An Analytical Report on Auschwitz Victims*

Xiyou Wang

2024-04-04

This report presents an analysis of data on the victims of Auschwitz, categorized by nationality and other demographic factors. The findings highlight the disproportionate impact on certain groups, with a focus on the number of deportees, victim percentages within categories, and the overall percentage of all victims.

Contents

1 Introduction

The Auschwitz concentration camp is an enduring symbol of the Holocaust's horrors. Between 1940 and 1945, it was the site of the largest mass murder in a single location. The data analyzed in this report reflects the nationalities and categories of the individuals who were deported to Auschwitz and those who became victims within its fences.

---

*Code and data are available at: https://github.com/wxywxy666/Auschwitz.

The Shiny app can be accessed ["outputs/app.R].

2 Data

The core of this report revolves around a detailed examination of a dataset that enumerates the victims of Auschwitz, broken down by nationality/category. This dataset is a historical compilation of records that quantifies the magnitude of human loss during one of history's darkest periods.

2.1 Data Cleaning

The dataset provided, titled "Auschwitz_data.csv," comprises six columns, each representing a different variable:

1. **Nationality/Category**: This categorical variable lists the identified groups of deportees, classified by either nationality or social classification used by the Nazi regime. Number of Deportees: A numeric estimate of individuals deported to Auschwitz within each category, expressed in text format with qualifiers such as "thousand" or "million" for ease of human readability. Percentage of the Total Number of Deportees: The proportion of each nationality/category relative to the total deportee population, reported in percentage terms.
2. **Number of Deportees** The estimated count of those who did not survive, formatted similarly to the "Number of Deportees" and including approximate indicators. Percentage of Murdered within the Category/Nationality: The victim rate within each group, indicating the lethality of Auschwitz's environment for that segment.
3. **Percentage of the Total Number of Deportees**: This column reflects each group's share of the total number of victims at Auschwitz. Upon importing the raw data into the R environment, we confronted the challenge of converting the human-readable numerical estimates into a format suitable for computation. The gsub() function from R's base package was employed to replace text-based quantifiers—transforming "thousand" into "e3" and "million" into "e6"—and to remove approximation indicators such as the tilde ("~"). The transformed strings were then converted into numeric data types to facilitate accurate calculations and visual representations.

2.2 Methodology

The initial stage of the analysis involved data cleansing and validation. The accuracy of numeric conversions was checked manually to ensure consistency with historical records. In instances of missing data or where approximations were given, conservative estimates were used to avoid overstatement of figures.

Subsequently, exploratory data analysis (EDA) was carried out to understand the distributions and relationships within the dataset. Using R and its associated packages (ggplot2 for visualizations and DT for interactive tables), we generated graphs to visualize the numbers and proportions of deportees and victims, allowing for a comparative analysis across the different groups.

The analysis was both quantitative and interpretative. Quantitative methods included calculating sums, percentages, and cross-tabulations to elucidate the relative impact on each nationality/category. The interpretative aspect involved contextualizing the figures within the broader historical narrative of World War II and the Holocaust, drawing on secondary sources for supplementary information where necessary.

Ethical considerations were paramount in this methodology. Data related to human loss requires a respectful approach that acknowledges the individuals behind the numbers. The interpretation of data sought not only to inform but also to honor the memory of the victims. The analysis was therefore conducted with a commitment to accuracy, sensitivity, and the educational value of remembrance.

3 Results

Demographics of Deportees Over 1.3 million individuals were deported to Auschwitz, with the majority being Jewish, followed by Poles, Roma, and Soviet POWs. This demographic distribution reflects the targeted persecution policies of the Nazi regime.

Victim Numbers and Percentages Approximately 1.1 million Jews were deported, with an estimated 90% murder rate, signifying the specific genocidal intent against Jewish populations. Poles, the second-largest group, had a lower percentage of deaths, suggesting a different pattern of persecution.

Comparative Analysis The analysis reveals stark contrasts between groups, shedding light on the varied experiences and fates within Auschwitz, influenced by Nazi racial policies.

4 Discussion

The data analysis provides a quantitative reflection of the tragedy of Auschwitz. It underscores the systematic nature of the Holocaust and the scale of its impact on different communities. The statistics are not mere numbers; they represent individual lives subjected to extreme brutality and are a testament to the atrocities of genocide.

5 Conclusion

The report concludes that the analysis of Auschwitz victims by nationality/category is more than a statistical overview; it's a necessary act of remembrance. As we delve into the depths of these numbers, we honor the memories of those who suffered and perished. Their stories, aggregated into data points, implore us to remember and educate, ensuring the permanence of "Never Again."

6 References