

A Guide On How to Use UrbanScope: A Shiny App

1. EDA

UrbanScope

EDA ▾

Explanatory Model ▾

Predictive Model ▾

Data Table

This tab enables the generation of histograms, choropleth maps, scatterplots as well as locations of interest.

1.1 Histograms

Continuous Categorical

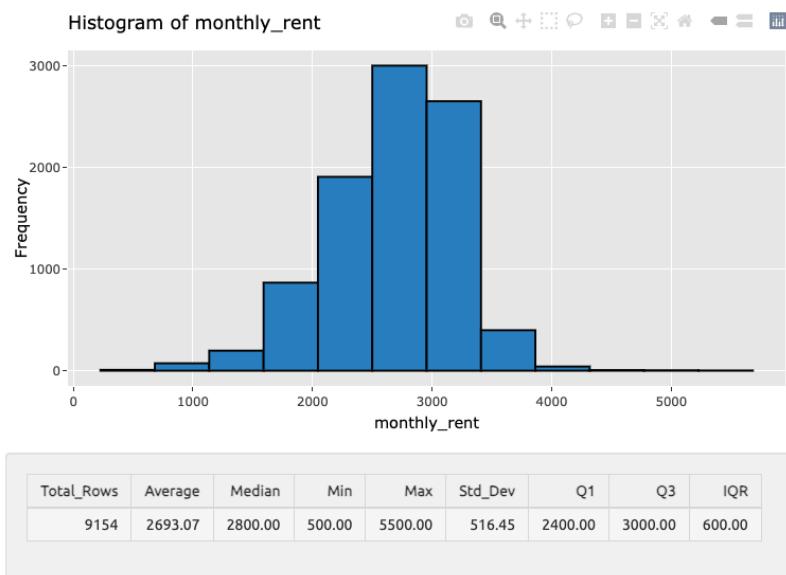
Select Plot Variable
monthly_rent

Month
2024 Jan to Sep

Select Flat Type
 3-ROOM 4-ROOM 5-ROOM

Town
All of Singapore

Number of Bins
5 7 9 11 12 13 15 17 19 20



For this, users can choose between categorical and continuous variables to plot for the histogram.

1.1.1 Continuous

Continuous Categorical

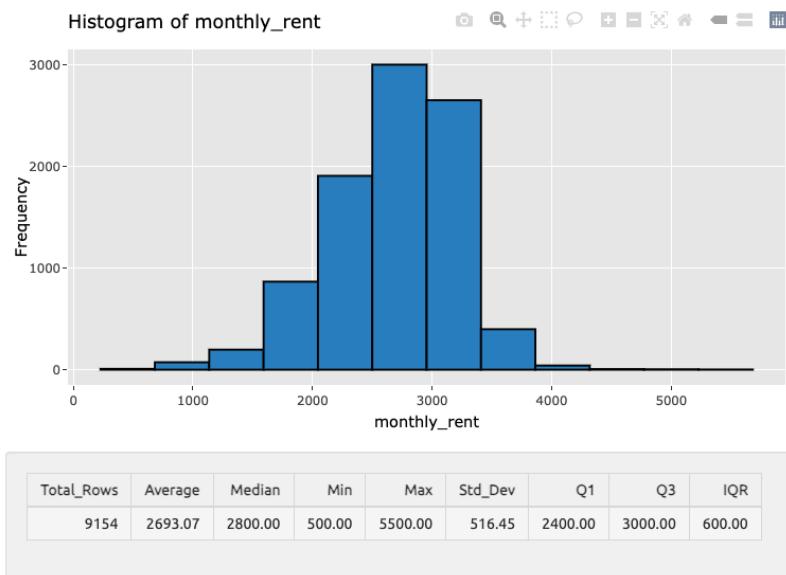
Select Plot Variable
monthly_rent

Month
2024 Jan to Sep

Select Flat Type
 3-ROOM 4-ROOM 5-ROOM

Town
All of Singapore

Number of Bins
5 7 9 11 12 13 15 17 19 20



1.1.1.1 Configuration Options

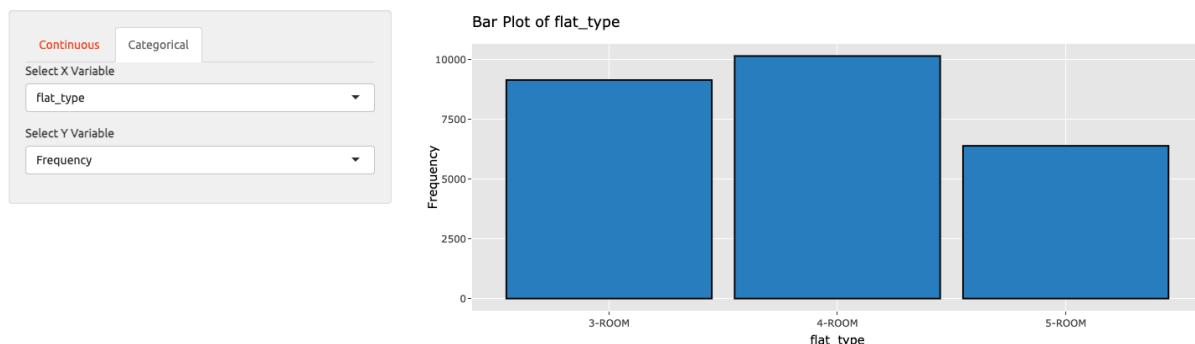
| Configuration Option | Choices | Description |
|----------------------|---|-------------------------------------|
| Plot Variable | [‘monthly_rent’], number of facilities within a set distance and proximity of nearest facilities (‘facilities’ covers different options like kindergartens, primary school etc etc) | Variable to plot for the histogram |
| Month | All months from Jan to Sep 2024 or any months from this range | Filter by time range |
| Flat Type | [‘3-ROOM’, ‘4-ROOM’, ‘5-ROOM’] | Filter by flat type |
| Town | All of Singapore or any one town | Filter by town |
| Number of Bins | Range from 5 to 20 | Choose number of bins for histogram |

