

# Mapping visiting pattern changes before and after COVID-19 using the Huff model

Yunlei Liang & Kyle McNair

GEOG 573 Final Project

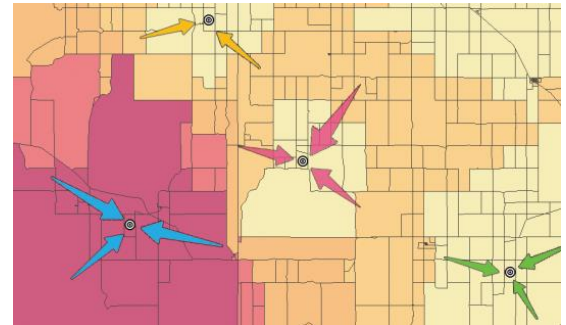


---



# Traditional Huff Model

- Idea: Estimate the shopping center trade area
  - a region containing potential customers for a store
- Two factors
  - Attractiveness: the merchandise offerings
  - Travel time/distance



# Traditional Huff Model

The probability of a customer  $i$  visiting a store  $j$  can be denoted as:

$$P_{ij} = \frac{\frac{S_j^\alpha}{D_{ij}^\beta}}{\sum_{j=1}^n \frac{S_j^\alpha}{D_{ij}^\beta}}$$

$S_j$  is the attractiveness of the store  $j$ ;  $D_{ij}$  is the physical distance between the customer  $i$  and the store  $j$ . The  $\alpha$  and  $\beta$  are parameters adjusting the model.

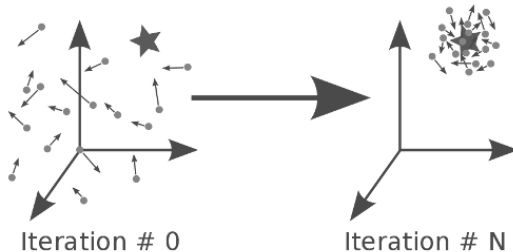
# Optimization

- Goal: find the parameters that maximize the Correlation ( $P_{\text{estimate}}$ ,  $P_{\text{actual}}$ )

$$P_{ijt} = \frac{\frac{S_j^\alpha}{D_{ij}^\beta}}{\sum_{j=1}^n \frac{S_j^\alpha}{D_{ij}^\beta}} * P_{jt}$$

- Particle Swarm Optimization

- inspired by the foraging behavior of bird flocking.
- each point needs to update its position and move towards the best location

