Week 8 assignment

May 14, 2025

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Optional: make plots look better
sns.set(style='whitegrid')
plt.rcParams['figure.figsize'] = (12, 6)
```

[2]:
```
# Load the CSV file
df = pd.read_csv("owid-covid-data.csv")

# Check the first few rows
df.head()
```

[2]:

	iso_code	continent	location	date	total_cases	new_cases	\
0	AFG	Asia	Afghanistan	2020-01-05	0.0	0.0	
1	AFG	Asia	Afghanistan	2020-01-06	0.0	0.0	
2	AFG	Asia	Afghanistan	2020-01-07	0.0	0.0	
3	AFG	Asia	Afghanistan	2020-01-08	0.0	0.0	
4	AFG	Asia	Afghanistan	2020-01-09	0.0	0.0	

	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	…	\
0	NaN	0.0	0.0	NaN	…	
1	NaN	0.0	0.0	NaN	…	
2	NaN	0.0	0.0	NaN	…	
3	NaN	0.0	0.0	NaN	…	
4	NaN	0.0	0.0	NaN	…	

	male_smokers	handwashing_facilities	hospital_beds_per_thousand	\
0	NaN	37.746	0.5	
1	NaN	37.746	0.5	
2	NaN	37.746	0.5	
3	NaN	37.746	0.5	
4	NaN	37.746	0.5	

	life_expectancy	human_development_index	population	\
0	64.83	0.511	41128772	

```
1            64.83                    0.511    41128772
2            64.83                    0.511    41128772
3            64.83                    0.511    41128772
4            64.83                    0.511    41128772

   excess_mortality_cumulative_absolute  excess_mortality_cumulative  \
0                                   NaN                          NaN
1                                   NaN                          NaN
2                                   NaN                          NaN
3                                   NaN                          NaN
4                                   NaN                          NaN

   excess_mortality  excess_mortality_cumulative_per_million
0               NaN                                      NaN
1               NaN                                      NaN
2               NaN                                      NaN
3               NaN                                      NaN
4               NaN                                      NaN

[5 rows x 67 columns]
```

[3]:
```
eastern_africa = ['Kenya', 'Uganda', 'Tanzania', 'Rwanda', 'Ethiopia',␣
 ↪'Somalia']

# Keep only rows from selected countries
df = df[df['location'].isin(eastern_africa)]

# Convert date column to datetime
df['date'] = pd.to_datetime(df['date'])

# Check the data again
df[['date', 'location', 'total_cases', 'total_deaths', 'total_vaccinations']].
 ↪head()
```

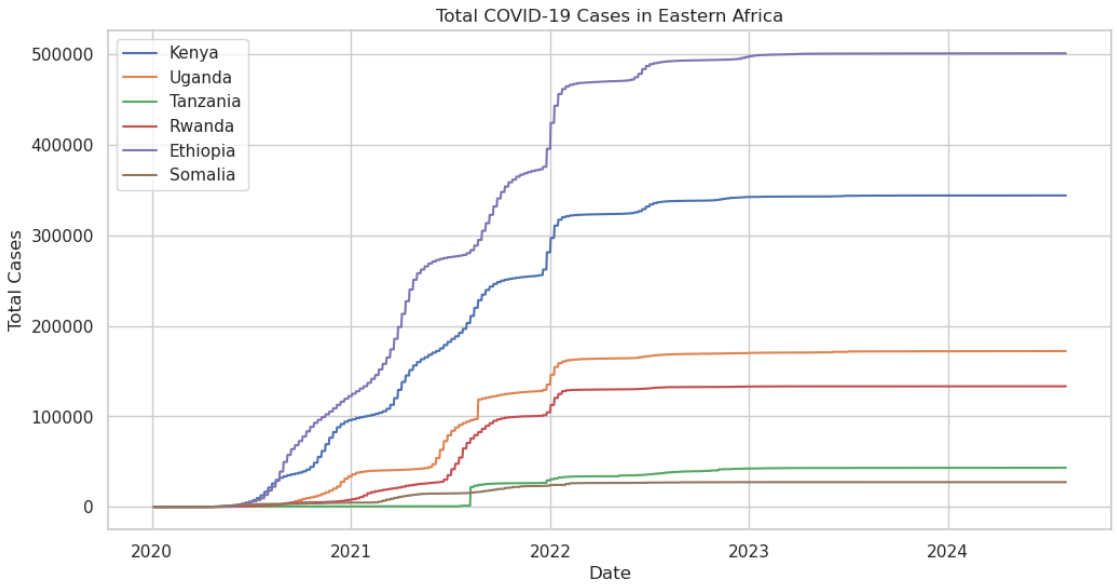
[3]:
```
             date  location  total_cases  total_deaths  total_vaccinations
116895 2020-01-05  Ethiopia          0.0           0.0                 NaN
116896 2020-01-06  Ethiopia          0.0           0.0                 NaN
116897 2020-01-07  Ethiopia          0.0           0.0                 NaN
116898 2020-01-08  Ethiopia          0.0           0.0                 NaN
116899 2020-01-09  Ethiopia          0.0           0.0                 NaN
```

[4]:
```
# Fill missing values with 0 (simple method for beginners)
df[['total_cases', 'total_deaths', 'total_vaccinations']] = df[['total_cases',␣
 ↪'total_deaths', 'total_vaccinations']].fillna(0)
```

[5]:
```
for country in eastern_africa:
    subset = df[df['location'] == country]
```

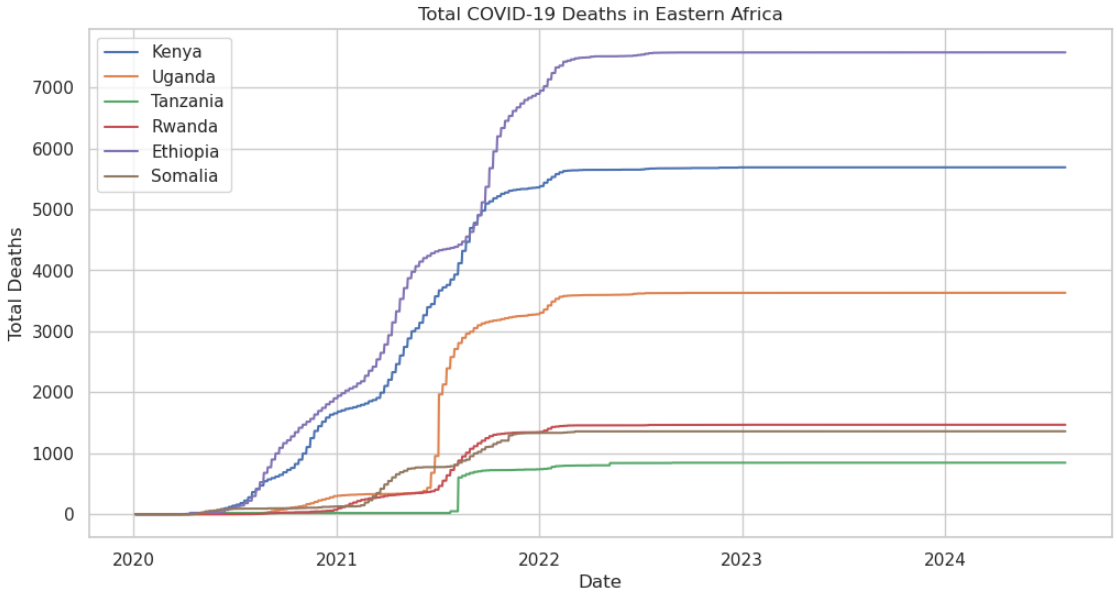
```
    plt.plot(subset['date'], subset['total_cases'], label=country)

