A comparative Study of Spatial and Non-spatial modelling in price prediction

A Case Study of Airbnb Price prediction in Amsterdam

Hassan Ali, Sahar Pourahmad, Wiktoria Libera

Utrecht Univeristy

INFOMSSML

Introduction

In recent years, the sharing economy and peer-to-peer activities have taken the world by storm. The concept of facilitating collaboration and the shared use of idle assets between individuals connected through community-based online platforms is gaining traction (Team, 2021; Team, 2020), with Airbnb being a prominent example. This company, founded in 2007, allows owners (hosts) to rent out their vacant rooms or houses to other people (guests) (Airbnb, 2023). It has grown rapidly since it was founded, recently reaching 1.4 billion visitors. Despite the fact that it can benefit local economies, Barker (2020) reports on the 'Airbnb effect,' which drives down house prices in tourist areas. Furthermore, research has shown that multiple factors can influence the price. They are divided into three groups: structural, quality-signalling, and neighbourhood variables (Contu et al., 2022). The first category contains property-related characteristics such as the number of bedrooms, bathrooms, and size (Gyódi & Nawaro, 2021; Contu et al., 2022). The second group concerns the host's reputation, cleanliness, rating, or communication abilities. The final category, 'neighbourhood variables', takes into account more spatial variables like distance to the city centre or train station (Kuby & Yoo, 2020). Furthermore, cultural amenities and tourist attractions are important in the area (Sada, 2021). We can also consider accessibility to nightlife, restaurants, and convention centres (Nault, 2019; Contu et al., 2022). Furthermore, tourists consider the area's greenness, which influences aesthetics and recreation (De Jonge, 2021). The last spatial factor is the safety and cost of the neighbourhood, which influences the price and tourist choice. All of those cause the neighbourhood area to be more attractive.

The "neighbourhood variables" group is especially important because it can have a spatial effect on the price as well as how the factors correlate and differ across locations. There are several approaches to factor analysis. Using the Hedonic Price Model (HPM) is one of the most popular methods. The intrinsic value of the attributes is used to calculate the transactional price (Monson, 2009). This approach is global in scope and does not take spatial dependency or heterogeneity into account (Gyódi & Nawaro, 2021). Scholars use Globally Weighted Regression (GWR) to account for variable local variation (Brunsdon et al., 2010).

The identification of price-influencing factors may assist policymakers in developing the necessary measures to limit the "Airbnb effect". As a result, this study focuses on the analysis of various 'neighbourhood variables' using GWR and HPM. With that said, the following research question will be investigated:

Is the GWR model better than the HPM model at predicting the price of Airbnb housing when 'neighbourhood variables' are considered?

This paper aims to present the data collection and processing process used to create the HPM and GWR models. Furthermore, it goes on to analyse the results and spatial effects of both models and finds insights for policymakers as well as the housing market.

Methodology

The Study Area

Amsterdam, the capital of the Netherlands, is in the province of North Holland. It is a popular tourist destination in North-Western Europe, with 4.5 million visitors each year (*General Information About Amsterdam | AmsterdamTourist.info*, 2020). It is a city that offers a wide variety of activities for all types of visitors, and everyone will find something to their liking within a short bike ride around the city (Norbert, 2023). **Error! Reference source not found.** shows the location of Amsterdam.

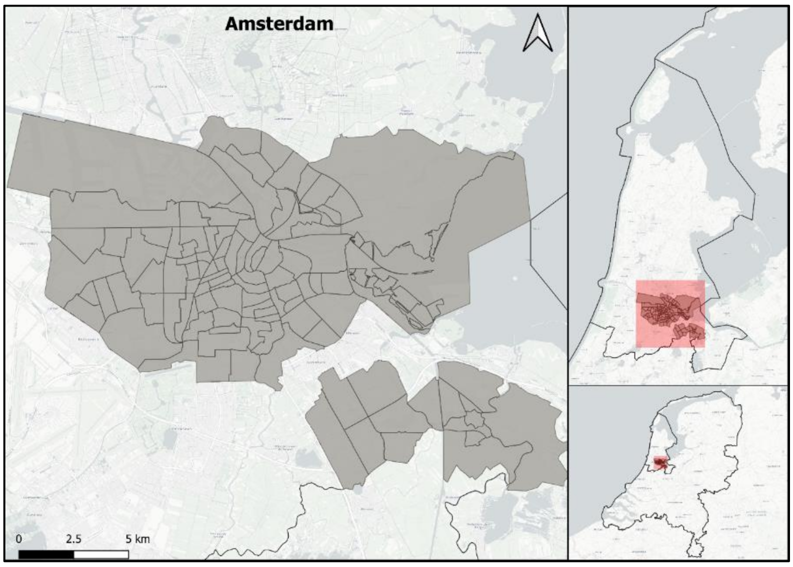

Figure 1 Amsterdam City, The Netherlands.

Data Collection and Aggregation

This project made use of data from the website Inside Airbnb, Amsterdam municipality, and OpenStreetMap. Inside Airbnb offered geographical coordinates as well as accommodation-specific information. OSM data was retrieved and the distance to the nearest facilities to each listing was determined. Additionally, the percentage of available greenspace, safety index, and cultural amenities was retrieved from Amsterdam municipality data at the neighbourhood level. To construct a single dataset, the variables were integrated at the listing level. Figure 2 depicts the complete procedure, and Appendix 1 has further information.

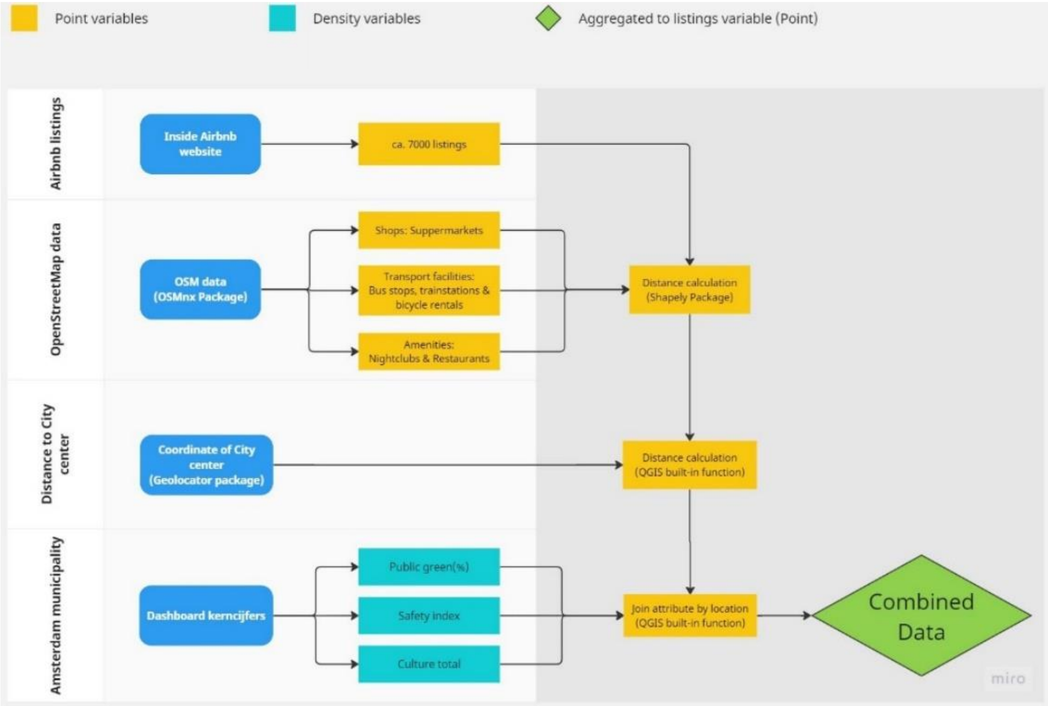

Figure 2. Data extraction and aggregation process.

