

Capstone Project

Airbnb Bookings Analysis

By

Siddhartha Pasayat

Pankaj Beldar

Data Science Trainee, AlmaBetter
Bengaluru, India

Introduction



Airbnb, as in “Air Bed and Breakfast,” is a service that allows property owners to rent out their spaces/condos to travelers looking for a place to stay.

Airbnb was started in 2008 by Brian Chesky and Joe Gebbia, based in San Francisco California. The platform is accessible via website and mobile app.

Key Takeaways

1. Peer to Peer (Economy Model)
2. Customization (Experiences)



Problem Statement



Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world.

These millions of listings generate a lot of data - data that can be analyzed and used for security, business decisions, understanding of customers' and providers' (hosts) behavior and performance on the platform, guiding marketing initiatives, implementation of innovative additional services.

Key Takeaways

1. Customer/Hosts Behaviour
2. Marketing Initiatives
3. Additional Services

DataSets



id

unique listing id

neighbourhood_group

group of area

name

represents
accommodation

neighbourhood

area under neighbourhood group

host_id

unique id for hosts

latitude

location of listing

longitude

location of listing

host_name

registered name
for hosts

minimum_nights

minimum nights stay
required for single visit

price

price of listing

number_of_reviews

total rating count of listings

last_reviews

last review given

room_type

3 unique rooms

calculated_host_listing_counts

total number of listings registered under hosts

Key Takeaways

1. 48895 x 16
2. Null values:
 - a) name = 16
 - b) host_name = 21
 - c) last_reviews = 10052
 - d) reviews_per_month = 10052

