CS 6630 - Project Proposal

- **Basic Info:**

  Project Title: Immigration Pattern for United States of America

  Repository: https://github.com/lordawak3n/dataviscourse-pr-WorldMigration.git.

  Team Members:

Name	Email	UID
Ankur Rathore	u0941534@utah.edu	u0941534
Garima Chhabra	garima.chhabra@utah.edu	u1143636
Rishabh Kaushik	rhul840@gmail.com	u1008988

- **Background and Motivation:**

  It is a well-known fact that the United States of America (US) is a country whose foundation is based on immigration of people from all across the world. With the help of visualization techniques we want to understand the immigration trend in US from all other countries over the course of years. The motivation comes from all of us being immigrants ourselves, who have moved to the US in pursuit of better career opportunities. Additionally, we found it fascinating how visualization helps us understand large quantities of immigration data in the most compact way possible.

- **Project Objectives:**

  The objective of this project is to get a detailed overview of the immigration pattern in the US. Through this project we plan on accomplishing a way to have a better understanding of how people immigrate to US over the years. This will help in regression analysis for anyone whose business model depends on immigration.

  The dataset we have obtained is from a statistics body called the Organisation for Economic and Co-operation and development (OECD). The dataset describes the inflows of foreign population based on nationality and based on continents from the year 2000 to 2016. Through this data we plan on creating representations such as inflow pattern of a country to the US on a per year basis and elements like a trend chart that showcases the pattern of increase or decrease of outflows out of any country to the US. The representation through animation demonstrates the rate of inflow into the US over the years.

- **Data**

  We will be using the dataset supplied by the Organisation for Economic and Co-operation and development (OECD). It provides the information of inflows of foreign population into the US by nationality. The dataset can be found at: https://stats.oecd.org/Index.aspx?DataSetCode=MIG and then selecting the variable from the

dropdown list as: Inflows of foreign population by nationality. The data has to be preprocessed to extract attributes which we want.

- **Data Processing**

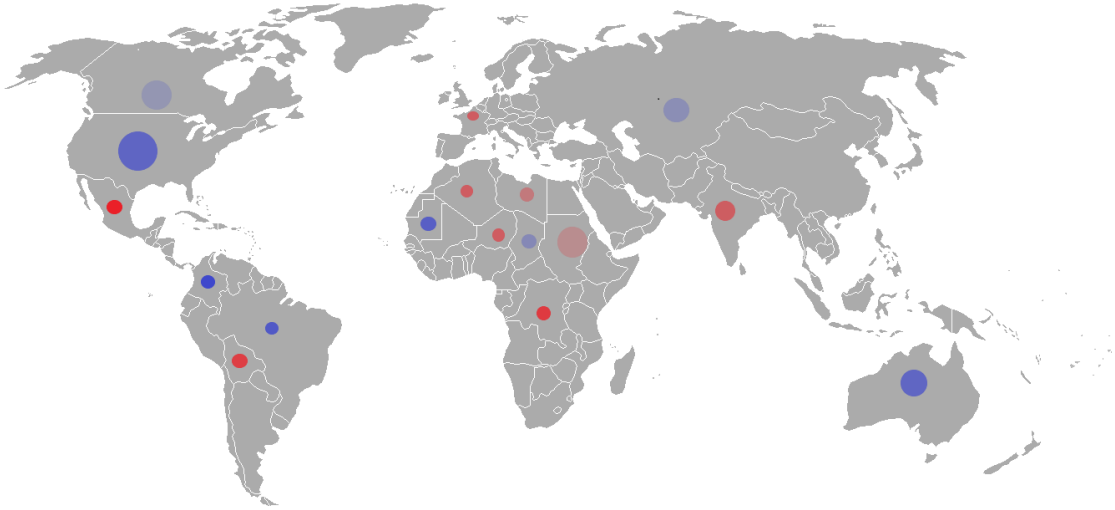
  The data is available in a .csv format and did not require a lot of cleanup. However, in order to get data on a per year basis we need to parse the data within an years range. Additionally, our dataset is a composition of many countries so in order to get the inflow data for the US we have to parse and cleanup the data for US only.

- **Visualization Design:**

  Brainstorm and initial design:

  Our initial draft establishes the major visualizations elements that will be present in the scene. Although, other visualization exists to capture the immigration inflow for multiple years but none explicitly visualizes the rate of inflow or outflow during a period of time. Thus, we wanted to make sure that our visualization does not just represent rate of inflow implicitly but user should be able to comprehend the difference in inflow during a period of time easily and compare it with other countries at the same time.

