

Altri indici di variabilità

Perché cercare altri indici di variabilità?

- A. Perché varianza e deviazione standard sono legate alla media e sono fortemente influenzate anche da un solo dato anomalo.**
- B. Perché varianza e deviazione standard richiedono calcoli lunghi, da eseguire con il computer se i dati sono numerosi.**

- ...

A. Come superare l'influenza di un dato anomalo?

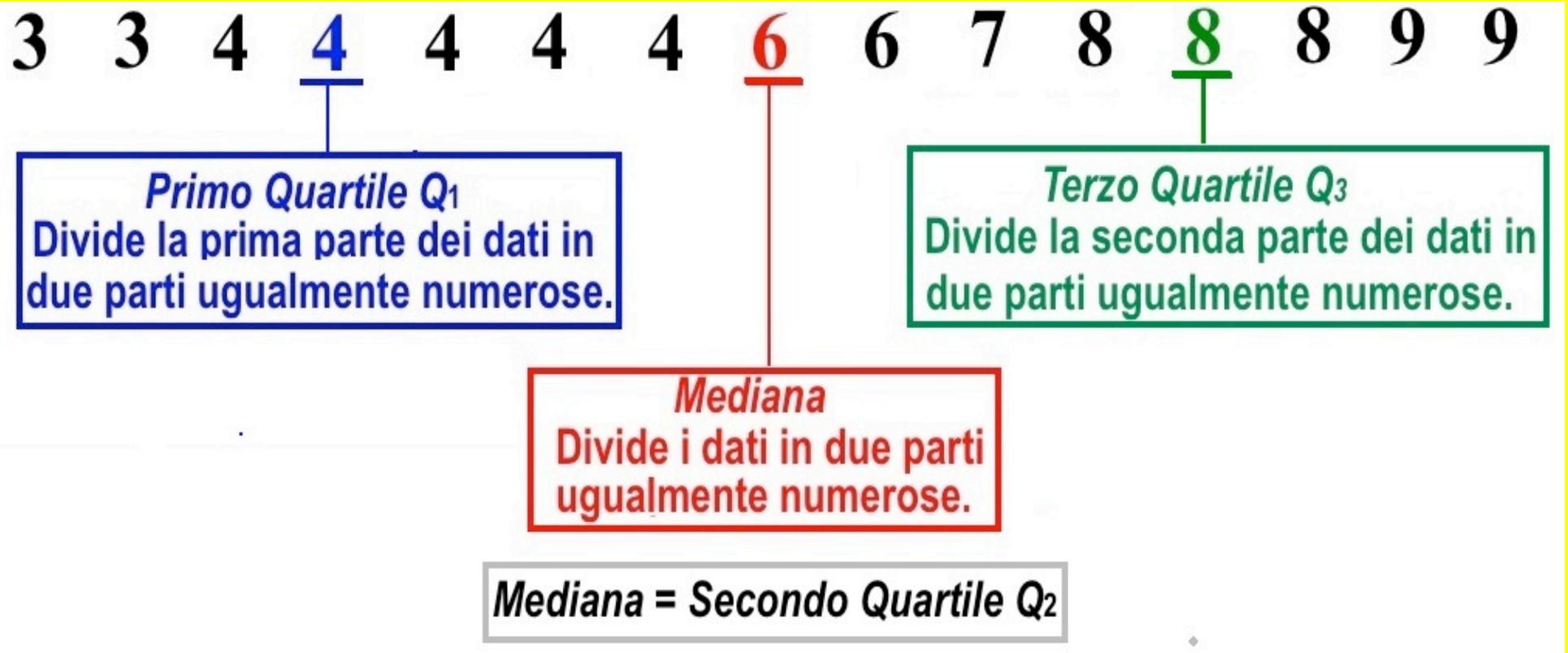
Un'idea per superare il primo problema: un dato anomalo non influenza la mediana.

La statistica propone perciò un procedimento per valutare la dispersione attorno alla mediana.

