New York Airbnb Analysis

DETAILED PROJECT REPORT

Shubham Tembhurne

PROJECT DETAIL

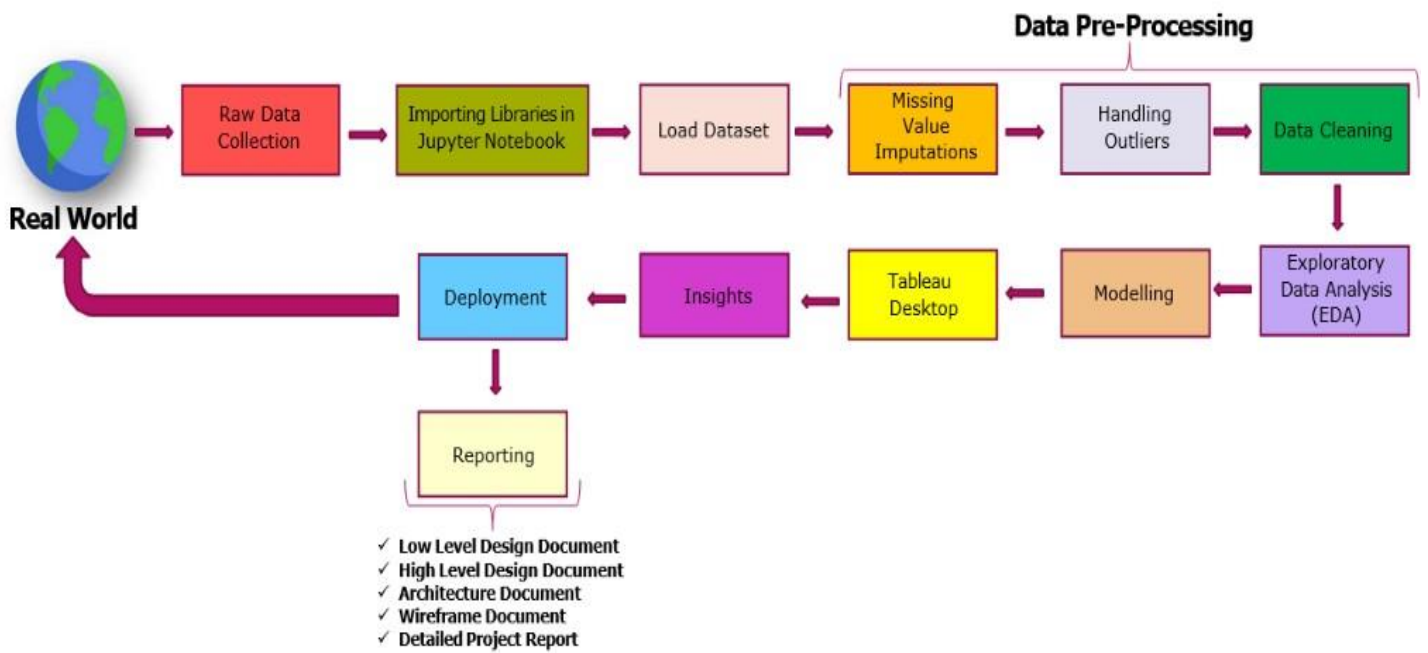| Project Title	US Pollution Analysis
Technology	**Business Intelligence**
Domain	**Travel and Hospitality Tech**
Project Difficulty level	**Advance**
Programming Language Used	**Python**
Tools Used	**Jupyter Notebook, MS-Excel, Tableau**

OBJECTIVE

- Airbnb, Inc. is San Francisco, California based company which act as a broker and provides an online marketplace for short-term homestays. The company charges a commission from each booking. Airbnb providing it premium service and experience to customers since 2008. Today Airbnb has millions of listings. These listings generate loss of data. Analyzing this data become crucial factor for the company. This data can be used for business decisions, marketing, implementations of initiatives, additional services and much more.

PROBLEM STATEMENT

- Given dataset include all information about host, listed properties, geographical location, prices reviews and all other required metrics. Analyze the given dataset makes different predictions and draw meaningful conclusion in order to grow the business. Also state what can we learn from different predictions.

