Battle of Neighborhoods

Best Miami Neighborhood to Open an Italian Restaurant

Daphne Cheung, Applied Data Science Capstone Project

Meet Matilda!

Matilda is an established chef and restaurateur in New York City who would like to open an Italian restaurant in Miami but has no existing knowledge of where to do so.

I will be clustering Miami venue data to determine which neighborhoods are most suitable for Matilda to open her restaurant!
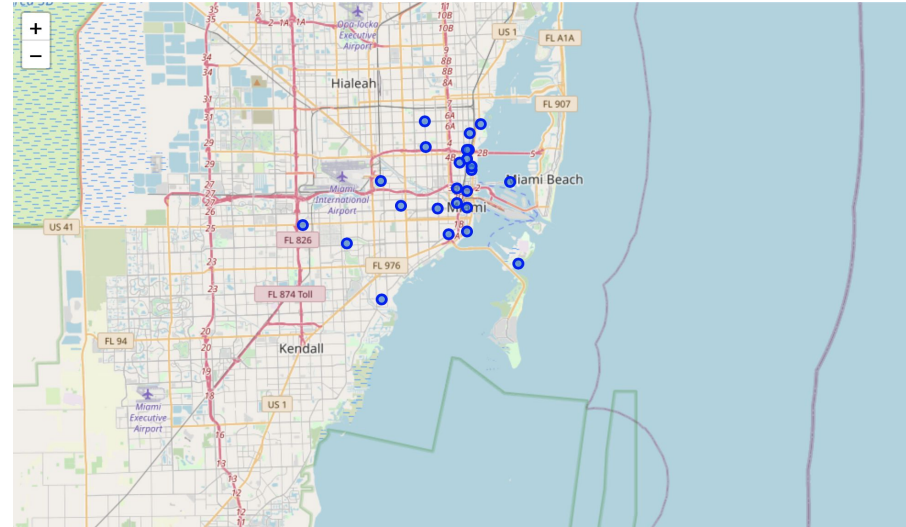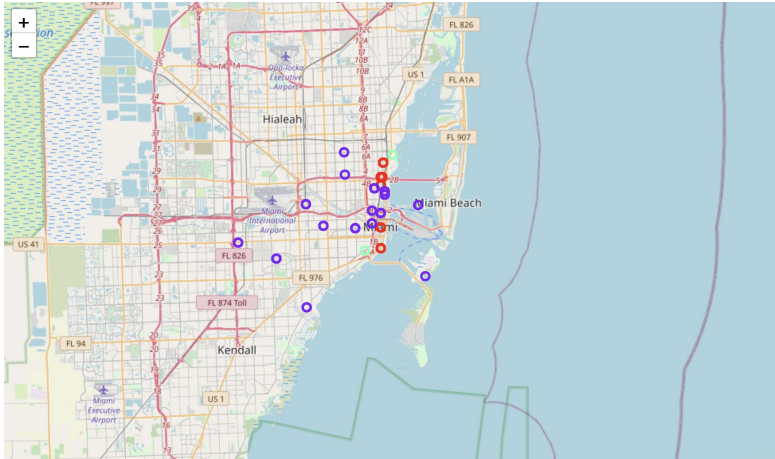
Data Collection and Cleaning

- Miami neighborhoods and coordinates gathered from Wikipedia
  - Transformed into Pandas DataFrame
  - 25 neighborhoods in total
  - Dropped irrelevant columns, only retaining neighborhood names and coordinates
- Top venues for neighborhoods extracted using Foursquare developer API
  - Venue coordinates and categories obtained
  - 137 unique venue categories
  - Transformed venue data into dummy variables to calculate frequency of each category

Data Exploration and Visualization

- Neighborhoods mapped out using Folium package
- Frequency of each venue category averaged for each neighborhood
  - Produced top 5 categories for each neighborhood
  - 9 of which contained Italian Restaurants

Applications of K-Means Clustering

- K-Means Clustering algorithm from scikit-learn applied on Miami neighborhoods with Italian restaurants
- Cluster labels produced and merged with existing neighborhood data
- Clusters visualized using the Folium package

Results

- In examining the individual clusters...
  - Cluster 2 had the highest frequency of Italian restaurants, containing neighborhoods: The Roads and Upper Eastside
  - Cluster 1 had the lowest frequency of Italian restaurants, containing neighborhoods: Park West, Wynwood, West Flagler, Virginia Key, Venetian Islands, Allapattah, Edgewater, Lummus Park, Coral Way, Arts & Entertainment District, Coconut Grove, Grapeland Heights, Liberty City, Little Havana, and Flagami

Next Steps for Matilda

- Explore real estate in The Roads or Upper Eastside
    - Research established and existing competitors
- Consider opening a restaurant in cluster 1 for fewer direct competitors