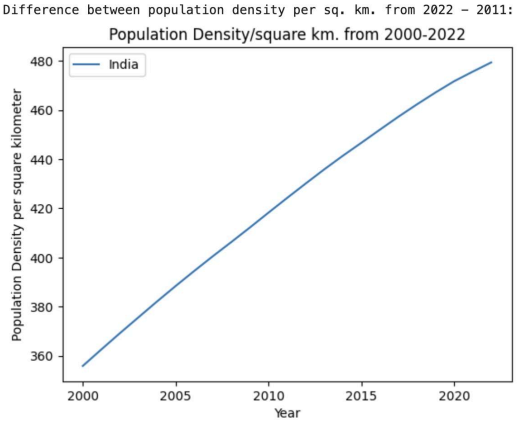
What is the Correlation Between Urbanization and Climate Change?

- **Background**: from sustainability presentation
    - 

    - We prioritize our economy (bottom line and cash flow) and tend to leave the environment for the backburner (conservation of resources and ecosystems)
    - There are already national and global frameworks/policies in place to regulate climate change (Paris Climate Accords) by targeting things like temperature, emissions and carbon pricing
    - Yet, there must still be a discrepancy or at least other factors playing a bigger role in climate change.
- **Our hypotheses:**
    - Urbanization and climate change are positively correlated
        - The more urbanized a country is, the greater effect is has on climate change
    - We also want to explore certain aspects of urbanization
        - Does the urban population of a country play a role?
        - Does access to and use of technology affect climate change?
- **Two datasets**
    - **Global Urbanization**: https://www.kaggle.com/datasets/bushraqurban/global-urbanization-and-climate-metrics
    - **Global Warming Trends**: https://www.kaggle.com/datasets/jawadawan/global-warming-trends-1961-2022/data
- **Notes on Data Cleaning**
    - ***Can't visualize all 182 countries***: overwhelming to analyze in given time
    - ***Instead***: one country from each continent (except Antarctica because it has no countries or measurable urbanization)

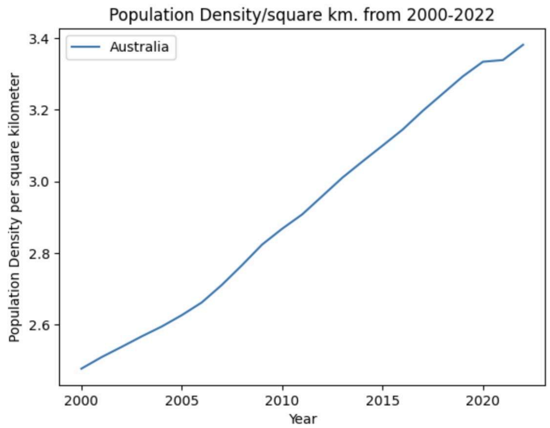
        - *North America*: United States of America
        - *South America*: Brazil
        - *Europe*: Germany
        - *Asia*: India
        - *Africa*: Nigeria
        - *Australia*: counting it as a country-continent
    - ***Filtering both datasets by year***
        - Lots of missing data before 2000s
        - ***Hence***: we will focus on 2000 - 2022, including 2011 as a midpoint
            - Full data
            - Relevancy
    - ***Temperature Units***: degrees Celsius
- **Global Urbanization Dataset**
    - **<u>First</u>**, we wanted to plot the growth in population density for each country over 2000 to 2022 since it is the best indicator urbanization
        - We also plotted the percent difference in urban population for each country from 2000 to 2022

```
Difference between population density per sq. km. from 2022 – 2011: 123.60477433329203
```



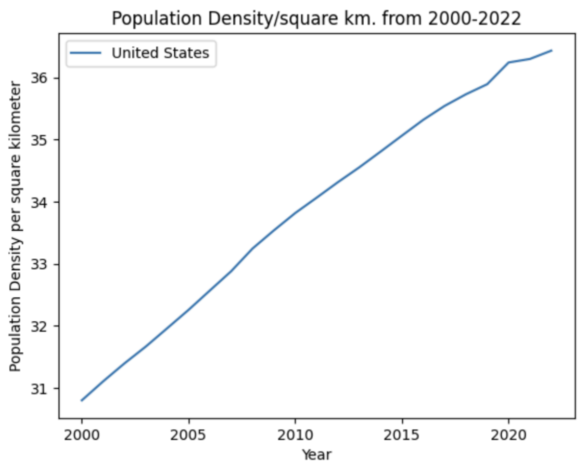
    -

```
Difference between population density per sq. km. from 2022 - 2011: 0.90503156293981
```

