

# Ulteriori Conoscenze di Informatica e Statistica

Carlo Meneghini

Dip. di fisica - via della Vasca Navale 84,  
st. 83 (I piano) tel.: 06 55 17 72 17

[meneghini@fis.uniroma3.it](mailto:meneghini@fis.uniroma3.it)

# Comportamento in laboratorio

## *Durante le esercitazioni in laboratorio evitare di:*

- utilizzare le macchine del laboratorio di calcolo per navigare in internet consultando pagine e siti non pertinenti al corso o alle esercitazioni
- scaricare, installare e utilizzare programmi non inerenti al corso o alle esercitazioni
- giocare al computer
- cambiare le impostazioni del sistema (es. home page, password di accesso)

Non utilizzare il **desktop** per salvare il proprio lavoro ma **creare la propria directory** e salvare i lavori in opportune sottodirectories.

*Alla fine della lezione effettuare correttamente lo shutdown del sistema e spegnere il PC*

# Note

I PC a disposizione nel centro di calcolo del Dip. di Fisica sono macchine *dual-boot*, possono cioè utilizzare piú di un sistema operativo. Nel nostro caso W2K e Linux. Dopo il caricamento del BIOS il computer aspetta per alcuni secondi in attesa che voi scegliate il sistema da usare, **Scegliete windows.**

Dopo il caricamento del sistema un finestra vi chiederà il nome utente e la password. Digitando il vostro nome utente e password entrate nel sul desktop di Windows 2000.

# Piano della lezione

## Statistica descrittiva:

- Strumenti software: **il foglio elettronico\***
  - Le variabili aleatorie: caratteristiche
  - distribuzioni di probabilità

**\*Nota:** programmi freeware (come OpenOffice della Sun) sono del tutto equivalenti e quasi totalmente compatibili con il pacchetto Excel della Microsoft. La versione integrale del pacchetto OpenOffice è gratuita per studenti e istituzioni accademiche.

# Statistica descrittiva

## ORDINATI

Possono essere ordinati naturalmente (ad esempio si possono ordinare gradualmente il carattere "Titolo di studio": licenza elementare, di scuola secondaria, diploma, ecc...)

## 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

## Unità statistica:

oggetto dell'osservazione  
individuale che costituisce il  
fenomeno collettivo in esame

## 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

# La rappresentazione dei dati statistici deve essere organizzata in modo da:

- semplificare i confronti
- sintetizzare i risultati

Esp. 1	
oss.	tipo
1	M
2	F
3	F
4	F
5	M
6	M
7	F
8	M
9	M
10	F
11	F
12	F
13	M
14	M
15	F

Esp. 2	
oss.	tipo
1	M
2	M
3	F
4	F
5	F
6	M
7	F
8	M
9	M
10	F
11	F
12	M
13	M
14	M
15	F
16	F
17	M
18	M
19	M
20	F
21	M
22	F
23	F
24	M
25	M
26	F

Esp. 1		Esp. 2	
M	F	M	F
7	8	14	12

Freq.  
Assoluta

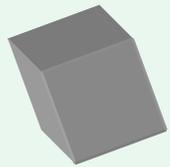
Freq.  
Relativa

Esp. 1		Esp. 2	
M %	F %	M	F
46.7	53.3	53.8	46.2

# Fenomeni

deterministici:

se ripetuti nelle medesime condizioni producono gli stessi risultati

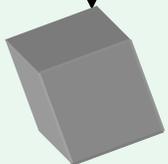


Caduta dei gravi

$$F = mg$$

$$v = mgt$$

$$x = \frac{1}{2} gt^2$$



$x, v =$  variabili deterministiche

aleatori:

pur ripetuti nelle medesime condizioni possono produrre risultati differenti

Lancio di dadi



Quale numero?

$N =$  variabile aleatoria

# Testa o croce?

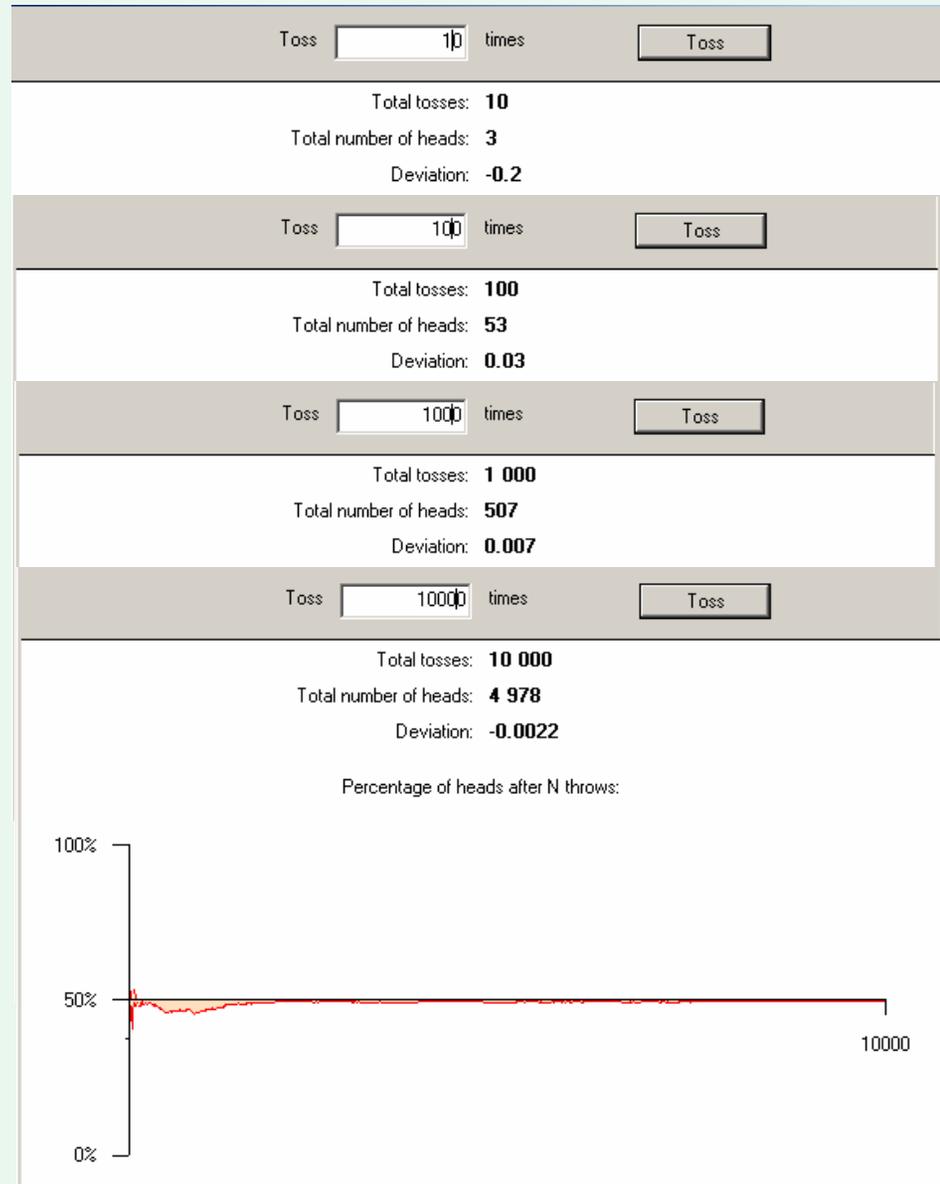


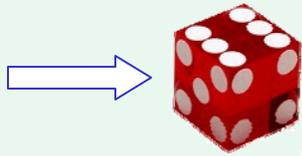
Lancia la moneta  
([moneta.exe](#))

