

# Global Trends in Regional GDP

Stellan Kilpatrick  
Spring 2024

## Introduction

As globalization connects the world's economies, understanding the patterns of GDP distribution becomes increasingly important. GDP (Gross Domestic Product) is a measure of the value of the final goods and services produced in a country.<sup>1</sup> GDP is an important indicator of a country's economic health and growth trajectory. The world GDP in 1970 was 2.96 trillion dollars, whereas in 2009 it was 60 trillion dollars. The question of interest is if the different regions of the world have had a statistically significant change in their mean proportion of the world GDP from 1970-1979 and 2000-2009. Hypothesis tests will be performed for each of the seven geographic world regions recognized by the World Bank, to compute a difference in the mean proportions of each region's average proportion of the world GDP between the two decades.

## Background

The data set comes from the World Bank and OECD and was accessed through datahub.io which collected the data and turned it into a CSV file with 11,507 rows and 4 columns. The key variables are the country name, year, and GDP. These represent the name of the country or region as a character, the year the data was taken (1960 - 2016) as a double, and the GDP of the country in US dollars for that given year as a double. A single row represents the GDP of a country or region in a single year. The regions used for the hypothesis tests are officially classified according to World Bank analytical grouping<sup>2</sup>. The other column is country code which is a character and is not too important for the project.

Data source is a CSV file retrieved from datahub.io<sup>3</sup> and sourced from the World Bank and OECD. The international standard for measuring world GDP was created by representatives of the IMF, EU, OECD, UN, and the World Bank.<sup>4</sup> Thus, the data is extremely reliable as it was collected by the World Bank and the OECD, so this data matches the international standard for measuring world GDP.

The time periods from 1970-1979 and 2000-2009 were selected to illustrate the economic changes experienced globally over 40 years. The focus will be on the data rows representing these periods for each region, as well as the world GDP rows for the same years, to compute the mean proportions. By using the GDP data from the world GDP rows and the corresponding regional GDP for each specified year, the proportion of each region's share of world GDP will be determined. With 20 years of data for seven regions, the resulting dataset will consist of 140 GDP proportions. This data will then be analyzed to calculate the mean proportion of world GDP share for each region and period, which will be used for hypothesis testing.

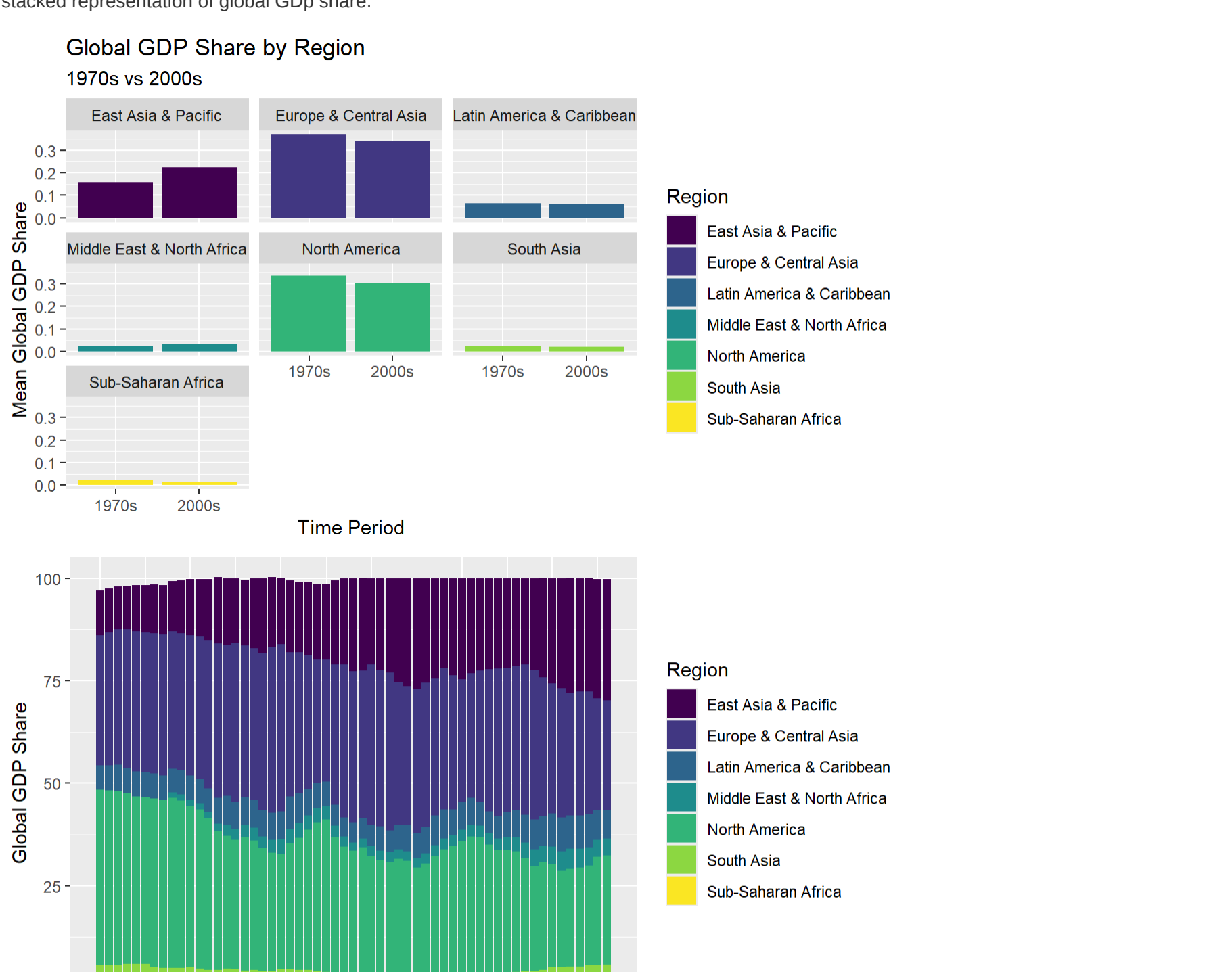
Some of the countries had GDP values that did not start to be reported until after 1970 which could lead to some missing GDP values for the regional data in 1970-1979 which can possibly affect the interpretation of the results as it could significantly alter the proportions.

