

Capstone Project - Grouping Neighborhoods by Proximity and Content

The Problem

- You have a successful business in one part of the city.
- You need to expand to a new area of the city, but don't know where or how similar other parts of the city are to yours

The Data Solution

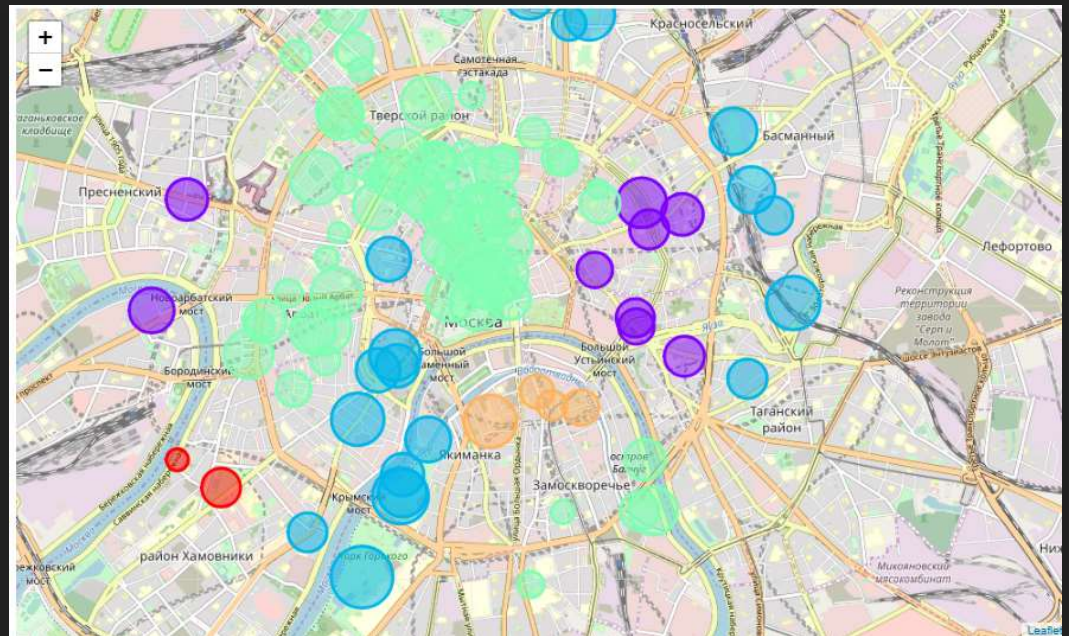
- Use data from Foursquare's Venues API to identify clusters of venues throughout the city and then classify them based on the content of their categories
- Also can scale these venues by the # of likes they have to get a feel for traffic/volume

Methodology

- Analyze different endpoints within Venues API.
 - Used
 - Explore
 - Likes
 - Categories
 - Analyzed, not used
 - Search
- Zip/postal codes may not be available &/or may artificially segment groups of venues.
 - Used clustering to identify "neighborhoods"
- Then cluster venue categories & subcategories in each neighborhood
- Also looked at clustering venues themselves based on their categories
- Explored supplementing with Twitter API info, decided the value-add wasn't worth the extra effort & complexity

Results

- There are several neighborhoods in the green sector at the top of the map, which have been clustered together by content
- You can see some clusters in opposite areas of the city which are similar. If you have a successful business in one color of area, another branch in a similar colored neighborhood may be the right location!



End