

EXECUTIVE REPORT: PRICE PREDICTION MODEL OF HOUSES IN NIGERIA

Prepared by: ADEBESIN ARAMIDE

Date: 5th of August 2023.

INTRODUCTION:

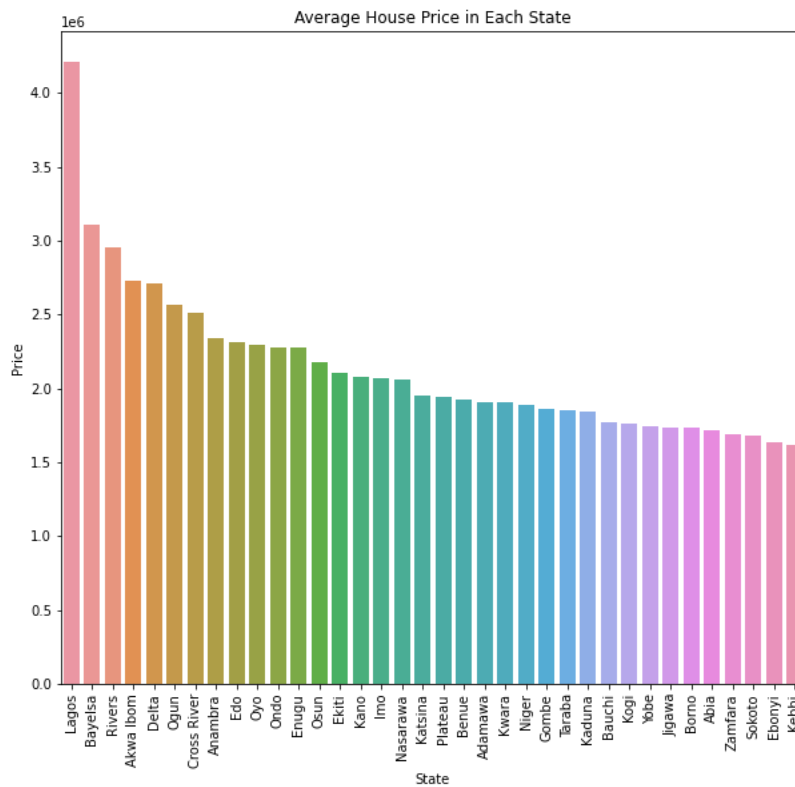
This report presents an analysis of the House Price Prediction Model, which aims to accurately estimate house prices based on features including location, house type, number of bedrooms, number of bathrooms and number of parking space. The model was developed to assist in making informed decisions for Wazobia Real Estate Limited to enhance its competitiveness in the market and deliver enhanced value to its customers.

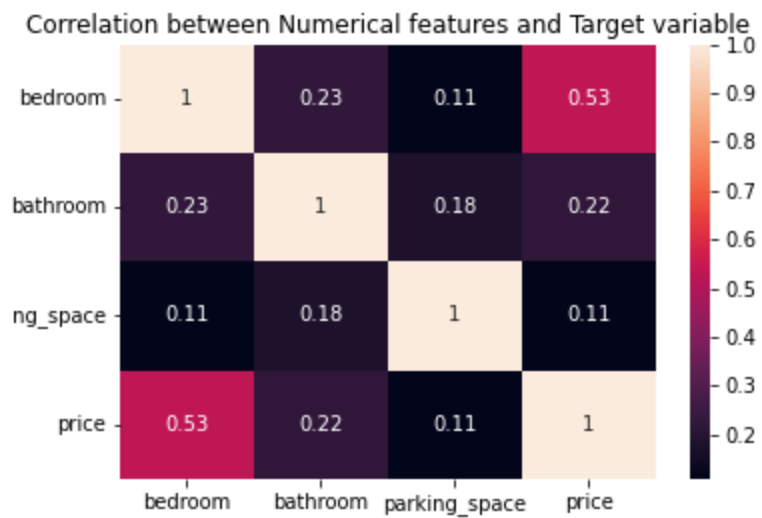
EXPLORATORY DATA ANALYSIS (EDA):

The dataset contains 14,000 rows and 6 columns. The following insights were derived from univariate and bivariate analysis

- **Location:** The location column contains the 36 states in Nigeria. Houses are distributed across different states. Certain states are associated with higher median prices. Lagos state was observed to be the state with the highest average house price. Cross river is the state with the highest number of houses.
- **House Types:** The house type column contains 10 unique house types: Semi-detached duplex, Apartment, Detached duplex, Terrace duplex, Mansion, Bungalow, Penthouse, Townhouse, Flat, Cottage. There is no significance difference in the relationship between house type and the average number of bedrooms, bathrooms and parking space, except for Cottage and this is because it is a very small house compared to the other house types. It is further observed that the most expensive house type is Mansion.
- **Bedroom:** The bedroom column contains a number of bedrooms from 1 – 9. There is a positive correlation between the number of bedrooms and house prices which means larger houses tend to command higher prices.

- Bathroom:** The bathroom column also contains a number of bedrooms from 1 – 7.
 There is a low positive correlation between the number of bathrooms and house prices.
- Parking Space:** The parking space column contains a number of bedrooms from 1 – 6.
 There is a very low positive correlation between the number of parking spaces and house prices. Houses with more parking space tend to have slightly higher prices on average.



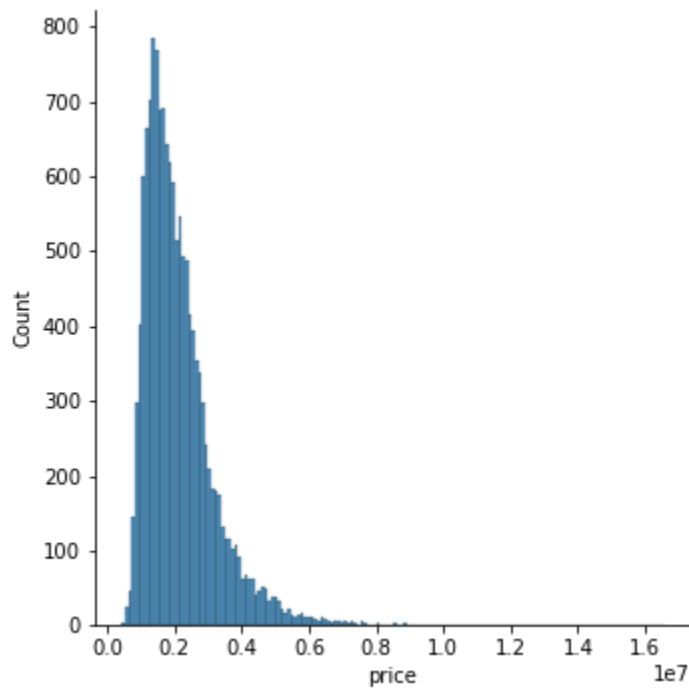


DATA PREPROCESSING:

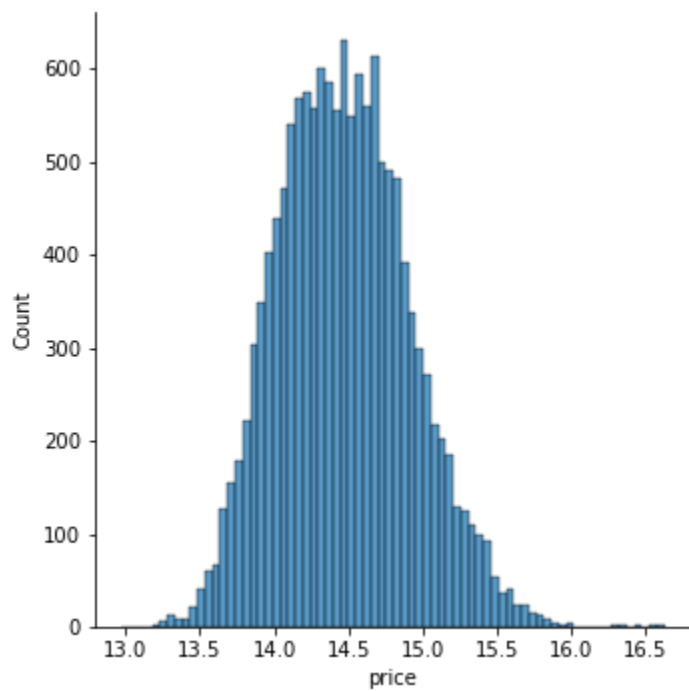
To prepare the data for modelling, the following preprocessing steps were applied:

- **Missing Values:** Missing values were observed in all of the columns except the price column. To handle it I filled the numerical columns with the mean of each column while the categorical column was filled with the mode.
- **Log Transformation:** The price column was observed to be rightly skewed, so I applied log transformation to normalize the data.

Before Applying Log Transformation



After Applying Log Transformation



- **Categorical Encoding:** Categorical features like location and house type were label encoded for compatibility with the model

MODEL OVERVIEW:

The LightGBM algorithm was chosen for its ability to capture complex relationships. Model hyperparameters were fine-tuned to optimize predictive accuracy.

MODEL EVALUATION:

Model Evaluation was assessed using the Root Mean Squared Error (RMSE) metric. The average RMSE was found to be 515430.51

BIASES AND ETHICAL CONSIDERATIONS:

- If the dataset contains houses from specific locations, the model might develop bias towards those areas.
- If the dataset is skewed towards certain type of houses, the model might not generalize well to other types.
- If the data collection process disproportionately favors certain types of houses or locations, this bias will affect the model predictions.
- Economic indicators, government policies and market trends which are external factor bias can as well affect the model.

CONCLUSION:

The House Price Prediction Model offers valuable insights for estimating house prices in Nigeria. While the model is a valuable tool for making informed decisions, it's important to acknowledge its limitations and utilize it in conjunction with domain expertise.