INFO-I 590 DATA VISUALIZATION

Prof. Sadamori Kojaku

Climate Change and Its Effects

Submitted By - Atharv Pramod Jangam (ajangam)

Luddy School of Informatics, Computing, and Engineering

Indiana University Bloomington

Table of Contents

1. Abstract Introduction and Keywords

1.1 Abstract

Climate change is a global phenomenon that is causing significant shifts in the Earth's climate, including increases in average temperature, changes in precipitation patterns, and rising sea levels. These changes are largely driven by human activities, such as burning fossil fuels and deforestation, which release large amounts of greenhouse gases into the atmosphere. The impacts of climate change are wide-ranging and include more frequent and intense heatwaves, droughts, extreme weather events, and disruptions to agriculture, natural ecosystems, and human health. To address the challenges of climate change, it is essential to visualize the climate change's impact and make people aware of it.

1.2 Introduction

There is a lot of controversy about is climate change is real and whether is it really affecting our environment this project will focus on the importance of climate change and figure out if is it real or not.

Data visualization is about telling the audience a story, which can help them understand and communicate the complex issues surrounding climate change. By presenting data in a visual format, we can gain insights and identify trends that would be difficult to see in raw data alone. This project will use data visualization to explore the various impacts of climate change and show how it affects different regions and populations of India and worldwide.

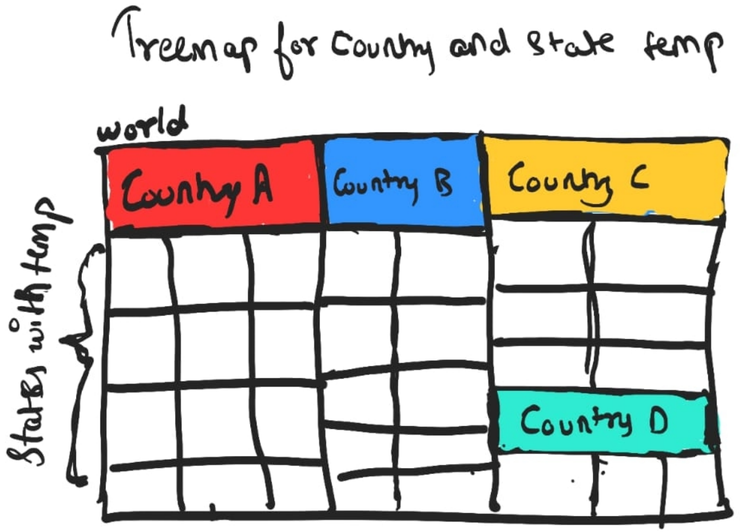
We are focusing on climate change in India and what factors are responsible for the change. There are various interactive visualizations such as Treemap, Time series graph, Bubble chart, Choropleth map, etc. Which will provide our audience with a deep understanding of climate change.

The data sources we have used are from various websites, and we will provide a link in a separate column.

1.3 Keywords

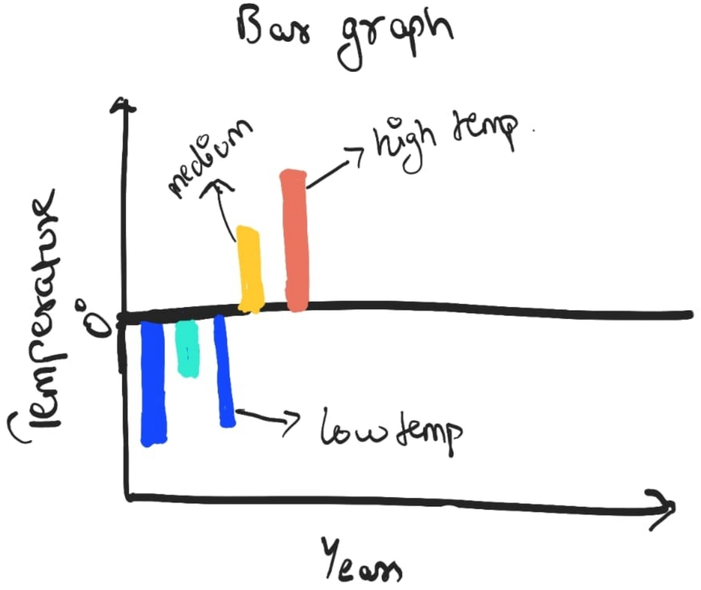
Average temperature change, Country, India, global land temperature, plotly, LOWESS trend, Tableau, Dataset, Treemap, Bubble char, Bar graph, Choropleth map, Area graph, GDP, Population, Greenhouse

Rough Visualization



I decided to create a treemap as it can give us a deep understanding of temperature change and the highest temperature in a country and a state.

A bar graph would compare temperature with the years; we can also color bars according to the temperature.



2. Data Description

The data we have used is taken from various sources, freely available on the website mentioned below. Also, since the data was extracted from government websites, it didn't require cleaning for most parts.

1. **Global Land Temperature:** This data consists of global land temperature from the years 1900 - 2022 and it was extracted from National Center for Environmental Information (click here for the dataset).

2. **Average Temperature in India:** This data consists of the average annual temperature in India from the years 1900 -2017 with a year and annual average temperature as columns and it was extracted from the Indian government website (click here for the dataset).

3. **Earth Surface Temperature data:** This data consists of the temperature change from the years 1744 - 2013 there are 645675 rows and 5 columns and we have performed data cleaning on this data set to remove null values and only extract data for Indian states. There were 25648 null values which we have dropped and also we have changed the data type of the date columns to Datetime format. It is freely available on Kaggle. (click here for the dataset).

4. **Greenhouse emission:** This data consists of the greenhouse gas emission from 1990 - 2019. It was extracted from the ClimateWatch website and needed some data cleaning which involves making the data tidy. (click here for the dataset).

5. **GDP vs. Carbon Emission:** This data consists of GDP vs. carbon emission in 2018. We have merged two datasets for carbon emission from point 4(i.e., ClimateWatch); we got the data for carbon emission and merged it with the GDP and population data we collected from Worldbank. (GDP data, Population data).

