

Introduzione alla STATISTICA CON EXCEL

Silvio Vitale
Istat Genova



INDICE

1. **Nota introduttiva**
2. **Lavorare con Excel**
 - 2.1 **L'ambiente di lavoro**
 - 2.2 **Accedere alle diverse funzioni e comandi**
 - 2.3 **Installazione di funzioni avanzate**
3. **Introduzione alla Statistica descrittiva**
 - 3.1 **l'unità statistica**
 - 3.2 **i caratteri dell'unità statistica**
4. **Rappresentazione dei dati**
 - 4.1 **le distribuzioni statistiche (cenni)**
 - 4.2 **distrib. di frequenze per caratteri quantitativi**
 - 4.2.1 **frequenze assolute e relative (inizio esercizi)**
 - 4.3 **costruire una Tabella di frequenza**
 - 4.4 **le rappresentazioni grafiche**
 - 4.4.1 **diagramma a barre**
 - 4.4.2 **istogramma**
 - 4.4.3 **diagramma degli estremi e quartili**
 - 4.4.4 **diagramma circolare**
5. **Indici di posizione**
 - 5.1 **Media**
 - 5.2 **Moda**
 - 5.3 **Mediana**
 - 5.4 **Quartili**
6. **Indici di dispersione**
 - 6.1 **Varianza**
 - 6.2 **Devianza standard**
 - 6.3 **Ampiezza**
 - 6.4 **Ampiezza interquartile**
7. **Dati bivariati**
 - 7.1 **Diagramma di dispersione**
 - 7.2 **Coefficiente di correlazione**
 - 7.3 **Retta di regressione**

GLOSSARIO

versione PDF per la stampa

crediti

BIBLIOGRAFIA

1. Nota introduttiva

Questo prodotto propone alcune esercitazioni sull'uso di Excel per la statistica.

Oggi la statistica è sotto l'occhio di tutti: è in continuo aumento l'utilizzo di dati statistici nei giornali e nei media per avvalorare o criticare scelte e posizioni di vario genere; ma per il cittadino comune non è quasi mai facile orientarsi, comprendere messaggi imprecisi, parziali o, addirittura sbagliati.

La statistica è uno strumento utile al cittadino consapevole che vuole capire i fenomeni che lo riguardano (naturali, sociali, economici e politici) svolgendo, così, anche un'attività di monitoraggio e controllo: nella vita di tutti i giorni è sempre più importante saper valutare le numerose informazioni statistiche che giungono da diverse fonti, comprenderne il significato e riconoscerne l'attendibilità.

La capacità di avvicinarsi al dato statistico in maniera critica può essere acquistata con un processo formativo mirato, che consenta di riconoscere l'informazione statistica ufficiale e la metodologia che ne è alla base.

La scuola è la sede naturale dove sviluppare questo processo formativo.



Salvare la cartella “corso excel” in “C:/" (inviare un “collegamento sul desktop” del file corso excel.ppt per facilitare i successivi accessi).

Durante il corso sarà possibile attivare un foglio di calcolo interattivo (link foglio.xls) ed eseguire le esercitazioni suggerite.

Attenzione: salvataggio delle esercitazioni

per salvare le esercitazioni svolte (ripeterle e/o eseguirle in un secondo momento) utilizzare il comando “salva con nome” in quanto il foglio di calcolo proposto è “temporaneo”.

È disponibile una [versione del corso in formato PDF](#) (se ne consiglia la stampa).

Questo lavoro è stato predisposto tenendo conto dell'esperienza dell'Istituto di statistica del Portogallo (www.ine.pt).

La [Bibliografia](#) oltre ad alcuni testi di riferimento utilizzati nella compilazione di questo prodotto, indica alcuni siti che forniscono ulteriori informazioni e approfondimenti.

Il comando “[Home](#)” riporta alla pagina di Avvio

Il comando “[Torna](#)”, quando presente, riporta all'ultima pagina precedentemente visualizzata.

Gli studenti e i docenti sono invitati a far pervenire osservazioni e commenti.



2. Lavorare con Excel

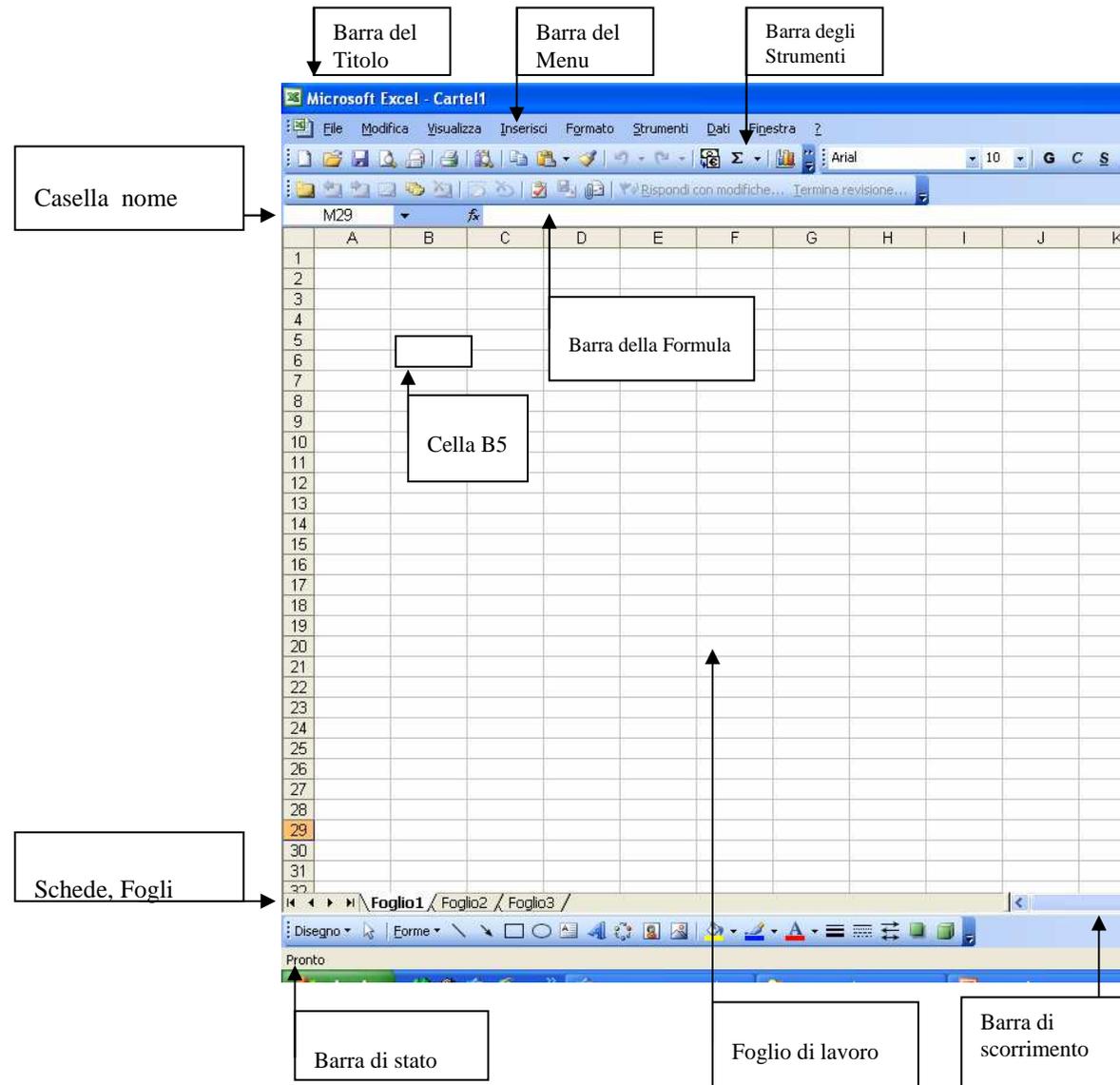
Excel è un programma di calcolo elaborato dalla Microsoft per l'ambiente Windows. Così come la maggior parte delle applicazioni informatiche, anche Excel si è via via perfezionato adattandosi alle nuove capacità dei PC.

Questo prodotto ha come base la versione 2000.

Per facilitare la lettura, diversi comandi e funzioni sono stati evidenziati con il carattere in grassetto.

Per chi possiede la versione in inglese del programma, sono riportati, [*in corsivo* e tra parentesi quadra], i comandi/funzioni corrispondenti.

2.1. L'ambiente di lavoro: il foglio di lavoro



Home

2.2 Accedere alle diverse funzioni e comandi:

Excel presenta diverse opzioni:

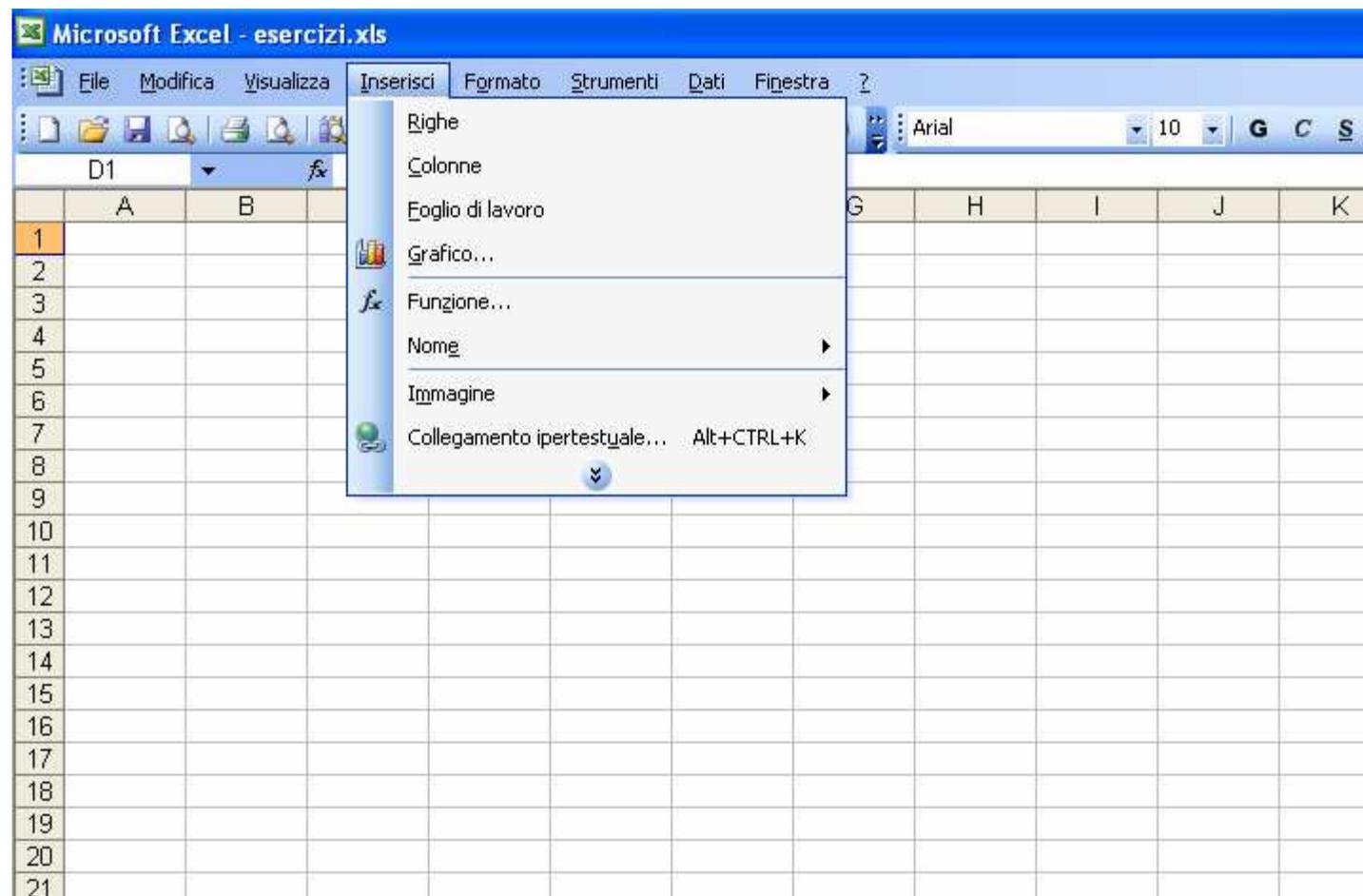
- 1) utilizzare i menù o sottomenù che possono essere selezionati dalla **Barra del menù**: cliccare con il mouse o utilizzare il tasto di scelta rapida. ([visualizza esempio](#))
- 2) utilizzare la **Barra degli strumenti**: cliccare con il mouse sull'icona corrispondente ([visualizza esempio](#))
- 3) utilizzare la **Barra della formula**: selezionare la cella di destinazione e cliccare sul tasto **modifica formula** per far comparire il menù delle funzioni ([visualizza esempio](#))

*Qui facciamo soprattutto riferimento alle funzioni statistiche; a questo proposito si sottolinea che, attraverso ciascuna delle tre modalità qui di fianco esposte, tra altri comandi e funzioni che è possibile attivare si potrà accedere al comando **Incolla Funzione [Function Wizard]**; le diverse funzioni sono raggruppate in 11 categorie. ([visualizza le funzioni](#))*

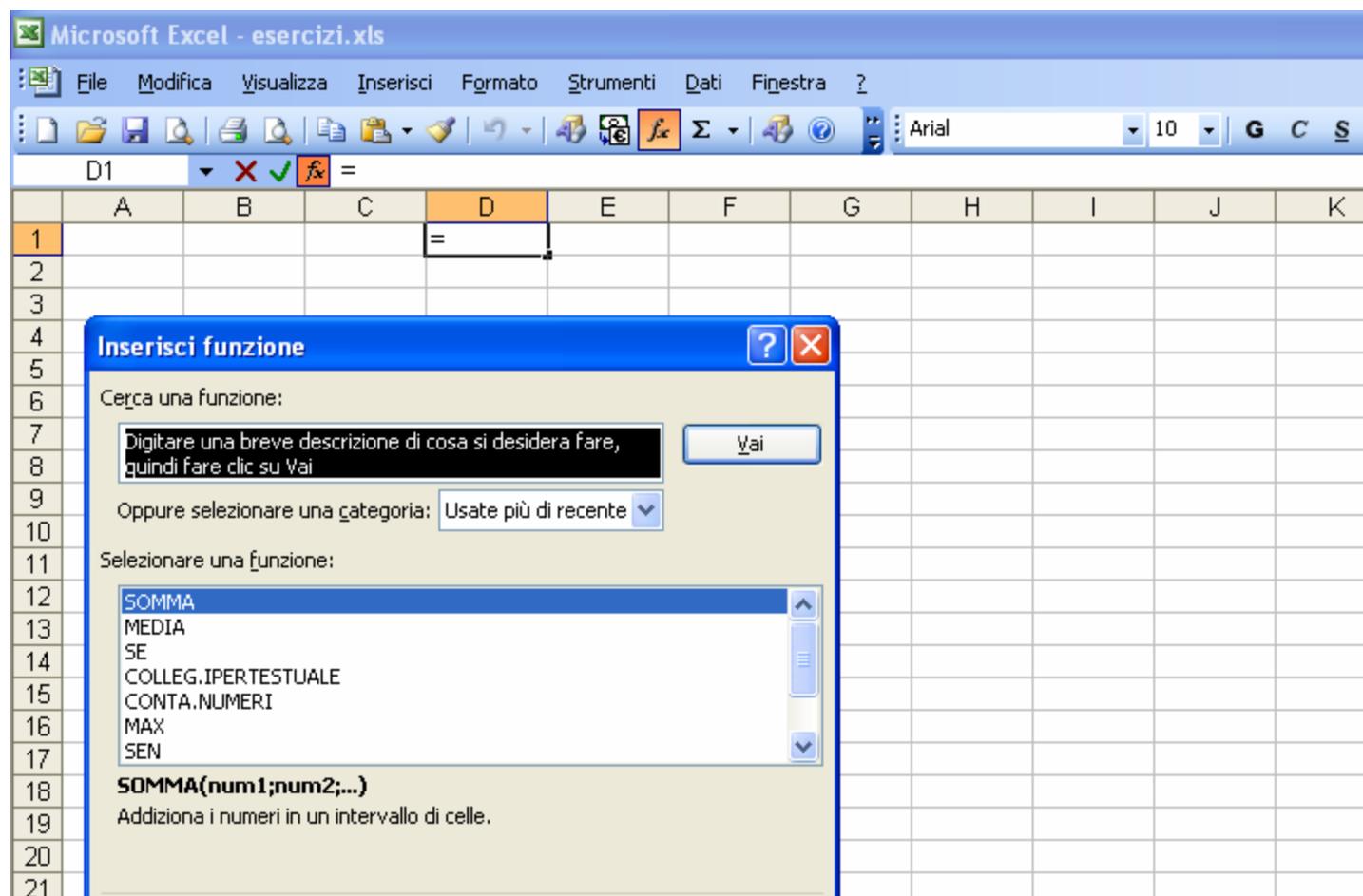
- 4) Per i più esperti: **digitare la funzione (formula) nella Barra della formula** o direttamente all'interno della cella di destinazione, ([visualizza esempio](#))

[Home](#)

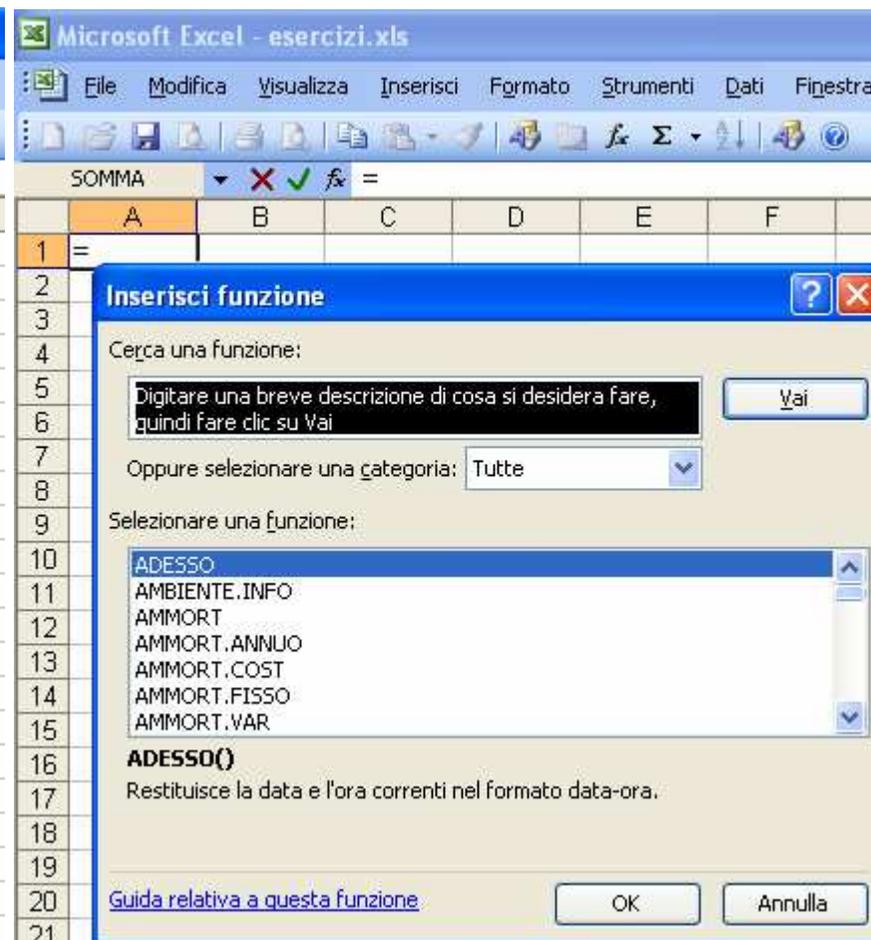
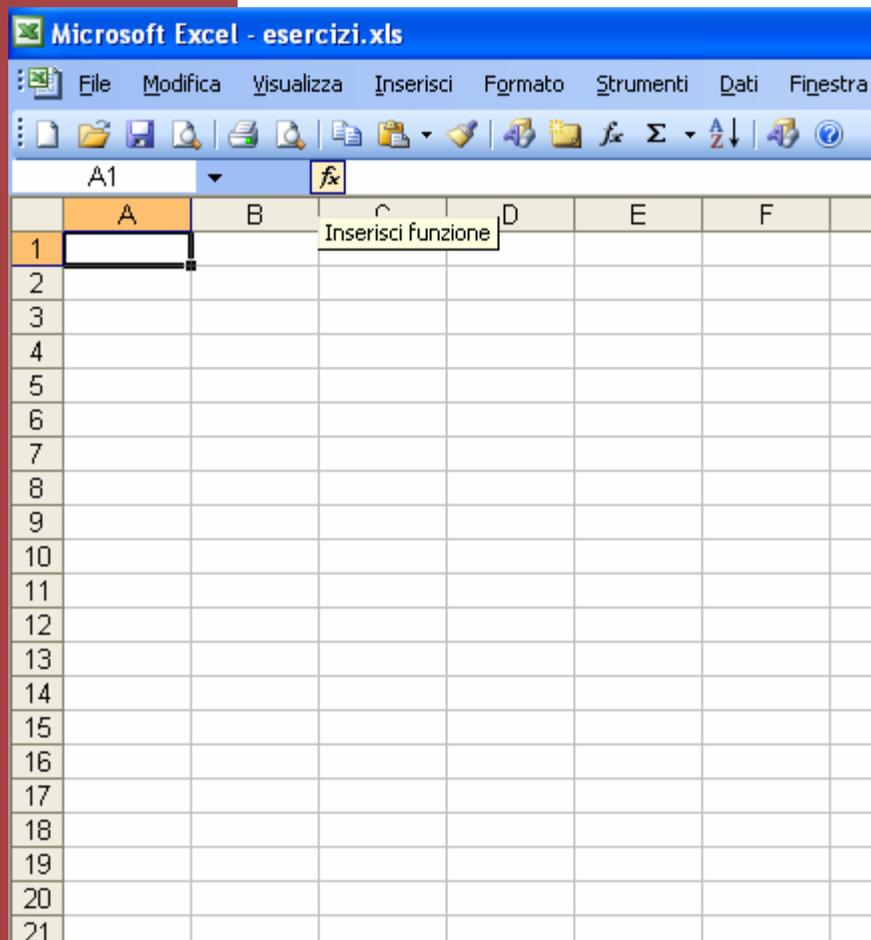
1) Utilizzare i **menù** o **sottomenù** che possono essere selezionati dalla **Barra del menù**:
cliccare con il mouse o utilizzare il tasto di scelta rapida.



2) Utilizzare la **Barra degli strumenti**: cliccare con il mouse sull'icona corrispondente.



3) Utilizzare la **Barra della formula**: selezionare la cella di destinazione e cliccare sul tasto **Inserisci funzione** per far comparire il menù delle **funzioni**.



Quando viene attivato, il comando **Incolla Funzione [Function Wizard]**, oltre alla possibile selezione di **“Usate più di recente”** e **“Tutte”**, presenta le diverse funzioni raggruppate nelle seguenti undici (11) categorie:

Categoria di funzione	(Function Category)
Finanziarie	Financial
Data e Ora	Date & Time
Matematiche e trig.	Math & Trig
Statistiche	Statistical
Ricerca e riferimento	Lookup & Reference
Database	Database
Testo	Text
Logiche	Logical
Informative	Information
Definite dall'Utente	User defined
Ingegneristiche	Engineering

[Home](#)



Per ogni categoria scelta il menù presenta le diverse funzioni, con una breve descrizione della funzione scelta e della sua sintassi.

The screenshot shows the Microsoft Excel interface with the 'Argomenti funzione' dialog box open for the 'MEDIA' function. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I
1	atleta	Tempi (min/sec)							
2	gianni	12,1							
3	maria	8,9							
4	luca	16,2							
5	fabio	8,2							
6	angela	9,8							
7	michele	15,1							
8	mirko	14,5							
9	daniela	13,4							
10	marco	14,7							
11	giuseppe	7,5							
12	stefano	8,8							
13	laura	12,4							
14	chiara	16,1							
15	davide	15,2							
16	luigi	13,5							
17	monica	14,6							
18	sara	15,5							
19	giovanni	7,8							
20	aldo	12,5							
21	serena	13,2							

The 'Argomenti funzione' dialog box shows the following details:

- Function: MEDIA
- Num1: B2:B12 (displayed as {12,1|8,9|16,2|8,2|})
- Num2: (empty, displayed as = num)
- Result: = 11,74545455
- Description: Restituisce la media aritmetica degli argomenti (numeri, nomi o riferimenti contenenti numeri).
- Syntax: Num1: num1;num2;... sono da 1 a 30 argomenti numerici di cui si calcola la media.
- Formula result: Risultato formula = 11,74545455
- Buttons: OK, Annulla
- Link: Guida relativa a questa funzione

[ritorna al corso \(senza chiudere\)](#)

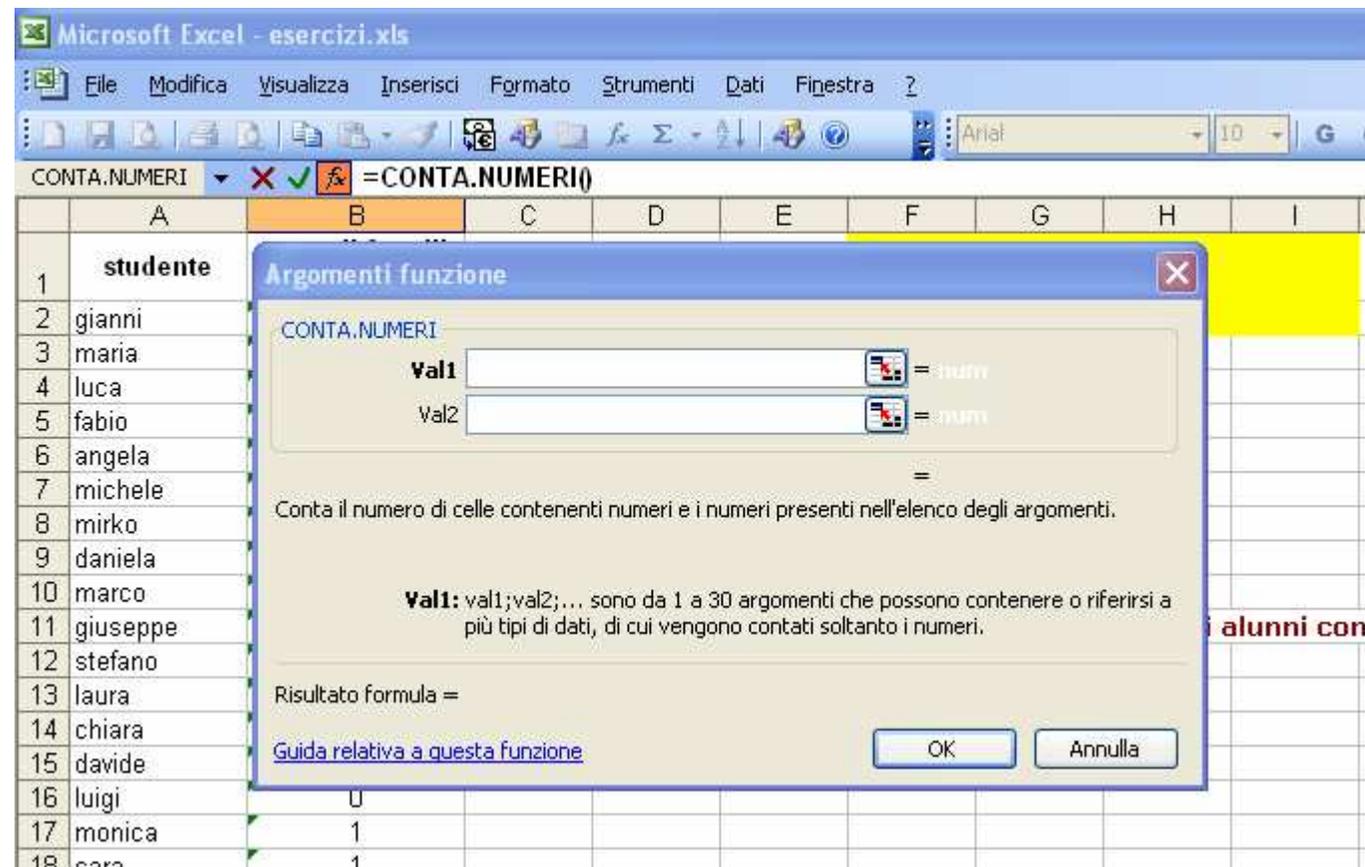


[Home](#)



Dopo aver selezionato la funzione desiderata, un secondo menù permetterà di definire i diversi argomenti della funzione.

Quando avrete acquisito maggiore familiarità con le diverse funzioni, potrete digitarle direttamente senza utilizzare il menù.



Microsoft Excel - esercizi.xls

File Modifica Visualizza Inserisci Formato Strumenti Dati Finestra ?

CONTA.NUMERI =CONTA.NUMERI()

	A	B	C	D	E	F	G	H	I
1	studente								
2	gianni								
3	maria								
4	luca								
5	fabio								
6	angela								
7	michele								
8	mirko								
9	daniela								
10	marco								
11	giuseppe								
12	stefano								
13	laura								
14	chiara								
15	davide								
16	luigi	U							
17	monica	1							
18	cara	1							

Argomenti funzione

CONTA.NUMERI

Val1 = num

Val2 = num

=

Conta il numero di celle contenenti numeri e i numeri presenti nell'elenco degli argomenti.

Val1: val1;val2;... sono da 1 a 30 argomenti che possono contenere o riferirsi a più tipi di dati, di cui vengono contati soltanto i numeri.

