



I veicoli Italiani



Un'analisi preliminare ed esplorativa



Una panoramica sul dataset

- Dati sul parco circolante dei veicoli italiani
- Pubblicato dal *Ministero delle Infrastrutture e dei Trasporti* sotto licenza **CC-BY**
- Contiene **52.165.189** records
- Peso di circa **7.7 GB** in .csv

Dati di ogni veicolo

- Tipologia e destinazione veicolo
- Residenza, età e sesso dell'intestatario
- Data immatricolazione
- Dettagli sul motore, emissioni, peso e marca
- Dettagli sull'assicurazione e revisione



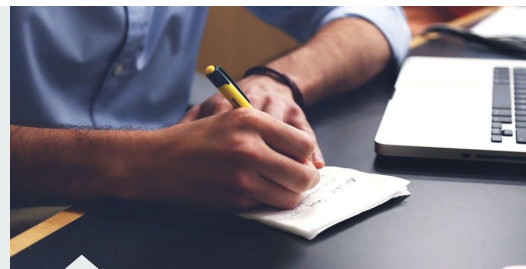


Obiettivi

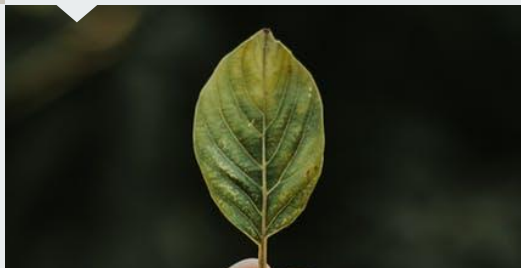


**Esplorare il
dataset**

**Valutare l'impatto
sull'ambiente**



**Sperimentare
nuove tecnologie**





Rapporto veicoli/persone

Dati ufficiali sulla popolazione presi dal sito ISTAT: **60.483.973** persone

Facendo il rapporto, viene **0,86 veicoli per persona**

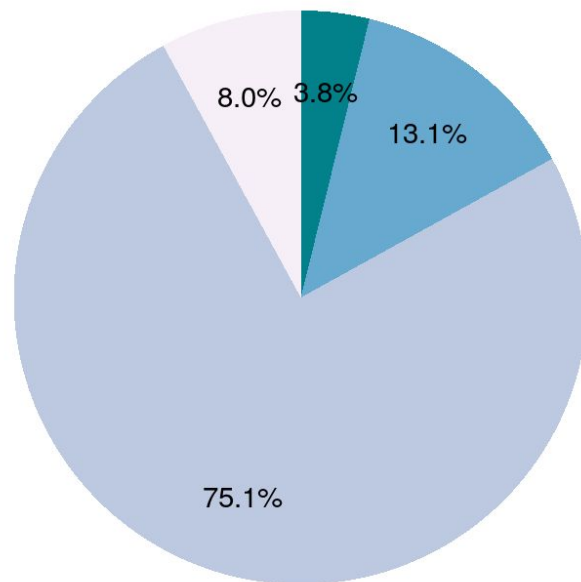
Tra i rapporti più alti d'Europa [1]



[1] https://en.wikipedia.org/wiki/List_of_countries_by_vehicles_per_capita







Tipologia veicoli



I mezzi di trasporto privato costituiscono più del'88% del parco vetture italiano

