

1) One paragraph description:

Our project will explore the intersections of food access and crime, using neighborhood-level data from neighborhoods in Chicago. Combining one member's interest in food security and another's interest in crime and safety, we will explore potential relationships between the two themes. Using data primarily from the Chicago Data Portal and Chicago Health Atlas, we will collect the data using APIs. Chicago Data Portal has neighborhood-level data we will primarily use for demographics, income and poverty rates, and crime rates. Chicago Health Atlas also has neighborhood-level data for several food and health related indicators such as food access, access to fruits/vegetables, and eligible households not receiving food stamps, amongst others. We hope to show these correlations using regression analysis using pandas and data visualizations.

2) Expected data source, point person, completion date

Chicago Data Portal (API): Ryoya Hashimoto (completed)

- The number of crimes
- Population data

Chicago Health Atlas (API): Takayuki Nitta (completed)

- Low Food Access rate
- Adult food and vegetable servings rate
- Poverty data

3) Brief sketch of the remaining work, associated point person, completion date

**Data processing:**

Creating a variable for crime rate: Ryoya (currently reported as an absolute number) by dividing the number of crimes by the population (estimated completion by March 6)

Joining/concatenating data files based on neighborhood id: Taka (estimated completion by March 6)

**Analysis:** Regression analysis and decision tree prediction analyses will allow us to explore the explanatory relationships between food access and crime in multiple ways. It will be interesting to compare the results of each prediction analysis, which can be done by splitting the data into training and test sets. Clustering analysis to categorize neighborhoods, which can then be used in visualizations as well.

Regression Analysis: Ryoya (estimated completion by March 6)

General regression concept:  $\text{Crime\_Rates} \sim \text{food\_insecurity} + \text{other demographic factors (eg race, age, gender, neighborhood)}$

Decision Tree: Liz (estimated completion by March 6)

Clustering: Ryoya (estimated completion by March 6)

**Visualizations:** Vincent (lead), with support from other group members (estimated completion by March 12)

Displaying mean, 25% quantile, 75% quantile, max, min for each indicator using bar charts

Mapping the level of each indicator by neighborhood with geopandas or Geodjango

**Virtual environment/shell script:** Liz (lead) with support from other group members if needed (estimated completion by March 12)

4) Additional Info:

Back-Up Data Sources if needed:

City Data

Greater Chicago Food Depository

Major Supermarket chains's location data

5) Additional Feedback: