The Impact of Different Location Factors on London Airbnb Listing Price Prediction

COMPGC27 - Individual Coursework

Weisi Han - 17028872 - 31 December 2017

Abstract

It is widely believed that location information could be the primary determining factor of housing price. This also works for the renting price for Airbnb listings. In this project, two linear regression models (Simple Linear Regressor and Gradient Boosting) will be built to predict the optimal Airbnb listing price based on a set of basic property features and additional property geographic features. Within the process, we will verify that the geographic features could improve the performance of our model and explore how different geographic factors contribute to the prediction of the listing price in London. After the analysis, it can be found that building locational features could indeed maximise predictive potency, in which the city centre proximity is the most promising factor with the highest importance, the borough the property belongs to also affects the listing price. However, the distance to tourist attraction/station generates the least influence.

Introduction

Airbnb is an online platform which provides peer to peer hospitality service, enabling property owners to rent out their spaces to travellers. Established in 2008, over 100 million visitors have booked property rentals via this platform. For Airbnb hosts, they can post a property for rent and meanwhile, decide the price to set. However, many hosts have less knowledge about how to calculate a proper rental price. Likewise, most guests also have no idea whether the property is worth renting. Therefore, making a price predictor to generate a fair price is really useful, which assists hosts in optimising their pricing and increasing profit potential.

Companies like AirDNA has provided analytics tools which allow people to explore how Airbnb is being used in cities. Apart from providing basic information, it also outputs the pricing of different types of estates based on comparable listings over millions of short-term rentals. This process involves analysing the property type, location, demand and other information.

In fact, the right price for Airbnb listings can be affected by various factors. Aside from some straightforward factors, such as property space and amenities, other significant elements could be location factors. With the aim of precisely prediction, it is important to figure out how different location factors affect the Airbnb listing price. The following parts of the report will focus on processing and comparing some valuable geographic parameters, and analysing the insights caused by using these features.

Methodology

1. Approach

As is mentioned above, our focus is to predict the Airbnb listing price and explore how different geographical factors can affect the prediction. To achieve this, I decide to create some additional geographic features at first. By applying this data together with other necessary information into regression models, we can predict the listing price.

2. Dataset (See Appendix)

- **Airbnb property dataset:** This dataset contains the details of each London listing information from 2008 to 2017. The data was collected from Airbnb website. There are 38818 listings in the dataset and each tuple illustrates a piece of property information, including room type, bedroom and bed number, borough, number of guests the property can accommodate, locations (latitude and longitude), price per night, review score, etc.

For each piece of listing record, two secondary dataset are used to generate additional geographic features.

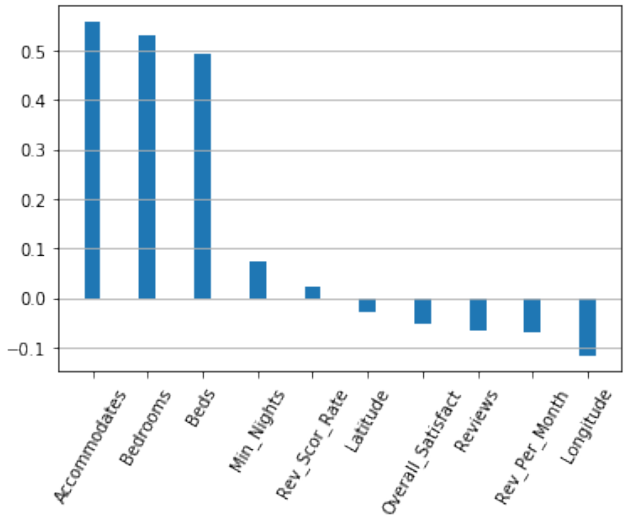
- **London tube station dataset:** This dataset contains 268 London tube stations with location information and zone information. This tube station list is provided by an OpenStreetMap project.

- **London tourist attraction dataset:** This dataset contains 32 top-ranked London famous attractions with location information. These famous landmarks was searched from google.

3. Data Preprocessing

- **Basic Listing Information Pre-processing:**

The first step of processing raw data should be eliminating features which may not be of great significance to a property's price. These features include "Room_id", "Host_id", etc. After that, I calculate the coefficients between each feature and the dependent variable "price". As is depicted in Figure1, there is a positive linear correlation between the listing price and "accommodate number", "bedroom" and "bed number". Other features like "reviews" or "overall satisfactions" are not strongly correlated with the price, which could be eliminated from our analysis.




```
Price                    1.000000
Accommodates             0.558162
Bedrooms                 0.529780
Beds                     0.493208
Minimum_Nights           0.074723
Review_Scores_Rating     0.021790
Latitude                -0.028997
Overall_Satisfaction    -0.051971
Reviews                 -0.065244
Reviews_Per_Month       -0.070074
Longitude               -0.115337
Name: Price, dtype: float64
```

Figure1. Coefficients between Each Independent Variable and Price

Apart from dealing with numerical variables, we have to interpret categorical variables like "room type" and "property type" into dummy variables. As every dummy variable represents a category of the explanatory variable, it will be marked as 1 if the case falls in this category or as 0 if not.

- **Geographic Feature Creation:**

I would like to divide the geographic information into three categories: the borough of London where the property is located, the proximity to local transportation or tourist attractions, and the closeness to the city centre.

i) Borough

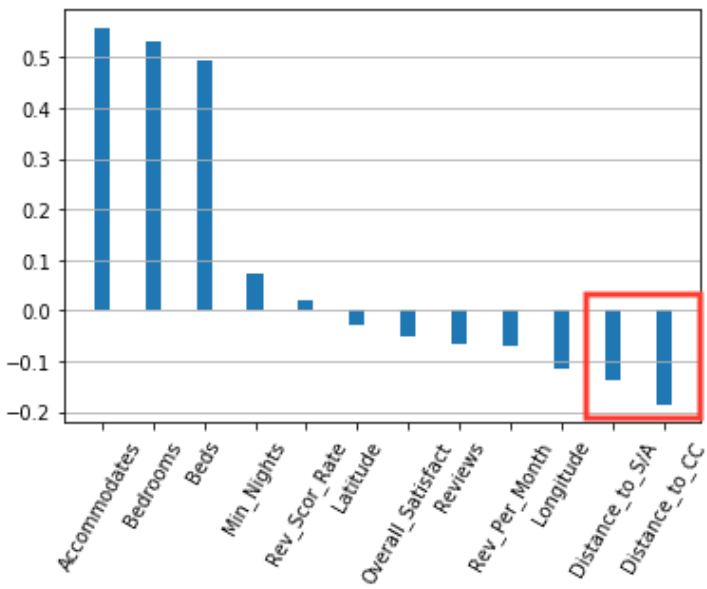
Certain London boroughs are more likely to have a relative higher average listing price. The borough information has already existed in the original dataset which appears as a categorial variable. Since there are 33 boroughs in London, I will create 33 dummy variables to represent each London borough.