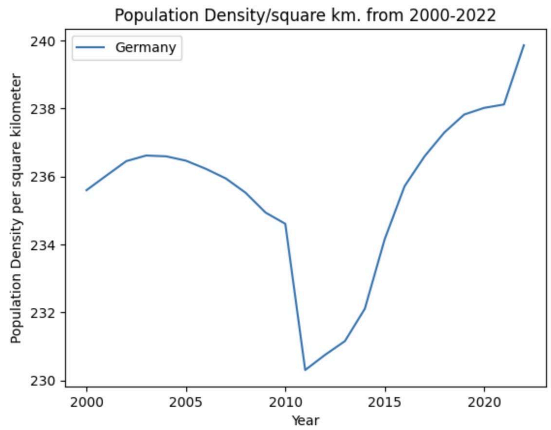


-

```
Difference between population density per sq. km. from 2022 - 2011: 5.6360767155419005
```

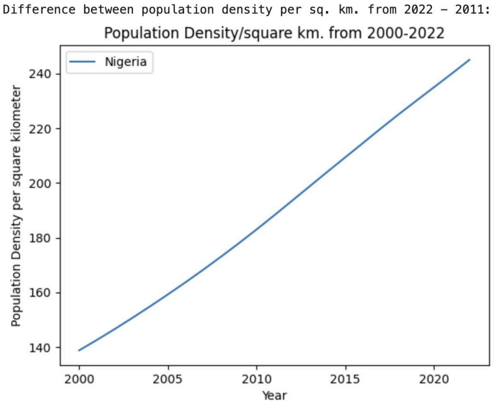


Difference between population density per sq. km. from 2022 - 2011: 4.264604724490994

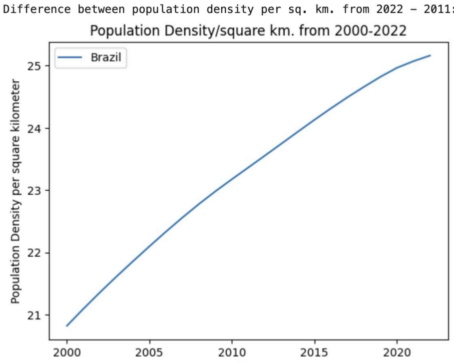


-

Difference between population density per sq. km. from 2022 - 2011: 106.249000296453



Difference between population density per sq. km. from 2022 - 2011: 4.341651731126802



-

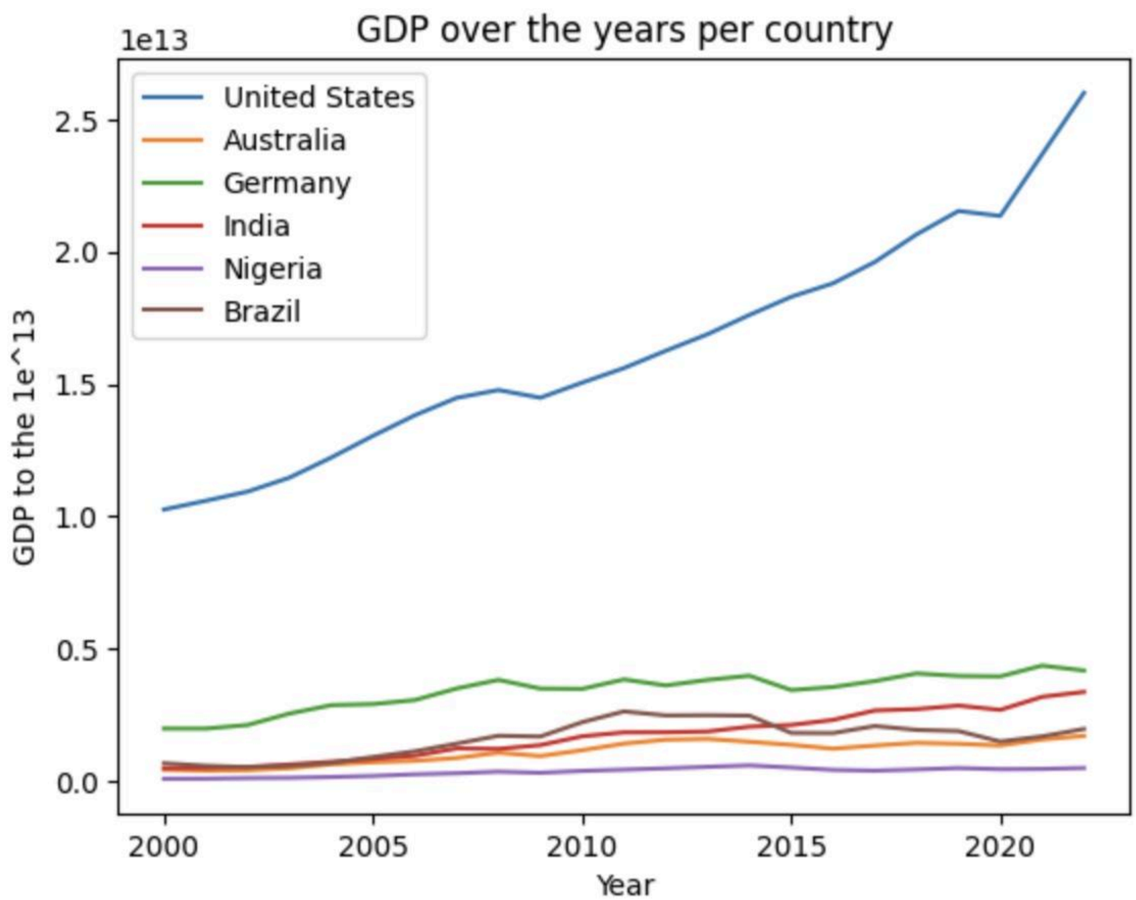
- Here is the code for the population growth graphs

```
years = range(2000,2023)
for country in countries:
  countrypop = []
  for year in years:
    sorted_val = sortcountry(country,df)
    sorted_val = sortyear(year,sorted_val)
    countrypop.append(float(sorted_val['pop_dens_sq_km'].iloc[0]))
  print(f'Difference between population density per sq. km. from 2022 - 2011: {countrypop[len(countrypop)-1] - countrypop[0]}')
  plt.plot(years,countrypop,label = str(country))
  plt.legend()
  plt.xlabel('Year')
  plt.ylabel('Population Density per square kilometer')
  plt.title('Population Density/square km. from 2000-2022')
  plt.show()
```

