# HDB Resale Price Analysis Report

## 1. Project Background

**Overview**

I aimed to analyse and forecast the resale prices of HDB (Housing Development Board) flats in Singapore. I have utilised historical data from 1990 to 2020.

The aim is to to predict future resale prices up to 2030. Ultimately, the goal is to offer married couples actionable insights for informed decision-making and strategic planning in their property investments.

**Business Objectives**

I aimed to use predictive modelling, to develop 2 time-series forecasting model that predicts future resale prices, that would help potential homebuyers and investors make informed decisions.

**Forecasting:**

Can we predict future resale prices using historical trends and patterns?

## Work Accomplished

**Data Preparation**

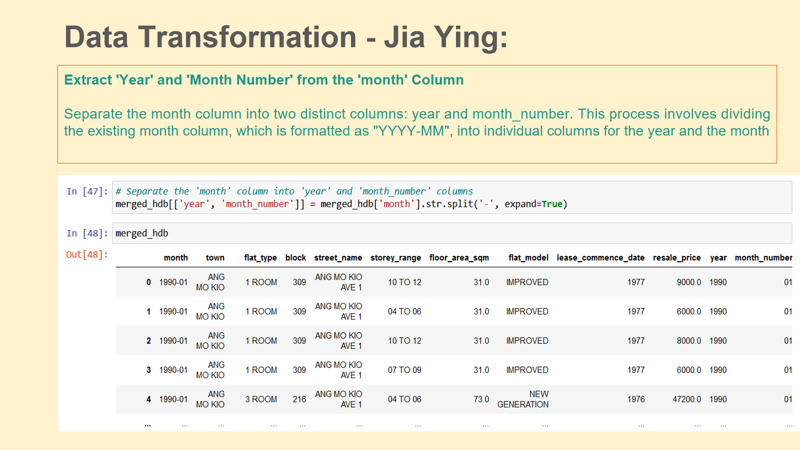
Historical HDB resale data from 1990 to 2020 was collected and consolidated into a single DataFrame.

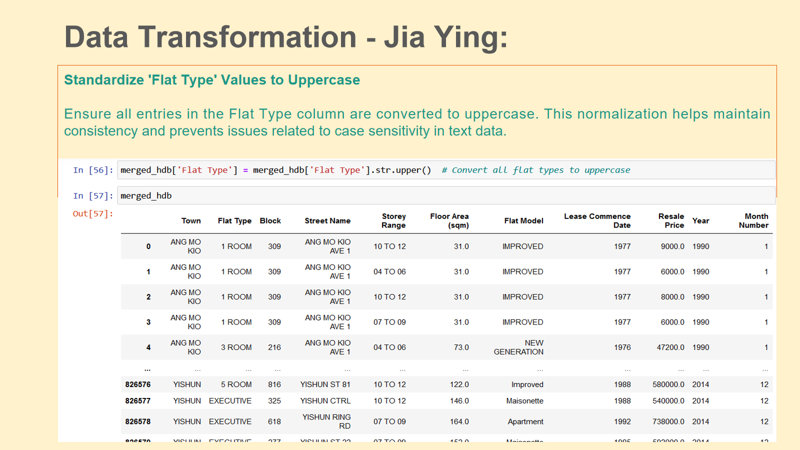
**Data Cleaning**

* Removed unnecessary files and handled missing values.
* Standardized column names and ensured consistency by converting text to uppercase.

**Feature Engineering**

* Extracted year and month from the 'month' column.
* Calculated 'Flat Age' as the difference between the current year and the lease commencement date.
* Renamed columns to align with forecasting model requirements.

