2020

Impact of COVID-19 on Customer Behaviors in the Retail Industry in Philadelphia

Authors:

Alex Fu, Edward He, Hao Deng,

Jiameizi Yao, Neel Kumtakar

12/2/2020

Table of Contents

1. Executive Summary

In this project, we focus on the impact of COVID-19 on people's shopping behavior in the retail industry in Philadelphia. We explored the business question from four angles based on the mobility data from 2018 to 2020 provided by SafeGraph, which are:

1. Average Visits per Visitor

   To analyze the average visits per visitor of each brand, represent each visitor's repeated visits to brands, and express the brand dependence on a specific brand to customers.

2. Visit Share

   To analyze each brand's visit share in the market, express the horizontal relationship among brands, and track the flows of visit share between brands. One brand's loss is the other brand's gain by setting a metric as visit share.

3. Brand Attraction Radius

   To analyze the distance from home to each brand, express the localization of customers' preferred brands during the pandemic.

4. Popular Associations

   To analyze the brands that customers visit together on the same day, express the associations and close relationship between two or more brands.

Our business question is the impact of COVID-19 on customer behaviors in the retail industry in Philadelphia. We are interested in this business question because we believe that due to government restrictions and safety concerns, customer behaviors of buying consumer goods change differently. We use the above four metrics to understand how shopping behaviors differ pre- and post-pandemic from different angles. According to the findings we get throughout the analyses, we also make recommendations under the severe COVID-19 situation to both businesses and consumers.

We made the following hypotheses about the changes in customer behaviors during the pandemic before the analyses:

1. The brands that customers are still willing to stick to during the pandemic focus either on necessities such as grocery or health-related products such as pharmacy.
2. Distance from home of stores in the suburb area should decrease during the pandemic.
3. People reduce the number of stores they go to on the same day to avoid getting more exposure to the COVID-19.

2. Introduction to the Business Challenge

SafeGraph is a data company founded in 2016, whose mission is to provide the most accurate Point-of-Interest and store location geofences for the U.S. SafeGraph's geofences offer the most precise understanding in the market on where stores are located. Building footprints are derived from satellite imagery and places are spatially hierarchical. With SafeGraph's Mobility data, we succeed in connecting customers with the stores they visit and researching on the impact of the COVID-19 on customers behavior in the retail industry in Philadelphia. There are 1045 retail stores in Philadelphia, which is the 20th highest among all cities in U.S. Philadelphia government also implemented Red Phase & Stay-at-Home Order during pandemic starting Mar 23 till Jun 5[i]:

1. Life Sustaining Businesses Only
2. Large Gatherings Prohibited
3. Only Travel for Life-Sustaining Purposes Encouraged

Mobility data show the activities at every point of interest daily, which not only reflect the popularity of each place, but also helps us understand each customer's shopping behavior. We explore customers' behavior with four angles, which are Average Visits per Visitor, Visit Share, Brand Attraction Radius and Popular Associations. The corresponding hypotheses are:

- Stores that focus on necessities or health-related products shall have a high customer retention rate during the pandemic.
- The median distance traveled by visitors from home to the store should decrease in the suburb area during the pandemic.
- The number of stores customers visit within one day should decrease since people tend to avoid getting more exposure to the COVID-19.

3. Literature Review

Our goal is to study how professional researchers conduct theories and methods objectively. In addition, our team can compare our research outcomes with two studies as below using different angles of perspective and dataset.

IBM Study: COVID-19 Is Significantly Altering U.S. Consumer Behavior and Plans Post-Crisis[1]

Business Challenge:

The research team from IBM wanted to explore how COVID-19 has altered people's shopping behavior and preferences.

Conduct method:

A team of IBM researchers conducted a survey and targeted over 25,000 U.S. adult citizens on the perspectives of issues, including retail spending, transportation, future attendance at events in large venues, and returning to work.

Findings:

1. Aversion to Public transportation.
2. Preference on a personal vehicle, yet purchases are delayed due to financial status.
3. Event attendance will be down for some time.
4. Retail payment options: contactless payment options are favored.
5. Retail experience: Shopping local & visiting for essential goods only.
6. Work from home has become the mainstream format.

McKinsey Study: The Great Consumer Shift: Ten Charts That Show How U.S. Shopping Behavior Is Changing[2]

Business challenge:

How profoundly has COVID-19 crisis changed our behavior as consumers?

Conduct method:

McKinsey & Company designed a COVID-19 U.S. Consumer Survey from 06/15/2020 to 06/21/2020. The survey was targeted at 2,006 citizens who are sampled and weighted to match the United States general population.

Findings:

1. Consumers tend to purchase 20+ categories of goods online post-COVID-19 than before.
2. Millennials and high-income earners are in the lead of shopping online.
3. Consumer stickiness to the brand has decreased.
4. Brands need to convey their value to consumers.
5. Hygiene transparency is a key factor in the retail experience.
6. Entertainment spending has decreased; Consumers seek value for their money mainly in essential categories.

4. Data Review

In this project, we mainly utilize two datasets from SafeGraph Inc., which are Monthly Datasets and Core Points of Interest Data. Point of interest is a specific point location that is useful and interesting to us. In SafeGraph data, it contains business information about retail stores, gas stations, restaurants, and other major business locations that people visit daily, such as brand name, business category, and open hours.

Dataset

Monthly Datasets

These datasets describe the mobility of the population in the United States every month at the store level. We decide to use data from January 2018 to September 2020 to compare consumers' mobility patterns in 2020 at the same time in 2019 and 2018. We select a part of the variables from all the 26 variables provided by the monthly datasets, based on the four metrics we use to analyze the business question. Geographic information such as city and region are used to help us narrow down the city we focus on to Philadelphia. Visits per visitor metric utilize data of visit counts and visitor counts; visit share metric utilizes data of visit counts; attraction radius metric utilizes data of distance from home and brands association metric use data of related same day brand. A detailed table of the selected variables is shown below:

Variable Names	Description
safegraph_place_id	Distinct places I.D. from Safegraph
safegraph_brand_ids	Distinct brands I.D. from Safegraph
brands	Brands Name
city	Geographic Info
region	Geographic Info
raw_visit_counts	Visit Variables
raw_visitor_counts	Visit Variables
Average_visits_per_visitor	New Variables From *raw_visit_counts/raw_visitor_counts*
distance_from_home	Distance from visitors' home to a store.
related_same_day_brand	Following brands the visitors visit on the same day
Related_brands	New Variables From *related_same_day_brand*
Related_brands_frequency	New Variables From *related_same_day_brand*

Core POI Data

The POI data includes information about the points of interest, such as location name, address, and brand association for top national brands. We select part of the variables from all the 19 variables provided by the core POI data to help us narrow down the industry we focus to

retail industry and visualize the locations on the map based on the geographic data such as address and postal code. A detail table of the selected variables is shown below:

Variable Name	Description
safegraph_place_id	Distinct places I.D. from Safegraph
safegraph_brand_ids	Distinct brands I.D. from Safegraph
brands	Brands Name
sub_category	Category
naics_code	Formal classification system
latitude	Geographic Informaiton
longitude	Geographic Information
street_address	Geographic Information
postal_code	Geographic Information

Data Manipulation

Narrow Down the Datasets

The first task is to narrow down the datasets by geographical level and industry level for both Monthly datasets and Core POI data. After filtering geographically by setting the city as "Philadelphia," we use the NAICS code to define the retail industry, which "is a classification within the North American Industry Classification System." [3]. Then, we get information about 1085 retail stores in Philadelphia out of over 360,000 retail stores national-wide. The first five brands are Rite Aid, Sunoco, 7-Eleven, CVS, and Wawa, which have the most stores in the city – 70, 53, 46, 45, and 37, respectively. Also, there are 49 categories that align with the NAICS code from 441110 to 453998.

Aggregation: Store's Level to Brand's Level

In the dataset, the records are for each store with a more general brand name. Since our analysis focuses on the brands, we took the median or sum of all stores' values of target variables within one brand based on the different analyses. For example, the Visit Share part is using the total visits of a brand in each month. Assume there are n brands $B_i,\ where\ i = 1, 2, \dots, n$. Each brand has m different stores $S_j, where\ j = 1, 2, \dots, m,$ with different values of total monthly visits, v_{ij} $(m, n\ \in \mathbb{Z}^{0+})$. The total visits V_i for the brand B_i is:

$$V_i = \sum_{j=1}^{m} v_{ij}\ , where\ i = 1, 2, \dots, n.$$

Besides the Visit Share, in the Brands Association part, we use the total values for each brand. The other two analyses, the Visits per Visitor and the Brand Attraction Radius use the median

value instead of each brand. Continue with the example, the median visit V_i for the brand B_i is $\widetilde{V_i}$.

