Project Goal

This notebook analyzes world population trends across countries and continents from 1970 to 2022 using visualizations and statistical summaries.

Dataset Source

Kaggle(World Population Dataset)

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

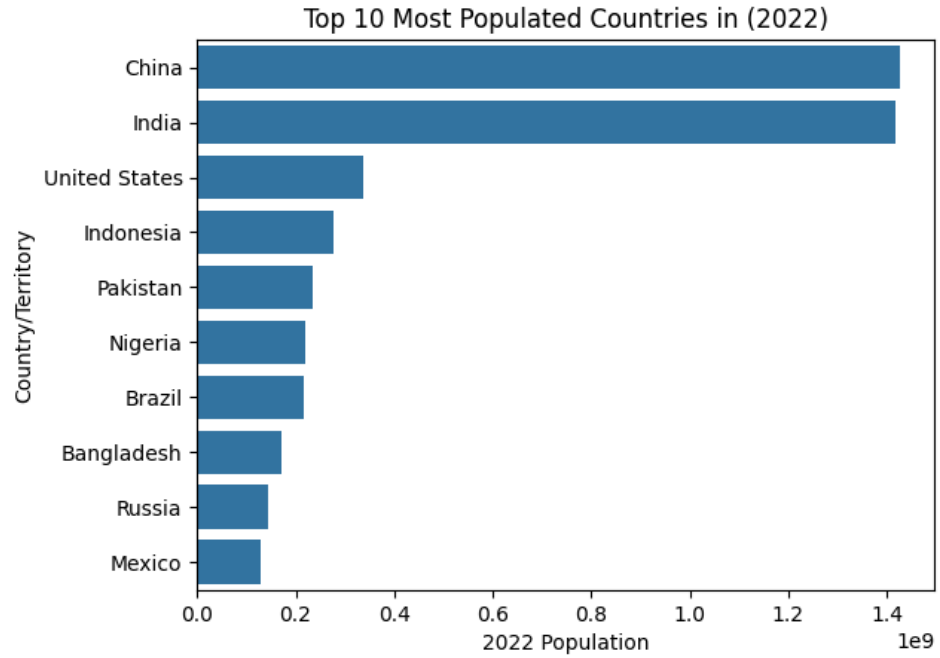
```
import pandas as pd
df = pd.read_csv("/content/drive/MyDrive/world_population.csv")
```

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

df = pd.read_csv("/content/drive/MyDrive/world_population.csv")
```

```
df.info()
df.isnull().sum()
```

Q1. Which countries have the highest population in 2022? Bar Chart

```
top10 = df.sort_values("2022 Population", ascending=False).head(10)

plt.figure()
sns.barplot(x="2022 Population", y="Country/Territory", data=top10)
plt.title("Top 10 Most Populated Countries in (2022)")
plt.show()
```

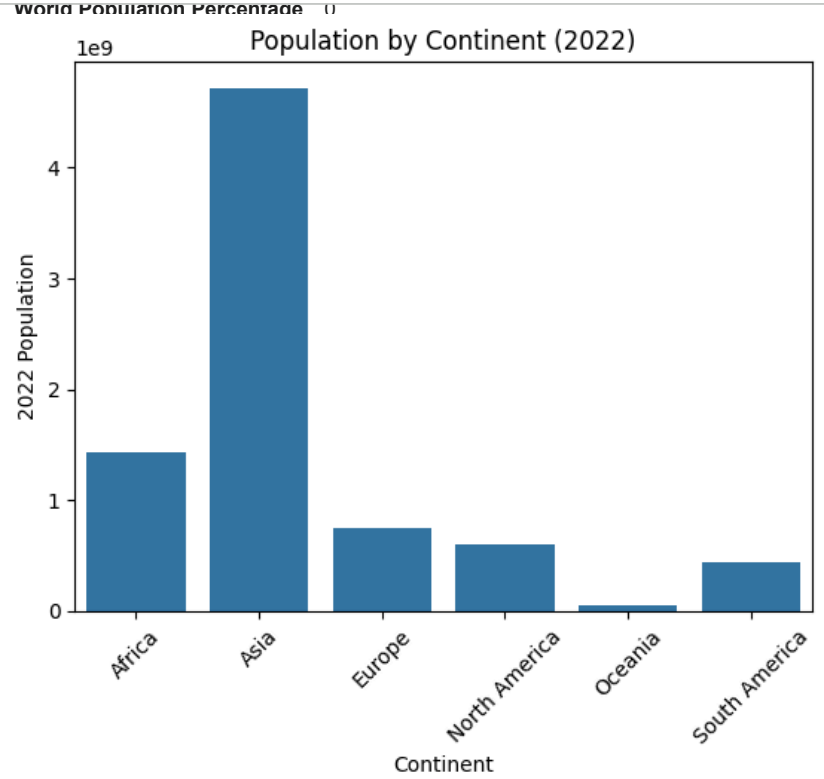


The bar chart shows that the countries like India and China have the highest populations in 2022. These countries dominate global population size compared to others. This indicates that global population is highly concentrated in a small number of countries.

Q2. How is world population distributed across continents? Bar chart