1.1.1.2 General Summary of Data

| Total Rows | Average | Median | Min | Max | Std_Dev | Q1 | Q3 | IQR |
|------------|---------|---------|--------|---------|---------|---------|---------|--------|
| 9154 | 2693.07 | 2800.00 | 500.00 | 5500.00 | 516.45 | 2400.00 | 3000.00 | 600.00 |

This table shows a general summary of the statistics based on the filters and configuration options selected.

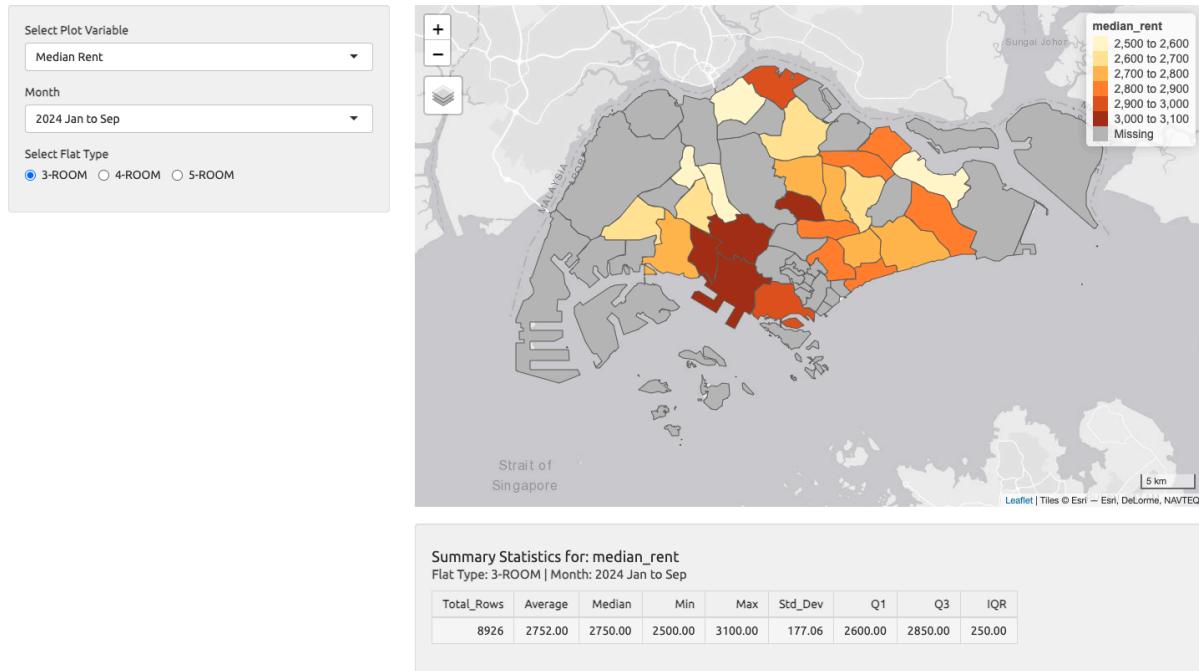
1.1.2 Categorical



1.1.2.1 Configuration Options

| Configuration Option | Choices | Description |
|----------------------|--------------------------|--------------------------------|
| Plot Variable X | [‘flat_type’, ‘region’] | Variable to plot on the x axis |
| Plot Variable Y | [frequency, median rent] | Variable to plot on the y axis |

1.2 Choropleth Map



This plot shows the distribution of rent in various towns in singapore and allows for filters based on month and flat type.

