Echoes of Fate: Analyzing Demographics and Last Words of Texas' Death Row Inmates

Group E

Digital Humanities: Tools and Methods

Dr. Federico Pianzola & Dr. Lucas van der Deijl

January 18, 2024

Gen Daquinan (s5722926)

Jennifer Dijkstra (s4117840)

Linchun Hou (s5484197)

Sijie Qiu (s5686865)

Yuhan Si (s5495415)

Xinyi Song (s5705258)

Yuning Xie (s5743508)

Table of Contents

Introduction

The death penalty, a subject of enduring controversy and profound significance within the American legal system, is especially pertinent in Texas, where the use of capital punishment is notably prevalent. Since the death penalty's reinstatement in 1976, Texas has surpassed all other states in the number of executions, cementing its position at the forefront of this debate (Death Penalty Information Center, 2019). Remarkably, 79.8% of these individuals offered final statements, providing a unique and potent insight into their thoughts and emotions, according to our calculations. This paper aims to dissect and understand these final words within the broader context of each inmate's demographic background. We pose the central research question: "What insights into the Texan criminal justice system can be derived from an analysis of demographic data and the last utterances of death row inmates?"

Our project employs a comprehensive methodology, integrating a variety of analytical tools to derive humanistic insights from data collected on Texas death row inmates. We scraped the data from the Texas Department of Criminal Justice (TDCJ) official website. The dataset 'last_statement_of_death_row.csv' contains detailed information about the inmates and their last statements from 1982 to 2023. It includes data such as TDCJ Number, Age, Execution Date, Race, County, availability of a Last Statement (Yes/No), the actual Last Statement text, as well as Date of Birth, Date Received, and Age at Reception. These elements were scraped using BeautifulSoup, a web scraping package for Python. However, for the extraction of Education Level, Date of Offense, and Prior Occupation, manual data collection methods were employed, as these variables were presented in photographic form on the website and were not conducive to automated web scraping techniques. This dual approach of data collection ensured a comprehensive and accurate dataset for our analysis.

We employed a range of tools to explore the data and text. This includes the use of Voyant for text analysis, ArcGIS for geographical distribution examination, and visualization tools like Flourish to elucidate demographic patterns, educational backgrounds, and so on. By correlating these elements, we seek to unveil the intricate interplay between societal structures and individual narratives, thus shedding light on the multifaceted nature of the Texas criminal justice system and its ultimate penalty.

The Analysis of Inmate Demographics

In this part, we use ArcGIS and Flourish to explore the inmates' data. We researched the geographic distribution of inmates, the composition of inmates, their age, race, educational background, previous occupation, as well as the length of their stay on the death row.

Geographic Analysis

When it comes to understanding crime, people often refer to statistics. While this is a perfectly fine way to show how crime has affected communities and people, graphs truly make these statistics come to life. Graphs help visualize the impact crime has on these aforementioned communities. For example, it is one thing to mention how many people go missing in a year, it is another to have a heat map that highlights the amount of people and where they have gone missing. Similarly, this research project wishes to highlight the importance of geographical data in relation to crime, as it can tell us quite a lot.

Therefore, it is important to consider why these violent offenders come to commit such crimes in the first place, and whether this might have something to do with their geographical

upbringing. As such, this section of the report will show how we used geographical data to create a couple of maps which paint several pictures of violent offenders in Texas. While there is a possibility that certain counties are more likely to sentence serious offenders to death, none of the data which has been collected makes any such implication. As such, the assumption has been made that the offenders in our dataset were sentenced to death based on the severity of their crime(s), not because of a predetermined bias in the Texas judicial system.

Lastly, before getting into the data, it is important to highlight how these maps were rendered. The data used to look at the county of origin for each offender were taken from the "Death Row Information" provided by the Texas Department of Criminal Justice (2023), and the 2020 census data for population and median household income were provided by the U.S. Department of Agriculture (2023). Moreover, the data was visualized using ArcGIS. The primary symbology for visualizing the upcoming maps uses 8 classes with a gradient color scheme. Most notably, the data is presented using geometric intervals rather than equal intervals. This choice was made to better represent the data visually.

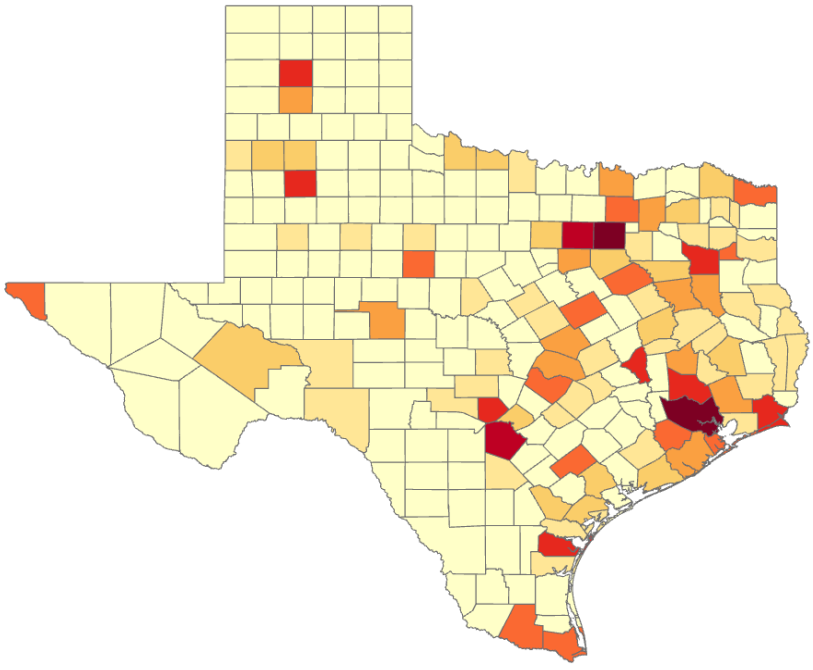


Figure 1. Number of Death Row Inmates in each Texas county

Figure 1 highlights the number of death row inmates from each Texas county. This map shows the counties of Dallas and Harris to have had the most death row inmates. While this provides a nice overview, the question remains as to why these counties in particular are more prone to have more violent offenders worthy of the death penalty than others, so to speak. Thus, it would be disingenuous to present a map solely based on the number of death row inmates in any given county, as this doesn't take into consideration the population nor economic affluence of a county.

In order to counteract this, the following figures use the previously referenced "death row inmate per county" data, as well as the census data from 2020. They normalize the number of inmates against the 2020 population size of each county as well as the 2020 median income of each county. Thus, both maps will highlight the outlier counties.

Figure 2 shows us that when comparing the number of death row inmates to the population size of the counties, certain counties are outliers. For example, Crockett County, which is the darkest county on the left-hand side, is an outlier because a single death row inmate had originated from there. While it is but one person, it is a lot in comparison to the only 3098 residents that lived in that country. Similarly, Dallas, which was highlighted prevalently in the very first figure, is now depicted as in the middle of the road, despite the fact that Dallas had 65 death row inmates. This is because Dallas' population is much higher, and thus the amount of death row inmates is "normal" for a county of that population size.

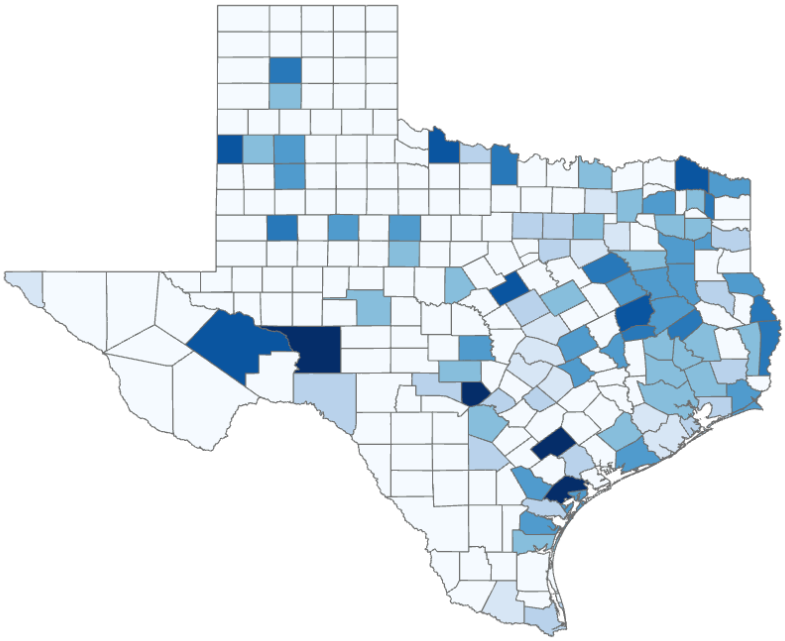


Figure 2. Death Row Inmates Ratios to County Population

Figure 3 follows this similar pattern: the death row inmate that is normalized against the median income of each county. When looking at Harris (the darkest county in the bottom right), one can see its median income is 61906, with 133 death row inmates. That doesn't just stand out in the map, but it does stand out in relation to the previous two maps. The first figure — solely based on inmate appearance — shows Dallas and Harris to have the most death row inmates. This coincides with the map that explores the median income.

