
Totale	5.134	3.868	9.002

Fonte: Indagine multiscopo «Aspetti della vita quotidiana»

Per una tabella a doppia entrata possiamo utilizzare il diagramma a barre affiancate



Fonte: Indagine multiscopo «Aspetti della vita quotidiana»

Diagramma a barre affiancate o cumulate per tabelle di contingenza

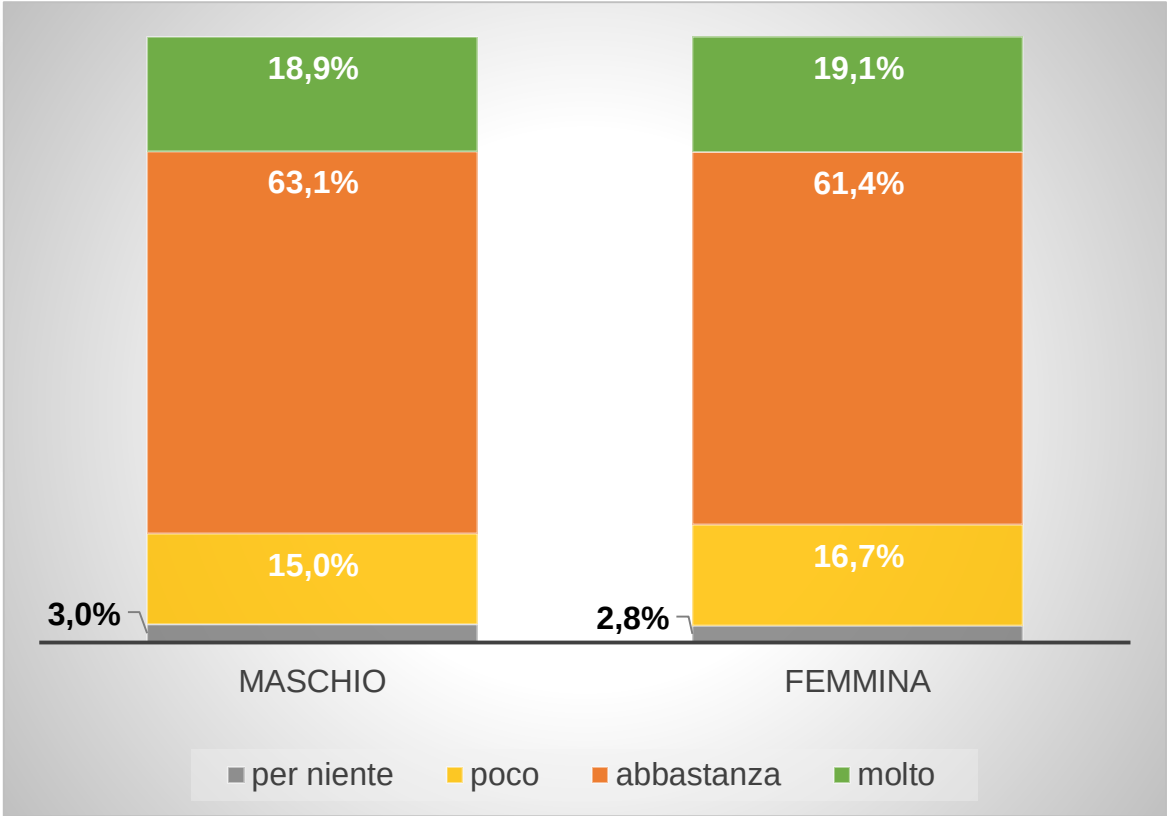
Persone tra 25 e 44 anni secondo la soddisfazione per il lavoro e genere. Italia. Anno 2022 (valori in migliaia)

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	971	738	1.709
abbastanza	3.239	2.376	5.615
poco	770	646	1.416
per niente	154	108	262
Totale	5.134	3.868	9.002

Fonte: Indagine multiscopo «Aspetti della vita quotidiana»

Oppure un diagramma a barre cumulate (in pila)

Persone tra 25 e 44 anni secondo la soddisfazione per il lavoro e genere. Italia. Anno 2022 (composizioni %)



Fonte: Indagine multiscopo «Aspetti della vita quotidiana»

Diagramma a barre affiancate o cumulate per tabelle di contingenza

La distribuzione bivariata di variabili **qualitative** può essere descritta con una **successione di rettangoli verticali o orizzontali**

Questa tipologia di grafico:

- Consente il confronto tra popolazioni differenti (es. maschio/femmina)
- Se cumulate, le barre sono suddivise in tanti rettangoli quante sono le modalità della seconda variabile di interesse
- Si utilizza se le modalità della seconda variabile non sono numerose (più di 5)

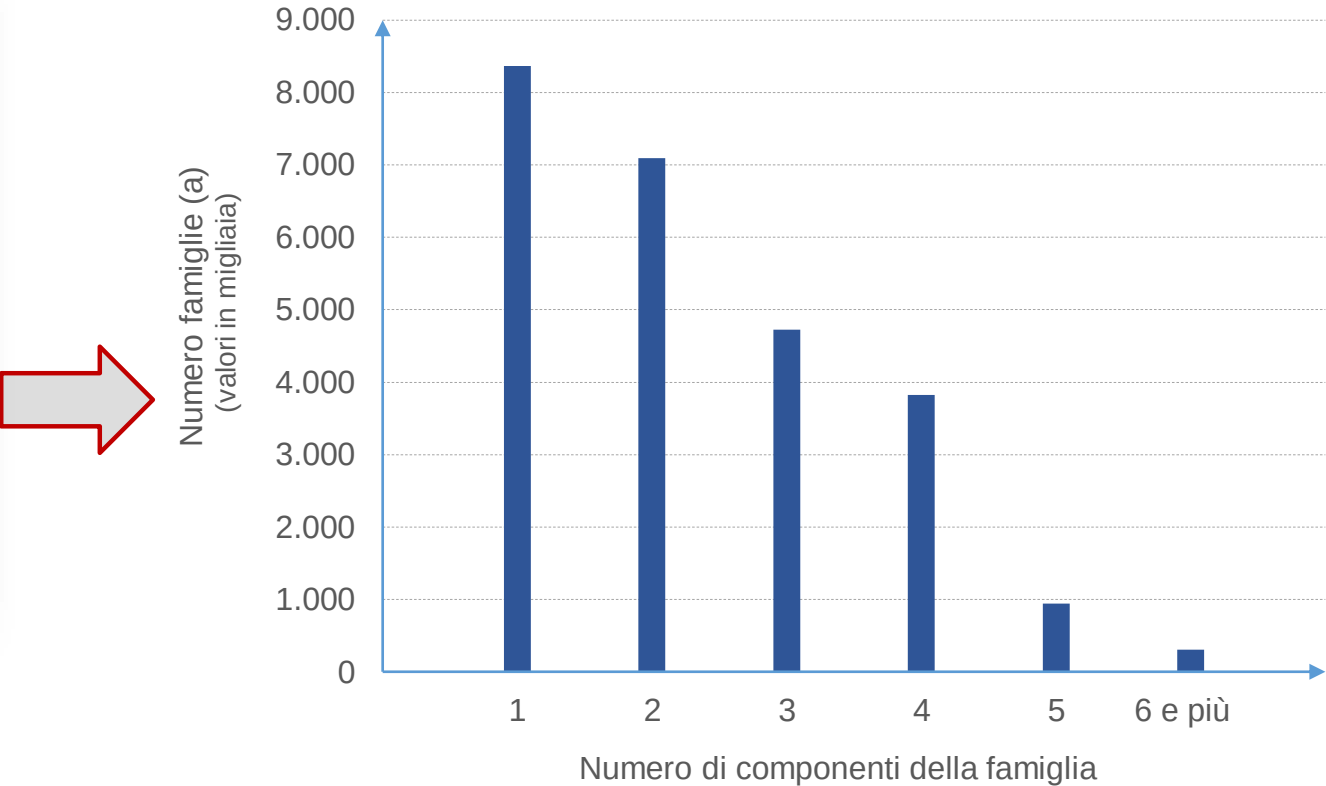
Valutare di volta in volta quale presentazione è più informativa, tra le barre affiancate o cumulate.

Diagramma a bastoncino o a canne d'organo

Famiglie per numero di componenti in Italia. Media biennale 2021-2022 (valori assoluti in migliaia)

Numero di componenti della famiglia	Famiglie (media biennale) (a)
1	8.362
2	7.092
3	4.723
4	3.827
5	946
6 e più	313
Totale	25.263

Fonte: Censimento permanente della popolazione e delle abitazioni
(a) Media biennale calcolata sui dati dell'anno corrente e quello che lo precede



Fonte: Censimento permanente della popolazione e delle abitazioni
(a) Media biennale calcolata sui dati dell'anno corrente e quello che lo precede

Diagramma a bastoncino o a canne d'organo

È utile per rappresentare la distribuzione semplice di frequenze di una variabile **quantitativa discreta**.

