UNIVERSITY OF MICHIGAN

SCHOOL OF INFORMATION

INFORMATION VISUALIZATION (SI 649)

STATIC VISUALIZATION PROJECT REPORT

EXPLORING UNAUTHORIZED IMMIGRATION THROUGH DATA VISUALIZATION

INSTRUCTED BY

DR. DALLAS CARD

February 21, 2024

Name	ID
YANZHUO CAO	yanzhuo@umich.edu

Contents

1 Introduction

This blog post delves into the intricacies of unauthorized immigration in the United States, presenting a series of data visualizations that shed light on different aspects of this complex issue.

2 Unauthorized Immigrant % of Labor Force in Each State

2.1 Tasks

The visualization, titled "Unauthorized Immigrant % of Labor Force in Each State," is designed to support several comparative and informative tasks:

- **Geographical Comparison:** It enables viewers to compare the percentage of unauthorized immigrants in the labor force across different states, highlighting geographical variations.
- **Identify Patterns:** Viewers can identify patterns of unauthorized immigration across the United States, particularly noting the concentration in specific regions.
- **Quantitative Analysis:** By providing exact percentage values for each state, the visualization supports quantitative analysis of unauthorized immigration's impact on the labor force.
- **Insightful Conclusion:** The summarizing statement in the visualization helps viewers to quickly grasp the overarching trend, which indicates a more severe unauthorized immigration problem in southern states compared to northern states.

From this visualization, viewers will learn about the distribution and magnitude of unauthorized immigrants in the U.S. labor force, gaining insights into how this issue varies regionally. The design aims to foster a deeper understanding of the unauthorized immigration issue, promoting informed discussions on the topic.

2.2 Data

- **Source:** The data for the visualization was sourced from "Pew Research Center estimates based on augmented U.S. Census Bureau data, American Community Survey 2021 (IPUMS)."
- **Questions or Concerns:** I don't know the meaning of z and the difference of column "industry with the most unauthorized immigrant workers" and column "Occupation with the most unauthorized immigrant workers".
- **Procession or Clean up:** Need to exclude the state with unknown industries. And this is reflected in my final version pictures.

2.3 Design Process

I decided to draw a map showing Unauthorized Immigrant % of Labor Force in Each State.

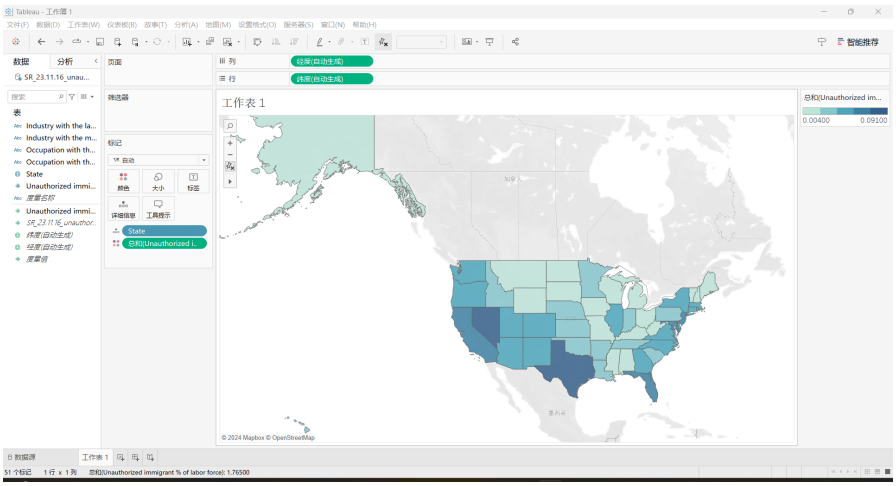

Figure 1: 1st Version

In order to emphasize the data of the states of the United States, I decided to delete all other countries in the background.

