EEP 153 – Project 1

Population, Agriculture and the Green Revolution

Project Team
Emma Azhan
Elan Trager
Ploy Sripungwiwat
Xinyu Zhou
Xi Wang

Agenda

01 Research Overview

02 Code Walkthrough

03 Population Visualisations

04 Agriculture Visualisations

05 Synthesis

Background Information

The Green Revolution (1940s – 1970s) refers to a period of agricultural transformation driven by **the adoption of high-yield crop varieties, improved irrigation systems, chemical fertilizers, pesticides, and modern farming techniques.**

In Asia, this technological change significantly impacted nations like India, Pakistan, and the Philippines, where staple crops such as rice and wheat were central to diets. Key innovations included the introduction of high-yielding varieties (HYVs), such as IR8 rice, and significant investments in agricultural infrastructure.

Research Question

To what extent did the Green Revolution impact population growth and agricultural production across Asia?

Metrics of consideration

Population

- Total Population
- Male versus Female Population
- Urban versus Rural Population

Agriculture

- Cereal Production Index
- Food Production Index
- Crop Production Index

Key Motivation

How is agricultural efficiency related to population growth?

Rooted in the Malthusian Problem, our team is curious to discover how implicitly agricultural production impacts population growth and development. Particularly, when considering technological advances related to production efficiency, we hope to understand the correlation between the rate of population growth in different sectors of society and major events of technological development (within the Green Revolution).

Regions of Interest

South Asia (SA)	Mainland Southeast Asia (SEA)	Maritime Southeast Asia (MSEA)	East Asia (EA)
Bangladesh	Cambodia	Brunei Darussalam	China
Bhutan	Lao People's Democratic Republic	Indonesia	Democratic People's Republic of Korea
India	Myanmar	Malaysia	Mongolia
Maldives	Thailand	Papua New Guinea	Republic of Korea
Nepal	Viet Nam	Philippines	
Pakistan		Timor-Leste	
Sri Lanka			

02

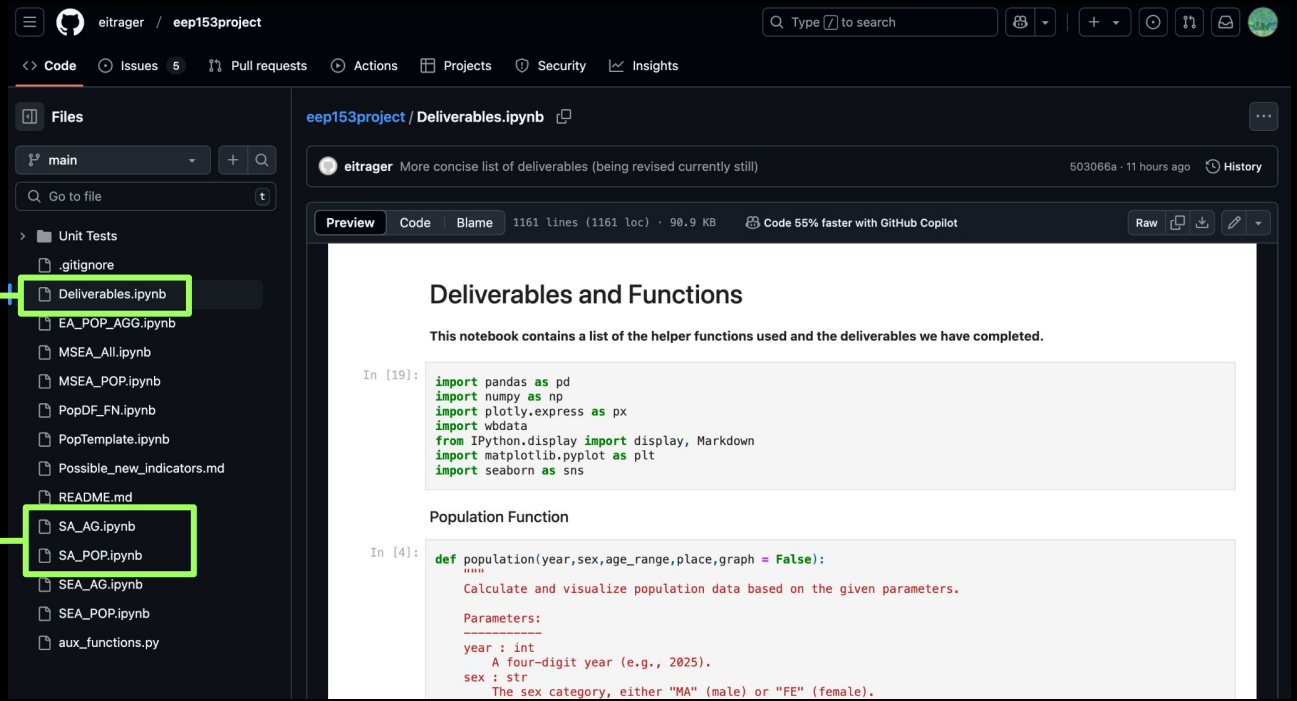
Our Code

Our Repo

We utilised GitHub as the primary platform to collaborate and store our code. Our repo houses files addressing both foundational deliverables such as the Population Function, as well as auxiliary visualisations and analysis which speak to our research question.

Core deliverables and functions consolidated in Jupyter Notebooks.

Population functions and data frames are pulled into other Notebooks to produce region-specific visualisations and analysis for population and agriculture growth.



Population Function

While developing the Population Function, we were able to produce helper functions such as age_ranges that were helpful in later analyses.

Additionally, the final function evolved to be a powerful tool in exploratory data analysis.

Population Function

In [4]:

```
def population(year,sex,age_range,place,graph = False):
    """
    Calculate and visualize population data based on the given parameters.

    Parameters:
    -----------
    year : int
        A four-digit year (e.g., 2025).
    sex : str
        The sex category, either "MA" (male) or "FE" (female).
    age_range : list of int
        A list of two integers specifying the min and max age (e.g., [5, 20]).
    place : str
        A 3-character country code (e.g., "CHN" for China).
    graph : bool, optional
        Whether to generate a line plot (default is False).

    Returns:
    --------
    int
        The total population within the specified age range.

    Example:
    --------
    population(2021, "FE", [7, 23], "JPN", graph=True)
    """
    first_term = "SP"
    second_term = "POP"
    if sex == "MA":
        age_ranges = generateageranges(first_term, second_term, age_range, "MA", place, f"{year}-01-01")
        agearrayspecific, popvalues = interpprep(age_ranges)
        age_array_all_years = interpfunc(popvalues, agearrayspecific, place, year, graph_values = graph)
        sliced_popvals = age_array_all_years[age_range[0]:age_range[1]]
        finaloutput = np.sum(sliced_popvals)
        if finaloutput<2000:
            return "No data for this year"
        return finaloutput
    else:
        age_ranges = generateageranges(first_term, second_term, age_range, "FE", place, f"{year}-01-01")
        agearrayspecific, popvalues = interpprep(age_ranges)
```

Population DataFrame Function

Another core aspect of our analysis was the Population DataFrame Function deliverable.

Variations of this function, and the output it generates, were utilised in every iteration of our regional analysis.

In [26]:

```python
def get_population_by_age_sex(countries):
    """
    Fetches population data by age and sex from the World Bank (source 40)
    and returns a pandas DataFrame indexed by Country/Region and Year.

    :param countries: List of country ISO codes or 'all' for all available.
    :param years: List of years as integers.
    :return: Pandas DataFrame with population counts by age and sex.
    """
    # Define indicators for population by age and sex (source 40)
    indicators = {
        "SP.POP.0004.MA": "Male 0-4",
        "SP.POP.0004.FE": "Female 0-4",
        "SP.POP.0509.MA": "Male 5-9",
        "SP.POP.0509.FE": "Female 5-9",
        "SP.POP.1014.MA": "Male 10-14",
        "SP.POP.1014.FE": "Female 10-14",
        "SP.POP.1519.MA": "Male 15-19",
        "SP.POP.1519.FE": "Female 15-19",
        "SP.POP.2024.MA": "Male 20-24",
        "SP.POP.2024.FE": "Female 20-24",
        "SP.POP.2529.MA": "Male 25-29",
        "SP.POP.2529.FE": "Female 25-29",
        "SP.POP.3034.MA": "Male 30-34",
        "SP.POP.3034.FE": "Female 30-34",
        "SP.POP.3539.MA": "Male 35-39",
        "SP.POP.3539.FE": "Female 35-39",
        "SP.POP.4044.MA": "Male 40-44",
        "SP.POP.4044.FE": "Female 40-44",
        "SP.POP.4549.MA": "Male 45-49",
        "SP.POP.4549.FE": "Female 45-49",
        "SP.POP.5054.MA": "Male 50-54",
```

