



***Università degli Studi della Basilicata***  
***Relazione Gruppo 1***  
***Progetto Big Data a.a 2021/2022***

*Antonio Rinaldi, 64767*

*Aurelia Santarsiere, 65124*

*Carminé Martinelli, 63795*

*Mario Di Lonardo, 66225*

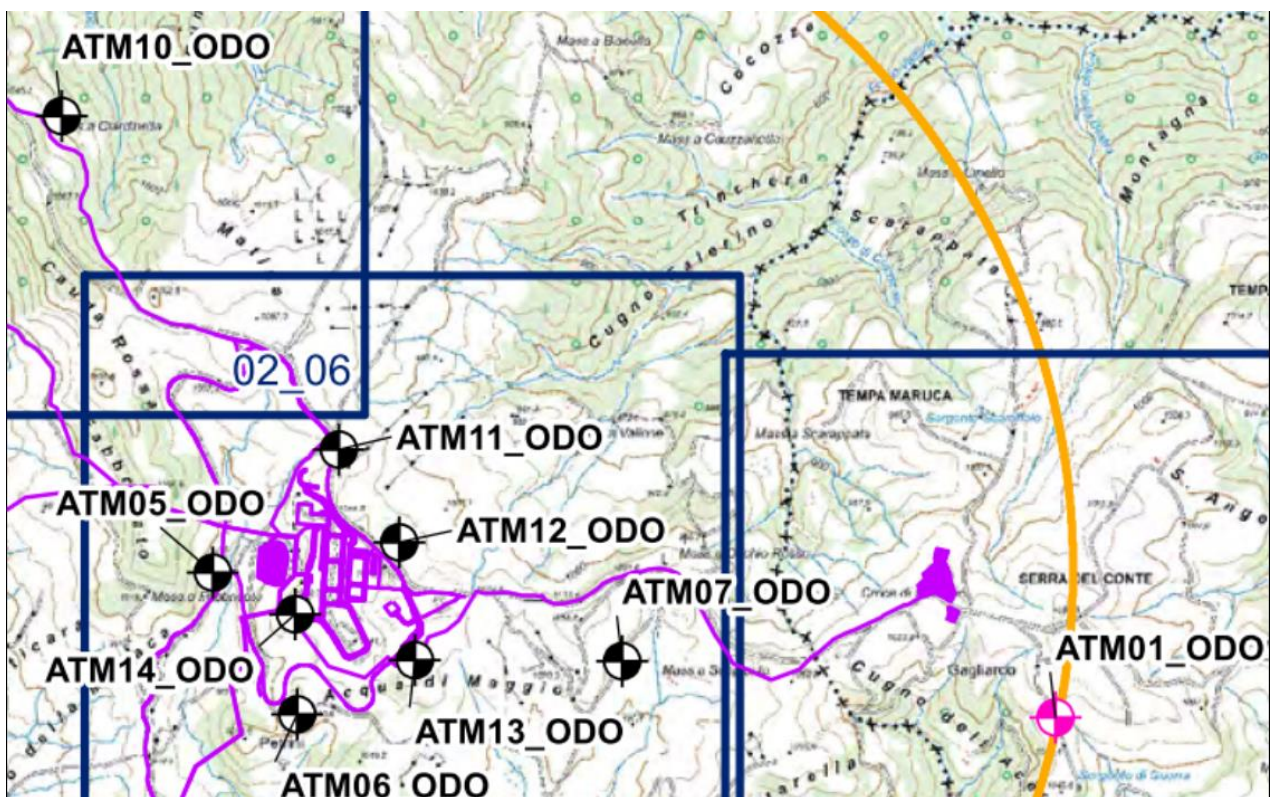
## Sommario

<b><i>INTRODUZIONE</i></b>	<b>3</b>
<b><i>STRUMENTI UTILIZZATI</i></b>	<b>3</b>
<b><i>QUERY PREFISSATE</i></b>	<b>4</b>
<b><i>ANALISI VISUALI</i></b>	<b>6</b>
• <i>ANDAMENTO DEI VARI SENSORI</i>	6
• <i>MEDIA ORARIA</i>	7
• <i>MEDIA MOBILE A 60 MINUTI</i>	7
• <i>MEDIA MOBILE A 120 MINUTI</i>	8
<b><i>CORRELAZIONI</i></b>	<b>8</b>
<b><i>ACCORGIMENTI!</i></b>	<b>13</b>
<b><i>ANALISI E RACCOMANDAZIONI</i></b>	<b>13</b>
<b><i>SOGLIE DI RISCHIO E INQUINANTI</i></b>	<b>13</b>
<b><i>INQUADRAMENTO TERRITORIALE E RISULTATI</i></b>	<b>15</b>

## Introduzione

Il progetto dell'a.a.2021/22 riguarda l'analisi di dati raccolti nel sito di estrazione di Tempa Rossa nella zona di Corleto Perticara. I dati provengono da sensori o nati elettronici posti nelle vicinanze dello stabilimento che hanno il compito di misurare la quantità di inquinanti nell'aria (benzene, acido solfidrico, composti organici). Il nostro gruppo si è concentrato sui dati dei sensori delle stazioni di monitoraggio: ATM05, ATM07, ATM10, ATM14.

Quest'analisi ha il fine di valutare se le emissioni dei composti chimici in questione potrebbero risultare cancerogeni per chi lavora all'interno del sito ma anche per chi vive nelle vicinanze (allevamenti, persone). L'analisi ha anche il fine di quantificare se gli odori prodotti durante l'estrazione potrebbero essere percepiti nelle zone limitrofe ed essere fastidiosi.



## Strumenti utilizzati

Gli strumenti utilizzati per condurre lo studio sono stati:

1. **Fogli di calcolo** (Excel);
2. **Tableau**: utilizzato per l'analisi visuale;
3. **Python** (Google Colab, Jupiter Notebook): impiegato per implementare e commentare il codice.
  - Librerie di Python utilizzate: **NumPy**, **Pandas**.

## Query prefissate

Prima di svolgere le query è stata effettuata una pulizia del Dataset al fine di migliorarne l'usabilità, consentendoci di lavorare con dati migliori di quelli ricevuti.