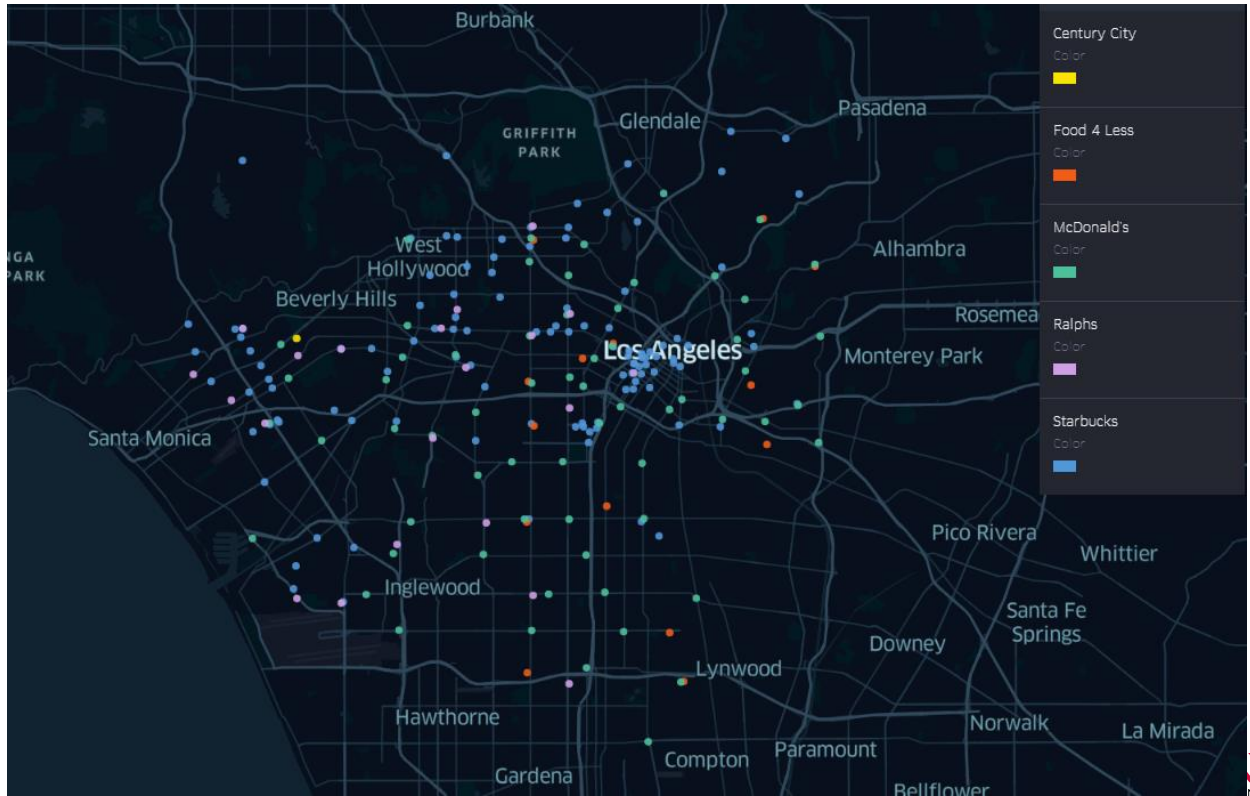


# Data

- POIs visiting data
  - Visiting information (# of visits, # of visitors)
  - The origin Census Block Group of visitors -> pairwise visits
  - The bucket dwell time: range of minutes and value is number of visits that were within that duration

location_name	city	raw_visit_counts	visitor_home_cbgs	bucketed_dwell_times
Garden Suite Hotel and Resort	Los Angeles	246	{'060372126101': 10, '060372126102': 5, '06037...	{'<5': 7, '5-20': 91, '21-60': 67, '61-240': 5...
Glabman Furniture	Los Angeles	12	}	{'<5': 0, '5-20': 4, '21-60': 0, '61-240': 2, ...
Four Points By Sheraton	Los Angeles	492	{'060372772001': 5}	{'<5': 4, '5-20': 120, '21-60': 85, '61-240': ...

# Study Area & Brands



# Process the bucket time

- Using the python packages pandas and matplotlib, we created a class that would help speed the process of analyzing different stores at different cities.
- With Safegraph data, we can collect and organize the dwell times in Los Angeles.
- Time Ranges:
  - < 5 Minutes
  - 5 – 20 Minutes
  - 21 – 60 Minutes
  - 61 – 240 Minutes
  - > 240 Minutes

