Final Capstone

By: Afnan

05/01/2022

Outline

- Introduction
- The problem
- The data
- Methodology
- Results and discussions
- Conclusion
- Appendix

Introduction

The final course of the Data Science Professional Certificate consist of a capstone project where in all the skills and relevant knowledge that one has gathered from this 9 intense courses has to be applied on a final capstone project.

The final problem as well as the analysis is the left for the reader to explore and decide. The idea uses location data with the help of the foursquare api that can be leveraged into coming up with a problem that the foursquare location data
to solve it or just in contrast to compare cities or neighbourhoods of ones own choice

The Problem

In this ever changing world of technology and reforms the use of AI will dominate and change most of the world and industries as we know so among the two busiest cities in the world which one would a person be willing to start a business in AI. Various factors would be included such as pricing, multiculturism, language barriers and so on would influence this decision.

Paris Dataset

Data

	Place Name	State	County	City	Latitude	Longitude
0	Paris 01 Louvre	Île-de-France	Paris	Paris	48.8592	2.3417
1	Paris 02 Bourse	Île-de-France	Paris	Paris	48.8655	2.3426
2	Paris 03 Temple	Île-de-France	Paris	Paris	48.8637	2.3615
3	Paris 04 Hôtel-de-Ville	Île-de-France	Paris	Paris	48.8601	2.3507
4	Paris 05 Panthéon	Île-de-France	Paris	Paris	48.8448	2.3471