### Query effettuate:

Inizialmente è stato creato un dataframe **df\_DatiSensori** direttamente dal file excel '**Dati\_gruppo1.xlsx**'. Successivamente è stato fatto un dataframe-copia: **df\_Dati** su cui svolgere le operazioni di pulizia.

Il primo step è stato rinominare le colonne. In seguito sono state eseguite operazioni per il riconoscimento dei fallimenti, ossia valori nulli o '**ND**' riportati nel file excel. Sono stati considerati fallimenti anche più di 24 zeri consecutivi dati dai sensori, quindi è stato effettuato un reset degli indici del dataframe per evitare salti tra gli indici.

- **Le 100 registrazioni con il maggiore livello di benzene**

Sono state estratte dal dataframe le colonne postazione, data e C6H6. In seguito, rimossi i duplicati, i dati sono stati ordinati in modo decrescente per prendere i primi 100 valori del benzene più alti.

- **Le 100 registrazioni con il maggior livello di acido solfidrico(H2S)**

In questo caso le colonne estratte dal dataframe utili per la query sono state: postazione, data e H2S. Dal dataframe ottenuto sono stati rimossi i duplicati e lo si è ordinato in modo decrescente. A quel punto sono stati presi i primi 100 valori di acido solfidrico.

- **Le 100 registrazioni con il maggior livello di acido solfidrico(H2SJ)**

Le operazioni effettuate sono analoghe a quelle della query precedente, con la differenza che l'ultima colonna estratta è stata quella corrispondente al sensore H2SJ.

- **Le 100 registrazioni con il maggior livello di VOC (sensore VOC)**

Analogamente alle query precedenti anche per il sensore VOC sono state estratte le colonne postazione e data oltre alla colonna corrispondente al sensore d'interesse. Rimossi i duplicati e ordinato in ordine decrescente il DataFrame risultante, sono stati presi i primi 100 valori.

- **Le 100 registrazioni con il maggior livello di VOC (sensore PIDVOC)**

Sono state prelevate dal dataframe le colonne postazione, data e PIDVOC. Rimossi i duplicati e ordinato in ordine decrescente il DataFrame risultante, anche qui, sono stati presi i primi 100 valori.

- **Le 50 ore con il livello medio più alto/basso di benzene**

Per questa query è stata realizzata una funzione che consente di eliminare i valori nulli, calcolare la media oraria e restituire un dataframe ordinato sull'ora. Questo DataFrame

contiene solo le colonne postazione, data e il sensore interessato. Dopodiché si è pensato di creare un dataframe richiamando la funzione **media\_oraria**, utilizzando i parametri **df\_DatiPuliti** e il sensore interessato, ovvero il benzene. Infine è stata calcolata la più alta e la più bassa media oraria di benzene. Sono stati mostrati solo i primi 50 valori, modificando solo il parametro **ascending** della funzione **sort\_values**.

- **Le 50 ore con il livello medio più alto/basso di H2S**

Sono state effettuate le stesse operazioni della query precedente, ma per il sensore H2S.

- **Le 50 ore con il livello medio più alto/basso di H2SJ**

Sono state effettuate le stesse operazioni della query precedente, ma per il sensore H2SJ.

- **Le 50 ore con il livello medio più alto/basso di VOC**

Sono state effettuate le stesse operazioni della query precedente, ma per il sensore VOC.

- **Le 50 ore con il livello medio più alto/basso di PIDVOC**

Sono state effettuate le stesse operazioni della query precedente, ma per il sensore PIDVOC.

- **Le 3 giornate con il maggior numero di fallimenti nell'invio**

Per questa query è stato creato un dataframe **df\_Fail\_ND** che ha come colonne la Data e i valori numerici dei sensori. E' stata creata una nuova colonna settando il tipo data a 'anno/mese/giorno'. Poi è stato effettuato il conteggio dei valori NaN sulla colonna **TRS\_ppb** (il conteggio risulta identico su qualunque altra colonna poiché il fallimento di una stazione coincide con il fallimento di invio dei dati di ogni sensore) facendo il groupby sul '**Giorno**', sommando e infine resettando gli indici. L'ultima operazione ha prodotto anche una nuova colonna '**count**'. E' stato effettuato, quindi, un ordinamento decrescente sulla colonna '**count**'. Infine sono state prese in considerazione solo le prime 3 righe.

- **Le 3 giornate con il minor numero di fallimenti nell'invio**

Per questa query sono state effettuate le stesse operazioni di quella precedente andando a modificare accuratamente i parametri interessati.

- **Il numero medio di fallimenti nell'invio per sensore**

Per questa query sono state create due liste, **inquinanti** e **stato\_inquinanti**, da usare nella funzione successiva. È stato sviluppato un metodo che permette di contare la somma dei fallimenti per ogni sensore andando poi ad inserirli in un nuovo dataframe **df\_NFails**. Innanzitutto è stata creata una lista di fallimenti, **fails**, e poi inizializzato il nuovo dataframe **df\_NFails** a due colonne titolate: **Sensori** e **Fallimenti**. Viene effettuato un ciclo sulla lunghezza di una delle due liste (quale è indifferente, avendo entrambe la stessa lunghezza). Nel ciclo vi è la creazione di un dataframe temporaneo contenente solo le colonne **inquinante** e **stato\_inquinante**. La variabile **errors** restituisce il numero di

fallimenti dovuti alla presenza di zero consecutivi (come fatto per la pulizia del dataset, abbiamo scelto un valore di 24 zeri consecutivi che equivalgono a 2 ore di invii di dati considerati falliti), mentre **num\_ND** restituisce il numero di valori **ND** presenti. Infine questi valori vengono sommati e aggiunti alla lista **fails**. Fuori dal ciclo for viene popolato il dataframe **df\_NFails** con la colonna sensori uguagliata alla lista inquinanti e la colonna fallimenti con **fails**. Viene aggiunta una colonna '**Dati TOT**' che comprende la lunghezza dei dati per ogni sensore ed infine viene creata un'ultima colonna chiamata '**Media**' contenente la media tra la colonna '**Fallimenti**' e '**Dati TOT**'.