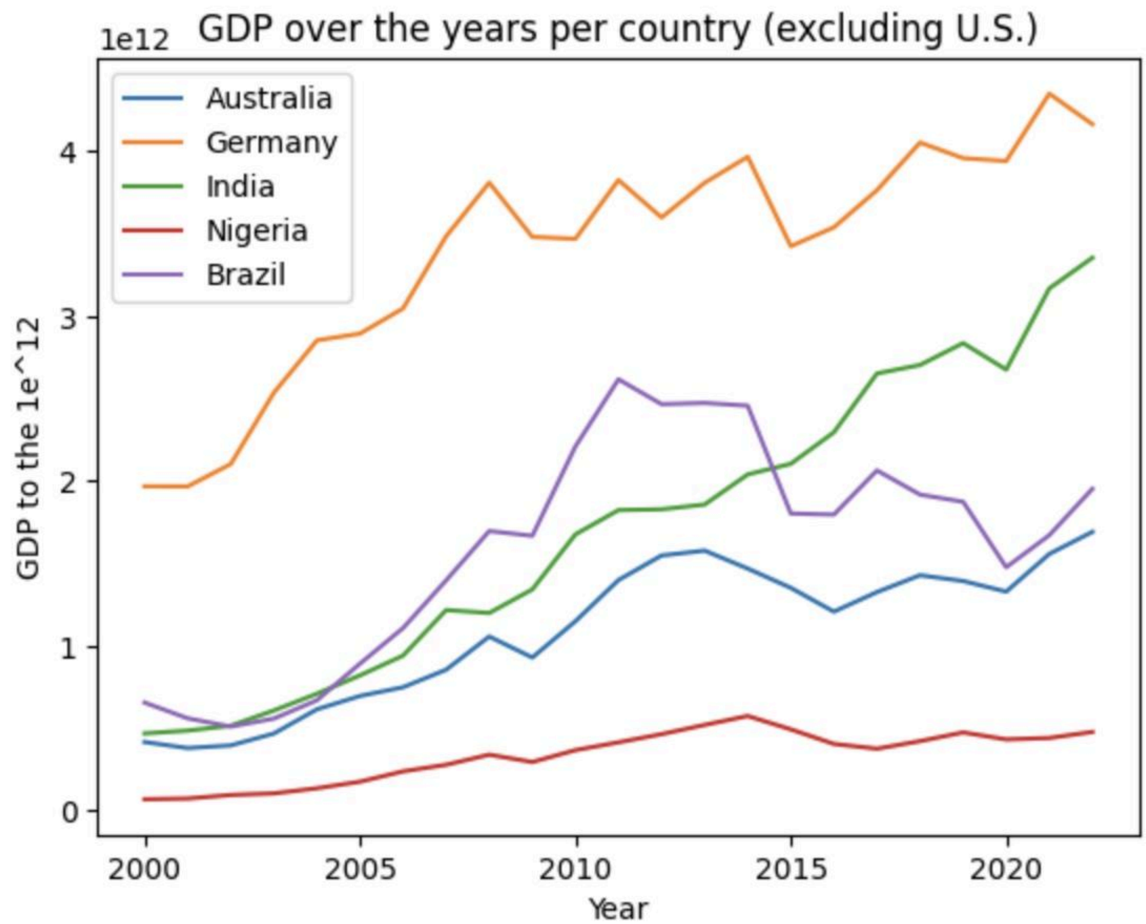
-

- The population has increased nearly linearly for every country

- Same applies for Germany, though there is a massive dip shown in the graph (this is due to a census gap)
- **Second,** we wanted to plot the growth of GDP for each country over 2000 to 2022



-



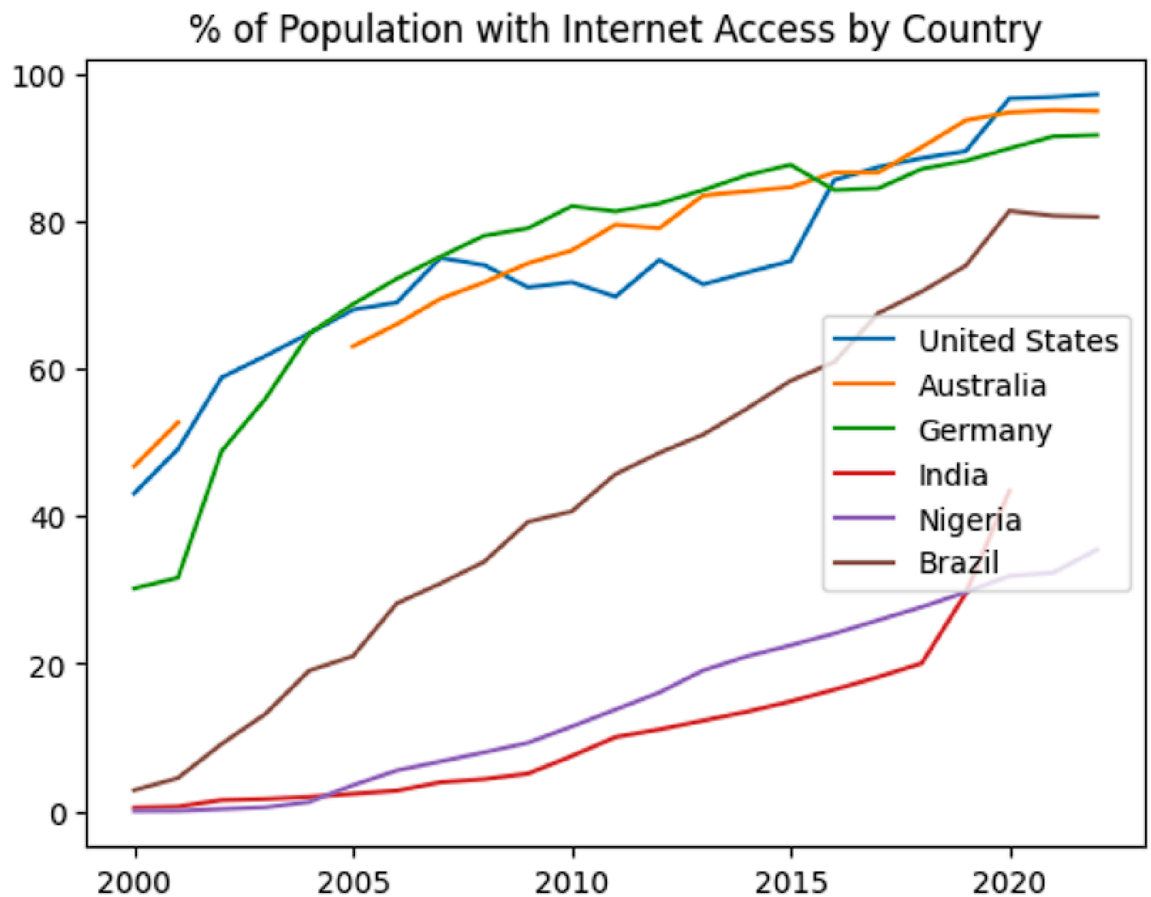
-
- Here is the code for the graphs

```
for country in countries:
  countrypop = []
  for year in years:
    sorted_val = sortcountry(country,df)
    sorted_val = sortyear(year,sorted_val)
    countrypop.append(float(sorted_val['co2_emiss_excl_lulucf'].iloc[0]))
  plt.plot(years,countrypop,label = str(country))
plt.legend()
plt.title("CO2 Emissions from 2000-2022 by Country")
plt.show()

for country in countries[1:6]:
  countrypop = []
  for year in years:
    sorted_val = sortcountry(country,df)
    sorted_val = sortyear(year,sorted_val)
    countrypop.append(float(sorted_val['co2_emiss_excl_lulucf'].iloc[0]))
  # print(f'Difference between population density per sq. km. from 2022 - 2011: {countrypop[len(countrypop)-1] - countrypop[0]}')
  plt.plot(years,countrypop,label = str(country))
plt.legend()
plt.title("CO2 Emissions from 2000-2022 by Country (Excluding USA)")
plt.show()
```

