

Project Assignment M6 – Final Report

Submitted By: Justin Ehringhaus and Lauren Todaro

ALY6015 | Fall A 2022

Roux Institute, Northeastern University

Prepared For: Dr. Steve Morin

October 29, 2022

Introduction

Though the details of this analysis are sensitive in nature, we have a shared interest in education, racial equality, and universal access to healthcare.

The institutional correctional facility system—specifically death row—is fraught with unique legal, clinical, and ethical complexities. Specifically, on death row, there are high rates of mental illness in incarcerated individuals prior to capital crime and significant trauma associated with confinement and awareness of impending execution. Mental health professionals serving incarcerated individuals routinely perform psychiatric assessments including, but not limited to: treatment for competency to be executed, “justifiable” depression, common disorders, types of therapies and medications, and use of religious figures.

As such, the **goal** of this project is to understand the relationships and dependencies between the attributes of individuals on death row including demographic (*Race*, *Level of Education*), case-specific (*Age When Received*, *Age When Executed*, *Years on Death Row*, *Total Number of Victims*) and execution-related (*Last Statement*) variables to inform psychiatric policies and psychological support for impacted parties, including families of incarcerated individuals and victims.

The area of focus and, therefore, our **questions** studied will help to inform psychiatry programs and policies for death row facilities. Additionally, we hope this will provide insights into the needs of psychological services for families of incarcerated individuals and victims, as well as support for the correctional officers who provide care and oversee executions. The questions included in the report include the following:

1. Is there a relationship between the *Race* (independent, nominal variable) of inmates and the *Length of Time on Death Row* (dependent, continuous variables)?
 - **Connection to Goal:** We were interested in understanding if there are significant differences in the treatment of individuals based on race, specifically, the length of time spent on death row. This may help to understand if there is potentially any bias or preferential treatment.
 - **Method Used:** Kruskal-Wallis Test
2. Is there a difference in *Age at Time of Execution* between Inmates with *Less Education* (i.e., 2 years of high school or less) versus those with *More Education* (i.e., 3 years of high school or more)?
 - **Connection to Goal:** We chose this question to better understand, from a statistical perspective, if there is a difference in *Age at Time of Execution* based on inmates educational levels. For this analysis, we binned *Level of Education* into two (2) categories, creating the most even split possible to compare those with two (2) years of high school or less versus those with three (3) years of high school or more. This information could inform psychiatric policies to ensure appropriate levels of care for individuals across a range of ages and education levels.
 - **Method Used:** Wilcoxon Rank Sum Test
3. Is there a relationship between the *Total Number of Victims* by *Sentiment of Last Statement*?
 - **Connection to Goal:** We selected this question to better understand the state of mind of individuals on death row on the day of execution (*Last Statement*) as it relates to sample means of the *Total Number of Victims*. For this analysis, we manually reviewed and binned *Last Statement* into fifteen (15) categories of emotion. We seek to understand if there are any correlations between the mental state and/or outlook of individuals based on the sample’s mean total number of victim(s).
 - **Method Used:** Kruskal-Wallis Test

This project utilizes the Texas Department of Criminal Justice death row dataset, which includes profiles of each inmate who is on, or has been on, death row from December 7, 1982 until the present. The **1,119** rows represent people who have been convicted for murder and sentenced to death, and the **20** variables contain ages, genders, races, heights, weights, counties, educational levels, counts of victims by gender, and last statements before execution. The code utilized to conduct the analysis and results is included in the accompanying R file.

Methods

As noted in the introduction, we utilized two methods for analysis: the Kruskal-Wallis Test and the Wilcoxon Rank Sum Test, using $\alpha = 0.05$ in all cases.

Prior to learning about non-parametric methods, we attempted to utilize the Chi-Square Test of Independence and ANOVA. Upon further assessment, we came to understand that, given the lack of normality in the distribution of our data, a difference in variances, and/or a difference in sample sizes, we did not meet the required assumptions. Though parametric methods are typically more accurate and carry higher power, data is assumed to have come from a normal distribution which is not appropriate for our dataset. Hence, the following non-parametric methods are best-suited for this study.

