

# Homework 1: Diagnostic

Jean Salac (salac@uchicago.edu)

April 3, 2018

GitHub Repo: <https://github.com/jeansalac/ml-ppol>

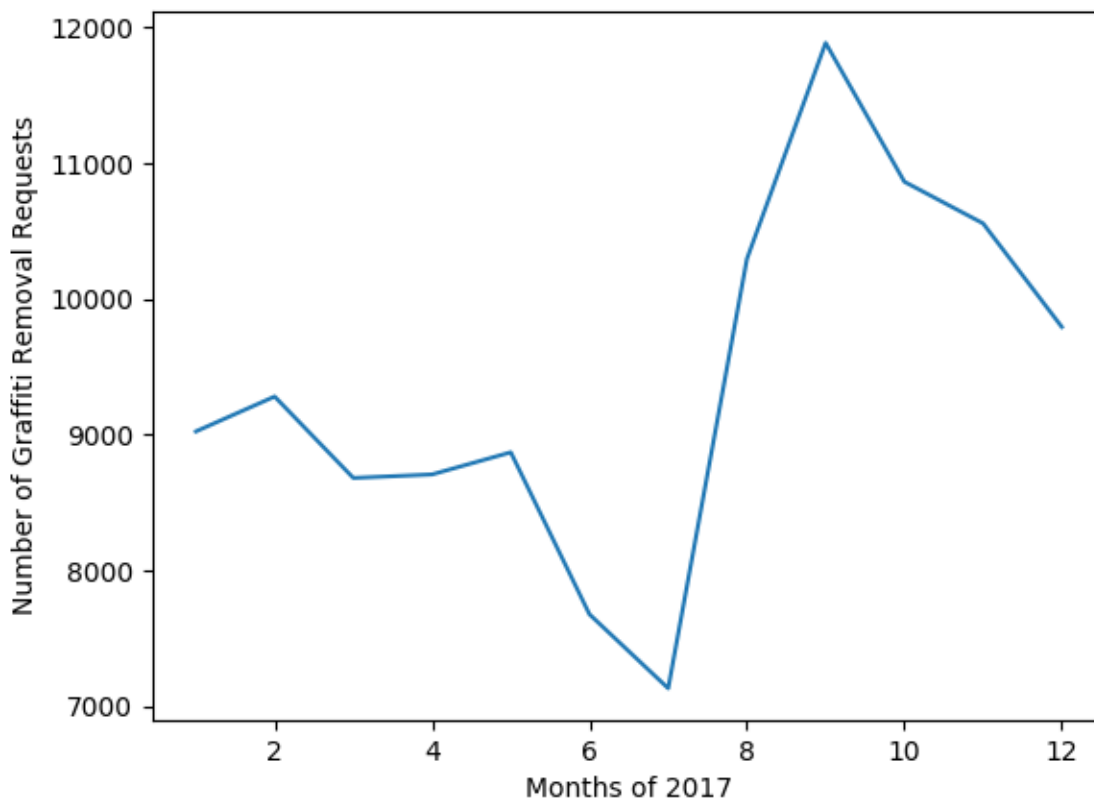
Collaborator: Yuliana Zamora

## 1 Data Acquisition and Analysis

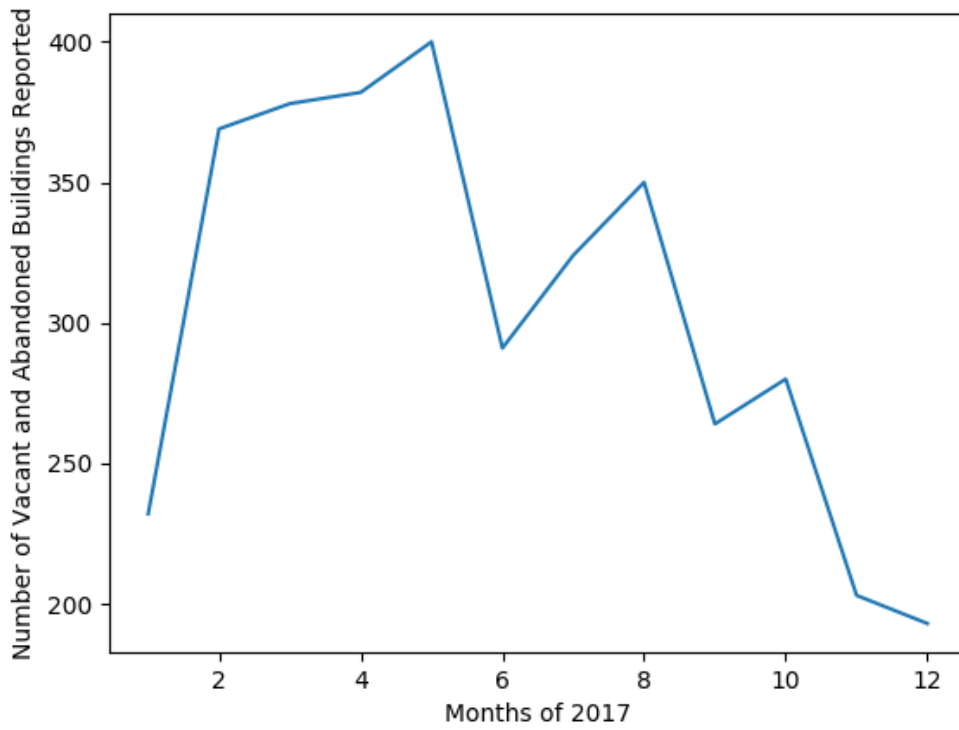
Requests per Month:

Requests	Mean Requests per Month	Standard Deviation
Graffiti Removal	9398.25	135.37
Vacant and Abandoned Buildings	305.5	72.35
Lights out in Alleys	2324.67	308.24