- **Il sensore con il massimo numero di fallimenti**

Per questa query è stato utilizzato lo stesso dataframe creato nella query precedente. E' stata fatta una chiamata al metodo **max()** sulla colonna '**Fallimenti**' del dataframe **df\_NFails**.

- **Il sensore con il numero minimo di fallimenti**

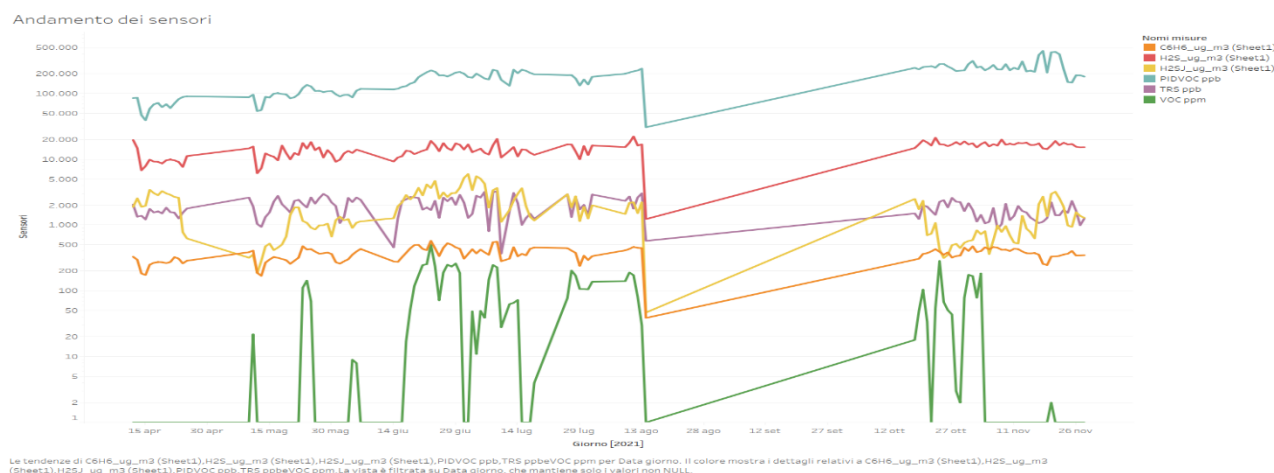
Per questa query è stato utilizzato lo stesso dataframe creato nella query precedente ed è stata fatta una chiamata al metodo **min()** sulla colonna '**Fallimenti**' del dataframe **df\_NFails**.

## Analisi visuali

Nella fase successiva abbiamo proseguito con l'analisi visuale dei dati tramite l'utilizzo di Tableau; in particolare calcolando l'andamento dei sensori, la media oraria e quella mobile a 60 e 120 minuti.

### Analisi visuali effettuate:

- **Andamento dei vari sensori**

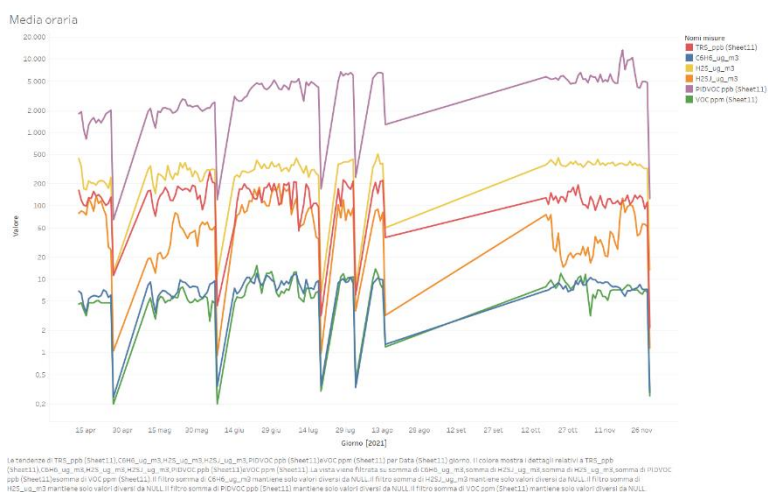


Il grafico, di cui sopra, mostra l'andamento dei valori registrati per ogni sensore. E' possibile notare come per tutti i sensori si abbia un calo nelle rilevazioni, in prossimità del 13 agosto. Si può osservare che l'andamento del sensore VOC appare alquanto altalenante a causa del verificarsi di ripetuti malfunzionamenti dovuti o alla mancata



ricezione dei dati (a causa di un blocco del server di acquisizione) o al mancato invio dei dati causato dall'assenza di connessione.

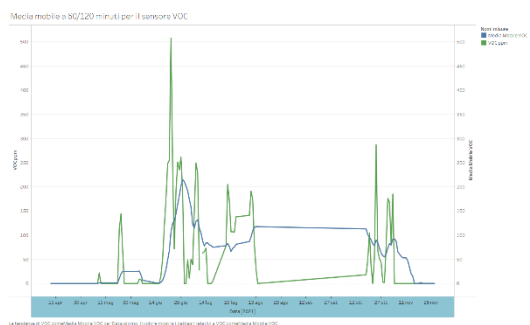
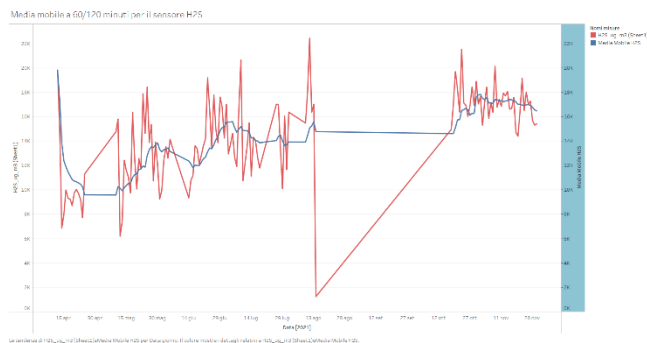
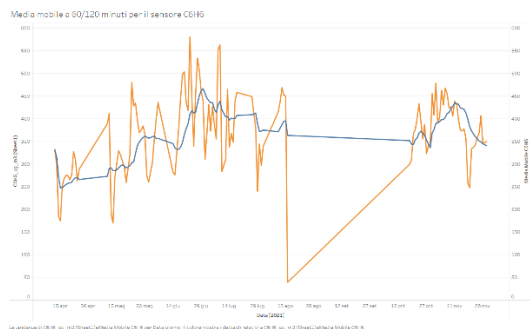
- **Media oraria**



Per la media oraria, come possiamo notare dal grafico, abbiamo dei periodi in cui i sensori vanno in down di rilevazioni. Sono quattro: 15-30 APR, 30-14 MAG-GIU, 14-29 LUG, 29-13 LUG-AGO.

