

Dealing with Holocaust Data in a Sensitive Way*

Rahul Gopeesingh

April 3, 2024

1 INTRODUCTION

The holocaust was a disturbing piece of history in which a humongous number of people were killed. At its core, the Holocaust was not just a statistic or a historical event—it was a profound assault on the very essence of humanity itself, stripping away the dignity, hope, and lives of millions of individuals.

2 SENSITIVITY WITH DATA

According to Bouie (2022), it is important to think about the ethical implications of dealing with data that came from atrocities. An parallel can be drawn from the slave trade discussed in that article as with the analysis of deaths of people at the Auschwitz camp. It is important when considering a dataset to think about the validity and accuracy of the data. These questions can be difficult to ask when dealing with a sensitive topic like this. Imagine telling someone that the document listing their ancestor as a victim of a heinous crime was not correct.

3 IMPLICATIONS FOR DATA SCIENCE

How does a data scientist approach these topics with sensitivity and without losing the meaning of the event? The mere fact that this data is available speaks to the way in which the people at Auschwitz camp were treated. They were seen as a number, a statistic and not a human

*Code and data are available at: <https://github.com/Rahul-Uoft/Shiny.git> and <https://6p27t3-rahul-gopeesingh.shinyapps.io/shiny/>

being. This is of huge concern as were the slaves according to Bouie (2022) As such we must be very careful when analyzing such data as to not diminish the historic meaning of these people's lives. There are these ethical considerations that must be made. In general, historians need to critically analyze many aspects of a dataset before accepting it and by doing must be ethically considerate whenever this is done.

4 INTERACTIVE WEBSITE

With the given dataset, I created a website in which we can see the graph of the victims at Auschwitz and we are also able to filter it based on the nationality/category of the victims. There is also a table included with lots of relevant information. This allows us to compare how the different ethnic groups were affected by this tragedy. We were able to do this with the help of R(R Core Team 2023), the shiny package (Chang et al. 2022) and the tidyverse (Wickham et al. 2019)

References

- Bouie, Jamelle. 2022. "We Still Can't See American Slavery for What It Was." <https://www.nytimes.com/2022/01/28/opinion/slavery-voyages-data-sets.html>.
- Chang, Winston, Joe Cheng, JJ Allaire, Yihui Xie, Jonathan McPherson, and Barbara Borges Ribeiro. 2022. *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- 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, 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>.