Architecture

Data Pre-Processing

Real World → Raw Data Collection → Importing Libraries in Jupyter Notebook → Load Dataset → Missing Value Imputations → Handling Outliers → Data Cleaning

Data Cleaning → Exploratory Data Analysis (EDA) → Modelling → Tableau Desktop → Insights → Deployment → Real World

Deployment → Reporting

- ✓ Low Level Design Document
- ✓ High Level Design Document
- ✓ Architecture Document
- ✓ Wireframe Document
- ✓ Detailed Project Report

DATASET INFORMATION

This dataset has round 48895 listings and 16 Columns. It is mix between categorical and numeric values. Given dataset contains null values as blanks well which we have to consider while doing analysis. Last_review and reviews_per_month has more null values. There are 5 neighborhood group in which all listings located. Nearly 80-85% of listings located in Manhattan and Brooklyn. In Manhattan booking price is bit higher as compared to other neighbourhood groups. There are 3 kinds of room type (i.e. Shared Room, Private Room, Entire home/Apt).

Columns Present in dataset are:

Id: Listing ID of the property
Name: Name of the listed property.
host_id: ID of the property owner.
host_name: Name of the property owner.
neighbourhood_group: Location at which property located.
Neighbourhood: Area in which property located.
Latitude: Latitude coordinate.
Longitude: Longitude coordinate.
room_type: Type of the room (Entire Home/ Appt, Private Room, Shared Room)
Price: Price in Dollars.
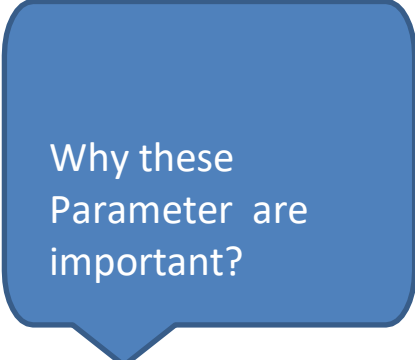**Minimum_nights:** Amounts of minimum night stay at property
number_of_reviews: No. Of reviews
last_review: last review on which date.
reviews_per_month: Numbers of reviews per months.
calculated_host_listings_count: Count of properties listed on that host.
availability_365: Number of days when listing is available for booking

Why these Parameter are important?

1.Location Metadata

Latitude, Longitude: Help uniquely identify the monitoring location.

Id, Name, host_id, host_name, neighbourhood_group, Neighbourhood: Provide human-readable location information for spatial analysis, policy targeting, and regional trends.

Importance: These are crucial for geospatial analysis and identifying pollution hotspots.

2.Date

last_review: Specifies the date of observation.

Importance: Enables temporal analysis to detect trends, seasonal patterns, or year-over-year changes in reviews, booking, number of customers.

INSIGHTS

1. In which Neighborhood group there is maximum number of properties listed?

Number of Property by Neighbourhood Group

Neighbourhood	Number of Property
Manhattan	2,77,073
Brooklyn	45,925
Queens	23,005
Bronx	2,437
Staten Island	865

- Manhattan has the maximum number of properties listed followed by Brooklyn over 2,77,073 by Manhattan and 45,925 by Brooklyn.

