

## Data Analytics with Tableau

By SmartIntenz

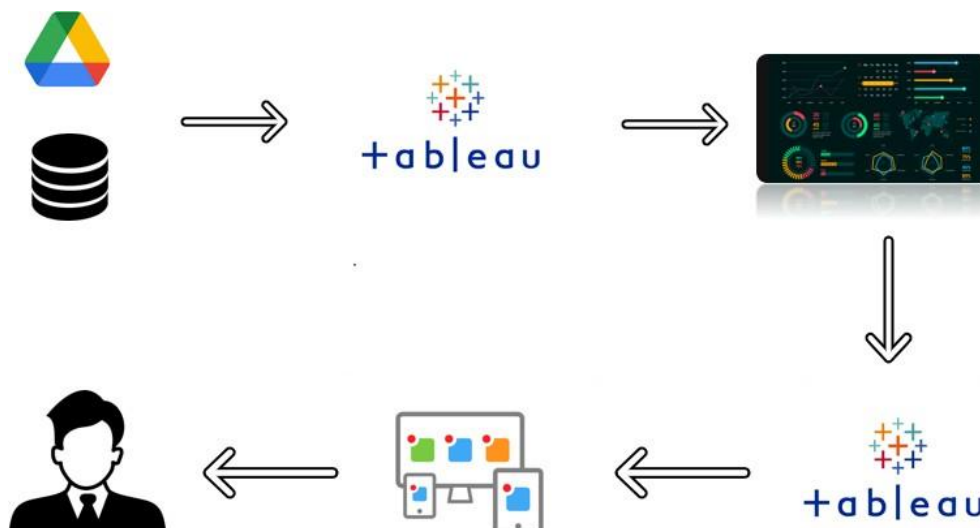
**GARUGU LOKESH DEEP (20A31A4433)**

### **Unearthing the Environmental Impact of Human Activity:A Global CO2 Emission Analysis**

Global warming is one of the biggest challenges currently being faced by the human race, although correlation is not causation, a likely cause of global warming is due to increased atmospheric carbon dioxide from human activities. **CO2 Emission** refers to the Carbon Dioxide emitted throughout the world. For this analysis we will be focusing on CO2 Emissions and its effect on the world we live in as well as some key factors and stats that may play a role in the emission of CO2 globally. Fossil fuel use is the primary source of CO2. The data throws light onto how much fossil fuels are burnt, per year per nation, which amounts to an increase in CO2 every year. This will help researchers and environment experts to predict global warming. So countries should set a goal to decrease this amount yearly.

Analyzing Global Co2 Emission across countries from 1975 to 2020. This dataset contains a record of Co2 Emission by each Country and Region of Earth, here we are going to analyze and visualize Country wise, Region wise and Overall Co2 Emission on Earth.

#### **Technical Architecture:**



## Project Flow

To accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding
  - o Specify the business problem
  - o Business requirements
  - o Literature Survey
  - o Social or Business Impact.
- Data Collection & Extraction from Database
  - o Collect the dataset,
  - o Storing Data in DB
  - o Perform SQL Operations
  - o Connect DB with Tableau
- Data Preparation
  - o Prepare the Data for Visualization
- Data Visualizations
  - o No of Unique Visualizations
- Dashboard
  - o Responsive and Design of Dashboard
- Story
  - o No of Scenes of Story
- Performance Testing
  - o Amount of Data Rendered to DB ‘
  - o Utilization of Data Filters
  - o No of Calculation Fields
  - o No of Visualizations/ Graphs
- Web Integration
  - o Dashboard and Story embed with UI With Flask
- Project Demonstration & Documentation
  - o Record explanation Video for project end to end solution
  - o Project Documentation-Step by step project development procedure

## **Milestone 1: Define Problem / Problem Understanding**

### **Activity 1: Specify the business problem**

Refer Project Description

### **Activity 2: Business requirements**

The business requirements for analyzing the Co2 Emission Globally over time, identifying affecting factors, creating interactive dashboards and reports, identifying areas for improvement, making data-driven decisions, comparing to countries average and creating forecasting models for future performance. The ultimate goal is to gain insights and reduce the emission through data visualization techniques.

### **Activity 3: Literature Survey (Student Will Write)**

A literature survey is a method of researching existing literature and studies related to a specific topic. In the context of analyzing the Global Co2 Emission, a literature survey would involve reviewing studies and articles that have been published on the topic of Emission, as well as studies specific to Co2. The literature survey would include sources such as academic journals, industry reports, and online articles. It would aim to identify different internal and external factors that are responsible and commonly used to determine Co2 Emission, as well as any best practices or strategies that have been identified for reducing emission. The literature survey would also explore any existing research on Co2 Emission specifically, and would aim to identify any challenges or opportunities that the Countries can opt to reduce emission.

### **Activity 4: Social or Business Impact.**

**Social Impact:** Carbon dioxide emissions are the primary driver of global climate change. It's widely recognized that to avoid the worst impacts of climate change, the world needs to urgently reduce emissions.

**Business Model/Impact:** By conducting an analysis the countries can identify areas for improvement and take steps to reduce factors that are responsible for Co2 Emission for environmental sustainability by improving the efficiency and transitioning to low carbon alternatives.

## **Milestone 2: Data Collection & Extraction from Database**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes and generate insights from the data.

### **Activity 1: Collect the dataset**

Please use the link to download the dataset:

[https://drive.google.com/file/d/1n764uDPT\\_ZF7kzGFLtpxkwBBsDBScbWm/view?usp=sharing](https://drive.google.com/file/d/1n764uDPT_ZF7kzGFLtpxkwBBsDBScbWm/view?usp=sharing)

### **Activity 1.1: Understand the data**

Dataset consists of CO2 emissions in metric ton per capita of every country around the world. The data is collected from 1975 to 2020. In this dataset Countries and regions are included. Data is initially pre-processed using excel.