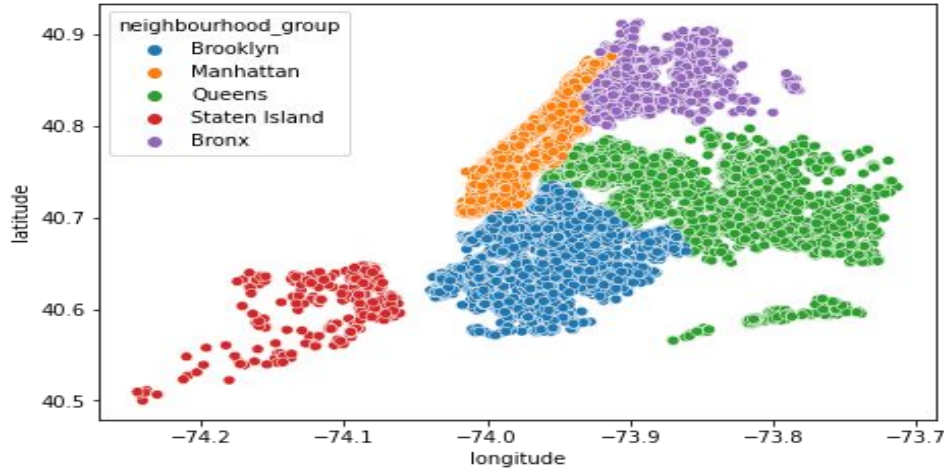
reviews_per_month

ratings received per month

availability_365

Number of days for which
host is available in a year

New York City Map



Key Takeaways

1. Location of various neighbourhood
2. Demographic distribution
3. Area wise presence

Objectives

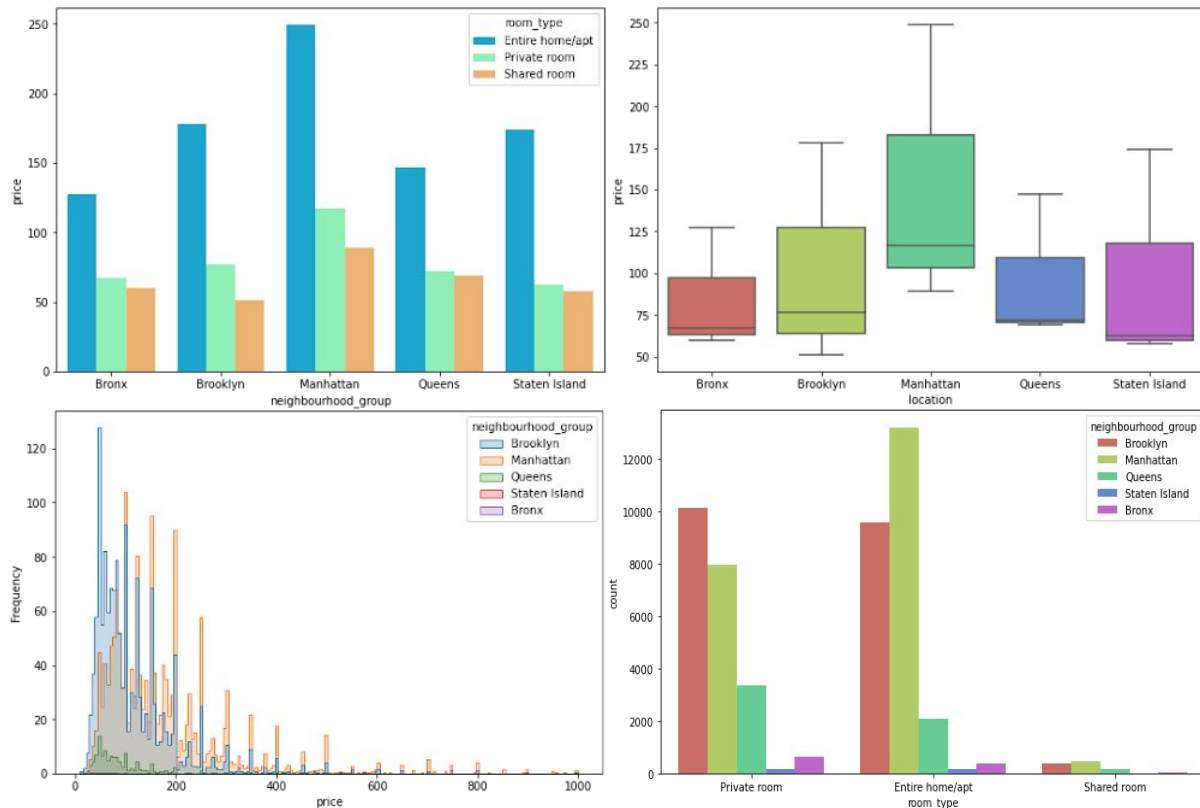


1. Which is the preferred location according to average best price?
2. Where are most of the hosts located?
3. The highest and lowest rent paying locations by customers
4. Most Popular/demanded host based on reviews and availability 365 days
5. Establishing relation between neighbourhood group and availability of rooms.
6. Which are the top hosts, neighbourhoods, neighbourhood groups based on their turnover?
7. Room type selection based on price, availability on 365 days.
8. Top ten neighbourhood based on listing price.
9. Distribution of properties based on Mandatory stays.
10. Type of Visit based on Mandatory stay allowed for single booking.

Key Takeaways

1. Customer/Hosts Behaviour
2. Marketing Initiatives
3. Additional Services

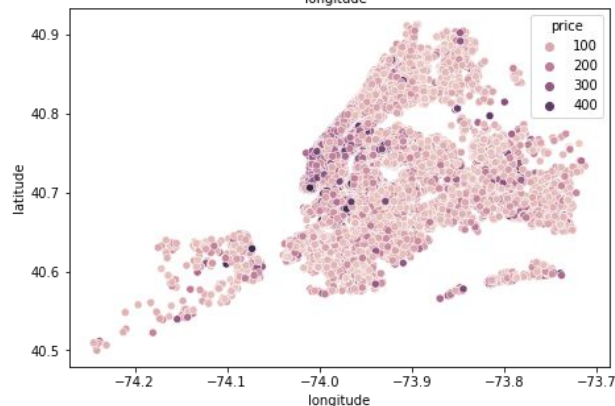
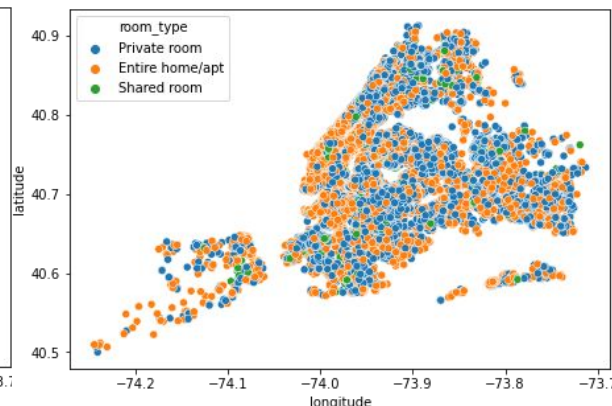
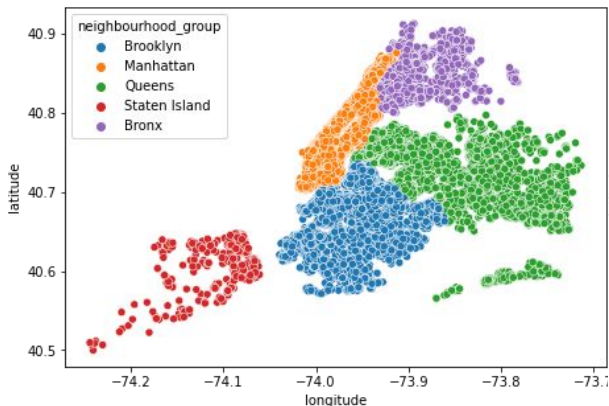
Which is the preferred location according to average best price?



Key Takeaways

1. Manhattan is the preferred location based on Pricing and Frequency of Customer Visit
2. Manhattan is preferred in all type of rooms.
3. Manhattan prices are higher as compared to other place.
4. Entire Home/Appt. are higher in Manhattan.
5. Private rooms are higher in Brooklyn.
6. Shared rooms are higher in Manhattan.

