

# INFS 5700 GROUP 2 TEAM ASSIGNMENT

Z5199975 JING YANG

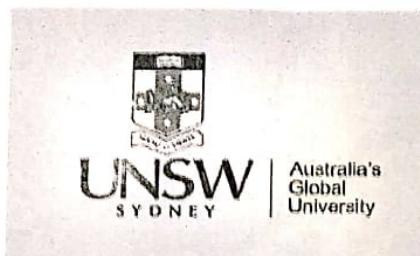
z5220102 Somadina Nnaji Chukwu

z5228145 Aarushi Gera

z5226236 Niketa Patel

z5251470 Mohamad Haekal Irsananda

Z5212497 Yixin Wu



UNSW Business School  
Information Systems and Technology Management

## ASSESSMENT COVER PAGE

Title of Assignment: INFS 5700 GROUP 2 TEAM ASSIGNMENT

Date Due: 4/04/2019

Date Submitted: \_\_\_\_\_

I/We declare that this assessment item is my/our own work, except where acknowledged, and has not been submitted for academic credit elsewhere, and acknowledge that the assessor of this item may, for the purpose of assessing this item:(i)Reproduce this assessment item and provide a copy to another member of the University; and/or,II)Communicate a copy of this assessment item to a plagiarism checking service (which may then retain a copy of the assessment item on its database for the purpose of future plagiarism checking).

I/We certify that I/we have read and understood the University Rules in respect of Student Academic Misconduct.

ID	Name	Signature	Date
z5199975	Jing Yang	BmYR	4th APR 2019
z5251470	Mohamad Haekal Irsananda	nnj	4th APR 2019
Z5212497	Yixin Wu	Yi	4th APR 2019
Z5220102	Somadina N. Chukwu	SN	5th APR 2019
Z5226236	NIKETA JAGDIPCHANDRA PATEL	N.G.Patel	4th APR 2019
Z5228145	AARUSHI GERA	Aarushi	4th APRIL, 2019

This cover sheet has to be completed and signed by ALL members in the group for the assignment submitted. Note: 10 percent of the marks available for the assessment will be deducted for assessments submitted without a fully completed and signed cover page.

# 1 Science Profile of the Members

Our group have 6 members with balance skill set for the adequate job of System analyst and data exploration of Airbnb Chicago. Among the six skills and talent required to create a good data analyst, group 2 purposes 4 specialists which are included in the table below:

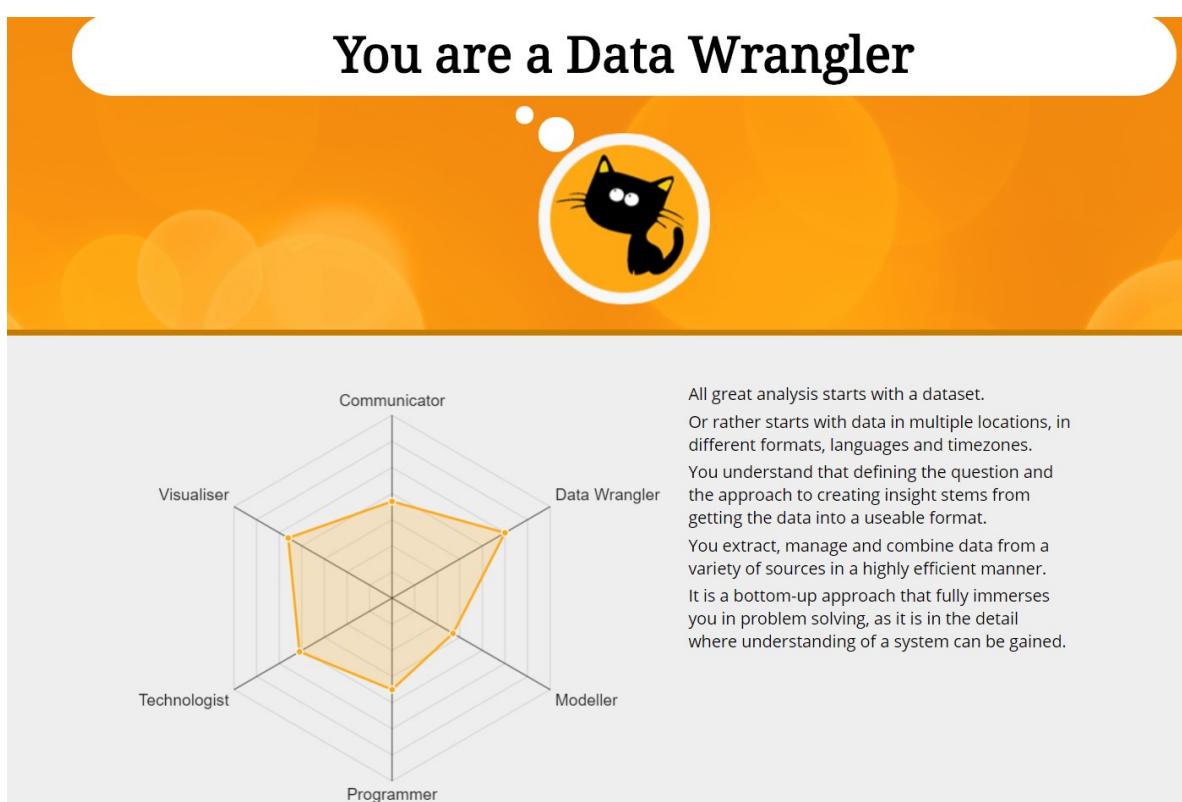
Group 2 Data Science portfolio Compilation with Skill Set strength			
	Group Member Name	Special Talent	Special assigned task
1	Somadina Nnaji Chukwu	Communicator	Task 1 Introduction
2	Jing Yang	Data wrangler	Task 2 Data exploration
3	Yixin Wu	Data wrangler	Task 2 Data exploration
4	Aarushi Gera	Communicator	Task 3 Recommendation
5	Mohamad Haekal Irsananda	Data modeler	Task 3 Recommendation
6	Niketa Patel	Data wrangler	Task 4 Slide-deck