Target Brands Selection

In the following analyses, we mainly focus on eight "top brands." They are picked based on the total visits for each brand in different months. We find the top five brands in February to May in 2020 and combine the four lists, then take the top eight, which are:

brands	subcategory
Walmart	All Other General Merchandise Stores
Sam's Club	All Other General Merchandise Stores
BJ's Wholesale Club	All Other General Merchandise Stores
ShopRite	Supermarkets and Other Grocery (except Convenience) Stores
Century 21 Department Stores	Family Clothing Stores
Lowe's	Hardware Stores
The Fresh Grocer	Supermarkets and Other Grocery (except Convenience) Stores
The Home Depot	Hardware Stores

Additionally, the Brands Association part starts with a combined list of top 100 associations each month, then focuses on the eight top brands.

5. Data Analysis

Business Question: What is the impact of COVID-19 on customer behaviors in the retail industry in Philadelphia?

Average Visits per Visitor

We analyze how shopping behaviors of customers change during the pandemic from several aspects. The first is to look at the repeat times that a customer visits a brand monthly. This value represents the dependence level of a brand to consumers, which is called **Brand Dependence.** The hypothesis is the brand dependence decreases since COVID-19 makes people stay at home and avoid shopping.

To check if the hypothesis is right, we analyze the changes in average visits per visitor in this part by utilizing the two columns in the Monthly dataset, *Raw Visits Counts* and *Raw Visitors Counts*. Then, the *Average Visits per Visitor* for a brand can be calculated by:

$$Average\ Visits\ per\ Visitor = \frac{Median\ Raw\ Visits\ Counts\ among\ all\ stores}{Median\ Raw\ Visitors\ Counts\ among\ all\ stores}$$

for February to May from 2018 to 2020.

Overview:

The overall situation of the distribution of retailers' brand dependence in Philadelphia is showing below. Each bar stands for each month. Most average visits are lower than four times, while some brands almost reach ten times repeated visits in one month for each customer. At the same time, the top brands located in the upper part of the bar each month.

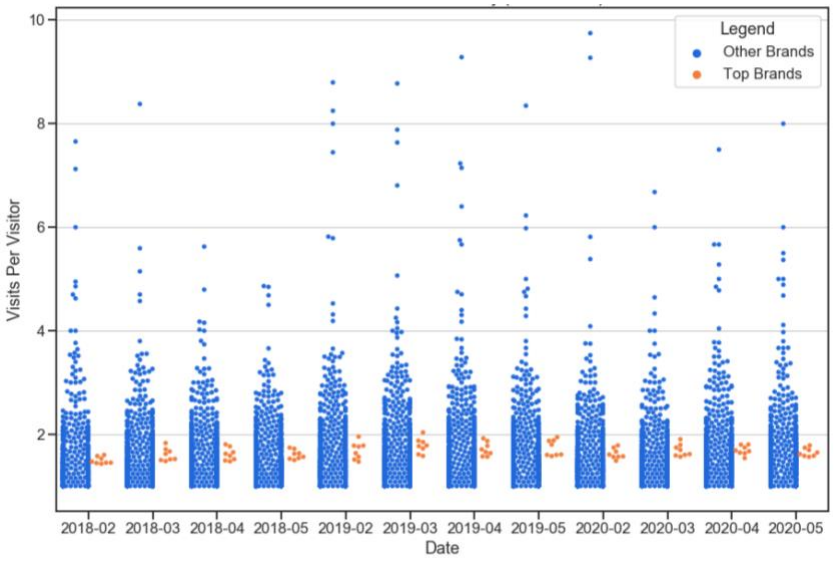


Figure 5.1.1 Visits Per Visitor in Feb – May (2018 – 2020)

The plot blow shows the variation of the visits per visitor among the 12 months. The line represents the median values within each month. As we can see in the plot, the top brands' fluctuation is more extensive than other brands'. To be more specific, the median value for other brands only increases from around 1.3 to around 1.4, and the overall trend is flat. In contrast, the top eight brands almost reached 1.8 from lower 1.3 in March 2019 and dropped rapidly in February 2020. This plot shows that the leading retailers' brand dependence with a larger scale of visits is more sensitive to the abnormal situation.

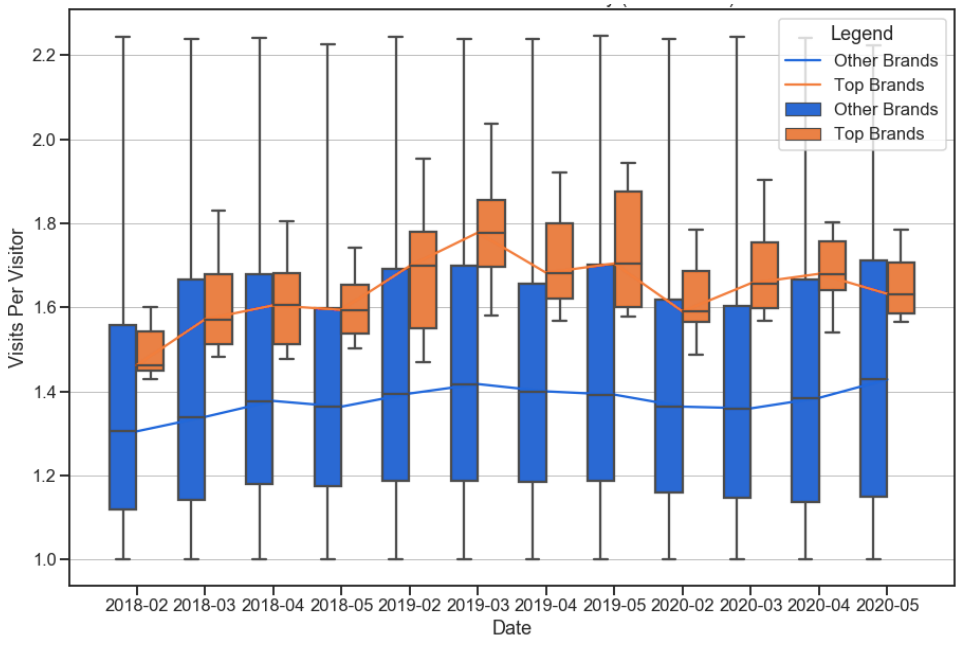


Figure 5.1.2 Trend of Visits Per Visitor in Feb – May (2018 – 2020)

Changes in the Adjacent Years:

The changes in Adjacent Year is calculated by:

$$\text{Changes in average visits per visitor} =$$

$$\frac{Average\ visits\ per\ visitor\ (in\ Year_{i+1} - in\ Year_i)}{Average\ visits\ per\ visitor\ in\ Year_i}, where\ i = 2018, 2019.$$

The two plots below are showing results for the top eight retailing brands.

- All Other General Merchandise Stores: **Walmart** had a significant increase in 2019 but had a more massive decrease due to the pandemic in February and March in 2020. In April and May, they recovered and back to the same changes in 2019. **Sam's Club** performed excellently during the most severe period – February and March in 2020 and had a more considerable increase than 2019. **B.J.'s Wholesale Club** got some impact from COVID-19, but not very severe.
- Family Clothing Stores: **Century 21 Department Store** got the most massive impact during the pandemic, which is most likely why they went bankrupt in September 2020.
- Hardware Stores: **Lowe's** recovered fast in May 2020 from the considerable decrease of average visits in February. **The Home Depot** had a tremendous increase in February, which is rare in the retail industry, especially when other months decreased or slightly increased. It is probably related to the stock up during the early stage of the pandemic.
- Supermarkets and Other Grocery (except Convenience) Stores: Both **ShopRite** and **The Fresh Grocer** decreased or slightly increased, but their trends are opposite in 2020. The reason maybe there are many ShopRite stores in Philadelphia, and people start to be familiar with the Fresh Grocer's delivery service, which is more convenient and safe than shopping in person.

