Too High, Too Low, or Just Right?

School of Information Systems (SIS)

IS415 – Geospatial Analytics and Application

Guided by Dr. Kam Tin Seong

Fu Weiyu, Chloe †
L.K.C. School of Business
Singapore Management University
Singapore

Jia Khee, Lim †
L.K.C. School of Business
Singapore Management University
Singapore

Xin Yuan, Cheng†
L.K.C. School of Business
Singapore Management University
Singapore

1 INTRODUCTION

Airbnb has been democratic in providing public access to its data for any forms of analysis. However, there is a lack of aggregated platforms to help distill this mass of data into information which can empower Airbnb hosts to make better economic decision makings.

While there are literature and methodology on pricing models out there, they commonly focus on the reviews and scores given by previous guests, or the interior design and amenities provided for in each apartment. Little geographically-related analysis is given to the data – i.e. the overall accessibility of the apartment to all places that a guest will be travelling to for during their duration of stay. Hence, our team sees the potential of filling in this gap.

We have chosen to delve into the chosen landscape of Downtown Seattle in Washington, United States to understand the spatial relationship between key places and listing locations, and how it affects each listing's price. Using **Geographical Accessibility** technique as a data exploration tool to help us better understand the overall accessibility of each listing to the key locations, we hypothesized that listings with higher accessibility scores should fetch higher prices. Subsequently, we used **Spatial Point Pattern Analysis** to understand the distribution patterns, density and clustering of listings within various small neighborhoods in Downtown Seattle. Lastly, we built a **Geographically Weighted Regression** model to better support hosts in listing their prices based off their listing's location, distance to different areas within Downtown Seattle and the type of accommodations they are providing.

To make our analysis easily available, customisable and understood by all end users, we have included a RShiny Application Tool that invites users to input different listing parameters into our application. Users can interact with it and observe how their chosen listing matches up with various listings of similar properties. Certain features include comparing listing prices and analysing the distribution of listings across Downtown Seattle.

Keywords

Spatial Analytics, Geographical Accessibility, Spatial Point Pattern Analysis, Geographically Weighted Regression, Airbnb, Seattle, Downtown

2 RELATED WORKS

Others have gone before us at analyzing Airbnb datasets, as well as adopted similar (or new) geographical concepts for separate sets of data; offering supplementary information to the tools and resources we would use. In this section, we highlight these great works.

2.1 Airbnb Rental Listings Dataset Mining

The review is an exploratory analysis of Airbnb's Data set to understand the rental landscape in New York City (NYC). [1] Like our datasets, these datasets used in this related work are sourced from Inside Airbnb as well. The datasets used are "listings.csv", "reviews.csv" and "calendar.csv". "Listings.csv" provided 96 detailed attributes of each listing for all listings within NYC, "reviews.csv" contained information on reviews given by guests with six attributes and "calendar.csv" provided daily details about bookings for the next year for each listing.

2.1.1 Key Assumption made. The authors of the paper ran into the problem of not knowing how many days (in a year) a listing is available for booking. This problem is common to everyone working with Airbnb datasets as it is a piece of information not made publicly available by Airbnb. To bridge this gap, the authors of the paper used 'number of reviews' as a proxy to estimate demand. As per the company, about 50% of the guests review host/listings. Hence, studying the number of reviews will give a good estimation of demand. We adopted this method when exploring our Seattle Airbnb data set.

2.1.2 Key Findings 1: Seasonality in Demand. From analyzing the demand patterns in year 2016 (left most graph), 2017 and 2018 (right most graph), it is apparent that there is a fixed pattern in

[1] Gupta, S., Peshin, A., & Agrawal, A. (2019, January 04). Airbnb Rental Listings Dataset Mining. Retrieved March 29, 2019, from https://towardsdatascience.com/airbnb-rental-listings-dataset-mining-f972ed08ddec

demand trend for listings. All three graphs show identical patterns of having monthly demand rates pick up across the year, with peak demand during the month of October. After which, demand starts to fall, and this cycle repeats again the following year. We postulate that this is due to the seasonality of NYC, as the season gradually shifts from Fall to Winter. From this exercise, we inferred that there exists a relationship between demand and seasonality for listings in NYC.

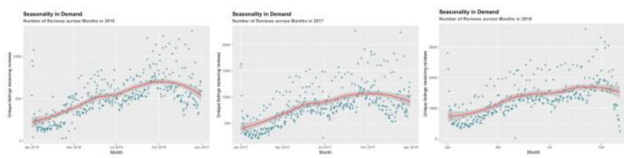


Figure 1: Demand across months of the year, 2016 to 2018

2.1.2 Key Findings 2: Price Fluctuations. Monthly average listing prices follows a similar trend to demand patterns across the year. The average prices tend to increase as one progresses along the year and spikes in December. The only exception exists in November and December, where demand rates starts falling since end October but listing prices do not, instead increasing with time. Hence, one hypothesis we came up with when analyzing listing prices and demand rates across the months of the year is that Airbnb owners do not factor in demand shifts when setting prices. It is likely that they set prices based on intuition.

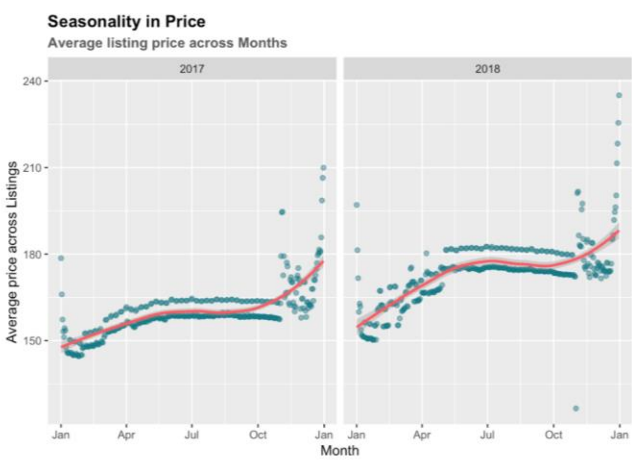


Figure 2: Average listing prices across months of the year, for year 2017 and 2018

2.1.3 Key Findings 3: Dealing with missing values. The authors chose to first construct a Visna plot to analyse the missing values for the variables that they will be using for their exploratory analysis. To preserve all the information, they imputed or dropped all rows and columns containing null values when conducting their exploratory data analysis.

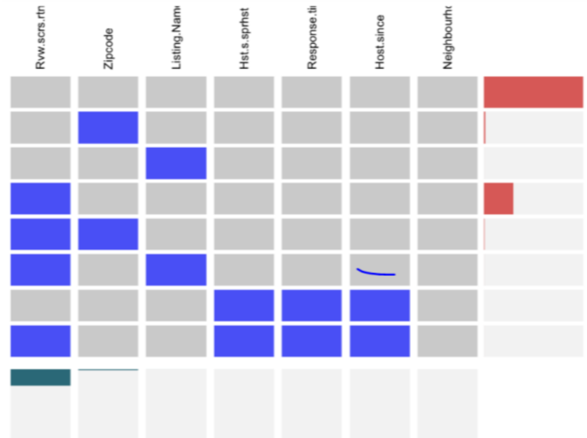


Figure 3: Dealing with missing values through visna plot

Throughout their exercise, a few observations were gathered. Firstly, most rows do not have any missing values, suggesting that the data is rather complete on first inspection. Secondly, 'review_scores_rating' variable have close to 30% missing fields across all rows. Lastly, 'zipcode', 'name', 'host_response_time', 'host_since' and 'host_is_superhost' only have a few missing values across all rows and hence are not reflected in the visna plot. We concluded that the dataset contains only minimal number of missing values and analysis can be performed on the data set without much loss of information.

2.1.4 Key Findings 4: Understanding Categorical Variables. This analysis aims to uncover any type of relationship between different categorical variables in the data. Specifically, the goal of this study was to see if regular hosts and super hosts have different listing policies and examine if there is a correlation about variables. The variables examined were 'instant_booking', 'cancellation_policy' and 'room_type'.

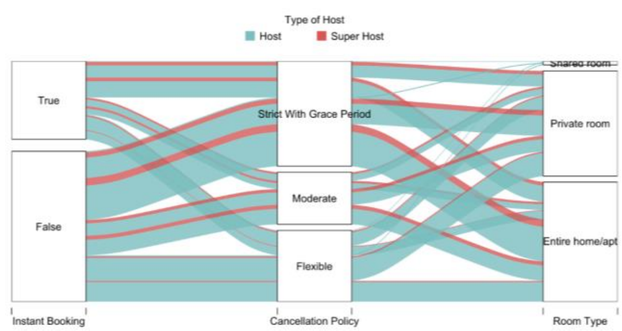


Figure 4: Analysis of Categorical Variables

There were five key takeaways from this exercise. First, majority of the listings are not available for instant bookings. For the ones available, majority of them tend to have strict cancellation policy. Secondly, properties that are entire homes or apartments tend to have stricter cancellation policy compared to shared and private rooms. This made sense as hosts would incur heavier loss by last-minute cancellations on entire homes/apartments booking type. Thirdly, the graph suggests that the shared rooms are available for instant booking and have more flexible cancellation policy. Fourth, owners of shared rooms are generally regular hosts, not super hosts. Lastly, based on the features in this analysis, there were no major discrepancies in behavior between regular and super hosts as they seem to display similar behavior types.

