

Refugee Crisis



**Eric Hagen
Rob Kuvinka
Reema Naqvi**

Project Goals

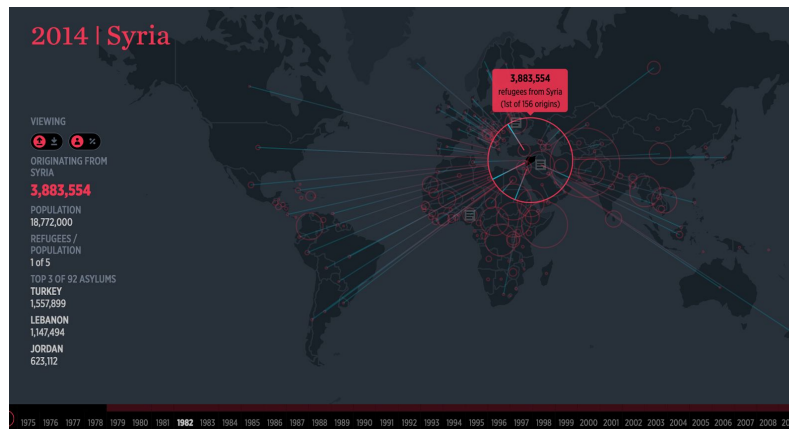
The goal of our project was to present information about the 'Refugee Crisis' that has been featured prominently in Western news media over the last half decade. We wanted to present information that would educate someone unfamiliar with the crisis. The goal for this user would be that once they had completed our narrative and visual walkthrough covering major trends, the user would have an understanding of where these refugees were coming from and to which countries they were commonly going to in order to seek asylum. In order to get to this point, we would begin by scoping the countries for inclusion in our analysis. We did this by selecting countries that were directly associated with Arab Spring protests and uprisings, beginning in 2010. The goal from there would be to persuade the users, through clear visual presentation, that there exists a small minority of countries that were contributing the majority of refugees to this crisis situation. After these countries of interest were established, we would drill down into conditions and facts of interest about the affected populations in these countries. Finally, we would present information to the user that would serve to educate as to the countries in which these people sought refuge.

It is important to note that we did not attempt to produce an argument as to what conditions triggered large increases in the number of refugees. We simply sought to present a narrative that discussed refugee increases and movement patterns, along with facts relevant to providing as much context as was reasonable to help understand these patterns.

Related Work

Stories and reporting about the refugee crisis regularly appear in news media. While we do much of our data acquisition through highly reputable sources, we were able to cast a wider net in terms of gaining inspiration for our visuals through various news and non news outlets.

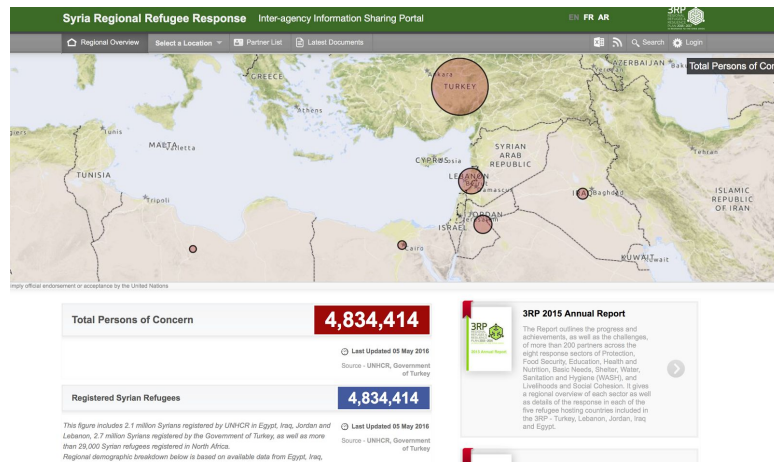
1. The Refugee Project



An inspiring visualization that effectively and comprehensively represented migration data was from The Refugee Project. Developed by Hyperakt Labs, they sought to make UNHCR’s data, the same data source we used, into “an interactive experience that illuminates where and when refugees emigrate, as well as the complex stories of political, social and economic turmoil behind each displacement.”

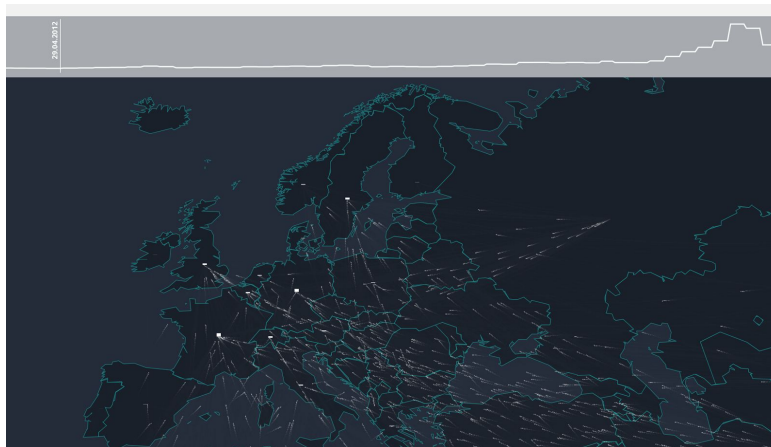
The biggest strength of this visualization is that it allows clicking on a country, and then arrows shoot out to all those countries where people are going/coming from. Hovering over any of these arrows gives the name of the other country and the number of people on that path. What is also really nice is that the viewing perspective can be switched from immigration to emigration. The tooltip also shows the rank of the country (e.g. Syria is 1st of 156 origins as seen in the screenshot). The circles serve the same function as the tooltip rank (bigger circle = higher rank) but they are actually quite confusing as many overlap and it takes careful hovering over them to trigger the right country. (<http://www.therefugeeproject.org/>).

2. Syria Regional Refugee Response



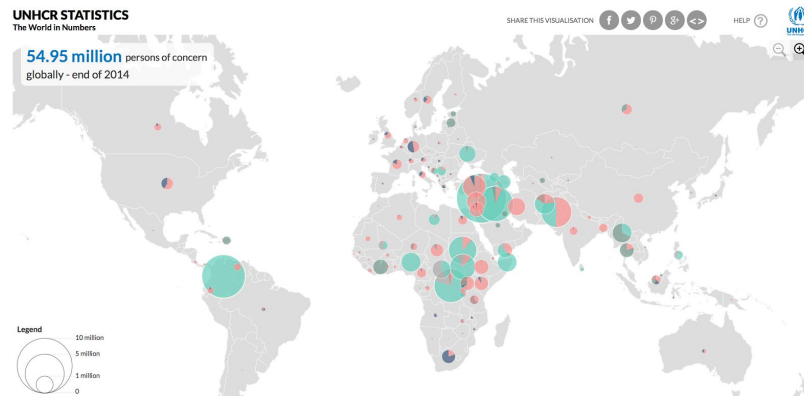
The Syria Regional Refugee Response provides a good starting place to look at the state of the Syrian crisis. This page includes visualizations on refugee growth trends as well as demographics (age and gender). This page is updated frequently by the UNHCR (May 5th 2016). <http://bit.ly/1jsBiUu>

3. The Flow Towards Europe



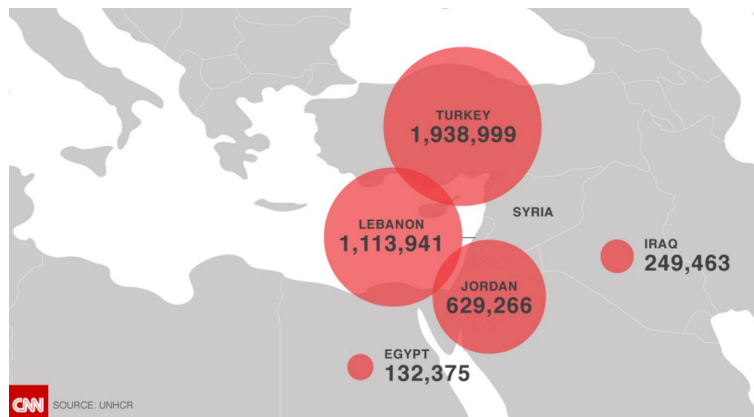
This impressive visualization does a nice job using animation to show the flight of refugee to Europe. First, it has a custom animation which effectively shows where the refugee are flowing from and to. Secondly, it uses a Sankey diagram to display the same information. Both these visualization were likely coded in d3. <http://www.lucify.com/the-flow-towards-europe/>

4. UNHCR Refugee Overview



This is the main tool used by the UNHCR to visualize the refugee crisis. Given that the number of refugees are coded by the radius of the circle and that most refugees are from the same part of the world, the bubbles have a tendency to overlap. Only when you zoom in does this get fixed, but consequently, you lose the global view. This is poor design, and we decided to stay away from population bubble maps. <http://popstats.unhcr.org/en/overview>