Unit Tests

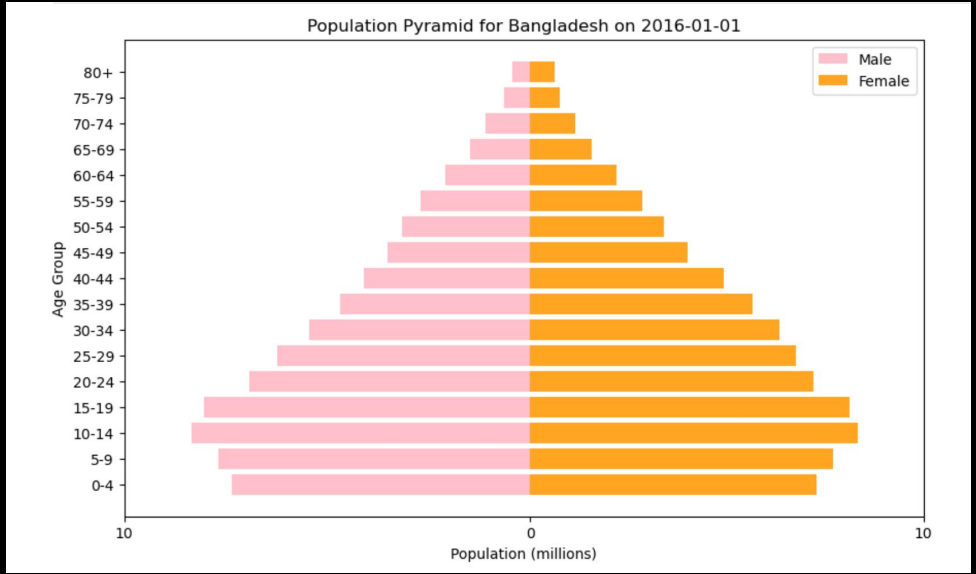
Unit Tests, while another mandatory deliverable, were extremely important in helping frame our thought processes while exploring and analysing data. Particularly, unit tests were helpful in gauging logical ranges for data values, note key outliers, and push us to consider external contexts/events that would explain abnormalities.

```python
import unittest
import numpy as np
from Deliverable import population

class TestPopulationFunction(unittest.TestCase):

    def test_valid_population_male(self):
        """Test valid population retrieval for males."""
        result = population(2025, "MA", [10, 20], "USA")
        self.assertIsInstance(result, int)  # Expected output should be an integer

    def test_valid_population_female(self):
        """Test valid population retrieval for females."""
        result = population(2025, "FE", [5, 18], "JPN")
        self.assertIsInstance(result, int)  # Expected output should be an integer

    def test_no_data_available(self):
        """Test case where no data is available (returning a string)."""
        result = population(1800, "MA", [0, 10], "CHN")  # Assuming data does not exist for 1800
        self.assertEqual(result, "No data for this year")

    def test_invalid_sex_input(self):
        """Test invalid sex input (should raise an error)."""
        with self.assertRaises(ValueError):
            population(2025, "INVALID", [10, 20], "USA")

    def test_invalid_year_input(self):
        """Test invalid year input (should raise an error)."""
        with self.assertRaises(TypeError):
            population("year", "MA", [5, 20], "USA")  # Year should be an integer

```

Population Pyramids

The function producing Population Pyramids called upon many similar, if not the same skills as the DataFrame and general Population function.

The visualisations produced by the defined function were extremely helpful in preliminary data analysis and exploration, ensuring we understood the scale and distribution of each nation’s population.



(A and B) Deliverables

Our core deliverables served as building blocks for our analysis.

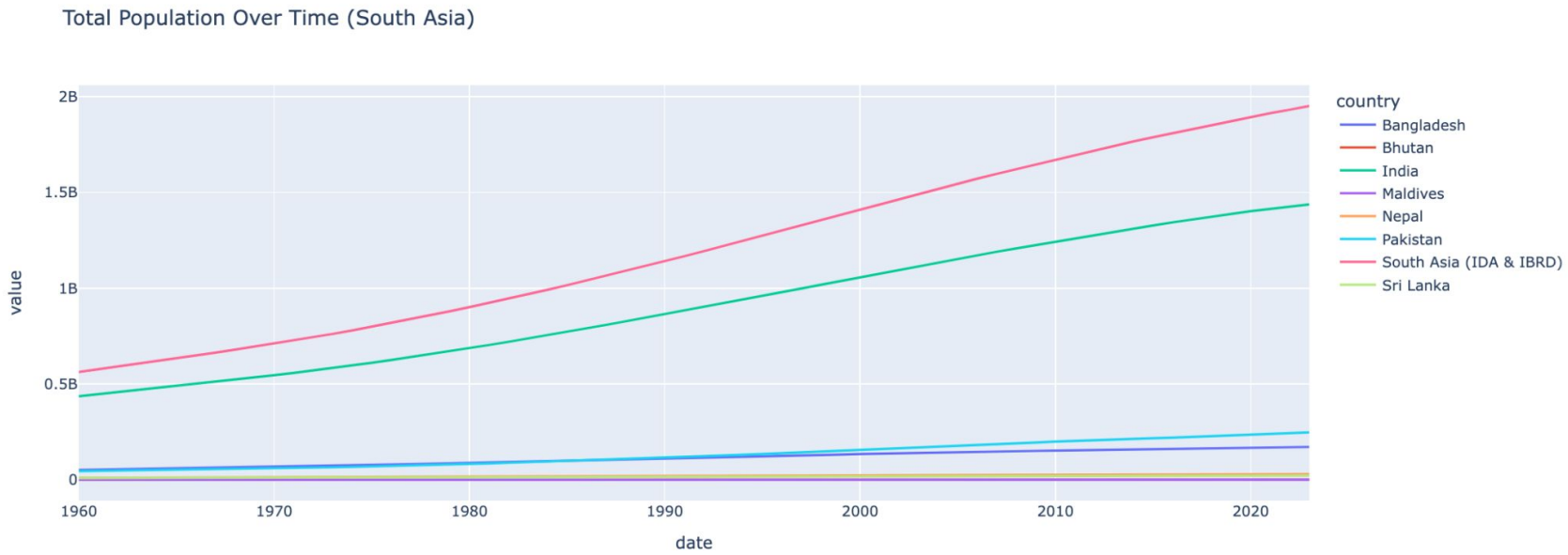
The tools developed for our A and B-Level deliverables allowed our team to better understand the data we were working with, and begin thinking about integrating different functions to produce our desired visualisations.

03

Population Visualisations

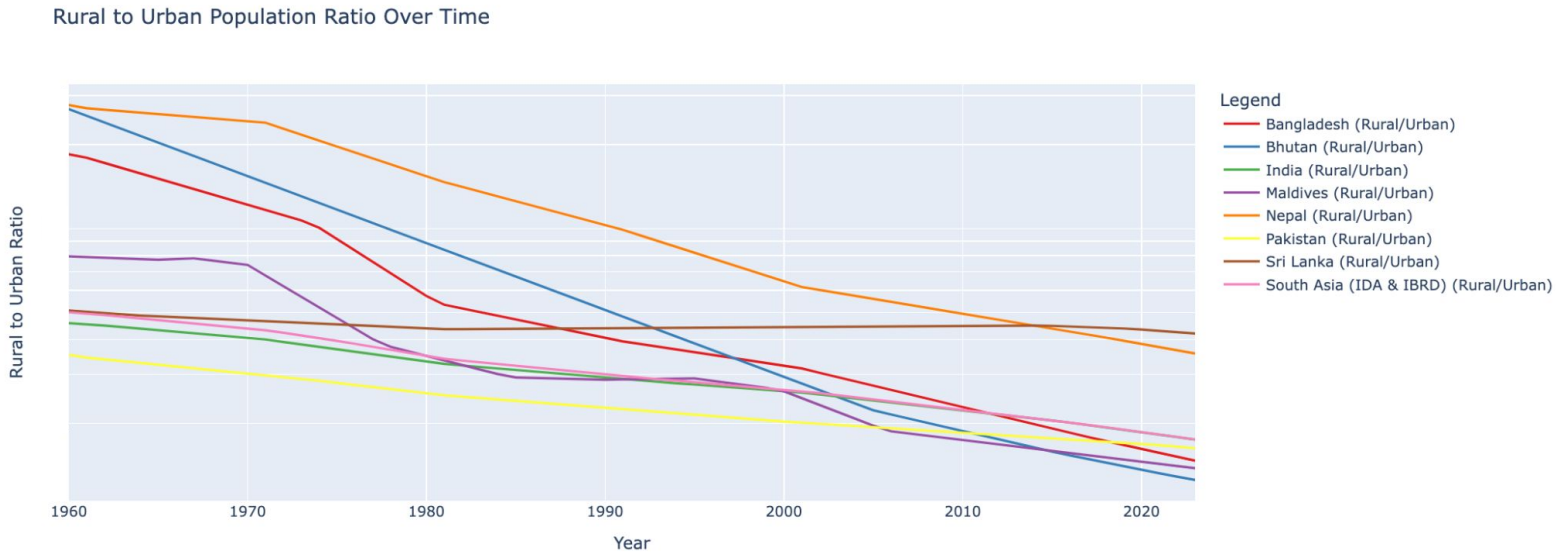
How has total population grown across Asia? Do we see trends of movement from rural to urban population? What context does this give us to interpret the timeline of the Green Revolution?

