

In [2]: `import pandas as pd`

```

# Caricamento del dataset
df = pd.read_csv("covid.csv")

# Dimensioni
print(f"Righe: {df.shape[0]}")
print(f"Colonne: {df.shape[1]}")

# Prime colonne
print("Colonne principali:", df.columns[:10].tolist())

# Metadati
print("Esempio di record:")
print(df.head(1).T)

```

Righe: 429435

Colonne: 67

Colonne principali: ['iso_code', 'continent', 'location', 'date', 'total_cases', 'new_cases', 'new_cases_smoothed', 'total_deaths', 'new_deaths', 'new_deaths_smoothed']

Esempio di record:

	0
iso_code	AFG
continent	Asia
location	Afghanistan
date	2020-01-05
total_cases	0.0
...	...
population	41128772
excess_mortality_cumulative_absolute	NaN
excess_mortality_cumulative	NaN
excess_mortality	NaN
excess_mortality_cumulative_per_million	NaN

[67 rows x 1 columns]

```

In [4]: # Pulizia dati: elimino righe senza continente
df_clean = df.dropna(subset=["continent"])

# Ottengo l'ultimo record per ogni nazione
latest = df_clean.sort_values("date").groupby("location").last()

# Calcolo i casi totali per continente
continent_cases = latest.groupby("continent")["total_cases"].sum().sort_values(ascending=False)

# Stampa del risultato
print("Totale casi per continente:")
print(continent_cases)

```

Totale casi per continente:

continent	
Asia	301499099.0
Europe	252916868.0
North America	124492666.0
South America	68809418.0
Oceania	15003352.0
Africa	13145380.0
Name: total_cases, dtype: float64	

```
In [5]: # Calcolo il totale mondiale
total_world = continent_cases.sum()

# Calcolo Le percentuali
percentuali = (continent_cases / total_world) * 100

print("\nPercentuale dei casi per continente:")
print(percentuali.round(2))
```

Percentuale dei casi per continente:

continent	
Asia	38.86
Europe	32.60
North America	16.05
South America	8.87
Oceania	1.93
Africa	1.69

