
EXAMINING THE CAUSES OF HUMAN MIGRATION

KATIE WALLIS, PHILIP FANG, AMAN JIDDEWAR, ALEX SU



"MIGRATION IS AN EXPRESSION OF THE HUMAN ASPIRATION FOR DIGNITY, SAFETY AND A BETTER FUTURE. IT IS PART OF THE SOCIAL FABRIC, PART OF OUR VERY MAKE-UP AS A HUMAN FAMILY."
-BAN KI-MOON

Tableau Server Link: [Story of Migration Tableau](#).

YouTube Link: [Story of Migration Tableau - Video](#)

Overview & Key Questions

Today the world population is 6,853,328,470. Total migrants in the world are 215,738,321. Almost 216 million or 3.15% of the world population live outside of their countries. These Immigrants also includes refugees and asylum seekers, made up 16.3 million, or 8%, of international migrants in 2017. In the last century, human migration has accelerated at an unbelievable pace. Through our visualization, we wanted to examine how migration is occurring worldwide. We started by thinking about the questions "Which countries have the most emigrants?" and "Which countries have the most immigrants?". Moreover, we wanted to consider how immigration had been changing over time. Did countries have a consistent number of immigrants and emigrants each year or were there certain trends that were occurring within countries? We found, of course, that there were different migrations trends within each country. This led us to consider the question, "What factors are influencing migration in the world?", which ultimately became the focus of our project.

We knew we were going to need to narrow our analysis down to look at a select handful of countries rather than all 195 countries in the world for the scope of the project, but we weren't sure exactly how to choose which countries were the most important. We started by looking at countries with the highest raw number of migrants for 2017 but then realized that the raw number of people migrating from each country would be highly influenced by the number of people in that country. Countries like China and India that hold a high percentage of the world population, would always have more emigrants than countries like Costa Rica. We then decided to look at which countries had the most migrants as a percentage of the population. This was good in that it showed us which countries were experiencing interesting, extreme migration, but when we looked at the causes of that extreme migration, they didn't seem to be the main factors that influenced migration worldwide. For instance, the country that had the highest percentage of migrants out of the population was the island of Monceratt, where the population had abruptly decreased from 13,000 to 4,000 (~ 70% decrease) due to a volcanic explosion on the island. Instead, we decided to look at which countries had the largest change in the number of migrants leaving the country in 1990 and in 2017. While this measure was still influenced by the population, it also identified countries which were experiencing large trends in migration. From this analysis, we decided to focus on India, China, the United States, Mexico and Syria.

Immigration flow Sankey Dashboard:

For this dashboard, our overarching question was "Where are migrants from India, China, The United States, Mexico and Syria going?". We chose to use the data to answer this question in several ways. We have chosen five migration countries based on their immigration variability in the last 15 years. First, from a geographic perspective, which continents did the migrants go to? From there, we decided it would also be interesting to examine the question of "where" from a socioeconomic perspective. Using the dropdown on this dashboard, the user can choose to change the destination to show income or level of development in addition to continent. The user can also watch the graph animate over time to see how migration from these countries is changing over time. Notice, for instance, that when you watch the visualization, the number of migrants from Syria spikes massively from 2010 to 2015. After watching that spike, a user may ask themselves "Why did the Syrian population jump so suddenly?". This is the question that led us to create page 2 and 3 of the storyboard.

Syria & the Middle East Dashboard:

Looking into the political situation in Syria, it is obvious that the Syrian Civil War and other political instability in the Middle East explains a large portion of the migration away from the region. For dashboards on the Middle East and Syria we wanted to dive into the number behind the situation. We compared five indicators between US and Syria. Even before the war broke out, the political condition in Syria was already much worse than in the US. After 2011, the indicators dropped even further, especially in the areas of “political stability” and “rule of law,” both of which a direct result of war.

Mexico Dashboard:

One of the most striking aspects about Mexico from the first Sankey graph is the fact that the vast majority of immigrants from Mexico go to North America. While it is striking in the visualization, it is probably not surprising to Americans who are frequently bombarded with news from the media about all of the “bad hombres” coming in from Mexico trying to steal jobs from Americans while evading taxes. One thing we wanted to show throughout the visualization is the fact that many immigrants to the United States and to all parts of the world are coming to escape extremely difficult and dangerous situations in their home countries as opposed to coming to the United States purely because they “couldn’t get a job.” With Mexico, there were several factors we looked at such as literacy, political stability and unemployment, but we discovered the greatest differences between Mexico and other countries like the United States was the differences in crime and corruption. First, looking at the bottom right-hand corner of the dashboard, you can see that Americans enjoy almost twice the freedom from corruption than Mexicans. Additionally, we noticed a difference in rates of violent crimes such as rape and homicide in particular. The more we researched, the more we discovered that the crime in Mexico was extremely connected to changes in cartel and political leadership. Incidentally, we noticed that frequently when things that would be considered positive improvements like a prominent cartel leader being arrested or a new president coming to power, crime rates would actually increase. After speaking to my Emory friends that grew up in Mexico, they explained that when a politician like President Enrique Pena Nieto comes to power, you see the crime rates spike because the police are actually making arrests that are then being reported in the statistics. On the other side of things, when a cartel member is arrested, the instability that comes from other cartel members trying to take the arrested person’s place will cause higher instances of violent crime. It’s a difficult situation for everyone in the country and often the only option for a safer life for those not involved in drugs is to migrate. On the lower left-hand corner of the dashboard for Mexico, we visualized the correlation between increased homicide rates and yearly migration. More points would be required to come to concrete conclusions, but even with these few points the trend suggests that higher homicide rates are contributing to higher numbers of emigrants from Mexico.

India Dashboard:

In order to show a global view of immigrants from India are spreading out to the whole world, we decided to use flight path on world map with color representing different categories as our first graph. Next, we want to show the main reasons for people to leave India, so we create a word cloud with the size of the word representing the amount of immigrant and also as a highlight function for our first

graph. To give more detail and quantify information about the growth of India, we look at 4 different indicators with line chart to show the growing trend and with dots to represent the value of each year.

Since religion reason has the largest number of immigrants, we would like to dive into that on our analysis. We decided to show the proportion of Indian religious preference and the country they move to for religion reason. To show the proportion, we come up with a pie chart at first because the total of all religious preference should be nearly 100% which fit the concept of pie chart. But for the majority of the opinion in our team, we decide to change to tree map to show the proportion of religion in India. The colors represent different religions. In order to show the countries, we create another map with a globe shape to look closely on the area which most of the countries that Indian move to for religion reason. The bubbles with different colors represent different religion and they are consisting with the graph on the left

Data: Source, scraping method, cleanup, etc.

