CAPSTONE PROJECT

Lloyd Tatenda Chimanga

Introduction

IN ORDER TO EARN THE DATA SCIENCE PROFESSIONAL CERTIFICATE, STUDENTS MUST COMPLETE A CAPSTONE PROJECT IN WHICH THEY MUST APPLY ALL OF THE KNOWLEDGE AND SKILLS THEY HAVE LEARNED DURING THE NINE RIGOROUS COURSES. THE READER IS FREE TO INVESTIGATE AND MAKE DECISIONS REGARDING THE FINAL ISSUE AS WELL AS THE ANALYSIS. THE CONCEPT MAKES USE OF LOCATION DATA USING THE FOURSQUARE API, WHICH MAY BE USED TO CREATE A PROBLEM AND USE FOURSQUARE LOCATION DATA TO SOLVE IT OR JUST TO CONTRAST OTHER CITIES OR NEIGHBORHOODS OF ONE'S CHOOSING.

Business Problem

WHICH OF THE TWO BUSIEST CITIES IN THE WORLD WOULD AN INDIVIDUAL BE WILLING TO ESTABLISH A BUSINESS IN AI IN? IN THIS CONSTANTLY EVOLVING WORLD OF TECHNOLOGY AND REFORMS, THE USAGE OF AI WILL DOMINATE AND CHANGE THE MAJORITY OF THE GLOBE AND INDUSTRIES, AS WE KNOW. PRICING, MULTICULTURALISM, LANGUAGE HURDLES, AND OTHER VARIABLES WOULD ALL PLAY A ROLE IN THIS CHOICE.