È un grafico che si basa su un riferimento cartesiano.

- L'asse orizzontale (ascisse) riporta i valori della variabile di interesse con la sua unità di misura
- L'asse verticale (ordinate) riporta le frequenze assolute/percentuali
- L'altezza del segmento associato ad ogni modalità è proporzionale alla frequenza osservata
- La distanza tra i bastoncini dipende dalle modalità, non è arbitraria

Istogramma

Istituzioni non profit con volontari per forma giuridica e classe di volontari. Italia.
Anno 2021 (valori assoluti) (dati definitivi)

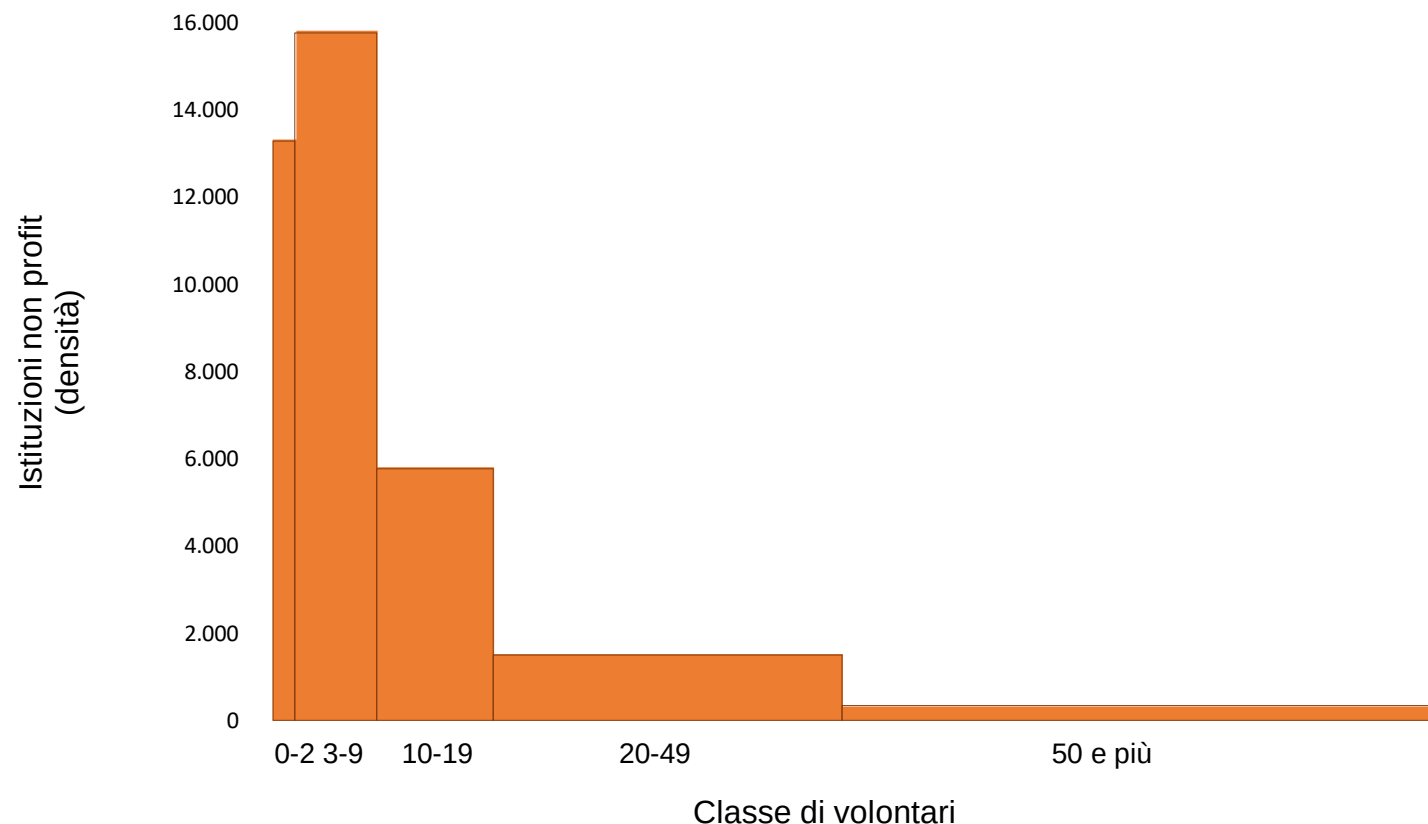
Classi di volontari	Forma giuridica				Totale istituzioni non profit con volontari (a)
	Associazione	Cooperativa sociale	Fondazione	Altra forma Giuridica	
1-2	20.369	3.416	690	2.158	26.633
3-9	99.833	2.745	2.862	5.209	110.649
10-19	53.856	775	567	2.847	58.046
20-49	41.879	297	389	2.954	45.519
50 e più	14.280	97	154	1.904	16.435
Totale (a)	230.217	7.331	4.663	15.072	257.282

