Group 10
ECEN 360 / STAT 315
Michael Cubriel, Gabe Dare, Joaquin Salas, Sandip Singh

Project Proposal

For the “Defining the Questions” section we ask “Are there any significant trends/patterns in the Airbnb data between different cities?” We plan on using visualization techniques and statistical analysis to explore our data, containing Airbnb information across six different cities.

For the “Data Collection” section, the data we are provided with is enough to solve our previously stated question, so we will gather the existing data and use it to analyze the trends, identify the patterns, and draw an appropriate conclusion.

Furthermore, for the “Data Cleaning and Preparation” section of our notebook there already exists some data cleaning that removes null values, but we can still look into the data and make sure that there are no outliers that might drastically skew our results.

As for “Data Analysis” we plan to use different visualization techniques such as dot plots, contour plots, and bar graphs to help illustrate any noticeable patterns. We will also compare summary statistics and analyze our models to identify any relationships between variables. This will allow us to conduct further analysis through linear regression, logistic regression, or any appropriate type of regression analysis.

To interpret the data for the “Interpretation and Reporting” section and share our results of the analysis, we will accumulate the results found in our data analysis and highlight any noticeable differences and trends between different cities. We will explain if there are any important factors influencing these trends in a clear and concise manner.

For the “Contribution Report” part, our team will plan to meet several times until the project due date to stay on schedule. We will evenly distribute the tasks and create deadlines for different parts of our project.

Finally, in order to create the Docker image and the GitHub repo that houses all of the important files for our project, the member with the most confidence in Docker, will create the Docker image. Additionally, a GitHub repo will be created and the files from the Docker image will be uploaded to the repo and instructions on how to run the image will be drafted to explain how to access and run our notebook. We will also try running the Docker image on each of our computers to make sure that it is reproducible and easily accessible for someone following the instructions.