**STAT-493: Statistical Modeling**

Name: Kevil Khadka

Project Topic:

Mining Airbnb dataset Using Machine Learning and R Concepts

Explanation:

Airbnb is fastly growing popular online marketplace which offer lodging facilities for tourists. Recently, they start offering Airbnb adventures service to traveler who wants to explore the places. Instead of staying in 5-star hotel, Airbnb is widely popular for providing home experience to travelers in affordable price. Started in 2008, Airbnb service is now all over the world.

Big cities like Chicago, New York, Shanghai, San Francisco, etc., Airbnb is mostly admired. This project is going to perform an exploratory analysis of the Airbnb dataset to understand the major factors why people prefer Airbnb. Main focus is to understand the dataset using Machine Learning technique. Some ideas are listed below:

Ideas:

1. Data Exploration

* Structure and features of dataset, exploring missing data

1. Data Visualization

* Price vs Location, Seasons, Neighborhood area, availability rate, room type
* Demand for Airbnb rentals over the year
* Growth of Airbnb in higher traffic cities vs low traffic cities
* Home sharing vs commercial use; and its earning revenue

1. Machine Learning

* Split data into train/test sets
* Predicting price
* Linear Regression model
* K-means clustering
* (Maybe Decision tree)

My main purpose is to learn different methods of machine learning to explore the dataset.

Resources: <http://insideairbnb.com/get-the-data.html>