

# PCTO in Coding & Data Science

## Modulo 1

Liceo Morgagni, Roma, 01/02/2022

Il Data scientist è il **terzo**  
profilo professionale più  
richiesto a livello globale

[fonte]

La differenza di stipendio tra uomini e  
donne nelle STEM è del **19,60%**

[fonte]

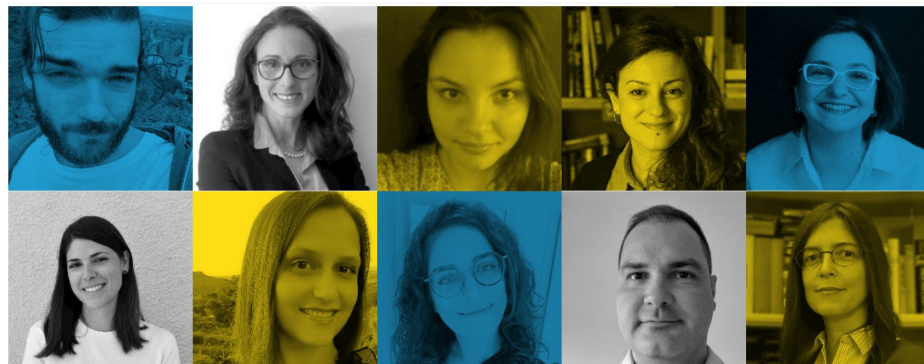
Si iscrivono a corsi di laurea STEM:

**6 ragazzi su 10**

**2 ragazze su 10**

[fonte]

**CD:50/50**  
**CODING DIVERSITY**



<https://cd5050.org/>

**CD:50/50**  
**CODING DIVERSITY**

# PROGRAMMA

- Modulo 1 - Introduzione: la data science e il coding
- Modulo 2 - Riassumere i dati: statistica e visualizzazione dei dati
- Modulo 3 - Costruire modelli: capire e predire il mondo
- Modulo 4 - Intelligenza Artificiale

# PROGETTO

**Ogni giorno vi chiederemo di svolgere dei compiti per il giorno successivo, costruendo un piccolo **progetto** a gruppi.**

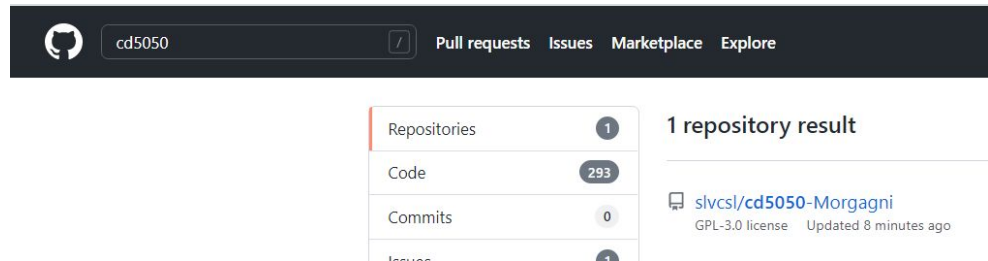
**Per lo più si tratterà di ripetere attività svolte in classe, ma su dati differenti.**

**L'ultimo giorno condivideremo i risultati e ne discuteremo insieme**

**GitHub- <https://github.com/>**

**<https://github.com/slvcsl/cd5050-Morgagni>**

- 1) Collegarsi a github.com
- 2) Cercare “cd5050-Morgagni”
- 3) Cliccare sull'unico risultato



**Discord:**

**<https://discord.gg/fJs4BGdW>**

**Telegram:**

**<https://t.me/+jtDRevz2KlCcxZDU0>**

# Presentiamoci con un dato

- Silvia: nata in Sicilia
- Riccardo: ha un fratello
- ...

# Cosa sono i dati?

- fatti o informazioni che possono essere usati per calcolare, analizzare o pianificare qualcosa.

<https://www.merriam-webster.com/dictionary/data>

# Cosa sono i dati?

- **fatti o informazioni** che possono essere usati per **calcolare, analizzare o pianificare** qualcosa.

