Step 4 Report

Group: Cockatiel

Group Members:
Tolga Başak 29074
Şimal Yücel 29420
Ulaş Meriç 29244
Arif Kemal Sarı 28999
Tolga Başak 29074
Emre Bülbül 28882

GITHUB LINK:
https://github.com/tolgabasak/cs306/tree/main/Step4

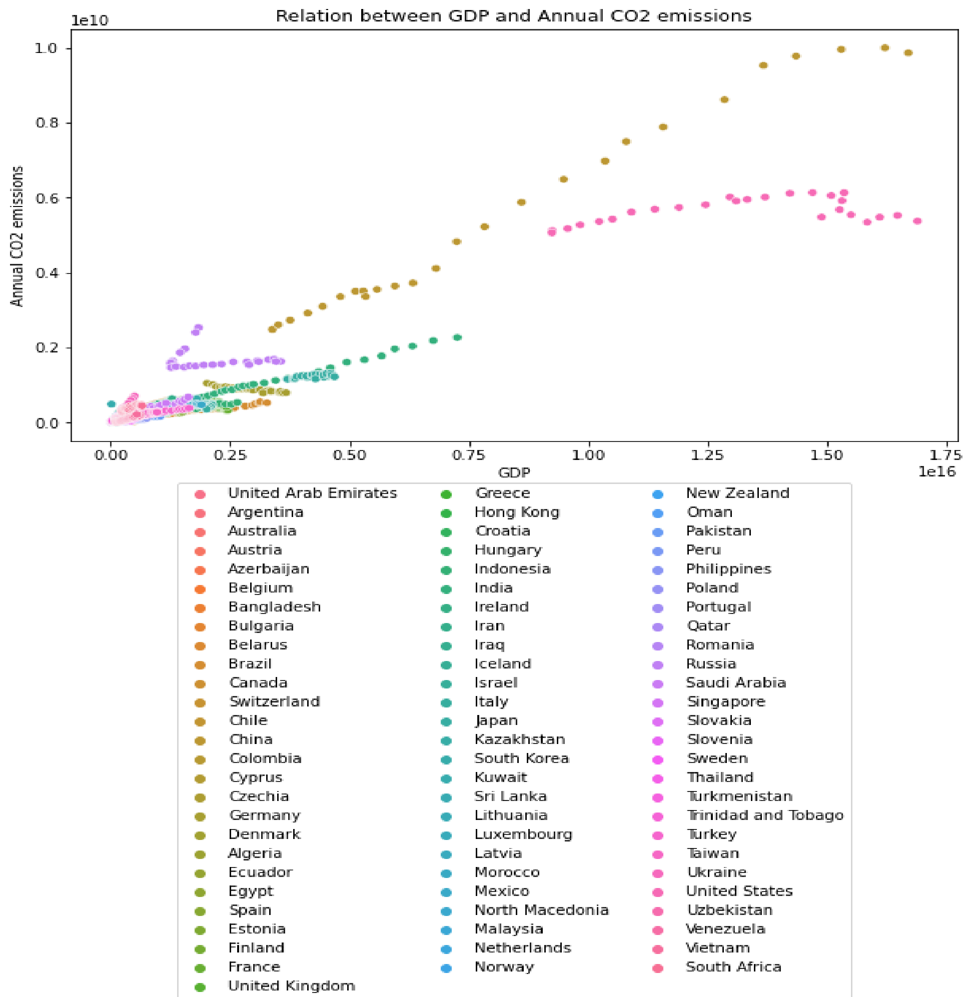
Python Code:
```
# -*- coding: utf-8 -*-
"""
Spyder Editor
"""
# Import libs
import mysql.connector
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Connect to the database
cnx = mysql.connector.connect(user='root', password='Tb843749',
                              host='127.0.0.1',
                              database='cs306')

# Define a SQL queries to fetch the data
query1 = "SELECT * FROM gdp_and_emissions"
query2 = "SELECT * FROM emissions_and_ozone"
query3 = "SELECT * FROM gdp_and_pollution"
query4 = "SELECT * FROM population_and_emissions"
query5 = "SELECT * FROM wind_and_airpollution"
query6 = "SELECT * FROM gdp_population_wind_energy"

# Queries for each view for Turkey
query7 = "SELECT * FROM emissions_and_ozone WHERE Code = 'TUR'"
query8 = "SELECT * FROM gdp_and_emissions WHERE Code = 'TUR'"
query9 = "SELECT * FROM gdp_and_pollution WHERE Code = 'TUR'"
```

```
query10 = "SELECT * FROM population_and_emissions WHERE Code = 'TUR'"
query12 = "SELECT * FROM gdp_population_wind_energy WHERE Code = 'TUR'"

# Use pandas to execute the query and store the result in a DataFrame
df1 = pd.read_sql_query(query1, cnx)
df2 = pd.read_sql_query(query2, cnx)
df3 = pd.read_sql_query(query3, cnx)
df4 = pd.read_sql_query(query4, cnx)
df5 = pd.read_sql_query(query5, cnx)
df6 = pd.read_sql_query(query6, cnx)

# dataframes for Turkey
df7 = pd.read_sql_query(query7, cnx)
df8 = pd.read_sql_query(query8, cnx)
df9 = pd.read_sql_query(query9, cnx)
df10 = pd.read_sql_query(query10, cnx)
df12 = pd.read_sql_query(query12, cnx)

# Close the connection
cnx.close()

# Scatter plot for GDP and emissions
plt.figure(figsize=(10,6))
sns.scatterplot(data=df1, x='GDP', y='Annual_CO2_emissions', hue='Country')
plt.title('Relation between GDP and Annual CO2 emissions')
plt.xlabel('GDP')
plt.ylabel('Annual CO2 emissions')

# Move the legend to below the plot
plt.legend(bbox_to_anchor=(0.5, -0.1), loc='upper center', borderaxespad=0., ncol=3)

plt.show()


# Line Chart for emissions and ozone
df2_filtered = df2[df2['Year'] == 2015] #adjust year
plt.figure(figsize=(10,6))
for country in df2_filtered['Country'].unique():
    country_data = df2_filtered[df2_filtered['Country'] == country]
    plt.plot(country_data['Annual_CO2_emissions'],
country_data['Ozone_concentration_StateofGlobalAir'], label=country)
plt.title('Relation between Annual CO2 emissions and Ozone concentration')
plt.xlabel('Annual CO2 emissions')
plt.ylabel('Ozone concentration')
plt.legend(bbox_to_anchor=(0.5, -0.1), loc='upper center', borderaxespad=0., ncol=3)
plt.show()
```

