**Problem Statement**

Chennai House Price Prediction (Regression)  
  
ChennaiEstate is a real estate firm based in Chennai that is involved in the property business for the past 5 years. Since they are in the business for so long, they have enough data of all the real estate transactions in the city. They decided to venture into Analytics and have now started a division called “Chennai Estate Analytics” to give consumers as much information as possible about housings and the real estate market in Chennai. A home is often the largest and most expensive purchase a person makes in his or her lifetime. Ensuring real-estate owners have a trusted way to monitor the asset is incredibly important. Hence, they have hired you as a consultant to help them give insights and develop a model to accurately predict real estate prices. Based on the train dataset, you will need to develop a model that accurately predicts the real estate price in Chennai.  
  
Note - Use RMSLE (root mean log squared error) as the evaluation metric.

File Description  
Train.csv (7109 observations):  
The dataset contains information related to all real estate transactions that have taken place under the ChennaiEstate. The following are the details relating to the columns in the dataset:

PRT\_ID – The Property Transaction ID assigned by ChennaiEstate   
AREA – The property in which the real estate is located   
INT\_SQFT – The interior Sq. Ft of the property\  
DATE\_SALE – The date the property was sold   
DIST\_MAINROAD – The distance of the property to the main road   
N\_BEDROOM – The number of Bedrooms   
N\_BATHROOM - The number of bathrooms   
N\_ROOM – Total Number of Rooms   
SALE\_COND – The Sale Condition   
Normal: Normal Sale   
Abnormal: Abnormal Sale - trade, foreclosure, short sale   
AdjLand: Adjoining Land Purchase   
Family: Sale between family members   
Partial: Home was not completed when last assessed   
PARK\_FACIL – Whether parking facility is available   
DATE\_BUILD – The date in which the property was built   
BUILDTYPE – The type of building   
House   
Commercial   
Others   
UTILITY\_AVAIL   
AllPub: All public Utilities (E,G,W,& S)   
NoSewr: Electricity, Gas, and Water (Septic Tank)   
NoSeWa: Electricity and Gas Only   
ELO: Electricity only   
STREET   
Gravel   
Paved   
No Access   
MZZONE   
A: Agriculture   
C: Commercial   
I: Industrial   
RH: Residential High Density   
RL: Residential Low Density   
RM: Residential Medium Density   
QS\_ROOMS – The quality score assigned for rooms based on buyer reviews   
QS\_BATHROOM – The quality score assigned for bathroom based on buyer reviews   
QS\_BEDROOM – The quality score assigned for bedroom based on buyer reviews   
QS\_OVERALL – The Overall quality score assigned for the property   
REG\_FEE – The registration fee for the property   
COMMIS – The Commission paid to the agent   
SALES\_PRICE – The total sale price of the property