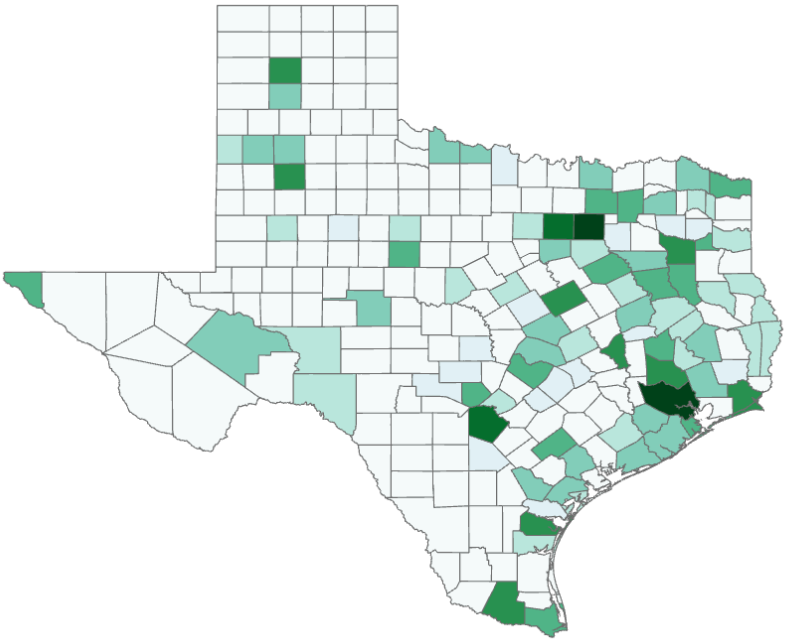


Figure 3. Death Row Inmates Ratios to the Median Income of Each County

Based on this preliminary research, we could argue that both population and income are part of the reason. High-income counties do seem to have more death row inmates. Although, an argument could be made that major metropolitan areas tend to have higher income in general. This, however, does raise another question. It is generally assumed that low-income areas are more prone to crime, but the number of death row inmates reflect a different story. A possible explanation might be that higher income counties are simply able to afford the death penalty more often. While it is often assumed that ending someone's life is more cost effective than keeping them alive for the unforeseeable future, the death penalty is extremely expensive. As per the Death Penalty Information Center (2023):

> Some of the reasons for the high cost of the death penalty are the longer trials and appeals required when a person's life is on the line, the need for more lawyers and experts on both sides of the case, and the relative rarity of executions. Most cases in which the death penalty is sought do not end up with the death penalty being imposed. And once a death

sentence is imposed, the most likely outcome of the case is that the conviction or death sentence will be overturned in the courts. Most defendants who are sentenced to death essentially end up spending life in prison, but at a highly inflated cost because the death penalty was involved in the process.

While it might seem that the convicted person would have to bear the brunt of these costs, but instead the county tends to pay most — if not all — of the costs associated with executions. Estimations about the final cost of a death sentence tend to range from 2 million to 7 million dollars, which can be up to 10 times more than life imprisonment (Baicker, 2001, 4). In Texas itself, Jasper County has claimed it raised property taxes by 8% to pay for the trial of three men in relation to a high-profile murder. Yet another Texas county allegedly tried to raise taxes for a similarly high-profile trial; instead, taxpayers voted for a tax rollback, causing the county to cut funding to the fire and ambulance services in the county (5). It is no wonder then that higher income counties have more instances of the death penalty being given, as lower income counties might simply not be able to afford sentencing someone to death. As such, geographical data provides an interesting perspective, but cannot be used to conclusively say whether more violent crimes happen more often in higher income counties. The only thing that can be tentatively concluded is that these counties are more likely to pursue the death penalty if a severe crime has been committed.

Age and Race (Composition of Inmates)

Age

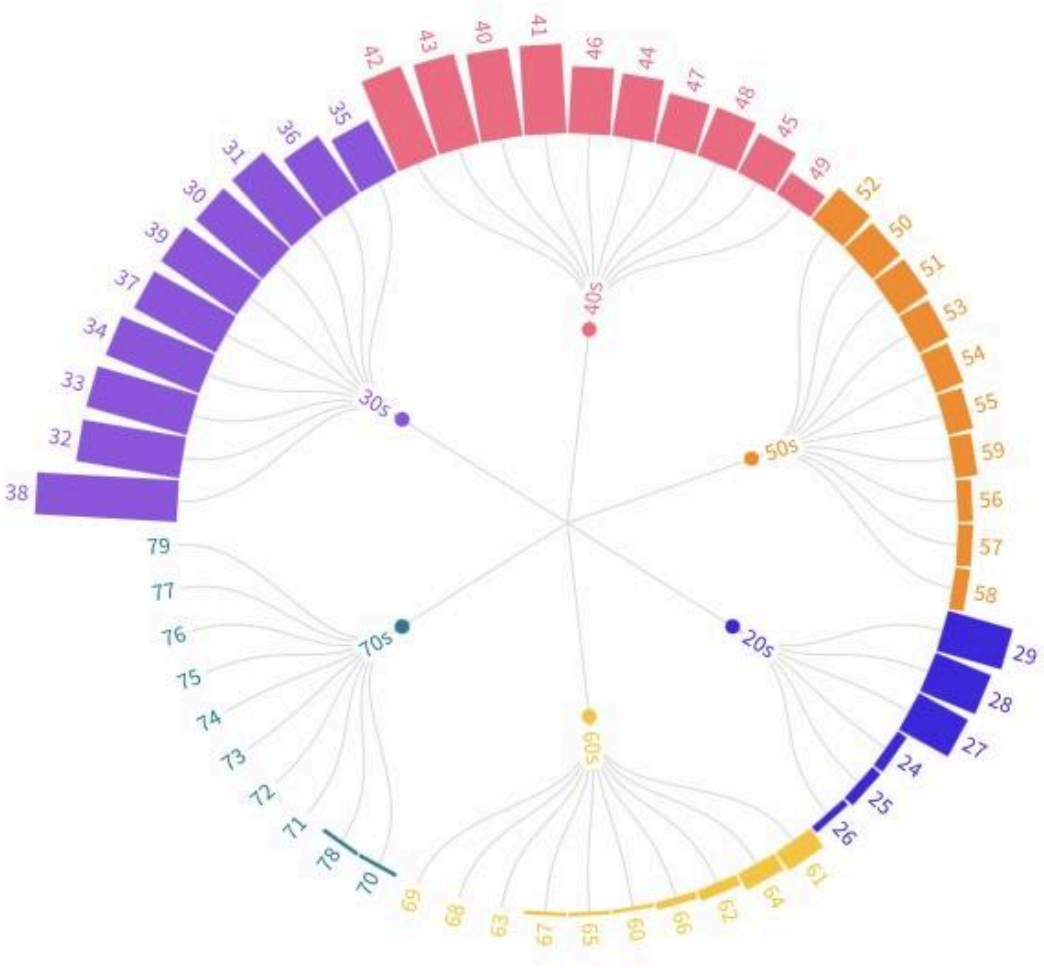


Figure 4. Age of Inmates When They Were Received

https://public.flourish.studio/visualisation/16438483

By analyzing the radial chart showing the age of inmates when received, it is evident that a significant number of them made their last statements in their thirties or around the age of thirty. A substantial portion of inmates were in their forties as well. Those older than 50 years old constitute only a small segment of the overall population.

Previous statistical data has indicated that the age range of inmates primarily falls between 26 and 45 (Federal Bureau of Prisons, 2019). During this age stage, individuals tend to experience comparatively heightened pressure from multiple sources. And individuals in this age range often possess the capability, particularly in terms of physical strength, to engage in criminal activities.

The concentration of inmates making their last statements in their thirties or forties aligns with the statistically identified peak in criminal activities, which may suggest a critical period in their lives, where circumstances or personal choices lead to significant consequences.

Race

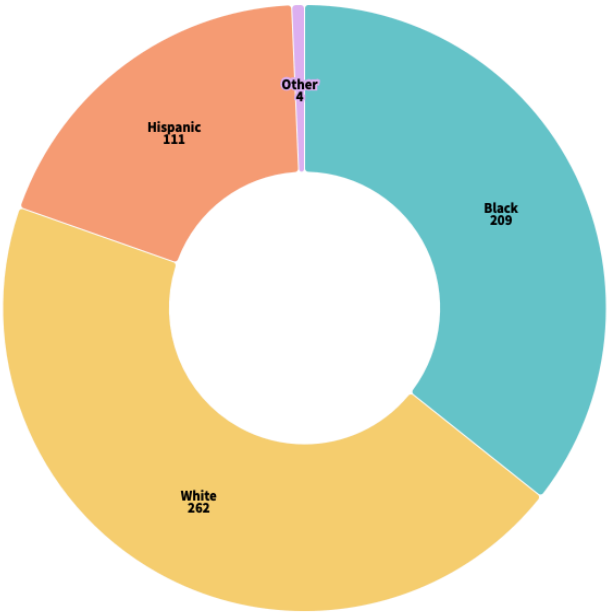


Figure 5. The Race of Inmates

https://public.flourish.studio/visualisation/16438627/

The doughnut chart highlights a majority of White inmates, followed by a slightly smaller proportion of Black inmates. Hispanic inmates constitute less than 20%, and there are only four inmates from other racial backgrounds.

