

RAPPRESENTAZIONI GRAFICHE

La Statistica Descrittiva è volta alla *rappresentazione* di uno o più fenomeni reali su cui si sia condotto uno studio: la rappresentazione grafica di uno o più fenomeni statistici, infatti, non aggiunge e non toglie alcuna informazione ai dati oggetto di studio, in quanto i grafici non contengono dati diversi rispetto a quelli riportati nelle tabelle statistiche, ma consentono di cogliere, con maggiore evidenza visiva, la struttura e l'andamento di uno o più fenomeni, permettendo di confrontarli rapidamente, aiutando così l'analisi ed il ragionamento. Metodo grafico e metodo numerico, dunque, risultano complementari tra loro.

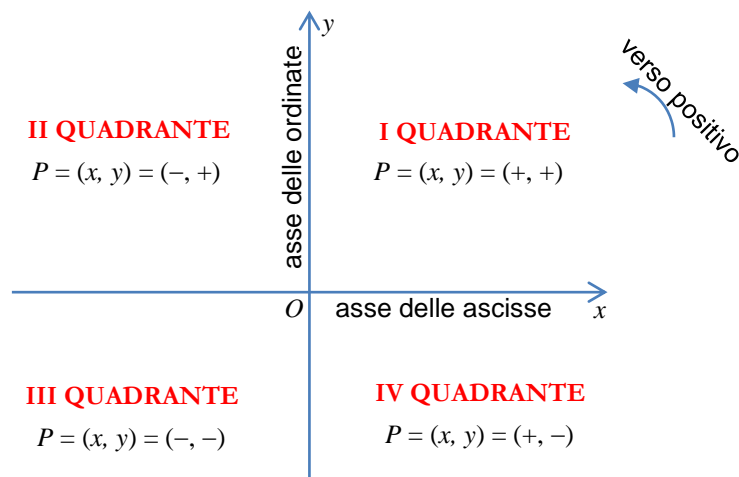
Nel seguito verranno riportate alcune delle molteplici rappresentazioni grafiche possibili.

- **Diagrammi in coordinate cartesiane.**

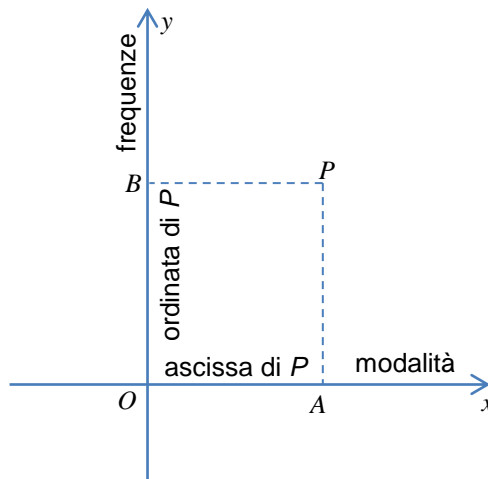
Sono usati generalmente per le serie storiche o temporali che misurano la dinamica di un certo fenomeno rilevata all'istante o in un periodo di tempo prefissato.

Si costruiscono tracciando una retta orizzontale x (*asse delle ascisse*), di verso positivo, fissando su di essa un'*origine* O e tracciando un'altra retta y (*asse delle ordinate*), anch'essa di verso positivo, perpendicolare alla prima e passante per l'origine. Le due rette, x ed y , formano un *sistema di assi cartesiani* che dividono il piano in quattro regioni, denominati *quadranti*, indicate, in senso antiorario, mediante numeri romani progressivi (I, II, III, IV). Il segno delle coordinate, quindi, dipende proprio dal quadrante in cui il punto in esame si trova, precisamente:

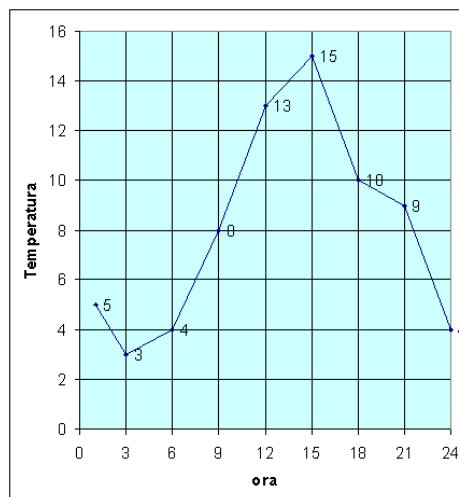
- primo quadrante: ascissa ed ordinata entrambe positive;
- secondo quadrante: ascissa negativa ed ordinata positiva;
- terzo quadrante: ascissa ed ordinata entrambe negative;
- quarto quadrante: ascissa positiva ed ordinata negativa.