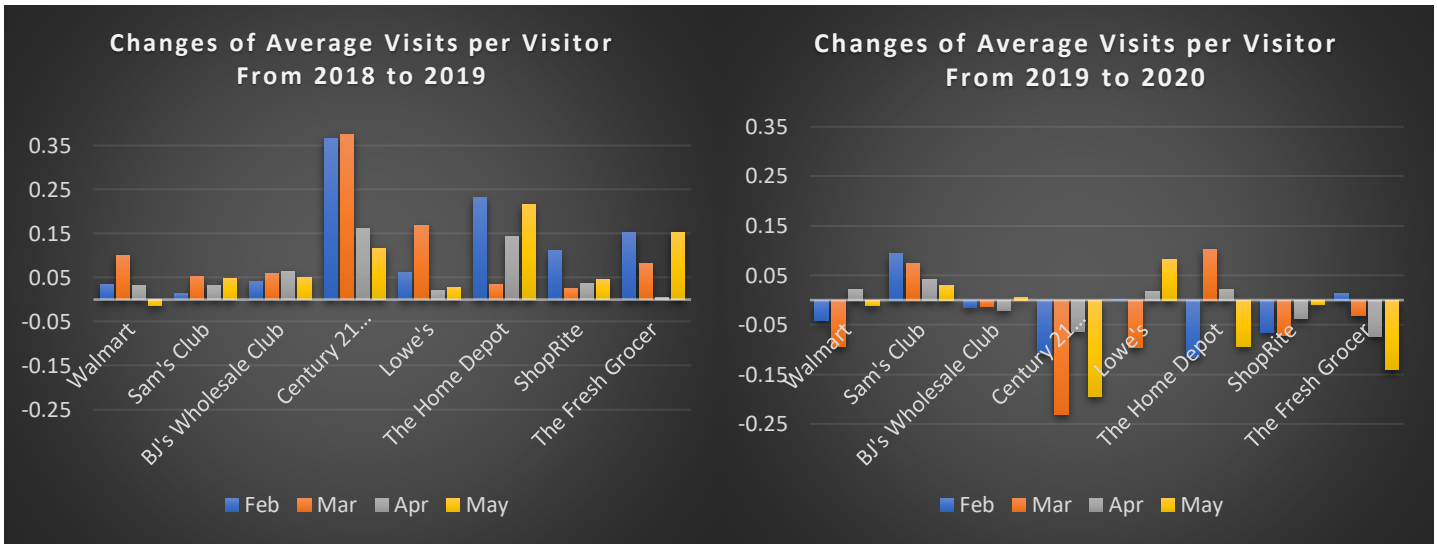


Figure 5.1.3 Changes of Average Visits per Visitor From 2018 (2019) to 2019 (2020)

Total Visit Counts and Visit Share

We are interested in if customers change their brand focus during the pandemic. To examine customers' brand preferences, we pick the top 8 brands by visit counts and calculate their visit share by total visits among over 600 brands (the pool) in the market. The equation that we use to calculate the visit share is:

$$Visit\ Share = \frac{Total\ Visits\ per\ Brand}{Total\ Visits\ in\ the\ Pool}$$

We choose to compare visit share from February to May 2018, 2019 and 2020 since visit share is a key indicator of market competitiveness, essentially one brand's gain in visit share is other brands' loss. It can help us to identify the trends of competitiveness among brands.

Figure 5.2.1 Total Visit Counts & Top Brand Visit Share Feb – May 2018 & 2019 & 2020

As shown from the above chart, the pool we use to calculate the visit share is over 600 brands available in Philadelphia. We pick the top 8 brands as stated in the data manipulation step, and the blue bars represent the sum of total visit counts of over 600 brands; the orange line represents the total visit share for the top 8 brands. There is a huge decrease of total visit counts around March to May in 2020 due to pandemic, however, the sum of total visit share of top 8 brands shows an opposite increase around the same time. The situation reveals that these leading brands are performing better than the rest brands in the pool in terms of visit share, and express the competitiveness during the pandemic. An brief conclusion is that customers pay relatively more visits to the leading brands compared to niche brands, and a reasonable explain is that leading brands normally have higher brand reputation, making customers feel safer with

the leading brands because they believe these leading brands have more experience in solving unexpected situation such as COVID-19.

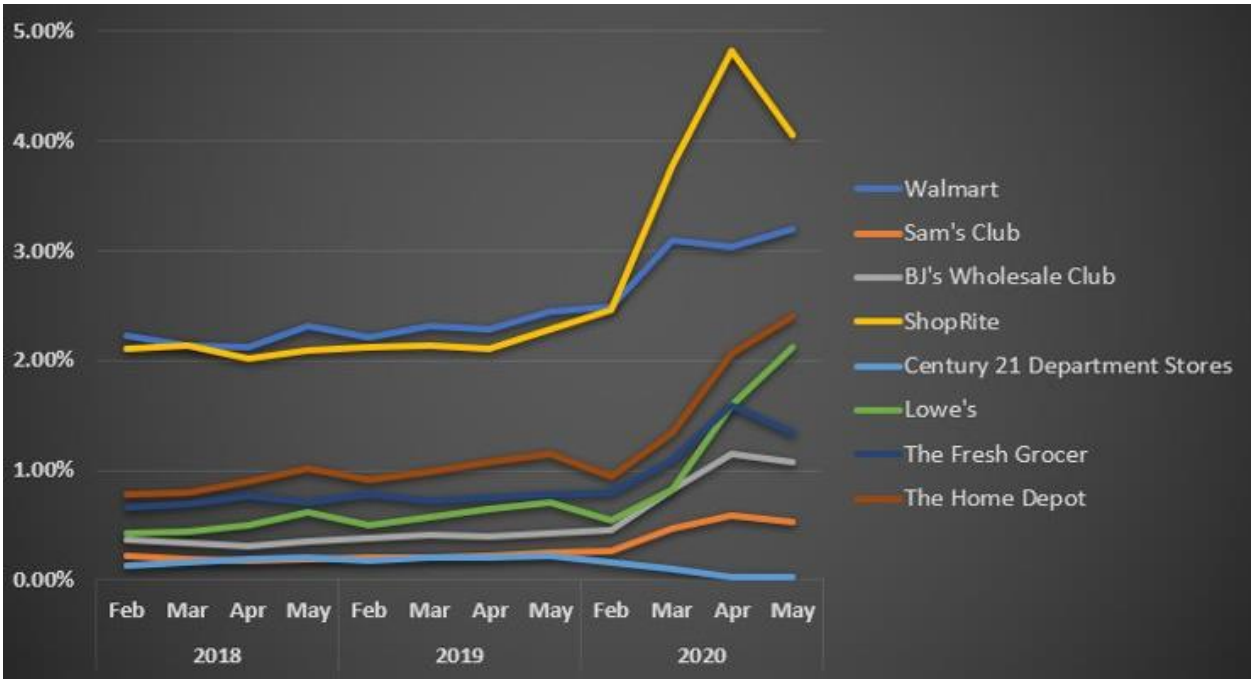


Figure 5.2.2 Visit Share Feb – May 2018 & 2019 & 2020

After knowing that leading brands turn to perform better than the rest brands in the pool in terms of visit share, we are interested that how these top 8 brands contribute to the increase in the sum of total visit share. By looking into the visit share of each top 8 brands, we are able to compare their visit share within these leading brands. The overall pattern of the visit share of each top 8 brands, as shown in the chart above, is following the result from the previous chart. Most of brands in the top list have the visits share stable during Feburary to May in 2018 and 2019, and incur a significant increase of visit share during Feburary to May which are the pandemic months in 2020. These leading brands are becoming competitive compare to the rest niche brands in terms of visits. Within all of these increases in visit share, the visit share of ShopRite, which is the yellow line, outperformes other brands a lot during the pandemic in terms of visit share. We can see that grocery stores especially ShopRite experience significant surge in demand, also because ShopRite has much more stores located in the city while other big box stores such as Wakmart and Lowe's locate in the edge of the city. The result tells us that stores such as ShopRite who have many store chains in the city remain competitive during the pandemic. A reasonable explain is that customers turn to choose the most convenient way to buy necessities to reduce the risk of getting COVID-19.

Brand Attraction Radius

For this part, we focus on the impact of the COVID-19 to the distance traveled by visitors from home to the stores-distance from home. There are three areas(South, Upper Northeast and Kensington Area)in Philadelphia where top 8 brands that we focused on are gathered, we research on the situation in 2020 and analyze how distance from home changed when compared to the situation in previous years.

