

**SMM635: Data Visualisation  
Final Course Project  
(Group 18)**

AirBnb Vs. Manhattan Hotels

Fung, Ka  
Laubenthal, Christoph  
Vijitratananakit, Pakatorn  
Bonaventura, William  
Low, Ming  
Han, Jiacheng

20 December 2019

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Audience . . . . .	2
1.3	Objective . . . . .	2
<b>2</b>	<b>Literature Review</b>	<b>3</b>
2.1	Cairo's Wheel of Visualization . . . . .	3
2.1.1	Density vs Lightness . . . . .	3
2.1.2	Novelty vs Redundancy . . . . .	3
2.1.3	Functionality vs Decoration . . . . .	4
2.1.4	Multidimensionality vs Unidimensionality . . . . .	4
2.1.5	Originality vs Familiarity . . . . .	4
2.1.6	Abstraction vs Figuration . . . . .	4
<b>3</b>	<b>Visualization</b>	<b>5</b>
3.1	AirBnb's market traction in New York City . . . . .	5
3.1.1	Context . . . . .	5
3.1.2	Description and Insights . . . . .	6
3.1.3	Dimensions of the Visualisation Wheel . . . . .	6
3.2	Importance of Different Attributes . . . . .	8
3.2.1	Context . . . . .	8
3.2.2	Description and Insights . . . . .	9
3.2.3	Dimensions of the Visualization Wheel . . . . .	10
3.3	Revenue by Hosts . . . . .	11
3.3.1	Context . . . . .	11
3.3.2	Description and Insights . . . . .	11
3.4	Dimensions of the Visualization Wheel . . . . .	12
3.5	Airbnb in NYC . . . . .	13
3.5.1	Context . . . . .	13
3.5.2	Description and Insights . . . . .	13
3.6	Dimensions of the Visualization Wheel . . . . .	14
	<b>Bibliography</b>	<b>15</b>

# **Chapter 1**

## **Introduction**

### **1.1 Introduction**

Our team was approached by the Hotel Association to investigate the rising competition from Airbnb and to understand the trends behind the development of Airbnb within the years. Airbnb is a platform that allows the general public to rent accommodation, providing short-term rental housing or room services. It allows travelers to discover and book a variety of unique properties around the world through their websites or mobile phones. It is one of the representatives of the sharing economy in recent years. The website was established in August 2008 and is headquartered in San Francisco, California. Currently, Airbnb has over 3,000,000 listings in 6191 cities in 191 countries.

### **1.2 Audience**

The audience for our project is the Manhattan hotel's industry association, which could utilize the information presented to improve their hotel services based on its competitor, understand the market and to lobby the government for favourable regulations for them. Due to being industry professionals, they would want graphics that are direct in the portrayal of insights. Statistical numbers will also be preferred as to provide concrete evidence behind our claims.

### **1.3 Objective**

The objective of our visualizations in this investigation will be to provide information on how Hotel's competitor i.e. Airbnb gain market traction over time, which areas of Manhattan have been affected the most by Airbnb, how hotels can learn from the development of Airbnb and We would use numerous Python libraries namely Bokeh to create a dashboard consisting of both interactive and static visualizations that would deliver valuable information towards the audience. The design specification of each visualization will be based on Cairo's Wheel of Visualization by Alberto Cairo where he talked about design-trade offs.

# Chapter 2

## Literature Review

### 2.1 Cairo's Wheel of Visualization

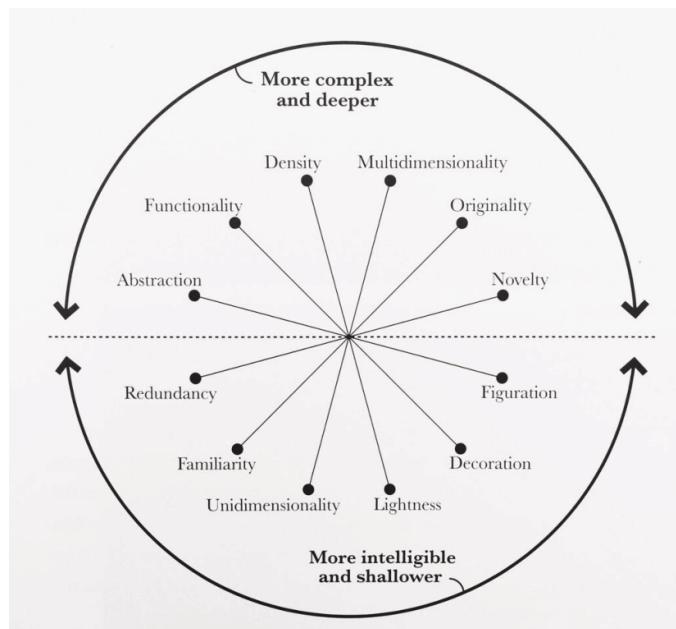


Figure 2.1: Cairo's Wheel of Visualization

Cairo (2013) introduced a tool which provided a basis for design trade-offs in visualizations. During the creation of our visualization we utilized this tool to ensure that the delivery of information is suitable for the audience. The visualization wheel is divided into two regions; the upper region refers to highly complex and deep visualizations in terms of readability while the bottom region refers to easily readable and shallow visualizations. The attributes of the visualization wheel is explained in the following.