Name: total_cases, dtype: float64

```
In [7]: # 1. Convertiamo la colonna 'date' in formato datetime
df["date"] = pd.to_datetime(df["date"])

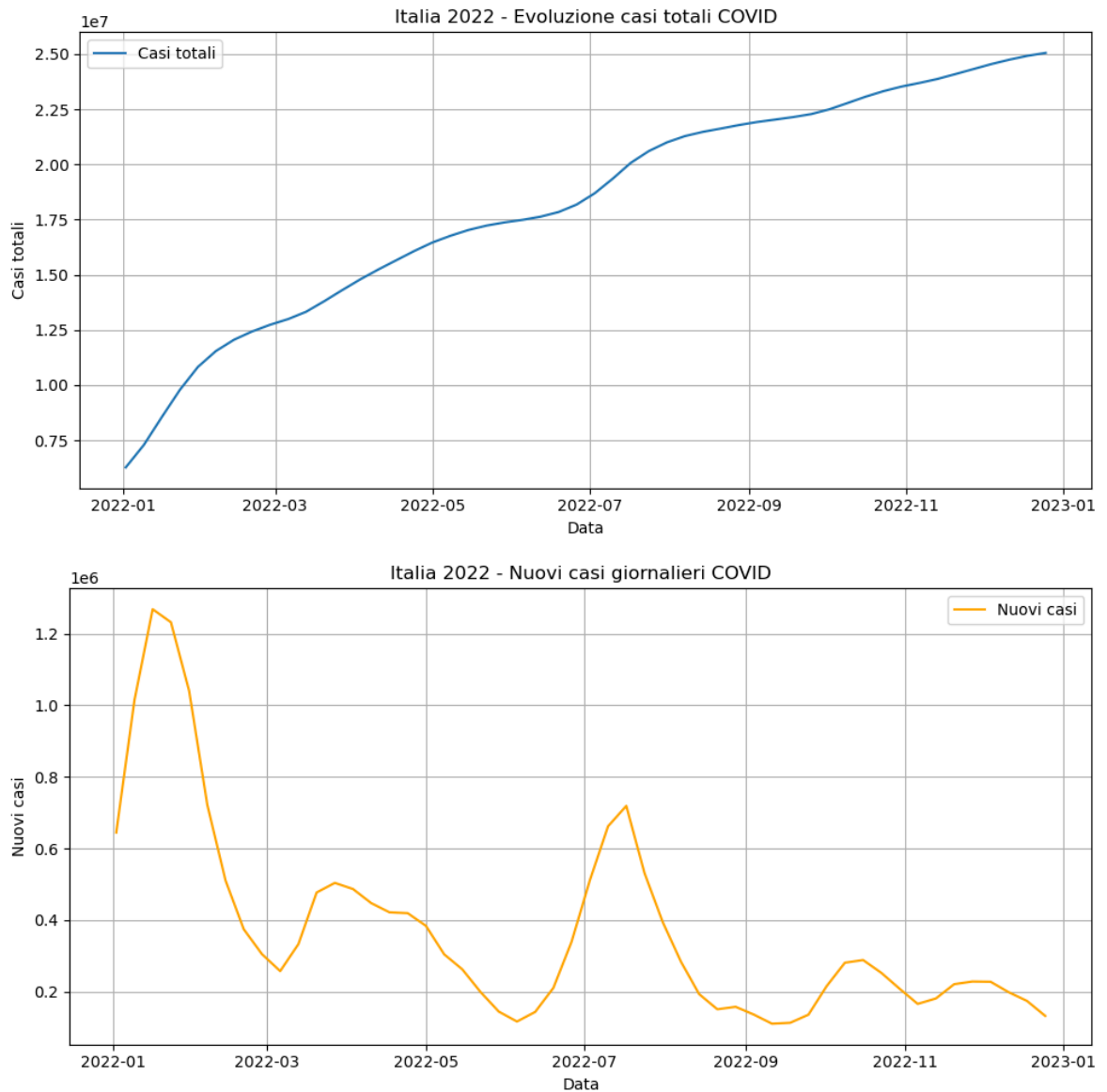
# 2. Selezioniamo solo i dati dell'Italia nel 2022
italy_2022 = df[(df["location"] == "Italy") &
                (df["date"].dt.year == 2022)]

# 3. Rimuoviamo le righe senza nuovi casi (NaN o 0)
italy_2022 = italy_2022[italy_2022["new_cases"].notna()]
italy_2022 = italy_2022[italy_2022["new_cases"] > 0]

import matplotlib.pyplot as plt

plt.figure(figsize=(10, 5))
plt.plot(italy_2022["date"], italy_2022["total_cases"], label="Casi totali")
plt.title("Italia 2022 - Evoluzione casi totali COVID")
plt.xlabel("Data")
plt.ylabel("Casi totali")
plt.grid(True)
plt.legend()
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 5))
plt.plot(italy_2022["date"], italy_2022["new_cases"], label="Nuovi casi", color=
plt.title("Italia 2022 - Nuovi casi giornalieri COVID")
plt.xlabel("Data")
plt.ylabel("Nuovi casi")
plt.grid(True)
plt.legend()
plt.tight_layout()
plt.show()
```



```
In [9]: # Filtro per le 3 nazioni
countries = ["Italy", "Germany", "France"]

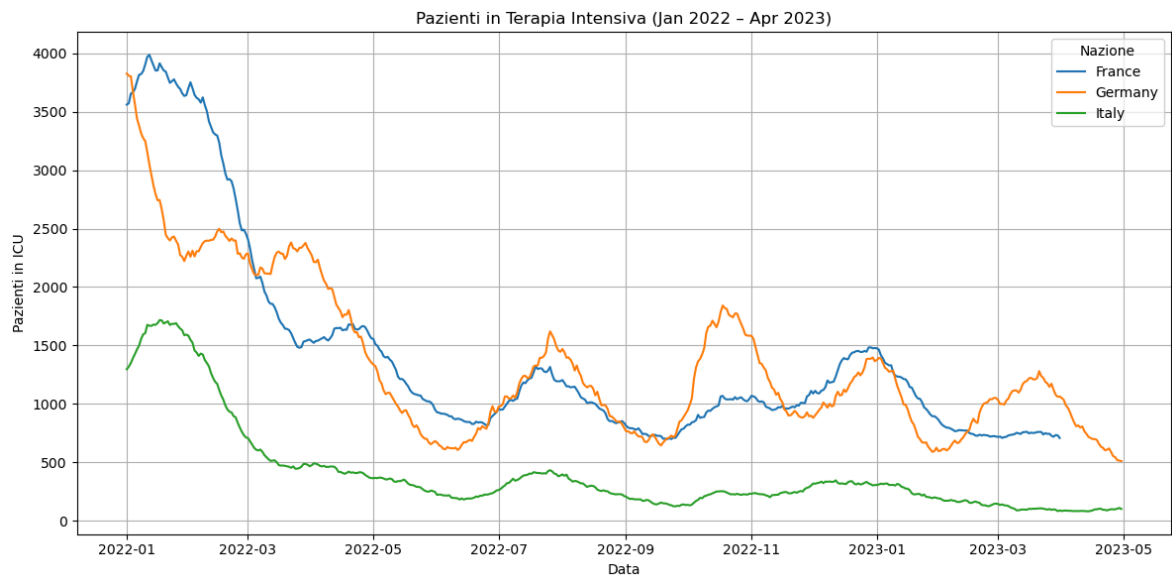
# Selezione intervallo di date
mask = (df["location"].isin(countries)) & \