Age & Race

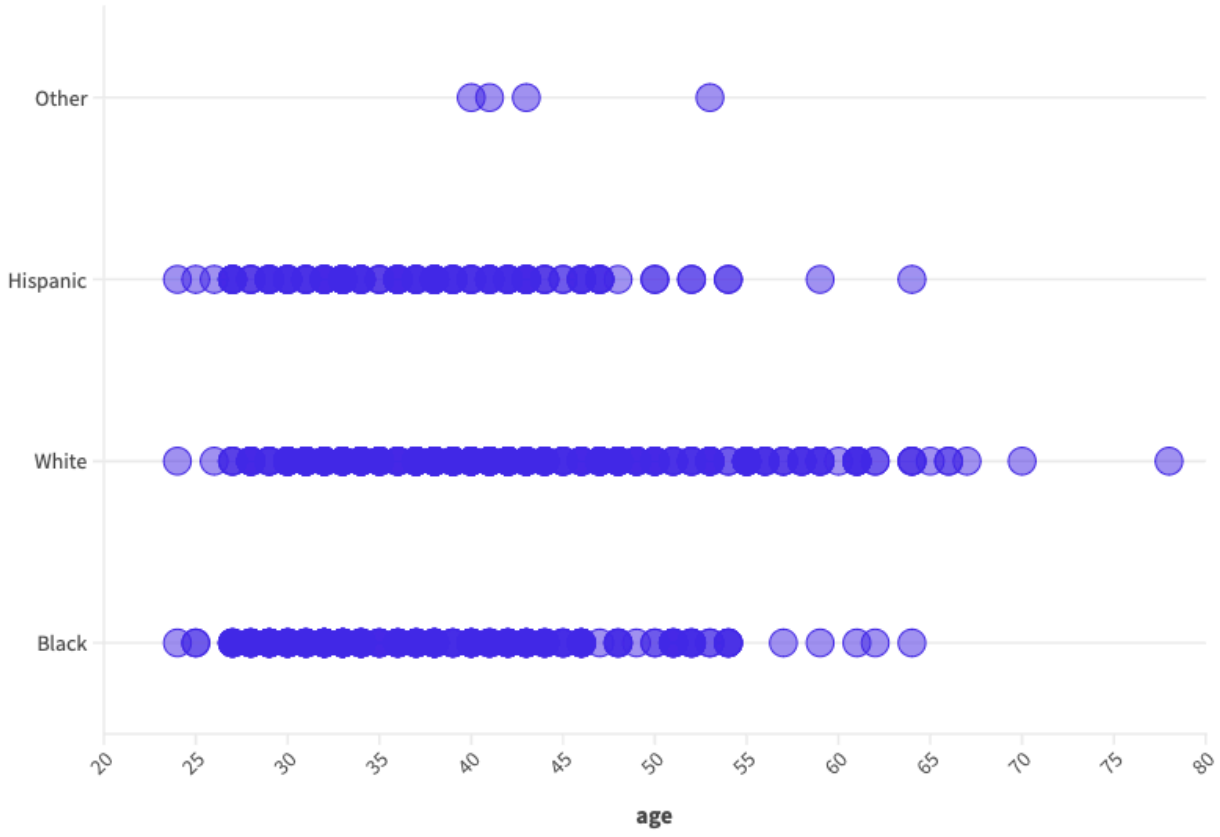


Figure 6. Age and Racial Distribution of Executed Inmates

https://public.flourish.studio/visualisation/16438589/

This dot plot shows that most inmates that are sentenced to death from Hispanic inmates are under 50 years old and Black inmates are under 55 years old, however, for white inmates, there is no significant reduction until the age of 65, at which point a noticeable decline becomes evident.

The observed age patterns suggest potential disparities influenced by legal practices, demographic composition, and socioeconomic factors, etc. This trend may stem from systemic biases within the criminal justice system and case-specific factors. A comprehensive understanding requires detailed analysis and consideration of these multifaceted influences.

Education and Occupation

Education

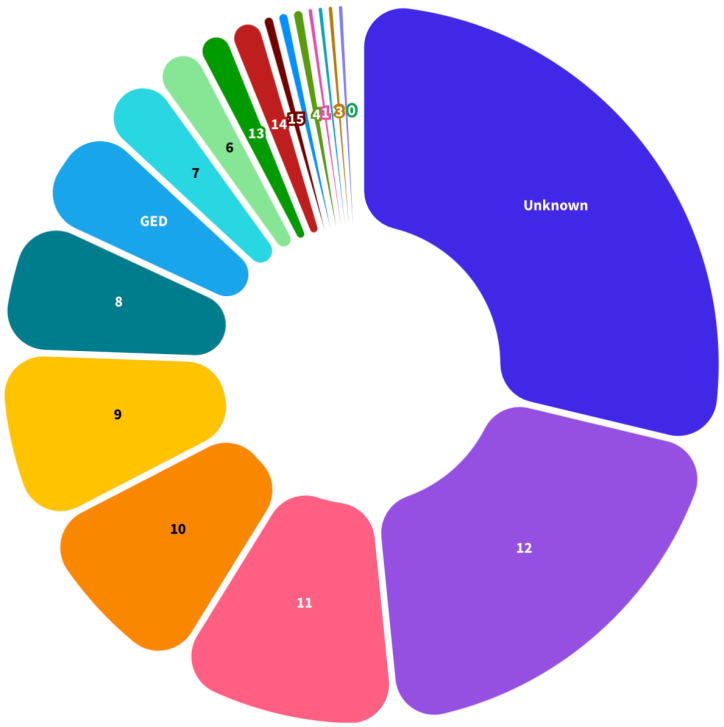


Figure 7. Years Spent in School of Those on Death Row

https://public.flourish.studio/visualisation/16385877/

The educational background, as shown in the pie chart, indicates that a majority of the inmates had completed high school, with fewer having attained higher education. Notably, based on the dataset, only 3 people finished 16 years of education. This raises questions about the correlation between educational attainment and propensity to be sentenced to death.

Occupation

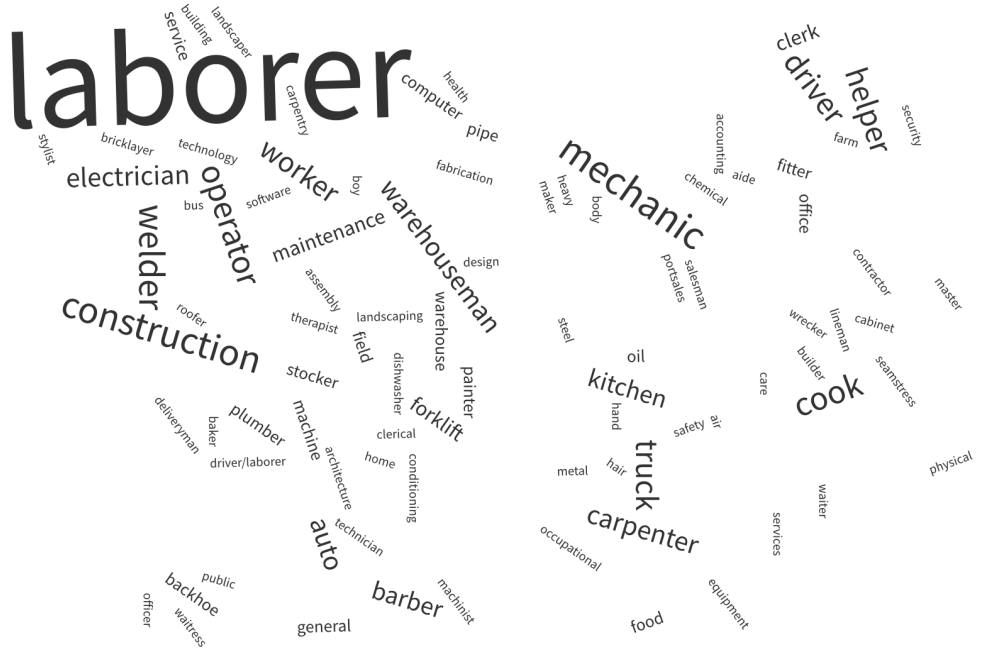


Figure 8. Word Cloud of Previous Occupation of Inmates

https://public.flourish.studio/visualisation/16368528/

It is notable that most inmates worked as laborers. They mostly did jobs in manual labor roles such as mechanics, drivers, welders, and construction workers.

The correlation between limited educational opportunities and increased crime rates is well-documented (Lochner and Moretti, 2004). Laborers and manual workers often have restricted access to quality education, a factor that significantly limits opportunities presented to them later in life (Gould, Weinberg, & Mustard, 2002). This educational deficit can lead to a cycle of poverty, where criminal behavior becomes a perceived necessity to survive.

Furthermore, it can be inferred that lower-income groups face a system inherently biased against them, diminishing their chances of obtaining justice, especially since they often depend on pro bono lawyers who are typically burdened with long waiting lists (Bright, 1994).