Ecco il procedimento.

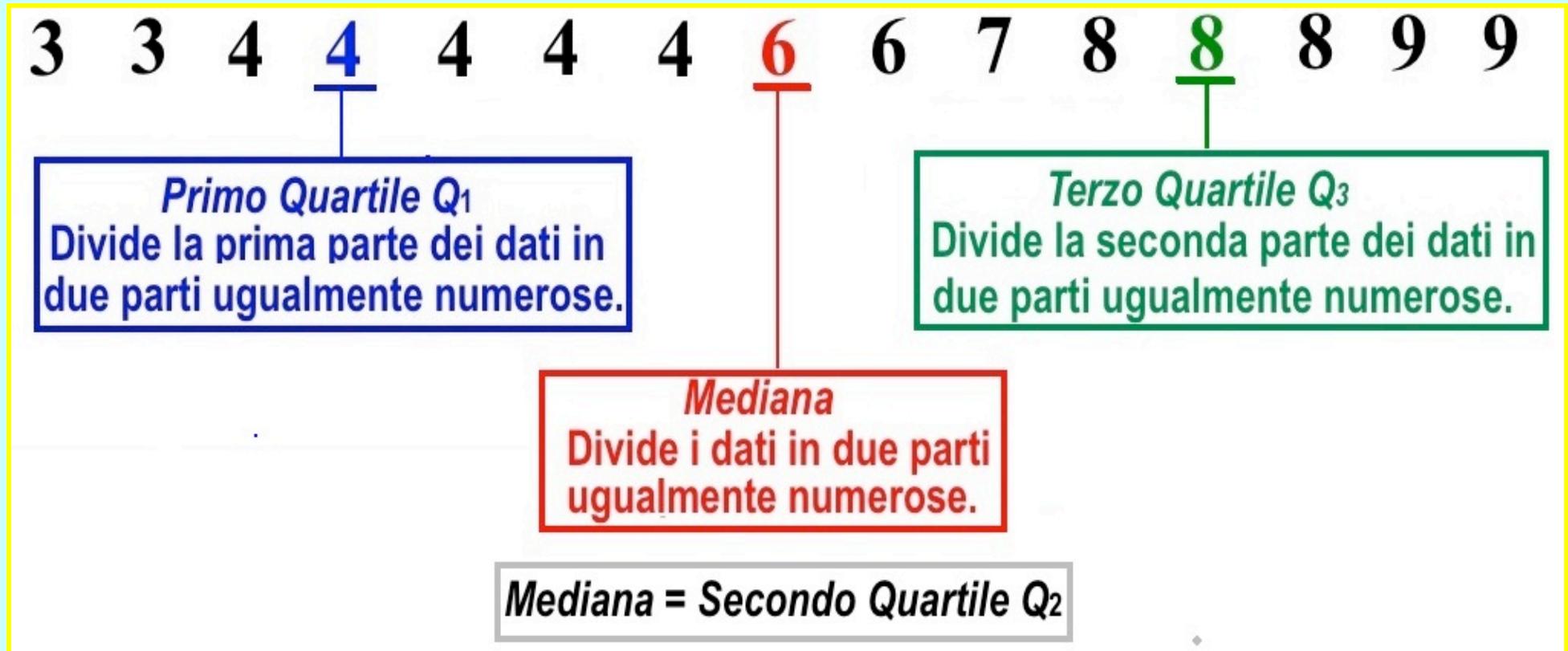
Mediana e quartili

Esamino i voti dei ragazzi



1. Determino il primo e il terzo quartile

Mediana e differenza interquartile

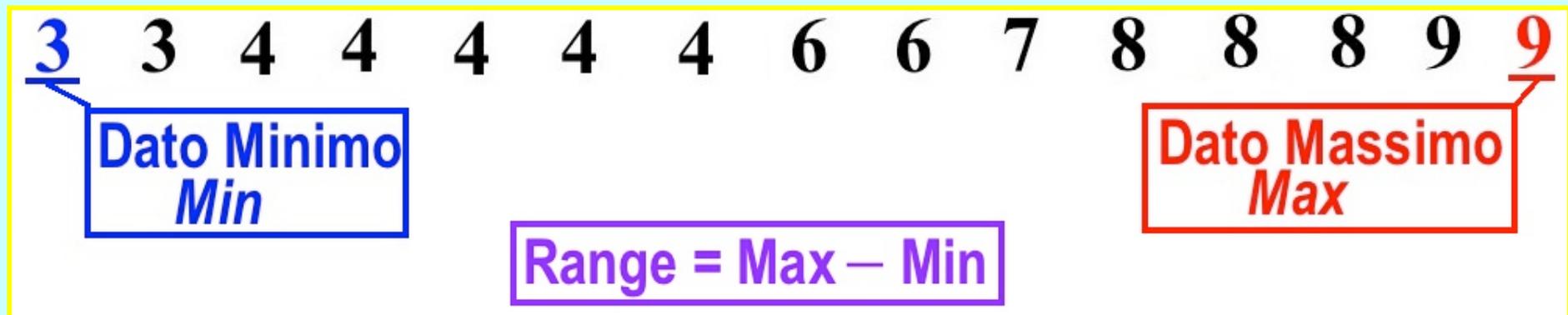


2. Valuto la variabilità attorno alla mediana con la **differenza interquartile** data da

$$Q_3 - Q_1 = 8 - 4 = 4$$

B. Come superare la lunghezza dei calcoli?

Ecco un'idea per superare il secondo problema.
Esamino ancora i voti dei ragazzi



Per valutare la variabilità calcolo il *Range*
Dato massimo – Dato minimo = 9 – 3 = 6

Il Range prende anche il nome di '*Campo di variazione*'

Attività

**Completa la scheda di lavoro per
confrontare i diversi indici di variabilità.**

Che cosa hai trovato?

A. Influenza di un dato anomalo

Quesito 1

1. Descrivi l'effetto dell'ultimo dato 'anomalo' sugli indici elencati qui sotto

- La Media è circa raddoppiata
- La Varianza è moltiplicata per circa 100
- La Dev. standard è moltiplicata per circa 10
- Il Range è moltiplicato per circa 15
- La Diff. interquartile rimane invariata
