THE DEPORTATION INFORMATICS PROJECT

Kendal Sandridge
Columbia University
kendal.sandridge@columbia.edu

May 14, 2019

1 Introduction

The Deportation Informatics Project is an effort to increase transparency and accountability in deportation statistics. In the current cultural climate, there is much misinformation around deportation and how it affects individuals. The main goal of this project is to provide information about deportation as it pertains to Latin America and how the institution of deportation negatively affects individuals. This project comprises of a timeline of deportation law, and also data visualizations from data gathered from ICE and Trac Immigration. This is to serve as a model for what a larger project on deportation collecting deportation statistics would look like, and to stress the importance of data-driven approaches to journalism and research.

2 Importance in Data-driven Journalism

One of the goals of data-driven journalism is to use data to guide stories and find interesting events to investigate. Much of journalism on the topic of deportation (particularly with right-leaning groups) is riddled with misinformation and fear-mongering practice. In order to combat the issue, this project calls for a data-driven journalism practice. Questions like, "How many people were deported within this time frame?", "What was the demographic of people affected by this shift?", and "What areas were most affected by this policy shift?" can be great questions that data-driven journalism can address. Note that I would like to distinguish a data-driven approach from simple, fact and figure lookup. Many basic numbers on deportation can be looked up through a Google search, but what data-driven approach look for are less apparent questions of how these figure change over time, and what groups are most affected.

Data-driven approaches can even scale to even larger use cases. Researchers who want to analyze oral histories can use spoken language processing instead of listening and transcribing each recording. The power of computers to find anomalies in data, interesting points of contention, has greatly increased over the years, so it is a vast new frontier of being able to process and review much of the data collected over years and years, and now more readily available as it is digitized. Many questions rely on accurate and consistent collection of data that is usually done by governmental or academic institutions. One big issue is the lack of access to this data by the general public, and the lack of public data sets for analyzing this data. ICE releases yearly deportation statistics within a .pdf table every year. This format might be good for anyone wanting to read the data to get initial scale, but for a researcher wanting to employ more sophisticated data analysis techniques, this poses a greater challenge. Not having accessible data in .csv or .xlsx format hinders researchers from finding trends, patterns, or anomalies in deportation practice. This project is a call to institutions to release open data sets on deportation statistics.

3 What the Data Shows

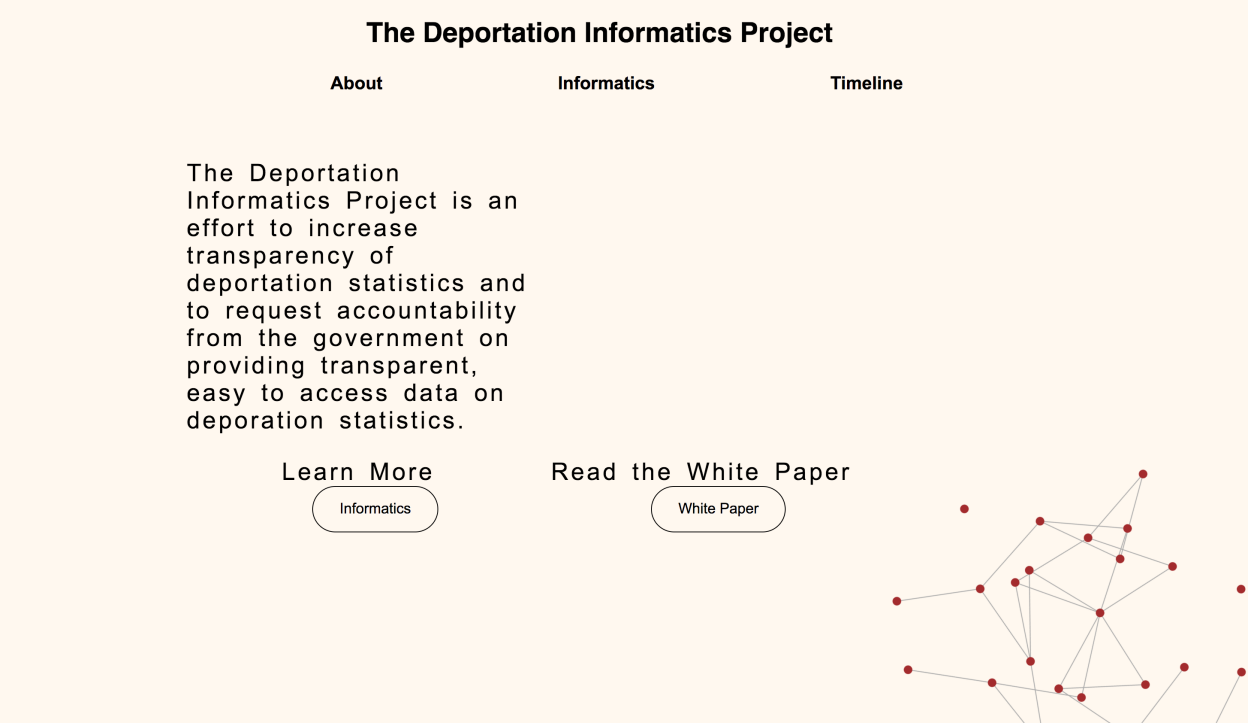
In doing preliminary data visualization, I came across various pot holes. One of the inhibiting things was the lack of publicly available data in .csv or .xlsx format. Much of the data on deportation provided by ICE is in .pdf tabular format. This is not readily available for data processing and visualization. In order to use the provided data, one must meticulously do data entry for each document which can easily scale to hundreds or thousands of data entries. This creates a barrier to what individuals can do with data. For this project I provide three interactive data visualizations which

illustrate (at just the surface-level) what individuals can do to with data on deportation statistics. These visualizations show the where most people are sentenced to deportation, the rates of deportation over time, as reported by ICE, and the age groups most deported. These visualizations were created using data from ICE reports, along with web scraped data from Trac Immigration. Trac Immigration submits FOIA requests to ICE and gets information back which it keeps in a database for academics to use. In order to access their data, one must be affiliated with an academic institution that agrees to partner with Trac. This element of the data access policy makes it difficult to use their data in data analysis. Trac does provideits own in-house tools for visualizing data without explicitly releasing the data. This allows researchers and journalists to quickly fact-check statistics, but does not allow data analysts to readily find patterns, or see anomalies. Trac has data of high dimensionality, including data on the age, gender, nationality, court location, for the people deported or undergoing asylum cases. This data is of high value of researchers wanting to use a numbers-based approach to research questions based on age, gender, or nationality of deportees and asylees. In fact, just by plotting deportation numbers by age group there is much to see.

We can see that the most deported group are people from ages 18-24 and the numbers deported decreases steadily with age, except that children ages 0-17 are deported much, much less. Many narratives on deportation focus on children, their particular vulnerability, and also the lack of culpability for them for crossing the border. Popular works on immigration like Valerie Liuselli's Tell Me How it Ends, or Susan Coutin's, Exiled Home, feature and focus on youth who migrate to the U.S., much of the time alone. Does the focus on the innocence of children, especially by shifting the onus of the migration on adults, make adults seem less vulnerable or innocent in comparison? From here we can as ourselves, have the deportation numbers been a reflection of this narrative, or has this narrative been a reflection of the deportation numbers? This is a question of causality that would require more data to actually research, but at this point, the data can raise significant questions regarding how we treat deportation, and how age can reflect innocence and culpability, when, given the reasons and contexts which drive young adults to migrate, they may not be any less in need as the children that migrate.