3. Methodologies Used

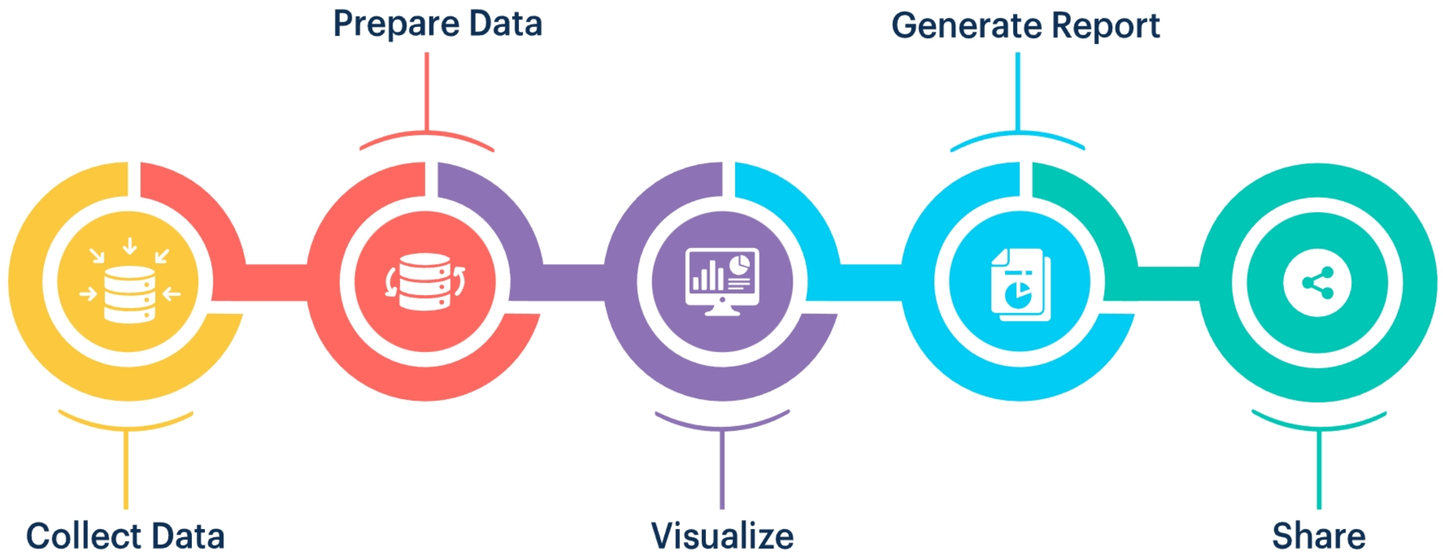


Image Source - https://www.sigmadatasys.com/wp-content/uploads/2019/10/05-data-visualization.svg?x88783

3.1 Data Cleaning

As mentioned earlier, most of the data we collected was clean and did not require much cleaning. But as the data was collected from different sources, we had to merge it into a single data file.

First, we need to analyze the data and perform EDA to check if there are any null values and if the datatype is correct. The only dataset that required data cleaning was the Global land temperature taken from Kaggle. The data set had multiple null values, which we dropped. Also changed the datatype of the date column, which was previously an object type and needed to be converted to Datetime format.

	dt	AverageTemperature	AverageTemperatureUncertainty	State	Country
0	1855-05-01	25.544	1.171	Acre	Brazil
1	1855-06-01	24.228	1.103	Acre	Brazil
2	1855-07-01	24.371	1.044	Acre	Brazil
3	1855-08-01	25.427	1.073	Acre	Brazil
4	1855-09-01	25.675	1.014	Acre	Brazil

Fig 1: Global average land temperature

```
dt                               object
AverageTemperature              float64
AverageTemperatureUncertainty   float64
State                            object
Country                          object
dtype: object
```

Fig 2: Global average land temperature Datatypes

```
dt                                  0
AverageTemperature              25648
AverageTemperatureUncertainty   25648
State                               0
Country                             0
dtype: int64
```

Fig 3: Total Null Values

```
state_df.shape
```

```
(645675, 5)
```

Fig 4: Shape of the data

	Date	AverageTemperature	State	Country
0	1855-05-01	25.544	Acre	Brazil
1	1855-06-01	24.228	Acre	Brazil
2	1855-07-01	24.371	Acre	Brazil
3	1855-08-01	25.427	Acre	Brazil
4	1855-09-01	25.675	Acre	Brazil

620027 rows × 5 columns

Fig 5: After dropping Null values

Here we can see that the name of the column 'dt' has also been changed to "Date"

```
Date                  datetime64[ns]
AverageTemperature           float64
State                         object
Country                       object
dtype: object
```

Fig 6: After changing the datatype

In this data, we didn't have the latitude and longitude of the states. We first only took values for Indian states and then merged different CSV files with the latitude and longitude to plot the choropleth map

```
Final_df = pd.merge(df_filtered, lat)
```

	Date	AverageTemperature	State	Country	latitude	longitude
79121	2012-06-01	30.833	West Bengal	India	22.58039	88.329947
79122	2012-07-01	28.667	West Bengal	India	22.58039	88.329947
79123	2012-08-01	28.842	West Bengal	India	22.58039	88.329947
79124	2012-09-01	28.496	West Bengal	India	22.58039	88.329947
79125	2012-10-01	26.182	West Bengal	India	22.58039	88.329947

Fig 7: Merging two CSV files

We can see here finally we got a our finally dataset with latitude and longitude for all the Indian states

Some of the data files need to be made tidy. Making a data tidy means: Each variable must have its own column, each observation must have its own row and each value must have its own cell.