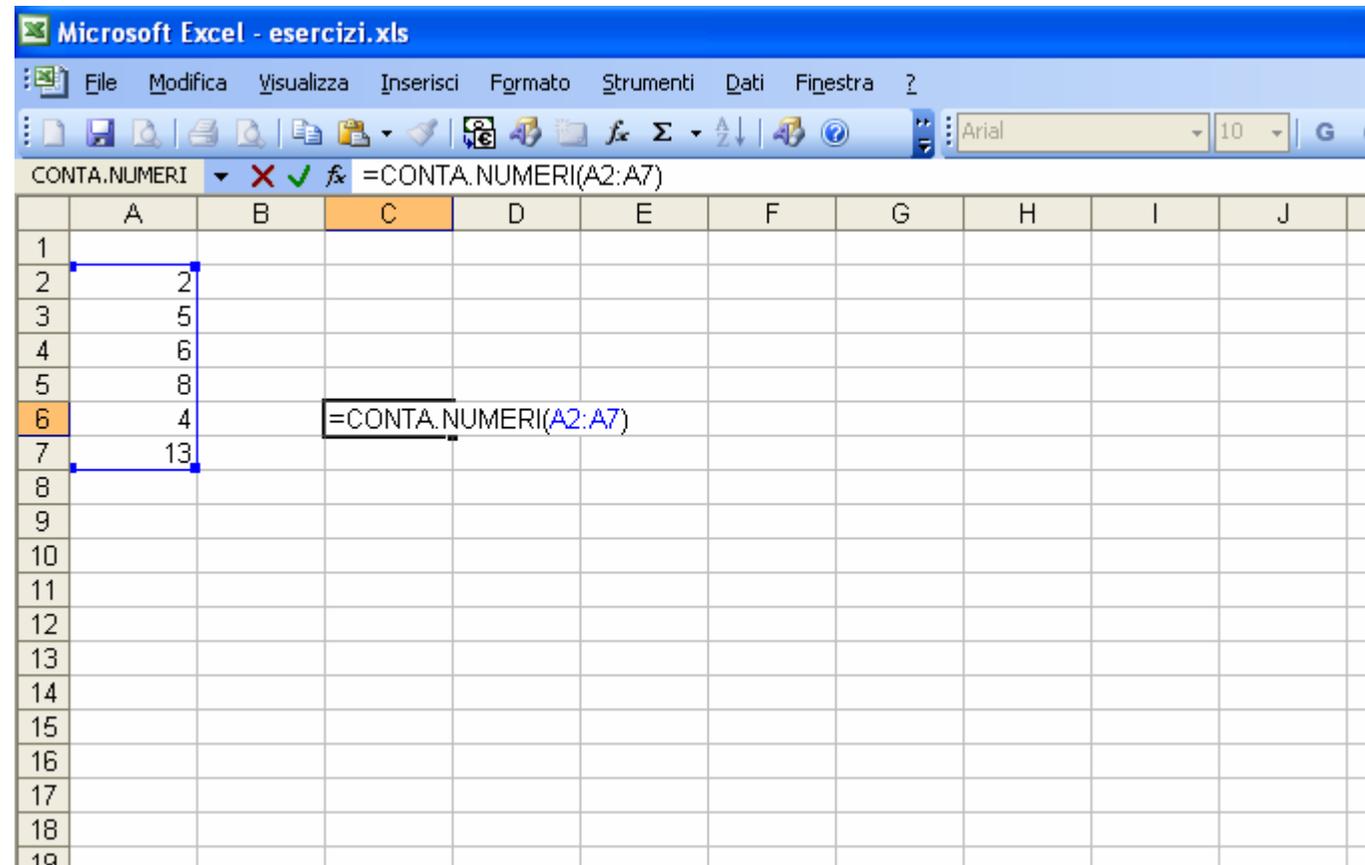
Risultato formula =

[Guida relativa a questa funzione](#)

OK Annulla



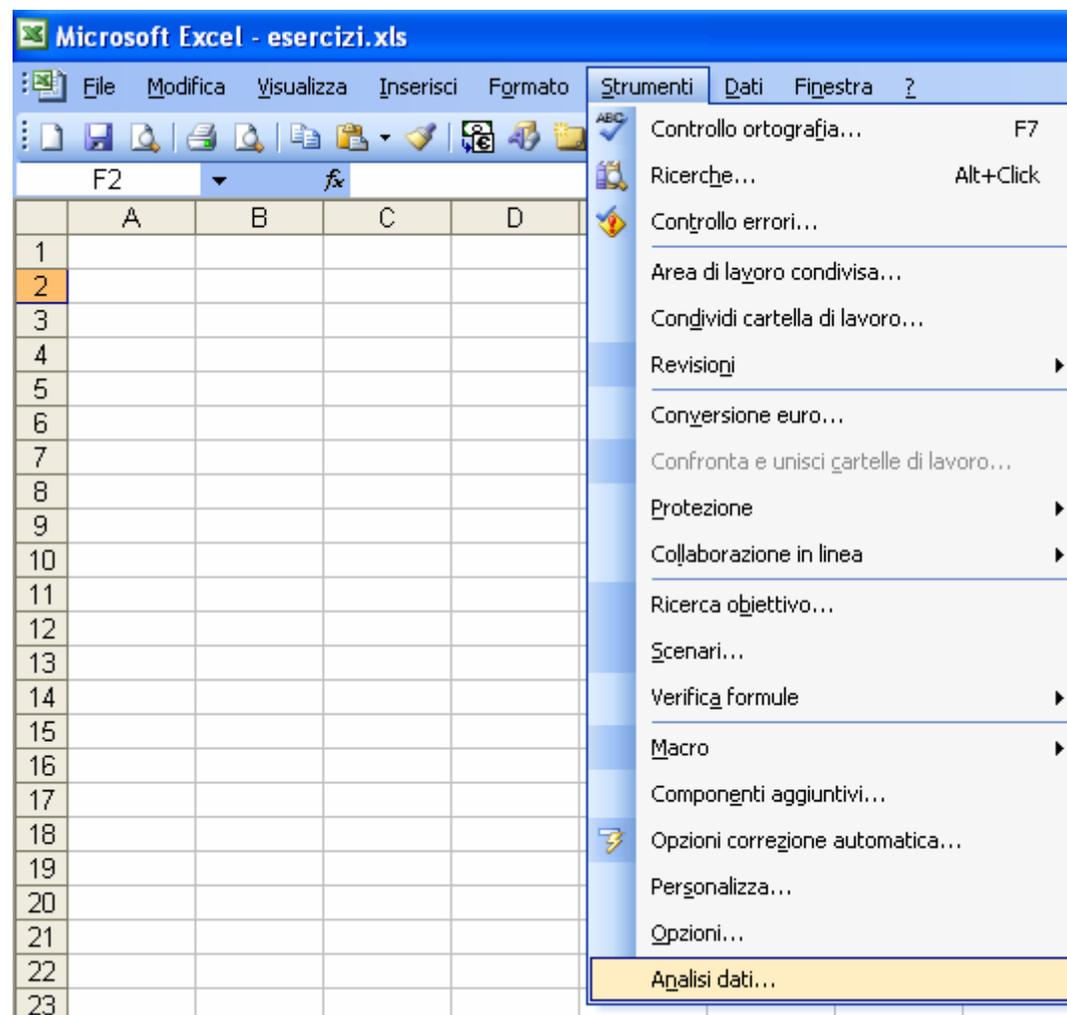
4) È anche possibile digitare la funzione (formula) direttamente nella **Barra della formula** come pure all'interno della **cella di destinazione**.



[Torna](#)

[Home](#)

2.3 Installazione delle funzioni avanzate.



Excel presenta diverse funzioni e procedure avanzate per l'analisi statistica dei dati, che si trovano sotto il comando **Analisi Dati** [Data analysis].

Prima di procedere con la presentazione delle diverse funzioni statistiche, verificare che queste siano già installate nel computer:

Come si fa:

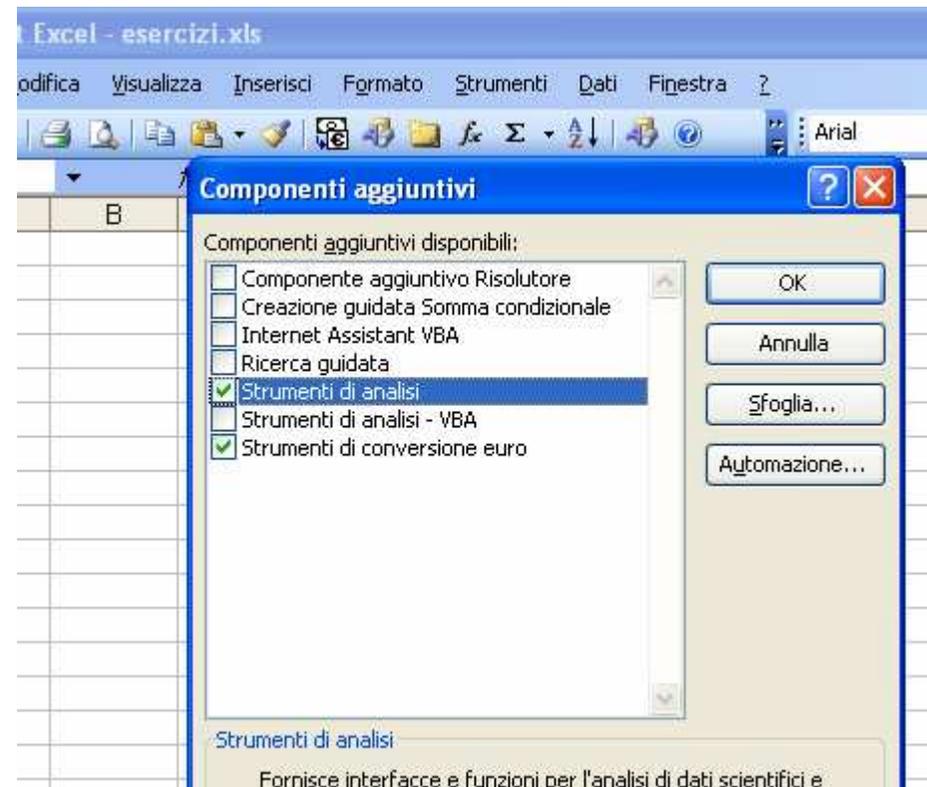
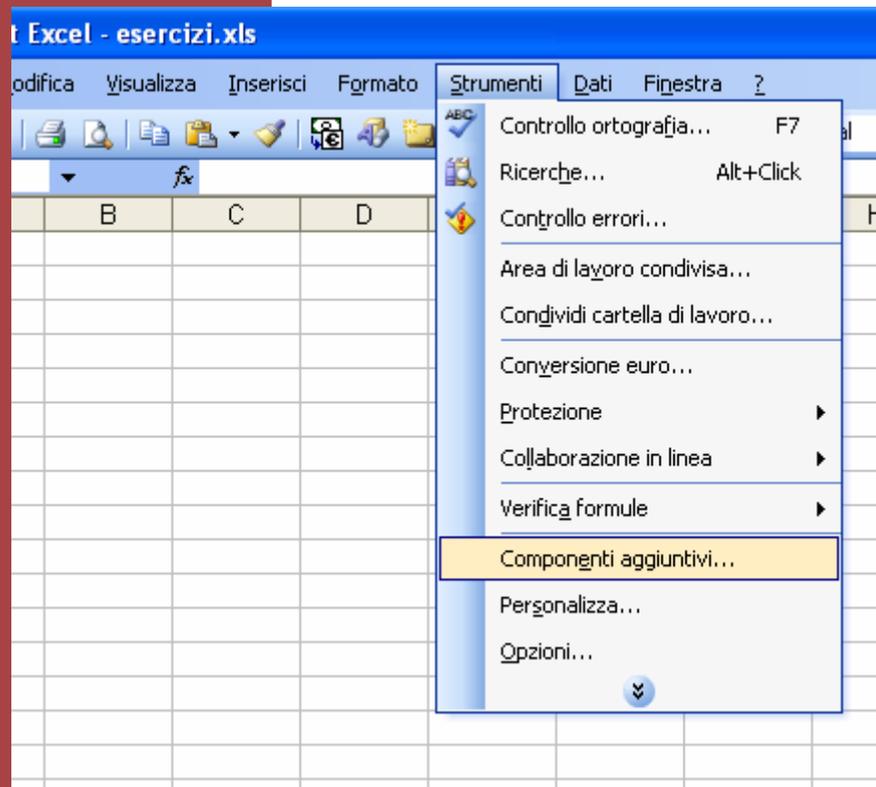
- 1) [Accedere al menù Strumenti](#) [Tools] e [verificare che il comando Analisi Dati](#) [Data analysis] sia presente nel menù.

[Torna](#) [Home](#)



2) se non fosse presente, [selezionare il comando Componenti aggiuntivi](#): dal menù **Strumenti** [Tools] e poi l'opzione **Strumenti di Analisi** [Analysis ToolPack] e cliccare su **OK**, per rendere disponibili le funzioni e le procedure avanzate.

Il comando **Analisi Dati** [Data analysis] sarà ora disponibile nel menù **Strumenti**.



3. Introduzione alla Statistica descrittiva

Il principale obiettivo della statistica è la conoscenza quantitativa dei fenomeni collettivi.



Statistica e fenomeni collettivi

Se si osserva un gruppo piuttosto folto di persone, ad esempio in fila per entrare allo stadio, non si è in grado di affermare di quanti individui è composto il gruppo.

Per saperlo si dovrà contarli uno ad uno, perché nessuno è in grado di valutare esattamente la consistenza di un folto gruppo di persone senza contarle.

L'operazione che si deve fare è forse la più importante, ma certo la più semplice, operazione della statistica: il *conteggio*; in base ad esso si perviene alla conoscenza del numero delle persone che costituiscono il gruppo.

La statistica è servita dunque per colmare una deficienza delle facoltà umane ed è quindi uno degli strumenti utilizzati a tal fine.

Anche per rimediare ad altre deficienze delle nostre facoltà vi sono appositi strumenti: ad esempio le limitazioni della vista vengono corrette con occhiali, quelle dell'udito con altoparlanti, eccetera.

La limitazione a cui si sopperisce con la statistica consiste nel fatto che, come si è visto a proposito della determinazione del numero delle persone che compongono un gruppo, la mente umana non sa sintetizzare quantitativamente i risultati di un elevato numero di osservazioni di fatti che invece percepisce singolarmente.



[Home](#)



I fenomeni che la nostra mente non può conoscere con una sola osservazione, ma che invece apprende tramite la sintesi delle osservazioni di fenomeni più semplici, vengono abitualmente chiamati *fenomeni collettivi o di massa*.

Essi, per essere conosciuti quantitativamente con precisione, richiedono dunque una collettività di osservazioni di fenomeni più semplici che sono detti quindi, per contrapposizione, *fenomeni individuali o singoli*.

Sono per esempio fenomeni collettivi la natalità, la nuzialità e la mortalità, le cui misure si ottengono contando l'insieme dei fenomeni individuali cui attengono (nascite, matrimoni o morti) di un certo periodo e, se necessario, mettendo a confronto valori ottenuti con quello relativo ad un altro fenomeno collettivo: l'ammontare della popolazione.

All'incapacità della mente di percepire con una sola osservazione i fenomeni collettivi, si provvede con la statistica, che ritornando alla prima definizione che ne abbiamo dato è, dunque, la tecnica che ha come scopo la conoscenza quantitativa dei fenomeni collettivi.



I fenomeni collettivi (o caratteristiche) relativi ad una certa popolazione possono manifestarsi con diverse modalità:

- *quantitative* (ad esempio: numero di figli, 0, 1, 2, ...)
- *qualitative* (ad esempio: il genere maschile e femminile)

Home



3.1 l'unità statistica

L'oggetto dell'osservazione di ogni fenomeno individuale che costituisce il fenomeno collettivo è detto unità statistica. Per esempio, è un'unità: ciascun individuo di una popolazione, ciascun nato, ciascun morto, ogni coppia di sposi, ciascun emigrato, eccetera, a seconda del fenomeno oggetto di indagine.

[Torna](#) [Home](#)



3.2 I Caratteri dell'unità statistica

Ciascuna unità statistica presenta delle caratteristiche, che più brevemente si indicano con il nome di **caratteri**. Ogni unità statistica è portatrice, generalmente, di infiniti caratteri. Si consideri un gruppo di studenti universitari. Per ognuno di essi si può determinare l'età, la statura, il peso, il sesso, il colore dei capelli, il gruppo sanguigno, il numero dei fratelli, la nazionalità, la religione, il luogo di residenza, la facoltà cui è iscritto, l'anno di corso, eccetera. Si dice che l'età, la statura, il sesso, eccetera, sono tutti *caratteri di ciascuna unità* (nell'esempio: studente).

Le modalità del carattere

Ciascun carattere è presente in ogni unità con una determinata **modalità**: per esempio, se uno studente è di nazionalità italiana ed un altro è di nazionalità greca, diciamo che nel primo il carattere nazionalità presenta la modalità "italiana", mentre nell'altro la modalità "greca".

Uguaglianza di due unità

Per qualunque carattere è sempre possibile riscontrare l'uguaglianza o la disuguaglianza delle modalità presentate dal carattere in due unità diverse. Due unità le diciamo *uguali rispetto ad uno o più caratteri* da esse posseduti se ognuno di quei caratteri è presente in ciascuna delle due unità con la stessa modalità. L'insieme di tutte le unità statistiche che sono uguali rispetto ad alcuni loro caratteri diciamo che è un "*collettivo statistico*" (o soltanto "*un collettivo*") o "*una popolazione*".

Caratteri quantitativi e caratteri qualitativi

I caratteri presenti in una **unità statistica** sono generalmente di natura assai diversa. Nei casi in cui questi possano essere espressi da numeri, ossia siano misurabili, ossia per essi sia possibile definire un'unità di misura, cosicché le modalità siano numeri che è possibile sommare o sottrarre fra loro (numeri cardinali), si dice siamo in presenza di **caratteri quantitativi** (ad esempio il peso, l'altezza, il numero dei fratelli, eccetera); in tutti gli altri casi si è generalmente in presenza di **caratteri qualitativi** (ad esempio: il colore dei capelli, o l'anno di corso, che pure è rappresentato da un numero ma di tipo ordinale, cioè che non esprime una quantità).

[Torna](#)

[Home](#)



Caratteri qualitativi sconnessi e ordinati

I **caratteri qualitativi**, a loro volta, possono distinguersi in **sconnessi** oppure **ordinati** a seconda che sia possibile o meno graduarne la diversità. Ad esempio, il carattere qualitativo “religione” è sconnesso: noi diciamo che le religioni ebraica e musulmana sono diverse e che diverse sono pure quella cristiana e musulmana, ma non possiamo graduare le diversità; invece per il carattere “titolo di studio” possiamo dire, per esempio, non solo che i titoli “licenza elementare” e “licenza media” sono diversi e che lo sono anche “licenza elementare” e “diploma di scuola secondaria superiore” ma anche che vi è una maggiore diversità tra “diploma” e “licenza elementare” che fra questo titolo e “licenza media”.

Un'altra importante distinzione:

Caratteri quantitativi continui e discreti

I caratteri **quantitativi** vengono a loro volta distinti in caratteri **continui** e in caratteri **discontinui** (e fra questi ultimi vengono individuati anche i caratteri **discreti**).

Un carattere quantitativo è detto **continuo** se, comunque si fissino due valori (entro l'intervallo in cui il carattere è osservabile), tutti i valori intermedi possono essere assunti come modalità del carattere (si pensi al “peso” e anche all’età” se misurata in anni, mesi, giorni, ore, minuti).

Un carattere che non sia continuo è detto discontinuo.

Un carattere discontinuo è denominato **discreto** se, comunque si fissi una sua modalità (interna all'intervallo in cui il carattere è osservabile), esiste tutto un intervallo - di cui la modalità è il centro - in cui, all'infuori di essa, nessun altro valore può essere assunto come modalità del carattere. Ad esempio, il “numero dei fratelli” è un carattere discreto: infatti, mentre si possono avere 3 fratelli, non se ne possono avere 2,7 o 2,8 o 2,9 o 3,1 e così via... e quindi questi ultimi valori non possono essere assunti come modalità del carattere “numero dei fratelli”.

[vedi schema riassuntivo](#)

[Torna](#)

[Home](#)



ORDINATI
Possono essere ordinati naturalmente
(ad esempio possiamo ordinare gradualmente il
carattere “Titolo di studio” che, infatti, può essere:
licenza elementare, di scuola secondaria, diploma,
eccetera)

CONTINUI
Possono assumere qualunque valore numerico
compreso nell’intervallo di variazione; non è
possibile elencare tutte le modalità che può
assumere la variabile ma occorrerà limitarsi a
contare quante unità manifestano la variabile
con modalità compresa in un certo intervallo
di valori reali

QUALITATIVI
Caratteri dell’unità statistica che
identificano qualità o categorie non
misurabili, ma soltanto classificabili
secondo modalità diverse

CARATTERI
Le diverse caratteristiche
di ciascuna unità
statistica; possono
distinguersi in *qualitativi*
e *quantitativi*

QUANTITATIVI
Caratteri dell’unità statistica che
possono essere misurati o espressi
mediante un numero e che possono
essere di natura discreta o continua.
Anche, si dicono *non trasferibili* se
non possono essere ceduti, del tutto
o in parte, ad un’altra unità statistica
(età, peso)

SCONNESSI
Non possono essere ordinati naturalmente
(ad esempio: la religione professata)

DISCRETI
Possono assumere soltanto un numero finito
intero entro l’intervallo di variazione (ad
esempio: numero di persone residenti in una città,
num. di vani in una abitazione, ecc.);
è sempre possibile elencare tutte le modalità
che può assumere la variabile

Torna

4. Rappresentazione dei dati

La presentazione dei dati statistici in una forma organizzata (enumerando e classificando **modalità** coincidenti) semplifica confronti e analisi successive oltre a fornire sintesi di migliore leggibilità per l'interpretazione dei risultati, anche a fini comparativi.

Ad esempio: anziché leggere la successione (M, M, F, M, M, F, F, M, F,) relativa al Sesso di 48 studenti universitari laureati in Economia in una certa sessione, è conveniente esaminare la tabella:

Maschi	27
Femmine	21
Totale	48

Talvolta, si preferisce un grafico, detto “a torta”, (*pie chart*) nel quale l'area del settore circolare attribuito ai Maschi sarà del 56,3% (ottenuto dal calcolo di $27/48$ per 100 e la rimanente parte, pari al 43,8% (ottenuta da $21/48$ per 100), è attribuita alle Femmine (esempi di tali rappresentazioni saranno esaminati nei capitoli seguenti).

In generale, i dati statistici si possono (rap)presentare in forma **enumerativa**, **tabellare** o **grafica** a seconda che essi siano **disponibili secondo una successione di informazioni sulle singole unità statistiche**, **secondo un raggruppamento di casi evidenziato in tabelle semplici o complesse**, oppure **secondo una rappresentazione grafica che evidenzi aspetti particolari nei dati**.



4.1 Le distribuzioni statistiche (brevi cenni)

L'effetto dell'operazione di determinazione della modalità con cui ognuno dei caratteri si presenta in ciascuna unità del collettivo è la “*distribuzione*” del collettivo secondo i caratteri considerati.

Il nome “*distribuzione*” deriva dal fatto che mediante essa si indica come le **modalità** dei **caratteri** si distribuiscono nelle unità del collettivo.

Si distinguono: distribuzioni di quantità e distribuzioni di frequenza.

Distribuzioni di quantità

La distribuzione di quantità è una organizzazione di dati in forma tabellare tale che ad ogni modalità di una certa variabile si fa corrispondere una quantità (assoluta o relativa), idealmente trasferibile tra le unità della popolazione.

Quindi la distribuzione di quantità esplicita come l'ammontare complessivo del fenomeno si distribuisce fra le modalità, le categorie, le unità amministrative, i gruppi sociali e così via.

Tali rappresentazioni sono molto diffuse nelle pubblicazioni periodiche dell'Istat nonché sui mezzi di informazioni di massa. Esempi: immatricolati per corso di laurea, incidenti stradali per mese dell'anno, reati denunciati per tipologia, eccetera.

Distribuzioni di frequenza

La **distribuzione di frequenza** è una organizzazione di dati in forma tabellare tale che ad ogni modalità di una certa variabile (qualitativa o quantitativa) si fa corrispondere la rispettiva **frequenza (assoluta o relativa)**.

Quindi la distribuzione di frequenza esplicita quante volte (in assoluto se trattasi di frequenza assoluta, oppure in rapporto al totale se trattasi di frequenza relativa) una determinata modalità si presenta nel collettivo in esame.

Nota: La distinzione tra distribuzione di quantità e distribuzione di frequenza non è sempre netta; esistono fenomeni che possono essere considerati sotto il punto di vista delle due tipologie consentendo evidentemente interpretazioni differenti.

[Tra le rappresentazioni più comuni, ai fini di questo corso ci apprestiamo ad esaminare le distribuzioni di frequenza](#)

[Home](#)



4.2 Distribuzioni di frequenza per caratteri quantitativi (discreti e continui)

Si è appena visto (cap. 3) come i caratteri quantitativi possano classificarsi in:

discreti

Nel caso di caratteri discreti sarà sempre possibile elencare tutte le modalità che può assumere la variabile (si pensi ai voti conseguiti da un gruppo di studenti con riferimento ad una certa prova d'esame)

continui

Nel caso di caratteri continui la rilevazione darà luogo a numeri quasi sempre differenti tra loro (si pensi ai tempi cronometrati per singolo atleta al termine di una corsa podistica); occorrerà quindi limitarsi a contare quante unità manifestano la variabile con modalità compresa in un certo intervallo di valori reali.

[ripassa lo schema riassuntivo](#)

Esemplificheremo adesso i diversi processi di rappresentazione per ciascuno dei due tipi caratteri quantitativi (fenomeni) considerati: →



	A	B	C	D
1	studente	num. di fratelli per studente		
2	gianni	1		
3	maria	1		
4	luca	2		
5	fabio	1		
6	angela	0		
7	michele	3		
8	mirko	4		
9	daniela	2		
10	marco	3		
11	giuseppe	1		
12	stefano	0		
13	laura	2		
14	chiara	1		
15	davide	1		
16	luigi	0		
17	monica	1		
18	sara	1		
19	giovanni	0		
20	aldo	3		
21	serena	2		
22				

Esempi di tipo 1: Caratteri discreti

Utilizzeremo un campione di 20 studenti che, alla domanda “*quanti fratelli (o sorelle) hai?*”, hanno risposto (rispettivamente):

1,1,2,1,0,3,4,2,3,1,0,2,1,1,0,1,1,0,3,2

Gli studenti sono elencati dalla cella A2 alla cella A21.

Nella cella B1 abbiamo digitato il nome del carattere: **numero di fratelli per studente**

Dalla cella B2 alla cella B21 abbiamo inserito i corrispettivi valori.



Home



	A	B	C
1	atleta	Tempi (min/sec)	
2	gianni	12,1	
3	maria	8,9	
4	luca	16,2	
5	fabio	8,2	
6	angela	9,8	
7	michele	15,1	
8	mirko	14,5	
9	daniela	13,4	
10	marco	14,7	
11	giuseppe	7,5	
12	stefano	8,8	
13	laura	12,4	
14	chiara	16,1	
15	davide	15,2	
16	luigi	13,5	
17	monica	14,6	
18	sara	15,5	
19	giovanni	7,8	
20	aldo	12,5	
21	serena	13,2	
22	claudia	11	
23	piero	10,5	

Esempi di tipo 2: Caratteri continui

Utilizzeremo un campione costituito da 22 atleti impegnati in una gara di corsa campestre; questi i tempi (in minuti, secondi) ottenuti da ciascuno:

12,1; 8,9; 16,2; 8,2; 9,8; 15,1; 14,5; 13,4;
 14,7; 7,5; 8,8; 12,4; 16,1; 15,2; 13,5;
 14,6; 15,5; 7,8; 12,5; 13,2; 11,0; 10,5.

Gli atleti sono elencati dalla cella A2 alla cella A23.

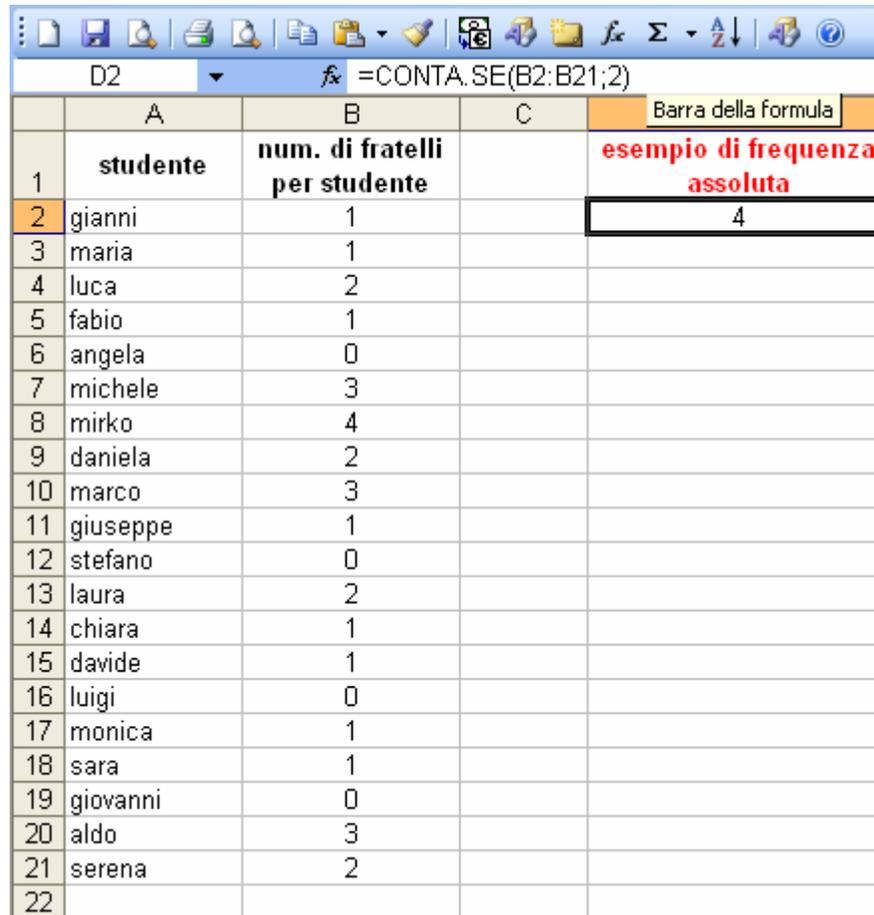
Nella cella B1 abbiamo digitato il nome del carattere: **Tempi (minuti, secondi)**.

Dalla cella B2 alla cella B23 abbiamo inserito i corrispettivi valori.



4.2.1 Frequenza assoluta e frequenza relativa

Frequenza assoluta



	A	B	C	Barra della formula
1	studente	num. di fratelli per studente		esempio di frequenza assoluta
2	gianni	1		4
3	maria	1		
4	luca	2		
5	fabio	1		
6	angela	0		
7	michele	3		
8	mirko	4		
9	daniela	2		
10	marco	3		
11	giuseppe	1		
12	stefano	0		
13	laura	2		
14	chiara	1		
15	davide	1		
16	luigi	0		
17	monica	1		
18	sara	1		
19	giovanni	0		
20	aldo	3		
21	serena	2		
22				

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Home](#)

La **frequenza assoluta** si determina mediante la funzione **CONTA.SE [COUNT IF]**.

CONTA.SE (sequenza di numeri o indirizzo della cella; criteri di selezione o indirizzo di cella contenente i criteri).

Esercizio 1.2 - determinare il numero di alunni con 2 fratelli:

Come si fa:

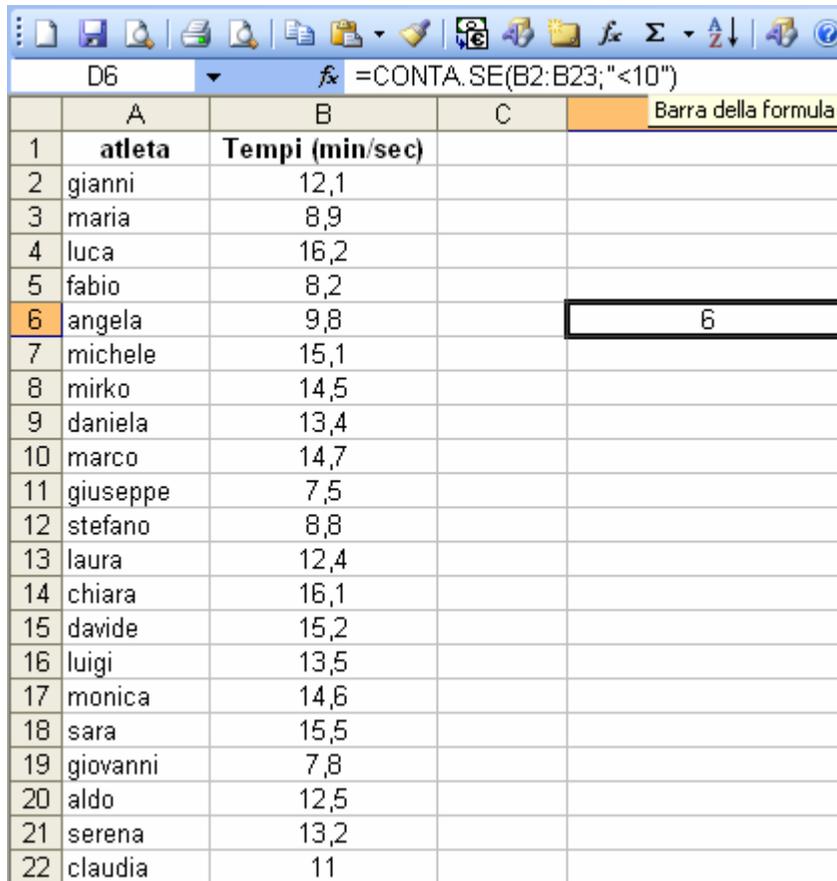
- Spostare il cursore nella cella di destinazione D2.
- Introdurre la funzione:
=CONTA.SE(B2:B21;"2").

Spiegazione: questa funzione calcola il numero di celle che non siano vuote e che rispondano ad un determinato criterio (nell'esercizio in esempio = 2) (Criterio: è la condizione, espressa in numero o testo, che definisce le celle da contare)

La procedura utilizzata per calcolare la frequenza assoluta delle classi di caratteri qualitativi, è identica a quella presentata per i caratteri discreti, sostituendo nei criteri il numero cardinale con il nome della modalità.



Frequenza assoluta, altri esempi:



	A	B	C	Barra della formula
1	atleta	Tempi (min/sec)		
2	gianni	12,1		
3	maria	8,9		
4	luca	16,2		
5	fabio	8,2		
6	angela	9,8		6
7	michele	15,1		
8	mirko	14,5		
9	daniela	13,4		
10	marco	14,7		
11	giuseppe	7,5		
12	stefano	8,8		
13	laura	12,4		
14	chiara	16,1		
15	davide	15,2		
16	luigi	13,5		
17	monica	14,6		
18	sara	15,5		
19	giovanni	7,8		
20	aldo	12,5		
21	serena	13,2		
22	claudia	11		

Esercizio 2.1 – tra i partecipanti alla gara podistica immaginata nell'esempio 2, determinare il numero di atleti che hanno ottenuto un tempo inferiore a 10 minuti

Come si fa:

- Spostare il cursore nella cella di destinazione (ad esempio D6)
- Introdurre la funzione:
=CONTA.SE(A2:A23;"<10").

Spiegazione: come nell'esempio precedente, questa funzione calcola il numero di celle che non siano vuote e che rispondano ad un determinato criterio (nell'esempio: < 10)

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



Frequenza assoluta, altri esempi:



	A	B	C	D
1	atleta	Tempi (min/sec)		
2	gianni	12,1		
3	maria	8,9		
4	luca	16,2		
5	fabio	8,2		
6	angela	9,8		
7	michele	15,1		
8	mirko	14,5		8
9	daniela	13,4		
10	marco	14,7		
11	giuseppe	7,5		
12	stefano	8,8		
13	laura	12,4		
14	chiara	16,1		
15	davide	15,2		
16	luigi	13,5		
17	monica	14,6		
18	sara	15,5		
19	giovanni	7,8		
20	aldo	12,5		
21	serena	13,2		
22	claudia	11		
23	piero	10,5		
24				
25				

Esercizio 2.2 - determinare il numero di atleti che hanno ottenuto un tempo di gara uguale o superiore a 10' ed inferiore ai 14' (ovvero compreso tra i 10' ed i 14').

Come si fa:

-Spostare il cursore nella cella di destinazione D3.

- Introdurre la funzione:
=CONTA.SE(A2:A23;"<14") - CONTA.SE(A2:A23;"<10").

Spiegazione: il risultato è dato dalla differenza tra:

- il numero di studenti che hanno ottenuto un tempo inferiore a 14'

Ed il

- numero di studenti che hanno ottenuto un tempo inferiore ai 10'

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Torna](#)

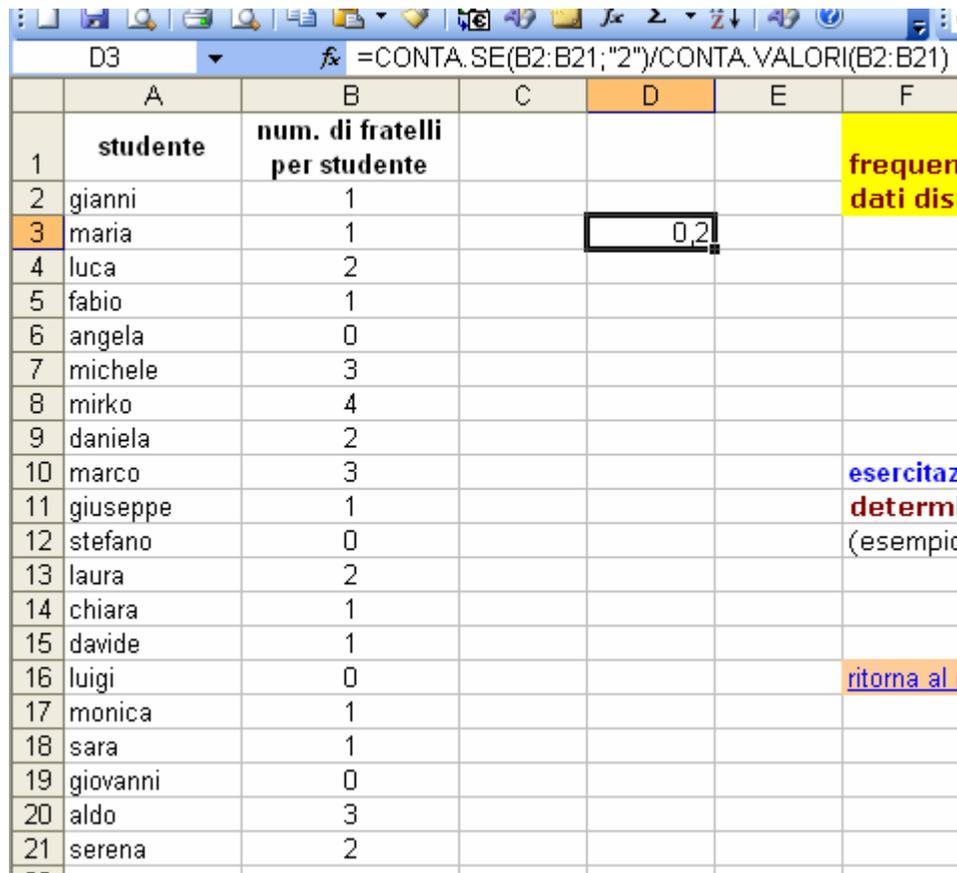
[Home](#)



Frequenza relativa

Per il calcolo della frequenza relativa è necessario determinare la dimensione del **campione**; a questo scopo possiamo utilizzare una delle seguenti funzioni:

- **CONTA.NUMERI [COUNT]** (sequenza di numeri o indirizzi di cella): questa funzione calcola il numero di celle totali che non siano vuote e che rappresentino valori numerici.
- **CONTA.VALORI [COUNTA]** (sequenza di numeri o indirizzi di cella): questa funzione calcola il numero di celle totali che non siano vuote ed è idonea al calcolo della dimensione dei campioni che contengano variabili qualitative.



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	studente	num. di fratelli per studente				frequen dati dis
2	gianni	1				
3	maria	1		0,2		
4	luca	2				
5	fabio	1				
6	angela	0				
7	michele	3				
8	mirko	4				
9	daniela	2				
10	marco	3				esercitaz determi
11	giuseppe	1				(esempic
12	stefano	0				
13	laura	2				
14	chiara	1				
15	davide	1				
16	luigi	0				ritorna al
17	monica	1				
18	sara	1				
19	giovanni	0				
20	aldo	3				
21	serena	2				

The formula bar shows: `=CONTA.SE(B2:B21;"2")/CONTA.VALORI(B2:B21)`

Esercizio 1.3 - determinare la frazione di alunni con 2 fratelli:

Come si fa:

-Spostare il cursore nella cella di destinazione D3.

-Introdurre la funzione:
`=CONTA.SE(A2:A21;"2")/CONTA.VA
LORI(A2:A21).`

(verificare che la cella di destinazione sia impostata a più decimali)

Spiegazione: il risultato è dato dal rapporto tra:

- il numero di studenti che hanno dichiarato di avere 2 fratelli

Ed il

- numero totale di studenti intervistati

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

Home



4.3 Costruire una Tabella di frequenza (esempi su dati discreti e/o continui)

Una Tabella di frequenza rappresenta la distribuzione della variabile, cioè, quali valori o modalità la variabile assume e con che frequenza assume questi valori o modalità.

Per costruire una tabella di frequenza dobbiamo combinare insieme le funzioni cui abbiamo accennato in precedenza.

Primo caso: caratteri discreti

Per quanto riguarda il caso dei caratteri discreti, la creazione della tabella di frequenza non presenta alcuna difficoltà (vedi slides immediatamente successive).

Tutti gli esercizi di tipo 1

Secondo caso: caratteri continui

In presenza di caratteri continui il processo è un po' più elaborato e occorrerà soffermarci su alcuni passi principali che verranno descritti più avanti e che saranno illustrati utilizzando i dati di base dell'esercizio 2.

Tutti gli esercizi di tipo 2

[ripassa lo schema riassuntivo](#)



[Torna](#)

[Home](#)



4.3.1 costruire una Tabella di frequenza su dati discreti

	A	B	C	D	E	F
1	N°.di Fratelli			Tabella di frequenza		
2	1		modalità	Freq. assoluta	Freq. Relativa	
3	1		0			
4	2		1			
5	1		2			
6	0		3			
7	3		4			
8	4		Totale			
9	2					
10	3					
11	1					
12	0					
13	2					
14	1					
15	1					
16	0					
17	1					
18	1					
19	0					
20	3					
21	2					

Esercizio 1.4

Come si fa:

Utilizziamo gli stessi dati dell'esercizio 1, (numero di fratelli per ciascun studente) elencati nella colonna A della figura a lato

1) Definizione delle modalità:

osserviamo come i risultati (in colonna A) possono verificarsi in 5 modalità (colonna B) distinte: 0, 1, 2, 3, o 4 (ovvero: ci sono studenti con zero fratelli, altri che ne hanno 1, 2, eccetera).

Come illustrato nelle precedenti slides, servendoci delle funzioni disponibili in excel possiamo distribuire i rispettivi valori alle diverse classi; ovvero: **conteggiare gli elementi che ricadono in ciascuna modalità**; calcoleremo quanti studenti sono figli unici (frequenza assoluta di classe 0), quanti hanno un fratello (frequenza assoluta di classe 1), eccetera.

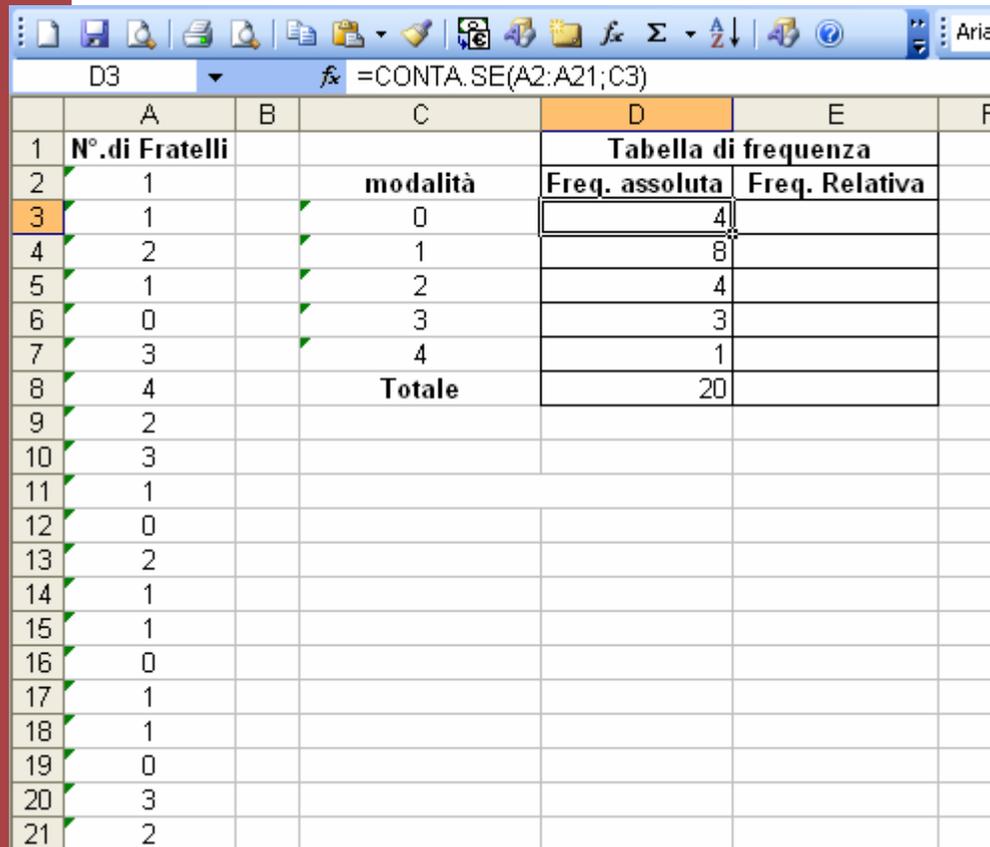
Rivediamo come →



[Torna](#) [Home](#)



2) Calcolo delle frequenze assolute:



	A	B	C	D	E	F
1	N°.di Fratelli			Tabella di frequenza		
2	1		modalità	Freq. assoluta	Freq. Relativa	
3	1		0	4		
4	2		1	8		
5	1		2	4		
6	0		3	3		
7	3		4	1		
8	4		Totale	20		
9	2					
10	3					
11	1					
12	0					
13	2					
14	1					
15	1					
16	0					
17	1					
18	1					
19	0					
20	3					
21	2					

Come si fa:

- Spostare il cursore nella cella di destinazione D3.

- Introdurre la funzione:
=CONTA.SE(A2:A21;C3).

- Ripetere il procedimento per le celle da D4 a D7.

Spiegazione:

stiamo verificando quante celle, comprese nell'intervallo A2:A21, soddisfano la condizione: valore corrispondente a C3 (poi C4, C5, ecc.).

La somma delle frequenze assolute deve corrispondere alla numerosità del campione:

- Inserire la funzione
=CONTA.VALORI(A2:A21) nella cella D8.

Quindi procediamo al calcolo delle frequenze relative:



[Torna](#)

[Home](#)



3) Calcolo delle frequenze relative:

	A	B	C	D	E	F
1	N°.di Fratelli			Tabella di frequenza		
2	1		modalità	Freq. assoluta	Freq. Relativa	
3	1		0	4	0,2	
4	2		1	8	0,4	
5	1		2	4	0,2	
6	0		3	3	0,15	
7	3		4	1	0,05	
8	4		Totale	20	1	
9	2					
10	3					
11	1					
12	0					
13	2					
14	1					
15	1					
16	0					
17	1					
18	1					
19	0					
20	3					
21	2					

Come si fa:

- Spostare il cursore nella cella di destinazione E3.

- Introdurre la funzione: **=D3/D8**.

- Ripetere il procedimento anche per le celle dalla E4 alla E7.

- Inserire la funzione: **=SOMMA(E3:E7)** **[SUM]** nella cella E8.

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



4.3.2 costruire una Tabella di frequenza su dati continui

Esercizio 2.3 Come si fa:

1) Definizione delle classi (numero e ampiezza)

a) anzitutto determinare l'ampiezza del campione (MAX – MIN vedere capitolo 6.3).

b) stabilire il **numero di classi (k)** impostando la seguente formula nella cella: {ARROTONDA.PER.ECC(LOG(n;2);0)}

c) calcolare: **ampiezza di classe (h) = ampiezza del campione diviso per il numero di classi (k)**

B10		=ARROTONDA.PER.ECC(LOG(B8;2);0)				
	A	B	C	D	E	F
1	Tempi (min/sec)	Minimo		Tabella di frequenze		
2	12,1	7,5		Classi	Freq. Assoluta	Freq. Relativa
3	8,9	Massimo		[7,5;9,3[
4	16,2	16,2		[9,3;11,1[
5	8,2	Ampiezza		[11,1;12,9[
6	9,8	8,7		[12,9;14,7[
7	15,1	n		[14,7;16,5[
8	14,5	22		Totale		
9	13,4	k				
10	14,7	5				
11	7,5	h				
12	8,8	1,74				
13	12,4	h approssimato				
14	16,1	1,8				
15	15,2					
16	13,5					
17	14,6					
18	15,5					
19	7,8					
20	12,5					
21	13,2					
22	11					
23	10,5					

d) considerare come **ampiezza di classe (h)** il suo valore approssimato in **eccesso (h approssimato)** con la formula:

=ARROTONDA.PER.ECC(B12;1)

per evitare che una misura coincida con un estremo della classe

e) Costruire le classi C1, C2, ..., Ck così che:

-abbiano tutte la stessa ampiezza

- la loro unione contenga tutti gli elementi del campione

In forma generica $C_i = [\text{min. del campione} + (i-1) \times h, \text{min. del campione} + i \times h]$ con $i=1, 2, \dots, k$.



[Torna](#)

[Home](#)



2) conteggio degli elementi delle classi (frequenze assolute)

Determinare il numero degli elementi del campione che appartengono a ogni classe.

Analogamente a quanto è stato fatto nel caso dei dati discreti ([rivedi in cap. 4.3.1](#)), tali valori saranno le frequenze assolute delle classi. Nel caso dell'esempio presentato:

E3 = =CONTA.SE(A1:A23;"<9,3")-CONTA.SE(A1:A23;"<7,5")						
	A	B	C	D	E	F
1	Tempi (min/sec)	Minimo			Tabella di frequenze	
2	12,1	7,5		Classi	Freq. Assoluta	Freq. Relativa
3	8,9	Massimo		[7,5;9,3[5	
4	16,2	16,2		[9,3;11,1[3	
5	8,2	Ampiezza		[11,1;12,9[3	
6	9,8	8,7		[12,9;14,7[5	
7	15,1	n		[14,7;16,5[6	
8	14,5	22		Totale	22	
9	13,4	k (numero classi)				
10	14,7	5				
11	7,5	h				
12	8,8	1,74				
13	12,4	h approssimato				
14	16,1	1,8				
15	15,2					
16	13,5					
17	14,6					
18	15,5					
19	7,8					
20	12,5					
21	13,2					
22	11					
23	10,5					
24						



Infine: 3) calcolo delle frequenze relative:

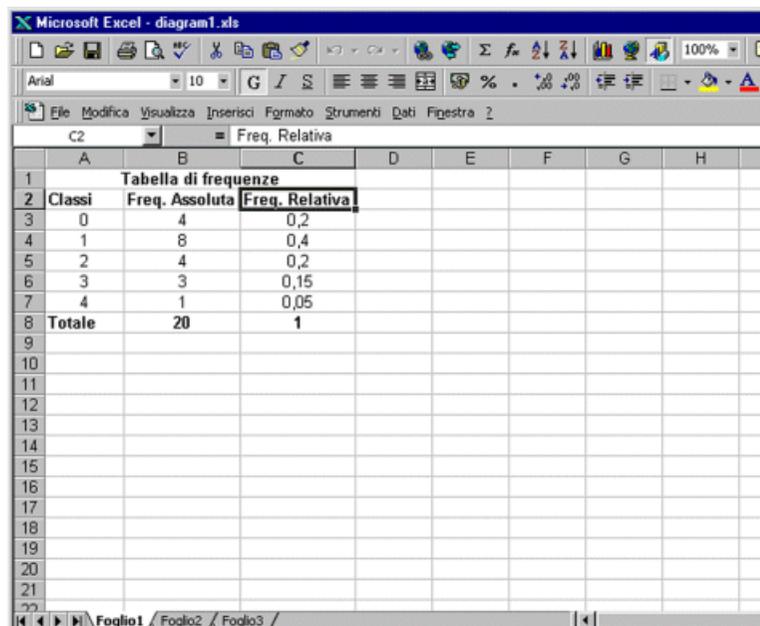
	A	B	C	D	E	F	G	H
1	Tempi (min/sec)	Minimo			Tabella di frequenze			
2	12,1	7,5		Classi	Freq. Assoluta	Freq. Relativa		
3	8,9	Massimo		[7,5;9,3[5	0,227272727		
4	16,2	16,2		[9,3;11,1[3	0,136363636		
5	8,2	Ampiezza		[11,1;12,9[3	0,136363636		
6	9,8	8,7		[12,9;14,7[5	0,227272727		
7	15,1	n		[14,7;16,5[6	0,272727273		
8	14,5	22		Totale	22	1		
9	13,4	k						
10	14,7	5						
11	7,5	h						
12	8,8	1,74						
13	12,4	h approssimato						
14	16,1	1,8						
15	15,2							
16	13,5							
17	14,6							
18	15,5							
19	7,8							
20	12,5							
21	13,2							
22	11							
23	10,5							
24								



Home

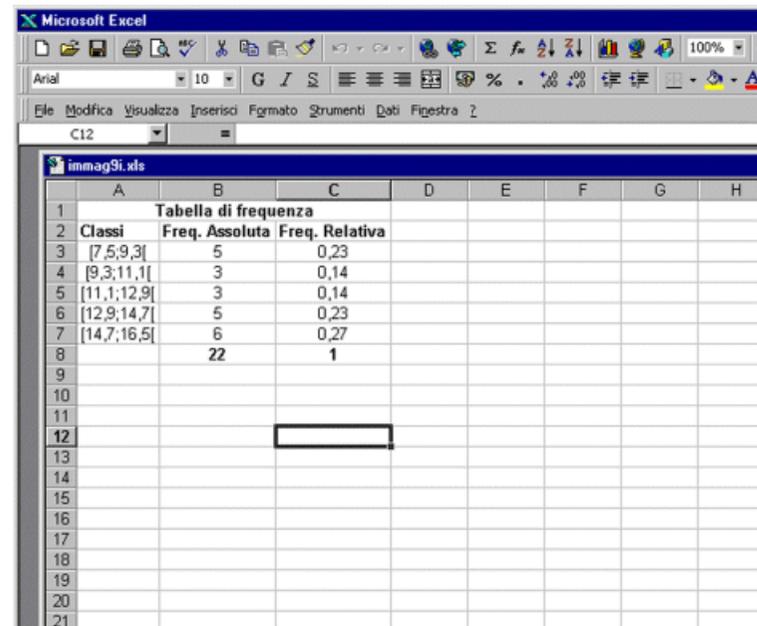


Esplorando le diverse funzioni di formattazione (“centra”, “diminuisce decimali”, ecc.), si potranno ottenere le seguenti tabelle:



Microsoft Excel - diagram1.xls

	A	B	C	D	E	F	G	H
1	Tabella di frequenze							
2	Classi	Freq. Assoluta	Freq. Relativa					
3	0	4	0,2					
4	1	8	0,4					
5	2	4	0,2					
6	3	3	0,15					
7	4	1	0,05					
8	Totale	20	1					



Microsoft Excel

	A	B	C	D	E	F	G	H
1	Tabella di frequenza							
2	Classi	Freq. Assoluta	Freq. Relativa					
3	[7,5;9,3]	5	0,23					
4	[9,3;11,1]	3	0,14					
5	[11,1;12,9]	3	0,14					
6	[12,9;14,7]	5	0,23					
7	[14,7;16,5]	6	0,27					
8		22	1					

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



4.4 Le rappresentazioni grafiche

Il cervello umano possiede, tramite il senso della vista, una velocissima, efficiente e duratura capacità di immagazzinare e confrontare informazioni, generando persino la memoria permanente di quelle informazioni. Per questo la rappresentazione grafica è un momento importante di ogni analisi statistica.

Ciò che da una tabella può dedursi solo con una certa fatica e soltanto da esperti è invece fornito immediatamente e a chiunque dalle rappresentazioni grafiche.

Queste sono sostanzialmente dei disegni ottenuti facendo corrispondere ai numeri di una tabella:

- enti geometrici elementari (come ad esempio punti, linee, angoli, superfici)

Oppure:

- diverse intensità di colore o di tratteggio.

Le rappresentazioni grafiche vengono oggi sviluppate con grande efficacia tramite Pc e software sofisticati, alcuni dei quali consentono anche rappresentazioni dinamiche.

Esamineremo i seguenti principali tipi di grafico:

- Diagramma a barre
- Istogramma
- Diagramma degli estremi e quartili
- Diagramma a settori circolari
- Diagramma di dispersione o a punti



[Home](#)



Diagramma a barre

È un grafico generalmente usato per caratteri qualitativi che si costruisce mediante un sistema di assi ponendo in orizzontale le frequenze del carattere e in verticale le modalità; le barre sono dei rettangoli di altezza costante e base proporzionale alla frequenza di ciascuna modalità.

Istogramma

È un grafico generalmente usato per caratteri quantitativi continui; si basa su di un sistema di assi cartesiani, in cui sull'asse delle ascisse si pongono le modalità e sull'asse delle ordinate le densità di frequenza (pari al rapporto fra le frequenze e l'ampiezza delle classi). È costituito da tanti rettangoli quante sono le modalità del carattere, con base pari all'ampiezza di ciascuna classe ed area pari alle frequenze relative (l'altezza è quindi data dalle densità di frequenza).

Diagramma a settori circolari (torta)

Si costruisce suddividendo un cerchio in tanti settori quante sono le modalità del carattere; l'angolo di ciascun settore è proporzionale alla frequenza assoluta (o relativa).

Box Plot: Diagramma degli estremi e dei quartili

Rappresentazione grafica in cui sono evidenziate alcune caratteristiche del campione. L'insieme dei valori campionari compresi tra il 1° e il 3° quartile (vedi cap. 5.4), Q25 e Q75, è rappresentato da un rettangolo (scatola) la cui mediana (vedi anche: mediana) è indicata da una barra. Due linee uniscono i lati dei rettangoli rispettivamente con il valore massimo e minimo.

Diagramma di dispersione o a punti

Rappresentazione grafica di valori bivariati (vedi anche: dati bivariati), in cui ogni coppia di dati (x_i, y_i) è rappresentata da un punto di coordinate (x_i, y_i) in un sistema di assi cartesiani (questi tipi di diagramma saranno illustrati nel cap. 7.1)



4.4.1 Diagramma a barre

Una volta ottenuta la tabella di frequenza si può ottenere con facilità un diagramma a barre, mediante la funzione *Creazione guidata grafico* [Chart Wizard]

The screenshot shows an Excel spreadsheet with a frequency table and the 'Creazione guidata Grafico - Passaggio 1 di 4 - Tipo di grafico' dialog box. The spreadsheet data is as follows:

	A	B	C
1	N°.di Fratelli		Tabella di
2	1	modalità	Freq. assoluta
3	1	0	4
4	2	1	7
5	1	2	4
6	0	3	3
7	3	4	1
8	4	Totale	20
9	2		
10	3		
11	1		
12	0		
13	2		
14	1		
15	1		
16	0		
17	1		
18	1		
19	0		
20	3		
21	2		
22			
23			
24			

The dialog box shows the 'Barre' option selected under 'Tipi standard'. The 'Barre non in pila' option is selected in the preview area. The dialog box also includes a list of other chart types and a 'Tenere premuto per visualizzare l'esempio' button.

Esercizio 1.5 – creare un grafico a barre

Come si fa:

Selezionare innanzitutto le celle contenenti i dati e i rispettivi titoli e cliccare sull'icona della Barra degli Strumenti.

Nella prima finestra di dialogo selezionare l'opzione **Barre** [Column]

[Home](#)

Per continuare la creazione del grafico e per passare al passaggio [Step] successivo, cliccare su **Avanti** **[Next]**.

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C
1	N°.di Fratelli		Tabella di
2	1	modalità	Freq. assoluta
3	1	0	4
4	2	1	7
5	1	2	4
6	0	3	3
7	3	4	1
8	4	Totale	20

The 'Creazione guidata Grafico - Passaggio 2 di 4 - Dati di origine del grafico' dialog box is open, showing a horizontal bar chart. The chart has a vertical axis labeled '0' to '4' and a horizontal axis labeled '0' to '8'. The bars represent the absolute frequencies for each number of brothers: 0 (3), 1 (7), 2 (4), 3 (1), 4 (1). The 'Intervallo dati' field contains the formula `=Foglio1!B3:C7`. The 'Serie in:' section has 'C colonne' selected.



Giunti al terzo passaggio, la Finestra di dialogo presenta varie opzioni che permettono di formattare il grafico:

- Nella funzione **Titolo [Title]** seguire l'esempio presentato.
- Nella funzione **Legenda [Legend]** disattivare l'opzione legenda.

Per continuare la creazione del grafico e per passare al Passaggio [Step] successivo, cliccare su **Avanti [Next]**.

	A	B	C	D	E	F	G	H	I	J
1	N°.di Fratelli		Tabella di frequenza							
2	1	modalità	Freq. assoluta							
3	1	0	4							
4	2	1	7							
5	1	2	4							
6	0	3	3							
7	3	4	1							
8	4	Totale	20							
9	2									
10	3									
11	1									
12	0									
13	2									
14	1									
15	1									
16	0									
17	1									
18	1									
19	0									
20	3									
21	2									
22										
23										
24										

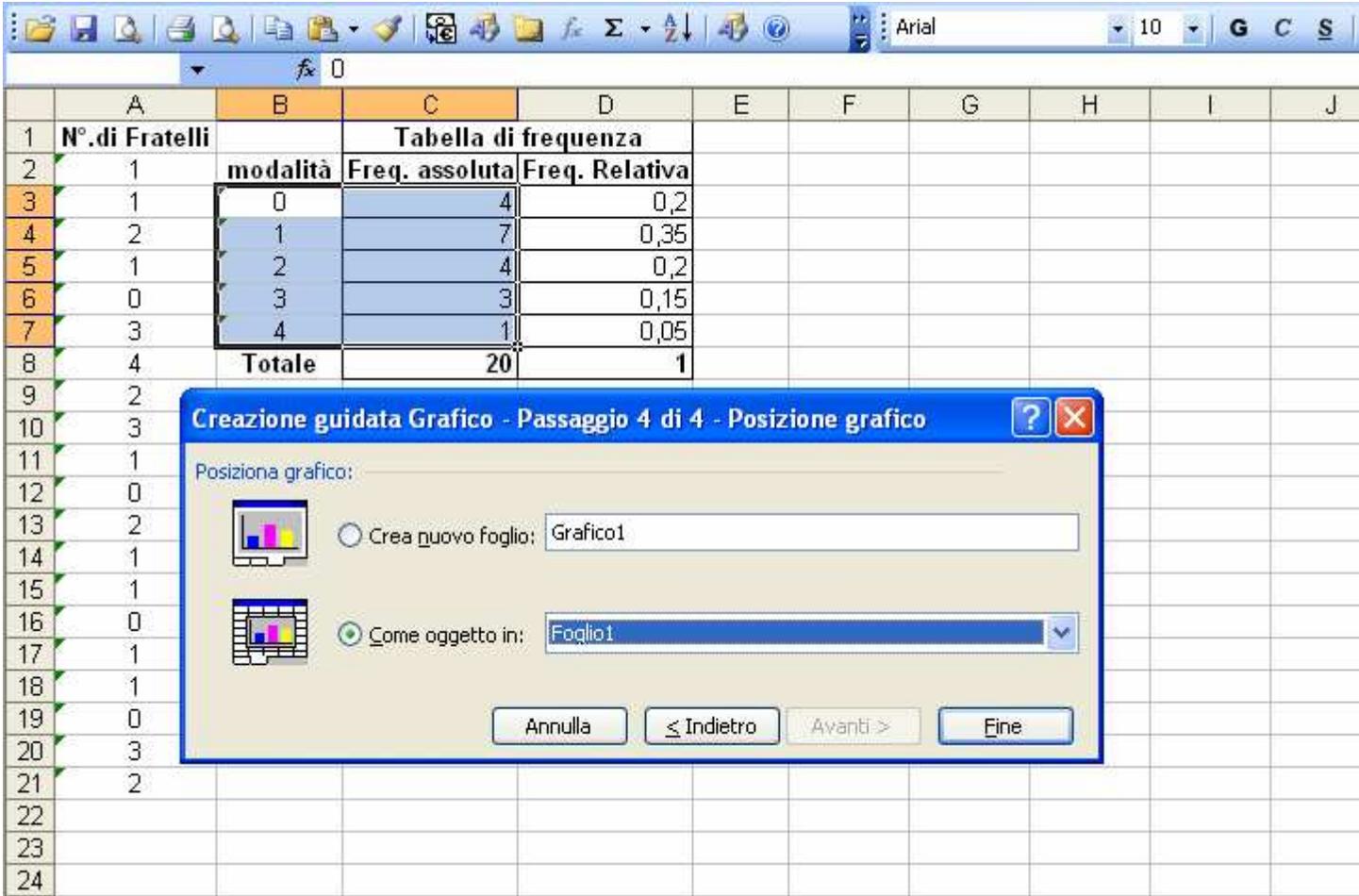


[Home](#)



Nell'ultimo passaggio si può scegliere se il grafico deve essere collocato su un Nuovo foglio di calcolo, o su uno già esistente.

