Clustering Technology Hubs Around the World

Veronica Rodarte

July 2, 2020

1. Introduction

1.1 Background

Technology is one of the fastest growing industries and there is no stopping it. The world is becoming more dependent on technology and there is a start-up around every corner looking to capitalize on new emerging technologies. There are many hubs for technology around the world, but few know about the up and coming hubs that are leading the way in their respective countries. These various hubs have a lot to offer the world and are gaining speed in their growth and reach. Many investors are eager to invest in these countries because of the untapped talent. This report focuses on a wide range of cities across the world.

1.2 Problem

This project seeks to find the up and coming technology hubs that many entrepreneurs do not know about and or may not be aware of the potential these markets hold for their future business. There are many hubs for different industries/technologies within the tech field that have emerging opportunities that can be beneficial to start-ups that may not have the capital to compete with the big tech giants. Also, many of these tech hubs are looking to attract entrepreneurs and talent to help grow industries within the respective country.

1.3 Stakeholders

It would be worthwhile to have this information for entrepreneurs and enthusiasts that seek cities and hubs that are up and coming and exciting in the fourth technological revolution. As well for other businesses that seek to follow the growth of these cities and centers.

2. Data Acquisition

2.1 Data Sources

The data for the most popular technology hubs around the world was extracted from multiple sources: Wikipedia Technology Hubs Worldwide for the Latitude and Longitude of each hub (City/General Location), Startup Genome for cities with high start-up rates divided into sectors, and Business Insider for countries with a booming tech scene. For the nearby venues to all the technology hubs and for each hub, Foursquare API was used for venue category and most common. The population density and other demographic information of each respective district or city was extracted from the UN data. For the property price-to-income ratio for each

respective city, the website Numbeo was used. Lastly, to map out the world the Coursera world GeoJSON file was downloaded.

3. Methodology

3.1 Data Cleaning & Feature Selection

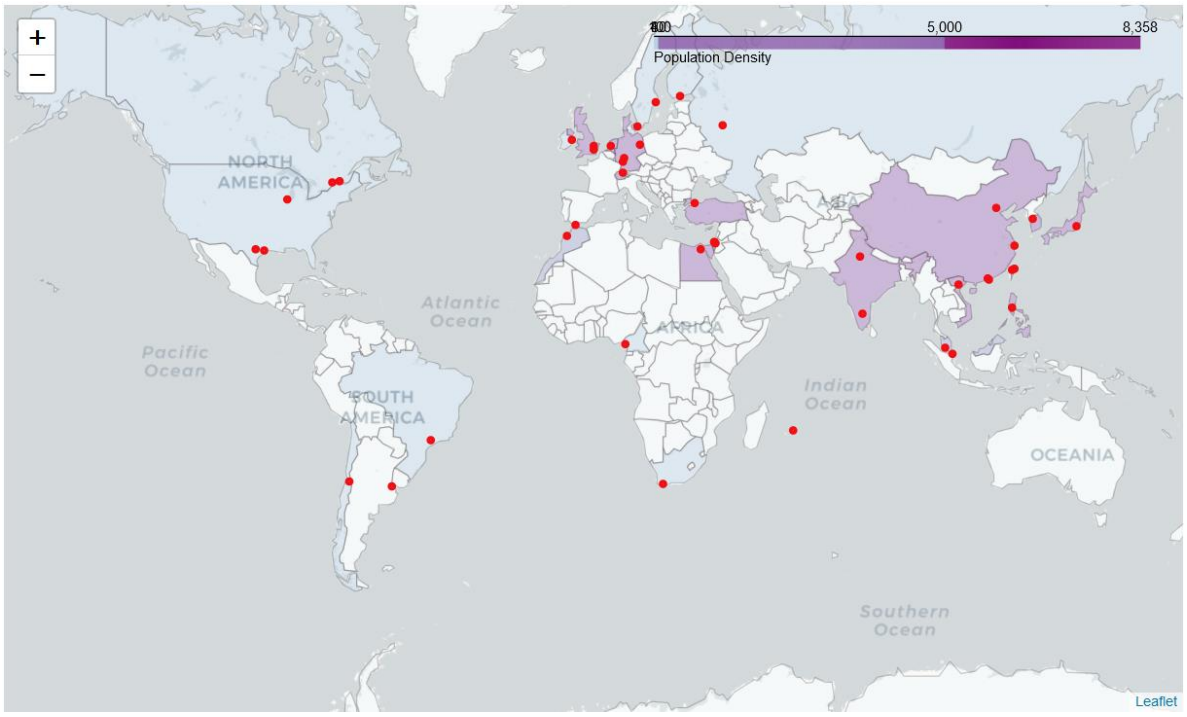
The data downloaded was created beforehand, with the features being pre-selected from the various data sources used and therefore there was no need for extensive data cleaning. All the data that was needed for the analysis was already present therefore I only needed to divide the technology hubs by region for further analysis.