*The dataset contains*

- Country- Country for which Co2 is Recorded
- Year- Year the data was recorded
- Co2 Emission (In Million Metric Tons)
- Co2 Growth per Capita
- Co2 Per Capita
- Cumulative Co2
- Several Fossil Fuels rate of Emission

### **Activity 2: Storing Data in DB & Perform SQL Operations**

Initially you have to download the MySQL Server and connect it with using tableau desktop. After connection create a data base using MySQL query tab. Then, have to extract the above dataset and perform the following action

### **Activity 3: Connect DB with Tableau**

This activity is mentioned in the above activity 2

## **Milestone 3: Data Preparation**

### **Activity: Prepare the Data for Visualization**

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. Since the

Data is initially pre-processed we can skip this step. Basically this process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency.

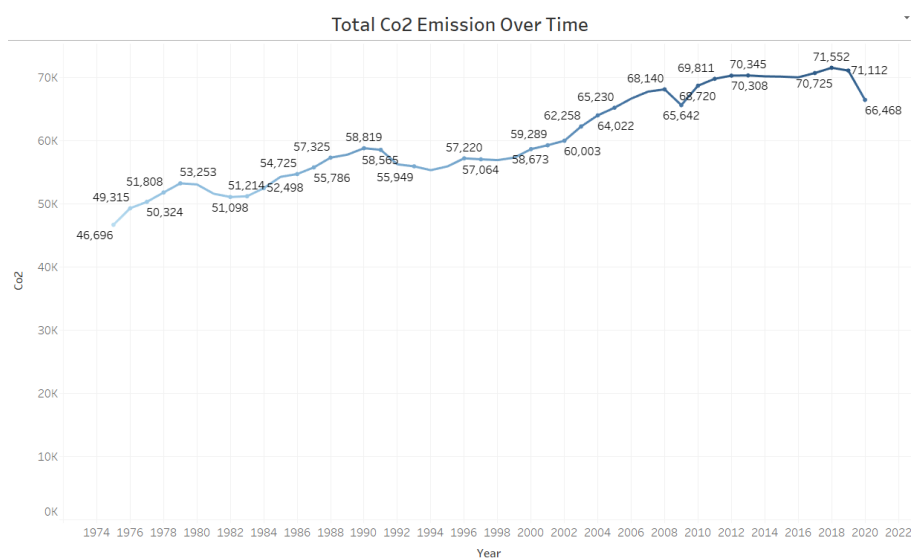
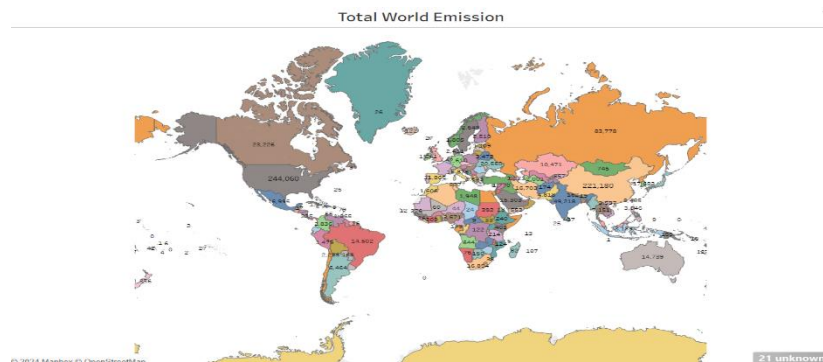
## **Milestone 4: Data Visualization**

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.

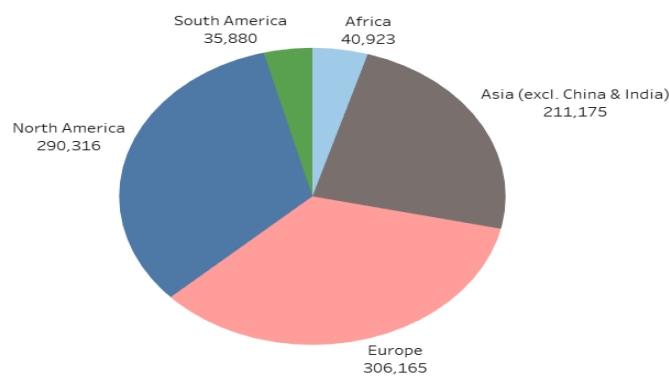
### **Activity 1: No of Unique Visualizations**

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.

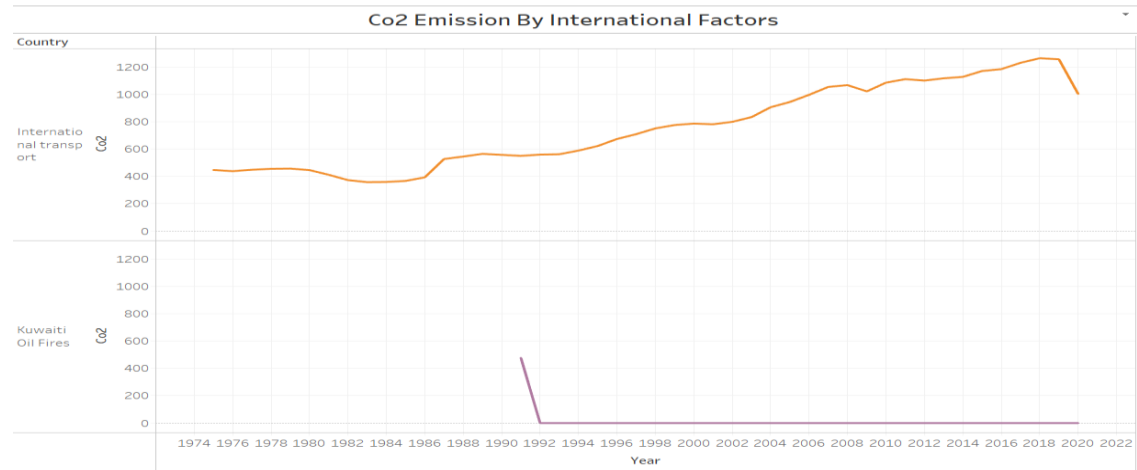
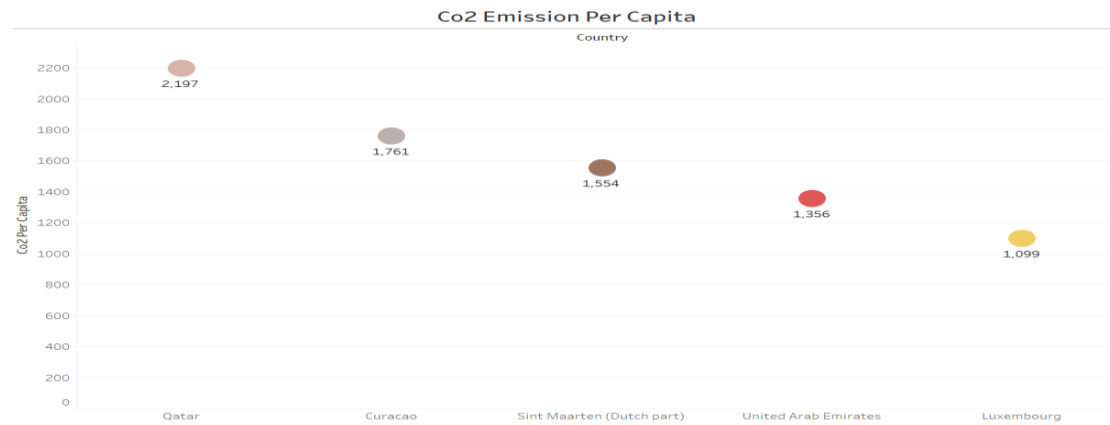
#### **Activity 1.1: To Understand- Total World Emission, Co2 Emission over Time, Total Emission by Continents**