Sotto la voce "Others" sono contemplati anche i mezzi pubblici o speciali

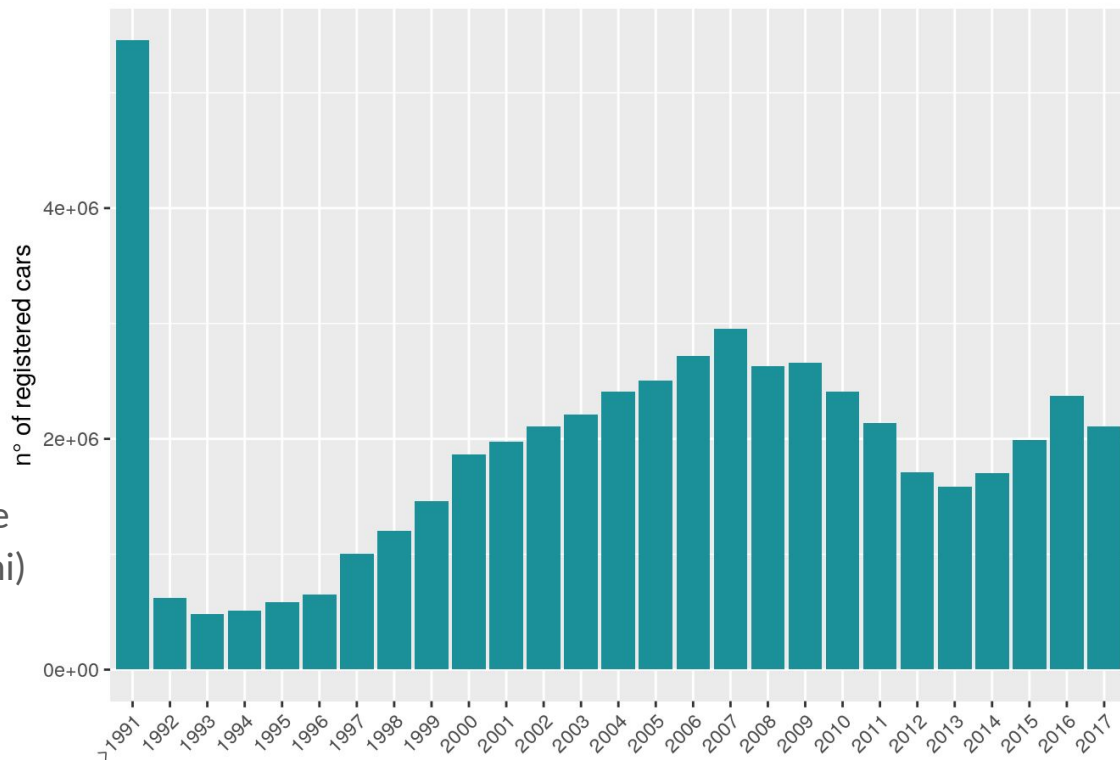
destinazione

-  AUTOCARRO PER TRASPORTO DI COSE
-  AUTOVETTURA PER TRASPORTO DI PERSONE
-  MOTOCICLO PER TRASPORTO PERSONE
-  OTHERS



Immatricolazioni per anno

Il parco macchine non sembra essere
giovannissimo (macchine con **25+** anni)

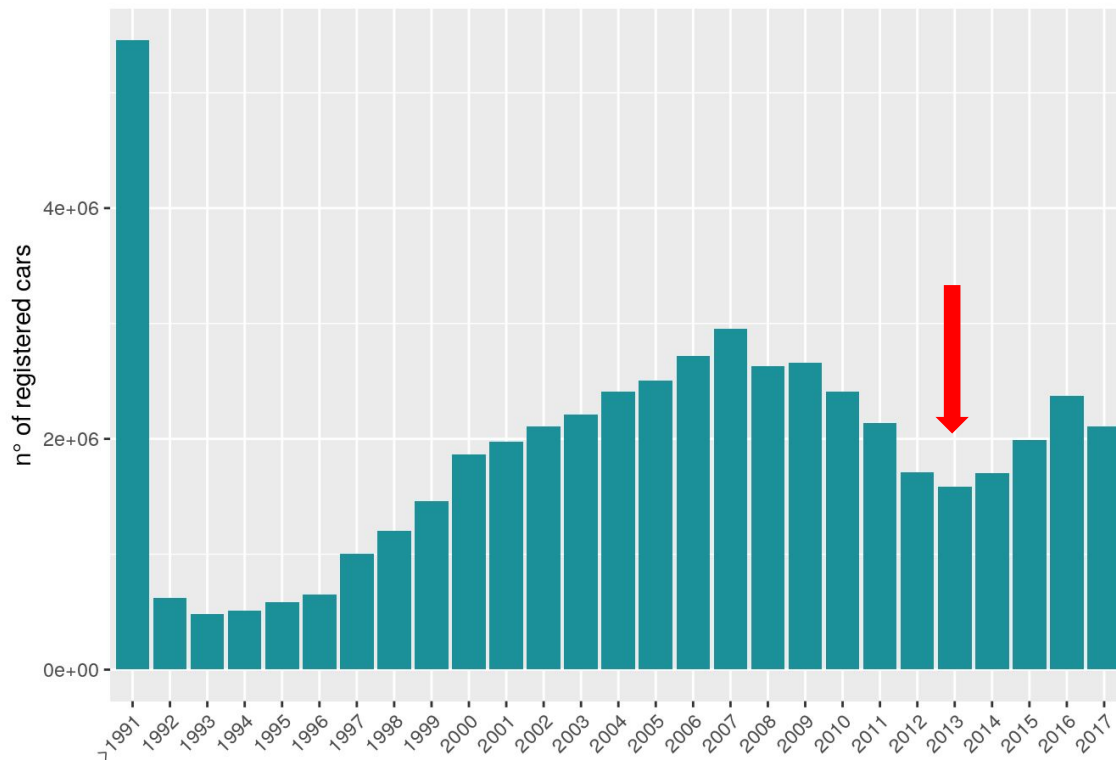




Immatricolazioni per anno

Il parco macchine non sembra essere giovanissimo (macchine con **25+** anni)

Possibile notare gli effetti della crisi economica

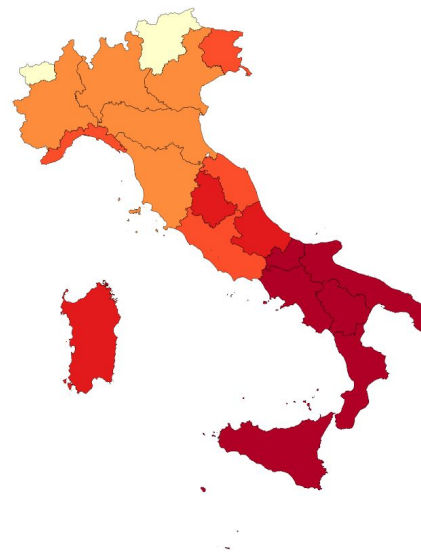
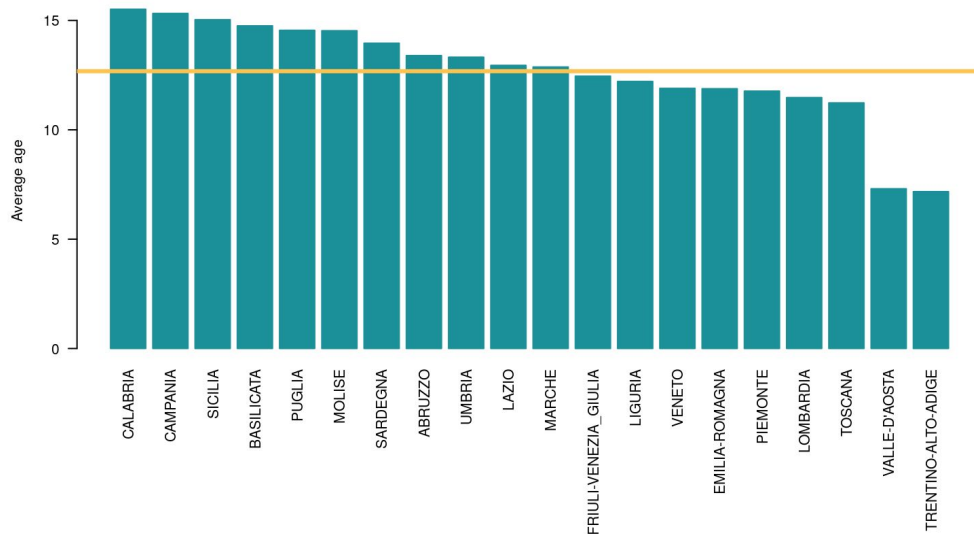