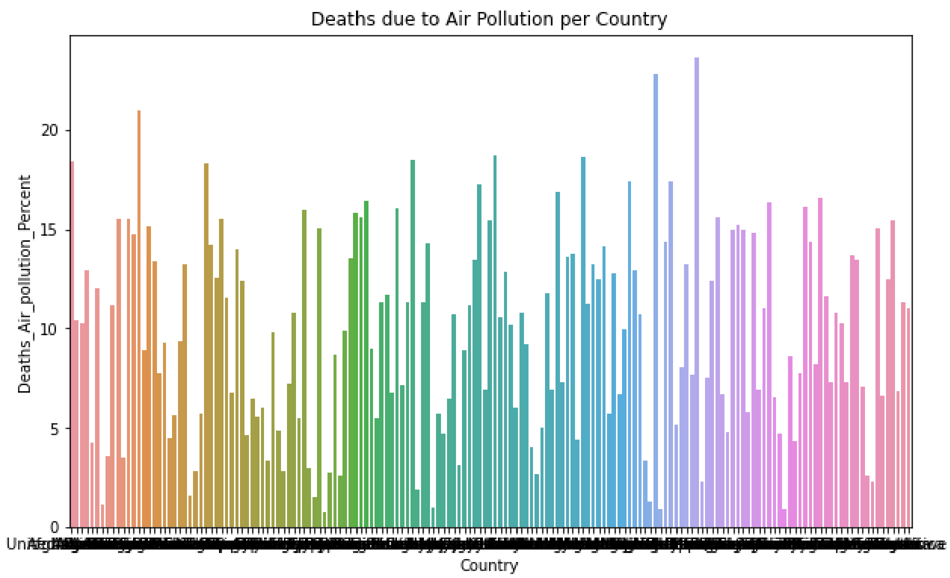
```
# Bar plot for gdp and pollution
df3_filtered = df3[df3['Year'] == 2015] #adjust year
plt.figure(figsize=(10,6))
sns.barplot(data=df3_filtered, x='Country', y='Deaths_Air_pollution_Percent')
plt.title('Deaths due to Air Pollution per Country')
plt.xlabel('Country')
plt.ylabel('Deaths_Air_pollution_Percent')
plt.show()

# Column Chart (using barplot for vertical bars) for population and emissions
df4_filtered = df4[df4['Year'] == 2015] #adjust year
plt.figure(figsize=(10,6))
sns.barplot(data=df4_filtered, x='Country', y='Population')
plt.title('Population per Country')
plt.xlabel('Country')
plt.ylabel('Population')
plt.show()

# Pie chart for wind and air pollution (showing proportion of wind energy per capita)
df5_filtered = df5[df5['Year'] == 2015] #adjust year
wind_sum = df5_filtered['Wind_per_capita_kWh_equivalent'].sum()
df5_filtered['wind_proportion'] = df5_filtered['Wind_per_capita_kWh_equivalent'] /
wind_sum
plt.figure(figsize=(12, 8))  # Increase the figure size to 12x8 inches
plt.pie(df5_filtered['wind_proportion'], labels = df5_filtered['Country'],
autopct='%1.1f%%')
plt.title('Proportion of Wind Energy per capita by Country')
plt.show()

# Area Chart for gdp, population and wind energy
df6_filtered = df6[df6['Year'] == 2015] #adjust year
plt.figure(figsize=(10,6))
plt.fill_between(df6_filtered['GDP'], df6_filtered['Population'], color="skyblue", alpha=0.4)
plt.plot(df6_filtered['GDP'], df6_filtered['Population'], color="Slateblue", alpha=0.6)
plt.title('Relation between GDP and Population')
plt.xlabel('GDP')
plt.ylabel('Population')
plt.show()



# Create visualizations for Turkey
# 1. Line chart for annual CO2 emissions over the years
plt.figure(figsize=(12,8))
sns.lineplot(data=df7, x='Year', y='Annual_CO2_emissions')
```

```
plt.title('Annual CO2 emissions Over Years in Turkey')
plt.show()

# 2. Bar chart for fossil fuel consumption over the years
plt.figure(figsize=(12,8))
sns.barplot(data=df8, x='Year', y='Fossil_fuels_TWh')
plt.title('Fossil Fuels Consumption Over Years in Turkey')
plt.show()

# 3. Scatter plot for GDP vs Ozone concentration
plt.figure(figsize=(10,6))
sns.scatterplot(data=df9, x='GDP', y='Ozone_concentration_StateofGlobalAir')
plt.title('GDP vs Ozone Concentration in Turkey')
plt.show()

# 4. Area chart for Population and Annual CO2 emissions
df10.set_index('Year')[['Population', 'Annual_CO2_emissions']].plot(kind='area',
alpha=0.4, stacked=False)
plt.title('Population and Annual CO2 emissions Over Years in Turkey')
plt.show()


# 6. Column chart for Wind per capita kWh equivalent over the years
plt.figure(figsize=(12,8))
df12.set_index('Year')['Wind_per_capita_kWh_equivalent'].plot(kind='bar')
plt.title('Wind per capita kWh Equivalent Over Years in Turkey')
plt.show()
```