Sources can be found below in References.

The UN Data originally had each destination country and then a column under the name of every origin country that contained the number of migrants to the destination country. As shown below:

The first thing we needed to do was transform the data which was in wide format to long format. We did this in R using the melt function. Consequently, we ended up with a very large dataset. Many of the rows in this dataset were blanks, indicating that there weren't any migrants from the destination country to the origin country in that year. To reduce the size of the dataset we removed these rows. Nonetheless, the dataset was still too large for Tableau to handle in an efficient way. From here, we did our analysis on

which countries had the largest change in the number of migrants away from that country using a pivot table in excel. We were able to aggregate the number of migrants from a given country to all other countries and compare that aggregate value in 1990 to the aggregate value for the country in 2017. Once we had a ranking, we chose the top five countries and put them into the dashboard.

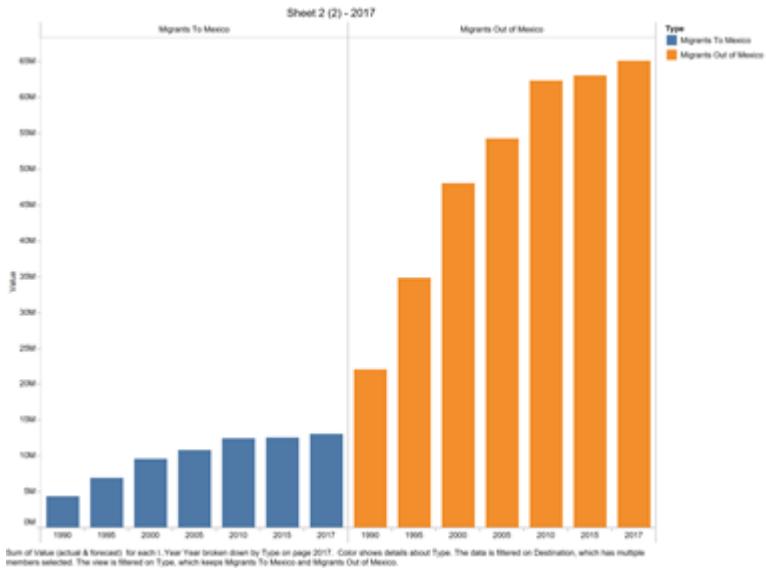
When creating the dashboards on individual countries, one common problem we ran into was the fact that we had to merge several datasets together to get the full picture of that country. For instance, crime data was frequently in a different dataset from information about corruption. Combining these datasets wasn't always easy. Different datasets would cover different time periods and sometimes the names of countries were inconsistent across datasets (some would say United States and others would say USA). Because we wanted to examine the interactions of different indexes, combining the data in a way where Tableau would let us show the measures on the same chart was a challenge. One other interesting challenge we ran into when creating the crime map for Mexico was that the names of the states and the names of the types of crimes in the dataset were in Spanish. Fortunately, Katie speaks Spanish and was able to translate, but we ran into issues with the Tableau map recognizing the location of the states. Ultimately, we solved the problem by changing the location of the workbook to Mexico so that it knew to interpret the geographical names in Spanish.

Exploratory Data Analysis:

Primarily, We looked mostly at bar graphs when comparing values from economic indexes because we felt that bar plots gave us a really easy quick feel for how values compared. When we were looking at time series data, we primarily started by looking at line plots so we could see the trend. Once we got a feel of how the indexes compared across countries and across time, we moved on to seeing correlation between the different indexes. We made many graphs that showed similar messages but needed to narrow down exactly what we wanted to show. For this reason, we decided to focus in on crime in Mexico.

There were several graphs that we played around with that did not ultimately get included. Those graphs are shown below. We decided that the best way to convey our quantitative message, however, was to focus on graphs that really made the point clear for each country. Our insights definitely pushed us to the conclusions of Crime, War and Religion and so our designs came about as a product of that evaluation.

This chart we played with explored the differences in the number of migrants entering Mexico versus leaving mexico.



One thing that we found interesting was the differences in happiness amongst the countries.

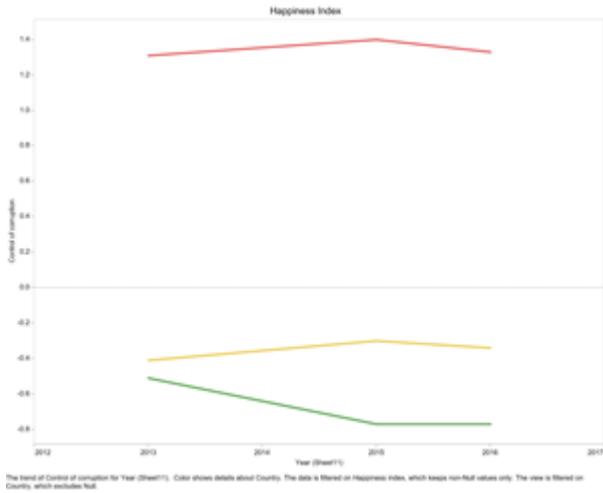


Tableau is pretty bad with big datasets. Hence, we cleaned our data in R and Excel by aggregations and filters. This reduced the size of data giving us the speed in visualization. In any visualisation, performance of the dashboards is as important as the visual part. Users like interactions with dashboard, but hates lagging and slow dashboards.

Design Evolution: What are the different visualizations you considered? Justify the design decisions you made using the perceptual and design principles you learned in the course. Did you deviate from your proposal?

For the first dashboard with the Sankey graph, we knew what we wanted to make one before we looked too much at the data. However, it wasn't until after we examined our data that we found which countries we wanted to show the migration for. Obviously, it would be infeasible to show all of the countries individually. One key consideration we had while making the Sankey graph was whether to include countries like China and Iraq which we did not create in depth dashboards for. Given our time restriction,

we knew we would not be able to create all 5 dashboards; regardless, we chose to include the fields because we thought it made the overall graph more interesting and informative.

We did deviate somewhat from our proposal. Our motivation and overarching questions were still the same in that we wanted to look at migration over time. However, we decided to examine countries that had high levels of migration out of the country as opposed to looking at the way that migration affected life in the United States. The main reason for this was because it is difficult to draw conclusions over how immigration is affecting the United States by just looking at how things correlate over time given the number of political and social factors that could be the root causation.

Overall Design Aspects:

Color: Each country in the Sankey graph was assigned a color in keeping with the flag of that country. The color chosen for each country on that graph was then used as the primary color of the in-depth analysis dashboard for the country. By doing this, users can more easily connect the dashboard together with the insights they get from the Sankey graph.

