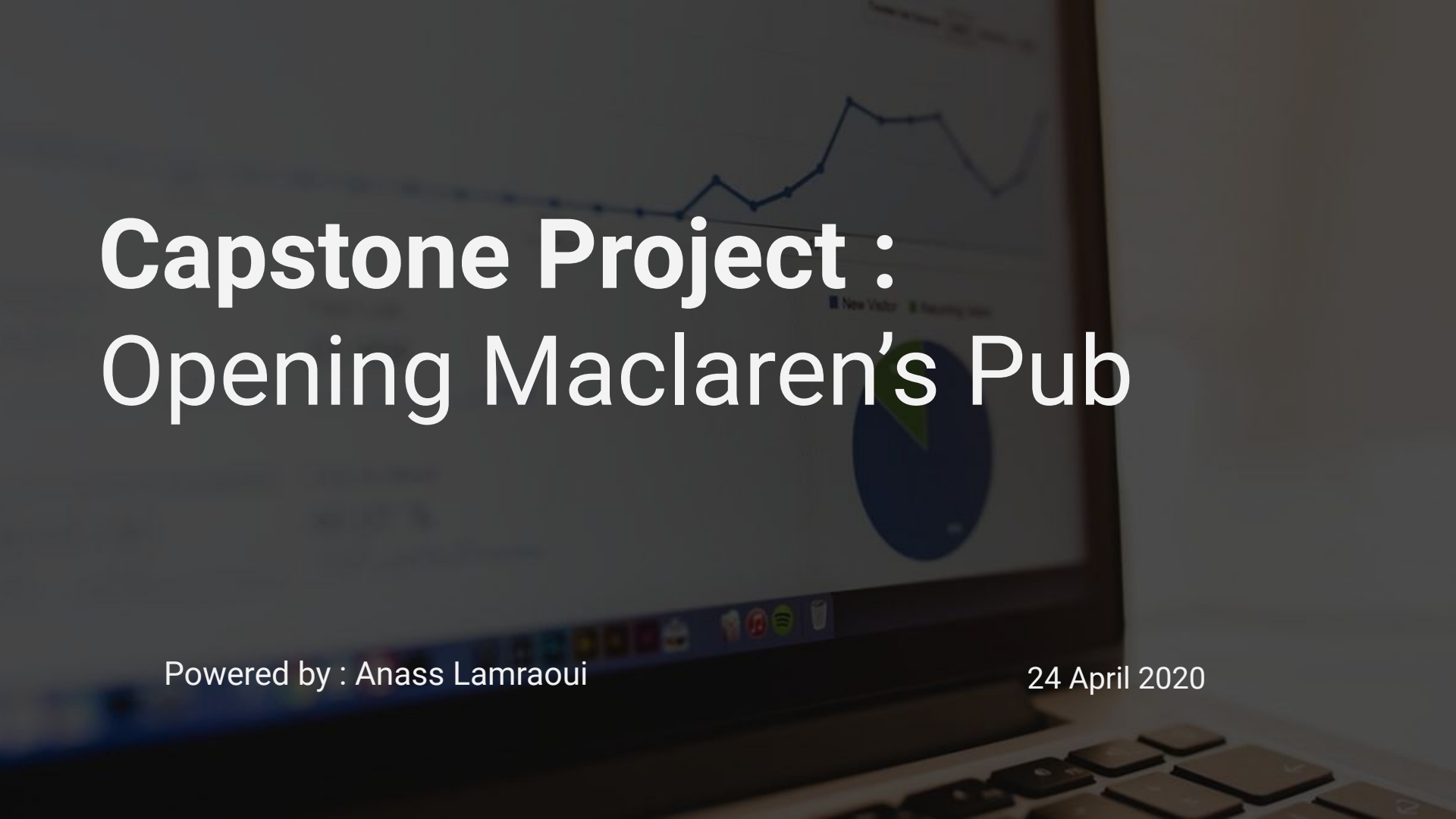


Capstone Project :
Opening Maclaren’s Pub

Powered by : Anass Lamraoui

24 April 2020

Problem Description

The goal of the project is that we want to invest on new pub in the state of Iowa. Where can we open it ?

