GWModel: A Visualisation Application to explore Geographical Weighted Regression model of Singapore Hawker Centres rental price trends

Chia Yong Qing, Choo Mei Xuan, Clara Chua Kiah Hwii

Abstract—Although information on past rentals are publicly available on the National Environmental Agency (NEA)'s website, it is not in a form that is easily accessible to users. The project aims to understand hawker centre rentals by visualising the rental price trends of different hawker centres, and to analyse the factors that affect stall rentals. Our group built an R Shiny app that allows users to explore Singapore's public hawker centre rental data, view price trends, and lookup prices of hawker stall rentals in selected individual hawker centres. Users may be able to use our application to see if their rental bid price is within the market range for the type of stalls in their desired market. More advanced users can use our application to perform a geo-weighted regression to understand how locations and different variables affect the rental price of the hawker stalls.

Index Terms — Hawker Centre, Bid Price, Rental, Trade, GWmodel, Geographical Weighted Regression, Cooked Food, Market

1 Motivation

Hawker centres are a staple in the Singapore food scene. In July 2018, there was a monthly rental bid of $10,028 for a drinks stall at Chomp Chomp Food Centre (Wei & Lee, 2018). The hawker in question has since ended the tenancy agreement, but this has led us to wonder: do hawkers know the market rate for hawker stalls and do they know how much to bid?

NEA manages and regulates 114 public hawker centres in Singapore. Every month, NEA puts up vacant stalls at various hawker centres for tender, which are then awarded to the highest bidders. Aspiring hawkers will need to submit a tender bid for a selected stall. If they are the winning bidder, they have a license to operate the stall for 3 years, after which they may need to renew or bid for a stall again. Bidders who fail to sign the Tenancy Agreement after being awarded the stall will have their deposit forfeited and be debarred from participating in all future government tenders.

Therefore, it is important for bidders to know what is a 'fair' market rate to pay for their stall, as well as what is sustainable for their business.

2 Objective

Although information on past tender bids is publicly available on NEA's website (Tender Notices , 2018), it is not in a form that is easily accessible to users. The project aims to understand hawker centre rentals by visualising the rental price trends of different hawker centres, and to analyse the factors that affect stall rentals.

Our group built an application that allows users to explore Singapore's public hawker centre bid price data. Users will be able to view price trends, and look up prices of hawker stall rentals at any selected hawker centre. Most importantly, users can use our application as a benchmark to see if their bid price is within a reasonable price range for the type of stalls at their desired hawker centres.

More advanced users can use our application to perform geo-weighted regression to understand how locations and different variables affect the rental price of the hawker stalls.

3 Critique On Existing Visualisations

Prior to designing our application dashboard, our group took reference from existing visualisations online to understand their limitations and to emulate best practices. We looked at two main sites: SG Charts and iChef Club SG.

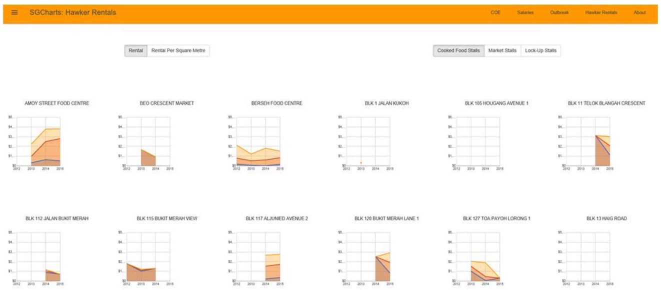


Figure 1: SG Charts dashboard of Hawker Centre Rental Prices

SG Charts' dashboard features hawker stall centre rentals from 2012 to mid-2015. It shows a trellis chart of the min, median and max rental prices, and rental per square meter (psm) for 3 main categories of stalls in each hawker centre. Though the chart allows for interactivity in exploring the distribution of rental prices across various categories of stalls and in different units of measurement, the creator has missing values in the dataset which results in empty graphs within the visualisation.

Figure 2: Infographic from iChef Club SG

iChef Club SG's visualisation shows a range of rentals and the price trends at selected hawker centres, for cooked food stalls. While the visualisation provides a clear snapshot of popular hawker centres, it does not allow viewers to explore price trends of different food stalls for individual hawker centres across Singapore.

Nonetheless, we have to consider that the visualisation was intended for a blog post and an interactive component is not required.