<https://www.merriam-webster.com/dictionary/data>

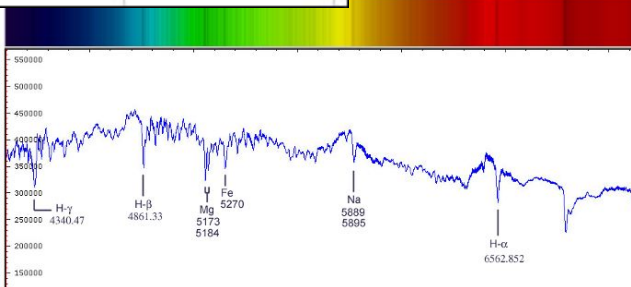
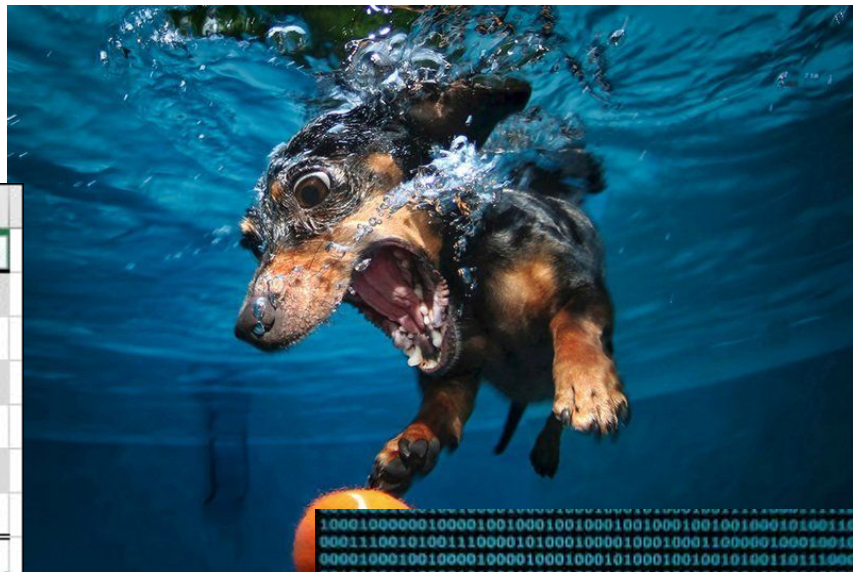
- 30: ?
- 30 anni: ?
- Riccardo ha 30 anni -> è un dato
- Riccardo ha 30 anni e guarda l'Attacco dei Giganti -> è un dato utile (per Netflix)

**Un dato diventa informazione in un contesto**



# Molti tipi di dati

	A	B	C	D
1	Prodotto ▼	Trim 1 ▼	Trim 2 ▼	Totale com ▼
2	Chocolade	€ 744,60	€ 162,56	€ 907,16
3	Gummibarchen	€ 5.079,60	€ 1.249,20	€ 6.328,80
4	Scottish Longbreads	€ 1.267,50	€ 1.062,50	€ 2.330,00
5	Sir Rodney's Scones	€ 1.418,00	€ 756,00	€ 2.174,00
6	Tarte au sucre	€ 4.728,00	€ 4.547,92	€ 9.275,92
7	Chocolate Biscuits	€ 943,89	€ 349,60	€ 1.293,49
8	Totale	€ 14.181,59	€ 8.127,78	€ 22.309,37



Quando è stata l'ultima volta  
che avete prodotto dei dati  
su di voi?

# Utenti passivi: generano i dati

## Modalità di personalizzazione degli annunci

Gli annunci sono basati sulle informazioni personali che hai aggiunto al tuo Account Google, sui dati derivanti dagli inserzionisti che collaborano con Google e sulla deduzione dei tuoi argomenti di interesse da parte di Google. Scegli un qualsiasi fattore per conoscere ulteriori informazioni o per aggiornare le tue preferenze. [Scopri come controllare gli annunci che vedi](#)



25-34 anni



Donna



Lingua: italiano e altre 1



Trello



JetBrains



Grammarly



Remarkable

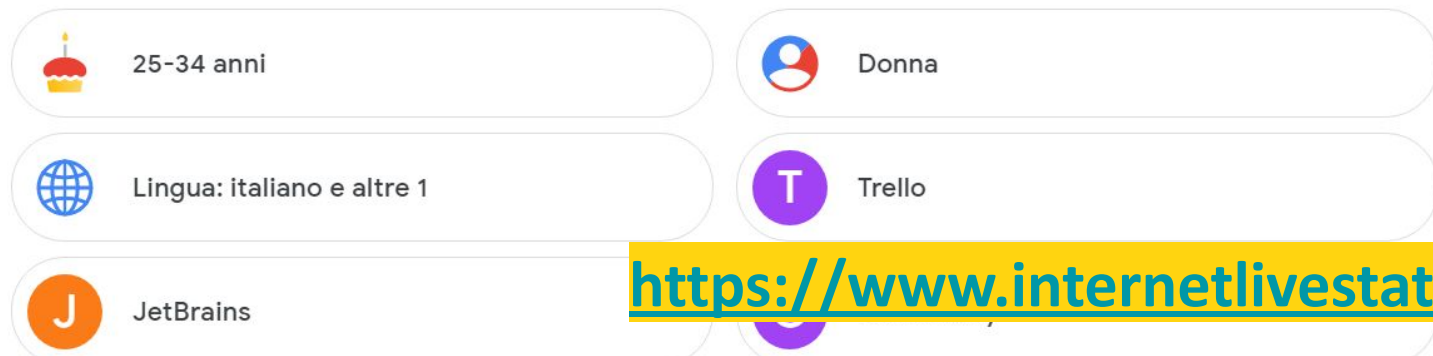


Google

# Utenti passivi: generano i dati

## Modalità di personalizzazione degli annunci

Gli annunci sono basati sulle informazioni personali che hai aggiunto al tuo Account Google, sui dati derivanti dagli inserzionisti che collaborano con Google e sulla deduzione dei tuoi argomenti di interesse da parte di Google. Scegli un qualsiasi fattore per conoscere ulteriori informazioni o per aggiornare le tue preferenze. [Scopri come controllare gli annunci che vedi](#)



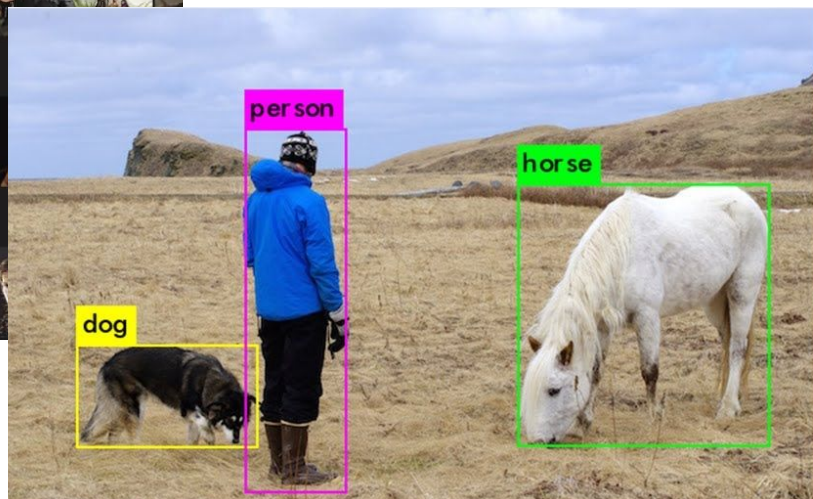
Quando è stata l'ultima volta  
che avete utilizzato dei dati?