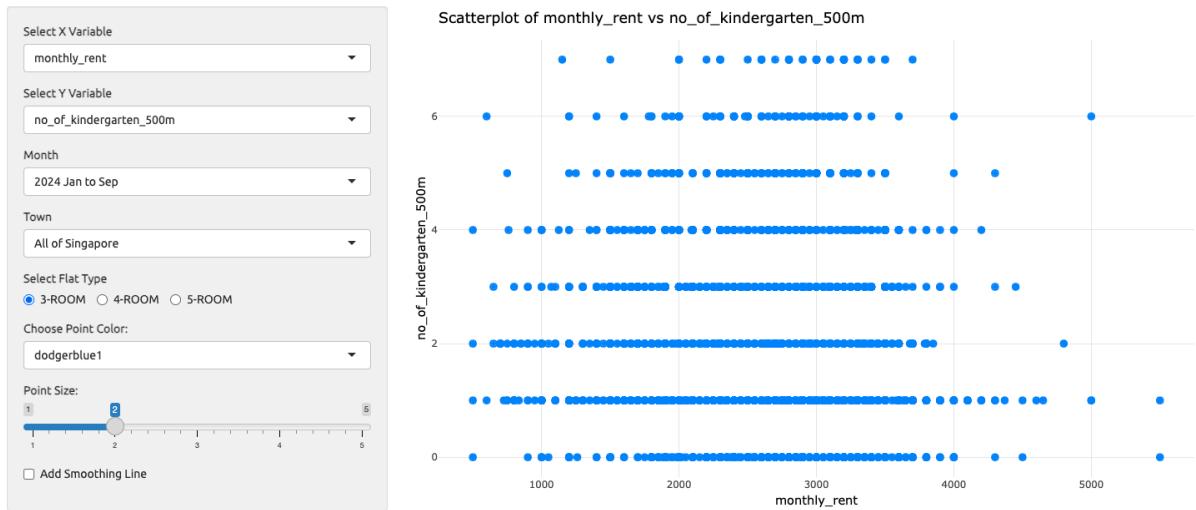
1.2.1 Configuration Options

| Configuration Option | Choices | Description |
|----------------------|---|--------------------------------|
| Plot Variable | [Median Rent, Monthly Rent] | Variable to plot on the x axis |
| Month | All months from Jan to Sep 2024 or any months from this range | Filter By Month |
| Flat Type | ['3-ROOM', '4-ROOM', '5-ROOM'] | Filter by flat type |

1.2.2 Summary Statistics

| Summary Statistics for: median_rent | | | | | | | | | |
|--|---------|---------|---------|---------|---------|---------|---------|--------|--|
| Flat Type: 3-ROOM Month: 2024 Jan to Sep | | | | | | | | | |
| Total_Rows | Average | Median | Min | Max | Std_Dev | Q1 | Q3 | IQR | |
| 8926 | 2752.00 | 2750.00 | 2500.00 | 3100.00 | 177.06 | 2600.00 | 2850.00 | 250.00 | |

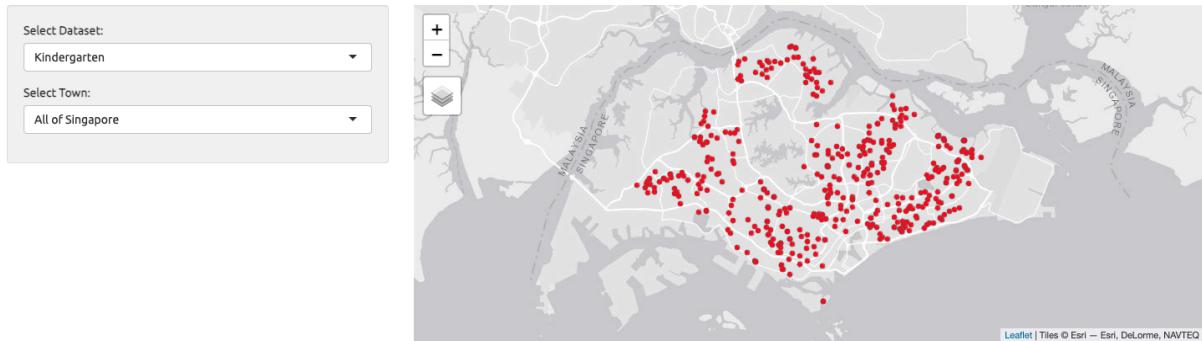
1.3 Scatterplot



This plot allows the users to plot the relationships between the different variables.

| Configuration Option | Choices | Description |
|----------------------|---|---|
| Plot Variable X | [‘monthly_rent’], number of facilities within a set distance and proximity of nearest facilities (‘facilities’ covers different options like kindergartens, primary school etc etc) | Variable to plot on the x axis |
| Plot Variable Y | [‘monthly_rent’], number of facilities within a set distance and proximity of nearest facilities (‘facilities’ covers different options like kindergartens, primary school etc etc) | Variable to plot on the y axis |
| Month | All months from Jan to Sep 2024 or any months from this range | Filter by time range |
| Flat Type | [‘3-ROOM’, ‘4-ROOM’, ‘5-ROOM’] | Filter by flat type |
| Town | All of Singapore or any one town | Filter by town |
| Point color | Range of colors | Select a color for the scatterplot points |
| Point Size | Range of 1 to 5 | Select size for the scatterplot points |
| Smoothing Line | Add/ Don’t Add | Choose whether to add the smoothing line that will be plotted in accordance to the points |

1.4 Locations of Interest



This map allows the users to visualise the locations of interest.

| Configuration Option | Choices | Description |
|----------------------|---|-----------------------------|
| Dataset | Datasets like kindergartens, primary school, hawker centres etc | Variable to plot on the map |
| Town | All of Singapore or any one town | Filter by town |

