

*Visualizing*

# US Green Card Applications

**By Violet Luo, Shreena Bhati, Tarun Mugunthan**

Project Link: <https://violet-luo.github.io/info-247-final-project/>

# Contents

1. Introduction	3
2. Related work	4
2.1 H-1B visualization	4
2.2 US visa applications	4
2.3 Wages of US work visa holders by visa category	5
2.4 Air Pollution Visualization	6
2.5 Kaggle Notebooks (3)	6
a. EDA US Permanent Visa with Feature Analysis	6
b. H1B Visa Analysis	6
c. US Permanent Visa Applications v_1.1	6
2.6 20 years of US visas	7
2.7 US Visa Analysis	8
3. Our Visualization	9
3.1 Total number of applications	9
3.5 Your Occupation Also Plays A Role	12
3.6 Employer Location	14
4. Data and Tools Used	17
4.1 Data	17
4.2 Tools	17
5. Usability Testing	18
5.1 Research Plan	18
5.2 Findings, Problems and Solutions	19
6. Member Contributions	26

# 1. Introduction

We are interested in visualizing patterns in US green card applications in the past years. As international students, this information would be relevant to us as we move towards our own new life in the US soon. Knowing trends across nationalities, visa type held and occupation may even impact our own and other international students' decisions about our career paths and futures more broadly.

We wanted the visualization to be an exploratory tool for anyone curious about knowing more about the patterns in green card approvals and denials based on the backgrounds and careers of applicants.

## **Tasks we wanted to facilitate:**

we wanted to enable viewers to understand and compare the trends related to US green card approvals (and denials) across:

- Geographic Distribution in the US
- Visa category
- Industry, Employer
- Nationality of applicant

We hope that those who wish to find a new life in this country will find our tool helpful in understanding the landscape of residence permissions granted by the US.

## 2. Related work

### 2.1 H-1B visualization

By Aquinov Mathappan

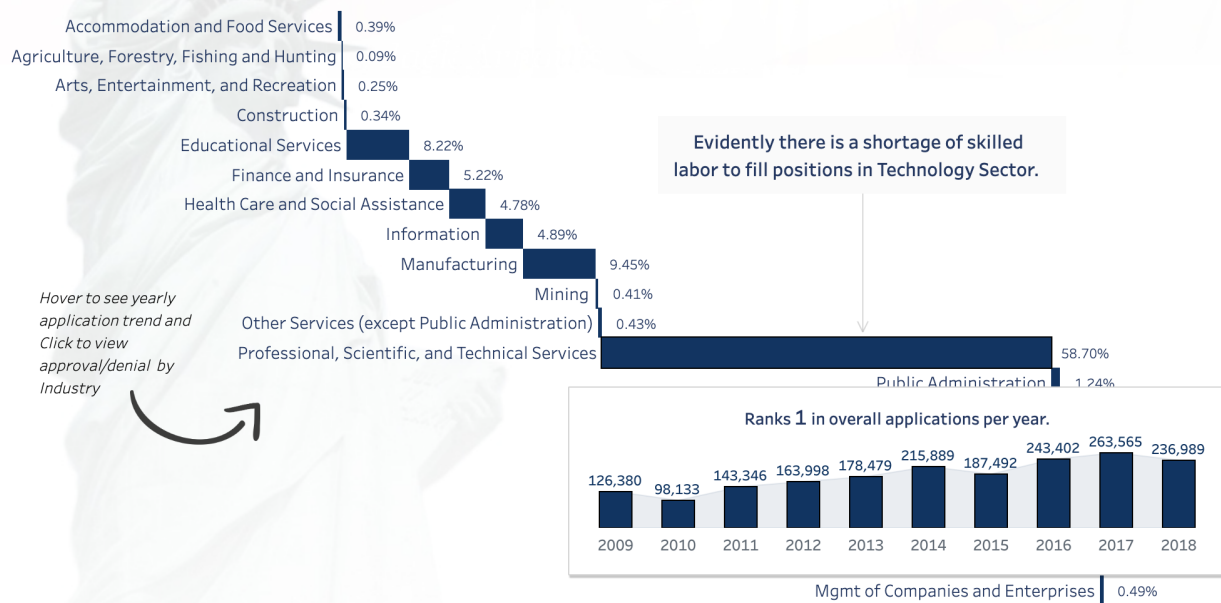
Apart from taking cues about how to present visa data, this visualization inspired us in two ways.

- Having a secondary bar chart in the tooltip on hover to show more data about a specific data point
- Using a US map visualization to show the acceptance rate of green card applications

#### Which Industries file the most number of H1B Visa?

(Fiscal Year 2009 - 2018)

Industries are categorized using NAICS (North American Industry Classification System)

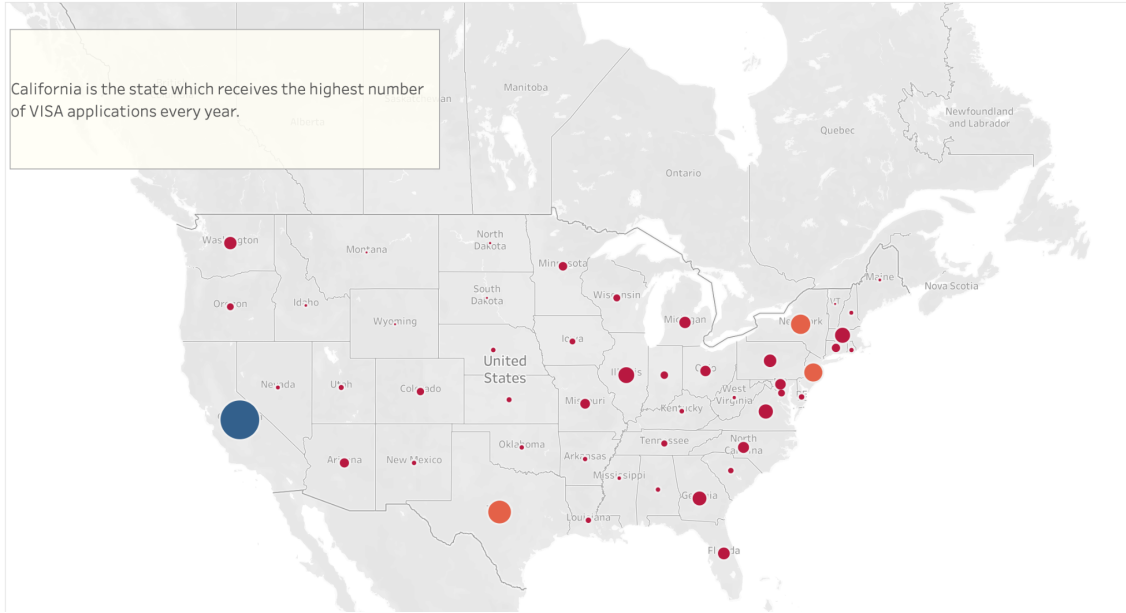


[https://public.tableau.com/shared/TRB8GPKQB?:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/shared/TRB8GPKQB?:display_count=n&:origin=viz_share_link)

### 2.2 US visa applications

By Sujan Shirol

The graph uses a map with marks instead of a choropleth to show quantities. The color encoding is something that we thought would be effective to communicate the approval rate.

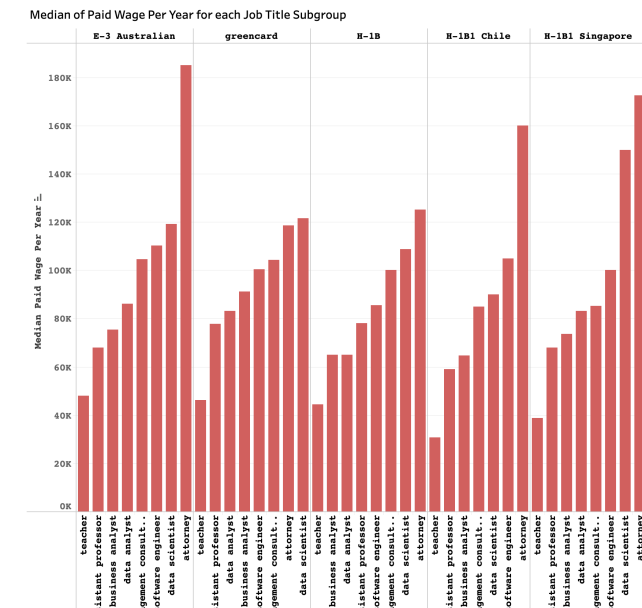


<https://public.tableau.com/app/profile/sujan.shirol/viz/USVISAApplicationsAnalysis/Story1>

## 2.3 Wages of US work visa holders by visa category

By Reshma

The faceting here inspired our own faceting in showing approval rates by industry. Faceting helped make comparisons easier as opposed to filters.



