**Methods of Advanced Data Engineering**

**Abdelkhalek Mohamed (23405460)**

## How has rainforest degradation in the Brazilian Amazon contributed to carbon emissions ($CO_2$) and climate change?

### Introduction:

The Brazilian Amazon, often called the "lungs of the Earth," is critical to global climate stability, yet it faces growing threats from deforestation and fire activity. These processes not only degrade the rainforest's ability to absorb carbon dioxide but also release massive amounts of $CO_2$, a key driver of climate change. Understanding how rainforest degradation contributes to carbon emissions is vital for shaping effective environmental policies and promoting sustainable forest management. This study explores the connection between firespot activity and carbon emissions in the Brazilian Amazon between 1999 and 2019, aiming to uncover trends and insights that highlight the broader implications of these changes on climate systems.

### Used Data:

This study integrates three key datasets to analyze emissions and firespot activity in the Brazilian Amazon:

1. **FAOSTAT and UNFCCC Emissions Data:**
    - **Metadata URL:** https://www.fao.org/faostat/en/#data/GT/metadata .
    - **Data URL:** https://bulks-faostat.fao.org/production/Emissions_Totals_E_Americas.zip .
    - **Description:** This dataset provides $CO_2$ emissions data segmented by reasons (e.g., forest fires, savanna fires) from 1999 to 2019. It includes information from both FAOSTAT and UNFCCC sources. The data was filtered to include only Brazil and further refined to align emissions with firespot data from the Amazon region. Combining UNFCCC data with FAOSTAT data increases dataset richness by providing additional segmentation
    - **Structure:** Tabular format with columns for emission sources, years, and $CO_2$ values (in metric tons).
    - **Licensing:** Available under CC BY 4.0 International. Attribution to the UNFCCC and FAOSTAT is maintained, https://creativecommons.org/licenses/by/4.0/.
2. **Amazon Fires Dataset (Kaggle):**
    - **Metadata URL:** Brazilian Amazon Rainforest Degradation 1999-2019.
    - **Data URL:** Not applicable (downloaded using KaggleHub).
    - **Description:** This dataset contains firespot counts in the Brazilian Amazon from 1999 to 2019, highlighting temporal patterns of fire activity.
    - **Structure:** Tabular format with columns for years and firespot counts.
    - **Licensing:** Licensed under CC0 1.0 Universal https://creativecommons.org/publicdomain/zero/1.0/ .

### Reasons for Choosing These Data Sources

- **Relevance:** The datasets focus specifically on emissions from land-use changes and fire activity in the Brazilian Amazon, directly aligning with the study's objective of analyzing rainforest degradation and its impact on $CO_2$ emissions.

- **Coverage Period:** Both datasets span the same timeframe (1999–2019), ensuring temporal alignment for comparative and correlation analyses.
- **Granularity:** The datasets provide detailed emissions data segmented by reasons (e.g., forest fires, savanna fires) and firespot counts, enabling a deeper exploration of specific contributors to carbon emissions.
- **Open Data:** All datasets are publicly available under open licenses (CC BY 4.0 and CC0 1.0), ensuring transparency, legal compliance, and unrestricted use for analysis.