```
continent_pop = df.groupby("Continent")["2022 Population"].sum().reset_index()

plt.figure()
sns.barplot(x="Continent", y="2022 Population", data=continent_pop)
plt.title("Population by Continent (2022)")
plt.xticks(rotation=45)
plt.show()
```

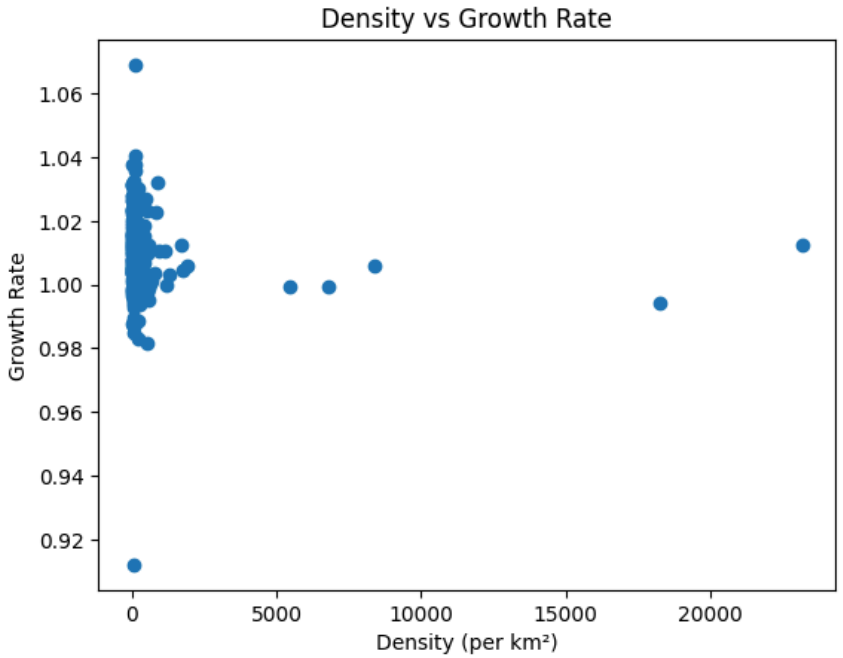
1990 Population	0
1980 Population	0
1970 Population	0
Area (km²)	0
Density (per km²)	0
Growth Rate	0



The continent-wise population comparison shows that Asia has the highest population among all the continents. Africa follows the second most populated continent. This highlights that population is not evenly distributed worldwide, with Asia contributing the largest share.

Q3. Which countries have the highest population density and high growth rate? scatter plot

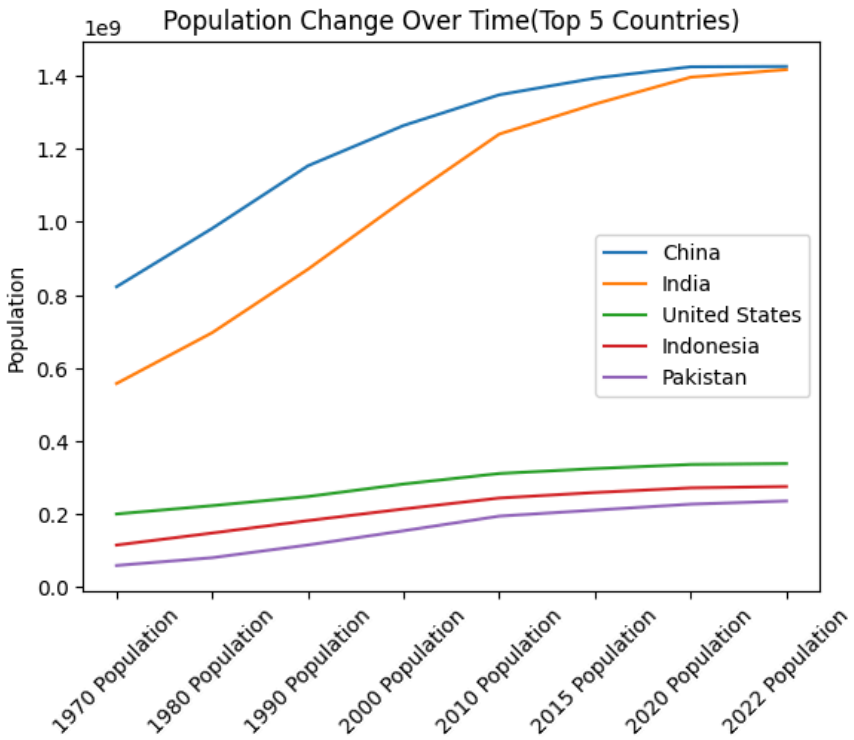
```