5. CNN

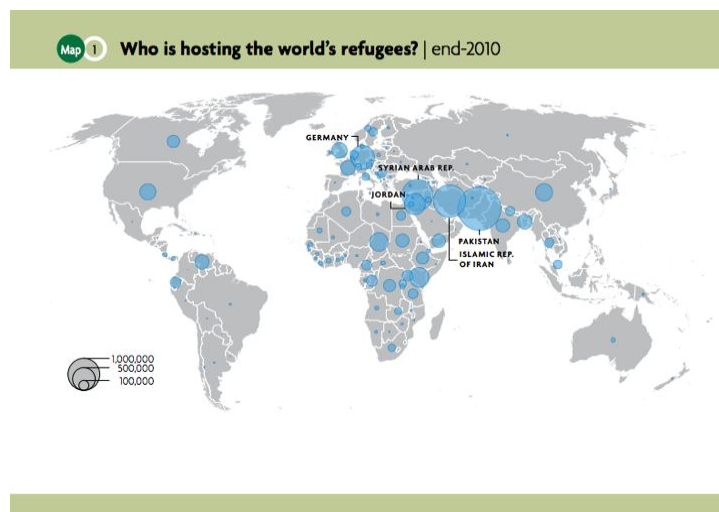


Though we scoped our project such that we did not exclusively focus on Syria, this was a very good reading from CNN that gives a very detailed description of where Syrian refugees are going, what countries have closed their doors off to them, what countries are expected to accept them but aren't (i.e. the Gulf Cooperation Council).

<http://www.cnn.com/2015/09/09/world/welcome-syrian-refugees-countries/>

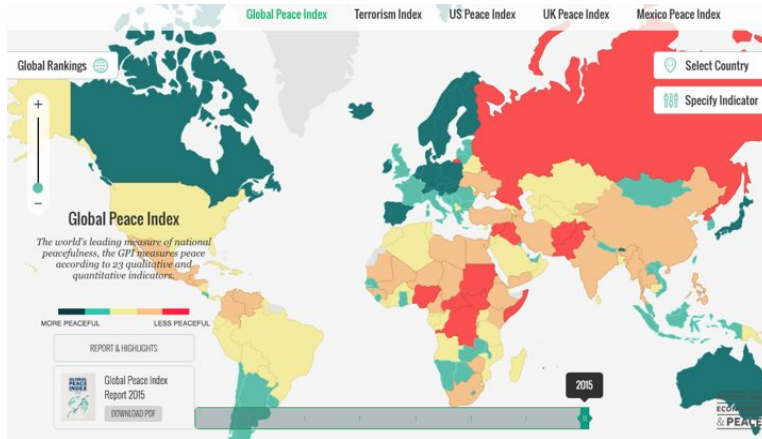
6. United Nations High Commissioner for Refugees, “UNHCR Mid-Year Trends 2015”

This graphic became the initial inspiration for the original map that was featured in our first Tableau dashboard. I thought this graphic did a nice job, through increasing the size of the bubbles to represent differences in quantity, of displaying a map that afforded comparisons between different countries in terms of the number of refugees the countries accept. I also thought it would've been nice to build a map like this that allowed for the use of tooltips to obtain the actual numbers underlying the sizing. This entire UNHCR report was highly informative through and through, and certainly was an essential read to understand the scope of our project. <http://www.unhcr.org/56701b969.html>



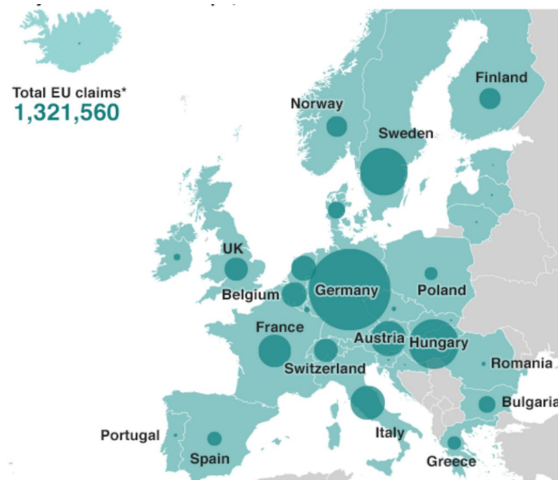
7. Vision of Humanity, “Global Peace Index 2015”, Institute for Economics and Peace

Again, the entire report from the Institute for Economics and Peace was an essential read to further understand this problem. While this source did not deal with refugees, it did display changing conditions over time. The slider at the bottom of the screen served to allow users to examine peace conditions in different years. A big part of our project would center on how refugee conditions changed dramatically with time. This visual, and the data supporting it, would serve to further develop our Tableau dashboard presenting an overview of the refugee crisis. <http://bit.ly/1ixJqjv>