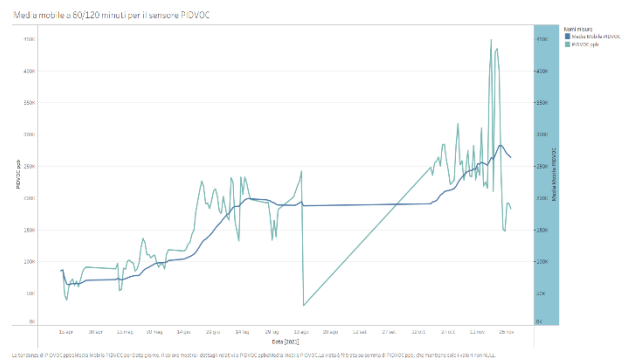
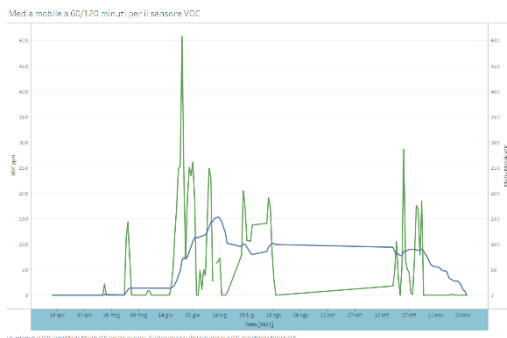
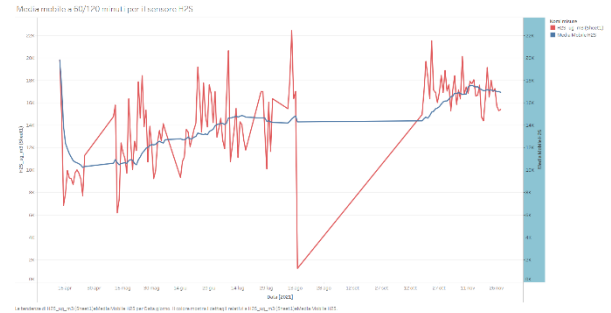
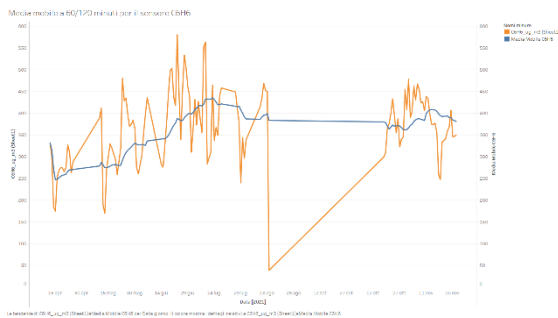
- **Media mobile a 60 minuti**

In arancio quella del sensore C6H6, in rosso per H2S, in verde e celeste rispettivamente per i sensori VOC e PIDVOC



- **Media mobile a 120 minuti**

*In arancio quella del sensore C6H6, in rosso per H2S, in verde e celeste rispettivamente per i sensori VOC e PIDVOC*



Come è possibile vedere dai grafici della media mobile a 60 e a 120 minuti. All'aumentare della finestra oraria presa in considerazione, questa tende ad attenuarsi. Quindi maggiore sarà il periodo in minuti calcolato e maggiore sarà l'attenuazione della media mobile.

## Correlazioni

Nella terza fase dello studio ci siamo occupati delle correlazioni, ossia relazioni tra due valori  $x$  e  $y$ . Le correlazioni possono essere positive, se all'aumentare (diminuire) della  $x$  aumenta (diminuisce) anche la  $y$ , o negative se all'aumentare (diminuire) della  $x$ , diminuisce (aumenta) la  $y$ . Una correlazione è tanto forte quanto più il suo valore si avvicina ad 1.

### Correlazioni effettuate:

Inizialmente è stata creata una funzione correlazione **corr\_Sensori** che permette di calcolare tre coefficienti di correlazione: quello di **Pearson**, quello di **Spearman** e quello di **Kendall**. Questa funzione prende in input il dataframe su cui effettuare le correlazioni, i due sensori da mettere in correlazione e la postazione come stringa. Quest'ultima è settata come **None** (poiché questa funzione è stata adattata per essere utilizzata per due tipi di correlazioni diverse). La funzione inizialmente verifica che il valore di **postazione** è diverso da **None**, così da poter creare un dataframe **df1** prendendo solo la colonna



postazione interessata. Altrimenti fa una copia dell'intero dataframe che è stato passato. Dopo la funzione effettua la regressione lineare, grafica i valori e calcola i coefficienti di correlazione.

- **Correlazione H2S e H2SJ in una data stazione**

Nella prima correlazione avviene la creazione di un dataframe `df_AcidoSolf` prendendo dal dataframe `df_DatiPuliti` le colonne di **postazione**, **Data**, **H2S\_ug\_m3** e **H2SJ\_ug\_m3**. Viene effettuata l'eliminazione dei valori nulli attraverso la funzione `dropna()` e infine richiamata la funzione `corr_Sensori` passandole il dataframe `df_AcidoSolf`, **H2S\_ug\_m3**, **H2SJ\_ug\_m3** e la **postazione** (viene effettuata per tutte le stazioni di monitoraggio). I valori restituiti dalla funzione vengono associati alle variabili **pearson**, **spearman** e **kendall** e poi stampati a video.

