ESERCIZIO DATASET COVID

1. Per l’esercizio dato che la fonte di dati comune a tutti i servizi di condivisione è la protezione civile italiana verrà usata come fonte la stessa.

-URL: [COVID-19/dpc-covid19-ita-andamento-nazionale-20230203.csv at 200ebdacedd3d24908c34302eb2b2b6894ff3e9f · pcm-dpc/COVID-19 (github.com)](https://github.com/pcm-dpc/COVID-19/blob/200ebdacedd3d24908c34302eb2b2b6894ff3e9f/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale-20230203.csv)

Questa fonte dati indica dati sulla pandemia Covid19 con riferimento alla zona geografica Italia dall’inizio al momento della visualizzazione del dataset.

Come campi sono usati tutte le descrizioni di ogni singola colonna, ogni campo può essere facilmente interpretato data la forma descrittiva usata per nomenclarli.

La licenza è CC-BY

Alla fonte può essere data una classificazione di \*\*\* stelle.

Al file creato può essere attribuita una classificazione \*\*\*\*\*stelle.

2. E’ stato creato il CSV Classe.

3. Nel CSV classe sono stati raccolti i dati dei singoli partecipanti, gli è stata creata una chiave primaria rappresentata dal campo ID, e sono stati creati i campi autodescrittivi.

4. La tabella avendo una sola riga non ha bisogno di chiave primaria collegata alla tabella Classe, la chiave primaria più sensata potrebbe essere il luogo di riferimento e cioè il campo Luogo.

5. Il confronto che può nascere dai due set di dati è la percentuale ricoveri/positivi.

Nel gruppo classe le ospedalizzazioni sono zero quindi per varie considerazioni quali il range di età ristretto non ci sono le condizioni per un confronto costruttivo.

6. Le figure che sono servite alla stesura di questo esercizio sono Data Engineer e Data Analyst.

Il Data Engineer è stata la figura che ha permesso la raccolta e la strutturazione dei dati.

Il Data Analyst ha permesso le analisi richieste.

Mancano il Data Scientist dato che non è richiesto trovare una soluzione o fare una previsione.

La figura di Data Journalist potrebbe essere stata rappresentata durante la fase di raccolta dei dati di classe per la formulazione del CSV classe.

7. I Metadati riconosciuti possono essere distinti tra Metadati fonte e Metadati file word esercizio.

I metadati della fonte sono tutte le informazioni relative al dataset pubblico sui dati covid come per esempio creatore, data, descrizione, ecc.

I metadati del file word esercizio sono le informazioni relative a questo file.docx e le parti descrittive usate nello stesso.