Saptarshi Ghose

Discussed concepts with Natasha and Andrea

Homework 1 Write-Up

Problem 1

**Based on these summary statistics, tell me 5 interesting things you learned (about Chicago and the different neighborhoods) using the 311 data.**

1. Among the 311 requests in Chicago in 2017 for the service types we are working with, there were 112,779 requests for graffiti removal, 27,896 for alley lights out, and just 3,666 for vacant/abandoned buildings. As such, there were more than 4X as many graffiti removal reports than alley lights out reports. And there were more than 30X as many graffiti removal reports than vacant/abandoned building reports.
2. Considering all three types of 311 requests in 2017 we are exploring in this assignment, September, October, and November were the months with the most total requests by creation month. In contrast, June and July had the fewest total requests. In our limited dataset for just 2017, there were far more 311 requests in colder months than warmer months.
3. Of the 3148 vacant/abandoned buildings 311 requests without relevant missing data, 1267 (40%) had people still using the property (i.e. homeless people, children, gangs).
4. Among all explored 311 requests in 2017, the average response time (completion date – creation date) was 6.4 days.
5. 40,329 (36%) out of the 112,669 requests for graffiti removal in 2017 were for graffiti on metal surfaces. Unpainted brick had the next highest number of graffiti removal requests with 19,482 (17%).

Problem 2

**Based on this augmented data, provide some descriptive statistics to describe:**

**1. What types of blocks get “Vacant and Abandoned Buildings Reported”?**

**2. What types of blocks get “Alley Lights Out”?**

**3. Does that change over time in the data you collected?**

**4. What is the difference in blocks that get “Vacant and Abandoned Buildings Reported” vs “Alley Lights Out”?**

1. FINISH

Problem 3

**Assume you are running the 311 call center for Chicago. You get a call from 3600 W Roosevelt Ave.**

**A. Of the three types of requests you have data for, which request type is the most likely given the call came from 3600 W Roosevelt Ave? What are the probabilities for each type of request?**

**B. Let’s now assume that a call comes in about Graffiti Removal. Which is more likely – that the call came from Garfield Park or Uptown? How much more or less likely is it to be from Garfield Park vs Uptown?**

**C. Now assume that you don’t have access to all the raw data and you know the following things:**

**There are a total of 1000 calls, 600 from Garfield Park and 400 from Uptown. Of the 600 calls from Garfield Park, 100 of them are about Graffiti Removal. Of the 400 calls from Uptown, 160 are about Graffiti Removal. If a call comes about Graffiti Removal, how much more/less likely is it that the call came from Garfield Park versus Uptown?**