Time Spent on Death Row

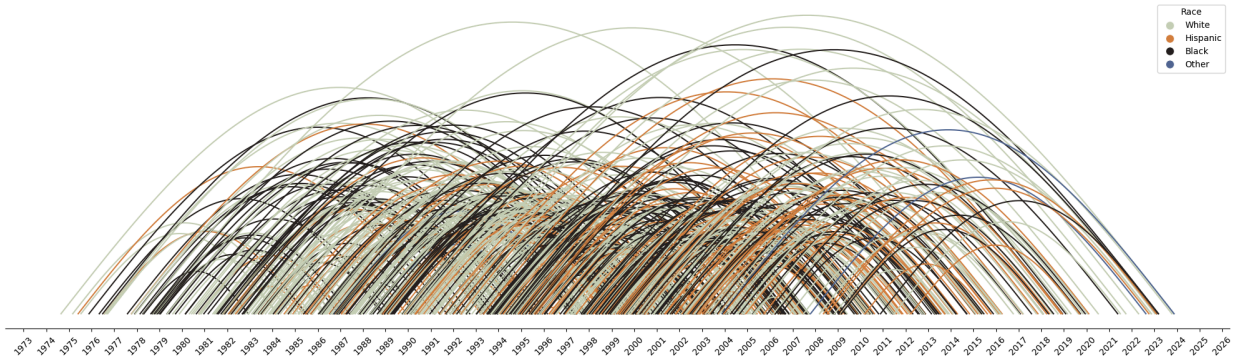


Figure 9. Time Spent on Death Row Individually by Race

The arc diagram presents a stark picture of the time spent by inmates on death row. The duration from the date of the offense to the execution date spans decades for many inmates, while for some, it can only take as short as one year.

The average time spent on death row prior to execution in this dataset is approximately 4,218 days, which is about 11.5 years. A Hispanic man from Potter County, holds the record for the shortest duration on death row. He was received on January 10, 1996, and executed on September 18, 1996, spending a total of 252 days. On the other a White man from Randall County, experienced the longest duration on death row. His time began on June 4, 1991, and ended with his execution on November 9, 2023, totaling 11,846 days.

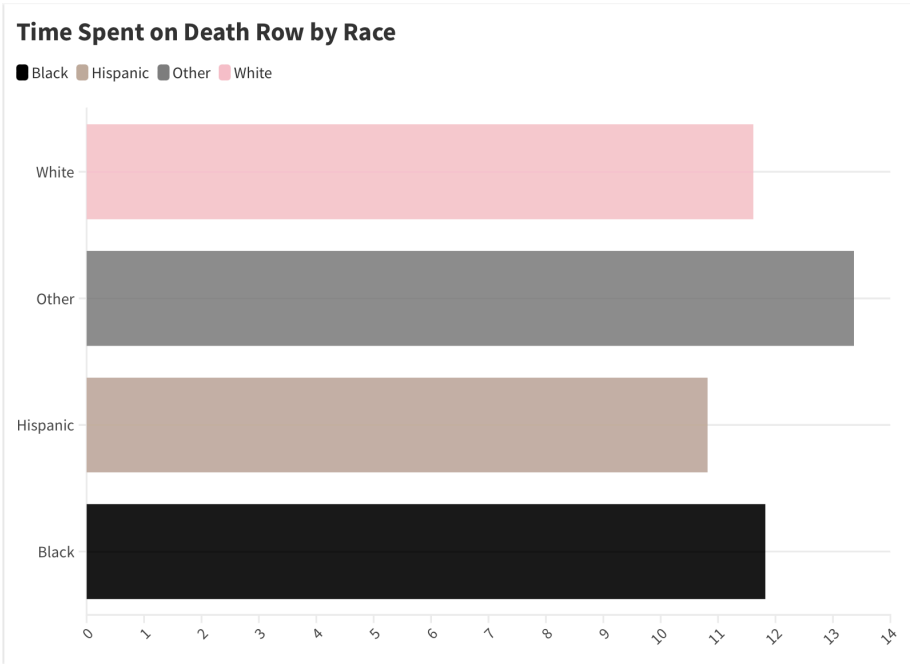


Figure 10. Time Spent on Death Row by Race

https://public.flourish.studio/visualisation/16386914/

According to research by Johnson (2020), the likelihood of seeking the death penalty against Latino offenders was found to be four times higher as compared to their white counterparts. Interestingly, based on our dataset, Hispanic people spend the least time on the death row.

Last Statements Analysis

In this section, we conduct text analysis on the last statements using Voyant. We begin by examining the longest occurring words in the last statements and studying the collocates of certain words to determine their specific references in the text. Subsequently, we classify the text based on age, race, and educational background, conducting comparative analysis in an attempt to identify differences and similarities among different groups.

Word Frequencies of Death Row Inmates's Last Statements

To research the most frequently mentioned words in the last statements of death row inmates, we utilized the "Cirrus" tool of Voyant to create a word cloud based on the words in our corpus. In Figure 1, larger words signify more frequent mentions, while smaller words are less common.

Figure 11. Word cloud of most frequent words in the whole corpus

The most striking observation is that in the last statements of Texas inmates, the word "love" is the most frequently used, appearing 823 times. This high frequency indicates that, in their last moments, prisoners choose to express profound affection and care towards their families. "Family" follows closely, with 383 appearances, underscoring the significance of familial ties; prisoners emphasize their connections with family at life's end. This is likely because family relationships are often the most cherished and emotionally invested parts of an individual's life, especially in the last stages (Amjad, I., & Rafi, M. S., 2017).

The phrase "thank you" also occurs 338 times, suggesting that expressing gratitude is a common sentiment among prisoners when facing their ultimate fate. "Sorry" appears 290 times, which may reflect their remorse for their crimes and apologies to the victims and their families. The presence of "sorry" and "forgiveness" suggests a sense of penitence and a search for absolution. This may indicate that prisoners, when facing the death penalty, hope to express regret and seek forgiveness (Eaton & Theuer, 2009).

Religion is another common theme in their last statements. We observe words like "God" (252 times), "Lord" (178 times), "pray" (68 times), "Christ" (62 times), and "Allah" (43 times) in the corpus. It shows that, for some inmates, religious faith is a source of comfort or a means of seeking redemption in their final moments. The word "father" (85 times) caught our attention. Does it refer to the real father or the holy father? We will use another tool in Voyant to research it in the next section.

Furthermore, negative and extreme emotional expressions like "death" (112 times) and "pain" (100 times) are less frequent in the last words. This choice of emotional expression may represent the prisoners' attempts to find peace and reconciliation as their lives come to a close. Notably, compared to the positive expressions, the word "hate" appears much less often (42 times), indicating that prisoners tend to express positive words like “love” rather than negative words like “hate” in their last statements.

Overall, the analysis of the frequency of these words paints a picture: faced with the grim reality of the death penalty, death row inmates employed a greater proportion of positive emotion expressions compared to negative emotion expressions (Hirschmüller, S., & Egloff, B., 2016). They more often chose to articulate their last thoughts through expressions of love, familial bonds, and gratitude, as well as seeking peace and forgiveness through religion.

Word Collocations of Death Row Inmates's Last Statements

The word cloud provides us with a clear glimpse of the theme of the inmates' last words and raises new questions, such as whom inmates are expressing remorse for. When they mention "family," does it refer to their own family or the family of the victim? Similarly, regarding the previously mentioned question: does the word "father" represent their biological father or the holy father? To address these questions, we utilized the "Collocate" and "Links" functions in Voyant, aiding our understanding of word relationships and how other words are connected to the theme.

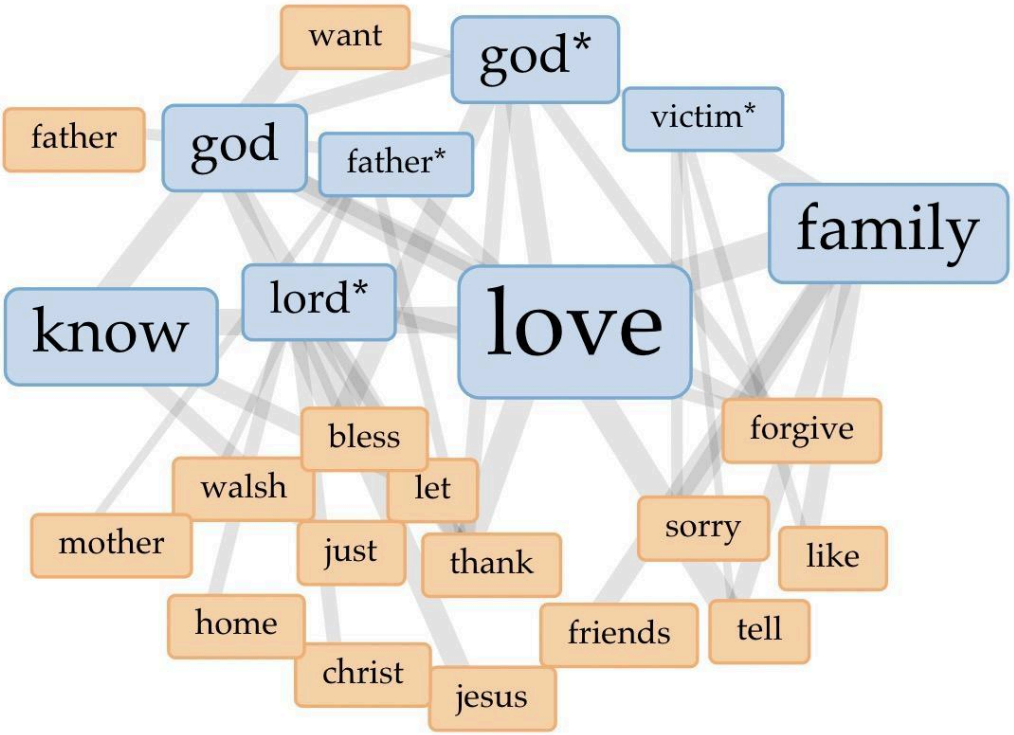


Figure 12. Collocates of the keywords

First, we selected the keywords "sorry," "family," and "victim" to investigate whether the family is their own family or the families of the victims, and to whom they are apologizing. In

our corpus, "family" is the second most frequently used word (383 times), while "sorry" is the fifth most frequently used word (290 times). The word "victim" appears 50 times in total. The "Collocate" tool reveals the most frequently associated words for the selected keywords in the corpus. For the word "sorry," some noteworthy collocates include "family" (46 times), "pain" (33 times), "love" (26 times), and "forgive" (15 times). The collocates of the word "pain," such as "caused" (47 times), "suffering" (16 times), "family" (15 times), "forgive" (7 times), "apologize" (7 times), and "sorrow" (5 times), suggest that when inmates mention pain, it typically refers to the pain they caused to the victim and their family.

