

APPLIED DATA SCIENCE FINAL PROJECT

BATTLE OF NEIGHBORHOODS IN LONDON

1 INTRODUCTION AND BUSINESS PROBLEM

The main benefits of living in London include access to a great public transport system, job opportunities, exceptional cultural, and musical entertainment. The main disadvantages of London include expensive housing, traveling in rush hour, pollution, and tourists. There is a high variation on budgets depending on the people's needs, and so it's not easy to find a suitable place to accommodate different kind of families. The price on London housing suffers from high cost so every client would want to know the difference in price and characteristics to make an optimal decision. Nonetheless, it's an important factor to consider near locations such as restaurants, malls, shops, hospitals, education centers.

Using data science techniques, we can know very useful information about the current pricing depending on location and the other factors mentioned above.

2 DATA

- Government of UK open data
 - HM Land Registry / Price Paid Data
 - Helps obtaining location coordinates
- Type
 - CSV
 - JSON
- Description
 - Price on property sales
 - Location: England
 - Includes: received transactions at HM Land Registry
 - Source:
 - <http://landregistry.data.gov.uk/>
 - <https://foursquare.com/>
- Platform
 - IBM Cloud for Data – Jupyter Notebooks
- Google Maps Geocoding API
- Google Cloud Platform

3 METHODOLOGY

Sale prices of properties at Price Paid Dataset in England and Wales submitted to HM Land Registry for registration. An open dataset which is updated monthly. This project focuses on investigating the most recent market prices of property locations in London and recommends various places where the prospective client can buy a property based upon their financial situation.

These are the steps given in the process considering the selection of city and budget, so that each client can view changes as desired-

- I. Extract data from the price paid dataset
- II. Cleanse and delete old data with more than 5 years from now
- III. Centre the data for only London city
- IV. Any name below a label which is unique were filtered
- V. Get location information with Google Maps Geocoding API
- VI. Calculate the mean on recent transactions of property sale on various streets
- VII. Given an established budget, the prices are compared and recommendations are given for the location selected
- VIII. Given the location, it determines nearby places of interest

4 RESULTS

After implementing the processes, the exploratory data analysis for London city with a budget of 2.8 to 3.2 million dollars, it is recommended by the machine learning algorithm that 41 streets in London are the options for any client or future owner to choose and but a property given the current financial study. Some locations are presented in the next table.

Street	Latitude	Longitude	Avg_Price
DULWICH WOOD AVENUE	51.425586	-0.082416	2.297000e+06
SOUTH HILL PARK	51.557134	-0.164343	2.466667e+06
TEIGNMOUTH ROAD	51.550139	-0.214496	2.295000e+06
BURNSALL STREET	51.489042	-0.166883	2.286500e+06
FORDWYCH ROAD	51.551511	-0.206736	2.290000e+06
PORTEN ROAD	51.498603	-0.214120	2.200000e+06
ALBERT BRIDGE ROAD	51.477861	-0.164743	2.383333e+06
EDITH VILLAS	51.491665	-0.206556	2.402500e+06
WESTBOURNE GROVE	51.514797	-0.197071	2.300000e+06
LADBROKE ROAD	51.508776	-0.203410	2.261250e+06

Afterwards, the recommended street names were plotted on the London map with the average prices.



Then, the nearby venues are given to make the optimal choice of location.