Data

Paris Dataset

	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

Data

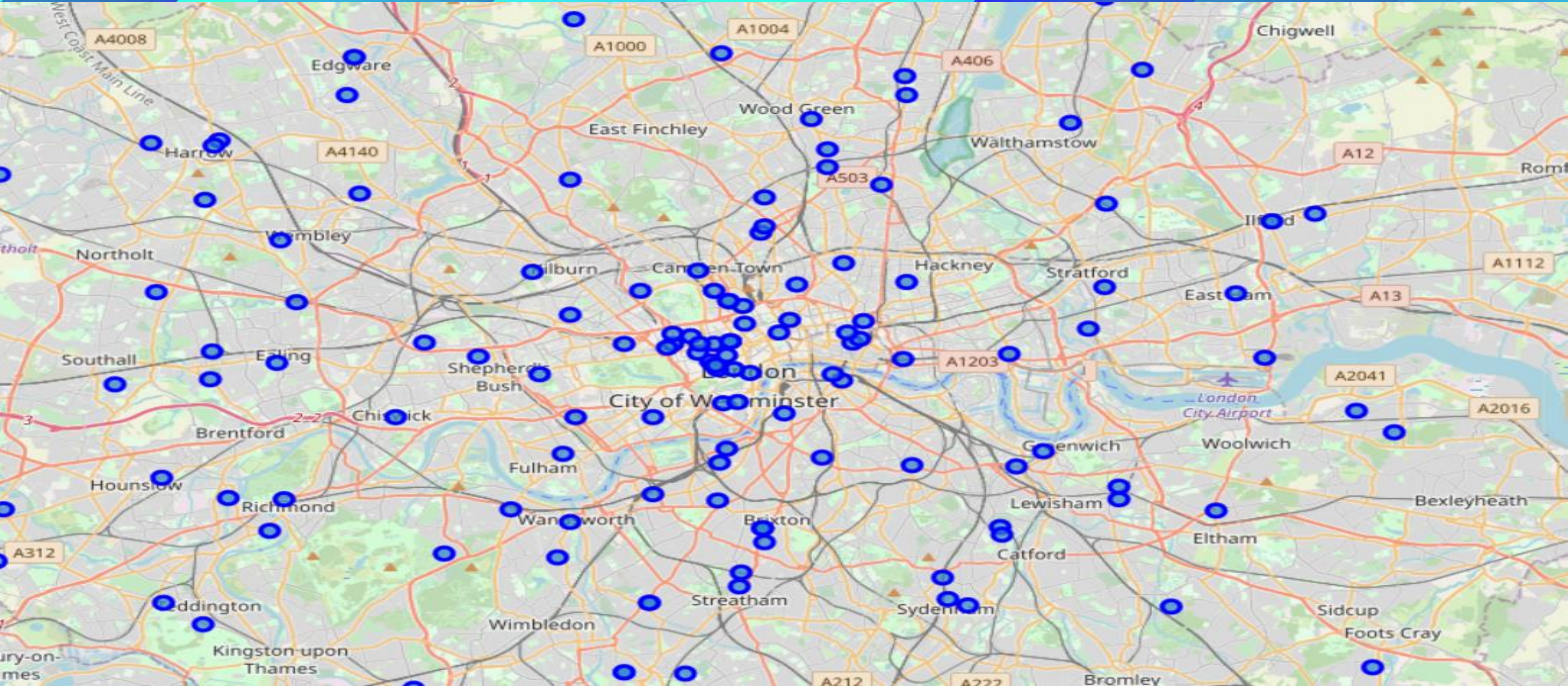
London Dataset

	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

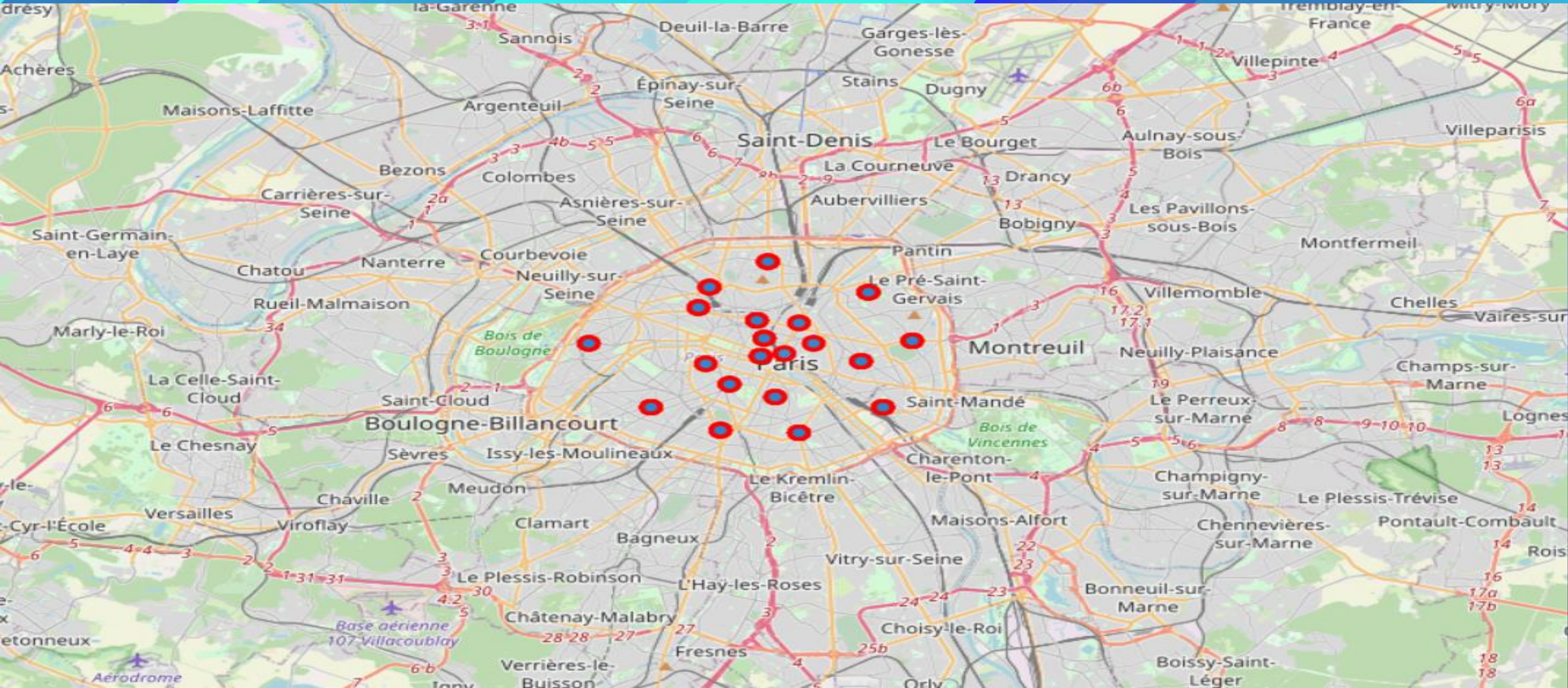
TO ACHIEVE THE HIGHEST LEVEL OF MODEL ACCURACY POSSIBLE, A DETAILED ANALYSIS OF THE DIFFERENT FEATURES AND METHODOLOGIES HAS BEEN CONDUCTED ALONG WITH AN IN-DEPTH STUDY OF THE DATASET. IN ORDER TO DISCOVER THE OPTIMUM CLUSTER FOR BOTH PARIS AND LONDON, THE AMOUNT OF FEATURES IN THE DATA FRAME WAS REDUCED BY REPLACING THEM WITH MORE USEFUL DATA. CORRELATION AND MANY OTHER VISUAL GRAPHS WERE THEN UTILIZED TO COMPARE THE TWO CITIES.