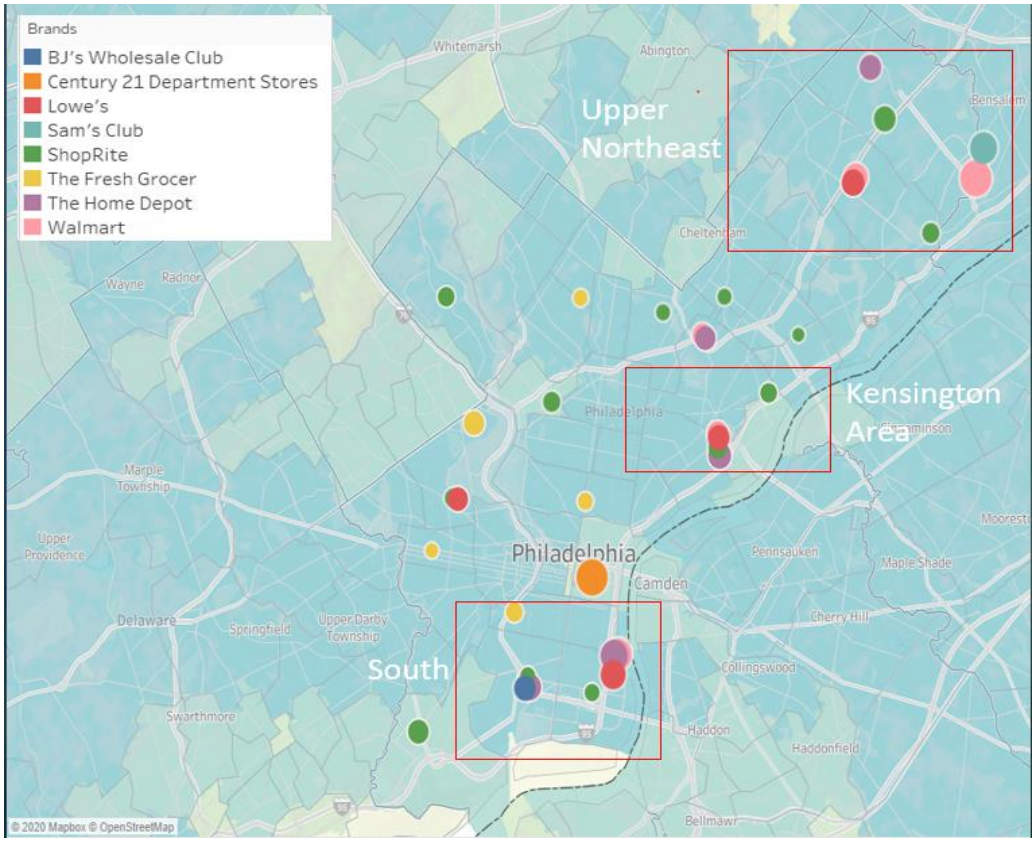


Figure5.3.1 distribution of top 8 brands stores in Philadelphia

Overview:

It can be observed from Figure below that for each brand, traveling distance between visitors' home and branches in South and Upper Northeast is further than that in Kensington. The reason for this phenomenon is that Kensington is closer to central Philadephia where people are more populated compares to the other two areas. In addition, it is obvious that traveling distance between visitors' home and branches of Supermarkets and grocery stores(ShopRite and The Fresh Grocer) is closest. This can be explained by the fact that Shoprite and the Fresh Grocer are the top 2 brands which own the most branches in Philadelphia and from figure above we can see

that their branches are scattered separately across Philadelphia. Customers can always find the closest branches of these two brands around them, which shorten the distance from visitors' home to the branches.

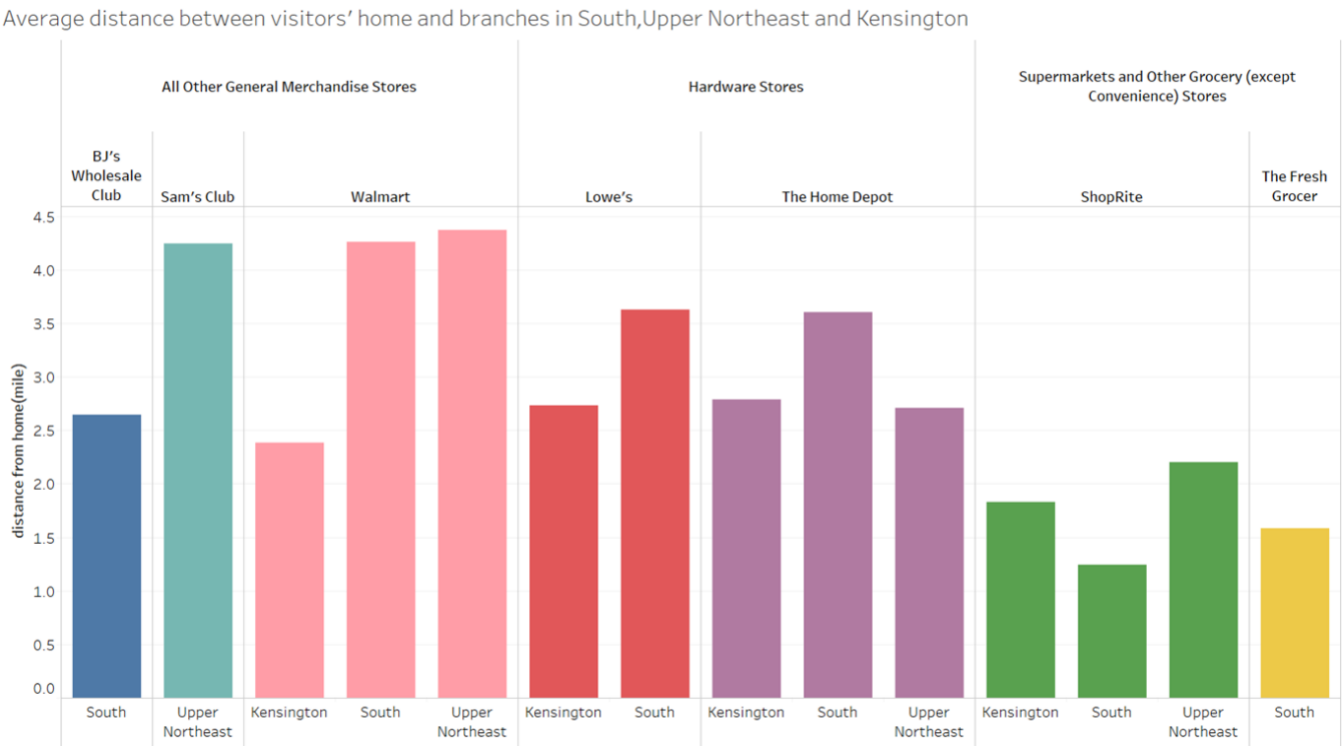


Figure 5.3.2 Average distance between visitors' home and branches in South,Upper Northeast and Kensington

Situation in 2020:

In order to analyze the situation in 2020, we need to separate the monthly data in 2020 from previous 2 years and see how it deviates. We introduce Flucatuation to perform comparation:

$$Fluctuation\ in\ month\ i = \frac{D_{2020i} - (D_{2019i} + D_{2018i})/2}{(D_{2019i} + D_{2018i})/2},$$

$$where\ D_{2020i}\ is\ distance\ from\ home\ in\ month\ i\ of\ 2020,\ etc.)$$

brands	Area	2020FEB fluctuation	2020March fluctuation	2020April fluctuation	2020May fluctuation
Lowe's	Kensington	3%	-5%	-5%	1%
ShopRite	Kensington	-7%	-1%	-1%	3%
ShopRite	Kensington	-2%	1%	-2%	4%
The Home Depot	Kensington	5%	4%	-7%	1%
Walmart	Kensington	3%	5%	-3%	8%
BJ's Wholesale Club	South	-5%	-5%	-5%	4%
Lowe's	South	10%	10%	-8%	-12%
ShopRite	South	-4%	-1%	-11%	-7%
ShopRite	South	6%	5%	0%	15%
The Fresh Grocer	South	-3%	-3%	3%	6%
The Home Depot	South	4%	-2%	-5%	-6%
The Home Depot	South	8%	11%	closed	10%
Walmart	South	5%	3%	0%	6%
Sam's Club	Upper Northeast	-5%	2%	0%	-4%
ShopRite	Upper Northeast	-4%	4%	closed	2%
ShopRite	Upper Northeast	-6%	-8%	closed	-1%
The Home Depot	Upper Northeast	10%	-1%	3%	4%
Walmart	Upper Northeast	-9%	-19%	-25%	-16%
Walmart	Upper Northeast	0%	3%	closed	1%

Figure 5.3.3 Fluctuation in February, March, April, and May 2020

From the table above, it can be observed that the overall fluctuation is minor before and after pandemic. However, in April, almost all the stores have negative fluctuation. Three branches in Upper Northeast and one branch in South area are even closed.These stores are closed because they are located in shopping center, due to the COVID-19 red phase regulations implemented by Philadelphia government, shopping centers are forced to close. In a nutshell, COVID-19 has impact on retail industry, especially those stores located in shopping centers. But its impact on people's traveling distance is minor.

