

BOSTON PHARMACY

JIA JIA SHEN, SHIWEI CHEN, YUANPEI WANG

JSHEN97@BU.EDU, LEOCHANS@BU.EDU, WANGYP@BU.EDU

CS 504 DATA MECHANICS

INTRODUCTION

It is said that we are living in an Age of Data. Data are literally omnipresent in our day-to-day life—Every Ad we see, every video or music recommendation we get, every choice we make—Data act behind the scene.

Project Pharmacy begins with a question that: What if we obtain one or two "host" datasets along with several "guest" datasets. Then before continue onto manipulating them, we first assume there is some kinds of relationships between the "host" and the "guests". Finally after performing some statistical or operational tests, will the outcomes bring us any interesting conclusions?

The project treated the CVS/ Walgreen pharmacy stores in central Boston area as the "host" datasets and several other interesting "guest" datasets as described in the following section.

DATASETS USED AND TRANSFORMATIONS

Google Places API (search Nearby)

- CVS, Walgreen, and 7Eleven stores within 15 km Boston area
- Boston supermarket or grocery store using google API

Data Mechanics IO

- Eviction Incidents and Crime Incidents in Boston Area
- Boston house_hold_income
- Boston poverty_rate

Transformation Implementations

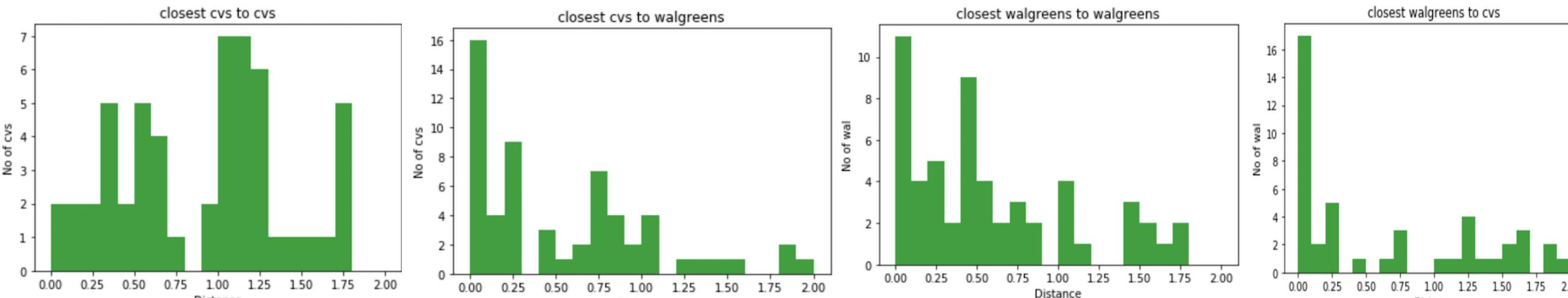
- walgreenCrime: combine the Walgreen dataset and crimes dataset
- walgreenEviction: combine the Walgreen dataset and eviction dataset
- correlationCVS: recorded the calculated results of the correlations between CVS ratings and Larceny cases, CVS ratings and Eviction cases
- countEvictionCrimeCVS: for every CVS pharmacy store, add two records. One for the number of Larceny cases nearby, and the other for the number of Eviction cases nearby
- cvsCrime: combine the Walgreen dataset and CVS dataset, filtering every crime except for Larceny
- cvsEviction: combine the CVS dataset and eviction dataset
- ratingCrime: recorded a pair of (ratings, number of larceny cases) for every CVS store
- ratingEviction: recorded a pair of (ratings, number of Eviction cases) for every CVS store

DATA VISUALIZATIONS AND ANALYSIS

We find that some CVS and Walgreen stores are next to each other. So for each CVS stores in Boston area, we find the closest Walgreen and CVS stores and record their distances, same for Walgreen stores. Results are shown below

CVS (miles)	Distance to closest Walgreen	Distance to closest CVS	Walgreen (miles)	Distance to closest Walgreen	Distance to closest CVS
Mean	1.32	1.06	Mean	0.74	0.57
Standard Deviation	1.40	0.66	Standard Deviation	0.70	0.59
Correlation	0.35	0.35	Correlation	0.23	0.23

