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

Capstone Weekly Status Report

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

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

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

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

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

<https://github.com/nat642/CS475.git>

**REPORT WEEK 5 (MARCH 22- MARCH 29)**

1. Weekly Accomplishments

* Jonathan: Moved towards analyzing data regarding crime rates in New York City and nationwide. When regarding the statewide crime, there were similarities between violent crime and the weapons used to commit it.
* Juan Pablo worked on a base format (.csv file), that will be processed in D3.js, that will consolidate the data of states and territories for each year.
* Natalie focused on reading the articles from the preliminary lit review and wrote an annotated bibliography for the article “Mass Incarceration: The Whole Pie 2020.” In addition, I went through previously written annotated bibliographies to proofread and adjust accordingly.
* Lauren read through two articles: “Jails, prisons, and the health of urban populations: A review of the impact of the correctional system on community health” for the literature review and has started to analyze how the sources relates to others read.
* Alex is helping setting up the dashboard with D3.js through GitHub, still studying a looking over example code of how this would be set up in GitHub

1. Problems/ Issues
2. Deciding on how to handle the states with a few missing statistics, or in a couple instances missing many categories.
3. Combining multiple datasets and fixing bugs is my primary issue. I am currently trying to do linear regression with the dataset that I currently have, but there are errors that prevent me from doing so.
4. Since GitHub is not a platform I use often, Alex is still trying to figure out the coding using GitHub D3.js and will coordinating with teammates on what the dashboard will look like.
5. Next Week’s Planned Work

* Finish consolidating the data tables for the remaining years.
* Continue reading through articles for the literature review while analyzing how the articles relate to one another as well as the project datasets.
* Continue writing annotated bibliographies for the found resources and incorporating them into a paper to give an understanding of the overview of our research project.
* To successfully analyze the nation-wide dataset based on crime activity from 1969-2016. This dataset may allow us to provide a perspective that will allow us to identify key factors of crime and why our prison population is so high.
* Implement code for D3.js onto GitHub for the dashboard.

1. Time Log

Hani: 3 hours

Wallace: 2 hours

Chacon: 2 hours

Thompson: 2 hours

Blank: 2 hours

Total Weekly Team Hours: 11 Hours

**REPORT WEEK 4 (MARCH 15- MARCH 21)**

1. Weekly Accomplishments

* Jonathan: Moved towards analyzing data regarding crime rates in New York City and nationwide. When regarding the statewide crime, there were similarities between violent crime and the weapons used to commit it.
* Juan Pablo created base template for data visualization with D3.js, representing general numbers from each state.
* Natalie focused on reading the articles from the preliminary lit review, and wrote an annotated bibliography for the article “Mass Incarceration: The Whole Pie 2020.”
* Lauren read through two articles: “Inmate Society in the Era of Mass Incarceration” and “Human Cattle: Prison Overpopulation and the Political Economy of Mass Incarceration” for the literature review and has started to analyze how these two sources relate to each other.
* Alex is helping setting up the dashboard with D3.js through GitHub, still studying a looking over example code of how this would be set up in GitHub

1. Problems/ Issues
2. Handling the different number of data sets (.csv files) and representing them with D3.js.
3. Combining multiple datasets and fixing bugs is my primary issue. I am currently trying to do linear regression with the dataset that I currently have, but there are errors that prevent me from doing so.
4. Since GitHub is not a platform I use often, Alex is still trying to figure out the coding using GitHub D3.js and will coordinating with teammates on what the dashboard will look like.
5. Next Week’s Planned Work

* Consolidate data sets into one.csv as well as include data from U.S. territorial areas into the consolidated data sets.
* Continue reading through articles for the literature review while analyzing how the articles relate to one another as well as the project datasets.
* Continue writing annotated bibliographies for the found resources and incorporating them into a paper to give an understanding of the overview of our research project.
* To successfully analyze the nation-wide dataset based on crime activity from 1969-2016. This dataset may allow us to provide a perspective that will allow us to identify key factors of crime and why our prison population is so high.
* Implement code for D3.js onto GitHub for the dashboard.

