

StatCities 2026

Misurare le Città, progettare il futuro

Roma, 10 e 11 settembre 2026

Business Intelligence per le politiche urbane e locali

Il caso di Parma



Finalità del progetto

- Dotare il Comune di strumenti statistici moderni, potenziando la funzione statistica interna come strumento di *data visualization* e *data science*
- Offrire un supporto *data-driven* alle decisioni degli amministratori e alle politiche dell'Ente
- Promuovere la cultura del dato all'interno dell'organizzazione e offrire consulenza agli altri settori

Attività

- Integrazione delle basi dati interne del Comune in un *data warehouse*, quale fonte unica a cui rivolgere le richieste di accesso ai *dataset* di interesse (di competenza della società *in-house* del Comune)
- Ricerca e creazione di *dataset* da fonti esterne su temi d'interesse per il Comune
- Creazione di *routine* di analisi dati e statistica multivariata
- Creazione di report e *dashboard* pubblicati su una piattaforma *web*, riservati a un pubblico interno selezionato
- Disseminazione: corsi di aggiornamento, seminari ...

Parti coinvolte

- **Ufficio di Statistica:** Lara Berzieri, Massimiliano Ferrari
- **Settore Transizione Digitale** | S.O. Infrastrutture Tecnologiche e Comunicazioni
- **It.City S.p.A.** (società *in-house* del Comune di Parma): Piero Santovito, Giovanni Corradi
- **Centro interdipartimentale *Robust Statistics Academy* (RoSA), Università di Parma:** Aldo Corbellini, Fabrizio Laurini, Gianluca Morelli, Marco Riani
- **Destinatari interni:** decisori politici, dirigenti, posizioni organizzative (P.O.)



Storia del progetto



Stack tecnologico

Power BI Desktop e Service (pro/premium)

- creazione di *dashboard* interattive e report;
- pubblicazione in *cloud*.

Python

- *data pre-processing, cleaning*, trattamento di errori e anomalie, normalizzazione;
- *machine learning* (scikit-learn, PyTorch);
- *web automation* e *web scraping* (Playwright).

MATLAB

- EDA (analisi esplorativa dei dati);
- statistica multivariata;
- *outlier detection*;
- uso *toolbox* <Flexible Statistics and Data Analysis> sviluppato da RoSA (Analisi delle Corrispondenze, *Cluster Analysis*, *Forward Search*, PCA ...).