The remainder of this report will analyze the regions to determine whether the change in their mean proportion of the world's GDP from 1970-1979 to 2000-2009 was statistically significant. A hypothesis test will be conducted to assess the difference in mean proportions for each region.

## Analysis

The first step involved calculating each region's proportion of the world GDP for the periods 1970-1979 and 2000-2009. This was achieved by using the original dataset to create a World GDP dataset for each year. The World GDP dataset was then joined with the original dataset and filtered to include only the seven regions and the specified periods, resulting in the Region dataset. A new column, GDP Share, was created by dividing each region's GDP by the World GDP for the corresponding year, providing the proportion of the World GDP for each region in the years 1970-1979 and 2000-2009. Additionally, a period column was introduced to distinguish between the two periods. The average GDP share was then calculated for each period and region.

The graph below shows these mean proportions split up into each region. The x-axis are the two periods in each region. The y-axis is the mean global GDP share for each region, or the mean proportion of the region's share of the global GDP, as well as a time-series graph displaying a stacked representation of global GDP share.



## Hypothesis Testing

For all of the regions a Hypothesis Test (Welch Two-Sample t-Test) will be performed where  $\mu_1$  is the mean proportion of the region's worldwide GDP share from 1970-1979 and  $\mu_2$  is the mean proportion of the region's worldwide GDP share from 2000-2009.

Statistical Model:

$$X_i \sim F_1(\mu_1, \sigma_1), \quad i = 1, \dots, n_1$$

$$Y_i \sim F_2(\mu_2, \sigma_2), \quad i = 1, \dots, n_2$$

Hypotheses:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

Test Statistic:

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

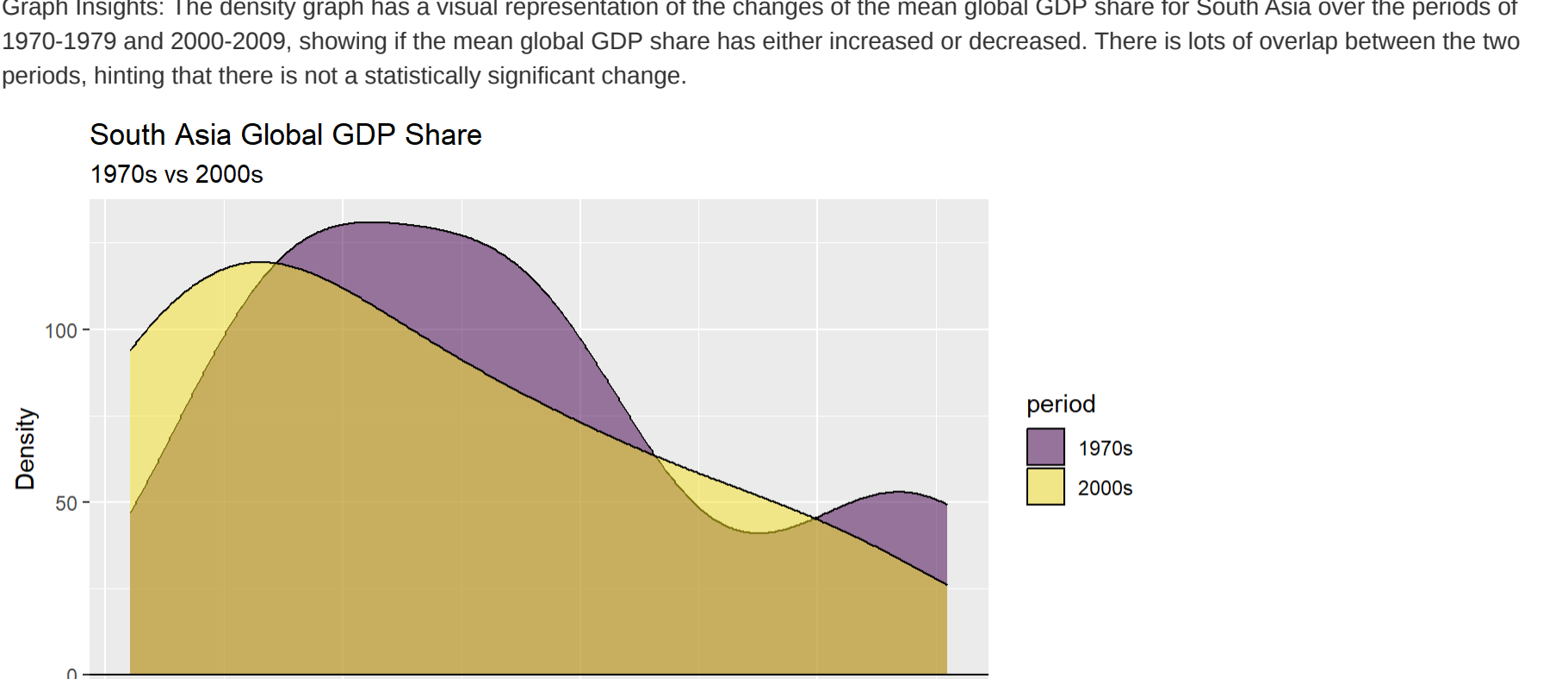
Sampling Distribution: T with 19 degrees of freedom

