CSE 564 : Visualization

Migration Visualization

Nidhi Mendiratta
110928535
nmendiratta@cs.stonybrook.edu

Shilpa Gupta
110948405
shilgupta@cs.stonybrook.edu

BACKGROUND :

Immigration is the international movement of individuals who are not natives or do not possess citizenship in order to settle, reside, study or to take-up employment. It has been a major source of population growth and cultural change throughout much of the history of the United States. The economic, social, and political aspects of immigration have caused controversy regarding ethnicity, economic benefits, jobs for non-immigrants, settlement patterns, impact on upward social mobility, crime, and voting behavior. According to U.S. Census Bureau, net international migration to the United States will become the primary driver of the nation's population growth between 2027 and 2038. Therefore our aim is to identify the patterns observed in immigration across different countries of the world.

PROBLEM STATEMENT :

Our aim is to project the world immigration data in a form in which we can visualize the different patterns of immigrants from different countries across the timeline. Immigration data can provide great insights into how the demographics of countries changes over time for example number of immigrants of a country greatly determines the economic status of the country, therefore deciphering these patterns just from the bilateral data is the challenge we are picking up.

DATASET :

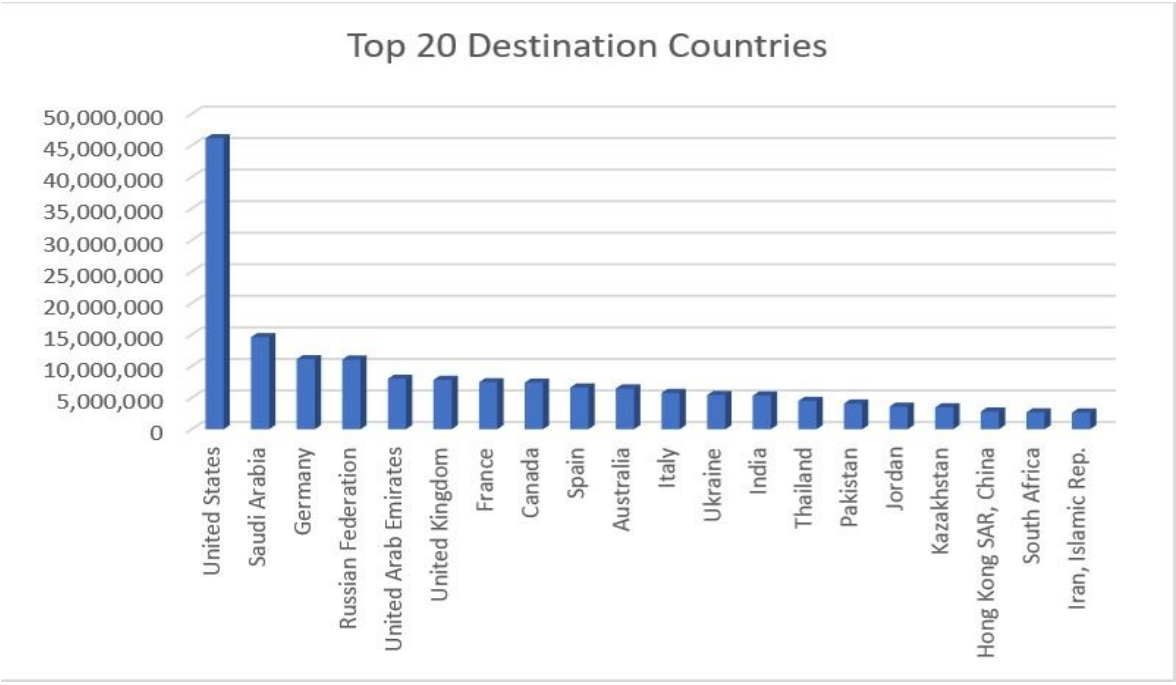
Our dataset comprises of a bilateral migration matrix, release by World Bank where rows represent the source country from where people are migrating whereas the

columns represent the destination matrix to which country the people are migrating. In all we have 216 countries, therefore a 216*216 matrix where any cell value at [i^{th} row, j^{th} column] signifies people travelling from i^{th} country to j^{th} country. Apart from this dataset, we are also considering other government release official data of changing demographics of the country.

