Neighborhood Analysis Of England
By
Vindhya R

Introduction

- The housing market in the UK has always been a topic of national attention, with news sites dedicating sections that report on key news that can potentially affect housing prices and the trends in recent months.

- For many people, buying a property is one of the most important decision and purchase in life. Besides the affordability of a house, other factors such as the desirability of the location and the long-term investment prospects also affect the decision-making process.

- England homebuyers who want to find out more about housing trends in various districts would therefore have little recourse.

- In the last decade, the housing market in England has been rapidly growing, with average housing prices increasing by more than 10% yearly on most years .

Business Problem

- My client, a resident in UK is looking to settle in England and wants to buy a property. They want a green area, that has good schools, easy access to hospitals and is children friendly.

- Since the England demography is so big, my client needs deeper insight from available data in other to decide where to buy a property. My client is asking a help from us to analyse demography of England and give some insight about the Neighbourhoods.

Target Audience

- Considering the diversity of UK, there is a high multicultural sense. England is a place where different shades live. As such, in the search for an property, there is a high demand. The target audience is broad, it ranges from people of England, and others who wants to settle in England.

Data

- **Description of Data**

  This project will rely on public data from UK government dataset and Foursquare.

- **Dataset:**

  Within the England Area, there are lot of Boroughs. The focus of this project will be the neighbourhoods. The England Area consists of 380 Boroughs.

  We extract venues from all neighbourhoods within a certain radius and pick top 5 venues in each neighbourhood and analyse which neighbourhoods are good for property investment depending on client's requirements.
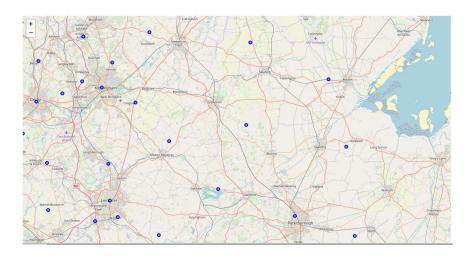
Data (contd)

- Cleaning data

1. Convert the data that are downloaded into a data frame.

2. Remove the unwanted attributes/columns that are not used in analysis

3. Rename few columns to give some meaningful names depending on the context
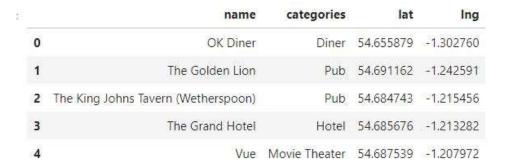
4. Check the data frame for missing values

Data (contd)

	Neighborhood	latitude	longitude	area	length
0	Hartlepool	55	-1	93,597,465	69,388
1	Middlesbrough	55	-1	53,878,997	42,086
2	Redcar and Cleveland	55	-1	244,838,828	96,190
3	Stockton-on-Tees	55	-1	204,935,660	115,447
4	Kingston upon Hull, City of	54	-0	71,455,773	63,330

Methodology

- As said before my client wants to buy a property and settle in England. I have analysed the neighbourhood of England. First, have extracted the latitude and longitude of England. The map of England is shown here

Methodology (contd)

	name	categories	lat	lng
0	OK Diner	Diner	54.655879	-1.302760
1	The Golden Lion	Pub	54.691162	-1.242591
2	The King Johns Tavern (Wetherspoon)	Pub	54.684743	-1.215456
3	The Grand Hotel	Hotel	54.685676	-1.213282
4	Vue	Movie Theater	54.687539	-1.207972

- Extract the first neighborhood from the data frame i.e. Haretlpool and using foursquare API fetch nearby venues in Hartelpool within a specified radius.

Methodology (contd)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
11999	Newport	51.582321	-2.89769	Tesco Extra	51.576474	-2.946330	Supermarket
12000	Newport	51.582321	-2.89769	The Snug	51.609064	-2.952901	Café
12001	Newport	51.582321	-2.89769	Frankie & Benny's	51.574729	-2.943908	American Restaurant
12002	Newport	51.582321	-2.89769	The Coldra (Toby Carvery)	51.596959	-2.937708	English Restaurant
12003	Newport	51.582321	-2.89769	B&Q	51.578311	-2.966849	Hardware Store
12004	Newport	51.582321	-2.89769	Elite Fast Car Hire Ltd	51.607210	-2.899592	Rental Car Location
12005	Newport	51.582321	-2.89769	WHSmith	51.588052	-2.836298	Bookstore
12006	Powys	52.348629	-3.43533	Ty Coed Eco Lodge	52.344313	-3.488063	Hotel
12007	Powys	52.348629	-3.43533	Happy Union Inn	52.331192	-3.389561	Bar
12008	Powys	52.348629	-3.43533	Mid Wales Inn	52.361325	-3.493690	Bar