3.2 Bar chat for Global land temperature anomalies from 1900-2022

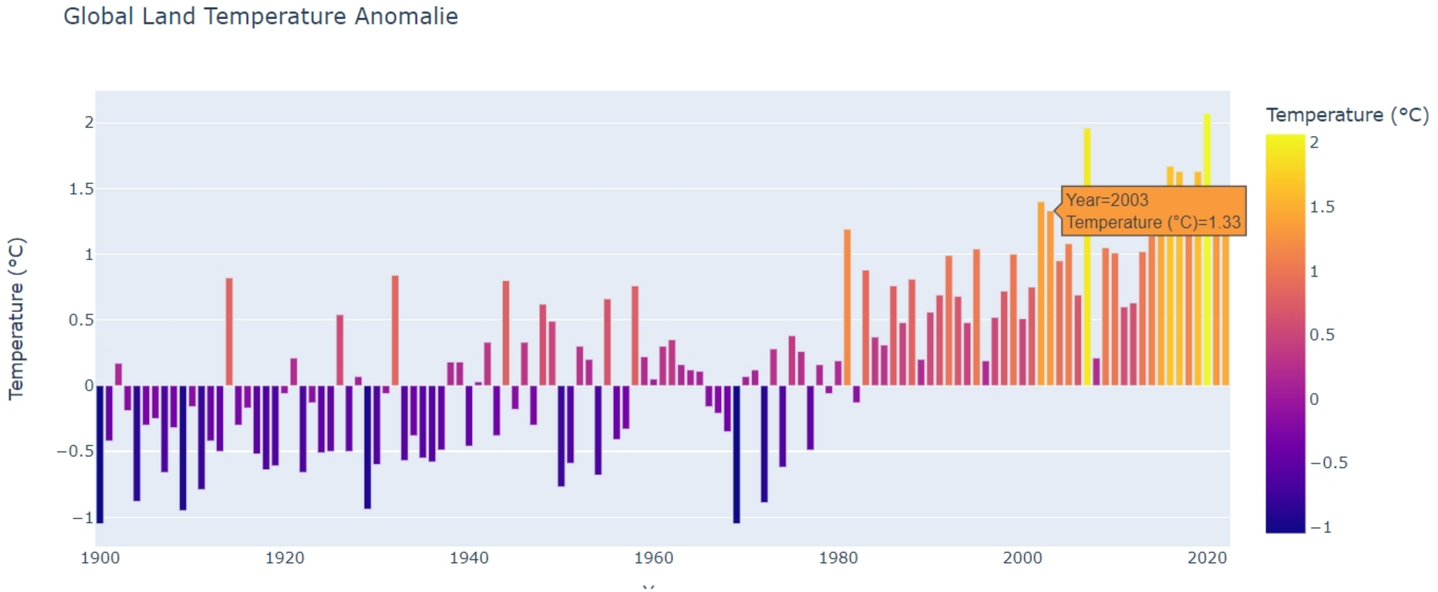


Fig 8: Interactive Bar Graph for Global Land temperature from 1900-2022

This bar graph is useful for comparing temperature anomalies over time. In this visualization, we have used plotly to plot a graph. Where the x-axis is the year and the y-axis is the temperature in °C.

From the Graph above, we can see the change in the global land temperature and say that from 1900-2022 there was a massive change in the temperature. The significant change occurred in 1980, and it is still increasing. Now let us find out for India.

3.3 Line graph and LOWESS for India Temperature analysis

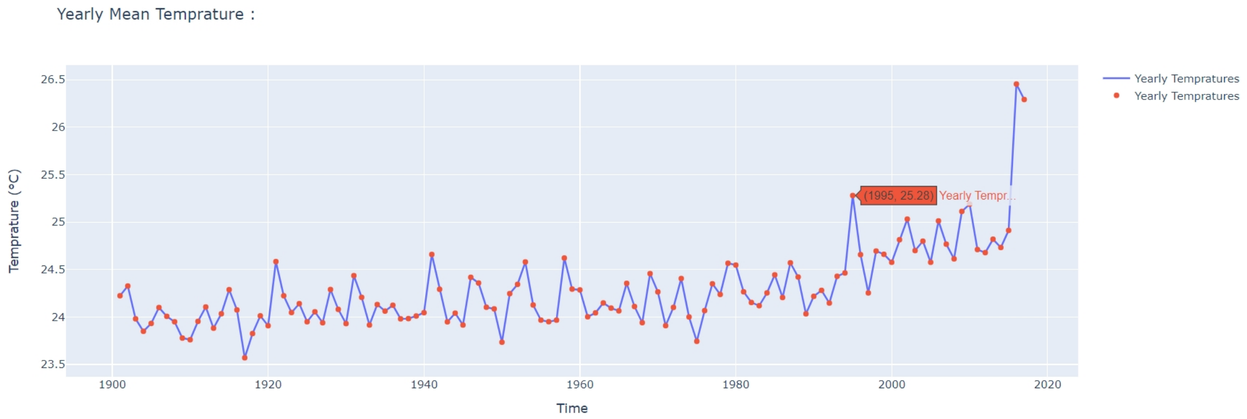


Fig 9: Interactive Scatterplot with mode as line

In this graph, we have used a scatterplot with mode as a line. It gives us an interactive plot for yearly mean temperature change in India from 1900-2017.

We are using Locally weighted scatterplot smoothing, also known as LOWESS (Locally Weighted Scatterplot Smoothing)

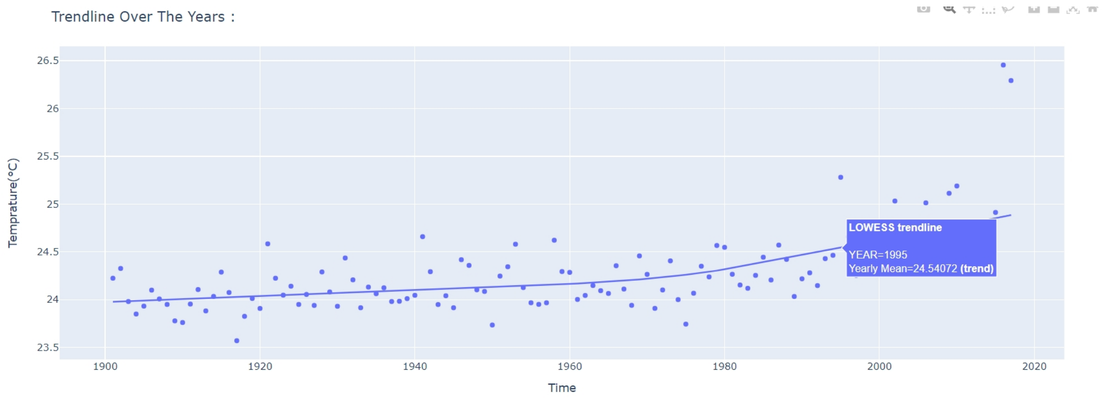


Fig 10: Interactive Scatterplot with LOWESS trend

These graphs show that after 1995 and 2015, there was a massive change in the yearly mean temperature. Also, the LOWESS trendline is increasing, which suggests an increase in temperature in India.

The average temperature in India has increased from 1990 to 2016 by 2 degree Celsius. This trend of increasing annual temperature for India is in line with the increasing global surface temperatures. A significant change occurred in 1980 it can be due to the urbanization that peaked after India's independence in 1950.