Map London



West Coast Main Line
A4008
A1000
A1004
A406
Chigwell
Wood Green
East Finchley
Walthamstow
A4140
A12
A503
Northolt
Hackney
Stratford
A1112
A13
Southall
A1203
A2041
London
City Airport
A2016
Brentford
Fulham
Woolwich
Lewisham
Bexleyheath
Eltham
A312
Catford
Streatham
Wimbledon
Sidcup
Foots Cray
Kingston upon
Thames
Bromley

Map Paris


Deuil-la-Barre
Sannois
Garges-lès-Gonesse
Stains
Dugny
Villepinte
Achères
Maisons-Laffitte
Argenteuil
Épinay-sur-Seine
Saint-Denis
Le Bourget
Aulnay-sous-Bois
Villeparisis
Bezons
Colombes
La Courneuve
Drancy
Carrières-sur-Seine
Asnières-sur-Seine
Aubervilliers
Les Pavillons-sous-Bois
Saint-Germain-en-Laye
Bobigny
Montfermeil
Courbevoie
Pantin
Chatou
Nanterre
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
Meudon
Le Kremlin-Bicêtre
Champigny-sur-Marne
Chaville
Le Plessis-Trévise
Versailles
Viroflay
Clamart
Maisons-Alfort
Chennevières-sur-Marne
Pontault-Combault
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