Fonte: Censimento permanente delle istituzioni non profit
(a) Eventuali differenze tra le somme ed il relativo totale sono da attribuire esclusivamente ad arrotondamenti nei decimali

In questa tabella doppia la variabile in fiancata è di tipo quantitativo mentre quella in testata è qualitativa

Istogramma

Istituzioni non profit con volontari per classe di volontari. Italia. Anno 2021 (valori assoluti) (dati definitivi)



L'area di ogni rettangolo è pari al numero di istituzioni non profit con volontari della classe

Area totale = 257.282

$$\text{Densità} = \frac{\text{frequenza}}{\text{ampiezza della classe}}$$

$$\begin{aligned} \text{Densità della classe 10-19} \\ &= \frac{58.046}{10} \end{aligned}$$

Fonte: Censimento permanente delle istituzioni non profit

(a) Eventuali differenze tra le somme ed il relativo totale sono da attribuire esclusivamente ad arrotondamenti nei decimali

Istogramma

È appropriato per una variabile **quantitativa continua** raggruppata in classi.

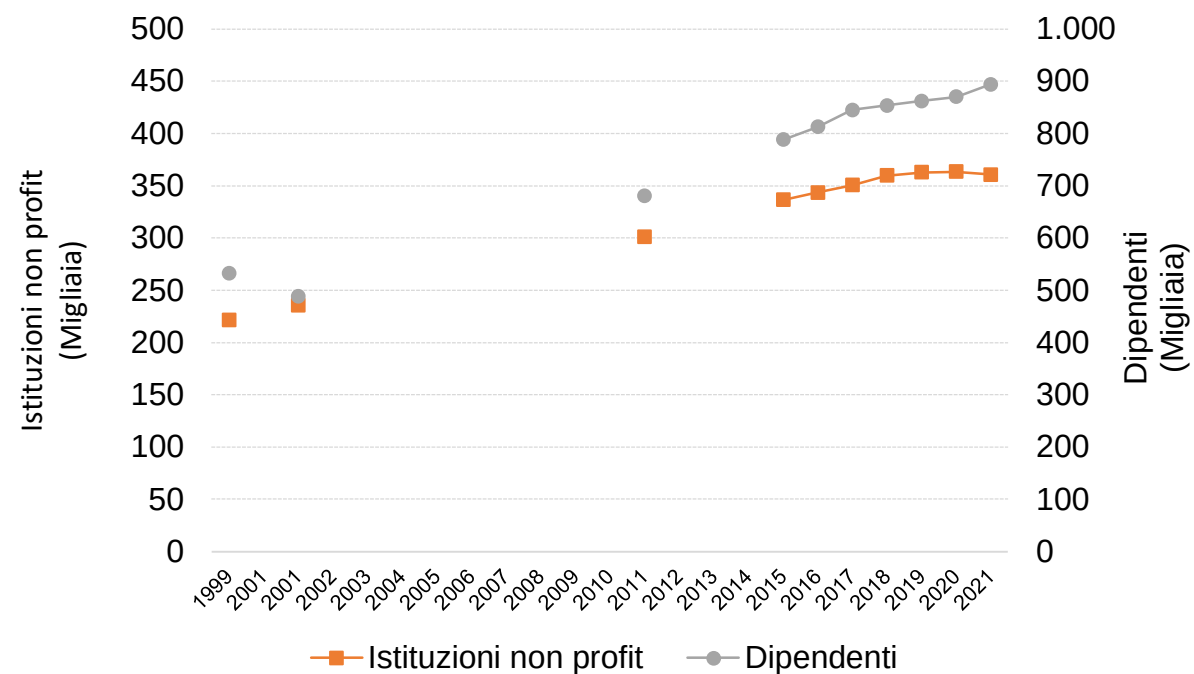
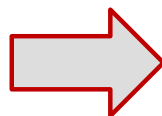
La rappresentazione si basa su un riferimento cartesiano.