3.4 Treemap for country and state temperature

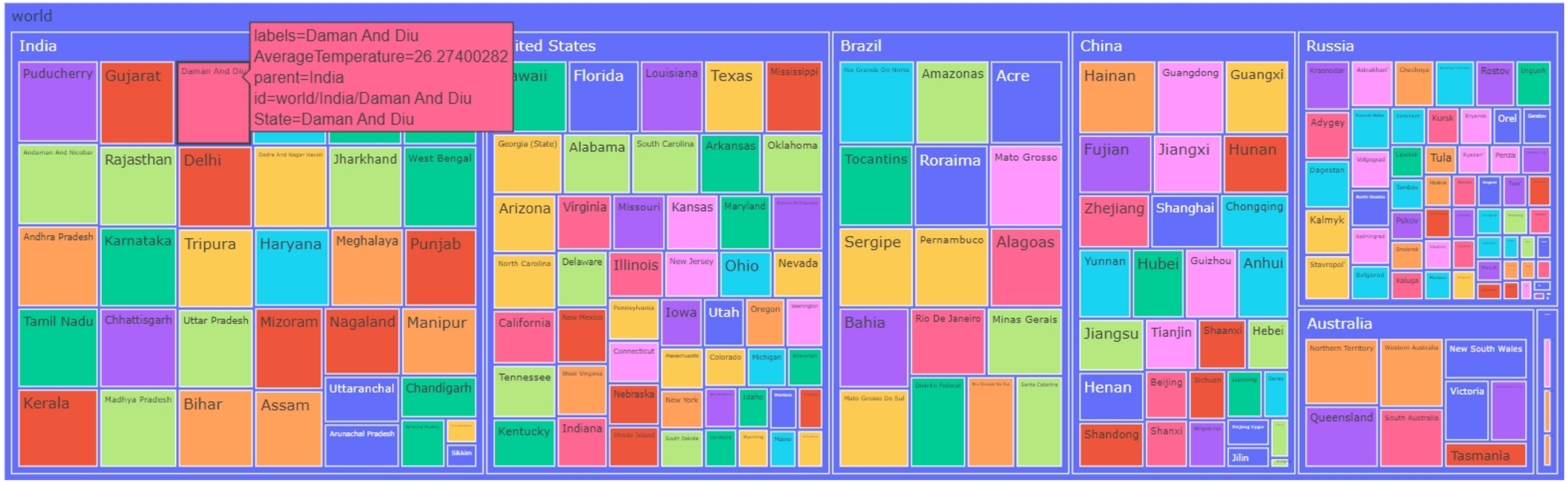
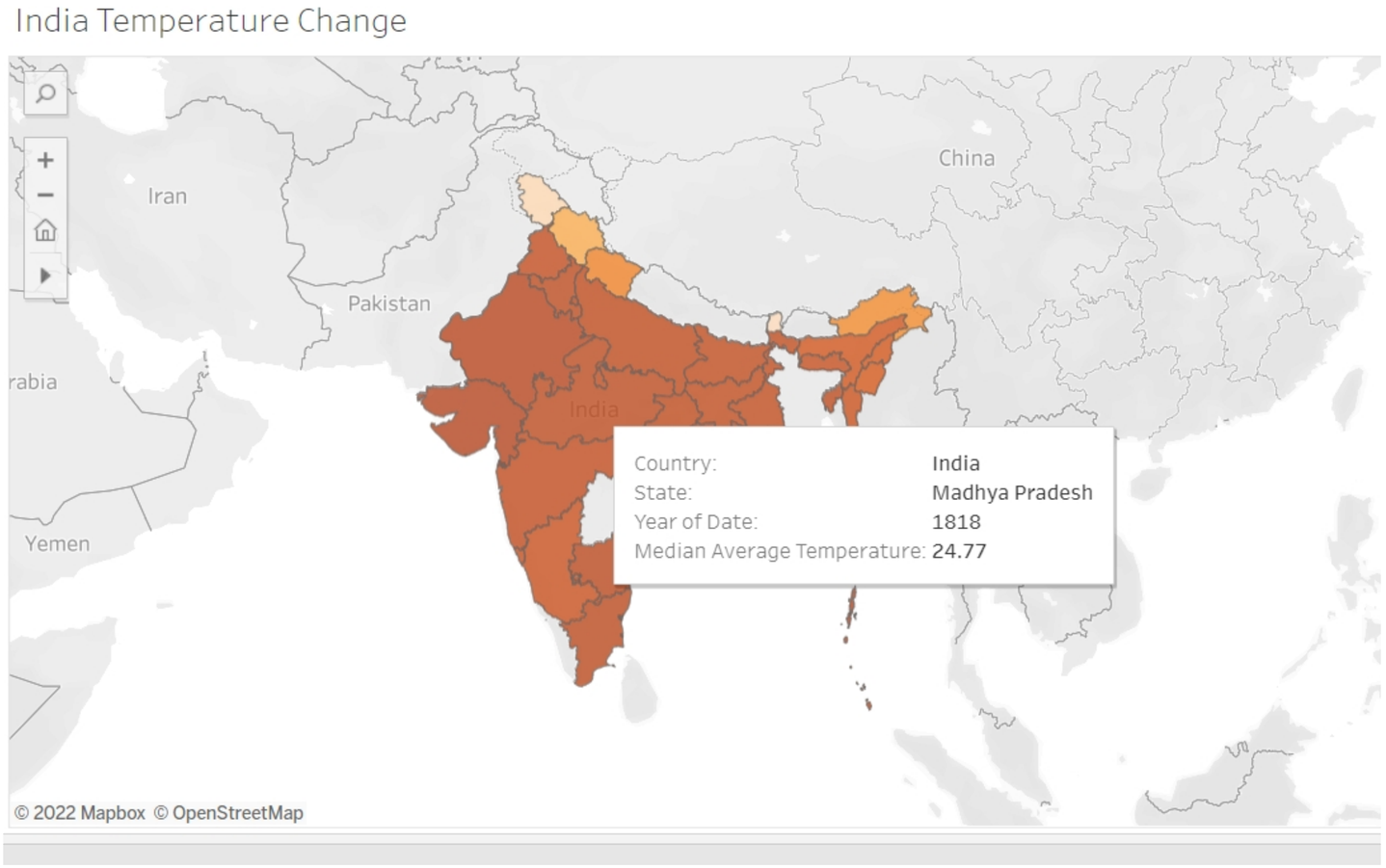


Fig 11: Interactive Treemap for countries and its states temperature

A treemap chart provides a hierarchical view of your data and makes it easy to spot patterns, such as which country and its states had the most average temperature. And we can see that Puducherry in India had the maximum

3.5 Time series Choropleth map using Tableau



[Please click here to play the video on YouTube.]