- La Mediana rimane invariata

Con il foglio di calcolo

Mediana: `Mediana[C2:C16]`

Range: `Max[C2:C16] - Min[C2:C16]`

Differenza interquartile:

`Q3[C2:C16] - Q1[C2:C16]`

RAGAZZI	RAGAZZI	VOTO NUOVO	VOTO NUOVO
3		3	
3		3	
4		4	
4		4	
4		4	
4		4	
4		4	
4		4	
6		6	
6		6	
7		7	
8		8	
8		8	
8		8	
8		8	
9		9	
9		90	
Media	5.8	Media	11.2
Varianza	4.56	Varianza	447.36
Dev.Stand.	2.14	Dev.Stand.	21.15
Range	6	Range	87
Q3 - Q1	4	Q3 - Q1	4
Mediana	6	Mediana	6



B. Tre serie di dati 'insoliti'

Quesito 2

2. Esegui il lavoro richiesto e rispondi ai seguenti quesiti:

a. Quando i dati sono tutti uguali,

- quanto valgono i valori medi? **Il dato ripetuto**

- quanto valgono le misure di variabilità? **Zero**

b. Confronta dati e indici statistici del gruppo 2 e del gruppo 3 rispondi alle domande

- Quale diversità trovi nei due gruppi di dati?

Il dato centrale in un caso è 2 e nell'altro 10

- Quale effetto ha la diversità sugli indici statistici?

Influenza solo i valori medi: la media aumenta di circa 1 e la mediana passa da 2 a 10.