La correlazione tra questi due sensori, in tutte le stazioni, ha valori vicini allo 0 quindi risulta **molto debole**.

- **Correlazione VOC e PIDVOC in una data stazione**

Per questa correlazione sono state effettuate le stesse operazioni della correlazione precedente andando ad effettuare le opportune modifiche per i composti interessati, **VOC** e **PIDVOC**.

La correlazione calcolata è stata classificabile come **molto debole**, ad eccezione però dell'ATM10 dove la correlazione risulta **debole**.

- **Correlazione TRS e H2S in una data stazione**

Per il calcolo di questa correlazione sono state effettuate operazioni analoghe a quelle precedentemente descritte andando a modificare opportunamente i campi relativi ai sensori inserendo: **TRS** e **H2S**.

In tutte le stazioni le correlazioni sono risultate **molto deboli**.

- **Correlazione TRS e H2SJ in una data stazione**

Le stesse operazioni delle correlazioni precedenti sono state eseguite per quest'ultima correlazione, modificando però i campi relativi ai sensori inserendo: **TRS** e **H2SJ**.

Anche la correlazione che intercorre tra i valori di questi sensori risulta **molto debole** in tutte le stazioni. Per quanto riguarda il TRS nella stazione ATM14 non è possibile effettuare correlazione a causa della mancanza di dati.

- **Correlazione VOC e C6H6 in una data stazione**

Per questa correlazione sono state effettuate le stesse operazioni delle correlazioni precedenti andando ad effettuare le opportune modifiche sui sensori. Questa volta sono stati usati **VOC** e **C6H6**.

Tra i valori di registrati da questi due sensori le correlazioni risultano **molto deboli** in tre stazioni, ad eccezione dell'ATM10 di tipo **moderato** con Pearson, e molto deboli per Kendall e Spearman.

- **Correlazione PIDVOC e C6H6 in una data stazione**

Per questa correlazione sono state effettuate le stesse operazioni delle correlazioni precedenti andando ad effettuare le opportune modifiche per usare i composti, **PIDVOC** e **C6H6**.

Qui accade già qualcosa di interessante, nelle stazioni ATM05, ATM07 le correlazioni sono **molto deboli**, nella ATM10 la correlazione è **moderata**, mentre per la stazione ATM14 la correlazione è **forte** (varia dallo 0.67 di Kendall allo 0.84 di Spearman).

- **Per ogni coppia di stazioni:**
  - **Correlazione valori stesso sensore, stazioni diverse**

Per questa correlazione è stata sviluppata una funzione che permette di creare e restituire un dataframe costituito dalle colonne necessarie. La funzione **creazione\_df** prende in input un dataframe **df**, le colonne da usare per la correlazione ed un composto. La funzione presenta anche un parametro di default, **df2 = None**, che rende riutilizzabile la funzione anche per altre correlazioni. Viene inizializzato il primo dataframe temporaneo **df\_1** prendendo le righe con postazione uguale al valore di **colonna1**. Quindi vengono eliminati i valori nan. Dopo viene fatta una proiezione sulla colonna composto e un reset degli indici. Infine si trasforma la serie ottenuta in un dataframe. Viene effettuato il controllo su **df2**, per verificare se è stato passato come parametro. Se non è stato passato, si copia al suo interno il df1. Viene creato il dataframe **df\_Full** concatenando i due dataframe. Si rinominano le colonne di quest'ultimo con **colonna1** e **colonna2**. Si sostituiscono i valori **NaN** con 0. Per effettuare la correlazione viene richiamata la funzione **corr\_Sensori** sul df restituito dalla funzione **creazione\_df** (a cui sono stati passati **colonna1** e **colonna2**). Si visualizzano le correlazioni di **Pearson** e **Spearman**. Questa operazione viene effettuata per tutte le combinazioni di sensori e stazioni.

#### **Postazioni ATM05 e ATM07**

Tutti i composti hanno correlazioni **molto deboli**. Solo per benzene (C6H6) la correlazione è **debole**. Per il PIDVOC con è **debole** la correlazione calcolata con Spearman.

#### **Postazioni ATM05 e ATM10**

Tutti i composti hanno correlazioni **molto deboli**. Per il benzene (C6H6) la correlazione è **debole** e lo stesso si vede per il PIDVOC con Spearman e Kendall.

#### **Postazioni ATM05 e ATM14**

Le correlazioni sono **molto deboli** per tutti i composti eccetto il benzene (C6H6) con una correlazione **debole**. Sul PIDVOC la correlazione è **moderata**. Per quanto riguarda il TRS, nella stazione ATM14, non è possibile effettuare correlazione perché abbiamo mancanza di dati.

#### **Postazioni ATM07 e ATM10**

H2S e C6H6 risultano avere una correlazione **debole**, per il resto dei composti le correlazioni sono **molto deboli**.

### **Postazioni ATM07 e ATM14**

E' **debole** la correlazione del PIDVOC con Spearman. Il restante dei componenti danno correlazioni **molto deboli**. Per quanto riguarda il TRS, nella stazione ATM14, non è possibile effettuare correlazione perché abbiamo mancanza di dati.

### **Postazioni ATM10 e ATM14**

Correlazione **debole** di C6H6, **moderata** per il VOC, **molto deboli** per il resto dei composti. Per quanto riguarda il TRS, nella stazione ATM14, non è possibile effettuare correlazione perché abbiamo mancanza di dati.

#### ○ **Correlazione temperatura per stazione e sensore**

Innanzitutto, carichiamo il dataset dei dati meteo relativi alla zona di analisi. Creiamo tre liste, una con tutte le postazioni, una con tutti i sensori e una con tutte le informazioni meteo necessarie. A questo punto viene richiamata la funzione **corr\_Sensori** per restituire i valori delle correlazioni di Pearson, Spearman e Kendal. Questa funzione prende in input un DataFrame costruito con la funzione **creazione\_df()**, la **colonna 1** e la **colonna2**. Per la creazione del dataframe sono stati passati alla funzione **creazione\_df()**, la funzione **pos()**, la quale ha come parametri il dataframe e il sensore di riferimento, che elimina i valori nulli, **stazione[i]**, **meteo[0]** per prendere in considerazione la temperatura, **sensore[i]** per indicare il composto da analizzare e infine il dataframe **df\_DatiMeteo**. Ottenuti i valori delle correlazioni di **Pearson** e **Spearman** da **corr\_Sensori**, sono stati mostrati a video.