$$p_T = \lim_{N \rightarrow \infty} \frac{n_T}{N}$$

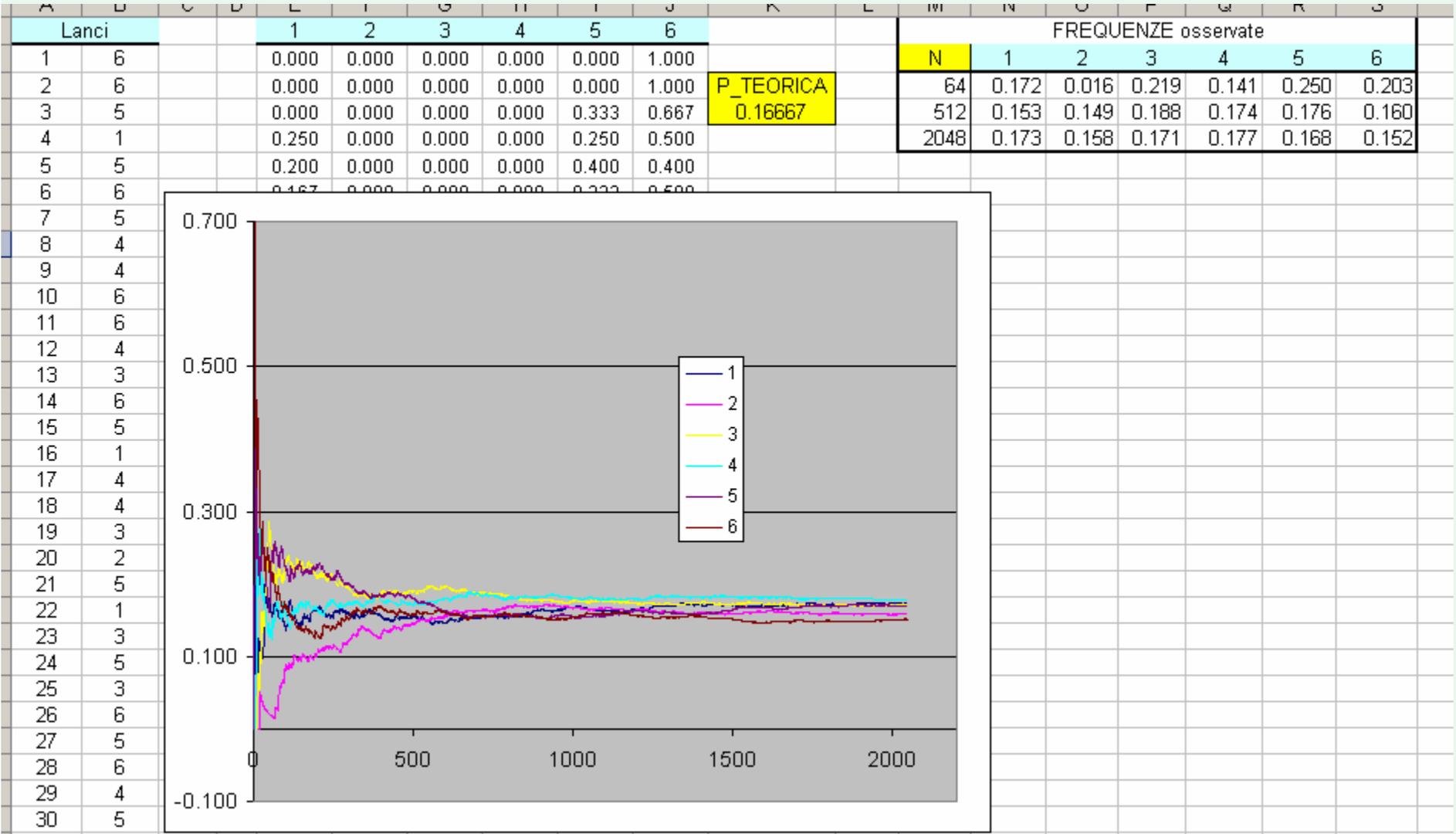
$$p_+ = \lim_{N \rightarrow \infty} \frac{n_T}{N}$$

$$p_T = 0.5 = p_+$$





# Lancio di un dado



# Frequenza: Variabili discrete

$X$  = variabili aleatoria,  $N$  = numero di osservazioni

$x_1, x_2, \dots, x_i$  : valori assunti dalla variabile  $X$

$n_1, n_2, \dots, n_i$  : numero di volte che si osserva il valore  $i$ -esimo  $x_i$

$n_i$ : frequenza assoluta della variabile  $x_i$  con:

$$\sum_{i=1}^V n_i = n$$

$f_i = \frac{n_i}{N}$  frequenza relativa della variabile  $x_i$

$$\sum_{i=1}^V f_i = 1$$

$$f_i \geq 0$$

# Frequenza: Variabili continue

$X$  = variabili aleatoria

$x$ : valori assunti dalla variabile  $X$ ,

$f(x)$ : densità di frequenza della variabile aleatoria  $X$

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

$f(x_i)$ : frequenza relativa nell'intervallo  $x_i < X < x_{i+1}$

$$f(x_i) = \int_{x_i}^{x_{i+1}} f(x) dx$$

$$f(x_i) = \frac{\text{numero di osservazioni tra } x_i \text{ e } x_{i+1}}{N}$$

La statistica descrittiva sintetizza l'informazione contenuta nell'insieme dei valori assunti da una variabile aleatoria (distribuzione) utilizzando:

- indici di posizione
- indici di dispersione (variabilità)
  - indici di forma
- istogrammi di frequenza
  - box plot

# Indici di "posizione" (indici di tendenza)

indice	definizione	funzione EXCEL
Media	$\frac{1}{N} \sum_{i=1}^N x_i = \bar{x} = \langle x \rangle$ $\frac{\int_A^B f(x) dx}{\int_A^B dx} = \bar{x} = \langle x \rangle$	MEDIA( <i>dati</i> )
Moda	Valore della variabile cui corrisponde la massima frequenza	MODA( <i>dati</i> )
Mediana	Valore della variabile che permette di dividere la distribuzione delle osservazioni in due parti uguali	MEDIANA( <i>dati</i> )
Quartili		QUARTILE( <i>dati</i> ; <i>q</i> )