Fissata, quindi, un'unità di misura per entrambi gli assi, si riportano, sull'asse delle ascisse, le modalità del carattere investigato (ove per *carattere* si intende il tipo di dati su cui si fonda lo studio e per *modalità* il modo in cui il carattere si manifesta) e su quello delle ordinate le corrispondenti frequenze. Per ogni coppia modalità-frequenza, pertanto, si individua un punto P del piano, le cui proiezioni, indicate con A (quella su Ox) e B (quella su Oy), ovvero le *coordinate del punto P* , rispettivamente l'*ascissa* e l'*ordinata*, non sono altro che le misure dei segmenti OA ed OB . Occorre osservare, a tal riguardo, però, che la Statistica utilizza esclusivamente il primo quadrante essendo, sia le modalità del carattere, sia le rispettive intensità o frequenze, quantità positive.



Il diagramma cartesiano di seguito riportato, ad esempio, illustra le temperature registrate nelle varie ore di una giornata in una determinata località:



I diagrammi in coordinate cartesiane, inoltre, possono essere:

- ✓ *per caratteri discontinui o a segmenti*: utilizzati per i caratteri discontinui (quantitativi o qualitativi), si avvalgono di segmenti rettilinei e paralleli aventi lunghezza proporzionale alle intensità o alle frequenze del carattere studiato;
- ✓ *per caratteri continui o logaritmici*: utilizzati quando la serie di osservazioni, avendo variazioni troppo ampie, non può essere riprodotta in scala naturale, si ottengono ponendo,

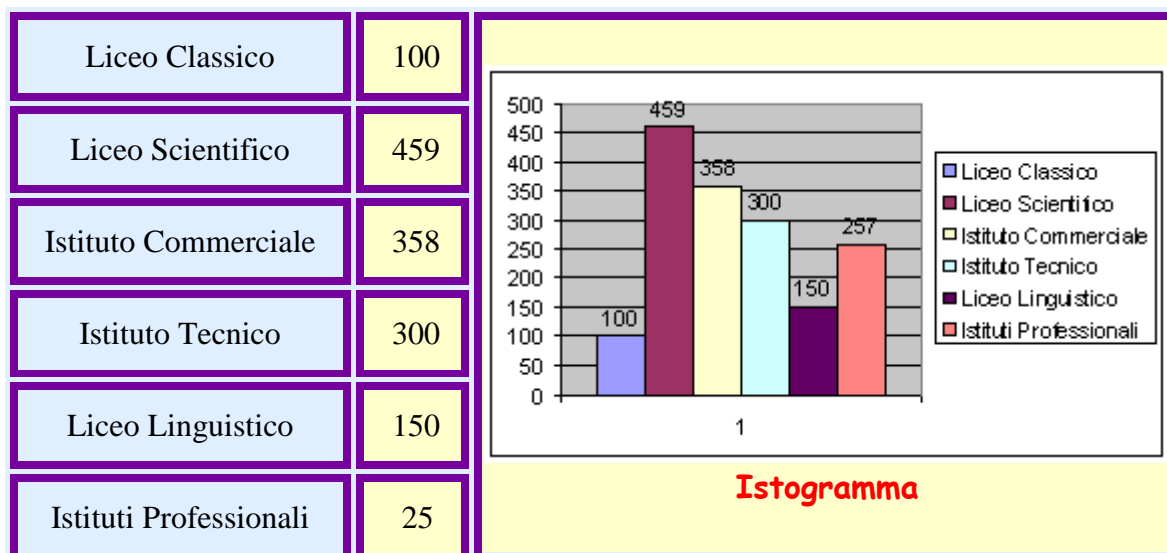
sull'asse x , le modalità del carattere in scala naturale e, sull'asse y , le relative frequenze in scala logaritmica, dando luogo al cosiddetto *poligono di frequenza*. Quando l'ampiezza delle classi, quindi, diviene infinitesima, la si approssima ad una curva regolare, detta *curva di frequenza*; se, invece, la distribuzione del carattere riporta, in corrispondenza delle diverse modalità, le frequenze relative, la rappresentazione prenderà il nome di *poligono di frequenze relative*. Il poligono delle frequenze cumulate, o *ogiva di frequenza*, dunque, rappresenta graficamente una distribuzione cumulativa di frequenza.

- **Istogrammi.**

Sono utilizzati per caratteri qualitativi ordinati e quantitativi continui con le modalità del carattere raggruppate in classi della medesima ampiezza, sì da evitare errori nella costruzione del grafico.