A choropleth map is a type of map that uses different colors or shades to represent different values or ranges of data. Choropleth maps are often used to show the spatial distribution of a particular variable; in our case, it shows the median average temperature change from 1750-2013. In the video above we have only included the visualization from the years 1950-2013 as there was very fewer data available from the past and the demographic of Indian states has changed. In the map, we can see that some states (i.e., Telangana and Kerela) are in grey because Telangana was considered a separate state after 2014, and as we have data till 2013, it is showing it in grey. We have passed latitude and longitude as x and y axis respectively and keeping median average temperature as our target

From the above Choropleth map of India, we can see that In Madhya Pradesh, the median average temperature in 1950 was 25.77°C, Rajasthan 27.35°C, and Uttaranchal 15.44°C. But in 2013, the temperature changed drastically. The median average temperature of Madhya Pradesh, Rajasthan, and Uttaranchal became 26.93°C, 29.64°C, and 18.75°C, respectively.

We might think 2°C is not more, but it is more than enough to cause glaciers to melt and can cause heatwaves. Without concomitant increases in precipitation, heatwaves can lead to water shortages and increased stress for plants, particularly in arid regions.

Check this article by NASA - https://climate.nasa.gov/news/2865/a-degree-of-concern-why-global-temperatures-matter/

From all the visualization above, we can see an apparent change in the temperature, but what could be the reason behind this change? We Dwell in a Greenhouse gas. Energy from the Sun is essential to life on Earth. Most of the light that reaches the Earth's atmosphere is absorbed by the atmosphere and clouds as it travels to the surface, where it is subsequently reflected upward as infrared heat. The greenhouse gases then absorb around 90% of this heat and skip it toward the surface.

We have data about greenhouse gas emissions. Let's find out if we can relate the change in the gas with the difference in the temperature.

In our dataset, there are four columns years(1990 - 2019), CH4(also known as methane), CO2(carbon dioxide), and N2O(Nitrogen oxide), and these are the gases that are considered responsible for the greenhouse effect.

	CH4	CO2	N2O
Year			
1990	511.25	342.41	147.25
1991	516.81	386.06	151.40
1992	518.28	405.13	155.55
1993	522.27	430.19	159.09
1994	525.93	464.74	164.80

Fig 12: Greenhouse gas data.

Here is a quick overview of the DataFrame.

3.6 Area graph for greenhouse gas emission in India

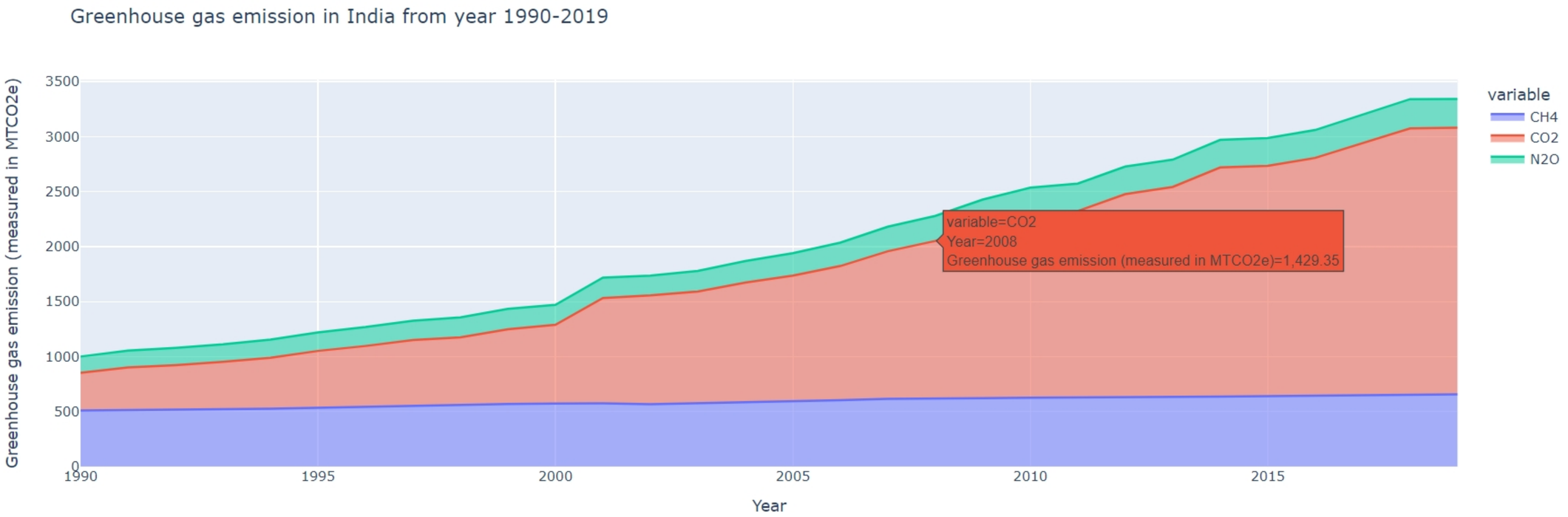


Fig 13: Greenhouse gas area graph.

Over the past 30 years, we can see that CO2, methane, and n20 have been rising. Due to increasing industrialization, globalization, and the combustion of fossil fuels, in India, carbon dioxide makes up the bulk of greenhouse gas emissions. Studies have revealed that nitrous oxide and methane emissions are far more destructive in terms of their influence on climate change, even though they are smaller.

3.7 Bubble chart for GDP, Population, and Carbon emission

We have collected this data from multiple sources and are required to merge it to get an attractive visualization known as a bubble chart. An extension of a scatterplot, a bubble chart, is used to visualize relationships between GDP, population, and carbon emission shown below. Graph showing the logarithmic relationship between GDP and carbon dioxide emissions for various nations in 2018. Metric tons of carbon dioxide equivalent, or MTCO2e, is used to measure carbon dioxide emissions. USD is used to measure GDP.