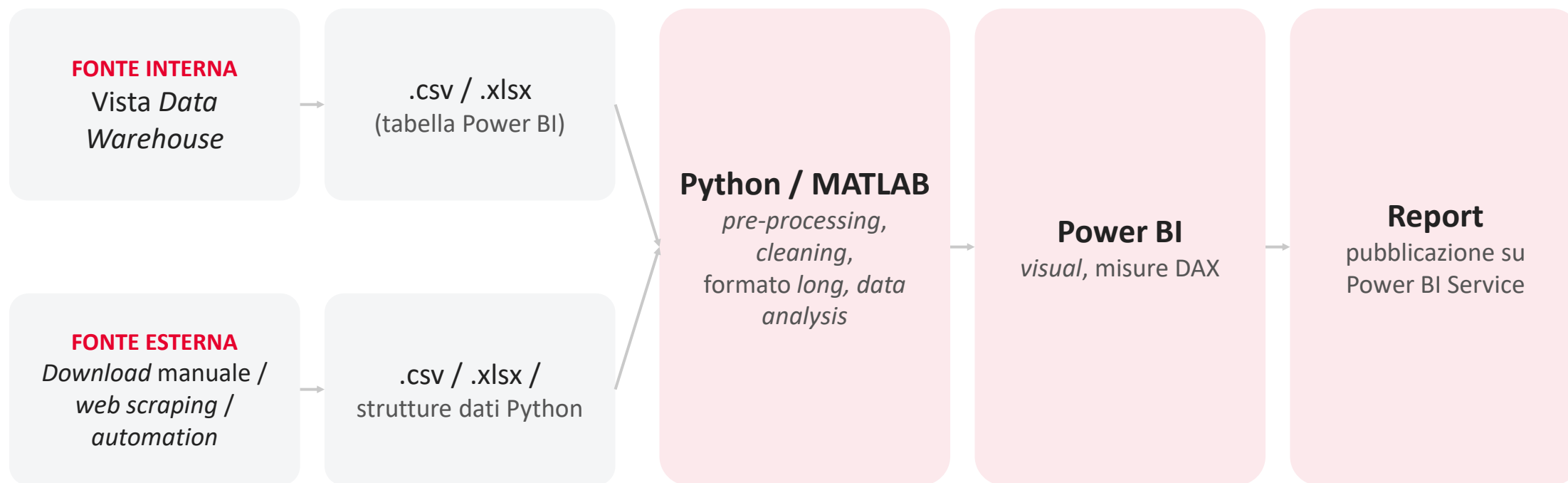
SQLite

- piccoli *database* locali.

Le fonti

- **Fonti esterne** istituzionali: ISPRA, Mef, Istat ...
- **Fonti interne:**
 - ❖ anagrafe,
 - ❖ catasto,
 - ❖ esercizi commerciali,
 - ❖ incidenti,
 - ❖ strutture ricettive
 - ❖ altri in fase di lavorazione ...
- **AIDA:** banca dati di Moody's contenente i dati di bilancio delle imprese italiane

Pipeline



- Si è scelto deliberatamente di non usare Power Query all'interno di Power BI per le trasformazioni, svolte invece in Python: anche quando l'*input* è già una tabella Power BI, si passa comunque a un altro ambiente di sviluppo (apparente controsenso).
- Questo implica alcuni passaggi manuali nella *pipeline*; tuttavia, in questo contesto, non sono presenti report in tempo reale e gli aggiornamenti hanno una periodicità bassa.
- Si ritiene che i benefici e le potenzialità offerte da Matlab/Python in termini di efficienza e qualità dell'analisi superino i costi della gestione semi-manuale in un contesto di indicatori con frequenza di aggiornamento bassa.
- Non è escluso che si adottino soluzioni più efficienti in futuro.