### South Asia

P-value Calculation:

- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = 0.829$ . The T-test for the difference of mean proportions of South Asia's worldwide GDP share between the two periods resulted in a p-value of 0.42.

Graph Insights: The density graph has a visual representation of the changes of the mean global GDP share for South Asia over the periods of 1970-1979 and 2000-2009, showing if the mean global GDP share has either increased or decreased. There is lots of overlap between the two periods, hinting that there is not a statistically significant change.

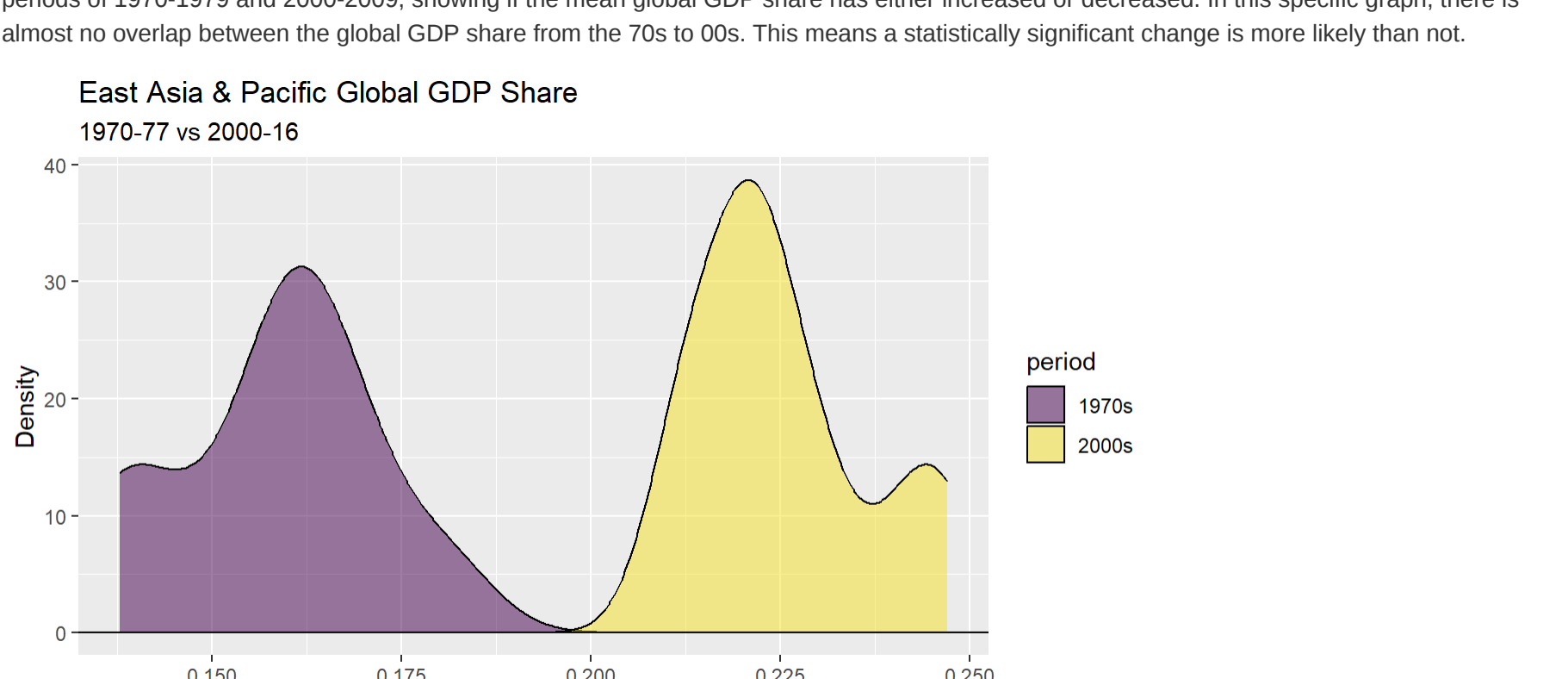


### East Asia & Pacific

P-value Calculation:

- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = -11.87$ . The T-test for the difference of mean proportions of East Asia & Pacific's worldwide GDP share between the two periods resulted in a p-value of 3e-10.

Graph Insights: The density graph has a visual representation of the changes of the mean global GDP share for East Asia & Pacific over the periods of 1970-1979 and 2000-2009, showing if the mean global GDP share has either increased or decreased. In this specific graph, there is almost no overlap between the global GDP share from the 70s to 00s. This means a statistically significant change is more likely than not.