4 Data Preparation

Column	Description	Example	Data Source
HAWKER_CENTRE	Name of hawker centre	Redhill Hawker Centre	NEA "List Of 5 Highest Tender Bids"
STALL_NO	Unit number of stall	01-66B	NEA "List Of 5 Highest Tender Bids"
STALL_AREA	Size of stall (sqm)	5.65	NEA "Tender Bids For Hawker Stalls From March 2012 To September 2018"
TRADE	Specific type of trade	Vegetables	NEA "List Of 5 Highest Tender Bids"
TRADE_GENERIC	Generic type of trade	Fresh Produce	Derived from "Trade"
TENDERED_BID	Successful bid price ($)	$2988	NEA "List Of 5 Highest Tender Bids"
AVERAGE_BID_PRICE	Average successful bid price of Hawker Centre ($)	$450	Derived from "Tendered bid"
MONTH	Month of successful bid	Apr	NEA "Tender Bids For Hawker Stalls From March 2012 To September 2018"
YEAR	Year of successful bid	2017	NEA "Tender Bids For Hawker Stalls From March 2012 To September 2018"
DATE	Month and Year of successful bid	Apr-2017	NEA "Tender Bids For Hawker Stalls From March 2012 To September 2018"
TYPE_OF_STALL	Type of business	Market Stalls	NEA website
POSTAL CODE	6 digits zip code representing postal location of hawker centre	560572	NEA website
LAT	Latitude in 7d.p.	1.2793399	Derived from postal code using R
LONG	Longitude in 7.d.p	103.8466525	Derived from postal code using R
BID_PRICE_PER_SQM	Tendered bid price of stall divided by stall size in sqm	528.85	Derived from tendered price and store size
LAST_RENOVATION	Reopening date after renovation	Feb-2003	NEA Hawker Centres Upgrading Programme
AGE_OF_HAWKER	Number of years past since reopening date	15.8	Derived from fraction of year since reopening date
TYPE_OF_UPGRADING	Reason for upgrading	Reconfiguration	NEA Hawker Centres Upgrading Programme

Table 1: Metatdata of main working file

Geocoding is performed to obtain the latitude and longitude of the hawker centres for geospatial analysis, leveraging on the OneMap API. Latitude and longitude for locations of MRT station and bus stops are extracted from databases and shapefiles available on Data.Gov using the sf package. The dataset for HDB location based on postal code is provided by Professor Kam Tin Seong.

We use the sp package to manipulate and convert our data into Spatial DataFrames, as the GWmodel package only accepts that format. We use the sf package to convert x, y coordinates to longitude and latitude and use the gw.dist function from the GWmodel package to calculate the distance between the hawker centres and each amenity.

Additional model parameters are derived as follows:

i. mindist_(input): Shortest distance to the input variable e.g. mindist_mrt
ii. (input)_0.35: number of input amenities within 350m e.g. mrt_0.35
iii. (input)_0.5: number of input amenities within 500m e.g. mrt_0.5
iv. (input)_0.75: number of input amenities within 750m e.g. mrt_0.75
v. (input)_1: number of input amenities within 1km e.g. mrt_1
vi. avepricepsm_cooked: average price per square metre of cooked food stalls at the selected hawker centre
vii. avepricepsm_market: average price per square metre of market stalls at the selected hawker centre
viii. medpricepsm_cooked: median bid price per square metre of cooked food stall at the selected hawker centre
ix. medpricepsm_market: median bid price per square metre of market stall at the selected hawker centre

5 Methodology

5.1 Geographical Weighted Regression

Our application generates a regression model of each individual hawker centre using the R package, GWmodel. Geographical Weighted Regression (GWR) models are most suitable for situations where the local difference is pronounced. For the scope of the project, users can investigate what are the input variables i.e. proximity to amenities which have a significant impact on the rental bid prices of hawker centres.

$$y_i = \beta_{i0} + \sum_{k=1}^{m} \beta_{ik} x_{ik} + \epsilon_i$$

5.2 Kernel and Bandwidth Selection

Spatial dependency between each hawker centre and nearby amenities can be explored by weighing each observation point in a diagonal matrix, W (ui, vi). The distance matrix varies according to the type of distance function, kernel function and its bandwidth (Lu, 2018).

The GWR model provides a local model of the rental prices by fitting a regression equation to every hawker centre in the dataset. It constructs these separate equations by incorporating the dependent and input variables falling within the bandwidth of each parameters.

Being distance decay in nature, the GWR model employs a moving window technique where nearer observations have more weightage in estimating the local set of regression coefficients than observations which are further away (Páez, Farber, & Long, 2008). Our application enables users to select their desired kernel function from 5 options – *Gaussian*, *Exponential*, *Box-Car*, *Bi-Squared*, *Tri-Squared*.

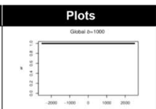
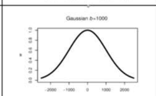
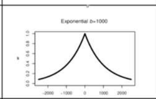
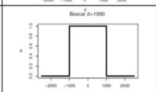
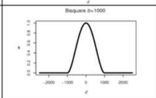
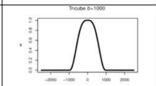
Kernel	Formula	Plots	Description
Basic	$w_{ij} = 1$		- Equal weightage for all observations.
Gaussian	$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$		- Continuous function. - Weight decreases according to Gaussian Curve as distance between observation and calibration points increases.
Exponential	$w_{ij} = \exp\left(-\frac{\lvert d_{ij}\rvert}{b}\right)$		- Continuous function. - Weight decreases according to Exponential Curve as distance between observation and calibration points increases.
Box Car	$w_{ij} = \begin{cases} 1 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$		- Discontinuous function. - Exclude observations that is some distance further from calibration point. - Useful for large dataset.
Bi-Square	$w_{ij} = \begin{cases} (1-(d_{ij}/b)^2)^2 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$		- Discontinuous function. - Weight decreases as distance between observation points increases.
Tri-Cube	$w_{ij} = \begin{cases} (1-(\lvert d_{ij}\rvert/b)^3)^3 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$		- Discontinuous function. - Weight decreases as distance between observation points increases.

Table 2: Types of Kernel Functions

The default distance function of our model is set as Euclidean Distance, i.e. distance between 2 observations are measured as a straight line drawn between them. (Euclidean Distance: Raw, Normalised and Double Scaled Coefficient, 2005). The optimal bandwidth is selected by the model.

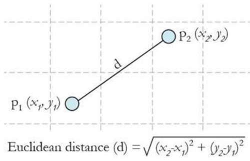


Figure 3: Euclidean Distance

6 Visual design framework

The following methodology was applied in the design of our visualisation dashboard. The first tab gives a mapped overview of the location and distribution of rental price trends across different hawker centres in Singapore. The second tab shows selected exploratory analysis across different hawker stall types and hawker trades in Singapore. Finally, the third tab allows users to perform geographically weighted regression to understand how availability and proximity of transport options affect the successful bid price of hawker centres.

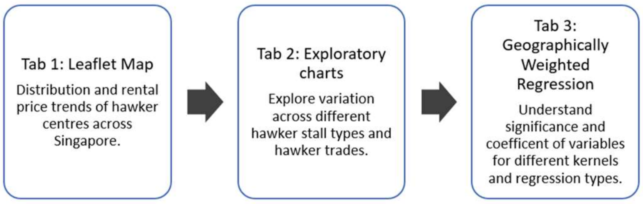


Figure 4: Overview of Visualisation Design Framework

6.1 Overall Visualisation of Hawker Centres and Stall Rentals across Singapore (Tab 1)

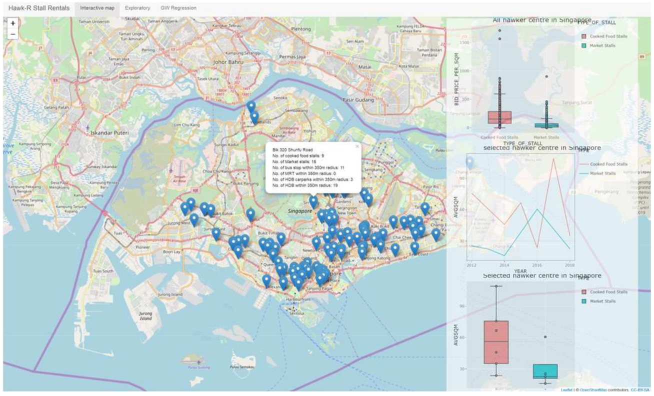


Figure 5: Interactive map of public hawker centres and rental prices

6.1.1 Visualising hawker centres on a map and corresponding price information

This interactive visualisation uses an AbsolutePanel function to display a translucent, draggable tab panel against a full view backdrop of Singapore from openstreetmap using the leaflet package. This allows users to explore the interactive plots without obscuring any data points, and also allows users to pan and zoom into various locations on the map.