8. BBC

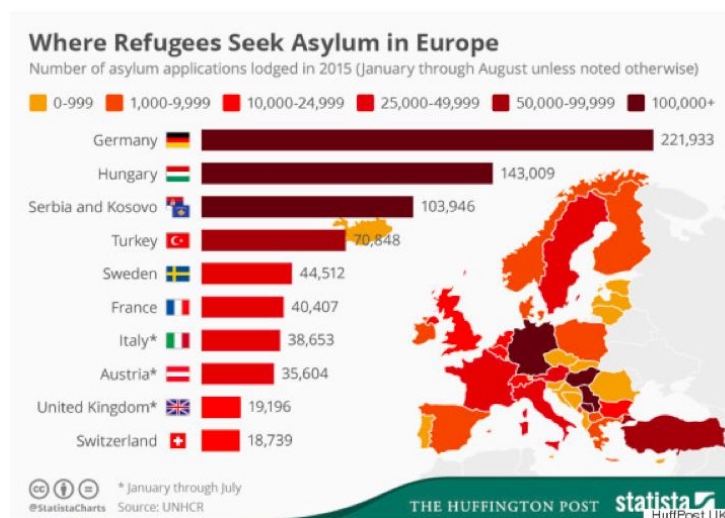
While most of the research in our topic is being covered by news agencies, the BBC had an article in March, 2016 that used a number of static visualizations accompanied by brief stories to encapsulate the refugee crisis (which BBC insists on calling the migrant crisis). This news story was a good starting point to start making sense of our data. The problem with this piece, like most others, is that it focuses too heavily on the burden on Europe. <http://bbc.in/1N7F6YJ>



9. Bowden, Greg. “Refugee Crisis in Europe Explained Through Infographics Laying Bare Current Situation” The Huffington Post. 27 October 2015.

There were a couple of main takeaways from this visual. First, it became obvious that using some type of chart (bar in this case) would dramatically improve a user’s understanding of the comparison of numbers between countries. While the heat maps were interesting and necessary to educating a user about the region of the world affected, the supporting graphs allowed the user to gain a deeper understanding of the underlying numbers. Second, I really liked the inclusion of the country flags in the visual. Somehow, I felt like this added a bit of a cultural touch, and forged a stronger connection to the countries involved.

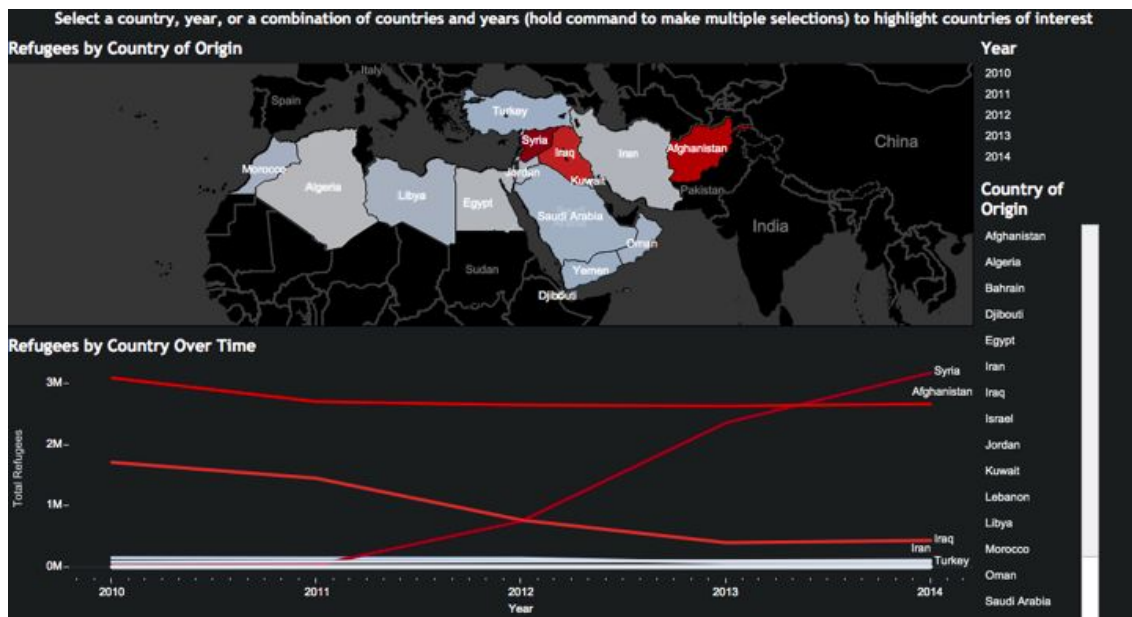
<http://huff.to/1MsOpVf>



Visualizations

We produced four main visualizations, supported by intermittent clarifying text. The goal was to tell the story of the refugee crisis in an order that moved from a big picture look at the entirety of the post-Arab Spring refugee crisis, to more narrowly focused examinations of refugees coming from countries that we identified as being of particular concern. Finally, we would conclude our narrative by visualizing the countries in which these refugees were seeking refuge.

