WHERE TO SET UP ENTERTAINMENT VENUES IN DOWNTOWN BIRMINGHAM?

Data Science Capstone - IBM Data Science Professional Certificate on Coursera

Introduction:

Citizens and tourists tend to entertain themselves after meals, and countries with high rates of eating out also tend to entertain outdoors. So we have reason to believe that the restaurant industry and the entertainment industry are closely related. Does that mean that if someone sets up an entertainment place near a restaurant, such as a movie theater or a video game hall, to make it easier for people coming out of the restaurant to have fun nearby, these entertainment places will get more people and more money after the meal time, so as to get higher income?

The project will focus on helping entertainment operators locate "after-dinner opportunities", using APIs and simple machine learning algorithms.

Data Section:

Starting from downtown Birmingham, we will use the Foursquare API to find the restaurant address nearby as the data source, and use the machine learning method (K-means) to find the location core of these restaurants. Note that since the API I use is free, the API limits the number of places to 50 at a time, which somewhat affects the accuracy of our results.

Method:

First I need to find the latitude and longitude coordinates of downtown Birmingham as the starting point of our search. This can be easily found

on geopy. At the same time, geopy's coordinates are compared with those of the city of Birmingham on Google maps, with negligible errors.For convenience, the coordinates provided by geopy are used here.
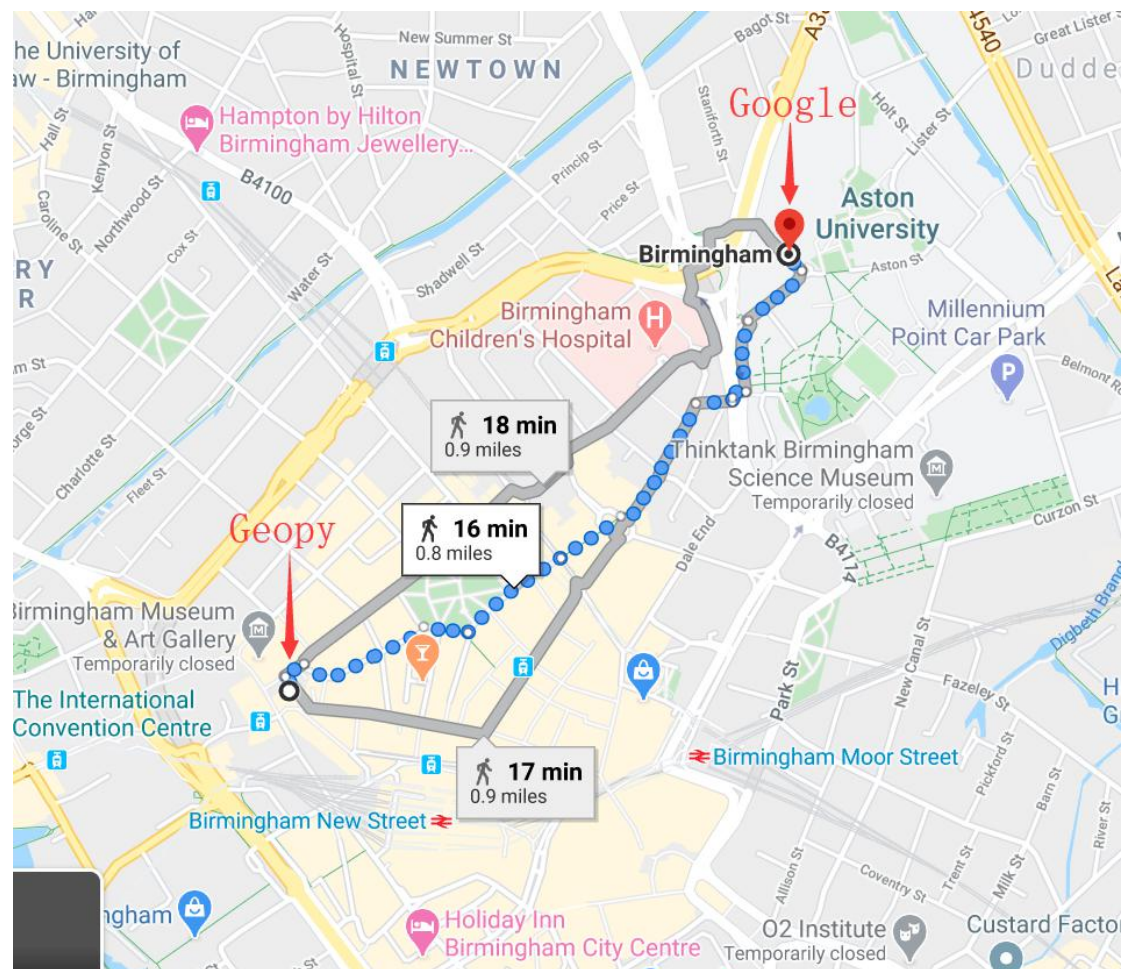

Second, we use the Foursquare API to search for nearby restaurants (with "restaurant" as the keyword). An example of my query:

https://api.foursquare.com/v2/venues/search?client_id=IG50DNURFEL51PCLICE1GHMPCLUR KVZF4T0LL4V0RBP5JG1W&client_secret=EW5E4KW1I3M4FM01JUVXM132X41UHJMIK C1TA2Q5NZH4CFU2&ll=52.4796992,-1.9026911&v=20180604&query=restaurant&radius=200 0&limit=100

At this point, we have the details of 50 restaurants and have converted them through a series of transformations (Jason data to dataframe).