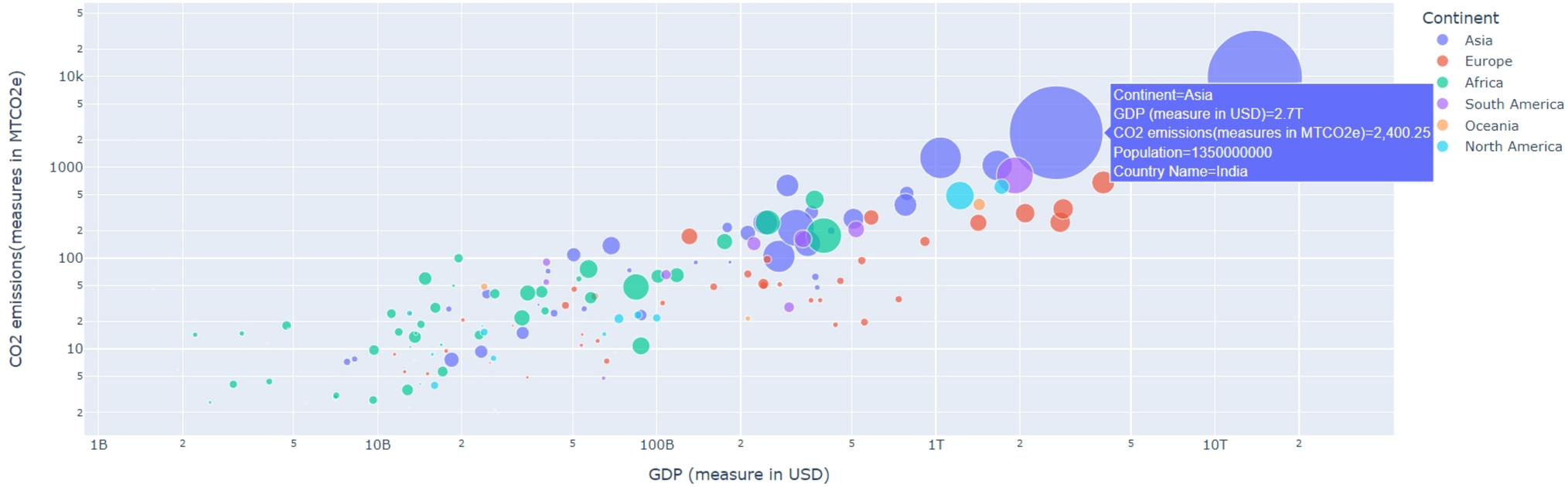


Fig 14: Bubble chart to analyze the relationship between GDP, population, and CO2 emission.

```
gdpcm_df.head()
```

	Continent	Country Name	CO2 emissions	GDP	Population
0	Asia	Afghanistan	7.59	18400000000	37171922
1	Europe	Albania	5.32	15100000000	2866376
2	Africa	Algeria	151.87	175000000000	42228415
3	Africa	Angola	62.93	101000000000	30809787
4	South America	Argentina	207.11	518000000000	44494502

Fig 15: GDP, population, and CO2 emission data.

Here is a quick overview of the Data.

The nation with greater Gross Domestic Product (GDP) contributes more to carbon dioxide emissions, as shown by the positive connection between GDP and carbon dioxide emissions for 2018. The correlation between the emissions and a nation's population serves as further evidence that greenhouse gas emissions are indeed caused by human activity. Also evident is how little the African countries contribute to global carbon dioxide emissions.

4. Discussion

All the above visualization shows that climate change is real and is slowly eating mother nature. Climate change is a long-term warming of the planet and the associated changes in weather patterns. This warming is primarily caused by the increased levels of greenhouse gases, such as carbon dioxide and methane, in the atmosphere, as shown in the above visualization. These gases trap heat from the sun and prevent it from escaping into space, causing the Earth's temperature to rise.

India is one of the top countries responsible for this change. It is not only affecting the people who live in India but also it can be the main factor behind the shift in the overall world. One of the most important factors includes an increase in the population and growth in the GDP. Also evident is how little the African countries contribute to global carbon dioxide emissions. Because of this, not all nations have the same share of the blame for greenhouse gas emissions. The Asian nations with a developing industrial sector and a shift toward industrialization and globalization are more responsible for the increase in carbon dioxide emissions. Many European nations with lower populations contribute significantly or nearly as much as those with bigger populations and comparable GDPs. However, the lack of emphasis on per-capita emissions in the climate agreements discourages developing nations from ratifying the treaty.

4.1 Existing Bad Visualization

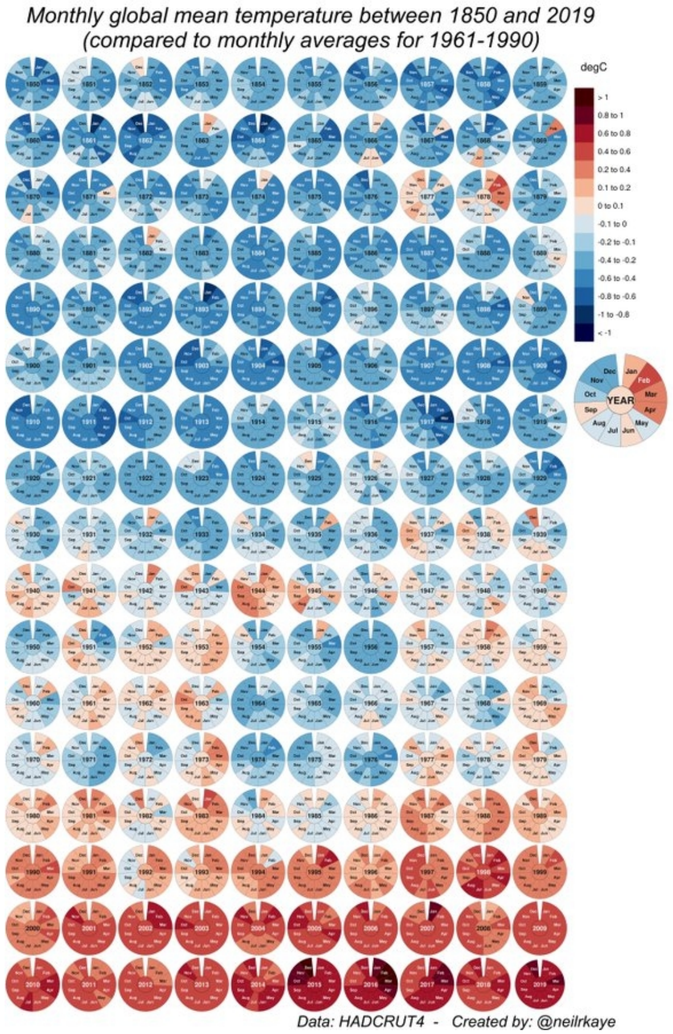


Image source: https://pstblog.com/2016/10/26/climate-change

This visualization is about the monthly global mean temperature between 1850 and 2019. However, this visualization uses a good color scheme. We can see in this visualization that most of the data is unclear, as a lot of data has been plotted in a single graph. It is also hard to see the hue at some point as it looks the same. Overall, the plot is good and conveys how the change has occurred. But it would have been better if the designer had used dynamic time series instead of plotting all the years in the same plot.