Età veicoli

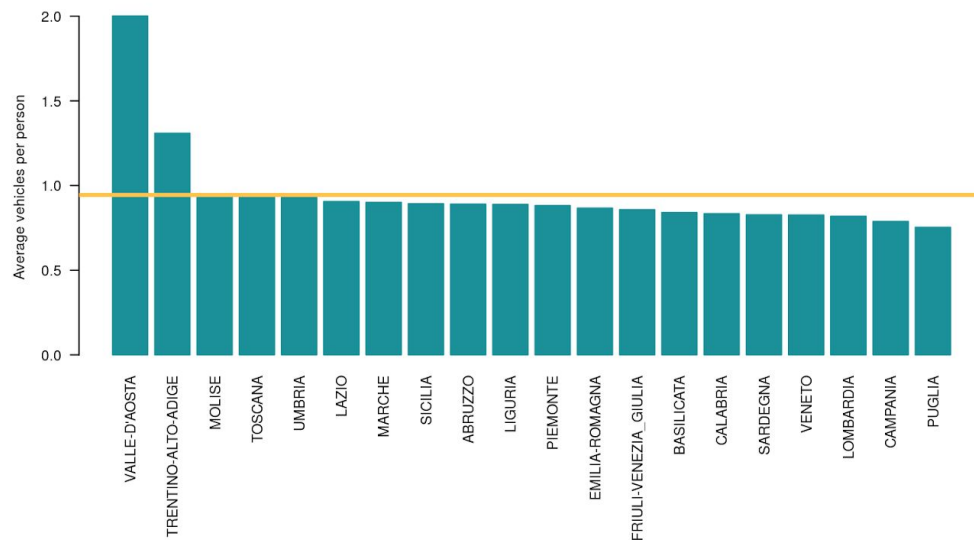
In **media**, l'età di un veicolo è circa **12.5 anni**

La mappa sembra ricalcare la **questione meridionale** e presenta **due outlier**





Numero di veicoli per abitanti



I dati normalizzati sono conformi alla media

Anche in questo caso i due outlier, la Valle d'Aosta sfiora i 2 veicoli per persona

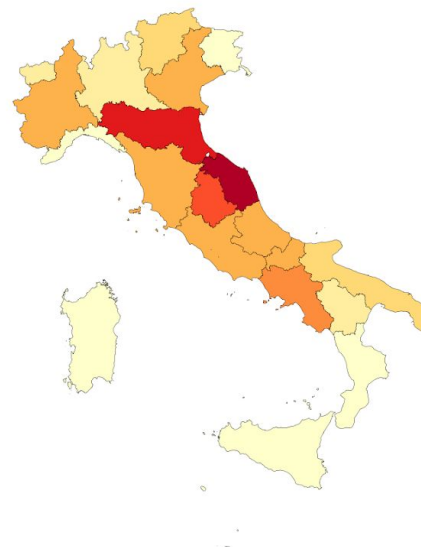
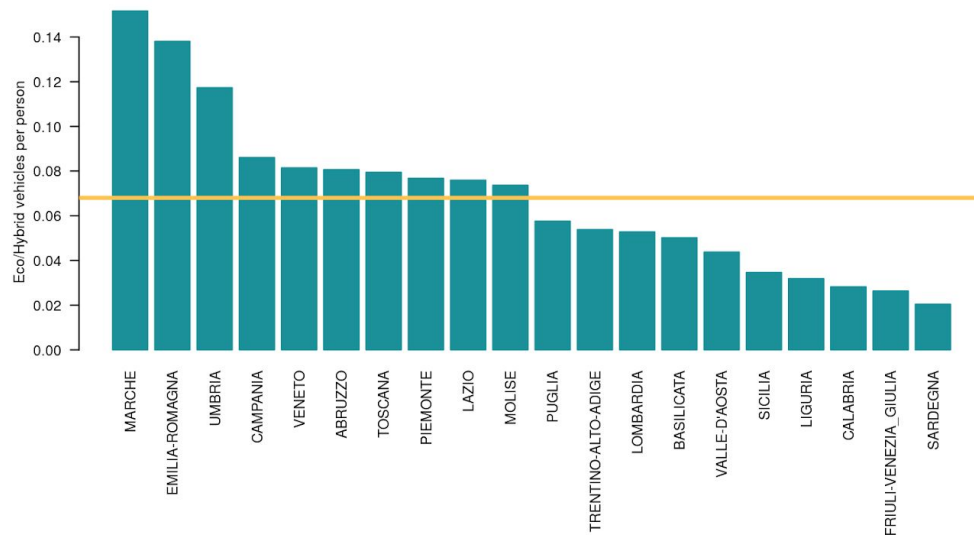




Veicoli eco-friendly

Concentrazione **normalizzata** di veicoli eco-friendly
(elettricità, gas, ibrido)