2. Explanatory Model

UrbanScope

EDA ▾

Explanatory Model ▾

Predictive Model ▾

Data Table

2.1 Correlation Matrix

Correlation Plot Type

 corrplot
 ggcorrmat

Method

Circle

Matrix Type

Full

Order

Original

Select Independent Columns

no_of_kindergarten_500m

prox_kindergarten

no_of_childcare_500m

prox_childcare

no_of_hawker_500m

prox_hawker

no_of_busstop_500m

prox_busstop

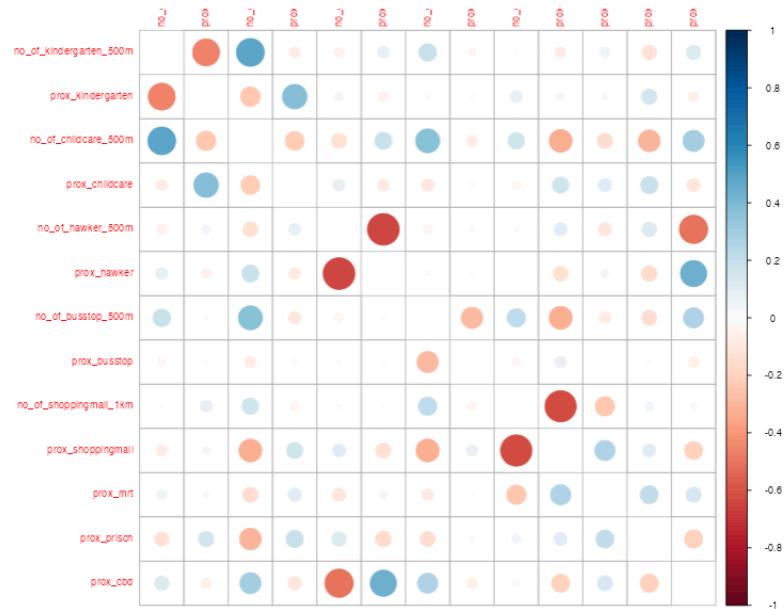
no_of_shoppingmall_1km

prox_shoppingmall

prox_mrt

prox_prisch

prox_cbd



Correlation Plot Type

 corrplot

 ggcorrmat

Matrix Type

Full

Select Independent Columns

no_of_kindergarten_500m

prox_kindergarten

no_of_childcare_500m

prox_childcare

no_of_hawker_500m

prox_hawker

no_of_busstop_500m

prox_busstop

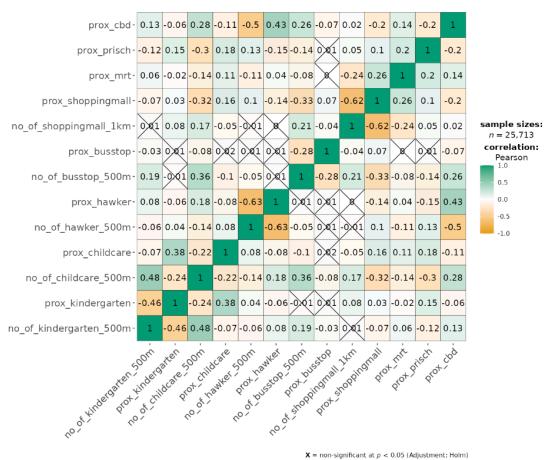
no_of_shoppingmall_1km

prox_shoppingmall

prox_mrt

prox_prisch

prox_cbd



| Configuration Option | Choices |
|-----------------------|---------------------------|
| Correlation Plot Type | ['corrplot', 'ggcorrmat'] |

Instructions

1) User select **Correlation Plot Type** (corrplot package or ggcormat)

Method

Circle

Circle

Square

Ellipse

Number

Pie

Shade

Color

Corrplot: Users are able to select the **Method**, **Matrix Type** and **Order**.

- If **Hierarchical Clustering Order** are selected, Hierarchical Clustering Method will be displayed.

Matrix Type

Full

Full

Upper

Lower

Select Independent Columns

Order

Original

Original

Angular Order of the Eigenvectors Order (AOE)

First Principal Component Order (FPC)

Hierarchical Clustering Order

Alphabetical Order

Order

Hierarchical Clustering Order

Hierarchical Clustering Method

Ward's Method (D)

Ward's Method (D2)

Single Linkage

Complete Linkage

Average Linkage

McQuitty Method

Median Method

Centroid Method

Matrix Type

Full

Upper

Lower

no_of_children_500m

Ggorrmat: Users are able to select **Matrix Type**.

- 2) Users then can select the independent columns to be shown in the correlation plot.

Select Independent Columns

- no_of_kindergarten_500m
- prox_kindergarten
- no_of_childcare_500m
- prox_childcare
- no_of_hawker_500m
- prox_hawker
- no_of_busstop_500m
- prox_busstop
- no_of_shoppingmall_1km
- prox_shoppingmall
- prox_mrt
- prox_prisch
- prox_cbd

Users can multiselect these.

2.2 Multiple Linear Regression

Dataset Selection

3-Room

Date Range (latest months)

2

Dependent Variable

Monthly Rent

Model Info
Publication Quality Table
Regression Diagnostics (olsrr Package)

Please build the MLR model first.

- Select Independent Variables
- no_of_kindergarten_500m
 - prox_kindergarten
 - no_of_childcare_500m
 - prox_childdcare
 - no_of_hawker_500m
 - prox_hawker
 - no_of_busstop_500m
 - prox_busstop
 - no_of_shoppingmall_1km
 - prox_shoppingmall
 - prox_mrt
 - prox_prisch
 - prox_cbd

Submit

| Configuration Option | Choices |
|-----------------------|---|
| Dataset Selection | ['3-ROOM', '4-ROOM', '5-ROOM'] |
| Date Range | [1-9], numbers refers to latest x months, our dataset is from Jan - Sep 2024 |
| Dependent Variable | ['Monthly Rent'] |
| Independent Variables | Number of facilities within a set distance and Proximity of nearest facilities ('facilities' covers different options like kindergartens, primary school etc etc) |

Instructions

Dataset Selection

3-Room

Date Range (latest months)

1

Dependent Variable

Monthly Rent

Select Independent Variables

no_of_kindergarten_500m
 prox_kindergarten
 no_of_childcare_500m
 prox_childcare
 no_of_hawker_500m
 prox_hawker
 no_of_busstop_500m
 prox_busstop
 no_of_shoppingmall_1km
 prox_shoppingmall
 prox_mrt
 prox_prisch
 prox_cbd

Submit

- 1) User select parameters
 - a) Users will be able to filter the Dataset based on their preference, based on Room Type and the Date Range. (**Note:** since our dataset was from Jan 2024 - Sep 2024, Date Range = 1 means Sep 2024, 2 means Aug 2024 - Sep 2024)
- 2) Users then can select the independent columns to be used to build the MLR model by pressing the **Submit** button.

