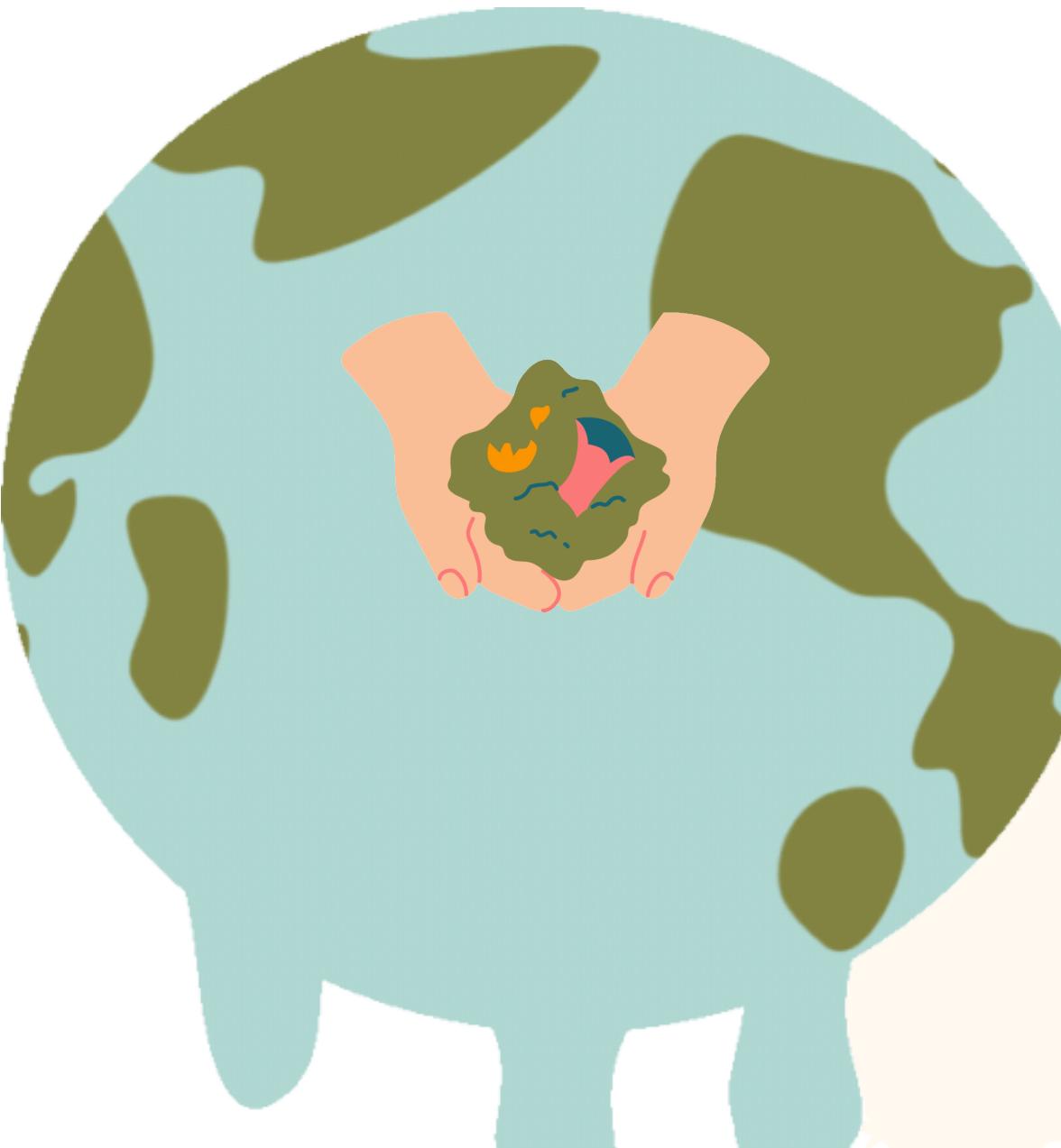


# Food & Climate Data-Driven Sustainability Project

Understanding the Causes, Impacts, and Solutions of  
Food Production and climate change  
(1961 - 2022)

Ivana Ruiz  
All Woman in Data (July 2025)



# Table of contents

Overview

I. Data Analysis

II. EDA: Exploring relationships

III. Hypothesis testing

V. Conclusions



# Datasets



📌 FAOSTAT - Food and Agriculture Organization of the United Nations.  
(1961-2023).  
Retrieved from: [FAO](#)



📌 Climate Data - Temperature Anomalies (FAOSTAT 1961 - 2022)  
Retrieve from: [Kaggle](#)

# Research questions

## Global Food production trends & Environmental Impact by CO<sub>2</sub> Emissions

1. How has food production evolved over time and which countries are the largest producers to both food production and its environmental impact?
2. Which foods contribute the most to greenhouse gas emissions (GHG CO<sub>2</sub>eq)?
3. Which countries have experienced the most significant temperature changes?
4. Is there a discernible relationship between food production, its environmental impact, and rising temperature anomalies?



# I. Data Analysis

I. Data Analysis:

ii. Descriptive statistics.

iii. Categorical and numerical variables

iv. Outliers

v. Key Performance Indicators (KPIs):

- Top 10 countries by food production (t)
- Top 10 countries by avg GHG emission in food production globally.
- Most produced food items globally

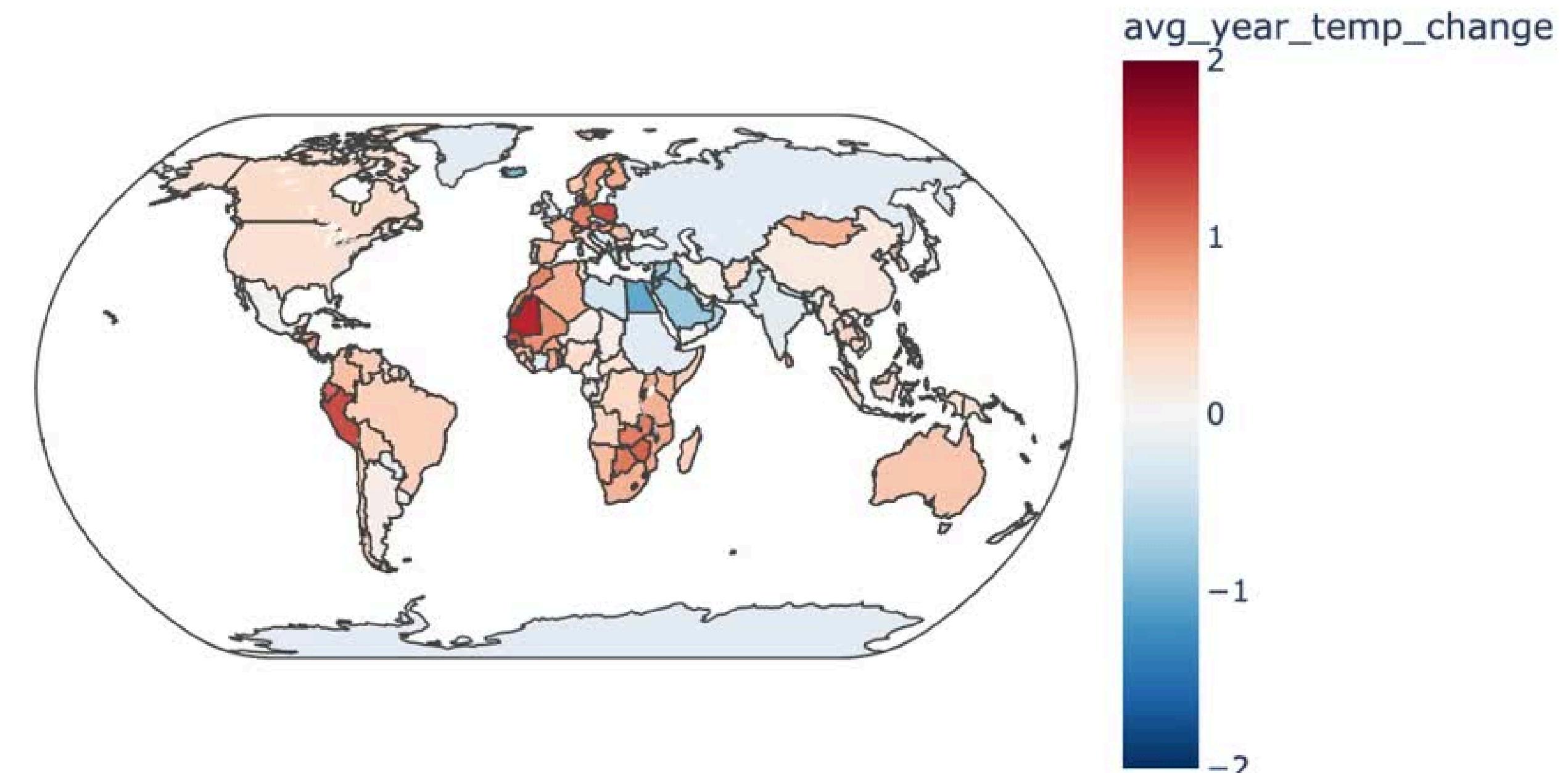


# What is Climate Change?

And how is  
related with  
food production?