We utilized the **Kruskal-Wallis Test** for Q1 and Q3. This test allows us to compare three or more samples to determine whether or not they originated from populations with the same distribution. In the case of Q1, we test the association between the *Length of Time on Death Row* by *Race*. In the case of Q3, we test the association between the *Total Number of Victims* by the *Sentiment of the Last Statement*. Significance would point to correlations between these variables, but this test would not reveal which specific groups are statistically and significantly different from one another. Nonsignificance, on the other hand, would reveal that there is no evidence of a correlation to be found given the respective test statistics and alpha cutoff values.

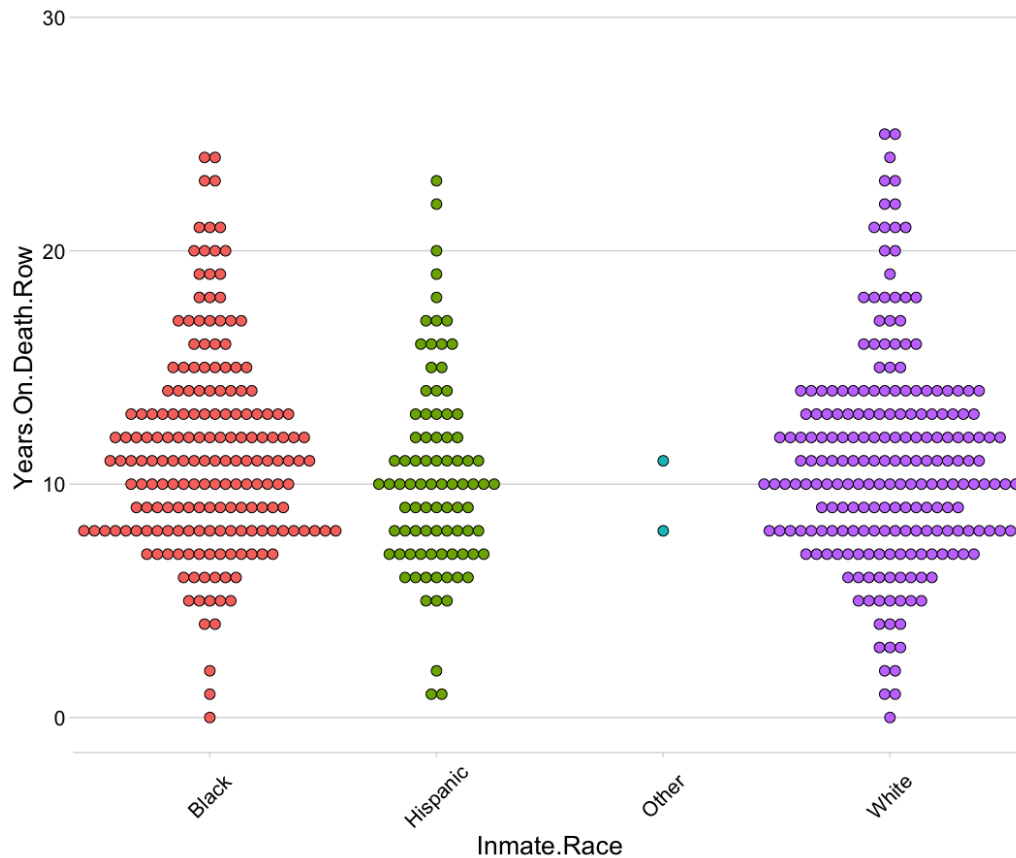
We utilized the **Wilcoxon Rank Sum Test** for Q2. This test uses ranks to determine whether two independent samples originate from populations with the same distribution. In this case, we binned *Level of Education* into two (2) categories, representing those with *Less Education* (i.e., two years of high school or less) and those with *More Education* (i.e., three years of high school or more), to compare this to *Age at Time of Execution*. Significance would point to there being a difference in *Age at Time of Execution* depending on *Level of Education*.

Detailed analysis as follows.

