



Tableau Visualization on Airbnb's New York Listings



BAMA 520 Customer Analytics

Assignment 1

Group 13

Giselle (Yihan) Chen

Yuming Liang

Agnes (Siyao) Peng

Cathy (Jing) Yan

Q1. Summary Statistics

- There are around *40,000* Airbnb listings in New York.
- Price is fair in New York, but variance is relatively high.
- Most of the listings requires more than 1 nights per stay, and the average availability is *1/3 year* annually. Those characteristics are different from hotels.
- There are around *1.1M* comments and *1.6M* stays from 2009 to 2019. It may indicate that New York is an important market for Airbnb.

Type of Statistics	Number
No. of Listings	39,937
Average Price Per Night	\$142
90 th Percentile of Price	\$250
Minimum Night > 1	73.6%
Average Availability Per Year	126 Days
Total Number of Reviews	1,115,253
Total Demand (Stays)	1,593,219

Note:

Total Demand = Total Number of Reviews/70%

According to Airbnb, 70% of guests leave reviews.

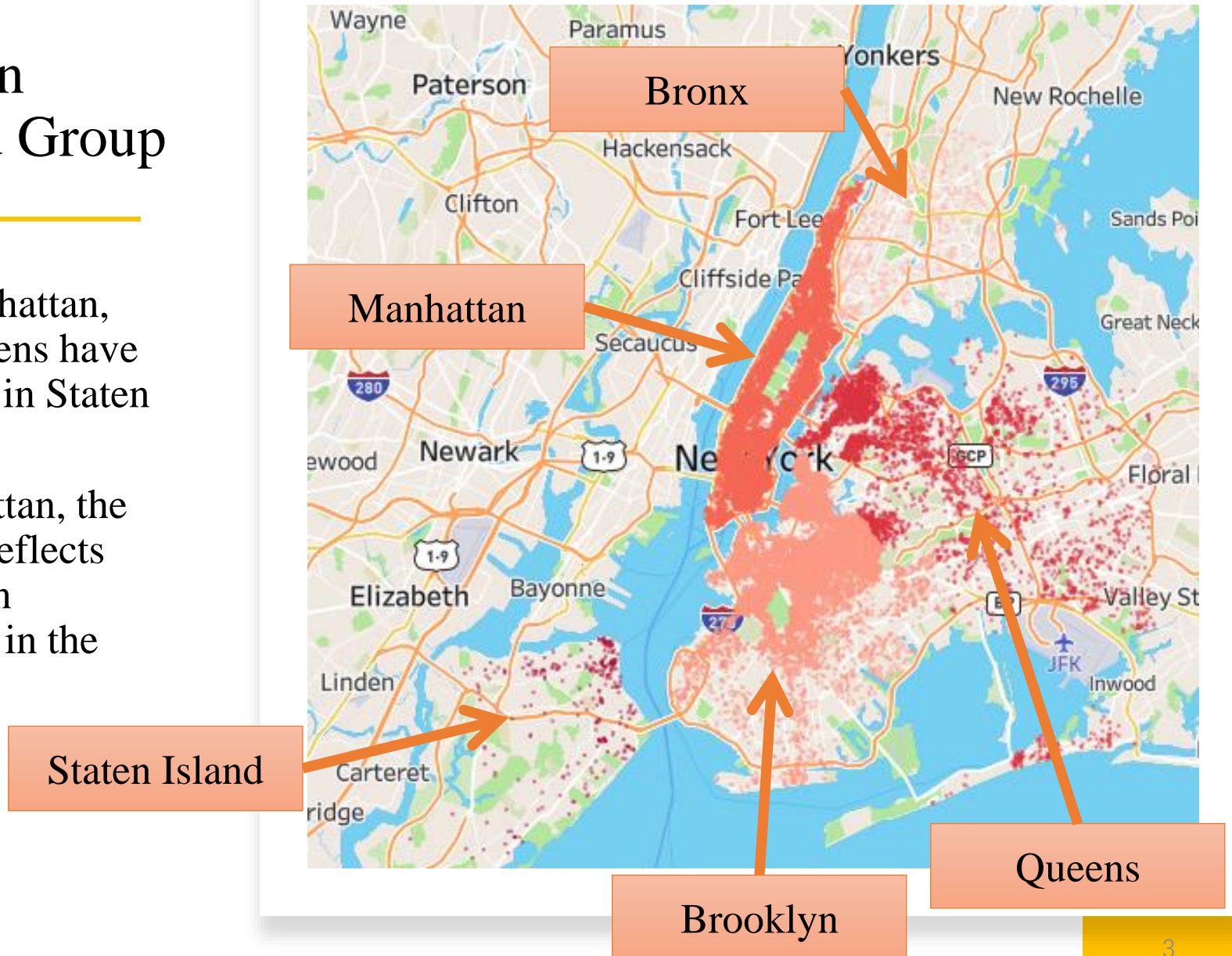
Therefore, we can estimate the total demand (in stays) by dividing reviews by 70%.