Popular Associations

Background

As we mentioned in the hypothesis, the purpose of studying associations is to find potential partners for top retail brands such as B.J.'s Wholestore and Walmart. The more frequent the customers visit the two brands on the same day, the closer the two brands' relationship is. To clarify the term "Brand" actually refers to "Brand store" which is not an actual store but representative of all the actual stores in Philadelphia. For instance brand "Dunkin" represents all physical Dunkin stores located in Philadelphia.

Methodology

Social Network Analysis

The primary methodology we used here is the Social Network Analysis (SNA). Edges and nodes are the two main elements in the algorithm. More specifically, we map the columns "brands" and "Related Brands Same Day" into different associations. The brands represent nodes, while the associations represent edges.

The column "brands" will be transformed into "brand1," and the column "related brands same day" will be transformed into multiple "brand2" with different frequencies. The associations always flow from brand1 to brand2. In other words, the data recorded visitors visited brand1 and then visited brand2 on the same day.

Due to the associations' direction, we measured the brands in two ways: outwardness and inwardness. The outwardness is a standardized metric for a brand's outward behavior, i.e., the number of customers going out. At the same time, inwardness is a metric for a brand's inward behavior: the number of people entering the establishment. Consider an association as A_{ij} where customers flow from brand i to brand j. The formulas for the Inward Score (I_m) and Outward Score (O_i) are shown below:

$$O_i = \sum_{m=1}^{n} A_{im}$$

$$I_m = \sum_{i=1}^{n} A_{im}$$

While inward and outward scores are not used in the association analysis, it is crucial for understanding in order to understand associations.

In terms of association level, we have two features in the metric: frequency and distance. The frequency is the average number of people and will be the weight of the association between two brands. Distance is defined by the average smallest geometric Manhattan distance of all stores in Philadelphia for brand1 away from brand2. The process is:

a. Calculate the Manhattan distance for all combinations of stores of brand1 and brand2.
b. Find the nearest store of brand2 for each brand1.
c. Calculate the average distance among all brand1's stores, which will be the association's final distance.

The more structured version of the explanation of the SNA is shown below:

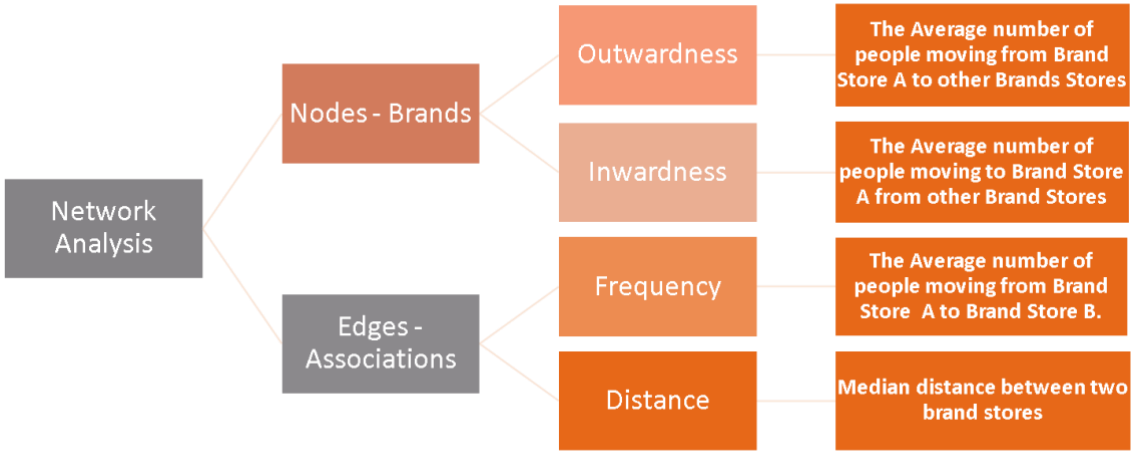


Figure 5.4.1 Social Network Analysis Structure

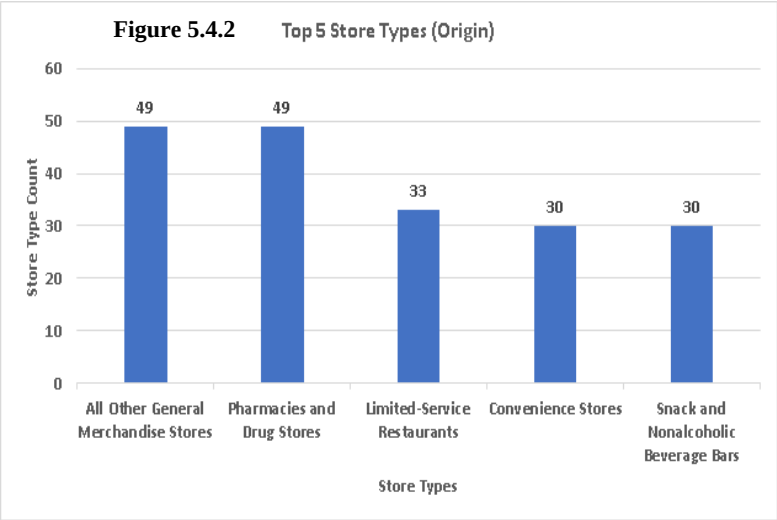
Reduce the impact of distance

Two factors led to the frequency we have right now -- the business side and convenience level. When the two brands' products complement each other, when the two brands have a similar culture that has attracted similar customer groups, etc., the same customer might go to the same brands on the same day. Of course, when the two brands are always neighbors, it will also increase the associations' frequency.

In certain circumstances, the distance impacts our analysis. If there is a connection between two brands, it is assumed that the brands are relatively close to each other. For instance, in a plaza, if there is a Dunkin Donuts and a Subway next to each other, it is possible that customers from Dunkin' could enter Subway after and vice versa. Recommendations by distance is dangerous since it may not always be the always the case that the two brands might be successful together even when they are partners. We want to focus on the **changing** factor known as frequency and reduce the impact of distance.

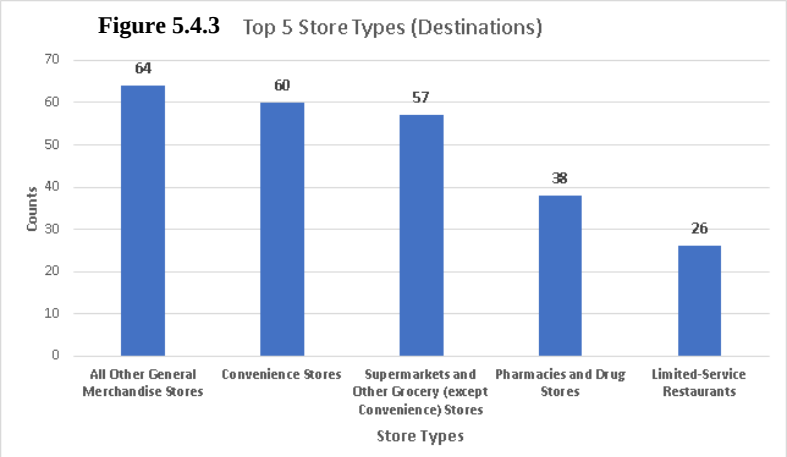
Analysis

As mentioned in the data manipulation, we find the top 100 associations each month and combine the list. In the data we found 335 unique associations. As mentioned before, as association consists of two brands. The first brand is known as the "origin brand" while the second brand is the "destination brand". When people move from one brand to another, this collective instance is considered an association.