2.2 Spatial Point Patterns in Plant Ecology

This review[2] studies the usage of spatial point pattern analysis (SPPA) methods for the purposes of plant ecology. Although it steers away from analyzing Airbnb datasets, it offers 2 insightful approaches which was worth looking into.

2.2.1 Key Findings 1: Bivariate Point Pattern Analysis. The usage of two variables helps answer the question if events of interest are occurring with any respect to separate types of events. The examples used in the review are such as comparing the points of newly emerged plants, and adult plants. Applying to our project's scope and objective, introducing this analysis brings us one step further in deriving the reasons behind the high Airbnb listing activities in Downtown Seattle; a possible causation.

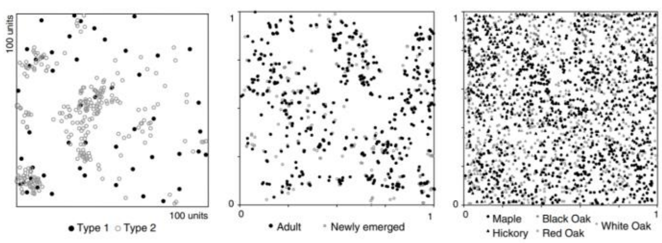


Figure 5: Maps of the 3 Bivariate Sets of Events

In the research paper, they used 3 bivariate sets of events: the Gibbsian point pattern, the Bramble canes event set (newly emerged vs. adults) and the Lansing Woods event set (hickory vs. maple, with other species having smaller grey symbols). For a first order effect analyses, they used a 2x2 contingency table which was recommended by Pielou (1961) and Dixon (1994).

Pattern	Pielou's S	S_{AA}	S_{BB}	P [S]	P [S_{AA}]	P [S_{BB}]
Gibbsian process	−0.014	−0.81	0.68	0.325	0.036	0.002
Brambles: Emerg vs. Adult	0.0061	0.120	−0.09	0.584	0.246	0.856
Lansing: Hickory vs. Maple	0.53	1.16	1.28	0.002	0.002	0.002

Figure 6: Summary Statistics

Interpreting the table gave information such as whether the Gibbsian process has a high association value with Dixon's SAA and SBB, and whether Pielou's S is any dissimilar from being completely random.

For a second order effect, they used Ripley's K and Neighbourhood Density Function. The overall analysis of the second order concluded that it supports the contingency table analyses.

The usage of a bivariate analysis certainly gives more room for our model to be built upon. Unfortunately, our lack of data hinders us from exploring other events that could suggest the high number of listings in Downtown Seattle. Future research can consider adopting this similar approach to identify a possible reason for dense Airbnb listings in Downtown Seattle. Knowing so could inspire new business strategies for home owners or Airbnb itself.

2.2.2 Key Findings 2: Spatial Analysis by Distance Indices (SADIE). This approach uses an algorithm in which observed events are iteratively displaced until they achieve a regular arrangement. As seen from the image, these distances are interpreted by totalling up the number of moves each event records until a regular pattern of events are achieved.

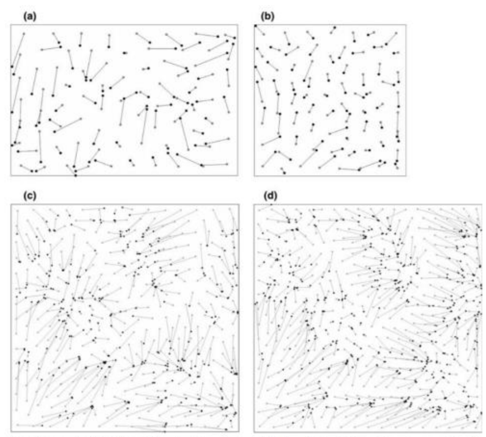


Fig. 8 'Initial-and-final' IAF plots for the univariate event sets: (**a**) NZ Trees, (**b**) Swedish Pines, (**c**) newly emergent bramble canes, (**d**) adult bramble canes. Black points represent the original location of the individuals and the grey points the regular pattern

Table 4 Summary statistics from the SADIE analyses; n=no. of events, OD=observed distance to regularity, RD = distance to regularity for CSR data, P_p=probability of as extreme aggregation under CSR, I_p=index of aggregation

Data set	n	OD	RD	P_p	I_p
NZ Trees	78	876.7	808.3	0.218	1.08
Swedish Pines	71	497.1	629.1	0.976	7.90
Brambles: emergent	359	27.1	14.4	0.002	4.03
Brambles: adult	464	33.6	15.8	0.002	1.89

Figure 7: Spatial Analysis by Distance Indices (SADIE)

[2] L. W. Perry, George & Miller, Ben & Enright, Neal. (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. Plant Ecology. 187. 59-82. 10.1007/s11258-006-9133-4.

Comparing the 2 of the plots, we can see from the top right plot, B, that its lines are not overlapping each other as frequently as the bottom left plot, C. More so, B has a shorter length of lines while C has a larger variance. In addition, observing the directions of the lines in C, they show some radiation outwards and away from distinct clusters.

Similarly, a summary statistic was produced in the literature. Pp is the probability as an extreme aggregation of CSR - offering a statistical result at the same time. Thus, the plots, along with the table, offer another way to observe the clustered distribution of events.

While this technique will not be carried into our research methodology, it makes aware the existence of other geospatial methods that could statistically interpret spatial point patterns.

3 DATA SETS

In order to perform our analysis, we have sourced the following data sets.

3.1 Airbnb Listings

This data[3] set comprises information on all Airbnb listings found within Seattle. They were last scrapped on 15 November 2018. We subsequently filtered for Downtown listings only – a total of 1,148 listings.

3.2 Common Place Name (CPN)

This data[4] was derived from the City of Seattle's Open Data Portal. It is a shapefile containing the names, addresses and coordinates of various types of locations. These CPNs range from key attraction sights, building names and natural attractions. It is from this dataset, where we derived the coordinate locations of our chosen attractions.

3.3 City Clerk Neighborhoods

A shapefile[5] featuring 20 large City Clerk neighborhood boundaries, along with their smaller neighborhood boundaries.

3.4 Zoning (Generalized)

A shapefile[6] featuring boundaries which are not labelled by neighborhood names, but rather, urban planning descriptions. i.e. Downtown, Major Institutions, Manufacturing/Industrial, Multifamily, Neighborhood / Commercial, Residential / Commercial and Single Family

4 METHODOLOGY

4.1 Distance Matrix

There were 2 attractions whose addresses were inaccurate from our verifications – Olympic Sculpture Park and Klondike Gold Rush. To overcome this hurdle, we used the "**OSMData**[7]" and "**tmaptools**[8]" packages to send an API query to derive the correct coordinates.

Subsequently, we used the "**OSRM**[9]" package to derive our distance matrix - distances between every listing to the 12 attractions. Certain requirements had to be met before executing our methods. Primarily, the listings and attractions objects needed to be in spatial data class – 'sp', and their projections needed to be in EPSG 4326 (a.k.a. WGS84).

Sending an API query, the OSRMRoute function helps pull the network distance between two points. Unfortunately, our team noticed that it pulls distances one at a time. There was potential to use an alternative function, OSRMTable, to conveniently derive the distance matrix. Unfortunately, the distance method is still in the 'beta' stages. This meant it could only derive the travel time matrix. We were considering the possibility of using travel time as a substitute for distance. However, in an openstreetmap forum[10], users mentioned that the travelling speeds which OSRM uses are mostly unknown. It was only at certain roads where travelling speed was tagged to the speed limit of that road.

Given the uncertainty, our team decided to stick with the OSRMRoute function. To conveniently pull our distance matrix, we used a recursive loop.

4.1.1 Collinearity between distances. After running a correlation plot between the distances, we gathered that there was a strong correlation between multiple attractions. Hence, we decided to group highly correlated attractions together and use the centroid of the group's attractions as the group's coordinates. This was derived using the gCentroid() function. This meant we now had 4 primary coordinates to work with:

1. **Northern Seattle Attractions (3):** Space Needle, Washington State Ferries, Olympic Sculpture Park
2. **Central Seattle Attractions (7):** Pike Place Market, Benaroya Concert Hall, Columbia Center, Seattle Art Museum, Seattle Public Library (Downtown), Seattle Great Wheel, The Seattle Aquarium
3. **Klondike Gold Rush**

[3] Inside Airbnb. Adding data to the debate.. (2019). Inside Airbnb. Retrieved from http://insideairbnb.com/get-the-data.html
[4] Data.seattle.gov. (2019). Retrieved from https://data.seattle.gov/Land-Base/Common-Place-Names-CPN-/599c-9ddc
[5] Seattle City GIS. (2019). City Clerk Neighborhoods. Retrieved from http://data-seattlecitygis.opendata.arcgis.com/datasets/b76cdd45f7b54f2a96c5e97f2dda3408_2
[6] Seattle City GIS. (2019). Zoning (Generalized). Retrieved from https://data-seattlecitygis.opendata.arcgis.com/datasets/a85e74dac41d43cab5a8b840558c4d77_3
[7] Cran.r-project. (2019). OSM data package. Retrieved from https://cran.r-project.org/web/packages/osmdata/osmdata.pdf
[8] Cran.r-project. (2019). Tmap tools package. Retrieved from https://cran.r-project.org/web/packages/tmaptools/tmaptools.pdf
[9] Cran.r-project. (2019). OSRM. Retrieved from https://cran.r-project.org/web/packages/osrm/osrm.pdf
[10] OpenStreetMap. (2019). [OSRM-talk] Accuracy of distance matrix calculation. Retrieved from https://lists.openstreetmap.org/pipermail/osrm-talk/2016-June/001254.html

4. **Washington State Convention Centre**

The metrics of our distance matrix were subsequently converted from kilometers to meters. This step was taken to create more variances in our data.