Reference: Rusteen, D. (2020, May 11). 5 Tips to 5-Star Airbnb Reviews.
Retrieved January 23, 2021, from

<https://optimizemybnb.com/5-tips-5-star-airbnb-reviews/#:~:text=According%20to%20Airbnb%2C%20about%2070,hurt%20your%20revenue%20for%20months.>

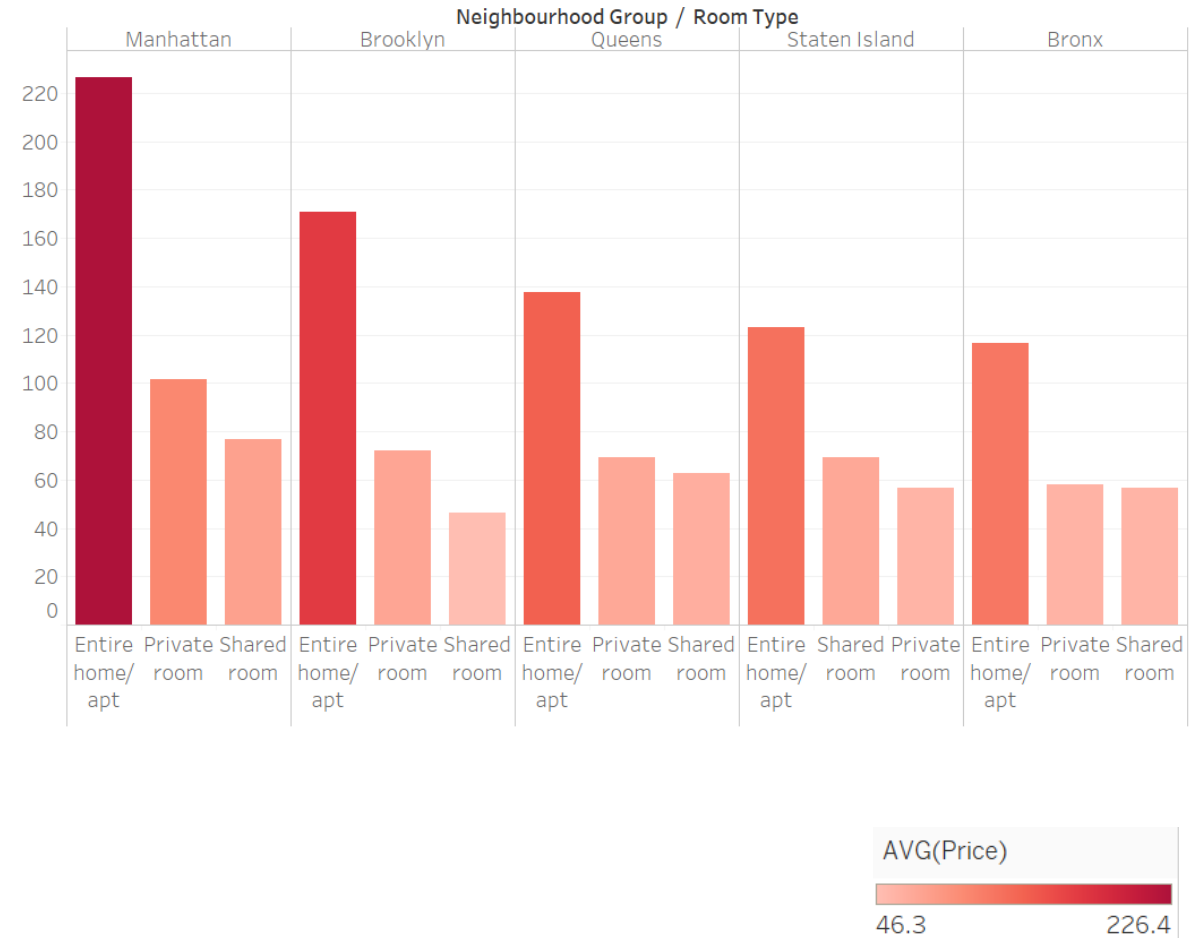
Q2(a). Listing Density in Different Neighborhood Group