Data Pre-Processing

Considering that the data is gathered from various sources, it must be cleaned. The first step is to deal with missing data. Some of them were not included in the model and thus left in the set to avoid limiting the number of observations. The statistical quantile ranges were used to identify the outliers, resulting in removing four outliers in price from the dataset. This was step followed by categorical variable encoding, data scaling and normalization and feature selection. It was accomplished through the use of the literature, Lasso, and correlation.

Modelling

To begin, the Hedonic Price Model (HPM) was used to model the impact of various price variables. This is a type of multiple regression model that is used to establish the influence of tangible, intangible and external factors on price (Monson, 2009). The following equation was created to represent the relationships. It contains both listing variables and spatial variables that are treated as normal. Later, the HPM was tested for the assumptions. To do so, graphs were created, and the variance inflation factor (VIF) test was used to check for multicollinearity.

$$
\begin{aligned}
Price = \; & \beta_0 + \beta_1 \, \text{neighbourhood}\_cleansed + \beta_2 \, property\_type + \beta_3 \, accommodates \\
& + \beta_4 \, review\_score\_rating + \beta_5 \, restaurants + \beta_6 \, \text{nightclub} \\
& + \beta_7 \, \text{suppermarket} + \beta_8 \, \text{park} + \beta_9 \, \text{trainstation} + \beta_{10} \, \text{busstop} \\
& + \beta_{12} \, \text{bicycle\_rental} + \beta_{13} \, \text{city\_center} + \beta_{14} \, \text{culture\_to} + \beta_{15} \, \text{public\_gre} \\
& + \beta_{16} \, \text{Safety\_ind} + \varepsilon
\end{aligned}
$$

The second model was Geographically Weighted Regression (GWR). GWR takes into account local autocorrelation and variable heterogeneity based on location. First, the weight matrix with k-nearest neighbours was created using HPM, with k=3. Morans' I test was then used to determine if the spatial autocorrelations are significant. Finally, the data was modelled using the fixed and adaptive kernel.

Results

Hedonic Price Model

The HPM results are unsatisfactory, with an R-squared of only 37%. Additionally, only a few variables associated with neighbourhoods are statistically significant (p-value less than 0.05) (Table 1[1]).

Table 1. Condense results from Hedonic Price Model

Characteristic	Beta	95% CI[1]	p-value
accommodates	0.45	0.42, 0.47	<0.001
review_scores_rating	0.06	0.04, 0.08	<0.001
restaurant	-0.02	-0.05, 0.02	0.3
nightclub	-0.04	-0.09, 0.02	0.2
supermarket	0.00	-0.03, 0.04	0.9
park	0.01	-0.02, 0.04	0.4
trainstation	0.05	0.01, 0.10	0.016
busstop	0.02	0.00, 0.04	0.082
bicycle_rental	-0.06	-0.12, 0.01	0.088
city_center	-0.26	-0.37, -0.15	<0.001
culture_to	0.00	-0.03, 0.03	>0.9
public_gre	0.02	-0.01, 0.06	0.2
Safety_ind	-0.04	-0.07, -0.01	0.014

[1]CI = Confidence Interval

Only accommodates, review_score_rating, trainstation, city_center, and Safety_ind are statistically significant with a p-value less than 0.05. However, the accommodation has the greatest influence on price; for every increase in the number of people that the place can accommodate, the price rises by 0.45. The second most influential factor is city_center, which means that as the distance to the centre increases, the price decreases by 0.24.

The model violates some of the main assumptions (Figure 3) specially homoscedasticity. Furthermore, the variance of the residuals is not constant, and vary a lot, which indicates heteroscedasticity. This can be clearly visible on the Normal Q-Q and Residuals vs Fitted plots.

[1] Because this model includes categorical variables such as property_type and neighbourhood_cleansed, the summary is too large to report in this section, and we are not as interested in them, so it is moved to Appendix 2.

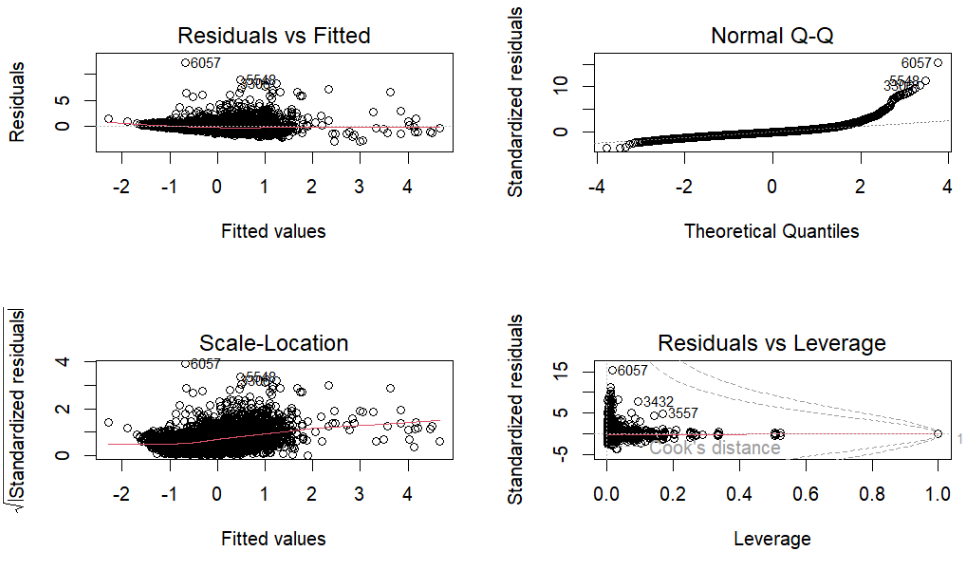

Figure 3. Graphs for the analysis of assumptions.

Another aspect is the multicollinearity, which indicates that some of the predictors can be highly correlated with other ones. With the VIF test, it turned out that this is the case for city_center and bicycle_rental variables.

To explore the presence of spatial autocorrelation in the data, the Monte Carlo test was done using the global Morons's I. As it can be seen in Table 1, the p-value is less than 0.05, so we reject the null hypothesis that there is no significant autocorrelations in the price variable, thus the spatial modelling needs to be done.