4.2 Geographical Accessibility Analysis

Geographical Accessibility Analysis was conducted as a form of data exploration to better understand a listing's location within Downtown Seattle.

While we had the individual distances between a listing and a key location, we wanted to find out the areas within Downtown Seattle that had higher accessibility based on all 12 key locations. Early on in our project, we hypothesized that areas with higher accessibility scores should be more attractive to visitors and ultimately, command higher Airbnb prices. Using Hansen's Potential Model and the distance matrix created earlier, we ran our analysis and visualised the accessibility score of all listings within Downtown Seattle.

To do so, the main package which we used was the "**SpatialAcc**" package.[11]

4.2.1 Concept. Hansen's Potential Model is defined as the potential of opportunities for interaction and is a measure of the intensity of the possibility for interaction. It originated from the gravitational principle and hypothesizes that the accessibility of a location is directly proportional to the size of the attraction and is inversely proportional to the distance to the attraction.[12]

$$P_i = \sum_j \frac{M_j}{d_{ij}^{\propto}} \qquad (1)$$

$\boldsymbol{P_i = Potential\ at\ point\ i}$
$\boldsymbol{M_j = The\ size\ (attraction) of\ centre\ j}$
$\boldsymbol{d_{ij}^{\propto} = the\ distance\ between\ i\ and\ j}$
$\boldsymbol{\propto = parameter, usually\ between\ 1\ and\ 2,}$
$\boldsymbol{reflecting\ the\ rate\ of\ increase\ of\ the\ friction\ of\ distance}$

4.2.2 Parameters. For our analysis, P_i is defined as the potential of each Airbnb listing within Downtown Seattle.

M_j is defined as the capacity of each key attraction within Downtown Seattle. To find the capacity of each key attraction, there were three factors that we considered. First, we looked at whether the location was enclosed or in the open. For areas that were enclosed, such as the Benaroya Concert Hall, calculated the capacity of the Concert Hall by multiplying the number of seats within each Concert Hall with the total number of Concert Halls within its vicinity. For open-air spaces, we searched for the highest demand ever achieved in that location and took a 10% markup as the maximum capacity the venue can have.

$d_{ij}^{\propto}$, the distance between listing i and location j was computed through the "**OSRM**" package mentioned under section 4.1.

Our selected parameter, ∝, for our analysis was the power function. We chose the power function over the exponential function due to the limitation of "SpatialAcc" package, as the package only allowed for computation using the power function. However, we believe that the exponential function was more applicable for our analysis. Given that the 12 locations selected within Downtown Seattle were key locations for visitors, we believe that these were places that a visitor would want to visit when in Seattle. Hence, the decrease in demand over distance should not be as sensitive and the exponential function should better reflect this decay in distance. With the power function, we used a value of $\frac{\propto}{d_\beta}$ = 0.5.

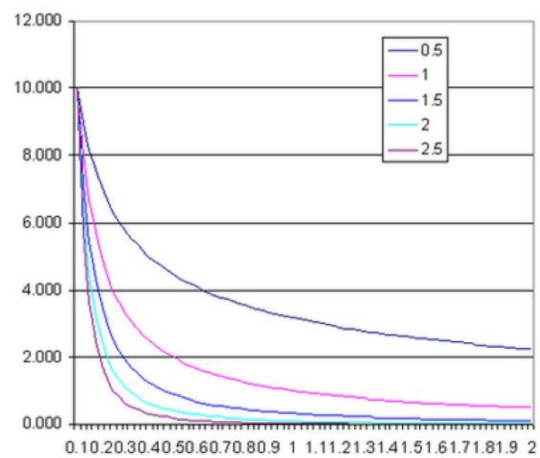


Figure 8: Graph of Inverse Distance Decay

4.3 Spatial Point Pattern Analysis

The use of spatial point pattern analysis (SPPA) is popular in the literatures of plant ecology and the studying of settlement distributions. For a moment in time, SPPA was neglected in geographical analysis purposes. But lately, with the developments of Geographical Information Systems (GIS) and the surge in geo-referenced databases, it has begun to attract renewed interest[13].

Applying to our review, the popularity of Downtown Seattle being densely littered with Airbnb listings makes it worthwhile to study the distribution pattern. From a research perspective, we hope to understand whether the distribution of Airbnb listings follows a random spatial distribution, or a regular or clustered pattern. From an application perspective, the utility of SPPA can help potential

[11] Kalogirou, S. (2017, August 04). Spatial Accessibility Measures [R package SpatialAcc version 0.1-2]. Retrieved from https://cran.r-project.org/web/packages/SpatialAcc/index.html

[12] Hansen, W. G. (1959, June). Accessibility and Residential Growth (pp. 1-51, Rep.). Massachusetts Institute of Technology.

[13] *Spatial point pattern analysis and its application in geographical epidemiology*. (1996). *Royal Geographical Society.*

Airbnb owners identify where best to locate their next listing location. Some Airbnb hosts may choose to own listings that are away from clusters of Airbnb listings, while some may prefer the company of a crowded Airbnb neighbourhood. Being in either sides triggers different business approaches to optimize the value earned in a listing. This could be in terms of prices, or non-price features such as amenities.

The main package which we will use is the "**spatstat**[14]" package.

4.3.1 Concept. SPPA uses hypothesis testing to statistically determine the behavior of point patterns over a map in one of the following 4 patterns depicted in figure 10. It is worth highlighting that a uniform distribution of points has a variance of zero. Whilst random and clustered distributions have variances that are close to one and greater than one respectively.

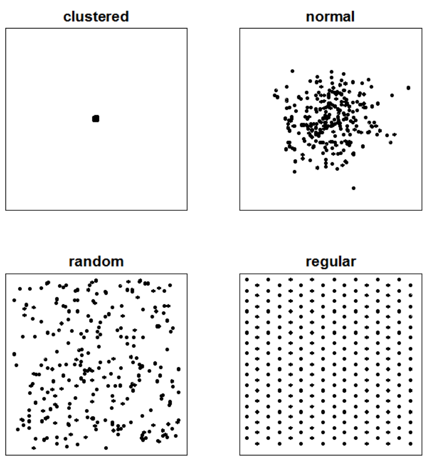


Figure 9: Four types of patterns (n = 256)

The null hypothesis usually describes the point patterns as following a complete spatial randomness (CSR)[15]. The alternate hypothesis then attempts to describe the points as not following a CSR. The actual set of point patterns are measured against a simulation of point patterns which follow a CSR pattern. A p-value is derived which is compared to the alpha value[16]. If the p-value is lesser than the alpha value, then we reject the null hypothesis of CSR and conclude that the points are not random.

Fundamentally, the SPPA is made of 2 forms of analysis - the first and the second order effect. Each has its own scopes which delivers separate types of observations. It is crucial to note that isolating the results of each is not a recommended approach as both are jointly used to arrive at a reasonable conclusion.

4.3.1.1 First Order. The first order effect analyses if observations vary from one place to another due to underlying properties. such as topology or other features pertaining to geography. For plant studies, the distribution of plant species could vary from the soil type. This involves the usage of the Quadrat Analysis and Kernel Density Estimation methods.

4.3.1.1.1 Quadrat Analysis. The use of quadrat analysis helps us to assess the extent to which point intensity is constant across a space that is independent of boundary restrictions. By retrieving the counts per quadrat[17], it helps to statistically prove if events are randomly distributed or not. Every quadrat analysis begins with a null and alternative hypothesis, and a level of significance. In our review, we will use the following:

H_0: The distribution of Airbnb listings are randomly distributed
H_0: The distribution of Airbnb listings are not randomly distributed
(Alpha-value = 0.001; Significance Level = 99.9%)

There is no formula that optimally determines the ideal number of quadrats. Too small a quadrat and each quadrat may contain very few points. Too large would be too many. This makes the determination of it rather arbitrary. To determine this sensibly, we have decided to go with 25 by 20 quadrats in the x and y directions respectively, and then compare the statistical results of the same analysis that is done with 20 by 15 and 30 by 25.