This visualization shows us the cities that are prone to the risk of climate change. The point to notice here is there are no labels shown x-axis and y-axis. It is hard for the reader to understand what the y-axis percentage suggests. Also, again there is a lot of data that has been tried to put in this bubble graph, making it look messy. If we had to show the cities, it would have been better to color according to the countries, not continents. We can also see that a lot of the bubbles are overlapping, making it difficult to read.

To improve this graph, we can create a dynamic visualization where a reader can select which country he wants to look at the change in climate risks.

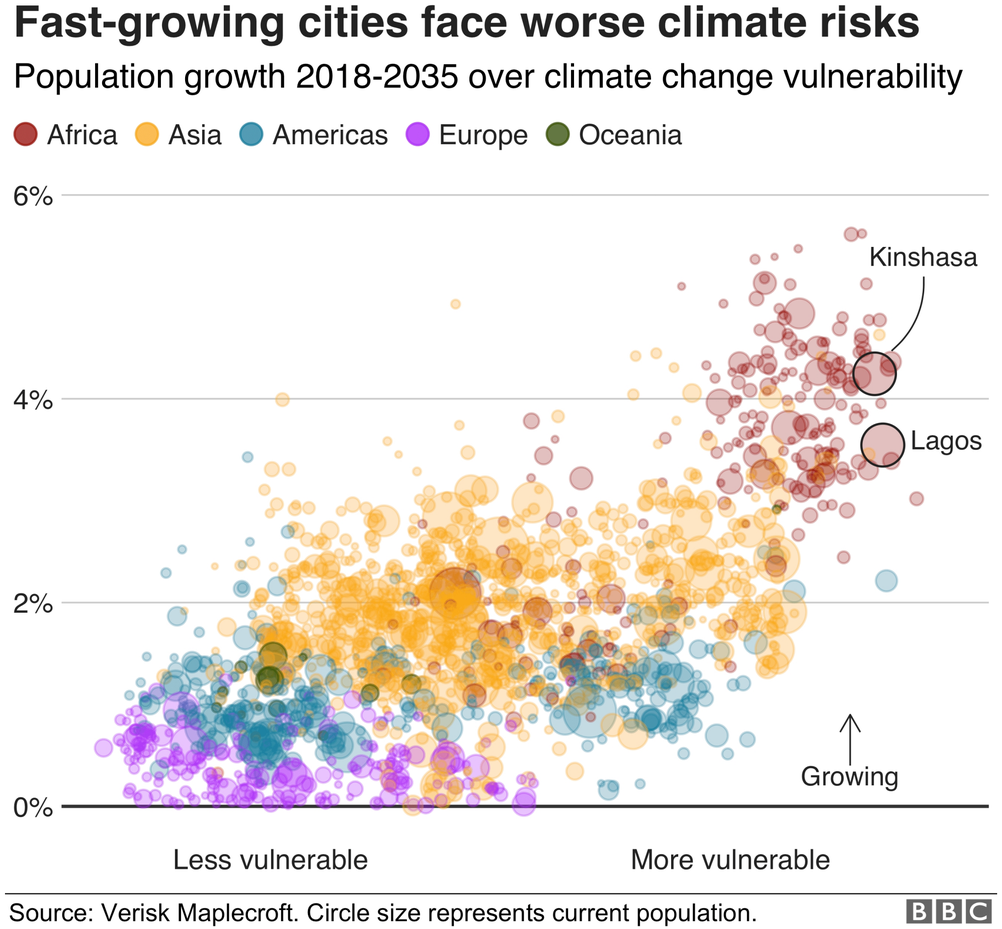


Image source: https://www.bbc.com/news/science-environment-46384067

4.2 Existing Good Visualization

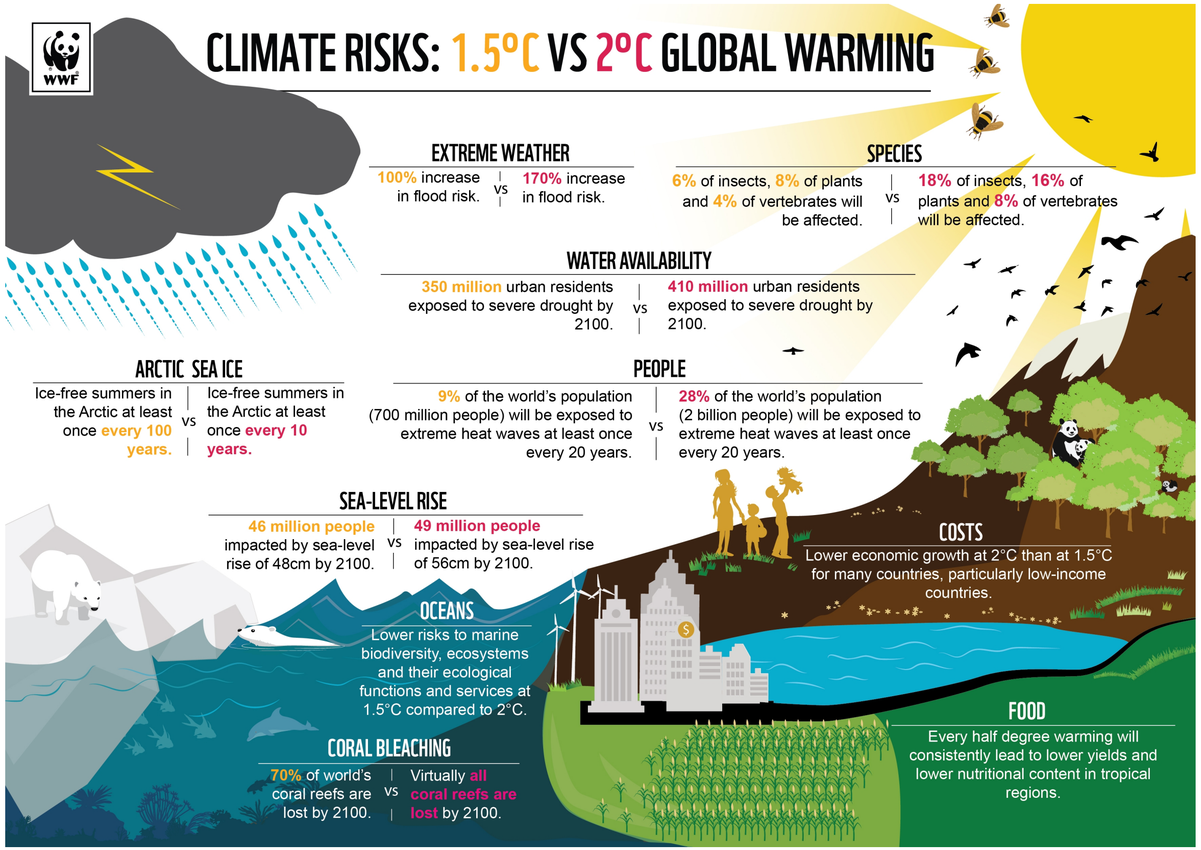
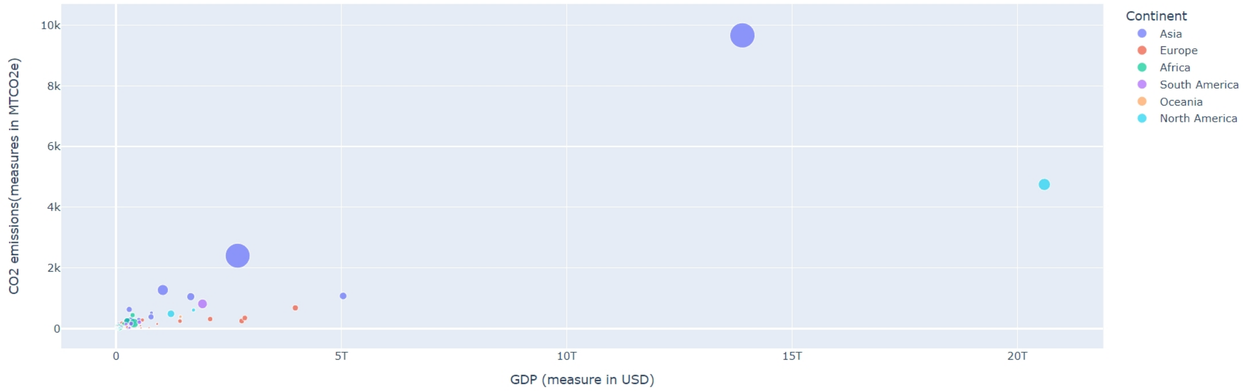