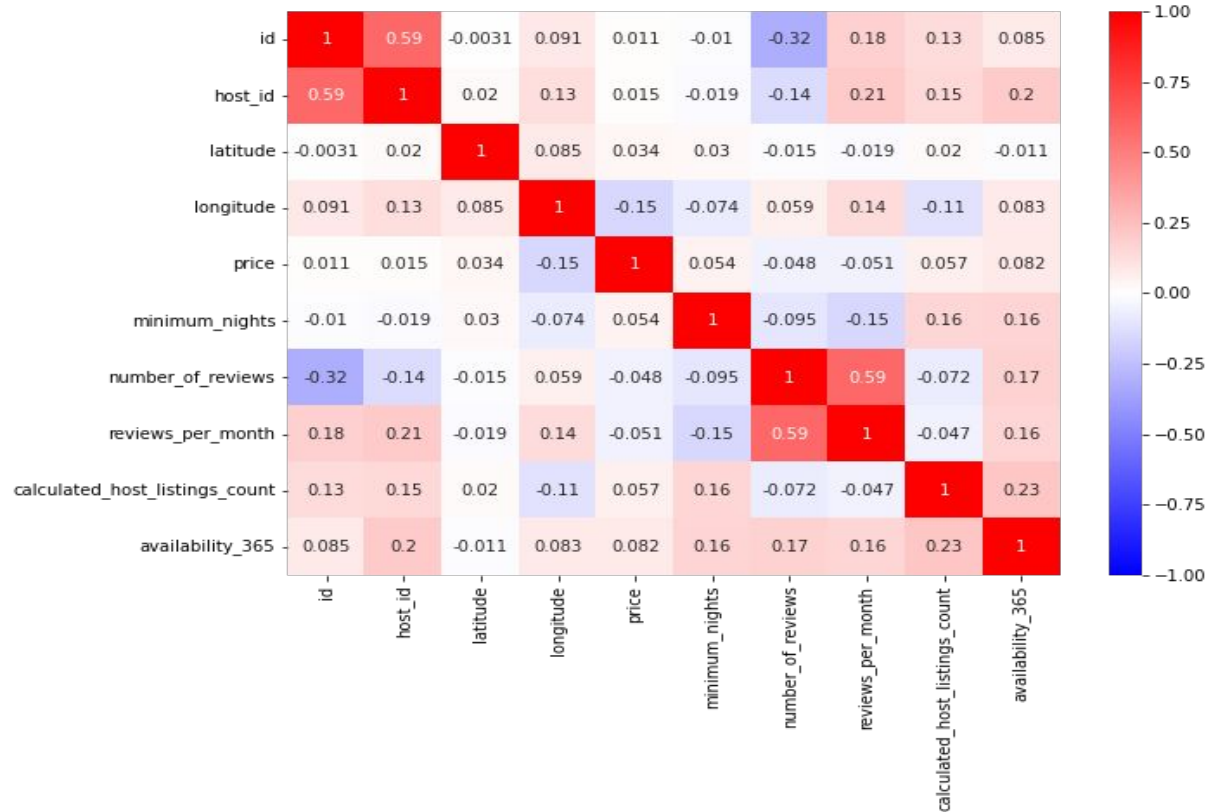
Why Manhattan?



Key Takeaways

1. Southern of Manhattan and Northern of Brooklyn mostly has expensive properties.
2. Manhattan has the maximum number of high price region.
3. Private rooms and Apartments are mostly preferred by customers over shared.
4. Manhattan has highest Private /Entire Apt rooms in combination followed by Brooklyn.

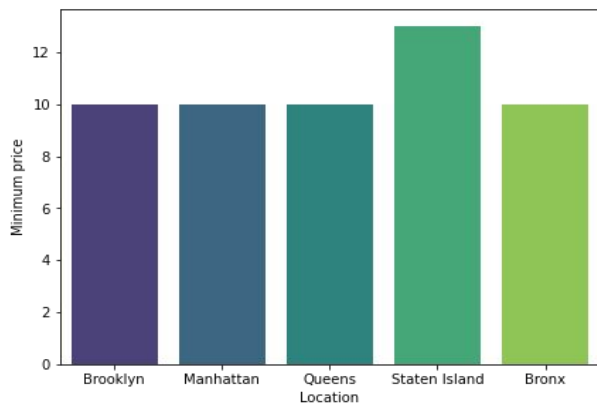
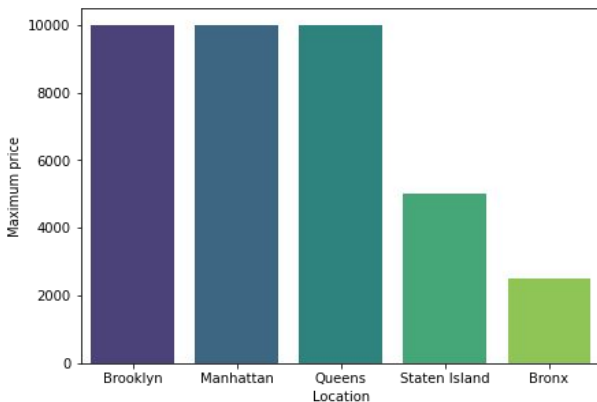
Correlation Matrix



Key Takeaways

1. High correlation number means high correlation between two variables.
2. Number_of_reviews and reviews_per_month has correlation of 0.59(high).
3. Host_id and minimum_nights has correlation of -0.019(low)

Highest and Lowest property prices and locations



	Location	Maximum price	Minimum price
0	Brooklyn	10000	10
1	Manhattan	10000	10
2	Queens	10000	10
3	Staten Island	5000	13
4	Bronx	2500	10

Key Takeaways

1. Customers pay highest tariff of 10,000 USD in Brooklyn, Manhattan and Queens.
2. Customers pay lowest tariff of 10 USD in Brooklyn, Manhattan, Queens and Bronx.