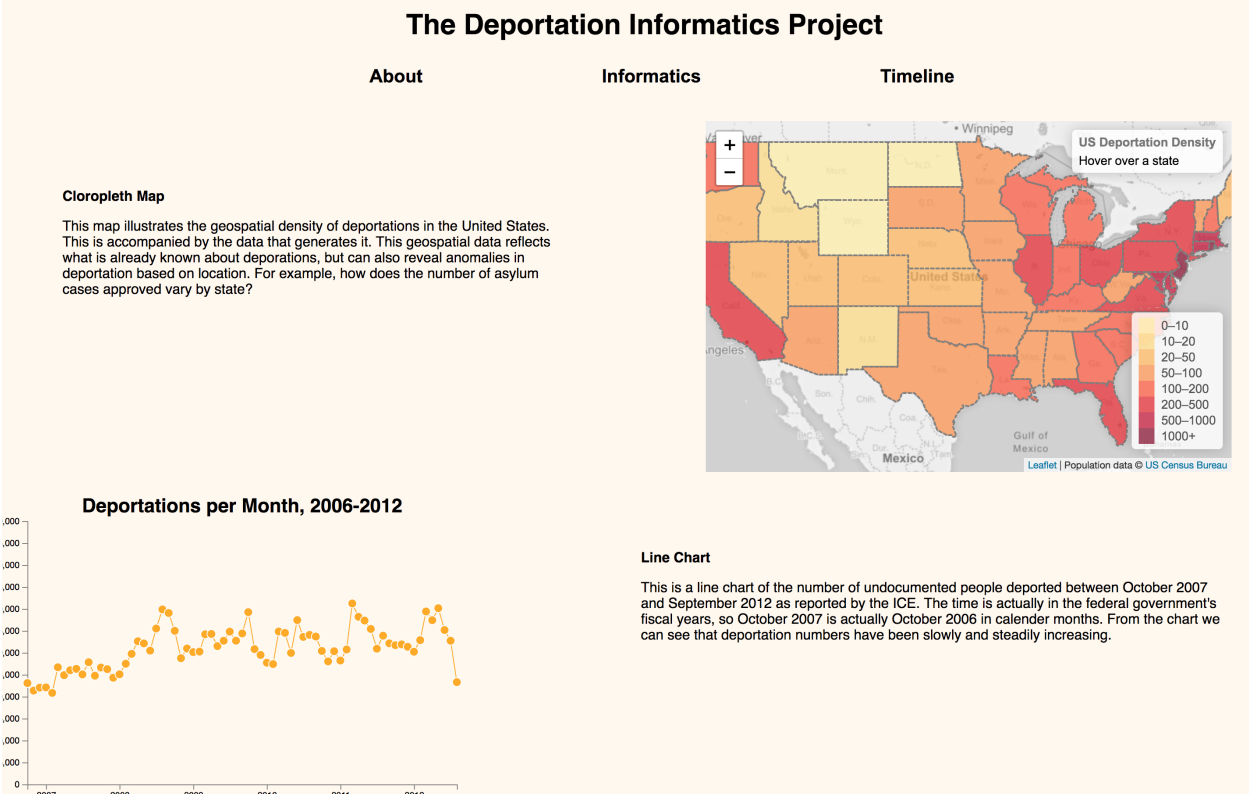
4 Website Design and Features

This website was designed to be engaging and interactive. The content of this website has been tailored for education on deportation legal history in the form of a timeline and analysis of deportation data in the form of interactive data visualizations.