plt.figure()
plt.scatter(df["Density (per km²)"], df["Growth Rate"])
plt.title("Density vs Growth Rate")
plt.xlabel("Density (per km²)")
plt.ylabel("Growth Rate")
plt.show()
```



The scatter plot shows the relationship between population density and growth rate across countries. Most countries are clustered around moderate density and moderate growth, while a few outliers have extremely high density but not necessarily high growth. This suggests that high population density does not always lead to higher growth rate, and growth is influenced by factors such as development level, migration, and birth rate.

Q4. How has population changed over time for major countries? Line Chart

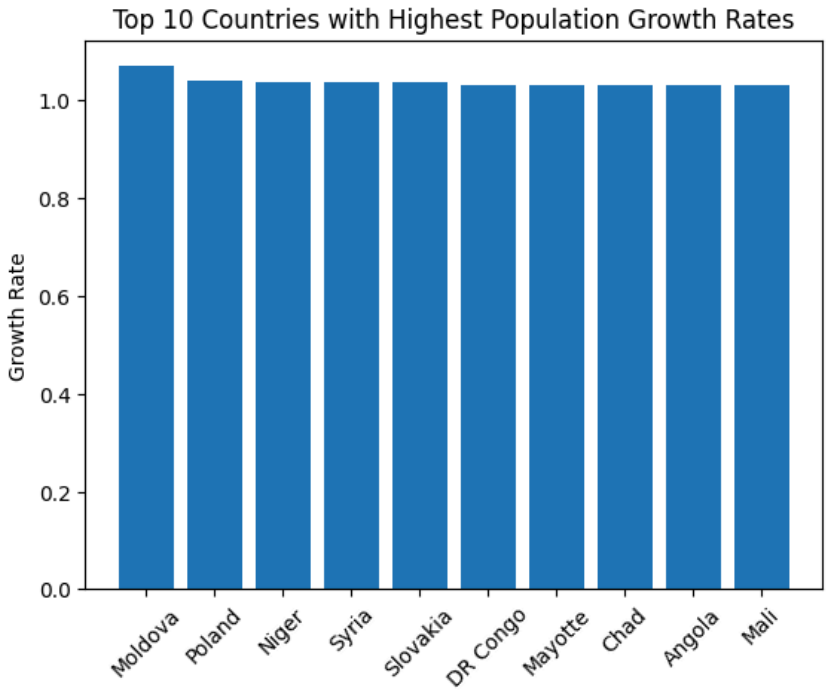
```
major = df.sort_values("2022 Population", ascending=False).head(5)