Table 2. Results of Monte Carlo test using Morans' I

Monte -Carlo simulation of Moran I	
data	sdf$price
Weights:	sdf_KNN_w
Number of simulations +1:	3000
statistic	0.2484
observed rank	3000
p-value	0.0003333
alternative hypothesis:	greater

To indicate the places where autocorrelations are present, the local autocorrelation analysis was done and plotted (Figure 4). Significant spatial clusters were discovered, mostly in the centre of the city.

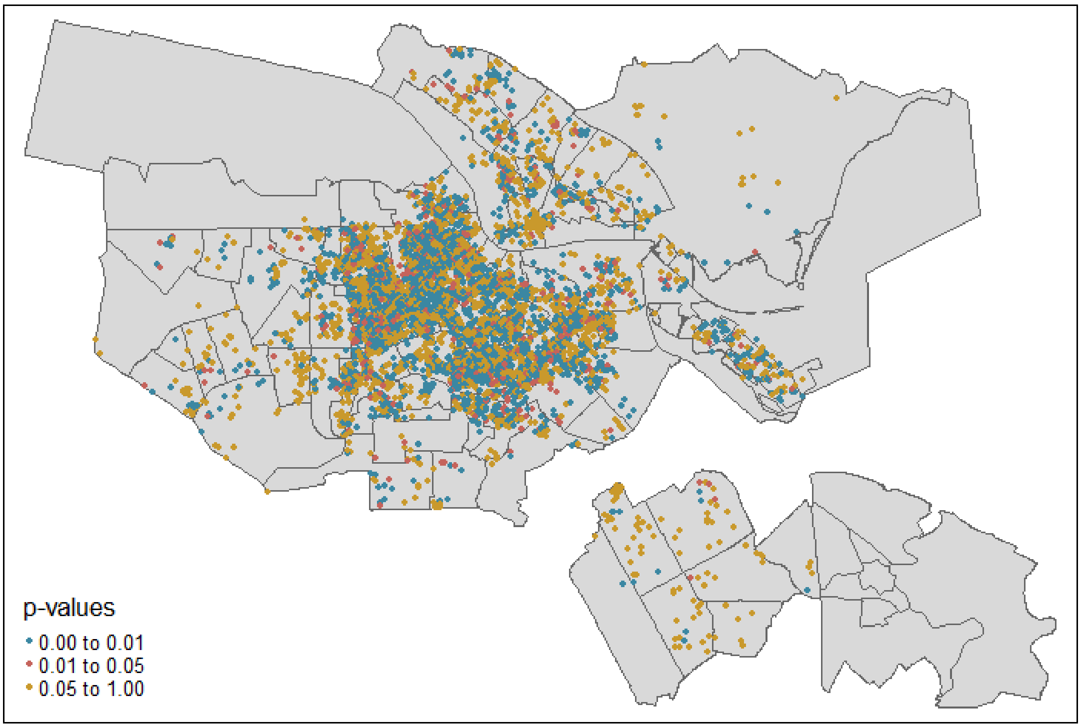

Figure 4. P-values for LISA.

Geographically Weighted Regression

GWR overcomes spatial nonstationary by allowing modelled relationships between a dependent variable and a set of predictors that vary across space. The GWR was fit using both fixed and adaptive kernels. The adaptive kernel performed better in terms of AIC and adjusted R-squared. An R-squared of 49% was obtained with the optimal number of neighbours of 215.

The results of GWR show that the variables vary across locations because the first and third quartiles are significantly different. To help understand the local R-squared, Figure 5 was created, which demonstrates that the variables do not explain the price in the same way in all locations.

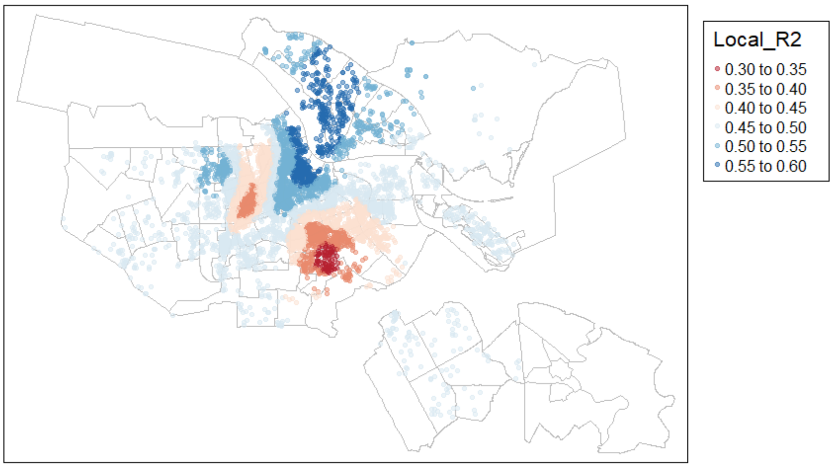

Figure 5. Local R-squared.

The GWR fits the data better in Amsterdam's centre and surrounding neighbourhoods. The area where GWR appears to perform the worst appears to be to the south-east of the city center, indicated by red.

Figure 6 shows presence of spatial nonstationarity for several variables, that were significant for both HPM and global regression[2].

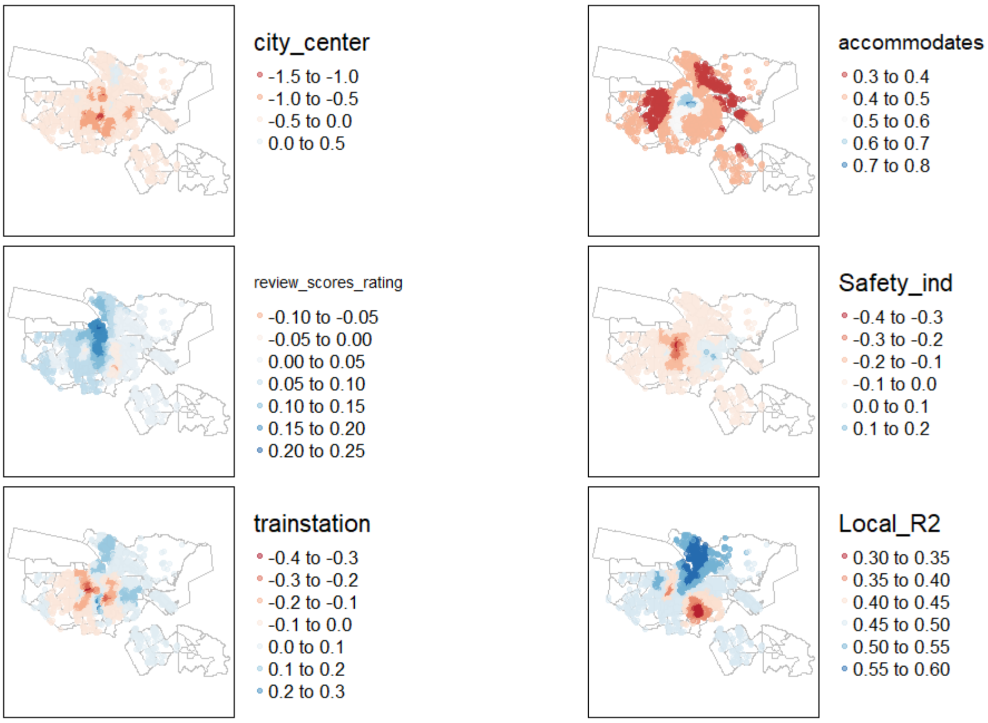

Figure 6. Visualisation of spatial variance of variables.