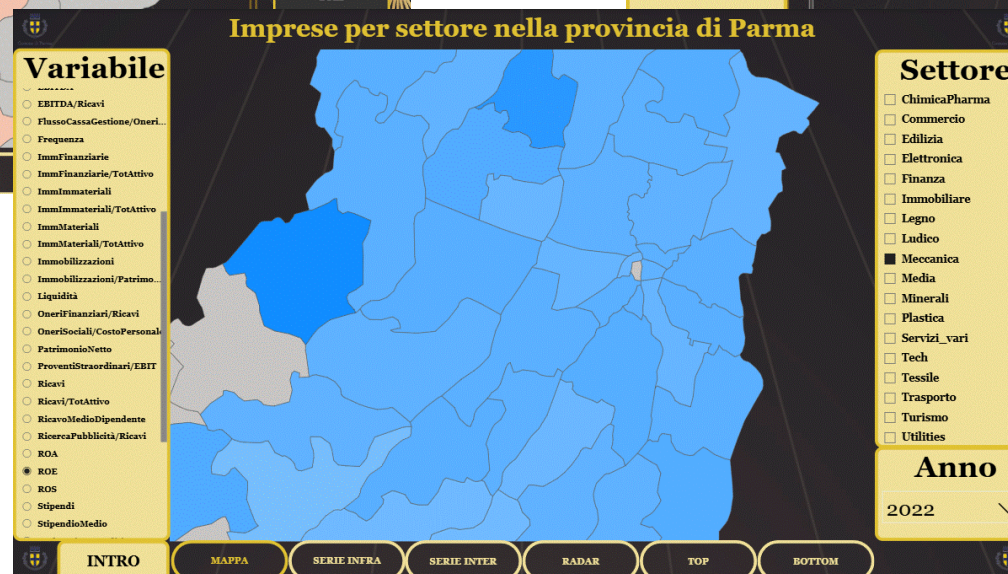
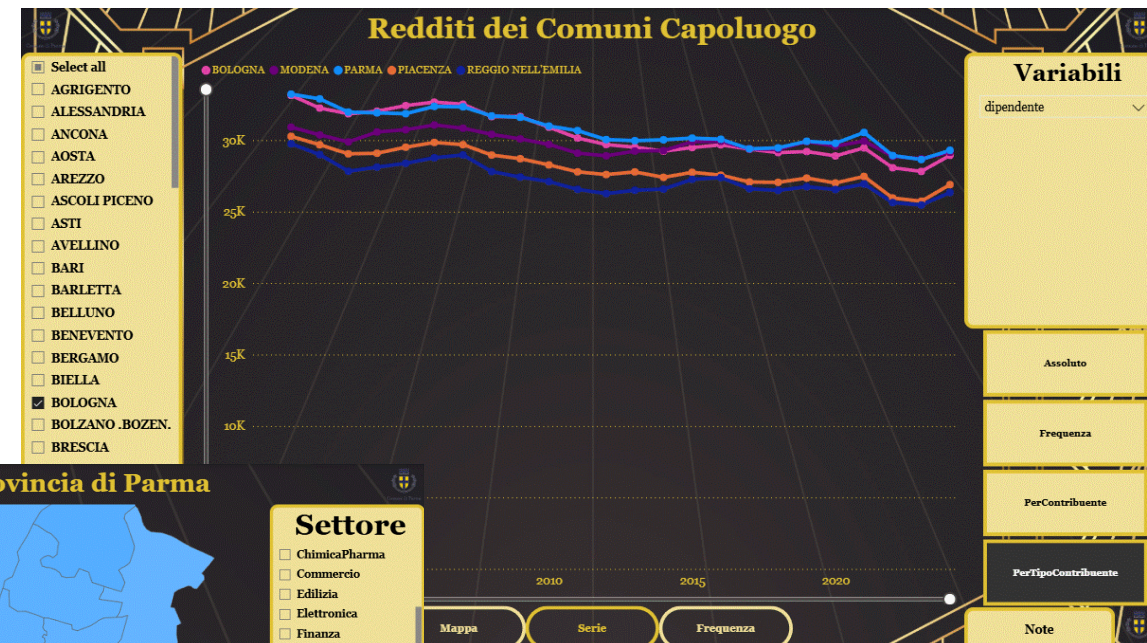
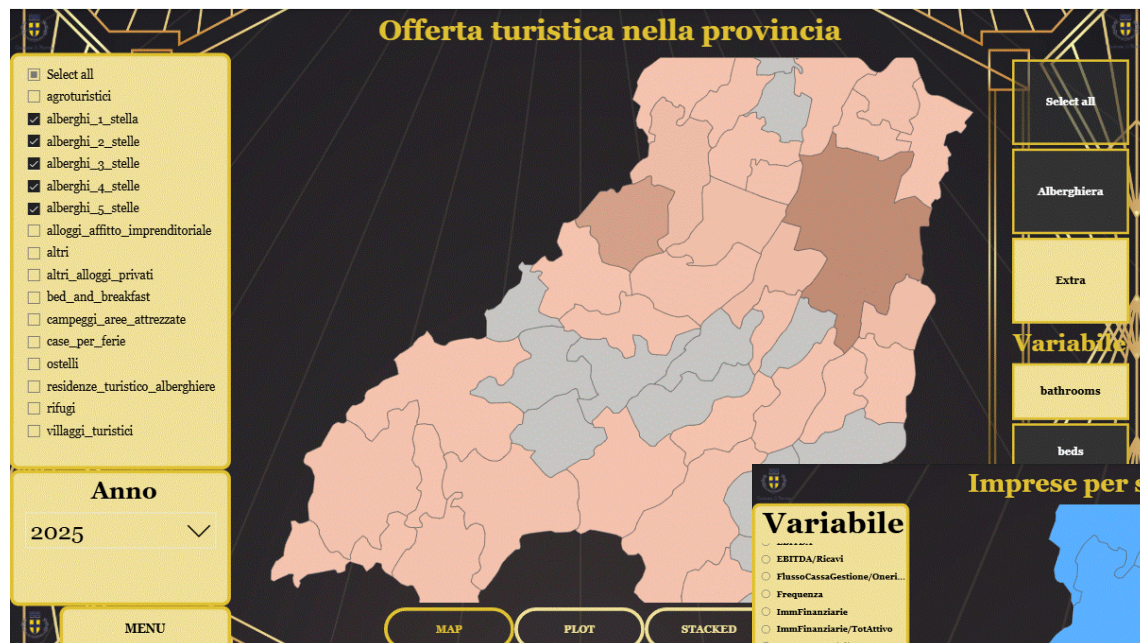
Lo stato del progetto