years = ["1970 Population","1980 Population","1990 Population",
         "2000 Population","2010 Population","2015 Population",
         "2020 Population","2022 Population"]

plt.figure()
for _, row in major.iterrows():
     plt.plot(years, row[years], label=row["Country/Territory"])

plt.xticks(rotation=45)
plt.ylabel("Population")
plt.title("Population Change Over Time(Top 5 Countries)")
plt.legend()
plt.show()
```



The line chart indicates that population has increased steadily over the years for the top 5 countries. Some countries show faster growth trends than others, especially after 2000. This suggests that population growth is uneven and depends on regional factors like fertility rate and development level.

Q5. Which countries have the highest population growth rates? Bar chart

```
top_growth = df.sort_values("Growth Rate", ascending=False).head(10)

plt.figure()
plt.bar(top_growth["Country/Territory"], top_growth["Growth Rate"])
plt.xticks(rotation=45)
plt.title("Top 10 Countries with Highest Population Growth Rates")
plt.ylabel("Growth Rate")
plt.show()
```

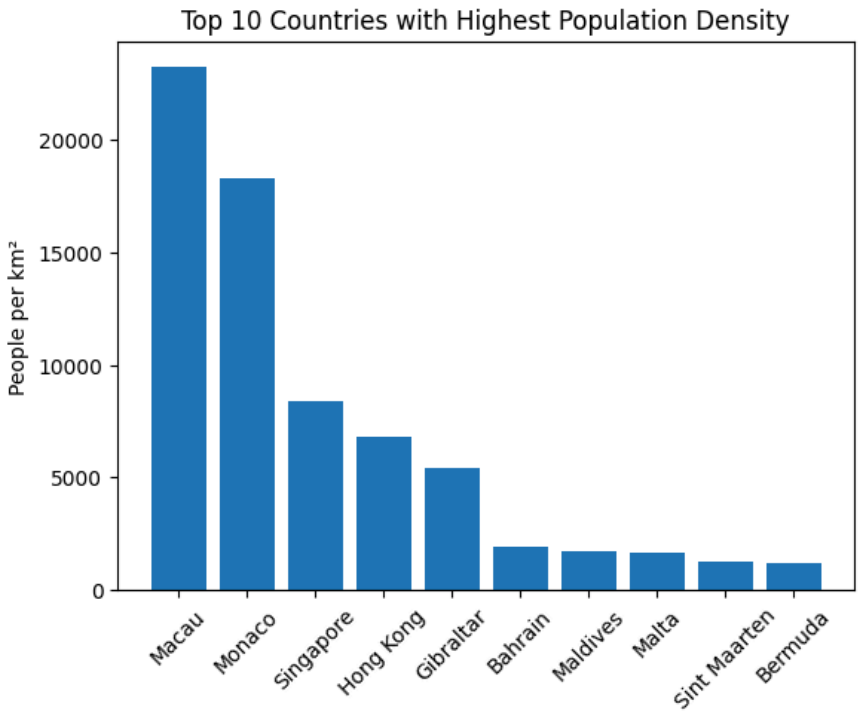


The chart highlights that countries with the highest population growth rates. Many of these countries belong to developing regions where birth rates remain high. This suggests that future population increase will be more significant in these countries compared to developed nations.

Q6. Which countries are the most densely populated? Bar chart

```
top_density = df.sort_values("Density (per km²)", ascending=False).head(10)

plt.figure()
```

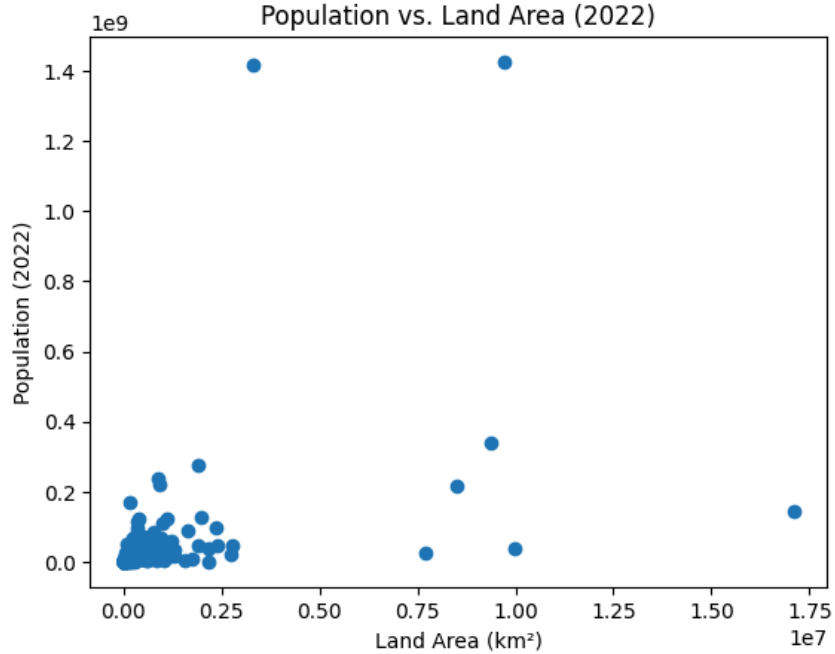
```
plt.bar(top_density["Country/Territory"], top_density["Density (per km²)"])
plt.xticks(rotation=45)
plt.title("Top 10 Countries with Highest Population Density")
plt.ylabel("People per km²")
plt.show()
```



The density chart shows that some countries have extremely high population density compared to others. These are generally smaller countries with high population concentration. This indicates that high density is not necessarily linked to total population but depends on land area and unbanization.

Q7. What is the relationship between land area and population? Scatter graph

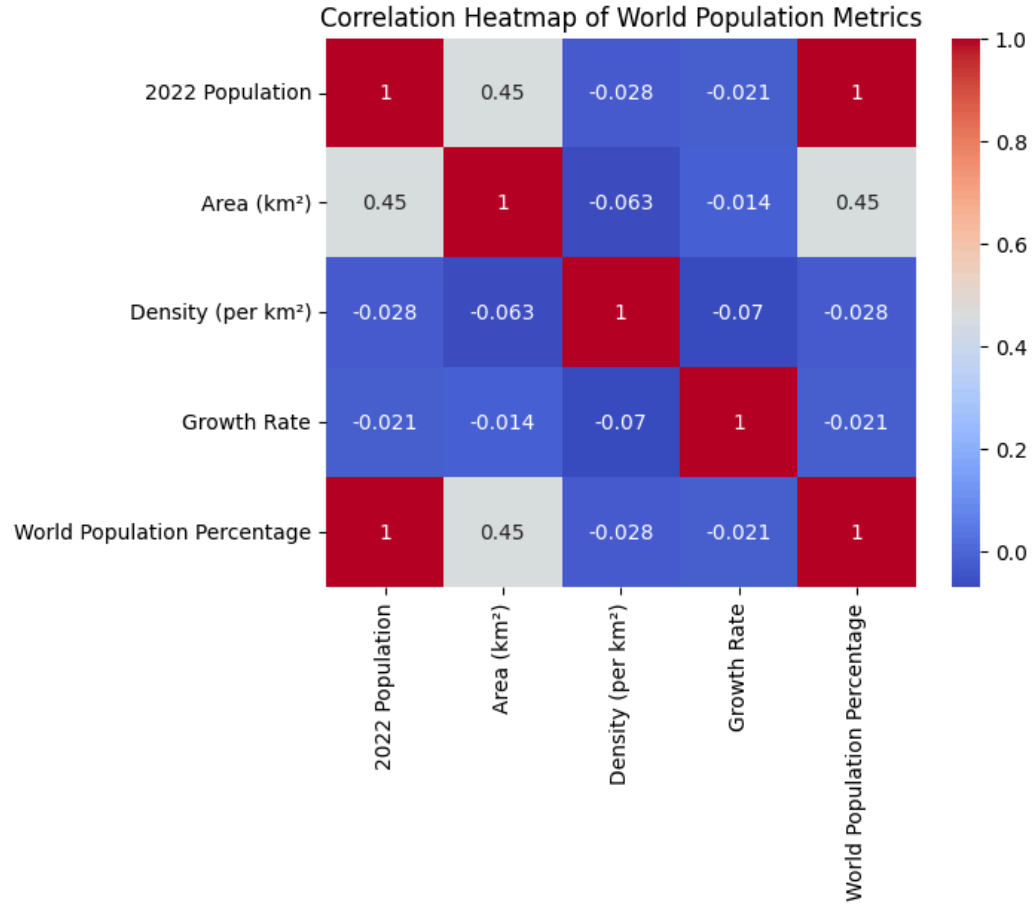
```
plt.figure()
plt.scatter(df["Area (km²)"], df["2022 Population"])
plt.xlabel("Land Area (km²)")
plt.ylabel("Population (2022)")
plt.title("Population vs. Land Area (2022)")
plt.show()
```



This scatter plot shows that there is no strong direct relationship between land area and population size. Some large countries have moderate population, while some smaller countries have very high population. This suggests that factors like economy, fertility, and migration influence population more than land size.

Q8. How are population, area, density, and growth rate correlated? Heatmap

