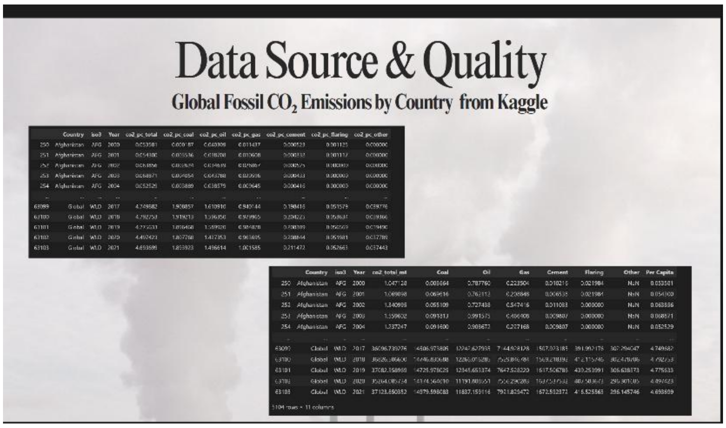
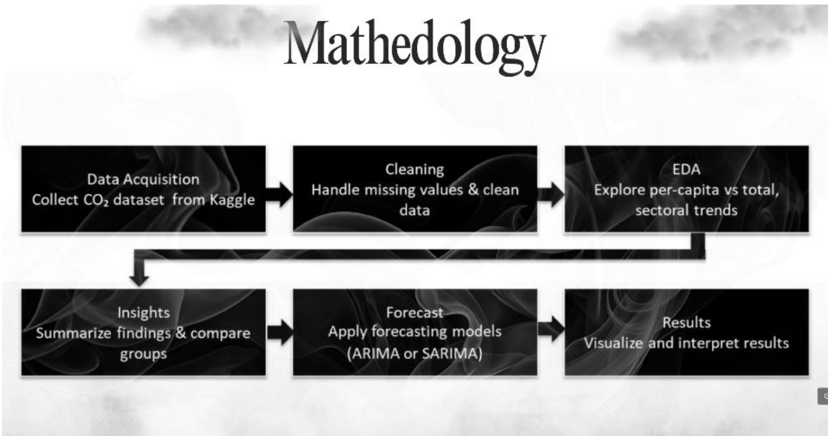


*"For our project, the data comes from a Kaggle dataset on global fossil CO_2 emissions by country, covering the years 2000 to 2022. The dataset includes both total and per-capita emissions, as well as emissions broken down by sources like coal, oil, gas, cement, flaring, and others.

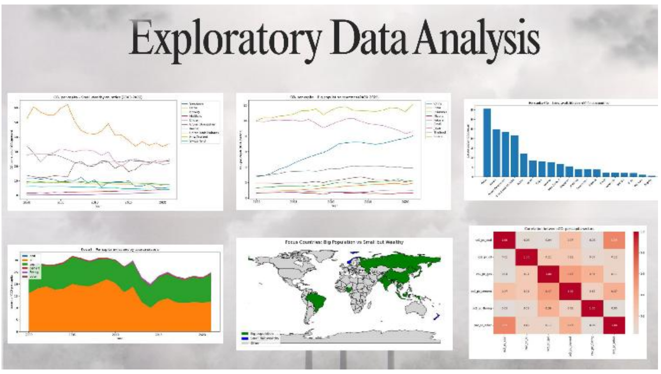
We can also see that the dataset provides global aggregates along with individual country data. This allows us to compare not only between countries but also against global trends.

In terms of quality, the dataset is comprehensive and consistent across years, though we did need to handle some missing values before moving forward.