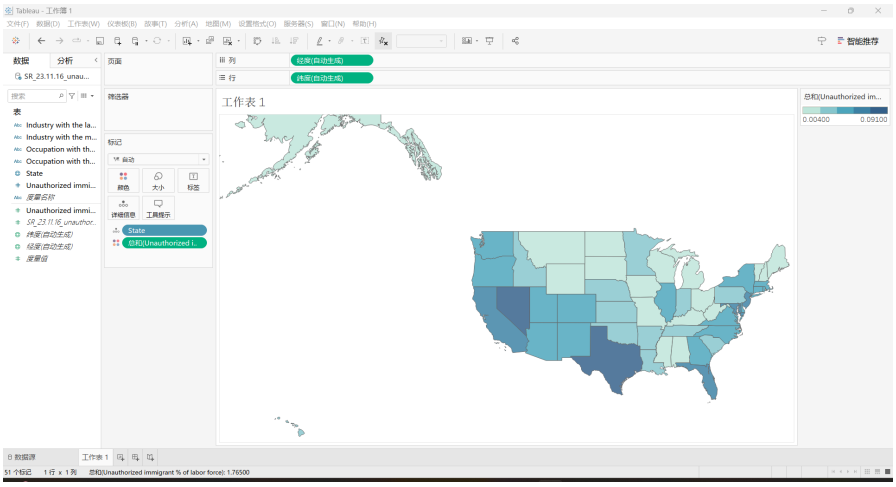

Figure 2: 2nd Version

Finally, I decided to add the actual number onto each state to make it clearer. And I also change the title and add a comment on the picture.

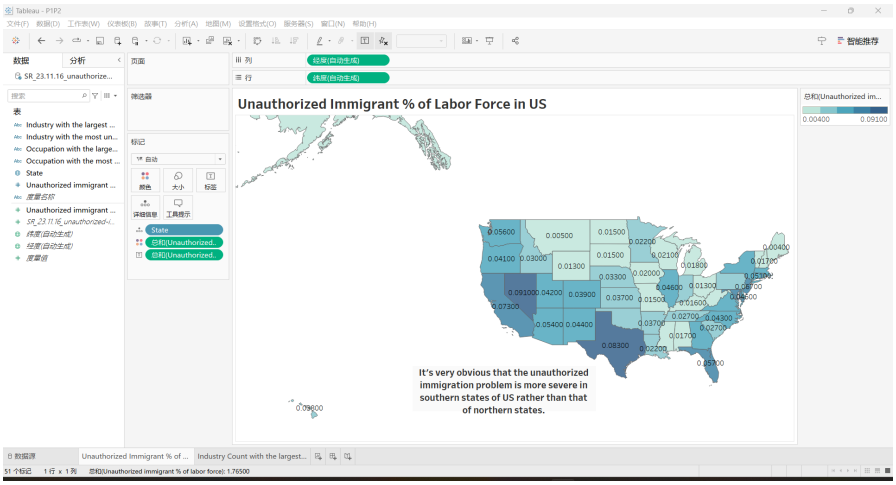

Figure 3: 3rd Version

2.4 Qualitative Self-Evaluation

- **Design Evaluation:** My design adheres to Tufte's principles, effectively using color gradients to show data without distortion, highlighting regional differences in unauthorized immigrant percentages.

- **Effectiveness:** The visualization is clear and informative, providing immediate understanding through direct data representation and quantitative labels for each state.
- **Improvement Areas:** Could enhance accessibility by considering color vision deficiencies and adding interactive elements for deeper data exploration.
- **Class Principles Connection:** Reflects our class focus on clarity, accuracy, and efficiency in data presentation, recognizing the need for continuous refinement.

3 Industry Count with the largest % of workers who are unauthorized immigrants in Different States

3.1 Tasks

The visualization, titled "UIndustry Count with the largest % of workers who are unauthorized immigrants in Different States," is designed to support several comparative and informative tasks:

- **Industry Comparison:** It facilitates a comparison of how frequently different industries appear as having the largest percentage of unauthorized immigrants across states, illustrating the varying impact of unauthorized immigration on industry sectors.
- **Highlight Predominant Industries:** By tallying the occurrences of each industry across states, the visualization identifies industries most affected by unauthorized immigration, spotlighting sectors like construction and agriculture.
- **Quantitative Insight:** The bar chart quantitatively conveys the count of states where each industry is most impacted by unauthorized workers, providing a clear numeric understanding of the issue's industry-specific distribution.
- **Pattern Recognition:** Viewers can recognize patterns in the prevalence of unauthorized immigration across industries, potentially guiding policy focus or public attention to the most affected sectors.

This visualization equips viewers with a clearer understanding of the relationship between unauthorized immigration and specific industries, facilitating informed discourse on labor and immigration policies by emphasizing the industries most influenced by unauthorized workers.

3.2 Data

- **Same** dataset as the first picture.

3.3 Design Process

I decided to draw a bar chart showing Unauthorized Immigrant % of Labor Force in Each State.

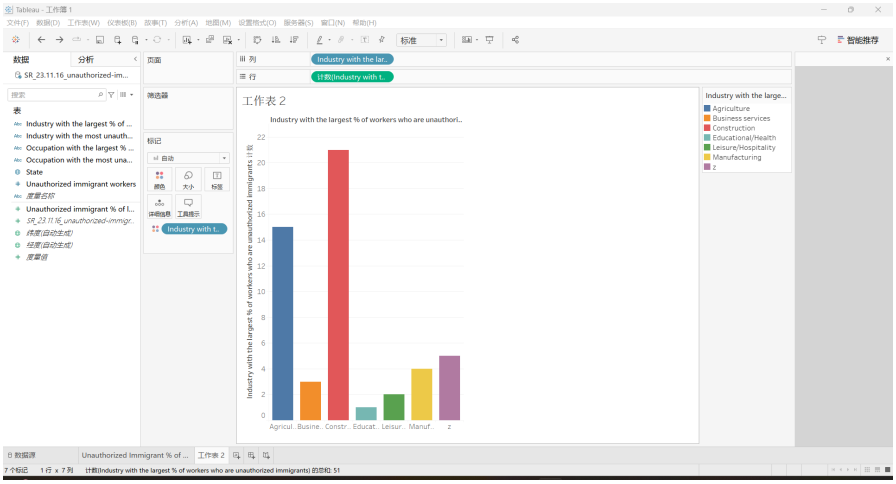

Figure 4: 1st Version

Originally, I decided to show the difference of the industries with the highest percentage of illegal immigrants the industries with the highest absolute numbers (actually they are not same for each state). But coincidentally the total counts for each industry are the same, so I gave up the following picture.

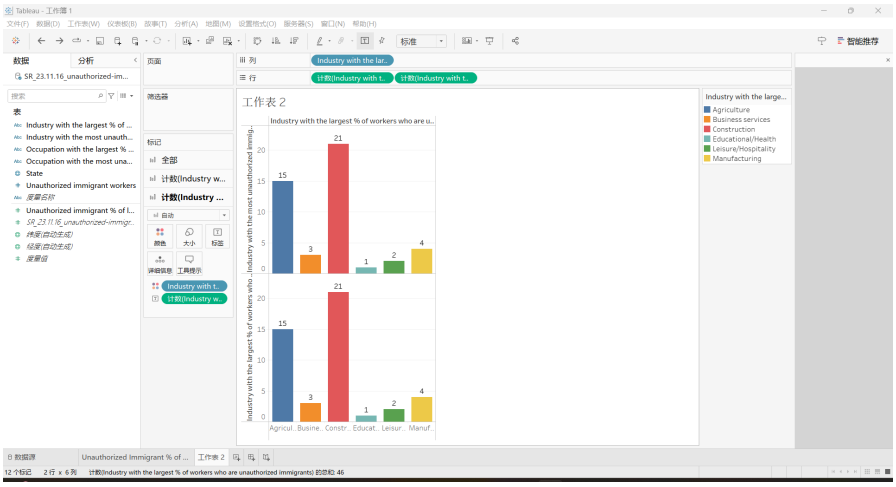

Figure 5: 2nd Version

By excluding the unknown industry z and adding the title and labels of the picture, I got the final version of my picture.

