Data Analysis

As stated in our introduction, our research question is: **How do population dynamics compare between countries of different income-levels?** To answer this question, we will examine three different indicators: life expectancy, adolescent fertility, and under-5 mortality rate.

We'll be using the pandas, matplotlib, and seaborn libraries for our analysis.

```
# Import Libraries and Read Cleaned Dataset

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

df = pd.read_csv('population_dynamics_clean.csv')
```

Life Expectancy Analysis

We'll begin by looking at life expectancy over time by country. Figure 1 is a plot of life expectancy years from 2000 - 2023.

Based on the different lines, we see a generally increasing trend in life expectancy for all 5 countries. Additionally, there is a signficant dip in life expectancy for all countries around 2020-2021. This is likely an effect of the COVID-19 global pandemic. According to the World Health Organization, global life expectancy deceased 1.8 years to 71.4 years from 2019 to 2021, which is similar to the life expectancy we saw in 2012. This, which we can see in Figure 1, is approximately true for our selection of countries.

```
# Life Expectancy Rate by Country Over Time

plt.figure(figsize=(10, 6))
sns.lineplot(
```

```
    data=df,
    x="year",
    y="life_expectancy",
    hue="country",
    marker="o"
)

plt.title("Life Expectancy Over Time by Country", fontsize=14)
plt.xlabel("Year")
plt.ylabel("Life Expectancy (Years)")
plt.legend(title="Country")
plt.tight_layout()
plt.show()
```

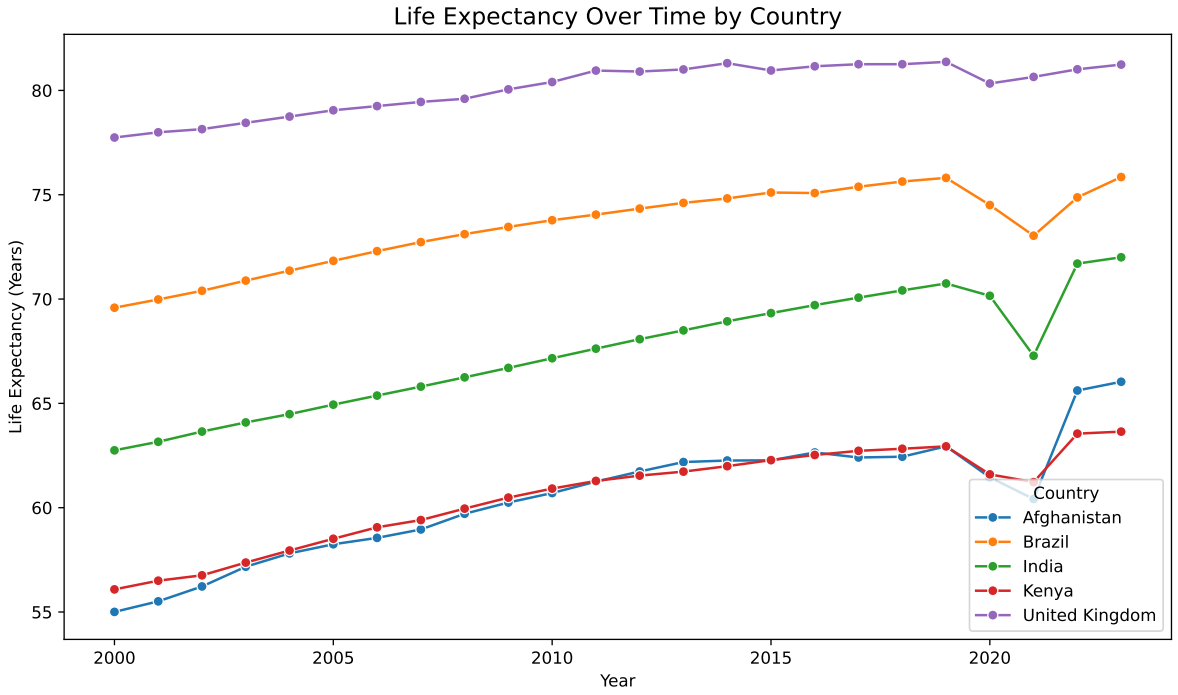

Figure 1: Life Expectancy from 2000 - 2023

While we see some general parallels, the life expectancy within these countries still vary drastically. Figure 2 depicts the selection of countries in order, from low-income, to lower-middle, to upper-middle, and to high. The height of the bars represents average life expectancy from more recent years, 2017 - 2023. People in high income countries have higher life expectancies