ii) the Closeness to the City Centre

In many cases, the distance to London city centre accounts for a large fraction of variations in housing prices. In order to find the influence caused by the proximity to central London, I searched the coordinates for London's city centre (51.507602, -0.127816) and iterate each property to obtain the geographic distance (in meters) on a Euclidean distance basis. After doing so, I append this feature to the dataset.

iii) the Proximity to Local Transportations and Tourist Attractions

The accessibility to famous tourist attractions or public transportations can always be a factor in affecting Airbnb listing prices. Especially amongst tourists, it is important for a property to provide traveling convenience. Therefore, I go through each tube station and famous tourist attraction for every listing in the dataset, then calculate the distance (in meters) from the Airbnb listing to the stops and attractions. For each iteration through listings, I keep track of the minimum distance to the underground or attraction for each individual property. Then, taking it as another feature to the dataframe.




```
Price                    1.000000
Accommodates             0.558162
Bedrooms                 0.529780
Beds                     0.493208
Minimum_Nights           0.074723
Review_Scores_Rating     0.021790
Latitude                -0.028997
Overall_Satisfaction    -0.051971
Reviews                 -0.065244
Reviews_Per_Month       -0.070074
Longitude               -0.115337
Distance_to_S/A         -0.136248
Distance_to_CC          -0.184168
Name: Price, dtype: float64
```

Figure2. Coefficients between Each Independent Variable (Updated) and Price

After adding these features, we can recalculate the coefficients (see Figure2). It can be seen that there is a small negative correlation between the distance and the listing price.

4. Model Building and Training

In this experiment, I train my regression models on the training set, which takes 80% of the whole listing data, and test them on the the rest 20%. Two types of regression models have been used, one is Simple Linear Regression and another is Gradient Boosting. In the first round, each of the models only takes basic numeric attributes (accommodates, bedrooms, beds), and categorial attributes (property type, room type) as X to predict the listing price, Y. In the second round, the three types of geographic factors will be added to X. After running both models, the performance of the model prediction will be compared.

Results and Interpretation

1. Overall Comparison

After training the models, I compared them on R^2 score, RMSE, and MAPE of the test set. The result is shown in Table1, from which we can find that the Gradient Boosting model generates a much better outcome than the Simple Linear Regressor. Besides, by introducing different locational parameters, the performance of both models are improved to a large extant. 54.41% of the variability in Y can be explained using X, which is just 47.89% if these location factors are not existed. In addition, the Gradient Boosting model could predict the value of every listing within $81.56 of the real price, which is more accurate than only use the basic features. The MAPE also decreased from 40.113% to 34.764% in Gradient Boosting model.

	R^2		RSME		MAPE	
	Without location factors	With location factors	Without location factors	With location factors	Without location factors	With location factors
Simple Linear Regression	0.4512	0.5033	89.4847	85.1348	41.014%	38.823%
Gradient Boosting	0.4789	0.5441	87.1965	81.5630	40.113%	34.764%

Table1. R^2, RSME and MAPE on Two Different Models.

2. Comparison among Different Geographic Parameters

As the additional geographic features could make a great contribution to the accuracy of the prediction model, I would like to further explore which exactly type of geographic features is the most important one and vice versa. The following feature-importance chart reveals that the most influential locational characteristic is the distance to city centre, which accounts for almost 20% of all the features. Distinct London boroughs also affect the

listing price to some degree, and the importance is 17.4%. The least important geographic feature is the distance to the nearest station or attraction, which is only 5.46%.

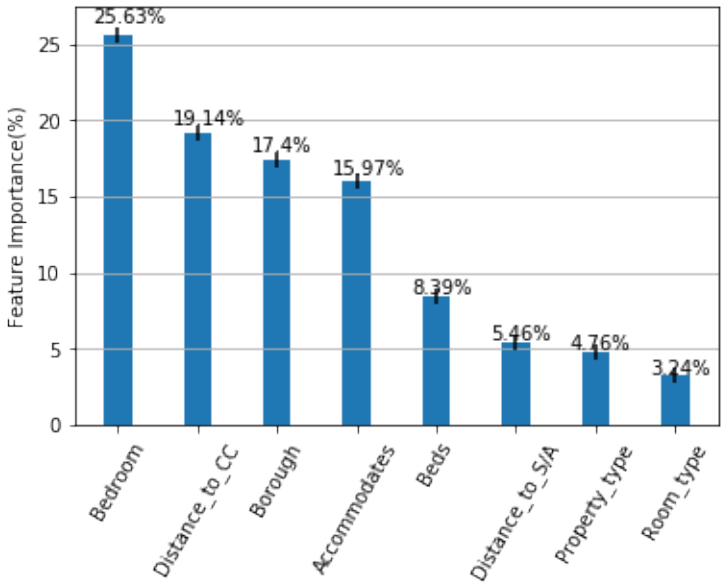


Figure3. Features by Importance

i) the Closeness to the City Centre

From the above processing, we can see that the proximity of a Airbnb listing to city centre is most correlated with the listing price. If we zoom in on their correlation scatter plot (see Figure4), we can find the central located properties are more likely to have a higher listing price compared with other properties in suburban area. This characteristic is very apparent as we can clearly see the downward trend in this diagram. Therefore, it is understandable that the listing price is largely dependent on the distance to city centre.