Report pubblicati, per ambito tematico:

- Demografia · *esterna*
- Residenti · *esterna*
- Previsioni demografiche per quartiere · *RoSA*
- Disagio subcomunale · *esterna*
- Emissioni · *esterna*
- Rifiuti · *esterna*
- Turismo · *esterna*
- Veicoli · *esterna*
- Delitti denunciati · *esterna*
- Qualità dell'aria · *esterna*
- Redditi · *esterna*
- Incidenti · *interna*
- Dati di bilancio delle imprese della provincia · *esterna*

IN CORSO Strutture ricettive · dati interni

Report



Criticità (1/2) — dataset e Power BI

Dataset esterni

- Sparsi, disponibili in singoli file (assenza di API)
- Assenza di microdati
- Livello di aggregazione elevato

Dataset interni

- Processo di raccolta lento e complesso: pluralità storica dei gestionali, digitalizzazione di procedure cartacee, interazioni con fornitori esterni
- Errori e valori mancanti che, oltre una certa soglia, rendono vana l'analisi (es. *dataset* sugli esercizi commerciali) -> il progetto auspicabilmente fungerà da catalizzatore per la *Data Governance* aziendale stimolando un miglioramento nell'attenzione alla qualità dei dati proprietari nel medio-lungo periodo laddove nel breve non è giudicata sufficiente
- Poche variabili quantitative
- Storicità limitata

Power BI

- Difficoltà a integrare *routine* di analisi complesse in Python/R, oltre ai semplici *visual*
- Difficoltà di *debug*
- Necessità frequente di pre-caricare i risultati
- Vincoli sull'ambiente *cloud* (pacchetti ammessi e relative versioni) e *timeout* nei tempi di esecuzione -> motivo per cui l'architettura prevede in genere il pre-caricamento dei dati elaborati (*long format*), usando Power BI principalmente per la visualizzazione e la navigazione dinamica (DAX), lasciando la fase intensiva di calcolo negli ambienti nativi (Python/Matlab)
- Inefficienza nel configurare ripetutamente elementi grafici, *bookmark*, ecc.

Criticità (2/2) — manutenzione

- Mantenere ha la stessa importanza di creare.
- I dettagli si dimenticano in fretta: perché si è trattato in un certo modo un *edge case*? Perché si sono scelte determinate trasformazioni?
- Un soggetto esterno, o lo stesso autore in futuro, deve poter comprendere e far funzionare il prodotto
- Concepire fin dall'inizio report e *script* che permettano un aggiornamento agevole, idealmente automatizzato
- Documentare a un certo punto, ove possibile, il lavoro svolto in passato
- Adottare il *version control* quando si lavora in *team*

Prospettive

- Circoscrivere l'insieme dei fruitori -> acquistare le licenze necessarie -> procedere al *deployment* effettivo dei report
- Creare nuovi report in base ai *dataset* che saranno disponibili nel *data warehouse* o da fonti esterne su temi di interesse
- Aumentare il grado di automazione nel reperimento e nella trasformazione dei dati, per semplificare il mantenimento dei report nel lungo termine
- Svolgere analisi più complesse laddove la quali-quantità dei dati lo permetta
- Accordo Attuativo della Convenzione Quadro Ufficio Statistica – Dipartimento Scienze Economiche UNIPR per una collaborazione strutturata
- Focus sul tema mobilità con i dati resi disponibili dalle società partecipate dal Comune grazie al lavoro del dottorando in Statistica C. Castrogiovanni