Table 1-1 Special talent and assigned task summary

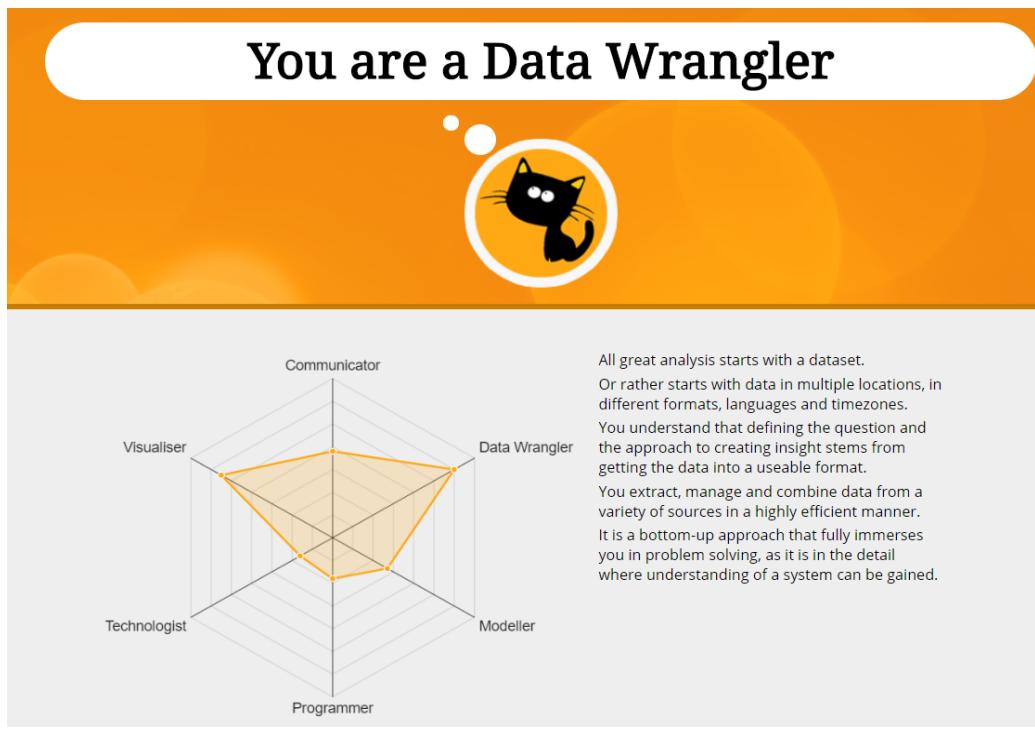
Using the data science profile grading, group 2 was able to assign tasks to individual members according to their potential strength. Figure 1-1 to figure 1-6 gives the detail of our members' special talents. This variations in members talent and skills create a balance on exploiting and performing data exploration with the Airbnb Chicago. The group 2 data science individual profiles are included below:



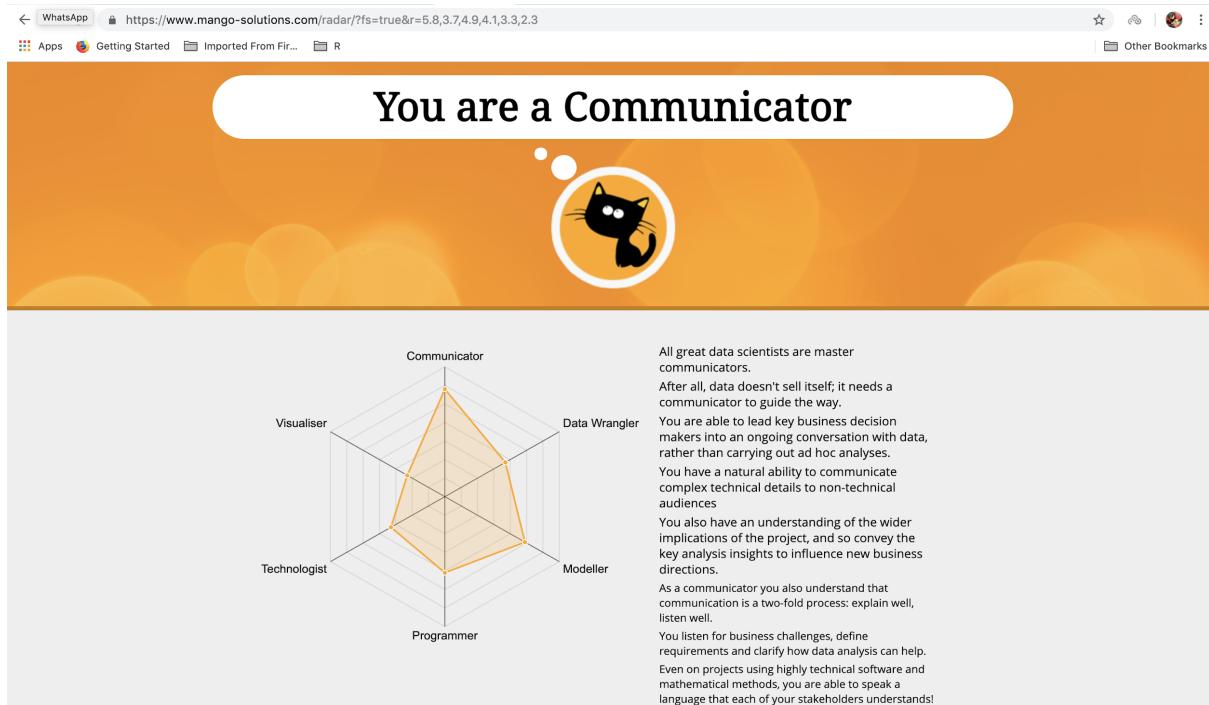
**Figure 1–1** Science Profile of Somadina Nnaji Chukwu