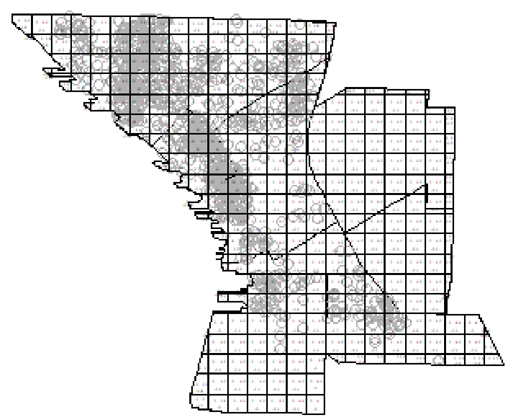


Figure 10: 25 by 20 quadrat count of listings in Downtown

From figure 10, there are multiple quadrats with zero events - a total of 167. This is not optimal. Each quadrat should have minimally 5 observations. Hence, we will not use the "chi-square" method, but rather, the "monte carlo" with nsim = 2999 (a total of 3000 simulations).

4.3.1.1.2 Kernel Density Estimation (KDE). The KDE is seen as an extension of the quadrat test. It calculates a local density for subsets of an area. What sets the KDE apart from the quadrat

[14] Cran.r-project. (2019). Spatstat package. Retrieved from https://cran.r-project.org/web/packages/spatstat/spatstat.pdf
[15] Complete Spatial Randomness: It assumes spatial events follow a homogeneous Poisson process over the area of study
[16] If Level of Significance = 99.9%, alpha value = 0.001.
[17] What is a quadrat? A simple grid shape that can be any shape and dimension.

intensity map is that the former has its subsets overlapping each other as it keeps moving. This moving frame is defined as a kernel.

What we end up with is a visualization that can correctly identify where hot spots of Airbnb listings are; a more accurate visual tool compared to the general observation of point patterns which gives different densities of listings when at different zoom levels. In the determination of its sigma[18], we have chosen bw.diggle as it determines the ideal bandwidth which minimises the mean-square error criterion using cross validation. Its formula is as follows:

$$M(\sigma) = \frac{MSE(\sigma)}{\lambda^2} - g(0) \qquad (2)$$

In determining the type of kernel function to use, the density function only offers four choices. We have decided to stick with "gaussian" because amongst the choices, it has the kernel smoothing that best generalizes the interaction between each listing.

4.3.1.2 Second Order. The second order effect analyses if observations (events) vary from one place to another due to its interaction with other events[19]; a distance-based method. For epidemiology, the distribution of contagious and non-contagious diseases could be dependent on one another.

There are multiple techniques under the second order effect. The G, F K, and L functions have results that are indifferent. Although in terms of execution, K and L uses a moving kernel. But what sets the L-function apart is its horizontal envelope which makes it easier to interpret. It is horizontal because it is Besag's transformation of Ripley's K-function. Hence, we have chosen to stick with the L-function to understand the distances where clusters or regular point patterns would be produced.

4.3.1.2.1 L – function. Since our analysis is over a set of points (pointwise), we would use a global = FALSE[20] in our envelope function. To ensure a 99.9% Significance Level, we set the number of simulations to be nsim = 1999. This follows the formula for the alpha of pointwise computations where alpha = 2 * nrank / (1 + nsim)[21]. Similar to the quadrat test, this is a statistical method that aims to disprove the null hypothesis of CSR. Hence, we'll once again adopt the same null and alternate hypothesis with the same level of significance.

H_0: The distribution of Airbnb listings are randomly distributed
H_0: The distribution of Airbnb listings are not randomly distributed
(Alpha-value = 0.001; Significance Level = 99.9%)

The results of an L-function plot would show us the distance between events where clusters, or regular patterns, would begin to form. In the interest to draw more insights about our data, we will perform this second order effect on a smaller neighborhood level to derive distinct distances for each.

4.4 Geographically Weighted Regression

Based on Tobler's First Law of Geography, a widely adopted principle is that everything is related with everything else, but *closer things* are *more related* to each other.

In geographically-weighted regression (GWR) models, heterogeneity in data relationships across space are examined. The geographical weighting of data implies that observations nearer each other have more influence in determining the local regression variables and hence the R^2. In the context of this project, we investigate the spatial variations in price explained by various attribute data of Airbnb listings and the OSM distances from the listings to attractions in Downtown Seattle.

The formula used for Geographically Weighted Regression is as follows:

$$y_i = \beta_{i0} + \sum_{k=1}^{m} \beta_{ik} x_{ik} + \varepsilon_i \qquad (3)$$

In this formula y_i represents the price of the listings at location i; x_{ik} represents the kth explanatory variable at location i; m represents the total number of explanatory variables; β_{i0} represents the intercept value at location i; β_{ik} is the local coefficient of the kth independent variable and ε_i represents the error term at location i[22].

For our project, the **'GWmodel'** in R was used to build the models. In the global regression model, which uses an Ordinary Least Square (OLS) method, a multiple linear regression was conducted, and all observations were weighted equally, without the influence of the listing's geographical location. However, in the local regression model where the spatial location of each listing plays a role, the output is determinant on the calibration of the model, which is discussed in the following sections.

4.4.1. Kernel and Bandwidth Selection. Since GWR measures the relationships for each location i, each set of regression coefficient is estimated by weighted least squares, which uses a matrix expression of:

$$\hat{\beta_i} = (X^T W_i X)^{-1} X^T W_{iy} \qquad (4)$$

[18] The standard deviation of an isotropic smoothing kernel.
[19] Gimond, M. (2019). *Chapter 11 Point Pattern Analysis | Intro to GIS and Spatial Analysis. Mgimond.github.io*. Retrieved 5 April 2019, from https://mgimond.github.io/Spatial/point-pattern-analysis.html
[20] Global = TRUE is the simultaneous method that is used to compute envelopes for polygons or lines spatial objects; objects that are not point objects.
[21] R Documentation. (2019). *envelope function | R Documentation*. Retrieved from https://www.rdocumentation.org/packages/spatstat/versions/1.59-0/topics/envelope
[22] Lu, B., Charlton, M., Harris, P. & Fotheringham, A.S. (2014). Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science*. Retrieved from http://dx.doi.org/10.1080/13658816.2013.865739

Here, X is the matrix of the independent variables and more important, W_i represents the diagonal matrix of the geographical weighting of each observed data for location i. This weighting scheme is dependent on the kernel function chosen by the user.

Bandwidth: Our model allows the user to select between fixed or adaptive bandwidth. Fixed bandwidth utilizes the same bandwidth in all local regression points, which may or may not capture the same number of points. On the other hand, adaptive bandwidth captures the same number of data points, allowing the bandwidth to adjust itself according to the density of the data. Looking at the distribution of listings in the figure below, it is evident that the distribution of listings is uneven across Downtown Seattle, indicating that a adaptive bandwidth would be more appropriate for the use of this project.

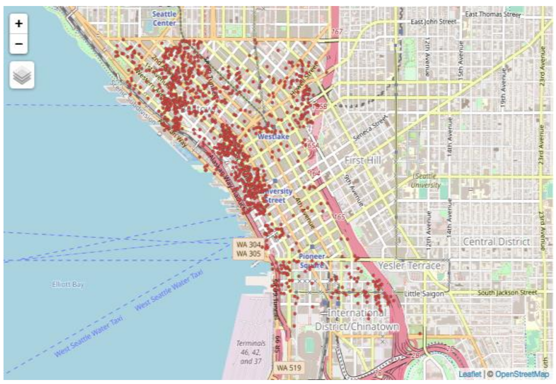
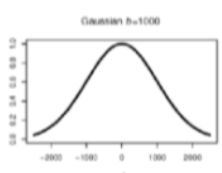
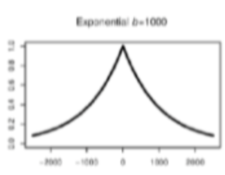
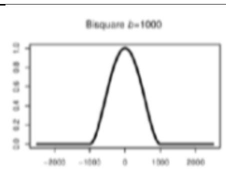
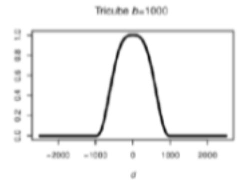
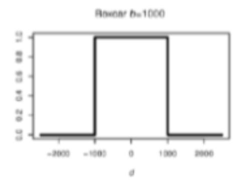


Figure 11: Distribution of Airbnb listings in Downtown Seattle

4.4.1.1 Kernel Functions.

Gaussian		Continuous function, decreases according to the Gaussian curve as distance between the observation points increase
Exponential		Continuous functions decreases according to the Exponential curve as distance between the observation points increase
Bisquare		Discontinuous function, weight decreases as the distance between observation points increase
Tricube		Discontinuous function, weight decreases as the distance between observation points increase
Boxcar		Simple discontinuous function that excludes observations further than distance b from the observation point

4.4.2. Variables and Feature Engineering

4.4.2.1 Dependent Variable - Price

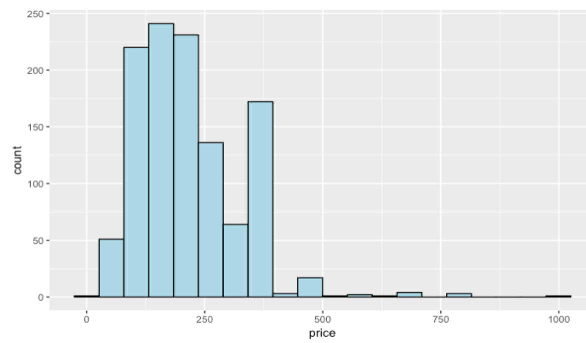


Figure 12: Histogram of Listing's Price

The histogram plot of the prices in the Downtown area is very much left-skewed. After removing a listing that is $0 in price, the summary statistics of Airbnb listing prices are as follows: *Min: $39.00, 25th percentile: $135.50, 50th percentile (Median): $199.00, 75th percentile: $275.00, Max: $999.000*

4.4.2.2 Independent Variables. Two key explanatory variable which are important to illustrate are: Accommodates and AmenitiesIndex. However, other independent variable such as property type, room type, number of bathrooms, number of reviews, minimum number of nights to book the listing, type of cancellation policy, neighbourhood and the walking distance to the centroid locations of 2 groups of attractions are also included in the model. These variables have been filtered out after testing for multi-collinearity.

4.4.2.2.1 Independent Variable - Accommodates

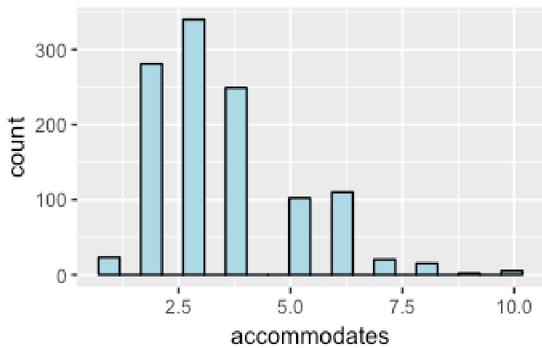


Figure 13: Histogram of Variable 'Accommodates'

This variable refers to the number of guests the listing can accommodate and serves as a proxy to replace the square-area of each listing, which was omitted since more than 80% of the observations were missing. Since the dependent variable for our model is price, it is essential to include elements which can suggest the floor area of the listing as a gauge.

4.4.2.2.2 Independent Variable – 'AmenitiesIndex'. Out of the 157 possible amenities found in downtown listings, we categorized the amenities into 3 groups to help with discriminating the prices in Airbnb listings better. This was inspired by GuestReady's article[23] on the must-haves and 'wow' factor extras. We then computed an index value 'AmenitiesIndex', which weighs the count of each category of amenities the listings have.

The 3 groups are namely;

(1) Basic Amenities: Includes essential amenities such as Wi-Fi, Heating, a Laptop friendly work space, Washer etc., including amenities that make up the criteria of what Airbnb defines as the minimal for a 'Business Travel Ready' listing. These basic amenities are present in at least 50% of listings.

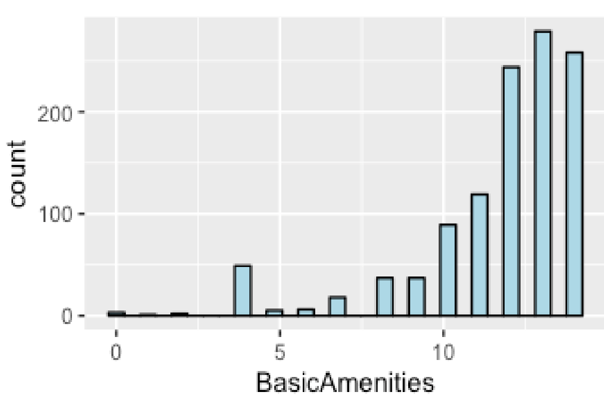


Figure 14: Histogram of the Basic Amenities in Listings

(2) Leisure Amenities: Beyond the basics, the article suggested that guests seeking a 'home away from home' comfort would require a few more amenities to make their stay more comfortable; which includes amenities such as cooking basics, a dishwasher, 24-hour check ins, a pool etc. In our categorization, we defined leisure amenities as those present in only 25-50% of listings

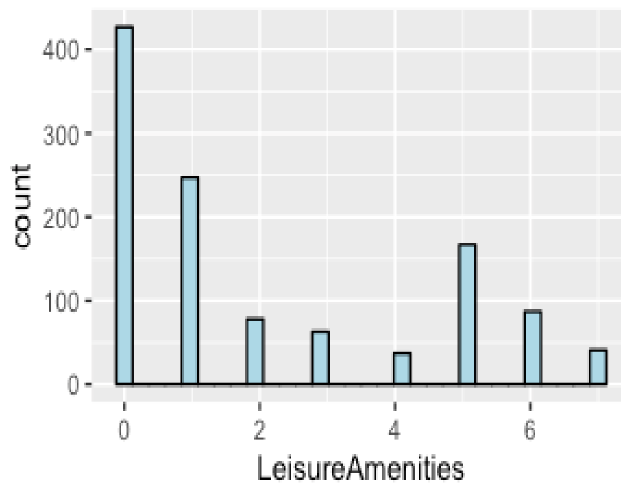


Figure 15: Histogram of the Leisure Amenities in Listings

(3) Luxury Amenities: This last group contain amenities and facilities which are rarer in nature, such as a hot tub, lock box, BBQ grill, and having a patio or balcony etc. These amenities are present in around 10-25% of the listings.

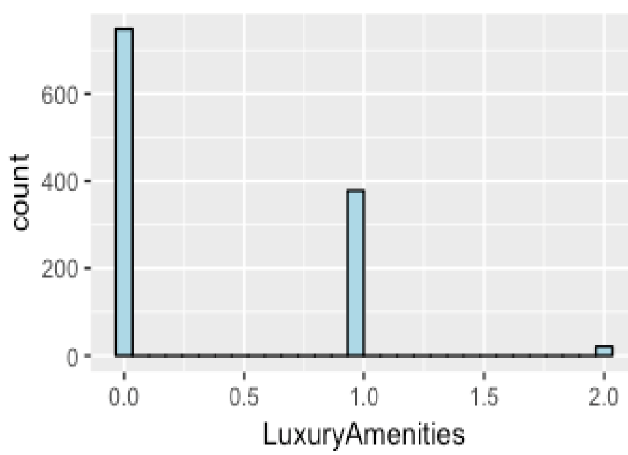


Figure 16: Histogram of the Luxury Amenities in Listings

From these three groupings, our understanding was that luxury amenities would best be able to differentiate the listings in terms of price. However, we cannot discount the presence of the basic and leisure amenities in providing a gauge for price as well. As such, we assigned a weight of 70% to luxury amenities, 20% to leisure

[23] GuestReady. (2017). Airbnb amenities: the must haves and the 'wow' factor extras. Retrieved from https://www.guestready.com/blog/airbnb-amenities-tips/

amenities and 10% to basic amenities in formulating the variable AmenitiesIndex.

5 RESULTS

5.1 Geographical Accessibility Analysis

5.1.1 *Geographical Accessibility Visualisation.* After defining the parameters for our analysis and running it with the "**SpatialAcc**" package, we visualised our results with R's "**tmap"** package.

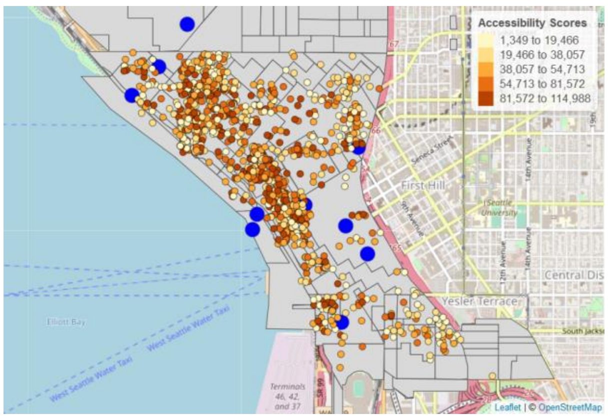


Figure 17: Geographical Accessibility Analysis of Downtown Seattle

From our visualisation, it is apparent that most of the darkest spots (dark brown colour) are found between the middle and upper areas of our plot, closer to the coastal areas of Downtown Seattle.

We believe that there are two reasons behind this. Firstly, the attractions found at the top of Downtown Seattle are largely open-air spaces that have capacities in the tens of thousands. The attractions found there are the Washington State Ferries, Space Needle and Olympic Sculpture Park. Within close proximity to these places and with these places having high capacities, it is expected that the listings surrounding them tend to have higher accessibility scores.

Secondly, most of the key attractions are found within central of Downtown Seattle. While these places individually do not have capacities as high as the ones above, but by having majority of attractions found within the area, the total capacity of the area is equally large. Hence, it is also expected that listings found within this area should have higher accessibility scores then others.

With our findings, we hypothesize that areas with higher accessibility scores should fetch higher listing prices, given their location superiority.

5.2 Spatial Point Pattern Analysis

5.2.1 Quadrat Test. The outcome of our quadrat tests produced significant results. Running a simulation of nsim = 2,999, we obtained a p-value of 0.0006667 for all 3 dimensions of quadrats which is lesser than our alpha value of 0.001. Hence, we can confidently reject the null hypothesis of CSR and accept the alternative hypothesis that the distribution of Airbnb listings follows either a regular or cluster pattern.