[https://public.tableau.com/app/profile/reshma1559/viz/USVisa\\_16287108328480/MedianVisaa](https://public.tableau.com/app/profile/reshma1559/viz/USVisa_16287108328480/MedianVisaa)

[ndJob](#)

## 2.4 Air Pollution Visualization

By Clara and Astoria

While this is not directly related to our topic, we were inspired by the use of icons when explaining categories of air pollutants. We used passport motifs to explain our visa categories as well, keeping in theme with the rest of our project.

### Carbon Monoxide (CO)

Released from mainly car, truck, and other vehicle emissions.



### Sulfer Dioxide (SO<sub>2</sub>)

Released from mainly burning fossil fuels in industry facilities.



### Nitrogen Dioxide (NO<sub>2</sub>)

Released from vehicle emissions and burning fossil fuels.



### Lead (Pb)

Released into the air from metal processing plants, waste incinerators, and aircrafts using leaded aviation fuel.



### Ground-Level Ozone (O<sub>3</sub>)

Created through reactions of sunlight with nitrogen oxides and VOCs emitted through combustion processes.



### Particulate Matter (PM)

Particles consist of a complex mix of organic and inorganic components. Used as a proxy for air pollution.



<https://astoriah.github.io/info247-final-aap/index.html>

## 2.5 Kaggle Notebooks (3)

### a. EDA US Permanent Visa with Feature Analysis

By Bukun

### b. H1B Visa Analysis

By Diganto

### c. US Permanent Visa Applications v\_1.1

By LUKASZ ZAWIESKA

These visualizations on kaggle helped us decide the structure of the narrative that we wanted

to tell. The notebooks started with general trends and eased the reader into more layers and specifics within themes. There are charts covering different aspects which call back to charts made earlier in the notebook.

We also decided to follow this top down approach when presenting our visualizations. Each layer builds on the previous one. We started with general approval and rejection rates and then added layers of detail by introducing nationality into the mix followed by occupation.

The notebook also does a good job of pairing traditional charts with map visualizations to combine quantitative and location data.

The screenshot shows a Kaggle notebook interface. The main content area displays a horizontal bar chart titled "Employer City and Visas" with the following data:

Employer City	Count Of Visa Applications
NEW YORK	16730
COLLEGE STATION	12118
SANTA CLARA	11113
SAN JOSE	8349
REDMOND	8804
MOUNTAIN VIEW	8259
HOUSTON	8446
SUNNYVALE	8285
SAN FRANCISCO	8254
PLANO	8836
CHICAGO	8658
SEATTLE	8132
EDISON	4222
DALLAS	3760
LOS ANGELES	3704
SAN DIEGO	3683
PHILADELPHIA	3529
CUPERTINO	3379
PALO ALTO	3239
IRVING	3144

Below the chart is a section titled "10 Map of Top 10 Cities with Visa Applications for Case Status as Certified". The text states: "The following map shows the Top 10 Cities with Visa Applications for Case Status as Certified." Below this is a map of the United States with red circles highlighting the top 10 cities: New York, College Station, Santa Clara, San Jose, Redmond, Mountain View, Houston, Sunnyvale, San Francisco, and Plano.

On the right side of the notebook, there is a "Table of Contents" with 17 items, including "Exploratory Data Analysis US...", "1 Introduction", "2 Case Status Values", "3 Top 20 Employees with High...", "4 Top 20 Employees with High...", "5 Top 20 Countries with High Vis...", "6 Top 20 Countries with High Vis...", "7 Top 20 Industries with High Vis...", "8 Top 20 Industries with High Vis...", "9 Top 20 Cities with Visa...", "10 Map of Top 10 Cities with Visa...", "11 Top 20 Employers in New York...", "12 PW Amount in 9089 distributi...", "13 Amazing Top 20 Median...", "14 Amazing Top 20 Median...", "15 Amazing Top 20 Median...", "16 Model Building", and "17 Trend Certified Visas".

- <https://www.kaggle.com/code/ambarish/eda-us-permanent-visas-with-feature-analysis>
- <https://www.kaggle.com/code/dignil/h1b-visa-analysis>
- <https://www.kaggle.com/code/elzawie/us-permanent-visa-applications-v1-1>

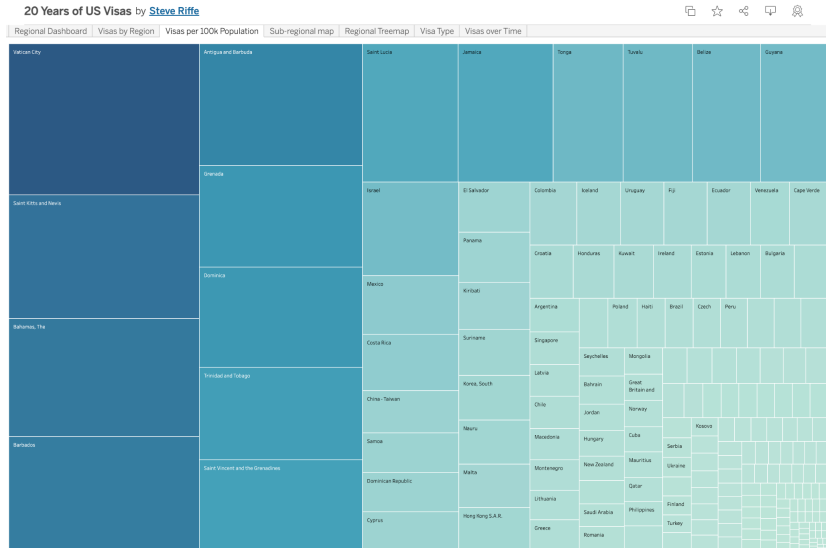
## 2.6 20 years of US visas

By Steve Riffe

This visualization taught us **what not to do**. We had similar problems as those seen in these visualizations where there were a lot of countries listed as the nationalities of the applicants.

The sheer number of countries made a treemap a meaningless chart as the boxes just became extremely small and their areas couldn't be compared to each other. We were then inspired to use a world map instead of a tree map or a list or any form of visualization where having the over 150 discrete elements from each country could be overwhelming.

The world map proved to be the least overwhelming for us because it didn't require users to learn new encodings. Everyone is already familiar with how countries are arranged on a map.

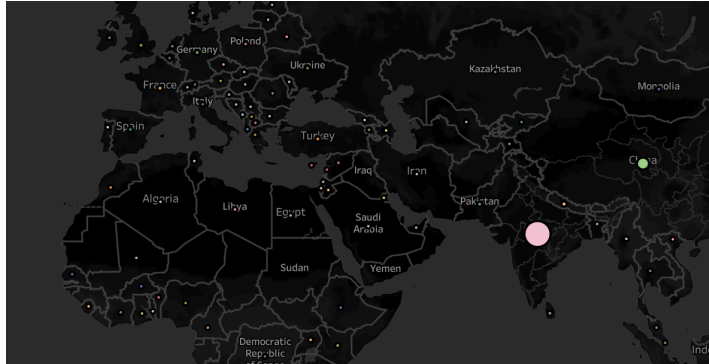


<https://public.tableau.com/app/profile/steveriffe/viz/USVisas1997-2016/RegionalDashboard>

## 2.7 US Visa Analysis

By Parameswaran

This visualization gave us the idea to use a world map to show nationalities instead of a list of filters or facets.



