

# Exploring Agricultural Impacts on Greenhouse Gas Emissions

Agriculture is essential for feeding the world's population and supporting economies. It provides some benefits but also accounts for massive greenhouse gas emissions (GHG) that accelerate global warming. The three largest crop producers worldwide, the USA, China, and India, also consume a lot of agricultural land whose productive practices have highly significant environmental effects. This work will attempt to understand how crop production and greenhouse gas emissions are related in these countries -and globally over 30 years (1990-2020). The key question is: **How does crop production influence greenhouse gas emissions?**

## Used Data

The data used for this analysis is within an SQLite Database named `merged.sqlite`, which is the final output of our created pipeline. The table within this database, which name is `merged_crop_emission` including data on greenhouse gas (GHG) emissions and crop production details for three different countries (China, USA, India) and the world over a period of 30 years between 1990 and 2020. this database is the output of the combination of two datasets for filtering and transformation, and other data engineering techniques are applied. The structure of this database is like the following:

#	Column	Non-Null Count	Dtype
0	Year	124 non-null	datetime64[ns]
1	Country	124 non-null	object
2	All GHG	124 non-null	float64
3	CH4	124 non-null	float64
4	N2O	124 non-null	float64
5	Maize	124 non-null	float64
6	Rice	124 non-null	float64
7	Soybean	124 non-null	float64
8	Wheat	124 non-null	float64
9	Total Crop Production	124 non-null	float64

dtypes: datetime64[ns](1), float64(8), object(1)

Notice:

- The considered unit for 'All GHG', CH4, and N2o is: MtCO2e.
- The the unit for considered for Maize, Soybean, Wheat, Rice and Total Crop Prodotion is: Thousand tonnes.

This project benefited from two datasets:

## Crop Production Dataset from OECD

This dataset includes information on crop production metrics world wide over 40 years between 1990 and 2030, focusing on major crops: Wheat, Maize, Rice, and Soybean. It includes data measured in tonnes per hectare and total production quantities Measured in thousand tonnes. This dataset is in CSV format and can be downloaded from Kaggle. The crop production dataset is provided by OECD, which generally allows usage for academic purposes [1].

## Historical Emissions Gases Dataset

Source: Climate Watch, 2024. Washington, DC: World Resources Institute. This dataset includes greenhouse gas (GHG) emissions data for all countries for the years between 1990 and 2020, covering multiple sectors and including emissions of various gases also it is in the CSV formation and can be downloaded from Kaggle. It is provided under the Creative Commons CC BY 4.0 license, permitting unrestricted reuse with proper attribution.

**Academic Use Disclaimer:** The purpose of this project is strictly for academic reasons and not for commercial use. The datasets used in this project are intended to provide educational insights and facilitate learning on the impact of crop production on greenhouse gas emissions.

## Analysis

### Method

The analysis phase involved finding the relationship between total emitted greenhouse gases (GHG) and total crop production comprising the sum amount of Wheat, Rice, Soybean, and Maize in hectare, measured across three different countries (China, USA, India) and globally over the period of 30 years, from 1990 and 2020. the following steps are performed:

- **ETL:** including three different phases(Extract, Transform, Load) explained in the previous task and can be found in the `data-report.pdf`.

