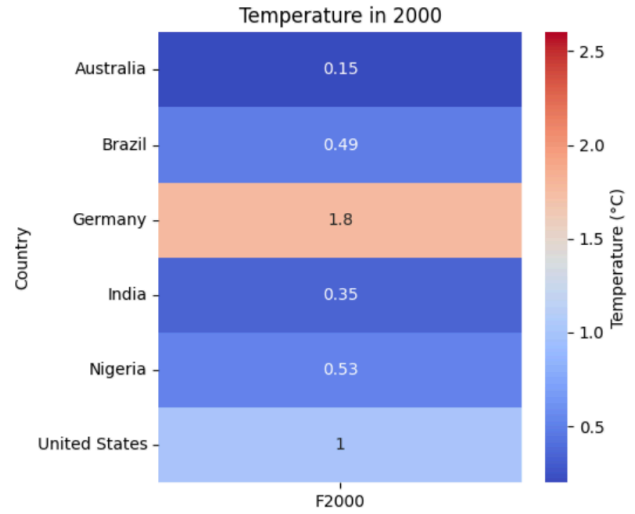
What is the Correlation Between Urbanization and Climate Change?

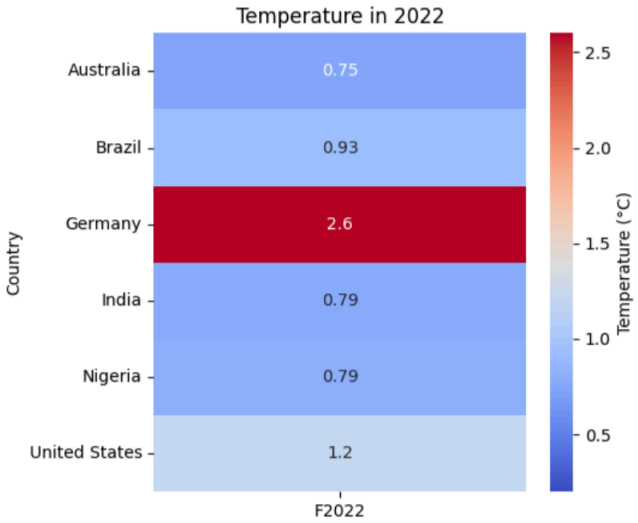
- **Background**: from sustainability presentation
  - 

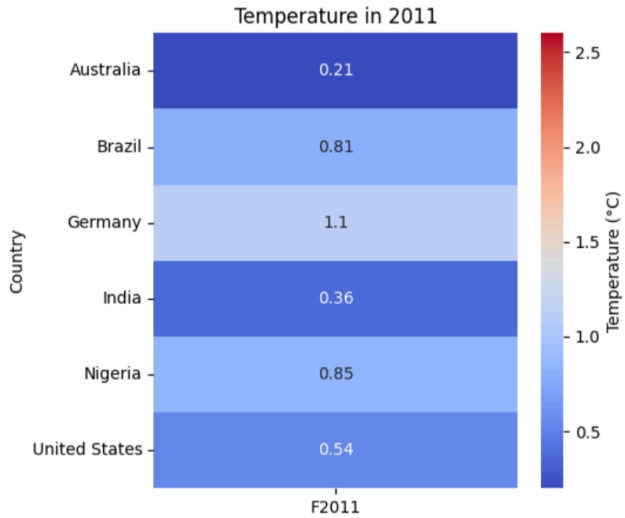
  - We prioritize our economy (bottom line and cash flow) and tend to leave the environment for the backburner (conservation of resources and ecosystems)
  - There are already national and global frameworks/policies in place to regulate climate change (Paris Climate Accords) by targeting things like temperature, 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**
  - The more urbanized a country is, the greater effect is has on climate change
  - Does the urban vs rural 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
  - **Climate 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**
    -

- **Climate Warming Trends Dataset**
    - **First**, 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 graph 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