Visualization 1: A Regional Perspective

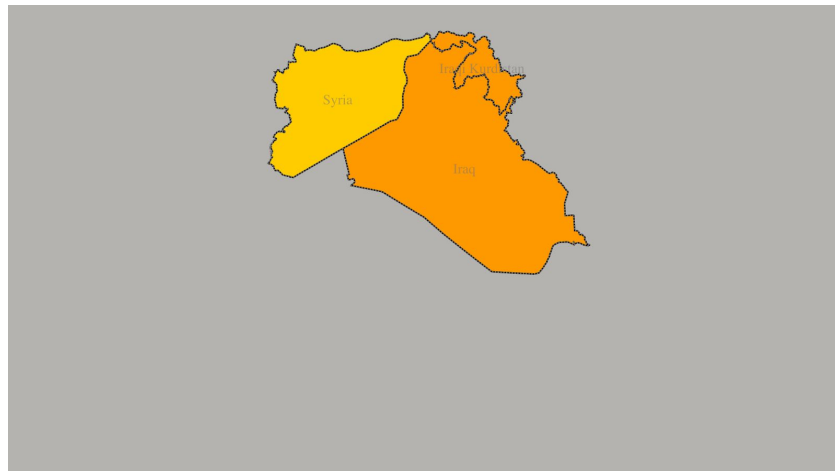


This dashboard serves to present an overview of the countries we examined for our project. In speaking with classmates, we learned that some people were not extremely familiar with the geography of the region we were studying. Thus we decided to include a map in our first Tableau dashboard in order to give a novice geographer an idea of where these countries were physically located. This would also serve to seed our final dashboard, which dealt with the destination of refugees. The color scheme of the map was originally based on the number of refugees leaving each country. However, this proved to be troublesome when comparing the countries to the line graph. Thus, this dashboard was altered to match distinct colors for each country. In order to learn about the number of refugees from each country, you need to explore the line graph. It is clear to see that what is being experienced in three countries was different than the rest of the countries included in our study. As we developed our project, it became

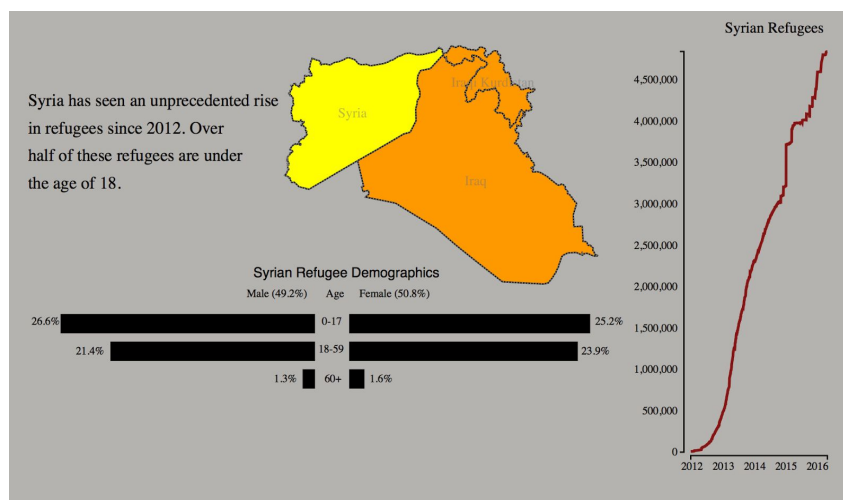
obvious that we wanted to focus most of our attention on Afghanistan, Iraq, and Syria. In order to set the stage for the rest of our project, we emboldened the color choices for these countries so that they would stand out and create a natural transition into the rest of the project

Visualization 2: Exploring Iraq & Syria

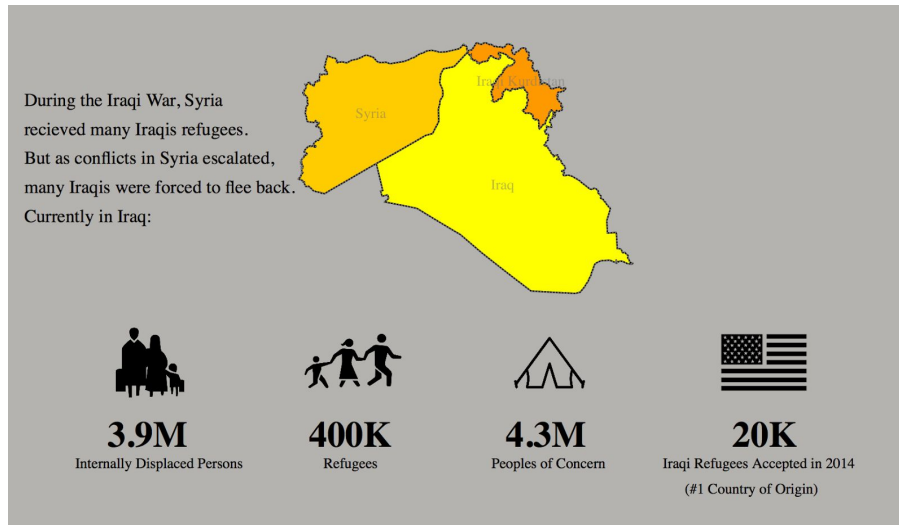
The Iraq and Syria section designed to give a high level view of notable facts and points of interest in the region. This visualization begins by the user scrolling to a map of Syria and Iraq. This map, and the animations on it, was created using d3, TopoJSON and the Natural Earth project. The Natural Earth project provides a public domain dataset of vector data. Converting these shapefiles into TopoJSON allows rendering of Syria and Iraq.