than those in low income countries. In other words, life expectancy seems to be positively linked with a country's GNI.

```
# Bar plot of average life expectancy per country
avg_life_exp = (
    df[df['year'].between(2017, 2023)]
    .groupby("country")["life_expectancy"]
    .mean()
    .reset_index()
)

country_order = ['Afghanistan', 'Kenya', 'India', 'Brazil', 'United Kingdom']

plt.figure(figsize=(8, 5))
sns.barplot(data=avg_life_exp, x="country", y="life_expectancy", order=country_order)
plt.title("Average Life Expectancy (2017-2023) by Country")
plt.ylabel("Average Life Expectancy (Years)")
plt.xlabel("Country")
plt.tight_layout()
plt.show()
```

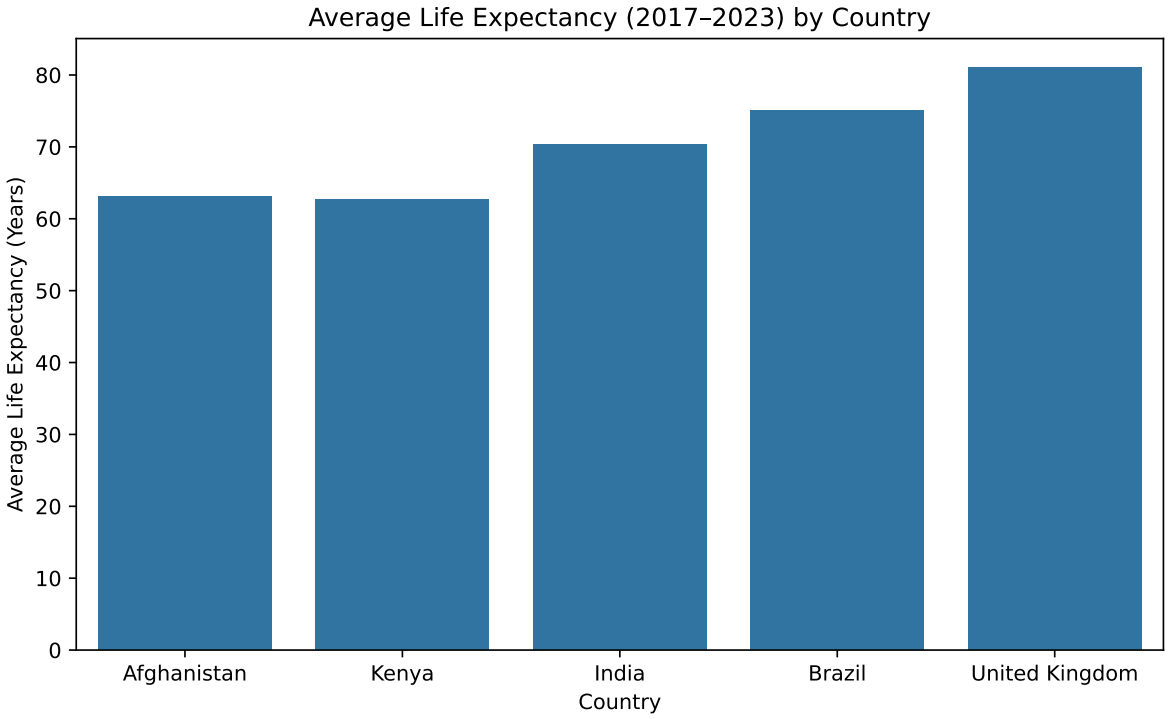

Figure 2: Average Life Expectancy from 2017 - 2023

A table of these average life expectancies is also shown in Table 1. High-income countries, such as the United Kingdom, have generally higher life expectancy. The United Kingdom life expectancy hovers near the top edge of the plot in Figure 1, and in Table 1, we see that it has a high average life expectancy of around 81.0 years. However, for low-income countries, such as Afghanistan, we see a significantly lower life expectancy of around 63.0 years. Thus, again, there appears to be a rather strong correlation between a country's income level and their life expectancy.

```
# Average Life Expectancy by Country

table_1 = (
    df[df['year'].between(2017, 2023)]
    .groupby('country', as_index=False)['life_expectancy']
    .mean()
    .round(3)
)
table_1.columns = ['Country', 'Average Life Expectancy (2017-2023)']
table_1
```

