

EINDHOVEN UNIVERSITY OF TECHNOLOGY

2IMV20

VISUALISATION

Visualising Immigration Patterns in the United Kingdom

Author
Stanley CLARK

January 28, 2018

Contents

1	Introduction	1
2	Dataset Selection	1
3	Initial Data Extraction and Problem Formulation	1
4	Visualisation Technique Choice	2
4.1	Bubble Chart	3
4.2	Bar Chart	4
4.3	Stacked Bar Chart	4
5	Application Construction	5
6	Data Analysis	6
7	Results	6
8	Extensions	6
9	Conclusions	7

1 Introduction

Recent elections in the western world have seen what can only be described as an anti-immigrant sentiment that is protruding its teeth at all who oppose it. Indeed, also in the United Kingdom one can see the rise of parties that advertise a curbing of the migrant flow and rethinking our policies on asylum as well as visa restrictions. A disadvantage of relying on news media and political figures for accurate information is the lack of readily available sources on the matter at one's disposal. Think of when was the last time that you saw a newspaper article stating a claim, and then failing to provide any links to where those particular numbers came from. Thus is the motivation for this paper. In it, we will attempt to show how one can use publicly available data on immigration figures to visualise information in a user friendly tool, that allows for exploration and the gathering of your own insights with regards to the immigration debate.

2 Dataset Selection

For this project, we choose to use publicly available datasets from the United Kingdom government regarding immigration. Since this report will be about the implementation of a visualisation tool, we will be focussing on one core dataset, from which we can build a visualisation tool, rather than aggregating many different datasets for deeper data analysis insights. Having said that, there were two datasets that were explored until the conclusion was obtained that the first dataset, simply did not have all the statistics needed in order to provide for a good enough visualisation. The first dataset found that was going to be used, [1], had breakdowns per quarter of total admissions to the United Kingdom in any particular quarter. It then proceeded to breakdown by country and by reason for entry. However, in the context of the European Union, no data was available, seeing as how citizens of the European Union do not, yet, have to give any reasons for coming to the United Kingdom.

This eventually led to me being fed up and looking for a different data source. As such the dataset [2], featuring immigration estimates, was found. Although this dataset does not allow for a breakdown per country to be seen, we do have the privilege of having access to more global regions, such as Asia, as well as having access to both inflow and outflow data. This is an interesting thing to visualise since it would be interesting to see how this changes over time.

3 Initial Data Extraction and Problem Formulation

As discussed, the data contains inflow and outflow figures for the United Kingdom on a per year basis. Not only this, but these figures are split into groups of countries, and this is further split into the reasons behind entering or exiting the United Kingdom. From

this we can start to formulate some questions that may be useful to ask when building the application.

- From what region of the world do migrants most heavily contribute to the United Kingdom work economy?
- What has the United Kingdom European Union referendum result done to migration from and to the European Union?
- Does the United Kingdom really have more migrants from the European Union than other regions of the world?
- How many people leave the United Kingdom each year, compared to how many people enter?
- Is there a return of British Citizens living abroad to the United Kingdom, or are they emigrating?
- In what year was immigration to the United Kingdom the highest it has been?
- Which regions have seen the highest growth of immigrants to the United Kingdom?
- From which regions to students come to the United Kingdom to study the most?
- Where are people going when they leave the United Kingdom?

From these questions, and the type of dataset that we are using, the tasks that a data analyst may want to accomplish using such a visualisation tool are numerous. For one, they will want to have a tool to quickly scrub through the different years available to see shifts in immigration dynamics right before their eyes. This could be executed in the form of a slider widget. Then, the data analyst will want to have an overview of particular regions where they can see the immigration figures for each, as well as analysing the purposes of entry/exit for particular regions. They will also want to monitor the inflow vs. outflow of people from the United Kingdom so some sort of switch for this information must be provided.

All in all, the goal of this application is to allow gathering of insights from a high level view of the immigration debate. If a user was to use this application, they could then pinpoint particular areas of interest, after which, using other tools, statistical methods could be used to analyse the data further. This is a valid visualisation use case and provides the greatest number of possible cool techniques that we can use to visualise the data on a high level.

