Neighborhood analysis for Airbnb listing in Bristol, United Kingdom

Introduction

Joanna and Jeff are couple lives in Bristol, United Kingdom.

Jeff is a football coach and Joanna works nearby Tesco.

They want to list their spare rooms in Airbnb, for additional income.

Before listing, they want to analyze existing Airbnb listings in Bristol area.

Typical questions in their mind are:

> What is the distribution of Airbnb listing in Bristol?
>
> What are the characteristics of highly concentrated listing, if there are any?
>
> What is the range of listing prices in Bristol?

They discussed their idea and questions in their mind with one of their friend Anderson.

Anderson is a Data scientist, agreed to help them out.

Data

Inside Airbnb is an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world.

By analyzing publicly available information about a city's Airbnb's listings, Inside Airbnb provides filters and key metrics so you can see how Airbnb is being used to compete with the residential housing market.

Bristol Airbnb listing data set consists of **2500+ listings**, which are compiled on 19 May, 2019.

It consists of many useful fields like:

- Neighbourhood
- Geo coordinate of the property
- Room type
- Price

Example from Bristol Airbnb dataset

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	number_of_reviews	last_review	reviews_per_m
0	70820	City View - Sarah's double room.	360195	Sarah	NaN	Windmill Hill	51.43994	-2.59173	Private room	28	7	138	2019-05-03	2.03
1	117122	City Centre-Waterside Retreat	591555	Marcus	NaN	Clifton	51.45051	-2.61054	Private room	65	1	131	2019-03-17	1.37
2	146407	Sunny Central Artist Cottage (Dbl)	708175	Orla	NaN	Southville	51.44131	-2.60271	Private room	38	2	65	2019-04-28	0.80

Fetch nearby venues using Forsquare APIs

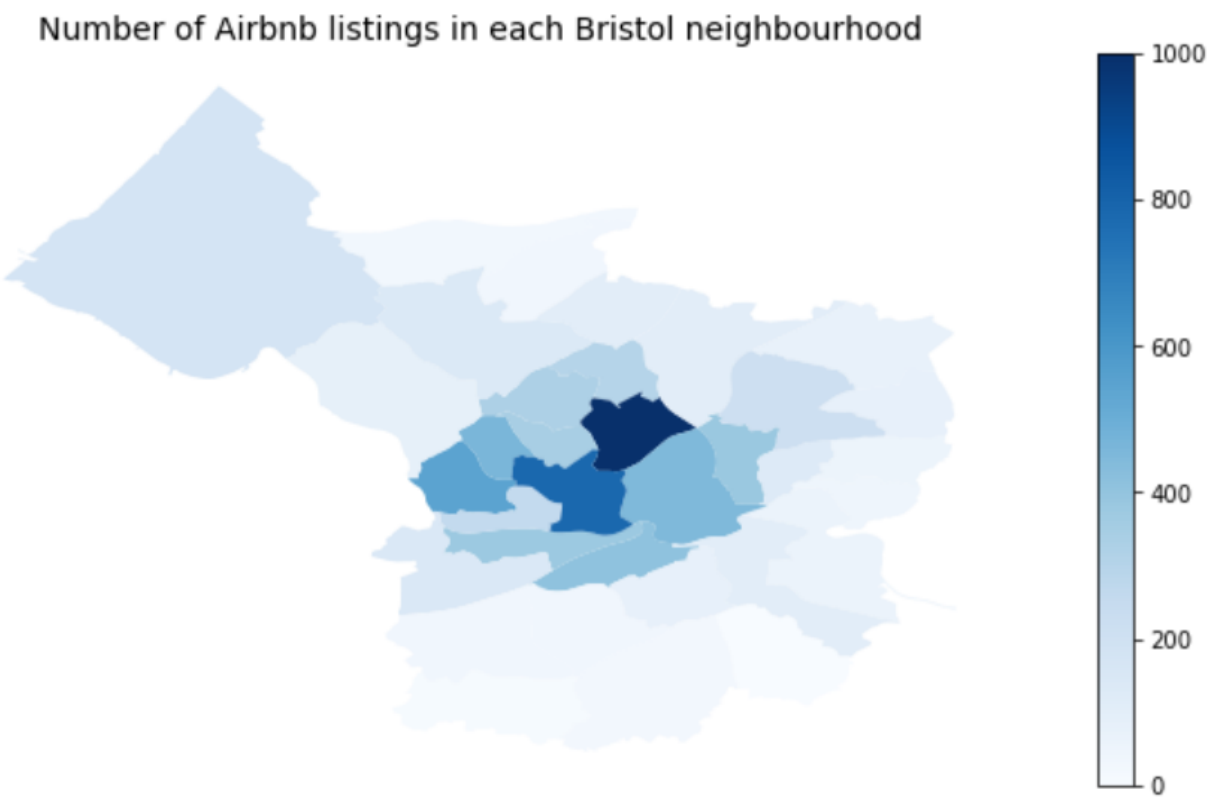
This data set is supplied with Neighbourhood json file.

By using Forsquare APIs, near by venues of each neighborhood can be fetched. Which will be used to provide characteristics for each listings.

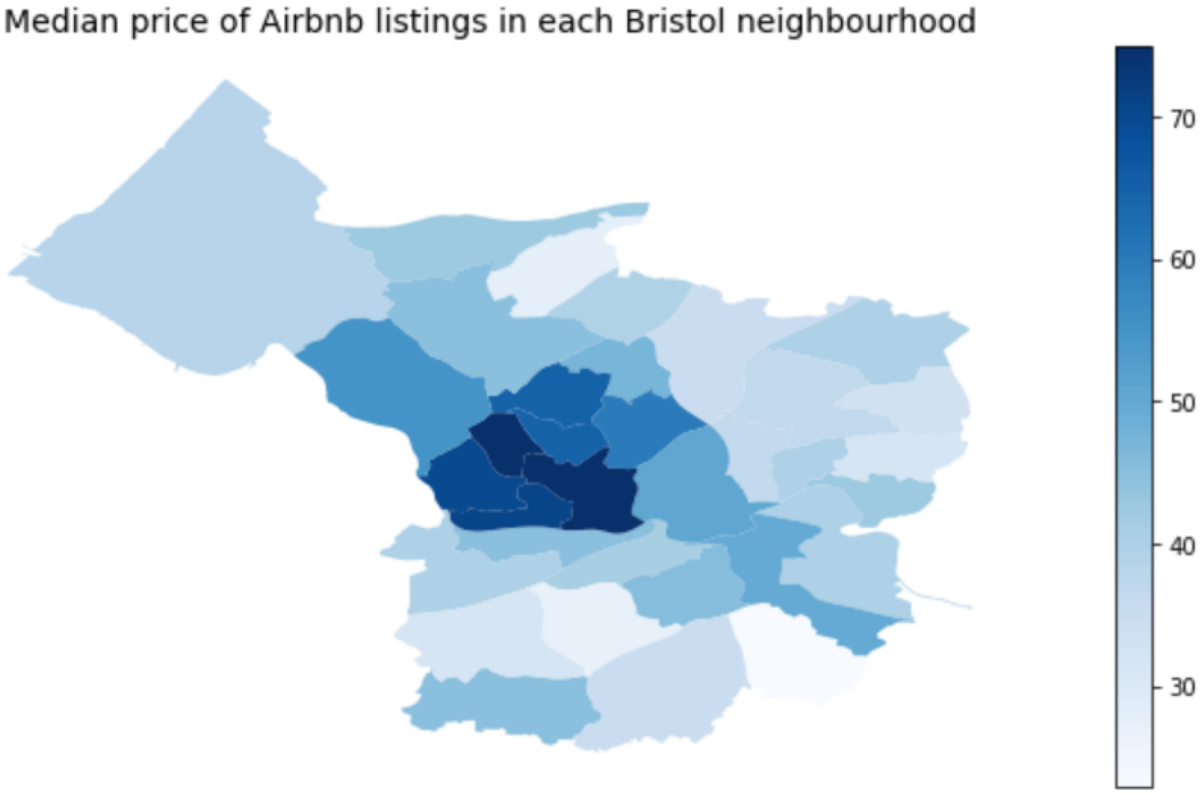
Methodology

Exploratory data analysis

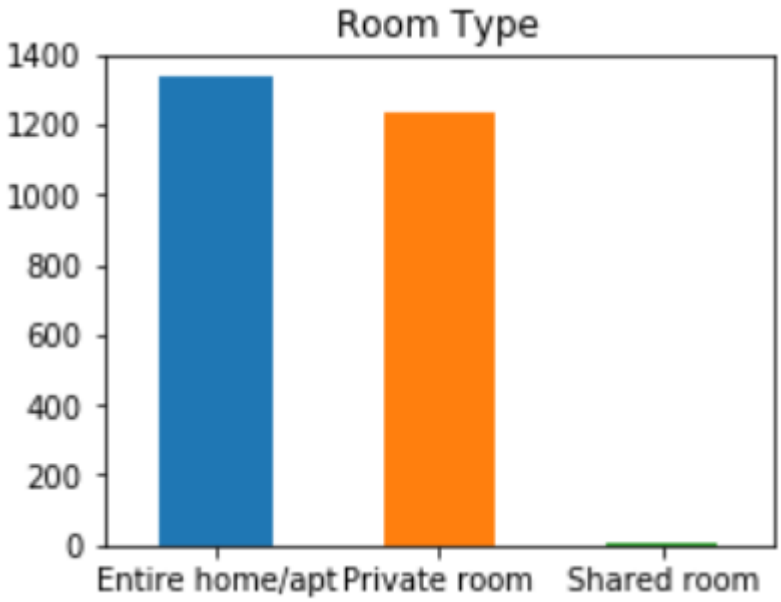
> **Which areas of Bristol have the most Airbnb properties?**
>
> Answer: Ashely, Central and Clifton