- 3) After model is built, users can select different tabs to view the model details

Model Info Publication Quality Table Regression Diagnostics (olsrr Package)

Model Summary

| R | 0.219 | RMSE | 486.268 |
|----------------|---------|-----------|------------|
| R-Squared | 0.048 | MSE | 240293.592 |
| Adj. R-Squared | 0.034 | Coef. Var | 18.104 |
| Pred R-Squared | 0.016 | AIC | 13339.944 |
| MAE | 357.896 | SBC | 13411.557 |

RMSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error
AIC: Akaike Information Criteria
SBC: Schwarz Bayesian Criteria

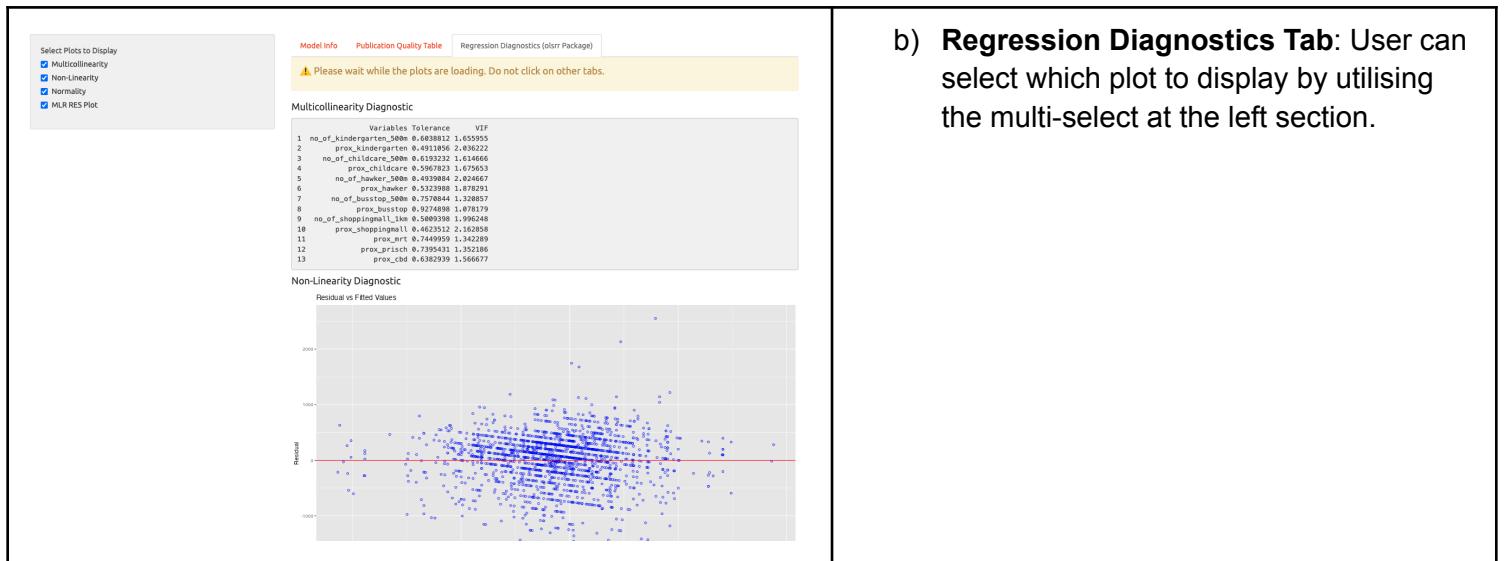
ANOVA

| | Sum of Squares | DF | Mean Square | F | Sig. |
|------------|----------------|-----|-------------|-------|-------|
| Regression | 10452772.933 | 13 | 804059.456 | 3.346 | 1e-04 |
| Residual | 206892782.618 | 861 | 246293.592 | | |
| Total | 217345555.543 | 874 | | | |

Parameter Estimates

| model | Beta | Std. Error | Std. Beta | t | Sig. | lower | upper |
|-------------------------|----------|------------|-----------|--------|-------|----------|----------|
| (Intercept) | 2812.397 | 137.891 | | 20.492 | 0.000 | 2541.933 | 3882.868 |
| no_of_kindergarten_500m | -0.135 | 16.574 | 0.007 | 0.153 | 0.878 | -0.995 | 35.865 |
| prox_kindergarten | 0.152 | 0.117 | 0.051 | 1.295 | 0.196 | -0.078 | 0.383 |
| no_of_childcare_500m | 0.177 | 7.540 | 0.052 | 1.217 | 0.224 | -5.622 | 23.977 |
| prox_childcare | -0.012 | 0.174 | -0.003 | -0.071 | 0.943 | -0.353 | 0.328 |
| no_of_hawker_500m | -20.239 | 25.115 | -0.038 | -0.806 | 0.421 | -69.533 | 29.855 |
| prox_hawker | 0.068 | 0.058 | 0.047 | 1.022 | 0.307 | -0.055 | 0.175 |
| no_of_busstop_500m | 1.163 | 4.705 | 0.009 | 0.247 | 0.885 | -8.071 | 10.397 |
| prox_busstop | -0.220 | 0.299 | -0.025 | -0.736 | 0.462 | -0.887 | 0.367 |
| no_of_shoppingmall_1km | 16.807 | 14.724 | 0.012 | 1.134 | 0.278 | -12.409 | 45.864 |
| prox_shoppingmall | 0.123 | 0.059 | 0.010 | 0.391 | 0.696 | -0.893 | 0.39 |
| prox_mrt | -0.144 | 0.059 | -0.093 | -2.425 | 0.016 | -0.261 | 0.427 |
| prox_prisch | 0.029 | 0.063 | 0.018 | 0.463 | 0.644 | -0.895 | 0.153 |
| prox_cbd | -0.023 | 0.005 | -0.108 | -4.499 | 0.000 | -0.033 | -0.013 |

a) Publication Quality Table Tab



b) **Regression Diagnostics Tab:** User can select which plot to display by utilising the multi-select at the left section.

