**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| **Name:** Indugopal Maity **Email-Id:** [ig.maity@gmail.com](mailto:ig.maity@gmail.com)  **Contribution:**   1. Data Wrangling    * 1. Booking Dataset      2. Host Name, Neighborhood group, Calculate Host Listing Count, Location, Number of Reviews and Price 2. Calculated relationship between “Neighborhood Group” and “Calculated Host Listings Count”. 3. Calculated relationship between “Neighborhood Group” and “Number of Reviews”. 4. Calculated relationship between “Price” and “Number of Reviews”. 5. Calculated relationship between “Neighborhood Group” and “Price”. 6. Found out what points to be given in the Presentation and found the relevant picture. 7. Found the disruption Economy with relevant Picture in presentation. 8. Made the steps and algorithm with a clear picture view in presentation. 9. Made the data set and graph to find relationship among “Hosts” and “Areas”. 10. Made the data set and graph to find relationship among “Groups” and “Reviews”. 11. Made the data set and graph to find relationship among “Price” and “Reviews”. 12. Made the data set and graph to find relationship among “Groups” and “Price”.   **Name:** Md. Tousif Ali **Email-Id:** [tousifali2018@gmail.com](mailto:tousifali2018@gmail.com)  **Contribution:**   1. Data Wrangling    * 1. Booking Dataset      2. Host name, Host Id, Room Type, Number of Reviews, Neighborhood Group and Minimum Nights 2. Calculated relation between “Number of Reviews” and “Name of the host”. 3. Calculated relation between “Neighborhood group” and “Minimum Nights Spend”. 4. Found the basic details about Airbnb, “Full Form”, “Where it operates”. 5. Detailed list of facilities provided by Airbnb 6. Where and How we can find Airbnb in presentation. 7. Found the busiest hosts with number of reviews. 8. Found the reason to find the busiest hosts. 9. Difference of traffic with different places with reason. 10. Found The graphical interpretation of traffic for different place. 11. Found a hypothetical conclusion based on the action performed. |
| **Please paste the GitHub Repo link.**  **https://github.com/indugopal1991/Airbnb-Booking-Analysis/blob/main/Indugopal\_Maity\_Airbnb\_Bookings\_Analysis\_Capstone\_Project.ipynb** |
| Github Link:- https://github.com/indugopal1991 |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**  As the first step, performed **“DATA Wrangling”** over the raw data, **“Airbnb NYC 2019**” and we can see that there are altogether **782320** data with a combination of **48895** rows and **16** columns. The chart is consisting of several necessary columns like **“Id”**, **“Name”**, **“Host\_id”**, **“Host\_name”**, **“Neighbourhood\_group”**, **“neighbourhood”**, **“room\_type”**, **“price”**, **“Minimum Nights”**, **“Number\_of\_reviews”**, **“calculated\_host\_listings\_count”** and **“availability\_365”** so we have kept those in the new file **“new\_df”**. Some of the columns like **“latitude”**, **“longitude”**, **“last\_review”** and **“review\_per\_month”** are not that much important to make any comparison in the report so we have not kept that in the file**” new\_df”**.  Farther we have divided the complete project into four different parts as per the **“Review”**, **“Location”**, **“Host area”**, **“Traffic”**, **“Price”**, **“Host\_name”, “Minimum Nights Spend”** and **“Busiest Host”**.  In the first part, we tried to find out relation between **“Host Name”**, **“Neighbourhood Group”** and **“Calculate Host Listing Counts”**. We have written the code such that we will get first ten data in descending order to know who are on the top of the list.  In the second part, we have reviewed three things **“Neighbourhood Group to Number of Reviews”**, **“Price to Number of Reviews”** and **“Neighbourhood Group to Price”** to understand the relationship between them. We understand few things from these, firstly which particular area`s customers like to give the most number of reviews, secondly where the most numbers of people like to stay and why and finally price difference between different **“Neighbourhood Group”**.  In the third part, we have checked what are the **“Busiest hosts”** in terms of **“Number of Reviews”** and what is the **reason** behind that.  In the final part, we have reviewed in which particular area is on the top based on **“Minimum nights Spent”**.  So, by the analysis we have figured out certain things, those are listed below:   1. Nature of the people who stay in **“Entire Home/ Apartment”**. 2. People staying in **“Private Room”** stay lesser number of days with respect to **“Entire Home/ Apartment”**. 3. Most people prefer to pay **less to stay**. 4. If more the numbers of reviews, then number of tourist are more as they come to spend leisure time. 5. If people are not staying more than one night, in most of the cases they are travelers. |
|  |