INFO - I590 Data Visualization
Project Report
Travel Visa Inequality

Team Members
Rohit Bapat
Yashaswini Dhatrika

December 17, 2019

1 Abstract

Visa policies have always been restrictive and controlled people from travelling and exploring the world. However, these policies have many times been against the concept of Globalisation which promotes cross-border transactions. Recent developments with respect to immigration, security have largely affected the visa policies of all the countries.The main aim of this project to visualize and analyze the factors that contribute to treating the equal citizens of different countries differently in terms of travelling. The visualizations will help to give insights regarding the factors which distinguish different visa policies across different countries. Interactive visualizations coupled with plots based on country specific ranks will help to drill down and identify the root causes that affect a country's visa policy.

2 Introduction

Through this project, we try to identify if the extent of visa free mobility is constant across the world or not? If not, the distribution of this mobility may explain certain other derivable insights regarding each country's progress and overall economy.

2.1 Motivation

Mobility is a key feature of modern times. Although the movement of people has always been a central feature of human history, it has never been as frequent and widespread as today. The number of people crossing borders

each year has increased about 50 times since World War II[6]. As travelers now can witness at international airports, citizenship and identification papers have a great influence on the ease of travel. While people holding passports of the wealthy or befriended countries can often breeze through custom checks, often using fully automated systems such as "e –gates" citizens of poor and politically fragile countries, particularly from South and South–East Asia, the Middle East and Africa, generally require a travel visa to enter and have to queue up in long rows at the border. This seems to mirror a structural inequality in immigration and travel rights. We find that, on average, visa–free mobility has increased over the past 40 years. However, not everybody has benefited from these developments. In fact, visa waivers are increasingly unequally divided: While citizens of OECD countries and rich countries have gained mobility rights, mobility rights for other regions have stagnated or even diminished, in particular for citizens from African countries[1]. Overall, we find a clear bifurcation in mobility rights, leading to a 'global mobility divide.So this piece of work aims to visualize the inequality of visa restrictions and how it is affecting the mobility of citizens of different countries.[5]

2.2 Background

The visa policies are changing rapidly with an intention to make the travel and exploration easier. However, our personal experiences made us wonder whether these policies actually promote the travel or restrict the same. This was the main motivation behind exploring the related work and datasets available to develop the visualizations.
the United Nations World Tourism Organization (UNWTO)[4] drafted a Global Code of Ethics for Tourism where it has proposed that "The prospect of direct and personal access to the discovery and enjoyment of the planet's resources constitutes a right equally open to all the world's inhabitants." Even though still being in the preliminary stage we are far away from this goal of establishing free travel as a right.
Another reading [7] describes that the freedom of travel for the people is directly affected by the citizenship they have. Additionally, factors like Economy of the home country, infrastructure, presence of consulates always correlate with overall visa policies. This can either encourage or limit the free travel opportunities.
Travel now is a large of share of global economy contributing to almost

9.5%.[7] The visa policies have actually restricted the travel economy considering the sovereignty of every country and the decisions taken by the respective governments.
Some of our inspirations are work done by Chritian Laessar and website (passportindex.org)

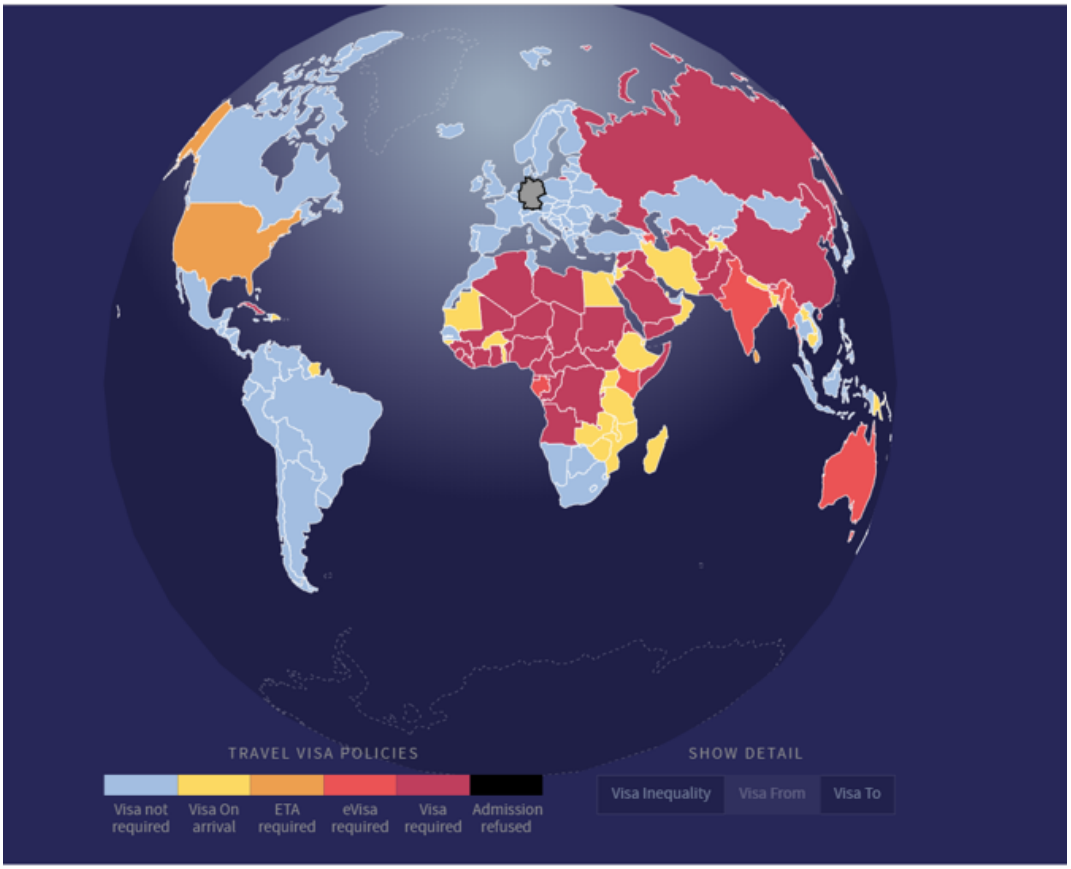

Figure 1: Country Specific Visa Policies

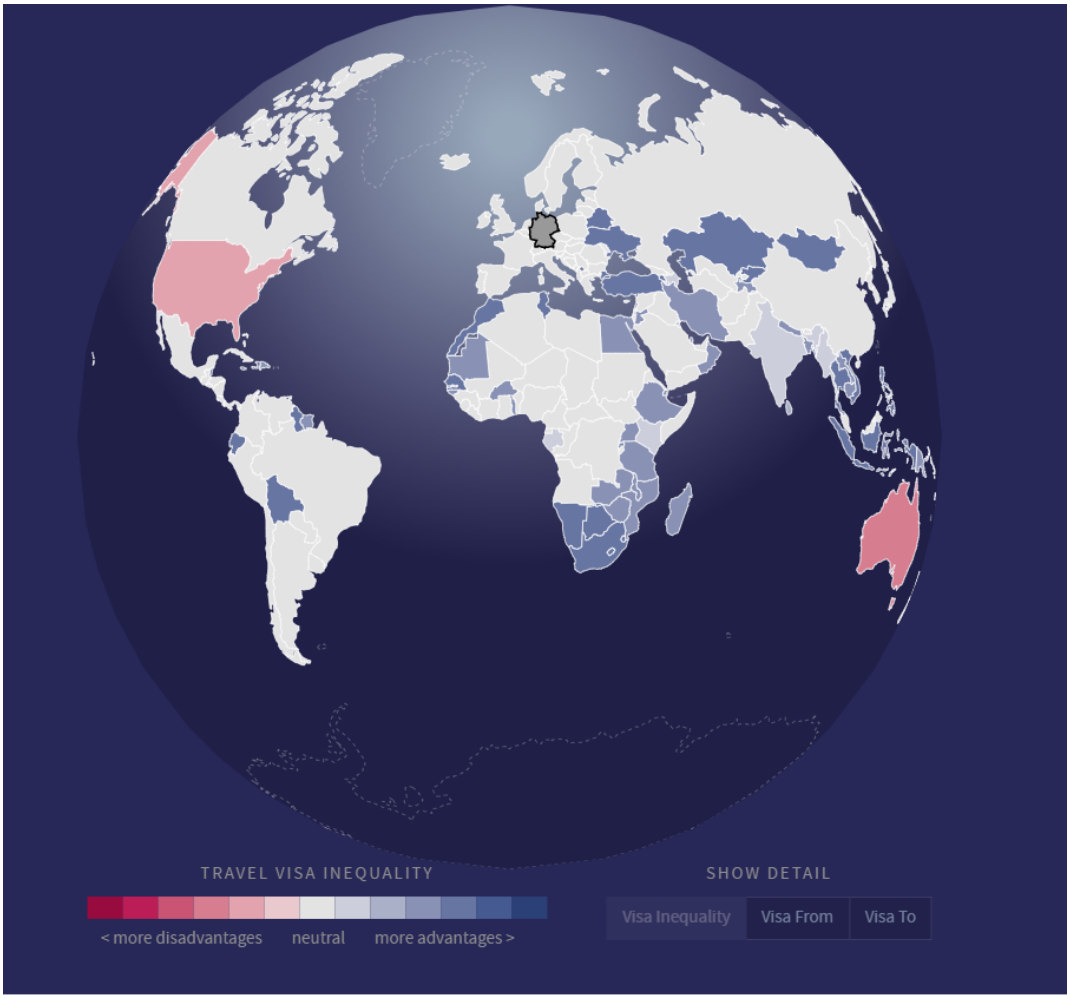

Figure 2: Advantages and Disadvantages for a selected country

In figure 1, the visualization describes what visa policy is applicable for German citizens. The colors in this visualization show the type of Visa required. In figure 2, the figure shows for German citizen which countries are advantages and disadvantage to travel in terms of visa policy. These visualization can be further improved as followings:

- The visualization can be projected into 2D space, because we know that 2D is better perceived by human eyes.
- As these who visualization is only restricted to only German citizens,

we can extend and make it more interactive that any citizen can view their respective visa policies.

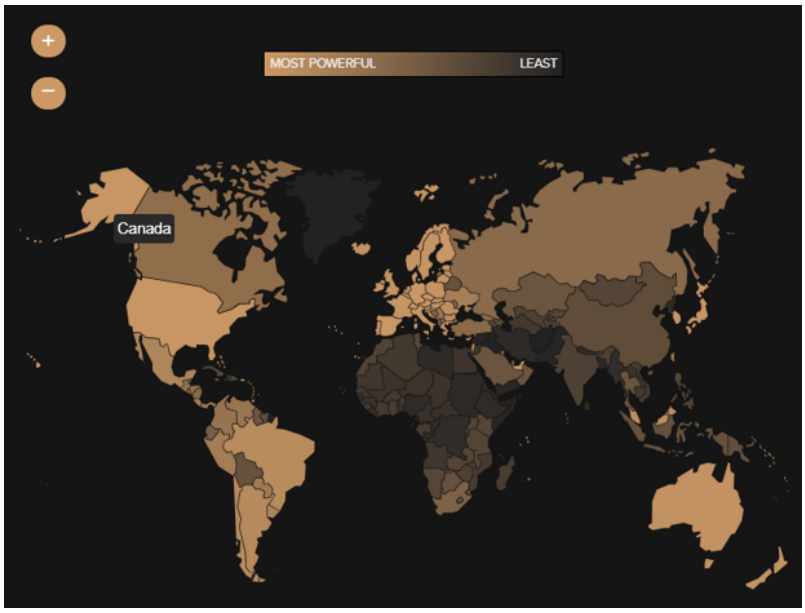

Figure 3: Visa Score Across the countries

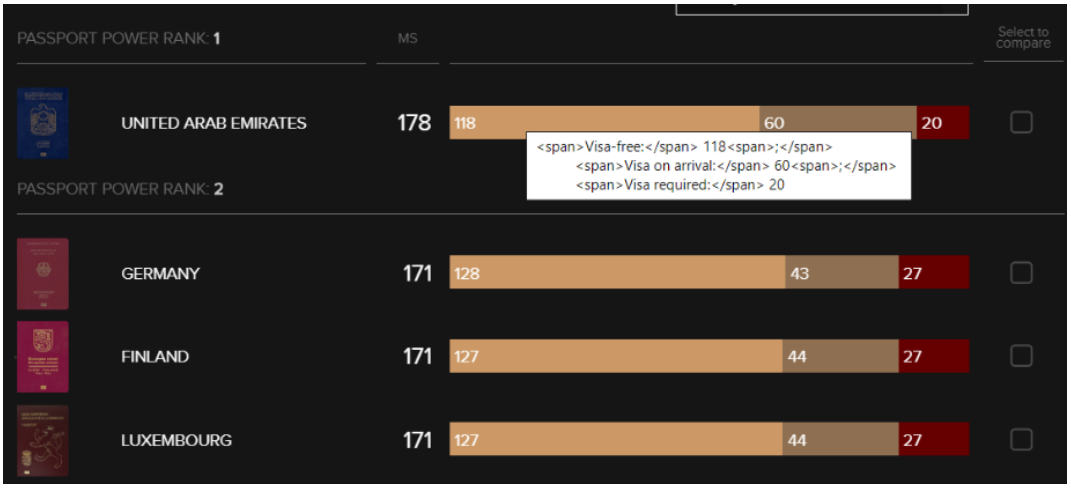

Figure 4: Visa Policies

In figure 3, the visualization describes the score gradient across all the countries the darker the shade less is the visa score and more restrictive are the visa policies In figure 4, the figure shows each countries, for how many countries it can travel without visa , with visa and etc. So the values are stacked in bar.

These visualization can be further improved as followings:

- Choice of colors could have been better in case of Figure 3
- When you hover on particular countries, more information could have been added, like score. Where as in case of 4 the tooltip has html format,so it could have been it plain language
- Figure 4 doesn't give information about which countries that a particular country can travel without visa or other visa policy, it just gives the count

So we would like to take inspiration from these visualization and improve further in getting ride of some of the loopholes and also make it more interactive. The other aspect we would like to look was to see how the visa score and visa policies have changed over years. All these sources will help us to identify the key factors affecting visa policies, and identify if they prove to be restrictive for travellers through a visualization perspective.

2.3 Objectives

1. The main objective of this project to put forth a visualization medium to show the mobility of people across the world. The visualization will help us analyze as to how visa policies restrict this mobility.
2. Diving deeper into this visualization to analyze the type of visas and their individual effects on the mobility.
3. Making an Interactive tool will help us better visualize and understand countries promoting or discouraging free visa travel.
4. We intend to add filters like country, region, location and passport color to increase the granularity of the data available.
5. We plan on getting rank based scores to explore every country's policy independently.

6. We intend to test the hypothesis : Restrictions in visa policies affect the travellers.

7. The visualization will help to study if GDP, resident population, relative economy, incoming and outgoing rates negatively affect the visa policies or vice versa.

8. Encourage discussions on visa policies, citizenship, country branding, global mobility, public security, philanthropy and foreign policy.

9. In today's world, becoming a Global Citizen is more important than ever. Having a second citizenship is a liberating and empowering privilege which comes with a responsibility to the world.

3 Datasets

Dataset1:

Source:https://projects.christianlaesser.com/travel-visa-inequality/
Overview of data:
The data stores the information about the visa policies from one country to another.

Country code1	Country code 2	Visa Type 1	Visa Type 2
AFG	BGD	1	4
AFG	AGO	4	4

Visa Type:
0 = Visa not required
1 = Visa on arrival
2 = ETA required
3 = eVisa Required
4 = Visa Required
5= for all instances where passport and destination are the same

Dataset2:

Source: https://www.passportindex.org/byRank.php
Overview of data:
Data set has features like country, GDP per capita, tourist incoming and

outgoing for year 2018.

Country code	GDP per capita	Incoming(M.)	Outgoing(M.)
CAN	42752.66	16.59 M	32.97 M
CZE	29017.84	9.04 M	5.3 M

We plan to use this dataset to include the effect of the factors like population, GDP etc on the visa policies.

Dataset3:

Source:https://www.henleypassportindex.com/passport

Overview of data:

The data is scrapped from website using python requests library . The data collected is for period 2006-2019. It contains the information about the yearly rank, visa score for each country

Country Code	Rank	Score	Year
DNK	1	130	2006
USA	1	130	2006

Dataset4:

Source:https://www.naturalearthdata.com/downloads/110m-cultural-vectors/

Overview of data:

As our visualization involves the maps, so shape file of each country is considered. A shapefile is a simple, nontopological format for storing the geometric location and attribute information of geographic features. Geographic features in a shapefile can be represented by points, lines, or polygons (areas).

Country	Code	Geometry
Canada	CAN	(POLYGON((-122.84 49.000001,-122.974...
United States	USA	(POLYGON((-122.84 49.000001,-120 49...

Dataset5:

Source:https://www.naturalearthdata.com/downloads/110m-cultural-vectors/
Overview of data:
As our visualization involves the maps, a centroid for each country is considered

Country	Country Code	Longitude	Latitude
Canada	CAN	-98.307	61.36
United States of America	USA	-112.46	45.67

4 Process

4.1 Analysis of Data

Based on the above mentioned data sources, we started with preprocessing of data. We emphasized on making the visa data tidy by melting the existing country to country travel matrix. With the motive to plot the locations on the maps we included the latitude longitude data in the visa data. Or resultant data was transformed into the following format:

Country 1	Country 2	Visa Type	Latitude 1	Longitude 1	Latitude 2	Longitude 2
Afghanistan	Albania	4	34.5167	69.1833	41.3275	19.8189

We wanted to analyze the general trend of the passport scores and the corresponding ranks of all the countries included in our data source 1. We used the Henley passport index data scraped from the website to collect the data for the time frame between 2006-2019. After aggregating all the data sources, we observed that the country names alone are not sufficient to match and merge all data. For example Tanzania was present by name 'Republic of Tanzania' in some of the data sources. Hence we have used the country codes to match the country and its corresponding rows.

4.2 Candidate Visualization Methods

For initial proposal phase we tried using existing visualization platforms available. However we decided to plot the data with the help bokeh packages from scratch.

4.2.1 Tableau

For the initial project proposal phase, we tried using the Tableau visualization too to explore and identify the scope of the data we had. We have plotted the data just to get the gist of the data.

4.2.2 Kepler.gl

We learnt about the kepler.gl during the guest talk arranged by the professor for web based visualization tools. We tried this Uber developed kepler.gl as this was known for its interactive map visualization.

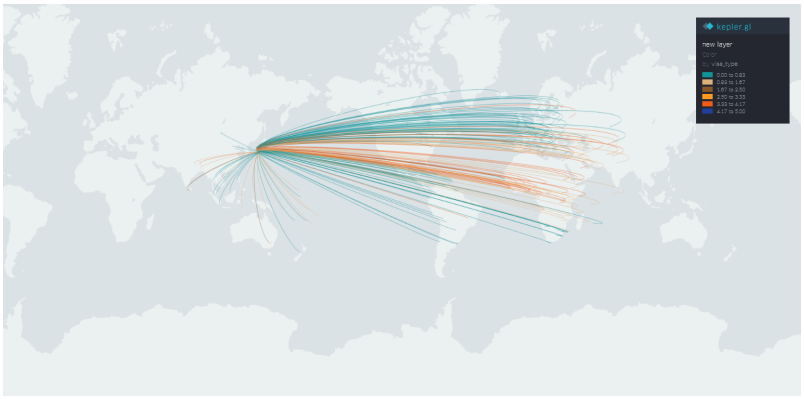

Figure 5: Canada Visa Mobility in Kepler.gl

The Figure 5 shows the initial experiments in kepler.gl.

4.2.3 Python Matplotlib

We also tried simple scatterplots in Python matplotlib or the scraped data points with respect to passport rank and score plotted against the GDP.

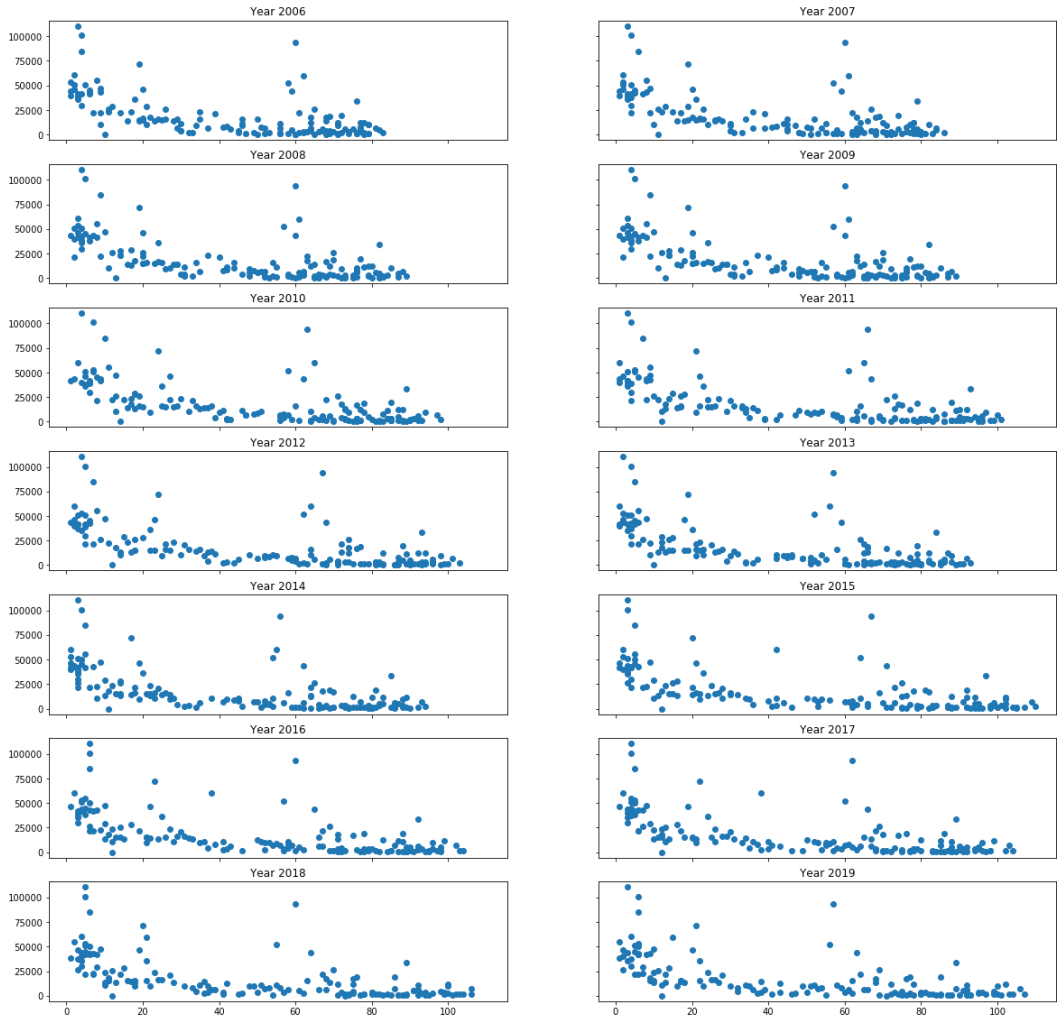

Figure 6: Scatter plots for GDP vs Passport Score

In Figure 6, we can see the circles represent the countries with their position based on the rank on X-axis while GDP on the Y-axis.

4.2.4 Finalized Visualization Techniques

Visualization 1: Country Comparison

1. Bar Chart visualization[3] representing the Top 5 and Bottom 5 Countries with highest or lowest score difference respectively.

2. Bokeh Library is used.Bokeh provides an extensive set of widgets and tools and makes it very simple to create rich, interactive visualizations.
3. Used ColumnDataSource which is like a dataframe in Bokeh to store and update the country data.
4. The HoverTool[3] displays the scores of 2006 and 2019 for every country.
5. All the colors have been used from Viridis10 palette.

Visualization 2: Biggest Climbers vs Biggest Fallers

1. MultiLine plot[3] is used to show the 7 biggest climbers and fallers from 2006 to 2019.
2. Bokeh Library is used.Bokeh provides an extensive set of widgets and tools and makes it very simple to create rich, interactive visualizations.
3. Used ColumnDataSource which is like a dataframe in Bokeh to store and update the country data.
4. Box Zoom tool is used to observe the trend of the score more close. We can drag and select a box in which we want to zoom and closely see the trend.
5. This visualization uses a callback to select the 7 Climbers or Fallers with a Dropdown selection.

Visualization 3: GDP Per Capita and Incoming Rate vs Rank

1. Scatter plot[3] visualization to compare the GDP per Capita or Incoming Rate with Overall rank where each dot represents a country. 2. Used ColumnDataSource[3] which is like a dataframe in Bokeh to store and update the country data.
3. Added the periodic Callback whenever the tab is clicked to stream the data between 2006-2018. When the tab is inactive the periodic callback is removed.

Visualization 4: Visa Score Gradient

1. Map visualization is used for this so that it is would be easy to compare across the countries.
2. Bokeh Library[3] is used.Bokeh provides an extensive set of widgets and tools and makes it very simple to create rich, interactive visualizations. Bokeh offers several ways to work with geographical data including Tile Provider Maps, Google Maps and GeoJSON data
3. Geopandas is used as it can read almost any vector-based spatial data format including ESRI shapefile. This step is a precursor to creating the GeoJSONDataSource.
4. As the visa score is continuous Bokeh's LinearColorMapper to set the palette and range of the colorbar. Bokeh's palette (YlGnBu) is used

5.Added a Bokeh HoverTool that displays data when hovering over a neighborhood.
6. We created the ColorBar using Bokeh's NumeralTickFormatter and ColorBar.

Visualization 5: GDP per capita and Visa Score

1. Bubble Map[3] visualization is used for this so that it is would be easy to compare score and GDP across the countries. And bubble size are used to represent the GDP per capita
2. Bokeh Library is used
3. Geopandas is used as it can read almost any vector-based spatial data format including ESRI shapefile. This step is a precursor to creating the GeoJSONDataSource.
4. As the visa score is continuous Bokeh's LinearColorMapper to set the palette and range of the colorbar.Bokeh's palette (YlGnBu) is used
5.Added a Bokeh HoverTool that displays data when hovering over a neighborhood.
6. We created the ColorBar using Bokeh's NumeralTickFormatter and ColorBar.
7. The bubbles are plotted using Bokeh's circles and the centroid points are passed to place bubble at centroid of every country and the GDP per capita is normalized so that size of circles is controlled.

Visualization 6:Incoming and Outgoing Tourist and Visa Score

1. Bubble Map visualization is used for this so that it is would be easy to compare score and Incoming and Outcoming across the countries. And bubble size are used to represent the flow of tourist
2. Bokeh Library is used.[3]
3. Geopandas is used as it can read almost any vector-based spatial data format including ESRI shapefile. This step is a precursor to creating the GeoJSONDataSource.
4. As the visa score is continuous Bokeh's LinearColorMapper to set the palette and range of the colorbar. Bokeh's palette (YlGnBu) is used
5.Added a Bokeh HoverTool that displays data when hovering over a neighborhood.
6. We created the ColorBar using Bokeh's NumeralTickFormatter and ColorBar.
7. The bubbles are plotted using Bokeh's circles and the centroid points are passed to place bubble at centroid of every country and the both incoming and outgoing flow is normalized so that size of circles is controlled. Choice of

bubbles are orange and grey as it is contrasting with respective to the pallete used for the visa score.

Visualization 7:Visa Policies: How can I travel?

1.Map visualization[3] is used for this so that it is would be easy to perceive that for a selected country what visa policy is required to travel to different countries

2. Bokeh Library is used

3. Geopandas is used as it can read almost any vector-based spatial data format including ESRI shapefile. This step is a precursor to creating the GeoJSONDataSource.

4. As the visa policy is categorical Bokeh's CategoricalColorMapper to set the palette for the categories in visa. Bokeh's palette (RdYlBu5) is used as we have 5 categories

5.Added a Bokeh HoverTool that displays data when hovering over a neighborhood.

6. Bokeh's circles are used to plot the legend for the visa categories

7. Adding a Bokeh Select Widget that enables a user to select the data based on country criteria

Visualization 8:Highlighting the difference

1.Map visualization[3] is used for this so that it is would be easy to perceive that for a selected country what countries are advantage or disadvantage in terms of visa policy

2. Bokeh Library is used.

3. Geopandas is used as it can read almost any vector-based spatial data format including ESRI shapefile. This step is a precursor to creating the GeoJSONDataSource.

4. As the visa policy is categorical Bokeh's CategoricalColorMapper to set the palette for the categories in visa. Bokeh's palette (RdYlBu4) is used as we have 4 categories

5.Added a Bokeh HoverTool that displays data when hovering over a neighborhood.

6. Bokeh's circles are used to plot the legend for the advantage / disadvantage /neutral tag

7. Adding a Bokeh Select Widget that enables a user to select the data based on country criteria

4.3 Failed Experiments

The following section describes the experiments based on the visualization methods we used in the earlier section.

4.3.1 Tableau

While exploring this tool we felt that the tool isn't flexible to customize the visualization based on the requirement

4.3.2 Kepler.gl

Even though we found that Kepler.gl is really helpful report to quickly analyze the map based data, it produces very generic visualizations with limited filters. We wanted to combine the map visualization with simple plots like bar graphs and scatterplots. This becomes difficult while using the Kepler.gl. The 3D visual map in kepler.gl is interesting but was not required for the scope of this project.

4.3.3 Python Matplotlib

Based on the scatterplots shown above we thought a streaming data based visualization should be more helpful in showing the trends throghout the year window of 2006 to 2018. Subplots were interesting but lack the interactivity we wanted to introduce in the visualization.

5 Results

We decided to use bokeh visualization package available in python. We used this method as it gave us a mix of interactive, simple and streaming data based visualizations. We were able to come up with 8 different kinds of visualization which effectively and interactively explain the data.

5.1 Tabbed Views on Bokeh

Based on the visualizations that we have come up with, tabbed windows for each visualizations are used. Every tb has been given a heading. The visualization loads when the tab is clicked.

5.2 Country Comparison

This visualization compares the 5 biggest climbers and biggest fallers between the period 2008-2018.

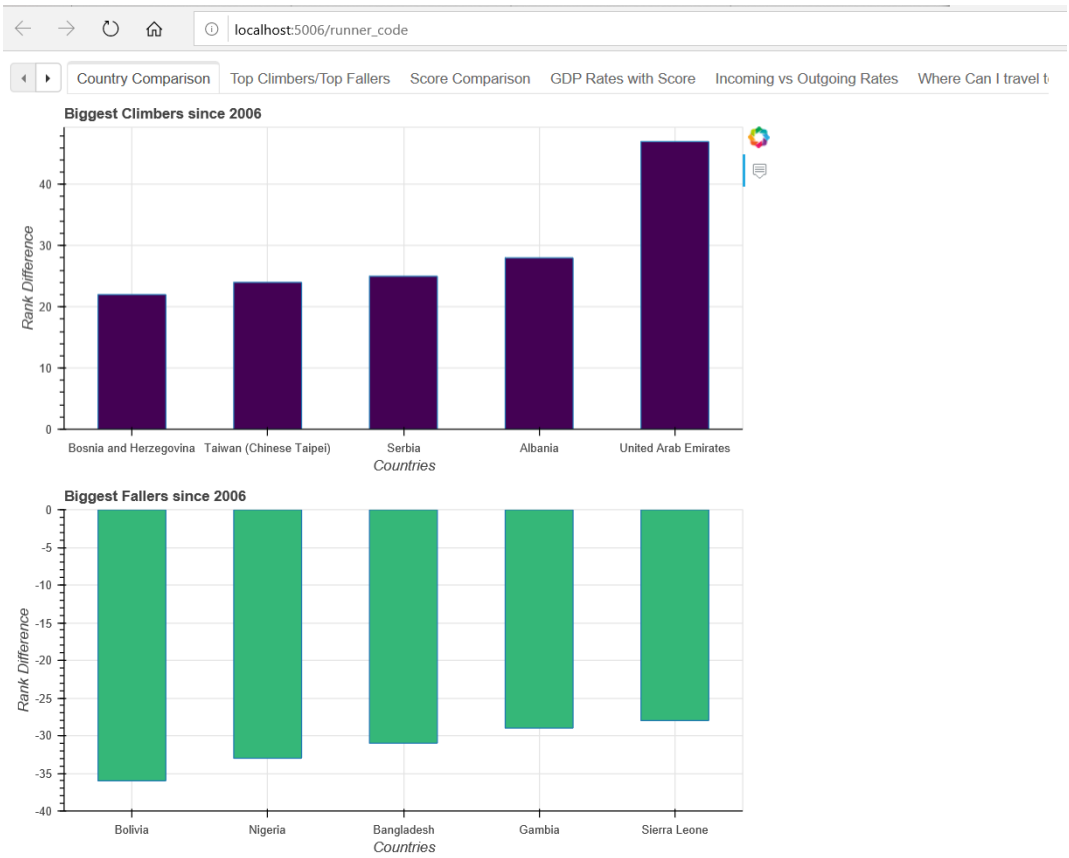

Figure 7: Country Comparison Bar Plot

The figure 7 is the visualization for biggest climbers and fallers. The countries Bosnia and Herzegovina, Taiwan, Serbia, Albania, United Arab Emirates are the biggest climbers with highest rank difference. While countries like Bolivia, Nigeria, Bangladesh, Gambia and Sierra Leone turned out to be biggest losers.

5.3 Top Climbers vs Top Fallers

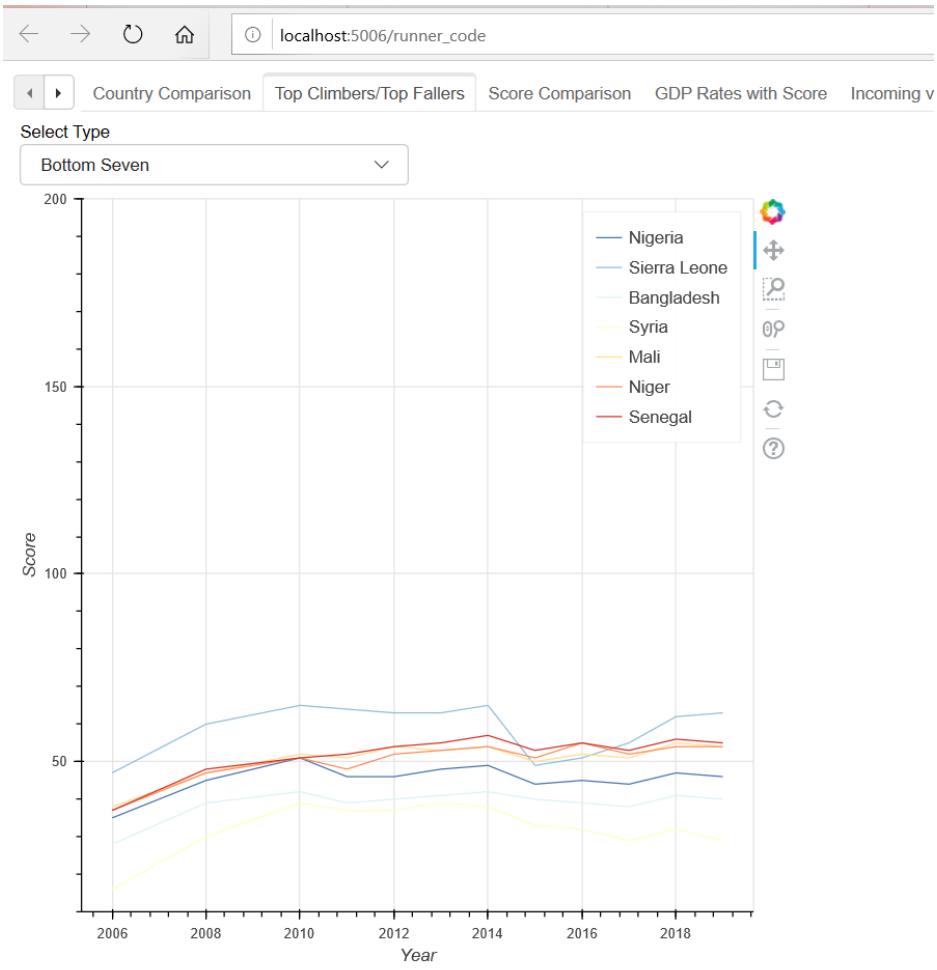

Figure 8: Bottom Seven Countries

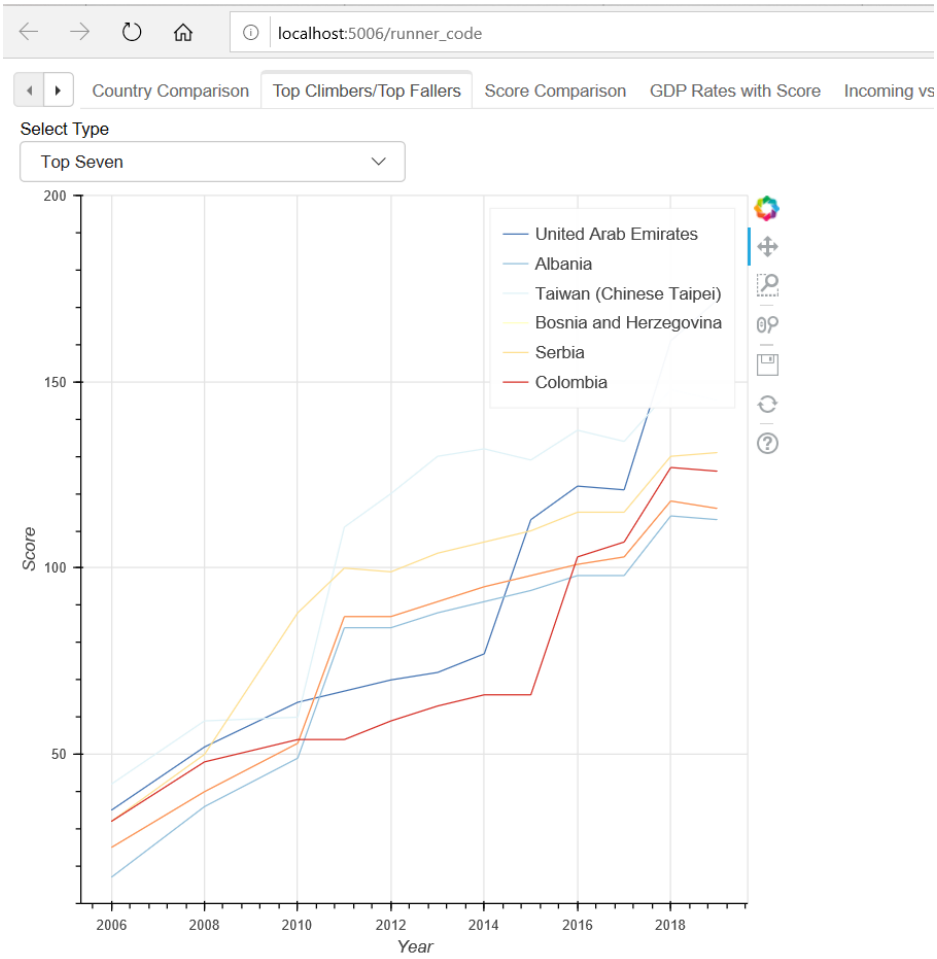

Figure 9: Top Seven Countries

We tried to see the progress of score of the top 7 and bottom 7 countries from 2006-2018. The bottom 7 countries have a stagnant scores while the top 7 shows a very steep gradient. Our inference from this visualization was that, the score improvement is concentrated only towards certain countries and regions. The score of a bottom tier country is relative to the improvement of score for top tier countries. This visualization is interactive with a drop down selection to view the Top 7 and Bottom 7 countires. This visualization uses the Viridis color map to plot the lines.

5.3.1 Insights

In Figures 9 and 8 we can see that the top tier countries continued to grow till 2018 with noticeable differences in their scores while the bottom tier countries had a stagnant score throughout 2006 till 2018. After learning this we tried to the features that may have affected the ever increasing or stagnant visa score in both these cases.

5.4 Visa Score Gradient

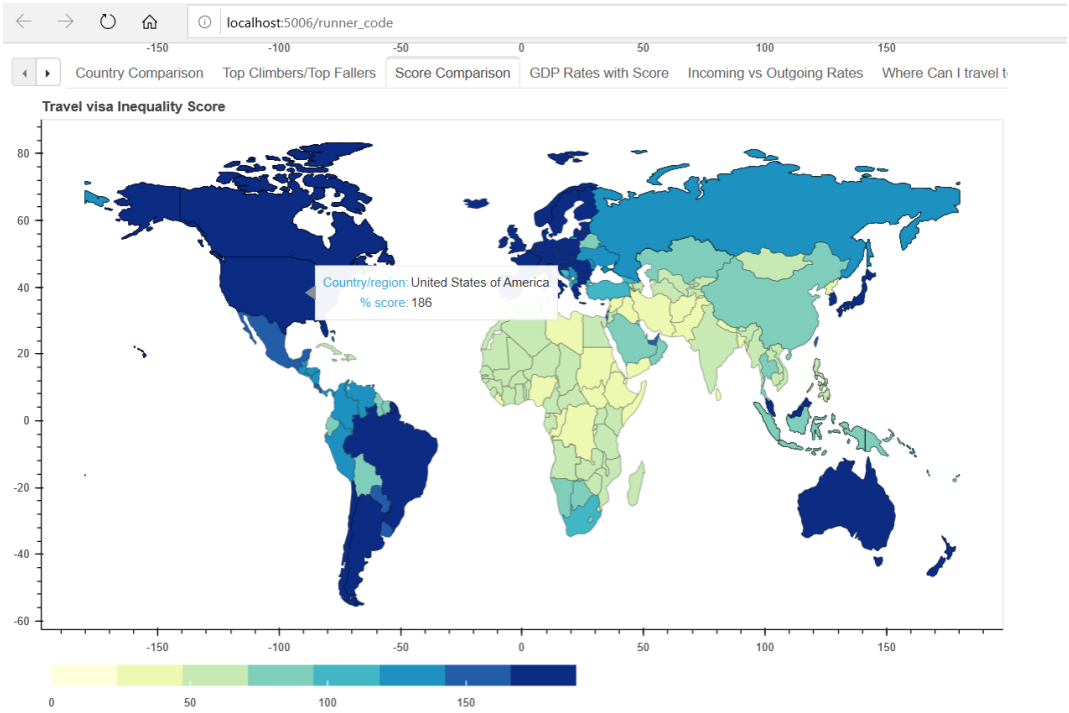

Figure 10: Visa Score Gradient

This visualization shows a overview of the countries and their corresponding visa scores ranging from 0 to 200. The shape file from the data sources helped us to plot the country boundaries and color code the regions with the help of palettes. We also used a ColorBar from bokeh to represent the color

scale.

5.4.1 Insights

In Figure 10 we visualized the scores of different countries. We can clearly see that there was a clear divide in the visa score for a continent based data. The African continent shows below average visa scores throughout. The European continent has very strong visa scores, Both North and South America have high visa scores. Except Japan, Asia has average scores. This can be supported from the fact that Japan has one of the most powerful passports in the world.

5.5 GDP Per Capita Comparison

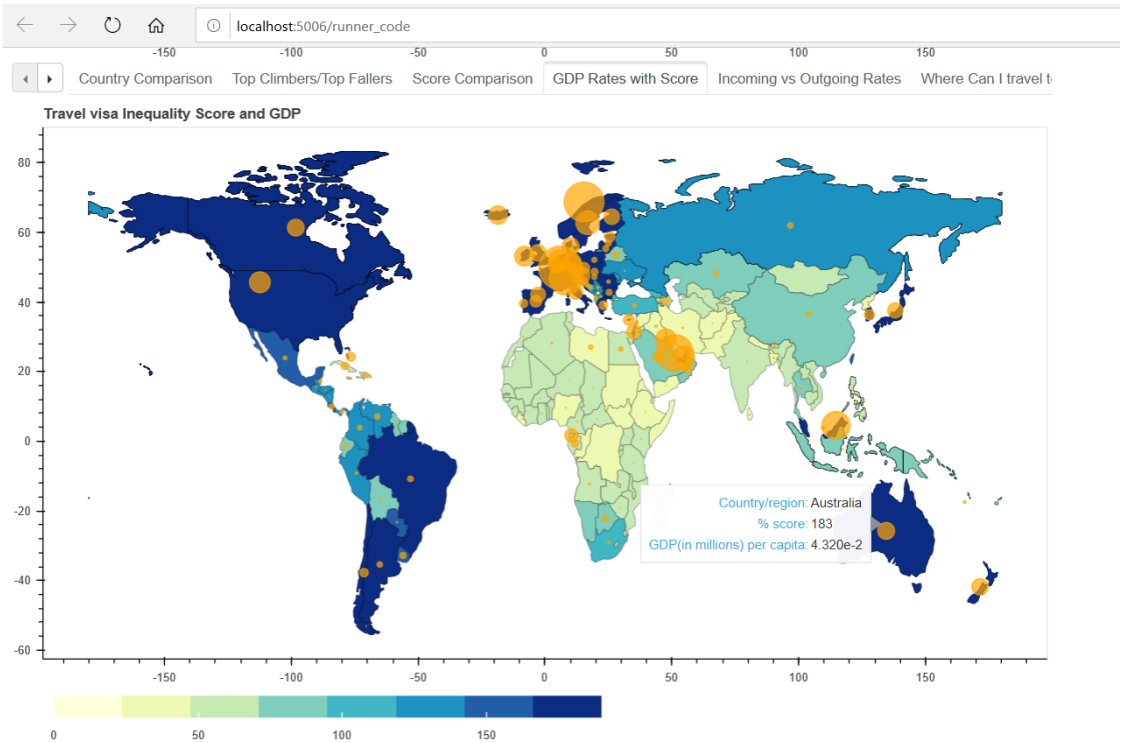

Figure 11: GDP Per Capita Comparison

We developed this visualization to visualize the relation between the GDP per capita and the Visa score of all the countries. We went through [2] article. We ensured that as the variable is GDP per capita the values will be in proportion to the country under consideration. We used the orange colored bubbles to distinctively show the GDP per capita of every country. We referred [2], to explore the ways to show additional features over the map visualizations.

5.5.1 Insights

After looking at score comparison we tried to find if the GDP per capita affects overall visa score of the country. In Figure 11 we can see that the Europe continent reflects a positive correlation between gdp per capita and a visa score. The gdp bubbles are not that relvant for countries like India and China as they have largest populations in the world. The GDP bubble can be clearly seen in case of Iceland.

5.6 Incoming vs Outgoing from Countries

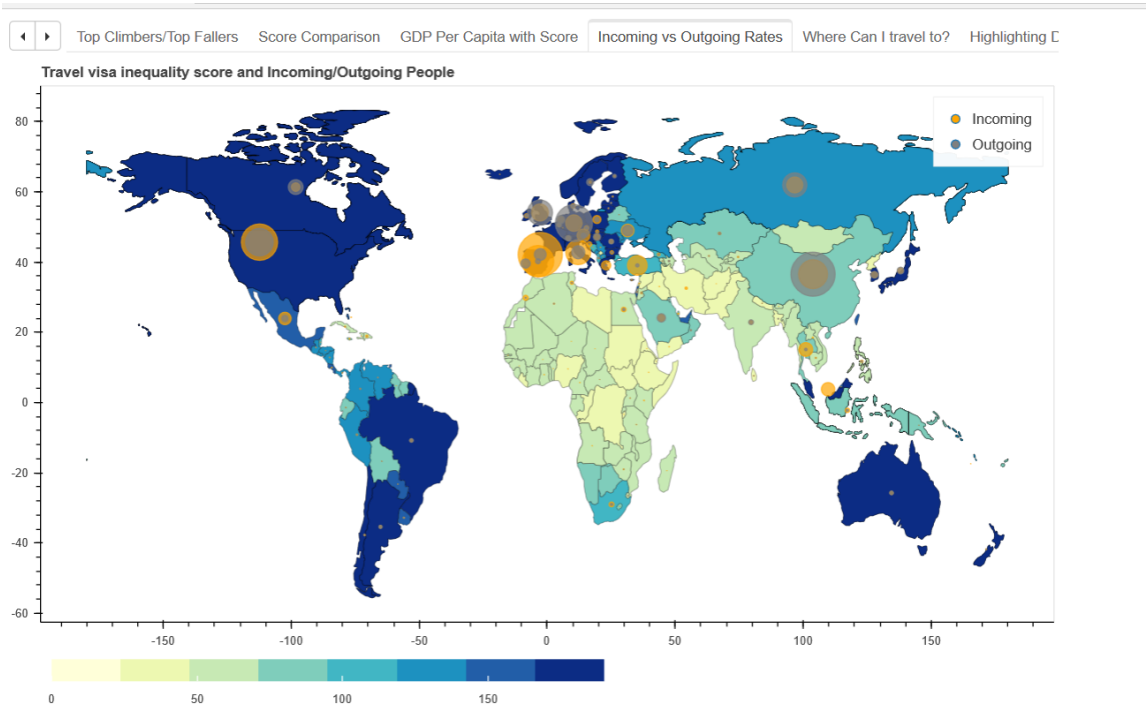

Figure 12: Incoming vs Outgoing Comparison

We used the same score comparison visualization, but added the incoming and outgoing features to it. We used two different color encoding (grey and orange) for representing incoming and outgoing counts respectively. The radius of the bubble helps to understand how the rates affect the overall score of the country.

5.6.1 Insights

On comparing the Incoming and Outgoing rates with the scores in Figure 12 we found that countries with high scores motivate people to travel to other countries. They also attract lot of incoming people into country as tourists, students and so on. For example United States and European has a high Incoming as well as Outgoing rates supported by a high Visa scores. Japan

has high outgoing rates which support the fact that citizens travel a lot given a stronger acceptance of passports in other countries.

5.7 Visa Polices: How can I travel?

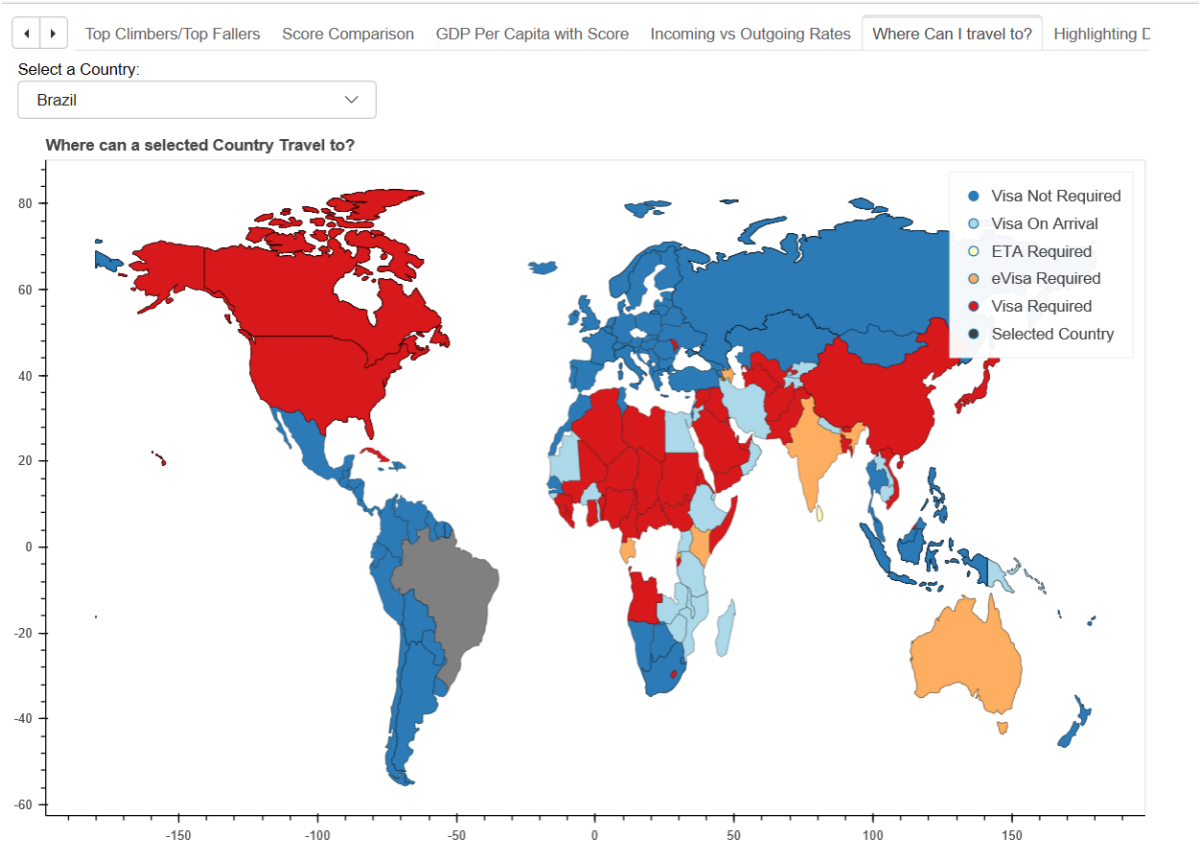

Figure 13: Visa Polices: How can I travel?

Based on the visa type country names from Data Source 1 and area data from the shape files we developed an interactive visualization shown in Figure 13. The user can select the country of citizenship and see which countries he can travel to by different visa types. We have color coded the countries based on the visa type.

- Red represents Visa Required

- Light blue represents Visa on Arrival
- Dark blue represents Visa Not Required
- Orange shows eVisa Required
- Yellow shows ETA required
- Black is the country Selected

So in the above visualization the country Canada is selected, so if you want to travel to USA visa isn't required where if you want to travel to Russia the Visa is required.

5.7.1 Insights

On visualizing the countries from the Biggest Climbers and Biggest Fallers from Figures 9 and 8 we can clearly see the difference between the visa restrictions and scope of visa mobility. Having the data for each year would have helped us to see the change in restrictions with respect to political and economic situations during those time frames.

5.8 Highlighting Differences

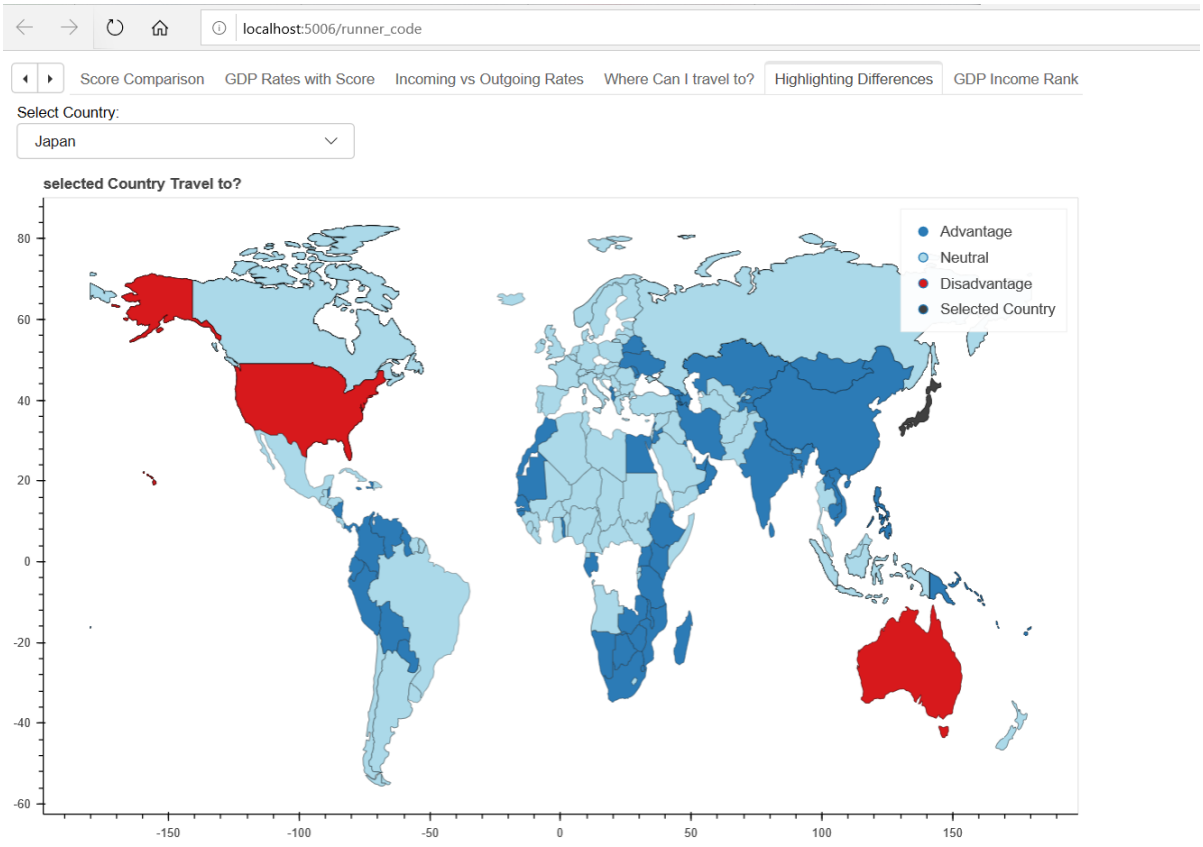

Figure 14: Highlighting Difference - Japan

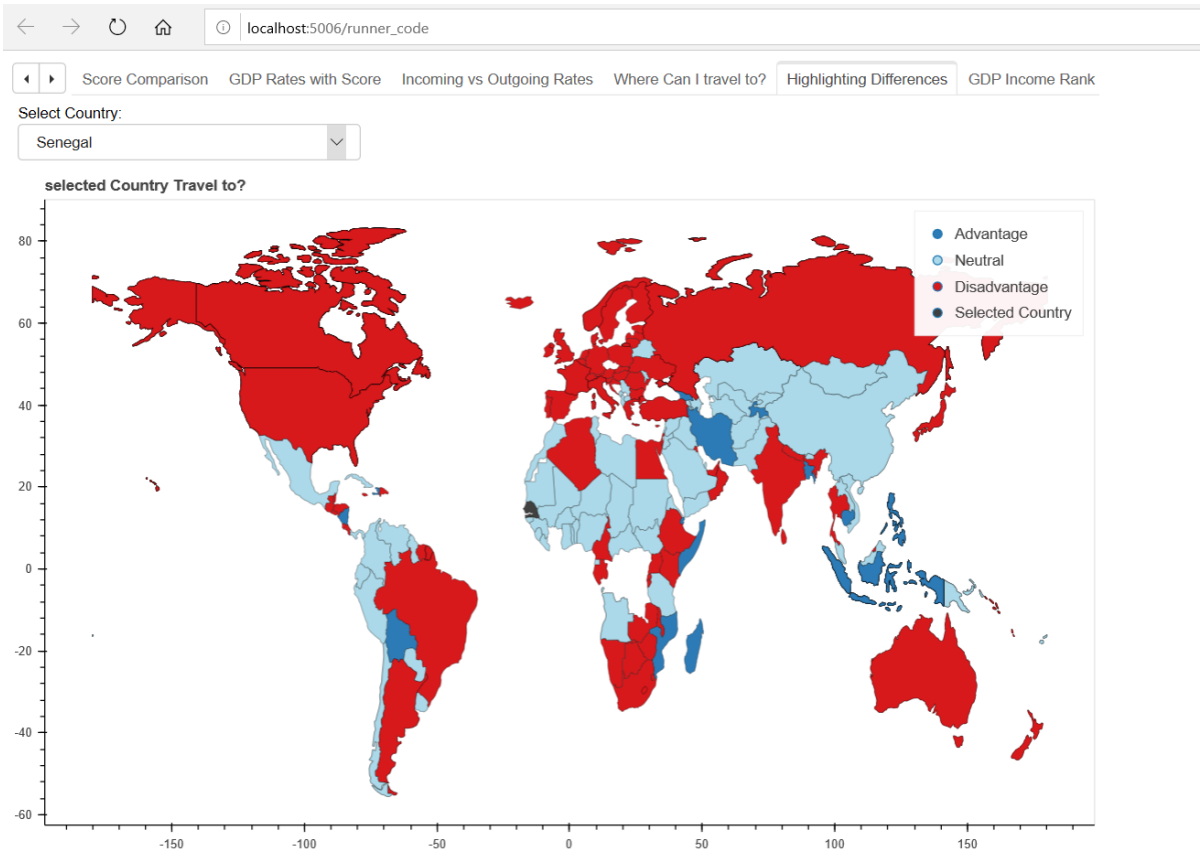

Figure 15: Highlighting Difference - Senegal

This visualization shows a advantage or disadvantage for a citizen when he wants to travel to different parts of the world.
If a person from country A wants to travel to country B we check what type of visa is required for country B to travel to country A. If the difference is positive the person is at a disadvantage as the difference here means more restriction. A negative difference represents a advantage of travelling to that country. We color coded the visualization based on these metrics.

- Blue represents advantage
- Light blue represents Neutral (Zero Visa Difference)
- Red shows a Disadvantage

- Black represents the selected country

5.8.1 Insights

In figures 15 and 14 we could clearly see difference, where being a Japanese national allows you to travel more with lesser visa restrictions. Senegalese national has a lot of disadvantages while travelling to other countries given their passport scores.
Overall, we see three major types of policies for countries namely:

- Restrictive
- Welcoming
- Neutral

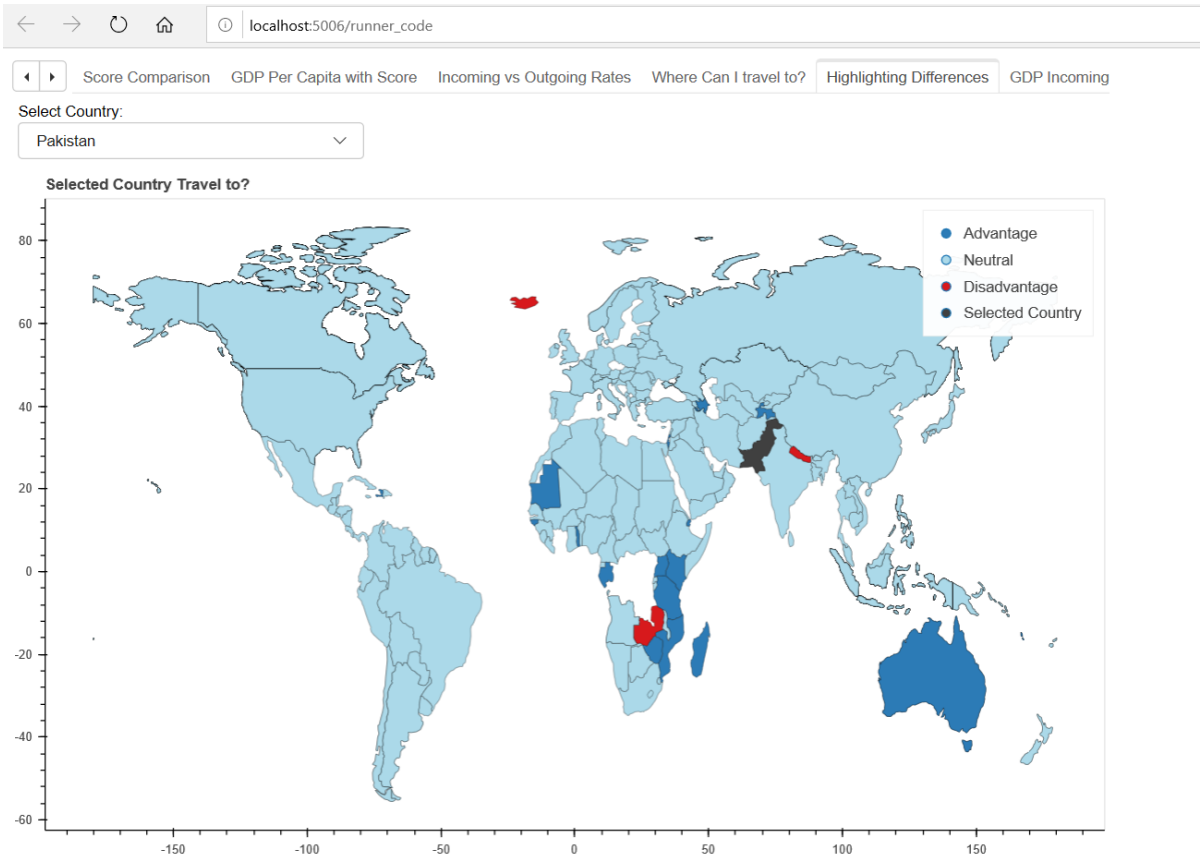

Figure 16: Highlighting Difference - Pakistan

For example, we can see from Figure 16, Pakistan might be practicing a neutral policy, where it imposes the same restrictions as the target country. On the other hand a country with lot of disadvantages, may imply that people from other countries find it easier to travel to that country.

5.9 GDP Per Capita and Incoming vs Visa Rank over Years

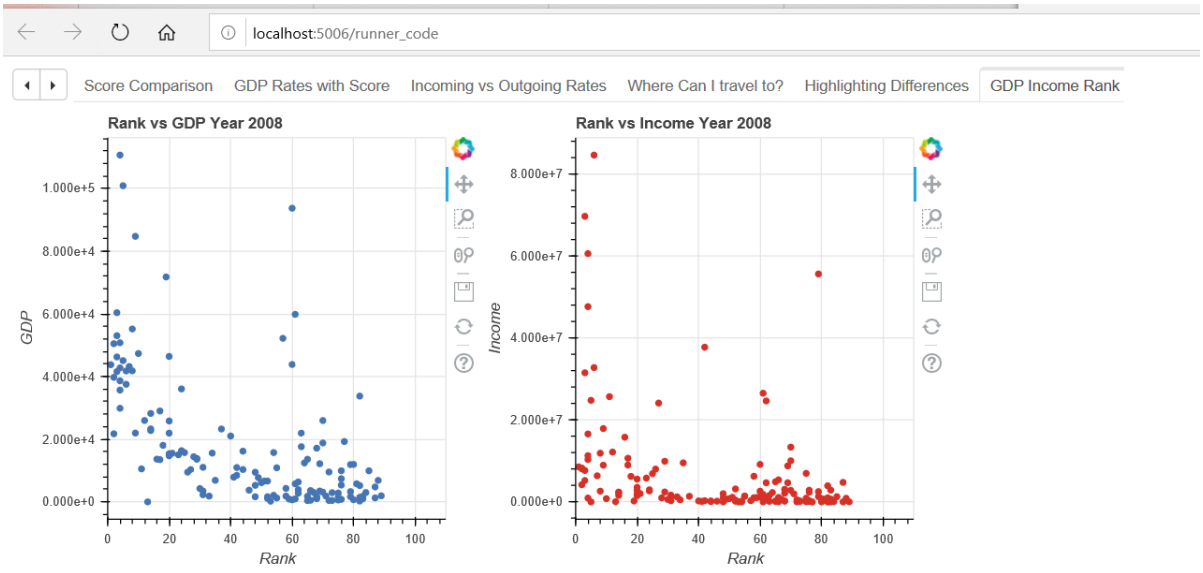

Figure 17: GDP and Incoming Rate vs Visa Rank over years

We scraped the rank data from Henley Passport index for all countries between 2006 and 2018. We used a scatter plot to show two feature pairs - Rank vs GDP Per Capita and Rank vs Incoming Rate. This visualization is based on streaming data. The visualization refreshes every 2 seconds. Every time it pulls in the year based data and plots it. Here each dot represents a country. The blue dots represent the GDP per capita while red dots represent the Incoming rates.

5.9.1 Insights

We could see from Figure 17 that for the first 20 ranks between 2006 and 2018, the countries have high GDP per capitas. Intially there isn't relation but over the period we could observe the strong correlation between the rank and GDP per capita. This correlation also exists in case of incoming rates and ranks, however it is not as strong as GDP per capita.So from this graph it is evident that inequality in visa is increasing rather than decreasing.

6 Conclusion and Future Work

- From all the visualizations we tried, we found that factors like GDP per capita, Incoming and Outgoing rates do affect the overall passport scores. Exploring some more features which may affect the visa score. It will be interesting to factors like education, literacy, health index and crime rates effect on the score.
- It is also observed that visa inequality is increase over the years instead of decreasing.
- There is a clear distinguishing factor which divides the world based on continents by using the visa score only,
- It will be interesting to learn about a country's visa policy with respect to duration of the Visa. For example United States issues a visitor visa which has a 10 year duration. Schengen group countries issue a visa of just 6 months.
- All these factors support a hypothesis a citizen is encouraged to travel due to his country's overall performance and the same factor also restricts his mobility.

7 Web Page URL

The Visualization dashboard can be viewed at the URL* given below

`http://pages.iu.edu/~rbapat/bokeh_visual/runner_code.html`

*The visualization on dashboard are not interactive visualizations as the javascripts are not embedded on the HTML page

References

[1] Roland J Fuchs and George J Demko. Geographic inequality under socialism. *Annals of the Association of American Geographers*, 69(2):304–318, 1979.

[2] Yan Holtz. *Bubble Map.*

[3] Jim King. *How to create an interactive geographic map using python and bokeh?*

[4] Lena Laube. Who may see the Acropolis? Global patterns of inequality and the right to tourism. *Völkerrechtsblog*, November 2017.

[5] Steffen Mau. Mobility Citizenship, Inequality, and the Liberal State: The Case of Visa Policies. *International Political Sociology*, 4(4):339–361, December 2010.

[6] Steffen Mau. *Social transnationalism: Lifeworlds beyond the nation-state*. Routledge, 2010.

[7] Lene Laube Steffan Mau, Fabian Gulzau and Natascha Zaun. The global mobility divide: How visa policies have evolved over time. *Journal of Ethnic and Migration Studies*, 41(8):1192–1213, 2017.