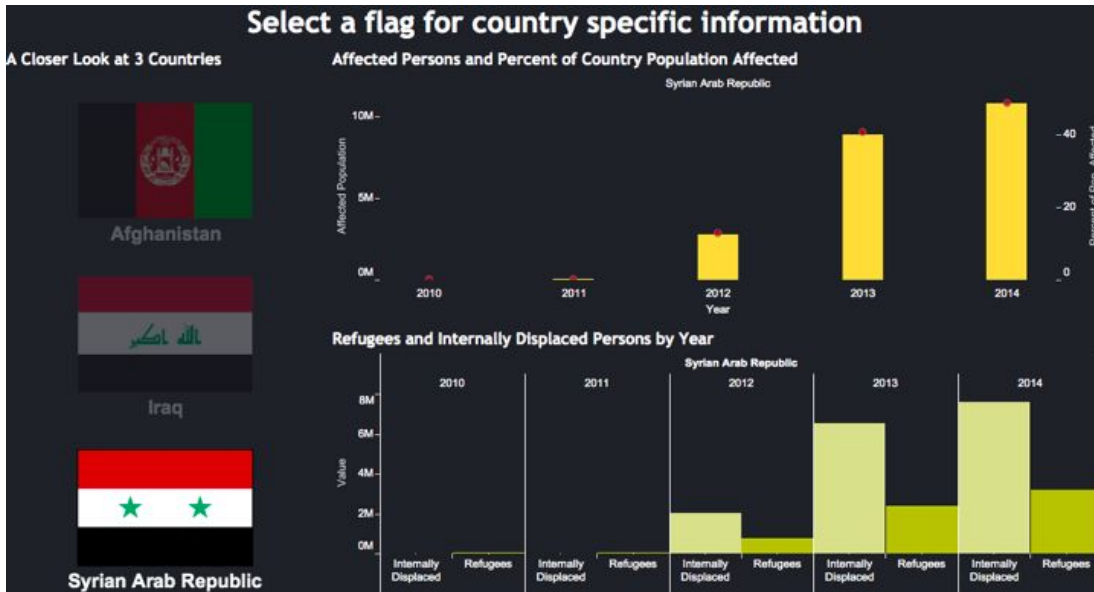
Using D3, mousing over Syria calls two animations to appear. First, a line graph grows with time to visualize the astonishing rise in Syrian refugee counts over the last four years. Secondly, a demographic breakdown by age and gender appears. This and the supporting text highlight how largely children have been affected in Syria. These animations serve to highlight these two major events, as a prelude to the more granular visualization to come after.



Mousing over Iraq calls an animated infographic of major trends in the country.



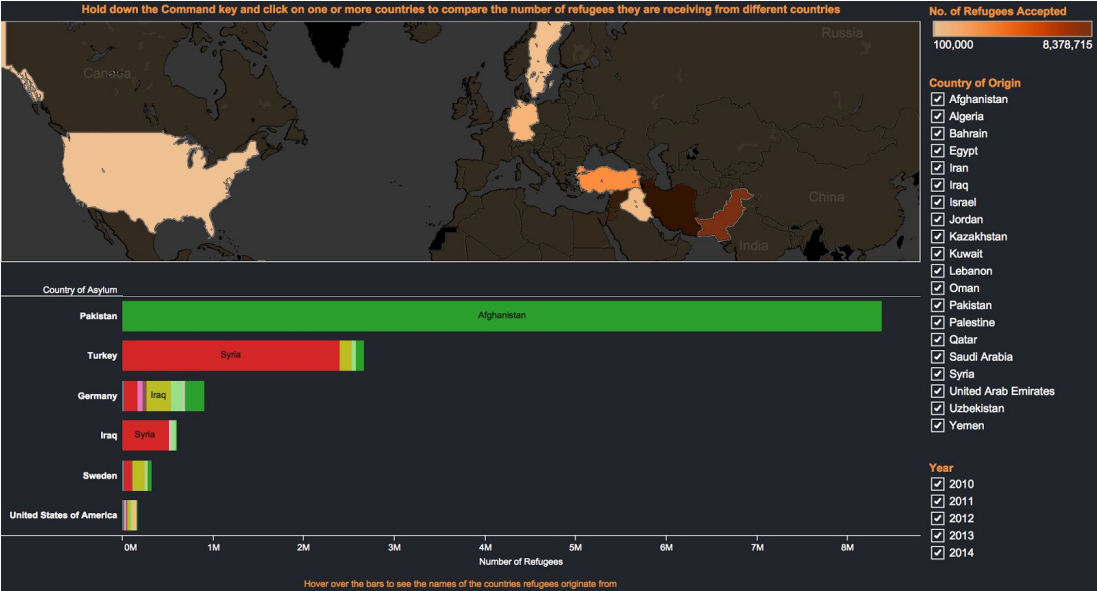
Visualization 3: Drill Down into Afghanistan, Iraq, and Syria



