Understanding of Denver restaurant business

CAPSTONE PROJECT PRESENTATION

VERONIKA V.

Presentation outline

- Introduction
- Methodology (workflow & software)
- Data summary & processing
- Exploratory data analysis
- Cluster analysis
- Summary of findings

Introduction

- **Background:** Denver is the largest city in Colorado with near 31 million visitors per year and tourism revenue at $6.5 billion (data for 2017)
- **Business issue:** the customer plans to open a new restaurant for city visitors
- **Business request:** identify optimal locations (touristic areas) and type of restaurant to open

Methodology and software

- Publicly available data will be used

- Exploratory data analysis will be performed to understand:
  (1) Restaurants and hotels distribution in Denver
  (2) Identify the most (un)common restaurant types to understand untapped opportunities

- Kmean Clustering will be done to identify ‘good’ and ‘bad’ locations to open a restaurant

- Analysis will be done on Python 3.6 available in IBM Watson Studio.

Data

- Denver Zip area data was downloaded and extracted from public repository https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude

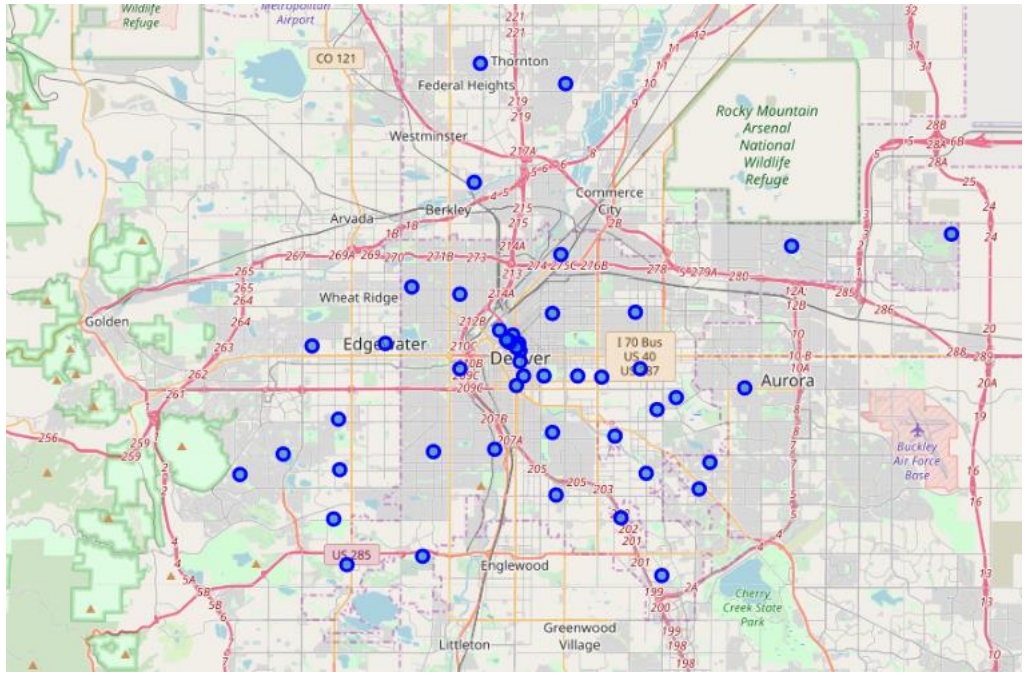
Snapshot of Denver Zip data

US Zip Code Latitude and Longitude

Information | Table | Map | Analyze | Export | API

	Zip	City	State	Timezone	Daylight savings time flag	Latitude	Longitude
1	80234	Denver	CO	-7	1	39.909	-105.008
2	80280	Denver	CO	-7	1	39.717	-104.907
3	80257	Denver	CO	-7	1	39.739	-104.408

Denver Zip data visualization with Folium (Python 3.6)



Data

- A Foursquare public database (https://foursquare.com/) was used to obtain information For each Zip area. Requests for hotels and restaurants were generated:

url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&categoryId=CategoryId '

