Project Documentation On Unearthing the Environmental Impact of Human Activity: A Global CO2 Emission Analysis

Full Name : Karnasula Sai Sri Pravallika

Roll no : 20A31A4213

Branch : Computer Science Engineering – Artificial Intelligence & Machine Learning

College : Pragati Engineering College, Surampalem

Register no : SBAP0018682

Problem statement

In the face of escalating global warming, attributed largely to increased atmospheric carbon dioxide levels from human activities, understanding the dynamics of CO2 emissions becomes paramount. This project aims to analyze and visualize global CO2 emissions from 1975 to 2020, investigating both country-specific and regional contributions. By scrutinizing the trends and patterns in CO2 emissions data, we seek to identify key factors influencing emission levels and their impact on the environment. The ultimate goal is to provide insights that can inform policy-making and facilitate the establishment of effective strategies for mitigating CO2 emissions, thereby combating global warming on a collective scale.

In addition to analyzing historical trends in CO2 emissions, this project seeks to forecast future emission trajectories based on identified patterns and factors. By leveraging predictive modeling techniques, we aim to anticipate potential shifts in CO2 emission levels and assess the efficacy of proposed mitigation measures. Furthermore, the analysis will explore the disparities in CO2 emissions among nations and regions, investigating the socio-economic, industrial, and policy-related determinants underlying these variations. The insights generated from this comprehensive examination will not only aid in setting realistic emission reduction targets for individual countries but also inform international cooperation efforts aimed at addressing the global challenge of climate change.

Business Requirements

1. **Data Acquisition:** Obtain comprehensive and reliable datasets spanning CO2 emissions across countries and regions from 1975 to 2020. Ensure data integrity and accuracy to facilitate robust analysis and decision-making.

2. **Analysis and Visualization Tools:** Implement advanced analytics and visualization tools capable of handling large-scale CO2 emissions data. The tools should enable thorough exploration of trends, patterns, and correlations in emissions data, facilitating insightful visual representations for stakeholders.

3. **Forecasting Capabilities:** Develop predictive modelling capabilities to forecast future CO2 emission trajectories based on historical data trends and identified influencing factors. The forecasting module should provide probabilistic estimates of future emission levels under different scenarios.

4. **Factor Identification:** Conduct in-depth analysis to identify key factors influencing CO2 emissions, including but not limited to industrial activities, energy consumption patterns, population dynamics, and policy interventions. Prioritize factors based on their significance and potential impact on emission levels.

5. **Disparity Analysis:** Explore disparities in CO2 emissions among countries and regions, considering socio-economic, geopolitical, and environmental factors. Identify hotspots of emissions and assess the underlying drivers contributing to disparities.

6. **Policy Recommendations:** Generate actionable insights and policy recommendations for governments, policymakers, and environmental agencies to formulate effective strategies for mitigating CO2 emissions. Emphasize the importance of setting realistic emission reduction targets and fostering international cooperation to address climate change collectively.

7. **Stakeholder Engagement:** Engage with relevant stakeholders, including government bodies, environmental organizations, industry representatives, and research institutions, to gather input, validate findings, and ensure alignment with stakeholders' priorities and objectives.

8. **Accessibility and User Interface:** Develop user-friendly interfaces and dashboards to facilitate easy access to CO2 emissions data, analysis results, and visualization outputs. Ensure compatibility with various devices and browsers to maximize accessibility for stakeholders.

9. **Compliance and Ethics:** Adhere to data privacy regulations and ethical standards throughout the data acquisition, analysis, and reporting processes. Ensure transparency and accountability in handling sensitive environmental data to maintain stakeholders' trust and confidence.

10. **Continuous Improvement:** Establish mechanisms for ongoing monitoring, evaluation, and refinement of the CO2 emissions analysis framework.

Literature Survey

1. **"Trends and Drivers of Global Carbon Dioxide Emissions" by Le Quéré et al. (2018)**
   This study provides a comprehensive analysis of the trends and drivers of global CO2 emissions, examining the contributions of various sectors, including energy, industry, transportation, and land use. It explores the implications of these trends for climate change mitigation efforts and highlights the importance of addressing key drivers such as population growth, economic development, and technological innovation.

2. **"Spatial and Temporal Patterns of Global CO2 Emissions: Integrating Data from Multiple Sources" by Boden et al. (2017)**
   This research paper presents spatial and temporal patterns of global CO2 emissions based on integrated data from multiple sources, including national inventories, energy statistics, and satellite observations. It discusses the challenges and opportunities in harmonizing heterogeneous datasets for comprehensive emissions analysis and emphasizes the need for improved data sharing and collaboration among stakeholders.

3. **"Determinants of CO2 Emissions Growth in Emerging Economies: A Comprehensive Review" by Khan et al. (2020)**
   This review article examines the determinants of CO2 emissions growth in emerging economies, synthesizing findings from empirical studies across various disciplines, including economics, environmental science, and policy analysis. It identifies factors such

as economic growth, industrialization, energy consumption patterns, and environmental policies as key drivers of emissions growth, highlighting the complex interactions among these factors.

4. **"Forecasting Global CO2 Emissions: A Review of Modeling Approaches and Challenges" by Wang et al. (2019)**
   This review paper evaluates existing modeling approaches for forecasting global CO2 emissions, comparing their strengths, limitations, and uncertainties. It discusses the role of socio-economic factors, technological advancements, and policy interventions in shaping future emission trajectories and identifies challenges such as data availability, model complexity, and scenario uncertainty that need to be addressed for more accurate and reliable forecasts.

5. **"Regional Disparities in CO2 Emissions: Causes, Consequences, and Policy Implications" by Li et al. (2021)**
   This research article investigates regional disparities in CO2 emissions, analyzing the underlying causes, consequences, and policy implications. It examines factors such as economic development, energy infrastructure, demographic trends, and environmental regulations that contribute to disparities in emissions levels among regions. The study also discusses strategies for addressing these disparities through targeted policy interventions and international cooperation.

6. **"Assessing the Effectiveness of Climate Change Mitigation Policies: Lessons from Global CO2 Emissions Reduction Initiatives" by Fuso Nerini et al. (2018)**
   This study assesses the effectiveness of climate change mitigation policies and initiatives aimed at reducing global CO2 emissions. It reviews case studies of diverse policy instruments, including carbon pricing mechanisms, renewable energy incentives, and emission trading schemes, to evaluate their impact on emissions reduction, economic competitiveness, and social equity. The findings provide valuable insights for designing and implementing effective climate policies at the national and international levels.

Social or Business Impact.

1. **Climate Change Mitigation:** By understanding the drivers and patterns of CO2 emissions, stakeholders can develop and implement effective mitigation strategies to combat climate change. Reducing CO2 emissions helps to mitigate the adverse impacts of climate change, such as extreme weather events, sea-level rise, and disruptions to ecosystems, agriculture, and human settlements.

