

import packages

Table of contents

| | | |
|----------|--------------------------|----------|
| 0.1 | connect to SQL | 1 |
| 1 | fertility_rate | 2 |
| 1.1 | table | 2 |
| 1.2 | plot | 2 |
| 2 | child_mortality | 3 |
| 2.1 | table | 3 |
| 2.2 | plot | 4 |
| 3 | pop_decline | 5 |
| 3.1 | table | 5 |
| 3.2 | plot | 6 |
| 4 | pop_growth | 7 |
| 4.1 | table | 7 |
| 4.2 | plot | 8 |

```
import pandas as pd
import sqlite3
import matplotlib.pyplot as plt
import seaborn as sns
```

0.1 connect to SQL

```
conn = sqlite3.connect('../data/wdi_data.db')
```

1 fertility_rate

1.1 table

```
# SQL query
query = """
SELECT
    country,
    fertility_rate,
    ROUND(adolescent_fertility_rate/1000, 4) AS ado_fertility_rate
FROM wdi_data
WHERE fertility_rate IS NOT NULL AND adolescent_fertility_rate IS NOT NULL
ORDER BY fertility_rate DESC;
"""
df1 = pd.read_sql_query(query, conn)

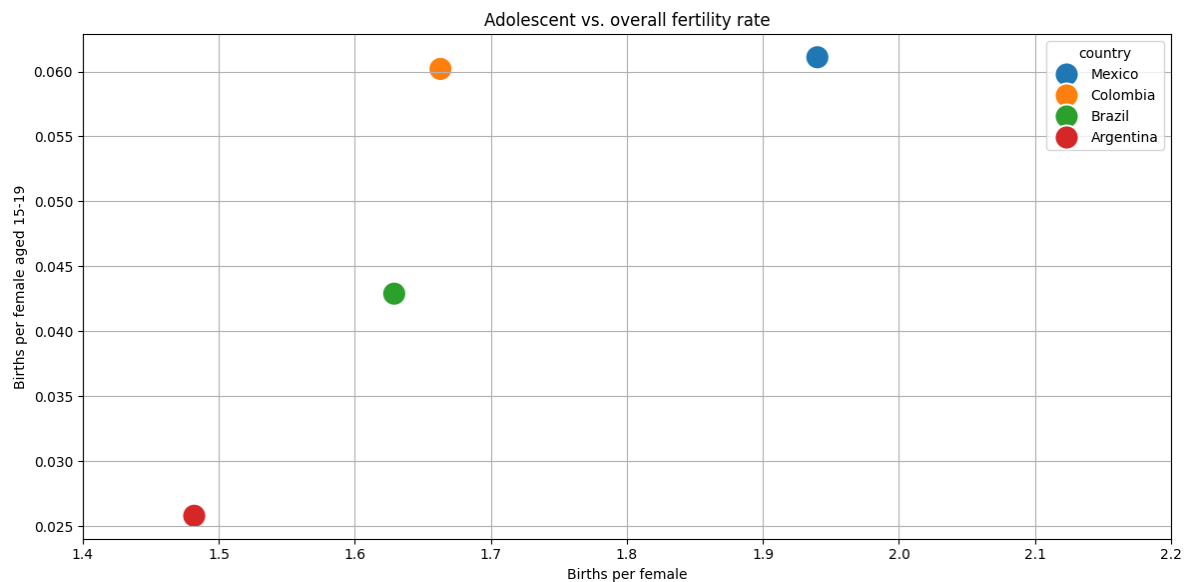
# save an image of the table:
fig, ax = plt.subplots(figsize = (6,2))
ax.axis("off")
table = ax.table(cellText = df1.values,
                 colLabels = ["Country", "Fertility Rate", "Adolescent Fertility Rate"],
                 rowLabels = df1.index,
                 loc = "center",
                 cellLoc = "center")
plt.savefig("../figures/fr_tab.png")
```

| | Country | Fertility Rate | Adolescent Fertility Rate |
|---|-----------|----------------|---------------------------|
| 0 | Mexico | 1.94 | 0.0611 |
| 1 | Colombia | 1.663 | 0.0602 |
| 2 | Brazil | 1.629 | 0.0429 |
| 3 | Argentina | 1.482 | 0.0258 |

1.2 plot

```
plt.figure(figsize=(12, 6))
sns.scatterplot(
    data=df1,
    x='fertility_rate', y='ado_fertility_rate',
    hue='country',
    s = 300
)

plt.grid(True)
plt.xlim(1.4, 2.2)
plt.title('Adolescent vs. overall fertility rate')
plt.ylabel('Births per female aged 15-19')
plt.xlabel('Births per female')
plt.tight_layout()
plt.savefig('../figures/fertility_rate.png')
```