Per quanto riguarda il PIDVOC abbiamo una correlazione quasi **moderata** con ATM5, mentre per il VOC risulta forte (ma non attendibile). Abbiamo una correlazione **debole** con ATM07 e VOC (ma non attendibile a causa dei pochi valori), una correlazione **moderata** con PIDVOC, e correlazioni deboli con H2S e H2SJ. Per la stazione ATM10 abbiamo una correlazione **debole** con VOC e **molto debole** per i restanti. Per l'ATM14 abbiamo una correlazione **moderata** con C6H6 e PIDVOC e i restanti sono **molto deboli**, ma con l'H2SJ è **debole** ma negativa.

#### ○ **Correlazione direzione vento per stazione e sensore**

Per questa correlazione sono state effettuate le stesse operazione per la correlazione precedente andando a modificare opportunamente i valori necessari, ovvero **meteo[0]** con **meteo[1]** per la direzione del vento.

Per l'ATM05 abbiamo una correlazione **molto debole** per tutti i sensori ad eccezione del VOC che risulta **moderata** ma presenta pochi valori per essere affidabile. Per l'ATM07 la correlazione con il TRS e il VOC è **debole** ma sono pochi valori, per il resto **molto deboli**. Per l'ATM10 abbiamo correlazioni **molto deboli** con tutti i sensori. Nell'ATM14 ci sono correlazioni tutte **molto deboli**, tranne per il sensore TRS il quale ha poche registrazioni.

- **Correlazione pressione atmosferica per stazione e sensore**

Per questa correlazione sono state effettuate le stesse operazione per la correlazione precedente andando a modificare opportunamente i valori necessari, ovvero sostituendo **meteo[1]** con **meteo[2]** per la pressione atmosferica.

Per l'ATM05 c'è una correlazione **moderata** con il VOC ma poco attendibile a causa della mancanza di valori, per il resto tutte **molto deboli**. Per l'ATM07 risultano correlazioni moderate per il TRS e VOC ma sono poco attendibili, **debole** negativamente per l'H2S. Per l'ATM10 ci sono correlazioni tutte **molto deboli** tranne l'H2SJ che con Spearman è **debole**. Per l'ATM14 la correlazione è **debole** con il PIDVOC é **nulla** con TRS, per il resto tutte **molto deboli**.

- **Correlazione intensità del vento per stazione e sensore**

Per questa correlazione sono state effettuate le stesse operazione per la correlazione precedente andando a modificare opportunamente i valori necessari, ovvero **meteo[2]** con **meteo[3]** per l'intensità del vento.

Per l'ATM05 è **moderata** con il VOC ma poco attendibile, per il resto tutte **molto deboli**. Per l'ATM07 la correlazione è **debole** con il TRS e VOC ma non attendibile per i pochi valori, per il resto tutte **molto deboli**. Per l'ATM10 le correlazioni sono tutte **molto deboli**, mentre per l'ATM14 la correlazione è nulla con il TRS e **molto debole** per il resto.

- **Correlazioni spurie**

- **Numero di fallimenti giornalieri con valori meteo**

Per questa correlazione è stato creato un nuovo dataframe **df\_nuovo** a partire da quello dei dati meteo. Poi è stato effettuato un **groupby** sul giorno. Infine, è stata fatta una giunzione tra il nuovo dataset e il dataset dei fallimenti giornalieri calcolato nella query 7. Abbiamo una correlazione **debole** con la temperatura, per il restante dei valori abbiamo correlazioni **molto deboli**. Ovviamente è da tenere in conto il periodo estivo in cui ci sono poche precipitazioni.

- **Correlazione con l'umidità**

Per l'ATM05 si ha una correlazione **forte** con il VOC, ma poco attendibile, per il resto **molto debole**. Per l'ATM07 risulta una correlazione **debole** con TRS e VOC, ma poco attendibile, mentre per H2S è **debole** con Spearman e per il resto **molto debole**. Per l'ATM10 la correlazione è debole e negativa con il VOC e per l'ATM14 si ha una correlazione **nulla** con il TRS e con l'H2SJ si ha una correlazione **debole** con Pearson, per il resto **molto deboli**.

## Accorgimenti!

Dopo l'analisi del dataset, dei sensori e dei composti e delle correlazioni ci siamo accorti che le stazioni più vicine al sito di Tempa Rossa non registrano per la quasi totalità del tempo determinati dati:

- **VOC e H2SJ** per **ATM05** (all'entrata della struttura);
- **TRS, VOC e H2SJ** per **ATM07** (poco fuori dalla struttura);
- **TRS e VOC** per **ATM14** (situata all'interno della struttura).

Stranamente per quanto riguarda **ATM10**, ossia la stazione più lontana dalla struttura, i dati vengono registrati correttamente per tutti i sensori.

## Analisi e Raccomandazioni

### Soglie di Rischio e Inquinanti

- **VOC (Volatile Organic Compounds)**

Possiamo definire i composti organici volatili come una classe che comprende composti chimici differenti, caratterizzati da una facile vaporizzazione a temperatura ambiente e in grado di reagire nella troposfera dando vita a composti inquinanti. Ne fanno parte gli **idrocarburi alifatici e aromatici** (tra cui il benzene), le **aldeidi** (come la formaldeide, altamente cancerogena) e **altri composti**. Non tutti i composti organici volatili presentano pericolosità specifiche per l'uomo: ne sono un esempio sia il metano che il propano. Altri invece sono stati addirittura classificati dalla **IARC** (International Agency for Research on Cancer) come cancerogeni: fra questi si annoverano la formaldeide ed il benzene, su cui approfondiamo nel seguito.

