# AirBnB Price Prediction Challenge

### Suggest the price of AirBnB listings in major U.S. cities

## Introduction

The aim of this case study is to predict the price of AirBnB listings in major U.S. cities.

## Problem Statement

Given a dataset with 24 variables such as number of bedrooms and a log-price indicator (greater than 0) for each observation in the training data, the objective is to suggest the log-price of a particular listing using the 24 features provided for the test observations. The suggested prices that are closest to the true prices, as calculated by the Root Mean Squared Error

## Data

The training data consists of the following columns:

1. id : Record identifier
2. log\_price : log(price) (dependent variable)
3. …. : 24 features for the listing

There are two major files provided for this case study:

1. train.csv : Training data
2. test.xlsx : Test data

## Submission

The submission file**(test.xlsx)** for the case study should consist of two columns:

1. id : Record identifier for the test observation
2. log\_price : Prediction score for the listing (Type: Float)

**We recommend using one classical machine learning model and one deep learning model.**

## Evaluation

You will be evaluated on the following:

1. **Model Accuracy** – The evaluation metric is the Root Mean Squared Error (RMSE) score on Test data.

You can send your submissions to validate your model results on Test data along with the presentation deck.

**Innovation –** Explores the Key drivers of the target variable(**log\_price**) and Innovative use of unstructured data and ML techniques

1. **Presentation deck** (below should be the points covered in the deck)
   1. Introduction and Problem Overview
   2. Data Summary and Hypothesis
   3. Variable creation/Feature engineering and Interesting findings
   4. Modeling approach and results
   5. Conclusion and Final thoughts

**Model accuracy on test.xlsx will be discussed upon submission and meeting the above criterions.**