Font: We chose to use Tableau SemiBold for our titles and Tableau Book for the non-heading elements. By keeping the fonts consistent across dashboards, the overall style of the story remained cohesive. These fonts were selected in particular because they were easy to read on the screen. Font sizes were also all kept consistent across dashboards.

Layout: All of the country level analysis had similarly designed headings with a bolded, colored title and the countries flag. This consistency added to the cohesion of the document as a whole.

Implementation: **Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.**

Sankey Graph:

The Sankey Graph allows the user to hover over a given country to highlight it and see where migrants are going to. It also allows for the converse where you hover over a destination to highlight where the migrants are coming from. Moreover, using the dropdown menu, you can choose to change the destination from continents to levels of development and income. Through this, you get to examine migration from multiple angles. Lastly, using the page control, the user can have the animation play over time and can see how migration is changing.

We used Sankey chart to show where are the immigrants from countries moving. We have the data for 1990, 1995, 2000, 2005, 2010, 2015, and 2017 years. The color palette is consistent across all the views.

We chose Sankey chart for showing migration because it gives a clean visual to show the flow of people intuitively. Other possible chart options to explore were flowcharts, stacked bar charts, and multiple bar charts.

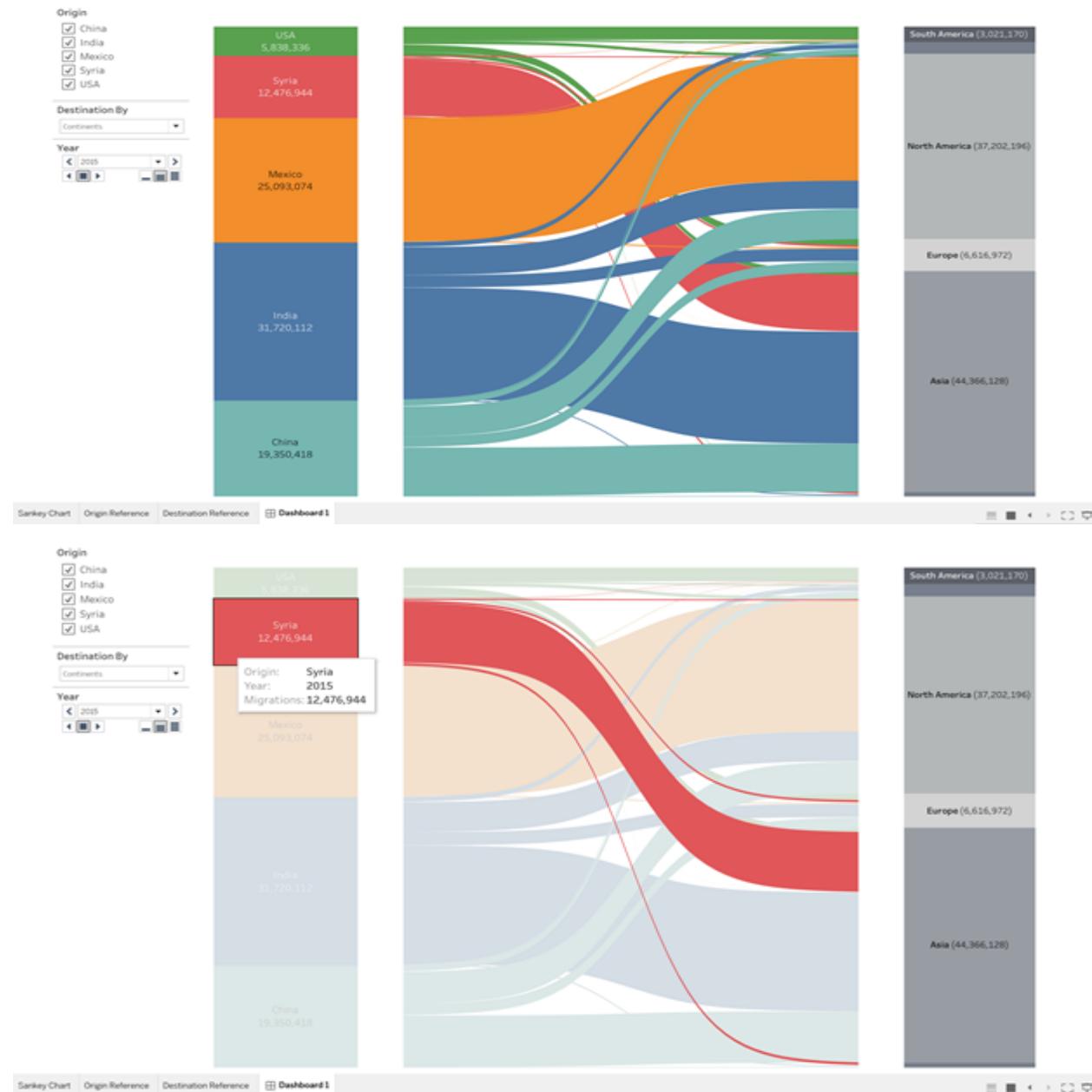
The chart puts the load on the tableau server, so we prepared the data in R to have only relevant rows. A sequence of calculated fields were created for showing the curved part of the Sankey chart and synchronizing both ends of the chart.