South Asia



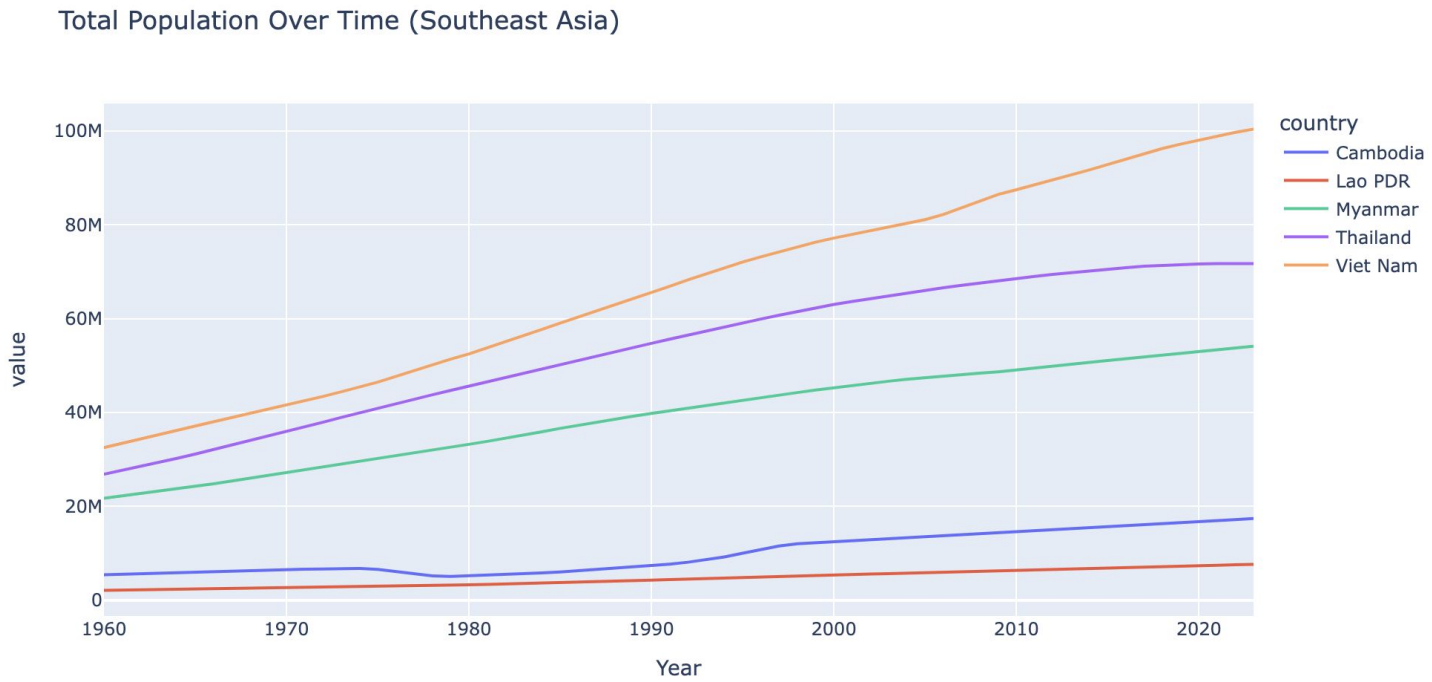
Overall trends see total population steadily increasing across South Asia with India contributing the largest proportion to total regional population.

South Asia



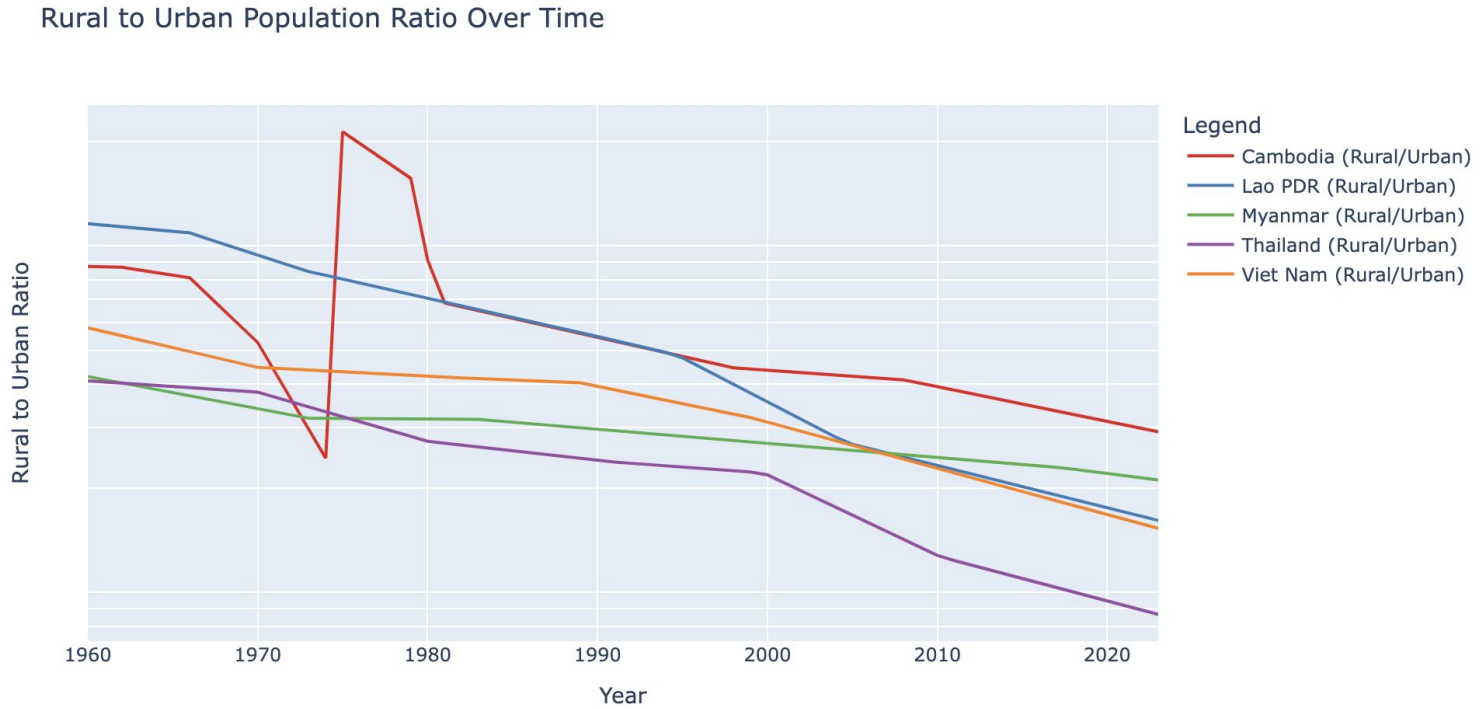
Over time we see that the ratio of rural to urban population decreases across South Asia, indicating that fewer individuals are involved in rural industries such as agriculture; even while noticing that overall production is increasing. This indicates that agriculture is becoming more efficient, coinciding with innovations from the Green Revolution.

Southeast Asia



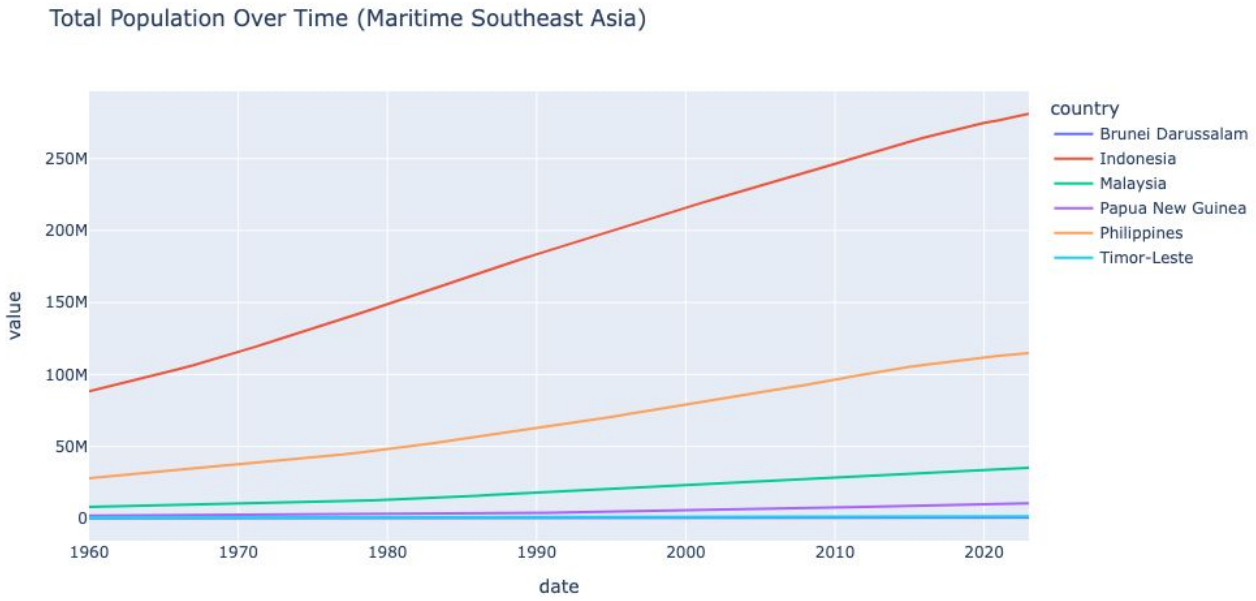
SouthEast Asian countries saw an overall positive trend for total population over time with Viet Nam contributing the most. A noticeable anomaly is Cambodia, which saw a decrease from mid 1970s to mid 1990s. This potentially can be attributed to the Khmer Rouge Regime (1975–1979), responsible for the Cambodian Genocide (1975).

Southeast Asia



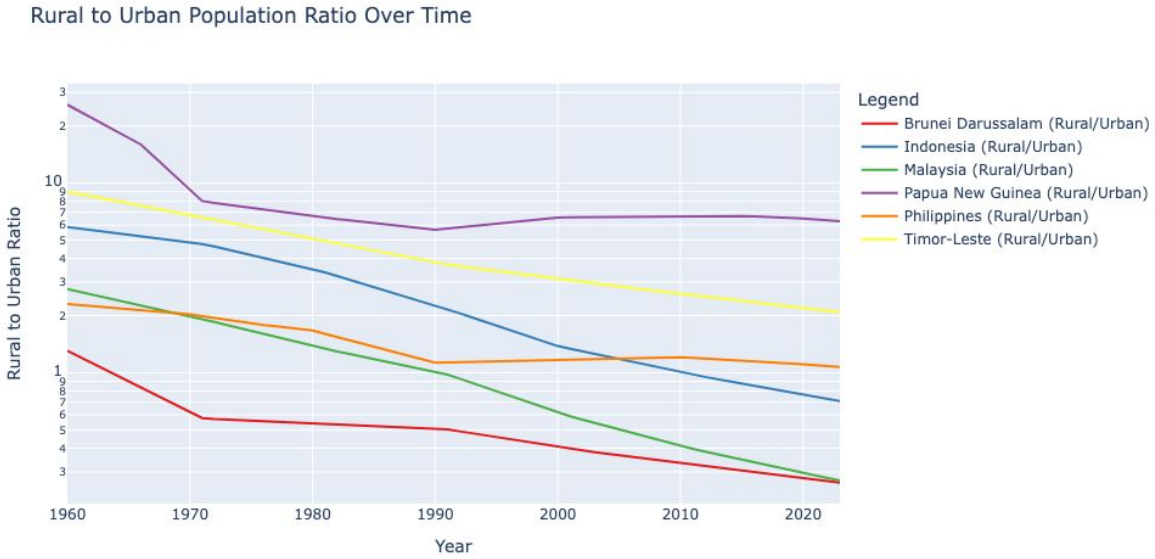
The general downward trend in rural to urban ratio for the SouthEast Asian countries follow the broader trend of rural to urban migration resulting from economic development in the region.