	Gruppo 1	Gruppo 1	Gruppo 2	Gruppo 2	Gruppo 3	Gruppo 3
	6		2		2	
	6		2		2	
	6		2		2	
	6		2		2	
	6		2		2	
	6		2		10	
	6		10		10	
	6		10		10	
	6		10		10	
	6		10		10	
	6		10		10	
	6		10		10	
	6		10		10	
	6		10		10	
Media		6	Media2	5.64	Media3	6.36
Varianza		0	Varianza2	15.87	Varianza3	15.87
Dev.Stand.		0	Dev.Stand2	3.98	Dev.Stand3	3.98
Range		0	Range2	8	Range3	8
$Q_3 - Q_1$		0	$(Q_3 - Q_1)_2$	8	$(Q_3 - Q_1)_3$	8
Mediana		6	Mediana2	2	Mediana3	10

C. Due gruppi di dati a confronto

Quesito 3

3. Esegui il lavoro richiesto e scrivi qui sotto le risposte alle domande.

a. Confronto i valori medi di ragazze e ragazzi.

Trovo la stessa media e la stessa mediana

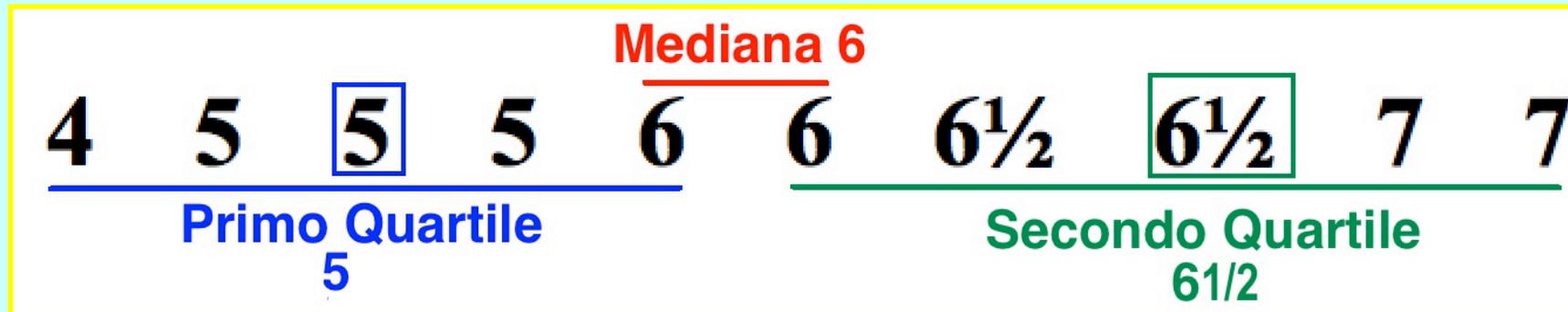
b. Confronto la variabilità di ragazze e ragazzi.

Tutte le misure di variabilità dei ragazzi sono più grandi di quelle delle ragazze; in particolare range e deviazione standard sono circa il doppio e la differenza interquartile quasi il triplo.

RAGAZZE	RAGAZZE	RAGAZZI	RAGAZZI
4		3	
5		3	
5		4	
5		4	
6		4	
6		4	
6.5		4	
6.5		6	
7		6	
7		7	
Media	5.8	8	
Varianza	0.91	8	
Dev.Stand.	0.95	8	
Range	3	9	
Q₃ - Q₁	1.5	9	
Mediana	6	Media	5.8
		Varianza	4.56
		Dev.Stand.	2.14
		Range	6
		Q₃ - Q₁	4
		Mediana	6

C. Determinare i quartili

Quesito 3c



Il numero dei dati (10) è pari.

La mediana è la media dei due dati centrali.