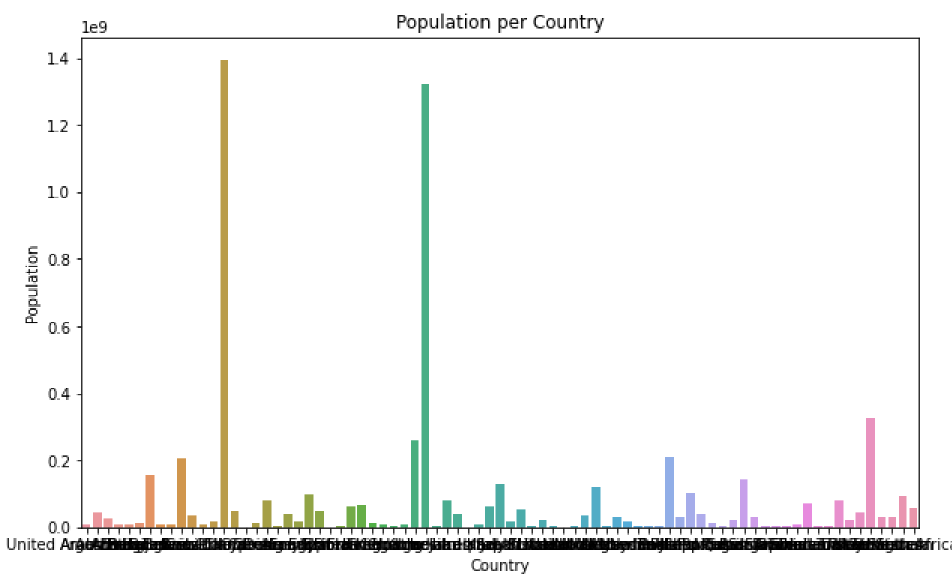
Reports for Visuals:



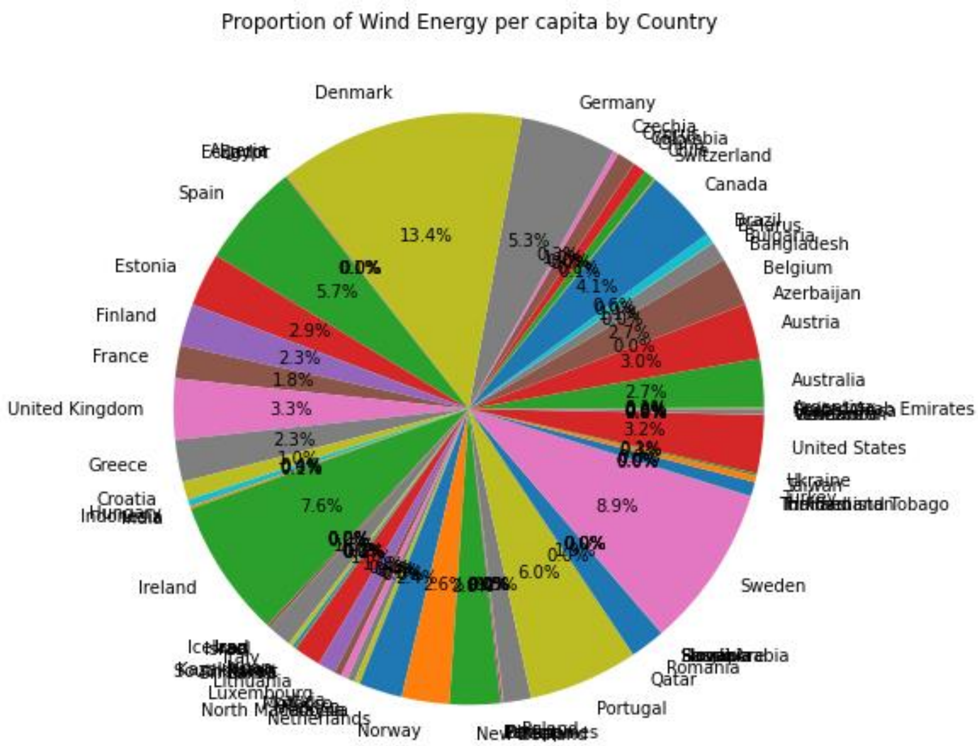
In the first scatter plot, we compared the GDP and Annual CO2 emissions for various countries. The visualization shows the relationship between economic prosperity (GDP) and carbon emissions. We aggregated the data for the year 2015 and observed that countries with higher GDP tend to have higher CO2 emissions.



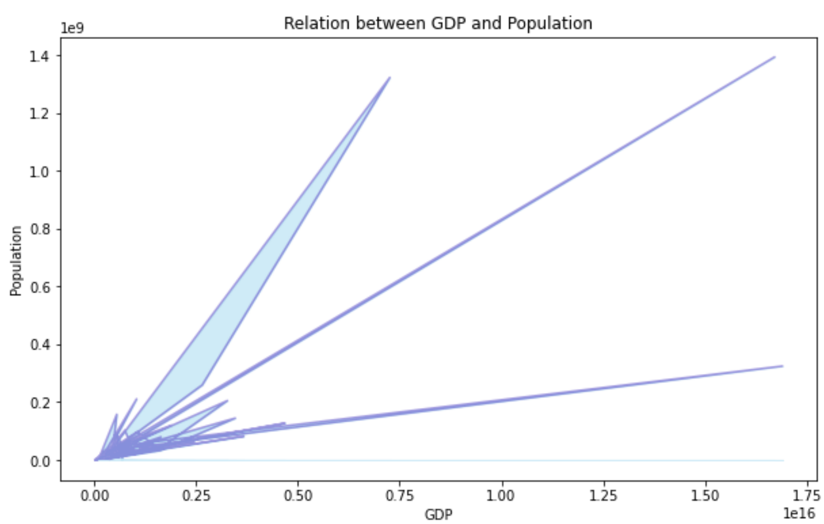
The bar plot displays the percentage of deaths due to air pollution in different countries. We aggregated the data for the year 2015, providing insights into the impact of air pollution on public health across nations.



