Top places-to-go in Taiwan

Introduction

'Isla Formosa' – Taiwan, is a beautiful island that is worth to discover. There are so many attractions to explore, from incredible natural scenery, lively traditions and cultures to delicious food. Tourists can have variety options to choose what places to visit based on their interest. Given there are numerous things tourists can do, my objective for this project is to cluster the venues in Taipei city into different segments which can help tourists to identify where to go based on their interests.

Data

The data I am going to use are Foursquare data for Taipei city. I am going to use Foursquare API to pull the venues data in a json file then clean-up the data into pandas data frame as shown below:

	name	categories	lat	lng
0	白石58莊園	Farm	25.106392	121.594388
1	Drift Wood Cafe Da	Café	25.104982	121.591988
2	龍船岩(石船）	Mountain	25.100527	121.597000

I also going to use Taipei city district geo location data for the latitude and longitude so we can merge with foursquare venues data for analysis. Below is an example of the geo location data:

	District	City	District (Chinese)	Latitude	Longitude
0	Zhongzheng	Taipei City	中正區	25.032400	121.518000
1	Datong	Taipei City	大同區	25.065986	121.515514
2	Zhongshan	Taipei City	中山區	25.064361	121.533468
3	Songshan	Taipei City	松山區	25.049698	121.577206
4	Daan	Taipei City	大安區	25.026389	121.534444

Then I merged two datasets into one data frame as shown below. This will be my analysis dataset to do the clustering for finding similar places for each district.

	District	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Zhongzheng	Kinfen Braised Pork Rice (金峰魯肉飯)	25.032194	121.518534	Taiwanese Restaurant
1	Zhongzheng	虎記商行	25.031744	121.519284	Café
2	Zhongzheng	樂田麵包屋 Gakuden Boulangerie	25.032757	121.517534	Bakery
3	Zhongzheng	National Theater (國家戲劇院)	25.035197	121.518188	Theater
4	Zhongzheng	豆味行甜不辣、豆花、芋圓	25.031303	121.517232	Snack Place