Capstone Project
Applied Data Science Capstone, IBM Data Science Certificate

Topic: To Find the Best Location to Operate A Futuristic High-Tech Bar
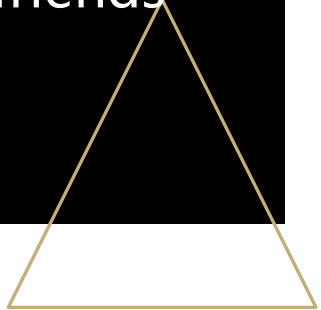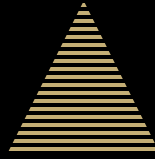
X FAN

25 APRIL 2020

1. Introduction

- In this project, we will explore Central London *('Western Central London' and 'Eastern Central London')* and try to find the best location to operate a futuristic high-tech bar.

1.1 Project Background

- The pace of work and life in London is very fast, especially in Central London.

- The advancement of technology brought convenience but also great pressure.

- Many pubs, cafes, and bars in Central London have built for a long history and lack leisure atmosphere that incorporate smart technology.

- We hope to open a futuristic bar where daily life and digital technology can coexist.

- Young and middle-age science and technology professionals can also enjoy the welfare and happiness of life brought by technological progress.

- They can conveniently find the bar, have a drink after work and make geek friends face-to-face, not just on GitHub, to achieve a real work-life balance.

1.2 Business Problem

Different boroughs and locations in Central London have distinct culture background, business environment, public cognition and functional infrastructures. Site selection is particularly important.

4 key factors needed to be considered for site selection:

- **Target working place cluster:** Industrial cluster of technology startups and Internet companies

- **Target people cluster:** People there have high income/ customer consumption level, but also with high working press

- **Accessibility:** Convenient access for people who work nearby

- **Culture of having a drink after work:** Several pubs or bars already existed, so people have habits to drink and relax there
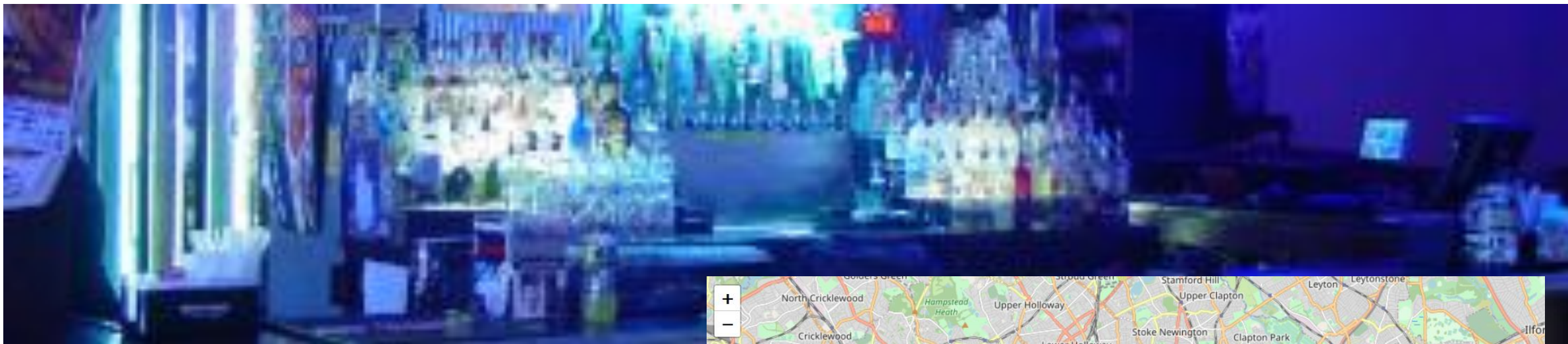
2. Data preparation

- Part1: Scrapping Location Name, London Borough, Post Town, Postcode District, etc for each location in London from Wikipedia. URL: https://en.wikipedia.org/wiki/List_of_areas_of_London

- Part2: Get latitude and longitude of these locations via Geocoder package.