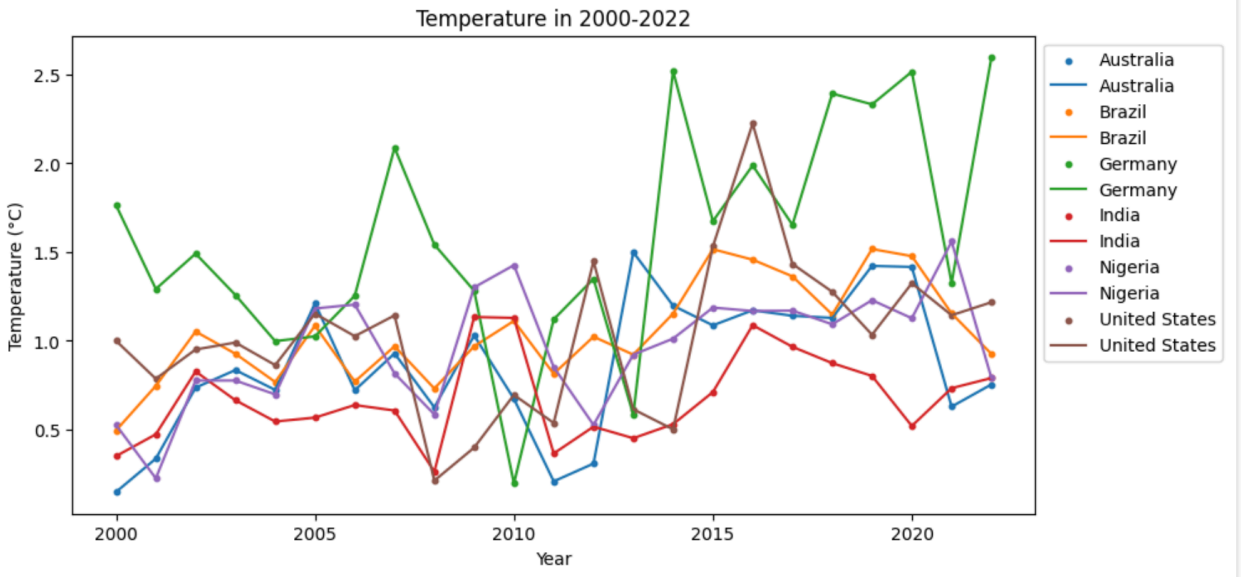


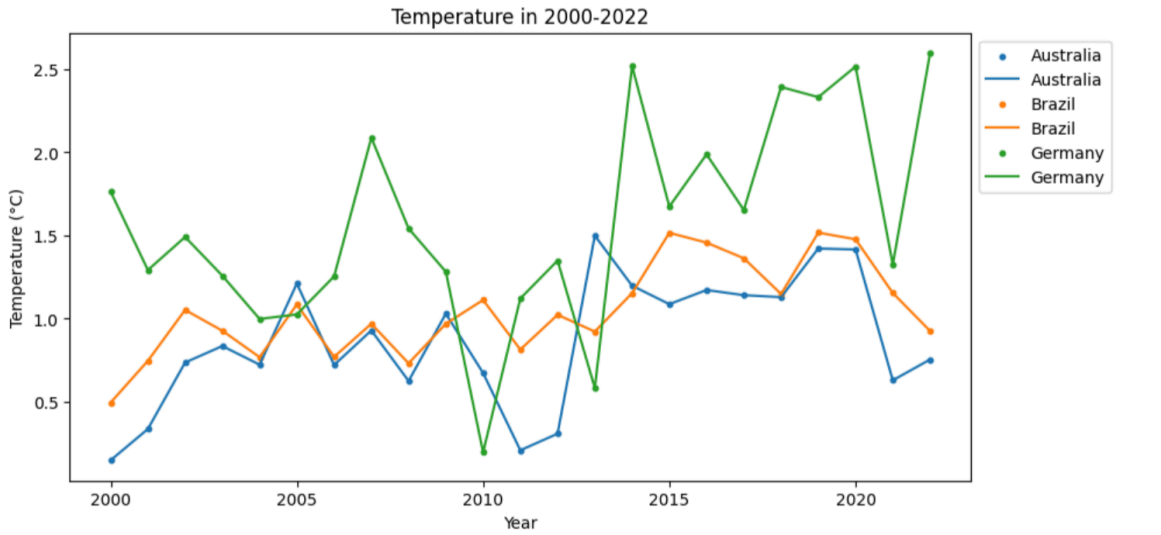
- From these graphs, we can see that **the average temperature of every country increased drastically at some point in time**

