



▪ Fonte 1 – INFN :

I grafici sono stati prodotti utilizzando dati forniti dall'Istituto Superiore di Sanità (ISS) con il quale l'INFN ha stabilito un accordo di collaborazione scientifica.

**Link:** <https://covid19.infn.it/iss/>

▪ Fonte 2 - HARVARD:

Descrizione: Dati statistici sulla malattia da coronavirus (COVID-19) del Dipartimento della Protezione Civile, Presidenza del Consiglio dei Ministri, Italia, ordinati per giorni e regioni italiane.

**Link:** <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/KDFYZW>

▪ **Fonte 3 – PCM DPC (Dati validi per effettuare il confronto):**

Lo stato Italiano per informare i cittadini a messo a disposizione i dati raccolti dal Dipartimento della Protezione Civile, con licenza **CC-BY-4.0**, che vengono aggiornati quotidianamente.

**Link:** <https://github.com/pcm-dpc/COVID-19>

**Link METADATA:** <https://github.com/pcm-dpc/COVID-19/blob/master/metadata/covid-19-ita-andamento-nazionale.xml>

▪ Fonte 4 – GIMBE:

Casi confermati dal Ministero della Salute nella settimana 27 gennaio-2 febbraio 2023. La Fondazione GIMBE aggiornerà questa pagina settimanalmente nella giornata di venerdì.

**Link:** <https://coronavirus.gimbe.org/>

- Fonte 5 -GEDI LAD:

Aggiornato al 2 febbraio 2023, dati del ministero della Salute. Dal 30 ottobre 2022 l'aggiornamento e la pubblicazione dei dati è settimanale (ogni venerdì).

**Link:** <https://lab.gedidigital.it/gedi-visual/2020/coronavirus-i-contagi-in-italia/>

---

**FONTE:** Stato italiano - Dipartimento della Protezione Civile

**DESCRIZIONE:** File CSV, Lo stato Italiano per informare i cittadini a messo a disposizione i dati raccolti dal Dipartimento della Protezione Civile, con licenza **CC-BY-4.0**, che vengono aggiornati quotidianamente.

**CAMPI:** ID, data, stato, ricoverati\_con\_sintomi, terapia\_intensiva, totale\_ospedalizzati, isolamento\_domiciliare, totale\_positivi, variazione\_totale\_positivi, nuovi\_positivi, dimessi\_guariti, deceduti

**CLASSIFICAZIONE:** Inizialmente Tre Stelle (★★★) perché i file originali erano in CSV dopo l'importazione in Excel sono passate a Due Stelle (★★)

**CONSIDERAZIONI:** Dopo aver trovate le 5 fonti analizzandole singolarmente ho notato che la fonte dei dati era sempre la stessa cioè quella del Dipartimento della Protezione Civile, perchè essendo un Dato Open (licenza **CC-BY-4.0**, si tratta della licenza "Creative Commons Attribuzione 4.0 Internazionale", questa licenza prevede che i contenuti di questo sito siano liberamente distribuibili e riutilizzabili, a condizione che sia sempre indicata la fonte) molte società hanno costruito analisi e report in base alle proprie esigenze.

---

Per inserire i dati della classe è stato creato un Form di Google che conteneva i campi: ID, Data inserimento, Nome, Cognome, Città, Positivo, Ricoverato e Giorno Degenza. (*File importato su GIT Esercizio\_Covid\_CSV\_1.csv*)

I dati scelti per il confronto sono quelli della Protezione Civile (**Fonte 3 - PCM DPC**) il file inizialmente riportava tanti campi io per gestire meglio il file ho scelto i campi essenziali: data, stato, ricoverati\_con\_sintomi, terapia\_intensiva, totale\_ospedalizzati, isolamento\_domiciliare, totale\_positivivariazione\_totale\_positivi,nuovi\_positivi, dimessi\_guariti e deceduti. (*File importato su GIT covid19-ita-dati-completi.csv*)

Per una lettura completa e un ordine strutturale i due file csv sono stati importati in un foglio di calcolo formato Excel. (*File importato su GIT covid19-ita-dati-import.xlsx*)

---

Le figure Professionali che intervengono durante il processo del nostro progetto, sono:

### **DATA ENGINEER**

Può intervenire con le seguenti responsabilità:

- Creazione di moduli e sondaggi online con diversi tipi di domande. Analisi dei risultati in tempo reale e da qualsiasi dispositivo.
- Supporto ai colleghi del settore Analyst;

### **DATA ANALYST**

Può intervenire con le seguenti responsabilità:

- Raccogliere i dati da fonti interne ed esterne;
- Analisi dei data nei formati arrivati Identificare gli obiettivi dell'analisi dei dati;
- Importare file e creare unica struttura per gestire i file;
- Organizzare e strutturare i dati, schemi e trend;
- Creare report e dashboard con i risultati delle analisi da presentare ai responsabili;
- Analizzare, progettare e mantenere le architetture dei dati, per garantire qualità, disponibilità e usabilità.

### **DATA SCIENTIST**

Può intervenire con le seguenti responsabilità:

- Ricercare e ideare modelli statistici innovativi per l'analisi dei dati, collaborando alla progettazione;
- Costruire modelli predittivi tramite algoritmi di machine learning e tecniche statistiche al fine di supportare le scelte strategiche;
- Presentare soluzioni con tecniche di data visualization;

### **DATA JOURNALIST**

Può intervenire con le seguenti responsabilità:

- Collabora con i colleghi della redazione per trovare e analizzare i dati che supportano le loro indagini e i rapporti con la popolazione.
- Utilizza i dati per creare visualizzazioni per report e grafici, che presentano informazioni complesse in modo più semplice e significative alla portata della pubblica popolazione.
- Racconta storie di grande impatto per le trasmissioni, il web e i social media.
- Utilizzo di dati, tecniche analitiche e grafica per trovare e raccontare storie, coprendo sia notizie che analisi più approfondite.

---

### **METADATI:**

Di seguito i metadati più basilari in riferimento all'esercitazione:

- Formato: EXCEL
- Nome del file: covid19-ita-dati-import
- date di creazione: 07/02/2023
- ultima modifica: 10/02/2023
- Autore: Pierluigi Venneri
- Descrizione: Dati Covid19 Italia e confronto con dati della classe Epicode.