**Figure 1–2** Science Profile of Jing Yang



**Figure 1–3** Science Profile of Yixin Wu



**Figure 1–4** Science Profile of Aarushi Gera

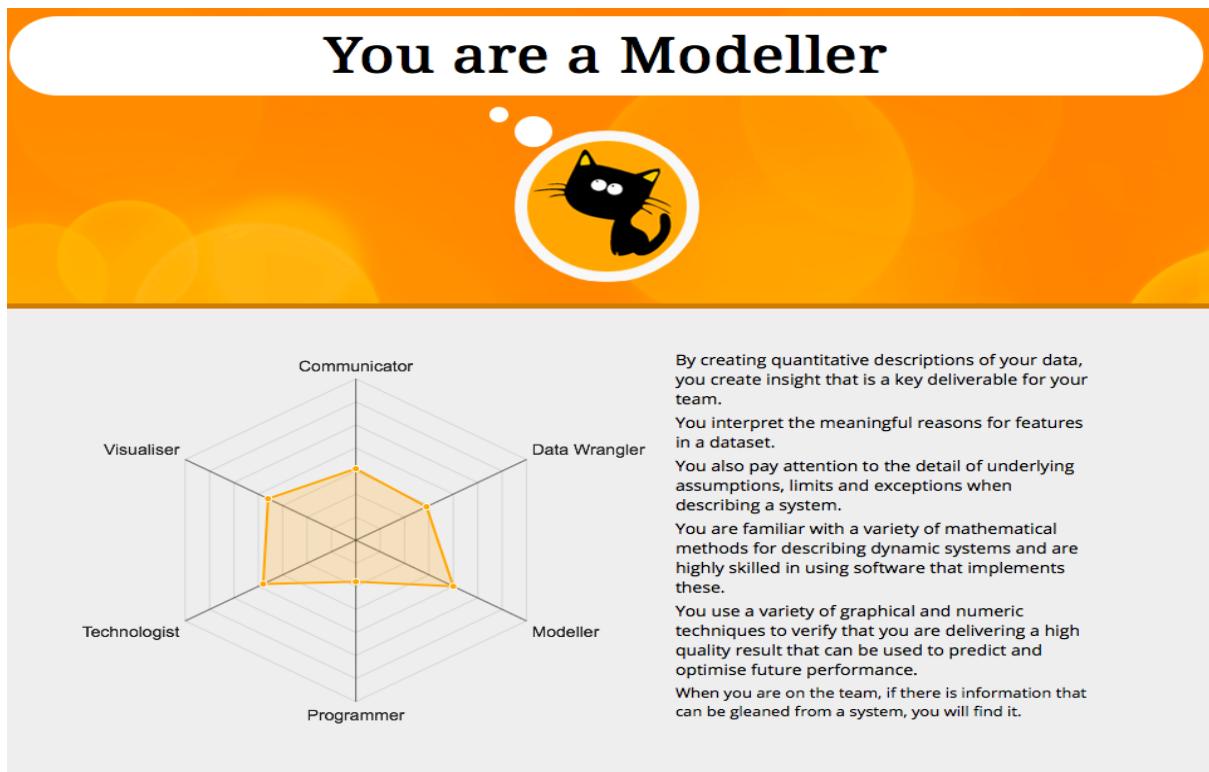


Figure 1–5 Science Profile of Haekal Irsananda

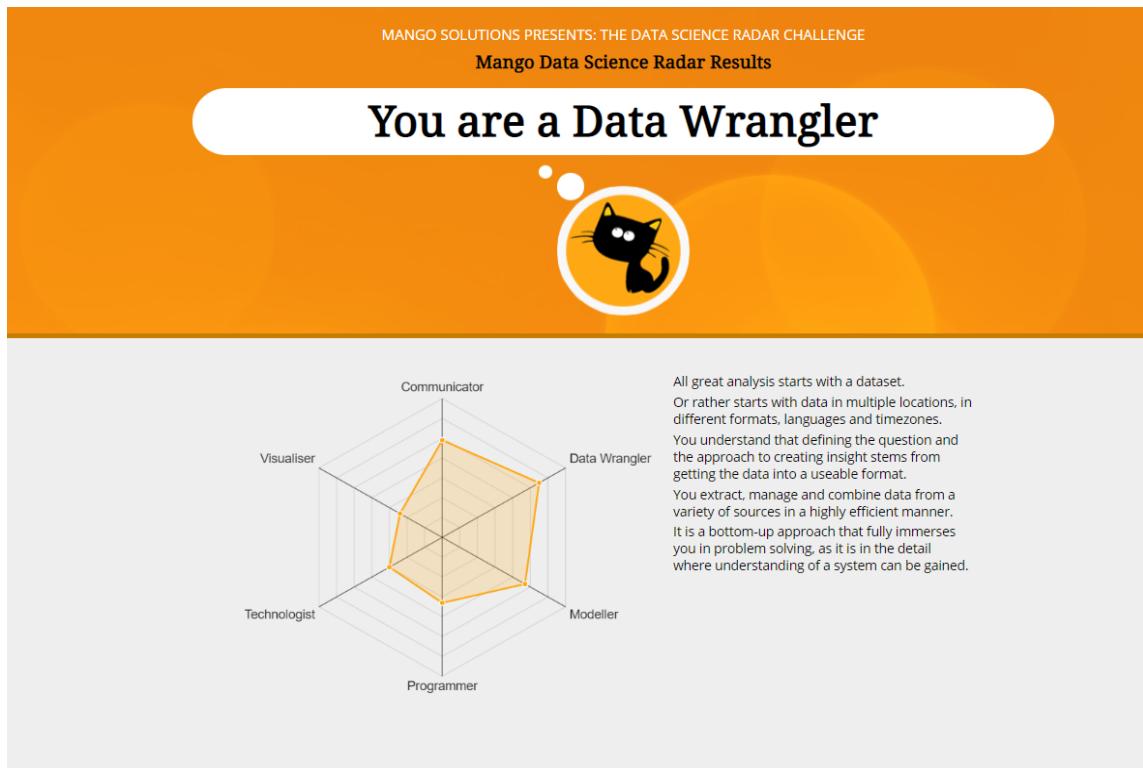


Figure 1–6 Science Profile of Niketa Patel

## 2 Introduction

Airbnb is a company that deals with properties as a multi-sided platform. They help owners to rent their properties (homes) to peoples (travellers, backpackers or bloggers) who want to stay in that home. Airbnb Host Year dataset implies that the business has been around since 2008 since the information provided in this dataset is consistent from 2008 to 2018.