1. Time Log

Hani: 3 hours

Wallace: 2 hours

Chacon: 4 hours

Thompson: 3 hours

Blank: 2 hours

Total Weekly Team Hours: Hours

**REPORT WEEK 3 (MARCH 8- MARCH 14)**

1. Weekly Accomplishments
2. Jonathan has downloaded a visualizations software, GeoDa. This software has allowed the visualization of the data and allowed it to display its relationships.
3. Juan Pablo: Began implementing datasets into D3.js
4. Natalie began analyzing sources for the literature review, beginning with research on the effects of overcrowding.
5. Lauren began to organize materials for the literature review by preforming a general overview and providing a rough annotated bibliography for six of the articles.
6. Alex started studying D3.js, as he is not entirely familiar with GitHub and JavaScript programs.
7. We worked on gaining knowledge on how the software is used, the abilities it has, and getting a basic understanding of how we can use this software to create a dashboard to display our data.
8. Problems/ Issues
9. Jonathan: GeoDa is a visualization software that allows users to plot, analyze, and interpret data in a way that allows users to easily understand it. The issue that is arising from it is the search for suitable shapefiles. Most shape files that are created do not use a standard identification column. I have been able to find two types of shapefiles that do work luckily, but now the issue is combining the data per county/state with the shapefiles data as well. Currently attempting to combine them through the excel sheets and csv files.
10. Next Week’s Planned Work
11. Jonathan: Combining the data from the csv and excel datasets with the shapefiles and their identification columns.
12. Reformat the data documents that will be used by D3.js.
13. Continue analyzing sources from the preliminary lit review for the literature review.
14. Start thoroughly reading articles in preparation for lit review.
15. Help set up the dashboard using D3.js.
16. Time Log
17. Hani: 2 hours
18. Jonathan: 4 hours
19. Juan Pablo: 3 hours
20. Thompson: 3 hours
21. Alex: 1 hour
22. Total Weekly Team Hours: 13 Hours

**REPORT WEEK 2 (MARCH 1- MARCH 7)**

1. Weekly Accomplishments
2. Natalie met with the group via Zoom to discuss goals, revise the proposal, set a plan for responsibilities, as well as set up a GitHub repository.
3. Jonathan: Found basic relations between prisons located throughout the country and their prisoners. This data is informational but needs more attention when analyzing.
4. Juan Pablo: Reformatting and cleaning up data tables. Began practicing with D3.js to experiment with different representations for the data.
5. Alex Blank: Worked alongside Juan and John to clean data tables, attempted to clean my own data that provided last week.
6. Lauren met with the group via Zoom to discuss goals, revise the proposal and summary sheet, and set a plan for responsibilities. Lauren also researched new tools for the team to explore for creating a dashboard.
7. Problems/ Issues
8. Natalie did not face any problems or issues this week.
9. Jonathan: Geopandas and the importing of shape files are creating error messages regarding the new update. Currently attempting to resolve this issue.
10. Juan Pablo: Still learning and familiarizing myself with the library.
11. Alex Blank: The data I provided last week was missing and the data that was provided was too messy to be used.
12. Lauren: The tools listed previously were not appropriate for displaying our dashboard.
13. Next Week’s Planned Work
14. Natalie will begin the lit review by analyzing sources and beginning the introduction of our research project.
15. Jonathan: To resolve the issue regarding Geopandas and begin executing more in-depth calculations.
16. Juan Pablo: Implement all current data sets using the D3.js library.
17. Alex Blank: Help Juan and Jon with implementing data, should be a good idea to start familiarizing myself with D3.js.
18. Lauren will begin the lit review and discover new sources for our research project.
19. Time Log
20. Hani: 2 hours
21. Jonathan: 4 hours
22. Juan Pablo: 2 hours
23. Alex Blank: 2 hours
24. Thompson: 2 hours
25. Total Weekly Team Hours: 12 hours

**REPORT WEEK 1 (FEB 22- FEB 28)**

1. Weekly Accomplishments
   1. Natalie completed the Project Proposal and Project Description portion of the project proposal assignment. In addition, cooperated with the team to organize a schedule and find a new topic for the research project to complete.
   2. Lauren: completed the Resource Requirements portion of the project proposal assignment. Actively communicating with the team through group chat and meetings about ideas and plans of action.
   3. Jonathan: Created three large datasets from Data.gov
   4. Alex: Located possible datasets regarding crimes specified by state and year, the files contain specific crime data that includes every state in 2019.
   5. Juan Pablo: Help Alex find data sets relating to incarceration rates and other prison inmate demographic details.
2. Problems/Issues
   1. Natalie struggled with brainstorming potential research project ideas that would be sufficient enough to complete a research project on.
   2. Lauren: struggled with maintaining the scope of the project until meeting with professor took place.
   3. Jonathan: Struggling to clean and process data for the upcoming data analysis.
   4. Alex: The data found itself is messy, some of the information goes by an ID, which it makes it difficult to decipher.
   5. Juan Pablo: Not much difficulty finding sources for datasets or for literature review.
3. Next week’s planned work
   1. Natalie will set up a repository for the team to share resources and keep up to date with the project.
   2. Lauren: will continue to provide support to team mates as well as begin review the articles for the Preliminary Literature Review
   3. Jonathan: Finish cleaning the data and execute it.
   4. Alex: Start to process the data in R and run datamining.
   5. Juan Pablo: Begin cleanup of data tables.
4. Time log

Total Number of Hours worked on project this week per team member: Hani: 2 hours; Thompson: ~ 2 hours; Jonathan: ~1.5; Alex: ~1 hour; Juan Pablo: ~2. Total Weekly Team Effort: ~6.5 hours