# Indici di "dispersione"

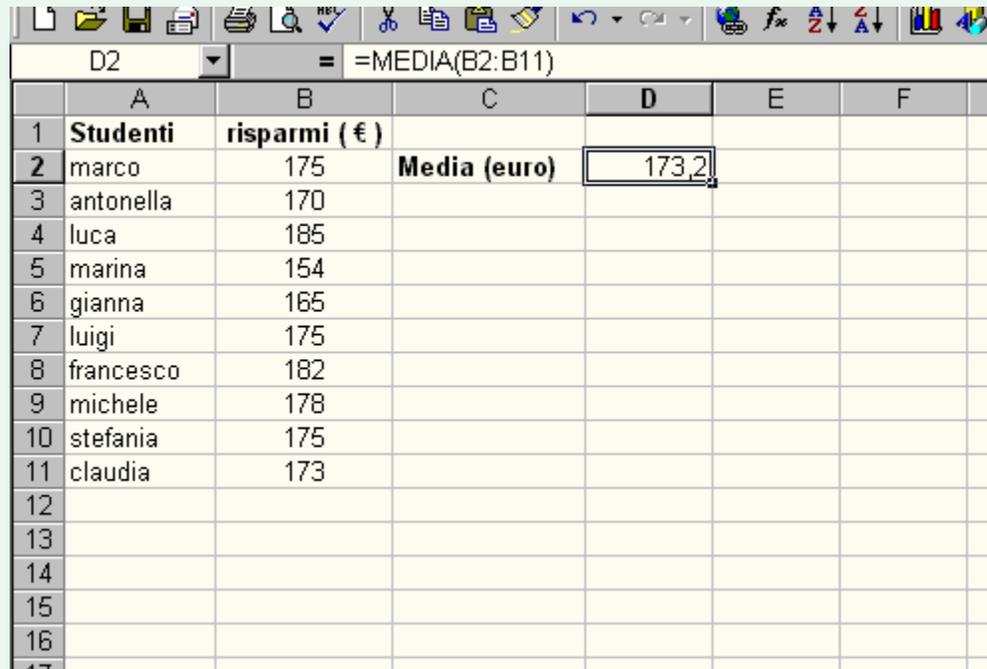
indice	definizione	funzione EXCEL
Varianza	$\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1} = \sigma^2$	<code>VAR(dati)</code>
Deviazione Standard	$\sqrt{\sigma^2} = \sigma$	<code>DEV.ST(dati)</code>
Interquartile	Q3-Q1	

# Media

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	<b>Studenti</b>	<b>risparmi (€)</b>				
2	marco	175	<b>Media (euro)</b>	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						

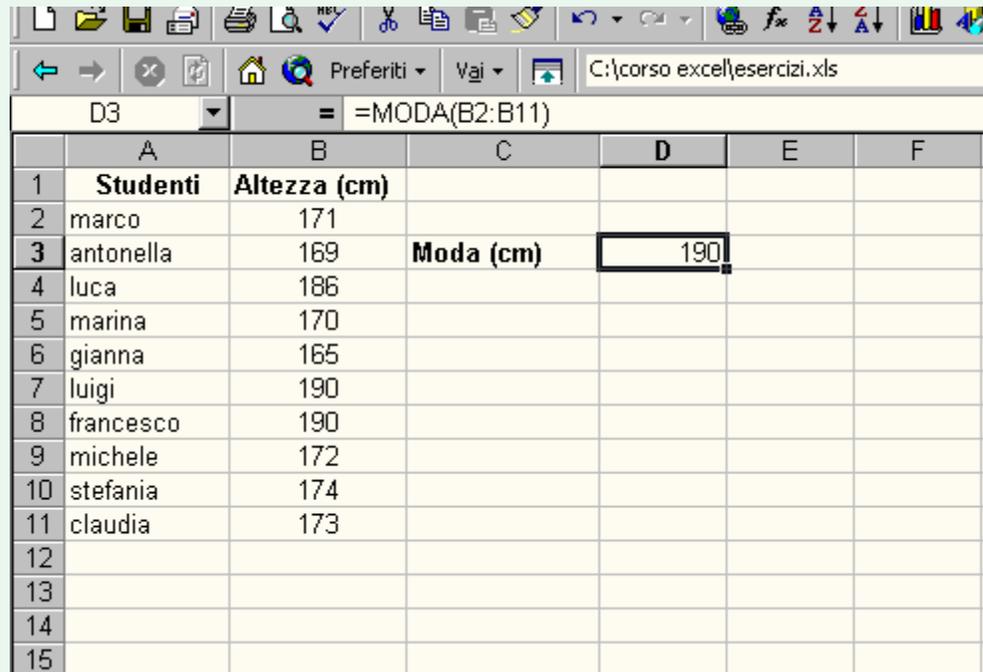
Come si fa:

**=MEDIA(B2:B11).**

# 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	<b>Studenti</b>	<b>Altezza (cm)</b>				
2	marco	171				
3	antonella	169	<b>Moda (cm)</b>	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						

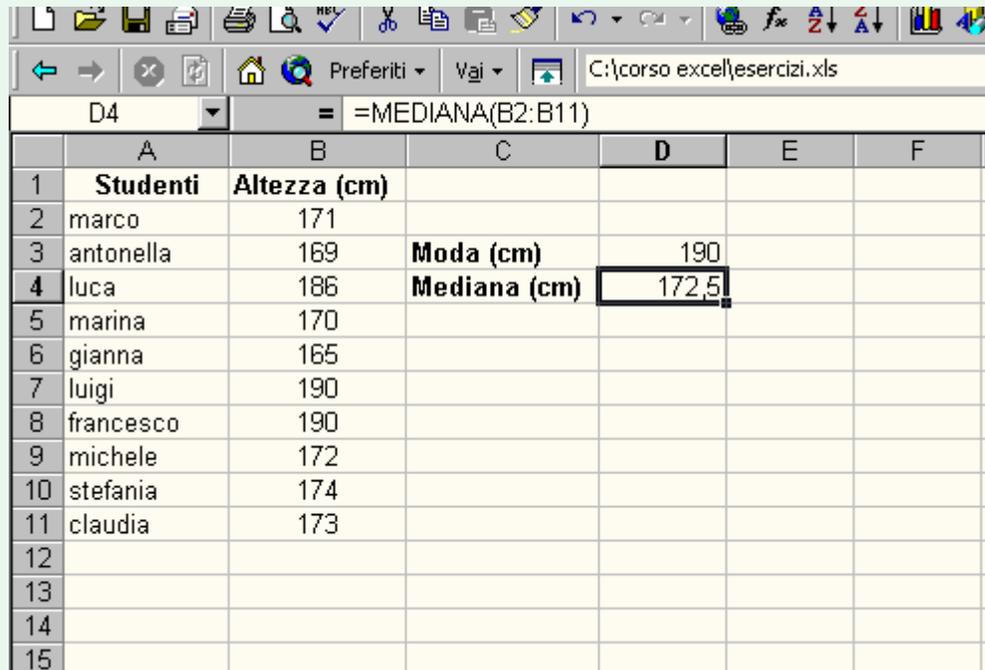
Come si fa:

**=MODA (B2:B11)**

# 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	<b>Studenti</b>	<b>Altezza (cm)</b>				
2	marco	171				
3	antonella	169	<b>Moda (cm)</b>	190		
4	luca	186	<b>Mediana (cm)</b>	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						

Come si fa:

**=MEDIANA (B2:B11).**

# Quartili

Si può dividere la distribuzione parti (percentili) contenenti ognuna la  $q$ -esima parte della quantità della distribuzione totale.