        (df["date"] >= "2022-01-01") & (df["date"] <= "2023-04-30")

# Filtriamo il dataframe
icu_df = df[mask][["date", "location", "icu_patients"]]

# Rimuoviamo i valori nulli
icu_df = icu_df.dropna()

import seaborn as sns

plt.figure(figsize=(12, 6))
sns.lineplot(data=icu_df, x="date", y="icu_patients", hue="location")
plt.title("Pazienti in Terapia Intensiva (Jan 2022 - Apr 2023)")
plt.xlabel("Data")
plt.ylabel("Pazienti in ICU")
plt.grid(True)
plt.legend(title="Nazione")
plt.tight_layout()
plt.show()
```



```
In [10]: # Paesi da analizzare
countries = ["Italy", "Germany", "France", "Spain"]

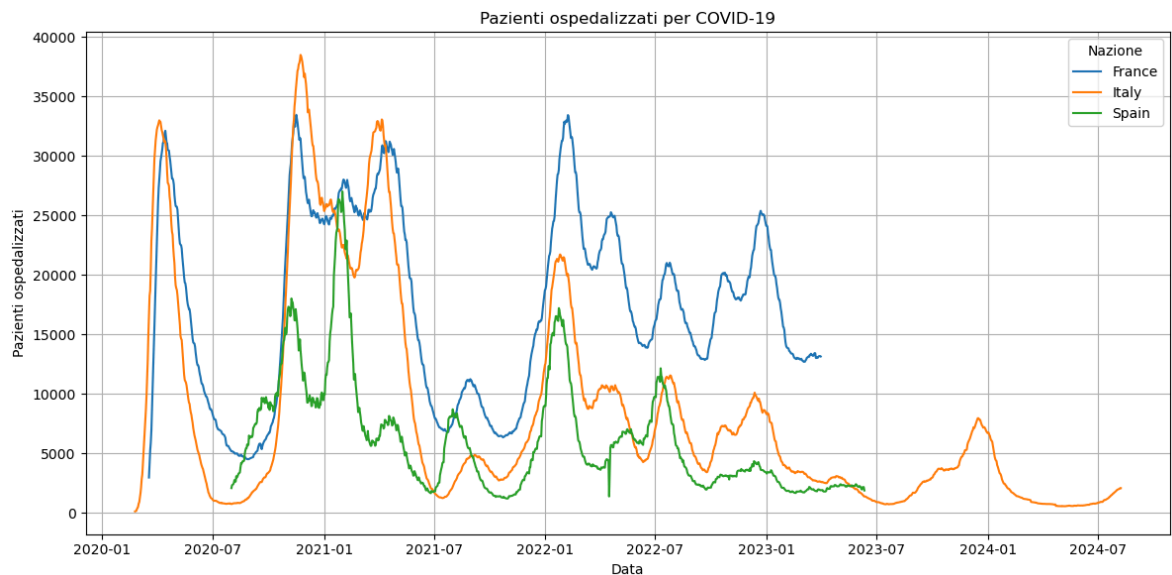
# Filtro per i paesi
mask = df["location"].isin(countries)

# Selezione colonne rilevanti
hosp_df = df[mask][["date", "location", "hosp_patients"]]
```

```
In [11]: import seaborn as sns
import matplotlib.pyplot as plt

# Rimuovo righe con hosp_patients mancanti
hosp_df_clean = hosp_df.dropna()

plt.figure(figsize=(12, 6))
sns.lineplot(data=hosp_df_clean, x="date", y="hosp_patients", hue="location")
plt.title("Pazienti ospedalizzati per COVID-19")
plt.xlabel("Data")
plt.ylabel("Pazienti ospedalizzati")
plt.grid(True)
plt.legend(title="Nazione")
plt.tight_layout()
plt.show()
```



