Beyond Numbers: The Ethical Implications of Quantifying Human Tragedy*

Ethical considerations in the analysis of Auschwitz victim data

Abbass Sleiman

April 1, 2024

1 Introduction

"A single death is a tragedy; a million deaths is a statistic". This quote, generally attributed to Joseph Stalin (Investigator 2010), highlights an important, and often neglected, aspect of data analysis and statistics as a whole. When dealing with large bouts of data, it is easy to forget the underlying meaning of each data point. When discussing the implications of economic policies on unemployment rates or homelessness, what a statistician may see as a simple change of a few percentage points is in reality the profound disruption of countless families and individuals, thrust into hardship and uncertainty. When seeking to evaluate wars or conflicts, what may simply appear as numbers on both sides is in reality the deaths of countless individuals and the destruction of homes and livelihoods.

2 Ethical Considerations in Data Analysis

As individuals looking to understand the world with data, it is our responsibility not only to conduct sound statistical analyses, but to make the appropriate ethical considerations necessary when deciding how to use this data. Dropping a data point for the sake of a statistical argument is in reality omitting someone's story, a story that, depending on the context of the study, can be harsh or heartbreaking. Each decision made when conducting analysis on data should be thoughtfully evaluated. It is not enough to make statistical decisions in the pursuit arbitrary measures of statistical power, but instead each decision should be made with the aim of authentically capturing the human narratives hidden within the data.

*Code and data are available at: https://github.com/AbbassSleiman/shiny_essay. The shiny web app can be accessed at: https://wklzec-abbass-sleiman.shinyapps.io/Auschwitz_interactive_data/

Utilizing data from the United States Holocaust Memorial Museum (*Holocaust Survivors and Victims Database*, n.d.), I worked to create a shiny web app providing an interactive graph and table to allow individuals to see the victims of Auschwitz by religion, birthplace, and residence. This analysis was done through R (R Core Team 2023), along with the aid of the following packages: tidyverse (Wickham et al. 2019), shiny (Chang et al. 2024), ggplot2 (Wickham 2016), dplyr (Hadley Wickham 2023), readr (Wickham, Hester, and RStudio 2024), DT (Xie et al. 2024), and rsconnect (Chang, Allen, and RStudio 2024).

Drawing upon Bouie's insights (Bouie 2022), I approached my analysis with a deep sense of respect and responsibility, recognizing the complex historical and cultural dimensions present in the data. By acknowledging the complexities and sensitivities inherent in working with Holocaust data, my analysis aimed to uphold the dignity and memory of the victims. Hence, as part of the ethical considerations I made when making use of this data, in line with the themes brought up by Bouie, I made the decision to avoid omitting data, even if certain entries were missing or had placeholder values. This approach was twofold: it aimed to honor the memory of each individual, acknowledging the tragic fate they endured, as well as highlighting the grim reality faced by many who were not deemed worthy of proper documentation, in turn amplifying their voices and stories in our collective narrative.

Moreover, following ethical considerations, I ensured the incorporation of all available information within the dataset. This encompassed details such as names, birthplaces, residences, dates of birth and death, and religious affiliations. This approach stemmed from my commitment to honoring the individual narratives behind each data point. By embracing this methodology, my aim was to uphold the dignity of the victims and acknowledge the significance of their identities within the historical context of the Holocaust.

3 Concluding Remarks

As we navigate the delicate landscape of data analysis and ethical considerations, it becomes increasingly evident that every data point represents more than just a number. By delving into the narratives of the Auschwitz victims with sensitivity and respect, we not only shed light on the human stories behind the statistics but also honor the memory of those who suffered unimaginable atrocities. This methodology centered around the thoughtful consideration of ethical principles and a commitment to inclusivity in data representation is one that every statistician and data scientist should aim to uphold. Each statistic and every number tells a story, reminding us of the human lives and experiences embedded within the data, waiting to be heard, understood, and remembered.

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.

Chang, Winston, Jeff Allen, and RStudio. 2024. *Rsconnect: Deployment Interface for r Markdown Documents and Shiny Applications*. https://CRAN.R-project.org/package=rsconnect.

Chang, Winston, Joe Cheng, JJ Allaire, Barret Schloerke, Jeffrey Horner, Vicent Marti, and RStudio. 2024. *Shiny: Web Application Framework for r*. https://CRAN.R-project.org/package=shiny.

Hadley Wickham, Lionel Henry, Romain François. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Holocaust Survivors and Victims Database. n.d. United States Holocaust Memorial Museum. https://www.ushmm.org/online/hsv/source_view.php?SourceId=49478.

Investigator, Quote. 2010. "A Single Death Is a Tragedy; a Million Deaths Is a Statistic." https://quoteinvestigator.com/2010/05/21/death-statistic/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and RStudio. 2024. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Xie, Yihui, Joe Cheng, Xianying Tan, and Guangchuang Yu. 2024. *DT: A Wrapper of the JavaScript Library 'DataTables'*. https://CRAN.R-project.org/package=DT.