The trainstation map, for example, shows that the closer the listing is to a train station, the higher the price (negative correlation in the city centre), but the link is inverse on the city outskirts. This is an excellent illustration of spatial heterogeneity. Furthermore, the majority of the negative relationships across the city are related to the distance to the city centre. That means that as the distance increases, the price decreases. The majority of locations have a negative relationship with the safety index, which means that when the safety index rises, the price falls.

Analysis and Discussion

The findings revealed that the HPM approach is unsuccessful for predicting the price of Airbnb accommodations in Amsterdam. For starters, the assumptions were violated and the R-squared was only 40%. The GWR approach is then used to detect autocorrelation. With a lower AIC, better error models, and a higher R-square, that model outperforms HPM. It better explains the data relationships because it considers spatial effects like spatial dependency and heterogeneity. Figures 4 and 5 clearly show these spatial effects. The effect on price varied by location, with 'neighbourhood variables' yielding interesting findings. Some variables have overwhelmingly positive effects, so as the predictor rises, so does the price. These were primarily the variables that were related to the listings, so review_scores_rating and accommodates. One particular result from the 'neighbourhood variables' was intriguing. The safety index demonstrated a negative relationship, indicating that the safer the area, the lower the price. Additionally, when it comes to trainstation varible, there is clear relationship that in the places that there is train station the price of the accommodation is higher. The only exception is the Noord (blue part on the map), as it is separated from the city by the river and there is no train station there. Furthermore, many of the spatial variables were insignificant, which could be because the difference within Amsterdam is too large to be established as a spatial effect. This was in contrast to the literature, which indicated that many tourists look for specific places and attractions, and thus they should be more

[2]Summary of the GWM is in the Appendix 3.

significant. The reason for this could be that Amsterdam is a small and compact city. There is public transportation and bikes available to get from one location to another, making commuting easier. As a result, the listing variables may be more important, as the tourist's place of residence is more important than the neighborhood. Furthermore, the reason could be that Amsterdam is an expensive city (*Cost of Living in Amsterdam: Average Cost & Top Expenses*, 2023), and guests may forego some of the area's attractiveness in exchange for a lower price, as indicated by the safety index variable and its results. Additionally, the negative relationship with safety_index can be due to the fact that the center of the city is not the safest place, according to Amsterdam Municipality (*Veiligheid in Beeld | Website Onderzoek En Statistiek*, 2022).

Nonetheless, GWR's results are far from perfect. Because of the point data and the adjustments required to aggregate it, the prediction may be poor. Furthermore, polygon data was not used because Amsterdam has too few neighbourhoods, which would lead to generalisations. One solution, as well as additional research, could be to use Multiscale Geographically Weighted Regression. Because that method finds a different bandwidth for each variable, the data can be modelled with greater precision. This would allow us to investigate one more time insignificant variable as well as issues with the safety index.

Conclusion

This study demonstrated that the GWR is a more accurate predictor of Airbnb accommodation prices than HPM. It performs better and can explain the local effects of the factors in greater detail. The factors belonging to the 'neighbourhood variable' category were not as significant as expected. To avoid the 'Airbnb effect,' policymakers can take distance to the city centre and nearest train station into account when establishing new regulations, as this effect can work similar to gentrification, which will increase the price even further in those places (Barker, 2020).

References

Airbnb. (2023, February 14). *About us - Airbnb Newsroom*. Airbnb Newsroom. https://news.airbnb.com/about-us/

Bakker, M. (2021, September 1). *Airbnb listing prices in Amsterdam and distance to tourist spots - Student Theses Faculty Spatial Sciences*. https://frw.studenttheses.ub.rug.nl/3694/

Barker, G. (2020, February 21). The Airbnb Effect On Housing And Rent. *Forbes*. https://www.forbes.com/sites/garybarker/2020/02/21/the-airbnb-effect-on-housing-and-rent/?sh=855b47722260

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (2010). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Contu, G., Frigau, L., & Conversano, C. (2022). Price indicators for Airbnb accommodations. *Quality & Quantity*. https://doi.org/10.1007/s11135-022-01576-6

Cost of Living in Amsterdam: Average cost & Top Expenses. (2023, April 13). Digit Insurance. https://www.godigit.com/guides/cost-of-living/cost-of-living-in-amsterdam#:~:text=How%20Expensive%20Is%20Amsterdam%20Compared,Amsterdam%20than%20that%20in%20Singapore.

De Jonge, M. D. J. (2021, July 20). *DO TOURISTS' VALUE URBAN GREEN SPACE? - Student Theses Faculty Spatial Sciences*. https://frw.studenttheses.ub.rug.nl/3578/

General Information about Amsterdam | AmsterdamTourist.info. (2020, December 14). AmsterdamTourist.info. https://www.amsterdamtourist.info/about-amsterdam/

Gyódi, K., & Nawaro, Ł. (2021). Determinants of Airbnb prices in European cities: A spatial econometrics approach. *Tourism Management*, *86*, 104319. https://doi.org/10.1016/j.tourman.2021.104319

Kuby, M., & Yoo, C. (2020). Analyzing Spatial Variance of Airbnb Pricing Determinants Using Multiscale GWR Approach. *Sustainability*, *12*(11), 4710. https://doi.org/10.3390/su12114710

Monson, M. (2009, July 1). *Valuation Using Hedonic Pricing Models*. https://ecommons.cornell.edu/handle/1813/70656

Nault, A. (2019, May 20). NIGHTLIFE AS AN ESSENTIAL FEATURE FOR THE TOURIST SECTOR. *Tourism Review*. https://www.tourism-review.com/tourist-sector-benefits-from-quality-nightlife-news11051

Norbert. (2023). The Most Visited Cities in Europe: Where You Should Go and Why. *GloboTreks*. https://www.globotreks.com/travel-tips/most-visited-cities-in-europe/

Presentation-Ready Data Summary and Analytic Result Tables. (n.d.). https://www.danieldsjoberg.com/gtsummary/

Sada, N. (2021, December 10). This is why an Amsterdam Airbnb Listing is Expensive. *Medium*. https://medium.com/@nesreensada/this-is-why-an-amsterdam-airbnb-listing-is-expansive-593029aec0df

Team, I. (2020, October 3). *Sharing Economy: Model Defined, Criticisms, and How It's Evolving*. Investopedia. https://www.investopedia.com/terms/s/sharing-economy.asp

Team, I. (2021, October 26). *Peer-to-Peer (P2P) Economy: Definition Vs. Capitalism*. Investopedia. https://www.investopedia.com/terms/p/peertopeer-p2p-economy.asp

Veiligheid in beeld | Website Onderzoek en Statistiek. (2022, October 31). https://onderzoek.amsterdam.nl/interactief/dashboard-veiligheid?gebied=52&index=criminaliteit&subindex=totaal&indeling=wijk&meting=2022-3

