**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| **G. MOHAMED LUQMAN (** [**lukiengr@gmail.com**](mailto:lukiengr@gmail.com) **)**  **Descriptive Analysis:**   * Dataframe Description * Dataframe shape   **Data Cleaning and Analysis:**   * Handling null values of ‘reviews per month’ * Handling null values of ‘last\_review’ * Handling null values of ‘name’ * Handling null values of ‘host\_name’ * Find zero price columns and remove them * Replacing NaN values from columns like name and Host Name as ‘un available’ * Replace NaN values and Imputing mean() value in ‘reviews per month’ column * Drop unwanted column [‘ID’,’last\_review’]   **Data Wrangling:**   * Relation between ‘Price’ and ‘reviews’ * Relation between ‘Room\_type’ and ‘minimum\_nights’ * Relation between ‘Room\_type’ and ‘number of reviews’ * Relation between ‘Neighbourhood group’ and ‘number of reviews’ * Calculation of average price per neighborhood group * Room types preferred in Manahattan and Brooklyn   **Data Visualization:**   * Relationship between Room\_type’ and ‘minimum\_nights’ * Relationship between Room\_type’ and ‘Number of reviews’ * Relationship between Neighbourhood\_group’ and ‘number of reviews’ * Relationship between ‘Neighbourhood\_group’ and ‘number of host * Relationship between Room\_type’ and ‘Neighbourhood\_group’ * Plotting Graph of busiest host * Plotting satellite map graph of price in neighbourhood\_group * Plotting word cloud of column ‘name’ |
|  |
| Github Link:- https://github.com/lukilearn/Capstone\_1\_AirBnb\_NYC\_2019 |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| **Airbnb (AirBed & Breakfast, where Air was nothing but mattresses) comes with a small idea ofearning rent by Brian Chesky and Joe Gebbia in 2007 turned hospitality and travel industry on itshead and gained massive recognition now. In this Capstone Project, we were given dataset of 49000observations distributed in 16columns.Columns are of two types categorical and numerical type.**  **Initially we started with analyzing the data. And after some initial analyzing , it was gathered that some columns are not useful and then we drop them. And some columns have NaN (null) values which also need removal. This whole process waster med as Data Cleaning and Analyses.**  **Next moving forward , started with Descriptive analysis over the raw data provided using box plot, correlation and other descriptive methods. The Project is now further broken into majorly four parts according to the question we need to answer.**  **At first, hosts and room\_types were analysed. From that it was concluded that ‘Entire home/apt’ were most preferred type of room. On analysis, using data visualization techniques a clear relationship between room\_type and minimum number nights spent was shown, which concluded that if the rooms type is ‘Entirehome/apt’ then customers spent more number of nights there.**  **Next after analyzing the neighbourhood group with respect to average price at each location, Manhattan was concluded as costliest state among others and possible reason could be the financialstateof entirecountry. People are not money conscious when spending money mainly in Manhattan and Brookyln as they want more privacy**  **And then analyzing was done based on ‘number of reviews and host listing’ which provided host were giving more reviews in most listed neighbourhood**  **Also we looked for the top three busiest host in US and also in top three busiest in each neighbourhood group**  **After exploring more about neighbourhood, we can conclude that shared room type were less likely preferred by US customers of Airbnb. And Bedford -studyvesent and Williams burgwere most populated from Manhattan state.**  **We can conclude that most of the Airbnb hosts are present in Manhattan and Brooklyn. After correlation between price and neighbourhood,** |