

Introduction

For this assignment, I decided to analyze the pros and cons regarding the privatization of prisons. Prison privatization can occur either through private companies offering services or management for government-owned prisons, or by a private company owning and operating the prison themselves. The roots of American private prisons can be traced back to the early 1800s where they allowed for the continuation of unpaid labor (often from minorities) while not technically being considered slavery. Prison owners would often be paid for the labor of their inmates who were imprisoned for extended periods of time for petty or non-existent crimes. More modern forms of private prisons were started in hopes of running large prisons more cheaply during the rising prison population during the war on drugs. As of 2019, 20 states do not use private prisons, while the other 30 do. The goal of this analysis was to identify popular words occurring in the arguments for and against prison privatization, while also performing sentiment analysis on the two texts.

Methods

The analysis was conducted on two separate text files: one containing three arguments supporting the privatization of prisons (pros), and one containing three arguments against the privatization of prisons (cons). The text files were taken from the website ProCon.org.

Once the data was collected, I performed a preliminary manual cleaning of the text. I removed headers from the arguments (Pro 1, Pro 2, Con 1, Con 2, etc...) because they only served to separate the arguments, which do not matter for this assignment. I also removed the bracketed numbers at the end of each argument as the sources for each argument were not necessary.

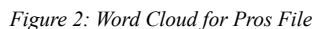
In the next phase, I used Python's Natural Language Toolkit library (NLTK) to perform a more thorough cleaning of the text. The first step in this process was tokenizing the text. Tokenization breaks the raw text down into words to allow the program to have a better understanding of the text. I then removed stopwords from the text. Stopwords do not add any value to the text, so removing them allows the program to focus on words that reveal more meaning in the text. I added the words "pro", "con", "public", "private", "prison", and "zealand" to the list of stopwords because they occurred too frequently in the files compared to other words without impacting the meaning of the text. I then removed all non-alphanumeric tokens to ensure that any special characters I missed during the preliminary cleaning did not interfere with the analysis. Lastly, I removed all tokens that were not longer than three characters because they most likely added no additional value to the text (and the assignment specified to do so).

The last phase involved performing sentiment analysis on each set of cleaned text, as well as producing word clouds that display the most frequently occurring words in each text. Sentiment analysis allows the program to identify what percentage of the text has a positive connotation, negative connotation, or is neutral. I used the Vader library to perform the sentiment analysis. I then used the Wordcloud library to produce the word clouds.

The results of the sentiment analysis were that the arguments for the privatization of prisons had a more positive sentiment, while the arguments against the privatization of prisons had a more negative sentiment.

Figure 1: Sentiment Analysis for Each File

The most frequently occurring words in the pros file are “government”, “recidivism”, “contract”, and “population”. The word cloud is shown below.



The most frequently occurring words in the cons file are “prisoner”, “staff”, “inmate”, and “cost.” The word cloud is shown below.



Figure 3: Word Cloud for Cons File

The words “prisoner” and “staff” occur frequently because the cons file focuses on how both prisoners and employees are exploited by private prisons. These prisons often benefit financially from holding more prisoners. This provides incentive to keep inmates behind bars, rather than providing them with opportunities to reintegrate into society. Furthermore, private prisons pay their employees less than public prisons in order to cut costs and increase profits. Lastly, private prisons also cut costs by leaving the most expensive prisoners to be handled by government-owned facilities. They only house minimum and medium security prisoners who are significantly cheaper than death row inmates, maximum security inmates, and inmates with serious medical conditions that public prisons have to house.