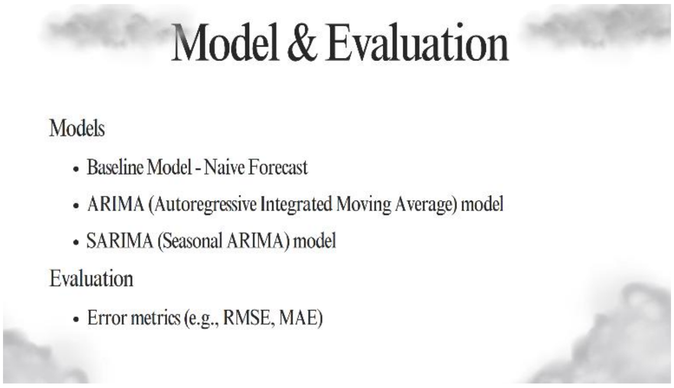
With this reliable data foundation, we then move on to our methodology—where we explain how the data was cleaned, explored, and later used for forecasting."*



"Our project follows a structured process. First, we collected the CO_2 dataset from Kaggle, covering the years 2000 to 2022. Then we cleaned the data, dealing with missing values to make sure it was reliable. Next, we performed exploratory data analysis to compare per-capita and total emissions, and also looked at sector-level patterns. After that, we summarized key insights and compared the two main groups of countries. Finally, we applied forecasting models, like ARIMA and SARIMA, to project emissions beyond 2022 and interpreted the results."



"In the exploratory analysis, we compared emissions from two focus groups: big population countries such as China, India, and the US, and small but wealthy countries such as Qatar, Singapore, and Norway.
The plots show that per-capita emissions are generally higher in small wealthy countries, while total emissions are dominated by big population countries.
We also mapped these focus countries and studied correlations between emission sources, like coal, oil, and gas. These insights reveal inequities in responsibility and highlight where reductions matter most."



"For modeling, we used three approaches. The baseline was a naïve forecast, which assumes the future will follow the last observed value. This gives us a simple benchmark.
Next, we applied ARIMA, the autoregressive integrated moving average model, which is widely used for time series forecasting. Finally, we tested SARIMA, which extends ARIMA to account for seasonality in the data.
To evaluate performance, we used error metrics like RMSE and MAE to measure prediction accuracy."

"From this project, we expect several outcomes. First, we provide clear evidence of emission inequality between countries, especially when comparing per-capita and total values. Second, we aim to raise awareness about responsibility sharing in climate change. The results also inform discussions on fairness—showing that small but wealthy nations cannot be overlooked just because their populations are small. Lastly, this work supports policy-making, offering insights that can guide negotiations on climate responsibility."

Possible Questions & Answers

Q1. Why did you choose ARIMA and SARIMA for forecasting?
A: We chose ARIMA because it is widely used for time series forecasting and can capture trends effectively. SARIMA was added to handle potential seasonality in the data, ensuring that periodic fluctuations are accounted for.

Q2. Why did you define two focus groups—big population vs. small but wealthy countries?
A: This distinction helps highlight global inequality. Big population countries have high absolute emissions, while small wealthy countries often have much higher per-capita emissions. Comparing the two gives a fairer picture of responsibility sharing.

Q3. How reliable are the forecasts beyond 2022?
A: Like any model, forecasts come with uncertainty. We used error metrics like RMSE and MAE to evaluate performance and ensure that the models are reasonably accurate. Still, forecasts should be interpreted as guidance rather than exact predictions.

Q4. What is the policy relevance of your findings?
A: Our results show that responsibility for reducing emissions should not only depend on total emissions but also per-capita contributions. This can support fairer climate policies and negotiations among countries.

Q5. What are the limitations of your study?
A: The main limitation is that the dataset only goes up to 2022, and we rely on historical data trends without accounting for sudden policy changes, technological shifts, or geopolitical events. These factors may affect future emissions in ways that models cannot fully capture.

- **Naive Forecasting**
  Naïve forecasting is an easy-to-implement approach that relies on your business's historical data. This method utilizes your past year's actual data as current period forecasting data. This way, you can quickly predict your future strategy based on your previous data. Due to its simplicity, it has various benefits such as being easy to implement, needing limited data, not being tricky for system integration, being an ideal technique for steady demand, and being appropriate for small businesses.
  Although this method is crucial for many organizations, it has its own limitations. For instance, it does not provide real-time data, lacks accuracy, is challenging to predict seasonal changes, and gives a more reactive approach than proactive decision-making.