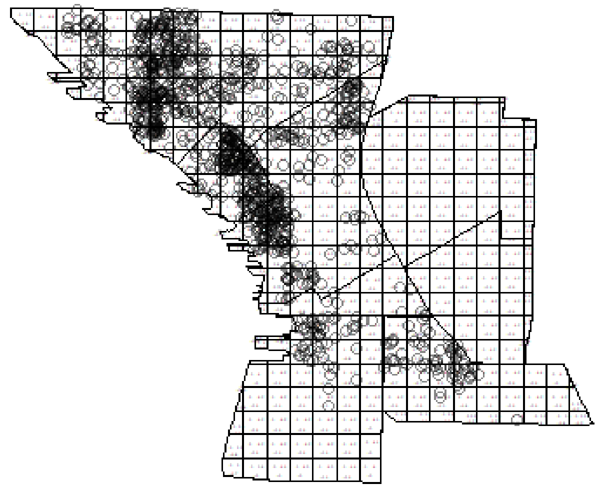


Figure 18: Quadrat test for 25 by 20 quadrats

5.2.2 Kernel density Estimation (KDE). Rescaling the distance of the shapefile to kilometres, we produced the following KDE raster file.

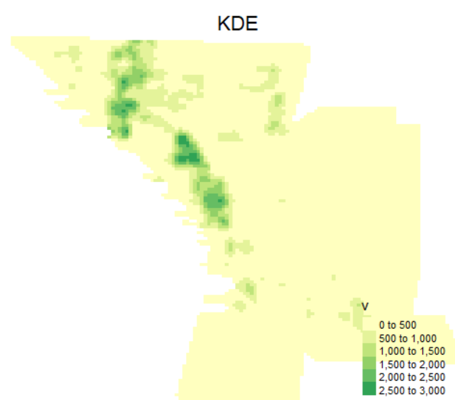


Figure 19: Kernel Density Estimation Map

5.2.3 L-function. After the necessary conversion of a "spatialpointdataframe" (spdf) to a "spatialpoint" (sp) only class, and then a "ppp" object, as well as converting the boundary base map shape file to an "owin" object, we obtained the following point pattern plots in each small neighbourhood (Figure 20).

From figure 20, International District displays the least clustered distribution of listings amongst the 5-small neighbourhood.

On the other hand, Belltown and Central Business District reflect one of the most clustered patterns. The clustering in Belltown tends to gravitate towards a linear vertical band while the Central Business District shows 2 possible cluster areas – each in the west and north. To derive our statistical results, we then performed the Lest envelope function. The results are seen in figure 21.

Here is how we interpret the L-function results. The middle dark grey band represents CSR. Its top boundary, known as the upper envelope, represents the 99.9% significance level in our review. Its bottom boundary, known as the lower envelope, is the 0.1% significance level. Any line which falls between the upper and lower envelops represents CSR and so we do not reject the null hypothesis. The key areas we focus in each plot are the points where the black line crosses above or below the upper or lower boundaries. The intersections mark the formation of clusters or regular patterns respectively.

Analysing each plot, we would notice that there is no significant pattern of regular distribution. Each neighbourhood has a different minimum distance of cluster formation and they are mostly very small. As expected, the least clustered region of International District has the largest cluster distance of approximately 20m. And the smallest is Belltown with approximately 2m; a reflection of stronger clustering patterns.

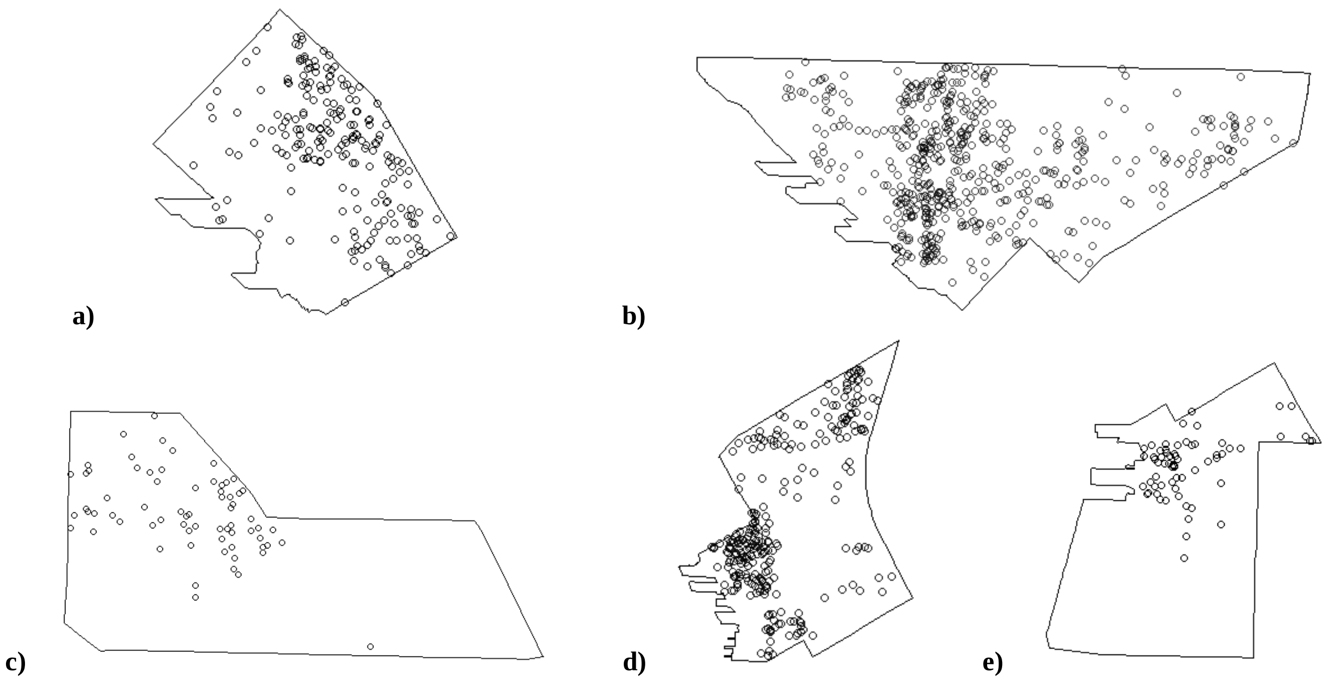


Figure 20: (First row, from left to right) Point pattern of a) Pike-Market (n = 203) and b) Belltown (n = 541). (Second row, from left to right) Point patterns of c) International District (n = 68), d) Central Business District (n = 272) and e) Pioneer square (n = 64)

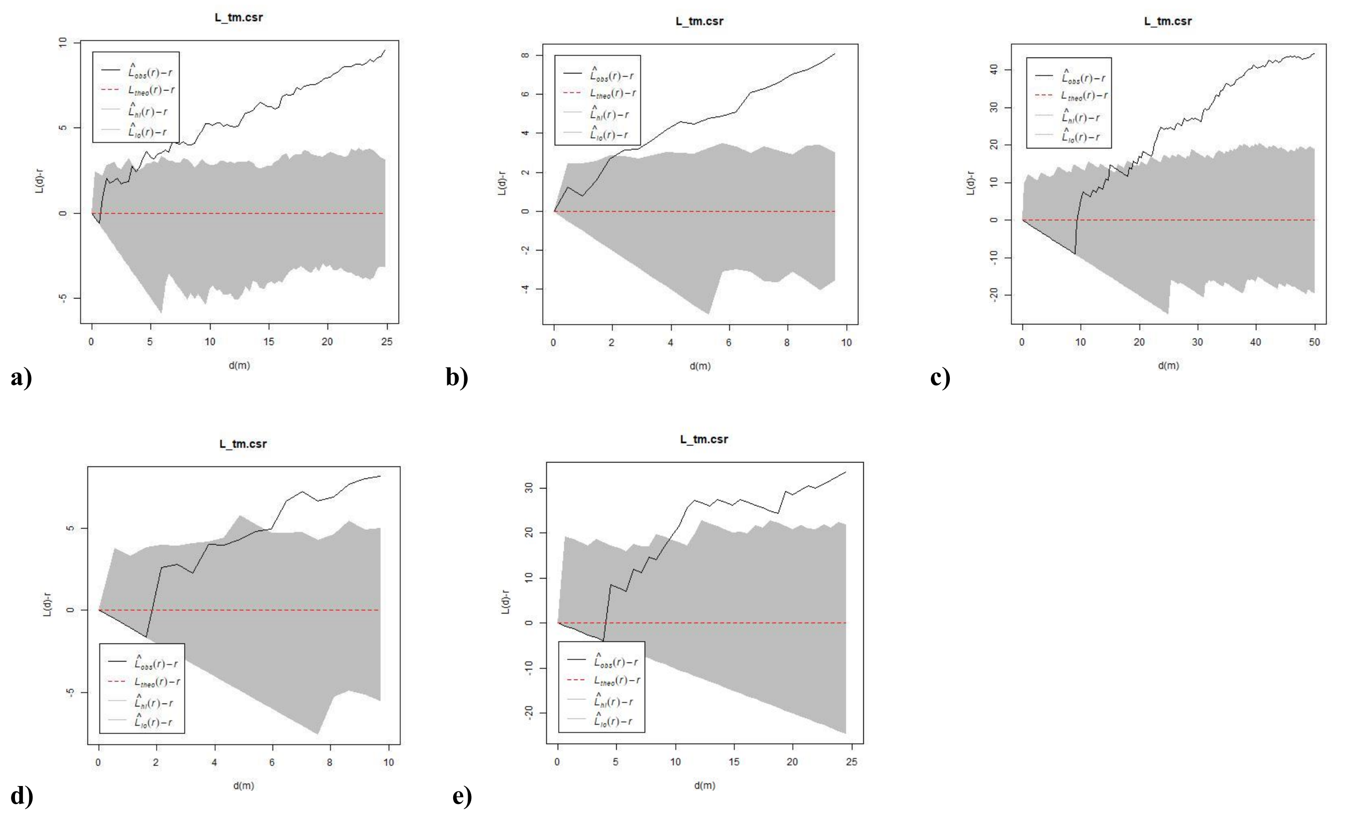


Figure 21: (First row, from left to right) L-Function of a) Pike-Market, b) Belltown, c) International District. (Second row, from left to right) L-Function of d) Central Business District and e) Pioneer square

