

01 – Rappresentazioni statistiche e grafiche

Unità n° 03

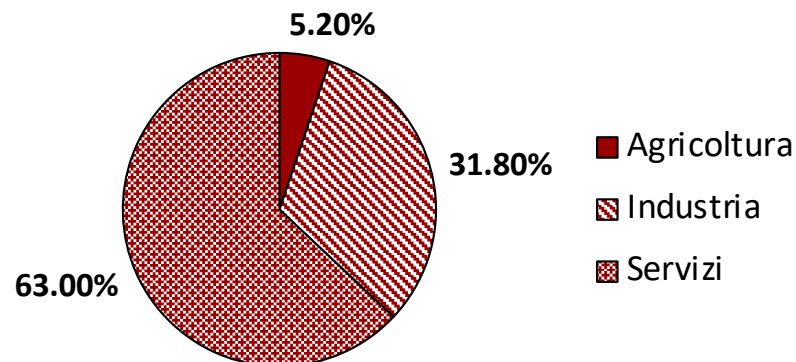
Per descrivere l'insieme delle modalità di un carattere osservato su un collettivo è possibile ricorrere alle distribuzioni o ad una loro opportuna rappresentazione grafica

Occupati per settore produttivo - 1971-2001

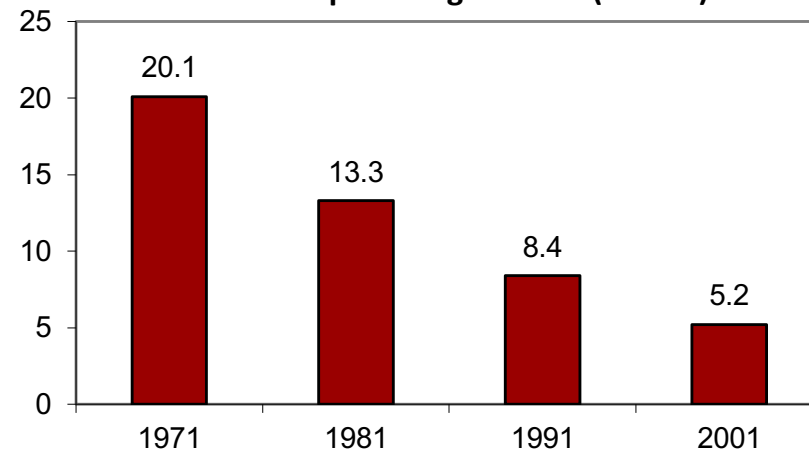
Settore	Anni			
	1971	1981	1991	2001
Agricoltura	20,1	13,3	8,4	5,2
Industria	39,5	37,2	32,0	31,8
Servizi	40,4	49,5	59,6	63,0

Vogliamo studiare la distribuzione per settore produttivo in Italia tra il 1971 e il 2001

Occupati per settore produttivo (2001)



Occupati in Agricoltura ('71-'01)



02 – Un diverso tipo di rappresentazione

Unità n° 03

L'utilizzo di tabelle è senza dubbio utile, sia perché abbiamo immediatamente un quadro generale delle manifestazioni del fenomeno nel collettivo/campione sia perché possiamo presentare a terzi i risultati della nostra rilevazione con un buon livello informativo (non tutti studiano la Statistica...)

Esiste un altro strumento molto importante per la rappresentazione dei dati che presenta talvolta un potere informativo superiore a quello delle distribuzioni: si tratta in sostanza di effettuare una ulteriore trasformazione dei dati e di rappresentarli in forma diagrammatica

Si soddisfano principalmente tre esigenze:

- 1) dare una visione d'insieme efficace del fenomeno**
- 2) rendere possibile il confronto tra fenomeni diversi, al fine di scoprire delle relazioni tra loro**
- 3) aumentare la leggibilità dei dati e delle loro elaborazioni facilitandone una lettura di sintesi**

03 – Saper scegliere la giusta rappresentazione

Unità n° 03

Esistono tantissime rappresentazioni grafiche: alcune sono più idonee a rappresentare i caratteri quantitativi (discreti o continui), altre a rappresentare i caratteri qualitativi (sconnessi o ordinali)

Ovviamente la scelta del tipo di rappresentazione influenza notevolmente la comprensione del fenomeno perché in assenza di una tabella esplicativa in alcuni casi non abbiamo più nessuna idea di come i dati si sono presentati nel collettivo, soprattutto se, come visto, abbiamo delle frequenze relative o percentuali

ESEMPIO



Grafici a nastri o a barre

caratteri qualitativi (e quantitativi)



Grafici a torta

caratteri qualitativi (e quantitativi)



Grafici radar

caratteri qualitativi ordinali ciclici



Grafici ad aree

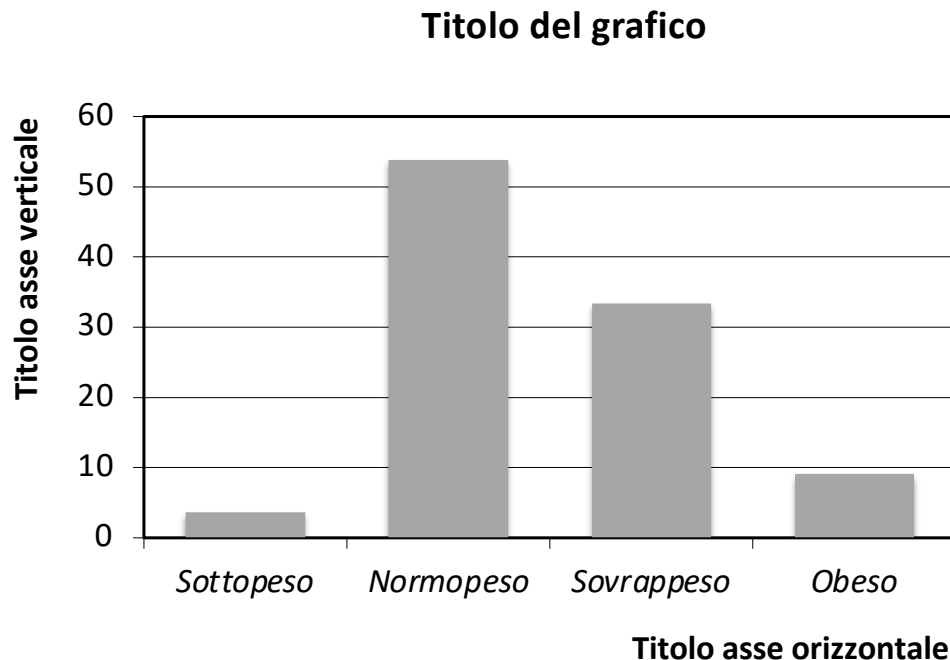
caratteri quantitativi continui nel tempo

...

04 – Disegnare correttamente un grafico

Unità n° 03

Affinché la rappresentazione grafica sia di chiara ed univoca interpretazione è necessario che nella sua costruzione vengano rispettate alcune regole



È necessario innanzi tutto che nel titolo del grafico ci sia l'indicazione del tipo di carattere e del collettivo:

“Percentuale di individui con età >17 per indice di massa corporea”

È poi necessario indicare le modalità del carattere che studiamo, e l'unità di misura se è quantitativo

Completano il grafico la legenda, la fonte dei dati ed eventuali note esplicative. Eventuali elementi decorativi non sono legati ai dati (e a volte diminuiscono la leggibilità del grafico...)

05 – Le componenti di supporto al grafico

Unità n° 03

Il **titolo del grafico** deve essere breve e conciso:

➔ un titolo **informativo** contiene le informazioni necessarie per comprendere i dati, e risponde alle tre domande: “Cosa?”, “Dove?” e “Quando?”

esempio: Tasso di disoccupazione in Italia (2009-2012)

➔ un titolo **descrittivo** sintetizza in poche parole il trend o il pattern rappresentato

esempio: Aumento della disoccupazione in Italia dal 2009 al 2012

I **titoli degli assi** identificano le variabili rappresentate dagli assi (se si possono evincere dal titolo del grafico non è necessario ripeterle)

Le **etichette degli assi** identificano le modalità o i valori rappresentati nel grafico

06 – Altre componenti

Unità n° 03

L'**unità di misura dei dati** (“in migliaia” , “%” etc.) è fondamentale per la leggibilità dei dati. Se l'unità di misura è ovvia, si può omettere

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 utili per fornire definizioni o informazioni sulla metodologia

07 – Grafici a nastri e a barre: gli ortogrammi

Unità n° 03

Gli **ortogrammi** sono impiegati per rappresentare distribuzioni unitarie semplici o distribuzioni di frequenza di caratteri qualitativi (sconnessi e rettilinei)

Possiamo distinguere grafici:

- **a barre**, costituiti da una successione di colonne, segmenti verticali o rettangoli (a base uguale) equidistanti, in numero pari alle unità o alle alle modalità del carattere, e con una *altezza* uguale o proporzionale all'intensità o alla frequenza della modalità
- **a nastri**, costituiti da tanti nastri sovrapposti ed equidistanti, in numero pari alle unità o alle alle modalità del carattere, e con una *lunghezza* uguale o proporzionale all'intensità o alla frequenza della modalità

Gli ortogrammi sono quindi caratterizzati dal fatto di avere un solo asse (verticale per quello a barre, orizzontale per quello a nastri) con scala graduata, secondo l'unità di misura che si è scelta per rappresentare intensità o frequenze. Sull'altro asse figureranno unità o modalità

08 – Gli ortogrammi: grafici a nastri e a barre (2)

Unità n° 03

GRAFICO A NASTRI



TOP 5 REGIONS IN THE EU
NIGHTS SPENT IN TOURIST ACCOMMODATION, 2016
(million)

