Hate Groups and Crimes Correlation with State Demographics

Leigh Wright – [lwrig25@lsu.edu](mailto:lwrig25@lsu.edu)

Jeremy Kelly – [jkell89@lsu.edu](mailto:jkell89@lsu.edu)

April 22, 2019

Table of Contents

1. Cover Page
2. Table of Contents

       3.  Introduction

                        Motivation

                        Project Description

Division of Roles

4.   System Design

                        Chosen Big Data Frameworks

                        Chosen Datasets

5. Detailed Description of Components

                        Each Component Description

            Evaluation & Test

Software Manual

                        Test Environment

                        Test Cases

(*only for projects using Raspberry Pi*) some photos showing your hardware configuration

6. Conclusion

Introduction

Our motivation for the project was to try and find out what the origin of hate crimes and groups. Given the increasing political polarization that has been going on in this country, we felt this was an important topic worthy of exploration. Specifically, we wanted to find correlations between demographics and various *types* of hate groups (i.e. black nationalists compared to the Ku Klux Klan), if possible. We also wanted to establish if there were any correlations between demographics of a state and the hate crimes committed there.

About halfway through the project, a group member decided to break off and do his own project. Jeremy Kelly was responsible for building our four-layer Raspberry Pi cluster. This includes setting up cluster environment, the operating system, installing fans for each node to ensure they do not overheat, and some of the programming as well. Building the layers was Jeremy’s most time-consuming task.

Leigh was responsible gathering, cleaning, and documenting our data in addition to programming. Some of the data sets were not readily usable and required a significant amount of cleaning; the data cleaning is what took up the overwhelming majority Leigh’s time.

System Design

We used a Raspberry Pi cluster in place of a big data framework. Our project was originally done in python within Jupyter Notebook before we covered Spark in class. Much of the data cleaning was done early on in Jupyter Notebook because python is still very efficient for cleaning data.

The datasets themselves were far more complicated; we used data from 538 that combined multiple datasets. Besides that data, we also used a large data set from the census bureau, in addition to a separate dataset from the FBI. The census bureau data especially needed to be cleaned extensively before it was useful. To establish where hate groups are located, and what type of hate group they are, we used data from the Southern Poverty Law Center’s Hate Map. The map is an interactive map available on the SPLC’s website.

Here are links for the datasets

FBI - <https://ucr.fbi.gov/hate-crime/2017/topic-pages/tables/table-12.xls>

FBI - <https://ucr.fbi.gov/hate-crime/2017/tables/table-12.xls>

SPLC Hate Map: <https://www.splcenter.org/hate-map>

American Community Survey (census bureau): <https://www.census.gov/programs-surveys/acs>

Description

Raspberry Pi 3 Model B+ (X4)