Most popular Hosts based on reviews and availability



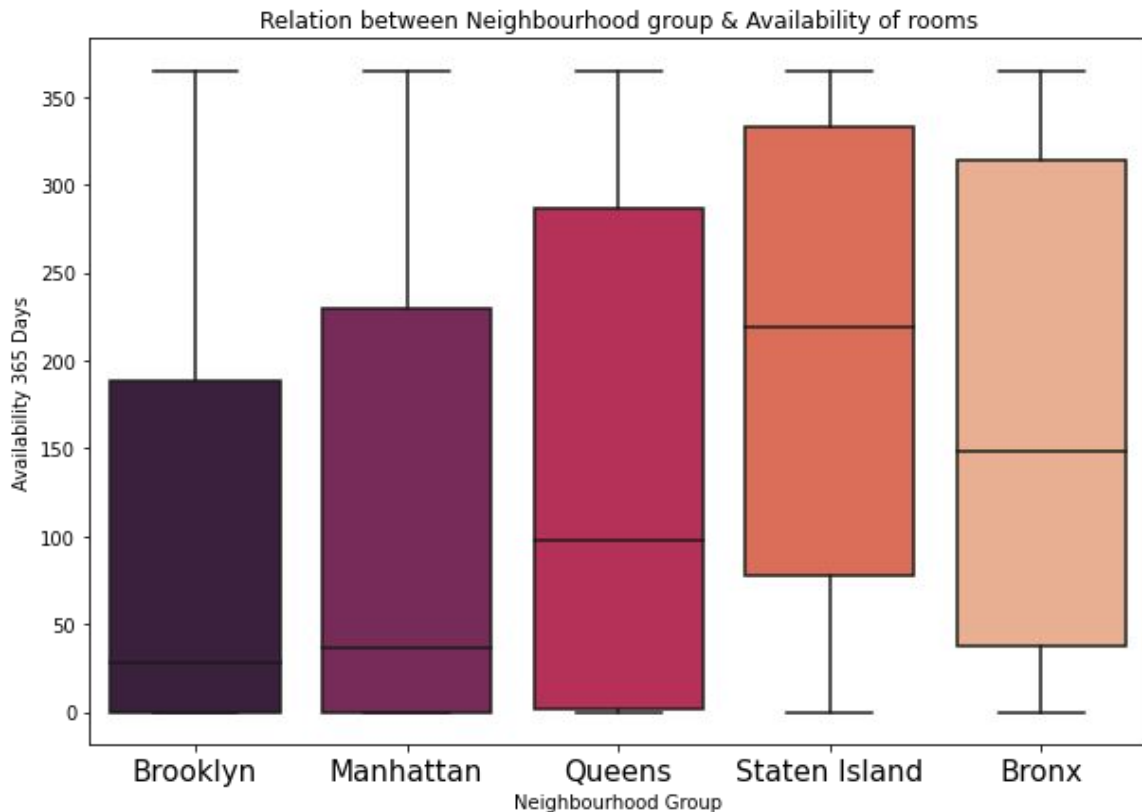
host_name	calculated_host_listings_count	number_of_reviews
Sonder (NYC)	327	1281
Blueground	232	29
Kara	121	65
Kazuya	103	87
Jeremy & Laura	96	138

Key Takeaways

Most popular hosts are :

1. **Sonder,**
2. **Blueground,**
3. **Kara,**
4. **Kazuya,**
5. **Jeremy & Laura**

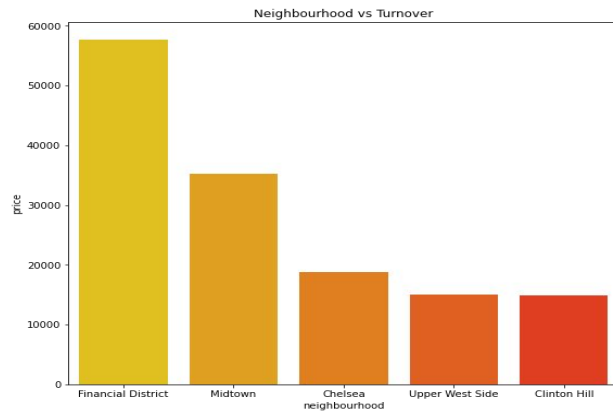
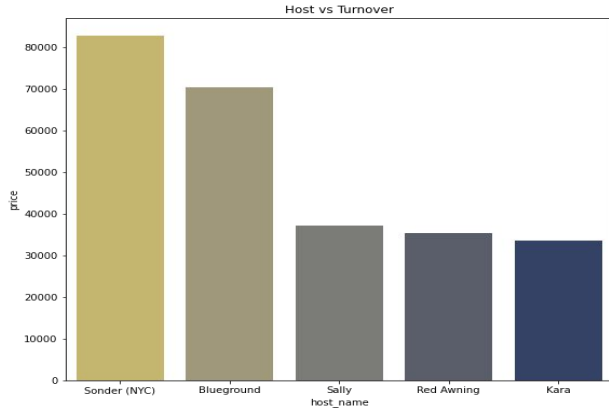
Relation between Neighbourhood group and availability of rooms