- We get the nearby venues from all neigbourhoods in the dataframe and put them in a dataframe along with the category of that venue

Methodology(contd)

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Aberdeen City	50	50	50	50	50	50
Aberdeenshire	3	3	3	3	3	3
Adur	50	50	50	50	50	50
Allerdale	6	6	6	6	6	6
Amber Valley	42	42	42	42	42	42
Angus	4	4	4	4	4	4
Argyll and Bute	2	2	2	2	2	2
Arun	20	20	20	20	20	20
Ashfield	50	50	50	50	50	50
Ashford	50	50	50	50	50	50
Aylesbury Vale	5	5	5	5	5	5
Babergh	6	6	6	6	6	6
Barking and Dagenham	50	50	50	50	50	50
Barnet	50	50	50	50	50	50
Barnsley	16	16	16	16	16	16
Barrow-in-Furness	30	30	30	30	30	30
Basildon	50	50	50	50	50	50
Basingstoke and Deane	8	8	8	8	8	8
Bassetlaw	11	11	11	11	11	11
Bath and North East Somerset	14	14	14	14	14	14

- The venues are counted for each neighborhood for ease of analysis

Modelling

gitude	area	length	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
1.27023	9.359746e+07	69387.723173	4	Supermarket	Fast Food Restaurant	Pub	Pizza Place	Bar	Hotel	Soccer Stadium
1.21099	5.387900e+07	42085.584812	0	Pub	Clothing Store	Coffee Shop	Pizza Place	Café	Hotel	Park
1.00611	2.448388e+08	96189.660709	4	Pub	Convenience Store	Supermarket	Coffee Shop	Train Station	Café	Beach
1.30669	2.049357e+08	115447.244867	0	Pub	Park	Clothing Store	Pizza Place	Gym	Italian Restaurant	American Restaurant
0.30380	7.145577e+07	63330.240277	0	Pub	Coffee Shop	Bar	Café	Nightclub	Italian Restaurant	Portuguese Restaurant

- We have used K-means clustering here because it is fast and efficient in terms of computational cost, is highly flexible to account for mutations in real estate market in England and is accurate.

- The resulting dataframe after clustering and merging is shown below

Modelling (contd)

- The map is shown after clustering

Result

	latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
16	52.914639	0	Pub	Café	Burger Joint	Indian Restaurant	Park	Restaurant	Coffee Shop	Japanese Restaurant	Bar	Gym
28	51.546711	0	Pub	Grocery Store	Fish & Chips Shop	Supermarket	Coffee Shop	Warehouse Store	Gym / Fitness Center	Fast Food Restaurant	Golf Course	Sandwich Place
35	51.447708	0	Pub	Doctor's Office	Boat or Ferry	Grocery Store	Zoo	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market
51	49.923321	0	Pub	Harbor / Marina	Grocery Store	Restaurant	Seafood Restaurant	Heliport	Tea Room	Garden	Café	Airport
54	52.196281	0	Racetrack	Pub	Gastropub	Grocery Store	Golf Course	Supermarket	General Entertainment	Water Park	Music Venue	Food Court
61	51.666199	0	Pub	Grocery Store	Café	Historic Site	Garden Center	Soccer Stadium	Gastropub	Other Great Outdoors	Supermarket	Cave
66	53.238750	0	Pub	Grocery Store	Fast Food Restaurant	Garden Center	Coffee Shop	Castle	Bakery	Train Station	Sandwich Place	Furniture / Home Store
69	52.108051	0	Pub	Café	History Museum	Zoo	Hotel	Garden Center	Museum	Park	Train Station	Historic Site
75	53.123261	0	Campground	Pub	Hotel	Construction & Landscaping	Scenic Lookout	Flower Shop	Farmers Market	Fast Food Restaurant	Field	Film Studio
78	53.028839	0	Pub	Gastropub	Grocery Store	Chinese Restaurant	Restaurant	Fish & Chips Shop	Doctor's Office	Bar	Gym / Fitness Center	Supermarket
79	52.938202	0	Pub	Grocery Store	Hotel	Gastropub	Supermarket	Café	Sandwich Place	Coffee Shop	Park	Pharmacy
80	53.385689	0	Pub	Trail	Mountain	Scenic Lookout	Train Station	Zoo	Fish Market	Farm	Farmers Market	Fast Food Restaurant
82	50.868820	0	Cricket Ground	Furniture / Home Store	Pub	Food	Farmers Market	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market

- The neighbourhood clustering results as 5 different clusters

- First Cluster

- Analysis: As we can see here in first cluster there are lot of pubs, coffee shops, restaurants and cafes. This may not be ideal for a family with children like my client. This is good for people who hang out often and enjoy partying.