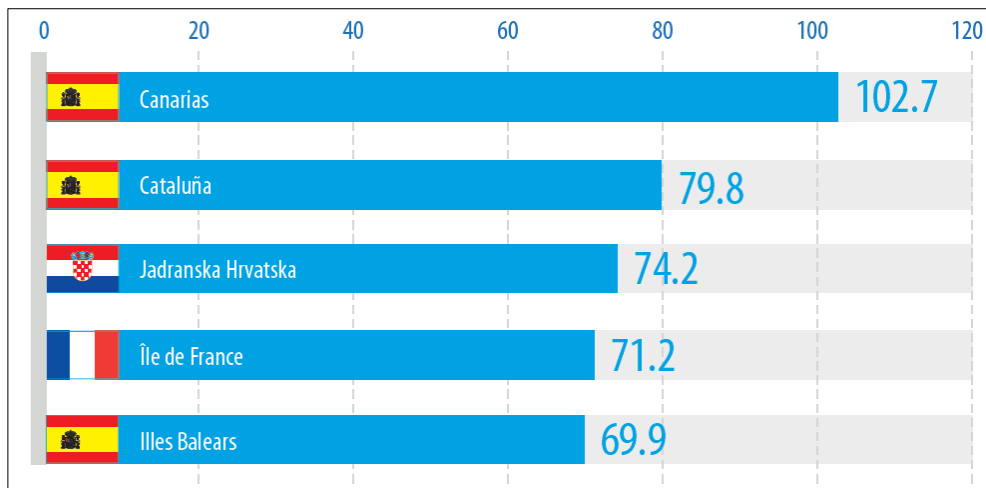
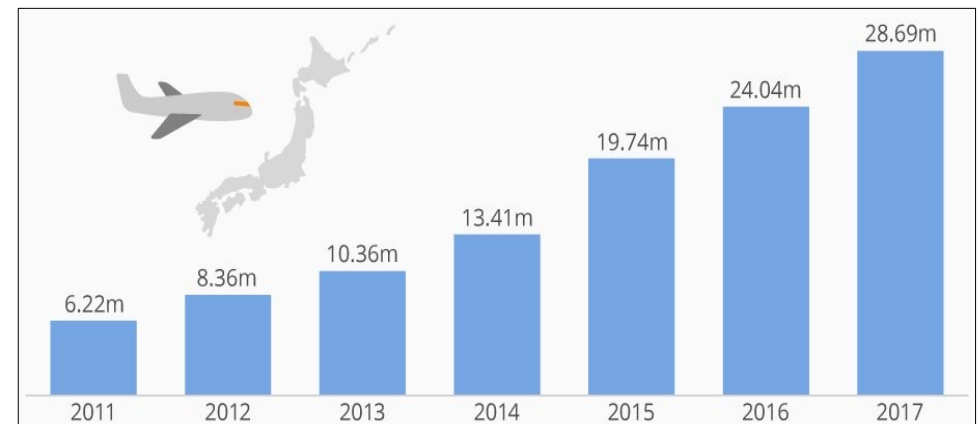


GRAFICO A BARRE



Number of international visitors to Japan from 2011 to 2017
(million)

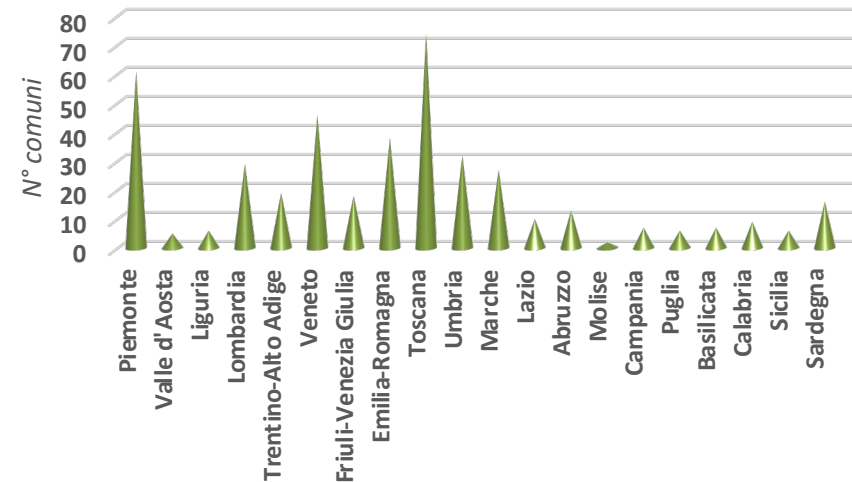
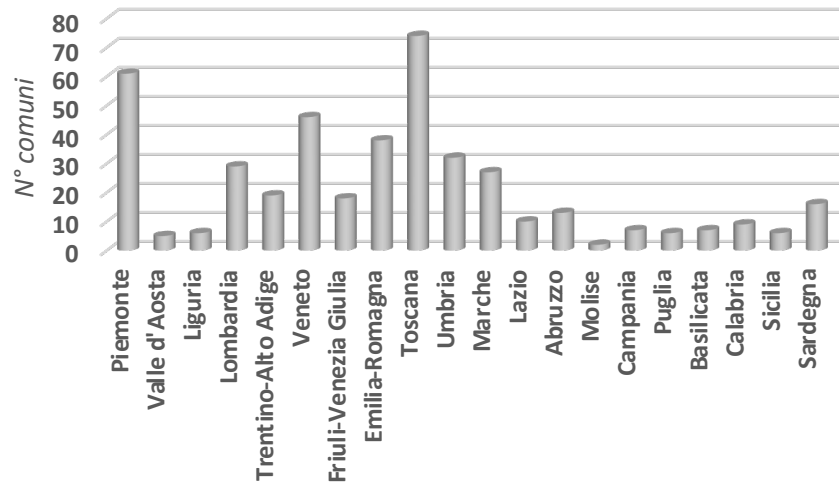
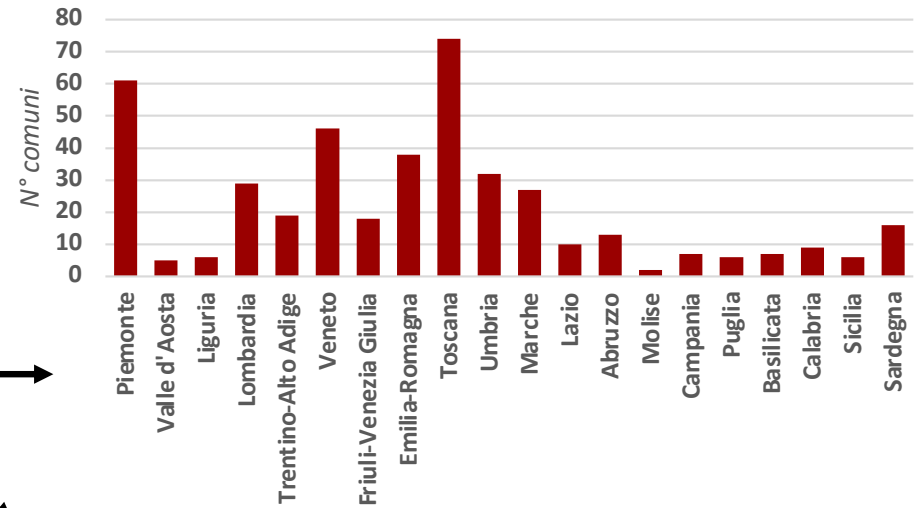


09 – Ortogrammi con diverso aspetto (2D vs 3D)

Unità n° 03

Possiamo modificare l'aspetto di un diagramma, ad esempio utilizzando una rappresentazione di tipo tridimensionale, senza alterare la lettura dei dati originari e l'interpretazione del fenomeno

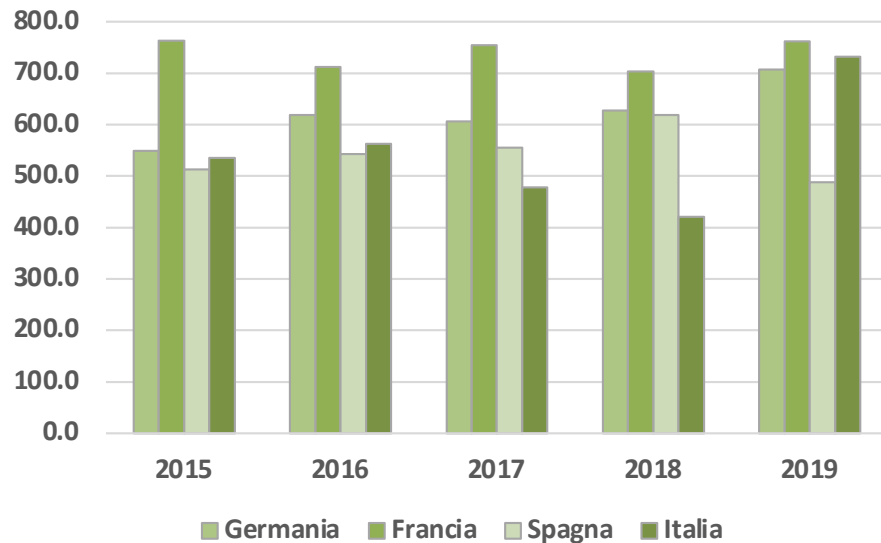
Distribuzione dei comuni a vocazione culturale, storica, artistica e paesaggistica per Regione (anno 2019)



10 – Altri ortogrammi (a nastri e a barre)

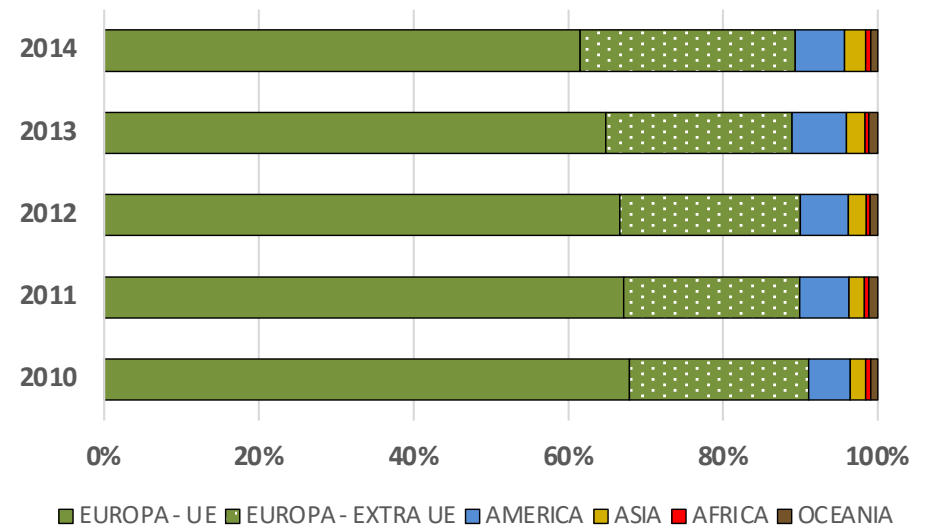
Unità n° 03

*Spesa (milioni di £) dei visitatori stranieri a Londra
(anni 2015-2019)*



Negli ortogrammi a barre appaiate è possibile confrontare le distribuzioni di caratteri diversi, la composizione di collettivi/campioni oppure di sottocollettivi/sottocampioni diversi

*Viaggiatori stranieri in Italia per continente di provenienza
(anni 2010-2014)*



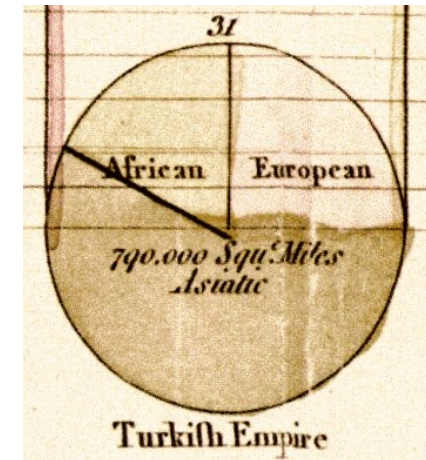
Negli ortogrammi a barre impilate è possibile confrontare la composizione (solitamente in termini percentuali) di collettivi/campioni o di sottocollettivi/sottocampioni

