

Through Shadows and Memories: Navigating the Holocaust with Data*

Chenyiteng Han

April 3, 2024

To preserve memory through data, our project leveraged the carefully cleaned Auschwitz concentration camp dataset to create the ‘HolocaustExplorer’ Shiny application. The tool is designed to make the tragic statistics of the Holocaust accessible and interactive, not only as a digital memorial but also as an educational resource that invites users to explore the personal stories behind the numbers. Our efforts aim to honor the victims while deepening understanding of this dark chapter in history and ensuring that its lessons are never forgotten.

Table of contents

1	Introduction	2
2	Data	2
2.1	Raw Data	2
2.2	Cleaned Data	3
3	Data and Memory	4
3.1	HolocaustExplorer	4
3.2	Remember the Suffering	4
4	Discussion	5
4.1	moral responsibility	5
4.2	A Humanistic Perspective on Data	5
	References	5

*Code and data from this analysis are available at: <https://github.com/Hhnxxxxxx/Holocaust.git>. Shiny app is available at: <https://puszw3-chenyiteng-han.shinyapps.io/holocaustexplorer/>.

1 Introduction

When embarking on this project, our primary aim was to illuminate personal stories that were obscured by the scale of the Holocaust tragedy. Using a carefully cleaned dataset from Auschwitz death certificates, we created the “HolocaustExplorer” shiny app. In order to strengthen the foundation of “HolocaustExplorer”, we began a meticulous data collection and cleaning process. The Auschwitz Death Certificate Dataset is derived from the United States Holocaust Memorial Museum and has been thoroughly refined. Designed to maintain accuracy while ensuring respect, our cleansing process involves distilling large amounts of data into a format conducive to analysis and interaction.

It’s a digital conduit into the murky depths of history. This effort is not just about data visualization; This is a strong commitment to ethical commemoration, educational outreach and the dignified preservation of memory. Through interactive exploration, we seek to provide a platform where the public can engage with data in meaningful ways, promoting a deeper understanding of the impact of the Holocaust. This introduction underpins the essence of our project, bridging the gap between historical data and human empathy, ensuring victims’ stories are not only remembered, but felt.

2 Data

In our project, we use several R packages (R Core Team (2022)). We utilized googlesheets4 (Bryan et al. (2022)) to download the dataset directly from Google Sheets, ensuring real-time access to data. The readr (Wickham and Hester (2022)) package facilitated efficient data reading into R. Data cleaning and transformation were accomplished with dplyr (Wickham et al. (2022)) and tibble (Wickham, Müller, and François (2022)), allowing us to refine and structure the dataset for analysis. The development of the interactive web application was achieved through shiny (Chang et al. (2022)) and enhanced with shinyWidgets (Perrier and Meyer (2022)) for improved UI elements. Finally, ggplot2 (Wickham (2022)) enabled the creation of compelling visualizations, making the data more accessible and understandable.

2.1 Raw Data

The raw dataset originates from the United States Holocaust Memorial Museum (United States Holocaust Memorial Museum (2022)), specifically from the collection titled “Auschwitz Death Certificates 1941-1943.” This poignant dataset contains records from the Death Books (Sterbebücher) of Auschwitz Concentration Camp, documenting the tragic end of 65,280 prisoners who perished at the camp between July 29, 1941, and December 31, 1943.

The dataset includes detailed personal information (see Table 1):

- **Last Name** and **First Name(s)** provide the full name of each individual.

- **Date of Birth** and **Date of Death** record the life span of each prisoner in a YYYY-MM-DD format.
- **Birthplace** and **Residence** columns list the origin and the last place of residence, respectively.
- **Religion** denotes the recorded religious affiliation, highlighting the diverse backgrounds of the victims.

This data is not just a mere collection of numbers and dates; it represents individual lives, their heritage, and the profound loss experienced by communities during the Holocaust.

Table 1: Sample of the Raw Dataset

Last Name	First Name(s)	Date of Birth	Date of Death	Birthplace	Residence	Religion
Aandagt	David	1906-04-26	1942-09-01	Amsterdam	Amsterdam	Jew
Aandagt	Hartog	1899-02-10	1942-10-27	Amsterdam	Amsterdam	Jew
Aandagt	Israel	1909-05-25	1942-10-10	Amsterdam	Amsterdam	Jew
Aandagt	Jacob	1918-03-10	1942-10-08	Amsterdam	Amsterdam	Jew
Aandagt	Jacob	1922-01-26	1942-08-22	Amsterdam	Amsterdam	Jew
Aandagt	Meyer	1914-05-31	1942-07-28	Amsterdam	Amsterdam	Jew

2.2 Cleaned Data

After an extensive cleaning process, the dataset was transformed to better facilitate analysis and understanding. The columns have been refined to include:

- **Last Name** and **First Name(s)** to maintain the individual identity of each record.
- **Birthplace** and **Residence** to illustrate the geographic origins and last known locations of the individuals.
- **Religion** to provide insight into the cultural and religious diversity of the victims.
- **Birth_Year** and **Death_Year** to focus on the life span without the granularity of exact dates.

These changes allow for a dignified yet analytical approach to examining the data, preserving the memory of each individual within the broader historical context (see Table 2).