- In terms of listing density, Manhattan, North Brooklyn, and West Queens have highest density. Listing density in Staten Island is very low.
- In general, the closer to Manhattan, the higher the listing density is. It reflects that most of the accommodation demands are highly centralized in the city center of New York.



Q2(b). Listings' Prices on Different Neighborhood Groups and Room Types

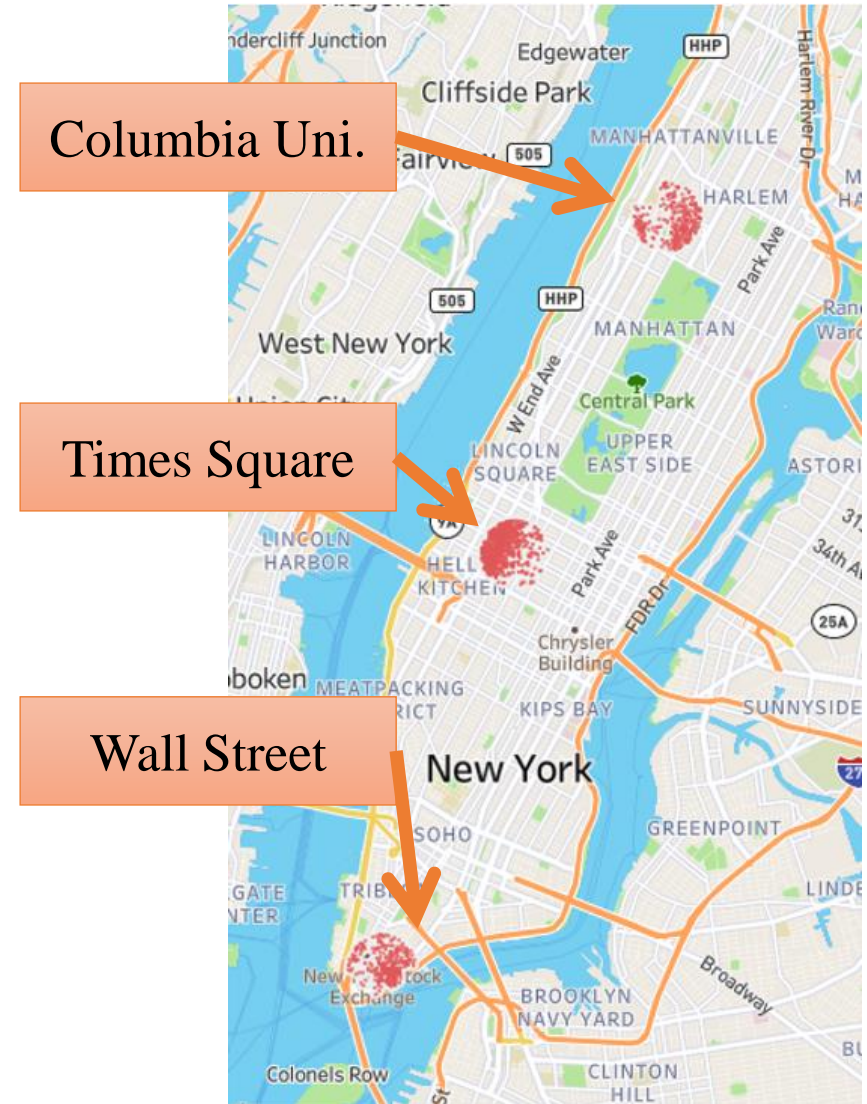
- Manhattan enjoys the highest average listing prices, followed by Brooklyn, Queens, Staten Island, and Bronx. It also reflects that demands are concentrated at downtown areas.
- Entire home/apartment is the most expensive room type, followed by private room, then shared room regardless of region. It indicates that larger and more comfortable listings can charge price premium.