We find that the first brand in most of the associations comes from the top 5 subcategories.

- General Merchandising Store
- Pharmacies and Drug Stores
- Limited-Service Restaurants
- Continence Stores
- Snack and Non-Alcoholic Beverage bars

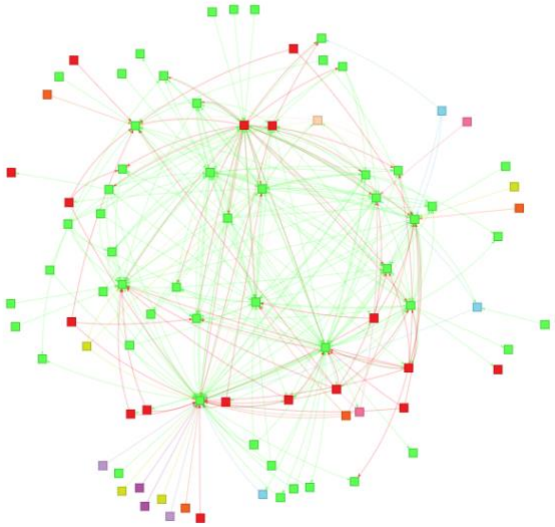


The Second brand however appears to derive from below categories:

- General Merch. Stores
- Pharmacies and Drug Stores
- Limited-Service Restaurants
- Convenience Stores
- Supermarkets

For the most part, brands 1 and 2 have the same categories, which is interesting. From these graphs, Philadelphia gets most of its business from General Merchandising,Pharmacies, Limited-Service Restaurants, and Continence stores.

All Philadelphia Associations



The network map on the left side lists all the nodes (Brands) and all the connections / Associations between them. The square nodes are colored according to the category, and the green, in particular, represents retail stores. From the network map, most of the associations are from Retail stores.

Figure 5.4.4: SNA of Phialdephia Associations

The above map is relatively large, and we need to condense it to do some proper analysis. We want to reduce the dataset in such a way that either Brand A or Brand B is from a retailer Brand. In particular we want to see the associations are connected to the following big retailer brands

- The Home Depot
- BJ's Wholesale
- Walmart
- The Fresh Grocer
- Shop Rite
- Lowe's

This implies that in an association, either Brand A or Brand B has to be one of the 6 brands. The resulting map is shown below.

Key Retail Associations

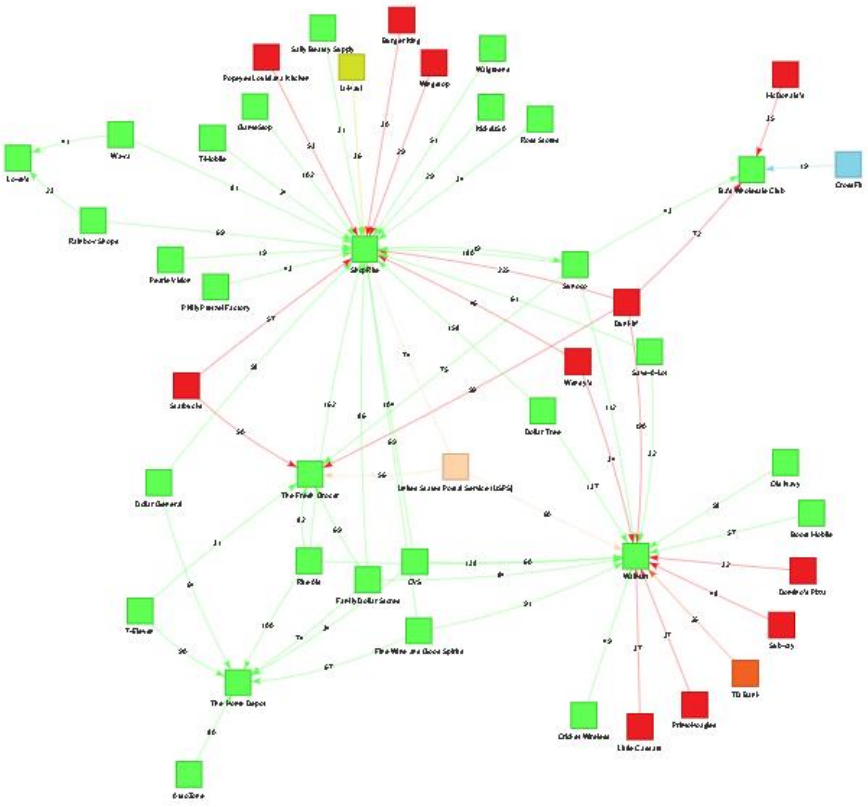


Figure 5.4.5: SNA of Key Retailer Philadelphia Assocaitions

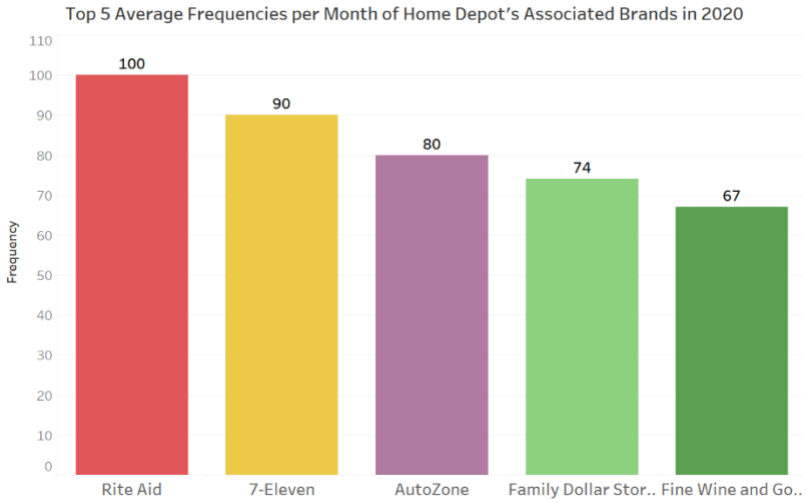
From this map, we see that retailer brands (represented by green squares) are associated with fast food and coffee shop brands (represented by red squares). A quick magnification shows that there are some retailers that serve as "magnets" to other brands including other retailers.

There are no arrows pointing out to other brands, and all arrows are pointing into the brand, hence the term. Interestingly enough, these Magnets are actually the big retailers mentioned earlier.

Business insight overview

There is a total of 7 big retailer brands and we are interested in what brands are associated with each and every one of them. In each key brand, the top 5 associated brands will be listed but it is important to note that not always a brand will have 5 associated key brands. It could be that a Key retailer brand has less than 5 associated brands. Thus the top n where $n \leq 5$ will be shown.

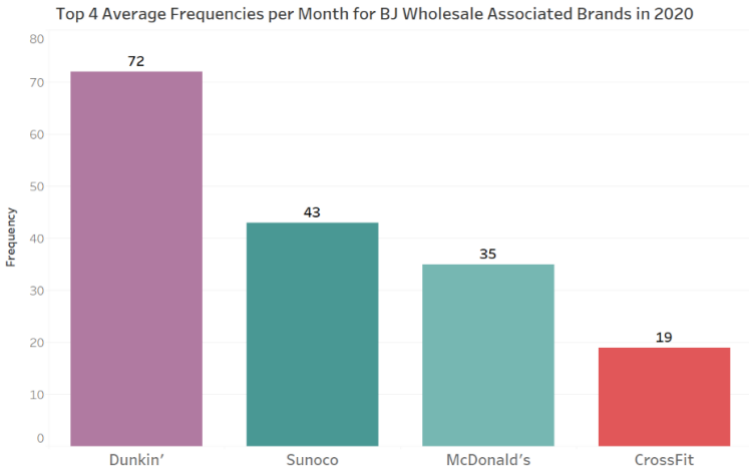
Home Depot: From the bar chart, we see that the top 5 brands that are most associated to Home Depot. We see that on average each month:



- 100 people come from Rite Aid to Home Depot
- 90 people come from 7-Eleven to Home Depot
- 80 People come from Auto Zone to Home Depot
- 74 people come from Home Depot from Fine Wine and Good Spirits to Home Depot
- 67 people come from Fine Wine and Good Spirits to HD

