**Parte 1**

**Fonte 1**

<https://github.com/CSSEGISandData/COVID-19>

Questa prima fonte presenta vari dataset a cui accedere. Nello specifico, vengono aggiornati giornalmente i casi, il totale delle vittime e le persone che si sono sottoposte al vaccino elencati per ogni nazione. E’ inoltre possibile cambiare la visualizzazione e confrontare i dati tra gli anni, decidere che sia settimanale, concernente solo gli ultimi 28 giorni o giornaliero. Possibile scaricare vari.csv e sono presenti metadati.

Scala Tim Berners Lee: 5 stelle

**Fonte 2**

<https://ourworldindata.org/explorers/coronavirus-data-explorer?zoomToSelection=true&time=2020-03-01..latest&facet=none&country=~ITA&pickerSort=asc&pickerMetric=location&Interval=7-day+rolling+average&Relative+to+Population=true&Color+by+test+positivity=false&Metric=Confirmed+cases>

Il contenuto del dataset di questa fonte è interattivo e accessibile, nonché scaricabile come file .csv. Al suo interno è possibile filtrare i dati per nazioni, decidere se visualizzare le “morti confermate”, “casi confermati”, “terapia intensiva”, nonché in quale intervallo di tempo. Presenta metadati. E’ possibile filtrare i dati da visualizzare per ogni nazione del globo e per ognuna scaricare il .csv interessato.

Scala Tim Berners Lee: 5 stelle

**Fonte 3**

<https://dati-covid.italia.it/>

Questa fonte presente un dataset completo, così da poter effettuare un’analisi a livello nazionale italiano. E’ possibile scaricare diversi .csv (in base a che tipo di indagine vogliamo condurre regionale/provinciale/nazionale).

I dati presenti dichiarano a livello nazionale e livello regionale quanti ricoverati, quanti deceduti, quanti con sintomi e le variazioni da un giorno analizzato all’altro ecc. Presenti metadati.

Scala Tim Berners Lee: 4 stelle

**Fonte 4**

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

Questa fonte ha una tipologia di visualizzazione interattiva, in forma tabellare e presenta la possibilità di visualizzare sia i dati ogni nazione del mondo e, nel caso in cui venisse selezionata “Italia”, applicabile un ulteriore filtro per ogni singola regione e città che vogliamo prendere in considerazione per la nostra indagine. Questi dati sono aggiornati al 15/09/2022.

I campi su cui poter operare sono il periodo di analisi (quindi selezionabili le voci come “dall’inizio”, “ultimi 7 giorni”, “ultimi giorni”), la voce “nazione”, dopo il quale è possibile (soltanto per la voce “Italia”) selezionare sia la regione (qualora volessimo operare una ricerca più specifica e non a livello generale in Italia) che la città interessata e quindi il numero dei casi e dei decessi. I dati relativi alle vaccinazioni, guariti e ricoveri sono disponibili solo su scala nazionale.

È possibile esplorare ogni dataset relativo all’indagine che vogliamo condurre. Presenti metadati.

Scala Tim Berners Lee: 4 stelle

**Fonte 5**

<https://covid19.who.int/data>

Questa fonte presente un dataset molto schematico. Scaricando il dataset a noi interessato, si apre una tabella che specifica ogni tipo di dato presente con la loro descrizione. Metadato immediato. Dataset parzialmente completo, in quanto mancano i dati concernenti i ricoveri e i guariti. Possibile verificare casi, deceduti e vaccinazioni sia su scala globale che per ogni singola nazione. In aggiunta, questi dati possono essere ulteriormente filtrati per periodo di tempo interessato.

Scala Tim Berners Lee: 5 stelle

|  |  |  |  |
| --- | --- | --- | --- |
| **Fonte** | **Positivi** | **Ricoverati** | **Deceduti** |
| **1.** GITHUB | 24.099.206 | 7.350 | 180.679 |
| **2.** Our World in Data | 25.870.833 | non chiaro | 190.517 |
| **3.** MID | 24.327.664 | 7.793 | 181.271 |
| **4.** Google Data | 22.096.450 | N/A | 176.404 |
| **5**. WHO | 25.870.833 | N/A | 190.517 |

**Attributi:**

1. **Fonte:** identifica la provenienza dei dati che andiamo a trattare
2. **Positivi:** identifica il totale delle persone che sono risultate positive al COVID-19 in Italia
3. **Ricoverati**:identifica il totale dei ricoveri in Italia
4. **Deceduti:** identifica quante vittime da COVID-19 ci sono state in Italia

**Parte 2**

Prendendo in considerazione la prima fonte: <https://github.com/CSSEGISandData/COVID-19> è possibile l’identificazione dei ruoli di:

**Data scientist:** in quanto è stato necessario analizzare i dati per tutto il periodo del COVID-19 e le sue evoluzioni giornaliere. In questo, è stato chiamato anche a fare previsioni sul possibile futuro andamento, basandosi su dati registrati in precedenza.

**Data Engineer:** in questa fonte specifica, dove sono indicati tutti i dati concernenti ogni singola nazione, il suo ruolo ha visto come assicurarsi che tali dati fossero attendibili e fruibili in sicurezza, nonché il loro attento aggiornamento.

**Data Analyst:** colui/colei che ha dovuto analizzare, ispezionare e trasformare i dati che provenivano da tutte le fonti attendibili riportanti quanto richiesto, ovvero i dati concernenti la diffusione del COVID-19.

**Metadati:** confrontando le fonti utilizzate in questo esercizio è possibile vedere come si inseriscono sia metadati descrittivi, in quanto sono volti a descrivere l’oggetto a cui sono associati, sia metadati di tipo amministrativi e gestionali dal momento che vengono conservati nel tempo e mantengono una loro integrità con il passare del tempo, sia metadati di tipo strutturale perché permettono sia una loro definizione veloce sulla struttura interna sia per l’identificazione delle relazioni che sussistono con altri documenti presenti.