2. **Public Health Improvement:** Decreasing CO2 emissions can lead to improvements in public health by reducing air pollution levels. Many of the sources of CO2 emissions, such as industrial facilities and vehicles, also emit pollutants harmful to human health, such as particulate matter and nitrogen oxides. By curbing CO2 emissions,

communities can experience cleaner air and lower rates of respiratory illnesses and cardiovascular diseases.

3. **Energy Transition and Innovation:** Addressing CO2 emissions necessitates a transition towards cleaner and more sustainable energy sources, such as renewable energy and energy efficiency measures. This transition drives innovation and investment in clean energy technologies, creating new business opportunities and jobs in the renewable energy sector while reducing dependence on fossil fuels.

4. **Economic Resilience and Sustainability:** Mitigating CO2 emissions promotes economic resilience and sustainability by reducing the vulnerability of economies to the impacts of climate change, such as crop failures, water scarcity, and infrastructure damage. Investing in low-carbon technologies and infrastructure enhances the long-term competitiveness and resilience of businesses and economies in a rapidly changing climate.

5. **International Cooperation and Diplomacy:** Addressing global CO2 emissions requires international cooperation and diplomacy to develop and implement coordinated climate policies and agreements. Collaborative efforts to reduce emissions foster trust and cooperation among nations, strengthen diplomatic ties, and promote shared responsibility for addressing the global challenge of climate change.

6. **Social Equity and Justice:** Efforts to mitigate CO2 emissions should prioritize considerations of social equity and justice to ensure that the burdens and benefits of emission reduction measures are distributed

fairly among different communities and socioeconomic groups. This includes ensuring access to clean energy, addressing energy poverty, and promoting inclusive decision-making processes that engage marginalized communities.

7. **Responsible Corporate Citizenship:** Businesses play a crucial role in reducing CO2 emissions through responsible corporate citizenship and sustainability initiatives. By adopting environmentally friendly practices, investing in clean technologies, and reducing their carbon footprint, companies can enhance their reputation, attract environmentally conscious consumers and investors, and contribute to a more sustainable future for society and the planet.

Collect the dataset

country	year	co2	co2_growth_prct	co2_per_capita	cumulative_co2	coal_co2	cement_co2	flaring_co2	gas_co2	oil_co2	other_industry_co2	cement_co2_per_ca	coal_co2_per_capit	flaring_co2_
Afghanistan	1975	2.121	10.88	0.167	21.287	0.399	0.069	0.304	0.476	0.874	0	0.006	0.031	
Afghanistan	1976	1.981	-6.62	0.153	23.267	0.425	0.079	0.293	0.3	0.883	0	0.006	0.033	
Afghanistan	1977	2.364	20.36	0.181	25.652	0.451	0.065	0.381	0.513	0.975	0	0.005	0.034	
Afghanistan	1978	2.153	-9.68	0.161	27.805	0.576	0.058	0.283	0.301	0.936	0	0.004	0.043	
Afghanistan	1979	2.233	3.69	0.166	30.038	0.352	0.064	0.267	0.385	1.165	0	0.005	0.026	
Afghanistan	1980	1.756	-21.34	0.132	31.794	0.316	0.023	0.305	0.187	0.925	0	0.002	0.024	
Afghanistan	1981	1.978	12.65	0.15	33.772	0.333	0.033	0.293	0.304	1.015	0	0.002	0.025	
Afghanistan	1982	2.095	5.87	0.163	35.867	0.385	0.039	0.282	0.396	0.993	0	0.003	0.03	
Afghanistan	1983	2.52	20.31	0.201	38.387	0.385	0.006	0.293	0.616	1.22	0	0	0.031	
Afghanistan	1984	2.822	11.97	0.231	41.209	0.393	0.048	0.316	0.932	1.134	0	0.004	0.032	
Afghanistan	1985	3.501	24.1	0.293	44.71	0.4	0.032	0.33	1.192	1.548	0	0.003	0.034	
Afghanistan	1986	3.134	-10.5	0.267	47.844	0.425	0.038	0.33	1.202	1.14	0	0.003	0.036	
Afghanistan	1987	3.114	-0.63	0.268	50.957	0.443	0.043	0.223	0.392	2.013	0	0.004	0.038	
Afghanistan	1988	2.857	-8.25	0.246	53.814	0.366	0.043	0.187	0.44	1.821	0	0.004	0.032	
Afghanistan	1989	2.765	-3.22	0.233	56.579	0.337	0.043	0.04	0.48	1.865	0	0.004	0.028	
Afghanistan	1990	2.603	-5.85	0.21	59.182	0.278	0.046	0.026	0.403	1.85	0	0.004	0.022	
Afghanistan	1991	2.427	-6.76	0.182	61.61	0.249	0.046	0.026	0.388	1.718	0	0.003	0.019	
Afghanistan	1992	1.379	-43.17	0.095	62.989	0.022	0.046	0.022	0.363	0.927	0	0.003	0.002	
Afghanistan	1993	1.333	-3.36	0.084	64.322	0.018	0.047	0.022	0.352	0.894	0	0.003	0.001	
Afghanistan	1994	1.282	-3.86	0.075	65.604	0.015	0.047	0.022	0.338	0.86	0	0.003	0.001	
Afghanistan	1995	1.23	-3.99	0.068	66.834	0.015	0.047	0.022	0.322	0.824	0	0.003	0.001	
Afghanistan	1996	1.165	-5.33	0.062	67.999	0.007	0.047	0.022	0.308	0.78	0	0.002	0	
Afghanistan	1997	1.084	-6.94	0.056	69.083	0.004	0.047	0.022	0.283	0.728	0	0.002	0	
Afghanistan	1998	1.029	-5.07	0.052	70.113	0.004	0.047	0.022	0.265	0.691	0	0.002	0	
Afghanistan	1999	0.81	-21.34	0.04	70.922	0.004	0.047	0.022	0.242	0.495	0	0.002	0	
Afghanistan	2000	0.758	-6.4	0.036	71.68	0.004	0.01	0.022	0.224	0.498	0	0	0	
Afghanistan	2001	0.798	5.32	0.037	72.478	0.07	0.007	0.022	0.209	0.491	0	0	0.003	
Afghanistan	2002	1.052	31.79	0.046	73.529	0.076	0.011	0	0.546	0.44	0	0	0.002	

A snapshot of the dataset

Understanding the dataset:

The dataset consists of CO2 emissions data spanning from 1975 to 2020, encompassing 20 columns and 10,920 rows. Each row represents a unique record containing information on CO2 emissions for a specific country and year. The dataset has been pre-processed using excel to ensure data integrity and accuracy.

