Evaluating Neighborhood in Manhattan

Capstone Project As Part of “Applied Data Science Capstone” course

By:

Saleh Ardameh

August 2020

Contents

1. Introduction

1.1. Significance of the study

As issues (e.g. traffic congestion and environmental degradation) emerging from urbanization and urban sprawl have attracted considerable attention, there is a growing consensus that integrating land use and transport is a vital pathway to urban sustainable development (Bertolini et al. 2005; Kenworthy and Laube 1996; Lautso et al. 2004; Weiss et al. 2018). The physical components of urban spatial structure, or more specially land use, are connected by transport systems, which facilitate the functioning of cities by passengers and freight movements. The implementation of both urban functions and transport construction need substantial land requirement and consumption. Since urban land is a scarce resource and must be used efficiently (Lambin and Meyfroidt 2011), urban development, both in land use and transport, should be planned and managed rather than develop spontaneously (Tian et al. 2017). The integration of land use and transport thereby is necessitated in urban planning and management. The definition and representation of land use systems and transport systems vary in research purpose and scale. Land use/land cover pattern can reflect the spatial characteristics of land use practices and location-specific transport capacity can be measured by accessibility, which refers to the ease with which anyplace of a certain area can be reached by individuals at a particular location using the mobility service of specific transport systems. The relationship between land use/land cover types and accessibility can provide new insights into the synergy of land use and transport planning at regional scale.[1]

The relationship between land use and transport is complex and dynamic with various interactive effects (Litman and Steele 2017), which results in the breadth of land use transport literature. There is a long tradition in research of integrated land use and transport modelling and land use transport interaction (LUTI) model has come to significant fruition (Acheampong and Silva 2015). However, on account of a series of challenges such as model transparency, operability, parameter validity, computational performance, LUTI model has not progressed in practice as far as it is expected (Waddell 2011). With the enhancement of quality and availability in land use data, a trend towards empirical studies on this relationship has arisen. These researches can generally be divided into: pedestrian bike infrastructure and land use (Cervero and Duncan 2003; Duncan et al. 2010; Lee and Moudon 2006; Rodríguez et al. 2009; Wang et al. 2015; Witten et al. 2011); motorized infrastructure and land use (Duranton and Turner 2012; Giuliano et al. 2012; Horner and Schleith 2012; Stanilov 2003; Vandenbulcke et al. 2009). Rodríguez et al. (2009) pointed out that higher levels of walking accessibility can intensify population density, availability of retail and land use mix. Stanilov (2003) proposed integral accessibility, which assesses the ease of access to high-class roads, to explore the impact of transport networks on land use pattern. The result suggests that the priorities of accessible locations allocation are: commercial, industrial, multi-family, medium-density single-family and low-density single-family uses. In addition to the above mentioned two categories, urban rail system (tram and metro) is taken into account in studies at metropolitan scale (King 2011; Ratner and Goetz 2013). Ratner and Goetz (2013) found that the extended light rail transit increased residential density in station areas of Denver. All of these attempts confirm the fact that transport is inextricably related to land use and development. However, the transport impacts vary in mobility service provided by transport infrastructure. Little attention has been paid to the integrated effects of pedestrian bike infrastructure and motorized infrastructure. Moreover, besides land use/land cover change driven by transport networks (Chaudhuri and Clarke 2015; Mothorpe et al. 2013; Müller et al. 2010; Patarasuk 2013), there is a lack of investigation into the transport characteristics of different land use/land cover types, which can bring the perspective of integrated land use transport to urban spatial planning.[1]

1.2. Problem Definition

Since the choice of place of residence is one of the most important issues for people in choosing their desired neighborhood, so this project was defined in such a way as to be able to study the amount and types of land uses that people in each neighborhood have access to. For this reason, in New York City and in the Manhattan borough, different neighborhoods were considered for the study. This borough is specified in the Figure 1-1 below.

