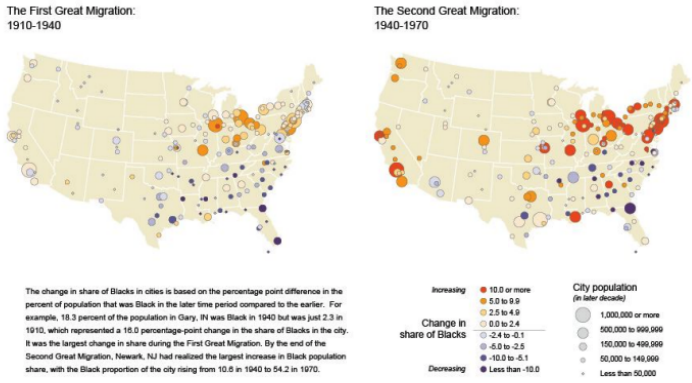
Immigration Tracking Over Time in the United States	
Contact	Names: Arnaud Harmange, Luke Staib Emails: arnaudh@bu.edu, ljstaib@bu.edu Cell Phone: 978-809-1682, 516-408-8668
Organization	N/A
Organization Description	N/A
Project Type	Data Science
Project Description	We would like to create a visualization of the movement of immigrants and non-native United States residents over time. We'd also like to use past datasets to create a prediction model. This prediction model might help show where immigrant populations are likely to move to and concentrate based on historical data. Our intent is that our visualization will be a compendium of existing tools and datasets that track immigration patterns in the US. The prediction model that we create could be used by lawmakers, historians, and analysts as a supplemental tool.
Data Sets & Sources	We can use datasets from the US Census Bureau and from the Department of Homeland Security (DHS) to compile necessary immigration data. For census data (*https://www.census.gov/data/datasets.html*), use the filter "Population Estimates" under the "Population" section on the left of the page. Many datasets over the years are available for the kind of data we want to use. For DHS data (*https://www.dhs.gov/immigration-statistics/yearbook*), choose a year from 1996 to 2020 using the sidebar on the left. Each year contains a set of immigration tables along with a description of each table.
Suggested Steps	1. Utilize the datasets made available by the US Census Bureau and by the Department of Homeland Security. 2. Aggregate all historical datasets and refactor them for use in one or more pandas dataframes.

	3. Use this aggregation to create an animated visualization of historical immigrant population movement within the United States over time. 4. Once completed, the historical data can then be used to train a population movement model, most likely similar to the one utilized in this paper (source below):  Source: http://snap.stanford.edu/class/cs224w-2015/projects_2015/Analyzing_and_Predicting_Internal_Migration_Patterns_in_the_USA.pdf
Questions to be answered in Analysis	How have immigrants moved throughout the United States historically? What are current immigration patterns by group in the United States? How are immigrants likely to continue moving throughout the United States in the coming years?
Ideal Output + Final Deliverable	We hope to deliver: 1. An immigrant group prediction model using census and DHS data that can freely be used for any future project. 2. An animated visualization of past and predicted immigrant population movements in the United states. 3. Our implementation of a population migration model that is reasonably accurate.
Additional Information	N/A

List of Limitations	1. It may be difficult to compile the extensive amount of data from the US Census and the US DHS. 2. We may have to look deeper into whether any data is being shared between the census and DHS in order to avoid duplicate data. This could end up being a difficult and tedious manual step. 3. A great deal of data cleaning will likely need to be done before working with the previously mentioned datasets. 4. DHS data spans from 1996 to 2020 while Census data spans from 1970-2021. Less data in earlier and very recent years (1970s-1980s and 2021-22) could potentially be an issue when creating an accurate model.