London Dataset

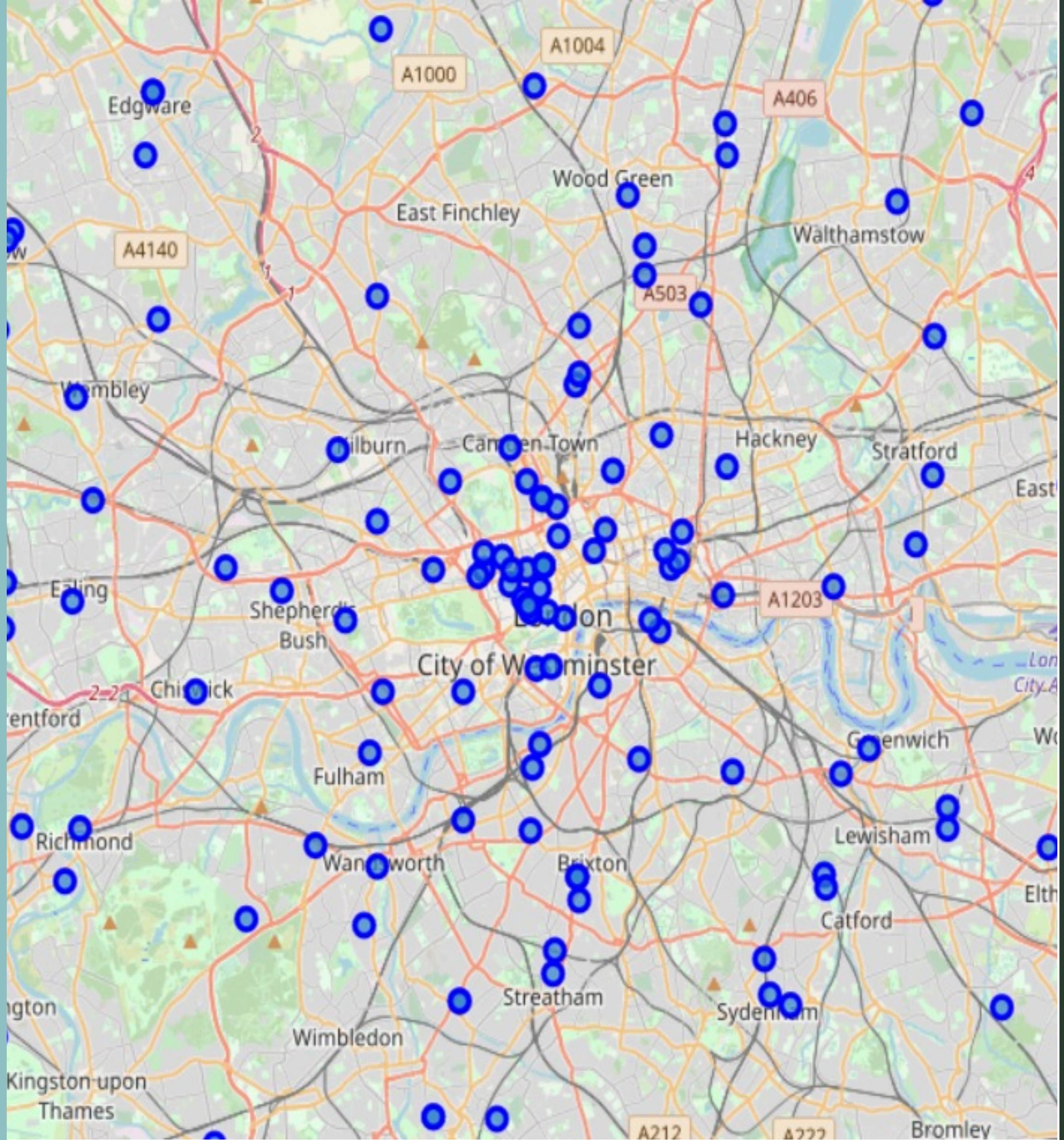
Data

	Postcode	Country	County	District	Latitude	Longitude
0	BR1 1AA	England	Greater London	Bromley	51.401546	0.015415
1	BR1 1AB	England	Greater London	Bromley	51.406333	0.015208
2	BR1 1AD	England	Greater London	Bromley	51.400057	0.016715
3	BR1 1AE	England	Greater London	Bromley	51.404543	0.014195
4	BR1 1AF	England	Greater London	Bromley	51.401392	0.014948