Columns in the dataset include:

Country: The country for which CO2 emissions data is recorded.

Year: The year in which the CO2 emissions data was recorded.

CO2 Emission (In Million Metric Tons): The total CO2 emissions in million metric tons for the corresponding country and year.

CO2 Growth per Capita: The rate of growth of CO2 emissions per capita.

CO2 Per Capita: The CO2 emissions per capita, typically measured in metric tons per person.

Cumulative CO2: The cumulative CO2 emissions for the country up to the recorded year.

Several Fossil Fuels rate of Emission: This column likely contains rates of emission for various fossil fuels such as coal, oil, and natural gas.

The dataset provides a comprehensive overview of CO2 emissions trends over time, allowing for detailed analysis and exploration of factors influencing emissions levels across different countries and regions. Researchers and analysts can utilize this dataset to investigate patterns, trends, and correlations in CO2 emissions data, as well as to assess the effectiveness of mitigation strategies and policy interventions aimed at reducing carbon emissions globally.

Data Preparation

Data Cleaning: Check for missing values, duplicates, and inconsistencies in the dataset. Handle missing values through imputation or removal, depending on the extent of missing data and its impact on the analysis. Eliminate duplicates and ensure data consistency across all columns.

Feature Selection: Identify relevant features for analysis based on the project objectives and research questions. Consider factors such as CO2 emissions, year, country, and per capita emissions for inclusion in the analysis. Exclude irrelevant or redundant features that do not contribute to the analysis.

Data Transformation: Convert categorical variables, such as country names, into numerical representations using techniques like one-hot encoding or label encoding. Normalize numerical variables to ensure uniform scales across different features, especially if using algorithms sensitive to feature scales, such as clustering or regression models.

Feature Engineering: Create new features or derived variables that capture additional insights from the data. For example, calculate the cumulative CO2 emissions for each country-year combination by aggregating emissions over previous years. Compute growth rates or trends in emissions to capture temporal patterns.

Handling Outliers: Identify and handle outliers in the dataset that may skew the analysis results. Use statistical methods or domain knowledge to detect outliers and apply appropriate techniques such as trimming, winsorization, or transformation to mitigate their effects on the analysis.

Data Integration: Integrate additional datasets containing relevant information, such as population demographics, economic indicators, energy consumption data, or environmental policies. Merge these datasets with the CO2 emissions dataset using common identifiers such as country names or ISO codes to enrich the analysis.

Temporal Alignment: Ensure temporal alignment of data across different variables and datasets. Verify that the time intervals and granularity of the data are consistent to facilitate meaningful comparisons and analysis over time.

Data Aggregation: Aggregate the data at different levels of granularity, such as country-level, regional-level, or global-level, depending on the scope of the analysis. Compute summary statistics or aggregate measures such as averages, sums, or percentiles to condense the data for easier interpretation and analysis.

Data Splitting: Split the dataset into training, validation, and test sets if performing predictive modeling or machine learning tasks. Reserve a portion of the data for model evaluation and testing to assess the generalization performance of the models on unseen data.

Data Visualization: Visualize the pre-processed data using descriptive statistics, histograms, box plots, time series plots, and heatmaps to explore patterns, trends, and relationships in the data. Visualizations help in identifying outliers, understanding data distributions, and gaining insights into the underlying structure of the data.

