

# Analyzing venues nearby the Airbnb listings in Staten Island, N.Y.

Oscar Aguilar  
July 2020

## 1 Introduction

### 1.1 Background

Since December 2019, the world has encountered several challenges in the healthcare sector due to the outburst of a novel virus: COVID-19. Furthermore, the impact of this situation has reached several major fields with economic, political and social implications; for instance, tourism. Since the commencement of the lockdown in many countries, different travel restrictions have emerged (locally, regionally and internationally). The latter has affected SMEs and globally recognized companies such as Airbnb (Gössling, 2020).

In late June 2020, Brian Chesky (CEO of Airbnb) publicly expressed: “It took us 12 years to build Airbnb, and we lost almost everything in four to six weeks” (Entrepreneur, 2020). He added that the future of the company is uncertain since the future of travelling will not be the same until there is a vaccine that can fight this pandemic. The company has published drastic statistics using cancellations and availability information of their listings worldwide: after the Wuhan lockdown, their booking activities showed a fall of 57.8% (Hu and Lee, 2020), therefore the company laid off 25% of their employees and raises a debt equivalent to \$2 billion in equities (Pavlovskaya, 2020).

Nowadays, the governments have started to reactivate the economy of their respective countries. Society has to learn to live a new normality with extreme precautions. Chesky said that since people is not getting on airplanes (or crossing borders), there is an urge to promote local travelling. The latter because people are showing a clear preference to travel by car if required. Hence, with some precautions in mind, people could start making small trips to near-local areas in order to support small businesses and the accommodation industry. Some precautions Chesky suggested were asking the owner of the property: when the listing was rented, what is being done to clean it, if it is being rented continuously, and local-or-regional restrictions (Krstic, 2020). In addition, it is preferable to wear masks upon arrival, bring sanitizer and disinfectants, as well as an air purifier, to mention a few.

### 1.2 Problem

Considering New York, N.Y. as one of the most preferred and concurred cities by tourists in the world, it will be of paramount importance to analyze the venues offer of a smaller-and-near local area such as Staten Island. Using the location of Staten Island’s Airbnb listings, it will be possible to find the kind of venues around them. In this way, local people near the zone from will be able to see which kind of businesses and places are around a potential Airbnb accommodation. This will benefit local tourists, Airbnb, listings’ owners and local businesses.

## 2 Data Acquisition and Preprocessing

### 2.1 Data Sources Description

In order to perform this analysis, two different sources of data will be utilized. In this section of the report, each data set will be described thoroughly. The description includes their respective source, features, and origin.

#### 2.1.1 New York's Airbnb Data Set

This first data set was extracted from Kaggle as a csv file. The data retrieval process started in 2008, with the aid of guests and hosts of Airbnb from N.Y. They contributed by providing information that describes the listings' generalities, availability and metrics. This continued until 2019, and three different version of the data set have been built. This project utilizes the most updated version. The included features are listed in the table below:

Table 1. N.Y. Airbnb Data Set Feature Description

<u>Feature</u>	<u>Type</u>	<u>Description</u>
Id	Integer	Identification number of the listing
Name	Object	Name of the listing
Host_id	Integer	Identification number of the host
Host_name	Object	Name of the host
Neighbourhood_group	Object	Boroughs of N.Y. (e.g. Manhattan, Queens, Staten Island)
Neighbourhood	Object	Neighborhoods within the boroughs
Latitude	Float	Latitude coordinate
Longitude	Float	Longitude coordinate
Room_type	Object	Type of listing: private, shared, etc.
Price	Integer	Price per night
Minimum_nights	Integer	Minimum required number of nights to book
Number_of_reviews	Integer	Total number of reviews of the listing
Last_review	Object	Date of the last review
Reviews_per_month	Float	Average number of reviews the listing gets per month
Calculated_host_listings	Integer	Number of listings that the host owns
Availability_365	Integer	Number of days/year available for booking

### **2.1.2 Foursquare location data**

The second source of information are the location-based services offered by Foursquare which consist of a RESTful service used to request JSON or XML data. Foursquare is a free location discovery app that allows users to find and share information about businesses and attractions in any part of the world. The Foursquare API allows developers to interact directly with their data platform. The different API methods allow developers to retrieve check-ins, venues, categories, tips, menus, among other data. Since these requests require authentication, programmers are required to create an account, which can be upgraded by a monthly fee. An upgraded account has more perks regarding the number of calls that can be done per day, the ability to make premium calls, etc.