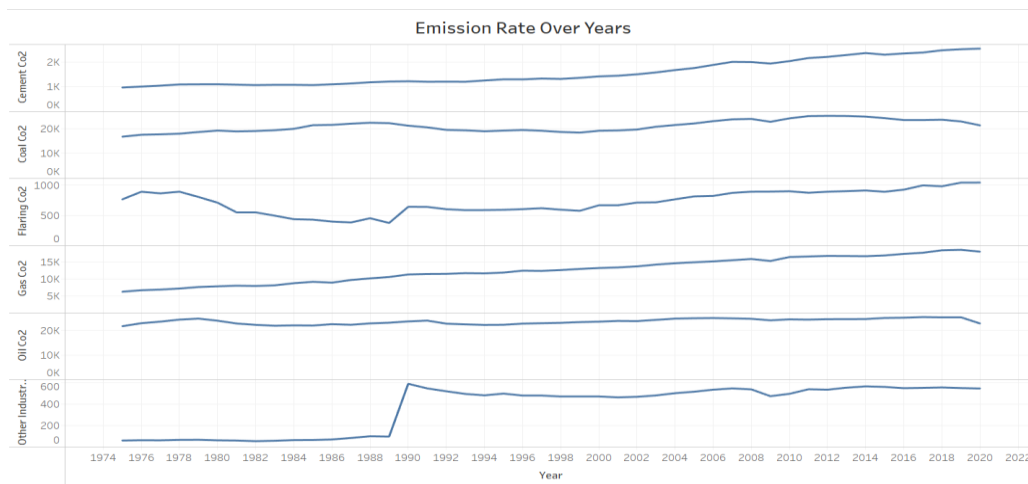


Total Emission By Continents



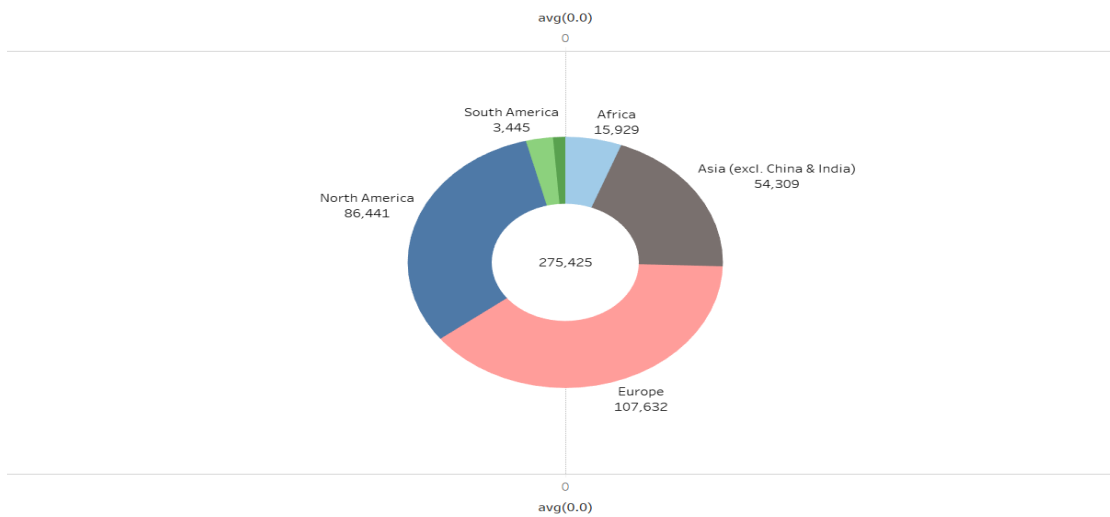
Activity 1.2: To Understand- Co2 Emission per Capita, Co2 Emission byInternational Factors, Emission Rate Over Years



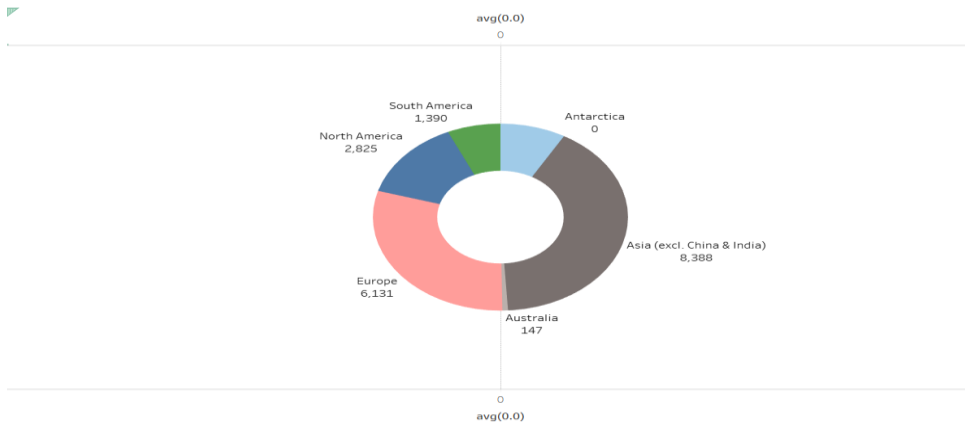


### Activity 1.3: To Understand- Co2 contribution by different fossil fuels, Co2 Emission over past 10 Years, Change in Co2 Emission and Co2 Emission in 2020

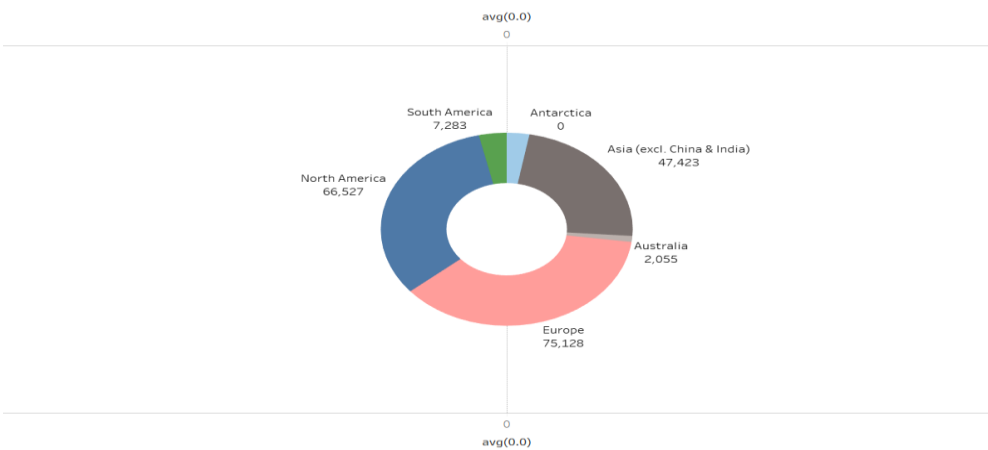
Donut Charts-Coal Co2



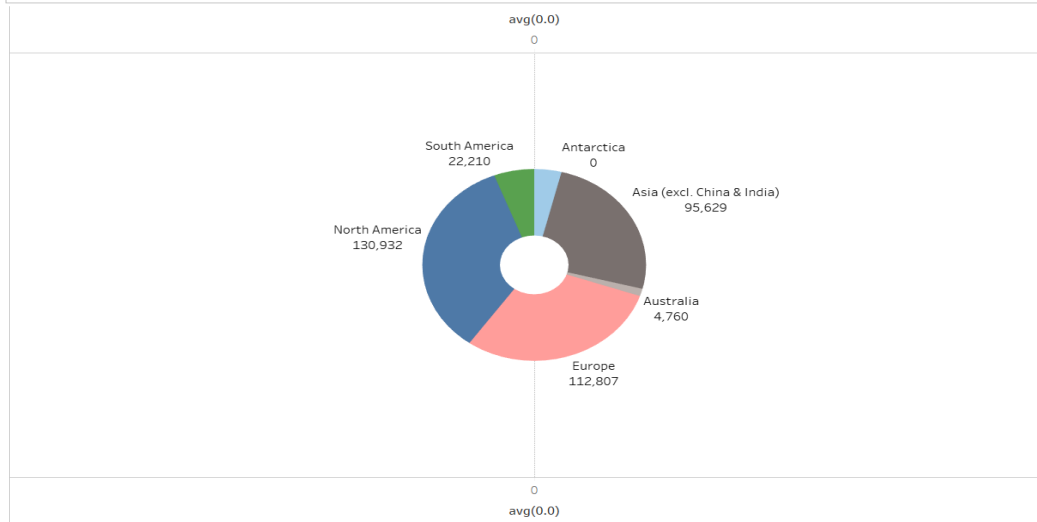
Donut Chart-Cement Co2



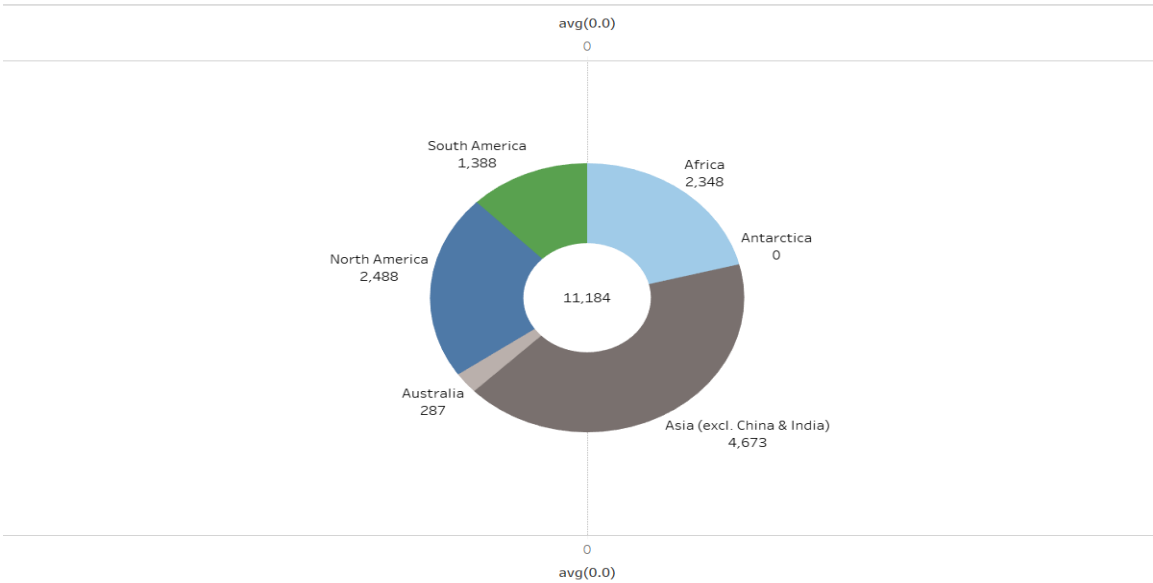
Donut Chart-Gas Co2



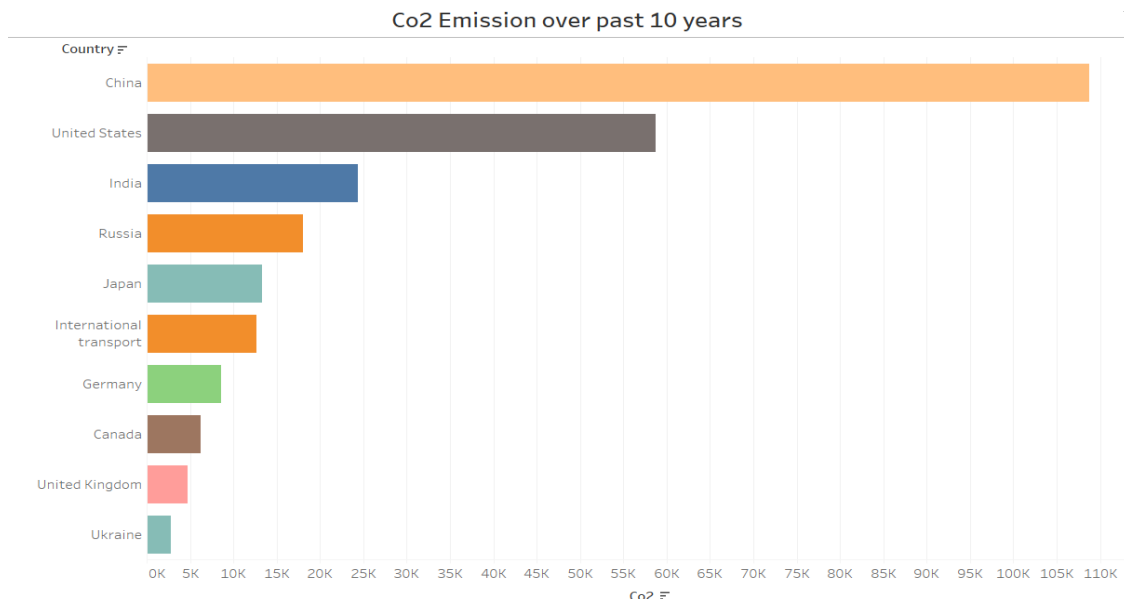
Donut Chart-Oil Co2



Donut Chart-Flaring Co2





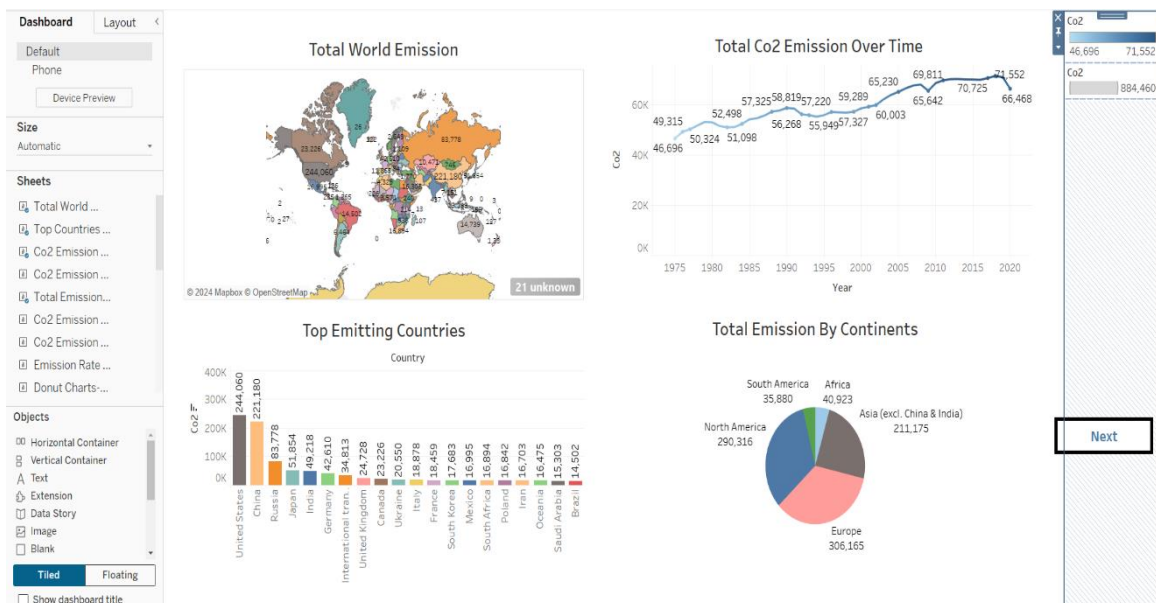


## Milestone 5: Dashboard

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data, and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

### Activity :1- Responsive and Design of Dashboard

The responsiveness and design of a dashboard for analyzing the globally Co2 Emission. It is crucial to ensure that the information is easily understandable and actionable. Key considerations for designing a responsive and effective dashboard include user-centred design, clear and concise information, interactivity, data-driven approach, accessibility, customization. Once you have created views on different sheets in Tableau, you can pull them into a dashboard.