- L'asse orizzontale (ascisse) riporta i valori della variabile di interesse con la sua unità di misura
- L'asse verticale (ordinate) riporta le densità di frequenza
- I rettangoli sono sempre contigui
- La loro area è proporzionale alla frequenza osservata
- Le classi possono essere di ampiezza differente

Diagramma a linee

Istituzioni non profit e dipendenti. Anni 1999, 2001, 2011, 2015-2021 (valori assoluti)

Anno	Istituzioni non profit	Dipendenti
1999	221.412	531.926
2001	235.232	488.523
2011	301.191	680.811
2015	336.275	788.126
2016	343.432	812.706
2017	350.492	844.775
2018	359.574	853.476
2019	362.634	861.919
2020	363.499	870.183
2021	360.625	893.741



Fonte: Registro statistico delle istituzioni non profit e Censimenti

Fonte: Registro statistico delle istituzioni non profit e Censimenti

Diagramma a linee

E' idoneo a descrivere l'evoluzione di un fenomeno nel tempo, si utilizza per le serie storiche di **variabili quantitative**.

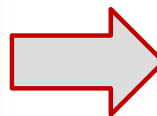
È un grafico che si basa su un riferimento cartesiano.

- Ogni punto dell'asse orizzontale è un istante del tempo considerato (giorni, mesi, anni)
- L'asse verticale riporta la misura del fenomeno osservato (valore assoluto, un indicatore ecc.)
- Se i punti sono congiunti da segmenti, si ottiene una spezzata (si ipotizza un andamento lineare tra un istante e il successivo)

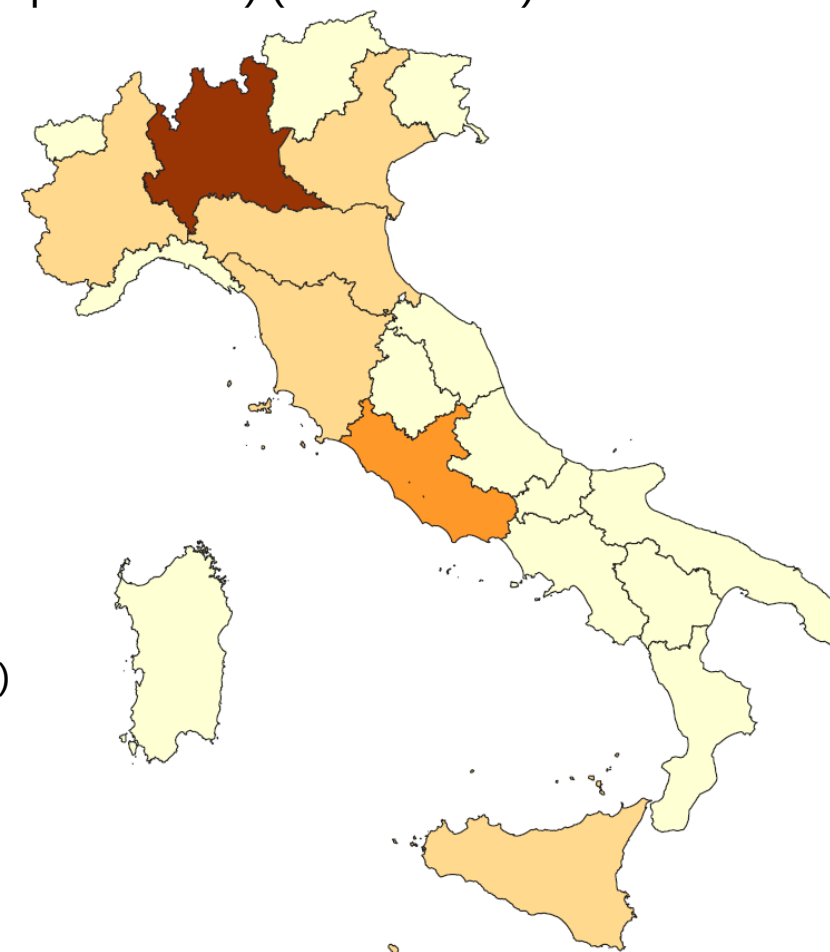
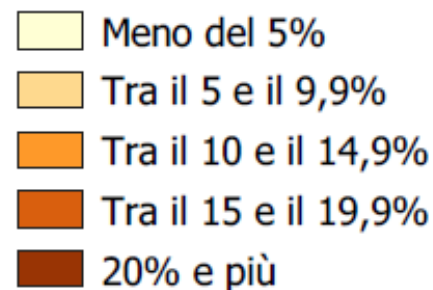
Cartogramma