**I quartili sono le  $n$  parti in cui è stata suddivisa una distribuzione.**

per  $q = 4$  (più usati) si parla di *quartili*

**I quartili dividono la distribuzione in quattro parti aventi ognuna il  $1/4$  (25%) della quantità totale;**

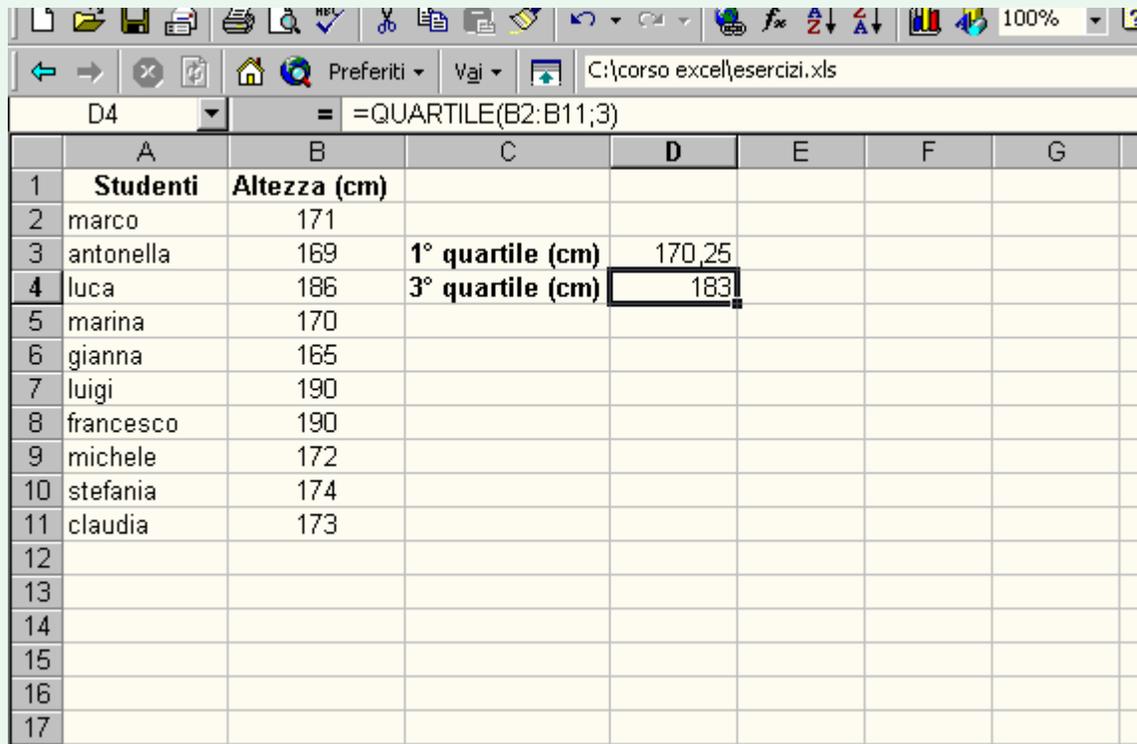
**Il I quartile (Q1)** è il limite superiore della distribuzione che ha il 25% della quantità totale;

**Il II quartile (Q2)** è il limite superiore della seconda distribuzione e quindi da solo separa nella distribuzione totale due distribuzioni che hanno ciascuna il 50% della quantità totale, il Q2 coincide con la mediana;

**Il III quartile (Q3)** è il limite superiore della distribuzione che ha il 75% dell'ammontare della distribuzione totale.

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	<b>Studenti</b>	<b>Altezza (cm)</b>					
2	marco	171					
3	antonella	169	<b>1° quartile (cm)</b>	170,25			
4	luca	186	<b>3° quartile (cm)</b>	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)`. The value 183 is displayed in cell D4.

Come si fa:

**Q1:**

**=QUARTILE (B2:B11,1).**

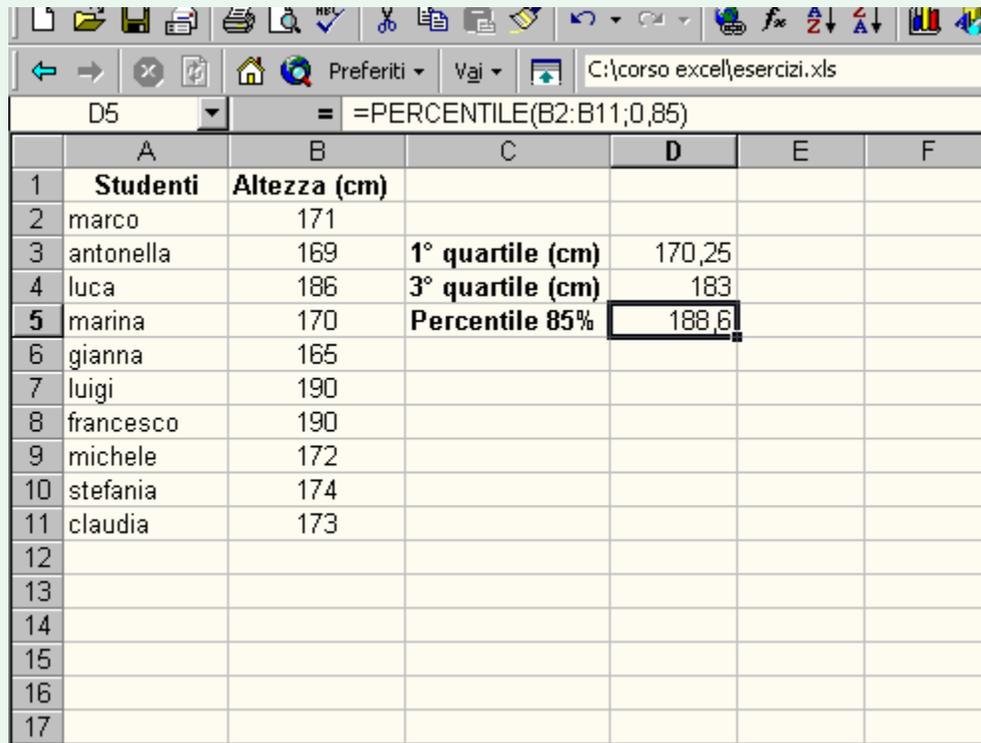
**Q2:**

**=QUARTILE (B2:B11,2).**

**Q3:**

**=QUARTILE (B2:B11,3).**

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



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	<b>Studenti</b>	<b>Altezza (cm)</b>				
2	marco	171				
3	antonella	169	<b>1° quartile (cm)</b>	170,25		
4	luca	186	<b>3° quartile (cm)</b>	183		
5	marina	170	<b>Percentile 85%</b>	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						