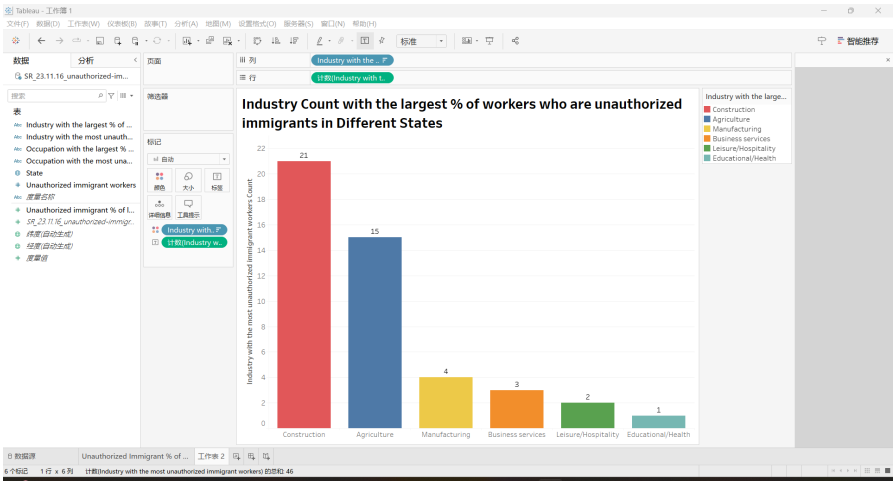

Figure 6: 3rd Version

3.4 Qualitative Self-Evaluation

- **Evaluation:** The bar chart effectively delineates industries with the highest percentage of unauthorized immigrants across states, adhering to principles of clarity and direct data representation.
- **Effectiveness:** This visualization simplifies complex labor market data,

enabling easy comparison and highlighting affected industries, showcasing the utility of targeted visual design.

- **Improvement Areas:** Future iterations could explore visual distinctions between industries with high percentages versus absolute numbers of unauthorized immigrants, enhancing analytical depth.
- **Class Principles Connection:** The design reflects learned principles of effective data visualization, with opportunities for increased interactivity and data granularity to foster deeper viewer engagement.

4 Number of Unauthorized Immigrants and Their Share of the Population in Each State

4.1 Tasks

The visualization, titled "Number of Unauthorized Immigrants and Their Share of the Population in Each State" is designed to support several comparative and informative tasks:

- **State-wise Comparison:** It facilitates a direct comparison between states in terms of the total number of unauthorized immigrants and their share of the state's total population, highlighting disparities.
- **Identify Concentration:** The plot allows viewers to identify states with high concentrations of unauthorized immigrants, with a specific focus on the magnitude (through the number of unauthorized immigrants) and the relative impact (through their share of the population).
- **Logarithmic Analysis:** By using the logarithm of the number of unauthorized immigrants, the visualization ensures that the data's wide range is manageable and comparisons are easier to make, particularly for states with large numbers.
- **Pattern Recognition:** Observers can recognize patterns of distribution among the states, such as the significant concentration of unauthorized immigrants in certain states like California and Texas, versus the majority of states clustered with lower values.

From this visualization, viewers will gain insights into the significant variances in unauthorized immigrant populations across the states. This disparity could be a contributing factor to the challenges in forming a unified national strategy for addressing immigration issues.

4.2 Data

- **Source:** The data for this scatter plot was sourced from the Migration Policy Institute (MPI) analysis of U.S. Census Bureau data from the 2015-19 American Community Survey (ACS) pooled, and the 2008 Survey of Income and Program Participation (SIPP). This analysis uses a methodology developed in consultation with James Bachmeier of Temple University and Jennifer Van Hook of The Pennsylvania State University, Population Research Institute.

Concerns and Processing: No need.

4.3 Design Process

Originally, I decided to draw a map like picture 1 to show the Unauthorized Immigrants Share of the Population in Each State.