-
- Over the years, we can see that each country's GDP followed a linear growth pattern.
- We created a separate graph without the United States in order to see other country's GDP in more detail.

- This shows us that each country that we analyzed prioritized economic growth.
- **<u>Third</u>**, we wanted to plot the percentage of the population with internet access in each country

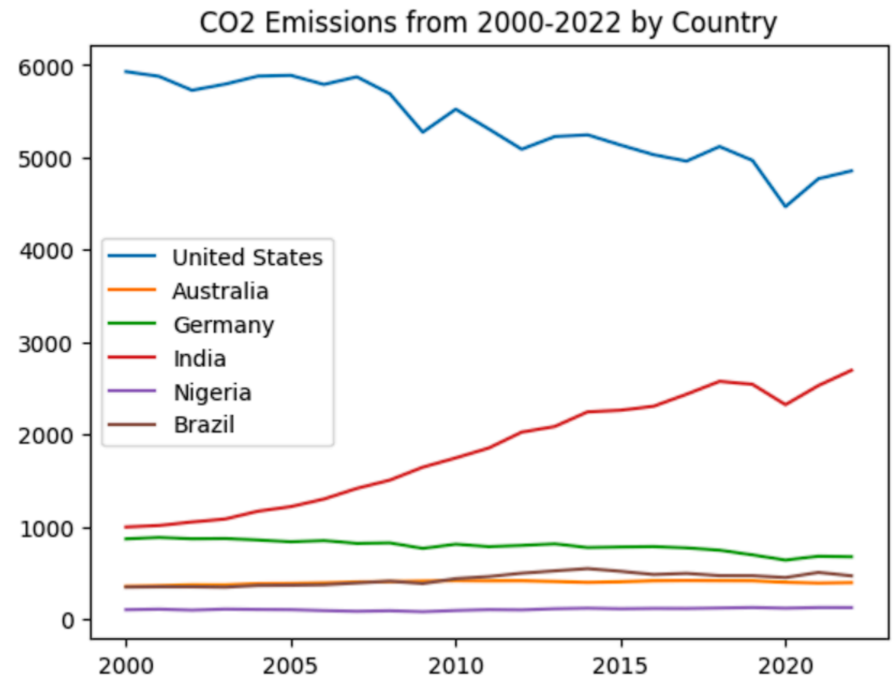


-
- Here is the code for the graph:

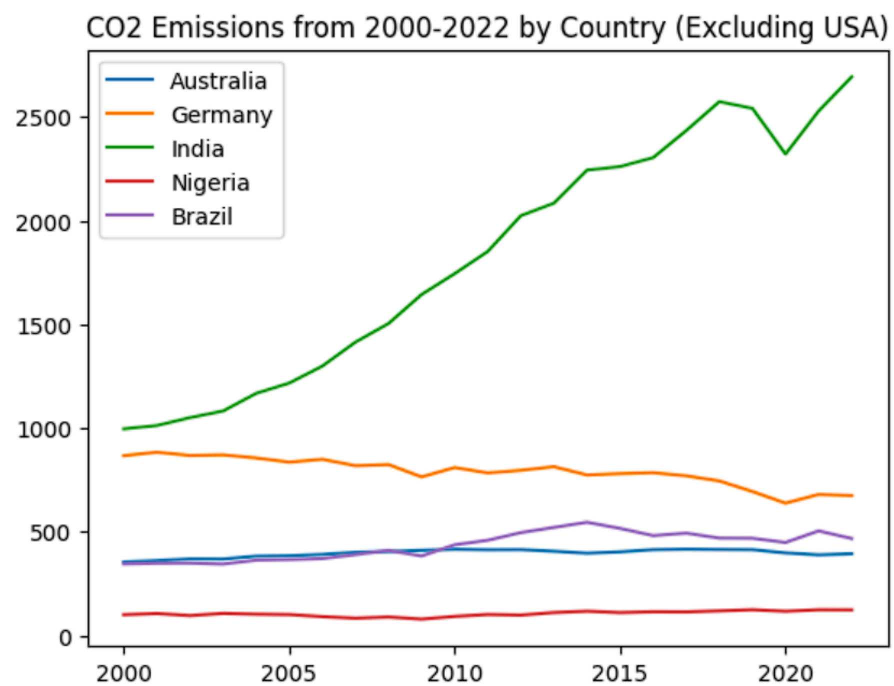
```
for country in countries:
  internetpercent = []
  for year in years:
    sorted_val = sortcountry(country,df)
    sorted_val = sortyear(year,sorted_val)
    internetpercent.append(float(sorted_val['internet_use_pop'].iloc[0]))
  plt.plot(years,internetpercent,label = str(country))
plt.legend()
plt.title("% of Population with Internet Access by Country")
plt.show()
```

- We can see that over time, every single country has had an increase in the population that is accessing the internet
- This undoubtedly contributes to climate change.

- Lastly, we created graphs showing the rise in CO_2 emissions between all six countries.