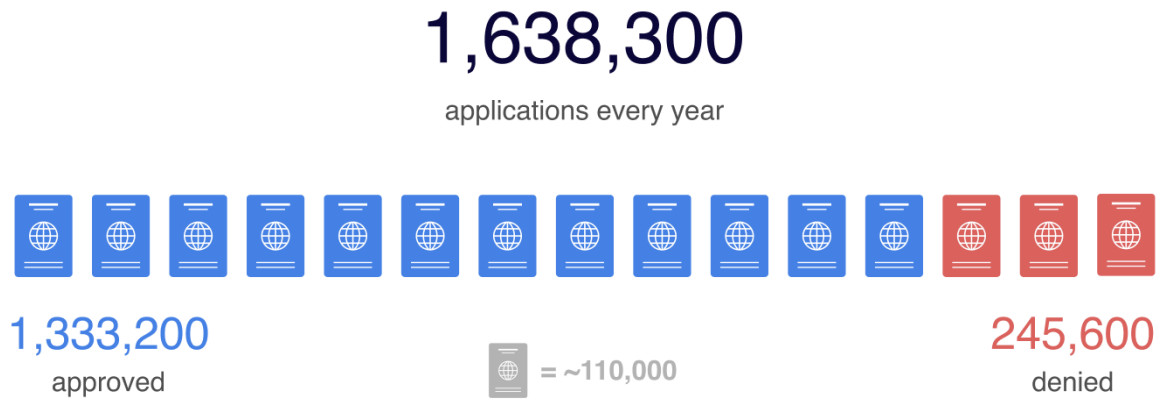
<https://public.tableau.com/app/profile/parameswaran5949/viz/USVisaanalysis/USVisaanalysis>

S



# 3. Our Visualization

## 3.1 Total number of applications



Some applications are withdrawn and not in either category

The very first visual we have on our website sets the broad context of the rest of the project. Throughout this project we're interested in showing how different factors affect people's likelihood of their applications getting approved. Thus approval rates are what we focus on for the most part.

This chart starts off showing the total number of applications every year. It then shows how many are approved and how many are denied. The passport motif is used to stay consistent with the theme of the project.

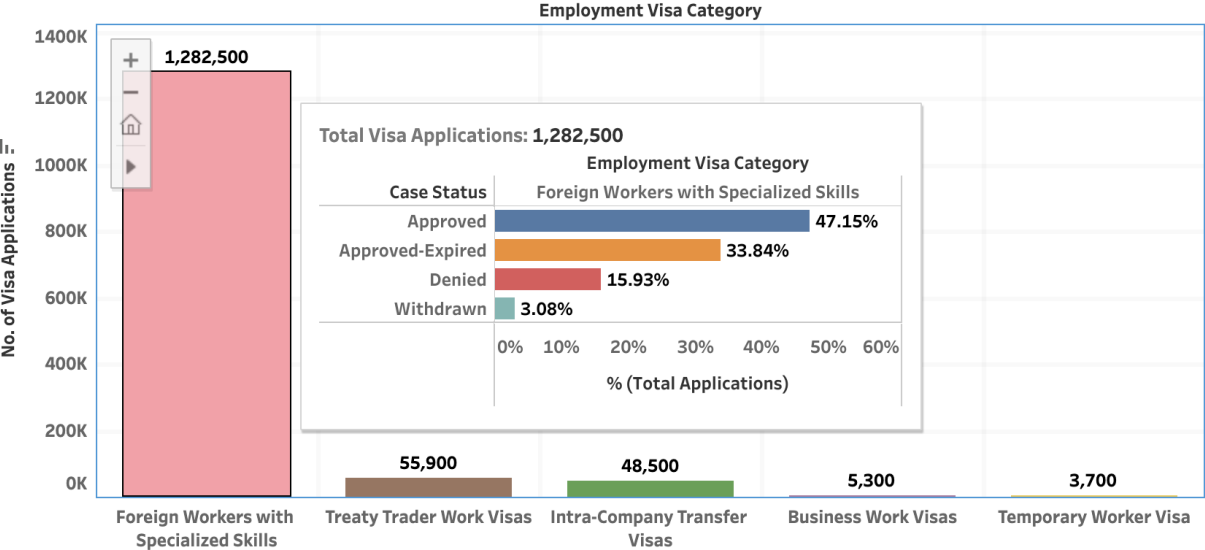
## 3.2 Visa Categories

Applications come from various types of work visas

<p><b>Foreign Workers with Specialized Skills</b></p> <p>H-1B</p> <p>Individuals in specialty occupations that require specialized knowledge or expertise.</p>	<p><b>Treaty Trader Work Visas</b></p> <p>E-1, E-2, E-3</p> <p>Individuals who are citizens of countries that have a qualifying treaty with the U.S.</p>	<p><b>Intra-Company Transfer Visas</b></p> <p>L-1, L-2</p> <p>Individuals who are transferred across offices within multi-national corporations into the US.</p>	<p><b>Business Work Visas</b></p> <p>B-1, B-2</p> <p>Individuals entering the United States for the purpose of engaging in business-related activities.</p>	<p><b>Temporary Worker Visas</b></p> <p>H-2A, H-2B, H-3</p> <p>Temporary workers hired in the US for temporary agricultural and non-agricultural work.</p>
--	--	--	---	--

We then explain the different visa types that we deal with in our project. The cards here have broad categories of visas along with the individual visa types under each category. We reduce the number of categories to five by focusing only on foreign workers trying to get green cards. This leaves out people like spouses of American citizens applying for green cards. We do this for the sake of simplicity and relevance to our target population – immigrant workers trying to find residence in the US.

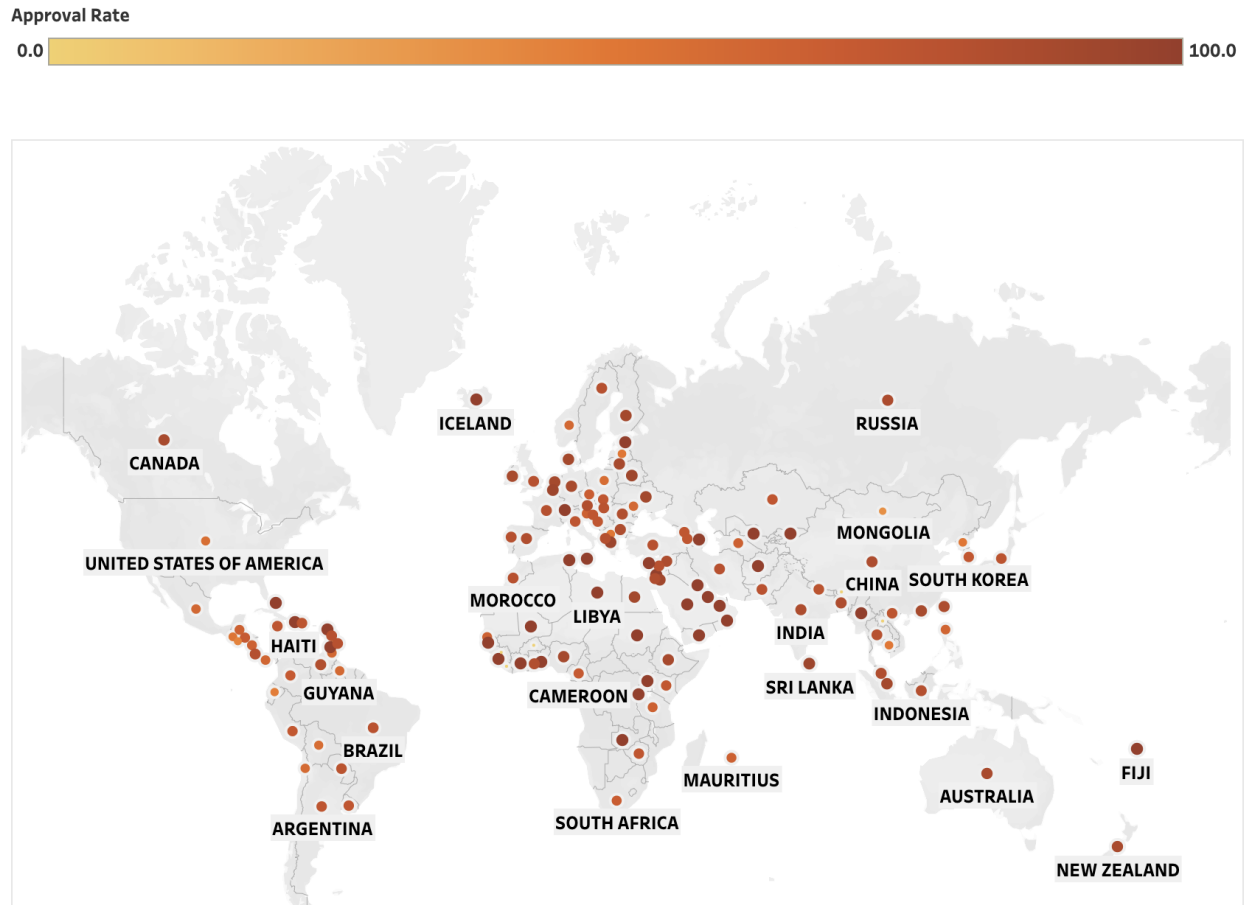
### 3.3 Approval Rate by Visa Category



This chart shows the number of green card applications coming from different work visa holders. When applying for a green card, a foreign worker must disclose their current visa. We explore how the visa they hold affects their likelihood of getting a green card. The graph first shows the overall trend of number of applications but as you hover on the individual bars, it shows a smaller chart which plots the approval rate for each type of visa. Users are able to first understand the number of applications by visa type and then look into the likelihood for each of them.

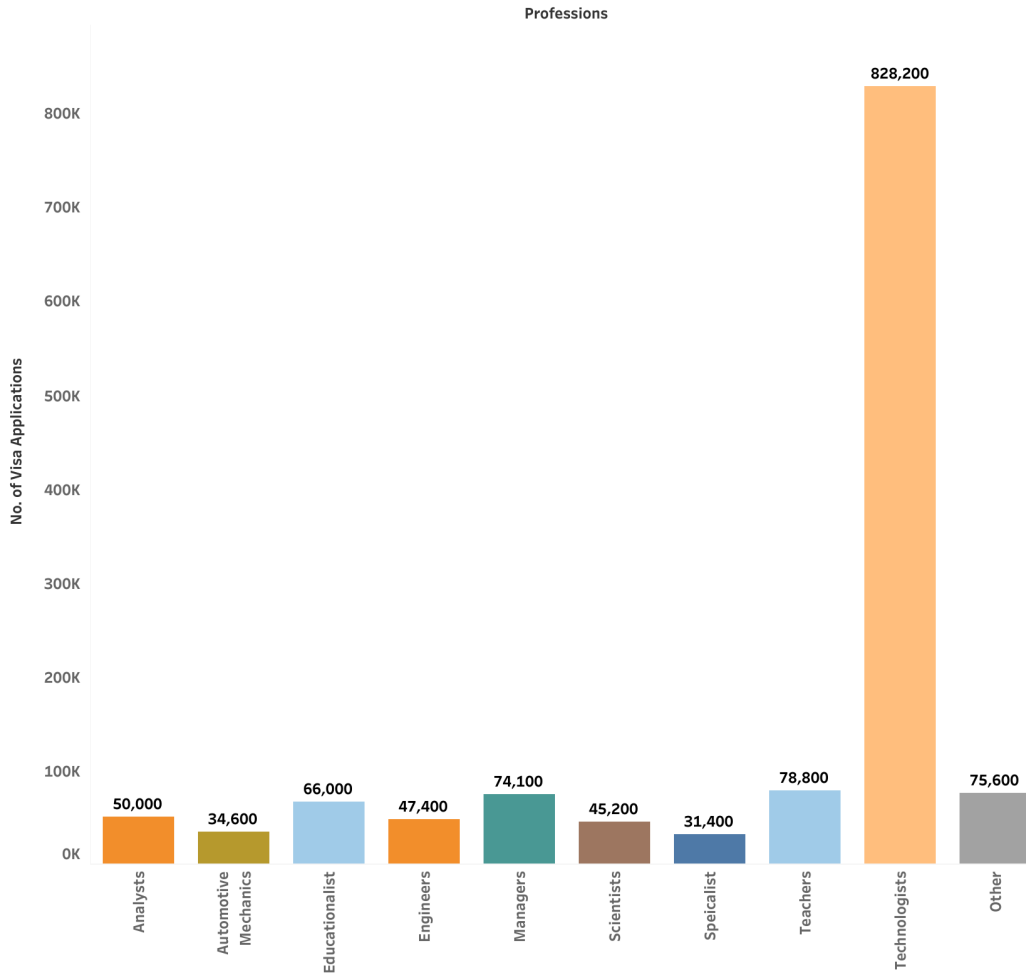
### 3.4 Approval Rate by Nationality

The next chart shows the approval rate as it varies between people from nationalities across the world. The nationality of an applicant is taken into consideration for the decision making process and thus we wanted to explore its effects on the likelihood to get an approval.



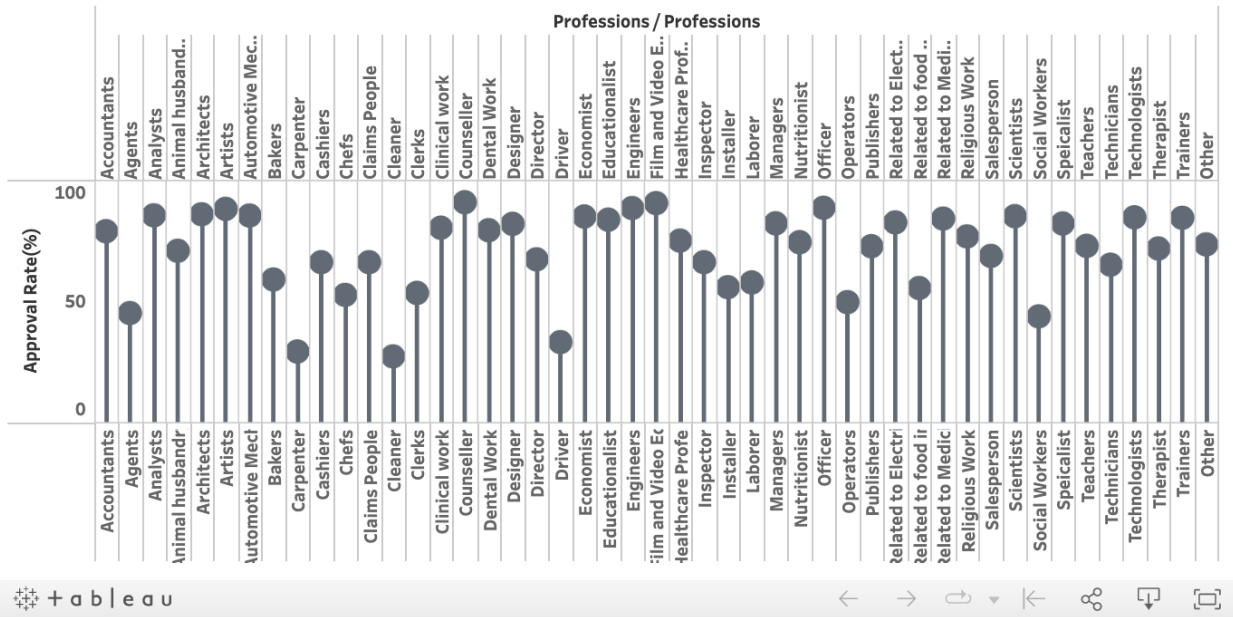
This chart shows that the acceptance rate does not vary drastically across the world. However, some parts of latin america and eastern asia seem to have lower approval rates than the rest of the world. There are several countries in the middle east which have high approval rates ruling out religious discrimination. Europe seems to be the middle of the pack in terms of approval rate.

