

CS 3300: Project 2 Writeup

Authors: Danyal Motiwalla (djm453), Riley Niu (hn263), Xinyu Zhao (xz293)

Datasets

Our main dataset was a collection of refugee data, obtained from Kaggle (<https://www.kaggle.com/unitednations/refugee-data>). We specifically used the time-series dataset, which included data from 1951-2016, but was more comprehensive from 1995, thus we only used data from 1995-2016. The dataset included the year of study, the country of asylum (known as *destination country* in our dataset), the country the refugees came from (*origin country*), the type of population that moved (we specifically chose *refugees*), and the number of refugees that ended up in the destination country.

After choosing to work with this dataset, we discussed how we could add in more data to make our visualization thoughtful and interesting. We decided to investigate relationships related to population, and found our population dataset from the World Bank (2016) World Development Indicators (<https://data.worldbank.org/indicator/SP.POP.TOTL>). Additionally, we knew we wanted to incorporate a map into our visualization as it could be interesting to see the refugee travel paths plotted on the graph. Additionally, we wanted to see if refugee paths typically ended up in countries with higher populations. To allow this information to come through in a visual manner, we decided to display each of the countries based on their populations, but with the same coordinates as the world map. We chose to use a dataset with the center coordinates of countries of the world, obtained from Socrata.com (<https://opendata.socrata.com/dataset/Country-List-ISO-3166-Codes-Latitude-Longitude/mnkm-8ram>).

Using the merge data feature of Excel, we started combining these datasets into a dataset that could be easily transformed into a JSON file, with information that we were highly likely to manipulate in our visualization. The hardest part of matching datasets was to ensure that the country names were consistent across datasets, otherwise important information may be excluded. Additionally, we had to ensure that missing data would appear as null in the json file. We produced two json files, both including an array of objects, for two different purposes. The first json file was used for plotting the “population” map, with keys including country, latitude, longitude and population. The other data set was used for our path data, with keys including year, destination country, origin country, number of refugees (travelling to destination countries), the destination country’s latitude, longitude and population (2016 numbers), and the origin country’s latitude, longitude and population.

The Map (for a specific year)

The crux of our story is represented on the map visualization, and we can break it down into several parts. Firstly, to clearly show the population of each country effectively, we decided to replace land mass with circles that were square root scaled based on population, while still situated in the same relative center positions as the countries on the world map.

Next, we incorporated paths between origin and destination countries, to show refugee travels patterns. We decided to display 20 paths - to avoid cluttering the map. These paths were the top 20 refugee movements with the highest number of refugees travelling these paths in the current year. Additionally, to show the variation in refugees numbers across the 20 paths, we used color density. The darker colors represented a higher number of refugees while the whiter-toned colors represented the lower end of the spectrum.

To allow the audience to differentiate between origin and destination countries, we added path animation to indicate the direction of the refugee flow, and mapped the flow animation speed to the amount of the total refugee in that transition in addition to the color density attribute of the path. We believe this mapping provides a more vivid visualization for the differences between mapped paths. Additionally, the countries featured in the top 20 refugee movement paths were highlighted green, to stand out from the mass of blue circles representing the rest of the countries on the world map.

To further enhance the user experience, especially since many users may not be familiar with the exact centers of each country - we devised a mouseover tool and display panel. As a mouse hovers a circle, the country's name would immediately appear on a panel on the same screen. When the mouse hovers over one of the countries (represented by the top 20 paths), all paths and all countries connected to that country are highlighted through a change in opacity. Additionally, the panel on the left displays the origin and destination countries connected to that country in different sections along with the number of refugees who have moved from and to those countries.

Lastly, a filter was added as a way of seeing how refugee travel patterns have changed over time. We introduced a slider between 1995-2016, and as it shifts, we see paths change on the map.

The Story

In an effort to find out how have global refugee movement developed and changed throughout the history since the end of the Cold War, we collected, filtered and mapped a global refugee dataset from 1995 to 2016 to produce this interactive visualization of top 20 movement path with highest number of refugees over this historical period. From the visualization, we notice some very interesting trends - there has always been large scale movement around the west African

countries, Germany used to have many refugees moving to the US before the 2000s, and there is a lot more movement in the Middle East as of late. We also notice that throughout time, the highest movement of refugees has always taken place in the Middle East, from Afghanistan to Pakistan or Iran and also from Syria outward.

Resources Used in the Code

1. jQuery: <https://d3js.org/d3.v4.min.js> is extensively used for dynamic visualizations in the code
2. We referenced the index.css file in <http://bl.ocks.org/nitaku/6354551> to create the flow animation of refugee movement paths.
3. Fonts: https://www.whatfontis.com/CT_Druk-MediumItalic-Reduced.font