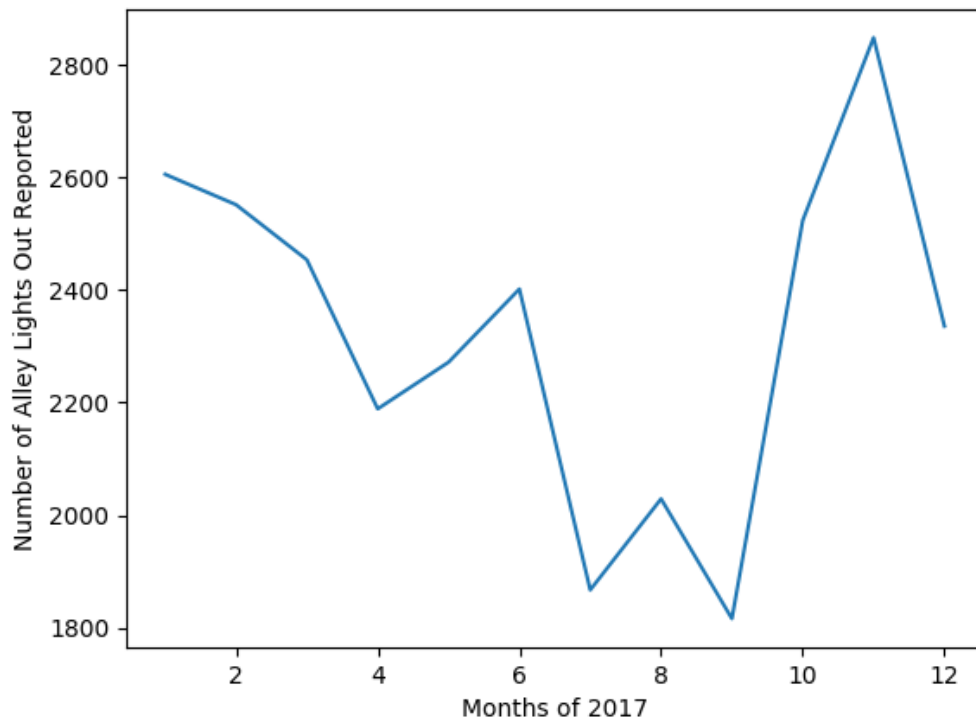
Number of Graffiti Removal Requests Every Month in 2017



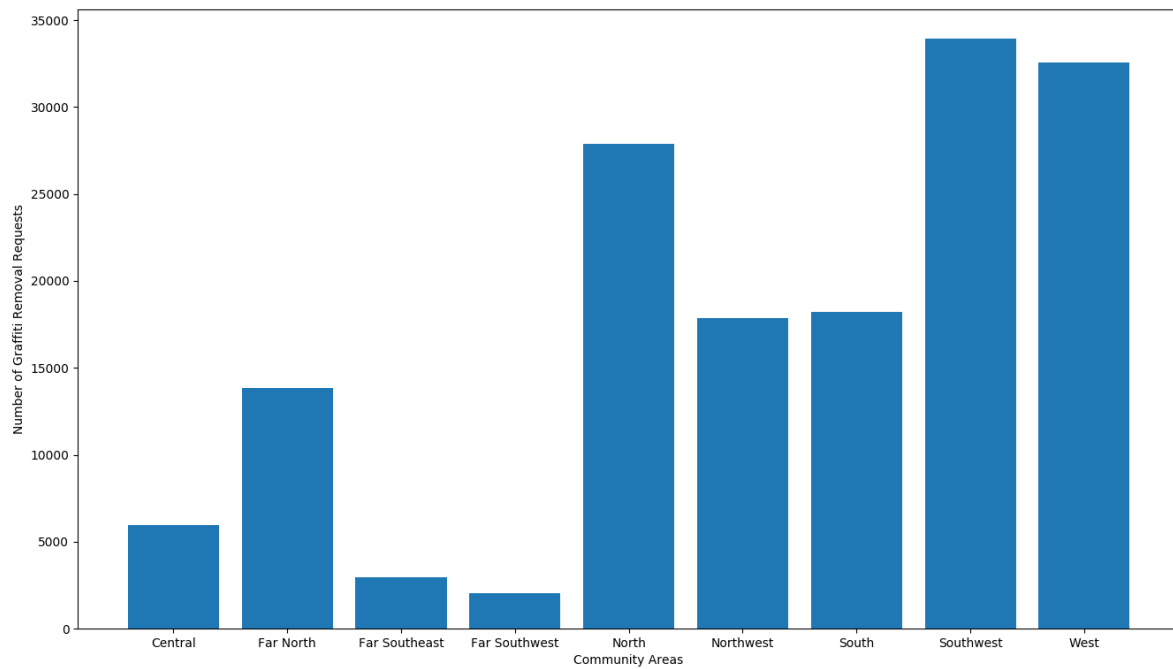
Number of Vacant and Abandoned Buildings Reported Every Month in 2017