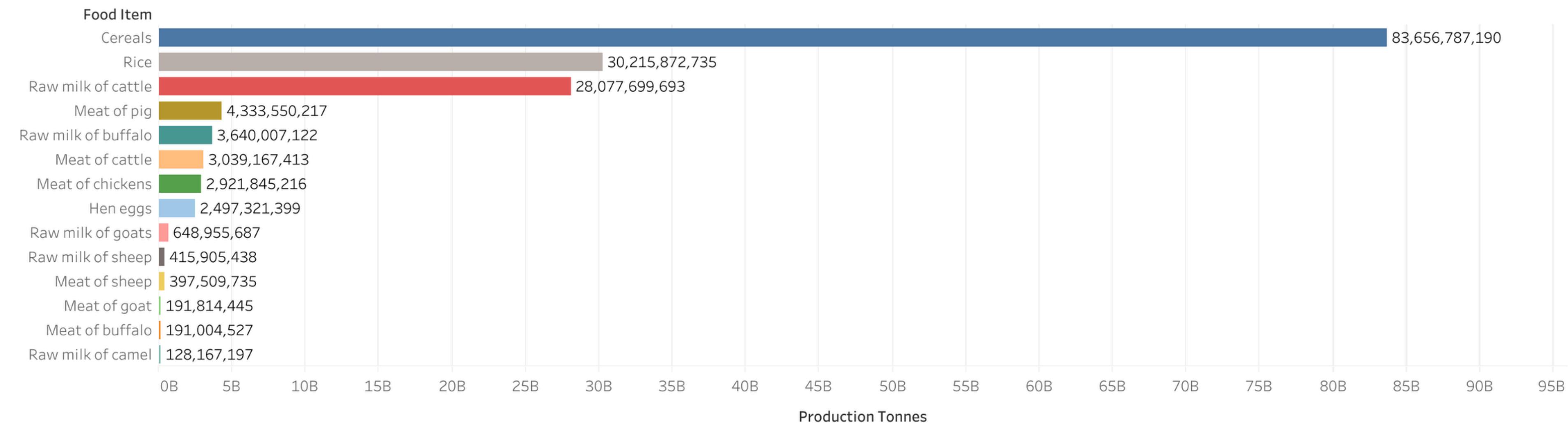


Temperature Change: 1983



- emissions\_CO2eq\_AR5: Carbon Dioxide Equivalent Emissions

***Most produced Food globally  
(1961 - 2022)***

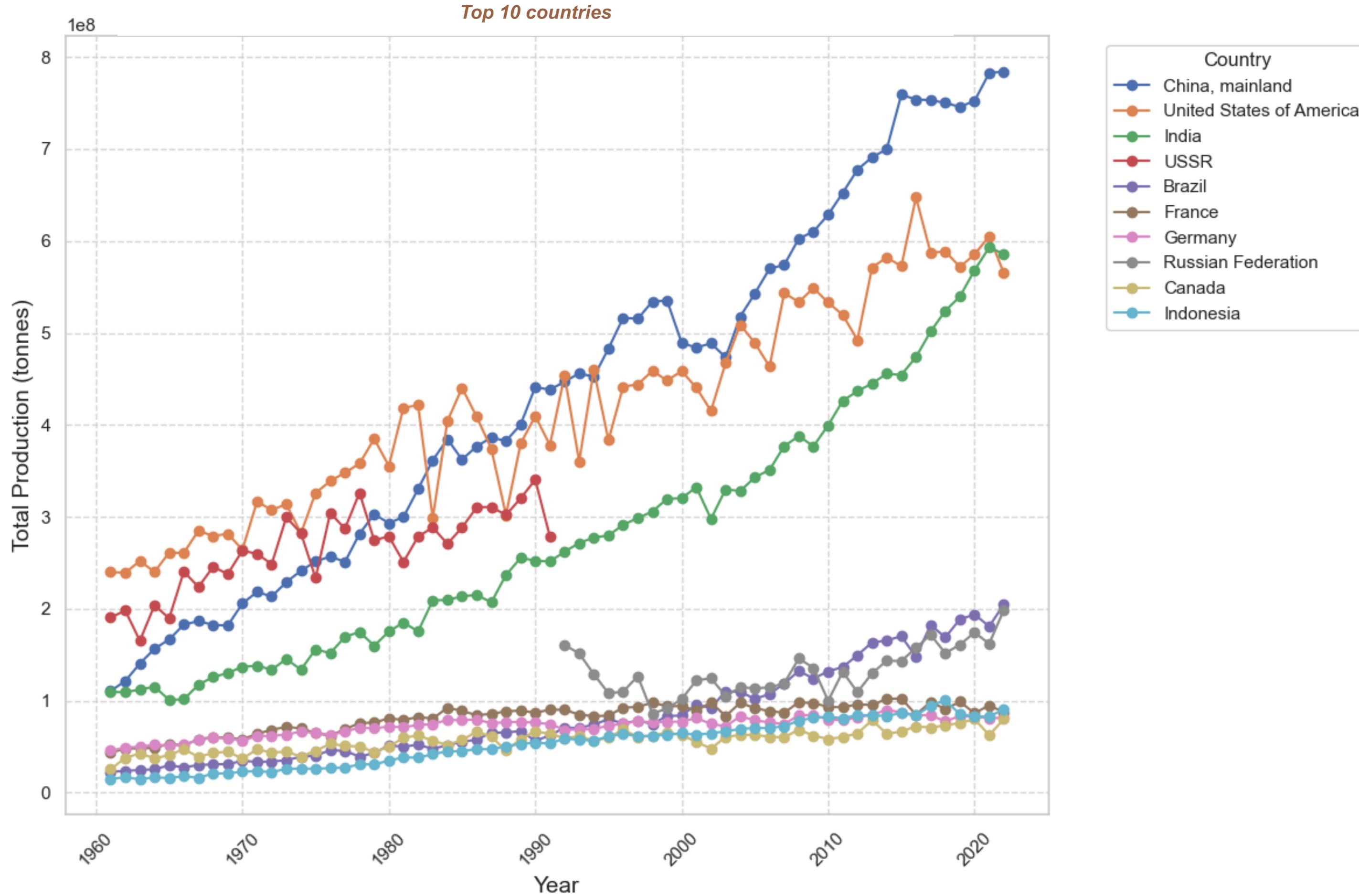


Even though plant-based products (like cereals and rice) are produced in greater volumes, animal-based products contribute far more to GHG emissions per unit. !!!

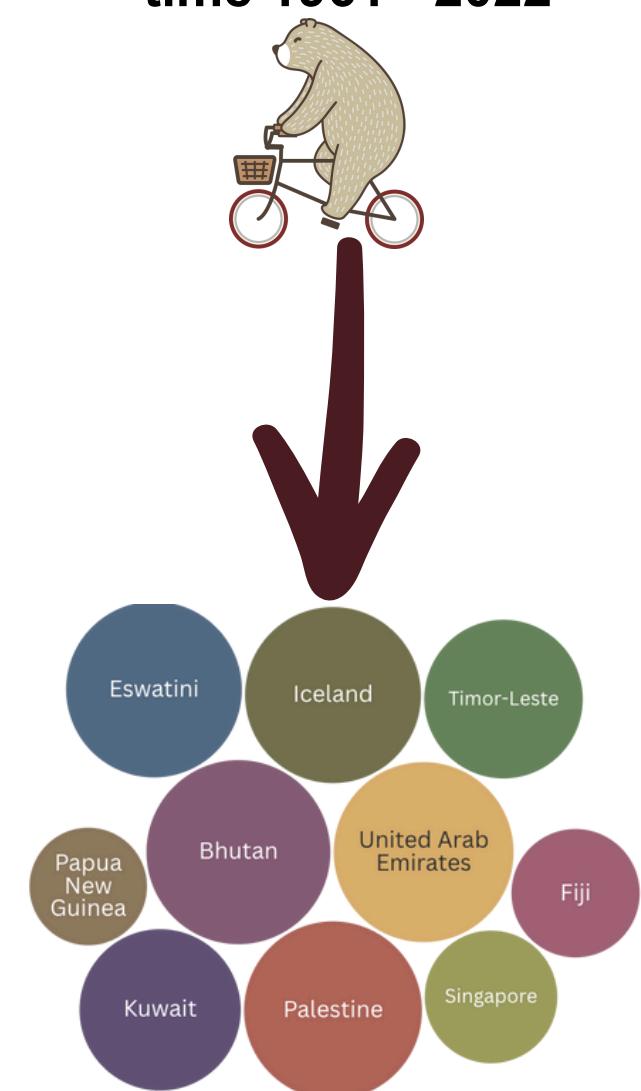
**“While plant foods lead in volume, animal products dominate in climate cost..”**

Data Analysis: How has food production evolved over time, and which countries are the largest producers?

