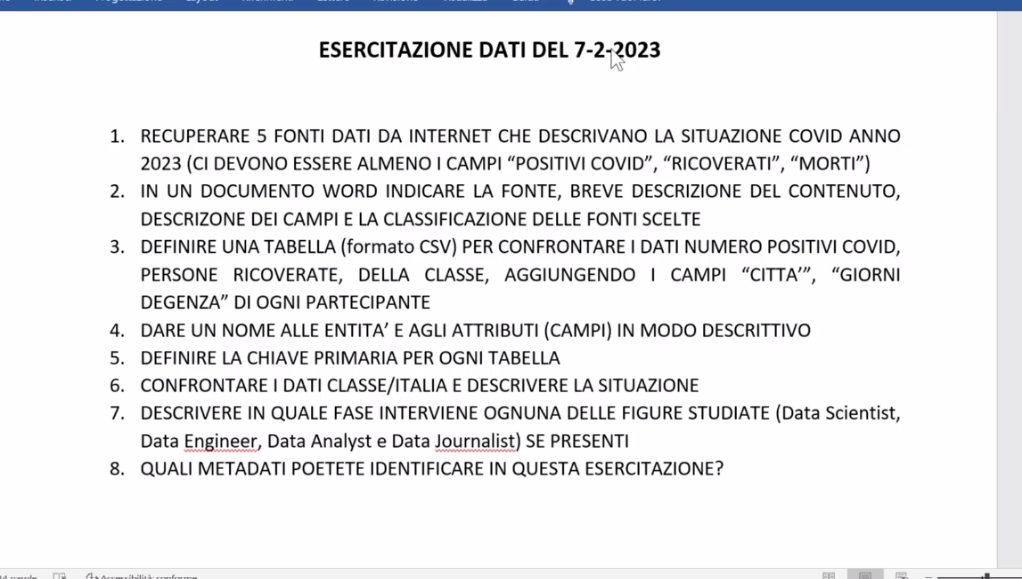
**Esercitazione Dati**



-**Fonte 1** : Sito del Dipartimento della Protezione Civile - Emergenza Coronavirus: la risposta nazionale (<https://github.com/pcm-dpc/COVID-19.git>)

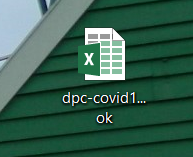
**-Descrizione:** Dataset riepilogativo riportante l’andamento nazionale della situazione Covid-19 da inizio pandemia alla data odierna ( licenza CC-BY-4.0)

In questa repository si trova una cartella zip dalla quale possiamo estrarre il primo open data in formato CVS , successivamente convertito in Excel per lavorarlo ed estrapolare alcuni dati ( N.positi covid, N.guariti,N.deceduti)

**-Classificazione :** ⭐ ⭐ ⭐ (Csv)



⭐ ⭐ ( Excel)



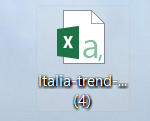
**La chiave primaria è il campo data**

Se il dato viene considerato come pagina tutte le informazioni sono dati. Sul file csv non sono presenti dati

-**Fonte 2**: Il sole 24 ore (https://lab24.ilsole24ore.com/coronavirus/?refresh\_ce)

**-Descrizione:** Il Dataset descrive i nuovi casi, attualmente positivi, guariti e morti dal primo rilevamento a oggi.

**-Classificazione :** ⭐ ⭐ ⭐ (Csv)



**La chiave primaria è il campo data**

**Metadati identificati:**

**Data ultimo aggiornamento:** 10 FEBBRAIO 2023

**Fonti:** I dati italiani utilizzati per i grafici sono forniti dalla protezione civile sulla base della segnalazioni delle regioni, con dettaglio provinciale. Sono disponibili qui. I dati sulla letalità e l’indice Rt sono diffusi dall’Istituto superiore di sanità (Iss). I dati aggregati per regione e provincia sono del Ministero della Salute. I dati sulle chiamate al 118 sono forniti da Areu. I dati internazionali sono invece della Johns Hopkins University. I dati internazionali sui vaccini sono di Our World in Data.I dati sull'efficacia dei vaccini sono forniti settimanalmente dall'Iss. I dati sulla circolazione delle varianti sono di Gisaid.