It is assumed that the company that is trying to understand the Airbnb market in Chicago is planning to invest in real estates then rent them on Airbnb. To find a good investment and earn profit from this business, they need suggestions about:

1. Where is the most popular place in Chicago on Airbnb?
2. What types of property have high demand?
3. After start renting on Airbnb, how to attract more customers?

After a series of research, our team provide a brief analysis of questions listed above through data exploration. We also provide a short recommendation on the way to extend the dataset with add-open data, unstructured data or social media data.

## 3 Data Exploration

### 3.1 Assumptions and processing methods

To conduct data exploration efficiently, the following assumptions are necessary:

1. The company we report to is planning to invest in real estate in Chicago and rent them on Airbnb.
2. All Airbnb hosts in Chicago stated their location as ‘Chicago, Illinois, U S’ in the database.
3. The review to order ratio is constant. It means that the more popular a house is, the more result it gets. Thus, review\_per\_month is a reflection of customer per month.
4. The number of guests included will match the capacity of the accommodation. That is, the accommodates will be equal to guests\_included. Thus, price per person indicates the price level of a property.
5. New hosts listed on Airbnb should have a profile picture and are business travel ready (the existing hosts meet the requirements).

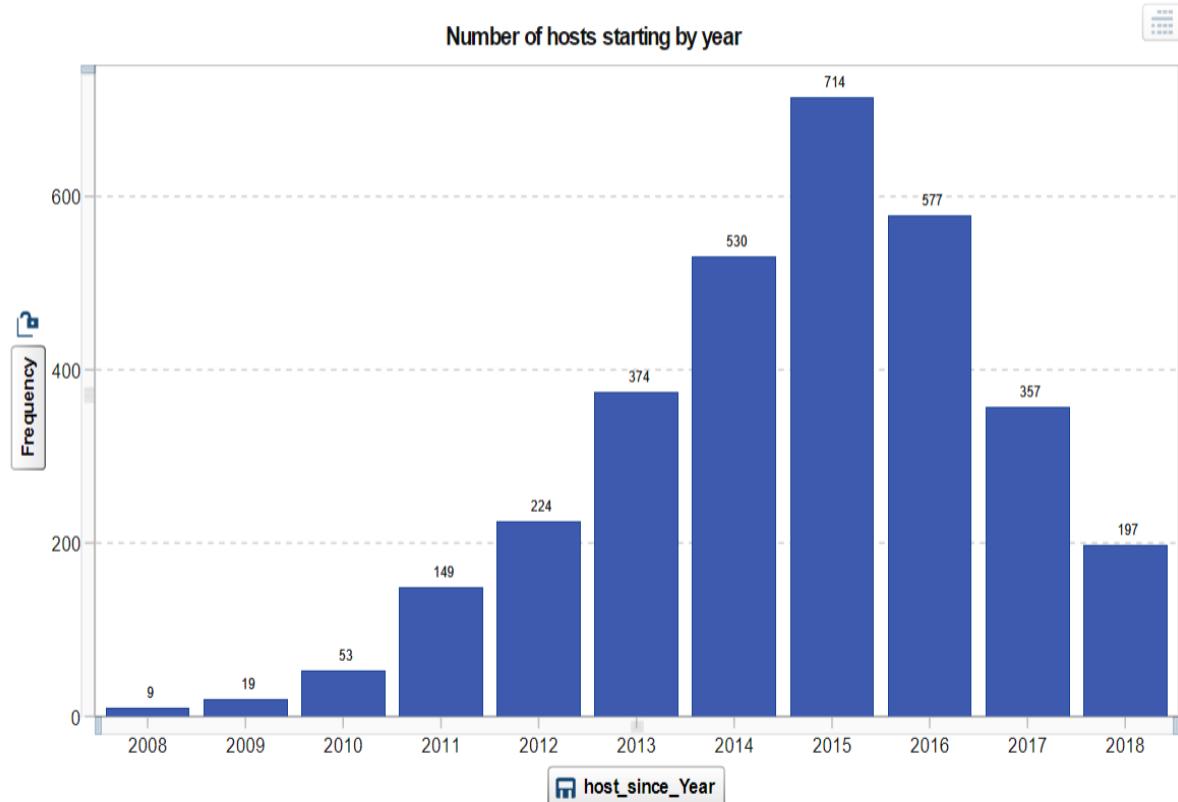
The listed processing methods apply to all visualisations:

1. Host\_Location was filtered, only “Chicago, Illinois, United States” was selected in the filter panel, to ensure only Chicago market would be included in the analysis.
2. Missing values were removed for all variables used in the visualizations.
3. Price per person is calculated by price divided by accommodates. Based on assumption 4, it can be concluded that the level of price per person is an indicator of the level of total price that customers undertake. Since given the number of guests, the higher the price per person, the higher the total price.

### 3.2 Feasibility of entering the Chicago Airbnb market

For those who would like to enter Airbnb market in Chicago, the first question they might ask is: Is 2019 a good time to enter this market? Are guests willing to choose accommodations from new Airbnb hosts?

The answer is affirmative. Figure 3-1 shows that the number of new hosts starting their Airbnb business peaked in 2015 and have kept decreasing since then. Although there are fewer hosts entering in Airbnb market, the new hosts got more orders as their review per month are higher, shown in figure 3-2. In fact, those hosts started in 2018 received the most order. Those two visualisations imply that there are low entry barriers in the Airbnb market.



**Figure 3-1:** Number of hosts starting their business on Airbnb every year. Drawn by select host\_since\_Year in the auto chart. Host\_since\_Year is a customer category of the host\_since variant, Created by grouping the years of the dates in host\_since.