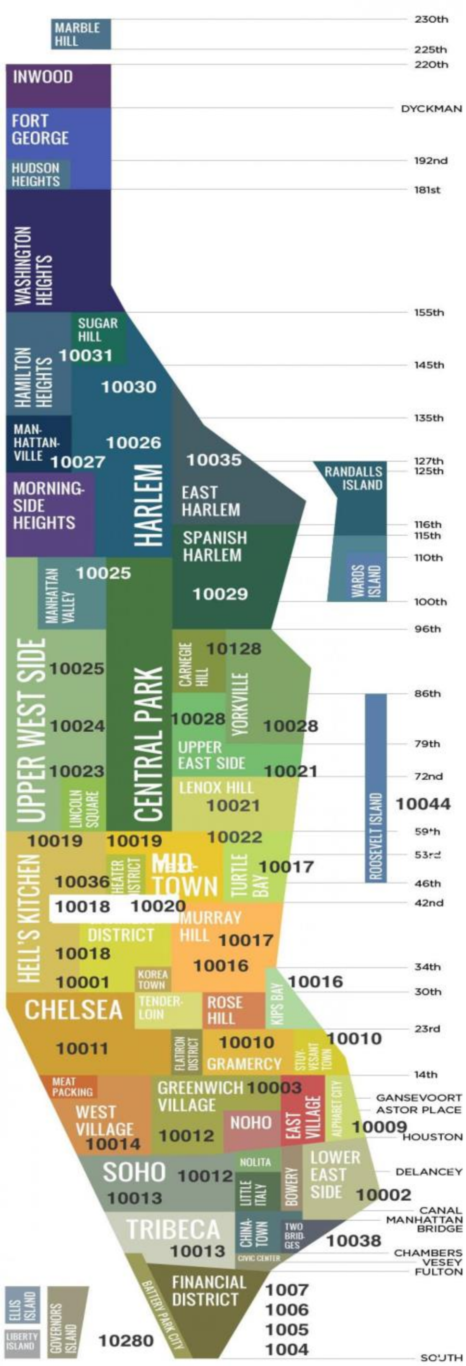

Figure 1-1: overview of Manhattan area map

1.3. Who would be interested in this project

The results of this project are suitable for all those who intend to choose the desired neighborhood to stay. Since the proposed method for this project is not limited to the study area, so the method can be generalized to other urban areas.

1.4. Study area

Manhattan (/mænˈhætən, mən-/), often referred to by residents of the New York City area as the City, is the most densely populated of the five boroughs of New York City, and coextensive with the County of New York, one of the original counties of the U.S. state of New York. Manhattan serves as the city's economic and administrative center, cultural identifier, and historical birthplace. The borough consists mostly of Manhattan Island, bounded by the Hudson, East, and Harlem rivers; as well as several small adjacent islands. Manhattan additionally contains Marble Hill, a small neighborhood now on the U.S. mainland, separated from the rest of Manhattan by the Harlem Ship Canal and later connected using landfill to the Bronx. Manhattan Island is divided into three informally bounded components, each aligned with the borough's long axis: Lower, Midtown, and Upper Manhattan [2].

Manhattan has been described as the cultural, financial, media, and entertainment capital of the world, and the borough hosts the United Nations Headquarters. Anchored by Wall Street in the Financial District of Lower Manhattan, New York City has been called both the most economically powerful city and the leading financial center of the world, and Manhattan is home to the world's two largest stock exchanges by total market capitalization: the New York Stock Exchange and NASDAQ. Many multinational media conglomerates are based in Manhattan, and the borough has been the setting for numerous books, films, and television shows. Manhattan real estate has since become among the most expensive in the world, with the value of Manhattan Island, including real estate, estimated to exceed US$3 trillion in 2013; median residential property sale prices in Manhattan approximated US$1,600 per square foot ($17,000/m2) as of 2018, with Fifth Avenue in Midtown Manhattan commanding the highest retail rents in the world, at US$3,000 per square foot ($32,000/m2) per year in 2017 [2].

2. Data acquisition and cleaning

2.1. Data Source

The information used in this project consists of two parts:

1- The first part is information in json format, which is used to determine the names of borough and their geographical location. This information was extracted from the third week of practice.

2- The second part of the information is extracted from the Foursquare, which includes places based on their land use classification. This information is used to classify land uses around our study sites.

2.2. Data preparation

The following steps are followed for this section:

1- Creating map of Manhattan using latitude and longitude values and add locations of the neighborhood along with the venues (Location of the neighborhood and venues are Displayed in red and blue). This issue is shown in Figure 2-1.

