# **TECHNICAL GUIDE DOCUMENT**

# House Price Prediction Challenge

## Overview

Welcome to the House Price Prediction Challenge, we will test our regression skills by designing an algorithm to accurately predict the house prices in India. Accurately predicting house prices can be a daunting task. The buyers are just not concerned about the size(square feet) of the house and there are various other factors that play a key role to decide the price of a house/property. It can be extremely difficult to figure out the right set of attributes that are contributing to understanding the buyer's behavior as such. This dataset has been collected across various property aggregators across India. In this competition, provided the 12 influencing factors our role as a data scientist is to predict the prices as accurately as possible.

Also, in this competition, we will get a lot of room for feature engineering and mastering advanced regression techniques such as Random Forest and various other ensembling techniques.

## Data Description:

Train.csv - 29451 rows x 13 columns  
Test.csv - 68720 rows x 12 columns

## Attributes Description:

INPUT COLUMNS:

| Column | Description |
| --- | --- |
| ID | House ID |
| POSTED\_BY | Category marking who has listed the property |
| UNDER\_CONSTRUCTION | Under Construction or Not |
| RERA | Rera approved or Not |
| BHK\_NO | Number of Rooms |
| BHK\_OR\_RK | Type of property |
| SQUARE\_FT | Total area of the house in square feet |
| READY\_ TO\_MOVE | Category marking Ready to move or Not |
| RESALE | Category marking Resale or not |
| ADDRESS | Address of the property |
| LONGITUDE | Longitude of the property |
| LATITUDE | Latitude of the property |

OUTPUT COLUMN:

| Column | Description |
| --- | --- |
| TARGET(PRICE\_IN\_LACS) | Predicting house prices in lacs |

**STEPS TO BE FOLLOWED**

Importing libraries and reading the dataset.

**↓**

Describing and counting the required values.

**↓**

Data preprocessing.

**↓**

Dropping quasi constant & constant features and splitting the data.

**↓**

Performing the (Exploratory Data Analysis) part.

**↓**

Showing the relations using Bar graph,line,pie-chart,catplot,grid etc.

**↓**

Data Analysis.

**↓**

Feature engineering and feature selection.

**↓**

Applying various regression model and doing the prediction.

**↓**

Selection of Model and passing the test data through this model

**↓**

Get Final result and Final dataset

**Group-10**

- Kushagra Kumar Rai (1906484)

- Shivam Kumar (1906504)

- Souhardya Pal (1906511)

- Tapabrata Roy (1906522)