Data analysis of Holocaust Victims*

Shiny app applied for interactive graph

Yingzhi Zhang

April 3, 2024

1 Introduction

Auschwitz concentration camp, operational from 1940 until 1945, stands as a grim symbol of the Holocaust and the atrocities committed by Nazi Germany during World War II. Initially intended as a detention center for political prisoners, Auschwitz evolved into a network of camps where millions were murdered under inhumane and cruel methods, including gas chambers or succumbed to starvation, forced labor, disease, and medical experiments. It was the largest of the German Nazi concentration camps and extermination centers, where approximately 1.1 million people were killed.(Piper 2000) The camp's liberation by Soviet troops in January 1945 revealed the extent of the Nazi regime's brutality and had a profound impact on global awareness of human rights atrocities, leading to the establishment of the Genocide Convention and the Universal Declaration of Human Rights. The Holocaust's sheer scale prompted a reevaluation of international laws regarding human rights and genocide, significantly influencing post-war international relations and the establishment of organizations committed to preventing such atrocities.

The victims of Auschwitz were primarily Jews, who constituted about 90% of the deaths, making it a central site of the Nazis' "Final Solution to the Jewish Question."(Hayes 2003) However, the camp also held Poles, Romani people, Soviet prisoners of war, and prisoners of various other ethnicities and nationalities. The ethnic diversity of the victims underscores the racially motivated ideology of the Nazi regime, which sought to exterminate what it considered "undesirable" groups. The targeting of Jews, in particular, was driven by deeply ingrained anti-Semitic beliefs, which aimed at their total eradication. The destruction of these communities left lasting scars on the surviving populations and profoundly affected the demographic and cultural landscapes of Europe and beyond.

In the following part of this essay, I will analyze the number of victims in Auschwitz concentration camp focusing the perspective of ethnic group.

---

*Code and data are available at:https://github.com/melissazhangyz/holocaust_victim. Interactive graphs are available at: https://melissazhang.shinyapps.io/holocaust_victim/

2 Data

A dataset of victim list is obtained from the official website of United States Holocaust Memorial Museum. The title of the date set is "AUSCHWITZ DEATH CERTIFICATES 1941-1943 (ID: 49478)". The raw data contains 7 columns and 65280 observations, recording the death book of Auschwitz Concentration Camp prisoners who passed away from July 29, 1941 to December 31, 1943. As an electronic transcript of the original death book, the dataset may not reveal the total number of death due to a lost of data or intentional omission due to political concerns.

When looking at the data of victims in the Holocaust, we should always remind ourselves that they are HUMAN.(Bouie 2022) When conducting data analysis to historical events, we may easily fall in the trap of dehumanization if we only consider data. It is important that we should always remember each observations in this dataset represent an individual who was murdered at that time. Therefore, when we try to analyze and understand the number of victims in the Auscheitz camp, we should pay attention to their ethnic group.

For the data cleaning procedures, I use the programming language R (R Core Team 2023) and packages "tidyverse" (Wickham et al. 2019), "lubridate" (Grolemund and Wickham 2011) and "kableextra" (Zhu 2021). Then I use "Shiny" (Chang et al. 2023) and "DT" (Xie, Cheng, and Tan 2024) to produce interactive graphs and tables for better illustration and data visualization. After initial cleaning of data, I also mutate two new columns to calculate the age when the victim died and the Year_Month that they died for next step. For the Religion column, I combined some similar religions to receive only: Jew, Protestant, Catholic, Orthodox, Buddhist, Hussite, Muslim, Atheist, Other and Unkown type. Finally, a cleaned version of the dataset is shown as Table 1.

Last_Name	First_Name	Date_of_Birth	Date_of_Death	Birth_Place	Residence	Religion	Age	Death_Year_Month
Aandagt	David	1906-04-26	1942-09-01	Amsterdam	Amsterdam	Jew	36	1942-09
Aandagt	Hartog	1899-02-10	1942-10-27	Amsterdam	Amsterdam	Jew	44	1942-10
Aandagt	Israel	1909-05-25	1942-10-10	Amsterdam	Amsterdam	Jew	33	1942-10
Aandagt	Jacob	1918-03-10	1942-10-08	Amsterdam	Amsterdam	Jew	25	1942-10
Aandagt	Jacob	1922-01-26	1942-08-22	Amsterdam	Amsterdam	Jew	21	1942-08
Aandagt	Meyer	1914-05-31	1942-07-28	Amsterdam	Amsterdam	Jew	28	1942-07
Aandagt	Mozes	1894-11-15	1942-09-04	Amsterdam	Amsterdam	Jew	48	1942-09
Aardewerk	Alex	1921-04-02	1942-08-24	Amsterdam	Amsterdam	Jew	21	1942-08
Aardewerk	Esther Wilhelmina	1921-12-03	1942-09-13	Amsterdam	Amsterdam	Jew	21	1942-09
Aardewerk	Marcus	1921-06-12	1942-08-11	Amsterdam	Amsterdam	Jew	21	1942-08

