In India, buying a house is one of the most important decisions a family takes. It is considered to be a process where a lot of consultation and thinking through happens. But what are the things that a family considers before purchasing a house? The location, the property size the house type, vicinity to schools and malls, vicinity to hospitals, the neighborhood? What about the price of the house? What exactly drives the price of a house? What gives confidence to a family that the price of a house they are buying is right?

Today we will simulate a problem where you will use Machine learning to predict the monetary value of a house based on certain house features as well as geography of the house location. Such a model would prove to be invaluable for both sellers and buyers. Real estate agents, brokers or real estate companies could make use of such information on a daily basis. If they’re going to sell a house, they can accurately estimate what price tag to put on it. Buyers can also have their peace of mind if they are able to see that they’ve got their money’s worth.

**Problem Statement:**

The data provided has information about 17597 houses sold in a period of 10 months. This is a simulated data and the postal codes shared are fictional. The data mostly includes various features related to a house such as number of bedroom and bathrooms, living area, number of floors, house condition, built year etc. The data also includes some information regarding the geography around it such as postal codes, distance from the airport and number of schools in the vicinity of the house. Lastly, the train data captures information about when the house was sold and what was the final price at which it was sold.

The test data captures the same amount of information for the houses sold in the next 3 months except details about the prices at which those houses were sold. The problem of this Hack of all Trades is to predict the price at which the house was sold. Feel free to use whatever machine learning model and common sense you think will be applicable for this problem.

**What is required from you:**

Train your machine learning model and run though the test data. Upload the predicted values (house prices) on the platform against their respective ID. To reiterate, your submission file will have 4001 rows (including the header rows) and two columns ‘ID’ and ‘Price’. The column ‘ID’ is taken from the test data and ‘Price’ is what you will predict for each house corresponding to that specific ‘ID’.