# Diventiamo attivi: la data science

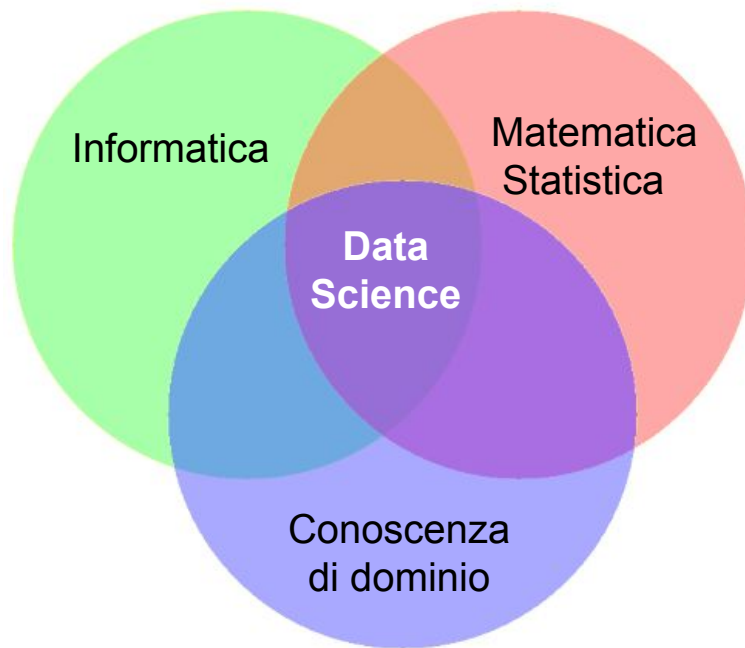




## A smiling woman with short curly hair, wearing glasses and a white lab coat over a blue shirt. She has a stethoscope around her neck and is giving a thumbs up. In the background, there is a collage of TV show posters including 'Killer e gialli TV', 'PENNY DREADFUL', 'NETFLIX BETWEEN', 'WHITECOLLAR', 'Breaking Bad', 'weeds', 'Sons of Anarchy', 'NETFLIX NO ALL'OSSO', and 'NON È IN PAESE'.



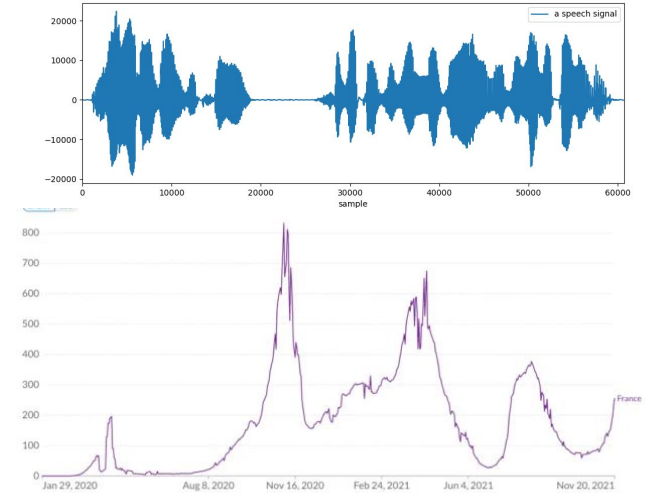
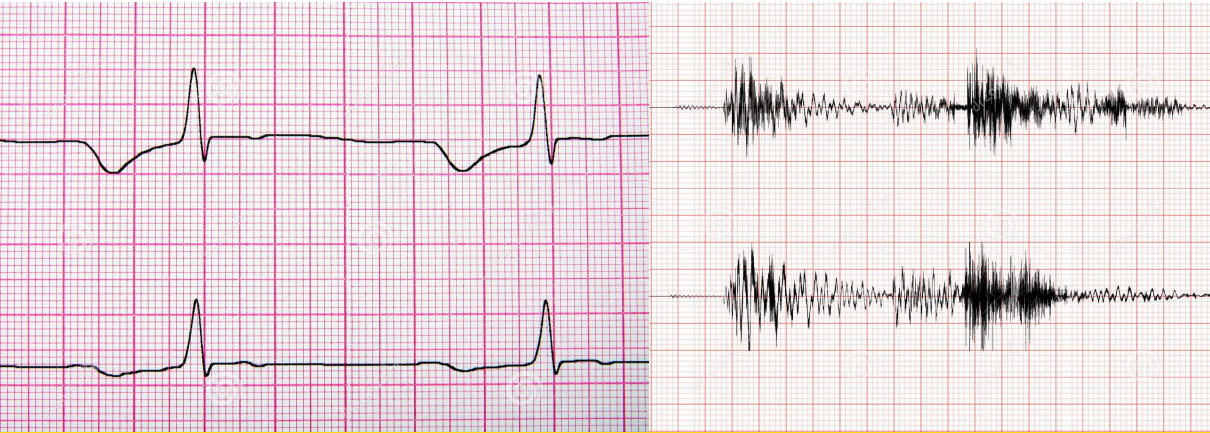
# Data science: gli ingredienti





# Conoscenza di dominio

La conoscenza del dominio è la conoscenza del **contesto** da cui provengono i dati e permette di analizzarli tenendo conto del loro **significato**.



