<div align="center">

Capstone Project Submission

</div>

<u>Instructions:</u>
i) Please fill in all the required information.
ii) Avoid grammatical errors.

Team Member's Name, Email, and Contribution:

1) Piyush Lanjewar
   Planjewar51@gmail.com
   - Data wrangling
     Dealing with price column
     Handling missing values in the reviews column
   - Top 3 Hosts counting
   - Top reviews
   - Price distribution with the different neighborhood groups
   - Using Plotly.express to define the traffic conditions
   - Defining Superhost
   - Defining busiest host

2) Pruthvi Raj
   Pruthviraj1698@gmail.com
   - Data wrangling
   - Handling missing values
   - Define outliers using boxplot
   - Defining Traffic conditions using Folium library
   - Finding most popular hosts
   - Defining the hosts who got the most reviews
   - Outliers plot concerning prices and area
   - Defining availability of hosts in different areas

3) Yogesh Reddy
   Yogeshreddy005@gmail.com
   - Data wrangling
   - Defining avg prices in the different neighborhood groups
   - Defining minimum nights in different areas
   - Showing visual reports of different areas
   - Conclusion Report

Please paste the GitHub Repo link.

Github Link:- https://github.com/Pruthviraj009/Airbnb-Capstone

Please write a summary of your Capstone project and its components. Describe the problem statement, your approaches, and your conclusions. (200-400 words)

Airbnb is one of the largest used companies for lodging primarily homestays for vacation rentals and tourism activities. Today Airbnb is one of the most used brands for giving a good experience to hosts and guests. The data is used to increase the understanding of every detail to make traveling easy and convenient. The data is utilized to show the required conclusions.

We have given data that have 16 columns and 49000 observations with null values and missing values. There was a presence of erroneous data. It was very important to understand the data and give insights out of it.

First, we have to find which columns have lots of missing values. Deal with the missing values by replacing zero values. The price column has some of the missing values, So, the zero values or missing values are replaced by the median of the column (In this case we have taken the median other than the mean of the column because there are too many outliers present in the data).

Replacement of zero in place of NA values had applied in the reviews_per_month column. The summary statistic was then observed by describing the method and then concluded that there are no insignificant values present after applying operations on the data.

The data is now ready to do visualizations through it.
Various results are concluded with the help of plotly.express and folium libraries.