Maritime Southeast Asia



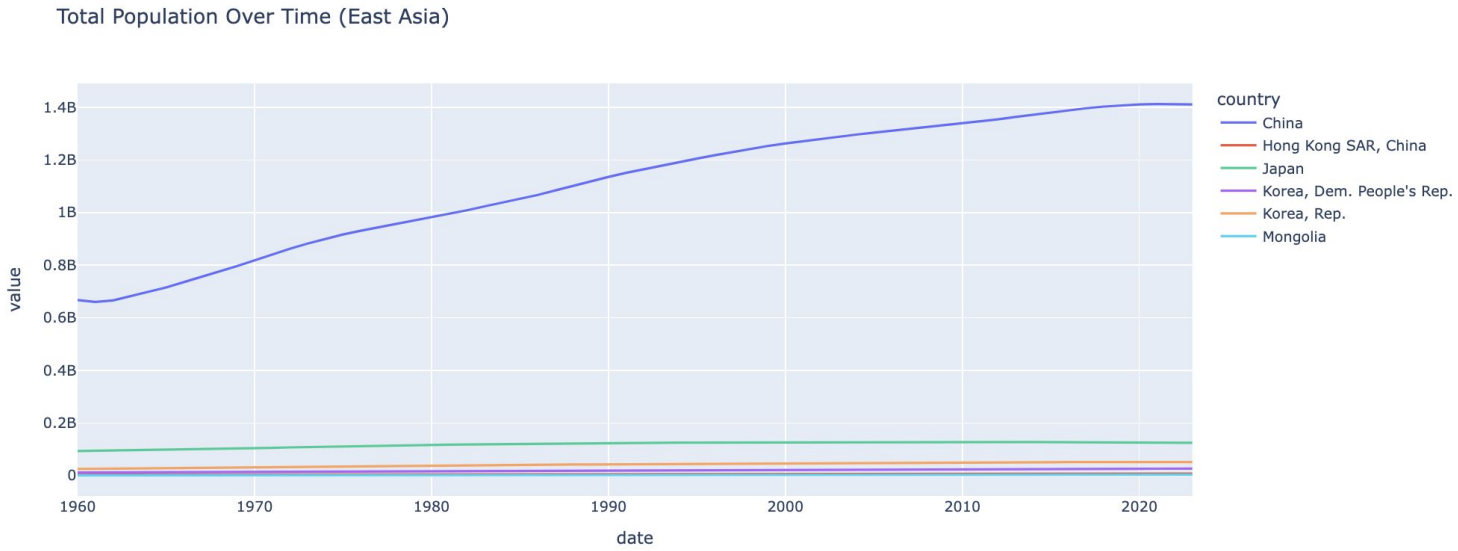
Indonesia and the Philippines have experienced significant population growth over time, with Indonesia leading by a wide margin, while other countries in the region show relatively smaller but steady increases.

Maritime Southeast Asia



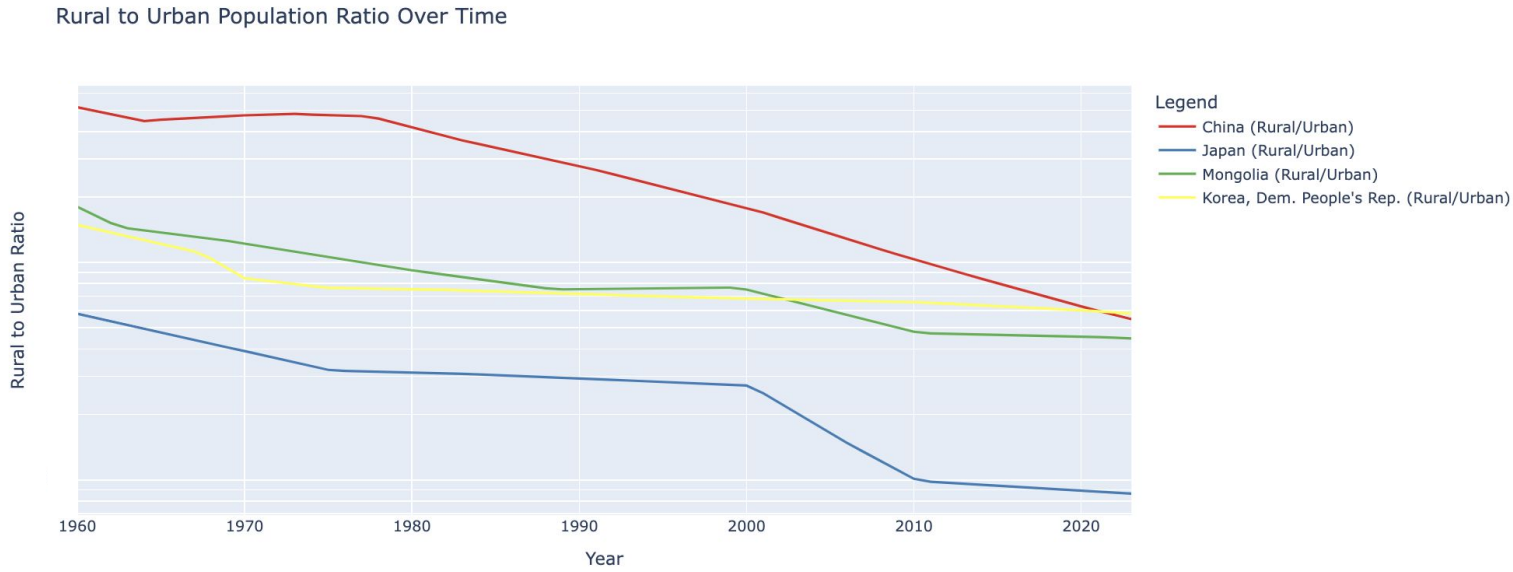
The decline in rural populations across all six countries reflects the broader trend of decreasing rural-to-urban population ratios in Maritime Southeast Asia, emphasizing the ongoing shift towards urbanization in the region.

East Asia



China has experienced significant population growth over time, compared to nearly flat growth for all other countries within this region. China's growth can be attributed to a range of factors such as high fertility rates, low mortality rates, and the increased food security created through the green revolution.

East Asia



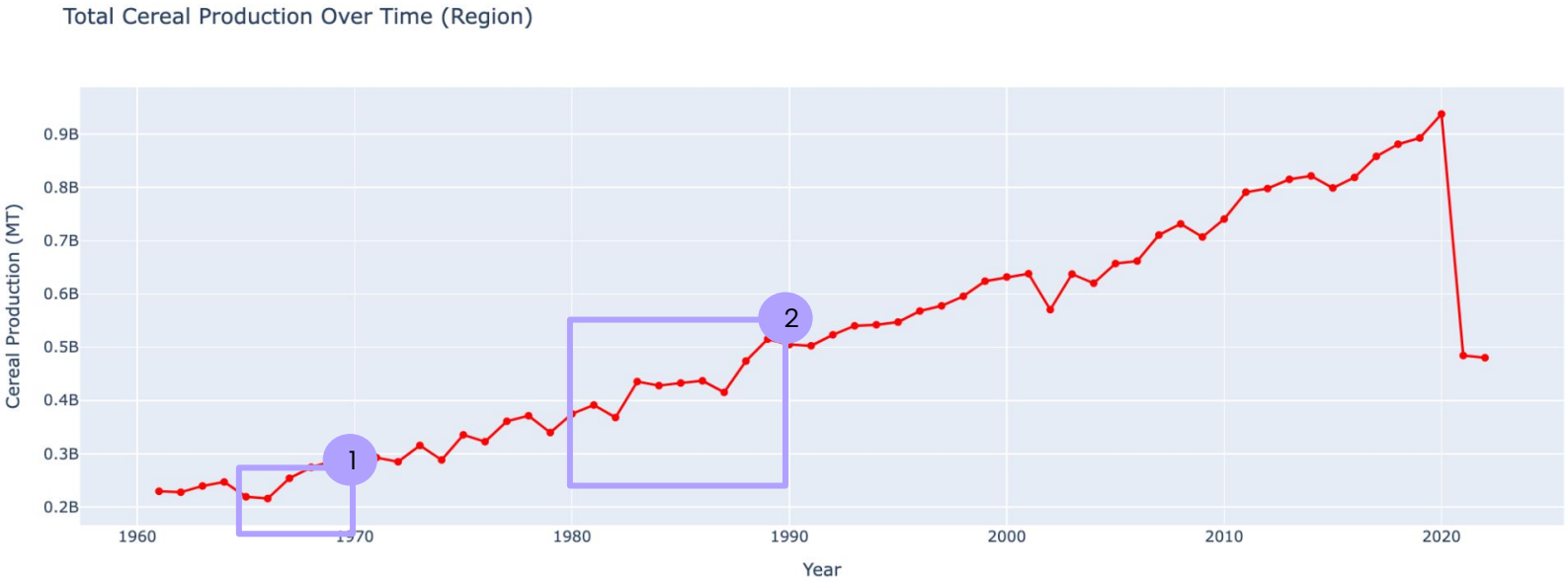
All countries outside of Korea in the above graph experience significant decline in the ratio of Urban to Rural populations, indicating development over time and correspondingly high rates of urbanization.