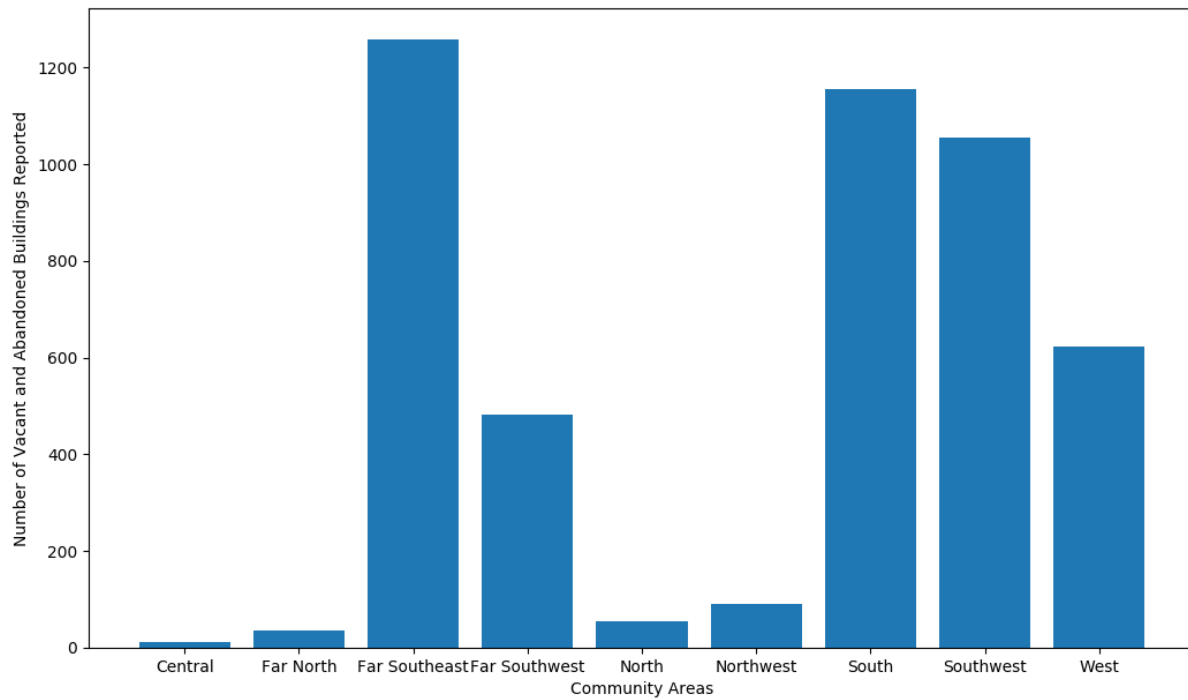
Number of Alley Lights Out Reported Every Month in 2017



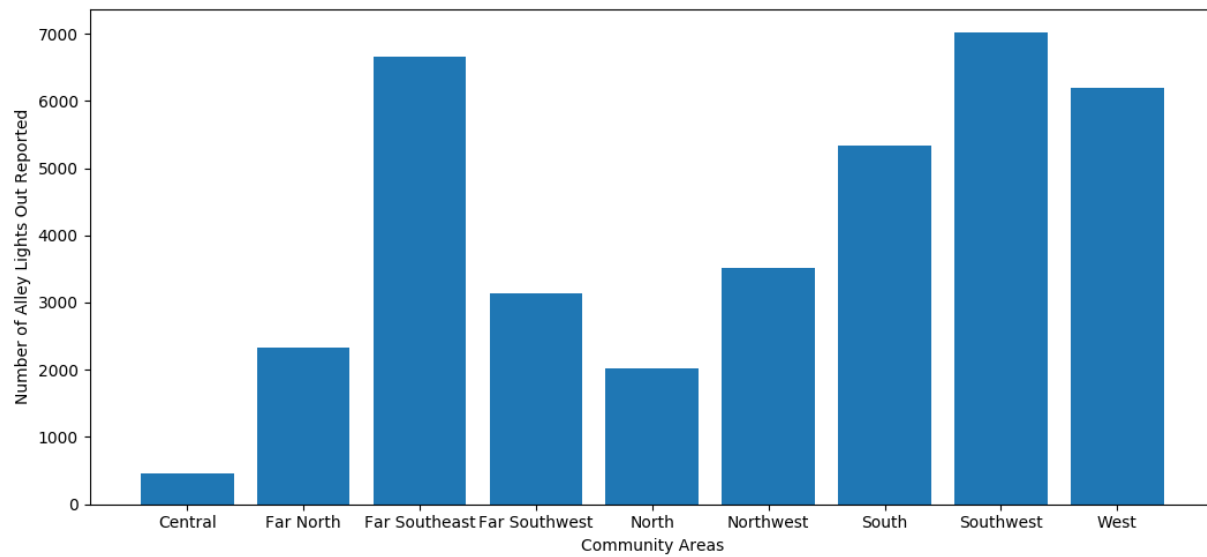
## Number of Graffiti Removal Requests for Each Community Area in 2017



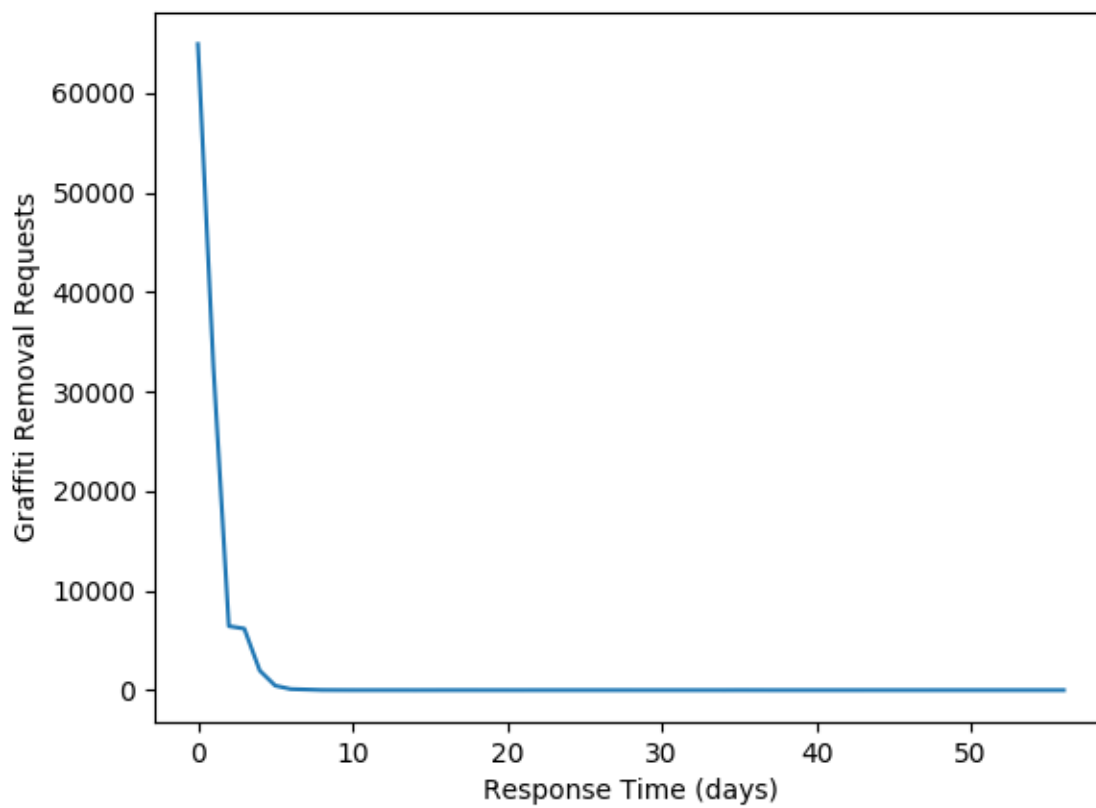
## Number of Vacant and Abandoned Buildings Reported for Each Community Area in 2017