The locations of Walgreen stores are intensive because it has lower mean value and standard deviation. Also, Walgreen are more likely to locate their stores closer to CVS stores

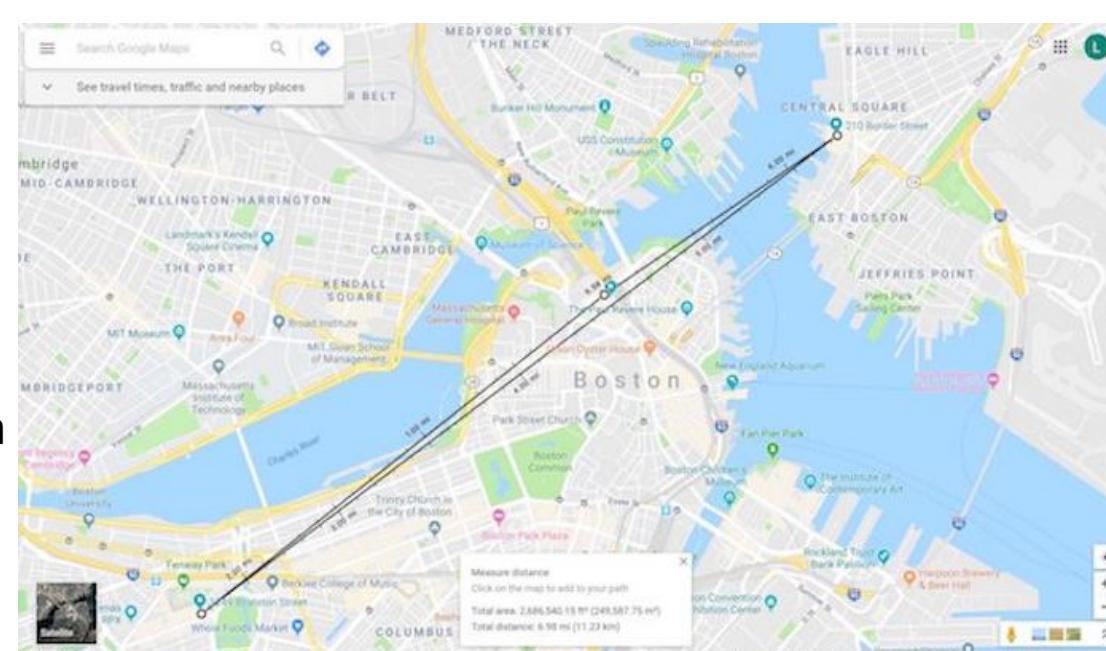


we plot the histograms shown above. It is interesting that there are around 16 pairs of CVS and Walgreen right next to each other which is consistent with our intuition. It is reasonable to believe that density of CVS and Walgreen stores is proportional with the density of population. Hopefully, we can have further investigation about the location selection in next part of the project