	Street	Street Latitude	Street Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	DULWICH WOOD AVENUE	51.425586	-0.082416	Beer Rebellion	51.424580	-0.083425	Bar
1	DULWICH WOOD AVENUE	51.425586	-0.082416	The Indian Dining Club	51.427795	-0.086488	Indian Restaurant
2	DULWICH WOOD AVENUE	51.425586	-0.082416	The Paxton	51.427880	-0.086168	Pub
3	DULWICH WOOD AVENUE	51.425586	-0.082416	Gipsy Hill Railway Station (GIP)	51.424530	-0.083959	Train Station
4	DULWICH WOOD AVENUE	51.425586	-0.082416	Brown & Green @ The Station	51.424425	-0.083836	Breakfast Spot
5	DULWICH WOOD AVENUE	51.425586	-0.082416	Manuel's Restaurant and Bar	51.427591	-0.086131	Italian Restaurant
6	SOUTH HILL PARK	51.557134	-0.164343	Daunt Books Hampstead	51.555513	-0.166277	Bookstore
7	SOUTH HILL PARK	51.557134	-0.164343	Hampstead Heath Ponds	51.559300	-0.165973	Lake
8	SOUTH HILL PARK	51.557134	-0.164343	Hampstead Heath	51.559622	-0.164921	Park
9	SOUTH HILL PARK	51.557134	-0.164343	Paradise	51.555476	-0.166312	Indian Restaurant
10	SOUTH HILL PARK	51.557134	-0.164343	Keats House	51.555745	-0.167975	Museum
11	SOUTH HILL PARK	51.557134	-0.164343	karma bread	51.554494	-0.165586	Bakery
12	SOUTH HILL PARK	51.557134	-0.164343	The Garden Gate	51.554733	-0.165697	Pub
13	SOUTH HILL PARK	51.557134	-0.164343	The Little Thai	51.554115	-0.164737	Thai Restaurant
14	SOUTH HILL PARK	51.557134	-0.164343	Parliament Hill	51.559661	-0.159639	Scenic Lookout
15	SOUTH HILL PARK	51.557134	-0.164343	Silverberry Deli & Kitchen	51.554174	-0.164975	Café
16	SOUTH HILL PARK	51.557134	-0.164343	The Freemasons Arms	51.556968	-0.168806	Pub
17	SOUTH HILL PARK	51.557134	-0.164343	Zara	51.554423	-0.165561	Greek Restaurant
18	SOUTH HILL PARK	51.557134	-0.164343	The Stag	51.553420	-0.161576	Gastropub
19	SOUTH HILL PARK	51.557134	-0.164343	M&S Simply Food	51.553956	-0.165119	Grocery Store
20	SOUTH HILL PARK	51.557134	-0.164343	Mimmo la Bufala	51.555340	-0.166230	Italian Restaurant

We add different places such as facilities to make the client consider health options.

	Street Name	Facility Name	Facility Category	Distance	Facility Latitude	Facility Longitude
0	DULWICH WOOD AVENUE	Tesco	Grocery Store	1311	51.436213	-0.090561
1	DULWICH WOOD AVENUE	Tesco	Supermarket	3911	51.397043	-0.049559
2	DULWICH WOOD AVENUE	M&S Simply Food	Grocery Store	4513	51.389183	-0.111039
3	DULWICH WOOD AVENUE	Tesco	Grocery Store	3677	51.445271	-0.124966
4	DULWICH WOOD AVENUE	Tesco	Grocery Store	5315	51.438880	-0.155979
5	DULWICH WOOD AVENUE	Sainsbury's Local	Grocery Store	3749	51.446901	-0.124255
6	DULWICH WOOD AVENUE	telferscot primary school	Elementary School	4250	51.442807	-0.137087
7	DULWICH WOOD AVENUE	M&S Streatham Hill Foodhall	Grocery Store	3419	51.439372	-0.126443
8	DULWICH WOOD AVENUE	The Co-operative Food	Grocery Store	2306	51.440017	-0.106273
9	DULWICH WOOD AVENUE	Tesco	Supermarket	3216	51.398649	-0.099187
10	DULWICH WOOD AVENUE	Golden Jubilee Wing	Hospital	4836	51.468522	-0.093121
11	DULWICH WOOD AVENUE	Tesco Express	Grocery Store	3634	51.401783	-0.118253
12	DULWICH WOOD AVENUE	The Co-operative Food	Grocery Store	5529	51.471577	-0.112531
13	DULWICH WOOD AVENUE	Tesco Express	Grocery Store	4642	51.465227	-0.103222
14	DULWICH WOOD AVENUE	M&S Foodhall	Grocery Store	3386	51.426259	-0.131190
15	DULWICH WOOD AVENUE	The Co-operative Food	Supermarket	4366	51.432608	-0.020514
16	DULWICH WOOD AVENUE	Aldi	Supermarket	5169	51.448675	-0.017781
17	DULWICH WOOD AVENUE	The Co-operative Food	Supermarket	4934	51.381969	-0.069717
18	DULWICH WOOD AVENUE	Sainsbury's Local	Grocery Store	3155	51.440406	-0.121180
19	DULWICH WOOD AVENUE	Tesco	Grocery Store	6212	51.470177	-0.028570
20	DULWICH WOOD AVENUE	Balgowan Primary School	Elementary School	3811	51.404480	-0.039182
21	DULWICH WOOD AVENUE	Tesco	Grocery Store	5228	51.468581	-0.112767
22	DULWICH WOOD AVENUE	Tesco	Grocery Store	4831	51.385912	-0.110613

5 DISCUSSION

With the information given based on 41 streets, variables such as prices, venues and medical care locations, the client can choose the best place for its own interest. There are many cases in which you could opt for different combinations such as kids near schools, people who suffers from a disease near a hospital or even students near university campus.

6 CONCLUSION

The final choice is given depending on each family composition or each individual requirement. Of course it would also be important to take care in a further investigation the security of each location and its behavior with the economical changes that happens constantly. This applies for any location desired and just data from location and prices would be changed.