Effects of Amazon deforestation on Forest fires and CO₂ emissions in Brazil.

Presented by: Muhammad Azam Khan Khattak

Methods of Advanced Data Engineering Project (Winter 2024/25)
FAU Erlangen Nürnberg, Department Mathematik
M.Sc. Data Science

Objectives

- **Introduction**

- **Datasets**

- **ETL Pipeline**

- **Analysis**

- **Conclusion**

Introduction: Why I choose this topic?

- The Amazon rainforest, Earth's largest tropical rainforest, is vital for regulating global carbon dioxide levels.

- Decades of extensive deforestation have significantly disrupted the Amazon's ecosystem. The region has become increasingly prone to forest fires.

- Human activities, such as slash-and-burn agriculture, are a primary cause.

- Deforestation and fires are interconnected, creating a feedback loop:

  a. Forest loss leads to drier conditions.
  b. Drier conditions, in turn, fuel more fire outbreaks.

Introduction: Objectives

- Investigate the relationship between:
    - Deforestation
    - Forest fires
    - CO_2 emissions

- Focus on the Amazon Basin states in Brazil.

- Provide a clearer understanding of how deforestation and fires impact CO_2 emissions.

- Utilize statistical analysis to uncover patterns and relationships.

Datasets Used for Analysis

Dataset 1: *Amazon_Deforestration*

- ○ Source: Kaggle

- ○ *Deforestation in Amazon (Brazil) Data*

- ○ Key Features:
  - ■ Annual deforestation data for Brazil's Amazon from 2004 to 2019.
  - ■ Reliable satellite-derived data from Brazilian government sources ensures accuracy.
  - ■ High-quality, well-structured CSV files with year, state, and area details.

Datasets Used for Analysis
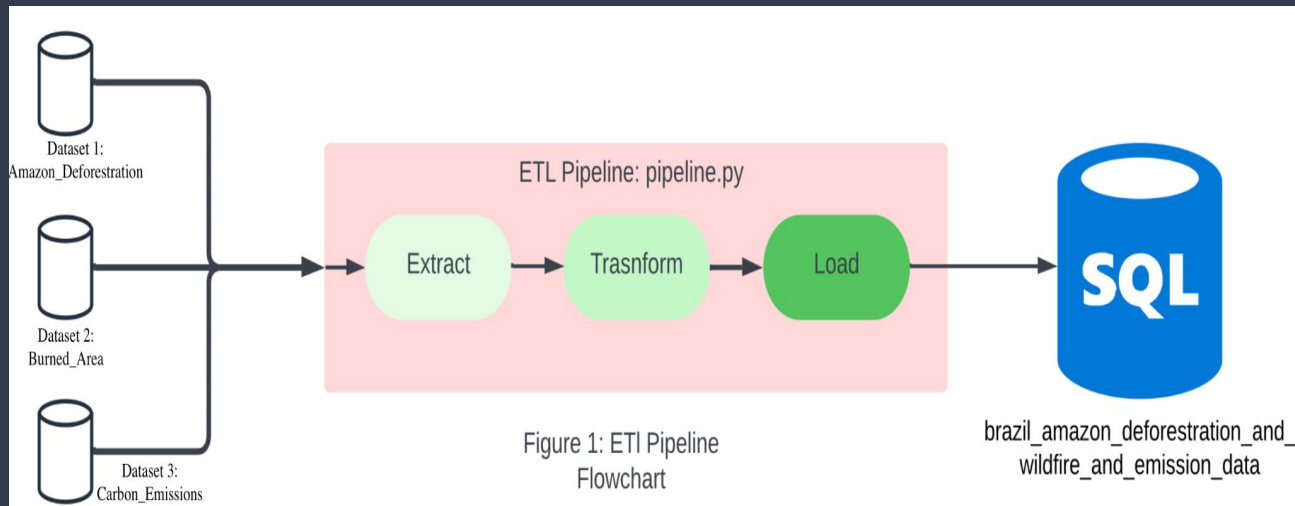
Dataset 2: *Burned_Area*

- ○ Source: Global Wildfire Information System

- ○ *Burned Area in Brazil*

- ○ Key Features:
  - ■ Monthly burned area data state wise from 2002 to 2023 .
  - ■ Categorized by land cover classes, including Amazon forest fire analysis.
  - ■ Consistently formatted CSVs with year, month, country, state, and land type.
  - ■ Enables correlation analysis between deforestation and forest fires in the Amazon.

Datasets Used for Analysis

Dataset 3: Carbon_Emissions

- Source: Global Wildfire Information System

- *Carbon Emissions in Brazil*

- Key Features:
  - Monthly emissions data from 2002 to 2023.
  - Categorized by pollutant type, including data for Brazil's Amazon.
  - Structured CSV files with year, month, country, and pollutant details.
  - High-quality, consistent format suitable for detailed environmental analysis.