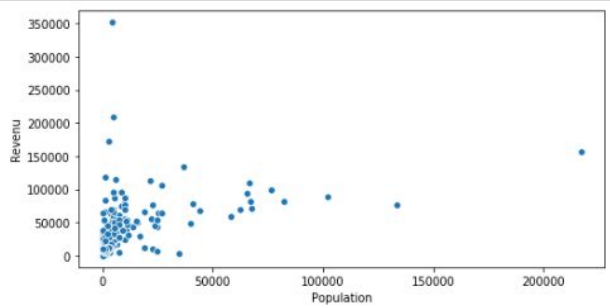
It is important that our investment should be optimal and profitable in the common years.

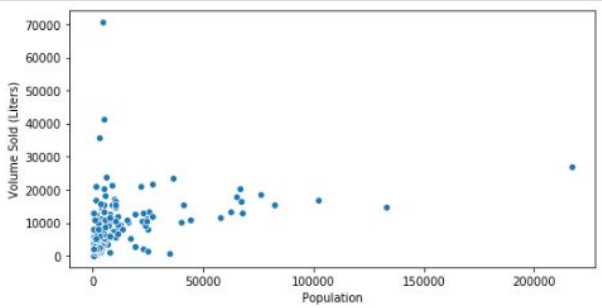
How can we chose the best city to open our new pub ?

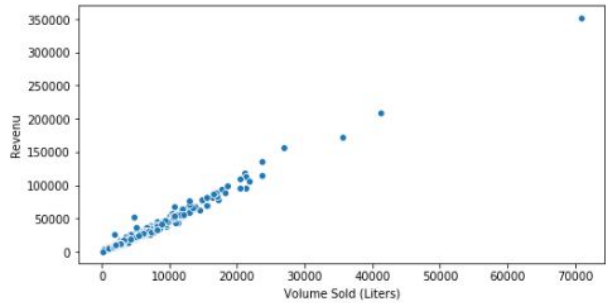
Data acquisition and cleaning

- Liquor Sales Data in the state of IOWA - 2019 : it's csv file taken from https://data.iowa.gov/
  - In total, 1966458 row and 24 features in the raw dataset.
- Cities population Data in the state of IOWA filtered in 2018. It contains the average population of each city. Source : https://data.iowa.gov/
- Rent Data in 2019 contains the average value of rent of each county in the state of IOWA. Source : https://www.rentdata.org/states/iowa/2019
- Our data is analyzed by each city in the state of IOWA

Exploratoring & Modeling I

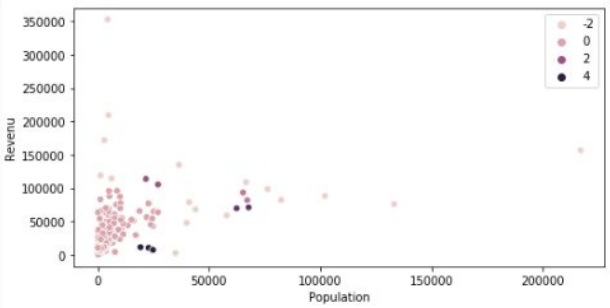


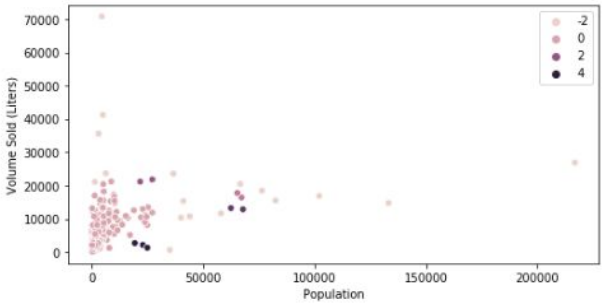


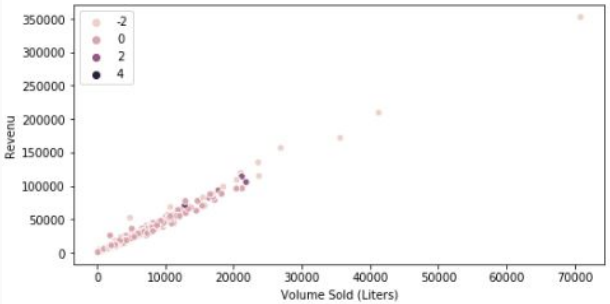


Our Dataset contains a lot of noise. For better clustering we used DBSCAN algorithm to eliminate the noise.