11 – Gli areogrammi circolari

Unità n° 03

Gli **areogrammi circolari** sono comunemente utilizzati per visualizzare la composizione di un collettivo/campione rispetto al carattere studiato o l'intensità del carattere tra le diverse unità del collettivo/campione

➔ Consentono di mettere in evidenza l'importanza relativa delle singole parti rispetto al totale



(William Playfair, 1801)

È possibile distinguere due tipologie di areogramma circolare:

- **a settori (o spicchi) proporzionali:** con angoli proporzionali al peso delle singole parti (e raggio fisso)
- **a settori (o spicchi) uguali:** con raggio proporzionale al peso delle singole parti (e angoli fissi)

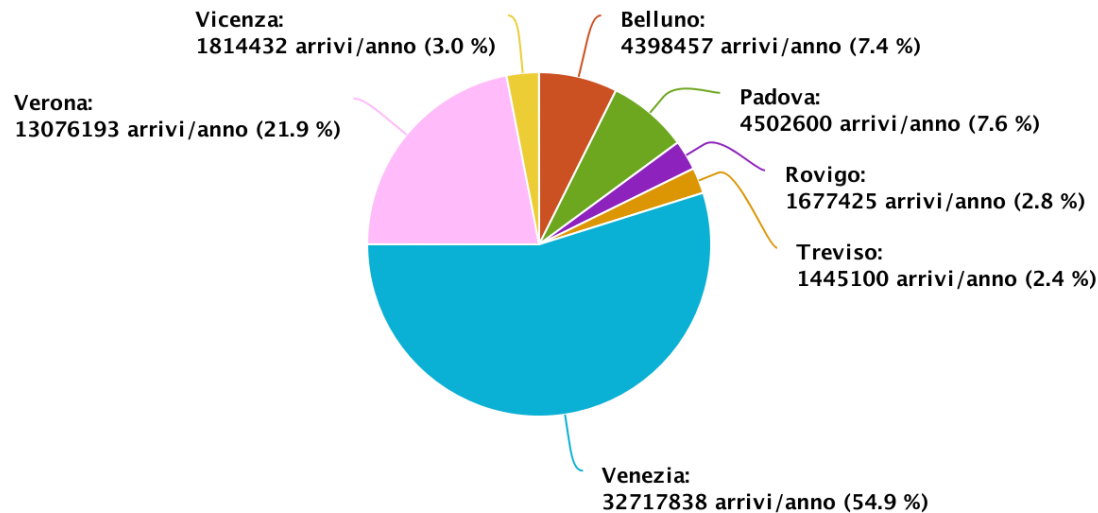
Per entrambe le versioni è possibile utilizzare differenti soluzioni grafiche

12 – Areogrammi a settori proporzionali

Unità n° 03

Gli areogrammi a settori proporzionali sono noti come *grafici a torta*: possiamo rappresentare la distribuzione di frequenza di un carattere qualitativo sconnesso o una distribuzione in classi

Distribuzione degli arrivi* per le Province del Veneto nell'anno 2015



* Numero di clienti, italiani e stranieri, ospitati negli esercizi ricettivi (alberghieri o complementari)

L'idea è che ogni "fetta" della torta è proporzionale all'importanza delle diverse manifestazioni del carattere.

Se volessimo tracciarlo manualmente:

$$\text{Fetta} : x^\circ = \text{Torta} : 360^\circ$$

Troviamo l'angolo di una fetta, noto il fatto che il cerchio ha un angolo di 360° (la fetta è il numero di unità con quella specifica modalità, la torta è il collettivo)

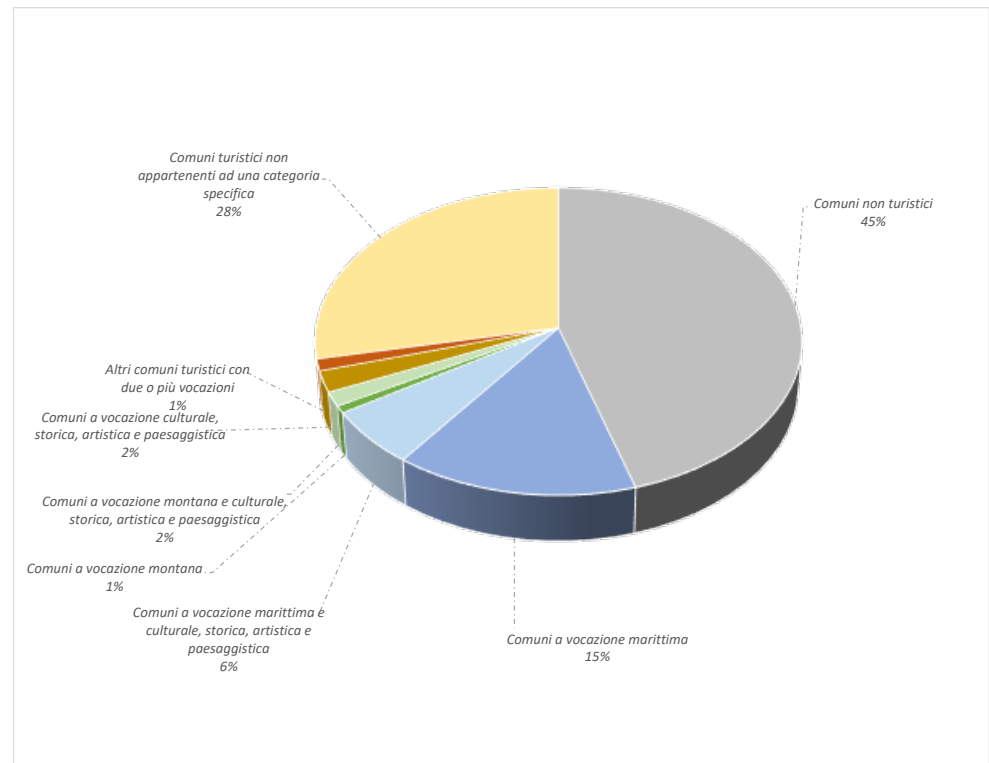
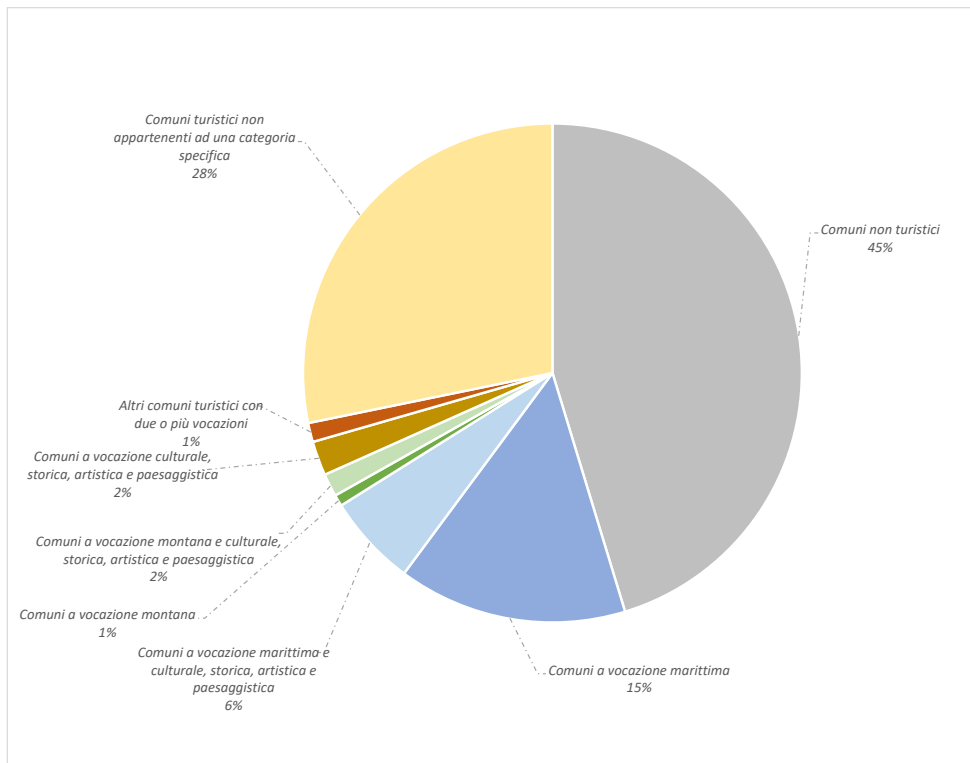
Può anche essere usato per distribuzioni unitarie di caratteri quantitativi quando ha senso sommare le diverse modalità

13 – Grafici a torta con diverso aspetto: 2D vs 3D

Unità n° 03

Anche per gli areogrammi si può usare una rappresentazione di tipo bidimensionale o di tipo tridimensionale, senza alterare la lettura dei dati originari e l'interpretazione del fenomeno

Distribuzione dei comuni turistici e non turistici della Calabria per tipologia (anno 2019)



<https://www.istat.it/it/archivio/247191>

Per migliorare la leggibilità del grafico è possibile separare una o più dette della torta

14 – Grafici a torta con diverso aspetto: diagramma ad anello

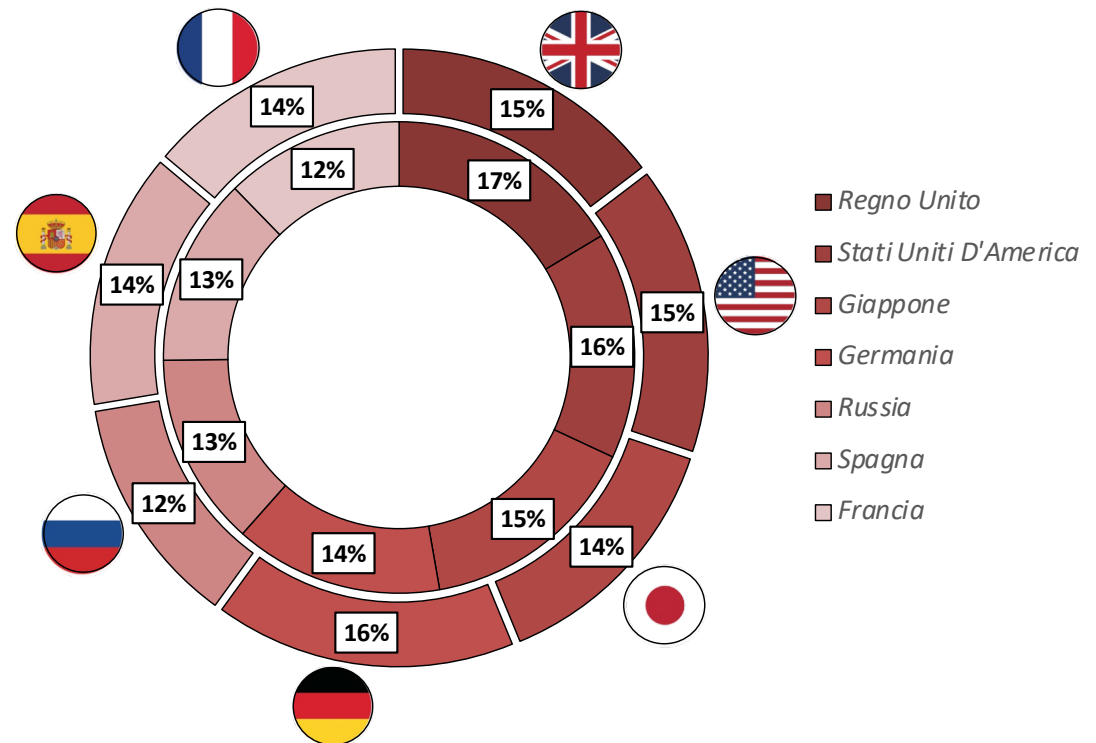
Unità n° 03

Una alternativa al diagramma a torta è il cosiddetto *diagramma ad anello*, con il quale è possibile anche confrontare collettivi differenti o effettuare dei confronti temporali su un fenomeno