## Total Food production over time 1961 - 2022



**Less 10 countries with food production over time 1961 - 2022**



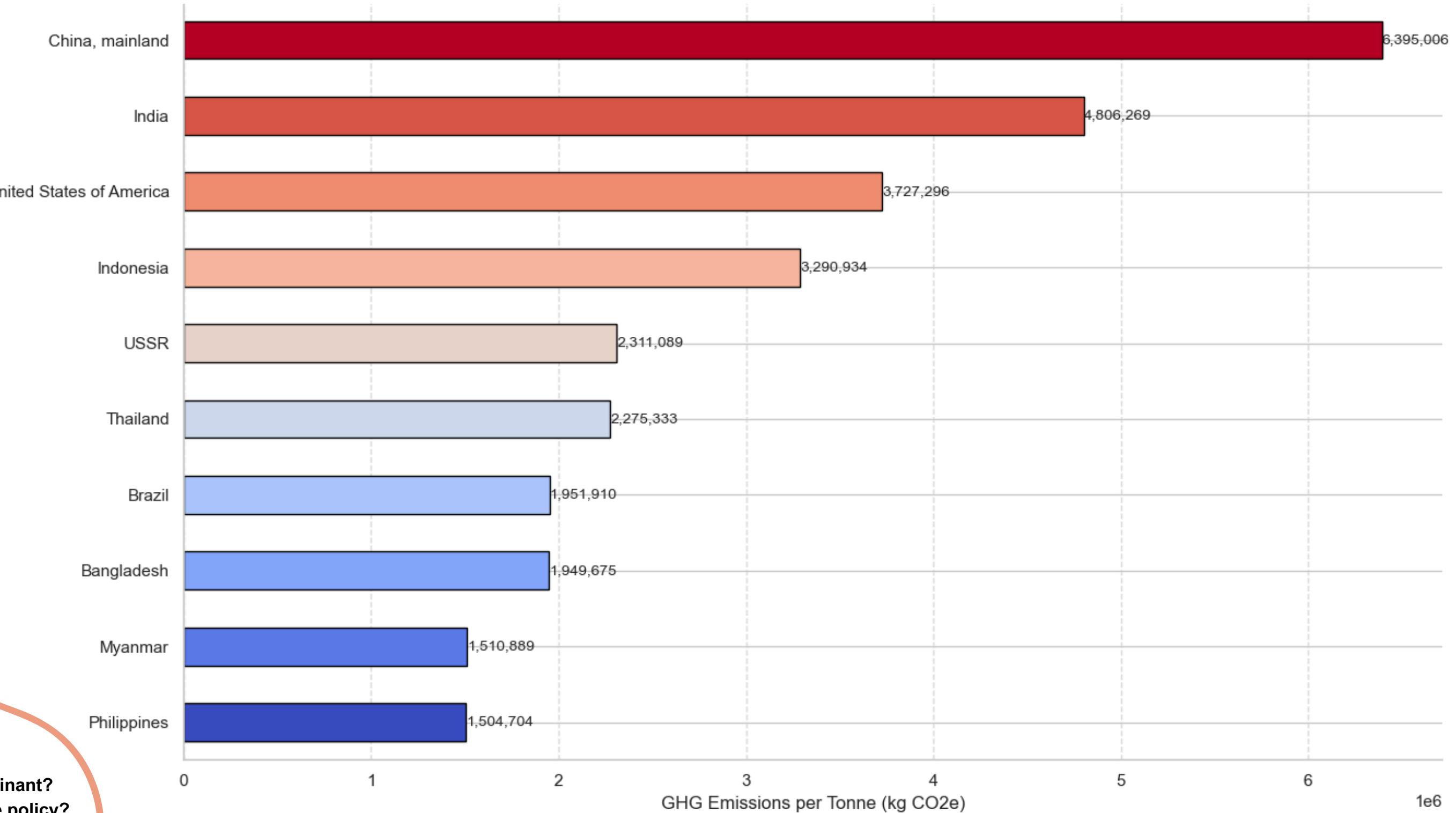
# Role of Greenhouse Gases



**Main Greenhouse Gases:** Carbon dioxide ( $\text{CO}_2$ ), methane ( $\text{CH}_4$ ), and nitrous oxide ( $\text{N}_2\text{O}$ ) **How It Works?**  
**The greenhouse effect trapping heat in the Earth's atmosphere**

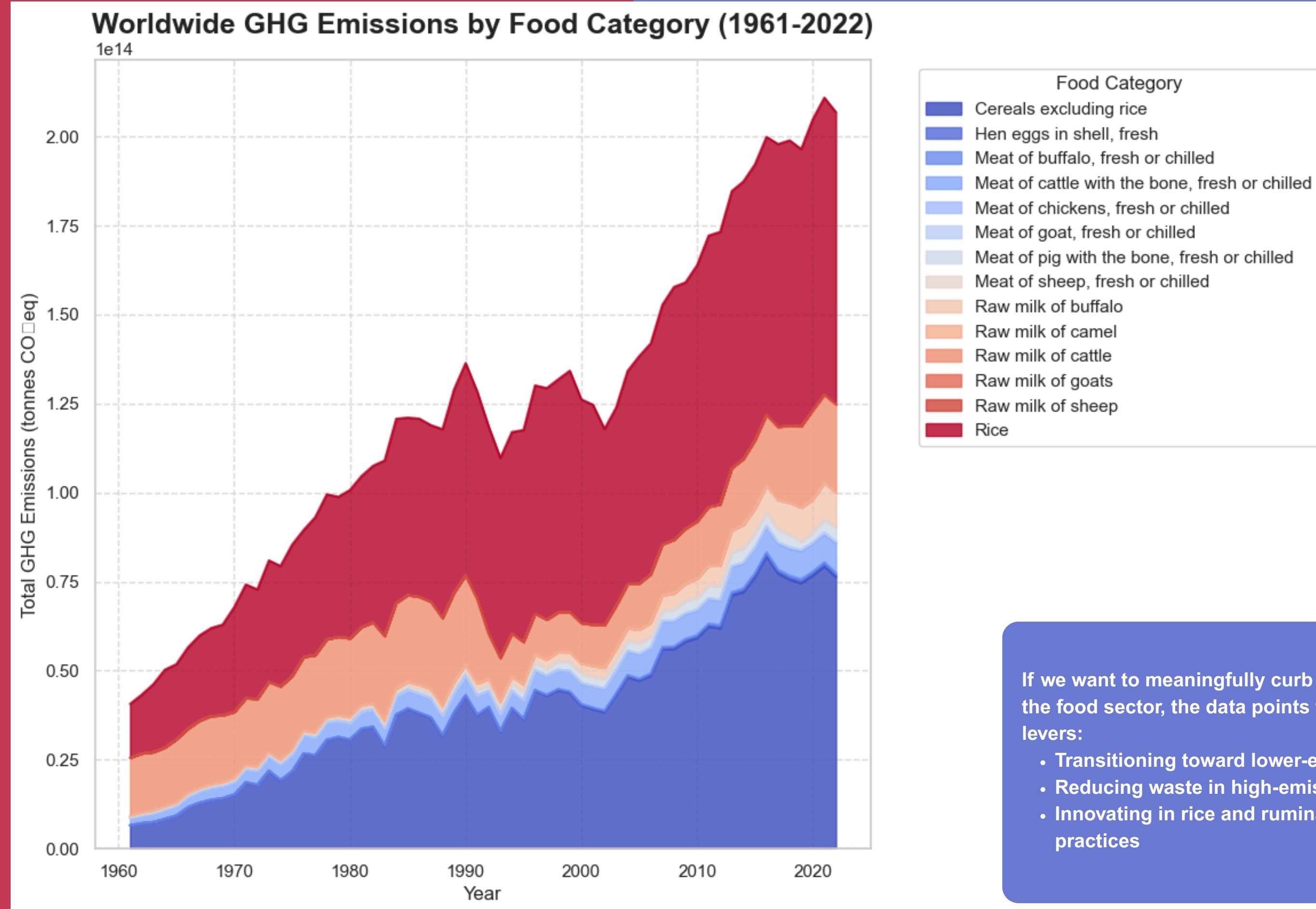
Are methane-heavy systems like rice or livestock dominant?  
Is there a lack of low-emission technology or land-use policy?

**Top 10 countries by avg GHG emissions per Food Production  
(1961 - 2022)**



Emissions are measured in carbon dioxide equivalents (CO<sub>2</sub>eq). This means non-CO<sub>2</sub> gases are weighted by the amount of warming they cause over a 100-year timescale.