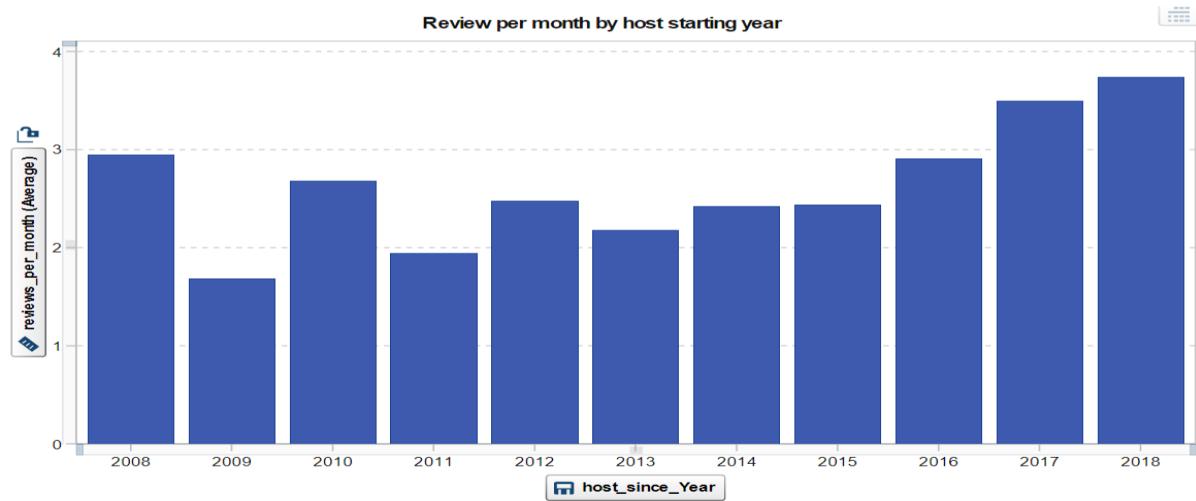


Figure 3-2: Average review per month by host starting year.

### 3.3 Preferred areas in Chicago for renting a house

After deciding to enter the Airbnb market in Chicago, the first task is to determine the place of the accommodations that would be purchased and rented on Airbnb. To earn a stable income, it is recommended that accommodations should be placed in a popular location, areas where the average number of monthly reviews and the average price is high. Figure 3-3 gives the correlation between place (zip code), popularity (reviews per month), and price. In figure 3-3, a larger bubble suggests a higher average monthly review and a deep green colour suggests a higher average price per accommodate. This figure implies that the places near the city and the lake, where there are many large green bubbles, could be optimal choices. Figure 3-4 gives ten zip codes, in which the properties have the highest number of monthly reviews.

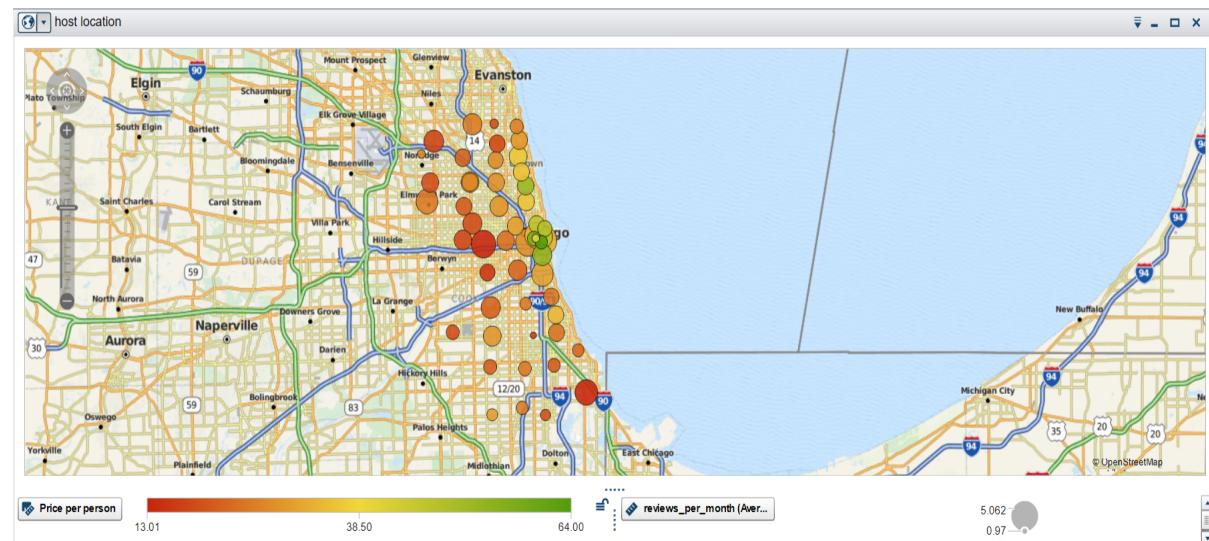
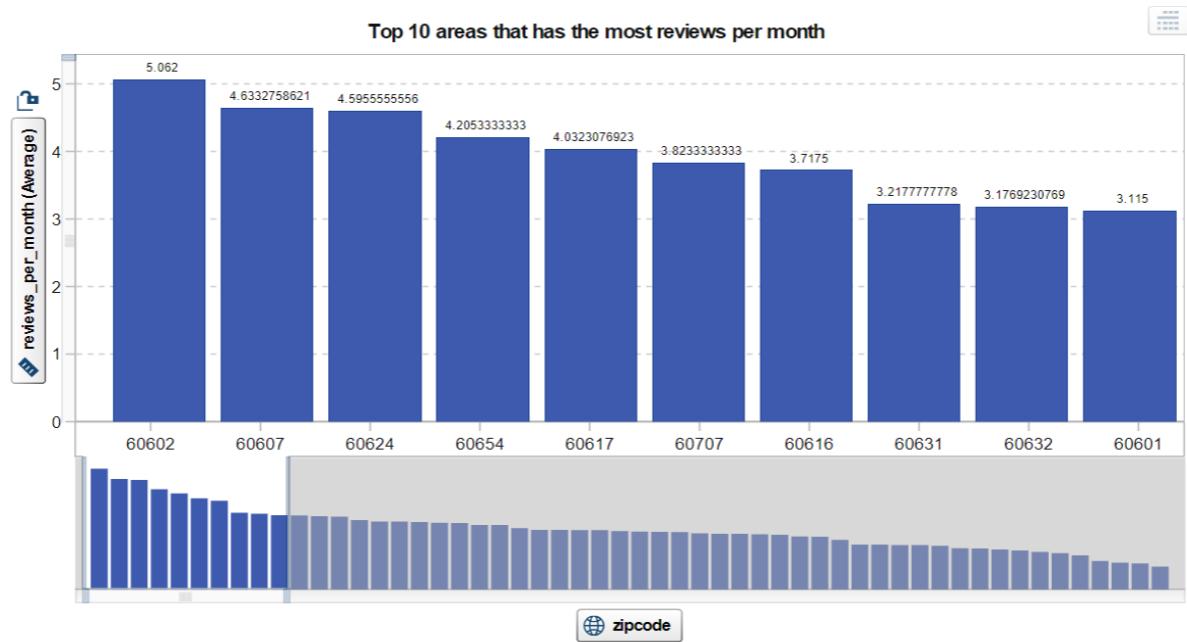