È possibile anche combinare un diagramma ad anello con un diagramma a torta

Presenze di turisti stranieri per Paese di provenienza nel Comune di Milano (primi 7 Paesi, anni 2008-2009)

	2008	2009
Regno Unito	327.704	294.457
Stati Uniti D'America	311.075	311.784
Giappone	306.115	274.284
Germania	286.381	325.931
Russia	265.026	249.077
Spagna	258.308	276.085
Francia	245.412	279.512

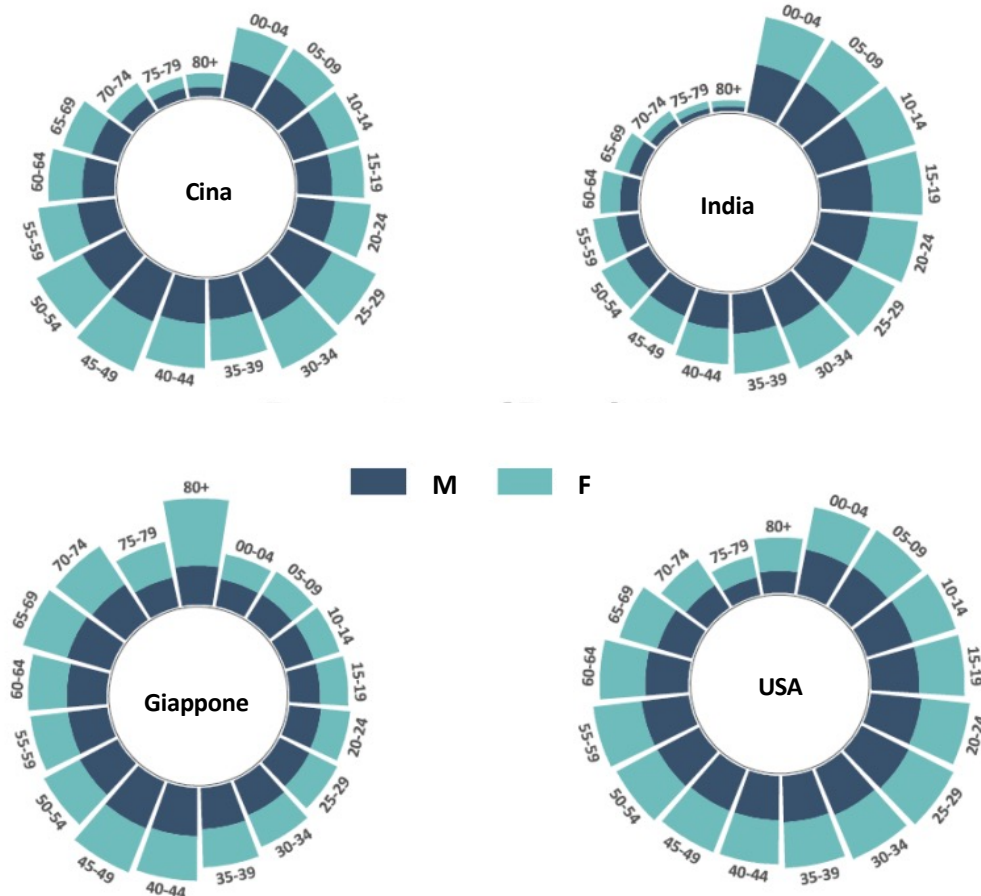


15 – Areogrammi a settori uguali

Unità n° 03

Quando i settori sono uguali, le “fette” della torta hanno tutte uguale ampiezza/angolo e un raggio proporzionale al diverso peso degli elementi considerati -> Diagrammi Coxcomb

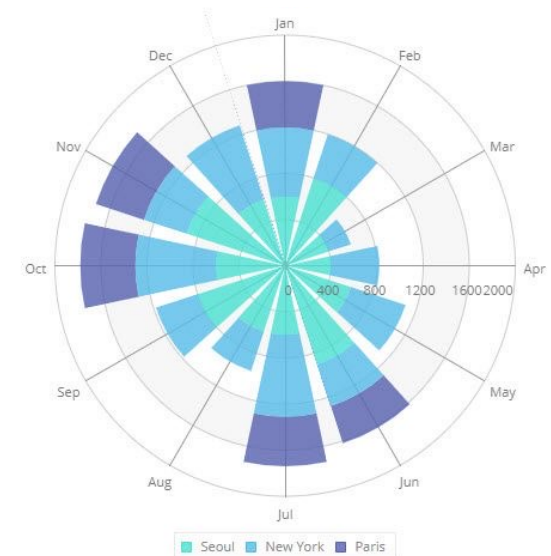
Distribuzione della popolazione in classi di età



È possibile rappresentare più caratteri o più sottocollettivi

Tali rappresentazioni possono anche essere utilizzate nel caso di caratteri qualitativi ciclici

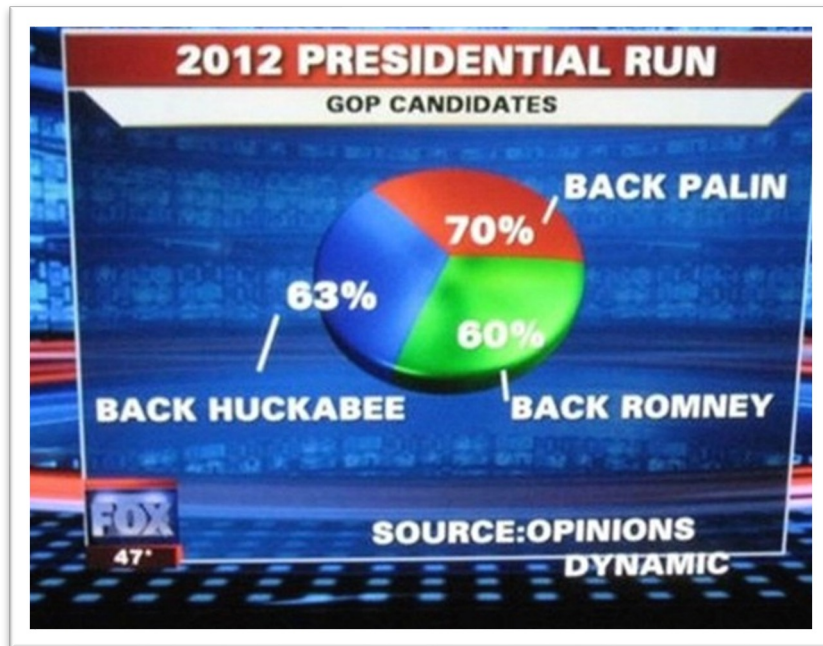
Piovosità (precipitazioni in mm)



16 – Non tutte le torte sono buone...

Unità n° 03

Spesso i dati statistici e le rappresentazioni grafiche sono utilizzate in modo erroneo o grossolano producendo effetti catastrofici nella comprensione del fenomeno



Sondaggio telefonico del 17-18 novembre 2009 sui candidati repubblicani alla Casa Bianca per le elezioni presidenziali del 2012 (Fox News)

DOMANDA:

“Le leggerò ora il nome di alcune persone. Mi dica se in generale ha una opinione favorevole o non favorevole per ciascuna di esse...”

Il giornalista nel commentare il grafico afferma che Sarah Palin è in testa nel sondaggio con il 70% delle preferenze, ma il totale sarebbe in questo modo del 193%!

Sarah Palin

	Favorable	Unfavorable	Can't say	Never heard
17-18 Nov 09	47%	42	10	2
Democrats	25%	62	9	4
Republicans	70%	21	8	1
Independents	49%	38	11	1

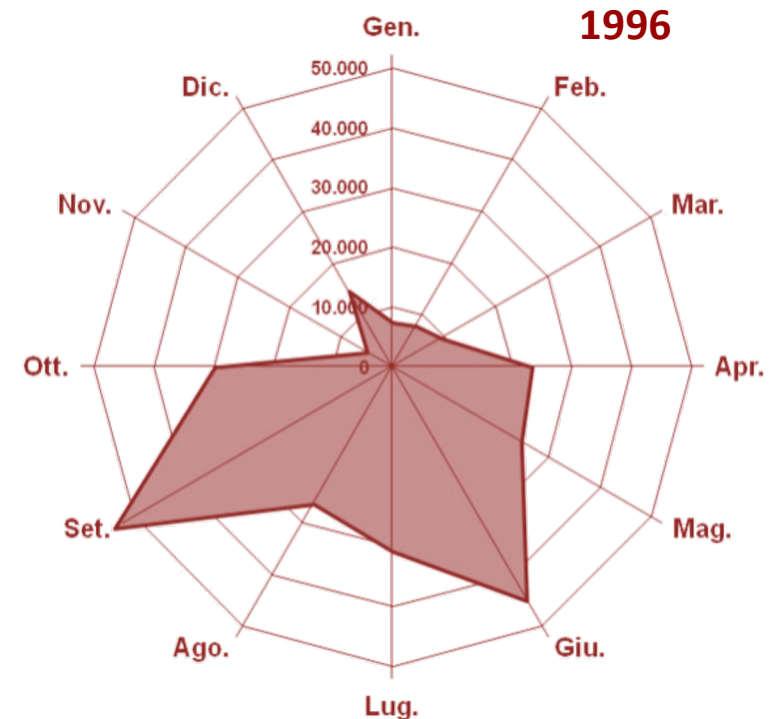
In realtà il dato corretto è che gli elettori repubblicani hanno in generale un'opinione favorevole più alta per Sarah Palin rispetto a quella che hanno degli altri due candidati