Explore Relationships between CVS, Larceny, and Eviction in the Central Boston

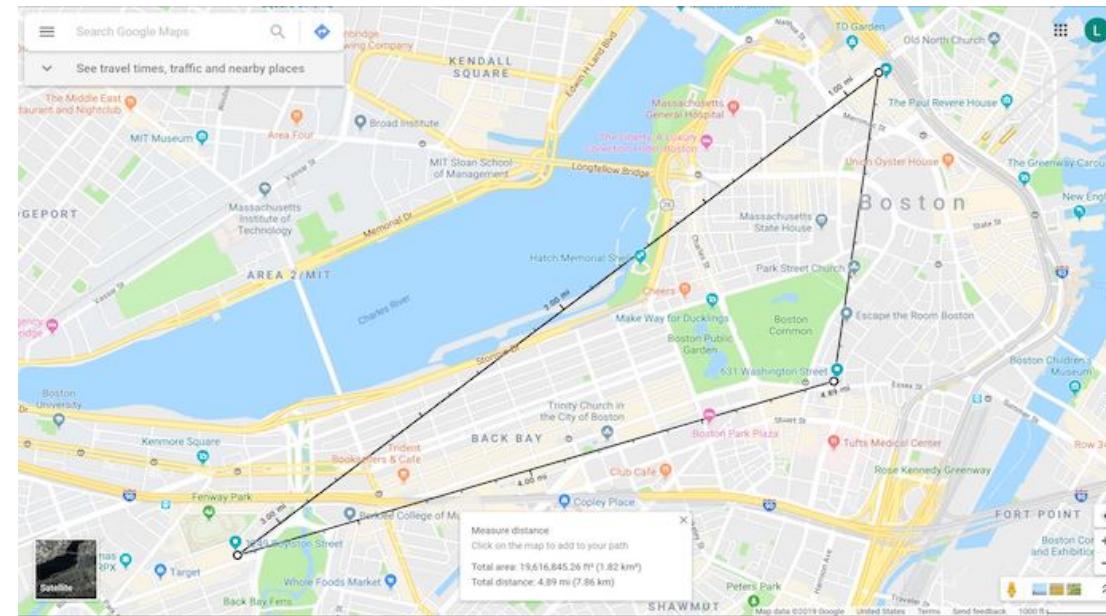
Solve the Optimization Problem

- Finding the model that has the maximum (total stability S, total accessibility A) pair
- Accessibility A is defined as the sum of the distances between each pair of salesmen, or each pair of stores that are assigned a salesman.
- $(S_1, A_1) > (S_2, A_2)$ if $S_1 > S_2$ and $A_1 > A_2$
- Results:
 - the 3 CVS are located at 210 Border St,
 - East Boston 101 Canal St suite A, Boston 1249 Boylston St, Boston
 - with a total stability $S = 0.4886$
 - with a total accessibility $A = 11.2147$ km



Solve the Constraint Satisfaction Problem

- If we have the chance to send 3 salesmen (1 per store), obtained the model that has the total stability $\geq S$, total accessibility $\geq A$?
- Results:
 - For $S := 0.6$; $A := 6.0$:
 - a possible set of 3 CVS are:
 - 101 Canal St suite A, 1249 Boylston St, 631 Washington St, Boston
 - with a total $S = 0.63646$
 - with a total $A = 7.91282$