Q3. Comparing Listings Between Columbia University, Times Square, and Wall Street

- Times Square has the highest prices, number of listings, and reviews per listing among three regions.
- Columbia University has higher reviews per listing than Wall Street, but Wall Street has higher average price than Columbia University.

Region	Average Price	Number of Listings	Number of Reviews	Reviews Per Listing
Columbia University	\$116	296	6,518	22.02
Times Square	\$196	736	27,637	37.55
Wall Street	\$192	290	5,513	19.01



Rank by review

Host Id	Host Name
37312959	Maya
344035	Brooklyn& Brea..
35524316	Yasu & Akiko
4734398	Jj
26432133	Danielle
40176101	Brady
51913826	The Bowery Hou..
102524686	Suzy
23591164	Angela
22959695	Gurpreet Singh
6885157	Randy
47621202	Dona
16677326	Alex And Zeena
20828747	Kathlyn
59529529	Han
137814	Waldemar
58391491	Juel
15718962	Samira
21641206	Veronica
7831209	Bobi
9922972	Memo & Iso
2680820	Linda

List per host

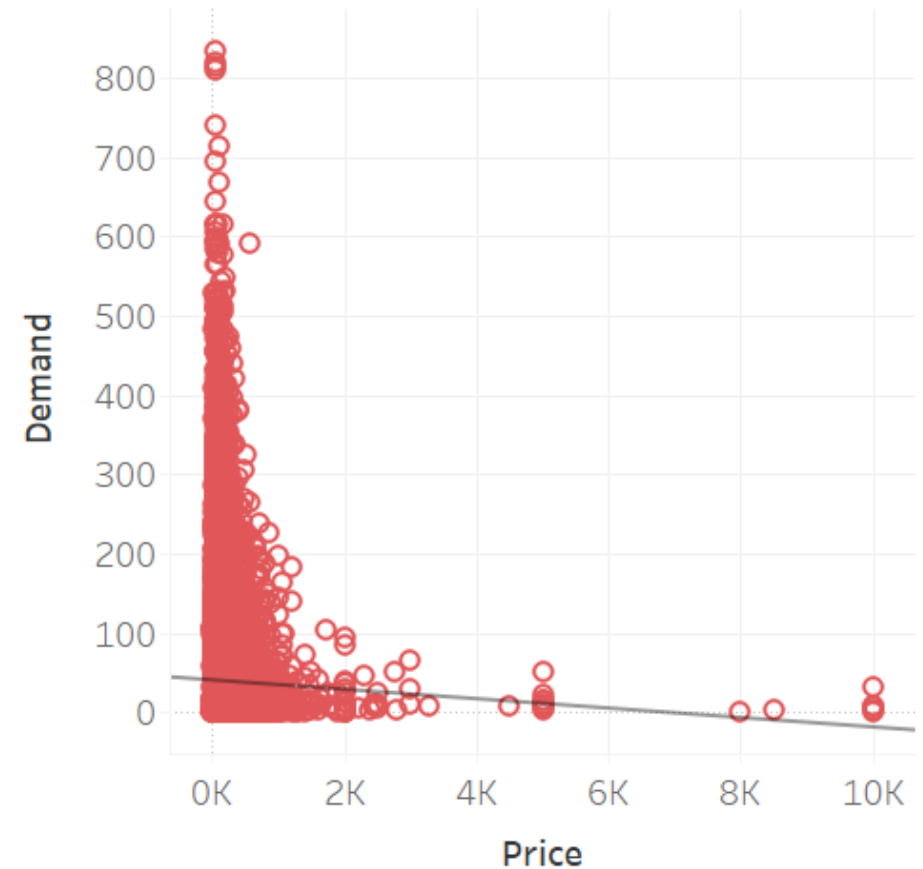
Host Id	Host Name
51501835	Jeniffer
61391963	Corporate Housi..
16098958	Jeremy & Laura
7503643	Vida
30283594	Kara
22541573	Ken
1475015	Mike
190921808	John
61396454	Ash
2119276	Host
119669058	Melissa
120762452	Stanley
25237492	Juliana
417504	The Box House H..
13347167	AFI Apartments
50760546	CRNY Monthly R..
3191545	Kyle
113805886	Yaacov
224414117	Gabriel
159091490	Alec
59917107	Rob
12243051	Sonder

