trumpR: An Analysis of Contributions to Trump's Presidential Campaign with R

Salman Quaiyum

1. Introduction

I was first introduced to Trump through the Apprentice tv show. Not that I was a great fan of the show, but I did manage to pick up bits and pieces of it while channel surfing. I remember Trump mostly for his catchphrase on the show: "You're fired!". So when Trump disclosed his intentions to run for the 2016 United States Presidental Election - I did not take it seriously. But here we are in August 2016 and **Donald John Trump** is definitely running to become **"Leader of the free world"**.

Anyways, I thought it would be interesting to explore queries such as where his contributions are coming from, which group is contributing the most etc. So, I collected information about Contributions to Trump's Presidential Campaign and started crunching the numbers.

Before procceding any further, I should warn you that I am not a political analyst nor I am knowledgeable about the socio-political nuances of the United States of America. The following analysis is simply based on what I found by exploring the dataset.

2. Dataset

The dataset contains contributions beginning from **01 June, 2015** to **31 July, 2016**. After importing the dataset in **R**, I did some initial cleaning - renamed columns, fixed data types, removed records containing contributions of 0 and below etc. I was left with **318,681** (three hundred eighteen thousand six hundred eighty-one) records. It contains informations such as contributors' name, occupation, employer, city, zipcode, state etc. Unfortunately, it does not contain any information about the contributors' gender or age - both of which might have answered some intereseting questions.
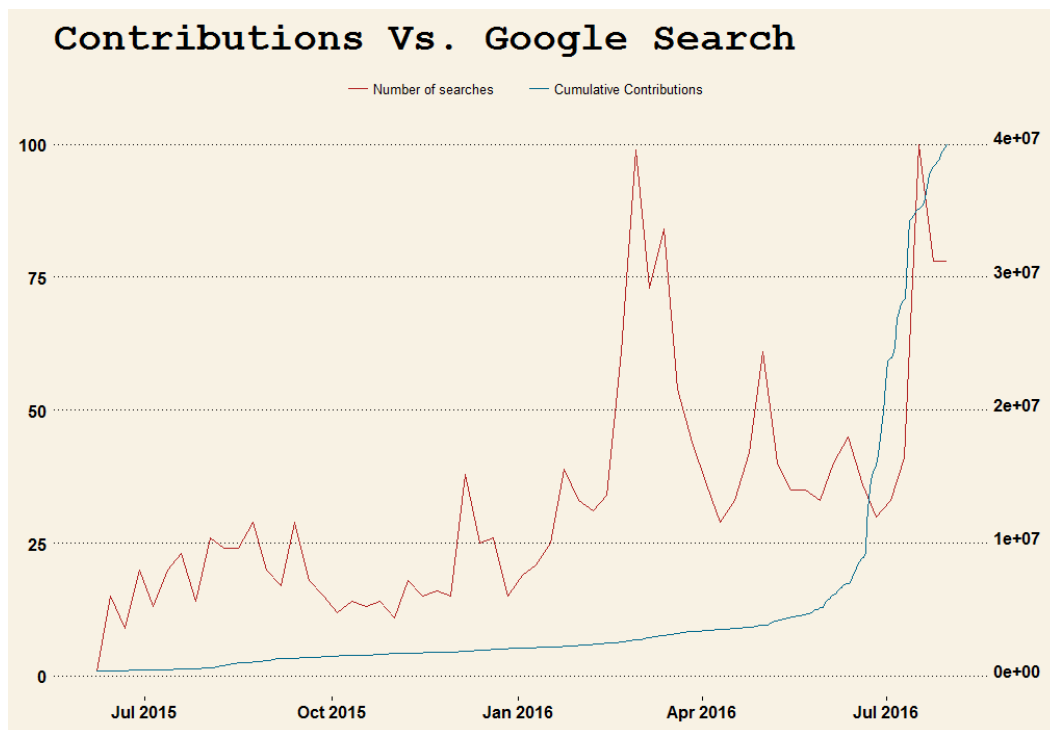
3. Total Contribution

According to the records, Trump has received **Thirty-Nine Million Four Hundred Sixty-Six Thousand Two Hundred Eighty-One** (39,466,281) US dollars as contributions to his campaign.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.8    25.0    40.0   123.8    80.0 86940.0
```

The contributions range from **0.8** to **86,940** USD, with an average of **123.84** USD. Let us take a look at how the amount of contributions changed with time.