The purpose of this visualization was to drill down deeper into the populations of affected persons in Afghanistan, Iraq, and Syria. In performing our EDA, we found that there were different trends in these main countries in refugee movement. A country like Afghanistan has been steadily exporting refugees, while a smaller number of people remain affected within the country as internally displaced. Afghanistan was included as a reference point that represented

the most consistent major origin of refugees over the last decade. However, the situation in Iraq and Syria has deteriorated rapidly over the last three years. We found that as both the refugee and internally displaced person populations in Syria exploded beginning in 2012, this created huge problems for Iraqi refugees. A once viable destination was no longer viable for Iraqi refugees. Thus you can select Iraq and see their internally displaced persons population surge. The purpose of the dual axis graph is to provide a number for the total number of people within these countries affected on the left axis, and combine that with a look at what percentage of the total population of the country is affected on the right axis. It is sobering to understand that in 2014, basically half of the Syrian population was internally displaced or forced to flee the country. The lower bar graph presented the breakdown between refugees and internally displaced persons, by year. This graph allows a user to gain a better understanding of the situation, and especially an awareness that if you see a country is producing a large number of refugees, it is also likely that there are even more people stuck in that country that have also been displaced.

Visualization 4: Asylum Landscape



We used the same data as therefugeeproject.org but a little differently. Our primary focus was on *enabling comparisons* between different countries. This dashboard also allows for a dual perspective view for both emigration and immigration (though the default is immigration). By

default, the map is shaded to reflect the number of refugees accepted by different countries. Clicking on any country dynamically changes the bar chart below to give a breakdown of where the refugees are coming from. Choosing multiple countries on the map allows allows comparison through the bar chart. Conversely, the user can use the country of origin filter on the right to see where people from one of those countries specifically are going since it changes the shading on the map and also the bar chart below.

Data

The primary source for most our data comes from the United Nations High Commissioner of Refugee (UNHCR), which is a United Nations program mandated to lead and coordinate international action to protect refugees and resolve refugee problems worldwide. The UNHCR Population Statistic Database includes time-series data for asylum seekers, refugees and internally displaced people. Information for asylum seekers destined for Europe was the most granular data set, with weekly breakdowns of number new asylum seekers by country of origin and by country of asylum. Refugee data by country of origin and country of asylum was less granular, and gave only yearly information (up to 2014). Other data sets from the UNHCR that we used was demographic (age and gender) and internally displaced peoples (yearly numbers).

Conflict data was obtained from Global Peace Index from Vision of Humanity. Total country population data was obtained from the World Bank. Lastly, we collected data from the United States Office of Refugee Resettlement, which gave us the United States' refugee arrival figures by country of arrival and state of settlement.

<http://popstats.unhcr.org/en/overview> (UNHCR Overview)

<http://www.visionofhumanity.org/#/page/indexes/global-peace-index> (Global Peace Index)

<http://www.acf.hhs.gov/programs/orr/resource/refugee-arrival-data> (US Refugee Arrival Data)

<http://data.worldbank.org/indicator/SP.POP.TOTL> (World Bank population data)

Tools

D3: D3.js was used in the Syria/Iraq map. The primary purpose of D3 in our case was to get animations to show the extreme growth in refugees and internally displaced peoples. Animation was also used to showcase the large percentage of children that make up the crisis.

Tableau: Tableau was used to create our remaining visualizations. Tableau was also used to merge data sets and create calculated fields from existing data. For example, although our data had a 'Year' column, Tableau did not register it as a date until we created a calculated field for it. Plus, Tableau's country names did not exactly match some from the UNHCR dataset. Discrepancies such as "Syrian Arab Republic" and "Palestinian" which were called "Syria" and "Palestine" in Tableau were resolved within the program. This issue only became apparent and was resolved when the numbers from Syria seemed lower than expected and were not corroborated by the news articles we had used for our research.

Excel: Before importing to Tableau, Excel was used to look at various csvs to see what was in them and make initial judgement on usefulness. Excel was also used to scope the data i.e. remove thousands of rows of data from the years before 2010.

Process

In order to create this visualization, we first had to do research on the topic. We didn't want to find ourselves in a positions where we produced a bunch of cool looking visuals that told a story based on what could have been a flawed data set. So, in order to limit scope, we researched trends in refugee movement over time beginning with the Arab Spring.

