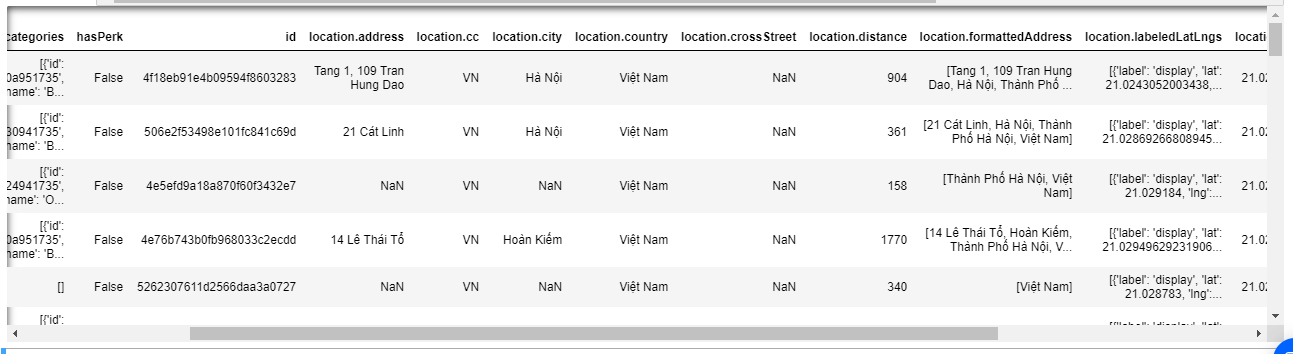
|  |  |  |
| --- | --- | --- |
| DS Objective Apply data science technical to support manage all transaction point: Branches, ATM and POS Merchant DATA Foursquare API  Transaction Network of bank  Folium data to visualization | |  | | --- | | LE CONG BINHBANKING TRANSACTION TEMINAL NETWORK OPTIMIZE |  introduction Ha Noi, the capital of Vietnam, where over 8 million people live and it has a population density of 2.398 people per square kilometer. The city have 30 district but all transaction only do at five centers district. So, each of bank need control transaction network. That include, branch, ATM and POS Merchant. There are more 40 bank in Ha Noi; customer of any bank can do transaction in every bank. Therefore, We need a system to easy control transaction network.  I am working in a Bank. We have data about all transaction point of my Bank. However, I do not have information of their bank. Other side, I need more information about each of place to decide which place should be setup new agent, which place need to remap for control network. I chose Foursquare data and unsupervised cluster to solve that problem. Methodologybusiness understand. We focus about collect all information about banking network to support deciding rebuilt transaction network data procecing As I show in notebook, There are many technical to preprocessing data. This is step, which need more time to complete. After processing data I use K-Mean to cluster and use that information merge with my branch network to decide next action. |

I use google service to define current coordinates. (Latitude = 21.027763

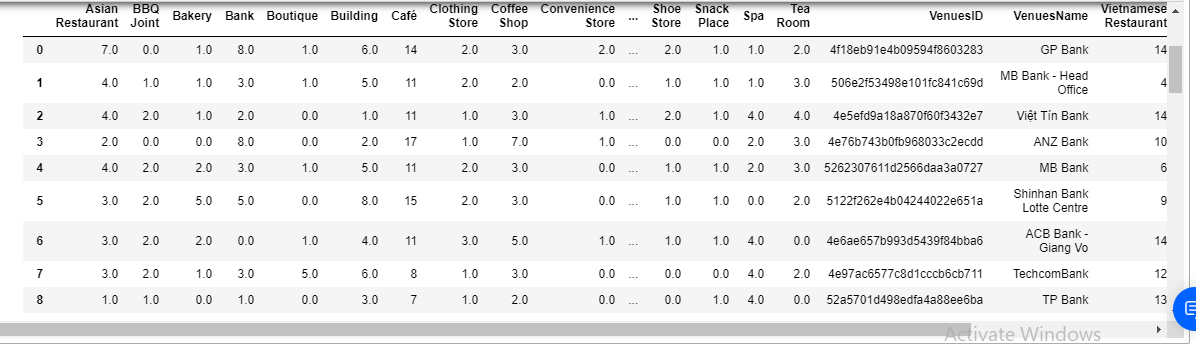
Longitude = 105.834160)

After that used Regular Call with search query to get json content.

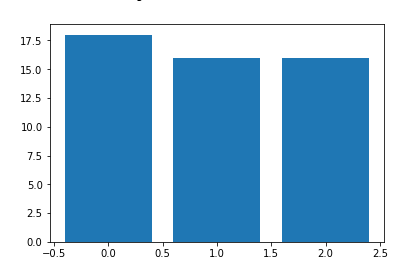


Although Foursquare return JSON format, but it only raw data sow I need use python script to process data to my work. I am attack full source code in Github to verify.

Each of Bank endpoint I use loop process to get all venues nearby bank location. Foursquare returns many categories. However, it has not much venues in Ha Noi. Therefore, I need sum technical to reduce missing feature and use total venues by category to cluster.

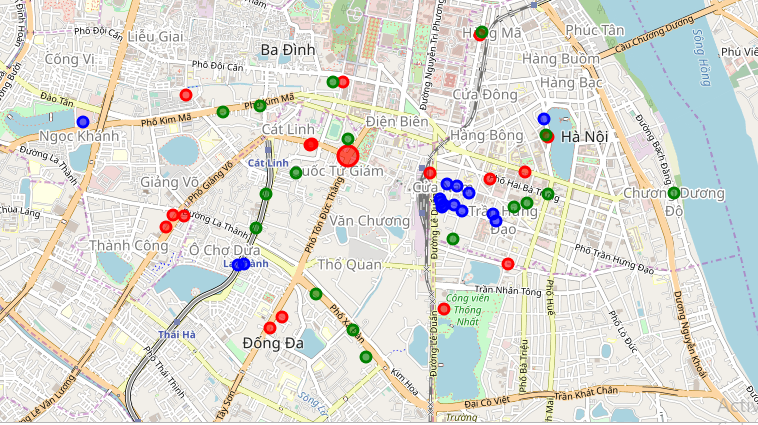


I will use this information to cluster each of place by K-Mean algorithm. Because of missing data in Vietnam from Foursquare, I only have 3-cluster segment. If I have enough data I can elbow method it ensured me the degree for optimum k of the K-Means. Before run K-Mean cluster, I need clean all input by transform, mapping and standard by Sklearn StandardScaler method



This is chart number of bank venue by cluster auto by K-Mean.

Now, I visual bank network (with cluster label) in map



From map, I can see all transaction place of bank locate at center. Blue terminal is place with many office and travel place. K-Mean algorithm some place it far away center. That is a signal to review network transaction at that place.

Base on list of transaction place auto cluster by K-Mean, I need merge with my data in the bank for adjust. Example: position of transaction place, marketing by place customer segment, characteristics of each cluster to support deciding transaction number, transaction position of bank in Ha Noi

### Conclusions

## In this project, I only demonstration my knowledge, that I take after nine courses of IBM Data Science Professional Certificate. It include methodology of data science, python technical to process data, visualization data, using model to support business, the machine learning technical and many tool support for data science: IBM Cloud, Foursquare, many library. It may be do not have enough to apply because missing of data, missing of business requirement, and not enough to research.

## Thank Coursera and Instructor to help me complete this course. That is giving me very interesting time and much love about data science.

## The best interesting thing of me in this course: Data Science is the sexiest job in 21st century.

## **òScience Professional Certificate**