

W205 - Final Project

Andrew Burrage

Chicago & San Francisco Open Data

Chicago Crime and 311 Data

San Francisco Crime and 311 Data

What is the relationship between crime rate of a district and time to close 311 requests (street lights out) in same district?

I will look at crime and 311 data from Chicago and San Francisco.

Acquisition and Organization of Info

City of Chicago Data Portal

- Crimes - 2001 to Present
- 311 Service Requests - Street Lights - All Out

San Francisco Open Data

- SFPD Incidents from 1 January 2003
- Case Data from San Francisco 311

Architecture

Very similar to Exercise 1

HDFS-backed Data Lake

Transform with Hive's SQL interface

Analysis and Results

Will use R for analysis.

Questions to answer:

- What districts in Chicago and San Francisco have highest and lowest crime rate?
- What districts have the best & worst closure times in Chicago and San Francisco?
- What is the relationship between crime rate and 311 closure time?
- How comparable are the results found in Chicago and San Francisco?