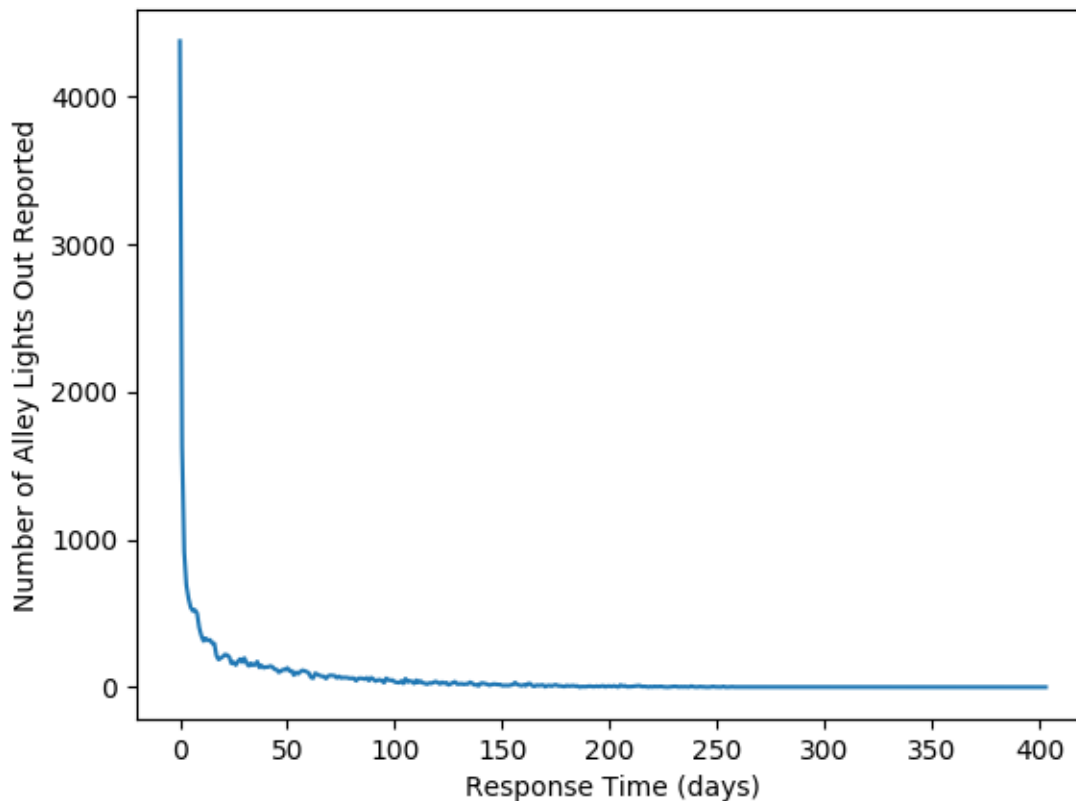
Number of Alley Lights Out Reported for Each Community in 2017