Una volta effettuata la scelta cliccate su **Fine [Finish]**



The screenshot shows an Excel spreadsheet with a frequency table and a dialog box for creating a chart. The frequency table is as follows:

	A	B	C	D	E	F	G	H	I	J
1	N°.di Fratelli		Tabella di frequenza							
2	1	modalità	Freq. assoluta	Freq. Relativa						
3	1	0	4	0,2						
4	2	1	7	0,35						
5	1	2	4	0,2						
6	0	3	3	0,15						
7	3	4	1	0,05						
8	4	Totale	20	1						

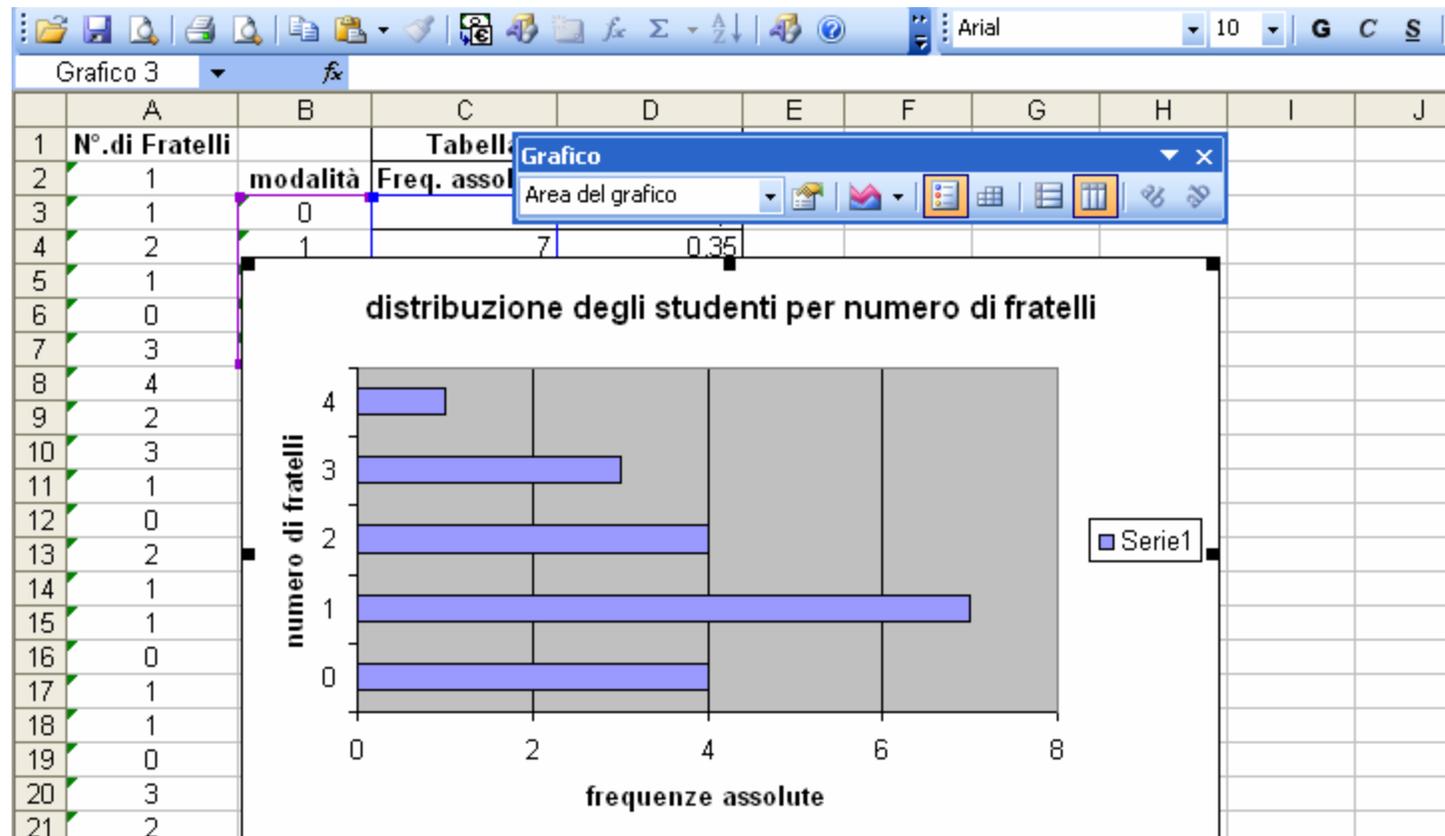
The dialog box, titled "Creazione guidata Grafico - Passaggio 4 di 4 - Posizione grafico", has the following options:

- Posizione grafico:
- Crea nuovo foglio: Grafico1
- Come oggetto in: Foglio1

Buttons at the bottom: Annulla, ≤ Indietro, Avanti >, Fine.



ed ecco il risultato:



Lo stesso procedimento potrà essere applicato alla creazione di un diagramma a barre con frequenze relative.

In questo caso, selezionare innanzitutto le due colonne contenenti i dati e i rispettivi titoli, relativi alle classi e alla frequenza relativa. Nel caso in cui le colonne non siano contigue, per selezionarle si cominci dalla prima colonna e, premendo il tasto <ctrl>, si selezioni la seconda colonna.

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Home](#)

4.4.2 Istogramma

Esercizio 2.4 – creare un Istogramma

	A	B	C	D	E
1	atleta	Tempi (min/sec)	Minimo		
2	gianni	12,1	7,5		Classi
3	maria	8,9	Massimo		[7,5;9,3[
4	luca	16,2	16,2		[9,3;11,1[
5	fabio	8,2	Ampiezza		[11,1;12,9[
6	angela	9,8	8,7		[12,9;14,7[
7	michele	15,1	n		[14,7;16,5[
8	mirko	14,5	22		Totale
9	daniela	13,4	k		
10	marco	14,7	5		
11	giuseppe	7,5	h		
12	stefano	8,8	1,74		
13	laura	12,4	h approssimato		Limiti superiori delle classi
14	chiara	16,1	1,8		9,3
15	davide	15,2			11,1
16	luigi	13,5			12,9
17	monica	14,6			14,7
18	sara	15,5			
19	giovanni	7,8			
20	aldo	12,5			
21	serena	13,2			
22	claudia	11			
23	piero	10,5			

Prima di iniziare la creazione di un istogramma, si dovranno definire:

numero di classi, ampiezza e limiti così come descritto nell' **esercizio 2.3** per la creazione della tabella di frequenza in caso di **dati (caratteri) continui**.

Se si vuole utilizzare l'opzione Istogramma [Histogram] di Excel, la frequenza assoluta di ogni classe dovrà essere calcolata considerando il limite inferiore della classe aperto e il limite superiore chiuso.

Pertanto, per creare le classi in modo che:

- abbiano tutte la stessa ampiezza
- la loro unione contenga tutti gli elementi del campione,

si deve cominciare dal valore massimo considerando le seguenti classi:

$C_{k+i+l} = [\max. \text{ del campione} - i \times h, \max. \text{ del campione} - (i-l) \times h]$ Com $i = 1, 2, \dots, k$

Come si fa:

Utilizzeremo i risultati dell'esempio 2, nel quale dovremmo aver ottenuto le seguenti classi:

[7,5; 9,3], [9,3; 11,1], [11,1; 12,9], [12,9; 14,7] e [14,7; 16,5]

Creare una tabella con i limiti superiori delle classi:

- digitare, ad esempio nella cella E13 il titolo **Limiti superiori delle classi**;

- dalla cella E14 alla cella E17 digitare ognuno dei valori dei limiti superiori di ognuna delle classi, con eccezione del limite dell'ultima classe.



Attraverso il comando **Analisi dati [Data Analysis]** presente nel menù **Strumenti [Tools]** (se non installato procedere come descritto nel capitolo [2.3](#)), selezionare l'opzione **Istogramma [Histogram]** e cliccare su **OK**.

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	atleta	Tempi (min/sec)	Minimo			
2	gianni	12,1	7,5		Classi	
3	maria	8,9	Massimo		[7,5;9,3[
4	luca	16,2	16,2		[9,3;11,1[
5	fabio	8,2	Ampiezza		[11,1;12,9[
6	angela	9,8	8,7		[12,9;14,7[
7	michele	15,1	n		[14,7;16,5[
8	mirko	14,5	22		Totale	
9	daniela	13,4				
10	marco	14,7				
11	giuseppe	7,5				
12	stefano	8,8				
13	laura	12,4				
14	chiara	16,1				
15	davide	15,2				
16	luigi	13,5				
17	monica	14,6				
18	sara	15,5				
19	giovanni	7,8				
20	aldo	12,5				
21	serena	13,2				
22	claudia	11				
23	piero	10,5				

The 'Analisi dati' dialog box is open, showing the following options under 'Strumenti di analisi':

- Analisi varianza: ad un fattore
- Analisi varianza: a due fattori con replica
- Analisi varianza: a due fattori senza replica
- Correlazione
- Covarianza
- Statistica descrittiva
- Smorzamento esponenziale
- Test F a due campioni per varianze
- Analisi di Fourier
- Istogramma**

The 'OK' button is highlighted.



[Home](#)



Compilate la finestra di dialogo: sia nella selezione dell'**intervallo di input** (tutti i tempi di gara) che in quella dell'**intervallo della classe** (questa sarebbe la tabella che abbiamo appena creato da E13 a E17 con i limiti superiori delle classi) sono da comprendersi le celle contenenti i Titoli. Biffate la casella relativa alle **Etichette** e quella di **Grafico in output**.

	A	B	C	D	E	F
1	atleta	Tempi (min/sec)	Minimo			
2	gianni	12,1	7,5		Classi	
3	maria	8,9	Massimo		[7,5;9,3[
4	luca	16,2	16,2		[9,3;11,1[
5	fabio	8,2	Ampiezza		[11,1;12,9[
6	angela	9,8	8,7		[12,9;14,7[
7	michele	15,1	n		[14,7;16,5[
8	mirko	14,5	22		Totale	
9	daniela	13,4	k			
10	marco	14,7	5			
11	giuseppe	7,5	h			
12	stefano	8,8	1,74			
13	laura	12,4	h approssimato		Limiti superiori	
14	chiara	16,1	1,8		9,3	
15	davide	15,2			11,1	
16	luigi	13,5			12,9	
17	monica	14,6			14,7	
18	sara	15,5				
19	giovanni	7,8				
20	aldo	12,5				
21	serena	13,2				
22	claudia	11				
23	piero	10,5				
24						
25						

Istogramma

Input

Intervallo di input:

Intervallo della classe:

Etichette

Opzioni di output

Intervallo di output:

Nuovo foglio di lavoro:

Nuova cartella di lavoro

Pareto (istogramma ordinato)

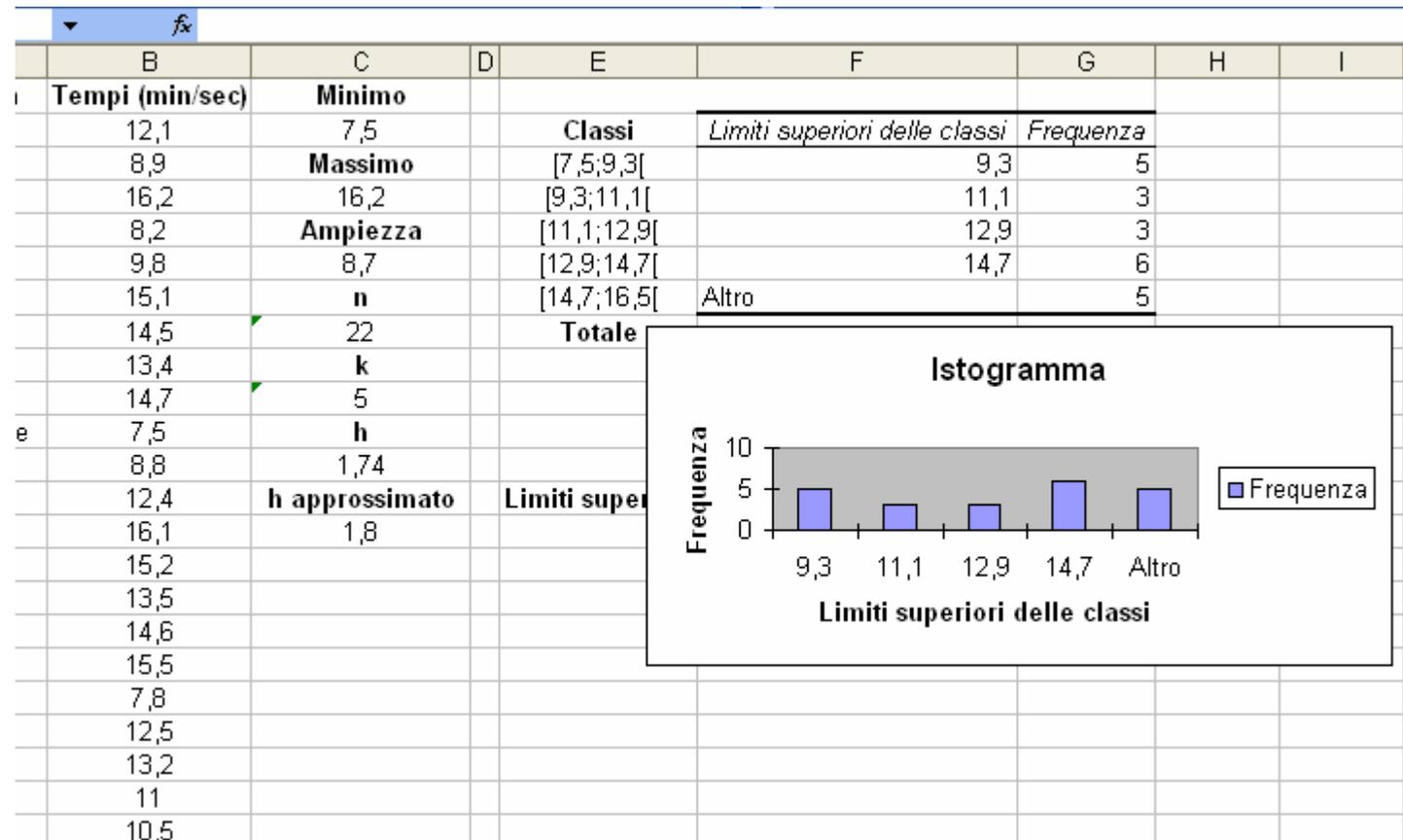
Percentuale cumulativa

Grafico in output

Dopo aver compilato la finestra di dialogo, **clickare su OK**.



Ed ecco il risultato:



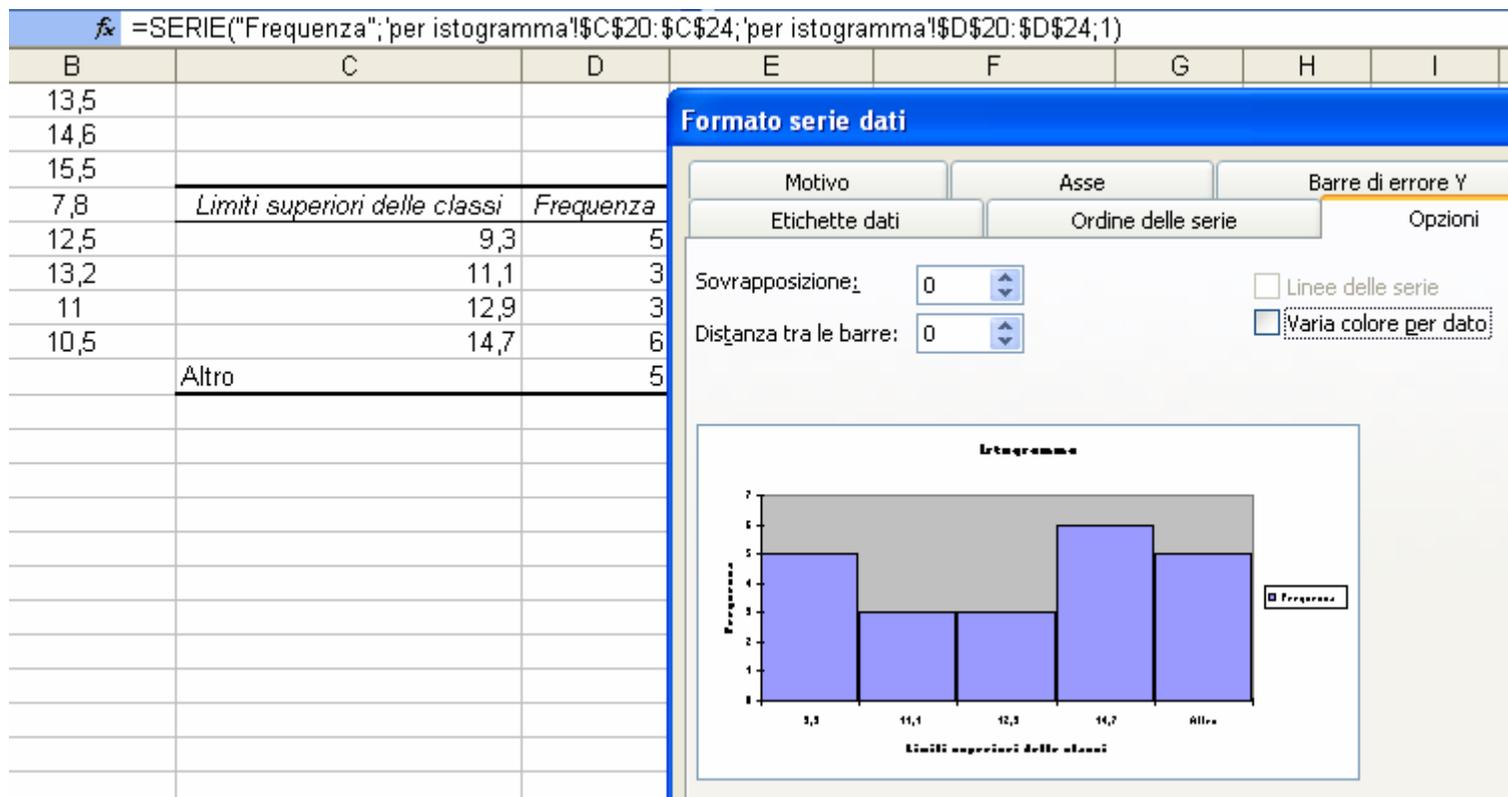
[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Home](#)

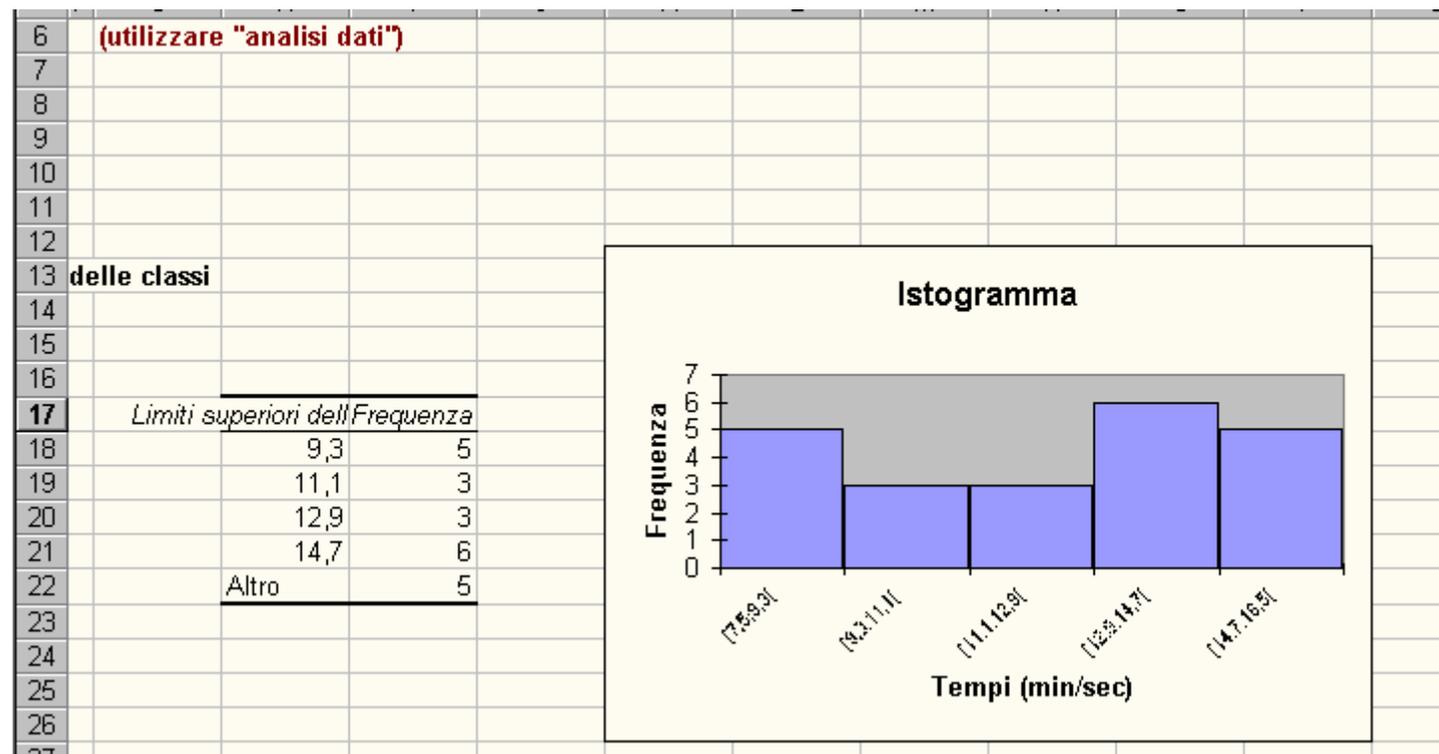
Di fatto, ciò che Excel offre è un diagramma a barre più che un vero e proprio istogramma.

Si può ridurre al minimo questo limite, eseguendo le seguenti operazioni:

- Selezionare il grafico e cliccare due volte sulla barra delle frequenze**, in modo che si apra il menù **Formato Serie Dati** [Format Data Series]. **Selezionare il sottomenù Opzioni** [Options] e **indicare** come valore di Distanza tra le barre [Gap width] il valore **0 (zero)**.
- Sostituire** nel foglio di calcolo i valori del limite superiore delle classi, con l'espressione degli intervalli di classe.



Ed ecco il risultato:



Attenzione: nella versione di Excel'97 l'opzione **Istogramma** [Histogram] non permette la creazione (disegno) di istogrammi con ampiezze di classe variabili. Dunque per visualizzare un istogramma elaborato in modo corretto con quella versione di Excel, bisogna accedere alla pagina del DISCUSS e aprire il modulo di statistica descrittiva (DESC.XLS) (http://www.mis.coventry.ac.uk/research/discus/discus_home.html).

[link foglio.xls](#) (clicca se vuoi ripetere l'esercitazione COMPLETA)

[Home](#)

4.4.3 Diagramma degli estremi e quartili

Si tratta di una rappresentazione grafica in cui sono evidenziate alcune caratteristiche del **campione**.

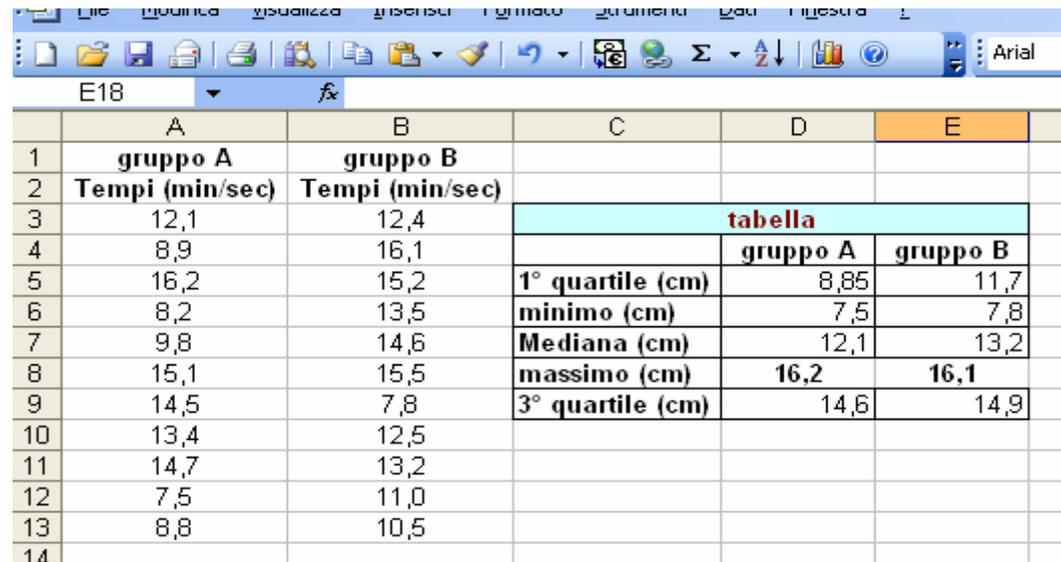
Per creare il diagramma degli estremi e quartili, sono necessari cinque parametri statistici: la **mediana** (vedi [cap. 5.3](#)), il **primo** ed il **terzo quartile** (vedi [cap. 5.4](#)), il **massimo** e il **minimo**.

Esercizio 2.5 – composizione del diagramma degli estremi e quartili

Immaginiamo che la corsa podistica (i cui tempi di percorrenza abbiamo utilizzato nell'esempio 1) si sia effettuata dividendo gli atleti in due batterie (A e B) ciascuna composta da 11 atleti.

Costruiamo una tabella nella quale riporteremo i tempi cronometrati per ciascun atleta di entrambe le batterie.

Utilizziamo i dati riportati nelle colonne A e B per ottenere i valori di 1° quartile, minimo, Mediana, massimo e 3° quartile.



	A	B	C	D	E
1	gruppo A	gruppo B			
2	Tempi (min/sec)	Tempi (min/sec)			
3	12,1	12,4	tabella		
4	8,9	16,1		gruppo A	gruppo B
5	16,2	15,2	1° quartile (cm)	8,85	11,7
6	8,2	13,5	minimo (cm)	7,5	7,8
7	9,8	14,6	Mediana (cm)	12,1	13,2
8	15,1	15,5	massimo (cm)	16,2	16,1
9	14,5	7,8	3° quartile (cm)	14,6	14,9
10	13,4	12,5			
11	14,7	13,2			
12	7,5	11,0			
13	8,8	10,5			
14					

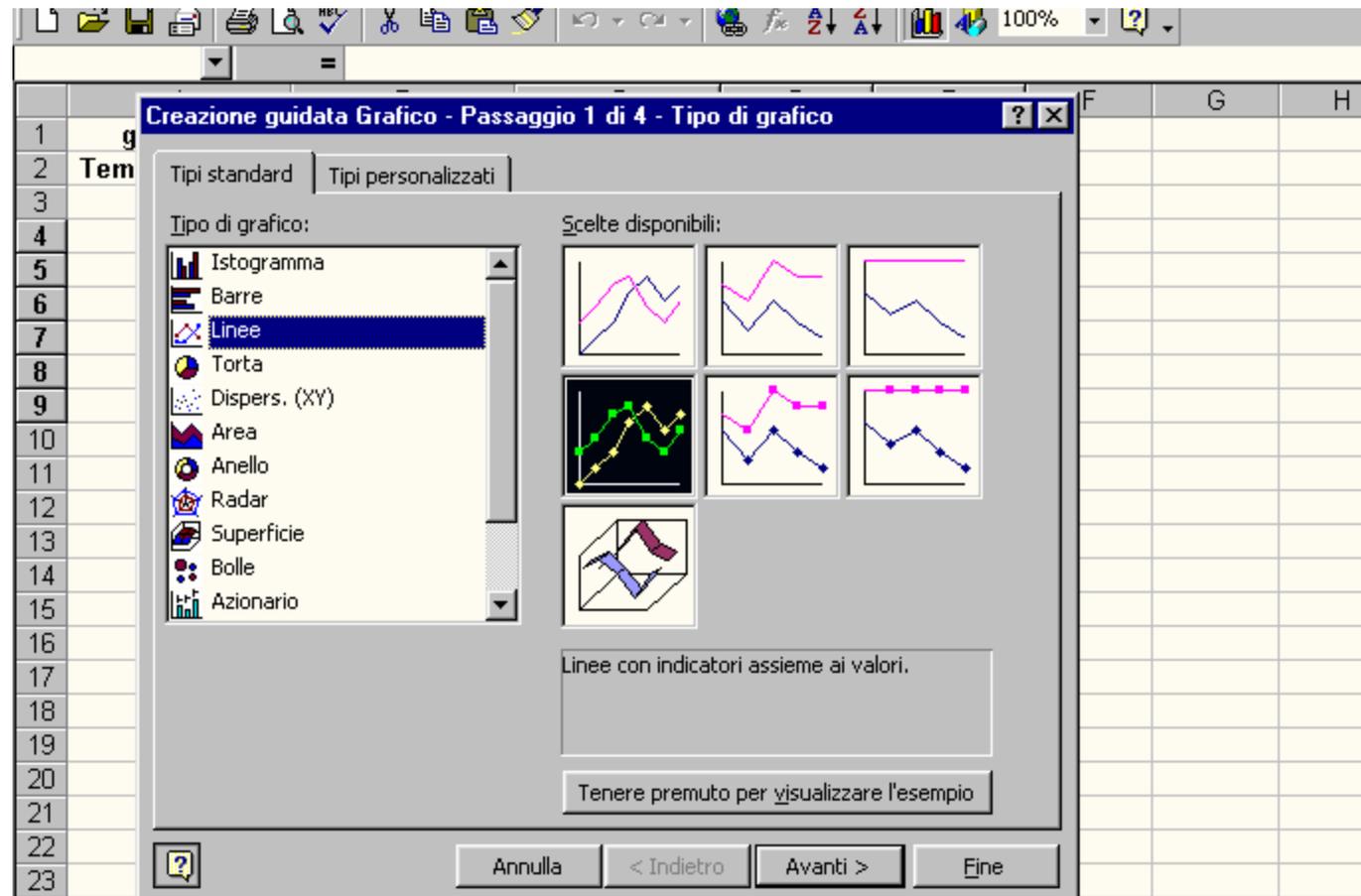
Dopo aver digitato la tabella si può facilmente ottenere il diagramma degli estremi e quartili, mediante la funzione **Creazione guidata grafico** [Chart Wizard].

Selezionare innanzitutto le celle contenenti i dati e i rispettivi titoli e **clickare sull'icona Creazione guidata di grafico** sulla Barra degli strumenti.



Nella prima Finestra di dialogo **selezionare l'opzione Linee [Line]**.

Per continuare la creazione del grafico e per andare al Passaggio [Step] successivo, **clickare su Avanti [Next]**.



Home



Per continuare la creazione del grafico, **selezionare l'opzione Serie in: Righe** [Series in: Lines] e **clickare su fine** [Finish].