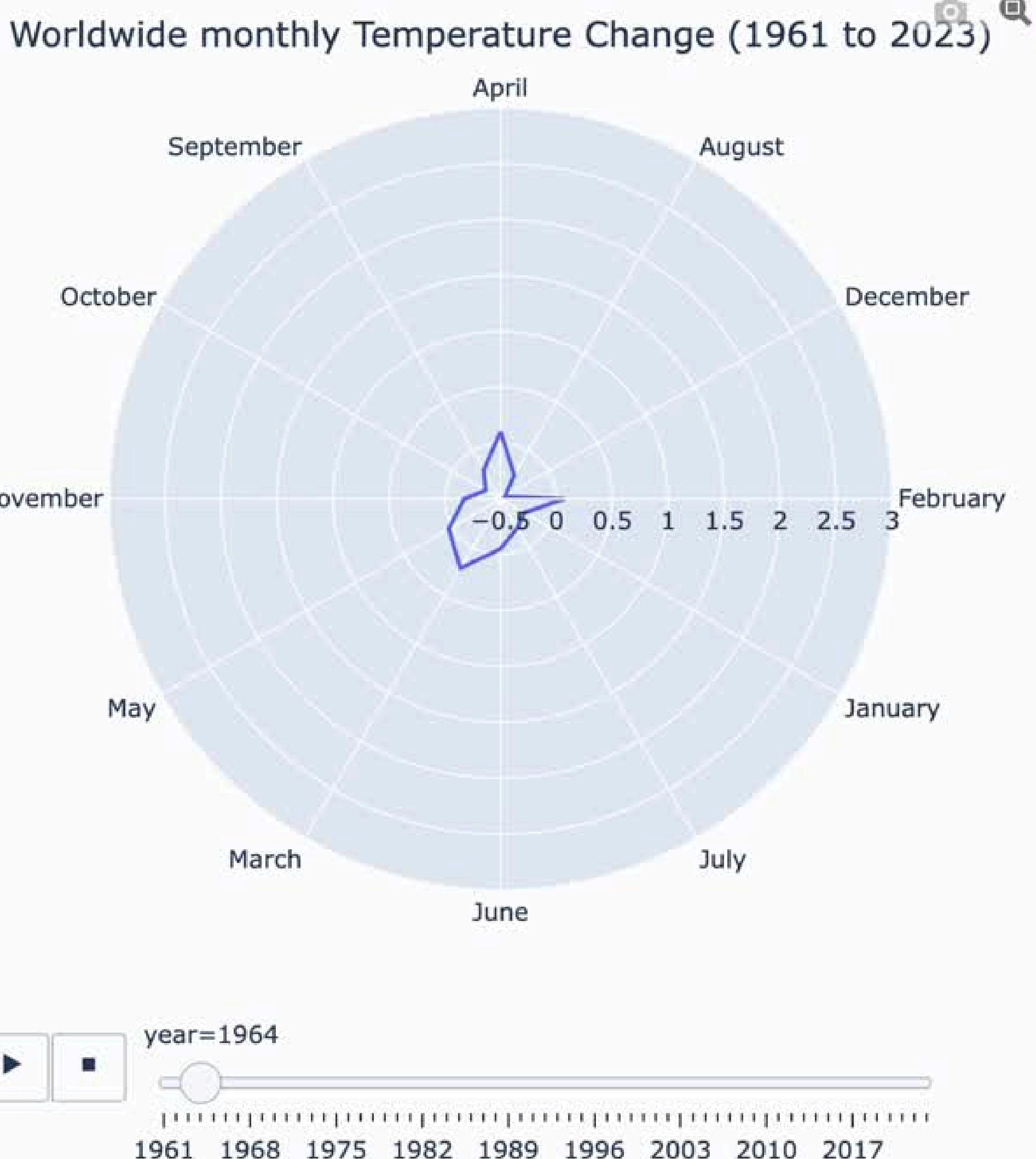
#### 4. Data Analysis: Which foods contribute the most to greenhouse gas emissions?



## II. Exploring Relationships

# Exploratory Data Analysis (applied statistics)

Is there a discernible relationship between food production, its environmental impact, and rising temperature anomalies?



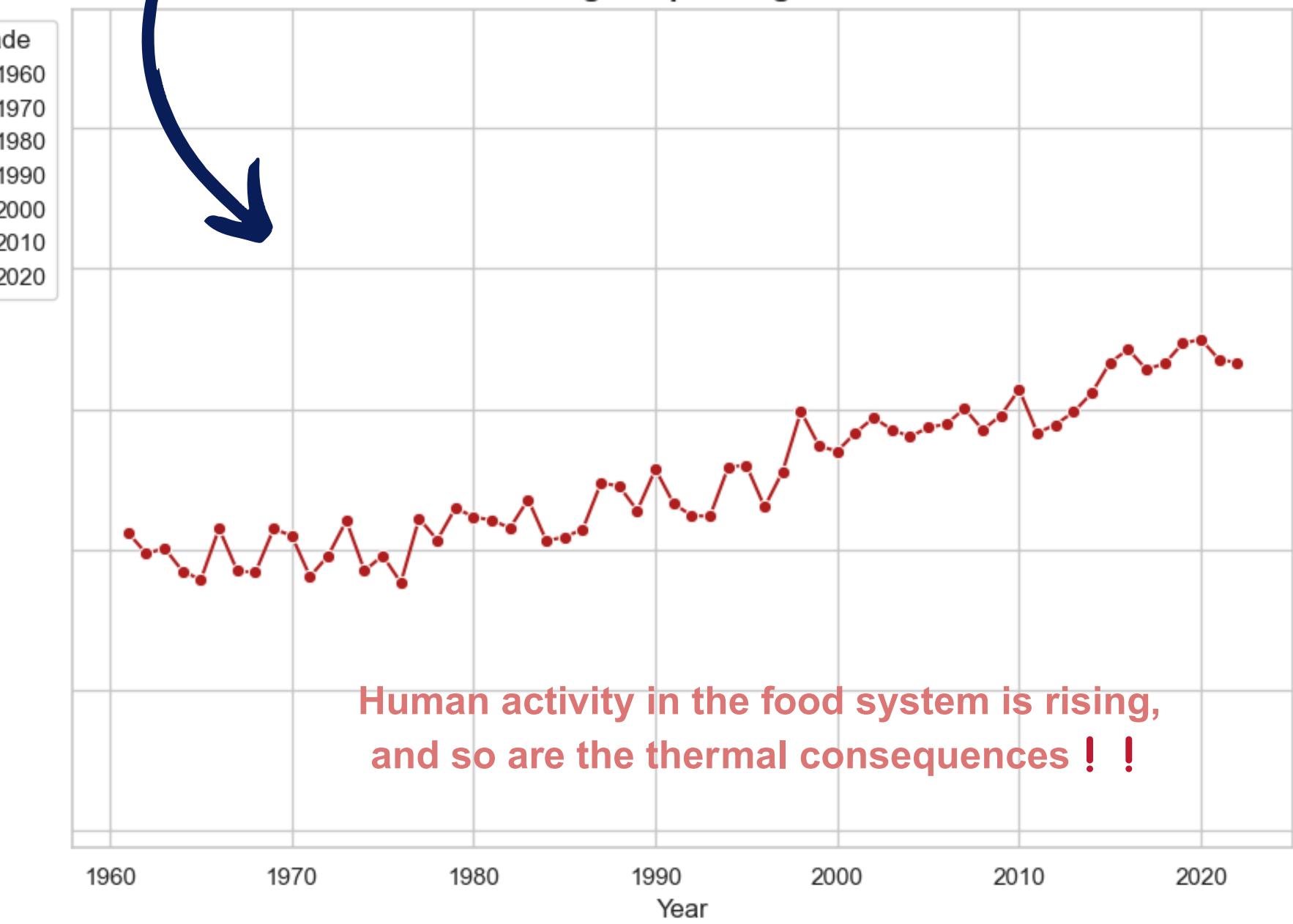
## Global Food System & Climate Change

A: Food Production vs Temp Change



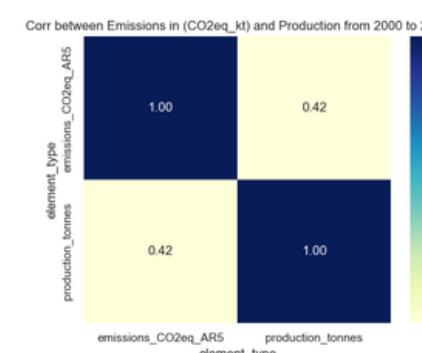
How has the global temperature varied over the years?

B: Global Avg Temp Change Over Time



Human activity in the food system is rising, and so are the thermal consequences ! !

- 1998 El Niño, 2016 warmest year\*



Corr between Emissions in CO<sub>2</sub>eq\_kt and food production (2000 - 2022) is 0.41

# III. Hypothesis Testing



# Hypothesis testing:

## Structured approach

📌 **Null Hypothesis ( $H_0$ ):** There is no significant difference between the mean emissions from meat & dairy products and other food categories.

🔴 **Alternative Hypothesis ( $H_1$ ):** Meat & dairy production generates significantly higher GHG emissions than plant-based food.

\* ✓ Z - Test.

```
z_statistic = -3.750  
p_value = 0.001350  
 $\alpha$  = 0.05
```

★ Reject the null hypothesis ! ★

I can confidently reject the idea that emissions are the same across food categories.



# Key Project Conclusions: Food Systems & Climate Impact

- Production Volume Drives Total Emissions.
- Emissions Intensity Varies by Food Type.
- Climate-Food Link is Complex.
- Outliers Reveal Critical Insights:
  - Data Anomaly: Extreme GHG values for Rice in Hong Kong SAR were identified as a data reporting error, requiring imputation.
  - Genuine Hotspot: High GHG from Raw Milk of Sheep in Russia reflects real, high-intensity production practices.
- Overall Implication: **Sustainable food policies !**

## Next steps:

- Analysis was limited by the available data's time range (1961-2022 for production/temperature) and the absence of per capita data. *Future work could involve:*
- Incorporating population data to enable per capita consumption and emissions analysis.
- Connect “Our World in Data API”(dataset) to the Jupyter notebook to explore GHG by food item and compare with the sample from food category.
- Group by countries with high production and vice versa to look up for outliers per category within each one of them.
  - And per category, do they produce more or less GHG per production in their own scale?
- \* Conducting more granular regional or climate zone-specific analyses to better capture the localized impacts of temperature change on agriculture.



thank  
you



# Data Engineering

## 📌 : Data Engineering:

- Python (Pandas),
- Kaggle APIs
- Tableau & Google Big Query (Google Cloud)
- GitHub Desktop to update the repository daily.
- ([Food-Climate-Data-Driven-Sustainability](#))