Response Time for Graffiti Reported in 2017



## Response Time for Alley Lights Out Reported in 2017

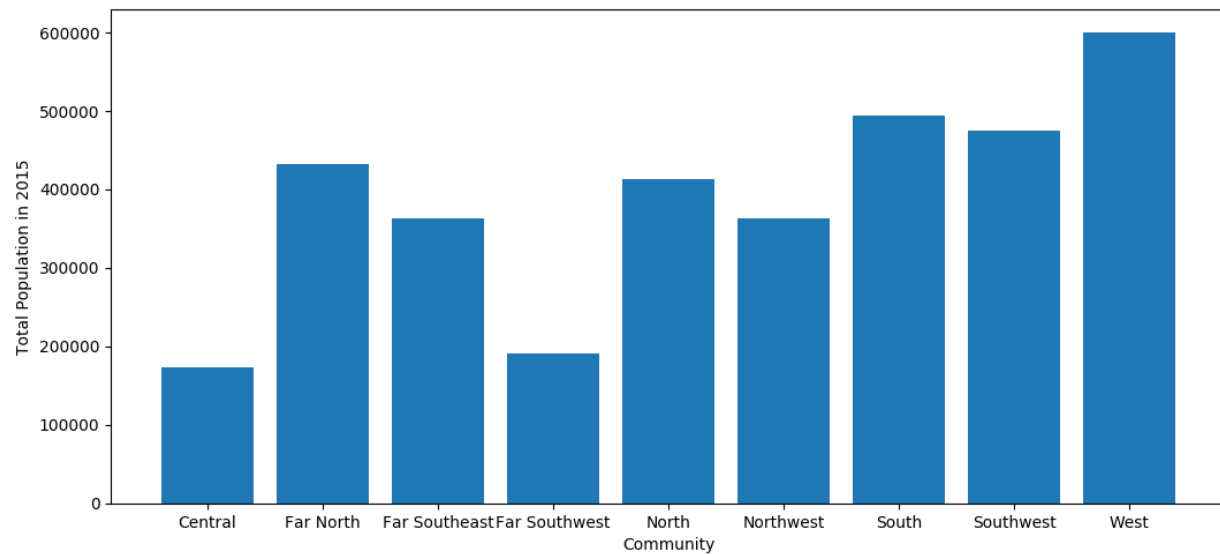


### Interesting Things Learned from the Data

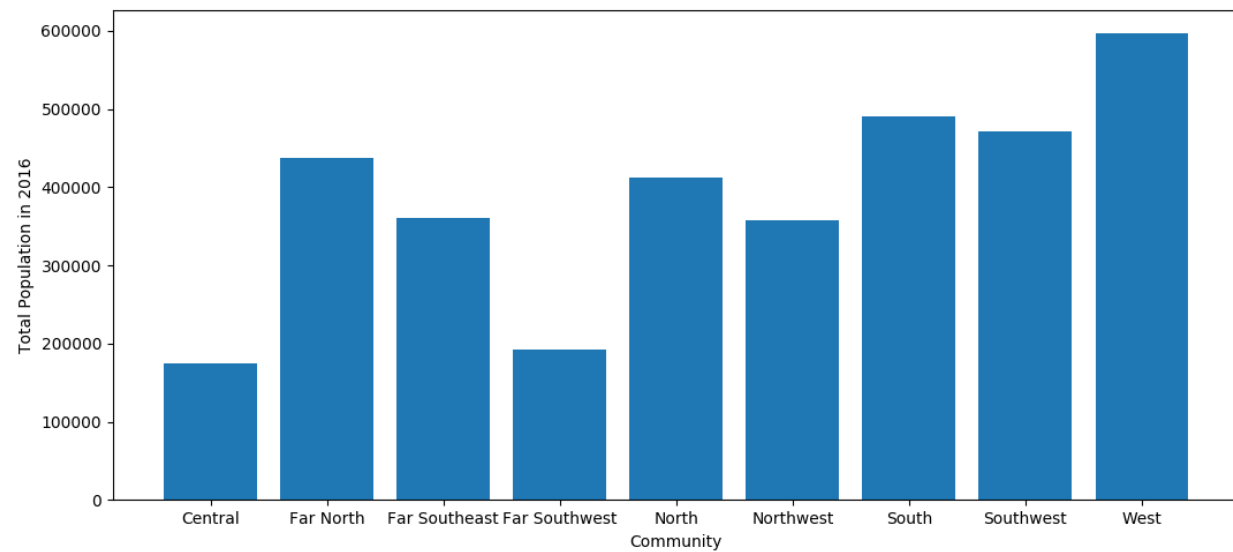
1. The number of graffiti removal requests and alley lights out reported dip in the summer months, while the number of vacant and abandoned buildings reported rise in the summer.
2. The greatest graffiti removal requests come from the North Side, Southwest Side, and West Side. This could be either a reflection of the amount of graffiti in those communities or the community's tolerance for graffiti.
3. The greatest number of vacant and abandoned buildings reported come from the Far Southeast, South, Southwest communities.
4. The greatest number of the alley lights reported are from the Far Southeast, South, and Southwest communities.
5. Most of the requests for graffiti removal were responded to in under 10 days, while most of the requests for the alley lights out reported were responded to in under 50 days. This discrepancy could be because it may be easier to clean up graffiti than to identify and replace which alley lights are out.

## 2 Data Augmentation and APIs

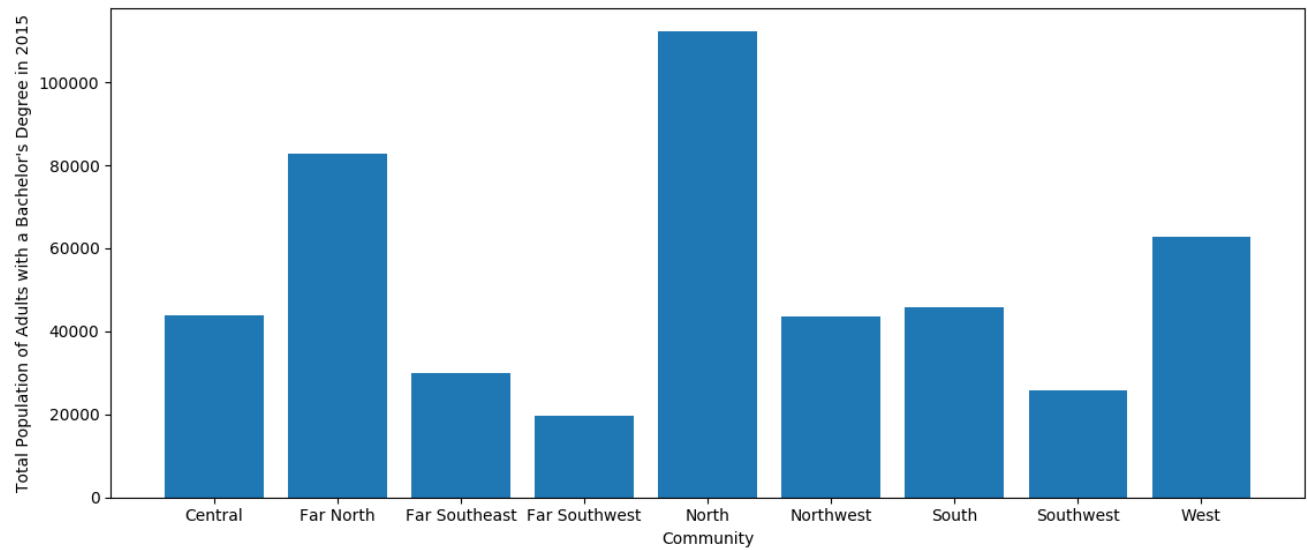
Total Population for Each Community in 2015



