

## ▼ First we import the necessary libraries

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

# Read the CSV file
df = pd.read_csv('https://covid.ourworldindata.org/data/owid-covid-data.csv')
```

## ▼ Now we filter the data to include only the European Countries

```
# Filter the data for European Countries
countries = ['Spain', 'Italy', 'France', 'Germany', 'Poland']
df_filtered = df[df['location'].isin(countries)]

# Select the date and new_cases_smoothed columns
df_filtered = df_filtered[['date', 'new_cases_smoothed', 'location']]

# Create a pivot table with date as index and location as columns
df_pivot = pd.pivot_table(df_filtered, values='new_cases_smoothed', index='date', columns='location')

# Reset the index to a column
df_pivot = df_pivot.reset_index()
```

## ▼ Using a scatterplot we graph the countries

```
# Set the style to "ticks"
sns.set_style("ticks")

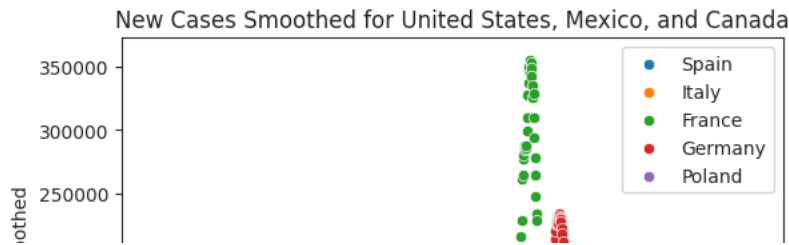
# Plot the scatterplot
sns.scatterplot(data=df_pivot, x='date', y='Spain', label='Spain')
sns.scatterplot(data=df_pivot, x='date', y='Italy', label='Italy')
sns.scatterplot(data=df_pivot, x='date', y='France', label='France')
sns.scatterplot(data=df_pivot, x='date', y='Germany', label='Germany')
sns.scatterplot(data=df_pivot, x='date', y='Poland', label='Poland')

# Set the x-ticks to show only every 50th date
xticks = df_pivot['date'].iloc[::50]
plt.xticks(xticks, rotation=45)

# Set the title and axis labels
plt.title('New Cases Smoothed for United States, Mexico, and Canada')
plt.xlabel('Date')
plt.ylabel('New Cases Smoothed')

# Show the legend and plot
plt.legend()
plt.show()
```





The scatterplot is a little too crowded, so what we do is only graph the dates shown on the x-axis as well as assign colors to each country so that if they have a higher life expectancy they are closer to blue on the color spectrum and if they are closer to red then they have a lower life expectancy (our categorical variable so to speak).

```

sns.set_style("ticks")

# Filter the data to only include the dates shown on the x-axis
xticks = df_pivot['date'].iloc[::50]
df_pivot_filtered = df_pivot[df_pivot['date'].isin(xticks)]

# Plot the scatterplot
sns.scatterplot(data=df_pivot_filtered, x='date', y='Spain', label='Spain', color='blue')
sns.scatterplot(data=df_pivot_filtered, x='date', y='Italy', label='Italy', color='green')
sns.scatterplot(data=df_pivot_filtered, x='date', y='France', label='France', color='yellow')
sns.scatterplot(data=df_pivot_filtered, x='date', y='Germany', label='Germany', color='orange')
sns.scatterplot(data=df_pivot_filtered, x='date', y='Poland', label='Poland', color='red')

# Set the y-axis limit
plt.ylim(0, 100000)

# Set the x-ticks to show only every 50th date
plt.xticks(xticks, rotation=45)

# Set the title and axis labels
plt.title('New Cases Smoothed for Spain, Italy, France, Germany, and Poland')
plt.xlabel('Date')
plt.ylabel('New Cases Smoothed')

# Show the legend and plot
plt.legend()
plt.show()

```