## 💻: Storage & Querying:

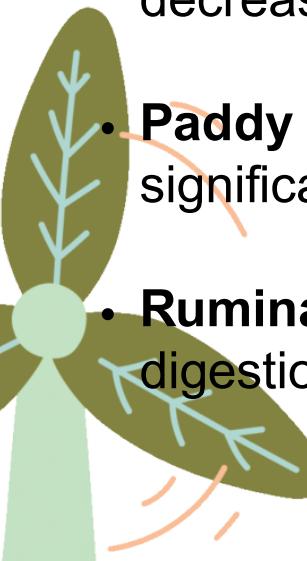
Google BigQuery and SQL

- project\_id = [food-climate-project-id](#)
  - Dataset: food\_climate\_data
    - Table: merge\_food\_climate
- Create tables with (PK-FK), join and create different relationships



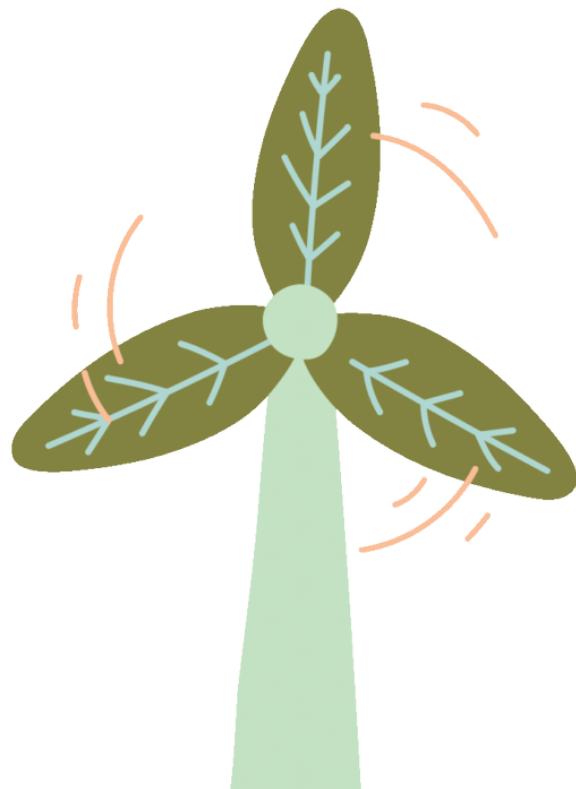
# Glossary

- **GHG Emissions (CO<sub>2</sub>eq):** Greenhouse Gas emissions, measured in Carbon Dioxide Equivalents. This standard unit accounts for the warming potential of different GHGs (like methane and nitrous oxide) over a 100-year period, converting them to an equivalent amount of CO<sub>2</sub>.
- **Emissions Intensity (GHG/kg):** The amount of greenhouse gas emissions (in CO<sub>2</sub>eq) generated per kilogram of a food product. This metric indicates the environmental efficiency or footprint of producing a specific food item or category.
- **Production Tonnes:** The total volume of food produced, measured in metric tonnes. This represents the absolute quantity of agricultural output.
- **Temperature Anomaly:** The deviation of a measured temperature (e.g., average annual temperature) from a historical baseline average. A positive anomaly indicates warmer-than-average conditions, and a negative anomaly indicates cooler-than-average conditions.
- **Food Category:** Broad groupings of food items (e.g., "Cereals & Grains," "Meat," "Dairy & Eggs") created through a custom mapping process to facilitate higher-level analysis of environmental impacts.
- **Outlier:** A data point that significantly deviates from other observations in a dataset. In this project, outliers were investigated to determine if they represented data errors or genuinely extreme (but valid) real-world phenomena.
- **Imputation:** The process of replacing missing or anomalous data values with substituted values. In this project, it was used to replace physically implausible GHG emissions data for Rice in Hong Kong SAR to ensure data integrity.
- **Correlation:** A statistical measure that describes the extent to which two variables tend to change together. A positive correlation means they increase or decrease in tandem, while a negative correlation means one increases as the other decreases.
- **Paddy Fields:** Flooded agricultural fields used for cultivating rice. These waterlogged conditions create anaerobic environments where microbes produce significant amounts of methane, a potent greenhouse gas.
- **Ruminant:** Mammals (such as cattle, sheep, goats, and buffalo) that have a specialised digestive system involving multiple stomach compartments. Their digestion process, particularly enteric fermentation, produces methane, contributing significantly to agricultural GHG emissions.



# References

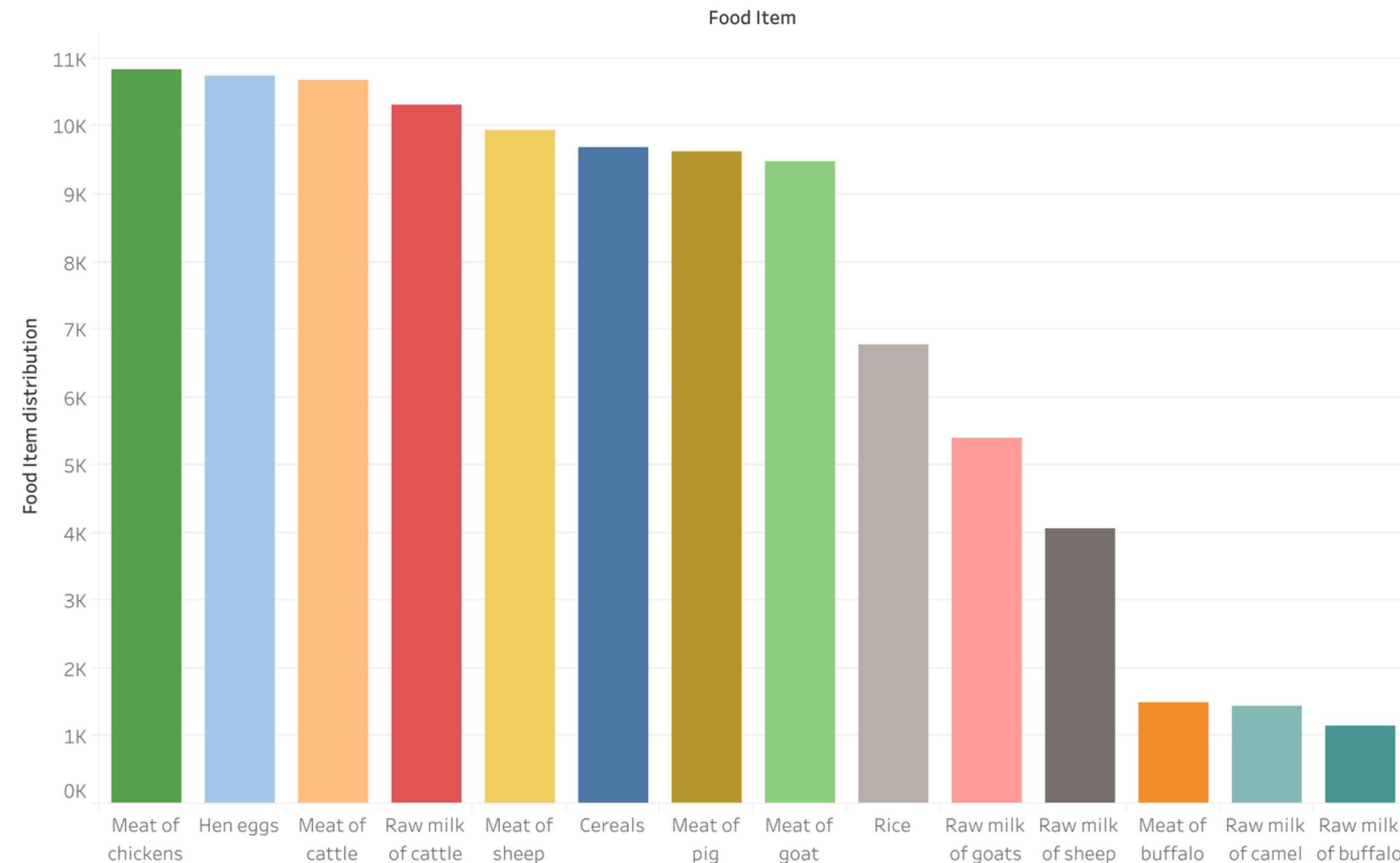
- FAO, 2023. FAOSTAT Climate Change : Agrifood systems emissions; Emissions intensities <http://www.fao.org/faostat/en/#data/EI>
- FAO, 2025. FAOSTAT Land, Inputs and Sustainability; Climate Change Indicators; Temperature change on land, available at: <http://www.fao.org/faostat/en/#data/ET>. FAO, Rome Italy.