04

Agricultural Visualisations

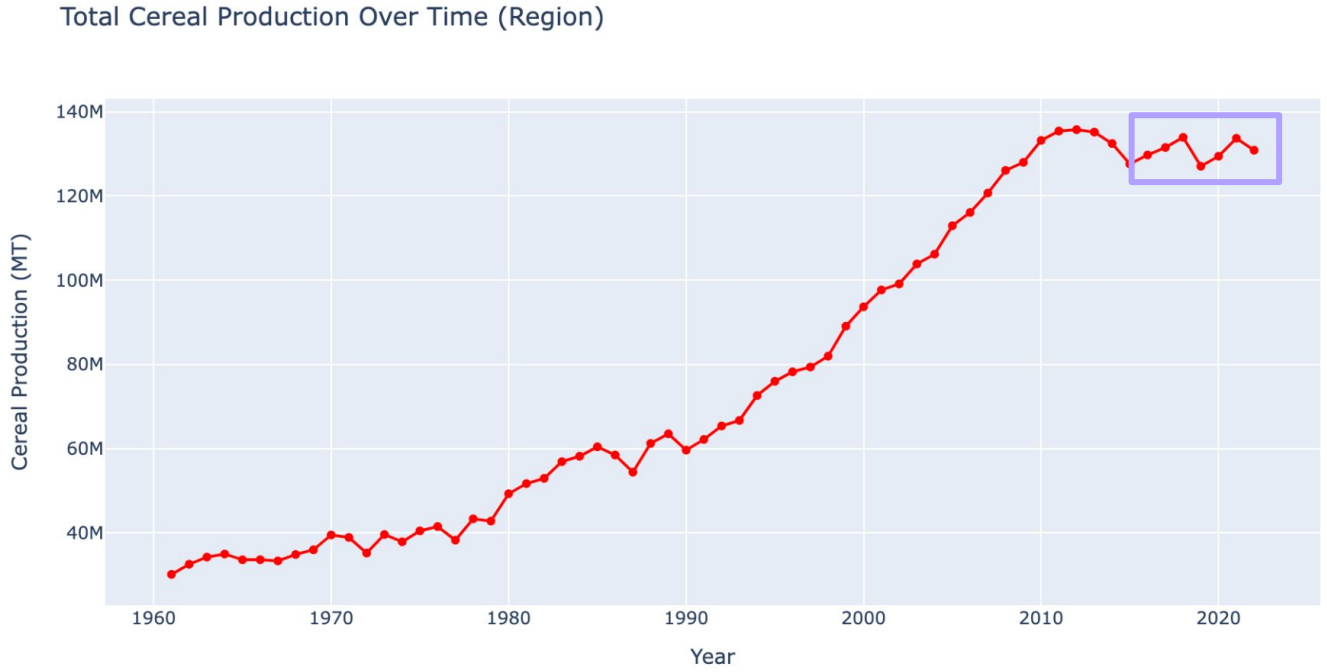
How has agricultural production changed across Asia? Do these changes in cereal production coincide with important Green Revolution developments?

South Asia



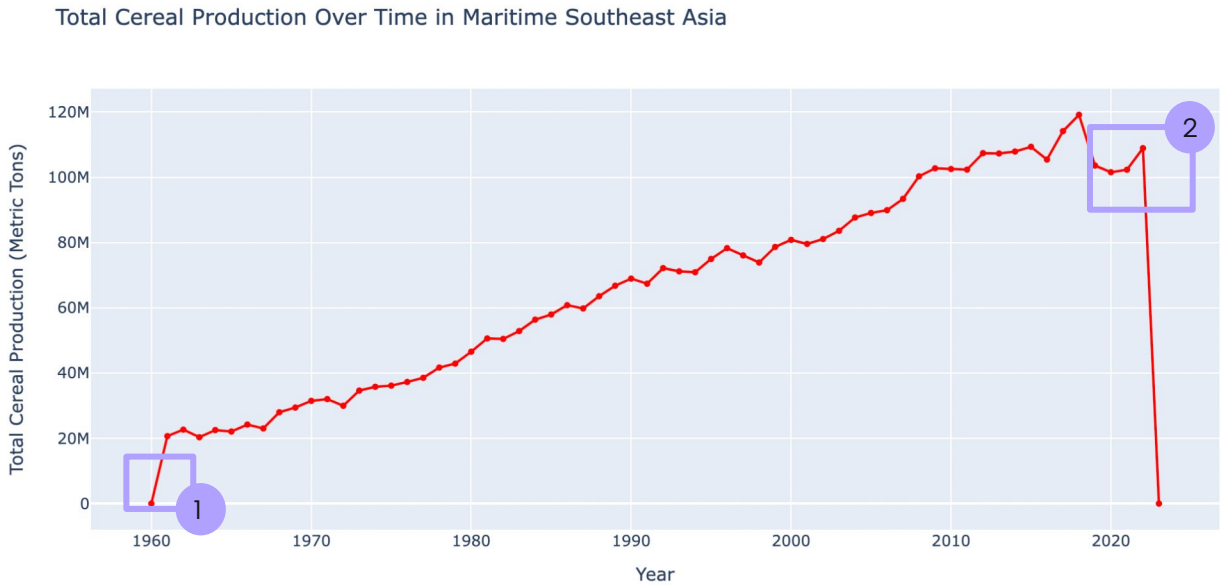
Total cereal production across South Asia seems to be steadily increasing overtime, until 2020 where we notice a sharp drop in total production. This coincides with the COVID-19 pandemic, but may also be attributed to gaps in the World Bank database. Key periods of growth include the **1965-1966 (1)** when India and Pakistan adopted high yield varieties of wheat, leading to record harvest and reducing food shortages. Additionally, sharp growth in the **1980s (2)** coincides with the spread of Green Revolution technology to Bangladesh.

Southeast Asia



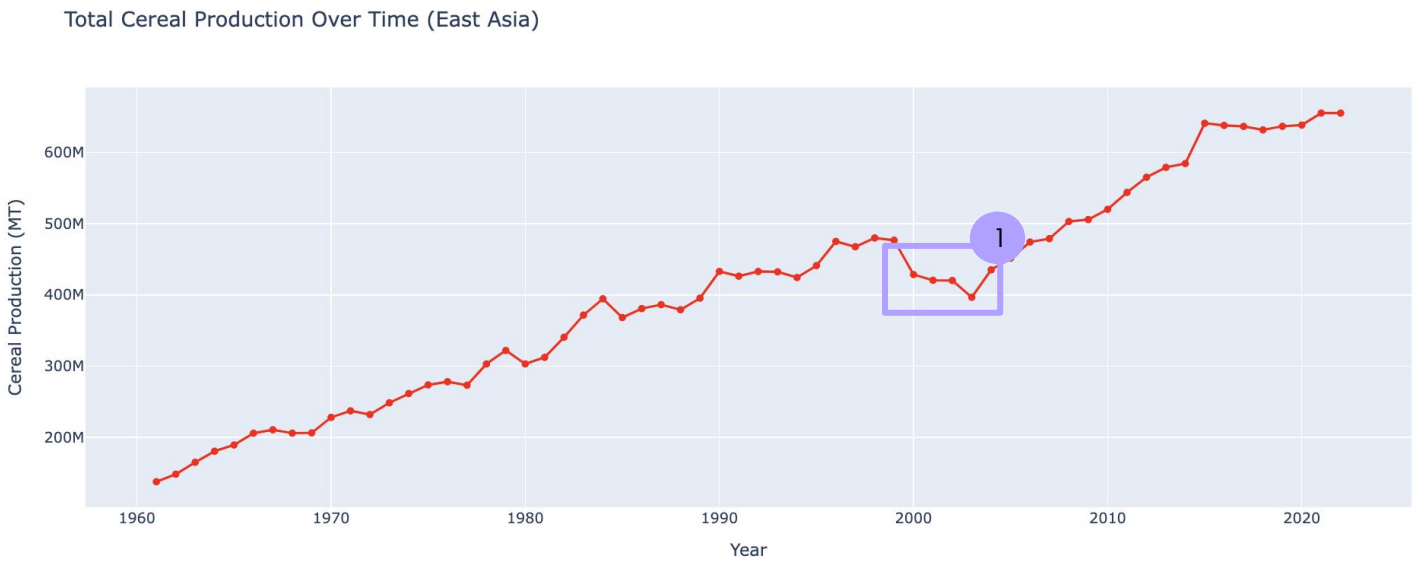
The total cereal production in SouthEast Asian countries see an overall positive trend that accelerated in the 1990s. The region prior to 1990s were relatively marred by political instabilities. This peaked in the mid 2010 and plateau before a fluctuation is seen after. This can be attributed to the periods of droughts, floods, and regional famine that greatly influence crop yields in SouthEast Asian countries.

Maritime Southeast Asia



Total cereal production in Maritime Southeast Asia increased over time from 1960 to 2020. The significant rise in production starting in the **1960s** is linked to the Green Revolution, which introduced high-yielding varieties and improved farming methods. However, the sharp drop in **2020** is likely due to the COVID-19 pandemic, which disrupted supply chains, caused labor shortages, and limited access to key agricultural resources, resulting in a decline in production.

East Asia



This graph demonstrates a greatly increasing cereal production over time, with some significant dips around 2000. These decreases can be attributed significantly to disastrous droughts, flooding, and a range of other environmental impacts experienced in China during this period. A primary event responsible was the 1988 China Floods**(1)**.