2 child_mortality

2.1 table

```

# SQL query
query = """
SELECT
    country,
    ROUND(population_density, 1) AS pop_density,
    under_5_mortality_rate
FROM wdi_data
WHERE population_density IS NOT NULL AND under_5_mortality_rate IS NOT NULL
ORDER BY under_5_mortality_rate;
"""
df2 = pd.read_sql_query(query, conn)

# save an image of the table:
fig, ax = plt.subplots(figsize = (6,2))
ax.axis("off")
table = ax.table(cellText = df2.values,
                 colLabels = ["Country", "Population Density", "Child Mortality Rate"],
                 rowLabels = df2.index,
                 loc = "center",
                 cellLoc = "center")
plt.savefig("../figures/cm_tab.png")

```

| | Country | Population Density | Child Mortality Rate |
|---|-----------|--------------------|----------------------|
| 0 | Argentina | 16.6 | 9.8 |
| 1 | Colombia | 46.6 | 12.4 |
| 2 | Mexico | 66.2 | 12.9 |
| 3 | Brazil | 25.2 | 14.6 |

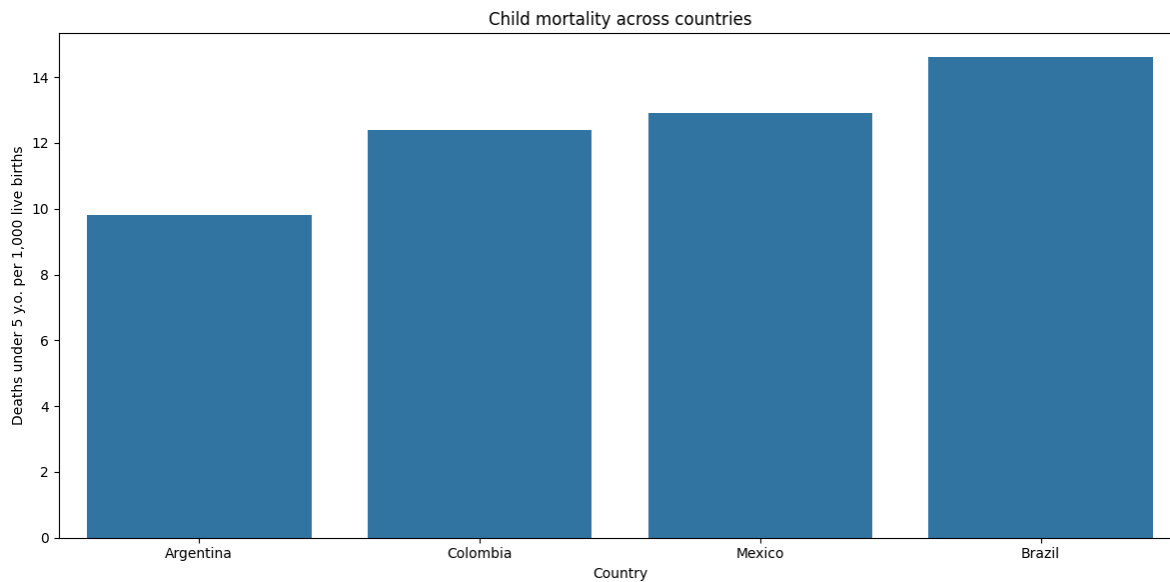
2.2 plot

```

plt.figure(figsize=(12, 6))
sns.barplot(
    data=df2,
    x='country', y='under_5_mortality_rate'
)

```

```
plt.title('Child mortality across countries')
plt.ylabel('Deaths under 5 y.o. per 1,000 live births')
plt.xlabel('Country')
plt.tight_layout()
plt.savefig('../figures/child_mortality.png')
```



3 pop_decline

3.1 table

```
# SQL query
query = """
SELECT
    country,
    birth_rate,
    death_rate,
    ROUND((birth_rate - death_rate), 3) AS circle_of_life
FROM wdi_data
WHERE birth_rate IS NOT NULL AND death_rate IS NOT NULL
ORDER BY circle_of_life ASC;
"""
df3 = pd.read_sql_query(query, conn)
```

```
# save an image of the table:
fig, ax = plt.subplots(figsize = (6,2))
ax.axis("off")
table = ax.table(cellText = df3.values,
                 colLabels = ["Country", "Birth Rate", "Death Rate", "Sum of Rates"],
                 rowLabels = df3.index,
                 loc = "center",
                 cellLoc = "center")
plt.savefig("../figures/pd_tab.png")
```