Istituzioni non profit e dipendenti per regione. Anno 2021 (valori assoluti e composizioni %) (dati definitivi)

Territorio	Valori assoluti		Composizioni %	
	Istituzioni non profit	Dipendenti	Istituzioni non profit	Dipendenti
Piemonte	30.049	73.666	8,3%	8,2%
Valle d'Aosta / Vallée D'Aoste	1.362	2.030	0,4%	0,2%
Liguria	11.199	22.762	3,1%	2,5%
Lombardia	57.927	197.563	16,1%	22,1%
Trentino-Alto Adige / Südtirol	12.201	24.933	3,4%	2,8%
- Bolzano	5.730	10.714	1,6%	1,2%
- Trento	6.471	14.219	1,8%	1,6%
Veneto	30.597	80.941	8,5%	9,1%
Friuli-Venezia Giulia	10.902	20.807	3,0%	2,3%
Emilia-Romagna	27.375	84.867	7,6%	9,5%
Toscana	27.005	55.173	7,5%	6,2%
Umbria	7.129	12.276	2,0%	1,4%
Marche	11.357	19.119	3,1%	2,1%
Lazio	34.266	114.546	9,5%	12,8%
Abruzzo	8.298	11.952	2,3%	1,3%
Molise	2.109	3.604	0,6%	0,4%
Campania	21.911	37.916	6,1%	4,2%
Puglia	18.966	42.266	5,3%	4,7%
Basilicata	3.687	6.422	1,0%	0,7%
Calabria	10.266	11.618	2,8%	1,3%
Sicilia	22.747	47.070	6,3%	5,3%
Sardegna	11.272	24.210	3,1%	2,7%
Italia	360.625	893.741	100%	100%



Dipendenti delle INP per regione (composizione %)



Cartogramma

E' indicato per confrontare **variabili quantitative nello spazio** e utilizza i metodi della cartografia per presentare i dati associati a specifici punti, linee o superfici.

Caratteristiche:

- Simultaneità tra spazio e valore
- Focus sulle relazioni tra indicatori
- Differenze e analogie tra aree
- Dati raggruppati in classi secondo la legenda
- Solitamente privo di etichette (visualizzazione approssimata)

Il più comune è la mappa coropletica (basata su differenti tonalità di colore)

È pensato per essere "visualizzato" piuttosto che "letto"