### 3.5 Your Occupation Also Plays A Role



This chart shows the top 10 professions by number of applications. The top 10 professions by number of applications are analysts, automotive mechanics, educationalist, engineers, managers, scientists, specialis, teachers, technologists, and others.

The number of applications from Technologists significantly outnumbered the rest, and is higher than all other occupations combined.

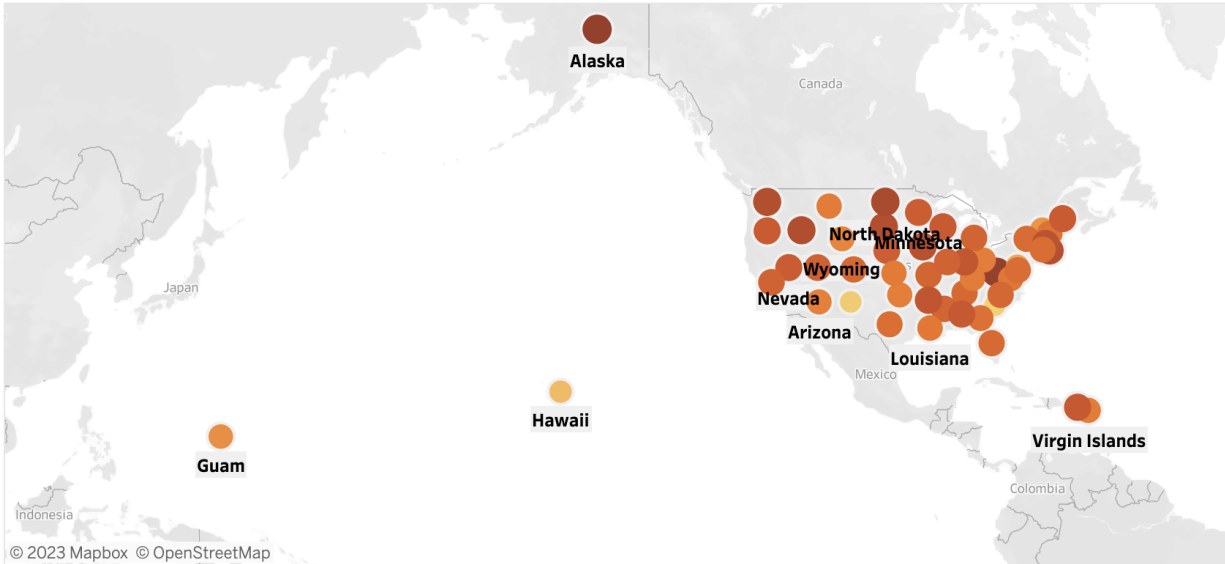


Then we display a world map to explore how approval rate of the same occupation changes across countries. Clicking on the country on the map displays the approval rate for the occupation in that country.

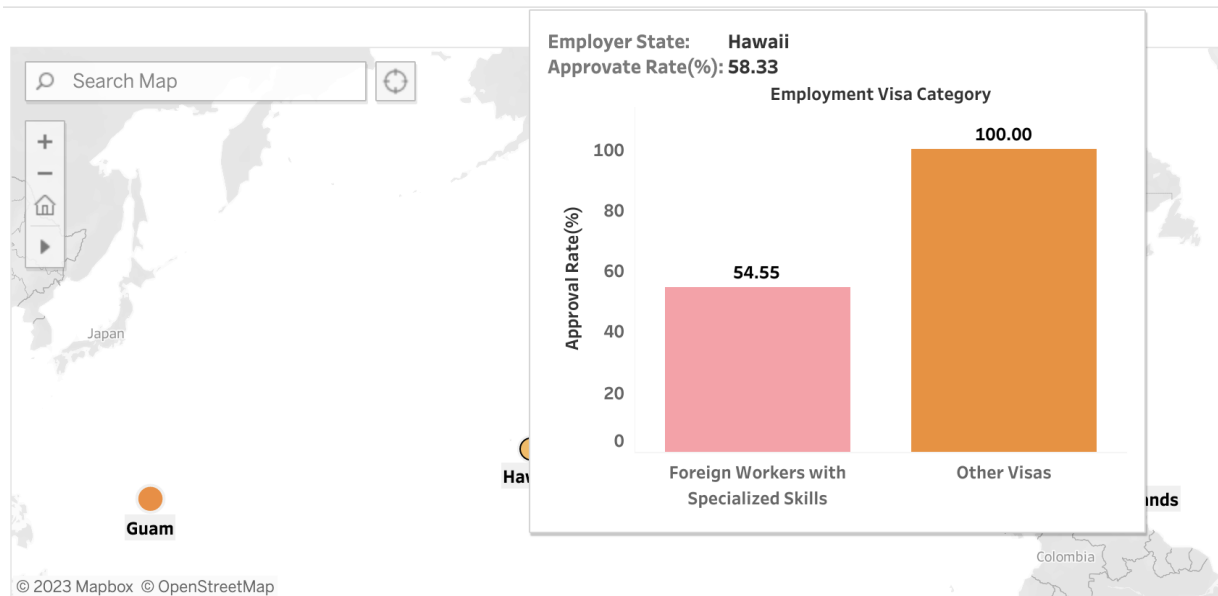
India, China, South Korea, Canada, Mexico, Philippines, UK, Taiwan, Pakistan, and France are 10 countries with the most applications. Out of the top 10 countries, approval rate of the same profession varies. Take technologists as an example. It varies from 93.7% approval rate for Chinese to 78.95% for French. Admittedly, when assessing approval rate, there are many other