4 Layer Raspberry Pi Cluster Case with Cooing Fans for each layer

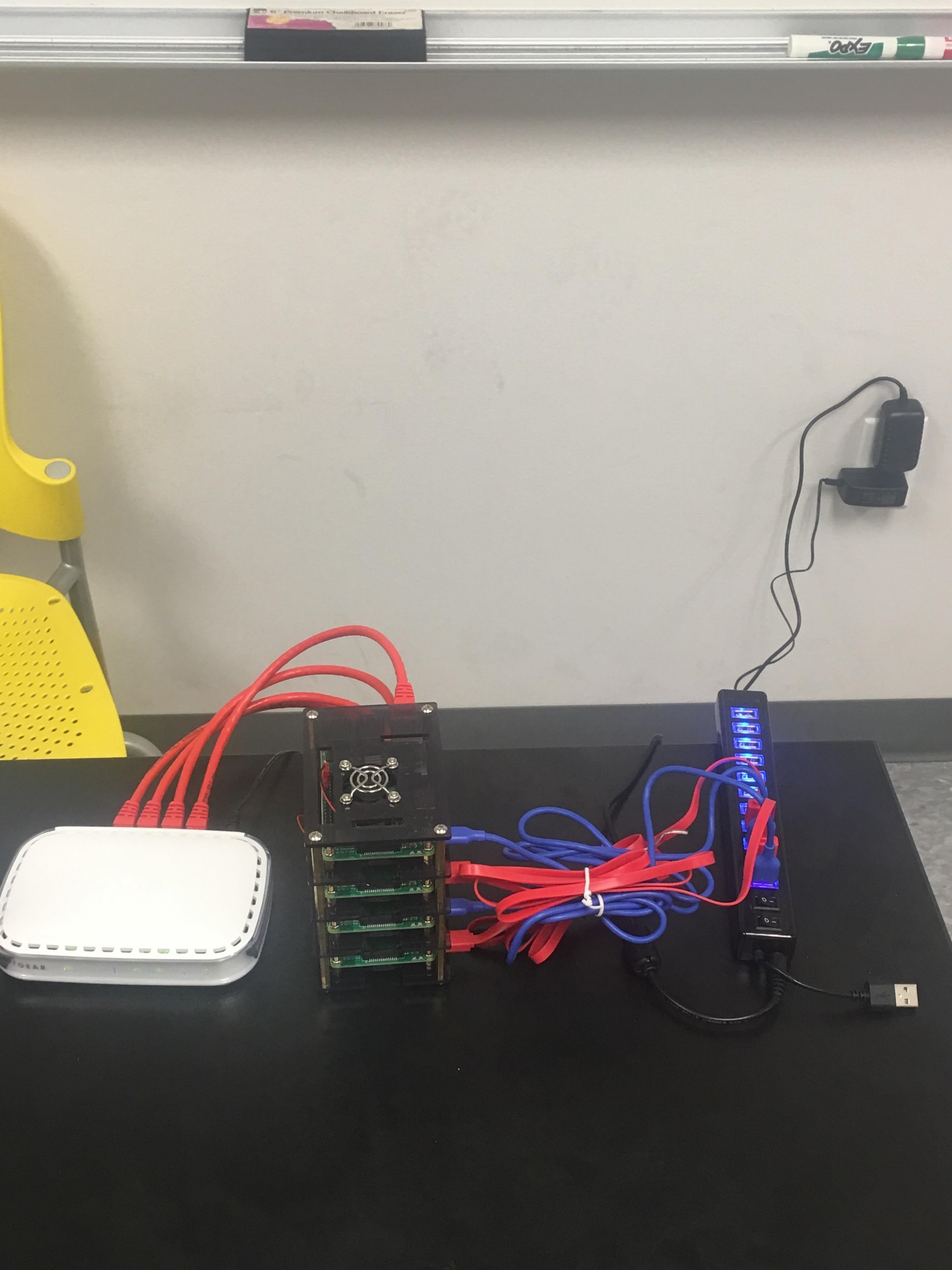
Satechi 12 port USB hub & 2 control switches

Ethernet Cable (x4)

Micro USB cables(x4)

Netgear 5 -port switch

See picture below



Evaluation and Test

We used our sources as csv’s and put them into Jupyter Notebook. That was to clean our data. We then made new csv’s and put them into Jupyter Notebook again in order to find correlations and to visualize the data.

Conclusion

To put it bluntly we do not consider our project to be a great success, although much of this is due to things outside of our control. Hawaii had some key data missing, and other states had populations that were so small we left them out because they skewed the data too much.

For some other states, such as Mississippi, we do not believe all the data was reported accurately. For example, the state of Mississippi had only one recorded hate crime. This is likely due to hate crimes not being reported and/or recorded, thus resulting in insufficient or inaccurate data. However, there were other times where we could see a correlation; the share of voters who voted for Donald Trump in 2016 had strong positive correlations between white people under the poverty line, as well KKK and Neo-Confederate hate groups.

Other times the data contradicted itself; the share of Trump voters had positive correlation with the above stated groups but had a negative correlation with White Nationalist hate groups. This could be for numerous reasons, which go beyond the scope of our project.

|  |  |
| --- | --- |
| Project.py (from the midterm) | ~75 lines of code |
| CSC\_4740\_Final\_Project | 20 lines of code |