# **The DataViz Challenge - Transforming Airbnb EDA Project to Dashboards**

**Project Type -** Product Dissection

**Contribution -** Individual

**Power BI dashboard**

**workbook link-** <https://drive.google.com/file/d/1-Tzm7laSLV_eqi-lzJuMhjo_S_aQs_e8/view?usp=sharing>

**Problem Statement:**

In the context of Airbnb operations, PowerBI's visual analytics capabilities are leveraged to uncover and illustrate the shared attributes, disparities, and distinctive patterns inherent to Airbnb's presence in New York citie, thus elevating the depth and insightfulness of the study. The goal is to create a comprehensive dashboard that highlights the shared attributes, disparities, and distinctive patterns within the Airbnb market, enabling stakeholders to make informed decisions and optimize their strategies based on a thorough understanding of the current market landscape.

**Dataset Selection:**

For this EDA project, I have chosen the "Airbnb Listings Data" dataset of New york city. This dataset provides a comprehensive snapshot of various attributes related to Airbnb listings, such as property type, neighbourhood, pricing, availability, and more. The dataset is ideal for conducting an in-depth exploration of the local Airbnb market and deriving actionable insights.

**Why Airbnb:**

First of all I have chosen Airbnb dataset as a project due to my interest in traveling and since 2008 Airbnb is the platform which is highly used to study the guests and hosts data effectively.

Secondly Airbnb, a prominent online platform, enables individuals to reserve accommodations spanning a spectrum from beds and rooms to apartments and entire homes across global locales. This user-centric platform serves as a conduit for seamless property rentals, negating the need for intricate intermediaries or substantial capital outlays. Notably, users can secure lodgings at significantly competitive rates relative to traditional hotels. Distinctively, Airbnb extends its reach to regions where convectional hotel presence might be limited, offering an avenue for lodging acquisition in underserved locales. Moreover, the inclination towards immersive local experiences often steers individuals towards selecting accommodations embedded within native communities, fostering a distinctive preference for authenticity and cultural engagement.

Airbnb Statistics • Over 4 million listings worldwide • 150 million users in 191 countries • Worldwide value is $32 billion • Global growth rate since 2009 - 153%

**Dataset Details:**

* Dataset Name: Airbnb Listings Data
* Source: [Link to dataset](https://drive.google.com/file/d/1-Tzm7laSLV_eqi-lzJuMhjo_S_aQs_e8/view?usp=sharing)
* Citie: New York
* Description: The Airbnb Listings Data contains information about different properties available for rent on Airbnb in a specific city. Each record represents a unique listing and includes attributes such as property type, neighbourhood, number of bedrooms, pricing, availability, host information, and more.

**Key Attributes**:

1. id: Unique identifier for each listing.
2. name: The title or name of the listing.
3. host\_id: Unique identifier for the host of the property.
4. host\_name: Name of the host.
5. neighbourhood\_group: The broader area or group that the neighbourhood belongs to.
6. neighbourhood: Specific neighbourhood where the property is located.
7. latitude: Latitude coordinate of the property.
8. longitude: Longitude coordinate of the property.
9. room\_type: Type of room (e.g., Private room, Entire home/apt, Shared room).
10. price: Price of the listing per night.
11. minimum\_nights: Minimum number of nights required for booking.
12. number\_of\_reviews: Total number of reviews received for the listing.
13. last\_review: Date of the last review.
14. reviews\_per\_month: Average number of reviews per month.
15. availability\_365: Number of days the listing is available for booking in a year.

**Problem Areas to Explore:**

1. Which are the popular neighbourhoods, their average prices and no. of listings?
2. What is the percent share of different property types and room types?
3. How the pricing is varying with location, property type, and reviews?
4. What are the different correlations between type of hosts and factors like- reviews & price?

Divided the visualisation findings into 3 categories:

* Property analysis according to area
* Pricing analysis
* Host analysis

**How to proceed with the dashboard:**

1. **Data Cleaning**

At the beginning I addressed the disorder and inconsistency within the dataset. Thereafter I Utilized Jupyter PowerBI Prep to systematically cleanse the data, rectifying discrepancies, eliminating duplicates, and standardising formats.

1. **Data Transformation**

In Data Transformation I Generated supplementary columns by utilising pre-existing categorical data. These columns will be derived from extensive descriptive text, which, in its original form, proved arduous to comprehend and unsuitable for visualisation purposes. The extra columns that I created gave a much clear sense of how to approach and make an effective visualisation.

1. **PowerBI**

Then Employed PowerBI Prep to leverage its distinctive "Remove Errors" and “Change data type” feature. There were some errors in properly loading the data which were removed by “Remove Error" feature of Power BI. There were some columns which did not have the right data type that were set by the “Change data type ” feature of Power BI.