Result (contd)

	latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	53.586449	1	Grocery Store	Bar	Gym Pool	Hotel	Zoo	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market
17	52.667648	1	Hotel	Garden	Pub	Scenic Lookout	Gastropub	Athletics & Sports	History Museum	Park	Café	Fish Market
48	53.163361	1	Hotel	Pub	Zoo	Ice Cream Shop	Golf Course	Historic Site	Furniture / Home Store	Design Studio	Dance Studio	Fast Food Restaurant
70	54.685230	1	Hotel	Zoo	Spa	Restaurant	Café	Flower Shop	Farmers Market	Fast Food Restaurant	Field	Film Studio
73	54.466190	1	Bed & Breakfast	Brewery	Hotel	Farmers Market	Field	Film Studio	Fish & Chips Shop	Fish Market	Flea Market	Flower Shop
92	50.664879	1	Gastropub	Reservoir	Hotel	Zoo	Flower Shop	Farm	Farmers Market	Fast Food Restaurant	Field	Film Studio
97	50.794971	1	Hotel	Train Station	Go Kart Track	Zoo	Farmers Market	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market
119	51.935921	1	Hotel	Train Station	Bar	Golf Course	Mobile Phone Shop	Flower Shop	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop
121	51.772549	1	Hotel	Bar	Bed & Breakfast	Zoo	Food	Field	Film Studio	Fish & Chips Shop	Fish Market	Flea Market
126	51.934849	1	Gastropub	Hotel	Bar	Pub	Golf Course	Zoo	Flower Shop	Fast Food Restaurant	Field	Film Studio
174	52.537659	1	Garden Center	Hotel	Fast Food Restaurant	Restaurant	Farm	Farmers Market	Field	Film Studio	Fish & Chips Shop	Fish Market
181	53.264832	1	Bar	Zoo	Gas Station	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market	Flea Market	Flower Shop
190	53.400440	1	Café	American Restaurant	Antique Shop	Hotel	Zoo	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market

- Second Cluster
- Analysis: Here there are pubs and café but compared to first cluster we see gardens, grocery stores, park, pharmacy, Zoo, bookstores. This is ideal for a family.

Result (contd)

	latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	54.676159	2	Pub	Fast Food Restaurant	Hotel	Supermarket	Pizza Place	Pharmacy	Clothing Store	Soccer Stadium	Movie Theater	Gas Station
8	53.334240	2	Grocery Store	Coffee Shop	Pub	Fast Food Restaurant	Gym / Fitness Center	Sandwich Place	Pizza Place	Hotel	Supermarket	Gastropub
13	53.523361	2	Grocery Store	Pub	Fast Food Restaurant	Italian Restaurant	Pizza Place	Chinese Restaurant	Supermarket	Hotel	Gas Station	Convenience Store
22	51.356041	2	Pub	Hotel	Bar	Rental Car Location	Theater	Construction & Landscaping	Gastropub	Grocery Store	Paintball Field	Snack Place
26	51.397072	2	Pub	Grocery Store	Hotel	Convenience Store	Supermarket	Coffee Shop	Gastropub	Gym / Fitness Center	Airport Lounge	Garden Center
47	53.167931	2	Gastropub	Grocery Store	Paintball Field	Bookstore	Farmers Market	Gym	Fast Food Restaurant	Nature Preserve	Sandwich Place	Clothing Store
59	51.900379	2	Grocery Store	Bar	Bed & Breakfast	Café	Fast Food Restaurant	Film Studio	Fish & Chips Shop	Fish Market	Flea Market	Flower Shop
60	51.559441	2	Gastropub	Grocery Store	Pub	Hotel	Coffee Shop	Film Studio	Park	Supermarket	Café	Bar
65	52.535439	2	Warehouse Store	Fast Food Restaurant	Construction & Landscaping	Restaurant	Zoo	Farmers Market	Field	Film Studio	Fish & Chips Shop	Fish Market
71	54.157372	2	Grocery Store	Pub	Hotel	Coffee Shop	Pizza Place	Train Station	Warehouse Store	Zoo	Supermarket	Gym / Fitness Center
89	50.757648	2	Supermarket	Gastropub	Hotel	Garden Center	Grocery Store	Warehouse Store	Hardware Store	Gas Station	Pub	Food Truck
90	50.610199	2	Grocery Store	Café	Brewery	Trail	Resort	Miscellaneous Shop	Hotel	Outdoor Supply Store	Zoo	Flea Market
104	51.916340	2	Supermarket	Coffee Shop	Restaurant	Gym	Golf Course	Shopping Mall	Sandwich Place	Video Game Store	Pharmacy	Gym / Fitness Center

