ADS ASSIGNMENT 02

MUSKAN PATEL (3110495)

**Part I: Exploratory Analytics and Visualization of Data via Information Dashboard**

**Domain**:

For this assignment, I have selected the domain of Airbnb rental listings in New York City (2011-2019). The dataset contains information about Airbnb rental listings in New York City, including information about the host, the location, the price, and the property.

**Data Source**:

The dataset I have acquired is the New York City Airbnb Open Data from Kaggle. The dataset contains 49,000 observations and 16 variables. The data is reliable as it was collected directly from Airbnb, and it is updated regularly.

Link to the dataset: <https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data>

**Variables Used in Data Set:**

* **id**: A unique identifier for each Airbnb rental listing.
* **name**: The name of the Airbnb rental listing.
* **host\_id**: A unique identifier for the host of the Airbnb rental listing.
* **host\_name**: The name of the host of the Airbnb rental listing.
* **neighbourhood\_group**: The borough in which the Airbnb rental listing is located.
* **neighbourhood**: The neighbourhood in which the Airbnb rental listing is located.
* **latitude**: The latitude of the location of the Airbnb rental listing.
* **longitude**: The longitude of the location of the Airbnb rental listing.
* **room\_type**: The type of room being listed (e.g., entire home/apartment, private room, shared room).
* **price**: The nightly price for the Airbnb rental listing.
* **minimum\_nights**: The minimum number of nights required to book the Airbnb rental listing.
* **number\_of\_reviews**: The number of reviews for the Airbnb rental listing.
* **date**: The date of the last review for the Airbnb rental listing.
* **reviews\_per\_month**: The average number of reviews per month for the Airbnb rental listing.
* **host\_listings\_count**: The number of Airbnb rental listings that the host has.
* **availability\_365**: The number of days in a year that the Airbnb rental listing is available.

**Relationship Between Variables:**

The variables in this dataset are mostly independent of each other, with the exception of price, which is dependent on the type of room, the location, and the availability of the listing. The variables last\_review and reviews\_per\_month are also dependent on the number\_of\_reviews variable.

**Questions to Explore in Dataset:**

1. What are the most common types of Airbnb rental listings in New York City?
2. Which boroughs have the most Airbnb rental listings?
3. What is the average price of an Airbnb rental listing in New York City?
4. What is the availability of Airbnb rental listings throughout the year?
5. What factors affect the price of an Airbnb rental listing?

**Problem statement:**

The objective of this project is to perform a comprehensive analysis of the Airbnb rental market in New York City with the goal of identifying trends, patterns, and factors that impact the pricing and availability of the listings. The growing popularity of Airbnb in New York City has resulted in a highly competitive and dynamic rental market that is constantly evolving, making it challenging for both hosts and guests to make informed decisions. Therefore, this analysis aims to provide valuable insights into the rental market that can be used by hosts to optimize their listings, by travellers to make informed decisions about their accommodation options, and by policy makers to regulate the market and ensure fair competition.

To achieve this objective, the project will gather and analyse a large dataset of Airbnb rental listings in New York City, including information on location, price, availability, amenities, and host characteristics. The analysis will focus on identifying patterns and trends in the data, such as seasonal fluctuations in demand and pricing, popular neighbourhoods and the impact of host characteristics on rental prices and availability.

The insights obtained from this analysis can have significant implications for a wide range of stakeholders. Hosts can use the findings to optimize their rental listings, such as adjusting prices to better meet the needs of guests. Travelers can make more informed decisions about their accommodation options based on the insights gathered from this analysis, such as choosing the best time to book or identifying the most affordable and desirable neighbourhoods. Policy makers can also use the findings to regulate the Airbnb rental market and ensure fair competition among hosts. Overall, this project has the potential to provide valuable insights that can benefit multiple stakeholders in the Airbnb rental market in New York City.

15 Information Dashboards:

* **Dashboard Name: New York Overview**

1. Target audience: General public
2. Purpose: To provide an overview of the Airbnb rental market in New York City
3. Information presented: Number of listings on map through neighbourhood groups, average price, average review and average availability of these neighbourhoods across New York.
4. Actionable insights: Users can use this dashboard to choose a neighbourhood and a type of listing based on their budget and availability, while hosts can use this dashboard to set the price of their listing and optimize their occupancy rate. Policy makers can use this dashboard to monitor the impact of Airbnb on the housing market and regulate its growth.