#### 2.1.1 Density vs Lightness

This set of feature refers to the amount of space the visualization utilize in comparison to the size of the data and information it shows. The larger the space used by the graphic in the visualization, the denser the visualization will be.

#### 2.1.2 Novelty vs Redundancy

Novelty alludes to when visualizations attempt to describe each information once in the graph while redundancy refers to when a visualization attempts to use various modalities to deliver the same information.

### 2.1.3 Functionality vs Decoration

Functional visualization does not have artistic embellishment and is direct in the presentation of information. It is important that the visualizations do not have decorative features that could inhibit the transfer of information to the readers.

### 2.1.4 Multidimensionality vs Unidimensionality

A multi-dimensional visualization invites the audience to observe and investigate various aspects of the phenomena.

### 2.1.5 Originality vs Familiarity

Original visualizations do not adhere to a large proportion of visualization's form while Familiar visualizations would conform to generally-understood visuals such as bar charts, line charts, distribution functions and pie charts.

### 2.1.6 Abstraction vs Figuration

Figurative visualization will deliver information using pictorial and visual representations of the phenomena. While abstract visualization will be a more conceptual graph and have less representation of pictorials describing the phenomena.

# Chapter 3

## Visualization

### 3.1 AirBnb's market traction in New York City

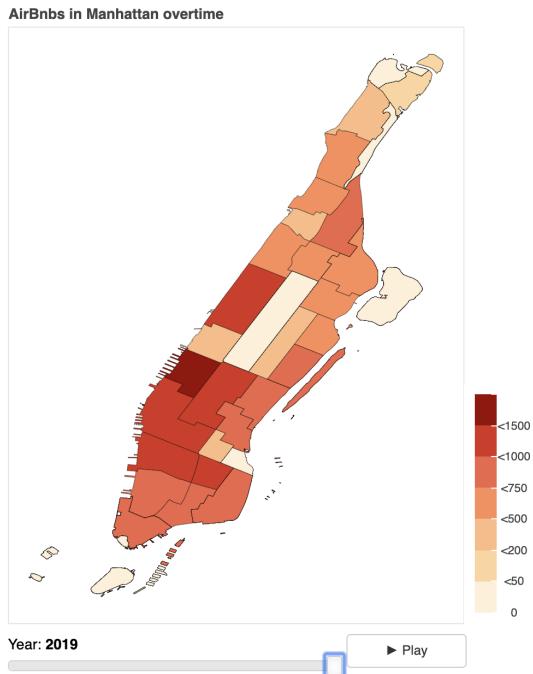


Figure 3.1

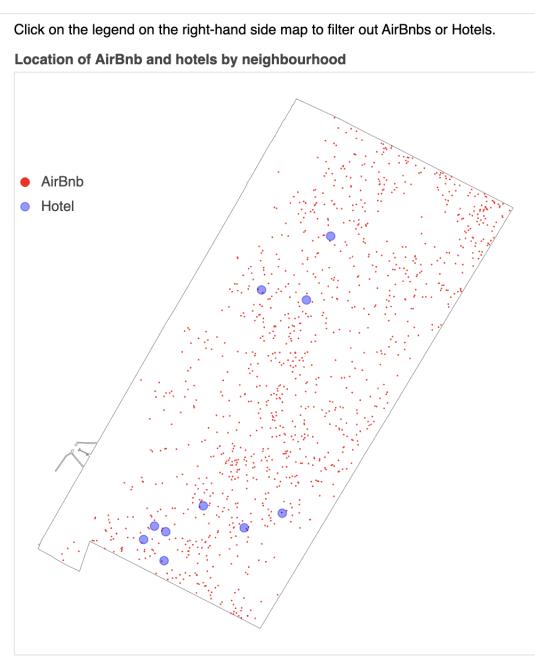


Figure 3.2

#### 3.1.1 Context

The concept of sharing economy in the distribution of goods and services have allowed consumers, through Airbnb, to offer spare bedrooms, houses and hotels for visiting consumers. This creates a huge convenience for consumers to find accommodation to stay and for hosts to create spare cash by leasing out its spare-room/apartment. However the rapid increase in Airbnb listings may cause disturbance to the community with reference to excessive noise, less properties available for long-term rentals which may drive up long-term rental prices for Manhattan locals which is also very high. Despite the massive homelessness crisis in the city, a large number of apartments are basically being listed out as hotel rooms rather than to help alleviate the cause (Sherwood, 2019). However, in 2017 a bill was passed that ruled home-owners to not be able to lease out entire properties for less than 30 days unless the owner is present in the property. This lasted for two years before the judge overruled the bill in 2019.