Intervallo dati:

Serie in:

- Righe
- Colonne

Intervallo dati:

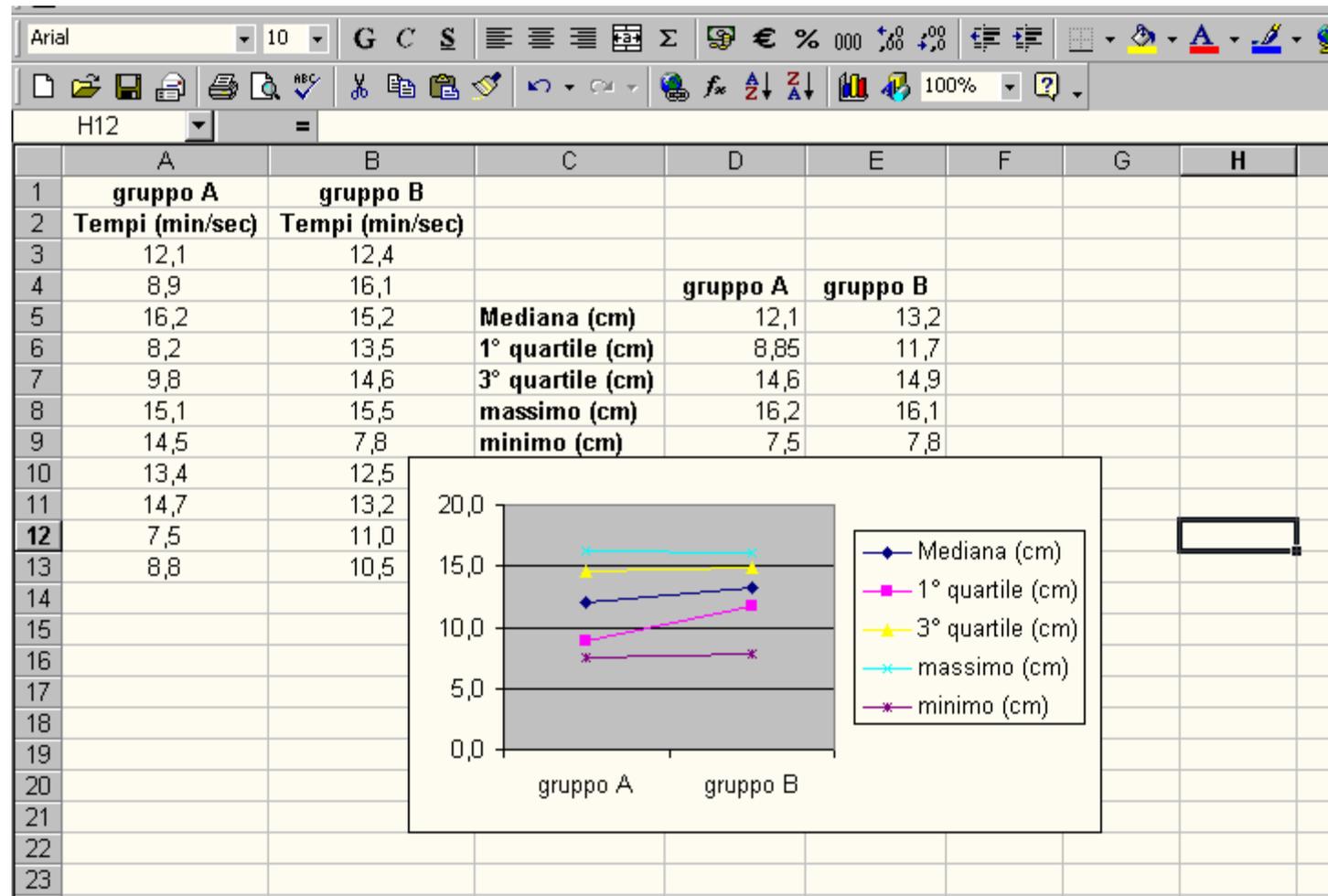
Serie in:

- Righe
- Colonne

Buttons: Annulla, < Indietro, Avanti >, Fine



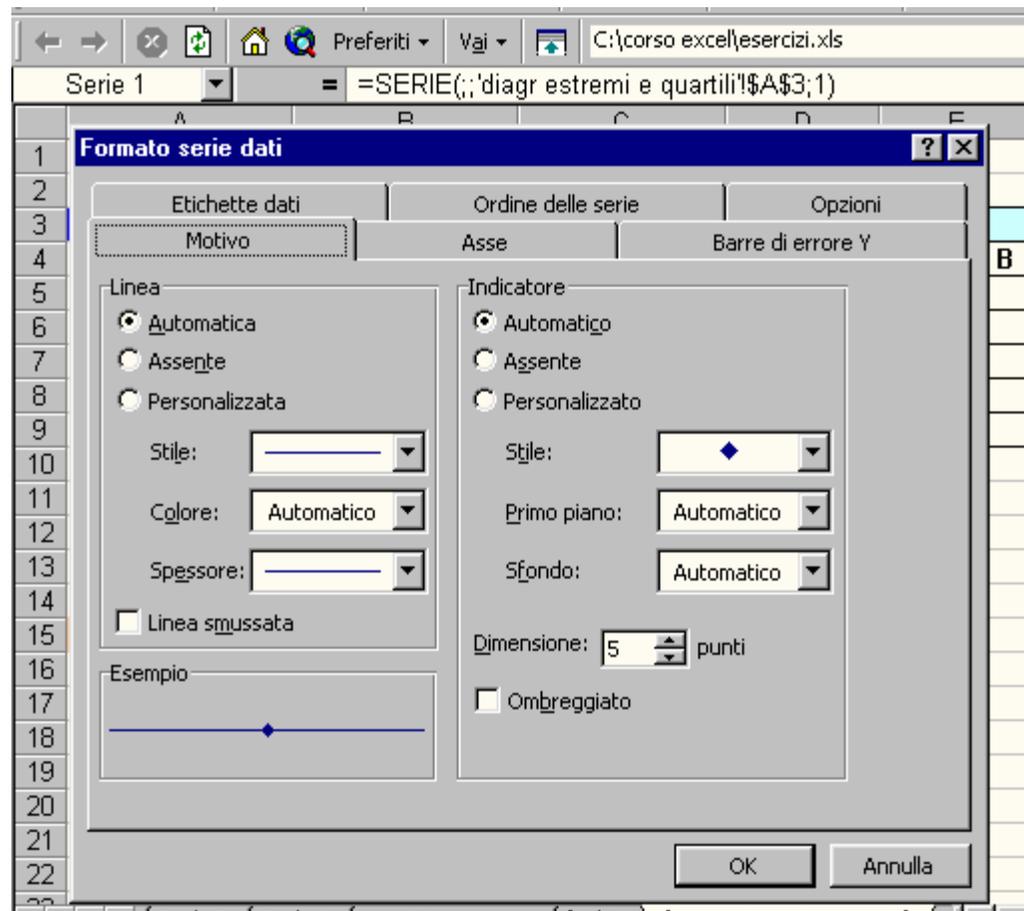
ed ecco il risultato:



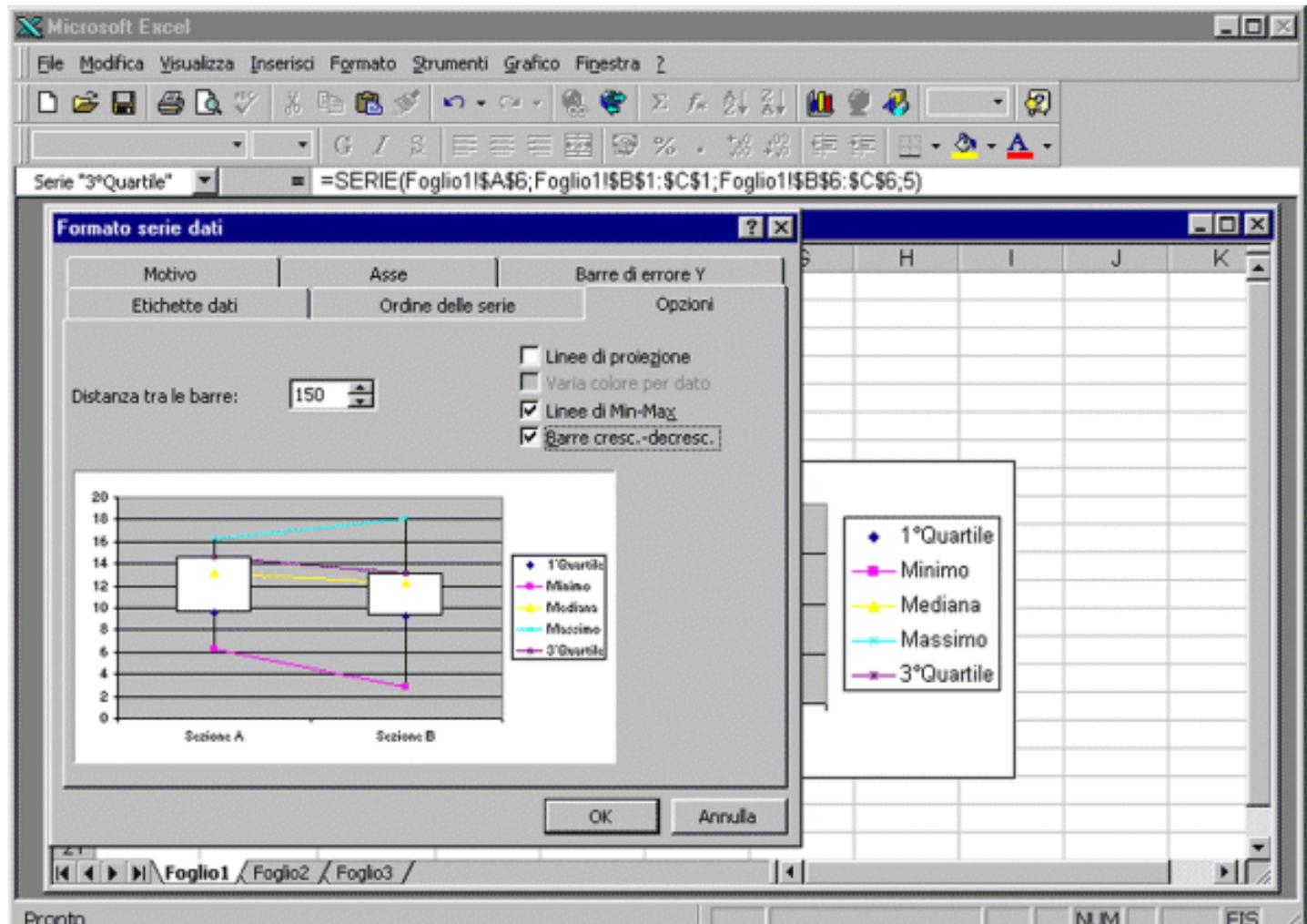
Ora **selezionare** il grafico e **clickare due volte** su una delle linee in modo che si apra il menù **Formato serie dati** [*Format Data Serie*].

[Home](#)

Selezionare il sottomenù **Motivo** [Patterns] e selezionare nell'opzione linea [Line] la voce **Assente** [None]. Ripetere il procedimento per tutte le serie.



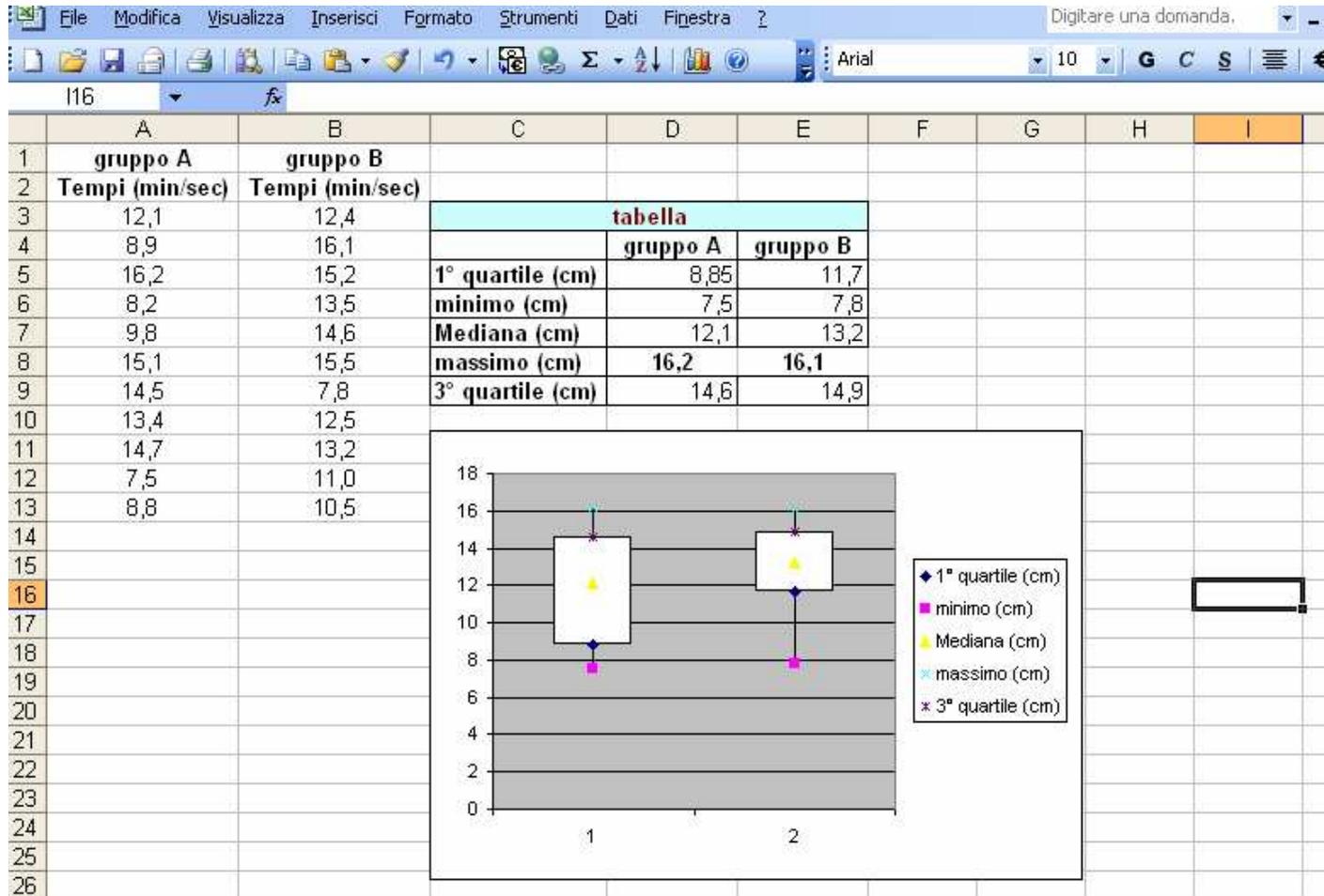
Selezionare di nuovo il grafico e **clickare due volte** su una serie in modo che si apra il menù **Formato serie dati** [*Format Data Series*]. **Selezionare il sottomenù Opzioni** [*Options*] e **scegliere le opzioni: Linee di Min-Max** [*High-Low lines*] e **Barre cresc.-decresc** [*Up-Down bars*].



Home



Provando le diverse funzioni di formato, si potrà ottenere il seguente grafico:



[link foglio.xls](#) (clicca per ripetere l'esercitazione)



4.4.4 Diagramma circolare

Analogamente a quanto detto circa l'elaborazione del diagramma a barre, una volta ottenuta una tabella di frequenza, si può ottenere facilmente un diagramma circolare utilizzando la finestra **Creazione guidata grafico** [Chart Wizard].

Esercizio 1.6 – creare un grafico circolare

Selezionare innanzitutto le celle contenenti i dati e il rispettivo titolo e **clickare** sull'icona della Barra degli strumenti.

Nella prima Finestra di dialogo **selezionare l'opzione Torta** [Pie].

Per continuare la creazione del grafico e per andare al Passaggio [Step] successivo e **clickare** su **Avanti** > [Next>].

	A	B	C	D
1	N°.di Fratelli		Tabella di frequenza	
2	1	modalità	Freq. assoluta	Freq. Relativa
3	1	0	4	0,2
4	2	1	7	0,35
5	1	2	4	0,2
6	0	3	3	0,15
7	3	4	1	0,05
8	4	Totale	20	1
9	2			
10	3			
11	1			
12	0			
13	2			
14	1			
15	1			
16	0			
17	1			
18	1			

Creazione guidata Grafico - Passaggio 1 di 4 - Tipo di grafico

Tipi standard | Tipi personalizzati

Tipo di grafico:

- Istogramma
- Barre
- Linee
- Torta**
- Dispers. (XY)
- Area
- Anello
- Radar
- Superficie
- Bolle
- Azionario

Scelte disponibili:

Torta. Visualizza il contributo di ciascun valore al totale.

Home

Per proseguire con la creazione del grafico, e per andare al Passaggio [Step] successivo **clickare** su **Avanti** > [Next>].

	A	B	C	D
1	N°.di Fratelli		Tabella di frequenza	
2	1	modalità	Freq. assoluta	Freq. Relativa
3	1	0	4	0,2
4	2	1	7	0,35
5	1	2	4	0,2
6	0	3	3	0,15
7	3	4	1	0,05
8	4	Totale	20	1
9	2			
10	3			
11	1			
12	0			
13	2			
14	1			
15	1			
16	0			
17	1			
18	1			

Creazione guidata Grafico - Passaggio 1 di 4 - Tipo di grafico

Tipi standard | Tipi personalizzati

Tipo di grafico:

- Istogramma
- Barre
- Linee
- Torta**
- Dispers. (XY)
- Area
- Anello
- Radar
- Superficie
- Bolle
- Azionario

Scelte disponibili:

Torta. Visualizza il contributo valore al totale.



Nel terzo passaggio, la Finestra di dialogo presenta varie opzioni che permettono di formattare il grafico:

- In **Etichette dati** [Data labels] **selezionare l'opzione Mostra percentuali** [Show percentages].

Tabella di frequenza				
	modalità	Freq. assoluta	Freq. Relativa	
1	0	4	0,2	
2	1	7	0,35	
3	2	4	0,2	
4	3	3	0,15	
5	4	1	0,05	

Creazione guidata Grafico - Passaggio 3 di 4 - Opzioni del grafico

Titoli | Legenda | **Etichette dati**

Etichette dati

- Assente
- Mostra valore
- Mostra percentuale**
- Mostra etichetta
- Mostra etichetta e percent.
- Mostra dimensioni bolla

Chiave legenda vicino etichetta

Mostra guide

Num. di fratelli per studente: distribuzione delle frequenze relative

Numero di fratelli	Frequenza Relativa (%)
0	5%
1	35%
2	20%
3	15%
4	5%

Per continuare la creazione del grafico e per andare al Passaggio [Step] successivo **clickare** su **Avanti >** [Next>].



[Home](#)



Nell'ultimo passaggio si può scegliere se il grafico deve essere collocato in un nuovo foglio di calcolo o in un foglio già esistente.

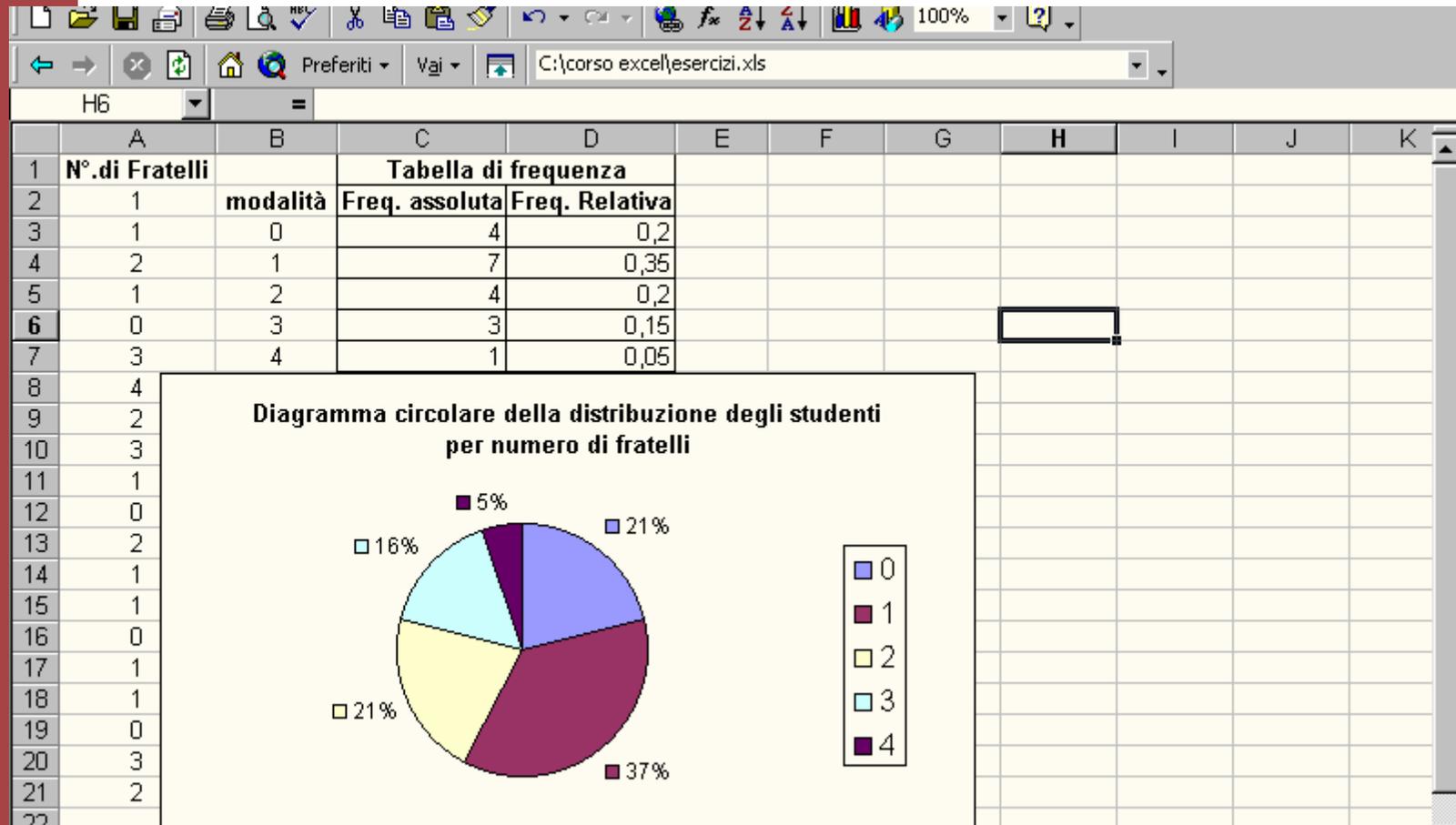
The screenshot shows an Excel spreadsheet with a frequency table and a dialog box for creating a chart. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	N°.di Fratelli		Tabella di frequenza								
2	1	modalità	Freq. assoluta	Freq. Relativa							
3	1	0	4	0,2							
4	2	1	7	0,35							
5	1	2	4	0,2							
6	0	3	3	0,15							
7	3	4	1	0,05							
8	4	Totale	20	1							
9	2										
10	3										
11	1										
12	0										
13	2										
14	1										
15	1										
16	0										
17	1										
18	1										
19	0										
20	3										
21	2										
22											

The dialog box "Creazione guidata Grafico - Passaggio 4 di 4 - Posizione grafico" is open, showing the "Posizione grafico:" section. The "Crea nuovo foglio:" option is selected, and the name "Grafico1" is entered in the text box. The "Come oggetto in:" option is also visible, with "Foglio1" selected in the dropdown menu. The dialog box includes buttons for "Annulla", "Indietro", "Avanti", and "Fine".

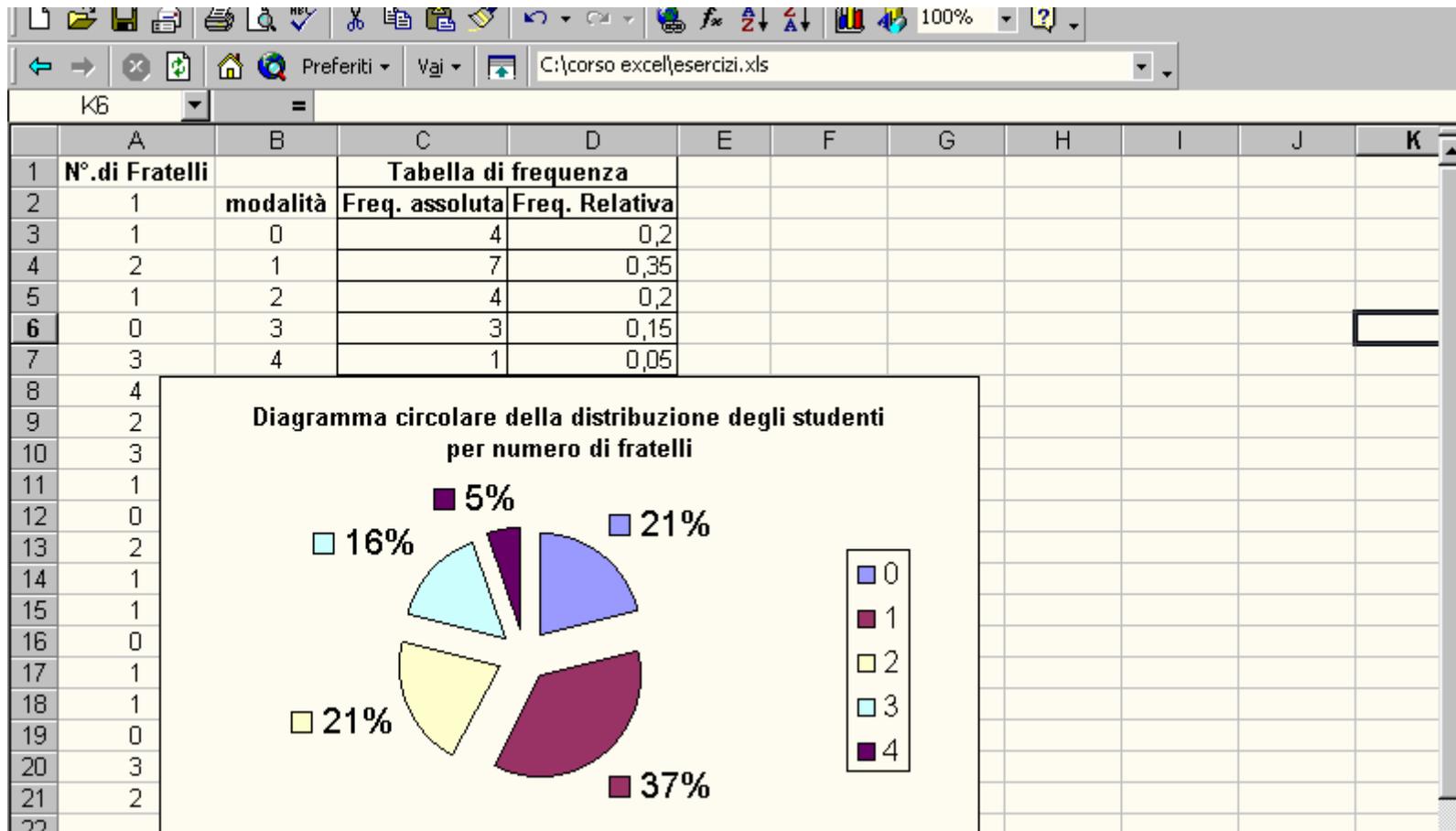


Cliccando su **Fine [Finish]** si otterrà il seguente risultato



[Home](#)

Provando le diverse funzioni di formattazione si potrà ottenere il seguente grafico



[link foglio.xls](#) (clicca per ripetere l'esercitazione)



5 Gli indici di posizione

Lo scopo principale di un valore medio è di fornire un valore numerico capace di rappresentare sinteticamente tutti i dati di un certo insieme.

I valori medi più utilizzati sono

- la media aritmetica
- la moda
- la mediana

[Home](#)

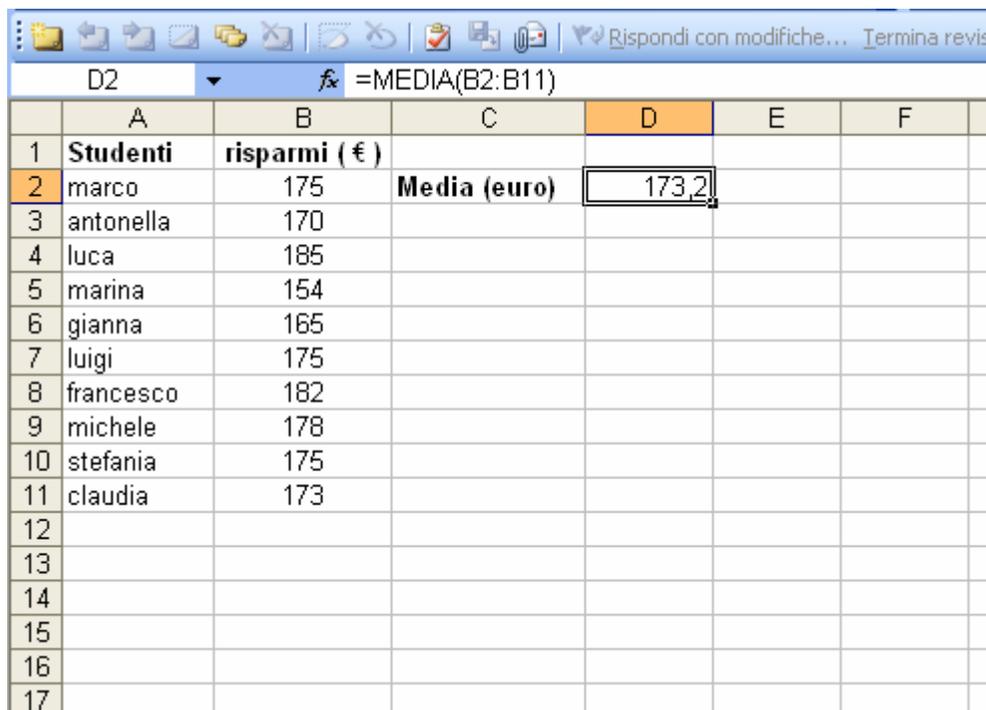


5.1 Media aritmetica

La **media** si determina attraverso la funzione **MEDIA** [AVERAGE].

Il risultato di questa funzione è la media aritmetica

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	Studenti	risparmi (€)				
2	marco	175	Media (euro)	173,2		
3	antonella	170				
4	luca	185				
5	marina	154				
6	gianna	165				
7	luigi	175				
8	francesco	182				
9	michele	178				
10	stefania	175				
11	claudia	173				
12						
13						
14						
15						
16						
17						

The formula bar shows the formula: `=MEDIA(B2:B11)`

Determinazione della media aritmetica:

Come si fa:

Sono elencati i risparmi (in euro) accantonati da un gruppo di dieci studenti in un periodo di sei mesi;

Spostare il cursore nella cella, ad esempio, C2 e digitare il Titolo: Media (euro)

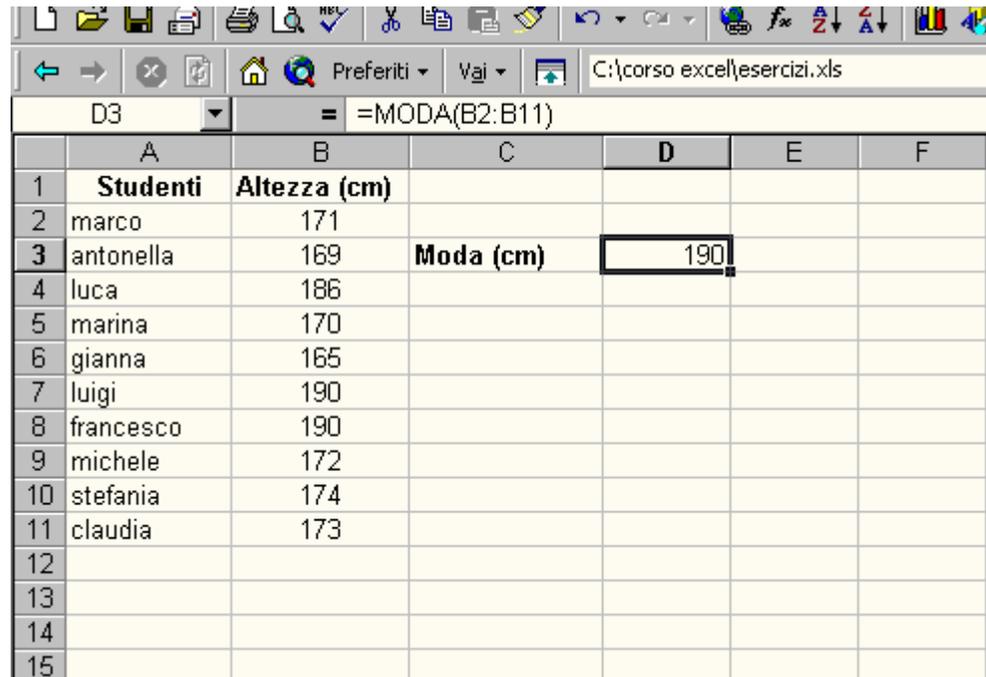
-- Spostare il cursore nella cella, ad esempio, D2 e inserire la funzione: **=MEDIA(B2:B11)**.



5.2 Moda

La **moda** di un collettivo, distribuito secondo un carattere, è la modalità prevalente del carattere ossia quella a cui è associata la massima frequenza.

Si determina mediante la funzione **MODA** [MODE].