Analysis

In this section, we consider visual prompts from our exploratory data analysis followed by analysis and interpretation of statistical findings to the three (3) questions that will help to reach our goal of understanding the relationships and dependencies between the attributes of individuals on death row to inform psychiatric policies and psychological support for impacted parties. Within each question outlined below, we discuss the method used, our justification for selecting it, and our interpretation of the results.

Figure 1. Years on Death Row by Inmate Race
Years on Death Row by Race



Question 1. Is there a relationship between the *Race* (independent, nominal variable) of the inmate and the *Length of Time on Death Row* (dependent, continuous variables)? Each dot in Figure 1 represents a person in the dataset, categorized by race on the x-axis and positioned by years on death row along the y-axis. Interestingly, each population appears approximately normally distributed with a slight right-skew. Although they visually look similar, statistical testing can determine whether there is a statistically significant difference between any of the variables.

Test Type: Kruskal-Wallis Test at $\alpha = 0.05$

Method Justification: Although populations were approximately normally distributed, variances of populations were unequal, and thus we selected the non-parametric Kruskal-Wallis Test over the parametric ANOVA test. We have one categorical variable (*Race*) and one continuous variable (*Length of Time on Death Row in Years*).

- H₀: There is no relationship between the *Race* of inmate and *Length of Time on Death Row*.
- H₁: There is a relationship between the *Race* of inmate and *Length of Time on Death Row (claim)*.

