

Progetto d'Esame – Laboratorio di Big Data, Data Mining e Data Analytics

Titolo del progetto: Analisi e correlazione tra inquinamento atmosferico e ricoveri ospedalieri per patologie respiratorie

Obiettivo del progetto

Lo scopo di questo lavoro è analizzare i dati relativi alla qualità dell'aria e confrontarli con i ricoveri ospedalieri per problematiche di natura respiratoria, al fine di individuare possibili correlazioni tra i picchi di inquinamento e l'incidenza di tali ricoveri.

L'idea di fondo è quella di esplorare i dataset disponibili, comprendere la struttura e le informazioni in essi contenute, e successivamente applicare tecniche di analisi e correlazione. In particolare, ci si concentra sulla ricerca di pattern ricorrenti e sull'individuazione di possibili relazioni tra i principali inquinanti atmosferici e i ricoveri ospedalieri filtrati per cause respiratorie.

Dataset utilizzati

- **Qualità dell'aria – Comune di Milano**
 - File: *qaria_gennaio_marzo_2025.csv*
 - Fonte: [Comune di Milano – Dati aperti](#)
 - Descrizione: contiene i dati giornalieri relativi alla qualità dell'aria a Milano, redatti da AMAT sulla base delle rilevazioni effettuate da ARPA Lombardia nelle stazioni istituzionali presenti sul territorio comunale.
- **Ricoveri ospedalieri per cause respiratorie – Milano**
 - File: *ricoveri_ospedalieri_Milano_gennaio_marzo_2025.csv*
 - Fonte: dataset simulato.
 - Motivazione: i dati ufficiali disponibili, come le Schede di Dimissione Ospedaliera (SDO) pubblicate annualmente dal Ministero della Salute ([link al documento PDF](#)), non sono rilasciati in formato elaborabile e non consentono un'estrazione diretta delle informazioni necessarie. Per questa ragione è stato generato un dataset ad hoc, ispirato ai dati reali, che permettesse di svolgere le analisi richieste.
- **Limiti normativi degli inquinanti**
 - File: *limiti_inquinanti_italia.csv*
 - Fonte: [Testo Unico Sicurezza – Tabella parametri di valutazione della qualità dell'aria](#)
 - Descrizione: contiene i valori soglia previsti dalla normativa italiana per i principali inquinanti atmosferici, utili come benchmark di riferimento per l'analisi.

Parametri monitorati

I principali inquinanti atmosferici considerati sono:

- **C6H6 (Benzene):** concentrazione media sulle 24 ore.
- **CO (Monossido di carbonio):** concentrazione massima media mobile su 8 ore (in mg/m³).
- **NO2 (Biossido di azoto):** concentrazione massima oraria sulle 24 ore.
- **O3 (Ozono):** concentrazione massima oraria sulle 24 ore.

- **PM10:** concentrazione media sulle 24 ore (polveri con diametro aerodinamico $< 10 \mu\text{m}$).
- **PM2.5:** concentrazione media sulle 24 ore (polveri con diametro aerodinamico $< 2,5 \mu\text{m}$).
- **SO2 (Biossido di zolfo):** concentrazione massima oraria.

Aspetti innovativi e finalità

L'approccio proposto consente di:

1. **Integrare fonti di dati eterogenee** (dati ambientali, dati sanitari e limiti normativi).
2. **Valutare potenziali correlazioni** tra l'andamento degli inquinanti atmosferici e i ricoveri ospedalieri per patologie respiratorie.
3. **Confrontare i dati reali con i limiti di legge** sugli inquinanti, per verificare se e in quali casi i picchi di inquinamento superano i valori consentiti.
4. **Simulare scenari di rischio** e fornire un contributo alla riflessione su tematiche di salute pubblica e politiche ambientali.