### 3.1.2 Description and Insights

This visualization gives insight on the market trajectory and traction of Airbnb and Hotels listings in Manhattan from 2009 until 2019. The plot is interactive in which you could zone in to certain neighbourhoods and the time-frame can be adjusted to specific years. This graph emphasizes the rapid growth Airbnb throughout the years where the numbers of Airbnb listings have changed drastically throughout the years. It will also highlight which neighborhood is more affected by the growth of Airbnb compared to others. Listings will appear on the map once its first review has been posted.

Neighbourhoods affected the most by the rise in Airbnb at the end of 2019 was Clinton with greater 1500 listings which was followed by the neighbourhoods of Upper West Side, Midtown-Midtown South, East Village, West Village, Hudson Yards-Chelsea-Flatiron and Union Square . In contrast, neighbourhoods that are least affected by the rise of Airbnb lies within the edges of Manhattan which includes Park Cemetery, Stuyvesant Town and Cooper Town. It is also interesting to see that a strategic location such as Lincoln Square has relatively lower listings than its surrounding neighborhoods. From this it can be inferred that investors planning to create new hotels may want to avoid neighbourhoods have high Airbnb listings in the area. However, high Airbnb listings could also be inferred as desirable areas to visit as it may be deemed strategically located to attractions and landmarks.

Also in most neighbourhoods, large number of Airbnb listings surround relatively very few hotel listings. For example in Clinton for the year 2019, there are only 17 hotels compared to slightly less than 1500 Airbnb listings in the neighbourhood. This could be used by hotel officials to lobby governments for favourable regulations to ensure healthy competition of Airbnb properties vs Hotels in a given area. An over-populated Airbnb listings may endanger the local community and long-term property rental market as owners may want to put the property on short-lease to Airbnb rather than long-term.

To solve this issue, the NY Government may follow London's policy in handling short-term leasing which is to put a restriction of 90 nights a year for short-term lease and a mandatory registration scheme for hosts(Sherwood, 2019). However difficulties persists to enforce such regulation in Manhattan.

One finding was that despite the bill that was passed in 2017(and lasted until 2019) there was still an increase in Airbnb listings which may indicate that private rooms and shared rooms may be more prevalent than entire homes and apartments.

### 3.1.3 Dimensions of the Visualisation Wheel

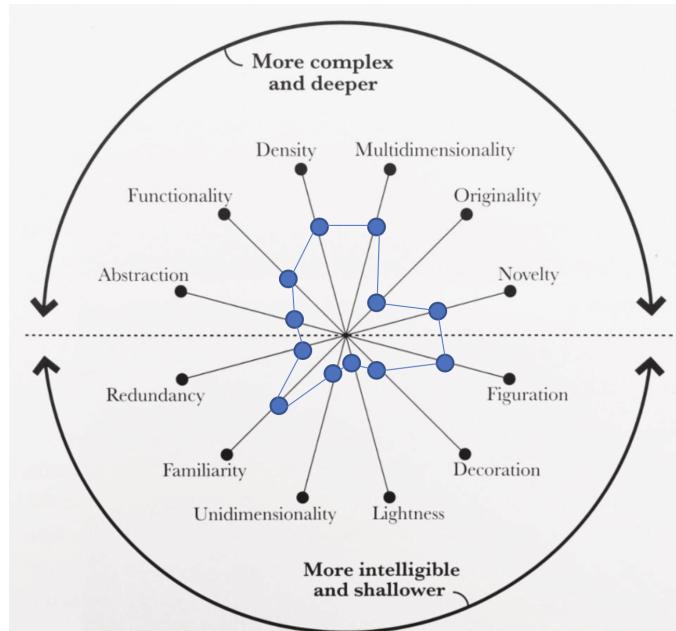


Figure 3.3: Cairo's Wheel of Visualization for figure 3.1 and figure 3.2

### Density vs Lightness

The plot is information dense as we attempt to plot each listings specific location on the map so it is easier to visualize rather than just trying to plot the number of listings per neighbourhood on a bar-chart. We argue that this amount of density is key to deliver as much insight as possible.

### Multidimensionality vs Unidimensionality

This plot is more multidimensional than unidimensional as it explores multiple aspects of the phenomenon and allows the user to free-ly interact with the visualization. Users could explore the time dimension as they could adjust different time periods of the Manhattan map as well as to zoom into certain neighbourhoods to see how Airbnb and hotel listings are scattered within the neighbourhood and the specific hotel name that lies on the neighbourhood. Through this they would understand the trajectory of how Airbnb grew throughout the years and could analyze each neighbourhood specifically.

### Originality vs Familiarity

We created the visualization that quite familiar which is a geo-map visualization so the audience could specifically visualize Manhattan as a whole and the number of listings each neighbourhood has. With a geo-map it would create the audience experience as enjoyable and easy as possible for the audience to read it and understand its insight.