Furthermore, the word "family" is found in proximity to the word "victim" 41 times. Based on this data, we suppose that when inmates say "sorry" in their last words, most of the time, they are apologizing to the victims and their families. In Texas, since January 12, 1996, close friends and family of victims have been allowed to witness executions. Although the victims' and the inmates' families and friends cannot see each other face to face, it provides an opportunity for the inmates to apologize, allowing the victims' families and friends to hear it. While an apology cannot bring the victim back to life, it holds significance for their family to hear remorse from the perpetrator (Eaton & Theuer, 2009).

To figure out whether the word "father" represents the real father or the holy father, we examined its collocates. The most frequently occurring collocates are "love" (18 times), "thank" (11 times), "mother" (7 times), "heaven" (7 times), and "ask" (7 times). While other religious words collocate with the term "father," their frequencies are not high, including "hallowed" (5 times), "spiritual" (4 times), "pray" (4 times), "god" (4 times), "prayer" (3 times), "Jesus" (3 times), "forgive" (3 times), "touch" (2 times), and "soul" (2 times). The results indicate that in our corpus, the term "father" usually represents the holy father. Before their execution, it is

common for inmates to pray and seek forgiveness from the heavenly father. It appears that they approach their impending death with a sense of calmness, pleading with the holy father to accept their souls.

The "Collocates" tool helps us further understand what the inmates want to express before their execution. It reveals that expressing love for family and friends, remorse towards victims and their families, and seeking forgiveness from God are the three most common themes in death row's last statements.

Comparative Analysis of Various Racial Groups

Furthermore, we utilize the "Trend" tool to explore the frequency of use of five keywords - "love", "family", "know", "thank", and "sorry" - in different corpora representing different racial, age and educational backgrounds groups. These keywords are the most frequently used words identified from the corpus of the overall corpus.

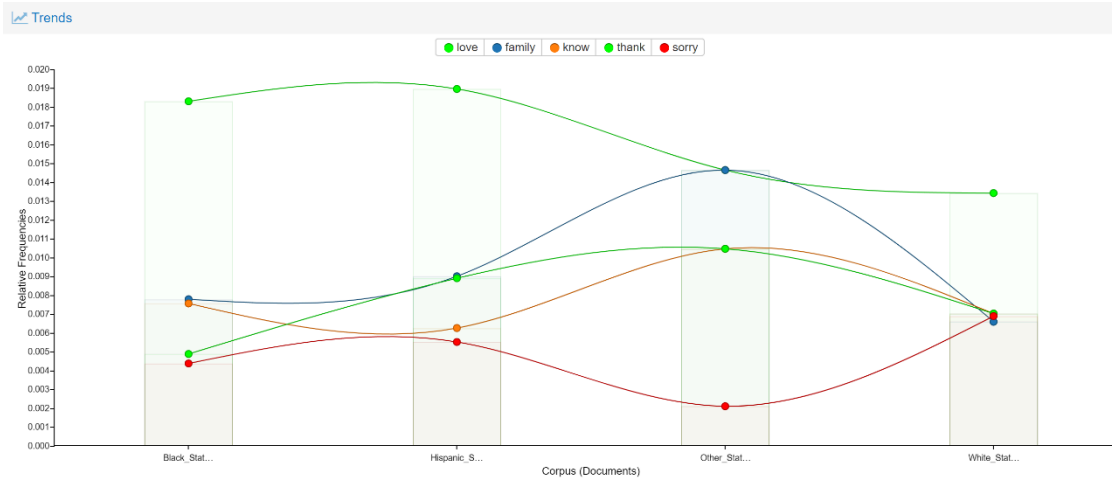


Figure 13. Trend of most frequent words between four racial groups

The horizontal axis details four groups: the Black community, the Hispanic community, the Other racial community, and the White community. The vertical axis quantifies the relative frequency of each keyword in the documents. The analysis highlights several significant trends:

The word "love", as a shared emotional expression, appears frequently in statements of all races, especially in statements of the Hispanic community, with its frequency being most prominent.

The term "family" is frequently mentioned in statements of the Other racial community, while it is relatively less mentioned in statements of the White community.

The word "know" is relatively common in statements from the Other racial community, but its frequency decreases in statements from the White community.

The term "thank", which expresses gratitude, is most frequently used in statements from the Other racial community, while it appears less frequently in statements from the Black community.

The word "sorry" is not very common in all statements, especially in statements of the Other racial community, and its use is almost negligible. In contrast, the use of it by the White community is relatively high.

By analyzing the frequency of using these keywords, we can preliminarily reveal the emotions that various ethnic groups may tend to express in their statements. The Black and Hispanic communities are particularly passionate about expressing love before being executed. Although other ethnic groups are very enthusiastic about expressing their feelings towards their families, they have almost no words related to apology. Additionally, the expression of emotions among the white population is very even, whether it is gratitude or apology.

Using Voyant's "Summary" tool allows us to extract and analyze four different document sets' Vocabulary Density and Average Words Per Sentence. In addition, we have also noticed significant differences in the use of Distinctive Words compared to the rest of the corpus. Given that Voyant does not provide a direct visualization option for this tool, we choose Python to

create line charts and word clouds to more intuitively display the differences and characteristics between these four sets of texts. This will deepen our understanding of the document set's characteristics and help highlight each group's unique language styles.

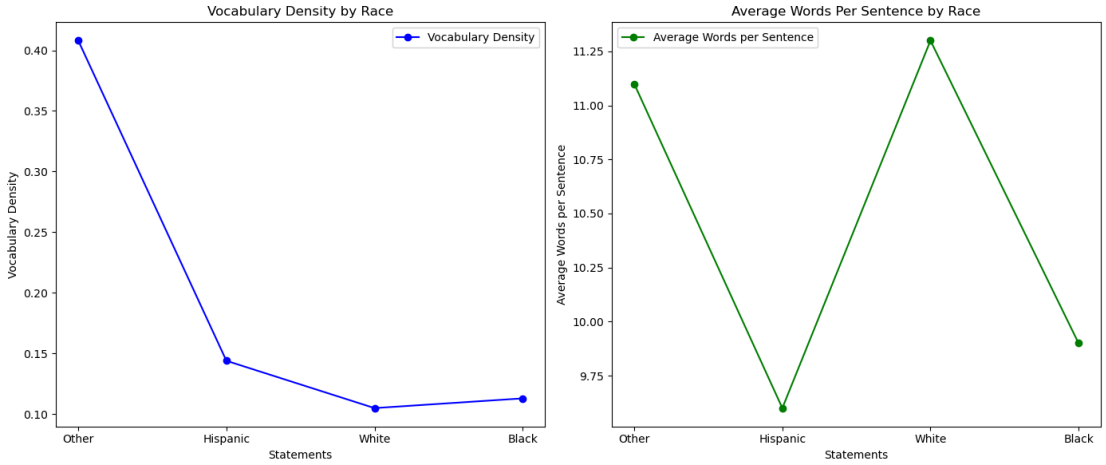


Figure 14. Vocabulary density and average words per sentence between four racial groups

The results showed that the word density of the statements from the Other racial community was the highest at 0.408, followed by the ones from the Hispanic community at 0.144. This indicates that these two documents have a more diverse use of vocabulary. In contrast, the vocabulary density of the White Community's statement and the Black Community's is relatively low, at 0.105 and 0.113, which may indicate that these statements are more concentrated in vocabulary usage.

Regarding the average words per sentence, the White community declared the highest average number, at 11.3 words, similar to the Other racial community, at 11.1 words. The sentence lengths for the Hispanic and Black community statements are relatively short, with an average of 9.6 and 9.9 words per sentence, respectively.

Figure 15. Word cloud of distinctive words between four racial groups

Furthermore, the distinctive words in the Black community's statements include "forward" (15 appearances), "tell" (61 appearances), "ya'll" (58 appearances), "lynching" (11 appearances), and "stay" (40 appearances). The Hispanic community's statements include "spanish" (11 times), "tell" (34 times), "Mexico" (7 times), "everyone" (25 times), and "waiting" (10 times). Unique words in statements from the Other racial community include "learned" (5 times), "walker's" (1 time), "transmitted" (1 time), "transgressed" (1 time), and "rich" (1 time). The distinctive words expressed by the White community include "tell" (52 times), "father" (47 times), "heart" (44 times), "ye" (9 times), and "coming" (18 times).