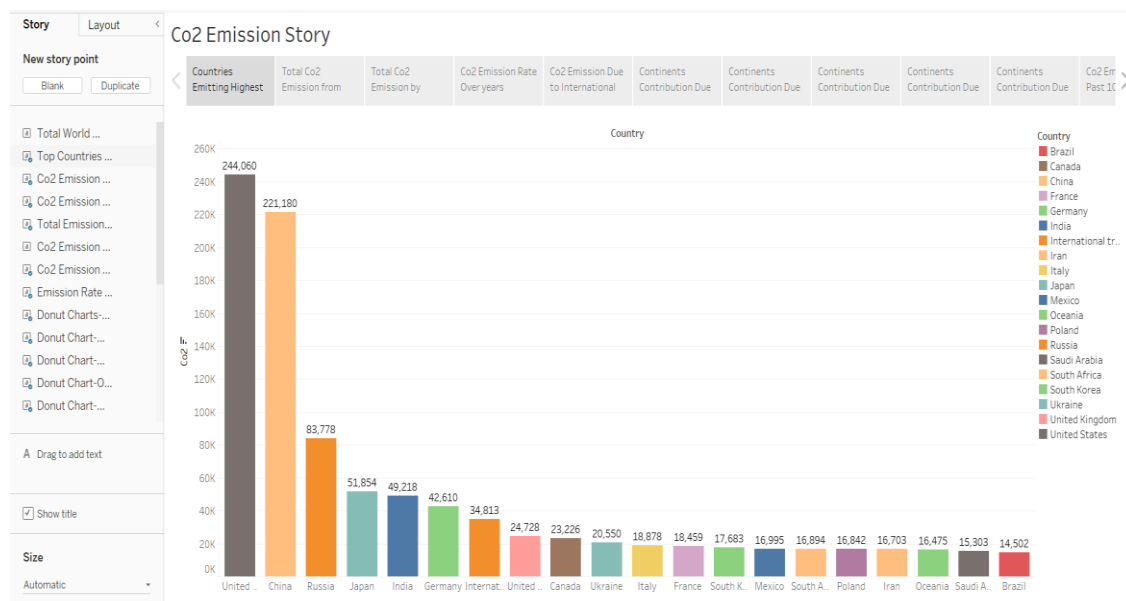


## Milestone 6: Story

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

### Activity: 1- No of Scenes of Story

The number of scenes in a storyboard for a data visualization analysis of the Co2 Emission will depend on the complexity of the analysis and the specific insights that are trying to be conveyed. A storyboard is a visual representation of the data analysis process and it breaks down the analysis into a series of steps or scenes.



## Milestone 7: Performance Testing

### Activity 1: Amount of Data Rendered to DB

- The amount of data that is rendered to a database depends on the size of the dataset and the capacity of the database to store and retrieve data.

- Open the SQL Server Management Studio, go to the database then click to expand the tables, select the table and Right click and select Properties

SQL Query Log - SQL Server Enterprise Edition (64-bit) - Microsoft SQL Server Management Studio

Object Explorer: Server: MSSQLSERVER (LOCAL) - Databases - CO2New

SQL Query Log - SQL Server Enterprise Edition (64-bit) - Microsoft SQL Server Management Studio

EXEC sp\_help 'dbo.CO2New';

Results: Messages

Line	Text	Severity	State	Text
1	CO2New	0	1	CO2New
2	CO2New	0	1	CO2New
3	CO2New	0	1	CO2New
4	CO2New	0	1	CO2New
5	CO2New	0	1	CO2New
6	CO2New	0	1	CO2New
7	CO2New	0	1	CO2New
8	CO2New	0	1	CO2New
9	CO2New	0	1	CO2New
10	CO2New	0	1	CO2New
11	CO2New	0	1	CO2New
12	CO2New	0	1	CO2New
13	CO2New	0	1	CO2New
14	CO2New	0	1	CO2New
15	CO2New	0	1	CO2New
16	CO2New	0	1	CO2New
17	CO2New	0	1	CO2New
18	CO2New	0	1	CO2New
19	CO2New	0	1	CO2New
20	CO2New	0	1	CO2New

Query executed successfully.

SQL Query Log - SQL Server Enterprise Edition (64-bit) - Microsoft SQL Server Management Studio

Object Explorer: Server: MSSQLSERVER (LOCAL) - Databases - CO2New

SQL Query Log - SQL Server Enterprise Edition (64-bit) - Microsoft SQL Server Management Studio

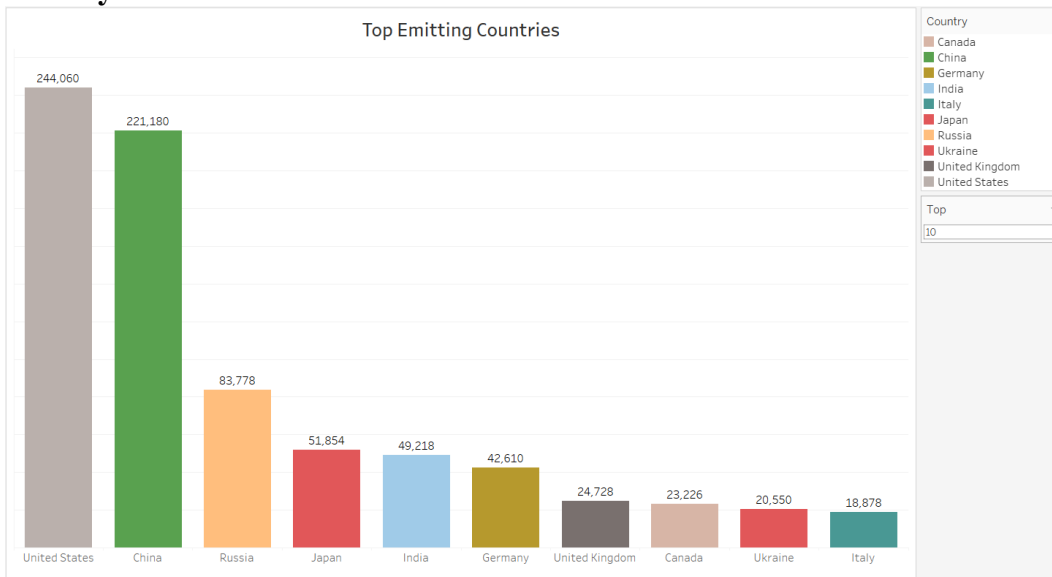
USE Proj2 SELECT \* FROM INFORMATION\_SCHEMA.COLUMNS

Results: Messages

Line	Text	Severity	State	Text
1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
2	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
3	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
4	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
5	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
6	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
7	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
8	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
9	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
10	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
11	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
12	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
13	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
14	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
15	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
16	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
17	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
18	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
19	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS
20	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS	0	1	USE Proj2 SELECT * FROM INFORMATION_SCHEMA.COLUMNS

Query executed successfully.