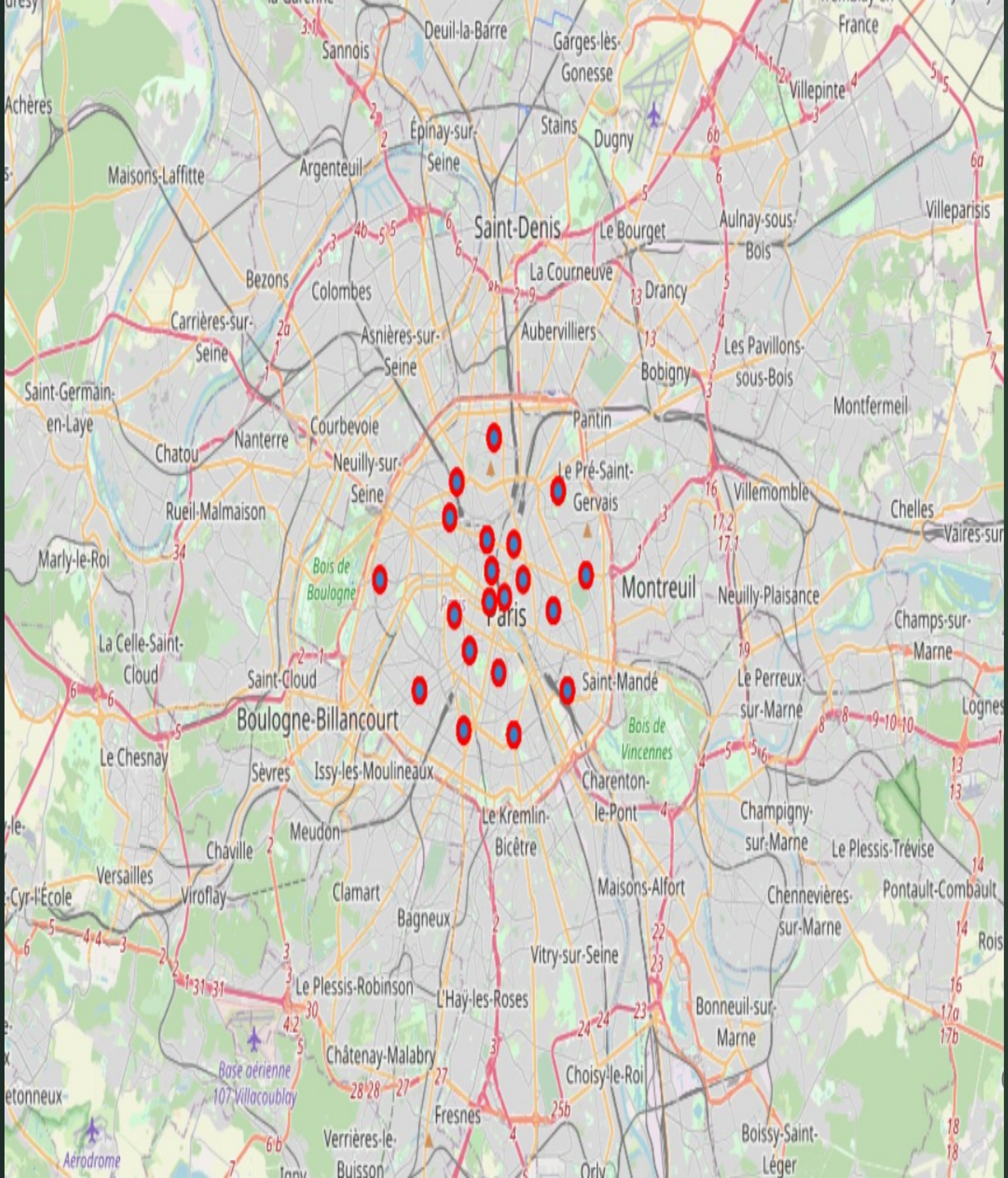
Methodology

An in-depth research of the dataset has been done and a thorough analysis of the various features and methods have been investigated to ensure the maximum accuracy of the model as possible.

After reduction of the number of features in the data frame by replacing them with more useful data cluster analysis was done to find the best cluster of both Paris and London and then correlation and various other visual graphs were used to compare the two cities.


A1004
A1000
A406
East Finchley
Walthamstow
A4140
A503
Hackney
Stratford
A1203
Fulham
Lewisham
Catford
Streatham
Wimbledon
Kingston upon
Thames
Bromley
A212
A222

London Map


Deuil-la-Barre
Garges-lès-Gonesse
Sannois
Villepinte
Achères
Épinay-sur-Seine
Stains
Dugny
Maisons-Laffitte
Argenteuil
Saint-Denis
Le Bourget
Aulnay-sous-Bois
Villeparisis
Bezons
Colombes
La Courneuve
Drancy
Carrières-sur-Seine
Asnières-sur-Seine
Aubervilliers
Les Pavillons-sous-Bois
Bobigny
Saint-Germain-en-Laye
Montfermeil
Courbevoie
Pantin
Nanterre
Chatou
Neuilly-sur-Seine
Le Pré-Saint-Gervais
Villemomble
Chelles
Rueil-Malmaison
Marly-le-Roi
Bois de Boulogne
Montreuil
Neuilly-Plaisance
Paris
Champs-sur-Marne
La Celle-Saint-Cloud
Saint-Cloud
Saint-Mandé
Le Perreux-sur-Marne
Boulogne-Billancourt
Bois de Vincennes
Le Chesnay
Sèvres
Issy-les-Moulineaux
Charenton-le-Pont
Champigny-sur-Marne
Le Kremlin-Bicêtre
Meudon
Chaville
Le Plessis-Trévise
Versailles
Viroflay
Maisons-Alfort
Chennevières-sur-Marne
Pontault-Combault
Clamart
Bagneux
Vitry-sur-Seine
Le Plessis-Robinson
L'Haÿ-les-Roses
Bonneuil-sur-Marne
Châtenay-Malabry
Base aérienne 107 Villacoublay
Choisy-le-Roi
Fresnes
Verrières-le-Buisson
Boissy-Saint-Léger
Aérodrome
Orly

