



Kenya House Price Estimator





CONTENT

01

Business
Understanding

02

Problem
Statement

03

Objectives

04

Data Analysis

05

Modeling

06

Conclusion

07

Participants

BUSINESS UNDERSTANDING

- As a leading data mining and statistical computation company, Chris Tech specializes in transforming raw data into valuable insights.
- Signature Realtors, recognizing our expertise, has entrusted us with a strategic assignment.
- Our collaborative venture aims to develop a sophisticated price estimation model for houses, leveraging data extracted from [jiji.com](https://www.jiji.com).



PROBLEM STATEMENT

- We recognize the importance of a comprehensive understanding of the project's context.
- Assessing the situation involves identifying the required resources, project requirements, potential risks, and conducting a cost-benefit analysis.
- Simultaneously, we've assessed the technological requirements and time commitments necessary for the successful completion of the project.
- Our risk assessment has identified potential challenges such as legal issues related to data scraping, data quality concerns, and model interpretability challenges.
- This detailed situational analysis, including a cost-benefit evaluation, ensures that our collaboration with Signature Realtors is grounded in a robust foundation.



OBJECTIVES



- Our first objective in this collaboration is to create a precise price estimation model for houses, a task that aligns seamlessly with our core competencies in data mining and statistical computation.



- Through in-depth discussions with Signature Realtors, we've established clear success criteria, emphasizing accurate price predictions, model interpretability, and seamless integration into their existing systems.



- Our commitment to understanding the business objectives sets the stage for a purposeful and effective data mining project.



DATA UNDERSTANDING

- In the initiation of the Data Understanding phase, our team diligently acquired data from jiji.com.
- The selection process ensured the inclusion of essential features crucial for the subsequent development of the price estimation model.
- By securing a comprehensive dataset, we laid the groundwork for an in-depth analysis that aligns with the project's objectives.



DATA PREPARATION

Handling
missing and
extreme
values



Removing
duplicate
entries



Standardizing
data formats



MODELING AND REGRESSION ANALYSIS

We developed four regression models with the data given and finally settled on a gradient boosting model.

We used the accuracy score of 86% and the RMSE of 0.0172 as the metrics of settling on the regression model.



RECOMENDATIONS

- 01** Improving house grades and conditions right from construction with better designs drives prices higher.
- 02** investing in high-end neighborhoods brings in good profits.



RECOMENDATIONS

- 03** The Signature relator should consider increasing the number of bathrooms as this drives the prices higher.
- 04** The Signature relator should consider building houses with more than four floors to have a good return on investment.



FUTURE PROPOSAL

We would propose to be provided with more data from other real estate agencies in order to have a robust model with improved model accuracy.