One possible reason for the high listing counts in Belltown (n = 541) is likely attributed to it being the most densely populated neighbourhood in Seattle, Washington[24] and whose rents of residential areas are also low[25]. Sitting on a plot of land that is artificially flattened, it has become a place for popular restaurants, social activities such as nightclubs, bars, art galleries, and an urban environment with new condominium blocks[26][27]. More so, it is the neighbourhood that is closest to one of the most attractive places in Seattle – the Space Needle. And because Belltown is known for popular streets such as First Avenue[28], it could explain its clustered pattern of listings mimicking a vertical corridor. These features of low costs, and abundance of tourists features, and an urban atmosphere, makes a strong case for Belltown's high listing counts.

On the other hand, International district (also known as Chinatown) sees the second lowest number of listing counts (n = 68), with the largest minimum distance for clusters. A possible reason for this lacklustre listing count is possibly because of its unpopularity. According to trip advisor's ratings and reviews, the location is largely reviewed as being less urbanized and with poor building conditions[29]. In addition, it is situated further away from most attractions. These are possible explanations for the low popularity and hence the low supply of listing.

5.3 Geographically Weighted Regression

Our application allows users to examine the distributions of each variable before selecting the variables to input into the GWR. As they select the variables, the application also updates a correlation matrix and data table for users to examine. After accounting for potential multi-collinearity, users can then perform a geographically weighted regression. Under the GWR tab, users can explore the R^2 and Intercept plots, VIF values between the

[24] Wikipedia. (2009). Belltown, Seattle. Retrieved from https://en.wikipedia.org/wiki/Belltown,_Seattle

[25] Seattle's Neighborhoods. Retrieved from http://www.nytimes.com/fodors/top/features/travel/destinations/unitedstates/washington/seattle/fdrs_feat_143_4.html?pagewanted=3

[26] Seattle's Neighborhoods. Retrieved from http://www.nytimes.com/fodors/top/features/travel/destinations/unitedstates/washington/seattle/fdrs_feat_143_4.html?pagewanted=3

[27] Wikipedia. (2009). Belltown, Seattle. Retrieved from https://en.wikipedia.org/wiki/Belltown,_Seattle

[28] Grambush, J. (2019). The 10 Coolest Streets in Seattle. Culture Trip. Retrieved 10 April 2019, from https://theculturetrip.com/north-america/usa/washington/articles/the-10-coolest-streets-in-seattle/

[29] Tripadvisor. (2019). Chinatown International District (Seattle) - 2019 All You Need to Know Before You Go (with Photos) - Seattle, WA. Retrieved from https://www.tripadvisor.com.sg/Attraction_Review-g60878-d269463-Reviews-Chinatown_International_District-Seattle_Washington.html

independent values, the local Moran I's test statistics as well as the output of the GWR.

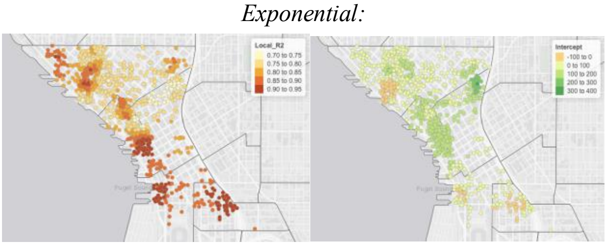
Amongst the variables provided, there are a few which are categorical whilst most are numerical. Studies show that caution needs to be exercised when including categorical data since there is a strong risk of encountering local multicollinearity issues where the categories cluster together spatially [30]. However, as these categorical variables are significant in explaining the variation in prices in the global regression model, we decided to include them in the application, which is only available for kernel functions gaussian and exponential.

	Without categorical variables	With categorical variables
Multiple R^2	31.23%	51.93%
Adjusted R^2	30.81%	50.64%

From the table above, it is evident that categorical variables such as neighbourhood (e.g. Pike Market, CBD), property type (e.g. condominium, RV), room type (e.g. entire apartment, private room) and type of cancellation policy (e.g. strict 30 days) play a significant role in explaining the variation in prices. In fact, R^2 increased by 20% to 50.64% with the inclusion of these variables.

The inclusion of these categorical variables was however only applicable to the gaussian and exponential kernel functions. Using an adaptive bandwidth, the plots of R^2 (left) and Intercept values (right) are as shown below:

Exponential:



Gaussian:

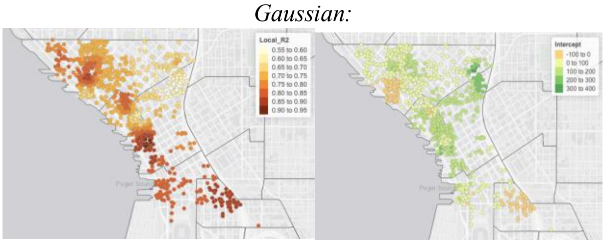


Figure 22: Geographically Weighted Regression Plots

Comparing the R^2 plots above, it is evident that the exponential kernel function is, on the whole, more adept at explaining the variation in prices in this local regression model, this can be seen from the lower bound of the R^2 values, where the model with exponential kernel function reflects a minimum of 70%, while the gaussian reflects a minimum of 55%.

Additionally, the plots also show a distinctive difference in the spread of R^2 values. In the gaussian plot, R^2 values tend to take on similar values over a wider distance compared to the exponential kernel function. This can be explained by the difference in the weighting functions used. Comparing the gaussian and exponential curve, the exponential weight function adopts a much steeper decrease compared to gaussian.

Shifting the focus to the Intercept plots, it can be observed that the intercept values are generally on the higher end for listings near to Washington Convention Center, and lower for listings near the pier. This could be justified as Seattle is a magnet for business travellers [31], and being the largest convention centre in Seattle, listings near this location where business travellers tend to congregate would give it prestige by nature of its location.

6 DISCUSSION

From our findings, users can expect Downtown Seattle to be a large neighbourhood that is densely populated with Airbnb listings. These listings show a noticeable cluster corridor pattern in each smaller neighbourhood. Interestingly, this pattern extends vertically, while being close to the harbour that is on the East of Seattle, as observed from the Kernel Density Plot (Figure 19). Complementing with the Geographical Accessibility analysis (Figure 17), Airbnb hosts who find themselves located in this cluster can take advantage of this accessibility information and describe themselves better by indicating their higher accessibility to nearby attractions in Seattle, as well as having homes with "a breeze-y view". These could be attractive features for a certain demographic of travellers.

Unfortunately, Airbnb hosts will find it a challenge to secure a listing location that is far or separated from any nearby listing. From the performed SPPA, the results exhibit a highly small minimum cluster distances for each neighbourhood. It almost always means that a location would certainly be a part of a significant cluster. This makes Downtown Seattle a competitive location for Airbnb hosts.

Fortunately, given the wide range of amenity information that has been collected by Airbnb, users can rely on our RShiny [32] platform to understand the neighbourhood around them; enabling users to differentiate themselves in terms of pricing or features.

One of the interesting findings worth mentioning is how the lack of and profound presence of various amenities has created a

[30] ArcGis (n.d.). Geographically Weighted Regression (GWR). Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/geographically-weighted-regression.htm
[31] Upside. (2018). Mixing Business with Leisure in Seattle. *Upside Business Travel*. Retrieved from https://upside.com/blog/category/destination-resources/seattle-bleisure

[32] Too High, Too Low, or Just Right?. (2019). Businessmafia-is415.shinyapps.io. Retrieved from https://businessmafia-is415.shinyapps.io/businessmafia_is415_rshiny_v2/

category of amenity features which our team differentiates as basic, leisure and luxury amenities. Their categorization is based on their availability percentage amongst the listings, and then refined by our human interpretation.