17 – Diagrammi in coordinate polari: i grafici a radar

Unità n° 03

Per quei fenomeni che è possibile rappresentare con caratteri qualitativi ciclici (ES.: quelli che osserviamo nei diversi mesi dell'anno o nei giorni della settimana) è idonea una rappresentazione detta "a radar" per la sua particolare forma

Su uno stesso grafico possiamo rappresentare più collettivi. L'area evidenziata è proporzionale alla dimensione del fenomeno nel collettivo



Matrimoni per mese di celebrazione

	Gen.	Feb.	Mar.	Apr.	Mag.	Giu.	Lug.	Ago.	Set.	Ott.	Nov.	Dic.
1994	8.441	8.188	8.606	28.253	27.386	40.762	36.473	24.715	54.816	32.892	4.973	16.102
1995	8.174	8.567	9.597	25.051	28.337	40.208	36.362	23.433	58.655	30.109	4.725	16.791
1996	7.451	7.911	9.560	23.423	24.953	45.000	30.745	26.391	53.837	29.810	4.898	14.632
1997	7.155	7.025	8.178	24.104	27.663	40.649	33.988	29.019	50.395	29.686	4.807	15.069
1998	7.189	8.647	9.048	20.653	31.023	38.732	33.659	29.392	51.106	30.938	4.663	14.984

18 – Rappresentazione grafica delle serie storiche

Unità n° 03

Nella tabella seguente sono riportati i *prezzi medi nazionali al consumo*, in Euro per Litro, della benzina senza piombo dal 1996 al 2011 (fonte: Ministero dello Sviluppo Economico)

ANNO	€/L.
1996	0.92
1997	0.94
1998	0.91
1999	0.96
2000	1.08
2001	1.05
2002	1.05
2003	1.06
2004	1.12
2005	1.22
2006	1.28
2007	1.30
2008	1.38
2009	1.23
2010	1.36
2011	1.55



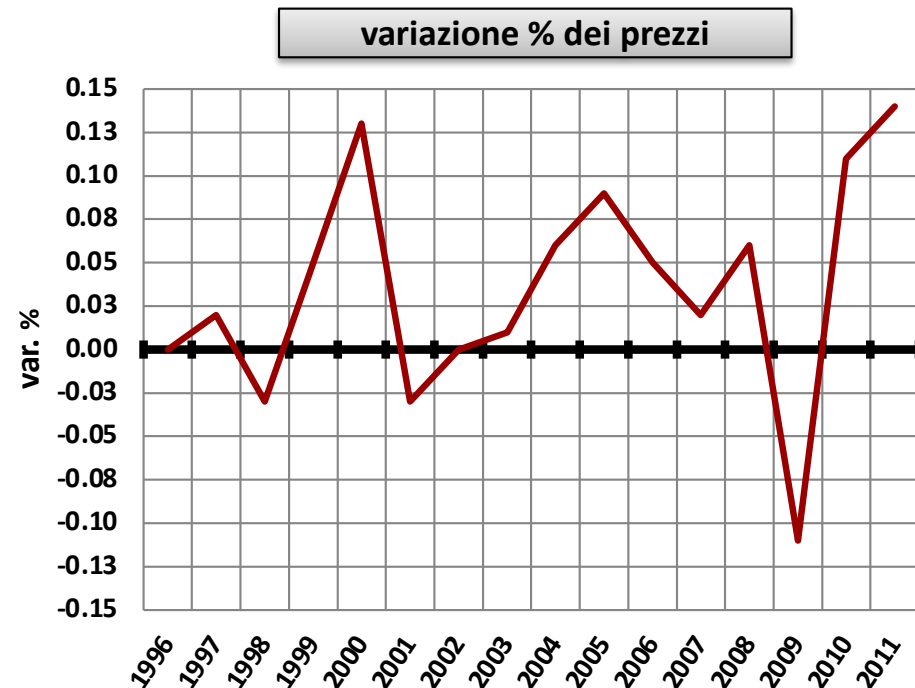
Qual è il collettivo?

19 – Dati grezzi e variazioni

Unità n° 03

Consideriamo i *prezzi medi nazionali al consumo*, in Euro per Litro, della benzina senza piombo dal 1996 al 2011, e costruiamo la distribuzione delle variazioni relative di prezzo

ANNO	€/L.	ANNO	var.rel	var.%
1996	0.92	1996/1997	1.02	0.02
1997	0.94	1997/1998	0.97	-0.03
1998	0.91	1998/1999	1.05	0.05
1999	0.96	1999/2000	1.13	0.13
2000	1.08	2000/2001	0.97	-0.03
2001	1.05	2001/2002	1.00	0.00
2002	1.05	2002/2003	1.01	0.01
2003	1.06	2003/2004	1.06	0.06
2004	1.12	2004/2005	1.09	0.09
2005	1.22	2005/2006	1.05	0.05
2006	1.28	2006/2007	1.02	0.02
2007	1.30	2007/2008	1.06	0.06
2008	1.38	2008/2009	0.89	-0.11
2009	1.23	2009/2010	1.11	0.11
2010	1.36	2010/2011	1.14	0.14
2011	1.55			

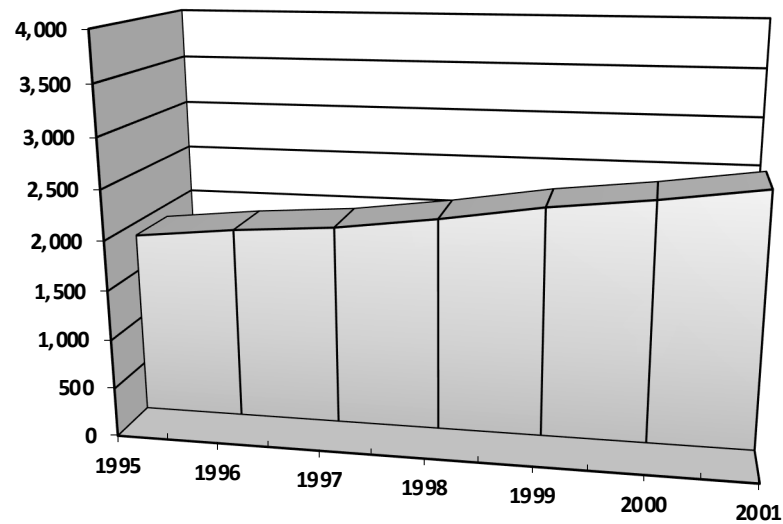


20 – Grafici ad aree

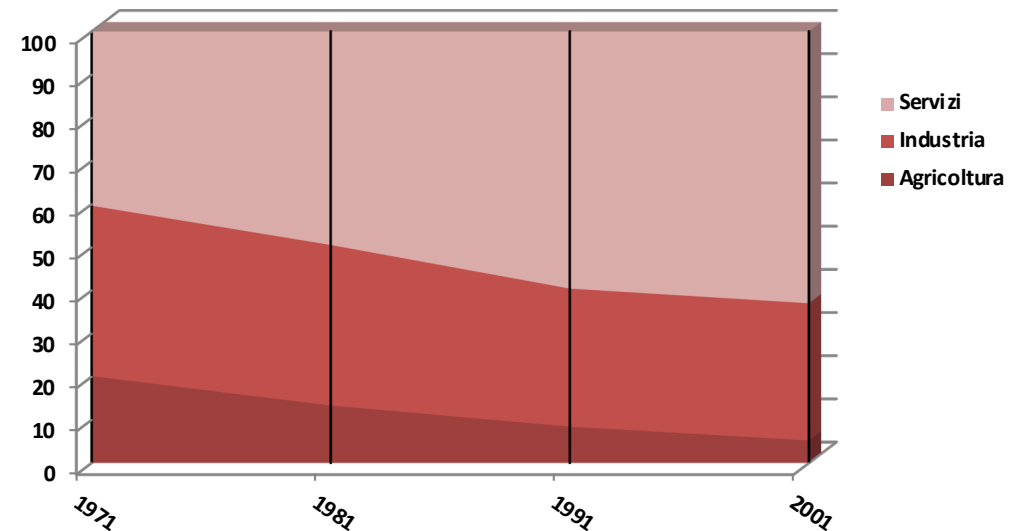
Unità n° 03

Quando dobbiamo rappresentare dei caratteri quantitativi continui nel tempo può essere utile utilizzare le rappresentazioni ad aree: anche in tal caso come nel precedente è possibile inserire nel grafico un unico collettivo o confrontare il fenomeno che ci interessa in diversi collettivo

Distribuzione del titolo di studio



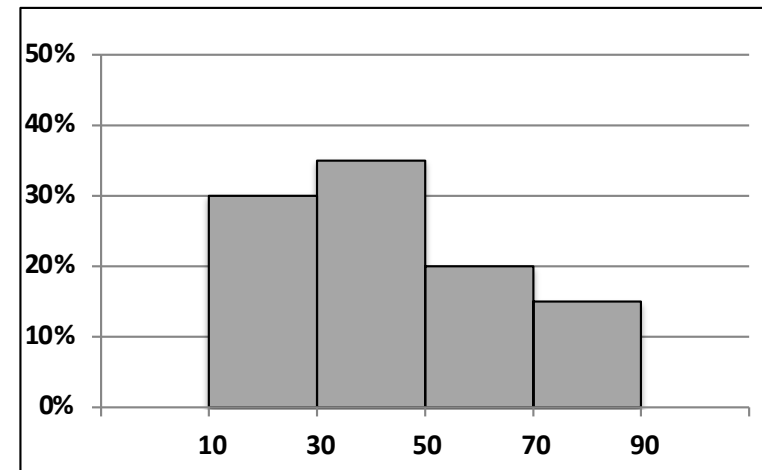
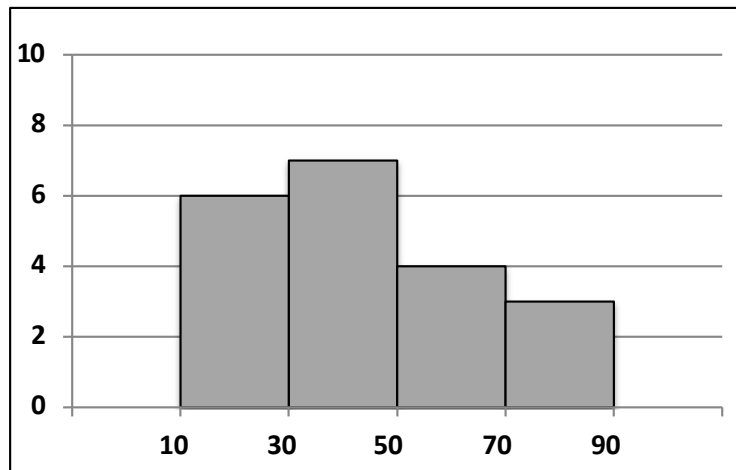
Occupati per settore di attività



