ESERCIZIO D3 - FONTI DATI, OPEN DATA, CLASSIFICAZIONE 5 STELLE, METADATI, CSV

1 FONTE :<https://github.com/pcm-dpc/COVID-19/blob/master/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale.csv> , 9 GIUGNO, ore 14:30

Questo dataset ci fornisce dati relativi all’andamento nazionale del covid su base giornaliera, fino all’8 di giugno 2023. E’ possibile scaricare il file in formato .csv. Come attributi sono presenti la data, ricoverati, ospedalizzati, totale positivi, deceduti e altre info di tipo giornaliero. Reputo la fonte di tipo 3 Stelle. I dati sono fino all’8 di giugno. <https://github.com/pcm-dpc/COVID-19/tree/master> fornisce un chiara definizione dei metadati

2 FONTE : <https://covid19.who.int/data> 9 GIUGNO, ore 14:40

Questa fonte è chiara nel fornire info riguardo i deceduti e i casi confermati. Come campi si hanno la data, il codice del paese, il nome della Nazione, ufficio regionale dell’Oms, nuovi casi, casi cumulativi, nuove morti e morti cumulativi. Reputo la fonte di tipo 3 stelle, in quanto disponibili in formato csv. Tuttavia, nonostante sia aggiornato al 7 di giugno, non vengono rilevati (al momento) ulteriori decessi o ulteriori casi (in confronto alla prima fonte). Anche qui vi è una descrizione delle variabili utilizzate.

3 FONTE: 9 GIUGNO, ore 15:00 <https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19>

Questo dataset fornisce i dati settimanali sul tasso di notifica a 14 giorni di nuovi casi e decessi Covid-19 segnalati dagli stati membri dell’ue/see al sistema europeo di sorveglianza. Ciò che emerge è che il dataset è aggiornato alla 22esima settimana del 2023 (29/05-04/06). Il numero dei casi e dei deceduti è superiore rispetto ai primi due. Reputo la fonte come Tre stelle. I metadati sono descrittivi ma di poche righe

FONTE 4: 9 GIUGNO, ore 15:10 Questo dataset, presente al sito dell’istituto superiore di sanità <https://www.epicentro.iss.it/coronavirus/sars-cov-2-sorveglianza-dati> scaricabile come file zip 2023 (sezione open data) è in formato excel giornaliero e aggiornato quotidianamentE (FILE DELL’ 8 GIUGNO). File da 2 stelle. Sono presenti diversi campi quali casi regioni, casi province, ricoveri, decessi ecc. Non sono riuscito ad eseguire la funzione somma per quanto riguarda i decessi e i ricoverati (se magari nella valutazione mi dicessi come fare sarebbe l’ideale, poiché avrei potuto inserire in tabella). Ma andando a valutare ogni singola data, in confronto con il dataset dell’oms, noto delle differenze nel numero di morti giornalieri. E’ presente una nota metodologica descrittiva.

FONTE 5: 9 GIUGNO ore 15:55 <https://ourworldindata.org/explorers/coronavirus-data-explorer?zoomToSelection=true&facet=none&country=~ITA&pickerSort=asc&pickerMetric=location&Metric=Confirmed+deaths&Interval=Cumulative&Relative+to+Population=false&Color+by+test+positivity=false> In questo sito, è possibile visualizzare un dataset in forma grafica, scaricando il csv. Anche qui possono essere visualizzati i deceduti, casi, quota test positivi. Reputo il dataset da 4 stelle. Metadati presenti.

TABELLA

|  | POSITIVI | RICOVERATI | DECEDUTI |
| --- | --- | --- | --- |
| FONTE 1 | 118319 | 1470 | 190625 |
| FONTE 2 |  |  | 190517 |
| FONTE 3 |  |  | 199222 |
| FONTE 4 |  | ? | ? |
| FONTE 5 |  |  | 190517 |
|  |  |  |  |

In conclusione, la mia sensazione è che ,in generale, ci sia abbastanza uniformità nel conteggio soprattutto dei deceduti (ad eccezione della fonte 3 (che vorrei una opinione al riguardo)).