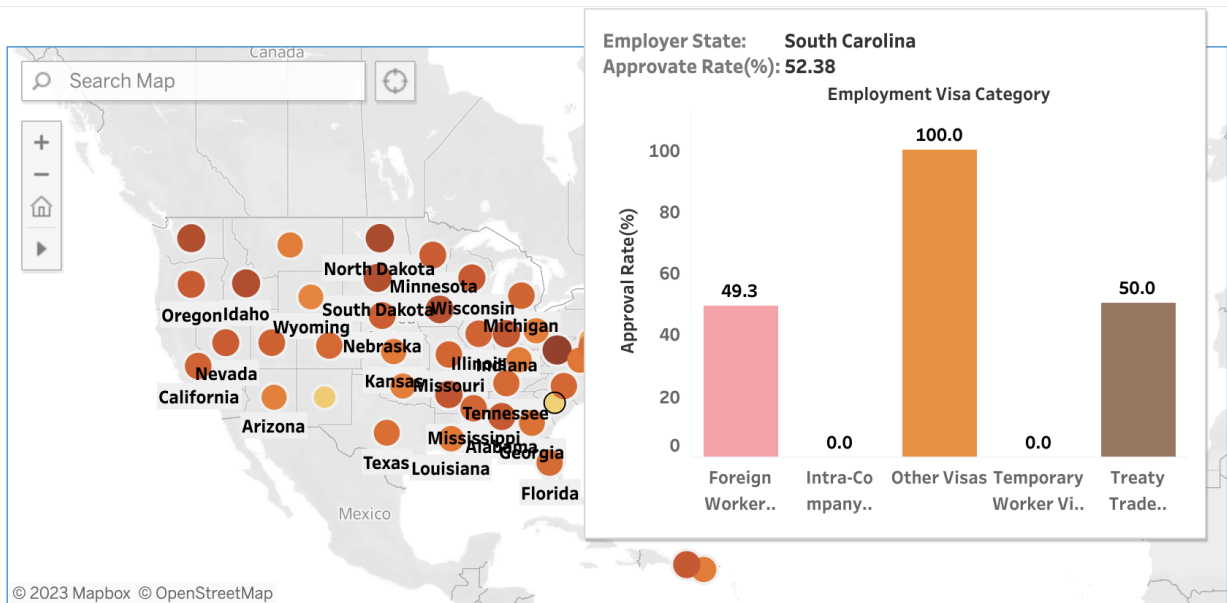
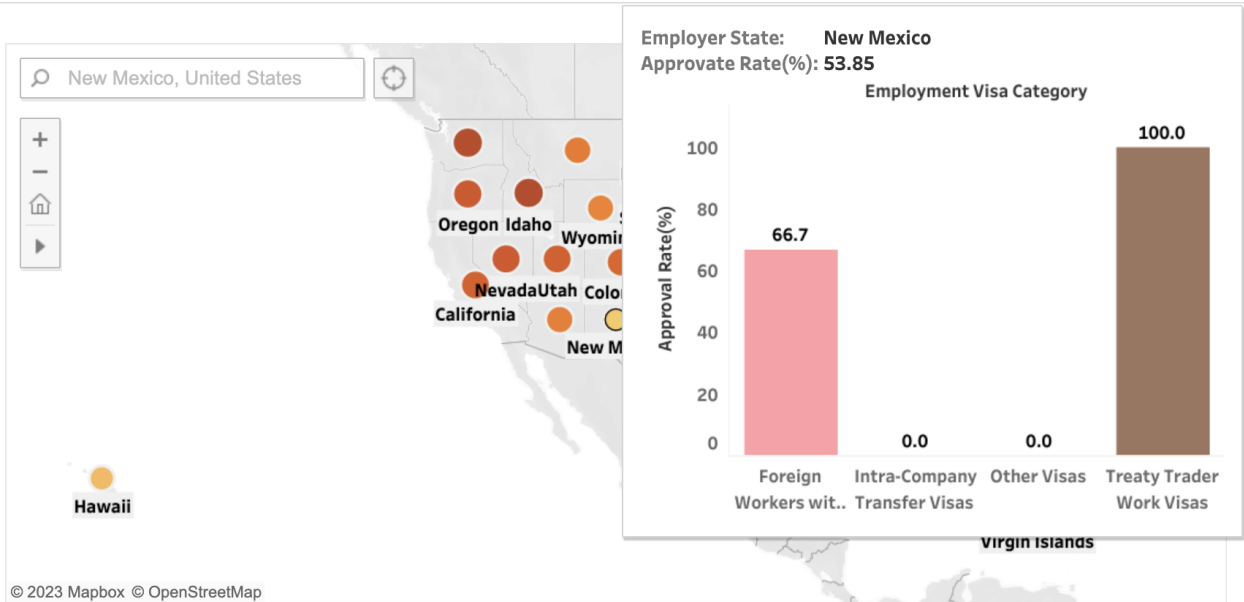
factors to consider, but we do see a wide difference between approval rates for the same occupation in different countries.

### 3.6 Employer Location

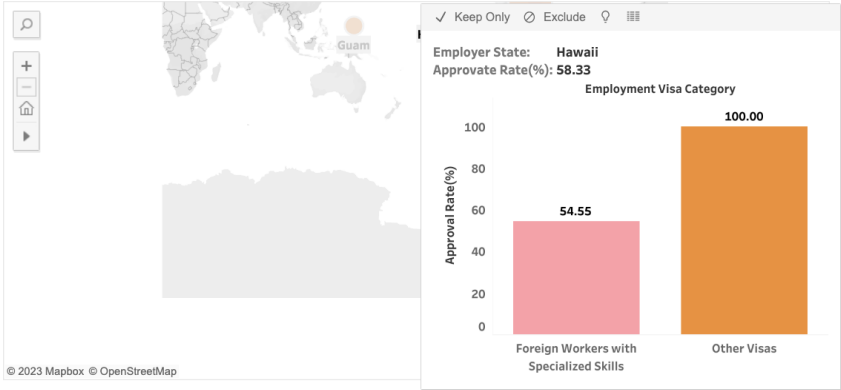


This United States map plots out the approval rates across the States. The darker the color, the higher the approval rates. It can be observed that in general the Midwest region has higher approval rates than both coasts.





We can click on the map to see the visa type breakdown. Three outliers of low approval rates are Hawaii, New Mexico, and South Carolina, whose overall approval rates were impaired by those of Foreign Workers with Specialized Skills. These individuals have a higher than 80% chance of obtaining a Green Card in most other states, but their chances are considerably reduced in these three states.



We can further uncover the underlying reasons for disparity by examining the approval rates across different industries in that state. Let's take Hawaii as an example. Previous charts show that the Education industry has one of the highest approval rates of 84%. However, it is significantly lowered to 50% in Hawaii. It can be inferred that for some professions, location could be a determining factor in the approval of Green Cards.



# 4. Data and Tools Used

## 4.1 Data

We used a kaggle dataset about US permanent visa applications as our source.

Link: <https://www.kaggle.com/datasets/jboysen/us-perm-visas>

Columns that we used are:

'case\_received\_date', 'decision\_date', 'country\_of\_citizenship', 'class\_of\_admission',  
'employer\_name', 'employer\_num\_employees', 'employer\_yr\_estab',  
'foreign\_worker\_info\_state', 'job\_info\_alt\_field\_name', 'job\_info\_alt\_occ',  
'job\_info\_alt\_occ\_job\_title', 'job\_info\_alt\_occ\_num\_months', 'job\_info\_foreign\_ed',  
'job\_info\_foreign\_lang\_req', 'job\_info\_work\_state'

## 4.2 Tools

- For exploratory data analysis and making our charts, we used Tableau.
- For building the website we used HTML and CSS.
- For making graphics and icons for our website we used Figma.
- For data processing and cleaning we used python, pandas and numpy.

# 5. Usability Testing

## 5.1 Research Plan

**Goal:** To test how well the information visualization we've created facilitates users' exploration. We want to understand how effectively users can understand how their background and professional situation might affect their process and outcomes when applying for a permanent US visa.

**Prototype being evaluated:** To test our visualization, we built a prototype of our final website using Notion. The prototype includes text and interactive data visualizations organized similar to how they would be in the final website.

We used convenience sampling to get our participants for this study – **iSchool international graduate students**. This was due to time and resource constraints. We argue that this sampling strategy is a valid one as international students all must deal with the permanent visa process to find employment. This visualization is directly relevant to their future needs.

We primarily wanted to test three tasks:

- Find information in the charts relevant for your nationality, profession and visa type.
- Understand how your nationality, profession and visa type affects your chances of getting approved
- Find overall patterns/trends in the data by looking at the charts and articulate them.

### Overall scenario for all three tasks

You're a German foreign worker in the United States working as a public administration official. You need to apply for an H-1B to continue your employment in the country. In order to make your decision and set your expectations, you decide to learn more about the landscape around H-1B applications.

**Task 1:** Find information in the visualization about your line of work, your nationality and visa type.

**Task 2:** Draw conclusions about your likelihood of securing your visa based on the above factors.

**Task3:** Compare your visa situation to some of your neighboring countries (France, Spain etc.)

## 5.2 Findings, Problems and Solutions

These are some of our most important findings from our usability test and how we fixed them.

### 5.2.1 Complicated Filters

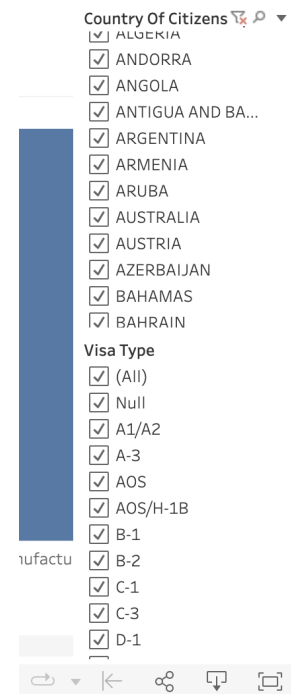
The filters were confusing and overwhelming to users because we had over 100 countries and over 20 visa types. Going through all these elements became extremely difficult.

#### What changes we made

The first change we made was to group the visa categories. All the visa categories were just arbitrary letters and numbers which meant nothing to the viewer without an explanation. We narrowed the scope to work visas and grouped them into broad categories (reduced number to 6) depending on the purpose.