4 Visualisation Technique Choice

For this report we choose to focus on 3 distinct visualisation techniques to be able to summarise and analyse the dataset that was chosen. It is rather a shame that there are

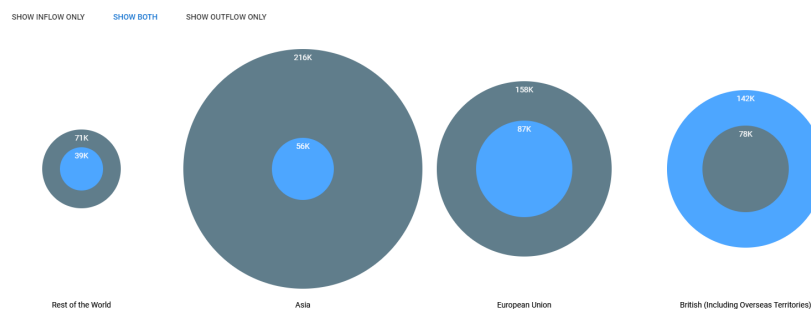
many instances where immigration estimates for a particular category are not available. The visualisation choices then all are displayed with a disclaimer when the data is not accurate, to indicate that such data may not be entirely reliable.

4.1 Bubble Chart

Because the key to this application is interactivity, a lot of analysis only makes sense when viewed over long periods of time. As such the bubble chart was chosen since, when scrubbing the years using the provided slider, a visual comparison of the areas of the bubbles will lead to far better assessments in what changed, compared to a bar chart, especially when we only have one continuous attribute for this portion of the data set; that is, the number of migrants flowing in or out of the United Kingdom in a given year.

In order to be able to have a better overview of both inflow and outflow of migrants, we overlayed the bubble chart with the other statistics when the "show both" option was selected. This also gives a good comparison via the size of the difference of the area between both bubbles, as to the total net migration, without calculating this for the user. This is explicitly not done since then a comparison and connection between the two variants can be made directly, as shown in figure 1.

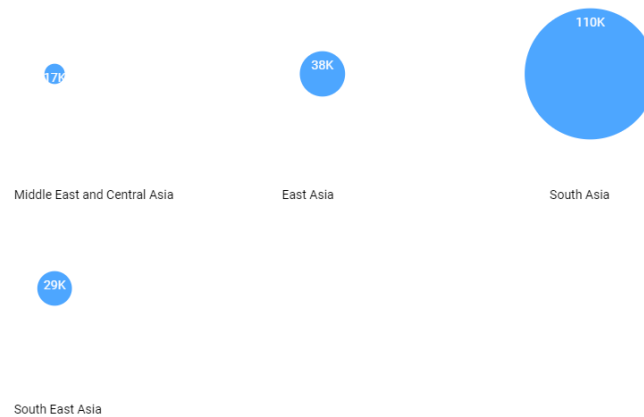
Figure 1: Bubble chart featuring migration inflow and outflow



Such a chart easily works because of the limited number of regions available. This is good since it allows the user to concentrate on the important details and not get inundated with too many meaningless graphics.

By clicking on one of these bubbles, you will be directed to the regional statistics breakdown page where the bubble chart makes a re-appearance in a display of the immigration figures of the sub-regions as shown in figure 2.

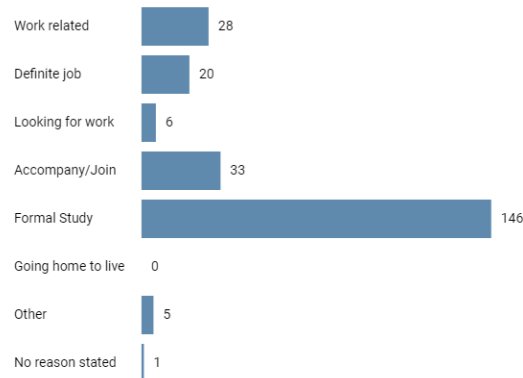
Figure 2: Bubble chart regional breakdown



4.2 Bar Chart

In order to analyse the specific reasons why people moved to and away from the United Kingdom, a simple bar chart could be used. Indeed we have a few categorical attributes such as "work", "study", etc. with assigned values to the whole region. Using this bar chart and a slider over time, we can see specific reasons for migration, as well as how these differ between outflow and inflow. A specific example in the case of Asian migrants can be seen in figure 3.