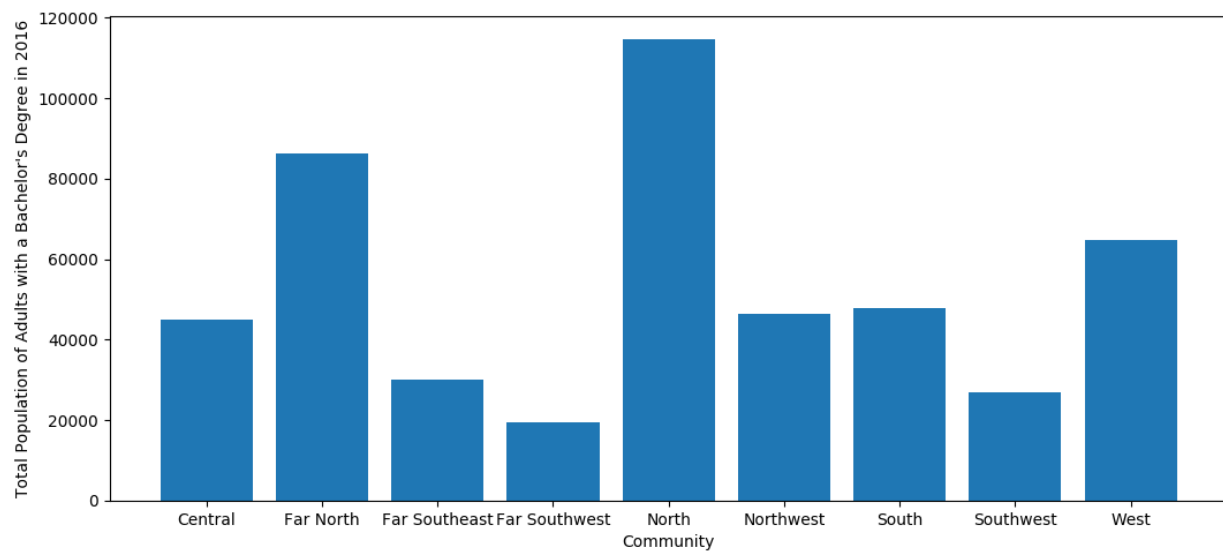
Total Population for Each Community in 2016



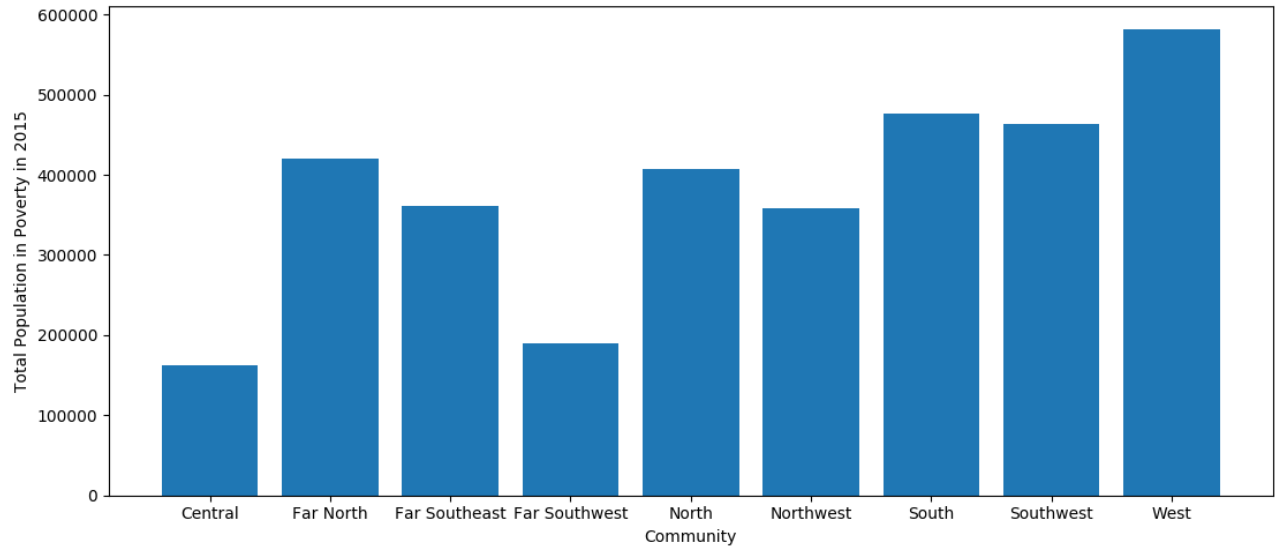
## Total Number of Adults with Bachelor's Degrees for Each Community in 2015



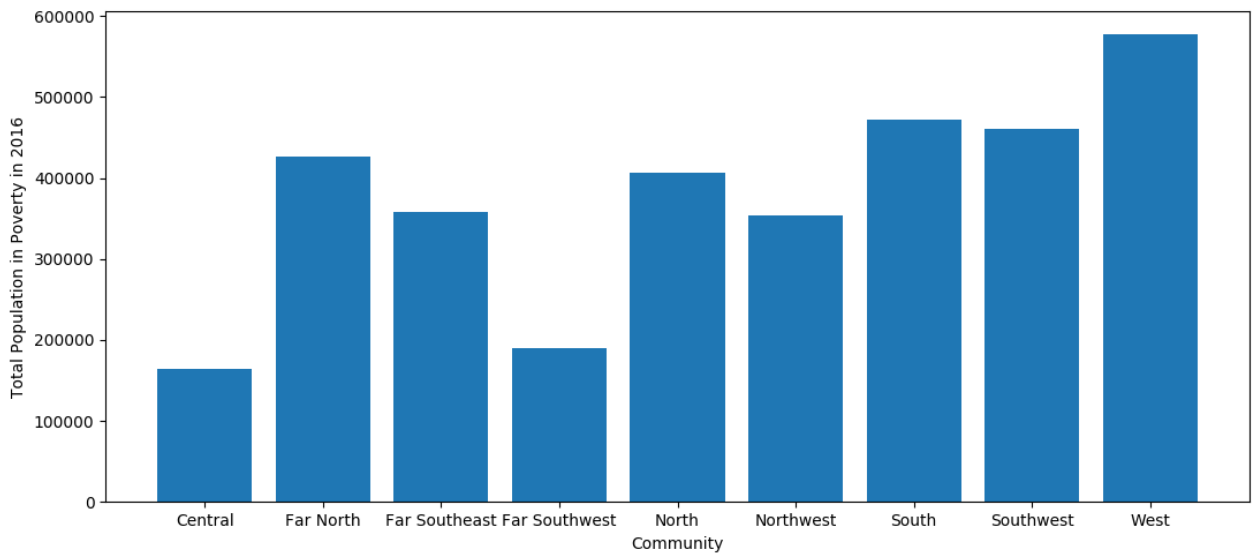
## Total Number of Adults with Bachelor's Degrees for Each Community in 2016



