D3-Esercizio COVID

1. RECUPERARE 5 FONTI DATI DA INTERNET CHE DESCRIVANO LA SITUAZIONE COVID AL 2023 (CI DEVONO ESSERE ALMENO I CAMPI “POSITIVI COVID”, “RICOVERATI”, “MORTI”)
2. IN UN DOCUMENTO TESTO INDICARE LA FONTE, BREVE DESCRIZIONE DEL CONTENUTO, DESCRIZIONE DEI CAMPI E LA CLASSIFICAZIONE 5 STELLE DELLE FONTI SCELTE
3. DEFINIRE UNA TABELLA PER CONFRONTARE I DATI “NUMERO POSITIVI” COVID, “PERSONE RICOVERATE”, “CITTA’”, “GIORNI POSITIVI” DI OGNI PARTECIPANTE
4. DARE UN NOME ALLE ENTITA’ E AGLI ATTRIBUTI (CAMPI) IN MODO DESCRITTIVO

**COVID – PARTE 2**

1. ~~DEFINIRE LA CHIAVE PRIMARIA PER OGNI TABELLA~~
2. CONFRONTARE I DATI CLASSE/ITALIA E DESCRIVERE LA SITUAZIONE
3. DESCRIVERE IN QUALE FASE INTERVIENE OGNUNA DELLE FIGURE STUDIATE (Data Scientist, Data Engineer, Data Analyst e Data Journalist) SE PRESENTI
4. QUALI METADATI POTETE IDENTIFICARE IN QUESTA ESERCITAZIONE?

**PARTE 1**

Fonte 1:

<https://statistichecoronavirus.it/continenti/coronavirus-europa/>

Il dataset è un open data, accessibile che fornisce informazioni sul coronavirus in Europa, con un focus specifico sulla situazione dei casi confermati, dei decessi e dei guariti nei paesi europei.

È possibile visualizzare diverse informazioni del dataset, tra cui:

* Paese: Elenco dei paesi europei inclusi nelle statistiche.
* Casi totali: Numero totale di casi confermati in ciascun paese.
* Nuovi casi: Numero di nuovi casi confermati riportati in ciascun paese.
* Decessi totali: Numero totale di decessi causati in ciascun paese.
* Nuovi decessi: Numero di nuovi decessi causati riportati in ciascun paese.
* Guariti totali: Numero totale di persone guarite in ciascun paese.
* Visualizzazione: possiamo scegliere tra un istogramma o forma mappale
* Il dataset può essere scaricato in formato CSV, ci sono disponibili i metadati
* Classificazione 5 stelle

Fonte 2:

<https://www.worldometers.info/coronavirus/>

In questa fonte troviamo un dataset molto complesso che contiene una vasta gamma di informazioni. Le più importante ci sono:

* Popolazione (la popolazione totale nella nazione desiderata)
* Decessi
* Nuovi casi
* Casi gravi (il numero di casi gravi o critici di COVID)
* Ricoverati
* Test effettuati (il numero totale di test effettuati nella nazione) ecc.

La classificazione è a 5 stelle indicando il loro grado di affidabilità e qualità dei dati forniti.

Fonte 3:

<https://ourworldindata.org/covid-vaccinations>

Questo dataset offre una vasta gamma di campi dati relativi alla pandemia di COVID-19, fornisce un'analisi approfondita dei dati relativi alle vaccinazioni a livello globale. Alcuni esempi di campi presenti nel tool di esplorazione includono:

* Continente
* Paese
* Nuovi casi
* Decessi
* Percentuale di popolazione vaccinata (la percentuale di popolazione nel paese che ha ricevuto almeno una dose di vaccino)
* Ricoverati in ospedale
* Tipo di vaccino (il nome o il tipo di vaccino utilizzato nel paese)

Il dataset può essere scaricato in un file csv, json

I dati sulle vaccinazioni vengono valutate con una classificazione a 5 stelle.

Fonte 4:

<https://data.humdata.org/dataset/coronavirus-covid-19-cases-and-deaths?>

Questo è dataset fornisce che fornisce un’analisi approfondita dei dati globali riguardanti i casi e le morti legate alla pandemia.

Alcuni esempi di campi presenti sul sito includono il paese, i casi confermati totali, nuovi casi nell’ultima settimana, morti confermati in totale, nuovi decessi nell’ultima settimana, vaccinazioni (non ho trovato i ricoverati)

Classificazione 5 stelle.

Fonte 5:

<https://health.google.com/covid-19/open-data/explorer?loc=IT>

Questa fonte, come la precedente, offre una visualizzazione interattiva dei dati in forma tabellare. È la prima fonte elencata che fornisce dati specifici per le singole città italiane, tuttavia questi dati sono aggiornati solo fino al 15/09/2022.

I campi disponibili nella versione interattiva includono il periodo di analisi (ultimi 7 giorni, 30 giorni, ecc.), la nazione (da cui è possibile selezionare la regione e la città) e il numero corrispondente di casi e decessi. Sono disponibili anche informazioni sulle vaccinazioni, guariti e ricoveri, ma solo a livello nazionale.

La fonte fornisce anche metadati e un file CSV per esplorare ulteriormente il dataset. Nel file CSV, sono presenti una quantità molto più ampia di dati, come il numero di infermieri per 1000 persone, l'umidità media, il numero di scuole chiuse e la social mobility.

Classificazione 5 stelle.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fonte | Paese | Numero positivi | Ricoverati | Deceduti |
| 1 | Italia | 25.547.414 | 25.174.036 | 187.850 |
| 2 | Francia | 40,105,208 | 39,887,439 | 167,424 |
| 3 | Cuba | 1114309.0 | 662,496,191 | 8530.0 |
| 4 | Cina | 99,273,266 |  | 99,273,266 |
| 5 | USA | 92,357,092 | 5,495,623 | 1,004,449 |

**PARTE 2**

Ho scelto questo dataset per spiegare le tre figure <https://ourworldindata.org/covid-vaccinations>

Data engineer si occupa con l’elaborazione e la gestione dei dati sulle vaccinazioni COVID-19, trasforma i dati garantendo la qualità e l'integrità dei dati stessi.

Data Scientist: Il Data Scientist utilizza i dati sulle vaccinazioni COVID-19 per eseguire analisi avanzate e sviluppare modelli predittivi. Potrebbe utilizzare tecniche di machine learning e algoritmi statistici per identificare pattern, tendenze e relazioni nei dati.

Data Analyst: lavora con i dati sulle vaccinazioni COVID-19 per estrarre informazioni significative e generare report o visualizzazioni. Analizza i dati per identificare i tassi di vaccinazione, confrontare i risultati tra paesi o regioni e valutare l'efficacia delle campagne vaccinali.

Data Journalist: Utilizza i dati sulle vaccinazioni COVID-19 per creare storie informative e coinvolgenti. Può creare visualizzazioni coinvolgenti, infografiche o articoli basati sui dati per trasmettere le informazioni relative alle vaccinazioni COVID-19 in modo chiaro e accessibile.

La maggior parte dell'esercizio può essere considerata un metadato in quanto fornisce tutte le informazioni di navigazione dei dataset.