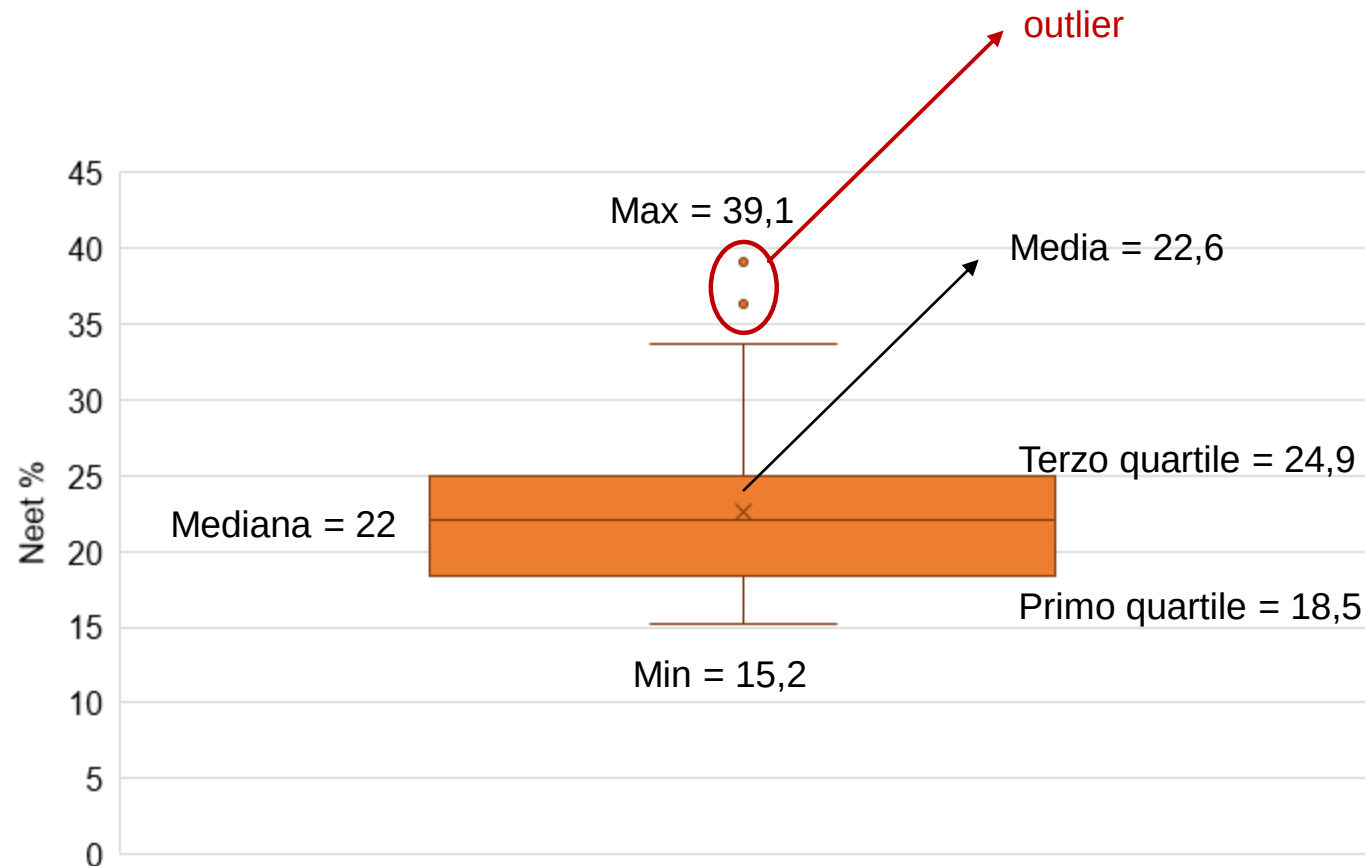
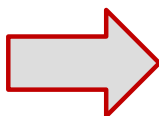
Box plot

Percentuale di Neet e tasso di disoccupazione per provincia. Anno 2019

Provincia	% Neet	tasso disoccupazione
Alessandria	23,9	11,0
Asti	21,5	9,7
Biella	22,8	9,2
Cuneo	18,5	6,6
Novara	22,5	10,3
Torino	22,3	10,8
Verbano-Cusio-Ossola	24,9	11,4
Vercelli	22,0	10,2

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Più è alto il rettangolo (box) maggiore è la variabilità del fenomeno



Box plot

Consente di rappresentare la **variabilità** di una **variabile quantitativa** e di individuarne i valori anomali.

Caratteristiche:

- Un simbolo che indica la media ($\bar{x}$)
- Una linea che indica la mediana
- Un rettangolo la cui altezza misura lo scarto tra il terzo e il primo quartile
- Gli estremi dei valori osservati (minimo e massimo)
- Eventuali outlier (valori anomali) rispetto alla distribuzione osservata