Figure 3: Bar Chart of Inflow of Migrants from Asia by Reasons in 2008



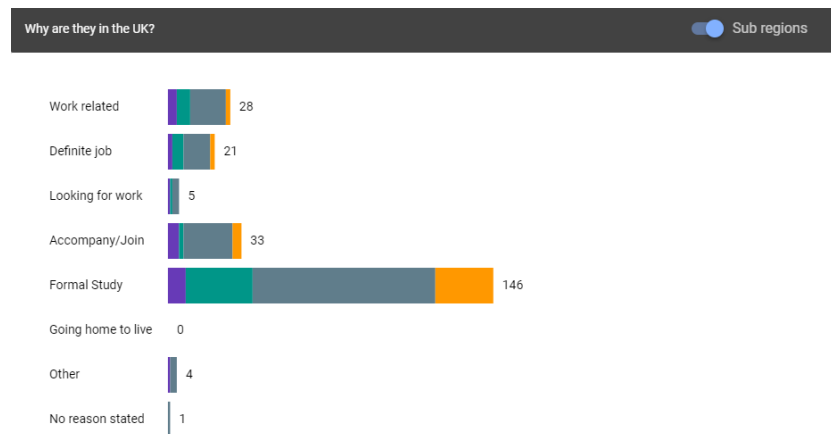
4.3 Stacked Bar Chart

The stacked bar chart may seem like a simple variation on the normal bar chart, but it introduces new insights and complications that would not be present with the old

technique. Indeed, since there is some rounding going on in the estimation figures for each sub region, we get a total of 0 for some categories by summing the values for each sub region, compared to the real estimate for the entire region, which is given as a total of 1K, in one specific case. This case is not isolated and it is interesting to note how the totals of the bar chart change because of the very dirty dataset.

Each colour of the stacked bar chart represents a different sub region and one can toggle whether or not to use this view by using a toggle provided in the interface.

Figure 4: Stacked Bar Chart of Inflow of Migrants from Asia by Reasons in 2008



5 Application Construction

The choice was made to build the application as a Web Application, due to my past experience as a front-end web developer. Since this is an application that will feature a lot of linked components, changing based on user selection, the natural choice was the Vue JS [3] framework, which does a lot of things for you, including acting as a templating engine. As such the initial choice to use D3.js in combination with Vue did not turn out to be the right course of action. Indeed, since D3.js is nothing more than an SVG templating engine, when not using force directed graph layouts, it was, in this case, sufficient to continue on with Vue.js.

In order to be able to quickly template the application, Vuetify.js [4] was used as a style framework in order to make the application look somewhat professional. As exemplified in the screen-cast, the advantage of using a framework such as Vue, is that transitions between sections are very easy to accomplish in a way that guides the user to where they are going, as well as state management, allowing the user to retrace their steps at any time.

Such state based management can be exemplified in the slider component (figure 5),

that allows the user to scrub between any year they want to quickly get a historical outlook on the data. The library also provides a numerous selection of widgets such as toggles (figure 6), and is completely card based, so a nice layout is obtained following the Google material design guidelines.

Figure 5: Slider Widget used to control years



Figure 6: Toggle Widget to toggle views



The end application comprises of two distinct screens, within which inflows, outflows and breakdowns can be viewed. It is always easy to get from one component to the other and transitions and animations are present in everything that is interactive. Though the process was not too complex, it still entailed rather a lot of work to keep good programming practices while developing the application, to make sure that it would be easy to be extended later on, which is the subject of discussion later on in section 8.

All the visual encodings produced were manually put together using SVG templates and Vue objects. Due to some knowledge of Vue.js in the past, it is not really the case that examples were copied and edited to one's liking, apart from the design framework widgets themselves.

6 Data Analysis

7 Results

8 Extensions

Due to time constraints, there are perhaps some items that were left out of this visualisation application that would otherwise have been nice to introduce. Firstly, the dataset that we obtained, while quite extensive in terms of the tasks that we defined, is a small chunk of the overall immigration story. There are no doubt numerous other factors at play, as can be seen by doing a search for "immigration" on the UK government's open data website. These extra datasets pertain to those people refused entry at the border,

citizenship applications and grants, asylum applications and much more. Given that this application only explores the immigration topic at a very high level of abstraction, it would be interesting in the future to extend the application by introducing further topics per region and per sub-region.

A hassle with this approach is that the data is not standardised between data sources and one would have to manually define a list of regions and aggregate found figures into these buckets. However, with such extensions the user would be able to ask new questions, such as what is the relation between those coming to study and future citizenship grants.

9 Conclusions

References

- [1] Immigration statistics: admissions. <https://data.gov.uk/dataset/immigration-statistics-admissions>. Accessed: 2018-01-28.
- [2] Provisional long term international migration estimates. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration>. Accessed: 2018-01-28.
- [3] vue.js. <https://vuejs.org/>. Accessed: 2018-01-28.
- [4] vuetify.js. <https://vuetifyjs.com/>. Accessed: 2018-01-28.