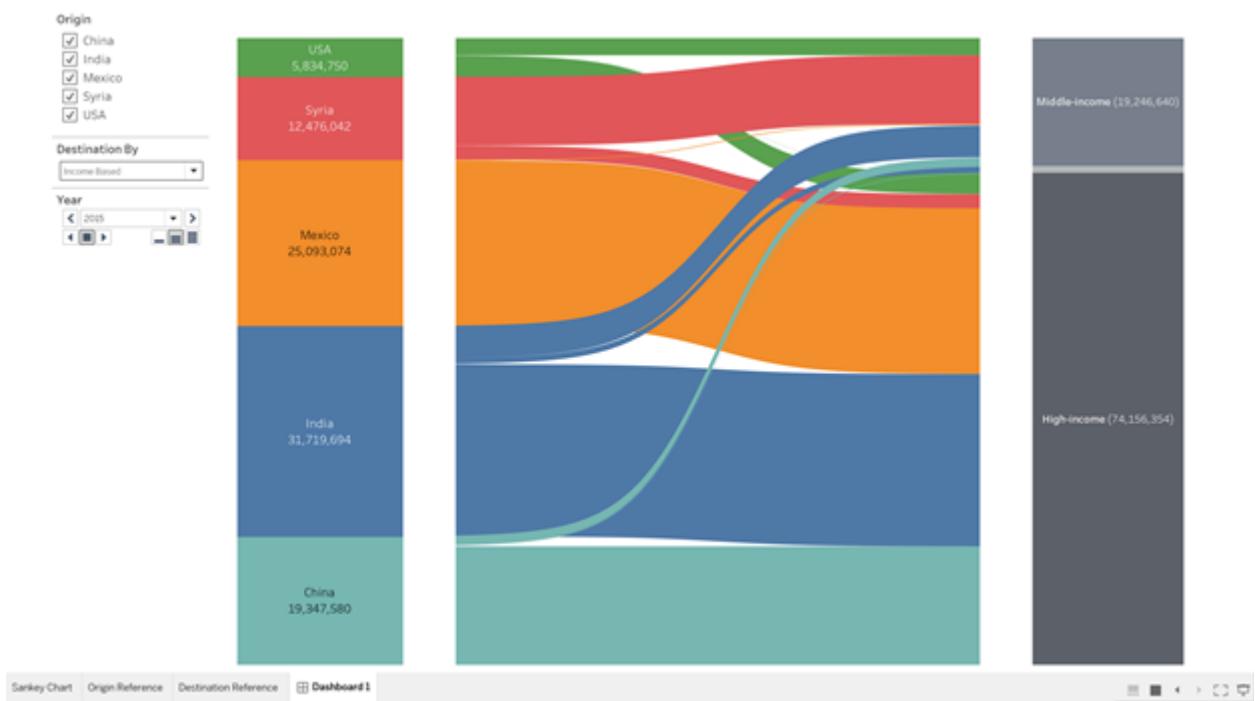
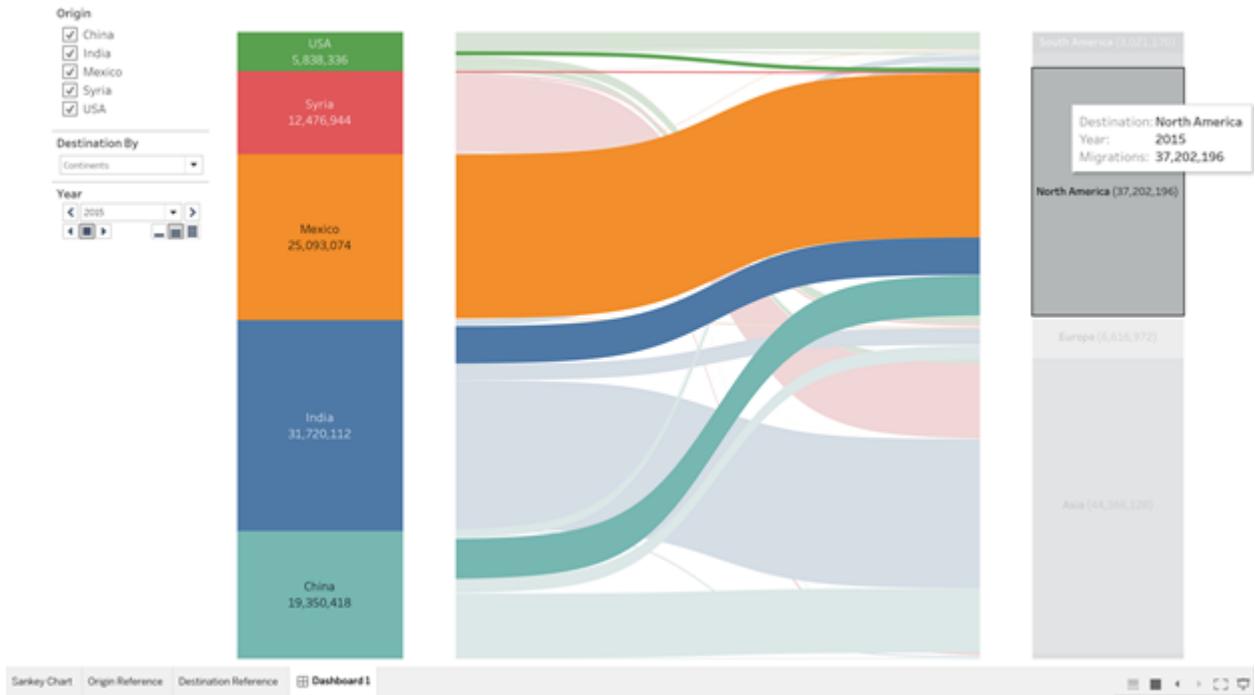
The filters on left are used to select the origin countries. The destination is grouped as the continents, income based, and development status.