* **Dashboard Name: Types of Listings**

1. Target Audience: Travelers and tourists who are looking for a place to stay in New York City.
2. Purpose: To present information on the different types of listings available on Airbnb in New York City.
3. Actionable Insights: The dashboard provides insights into the various types of listings available, which can help travellers understand the options they have while planning their trip to New York City. The dashboard can also help travellers determine the range of prices they can expect to pay for different types of listings, which can be helpful for budgeting purposes.
4. Audience: Travelers and tourists who are planning a trip to New York City and are interested in exploring their options for accommodation.

* **Dashboard Name: Price Distribution**

1. Target audience: Airbnb hosts and potential renters
2. Purpose: To show the distribution of nightly prices for all listings in NYC.
3. Information presented: This dashboard shows average price distribution by neighbourhood groups and property type. It also shows average reviews of neighbourhoods and review across the years 2011 - 2019.
4. Actionable insights: Hosts can use this dashboard to determine what price range their listing falls into and adjust accordingly. Potential renters can use this dashboard to determine what price range is common for listings in their desired neighbourhood.

* **Dashboard Name: Reviews by Month**

1. Target audience: Airbnb hosts
2. Purpose: To show how the number of reviews for listings change over time
3. Information presented: This dashboard shows the reviews trend by property type, Monthly review , top monthly reviews and Review distribution over years.
4. Actionable insights: Airbnb hosts can use this information to determine the busiest times of the year and adjust their prices accordingly.

* **Dashboard Name: Price vs. Availability**

1. Target audience: Airbnb hosts
2. Purpose: To show how the price of a listing is affected by its availability
3. Information presented: This dashboard shows the relationship between price and availability, Heat Map of availability, Average price by availability and price distribution of availability.
4. Actionable insights: Airbnb hosts can use this information to optimize their pricing strategy based on availability.

* **Dashboard Name: Popular Hosts**

1. Target audience: Airbnb guests
2. Purpose: To show the most popular Airbnb hosts in New York City
3. Information presented: This dashboard shows the top 10 hosts by listing, average rating, host rating across neighbourhood groups and map to locate the hosts along with host details.
4. Actionable insights: Airbnb guests can use this information to book with the most popular and highly rated hosts

* **Dashboard Name: Room Type by Neighbourhood**

1. Target audience: Airbnb hosts and potential renters
2. Purpose: To show the distribution of room types across different neighbourhoods in NYC.
3. Information presented: This dashboard shows the room types (private room, entire home/apt, shared room) for each neighbourhood/neighbourhood group, Heat map of room type in each neighbourhood, comparison of room type across neighbourhoods and Average price of room types.
4. Actionable insights: Hosts can use this dashboard to determine what types of rooms are in high demand in specific neighbourhoods and adjust their listings accordingly. Potential renters can use this dashboard to determine which neighbourhoods offer the type of room they are looking for.

* **Dashboard Name: Price Distribution by Borough**

1. Target audience: Airbnb hosts
2. Purpose: To provide an overview of the distribution of Airbnb listing prices by borough
3. Information presented: This dashboard shows the price distribution of Airbnb listing by borough, price comparison by neighbourhoods.
4. Actionable insights: Airbnb hosts can use this information to determine the appropriate price for their listings based on the prices of similar listings in their borough.

* **Dashboard Name: Top 10 Hosts**

1. Target audience: Airbnb hosts
2. Purpose: To show the top 10 hosts with the most listings and their average nightly prices.
3. Information presented: This dashboard shows top 10 hosts by listing and average price, relationship between listing and price for top 10 hosts, Top 10 hosts by average price and top 10 hosts by listing counts.
4. Actionable insights: Hosts can use this dashboard to determine what they can do to increase their number of listings or their average nightly price, by looking at what the top hosts are doing.

* **Dashboard Name: Host Review Rating**

1. Target audience: Airbnb hosts and potential renters
2. Purpose: To show the percentage of hosts who have different ratings of review and their average prices.
3. Information presented: This dashboard shows Listings count, average price by review rating for neighbourhood groups, relationship between review rating and price and count of listings by review rating.
4. Actionable insights: Hosts can use this dashboard to determine if increasing their review rating could lead to an increase in their average price. Potential renters can use this dashboard to determine if they prefer to rent from hosts with more rating.

