**1) Project Title**

An In-Depth Analysis with Visual Representation of the Criminal Justice Systems throughout the United States

**2) Current Date and Semester**

2/12/2021

Spring 2021

**3) Team Members**

Natalie Hani [nnh2@hood.edu](mailto:nnh2@hood.edu)

Lauren Thompson [lmt18@hood.edu](mailto:lmt18@hood.edu)

Jonathon Wallace [jw30@hood.edu](mailto:jw30@hood.edu)

Juan Pablo Chacon [jc34@hood.edu](mailto:jc34@hood.edu)

Alex Blank [adb19@hood.edu](mailto:adb19@hood.edu)

**4) Project Type**

Research

**5) Project Description**

The prison system is known to have many different sources of problems. Our goal is to research the prison system as a whole and go into detail analyzing the different problems that occur within it. For example, overcrowding in prisons, people staying beyond their sentence, separation of gender, and different types of charges and their severity. All these factors have created an inefficient prison system that subsequently impacts the entire system and relations to it. We would like to do a deeper study of these issues and their correlation within the prison system. Our solution is to develop an interactive visualization dashboard enabling dynamic analysis of data, highlighting the issues that occur within the prison system. Our goal is researching and developing data analytics to study our research topic. We will use specific data handling the number of cases specifically in the United States that are entangled with the issues of overcrowding, separation of gender, extended stays after sentencing, and specific criminal charges. We are well equipped to provide the necessary data analysis infrastructure for anyone to easily fuse and summarize as well as visualize vast amounts of disparate data. We will offer a basic statistical analysis and measurements of the data to provide a clear understanding of the subject our research is based upon to demonstrate that our research project was successful in developing a sound data analysis infrastructure. The report will present the purpose of the research, the means and methods used to gather data, the means and methods used to analyze the data, and the conclusions drawn from our analysis.

**6) Resource Requirements**

R-Studio will be used to organize data collected through the literature review and develop graphs to further aid our research. Through the tools R-Studio presents our team will have the ability to highlight key findings in the form of graphs and data tables. R-Studio will allow the team to detect relationships between datapoints within the dataset. The dataset consists of information regarding the prison system throughout the United States. This programming environment will allow us to find key values such as p-values, linear regression, highlighting irregularities, and information that will allow us to effectively analyze patterns and correlations within prison systems.

Python: Python is a high-level programming language which is used widely throughout computer and machine learning. This programming language will be used to analyze and find trends within the dataset given. Python provides the users the ability to use libraries such as Matlab, PySal, and Geopandas. These libraries provide both visual and mathematical analysis of the data provided. Matlab and PySal allow for the user to find trends and key values such as p-values, s-values, and linear regression. Geopanda will allow the users to have a visual aid of the general information such as what states, and facilities within, are overcrowded to further support their understanding of the analysis.

Github: Github is a code hosting platform for version control and collaboration. It lets users and others work together on projects. By having a collective interactive dashboard through Github the team will be able to stay organized and actively provide input on different areas of the project such as datasets and source code to further analyze datasets.

\*D3.js: D3.js is a JavaScript library for manipulating documents that are data based. D3.js will allow the team to have interactive data visualizations in web browsers. This data-driven documents library will allow the team to bind arbitrary data to DOM. The API contains functions such as math, arrays, geometry, color, etc. to aid in the findings of data based documents. With massive amounts of data D3.js will provide a way for said data to be visually represented.

\*As a team we are nervous to use this library as various team members have never worked in JavaScript before and no one has heard of D3.js so this is out of the teams comfort zone. We are unclear what other libraries would be appropriate and so as a result we will move forward with D3.js as you have strongly recommended.

**7) Tools**

Python, PySal, Matlab, R-Studio

**8) Processes**

Not Applicable

**9) Experimentation**

The steps we have taken include identifying a research area and writing a research proposal. Next, we will write a literature review, do data collection, data analysis, our first draft, second draft, the final paper, and share our findings through a presentation. The data we will use is regarding prison systems and the different aspects of prisons. We plan to experiment and analyze by using many resources to gather the data in terms of overcrowding in prisons, separation by gender, and extended sentences. We plan to provide a large basis of research and data analysis while using graphs to display the data.

[https://crime-data-explorer.fr.cloud.gov/](https://crime-data-explorer.fr.cloud.gov/downloads-and-docs)

[1] “Crime Data Explorer,” Federal Bureau of Investigation. [Online]. Available: <https://crime-data-explorer.fr.cloud.gov/>, Accessed on: February 28, 2021.

A detailed collection of data related to crime within the U.S. The data is filter between each state and year. The data contains information such as the demographics of the criminal, type of crime committed, and detailed information for the law enforcement as well. There is also non-filtered data related to specific crimes along with the arrest data of certain demographics.

<https://www.bjs.gov/index.cfm?ty=pbdetail&iid=7106>

[2] A.Carson, “Prisoners in 2019,” Bureau of Justice Statistics, Washington, District of Colombia, United States. [Online]. Available: <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=7106>, Accessed on: February 28, 2021.

A report from the Bureau of Justice that contains a collection of tables that track change in prison populations from 2019, showing changes from 2009-2019, that contains information pertaining to counts of prisoners on both the federal and state levels. The report also has other pertinent information such as the demographics of the prisoner’s, offense characteristics, and even imprisonment rates across the states.

<https://www.bjs.gov/index.cfm?ty=pbdetail&iid=7226>

[3] A. Beck, “Race and Ethnicity of Violent Crime Offenders and Arrestees, 2018,” Bureau of Justice Statistics, Washington, District of Colombia, United States. [Online]. Available: <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=7226>, Accessed on: February 28, 2021.

A report on of nonfatal violent crime in 2018 that compares data from the Bureau of Justice Statistics to data from the FBI’s Uniform Crime Reporting to examine the relationship between crime involvement along with race and ethnicity.

<https://www.bjs.gov/index.cfm?ty=pbse&sid=5>

[4] L. Maruschak, “Correctional Populations in the United States, 2017-2018,” Bureau of Justice Statistics, Washington, District of Colombia, United States. [Online]. Available https://www.bjs.gov/index.cfm?ty=pbdetail&iid=7026:, Accessed on: February 28, 2021.

A report from the Bureau of Justice that tracked information on populations supervised by adult correctional systems across the USA from 2017-2018. It contains information regarding incarceration rates for each state as well as showing change in the rate from 200-2018.

**10) Expected Deliverables**

A complete literature review of all sources used throughout the project.

An initial and secondary draft of the report will be provided along with the final report that will present the need & purpose of the research, present a summary of all the data used in the report, a brief explanation of how we built our interactive dashboard and how it manages all the data, a demonstration of how our dashboard is used to analyze said data, and the conclusions drawn from our analysis.