These data reflect the language features in different statements and may imply the themes and cultural backgrounds different groups are concerned about when expressing themselves. The

words "Mexico" in the Hispanic community statements represent themes related to cultural identity. Besides that, the words "spanish" show that the common practice of translating statements from Spanish to English for documentation purposes, clearly indicating the original use of the 'Spanish' language. This choice to express their final sentiments in Spanish suggests a desire to convey genuine emotions in their native language rather than opting for English for wider comprehension. Additionally, the words "tell", "father", and "heart" in the White community statements may reveal more traditional language tendencies in family values, emotional expression, and religious connotations.

Comparative Analysis of Different Age Groups

The following research shifts towards an in-depth comparison of the last statements from different age groups. To make this comparison, we divided death row inmates into six age groups: 20s, 30s, 40s, 50s, 60s, and 70s. By comprehensively analyzing the last statements of these six groups, we aim to reveal the potential connection between age and language use.

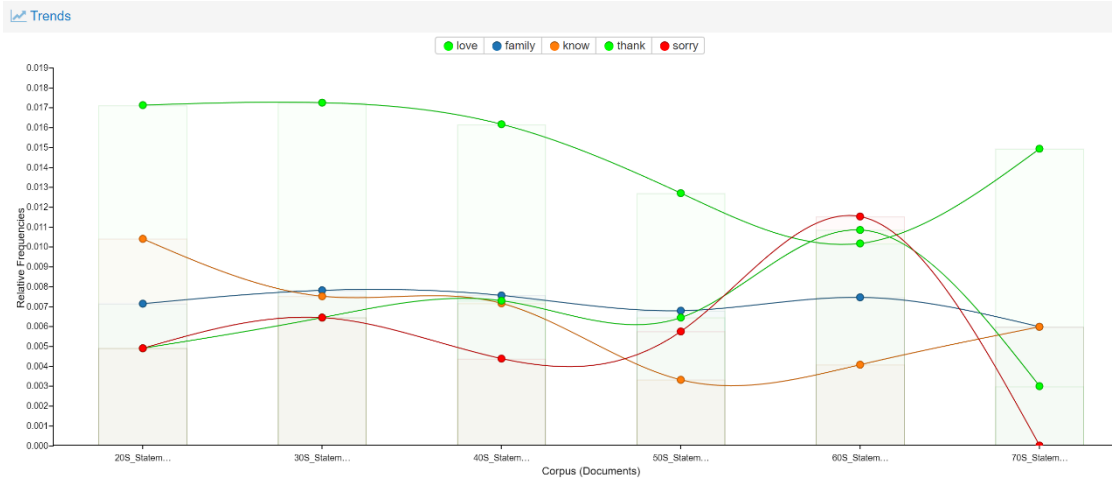


Figure 16. Trend of most frequent words between six age groups

The horizontal axis details the last statements for different age groups: 20s, 30s, 40s, 50s, 60s, and 70s. The vertical axis quantifies the relative frequency of each keyword in the document. The analysis results highlight several significant trends:

The word"love", as a shared emotional expression, frequently appears in statements of all age groups, especially in statements of the 20s and 30s generations, with its frequency being most prominent.

The term "family" is mentioned relatively evenly in the statements of the six groups, with little difference.

The word "know" is used most among the 20s generation, but in statements from the 50s generation, its frequency significantly decreased.

The term "thank" is most frequently used in statements from the 60s generation, while it appears very infrequently in the 70s generation.

The word "sorry" is relatively common in statements from the 60s generation, while its frequency is 0 in statements from the 70s generation.

The frequency of these keywords shows that death row inmates of different age groups vary in their willingness to express emotions before death. The frequency of using the word love has almost continuously decreased from the 20s to the 60s generation, reflecting that young death row inmates are more willing to express love in their last statements than older ones. This phenomenon is consistent with the survey results of other scholars, where younger adults are more inclined to express love than older adults (Montepare & Dobish, 2014).

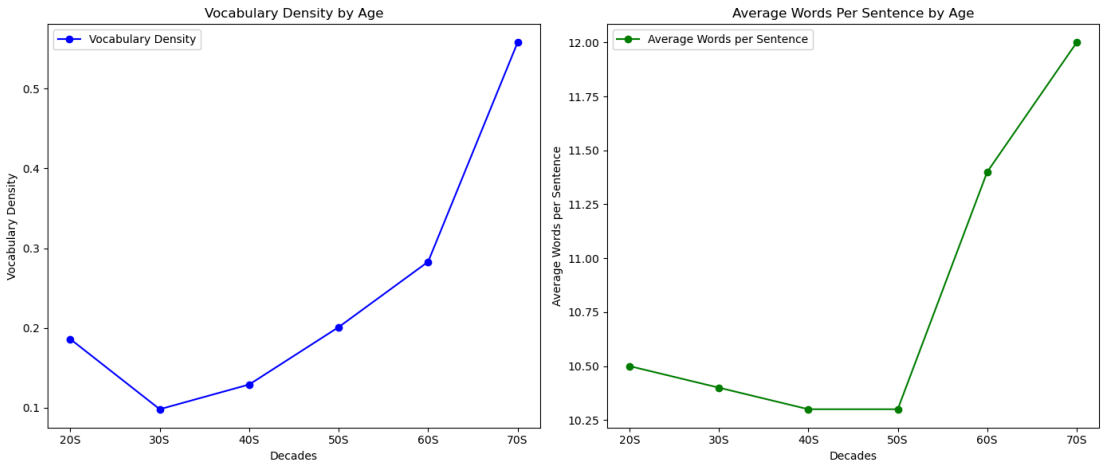


Figure 17. Vocabulary density and average words per sentence between six age groups

Regarding vocabulary density, the last statements of the 70s generation of death row inmates showed the highest lexical diversity, with a value of 0.558, followed closely by the 60s generation, with a lexical density of 0.283. The density of the 50s generation is 0.201, that of the 20s generation is 0.186, and that of the 40s generation is 0.129. The last statement vocabulary density of the 30s generation is the lowest, only 0.098, indicating that individuals may use more diverse and varied vocabulary in their last statements as age increases. In contrast, the vocabulary choices of the 30s and 40s generations appear more limited and concentrated.

In terms of average words per sentence, the last statement of the 70s generation contains an average of 12.0 words per sentence, the highest among all age groups. The 60s generation has 11.4 words, while the average words per sentence of the 20s and 30s generations is slightly shorter, at 10.5 and 10.4 words, respectively. The average words per sentence of the last statements of the 40s and 50s generations is the shortest, at 10.3 words. This trend reflects that elder people use more complex and detailed sentence structures when expressing themselves.

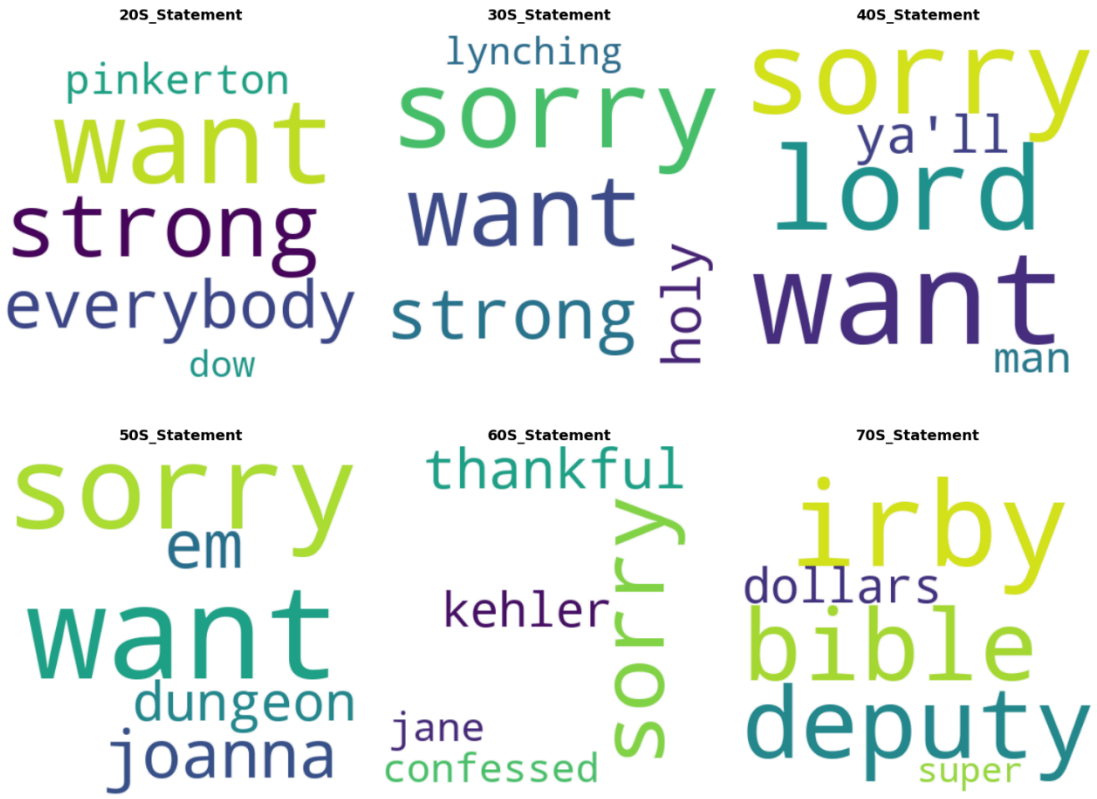


Figure 18. Word cloud of distinctive words between six age groups

About the usage of distinctive words in the last statements of each age group, the 20s generation prefers "strong" (18 times), "pinkerton" (4 times), "want" (36 times), "everyone" (14 times), and "do" (3 times). The 30s generation prefer to use "Holly" (25 times), "Sorry" (150 times), "Want" (115 times), "Lynching" (11 times), and "Strong" (48 times). There are "ya'll" (46 times), "want" (77 times), "man" (31 times), "sorry" (66 times), and "lord" (63 times) in the statements of the 40s generation. The 50s generation's statements include "Joanna" (4 times), "em" (4 times), "dungeon" (4 times), "want" (37 times), and "sorry" (33 times). The 60s generation like using "kehler" (3 times), "thankful" (4 times), "sorry" (17 times), and "Jane" (2 times). The 70s generation prefers "irby" (3 times), "consumption" (3 times), "daughters" (2 times), "Bible" (3 times), and "super" (1 time).