> **Which are the most expensive areas?**
>
> Answer: Central, Clifton Down and Hotwells & Harbourside



> **What are the most common property and room types?**



> **Most common venues in Bristol Neighborhood - Fetched using Foursquare API**

	Neighborhood	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	Ashley	Pub	Food & Drink Shop	Indian Restaurant	Food	Train Station	Grocery Store	Park	Mediterranean Restaurant	Dessert Shop	Electronics Store
1	Avonmouth & Lawrence Weston	Coffee Shop	Wine Shop	Convenience Store	Deli / Bodega	Dessert Shop	Electronics Store	English Restaurant	Falafel Restaurant	Fast Food Restaurant	Fish & Chips Shop
2	Bedminster	Soccer Stadium	Hardware Store	Electronics Store	Fast Food Restaurant	Supermarket	Wine Shop	Cosmetics Shop	Deli / Bodega	Dessert Shop	English Restaurant
3	Bishopston & Ashley Down	Pub	Café	Grocery Store	Fried Chicken Joint	Chinese Restaurant	Fish & Chips Shop	Middle Eastern Restaurant	Bakery	Coffee Shop	Food
4	Bishopsworth	Fast Food Restaurant	Grocery Store	Business Service	Bus Stop	Wine Shop	Food & Drink Shop	Dessert Shop	Electronics Store	English Restaurant	Falafel Restaurant

Preparing data for Airbnb listing clustering

To cluster property listing, data is prepared by merging property listing data (sourced from inside airbnb dataset) with venue details from foursquare apis.

Each property listing is enriched by adding venues from the neighborhood, where the property is belongs to.

Final data frame consists of 2558 rows and 117 columns

Clustering algorithm used:

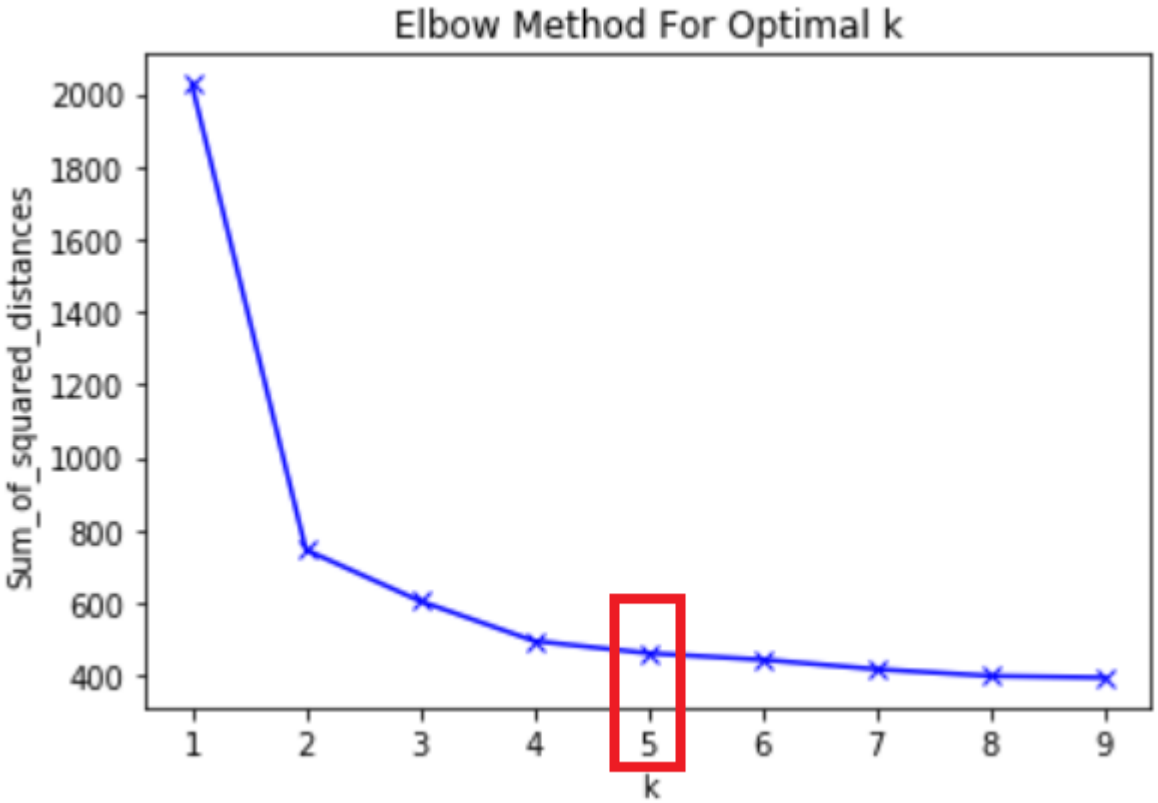
K-Means

Reason for choosing:

Data set consists of continuous variables.