- Part3: Using Foursquare API to get venue data related to these locations.
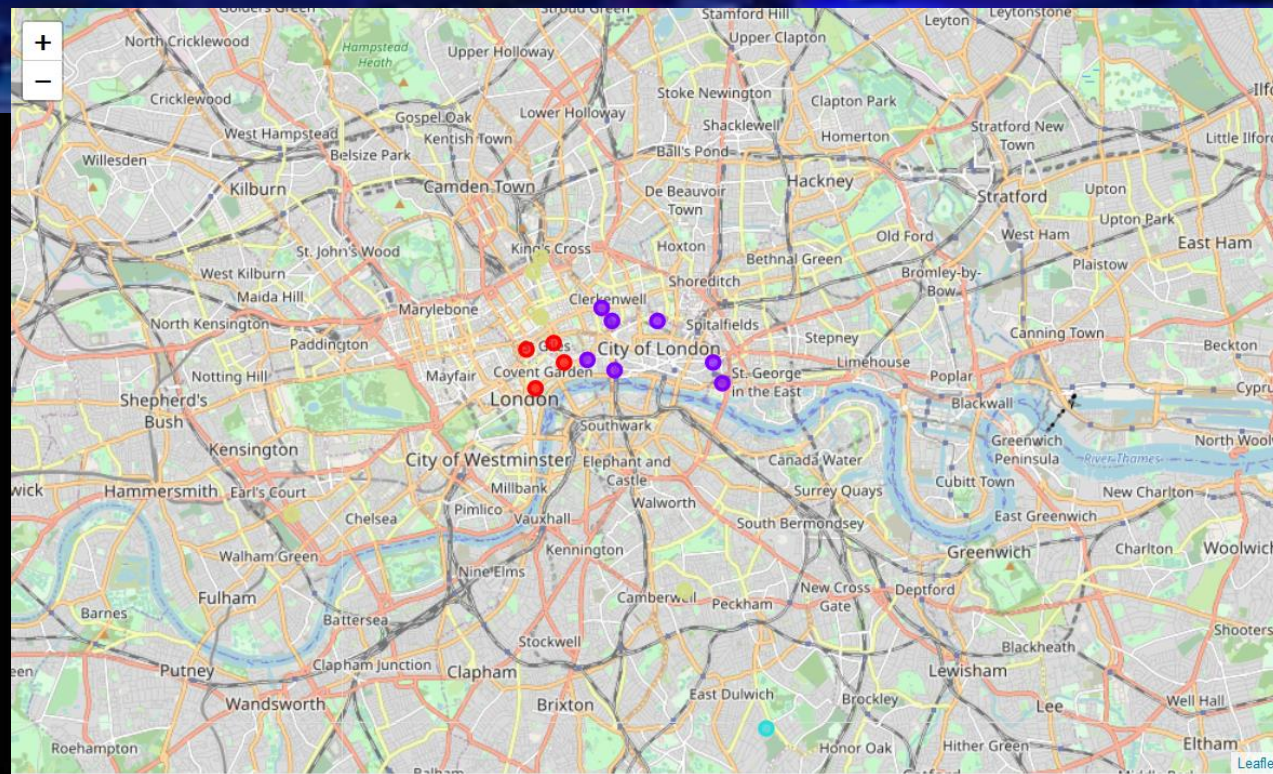
3. Methodology

- We use K-means clustering to find similar locations in Central London that meet our requirements on site selection based on the venues and amenities around each location. K - means clustering is a modelling way of unsupervised machine learning that clusters data based on predefined cluster size.

4. Analysis

- We analyze these 4 clusters to identify the characteristics of each location in Central London.

Cluster 1: Social activities; Culture and Art; Tourism

- Made of 4 locations in Western Central London.

- Most common venues are **theater, coffee shop. hotel, dessert and bakery shop**.

- We can image that people and tourists gather together to watch operas, buy ice creams and desserts in the street shops and settle down in nearby hotels.

	Location	Borough	Postcode	Post town	Latitude	Longitude	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
0	Aldwych	Westminster	WC2	LONDON	51.513307	-0.117092	0	Theater	Hotel	Ice Cream Shop	Hotel Bar	Event Space	Pub	Lebanese Restaurant
2	Charing Cross	Westminster	WC2	LONDON	51.509010	-0.124720	0	Theater	Ice Cream Shop	Dessert Shop	Hotel	Cocktail Bar	Liquor Store	Burger Joint
3	Covent Garden	Westminster	WC2	LONDON	51.515440	-0.126930	0	Theater	Coffee Shop	Bakery	Steakhouse	Cocktail Bar	Performing Arts Venue	Pizza Place
4	Holborn	Camden	WC1	LONDON	51.516510	-0.119680	0	Coffee Shop	Theater	Restaurant	Hotel	Café	Burger Joint	English Restaurant

Cluster 2: Pub; Creative lifestyle

- Made of 7 locations in Eastern Central London.

- They present a **creative, energetic, young and vigorous lifestyle**.

- Most common venues are **pub, coffee shop, hotel, gym/fitness center, garden/park, salad/falafel restaurant**, as well as venues that are not common in other clusters such as **indie movie theater, performing arts venue, cocktail bar**.