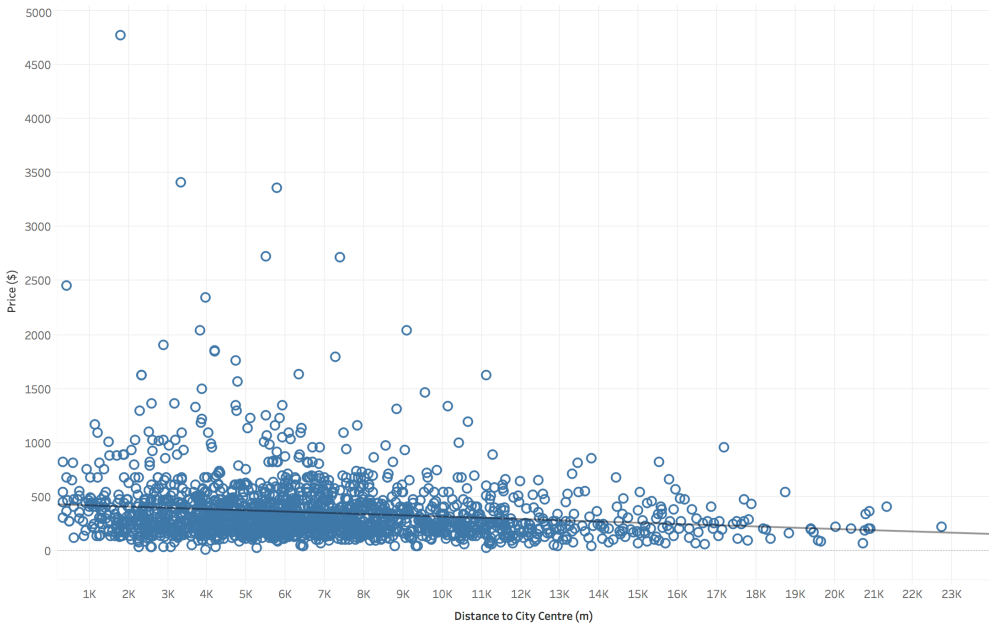


Figure4. The Relationship between Listing Price and Distance to City Centre

ii) Borough

Different borough information could generate a major effect on Airbnb listing price. The following map shows patterns of average property price on a borough scale. Since boroughs which have higher average listing price are represented in deeper blue colour, it is clear that properties located in inner London boroughs, especially the ones alongside the river Thames, tend to be more expensive than those in outer London boroughs (average listing price per borough ranges from $65.9 to $210.2). This may root in the fact that most historic, finance and business centres are located in inner London boroughs. Moreover, the different population density also strengthens the feature of individual borough's listing price.

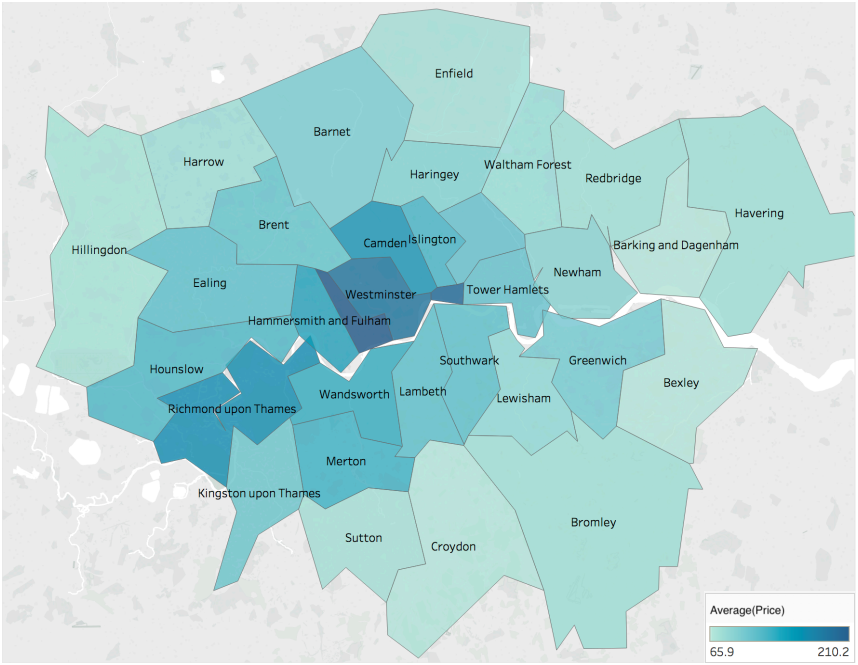


Figure5. Average property price of different London Boroughs

iii) the Proximity to Local Transportations and Tourist Attractions

Although we tend to believe that tourist attractions and convenient public transportation will catch the attention of potential visitors looking for an Airbnb listing, this is the least important factor in our experiment. Figure6 illustrates the relationship between listing price and the distance to its nearest station/attraction, which is categorised by inner London zones and outer London zones. We could observe a negative correlation between the two variables for both of the inner and outer London listings. The trend for inner London listings is more obvious than the outer ones since most London tube stations and tourist attractions are distributed in London zone one and two. Surprisingly, over 85% of Airbnb listings are located within 2000m of a tube station, which is considered to be a suitable distance for a property. Also, the average distance between the properties and the nearest station is around 1100m. This common high accessibility might be a reason that undermines the importance of this geographic feature to make prediction.

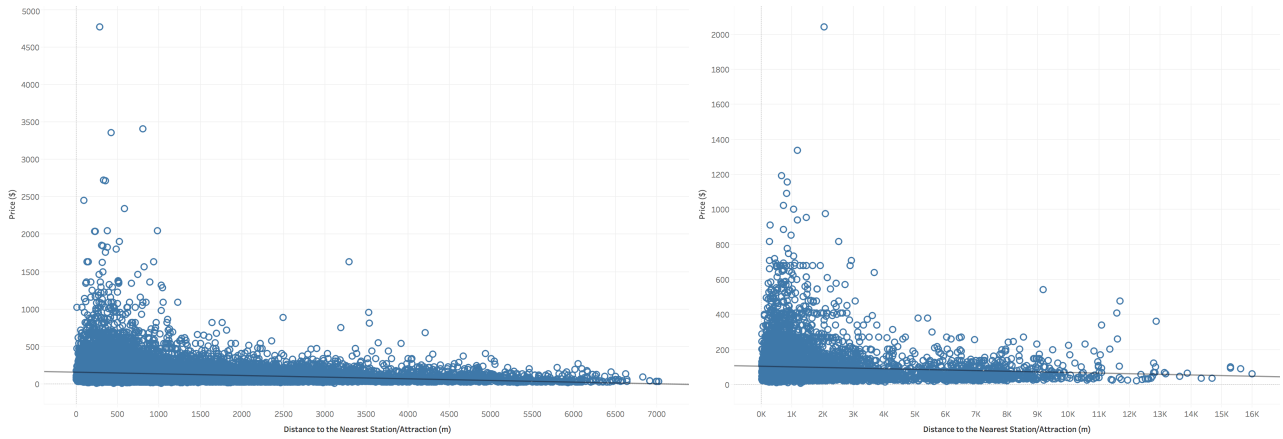


Figure6. The Relationship between Listing Price and Distance to the nearest Station/ Attraction

Conclusion and Insight

In this project, we apply two models to predict Airbnb listing price. The ensemble method, Gradient Boosting, gives us much more predictive accuracy compared to the Simple Linear Regressor, as it produces a higher R^2 value and lower error. During the experiment process, our focus is placed on how different geographic parameters could influence the listing price. Therefore, we create new locational features and append them to the original dataset. By looking at the feature importance bar chart, we could find these new features account for a large proportion of the importance (42% in total), which enables our model to predict price far better.

Among the three locational factors, the distance to city centre matters most to our prediction. The listing values tend to peak at the downtown area, then decline outward from the city centre. Similarly, distinct London boroughs also make a difference as properties built within same boroughs tend to have similar listing prices. The last factor is the accessibility to public transportation and scenic spots. This factor is not that influential compared with other factors. The reason behind may relate to the high spatial coverage of London underground. As the stations are highly accessible to a vast majority of Airbnb listings, the premium associated with living near underground or scenic spots is not that obvious.

