**Project Increment 2**

**Video Link :-** <https://drive.google.com/file/d/1s3Rtx7Y1BONFGPRtrsnYNhKuJyWsHFJe/view?usp=sharing>

**Github :-** <https://github.com/AnivChakravarty/global_inflation_viz>

• **INTRODUCTION:-**

o **Domain -**

Project aims at covering the geospatial aspects of inflation at a global level and tries to provide users with an interface to understand the inflation data for various countries and sectors for the years 1970 through 2022. There are a total of 211 countries that need to be represented throughout all the years. The main form of data is quantitative and categorical data. The plots used are map plots, pie charts, bar plots, line plots and bubble plots.

o **Workflow diagram with explanation -**

WorkFlow:-

We first start with cleaning the dataset. Once the data is ready we perform exploratory analysis on different visualization forms to pick a set of visualizations to have in the application. Figure 1 shows the workflow of our project plan. The story boarding process in the user interface and experience would determine what the user sees and how the user interacts with the main dashboard and respective features as they are added and would be tied to development. We will build a core set of features for the initial phase 1 of the project so that the app works first, we then proceed with the iterative cycles to add in further features. The analysis portion would be to determine which visualizations provide the details in an understandable format in order to avoid excessive clutter on the dashboard. We plan to first start with barcharts, spheres, scatterplots and then branch out to more complex plots.

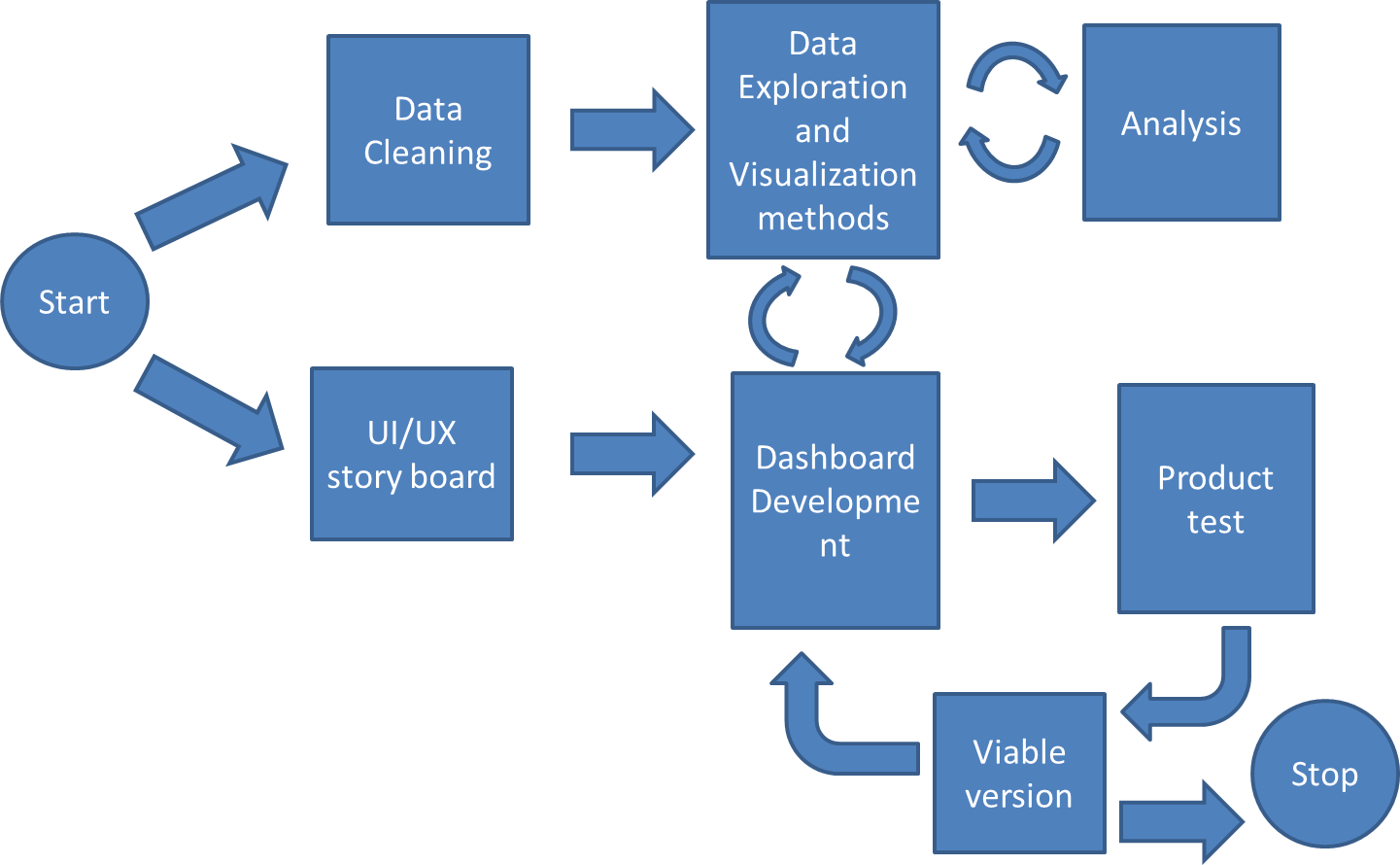


Figure 1: Project Workflow

In terms of data workflow, our dashboard will be fetching quantitative data from a storage repository while having an interactive module and sub visualizations that would form supplementary visualizations on context based on Figure 2. While this is an initial proposal our workflow may change based on how our ui/ux storyboard goes through and update based on that.