Each hawker centre's location is represented by a blue marker on the Singapore map. When the user clicks on a marker, a pop-up box will display the name of the hawker centre, the number of available cooked food stalls and market stalls. Other information displayed includes the number of bus stops, HDB blocks, HDB carparks and MRT stations within a 350m radius of the hawker centre.

All charts on the panel are plotted using the ggplot2 package and made interactive using the plotly package.

6.1.2 Comparing successful bid price of all hawker centres vs. individual hawker centres

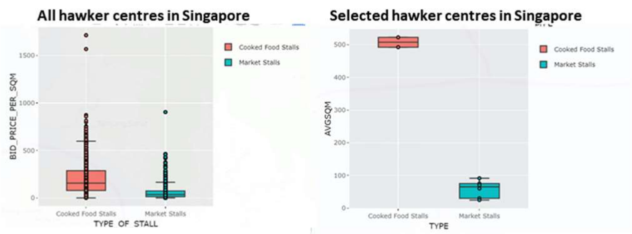


Figure 6: Box plot of bid price per sqm for all vs selected hawker centres in Singapore

A static box plot of the average bid price per sqm of all hawker centres in Singapore stall type is embedded at the top of the side tab panel. This allows users to have a visual comparison between the bid price spread of selected hawker centres and all hawker centres in Singapore.

An interactive box plot embedded at the bottom of the panel shows the bid price per sqm by stall type in a selected hawker centre. Unlike the first plot, this box plot is interactive, and updates when a different hawker centre is selected.

6.1.3 Time series of average bid price (psm) for individual hawker centres

Similarly, an interactively linked time series line graph of average bid price per sqm of the selected hawker centre is shown as the second plot on the panel. Users can view the changes in bid prices of market and cooked food stalls in the selected hawker centre over the years.

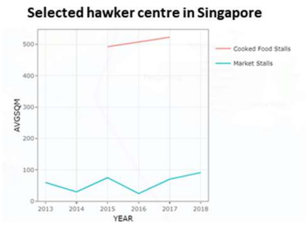


Figure 7: Line graph of selected hawker centre

6.2 Exploratory Visualisation on Hawker Types and Trades Across Singapore (Tab 2)

The exploratory tab contains 4 independent interactive plots to assist users in understanding the distributions of bid price and bid counts across different types of hawker stalls and trade types in Singapore.

6.2.1 Time Series of all hawker centres in Singapore

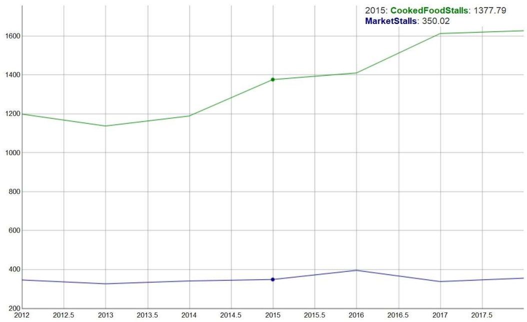


Figure 8: Time series chart of average bid price of all hawker centres by stall type, 2012 - 2018

An interactive time series chart of the average bid price of market and cooked food stalls across all hawker centres is plotted using the Dygraph package. Users can hover over different points to return the total number of successful bids for cooked food and market stalls in that year at the top right corner of the plot.

From 2012-2018, we observed a gradual increase in bid prices of cooked food stalls where average bid prices of market stalls have been relatively stagnant across the years.

6.2.2 Relationship of bid price and age of hawker centres

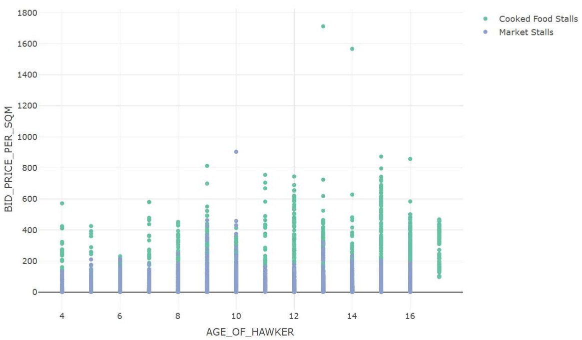


Figure 9: Scatterplot of bid price per sqm by age of hawker

The age of hawker centres (age_of_hawker) is derived as a fraction of year (as of 1 November 2018) from the date the hawker centre was reopened after the most recent renovation. Using the ggplot2 package, we plotted the variance of bid price per sqm across hawker centres of different ages. Counterintuitively, the age of hawker centres does not have a visible impact on the bid price of the stalls. Hawker centres which were last renovated more than 10 years ago are observed to have similar or higher bid price per sqm compared to its newly renovated counterparts.

6.2.3 Breakdown of bid prices and bid counts by hawker trades

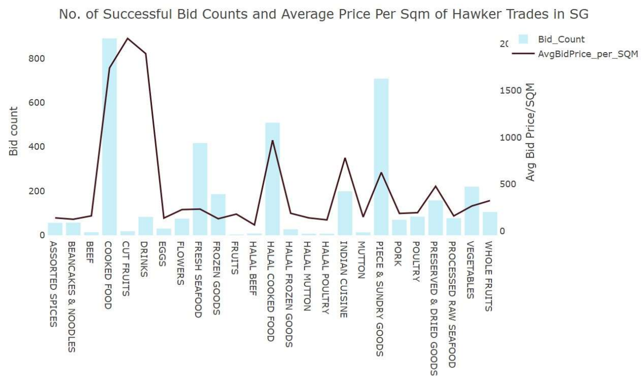


Figure 10: Bar chart of Cumulative Successful Bid Counts and Average Price per sqm by hawker trade, 2012 – 2018

Using the ggplot2 package, the number of successful bid counts and average bid price per sqm are plotted for the 25 hawker trades in a dual axis bar/line chart. The bid count of each trade is represented by the height of the bar in a bar chart and its corresponding average bid price is plotted as a line graph overlapping it.

6.2.4 Visualisation of proportion of bid counts and average bid prices

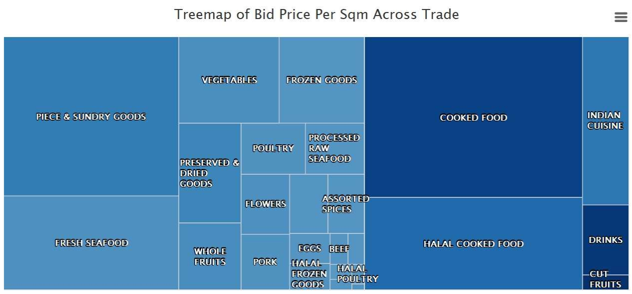


Figure 12: Treemap of bid price per square meter and bid counts across hawker trades for all hawker centres

A treemap is plotted using the Highcharter package which incorporates an interactive hover and drill-down function to the plot. A larger proportion of bid counts is represented by the size of the area. The average price per sqm of a trade is represented by the different intensity of the colour. A darker shade denotes a higher average price per sqm.

Collectively, there are more successful bids for market stalls compared to cooked food stalls. Users can click on each overarching type of hawker stall – market stalls and cooked food stalls to zoom in on the corresponding hawker trades.

6.3 Visualisation of Geographical Weighted Regression Modelling (Tab 3)

The application provides a platform for users to explore the variables in depth using a geographically weighted regression model without having to understand the code behind the GWmodel package. Specifically it allows users to perform the following:

Geographically Weighted Regression (GWR) Modelling – Explore the correlations between the different inputs, generate the optimal bandwidth, generate regression results based on user selections.

Visualisation of regression results – View the results of the regression for all hawker centres on a map of Singapore.

6.3.1 Exploring multicollinearity of input variables

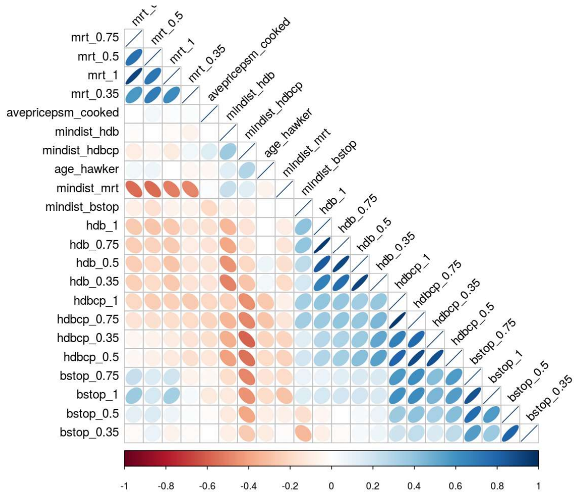


Figure 13: Correlation matrix of all variables

Using the corrplot package, users can explore the correlation between different inputs. Users can select different combinations of variables and interpret the correlations through the colour gradient from red (negative correlation) to blue (positive correlation). Users can choose to remove variables that are highly correlated to change the performance of the regression.

6.3.2 Fit Statistics and Significant Variables

Users can perform Geographically Weighted Regression (GWR) through selection of type of stalls, dependent variable, type of kernel

Significance Level

99.9% 99% 95% 90% All

Model Comparison

	LM model	GW model
RSS	1029984	793858.576454334
AIC	1142.481	1105.648070147
AICc	1159.208	1182.22850467186
R2	0.1767	0.365467958369127
Adjusted R2	-0.07751	-0.10959908609807

Global GW Regression Model Coefficients & p-Values

Search:

	input$indepvar	F3 statistic	Numerator DF	Denominator DF	Pvalue
1	mindist_bstop	1.890671397793	21.2913369907126	60.182761404899	0.0279748294736418
2	bstop_0.5	1.84307813444723	29.474350392619	60.182761404899	0.0226136742610369
3	bstop_1	1.91503742090645	28.8578930708377	60.182761404899	0.0171352566936511
4	mrt_0.5	2.42737667347024	37.8505497275635	60.182761404899	0.00102369277382283
5	hdb_0.5	2.02777301057863	29.1391878514768	60.182761404899	0.0104477656718385

Showing 1 to 5 of 5 entries

Figure 11: Fit statistics and global GWR results

functions, independent variables and regression type

A model comparison table shows the fit statistics of both the global linear model (LM) and GWR model. Once the regression is run, users can see which variables that are significant at 90%, 95%, 99% and 99.9% level.

6.3.3 Visualisation of Regression Results by Hawker Centre

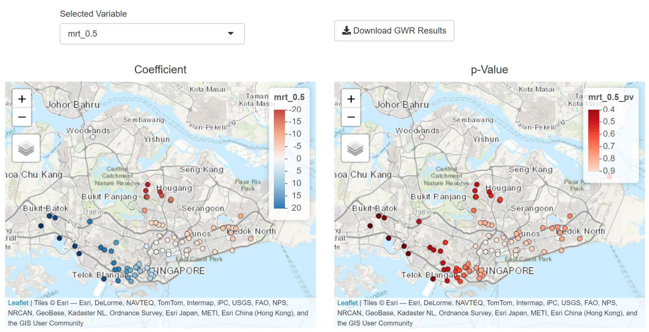
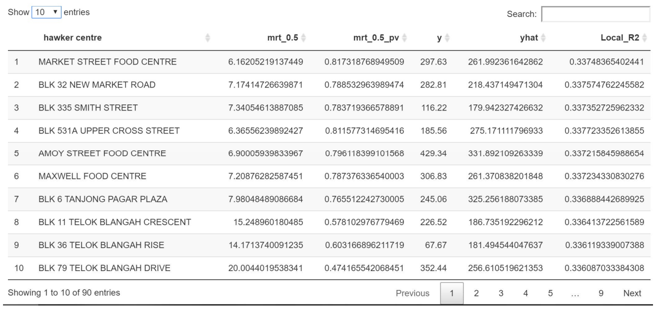


Figure 14: Plot of coefficient and p-values of selected dependent variable for each hawker centre

The GWR gives a localised coefficient and p-value for each hawker centre, and we can identify clusters of hawker centres that show similar characteristics. Users can select the variable to explore and see how that impacts each hawker centre.

The plot above shows a side-by-side view of the coefficient and p-values of each variable for all hawker centres. The variable shown (No. of MRT stations within a 500m radius), is 99% significant at the global level, and we can see how that varies from location to location. It is more significant in the west regions (from the p-value figure on the right) and we can see that the coefficients of those hawker centres are also larger, which means that the number of MRT stations within a 500m radius has more impact on the hawker centres in that area, than the others.

Additionally, the results of the localised regression for each hawker centre, from the local R^2 of the regression (Local_R2) to the actual and predicted bid price psm (y, yhat) can be seen from the data table under the figures. Users can also download the regression results for further reference or analysis.