### Abstraction vs Figuration

The visualization is figurative as it plot Airbnb listings specifically on where it is located on an actual Manhattan map based on its longitude and latitude. This helps the audience understand the significance of how populated certain neighbourhood is compared to other neighbourhoods.

### Novelty vs Redundancy

The graph is quite slanted towards novelty as it utilizes one method to deliver the information and does not have redundant features.

### Functionality vs Decoration

The visualization is mainly functional in the delivery of the main insight . But it also has decorative embellishment such as the colour scheme that is key in the delivery of information as it differentiates the range of listings that each neighbourhood has.

## 3.2 How can Hotels learn from Airbnb to recapture competitive advantage? Importance of Different Attributes with Various Airbnb Home-Types Based on User Review

Click on the legends to show/hide attributes.

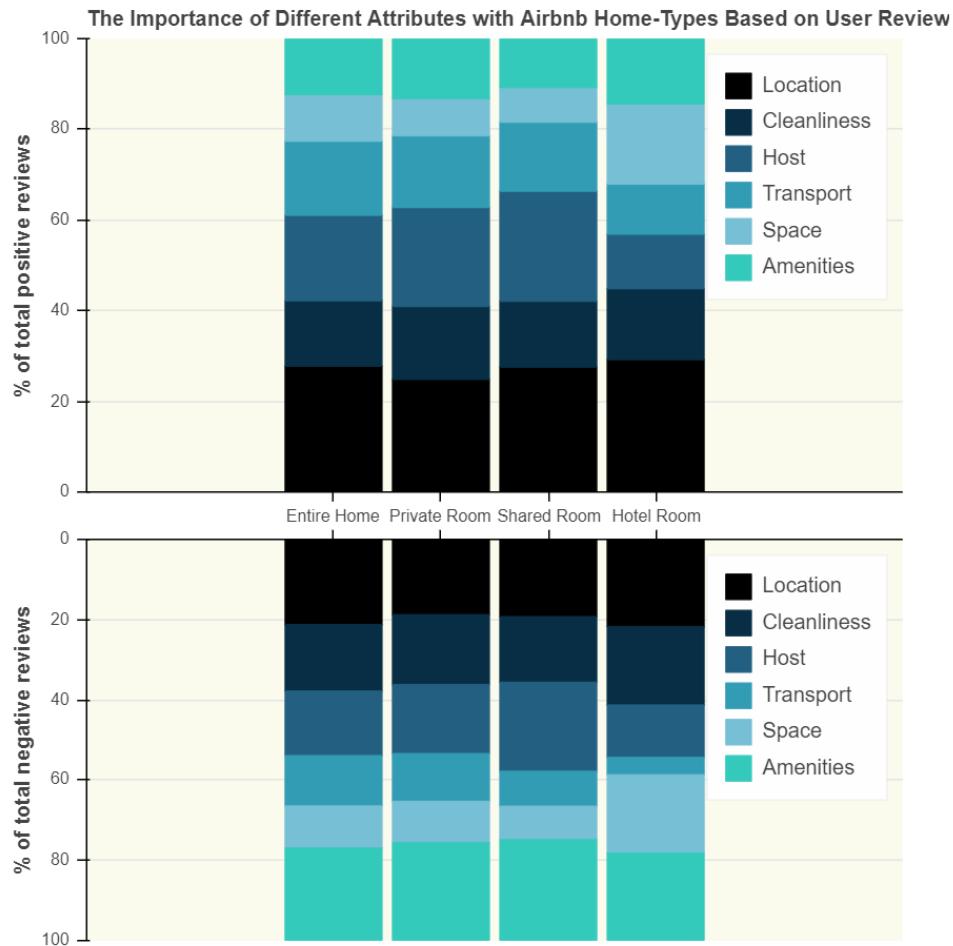


Figure 3.4

### 3.2.1 Context

In 11 years, Airbnb managed to grow from nothing into a \$30 billion company. Airbnb's rapid market development can be used by hotel officials as a benchmark in how to improve its day-to-day hotel operation to be able to effectively compete against Airbnb and other competitors. Currently, Airbnb outperforms hotels in personalization due to its wide-array of homes and locations, enabling genuine micro-segmentation and the "perfect match" between guest and host. Guests are able to select the type of property and price ranges typically vary more for Airbnb than hotels, especially in a busy area such as Manhattan. Airbnb hosts directly communicating with customers may also generate authenticity and genuine experience for the customers.

To diversify its range of products and services, numerous hotel groups have started to participate in home-sharing business model to be able to have a more personalized product and to integrate it to their current business model. This includes Marriott's home sharing platform which is Homes & Villas, Accor's purchase of onefinestay and Hyatt's Oasis. By doing so it could appeal to Airbnb's customer segment which prefers the authentic experience.

### 3.2.2 Description and Insights

This graph investigates the importance of attributes within different Airbnb home types by using Airbnb's review data from 2009 until 2019. We wanted to specifically explore if different attributes would score differently for different property types as each type would have its own customer segments. Entire homes would be more suited for group travellers, family while private rooms would be for private travellers or couples that would have different preferences. The attributes explored in this investigation were Location, Cleanliness, Host, Transport, Space and Amenities.