Weng, S. C. (2018). *Hotel vs. Airbnb: What Attributes Do Travelers Value The Most?* ScholarWorks. https://scholarworks.calstate.edu/concern/theses/h989r568p

Zhang, Z., Chen, R. R., Han, L. D., & Yang, L. (2017). Key Factors Affecting the Price of Airbnb Listings: A Geographically Weighted Approach. *Sustainability*, *9*(9), 1635. https://doi.org/10.3390/su9091635

Appendix 1. Complete list of all the variables

Variables	Short Description	Source
X	Row number	Inside Airbnb Listings
id	ID of each property	Inside Airbnb Listings
name	Name of the Airbnb accommodation	Inside Airbnb Listings
neighbourhood_cleansed	Neighbourhood in which the accommodation is located	Inside Airbnb Listings
latitude	Latitude coordinates of the accommodation	Inside Airbnb Listings
longitude	Longitude coordination of the accommodation	Inside Airbnb Listings
property_type	Type of property (e.g., apartment, house, etc.)	Inside Airbnb Listings
accommodates	Number of guests the property can accommodate	Inside Airbnb Listings
beds	Number of beds in the property	Inside Airbnb Listings
price	Nightly price of the property	Inside Airbnb Listings
number_of_reviews	Number of reviews for the property	Inside Airbnb Listings
review_scores_rating	Average review score for the property	Inside Airbnb Listings
geometry	Geometries of each neighbourhood	OSM
restaurant	Number of restaurants within a certain radius	OSM
nightclub	Number of nightclubs within a certain radius	OSM
supermarket	Number of supermarkets within a certain radius	OSM
park	Number of parks within a certain radius	OSM
trainstation	Number of train stations within a certain radius	OSM
busstop	Number of bus stops within a certain radius	OSM
bicycle_rental	Number of bicycle rentals within a certain radius	OSM
city_center	Distance in meters from each property to the city centre	OSM
Wijkcode	Code for the corresponding neighbourhood	Municipality
Wijk	Name of the corresponding neighbourhood	Municipality
culture_to	Number of cultural institutions within a certain radius	Municipality
public_gre	The total area of public green space in each neighbourhood	Municipality
Safety_ind	Safety index for each neighbourhood	Municipality

Variable Type	Description	Number of Columns
Non-spatial	Accommodation specifications (e.g., number of beds, capacity, neighbourhood, review score rating, property type)	17
Distance variables	Distance of the accommodation from certain places of interest to temporary visitors (e.g., city centre, restaurants, bus stops, train stations, nightclubs, bicycle rental places)	5
Neighbourhood-level Density variabkes	Neighbourhood characteristics (e.g., safety index, public green scale, cultural score related to the neighbourhood)	3

Source	Description	Data Collected	Data Type
Inside Airbnb	This is platform offers a detailed dataset of Airbnb listings for cities worldwide. Price, number of rooms, availability, host ratings, amenities, and other variables are included in the dataset. Inside Airbnb scraped the dataset on December 5, 2022.		Point data
OpenStreetMap	The Python OSMnx module was used to extract OpenStreetMap data, allowing for the retrieval of geographical data on several spatial variables within Amsterdam city. The retrieved variables can be classified as shops, transportation facilities, or amenities. We then imported the extracted variables into QGIS for inspection and projection to a single coordinate reference, Amersfoort / RD New.	Shops (supermarkets); transport facilities (bus stops, train stations, bicycle rentals); amenities (nightclubs, restaurants)	Point Data
Amsterdam Municipality	The dashboard of the Amsterdam municipality contains indicators at different geographies. These are as follows: culture_total, public green (%), and safety_index.	Culture_total Publicgreen (%); safety_index	Density Variables

Appendix 2. Full summary of Hedonic Price Model.

Characteristic	Beta	95% CI[1]	p-value
neighbourhood_cleansed			
1	—	—	
2	0.02	-0.09, 0.12	0.8
3	-0.31	-0.47, -0.14	<0.001
4	-0.32	-0.46, -0.18	<0.001
5	-0.18	-0.37, 0.00	0.053
6	-0.16	-0.29, -0.03	0.016
7	0.09	-0.05, 0.22	0.2
8	0.25	-0.05, 0.55	0.10
9	-0.07	-0.22, 0.07	0.3
10	-0.25	-0.48, -0.01	0.043
11	-0.17	-0.45, 0.10	0.2
12	-0.44	-0.67, -0.20	<0.001
13	-0.36	-0.54, -0.19	<0.001
14	0.94	0.39, 1.5	<0.001
15	0.23	-0.21, 0.66	0.3
16	-0.03	-0.29, 0.23	0.8
17	0.67	0.17, 1.2	0.009
18	0.48	0.08, 0.87	0.017
19	0.65	0.26, 1.0	0.001
20	-0.02	-0.26, 0.22	0.9
21	0.11	-0.06, 0.27	0.2
22	0.16	-0.13, 0.46	0.3
property_type			
1	—	—	
2	0.45	0.37, 0.53	<0.001
3	0.48	0.28, 0.68	<0.001
4	0.20	0.02, 0.37	0.026
5	1.0	0.87, 1.2	<0.001
6	0.94	0.78, 1.1	<0.001
7	0.70	0.58, 0.82	<0.001
8	0.16	0.04, 0.28	0.009

Characteristic	Beta	95% CI[1]	p-value
9	0.77	0.64, 0.91	<0.001
10	0.49	0.32, 0.66	<0.001
11	0.05	-0.13, 0.22	0.6
12	0.22	-0.19, 0.63	0.3
13	0.87	0.78, 0.96	<0.001
14	0.01	-0.18, 0.19	>0.9
15	0.22	0.06, 0.39	0.007
16	0.30	-0.04, 0.64	0.088
17	0.22	0.06, 0.39	0.006
18	0.57	0.37, 0.78	<0.001
19	0.68	0.26, 1.1	0.002
20	0.08	-0.26, 0.42	0.6
21	0.07	-0.41, 0.55	0.8
22	0.01	-0.25, 0.26	>0.9
23	0.60	0.12, 1.1	0.014
24	0.02	-0.62, 0.66	>0.9
25	0.33	0.15, 0.52	<0.001
26	0.33	-0.47, 1.1	0.4
27	0.14	-0.97, 1.3	0.8
28	0.04	-0.87, 0.94	>0.9
29	0.03	-0.19, 0.25	0.8
30	0.36	-0.05, 0.78	0.085
31	0.98	0.53, 1.4	<0.001
32	1.1	0.51, 1.8	<0.001
33	2.9	1.3, 4.4	<0.001
34	-0.01	-0.42, 0.40	>0.9
35	-0.08	-0.42, 0.26	0.6
36	1.6	1.1, 2.0	<0.001
37	0.14	-0.42, 0.71	0.6
38	0.26	-0.27, 0.79	0.3
39	1.2	0.68, 1.8	<0.001
40	0.23	-0.47, 0.93	0.5
41	0.67	0.01, 1.3	0.046
42	0.53	-0.25, 1.3	0.2