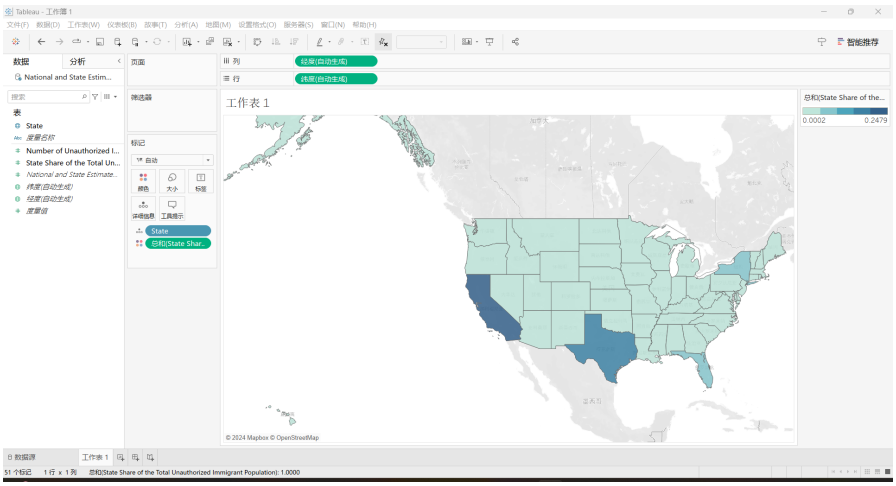

Figure 7: 1st Version

But I found that this can only show very limited information, so I decided to use a scatter plot with X-axis showing the proportion and Y-axis showing the absolute value of illegal immigrants of each state.

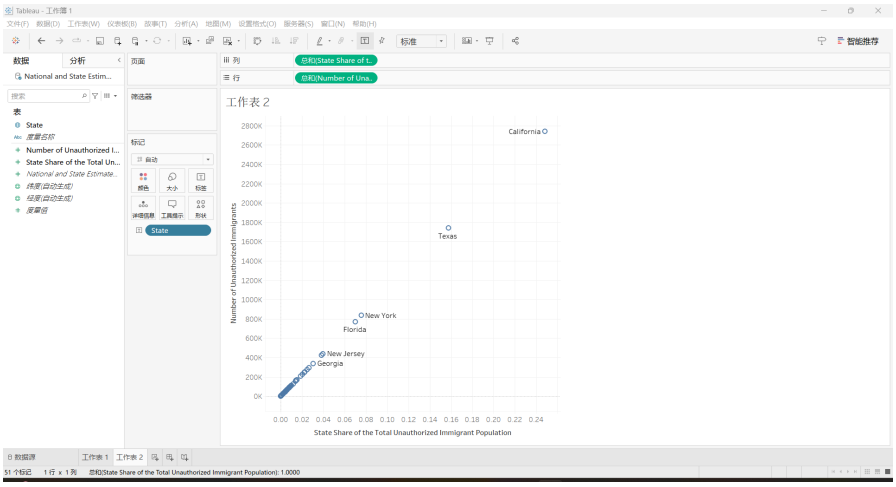

Figure 8: 2nd Version

Now I found that the Number of Unauthorized Immigrants for each state is too large, and most cities are gathering in the lower left corner of the picture. Therefore, I decided to take lg function on the y-axis and do some picture beautification to get the final version.

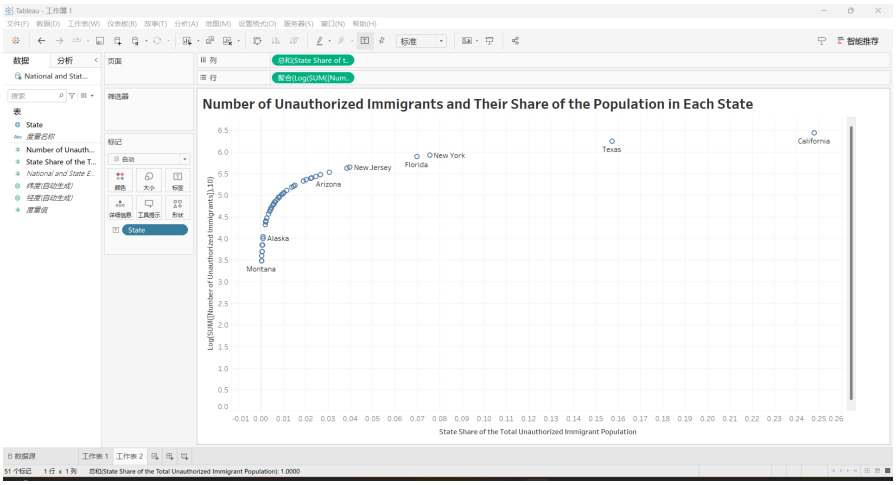

Figure 9: 3rd Version

4.4 Qualitative Self-Evaluation

- **Evaluation:** My scatter plot visualization effectively demonstrates the distribution of unauthorized immigrants across states, employing logarithmic

scaling to manage the wide data range and enhance readability, reflecting Tufte's principle of data integrity and clear presentation.

- **Effectiveness:** The plot's design facilitates easy comparison and pattern recognition, directly aligning with the principle of presenting complex information clearly and efficiently, allowing viewers to discern trends in immigration data.
- **Improvement Areas:** While the visualization successfully conveys the intended message, incorporating interactivity could improve user engagement and understanding, offering a deeper exploration of the data.
- **Class Principles Connection:** This design reflects class principles emphasizing clarity, precision, and the importance of efficient data representation. Future iterations could benefit from further reducing non-data ink and exploring alternative visualization methods to accommodate diverse data aspects.

5 Proportion of Different Nationalities in All Border Encounters During 2020-2023

5.1 Tasks

The visualization, titled "Proportion of Different Nationalities in All Border Encounters During 2020-2023" is designed to support several comparative and informative tasks:

- **Comparative Analysis:** It enables a straightforward comparison of the frequency of border encounters for individuals from different countries, highlighting disparities in the volume of encounters.
- **Pattern Recognition:** Viewers can identify significant patterns in border encounters, notably the predominant proportion of encounters involving individuals from Mexico compared to other countries.
- **Quantitative Understanding:** The chart provides a quantitative breakdown, offering precise counts of border encounters by nationality, which helps in assessing the relative magnitude of encounters per country.

- **Insight on Immigration Dynamics:** By showing the concentration of encounters among a few nationalities and the relative rarity among others, the visualization underscores the varying impact of immigration issues across different nationalities. This may hint at underlying factors such as geographic proximity, socio-economic conditions, or political situations driving these patterns.
- **Highlighting Diversity and Disparity:** The visualization draws attention to the diversity of nationalities involved in border encounters and the significant disparities among them. This insight can prompt discussions on border management policies, resources allocation, and international cooperation on immigration matters.

Viewers will gain an understanding of the distribution and scale of border encounters by nationality, recognizing the complexity of immigration issues and the need for nuanced policy responses.

5.2 Data

- **Source:** The data for this visualization was sourced from the U.S. Customs and Border Protection (CBP) official website, under the ”Nationwide Encounters” statistics section (https://www.cbp.gov/document/stats/nationwide-encounters).
- **Concerns and Processing:** Actually, I don't know the meaning for some of the columns, like “AOR (Abbv)” “Demographic”. But since I won't use that data, I don't need to process it.

5.3 Design Process

Originally, I decided to draw a pie chart like to show the Proportion of Different Nationalities in All Border Encounters During 2020-2023. But then I found that there are too many countries, which made me change to the bar chart.