2.3 Multiple Linear Regression - Stepwise

| Configuration Option | Choices |
|--|---|
| Dataset Selection | ['3-ROOM', '4-ROOM', '5-ROOM'] |
| Date Range | [1-9], numbers refers to latest x months, our dataset is from Jan - Sep 2024 |
| Dependent Variable | ['Monthly Rent'] |
| Independent Variables | Number of facilities within a set distance and Proximity of nearest facilities ('facilities' covers different options like kindergartens, primary school etc etc) |
| P-value threshold for Stepwise Selection | Range of 0.01 to 1 |
| Stepwise Method | ['Forward', 'Backward', 'Both'] |

Instructions

Dataset Selection

3-Room

Date Range (latest months)

1

Dependent Variable

Monthly Rent

Select Independent Variables

- no_of_kindergarten_500m
- prox_kindergarten
- no_of_chilcare_500m
- prox_chilcare
- no_of_hawker_500m
- prox_hawker
- no_of_buststop_500m
- prox_buststop
- no_of_shoppingmall_1km
- prox_shoppingmall
- prox_mrt
- prox_prisch
- prox_cbd

Select p-value threshold for Stepwise Selection:

0.05

Choose Stepwise Method:

Forward

Submit

Stepwise Method Stepwise Comparison Regression Diagnostics (Performance Package) - Forward
 Regression Diagnostics (Performance Package) - Backward Regression Diagnostics (Performance Package) - Both

Please build the Stepwise model first.

- 1) Users will need to configure all the parameters based on their preference and then press submit to built the model

Dataset Selection

3-Room

Date Range (latest months)

1

Dependent Variable

Monthly Rent

Select Independent Variables

- no_of_kindergarten_500m
- prox_kindergarten
- no_of_chilcare_500m
- prox_chilcare
- no_of_hawker_500m
- prox_hawker
- no_of_buststop_500m
- prox_buststop
- no_of_shoppingmall_1km
- prox_shoppingmall
- prox_mrt
- prox_prisch
- prox_cbd

Select p-value threshold for Stepwise Selection:

0.05

Choose Stepwise Method:

Forward

Submit

Stepwise Method Stepwise Comparison Regression Diagnostics (Performance Package) - Forward
 Regression Diagnostics (Performance Package) - Backward Regression Diagnostics (Performance Package) - Both

Model Information - Forward

| Stepwise Summary | | | | | |
|------------------|------------|-----------|-----------|---------|---------|
| Step | Variable | AIC | BIC | R2 | Adj. R2 |
| 0 | Base Model | 13307.471 | 13366.459 | 0.00000 | 0.00000 |
| 1 | prox_cbd | 13333.008 | 13392.012 | 0.02059 | 0.02248 |
| 2 | prox_art | 13328.284 | 13347.381 | 0.04508 | 0.03466 |

| Final Model Output | | | | | |
|--------------------|---------|-----------|------------|--|--|
| Model Summary | | | | | |
| R | 0.152 | RSE | 489.119 | | |
| R-Squared | 0.037 | MSE | 240869.762 | | |
| Adj. R-Squared | 0.035 | Coef. Var | 18.896 | | |
| Pred R-Squared | 0.038 | AIC | 13328.284 | | |
| MSE | 355.249 | BIC | 13347.381 | | |

RSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error
AIC: Akaike Information Criteria
BIC: Schwarz Bayesian Criteria

| ANOVA | | | | | |
|------------|----------------|-----|-------------|--------|--------|
| | Sum of Squares | DF | Mean Square | F | Sig. |
| Regression | 8812571.142 | 2 | 4886285.571 | 16.689 | 0.0000 |
| Residual | 209332984.401 | 872 | 240869.762 | | |
| Total | 217345555.543 | 874 | | | |

| Parameter Estimates | | | | | |
|---------------------|--|--|--|--|--|
|---------------------|--|--|--|--|--|

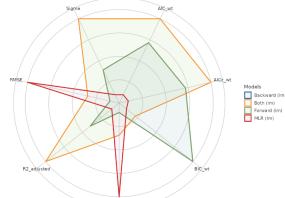
- 2) Model will be built and Stepwise Method model detail will be shown (based on the methods that were chosen)

- 3) Users then are able to explore the Model detail by going to different Tabs.

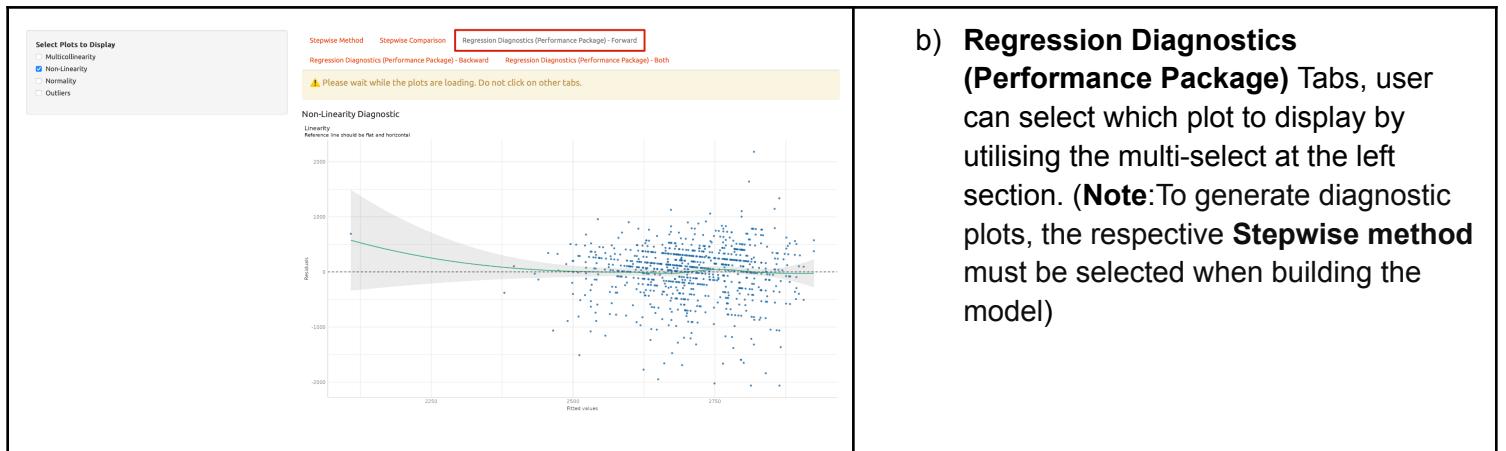
Information Only
This tab is for display purposes only. No input is required because it provides a comparison view based on pre-calculated data.

Stepwise Method Stepwise Comparison Regression Diagnostics (Performance Package) - Forward
 Regression Diagnostics (Performance Package) - Backward Regression Diagnostics (Performance Package) - Both

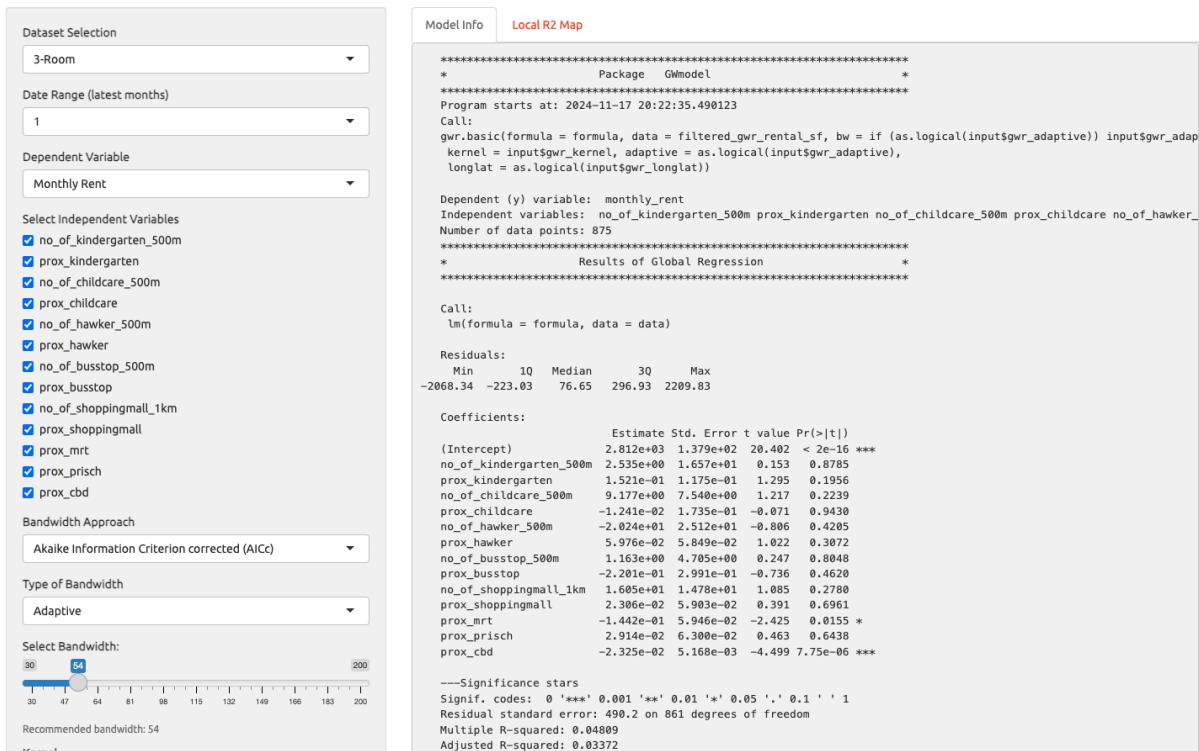
Comparison of Model Indices



- a) Stepwise Comparison Tab



2.4 GWR



| | |
|--------------------|---|
| Stepwise Selection | |
| Stepwise Method | [‘Forward’, ‘Backward’, ‘Both’] |
| Bandwidth Approach | [‘Cross-validation CV’, ‘Akaike Information Criterion Corrected’] |
| Type of Bandwidth | [‘Adaptive’, ‘Fixed’] |
| Bandwidth | Range of 30 to 200 |
| Kernel | [‘Gaussian’, ‘Exponential’, ‘Bisquare’, ‘Tricube’, ‘Boxcar’] |
| Distance Measure | [‘Euclidean Distance’, ‘Great Circle Distance’] |