Characteristic	Beta	95% CI[1]	p-value
43	0.57	0.19, 0.96	0.003
44	1.1	0.30, 1.9	0.007
45	0.27	-1.3, 1.8	0.7
46	0.14	-0.76, 1.0	0.8
47	0.33	-0.34, 1.0	0.3
48	0.75	-0.96, 2.5	0.4
49	0.40	-0.74, 1.5	0.5
50	0.38	-1.2, 2.0	0.6
51	0.01	-1.1, 1.1	>0.9
52	-0.07	-1.6, 1.5	>0.9
53	0.42	-1.1, 2.0	0.6
54	-0.21	-1.8, 1.3	0.8
accommodates	0.45	0.42, 0.47	<0.001
review_scores_rating	0.06	0.04, 0.08	<0.001
restaurant	-0.02	-0.05, 0.02	0.3
nightclub	-0.04	-0.09, 0.02	0.2
suppermarket	0.00	-0.03, 0.04	0.9
park	0.01	-0.02, 0.04	0.4
trainstation	0.05	0.01, 0.10	0.016
busstop	0.02	0.00, 0.04	0.082
bicycle_rental	-0.06	-0.12, 0.01	0.088
city_center	-0.26	-0.37, -0.15	<0.001
culture_to	0.00	-0.03, 0.03	>0.9
public_gre	0.02	-0.01, 0.06	0.2
Safety_ind	-0.04	-0.07, -0.01	0.014

[1]CI = Confidence Interval

Appendix 3. Results of the GWR.

```
 ***********************************************************************
   *                       Package   GWmodel                          *
   ***********************************************************************
   Program starts at: 2023-04-14 12:50:25
   Call:
   gwr.basic(formula = equation_1, data = sdf_sp, bw = abw, kernel = "gaussian",
    adaptive = TRUE)

   Dependent (y) variable:  price
   Independent variables:  neighbourhood_cleansed property_type
accommodates review_scores_rating restaurant nightclub suppermarket
park trainstation busstop bicycle_rental city_center culture_to
public_gre Safety_ind
   Number of data points: 6069
   ***********************************************************************
   *                    Results of Global Regression                   *
   ***********************************************************************

   Call:
    lm(formula = formula, data = data)

   Residuals:
    Min      1Q  Median      3Q     Max
-2.8023 -0.4100 -0.0972  0.2520 12.0813

   Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
   (Intercept)             -0.4234054  0.0664702  -6.370 2.03e-10 ***
   neighbourhood_cleansed2   0.0165356  0.0545312   0.303 0.761724
   neighbourhood_cleansed3  -0.3075307  0.0843457  -3.646 0.000269 ***
   neighbourhood_cleansed4  -0.3213810  0.0731220  -4.395 1.13e-05 ***
   neighbourhood_cleansed5  -0.1817275  0.0937974  -1.937 0.052738 .
   neighbourhood_cleansed6  -0.1592583  0.0659421  -2.415 0.015760 *
   neighbourhood_cleansed7   0.0869017  0.0673928   1.289 0.197281
   neighbourhood_cleansed8   0.2494624  0.1508857   1.653 0.098318 .
   neighbourhood_cleansed9  -0.0737657  0.0751687  -0.981 0.326467
   neighbourhood_cleansed10 -0.2456740  0.1213879  -2.024 0.043027 *
   neighbourhood_cleansed11 -0.1734795  0.1395205  -1.243 0.213770
   neighbourhood_cleansed12 -0.4371663  0.1193995  -3.661 0.000253 ***
   neighbourhood_cleansed13 -0.3623954  0.0888738  -4.078 4.61e-05 ***
   neighbourhood_cleansed14  0.9401150  0.2781319   3.380 0.000729 ***
   neighbourhood_cleansed15  0.2281416  0.2216627   1.029 0.303414
   neighbourhood_cleansed16 -0.0294074  0.1332778  -0.221 0.825374
   neighbourhood_cleansed17  0.6748323  0.2594672   2.601 0.009323 **
   neighbourhood_cleansed18  0.4753500  0.1999260   2.378 0.017456 *
   neighbourhood_cleansed19  0.6464443  0.1994897   3.240 0.001200 **
   neighbourhood_cleansed20 -0.0168614  0.1229084  -0.137 0.890888
   neighbourhood_cleansed21  0.1058633  0.0822585   1.287 0.198158
```

```
neighbourhood_cleansed22  0.1634686  0.1504363   1.087 0.277244    
property_type2            0.4484241  0.0415337  10.797  < 2e-16 ***
property_type3            0.4829174  0.1023348   4.719 2.42e-06 ***
property_type4            0.1983543  0.0889044   2.231 0.025712 *  
property_type5            1.0119992  0.0724963  13.959  < 2e-16 ***
property_type6            0.9403347  0.0820791  11.456  < 2e-16 ***
property_type7            0.6985758  0.0599560  11.651  < 2e-16 ***
property_type8            0.1579436  0.0605625   2.608 0.009131 ** 
property_type9            0.7737611  0.0706001  10.960  < 2e-16 ***
property_type10           0.4912481  0.0863139   5.691 1.32e-08 ***
property_type11           0.0482736  0.0899307   0.537 0.591435    
property_type12           0.2193247  0.2085302   1.052 0.292950    
property_type13           0.8677908  0.0451464  19.222  < 2e-16 ***
property_type14           0.0090166  0.0940340   0.096 0.923614    
property_type15           0.2233316  0.0834065   2.678 0.007435 ** 
property_type16           0.2991663  0.1754461   1.705 0.088214 .  
property_type17           0.2243522  0.0820325   2.735 0.006258 ** 
property_type18           0.5741291  0.1058500   5.424 6.06e-08 ***
property_type19           0.6830838  0.2160342   3.162 0.001575 ** 
property_type20           0.0818890  0.1740343   0.471 0.637991    
property_type21           0.0669638  0.2439316   0.275 0.783695    
property_type22           0.0069884  0.1310858   0.053 0.957485    
property_type23           0.5973801  0.2429770   2.459 0.013977 *  
property_type24           0.0159690  0.3266712   0.049 0.961013    
property_type25           0.3322114  0.0948720   3.502 0.000466 ***
property_type26           0.3279621  0.4054503   0.809 0.418614    
property_type27           0.1404282  0.5661463   0.248 0.804110    
property_type28           0.0361100  0.4626926   0.078 0.937796    
property_type29           0.0297480  0.1139156   0.261 0.793993    
property_type30           0.3630243  0.2107196   1.723 0.084979 .  
property_type31           0.9811956  0.2323655   4.223 2.45e-05 ***
property_type32           1.1486230  0.3267975   3.515 0.000443 ***
property_type33           2.8543033  0.7970190   3.581 0.000345 ***
property_type34          -0.0110457  0.2090189  -0.053 0.957857    
property_type35          -0.0804359  0.1722188  -0.467 0.640477    
property_type36           1.5690398  0.2303761   6.811 1.07e-11 ***
property_type37           0.1440610  0.2861692   0.503 0.614693    
property_type38           0.2587444  0.2687501   0.963 0.335702    
property_type39           1.2372237  0.2843141   4.352 1.37e-05 ***
property_type40           0.2341924  0.3572610   0.656 0.512157    
property_type41           0.6673539  0.3350354   1.992 0.046428 *  
property_type42           0.5288232  0.3991599   1.325 0.185275    
property_type43           0.5724552  0.1951755   2.933 0.003369 ** 
property_type44           1.0833034  0.4021013   2.694 0.007077 ** 
property_type45           0.2688521  0.7983504   0.337 0.736310    
property_type46           0.1410185  0.4611471   0.306 0.759768    
property_type47           0.3311566  0.3432920   0.965 0.334759    
property_type48           0.7539452  0.8743548   0.862 0.388564    
```

```
property_type49          0.4042266  0.5833347   0.693 0.488363
property_type50          0.3787405  0.8096192   0.468 0.639944
property_type51          0.0055796  0.5684777   0.010 0.992169
property_type52         -0.0662905  0.8045540  -0.082 0.934336
property_type53          0.4208651  0.7956879   0.529 0.596872
property_type54         -0.2141174  0.7953337  -0.269 0.787772
accommodates             0.4460638  0.0113520  39.294  < 2e-16 ***
review_scores_rating     0.0579540  0.0106886   5.422 6.12e-08 ***
restaurant              -0.0176327  0.0187217  -0.942 0.346319
nightclub               -0.0381106  0.0286228  -1.331 0.183082
suppermarket             0.0027967  0.0171227   0.163 0.870261
park                     0.0112575  0.0139709   0.806 0.420402
trainstation             0.0533861  0.0222104   2.404 0.016262 *
busstop                  0.0198609  0.0114156   1.740 0.081946 .
bicycle_rental          -0.0561347  0.0328618  -1.708 0.087651 .
city_center             -0.2599356  0.0548176  -4.742 2.17e-06 ***
culture_to              -0.0008373  0.0160626  -0.052 0.958430
public_gre               0.0233385  0.0182055   1.282 0.199911
Safety_ind              -0.0389870  0.0158726  -2.456 0.014068 *

---Significance stars
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.7933 on 5981 degrees of freedom
Multiple R-squared: 0.3797
Adjusted R-squared: 0.3707
F-statistic: 42.08 on 87 and 5981 DF,  p-value: < 2.2e-16
***Extra Diagnostic information
Residual sum of squares: 3764.121
Sigma(hat): 0.7876709
AIC:  14502.04
AICc:  14504.72
BIC:  9805.589
***********************************************************************
*          Results of Geographically Weighted Regression             *
***********************************************************************

*********************Model calibration information*********************
Kernel function: gaussian
Adaptive bandwidth: 215 (number of nearest neighbours)
Regression points: the same locations as observations are used.
Distance metric: Euclidean distance metric is used.

****************Summary of GWR coefficient estimates:******************
                               Min.     1st Qu.      Median     3rd Qu.        Max.
Intercept               -1.3519e+00  -8.1072e-01  -5.8728e-01  -3.7572e-01  4.7110e-01
neighbourhood_cleansed2   -6.4745e-01  -3.2287e-02   7.0651e-02   1.8070e-01  7.4890e-01
neighbourhood_cleansed3   -1.1819e+00  -4.9771e-01  -3.7248e-01  -1.8107e-01  4.0870e-01
neighbourhood_cleansed4   -1.0810e+00  -3.7367e-01  -2.4309e-01  -7.0471e-02  3.6310e-01
```