Contributions Over Time
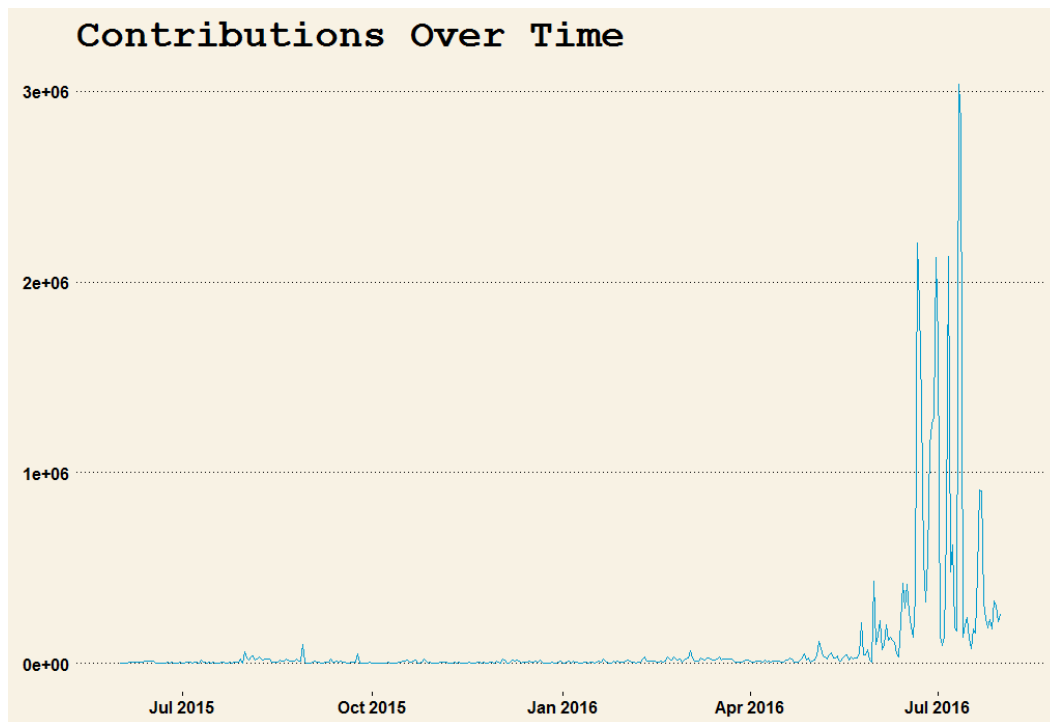
Trump enjoyed a steady stream of contributions from the beginning. But suddenly it was pouring in at the end of July 2016. Comparing the contributions to Trump's popularity in Google searches might reveal the spikes in the graph.

Contributions Vs. Google Search

And indeed it does. We can clearly see that peak number of Google searches coincide with the bumps in the contribution amount.

What events could influence the emergence of public interest in Trump? A peak at the US presidential campaign timeline reveals the following:

1. February 2016 was pretty eventful for Trump. He had debates with both Jeb Bush and Marco Rubio. During this time he also received his first congressional endorsement.

2. 01 March, 2016 was Super Tuesday. Some prominent Republicans such as Mitt Romney, Ted Cruze and Marco Rubio spoke out against Trump during this month. However, he also received endorsement from his formal rival Ben Carson in March.

3. Trump won Indiana Primary over Ted Cruz and inched closer to ensuring his nomination in May 2016.

4. Republican National Convention was held on 18 July, 2016. Trump accepted Presidential Nomination from the Republican party on 21 July 2016. So it is understandable, why suddenly so much money kept pouring in July.

4. Profession of Contributors

4.1 Top 10 Contributing Professions

I grouped the contributors by their profession and found out the top professions with most contributions.