The reviews were divided into positive and negative reviews using the NLTK Sentiment Analysis Library which gave a compound score of each review posted. The compound score is given by compiling the valence scores of each word in the vocabulary, which is then normalized to be positive 1 (extreme positive) and negative 1 (extreme negative). The threshold for positive sentiments was if the compound score is greater than 0.05 while negative sentiments was if the compound score is less than -0.05 . Positive and Negative reviews were used so attributes' importance can be segmented and compared whether an excellent location is more frequently mentioned than a terrible location. Neutral review (compound scores between 0.05 and -0.05) was excluded as it does not add value to the whole analysis.

Through this graph, hotels could re-capture competitive advantage by understanding certain attributes that Airbnb users pay the most attention to, with the amount of attention being defined by the number of times the certain attribute was mentioned in the review.

On a holistic level, it can be seen that Location represents the most important attribute within positive reviews for all types of Airbnb properties. Location also was represented by words such as "neighbourhood", "area" and "location". Hotels could leverage these points when managing its portfolio homes. With reference to hotels trying to enter the home-sharing business model, they should focus on homes that is strategically located to transportation, landmarks and tourist attractions would be key to attract customer compared to other attributes that they could focus on. A smaller apartment near an underground station in Central Manhattan may prove to be more desirable than a larger apartment that lies slightly more outside of Central Manhattan. It is key to maintaining a good image and customer loyalty. Due to how Airbnb can be easier spread than the conventional hotels, it is apparent that location was one of the key factors in Airbnb's as there are more options for customers to select the perfect Airbnb.

Host also signifies the second most important attribute for positive reviews of Entire Home, Private Room and Shared Room. Thus, it can be summed that good communication and relationship of the host with the customer is well regarded. Airbnb customers value a good relationship with the host as it may feel a sense of authenticity, self and sense of place as good hosts could be original, genuine and sincere.

It can also be seen that Amenities is the most prevalent attribute within negative reviews for all type of Airbnb properties. This highlights how Airbnb consumers are most disturbed and appalled by Airbnb listings that lack in sufficient amenities; this may include lack of Wi-fi, towels, kitchenware, linen and other features that could improve the customer experience. By ensuring a sustained quality through certain amenities to be provided and in good condition will allow hotels build a good brand image and enhance customer loyalty. This an important point that hotels could leverage in the creation of portfolio-homes where sufficient amenities could be key in maintaining loyalty.

### 3.2.3 Dimensions of the Visualization Wheel

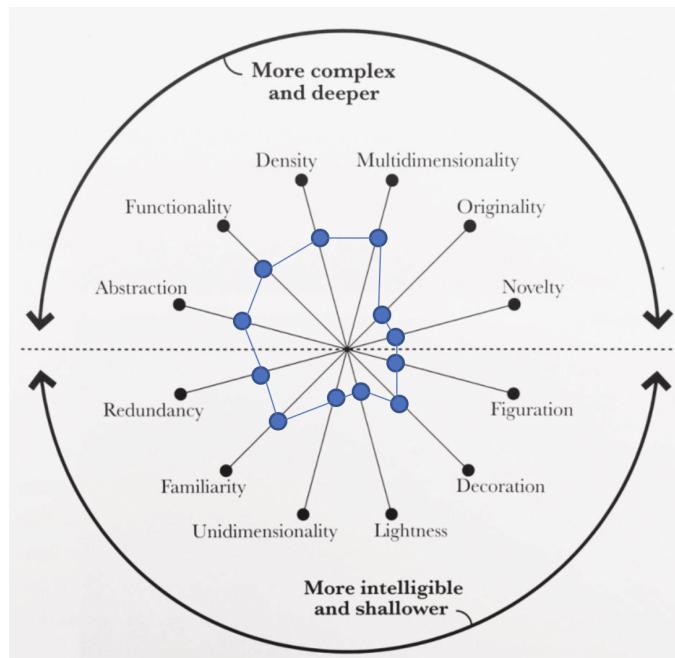


Figure 3.5: Cairo's Wheel of Visualization for figure 3.4

#### Density vs Lightness

The graph is information dense as it includes three variables: Type of Airbnb properties, Different attributes and the percentage the attribute was mentioned within the reviews. We argue that this amount of density is appropriate to deliver as much insight as possible. If, for example, we had only include all reviews instead of segregating it to positive and negative we would leave out an important aspect whether the attribute was tend to be highlighted in a good or bad light.

#### Multidimensionality vs Unidimensionality

This plot is more multidimensional as it explores multiple dimensions due to the high information density that the data has. Allows audience to understand more specifics of the data.