The frequent use of "holy" and "lord" by the 30s and 40s generation may indicate that this age group is seeking spiritual solace before death. The use of the "Bible" in the 70s generation demonstrates a dependence on religion. It is worth mentioning that both the younger and older groups cited religious content in the last statement. That is because seeking spiritual comfort through religion is a common practice among death row inmates (Schuck, A. R. T., & Ward, J., 2008).

Comparative Analysis of Distinct Educational Background Groups

Finally, our research compares the last statements with different educational backgrounds. For this purpose, we divide each death row inmate into three educational level groups based on the highest grade they completed: basic education (0-8 years), partial or complete high school education (9-12 years), and further education (13-16 years) (Nanyonjo et al., 2008). Through the analysis, we aim to explore the correlation between educational background and language expression.

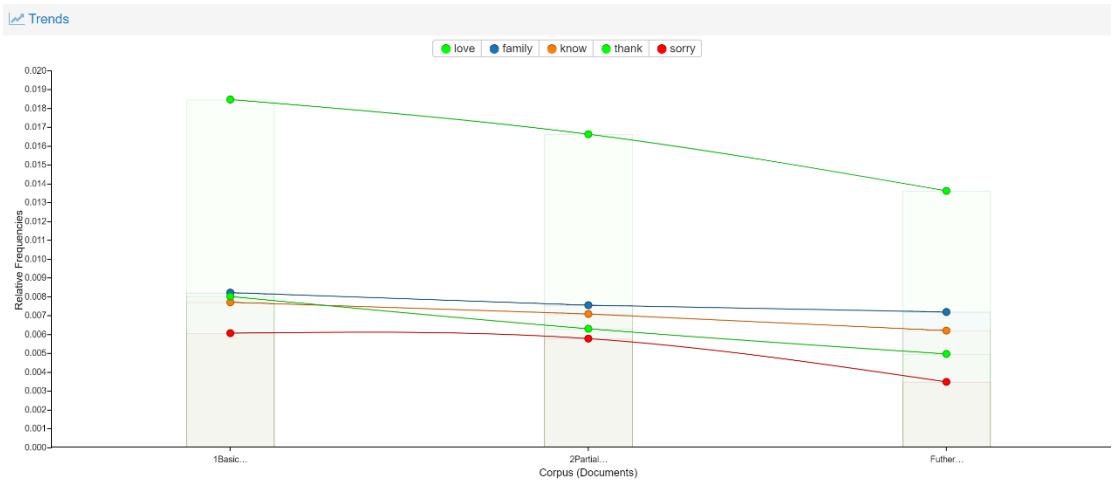


Figure 19. Trend of most frequent words between three educational groups

The horizontal axis details the different educational level groups based on the highest completion years: basic education (0-8 years), partial or complete high school education (9-12 years), and further education (13-16 years). The vertical axis quantifies the relative frequency of each keyword in the document. The analysis results show the following trends:

Except for the word "sorry", the frequency of use of the other four words continues to decrease with the hint of educational background.

The word "sorry" has the highest relative frequency in the basic education group, while it has a lower frequency in the continuing education group.

The above trends indicate that there may exist a negative correlation between educational level and the tendency to express emotions: in the basic education group. For example, the word "love" is used the most frequently in the basic education group, while with the improvement of the education level, the frequency of mentioning "love" gradually decreases.

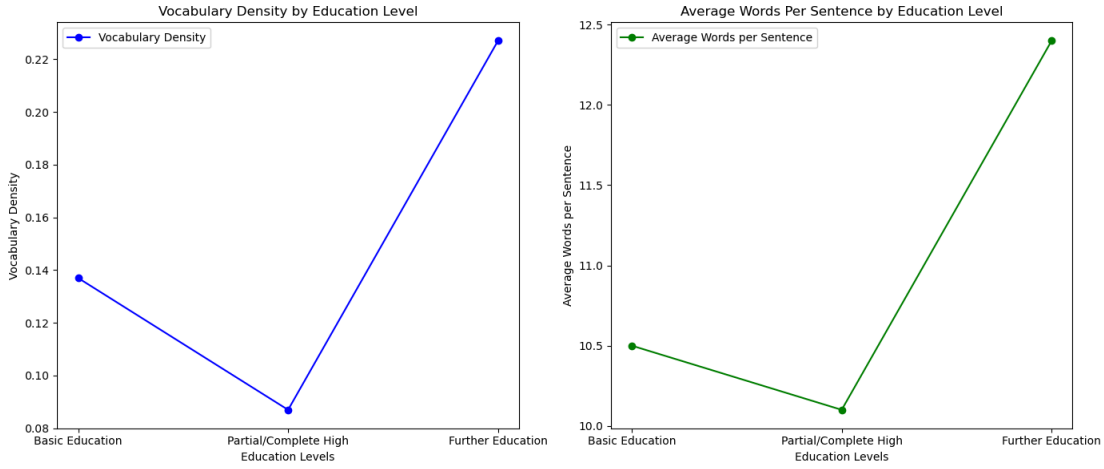


Figure 20. Vocabulary density and average words per sentence between three educational groups

The analysis of vocabulary density shows that the group receiving further education has the highest vocabulary density at 0.227. In contrast, the partially or fully educated high school group has the lowest vocabulary density at only 0.087, and the basic education group has 0.137.

This may mean that individuals with higher education use richer vocabulary in their last statements (Twenge et al., 2019).

In terms of the average words per sentence, the further education group also ranked highest with an average of 12.4 words, while the average words per sentence of the partial or complete high school education group was the shortest, at 10.1 words, and the basic education group was 10.5 words. This suggests that a higher level of education may be related to a tendency to use longer sentences.

Figure 21. Word cloud of distinctive words between three educational groups

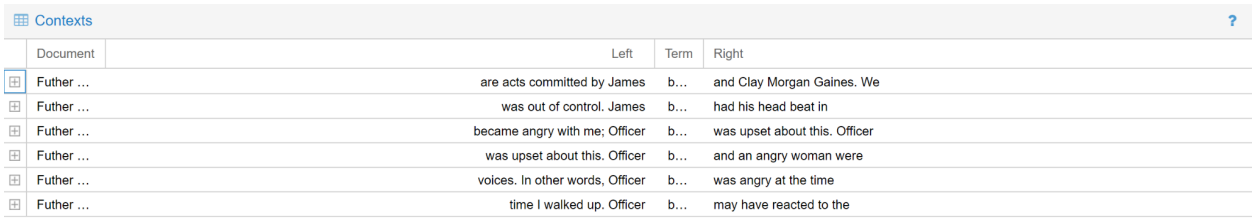
The use and frequency of distinctive words in the last statements of death row inmates with different levels of education demonstrate distinct patterns:

For inmates who have received basic education, the standard vocabulary used in their statements are "y'all" (24 times) and "mom" (20 times). In addition, the vocabulary "today" (16 times), "perfect" (5 times), and "soul" (13 times) reflect attention to the current state of life and the inner world.

For inmates with partial or complete high school education backgrounds, frequent use of vocabulary related to religion and social issues, such as "crime" (19 times), "black" (16 times), "cry" (15 times), and "allah" (38 times), is used in their statements, suggesting a profound

reflection on criminal behavior and religious beliefs. Meanwhile, the repeated use of "mom" (36 times) again emphasizes the family's central position in their lives.

For inmates who have received further education, the unique words that appear in their final statements include "boswell" (6 times), "ros" (3 times), "robinson" (3 times), "irby" (3 times), and "gene" (3 times).



Contexts

Document	Left	Term	Right
Futher ...	are acts committed by James	b...	and Clay Morgan Gaines. We
Futher ...	was out of control. James	b...	had his head beat in
Futher ...	became angry with me; Officer	b...	was upset about this. Officer
Futher ...	was upset about this. Officer	b...	and an angry woman were
Futher ...	voices. In other words, Officer	b...	was angry at the time
Futher ...	time I walked up. Officer	b...	may have reacted to the

Figure 22. Context of the key word boswell in the further education group

Although "boswell" was mentioned six times, according to the "Context" tool, this word only appeared in one person's last statement, and the death row inmate was accusing Boswell of being the real culprit. So, the use of the word "boswell" is not universal.

Through Voyant, we are able to quickly identify the most common and prominent vocabulary in the text, as well as changes in different parts. This method can reveal the important themes of the text and the author's stylistic tendencies, providing a macro-understanding of the text. However, Voyant's analysis fails to capture the subtle aspects of the text, such as the specific details of last statements and the metaphors in the text. Therefore, future research should still combine distant reading (Moretti, 2000) and close reading to thoroughly understand the complexity of the text.