Sources: OCHA

05

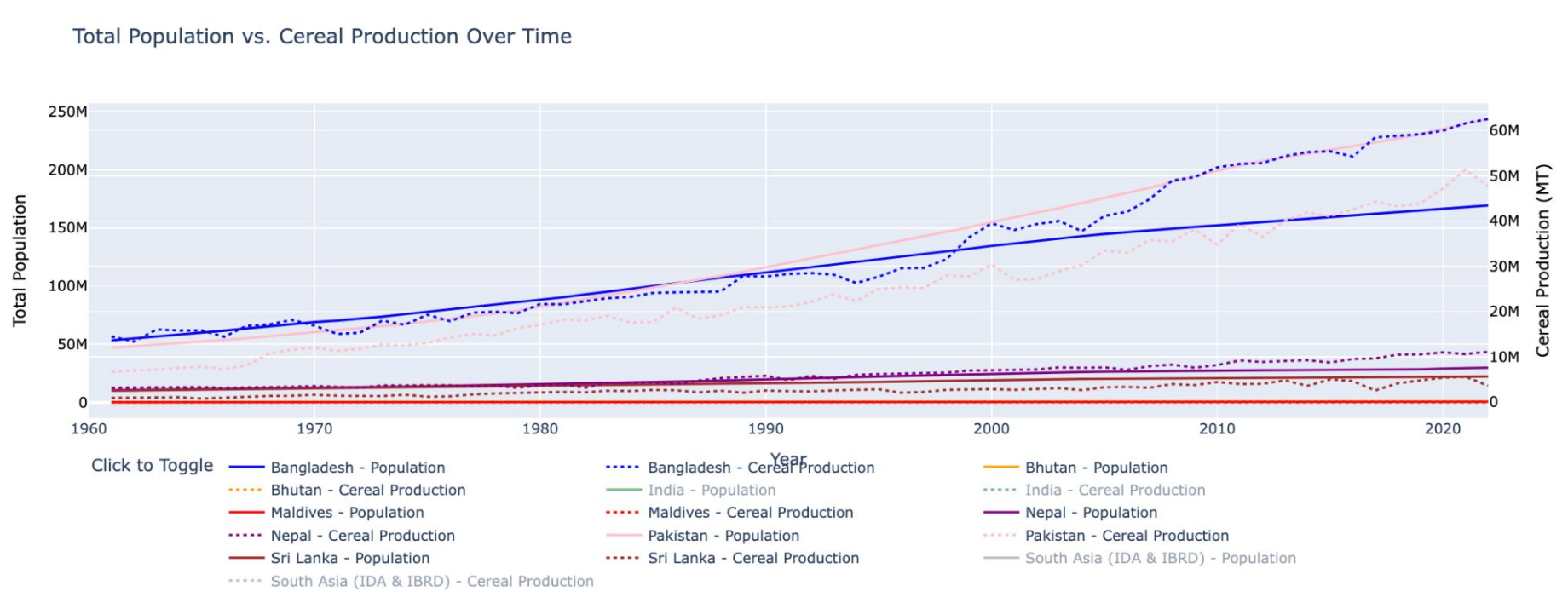
Synthesis

Putting things together

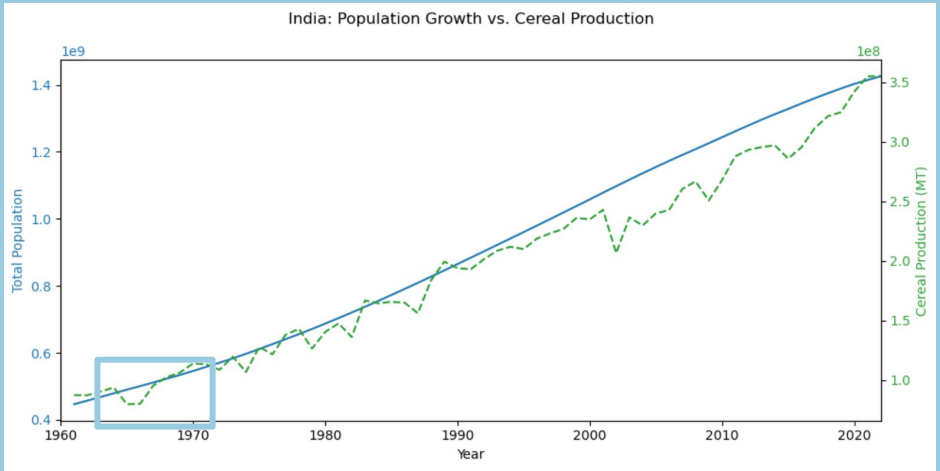
Do trends in population growth, particularly periods of sharp growth or decline, coincide with trends seen in agricultural production?

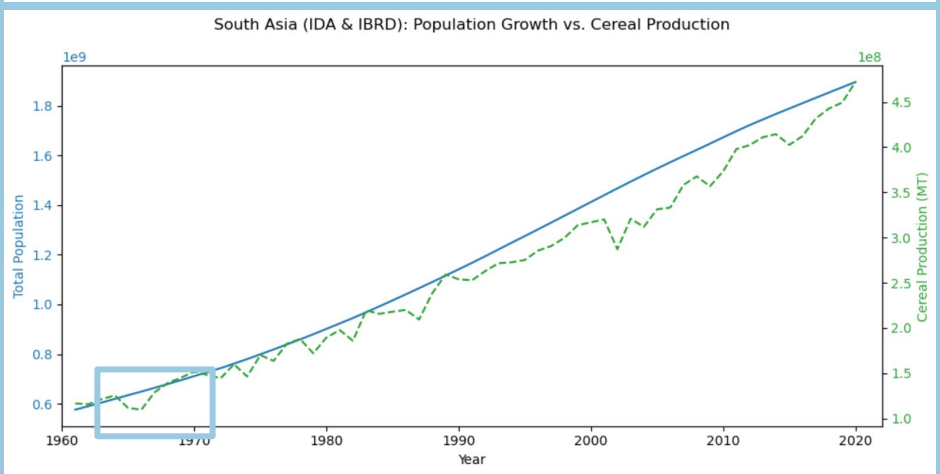
By overlaying total population and total cereal production data for each region, our team hopes to determine if there are any correlated trends of growth that align with key dates in the Green Revolution.

Population Against Cereal Production Over Time (South Asian Countries barring India)



Most South Asian countries show a closely correlated relationship between total population growth and increased agricultural (specifically cereal) production. Notably, Bangladesh leads in both fields with cereal production overtaking population around 1998, roughly 18 years after the country was introduced to Green Revolution technologies.



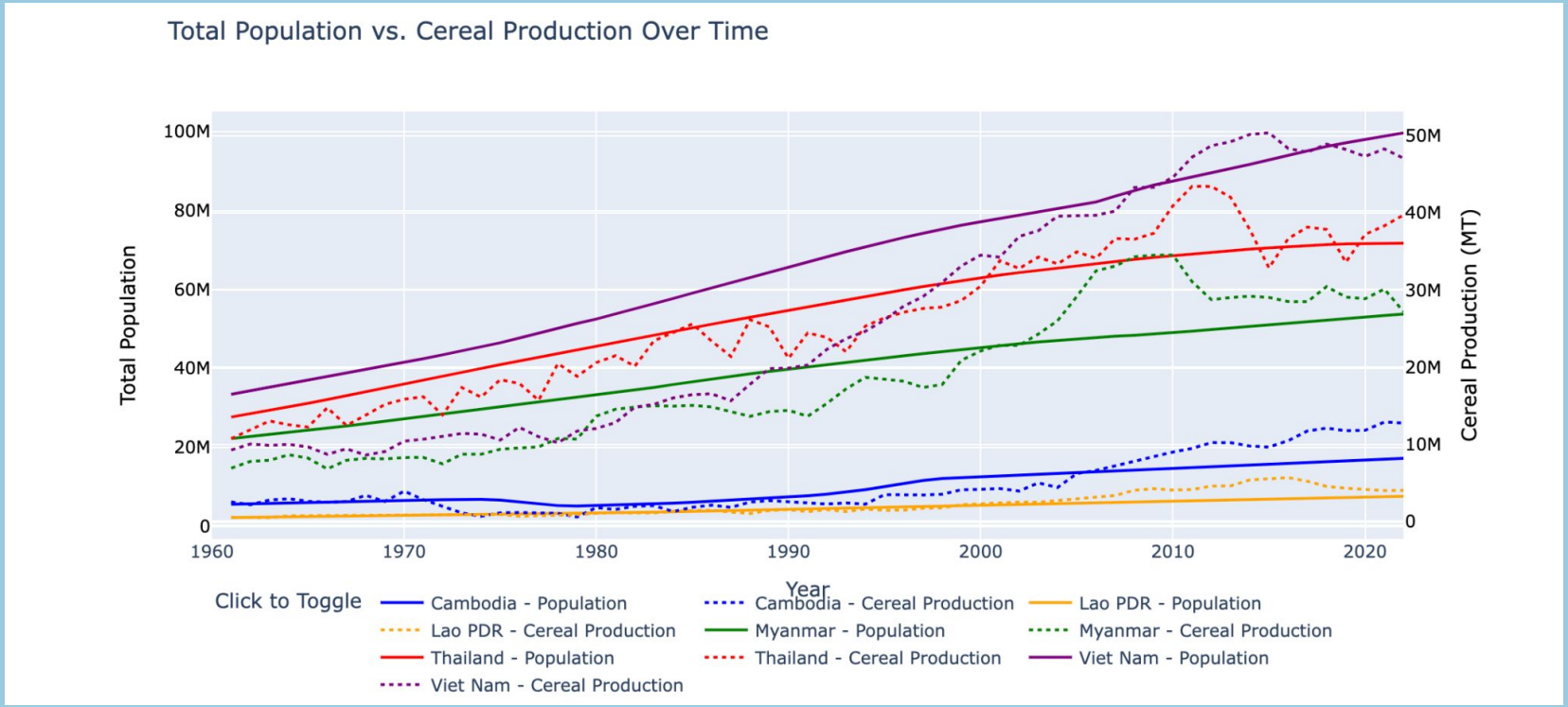


India and Aggregate South Asia

Specific analysis of India and the aggregated data for South Asia show a **similar pattern of close population and cereal production growth since 1960.** This nearly overlaid growth pattern starts to split in the early 2000s when cereal production in both India and across the region begins to slow in contrast to total population.