Venues

London

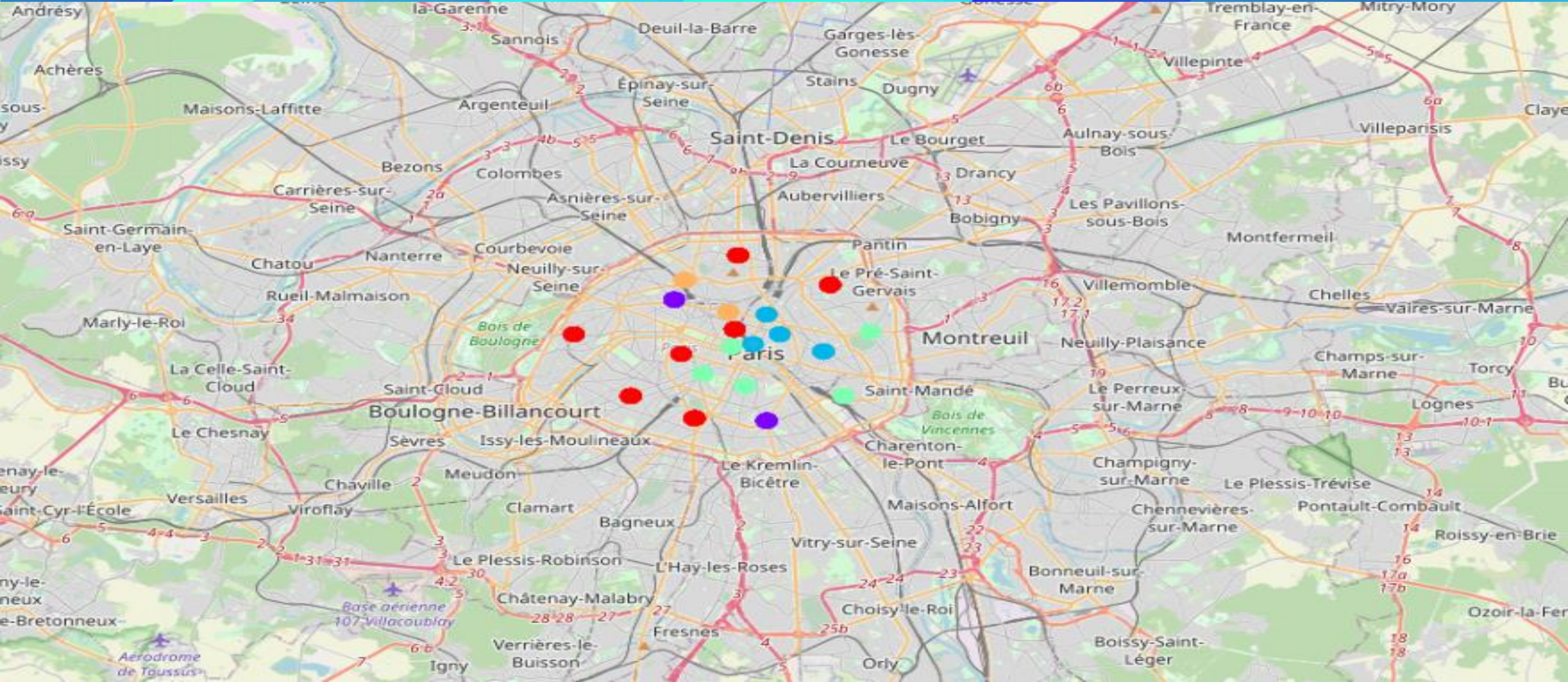
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Barnet	Turkish Restaurant	Italian Restaurant	Sushi Restaurant	Grocery Store	Indian Restaurant	Bakery	Deli / Bodega	Portuguese Restaurant	Coffee Shop	Gym / Fitness Center
1	Brent	Pub	Coffee Shop	Park	Platform	Indian Restaurant	Eastern European Restaurant	Supermarket	Food Truck	Japanese Restaurant	Deli / Bodega
2	Bromley	Pizza Place	Supermarket	Coffee Shop	Grocery Store	Pub	Stationery Store	Indian Restaurant	Fish & Chips Shop	Pharmacy	Café
3	Camden	Japanese Restaurant	Pizza Place	Coffee Shop	Beer Bar	Italian Restaurant	Tapas Restaurant	Malay Restaurant	Market	Hotel	Mexican Restaurant
4	City of London	Boxing Gym	Hotel	Burrito Place	Steakhouse	Department Store	Pizza Place	Indie Movie Theater	Event Space	French Restaurant	Botanical Garden