ETL Pipeline

Figure 1: ETl Pipeline Flowchart

- Extraction
  - Kaggle API for dataset retrieval
- Transformation
  - Data cleaning, Transformation and Merging
- Loading
 - SQLite database for storing merged dataset

Analysis
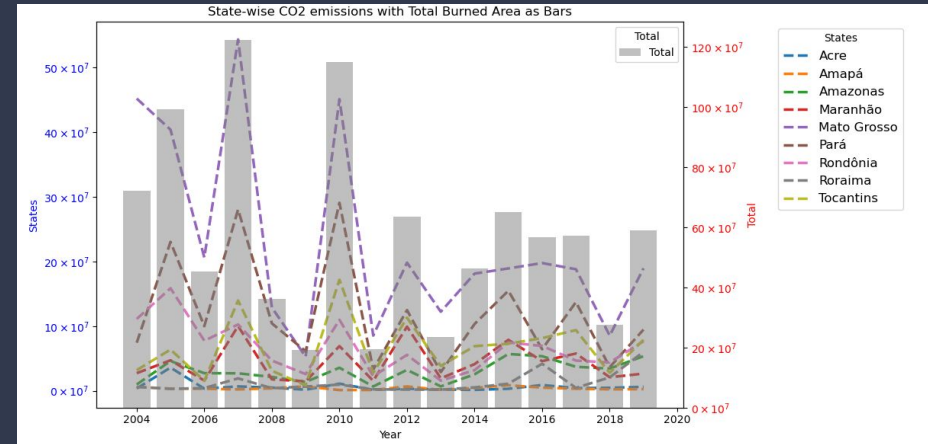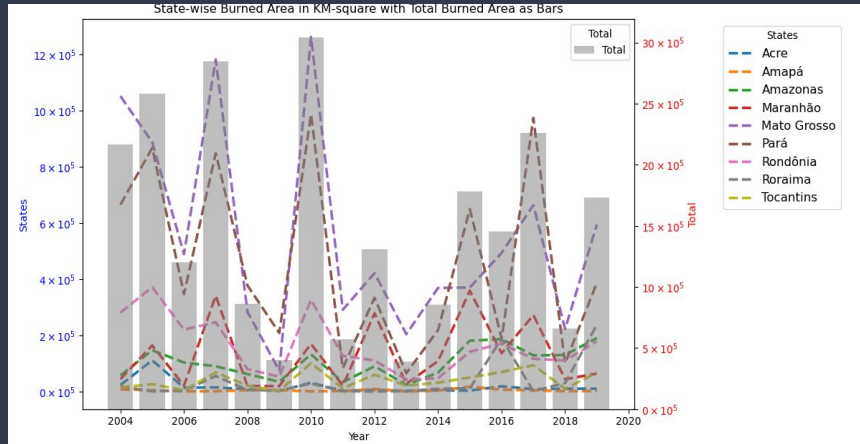
- Line graph: State-wise distribution; bar chart: total burned area/emissions.

- Mato Grosso, Para lead in forest burned area and CO_2 emissions.
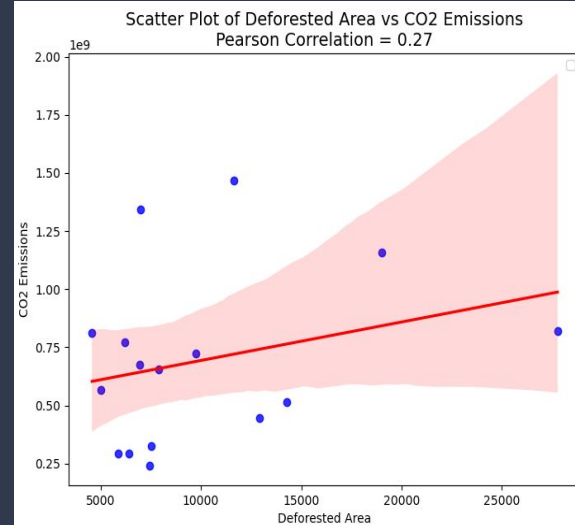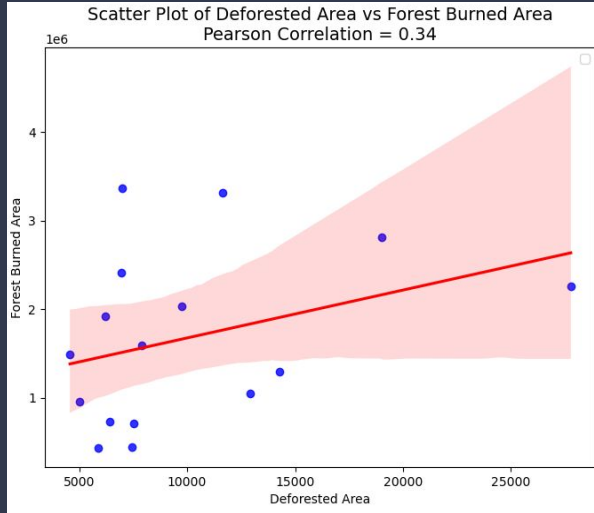
Analysis