Il centro-nord Italia la fa da padrone

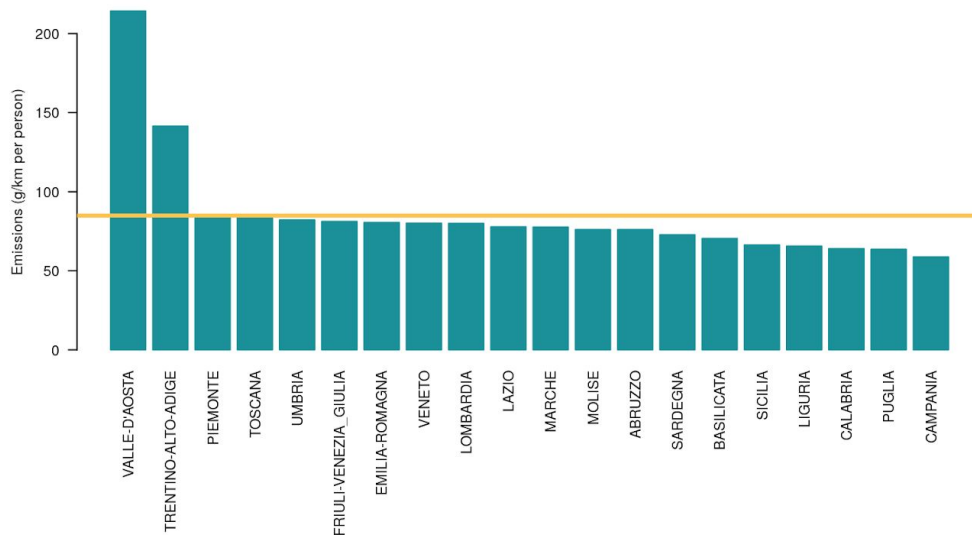




Emissioni CO₂

Le emissioni **normalizzate** di CO₂ sembrano proporzionali alla densità di veicoli per persona

Chiaramente i due outlier sono fortemente esposti



Immatricolazione veicoli nel tempo

È possibile notare le **evoluzioni** e i **trend** nel tempo

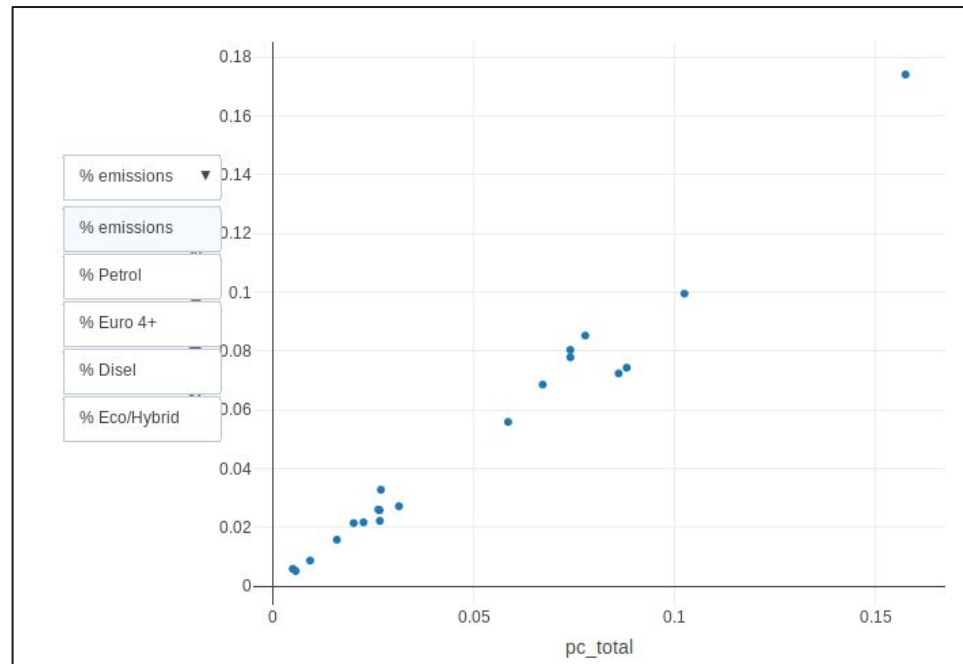
Le soluzioni eco-friendly **non** hanno una rilevanza nel panorama attuale



Esaminare le correlazioni

Tool interattivo realizzato con *plotly* che permette di **mettere in relazione** per ogni regione la percentuale di macchine sul totale con la percentuale di macchine che rispettano una condizione selezionabile.