Venues

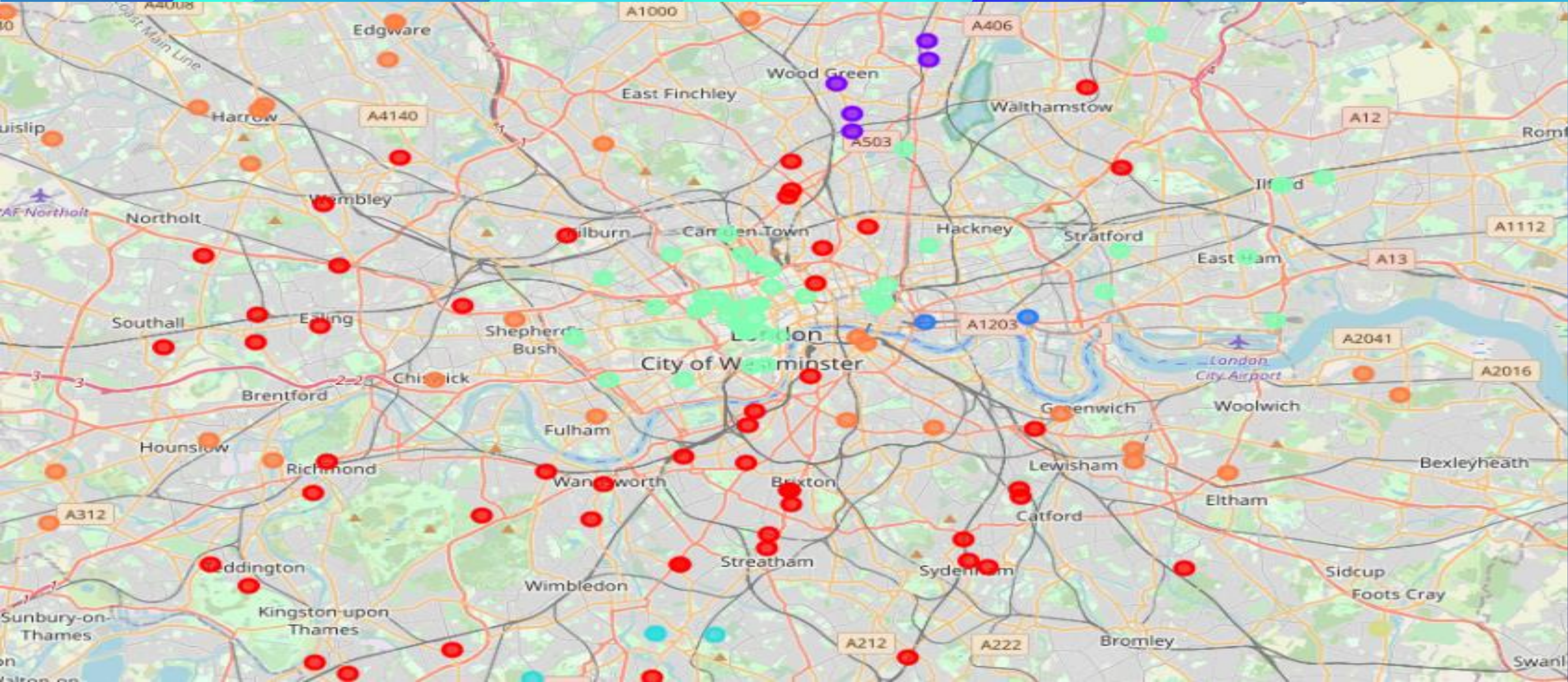
Paris

	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

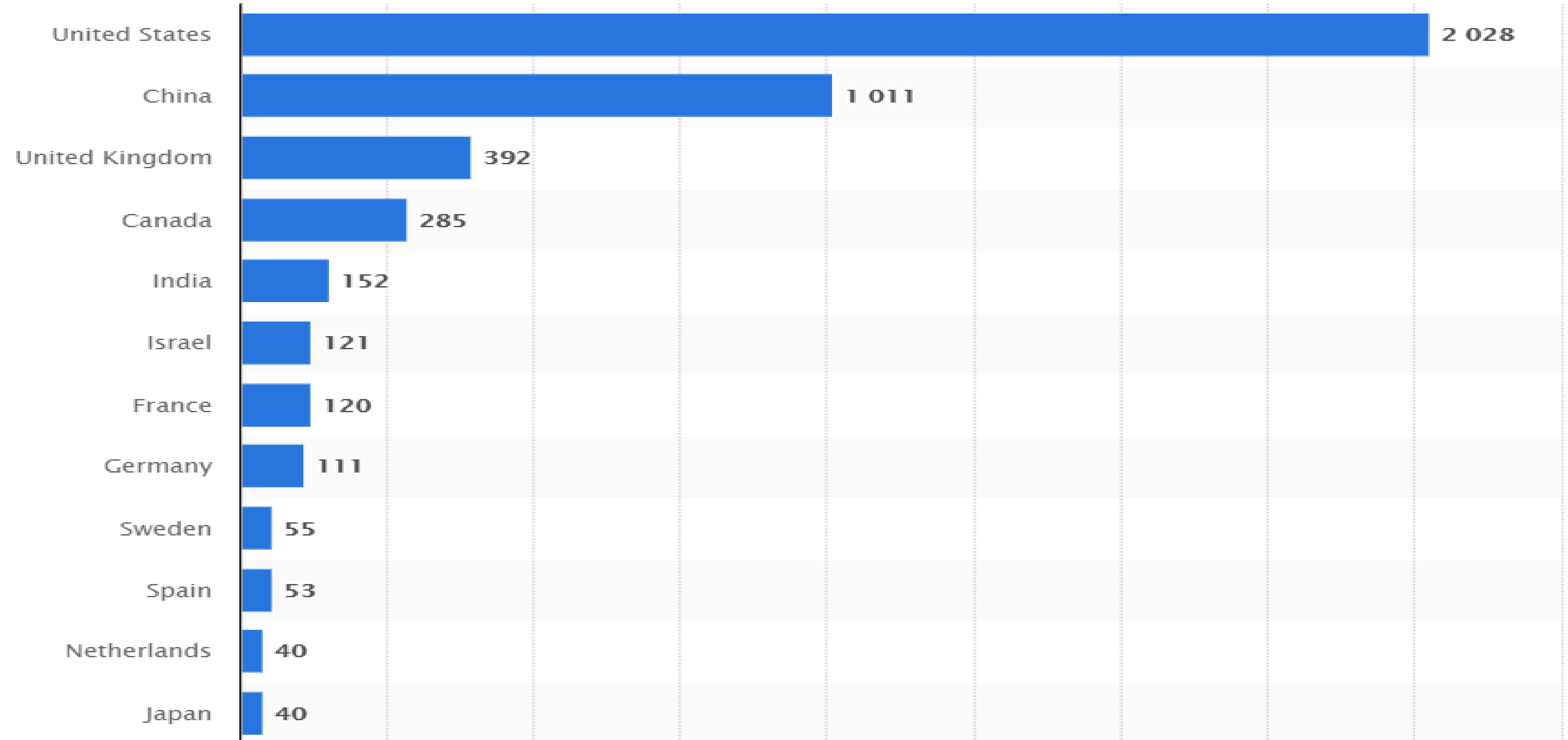
K Means Clustering Map - Paris



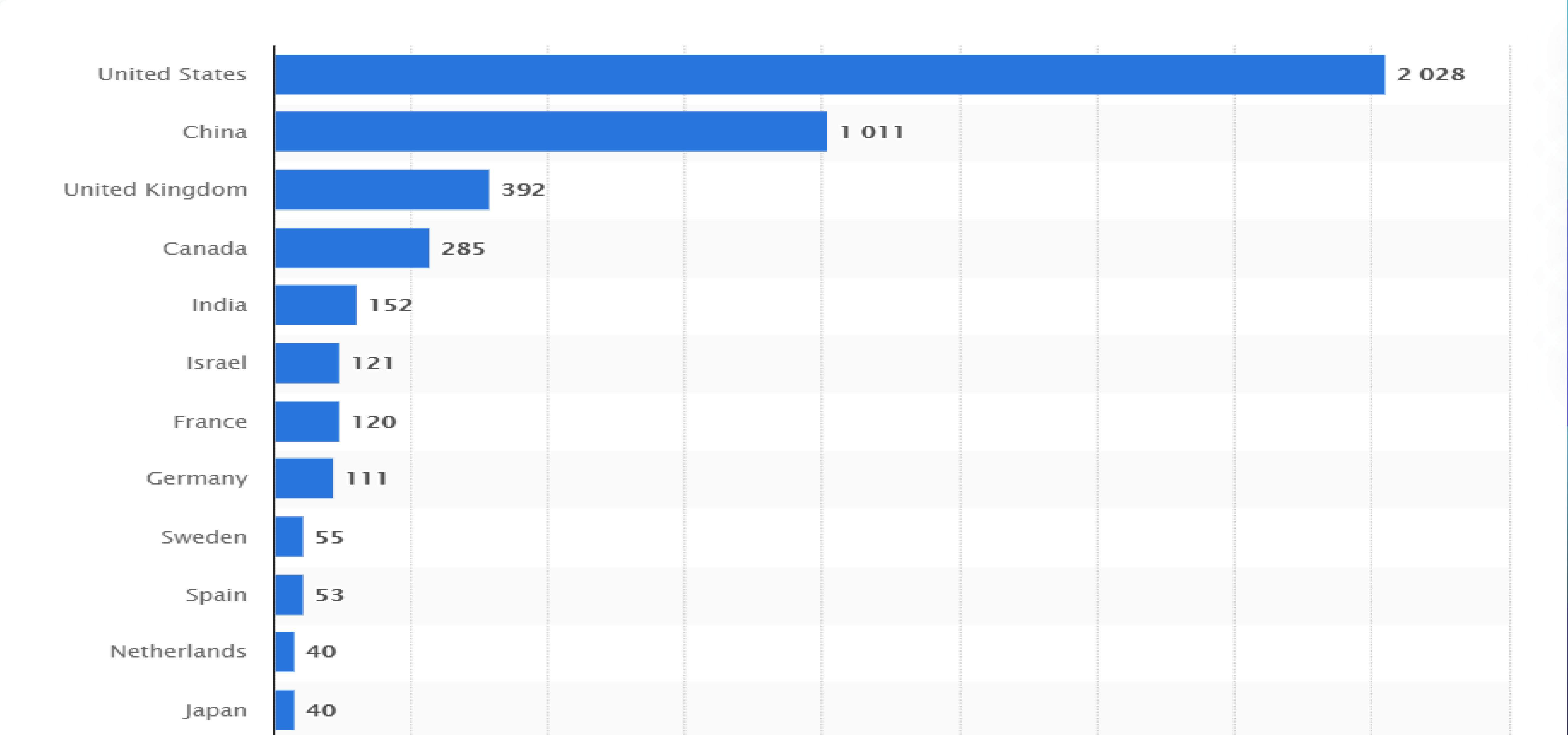
K Means Clustering Map - London



Artificial Intelligence



Artificial Intelligence



Results and Discussion

SIMILARITIES:
BOTH CITIES SHARE A RICH HISTORY AND ARE ETHNIC AND DIVERSE IN THEIR OWN UNIQUE WAYS. A RESTAURANT IS THE TOP MOST FREQUENTED VENUE IN THE MAJORITY OF NOTABLE NEIGHBORHOODS.

DIFFERENCES:
LOOKING AT THE MAPS, ONE CAN SEE THAT PARIS IS MORE COMPACT AND THAT WALKING AROUND ON ONE'S OWN IS MUCH MORE LIBERATING THAN USING A VEHICLE. PARIS DWARFS LONDON BY A FACTOR OF 4:1 IN TERMS OF POPULATION DENSITY.

