GDP Analysis and Trend

Kaidi Chen, Ziliang Huang, Weixiao Sun

2/16/2022

Abstract

Our analysis mainly focus on the impact of **Inflaction, Unemployment, and Population** to GDP in United States. GDP is one of the key indicators of the overall economic strength of a country, or even the global economy. Therefore, since we have majors in financial-related fields, we are very interested in this topic and eager to discover what factors affect GDP and how much they can interact.

##Introduction Our study project will be focused on the issue of globalized Gross Domestic Product, which is a key indication of national economic data and a common gauge of a region's economic status and overall degree of development. In general, the GDP of a country may be split into three categories: value, income, and product. It is the difference in value between the total value of all products and services generated by all resident units in a given period and the total value of all non-fixed assets, goods, and services invested in the same period. More specifically, we will investigate the influence and link between population, inflation, and unemployment on GDP. Meanwhile, using data from 2000 to 2021, we will analysis the economic environment across all region and for forecast for the future.

Design Situtaion

Table

For the aggregated table below, it calculates the average GDP in each Country from 2000 to 2021. The table displays only the Top 10 country with the highest GDP in *Billions* of the past twenty years across all regions.

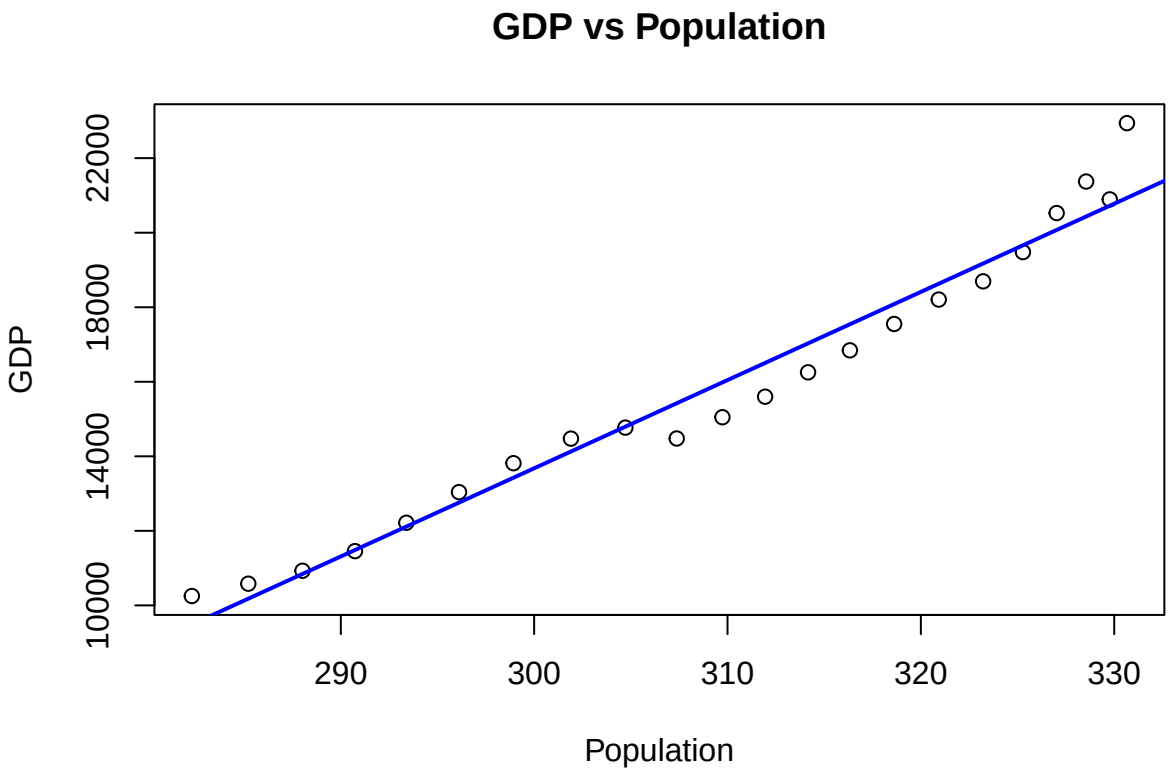
Country	Average_GDP
United States	15883
China	7391
Japan	5019
Germany	3294
United Kingdom	2581
France	2427
Italy	1918
India	1633
Brazil	1567
Canada	1446

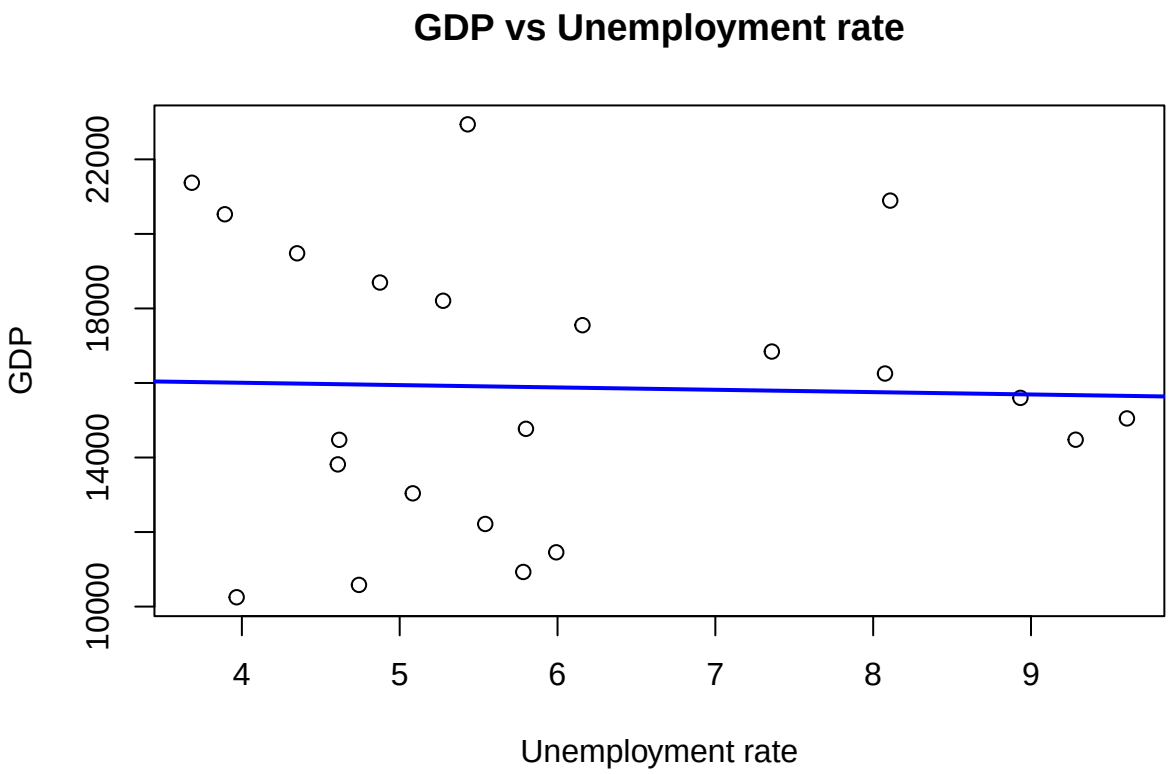
Chart1

How the population, the unemployment rate, and the inflation and end of period consumer prices relate to the GDP in the USA? Purpose: Scatter plot — uses to observe and show the

relationships between two numeric variables.

Insight: were created to compare the overall impact of population, unemployment, inflation and end-of-period consumer prices on GDP in the US. The scatter plot regression data analysis provides a plot of the distribution of data points in the plane for population, unemployment, inflation and end-of-period consumer prices to determine the general trend of GDP with respect to these independent variables.




GDP vs Unemployment rate
GDP
10000
14000
18000
22000
4
5
6
7
8
9
Unemployment rate

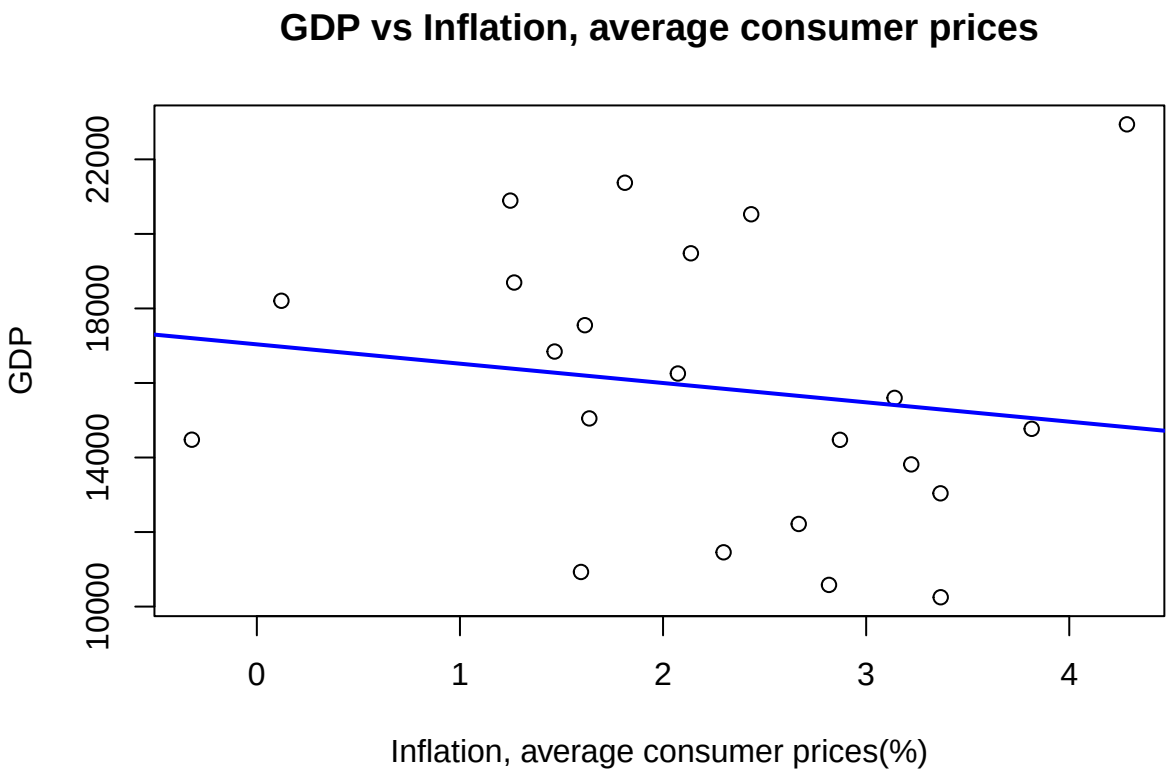
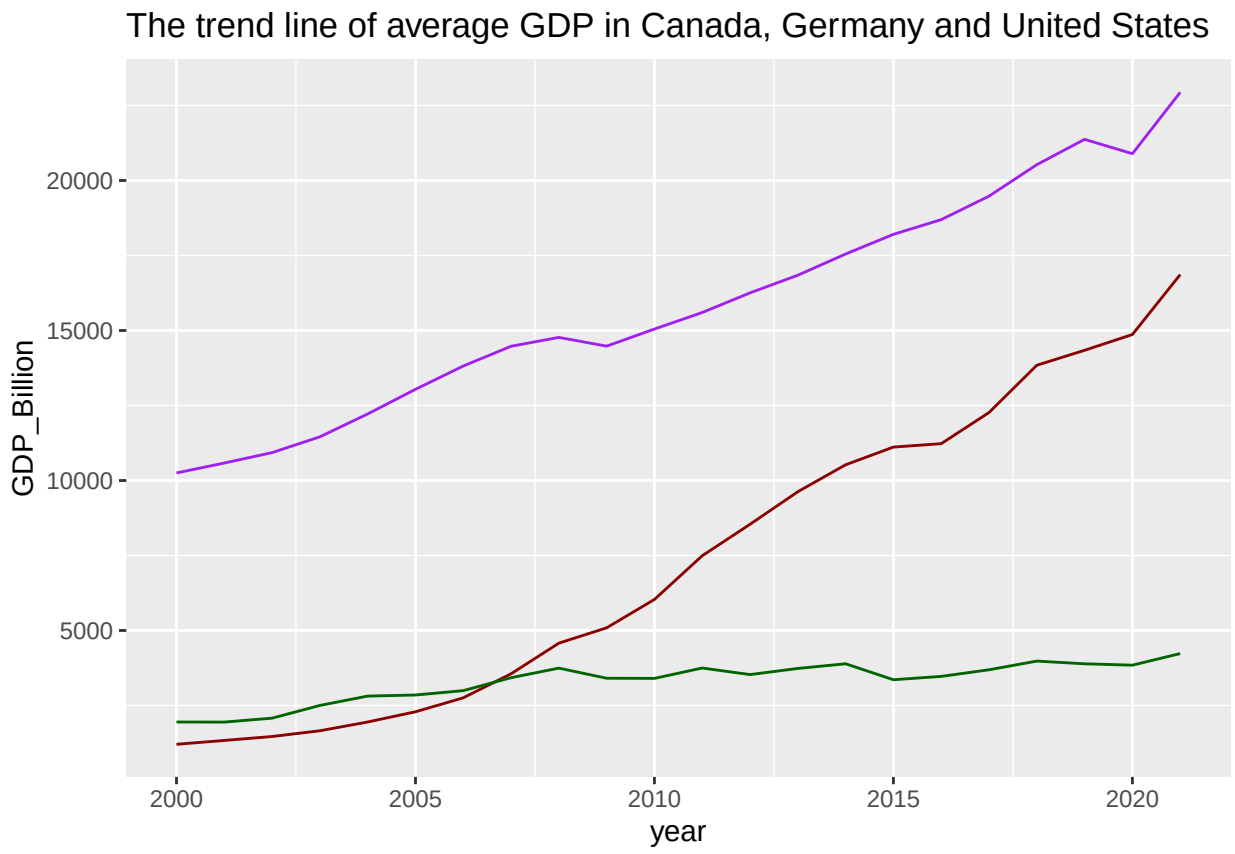


Chart2

What are the comparative GDP trends of China, Germany, United States since year of 2000?
Purpose: Line Graph — uses lines to connect individual data points and displays quantitative values over a specified time interval.

Insight: Was created to compare the GDP trends for the three countries. Because line graphs show the trend of continuous average GDP data since the year of 2000 in Canada, Germany and the United States with an upward trend in overall GDP for all three countries.



The Dataset

The origin of our data is from the International Monetary Fund (IMF), an organization working to foster global monetary cooperation, secure financial stability, facilitate international trade, promote high employment and sustainable economic growth and reduce poverty around the world. The data set we are going to use for this project is 8823 rows and 57 rows, and covers all economic data (e.g. Unemployment, Inflation, GDP, Population) for 190 countries in the world from 1980 to 2021. However, we will only use certain structure of data from the whole set, the population,unemployment, inflation and GDP of all regions from 2000 to 2021 to complete our research question. We only extract these data and compare to United States since the US has changed dramatically in the last 20 years. I think the difficulty with this dataset is whether we can extract the right economic data we need correctly and efficiently. However, there are certain data that are missing in this dataset, which input as "n/a". Moreover, these data are printed into different types of units, including index, number of people, US dollars, and national currency. Hence, we may end up with different numbers of X and Y arguments when we plot them into our graph. Additionally, the class of all the values in this dataset is actually characters, which will let us make mistakes when calculating and graphing this dataset. Thus, we have to convert all the data that we need to calculate into numerical class. As we mentioned above, since we only need two types of data, population, and GDP, which are presented horizontally, their classification is written in a "subject descriptor" column and even the GDP data has several different classifications for each country. Besides, the values in the "subject descriptor" column are complex and are difficult for us to extract corresponding columns for our analysis. This may lead to additional difficulties in the process of data extraction, depending on whether we can find consistent patterns in them. What's more, we discover that the "Country" and the "Region" columns are not factors, thus making our analysis difficult to proceed. In addition, there are too many features in this database and many of them are not related to our project's topic, so we have to reorganize this dataset to make it cleaner.

Expected Implications

For data visualization, technologists must utilize R. Simultaneously, in terms of the link between GDP and its connected elements, we expect to compute the correlation coefficient, r value, and p value to see whether there is a linear relationship that will assist us in forecasting. Designers are required to utilize R markdown to create a dynamic report that better displays our analyzing results, allowing our analysis to be understood by persons other than those in such a professional sector. Policymakers are obligated to safeguard our data analysis findings and guarantee the transparency of our data resources.

Limitations

Our research will be largely focused on the impacts of population, inflation, and unemployment on GDP, and because GDP is impacted by multiple factors and data restrictions, the research project's conclusions may not acquire the predicted trends. The unemployment rate and GDP growth seem to have a reasonably consistent connection. In any event, this link is not particularly rigid and should be evaluated based on the circumstances of the country or region.