21 – Gli istogrammi

Unità n° 03

Lo strumento grafico utilizzato per visualizzare le distribuzioni di frequenza in classi è il cosiddetto **istogramma**: sull'asse orizzontale sono rappresentate le classi, su quello verticale è invece espressa l'intensità del fenomeno (in termini di frequenza assoluta, relativa o %)



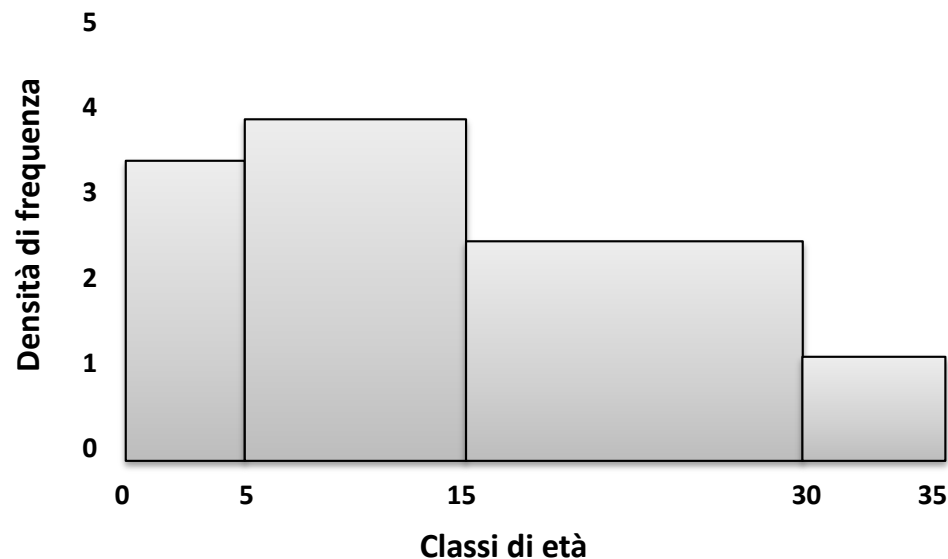
L'area di ciascuna barra sarà proporzionale all'intensità: la proporzione ovviamente è la stessa sia che l'intensità è misurata con le frequenze assolute che l'intensità è espressa, ad esempio, in termini percentuali. Le barre non sono tra loro distanziate per dare un'idea di continuità nella rappresentazione del carattere

22 – Istogramma per classi non equiampie

Unità n° 03

Nel caso di classi con ampiezza diversa è ancora possibile l'utilizzo dell'istogramma come rappresentazione grafica

In questo caso assume però un significato differente il fatto che l'intensità in una specifica classe sia maggiore o minore, perché i possibili valori osservati sulle unità statistiche sono "pochi" o "molti" a seconda che l'ampiezza della classe sia minore o maggiore



Per poter allora ottenere delle barre proporzionali all'intensità espressa nella classe, e allo stesso tempo considerare l'ampiezza della stessa, si utilizza come misura dell'intensità non più la frequenza ma la cosiddetta *densità di frequenza* :

$$\text{Area} = \text{base} \times \text{altezza}$$

Area -> frequenza

base -> ampiezza

altezza -> densità di frequenza

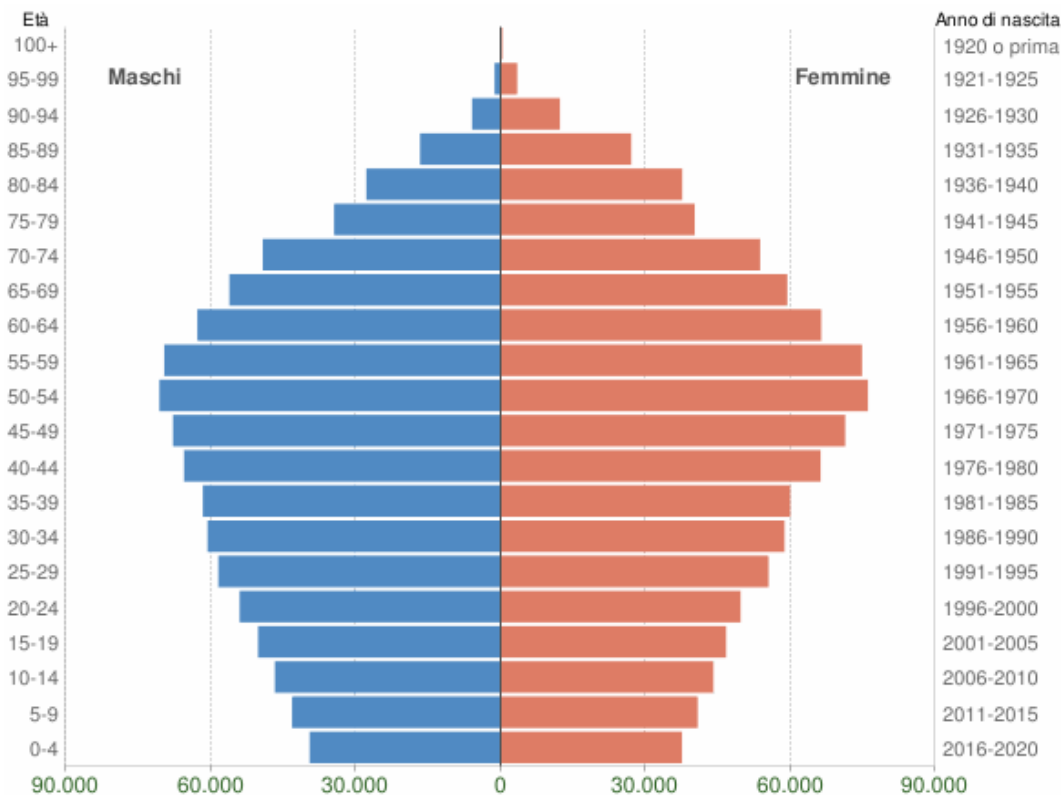
23 – La piramide delle età

Unità n° 03

È un modo per rappresentare la distribuzione per età di una popolazione. Si costruisce accostando:

- ➔ l'istogramma della distribuzione in classi di età degli uomini (a sinistra del grafico)
- ➔ l'istogramma delle distribuzione in classi di età delle donne (a destra del grafico)

Popolazione della Calabria per età e genere (al 1 gennaio 2020)



Prende il nome di piramide perché in passato aveva solitamente la forma approssimata di un triangolo isoscele: le popolazioni più giovani, e quindi le basi, dovrebbero essere più numerose per poi via via diminuire per effetto della mortalità

Dalla forma si può dedurre l'andamento demografico a cui sta tendendo un Paese:

- forma tendente a una *piramide*: popolazione in crescita
- forma tendente a un *rettangolo*: crescita nulla
- forma tendente a un *trapezio*: decremento

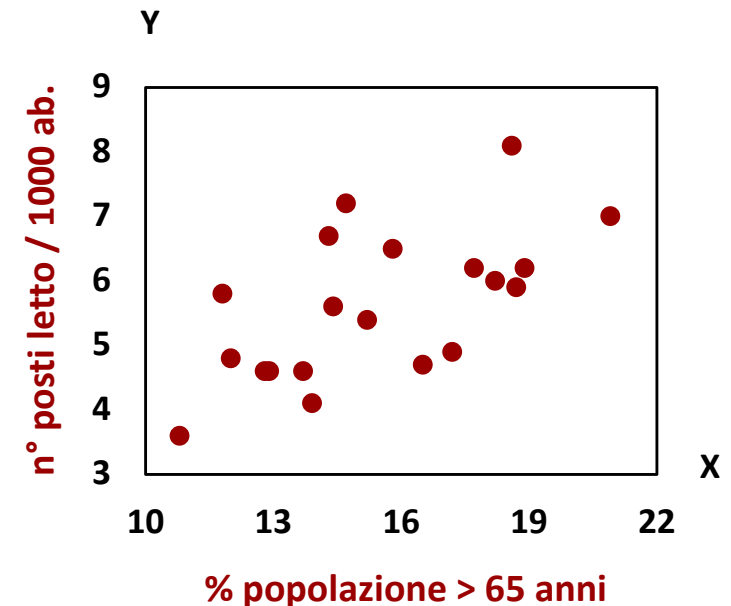
24 – Il diagramma di dispersione

Unità n° 03

Le variabili doppie quantitative possono essere rappresentate graficamente su un piano cartesiano che prende il nome di **diagramma di dispersione** o **scatterplot**

Consideriamo per ciascuna Regione la percentuale di popolazione residente con una età maggiore di 65 anni e l'offerta di posti letto ospedalieri per 1000 abitanti

Regione	X	Y	Regione	X	Y
<i>Piemonte</i>	17.2	4.9	<i>Marche</i>	17.7	6.2
<i>Valle d'Aosta</i>	15.2	5.4	<i>Lazio</i>	13.7	4.6
<i>Lombardia</i>	14.4	5.6	<i>Abruzzo</i>	15.8	6.5
<i>Trentino-Alto Ad.</i>	14.3	6.7	<i>Molise</i>	16.5	4.7
<i>Veneto</i>	14.7	7.2	<i>Campania</i>	10.8	3.6
<i>Friuli-Venez.-Giul.</i>	18.6	8.1	<i>Puglia</i>	11.8	5.8
<i>Liguria</i>	20.9	7.0	<i>Basilicata</i>	13.9	4.1
<i>Emilia Romagna</i>	18.9	6.2	<i>Calabria</i>	12.8	4.6
<i>Toscana</i>	18.7	5.9	<i>Sicilia</i>	12.9	4.6
<i>Umbria</i>	18.2	6.0	<i>Sardegna</i>	12.0	4.8



25 – Altre rappresentazioni grafiche

Unità n° 03

speedometer o gauge plot

Incidenza % degli arrivi di turisti stranieri
(anno 2018)