Paris Map

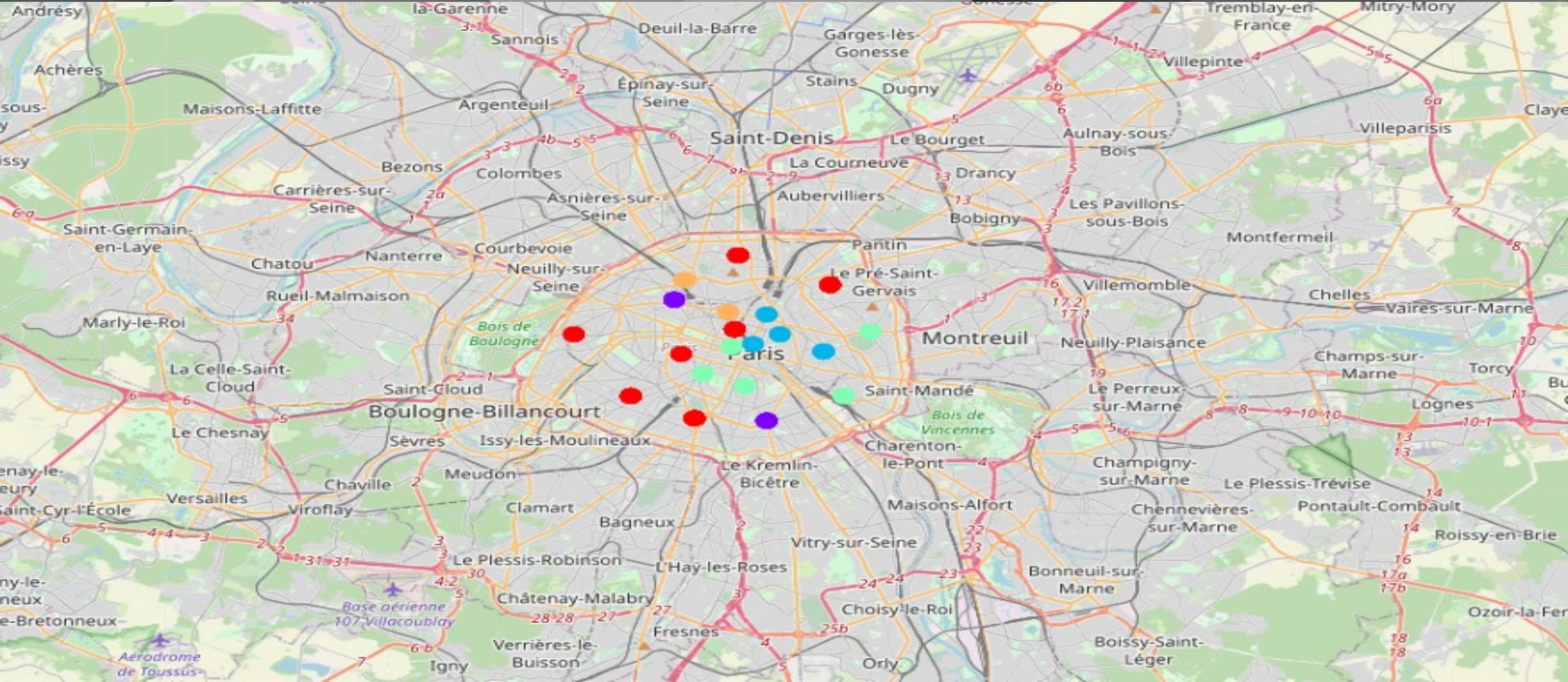
Venues in London

	Cluster Labels	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	3	Paris 01 Louvre	Plaza	French Restaurant	Cocktail Bar	Church	Pedestrian Plaza	Chinese Restaurant	Park	Coffee Shop	Art Gallery	Garden
1	0	Paris 02 Bourse	French Restaurant	Plaza	Bakery	Ramen Restaurant	Restaurant	Souvlaki Shop	Perfume Shop	Bookstore	Farmers Market	Coffee Shop
2	2	Paris 03 Temple	Sandwich Place	Wine Bar	Park	Tea Room	Burger Joint	Restaurant	Cocktail Bar	Seafood Restaurant	Farmers Market	Wine Shop
3	2	Paris 04 Hôtel-de-Ville	Ice Cream Shop	Souvenir Shop	Art Gallery	Art Museum	Cocktail Bar	Fountain	Gourmet Shop	Lebanese Restaurant	Pub	Alsatian Restaurant
4	3	Paris 05 Panthéon	Plaza	French Restaurant	Bar	Korean Restaurant	Monument / Landmark	Science Museum	Ice Cream Shop	Bakery	Creperie	Grocery Store