Q4. The “Superhosts” in New York

- Maya's listings have 1,992 reviews, and thus she is the host which most reviews in New York.
- Jennifer has a total of 79 listings, which makes her the top host in terms of the number of listings in New York.

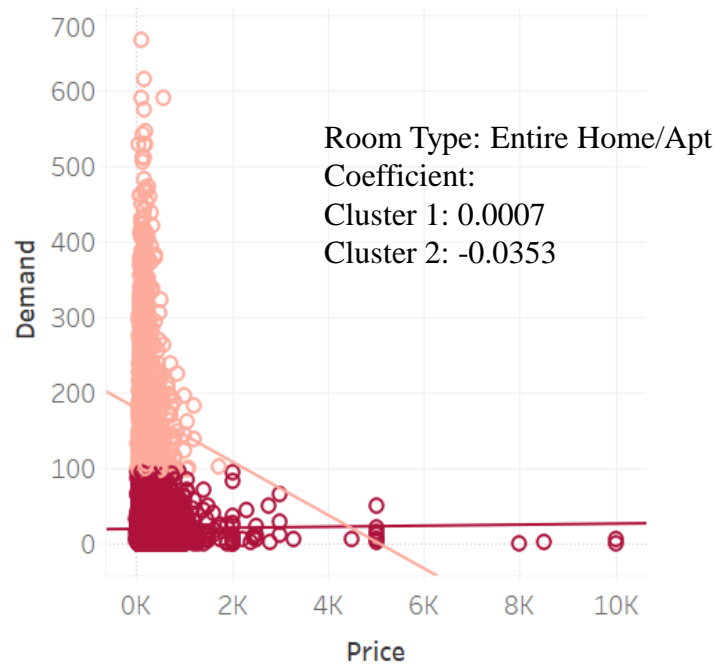
Q5(a). Demand Estimate for Airbnb's Listings in New York

- Using linear regression model, we conclude that the demand of Airbnb accommodations in New York is price inelastic. For every 10% increase in price, the demand drops by 0.59%.
- However, although this model is significant ($P\text{-value} < 0.05$), it is not accurate enough. It can only explain 0.033% of variation in demand.
- We therefore infer that price is not a determinant factor for demand, and we may find a better model with more variables to predict the demand in the future.