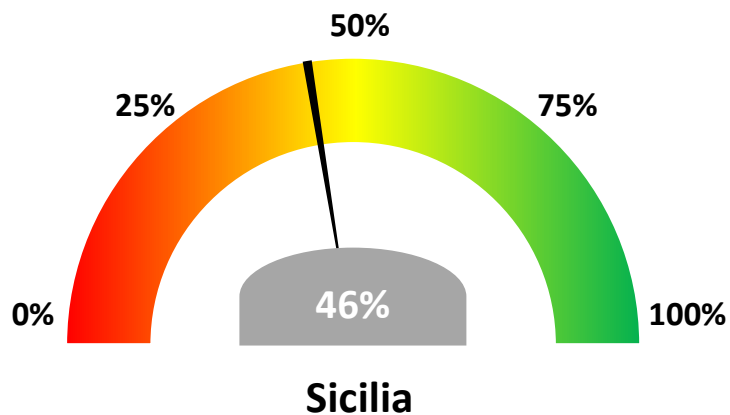
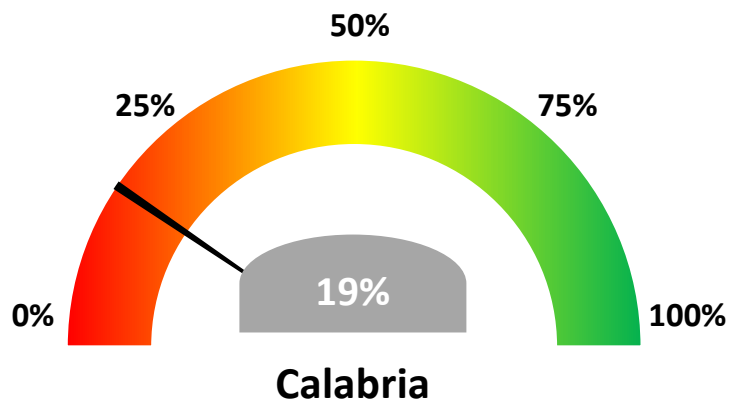
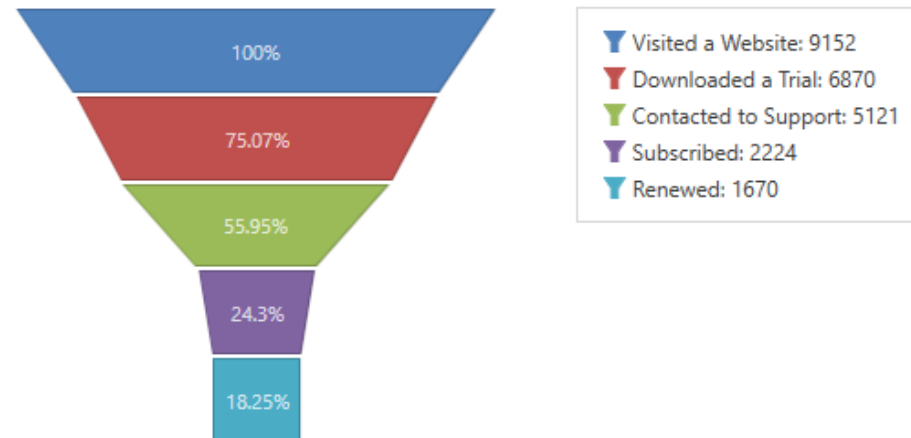


diagramma a imbuto

Visitatori del sito www.esempio.it per
profondità del contatto



racetrack plot

Top 10 prodotti acquistati sul Web
(Nielsen, Q3 2012)



26 – Rappresentazione grafica delle serie territoriali

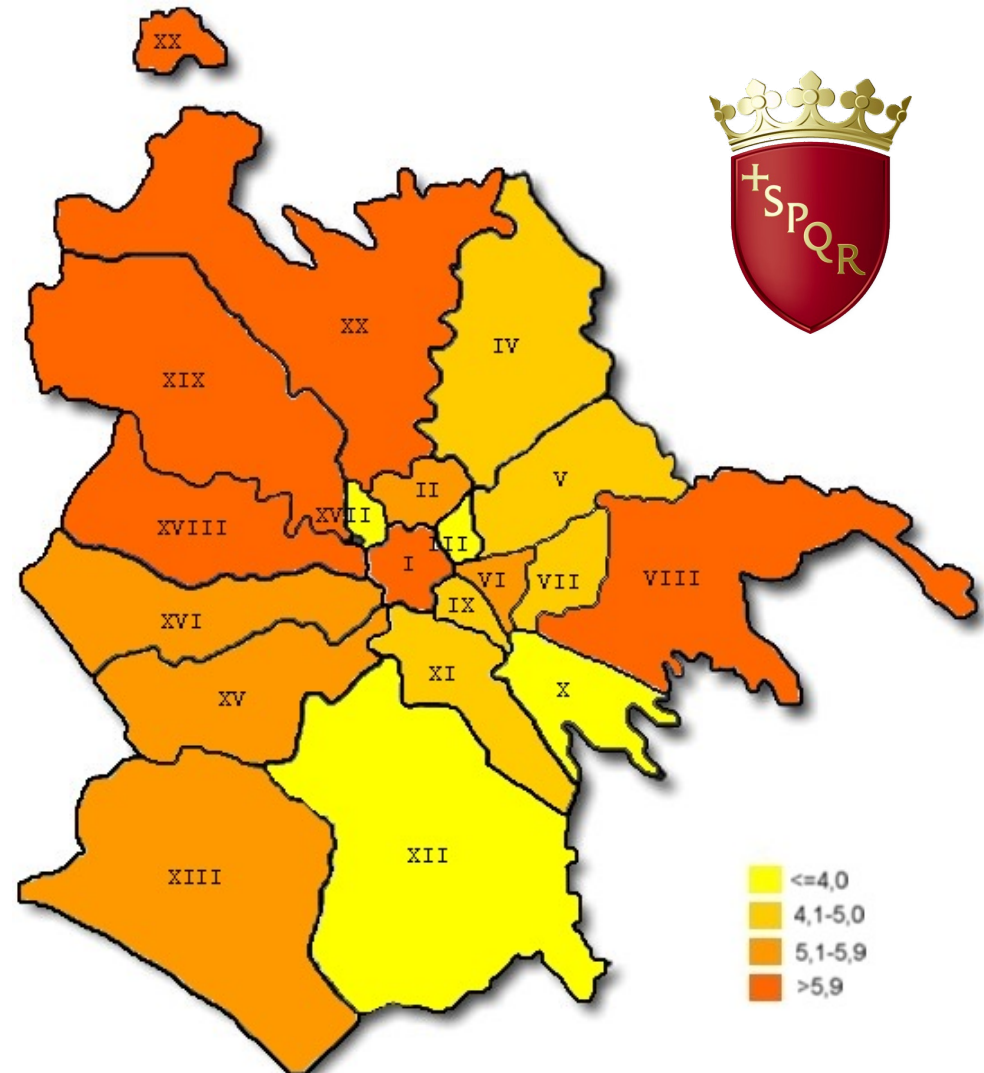
Unità n° 03

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

ESEMPIO

Incidenza della popolazione straniera sul totale della popolazione residente nel Comune di Roma per Municipio (valori percentuali, Censimento 2001)

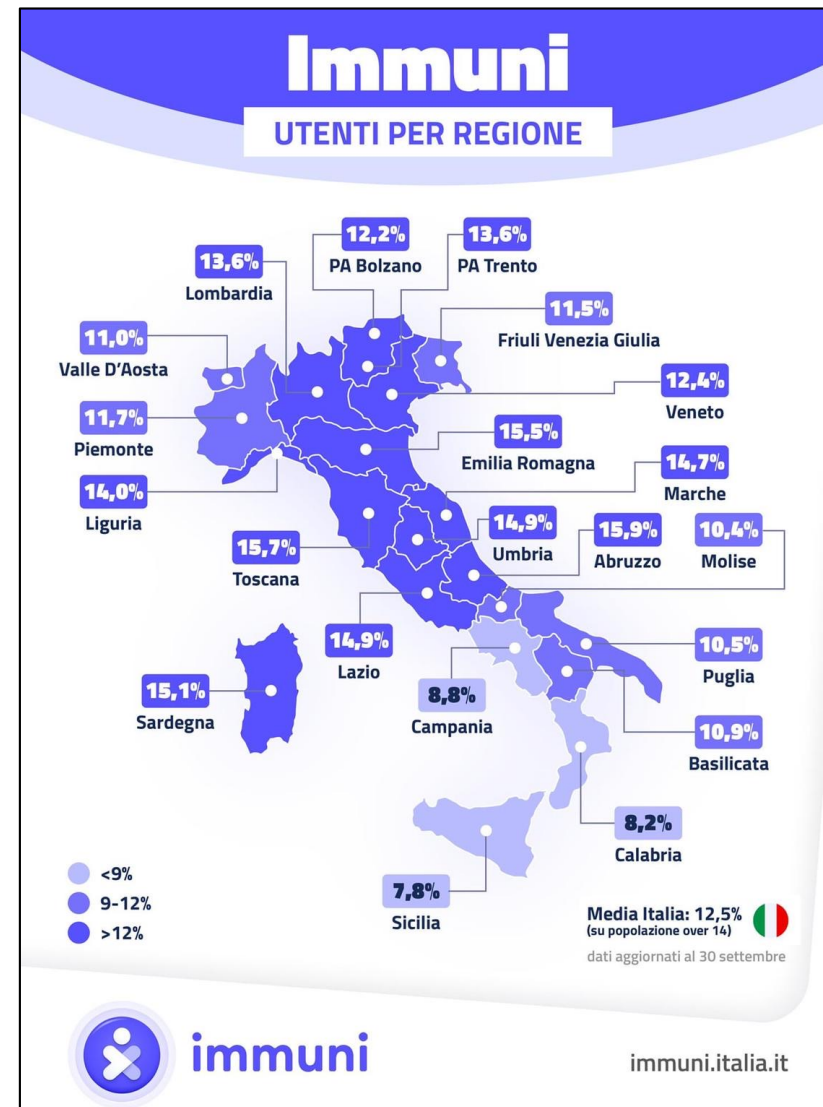
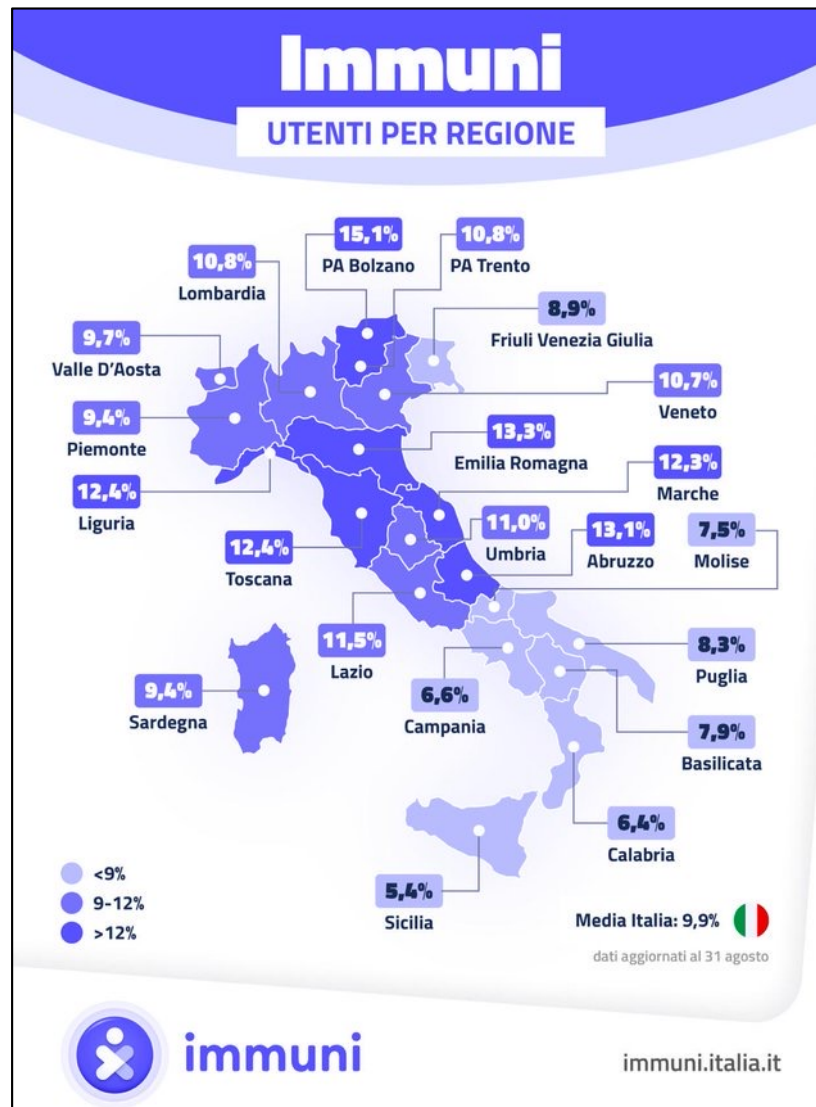