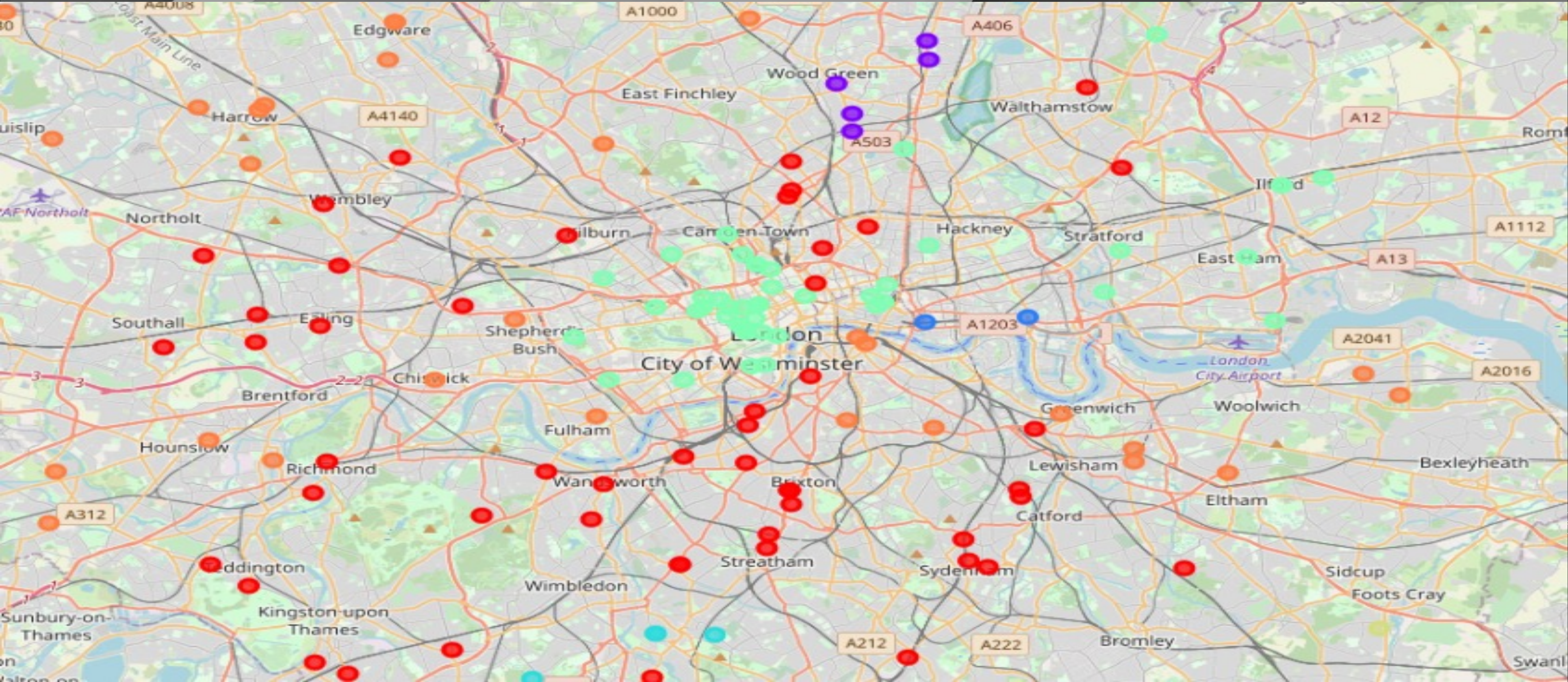
Venues in Paris

	Cluster Labels	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	3	Paris 01 Louvre	Plaza	French Restaurant	Cocktail Bar	Church	Pedestrian Plaza	Chinese Restaurant	Park	Coffee Shop	Art Gallery	Garden
1	0	Paris 02 Bourse	French Restaurant	Plaza	Bakery	Ramen Restaurant	Restaurant	Souvlaki Shop	Perfume Shop	Bookstore	Farmers Market	Coffee Shop
2	2	Paris 03 Temple	Sandwich Place	Wine Bar	Park	Tea Room	Burger Joint	Restaurant	Cocktail Bar	Seafood Restaurant	Farmers Market	Wine Shop
3	2	Paris 04 Hôtel-de-Ville	Ice Cream Shop	Souvenir Shop	Art Gallery	Art Museum	Cocktail Bar	Fountain	Gourmet Shop	Lebanese Restaurant	Pub	Alsatian Restaurant
4	3	Paris 05 Panthéon	Plaza	French Restaurant	Bar	Korean Restaurant	Monument / Landmark	Science Museum	Ice Cream Shop	Bakery	Creperie	Grocery Store

K Means Clustering in Paris Map



K Means Clustering in London Map



Results and Discussion

Similarities:
Both cities are multicultural and diverse in their own ways and share a rich history of their own.
Most of the famous neighbourhoods have a restaurant as its top most Common Venue.

Differences:
While looking at the maps one can observe that Paris is more compact and one can walk around much more freely without the use of transport
In terms of population density Paris definitely outweighs London by a ratio of 4:1.

Results and Discussion

Tech hub							
	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**Total**