#### Originality vs Familiarity

The graph lies more towards familiarity represents a common form of graph; the stacked bar chart that shows attributes within both positive and negative reviews. However, it does also have slight originality where the negative reviews were plotted underneath so the two reviews can be easily differentiated for readability

#### Abstraction vs Figuration

The graph is more abstract than figurative as it does not include visual representations of the attributes explored in the bar-chart that could add value to the graph. The graph is direct in its attempt to portray the most important attributes among positive and negative reviews of different Airbnb property type. To ensure this we utilized a familiar graph; the stacked bar chart, to ensure readability of the graph.

#### Novelty vs Redundancy

Due to information density, it attempts to include some redundant elements that increases the readability and emphasize the main point of the graph.

### Functionality vs Decoration

The plot is also direct and functional in the delivery of information to ensure readability of the graph to hotel officials. It has minimal embellishments with only reference to the colouring scheme that is important to differentiate various attributes explored in the visualization. It also has functional features such as widgets where when users hover on top of certain attributes, it will allow them to identify the exact percentage of the attribute within its property type. It also has a feature where it allows users to include and exclude attributes from the graph.

## 3.3 Revenue Generated by Hosts of Airbnb listings in Manhattan, NYC

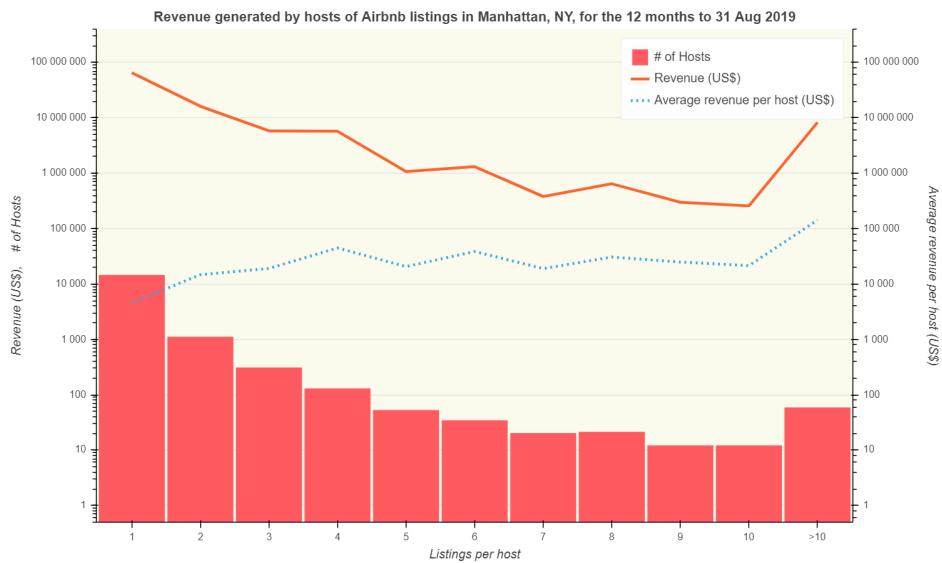


Figure 3.6

### 3.3.1 Context

The appeal of Airbnb lies in its narrative that it enables the everyday homeowner to make a supplemental income through the listing of a spare room or property via the Airbnb platform. However, this has been subject to abuse and controversy, as enterprising hosts have developed entire business models around the listing and letting of Airbnb rooms and homes without being subject to the same regulatory compliance levied upon the hotel industry.

### 3.3.2 Description and Insights

This visualisation aims to provide the user an understanding of how the revenue generated by bookings on the Airbnb platform are distributed amongst hosts with varying numbers of listings attributed to them.

Whilst the majority of revenue generated on Airbnb in the last 12 months accrue to hosts with only 1 listing in Manhattan, there is a small group of hosts with more than 10 listings (0.36% of all hosts in Manhattan) whom an exceedingly disproportionate amount of revenue (7.86%) can be attributed to. Tellingly, a host within this group earns an average of US\$142,680 in income for the past 12 months to 31st August 2019 which is significantly higher than the median per capita income of US\$72,832 (United States Census Bureau, 2018).

Another revealing insight is the weak relationship between the number of properties owned by hosts and the revenue they generate. For example, hosts with five listings (\$20,646) generate on average comparable income to hosts with ten listings (\$21,463). This indicates that occupancy rates are not being fully

utilized which could suggest the possibility of an oversupply of Airbnb listings in Manhattan.

With these insights, hotels may look to differentiate themselves from Airbnb listings by providing a more differentiated offering that competes on attributes beyond price as mentioned in the previous analysis above. Secondly, hotels' lobbyist groups could utilise the above insights regarding the presence of professional landlords to lobby regulatory bodies for stricter legislation to crack down on these large scale short-term rental businesses.