In [12]: *#Il dataset fornito da Our World in Data contiene informazioni su casi e vaccina*

Righe totali: 429.435

Colonne totali: 67

Le colonne includono indicatori sanitari, demografici e temporali come total_cas

#È stata calcolata la somma di total_cases per ciascun continente e la relativa
Questo permette di comprendere la distribuzione geografica della pandemia

IT 3. Italia - Anno 2022

a. Evoluzione settimanale dei casi totali

Sono stati selezionati i dati riferiti all'Italia nel 2022, convertiti in aggreg
Il grafico mostra chiaramente l'andamento delle ondate pandemiche durante l'anno

b. Nuovi casi rispetto alla data **ITDEFR 4. Terapie intensive - Italia, Germania, F**

a. Differenza nel numero di pazienti in terapia intensiva (icu_patients)

Dati aggregati da gennaio 2022 ad aprile 2023 hanno evidenziato che:

L'Italia ha avuto picchi iniziali più alti.

La Germania ha mantenuto un valore più costante nel tempo.

La Francia ha avuto un andamento simile all'Italia ma con fluttuazioni meno marc

b. Osservazioni

Dal grafico si può osservare un progressivo calo delle ospedalizzazioni intensiv

Sono stati visualizzati i new_cases giorno per giorno. Le settimane senza misura

#ITDEFR 4. Terapie intensive - Italia, Germania, Francia

a. Differenza nel numero di pazienti in terapia intensiva (icu_patients)

Dati aggregati da gennaio 2022 ad aprile 2023 hanno evidenziato che: *#ITDEFRES 5. F*

a. Ospedalizzazioni (hosp_patients) per paese

Il grafico ha mostrato l'evoluzione temporale dei pazienti ospedalizzati per Ita
Le curve evidenziano l'andamento delle ondate pandemiche e la differenza di gest

b. Gestione dei valori nulli

I dati hosp_patients presentano valori nulli. Non sono stati sostituiti con zeri
Si è optato per l'uso di dropna() o eventuale interpolazione selettiv

L'Italia ha avuto picchi iniziali più alti.

La Germania ha mantenuto un valore più costante nel tempo.

La Francia ha avuto un andamento simile all'Italia ma con fluttuazioni meno marcate.

b. Osservazioni

Dal grafico si può osservare un progressivo calo delle ospedalizzazioni intensive.

#Conclusione

L'analisi ha permesso di:

Verificare la qualità del dataset.

Comprendere la distribuzione globale e continentale dei casi.

Valutare l'andamento italiano nel 2022.

Confrontare i sistemi sanitari europei in termini di ricoveri e terapie intensive.

Cell In[12], line 11

IT 3. Italia - Anno 2022

^

SyntaxError: invalid character 'i' (U+1F1EE)

In []: