Battle of the Midsize University Towns in US

Introduction

A university town or college town is a community that is dominated by its university population. In US, it is often a separate town or city, but in some cases it can also be a city neighborhood or district. University towns are usually considered to be some of the best places to live in because of their low cost of living, rich cultural activities, educational opportunities, fun and sports, etc. With the stable consumption power coming from students, faculties, staff, and retirees, university towns are also favored by numerous investors. Because university towns usually have students forming a major fraction of their population, businesses that typically cater to students are in huge demand in such places.

In this project, I will explore the midsize university towns in US to study their similarities and dissimilarities. Here by "midsize", I meant to exclude the big cities with large populations. The reason is that cultures and business/life styles in large cities are influenced by many more factors compared with "university towns" where the university/institution plays important roles throughout the community.

This study will be helpful to both investors (for example those who want to open restaurants or bookstores in one of the university towns) and students who are going to apply colleges in the near future to make their decision.

Data

The data used in this study include three main parts. First, we need a list of university towns in the United States. Second, we need some basic information about those towns, such as population, location (latitude, longitude coordinates), etc. Once we have a database of the university towns and their location data, we can use Foursquare API to explore the nearby venues. By analyzing the three data sources, it is possible to find the relation between the common venue types and the town locations.

After an internet search, I found a list of university towns around the world in Wikipedia [1]. The US part can be scrapped using the BeautifulSoup package [2]. As for the population (scrapped from [3]) and location data (downloaded from [4]), online databases are used. One note on the location data, I first tried to use geopy package [5] to extract their location coordinates using name address, but it took quite a few tries to get a complete list (due to limitations of their service) and the return results are not always accurate. In turn, I used the online resources with all the US cities and their locations tabulated.

References

1. University town list from Wikipeida, https://en.wikipedia.org/wiki/List_of_college_towns#College_towns_in_the_United_States
2. https://www.crummy.com/software/BeautifulSoup/
3. 2017 Population data for US cities, https://www.biggestuscities.com/
4. Location data for US cities, https://simplemaps.com/data/us-cities
5. https://geopy.readthedocs.io/en/stable/