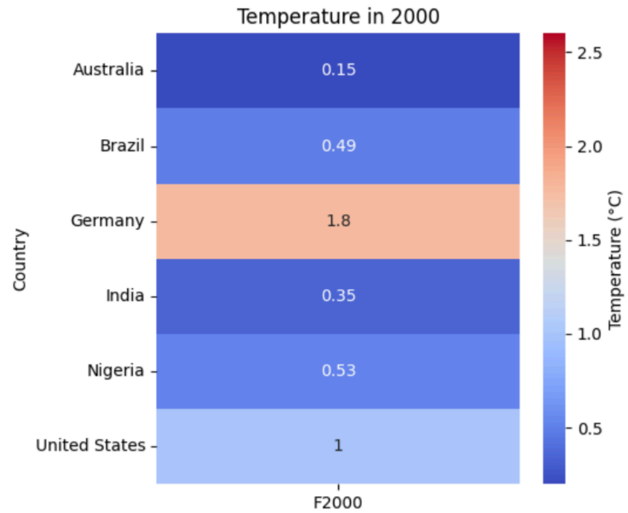


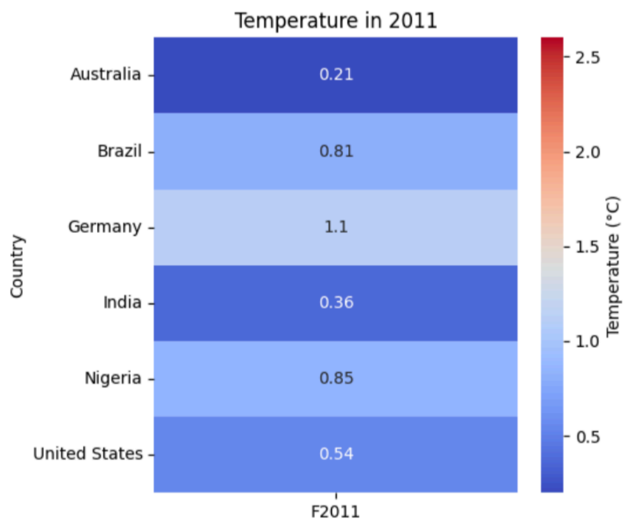
-

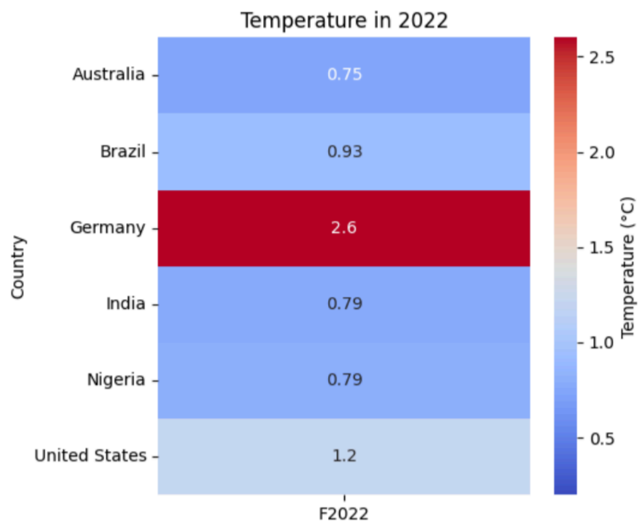


- **Climate Warming Trends Dataset**
  - <u>**First**</u>, we wanted to have a climate map (global map with our countries highlighted) of our countries in each of the three years, but this dataset did not have longitude/latitude data

- So we decided to create a heat map showing the average temperature of each country (in Celsius) in 2000, 2011, and 2022
- The set of three graphs below show the average climate of each country in 2000, 2011, and 2022







- This is the code for these graphs

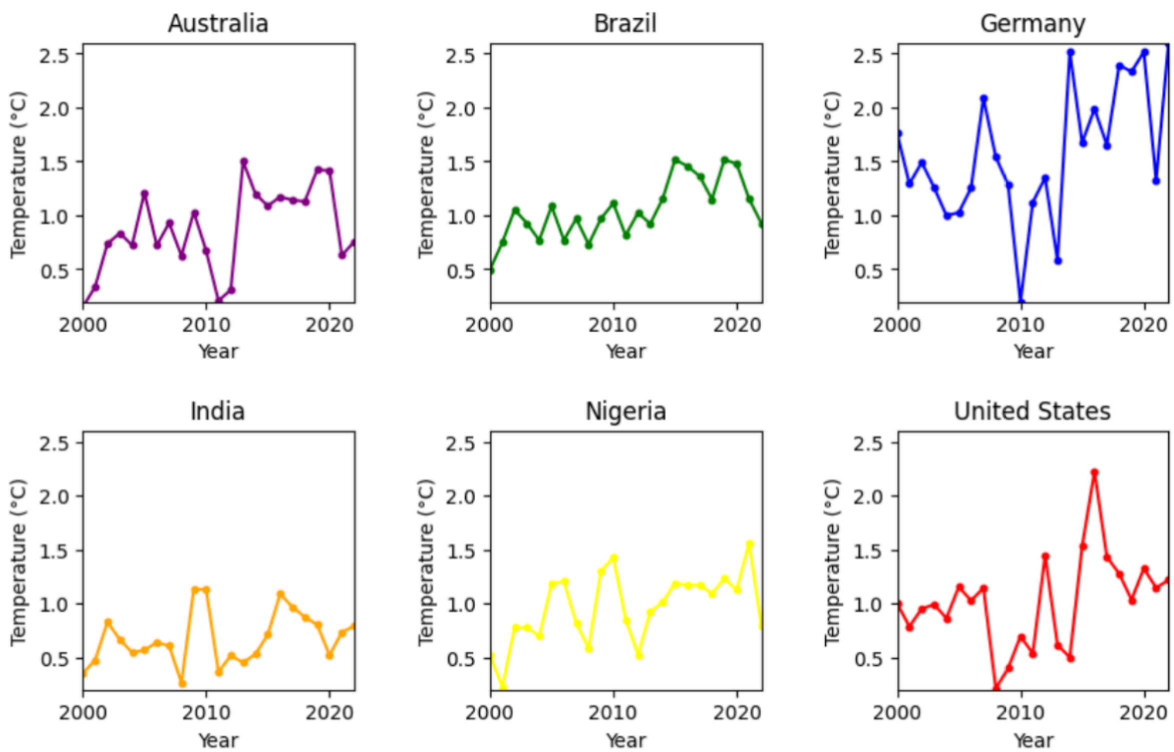
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
asfw = pd.read_csv('wide_format_annual_surface_temp.csv')
asfl = pd.read_csv('long_format_annual_surface_temp.csv')
colors = {
    'United States': 'red',
    'Brazil': 'green',
    'Germany': 'blue',
    'Nigeria': 'yellow',
    'India': 'orange',
    'Australia': 'purple'
}