Key Takeaways

1. **Staten Island** has the highest availability of rooms over 365 days followed by Bronx.
2. **Brooklyn and Manhattan** seems to be very much occupied during the season.

Top Hosts, Neighbourhood based on their turnover



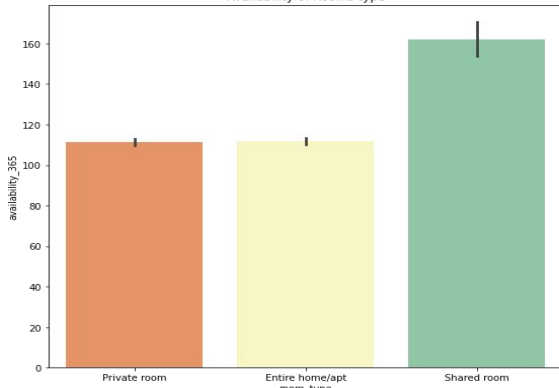
Key Takeaways

1. Top 5 hosts based on Turnover are:
 - a) Sonder
 - b) Blueground
 - c) Sally
 - d) Red Awning
 - e) Kara
2. Top 5 neighbourhood based on Turnover are :
 - a) Financial District
 - b) Midtown
 - c) Chelsea
 - d) Upper West Side
 - e) Clinton Hill