2. Which host has maximum number of properties listed?

Host Name	Price	Availability	Reviews	Listings
Sonder (NYC)	$82,795	98,588	1,281	106,929
Blueground	$70,331	58,884	29	53,824
Kara	$36,723	39,273	324	14,679
Kazuya	$4,514	20,237	87	10,609
Sonder	$20,451	26,555	43	9,216
Jeremy and Lau..	$20,060	28,063	138	9,216
Corporate Hou..	$13,308	22,015	417	8,281
Ken	$24,258	32,244	881	7,645
Pranjal	$18,865	21,972	1	4,225
Mike	$29,355	30,498	4,039	2,910
John	$41,892	40,249	7,223	2,792
Vida	$7,908	15,512	282	2,705
Stanley	$10,839	20,517	219	2,542
Ruchi	$12,199	17,380	12	2,402
Red Awning	$35,294	10,796	127	2,401
Stat	$2,070	12,471	52	1,849
Melissa	$16,583	27,049	2,795	1,594
Host	$7,440	10,598	335	1,521
Nina	$11,579	27,313	1,072	1,426
Hiroki	$1,330	456	22	1,369
Juliana	$7,366	10,904	470	1,181
Yaacov	$7,180	10,565	123	1,089
Rated	$7,273	11,353	3	1,089
Anting	$5,529	8,705	32	1,089
Michael	$66,895	38,888	11,081	1,043
Eyal	$2,388	10,283	233	1,028
Jeniffer	$4,313	9,750	144	965
CRNY Monthly ..	$5,676	4,588	72	961
Gabriel	$13,026	11,987	780	946
Max	$14,606	15,338	1,506	917
David	$65,844	44,171	8,103	907
Anthony	$18,065	18,968	3,877	858
AFI Apartments	$3,537	8,437	100	841
The Box House ..	$6,962	1,813	935	784
Emily	$24,130	12,991	1,887	782
Adam	$23,404	13,147	2,380	683
Zach	$5,486	8,924	597	667
Sergii	$4,596	28,852	129	651
Kyle	$10,941	8,842	1,352	591
Yuval	$1,706	6,185	366	538
Benjamin	$12,645	10,534	1,688	502
Alex	$52,563	30,031	6,204	475
Erin	$22,306	9,287	1,390	470
Jessica	$50,697	21,520	4,074	453
Laura	$24,683	15,059	2,992	434

- Sonder NYC Has the maximum number of properties listed followed by blue ground. which Have 106,929 listings highest amongst all.

3. Which host has maximum properties listed in neighborhood groups having maximum properties listed?

Host Name	Bronx	Brooklyn	Manhatt..	Queens	Staten Isl..
Sonder (NYC)			1,06,929		
Blueground		464	53,360		
Kara		8	14,669	2	
Kazuya		618	1,854	8,137	
Sonder			9,216		
Jeremy and Lau..			9,216		
Corporate Hou..			8,281		
Ken		108	7,500	37	
Pranjal			4,225		
Mike	2	66	2,824	18	
John	3	370	2,340	74	5
Vida		2,705			
Stanley		35	2,506	1	
Ruchi			2,402		
Red Awning			2,401		
Stat		1,505		344	
Melissa	1	1,523	51	19	
Host			1,521		
Nina		1,062	249	66	49
Hiroki	37	74	185	1,073	
Juliana		10	1,170	1	
Yaacov			1,089		
Rated			1,089		
Anting		759	297	33	
Michael	5	225	756	53	4
Eyal		1	1,027		
Jeniffer			961	4	
CRNY Monthly ..		31	930		
Gabriel	1	26	915	4	
Max	1	49	830	37	
David	4	184	558	155	6
Anthony	1	608	244	4	1
AFI Apartments			841		
The Box House ..		784			
Emily	1	75	689	17	
Adam	1	55	617	10	
Zach		645	20	2	
Sergii		521	94	36	
Kyle		32	557	2	
Yuval		5	4	529	
Benjamin	1	475	23	3	
Alex	4	264	182	16	9
Erin		42	426	2	
Jessica	1	237	154	57	4
Laura	1	99	318	16	

- The host which has the maximum properties listed is Sonder (NYC) in Manhattan followed by blue ground.