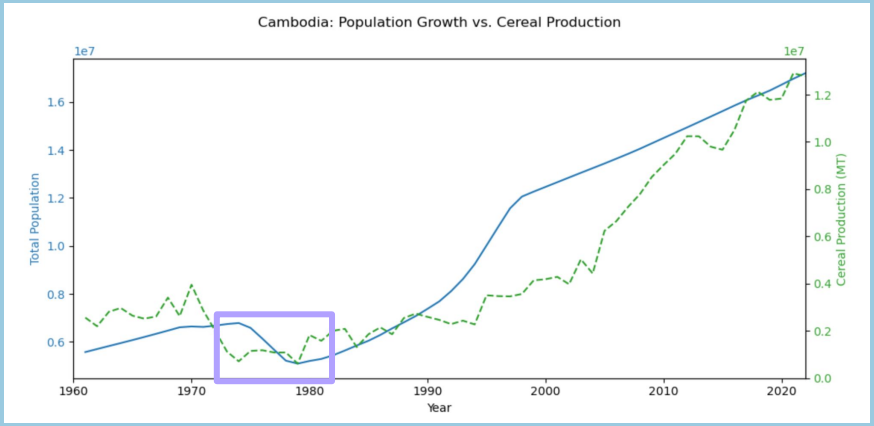
As indicated by the blue boxes, **the introduction of high yield varietal grains to India around 1965 are shown to jump cereal production** to match levels of total population growth; however, do not seem to impact population visibly.

Population Against Cereal Production Over Time (Southeast Asian Countries)

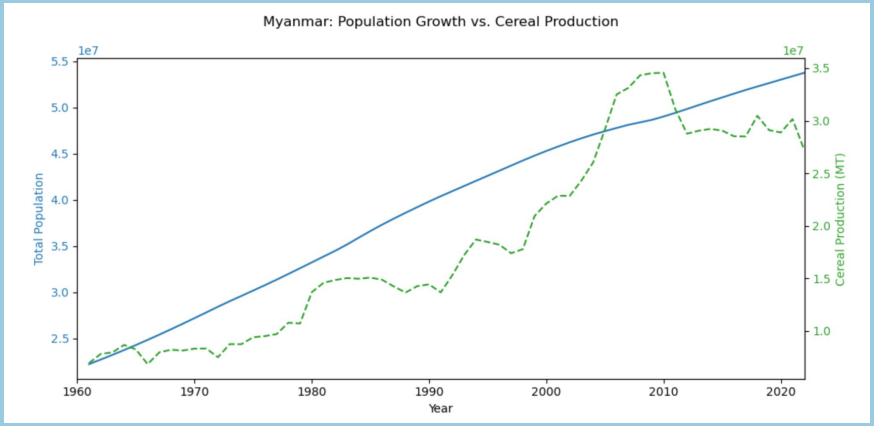


All countries in Southeast Asia follow the general trend of population growth and cereal production with the beginning of the 21st century showing a higher growth in cereal production than population growth for most countries.

Outliers: Cambodia and Myanmar

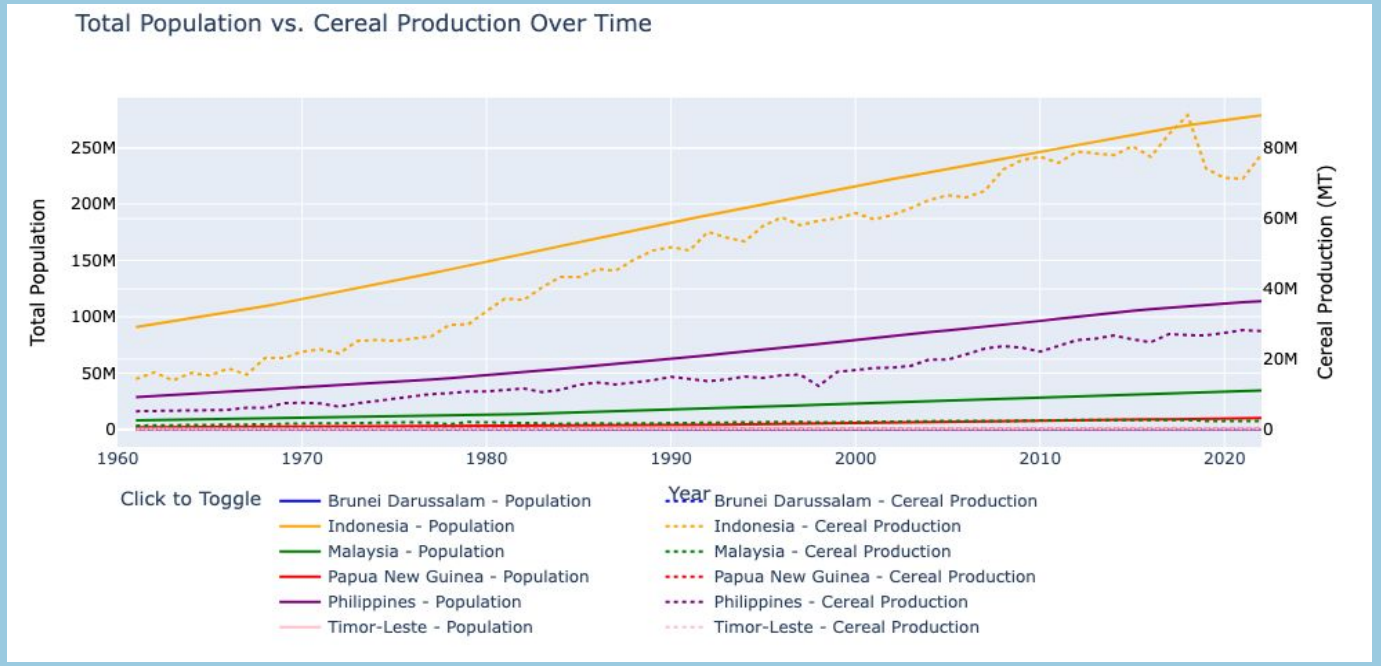


Southeast Asian countries were plagued by political instability: Vietnam War in the 1960s, Cambodia Khmer Rouge in the 1970s, Myanmar Coup D'etat, etc. Those stalled rice production, the main output for SEA countries.



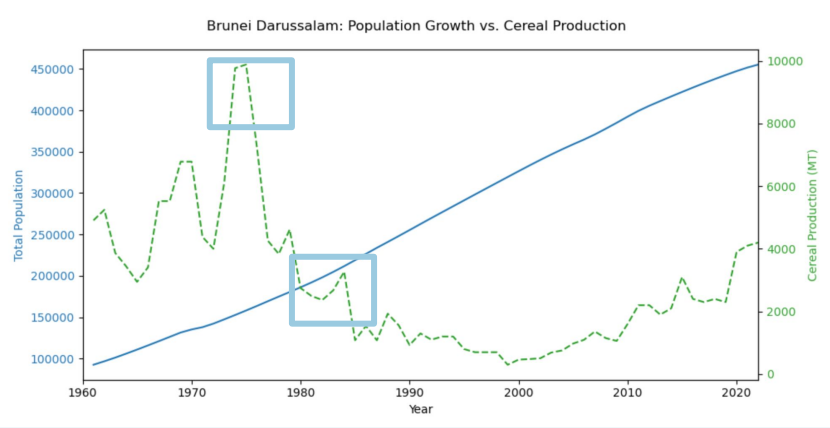
The improvement in agricultural technology post 1990s–2000s saw noticeable improvements in cereal production. This is accompanied by relative political stabilities for SEA countries that enable technological improvements in the agricultural sector. There is a decrease in cereal production post 2015 due to droughts and famine.

Population Against Cereal Production Over Time (Maritime Southeast Asia)

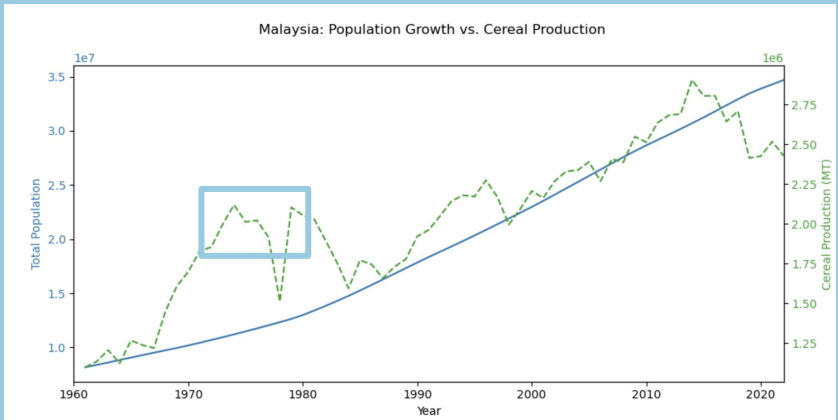


The visualization analyzes the relationship between the total population and cereal production over time for six countries in Maritime Southeast Asia. By comparing the changes in population and cereal production over time for each country, it can be observed that crop yields are linked to population growth.

