

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

2 Data Augmentation and APIs

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

- $P(GP) = \frac{3274}{40109} \times 100 = 8.16\%$
- $P(U) = \frac{36835}{40109} \times 100 = 98.14\%$
- 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

- $P(Garfield|Graffiti) = \frac{100}{100+160} \times 100 = 36.46\%$
- $P(Uptown|Graffiti) = \frac{160}{100+160} \times 100 = 61.54\%$
- A graffiti removal request is $\frac{160}{100} = 1.6$ times more likely to be from Uptown than Garfield Park.