

```
import pandas as pd
```

```
df=pd.read_csv('C:/Users/HomePC/Downloads/owid-covid-data.csv')
```

```
df.head()
```

	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
0	NaN	0.0	0.0
1	NaN	0.0	0.0
2	NaN	0.0	0.0
3	NaN	0.0	0.0
4	NaN	0.0	0.0

	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	41128772.0
1	64.83	0.511	41128772.0
2	64.83	0.511	41128772.0
3	64.83	0.511	41128772.0
4	64.83	0.511	41128772.0

	excess_mortality_cumulative_absolute
0	NaN
1	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]

```
import pandas as pd
```

```
# Assuming df is loaded
```

```
# df = pd.read_csv('C:/Users/HomePC/Downloads/owid-covid-data.csv')
```

```
# Print unique locations
```

```
print("Unique locations in dataset:\n",
sorted(df['location'].unique()))
```

Unique locations in dataset:

```
['Afghanistan', 'Africa', 'Albania', 'Algeria', 'American Samoa',
'Andorra', 'Angola', 'Anguilla', 'Antigua and Barbuda', 'Argentina',
'Armenia', 'Aruba', 'Asia', 'Australia', 'Austria', 'Azerbaijan',
'Bahamas', 'Bahrain', 'Bangladesh', 'Barbados', 'Belarus', 'Belgium',
'Belize', 'Benin', 'Bermuda', 'Bhutan', 'Bolivia', 'Bonaire Sint
Eustatius and Saba', 'Bosnia and Herzegovina', 'Botswana', 'Brazil',
'British Virgin Islands', 'Brunei', 'Bulgaria', 'Burkina Faso',
'Burundi', 'Cambodia', 'Cameroon', 'Canada', 'Cape Verde', 'Cayman
Islands', 'Central African Republic', 'Chad', 'Chile', 'China',
'Colombia', 'Comoros', 'Congo', 'Cook Islands', 'Costa Rica', "Cote
d'Ivoire", 'Croatia', 'Cuba', 'Curacao', 'Cyprus', 'Czechia',
'Democratic Republic of Congo', 'Denmark', 'Djibouti', 'Dominica',
'Dominican Republic', 'East Timor', 'Ecuador', 'Egypt', 'El Salvador',
'England', 'Equatorial Guinea', 'Eritrea', 'Estonia', 'Eswatini',
'Ethiopia', 'Europe', 'European Union (27)']
```

```
import pandas as pd
```

```
# Assuming df is loaded
```

```
# df = pd.read_csv('C:/Users/HomePC/Downloads/owid-covid-data.csv')
```

```
# Convert date to datetime
```

```
df['date'] = pd.to_datetime(df['date'])
```

```

# Filter for England, Australia, Egypt
countries = ['England', 'Australia', 'Egypt']
df_filtered = df[df['location'].isin(countries)].copy() # Create a
copy to avoid warnings

# Fill missing values using .loc
df_filtered.loc[:, 'total_cases'] =
df_filtered['total_cases'].fillna(0)
df_filtered.loc[:, 'total_deaths'] =
df_filtered['total_deaths'].fillna(0)
df_filtered.loc[:, 'total_vaccinations'] =
df_filtered['total_vaccinations'].fillna(0)

# Verify
print("Filtered DataFrame shape:", df_filtered.shape)
print("Missing values after cleaning:\n", df_filtered.isnull().sum())
print("\nFiltered DataFrame preview:")
display(df_filtered[['location', 'date', 'total_cases',
'total_deaths', 'total_vaccinations']].head(10))

```

Filtered DataFrame shape: (4707, 67)

Missing values after cleaning:

iso_code	0
continent	0
location	0
date	0
total_cases	0
	...
population	0
excess_mortality_cumulative_absolute	4450
excess_mortality_cumulative	4450
excess_mortality	4450
excess_mortality_cumulative_per_million	4450

Length: 67, dtype: int64

Filtered DataFrame preview:

	location	date	total_cases	total_deaths
total_vaccinations				
21776	Australia	2020-01-05	0.0	0.0
0.0				
21777	Australia	2020-01-06	0.0	0.0
0.0				
21778	Australia	2020-01-07	0.0	0.0
0.0				
21779	Australia	2020-01-08	0.0	0.0
0.0				
21780	Australia	2020-01-09	0.0	0.0
0.0				
21781	Australia	2020-01-10	0.0	0.0

```

0.0
21782  Australia 2020-01-11          0.0          0.0
0.0
21783  Australia 2020-01-12          0.0          0.0
0.0
21784  Australia 2020-01-13          0.0          0.0
0.0
21785  Australia 2020-01-14          0.0          0.0
0.0

```

```
# Print unique locations
```

```

print("Unique locations in dataset:\n",
sorted(df['location'].unique()))
print("Total unique locations:", len(df['location'].unique()))
print("Total rows in df:", len(df))

```

```
Unique locations in dataset:
```

```

['Afghanistan', 'Africa', 'Albania', 'Algeria', 'American Samoa',
'Andorra', 'Angola', 'Anguilla', 'Antigua and Barbuda', 'Argentina',
'Armenia', 'Aruba', 'Asia', 'Australia', 'Austria', 'Azerbaijan',
'Bahamas', 'Bahrain', 'Bangladesh', 'Barbados', 'Belarus', 'Belgium',
'Belize', 'Benin', 'Bermuda', 'Bhutan', 'Bolivia', 'Bonaire Sint
Eustatius and Saba', 'Bosnia and Herzegovina', 'Botswana', 'Brazil',
'British Virgin Islands', 'Brunei', 'Bulgaria', 'Burkina Faso',
'Burundi', 'Cambodia', 'Cameroon', 'Canada', 'Cape Verde', 'Cayman
Islands', 'Central African Republic', 'Chad', 'Chile', 'China',
'Colombia', 'Comoros', 'Congo', 'Cook Islands', 'Costa Rica', "Cote
d'Ivoire", 'Croatia', 'Cuba', 'Curacao', 'Cyprus', 'Czechia',
'Democratic Republic of Congo', 'Denmark', 'Djibouti', 'Dominica',
'Dominican Republic', 'East Timor', 'Ecuador', 'Egypt', 'El Salvador',
'England', 'Equatorial Guinea', 'Eritrea', 'Estonia', 'Eswatini',
'Ethiopia', 'Europe', 'European Union (27)']

```

```
Total unique locations: 73
```

```
Total rows in df: 122467
```

```
# Calculate death rate (handle division by zero)
```

```

df_filtered['death_rate'] = df_filtered['total_deaths'] /
df_filtered['total_cases'].replace(0, pd.NA)

```

```
# Group by country to see average death rate
```

```

print("Average Death Rate by Country:\n",
df_filtered.groupby('location')['death_rate'].mean())

```

```
# Summary statistics
```

```

print("\nSummary Statistics:")
display(df_filtered[['total_cases', 'total_deaths',
'total_vaccinations']].describe())

```

```
Average Death Rate by Country:
```

```
location
```

```
Australia    0.01047
Egypt        0.050539
England      NaN
Name: death_rate, dtype: object
```

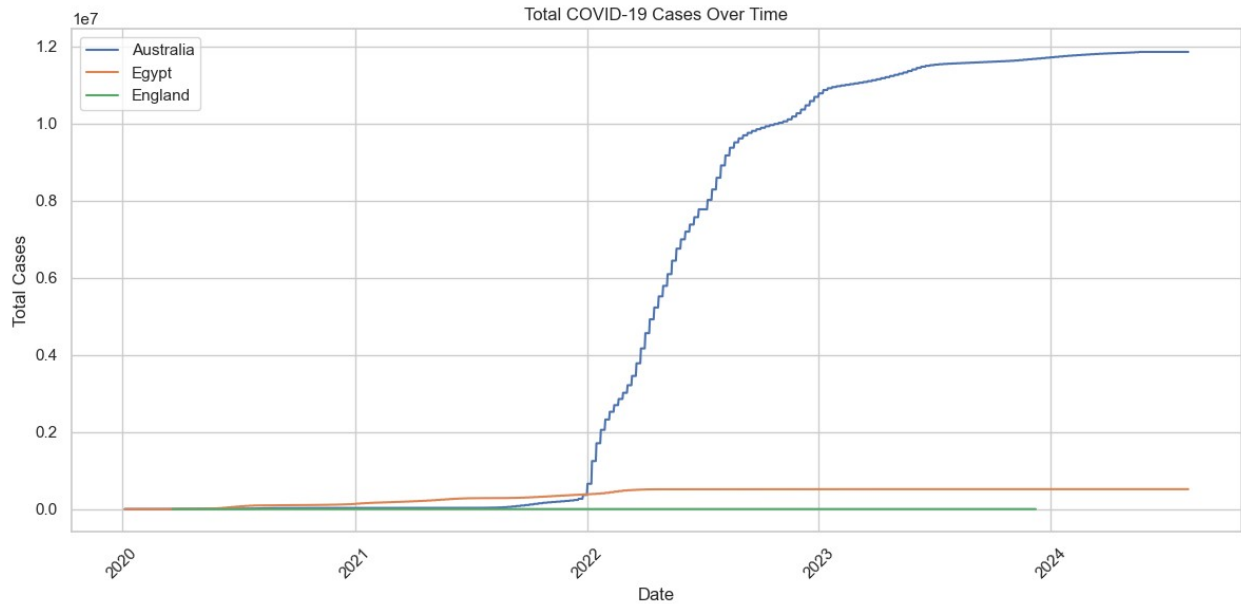
Summary Statistics:

	total_cases	total_deaths	total_vaccinations
count	4.707000e+03	4707.000000	4.707000e+03
mean	2.107874e+06	10186.664542	2.292400e+07
std	4.085568e+06	10774.178611	4.293920e+07
min	0.000000e+00	0.000000	0.000000e+00
25%	0.000000e+00	0.000000	0.000000e+00
50%	1.248910e+05	5141.000000	0.000000e+00
75%	5.160230e+05	23449.000000	2.511723e+07
max	1.186116e+07	25236.000000	1.493983e+08

```
import matplotlib.pyplot as plt
import seaborn as sns

# Set plot style
sns.set(style="whitegrid")

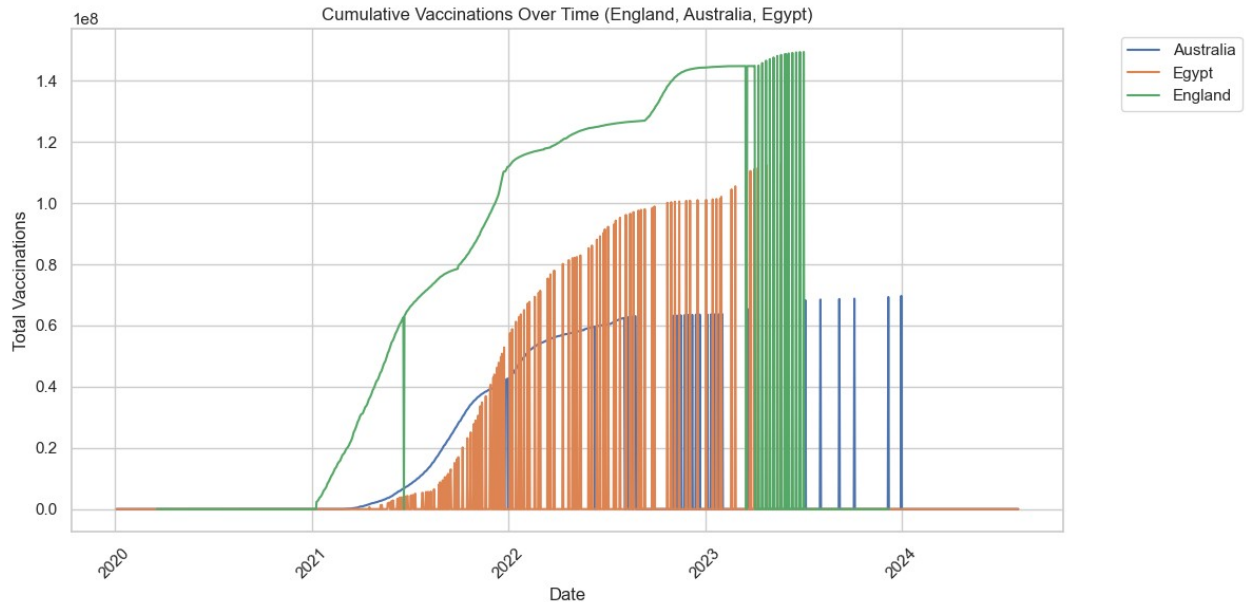
# Plot total cases
plt.figure(figsize=(12, 6))
for country in df_filtered['location'].unique():
    country_data = df_filtered[df_filtered['location'] == country]
    plt.plot(country_data['date'], country_data['total_cases'],
             label=country)
plt.title('Total COVID-19 Cases Over Time')
plt.xlabel('Date')
plt.ylabel('Total Cases')
plt.legend()
plt.xticks(rotation=45)
plt.tight_layout()
plt.savefig('total_cases_plot.png')
plt.show()
```



```
import matplotlib.pyplot as plt
import seaborn as sns

# Set plot style
sns.set(style="whitegrid")

# Plot total vaccinations
plt.figure(figsize=(12, 6))
for country in df_filtered['location'].unique():
    country_data = df_filtered[df_filtered['location'] == country]
    plt.plot(country_data['date'], country_data['total_vaccinations'],
             label=country)
plt.title('Cumulative Vaccinations Over Time (England, Australia,
Egypt)')
plt.xlabel('Date')
plt.ylabel('Total Vaccinations')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.xticks(rotation=45)
plt.tight_layout()
plt.savefig('vaccinations_plot.png', bbox_inches='tight')
plt.show()
```