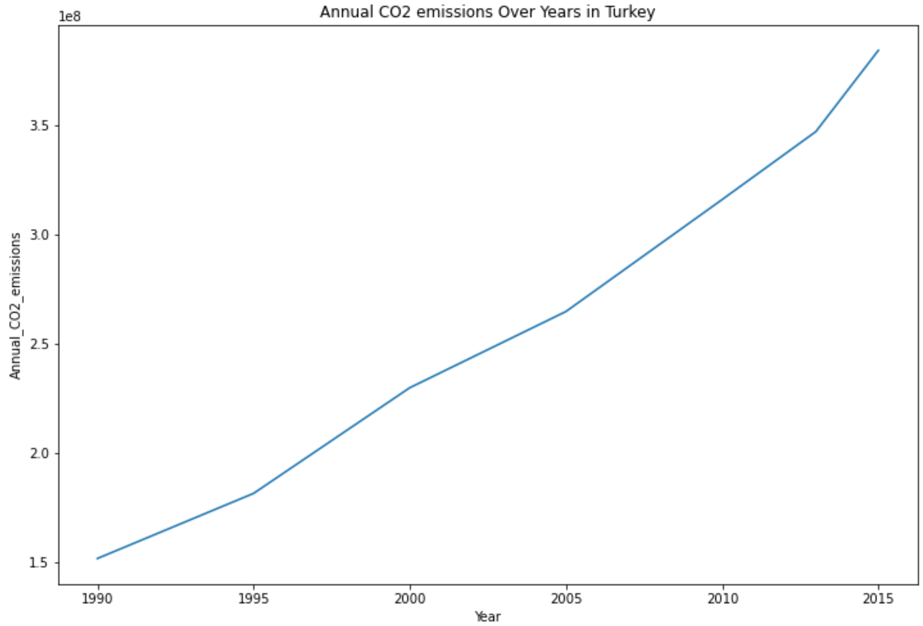
The column chart showcases the population of each country in 2015. By comparing the heights of the bars, we can observe the population differences among countries.



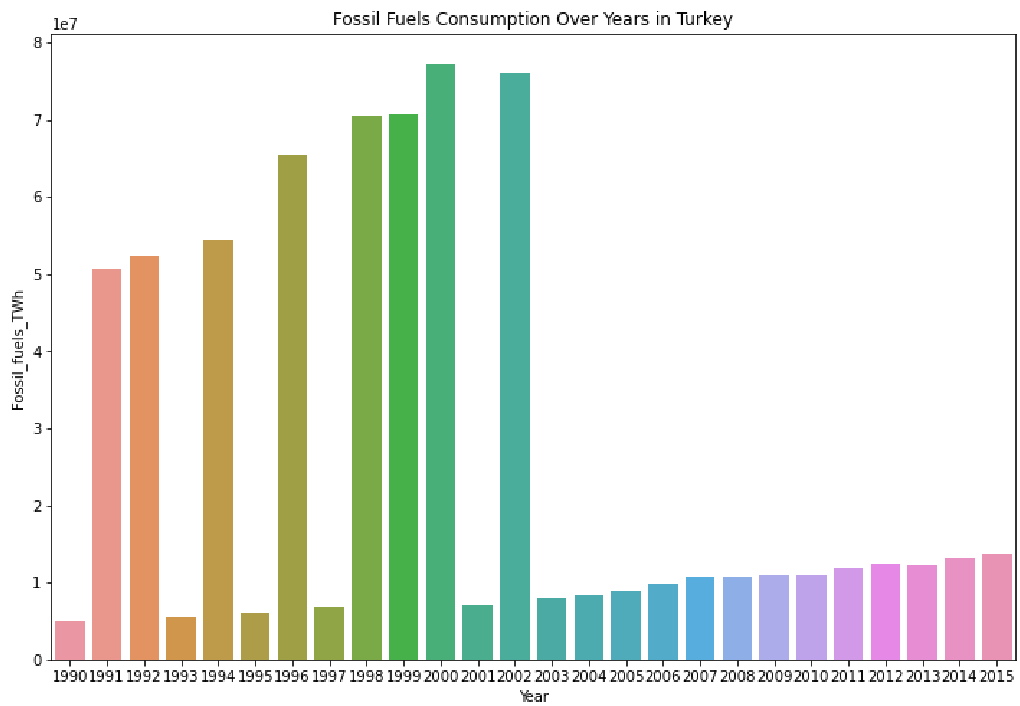
The pie chart represents the proportion of wind energy per capita by country in 2015. Each slice represents a country, and the size of the slice indicates the proportion of wind energy consumption relative to the total. The visualization offers an overview of the distribution of wind energy usage across countries.