Table 2: Sample of the Cleaned Dataset

Last Name	First Name(s)	Birthplace	Residence	Religion	Birth_Year	Death_Year
Aandagt	David	Amsterdam	Amsterdam	Jew	1906	1942
Aandagt	Hartog	Amsterdam	Amsterdam	Jew	1899	1942

Table 2: Sample of the Cleaned Dataset

Last Name	First Name(s)	Birthplace	Residence	Religion	Birth_Year	Death_Year
Aandagt	Israel	Amsterdam	Amsterdam	Jew	1909	1942
Aandagt	Jacob	Amsterdam	Amsterdam	Jew	1918	1942
Aandagt	Jacob	Amsterdam	Amsterdam	Jew	1922	1942
Aandagt	Meyer	Amsterdam	Amsterdam	Jew	1914	1942

3 Data and Memory

3.1 HolocaustExplorer

After completing the Auschwitz data set, it became imperative for us to reveal personal narratives. Recognizing that each entry is not just a data point but a representation of a human life led us to develop the “HolocaustExplorer” Shiny application. Our goal was to create an accessible, interactive medium through which the public could understand the scale of the Holocaust and the personal stories of its victims.

The “HolocaustExplorer” Shiny app features a main panel with interactive bar charts that vividly display data based on user-defined filters such as place of birth or religious affiliation. To respect privacy and the sensitivity surrounding individual identities, the app deliberately omits the display of specific victims’ names, thus preventing potential distress to relatives or individuals connected to the victims. Sidebars allow users to customize views of data to reflect real-time changes in the visualization.

Behind the scenes, the app leverages cleaned datasets to enable users to meaningfully engage with the information, encouraging a deeper understanding of each victim’s history through the lens of aggregated data. It is a digital memorial and educational tool that deepens users’ understanding and memory of tragic events, ensuring that such history is neither abstracted nor forgotten.

3.2 Remember the Suffering

Reflecting on Bouie’s paper (Bouie (2022)), it’s clear that the essence of utilizing digitized historical records goes beyond mere access; it’s about reconstructing narratives and understanding the human aspects of historical events. The “HolocaustExplorer” Shiny app, inspired by such principles, endeavors to make the Holocaust’s complex history more accessible, offering a deeper, more nuanced understanding of the data. This aligns with Bouie’s emphasis on the educational significance of digital archives, highlighting our app’s role in promoting historical awareness and empathy through interactive data exploration.

In developing HolocaustExplorer, we worked to address the ethical questions raised by Bouie about using data to tell stories of immense human suffering. Our application aims to responsibly use quantitative analysis to bring personal stories to the forefront, ensuring that the humanity of Holocaust victims is respected and remembered. In the face of the heavy history, escaping is useless. It may only bring about the recurrence of more tragedies. Only by remembering history can we always beat the suffering and move towards the light.

4 Discussion

4.1 moral responsibility

Handling and presenting Holocaust data requires navigating ethical considerations with sensitivity and respect. It's crucial to honor the memory and humanity of the victims, avoiding any form of sensationalism. Transparency about sources and methodologies is vital, as is a commitment to accuracy. Engaging with this data demands a careful balance between statistical analysis and the acknowledgment of individual tragedies, ensuring that the representation of such data does not diminish the personal experiences and atrocities faced by the victims.

Healy's insights (Healy (2020)) underscores the importance of contextualizing data within its human narrative. In creating the "HolocaustExplorer" Shiny app, we adhered to this principle by ensuring our data analysis and presentation remained ethically grounded. By avoiding the display of specific victims' names, we mitigate the risk of dehumanizing the data or causing distress. Furthermore, our approach emphasizes the individual stories behind the statistics, aligning with Healy's advocacy for maintaining the humanistic context of data, especially in representing historical atrocities like the Holocaust.

4.2 A Humanistic Perspective on Data

Ensuring data presentation respects the memory and humanity of victims involves several key approaches. Firstly, it's essential to present data with sensitivity and context, acknowledging each statistic represents a human life lost. Using respectful language, avoiding sensationalism, and providing educational context helps honor the victims' memory. Additionally, involving historians or experts in the Holocaust in the design process can ensure accuracy and respectfulness. Lastly, offering narratives or stories alongside data, where appropriate and done with care, can help humanize the statistics, making the human cost more palpable and real.

References

Bouie, Jamelle. 2022. "We Still Can't See American Slavery for What It Was." *The New York Times*. <https://www.nytimes.com/2022/01/28/opinion/slavery-voyages-data-sets.html>.

- Bryan, Jennifer et al. 2022. *Googlesheets4: Access Google Sheets Using the Sheets API V4*. <https://CRAN.R-project.org/package=googlesheets4>.
- Chang, Winston, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson. 2022. *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- Healy, Kieran. 2020. “The Kitchen Counter Observatory.” <https://kieranhealy.org/blog/archives/2020/05/21/the-kitchen-counter-observatory/>.
- Perrier, Victor, and Fanny Meyer. 2022. *shinyWidgets: Custom Inputs Widgets for Shiny*. <https://CRAN.R-project.org/package=shinyWidgets>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org/>.
- United States Holocaust Memorial Museum. 2022. “Auschwitz Death Certificates 1941-1943.” https://www.ushmm.org/online/hsv/source_view.php?SourceId=49478.
- Wickham, Hadley. 2022. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://CRAN.R-project.org/package=ggplot2>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, and Jim Hester. 2022. *Readr: Read Rectangular Text Data*. <https://CRAN.R-project.org/package=readr>.
- Wickham, Hadley, Kirill Müller, and Romain François. 2022. *Tibble: Simple Data Frames*. <https://CRAN.R-project.org/package=tibble>.