Figure 5.4.6: Average Monthly Frequency of Home Depot's Top 5 Associated Brands

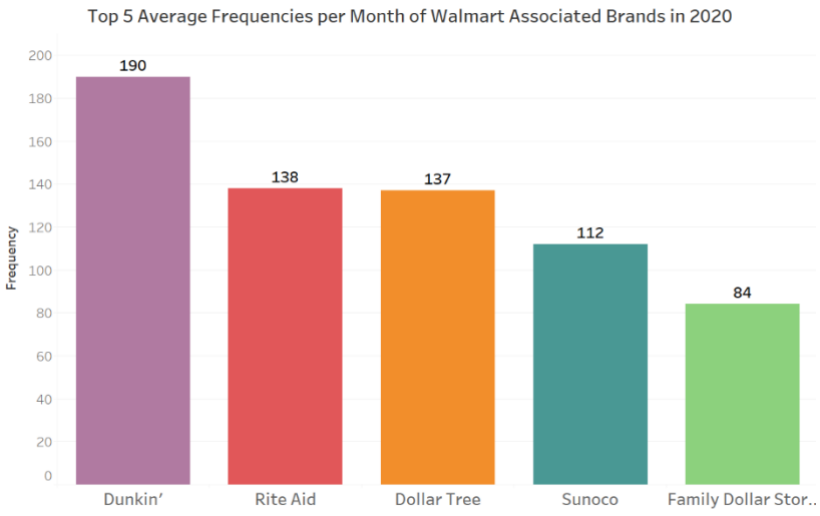
BJ's Wholesale: From the bar chart, we see that the top 4 brands that are most associated to B.J.'s Wholesale. We see that on average each month:



- 72 people come from Dunkin to BJ's
- 43 people come from Sunoco to BJ's
- 35 people come from McDonald's to B.J.'s
- 19 People come from Crossfit to BJ's

Figure 5.4.7: Average Monthly Frequency of BJ's Wholesale's Top 4 Associated Brands

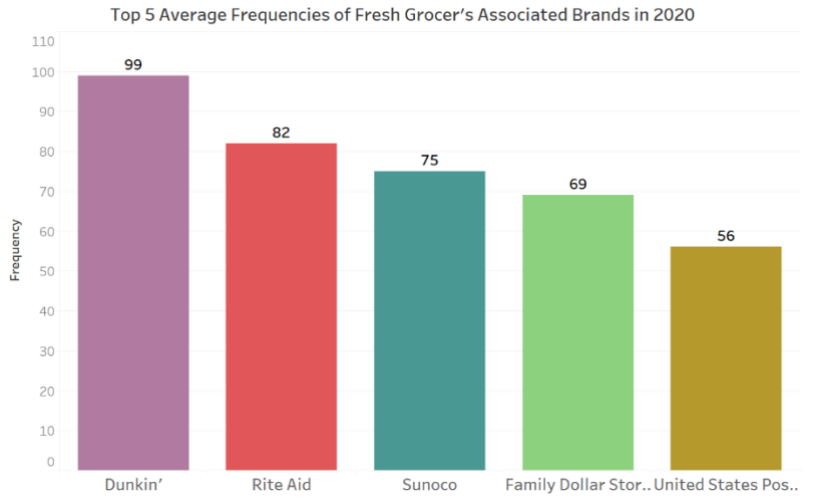
Walmart: From the bar chart, we see that the top 5 brands that are most associated to Walmart. We see that on average each month:



- 190 people come from Dunkin' to Walmart
- 139 people come from Rite Aid to Walmart
- 137 people come from Dollar Tree to Walmart
- 112 people come from Sunoco to Walmart
- 84 people come from Family Dollar Store to Walmart

Figure 5.4.8: Average Monthly Frequency of Walmart's Top 5 Associated Brands

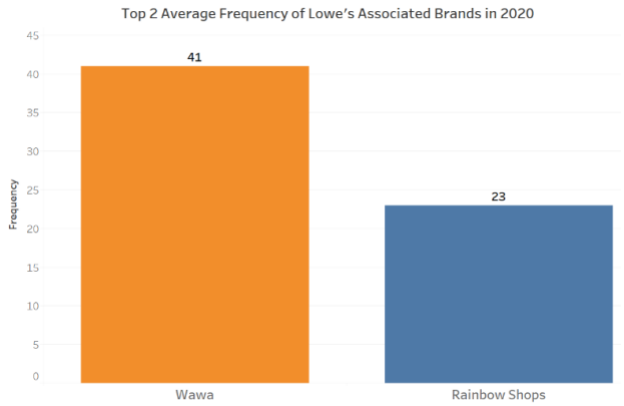
The Fresh Grocer: From the bar chart, we see that the top 5 brands are most associated to Walmart. We see that on average each month:



- 99 people come from Dunkin to Fresh Grocer
- 82 people from Rite Aid to Fresh Grocer
- 75 people come from Sunoco to Fresh Grocer
- 69 people come from Family Dollar Store to Fresh Grocer
- 56 people come from US Postal Service (USPS) to Fresh Grocer

Figure 5.4.9: Average Monthly Frequency of Fresh Grocer's Top 5 Associated Brands

Lowe's: From the bar chart, we see that only 2 brands that are most associated to Lowe's.

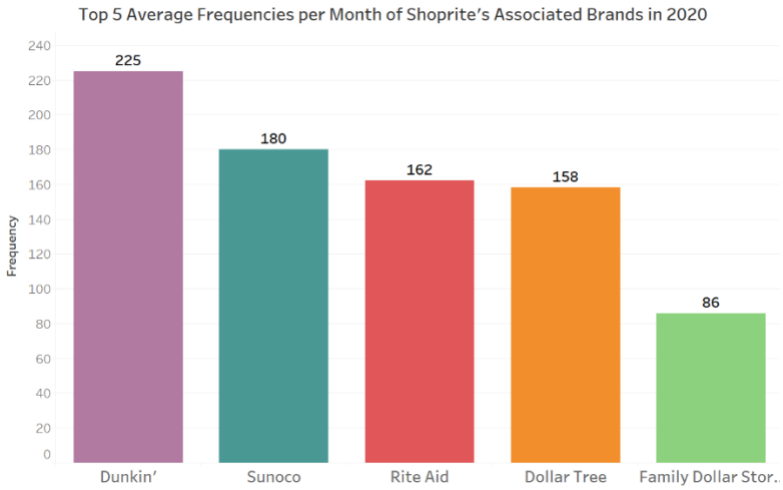


We see that:

- 41 people move from Wawa to Lowe's
- 23 people move from Rainbow Shops to Lowe's

Figure 5.5: Average Monthly Frequency of Lowe's Top 2 Associated Brands

Shop Rite: From the bar chart, we see that the top 5 brands that are most associated to Shop Rite. We see that:



- 225 people come from Dunkin' to Shop Rite
- 180 people come from Sunoco to Shop Rite
- 162 people come from Rite Aid to Shop Rite
- 158 people come from Dollar Tree to Shop Rite
- 162 people come from Family Dollar Store to Shop Rite

Figure 5.5.1: Average Monthly Frequency of Shoprite's Top 4 Associated Brands

Bringing it all together

The chart on the next page shows all the big retailers and the top 2 associated brands with respect to each one in 2020. It is interesting to note that **Dunkin, Rite Aid and Sunoco** are three brands that correspond to more than one main retailer. For instance, Dunkin' is associated with B.J.'s Wholesale, Fresh Grocer, ShopRite and Walmart. Rite Aid is connected to Fresh Grocer,