Costituiti da rettangoli adiacenti (l'adiacenza del rettangolo rappresenta proprio la continuità del carattere) che si sviluppano su un asse orizzontale, orientato e dotato di un'unità di misura (l'asse ha l'unità di misura del carattere e può essere inteso come l'asse delle ascisse), hanno la base di lunghezza pari all'ampiezza della corrispondente classe, l'altezza uguale o proporzionale alle frequenze (assolute o relative) delle unità statistiche di ciascuna classe, ovvero proporzionale al rapporto tra la frequenza assoluta associata alla classe e l'ampiezza della classe stessa, l'area uguale alle frequenze delle classi e la somma delle aree dei rettangoli uguale alla somma delle frequenze dei valori appartenenti alle varie classi. Nel caso in cui le classi non abbiano la stessa ampiezza, invece, il grafico si ottiene costruendo rettangoli contigui con altezza pari alla densità di frequenza, ovvero al rapporto tra la frequenza della classe e l'ampiezza corrispondente.

L'istogramma di seguito riportato, ad esempio, illustra i dati relativi agli alunni che frequentano le scuole superiori in una città italiana:



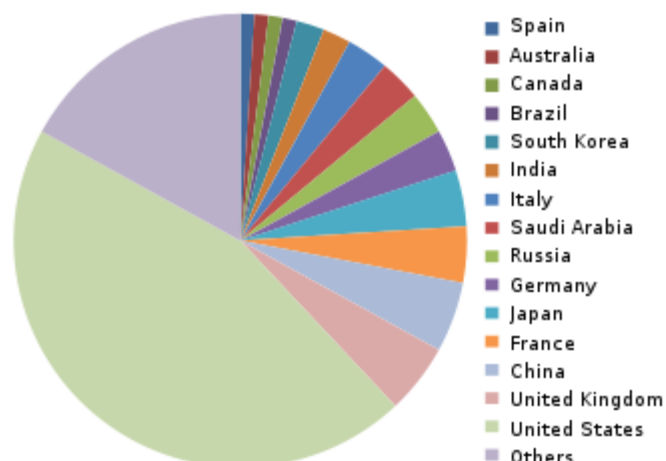
Gli istogrammi non devono essere confusi con i grafici a colonne (cfr. i **diagrammi a barre**): questi ultimi infatti, a differenza dei primi, hanno altezza proporzionale alla frequenza e sono costituiti da rettangoli separati tra loro.

Gli istogrammi, inoltre, possono essere:

- ✓ *istogrammi per classi di modalità di uguale ampiezza*: i valori del carattere sono raggruppati in classi di uguale ampiezza ed i rettangoli contigui, che si sviluppano dall'asse orizzontale, hanno le basi pari all'ampiezza della classe e l'altezza uguale o proporzionale alla frequenza della classe corrispondente;
- ✓ *istogrammi per classi di modalità di diversa ampiezza*: l'area dei rettangoli che insistono sulle classi è pari alla frequenza della classe corrispondente; essendo, poi, l'area del rettangolo pari al prodotto tra l'ampiezza di classe e l'altezza, ne segue che, in presenza di classi di diversa ampiezza, l'altezza dei rettangoli è uguale al rapporto tra la frequenza di classe e l'ampiezza corrispondente (*densità di frequenza*); in tal caso si congiungono i punti medi dei lati superiori dei rettangoli ottenendo una spezzata, denominata *poligono delle frequenze*, in cui ogni vertice corrisponde al valore centrale di una classe.

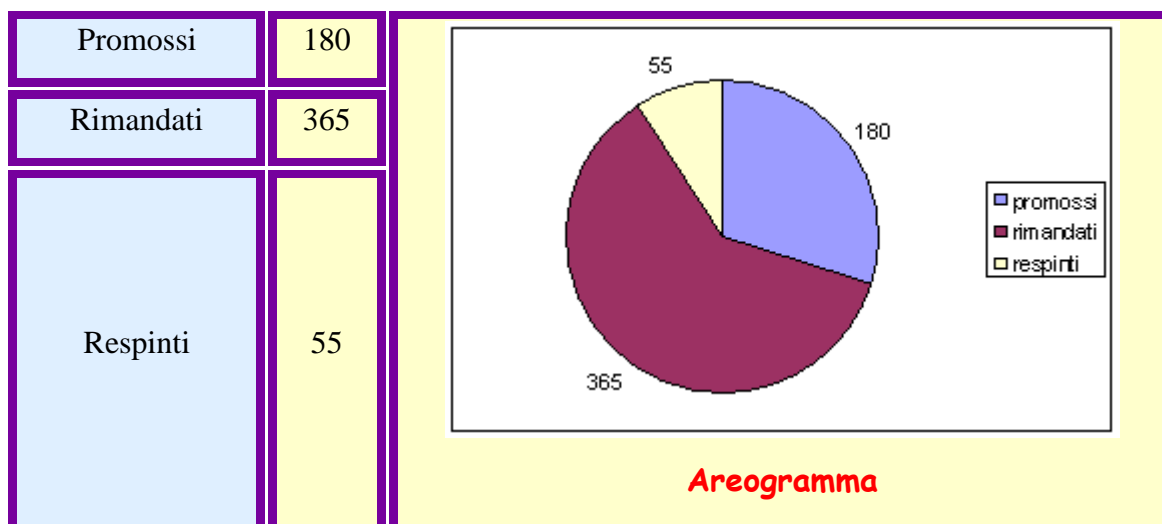
- **Diagrammi circolari o a torta o areogrammi.**

Utilizzati per i caratteri qualitativi, tali rappresentazioni mostrano le proporzioni delle parti rispetto all'insieme: le diverse percentuali di un'indagine statistica, visualizzate attraverso aree proporzionali di una figura geometrica piana, infatti, consentono di effettuare confronti fra diverse grandezze, senza necessariamente riferirsi ad un totale, al fine di evitare di stabilire, anche involontariamente, un ordine che non esiste nelle categorie, cosa che accadrebbe utilizzando un istogramma. Ciascun diagramma circolare, quindi, è caratterizzato dalla corrispondenza tra intensità o frequenze da una parte e superfici di cerchi o di settori circolari dall'altra, come riportato nella seguente figura:



Un diagramma circolare viene costruito, pertanto, dividendo un cerchio in spicchi le cui ampiezze angolari sono proporzionali alle classi di frequenza, ovvero le aree sono proporzionali alle frequenze o equivalentemente gli angoli al centro dei diversi settori circolari hanno ampiezza proporzionale alle frequenze percentuali.

Il diagramma circolare di seguito riportato, ad esempio, illustra il numero dei promossi, rimandati e respinti in un liceo frequentato da 600 allievi:



I diagrammi circolari, inoltre, possono essere:

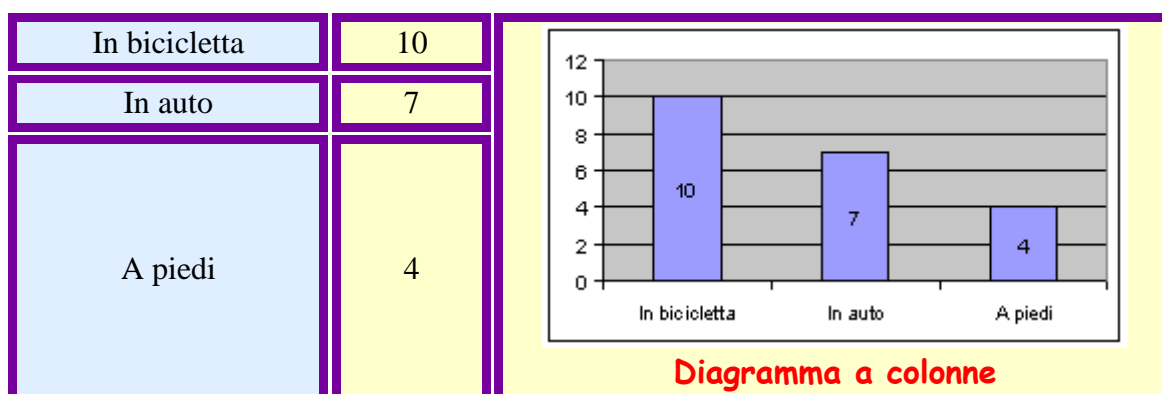
- ✓ *areogrammi per cerchi o grafici a bolle*: le misure delle superfici dei cerchi devono essere proporzionali alle intensità o frequenze da rappresentare. In tal caso, quindi, si assume un cerchio base che indica l'intensità o la frequenza più bassa e, per rappresentare un'intensità o una frequenza doppia o pari ad x volte quella minima, si ricorre a cerchi aventi area doppia o pari ad x volte quella minima;
- ✓ *areogrammi per settori circolari o grafici a torta*: l'area del cerchio esprime, in termini assoluti o percentuali, l'intensità o la frequenza totale, mentre i settori circolari, individuati dividendo l'intero angolo giro in parti proporzionali alle intensità o frequenze da rappresentare, esprimono le intensità o le frequenze delle varie modalità del carattere preso in considerazione.