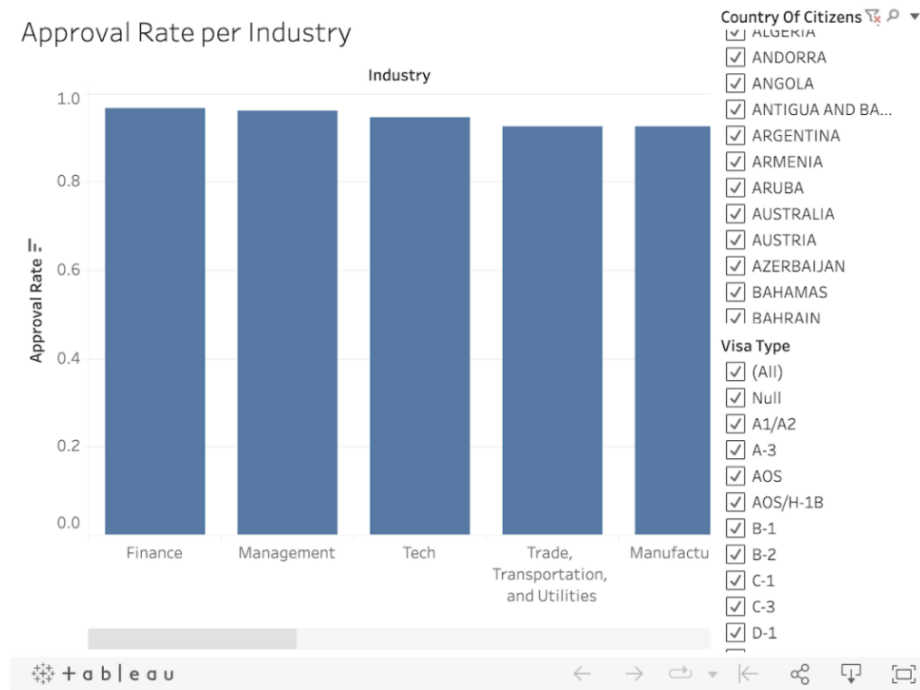
For countries, we realized that the best way to show all the nationalities would be using a map. A world map utilizes the existing mental models of the world and geography within the minds of users rather than trying to teach them a new form of encoding.

#### How it looks now



## 5.2.2 Confusing Narrative

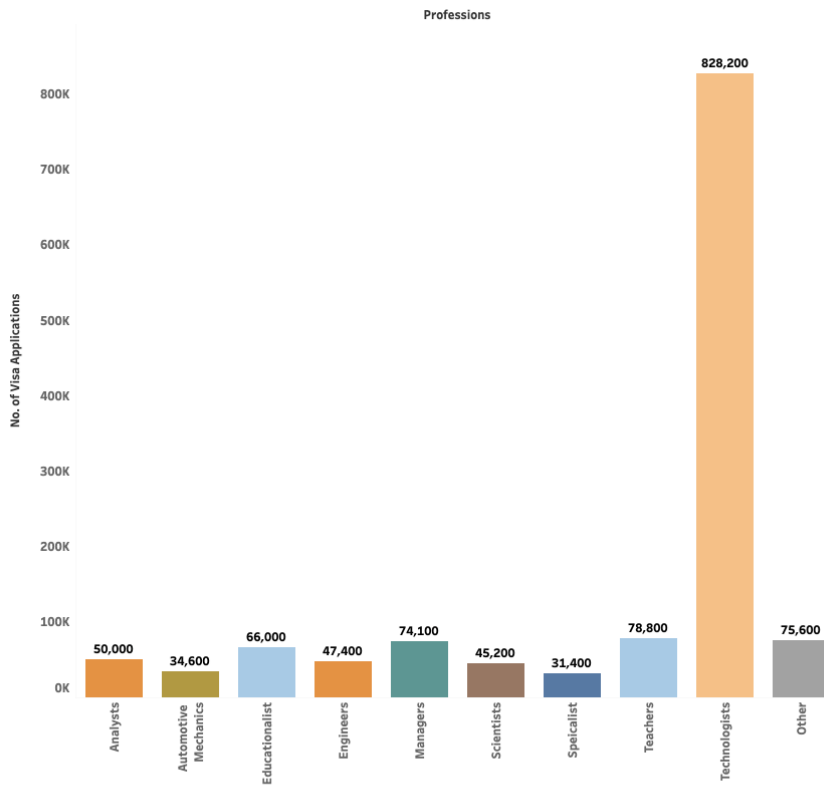
Our first prototype threw complicated visualizations at the users right off the bat. We expected users to be able to navigate the charts with the filters and find out patterns and trends for themselves. Having an exploratory tool as our project was always the idea but it proved to be challenging for new users. They were unsure of what they were looking at.



### What changes we made

For some sections, we started with a simple bar chart or other simple visualizations that showed the broad trends in the domain. For example, for the section about how profession relates to approval rates of green cards, we first show a simple bar chart displaying macro trends of the 10 most popular professions. Then we urge users to get down into the details of how this changes with different nationalities and more professions.

### How it looks now



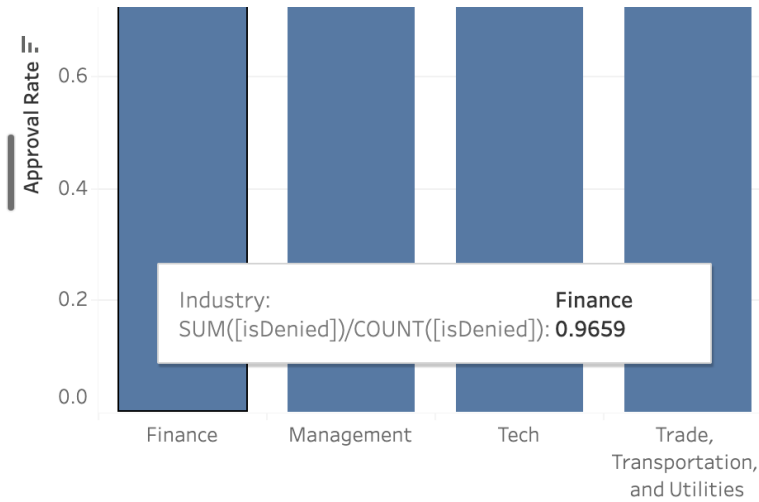
Credit: Top 10 Professions by Application - Dashboard by Shreena Bhati

's dig deeper into how this changes with nationality.



### 5.2.3 Confusing Tooltips

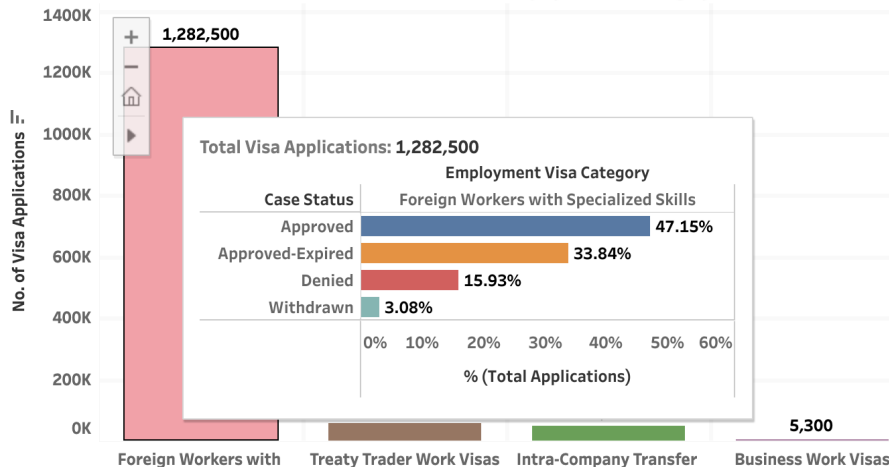
We found that users relied on tooltips a lot to make sense of the charts and see more information. The tooltips we had previously showed the formula of the value being calculated and not the actual name of the value. Users did not like this bug at all and we were asked to fix by all three of them.