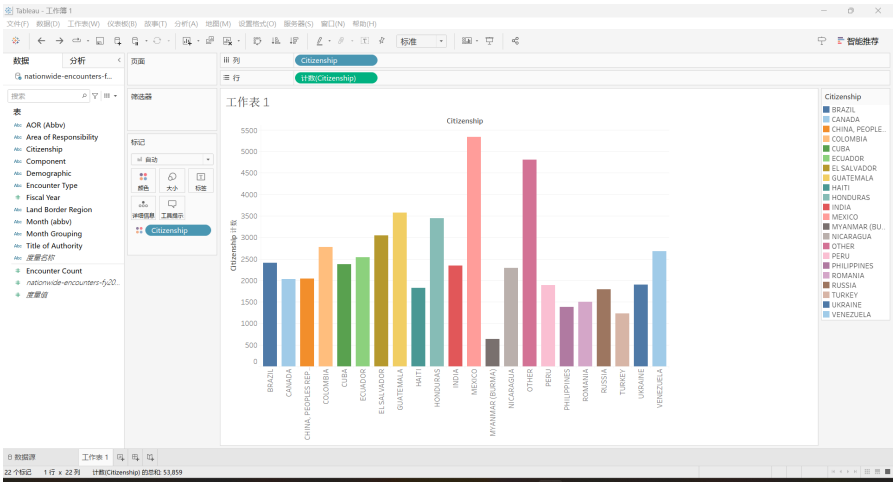

Figure 10: 1st Version

To make the result more intuitive, I delete the 'Other' group and sort the values. After adding the title and labels, I got the final version of my plot.

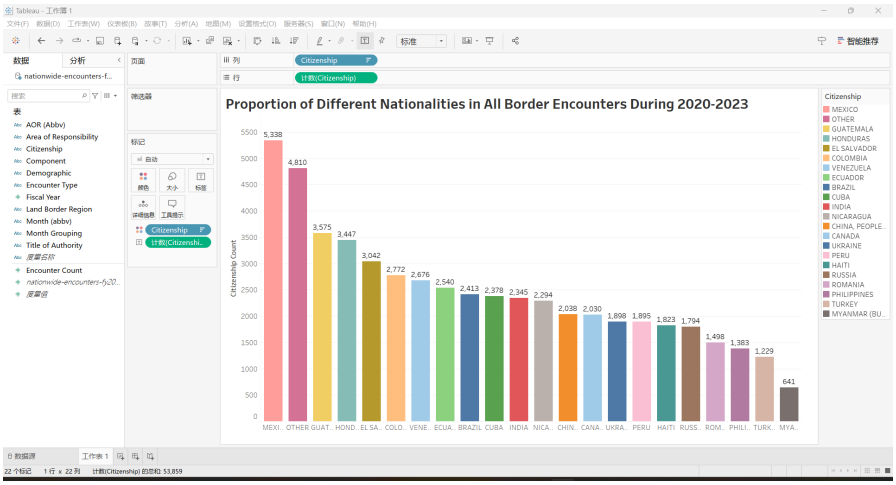

Figure 11: 2nd Version

5.4 Qualitative Self-Evaluation

- **Evaluation:** The bar chart clearly shows nationality proportions in border encounters, aligning with Tufte's clarity and detail principles.
- **Effectiveness:** It simplifies complex data, enhancing comparative analysis and pattern recognition in line with effective data presentation principles.

- **Improvement Areas:** Interactive features and a broader representation of 'Other' nationalities could provide deeper insights and inclusivity.
- **Class Principles Connection:** Adheres to minimizing non-data ink, with room for interactive enhancements to improve viewer engagement and data exploration.

6 Change of Different Encounter Type Counts with Years

6.1 Tasks

The visualization, titled "Change of Different Encounter Type Counts with Years" is designed to support several comparative and informative tasks:

- **Trend Analysis:** It facilitates an in-depth examination of the yearly trends in different types of border encounters, namely Apprehensions, Expulsions, and Inadmissibles, from 2020 to 2023.
- **Comparative Insight:** By presenting the counts of each encounter type as separate lines on the graph, the visualization enables viewers to compare the frequency of each encounter type directly across the given years.
- **Aggregate Analysis:** The observation that the sum of all types of encounters has been increasing annually provides a macro-level insight into the growing trend of border encounters overall.
- **Dynamic Observation:** Highlighting the specific trend of each encounter type—Apprehensions and Inadmissibles increasing rapidly, whereas Expulsions increase initially but then decrease— allows for a nuanced understanding of how policies or external factors might be influencing these patterns.

Viewers will gain a comprehensive understanding of how border encounter trends have evolved over the specified period, facilitating informed discussions on border security, immigration policies, and their implications.

6.2 Data

- **Same** dataset as the 4th picture

6.3 Design Process

Originally, I decided to draw a line chart like to show the Total Encounter Counts with Years.

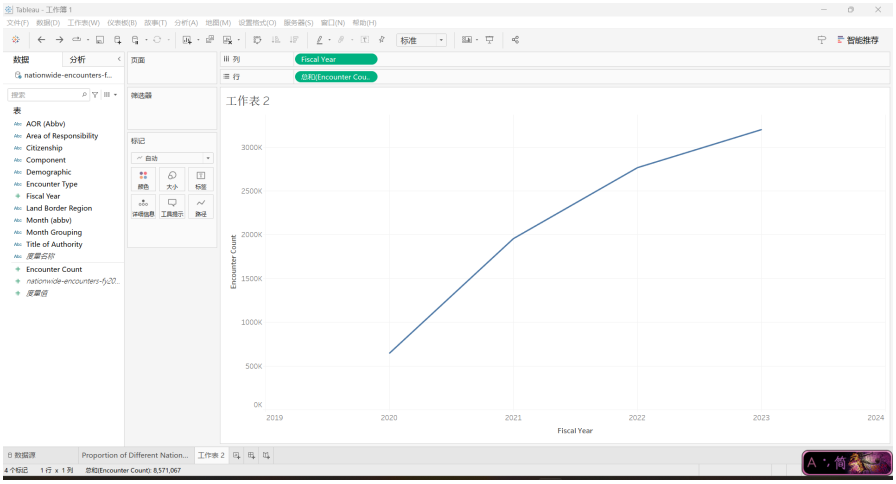

Figure 12: 1st Version

But then I found that the information it showed is too limited, which made me decided to show the details of different encounter types.

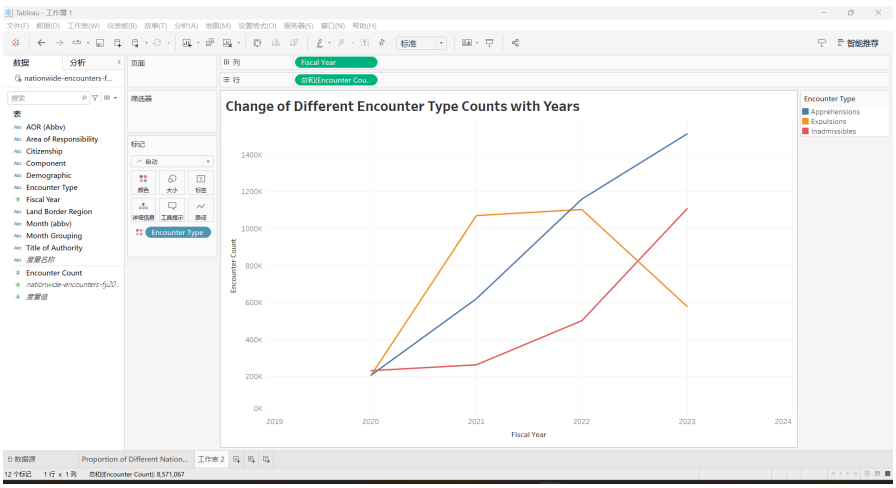

Figure 13: 2nd Version

6.4 Qualitative Self-Evaluation

- **Evaluation:** My design clearly visualizes the trend of border encounters over four years, effectively distinguishing between Apprehensions, Expulsions, and Inadmissibles. This approach aligns with the principle of clear data representation and comparison.
- **Effectiveness:** The line chart enables viewers to easily discern trends and compare encounter types, showcasing the effectiveness of data visualization in revealing patterns and insights.
- **Improvement Areas:** Incorporating interactive features could enhance user engagement, allowing for personalized data exploration. Adjustments for accessibility, such as color-blind-friendly palettes, would make the visualization more inclusive.
- **Class Principles Connection:** The design reflects principles of clarity, efficiency, and accurate data representation, emphasizing the importance of thoughtful visualization in conveying complex information. Future improvements could focus on enhancing interactivity and accessibility, further aligning with best practices in data visualization.