DATA VISUALIZATION PROJECT FINAL REPORT

Introduction:

Inflation, the term that we are listening more often these days and the word that people around the world, especially international students in United States need no introduction. As we know that the world is recovering from pandemic of covid-19 and several other pandemics in the past, these pandemics always have affected the world in many adverse ways. One of them is inflation which shatters the lives of people. That is the reason I want to show it in the lenses of data visualization to help how GDP growth is affecting inflation of different countries around the world, how special events are affecting inflation rates etc.

The absence of comprehensive, accessible, and actionable insights hinders efforts to formulate effective policies, anticipate economic trends, and mitigate risks associated with inflation. This project aims to address this gap by leveraging data visualization techniques to explore and analyze inflation and seeks to empower stakeholders with the knowledge needed to navigate the intricacies of inflation and foster sustainable economic growth.

Research Questions:

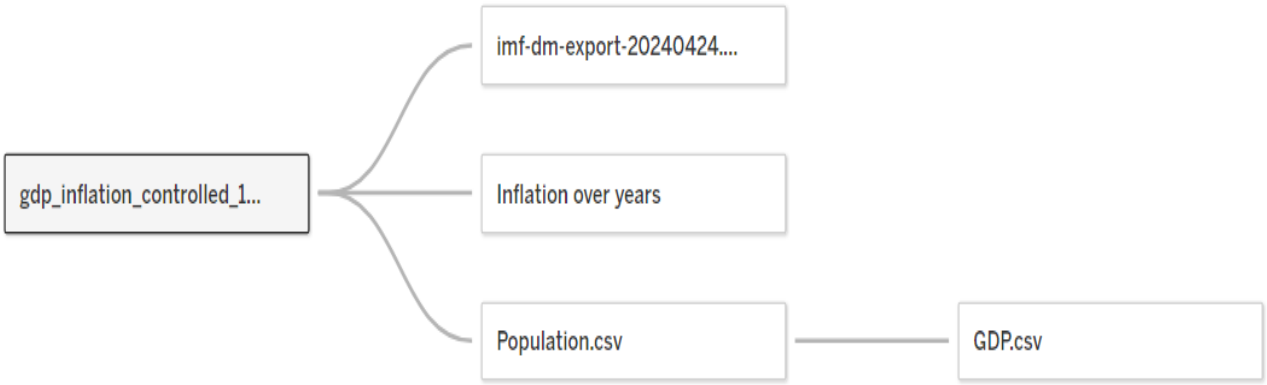
1. What are the top economies by GDP and by year? Are top economies concentrated in a particular continent?
2. What are the top populated countries and is the ranking changing over years?
3. What is the Per Capita Income of different countries during various years?
4. Does top GDP economies also have top per capita income? And what is their population during a particular year we are trying to compare?
5. What is the inflation rate and per capita income of each country during the period we are comparing?
6. What is country's inflation rate and per capita income along with GDP over this period?
7. What are the events that majorly affected inflation rates in the US during different years and does any of the event repeated for different years?
8. How is the GDP varying globally as well as for a country from 1960 to 2020?
9. What are top countries that are taking care of price levels in country to control GDP and inflation in 2020?

Methodology:

In the initial phase of the project, attention was directed towards enhancing the dataset's usability by restructuring a column in the Excel file. By meticulously splitting the column into two distinct attributes, namely the year and country abbreviation, data organization was significantly improved, laying a robust foundation for subsequent analyses. This crucial step not only streamlined data manipulation processes but also facilitated clearer insights into temporal and geographical trends. Through this meticulous approach to data preparation, the project ensured that subsequent visualizations and analyses would be both accurate and informative.

Furthermore, meticulous data cleaning procedures were undertaken to rectify missing or erroneous values within one of the tables. Null entries were systematically replaced with zero values, mitigating the risk of biased analysis, and ensuring the integrity of subsequent visualizations. By adopting rigorous data cleaning protocols, the project upheld the standards of accuracy and reliability, thereby enhancing the credibility of findings derived from the visualized data. This meticulous attention to detail underscores the project's commitment to producing actionable insights based on sound, high-quality data.

Below attached is the final data source I have created in tableau-



DATASOURCE

Datasets:

1. **IMF data:**

   This dataset provides data of inflation rate of countries over time with country name and then each as columns in the dataset we took. With the help of this we are going to analyze countries with inflation data. The primary data source is from international monetary fund website.

   https://www.imf.org/external/datamapper/PCPIPCH@WEO/OEMDC/ADVEC/WEOWORLD

2. **Inflation over years:**

   Inflation over years is the table that contains global inflation data such as year, global inflation rate, federal funds rate if applied, business cycle that caused GDP growth and events that affected inflation. The dataset is taken from Investopedia website
   Which is derived from various websites such US inflation calculator, US bureau of economic analysis etc.

   https://www.investopedia.com/inflation-rate-by-year-7253832

3. **Population:**

   This dataset is taken from world bank data website but the primary data source is mainly UN population data and some other websites. It basically contains the population of countries with regions also over the years.

   https://data.worldbank.org/indicator/SP.POP.TOTL?locations=1W

4. **GDP:**

   The primary source for this data is world bank data website and it contains the GDP of countries with regions also over the years.

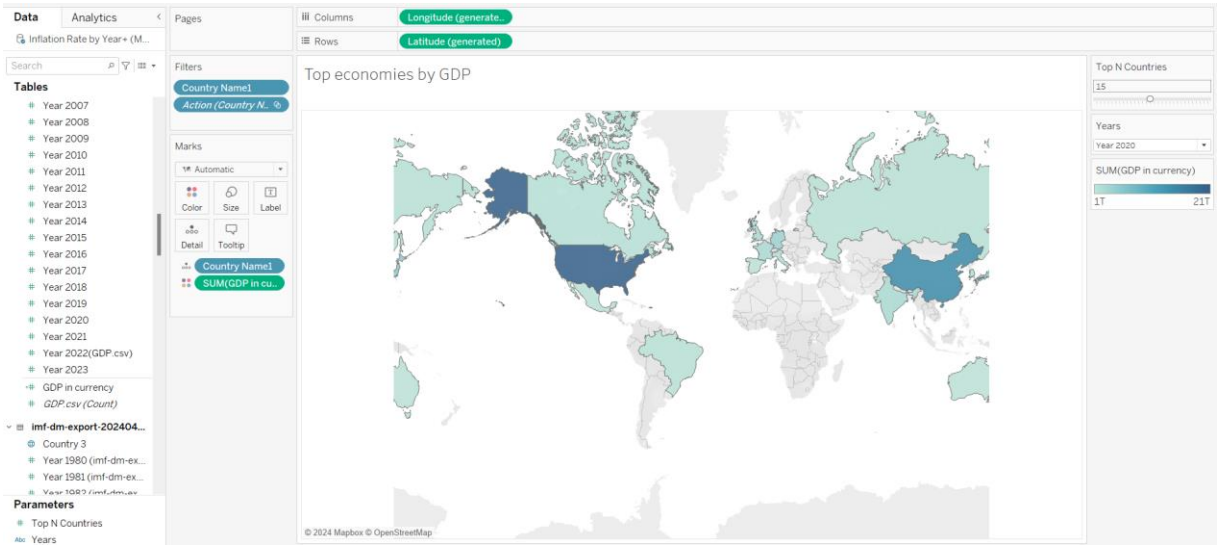
   https://wits.worldbank.org/CountryProfile/en/country/by-country/startyear/ltst/endyear/ltst/indicator/NY-GDP-MKTP-CD

5. **GDP inflation-controlled data:**

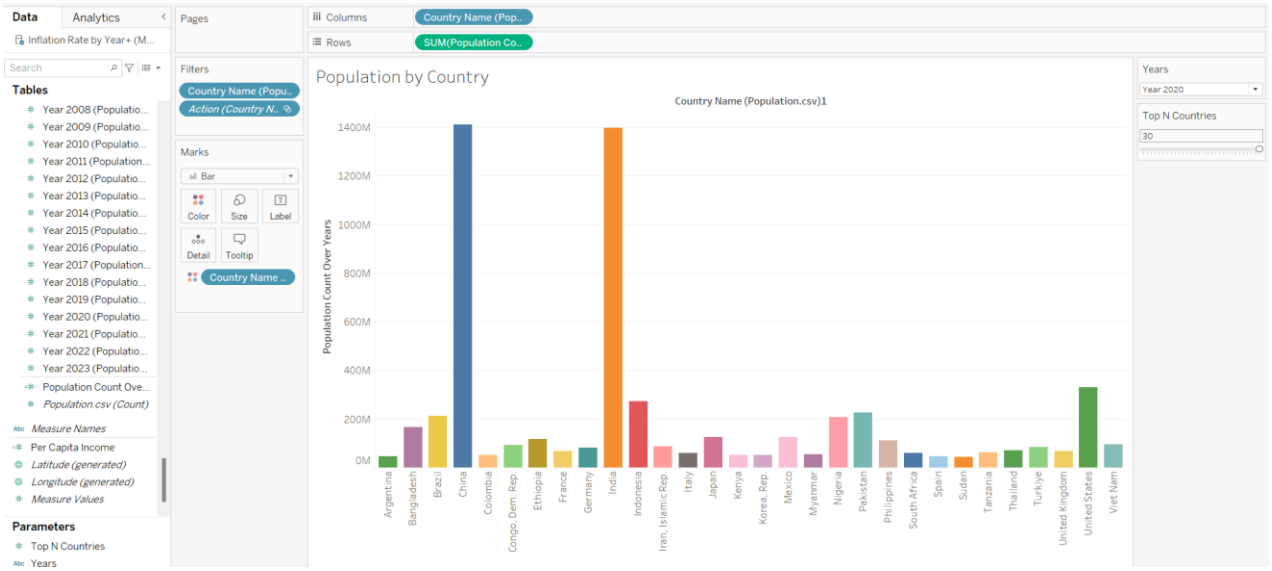
   It contains the data of different economic factors such as GDP deflator, inflation percent, global GDP percent of countries across different years. I took it from the Kaggle website.
   https://www.kaggle.com/datasets/steelcrossx/gdp-inflation

Analysis:

1.Map showing top economies by GDP in a selected year using GDP data set.

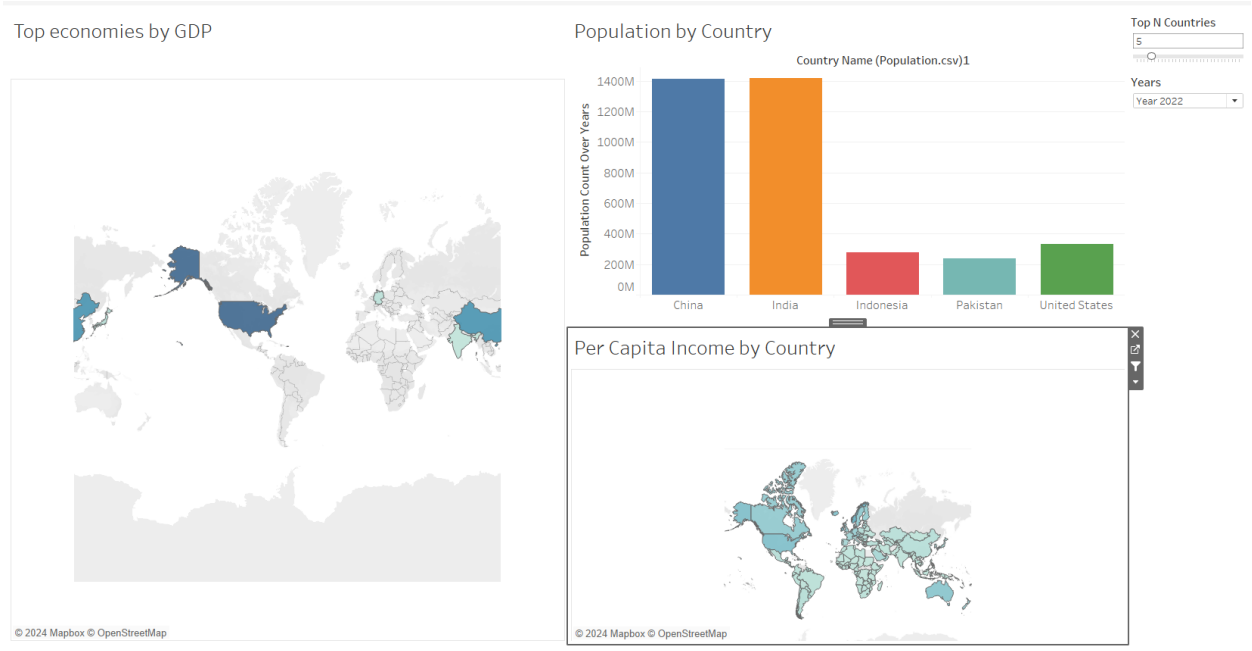


2. This bar graph shows the population of top N countries using data from population data set.

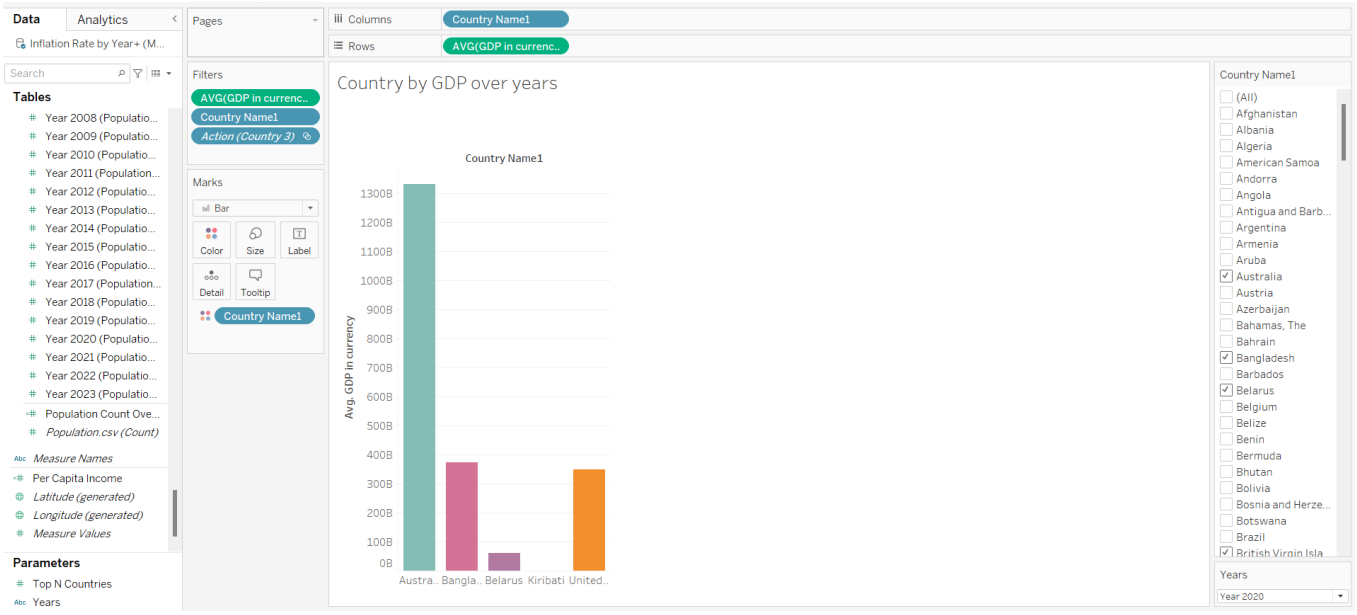


3. This chart shows the per capita income by country and region as we are using a geographical map with the help of GDP and population data sets.

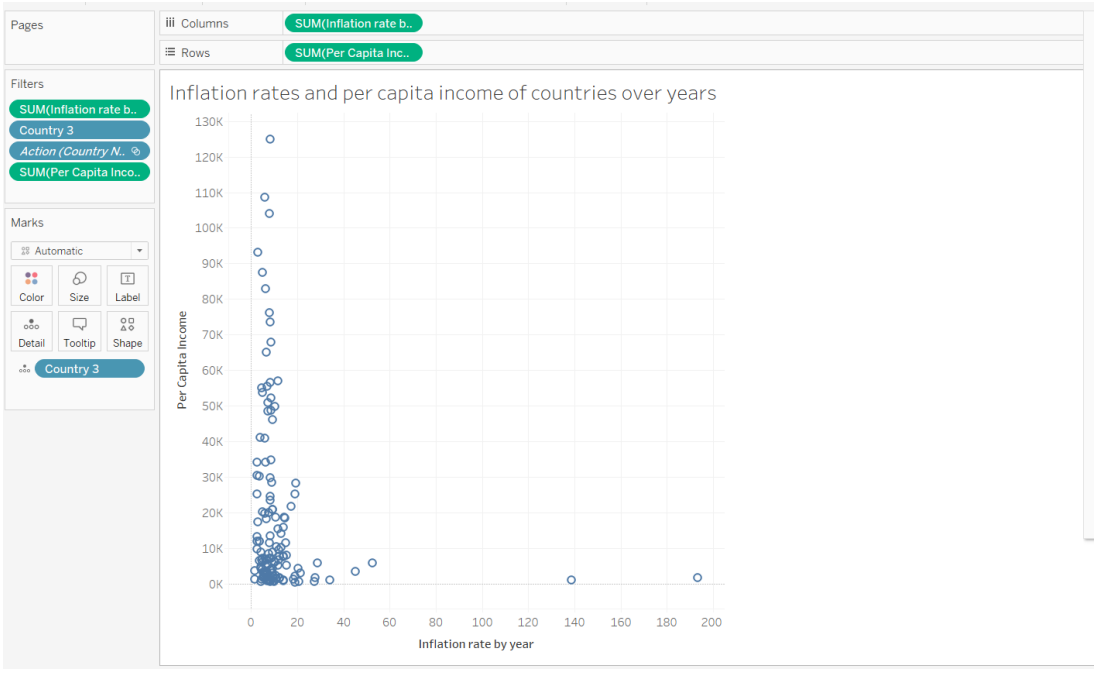
4. This dashboard uses previous three visualizations to find relation between GDP, population, and per capita income of top countries in their category.



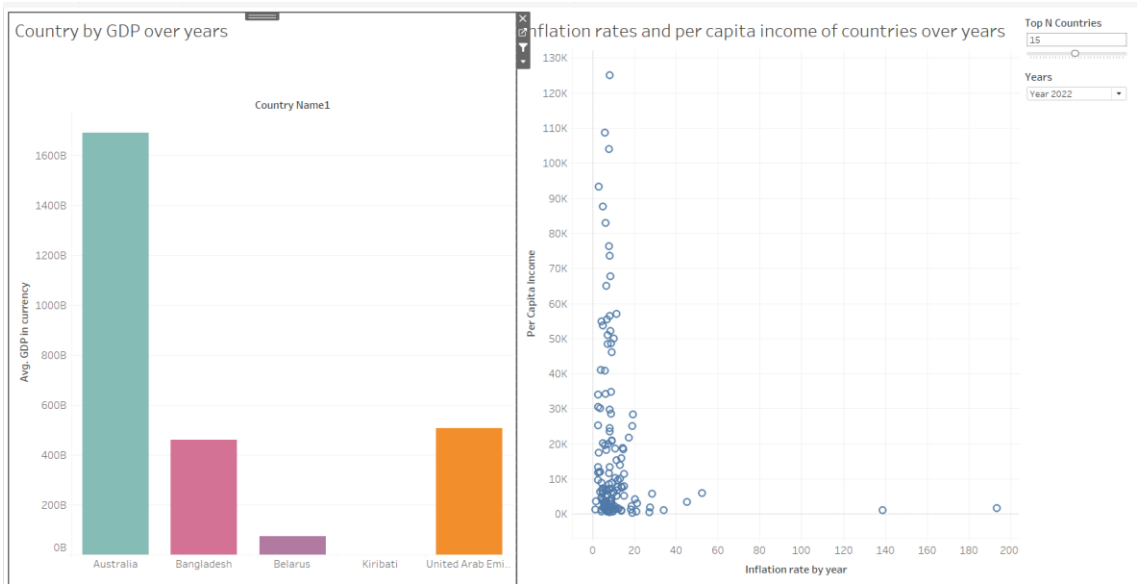
5. This visualization shows GDP of all countries in a particular year. We can select countries of our choice to be displayed. This uses GDP data set.



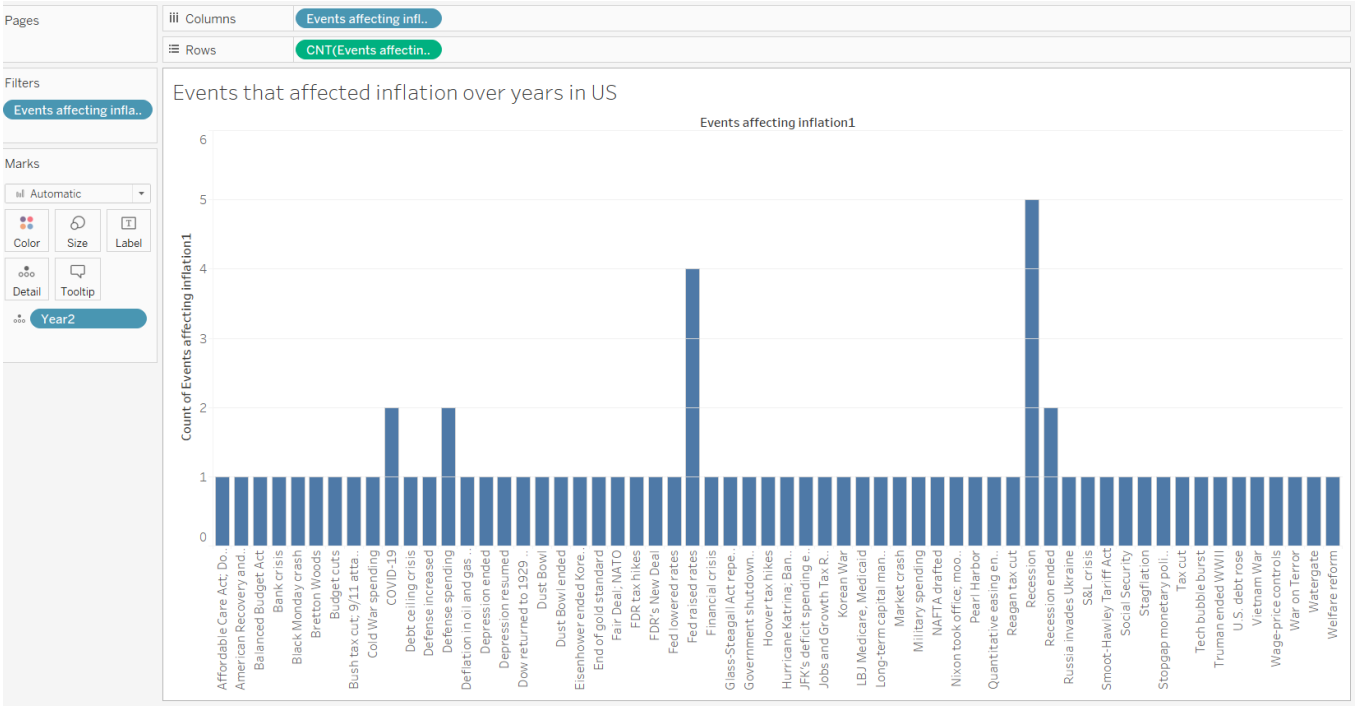
6. This scatter plot gives the inflation rates and per capita income of different countries in a particular year selected using GDP, population, and IMF data sets.



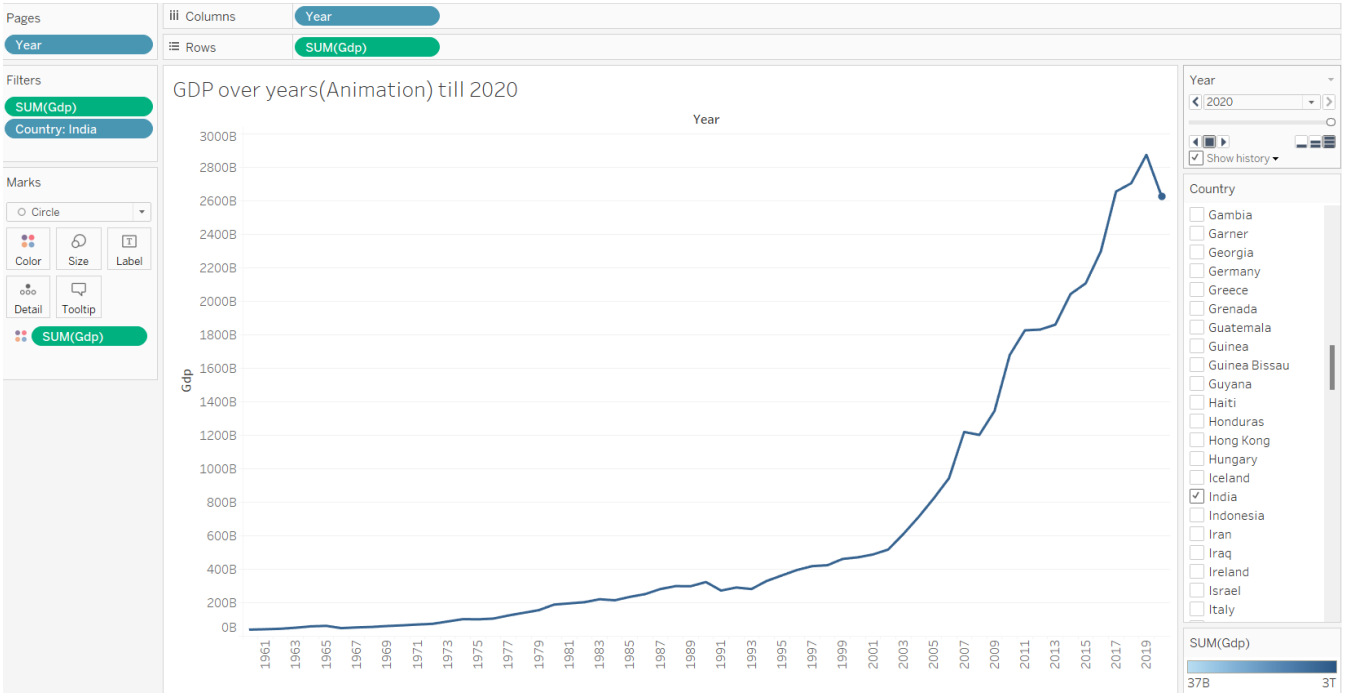
7. This dashboard uses the previous two visualizations to know GDP, inflation rates and per capita income of countries of our choice in a year from the data sets GDP, IMF, and population data sets.



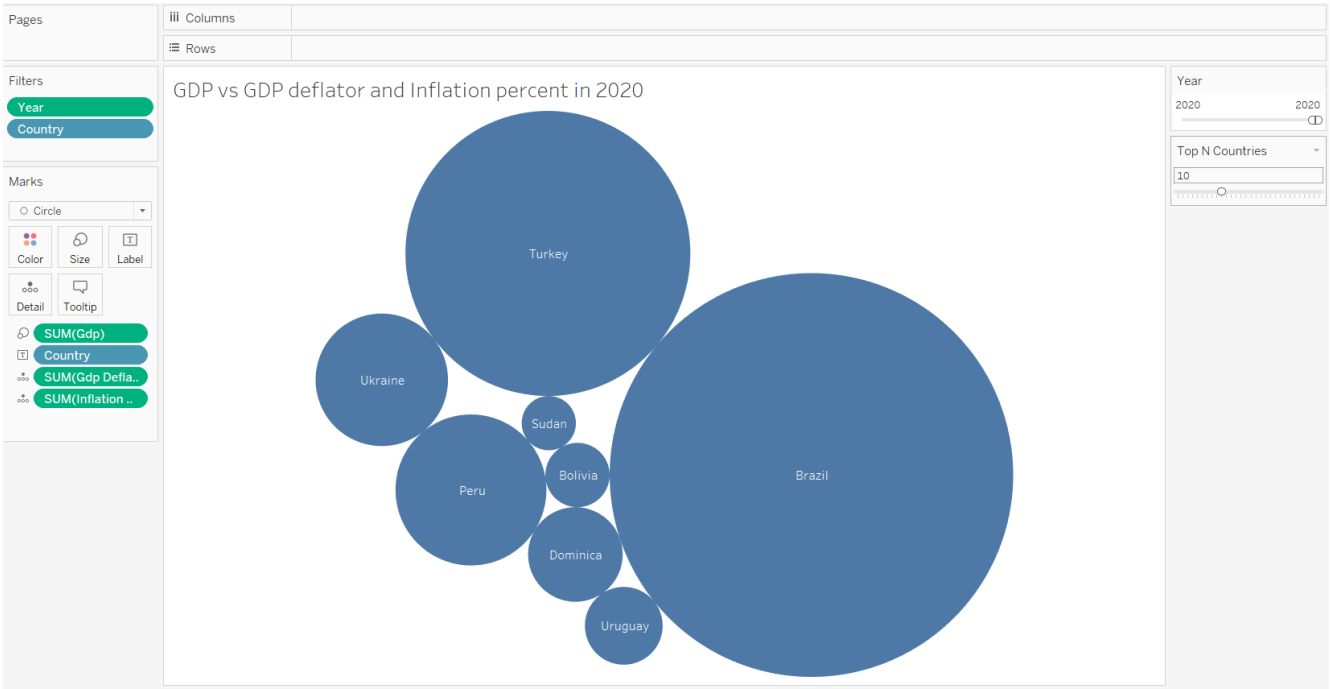
8. Visualization showing events that affected inflation in US in a wide range of years and number of times a particular event caused inflation to vary using the inflation over years data set and count of events calculated field.



9. Animation showing GDP of single country (we can select whole world or a continent's countries according to the requirement). I used GDP, country and year columns from GDP controlled data set.



10. Graph showing all related terms GDP deflator which influences both GDP and inflation across different countries. It is derived from GDP controlled data set.



Conclusion:

From the visualizations I succeed to make, they helped me as follows in answering research questions I formed at first-

R1) The interactive map using **visualization** 1 enables to represent data as it gives easy interpretation of data for my research question which is checking for distribution of top economies among continents. Secondly, the parameter top N helps in finding top economies with the country name field in filter. The interpretation for research question is – **India entered top 5 economies list in 2022 and there is no concentration of top economies in particular continent.**

R2) Visualization 2 helps in understanding top populated countries whether the rank is changing over years and does the most populated countries have higher GDP. **The top most populated country does not vary much over recent years.**

R3) In visualization 3, I used a calculated field called per capita income to normalize data for proper data insights and calculated it using the formula GDP/population. I took bar chart for the same reason I mentioned above. But I did not use top N parameter as I want to use this visualization in the dashboard and do not want to miss data by restricting it here. For example, if Egypt is in top economy but does not fall in top 30 per capita income, we cannot get its data. I found answer to the third research question - **per capita income of most of the countries in Africa, Asia, and South America is less than that of other continents** (as of data we have i.e., we do not have data of Antarcitica).

R4) The dashboard which is visualization 4 implies that there is no direct correlation between GDP and Population i.e., but in 2022 there are 3 top populated countries among top 5 economies and there is no necessity like higher populated economies will have higher GDP if we see look at the whole countries. Surprisingly, **the top 5 economies do not come under the top per capita income countries.** To see this, we must include top N factor in visualization 3 and see it in dashboard.

R5) In this research we just seeing at the data for any given country. The data is inflation and per capita income of country shown in scatter plot. This uses visualization 6

R6) For this research questions, I have used visualization 5 and visualization 6 and combined them into visualization 7 as dashboard. For a selected country it shows the GDP, inflation rate and per capita income during that particular year. From this we can see the data of each country we wish to know unlike dashboard which only compares top economies.

R7) US inflation is most affected by two events at different times. Those are fed rates raise and recession caused inflation change at different periods. As we covid pandemic adversely affected whole world. US seen inflation change because of it in the continuous years 2020 and 2021. We can see this in the visualization 8.

R8) Visualization 9 which is a animation justifies the 8th research question that shows the trend of GDP raise over the years globally and country wise as well. By selecting multiple countries, for example selecting all countries in Asia we can see how the GDP trend is going on in any continent.

R9) Visualization compares the top countries by GDP deflator. We can see that Brazil has more change in price levels in 2020 that is controlling its own GDP and inflation.

Future scope or other research questions:

- Future scope can be we can include political stability data and socio-economic data such as unemployment and other factors affecting the dynamics of country's GDP and inflation.

- We can also combine climatic data which might have affected GDP because of natural calamities. For example, Japan is such a country which faces tsunamis frequently. Simillarly, US faces hurricanes many times a year. So, we can research if it affecting any tourism revenue of the country and check if it really has impact on GDP of that country. Because, small countries like Maldives primarily rely on tourism.