### What changes we made

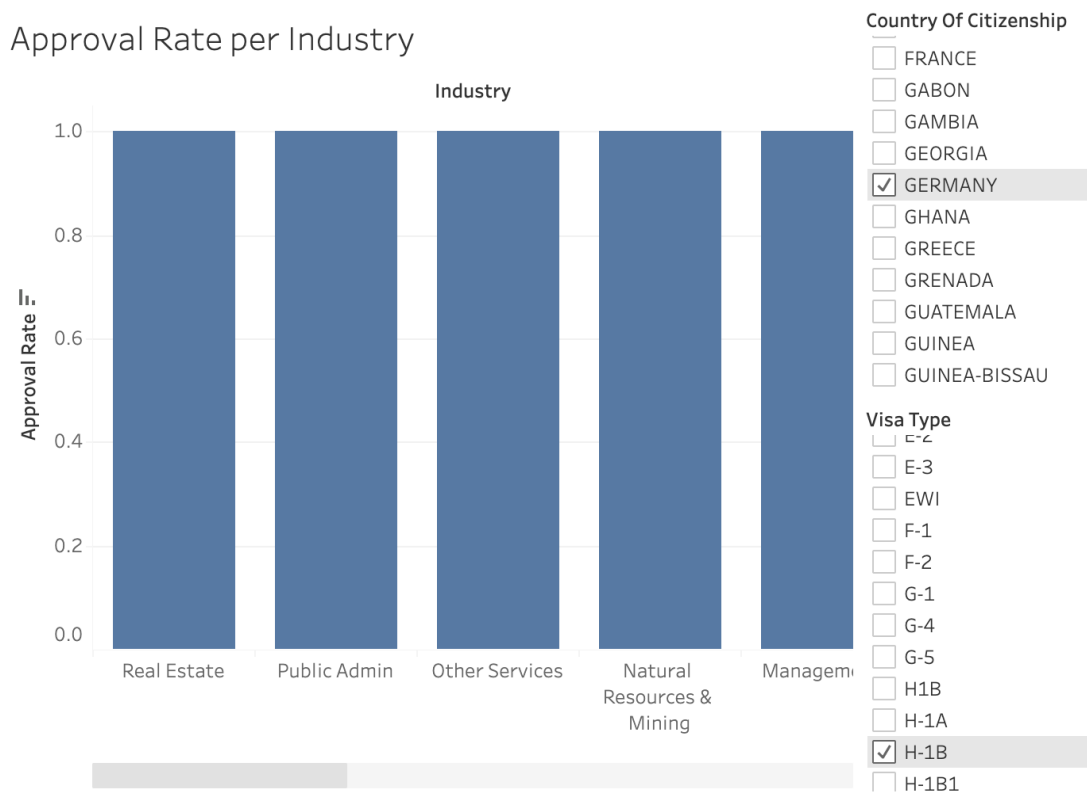
We fixed the tooltips to be more meaningful with appropriate names for the variables being displayed. However, we also noted how much users relied on tooltips and found an opportunity to utilize their behavior to communicate more information through interactivity. We embedded simple charts within the tooltip for an element of a parent chart. This provided us with another layer of information that appeared only when called for. These tooltips worked well in communicating information contextually based on some peer feedback we got on the updated charts.

### How it looks now



### 5.2.4 Difficult to compare across charts using filters

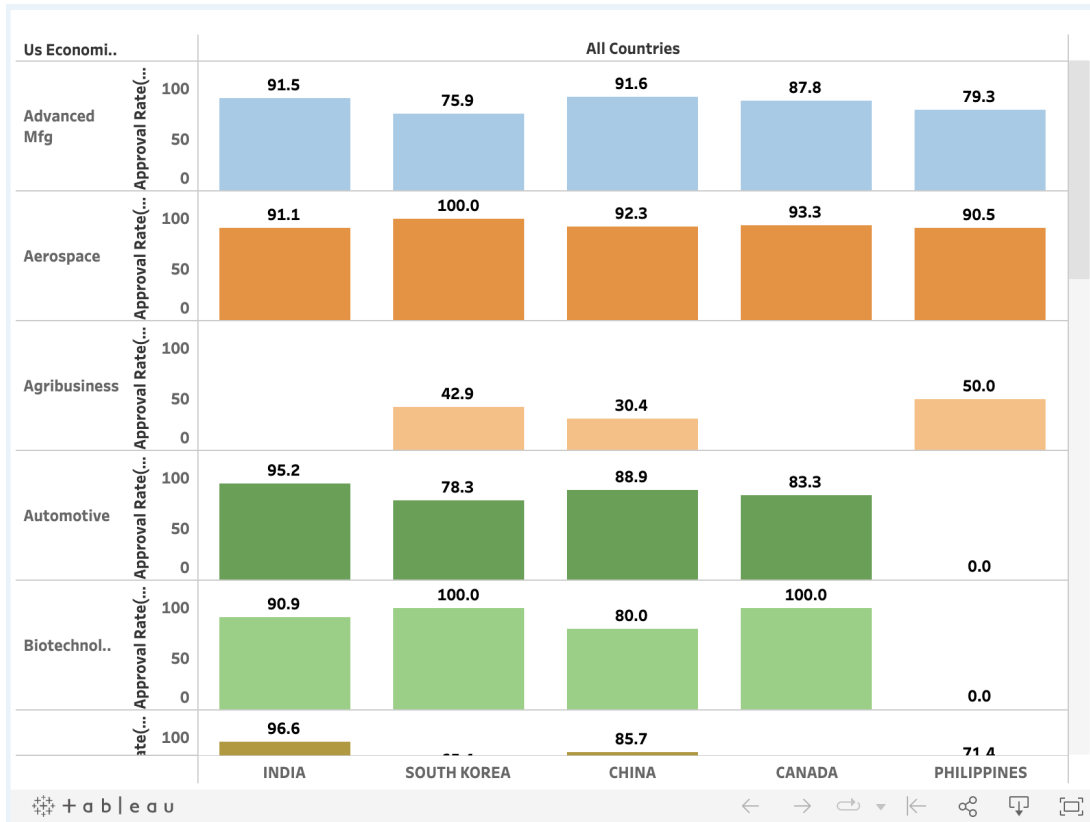
Our third task was about comparing charts between countries. During this task, users found it annoying to switch between the two countries through a filter. Using the principle of recognition over recall, we could see that this design put a lot of cognitive load on the users as it didn't allow users to compare values side by side. They had to look at a chart, remember the previous value and then switch to the next graph and check the two values in their head.



#### What changes we made

Instead of making users switch between parallels to compare values, we wanted to show important comparisons side by side. Thus we used faceting instead of filters to show comparisons wherever required. The filters were now reserved only for narrow selection and not for comparison.

#### How it looks now

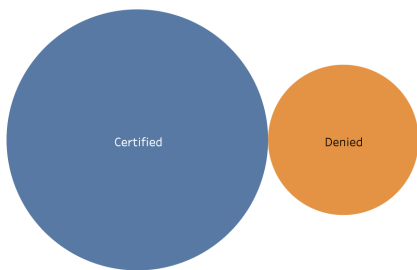


### 5.2.5 Bubble charts didn't communicate value comparisons well enough

The visualization for the number of applications accepted and denied started out as a bubble chart. Users pointed out that it was enough to give a rough sense of how much the quantities differed between categories. However, it didn't allow for precise comparisons. They could not infer the overall approval rate from this visualization.

Did you know how many visa applications were accepted between 2011-2016?

Case Status  
■ Certified  
■ Denied





## What changes we made

We changed the chart to a bar chart to enable more precise comparisons. Length encodings work better than area for numerical quantities.

## How it looks now



## 6. Member Contributions

<b>Task</b>	<b>Tarun</b>	<b>Shreena</b>	<b>Violet</b>
Project proposal	33%	34%	33%
Exploratory Data Analysis	25%	30%	45%
Sketching, Design, Iteration	45%	30%	25%
Prototyping on Notion	30%	35%	35%
Usability Testing	50%	25%	25%
Final Website	20%	50%	30%
Final Report	45%	25%	30%