Show 10 entries · Search:

	hawker centre	mrt_0.5	mrt_0.5_pv	y	yhat	Local_R2
1	MARKET STREET FOOD CENTRE	6.16205219137449	0.817318768949509	297.63	261.992361642862	0.33748365402441
2	BLK 32 NEW MARKET ROAD	7.17414726639871	0.788532963989474	282.81	218.437149471304	0.337574762245582
3	BLK 335 SMITH STREET	7.34054613887085	0.783719366578891	116.22	179.942327426632	0.337352725962332
4	BLK 531A UPPER CROSS STREET	6.36556239892427	0.811577314695416	185.56	275.17111796933	0.337723352613855
5	AMOY STREET FOOD CENTRE	6.90005939833967	0.796118399101568	429.34	331.892109263339	0.337215845988654
6	MAXWELL FOOD CENTRE	7.20876282587451	0.787376336540003	306.83	261.370838201848	0.337234330830276
7	BLK 6 TANJONG PAGAR PLAZA	7.98048489086684	0.765512242730005	245.06	325.256188073385	0.336888442689925
8	BLK 11 TELOK BLANGAH CRESCENT	15.248960180485	0.578102976779469	226.52	186.735192296212	0.33641372256159
9	BLK 36 TELOK BLANGAH RISE	14.1713740091235	0.603166896211719	67.67	181.494544047637	0.336119339007388
10	BLK 79 TELOK BLANGAH DRIVE	20.0044019538341	0.474165542068451	352.44	256.610519621353	0.336087033384308

Showing 1 to 10 of 90 entries · Previous 1 2 3 4 5 … 9 Next

Figure 15: Datatable of localized regression results for selected variable

The tmap package is used to create the side-by-side map, rgdal was used to associate the spatial data set with the relevant coordinate reference system (WGS84) and the datatable package (DT) was used to enable users to filter and also sort results as needed.

7 Results and Insights

Type of stall	Cooked food			Market		
Bid price per SQM	Average bid price per SQM			Median bid price per SQM		
Model	Global	Basic	Robust	Global	Basic	Robust
AIC	1,142	1,106	1,106	845	802	802
AICc	1,159	1,182	1,182	864	885	885
R^2	0.18	0.37	0.37	0.29	0.50	0.50
Adjusted R^2	(0.08)	(0.11)	(0.11)	0.04	0.07	0.07

Table 3: Comparison of global LM model and GWR models

From the results derived using our application, we discovered that GWR model generally performs better on various measures when benchmarked against the global model for both cooked food and market stalls (Brunsdon, 2015). Specifically,

i. R^2 which measures goodness of fit is highest for Basic model for both cooked food and market stalls.
ii. Lower AIC observed for Basic and Robust model for both cooked food and market stalls.
iii. After compensating for number of variables in the model, Basic and Robust model still outperform Global model for cooked food and market stalls i.e. higher adjusted R^2.

The improvement to adjusted R value when using the Basic and Robust model allude to differences between local and global environment which results in variance in bid price of hawker stalls at different locations.

	Cooked Food Stalls		Market Stalls	
Model	Basic / Robust (Filtered)		Basic / Robust (Filtered)	
Kernel	Bisquare		Bisquare	
Optimal Bandwidth	88		81	
No. Observations	90		83	
Dependent Variable	Average Price		Median Price	
	Significant Variables			
	Variable	P value	Variable	P value
99% significance	No. of MRT stations within 500m	0.001024	No. of Bus Stops within 1km	0.000631
			No. of HDB blocks within 1km	0.000160
			No. of HDB blocks within 750m	0.003601
95% significance	No. of HDB carparks within 500m	0.010448	Shortest Distance to MRT	0.046301
	No. of Bus Stops within 1 km	0.017135	No. of MRT stations within 500m	0.049937
	No. of Bus Stops within 500m	0.022614		
	Shortest Distance to Bus Stop	0.027975		
90% significance	Age of Hawker Centre	0.052305	No. of MRT stations within 750m	0.053912
			No. of Bus Stops within 500m	0.094862

Table 4: Significant Variables by Stall Type

A total of 22 variables are included as parameters for our GWR model – Number of HDB within 350m/500m/750m/1km, Number of HDB carpark within within 350m/500m/750m/1km, Number of MRT stations within within 350m/500m/750m/1km, Number of Bus Stops within within 350m/500m/750m/1km, Minimum distance to HDB/HDB carpark/MRT stations/ Bus Stops and age of hawker centre.