Negative Population and Cereal Production Growth (Maritime Southeast Asia)

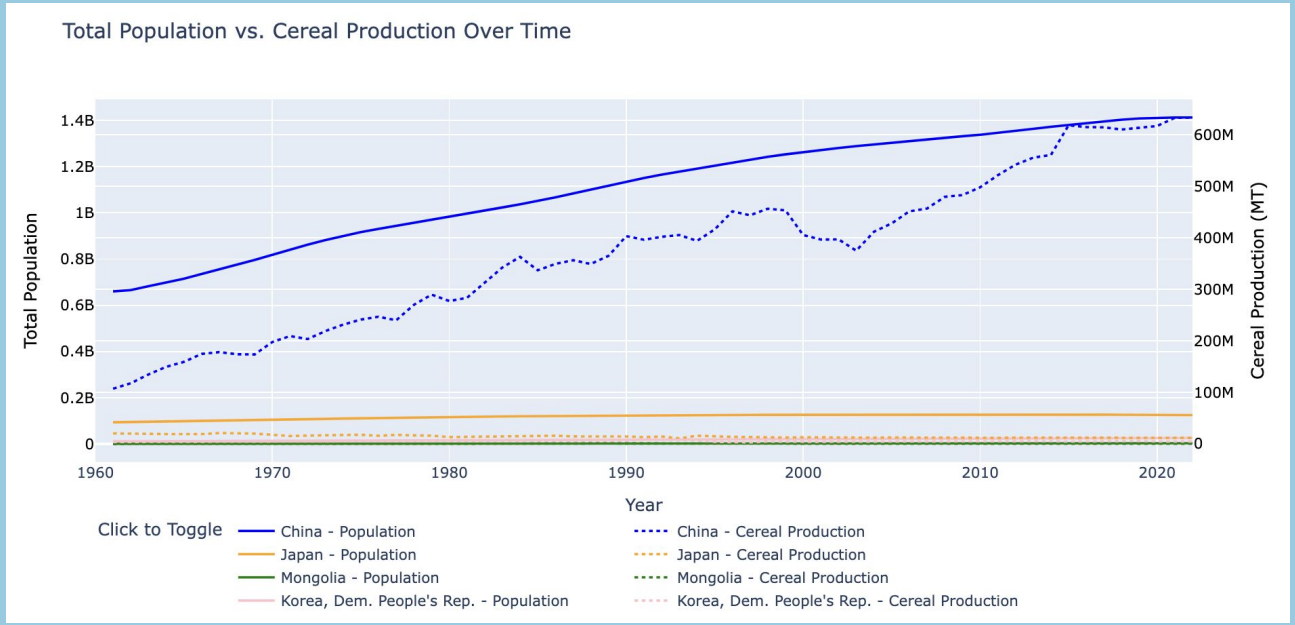


The sharp drop in cereal production around **1975** in Brunei Darussalam was due to the country shifting focus to oil and gas production. This shift, along with political changes following Brunei's full independence from Britain in **1984**, led to less emphasis on agriculture and a decline in cereal production.

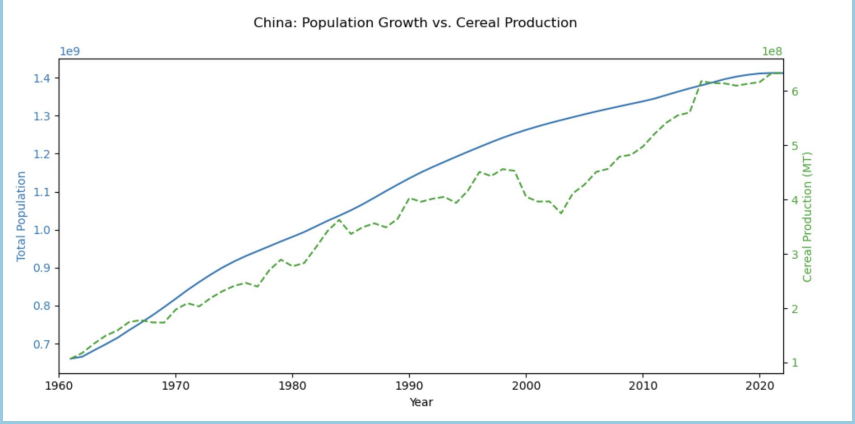


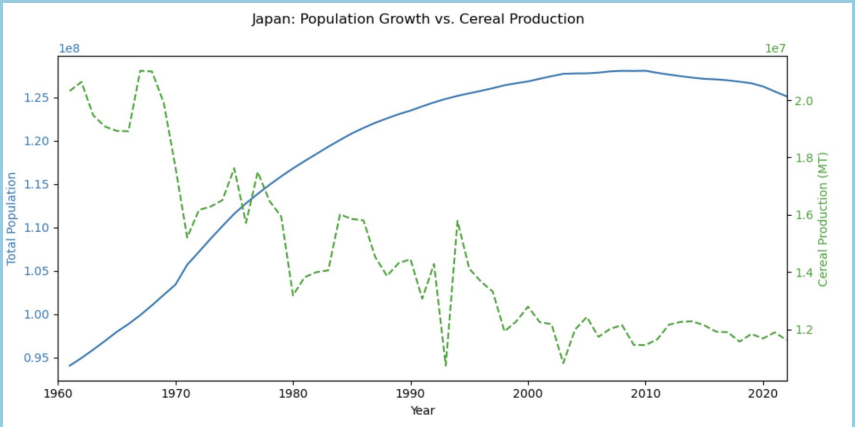
The decline in cereal production in Malaysia **from 1970 to 1980,** despite population growth, can be attributed to the country's shift towards diversifying its economy. During this period, Malaysia focused on export-driven sectors like palm oil and rubber, diverting resources and attention away from local food production, including cereal farming.

Population Against Cereal Production Over Time (East Asia)



The visualization analyzes the relationship between the total population and cereal production over time for 4 countries with the available data for East Asia. While this graph demonstrates the clear linkage between cereal production and population, let's get more granular in viewing population growth and cereal comparisons for some individual countries.



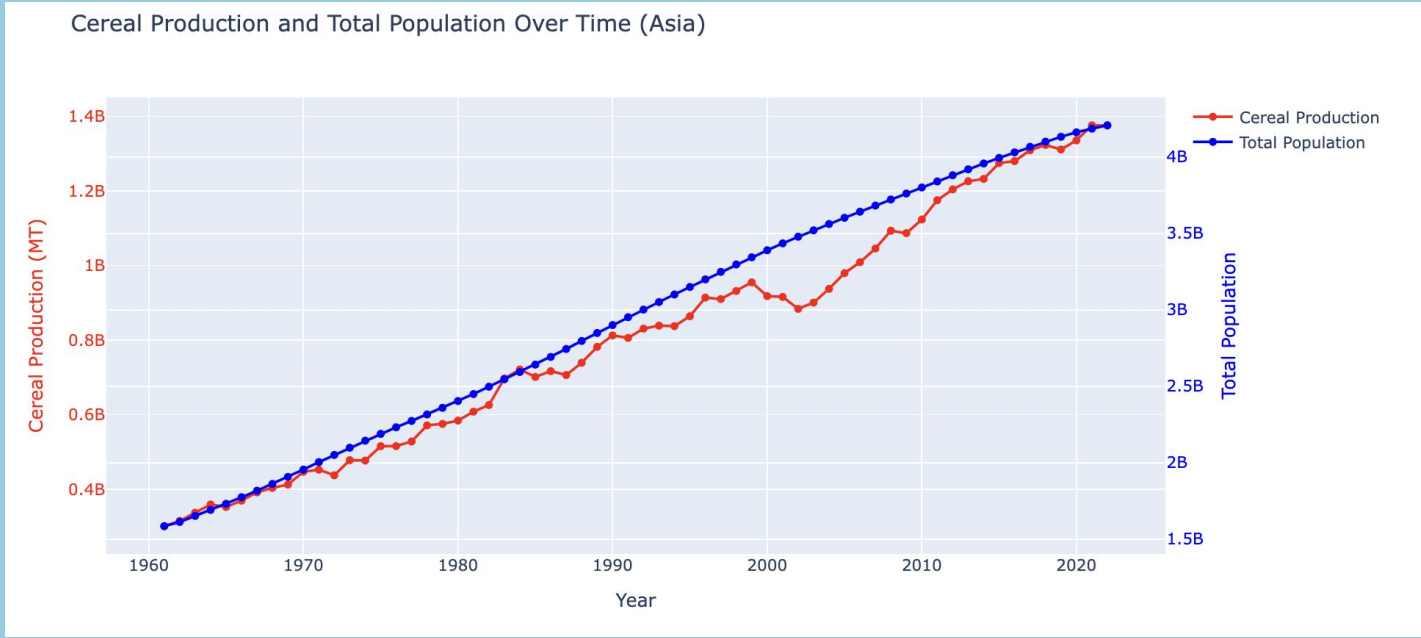


Evidence for and Against a Cereal / Population Relationship

When specifically looking at china, there is a clear alignment between population growth and Cereal Production over time, with both on an upward trajectory.

At the same time, when looking at Japan these trends are nearly inverse, with increasing population growth rates until the mid 2000s and decreasing Cereal Production rates. This is likely because population is slow to respond to changes in cereal production and Japan was one of the earliest countries to **industrialize** and move away from an Agrarian economy.

Zooming Out: Population Against Cereal Production Over Time (Asia)



From a view of Asia in its entirety, the there is a significant correlation between cereal production and population growth (Pearson correlation coefficient of ~ 0.99), demonstrating the impacts the green revolution has had on promoting the growth of population.

Omitted Variables

Globalisation

The spread of global trade and cooperation subsequently allowed for nations in the developing world to grow their economies, creating an environment conducive to population growth.

Health Innovations

Health advancements limiting infant mortality and providing treatment to previously life threatening illness removed blockers to rapid population expansion.

Welfare Laws

Advancements in welfare laws increased the standard of living in developing nations, allowing for longer, healthier lives and thus overall population growth.

Synthesis

Even while considering the omission of other variables that might, and have, influenced population growth across Asia, it is clear from our visualisations that **agricultural production (particularly cereal production) is closely correlated with population growth**. This provides evidence **supporting** the claim that agricultural development drives population growth, especially when noticing the spikes in production that coincide with events in the Green Revolution, such as the introduction of high yield variety grains.

While we cannot *definitively* relate total population growth across Asia with the Green Revolution, we can determine that agriculture is a key driver in increasing the amount of people in a region. Furthermore, noticing that the rural to urban proportion of each region's population is decreasing over time, we can assume that agriculture is becoming more efficient (as people move away from rural industry). Thus, we can conclude that population growth across Asia was **not only driven by increased production from the Green Revolution, but the socio-economic restructuring that occurred as a result of increased efficiency**, indicating a direct relationship between agricultural advancements (like the Green Revolution) and population growth.