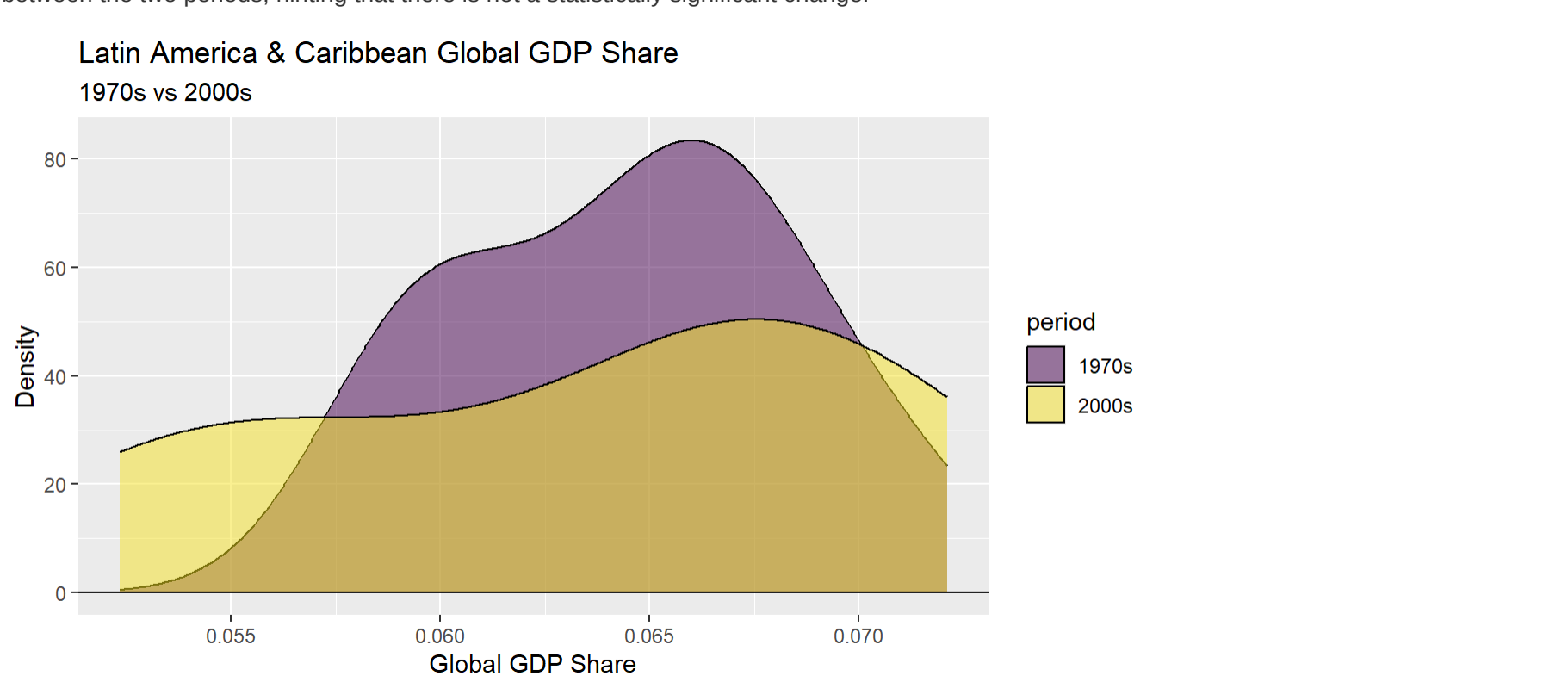


### Latin America & Caribbean

P-value Calculation:

- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = 0.446$ . The T-test for the difference of mean proportions of Latin America & Caribbean's worldwide GDP share between the two periods resulted in a p-value of 0.66.