27 – Esempio: Utenti App *Immuni*

Unità n° 03

Dati al 31 agosto 2020

Dati al 30 settembre 2020



28 – Cartodiagrammi

Unità n° 03

I **cartodiagrammi** non sono altro che dei cartogrammi con cui rappresentare delle serie territoriali di due o più caratteri, oppure delle distribuzioni di frequenza su collettivi che riguardano unità appartenenti alle diverse aree di un territorio

ESEMPIO

Nati vivi e nati morti per 1.000 abitanti nelle 20 Regioni italiane (anno 1986)



È possibile utilizzare sulla mappa diversi diagrammi, a seconda di ciò che si vuole rappresentare: solitamente si possono considerare sia grafici a barre sia grafici a torta, ma in entrambi i casi è necessario limitare il numero di elementi per avere una buona leggibilità della mappa

Talvolta nella cosiddette *carte tematiche* si sovrappone un cartodiagramma riferito alle diverse aree di un territorio al cartogramma riferito ad una particolare serie territoriale

29 – Esempio di carta tematica

Unità n° 03



Nella carta tematica qui riportata è rappresentata una fotografia al 2011 della ricerca di idrocarburi in Italia, a terra e in mare, includendo permessi, concessioni e richieste pendenti per Regione

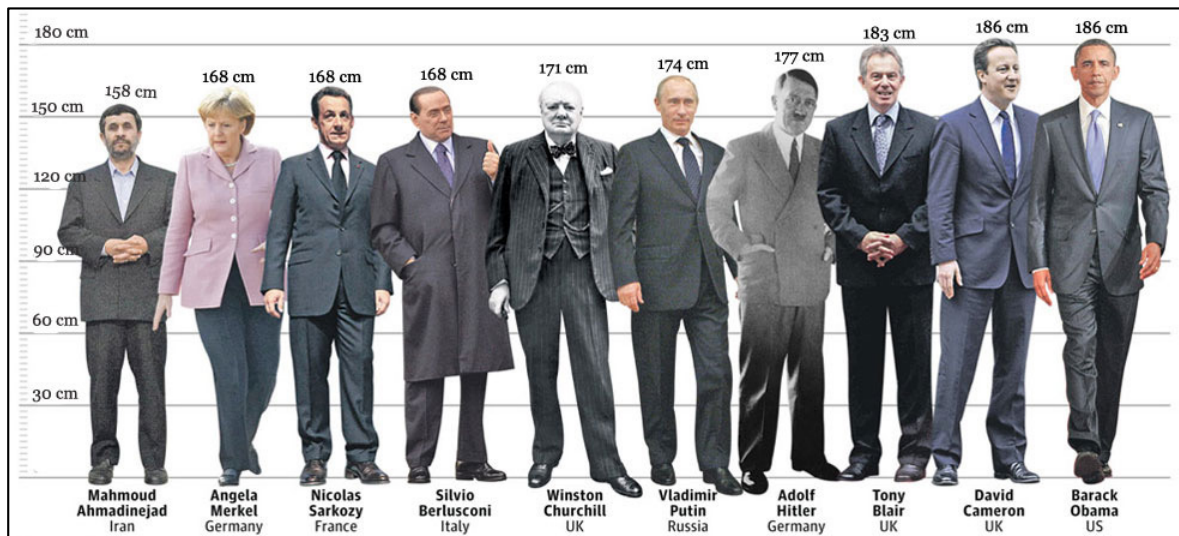
Questa carta è completa o manca qualche elemento necessario alla sua corretta interpretazione?

Che cosa aggiungereste per rendere la rappresentazione più informativa?

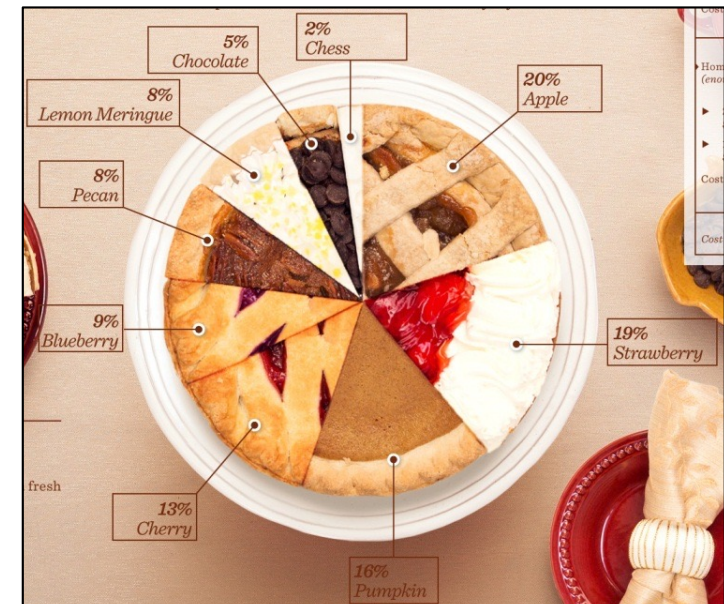
30 – Pictogrammi / ideogrammi

Unità n° 03

Con i **pictogrammi** è possibile combinare il potere informativo di una rappresentazione grafica a quello comunicativo dell'immagine: sono "divulgativi" ma spesso poco rigorosi...



L'altezza dei più famosi statisti di ieri e di oggi a confronto



Le torte preferite dagli americani nel Giorno del Ringraziamento

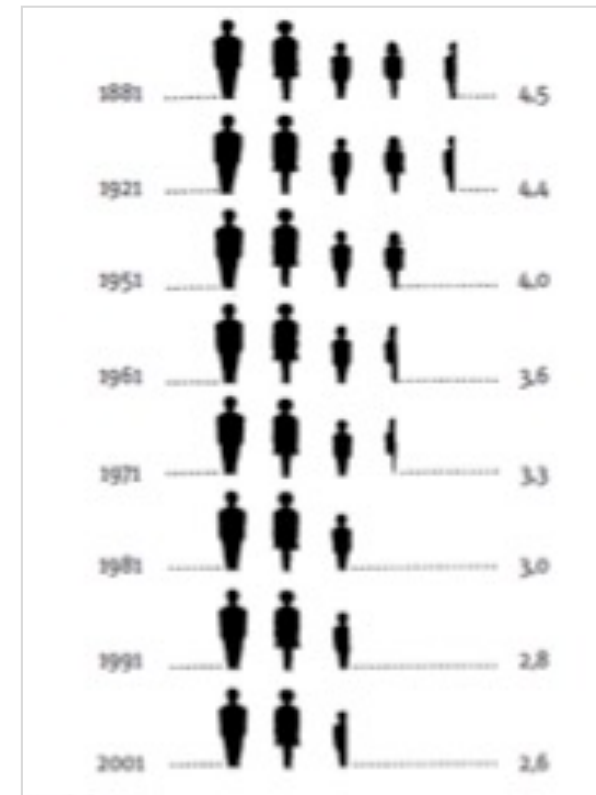
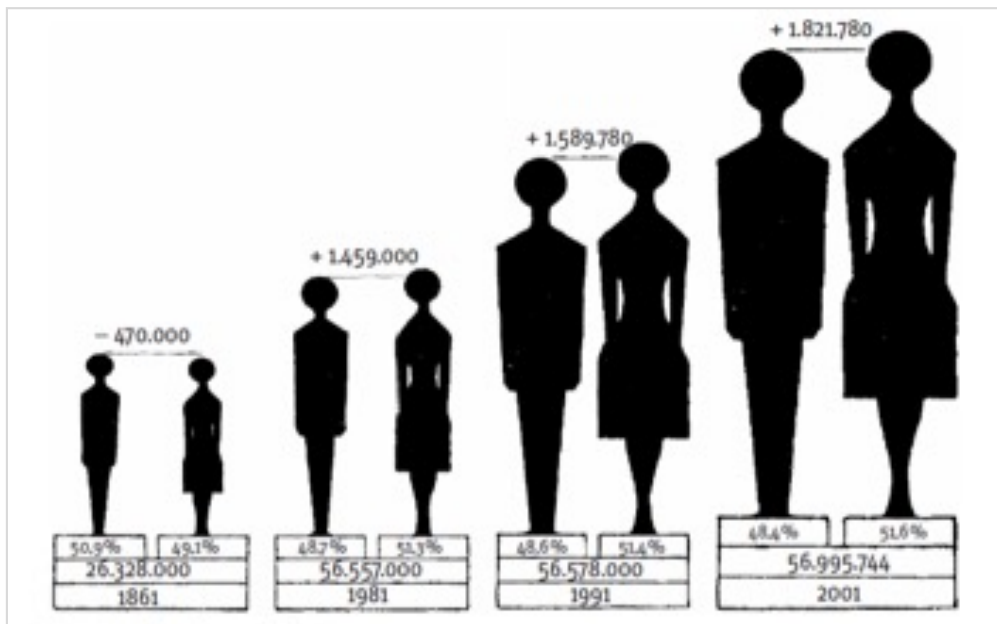
31 – Ambiguità dei pictogrammi

Unità n° 03

Questi diagrammi hanno il grave inconveniente di trarre in inganno con facilità

Numero medio di componenti per famiglia in Italia dal 1881 al 2001

Composizione della popolazione in Italia per genere dal 1861 al 2001



L'occhio coglie complessivamente non l'altezza di ogni figura ma la sua superficie

32 – Infogrammi/infografiche

Unità n° 03

Sono l'ultima "frontiera" nella rappresentazione dei dati: si considerano i diversi aspetti peculiari di un fenomeno complesso utilizzando diversi tipi di grafico con immagini e/o foto d'impatto