Figure 3-3: The correlation between hosts' zip code, average review per month, and price per accommodate. Drawn by selecting zipcode as geography, bubble size as review\_per\_month and colour as Price per person. Price per person is calculated by using price divided by accommodates and is an indicator of the price level.



**Figure 3-4:** Ten zip codes where the number of the monthly review is the highest.

### 3.4 Supply side: Accommodates, property and room types analysis

The property type and numbers of accommodates allowed could be important factors that a customer would consider when choosing places on Airbnb. Figure 3-5 gives the distribution of property types and numbers of accommodations in Chicago Airbnb market. It shows that the majority types of accommodations available on Airbnb are apartments, condominiums and houses. Also, there is more small-size property (where accommodates  $\leq 4$ ) in Airbnb market.

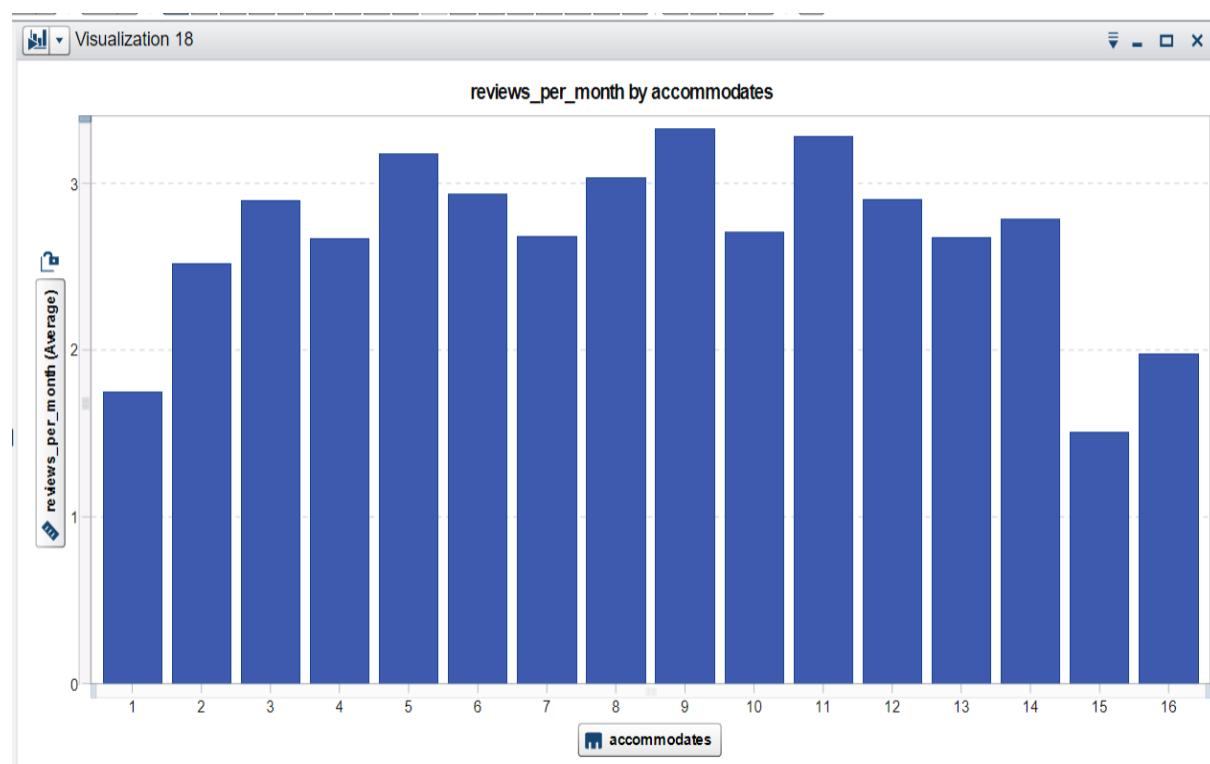
accommodates grouped by property_type																	
property_type	Apartment	Bed and breakfast	Boat	Boutique hotel	Bungalow	Castle	Condominium	Cottage	Guest suite	Guest house	Hostel	House	Loft	Other	Serviced apartment	Tiny house	Townhouse
accommodates	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency
2	578	3	1	2	18	.	183	1	22	8	3	219	27	2	1	.	41
4	339	1	1	2	5	.	105	1	18	5	1	56	16	.	8	1	10
6	246	1	.	.	3	.	57	1	6	1	.	41	8	.	5	.	11
5	159	.	.	.	.	.	43	.	1	3	.	19	7	.	1	.	3
3	157	.	.	1	1	.	29	.	11	.	1	26	8	.	.	.	6
8	93	.	.	.	1	.	32	.	.	2	.	21	5	.	1	.	11
1	92	.	.	.	2	.	20	.	.	.	6	74	8	3	.	.	3
7	52	.	.	.	.	.	15	.	.	.	.	10	3	.	2	.	2
10	50	.	.	.	.	.	10	.	.	.	.	13	.	.	.	.	5
9	13	.	.	.	1	.	6	.	.	.	.	8	.	.	1	.	1
16	13	.	.	.	.	1	3	.	.	.	.	20	2	.	.	.	.
12	8	.	.	.	1	.	2	.	.	1	.	10	1	.	.	.	2
11	5	.	.	.	.	.	2	.	.	.	.	3	.	.	1	.	1
14	5	.	.	.	.	.	2	.	.	.	.	4	1	.	1	.	4
13	3	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	2
15	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2

**Figure 3-5:** Number of accommodates grouped by property\_type. Drawn by select property\_type in the columns and accommodates in the rows in Crosstab.