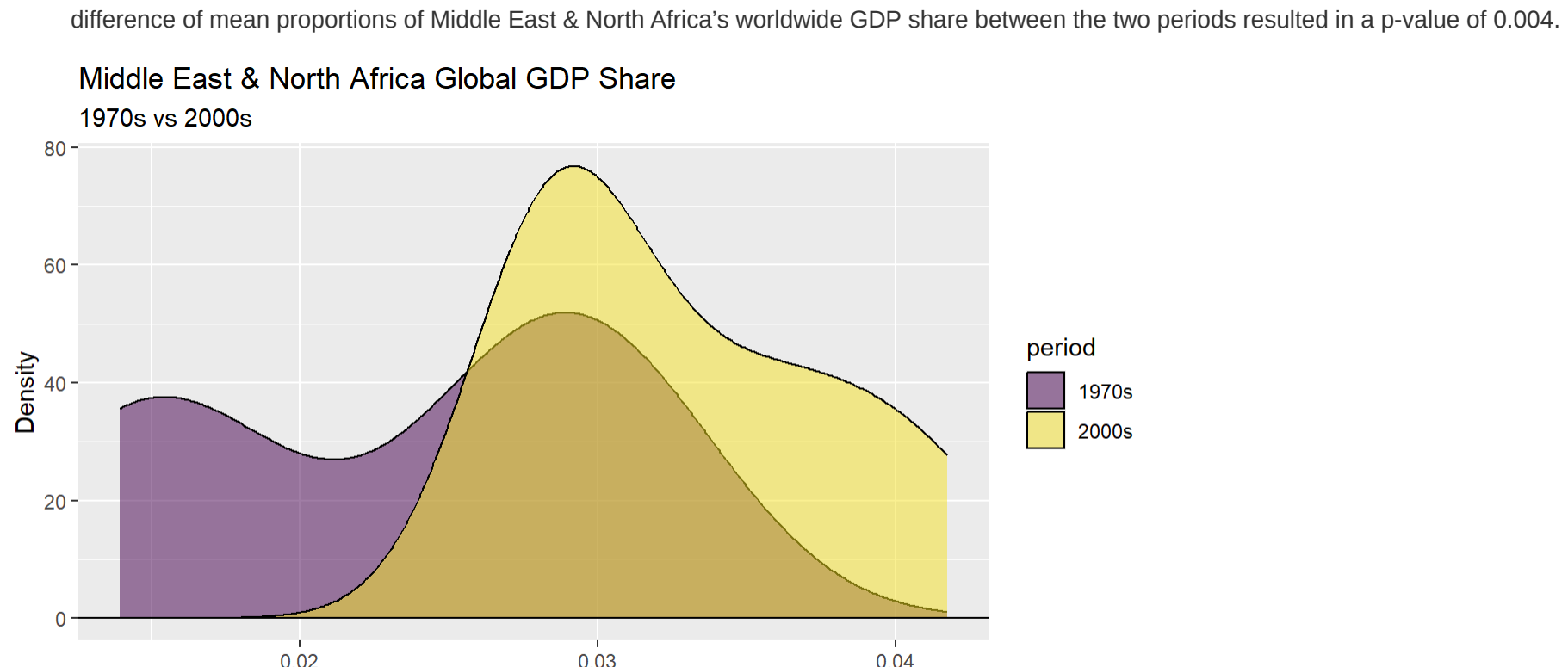
Graph Insights: The density graph has a visual representation of the changes of the mean global GDP share for Latin America & Caribbean over the periods of 1970-1979 and 2000-2009, showing if the mean global GDP share has either increased or decreased. There is lots of overlap between the two periods, hinting that there is not a statistically significant change.



### Middle East & North Africa

P-value Calculation:

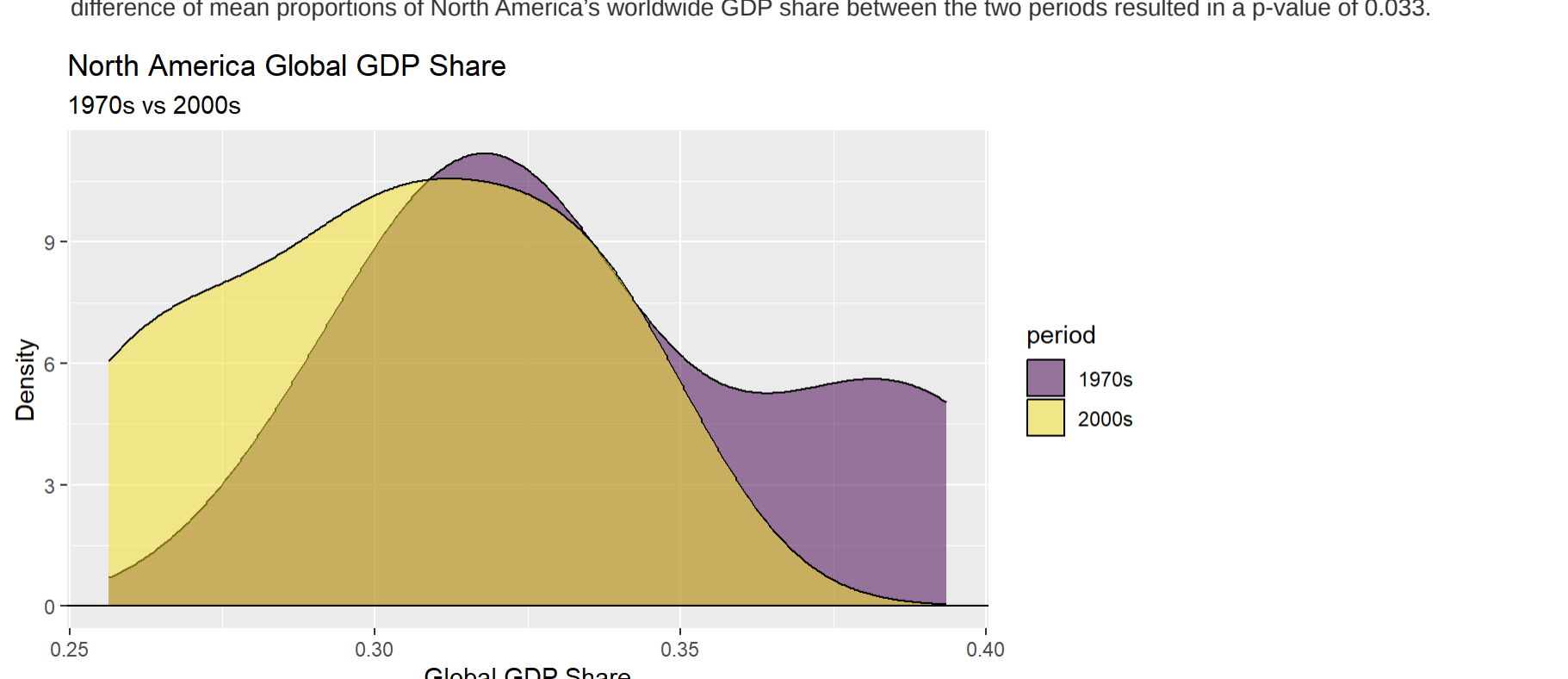
- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = -3.300$ . The T-test for the difference of mean proportions of Middle East & North Africa's worldwide GDP share between the two periods resulted in a p-value of 0.004.