Table 1: First ten rows of the cleaned data

3 Results

We can first take a look at the ethnic group distribution of the whole dataset. As shown in Figure 1, Catholic takes the largest proportion where 29871 of the total victims are Catholic. Jews, which has long been recognized as the main purpose of the Holocaust, follows to be the second largest proportion as

27980 victims are Jews. Other religion groups have much fewer victims. Though Atheist, Othodox and Protestant still have more than 1000 victims, Buddist, Muslim and Hussite have only 1 or 2 victims.
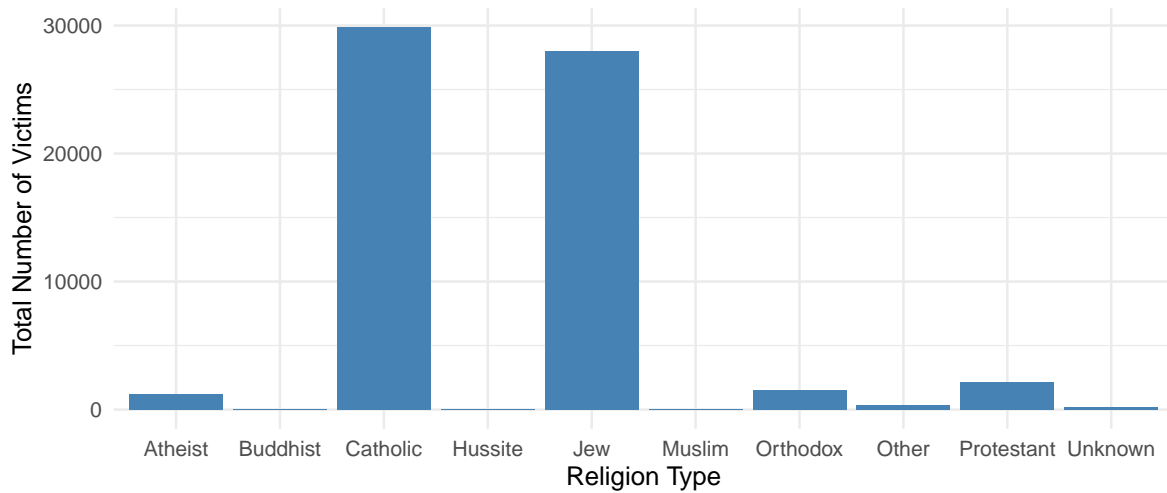

Figure 1: Number of total victims by ethnic group

Within each ethnic group, we can obtain the death number in each month. As shown in Figure 2, murder to the Catholic group in the Holocaust varies over the time. Most were killed in Febuary and March, 1943. This may because the rebut to German by English and French army and the expansion of war in the European area. On the other side, most Jews in Auschwitz died in August and September, 1942. Combining the information of both group, we can take a guess that the Holocaust targeting on Jews were comitted and finished in the fall of 1942. Later victims in the camps varied more in their ethnicities.
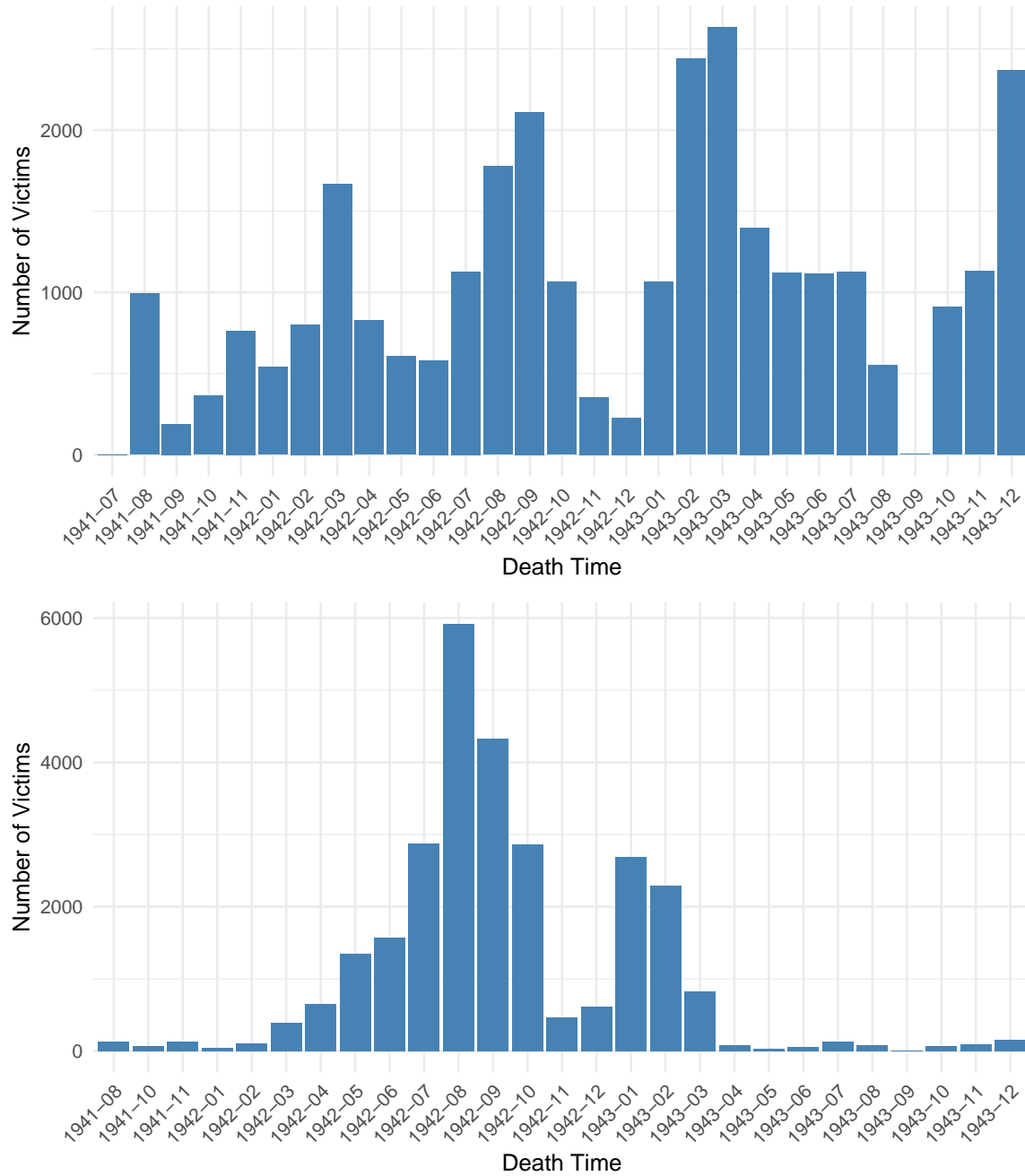
Figure 2: Catholic Victim by month