A second set of data was downloaded, this data contained the cities and the sector that they are becoming dominant in. The data was divided by sector and the subsequent tables were cities focused in on Blockchain, Advanced manufacturing & Robotics, and AI, Big Data & Analytics. These tables were used to visualize their geographic location in a folium map.

Using the first table that was created, each technology hub was divided by their respective region to better facilitate k-means clustering and to have a better visualization of the regions and their clusters. The regions were: East Asia, Southeast Asia, Middle East, Europe, North America, South America, and Africa. These regions were the best fit for the general locations of the technology hubs. Theses clusters were visualized and examined to determine similar common venues amongst all the regions and their neighboring cities. For each regional cluster, the optimal k was found using Elbow method.

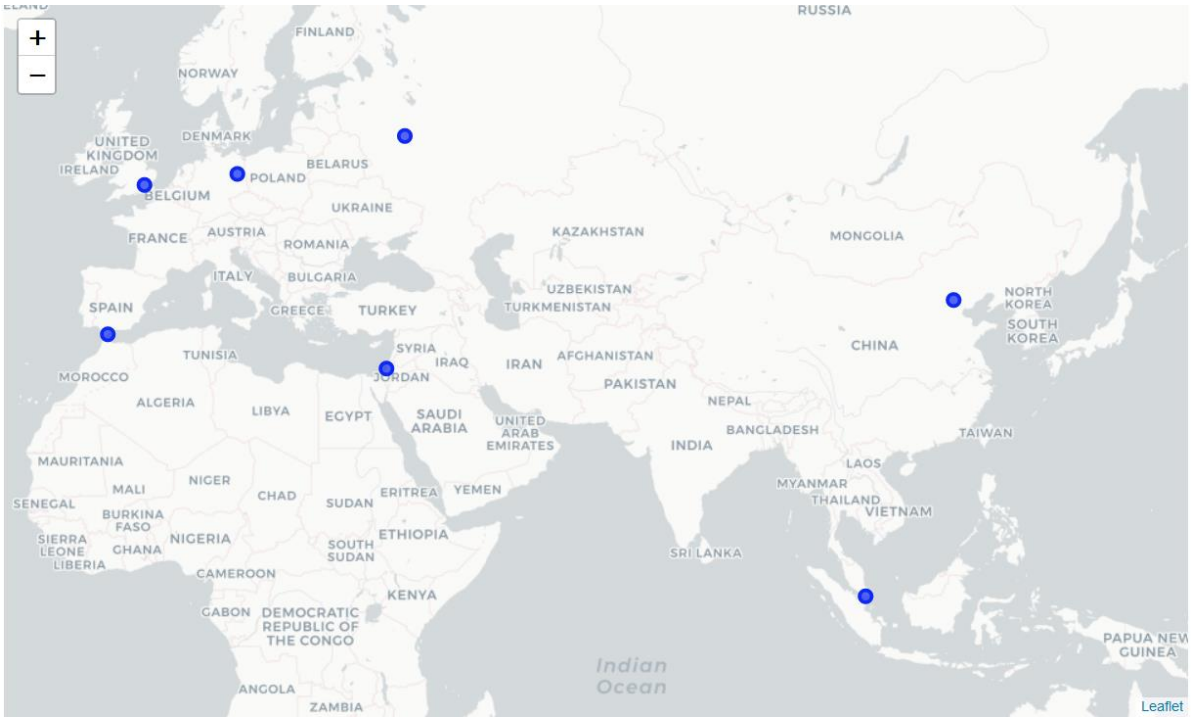
The first visualization was that of all features present in the first table: Technology Hub, City, Country, Population density, Price-to-income ratio. The purpose of the visualization was to show the population density (persons per km) and the affordability of residential areas around these technology hubs. Population density helps to gage the market size for a potential entrepreneur and the affordability helps to determine the approximately the overall living costs that may be encountered. Some cities are more affordable than others; Singapore, Hong Kong, Beijing, and Shenzhen are among some of the most expensive cities that are presented. Some countries did not have price-to-income values because of a lack of available information, but I decided to keep them since they are reputable technology hubs within their respective countries (Ebene, Mauritius & Buea, Cameroon).