4. What is the average price in different properties listed?

Price Difference in Different Properties
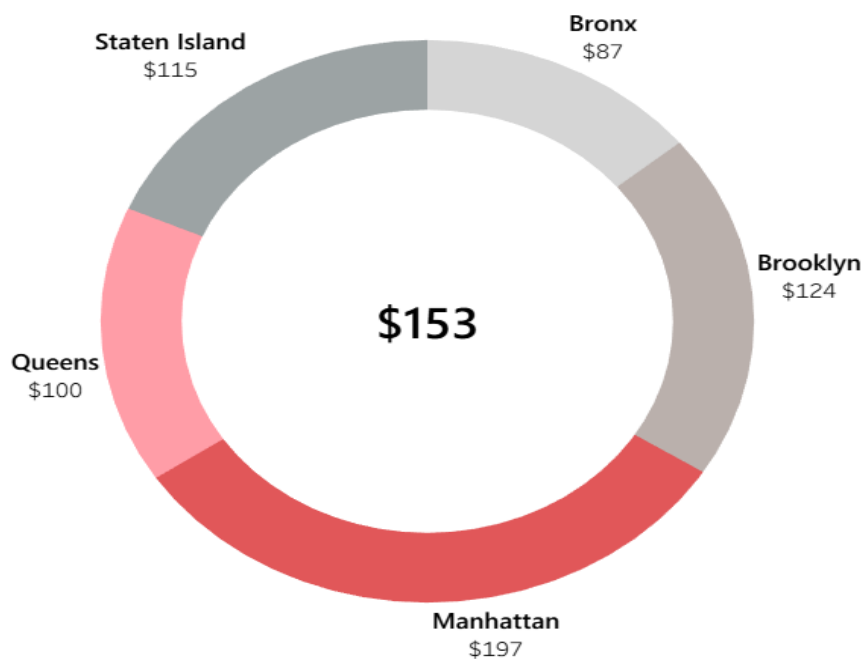
- The average price in Manhattan is $197 which is highest, then followed by Brooklyn $124, Staten Island $115, Queens $100 and Bronx $87 which is lowest amongst all.

5. What is the most preferred room type in every neighborhood groups?

Most Prefered Room in Neighbourhood Group

Room type	Neighbourhood Group	Count
Entire home/apt	Manhattan	13,199
	Brooklyn	9,559
	Queens	2,096
	Bronx	379
	Staten Island	176
Private room	Brooklyn	10,132
	Manhattan	7,982
	Queens	3,372
	Bronx	652
	Staten Island	188
Shared room	Manhattan	480
	Brooklyn	413
	Queens	198
	Bronx	60
	Staten Island	9
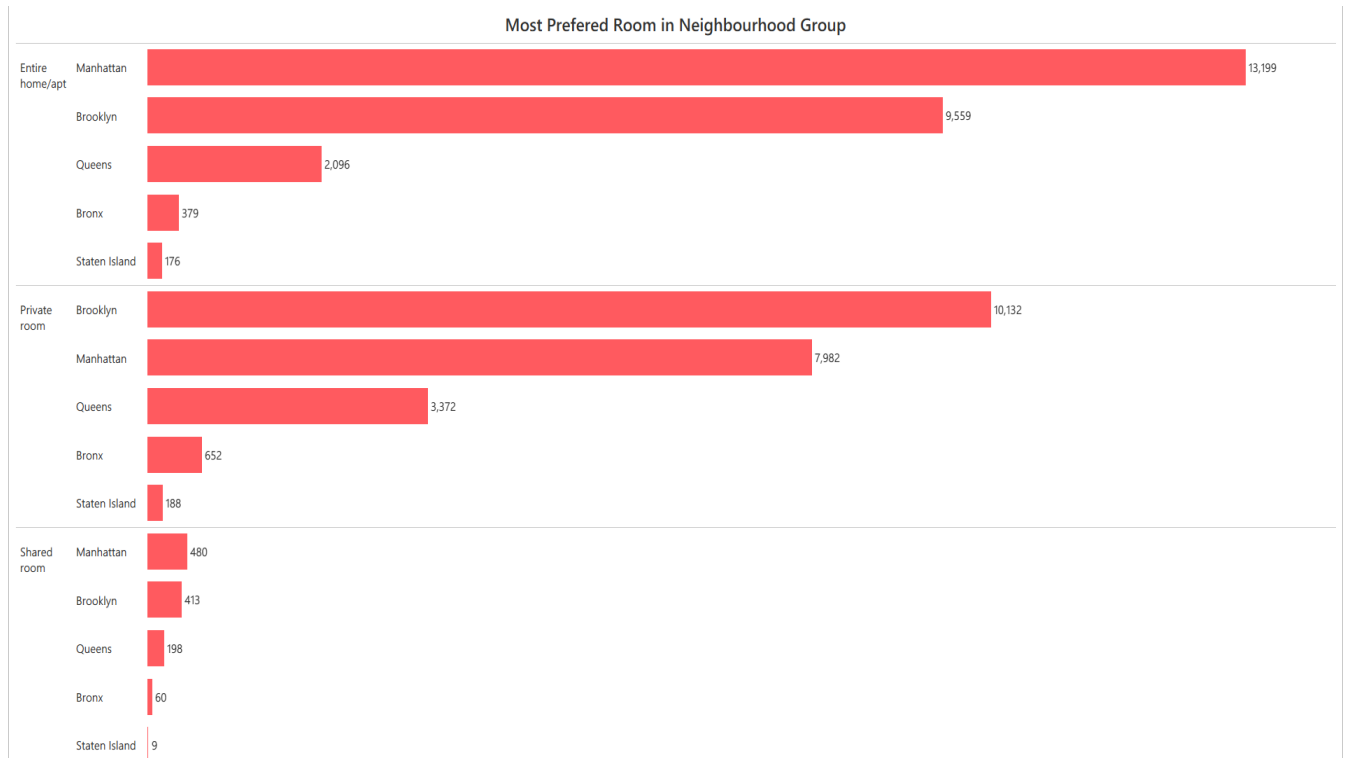
- Entire home/apt is most common in Manhattan and Brooklyn, while private rooms are also popular.