## Total Population in Poverty for Each Community in 2015

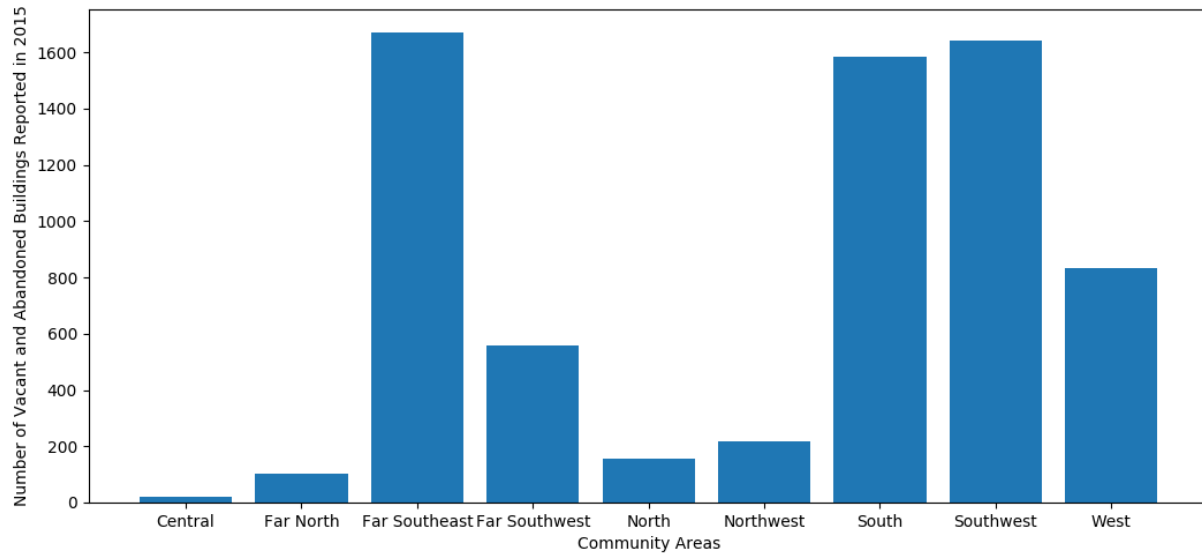


## Total Population in Poverty for Each Community in 2016

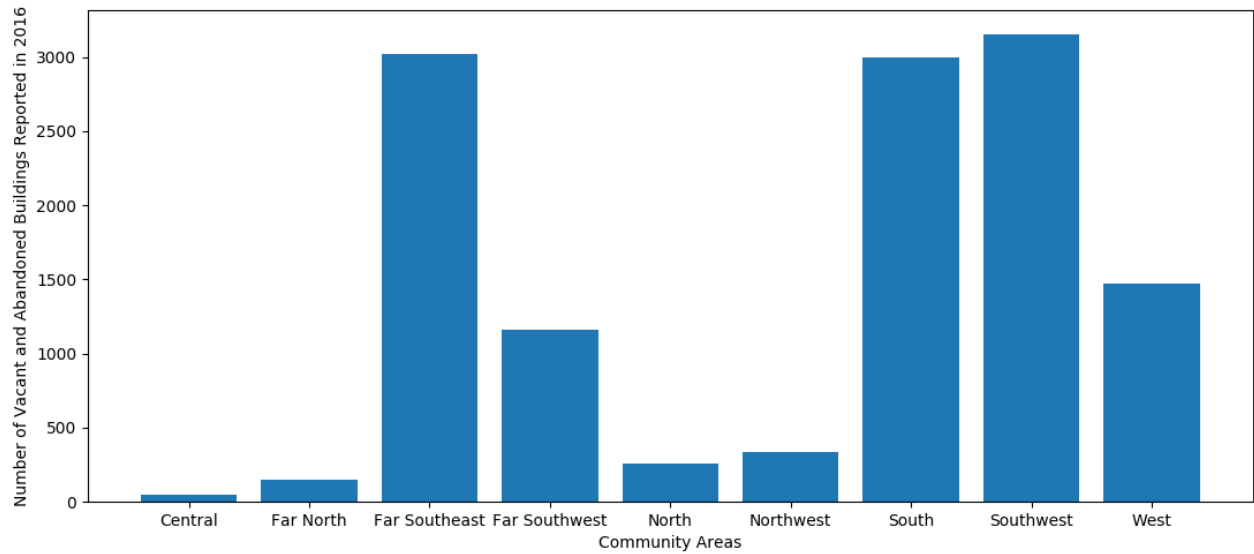




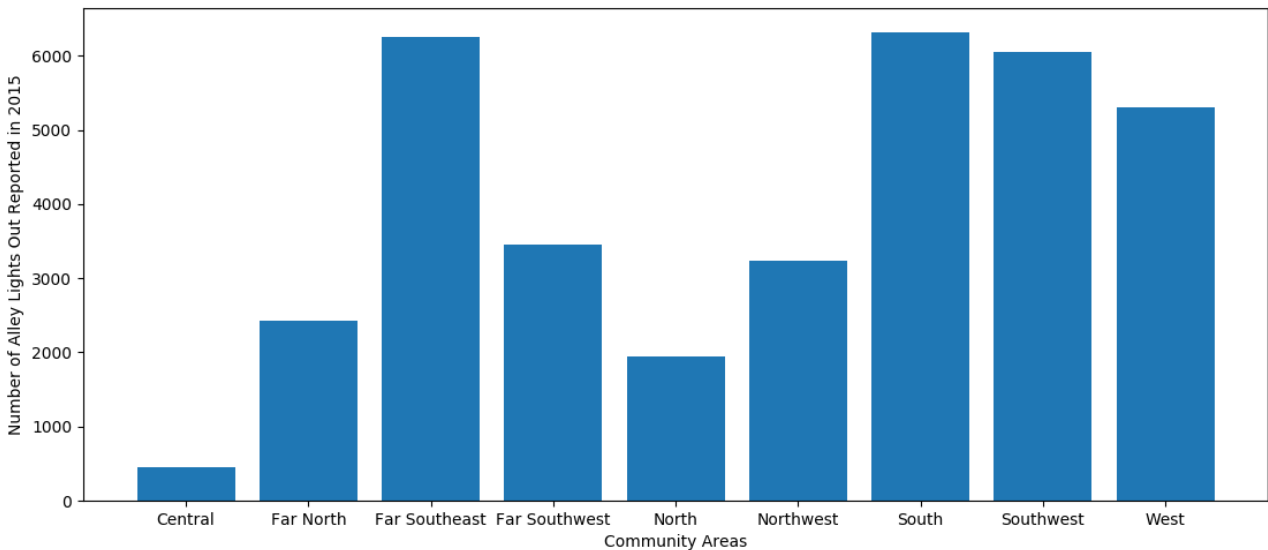
## Number of Vacant and Abandoned Buildings for Each Community in 2015



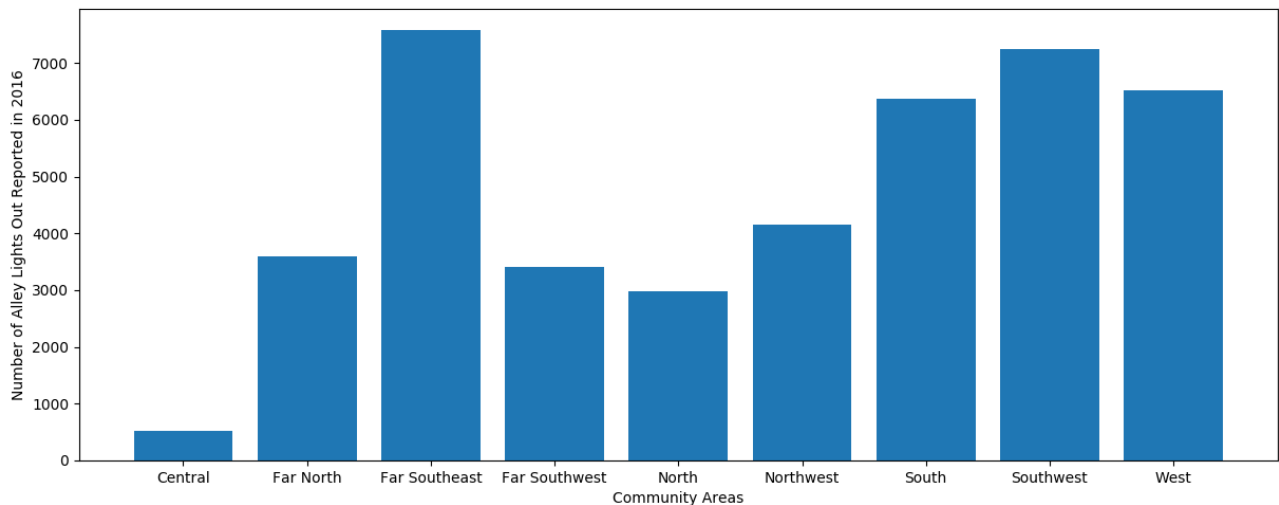
## Number of Vacant and Abandoned Buildings for Each Community in 2016



## Number of Alley Lights Out Reported for Each Community in 2015



## Number of Alley Lights Out Reported for Each Community in 2016



### Description

1. The types of communities that get more "Vacant and Abandoned Buildings" are those with a lower level of educational attainment and a higher population in poverty.
2. The types of communities that get more "Alley Lights Out" are also those with a lower level of educational attainment and a higher population in poverty. However, the discrepancy between the communities that get the most "Alley Lights Out" and the communities that get the least "Alley Lights Out" seem to be smaller than that of the communities with "Vacant and Abandoned Buildings".
3. The pattern does not seem to change significantly between 2015 and 2016. The graphs look very similar.
4. The community profiles that get "Vacant and Abandoned Buildings" and "Alley Lights Out" seem to be similar, but we could always use more demographic data to better classify these communities.