The analysis above gives an understanding of the types of accommodation provided by hosts in Chicago. However, the types demanded by customers are slightly different from the types supplied.

### 3.5 Demand side: Accommodates, property and room types analysis

Knowing what customers demanding is essential in an analysis. The number of reviews per month is used to indicate the number of orders per month. Thus the review number will also indicate the level of demands for each property. Figure 3–6 shows that properties that could accommodate more than four have more reviews per month. This implies that there is a higher demand for large-size properties which can accommodate 5 to 14 people. There is insufficient demand for extra-large properties where it can accommodate more than 15 and small or shared properties where capacity is below 2.



**Figure 3–6:** Average number of reviews per month for different numbers of accommodates. The Accommodates variant was set as a category variant.

Figure 3–7 indicates the welcomed type of properties. The guest houses and guest suite has the most reviews per month, implying that this kind of property is scarce and have an insufficient supply. Among the most popular type of properties supplied in the market, the overall demand is decent although not outstanding. It is worth mentioning that the demand for apartments is significantly larger than that of houses and condominiums.

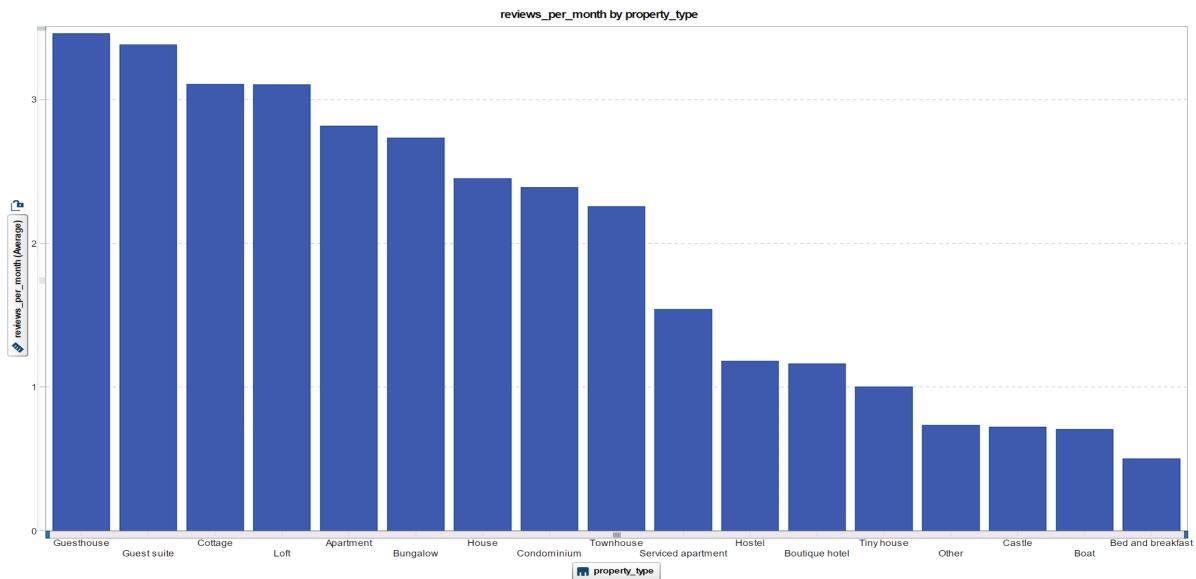


Figure 3-7: Average number of reviews per month for different property types.

In summary, there is a gap between the supply and the demand for property types and sizes (number of accommodates). The hosts in Chicago provides many accommodations for less than 4 people, while the number in the guest groups is sometimes greater than 4. Thus, we have an oversupply condition for small accommodation and a shortage for providing larger rentals to guests.

### 3.6 Superhost effect

Some hosts on Airbnb are highlighted as superhost by the website. They are the hosts who are top-rated and most experienced and are believed to provide better services to their customers. Those superhosts have a badge on their listing and profiles so that customers will recognise it when picking accommodations. Figure 3-8 shows the benefits of getting a superhost certification.

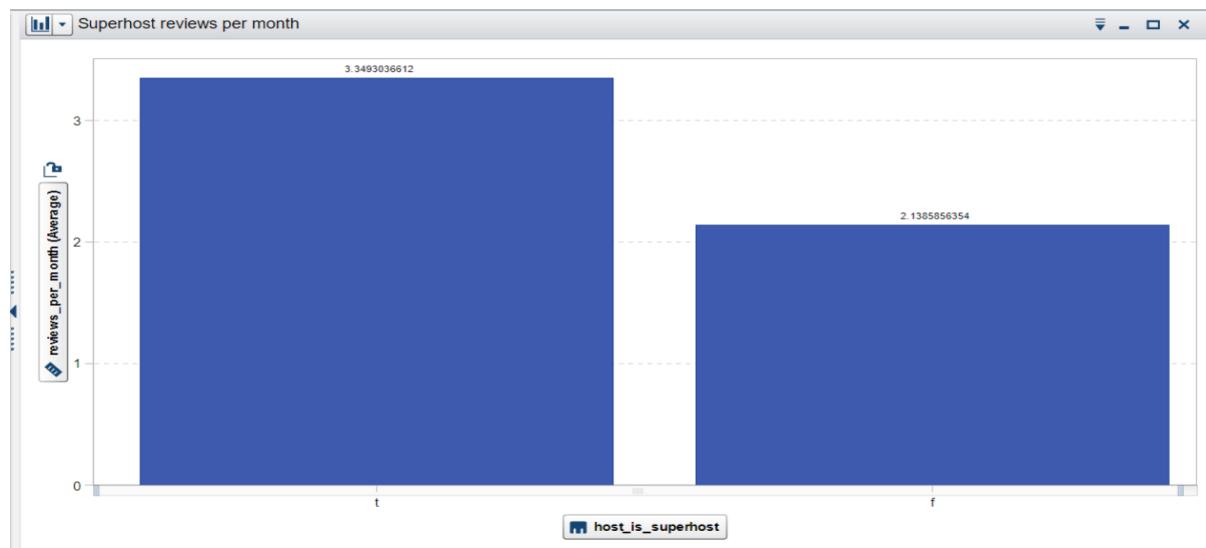


Figure 3-8: Average monthly Reviews for superhost and non-superhost.