Table 1

	Country	Average Life Expectancy (2017–2023)
0	Afghanistan	63.045
1	Brazil	75.013
2	India	70.338
3	Kenya	62.644
4	United Kingdom	81.016

Average Life Expectancy from 2017 - 2023

Adolescent Fertility Analysis

Our second indicator for population dynamics is adolescent fertility rate. Figure 3 shows adolescent fertility rates over time for our selection of countries. There appear to be drastic decreases for lower and lower-middle countries, such as India. Notably, India has high adolescent fertility rates, comparable to those of Afghanistan and Kenya in 2000, but decreased quickly, and currently has a low adolescent fertility rate, one similar to that of the United Kingdom. India's improvement, as well as the decreases in adolescent fertility rates for other non-high-income countries (Afghanistan, Kenya, and Brazil) demonstrates a level of global improvement on this indicator.

```
# Adolescent Fertility Rate by Country Over Time

# Prepare the data for plotting
pivot_df = df.pivot(index='year', columns='country', values='adolescent_fertility')

# Plot the line graph
plt.figure(figsize=(12, 8))
pivot_df.plot(ax=plt.gca(), marker='o')
plt.title('Adolescent Fertility Rate by Country Over Time')
plt.xlabel('Year')
plt.ylabel('Adolescent Fertility Rate')
plt.legend(title='Country', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.xticks(pivot_df.index, rotation=45)
plt.tight_layout()
plt.show()
```

