A Nowcasting Approach to Singapore's Property Price Index

1 Issues

Forecasting Singapore's property price index presents significant challenges due to the intricate relationships among a broad array of variables, including property resale indices, construction material costs, and economic and financial indicators. The complexity is amplified by the delayed availability of critical data, usually lagging by one to three months, which poses substantial hurdles to achieving precise and timely predictions of private housing prices.

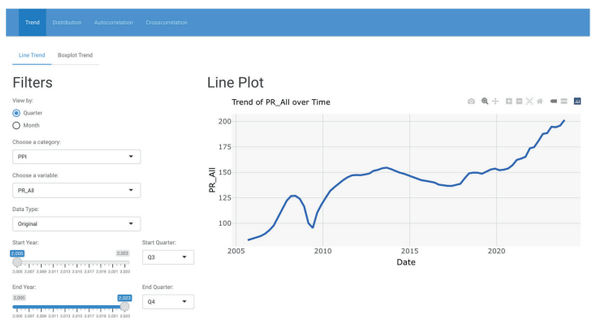
2 Motivation

The motivation to improve forecasts of Singapore's property price index stems from its significant influence on economic planning and decision-making. Precise predictions enable timely policy interventions, informed investment decisions, and optimized buying and selling strategies.

3 Approach

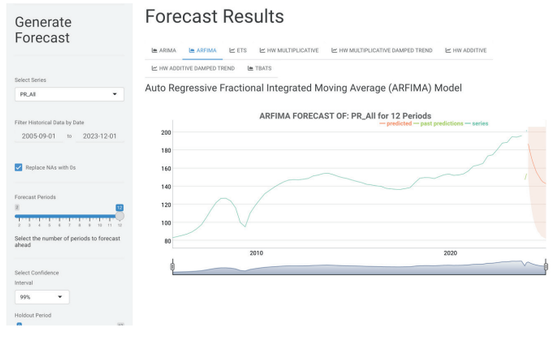
Our strategy encompasses Exploratory Data Analysis (EDA), conventional time series forecasting, and cutting-edge Machine Learning (ML) models for nowcasting, leveraging an extensive dataset that reflects the property market's complexity. We meticulously analyze Property Resale Indices, their Percentage Changes, Construction Material Costs, Economic and Financial Indicators (including Bonds), and Prices Per Square Foot (PSF) to provide a nuanced understanding and accurate forecasts.

4 Exploratory Data Analysis



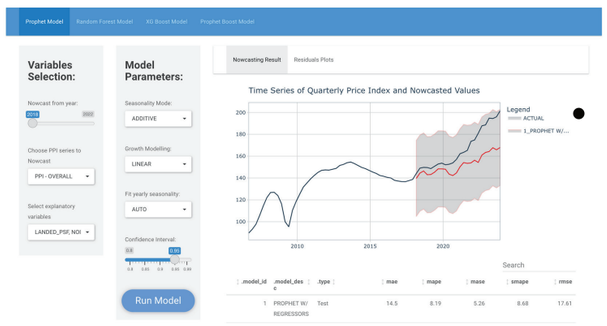
- **Trend Analysis:**
  - Options to view line trends or boxplot trends.
  - Filters for time frequency (monthly or quarterly data).
  - Category and variable selection.
  - Data type options (original values or changes).
  - Date range selection (from start year and quarter to end year and quarter).
- **Distribution Analysis:**
  - Histograms to visualize variable distributions.
  - Filterable by category and variable.
  - Adjustable date range for targeted analysis.
- **Autocorrelation Analysis:**
  - Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots.
  - Category and variable selection, with data transformation options (none, log, square root).
  - Customizable lag settings for detailed lag analysis.
  - Confidence level adjustment for statistical significance.
  - Differencing order to identify data stationarity.
- **Crosscorrelation Analysis:**
  - Correlation Matrix to assess relationships between multiple variables.
  - Variable selection for comparative analysis.
  - Correlation type choices (Pearson, Spearman, Kendall).
  - Data transformation options to normalize or adjust variable scales.

5 Traditional Time Series Forecasting



- **Forecasting Models**:
  - ARIMA (Autoregressive Integrated Moving Average).
  - ARFIMA (Autoregressive Fractionally Integrated Moving Average).
  - ETS (Exponential Smoothing State Space Model).
  - HW Multiplicative (Holt-Winters Multiplicative).
  - HW Multiplicative Damped Trend.
  - HW Additive (Holt-Winters Additive).
  - HW Additive Damped Trend.
  - TBATS (Trigonometric, Box-Cox transformation, ARMA errors, Trend and Seasonal components).
- **User Customization**:
  - Variable filtering for targeted forecasting.
  - Start and end date selection for time range specificity.
  - Option to replace missing values (NAs) with zeros for continuity.
  - Forecast period adjustment (ranging from 2 to 12 periods).
  - Confidence interval settings for forecast precision.
  - Holdout period configuration (ranging from 2 to 12 periods) for model validation.

6 ML Models Nowcasting



- **Prophet Model:**
  - Year selection for the start of nowcasting.
  - Choice of PPI series variables for nowcasting.
  - Option to select explanatory variables.
  - Adjustable model parameters:
    - Seasonality mode (additive or multiplicative).
    - Growth model (linear or logistic).
    - Yearly seasonality (auto-detect, enable, or disable).
- Outputs include nowcasting results and residual plots.
- **Random Forest Model:**
  - Year selection for the start of nowcasting.
  - Choice of PPI series variables for nowcasting.
  - Option to select explanatory variables.
  - Customizable model parameters:
    - Number of trees in the forest.
    - Number of variables considered for splitting at each leaf node.
    - Maximum size of terminal nodes.
    - Confidence interval settings.
  - Outputs include nowcasting results, variable importance, and residual plots.
- **XGBoost Model:**
  - Year selection for the start of nowcasting.
  - Choice of PPI series variables for nowcasting.
  - Option to select explanatory variables.
  - Tunable model parameters:
    - Number of gradient boosted trees.
    - Learning rate for weight updates.
    - Minimum observations required in a child.
    - Fraction of observations to be sampled for each tree.
    - Number of variables to consider when looking for the best split.
    - Maximum depth of the tree.
    - Confidence interval settings.
  - Outputs include nowcasting results and residual plots.
- **Prophet Boost Model:**
  - Year selection for the start of nowcasting.
  - Choice of PPI series variables for nowcasting.
  - Option to select explanatory variables.
  - Combination of Prophet and XGBoost model parameters as listed above.
  - Outputs include nowcasting results and residual plots.

7 Future Work

Real-Time Data Integration:
Boosting nowcasting accuracy by integrating real-time data sources to reflect immediate market shifts.