Ethical Consideration

In conducting our research on the last statements of Texas' death row inmates, we navigated critical ethical terrain, notably regarding respect for the deceased and their families, maintaining objectivity and bias, and addressing privacy concerns.

Respect for the Deceased and Their Families: Our research intrinsically intersects with the sensitive realm of mortality and the justice system. It is imperative to approach this with a profound respect for the deceased and their families. This respect is manifested in our handling of the data, where we prioritize dignity and privacy. This approach aligns with the ethical standards set by American Psychological Association (APA) (2002), which emphasize the importance of respecting participants and their families in research, especially in studies involving vulnerable populations or sensitive topics. Our analysis avoids sensationalism and instead aims to provide a dignified portrayal of the inmates, recognizing them not just as subjects of study but as individuals with inherent worth and dignity.

Objectivity and Bias: The credibility and validity of our findings hinge on our commitment to objectivity. The neutrality and impartiality are not just about the accuracy of a single study, but about maintaining the integrity and trustworthiness of the entire research process and its impact on society (Nii Laryeafio & Ogbewe, 2023). In line with scholarly standards of unbiased research, we actively identify and mitigate personal biases and preconceptions that could color our analysis (Johnson et al., 2020). This commitment to objectivity ensures that our conclusions are data-driven and not unduly influenced by our personal beliefs or societal biases.

Privacy Concerns: Despite working with publicly available data, we adhered to strict privacy protocols. Ethical research standards uphold the dignity and worth of all individuals,

regardless of their legal status or actions (Jameel & Majid, 2018). We carefully considered the implications of our research on the privacy and dignity of the individuals involved, ensuring that our data handling and reporting practices protected their identity and respected their privacy. This was particularly crucial given the nature of our research and the potential for public interest. Furthermore, we consider that while legal considerations of privacy often focus on living individuals, ethical considerations extend to the deceased, recognizing that the release of certain information could lead to reputational harm or distress among surviving relatives and communities.

Conclusion

At the beginning of our research project we posed the following question: "What insights into the Texan criminal justice system can be derived from an analysis of demographic data and the last utterances of death row inmates?" While the topic of the death penalty has, in our opinion, only become more poignant over the course of our research, we have also managed to make some discoveries that might not have been obvious from the onset.

First, we used a visualization of the counties of origins of the death row inmates in our dataset to examine if certain counties were perhaps more prone to violent and/or extreme offenders than others. While some counties did seem to have had significantly more death row inmates than others, by comparing data from our dataset to 2020 census data, as well as considering the cost of sentencing someone to death, the reason as to why these specific counties had more death row inmates became clear. These counties were not inherently more inclined to violent or repeated crimes, they are simply more willing to serve the death penalty, as the death penalty is extremely expensive. Therefore, a GIS visualization is interesting for research

purposes, but the data on-hand cannot make any claims as to whether certain counties are more prone to have more crimes as well. Instead, our visualization simply reinforced the idea that there is a cost associated with the death penalty that some counties are unable, or unwilling, to bear.

Secondly, we looked at the data on age, race, education, and occupation. Based on this data, we could conclude that the ages at which most death row inmates were incarcerated aligns with the national average age of offenders — regardless of the severity of the crime. We also found that most offenders are White, and that White inmates are more likely to commit crimes at an older age as well. On the other hand, Hispanic inmates spend the least amount of time on death row, possibly because the death penalty gets sought against them way more frequently. Similarly, we found that most death row inmates did not complete any type of further education, which explains why most of the inmates tended to have had manual labor jobs. This finding correlates with previous studies that have shown that a lack of education can increase the likelihood of crime. However, this also raised the question of whether or not the system — which, as previously discussed, is heavily dependent on money — is inherently biased towards lower educated people and/or manual laborers.

Finally, we conducted text analysis on the last statements of the Texas death row inmates. During this analysis, we found that "love", "family", "thank you", and "sorry" were the most frequently used words in final statements. These words seem to indicate regret. Moreover, religion was also a frequent topic in the final statements, indicating that people often find comfort in faith in the face of death. Lastly, words such as "death", "pain", and "hate" were least frequent across the board, which might indicate that people are more inclined to think either positively or wistfully, rather than negatively, in their final statements to the word. These

findings are in line with the findings from our look at collocations in the statements as well as the comparative analyses on racial groups, age groups, and educational background groups.

As such, our findings conclude something very human. It is easy to forget that these people we have datafied were people sometimes, as there is a great distance between us as data analysts and a person who was sentenced to death however many years ago. However, looking at the final statements of these individuals could be considered a journey through the final moments of individuals facing their ultimate fate. Each statement serves as a window into the complexities of human emotions, societal judgments, and the irreversible nature of the death penalty. The analysis suggests a complex interplay between societal, legal, and personal factors that culminate in the ultimate penalty. As such, this study not only contributes to the ongoing debate about the death penalty but also humanizes those often reduced to mere numbers in criminal statistics.

References

Association (APA). (2002). *Ethical principles of psychologists and code of conduct*. American Psychological Association. [Amended 2010 and 2016]. https://www.apa.org/ethics/code/ethics-code-2017.pdf

Amjad, I., & Rafi, M. S. (2017). A Phenomenological Analysis of Death Row Inmates' Last Words. *Pakistan Journal of Criminology*, *9*(3), 114-127. https://ssrn.com/abstract=3827757

Baicker, K. (2001, July). The Budgetary Repercussions of Capital Convictions. *National Bureau of Economic Research, 4*(1), 1-16. https://doi.org/10.3386/w8382

Bright, S. B. (1994). Counsel for the Poor: The Death Sentence Not for the Worst Crime but for the Worst Lawyer. *The Yale Law Journal, 103*(7), 1835–1883. https://doi.org/10.2307/797015

Death Penalty Information Center. (2019, May 26). Why is Texas #1 Executions? *Death Penalty Information Center*. https://deathpenaltyinfo.org/stories/why-is-texas-1-executions

Death Penalty Information Center. (2023, September 29). Costs. *Death Penalty Information Center*. https://deathpenaltyinfo.org/policy-issues/costs

Eaton, J., & Theuer, A. (2009). Apology and remorse in the last statements of death row prisoners. *Justice Quarterly*, *26*(2), 327-347.

Federal Bureau of Prisons. (2019). BOP Statistics: Average Inmate Age. *Bop.gov*. https://www.bop.gov/about/statistics/statistics_inmate_age.jsp

Gould, E. D., Weinberg, B. A., & Mustard, D. B. (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1977-1997. *Review of Economics and Statistics, 84*(1), 45.

Hirschmüller, S., & Egloff, B. (2016). Positive Emotional Language in the Final Words Spoken Directly Before Execution. *Frontiers in psychology*, *6*, 1985. https://doi.org/10.3389/fpsyg.2015.01985

Johnson, J. L., Adkins, D., & Chauvin, S. (2020). A Review of the Quality Indicators of Rigor in Qualitative Research. *American Journal of Pharmaceutical Education*, *84*(1), 7120.

Johnson, S. (2020). The Influence of Latino Ethnicity on the Imposition of the Death Penalty. *Annual Review of Law and Social Science, 16*, 421-431. https://doi.org/10.1146/annurev-lawsocsci-042220-111211

Jameel, B., & Majid, U. (2018). Research fundamentals: Data collection, data analysis, and ethics. *Undergraduate Research in Natural and Clinical Science and Technology Journal*, *2*, 1-8.

Lochner, L., & Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *The American Economic Review, 94*(1), 155–189. http://www.jstor.org/stable/3592774

Moretti, F. (2000). Conjectures on World Literature. *New Left Review*, *2*(1), 54-68.

Montepare, J. M., & Dobish, H. (2014). Younger and older adults' beliefs about the experience and expression of emotions across the life span. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *69*(6), 892-896.

Nanyonjo, R. D., Montgomery, S. B., Modeste, N., & Fujimoto, E. (2008). A Secondary Analysis of Race/Ethnicity and Other Maternal Factors Affecting Adverse Birth Outcomes in San Bernardino County. *Maternal and Child Health Journal*, 12, 435-441.

Nii Laryeafio, M., & Ogbewe, O. C. (2023). Ethical Consideration Dilemma: Systematic Review of Ethics in Qualitative Data Collection Through Interviews. *Journal of Ethics in Entrepreneurship and Technology*, *3*(2), 94-110.

Schuck, A. R. T., & Ward, J. (2008). Dealing with the inevitable: strategies of self-presentation and meaning construction in the final statements of inmates on Texas death row. *Discourse & Society*, *19*(1), 43–62. http://www.jstor.org/stable/42889177

Texas Department of Criminal Justice. (2023, November 20). Texas Death Row Information. *Texas Department of Criminal Justice.* https://www.tdcj.texas.gov/death_row/dr_executed_offenders.html

Twenge, J. M., Campbell, W. K., & Sherman, R. A. (2019). Declines in vocabulary among American adults within levels of educational attainment, 1974–2016. *Intelligence*, *76*, 101377.

U.S. Department of Agriculture. (2023, June 16). County-level Data Sets. *Economic Research Service.* https://www.ers.usda.gov/data-products/county-level-data-sets/

United States Government. (2023). U.S. Census Bureau QuickFacts: Texas. *United States Census Bureau.* https://www.census.gov/quickfacts/fact/table/TX#