Idealmente i punti dovrebbero stare sulla bisettrice



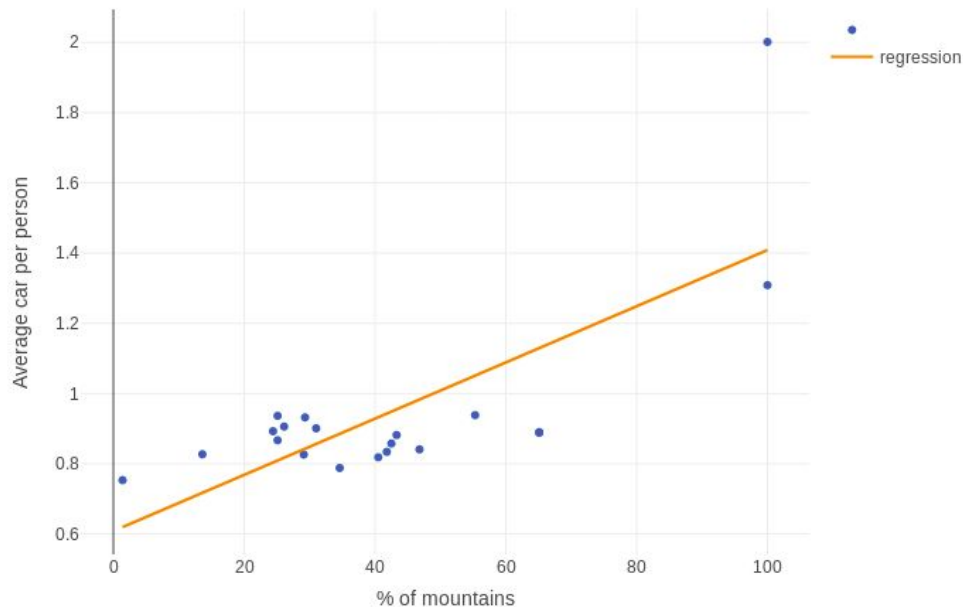
CURIOSITÀ

I problemi del Trentino Alto-Adige e della Valle d'Aosta

Problema di montuosità?

È la prima intuizione che potrebbe venire in mente considerando le particolarità delle regioni [2]

Di fatto ha un *tiepido riscontro*



[2] https://it.wikipedia.org/wiki/Zone_altimetriche_d%27Italia#Le_regioni_in_cifre



CURIOSITÀ

Il veicolo più antico



Immatricolato nel 1854

È una **motocicletta** e appartiene a un
nonnetto di **75 anni** che vive in
provincia di Parma





SFIDE

Gestire 52M di record

SQL + statistica

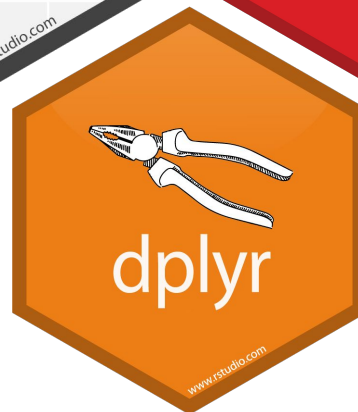
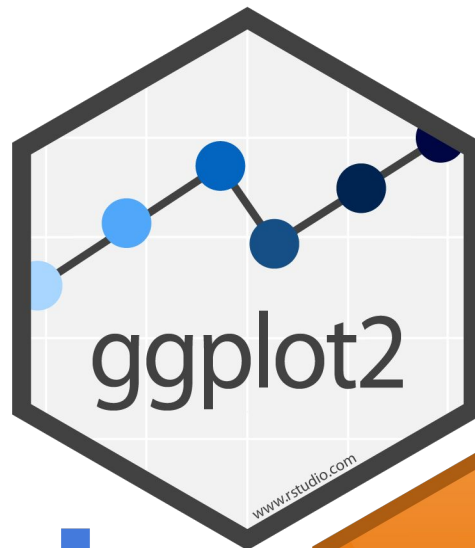
È possibile effettuare un **subsampling** del dataset prendendo record a caso

Facendolo per un numero abbastanza grande di volte, si otterrà un dataset che rispetta le proprietà statistiche di quello originale ma **più ridotto**

Il calcolo di una semplice query di *count* è passato da *4 minuti* a *10 secondi*



Tecnologie





Grazie per l'attenzione



Link al progetto

Data Management for Big Data | Gennaio 2019

Francesco Franchina