### North America

P-value Calculation:

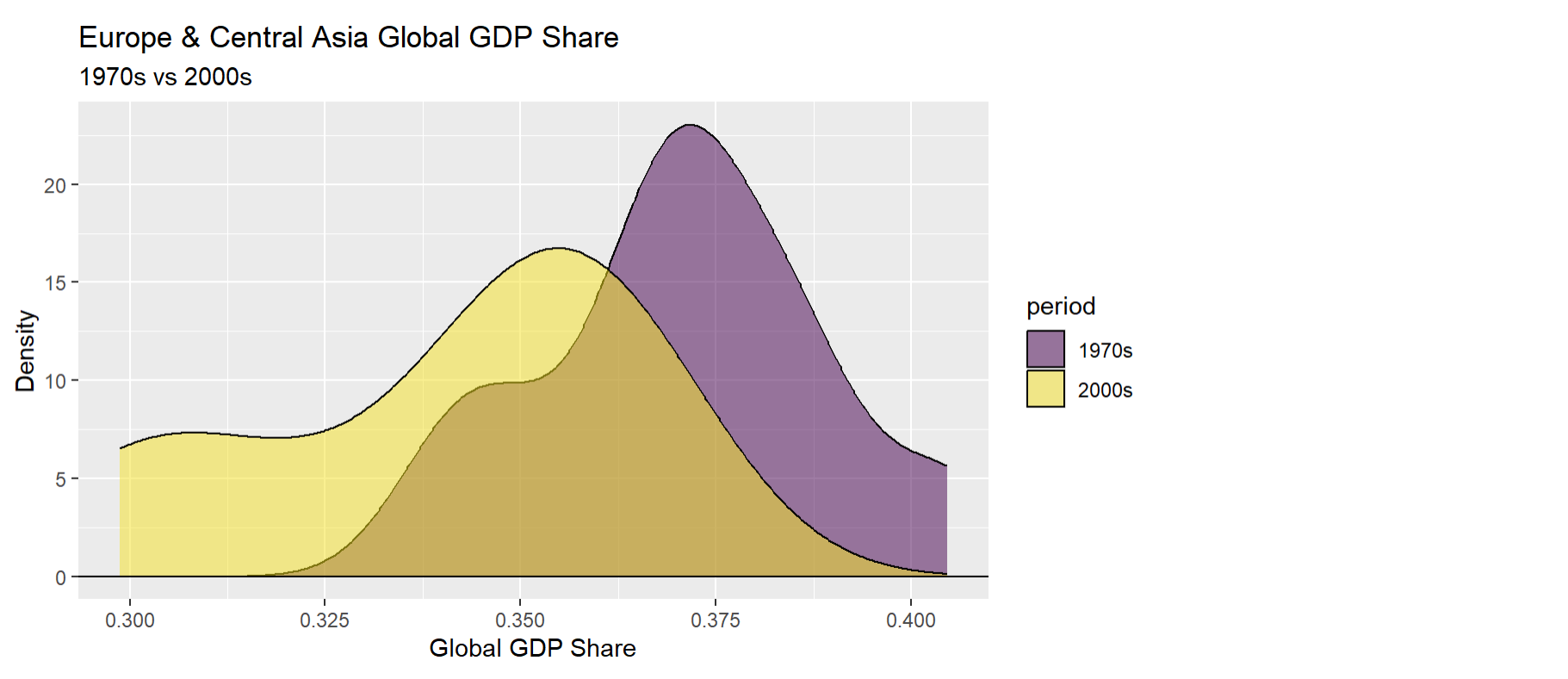
- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = 2.299$ . The T-test for the difference of mean proportions of North America's worldwide GDP share between the two periods resulted in a p-value of 0.033.



### Europe & Central Asia

P-value Calculation:

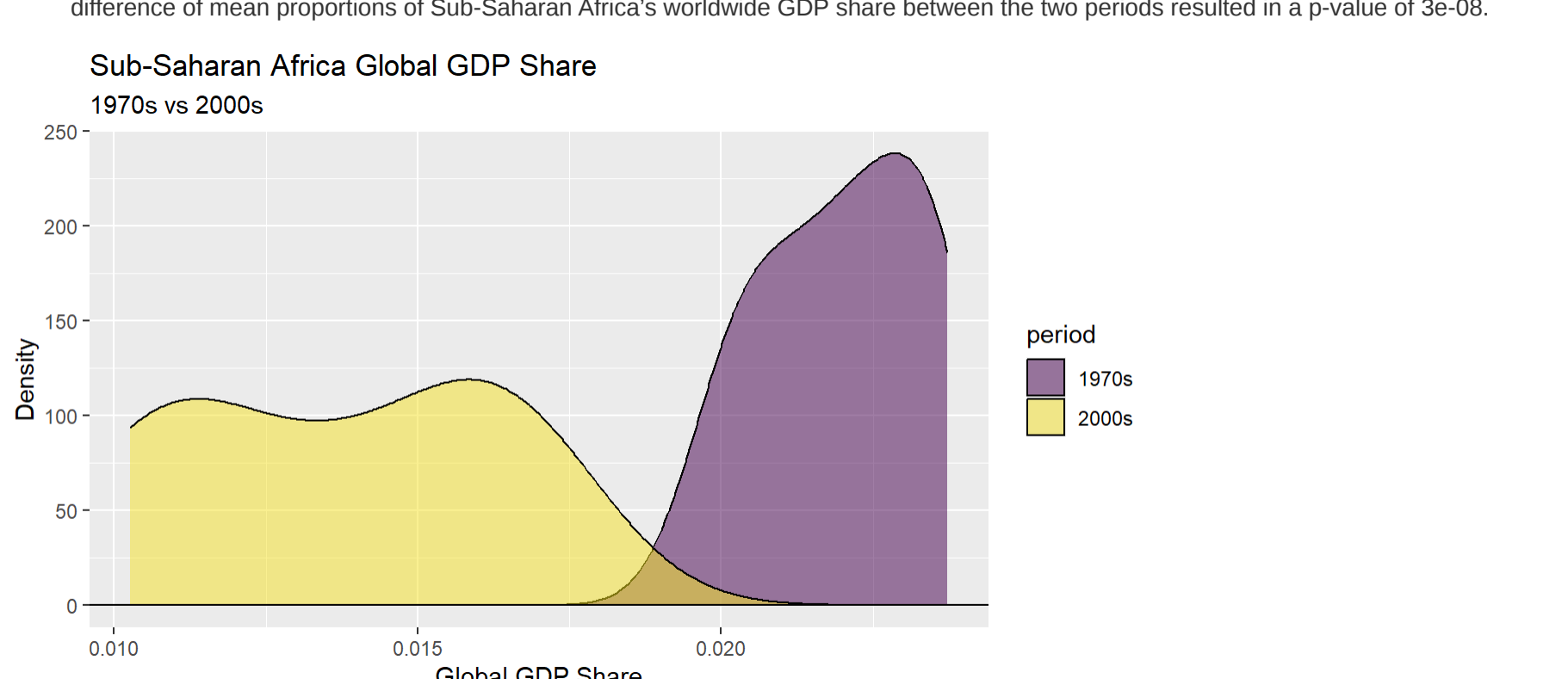
- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = 3.12$ . The T-test for the difference of mean proportions of Europe & Central Asia's worldwide GDP share between the two periods resulted in a p-value of 0.006.



### Sub-Saharan Africa

P-value Calculation:

- Statistics which are at least as favorable to the alternative hypothesis are those values less than or equal to  $T_{obs} = 8.87$ . The T-test for the difference of mean proportions of Sub-Saharan Africa's worldwide GDP share between the two periods resulted in a p-value of 3e-08.



## Discussion

### Interpretation

Based on the density graph and hypothesis test, there is not evidence that the mean proportion of the world GDP of South Asia and Latin America & Caribbean in 2000-2009 is larger than it's mean proportions of the world GDP in 1970-1979 ( $p=0.41$ ,  $0.66$ , respectively,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). The World Bank attempted to justify the lack of growth of South Asia by citing how "persistent structural challenges threaten to undermine sustained growth, hindering the region's ability to create jobs and respond to climate shocks."<sup>5</sup>

There is strong evidence that the mean proportion of the world GDP of North America in the 2000s is smaller than in the 1970s ( $p=0.033$ ,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). The decrease is a result of globalization and increased competition for imports and exports around the world, along with technological developments that have allowed other world regions to catch up to the economic power that was North America, more specifically the United States, in the 1970s.

There is very strong evidence that the mean proportion of the world GDP of East Asia & Pacific in the 2000s is larger than in the 1970s ( $p=3e-10$ ,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). This increase could be due to Asia's major industrial growth which allowed for their GDP to skyrocket.

Based on the density graph and hypothesis test, there is strong evidence that the mean proportion of the world GDP of Europe & Central Asia in the 1970s is smaller than in the 2000s ( $p=0.006$ ,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). This decrease could be due to Europe not increasing at the rate of other industrial powers like China who have stolen some of Europe's share of the world GDP.

There is strong evidence that the mean proportion of the world GDP of Sub-Saharan Africa in the 1970s is smaller than in the 2000s ( $p=3e-08$ ,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). According to African Development Bank Group, this is caused by "offset population increases, leading to minimal gains in per capita GDP. Structural transformation has been limited, with economies heavily reliant on traditional, low-productivity sectors like agriculture or low-skilled services for growth and employment."<sup>6</sup>

Finally, there is very strong evidence that the mean proportion of the world GDP of Middle East & North Africa in the 2000s is larger than in the 1970s ( $p=0.004$ ,  $df = 19$ , two-sided t-test, Welch Two-Sample t-Test). The explosion of GDP in this area can be attributed to the mass oil production and export.