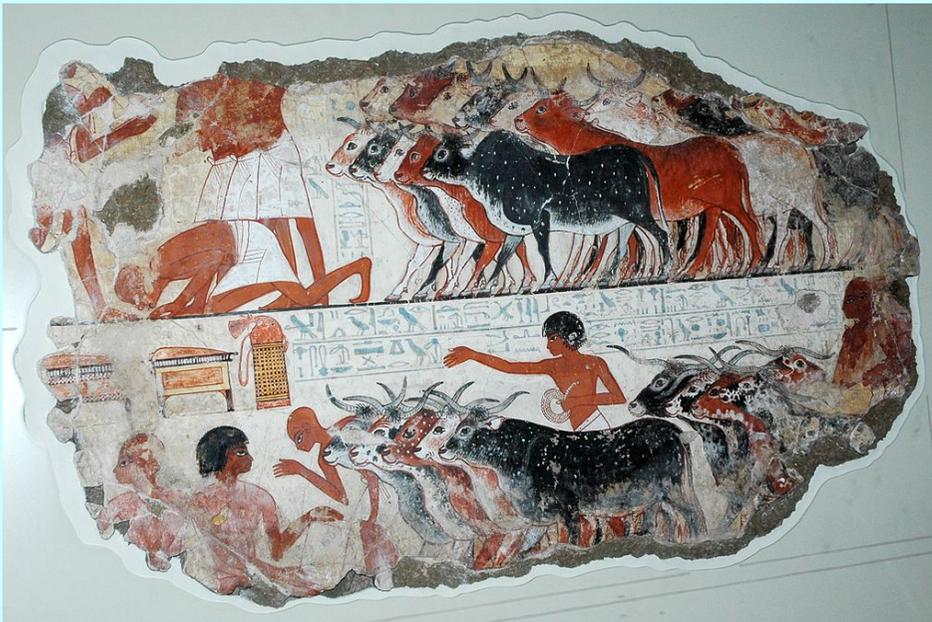
Per determinare i quartili:

- Q₁ è la mediana dei primi 5 dati;
- Q₂ è la mediana degli ultimi 5 dati.

Uno sguardo alla storia

Le antiche civiltà

I censimenti per contare la popolazione, gli animali, le abitazioni, ...



