Visualizing LGBTI Survey Data

Mattew Yang

December 2024

1 Introduction

As awareness and acceptance of LGBTI individuals grow, more people are coming out to embrace their true selves. It is crucial that individuals feel free to express their identity without fear of legal or moral consequences. However, despite increasing acceptance, resistance toward LGBTI people persists. The level of acceptance, however, varies significantly across countries, depending on the specific identity within the LGBTI spectrum and the local context.

In Europe, LGBTI acceptance has steadily increased since the beginning of the 21st century, but there is still much work to be done before LGBTI individuals are fully accepted by society. A significant stigma remains, especially in Eastern European countries.

Another factor to consider is the challenge of recognizing oneself as LGBTI. Societal pressures and the dominance of heteronormativity discourage people from exploring identities outside the traditional binary of "man" or "woman" or the expectation to form relationships with someone of the opposite gender.

The goal of this project is to create a website that showcases the lives and perspectives of LGBTI people interviewed across the European Union. Using data from a 2020 survey, we aim to examine how acceptance levels differ across countries and identify correlations between responses to various questions to better understand the patterns of opinion and experience within the LGBTI community.

2 Dataset Collection

The dataset used for this visualization is the EU LGBTI Survey 2020, an aggregated collection of data from a survey conducted by the European Union Fundamental Rights Agency. The survey covered the entire European Union. The data is anonymized, with respondents grouped by country, romantic orientation, and gender identity. It consists of six CSV files, each focusing on a different topic. Each file includes questions, with respondent answers first categorized by country and then by specific subgroups.

The survey asked participants about their views on the government's and society's acceptance of LGBTI people, their comfort with disclosing their LGBTI identity to others (such as at work, school, or within their families), their experiences with discrimination, and the overall quality of their daily lives.

3 Creating the Visualization

3.1 Development Details

The webpage was developed using D3.js, a JavaScript library designed to create data visualizations in web browsers. It is particularly effective for generating Scalable Vector Graphics (SVGs), which, as the name suggests, can be resized without losing quality or resolution. This makes SVGs ideal for webpage development, as they adapt seamlessly to various screen sizes. Unlike more abstracted tools, D3.js provides fine-grained control over the placement and appearance of visuals. It was chosen for this project because it offers an efficient way to generate visuals while maintaining flexibility in customization and modification.

We also used a map to visualize the data. Initially, the map was much larger, covering Europe, North Africa, the Middle East, and West Asia. Since this survey only focused on the European Union, regions outside of Europe were unnecessary. As a result, we had to crop the map to include only Europe.

The map consisted of three separate files: borders, regions, and graticules. The borders define the boundaries between countries and the ocean. The regions color the country areas, while the graticules mark specific latitudes and longitudes. In this project, the graticules were primarily used to set the correct map projection. Essentially, the map was zoomed in to focus solely on Europe.

Country borders and regions are made up of hundreds of coordinates that form their shapes. By calculating the minimum and maximum values of the X and Y coordinates for the borders, regions, and graticules, we

determined the bounding box for each file. We then cropped the coordinates outside of this bounding box. Some countries were completely removed, while parts of others were trimmed. An exception was made for Greenland, located at the top-left corner of the map. Its coordinates were anchored to the original bounding box to ensure it remained properly aligned with the edge of the map. After these adjustments, the new files were saved in the same format as the original, resulting in three updated files.

The coordinate culling process was done offline to avoid slowing down the webpage. Since it is computationally intensive—checking around 50,000 coordinates to see if they fall within the reduced bounding box—it would have caused long loading times if done during page load.

The page is divided into four sections, two of which are relevant to the visualization: the title, the country-specific data section, the global data section, and the footer. The title and footer provide padding around the visualizations, ensuring the other sections don't touch the webpage borders. The title introduces the topic, clearly indicating what the webpage is about, while the footer credits the data source.

The country-specific data section displays the visualizations for individual countries. The data here is filtered to focus on the selected country. On the left, the map is displayed, and on the right, the data visualizations are shown. The global data section, on the other hand, presents data for all countries, grouped by country to illustrate where each one stands within a given graph.

Additionally, there is a menu fixed in the top-left corner. This menu is not part of any specific region, as it applies to both the country-specific and global sections. It allows users to further subdivide the data by different LGBTI categories.

3.2 Country-Specific Data

3.2.1 Choropleth Map

Users can choose a country on the map in two ways: by selecting from the dropdown menu below the map or by clicking directly on a country.

Upon loading the page, one of the first visuals users encounter is the choropleth map. It stands out due to its design. Unlike the rest of the website, which follows a blue color palette, the map uses a red-to-yellow-to-green gradient to represent whether respondents believe their government is doing enough to combat prejudice against LGBT people. This color scheme makes the map visually striking, especially on standard screens.

The map's dominance of dark red, in particular, draws attention. Red, commonly associated with negative connotations, creates an immediate sense of criticism, even before reading the title. While general dissatisfaction with the government is not unusual, the widespread discontent among LGBT people—especially in Eastern European countries—suggests that much work remains to be done to achieve greater acceptance of LGBT individuals.

Choropleth maps effectively display data geographically, enabling users to identify regional patterns. In this case, the map highlights geographical trends in public opinion, such as the relatively weak efforts in Eastern Europe to address prejudice against LGBT people. This visual format allows users to easily compare neighboring countries and spot regional similarities or differences, without needing to recall each country's location. Additionally, the map offers a welcome contrast to the more common bar and point charts in data visualization, providing a refreshing change from circles and squares.

3.2.2 Discrimination Bar Chart

One of the most significant concerns for LGBT individuals is whether they are accepted by society. However, this acceptance is not consistent across different situations. In places like third spaces, where people from various backgrounds interact, discrimination tends to be more prevalent. In contrast, environments such as job searches, where a certain level of professionalism and impersonality are present (e.g., recruiters see only names and resumes, which generally do not disclose sexual orientation or gender identity), experience lower levels of discrimination. Comparing these situations reveals the varying degrees of discrimination that LGBT people face.

A bar chart is an effective tool for comparing these levels of discrimination. It clearly highlights which values are higher and which are lower, making it easier to interpret differences between various situations.

3.2.3 CDF Chart: When LGB People Realized They Were LGB

Another important aspect to consider is when LGB individuals come to realize or disclose their sexual orientation. Societal pressures and heteronormativity can delay this realization, causing individuals who feel "different" or "out of place" to question their identity. Many may not recognize that they are not alone or that there are alternative ways to understand gender and sexuality. This is especially common during physical and emotional development, a time when individuals often question their beliefs and their sense of self. It is during this period that many first discover their non-heterosexual orientation or gender identity.

The timing of when people disclose their identity to others is equally important. Understanding one's sexual orientation is one thing, but sharing that identity with others can be a difficult step, particularly in the face of societal stigma. Comparing when people realize their orientation with when they disclose it to others can shed light on the relationship between internal realization and external expression. This comparison can be effectively visualized with a cumulative distribution function (CDF) chart.

A CDF chart highlights the continuous nature of this process. It shows how more individuals recognize their sexual orientation as time progresses, unlike bar charts, which present discrete data without implying continuity. The CDF also allows for a comparison between the timing of realization and disclosure, demonstrating the close relationship between these two events.

3.3 Global Data

3.3.1 Violin Plot: Average Satisfaction Comparison

LGBTI individuals face various challenges, particularly regarding discrimination and government policies. Despite these obstacles, many strive to live fulfilling lives. One way to assess this is by examining the average satisfaction of LGBTI respondents, grouped by country. It is important to determine whether satisfaction levels differ across the EU, especially since many LGBT rights, such as same-sex marriage and anti-discrimination laws, have been legalized. While these laws may suggest a general improvement in satisfaction, it is crucial to assess whether this trend holds across the entire population. A violin plot provides a useful way to examine whether average satisfaction has increased across the EU due to standardized laws.

A violin plot is ideal for this analysis because it shows the distribution of satisfaction scores. The shape of the distribution reveals how closely the data points are clustered together, with a slightly longer tail than the head. The height of the distribution and the length of the tail provide insights into the mean and standard deviation of satisfaction scores, allowing us to assess whether the data is generally high or low.

3.3.2 Diverging Bar Chart: Outsider Reactions to LGBTI People

LGBTI individuals, like anyone else, interact with a variety of people daily—at work, school, and in public spaces. These interactions can result in either support or opposition based on their sexual orientation or gender identity. The goal here is to investigate whether individuals who report receiving more support also experience less opposition.

A diverging bar chart effectively visualizes this relationship. By sorting the data on support in descending order, we can compare it with the opposing responses. This allows us to identify any trends or patterns that may emerge in the data.

3.4 Scatter Plots

Scatter plots are the ideal choice for visualizing the relationships between multi-valued variables. They help identify trends and easily reveal whether a regression line can be drawn. Additionally, scatter plots allow us to assess the quality of the regression line by showing how closely the data points align with it and how far they deviate from the line.

The following subsections explain the rationale behind comparing these particular datasets.

3.4.1 Scatter Plot Comparing Discrimination to Government Dissatisfaction

It is well-known that LGBT individuals face discrimination in their daily lives. It is also likely that they hold the government accountable for its inadequate or delayed response to enacting laws that ban discrimination or support LGBT rights. However, it is valuable to explore the strength of the correlation between discrimination and dissatisfaction with the government. This analysis can reveal whether a clear trend exists and, if so, how steep that trend is. Additionally, it is useful to assess whether this correlation holds across different scenarios, such as discrimination in job searching versus discrimination in public spaces.

3.4.2 Scatter Plot Comparing LGBT Openness at School to Work

As discussed in Section 3.2.2, discrimination varies depending on the context. Two significant environments where people spend much of their time are schools and workplaces. By comparing the openness of LGBT individuals in these settings, we can determine whether people are more likely to express their identity at school or at work. This comparison sheds light on the factors influencing LGBT individuals' comfort levels in different environments.

3.4.3 Scatter Plot Comparing LGBT Openness Between People Groups

LGBT individuals do not necessarily disclose their identity to everyone. Instead, they may choose to share it with certain people based on trust. Analyzing patterns in who they are more likely to open up to—whether friends, family, colleagues, or medical professionals—can reveal underlying trends. For instance, it may be more common for individuals to confide in people they know well rather than strangers. Such patterns provide insights into the level of trust LGBT individuals place in different groups.

4 Conclusion and Future Work

In this paper, we introduced a webpage that allows users to explore survey data by country, LGBTI subset, and global perspective. The webpage offers various visualizations to help users understand the diverse opinions and experiences of LGBTI individuals. It highlights the varying levels of acceptance towards LGBTI people across different countries.

There are several areas where the webpage can be improved. Notably, the subsets are not equally distributed—some questions received fewer responses than others, and some were left unanswered. Additionally, there is uncertainty in the responses from individuals who indicated they were unsure or did not know how to answer certain questions. It would be valuable to visualize this uncertainty, which could provide insight into potential variations and the overall confidence in the data.