

Opinion: Data Fails to Fully Capture the Horrors of the Holocaust*

Fares Alkorani

April 3, 2024

The Holocaust, perpetrated by Nazi Germany during World War II from 1933 to 1945, stands as one of the most abominable atrocities in human history, resulting in the systematic murder of six million Jews. Recently, I came across the Sterbebücher database, one of the sources that contains personally identifiable information 65,280 Auschwitz concentration camp prisoners who were murdered between July 29, 1941 and December 31, 1943. More specifically, this database noted the “last name, first name(s), date of birth, date of death, birthplace, residence, and religion” of prisoners (U. S. H. M. Museum, n.d.).

Using the R programming language (R Core Team 2023) and the Shiny package (Chang et al. 2023), I created an [interactive application](#) that allows people to visualize the data in the Sterbebücher data set. I began using my own application and to my surprise, I felt a bit uncomfortable. Looking strictly at the dataset, we see that there is a limited amount of information regarding the deaths of the prisoners in the Auschwitz concentration camp. Sure, the names of prisoners are captured, along with some other demographic information - so we know a bit more about the individual beyond them being a number contributing to the death toll in the Holocaust. However, this data set fails to capture the individuality of each prisoner - their experiences, their trauma, and more importantly, their humanity. Their birthplace, residence, and religion can only unveil so much about the attitudes in the Holocaust.

In addition, to put things into perspective, the severity of the holocaust is understated by using a dataset like this. The death certificates only covered only registered prisoners from the July 1941 to the end of 1943 as mentioned above, but “overwhelming majority of victims, mostly Jewish, perished in the gas chambers immediately after arrival, without being entered in the camp records, and without their deaths being noted in the German documents” (A.-B. S. Museum, n.d.). To put things into perspective, one can only visualize the data of 65,280 Auschwitz concentration camp prisoners, approximately one percent of the total six million murdered Jewish people during the Holocaust. For someone using this interactive application

*Code and data are available at: https://github.com/faralk/auschwitz_analysis. Shiny application available at: https://dxhyici-fares-alkorani.shinyapps.io/auschwitz_app/

without prior knowledge of the Holocaust, it would severely underestimate the true impact of the holocaust. However, in theory, even if all of the murdered Jewish people murdered in the holocaust were added to a dataset, this would only fix the issue so much.

As a computer science and statistics student, I feel compelled to advocate for the visualization of data through applications as such - which leverage both computer science and statistics to a high degree. However, interacting and the aggregation of data for topics like this can be dehumanizing and take away from the true horrors of the genocide.

In the Holocaust, other genocides, and even other sensitive events like this, data should always be supplementing the experiences and stories communicated by the survivors of these events. Numbers may paint a general picture, but numbers are not tell all: they will never capture the specific actions undertaken by the perpetrators in the genocide, the level of suffering for each individual, and the underlying attitudes that carry forward from events like the Holocaust to this day.

References

- Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2023. *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- Museum, Auschwitz-Birkenau State. n.d. “Sterbebücher.” <https://www.auschwitz.org/en/museum/about-the-available-data/death-records/sterbebucher/>.
- Museum, United States Holocaust Memorial. n.d. “Holocaust Survivors and Victims Database.” https://www.ushmm.org/online/hsv/source_view.php?SourceId=49478.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.