**Censimento di bovini nell'Antico Egitto
circa 1350 a.C.**

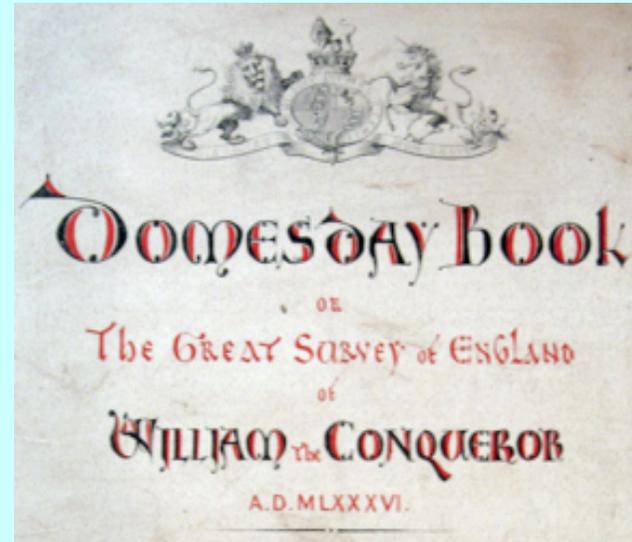


**Censimento a Roma
II secolo a.C.**

Dal Medio Evo al Rinascimento

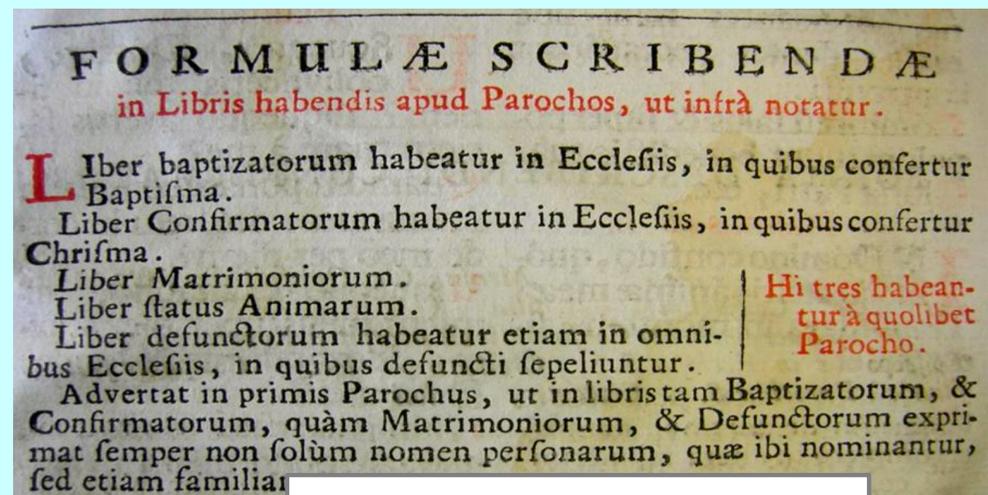
Censimenti, registri di matrimoni, battesimi e morti, ...

Carlo Magno
Censimento
808



Guglielmo il
Conquistatore
Censimento
1086

Elenchi di nascite,
matrimoni, morti,...



Registro dei battesimi
introdotto dal 1565

Nel secolo di Galileo

Origini della statistica come scienza

John Graunt



CAPTAIN JOHN GRAUNT

Natural and Political
OBSERVATIONS
Mentioned in a following INDEX,
and made upon the
Bills of Mortality.

By *JOHN GRAUNT*,
Citizen of
LONDON.

With reference to the Government, Religion, Trade,
Growth, Age, Diseases, and the several Changes of the
said CITY.

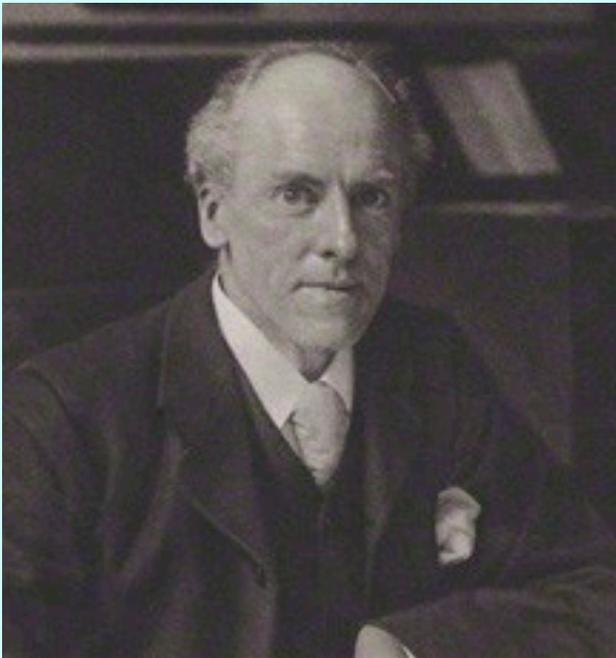
— *Nas, ne ut miratur Turba, labere.*
Contentus pariter Litteribus —

LONDON,
Printed by Tho: Roycroft, for John Martin, James Allcock,
and Tho: Dineen, at the Sign of the Bell in St. Paul's
Church-yard, MDCLXII.

Non si limita a contare nascite e morti, ma applica metodi scientifici per scoprire regolarità nei dati.

Dal 1700 a oggi

La statistica studia *con metodi matematici* fenomeni dotati di variabilità in qualsiasi campo del sapere: botanica, agricoltura, zoologia, medicina, fisica....



**K. Pearson (1857 – 1936)
Introduce la deviazione standard nel 1893**



**R. Fischer (1890 – 1962)
Introduce la varianza nel 1920**