

# Girl Scouts Cookie Booth Locations Project

Anju Nair

February 26, 2021

## 1 Business Understanding

### 1.1 Background

Girl Scouts is a pre-eminent leadership development organization for girls. And with programs from coast to coast and across the globe, Girl Scouts offers every girl a chance to practice a lifetime of leadership, adventure, and success. Their mission is to build girls of courage, confidence, and character, who make the world a better place.

Their extraordinary journey began more than 100 years ago with the original G.I.R.L.(Go-getter, Innovator, Risk-taker, Leader), Juliette Gordon “Daisy” Low. On March 12, 1912, in Savannah, Georgia, she organized the very first Girl Scout troop, and every year since, we’ve honored her vision and legacy, building girls of courage, confidence, and character who make the world a better place. The organization is 2.5 million strong—more than 1.7 million girls and 750,000 adults who believe in the power of every G.I.R.L. (Go-getter, Innovator, Risk-taker, Leader) <sup>™</sup> to change the world.

### Girl Scouts cookies

When you buy Girl Scout Cookies, you feed joy and make a difference with every box. The scout may help the local hospital or animal shelter across town, climb her first mountain or start a nonprofit. It’s up to her and her troop. Proceeds stay local to power life-changing programs, experiences, and learning all year long! When girls participate in the Girl Scout Cookie Program, they also develop important life skills—goal setting, decision making, money management, people skills, and business ethics—that will set girls up for success beyond anything they can imagine.

That’s why buying the delicious cookies you crave from your local Girl Scout Cookie business pro is important.

### Meet the Cookies

- *Thin Mints® - Crisp wafers covered in chocolaty coating. Made with natural oil of peppermint*
- *Samoas® - Crisp cookies, coated in caramel, sprinkled with toasted coconut, and striped with dark chocolaty coating*

- *Tagalongs® - Crispy cookies layered with peanut butter and covered with a chocolaty coating*
- *Trefoils® - Delicate-tasting shortbread that is delightfully simple and satisfying*
- *Girl Scout S'mores® - Crunchy graham sandwich cookies with creamy chocolate and marshmallowy filling*

Most people have a special place in their heart for Girl Scout Cookies. Not only because they're spectacularly delicious but also because the Girl Scout Cookie Program helps girls fulfill their dreams, follow their passions, and change the world!

The program, which is the largest girl-led entrepreneurial program for girls in the world, helps Girl Scouts fund unique adventures for themselves and their troops all year long. It also allows them to give back to the causes they're most passionate about—it's the Girl Scout way!

## **1.2 Problem**

Every year the girls have gone door-to-door to realize their individual and troop goals and contribute with confidence toward their community and troop. As you can gather the sale of girl scouts' cookies are important business and something that each girl scout looks forward to every year.

With COVID-19, door-to-door sales have become next-to-impossible.

On request from my very own fifth grader, a proud girl scout trooper, is a request to find the next possible solution.

Cookie booth sales at the local neighborhood would be an excellent solution to help my daughter's troop and the other troops in town to achieve their goals.

The problem is to find the top 10 prospective locations to set up cookie booths in our town, Grasmere, Staten Island, NY and one or two neighboring towns within a 3-mile radius to maximize sales. And, to make the Girl Scouts troop a happier lot in the process

## **1.3 Interest**

This analysis will be of great interest to my daughter's Girl Scouts troop and other troops in town to realize their troop's cookie goals. This information may turn out to be helpful while planning any other fund-raising initiatives which might require reaching out to a wider community.

## 2 Data Understanding

### 2.1 Data Sources

The data sources used for analyzing and gathering insights are:

Access to geospatial coordinates (latitude and longitude) will be required. This is provided by the course as a JSON file.

Neighborhoods to target and compare will be fixed to my town, Grasmere, Staten Island, NY, and the neighboring towns within a 3-mile radius.