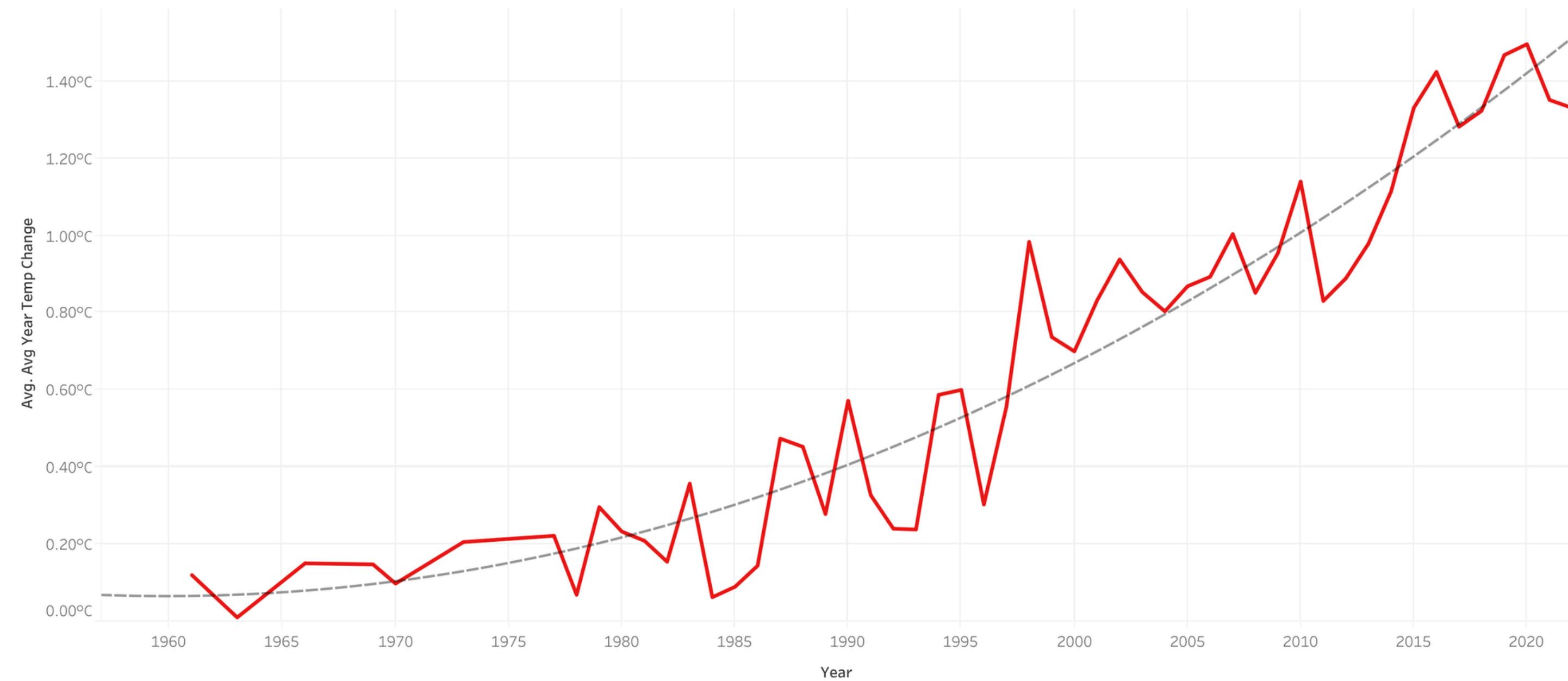
Othe viz



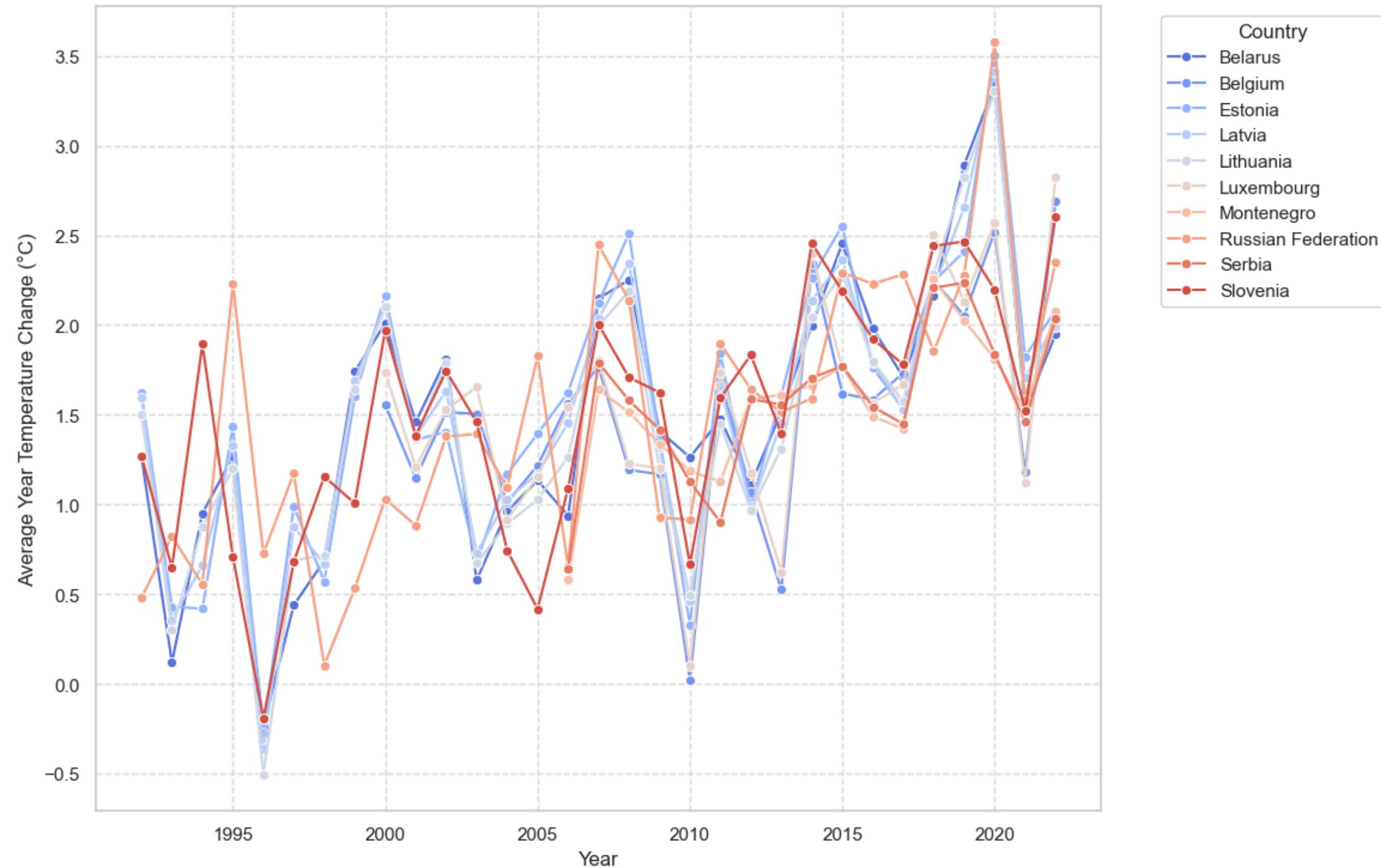
Food Product distribution globally  
(1961 - 2022)

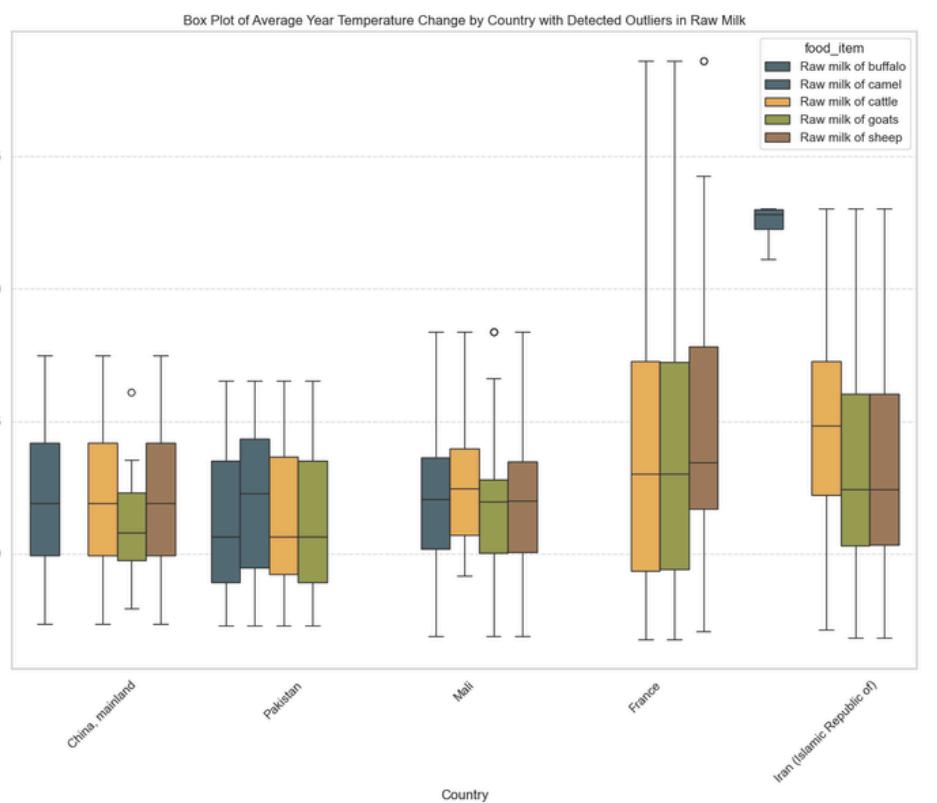
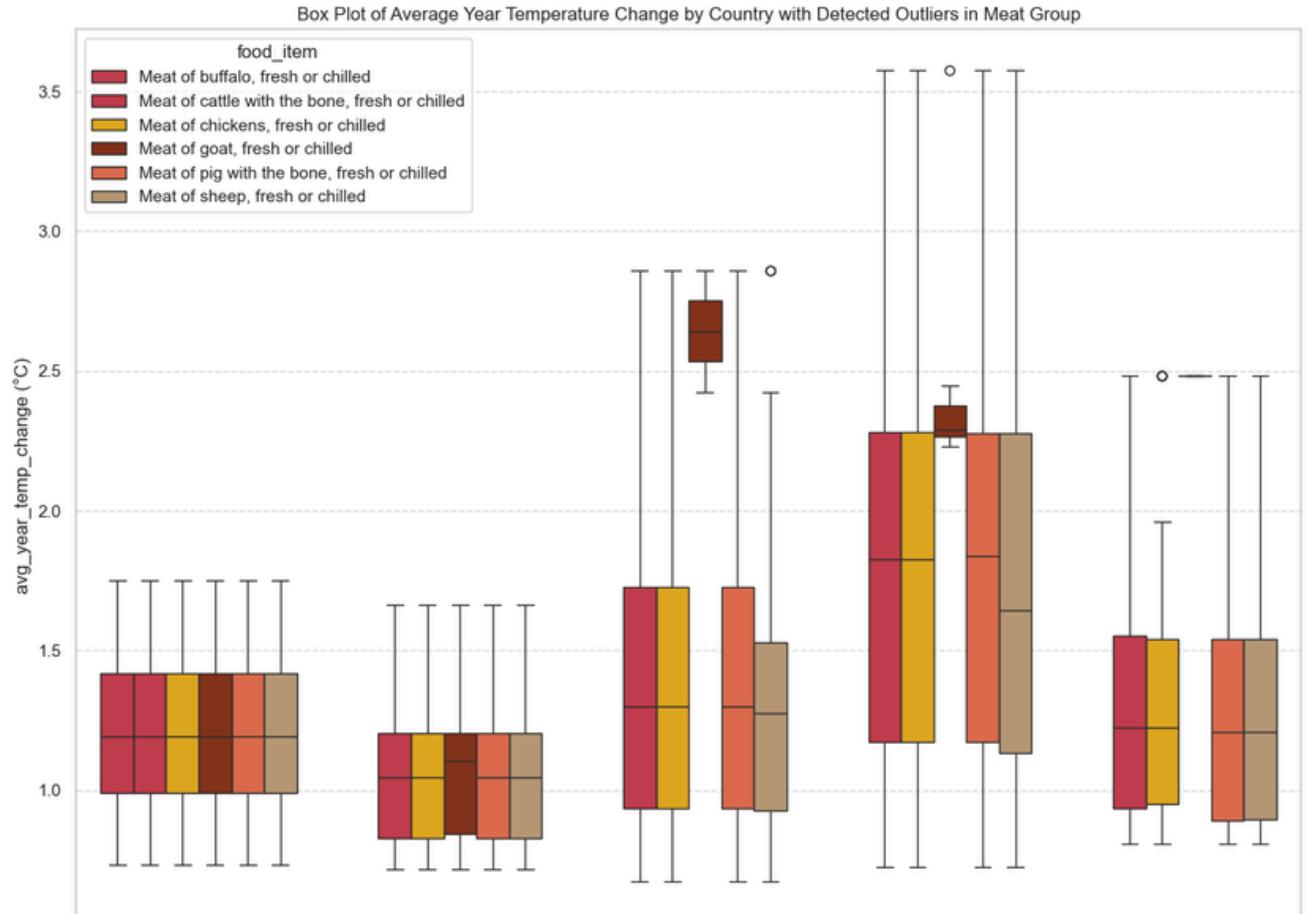
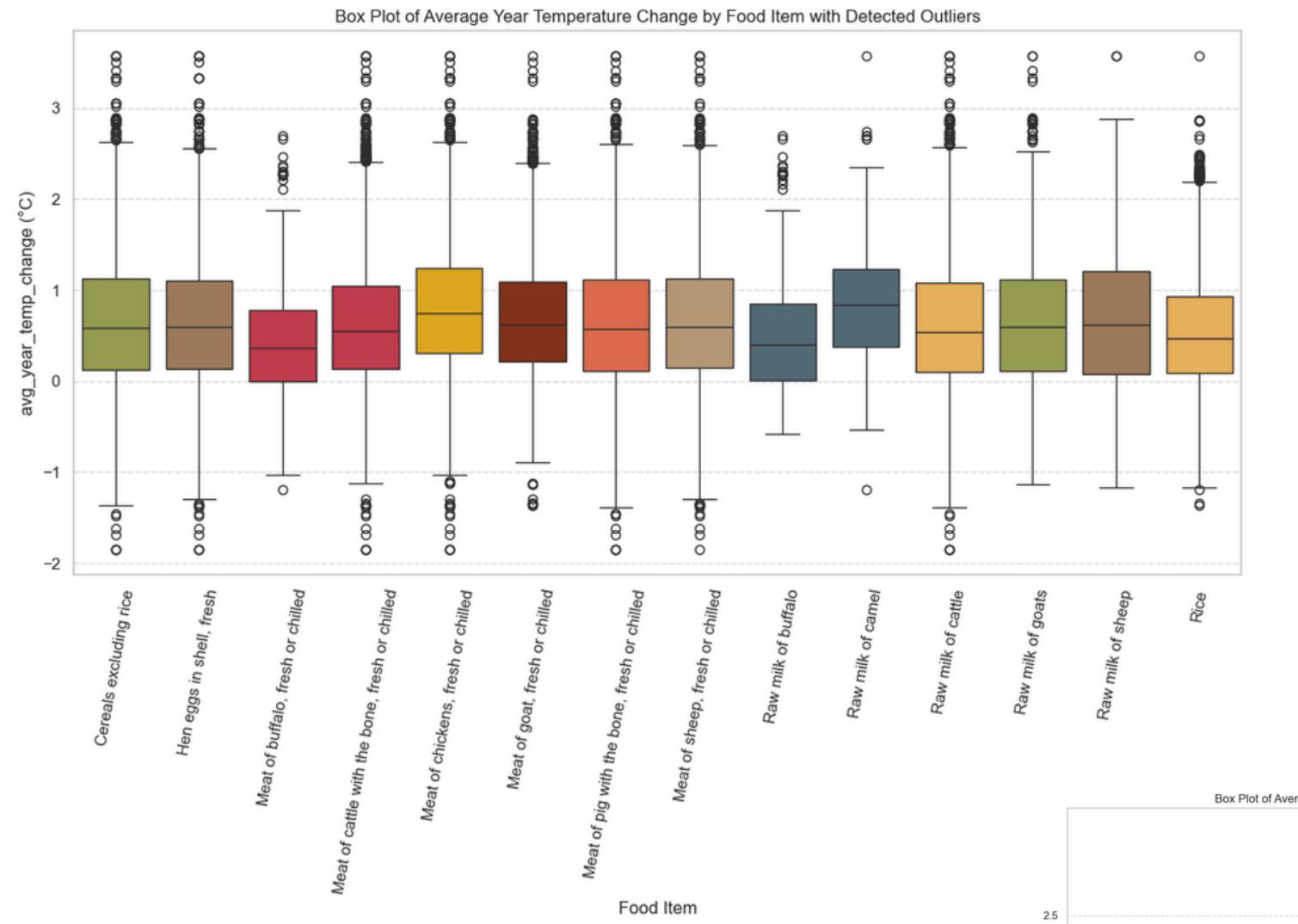


Average temperature change globally  
(1961 - 2022)