From what has been discussed above, we can find that identifying useful geographic parameters helps us better create a model that accurately calculates the price for which an Airbnb listing should be priced at. These data insights enable people to get an idea of market price and estimate the value of a rental property. It can also be offered to real estate industry as well as hospitality industry for efficient marketing.

Retrospective

In this experiment, both of the numerical and categorial predictors have been effectively applied in our prediction. Compared with the baseline exploration, our additional geographic feature construction is proved to be successful since it enhances the performance of all models. However, there are still some insufficiencies in our prediction. For one thing, only two models have been used in this project, and this is not enough to generate the best result. Thus, I will try different linear models for the next time which can be easily achieved by using sklearn packet. For another, as I combined the tourist attraction and public transportation information in one feature, the influence caused by attractions has been weakened. However, due to the scenic views the tourist spots could offer, properties that line outside of the top tourist destinations should affect the listing prices theoretically. Therefore, I will process this parameter separately for future exploration.

References

Bansal, S. (2017). *Natural Language Processing Made Easy - using SpaCy (in Python).* [online] Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2017/04/natural-language-processing-made-easy-using-spacy-%E2%80%8Bin-python/ [Accessed 23 Dec. 2017].

Li, S. (2017). *Linear Regression in Python; Predict The Bay Area's Home Prices.* [online] DataScience+. Available at: https://datascienceplus.com/linear-regression-in-python-predict-the-bay-areas-home-prices/ [Accessed 23 Dec. 2017].

Inside Airbnb. (2017). *Inside Airbnb. Adding data to the debate.*. [online] Available at: http://insideairbnb.com/get-the-data.html [Accessed 7 Dec. 2017].

Opeyemi, A., Davis, C. and Rodriguez, D. (2016). *OptiBnB*. [online] GitHub. Available at: https://github.com/d1no007/OptiBnB/blob/master/data_exploration/exploration.md [Accessed 25 Dec. 2017].

Slee, T. (2017). *Airbnb Data Collection: Get the Data*. [online] Available at: http://tomslee.net/airbnb-data-collection-get-the-data [Accessed 7 Dec. 2017].

Appendix

My interest of exploring the geographic features is derived from performing analysis on the listing names of Airbnb properties. We know that listing titles can reflect the essence of the houses to some extent. More attractive titles tend to get more clicks and page views and thus, positively affect listing price points.

After conducting a basic processing by using Natural Language Processing(NLP), the listing names have been split into separated words (unigrams). By accumulating the number of occurrence of the repeated unigrams, we can find one third of the top 50 listing title unigrams are location related, which involves which borough or zone the property is in (eg. 'zone', 'Kensington'), the general location within the city (eg. 'central', 'west'), and the accessibility to public transportations (eg. 'tube'). In addition, since most Airbnb hosts rent out their residences as tourist accommodations, properties located near tourist attractions (eg. 'hill', 'park') can also be a main reason which attracts renters and affects the listing price. Therefore, the geographical information of Airbnb listings is a complex issue to tackle which influences property value in different aspects.

```
('room', 7770), ('flat', 5578), ('double', 5453), ('london', 4759), ('bedroom', 3881),
('bed', 2943), ('apartment', 2578), ('house', 2433), ('central', 1932), ('spacious', 1886),
('lovely', 1739), ('modern', 1609), ('garden', 1582), ('home', 1576), ('cosy', 1507),
('bright', 1412), ('beautiful', 1406), ('large', 1385), ('private', 1348), ('near', 1275),
('studio', 1147), ('family', 997), ('single', 935), ('zone', 865), ('luxury', 839),
('kensington', 783), ('close', 757), ('hill', 747), ('victorian', 721), ('chelsea', 678),
('great', 648), ('quiet', 640), ('bathroom', 629), ('park', 622), ('new', 613),
('ensuite', 611), ('charming', 597), ('location', 585), ('bridge', 563), ('stylish', 559),
('west', 559), ('tube', 554), ('suite', 541), ('brixton', 531), ('south', 526),
('notting', 518), ('wimbledon', 508), ('fulham', 508), ('cozy', 482), ('stunning', 477)
```

Figure7. Natural Language Processing on Airbnb Listing Titles

- Airbnb property dataset:

1	Accommoda	Bed_Type	Bedrooms	Beds	Borough	Host_Id	Id	Latitude	Longitude	Minimum_N	Name	Neighborhoo	Overall_Satis	Price	Property_Ty	Review_Scor	Reviews	Reviews_Per	Room_Type
2	1	Real Bed	1	1	Kingston upo	69018624	15896822	51.410036	-0.306323	7	Double room	Grove	0	32	Apartment	100	1	0.32	Private room
3	2	Pull-out Sofa	1	1	Kingston upo	18154504	4836957	51.411482	-0.290704	1	Modern roor	Norbiton	4.5	47	Apartment	91	15	0.67	Private room
4	2	Real Bed	1	1	Kingston upo	75741819	13355982	51.415851	-0.286496	1	Lovely doubl	Canbury	0	32	Apartment	80	2	0.25	Private room
5	2	Real Bed	1	1	Kingston upo	77078182	13472704	51.415723	-0.292246	2	Double Roon	Canbury	0	54	House		0		Private room
6	2	Real Bed	1	2	Kingston upo	3430124	16374418	51.410154	-0.278575	1	Cosy Tudor H	Coombe Vale	4.5	65	House	100	5	2	Private room
7	5	Real Bed	2	4	Kingston upo	33096320	12296555	51.421499	-0.269354	2	The Little Co	Coombe Hill	5	184	House	94	21	1.04	Entire home/apt
8	4	Real Bed	1	1	Kingston upo	54385939	14108797	51.401395	-0.266498	2	Location Loc	Beverley	0	47	House		0		Private room
9	4	Real Bed	2	2	Kingston upo	107474644	16446105	51.418258	-0.282628	1	Rooms in a lu	Coombe Hill	4.5	123	House		27		Private room
10	2	Real Bed	1	1	Kingston upo	3822323	735298	51.417095	-0.300863	2	Sleep well in	Canbury	0	53	House	90	2	0.11	Private room
11	10	Real Bed	4	7	Kingston upo	56080333	13712199	51.422364	-0.299202	5	Between the	Tudor	0	228	House	100	1	0.16	Entire home/apt
12	2	Real Bed	1	1	Kingston upo	27697324	5344655	51.40176	-0.255244	1	Lovely doubl	Beverley	5	43	House	98	28	1.64	Private room
13	2	Real Bed	1	1	Kingston upo	72718519	14912894	51.431117	-0.309713	2	Comfy sunny	Tudor	4.5	21	Apartment	90	10	1.03	Private room
14	1	Real Bed	1	1	Kingston upo	8581542	13857967	51.398832	-0.293539	3	Small well de	St. Mark's	0	47	Apartment		0		Private room
15	2	Real Bed	1	1	Kingston upo	65568682	14541148	51.397621	-0.308129	1	1 double roo	Grove	0	68	Apartment	80	2	0.3	Private room
16	1	Real Bed	1	1	Kingston upo	34769575	14296637	51.421323	-0.278367	1	Double room	Coombe Hill	0	31	Apartment	100	2	0.34	Private room
17	2	Real Bed	1	2	Kingston upo	47654068	13844847	51.392789	-0.303679	1	Luxury flat in	St. Mark's	3	39	Apartment	70	5	0.71	Private room
18	3	Real Bed	1	1	Kingston upo	32264660	6222799	51.419415	-0.288245	1	Loft room in	Canbury	5	61	House	100	28	1.34	Private room
19	2	Real Bed	1	1	Kingston upo	4755566	1765414	51.422868	-0.274883	1	Only 20 minu	Coombe Hill	5	74	House	100	8	0.25	Private room
20	7	Real Bed	2	4	Kingston upo	45469611	13577025	51.411893	-0.306083	3	Entire 2 bed/	Grove	5	204	Apartment	100	3	0.3	Entire home/apt
21	2	Real Bed	1	1	Kingston upo	90757380	16716104	51.417613	-0.281522	1	Kingston Dou	Coombe Hill	5	54	House	100	5	0.81	Private room
22	2	Airbed	1	1	Kingston upo	26831951	7283027	51.406567	-0.269026	2	LARGE TENT	Coombe Vale	5	30	Other_propt	95	13	0.6	Private room
23	2	Real Bed	1	1	Kingston upo	66626449	15926622	51.380422	-0.289297	2	Nice room, n	Tolworth and	0	61	House		0		Private room
24	3	Real Bed	1	2	Kingston upo	15082436	10085459	51.409866	-0.291875	5	Warm mode	Norbiton	4.5	115	Apartment	93	30	1.68	Private room
25	2	Real Bed	1	2	Kingston upo	96188094	16048507	51.401024	-0.258977	4	Opp pond be	Beverley	0	54	House		0		Private room
26	2	Real Bed	1	1	Kingston upo	89813442	14527936	51.422011	-0.277328	2	Airport Kings	Coombe Hill	4.5	44	Apartment	97	20	0.93	Private room
27	2	Real Bed	1	1	Kingston upo	32033043	6179409	51.418537	-0.282554	1	Double bedr	Coombe Hill	5	70	House	100	39	1.18	Private room
28	2	Pull-out Sofa	1	1	Kingston upo	47094767	9020269	51.419453	-0.286342	1	Spacious Roo	Canbury	0	40	House	80	1	0.06	Private room
29	8	Real Bed	4	4	Kingston upo	35215305	16868346	51.397563	-0.295201	4	4 Bedroom S	St. Mark's	0	408	House		0		Private room
30	2	Real Bed	1	1	Kingston upo	28439533	5483827	51.406203	-0.263098	5	Sunny flat be	Coombe Vale	0	98	Apartment		0		Entire home/apt
31	7	Real Bed	4	1	Kingston upo	65800295	12250201	51.405925	-0.306588	7	Large 4 bed h	Grove	0	271	House		0		Entire home/apt
32	4	Real Bed	1	2	Kingston upo	1889832	375799	51.402958	-0.250436	7	Luxury apart	Beverley	5	109	Apartment	98	52	0.87	Entire home/apt
33	8	Real Bed	4	5	Kingston upo	33889201	10898495	51.415807	-0.2585	3	Up to 20% of	Coombe Hill	0	546	House		0		Entire home/apt
34	4	Real Bed	2	2	Kingston upo	56469941	10953690	51.39189	-0.244506	2	23MinutsTo	St. James	0	76	House		0		Entire home/apt
35	2	Real Bed	1	1	Kingston upo	93463316	14894399	51.374544	-0.303409	1	Large D/roor	Tolworth and	5	43	House	100	8	0.71	Private room
36	9	Real Bed	5	2	Kingston upo	8161941	1530163	51.404114	-0.249173	3	Private room	Beverley	5	80	House	95	21	0.41	Private room
37	3	Real Bed	1	1	Kingston upo	8129765	16909573	51.411093	-0.306617	1	Cosy apartm	Grove	5	101	Apartment	97	27	6	Entire home/apt
⋮																			
38805	2	Real Bed	1	1	Barking and I	39501036	15210167	51.536487	0.081224	1	Double Roon	Abbey	0	47	Apartment	100	2	0.19	Private room
38806	1	Real Bed	1	2	Barking and I	21309371	13931684	51.533691	0.092099	2	Comfy room	Gascoigne	0	34	House		0		Private room
38807	1	Real Bed	1	1	Barking and I	76911791	13970414	51.538232	0.08496	3	Sing bed in r	Abbey	3.5	19	Apartment	73	3	0.39	Shared room
38808	3	Real Bed	1	2	Barking and I	48310806	9306445	51.535292	0.080316	4	1 Bedroom A	Abbey	0	68	Apartment	100	2	0.25	Entire home/apt
38809	4	Real Bed	1	2	Barking and I	71336321	12977957	51.532213	0.075841	1	1 bedroom a	Gascoigne	4.5	81	Apartment	80	18	1.03	Private room
38810	5	Real Bed	2	3	Barking and I	34577870	13763886	51.554754	0.127603	1	Cosy home, r	Parsloes	4.5	81	House	87	3	0.41	Entire home/apt
38811	2	Real Bed	1	1	Barking and I	1743811	15074540	51.540702	0.173105	5	Double room	Village	0	30	House	100	2	0.59	Private room
38812	6	Real Bed	3	4	Barking and I	9418622	16384036	51.538222	0.075944	5	3-BED APART	Abbey	5	148	Apartment		3		Entire home/apt
38813	4	Real Bed	2	2	Wandsworth	72755987	17265088	51.483053	-0.138129	1	15min away	Queenstown	4.5	191	Apartment		10		Entire home/apt
38814	4	Real Bed	2	2	Wandsworth	68973065	17452124	51.482859	-0.139614	1	Modern Cent	Queenstown	5	394	Apartment		20		Entire home/apt
38815	6	Real Bed	2	3	Lambeth	28826463	17524310	51.485909	-0.126948	1	London Eye \	Oval	0	338	Apartment		0		Entire home/apt
38816	1	Real Bed	1	1	Barking and I	115148705	17131270	51.529607	0.076597	5	Modern Roo	Gascoigne	0	34	Apartment		2		Private room
38817	1	Real Bed	1	1	Greenwich	48746417	17413813	51.495944	0.046958	2	Double Roon	Woolwich Ri	0	53	Apartment		0		Private room
38818	6	Real Bed	3	5	Wandsworth	29211630	17248546	51.484056	-0.131815	10	Luxury Londo	Queenstown	0	293	Apartment		1		Entire home/apt
38819	4	Real Bed	2	2	Hounslow	115648419	17184388	51.484367	-0.299983	3	Luxury Two E	Syon	0	204	Apartment		1		Entire home/apt

