Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

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 was conceived years ago by two roommates who rented out an air mattress in their living room. This turned their whole apartment into a bed and breakfast. This was done to sustain the high-priced living in San Francisco. This gave the company its name Air-bed-and-breakfast. This San-Francisco-based start-up offers you someone's home as a place to stay instead of a hotel. Airbnb was started in 2008. Airbnb does not own properties. It acts as an intermediary between those who want to rent out space and those who are looking for space to rent.

Starting with raw data, perform data wrangling over it. In addition, we divided this project into different sections, such as Data Exploration and Variable Identification, Understanding Data, Handling NaN Values, Exploring and Visualizing Data, Single Variable Analysis, Bi-Variable Analysis.

For managing NaN values mainly focuses on important elements like neighborhoods, neighborhood groups, host ids, counting, etc. Several features were dropped from the analysis because they were of no value to us. Depending on the feature, we replace all null values with '0', 'unknown', etc.

During the exploration and visualization of data, we performed single and bi-variable analysis according to the requirements. We found out correlation between various features. We explored relationship between various aspects such as which city has the most listings, guest's preferred which types of room also what is effect of room location and availability of room on guest's preferences etc.

At the end we examined the Price feature where we detected and removed outliers from it by using the quantile method. Also, we found out cheapest and costliest listings in entire NYC, average prices of rooms in each neighbourhood group.

Please paste the GitHub Repo link.

GitHub Link: - https://github.com/Avinash-Saudagar/AirBnb-Analysis

Team Member's Name, Email and Contribution:

Avinash Saudagar | avinashsaudagar5.as@gmail.com **|**

Contribution Role:

- Data Wrangling
    1) Acquiring Data – Airbnb NYC 2019
    2) Variable Identification
- Data Analysis
    1) Co-relation matrix
    2) Single-variable Analysis
        a. Neighbourhood group Vs Number of listings
        b. Top 10 reviewed hosts on basis of review per month
    3) Bi-Variable Analysis
        a. Count of each room type in neighbourhood group in entire NYC
        b. Monthly review variations with room types in each neighbourhood groups
    4) Price Feature
        a. Room types Vs Price in different neighbourhood groups
        b. The costliest and cheapest listing & their respective hosts in entire NYC

Chaitanya Sawant | chaitanyasawant1298@gmail.com **|**

Contribution Role:

- Data Wrangling
    1) Acquiring Data – Airbnb NYC 2019
    2) Variable Identification
- Data Analysis
    1) Handling Null Values
    2) Single-variable Analysis
        a. Top 10 neighbourhood entire NYC on the basic of count of listings
        b. Minimum stays in different room types
        c. Top 10 host (id) with the greatest number of listings
    3) Bi-Variable Analysis
        a. room type and their relationship with availability and with different neighbourhood groups
    4) Price Feature
        a. Detecting outliers and removing them
        b. Top neighbourhood groups in NYC with respect to average price/day of Airbnb listings