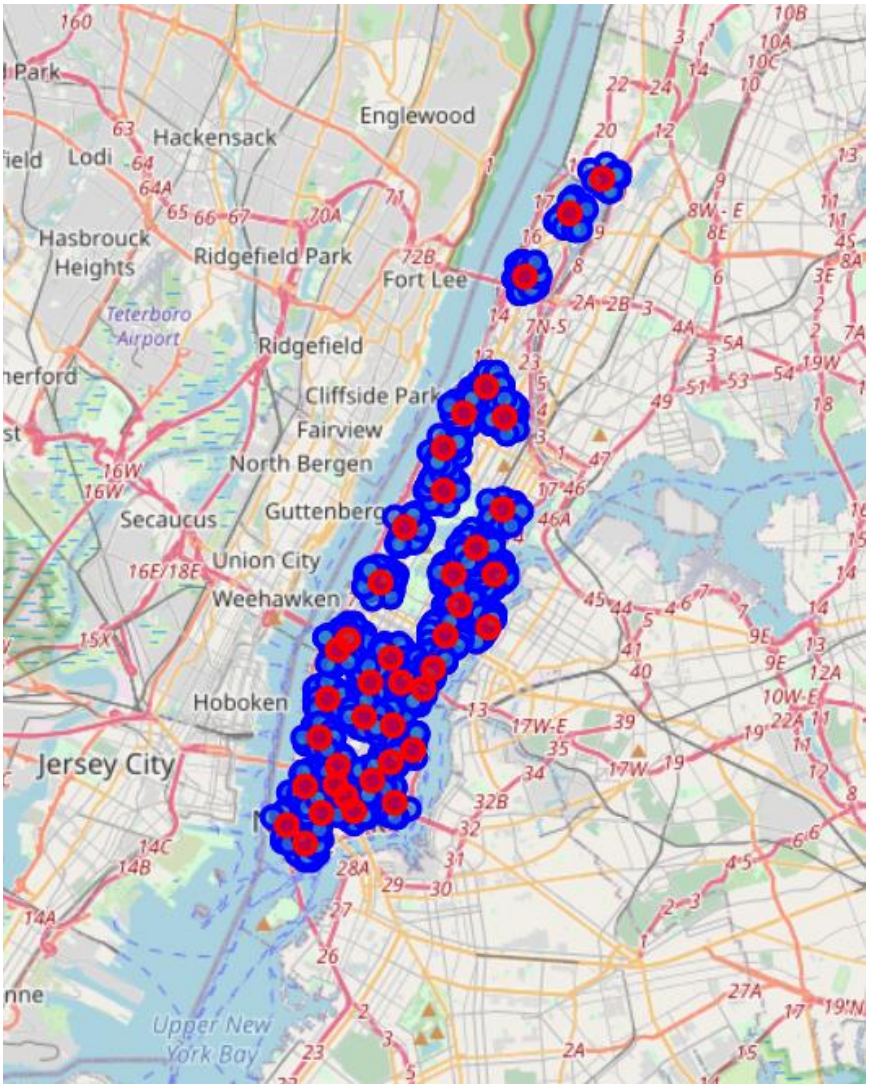

Figure 2-1: overview of Manhattan area map with locations of the neighborhood along with the venues

2- Add a column named "Category". In this column, we create our classification using the classification in the Jason file. In this form of classification, we put all the items that are related to a type of land use in one category. For example, "shop" and "store" land use are put into the category of "Shop & Store". In order to classify land uses, 11 categories have been considered. These classifications are:

- Neighborhood Latitude
- Neighborhood Longitude
- Arts
- Bar
- Cafe
- Entertainment
- Health & Beauty
- Office
- Restaurant & Food
- School
- Shop & Store
- Sport
- Transit

3- Classification results, After performing the classification, the number of classifications that do not fit in our category will be determined. The reason for the existence of unclassified cases is not to go into the details of land use, and the study of land use is essential. Number of venue category that categorized: 2777 ,and Number of uncategorized: 409.

4- Calculating the number of land uses in each Neighborhood, By summing the number of land uses in each neighbourhood, the total number of land use is determined according to the type and neighbourhood location.

3. Analyze Data

3.1. Distributing of each land use in Neighborhoods

In order to investigate this, a box plot chart is drawn. As can be seen in the diagram below, the land use of "food and restaurant" and "shopping and store" in the distribution network is higher. Still, the amount of this distribution varies from region to region. This is well illustrated in the diagram. Other land uses that are found in these areas and have a high distribution. The box plot is shown in Figure 3-1.

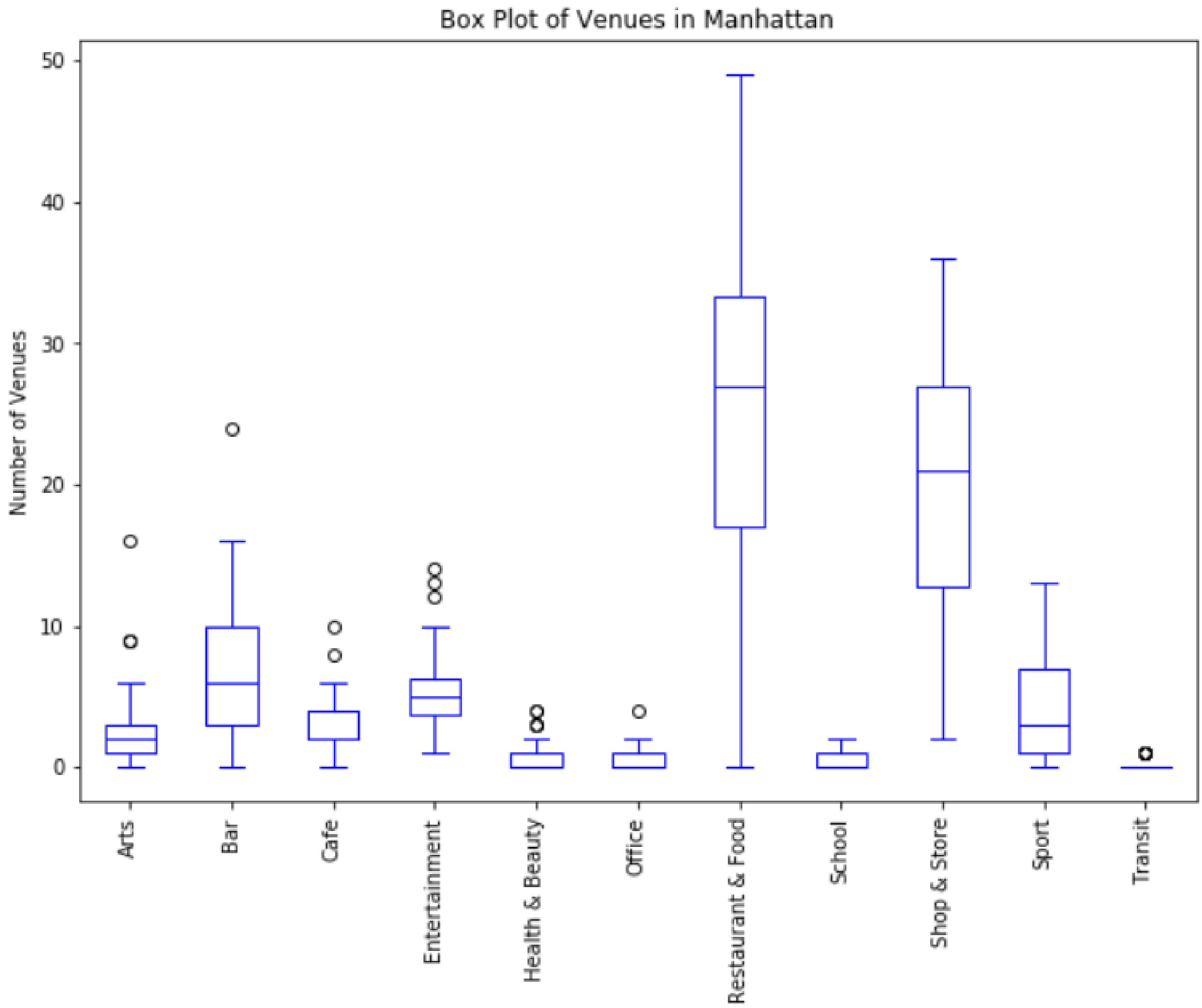

Figure 3-1: Box Plot of Venues in Manhattan

3.2. Show Manhattan Venues for each Neighborhood

To show the distribution of venues in a different neighbourhood, the map of Manhattan has been used, and in it, each of the venues is shown on the map with its specific land use color. The map is shown in Figure 3-2.

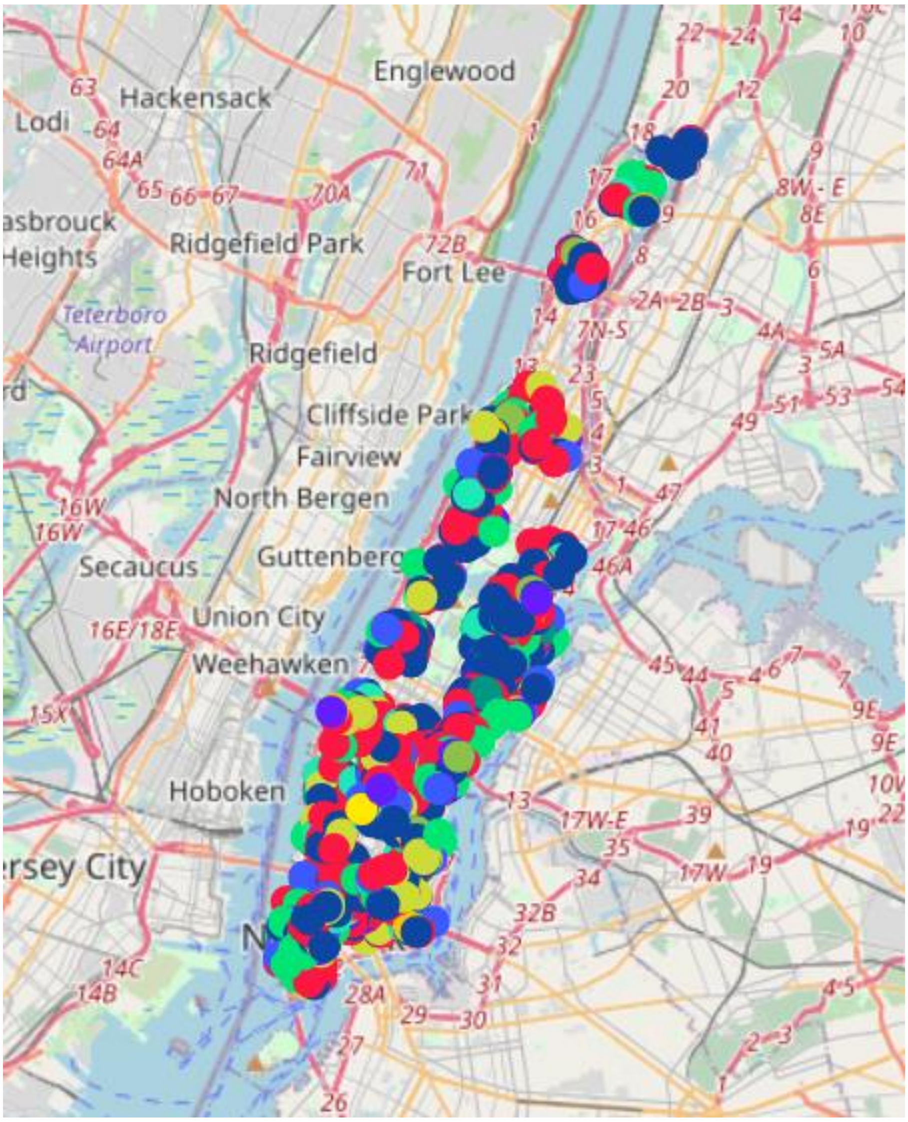

Figure 3-2: Manhattan area venues distribution

3.3. Clustering different Neighborhood based on the proximity of their land use to each other

In order to select the best neighbourhood that has the most access to different land uses, we use clustering. In this method, considering 5 clusters for the neighbourhoods in the Methane area, we perform the clustering operation. Figure 3-3 displayed clusetering distribution of manhattan area.

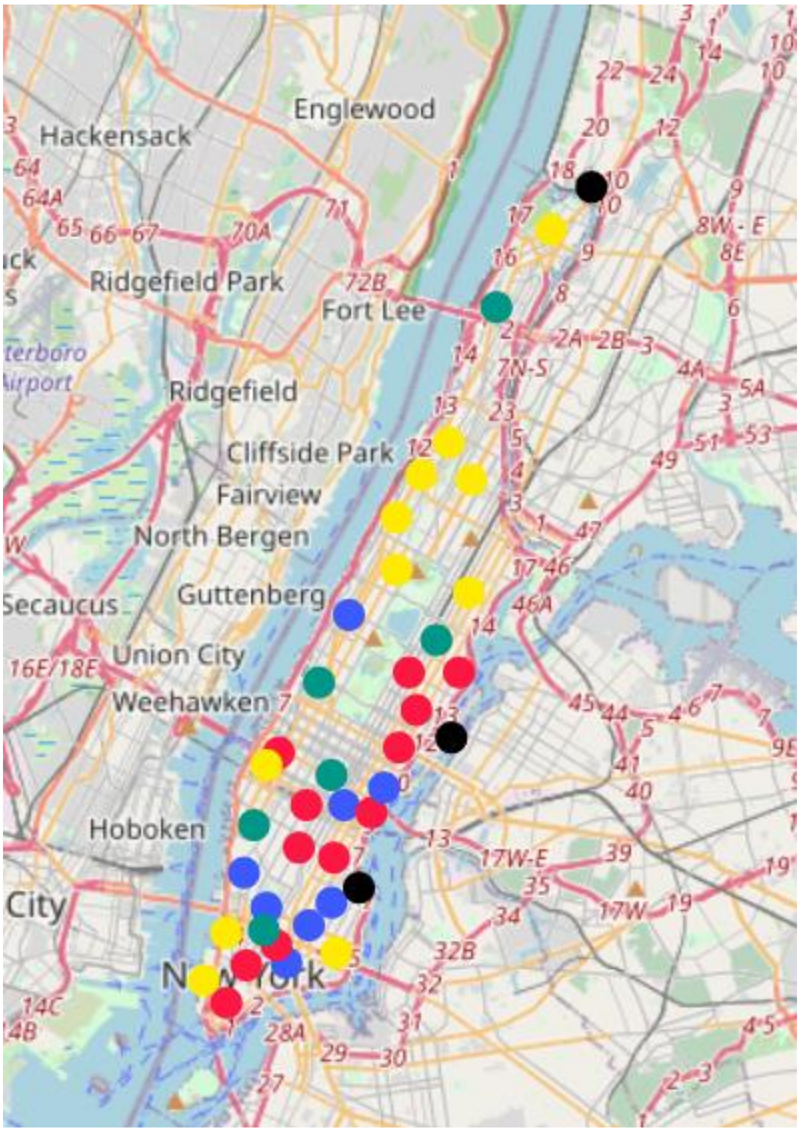

Figure 3-3: Manhattan area clusters distribution

In the two bar graphs below, the amount of each land use is shown separately for each cluster. As can be seen in this figure, cluster number 1, which includes Neighborhoods 'Marble Hill', 'Roosevelt Island', 'Stuyvesant Town', has the least access to different uses. Other clusters are better in terms of access than the neighbourhoods in the first cluster but differ in the amount of access to each of the uses. Figure 3-4 and Figure 3-5 displayed bar chart of venues in each cluster in manhattan area.