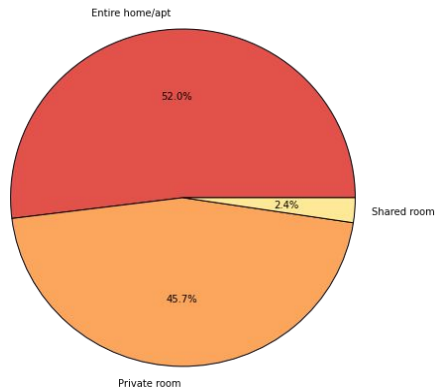
Room type selection based on price



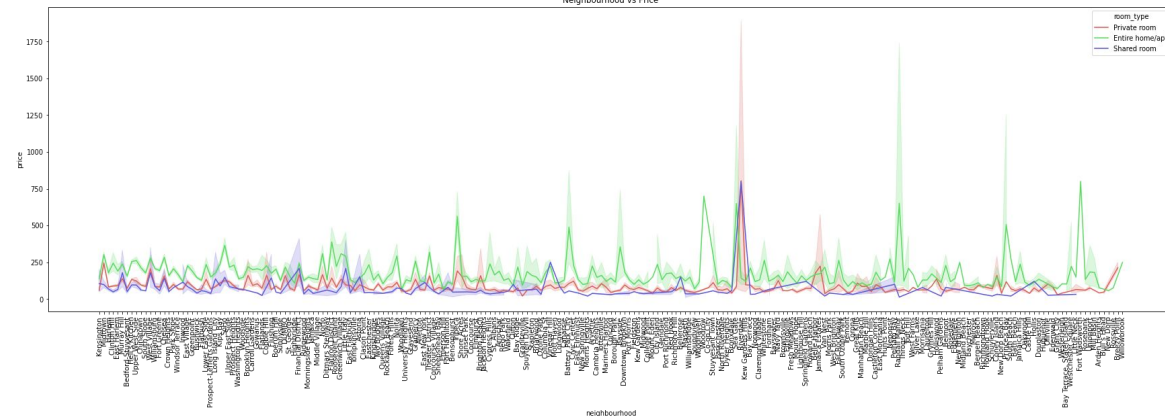
Availability of Rooms type



Proportion of Room Types



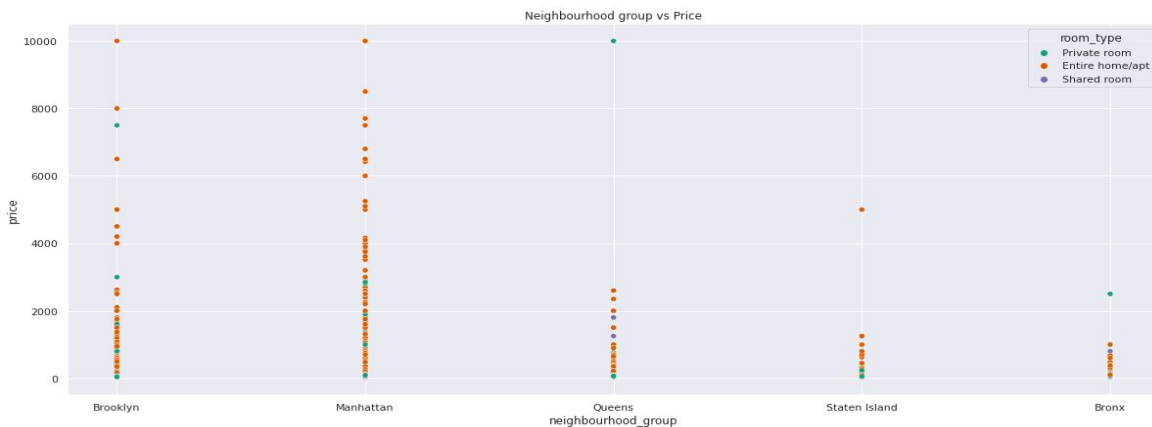
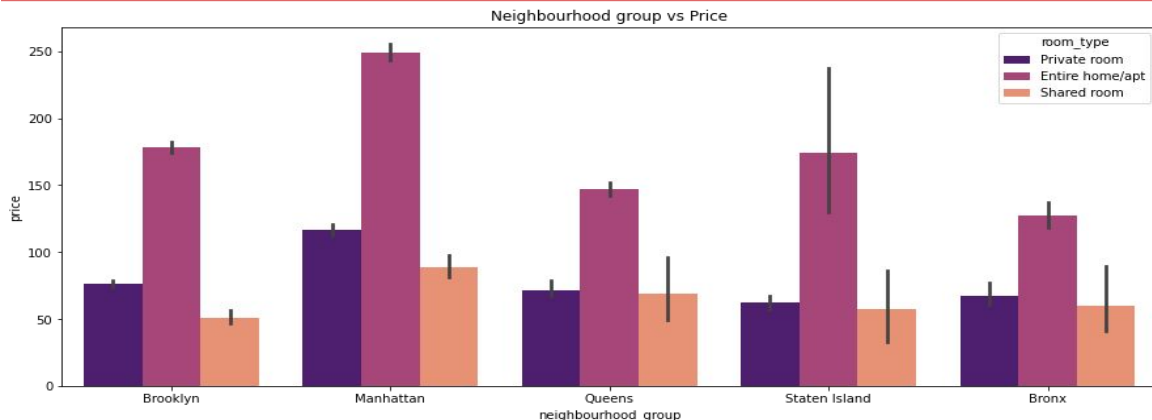
Neighbourhood vs Price



Key Takeaways

1. Shared rooms are available for most of the times as compared to private and entire home/ apartment
2. Room type distribution :
 - a) Entire Home/Apt - 52%
 - b) Private Room - 45.7%
 - c) Shared Room - 2.4%
3. Price of Entire Home/apt is usually higher as compared to Private rooms and Shared ones.

Room type selection based on price [Contd.]



Key Takeaways

1. Manhattan and Brooklyn are posh areas with high end properties available.
2. The high end properties are mostly Entire Home/Apt.

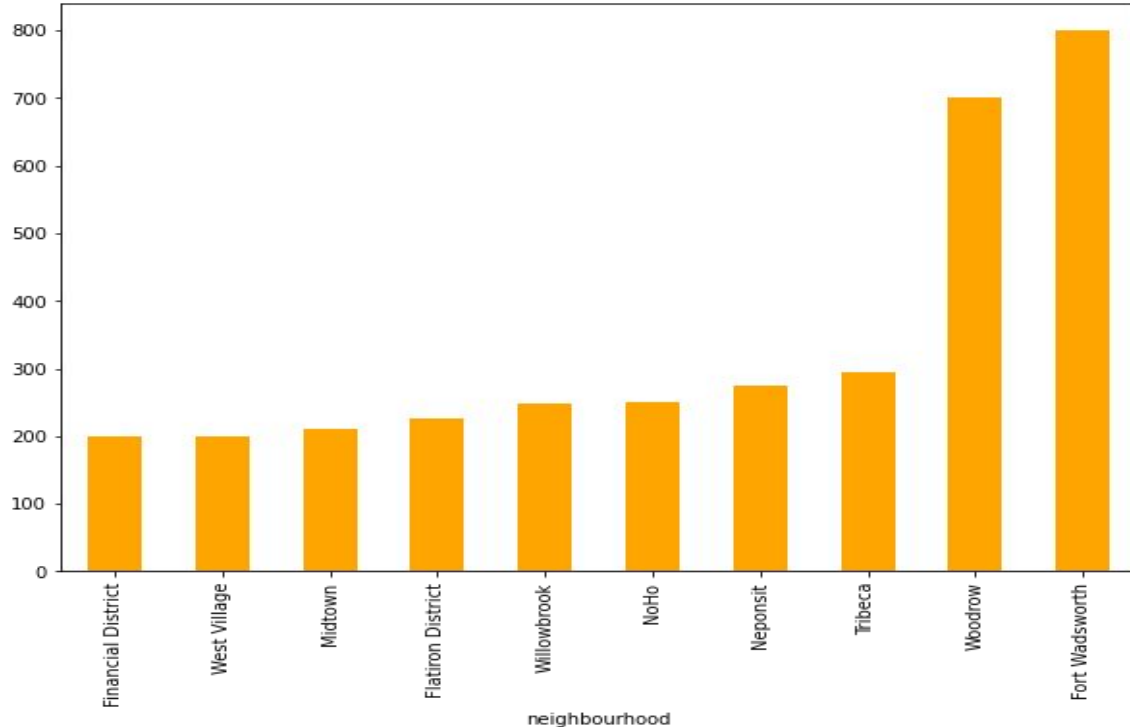
Top Neighbourhood based on listing price.