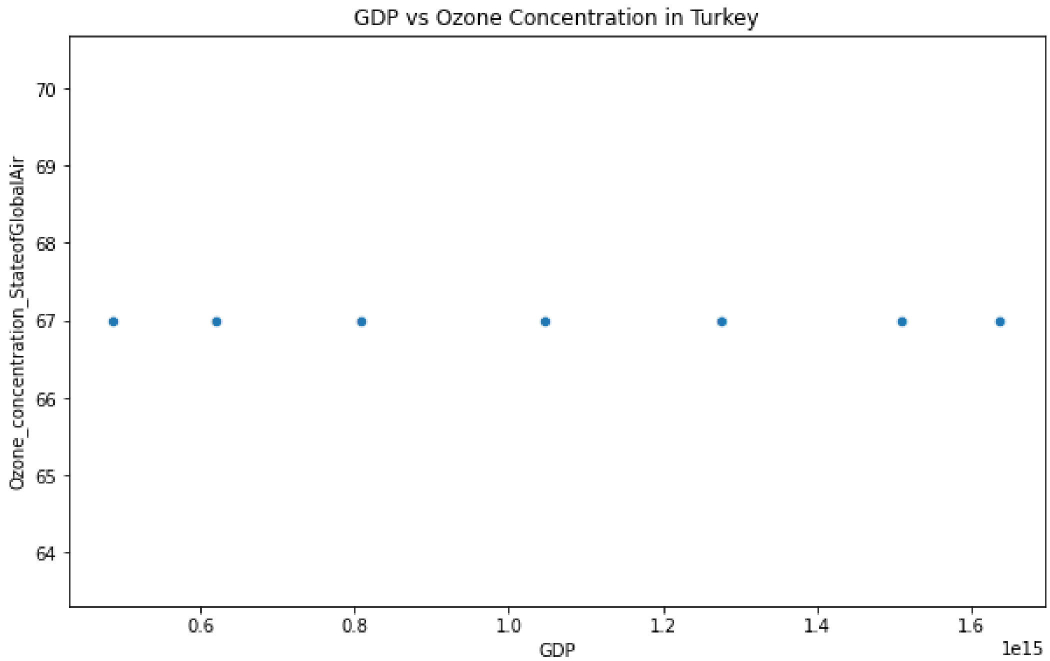
The area chart showcases the relationship between GDP and population in 2015. The filled area represents the population, while the line represents the GDP. By comparing the two, we can analyze the correlation between economic development and population size.