Where ll – latitude and longitude of each Zip area
CategoryId = 4d4b7105d754a06374d81259 (restaurants), 4bf58dd8d48988d1fa931735 (hotels)
Radius was set 500 m

Information about venue latitude and longitude, name and category was extracted and summarized in the table:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Group
0	80235	39.648328	-105.08431	Hampton Denver Southwest	39.648626	-105.082271	Hotel	Hotels
1	80235	39.648328	-105.08431	Comfort Suites Southwest	39.649333	-105.079389	Hotel	Hotels

Data processing

Data contained information about different venue categories:

```
(['Hotel', 'Bar', 'General Travel', 'Hostel', 'Hotel Pool',
  'American Restaurant', 'Bed & Breakfast',
  'Residential Building (Apartment / Condo)', 'Hotel Bar',
  'College Academic Building', 'Café', 'Pet Service', 'Resort'],
```

```
['Pizza Place', 'Mexican Restaurant', 'Fast Food Restaurant',
 'Asian Restaurant', 'Sushi Restaurant', 'Diner', 'Restaurant',
 'American Restaurant', 'Seafood Restaurant', 'Steakhouse',
 'New American Restaurant', 'Burger Joint', 'Breakfast Spot',
 'Food Truck', 'Modern European Restaurant', 'Bakery',
 'Fried Chicken Joint', 'Food', 'Sandwich Place',
 'Mongolian Restaurant', 'Taco Place', 'Italian Restaurant',
 'Bagel Shop', 'Gastropub', 'Soup Place', 'Brazilian Restaurant',
 'Deli / Bodega', 'Salad Place', 'French Restaurant', 'BBQ Joint',
 'Japanese Restaurant', 'Ramen Restaurant', 'Café', 'Snack Place',
 'Chinese Restaurant', 'Food Court', 'Greek Restaurant',
 'Noodle House', 'Vietnamese Restaurant',
 'Cajun / Creole Restaurant', 'Poke Place', 'Theme Restaurant',
 'Burrito Place', 'Middle Eastern Restaurant',
 'Mediterranean Restaurant', 'Hot Dog Joint', 'Thai Restaurant',
 'Donut Shop', 'Wings Joint', 'Indian Restaurant',
 'Korean Restaurant', 'North Indian Restaurant',
 'Vegetarian / Vegan Restaurant', 'Pet Café', 'Cuban Restaurant',
 'Comfort Food Restaurant', 'Caribbean Restaurant',
 'Southern / Soul Food Restaurant'], dtype=object)
```

For hotel group we filtered out data for hotel bars, pet service, American restaurant café and academic buildings.
For restaurant group we focused on restaurants only (e.g. food courts were excluded)

Exploratory data analysis

Firstly, we explored top hotel and restaurant categories in Denver:

N of venues for each hotel venue category

Venue Category	Group	Venue
Hotel	Hotels	254
Hostel	Hotels	8
Residential Building (Apartment / Condo)	Hotels	1
Resort	Hotels	1