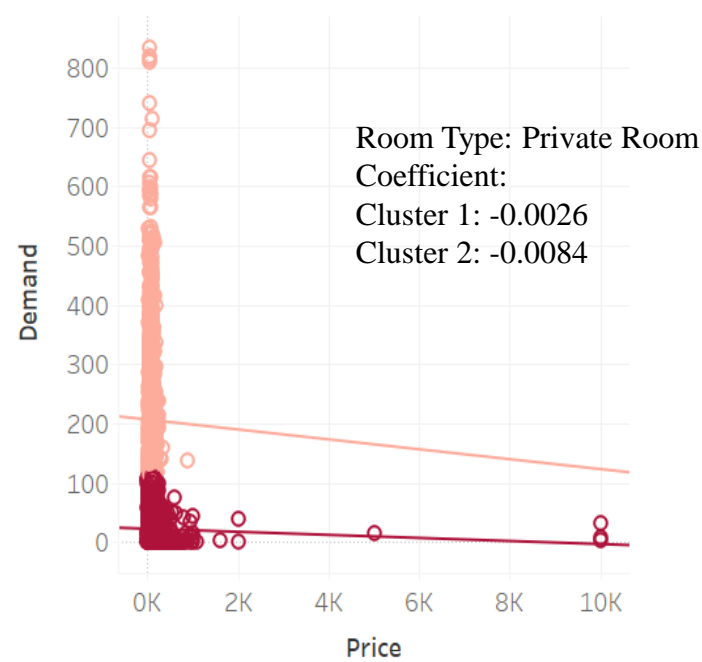


$$\text{Demand} = -0.00592419 * \text{Price} + 40.7369$$

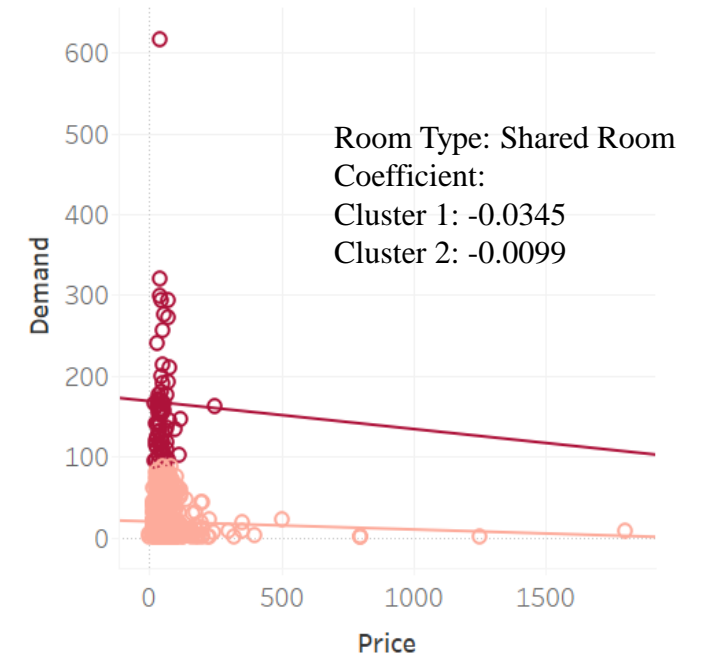
R-Squared: 0.0003274
P-value: 0.0002987



Clusters	Number of Items	Centers	
		Avg. Price	Sum of Demand
Cluster 1	18662	197.95	20.231
Cluster 2	2736	179.91	173.62
Not Clustered	0		



Clusters	Number of Items	Centers	
		Avg. Price	Sum of Demand
Cluster 1	15936	82.05	22.186
Cluster 2	1759	79.784	205.79
Not Clustered	0		



Clusters	Number of Items	Centers	
		Avg. Price	Sum of Demand
Cluster 1	62	51.371	166.64
Cluster 2	782	63.12	18.867
Not Clustered	0		

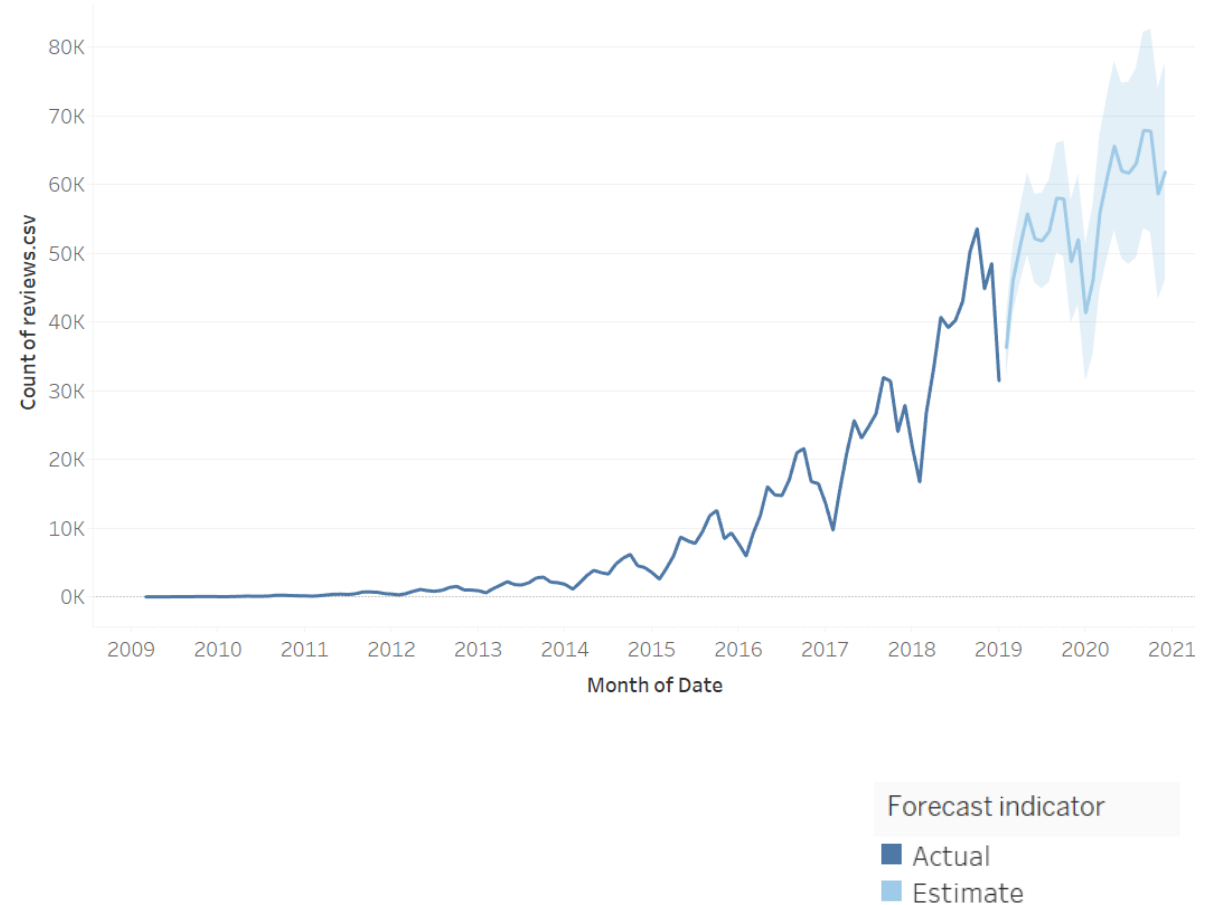
Q5(b). 2-Group Clustering Analysis for Different Room Types

Clusters
■ Cluster 1
■ Cluster 2

- For each room type, there is one cluster which number of listings is significantly smaller than another one, but demand is much higher. Average prices between the two groups are not very different.
- As for coefficient, all the coefficients expect Cluster 2 of Entire Home/Apt are not significantly different from 0. It indicates that in most cases, price is not a determinant factor (same conclusion as Q5(a)).

Q6. Time Series of the Number of Reviews and Predictions

- Based on historical data of reviews, we expect that the number of reviews on Airbnb will be increasing in the coming up 2 years.
- We also observed seasonality in the growth pattern.
- The number of reviews in November 2019 is expected to be 48,775.



Thank you.
See you in the next assignment 😊