## Activity 2: Utilization of Data Filters



### Activity 3: No of Visualizations/ Graphs

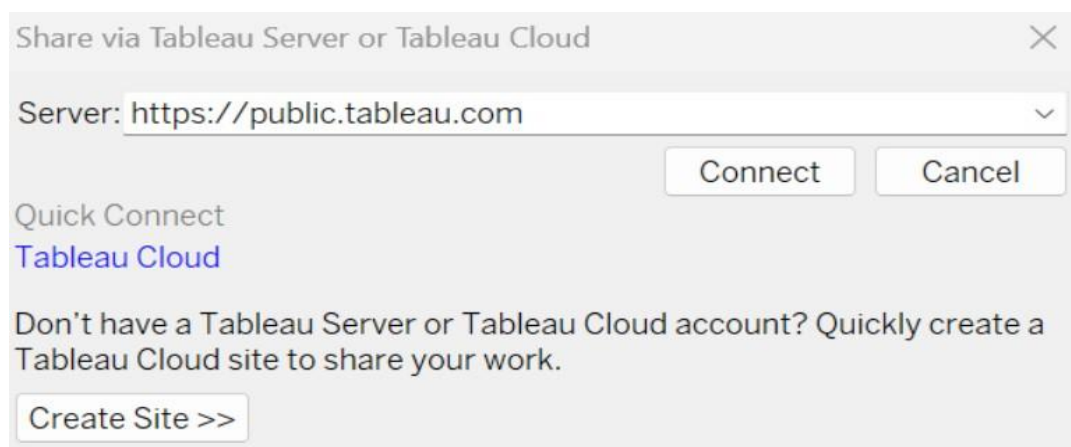
1. Top World Emission
2. Top Emitting Countries
3. Co2 Emission over Time
4. Co2 Emission India vs USA
5. Total Emission by Continents
6. Co2 Emission per Capita
7. Co2 Emission by International Factors
8. Emission Rate over Years
9. Donut Charts-Coal Co2,Cement Co2,Gas Co2,Oil Co2
10. Co2 Emission over past 10 years
11. Continent Contribution in Co2 Emission
12. Cumulative Co2 and Co2 per Capita
13. Co2 Emission in 2020
14. China vs India Co2 emission due to internal factors
15. Overall Contribution by China in Co2 Emission

### Milestone 8: Web integration

Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others.

#### **Publishing dashboard and reports to tableau public**

Step 1: Go to Dashboard/story, click on share button on the top ribbon

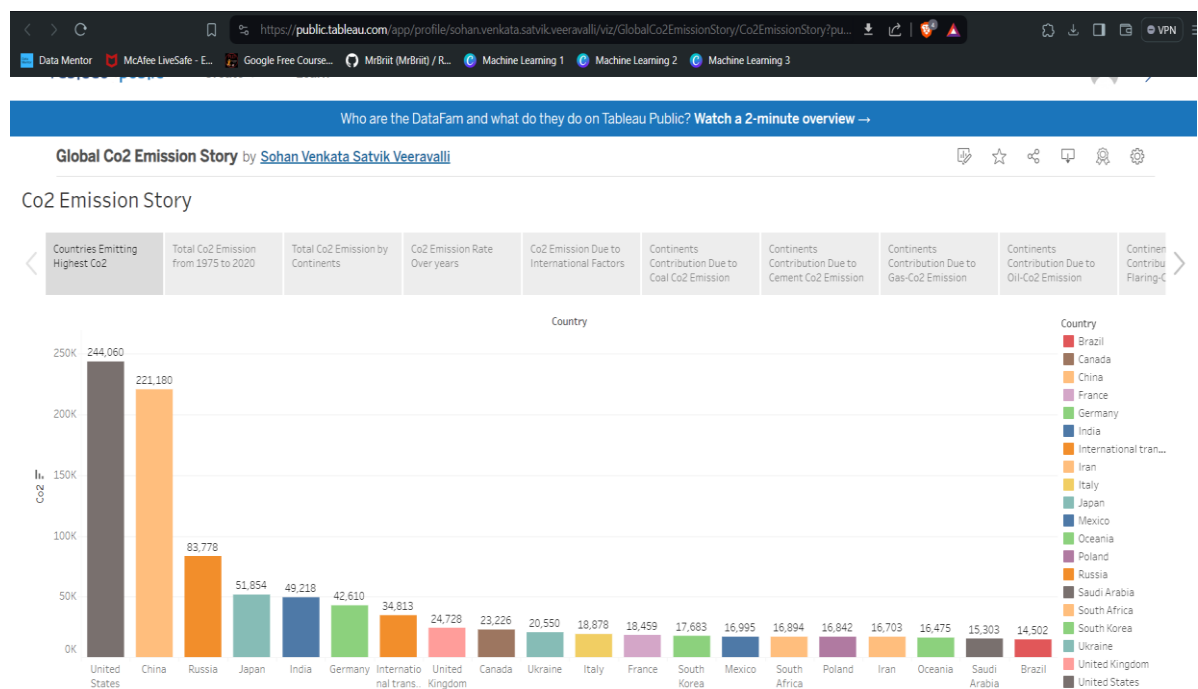


Give the server address of your tableau public account and click on connect.