Population Density



The next three visualizations are the maps that represent the cities that are gaining traction in the technology sectors chosen for this report. The map does not present a specific technology hub but rather the city because these hubs house a vast number of companies.

Blockchain

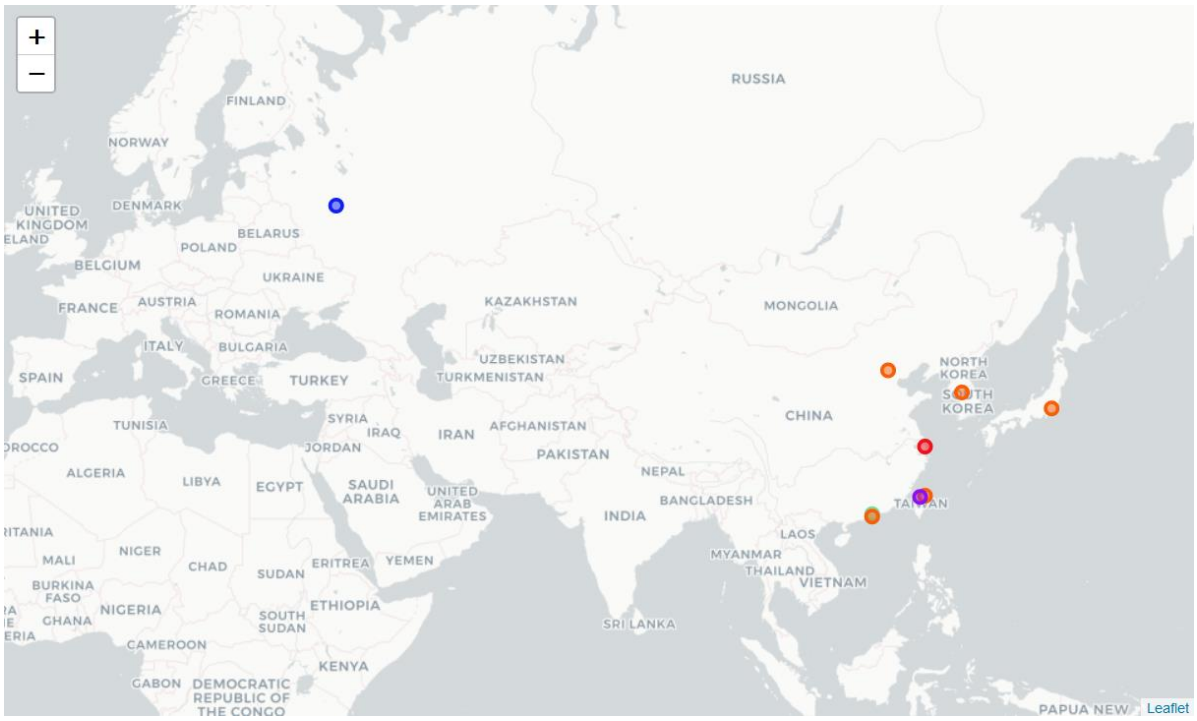


Advanced manufacturing & Robotics

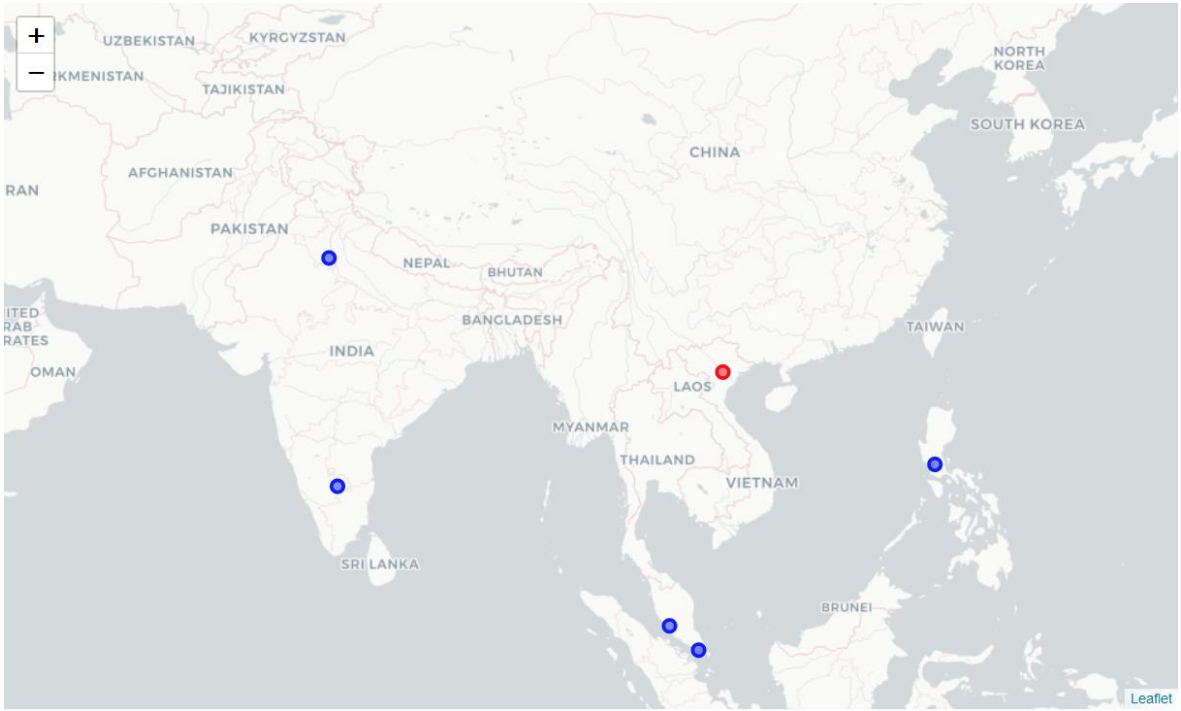
AI, Big Data & Analytics

The clusters for each region were mapped and then the clusters were examined for similarities amongst their neighbors and other regional clusters.