The formula bar shows the formula: `=PERCENTILE(B2:B11;0,85)`

**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)**.

# Indicatori di variabilità (dispersione)

Misurano la dispersione dei valori di una distribuzione

Varianza

Deviazione standard

Ampiezza

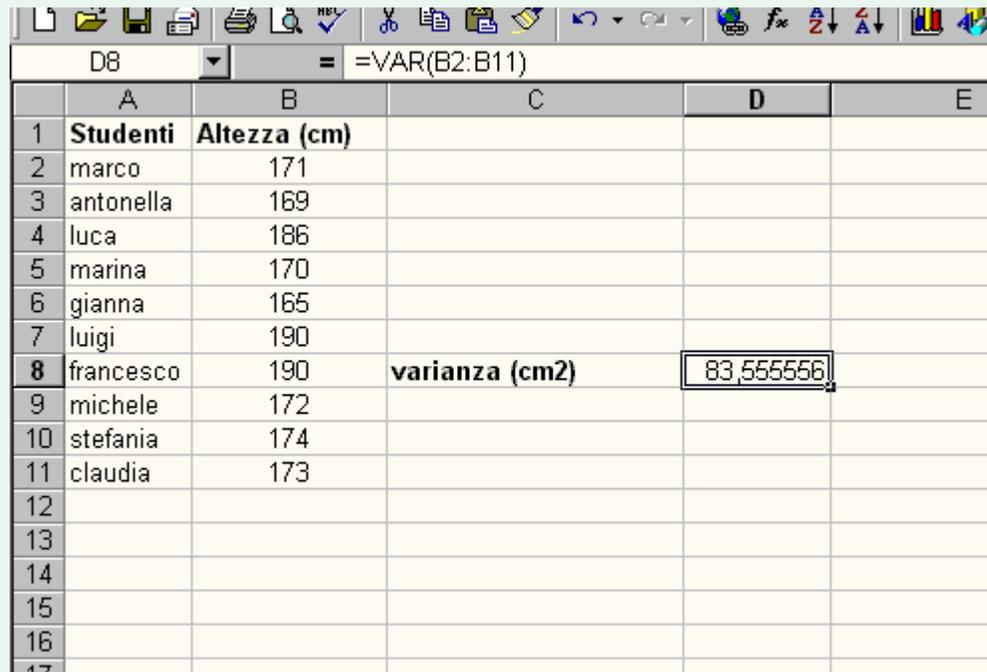
Interquartile

# Varianza

La varianza si determina attraverso la funzione **VAR** [VAR].

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

Come si fa:

**=VAR(B2:B11)**

# Deviazione standard

La **deviazione standard** si determina attraverso la funzione **DEV.ST** [STDEV].

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



The screenshot shows an Excel spreadsheet with the following data:

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

**=DEV.ST(B2:B11)**

## Ampiezza del campione

si ottiene come differenza tra l'estremo superiore e quello inferiore dei valori osservati del campione.

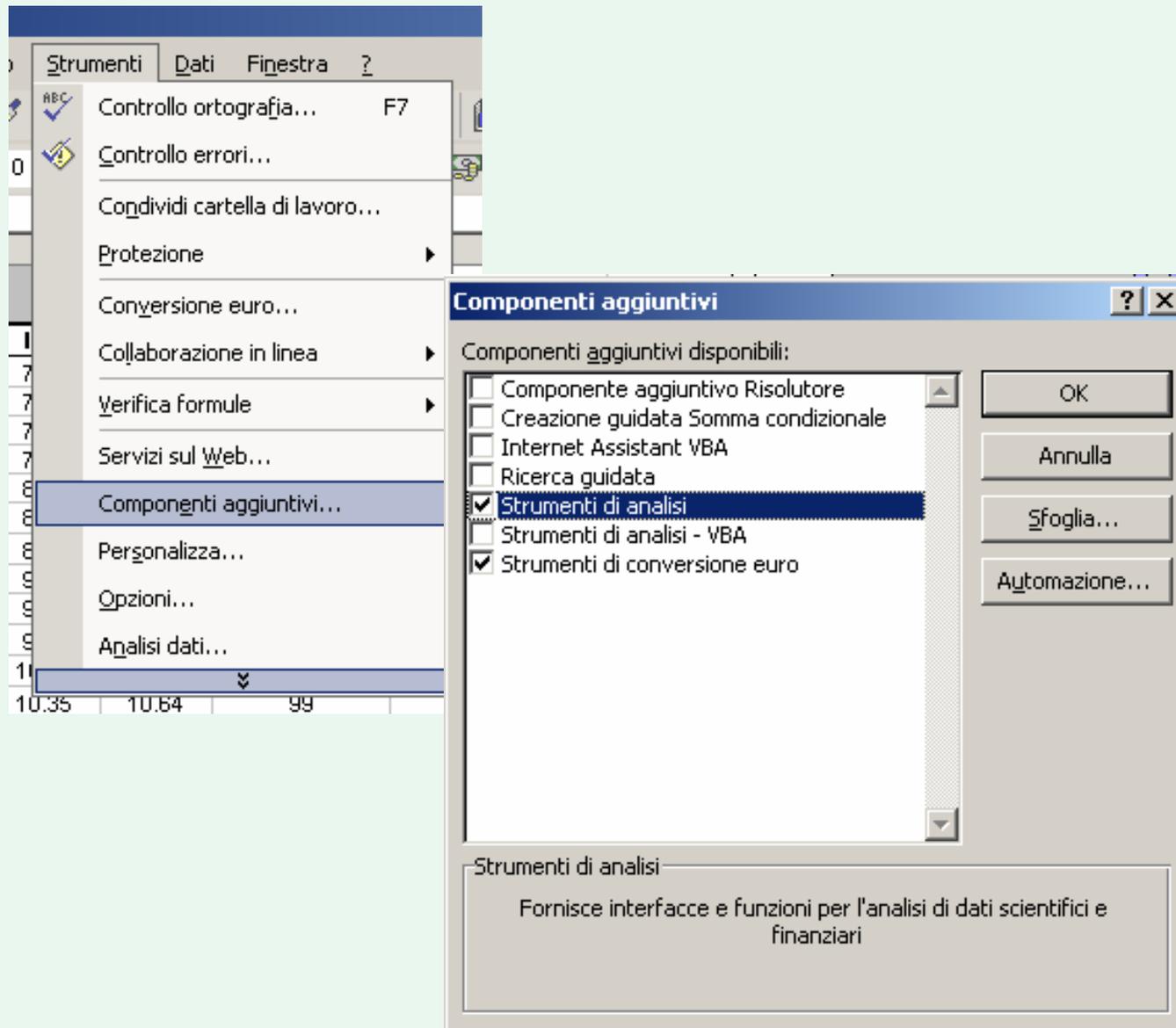
$$= \text{MAX}(\text{dati}) - \text{MIN}(\text{dati}).$$

## Ampiezza interquartile

si ottiene come differenza tra il terzo e il primo quartile

$$= \text{quartile}(\text{dati}, 3) - \text{quartile}(\text{dati}, 1).$$

# Componenti Aggiuntivi di Excel



J	K
	dati
1	93
2	84
3	98
4	99
5	108
6	92
7	102
8	80
9	91
10	94
11	91
12	98
13	79
14	91
15	108
16	95
17	104
18	92
19	98
20	112
21	96
22	97
23	79
24	102
25	88
26	102
27	118
28	98
29	110
30	87

Strumenti | Dati | Finestra | ?