- Third Cluster
- Analysis: There are lot of markets, Historic places, Train station, park so this is a good choice for a family with children.
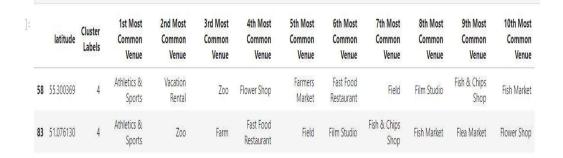
Result (contd)

	latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	54.544670	3	Pub	Clothing Store	Coffee Shop	Pizza Place	Café	Hotel	Supermarket	Park	Portuguese Restaurant	Bistro
2	54.567520	3	Pub	Coffee Shop	Convenience Store	Supermarket	Hotel	Beach	Grocery Store	Café	Seafood Restaurant	Gastropub
3	54.556911	3	Pub	Park	Clothing Store	Pizza Place	Gym / Fitness Center	Gym	Italian Restaurant	Gastropub	American Restaurant	Caribbean Restaurant
4	53.769791	3	Pub	Coffee Shop	Bar	Café	Nightclub	Park	Italian Restaurant	Portuguese Restaurant	Movie Theater	Spanish Restaurant
5	53.881222	3	Supermarket	Fast Food Restaurant	Gas Station	Pub	Indian Restaurant	Outdoors & Recreation	Flower Shop	Food	Field	Film Studio
6	51.891022	3	Pub	Grocery Store	Clothing Store	Gym / Fitness Center	Hotel	Fast Food Restaurant	Supermarket	Park	Indian Restaurant	Coffee Shop
7	54.535351	3	Pub	Italian Restaurant	Hotel	Coffee Shop	Grocery Store	Clothing Store	Department Store	Athletics & Sports	Pizza Place	Supermarket
9	52.635921	3	Pub	Coffee Shop	Bar	Indian Restaurant	Portuguese Restaurant	Park	Sandwich Place	Shopping Mall	Movie Theater	Pizza Place
10	53.391628	3	Pub	Coffee Shop	Bar	Gym / Fitness Center	Italian Restaurant	Grocery Store	Chinese Restaurant	Gastropub	Tapas Restaurant	Music Venue
11	53.700802	3	Pub	Supermarket	Coffee Shop	Clothing Store	Sandwich Place	Fast Food Restaurant	Grocery Store	Gym	Diner	Chinese Restaurant
12	53.821610	3	Pub	Coffee Shop	Café	Ice Cream Shop	Theater	Fish & Chips Shop	Hotel	Italian Restaurant	Pizza Place	Beach
14	53.965820	3	Pub	Historic Site	Coffee Shop	Bar	Café	Hotel	Park	Movie Theater	Tea Room	Beer Store
18	52.954189	3	Pub	Coffee Shop	Indian Restaurant	Bar	Bookstore	Café	Burger Joint	French Restaurant	Movie Theater	Lounge

- Fourth Cluster

- Analysis: As we can see here in first cluster there are lot of pubs, coffee shops, restaurants and cafes. This may not be ideal for a family with children like my client. This is good for people who hang out often and enjoy partying.

Result (contd)

	latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
58	55.300369	4	Athletics & Sports	Vacation Rental	Zoo	Flower Shop	Farmers Market	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market
83	51.076130	4	Athletics & Sports	Zoo	Farm	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market	Flea Market	Flower Shop

- Fifth Cluster
- Analysis:There are family friendly places here. As we see Chesterfield, Doncaster and others have good greenery, supermarkets.

Discussion

- First of all, even though the England Housing Market is always in demand in people who want to settle in United Kingdom.

- We may discuss our results under two main perspectives.