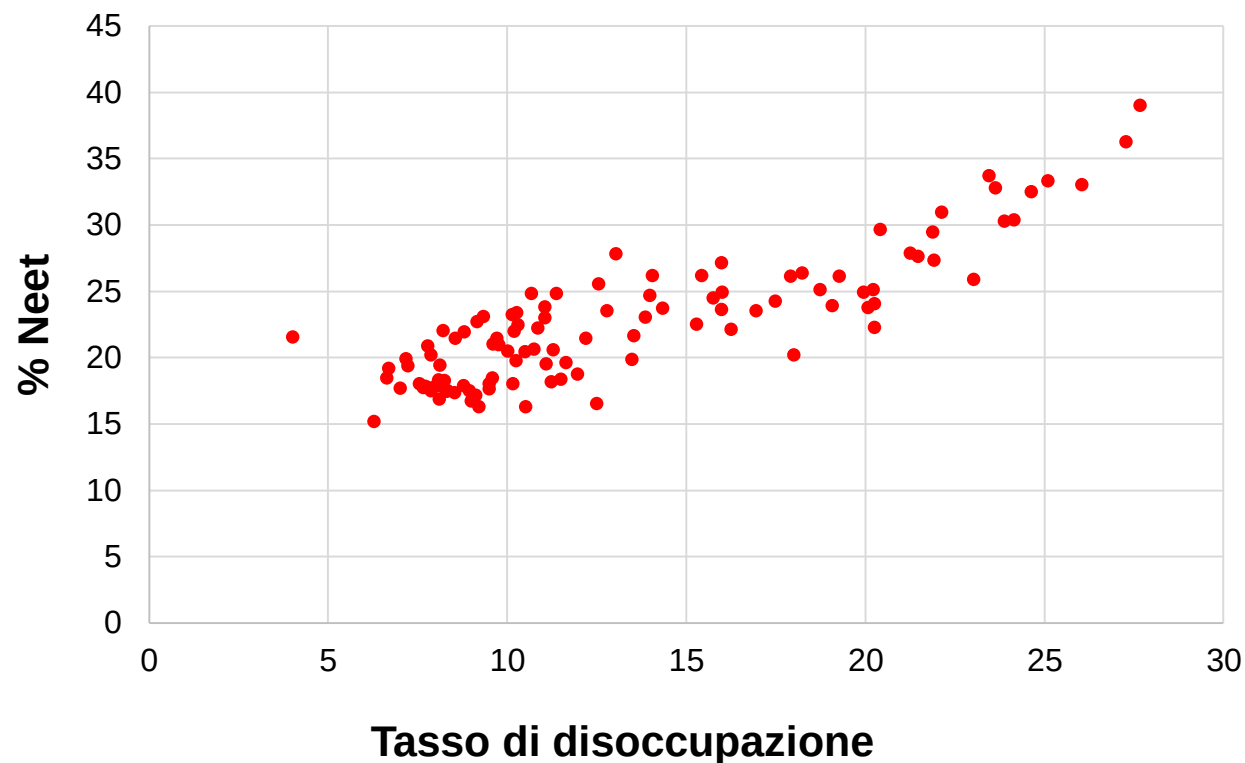
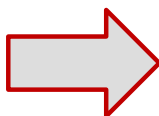
Grafico a dispersione

Percentuale di Neet e tasso di disoccupazione per provincia. Anno 2019

Provincia	% Neet	tasso disoccupazione
Alessandria	23,9	11,0
Asti	21,5	9,7
Biella	22,8	9,2
Cuneo	18,5	6,6
Novara	22,5	10,3
Torino	22,3	10,8
Verbano-Cusio-Ossola	24,9	11,4
Vercelli	22,0	10,2

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Analisi preliminare tra due variabili: forma, direzione e forza della relazione



Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Diagramma a dispersione

Detto anche scatter o nuvola di punti, consente di descrivere la distribuzione unitaria doppia di due **variabili quantitative** mettendo in **evidenza** un'eventuale **relazione**.

Per ciascuna unità statistica osservata si conoscono i valori delle due variabili oggetto di analisi.

È un grafico che si basa su un riferimento cartesiano.

Caratteristiche:

- Gli assi servono per indicare le variabili; tasso di disoccupazione (x) e %Neet (y)
- I dati sono rappresentati come punti nel grafico, coordinata (x, y)

Per approfondire

Sul tema della rappresentazione grafica dei dati un sito molto concreto e curato è *Data Visualization Catalogue*.

Si tratta di un progetto privato, ma accessibile gratuitamente a tutti, per creare un catalogo delle diverse tipologie di visualizzazione delle informazioni.

Per ogni tipo di grafico viene fornita una breve descrizione accompagnata da un esempio e da alcuni suggerimenti per l'utilizzo.

Bibliografia e sitografia

(2012) Bocci L., Mingo I., Statistiche, istruzioni per l'uso. Roma. Edizioni Nuova Cultura

(2022) Stoppiello S., Della Queva S., Nicosia M., Il nuovo censimento permanente delle istituzioni non profit: approfondimenti tematici e prospettive di analisi future

(https://www.istat.it/it/files//2022/09/34_ISTAT_MILeS2022_Stoppiello_DellaQueva_Nicosia.pdf)

<https://datavizcatalogue.com/>

IstatData. <https://esploradati.istat.it/databrowser/>

Censimento permanente delle istituzioni non profit: <https://www.istat.it/it/archivio/296339>

Grazie

utenti.territorio@istat.it