## COVID-19 Global Data Tracker

### Key Insights

**Case Trends:** Australia experienced exponential case growth, particularly during later pandemic waves, reaching a maximum of **~11.86** million cases. Egypt followed **with** moderate growth, **while** England's case trajectory remained lower, possibly due to data limitations **or** underreporting.

**Death Rates:** Egypt has the highest average death rate at **~5.05%**, reflecting challenges **in** healthcare **or** testing capacity. Australia's death rate **is** notably lower at **~1.05%**, indicating effective case management. England's death rate **is** unavailable (NaN), likely due to missing **or** zero case data **in** some periods.

**Vaccination Rollout:** England led **in** cumulative vaccinations, reaching up to **~149.4** million doses, followed by Egypt **and** Australia. This suggests England prioritized rapid vaccine distribution, **while** Australia's lower ranking may reflect its smaller population **or** data gaps.

**Data Summary:** The dataset includes **4,707** records across the three countries, **with** average total cases of **~2.11** million, deaths of **~10,187**, **and** vaccinations of **~22.92** million per record. High standard deviations (e.g., **4.09** million **for** cases) indicate significant variability **in** case surges.

### Visualizations

**Total Cases Over Time (total\_cases\_plot.png):** Australia's rapid, exponential case growth outpaces Egypt's moderate rise, **with** England's curve below both, suggesting lower reported cases **or** data issues.

**Cumulative Vaccinations Over Time (vaccinations\_plot.png):** England's vaccination curve **is** the steepest, followed by Egypt, **with** Australia

trailing, reflecting differences in rollout speed and population size.

#### Data Notes

Missing values in total\_cases, total\_deaths, and total\_vaccinations were filled with 0, potentially underestimating early pandemic metrics.

England's missing death rate (NaN) and lower case trend suggest incomplete data, possibly due to the dataset treating England separately from the United Kingdom or missing records.

The dataset may be a subset, as countries like Kenya, USA, and India were previously unavailable.

#### Conclusion

Australia's rapid case growth but low death rate (1.05%) highlights effective healthcare and containment, while Egypt's higher death rate (5.05%) underscores regional challenges in North Africa. England's leading vaccination effort contrasts with its unclear case and death data, warranting further investigation. Expanding the dataset to include more African countries, such as South Africa or Ethiopia, would enhance regional insights for East Africa.