- **C6H6 (Benzene)**

Il benzene (comunemente indicato anche come benzolo) è una sostanza chimica organica altamente infiammabile che si presenta in forma liquida ed inodore. Evapora nell'aria (volatile) a temperatura ambiente, con un caratteristico odore dolciastro ed aromatico. Noto per la sua tossicità, il benzene è stato classificato dall' IARC nel gruppo 1. Nel gruppo 1 vi sono le sostanze per cui ci sono sufficienti evidenze di cancerogenicità negli esseri umani.

L'esposizione al benzene avviene essenzialmente per inalazione (circa il 99% del benzene assunto), per contatto cutaneo o per ingestione (consumo di cibo o di bevande contaminate).

Gli effetti tossici sull'uomo:

- Effetti tossici acuti i quali sono dovuti a contatto per inalazione o per via cutanea. Brevi esposizioni di 5-10 minuti a livelli molto alti di benzene nell'aria possono condurre alla morte, mentre livelli di concentrazione più bassi possono causare giramenti, sonnolenza, aumento del battito cardiaco, tremori, confusione e perdita di coscienza. Concentrazioni minori ma più prolungate nel tempo possono alterare la memoria e capacità psichiche. L'ingestione di cibi e bevande che contengono alti livelli di benzene possono causare il vomito, irritazione allo stomaco, giramenti, sonnolenza, convulsioni, aumento del battito cardiaco, coma e morte.
- Effetti tossici cronici sono invece dovuti a periodi di esposizione molto lunghi e a basse concentrazioni.

*Il benzene è molto tossico per l'uomo. Ciò che preoccupa di più, sia a livello professionale che ambientale, è la comparsa del cancro del sangue dovuta all'esposizione ripetuta a concentrazioni di benzene di qualche ppm per più decine di anni. L'insorgere di queste malattie si manifesta più frequentemente in seguito ad esposizioni basse e continue piuttosto che elevate e intermittenti.*

*Al fine di tutelare la salute dei lavoratori, che sono esposti ad agenti cancerogeni o mutageni sul luogo di lavoro, la direttiva 2004/37/CE del Parlamento europeo e del Consiglio impone al datore di lavoro di garantire un certo livello di salubrità del luogo in cui il dipendente abitualmente svolge le proprie mansioni. Con la DIRETTIVA (UE) 2022/431 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO del 9 marzo 2022, che modifica la direttiva 2004/37/CE, si stabilisce che il benzene risponde ai criteri di classificazione come sostanza cancerogena (categoria 1 A) di cui al regolamento (CE) n. 1272/2008 ed è pertanto un agente cancerogeno ai sensi della direttiva 2004/37/CE. Alla luce di dati scientifici più recenti, è opportuno rivedere i valori limite di cui all'allegato III della direttiva 2004/37/CE per il benzene. Tuttavia, per il benzene potrebbe essere difficile rispettare un valore limite di 0,2 ppm ( $0,66 \text{ mg/m}^3$ ) in alcuni settori nel breve termine. È pertanto opportuno introdurre un periodo transitorio di quattro anni dall'entrata in vigore della presente direttiva. Come misura transitoria:*

- Il valore limite di 1 ppm ( $3,25 \text{ mg/m}^3$ ) dovrebbe continuare ad applicarsi fino al 5 aprile 2024.*
- Il valore limite di 0,5 ppm ( $1,62 \text{ mg/m}^3$ ) dovrebbe applicarsi dal 5 aprile 2024 fino al 5 aprile 2026*

*Valori limite che fanno riferimento anche ad un limite di esposizione di 8 ore.*

- H<sub>2</sub>S (Acido Solfidrico)**

*L'acido solfidrico (H<sub>2</sub>S) è un gas che si trova comunemente durante la perforazione e la produzione di petrolio greggio e gas naturale, oltre che nel trattamento delle acque reflue e nelle fogne. Incolore, infiammabile, velenoso e corrosivo, il gas H<sub>2</sub>S si riconosce per il suo odore di uova marce. Con una tossicità simile al monossido di carbonio, che impedisce la respirazione cellulare, il monitoraggio e la rilevazione precoce dell'H<sub>2</sub>S potrebbero significare la differenza tra la vita e la morte.*

*Coloro che hanno un'esposizione prolungata a livelli abbastanza alti di gas H<sub>2</sub>S da causare incoscienza possono continuare a sperimentare mal di testa, riduzione della capacità di attenzione e delle funzioni motorie. Gli effetti polmonari dell'esposizione al gas H<sub>2</sub>S possono non essere evidenti fino a 72 ore dopo la rimozione dall'ambiente colpito. L'edema polmonare ritardato, un accumulo di liquido in eccesso nei polmoni, può anche verificarsi in seguito all'esposizione ad alte concentrazioni.*

*L'H<sub>2</sub>S non si accumula nel corpo, ma l'esposizione ripetuta/prolungata a livelli moderati può causare bassa pressione sanguigna, mal di testa, perdita di appetito e perdita di peso. L'esposizione prolungata a bassi livelli può causare eruzioni cutanee dolorose e occhi irritati. L'esposizione ripetuta nel tempo ad alti livelli di H<sub>2</sub>S può causare convulsioni, coma, danni al cervello e al cuore, persino la morte. Di seguito viene mostrata una lista di effetti sull'uomo in base alla concentrazione a cui si è esposti:*

- 0.05 ppm – soglia attivazione odore*
- 3 ppm – odore offensivo*
- 20 ppm – irritazione delle mucose, degli occhi e dell'apparato respiratorio*
- 50ppm – soglia dei danni alla vista*
- 100 ppm paralisi olfattoria*



- 250ppm – possibile edema polmonare
- 500 ppm - sonnolenza, stordimento, cefalea, andatura incerta
- 700 ppm – perdita di conoscenza e stato d'ansia
- 1000/2000 ppm - Iperpnea, seguita da un rapido collasso e conseguente inibizione dell'apparato respiratorio
- 2000/5000 ppm - Effetto immediato paralizzante nei centri della respirazione
- >5000 ppm – morte rapida

*La soglia sul valore limite per gli odori molesti è invece posta ad un massimo di 7  $\mu\text{g}/\text{m}^3$  per un periodo di esposizione di 30 minuti e a 150  $\mu\text{g}/\text{m}^3$  per la media giornaliera.*