neighbourhood_cleansed5	-1.0225e+00	-3.8675e-01	-2.1798e-01	-6.7350e-02	4.9310e-01
neighbourhood_cleansed6	-1.0716e+00	-2.0252e-01	-1.2853e-02	1.9814e-01	7.3410e-01
neighbourhood_cleansed7	-1.1747e+00	3.0274e-02	1.5600e-01	3.6375e-01	9.4680e-01
neighbourhood_cleansed8	-8.0389e+36	-3.8143e-01	1.1965e-02	6.7863e-01	4.9726e+38
neighbourhood_cleansed9	-1.3926e+00	-2.1505e-01	-8.9006e-02	6.0307e-02	6.6150e-01
neighbourhood_cleansed10	-1.3035e+02	-3.4403e-01	-1.1709e-01	2.3623e-01	1.5684e+00
neighbourhood_cleansed11	-1.2552e+08	-8.1738e-01	-4.5360e-01	-2.6130e-01	5.6985e+06
neighbourhood_cleansed12	-2.6607e+00	-4.9512e-01	-2.4629e-01	2.0680e-01	1.2280e+00
neighbourhood_cleansed13	-1.0286e+00	-4.2372e-01	-1.7357e-01	2.1767e-01	1.3775e+00
neighbourhood_cleansed14	-5.1194e+80	8.0746e-01	1.6543e+00	2.5451e+00	8.3608e+79
neighbourhood_cleansed15	-1.2654e+34	-2.4252e-03	4.3799e-01	1.1143e+00	2.7028e+36
neighbourhood_cleansed16	-3.5940e+08	-2.6240e-01	-1.7026e-02	3.9483e-01	1.0194e+08
neighbourhood_cleansed17	-1.0706e+60	5.7025e-01	1.2899e+00	2.0375e+00	1.7070e+61
neighbourhood_cleansed18	-7.4727e+37	3.0117e-01	7.1138e-01	1.3478e+00	3.9353e+35
neighbourhood_cleansed19	-3.6981e+24	4.6310e-01	9.0025e-01	1.3769e+00	2.7752e+23
neighbourhood_cleansed20	-9.7872e-01	-1.4838e-01	5.0506e-02	3.6828e-01	1.2307e+01
neighbourhood_cleansed21	-6.7292e-01	-8.9584e-03	1.2478e-01	3.1750e-01	8.0840e-01
neighbourhood_cleansed22	-7.4458e-01	1.2419e-01	4.3635e-01	7.6467e-01	2.1482e+00
property_type2	1.9078e-01	3.7018e-01	4.6176e-01	5.5242e-01	7.7840e-01
property_type3	-1.9003e-01	2.9929e-01	4.6538e-01	6.7620e-01	1.8424e+00
property_type4	-2.2741e-01	9.5122e-02	1.9211e-01	3.1344e-01	6.1400e-01
property_type5	4.3877e-01	7.8788e-01	9.6089e-01	1.1519e+00	1.7880e+00
property_type6	-4.3623e-01	8.2728e-01	9.7591e-01	1.1951e+00	1.9744e+00
property_type7	4.1977e-01	6.1192e-01	7.6414e-01	1.0178e+00	1.6393e+00
property_type8	-9.8639e-02	6.4257e-02	1.3759e-01	2.5799e-01	5.1770e-01
property_type9	1.4236e-01	6.6169e-01	8.7110e-01	1.2573e+00	2.4985e+00
property_type10	-6.2659e-01	5.2980e-02	4.3981e-01	6.0978e-01	1.1294e+00
property_type11	-2.7958e-01	-6.6037e-02	2.4120e-02	1.2882e-01	3.4980e-01
property_type12	-7.5461e-01	-4.4651e-02	1.3403e-01	2.8736e-01	6.7500e-01
property_type13	4.8954e-01	7.7207e-01	8.8673e-01	1.0203e+00	1.3044e+00
property_type14	-4.2730e-01	-1.9301e-01	-1.1156e-01	-3.5144e-02	1.4340e-01
property_type15	-4.6028e-01	-1.2031e-01	-4.1952e-03	1.3213e-01	5.3840e-01
property_type16	-6.3059e-01	4.7091e-02	1.6060e-01	2.3824e-01	5.9210e-01
property_type17	-1.5355e-01	3.5936e-02	1.2964e-01	2.5710e-01	1.3031e+00
property_type18	-4.0160e-02	4.1135e-01	5.2988e-01	6.9533e-01	1.6932e+00
property_type19	-3.6973e-01	1.4693e-01	4.4322e-01	9.4828e-01	1.6892e+00
property_type20	-6.2846e-01	-9.0316e-02	9.7476e-02	2.1646e-01	8.2370e-01
property_type21	-1.0984e+00	-4.8590e-01	-1.5255e-01	2.0620e-01	5.5060e-01
property_type22	-2.1536e-01	-3.8414e-02	7.3778e-02	1.5826e-01	5.3550e-01
property_type23	-3.8959e-01	-9.1171e-02	2.5344e-01	1.2578e+00	6.6911e+00
property_type24	-1.3228e+00	-2.5525e-01	-9.5603e-02	1.2460e-01	9.9370e-01
property_type25	-5.7779e-01	8.2403e-02	2.8328e-01	4.3220e-01	1.3308e+00
property_type26	-3.2634e-01	2.8318e-01	5.2271e-01	7.2021e-01	2.4688e+00
property_type27	-3.9570e-01	1.2304e-01	3.0143e-01	4.8296e-01	1.0900e+00
property_type28	-5.6850e-01	-1.0638e-01	9.8868e-03	1.3191e-01	4.3470e-01
property_type29	-3.1523e-01	-8.2112e-02	6.0734e-03	1.2282e-01	5.8240e-01
property_type30	-5.3991e-01	-1.4689e-01	2.4163e-01	4.0260e-01	7.0360e-01
property_type31	1.3810e-01	6.2355e-01	8.5159e-01	9.9584e-01	1.6259e+00

```
property_type32          1.2691e-01   3.9532e-01   8.1664e-01   1.9070e+00  5.0280e+00