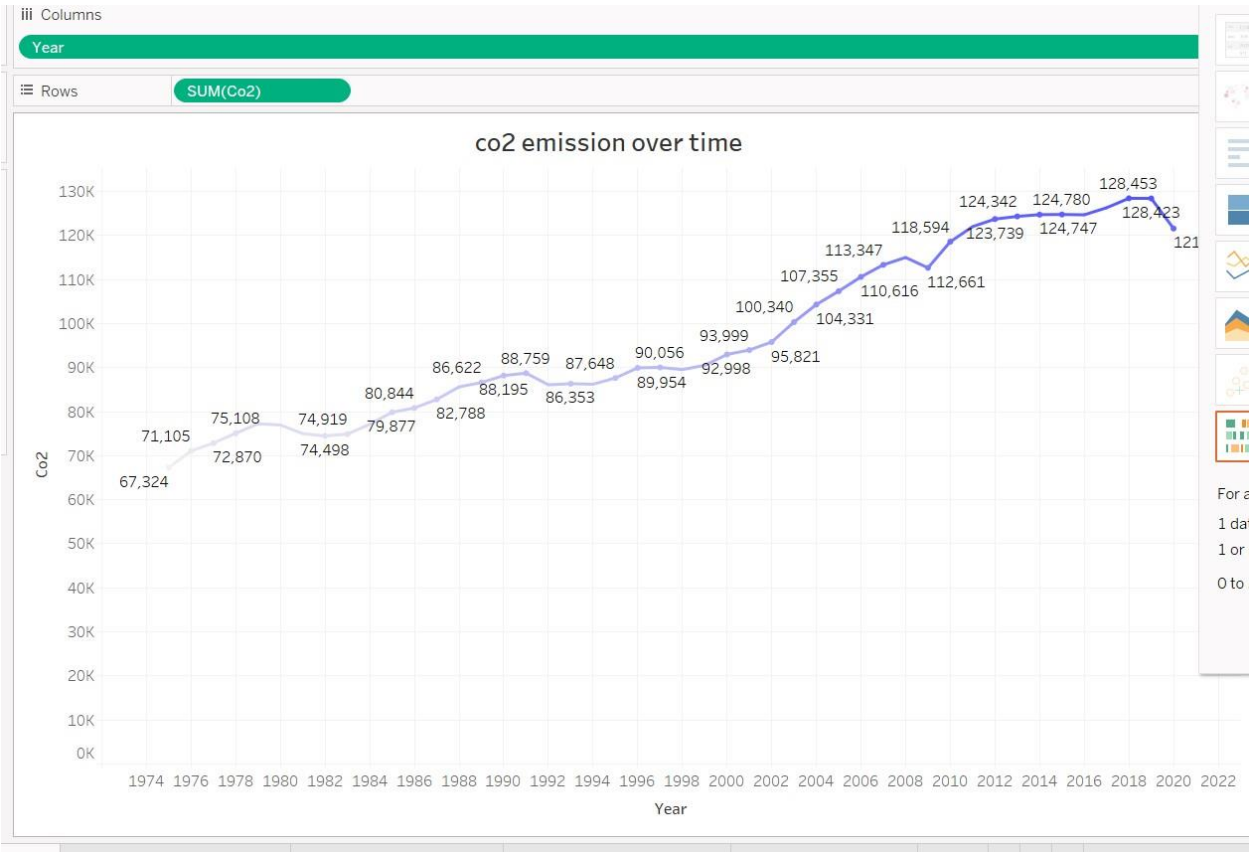
Data Visualization

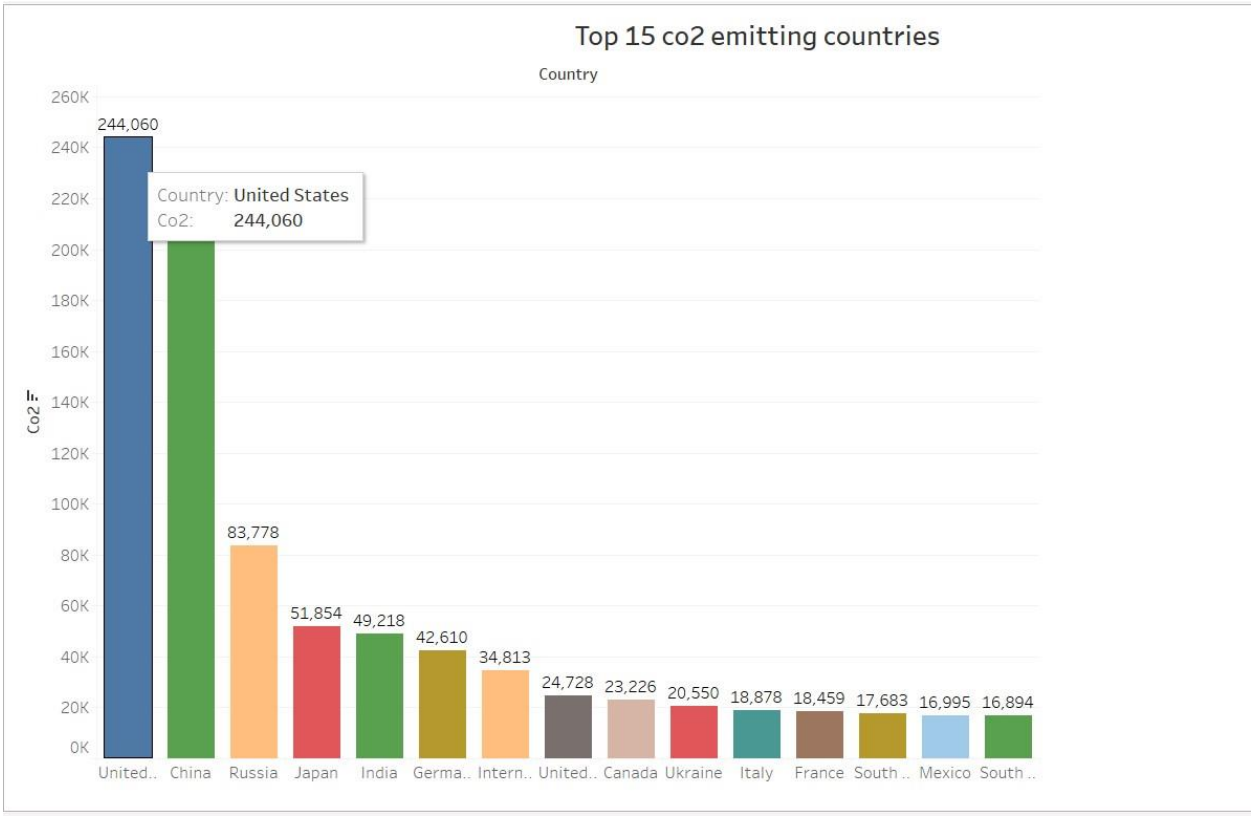
Data visualization is the process of creating graphical representations of data in order to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyze the Co2 Emission include bar charts, line charts, Tree Map, scatter plots, pie charts, Maps etc. These visualizations can be used to compare performance, track changes over time, show Emission, and relationships between variables, breakdown of factors and emission by countries and continents.