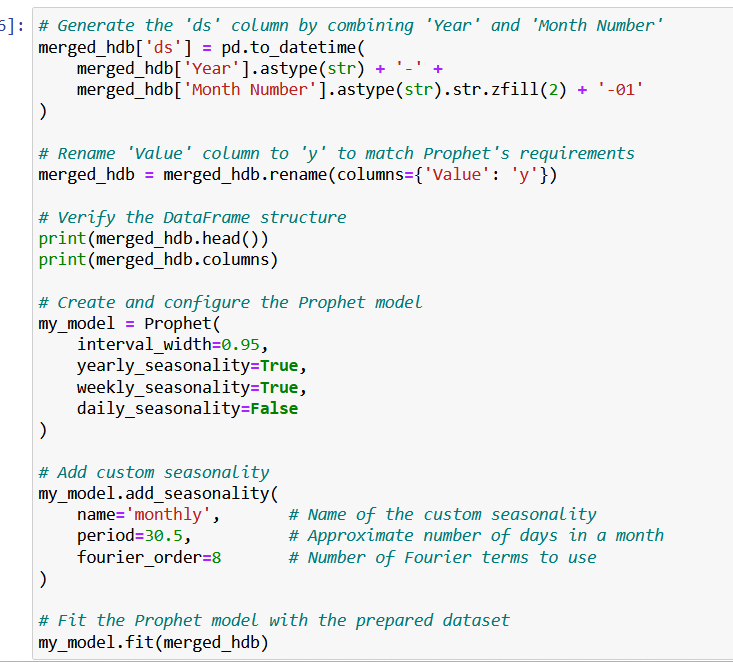


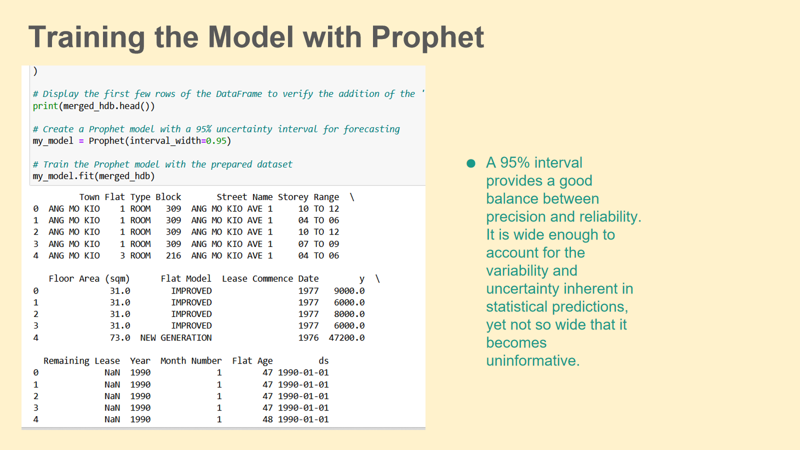


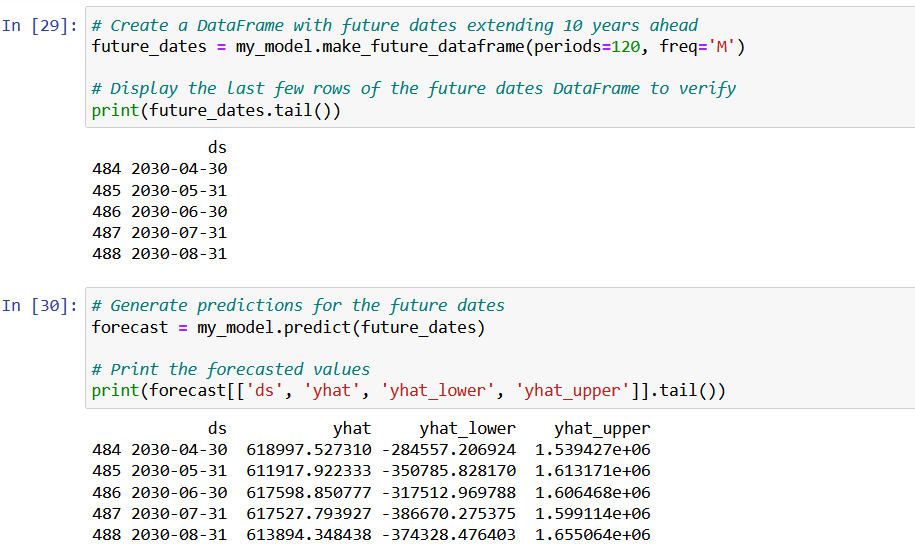
**Models**

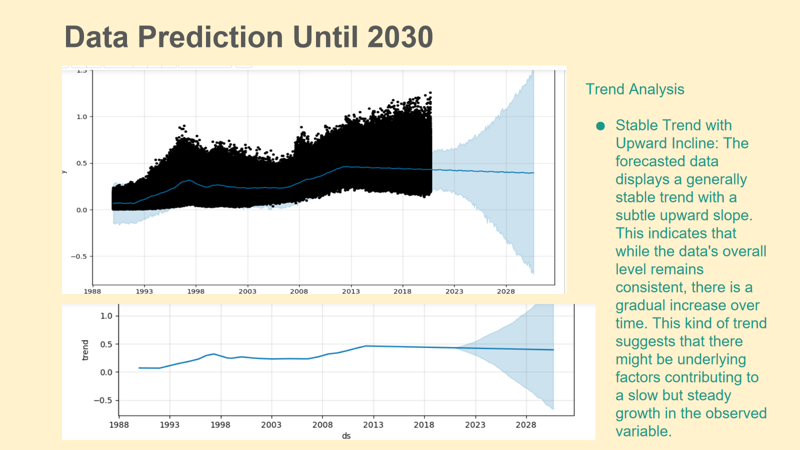
*Prophet Model*

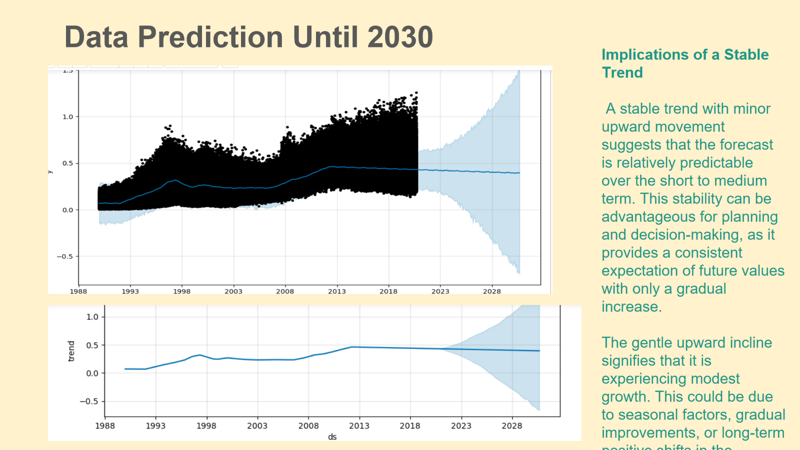
* Added seasonality
* Prepared the dataset for time series forecasting using the Prophet library, which requires columns named 'ds' (date) and 'y' (target variable). In this case, 'year' was used as the date and 'Resale Price' as the target variable.





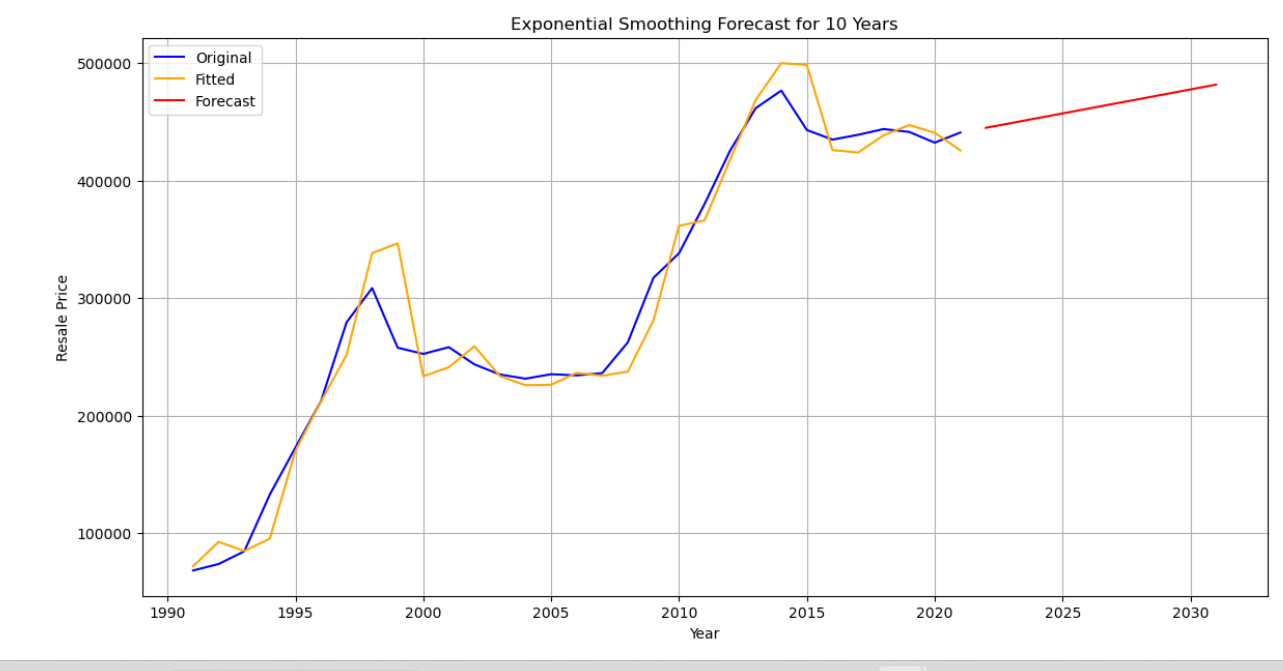


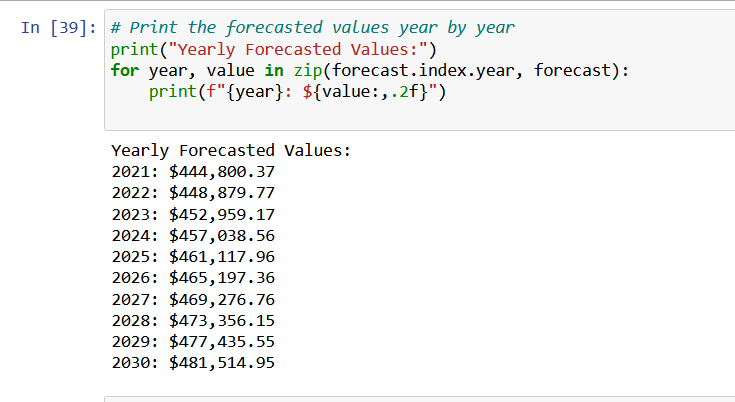




**Exponential Smoothing Model**

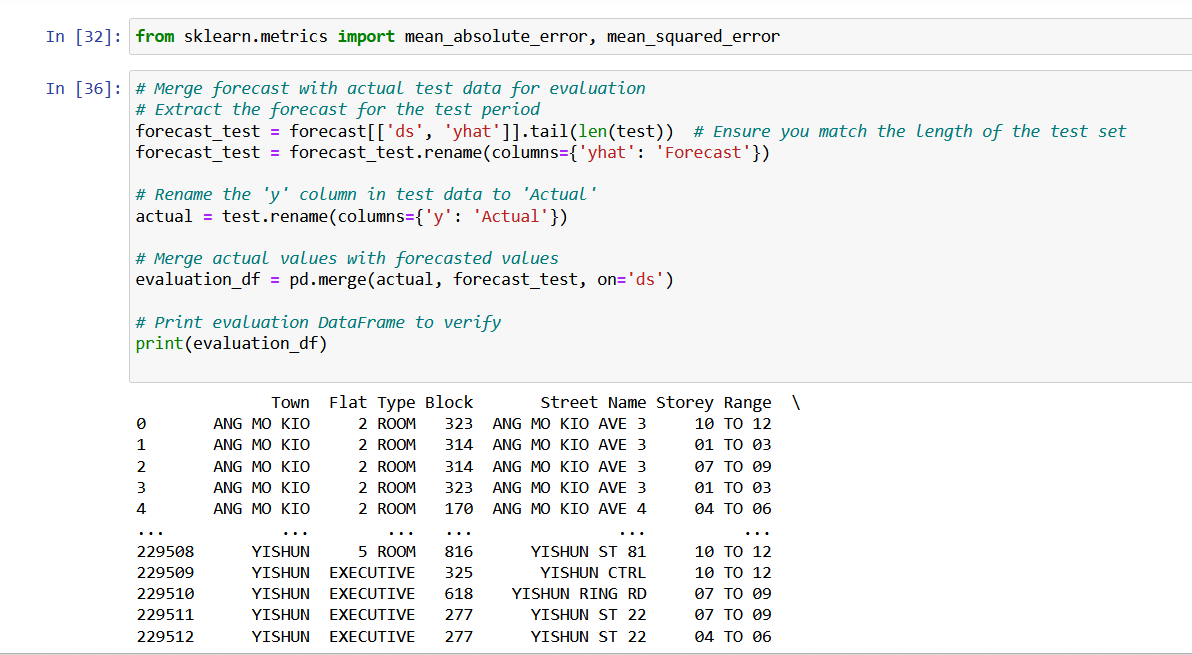


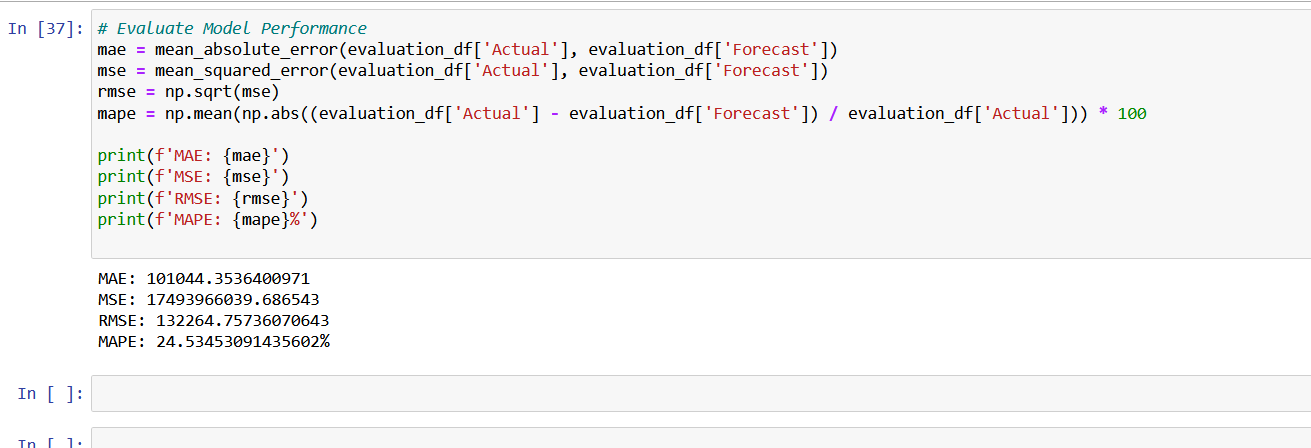




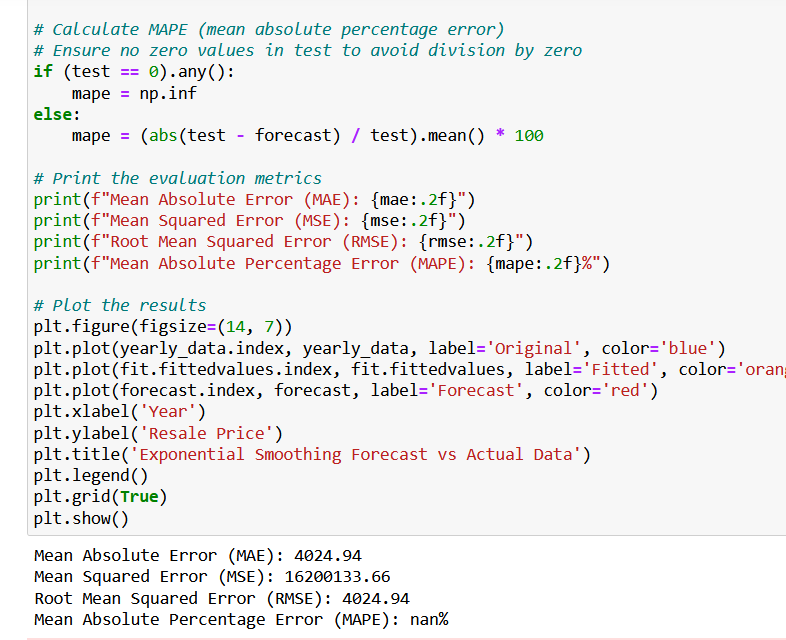
**Evaluation for Prophet**

(using code from my other file as my original file was not loading)





**Evaluation for Exponential Smoothing**



### 3. Recommendation and Analysis

**Comparing Models**

**Smoothing Model:**

MAE: 4024.94

MSE: 16200133.66

RMSE: 4024.94

MAPE: nan% (not available)

**Prophet Model:**

MAE: 101044

MSE: 17493966039

RMSE: 132264

MAPE: 24.53%

**Analysis**

**MAE**

The smoothing model has a much lower MAE (4024.94) compared to the Prophet model (101044). This means that the smoothing model's predictions are closer to the actual values.

**MSE and RMSE**

Both MSE and RMSE for the smoothing model are lower than those for the Prophet model. This suggests that the smoothing model has smaller prediction errors overall and less variance in the magnitude of these errors compared to the Prophet model.

