

Crime and Education in Chicago 2002-2017

Joe Rickard Ryan Rouleau
Kylee Budai

University of Colorado Boulder

February 2017

Project Description

- General spatial trends of crime over time
- Trends in categorical frequency
- Supervised learning to predict trends in areas
- Correlation between crime and education

Prior Work

- Governmental reports
- https://en.wikipedia.org/wiki/Crime_in_Chicago

Chicago Datasets

- Crime Data

<https://catalog.data.gov/dataset/crimes-2001-to-present-398a4>

- Census Data

<https://data.cityofchicago.org/Health-Human-Services>

Proposed Work

- Reduce dimension of database.
- Encode categorical attributes
- Store data in well-indexed relational table.
- Ensure consistent formatting for analysis.
- Run preliminary analysis and remove outliers/high leverage points.
- Analyze how categorical frequency is correlated to time.
(Dependent)

Tools

- git
- Python
- R
- C
- C++
- an SQL distro
- Excel
- Slack
- Tableau



How to Evaluate Results

- Generate plots that clearly demonstrate our statistical results.
- Randomly subset data and see how well our model can predict those data.
- Publish them, obviously.

Questions?