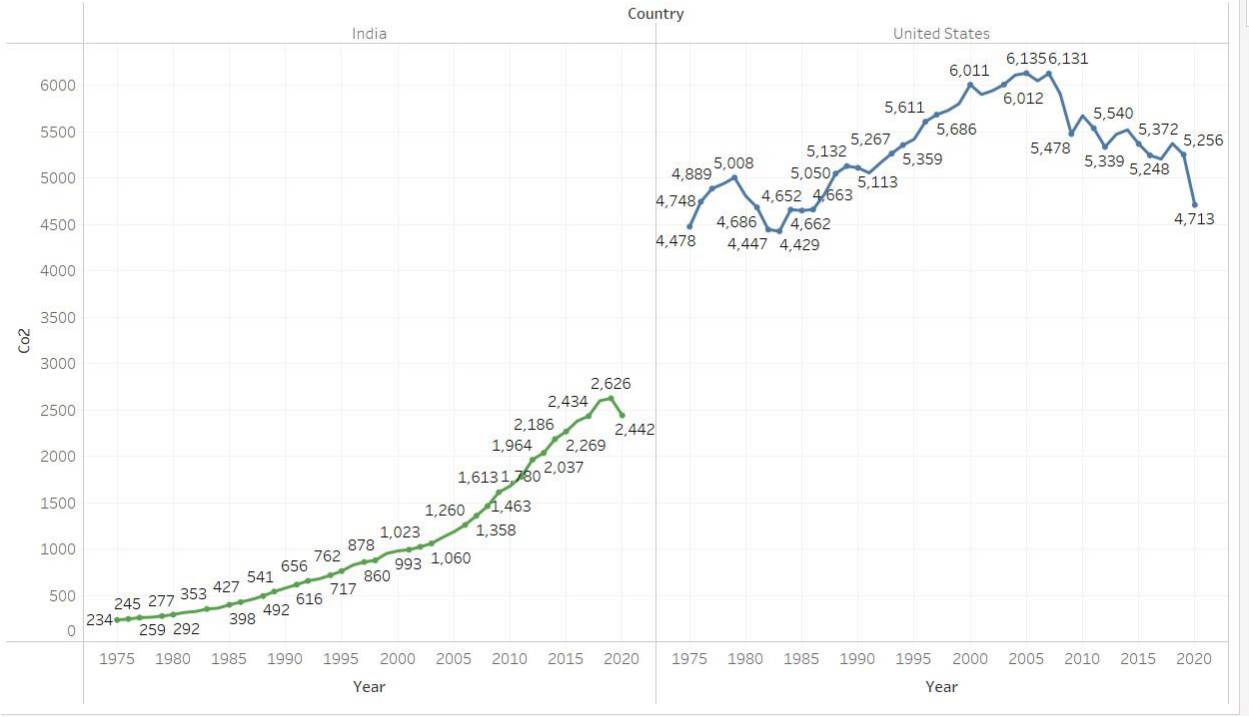
My Tableau Data Visualizations

Total CO2 emissions by each country

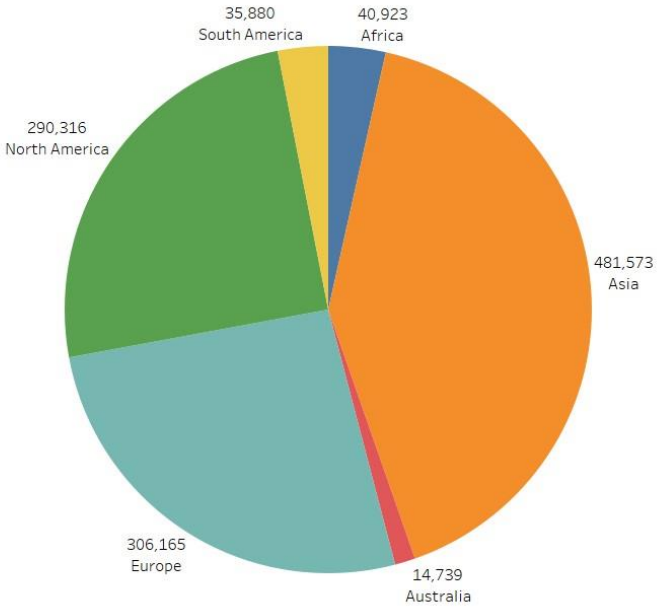

Columns
Year
Rows
SUM(Co2)
co2 emission over time
Co2
Year
130K
120K
110K
100K
90K
80K
70K
60K
50K
40K
30K
20K
10K
0K
1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018 2020 2022
67,324
71,105
72,870
75,108
74,498
74,919
79,877
80,844
82,788
86,622
88,195
88,759
86,353
87,648
89,954
90,056
92,998
93,999
95,821
100,340
104,331
107,355
110,616
113,347
112,661
118,594
123,739
124,342
124,747
124,780
128,423
128,453
121

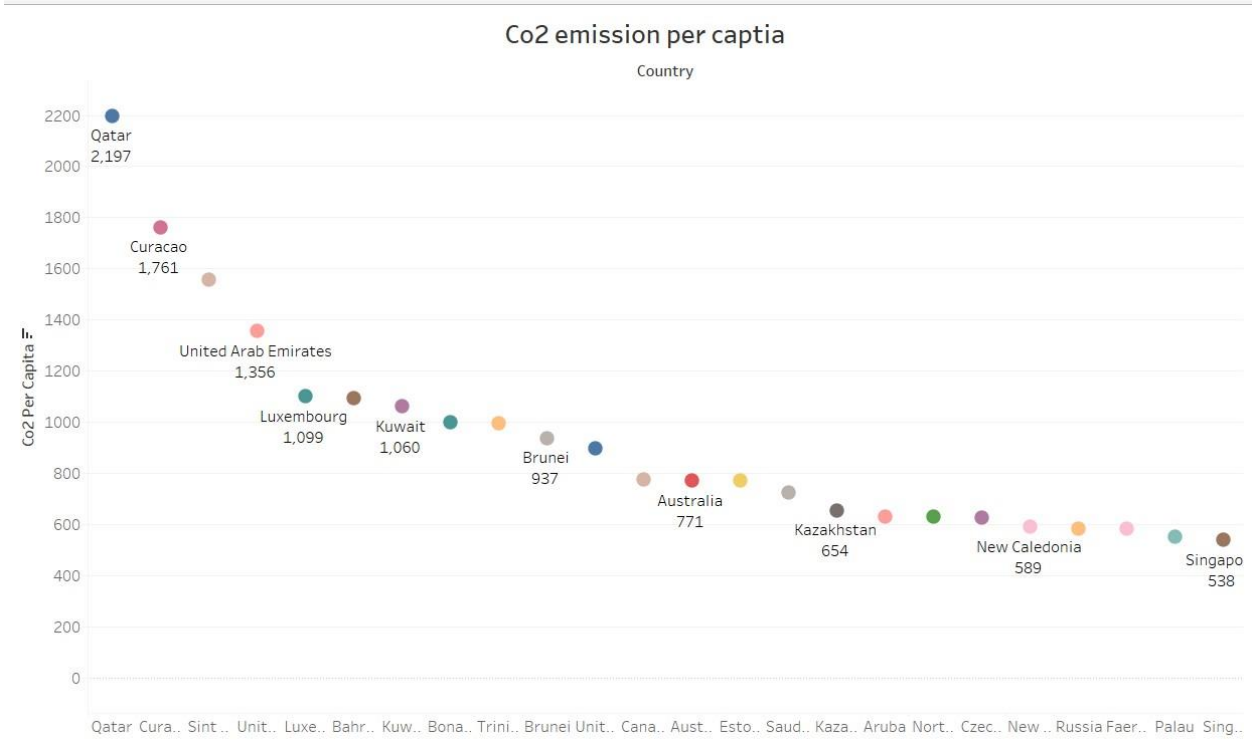

Top 15 co2 emitting countries
Country
Co2
260K
240K
220K
200K
180K
160K
140K
120K
100K
80K
60K
40K
20K
0K
Country: United States
Co2: 244,060
244,060
83,778
51,854
49,218
42,610
34,813
24,728
23,226
20,550
18,878
18,459
17,683
16,995
16,894
United..
China
Russia
Japan
India
Germa..
Intern..
United..
Canada
Ukraine
Italy
France
South ..
Mexico
South ..