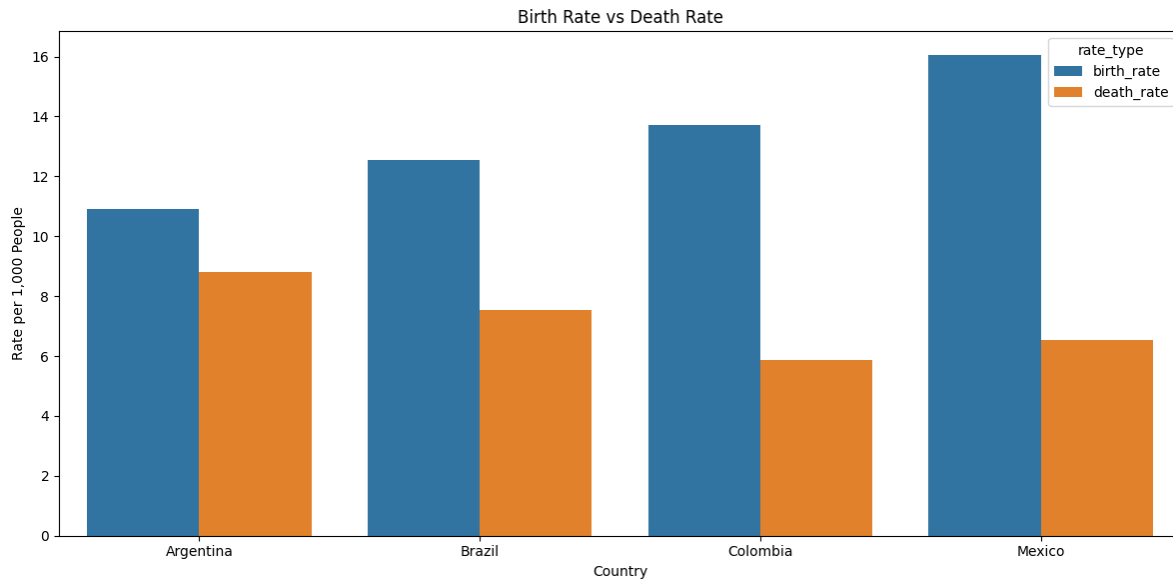
| | Country | Birth Rate | Death Rate | Sum of Rates |
|---|-----------|------------|------------|--------------|
| 0 | Argentina | 10.926 | 8.797 | 2.129 |
| 1 | Brazil | 12.555 | 7.539 | 5.016 |
| 2 | Colombia | 13.731 | 5.881 | 7.85 |
| 3 | Mexico | 16.041 | 6.521 | 9.52 |

3.2 plot

```
# Melt the dataframe to long format
df_melted = pd.melt(
    df3,
    id_vars='country',
    value_vars=['birth_rate', 'death_rate'],
    var_name='rate_type',
    value_name='rate'
)

# Plot
plt.figure(figsize=(12, 6))
sns.barplot(
    data=df_melted,
    x='country', y='rate',
    hue='rate_type'
)
```

```
plt.title('Birth Rate vs Death Rate')
plt.ylabel('Rate per 1,000 People')
plt.xlabel('Country')
plt.tight_layout()
plt.savefig('../figures/pop_decline.png')
```



4 pop_growth

4.1 table

```
# SQL query
query = """
SELECT
    country,
    ROUND(urban_population_percent, 1) AS urban_population_percent,
    ROUND(rural_population_percent, 1) AS rural_population_percent,
    ROUND(population_growth, 3) AS population_growth
FROM wdi_data
WHERE
    urban_population_percent IS NOT NULL
    AND population_growth IS NOT NULL
"""
```

```

df4 = pd.read_sql_query(query, conn)

# save an image of the table:
fig, ax = plt.subplots(figsize = (6,2))
ax.axis("off")
table = ax.table(cellText = df4.values,
                 colLabels = ["Country", "Urban Pop. %", "Rural Pop %", "Pop. Growth"],
                 rowLabels = df4.index,
                 loc = "center",
                 cellLoc = "center")
plt.savefig("../figures/pg_tab.png")

```

| | Country | Urban Pop. % | Rural Pop % | Pop. Growth |
|---|-----------|--------------|-------------|-------------|
| 0 | Colombia | 82.1 | 18.0 | 1.068 |
| 1 | Argentina | 92.3 | 7.7 | 0.211 |
| 2 | Mexico | 81.3 | 18.7 | 0.753 |
| 3 | Brazil | 87.6 | 12.4 | 0.36 |

4.2 plot

```

plt.figure(figsize=(10, 6))
sns.scatterplot(
    data=df4,
    x="urban_population_percent",
    y="population_growth",
    hue="country",
    s = 300
)

plt.xlabel("Urban Population (%)")
plt.ylabel("Population Growth Rate (%)")
plt.title("Urbanization vs Population Growth Rate")
plt.tight_layout()
plt.savefig('../figures/pop_growth.png')

```