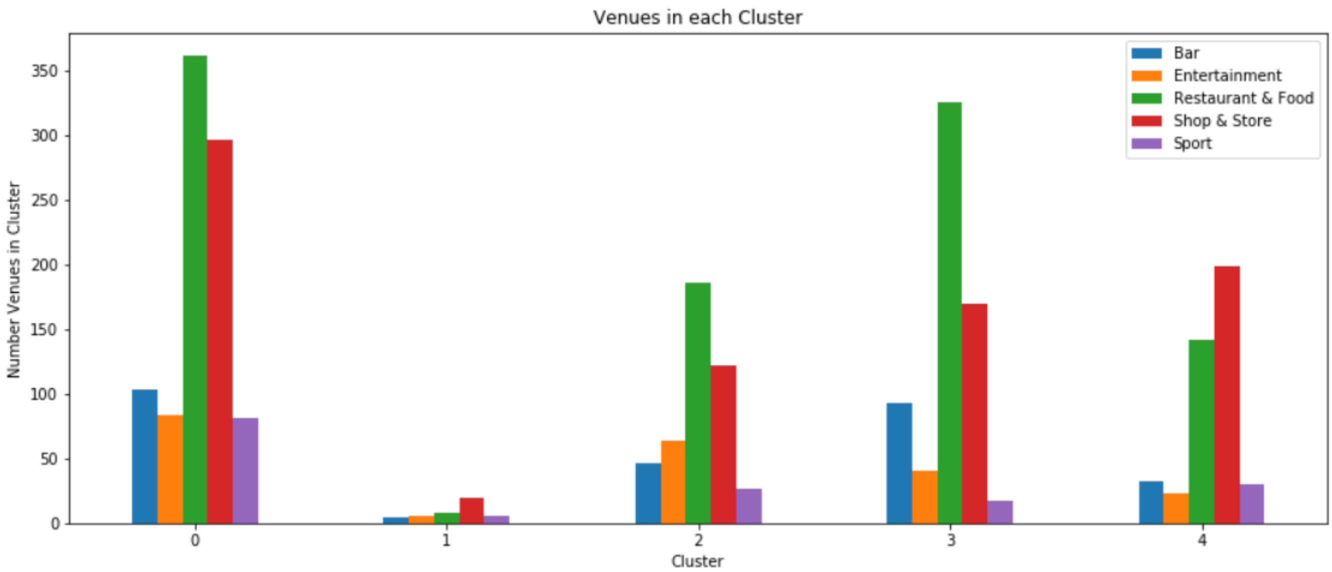

Figure 3-4: Venues in each cluster

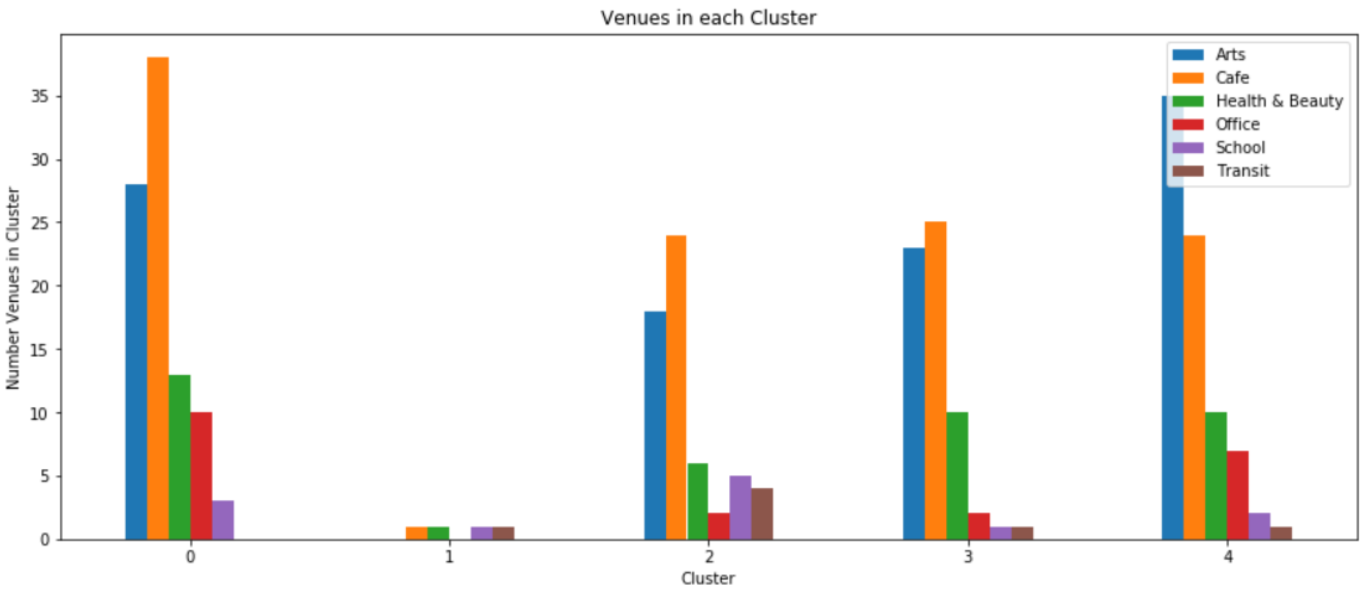

Figure 3-5: Venues in each cluster

Bibliography

[1] Wang, Z.; Han, Q.; de Vries, B. Land Use/Land Cover and Accessibility: Implications of the Correlations forLand Use and Transport Planning.Appl. Spat. Anal. Policy2018, 1–18

[2] Https://en.wikipedia.org/wiki/Manhattan