```
num_cols = [
    "2022 Population",
    "Area (km²)",
    "Density (per km²)",
    "Growth Rate",
    "World Population Percentage"
]

corr = df[num_cols].corr()
plt.figure()
sns.heatmap(corr, annot=True, cmap="coolwarm")
plt.title("Correlation Heatmap of World Population Metrics")
plt.show()
```

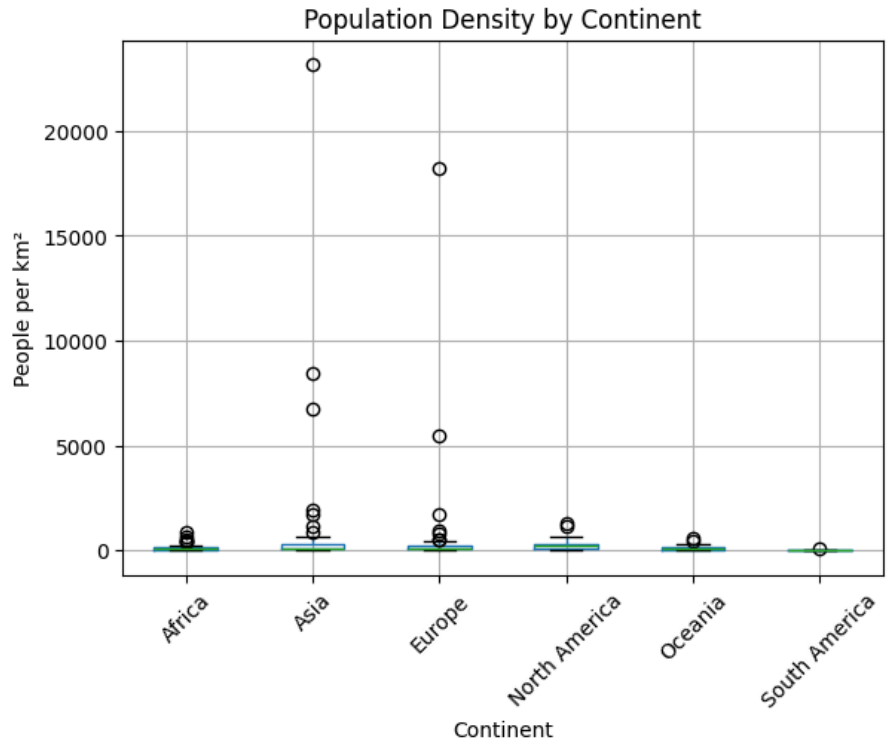


This heatmap shows strong correaltion between 2022 population and world population percentage, meaning countries with high population contribute more to global share. Density and land area tend to have weak or negative correlation, showing that large countries are not always densely populated. This analysis helps indentify how population-related variables interact with each other.

Q9. How does population density vary across continents? Box plot

