# Project Proposal: HDB Resale Price Forecasting & Profitability Analysis

## Team Members (Group 10)

1. Daniel Yuen 1006380
2. Ng Xue Min 1006127
3. Lim Jie Han 1006246

## Topic & Problem Statement

Prospective HDB buyers often seek insights into the potential returns of their property investment when selling after the Minimum Occupation Period (MOP). This project aims to build a predictive model that forecasts the resale price of an (Resale type) HDB flats upon reaching its MOP (5 years), helping buyers make informed decisions about their purchase. By leveraging historical resale data, the model will suggest which town and HDB classification (Mature or Non Mature) will likely yield the highest profit.

## Expected Inputs & Outputs

**Inputs**

* Year of intended (Resale type) HDB purchase
* Budget
* **Town**

**Features:**

* Town
* Flat Type (e.g., 3 Room, 4 Room, Executive)
* Storey Range (Level 1 to n)
* Floor Area (sqm)
* Lease Commencement Date­
* Remaining Lease (years)
* HDB Classification (Mature or Non Mature)
* Resale Price

**Outputs**

1. Identification of the most profitable street ( located in user input: Town %) when sold after 5,6,7 years after (user input: year of intended HDB purchase)
2. Forecasted resale price of streets (located in user input:Town) at (user input: year of intended HDB purchase)
3. Estimated profit based on point 2 + 5 years

## Dataset

We will use two primary datasets sourced from official HDB records:

**Resale Flat Prices (2017-2025)**

1. Size: 200.6K rows
2. Features:
   1. Month
   2. Town
   3. Flat Type
   4. Block
   5. Street Name
   6. Storey Range
   7. Floor Area (sqm)
   8. Flat Model
   9. Lease Commencement Date
   10. Remaining Lease
   11. Resale Price

**Median Resale Prices by Town and Flat Type (2007-2024)**

1. Size: 11.1K rows
2. Features:
   1. Quarter
   2. Town
   3. Flat Type
   4. Median Price

## Architecture

We will implement a deep learning-based time series forecasting model:

Primary Model:

1. LSTM (Long Short-Term Memory) network

Alternative Models for Evaluation:

1. GRU (Gated Recurrent Unit)
2. Transformer-based models

Evaluation Metrics:

1. RMSE (Root Mean Squared Error)
2. MAPE (Mean Absolute Percentage Error)
3. R-squared (R²) for model performance assessment

## Deliverables

We will submit the following components:

**Codebase**

1. Code for training the model from scratch
2. Code for recreating the trained model from a saved file

**Performance Analysis**

1. Metrics such as RMSE, MAPE, and R-squared
2. Visualization of performance curves

**Final Report (PDF)**

1. Project background and objectives
2. Data preprocessing steps
3. Model architecture and training methodology
4. Experimental results and analysis
5. Conclusion and future improvements

A document with text on it

AI-generated content may be incorrect.