



Data Analysis Project

Analyzing neighborhoods in Madrid



Introduction

- Detecting the most beneficial neighborhood in Madrid for a potential coffee opening



Data

- Our Data source is from wikipedia:
https://en.wikipedia.org/wiki/Districts_of_Madrid
- Data cleaning was needed - data came raw
- Only needed columns: District names
- Later on pulling longitud and lattitude with the help of Geocoder library

Methodology

- Importing the necessary libraries
- Showing map with the help of Folium
- Using Foursquare API
- Pulling in the most common venues for each neighborhood
- Clustering
- Finding out the best k means



Results

- The best area to open a coffee shop is the Center of Madrid
- This was decided by the cluster function



Discussion

- Following questions: Transport? Rent? Parking space?
- Project biggest aspect is done but lot of question has to be answered