Total Deforested Area by State
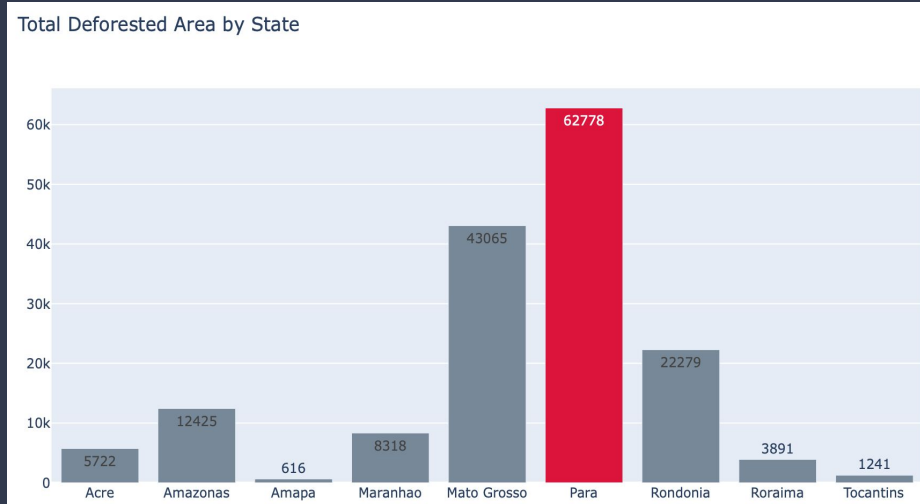
- These graphs show statewise and total deforestation that occurred from 2004 till 2019 in Brazil states that come under Amazon Basin.

- The most deforestation of the rainforest occurred in Para.

Analysis

Scatter Plot of Forest Burned Area vs CO2 Emissions
Pearson Correlation = 0.95
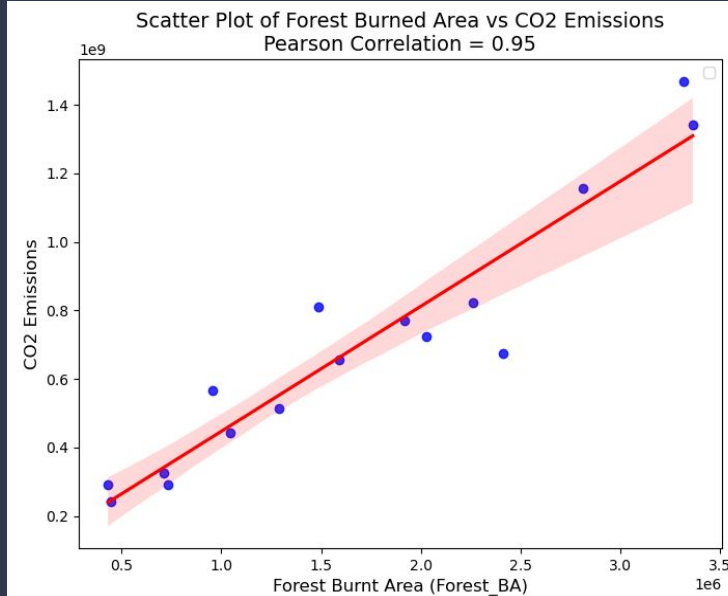
- Pearson correlation (0.95) shows strong link between burned area and CO_2 emissions.

- We have established link between CO_2 emissions and forest fires.

Analysis

Scatter Plot of Deforested Area vs Forest Burned Area
Pearson Correlation = 0.34

Scatter Plot of Deforested Area vs CO2 Emissions
Pearson Correlation = 0.27

- Predicted link between deforestation, fires, and CO_2 emissions was partially incorrect.

- Graphs show a positive Pearson correlation, but it's relatively weak.

- Deforestation's role in rising burned areas and CO_2 emissions is inconclusive.

Conclusion

- Para had the highest deforestation, but Mato Grosso led in emissions/burned area.

- Weak correlation suggests deforestation isn't the sole driver of emissions/fires.

- Other factors like climate, policies, and human activities significantly influence outcomes.

- Aggregated annual data may miss temporal variations and long-term deforestation impacts.

- Limited data and yearly aggregation reduced ability to generate meaningful correlations.

Limitations

- Analysis lacks details on wildfire causes like human activities and climatic factors.

- Annual data aggregation overlooks temporal variations and long-term deforestation impacts.

- Limited dataset prevented creation of a meaningful correlation heatmap for variables.