### **Inquadramento Territoriale e Risultati**

*Al fine di valutare se le attività svolte presso il Centro Olio Tempa Rossa, sito nel comune di Corleto Perticara (PZ), rappresentano un rischio per la salute dei lavoratori, degli abitanti dei paesi limitrofi e degli allevamenti prossimi al sito sono state effettuate delle ricerche. Queste ci hanno permesso di fare un inquadramento territoriale. I centri abitati più prossimi allo stabilimento sono quello di Corleto Perticara (PZ), di Guarda Perticara (PZ) e di Gorgoglione (MT) distanti dai 4 km ai 4,6 km dal Centro Olio.*

*Le aziende agricole presenti, invece, sono di Lombardi G. e di Filippo V. e distano rispettivamente circa 200m (in direzione Sud) e 875m (in direzione Est) in linea d'aria dal sito. Aziende queste che svolgono attività zootecniche e colture. E' stato possibile stabilire i livelli medi di benzene presso tali allevamenti poiché situati in prossimità delle stazioni di monitoraggio ATM05 e ATM07. Calcolando il livello medio di benzene per queste due postazioni è stato possibile constatare come per l'ATM05 il livello medio è circa 0,00026  $\text{mg}/\text{m}^3$  mentre per l'ATM07 è di circa 0,0002  $\text{mg}/\text{m}^3$ , valori che non solo rispettano i limiti governativi ma sono anche molto minori.*

*Una conclusione analoga viene dedotta calcolando i livelli medi di benzene anche per le restanti stazioni l'ATM10 e l'ATM14, che sono rispettivamente la stazione più lontana (1,7 km sita nelle prossimità di una pietraia) e la stazione più vicina posta sul perimetro esterno dello stabilimento. Per l'ATM10 il livello medio è di circa 0,00038  $\text{mg}/\text{m}^3$  per l'ATM14 è di 0,00046  $\text{mg}/\text{m}^3$ .*

*Viste le attività svolte presso queste ultime due stazioni di monitoraggio, è stato effettuato un controllo prendendo in considerazione la settimana dal 14 al 21 Maggio 2021 (settimana in cui non si sono registrati fallimenti nell'invio dei dati) per verificare se i livelli medi di benzene fossero influenzati dal traffico veicolare (prodotto dalle macchine lavoratrici e dal transito degli impiegati). Dal controllo è emerso che il livello di benzene ha dei picchi dopo periodi di tempo non comparabili. Però non sapendo quando avvengono i cambi turno e se questi avvengono spesso, non si è riuscito a stabilire se i picchi di benzene sono dovuti effettivamente dal traffico veicolare.*

*In conclusione, in base ai dati inviati, le emissioni di benzene dovute alle attività effettuate presso il Centro Oli Tempa Rossa non rappresentano un rischio per la salute dei cittadini e degli allevamenti vicini. Purtroppo, non è possibile effettuare una valutazione analoga per gli altri inquinanti in quanto, come sottolineato precedentemente, stranamente solo per quanto riguarda l'ATM10, ossia la stazione più lontana dalla struttura, si ha che i dati vengono registrati correttamente per tutti i sensori. In più come dichiarato nella direttiva 2004/37/CE per il benzene potrebbe essere difficile rispettare un valore limite di 0,2 ppm (0,66  $\text{mg}/\text{m}^3$ ) in alcuni settori nel breve termine, e ci viene difficile credere come che un'azienda petrolchimica non solo riesca a rispettarli nel breve termine ma addirittura a raggiungere anche un livello medio massimo di 0,00046  $\text{mg}/\text{m}^3$ . Inoltre i dati inviati*

*dall'A.R.P.A.B non sembrerebbero essere affidabili, a causa dell'alta percentuale di fallimenti in determinati periodi e la notevole mancanza di dati per addirittura due mesi. Ad avvalorare tale tesi sono le numerosissime lamentele degli abitanti delle aree comprese nelle concessioni petrolifere. Gli abitanti dei centri limitrofi dichiarano di avere disagi per gli odori molesti e rivendicano il diritto di sapere cosa respirano, senza puntare il dito contro nessuno, ma chiedendo solo i dati corretti. Durante i giorni più critici, dal punto di vista olfattivo, alcuni di essi si sono rintanati in casa, altri hanno persino lasciato la propria abitazione e le scuole sono state costrette a chiudere ermeticamente le finestre per proteggere i bambini dall'odore insopportabile di zolfo e uova marce. Analizzando accuratamente i valori rilevati dai sensori H<sub>2</sub>S e H<sub>2</sub>SJ in tutte le stazioni è stato constatato che il valore di soglia massimo per il cattivo odore (0,418 µg/m<sup>3</sup>) viene superato. Il sensore H<sub>2</sub>S per il 100% delle rilevazioni supera la soglia, mentre secondo i dati registrati dal sensore H<sub>2</sub>SJ la soglia viene superata dal 69% delle rilevazioni.*

*Una delle aree che si trova ad affrontare maggiormente questi disagi è il comune di San Martino d'Agri, a metà strada tra gli impianti in Val d'Agri e quelli siti in Corleto Perticara. L'amministratore comunale Conte F. insieme ad alcuni Istituti di ricerca ha tentato di elaborare un sistema di monitoraggio dell'aria, per fare chiarezza. Tuttavia, per fondi non sufficienti non si è mai riusciti.*

*Nonostante l'attivazione del sistema torcia sia frequente, ogni qual volta si richiedono a Total spiegazioni questa, attraverso comunicati stampa, tiene a rassicurare che dalla rete di monitoraggio costituita dalle centraline della qualità dell'aria posizionate nelle immediate vicinanze dell'impianto, non si registrano anomalie.*

*Nonostante tali "rassicurazioni" "ciò che preoccupa, è che la compagnia Total reiteri la violazione delle prescrizioni AIA con emissioni in atmosfera decisamente superiori ai limiti di legge" come dichiarato dall'allora Direttore Generale dell'ARPAB: Antonio Tisci.*