Key Takeaways

1. Fort Wadsworth and Woodrow are the top neighbourhood based on listing price by a huge margin than their contemporaries.
2. They belong to Staten Island neighbourhood

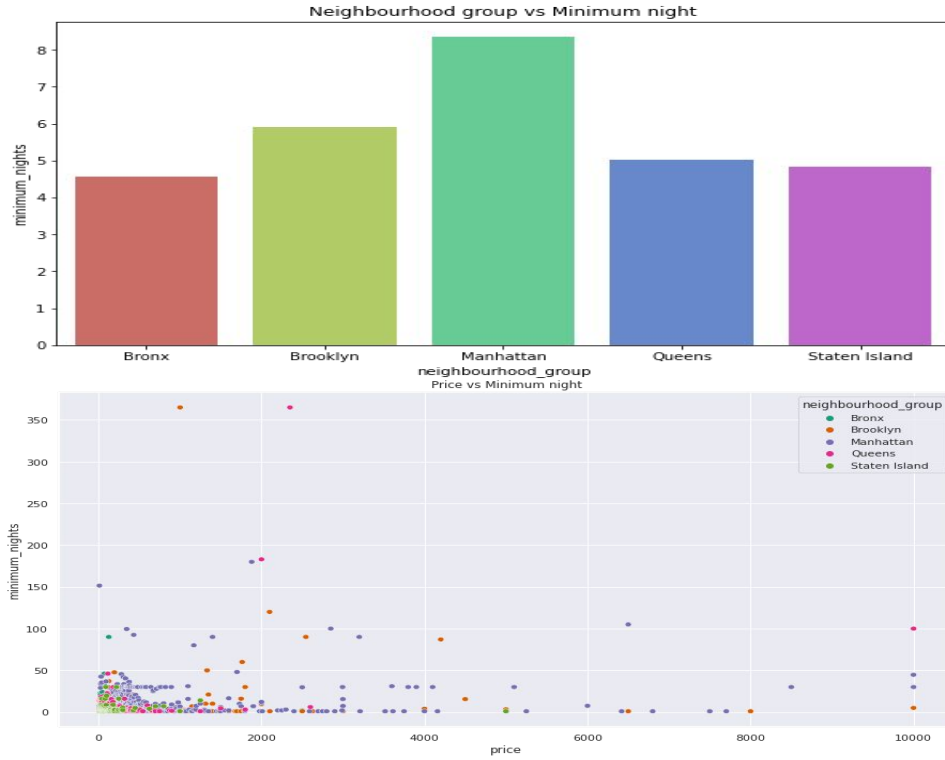
Top Ten Neighborhoods Median Listing Price by Neighborhood



Distribution of neighbourhoods based on Mandatory stay allowed for a single booking

Key Takeaways

1. Most hosts allow Mandatory stay for less than 5 days.
2. Manhattan has higher requirement for Mandatory stay followed by Brooklyn and Queens.
3. Customers prefer to stay in places where mandatory stay is minimum and budget friendly.
4. Manhattan has a highest range of offerings both in terms of higher mandatory stay and expensive listed properties.

