**Air BnB Case Study**

**Methodology Document**

* **Mr. Justin Dias**
* **Mandar Dhumal**

**Understanding the Case:**

* We analysed the given AirBnB dataset and tried to understand the various data fields from csv file
* We visualised a rough skeleton and flow of the presentation to tell the data story insights which will cover an important aspect of our story and that will make insightful presentation to the stakeholders

**Reading the Data:**

* In Python, we read the data and checked for its various attributes and fields for null values and outliers
* Null values for reviews per month were replaced with median and column the last review was dropped as it did not bring in any important insights for our case study
* We checked for outliers, but we did not deal with it as our case study did not necessarily require us to deal with it

**Analysis and Visualisation:**

* We analysed the data set in python as well as in tableau to draw meaningful insights by conducting univariate and bivariate analyses and tried to draw a conclusion from the correlation analysis.
* We created many charts like pie, stacked bar charts, group bar charts, boxplot maps, etc. in tableau which enhanced our data storytelling efforts for this case study

**Meaningful Insights:**

* We observed that people preferred to visit the center of New York to enjoy the serene view of the city. Hence, **Manhattan and Brooklyn** boroughs are mostly preferred, have the majority of the listings, and have high average prices.
* As the number of listings of shared rooms is limited, though its average price is low, its availability is high. High availability days mean not enough bookings for such properties.
* Acquiring more properties and pushing the hosts to go for competitive pricing in prime locations like **Manhattan and Brooklyn** will increase the market for Air BnB.
* Properties in **Staten Island** should come up with ways and means to market themselves to attract more customers. Staying near the Statue of Liberty can be marketed
* Entire apartments are popular and so are private rooms amongst the customers.
* Outliers and some errors like the Price of some properties are 0 should be dealt with at the time of the data collection stage.