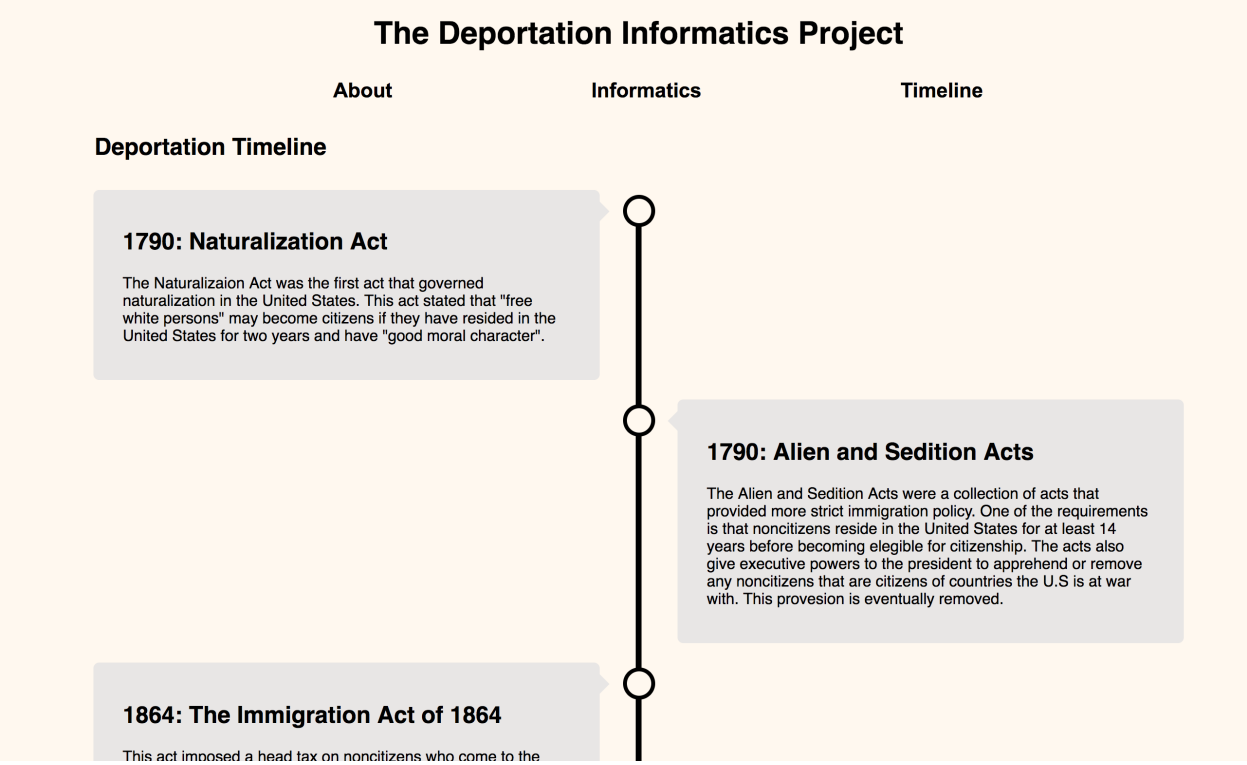
4.1 Data Visualization

Data Visualization offers a way of easily communicating data to people. Geospatial maps, bar graphs, and line charts can all be tools for communicating data that can serve to dispel myths and communicate ideas. One of the goals of the Deportation Informatics Project is to be able to present data in a clear way that shows trends in data and allows users to investigate how different demographics are affected by deportation. This is done by using D3.js a popular software library for data visualization on the web. This works by binding data to html elements and animating them using JavaScript. This project also makes use of Mapbox, a library and tool for geospatial visualization, used to make density maps.



4.2 Timeline

Timelines are a way of visualizing temporal data in a way that can highlight the order and temporality of events. This project features a timeline of immigration law in the U.S. that is meant to highlight the changes in practice in migration practice and impact on deportation.



5 Conclusion

This project is more of a seed, or prototype of what a larger project would undertake. In the future, it would be nice to have automated FOIA requests sent regularly to the government and have that data populated into a database seamlessly for the public to view and use. It is important for the public to be able to access a source of coherent, collected, material to understand issues of deportation better. In the future I would like to see that this project is expanded to include articles on how visas work, how the deportation process typically works, and the current work and research that academics and activists are doing.

References

[1] Coutin, Susan Bibler. 2016. Exiled Home: Salvadoran Transnational Youth in the Aftermath of Violence. Durham London: Duke University Press Books.

[2] Parasie, Sylvain, and Eric Dagiral. 2012. “Dagiral and S’ Programmer-Journalists in Chicago Data-Driven Journalism and the Public Good: ’ Computer-Assisted-Reporters ’.” In .

[3] “Tell Me How It Ends.” n.d. Coffee House Press. Accessed May 14, 2019. https://coffeehousepress.org/products/tell-me-how-it-ends.