# Process the bucket time

```

class Bucketed_Dwell_Time:
    def __init__(self, lst):
        self.lst = lst

    def average(self, lst):
        return sum(lst) / len(lst)

    def listSum(self, lst):
        total = 0
        item = 0
        while(item < len(lst)):
            total = total + lst[item]
            item += 1
        return total

    def findMedian(self, lst):
        n = len(lst)
        lst.sort()
        if n % 2 == 0:
            median1 = lst[n//2]
            median2 = lst[n//2 - 1]
            median = (median1 + median2)/2
        else:
            median = lst[n//2]
        return median

    def organizeDwells(self):
        dwell1 = []
        dwell2 = []
        dwell13 = []
        dwell14 = []
        dwell15 = []
        for b in self.lst:
            for k, v in b.items():
                if k == '<5':
                    dwell1.append(v)
                elif k == '5-20':
                    dwell2.append(v)
                elif k == '21-60':
                    dwell13.append(v)
                elif k == '61-240':
                    dwell14.append(v)
                else:
                    dwell15.append(v)
        self.dwell1 = dwell1
        self.dwell2 = dwell2
        self.dwell13 = dwell13
        self.dwell14 = dwell14
        self.dwell15 = dwell15

    def bDwell_avg(self):
        self.organizeDwells()
        bcktDwell_avg = {'<5': self.average(self.dwell1), '5-20': self.average(self.dwell2),
                        '21-60': self.average(self.dwell13), '61-240': self.average(self.dwell14),
                        '>240': self.average(self.dwell15)}
        return bcktDwell_avg

    def bDwell_median(self):
        self.organizeDwells()
        bcktDwell_mdn = {'<5': self.findMedian(self.dwell1), '5-20': self.findMedian(self.dwell2),
                        '21-60': self.findMedian(self.dwell13), '61-240': self.findMedian(self.dwell14),
                        '>240': self.findMedian(self.dwell15)}
        return bcktDwell_mdn

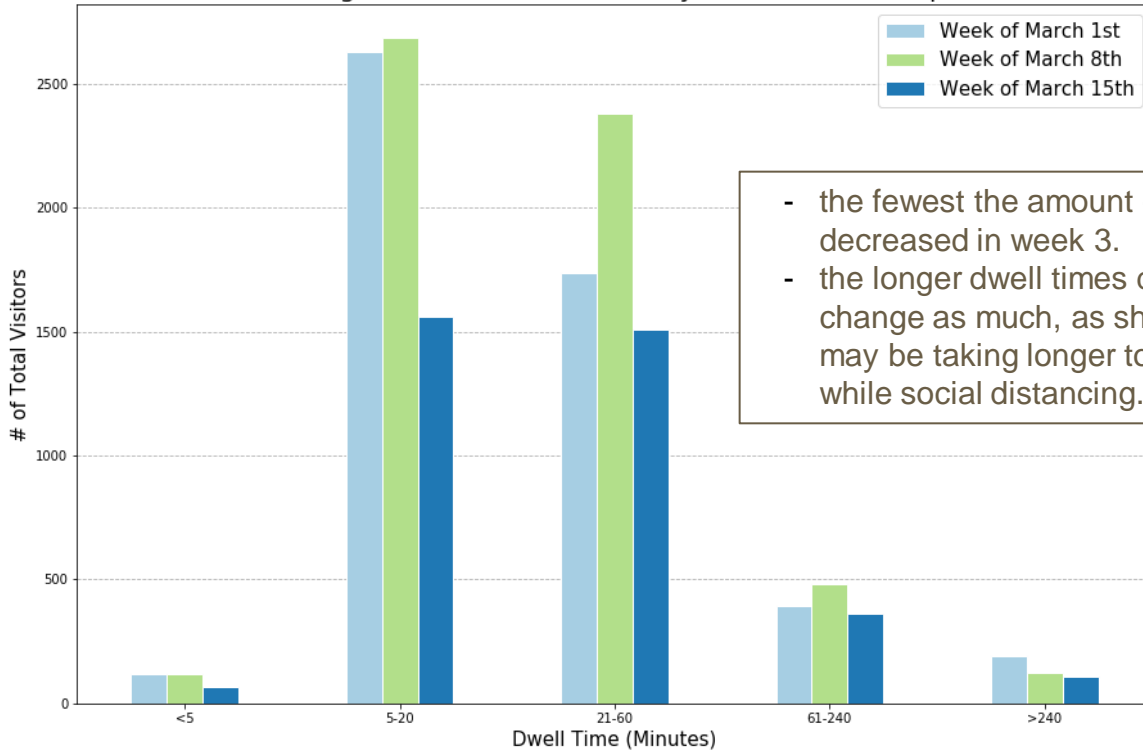
    def bDwell_count(self):
        self.organizeDwells()
        bcktDwell_count = {'<5': self.listSum(self.dwell1), '5-20': self.listSum(self.dwell2),
                           '21-60': self.listSum(self.dwell13),
                           '61-240': self.listSum(self.dwell14), '>240': self.listSum(self.dwell15)}
        return bcktDwell_count