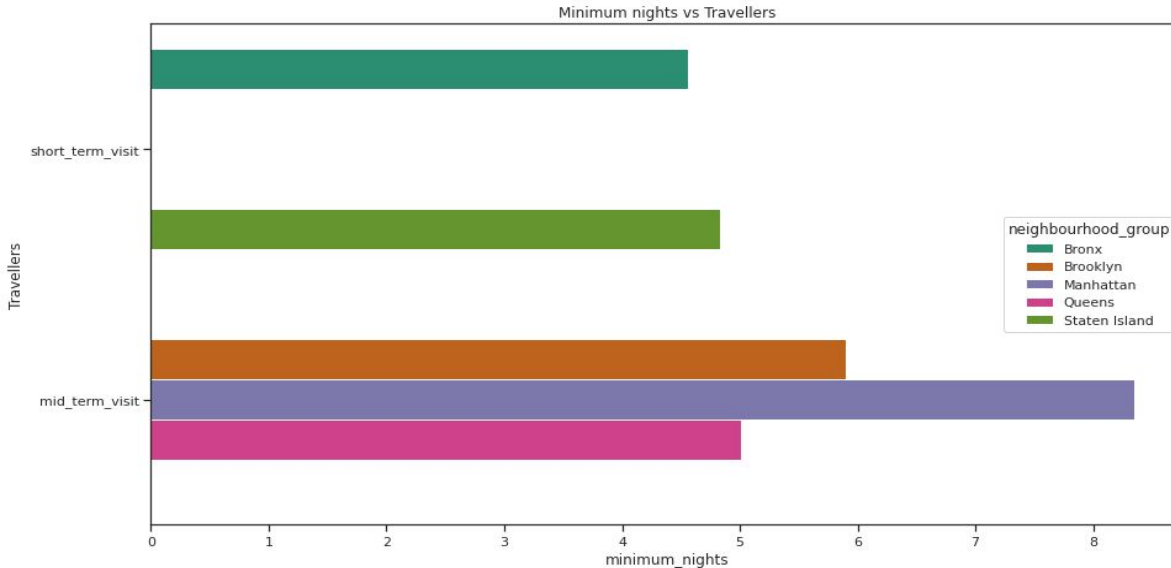


Types of Visit based on Mandatory stay allowed for a single booking



Key Takeaways

1. On the basis of hosts allowing minimum mandatory stay Manhattan, Queens and Brooklyn hosts prefer customers having a minimum 'Mid-term visit' whereas hosts in Bronx and Staten Island prefer customers having a minimum 'Short-term visit'.
2. Bronx and Staten Island can be preferred for shorter stays over other neighbourhoods making it budget friendly to some extent.
3. Manhattan and Brooklyn being posh areas and the implementation of higher mandatory stays for single booking will make these trips/visits expensive.



Short Term Visit - Upto 5 Days

Mid Term Visit - Upto 90 Days

Long Term Visit - More than 90 Days

Limitations & Scope of Improvement



1. Datasets have limiting attributes to classify various categories of properties.
2. Customer experiential and Category wise ratings for Hosts seemed to be missing which could have played an important role in identifying Star Hosts.
3. A lot of guest information were missing like Purpose of Visit, Number of Guests, which could have given a sense of understanding about the relation of customer footfall and neighbourhoods.
4. Key attributes of properties like Number of Beds, Closets, Bathrooms, Gym, Sauna, Property Age, Distances from nearest Hospitals, Shopping Complexes, Airport, Station were missing.

Key Takeaways

1. Key attributes related to Property
2. Key attributes related to customer experience with Hosts
3. Key attributes related to Guests

Conclusion



1. Manhattan and Brooklyn are the posh areas in NY as there is maximum footfall and properties based on prices and listings are on the higher side.
2. Manhattan and Brooklyn have the highest number of hosts.
3. Manhattan has highest number of Private rooms and Entire House/Apt. in culmination followed by Brooklyn.
4. Highest accommodations of 10,000 USD are available in Manhattan, Brooklyn and Queens.
5. Most popular hosts are Sonder, Blueground, Kara to name a few based on number of reviews and calculated host listing counts.
6. Staten Island seems more available for booking throughout the year compared to other neighbourhoods.

Key Takeaways

1. Manhattan and Brooklyn are the posh areas of NYC having high end homes and maximum number of hosts renting out spaces.
2. Staten Island seems more available for booking throughout the year.

Conclusion [Contd.]



1. Sonder, Blueground, Sally are some of the top hosts based on their turnover.
2. Financial District, Midtown, Chelsea are some of the top neighbourhoods based on their turnover.
3. Shared rooms are mostly available over other room types and Entire Home / Apt which has the highest proportion of room share are mostly on the expensive ends.
4. Fort Wadsworth and Woodrow are expensive neighbourhoods based on median listed price belonging to Staten Island.
5. Most hosts allow a minimum 5 nights mandatory stay for single booking but the average increases in case of Manhattan, Brooklyn and Queens.
6. Bronx and Staten Island are mostly preferred for shorter visits and onwards and others are for slightly longer stays.

Key Takeaways

1. Sonder and Blueground seem to be the most popular hosts in terms of reviews received, properties listed and turnover generated.
2. Financial District which is the buzz of Manhattan has high turnover from properties listed.
3. Staten Island though has fewer properties but the median listing price is more in its neighbourhood. Fort Wadsworth and Woodrow.
4. Most hosts allow less than 5 days mandatory stay. The average is higher in Manhattan, Brooklyn and Queens.

Marketing Initiative



1. Manhattan being the star neighbourhood we can roll down a lot of festive offers encouraging longer stays during Christmas and coming up with loyalty cards for frequent visitors which can also work for Brooklyn and Queens.
2. The star hosts should be incentivised to encourage them to maintain the properties and services as per the company standards.
3. Staten Island and Bronx can have discounts encouraging students to stay for a longer period. It can also have frequent check in cards for people from low income groups visiting regularly by offering good incentives and coupons.
4. Some local tours can be clubbed during longer visits encouraging customers to stay longer and prefer their stays.

Key Takeaways

1. Festive Offers and Loyalty Cards
2. Discounts and Frequent Check In Cards
3. Incentivising Hosts
4. Local tours and experiences

Thank You



Travel Like a Human

EDA Capstone Project



maBetter