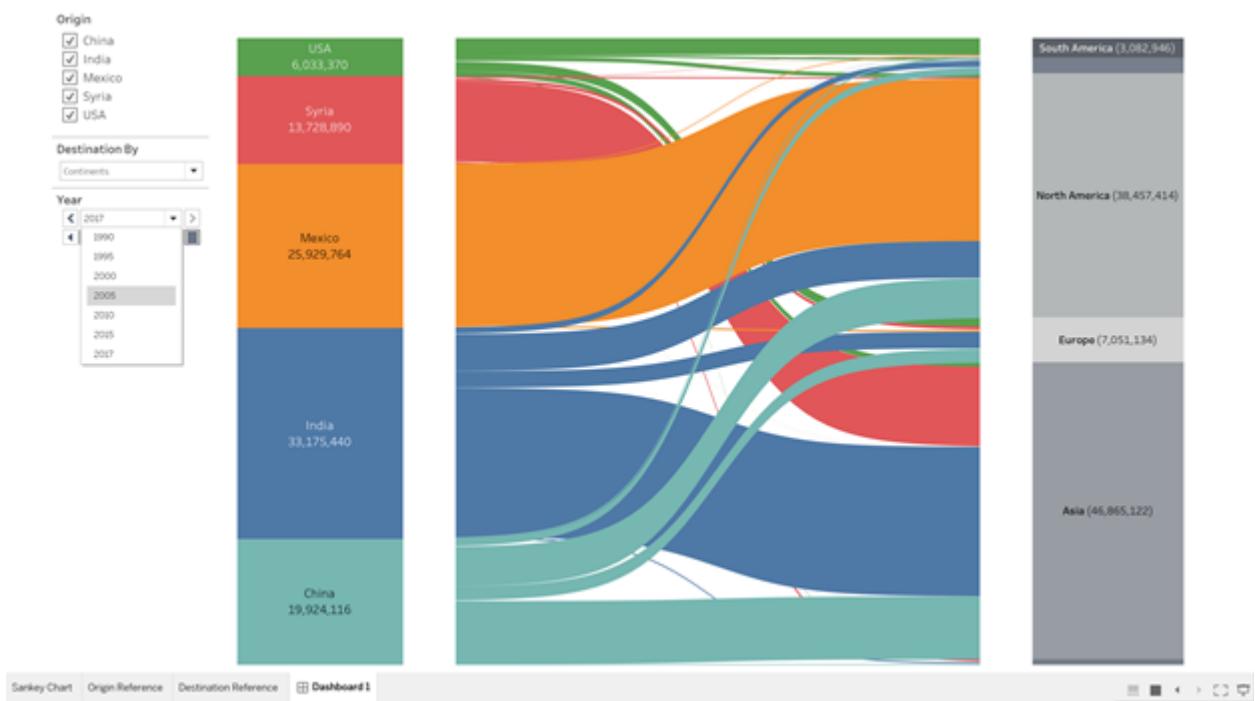
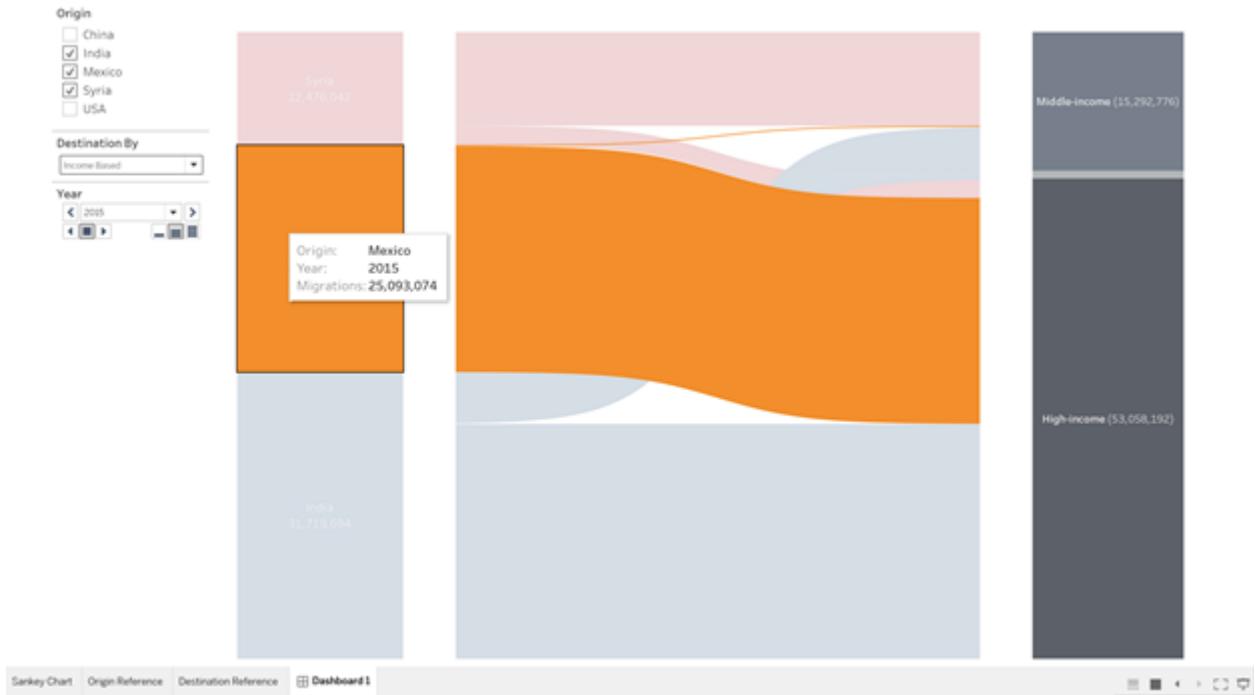
Continents: South America, North America, Asia, Europe, Africa, and Oceania.

Income Based: High-Income, Middle-Income, and-Low income.

Development Status: More developed, Less developed, and Least developed.







Syria & the Middle East:

As we see war as an important factor for some countries, we try to take countries in Middle East as an example.

First, we took a dataset called “Conflict Site Dataset” [1] from PRIO which provides the year and location information on war that happened from 1989 to 2008. After filtering out the Middle East region, we found that there’s an average of 40 wars going on every year. By saying war, we haven’t even talked about each single battle that went on during each war. For each country, there are several regions that often run into conflicts. The radius shows the affected region of each war in the unit of kilometers. Iran and Afghanistan are two countries that are most severely devastated by war. They almost suffered from war every year during the period.

The color we used here is red. We used the color and size to show the influence of each war. The interaction we put in this dashboard is Year to show the different countries where each war occurred.

In order to take a closer look on one country and the near future, we take Syria as the target. Syria Civil War broke out in 2011 and has continued till today. It has now become a worldwide problem since the war has led to the appearance of many refugee and has created instability even in many developed countries. This time, we want to have a glance at the immigrants.

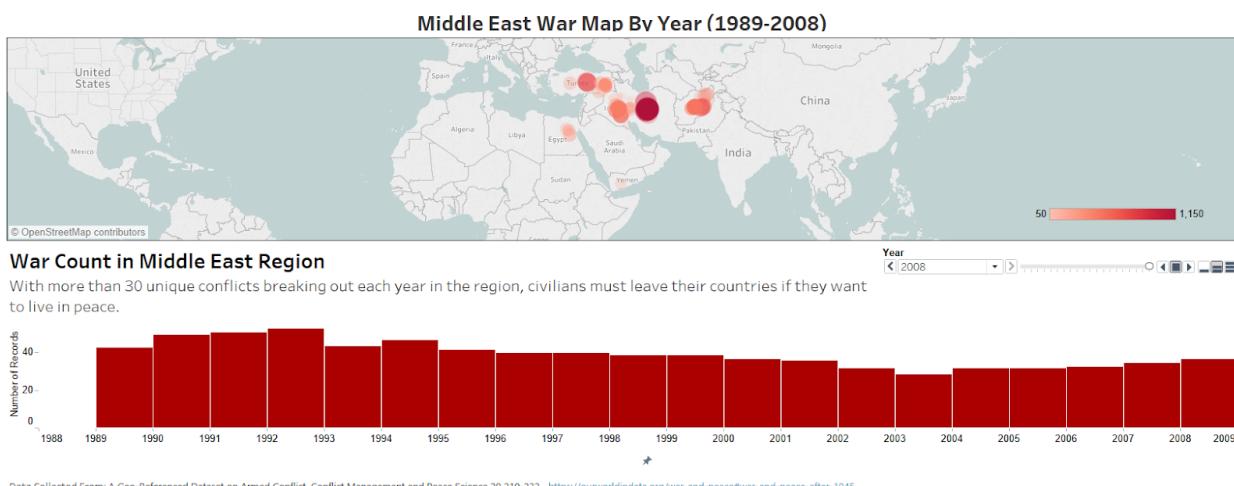
The first aspect we looked at is the politics. We compared five indicators between US and Syria. Even before the war broke out, the political condition in Syria was already much worse than in the US, but rather stable. After 2011, the indicators deteriorated, especially for political stability and rule of law. These two indicators are the direct victim from war.