	Location	Borough	Postcode	Post town	Latitude	Longitude	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
8	Aldgate	City	EC3	LONDON	51.513308	-0.077762	1	Hotel	Coffee Shop	Gym / Fitness Center	Cocktail Bar	Salad Place	Street Food Gathering	Garden	French Restaurant
10	Barbican	City	EC1	LONDON	51.520050	-0.092470	1	Gym / Fitness Center	Coffee Shop	Indie Movie Theater	Food Truck	Performing Arts Venue	Sushi Restaurant	Lake	Salad Place
11	Blackfriars	City	EC4	LONDON	51.512040	-0.103560	1	Coffee Shop	Falafel Restaurant	Hotel	Cocktail Bar	Pub	Gym / Fitness Center	Garden	Indian Restaurant
12	Clerkenwell	Islington	EC1	LONDON	51.522165	-0.107152	1	Pub	Coffee Shop	Hotel	Falafel Restaurant	Beer Bar	Food Truck	Flea Market	Electronics Store
13	Farringdon	Islington & City	EC1	LONDON	51.520150	-0.104510	1	Coffee Shop	Hotel	Wine Bar	Gym / Fitness Center	Pub	Church	Bar	Donut Shop
16	Temple	City, Westminster	EC4	LONDON	51.513679	-0.110885	1	Pub	Coffee Shop	History Museum	Hotel	Garden	Park	Distillery	Restaurant
17	Tower Hill	Tower Hamlets	EC3	LONDON	51.509788	-0.075153	1	Hotel	Gym / Fitness Center	Garden	Castle	French Restaurant	Hostel	Italian Restaurant	Park

Cluster 3: Far away from the center of Central London

- Made of 1 location in Eastern Central London but far away from the center of Central London.

- We could obviously **leave out this cluster.**

	Location	Borough	Postcode	Post town	Latitude	Longitude	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
9	Angel	Islington	EC1	LONDON	51.45302	-0.0634	2	Pub	Café	Garden	Park	Grocery Store	Indian Restaurant	Fish & Chips Shop	Business Service

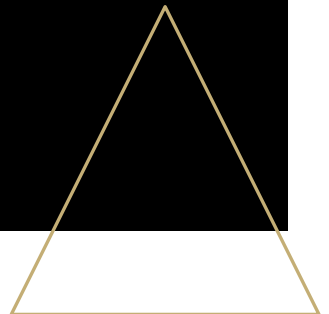# Cluster 4: Traditional lifestyle; Transportation

- Made of 6 locations in both Eastern and Western Central London.

- Leave out St Luke, St Giles and Finsbury; far away from the center of Central London.

- Most common venues in the remaining three locations are **café, coffee shop, hotel, restaurant,** as well as **train station**.

- They present a **traditional European lifestyle**.

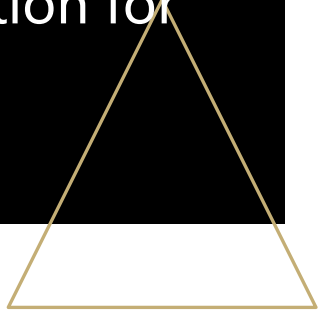|	Location	Borough	Postcode	Post town	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
1	Bloomsbury	Camden	WC1	LONDON	51.52074	-0.123100	3	Coffee Shop	Exhibit	Plaza	Bookstore	Park	Japanese Restaurant	Hotel Bar
5	King's Cross	Camden and Islington	WC1	LONDON	51.53043	-0.123100	3	Hotel	Burger Joint	Breakfast Spot	Café	Train Station	Coffee Shop	English Restaurant
7	St Pancras	Camden	WC1	LONDON	51.52811	-0.125240	3	Café	Hotel	Coffee Shop	Grocery Store	English Restaurant	Falafel Restaurant	Breakfast Spot

5.   Results and Discussion

- **4 Eastern Central London locations in Cluster 2** (Aldgate, Barbican, Blackfriars, Clerkenwell) may meet our requirements.

  ◦ First, there are many pubs so that people there already have the habits on drinking after work, know there are bars and pubs in that borough and will be more likely to try a new bar.

  ◦ Second, the number of gym/fitness center, salad/falafel restaurant and juice bar in Cluster 2 is significantly larger than that of other clusters, indicating that people there choose to keep an energetic and healthy living habit.

  ◦ Moreover, venues for young and fashion people such as indie movie theater, performing arts venue, cocktail bar are especially common in Cluster 2, meaning that people there will try modern and futuristic things.

6.   Conclusion

- The aim of this project is to assist businessmen or investors in making decisions to select a propriate location to operate or invest a futuristic high-tech bar in Central London, by exploring venues around target locations and get preliminary portrait and business environments around them.

- We find locations in Cluster 2 present a creative, energetic, young and vigorous lifestyle so that preliminarily match the theme - 'balancing work and life in high-technology interaction' – of our futuristic high-tech bar.

- Further work should be done to take other factors, such as the number and distribution of high-tech companies, rental fee, theme of pub/bar's nearby, into consideration to shortlist the locations, and present more valuable information for site selection reference.

'Balance work and life in high-technology interaction'

Thanks.