```



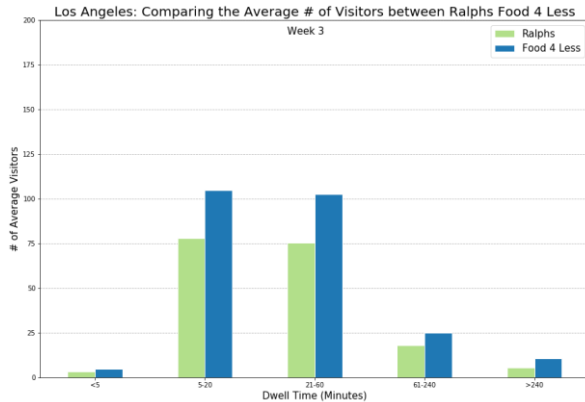
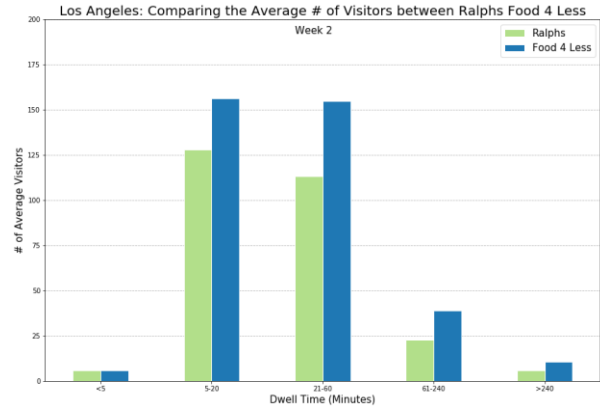
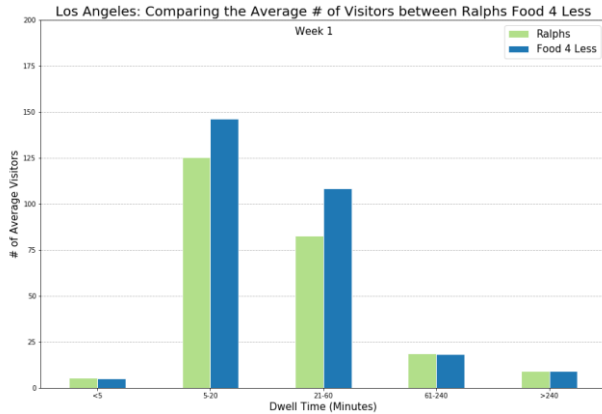
# Results – Ralphs in Los Angeles

Los Angeles: Total # of Visitors by Dwell Time at Ralphs



- the fewest the amount of visitors decreased in week 3.
- the longer dwell times do not change as much, as shoppers may be taking longer to shop while social distancing.

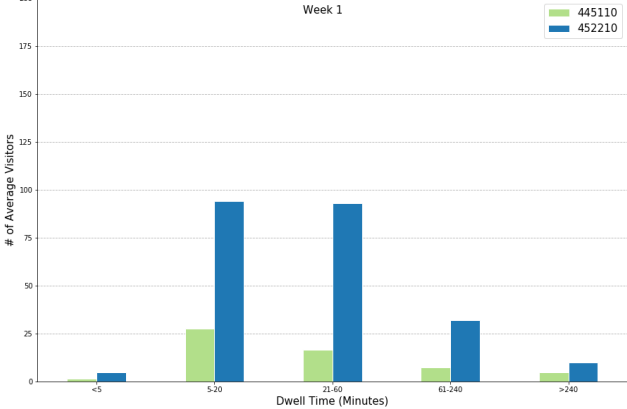
# Results – Comparing between Ralphs and Food 4 Less



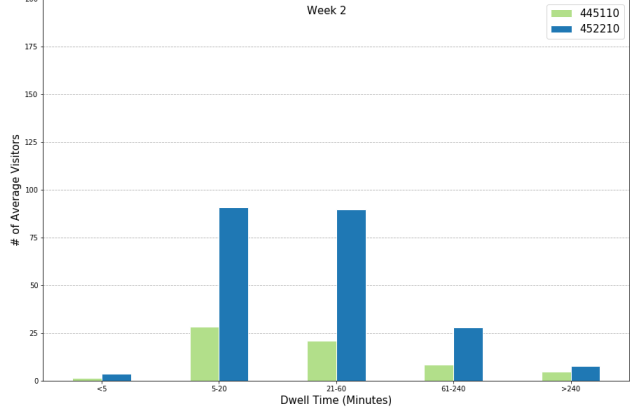
Comparing Ralphs and Food 4 Less shows the same trend of people who visit and stay at the store in the categorized dwell time. The steady decrease from week 2 to week 3 isn't as significant as expected, as people still need to buy food.