CONCLUSION OF THE PROBLEM

MAP VISUALIZATIONS

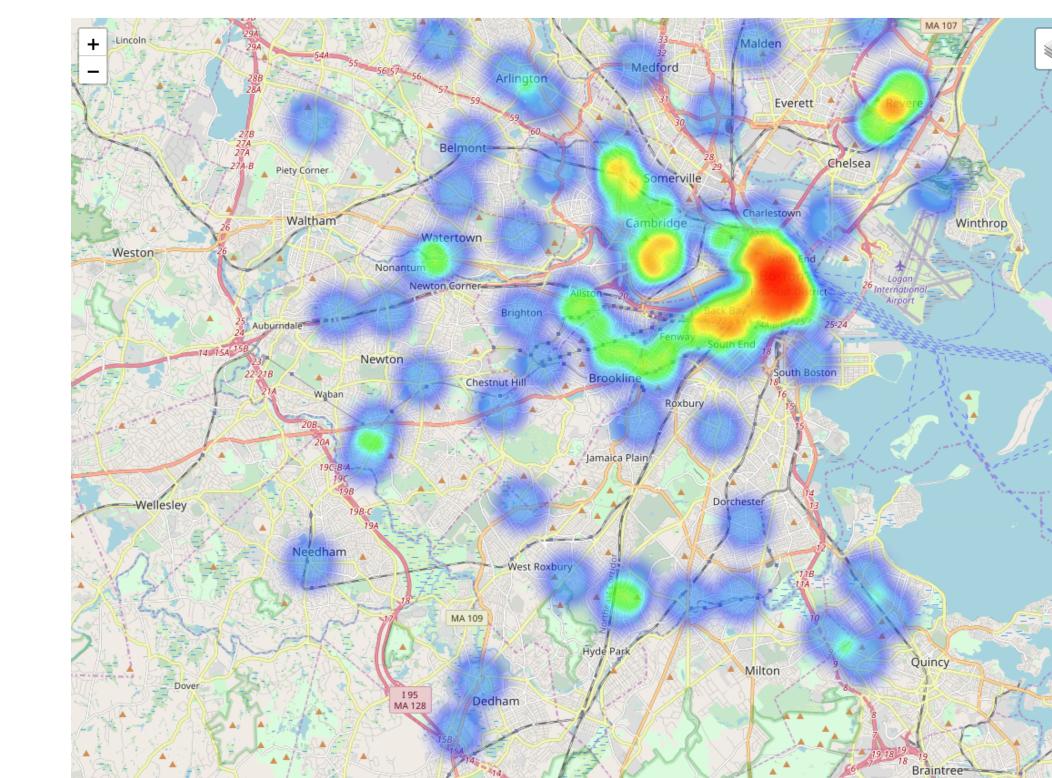


Fig 1: heat map of CVS and Walgreen stores

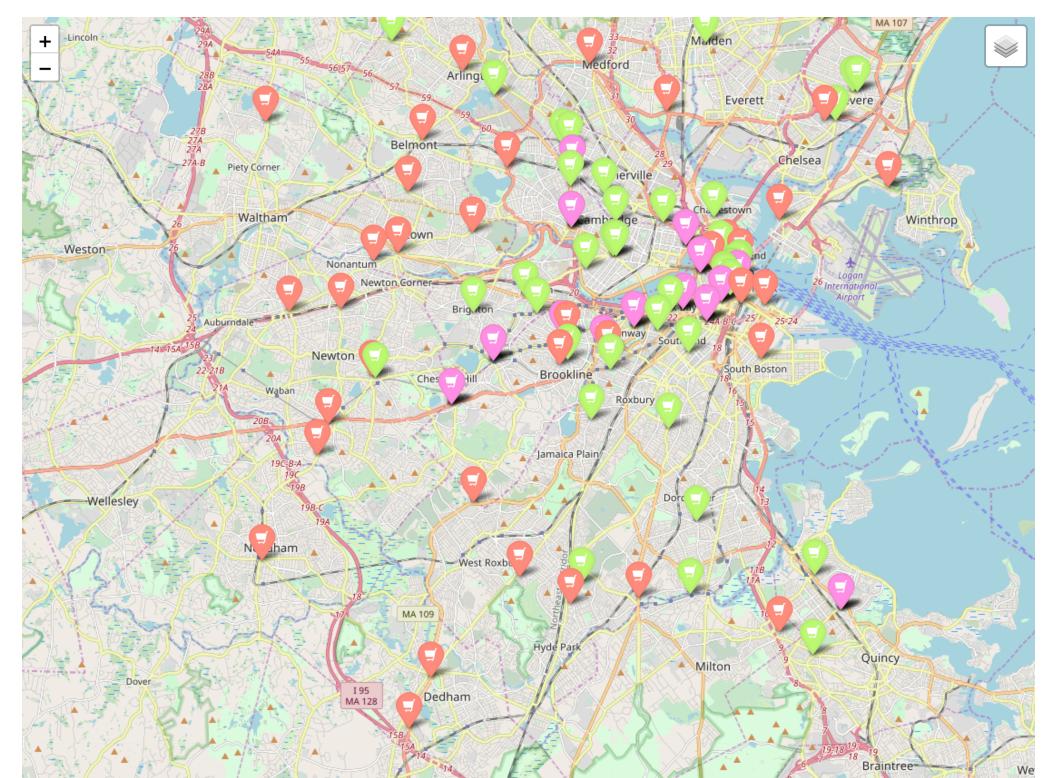


Fig 2: Distribution of CVS and Walgreen stores

Fig 1 shows that CVS and Walgreen are gathered in the center of Boston Metro area, and spread out outside Boston. Fig 2 indicates the specific distribution of CVS and Walgreen stores. The red icon represents CVS while the green icon represents Walgreen and the pink icon represents the location where CVS and Walgreen are close to each other. From those two figures, we can tell that beside the center area of Boston where both CVS and Walgreen are compacted, CVS spreads out its stores more and almost each town contains at least one CVS while Walgreen focuses more on targeting the center of Boston.

LIMITATIONS AND FUTURE RESEARCH

We initiate our investigation on the brand competition between CVS and Walgreen. Due to the lack of dataset we can access, our analysis merely rely on the location distribution. If we can get more information such as poverty level, income rate, age composition, race and so on, perhaps we can have more understanding on how CVS/Walgreen and other competitive brands in general select their locations based on their brand strategy and what is their target population.