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	Studenti	Altezza (cm)				
2	marco	171				
3	antonella	169	Moda (cm)	190		
4	luca	186				
5	marina	170				
6	gianna	165				
7	luigi	190				
8	francesco	190				
9	michele	172				
10	stefania	174				
11	claudia	173				
12						
13						
14						
15						

The formula bar shows the formula: `=MODA(B2:B11)`. The cell D3 contains the result: 190.

Determinazione della moda:

Come si fa:

-Spostare il cursore nella cella C3 e digitare il Titolo: Moda (cm).

- Spostare il cursore nella cella D3 e inserire la funzione: =MODA (B2:B11).



[Torna](#)

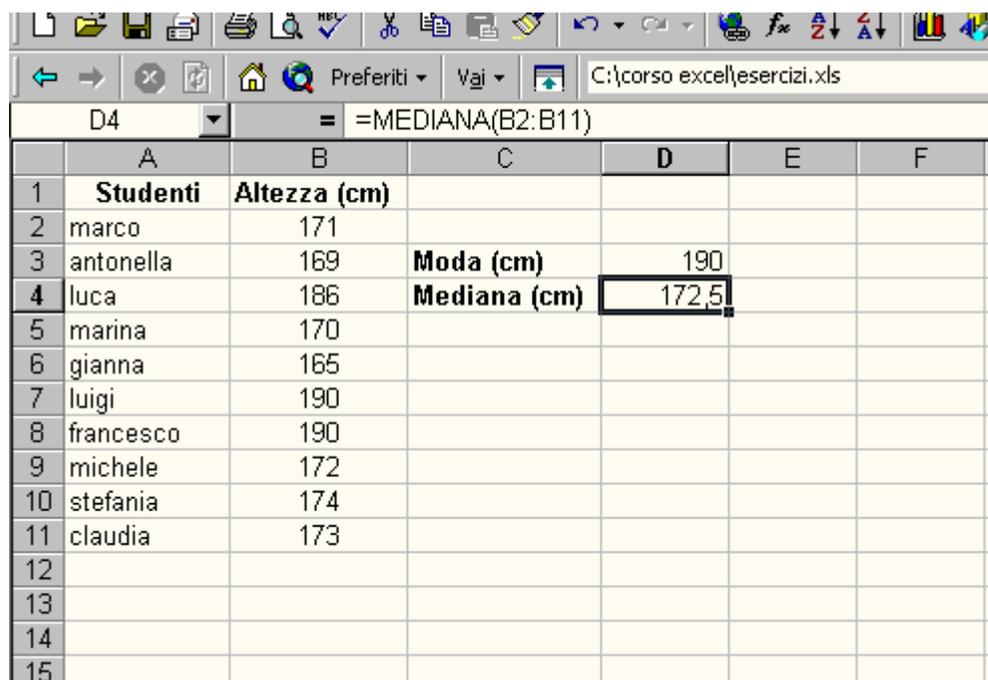
[Home](#)



5.3 Mediana

La **mediana** suddivide ogni distribuzione ordinata in due distribuzioni aventi ciascuna una numerosità (o una quantità) che è il 50% della numerosità (o della quantità) della distribuzione totale.

Si determina mediante la funzione **MEDIANA** [*MEDIAN*].



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	Studenti	Altezza (cm)				
2	marco	171				
3	antonella	169	Moda (cm)	190		
4	luca	186	Mediana (cm)	172,5		
5	marina	170				
6	gianna	165				
7	luigi	190				
8	francesco	190				
9	michele	172				
10	stefania	174				
11	claudia	173				
12						
13						
14						
15						

Determinazione della mediana:

Come si fa:

- Spostare il cursore nella cella C4 e digitare il Titolo: Mediana (cm).
- Spostare il cursore nella cella D4 e inserire la funzione: **=MEDIANA (B2:B11)**.

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



[Torna](#) [Home](#)



5.4 Quantili – Quartili - Percentili

Si è visto che la mediana suddivide ogni distribuzione ordinata in due distribuzioni aventi ciascuna una numerosità (o quantità) pari al 50% della numerosità (o della quantità) della distribuzione totale.

Talvolta può essere utile suddividere la distribuzione in q distribuzioni parziali, aventi la q -esima parte della numerosità (o quantità) della distribuzione totale.

I quantili sono le n parti in cui è stata suddivisa una distribuzione.

Per $q = 3$ si parla di **terzili**, per $q = 4$ di **quartili**, per $q = 5$ di **quintili**, per $q = 6$ di **sestili**, per $q = 10$ di **decili**, per $q = 100$ di **centili**.

Di questi, i più usati sono i quartili, che dividono la distribuzione in quattro parti aventi ognuna il 25% della numerosità (o della quantità) totale, mentre la distribuzione rimanente è il 75% del totale;

Il **I quartile** è il limite superiore della distribuzione che ha il 25% della numerosità (o della quantità) totale, mentre la distribuzione rimanente è il 75% del totale;

Il **II quartile** è il limite superiore della seconda distribuzione e quindi da solo separa nella distribuzione totale due distribuzioni che hanno ciascuna il 50% della numerosità (o quantità) della distribuzione totale, e perciò il II quartile coincide con la [mediana](#);

Il **III quartile** può anche essere visto come il limite superiore della distribuzione che ha il 75% dell'ammontare della distribuzione totale e il limite inferiore della distribuzione relativa al 25% del totale.



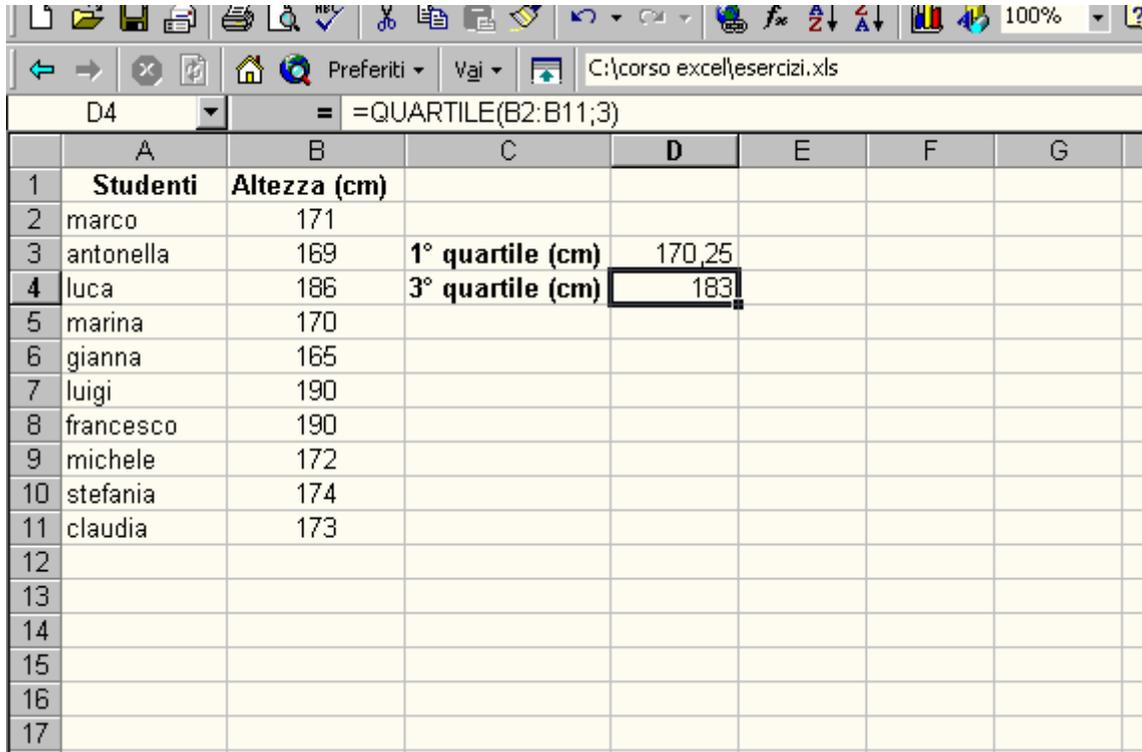
[Torna](#)

[Home](#)



I **quartili** si determinano mediante le funzioni **QUARTILE [QUARTILE]** e **PERCENTILE [PERCENTILE]**.

QUARTILE (sequenza di numeri o indirizzo di cella; 0 o 1 o 2 o 3 o 4) {0 =minimo; 1 = 1° quartile; 2 = mediana; 3 = 3° quartile; 4 = massimo}



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1	Studenti	Altezza (cm)					
2	marco	171					
3	antonella	169	1° quartile (cm)	170,25			
4	luca	186	3° quartile (cm)	183			
5	marina	170					
6	gianna	165					
7	luigi	190					
8	francesco	190					
9	michele	172					
10	stefania	174					
11	claudia	173					
12							
13							
14							
15							
16							
17							

The formula bar shows the formula for cell D4: `=QUARTILE(B2:B11;3)`.

Determinazione dei quartili:

Come si fa:

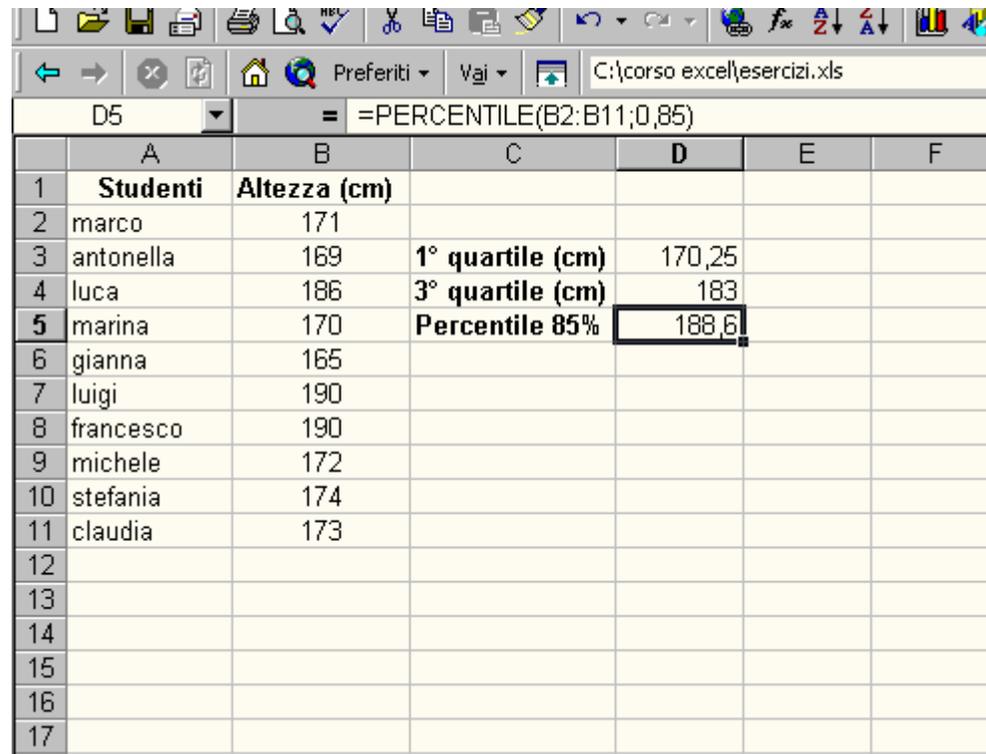
- Spostare il cursore nella cella C3 e digitare: 1° Quartile (cm).
- Spostare il cursore nella cella D3 e inserire la funzione: **=QUARTILE (B2:B11;1)**.
- Spostare il cursore nella cella C4 e digitare: 3° Quartile (cm).
- Spostare il cursore nella cella D4 e inserire la funzione: **=QUARTILE (B2:B11;3)**.



[Torna](#) [Home](#)



PERCENTILE (sequenza di numeri o indirizzo di cella; numero compreso tra 0 ed 1) {percentile p%: inserire il numero p}



	A	B	C	D	E	F
1	Studenti	Altezza (cm)				
2	marco	171				
3	antonella	169	1° quartile (cm)	170,25		
4	luca	186	3° quartile (cm)	183		
5	marina	170	Percentile 85%	188,6		
6	gianna	165				
7	luigi	190				
8	francesco	190				
9	michele	172				
10	stefania	174				
11	claudia	173				
12						
13						
14						
15						
16						
17						

Determinazione dei percentili:

Come si fa:

- Spostare il cursore nella cella C5 e digitare: Percentile 85% (cm)
- Spostare il cursore nella cella D5 e inserire la funzione: **=PERCENTILE (B2:B11;0,85).**

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Home](#)



6. Misure di dispersione

Indici che danno una misura o della variabilità dei valori della distribuzione rispetto a una media (o di quanto i valori stessi differiscono tra di loro).

Esamineremo brevemente:

- 6.1 [Varianza](#)
- 6.2 [Devianza standard](#)
- 6.3 [Ampiezza](#)
- 6.4 [Ampiezza interquartile](#)

[Torna](#)

[Home](#)

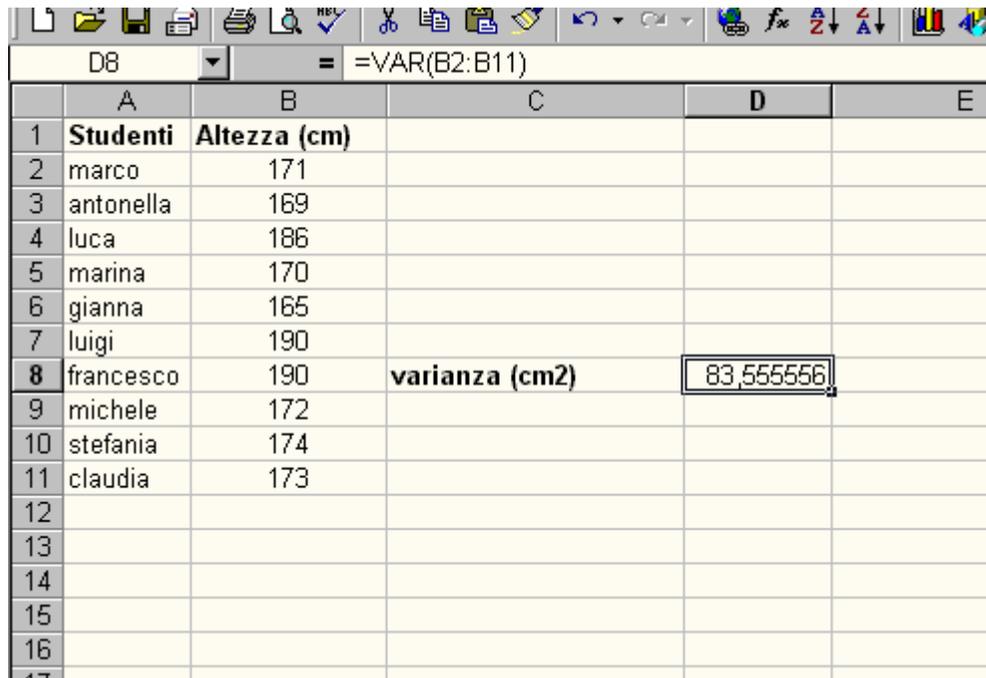


6.1 Varianza

La varianza si determina attraverso la funzione **VAR [VAR]**.
(VAR sequenza di numeri o indirizzi di cella)

Il risultato di questa funzione è la varianza campionaria (s^2) dei valori introdotti
come argomento:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E
1	Studenti	Altezza (cm)			
2	marco	171			
3	antonella	169			
4	luca	186			
5	marina	170			
6	gianna	165			
7	luigi	190			
8	francesco	190	varianza (cm2)	83,555556	
9	michele	172			
10	stefania	174			
11	claudia	173			
12					
13					
14					
15					
16					
17					

Come si fa:

Sono elencate le altezze (in cm) di 10 studenti;

- Spostare il cursore nella cella, ad esempio, C8 e digitare il Titolo: Varianza (cm2)

- Spostare il cursore nella cella, ad esempio, D8 e inserire la funzione: **=VAR(B2:B11)**



[Home](#)



6.2 Devianza standard

La **devianza standard** si determina attraverso la funzione **DEV.ST** [STDEV].
(DEV.ST sequenza di numeri o indirizzi di cella)

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

	A	B	C	D	E
1	Studenti	Altezza (cm)			
2	marco	171			
3	antonella	169			
4	luca	186			
5	marina	170			
6	gianna	165			
7	luigi	190			
8	francesco	190	varianza (cm2)	83,555556	
9	michele	172	devianza standard	9,1408728	
10	stefania	174			
11	claudia	173			
12					
13					
14					
15					
16					
17					

Come si fa:

- Spostare il cursore nella cella, ad esempio, C9 e digitare il Titolo: devianza standard (cm)

- Spostare il cursore nella cella, ad esempio, D8 e inserire la funzione: **=DEV.ST(B2:B11)**



6.3 Ampiezza

Ampiezza del campione

È un'altra misura di variabilità: si ottiene come differenza tra l'estremo superiore e quello inferiore dei valori osservati del campione.

Si determina ricorrendo alla funzione **MAX [MAX] MIN [MIN]**.

(MAX sequenza di numeri o indirizzi di cella)

(MIN sequenza di numeri o indirizzi di cella)

	A	B	C	D	E	F	G
1	Altezza (cm)						
2	175						
3	170						
4	185						
5	154						
6	165						
7	175						
8	182	varianza (cm2)	77,28889				
9	178	devianza standard	8,79141				
10	175	ampiezza	31				
11	173						
12							
13							
14							
15							
16							
17							

Come si fa:

- Spostare il cursore nella cella B10 e digitare: Ampiezza (cm).

- Spostare il cursore nella cella C10 e inserire la funzione:

=MAX(A2:A11) - MIN(A2:A11).

Spiegazione: determiniamo gli estremi del campione (massimo e minimo); la differenza tra codesti valori corrisponde all'ampiezza del campione.

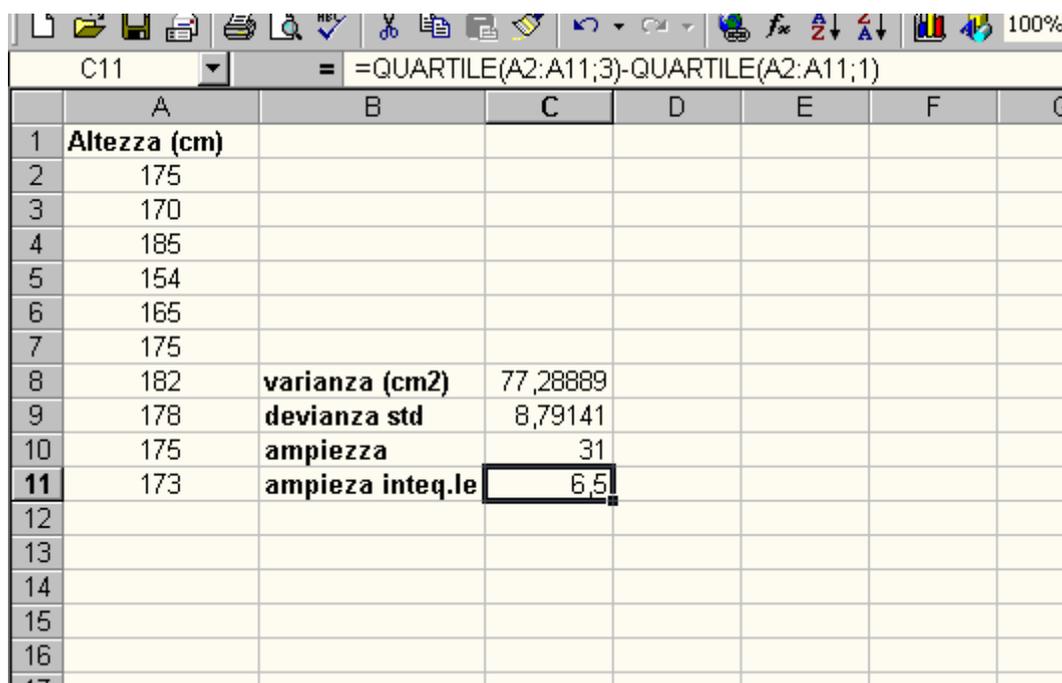


[Home](#)



6.4 Ampiezza interquartile

L'**ampiezza interquartile** si determina ricorrendo alla funzione **QUARTILE [QUARTILE]**, descritta nel capitolo **4.4**.



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1	Altezza (cm)						
2	175						
3	170						
4	185						
5	154						
6	165						
7	175						
8	182	varianza (cm2)	77,28889				
9	178	devianza std	8,79141				
10	175	ampiezza	31				
11	173	ampiezza inteq.le	6,5				
12							
13							
14							
15							
16							
17							

The formula bar at the top shows the formula for cell C11: `=QUARTILE(A2:A11;3)-QUARTILE(A2:A11;1)`.

Come si fa:

- Spostare il cursore nella cella B11 e digitare: Ampiezza interquartile (cm).

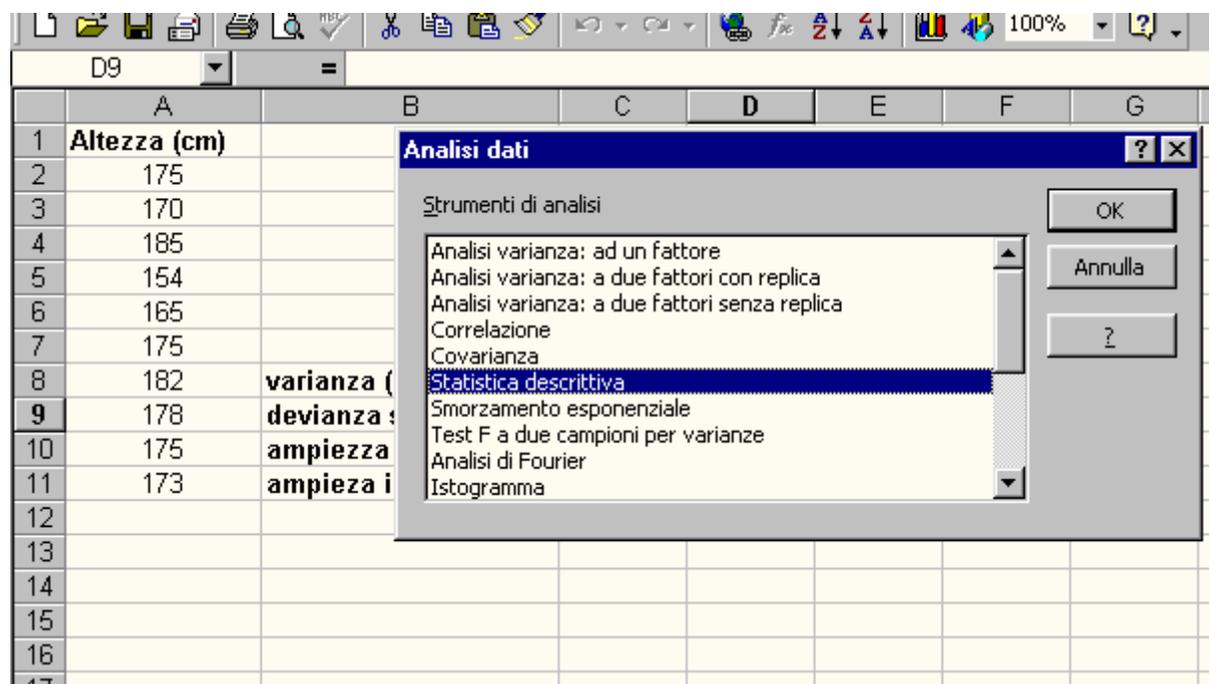
- Spostare il cursore nella cella C11 ed inserire la funzione:
**=QUARTILE(A2:A11;3) -
QUARTILE (A2:A11;1).**

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



Excel permette di calcolare i diversi valori di posizione e di dispersione, **in modo più rapido**, mediante il comando **Analisi dati** [Data analysis], presente nel menù **Strumenti** [Tools].

Accedendo al comando **Analisi dati** [Data analysis], selezionare l'opzione **Statistica descrittiva** [Descriptive statistics] e successivamente cliccare su **OK**.



[Home](#)



Compilate la Finestra di dialogo avendo cura di biffare su:

“Etichette nella prima riga”

“Intervallo di output” – selezionare una cella –

“Riepilogo statistiche”

clickare su OK.

The screenshot shows the 'Statistica descrittiva' dialog box in Microsoft Excel. The background spreadsheet has the following data:

Altezza (cm)
175
170
185
154
165
175
182
178
175
173

The dialog box settings are as follows:

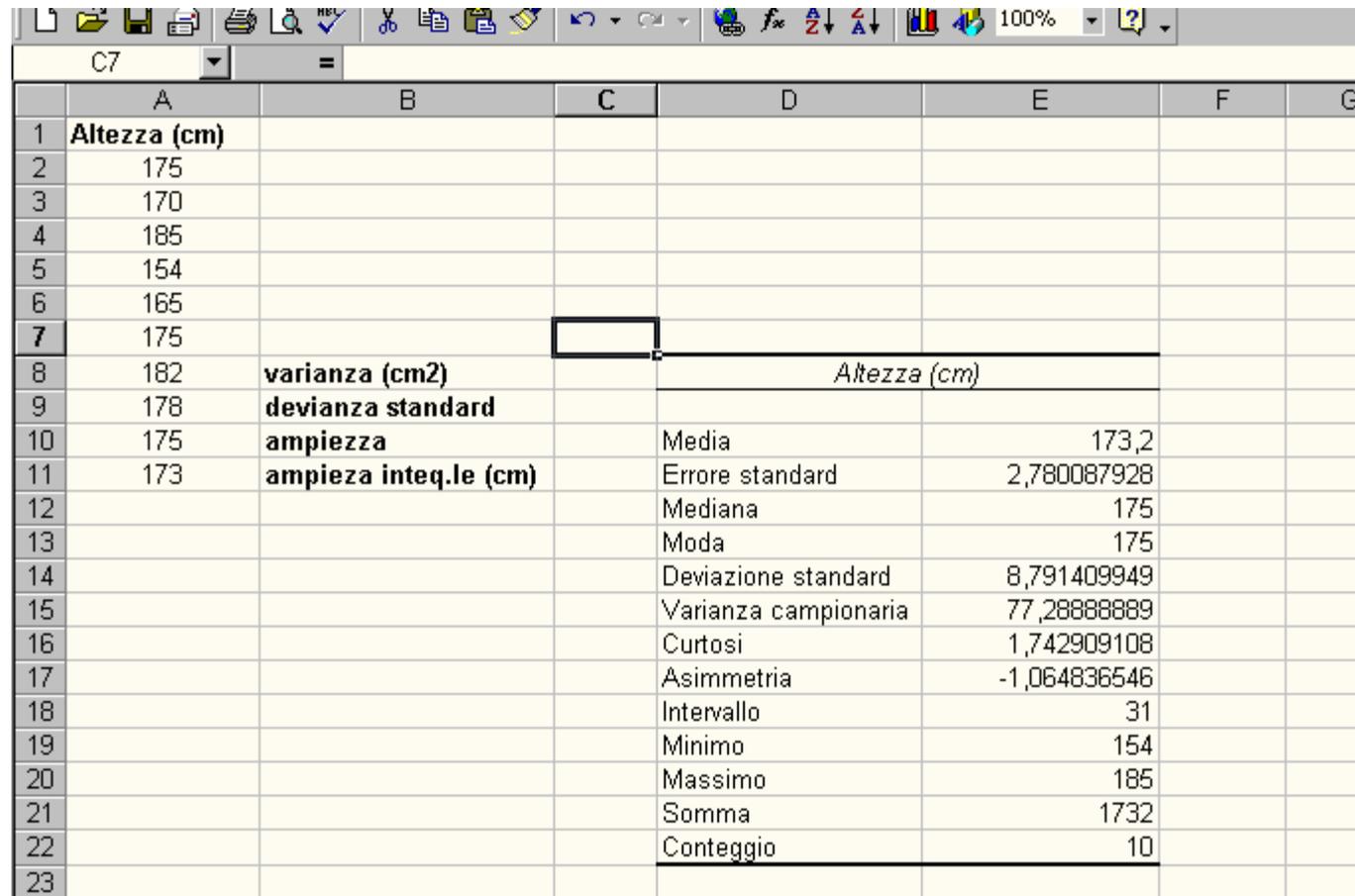
- Input: Intervallo di input:
- Dati raggruppati per: Colonne, Righe
- Etichette nella prima riga
- Opzioni di output: Intervallo di output:
 Nuovo foglio di lavoro:
 Nuova cartella di lavoro
- Riepilogo statistiche
- Livello di confidenza per media: %
- K-esimo più grande:
- K-esimo più piccolo:

Buttons: OK, Annulla, ?



Per evidenziare i risultati di questo comando formattare la tabella dei risultati.

Come si può notare, la lista delle statistiche fornita dall'opzione **Statistica descrittiva** [*Descriptive statistics*] include misure non descritte in precedenza (per esempio, curtosi, asimmetria), **ma non include i quantili** (ad eccezione della mediana).



The screenshot shows an Excel spreadsheet with a list of heights in column A and their descriptive statistics in columns B through G. The data is as follows:

	A	B	C	D	E	F	G
1	Altezza (cm)						
2	175						
3	170						
4	185						
5	154						
6	165						
7	175						
8	182	varianza (cm2)		Altezza (cm)			
9	178	devianza standard					
10	175	ampiezza		Media		173,2	
11	173	ampiezza inteq.le (cm)		Errore standard		2,780087928	
12				Mediana		175	
13				Moda		175	
14				Deviazione standard		8,791409949	
15				Varianza campionaria		77,28888889	
16				Curtosi		1,742909108	
17				Asimmetria		-1,064836546	
18				Intervallo		31	
19				Minimo		154	
20				Massimo		185	
21				Somma		1732	
22				Conteggio		10	
23							

[link foglio.xls](#) (clicca per ripetere l'esercitazione)

[Home](#)

Si può ottenere il rango e il **quantile** (percentile) di ognuno dei dati presentati, mediante l'opzione **Rango e percentile** [Rank and Percentile], dal comando **Analisi dati** [Data analysis].

	A	B	C	D	E
1	Altezza (cm)				
2	175				
3	170				
4	185				
5	154				
6	165				
7	175				
8	182				
9	178				
10	175				173,2
11	173				087928
12					175
13					175
14				Deviazione standard	8,791409949
15				Varianza campionaria	77,28888889
16				Curtosi	1,742909108
17				Asimmetria	-1,064836546
18				Intervallo	31
19				Minimo	154
20				Massimo	185
21				Somma	1732
22				Conteggio	10
23					

Dopo aver compilato la Finestra di dialogo, cliccare su OK.



Rango e percentile

Input

Intervallo di input:

Dati raggruppati per: Colonne Righe

Etichette nella prima riga

Opzioni di output

Intervallo di output:

Nuovo foglio di lavoro:

Nuova cartella di lavoro

	A	B	C	D	E	F
1	Altezza (cm)					
2	175					
3	170					
4	185					
5	154					
6	165					
7	175					
8	182					
9	178					
10	175					173,2
11	173					0087928
12						175
13						175
14						1409949
15				Varianza campionaria		77,28888889
16				Curtosi		1,742909108
17				Asimmetria		-1,064836546
18				Intervallo		31
19				Minimo		154
20				Massimo		185
21				Somma		1732
22				Conteggio		10
23						

[Home](#)

ed ecco il risultato:

	<i>Altezza (cm)</i>		<i>Dato</i>	<i>Altezza (cm)</i>	<i>Rango</i>	<i>Percento</i>
8	varianza (cm²)					
9	devianza standard		3	185	1	100,00%
10	ampiezza	Media	7	182	2	88,80%
11	ampiezza inteq.le (cm)	Errore standard	8	178	3	77,70%
12		Mediana	1	175	4	44,40%
13		Moda	6	175	4	44,40%
14		Deviazione standard	9	175	4	44,40%
15		Varianza campionaria	10	173	7	33,30%
16		Curtosi	2	170	8	22,20%
17		Asimmetria	5	165	9	11,10%
18		Intervallo	4	154	10	,00%
19		Minimo				
20		Massimo				
21		Somma				
22		Conteggio				
23						

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



Home



7. Dati bivariati

Coppia di caratteri corrispondenti ad un determinato individuo o risultato sperimentale.

Per illustrare lo studio di **dati bivariati** ricorreremo (qui di seguito in 7.1) all'esercizio già presentato nel **capitolo 5.2** aggiungendo, per ognuno dei 10 individui descritti (altezza) in precedenza, il **peso** espresso in Kg. :

72, 65, 80, 57, 60, 77, 83, 79, 67 e 68, rispettivamente

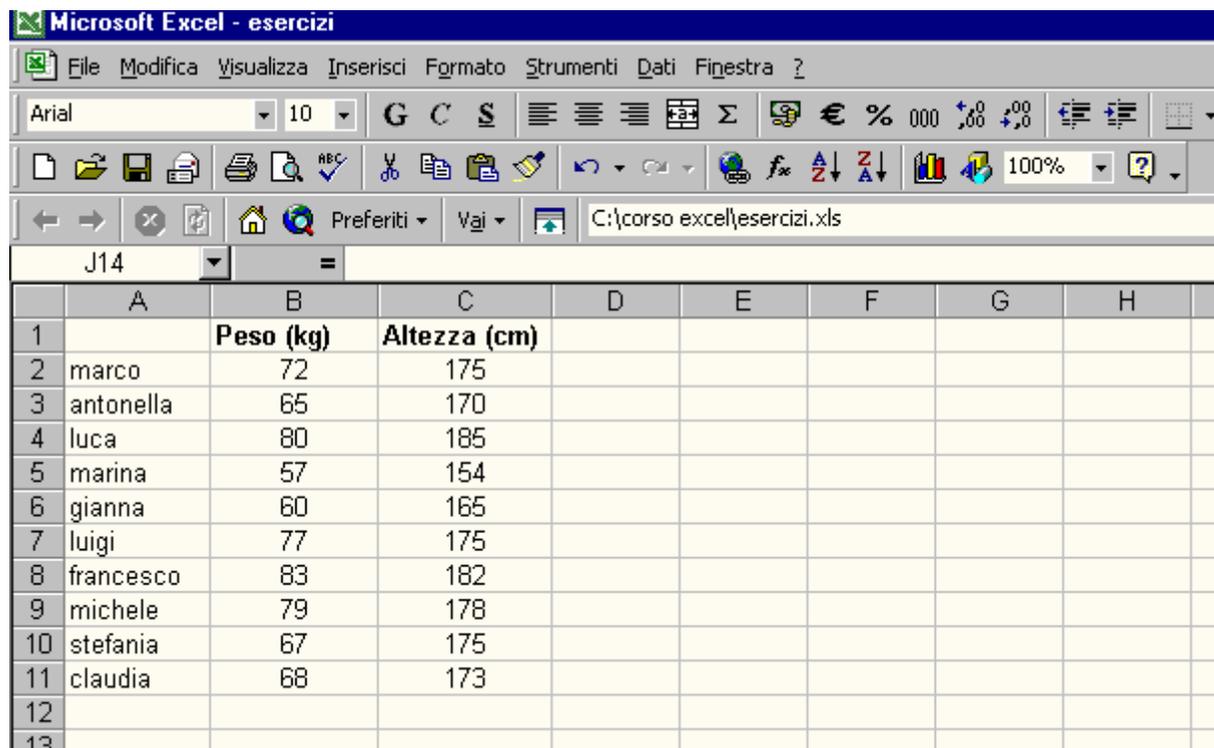
[Torna](#)

[Home](#)



7.1 Diagramma di dispersione

La rappresentazione grafica dei dati bivariati può essere ottenuta facilmente mediante la funzione **Creazione guidata Grafico** [Chart Wizard].



The screenshot shows the Microsoft Excel interface with a spreadsheet containing the following data:

	A	B	C	D	E	F	G	H
1		Peso (kg)	Altezza (cm)					
2	marco	72	175					
3	antonella	65	170					
4	luca	80	185					
5	marina	57	154					
6	gianna	60	165					
7	luigi	77	175					
8	francesco	83	182					
9	michele	79	178					
10	stefania	67	175					
11	claudia	68	173					
12								
13								

Come si fa:

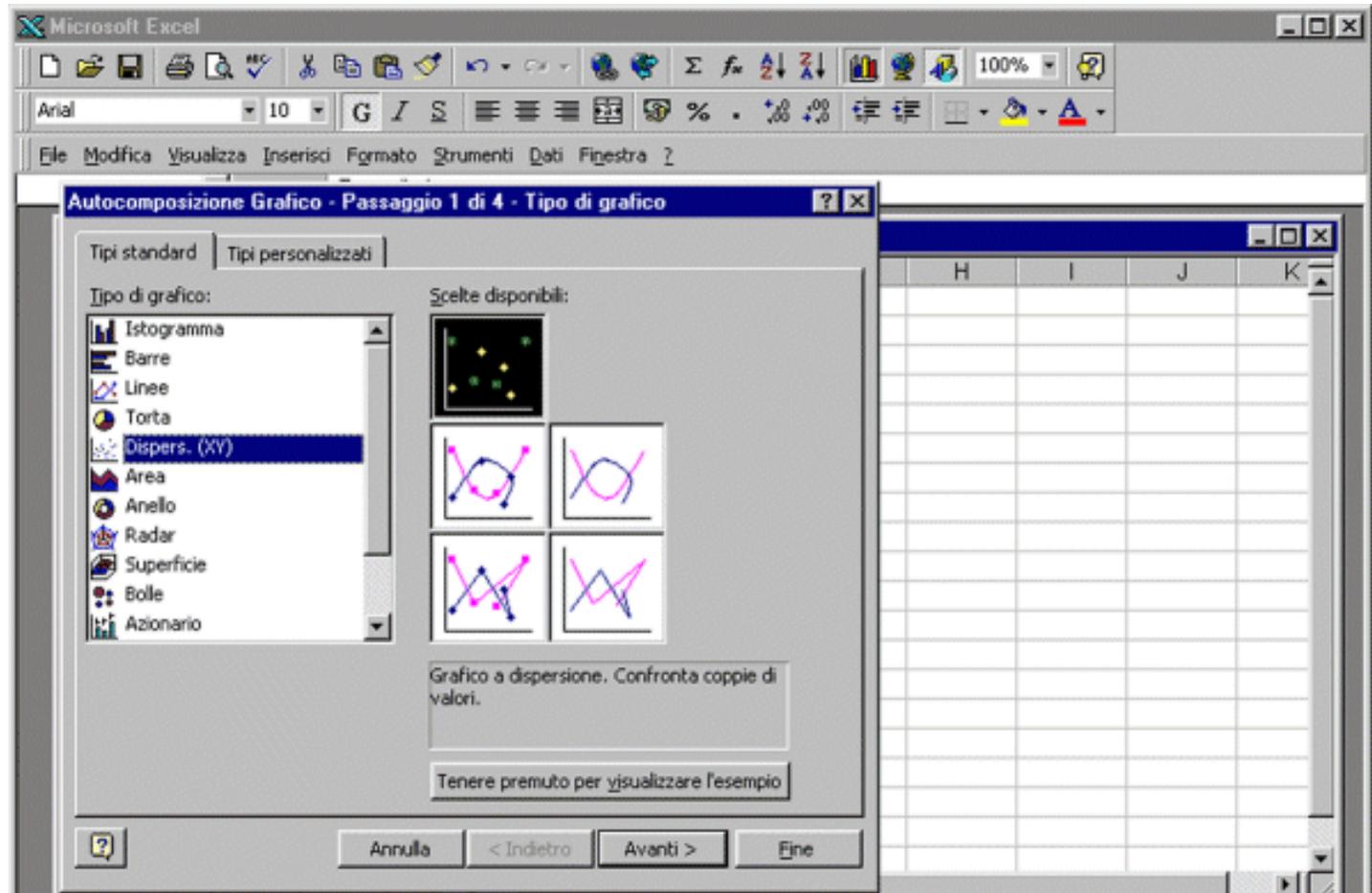
- Selezionare innanzitutto le celle contenenti i dati e i rispettivi titoli
- cliccare sull'icona **Creazione guidata Grafico** [Chart Wizard] (il grafico stilizzato) dalla Barra degli strumenti.



[Torna](#) [Home](#)



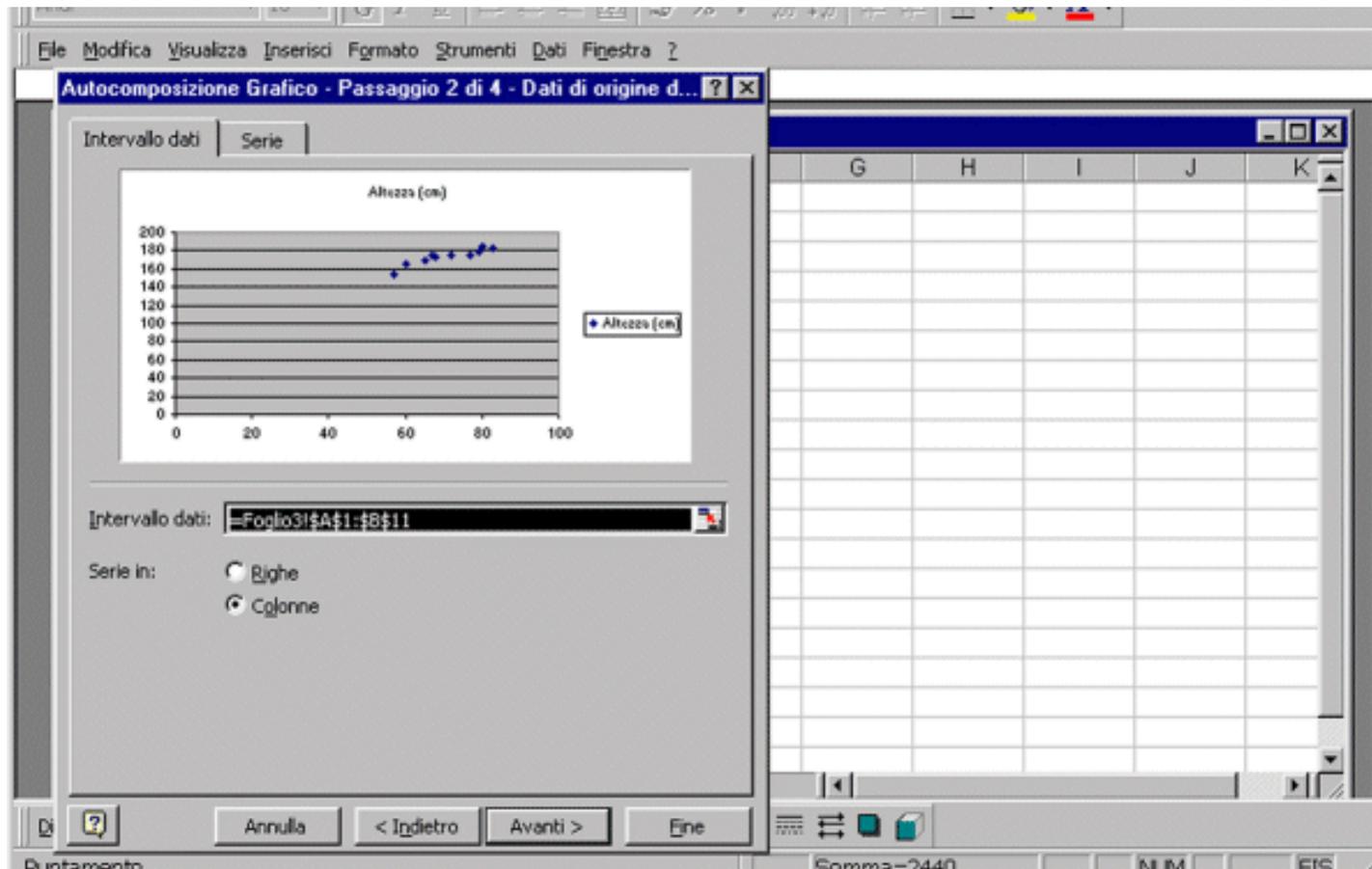
Nella prima Finestra di dialogo selezionare l'opzione **Dispers.(xy)** [XY (Scatter)].
Per continuare la creazione del grafico e per andare al Passaggio [Step] successivo cliccare su **Avanti >** [Next>].



Home



Per proseguire con la creazione del grafico, e per andare al Passaggio [Step] successivo cliccare su **Avanti** > [Next>].



Nel terzo passaggio la Finestra di dialogo presenta varie opzioni che permettono di formattare il grafico:

- Nella scheda **Titoli** [Titles] seguire l'esempio dato.
- Nella scheda **Griglia** [Gridlines], disattivare l'opzione Griglia.
- Nella scheda **Legenda** [Legend], disattivare l'opzione legenda.

Per proseguire con la creazione del grafico, e per andare al Passaggio [Step] successivo cliccare su **Avanti** > [Next>].

	A	B	C	D	E	F	G	H	I	J
1	Peso (kg)	Altezza (cm)								
2	72	175								
3	65	170								
4	80	185								
5	57	154								
6	60	165								
7	77	175								
8	83	182								
9	79	178								
10	67	175								
11	68	173								
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										



[Home](#)



Nell'ultimo passaggio si può scegliere se collocare il grafico in un nuovo foglio di calcolo o in uno già esistente.

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J
1	Peso (kg)	Altezza (cm)								
2	72	175								
3	65	170								
4	80	185								
5	57	154								
6	60	165								
7	77	175								
8	83	182								
9	79	178								
10	67	175								
11	68	173								
12										
13										
14										
15										
16										
17										
18										
19										

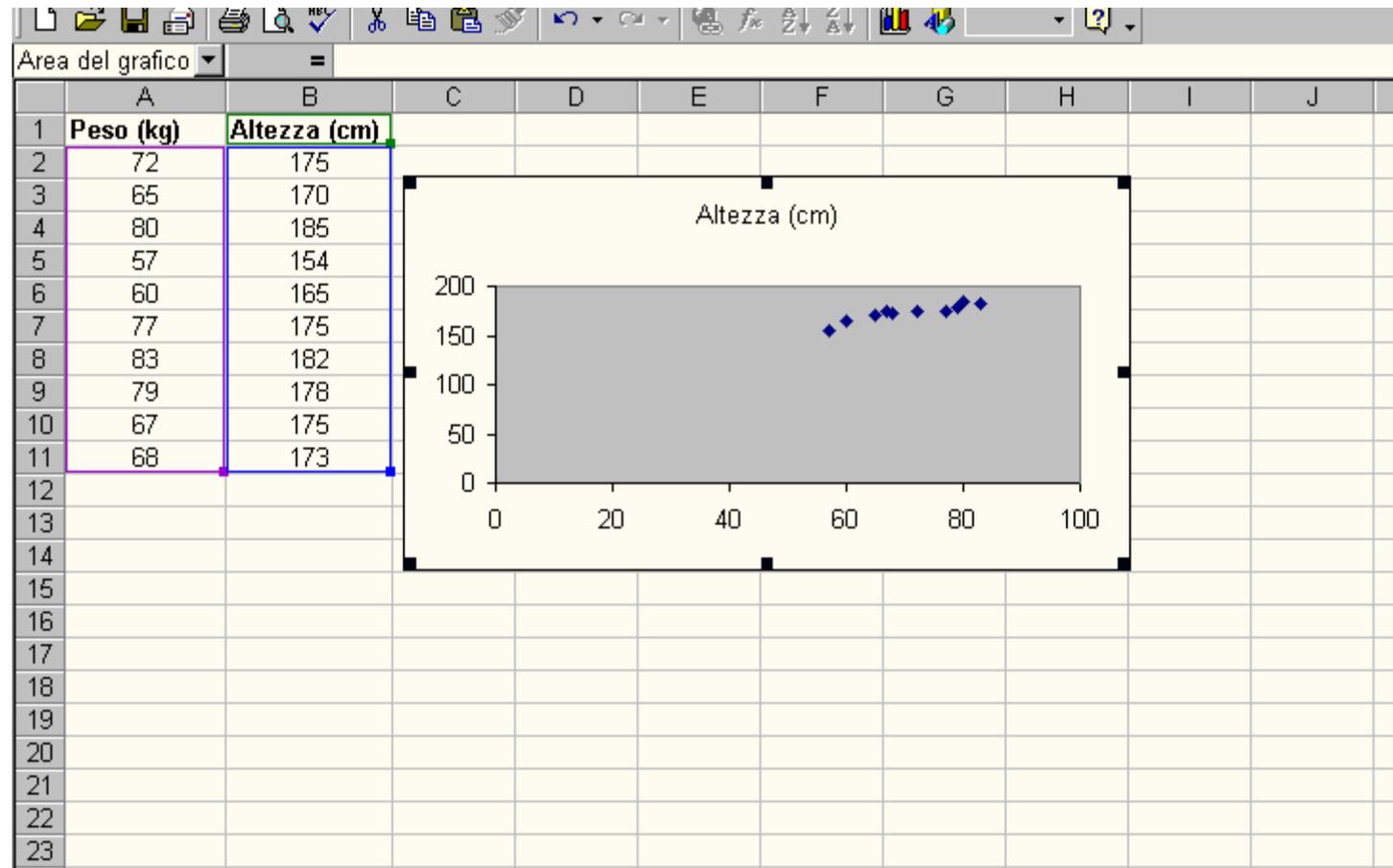
The dialog box 'Creazione guidata Grafico - Passaggio 4 di 4 - Posizione grafico' is open, showing the following options:

- Posiziona grafico: _____
- Crea nuovo foglio: Grafico1
- Come oggetto in: bivariati

Buttons: Annulla, ≤ Indietro, Avanti >, Fine



Cliccando su **Fine** [*Finish*] si otterrà il seguente risultato:

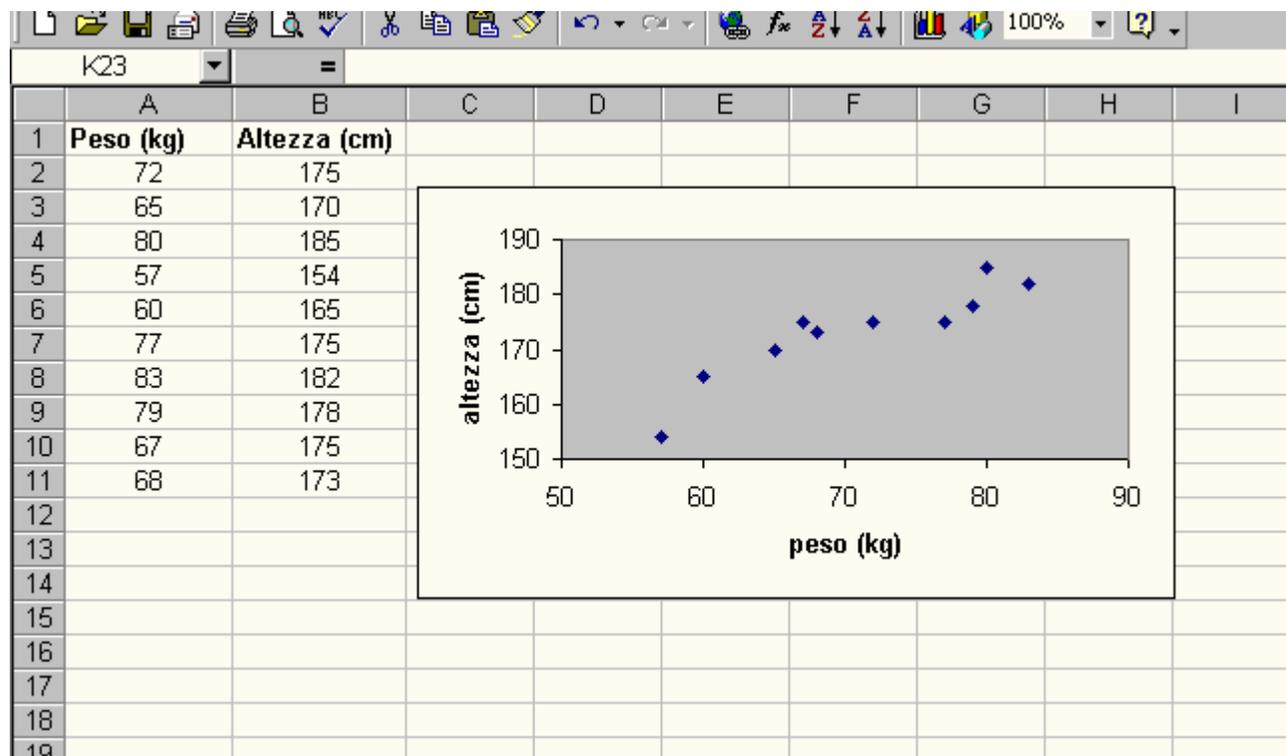


Home



Esistono molteplici opzioni per formattare i grafici di Excel, dall'aspetto generale al tipo di carattere, alla formattazione degli assi, ecc.

Ecco un esempio di che cosa si potrà ottenere.



[link foglio.xls](#) (clicca per ripetere l'esercitazione)



7.2 Coefficiente di correlazione

Quando, attraverso un grafico di dispersione, si constata l'esistenza di un'associazione lineare tra variabili, si può misurare la maggiore o minore forza con cui le variabili si associano, attraverso il coefficiente di correlazione, che in Excel corrisponde alla funzione **CORRELAZIONE [CORREL]**. (CORRELAZIONE sequenza di numeri o indirizzi di celle relativi alla prima variabile; sequenza di numeri o indirizzi di celle relativi alla seconda variabile).

Il risultato di questa funzione è il coefficiente di correlazione (r) tra i due insiemi di valori introdotti come argomento:

$$r = \frac{s_{xy}}{\sqrt{s_{xx}} \sqrt{s_{yy}}}$$

dove:

$$s_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

è *n* volte la covarianza fra X e Y.



[Torna](#)

[Home](#)



- Spostare il cursore nella cella D9 e digitare il Titolo: Correlazione.
- Spostare il cursore nella cella D10 e inserire la funzione:
=CORRELAZIONE (B2:B11; C2:C11).

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K
1		Peso (kg)	Altezza (cm)								
2	marco	72	175								
3	antonella	65	170								
4	luca	80	185								
5	marina	57	154								
6	gianna	60	165								
7	luigi	77	175								
8	francesco	83	182								
9	michele	79	178	correlazione							
10	stefania	67	175	0,906819							
11	claudia	68	173								
12											
13											
14											
15											
16											
17											
18											
19											
20											



In alternativa, accedere al comando **Analisi dati** [Data analysis] e selezionare l'opzione **Correlazione** [Correlation].

Riempire la Finestra di dialogo seguendo attentamente le istruzioni e cliccare su **OK**.

	A	B	C	D	E	F
1						
2	marco	72	175			
3	antonella	65	170			
4	luca	80	185			
5	marina	57	154			
6	gianna	60	165			
7	luigi	77	175			
8	francesco	83	182			
9	michele	79	178			
10	stefania	67	175			
11	claudia	68	173			
12						
13						
14						
15						
16						
17						
18						

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



[Home](#)

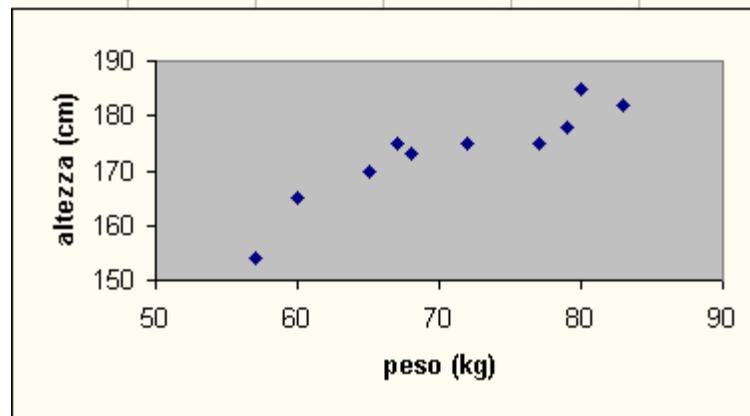


7.3 Retta di regressione

Nel caso di un insieme di dati bivariati, la cui rappresentazione sotto forma di grafico di dispersione suggerisce l'esistenza di una relazione lineare tra le variabili presentate, **sarà possibile approssimare una retta a questo insieme di punti**.

Tornando all'esempio dell'altezza e del peso degli individui e al relativo grafico di dispersione visto nei capp. 7 e 7.1, **si può osservare una relazione lineare tra peso e altezza**.

Sarà dunque possibile ricavare l'altezza di un individuo che pesa 70 Kg?



Di fronte ad una situazione analoga, in cui si è in presenza di un insieme di dati bivariati (x_i, y_i) , $i=1, \dots, n$, che seguono uno standard lineare, individuiamo una retta di equazione:

$$y = a + bx$$

che permette di descrivere come si riflettono in y – *variabile della risposta*, le variazioni processate in x – *variabile esplicativa*.

Tra i metodi più conosciuti per approssimare un insieme di dati con una retta vi è quello dei **minimi quadrati** che consiste nel determinare la retta che rende minima la somma dei quadrati degli scarti $[y_i - (a+bx)]$ tra i valori reali di y e quelli ottenuti dalla retta che si intende tracciare.

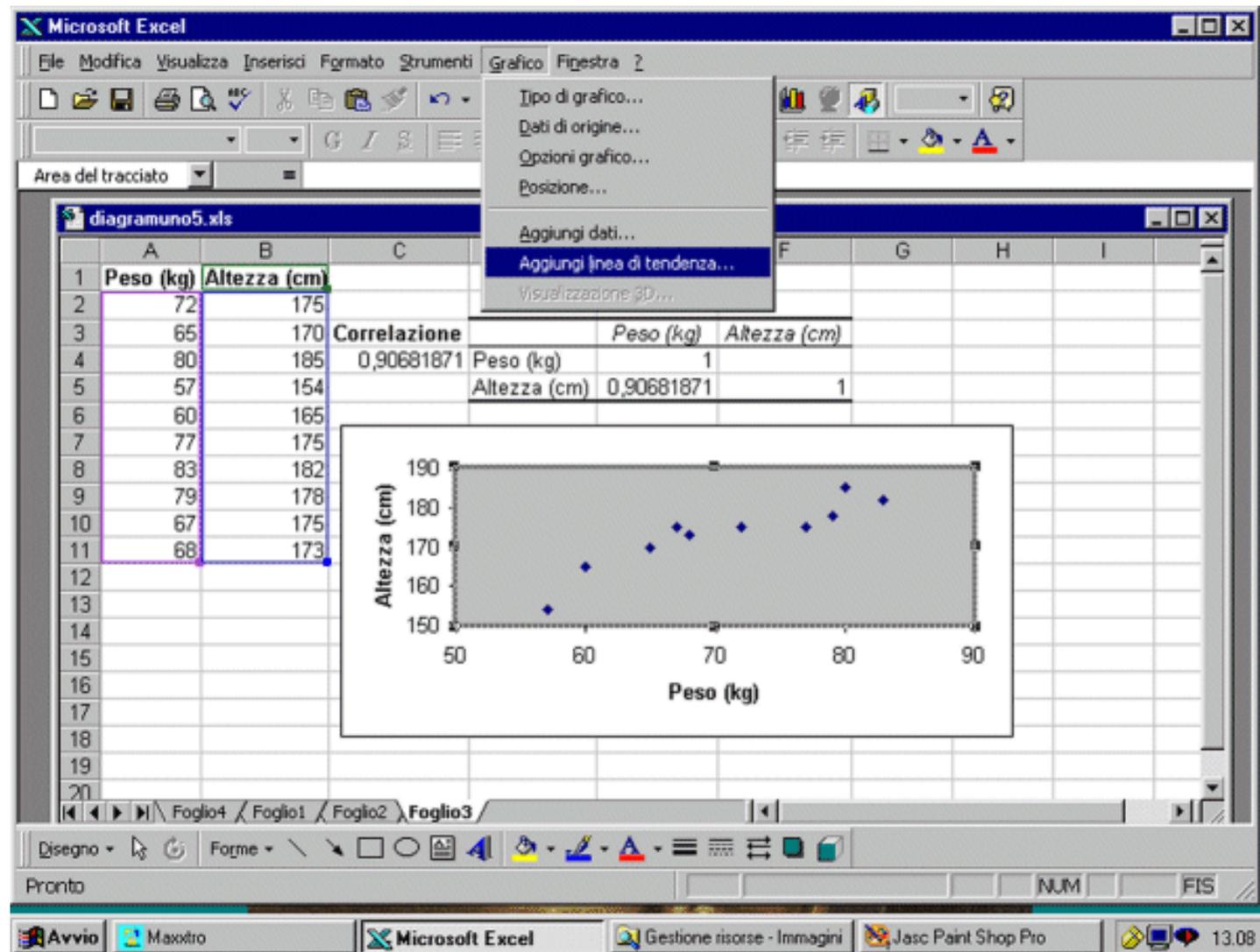
Questa retta prende il nome di retta di regressione o dei minimi quadrati. Per osservare una esemplificazione in Excel di quanto sopra riutilizziamo il grafico di dispersione visto nell'esercizio precedente →



[Torna](#) [Home](#)

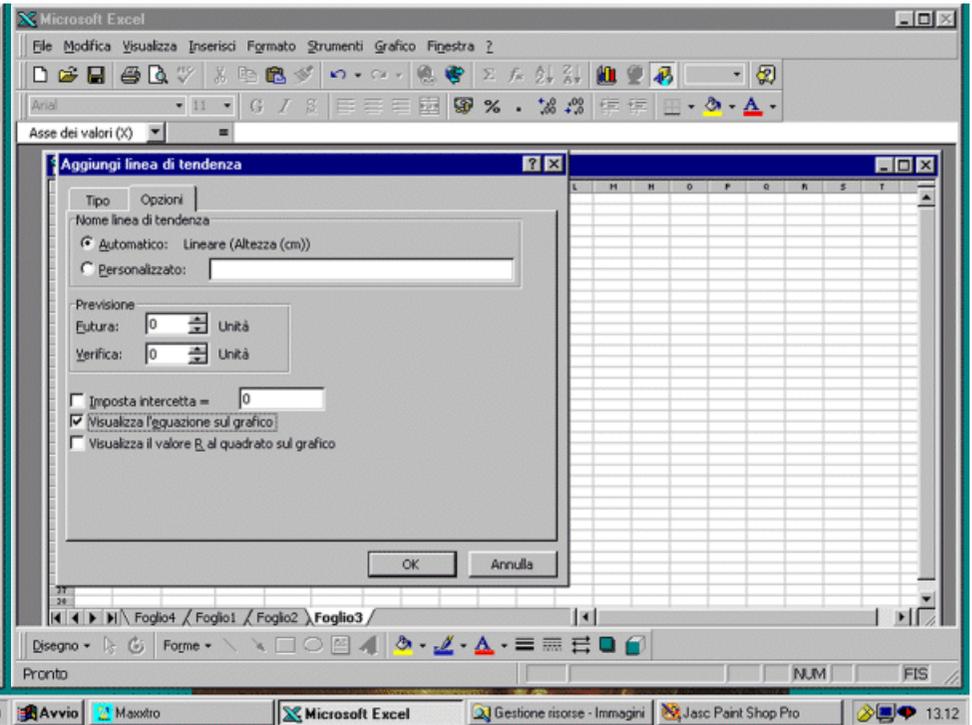
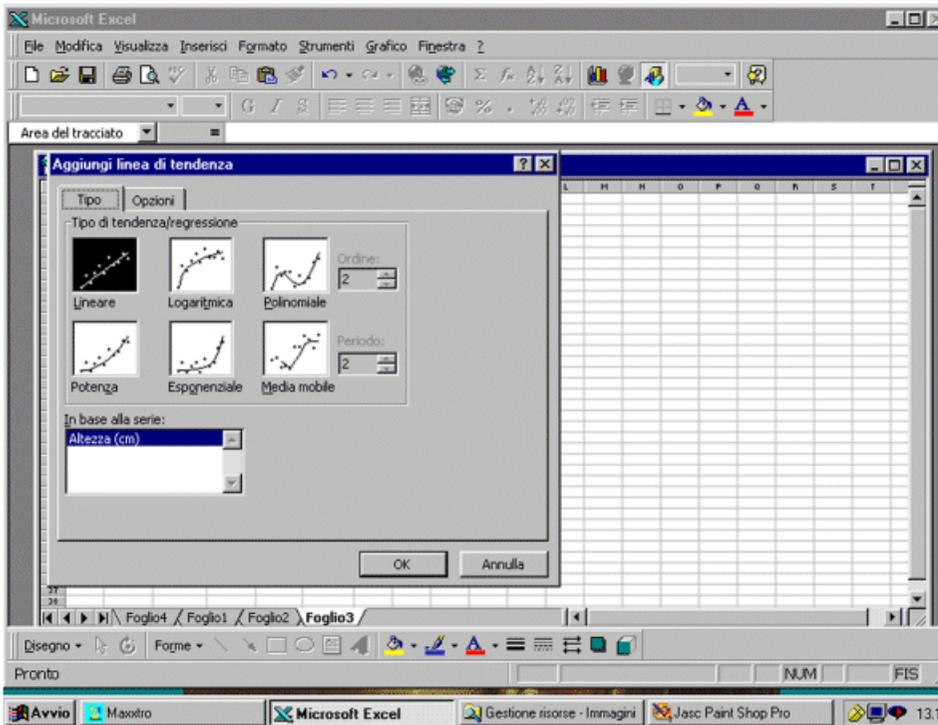


dopo aver **selezionato il grafico**, basterà semplicemente cliccare sul menù **Grafico** [Chart], selezionare il comando **Aggiungi linea di tendenza** [Add trendline] e continuare con le opzioni.