**Analysis**

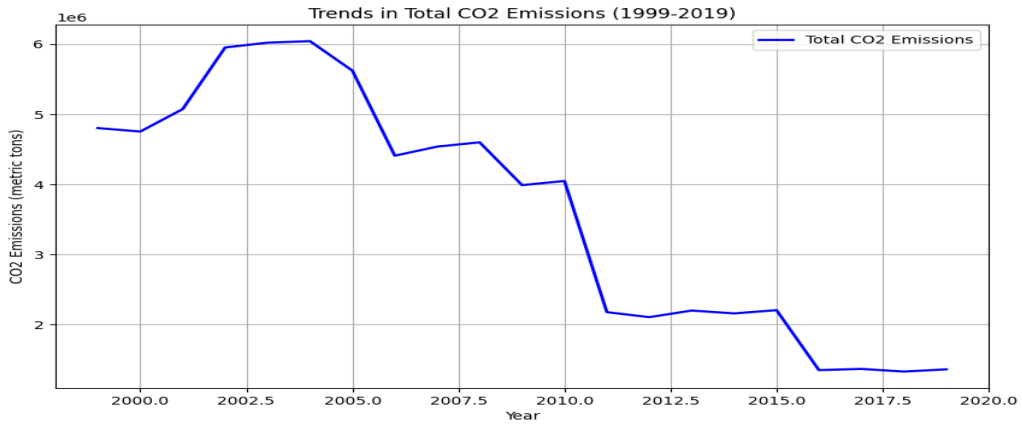
**Data Summary:**

This data summarizes the yearly changes in both **CO2 emissions** (aggregated across various emission sources) and **firespot counts** in the Brazilian Amazon region from **1999 to 2019**. The emissions data includes contributions from activities such as forest fires, crop residue burning, and deforestation, while the firespot data represents observed fire activities in the region. As shown in the table, the dataset provides detailed information on emission sources and reasons, making it more understandable for the reader.

| | Reasons of Emission | Element | Source | Year | Emissions | firespots |
|---|---|---|---|---|---|---|
| 2 | Burning - Crop residues | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 1999 | 1314.905 | 17.2043 |
| 3 | Burning - Crop residues | Emissions (CO2eq) from CH4 (AR5) | UNFCCC | 1999 | 3172.4 | 41.507875 |
| 4 | Burning - Crop residues | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 1999 | 322.6375 | 4.2214087 |
| 5 | Burning - Crop residues | Emissions (CO2eq) from N2O (AR5) | UNFCCC | 1999 | 779.1 | 10.193792 |
| 6 | Burning - Crop residues | Emissions (CO2eq) (AR5) | FAO TIER 1 | 1999 | 1637.543 | 21.425708 |
| 7 | Burning - Crop residues | Emissions (CO2eq) (AR5) | UNFCCC | 1999 | 3951.5 | 51.701666 |
| 8 | Drained organic soils | Emissions (CO2eq) (AR5) | FAO TIER 1 | 1999 | 420.4324 | 5.5009631 |
| 9 | Net Forest conversion | Emissions (CO2) | FAO TIER 1 | 1999 | 1385209 | 18124.155 |
| 10 | Net Forest conversion | Emissions (CO2) | UNFCCC | 1999 | 979523.6 | 12816.147 |
| 11 | Net Forest conversion | Emissions (CO2eq) (AR5) | FAO TIER 1 | 1999 | 1385209 | 18124.155 |
| 12 | Net Forest conversion | Emissions (CO2eq) (AR5) | UNFCCC | 1999 | 979523.6 | 12816.147 |
| 13 | Savanna fires | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 1999 | 5578.219 | 72.985754 |
| 14 | Savanna fires | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 1999 | 4820.297 | 63.069059 |
| 15 | Savanna fires | Emissions (CO2eq) (AR5) | FAO TIER 1 | 1999 | 10398.52 | 136.05481 |
| 16 | Forest fires | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 1999 | 16533.94 | 216.33109 |
| 17 | Forest fires | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 1999 | 4618.977 | 60.43497 |
| 18 | Forest fires | Emissions (CO2eq) (AR5) | FAO TIER 1 | 1999 | 21152.92 | 276.76606 |
| 19 | Burning - Crop residues | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 2000 | 1321.821 | 13.389696 |
| 20 | Burning - Crop residues | Emissions (CO2eq) from CH4 (AR5) | UNFCCC | 2000 | 2940 | 29.781415 |
| 21 | Burning - Crop residues | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 2000 | 324.3335 | 3.2854118 |
| 22 | Burning - Crop residues | Emissions (CO2eq) from N2O (AR5) | UNFCCC | 2000 | 720.8 | 7.3015116 |
| 23 | Burning - Crop residues | Emissions (CO2eq) (AR5) | FAO TIER 1 | 2000 | 1646.155 | 16.675108 |
| 24 | Burning - Crop residues | Emissions (CO2eq) (AR5) | UNFCCC | 2000 | 3660.8 | 37.082927 |
| 25 | Drained organic soils | Emissions (CO2eq) (AR5) | FAO TIER 1 | 2000 | 417.4538 | 4.2286956 |
| 26 | Net Forest conversion | Emissions (CO2) | FAO TIER 1 | 2000 | 1385209 | 14031.794 |
| 27 | Net Forest conversion | Emissions (CO2) | UNFCCC | 2000 | 979523.8 | 9922.3151 |
| 28 | Net Forest conversion | Emissions (CO2eq) (AR5) | FAO TIER 1 | 2000 | 1385209 | 14031.794 |
| 29 | Net Forest conversion | Emissions (CO2eq) (AR5) | UNFCCC | 2000 | 979523.8 | 9922.3151 |
| 30 | Savanna fires | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 2000 | 2245.541 | 22.746733 |
| 31 | Savanna fires | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 2000 | 1940.436 | 19.656099 |
| 32 | Savanna fires | Emissions (CO2eq) (AR5) | FAO TIER 1 | 2000 | 4185.977 | 42.402832 |
| 33 | Forest fires | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 2000 | 2436.249 | 24.678554 |
| 34 | Forest fires | Emissions (CO2eq) from N2O (AR5) | FAO TIER 1 | 2000 | 684.813 | 6.9369729 |
| 35 | Forest fires | Emissions (CO2eq) (AR5) | FAO TIER 1 | 2000 | 3121.062 | 31.615527 |
| 36 | Burning - Crop residues | Emissions (CO2eq) from CH4 (AR5) | FAO TIER 1 | 2001 | 1359.089 | 18.655792 |

