Coursera Capstone – Week 5

Capstone Project Report - The Battle of Neighborhoods

Author: Haitham Mostafa

1. Introduction
a. Problem and a discussion of the background

Since London UK is one of the most expensive places to live in, the project is established in order to help the property seekers in order to get the proper recommendation based on their budget, using existing open source data officially provided by the UK government, data science can be used in order to provide the recommendations together with the most popular venues related to this recommendation.

The project is generic and can be used for any place with available data for properties purchase transactions.

2. Data Source

a. Description of the data and how it will be used to solve the problem

Data provided by UK government in the form of CSV file with all the transactions for property purchasing with London and Wales, can be used in order to filter the proper districts for each property seekers budget and accordingly they can get all related facilities and point of interest using foursquare API.

Data Set Source: http://landregistry.data.gov.uk/

{7E86B6FA-7890-458C-E053-6B04A8C0C84C}	635000	7/6/2006 0:00	SW1X 0DH	F	N	L	40A		LENNOX GARDENS	LONDON	KENSINGT	GREATER I	A		A	
{7E86B6FA-90F9-458C-E053-6B04A8C0C84C}	193000	9/21/2006 0:00	TR15 3TP	D	N	F	5		HIGHER WEST TOLGU	REDRUTH	CORNWAI	CORNWAI	A		A	
{7E86B6FA-A144-458C-E053-6B04A8C0C84C}	110000	9/29/2006 0:00	EX34 8HT	S	N	F	4		KNELLER CLOSE	ILFRACON	NORTH DE	DEVON	A		A	
{7E86B6FA-A1C4-458C-E053-6B04A8C0C84C}	260000	10/9/2006 0:00	EX34 9HX	S	N	F	AMBERLEY		CASTLE HILL	ILFRACON	NORTH DE	DEVON	A		A	
{7E86B6FA-B219-458C-E053-6B04A8C0C84C}	390000	4/20/2006 0:00	S18 7SN	D	N	F	80		VALLEY RC	BARLOW	DRONFIEL	NORTH FA	DERBYSHII	A		A

3. Methodology

The project aims to give a brief data on the preferable streets where a person can find his affordable properties based on his pre-assigned budget, also it provides list of utilities and points of attraction within these locations.

Certain steps where followed in order to reach the required results:
- Locating of data source & analysis of the data structure and missing required data.
- Looping on google API in order to get the missing coordinates of the street names.
- Limiting the results based on the required budget.
- Map representation of the concluded results.
- Calling foursquare API to get related venues and famous locations.
- Listing the outcome.

4. Results

Based on the pre-assigned budget range of 2.1 GBP to 2.4 GBP we have the following list of streets where it can be affordable for a person searching for a property within this range

	Street	Avg_Price	latitude	longitude
71	ALMEIDA STREET	2.302000e+06	51.539575	-0.103625
151	BALHAM HIGH ROAD	2.183333e+06	51.447621	-0.148760
152	BALHAM HILL	2.375000e+06	51.452020	-0.147747
156	BALLINGDON ROAD	2.365000e+06	51.454189	-0.158856
265	BICKENHALL STREET	2.300000e+06	51.521197	-0.158934
342	BRAMPTON GROVE	2.400000e+06	51.570365	-0.283394
393	BRONDESBURY PARK	2.400000e+06	51.540439	-0.210419
620	CHILTERN STREET	2.225000e+06	51.521626	-0.156205
631	CHIVALRY ROAD	2.100000e+06	51.458021	-0.171349

Also, a map representation was conducted in order to give a better view of the outcome results

Based on the outcome below table represents the set of venues within these affordable streets

	Street	Street Latitude	Street Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	ALMEIDA STREET	51.539575	-0.103625	Ottolenghi	51.539716	-0.102314	Mediterranean Restaurant
1	ALMEIDA STREET	51.539575	-0.103625	Almeida Theatre	51.539620	-0.102845	Theater
2	ALMEIDA STREET	51.539575	-0.103625	The Taproom	51.540782	-0.102704	Pub
3	ALMEIDA STREET	51.539575	-0.103625	MEAT Liquor	51.539154	-0.102844	Burger Joint
4	ALMEIDA STREET	51.539575	-0.103625	Little Angel Theatre	51.538759	-0.101295	Theater
5	ALMEIDA STREET	51.539575	-0.103625	Zaffrani's	51.539255	-0.101202	Indian Restaurant
6	ALMEIDA STREET	51.539575	-0.103625	Amorino	51.538034	-0.102583	Ice Cream Shop
7	ALMEIDA STREET	51.539575	-0.103625	King's Head Theatre Pub	51.538249	-0.102404	Pub
8	ALMEIDA STREET	51.539575	-0.103625	The Regent	51.538952	-0.106809	Pub
9	ALMEIDA STREET	51.539575	-0.103625	Le Mercury	51.539697	-0.102715	French Restaurant
10	ALMEIDA STREET	51.539575	-0.103625	Hoxley and Porter	51.540361	-0.102983	Cocktail Bar
11	ALMEIDA STREET	51.539575	-0.103625	Gallipoli	51.537673	-0.102828	Middle Eastern Restaurant

5. Discussion

Based on the prepared project, it is noticed that Data science can be helpful in such cases of comparison and data representation.

Where Data science can be used in fields like stock market, property market, population analysis and much more.

The use of map representation and APIs like google and foursquare opens the door to a lot of innovative possibilities not only straight forward exercise.

6. Conclusion

As a conclusion the project can be used as a generic model for any other City , where the inputs are the data sets for transactions of properties purchasing in addition to the budget assigned for the new purchase , outcome will be set of affordable areas together with the set of venues available within these areas.

---End of Report--