### 3 Probability

- Zip code for 3600 W Roosevelt Ave: 60624
- Zip codes for Uptown: 60613 and 60640
- The values came from the number of requests for the zip code 60624 over all the years.

#### A Probabilities of Requests

- Number of Graffiti Removal Requests (G) = 3274
- Number of Vacant and Abandoned Buildings Requests (B) = 3069
- Number of Lights Out in Alleys Requests (L)= 3244
- Total Number of Requests = 9587
- $P(G) = \frac{3274}{9587} \times 100 = 34.15\%$
- $P(B) = \frac{3069}{9587} \times 100 = 32.01\%$
- $P(L) = \frac{3244}{9587} \times 100 = 33.84\%$
- Thus, graffiti removal is the most likely request from Garfield Park.

#### B Garfield vs Uptown - with raw data

- Number of Graffiti Removal Requests from Garfield Park (GP) = 3274
- Number of Graffiti Removal Requests from Uptown (U)= 19182+17653=36835
- Total number of Graffiti Removal Requests = 40109
- $P(GP) = \frac{3274}{40109} \times 100 = 8.16\%$
- $P(U) = \frac{36835}{40109} \times 100 = 91.84\%$
- A graffiti removal request is  $\frac{36835}{3274} = 11.25$  times more likely to be from Uptown than Garfield Park.

#### C Garfield vs Uptown - without raw data

- $P(\text{Garfield}|\text{Graffiti}) = \frac{100}{100+160} \times 100 = 38.46\%$
- $P(\text{Uptown}|\text{Graffiti}) = \frac{160}{100+160} \times 100 = 61.54\%$
- A graffiti removal request is  $\frac{160}{100} = 1.6$  times more likely to be from Uptown than Garfield Park.