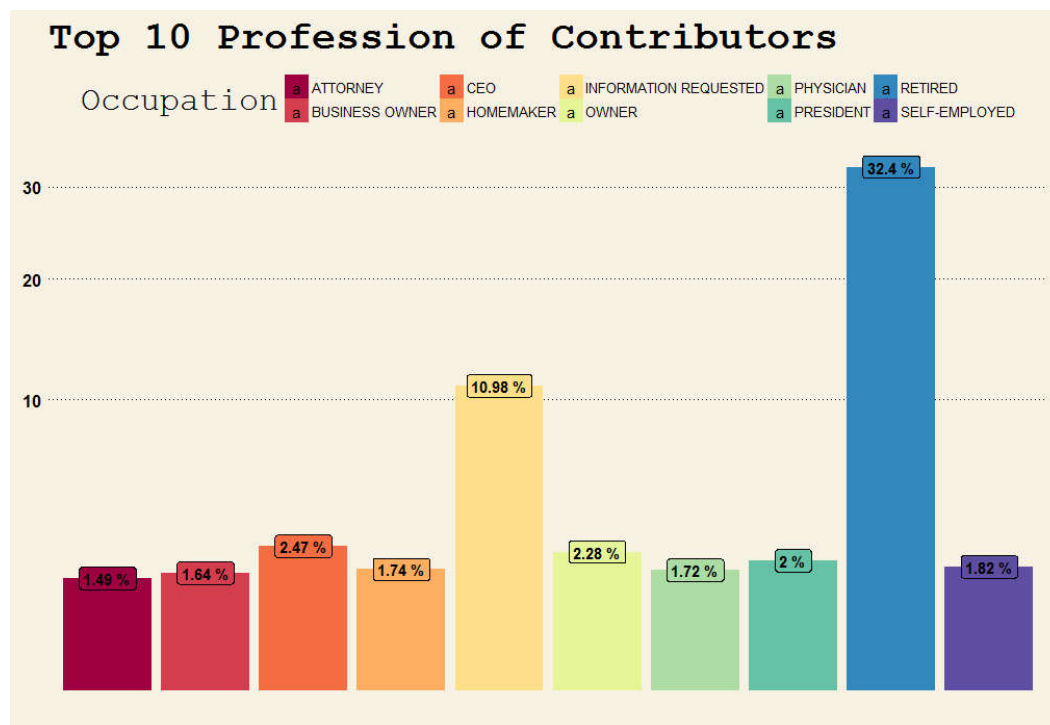

Now, here comes a big surprise. Around **32.4%** of Trump's contributions came from people who are **RETIRED**. They are not only the biggest contributors to the campaign, but also they eclipse contributors from all other professions. The second largeset group seems to be **INFORMATION REQUESTED** with **10.98%** of total contributions. I respect

their decision to keep their occupation private, but from an analyst's perspective it is rather disheartening. The remaining professionals are Chief Executive Officer, Business Owner, President, Attorney, Physician - which are not that surprising. Together they contribute around **15.16%**. But **HOMEMAKER** definitely is a surprise entry in this list. Specially considering that this group has contributed more than Physicians, Business Owners and Attorneys. They must really be ardent supporters of Trump like the retirees.

## 4.2 Distribution of Contributions

Let us look at the summary statistics of amount contributed by each group.
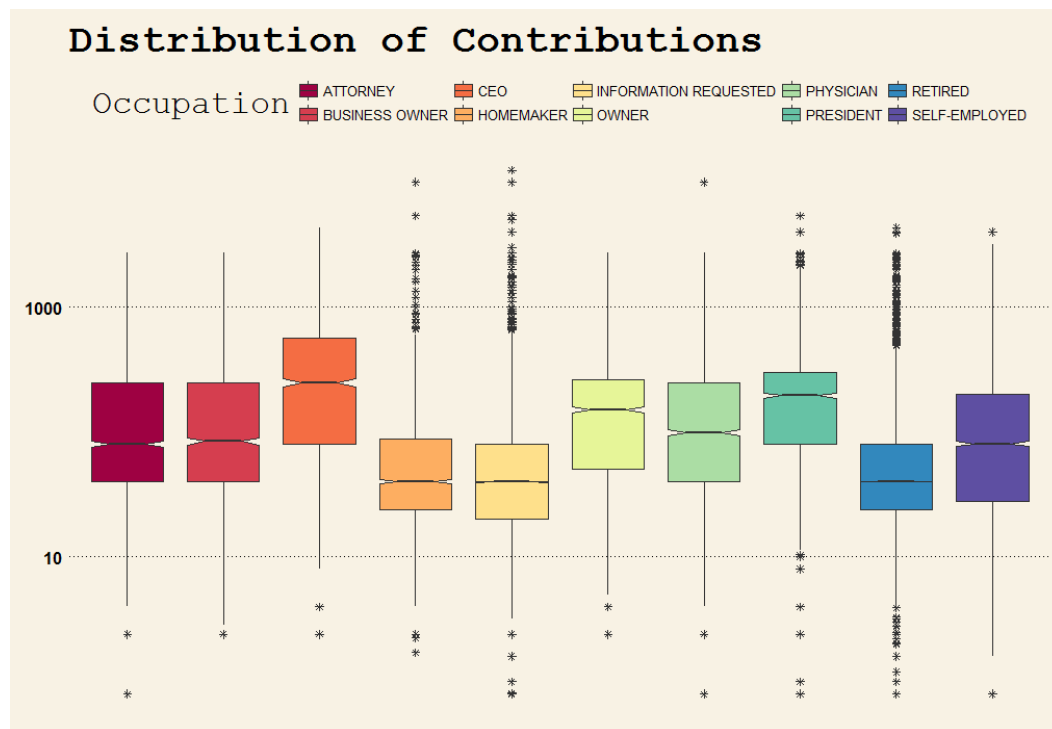
Although there are some outliers present, the distribution of values in each group does not appear to be significantly different. This is also evident by the overlapping 'notches' of the boxplots. Then why is there a huge difference between the leading group (**RETIRED**) and the trailing ones? Specially considering this group has the lowest IQR (Interquartile Range) of **56**. Let us explore their distributions a little more.

A Detailed Distribution
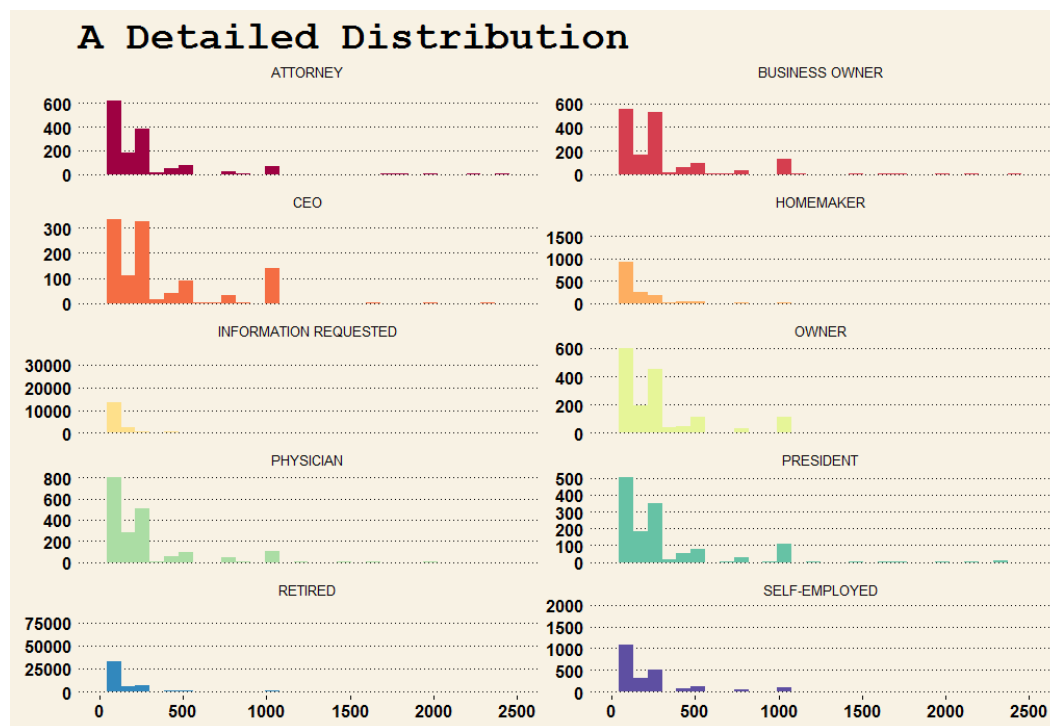
Now we see the underlying reason behind this anomaly. The Retired group wins by sheer number of contributors. Though their contributions range from **0.8** to **4,320** USD with an average of **89.5** USD, there are more contributors in this group than rest of the others combined. Let us sum up our findings with a nice graph.

Summary: Top 10 Professions
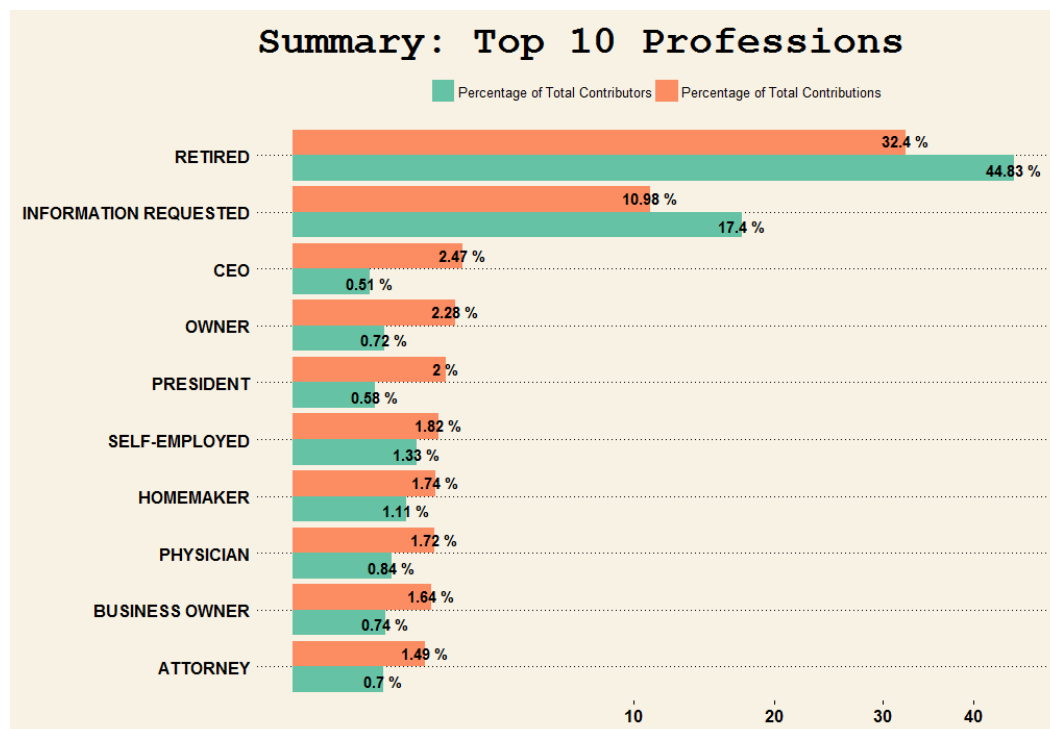
4.3 Top Individual Contributors

Next I thought of looking at individual top contributors. As I respect other people's privacy, I have decided to leave all identifying information out of the analysis and mention the contributors only through their occupation.