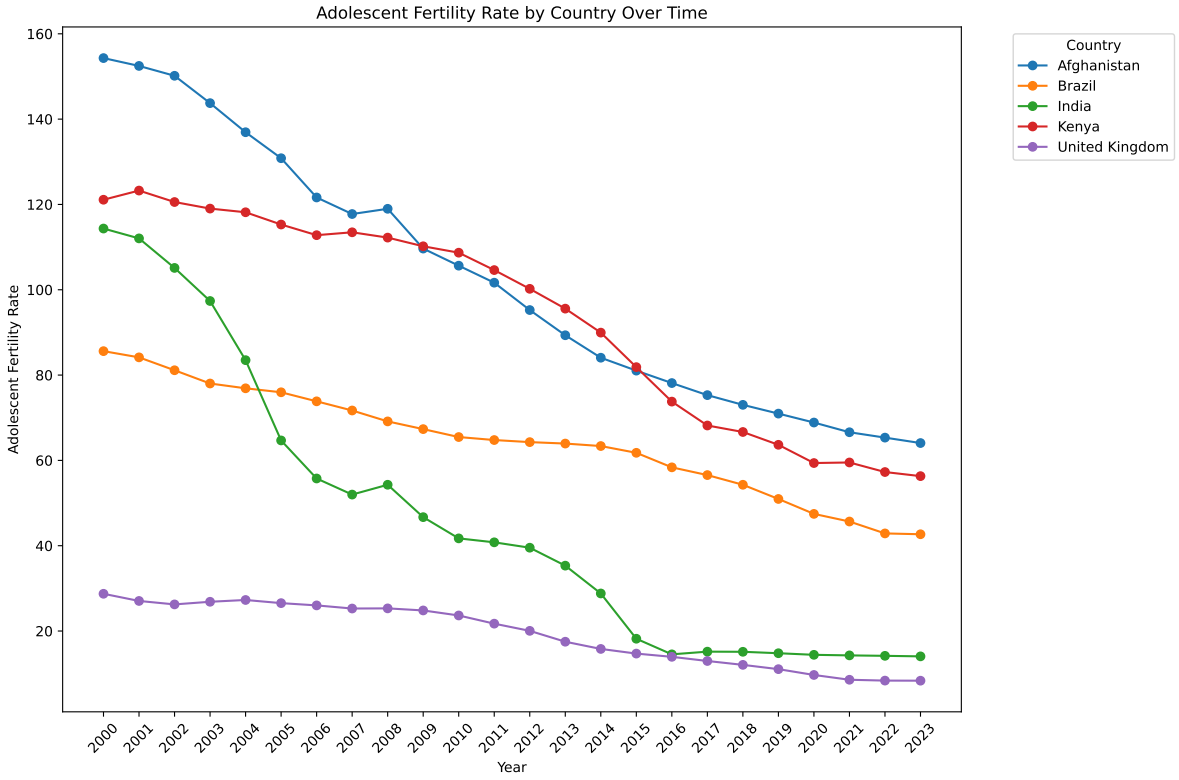

Figure 3: Adolescent Fertility Rates from 2000 - 2023

Adolescent Fertility and Secondary School Enrollment

Still, as shown in Table 2, there are signficantly wide gaps in adolescent fertility rates between lower- and higher-income countries. The United Kingdom has an average rate of approximately 10.1, whereas Afghanistan has a rate of around 70. This is a drastic difference, and one that still needs contionous attention and amendment.

```
# Average Adolescent Fertility Rate by Country

table_2 = (
    df[df['year'].between(2017, 2023)]
    .groupby('country', as_index=False)['adolescent_fertility']
    .mean()
    .round(3)
)
table_2.columns = ['Country', 'Average Adolescent Fertility (2017-2023)']
table_2
```

Table 2

	Country	Average Adolescent Fertility (2017–2023)
0	Afghanistan	69.167
1	Brazil	48.643
2	India	14.579
3	Kenya	61.568
4	United Kingdom	10.161

Average Adolescent Fertility Rates from 2017 - 2023

In addition to looking at adolescent fertility rates over time, we will examine how these rates interact with secondary school enrollment.

Adolescent fertility rates consider girls ages 15 - 19. Most of these years align with secondary-school-enrollment years, which is typically around the 14 - 18 age range. Education can better knowledge of reproductive health and the importance of contraception, and empower women to focus on their academics and career, which may delay child marriage. So, we will observe how education (secondary school enrollment, in particular) interacts with adolescent fertility rates in different countries.

In addition, it's important to note that a country's access to education generally correlates with their GNI, as more money can be invested in schools and teachers. On the other hand, greater access to education is known to improve a country's economy as well, by increasing human capital and innovation.

Figure 4 shows the adolescent fertility and secondary-school-enrollment rates over time. We chose the countries of India and the United Kingdom based on their income-level designations (lower-middle and high, respectively) as well as their availablity of secondary-school-enrollment data compared to other countries.

For India, as secondary school enrollment steadily increases, the adolescent fertility rate decreases quite drastically, suggesting an inverse correlation between these two variables. For the United Kingdom, we see an increase in enrollment rates through the years, as well as a slight decrease in adolescent fertility, reflecting said inverse correlation. Additionally, enrollment rates have been relatively high and adolescent fertility rates low from 2000 - 2023, further emphasizing this inverse correlation. Considering both countries' plots, increasing secondary school enrollment is likely highly associated with decreasing adolescent fertility rates.

```
# Adolescent Fertility Rate and Secondary School Enrollment

# Prepare the data for plotting
pivot_df_adolescent_fertility = df.pivot(index='year', columns='country', values='adolescent
pivot_df_secondary_school_enrollment = df.pivot(index='year', columns='country', values='sec
```

```
# Plot the graphs
fig, axes = plt.subplots(1, 2, figsize=(14, 6), sharex=True, sharey=True)

# India
axes[0].plot(
    pivot_df_adolescent_fertility.index,
    pivot_df_adolescent_fertility['India'],
    marker='o',
    label='Adolescent Fertility Rate',
    color='tab:blue'
)
axes[0].plot(
    pivot_df_secondary_school_enrollment.index,
    pivot_df_secondary_school_enrollment['India'],
    marker='s',
    label='Secondary School Enrollment',
    color='tab:orange'
)
axes[0].set_title('India')
axes[0].set_xlabel('Year')
axes[0].set_ylabel('Rate')
axes[0].legend()
axes[0].tick_params(axis='x', rotation=45)

# United Kingdom
axes[1].plot(
    pivot_df_adolescent_fertility.index,
    pivot_df_adolescent_fertility['United Kingdom'],
    marker='o',
    label='Adolescent Fertility Rate',
    color='tab:blue'
)
axes[1].plot(
    pivot_df_secondary_school_enrollment.index,
    pivot_df_secondary_school_enrollment['United Kingdom'],
    marker='s',
    label='Secondary School Enrollment',
    color='tab:orange'
)
axes[1].set_title('United Kingdom')
axes[1].set_xlabel('Year')
axes[1].legend()
```

```
axes[1].tick_params(axis='x', rotation=45)

fig.suptitle('Adolescent Fertility Rate and Secondary School Enrollment\n(India and United K
plt.tight_layout(rect=[0, 0.03, 1, 0.95])
plt.show()
```