# Results – Comparing NAICS Codes

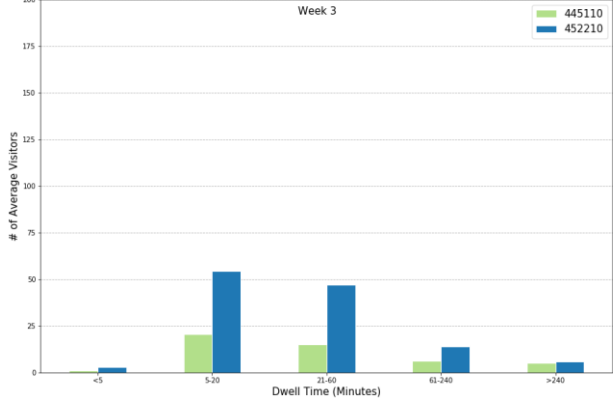
Los Angeles: Comparing the Average # of Visitors between NAICS Code 445110 452210



Los Angeles: Comparing the Average # of Visitors between NAICS Code 445110 452210



Los Angeles: Comparing the Average # of Visitors between NAICS Code 445110 452210



**445110 – Supermarkets & Grocery Stores**

**452210 – Department Stores**

# Optimization Results

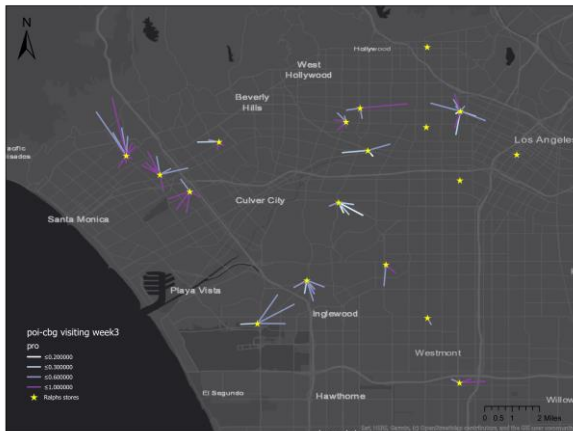
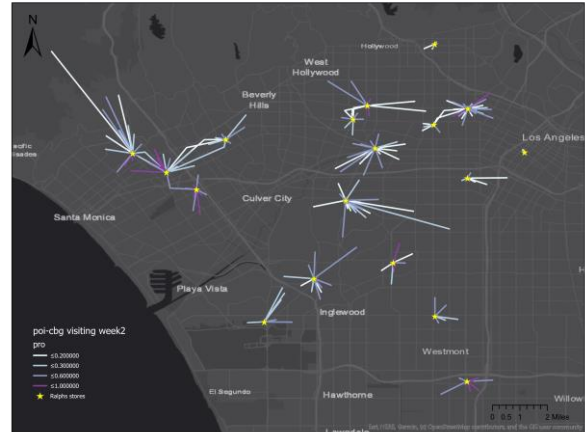
Table 1. The optimal correlation for each brand in each week

Brands	Week 1	Week 2	Week 3
Ralphs	0.976	0.978	0.989
Food 4 less	0.987	0.988	0.987
McDonald's	0.985	0.981	0.972
Century City	0.966	0.967	0.998
Starbucks	0.977	0.958	0.982

Table 2. The optimal beta for each brand in each week

Brands	Week 1	Week 2	Week 3
Ralphs	0.158	0.200	0.162
Food 4 less	0.004	0.221	0.284
McDonald's	0.172	0.153	0.167
Century City	0.387	0.340	0.542
Starbucks	0.274	0.303	0.150

# The visiting probability maps for Ralphs



- more visits in week 2 compared with week 1
- a dramatic decrease in the visits in week 3 compared with week 1 and week 2.
- the visits that disappeared in week 3 are those having a long distance or a low probability in week 2's map

## Brands comparison

Table 3. Sum of visiting probabilities\* pop in each cbg (market share)

Brands	Week 1	Week 2	Week 3
Ralphs	4800	7473	5496
Food 4 less	5977	9643	6695
McDonald's	5129	2976	1643
Century City	4816	3853	372
Starbucks	2806	2526	2598

- For Ralphs and Food 4 less
  - the market share increased from week 1 to week 2 and dropped from week 2 to week 3
  - many people went grocery in week 2 (March 8th to March 14th) and this phenomenon continued during week 3.
- For McDonald's and Century City, they both closed some normal functionalities in response to the COVID-19, so their market share decreased over the three weeks.
- For Starbucks, its market share keeps stable over time, it is possible that the transformation from dining to taking out did not affect it too much.

# Future work

- Extend the study area to more cities to detect regional variability and compare the same chain store performance in different cities.
- Compare the visiting probability between supermarkets and small local grocery stores to discover how COVID-19 affects their visiting probability.
- Analyze more types of brands such as restaurants.

**Thanks!**