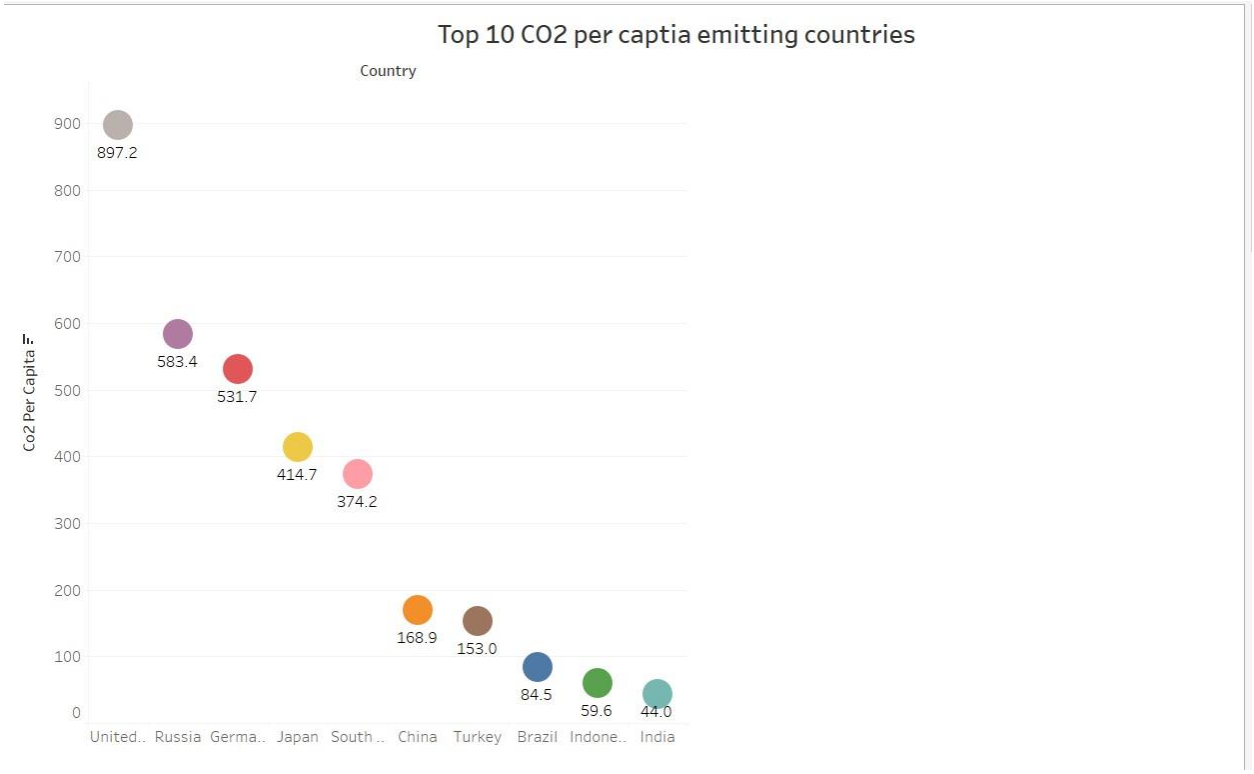
Co2 emissions India vs USA



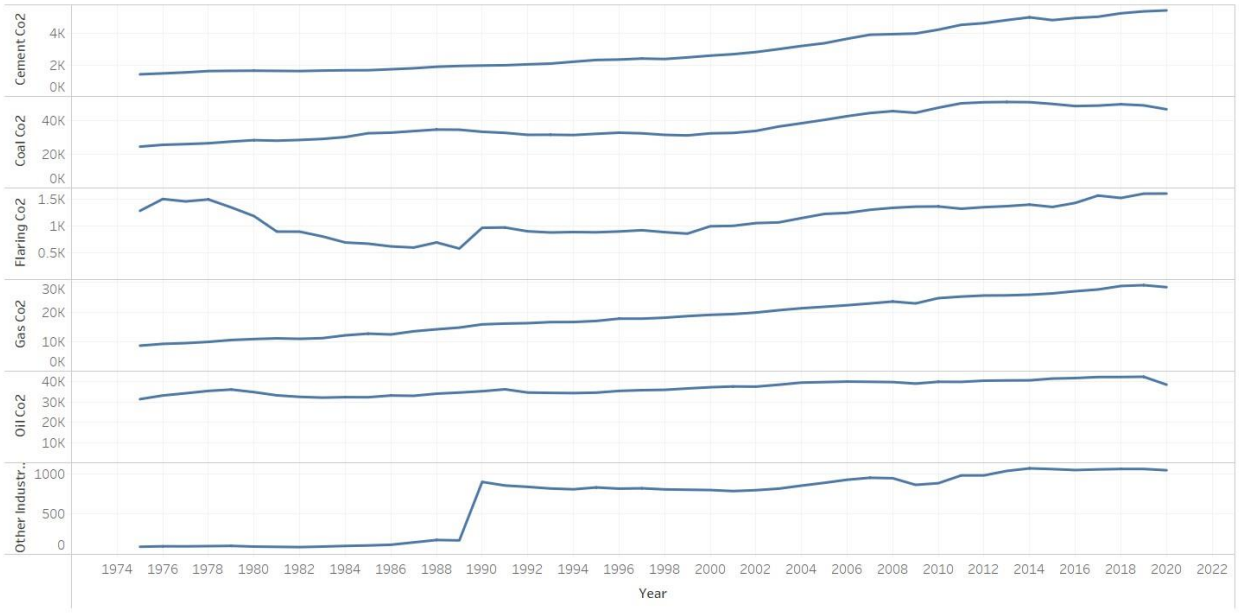
Total emissions by continents



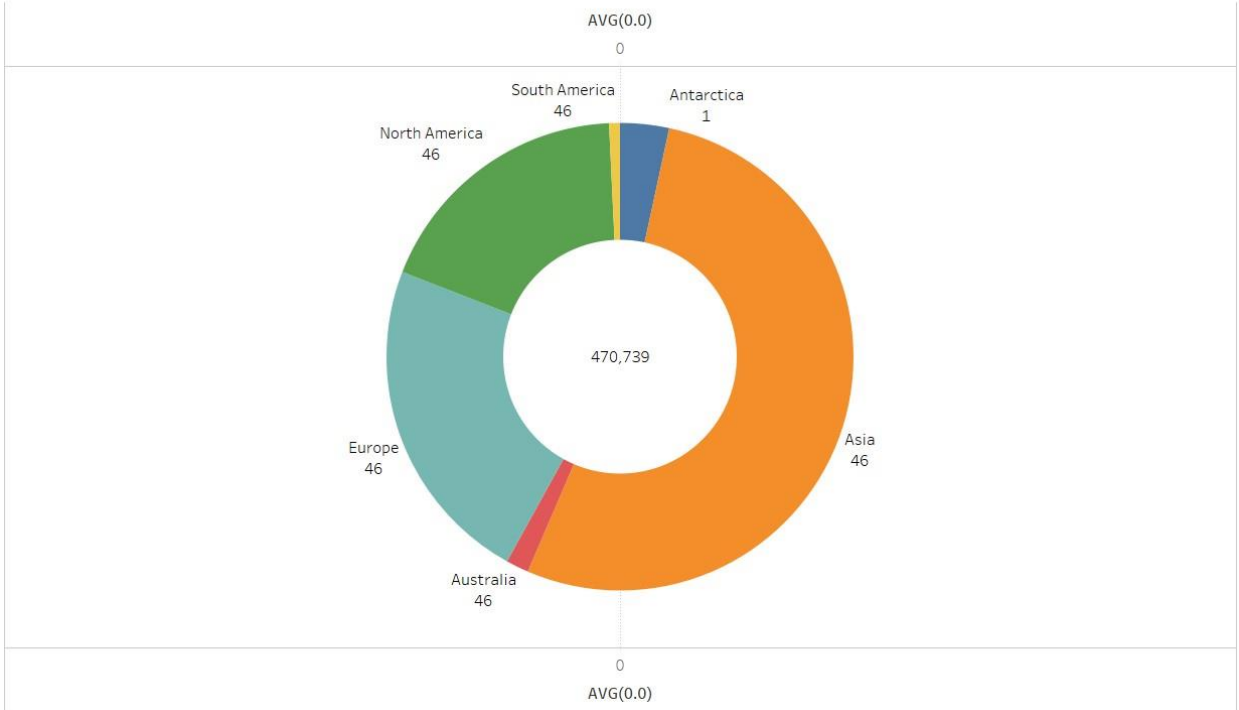

Co2 emission per captia
Country
Co2 Per Capita
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Qatar
2,197
Curacao
1,761
United Arab Emirates
1,356
Luxembourg
1,099
Kuwait
1,060
Brunei
937
Australia
771
Kazakhstan
654
New Caledonia
589
Singapo
538