# Matematica

Una penna e una caramella costano insieme 1.10 €, e la penna costa un euro in più della caramella. Quanto costa la caramella?



# Matematica

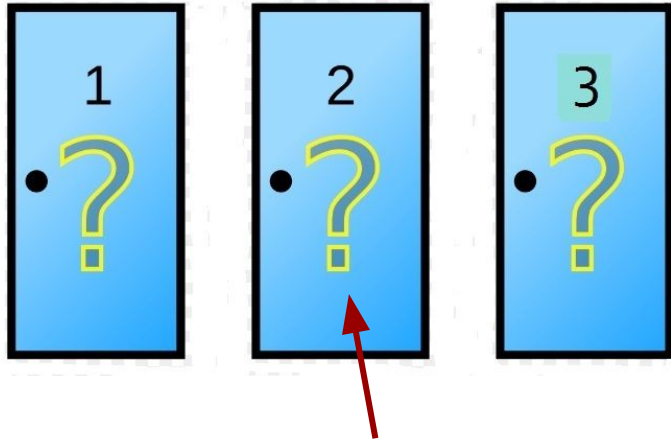
Una penna e una caramella costano insieme 1.10 €, e la penna costa un euro in più della caramella. Quanto costa la caramella?



$$\begin{cases} x_1 + x_2 = 1.1 \\ x_1 = x_2 + 1 \end{cases} \quad \begin{aligned} (x_2 + 1) + x_2 &= 1.1 \\ 2x_2 &= 0.1 \\ x_2 &= 0.05 \end{aligned}$$

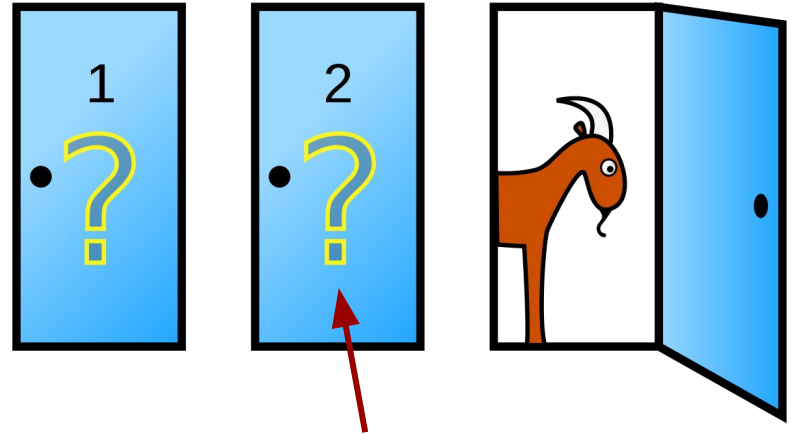
# Matematica

In un quiz televisivo abbiamo 3 porte uguali che nascondono ognuna un premio. Dietro una c'è una automobile, mentre dietro le altre due ci sono delle capre. All'inizio ci è chiesto di scegliere una porta qualsiasi.