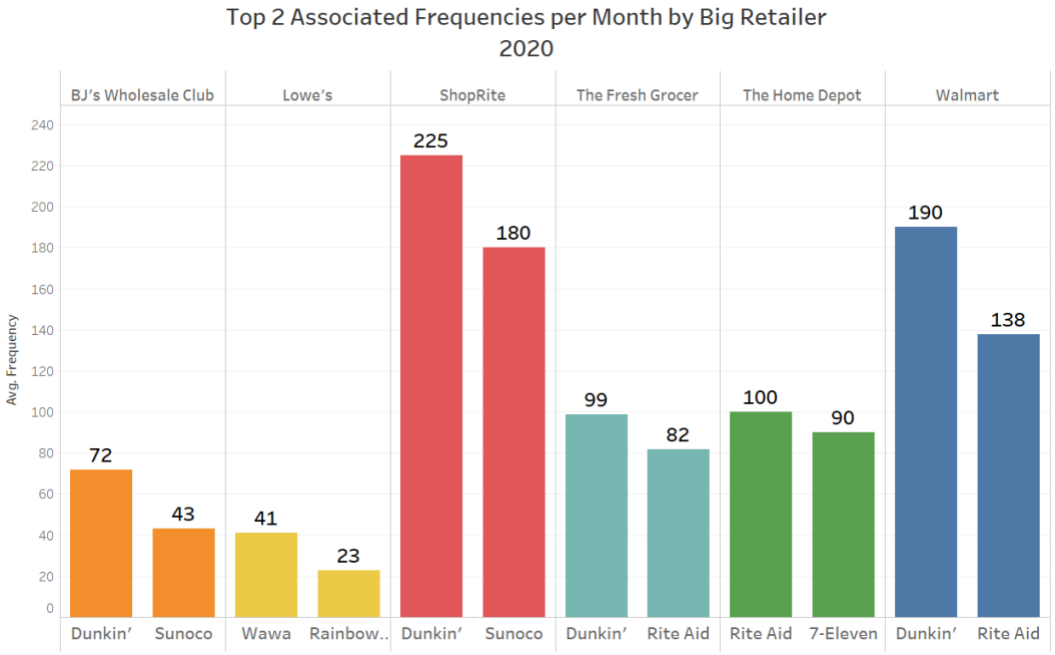


Figure 5.5.2: Average Monthly Frequency of every Big Retailer's top 2 Associated brands In 2020

When we compare with 2018/2019 as seen in the chart below, we still notice that **Sunoco and Rite Aid** are still the brands that are associated with more than one Retailer but the associations with the big retailer has been changed. For instance, Sunoco was associated with B.J.'s and Shoprite in 2020 however in 2019/2018, it was connected with Lowe's Fresh Grocer and Home Depot

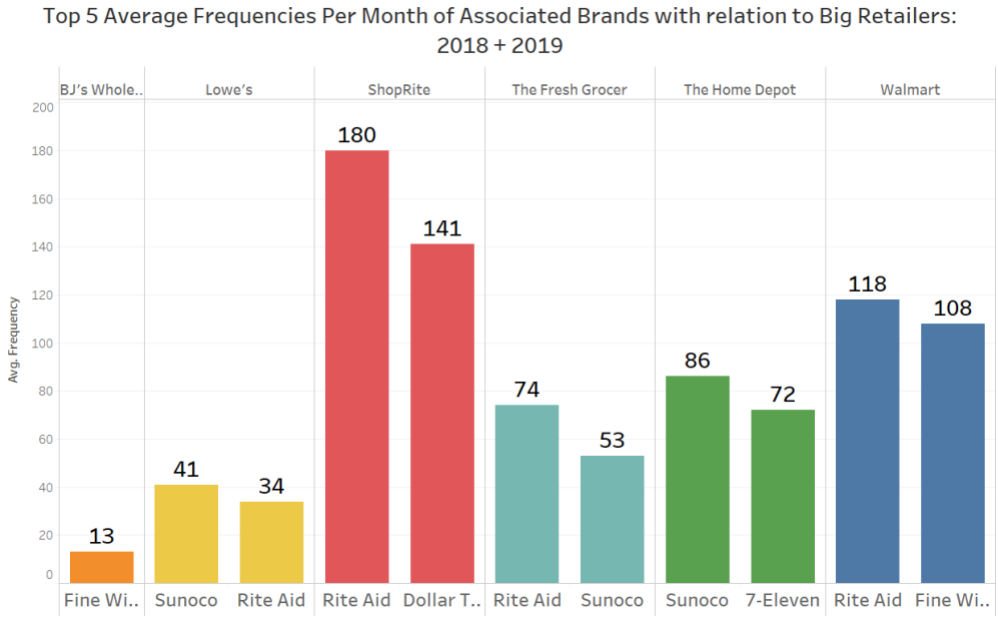


Figure 5.5.3: Average Monthly Frequency of every Big Retailer's top 2 Associated brands in 2018 and2019

To summarize, each big retailer has its two key brands that they need to focus on partnering to increase sales. For instance, The Fresh Grocer can focus on reaching out to Rite Aid and Sunoco. Likewise, brands like Sunoco and Rite Aid should not limit themselves to one main key Retailer.

It's interesting to note that these brands (not key Retail brands) are smaller retailer stores, gas stations, convenience stores, and coffee shops. These brand types most likely contribute to the majority of business in Philadelphia. Since the specific brands appear to be well known in their respective type (i.e. Dunkin is a well-known coffee brand, 7-Eleven is a well-known convenience store), the associations do not seem surprising in the grand scheme of things.

Business Insights/Conclusions/Recommendations

Conclusions:

- According to the overall decrease in average visits per visitor from February to May in 2020, the pandemic does impact brand dependency a lot, but different category shows different results. The trends are not similar, even in the same category. Locations, services, the number of stores, and policies are all playing roles in this situation.
- Customers pay relatively more visits to leading brands than to niche brands during the pandemic, and grocery stores especially brands that have many store chains in Philadelphia City have competitive advantages during the pandemic.
- COVID-19 has impact on retail industry, especially those stores located in shopping centers. But its impact on people's traveling distance is minor.
- The major retail brands are mostly associated with the following brand types: Convenience stores, Coffee stores, Gas stations and smaller retail stores via customer behavior.

Recommendations:

Some recommendations we make to businesses according to our findings in this report:

- Businesses should find alternative ways to increase the brands' dependency during abnormal situations to decrease visiting fluctuations.
- Niche brands could seek cooperation or promotion opportunities with leading brands to benefit from their brand reputation during the pandemic.
- Businesses could seek for more convenient ways to deliver their services to customers such as pick-up and home delivery.
- Each main retail brand has two key brands to partner with, in order to drive sales during 2020.

Limitations

About NAICS code

Even though the first two starting number are 44-45, which represents retail trade industry, most of common people's go-to stores are categorized in different industry NAICS codes, which is hard to filter and compare from SafeGraph dataset. It is easy to mis-identify all grocery stores are in the same industry. For example, Wholefoods market is in Supermarkets and other grocery (except convenience); Costco is in All other miscellaneous store retailers (except Tobacco stores) category; Wegmans is categorized in General warehousing and storage, whereas Walmart and Target are categorized in Department stores. As we may consider businesses mentioned above are retail brands competing, they are categorized under adifferent NAICS code with different competitors. Taking NAICS code of 452210 as an example, we can see Target is categorized as Department store and it competes with businesses that arelocated in malls. Indeed, except selling groceries, Target also sells apparel, household goods, and electronics.

Moreover, most businesses have multiple NAICS codes. Data from SafeGraph only capture the primary NAICS code of the business. Walmart and Target have the same NAICS code of 452210 according to the official NAICS association website. However, Walmart has been categorized as 452319 in SafeGraph dataset, which is easily to neglect if we compare by only one NAICS code and hard to filter desire businesses to form our list.

In conclusion, the NAICS code approach has its advantage yet not suitable for our objective.

About the sales and financial report

We couldn’t find the local sales or financial report in Philadelphia for each brand. It is a pity that we cannot combine the information to offer a more comprehensive business recommendation.

References

[1] *IBM Study: COVID-19 Is Significantly Altering U.S. Consumer Behavior and Plans Post-Crisis*. IBM Newsroom, newsroom.ibm.com/2020-05-01-IBM-Study-COVID-19-Is-Significantly-Altering-U-S-Consumer-Behavior-and-Plans-Post-Crisis. Accessed Dec 2, 2020.

[2] *McKinsey: 75% of Americans Have Changed Brands during the Pandemic* Vala Afshar, www.zdnet.com/article/mckinsey-three-factors-drive-consumer-loyalty-in-the-next-normal/. Accessed Dec 2, 2020.

[3] *What Is a NAICS Code and Why Do I Need One?* NAICS Association, www.naics.com/what-is-a-naics-code-why-do-i-need-one/#:~:text=A%20NAICS%20(pronounced%20NAKES)%20Code. Accessed Dec 2, 2020.

[4]*NAICS Code Description.* NAICS Association, www.naics.com/naics-code-description/?code=44-45. Accessed Dec 2, 2020.

[5] *New Brand Dependence Index Finds Consumers "Can't Live Without" Samsung Electronics* Amy Gesenhues, https://marketingland.com/new-brand-dependence-report-claims-consumers-cant-live-without-samsung-microsoft-electronics-70150. Accessed Dec 2, 2020.