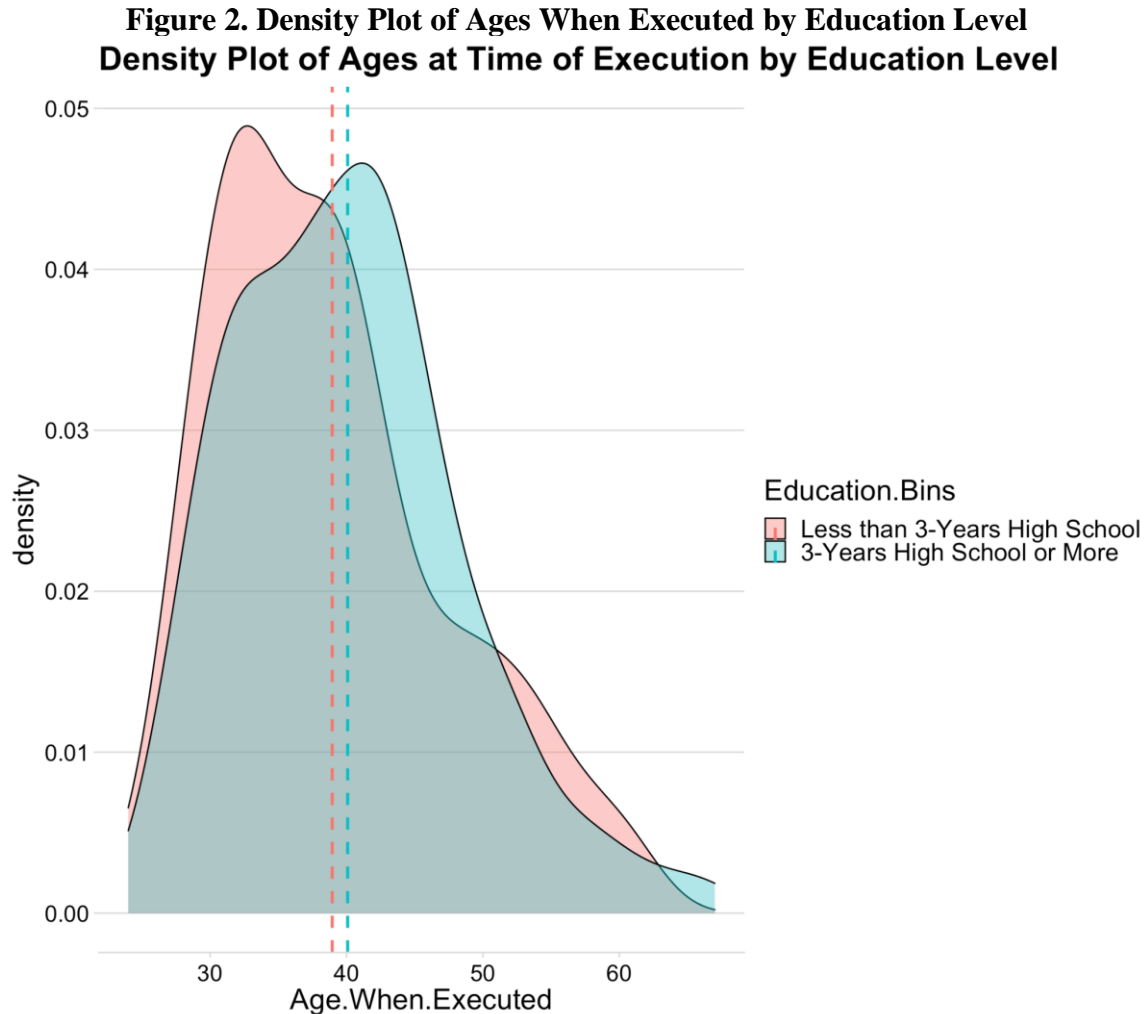
Kruskal-Wallis Chi-Squared: 2.3447

Test Statistic (p-value): 0.504

Decision: Do not reject the null hypothesis.

Summary and Interpretation: There is not enough evidence to support the claim that there is a relationship between the *Race* of the inmate and the *Length of Time on Death Row in Years*. Though there is not sufficient evidence to support our claim, there is further research to be done on potentially biased treatment of inmates

based on race. There are, of course, additional factors—such as severity of crime, number of victims, repeat offenses—which could be explored further to check for correlations.



Question 2: Is there a difference in *Age at Time of Execution* between inmates with *Less Education* (i.e., two years of high school or less) versus those with *More Education* (i.e., three years of high school or more)? The density plot in Figure 2 positions *Age at Time of Execution* along the x-axis and densities along the y-axis, comparing those with *Less Education* (red) and those with *More Education* (blue) as defined by the number of years of education. The dashed lines represent the mean *Ages at Time of Execution* for each group. Interestingly, the sample means are quite similar, but the skewness of the distributions differ. Statistical testing can determine whether there is a difference in the distributions of the populations from which these samples came.

Test Type: Wilcoxon Rank Sum Test, $\alpha = 0.05$

Method Justification: Populations were not normally distributed, and thus we selected the non-parametric Wilcoxon Rank Sum Test as we have one categorical variable binned into two groups (*Education Level*, those with *Less Education* versus those with *More Education*) and one continuous variable (*Age When Executed*).

- H₀: There is no difference in the *Age at Time of Execution* for inmates with *Less Education* versus inmates with *More Education*.
- H₁: There is a difference in the *Age at Time of Execution* for inmates with *Less Education* versus inmates with *More Education* (*claim*).

Test Type: Kruskal-Wallis Test at $\alpha = 0.05$

Method Justification: Populations were not approximately normally distributed, and thus we selected the non-parametric Kruskal-Wallis Test over the parametric ANOVA test. We have one categorical variable (*Sentiment of Last Statement*) and one continuous variable (*Total Number of Victims*).

- H0: There is no relationship between the *Sentiment of Last Statement* and *Total Number of Victims*.
- H1: There is a relationship between the *Sentiment of Last Statement* and the *Total Number of Victims* (*claim*).

Kruskal-Wallis Chi-Squared: 23.81

Test Statistic (p-value): 0.0483

Decision: Reject the null hypothesis.

Summary and Interpretation: There is enough evidence to support the claim that there is a relationship between the *Sentiment of Last Statement* and *Total Number of Victims*. These insights could provide significant value in understanding and/or estimating the state of mind of future incarcerated individuals on death row so they can receive appropriate treatment as it pertains to their likely state of mind. Although this test does not isolate the specific groups that are significantly different from one another, further research could be done to determine the way in which *Total Number of Victims* correlates with the *Sentiment of Last Statement*.