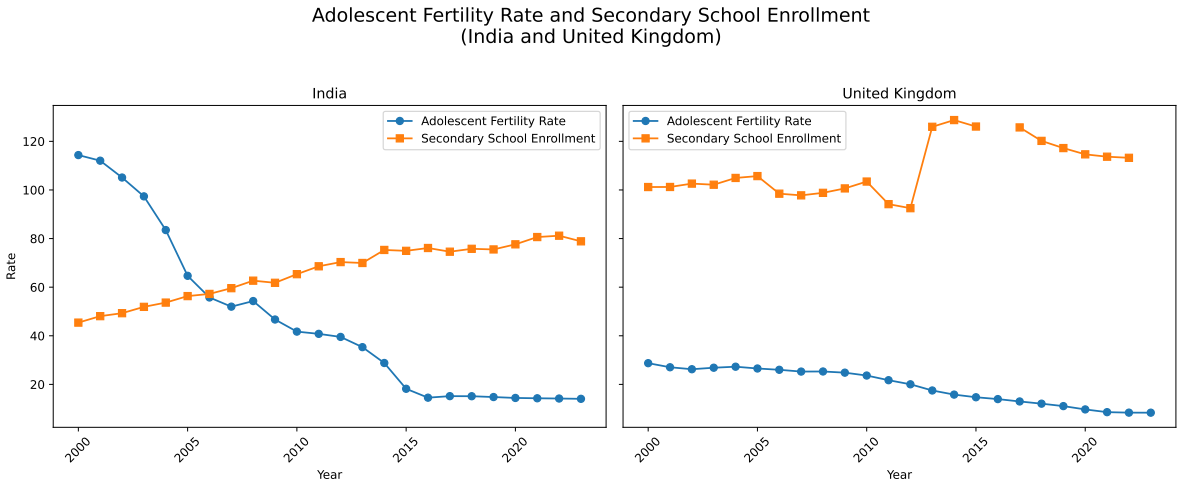

Figure 4: Adolescent Fertility and Secondary School Enrollment Rates: India and United Kingdom

Under-5 Mortality Rate Analysis

Our final indicator is under-5 mortality rate. Figure 5 shows the under-5 mortality rate for our selected countries over time. We see a general decrease in mortality rate for all countries, demonstrating that we're improving this indicator on a global level.

```
#Under 5 Mortality Rate by Country Over Time

# Prepare the data for plotting
pivot_df = df.pivot(index='year', columns='country', values='under5_mortality')

# Plot the line graph
plt.figure(figsize=(12, 8))
pivot_df.plot(ax=plt.gca(), marker='o')
plt.title('Under 5 Mortality Rate by Country Over Time')
plt.xlabel('Year')
plt.ylabel('Under 5 Mortality Rate')
```

```
plt.legend(title='Country', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.xticks(pivot_df.index, rotation=45)
plt.tight_layout()
plt.show()
```