# 2000 US, Brazil, Germany, Nigeria, India, Australia
asfw2000 = asfw[['Country','F2000']]
plt.figure(figsize=(5, 5))
updated = asfw2000[asfw2000['Country'].isin(['United States','Brazil','Germany','India','Nigeria','Australia'])]
updated.set_index('Country', inplace=True)
sns.heatmap(updated, annot = True, cmap='coolwarm', cbar_kws={'label': 'Temperature (°C)'}, vmin=0.2, vmax=2.6)
plt.title('Temperature in 2000')
plt.show()
#2011 US, Brazil, Germany, Nigeria, India, Australia
asfw2011 = asfw[['Country','F2011']]
plt.figure(figsize=(5, 5))
updated2 = asfw2011[asfw2011['Country'].isin(['United States','Brazil','Germany','India','Nigeria','Australia'])]
updated2.set_index('Country', inplace=True)
sns.heatmap(updated2, annot = True, cmap='coolwarm', cbar_kws={'label': 'Temperature (°C)'}, vmin=0.2, vmax=2.6)
plt.title('Temperature in 2011')
plt.show()
# 2022 US, Brazil, Germany, Nigeria, India, Australia
asfw2022 = asfw[['Country','F2022']]
plt.figure(figsize=(5, 5))
updated2 = asfw2022[asfw2022['Country'].isin(['United States','Brazil','Germany','India','Nigeria','Australia'])]
updated2.set_index('Country', inplace=True)
sns.heatmap(updated2, annot = True, cmap='coolwarm', cbar_kws={'label': 'Temperature (°C)'}, vmin=0.2, vmax=2.6)
plt.title('Temperature in 2022')
plt.show()
```

- From these graphs, we can see that **the average temperature of every country increased drastically at some point in time**
    - For example, Australia went from 0.15 C in 2000 to 0.75 C to 2011, and Germany went from 1.8 C in 2000 to 2.6 in 2011
- **However, every country's average temperature, except for Nigeria, fluctuated as well**
    - In the first interval (2000 to 2011), every country's average temperature increased
    - In the second interval (2011 to 2022), every country's average temperature, except for Nigeria, then decreased quite dramatically

- For example, Australia went from 0.75 C in 2011 to 0.21 C in 2022, and the United States when from 1.2 C in 2011 to 0.5 C in 2022
- ***Interesting mention***: India's temperature doubled in the first interval, then went back to its original temperature over the intervals (0.36 C to 0.79 C back to 0.36 C)
  - The only country in our set with such a pattern
- ***Interesting mention***: Nigeria is the only country, from our selection of countries, whose average temperature continued to increase over the years
  - Is this due to any environmental or socioeconomic situations?

- **Second**, we plotted each country's average temperature as a line graph over the years 2000 to 2022 to compare trends between countries
- Below we have 6 line graphs, one for each country, showing the trends in climate change over the course of 2000 to 2022



- This is the code for these graphs

```python
#Individual Graphs
plt.figure(figsize=(10, 6))
i = 1
for c in updated4['Country'].unique():
    country_data = updated4[updated4['Country'] == c]
    plt.subplot(2,3,i)
    plt.xlabel('Year')
    plt.ylabel('Temperature (°C)')
    plt.title(f'{c}')
    plt.xlim(2000,2022)
    plt.ylim(0.2,2.6)
    plt.scatter(country_data['Year'],country_data['Temperature'],label=c,color = colors[c],s=10)
    plt.plot(country_data['Year'],country_data['Temperature'],label=f'{c}',color = colors[c])
    i+=1
plt.subplots_adjust(hspace=0.5, wspace = 0.5)
plt.show()
```

  - While **every country had some fluctuations, most were able to stay relatively stable, despite their being an overall increase in temperature** over time
  - It is clear from these line graphs that **the United States and Germany had temperature dips and spikes over time compared to the other countries**
    - Can this be attributed to their urbanization and/or socioeconomic status?
- Comparing Data From Both Datasets
  - From our charts from our global urbanization dataset we see a steady pattern of different parts of urbanization increasing such as population density, GDP. C02 emissions, % internet access. From our charts of temperature change we also see a lot of steady increase across all countries, with some drastically more than others. As a result we have identified a likely correlation between urbanization and global warming.
- Final Conclusions
  - Trends of increasing urbanization in the 6 countries we analyzed corresponded to an increasing trend in temperature
  - Therefore we believe that urbanization is a contributing factor to climate change
  - Even more specifically, aspects of urbanization, including urban population and access to internet contribute play roles in urbanization contributing to climate change