Description: Logo

Description automatically generated

**Capstone Projects for Data Analysis**

**Feature Engineering on House Price Prediction**

House price prediction refers to the process of estimating the value or selling price of a property based on various factors such as location, size, condition, and market trends. It involves using statistical models and

help the developer determine the selling price of a house and can help the customer to arrange the right time to purchase a house. There are three factors that influence the price of a house which include physical conditions, concept and location.



Datasets is given in with this Project

In Data Analysis we will analyse to find the below questions:--

1. Missing Values
2. All the numerical Variables
3. Distribution of the numerical Variables
4. Categorical Variables
5. Outliers
6. Relationship between Dependent and Independent Features
7. Correlations between Discrete vs Sales Price && Continuous vs Sales Price

**Task 1 :-**

1. Import necessary and essential libraries
2. Display all the Columns of DataFrames
3. Read the data and display the first 100 rows from the data
4. Give the column insights

**Task 2 :-**

Q1) Checking for Missing Values

Q2) Features with NAN Values

Q3) Calculate with mean sales Price where the information is present or Missing

Q4) Gives the Count of Numerical features

Q5) Prints the first five rows of numerical values

Q6) We will Compare the difference between all the years features with SalesPrice

Q7) On the Discrete Variable Find the relationship between Discrete and Sales price

Q8) On the Continous Variable Find the relationship between Discrete and Sales Price

Q9) Analyse the Continous values by creating the histogram to understand the distribution.

Q10) Apply the Lograthmic Transformation

**Task 3 :-**

Q1) Find the Outliers

Q2) Find the relationship between Categorical feature and Sales Price

Q3) Find the Correlation between Numerical Features and Sales Price

Q4) Find Continous Features vs Sales Price

Q5) Do the Feature Engineering for the

a)Handle the Missing Values

b) Handle the Categorical Variables

c) Handle the Numerical Variables

d) Handling the Temporal Varaibles