**1. Trends in CO2 Emissions:**

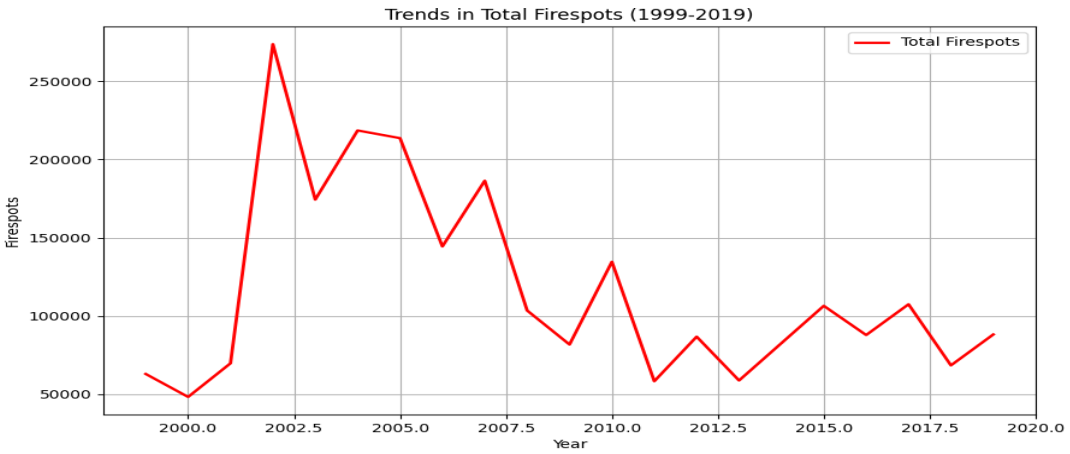
- As illustrated in the line graph, total CO2 emissions experienced a noticeable decline between 2005 and 2015, reaching their lowest levels in 2015. This decline was likely due to deforestation control policies and international environmental agreements. After 2015, emissions showed signs of stabilization, with minor fluctuations until 2019.