Qatar Cura.. Sint .. Unit.. Luxe.. Bahr.. Kuw.. Bona.. Trini.. Brunei Unit.. Cana.. Aust.. Esto.. Saud.. Kaza.. Aruba Nort.. Czec.. New .. Russia Faer.. Palau Sing..

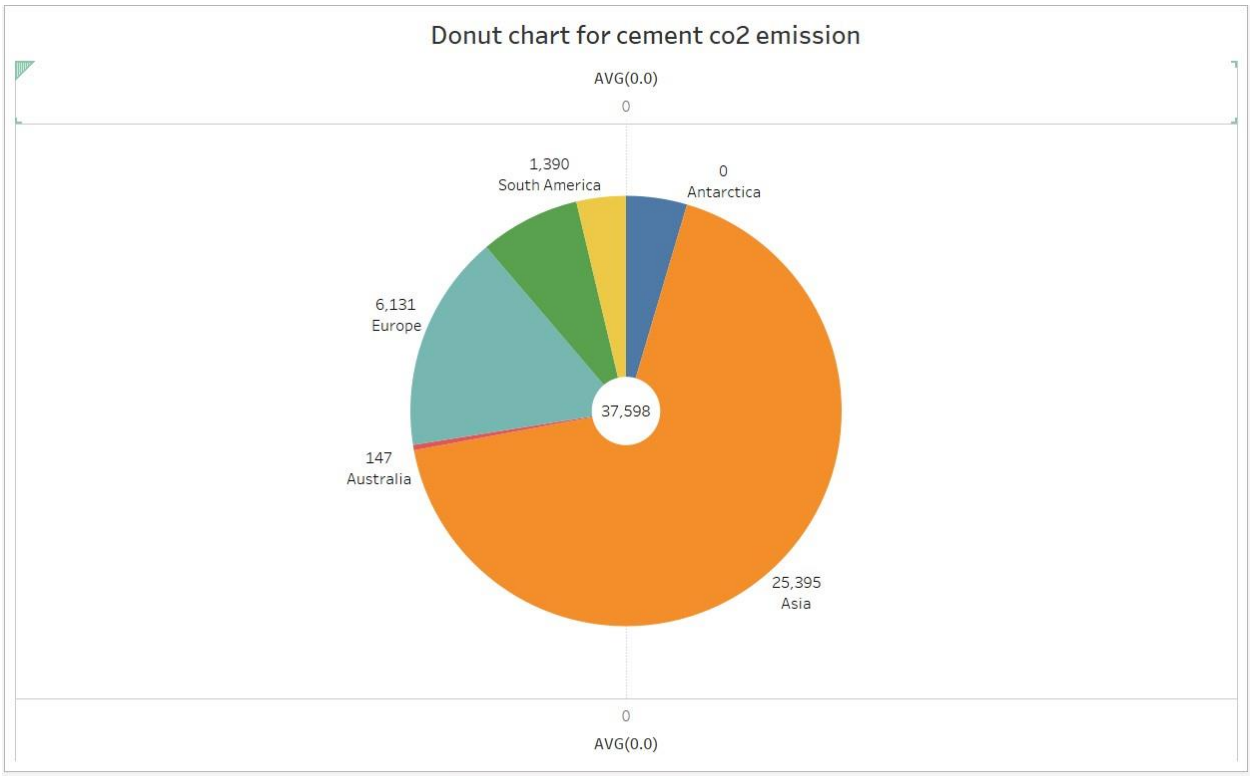

Top 10 CO2 per captia emitting countries
Country
Co2 Per Capita
900
800
700
600
500
400
300
200
100
0
897.2
583.4
531.7
414.7
374.2
168.9
153.0
84.5
59.6
44.0
United.. Russia Germa.. Japan South .. China Turkey Brazil Indone.. India