- Controllo ortografia... F7
- Controllo errori...
- Condividi cartella di lavoro...
- Protezione
- Conversione euro...
- Collaborazione in linea
- Verifica formule
- Servizi sul Web...
- Componenti aggiuntivi...
- Personalizza...
- Opzioni...
- Analisi dati...**

**Statistica descrittiva**

Input

Intervallo di input: \$K\$1:\$K\$30

Dati raggruppati per:
 

- Colonne
- Righe

Etichette nella prima riga

Opzioni di output:

Intervallo di output: \$M\$3

Nuovo foglio di lavoro:

Nuova cartella di lavoro

Riepilogo statistiche

Livello di confidenza per media: 95 %

K-esimo più grande: 1

K-esimo più piccolo: 1

OK  
Annulla  
?

**Analisi dati**

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

OK  
Annulla  
?

Colonna1	
Media	96.2
Errore standard =dev.st/n <sup>0.5</sup>	1.75
Mediana	96.5
Moda	98
Deviazione standard	9.59
Varianza campionaria	91.89
Curtosi	-0.0662
Asimmetria	0.13
Intervallo	39
Minimo	79
Massimo	118
Somma	2886
Conteggio =n	30
Livello di confidenza(95.0%)	3.58

J	K
	dati
1	93
2	84
3	98
4	99
5	108
6	92
7	102
8	80
9	91
10	94
11	91
12	98
13	79
14	91
15	108
16	95
17	104
18	92
19	98
20	112
21	96
22	97
23	79
24	102
25	88
26	102
27	118
28	98
29	110
30	87

Strumenti Dati Finestra ?

- Controllo ortografia... F7
- Controllo errori...
- Condividi cartella di lavoro...
- Protezione
- Conversione euro...
- Collaborazione in linea
- Verifica formule
- Servizi sul Web...
- Componenti aggiuntivi...
- Personalizza...
- Opzioni...
- Analisi dati...**

### Analisi dati

#### Strumenti di analisi

- Smorzamento esponenziale
- Test F a due campioni per varianze
- Analisi di Fourier
- Istogramma
- Media mobile
- Generazione di un numero casuale
- Rango e percentile**
- Regressione
- Campionamento
- Test t: due campioni accoppiati per medie

**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

OK Annulla ?

Dato	Colonna1	Rango	Percento
27	118	1	100.00%
20	112	2	96.50%
29	110	3	93.10%
5	108	4	86.20%
15	108	4	86.20%
17	104	6	82.70%
7	102	7	72.40%
24	102	7	72.40%
26	102	7	72.40%
4	99	10	68.90%
3	98	11	55.10%
12	98	11	55.10%
19	98	11	55.10%
28	98	11	55.10%
22	97	15	51.70%
21	96	16	48.20%
16	95	17	44.80%
10	94	18	41.30%
1	93	19	37.90%
6	92	20	31.00%
18	92	20	31.00%
9	91	22	20.60%
11	91	22	20.60%
14	91	22	20.60%
25	88	25	17.20%
30	87	26	13.70%
2	84	27	10.30%
8	80	28	6.80%
13	79	29	0.00%
23	79	29	0.00%

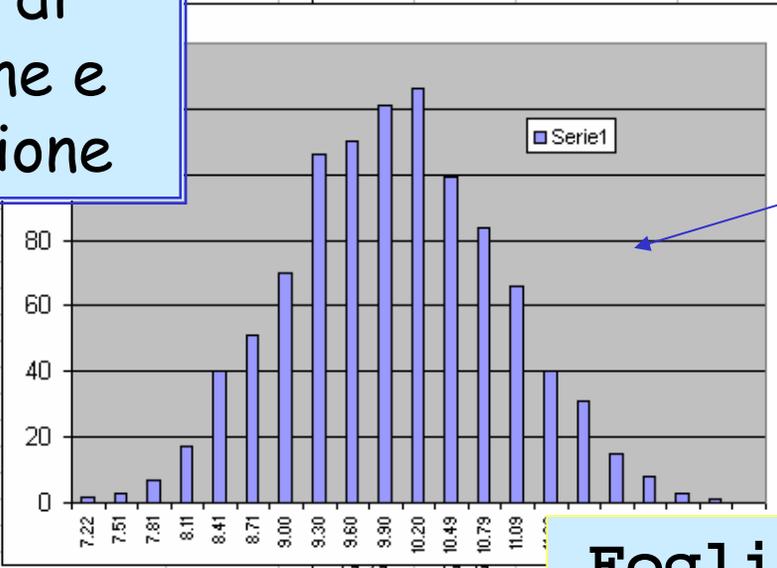
# Istogrammi di frequenza e indici statistici

A	B	C	D	E	F	G	H	I	J
1	<b>Dati</b>	<b>Nota:</b> inserire i dati nella colonna A del foglio <b>Dati</b> , modificare <b>k</b> e il <b>grafico</b>			<b>Limiti Classi</b>		<b>Tabella di frequenze</b>		
2	9.3403				<b>low</b>	<b>up</b>	<b>Freq. Ass.</b>	<b>Freq. Rel.</b>	<b>etichett</b>
3	8.4433				7.07	7.36	2	0.002	7.07
4	9.8992				7.36	7.66	3	0.003	7.36
5	9.9564				7.66	7.96	7	0.007	7.66
6	10.8832	<b>Numero di dati:</b>	1000		7.96	8.26	17	0.017	8.26
7	9.2346	<b>Minimo</b>	7.07		8.26	8.56	40	0.040	8.56
8	10.2698	<b>Massimo</b>	12.97		8.56	8.85	51	0.051	8.85
9	8.0678	<b>Ampiezza</b>	5.90		8.85	9.15	70	0.070	9.15
10	9.1191	<b>k (numero classi)</b>	20		9.15	9.45	106	0.106	9.45
11	9.4654	<b>h: ampiezza classe</b>	0.30		9.45	9.75	110	0.110	9.75
12	9.1337				9.75	10.05	121	0.121	10.05
13	9.8474	<b>controllo numero di dati:</b>	1000		10.05	10.35	126	0.126	10.35
14	7.9820				10.35	10.64	99	0.099	10.64
15	9.1692				10.64	10.94	84	0.084	10.94
16	10.8919	<b>Informazioni statistiche</b>			10.94	11.24	66	0.066	11.24
17	9.5542	<i>Media:</i>	9.9881		11.24	11.54	40	0.040	11.54
18	10.4926	<i>deviazione standard:</i>	0.9571		11.54	11.84	31	0.031	11.84
19	9.2772	<i>Varianza:</i>	0.9160		11.84	12.13	15	0.015	12.13
20	9.8439				12.13	12.43	8	0.008	12.43
21	11.2118				12.43	12.73	3	0.003	12.73
22					12.73	13.03	1	0.001	13.03
23									
24									
25									
26									
27									
28									
29									
30									
31									
32	10.6695								
33	9.7707								
34	10.2515								
35	9.2857								
36	10.8306								
37	10.5284								
38	10.6583								
39	9.4766								
40	9.6384								
41	10.6690								
42	9.1617								
43	9.4933								
44	11.3443								
45	9.0589								