We used two different colors other than red and a bar chart to show the comparison between two countries. A filter dropdown is applied to demonstrate to choose which year to look at.

Next, we demonstrated aggregate number of deaths from 2011 to 2016 and per 1000 habitant deaths every year. The plot on the left tried to tell the number of deaths of each region by the size of the circle. The one to the right is more like a rate over time. The slope is the steepest around 2011 when the war broke out.

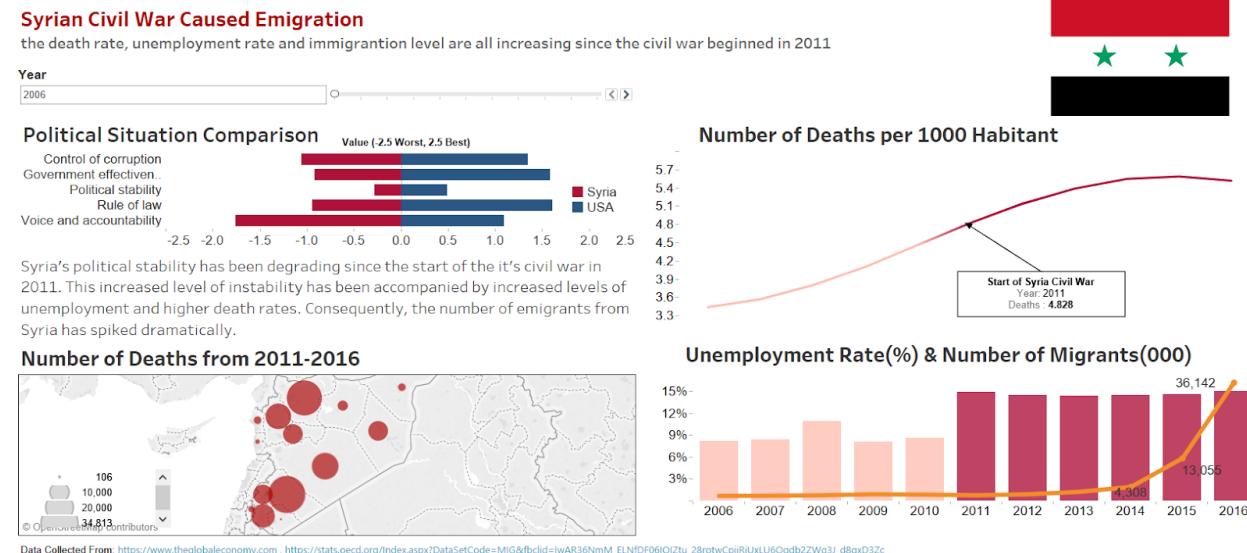
Conflict in the Middle East leads to Migration in the Region

The Middle East region has always been suffered from war for decades. More than 30 major conflict events break out every year. Consequently, many choose to leave the region.



One significant change after 2011 is the unemployment rate. Before 2011, the rate was around 8% while almost doubled to 15% after that, which is pretty high.

Finally, we plotted the number of immigrants by year. The number was below 5 thousand, until 2015 when the number significantly grew. This is an interesting result since the war started in 2011. The insight behind this is that, when the war begins, the number of refugee grows immediately. But they are not directly accepted as immigrants to other countries until a year or two later when all the process has been finished. Also for those people whose life is not in danger but are affected by the unstable politics and business, they have to wait for several years before they can transfer all their properties abroad and become an immigrant.