San Francisco	£418.08m	£1.83bn	£2.07bn	£4.46bn	£806.03m	£1.84bn	**£11.44bn**
Beijing	£11.66m	£53.75m	£197.32m	£599.56m	£1.63bn	£1.07bn	**£3.57bn**
New York	£79.43m	£165.62m	£318.28m	£667.51m	£593.85m	£1.2bn	**£3.05bn**
Shanghai	–	£1.28m	£400.93m	£16.10m	£1.6bn	£453.61	**£2.47bn**
London	£9.85m	£41.16m	£67.04m	£166.04m	£228.97m	£326.90m	**£839.96m**
Paris	£1.92m	£2.83m	£23.49m	£61.49m	£99.45m	£132.40m	**£321.48m**
Singapore	£13.76m	£13.89m	£70.92m	£55.59m	£106.52m	£30.81m	**£291.49m**
Tel Aviv	£14.80m	£17.12m	£5.49m	£39.04m	£112.25m	£89.01m	**£277.71m**
Berlin	£7.09m	£0.79m	£23.60m	£17.41m	£17.67m	£21.06m	**£87.62m**
Bangalore	£1.31m	£32.29m	£45.75m	£1.96m	£36.71m	£18.65m	**£136.67m**

Conclusion

* Artificial Intelligence is a booming topic and recently more people have started investing into it as well as companies automating their processes

* Both cities offer a wide range of opportunities for anyone starting to invest in Artificial Intelligence or even start a company and various factors were shown.

* Finally a better model could be made by various other methods and much stronger Machine Learning Algorithms like KD Tree which have a much faster run time algorithm.

* Furthermore, clustering however did help us to highlight the most optimal venues and areas.

* Finally correlation does not imply causation and so any result here is subject to change on various other trends and opinions and datasets.

Appendix

Github Repository:

Thank you!