**Coordinamento:** Luca Salvioli

**Design director:** Laura Cattaneo

**Design:** Andrea Marson

**Data visualization:** Federico Barbara, Alice Calvi, Luca Galimberti

**Sviluppo:** Marina Caporlingua, Renato Zitti Pozzi

**-Fonte 3 :** CovidStat INFN ([CovidStat INFN - Dati dell'ISS](https://covid19.infn.it/iss/))

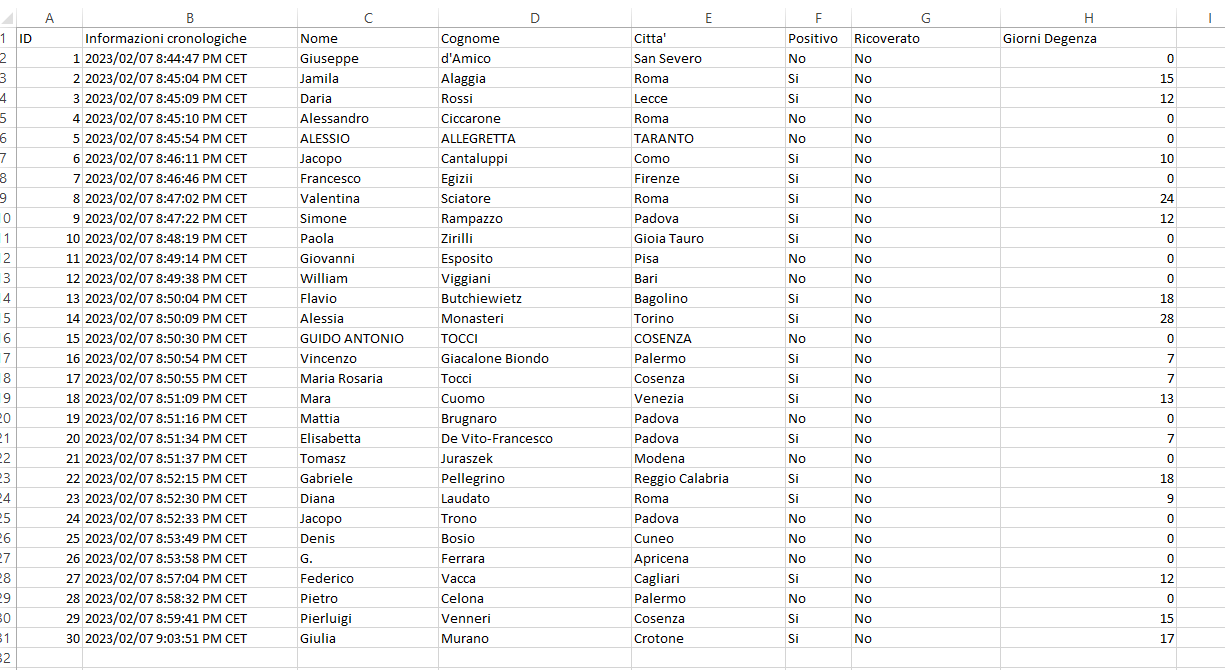
**-Descrizione:** Il DataSet nasce come un progetto realizzato dal Gruppo di Lavoro CovidStat INFN, la cui costituzione è stata promossa in seno all’Istituto Nazionale di Fisica Nucleare con l’obiettivo di mettere a disposizione dell’Unità di Crisi Covid19 interna all’INFN una analisi statistica dei dati forniti quotidianamente dalla Protezione Civile sulla diffusione della pandemia (Licenza )

**-Classificazione:** ⭐ ⭐ ⭐ (Csv)

****

**La chiave primaria è il campo data**

**Tabella dati classe Epicode**



I seguenti dati sono stati raccolti a seguito di un sondaggio effettuato su Google dove veniva richiesto di inserire i seguenti campi :

**Nome, Cognome,Città, Positività si o no,Ricoverato si o no,Giorni degenza**

Dopo un attenta analisi, possiamo affermare che il 60% della classe ha avuto il covid ed è una media piu alta rispetto a quella nazionale pari al 40%. Possiamo supporre che questa media cosi alta,sia dovuta anche all’età dei componenti della classe che essendo molto giovani hanno avuto una vita piu attiva.Non ci sono stati ricoverati nella nostra classe e anche la media dei ricoveri in Italia è molto bassa rispetto alla media dei positivi; ovviamente non si sono registrati decessi.

**Metadati:**

-Informazioni cronologiche

**Chiave Primaria:**

-ID

Dopo aver analizzato e confrontato le varie fonti, possiamo dire che è stato ed è compito del Data Engineer raccogliere i dati provenienti dalle varie regioni di italia, verificare la loro veridicità, elaborarli e consegnarli al data analyst e al data scientist che a loro volta si occuperanno rispettivamente di analizzare i dati in modo da poter elaborare soluzioni strategiche che insieme alle analisi predettive del data scientist servivano e serviranno per organizzare strutture sanitarie idonee ad affrontare la pandemia e a farsi trovare pronti in caso di un aumento di persone positive al covid-19. Infine il Data journalist si occuperà di diffondere i dati relativi alla pandemia in modo comprensibile a tutti.