  Initially we had drafted the visualization of population inflow as opacity. We felt that it wasn't really effective in establishing movement as a form of rate of flow of people. Although, rate of flow can be visualized as a line chart, it would not be able to capture different rates from different countries at the same time. Also, we wanted our visualization to be on the map so that user can gather secondary information such as geographical location, distance as a factor in immigration etc. at the same time. Since we are trying to visualize change in data for a period of time, we knew there has to be a time slider.

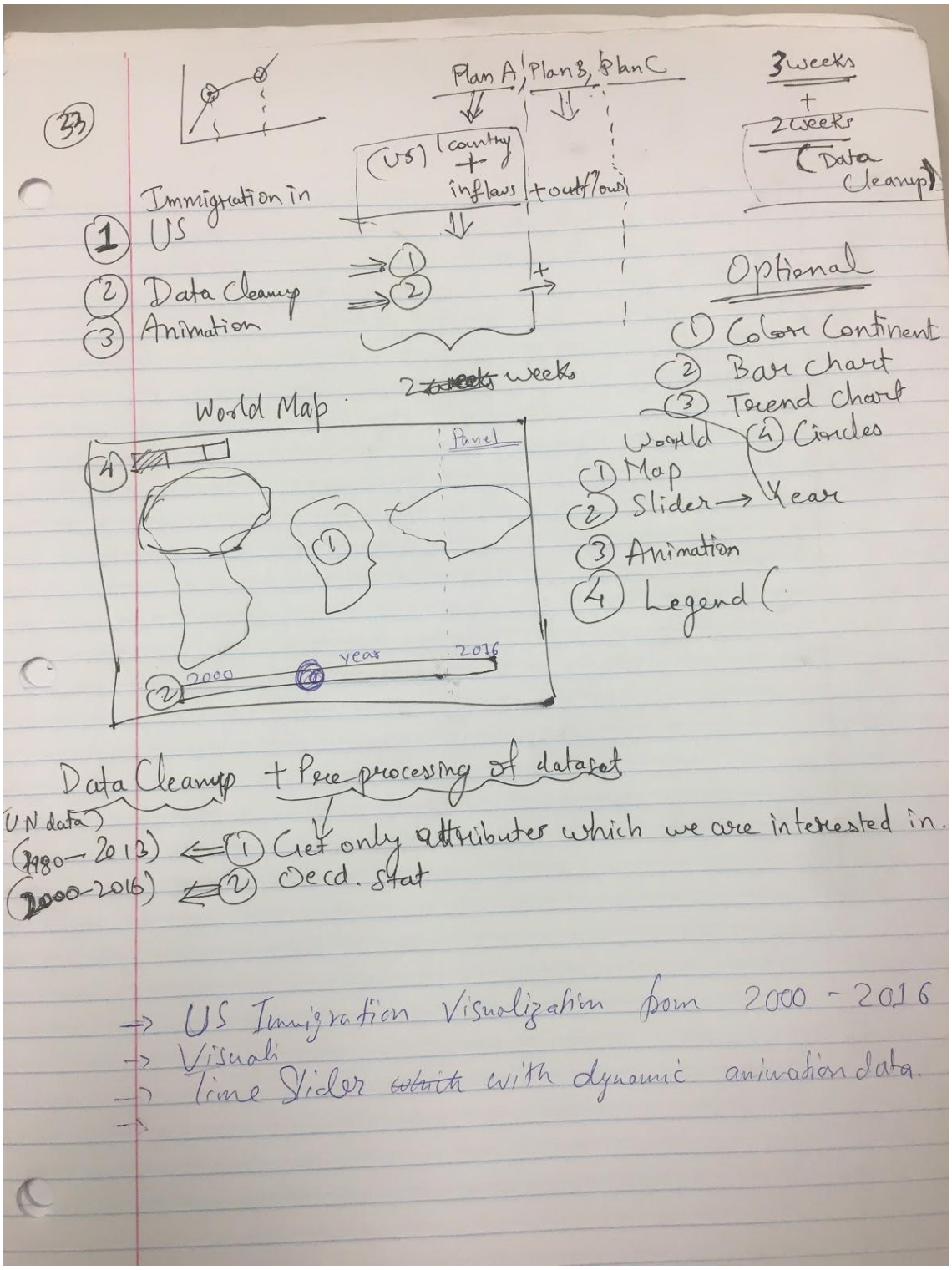


Discarded design 1. Representing net migration by radius and rate of population inflow by changing opacity.

At first, we thought of visualizing net migrating for each country by the radius of the circle and rate of population inflow by changing opacity of the circle as we move the time slider. However, we realized that change in radius is more pronounced than change in opacity so we thought of switching the role. Even with these adaptations, this approach of visualizing had one more shortcoming. With this visualization, user would have difficult time comparing rate of inflow for countries that are not near each other, even if we vary radius to represent rate. Which is why we decided to actually show connecting lines between migrating countries and try to apprehend the rate of population inflow by animating line density or by varying speed of animation.