## 3.4 Dimensions of the Visualization Wheel

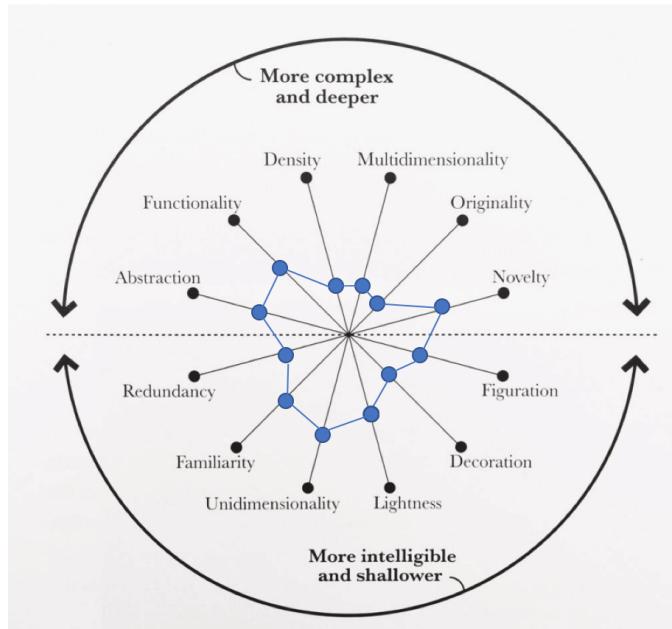


Figure 3.7: Cairo's Wheel of Visualization for figure 3.6

### Density vs Lightness

The graph is more light than it is dense. It provides a high level overview of the revenue earned by hosts on the Airbnb platform and is easy to understand.

### Multidimensionality vs Unidimensionality

The plot is more unidimensional than multidimensional. The main objective explored is to make it visible to the user about how listings are distributed amongst groups of hosts and how revenue is generated by these groups.

### Originality vs Familiarity

The visualisation can be categorised as familiar as it utilises vertical bar charts and line graphs for easy interpretation of the data.

### Abstraction vs Figuration

The graph features more highly in terms of abstraction as compared to figuration as it uses bar charts and line graphs to represent number of hosts and revenue respectively.

### Novelty vs Redundancy

The graph is highly novel as it does not feature any redundancy and only describes each piece of information in one way such as the height of the bar chart and the points of the line graphs to present different features of the graph.

### Functionality vs Decoration

The plot is highly functional as it uses a line graphs to represent the revenue generated and bar charts to represent number of hosts. Embellishments have not been used to decorate the graph. Furthermore, this enhances the ‘familiarity’ attribute as it conforms to the most common visualisation patterns such as bar charts and scatter plots.

## 3.5 Airbnb in New York City – The Story of Lawlessness

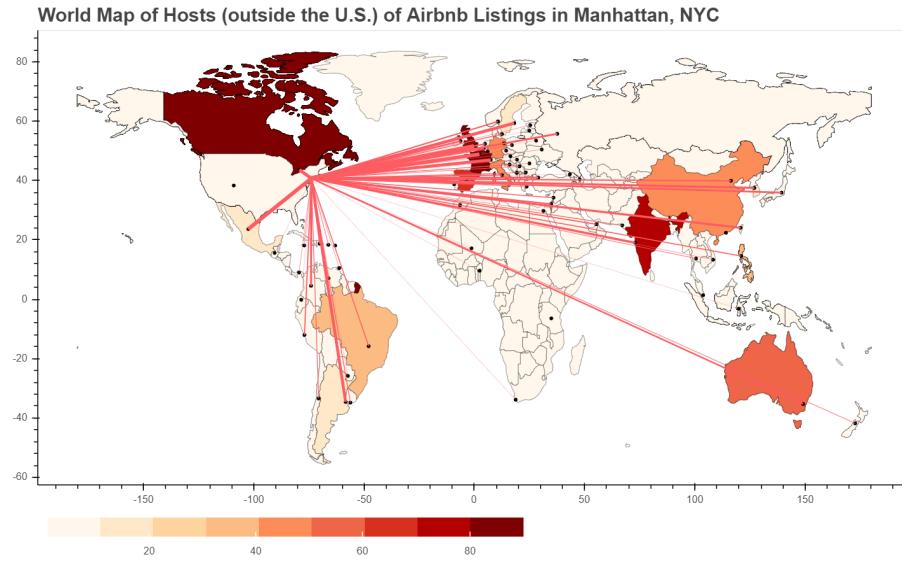


Figure 3.8

### 3.5.1 Context

Airbnb has faced increasing scrutiny in New York City for driving up rental prices due to individuals and companies operating effectively large scale illegal hotels in the area as they looked to increased rental revenue via Airbnb platform.

### 3.5.2 Description and Insights

This graph describes the home location of Airbnb hosts in Manhattan, New York City, that originated from outside the United States. The countries of the world map are colour-coded using a sliding scale based on the number of Airbnb hosts that have listed their address as within the country.

The graph shows that the area with the highest number of Airbnb hosts is Canada as there are over 80 hosts from the country. It was then followed by countries such as the UK and India which have slightly over 60 hosts from each of these countries.

A small but significant number of hosts that have listings in Manhattan appear to originate from overseas with a large number of them from as distant as India. As such, this creates increased competition for housing within the Manhattan housing market as a result of foreign ownership. This serves to add to the debate that Airbnb is not a platform that only serves the everyday homeowner looking to make a supplemental income but also profits rent-seeking, enterprising individuals that have created professional businesses with listings that may span Manhattan and other foreign markets.