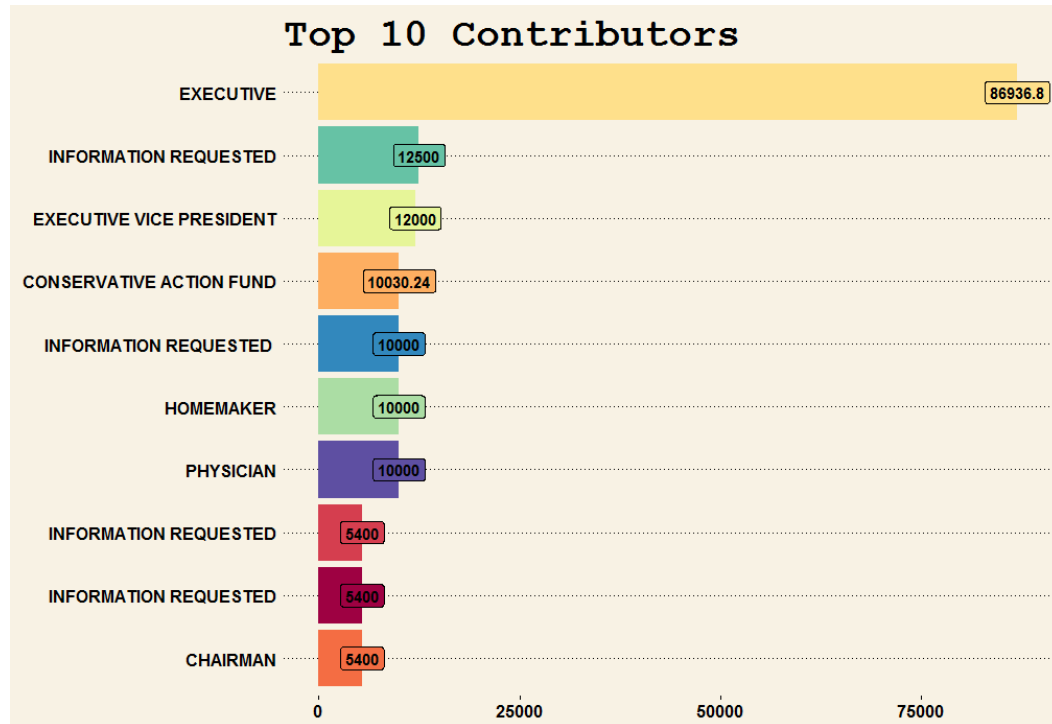
Well, this is no surprise. The top contributor is an **EXECUTIVE** from Evil Corp!

That was a joke.

Moving on, the mysterious profession of **INFORMATION REQUESTED** constitutes **40%** of the list. I wish they were not so secretive with their occupation and be more like **#1 Trump Fan**.

One person lists his occupation as **#1 Trump Fan**. He contributed **514.23 USD** to Trump's campaign; which is well above the mean value. He might not have been the higest contributor, but his zeal more than makes up for it.

Number 4 on the list, **CONSERVATIVE ACTION FUND**, as the name suggests is not an individual.

I am again amazed by the inclusion of a **HOMEMAKER** at number 6.

Anyways, here is another list illustrating the top contributing professions.

5. Contributions by State

We have seen top contributors: both by individual and group. Let us now go few levels up and turn our attention to states.

If you are unfamiliar with tree maps, here is what you need to know:

- The area of the rectangle represents number of contributors in that state.

- The color of the rectangle represents total contribution of that state.

Perhaps, the following map will help you better understand the concept.
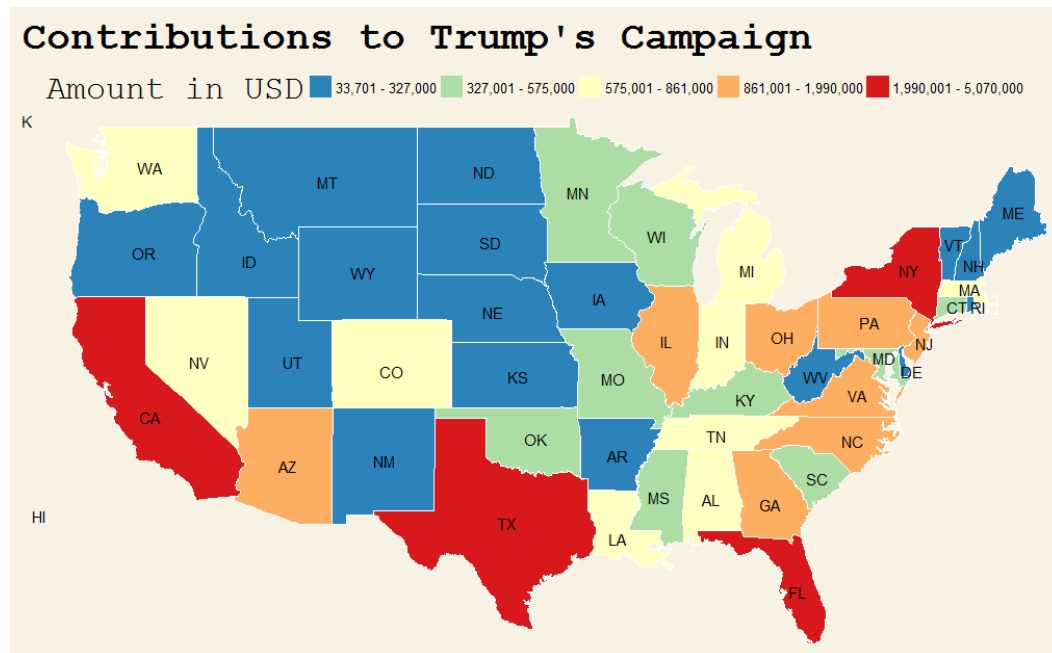
The leading state is Texas with **5,070,163** US dollars in contribution. Closely followed by California and Florida. The other states in the top 10 are: New York, Georgia, Pennsylvania, Arizona, Ohio, Illinois and Virgina.