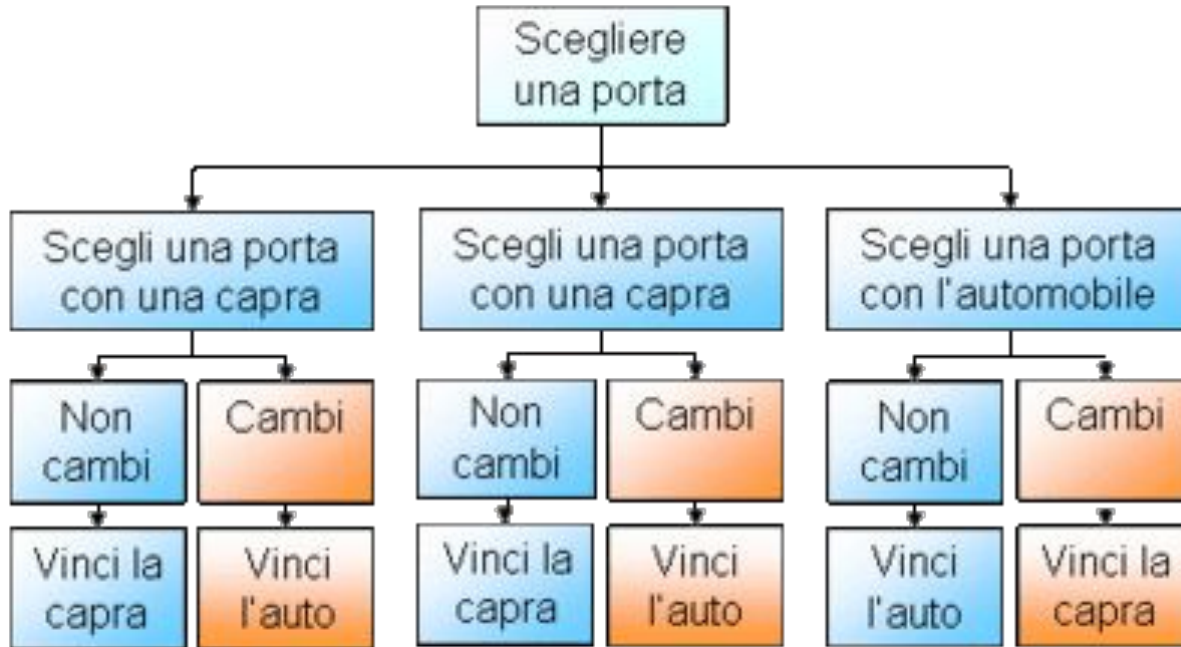


Dopo averla scelta, il conduttore apre una delle altre due porte, e mostra la presenza di una capra. Ora ci viene data la possibilità di cambiare porta.

Cosa fareste? Cambiate o restate sulla scelta iniziale?



# Matematica



Se cambi, due volte su tre vinci l'auto: quindi conviene sempre cambiare!

# Informatica

Nel censimento del 1880 degli Stati Uniti, l'analisi dei dati richiese **8 anni**.

Herman Hollerith (che fondò poi l'IBM) progettò la tabulating machine, usata nel censimento del 1890 che risolse l'analisi in **2 anni** di calcoli.

Nel 1930 il tempo scese a meno di **due mesi**.

Fonte: The History of Census Tabulation - Keith S. Reid-Green



# Informatica



Il SuperComputer più potente al mondo (il Fugaku) punta a risolvere alcune delle più complesse sfide al mondo.

Tra queste, predire il clima con anni di anticipo, e fornire nuove risposte al cambiamento climatico.

# Cominciamo a vedere qualche dato concreto

[Dati ISTAT scuole di secondo grado](#)



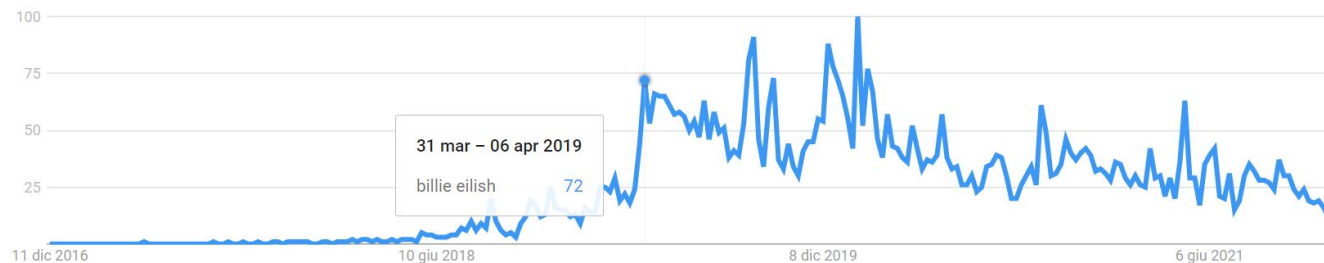


billie eilish

Termine di ricerca

+ Confronta

Interesse nel tempo ?