## 3.6 Dimensions of the Visualization Wheel

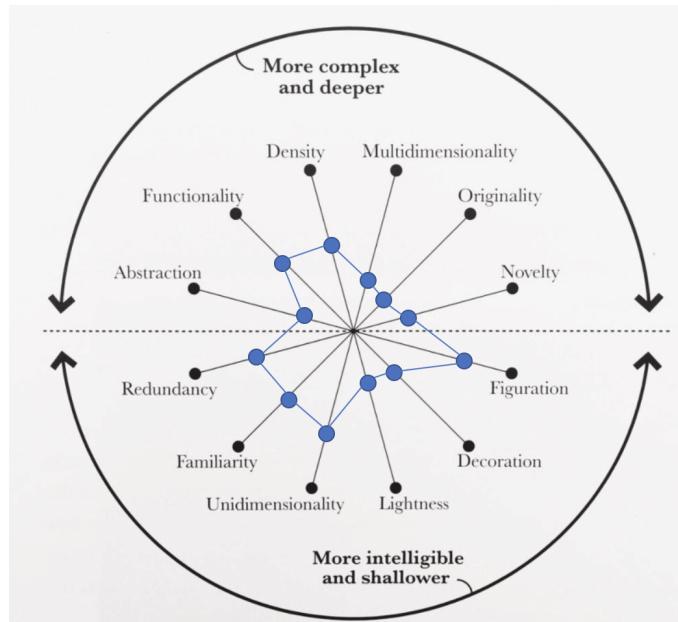


Figure 3.9: Cairo's Wheel of Visualization for figure 3.8

### Density vs Lightness

The graph is more dense than it is light. We argue that this amount of density is appropriate to get as much insight as possible.

### Multidimensionality vs Unidimensionality

The plot is more unidimensional than multidimensional. The main objective explored is to identify the number of hosts from each country and decided to not add more variables into the map as it is quite already information dense.

### Originality vs Familiarity

We created a visualization that is quite familiar which is a geo-map visualization so the audience could specifically see the area of where the hosts are from. With a geo-map it would create the audience experience as enjoyable and easy as possible for the audience to read it and understand its insight.

### Abstraction vs Figuration

The graph is figurative in the delivery of information as it utilizes a visual representation of the world map to explain the phenomenon. It will help the readers understand the extent of the host location through exact locations on the world map.

### Novelty vs Redundancy

The graph have redundant features that are essential to increase the readability of the plot and emphasize the main insights. These include the colour scheme and the dynamic line width that exhibit the same point which is the count of hosts in a country.

### Functionality vs Decoration

The graph is functional as it is direct in the delivery of the information towards the audience. However, it does have certain embellishment such as the colour scheme that is important to emphasize the difference

in the number of hosts there are from each country. Thus, we can argue that it is a necessary redundancy. We want to ensure the readability of the graph to the audience.

# Bibliography

- [1] Cairo, A., 2013., The Functional Art, Pearson Education
- [2] Coles, P.A., Egesdal, M., Ellen, I.G., Li, X. and Sundararajan, A., 2017. Airbnb usage across New York City neighborhoods: geographic patterns and regulatory implications. Forthcoming, Cambridge Handbook on the Law of the Sharing Economy.
- [3] Cox, M. and Slee, T., 2016. How Airbnb's data hid the facts in New York City. Inside Airbnb.
- [4] Dudás, G., Vida, G., Kovalcsik, T. and Boros, L., 2017. A socio-economic analysis of Airbnb in New York City. *Regional Statistics*, 7(1), pp.135-151.
- [5] Jefferson-Jones, J., 2014. Airbnb and the housing segment of the modern sharing economy: Are short-term rental restrictions an unconstitutional taking. *Hastings Const. LQ*, 42, p.557.
- [6] Kaplan, R.A. and Nadler, M.L., 2015. Airbnb: A case study in occupancy regulation and taxation. *U. Chi. L. Rev. Dialogue*, 82, p.103.
- [7] Schneidermann, E.T., 2014. Airbnb in the city. Office of the New York State Attorney Eric T. Schneidermann.
- [8] Sherwood, H., 2019. How Airbnb took over the world. *The Guardian*, 2019/05/05
- [9] Stabrowski, F., 2017. 'People as businesses': Airbnb and urban micro-entrepreneurialism in New York City. *Cambridge Journal of Regions, Economy and Society*, 10(2), pp.327-347.
- [10] Streitfeld, D., 2014. Airbnb listings mostly illegal, New York State contends. *New York Times*, 2017/06/01.
- [11] Wachsmuth, D., Chaney, D., Kerrigan, D., Shillolo, A. and Basalaev-Binder, R., 2018. The high cost of short-term rentals in New York City. School of Urban planning, McGill University: Montreal, QC, Canada.