- **Diagrammi a barre (o strisce) o ortogrammi.**

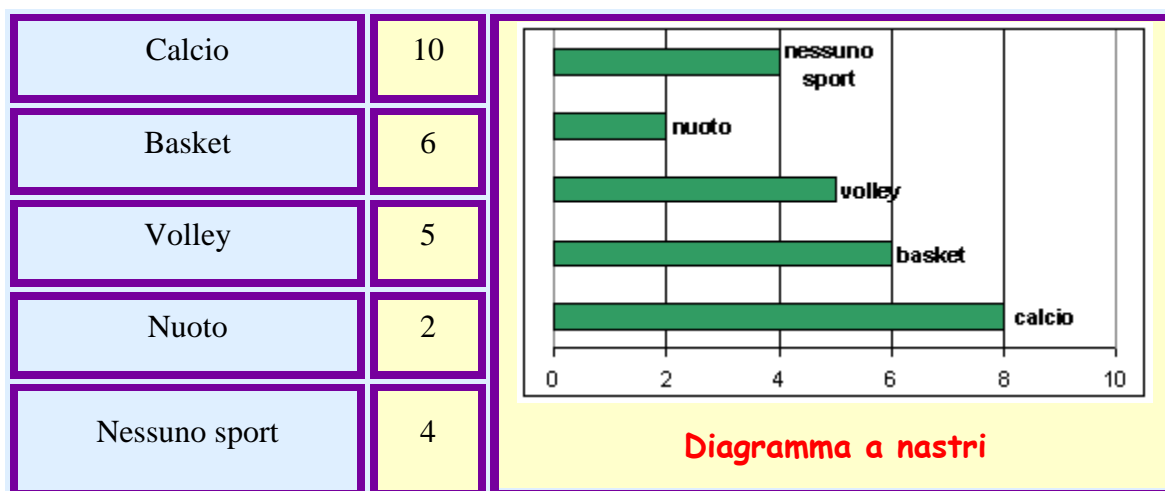
Utilizzati per rappresentare caratteri qualitativi ordinati rettilinei, sono costituiti da rettangoli, o barre, equidistanti, di larghezza arbitraria ma costante, di altezza proporzionale alla frequenza (assoluta o relativa) dell'attributo corrispondente ed in numero pari agli attributi del carattere. Tali diagrammi, che possono essere *a colonne* o *a nastri*, a seconda che si sviluppino rispettivamente in base ad un asse verticale o ad un asse orizzontale, ovvero che siano costituiti da una successione di colonne o di nastri, si servono, anche se non necessariamente, di un sistema di riferimento

cartesiano ortogonale, motivo per cui, se occorre rappresentare valori negativi, è possibile utilizzare anche il secondo ed il quarto quadrante di riferimento, a seconda che ci si trovi di fronte ad ortogrammi a nastri o a colonne. Gli ortogrammi, pertanto, sono caratterizzati dall'aver un solo asse (verticale nel caso "a colonne" ed orizzontale nel caso "a nastri") in scala graduata secondo l'unità di misura che si è scelta per rappresentare le frequenze o le intensità; sull'altro asse, invece, figurano le modalità (qualitative) equidistanti per convenzione.

Il diagramma a colonne di seguito riportato, ad esempio, illustra quali mezzi sono prevalentemente usati dai ragazzi di una classe per recarsi a scuola:

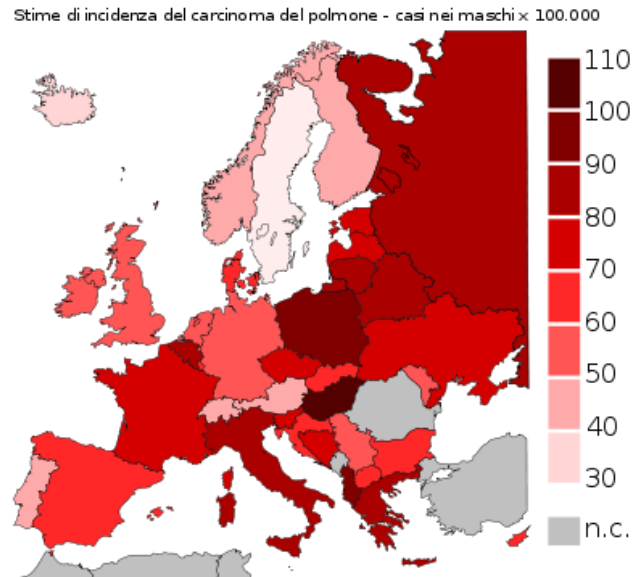


Il diagramma a nastri di seguito riportato, invece, illustra i dati relativi ad un'indagine svolta in una classe sugli sport praticati dagli allievi:



- **Cartogrammi.**

Tali rappresentazioni grafiche, che si avvalgono di carte geografiche o topografiche dei luoghi in cui si manifesta il fenomeno investigato, sono caratterizzate, a seconda delle varie zone, da una diversa colorazione, in base all'intensità del fenomeno stesso: vengono, infatti, aggiunti dei segni (a seconda dell'intensità), tratteggiate le zone interessate, utilizzate linee concentriche simili a curve di livello o, più semplicemente, diverse gradazioni del medesimo colore, facendo corrispondere ad ogni colore una ben precisa intensità, come illustrato nel seguente diagramma:



Al cartogramma, che consente di mettere in evidenza le relazioni esistenti tra le varie zone geografiche a maggiore o minore intensità del fenomeno, quindi, vengono spesso affiancati, per un'ulteriore chiarezza, altri grafici contenenti informazioni aggiuntive, in modo tale che i dati inseriti siano maggiormente comprensibili a tutti.

Il cartogramma di seguito riportato, fornito dall'Istituto Centrale di Statistica, illustra il numero di biglietti per cinematografo venduti, per abitante, nell'anno 1975:



- Ideogrammi o pittogrammi.

In tale rappresentazione grafica, molto approssimativa ma di comprensione abbastanza facile ed immediata, anche per i non esperti, e pertanto utilizzata per scopi pubblicitari e come mezzo di divulgazione, i dati vengono rappresentati attraverso oggetti che ricordano il fenomeno da studiare, ovvero l'unità di misura viene sostituita da un'icona attinente al fenomeno oggetto di analisi.

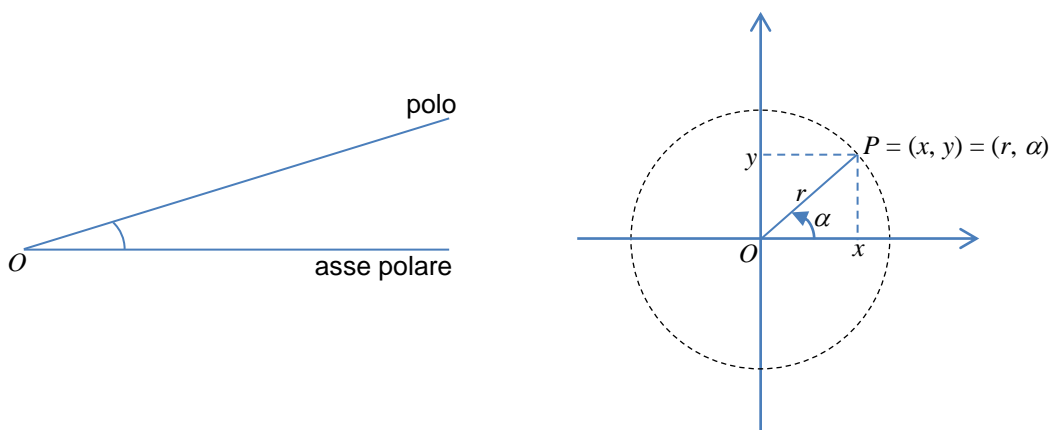
L'ideogramma di seguito riportato, ad esempio, illustra i dati relativi alla produzione, in un certo anno, di quattro case automobilistiche:



- **Diagrammi in coordinate polari.**

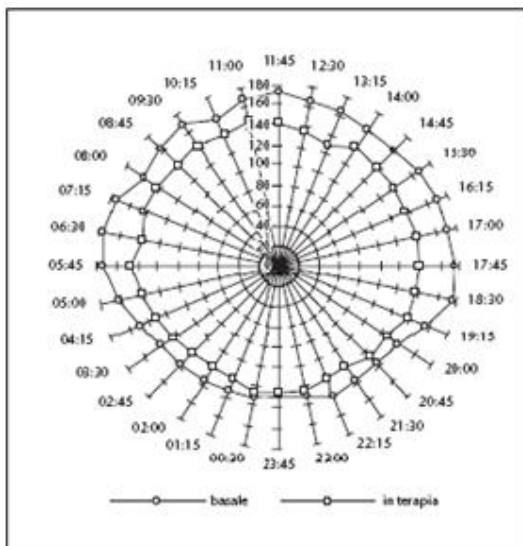
Utilizzati per la rappresentazione di particolari serie storiche, dette *cicliche*, in quanto le modalità del carattere si ripetono dopo un certo periodo di tempo (giornaliere, settimanali, mensili, ...), quali i fenomeni stagionali (quantità di pioggia, temperatura media, numero di matrimoni celebrati nell'ultimo anno, ...), che ben si prestano a questo scopo, tali diagrammi, che mostrano il cambiamento e le frequenze di dati in relazione, sia ad un *polo* (o *punto centrale*), che tra loro, si costruiscono attraverso le seguenti quattro fasi:

- I) dal punto centrale, ovvero dal polo, si traccia una serie di cerchi concentrici, la cui rispettiva distanza dal centro esprime la misura dell'intensità del fenomeno;
- II) si divide l'angolo giro in tante parti quante sono le modalità della serie;
- III) segnando dei punti nei cerchi, si individuano le modalità e la frequenza del fenomeno;
- IV) sia che il fenomeno sia discreto sia che esso sia continuo, si uniscono i punti con segmenti di retta,



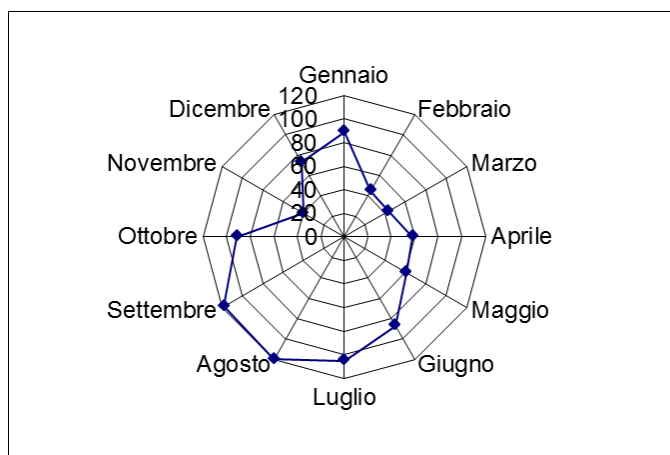
dove (x, y) rappresentano le *coordinate cartesiane* del punto P e (r, α) le corrispondenti *coordinate polari*.

Nel seguente diagramma, ad esempio, è riportato l'andamento medio giornaliero della pressione arteriosa sistolica di 45 individui, prima e dopo un trattamento antiipertensivo:



Nel successivo diagramma, invece, è riportato l'andamento del numero di viaggi distribuiti nei vari mesi dell'anno secondo la seguente tabella:

MESI	NUMERO VIAGGI
Gennaio	89
Febbraio	45
Marzo	43
Aprile	58
Maggio	61
Giugno	87
Luglio	105
Agosto	120
Settembre	118
Ottobre	91
Novembre	40
Dicembre	73



- **Grafici a ramo e foglia.**



Tale tipologia di rappresentazione, detta anche *Stem and leaf plot*, introdotta da John Wilder Tukey (16 giugno 1915 – 26 luglio 2000), usata per caratteri quantitativi, consiste in una forma “ibrida” tra una tabella di dati ed il corrispondente grafico: presentando i valori numerici delle distribuzioni in ogni singola frequenza, infatti, permette di esaminare le caratteristiche di distribuzioni di frequenza costituite da uno scarso numero di osservazioni.

Per realizzare tale grafico occorre dividere, dapprima, ogni singolo dato in due parti, una più significativa, detta *ramo*, ed una meno significativa, detta *foglia*, e successivamente costruire una tabella a due colonne, riportando, nella prima, la cifra più significativa (*centinaia e decine*) e, nella seconda, quella meno significativa (*unità*).

Ad esempio, se si considerano le seguenti cifre:

12 13 32 33 45 56

è possibile, dapprima, dividerle, separando le decine dalle unità, secondo la seguente tabella:

RAMO	FOGLIA
1	2
1	3
3	2
3	3
4	5
5	6

e poi raggruppare i rami posizionando le relative foglie una accanto all'altra, come di seguito riportato:

RAMO	FOGLIA
1	23
3	23
4	5
5	6

Se ora, invece, si considera la seguente serie di dati:

12.3 13.5 12.7 14.9 14.8 15.7 19.5 11.4 17.6 18.8 17.6 18.4 16.2 16.3 16.2

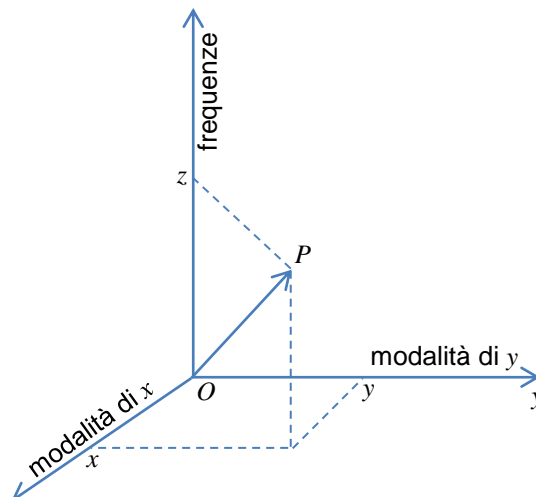
allora la rappresentazione ramo-foglia sarà data da:

RAMO	FOGLIA
11	4
12	37
13	5
14	89
15	7
16	223
17	66
18	48
19	5

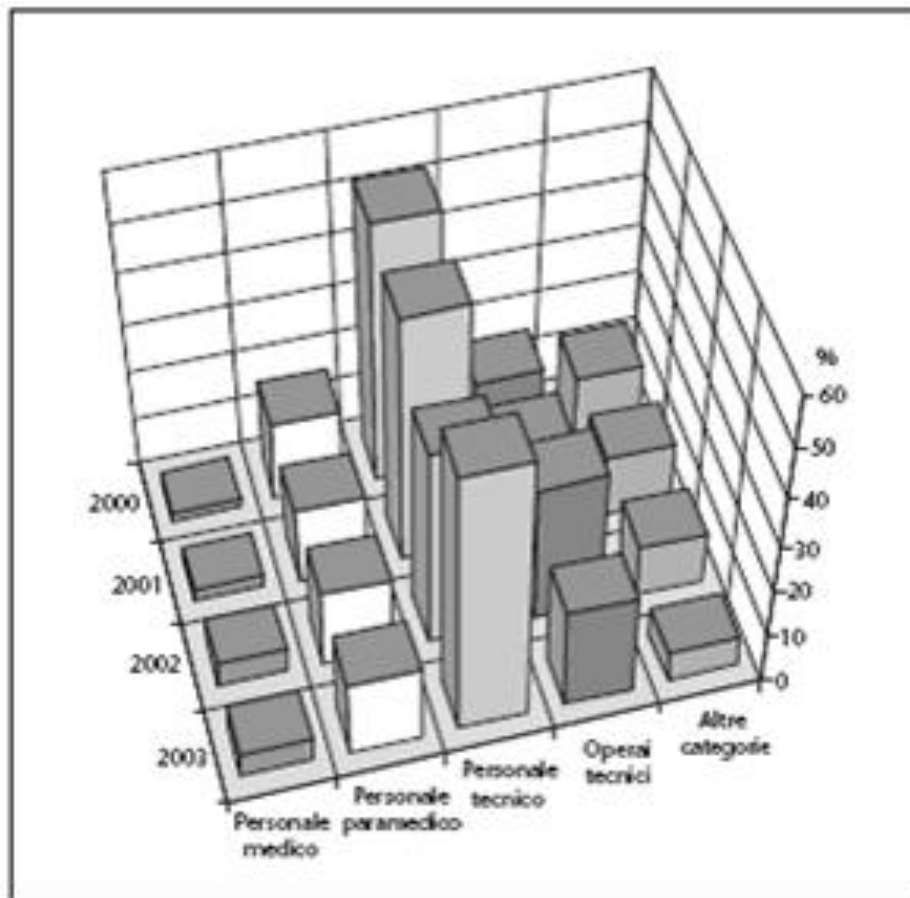
- **Stereogrammi o diagrammi in 3D.**

È la più nota rappresentazione grafica di una distribuzione doppia dalla quale si desumono le frequenze riguardanti classi di modalità dei due caratteri x ed y .

Per ottenere lo stereogramma si riportano su un asse le classi di modalità relative ad x , su un altro asse quelle relative ad y e su un terzo asse, disposto perpendicolarmente al piano cartesiano definito dagli altri due assi, le frequenze corrispondenti, ottenendo alla fine un insieme di parallelepipedi contigui in cui le dimensioni delle basi sono pari all'ampiezza delle classi ed i volumi uguali o proporzionali alle frequenze:



Nel seguente diagramma, ad esempio, sono riportati i numeri del personale medico, personale paramedico, personale tecnico, operai tecnici e personale di altre categorie di un ospedale di una determinata città:



Nel diagramma successivo, invece, sono riportati i laureati in Economia secondo il voto di laurea e secondo il tempo impiegato per trovare lavoro, in base alla seguente tabella:

VOTO	TEMPO IMPIEGATO			
	6 mesi	12 mesi	18 mesi	24 mesi
81-85	6	5	3	3
86-90	14	5	3	4
91-95	8	2	1	5
96-100	11	8	3	2
101-105	2	3	2	4
106-110	11	10	2	8
110 e lode	6	2	0	2