Interpretations and Conclusion

Based on our analysis, we have found that there is significant value in continuing to explore relationships and dependencies between the attributes of individuals on death row to inform psychiatric policies and psychological support for inmates and impacted parties. The findings summarized below inform policies to ensure the comprehensive treatment of individuals on death row, as well as potential opportunities to support families and correction facility staff.

Findings and Recommendations:

- 1) Although there is no statistical evidence to support the claim for Q1, a nonsignificant result does not prove that a correlation does not exist. A disadvantage of non-parametric testing is larger differences are required to reject the null hypothesis. If able to meet the assumptions of parametric testing by normalizing our data through random sampling (i.e., the assumption of normal distribution) and/or by collecting more data through further aggregation of death row statistics from other states (i.e., the assumption of equal variances), we could apply the parametric versions of these tests to obtain more accurate, sensitive results.
- 2) Although there is statistical evidence to support the claims for Q2 and Q3, a significant result marks the beginning for future research. Isolating the exact reasons and/or predictor variables that make the result significant would be the next step, and then using that understanding to improve psychiatric policies and psychologist support systems for inmates and/or impacted parties would be the ultimate goal.
- 3) Additional data, such as type and execution of crime, presence of pre-existing mental illness, history of familial involvement in crime, and previous occupation, may not only lend to a more in-depth understanding of how to best support inmates and related parties but may help predict (and potentially even prevent) individuals with a similar statistically significant predisposition from engaging in similar crimes in the future.

Thank you for the opportunity to apply our learnings from ALY6015 in a meaningful, real-world example.

References

- Bluman, A. G. (2018). *Elementary Statistics: A Step by Step Approach* (10 ed.). New York: McGraw-Hill Education.
- Grassian, S. (1983, November 11). Psychopathological effects of solitary confinement. New Mexico Legislature. Retrieved October 23, 2022, from <https://www.nmlegis.gov/handouts/CCJ%20102716%20Item%203%20Dr%20Grassian%20Psychopathological%20Effects%20of%20Solitary%20Confinement.pdf>
- Hamilton, S., & Goodwin, R. (2003, February 15). Lifetime comorbidity of antisocial personality disorder and anxiety disorders among adults in the community. *Psychiatry research*. Retrieved October 23, 2022, from <https://pubmed.ncbi.nlm.nih.gov/12606017/>
- Haney, C. (2003, January). Mental health issues in long-term solitary and “Supermax” confinement. *Sage Journals*. Retrieved October 23, 2022, from <https://journals.sagepub.com/doi/abs/10.1177/0011128702239239>
- Hoskin, T. (n.d.). Parametric and nonparametric: Demystifying the terms. Mayo Clinic. Retrieved October 23, 2022, from <https://www.mayo.edu/research/documents/parametric-and-nonparametric-demystifying-the-terms/doc-20408960>
- Ianmobbs. (2017, June 9). Texas death row executions info and last words. Kaggle. Retrieved September 29, 2022, from <https://www.kaggle.com/datasets/ianmobbs/texas-death-row-executions-info-and-last-words>
- Jones SJ, Beck E. (2006-2007). Disenfranchised grief and nonfinite loss as experienced by the families of death row inmates. *Omega (Westport)*. Retrieved October 26, 2022, from <https://pubmed.ncbi.nlm.nih.gov/18186424/>
- Long, W. (n.d.). Trauma therapy for death row families. *Journal of Trauma and Dissociation*. Retrieved October 23, 2022, from <https://pubmed.ncbi.nlm.nih.gov/21967176/>
- Nguyen, M. K. (2017, December 31). Last words of death row inmates. Kaggle. Retrieved September 29, 2022, from <https://www.kaggle.com/datasets/mykhe1097/last-words-of-death-row-inmates>
- Texas Department of Criminal Justice. (n.d.). Death Row Information. Retrieved September 29, 2022, from https://www.tdcj.texas.gov/death_row/index.html
- Yanofski, J. (2011, February). Setting up a death row psychiatry program. *Innovations in clinical neuroscience*. Retrieved October 23, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3071089/>