```
plt.figure()
df.boxplot(column="Density (per km²)", by="Continent")
plt.title("Population Density by Continent")
plt.suptitle("")
plt.ylabel("People per km²")
plt.xticks(rotation=45)
plt.show()
```

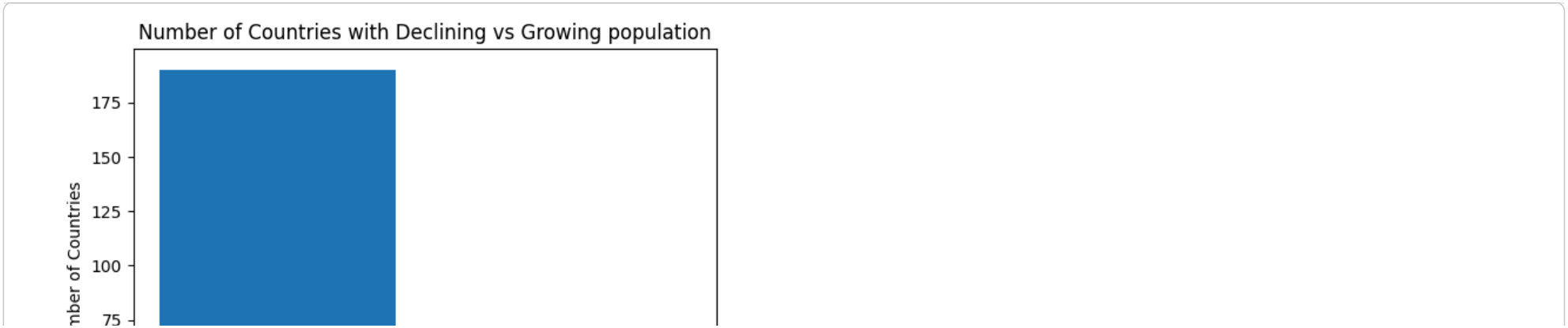
```
<Figure size 640x480 with 0 Axes>
```



The box plot shows that continents differ greatly in population density distribution. Some continents have a wider spread, meaning density varies strongly between countries. The presence of ouliers indicates that certain countries have extremely high density compared to the rest of the continent.

Q10. How many countries have declining vs growling population? Bar chart

```
df["Growth Category"] = df["Growth Rate"].apply(lambda x: "Declining" if x < 1 else "Growing")

growth_counts = df["Growth Category"].value_counts()

plt.figure()
plt.bar(growth_counts.index, growth_counts.values)
plt.xlabel("Population Trends")
plt.ylabel("Number of Countries")
plt.title("Number of Countries with Declining vs Growing population")
plt.show()
```


Number of Countries with Declining vs Growing population
175
150
125
100
75