- London tube station dataset:

1	In service Station	Latitude	Longitude	Zone	Postcode
2	Acton Town	51.5030714	-0.2803027	3	W3 8HN
3	Aldgate	51.5143424	-0.0756269	1	EC3N 1AH
4	Aldgate East	51.5150816	-0.0730015	1	E1 7PT
5	Alperton	51.5412093	-0.2995158	4	HA0 4LL
6	Amersham	51.6741293	-0.6065138	9	HP6 5AZ
7	Angel	51.532968	-0.1055814	1	N1 8XB
8	Archway	51.5654905	-0.1351216	2,3	N19 5RQ
9	Arnos Grove	51.6162301	-0.1342702	4	N11 1AN
10	Arsenal	51.5585414	-0.1054999	2	N5 1LP
11	Baker Street	51.5231296	-0.1569042	1	NW1 5LA
12	Balham	51.4431826	-0.1526852	3	SW12 9SG
13	Bank	51.5133471	-0.0890001	1	EC3V 3LA
14	Barbican	51.5202153	-0.0977216	1	EC1A 4JA
15	Barking	51.5395211	0.08083161	4	IG11 8TU
16	Barkingside	51.5858506	0.08847664	4	IG6 1NB
17	Barons Court	51.4902287	-0.2134301	2	W14 9DP
18	Bayswater	51.5121113	-0.1879016	1	W2 4QH
19	Becontree	51.5395854	0.12688029	5	RM9 4TP
20	Belsize Park	51.5501909	-0.1639745	2	NW3 2AL
21	Bermondsey	51.4979608	-0.0643301	2	SE16 4RX
22	Bethnal Green	51.5271917	-0.0553921	2	E2 0ET
23	Blackfriars	51.5115869	-0.1029954	1	EC4V 4DD
24	Blackhorse Road	51.5857766	-0.0396257	3	E17 6JJ
25	Bond Street	51.5137975	-0.1496136	1	W1C 2HU
26	Borough	51.5010489	-0.0941973	1	SE1 1JX
27	Boston Manor	51.4957655	-0.3247472	4	TW8 9LQ
28	Bounds Green	51.6070319	-0.1241771	3,4	N11 2EU
29	Bow Road	51.5270418	-0.0243596	2	E3 4DH
30	Brent Cross	51.5765992	-0.2133358	3	NW11 9UA
31	Brixton	51.4627368	-0.1145521	2	SW9 8HE
32	Bromley-by-Bow	51.5248402	-0.0116824	2,3	E3 3BT
33	Buckhurst Hill	51.6265767	0.04671254	5	IG9 5ET
34	Burnt Oak	51.6025519	-0.2643307	4	HA8 0LA
35	Caledonian Road	51.5478854	-0.1183609	2	N7 9BA
		⋮			
259	Whitechapel	51.5195878	-0.0594205	2	E1 1BY
260	Willesden Green	51.5493733	-0.2224076	2,3	NW2 4QT
261	Willesden Junction	51.532234	-0.2439091	3	NW10 4RA
262	Wimbledon	51.4211078	-0.2066628	3	SW19 1PS
263	Wimbledon Park	51.434581	-0.1991863	3	SW19 7DZ
264	Wood Green	51.5974537	-0.1095266	3	N22 8HH
265	Wood Lane	51.5096582	-0.2244726	2	W12 7DS
266	Woodford	51.6072018	0.0340562	4	IG8 7QE
267	Woodside Park	51.6178684	-0.1854255	4	N12 8SE
268	Hammersmith (District)	51.4918198	-0.2228609	2	W6 9YA
269	Hammersmith (Met.)	51.4934834	-0.2246402	2	W6 0ED

- London tourist attraction dataset:

1	Tourist Attraction	Latitude	Longitude
2	Big Ben	51.500729	-0.124625
3	British Museum	51.519413	-0.126957
4	Tower of London	51.508112	-0.075949
5	Buckingham Palace	51.501364	-0.14189
6	Westminster Abbey	51.499292	-0.12731
7	Hyde Park	51.50727	-0.16573
8	Churchill's War Rooms	51.502217	-0.129329
9	The London Eye	51.503324	-0.119543
10	Hampton Court Palace	51.403613	-0.337762
11	The Victoria and Albert Museum	51.496639	-0.17218
12	Piccadilly Circus and Trafalgar Square	51.508039	-0.128069
13	St Paul's Cathedral	51.513845	-0.098351
14	Covent Garden	51.511732	-0.12327
15	Tate Modern	51.507595	-0.099356
16	Tate Britain	51.491062	-0.127789
17	Kew Gardens	51.4667	-0.2833
18	National Gallery	51.5089	-0.1283
19	Harrods	51.499405	-0.163234
20	Borough Market	51.505487	-0.090733
21	Victoria and Albert Museum	51.496639	-0.17218
22	Selfridges	51.51463	-0.15257
23	Science Museum	51.4978	-0.1745
24	Kensington Palace	51.505837	-0.187724
25	Palace of Westminster	51.499479	-0.124809
26	UCL	51.524559	-0.13404
27	the Shard	51.5045	-0.0865
28	Kensington Palace	51.505837	-0.187724
29	Camden Town	51.539011	-0.142555
30	Shakespeare's Globe	51.508076	-0.097194
31	Chinatown London	51.511423	-0.131753
32	Millennium Bridge	51.509529	-0.098543
33	Trafalgar Square	51.508039	-0.128069
34	Royal Observatory Greenwich	51.476853	-0.0005

- Final Airbnb property dataset:

1	Accommoda	Bed_Type	Bedrooms	Beds	Borough	Distance_to_CC	Distance_to_S/A	Host_Id	Id	Latitude	Longitude	Minimum_N	Name	Neighborhoc	Overall_Satis	Price	Property_Ty	Review_Sc	Reviews	Reviews_Per	Room_Ty
2	1	Real Bed	1	1	Kingston upc	16451.52068	2294.648315	69018624	15896822	51.410036	-0.306323	7	Double roor	Grove	0	32	Apartment	100	1	0.32	Private ro
3	2	Pull-out Sofa	1	1	Kingston upc	15543.17042	3379.229895	18154504	4836957	51.411482	-0.290704	1	Modern roor	Norbiton	4.5	47	Apartment	91	15	0.67	Private ro
4	2	Real Bed	1	1	Kingston upc	14997.82383	3807.186562	75741819	13355982	51.415851	-0.286496	1	Lovely doubl	Canbury	0	32	Apartment	80	2	0.25	Private ro
5	2	Real Bed	1	1	Kingston upc	15301.72028	3432.077481	77078182	13472704	51.415723	-0.292246	2	Double Roon	Canbury	0	54	House		0		Private ro
6	2	Real Bed	1	2	Kingston upc	15050.29825	4169.250888	3430124	16374418	51.410154	-0.278575	1	Cosy Tudor H	Coombe Vale	4.5	65	House	100	5	2	Private ro
7	5	Real Bed	2	4	Kingston upc	13704.112	4347.22112	33096320	12296555	51.421499	-0.269354	2	The Little Co	Coombe Hill	5	184	House	94	21	1.04	Entire hor
8	4	Real Bed	1	1	Kingston upc	15225.16476	4693.212168	54385939	14108797	51.401395	-0.266498	2	Location Loc	Beverley	0	47	House		0		Private ro
9	4	Real Bed	2	2	Kingston upc	14619.1405	4156.16833	107474644	16446105	51.418258	-0.282628	1	Rooms in a lu	Coombe Hill	4.5	123	House		27		Private ro
10	2	Real Bed	1	1	Kingston upc	15652.52646	2965.946661	3822323	735298	51.417095	-0.300863	2	Sleep well in	Canbury	0	53	House	90	2	0.11	Private ro
11	10	Real Bed	4	7	Kingston upc	15191.78103	3390.993071	56080333	13712199	51.422364	-0.299202	5	Between the	Tudor	0	228	House	100	1	0.16	Entire hor
12	2	Real Bed	1	1	Kingston upc	14712.9157	3997.615149	27697324	5344655	51.40176	-0.255244	1	Lovely doubl	Beverley	5	43	House	98	28	1.64	Private ro
13	2	Real Bed	1	1	Kingston upc	15201.22064	3613.62749	72718519	14912894	51.431117	-0.309713	2	Comfy sunny	Tudor	4.5	21	Apartment	90	10	1.03	Private ro
14	1	Real Bed	1	1	Kingston upc	16677.67788	3113.491316	8581542	13857967	51.398832	-0.293539	3	Small well de	St. Mark's	0	47	Apartment		0		Private ro
15	2	Real Bed	1	1	Kingston upc	17483.26803	2160.961644	65568682	14541148	51.397621	-0.308129	1	1 double roo	Grove	0	68	Apartment	80	2	0.3	Private ro
16	1	Real Bed	1	1	Kingston upc	14170.81059	4565.760239	34769575	14296637	51.421323	-0.278367	1	Double roor	Coombe Hill	0	31	Apartment	100	2	0.34	Private ro
17	2	Real Bed	1	2	Kingston upc	17649.37264	2653.204347	47654068	13844847	51.392789	-0.303679	1	Luxury flat in	St. Mark's	3	39	Apartment	70	5	0.71	Private ro
18	3	Real Bed	1	1	Kingston upc	14821.42262	3857.646335	32264660	6222799	51.419415	-0.288245	1	Loft room in	Canbury	5	61	House	100	28	1.34	Private ro
19	2	Real Bed	1	1	Kingston upc	13876.79682	4734.360688	4755566	1765414	51.422868	-0.274883	1	Only 20 minu	Coombe Hill	5	74	House	100	8	0.25	Private ro
20	7	Real Bed	2	4	Kingston upc	16303.30084	2382.369325	45469611	13577025	51.411893	-0.306083	3	Entire 2 bed/	Grove	5	204	Apartment	100	3	0.3	Entire hor
21	2	Real Bed	1	1	Kingston upc	14612.09906	4199.765059	90757380	16716104	51.417613	-0.281522	1	Kingston Dou	Coombe Hill	5	54	House	100	5	0.81	Private ro
22	2	Airbed	1	1	Kingston upc	14897.62036	4617.315	26831951	7283027	51.406567	-0.269026	2	LARGE TENT	Coombe Vale	5	30	Other_propt	95	13	0.6	Private ro
23	2	Real Bed	1	1	Kingston upc	18034.40587	4237.648212	66626449	15926622	51.380422	-0.289297	2	Nice room, n	Tolworth and	0	61	House		0		Private ro
24	3	Real Bed	1	2	Kingston upc	15725.9594	3257.878159	15082436	10085459	51.409866	-0.291875	5	Warm mode	Norbiton	4.5	115	Apartment	93	30	1.68	Private ro
25	2	Real Bed	1	2	Kingston upc	14934.48149	4260.476811	96188094	16048507	51.401024	-0.258977	4	Opp pond be	Beverley	0	54	House		0		Private ro
26	2	Real Bed	1	1	Kingston upc	14065.98994	4663.893551	89813442	14527936	51.422011	-0.277328	2	Airport Kings	Coombe Hill	4.5	44	Apartment	97	20	0.93	Private ro
27	2	Real Bed	1	1	Kingston upc	14594.286	4173.11936	32033043	6179409	51.418537	-0.282554	1	Double bedr	Coombe Hill	5	70	House	100	39	1.18	Private ro
28	2	Pull-out Sofa	1	1	Kingston upc	14720.00933	3977.465118	47094767	9020269	51.419453	-0.286342	1	Spacious Roc	Canbury	0	40	House	80	1	0.06	Private ro
29	8	Real Bed	4	4	Kingston upc	16859.41827	3028.186928	35215305	16868346	51.397563	-0.295201	4	4 Bedroom S	St. Mark's	0	408	House		0		Private ro
30	2	Real Bed	1	1	Kingston upc	14662.28648	4250.311293	28439533	5483827	51.406203	-0.263098	5	Sunny flat be	Coombe Vale	0	98	Apartment		0		Entire hor
31	7	Real Bed	4	1	Kingston upc	16770.4584	2177.614714	65800295	12250201	51.405925	-0.306588	7	Large 4 bed h	Grove	0	271	House		0		Entire hor
32	4	Real Bed	1	2	Kingston upc	14407.57683	3645.459358	1889832	375799	51.402958	-0.250436	7	Luxury apart	Beverley	5	109	Apartment	98	52	0.87	Entire hor
33	8	Real Bed	4	5	Kingston upc	13643.81111	3642.616451	33889201	10898495	51.415807	-0.2585	3	Up to 20% of	Coombe Hill	0	546	House		0		Entire hor
34	4	Real Bed	2	2	Kingston upc	15196.59492	3630.096	56469941	10953690	51.39189	-0.244506	2	23MinutsTo	St. James	0	76	House		0		Entire hor
35	2	Real Bed	1	1	Kingston upc	19157.77302	4016.221831	93463316	14894399	51.374544	-0.303409	1	Large D/roor	Tolworth and	5	43	House	100	8	0.71	Private ro
36	9	Real Bed	5	2	Kingston upc	14252.09676	3501.809252	8161941	1530163	51.404114	-0.249173	3	Private room	Beverley	5	80	House	95	21	0.41	Private ro
37	3	Real Bed	1	1	Kingston upc	16389.53464	2314.839139	8129765	16909573	51.411093	-0.306617	1	Cosy apartm	Grove	5	101	Apartment	97	27	6	Entire hor
⋮																					
38805	2	Real Bed	1	1	Barking and I	14815.19916	338.4676168	39501036	15210167	51.536487	0.081224	1	Double Roon	Abbey	0	47	Apartment	100	2	0.19	Private roo
38806	1	Real Bed	1	2	Barking and I	15489.80005	836.952111	21309371	13931684	51.533691	0.092099	2	Comfy room	Gascoigne	0	34	House		0		Private roo
38807	1	Real Bed	1	1	Barking and I	15109.91391	319.4868894	76911791	13970414	51.538232	0.08496	3	Sing bed in r	Abbey	3.5	19	Apartment	73	3	0.39	Shared roo
38808	3	Real Bed	1	2	Barking and I	14725.70666	471.6061135	48310806	9306445	51.535292	0.080316	4	1 Bedroom A	Abbey	0	68	Apartment	100	2	0.25	Entire hom
38809	4	Real Bed	1	2	Barking and I	14354.35732	882.8983529	71336321	12977957	51.532213	0.075841	1	1 bedroom a	Gascoigne	4.5	81	Apartment	80	18	1.03	Private roo
38810	5	Real Bed	2	3	Barking and I	18429.66018	1687.411496	34577870	13763886	51.554754	0.127603	1	Cosy home, r	Parsloes	4.5	81	House	87	3	0.41	Entire hom
38811	2	Real Bed	1	1	Barking and I	21141.68778	634.8841236	1743811	15074540	51.540702	0.173105	5	Double roor	Village	0	30	House	100	2	0.59	Private roo
38812	6	Real Bed	3	4	Barking and I	14502.61164	367.6061133	9418622	16384036	51.538222	0.075944	5	3-BED APART	Abbey	5	148	Apartment		3		Entire hom
38813	4	Real Bed	2	2	Wandsworth	2821.545734	802.0136039	72755987	17265088	51.483053	-0.138129	1	15min away	Queenstown	4.5	191	Apartment		10		Entire hom
38814	4	Real Bed	2	2	Wandsworth	2869.968389	870.0220809	68973065	17452124	51.482859	-0.139614	1	Modern Cen	Queenstown	5	394	Apartment		20		Entire hom
38815	6	Real Bed	2	3	Lambeth	2412.901346	216.7572267	28826463	17524310	51.485909	-0.126948	1	London Eye \	Oval	0	338	Apartment		0		Entire hom
38816	1	Real Bed	1	1	Barking and I	14353.88003	1140.646581	115148705	17131270	51.529607	0.076597	5	Modern Roo	Gascoigne	0	34	Apartment		2		Private roo
38817	1	Real Bed	1	1	Greenwich	12166.7275	3040.369187	48746417	17413813	51.495944	0.046958	2	Double Roon	Woolwich Ri	0	53	Apartment		0		Private roo
38818	6	Real Bed	3	5	Wandsworth	2632.792608	583.6100409	29211630	17248546	51.484056	-0.131815	10	Luxury Londo	Queenstown	0	293	Apartment		1		Entire hom
38819	4	Real Bed	2	2	Hounslow	12195.3492	1316.248269	115648419	17184388	51.484367	-0.299983	3	Luxury Two E	Syon	0	204	Apartment		1		Entire hom