After researching the topic area, we obtained a series of reputable data sets. As listed above, the primary data set we used came from the UNHCR, and it was supplemented as necessary by data sets from the Institute for Economics & Peace, the World Bank, and the U.S. Department of Human Health and Services. Because this data did not merge cleanly in Tableau, there was significant work required, and mostly done in Excel, in order to combine the data sets into useful sources.

After a lengthy period of Exploratory Data Analysis, we determined that we wanted to focus on and emphasize trends in Afghanistan, Iraq, and Syria. Thus, we divided up the work of creating visuals based on our own personal interests. We decided to create a number of Tableau dashboards that would serve to both develop our narrative as well as allow user interaction with the data. We also decided to develop a d3 section that would allow us to control the information conveyed in an aesthetically pleasing, interactive animation.

Once our visuals were developed, we mapped out the design of our narrative, determining where text was needed to contextualize a visual. We decided to included text at different points: introducing the refugee crisis, justifying our focused scope on countries of particular interest, providing context on countries of asylum being sought by these refugees, and also text to conclude the visualization.

After coding our visualization, we conducted usability testing and found that there was concern about noisy visuals that seemingly detracted from efforts to produce a cohesive narrative. Other comments, obtained during the project showcase, led us to change different aspects of the presentation in order to more closely conform to established best practices.

We took the results of the usability testing and other comments into consideration, and revamped basically every one of our visuals. We were unable to do a second round of usability testing, but acknowledge that a second round would be very helpful in determining whether or

not we succeeded in addressing user concerns. We also made addition to the supporting text of the visualization in order to further develop the narrative voice of the project.

Usability Testing Results

We conducted usability testing with classmates, and combined the results of that testing with comments from the TA's and professor during the project showcase. The recurring theme from this testing was that there was too much 'noise' in the presentation, and people were unsure of what they were supposed to be taking away from the project. We expand on the results below.

What is the purpose of the Global Peace Index in the first dashboard?

During our project showcase, our first dashboard included a separate table, which had a score from the 'Global Peace Index' for each country, by year. This table was colored to depict the less peaceful countries as a darker shade of red than the more peaceful countries. Based on the scoring scale (1-5 with 5 being least peaceful), basically no user found any value in the table. They actually found the scale to be confusing and distracting. The goal of using the peace index had been to compare countries with a large number of refugees to countries to the peace index in order to determine if there were any interesting correlations. Scatterplots indicated that the correlation effect wasn't very interesting, but the peace index table was originally included nonetheless. However, this table bombed during the presentations and was thus removed.

Why does the line graph in the first dashboard use a different color scheme than the map?

As the map in our original dashboard was shaded red according to total number of refugees originating from those countries, we thought that using the same scheme for our line chart would make it difficult to parse through the different countries presented in the chart. However, the arbitrary color scheme we chose ended up being more confusing. There were questions of why we would even include both of these visuals in the same dashboard. Did it make more sense to split these apart as separate, stand-alone visuals? We decided to match the color scheme in the line graph to a new color scheme developed for the map in our revamped first dashboard, as discussed in the next paragraph.

I don't understand what I'm supposed to be taking away from the first dashboard.

The first dashboard ended up being something of a disaster. It was the first visual we presented, and people were getting hung up on it. They were just confused as to what they were supposed to be taking away from the visual. As we had decided that our goal was to present information that logically led users to the conclusion that Afghanistan, Iraq, and Syria were the most significant countries to study in this project, we decided to change the color scheme of the dashboard to accentuate those countries. Looking at the line graph, as well as data from our EDA, those three countries dominate the refugee story. We changed the color scheme to represent this by selecting a light blue scale for all of the other countries involved in the study. Then, we chose different shades of red to associate with Afghanistan, Iraq, and Syria. The resulting dashboard displays less information (elimination of the peace index) but in a fashion that makes the highly relevant actors in our narrative jump off the page at the user.

What is the purpose of the boxplots in the second Tableau dashboard?

In our second dashboard, we sought to drill down into refugee and internally displaced person numbers in our three chosen countries of interest. A boxplot view was included in this dashboard. The purpose was to highlight the hugely unpredictable and explosive trend in the increase of Syrian affected persons over a small period of time, especially compared to the relative 'stability' of affected persons in Afghanistan. However, it was not intuitive to any user that the different data points associated with the countries represented different years, and not a mix of refugee related variables. Thus, the boxplots were eliminated from the dashboard.

Task Table

	Eric	Reema	Rob
Research	33%	33%	33%
Data Cleaning and Manipulation	40%	20%	40%
EDA	33%	33%	33%
Textual Narrative	40%	20%	40%
Tableau dashboards	60%	40%	-
D3 visualization	-	-	100%
Website design and development	-	100%	-
Final Report	33%	33%	33%