Correlation of venue composition and political affiliations of districts in Vienna

1. Introduction

1.1. Background

When we think about cities like Paris we have certain images in our heads like fashion boutiques, cafes on every corner, wineries and many museums. However, every city has unique areas that are very different from other parts of that same city. Because of that, some districts or neighbourhoods are more popular than others and thus generate more money because of a higher consumption of services and goods. That is why it would be useful to find some similarities of these more successful districts so it might be applied to the less popular districts to make them more prosperous.

1.2. Problem

For this project, we will look at the districts of Vienna. We know that some districts generate more revenue from the service industry because people from other districts or even tourists prefer these districts over others.

The main questions for this project are whether these successful districts have something in common and which factors are influencing their positive development. For the latter part, we will look at the leading political party of every district to determine if they have an influence regarding the development of their respective districts. The reason is that every district partially managed by a political party and they have some power over the development of their districts. Of course they are boasting of their great management and that is why we want to see if there are any correlation between the leading party of a district and their composition of their venues.

1.3. Interest

The results of this analysis might be of interest to political parties if we do find any positive correlations between the success of a district and its affiliation to a certain political party. Real estate developers might also be interested to find out about emerging districts that are still cheap compared to the most popular ones. Various types of businesses such as restaurant operators might also be interested for the same reason as real estate operators.

2. Data

2.1. Data Sources

The main data source will be Foursquare's database and the coordinates of Vienna's districts. Since I could not find any online sources with the coordinates of all 23 districts of Vienna, I have decided to get the data myself. For that I have created a dataframe containing all district names of Vienna and 2 extra columns for the latitude and longitude. The district names were filled manually and for the other data I have used the geocode function to fetch the longitude and latitude

information. After retrieving this data it was possible to call the “venues/explore” API-Endpoint of Foursquare for every district and to get information about the surrounding venues. The only relevant information from Foursquare for this project is the “Venue Category”. With this information it is then possible to generate a list of venues for every district and therefore a venue profile based on the venue composition. With this venue profile it is then possible to cluster similar ones together and link them to their districts.