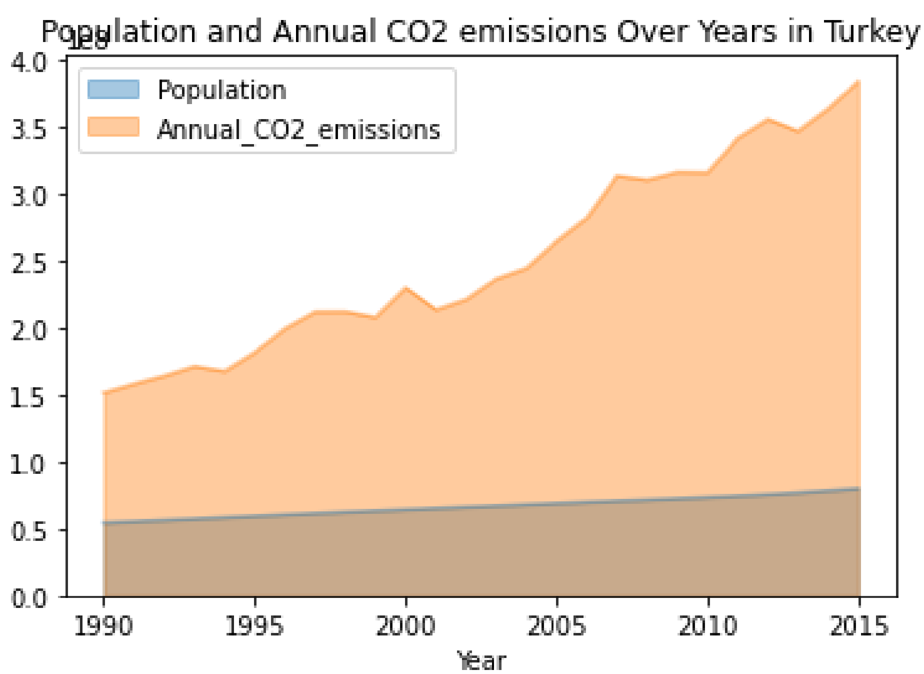
For the visualizations specific to Turkey, we explored various aspects. The line chart illustrates the annual CO2 emissions over the years in Turkey, indicating the trends and changes in emissions levels.



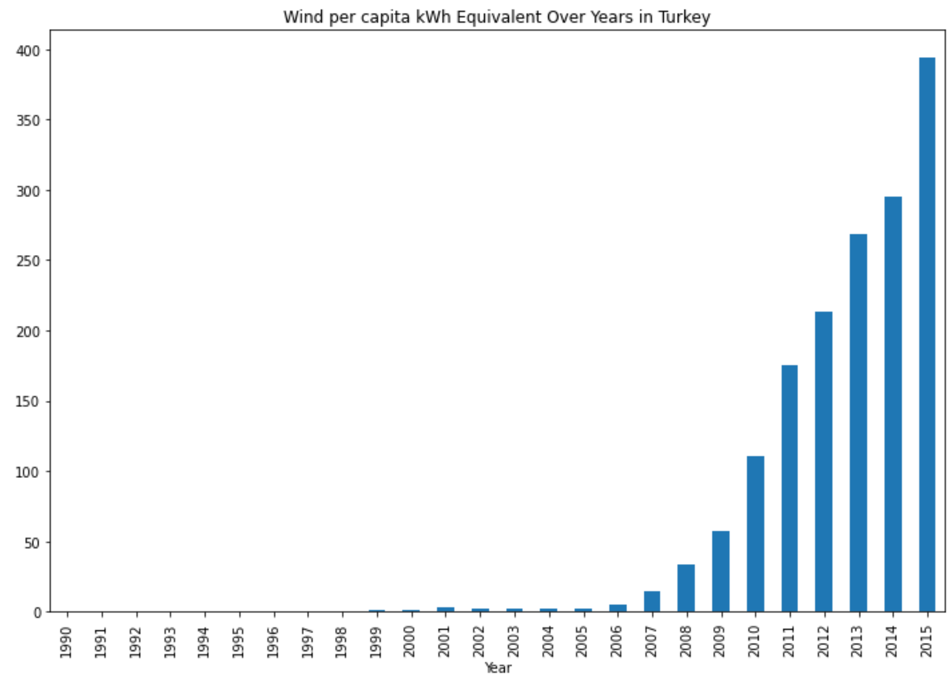
The bar chart represents the consumption of fossil fuels over the years in Turkey, providing insights into the energy consumption patterns and trends in the country.



The scatter plot shows the relationship between GDP and ozone concentration in Turkey. By analyzing the data points, we can understand the association between economic growth and environmental factors.



The area chart showcases the population and annual CO2 emissions in Turkey over the years. By comparing the two areas, we can identify population growth trends and its impact on CO2 emissions.



The column chart presents the wind per capita kWh equivalent over the years in Turkey, providing insights into the adoption and utilization of wind energy in the country.