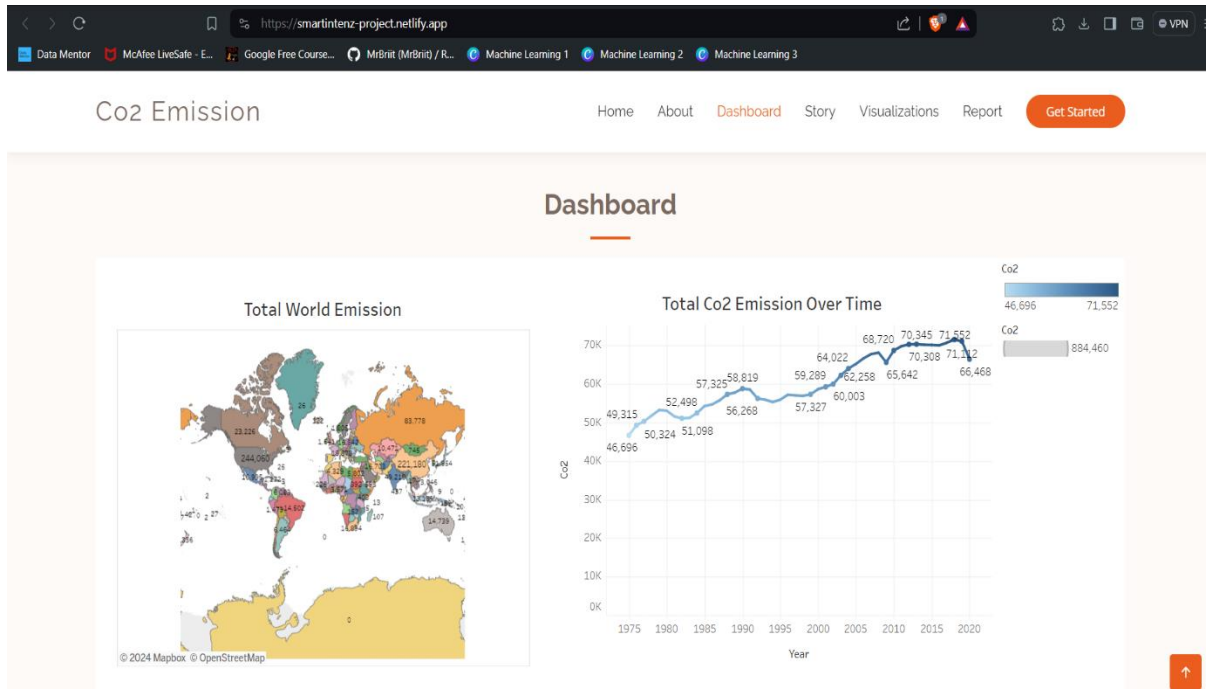
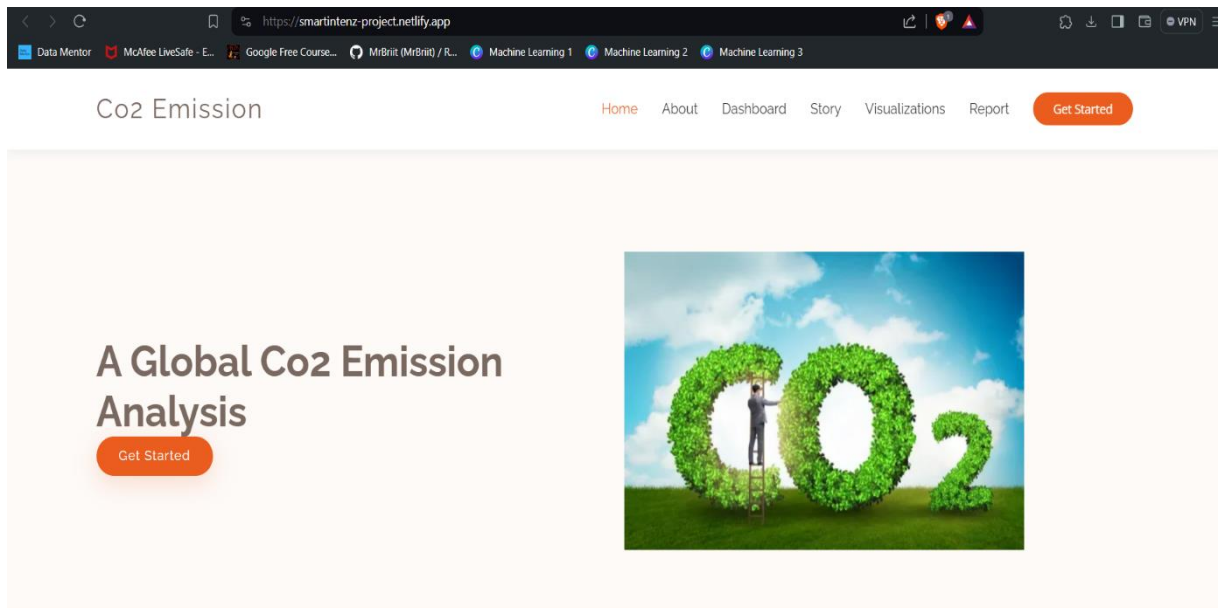
**Step 2:** Once you click on connect it will ask you for tableau public user name and password



**Note: While publishing the visualization to the public, the respective sheet will get published when you click on share option.**

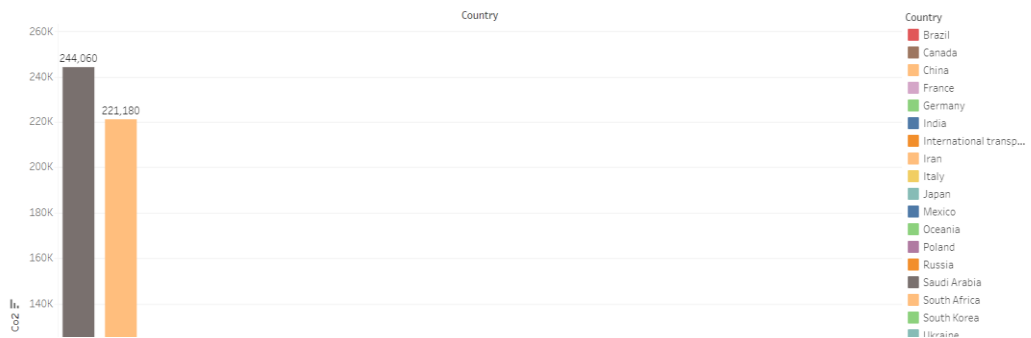


## Activity 1: Dashboard and Story embed with UI



## Story

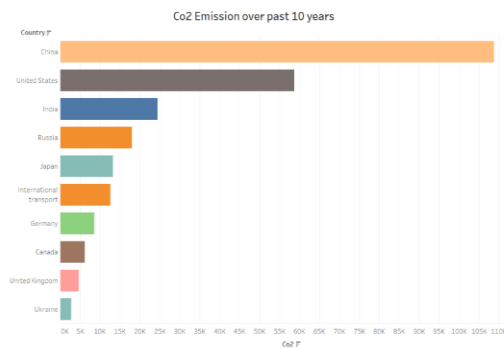
### Co2 Emission Story



## Visualizations



## Report



### Top Co2 Emitting Countries for Past 10 Years

China is the highest Co2 Emitting country among the other countries.

United States is the second highest Co2 Emitting country.

India is the Third highest Co2 Emitting country.