- 

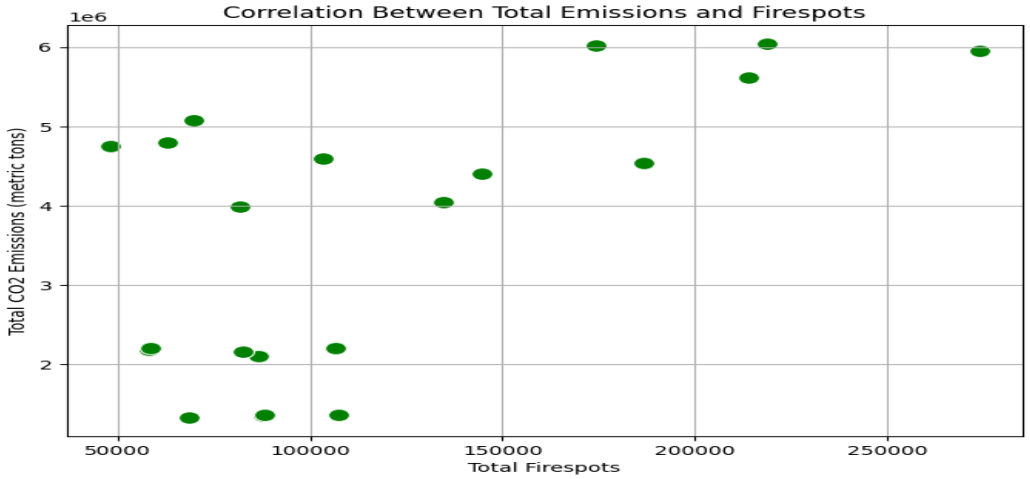
## 2. Trends in Firespots:

- Firespots showed significant volatility over the analyzed period. Peaks occurred in 2002 and 2004, followed by a general decline after 2010. This decline may be attributed to stricter enforcement of deforestation laws. However, intermittent surges in firespots were observed in later years, possibly due to policy relaxations or increased deforestation activit




## 3. Correlation Analysis:

- The scatterplot reveals a weak positive correlation (Pearson correlation coefficient of ~0.2) between total CO2 emissions and firespots.
- While years with high firespots correspond to relatively higher CO2 emissions, the relationship is not linear. The scattered data points suggest that other factors, such as industrial emissions, agricultural practices, and fire intensities, may play a significant role in influencing CO2 levels.



**Interpretation:**

- The analysis indicates that firespots alone cannot fully explain variations in CO2 emissions. Although both datasets reflect environmental activities in the Amazon, their weak correlation highlights the complexity of emission sources. Factors like industrial activity, urban expansion, and policy changes significantly contribute to the observed trends.
- This analysis provides a foundation for exploring more granular datasets, such as regional emissions or fire intensity data, to gain deeper insights into the interplay between firespots and CO2 emissions.

## Conclusions:

The analysis aimed to explore the relationship between **total CO2 emissions** and **firespots** in the Brazilian Amazon from **1999 to 2019**. The findings reveal the following:

1. **Key Findings**:
    - While firespots are one of the contributors to CO2 emissions, the weak positive correlation (Pearson coefficient ~0.2) suggests that they are not the sole or dominant factor driving CO2 emission levels.
    - Both datasets exhibit individual trends, with emissions generally declining during the period studied, while firespots displayed a more volatile pattern, influenced by deforestation policies and enforcement.
2. **Answer to the Question**: The posed question regarding the relationship between firespots and CO2 emissions was partially answered. While firespots contribute to emissions, the weak correlation suggests that other significant factors, such as industrial processes, transportation, agricultural practices, and potentially unmeasured variables, also play a critical role in determining CO2 levels. Additionally, the data used in this analysis may not comprehensively capture all relevant emissions sources or regional variations, leaving gaps that could impact the findings. This highlights the complexity of emissions dynamics and the need for more granular and extensive datasets to fully explore these relationships.
3. **Limitations**:
    - **Dataset Scope**: The analysis relied on aggregated annual data, which may mask finer temporal (monthly/seasonal) patterns or regional variations within the Amazon.
    - **Exclusion of Other Factors**: Emissions from non-fire-related activities, such as fossil fuel combustion or industrial emissions, were not separately analyzed.
    - **Data Cleaning Impact**: Rows with missing or zero values were removed, which could affect the completeness of the total yearly emissions data and potentially lead to underestimation of emissions.
    - **Data Quality**: The data's reliance on open-source and self-reported emissions could introduce inaccuracies or inconsistencies.
4. **Uncertainties**:
    - The specific contribution of each deforestation driver (e.g., agriculture, illegal logging) to emissions remains unclear.
    - Policy changes and enforcement variations over time might not be fully captured in the datasets.