- First, we may examine them according to neighbourhoods. It is interesting to note that, although Notting Hill, Kensington, Chelsea, Marylebone and might be considered highly profitable venues to purchase a real estate according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores, Basingstoke, Kent, Chesterfield, Doncaster are arising as next future elite venues with a wide range of amenities and facilities. Accordingly, one might target under-priced real estates in these areas of England in order to decide on settling down.

Discussion(contd)

- Second, we may analyse our results according to the five clusters we have produced. Even though, all clusters could praise an optimal range of facilities and amenities, we have found two main patterns. The first pattern we are referring to, i.e. Clusters 0, 3 and 4, may target home buyers prone to live with children areas with supermarkets, parks, waterfronts. Instead, the second pattern we are referring to, i.e. Clusters 1 may target individuals who love pubs, theatres and soccer.

- As my client is an Asian, I would suggest areas around Birmingham, Reading, Kent where more of Asian population are present with more Asian Restaurants and other cultural amenities.
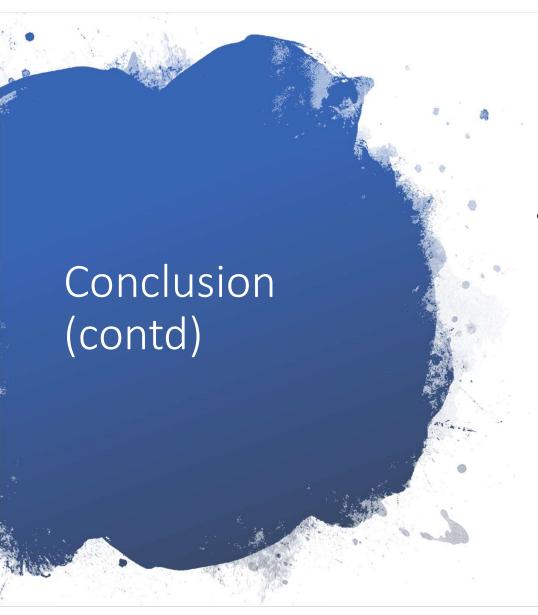
Conclusion

- UK housing prices is facing several different headwinds, including the prospect of higher taxes and a warning from the Bank of England that U.K. home values could fall as much as 30 percent in the event of a disorderly exit from the European Union. In this scenario, it is urgent to adopt machine learning tools in order to assist homebuyers in England to make wise and effective decisions. As a result, the business problem we were posing was : Helping one of my client to select a place to settle with family and children.

- To solve this business problem, we clustered England neighbourhoods in order to recommend venues where homebuyers can make a real estate investment. We recommended profitable venues according to amenities and essential facilities surrounding such venues i.e. grocery stores, parks, restaurants.

Conclusion (contd)

- First, we gathered data on England neighbourhood from the UK govt website. Moreover, to explore and target recommended locations across different venues according to the presence of amenities and essential facilities, we accessed data through Four Square API interface and arranged them as a data frame for visualization. By merging data on Neighbourhoods and data on amenities and essential facilities surrounding such properties from Four Square API interface, we were able to recommend profitable real estate investments.

- Second, The Methodology section comprised four stages:

1. Collect Inspection Data

2. Explore and Understand Data

3. Data preparation and pre-processing

4. Modelling. In the modelling section, we used the k-means clustering technique as it is fast and efficient in terms of computational cost, is highly flexible to account for mutations in real estate market in England and is accurate.

Conclusion (contd)

- We drew conclusion that England Housing Market is always in demand in people who want to settle in United Kingdom.

- We may discuss our results under two main perspectives.

- First, we may examine them according to neighbourhoods. It is interesting to note that, although Notting Hill, Kensington, Chelsea, Marylebone and might be considered highly profitable venues to purchase a real estate according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores, Basingstoke, Kent, Chesterfield, Doncaster are arising as next future elite venues with a wide range of amenities and facilities. Accordingly, one might target under-priced real estates in these areas of England in order to decide on settling down.

- Second, we may analyze our results according to the five clusters we have produced. Even though, all clusters could praise an optimal range of facilities and amenities, we have found two main patterns. The first pattern we are referring to, i.e. Clusters 1, 2 and 4, may target home buyers prone to live with children areas with supermarkets, parks, waterfronts. Instead, the second pattern we are referring to, i.e. Clusters 1 may target individuals who love pubs, theatres and soccer.

Conclusion (contd)

- As my client is an Asian, I would suggest areas around Birmingham, Reading, Kent where more of Asian population are present with more Asian Restaurants and other cultural amenities.