- Calculation of Distance (VBA Code):

Distance to the Nearest Station/Attraction

```
Attribute VB_Name = "__1"
Sub Example02()
    Dim brr(1 To 268) As Double
    Dim drr(1 To 64145, 1 To 4)
    Dim Distance As Double
    Dim lat1 As Double
    Dim lon1 As Double
    Dim lat2 As Double
    Dim lon2 As Double
    arr = Sheet2.UsedRange
    crr = Sheet1.Range("C1:D64145")

For n = 2 To 64145
    lat1 = crr(n, 1)
    lon1 = crr(n, 2)
    For x = 2 To 269
        lat2 = arr(x, 2)
        lon2 = arr(x, 3)
        brr(x - 1) = CalcDistance(lat1, lon1, lat2, lon2)
    Next

    Distance = Application.WorksheetFunction.Min(brr)

    For i = 1 To 268
        If brr(i) = Distance Then
            x = i + 1
            Exit For
        End If
    Next
    n1 = n - 1
    drr(n1, 4) = Distance
    drr(n1, 1) = arr(x, 1)
    drr(n1, 2) = arr(x, 2)
    drr(n1, 3) = arr(x, 3)
Next
Sheet1.Range("F2").Resize(64145, 4) = drr

End Sub
Function CalcDistance(ByVal lat1 As Double, ByVal lon1 As Double,
ByVal lat2 As Double, ByVal lon2 As Double) As Double
    CalcDistance = 6371004 * 2 * Application _
    .Asin(Sqr(SumSq(Sin((Radians(lat1) - Radians(lat2)) / 2)) + Cos(Radians(lat1)) * _
    Cos(Radians(lat2)) * SumSq(Sin((Radians(lon1) - Radians(lon2)) / 2))))
End Function
Function Radians(latORlon As Double) As Double
    PI14 = 3.14159265358979
    Radians = latORlon * PI14 / 180
End Function
Function SumSq(xx As Double) As Double
    SumSq = xx * xx
End Function
```

Distance to City Centre

```
Sub Example02()
Dim drr(1 To 64145, 1 To 1)
Dim Distance As Double
Dim lat1 As Double
Dim lon1 As Double
Dim lat2 As Double
Dim lon2 As Double
crr = Sheet1.Range("C1:D64145")
lat2 = Sheet1.Cells(2, 6).Value
lon2 = Sheet1.Cells(2, 7).Value
For n = 2 To 64145
lat1 = crr(n, 1)
lon1 = crr(n, 2)

drr(n - 1, 1) = CalcDistance(lat1, lon1, lat2, lon2)

Next
Sheet1.Range("E2").Resize(64145, 1) = drr

End Sub
Function CalcDistance(ByVal lat1 As Double, ByVal lon1 As Double, ByVal lat2 As
    Double, ByVal lon2 As Double) As Double
CalcDistance = 6371004 * 2 * Application _
.Asin(Sqr(SumSq(Sin((Radians(lat1) - Radians(lat2)) / 2)) + Cos(Radians(lat1)) * _
Cos(Radians(lat2)) * SumSq(Sin((Radians(lon1) - Radians(lon2)) / 2))))
End Function
Function Radians(latORlon As Double) As Double
PI14 = 3.14159265358979
Radians = latORlon * PI14 / 180
End Function
Function SumSq(xx As Double) As Double
SumSq = xx * xx
End Function
```