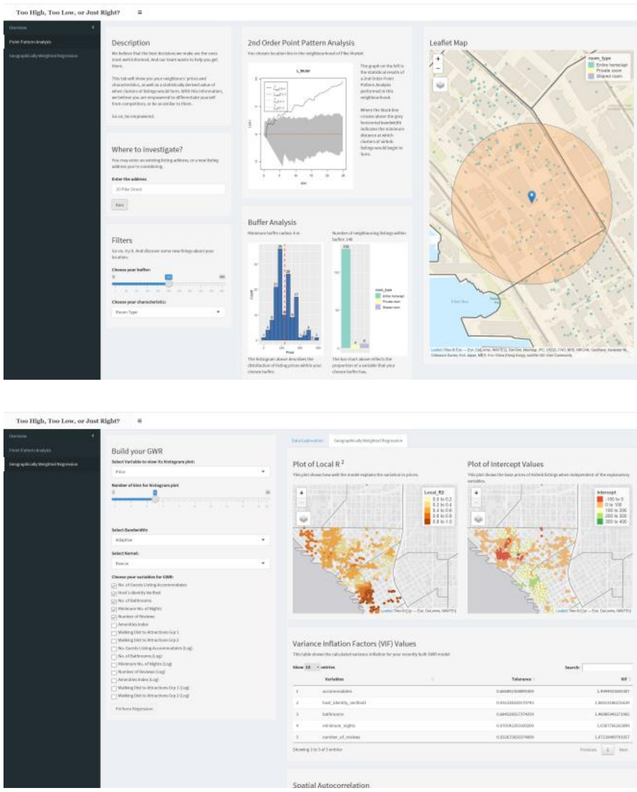


Figure 23: R Shiny Web Based Application Platform

From the geographically weighted regression model, it is also evident how big a role the location plays in determining the price Airbnb owners can set. From the GWR, Airbnb hosts would be able to gain a better understanding of how they can potentially price their listings, by virtue of its location (for example, being near prime locations such as the Washington Convention Centre) and the different amenities they can provide guests with.

By exploring more about the features which one has or lacks, we believe our users can have a more perfect information about the Airbnb market around them. The should allow hosts to compete on better terms which would in turn spill over benefits to potential travellers heading to Downtown Seattle.

FUTURE WORK

Given the opportunity, we believe there is ample headroom for additional analysis to be built upon our model. Presently, our model includes the fixed locations of attractions. However, across our research, we came across an interesting paper[33] which used the density of the geolocation of photos to indicate where streets or clusters of attractions are located at. Our team finds this as a novel approach in including more variables into our model. It could offer new insights and ensure a more accurate methodology. Unfortunately, the site[34] used to extract this data is no longer available – it has been purchased by Google. If alternative sites are sourced, we highly encourage users to adopt such an approach to broaden the definition of attractions.

Currently, the GWR uses the listings' features and distances to centroid locations of attractions to evaluate the listing's price. However, an extension of the GWR could include adding the proximity of the listings to even more features such as gyms, supermarkets and transport facilities to give a more in depth review of the listings' potential prices.

Lastly, in terms of transferability, we believe our geospatial model can be used to also support the housing rental market in Singapore. Although Singapore has regulations which prohibit the setting up of Airbnb listings, she is expected to see higher retail rents in the next couple of years due to the higher tourism and domestic spending in 2018[35]. With this foreseeable demand, learning how geospatial properties, if any, affects pricing and sets rental listings apart in Singapore could empower these local owners.

ACKNOWLEDGMENTS

We would like to thank Dr. Kam Tin Seong, the Associate Professor of Information Systems (Practice), who is also our guiding professor.

Dr Kam, thank you for your endless support and dedication to your students and the discipline. Your wisdom, knowledge and zeal for the geospatial discipline inspires us to not only do more, but *be more*. And along with your wealth of experience, you continually provoke us to consider new perspectives and approaches for our research. You have paved a way for us and it is truly an honor to be students under your care. We wish you the very best in your career and may you continue to inspire others along the way.

REFERENCES

ArcGis (n.d.). Geographically Weighted Regression (GWR). Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/geographically-weighted-regression.htm

Arvix. (2019). Airbnb in tourist cities: comparing spatial patterns of hotels and peer-to-peer accommodation. Retrieved from https://arxiv.org/ftp/arxiv/papers/1606/1606.07138.pdf

[33] Arvix. (2019). Airbnb in tourist cities: comparing spatial patterns of hotels and peer-to-peer accommodation. Retrieved from https://arxiv.org/ftp/arxiv/papers/1606/1606.07138.pdf

[34] Panoramio. (2019). Panoramio. Retrieved from https://www.panoramio.com/

[35] Mui, R. (2019). Singapore office, retail rents to see strong growth over next few years: report. The Business Times. Retrieved from https://www.businesstimes.com.sg/real-estate/singapore-office-retail-rents-to-see-strong-growth-over-next-few-years-report

Belltown, Seattle. (2009). En.wikipedia.org. Retrieved from https://en.wikipedia.org/wiki/Belltown,_Seattle

Cran.r-project. (2019). OSM data package. Retrieved from https://cran.r-project.org/web/packages/osmdata/osmdata.pdf

Cran.r-project. (2019). OSRM. Retrieved from https://cran.r-project.org/web/packages/osrm/osrm.pdf

Cran.r-project. (2019). Spatstat package. Retrieved from https://cran.r-project.org/web/packages/spatstat/spatstat.pdf

Cran.r-project. (2019). Tmap tools package. Retrieved from https://cran.r-project.org/web/packages/tmaptools/tmaptools.pdf

Data.seattle.gov. (2019). Retrieved from https://data.seattle.gov/Land-Base/Common-Place-Names-CPN-/599c-9ddc

Gimond, M. (2019). Chapter 11 Point Pattern Analysis | Intro to GIS and Spatial Analysis. Mgimond.github.io. Retrieved from https://mgimond.github.io/Spatial/point-pattern-analysis.html

Grambush, J. (2019). The 10 Coolest Streets in Seattle. Culture Trip. Retrieved from https://theculturetrip.com/north-america/usa/washington/articles/the-10-coolest-streets-in-seattle/

GuestReady. (2017). Airbnb amenities: the must haves and the 'wow' factor extras. Retrieved from https://www.guestready.com/blog/airbnb-amenities-tips/

Gupta, S., Peshin, A., & Agrawal, A. (2019, January 04). Airbnb Rental Listings Dataset Mining. Retrieved from https://towardsdatascience.com/airbnb-rental-listings-dataset-mining-f972ed08ddec

Hansen, W. G. (1959). Accessibility and Residential Growth (pp. 1-51, Rep.). Massachusetts Institute of Technology.

Inside Airbnb. Adding data to the debate. (2019). Inside Airbnb. Retrieved from http://insideairbnb.com/get-the-data.html

Kalogirou, S. (2017). Spatial Accessibility Measures [R package SpatialAcc version 0.1-2]. Retrieved from https://cran.r-project.org/web/packages/SpatialAcc/index.html

L. W. Perry, George & Miller, Ben & Enright, Neal. (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. Plant Ecology. 187. 59-82. 10.1007/s11258-006-9133-4.

Lu, B., Charlton, M., Harris, P. & Fotheringham, A.S. (2014). Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. International Journal of Geographical Information Science. Retrieved from http://dx.doi.org/10.1080/13658816.2013.865739

Mui, R. (2019). Singapore office, retail rents to see strong growth over next few years: report. The Business Times. Retrieved from https://www.businesstimes.com.sg/real-estate/singapore-office-retail-rents-to-see-strong-growth-over-next-few-years-report

OpenStreetMap. (2019). [OSRM-talk] Accuracy of distance matrix calculation. Retrieved from https://lists.openstreetmap.org/pipermail/osrm-talk/2016-June/001254.html

Panoramio. (2019). Panoramio. Retrieved from https://www.panoramio.com/

R Documentation. (2019). envelope function | R Documentation. Retrieved from https://www.rdocumentation.org/packages/spatstat/versions/1.59-0/topics/envelope

Royal Geographical Society. (1996). Spatial point pattern analysis and its application in geographical epidemiology. (1996).

Seattle City GIS. (2019). City Clerk Neighborhoods. Retrieved from http://data-seattlecitygis.opendata.arcgis.com/datasets/b76cdd45f7b54f2a96c5e97f2dda3408_2

Seattle City GIS. (2019). Zoning (Generalized). Retrieved from https://data-seattlecitygis.opendata.arcgis.com/datasets/a85e74dac41d43cab5a8b840558c4d77_3

Seattle's Neighborhoods. Retrieved from http://www.nytimes.com/fodors/top/features/travel/destinations/unitedstates/washington/seattle/fdrs_feat_143_4.html?pagewanted=3

Too High, Too Low, or Just Right?. (2019). Retrieved from https://businessmafia-is415.shinyapps.io/businessmafia_is415_rshiny_v2/

Tripadvisor. (2019). Chinatown International District (Seattle) - 2019 All You Need to Know Before You Go (with Photos) - Seattle, WA. Retrieved from https://www.tripadvisor.com.sg/Attraction_Review-g60878-d269463-Reviews-Chinatown_International_District-Seattle_Washington.html

Upside. (2018). Mixing Business with Leisure in Seattle. Upside Business Travel. Retrieved from https://upside.com/blog/category/destination-resources/seattle-bleisure

Wikipedia. (2009). Belltown, Seattle. Retrieved from https://en.wikipedia.org/wiki/Belltown,_Seattle