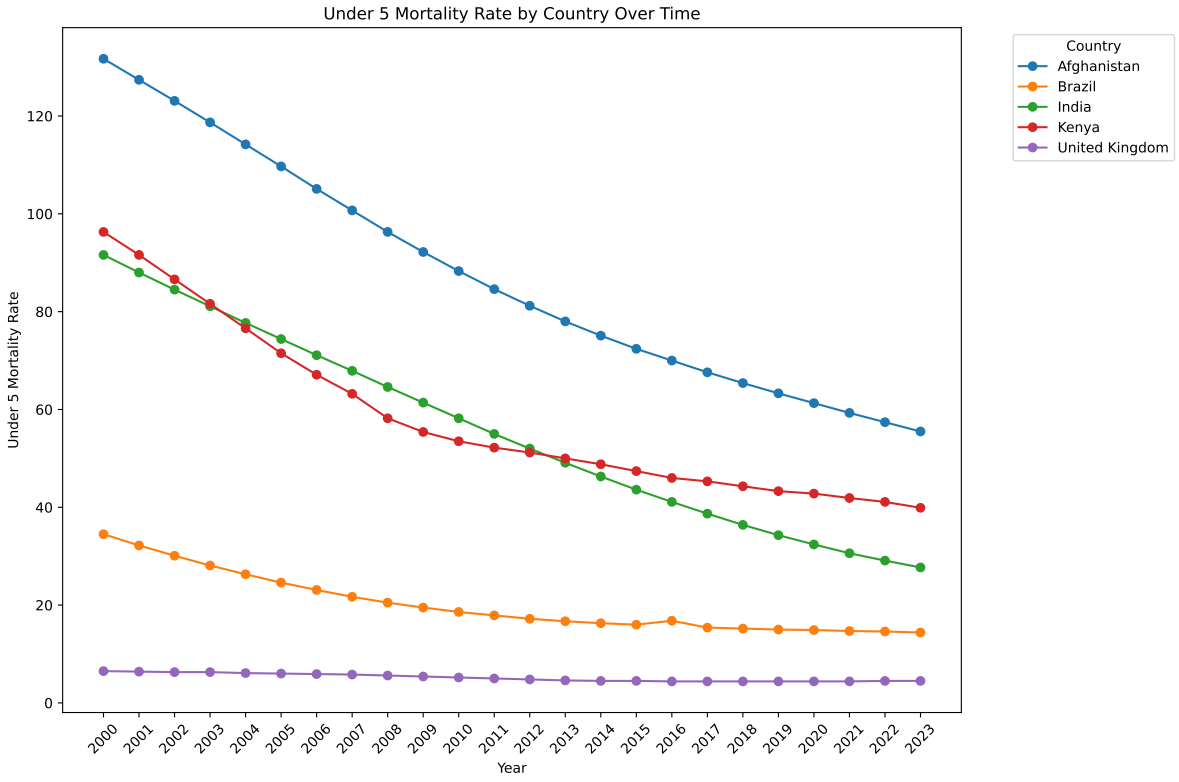

Figure 5: Under-5 Mortality Rates from 2000 - 2023

However, like previous indicators, there appears to be signficant differences in the mortality rate as shown in Table 3. Afghanistan, at the lowest income-level, has a very high under-5 mortality rate of 61.4. Kenya and India, both lower-middle income countries, have mortality rates of 42.7 and 32.7, respectively. We see more improvement as income-level increases. For Brazil, an upper-middle-income country, the under-5 mortality rate is 14.9, and for the United Kingdom, a high-income country, the rate is very low, at 4.4. A country's GNI, then, seems to have a strong bearing on their child mortality rate, which is generally similar to what we previously saw with GNI and adolescent fertility in Table 2.

```
# Average Under 5 Mortality Rate by Country
```

```
table_3 = (
    df[df['year'].between(2017, 2023)]
    .groupby('country', as_index=False)['under5_mortality']
    .mean()
    .round(3)
)
table_3.columns = ['Country', 'Under-5 Mortality (2017-2023)']
table_3
```

Table 3

	Country	Under-5 Mortality (2017–2023)
0	Afghanistan	61.400
1	Brazil	14.886
2	India	32.743
3	Kenya	42.657
4	United Kingdom	4.429

