Contributors:

**Akshad Goyanka**

* Data Discovery - Data import, data description , understanding variables.
* Data Wrangling - removing unnecessary columns , null values imputation
* Exploratory Data Analysis - bar charts, pie charts, scatterplots , histograms etc.

GitHub Link : <https://github.com/Aks18had/Airbnb_2019_NYC>

Google Colab Link: <https://colab.research.google.com/drive/1AeaACfGtWMoWC-Z2EPJ_DxrI5QQYkD-E?usp=sharing>

Google Drive Folder link: <https://drive.google.com/drive/folders/1_9aQWghSiKWLQvulA4Tz1-plKT8WJZvH?usp=sharing>

**Summary**

AirBnB is an online marketplace for rental properties founded in America. We analysed the given data consisting of around 49000 columns. We were given a dataset of 49000 entries & 16 variables. Our objective is to analyse the data & find the key factors that influence the property listings. We will use the techniques of Data Discovery, wrangling & exploratory data analysis to analyse the data to gain insights that will help the organisations make better informed decisions so that they could work efficiently.

The Project is divided into various parts

Data discovery - where we import the data & get some information about the data,

Data cleaning / wrangling - where we will check the data for any inconsistencies & clean the data to make it more useful for later steps,

EDA or exploratory data analysis - where we will analyse the data using charts & statistics [if necessary]. This can help us finding valuable insights that we can use to make decisions that will help our organisation to increase their profits.

# **Conclusion**

* *Summarising the key insights and their implications for Airbnb: Pricing strategies might need to be neighbourhood-specific.*
* Marketing efforts could focus on promoting less popular room types or areas.
* Hosts with multiple listings contribute significantly to the platform's supply.
* Reviews are crucial for understanding listing popularity and guest satisfaction
* The dataset contains around 49,000 observations and 16 columns, with a mix of categorical and numeric values.
* The columns likely include information such as listing ID, name, host ID, neighbourhood, room type, price, minimum nights, number of reviews, last review date, reviews per month, number of listings per host, availability, and more.
* From the above data analysis process we can deduce that there are certain factors that impact the properties of listings , like geographical locations , availability , minimum nights etc. The above visualisations displays these very factors & how they impacts the listings.