Instructions

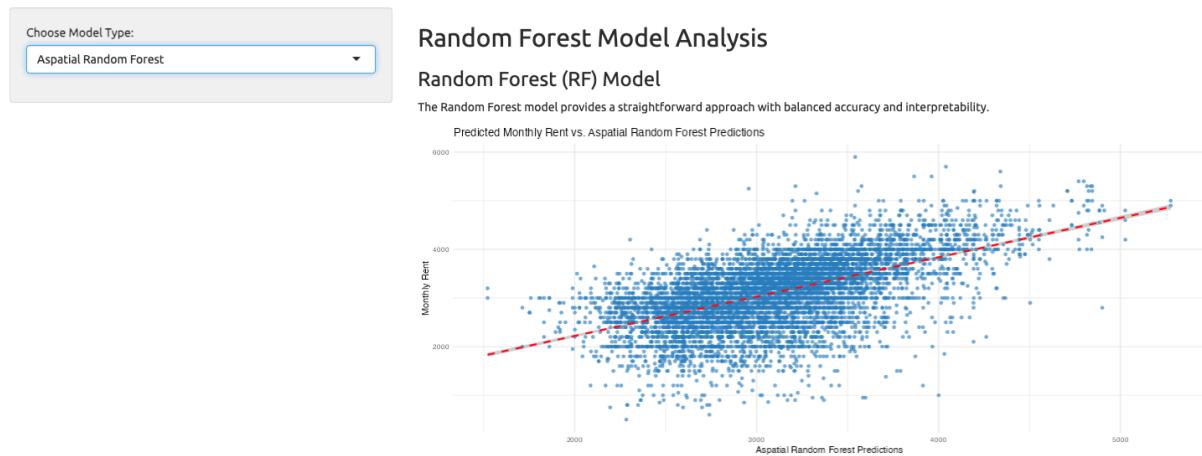
- 1) Users will need to configure all the parameters based on their preference and then press submit to built the model



- 2) Users can go to the **Local R2 Map** Tab to view the Local R2 Map.

3. Predictive Model

3.1 Scatterplot



| Configuration Option | Choices | Description |
|----------------------|---|---------------|
| Model Type | [‘Aspatial Random Forest’, ‘Tuned Random Forest’, ‘Geospatial Random Forest’] | Type of model |

3.2 Models

This predictive model tool allows you to estimate HDB rental prices in Singapore. Select a model type from the dropdown list on the left. If you choose an Aspatial Model, provide relevant inputs in the Aspatial Model Inputs section. For the Geospatial Model, select a location on the map to set coordinates along with additional inputs.

Once you've entered all required inputs, the predicted rental price will be displayed below.

Model Selection Guidance:

- **Aspatial Random Forest:** Provides balanced accuracy and interpretability, useful for general predictions.
- **Tuned Random Forest:** Optimized for scenarios where high accuracy is critical.
- **Geospatial Random Forest:** Ideal for predictions where spatial factors play a major role, such as in real estate.

| Configuration Option | Choices |
|---|--------------------|
| Number of Kindergartens within 500m | Range from 0 to 20 |
| Number of Childcare Centers within 500m | Range from 0 to 30 |
| Distance to Nearest Hawker Center (m) | Positive Numbers |

| | |
|---|------------------|
| Distance to Nearest Shopping Mall (m) | Positive Numbers |
| Distance to Nearest MRT Station (m) | Positive Numbers |
| Distance to Central Business District (CBD) (m) | Positive Numbers |

3.2.1 Aspatial & Tuned Random Forest

Choose Model Type:

Aspatial Random Forest

Aspatial Model Inputs

Number of Kindergartens within 500m:

Number of Childcare Centers within 500m:

Distance to Nearest Hawker Center (m):

Distance to Nearest Shopping Mall (m):

Distance to Nearest MRT Station (m):

Distance to Central Business District (CBD) (m):

How to Use This Page

This predictive model tool allows you to estimate HDB rental prices in Singapore. Select a model type from the dropdown list on the left. If you choose an Aspatial Model, provide relevant inputs in the Aspatial Model Inputs section. For the Geospatial Model, select a location on the map to set coordinates along with additional inputs.

Once you've entered all required inputs, the predicted rental price will be displayed below.

Model Selection Guidance:

- Aspatial Random Forest: Provides balanced accuracy and interpretability, useful for general predictions.
- Tuned Random Forest: Optimized for scenarios where high accuracy is critical.
- Geospatial Random Forest: Ideal for predictions where spatial factors play a major role, such as in real estate.

Model Prediction Output

3774.07

3.2.2 Geospatial Random Forest

Choose Model Type:

Geospatial Random Forest

Geospatial Model Inputs

Flat Type:

3-room

Number of Kindergartens within 500m:

Number of Childcare Centers within 500m:

Distance to Nearest Hawker Center (m):

Distance to Nearest Shopping Mall (m):

Distance to Nearest MRT Station (m):

Distance to Central Business District (CBD) (m):

How to Use This Page

This predictive model tool allows you to estimate HDB rental prices in Singapore. Select a model type from the dropdown list on the left. If you choose an Aspatial Model, provide relevant inputs in the Aspatial Model Inputs section. For the Geospatial Model, select a location on the map to set coordinates along with additional inputs.

Once you've entered all required inputs, the predicted rental price will be displayed below.

Model Selection Guidance:

- Aspatial Random Forest: Provides balanced accuracy and interpretability, useful for general predictions.
- Tuned Random Forest: Optimized for scenarios where high accuracy is critical.
- Geospatial Random Forest: Ideal for predictions where spatial factors play a major role, such as in real estate.

Select Location on Map for Coordinates

Model Prediction Output

Please select a location on the map to get coordinates.

Besides the usual configurations, users are also required to indicate a location on the map to specify a location for the model to generate predictions for. This feature is specific to the “Geospatial Random Forest” model