

Matej Popovski

STAT 436

Homework 4: Social Data Visualization

Github link: <https://github.com/matejpopovski/hw3-data-visualization>

Statistics on Illegal Immigration to the USA

(a) Before searching online, I considered different ways to visualize illegal immigration in the US statistics. Some key aspects I wanted to explore include:

- “How has illegal immigration changed over the past few years?": A line chart (regression) showing the number of illegal immigrants by year to understand patterns of increase or decrease.
- “What are the major source countries?": A bar chart, pie chart or a world map illustrating the origin countries of illegal immigrants.
- “Which states are most affected by illegal immigration?": A heatmap of the USA highlighting states with the highest number of illegal immigrants.
- Border Crossings vs. Deportations: A dual-axis chart comparing the number of illegal border crossings and deportations over time.

Challenges and Solutions:

- Data Availability – Finding comprehensive datasets that include reliable statistics on illegal immigration might be difficult. I could use sources such as Pew Research, DHS, or the Center for Immigration Studies.
- Bias in Data Representation – Different sources may present data differently depending on political perspectives. I should cross-reference sources and ensure a neutral approach.
- Overcrowded Visualization – If too many variables are included in a single chart, it might be overwhelming. Using multiple simple charts instead of one complex figure could improve clarity.

(b) Analysis of Publicly Posted Submissions

To explore the visualization of illegal immigration statistics in the U.S., I examined two primary sources and one supplementary source:

- U.S. Customs and Border Protection (CBP) Nationwide Encounters:
<https://www.cbp.gov/newsroom/stats/nationwide-encounters>
- U.S. Border Patrol Encounters by Tony Galvan:
https://gdatascience.github.io/us_border_patrol_encounters/us_border_patrol_encounters.html
- Supplementary: USAFacts on Unauthorized Immigration:
<https://usafacts.org/articles/what-can-the-data-tell-us-about-unauthorized-immigration/>

Data Presentation and Accessibility:

CBP Nationwide Encounters:

- The CBP provides official statistics on encounters, including U.S. Border Patrol apprehensions and Office of Field Operations inadmissibles. The data is presented in tabular format, allowing users to perform custom analyses. The data is categorized by factors such as area of responsibility, state, and encounter type, making it easier to analyze specific aspects. However, this visualization might be a little overwhelmed with too many selection options. The CBP's tabular data uses textual and numerical encodings, which, while precise, require users to interpret trends without visual aids. This design is suitable for detailed analysis but may not cater to audiences seeking quick visual insights. Providing summary statistics or key insights alongside raw data tables might also aid users in quickly grasping essential information.

Tony Galvan's Visualization:

- Galvan offers analysis of CBP data featuring plots that depict trends over time. The visualizations are user-friendly, allowing readers to grasp patterns quickly. For instance, the 'Inadmissible Aliens' plot distinguishes encounters (by using histograms) involving citizens from Cuba, Haiti, Nicaragua, and Venezuela (CHNV) across different border regions. Galvan employs curve and bar charts to represent encounter trends over time, effectively using position and length encodings to depict changes. Color differentiation highlights interior vs southern border. Incorporating interactive visualizations could enhance user engagement and make data trends more immediately apparent. Lastly, his visualizations lack a plot that shows only the trend of the inadmissible encounters (he has shown Admissible, All and CHNV Encounters).

USAFacts Article:

- USAFacts provides an article discussing unauthorized immigration statistics, including the number of border encounters and demographic information. The article shows great visualization for Apprehensions, Inadmissibles and Expulsions. It also show great reversed bar plot where on the y-axis we can see the top 22 countries from where the most of the illegal immigrant come form. Lastly, USAFacts shows a shaded map od the USA where most of the illegal immigrant reside. The article utilizes static charts and graphs to present data trends, employing color and spatial positioning to differentiate between categories. Some technical terms and concepts could deter readers who aren't familiar with immigration-related jargon.

Conclusion – response to prompt

Illegal immigration into the United States is a complex issue shaped by policy changes, economic conditions, and global migration patterns. Using data from TidyTuesday (U.S. Customs and Border Protection), I developed interactive visualizations to explore key trends.

A time-series analysis of monthly encounters from 2020 to 2024 reveals fluctuations influenced by policy shifts and seasonal migration patterns, with a LOESS smoothing curve and polynomial regression (including demographic) capturing long-term trends. Starting from 2020 the illegal immigrations have been increasing until late 2024, when it slightly decreased. That could be a result of the change of the president at the end of 2024. Linear regression here wouldn't be very appropriate because, we would not be able to recognize the decrease in the last year.

A global heatmap (world map) highlights Mexico as the dominant source of illegal immigration, followed by countries in Central and South America, such as Venecuela, Guatemala, Honduras, Cuba and so on, while a U.S. state-level map illustrates regional disparities in immigrant distribution, with border states experiencing the highest concentrations. We can visually see that the southern states have drastically more illegal immigrants compared to the northern. Some of the states with the highest concentrations of illegal immigrants include Texas, which accounts for nearly 50%, followed by California at almost 15% and Arizona at 19%. This aligns directly with the previous visualization, highlighting Mexico as the primary source of illegal immigration. Furthermore, the map visualizations are designed to be user-friendly and easily interpretable by both STEM and non-STEM audiences. These visualizations provide a data-driven perspective on immigration, helping to contextualize trends beyond headlines and political discourse.