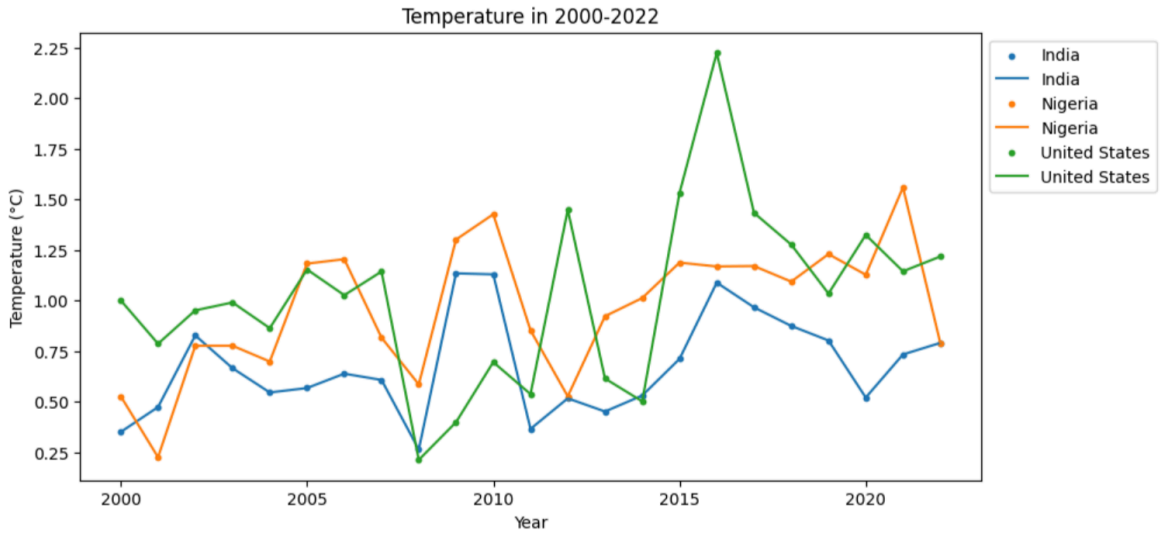
    - For example, Australia went from 0.15 C in 2000 to 0.75 C to 2011, and Germany went from 1.8 C in 2000 to 2.6 in 2011
- **However, every country's average temperature, except for Nigeria, fluctuated as well**
    - In the first interval (2000 to 2011), every country's average temperature increased
    - 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)
    - ***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 socio-economic situations?

```python
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')

# 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()
# all 6 countries over 2000-2022
asfl['Year'] = asfl['Year'].str[1:].astype(int)
asfl_u = asfl[(asfl['Year'] >= 2000) & (asfl['Year'] <= 2022)]
updated4 = asfl_u[asfl_u['Country'].isin(['United States','Brazil','Germany','India','Nigeria','Australia'])]
plt.figure(figsize=(10, 5))
for c in updated4['Country'].unique():
    country_data = updated4[updated4['Country'] == c]
    plt.scatter(country_data['Year'],country_data['Temperature'],label=c,s=10)
    plt.plot(country_data['Year'],country_data['Temperature'],label=f'{c}')
plt.title('Temperature in 2000-2022')
plt.xlabel('Year')
plt.ylabel('Temperature (°C)')
plt.legend(bbox_to_anchor=(1, 1), loc='upper left')
plt.show()
```




Temperature in 2000-2022
Temperature (°C)
Year
2000
2005
2010
2015
2020
0.5
1.0
1.5
2.0
2.5
Australia
Australia
Brazil
Brazil
Germany
Germany


Temperature in 2000-2022
Temperature (°C)
Year
2000
2005
2010
2015
2020
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
India
India
Nigeria
Nigeria
United States
United States

```
# Australia, Brazil, Germany
updated5= asfl_u[asfl_u['Country'].isin(['Brazil','Germany','Australia'])]
plt.figure(figsize=(10, 5))
for c in updated5['Country'].unique():
    country_data = updated5[updated5['Country'] == c]
    plt.scatter(country_data['Year'],country_data['Temperature'],label=c,s=10)
    plt.plot(country_data['Year'],country_data['Temperature'],label=f'{c}')
plt.title('Temperature in 2000-2022')
plt.xlabel('Year')
plt.ylabel('Temperature (°C)')
plt.legend(bbox_to_anchor=(1, 1), loc='upper left')
plt.show()
# India, Nigeria, US
updated6 = asfl_u[asfl_u['Country'].isin(['United States','India','Nigeria'])]
plt.figure(figsize=(10, 5))
for c in updated6['Country'].unique():
    country_data = updated6[updated6['Country'] == c]
    plt.scatter(country_data['Year'],country_data['Temperature'],label=c,s=10)
    plt.plot(country_data['Year'],country_data['Temperature'],label=f'{c}')
plt.title('Temperature in 2000-2022')
plt.xlabel('Year')
plt.ylabel('Temperature (°C)')
plt.legend(bbox_to_anchor=(1, 1), loc='upper left')
plt.show()
```

- Comparing Data From Both Datasets
- Final Conclusions