6. Total availability of properties having different room type?

Total Availability of Properties by Room types

Entire home/apt	28,43,783
Private room	24,82,739
Shared room	1,87,921

- Entire home/apt have the highest number of available properties across.

7. Which one is the busiest host?

Host Name	Price	Availability	Reviews	Listings
Michael	$66,895	38,888	11,081	1,043
David	$65,844	44,171	8,103	907
John	$41,892	40,249	7,223	2,792
Jason	$19,381	19,338	6,522	312
Alex	$52,563	30,031	6,204	475
Chris	$26,948	18,056	5,028	228
Anna	$27,003	20,580	4,799	269
Eric	$20,032	11,284	4,733	175
Daniel	$33,046	17,015	4,723	276
Sarah	$29,378	12,183	4,579	309
Kevin	$35,552	10,001	4,549	205
Michelle	$19,949	14,017	4,417	208
Maria	$24,328	24,920	4,405	338
Andrew	$33,694	14,832	4,150	264
Jessica	$50,697	21,520	4,074	453
Mike	$29,355	30,498	4,039	2,910
Maya	$5,758	5,140	3,938	78
Anthony	$18,065	18,968	3,877	858
Amy	$32,429	12,434	3,810	376
Robert	$17,454	10,065	3,781	154
Mark	$27,669	19,637	3,780	237
Brian	$24,466	11,934	3,705	220
James	$33,398	13,184	3,645	201
Danielle	$9,828	7,488	3,484	110

- Michael is the busiest host which has given almost 11081 reviews.

8. Which property has maximum number of reviews?

Property with Maximum Number of Reviews

Neighborhood	Reviews
BedfordStuyvesant	1,10,352
Williamsburg	85,427
Harlem	75,962
Bushwick	52,514
Hell's Kitchen	50,227
East Village	44,670
East Harlem	36,446
Crown Heights	36,408
Upper West Side	36,058
Upper East Side	31,686
Lower East Side	24,161
Chelsea	23,641
Midtown	19,444
Greenpoint	19,429
Astoria	19,310
Washington Heights	17,161
East Elmhurst	15,107
West Village	14,885
Flushing	14,818
Park Slope	14,638
Clinton Hill	14,586
ProspectLefferts Gard..	14,051
Flatbush	12,787
East Flatbush	12,448
Long Island City	12,256
Prospect Heights	10,875
Fort Greene	10,608
South Slope	10,405
Chinatown	9,941
Jamaica	9,910
Sunnyside	8,070
Sunset Park	7,882
Ditmars Steinway	7,852
Ridgewood	7,778
Gowanus	7,709
Gramercy	7,682
SoHo	7,235
Financial District	6,931
East New York	6,759
Greenwich Village	6,526
Carroll Gardens	5,897
Springfield Gardens	5,873
Kips Bay	5,837
Nolita	5,634
Elmhurst	5,554
Canarsie	5,338

- Bedford Stuyvesant has the maximum numbers of reviews almost 1,10,352.