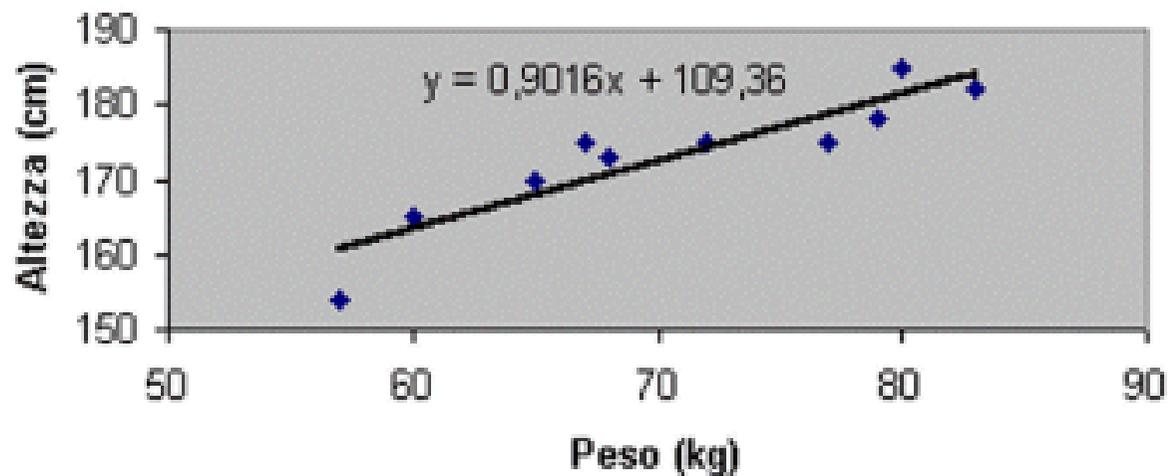
[Home](#)





L'equazione di questa retta sarà:

$$\text{Altezza} = 109,36 + 0,9016 \times \text{peso}$$



Sostituendo nell'equazione il Peso con 70, si otterrà il valore di 172,472 per cui l'altezza prevista per un alunno che pesa 70 kg, sarà di circa 172,5 cm.

[link foglio.xls](#) (clicca per ripetere l'esercitazione)



[Home](#)

Glossario

A B C D E F **G** H I **L** M N O P Q R S T U V Z

[Home](#)



A

Abitazione o Alloggio

L'insieme di vani, o anche un vano solo, destinato funzionalmente a uso di alloggio, che dispone di un ingresso indipendente su strada, pianerottolo, cortile, terrazza, ballatoio e simili, che alla data del censimento è occupato da una famiglia o da più famiglie coabitanti.

Abitazione occupata

L'abitazione in cui abitano una o più famiglie, i cui membri vi dimorano abitualmente, anche se temporaneamente assenti alla data del censimento.

Altre abitazioni

Abitazioni non occupate oppure abitate solamente da persone che non hanno dimora abituale nelle abitazioni stesse.

Altro tipo di alloggio

Alloggio non classificabile come abitazione presso il quale, al momento del censimento, dimorano abitualmente o temporaneamente una o più persone. Ne sono esempi: le roulotte, le tende, i caravan, i camper, i container; le baracche, le capanne, le casupole, le grotte; le rimesse, i garage, le soffitte, le cantine; gli alloggi contenuti in costruzioni che non sono edifici (secondo la definizione di edificio).

Ampiezza del campione

Valore riferito alla variabilità di un campione (vedi anche: campione), pari alla differenza tra l'estremo superiore e quello inferiore dei valori osservati del campione.

Ampiezza della classe

Differenza tra l'estremo superiore e quello inferiore della classe (detto anche valore).

Ampiezza interquartile

Indice di variabilità di un campione (vedi anche: campione), uguale alla differenza tra il terzo e il primo quartile, che fornisce informazioni circa l'ampiezza dell'intervallo in cui si trova il 50% delle osservazioni centrali.

[Torna](#)

[Home](#)

B

Box Plot o Diagramma degli estremi e dei quartili

Rappresentazione grafica in cui sono evidenziate alcune caratteristiche del campione (vedi anche: campione). L'insieme dei valori del campione compresi tra il 1° ed il 3° quartile, Q25 e Q75, è rappresentato da un rettangolo (scatola) la cui mediana (vedi anche: mediana) è indicata da una barra.

Due linee si allungano dai bordi della scatola fino ai punti estremi, i cosiddetti valori adiacenti.

[Torna](#) [Home](#)

C

Campione

Un campione è una riproduzione in scala ridotta dell'intera popolazione di riferimento.

Carattere (o variabile) dell'unità statistica

Ciascuna unità statistica presenta delle **caratteristiche**, che più brevemente si indicano con il nome di caratteri. Si consideri un gruppo di studenti universitari. Per ognuno di essi si può determinare l'età, la statura, il peso, il sesso, il colore dei capelli, il gruppo sanguigno, il numero dei fratelli, la nazionalità, la religione, il luogo di residenza, la facoltà cui è iscritto, l'anno di corso, eccetera. Si dice che l'età, la statura, il sesso, eccetera, sono tutti caratteri di ciascuna unità (studente) che si vuole osservare.

Caratteri qualitativi

Caratteri dell'unità statistica che identificano qualità o categorie non misurabili, ma soltanto classificabili secondo modalità diverse. Possono essere di natura ordinata (caratteri ordinati) o sconnessa (caratteri sconnessi).

Caratteri quantitativi

Caratteri dell'unità statistica che possono essere misurati o espressi mediante un numero e che possono essere di natura discreta (caratteri discreti) o continua (caratteri continui). Vedi dato statistico

Caratteri qualitativi ordinati

Caratteri dell'unità statistica che si possono ordinare naturalmente (es. titolo di studio: licenza elementare, di scuola secondaria, diploma ecc.).

Caratteri qualitativi sconnessi

Caratteri dell'unità statistica che non si possono ordinare naturalmente (es. luogo di residenza).

Caratteri quantitativi continui

Caratteri dell'unità statistica che possono assumere qualunque valore numerico compreso nell'intervallo di variazione (estensione di un territorio, densità di un gas); non è possibile elencare tutte le modalità che può assumere la variabile ma occorrerà limitarsi a contare quante unità manifestano la variabile con modalità compresa in un certo intervallo di valori reali.

Caratteri quantitativi discreti

Caratteri dell'unità statistica che possono assumere soltanto un numero finito intero entro l'intervallo di variazione (numero delle persone residenti in una città, numero dei vani di un'abitazione, numero delle stelle visibili in cielo in un dato periodo); è sempre possibile elencare tutte le modalità che può assumere la variabile.

[Torna](#)

[Home](#)

Censimento

Indagine diretta, individuale, totale e generalmente periodica sull'intera popolazione: è diretta e individuale perché le notizie non vengono desunte indirettamente ma dall'osservazione specifica delle singole unità; è totale perché tutte le unità debbono essere rilevate; è periodica perché viene ripetuta a intervalli regolari (generalmente decennali o quinquennali).

Cittadinanza

Vincolo di appartenenza a uno stato, richiesto e documentato per il godimento di diritti e l'assoggettamento a particolari oneri.

Coefficiente di correlazione lineare

Indice che misura il grado di relazione lineare (che varia da -1 a $+1$) tra due variabili.

Collettivo statistico

Vedi popolazione o collettivo statistico

Convivenza

Insieme di persone che, senza essere legate da vincoli di matrimonio, parentela, affinità e simili, conducono vita in comune per motivi religiosi, di cura, di assistenza, militari, di pena e simili. Le persone addette alla convivenza per ragioni di lavoro, se vi convivono abitualmente, sono considerate membri permanenti della convivenza purché non costituiscano famiglia a sé stante. I principali tipi di convivenza sono: istituti d'istruzione, istituti assistenziali, istituti di cura pubblici e privati, istituti penitenziari, convivenze ecclesiastiche, convivenze militari e di altri corpi accasermati, alberghi pensioni, locande e simili, navi mercantili, altre convivenze (ad esempio, case dello studente).

[Torna](#) [Home](#)

D

Dato statistico

Il dato statistico è il risultato dell'operazione di determinazione della modalità con cui un carattere è presente in ciascuna unità del collettivo.

Dati statistici bivariati

Coppia di caratteri corrispondenti ad un determinato individuo o risultato sperimentale.

Demografia

Il complesso di analisi descrittive e investigative volte a studiare le caratteristiche strutturali e dinamiche delle popolazioni umane, nei loro aspetti biologici e sociali e nelle loro interazioni.

Densità della popolazione

Concentrazione di individui espressa dal rapporto tra il numero di abitanti e la superficie del territorio (numero di abitanti per chilometro quadrato).

Devianza standard

Misura della variabilità dei dati (vedi anche: campione) rispetto a un valore medio, corrispondente alla radice quadrata della varianza (vedi anche: varianza), ed espressa nella stessa unità dei dati originali.

Diagramma a barre o distribuzione delle frequenze

Rappresentazione grafica, costituita da un sistema di assi cartesiani sulla cui ascissa viene indicato il valore centrale delle classi e su questi punti vengono disegnate delle colonne verticali la cui altezza rappresenta la frequenza assoluta o quella relativa.

Diagramma circolare (o *grafici circolari*, o *diagrammi a torta*)

Rappresentano le distribuzioni di frequenze relative o di percentuali utilizzando cerchi divisi in settori le cui aree sono proporzionali ai valori delle sequenze relative o delle percentuali. Se le categorie (o le classi) della distribuzione sono disposte secondo un ordine, di solito l'ordine si conserva procedendo in senso orario a partire dalle ore 12.

[Torna](#)

[Home](#)

Diagramma degli estremi e dei quartili (vedi anche Box plot)

Rappresentazione grafica in cui sono evidenziate alcune caratteristiche del campione. L'insieme dei valori campionari compresi tra il 1° e il 3°quartile (vedi anche: quartile), Q25 e Q75, è rappresentato da un rettangolo (scatola) la cui mediana (vedi anche: mediana) è indicata da una barra. Due linee uniscono i lati dei rettangoli rispettivamente con il valore massimo e minimo.

Diagramma di dispersione o a punti

Rappresentazione grafica di valori bivariati (vedi anche: dati bivariati), in cui ogni coppia di dati (x_i, y_i) è rappresentata da un punto di coordinate (x_i, y_i) in un sistema di assi cartesiani.

Dimensione del campione

Numero degli elementi che costituiscono il campione.

Distribuzione asimmetrica

Istogramma (vedi anche: mediana) che presenta una distribuzione delle frequenze in forma accentuatamente asimmetrica, con i valori sostanzialmente più piccoli su un lato, rispetto all'altro.

Distribuzione simmetrica

Istogramma (vedi anche: mediana) che presenta una distribuzione delle frequenze in forma approssimativamente simmetrica, rispetto a una classe media.

Distribuzione delle frequenze (diagramma a barre)

La distribuzione di frequenza è una organizzazione di dati in forma tabellare tale che ad ogni modalità di una certa variabile (qualitativa o quantitativa) si fa corrispondere la rispettiva frequenza (assoluta o relativa).

Quindi la distribuzione di frequenza esplicita quante volte (in assoluto se trattasi di frequenza assoluta, oppure in rapporto al totale se trattasi di frequenza relativa) una determinata modalità si presenta nel collettivo in esame.

(esempio di rappresentazione grafica: diagramma a barre)

[Torna](#) [Home](#)

E

Edificio

Si intende per edificio una costruzione di regola di concezione ed esecuzione unitaria; dotata di una propria indipendente struttura; contenente spazi utilizzabili stabilmente da persone per usi destinati all'abitazione c/o alla produzione di beni e/o di servizi, con le eventuali relative pertinenze; delimitata da pareti continue, esterne o divisorie, e da coperture; dotata di almeno un accesso dall'esterno.

Edificio per abitazione

Si intende un edificio progettato, costruito ed utilizzato (anche in seguito ad una variazione d'uso, od anche solo per un periodo) solo o principalmente a fini abitativi.

Eurostat - Ufficio statistico dell'Unione Europea

L'Ufficio statistico della Comunità europea costituito nel 1953, con sede a Lussemburgo.

Età media delle morti causate dal virus dell'HIV

Età in cui avviene il 50% delle morti totali causate da questa malattia.

[Torna](#)

[Home](#)

F

Fenomeni collettivi

Fenomeni che possono essere analizzati con metodi statistici e sono conoscibili soltanto attraverso una pluralità di osservazioni (es. consumo di un bene in un dato periodo, reddito in un settore in un dato periodo).

Flusso migratorio o saldo migratorio

Differenza tra immigrazioni ed emigrazioni di una popolazione, osservata in un dato periodo di tempo.

Frequenza

È il numero delle volte che una determinata modalità si verifica nel collettivo di riferimento (numero di elementi appartenenti a una determinata classe).

Frequenza assoluta

Numero di elementi appartenenti a una determinata classe quando la frequenza è un numero intero non negativo (esempio: Liguria - numero di famiglie con 2 componenti al censimento 2001 = 219.907).

Frequenza cumulata

Numero delle unità in cui il carattere è presente con intensità non superiore all'*i*-esima modalità.

Frequenza relativa

Frequenza assoluta rapportata al totale delle unità statistiche della popolazione; anche: rapporto tra il numero di elementi di una determinata classe e il totale degli elementi dell'insieme di dati da analizzare.

Esempio di frequenza relativa della modalità "2 componenti" sopra esposta = 0,3089 cioè il rapporto tra la frequenza assoluta (= 219.907) e il numero complessivo di famiglie della "classe" Liguria (= 711.947)

Funzione "distribuzione empirica" o cumulata

Funzione $F(x)$ definita per tutti i valori di x in R e che, per ogni x , dà la quantità degli elementi del campione minori o uguali ad x .

H

HIV

Sindrome da immunodeficienza acquisita.

Torna

Home

Indagine campionaria

Indagine in cui si osserva una parte della popolazione, che è detta campione, con lo scopo di ottenere da questa osservazione una descrizione della popolazione intera.

Indice di dipendenza totale

Rapporto esistente tra popolazione giovane e anziana da un lato, e popolazione attiva dall'altro, ogni 100 individui. Corrisponde alla somma degli indici di dipendenza dei giovani e degli anziani.

*Indice di Dipendenza Totale = (Indice Dipendenza Giovani + Indice Dipendenza anziani)/Popolazione di 15-64 anni) *100*

Indice dei prezzi al consumo per l'intera collettività

La variazione nel tempo dei prezzi, che si riferiscono alle vendite al dettaglio di beni e servizi effettuate dal settore delle imprese all'intero settore delle famiglie.

Indice dei prezzi al consumo per le famiglie di operai e impiegati

La variazione nel tempo dei prezzi al dettaglio, dei beni e servizi correntemente acquistati dalle famiglie di lavoratori dipendenti non agricoli (operai e impiegati).

Indice armonizzato dei prezzi al consumo nei paesi dell'Unione europea

È riferito alla stessa popolazione dell'indice dei prezzi al consumo per l'intera collettività ma è calcolato rispetto a un paniere di prodotti e a una struttura di ponderazione armonizzati nei diversi paesi per poter assicurare la migliore comparabilità internazionale.

Indice di posizione

Valore che tende a cadere centralmente all'interno di un insieme di dati ordinati.

Inflazione

Aumento generale e sostenuto dei prezzi. Il concetto di inflazione generalmente viene associato all'aumento dell'indice dei prezzi al consumo.

Istogramma o diagramma di area

Rappresentazione grafica dei caratteri quantitativi continui (vedi anche: caratteri quantitativi continui), costituita da una successione di rettangoli adiacenti, ognuno dei quali ha, come base, l'ampiezza di una classe e, come area, la frequenza relativa (vedi anche: frequenza relativa) o la frequenza assoluta (vedi anche: frequenza assoluta). Pertanto, l'area totale coperta dall'istogramma è uguale a 1 o a n (dimensione del campione).

[Torna](#) [Home](#)

M

Media

Indice di posizione del campione più utilizzato, che si ottiene sommando tutti gli elementi del campione e dividendo il risultato per la dimensione dello stesso.

Mediana

Indice di posizione di una distribuzione di dati disposti in ordine crescente e corrispondente al valore che divide il campione in due parti, per cui, una metà degli elementi dell'insieme di dati sarà minore o uguale alla mediana, mentre la restante sarà superiore o uguale.

Misura di dispersione o di variabilità

Indice che dà una misura della variabilità dei valori della distribuzione rispetto a una media o di quanto i valori stessi differiscono tra di loro.

Moda

Indice che si presenta con la massima frequenza in un insieme di dati, nel caso di caratteri quantitativi discreti, o in una classe modale, nel caso di caratteri quantitativi continui.

Modalità

I diversi stati o proprietà che i caratteri o variabili possono assumere (es. per lo stato civile: nubile, coniugato, divorziato).

[Torna](#)

[Home](#)

N

Nato vivo

Il prodotto del concepimento che, una volta espulso o completamente estratto dal corpo materno, indipendentemente dalla durata della gestazione, respiri o manifesti altro segno di vita.

Numero medio di componenti per famiglia

È calcolato dividendo il totale dei residenti in famiglia per il numero delle famiglie.

NUTS

Nomenclatura delle unità territoriali a fini statistici, introdotta da Eurostat per fornire una suddivisione unificata e a più livelli del territorio dell'Unione europea per la produzione e la diffusione di statistiche.

[Torna](#) [Home](#)

O

Occupato

Persona di 15 anni e più che all'indagine sulle forze di lavoro dichiara:

- 1) di possedere un'occupazione, anche se nel periodo di riferimento non ha svolto attività lavorativa (occupato dichiarato);
- 2) di essere in una condizione diversa da occupato, ma di aver effettuato ore di lavoro nel periodo di riferimento (altra persona con attività lavorativa).

Outlier

Vedere: valori anomali

[Torna](#)

[Home](#)

P

Parametro

Valore non noto che descrive la popolazione; si calcola mediante lo stimatore (vedi anche: stimatore).

Persona in cerca di occupazione

La persona di 15 anni e più che all'indagine sulle forze di lavoro dichiara:

- 1) una condizione professionale diversa da quella di occupato;
- 2) di non aver effettuato ore di lavoro nel periodo di riferimento;
- 3) di essere alla ricerca di un lavoro;
- 4) di aver effettuato almeno un'azione di ricerca di lavoro nei trenta giorni che precedono il periodo di riferimento;
- 5) di essere immediatamente disponibile (entro due settimane) ad accettare un lavoro, qualora gli venga offerto.

Poligono di frequenza

Linea che unisce gli estremi delle barre di un diagramma a barre (vedi anche: diagramma a barre).

Popolazione o collettivo statistico

L'insieme delle unità statistiche (persone, fenomeni, oggetti) oggetto dell'indagine, aventi una o più caratteristiche in comune.

[Torna](#) [Home](#)

Popolazione attiva

La somma delle persone occupate, di quelle disoccupate alla ricerca di nuova occupazione e delle persone in cerca di prima occupazione.

Popolazione non attiva

La popolazione in condizione non professionale meno le persone in cerca di prima occupazione.

Popolazione presente

Per ciascun comune è costituita dalle persone presenti nel comune alla data del censimento ed aventi in esso dimora abituale nonché dalle persone presenti nel comune alla data del censimento ma aventi dimora abituale in altro comune o all'estero.

Popolazione residente

Per ciascun comune è costituita dalle persone aventi dimora abituale nel comune, anche se alla data del censimento sono assenti perché temporaneamente presenti in altro comune italiano o all'estero.

Prodotto Interno Lordo ai prezzi di mercato (PIL pm)

Il risultato finale dell'attività di produzione delle unità produttrici residenti. Corrisponde alla produzione totale di beni e servizi dell'economia, diminuita dei consumi intermedi e aumentata dell'Iva gravante e delle imposte indirette sulle importazioni. È altresì pari alla somma dei valori aggiunti ai prezzi di mercato delle varie branche di attività economica, aumentata dell'Iva e delle imposte indirette sulle importazioni, al netto dei servizi di intermediazione finanziaria indirettamente misurati (SIFIM). (*Sistema europeo dei conti, SEC 95*).

[Torna](#)

[Home](#)

Q

Quantili - Quartili

Quantili

sono le n parti uguali in cui si suddivide l'intera distribuzione di frequenza, dopo aver ordinato i valori.

Quartili, quintili, ecc.

quando la distribuzione si suddivide in 4 parti uguali ($q = 4$) si parla di quartili (ognuno ha il 25% della numerosità (o della quantità) del totale)

- in 5 parti uguali ($q = 5$) si parla di quintili
- in 6 parti uguali ($q = 6$) si parla di sestili
- in 10 parti uguali ($q = 10$) si parla di decili
- in 100 parti uguali ($q = 100$) si parla di percentili.

[Torna](#)

[Home](#)

R

Retta di regressione o dei minimi quadrati

La retta che rende minima la somma dei quadrati degli scarti $[y_i - (a+bx)]$ tra i valori reali di y e quelli ottenuti dalla retta che si intende tracciare.

Ripartizioni geografiche

Nei testi e nelle tavole statistiche che illustrano i risultati dei Censimento dell'industria e servizi si è adottata la seguente ripartizione geografica.

- Italia Nord-occidentale: comprende Valle d'Aosta, Piemonte, Lombardia e Liguria;
 - Italia Nord-orientale: comprende Friuli-Venezia Giulia, Provincia autonoma di Trento, Provincia autonoma di Bolzano, Veneto, Emilia-Romagna;
 - Italia Centrale: comprende Toscana, Marche, Umbria, Lazio;
 - Italia Meridionale: comprende Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria;
 - Italia Insulare: comprende Sicilia, Sardegna.
- Il Mezzogiorno comprende le regioni dell'Italia meridionale e dell'Italia insulare.

[Torna](#)

[Home](#)

S

Speranza di vita alla nascita

Numero medio di anni che sono da vivere per un neonato.

Statistica (1)

Scienza che si occupa di raccogliere, classificare, analizzare e interpretare dati attraverso l'uso di metodi scientifici.

Statistica (2)

Valore che descrive il campione (vedi anche: campione). È una funzione dei valori osservati sul campione e viene utilizzata per valutare un parametro sconosciuto della popolazione osservata.

Statistica descrittiva

Ramo della statistica che si occupa di evidenziare, attraverso sintesi numeriche o grafiche, le caratteristiche dei fenomeni collettivi di interesse. L'obiettivo è quello di ricavare dalle osservazioni spesso numerose e in forma disordinata, informazioni di sintesi.

Statistica induttiva o inferenziale

Ramo della statistica che si occupa di interpretare le osservazioni in termini di modelli teorici che spiegano il meccanismo secondo il quale si producono i dati esaminati. Questi modelli permettono di generalizzare in modo induttivo i risultati, ottenuti da campionamento o da osservazioni indirette, dall'insieme dei dati osservati, alla popolazione di riferimento.

[Torna](#) [Home](#)

Stima

Risultato dello stimatore (vedi anche: stimatore) che si basa su un campione (vedi anche: campione).

Stimatore

Statistica campionaria (variabile aleatoria – vedi anche: variabile aleatoria) i cui valori specifici costituiscono le stime dei parametri in oggetto. Vedi Statistica (2).

Straniero residente

Cittadino straniero (residente in Italia) che ha dimora abituale nell'alloggio o nella convivenza ed è in possesso dei requisiti per l'iscrizione in anagrafe.

Straniero non residente

Cittadino straniero che non ha dimora abituale nell'alloggio o nella convivenza.

[Torna](#)

[Home](#)

T

Tabella di contingenza

Tabella che rappresenta una distribuzione di dati bivariati di tipo sia qualitativo, sia quantitativo. L'aspetto è quello di una tabella suddivisa per righe e colonne secondo le modalità delle variabili.

Tabella di frequenza

Tabella che rappresenta la distribuzione della variabile, cioè, quali valori o modalità la variabile assume e con che frequenza assume questi valori o modalità.

Tasso di disoccupazione

Rapporto tra le persone in cerca di occupazione e le forze di lavoro.

Tasso di mortalità

Rapporto tra il numero delle morti nell'anno e l'ammontare della popolazione residente (per 1000).

Tasso di mortalità infantile

Il rapporto tra il numero di morti avvenute nel primo anno di vita e il numero di nati vivi (per 1000).

Tasso di mortalità per HIV

Rapporto tra il numero di morti per HIV avvenuti in un determinato anno e la popolazione media dello stesso anno, ogni 100.000 individui.

Tasso di natalità

Rapporto tra il numero di nati vivi dell'anno e l'ammontare medio della popolazione residente (per 1000).

Tasso di scolarità e di iscrizione

Rapporto tra gli studenti iscritti al livello di istruzione considerato e la popolazione residente appartenente alla corrispondente classe teorica di età (per 100). Per la scuola secondaria superiore l'età teorica considerata è 14-18 anni, per l'università 19-25.

[Torna](#) [Home](#)

U

Unità statistica

L'oggetto dell'osservazione di ogni fenomeno individuale che costituisce il fenomeno collettivo è detto unità statistica. Per esempio, è un'unità: ciascun individuo di una popolazione, ciascun nato, ciascun morto, ogni coppia di sposi, ciascun emigrato, eccetera, a seconda del fenomeno oggetto di indagine.

[Torna](#)

[Home](#)

V

Valori anomali

Valori che si distinguono dagli altri, come se non appartenessero allo stesso gruppo di dati.

Variabile o carattere

Gli aspetti delle unità statistiche (stato civile, titolo di studio) che si vogliono osservare.

Variabile aleatoria

Variabile che può assumere un qualunque valore di un insieme e a ciascuno dei suoi valori è associata una probabilità.

Varianza

Misura di variabilità che si ottiene sommando i quadrati degli scarti dei valori osservati dalla loro media e dividendo il valore della somma per il numero di osservazioni.

Varianza campionaria

Misura di variabilità che si ottiene sommando i quadrati degli scarti dei valori osservati dalla loro media e dividendo il valore della somma per il numero di osservazioni meno una.

Variazione intercensuaria

La variazione fra il dato di due censimenti consecutivi. La variazione può essere assoluta, ed è calcolata come la differenza fra il dato del censimento al tempo t meno il dato del censimento al tempo $t-1$; può essere relativa percentuale ed in questo caso la differenza assoluta è rapportata al dato del censimento al tempo $t-1$ moltiplicata per 100.

[Torna](#) [Home](#)

Z

Zona altimetrica

Ripartizione del territorio nazionale in zone omogenee derivanti dall'aggregazione comuni contigui costruite di norma sulla base di valori soglia altimetrici. Si distinguono zone altimetriche di montagna, di collina e di pianura.

Zona altimetrica di collina

Il territorio caratterizzato dalla presenza di diffuse masse rilevate aventi altitudini, di regola, inferiori a 600 metri nell'Italia settentrionale e 700 metri nell'Italia centro-meridionale ed insulare. Eventuali aree di limitata estensione aventi differenti caratteristiche, intercluse, si considerano comprese nella zona di collina.

Zona altimetrica di montagna

Il territorio caratterizzato dalla presenza di notevoli masse rilevate aventi altitudini, di norma, non inferiori a 600 metri nell'Italia settentrionale e 700 metri nell'Italia centro-meridionale e insulare. Gli anzidetti livelli altitudinali sono suscettibili di spostamento in relazione ai limiti inferiori delle zone fitogeografiche dell'Alpinetum, del Picetum e del Fagetum, nonché in relazione ai limiti superiori delle aree di coltura in massa della vite nell'Italia settentrionale e dell'olivo nella Italia centro-meridionale e insulare. Le aree intercluse fra le masse rilevate, costituite da valli, altipiani ed analoghe configurazioni del suolo, s'intendono comprese nella zona di montagna.

Zona altimetrica di pianura

Il territorio basso e pianeggiante caratterizzato dall'assenza di masse rilevate. Si considerano nella zona di pianura anche le propaggini di territorio che nei punti più discosti dal mare si elevano ad altitudine, di regola, non superiore ai 300 metri, purché presentino nell'insieme e senza soluzione di continuità, inclinazione trascurabile rispetto al corpo della zona di pianura. Si escludono dalla pianura i fondovalle aperti ad essa oltre l'apice delle conoidi fluviali ancorché appiattite e si escludono, altresì, le strisce litoranee pianeggianti di modesta estensione. Eventuali rilievi montagnosi o collinari, interclusi nella superficie pianeggiante e di estensione trascurabile, si considerano compresi nella zona di pianura.

[Torna](#)

[Home](#)

Bibliografia e note

A. GAMBINI, 2007 *“Argomenti di statistica descrittiva”* (seconda edizione), G. Giappichelli Editore -Torino

A. GAMBINI, 2007 *“Argomenti di probabilità ad uso degli statistici”*, ed. ECIG - Genova

G. LETI, 1983 *“Statistica descrittiva”*, il Mulino

D. PICCOLO, 1998 *“Statistica”*, il Mulino

Le pagine del Centro Interuniversitario di Ricerca per la Didattica delle Discipline Statistiche <http://cirdis.stat.unipg.it/index.php?lang=ita> e <http://cirdis.stat.unipg.it/index.php?canale=89&lang=ita> propongono alcune interessantissime risorse didattiche utili all'insegnamento della statistica nelle scuole medie e superiori.

La pagina DISCUSS (Discovering Important Statistical Concepts Using Spreadsheets), dell'Università di Coventry (http://www.mis.coventry.ac.uk/research/discus/discus_home.html) presenta una serie di fogli di calcolo interattivi, che lavorano sotto Excel, ideati come strumento di sostegno all'insegnamento della statistica nelle scuole secondarie e/o nel corso propedeutico Universitario.

E, ancora, si segnala <http://www.matematicamente.it/statistica/index.htm> (interessante, tra l'altro, la sezione su: “definizioni di probabilità e simulazioni con Excel”).

Vi sono poi diversi programmi che supportano Excel, nel campo della statistica. Di seguito elencheremo a scopo esemplificativo i nomi di alcuni siti dove è possibile il download di versioni demo:

- 1) WinSTAT Statistics Add-In for Microsoft Excel <http://www.winstat.com/>
- 2) XLSTAT in Excel Add-in for statistics and data analysis + several Excel utilities + XLControl <http://www.xlstat.com/>
- 3) XLStatistics Home Page - Excel, Statistics, Data Analysis <http://www.man.deakin.edu.au/rodneyc/XLStats.htm>

[Torna](#) [Home](#)



Crediti:

Introduzione alla Statistica con Excel
è stato realizzato da:

Silvio Vitale
ISTAT – Istituto Nazionale di Statistica
Ufficio Regionale per la Liguria

Via San Vincenzo, 4
16121 Genova
Tel. 010 5849701
Email: urge@istat.it

Questo prodotto è uno sviluppo del manuale *L'uso di excel per la statistica* (progettazione e redazione: Danila Fulgenzio; *revisione testi*: Paola Giacché, Francesca Loporcaro. ISTAT - Direzione centrale per la diffusione della cultura e dell'informazione statistica - binariodieci www.istat.it).

Il manuale tiene conto dell'esperienza dell'Istituto di statistica del Portogallo (www.ine.pt).