

Book Bans In America

Abstract

Book bans are rapidly on the rise throughout America, and with little time to analyze the main contributing factors, creating a dashboard to quickly deduce correlations to take actionable steps is extremely important. The main tasks that are investigated are the geographical spread of book bans across the nation, the location of public libraries for access to banned books, the academic success of book banning states, the financial investments given to the education system in states, and the overall demographic makeup of the high schools. To foster both general discovery and the ability to look at data granularly, the dashboard created has two scope settings, examining all fifty states at once and a platform to choose states to compare. The data demonstrated that while there was a higher concentration of book bans in the southern states, book bans were present across the nation. It was concluded that while most states with high book bans were poor at funding education, the funding levels alone could not predict book ban levels. Lastly, the demographic makeup of the high school enrollment for high book ban states were very low in diversity, which could be a proponent for censoring marginalized communities. The dashboard provided key insights to the fight against book bans, and will help activists against bans understand the problem and take action.

Introduction

A unique and important feature of American society is the importance of freedom of speech and all the diverse opinions of its citizens. To develop the next generation of Americans into intelligent and well-rounded citizens, introducing each child to a variety of viewpoints on a continuous spectrum of topics is critical. This exposure teaches each child about cultures outside their community, and allows them to grow into the unique individual that they are meant to be. Censorship of information prevents children from being able to make their own decisions and make sense of the world around them. Censorship, through the banning of books in libraries, has been on the rise in the past few years, particularly targeting books exploring diverse sexual

orientations and racial backgrounds. These stories are important to share because it gives a voice to marginalized communities, validates people's experiences, and allows children to learn about lived experiences outside of their own. While many in our community understand censoring information is harmful, there have been hundreds of book bans throughout America, just in the last few years. The primary people fighting against these bans are city council lawmakers, librarians, and anti-book ban parents. Previously, the main visualizations lawmakers or public citizens had to understand the problem more and build plans were infographics (Admin. "Infographics.") displaying metrics or simple charts. This project is focused on providing the fighters against book bans a simple yet highly effective visualization dashboard to both extract data and discover unique and new trends emerging.

Dataset

To investigate the book banning trend and the makeup of those communities, an accumulation of different datasets are needed. The Center of Data Innovation has collected book ban data between July 2021 and March 2022 in America. Per book ban, this dataset provides the school district, state, book, author, and information on the ban's status. This information will provide the necessary information to visualize the prevalence of book bans per state, across the county. To understand each state's financial investments in education, their demographic makeup, and each state's academic success, a multi-dataset collection of US Census data was used from a Kaggle website. The latest school year of the dataset (2016-2017) was used. The dollar amount that the state budgets per student provides insights on the state government's priorities. The US Census provides data on total revenue and enrollment metrics to be able to calculate the average dollar investment into each student enrolled per state. The census data also provides the demographic makeup of the high school students per state. To understand the academic performance of the schools, The US Census provides the average 8th grade math and reading levels per state. To provide an actionable outlet to the visualization, a Kaggle dataset of

all the locations of public libraries in the nation was used, to provide the viewer with a location to check out any banned books if they are in an affected area. Lastly, to be able to merge these large datasets, a small dataset from Kaggle with the state names and state codes was used. The main pre-processing that occurred was aggregation and merging of datasets using the state names and state codes.

Tasks

This visualization's goal is to provide anti-ban policy makers and parents with an easy to use dashboard that they can pinpoint locations of high banning rates and understand the affected community. Specifically, the dashboard will provide users the ability to answer the following questions:

1. Which states are the main culprits of book bans? Are they geographically segmented or spread out?
2. Where are the locations of public libraries to get access to banned books?
3. Are there financial investment differences between book banning states and not? How do they differ?
4. Are there major academic achievement differences between states with book bans and without?
5. What is the racial demographics of the states that ban the most books?

These goals are best achieved by looking at the country as a whole, as well as being able to parse out specific states, or sets of states, for a more granular view.

Solution

To create a user-friendly UI that explored many facets of the book ban ecosystem, Plotly and Dash was the chosen software as it is easy to use and creates elegant designs. The visualization has two main dashboards, each unique by the scope of data they display. These displays are separated through the use of tabs at the top of the dashboard for easy navigation. The

first dashboard displays all fifty states together, while the second dashboard allows for choosing individual states to be displayed for optimal comparison and insight. The order of the tabs was intentional. Being initially prompted with clear visualizations of all the data forces the viewer to examine the data in its totality, leading to surprising insights or clear validation of predicted results. After looking at the wide view of the data and highlighting states they want to investigate further, the viewer has easy access, in the second tab, to do in depth analysis between specific states.

The first task is to investigate the spread of book bans and be able to highlight the main culprits of book bans. To understand the geographic distribution of book bans, a choropleth map was used, with color as the channel of book ban counts. Color is a very clear way to compare the quantitative value of counts. A white to dark red continuous color map amplified the intensity and negative emotion towards high ban states. While major color differences can be easy to differentiate, minor hue differences can be difficult. By adding a tooltip with the concrete metric of bans in the state, it provides further validation to the viewer about their assumptions from the colors. In addition to the bans count, the tooltip lists the five most common books banned in the state. While looking into each state, the viewer can see trends of prominent books being banned. To investigate the locations of the public libraries (Task 2), as well as adding actionable information to the graphic, a toggle switch, labeled “Show Libraries” at the top right of the dashboard, overlays a scatter map of the locations of all the public libraries across the nation. To equip the viewer with easy steps on how to get their affected students access to banned books, the addition of a tooltip with the address of the library and the number of young adult programs, this. The ability to toggle the scatter map overlay on and off allows for the information to be absorbed but not become overwhelming. This information is less about insights and more about actions, so the ability to mute the information is highly beneficial during an analysis of the map. This plot is our launching point to compare potential exterior influences on book bans in the next plots.

To navigate to the next plot, there is a drop down menu at the top of the dashboard. The next plot is looking at the Funding per Student vs the number of bans. The goal of this plot is to analyze whether the relationship of financial investment in the education system has any correlation with the number of bans (Task 3). The scatter plot is an effective visualization of these two quantitative metrics because the viewer can clearly see the distribution of data and see if there are any trends due to clumping or spread. A tooltip is added to give greater precision on the metrics.

The next plot is analyzing the Education scores of each state (Task 4). The average score for 8th grade reading and math did not vary that drastically between the states. While the score values do not vary greatly, any difference is substantial. To emphasize the differences between states, data augmentation to transform the data around the national average was performed. This will allow the viewer to easily distinguish between relatively high performing states and lower performing states. This data was effectively displayed in a double bar graph, displaying both the math and reading scores, and sorting the states by scores. Because of the sorting, the viewer can easily see each state's rank without needing to analyze the height channel for proper comparison. This graphic is easy to digest and esthetically pleasing.

Lastly, the investigation of the demographic makeup of the states (Task 5) is displayed in a stacked bar chart of the percentage per race of the state's population. The stacked bar chart is an effective plot to analyze the different racial ratios per state. A useful interaction element of this graph is the ability to double click on a race in the legend and it will solely display that race as a bar chart. This decreases the complexity of the chart and is highly digestible to the viewer.

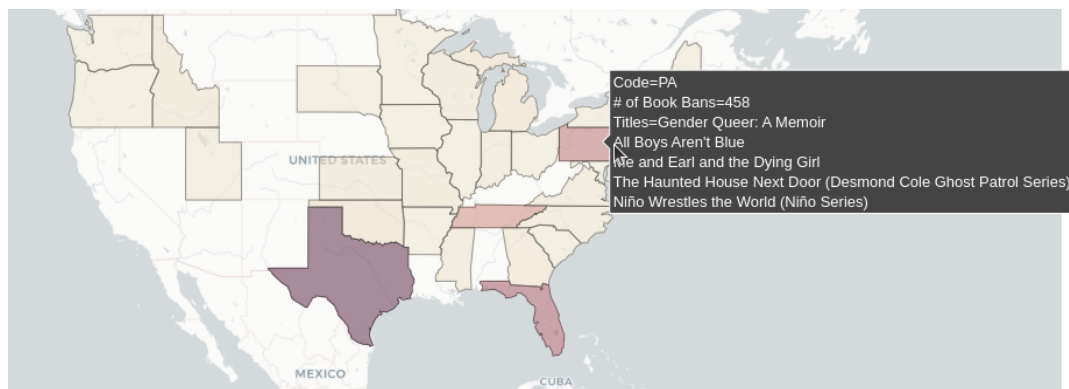
Through the use of the first tab, the viewer has made some assumptions and breakthroughs but more granular analysis would be beneficial. The second tab, Compare States, is a replica of the initial graphs but allows for custom selection of states to compare. With the easy dropdown button at the top, the viewer can select as many states as they want to investigate and populate the graphs solely with those states. This visualization is great for digging deeper

into hunches because it allows close comparison between states and the ability to see all the different facets of data in one screen.

Results

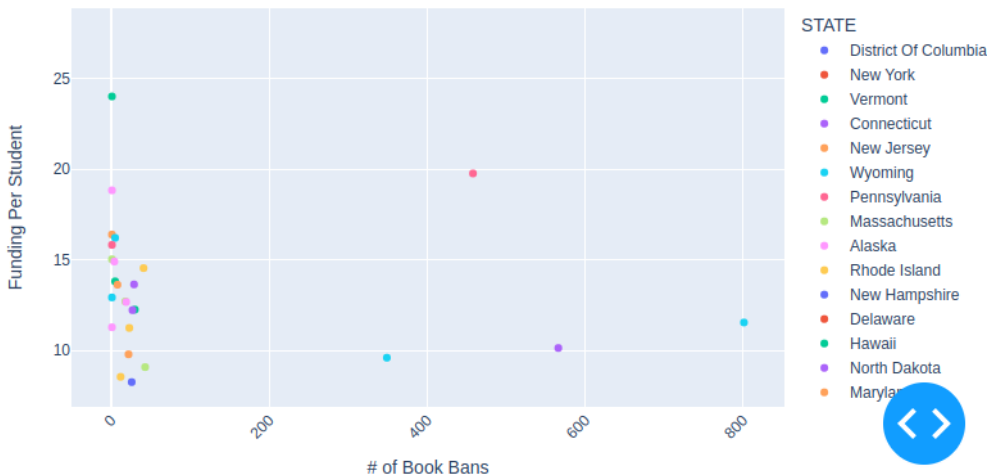
The dashboard validated both predicted hypotheses, as well as some surprising discoveries. With the pairing of the overview dashboard with the granular visualization, these discoveries were able to be accurately validated.

The investigation of the geographical spread of book bans had some unique findings. The hypothesis prior to the project was that states in the south would hold the highest summation of book bans. While Texas, Tennessee, and Florida were unsurprising culprits of bans, Pennsylvania was a surprisingly high contributor to book bans. Another surprising result was the prevalence of book bans in Oregon and Washington, both very democratic and western states. While they do not have many book bans, any bans in those states were surprising.

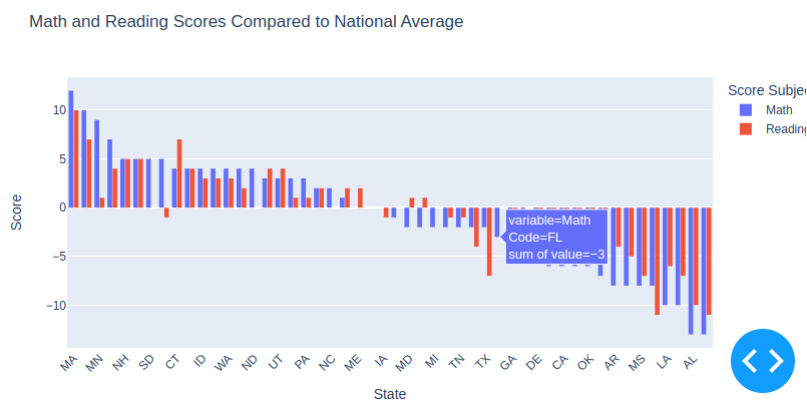


Next, the insights gained from the financial investments plots were unique. The plot effectively demonstrated that states with high book bans, with the exception of Pennsylvania, had lower investments in their education systems. In addition, the plot also clearly displayed that low investments in the education system did not explicitly predict higher book bans. This is shown

through the group of states at \$10-\$15 investments per student per day, but had barely any bans.



Next, similar to the financial plot, the academic success scores plot did not have clear conclusive findings. While high book ban states, such as Texas and Florida, had below average academic scores, Pennsylvania and Idaho, who also had substantial bans, had good academic performance. This plot unfortunately does not have the book ban metrics integrated into the plot. An improvement could be to average the math and reading scores to then use the color channel to categorize the states by high ban counts, low ban counts, and zero book bans. This will provide more informative insights on any correlations.



Lastly, the demographic makeup of high book ban states was majority white and hispanic. The findings that the high book ban states had low rates of diversity was a predicted assumption the visualizations validated this. The books that were targeted shared topics of

DSE 241, Kate O’Laughlin
Final Project

inclusion and stories of marginalized communities. The lack of diversity can create isolated communities that are not accepting of outside ideas. These graphics support this conclusion.



This dashboard provides lawmakers easy access to insights about their communities and what can perpetrated book bans. It did this through an investigative and highly customizable lens, which encourages viewers to look at the data from different angles, instead of summarizing the data for them, and providing conclusions, like previous visualizations have done.

Bibliography

Stevens, Morgan. "Tracking Banned Books." *Center for Data Innovation*, 7 July 2022,
<https://datainnovation.org/2022/07/tracking-banned-books/>.

Institute of Museum and Library Services. "2014 Public Libraries Survey." *Kaggle*, 26 Jan. 2017,
<https://www.kaggle.com/datasets/imls/public-libraries>.

Bown, Jonathan. "State Abbreviation Reference." *Kaggle*, 14 Feb. 2020,
<https://www.kaggle.com/datasets/jonbown/state-code>.

MuonNeutrino. "US Census Demographic Data." *Kaggle*, 3 Mar. 2019,
https://www.kaggle.com/datasets/muonneutrino/us-census-demographic-data?select=acs2015_county_data.csv.

Admin. "Infographics." *Advocacy, Legislation & Issues*, 3 Apr. 2021,
<https://www.ala.org/advocacy/bbooks/frequentlychallengedbooks/statistics>.