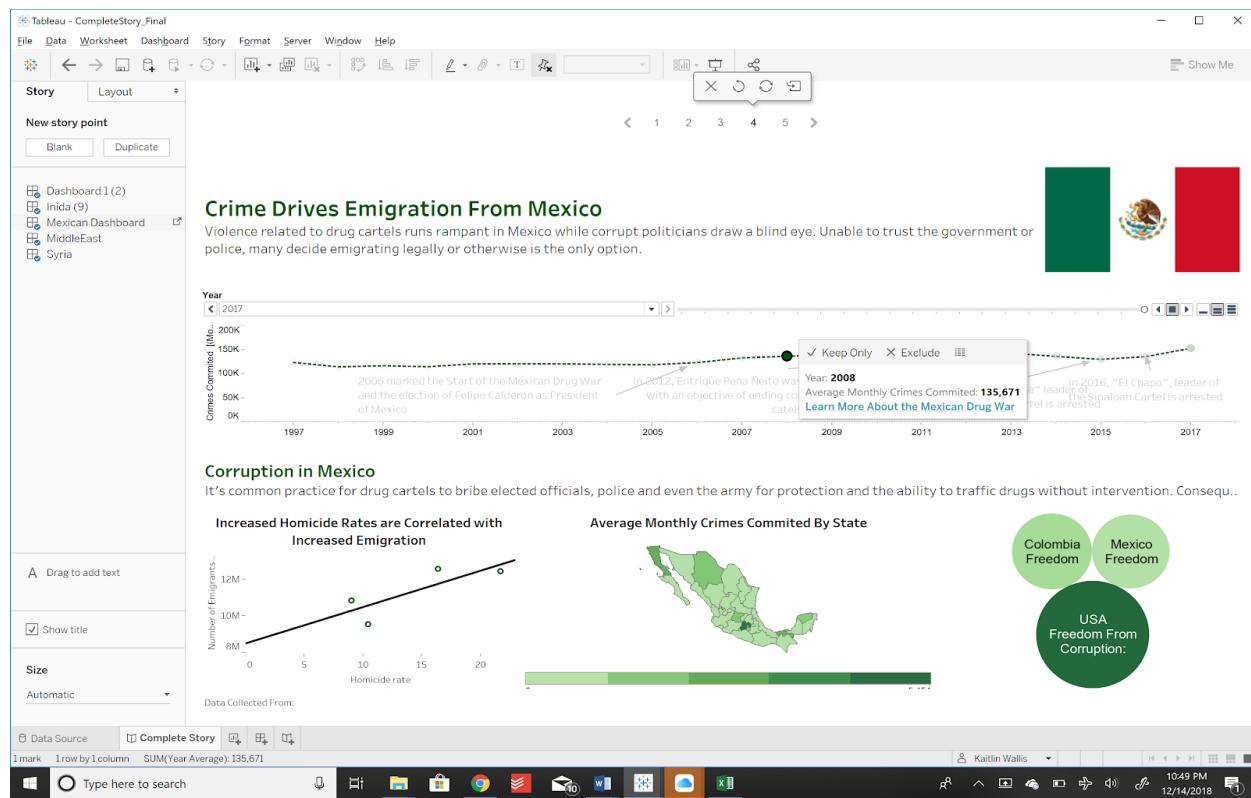
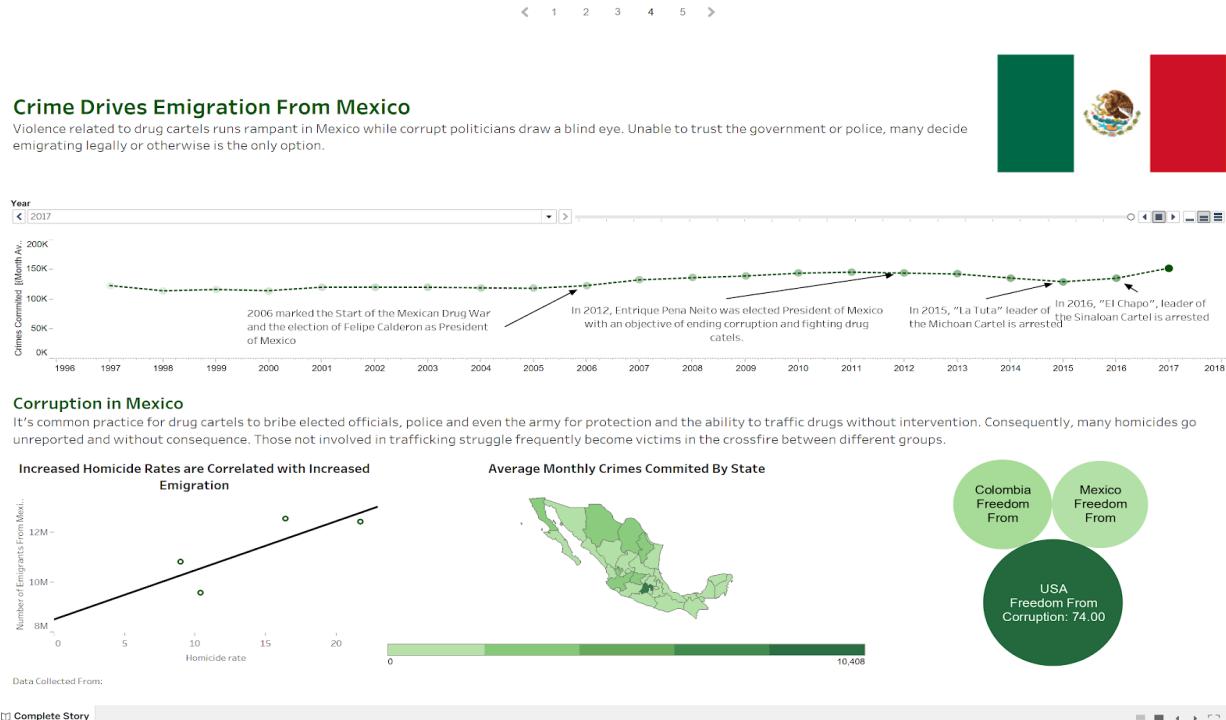
The color we chose here is red to show distinction from the first dashboard. The lighter color is chosen for the data before 2011 and darker red for after 2011.

It is also interesting to see how the preferences of migration destination for Syria is different from other countries. Our data shows that Syrian people are migrating to low-income and least developed countries as much as to countries which are better in terms of economy.

Mexico:

On this dashboard, we first see that the line graphs creates itself over time. We did this using the show history option on the page control. Additionally, as the line grows, different annotations appear explaining the history behind what may be causing changes in crime. You can also click on any of the points on the timeline and get an option called “Learn more about the Mexican Drug War” by clicking this link, you are directed to a Wikipedia timeline of the Mexican Drug war. If you are clicking from a point after 2006, the page opens to the year you’ve selected on the Wikipedia timeline. If you’re clicking on a time before 2006 (before the war started) the page opens to the top of the page. As the timeline animates by year, the crime map of Mexico at the bottom also animates in conjunction. Lastly, on the bottom left of the page, you see a correlation plot for number of emigrants out of Mexico and the homicide rate in Mexico. This graph shows the user that as homicide rates increase, so do the number of emigrants proving that there is a relationship between crime and migration. There is also a reference line on the chart showing the homicide rate in the United States in 2017. This reference indicates that even on the years where the homicide rate in Mexico is at its lowest, it is still more than double that of the United States. On the bottom right, we also see a bar chart which shows a Freedom from corruption

index for the US, Mexico and Colombia. Here, you can see that corruption in Mexico is much worse than the US and even slightly worse than Colombia.



India:

We want to look at India because the number of immigrants from India increased dramatically since 1990 and it is the top source of international migrants, with one-in-twenty migrants worldwide born in India. So, we want to find out the reasons behind this huge number of immigrants.

India has been among the world's top origin countries of migrants since the United Nations started tracking migrant origins. The number of international Indian migrants has more than doubled over the past 25 years, growing about twice as fast as the world's total migrant population.

We found that different from other countries that has huge number of immigrants, India remains positive growing in the past 25 years on GDP and education rate. It seems like more Indian were leaving countries because of becoming wealthier and well-educated. An estimated 17 million Indians were living abroad in 2017, making India the largest source country for international migrants globally.

Extending on first dashboard, which showed the migration from India to places with high or middle-level average income is high. According to the research and the data visualization, we categorize the reasons for Indian to leave their country into 3 categories and other. First is for work, India receives more remittances from migrants than any other country. About \$69 billion was sent by Indian migrants to family and friends in India, amounting to roughly 3% of the country's gross domestic product, according to the World Bank estimates. Most of the money comes from Indians living in Persian Gulf countries as well as the U.S., the UK and Canada.

Second is for religion, which is also the generating most number of emigrants, India's religious minorities have been more likely to migrate internationally. Religious minorities make up a larger share of India's international migrant population than they do among the nation's domestic population, about 20% of the Indian international migrant population was Christian, compared with only 5% of the population in India. Similarly, an estimated 27% of the Indian international migrant population was Muslim, compared with 14% of the population in India.

Third is for education. Take Australia as an example, nearly three lakh Indians migrated and settled in Australia, according to official figures, which said India contributed the maximum number of individuals on skilled visa. As many as 291,916 Indians migrated to Australia within recent 6 years, of which 154,012 individuals have acquired Australian citizenship.