Average Under-5 Mortality Rates from 2017 - 2023

Under-5 Mortality and Adolescent Fertility

In order to better understand the relationship between adolescent fertility and under-5 mortality, we plotted the two variables over time for two countries with very different income levels: Afghanistan and the United Kingdom. This is shown in Figure 6. Based on the parallel downward curves we see in both plots, especially that of Afghanistan, adolescent fertility and under-5 mortality are strongly linked. Additionally, we can clearly see the differences in both of these indicators for low- and high-income countries, as Afghanistan's rates from 2023 are still higher than those of the United Kingdom in 2000.

```
# Under 5 Mortality Rate and Adolescent Fertility Rate

# Prepare the data for plotting
pivot_df_under5_mortality = df.pivot(index='year', columns='country', values='under5_mortalit
pivot_df_adolescent_fertility = df.pivot(index='year', columns='country', values='adolescent_

# Plot the graphs
fig, axes = plt.subplots(1, 2, figsize=(14, 6), sharex=True, sharey=True)

# Afghanistan
```

```
axes[0].plot(
    pivot_df_adolescent_fertility.index,
    pivot_df_adolescent_fertility['Afghanistan'],
    marker='o',
    label='Adolescent Fertility Rate',
    color='tab:blue'
)
axes[0].plot(
    pivot_df_under5_mortality.index,
    pivot_df_under5_mortality['Afghanistan'],
    marker='s',
    label='Under-5 Mortality Rate',
    color='tab:orange'
)
axes[0].set_title('Afghanistan')
axes[0].set_xlabel('Year')
axes[0].set_ylabel('Rate')
axes[0].legend()
axes[0].tick_params(axis='x', rotation=45)

# United Kingdom
axes[1].plot(
    pivot_df_adolescent_fertility.index,
    pivot_df_adolescent_fertility['United Kingdom'],
    marker='o',
    label='Adolescent Fertility Rate',
    color='tab:blue'
)
axes[1].plot(
    pivot_df_under5_mortality.index,
    pivot_df_under5_mortality['United Kingdom'],
    marker='s',
    label='Under-5 Mortality Rate',
    color='tab:orange'
)
axes[1].set_title('United Kingdom')
axes[1].set_xlabel('Year')
axes[1].legend()
axes[1].tick_params(axis='x', rotation=45)

fig.suptitle('Under-5 Mortality and Adolescent Fertility Rates\n(Afghanistan and United Kingd
plt.tight_layout(rect=[0, 0.03, 1, 0.95])
```

```
plt.show()
```

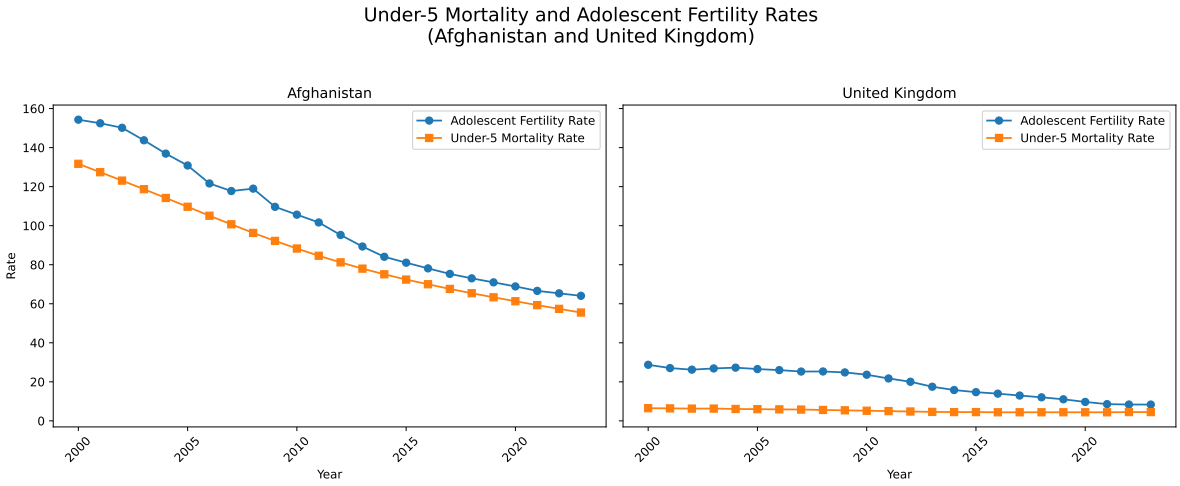

Figure 6: Under-5 Mortality & Adolescent Fertility Rates: Afghanistan and United Kingdom

Now, we'll discuss our results in the next section.