As shown in the figure, the average reviews per month made by customers for superhosts are almost 50% higher than that of non-superhosts, indicating that superhosts could attract more customers. Therefore, if a host wants to make more money from Airbnb, he or she would be better to get a superhost certification.

## 4 Recommendation

### 4.1 Recommendation on the investment

The first recommendation is about the location of properties. Figure 3-3 and figure 4-1 show the location of 10 areas (determined by zip code) that have the highest number of review per month and prices of the properties (determined by colour) according to their location respectively.

As shown in figure 4-1, 5 of these regions are located near the centre part of the city. 4 of the rest regions are located close to a motorway, and one is near a park. Sableman (2018) suggests that Chicago is placed 24th in the United States crime rate in comparison to other cities. Hence, to focus on safety and to reduce the risk of property damaged due to crime, we recommend investing in properties that are located near the centre of the city.

The suburb that gained the most reviews per month is ZIP code 60602, located in northern Illinois and covers slightly less average land area compared to other ZIP codes in the United States.

Figure 3-3 shows that prices of Airbnb are higher near the beach than in the city. This fact might due to the sea view near the beach. The average nightly stay around the beach per person is \$64 compared to \$18 - \$38.50 around the CBD. Hence, we recommend, if you want to earn more from renting Airbnb, invest in properties near the beach as people are willing to pay more for the view. To be specific, there seems to have high demand accommodations along the Lake Shore Boulevard between Lee street beach and Greenwood street beach.



**Figure 4-1:** 10 places that have the highest number of monthly review. The size of the bubble indicates the number of monthly reviews.

The second recommendation is about the size and type of property. As mentioned above, we consider the number of views to indicate demand for each property. There is an over supply for accommodations that can accommodate less number of people ( $\leq 4$ ) and a shortage of properties that can accommodate more number ( $> 4$ ) of guests. We can see that the market for properties accommodating people less than equal to 4 has saturated and thus, we recommend that it would be advantageous to invest in properties that can accommodate greater than 4 number of people.

Since the highest number of reviews for Guest Houses and Guest Suites, as shown in figure 3-7, and the lowest number of reviews for Boat and Bed and Breakfast, according to our assumption, most people book Guest Houses and Guest Suites. Therefore, if the company wants to invest in property, they are recommended to invest in Guest Houses and Guest Suites since they are booked the most and it should not invest in Boat and Bed n Breakfast.

The third suggestion is about daily operation and maintenance. Since becoming a superhost can attract many customers, we highly recommend that all hosts on the multi-sided platform should try to get a superhost certification. According to Airbnb official web site (2019), to become a superhost, hosts will need to meet to keep an overall rating of 4.8 stars based on reviews from at least 50% of their Airbnb guests in the past year. This is done to let guests know they can expect outstanding hospitality. Another factor to be considered is the experience level of the hosts. Hosts will have to host at least ten nights or have had longer reservations that amount to 100 nights over the past three stays. Peace of mind is key for customers; hence, hosts will not be allowed any cancellations in the past year. The last criteria are response and reliability. Hosts will need to respond to 90% of new messages within 24 hours.

## 4.2 Recommendation on extending the dataset

Also, the firm can use extend the dataset by adding open data, unstructured data and social media data. With the help of sentiment analysis algorithms, this unstructured information could be automatically transformed into structured data of public opinions about products, service and brands which relate to Airbnb and can be added to the dataset. Sentimental analysis on social media can be used for commercial applications like marketing analysis, public relations, product reviews, net promoter score and customer service.

Open data refers to any data that are available and can be freely used for any purpose. Locations, events and external activities information can be obtained through the use of open data. Doing so, they can personalise their offering, highlight which events are upcoming and optimise their platform.

Unstructured data refers to data that does not fit into traditional row and column structure of relational databases. In the case of Airbnb, data such as videos, web pages and pictures of ID proof, profile pictures and property can be added to the current dataset to have a better understanding of the client preference and also add the level of security.

## 5 Conclusion

The company need an understanding of Airbnb market in Chicago to make decisions about investing in real estate and rent them on Airbnb. The analysis result shows that the company can enter the Airbnb market in Chicago in 2019. Consider multiple issues, the properties located near the centre of Chicago and the beach is recommended. The analysis suggests that large-size accommodation and guest houses are preferred since they have a high demand in the historical data. Finding that becoming a superhost can attract many customers, we recommend them to become a superhost and gives short advice on how to become a superhost. Finally, we explain how to make use of add-open data, unstructured data and social media data to extend the dataset.

## Reference:

Airbnb Inc. 2019, *How do I become a super host*, USA, viewed 3 April 2019, <<https://www.airbnb.com/help/article/829/how-do-i-become-a-superhost>>

Sableman, P. 2018, *The 30 cities with the highest murder rates in the US*, Madison, USA, viewed 3 April 2019, <[https://madison.com/news/national/the-cities-with-the-highest-murder-rates-in-the-us/collection\\_33b57fa7-d832-502d-ad55-aa29609b9a57.html#1](https://madison.com/news/national/the-cities-with-the-highest-murder-rates-in-the-us/collection_33b57fa7-d832-502d-ad55-aa29609b9a57.html#1)>