The coordinates that we're going to get are shown by folium.
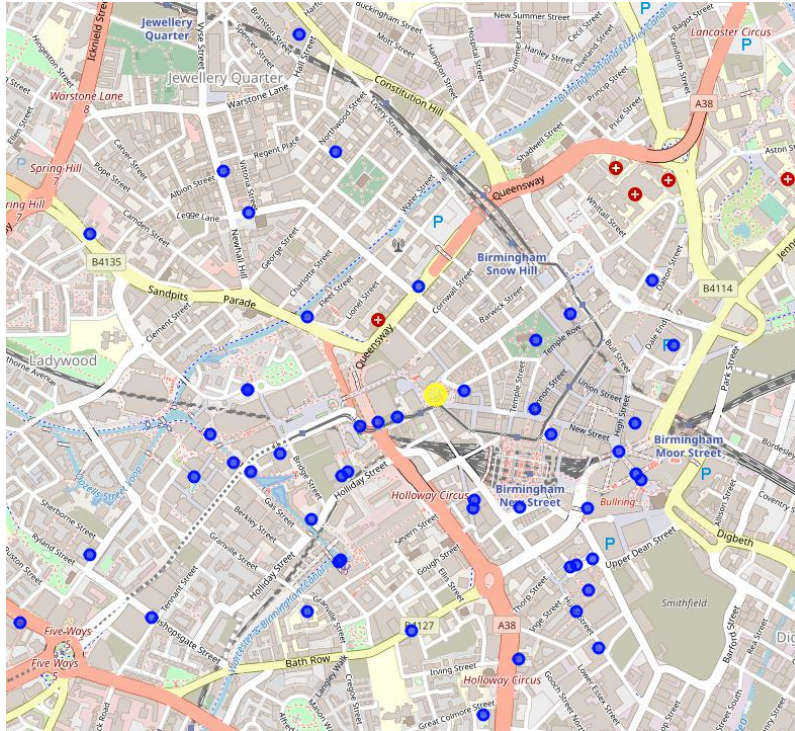
We create a new data set that includes the coordinates of only 50 restaurants. The k-mean algorithm library of sklearn was imported, and the number of k-mean clusters was set to 4 (which could be adjusted according to the actual needs). The coordinates of the four clusters were finally obtained after 200 repeated operations. Run the fitting program to get the coordinates of the four clusters.Use folium to visualize it.

Results:

The coordinates of the four clusters are shown below.

	lat	lng
0	52.477241	-1.915837
1	52.477392	-1.907625
2	52.477227	-1.897066
3	52.484841	-1.908219

The display on the map is shown below.
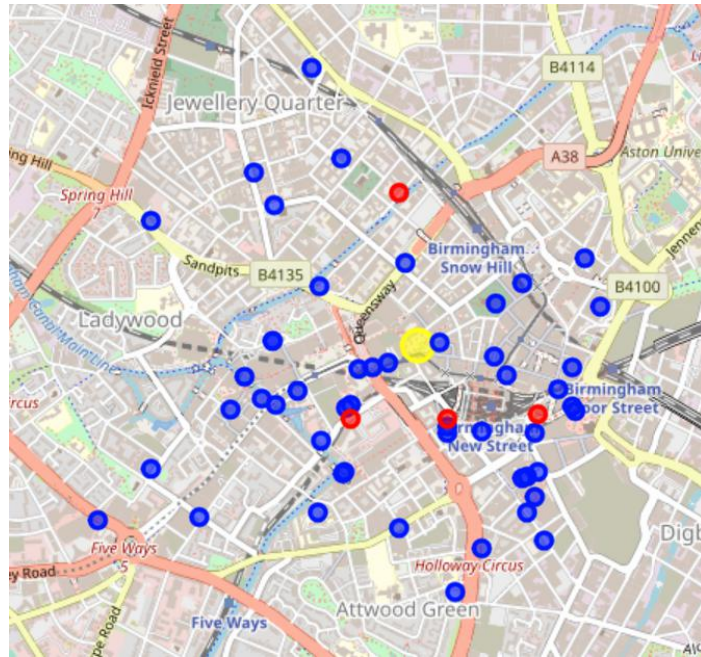
As a result, we can see that restaurants are concentrated in the south of the city center.

Discussion:

There are undoubtedly more than 50 restaurants within two miles of Birmingham city centre, but due to our free API we can only return 50 restaurants at a time. If we had location information for 100 or even hundreds of restaurants, we would end up with more accurate results. Similarly, the calculation times of k-mean algorithm also have a great impact on the final result, but this requires more powerful computing resources.

Conclusion:

Now that, we recommend the establishment of entertainment venues near the following four locations.:
1. Holiday Street (Near the Parking lot)
2. The junction of Lower Seven Street and Hill Street
3. St Martins Queens Way (the southeast of Newstreet Railway Station)