**Recommendations**

Given that the smoothing model has consistently lower MAE, MSE, and RMSE, the smoothing model is more accurate and married couples should use it to predict future prices. The Prophet model has larger variance.

**4. AI Ethics**

**Privacy:**

1. Fairness

Assessment Based on Report:

My model evaluations (MAE, MSE, RMSE) do not provide direct insights into fairness, but high error metrics in Prophet (e.g., MAE: 101044) may imply potential issues with model performance which could impact fairness if the model is biased.

I can improve by making sure that the data used is representative of all relevant demographic groups to avoid unintentional bias. I can continuously monitor the model’s performance to ensure fairness over time.

2. Transparency

Assessment Based on Report:

While I have detailed error metrics, my report lacks information on model explainability and transparency of the decision-making process. I can incorporate methods to explain the predictions of your model, such as feature importance or SHAP values. I can provide detailed documentation on how the model works, including the data sources and rationale behind model selection and training.

3. Accountability

I can establish clear processes for monitoring and improving model performance and handling issues.

Assessment Based on Report:

The evaluation metrics do not directly address data privacy, but if the model uses personal data, privacy considerations are crucial. I can ensure compliance with data protection laws, such as anonymizing personal data and securing data storage. I can conduct assessments to understand the impact of data usage on privacy and implement necessary safeguards.I can stay updated on global best practices and incorporate them into your project where applicable.

**Conclusion**

The project successfully developed a time-series forecasting model to predict HDB resale prices up to 2030. Insights derived from this model provide valuable information for homebuyers, investors, and policymakers. The ethical considerations were addressed to ensure responsible data use and modeling practices. This analysis supports strategic decision-making and planning for property investments, particularly for married couples aiming to make informed choices in the real estate market.