By combining data from Migration Policy Institute (MPI) and Wikipedia, I have come up with the following chart.

Immigrant Vs. Native Population
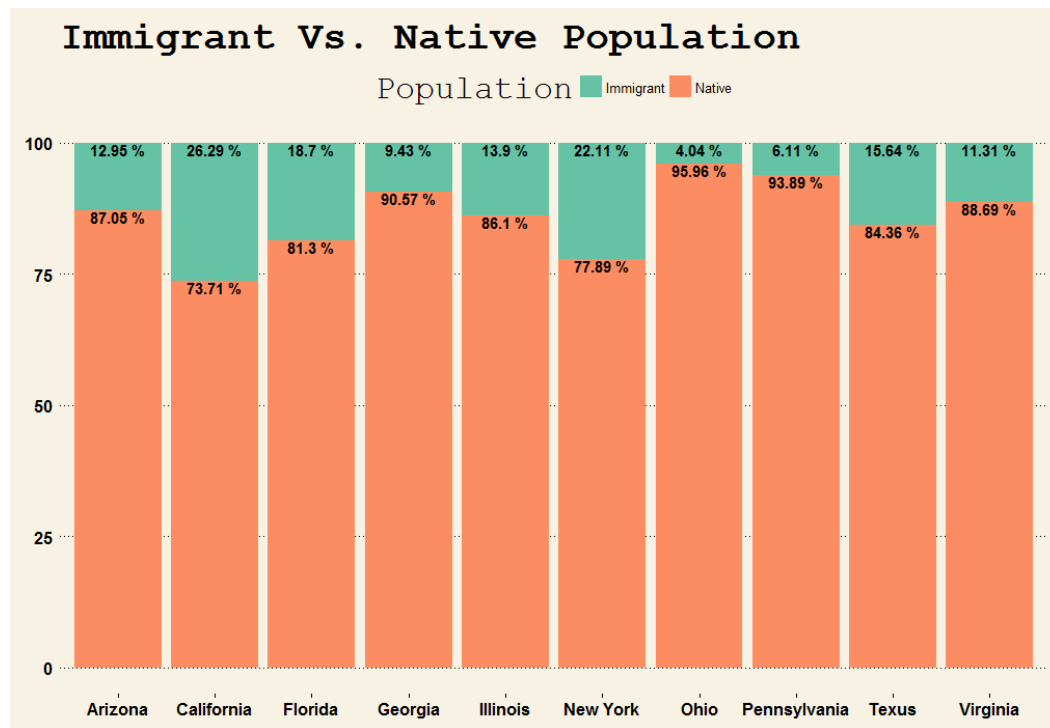
The term native might be a misnomer here. Nevertheless, This chart illustrates the immigrant population of each state with respect to their total population. On a cursory view, the immigrant population does not seem to be a significant proportion of the aggregate. However, according to Pew Research Center: Texas, California, Florida, New York, Arizona, Illinois and Virginia are among the top **15** states in America in terms of immigrant population. This might be one of the reasons for the populace of these states wishing to see an evidently **anti-immigrant** President sitting on the White House.

6. Conclusion

These are all the information I have extracted out of the dataset. I was more interested in telling a visually compelling story more than anything. While doing this analysis, I realized that **domain knowledge** is an invaluable tool to have in an analyst's toolbox. Being a fairly new user of R, I spent hours on the internet to find solutions for the issues I was facing. But my major handicap, undoubtedly, was lack of knowledge about the US election system.

I would be interested to know what else I could have done and whether I have added something unnecessary to the analysis.

https://about.me/salmanq