plt.title("Total COVID-19 Cases in Eastern Africa")
plt.xlabel("Date")
plt.ylabel("Total Cases")
plt.legend()
plt.show()
```



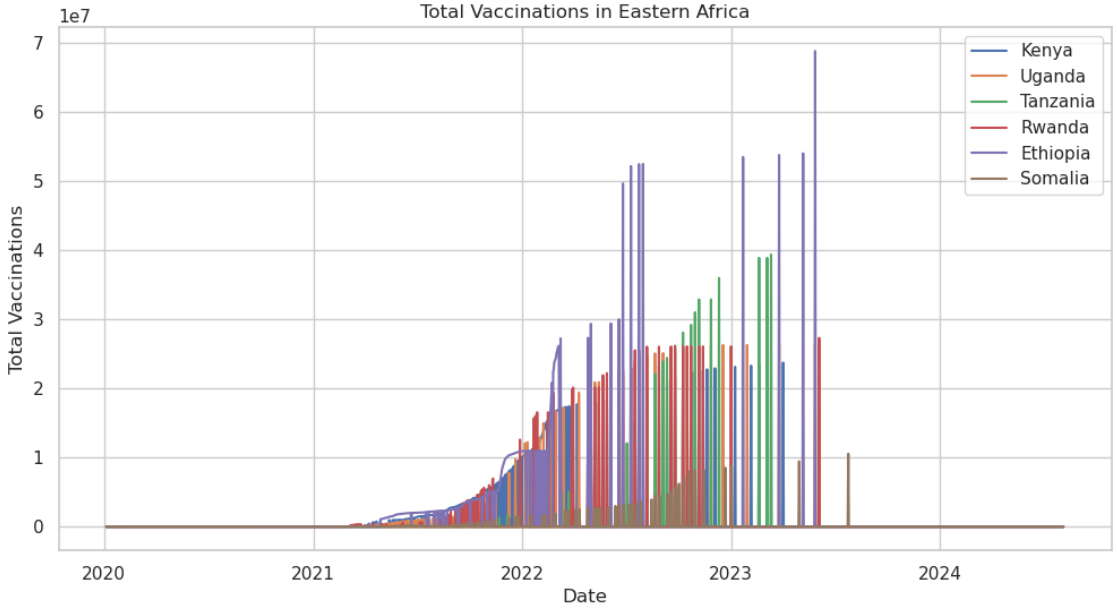
[6]:
```
for country in eastern_africa:
    subset = df[df['location'] == country]
    plt.plot(subset['date'], subset['total_deaths'], label=country)

plt.title("Total COVID-19 Deaths in Eastern Africa")
plt.xlabel("Date")
plt.ylabel("Total Deaths")
plt.legend()
plt.show()
```



[7]:
```
for country in eastern_africa:
    subset = df[df['location'] == country]
    plt.plot(subset['date'], subset['total_vaccinations'], label=country)

plt.title("Total Vaccinations in Eastern Africa")
plt.xlabel("Date")
plt.ylabel("Total Vaccinations")
plt.legend()
plt.show()
```



[ ]:
```
Key Insights
- Kenya had the highest total vaccinations by late 2021.
- Uganda and Ethiopia experienced spikes in cases around mid-2021.
- Somalia had fewer reported vaccinations compared to others.
```