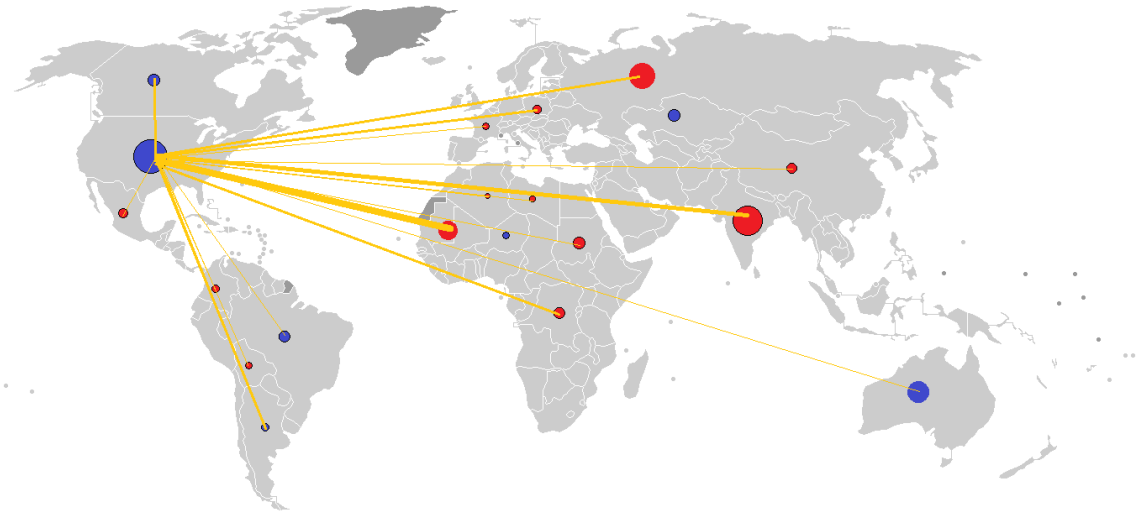
We also considered visualizing our immigration using wrap up chord diagrams. It was a good choice as it can show both migration and immigration between multiple countries in a single diagram but it suffers from same limitations as line charts and could easily become cumbersome as we increase number of datapoints. We thought of animating the diagram as we change the selected year. Cord's width will then increase or decrease based on the rate of inflow but chord diagram is difficult to grasp as it is and with animations if would make it very difficult to understand. So we decided to let go of this idea and focus on visualizing data on the map.

Discarded design 2. Wrap up Chord Diagram for immigration between limited countries.


23
Plan A
Plan B
Plan C
3 weeks
+
2 weeks
(Data Cleanup)
(US) country + inflows + outflows
Immigration in
1 US
2 Data Cleanup
3 Animation
2 weeks
Optional
1 Color Continent
2 Bar chart
3 Trend chart
4 Circles
World Map
Panel
1 Map
2 Slider → Year
3 Animation
4 Legend
year
2000
2016
Data Cleanup + Pre-processing of dataset
(UN data)
(1980–2013) ← 1 Get only attributes which we are interested in.
(2000-2016) ← 2 Oecd. stat
→ US Immigration Visualization from 2000 - 2016
→ Visuali
→ line Slider with dynamic animation data.
→

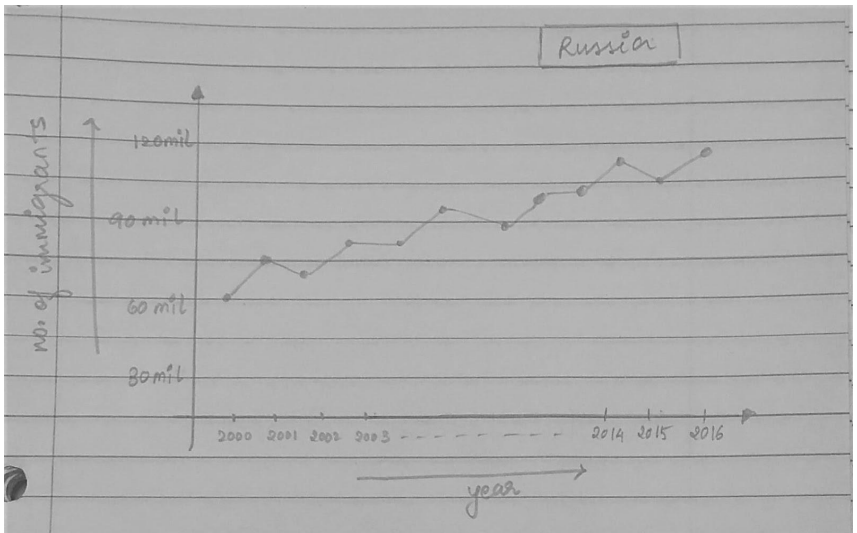
Realisation design:

We decided to actually show connecting lines between migrating countries and try to apprehend the rate of population inflow by animating line density or by varying speed of animation. As seen in the diagram below, the radius of circle represents the net migration for that country on the selected year. The color represents whether population inflow is more than population outflow. As we slide right the time-slider, the lines connecting migrating countries will animate to show the rate of population inflow.



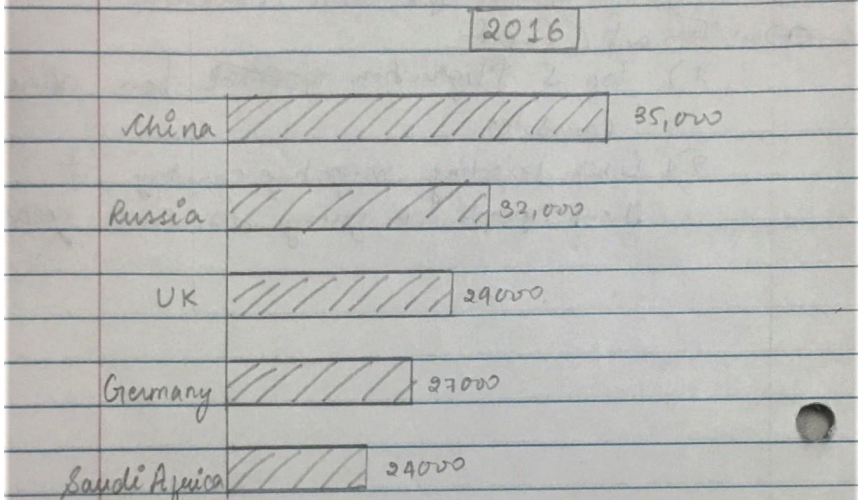
Final Draft.

On hover we can show migrating populating from selected country to United States while on the side we can show trend charts. Line chart would represent the trend of migration for the selected country. The data shown in the line chart would be the trends of migration over all the years for a given country. While the bar chart can show the top 5 countries from which the migration to US has been maximum for the selected year.



For country selected the entire trend across all years is shown.
This is a good way to analyze the change across the years of the migrants going out from their country to USA.

Top 5 Countries with highest migration of people to USA



On the slider when the year change the countries and data would get updated

- **Must-Have Features**

  Trends of Immigration in United States from year 2000 to 2016 is the sole objective of the project.

  - ➔ The most important feature in visualizing this will be to show by animation how people are migrating from other countries to United States. This would be represented by the animation of dots moving from a country to US.
  - ➔ The **animation** would be based on the rate at which the no. of immigrants are changing over the year. We still need to decide what the rate would be representing. We have two options of changing the animation based on density of points or the speed of the points.
  - ➔ **Year Slider-** this slider would dynamically update the animation and the bar chart based on the year selected.
  - ➔ **Tooltip**- would be there when you hover over the data which will show the exact numbers of immigrants for the selected country for a given year.
  - ➔ **Line Chart-** which would represent the trend of migration for the selected country. The data shown in the line chart would be the trends of migration over all the years for a given country.
  - ➔ **Bar Chart-** For a selected year on the year slider we would have the top 5 countries from which the migration to US has been maximum.

- **Optional Features**

  Here is the list of some features that we might incorporate in our system if the time permits-

  - ➔ We might include the immigration data for more the countries in the world.
  - ➔ We would add colors to the animation of dots coming from a specific continent.
  - ➔ We would include the migration from United States to other countries and design the animation based on the outflow..

- **Project Schedule**

  Week 1:

  - Rishabh: Import datasets and cleanup and processing of the data.
  - Garima: Create data structures to store the processed data.
  - Ankur: Create data structures to store the processed data.

  Week 2:

  - Rishabh: Layout a framework for project webpage.
  - Garima: Implement map.
  - Ankur: Implement map.

Week 3:

- Rishabh: Create year slider visualization
- Ankur: Create animation for inflow of foreign population by nationality
- Garima: Create linkage between year slider and inflow animation.

Week 4:

- Rishabh: Tooltip with updated data of net immigration on hover over countries.
- Ankur: Line chart to visualize migration trend, linking of line chart with other components.
- Garima: Bar chart to visualize top 5 migrating countries on an yearly basis, linking of bar chart with year slider.

Week 5:

- Final Project Due
- Finalization and cleanup**.**