9. What is the average price in each neighborhood group by room type?

Average Price by each Neighbourhood Group by room type
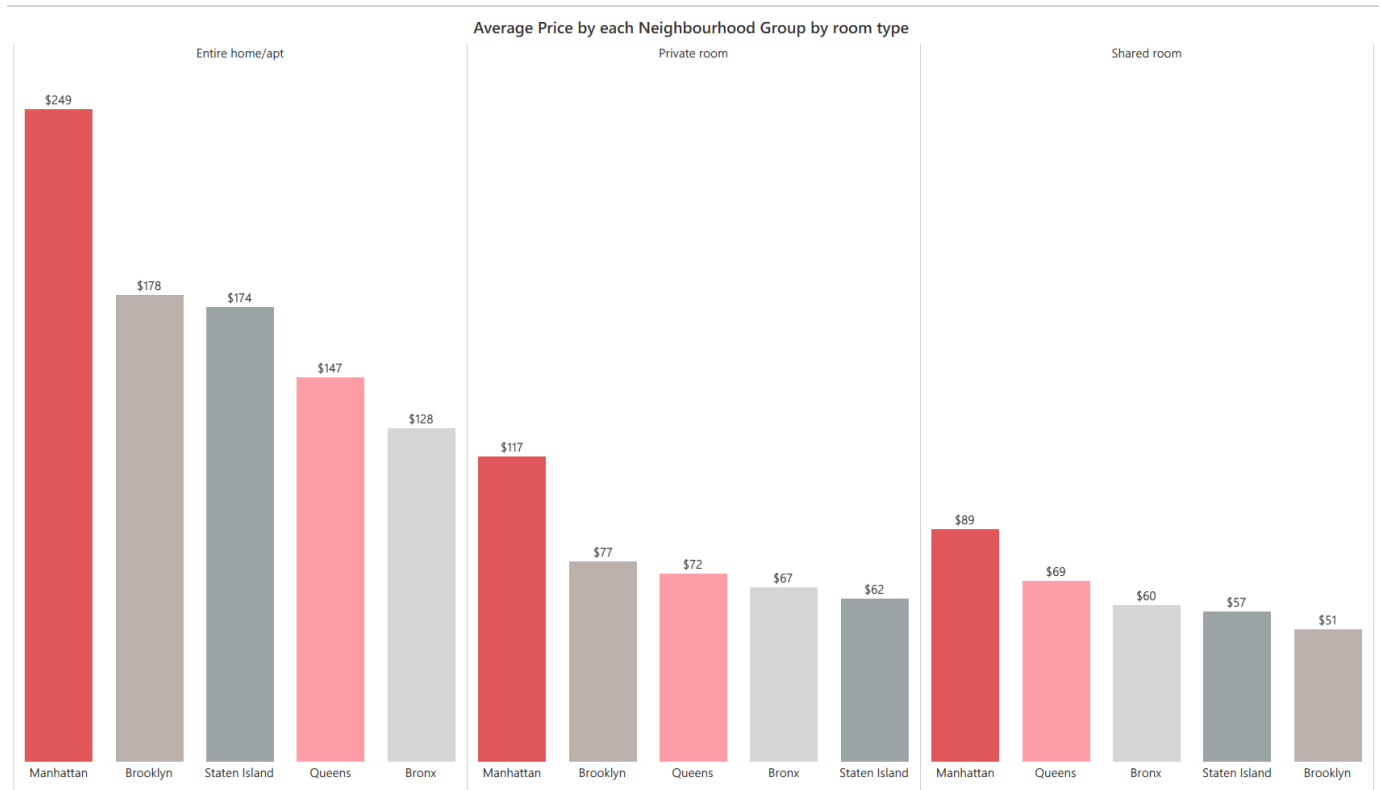
- Entire home /apt is the most expensive in the Manhattan neighborhood group the average price is $249 while Bronx is lowest with $128
- Private room is also expensive in the Manhattan neighborhood group the average price is $117 while Staten Island is the lowest with $62
- Shared room is cheap but highest in the Manhattan neighborhood group the average price is $89 While Brooklyn is the lowest with $51 .