< 1 2 3 4 5 6 7 8 9 >

Most Indian Move for Religious Reason

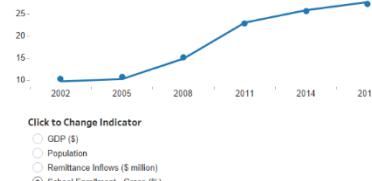
As India has become wealthier, more of its citizens are leaving its shores for better job or education. An estimated 17 million Indians were living abroad in 2017, making India the largest source country for international migrants globally, up from seven million in 1990 and a 143% increase.



Religion

Work
education other

School Enrollment - Gross (%)



Click to Change Indicator

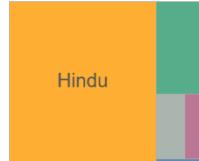
- (○) GDP (\$)
- (○) Population
- (○) Remittance Inflows (\$ million)
- (●) School Enrollment - Gross (%)

<<<<<

Religion of Majority

Main countries that Indians move to for religious reason

>>>>>



India's religious minorities migrate internationally

Religious minorities make up a larger share of India's international migrant population than they do among the nation's domestic population, an estimated 27% of the Indian international migrant population was Muslim, compared with 14% of the population in India. The reverse is true for Hindus: Only 45% of India's international migrant population was Hindu, compared with 75% of the population in India.

Data Collected From: <https://ourworldindata.org/>, <http://www.pewforum.org/>

Complete Story Dashboard 6 Dashboard 7 Dashboard 8 Dashboard 9

< 1 2 3 4 5 6 7 8 9 >

Most Indian Move for Religious Reason

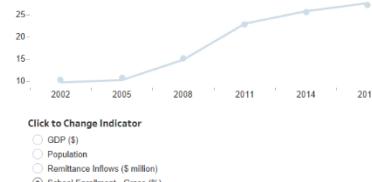
As India has become wealthier, more of its citizens are leaving its shores for better job or education. An estimated 17 million Indians were living abroad in 2017, making India the largest source country for international migrants globally, up from seven million in 1990 and a 143% increase.



Religion

Work
education other

School Enrollment - Gross (%)



Click to Change Indicator

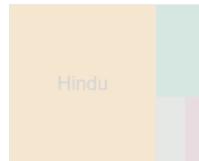
- (○) GDP (\$)
- (○) Population
- (○) Remittance Inflows (\$ million)
- (●) School Enrollment - Gross (%)

<<<<<

Religion of Majority

Main countries that Indians move to for religious reason

>>>>>



India's religious minorities migrate internationally

Religious minorities make up a larger share of India's international migrant population than they do among the nation's domestic population, an estimated 27% of the Indian international migrant population was Muslim, compared with 14% of the population in India. The reverse is true for Hindus: Only 45% of India's international migrant population was Hindu, compared with 75% of the population in India.

Data Collected From: <https://ourworldindata.org/>, <http://www.pewforum.org/>

Complete Story Dashboard 6 Dashboard 7 Dashboard 8 Dashboard 9

Technique:

The first map we basically put two charts together, one is by putting dot on the destination and the other is to connect the dot together by creating a path. In order to do so, we have to set the longitude and

latitude for the destination and set the order for each countries so that Tableau knows how to connect the dots in a certain order. Next we import a new map in the Mapbox to get more options of the map. For the circle Map on the below left side, we used the Seattleite map in Mapbox and create a transparent circle with a white border to cover the rest of the map. For the line chart on the upper right side, we did a lot of data combining and transforming to combine all the indicator's data together. So that we can show it in a single chart.



Insights and Self Evaluation: What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

We really enjoyed the challenge of this final project. Something that we learned more so in the final project that the rest of the class was how important data visualization is in the exploratory process. It's easy to think of reasons why one might use data visualization to convey insights once they already knew what they wanted to show, but we found through this process it was just as important to use data visualization to discover what those insights were in the first place. Our dashboard did a good preliminary job of answering the questions that we asked. Our questions were really large, important questions that can't be answered very easily in one visualization. They involve a lot of different opinions and every migrants story is different. One thing that would have been interesting would be to look at migration more by percentage of the population and to obtain data from migrants themselves as to why they were migrating. We did look for this causal data, but were unable to find any reliable sources. A lot of our insights came from the conjunction of analyzing data and conducting academic research. The process was really important for us to go through in such a technical program. Had we had time to do causal analysis that would have been good, but looking at the political science research as to the causes

of migration also pushed us to think critically about people's motivations and not rely exclusively on numbers.

References:

1. Dittrich Hallberg, Johan 2012. PRIO Conflict Site 1989–2008: A Geo-Referenced Dataset on Armed Conflict. *Conflict Management and Peace Science* 29 219-232
2. Global Economy data: <https://www.theglobaleconomy.com/download-data.php>
3. Our World In Data: <https://ourworldindata.org/war-and-peace#war-and-peace-after-1945>
4. OECD for immigration data:
https://stats.oecd.org/Index.aspx?DataSetCode=MIG&fbclid=IwAR36NmM_ELNfDF06IOIZtu_28rotwCpjRiUxLU6Oqdb2ZWq3J_d8qxD3Zc
5. Crime in Mexico Data
https://l.facebook.com/l.php?u=https%3A%2F%2Felcri.men%2Fen%2Fdatos.html%3Ffbclid%3DIwAR2kUdnf3DAzS-okOnSZIF3Ao8GATwkg8mACsBn8dov1YuljUAu9Z_-YwM&h=AT1hqsYd-IlfpoxW5lgHYRU60Rot110HRhThbOnZK9Cd9AzZfLwpWlvjcKRfeGFqPaiUoTKu-plXjonPYprj5RaJm2l9s5BrYrzPryoO3aBGOfeR8oUaPVr7gNqpJaGJBGg
6. Mexican Drug War Wiki:
https://l.facebook.com/l.php?u=https%3A%2F%2Fen.wikipedia.org%2Fwiki%2FTimeline_of_the_Mexican_Drug_War%3Ffbclid%3DIwAR0CnjGL3tDJQxf4e6xosROSn5mom80HrPktyRXryTF4hpr9aTDsxsyz2-4%232006&h=AT1hqsYd-IlfpoxW5lgHYRU60Rot110HRhThbOnZK9Cd9AzZfLwpWlvjcKRfeGFqPaiUoTKu-plXjonPYprj5RaJm2l9s5BrYrzPryoO3aBGOfeR8oUaPVr7gNqpJaGJBGg