### Shortcomings

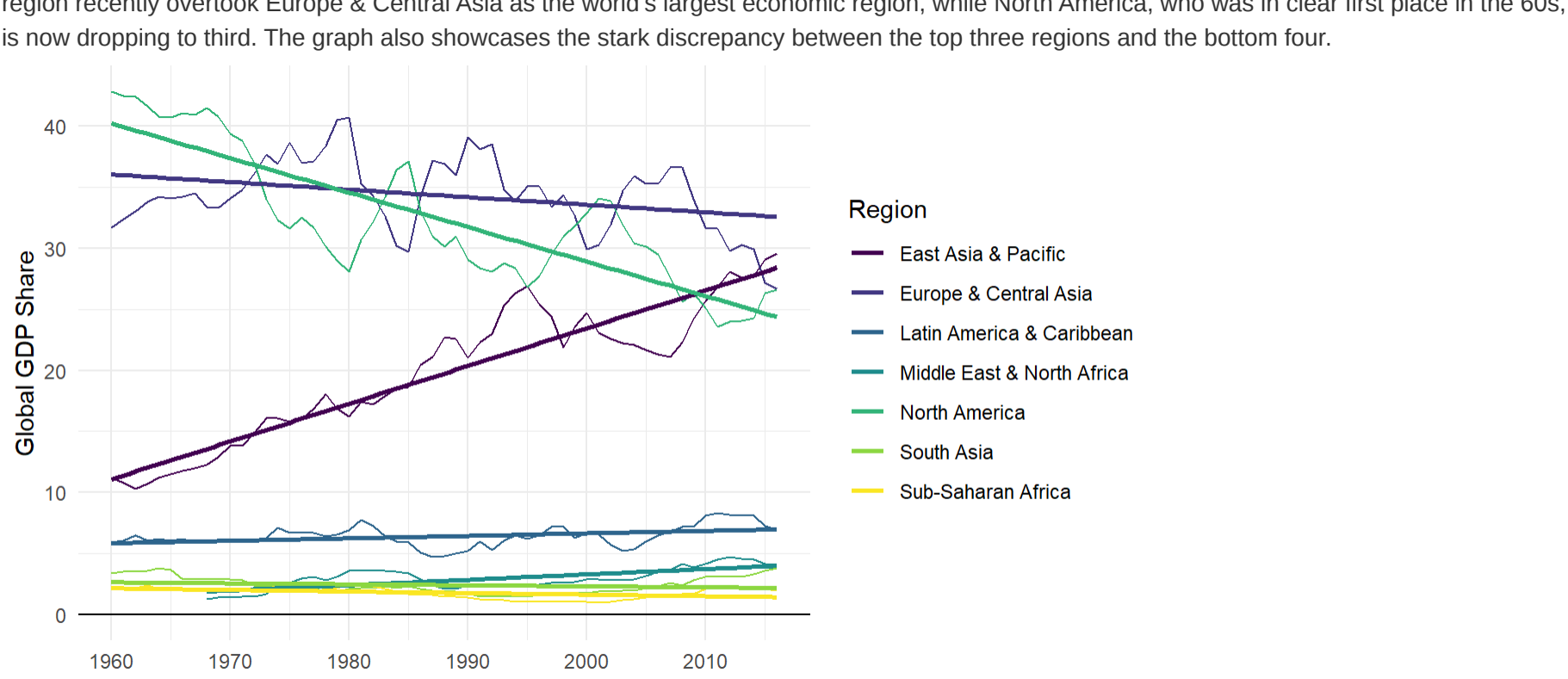
A potential short-coming of the analysis is some countries within the regions did not have GDP data until after 1970. This could alter the total GDP of regions because country data could be added to the region in the middle of a period, which would increase its GDP share.

### Further Research

Another question for the data is if the global GDP for the regions in the year 1970 versus 2009 changed by a statistically significant amount. That would require a difference in proportions test. A different method to address the overarching questions is a z-test instead of a t-test and using a normal approximation instead of a t-distribution. This new data could reveal more relationships between the GDP change of different regions, and give more information into the shift in economic power throughout the world that has happened within the last 50 years.

### Looking Ahead

Not only should the past be analyzed, but future projections can say a lot about how the world is evolving and it may look like in a few years or decades. By running a linear regression model on the 56 years (1960-2016) of GDP data, some interesting trends emerge. East Asia & Pacific region recently overtook Europe & Central Asia as the world's largest economic region, while North America, who was in clear first place in the 60s, is now dropping to third. The graph also showcases the stark discrepancy between the top three regions and the bottom four.



## Conclusion

After analyzing the findings, it is clear that the economic makeup of the world did change notably between the two time periods. Of the seven regions, two grew their share in global GDP at a statistically significant level (Middle East & North Africa, East Asia & Pacific), two regions did not have a statistically significant change to their global GDP share (South Asia, Latin America & Caribbean), and three regions saw their share of global GDP decline (North America, Europe & Central Asia, Sub-Saharan Africa). Combining the data from previous decades with the linear regression model that illustrates future growth trends, it is evident that these shifts will continue to occur and influence global economic dynamics.

## References

- <https://www.investopedia.com/terms/g/gdp.asp#1>
- <https://datatopics.worldbank.org/sdgdattas/archive/2017/the-world-by-region.html#1>
- <https://datahub.io/core/gdp#gdp-1>
- [https://en.wikipedia.org/wiki/Gross\\_domestic\\_product](https://en.wikipedia.org/wiki/Gross_domestic_product)
- <https://www.worldbank.org/en/region/sar/overview#1>
- <https://www.afdb.org/en/knowledge/publications/african-economic-outlook#1>