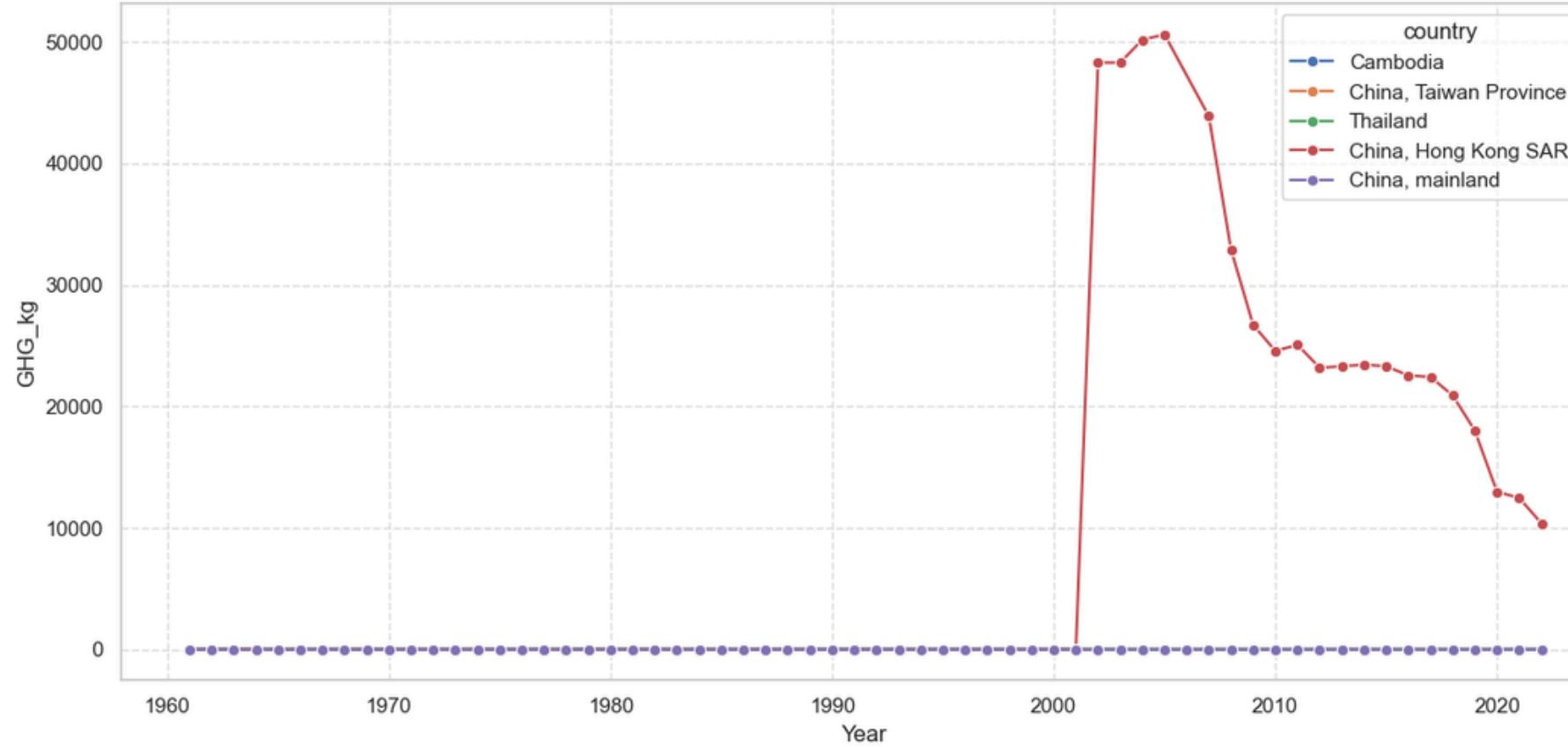


## Average Year Temperature Change over time for Top 10 Countries

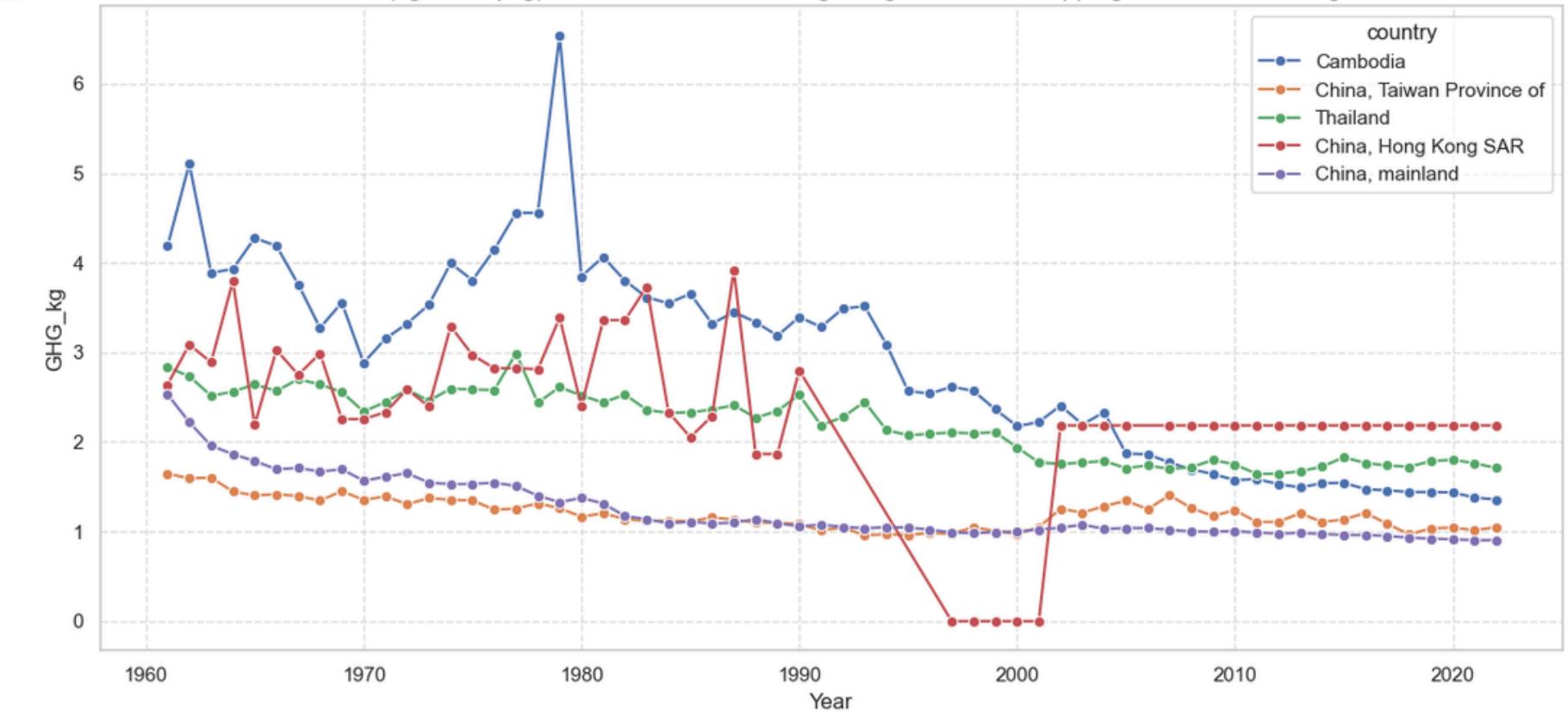




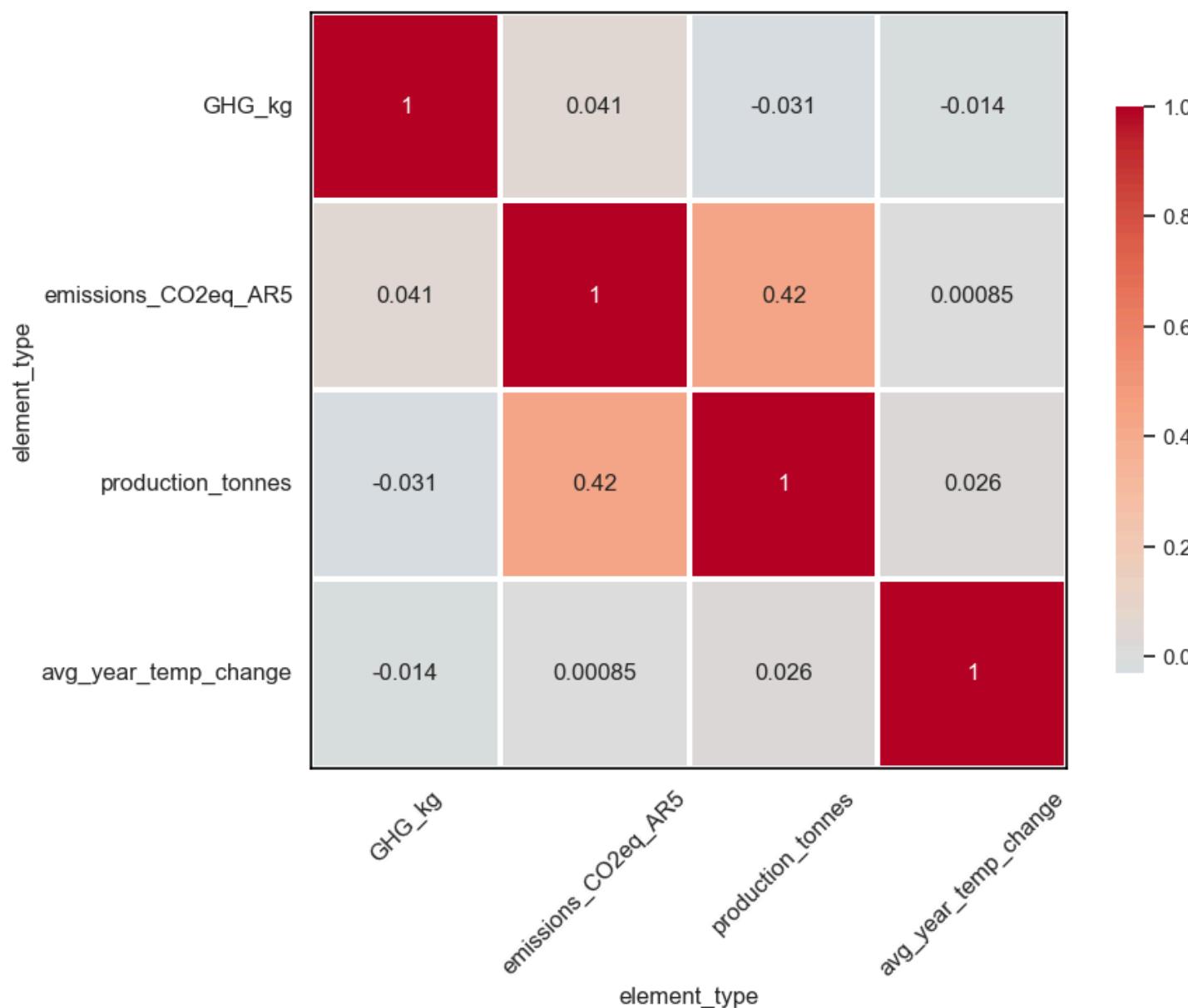
GHG Emissions (kgCO<sub>2</sub>eq/kg) for Rice in China & Hong Kong SAR Over Time



GHG Emissions (kgCO<sub>2</sub>eq/kg) for Rice in China & Hong Kong SAR after Dropping Outliers and Filling with Mean



**Corr Matrix per in Food Climate Data**



**Corr between Emissions in (CO2eq\_kt) and Production from 2000 to 2022**

