ESERCITAZIONE DATI 07.02.2023

Immagine che contiene testo

Descrizione generata automaticamente

# PUNTI 1-2

1. La fonte è: Istituto superiore di sanità

[Infografica web - Dati della Sorveglianza integrata COVID-19 in Italia (iss.it)](https://www.epicentro.iss.it/coronavirus/sars-cov-2-dashboard)

In questo dataset l’istituto superiore di sanità riporta i dati aggiornati giornalmente sulla situazione COVID-19 permettendo il download del dataset in formato Excel.

I campi riportano informazioni su casi, decessi e ricoveri a livello nazionale e regionale e gli attributi sono chiari.

[covid\_19-iss.xlsx (live.com)](https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.epicentro.iss.it%2Fcoronavirus%2Fopen-data%2Fcovid_19-iss.xlsx&wdOrigin=BROWSELINK)

classificazione 2 stelle in quanto il database è riportato in Excel, questo implica un software proprietario per la lettura ed elaborazione dei dati.

1. La fonte è: Ministero della salute

[Covid-19 (salute.gov.it)](https://www.salute.gov.it/portale/nuovocoronavirus/homeNuovoCoronavirus.jsp)

In questo dataset il ministero della salute aggiorna periodicamente l’andamento della pandemia. Il dataset che forniscono mediante GitHb permette di scaricare un file CSV.

I campi riportano informazioni su casi, decessi e ricoveri a livello nazionale e regionale e gli attributi sono chiari.

[COVID-19/dpc-covid19-ita-andamento-nazionale-20221119.csv at cae4442f4dd4e157fe18677a372d7be1c915e083 · pcm-dpc/COVID-19 (github.com)](https://github.com/pcm-dpc/COVID-19/blob/cae4442f4dd4e157fe18677a372d7be1c915e083/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale-20221119.csv)

Classificazione 3 stelle del dataset strutturato in quanto il file CSV non necessita di un software proprietario per la lettura ed elaborazione dei dati.

1. La fonte è Covid-19 Italia

[COVID-19 Italia - Desktop (arcgis.com)](https://opendatamds.maps.arcgis.com/apps/dashboards/0f1c9a02467b45a7b4ca12d8ba296596)

In questo dataset si fa riferimento agli stessi dati del ministero della salute permettendo di scaricare un file CSV

I campi riportano informazioni su casi, decessi e ricoveri a livello nazionale e regionale e gli attributi sono chiari.

[COVID-19/dpc-covid19-ita-andamento-nazionale-20221119.csv at 833dd35fda5cd6941fd5f06a7cc31caed1b3592e · pcm-dpc/COVID-19 (github.com)](https://github.com/pcm-dpc/COVID-19/blob/833dd35fda5cd6941fd5f06a7cc31caed1b3592e/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale-20221119.csv)

Classificazione 3 stelle del dataset strutturato in quanto il file CSV non necessita di un software proprietario per la lettura ed elaborazione dei dati.

1. Fonte: Organizzazione Mondiale della Sanità (WHO)

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

Il dataset presentato fa riferimento a tutti i continenti ma filtrando i dati si riesce a selezionare quelli italiani. Il dataset è scaricabile come CSV.

I campi riportano informazioni su casi, decessi e ricoveri a livello mondiale e mediante filtri si possono ottenere quelli a livello nazionale e gli attributi sono chiari.

<https://covid19.who.int/WHO-COVID-19-global-data.csv>

Classificazione 3 stelle del dataset strutturato in quanto il file CSV non necessita di un software proprietario per la lettura ed elaborazione dei dati.

1. Fonte. Protezione Civile

<https://data.world/ondata/covid-19-italia-dati-dipartimento-protezione-civile>

In questo dataset si fa riferimento agli stessi dati del ministero della salute permettendo di scaricare un file CSV

I campi riportano informazioni su casi, decessi e ricoveri a livello nazionale e regionale e gli attributi sono chiari.

[COVID-19/dpc-covid19-ita-andamento-nazionale-20221119.csv at 833dd35fda5cd6941fd5f06a7cc31caed1b3592e · pcm-dpc/COVID-19 (github.com)](https://github.com/pcm-dpc/COVID-19/blob/833dd35fda5cd6941fd5f06a7cc31caed1b3592e/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale-20221119.csv)

Classificazione 3 stelle del dataset strutturato in quanto il file CSV non necessita di un software proprietario per la lettura ed elaborazione dei dati.

# PUNTI 3-4-5

Mediante google form la classe ha raccolto i dati salvati in formato CSV

<https://github.com/JacopoCantaluppi/EPICODE/blob/ESERCITAZIONE-3/Esercizio_Covid_CSV.csv>

Gli attributi sono chiari e descrittivi.

Come chiave primaria è stato attribuito un ID univoco ad ogni registrazione.

# PUNTO 6

Confrontando quanto emerge dalla classe con i dati raccolti a livello nazionale si evince che il tasso di positività della classe è superiore a quello nazionale.

# PUNTO 7

DATA ENGINEER: nella prima fase di selezione delle fonti e ricerca

DATA ANALYST: nel momento in cui abbiamo inizizto ad analizzare i dati e confrontarli con quelli della classe

DATA JURNALIST: in questa attività di rendicontazione e presentazione dei dati agli stakeholder interessati

DATA SCIENTIST: no in quanto non è stato richiesto di fare analisi predittivo/statistiche sui dati raccolti

# PUNTO 8

I metadati ai quali si può fare riferimento in questa esercitazione dipendono dal soggetto che si prende in considerazione.

Se partiamo dai dati raccolti allora i metadati a cui facciamo riferimento sono quelli che descrivono e specificano cosa troviamo nei record. Per quanto riguarda i file scaricati invece possiamo fare riferimento a tutte le informazioni del documento stesso che si porta dietro al momento del download.