**Problem Statement**

**About House sales Data**

Dataset contains house sales data of King County.

List of columns is as follows:-

* id
* date
* price
* bedrooms
* bathrooms
* sqft\_living
* sqft\_lot
* floors
* waterfront
* view
* condition
* grade
* sqft\_above
* sqft\_basement
* yr\_built
* yr\_renovted
* zipcode
* lat
* long
* sqft\_living15
* sqft\_lot15

**Problem Statement:-**

You are looking to buy a property in Kings County but you are not sure about what are the factors that affect the prices of houses. You have heard from you colleges that brokers sometimes charge more than the actual cost of the house. Also, Since this is your first house you are not aware of the factors that affect the cost of the house. So in order to make a smart decision you decided to get some insights using the dataset provided.

1. Can you predict house prices using the features and data available of King County.
2. List the top 10 Zip Codes that have the highest house prices and the lowest prices.
3. List top five features that contribute to house pricing.
4. Is there a correlation between the location and price?

**Context:-**

**Criteria of Success:-**

Building a model that can predict the price of the house.

**Scope of Solution Space:-**

Predicting the price of the house as well as finding the features that affect the price of the houses.

**Constraints within Solution Space:-**

**Stakeholders:-**

**Data Source**

.csv file.