METHODOLOGY :

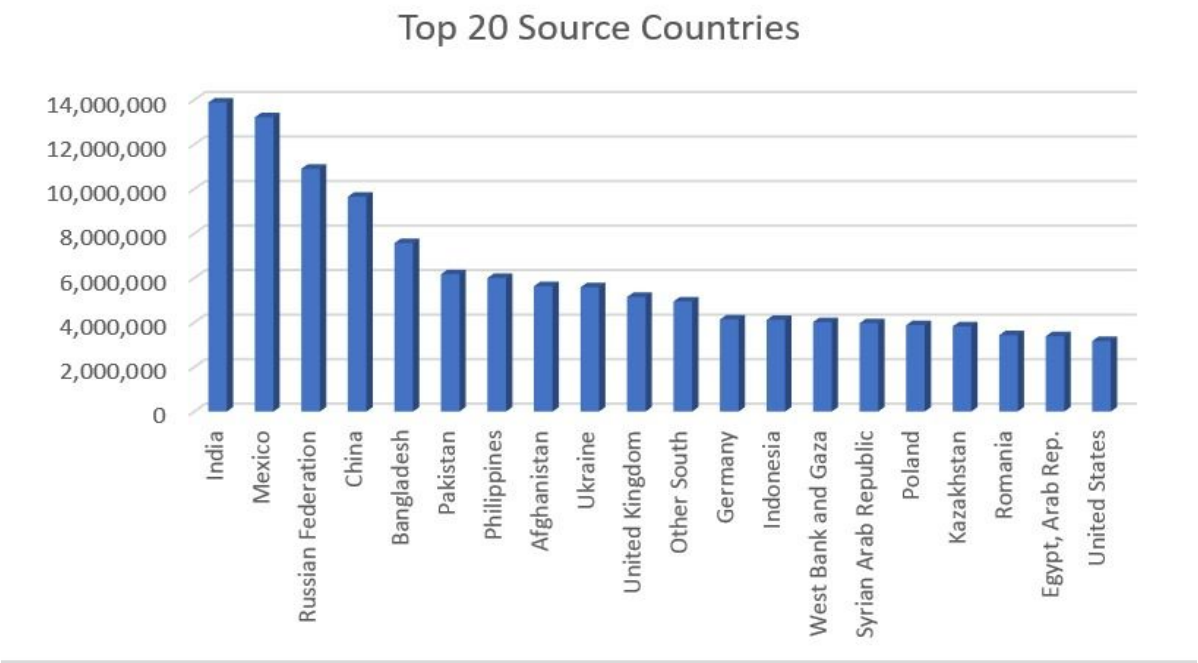
Making sense out of tabular data , can be a tedious job and one can miss the obvious and the not-so obvious patterns, therefore we aim to plot double vertical bipartite or any other d3 plot where we can clearly show the the relationship of people migrating between the countries. Also we aim to plot this complex data into a world map, to clearly depict and easily understand the positions of countries and also show graphically the movement of people from source to destination with simple d3 animations.

Along the main graphs , we aim to show trend observed in the data like top migration destinations , top emigration countries, top migration corridors etc. The dataset that we are using for this purpose has the number of people migrating from one country. To just get the gist of the data we plotted the top source and destination countries according to our dataset, below are the plots.



The above chart shows a sample of the dataset, considering the top 20 countries from a dataset of 216 countries , we show a simple visualization of the destination countries of immigrants and as can be seen US is the leading choice of people and that too with a huge margin. This is just a glimpse of what we wish to achieve in this project.

Similar to the first, below chart depicts the bar plot of the top 20 countries from where



most people migrate. As we can see in the plot india and Mexico are the countries from where most number of people are migrating. This can reveal the information hidden in the tabular data in more visible form.

Considering the dataset, there can be challenges of representing all the 216 countries in one single plot and might not be easy to interpret. Therefore, we aim to ensure that the task of visualization is done in such a manner that it is understandable and knowledgeable at the same time.

PROGRESS REPORT :

The project proposal mentioned above depicts our plan to proceed with problem at hand i.e to depict the migration patterns in a clean and understandable scenario. Here in this section, we wish to highlight what all we have achieved so far, the steps as to how we approached the problem and also the future work of what we wish to do next with the data.

DATASET ANALYSIS :

From the world bank website we got the migration data about the countries, which included datasets for different timelines and different features, so as a first step we reduced the dimensions of the data since we had was a sparse bilateral matrix, as shown below.

Destination country (across) - Source country (down)	Afghanistan	Albania	Algeria	American Samoa	Andorra
Andorra	0	0	0	0	0
Angola	0	0	0	0	0
Antigua and Barbuda	0	0	0	0	0
Argentina	0	0	0	0	708
Armenia	0	0	0	0	0
Aruba	0	0	0	0	0
Australia	0	0	167	0	61
Austria	0	0	167	0	0
Azerbaijan	0	0	0	0	0
Bahamas, The	0	0	0	0	0
Bahrain	0	0	215	0	0
Bangladesh	0	0	277	0	0

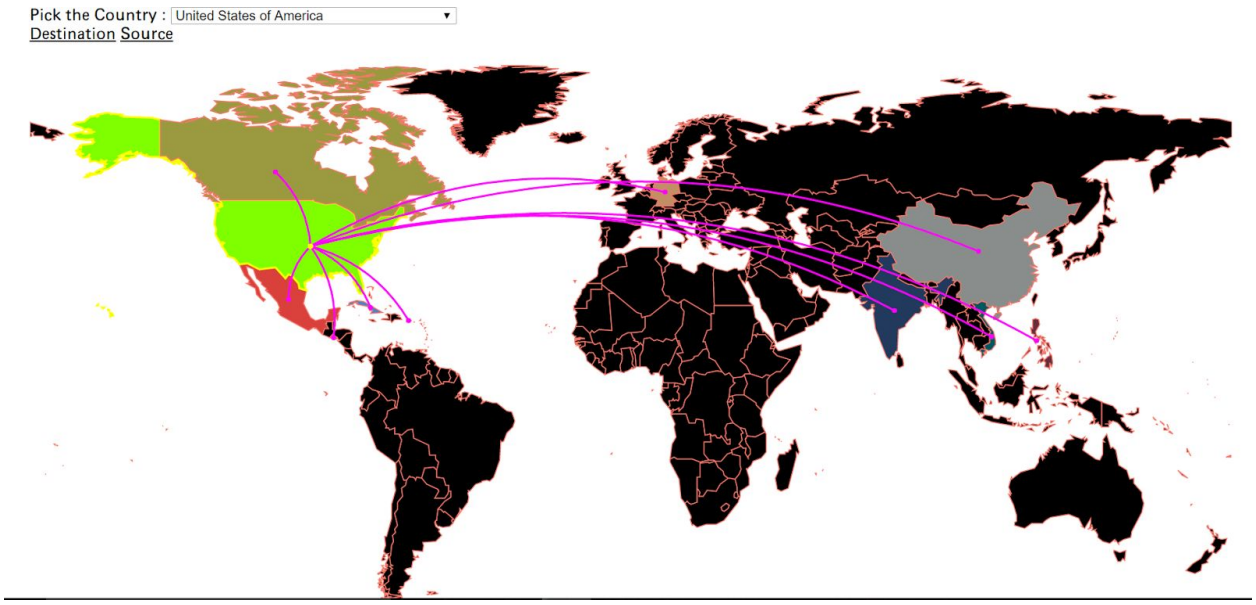
We also did clustering of data , since there are too many countries and all of them vary based on their sizes , population etc, therefore we thought of making a chord diagram i.e. clustering countries by their regions so that a top level inter regional migrational pattern can be observed and we can expand any region we want.

APPROACHES:

Visualising the immigration pattern in the world map :

To visualize this given immigration pattern from one country to another we thought it is better to view these patterns on a world map. So we tried to visualize the top 10 countries from where the population comes and top 10 countries to which people of the country migrate to. Along with this

data we are also showing the information about the countries based on the data provided by the world bank which includes country population, GDP, HIV prevalence, TB prevalence and Mortality under 5. As we can see in the screenshot below



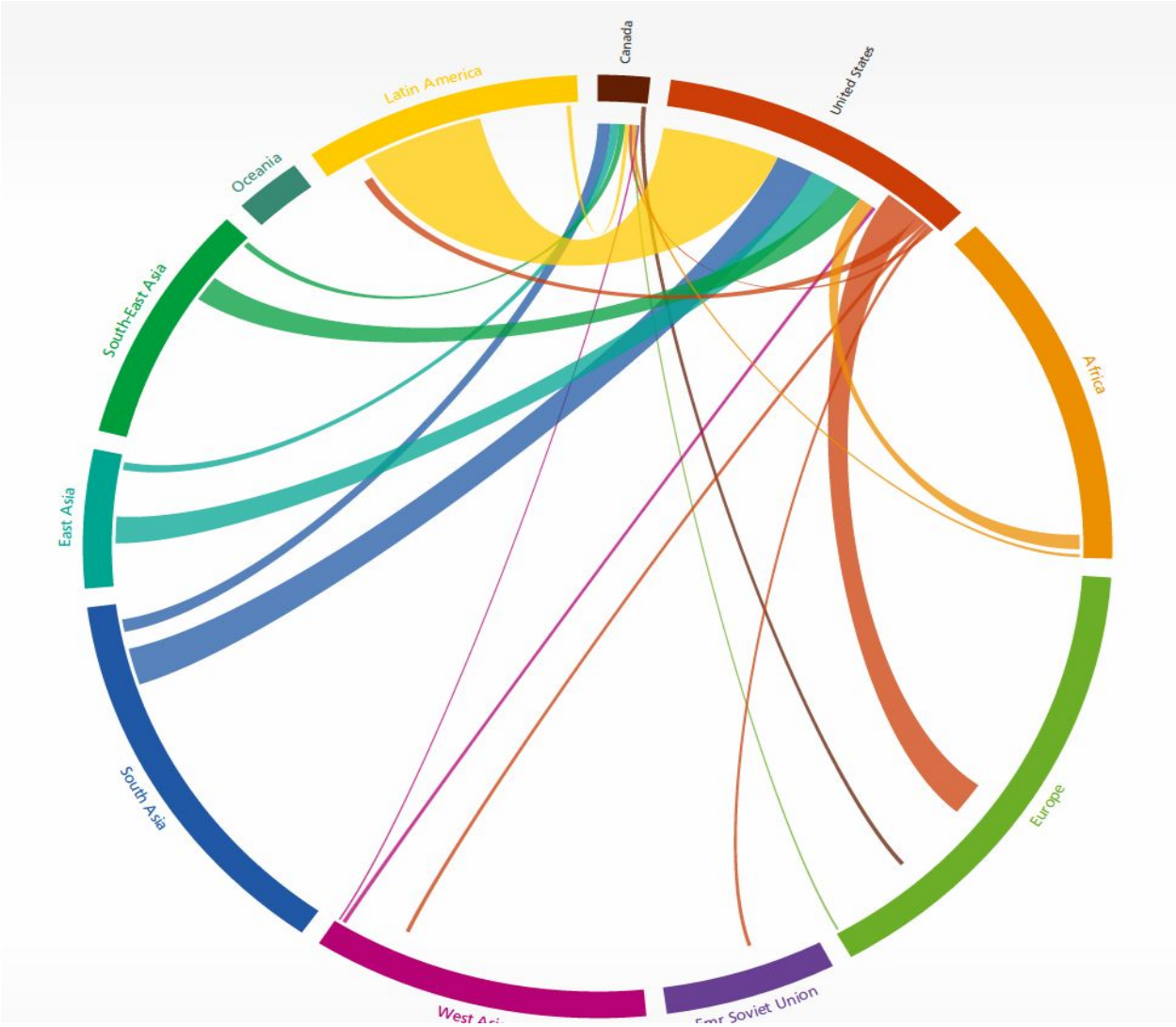
There is a menu in the top left which has the drop down containing the country names, a user either can select the country from the dropdown or can click on any of the country in the map, and later click on either source or destination to see the top 10 country from where people arrive to this country or top 10 countries from where people migrate to this country.
On the hover over the country in the map other information are also visible about that particular country.
In this visualization we are not showing the information about all the countries all the time which can make the visualization cluttered instead we are showing the top 10 destination/source for the country user selected. Which falls into the idea of engaging the user into the interaction and making the visualization more interactive.

Visualising immigration patterns by grouping countries in a chord diagram :

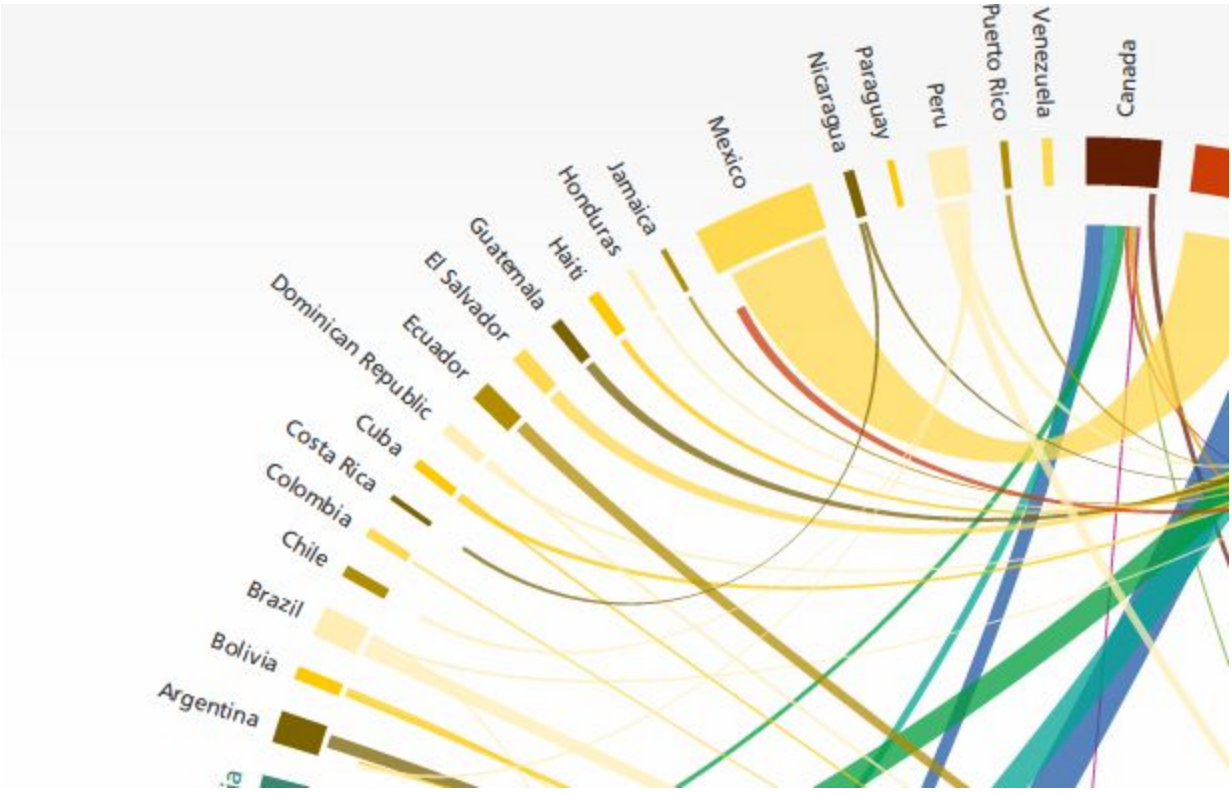
The above map shows location of every country and its migration patterns i.e people migrating to/from the selected country. To get high level idea of what top 10 source and destination countries are for any user selected country. Above visualization does not give the details about how many people are migrating from one country to another as well as it does not show every country to which people migrate, it shows only top 10. To overcome this and to present more detailed information to the user we have come up with the chord diagram.

Here we aimed to describe more on the patterns of migrations for close by regions, so we grouped countries , so as to enable migration patterns as a whole of the region as well as of individual countries , when the region is expanded by a simple click.

For this approach, we grouped countries according to regions and plotted on the chord diagram, to ensure that all countries are plotted in a structured manner. Whenever mouse is hovered on the arc , the corresponding country's intake and outtake values are shown on tool tip. Whenever the arc is clicked the region is expanded into the countries it composes of and clicking on any of the country would show the migration patterns of that country. Also the size of the chord connecting two regions or countries signifies the number of people migrating. Screenshots of all of the following are as follows :



This is the basic look of the chord diagram where an user can see the overall picture of migration pattern based on the different regions. On hovering over the different chords user can see the count in the tooltip as well.
Now if the user is interested in a particular region, he/she can click on it and expand to see the immigration pattern for the different countries in that region. As can be seen below



Now the user has been provided with more details and patterns at more detailed level where he/she can see the number of people migrating among different countries. The concept of brushing is being used here to provide the details only on expansion. So that user can focus on a particular region at a time. Hovering on an arc shows the total in and out of the country and hovering on any chord shows the number of people travelling between the two countries connected via the chord as shown below.

FUTURE WORK :

- We are aiming to combine the approaches that we have been implemented into an interactive visualization experience for the user where the user can select the type of visualization he/she wants to use as well as level of detailing they want.
- Along with it we are also thinkinking about if we can include more information about different countries in our visualization and make it more engaging for the user by providing more interactive options to them.