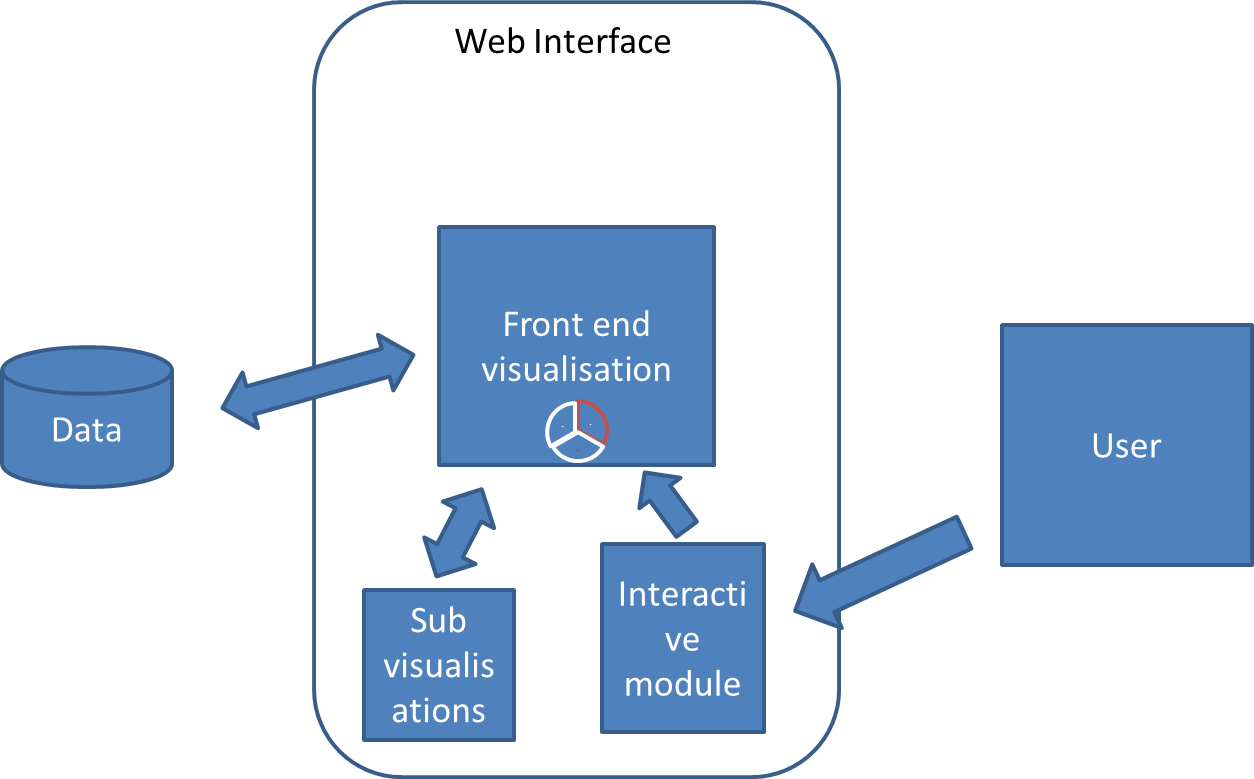


Figure 2: System Workflow

• **Background**

o Related work for your topic with linked references

Starting with the kaggle repository [[1](https://www.kaggle.com/datasets/belayethossainds/global-inflation-dataset-212-country-19702022)] for the dataset, there are numerous tutorials on kaggle for exploring the dataset. We observed various different plots created visualizing different aspects of the inflation data through different line, bar and geographical plots [[2](https://inflationdata.com/inflation/images/)] as well as customized line plots for depicting income as in [[3](https://www.anychart.com/blog/2021/11/12/visualizing-data-prices-emissions-workforce-palm-oil/#more-13940)]. The closest dashboard we found online was from the international monetary fund [[4](https://www.imf.org/external/datamapper/PCPIPCH@WEO/OEMDC/ADVEC/WEOWORLD)] which provides an interactive interface for geographic heatmap with a line scale trend for displaying average economic data but does not break down by different economic sectors and visual and the global inflation rate tracker [[5](https://www.ft.com/content/088d3368-bb8b-4ff3-9df7-a7680d4d81b2)] which made use of line and geographic plots for fixed years and a few major countries. We aim to provide a dashboard with few more visualizations compared to these. These visualizations do not provide much room to drill down and drill up on the data with limited interactivity in terms of mouse clicks. While these front end dashboards provide great visualizations our project aims to provide a more interactive approach with more ways to view the data.

• **Data Abstraction**

In order to make the data easier to read for time series the data is pivot and grouped for years.

This gives us an easier dataset to work with.

We abstract with a world map visualization with a choropleth map where the values are the average inflation of each country for that year.

o **Dataset (Type and Attributes)**

The details of the dataset in terms of types (Figure 4) consists of string countries and float inflation rate values. Categorical data for sub category names and some notes at the end. The derived aggregate dataset (Figure 6) consists of nominal data of countries and quantitative value of type float representing the inflation rate and integer id value.

o **Detailed description of Dataset**

Our data consists of numerical data across multiple sub categories and years. It is made public on kaggle and is free to download and explore. It is updated by [worldbanks.org](http://worldbanks.org) at a frequency of 2 months.

Our data (Figure 3) consists of

1. Country code: string, categorical data type holds the country code of the country
2. IMF Country code: float64 numeric value representing the country code.
3. Series Name: This field represents string, categorical data to the sub category domains of inflation
4. 1970-2022: these float64 fields represent the values among the countries and subcategories through the years 1970 to 2022
5. Notes: these are string data of extra notes for the subcategories.

Data abstraction is done by first getting the average of inflation throughout all the subcategories for each country for each year.



Figure 3: Dataset header

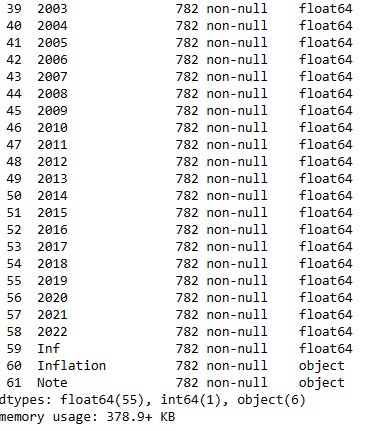
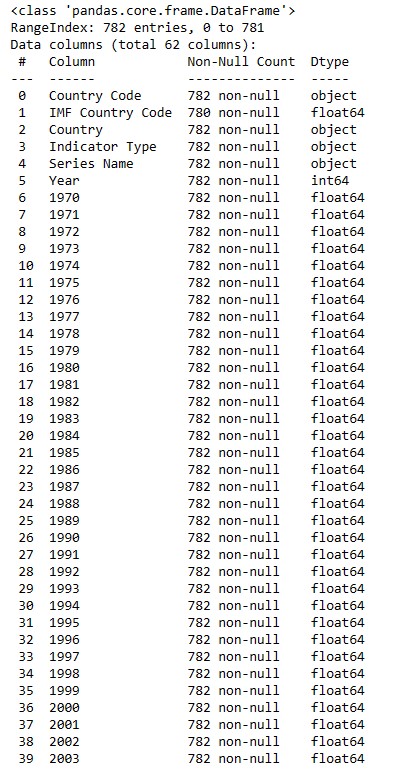


Figure 4: Dataset description

From the dataset the average inflation rate for each country is calculated and stored in a dataframe (Figure 5) which is mainly used for top view visualizations which consist of countries, a unique id from 1 to 211 and the average value for the years.

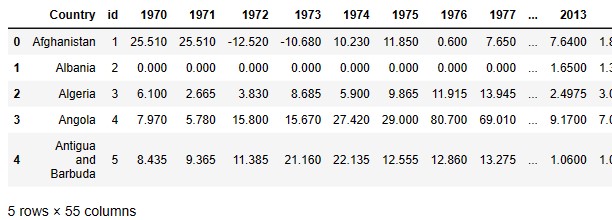


Figure 5: Average inflation rate by country

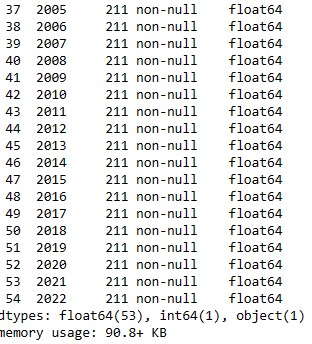
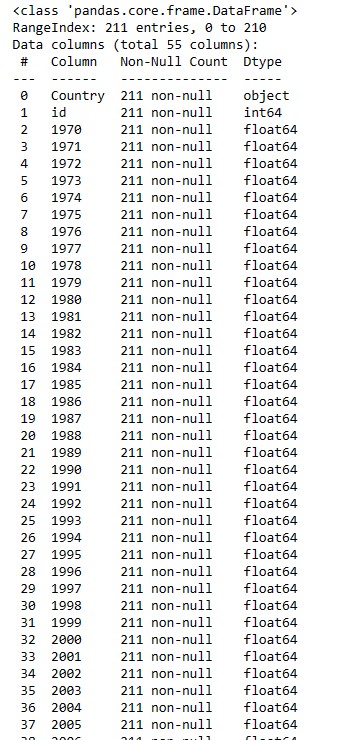


Figure 6: Average inflation rate dataset description

o **Detail design of Features with diagram**

The dataset [[1](https://www.kaggle.com/datasets/belayethossainds/global-inflation-dataset-212-country-19702022)] contains inflation data of over 206 countries in the time frame of 1970 to 2022 across four domains of consumer, producer, food, energy costs. This allows us a lot of room in terms of visualization of the data. So we plan to try and incorporate the feature listed below in our tool in an iterative method starting with core visualization parts first and adding features on top.

* Drill up and drill down options
* Navigation
* Interactive visualization
* Dashboard user interface
* Dynamic visualization changer
* Annotation options
* Multiple panels
* Colorblind features
* Customization options
* Web application
* Feature engineering options to perform grouping, numerical operations on the data.
* Animations
* Print report

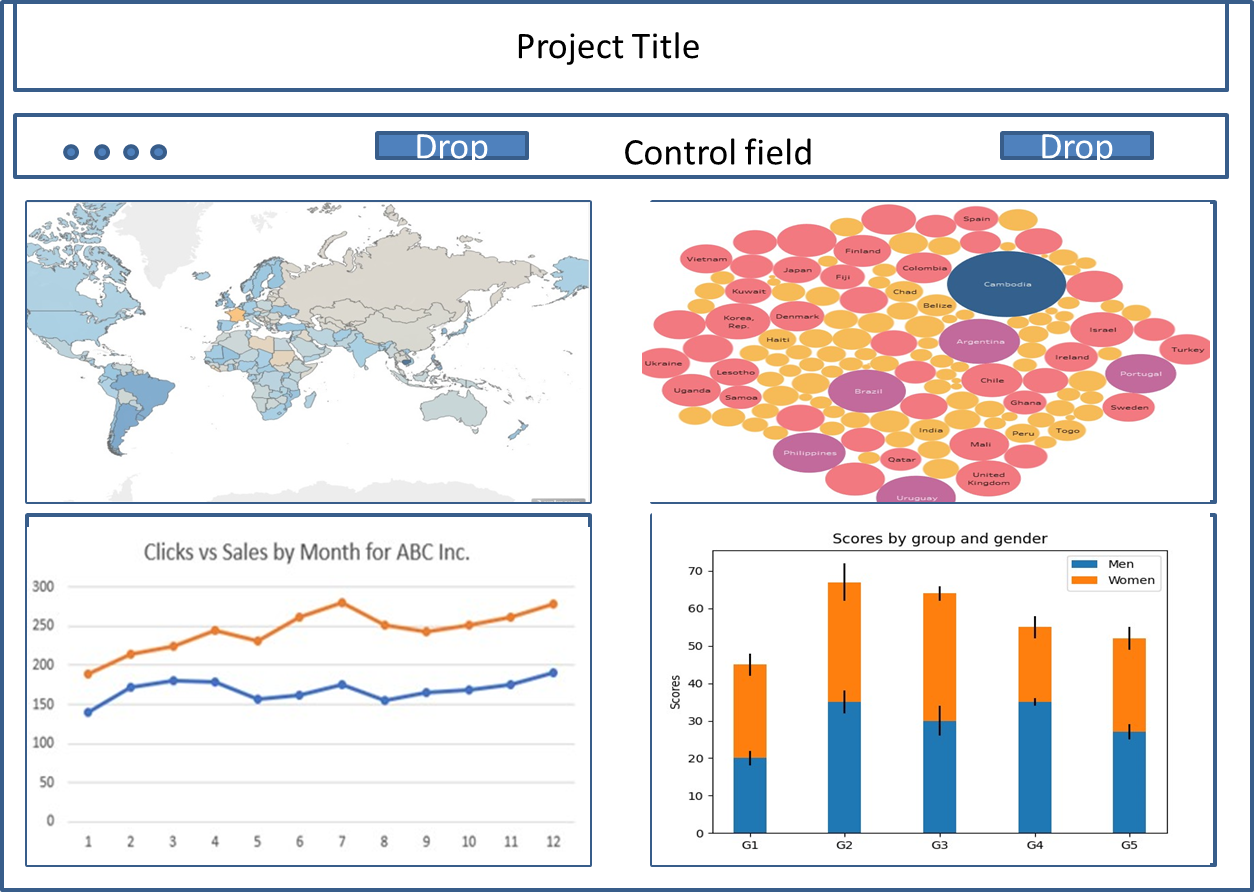


Figure 7: Dashboard design prototype

Initially we plan of having a web browser based dashboard as Figure 7 with a primary world map based spatial choropleth with additional plots on windows for added data refinement.

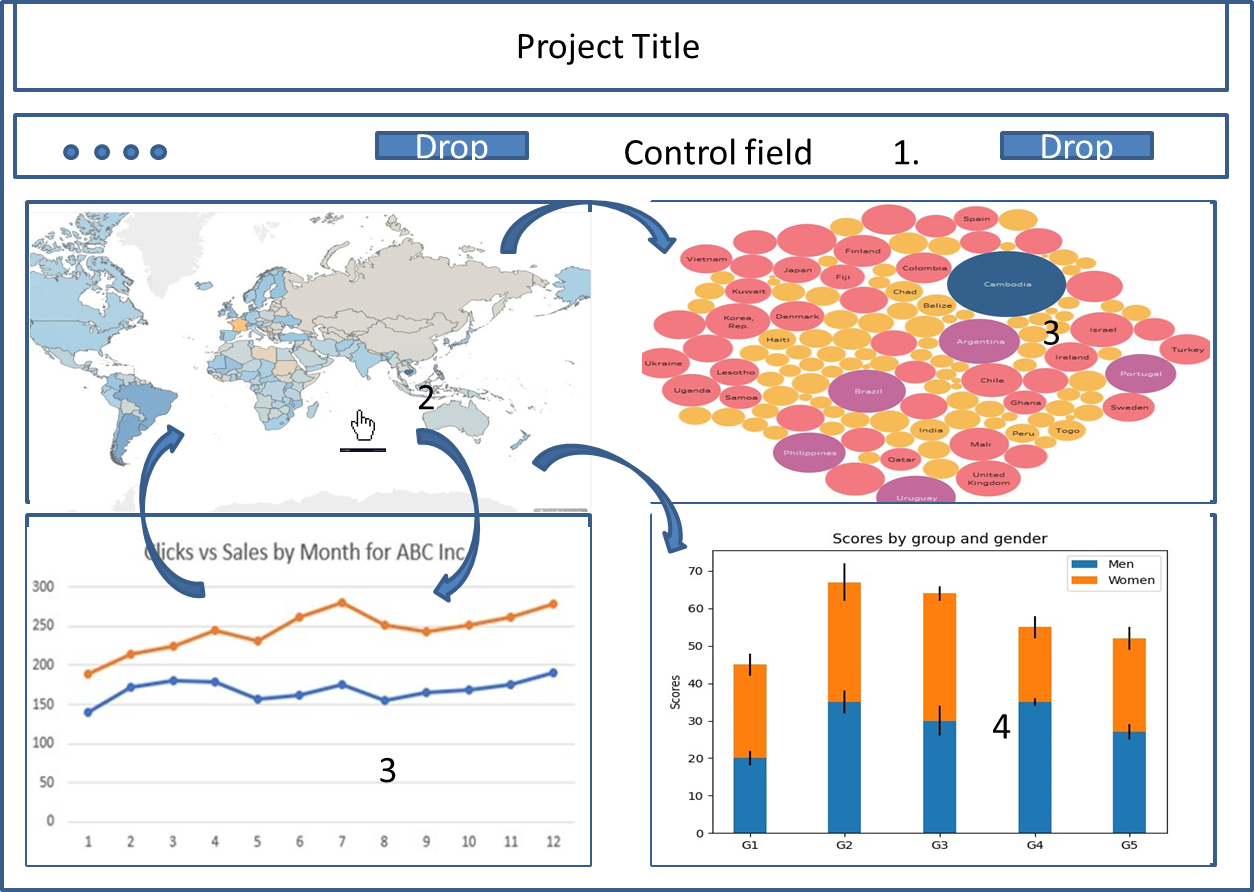


Figure 8: storyboard interaction view of dashboard

Figure 8 represents how the user would interact with the dashboard. The user starts with loading a map year from the drop down in step 1. This populates the map and starts the interaction of the map in step 2. The time plot populates and can be used to update the map as well from step 3. Step 4 and 5 populate from actions of 1,2.

o **Data Transformation**

For Data transformation we converted the country names to ids from 1 to 211 apart from the standard country ids. The columns are transformed to an array of years and stored as a list and fetched from the dataframe. When aggregating the average value any null values are either dropped if the entire data row or column is null or converted to 0.

• **Task Abstraction**

Our plan was to initially build the ui dashboard and embed our visualizations into the dashboard windows.

**Task 1**: Geographical map with inflation rates

Description: Generate an interactive world map that shows inflation rates by year

Action: the map is a spatial choropleth map from 1970 to 2022 using a reusable plot using a dropdown for changing map plots.

**Task 2**: Generate a pie chart for the top 10 countries where inflation is the most

Description: Build a pie chart to visualize the top 10 countries where inflation is high.

Action: interactivity needs to be able to show the information on hover.

**Task 3**: Build a bar chart for top 10 high inflation countries.

Description: The bar chart is a visual for top 10 countries where inflation is high.

Action:Based on the inflation rates obtained for a particular year, we create a bar chart to view the amount of inflation in the top 10 countries.

In terms of data processing our tasks are

**Task 4**: Clean data

Description: The dataset consists of extra fields that hold no value or empty values.

Action: Clearing out unnecessary value, pivot and grouping data as well as aggregating average fields of each country while replacing null average to 0.

For Dashboard UI,

**Task 5**: build a dashboard

Description: we need to build a dashboard that can be accessed via web browser.

Action: We made use of the python dash library to build our dashboard in a jupyter notebook that is accessed via browser url.

o Task (Target and Actions)

**Task 1**: Geographical map with inflation rates

Target: Identify what is the average inflation rate in the world for a year

Action: Have a choropleth map with boundary marks colored with a continuous spectrum color pallet in order to identify outliers and high inflation rate increases and their respective country.

**Task 2**: Generate a pie chart for the top 10 countries where inflation is the most

Target: Identify the top 10 countries with inflation

Action: For this task we initially have a pie chart to represent this data but later also incorporated a spiral bubble plot in order to clear up pie charts for another purpose.

Interactivity through tooltips were added to all plots

**Task 3**: Build a bar chart for top 10 high inflation countries.

Target: Show detailed level of the inflation rate in a bar plot.

Action: The bar plot on a positive and negative axis allows us to filter and look at the inflation rate across countries but we limit the plot to 10 for preventing excessive clutter.

**Task 4**: Clean data

Target: Have valid data in the dataset and no null values

Action: We loaded the dataset csv in pandas and made use of pandas operations and cleaned null values and created another dataset having the aggregate average value of each country for each year.

**Task 5**: build a dashboard

Target: Have a web accessible dashboard for the plots.

Action: We made use of the python dashboard library and plotly in order to build the dashboard and test on a local server.

• **Implementation using tools**

The tools we used for the purpose of this project are jupyter notebook dashboard, python libraries like plotly for visualizations and certain data processing libraries like pandas and numpy.

o **Describe the tools used for each visualization**

1. Choropleth map :- In order to visualize inflation data for a selected year on a world map, we made use of the choropleth visualization offered by python’s ‘plotly’ library. In the function to create a choropleth map, we pass countries as locations and inflation as our labels for the selected year. We’ve set the projection of the choropleth map as equirectangular to attain a better understanding of the data in view in a 2D plane.

2. Pie chart :- This pie-chart informs the user on the top 10 countries with highest inflation rates for the particular year that has been selected. We again make use of ‘pie’ method offered by ‘plotly’ and pass Country names with top 10 highly inflated countries to the function.

3. Bar chart :- This inverted bar chart informs the user about the countries where the inflation has been lowest for the selected year. Plotly’s bar method enables plotting of such a bar chart.

4. Spiral Bubble plot :- Finally, we conclude the dashboard by plotting a bubble plot. Interesting thing about the bubble plot is that it shows data with top 20 countries in accordance to the inflation rate and the bubbles spiral with the least inflated countries at the center and the most inflated country at the end of spiral for the selected year.

• **Results for Analysis**

From the data and choropleth plots from 1970 to 1980 we observed that the countries of Brazil, Argentina, Congo, Sudan have had a continuous increase in inflation rate.Chile saw high inflation rates in the 70s. Turkey and Israel had spikes in the 80s along with Laos and Vietnam. Russia and most of Asia saw Increase in rate in the late 90s and Zimbabwe saw a constant increase in inflation rate in the 2000s. 2009 and 2010 saw worldwide increase in inflation rates. The 2010s saw an increase in inflation rates in Ethiopia, Iran and Venezuela. The inflation rates of the USA, Canada, China and most Asian countries have remained relatively steady compared to the other countries mentioned above.

o Visualization graphs with detailed explanation

The key visualization plots start with the choropleth map plots using the magma color scheme (Figure 9).

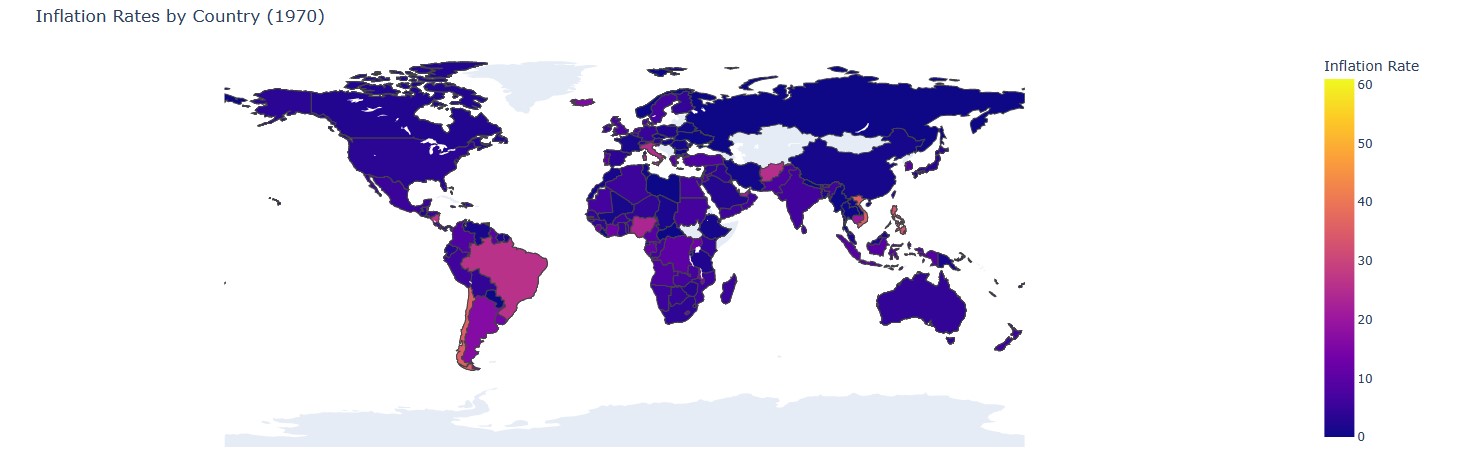


Figure 9: Choropleth

The next plot is the pie plot (Figure 10) consisting of the countries of highest inflation rate.

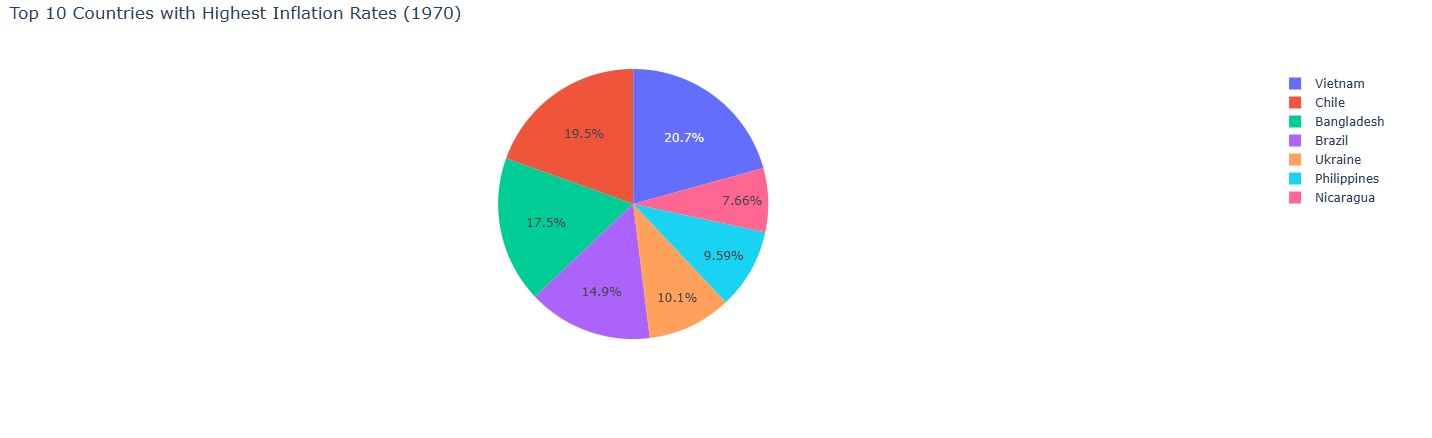


Figure 10: Pie Chart plot

The next plot is the bar plot representation shown in Figure 11.

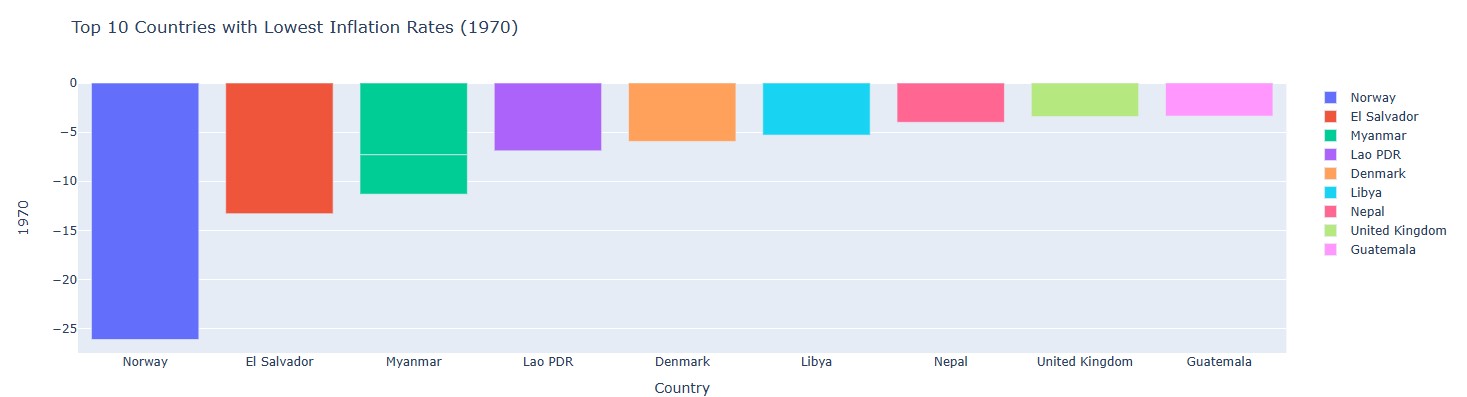


Figure 11: Bar plot

Next plot is the spiral bubble plot with circle marks that vary with size proportional to the value of the inflation rate (Figure 12).

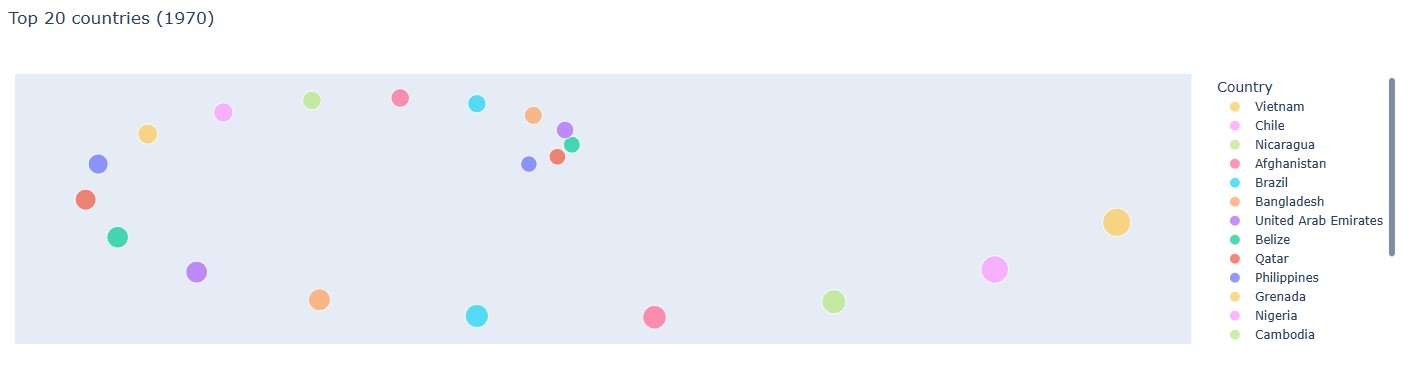


Figure 12: Spiral Plot

**Storytelling -**

Chapter 1: Prices going up:-

Inflation affects everyone no matter which country you live in. Whether it's the rent you pay or the gas you put in your car or the groceries you buy. Prices of goods and services are directly related in one way or another to inflation. But what causes inflation? Global events tend to cause a cascading effect as well as how your national government runs the country through its bills. An example would be the supply of oil through Europe from Russia or the export of grain from Ukraine were hit due to the war that caused global inflation in prices globally but more so in certain countries. So each country has its own way of handling how its resources are managed in terms of imports and exports as well as economic bills being passed to determine the rate at which prices rise or fall. So far we know who is affected by inflation (everyone), how it affects people (rising price of goods and services) and why (managing a nation's resources in terms of a nation's growth). This data represents the current condition based on the various economic sections which provides those in the government bodies or economic power to understand and make necessary decisions through bills or motions to navigate a country's economy. When provided with data from multiple years we start to notice trends and patterns and better set contingency plans or divert resources to avoid disasters. But merely looking at bug numbers and large portions of tabular data isnt as intuitive as visualizations. Hence, we decided to build an interactive dashboard with visualizations to better ingest the data and analyze the overall pattern to find meaning in the numbers.

CHAPTER2: Let's look at the data:-

When we think of leveraging data to capture the information that is generated in millions of bits every second, one thing that quickly comes to our mind is sampling and collecting. Based on our interest, we can decide on how frequently we want to sample the data. The data which is of interest to us in the context of this project is about Inflation on a global level. Inflation being an entity which sits on top of the gross domestic product, naturally needs to be sampled at a lower frequency than GDP. Sampling yearly can be one such suitable frequency to record the inflation. Also, another aspect of sampling that needs to be considered when it comes to terms like Gross Domestic Product and Inflation rate is the geography that we want to limit it to. In this case, the geographic limit is the entire globe. The data has been sampled for all the countries throughout the globe. Inflation, a result of the economic decline of the country, can provide valuable insights to government bodies. The data has been collected over the timeline of 52 years starting from 1970 through 2022. Moreover, the dataset contains countries and various sectors affecting inflation rate, hence it is possible to drill down to a country’s cause of inflation. To answer the question about how? Let's take an example of The Great Depression event of 2009. This event impacted the world economy and central banks went broke due to their massive corporate debts. This consequently increased the inflation rate in major parts of the world and this can be observed from the dashboard that’s been created.

Who: Economists and Data Scientists

What: Inflation rate increase across domains

Where: Data collected from all countries

Why: To keep track and cover trends

When: 1970 to 2022

How: Data collected for official sources from government

Chapter 3: See the trends:-

Our dashboard design is mainly aiming to help economists as well as normal people to understand the trends in the data. While we do not have in-depth domain knowledge in economics, the intuitive use of the world map with the choropleth gives an overview of the average inflation rates through the use of channels on country shapes. The pie charts help show parts of whole contextual visualizations of more fine grain data of inflation rates of the different economic sectors of a selected country along with bubble plots for year based country visualization, a bar plot for a more accurate rendition of the pie plots and a line plot covering the time frame for showing trends. But choropleths and line plots are already used effectively for plots. Our aim is to provide a more intuitive interactive experience for visualizing plots and navigation while keeping it as simple as possible on the front end. By making it a browser dashboard anyone with an internet browser should be able to open and view the plots on the data to gain useful insight for both exploratory as well as analytical purposes. The addition of animations along with the interactivity and the ability to save and download your plots provide further benefits. Finally as mentioned these plots provide valuable insight on trends of global inflation as well as the countries affected year over year allowing governments and officials to make necessary decisions as well as informing the general consumer of the sectors being hit the hardest and in which region on a global scale.

Who: The general public and officials will be looking at this dashboard

What: Choropleth, pie, bar, line and spiral plots

Where: Online web portal so any device that can use a browser

Why: To identify trends in the inflation rates on a global scale for discovery and analysis

When: Anytime the server is accessible

How: With a stable internet connection the person logs into the browser to view the plots

• **Project Management**

In terms of our progress we provide two sections of work completed and work to be completed along with the respective person who completed said work or is responsible for completing it in brackets next to the task. Given that we have a working dashboard with some of the key plots being visualized we can start to focus a bit on interactivity while wrapping up a few remaining design points.

o **Implementation status report**

§ **Work completed:**

* Data cleaning: Data cleaning involves removal of unnecessary columns, grouping and preparation of aggregate values of inflation rate.(Aniv)
* Visualization exploration: After looking through multiple visuals we saw choropleth global map, line plots were commonly used for this type of spatial data. Along with these plots we saw additional plots of bubble plots, bar plots, boxplots can also be used. Further we saw that pie charts can also be used to represent the various subcategories.(Aniv)
* Designing a storyboard and interaction plan as well as basic dashboard design.(Aniv)
* Building dashboard(Shridhar)
* Testing web accessibility(Shridhar)
* Dynamic plots(Shridhar)
* Save plots(Shridhar)
* Building spiral bubble plot(Aniv)

• **Description**

We managed to complete building the web dashboard accessed on a test local server through the jupyter notebook and have rudimentary interactivity through linking the plots with a drop down list on the different years and update all the plots with the data and have tooltips to further show details on the marks. Plotly also allowed us to apply zoom features and download the plots with the command bar on the plots. The data was also cleaned and processed to allow us further flexibility on visualizations.

• Responsibility (Task, Person)

Aniv Chakravarty ([anivchakravarty@my.unt.edu](mailto:anivchakravarty@my.unt.edu)):-

* Data Cleaning
* Analyzing data using basic visualizations.
* Researching possible visualization strategies for dashboard.
* Implementation of certain parts of dashboard

Shridhar Kshirsagar ([shridharkshirsagar@my.unt.edu](mailto:shridharkshirsagar@my.unt.edu)):-

* Research on data to be presented to the user via dashboard which will provide meaningful domain/business insights.
* Making the dashboard more interactive and user friendly.
* Implementation of Dashboard using a programming language.

• **Contributions (members/percentage)**

Aniv Chakravarty - 50% contribution with plots, workflows, integration and documentation.

Shridhar Kshirsagar - 50% contribution with dashboard development, web execution, testing, integration and documentation .

**Future Work and Scope**

There are more interactive features that could be worked on to improve interactivity and create a pdf report and add more filters. While we tried to implement some more of these features, the limitations of the library required more time to create new approaches and troubleshooting to get to work. Overall we got the core features needed to analyze the plots first.

• References/Bibliography

References:

[1]. “Global Inflation Dataset - (1970~2022).” Kaggle, <https://www.kaggle.com/datasets/belayethossainds/global-inflation-dataset-212-country-19702022>. Accessed 16 April 2023.

[2].“Global inflation tracker: see how your country compares on rising prices.” Financial Times, 12 April 2023, <https://www.ft.com/content/088d3368-bb8b-4ff3-9df7-a7680d4d81b2>. Accessed 16 April 2023.

[3].McMahon, Tim. “Inflation Charts and Graphs.” InflationData.com, 14 March 2023, <https://inflationdata.com/inflation/images/>. Accessed 16 April 2023.

[4].“Visualizing Data on Prices, Emissions, Workforce, Palm Oil | DataViz Weekly.” AnyChart, 12 November 2021, <https://www.anychart.com/blog/2021/11/12/visualizing-data-prices-emissions-workforce-palm-oil/#more-13940>. Accessed 16 April 2023.

[5].“World Economic Outlook (April 2023) - Inflation rate, average consumer prices.” International Monetary Fund, <https://www.imf.org/external/datamapper/PCPIPCH@WEO/OEMDC/ADVEC/WEOWORLD>. Accessed 16 April 2023.