K-Means is one of the simplest model based on centroids

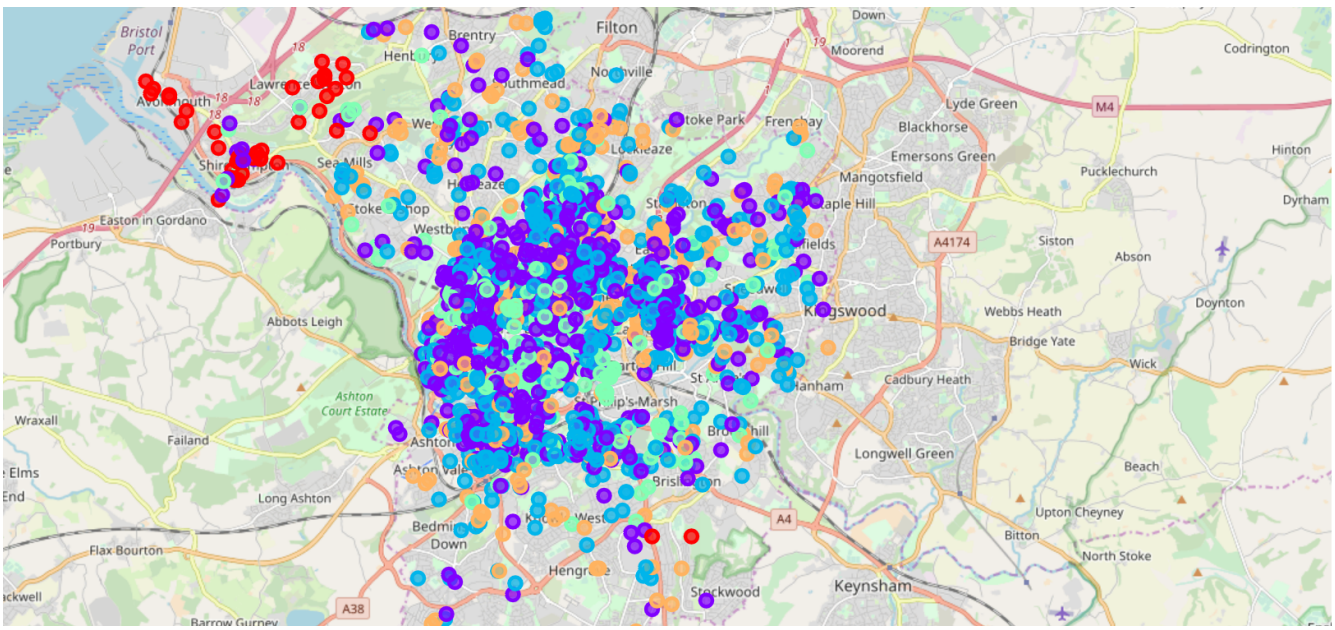
> **What is the optimal no. of clusters?**
>
> Find the optimal no of clusters by using 'Elbow method'



No. of features used for clustering: 151

Result

After applying clustering on the data set, Airbnb listing is mapped on Bristol



Discussion

With above clustering result, east clusters are examined further and provided a summary of each clusters.

Cluster Analysis

Cluster	Price	Room Type	Minimum Nights	Avg. yearly availability	Most common Venues
0	Low priced listings.Mostly below £40	All Private rooms	Mostly 1 night	162 days	Very few venues. Only Cafes are available.
1	Medium priced listings. Peaks nearly £100	Whole property	Mostly 2 nights	47 days	Most happening place with lots of pubs and parks
2	Low priced listings mostly less than £50	99% are private rooms	Mainly 1 night	42 days	Most happening place with lots of pubs and parks
3	Mix of pricings. Peak near £50. Many listings spotted above £100	Whole property	1 & 2 nights are equal	302 days	Lots of venues found. Eateries, parks and playgrounds.
4	Low priced listings mostly less than £50	99% are private rooms	Mostly 1 night. Also peaks in 2 nights	276 days	Lots of venues found. Pubs, Parks, Indian restaurants.

Conclusion

With above clustering analysis, it is noted that Clusters 1,2 and 3 are most sought after listings in Bristol with plenty of amenities around.

Property type of listings in Cluster 3, are whole properties and it is available for most of the time in an year.

If someone need whole property rent, they can look into Cluster 3.