To better understand the victims, we can also review the age. Using the birth year and death year recorded in the dataset, we can calculate the age when they passed away. In particular, we control for

those age less than 100 to avoid bias of data due to errors. Using box plot in Figure 3, we can see that age variation is similar between Catholic and Jew group. The average age is around 37. Most victims are middle age adult from 25 to 45. More children and babies were murdered in the camp in Catholic group than Jew.
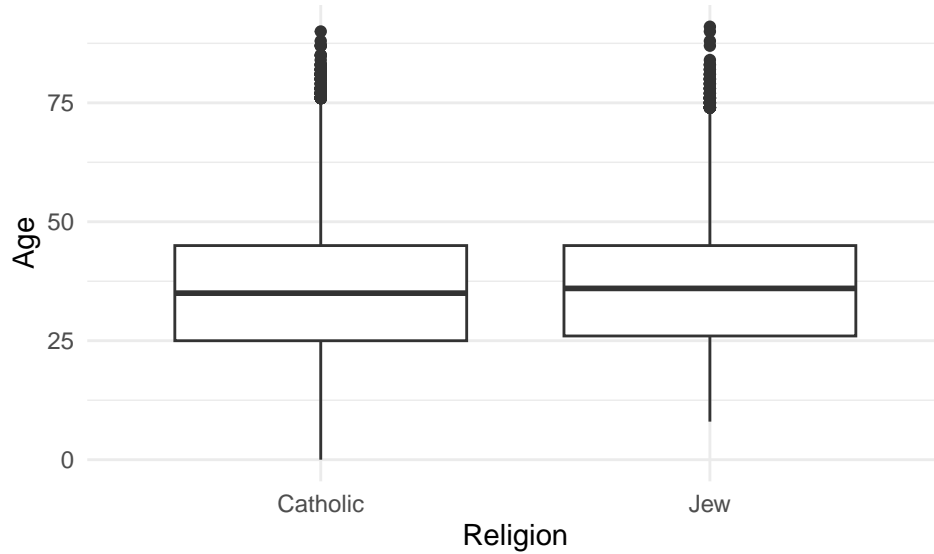

Figure 3: Age of Jew victims

For better visualization of the number of victims by ethnic groups, an interactive graph is published through Shiny.io at the following link:https://melissazhang.shinyapps.io/holocaust_victim/.

References

Bouie, Jamelle. 2022. "We Still Can't See American Slavery for What It Was." *International New York Times*, NA–.

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.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Hayes, Peter. 2003. "Auschwitz, Capital of the Holocaust." *Holocaust & Genocide Stud.* 17: 330.

Piper, Franciszek. 2000. "Długoborski, Wacław; Piper, Franciszek." *Auschwitz, 1940–1945. Central Issues in the History of the Camp* 3: 978–8385047872.

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.

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

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.