N of venues for each restaurant venue category

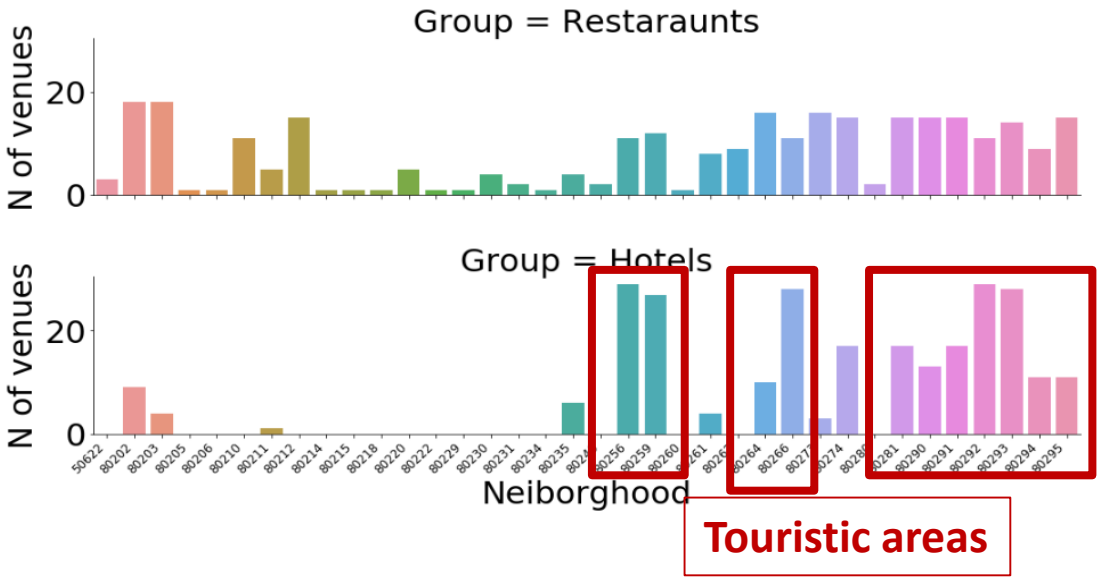
Venue Category	Group	Venue
American Restaurant	Restaraunts	64
Mexican Restaurant	Restaraunts	46
Italian Restaurant	Restaraunts	34
Restaurant	Restaraunts	17
Japanese Restaurant	Restaraunts	12
Brazilian Restaurant	Restaraunts	11
French Restaurant	Restaraunts	11
New American Restaurant	Restaraunts	11
Sushi Restaurant	Restaraunts	11
Asian Restaurant	Restaraunts	10
Cajun / Creole Restaurant	Restaraunts	7
Indian Restaurant	Restaraunts	6
Thai Restaurant	Restaraunts	6
Theme Restaurant	Restaraunts	6
Chinese Restaurant	Restaraunts	5
Mediterranean Restaurant	Restaraunts	4
Ramen Restaurant	Restaraunts	4
Vietnamese Restaurant	Restaraunts	4
Fast Food Restaurant	Restaraunts	3
Korean Restaurant	Restaraunts	3
Middle Eastern Restaurant	Restaraunts	3
Seafood Restaurant	Restaraunts	2
Vegetarian / Vegan Restaurant	Restaraunts	2
Caribbean Restaurant	Restaraunts	1
Comfort Food Restaurant	Restaraunts	1
Cuban Restaurant	Restaraunts	1
Greek Restaurant	Restaraunts	1
Modern European Restaurant	Restaraunts	1
Mongolian Restaurant	Restaraunts	1
North Indian Restaurant	Restaraunts	1
Southern / Soul Food Restaurant	Restaraunts	1

- **Most part of visitors live in hotels**
- **A lot of American, Mexican and Italian restaurants are already open, so, better to open vegetarian or seafood restaurant**

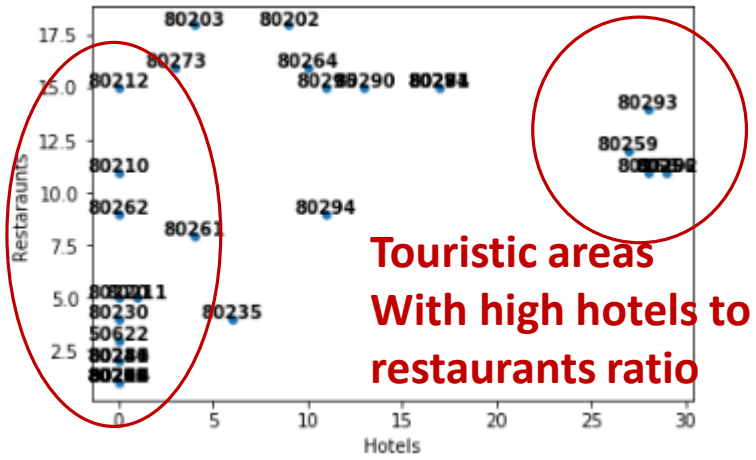
Exploratory data analysis

Secondly, we explored distribution of hotels and restaurants across the areas. There are a lot of hotels in some areas, howbeit, N of restaurants in these areas is high too:

N of venues in different city areas



Scatterplot of hotels and restaurants (N indicates Zip)

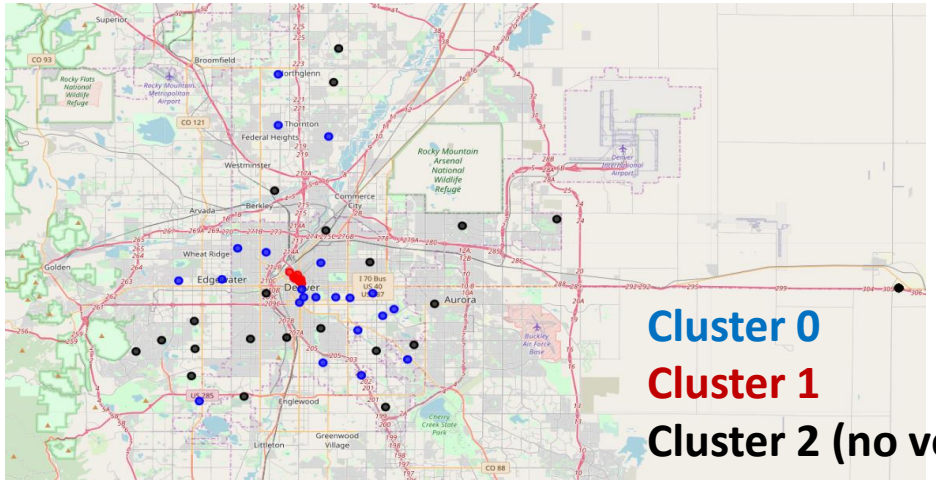


Non-touristic areas with restaurants for local people

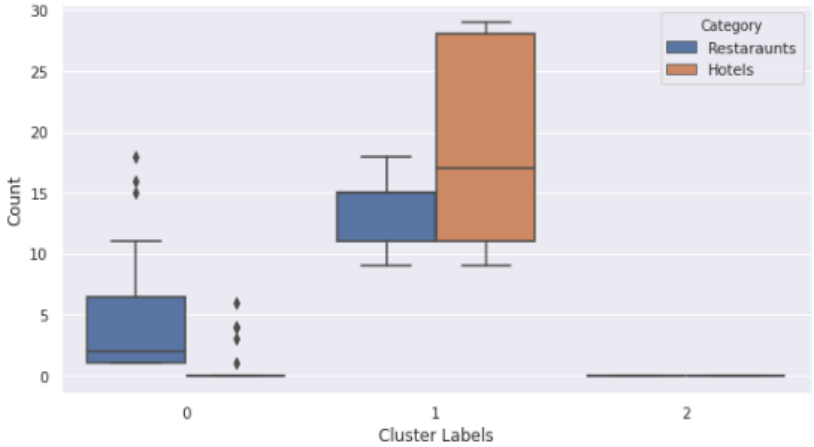
Clustering

Kmean clustering was done (N=2) to identify “good” and “bad locations” + informations about Zip areas without any venues was included:

Denver map with identified clusters



Distribution of hotel and restaurant venues across the clusters



Cluster 1 represents the center of the city and it is the target cluster for opening restaurants for visitors

Summary

- Data from public repositories were used to understand what type of restaurant would be optimal to open in Denver and what city areas should we focus
- Based on the exploratory analysis we would recommend to open a vegetarian or seafood restaurant as very few venues of these categories are presented
- Based on the exploratory data analysis and clustering we would recommend to open restaurant in in the city center (e.g. Zip areas 14199 or 3043)

Jupiter notebook with markdown code is available on Github:
https://github.com/VeronikaVor/Coursera_Capstone/blob/master/Denver_restaraunts.ipynb
Report is available on Github:
https://github.com/VeronikaVor/Coursera_Capstone/blob/master/Denver_restaraunts_report.docx