Image source: https://www.wwf.org.uk/updates/our-warming-world-how-much-difference-will-half-degree-really-make

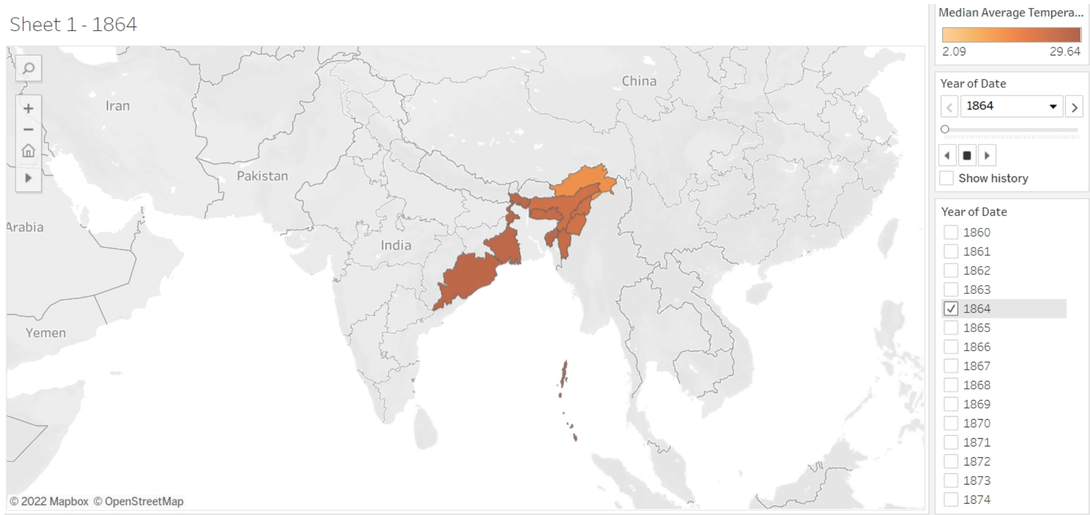
This visualization gives a broad picture of the 1.5°C VS 2°C change in temperature and how it can affect the different aspects of life on the earth. It is a very informative visualization wrapped in a single picture. And at the same time, it also conveys how only a .5°C temperature change can seriously hurt us. This is probably the best visualization available on the internet for climate change and its risks.

5.Failed Visualization



This is the bubble chart that I tried to plot but failed as I didn't give the x and y axis as log scale as the values were large. We should provide it with the log scale. Also, the size of the bubble matters, and I forgot to set the size. But in the end, you see that fig 14 (Page 16) is the correct version of this.

When I was trying to create a time series choropleth map for India in Tableau, I got to know some years in 1850-1890 did not have the values for the average temperature for some specific states. So I applied the filter from 1950 - 2013. It might be because we dropped the NA values while cleaning the dataset.



6. Conclusion

Climate change is impacting the planet significantly, including more frequent and severe heatwaves, droughts, and storms, as well as rising sea levels. These changes affect ecosystems and human communities worldwide and are likely to continue unless we reduce our greenhouse gas emissions.

One of the main ways to address climate change is to transition to renewable energy sources, such as wind and solar power, and to reduce our overall energy consumption. This can be achieved through individual actions, such as using energy-efficient appliances and reducing our use of single-use plastics, and through collective action, such as implementing policies that support renewable energy and encourage sustainable development.

Last but not least, as discussed above, data visualization is all about telling the story, and with attractive visualization, we have successfully delivered our point of view and the importance of climate change

7. References

1. https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/land/12/1/1900-2022?trend=true&trend_base=10&begtrendyear=1880&endtrendyear=2022

2. https://data.gov.in/catalog/all-india-seasonal-and-annual-minmax-temperature-series

3. https://www.climatewatchdata.org/ghg-emissions?breakBy=gas&end_year=2019&gases=co2%2Cn2o%2Cch4&regions=IND&start_year=1990

4. https://climate.nasa.gov/news/2865/a-degree-of-concern-why-global-temperatures-matter/

5. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=IN

6. https://www.epa.gov/climate-indicators/greenhouse-gases#:~:text=An%20increase%20in%20the%20atmospheric,atmosphere%20increased%20by%2045%20percent.

7. https://plotly.com/python/

8. https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data?resource=download

9. https://help.tableau.com/current/pro/desktop/en-us/maps_build.htm

10. https://www.un.org/en/climatechange/science/causes-effects-climate-change