To compare prospective locations in town and neighboring towns, access to location data is imperative. Location data is data describing places and venues, such as their geographical location, their category, working hours, full address, and so on, such that for a given location given in the form of its geographical coordinates (or latitude and longitude values), one is able to determine what types of venues exist within a defined radius from that location. There are many location data providers like Foursquare, Google Places, and Yelp. Providers like Foursquare gives a very wide geospatial coverage free of cost and has rate limits of 95,000 regular API calls and 500 premium calls in a day. They are known to provide comprehensive data that is both accurate and updated. **Foursquare will be the preferred choice for extracting location data.**

To target venues to achieve maximize sales, knowing the venues that are trending at the time of setting the booths will be most crucial. Trending venues means places with the highest foot traffic. The Foursquare API provides this data and the venues with the highest foot traffic is fetched live. The only drawback is since it is real-time the data will depend on when the code is run. To take care of this issue, the following method can be used.

As COVID-19 stay-at-home orders begin to lift, the venues people will start to visit more often and their order of preference will be worth taking into account. Foursquare Recovery Index shows the places consumers go to first as the restrictions are lifted in near real-time. The venues listed by the Foursquare Recovery Index will be considered.

### 2.2 Data Cleaning

The geospatial coordinates of New York(NY) was downloaded and loaded into the dataframe.

The dataframe has 5 boroughs and 306 neighborhoods.

Out[5]:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

The above data with Borough, Neighborhood and respective latitude and longitude was extracted from the given JSON file.

To restrict this analysis to my town of Grasmere, Staten Island, NY and the neighboring towns, the dataframe was filtered on the neighborhoods of interest.

---

The neighborhoods of interest are 8 neighborhoods.

Out[19]:

	Borough	Neighborhood	Latitude	Longitude
2	Staten Island	Stapleton	40.626928	-74.077902
7	Staten Island	South Beach	40.580247	-74.079553
24	Staten Island	Park Hill	40.609190	-74.080157
28	Staten Island	Arrochar	40.596313	-74.067124
29	Staten Island	Grasmere	40.598268	-74.076674
30	Staten Island	Old Town	40.596329	-74.087511
48	Staten Island	Shore Acres	40.609719	-74.066678
50	Staten Island	Concord	40.604473	-74.084024

Using the Foursquare API, I was able to extract the location data for each of the neighborhoods of interest. A sample of the data is shown below.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Stapleton	40.626928	-74.077902	Lakruwana	40.625654	-74.075174	Sri Lankan Restaurant
1	Stapleton	40.626928	-74.077902	Defonte's Sandwich Shop	40.627202	-74.077257	Sandwich Place
2	Stapleton	40.626928	-74.077902	The Hop Shoppe	40.629034	-74.079758	Beer Bar
3	Stapleton	40.626928	-74.077902	Every Thing Goes Cafe and Bookstore	40.636920	-74.076661	Café
4	Stapleton	40.626928	-74.077902	Dunkin'	40.631129	-74.076879	Donut Shop

Taking the new pattern of how people are returning to shop as the COVID-19 stay-at-home orders begin to lift, the top 30 venue categories from the Foursquare Recovery Index is picked for the week that we plan to setup the booth. The above dataframe will need to be filtered for these 30 venue categories. Below is a sample of the venue categories listed.

## New York by All Categories

DATA IS GROUPED WEEK-OVER-WEEK ON A 7-DAY ROLLING AVERAGE

Visits		
Categories ↑↓	Week over Week ↓	vs. pre-COVID-19 / Feb 2020 ↑↓
Beaches	14.3%	46.4%
Outdoors	13.4%	25.4%
Bank	12.9%	11.5%
Skiing	12.3%	120.6%
Medical	12.3%	12.4%
Metro	11.9%	-37.2%
Auto Shops	11.1%	33.4%
Bars	10.8%	-7.5%
Transit	10.2%	-27.9%
Airports	10.1%	-48.8%
School	9.7%	-5.6%

Considering we are setting up cookie booths for girl scouts in March, some of the irrelevant venue categories will need to be removed. For example, beaches, skiing, medical, airports and bars.

Also, looking closely at the above data and comparing it to the venue categories in the neighborhoods of interest, there were some problems. For example, the category, Restaurant, is in the recovery index list but the venue categories from Foursquare is more drilled down to each type of restaurant, for example, 'Sri Lankan Restaurant', 'Tapas Restaurant' and 'Seafood Restaurant'. The data was fixed to reflect 'Restaurant' for every type of restaurant.

I checked for missing data but there were no rows found.

The features that would be of interest would be Neighborhood, Venue name, Venue category, and Venue location.