Tabella di frequenze

Indici di posizione e dispersione

Istogramma



Foglio\_istogramma.xls

# Dati bivariati

K	L	M
	altezza	peso
	cm	Kg
1	167.4	66.4
2	151.2	60.0
3	176.4	70.0
4	178.2	70.7
5	194.4	77.1
6	165.6	65.7
7	183.6	72.9
8	144.0	57.1
9	163.8	65.0
10	169.2	67.1
11	163.8	65.0
12	176.4	70.0
13	142.2	56.4
14	163.8	65.0
15	194.4	77.1
16	171.0	67.9
17	187.2	74.3
18	165.6	65.7
19	176.4	70.0
20	201.6	80.0
21	172.8	68.6
22	174.6	69.3
23	142.2	56.4
24	183.6	72.9
25	158.4	62.9
26	183.6	72.9
27	212.4	84.3
28	176.4	70.0
29	198.0	78.6
30	156.6	62.1



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

Scelte disponibili:

Indietro | Avanti > | Fine

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

Titoli | Assi | Griglia | Legenda | Etichette dati

Titolo del grafico:

Asse dei valori (X):

Asse dei valori (Y):

Asse delle cat. (X) secondario:

Asse dei valori (Y) secondario:

Indietro | Avanti > | Fine

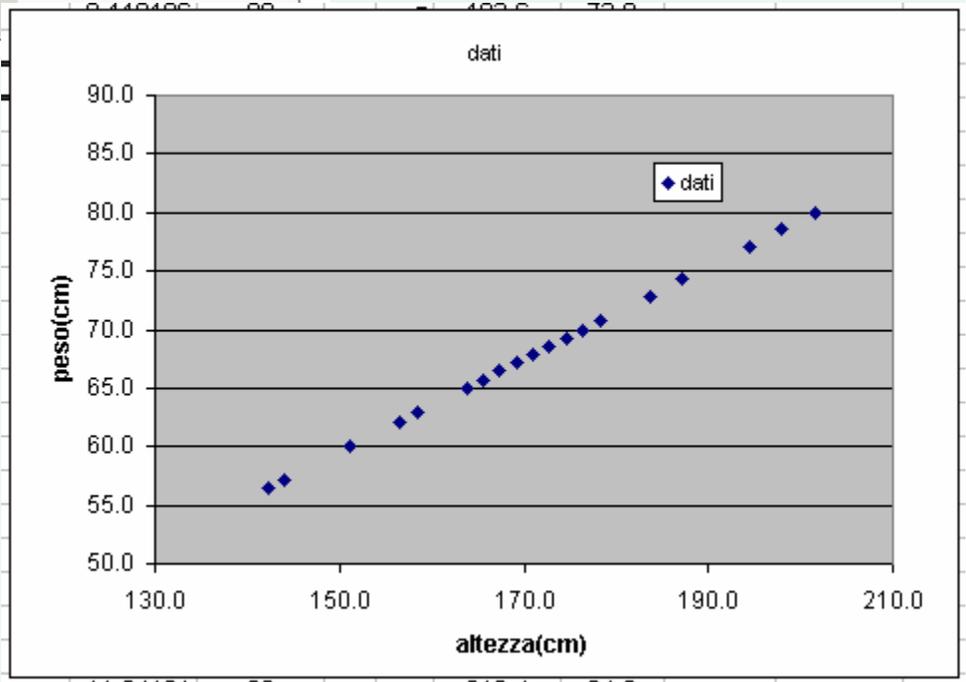
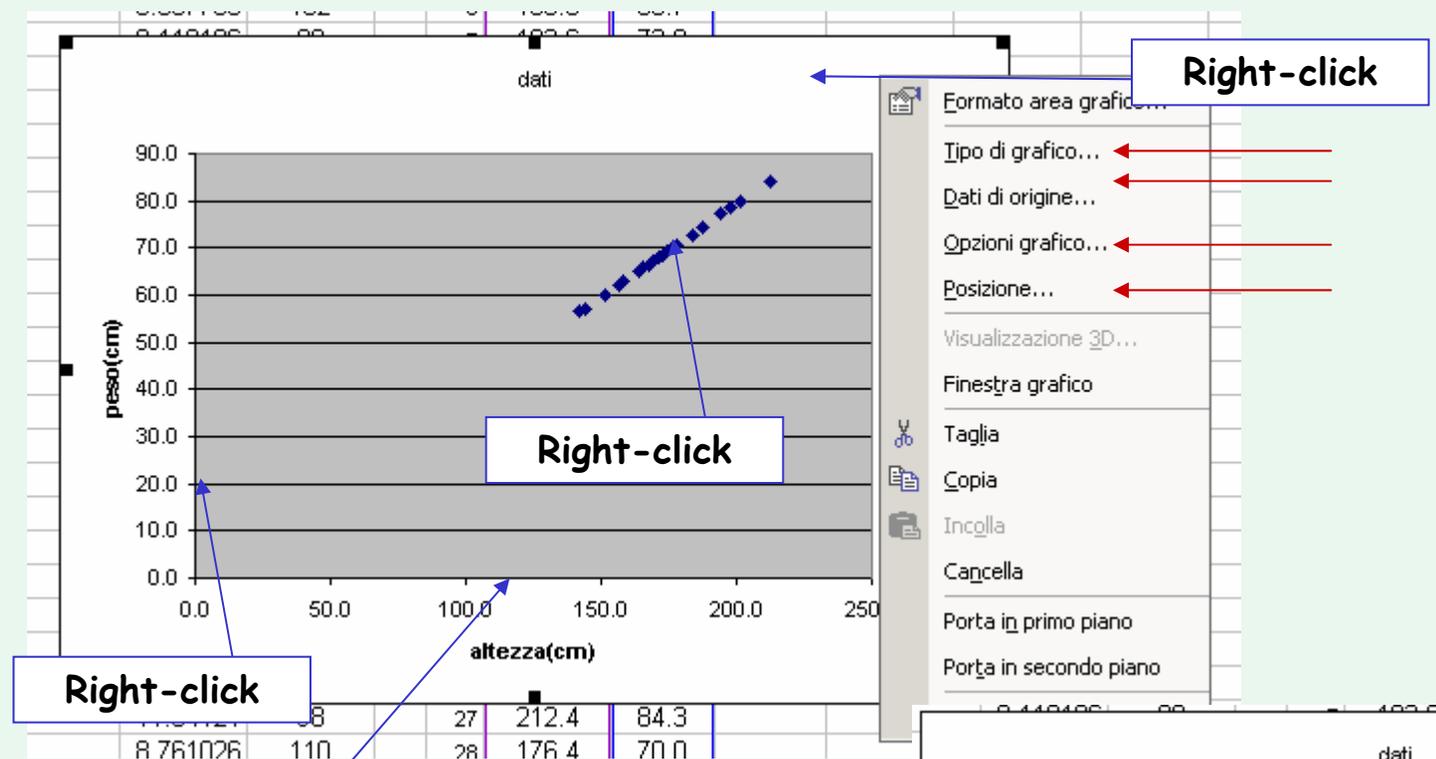
Creazione guidata Grafico - Passaggio 4 di 4 - Posizione grafico

Posizione grafico:

Crea nuovo foglio:

Come oggetto in:

Indietro | Avanti > | Fine



# Coefficiente di correlazione lineare

Misura la correlazione tra due variabili. In Excel si usa la funzione **CORRELAZIONE** [**CORREL**].