* **Dashboard Name: Availability by Month**

1. Target audience: Airbnb hosts and potential renters
2. Purpose: To show the availability of listings throughout the year.
3. Information presented: This dashboard shows availability of listing by property type, Monthly availability by borough/neighbourhood, availability by month and heat map of availability.
4. Actionable insights: Hosts can use this dashboard to determine what times of year are in high demand and adjust their prices accordingly. Potential renters can use this dashboard to determine what times of year they may be able to find the best deals.

* **Dashboard Name: Amenities**

1. Target audience: Airbnb hosts and potential renters
2. Purpose: To show the most common amenities offered in listings and how they affect prices.
3. Information presented: This dashboard shows popular features by neighbourhoods, Pricings and listing details, average price by feature, and popular features.
4. Actionable insights: Hosts can use this dashboard to determine what features they should offer to increase their price. Potential renters can use this dashboard to determine what amenities they should expect.

* **Dashboard Name: Room type distribution by neighbourhood group**

1. Target Audience: Hosts and Airbnb management team
2. Purpose: To show the distribution of room types (entire home, private room, shared room) across different neighbourhood groups
3. Information presented: This dashboard shows neighbourhood and room type distribution, heat map of room type distribution, room type availability across neighbourhoods.
4. Actionable Insights: This dashboard helps hosts and Airbnb management team to understand the popularity of different room types in different neighbourhood groups. They can use this information to adjust their pricing strategy and focus on the room types that are in demand in each neighbourhood group.

* **Dashboard Name: Relationship between price & availability**

1. Target Audience: Hosts and Airbnb management team
2. Purpose: To show the relationship between the price of listings and their availability
3. Information presented: This dashboard shows price and availability of rooms for minimum nights, price and availability details, price vs availability and distribution of availability by price.
4. Actionable Insights: This dashboard helps hosts and Airbnb management team to understand the relationship between price and availability of listings. They can use this information to adjust their pricing strategy and make sure their listings are priced appropriately based on their availability.

* **Dashboard Name: Number of listings by host & neighbourhood group**

1. Target Audience: Airbnb management team
2. Purpose: To show the number of listings for each host and neighbourhood group
3. Information presented: This dashboard shows Top neighbourhoods by no of listings, Listing count per host, top hosts by listing count and heatmap of listing count by neighbourhoods.
4. Actionable Insights: This dashboard helps Airbnb management team to understand the distribution of listings across hosts and neighbourhood groups. They can use this information to identify hosts with a large number of listings and neighbourhood groups with a high concentration of listings, and take appropriate actions to manage the supply of listings in those areas.

ADS ASSIGNMENT 03

MUSKAN PATEL (3110495)

**PART 2: Predictive Analytics and Visualisation of Insights**

We need to first perform exploratory analytics on the New York City Airbnb Open Data and identify the variables that might have a relationship. Below are some of the questions that the users of this dataset might have:

* What is the average price of Airbnb rentals in different neighbourhoods in New York City?
* Which neighbourhoods have the most listings and highest occupancy rates?
* What is the relationship between price and the number of bedrooms or bathrooms in a listing?
* How do the prices of Airbnb rentals vary across different seasons?
* Are there any correlations between the prices of Airbnb rentals and the reviews given by guests?

After performing exploratory analytics, we can identify at least four sets of variables where there might be a relationship between the variables. These relationships can lead to actionable insights. Below are some of the sets of variables that we can explore:

* **Relationship between price and neighbourhood**: We can explore the relationship between the average price of Airbnb rentals and the neighbourhood they are located in. We can create a dashboard that displays the average price of rentals in different neighbourhoods, the number of listings in each neighbourhood, and the occupancy rates.

1. Dashboard Name: Neighbourhood VS Price
2. Target Audience: Airbnb hosts and potential guests looking to rent in specific neighbourhoods
3. Purpose: To display the average price of rentals in different neighbourhoods, the number of listings in each neighbourhood, and the occupancy rates.
4. Information Presented: Average price of property type in neighbourhood, Average price in neighbourhood, average price of top neighbourhoods and average price of neighbourhoods over the years.
5. Actionable Insights: Airbnb hosts can adjust their pricing based on the average price of rentals in their neighbourhood and the number of listings available. Potential guests can make informed decisions on which neighbourhoods to rent in based on the average price and occupancy rates.