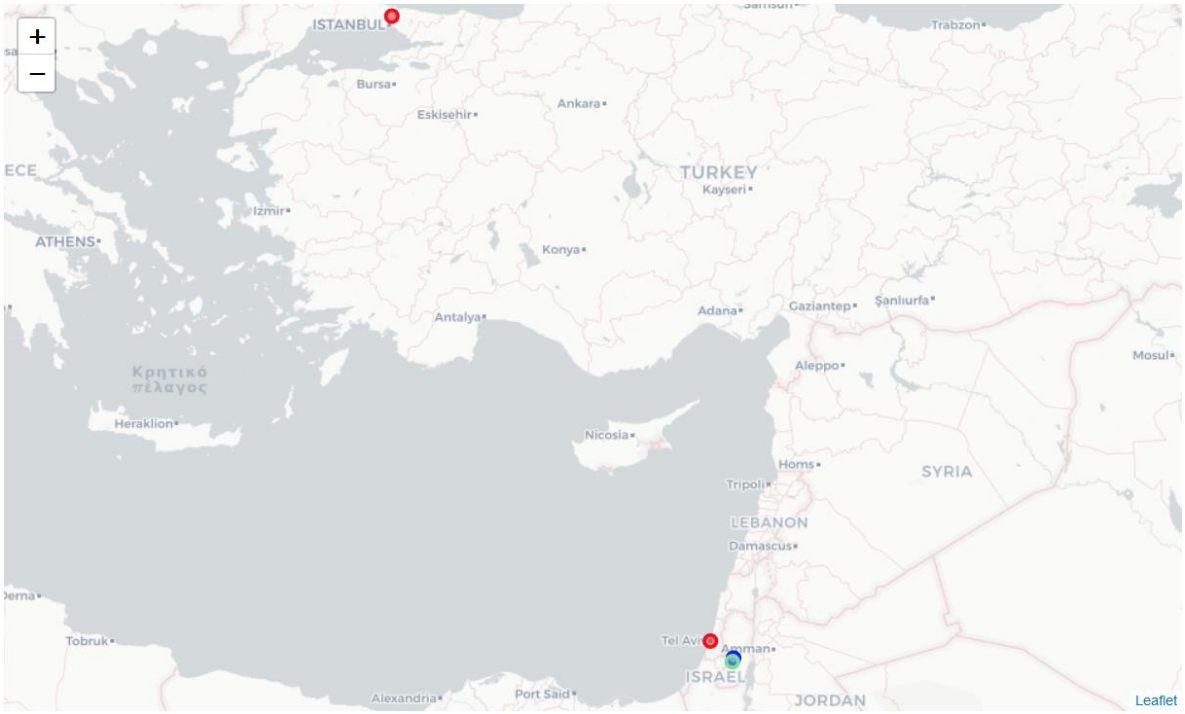
East Asia Clusters



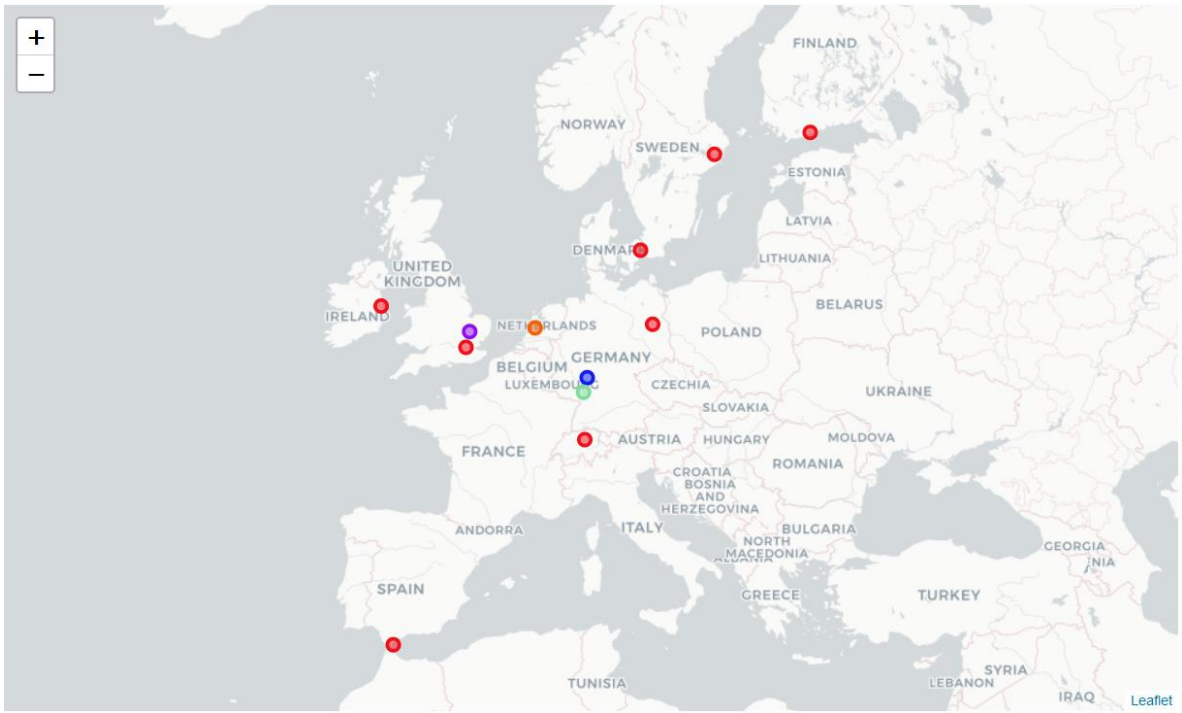
Southeast Asia Clusters



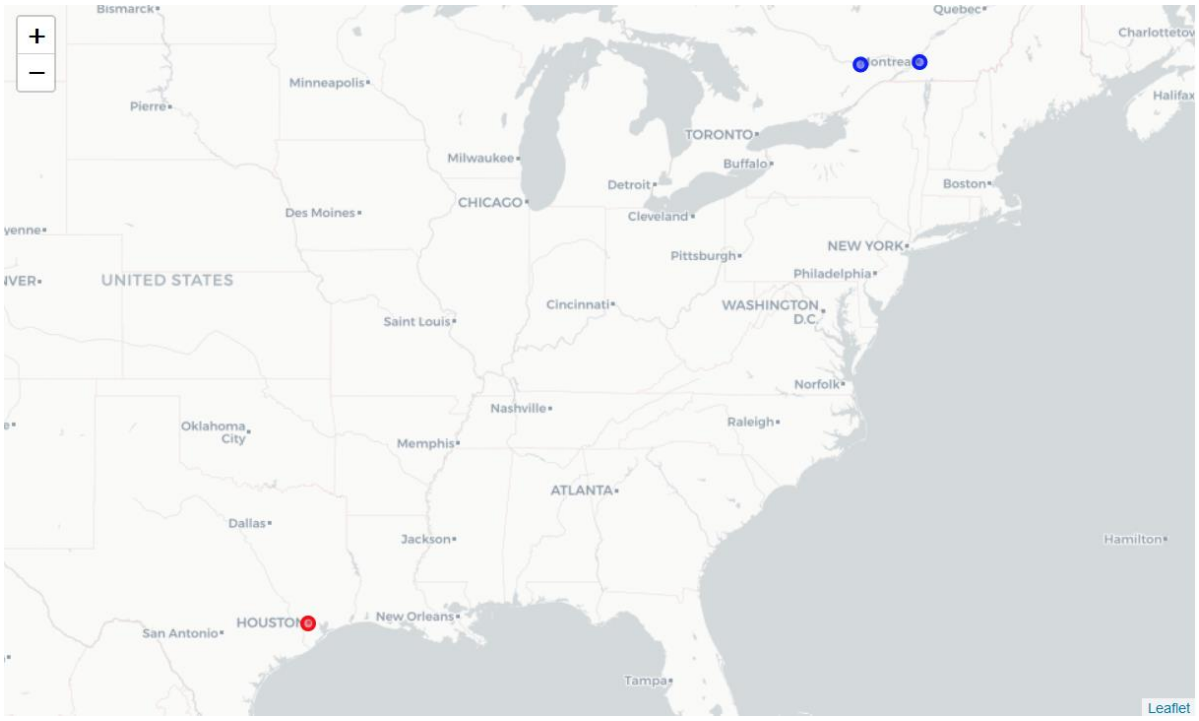
Middle East Clusters



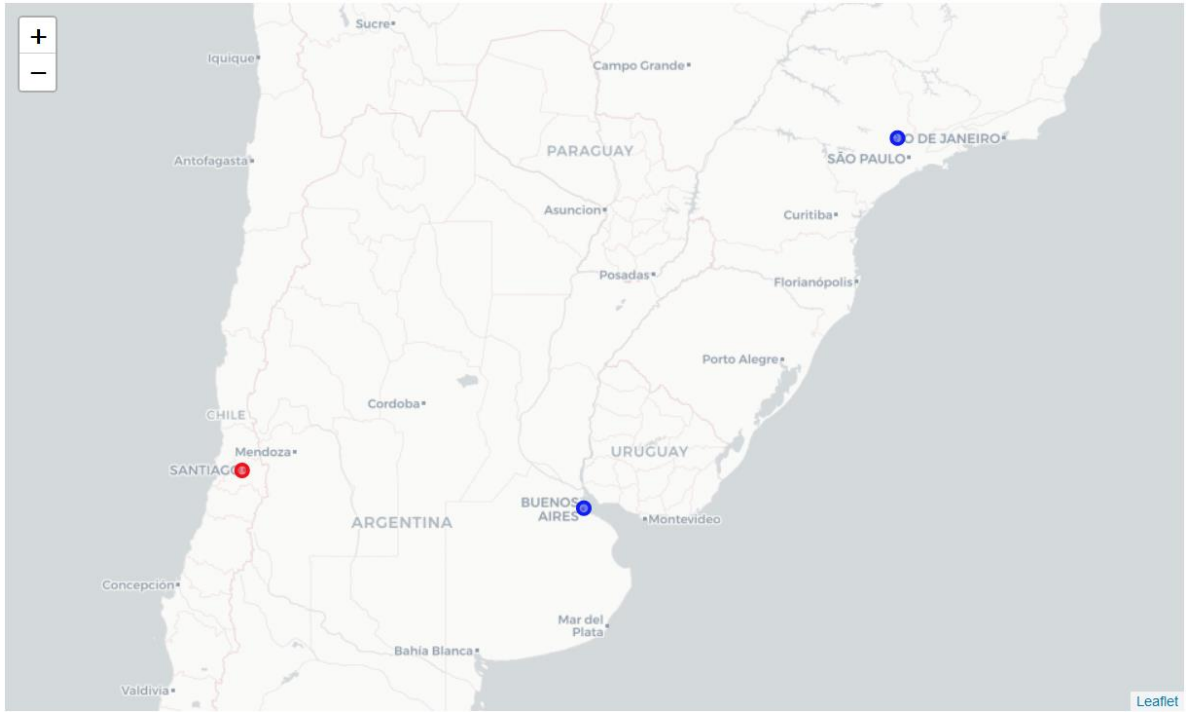
Europe Clusters



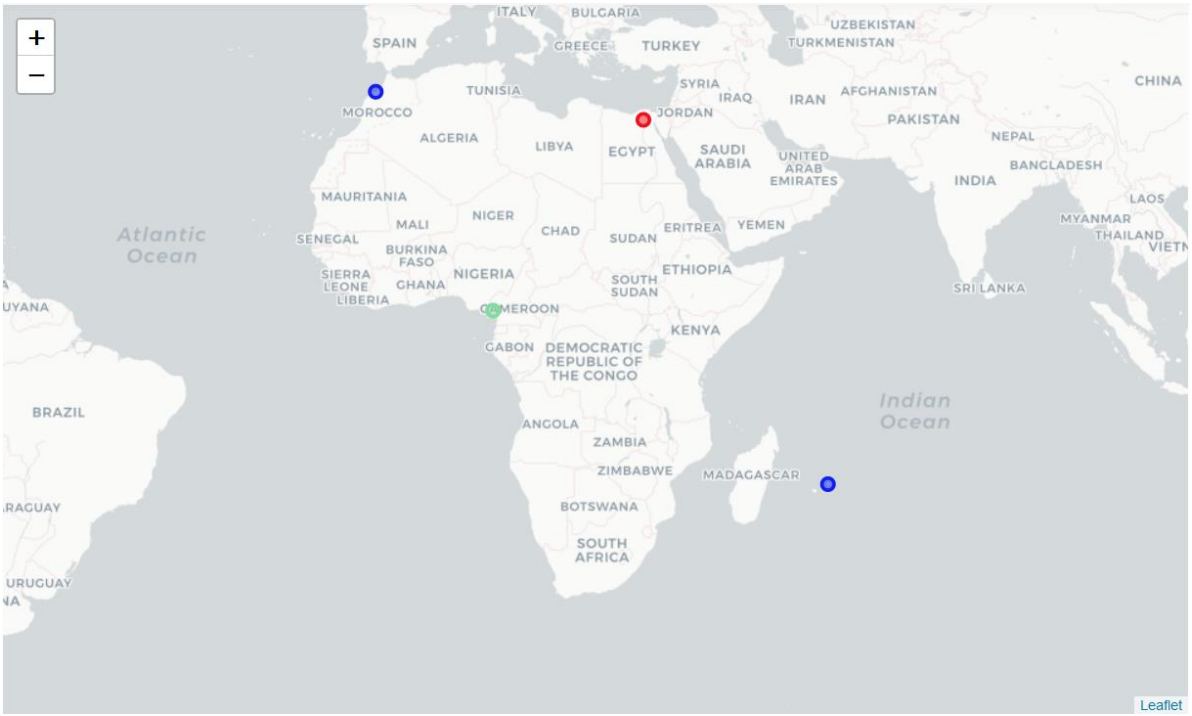
North America Clusters



South America Clusters



Africa Clusters



4. Results

4.1 Examine Clusters

Upon examination of all the clusters, I chose to focus in on the largest clusters, clusters the metropolitan areas of the cities. The most common venues were all directly related to food. The table (Table 1) that follows presents the percentage of food related venues for all clusters grouped by region. The average of food related venues overall is 60%. For the major metropolitan clusters, the average of food related venues overall is 69%. Another common venue similar across regional clusters is coffee shops/cafes. The similar restaurants within the food related venues range from: Italian, Chinese, Japanese, Middle Eastern, and Mexican.

Table 1: Percentage of Food Related Venues for each Cluster