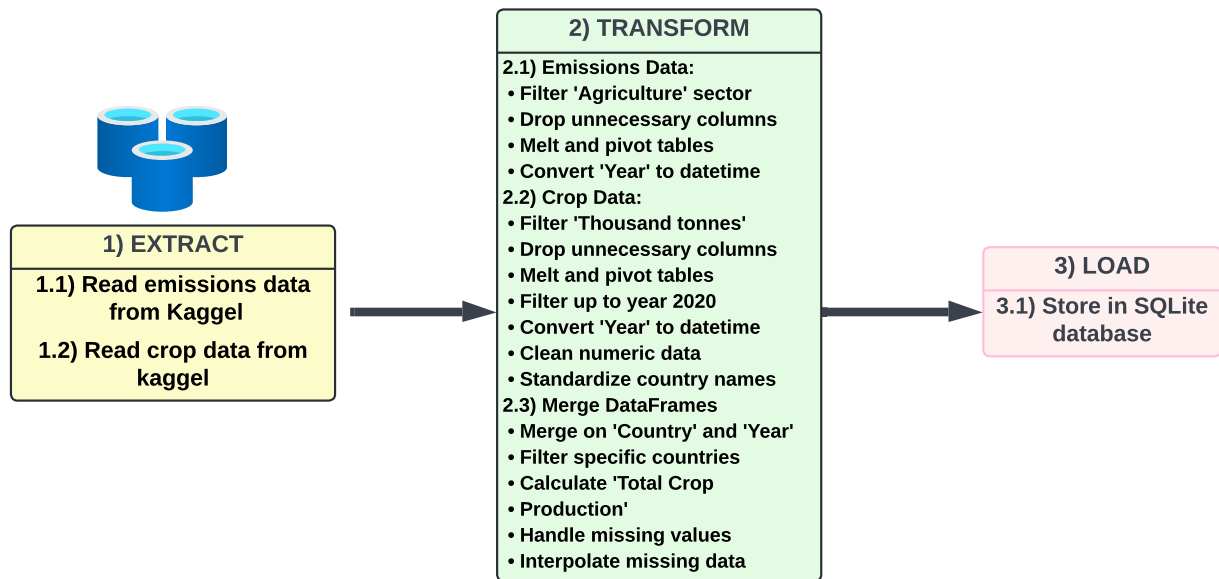


Figure 1: ETL Diagram of the project

- **Visualization:** Line plots are created to have an intuitive insight into our output data.
- **Corollation Analysis:** to quantify the relationship between GHG emissions and total crop production on the scale of global, a scatter plot and corollation coefficient are calculated.

## Interpretation



Figure 2: GHG emissions and crop production graphs

- **China:** GHG emissions and total crop production, which gradually increased Over the period 1990 to 2020, emissions showed quite some ups and downs, with very notable dips around 1996 and again between approximately 2016.
- **India:** Both GHG emissions and crop production have increased over the period of 30 years, with a strong alignment between the two indicators.
- **United States:** GHG emissions indicated more variability compared to a more consistent increase in total crop production.
- **World:** There is an upward trend in both GHG emissions and total crop production globally, indicating a parallel increase over the period of 30 years. Also there is a high correlation between global GHG emissions and crop production, as indicated by the scatter plot with a correlation coefficient of approximately 0.9788.

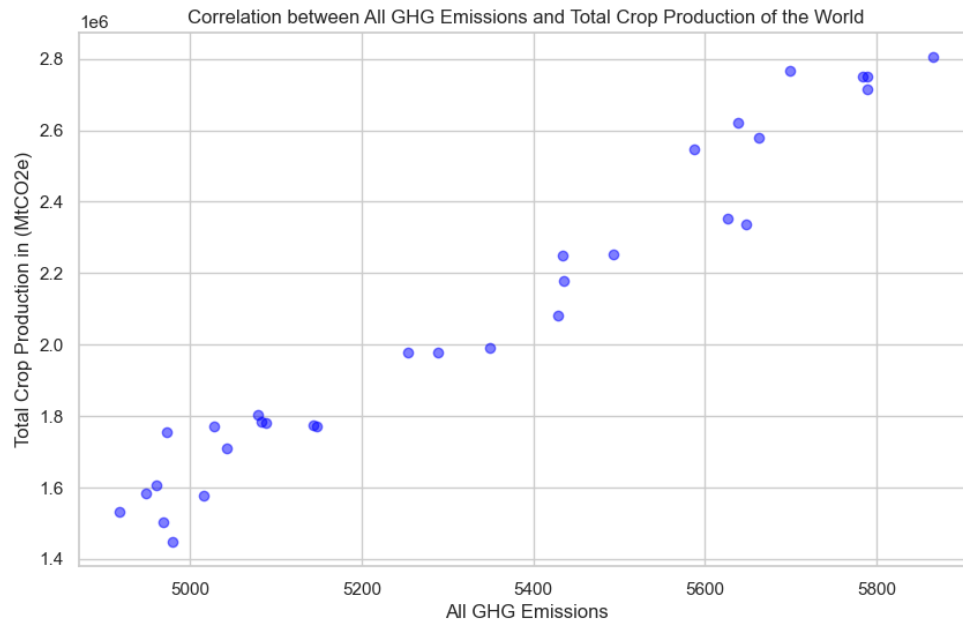


Figure 3: Scatter plot

## 0.1 Conclusion

Analysis of the data demonstrates a significant relationship between crop production and GHG (greenhouse gas) emissions for China, India, and the USA, as well as at a global scale. Agricultural emissions will also grow with the growth in crop production. This is significantly more pronounced in the global scale, where a correlation coefficient of around 0.9788 reflects an approximately linear relationship. The observed trends suggest that agriculture practices including the utilization of fertilizers, pesticides, irrigation, and mechanization, significantly are leading causes of GHG emissions.

## Bibliography

- [1] "OECD (2024), Crop production (indicator). doi: 10.1787/49a4e677-en (Accessed on 31 May 2024)." [Online]. Available: <https://data.oecd.org/agroutput/crop-production.htm#indicator-chart>