* **Relationship between price and reviews:** We can explore the relationship between the price of Airbnb rentals and the reviews given by guests. We can create a dashboard that displays the average price of rentals based on the review scores for different categories such as cleanliness, accuracy, and communication.
  1. Dashboard Name: Reviews VS Price
  2. Target Audience: Airbnb hosts and potential guests looking for high-quality rentals.
  3. Purpose: To display the average price of rentals based on the review scores for different categories such as cleanliness, accuracy, and communication.
  4. Information Presented: Average price based on review scores for different categories for each neighbourhood/ neighbourhood group over months.
  5. Actionable Insights: Airbnb hosts can adjust their pricing based on the quality of their rental based on the review scores, and potential guests can make informed decisions on which rentals to choose based on the review scores and price.
* **Relationship between price and time**: We can explore the relationship between the price of Airbnb rentals and the time of the year. We can create a dashboard that displays the average price of rentals for each month of the year. We can also explore the relationship between the price and the day of the week, the minimum number of nights, and the availability of the rental.
  1. Dashboard Name: Time VS Price
  2. Target Audience: Airbnb hosts and potential guests looking to rent during specific times of the year.
  3. Purpose: To display the average price of rentals for each month of the year, and the relationship between price and day of the week, minimum number of nights, and availability.
  4. Information Presented: Average price per year, relationship between price and years.
  5. Actionable Insights: Airbnb hosts can adjust their pricing based on the time of year, and potential guests can make informed decisions on which times of the year to rent based on the average price and availability.
* **Relationship between price and host characteristics:** We can explore the relationship between the price of Airbnb rentals and the characteristics of the hosts. We can create a dashboard that displays the average price of rentals based on host factors such as their response rate, acceptance rate, and the number of listings they have. We can also explore the relationship between the price and the host's verified identity, experience level, and super host status.

1. Dashboard Name: Host VS Price
2. Target Audience: Airbnb hosts, potential guests, and Airbnb management team
3. Purpose: To explore the relationship between the price of Airbnb rentals and the characteristics of the hosts and to identify trends and patterns in the data to improve their service offerings.
4. Information Presented: Top 10 hosts per neighbourhood by average price & availability, Review ratings and listing count of hosts by price, Top hosts, top 30 hosts by neighbourhood price.
5. Actionable Insights: Hosts can use this dashboard to determine how their host characteristics impact their rental prices and make adjustments to improve their pricing strategy. Guests can use this dashboard to understand how different host characteristics can impact rental prices and make informed decisions when booking a rental. The management team can use this dashboard to identify trends and patterns in the data to improve their service offerings and identify areas for improvement.

These dashboards can help decision-making processes by providing actionable insights to Airbnb hosts and potential guests. Hosts can adjust their pricing and amenities based on the data presented, and potential guests can make informed decisions on which rentals to choose. Improved decision-making can lead to more satisfied guests and increased revenue for hosts.

**Conclusion:**

In conclusion, the analysis of Airbnb rental listings in New York City has provided valuable insights into the trends, patterns, and factors that affect the pricing and availability of listings. The findings from the dashboards created so far reveal several key trends, such as the strong seasonal demand for rentals, particularly in popular tourist neighbourhoods like Manhattan and Brooklyn. The data also highlights the significant impact of location on rental prices, with listings in central Manhattan commanding the highest prices.

Moreover, the dashboards showed that the availability of listings is heavily influenced by various factors, such as the time of year, the type of property, and the number of amenities offered. Additionally, the analysis reveals that certain amenities, have become increasingly important to guests, as they significantly impact the demand for rentals.

The insights obtained from this analysis can be used by hosts to optimize their listings, by travellers to make informed decisions about their accommodation, and by policy makers to regulate the Airbnb rental market. Hosts can adjust their pricing and amenities to better meet the needs of guests, while travellers can use the insights to choose the best time and location to book their rentals. Policy makers can use the findings to implement regulations that ensure fair competition among hosts and promote the sustainability of the rental market.

Overall, the analysis of Airbnb rental listings in New York City has provided valuable insights into the rental market, and the findings have the potential to benefit multiple stakeholders in the industry.