Some general observations observed are that Cooked Food stalls depend largely on the availability of transport options, in the form of HDB carparks and MRT stations. The number of MRT stations are 99% significant within 500m.

The availability and number of bus stops and HDB blocks within 1km is significant at 99.9% level of confidence for Market Stalls. This could signal dependency of Market Stalls on a larger catchment area.

8 Future Work

One major limitation to our current dataset is the lack of complete information on rental prices. There was no data available for all tender bids (NEA only had the top 5 bids available) and the pricing also does not take into consideration rentals of hawker stalls under older schemes. Future work can also include information on running costs and operating costs, which are only available for selected stalls and are not in easily ingestible format. Additionally, we can explore a larger set of data, including non-NEA-regulated hawker centres across a longer period of study.

Though our results show that adjusted R^2 value for basic and robust models are relatively better than that of the Global regression model, it is considered low at 0.07 for market stalls and -0.11 for cooked food stalls. Other independent variable inputs can be explored to improve the fit of the model by increasing adjusted R^2 value.

A potential replication of our work can be achieved using data from HDB shophouses, coffee shops and shopping mall tender bid prices for appreciation of bid price trends impact of local amenities on the demand across various retail formats.

9 Learning Experience

Chia Yong Qing – The overall look of the whole app and making the side panel translucent was difficult to achieve without the use of style.css, which we had not had prior experience of. We were luckily able to seek the advice of our seniors who had some experience in this area. We also had other visualisations intended but after going in depth into the dataset, we had to come up with alternative ways to achieve our project objectives.

Choo Mei Xuan – We had tried different packages (from d3tree to treemap to highcharter), but we were not able to link the treemap to the breakdown of the bid price and bid counts, which would have made our app more interactive. It was also challenging to link the map markers on the first tab to the charts in the draggable side panel. We managed to figure this out through an iterative process.

Clara Chua – It was a challenge to translate our previously coded geoweighted regression results into an R shiny app. We had to rethink the structure of our data and code to make the geoweighted regression reactive to user inputs. In addition, we had to use different workarounds to show the visualisations we intended for the side-by-side map of the coefficients and p-value plots. Using shiny and leaflet, it was not possible to show a legend for the size of a bubble plot, which meant that we were unable to show both coefficient and p-value of each variable in one single map. In addition, we were unable to sync the tmaps together in the shiny app, unlike in the normal R environment. In conclusion, our learning point is really to explore different ways of presenting our visualisation and to look at alternative packages.

10 Conclusion

Our group designed and produced an exploratory visualization application for aspiring hawkers and the curious alike to explore past trends in tender bid prices at various hawker centres.

Through the R shiny app, we achieved our objectives to deliver tender bid information available on NEA website in a concise and easily understandable format with the use of R visualisation packages and interactively linked mechanisms.

Our app also provides additional analysis by incorporating proximity to nearby transport options, which may affect tender bid price at a hawker centre. The impact of each option on hawker stall prices are tested as parameters in a Geographically Weighted Regression model at various levels of significance.

The results derived from our application shed some light on the possible reasons for variation in tender bid prices across various hawker centre locations across Singapore.

11 Acknowledgments

The authors wish to thank Professor Kam Tin Seong for his guidance and sharing for this project. Professor Kam also provided a list of geocoded HDB postal code as one of the parameters for the GWR model.

We would also like to thank Mr Raymond Foo for his advice in coding the UI and using CSS.

12 References

[1] Brunsdon, C. (2015, Aug 10). Geographically Weighted Summary Statistics. Retrieved from RPubs.

[2] Euclidean Distance: Raw, Normalised and Double Scaled Coefficient. (2005, September). Retrieved from Paul Barrett: pbarrett.net

[3] Lu, B. (2018, 08 20). Geographically-Weighted Models. Retrieved from Package GWmodel: http://gwr.maynoothuniversity.ie/

[4] Páez, A., Farber, S., & Long, F. (2008). Moving Window Approaches for Hedonic Price Estimation: An Empirical Comparison of Modelling Techniques. *Urban Studies*, 1565-1581.

[5] *Tender Notices* . (2018). Retrieved from National Environmental Agency: https://www.nea.gov.sg/corporate-functions/resources/tender-notices

[6] Wei, T. T., & Lee, J. (2018, August 29). https://www.straitstimes.com/singapore/10028-monthly-rental-bid-for-chomp-chomp-hawker-stall-highest-in-3-years-bidder-walked. *The Straits Times*.

[7]