Results and Discussion

ARTIFICIAL INTELLIGENCE

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**

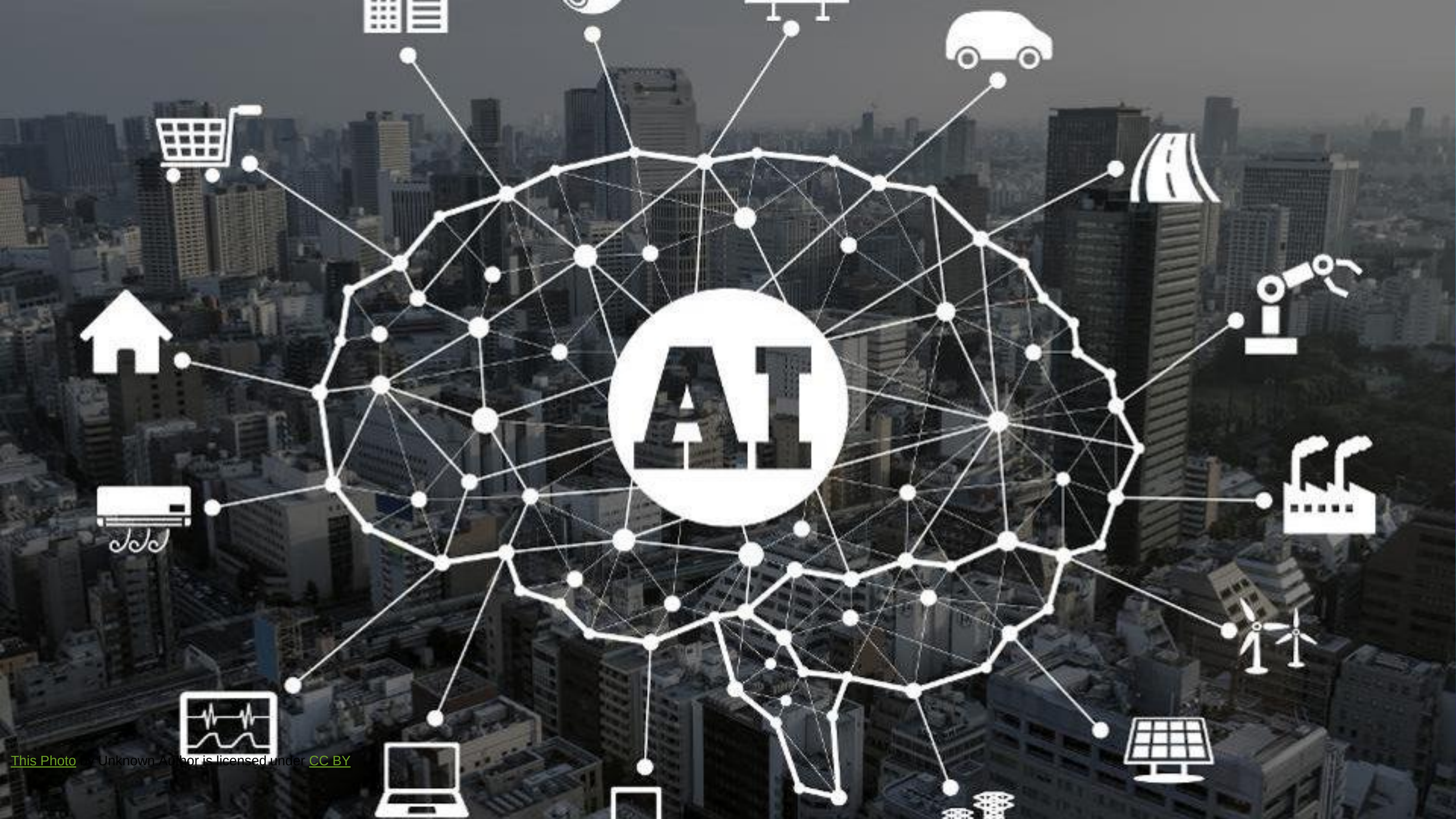
Results and Discussion

ARTIFICIAL INTELLIGENCE

BECAUSE THE DATASET FOR THE ARTIFICIAL INTELLIGENCE WASN'T EASILY AVAILABLE, IT HAD TO BE CULLED FROM VARIOUS SOURCES, WHICH FREQUENTLY RESULTS IN INACCURACIES AS WELL AS INCONSISTENCIES. BECAUSE OF THE DISTRICTS' EXCESSIVELY COMPLICATED LAYOUT, IF THE VENUES ARE TOO CLOSE TO ONE ANOTHER, OUR ANALYSIS WILL BE FLAWED. EACH TIME AN API REQUEST IS MADE, A DISTINCT SET OF RESULTS IS RETURNED USING THE DATA. TO ACHIEVE THE INTENDED OUTCOME, NUMEROUS TRIALS AND ERROR RUNS ARE NECESSARY.

Conclusion

RECENTLY, MORE PEOPLE HAVE BEGUN INVESTING IN ARTIFICIAL INTELLIGENCE, AND BUSINESSES HAVE BEGUN AUTOMATING THEIR PROCEDURES. FOR THOSE LOOKING TO INVEST IN ARTIFICIAL INTELLIGENCE OR EVEN LAUNCH A BUSINESS, BOTH CITIES PROVIDE A WIDE RANGE OF CHANCES, AS NUMEROUS VARIABLES WERE DEMONSTRATED. FINALLY, NEW TECHNIQUES AND STRONGER MACHINE LEARNING ALGORITHMS LIKE KD TREE, WHICH HAS A FASTER RUN TIME ALGORITHM, CAN BE USED TO CREATE A BETTER MODEL. HOWEVER, CLUSTERING DID ENABLE US TO EMPHASIZE THE BEST VENUES AND NEIGHBORHOODS. FINALLY, SINCE CORRELATION DOES NOT INDICATE CAUSALITY, ANY CONCLUSION DRAWN FROM THIS DATA MAY VARY IN LIGHT OF DIFFERENT OTHER TRENDS, VIEWPOINTS, AND DATASETS.


AI