Emission rate by internal factors

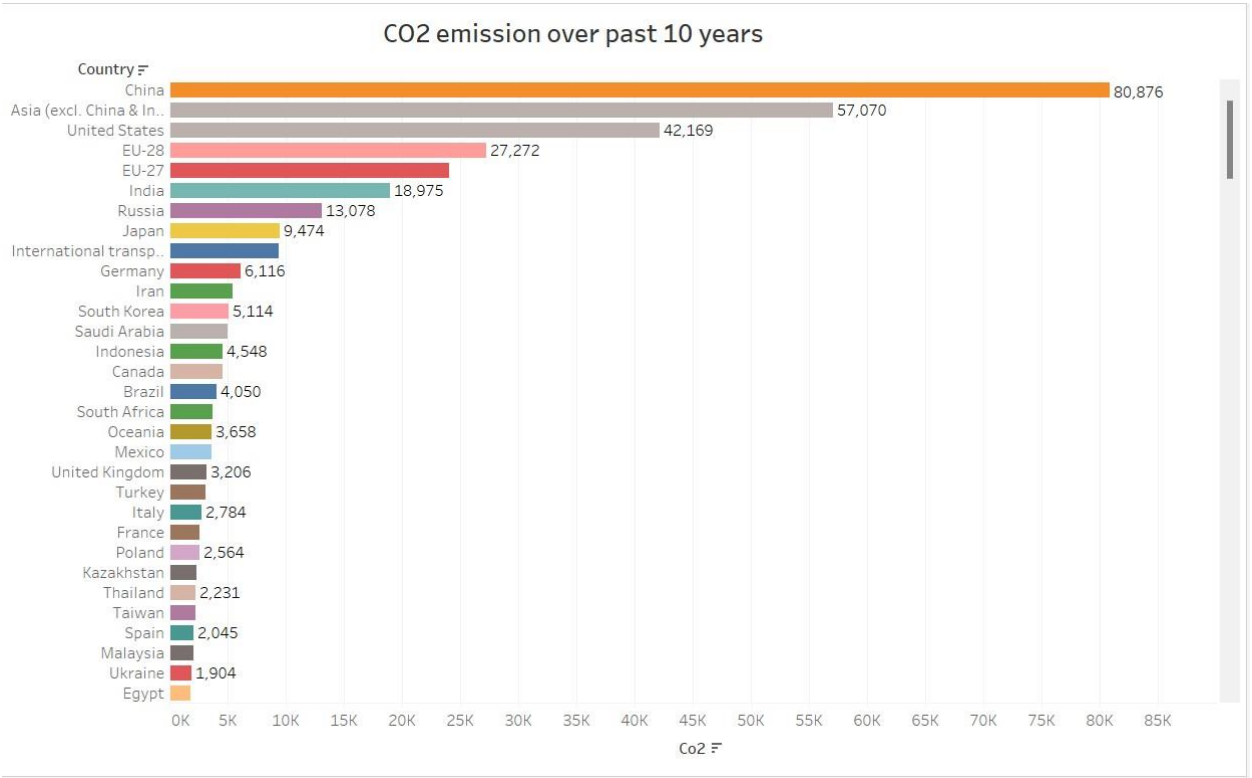


Donut chart for coal co2 emission

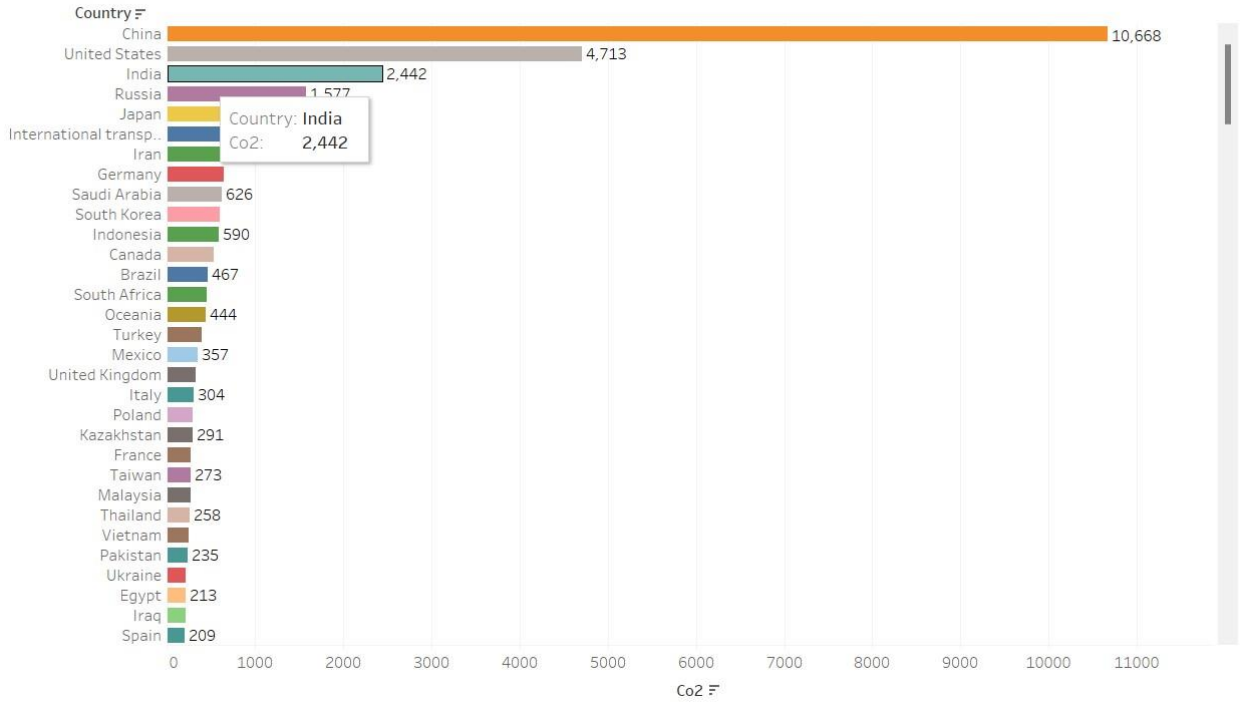




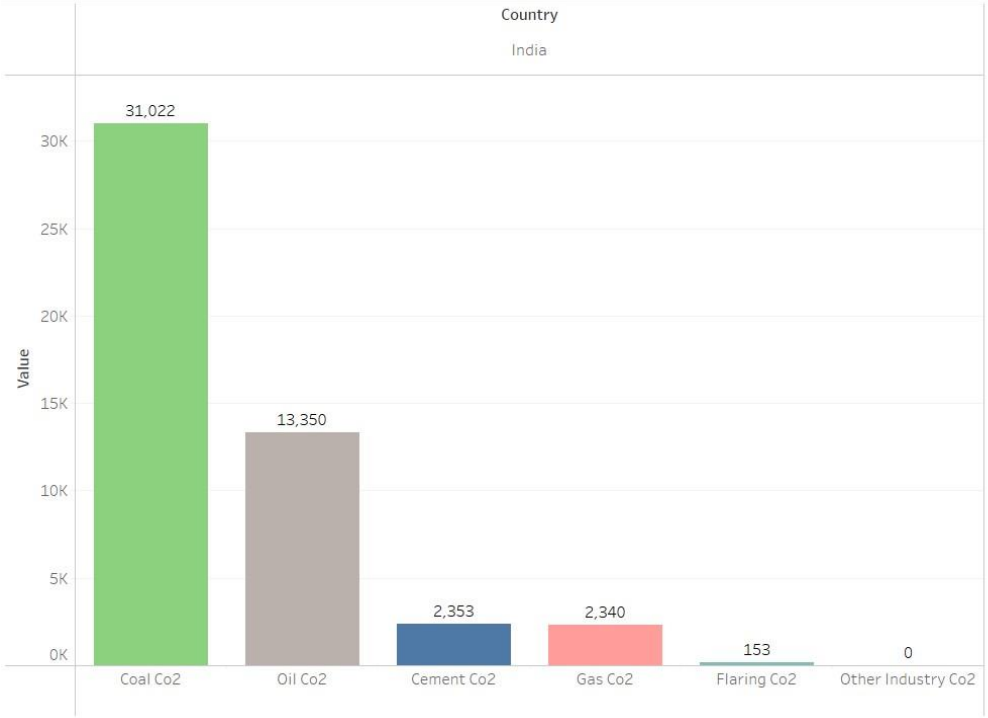
on rate by internal factors | Donut chart for coal co2 emiss... | **Donut chart for cement co2 e...** | CO2 emission over past 10 years | CO2 emission in 2020 | Overall contributic



CO2 emission in 2020



Overall contribution of india in co2 emission



Project Demonstration & Documentation

References:

The tableau workbooks are hosted in the following github link.

Github link: https://github.com/pravallika1309

Testimonial video link:
https://drive.google.com/file/d/1gEpVYagLji5p9aa6_0aLqudngxCh25BG/view?usp=drivesdk