Il 30 gennaio dello stesso anno viene annunciato che l'album di debutto *When We All Fall Asleep, Where Do We Go?*, contenente quattordici tracce, sarebbe stato pubblicato il 29 marzo.<sup>[41]</sup> Parallelamente a questo annuncio, viene estratto il

## Album in studio [\[ modifica \]](#) [\[ modifica wikitesto \]](#)

- 2019 – *When We All Fall Asleep, Where Do We Go?*
- 2021 – *Happier than Ever*

# Coding

Ottenere un risultato di calcolo, attraverso una serie di **comandi impartiti al computer**; trovare la sequenza di istruzioni (algoritmo) che permetta di automatizzare un certo processo.

## La moltiplicazione in colonna:

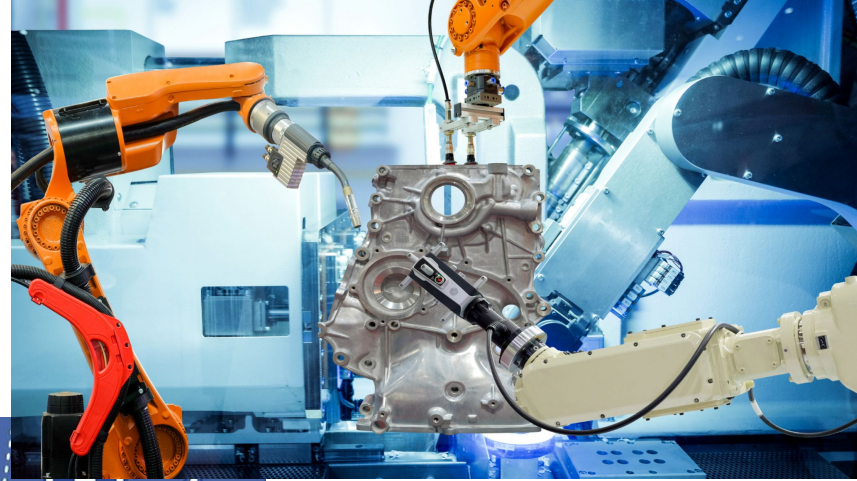
Dati due numeri, il processo è sempre uguale. Programmare significa tradurre un ragionamento\algoritmo in modo che un computer lo possa attuare.

$$\begin{array}{r} 147 \times \\ 25 = \\ \hline 735 + \\ 2940 = \\ \hline 3675 \end{array} \rightarrow \text{Risultato finale}$$

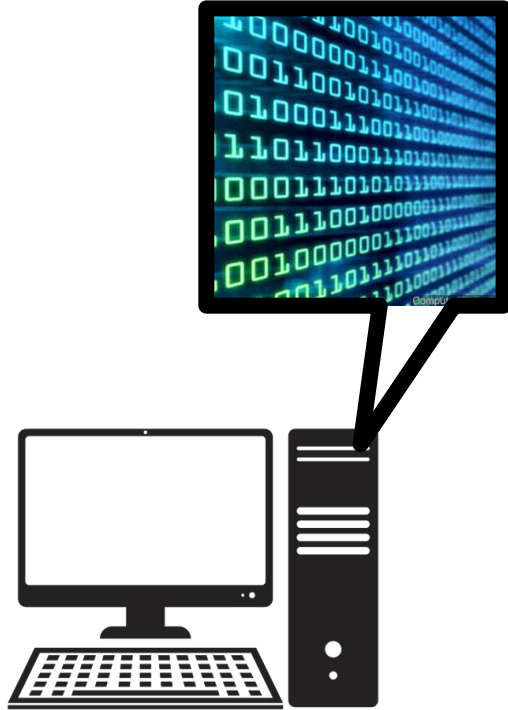
*Il genio della matematica*

# Coding

Ha molte applicazioni diverse!



# Coding



```
n=20  
m=30  
p=n+m  
print(p)
```

Dimmi quanto fa  $20 + 30$



# Coding

Top 10 Data Science Programming Language by % of Job Ads in which the Language is Mentioned