property_type33         -1.0645e+13   2.8296e+00   3.1197e+00   3.4484e+00  4.1667e+00
property_type34         -1.3043e+00  -7.0256e-01   1.2688e-02   1.7844e-01  4.9470e-01
property_type35         -1.1218e+00  -2.0281e-01  -9.2580e-02   8.3884e-03  4.0210e-01
property_type36          1.3137e+00   2.2188e+00   2.4749e+00   2.7581e+00  3.8025e+00
property_type37         -1.1428e+00  -2.1892e-01   5.4619e-02   3.0769e-01  9.0700e-01
property_type38         -1.6283e+00  -1.8063e-01  -1.5889e-02   2.0815e-01  5.1350e-01
property_type39          4.0739e-01   6.5177e-01   7.3841e-01   8.1766e-01  3.5980e+00
property_type40         -6.5991e-01   1.5710e-02   1.7386e-01   4.0750e-01  6.3030e-01
property_type41         -1.9177e+01   5.4485e-02   3.0554e-01   7.3810e-01  2.6845e+04
property_type42          1.4768e-01   4.9674e-01   5.9817e-01   6.9965e-01  1.2256e+00
property_type43         -1.2294e+00  -6.5616e-01   5.1740e-01   7.7071e-01  2.7210e+00
property_type44          7.1801e-01   1.0491e+00   1.2423e+00   1.3686e+00  1.9312e+00
property_type45         -4.8500e-01   2.2791e-01   3.8867e-01   7.2743e-01  4.5085e+05
property_type46         -3.8175e-01  -1.3589e-01   6.8529e-03   1.7515e-01  1.1563e+00
property_type47         -3.5993e+16   4.3926e-01   1.0042e+00   1.7926e+00  3.4130e+40
property_type48        -2.6391e+166   4.2846e-01   2.3697e+00   4.2881e+00 4.7302e+157
property_type49         -1.6427e+48   1.6943e-01   8.7271e-01   1.9099e+00  1.8394e+30
property_type50         -4.6575e+14   1.8213e-01   3.9468e-01   8.4313e-01  1.1473e+13
property_type51         -1.4353e+09   4.1326e-02   1.5289e-01   3.3872e-01  1.2975e+00
property_type52         -1.2098e+23  -4.8231e-02   3.5287e-01   7.2911e-01  4.0953e+10
property_type53          1.4659e-01   3.3481e-01   4.3391e-01   5.5317e-01  1.5974e+00
property_type54         -1.4340e+00  -5.3771e-01  -3.4055e-01  -1.6032e-01  4.0810e-01
accommodates             3.5347e-01   4.1030e-01   4.6600e-01   5.2788e-01  7.3550e-01
review_scores_rating    -5.3252e-02   3.3322e-02   6.5270e-02   1.2381e-01  2.0950e-01
restaurant              -3.9927e-01  -6.6195e-02  -2.1904e-02   1.3722e-02  2.9510e-01
nightclub               -6.6509e-01  -1.9732e-01  -6.4463e-02   1.8874e-02  4.6930e-01
suppermarket            -2.3502e-01  -2.2160e-02   3.1989e-02   8.4957e-02  1.8590e-01
park                    -1.7304e-01  -1.6231e-02   9.5071e-03   5.0859e-02  2.2410e-01
trainstation            -3.3398e-01  -1.1006e-01  -1.7307e-02   6.7729e-02  2.7180e-01
busstop                 -1.5354e-01  -2.3609e-02   3.0692e-03   2.7383e-02  1.1560e-01
bicycle_rental          -4.2982e-01  -1.6742e-01  -1.0014e-01  -3.1695e-02  3.4460e-01
city_center             -1.2107e+00  -5.3557e-01  -3.6660e-01  -2.2033e-01  4.4530e-01
culture_to              -2.6187e-01  -3.9804e-02  -4.6672e-03   3.4470e-02  1.0950e-01
public_gre              -2.9681e-01   1.0786e-03   4.6104e-02   1.0064e-01  3.5400e-01
Safety_ind              -3.4620e-01  -1.1000e-01  -6.0109e-02  -1.4111e-02  1.2720e-01
************************Diagnostic information*************************
Number of data points: 6069
Effective number of parameters (2trace(S) - trace(S'S)): 662.4113
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 5406.589
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 14287.79
AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 13682.55
BIC (GWR book, Fotheringham, et al. 2002,GWR p. 61, eq. 2.34): 11540.96
Residual sum of squares: 3113.951
R-square value:  0.4868242
Adjusted R-square value:  0.4239386

***********************************************************************
```

Group 15: Hassan Ali, Sahar Pourahmad, Wiktoria Libera