Region	Clusters	% Food Related Venues
East Asia	Cluster zero	60%
	Cluster one	50%
	Cluster two	70%
	Cluster three	74%
	Cluster four	50%
Southeast Asia	Cluster zero	70%
	Cluster one	80%
Middle East	Cluster zero	40%
	Cluster one	85%

	Cluster two	50%
Europe	Cluster zero	40%
	Cluster one	56%
	Cluster two	70%
	Cluster three	50%
	Cluster four	50%
North America	Cluster zero	75%
	Cluster one	80%
South America	Cluster zero	55%
	Cluster one	50%
Africa	Cluster zero	65%
	Cluster one	40%
	Cluster two	50%

Other venues that were similar for the regional clusters were related to recreational/attraction venues. From the analysis, many of these hubs are very similar with respect to the common venues that are present and found around their general location. It would seem that many mimic each other in their surroundings or in other words, the general environments that grow around these hubs tend to resemble each other.

5. Discussion

The analysis of the clusters indicates that not only are these technology hubs hotspots for tech companies but as well for food businesses. The growth of these industries and cities attracts newcomers and it would seem in turn that they attract food businesses. I make this observation solely based on the data that I gathered after clustering and analysing the most common venues amongst all these regions. It was peculiar to find foreign restaurants in some of these technology hubs, such as a Mexican restaurant in the East Asian region. This goes to show the vast reaches of globalism as well as the opportunities for other cuisines to appear soon while these markets are still young.

6. Conclusion

From the analysis of the clusters in each region the most common venues near and around all technology hubs presented were restaurants and overall venues related to food. This is good insight into the venues that would have most success in and around technology hubs, from this primarily analysis it could be concluded that with larger technology hubs and growing cities focused on technology, food is an important and recurring venue across the globe. There are many options for upcoming entrepreneurs not only in the technology sector but also in other sectors that work with or are affected by the sector, in this case the food industry.

This analysis goes to show that there are many technology hubs that are growing and will continue to grow as the industry grows. They provide various opportunities not only to locals but to foreigners. The industries are young and suitable for many that may not have extensive capital to compete with larger firms.

I hope this report helps to shine the light on a potential headquarter for your next company!

7. Data Sources

6.1 Location Data

https://www.bloomberg.com/news/articles/2018-04-19/global-cities-challenge-america-s-high-tech-dominance

https://www.weforum.org/agenda/2017/04/these-are-the-22-best-cities-in-the-world-for-tech/

https://en.wikipedia.org/wiki/List_of_technology_centers

6.2 Property price-to-income Data

https://www.numbeo.com/property-investment/

6.3 Population density

https://worldpopulationreview.com/country-rankings/countries-by-density

6.4 GeoJSON file

Coursera Labs

6.5 Foursquare Data

https://foursquare.com/developers/apps