KEY PERFORMANCE INDICATOR(KPI)

1. Maximum number of properties listed in neighborhood groups
2. Host with Maximum number of properties listed.
3. Host with Maximum number of properties listed having maximum properties listed.
4. Average price in different properties listed.
5. Most preferred room type in every neighborhood groups.
6. Total availability of properties having different room type.
7. Properties with maximum number of reviews.
8. Average price for each neighborhood group for room type shared room.

CONCLUSION

- Manhattan and Brooklyn dominate the market with more than 2 lakhs of total listings, making it the most active neighborhood group.
- It also records the highest average price $197 for shared rooms.
- Bedfordstuyvsant stands out for having the most reviews signaling high guest engagement and popularity in that sub-region
- **S**onder (NYC) is the top-performing host in terms of property count, especially in Manhattan, while Michael property has the most reviews showing strong guest interest and trust.
- Other top-reviewed hosts Michael, David, John, Jason indicates potential engagement, customer satisfaction, dissatisfied with service for that particular location**.**
- Entire home/apartment is both the most expensive $249 average and widely available, especially in Manhattan and Brooklyn.
- Private rooms ($62) and shared rooms ($51.) offer budget-friendly alternatives and also show steady booking patterns which is in Staten Island and Brooklyn.
- Consistent review patterns 1–2 reviews per month across listings and neighborhoods help customers judge reliability and performance.

Strategic Recommendation:

1) **Market and demand:** Manhattan and Brooklyn are most in demand for the most listing and having lots of reviewers and guest are willingly to pay higher price specially for entire home/apt.
   - We can prioritize this place and to proper marketing, run campaigns also improve services with customer satisfaction and will increase more tourist's attraction.
   - Slowly increase the price, so we can get good ROI

2) **Product and Inventory: for Room type**
   For entire room/apt: Treat as premium product, do better listing description, upselling services like late checkout, luggage storage improve room interior for customer satisfaction.
   Private room: Clearly position a budget but quality, Promote cleanliness and safety for the tourist's attraction
   Shared room: Don't invest to much on this niche's, just keep it clean for used because people are not preferring it. Later can upgrade it, in a private room for better revenue and pricing.

3) **Host Performance:**
   - Simplify onboarding short guide on room, services, pricing and responding to the reviews.
   - Improve for better guest experience fast response with high rating or just volume in reviews.

4) **Pricing and revenue:**

   - Increase daily average rate where demand is high
   - Stop overpricing in low demanded area will lose lot of customers

5) **Demand and engagement (reviews)**

   - Built the trust with the customers So they can give good reviews, which might lead to drive more customers in the locality.
   - Increase the rate of conversion with the customer that will directly affect the revenue.

Q & A

Q1) What's the source of data?

Ans) The Dataset was taken from Kaggle

Q2) What was the type of data?

Ans) The data was a combination of numerical and Categorical values.

Q3) What's the complete flow you followed in this Project?

Ans) Refer slide 5th for better Understanding

Q4) What techniques were you using for data?

Ans) -Removing unwanted attributes

-Visualizing relation of independent variables with each other and output variables

-Removing outliers

-Cleaning data and imputing if null values are present.

-Converting Numerical data into Categorical values.

Q6) What were the libraries that you

used in Python?

Ans) I used excel , pivot tables and tableau for visualization.

THANK YOU