= **correlazione (dati\_x, dati\_y)**

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

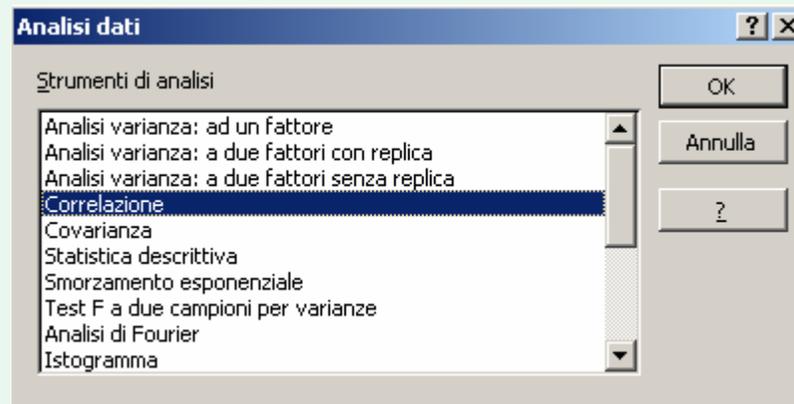
$$r = \frac{s_{xy}}{\sqrt{s_{xx}} \sqrt{s_{yy}}}$$

dove:

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

è  $n$  volte la covarianza fra  $X$  e  $Y$ .

K	L	M
	altezza	peso
	cm	Kg
1	167.4	66.5
2	151.2	60.2
3	176.4	70.8
4	178.2	71.3
5	194.4	77.8
6	165.6	66.2
7	183.6	73.6
8	144.0	58.1
9	163.8	65.5
10	169.2	67.8
11	163.8	65.5
12	176.4	70.8
13	142.2	57.1
14	163.8	65.1
15	194.4	77.7
16	171.0	68.6
17	187.2	74.7
18	165.6	65.8
19	176.4	70.2
20	201.6	81.0
21	172.8	69.3
22	174.6	69.4
23	142.2	56.5
24	183.6	73.7
25	158.4	63.6
26	183.6	73.6
27	212.4	84.7
28	176.4	70.1
29	198.0	78.8
30	156.6	62.9



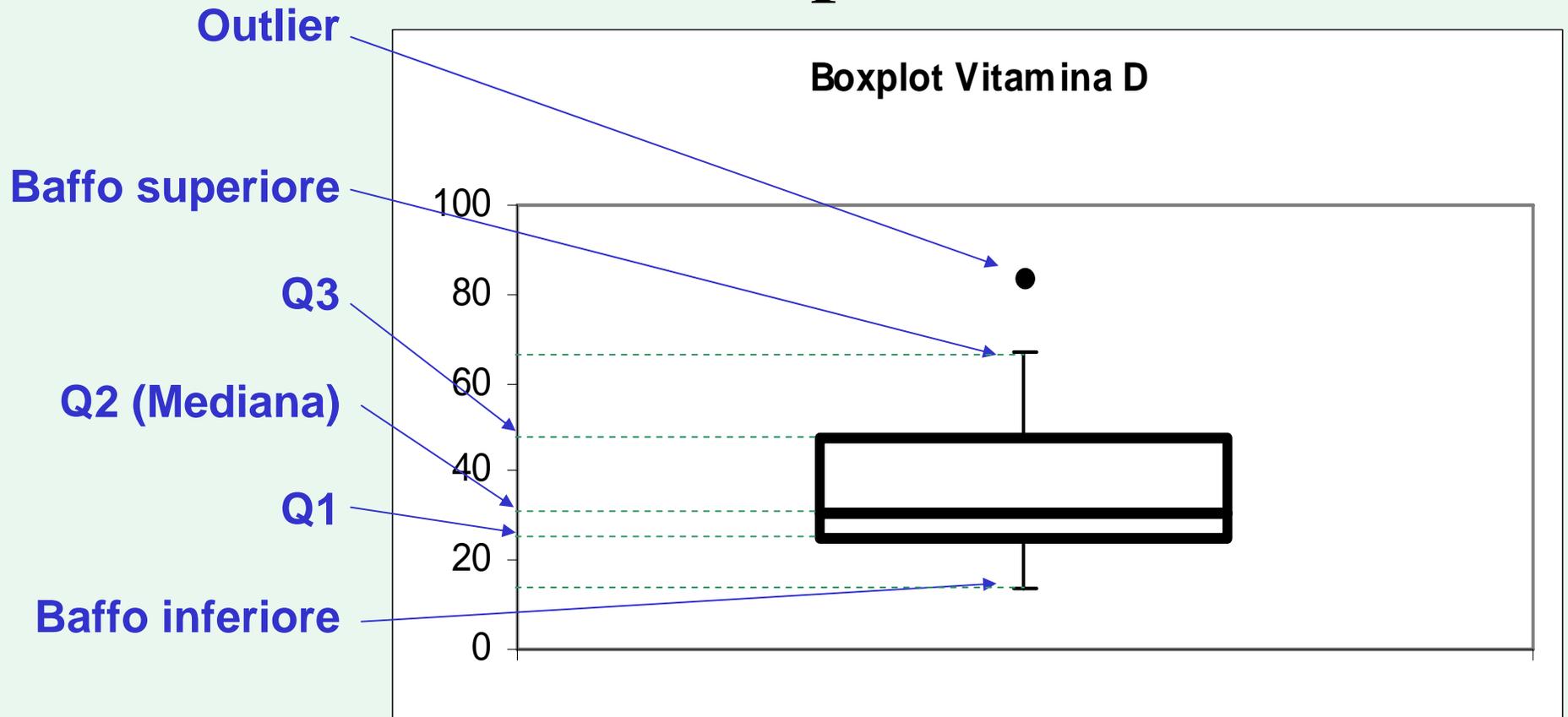
	<i>Colonna 1</i>	<i>Colonna 2</i>
<i>Colonna 1</i>	1	
<i>Colonna 2</i>	0.99918	1

# Box plot

E' una “scatola” in cui

- I bordi corrispondono a Q1 e Q3
- Una linea fra di essi indica il valore di Q2 (mediana)
- All'esterno vengono aggiunti:
  - Un “baffo superiore” = distanza da Q3 del più grande valore inferiore a  $Q3 + 1.5(Q3 - Q1)$
  - Un “baffo inferiore” = distanza da Q1 del più piccolo valore minore di Q1 ma maggiore di  $Q1 - 1.5(Q3 - Q1)$
- I valori esterni all'intervallo compreso tra i due “baffi”, detti “outliers”, vengono rappresentati individualmente

# Box plot



	A	B	C	D	E	F	G	H	I	J
1		Gruppo A	Gruppo B							
2		19	21							
3		16	19							
4		20	22							
5		23	24							
6		23	24							
7		25	25							
8		13	18							
9		19	21							
10		23	23							
11		16	20							
12		17	20							
13		14	18							
14		14	18							
15		17	20							
16		17	20							
17		13	18							
18		18	20							
19		18	21							
20		20	22							
21		18	21							
22		19	21							
23		18	21							
24		24	24							
25		19	21							
26		19	21							
27		18								
28		25								
29		22								
30		27								
31		18								
32	media	19.07	20.92							
33	q3	21.5	22							
34	max	27	25							
35	mediana	18.50	21.00							
36	min	13	18							
37	q1	17	20							