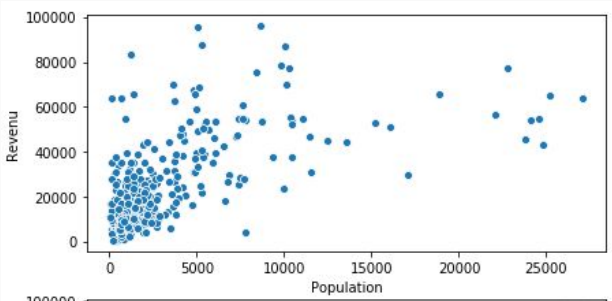
Exploratoring & Modeling I

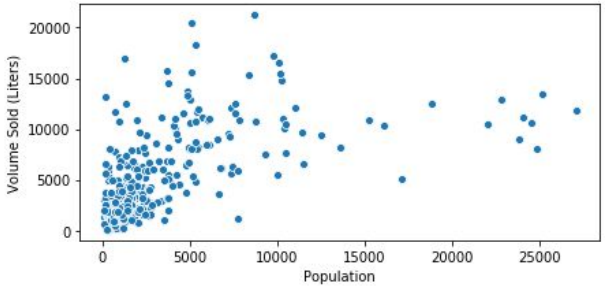


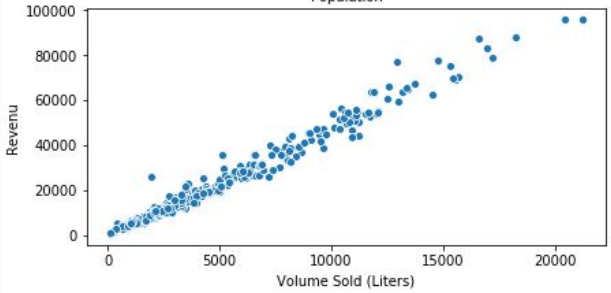




Exploratoring & Modeling I

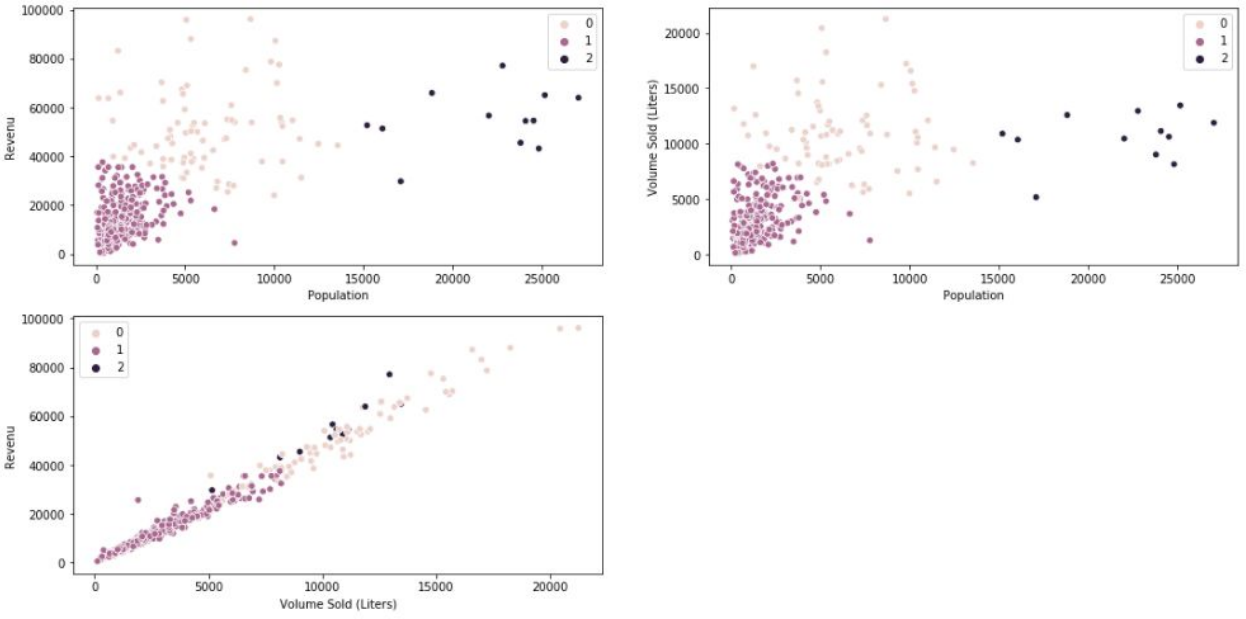






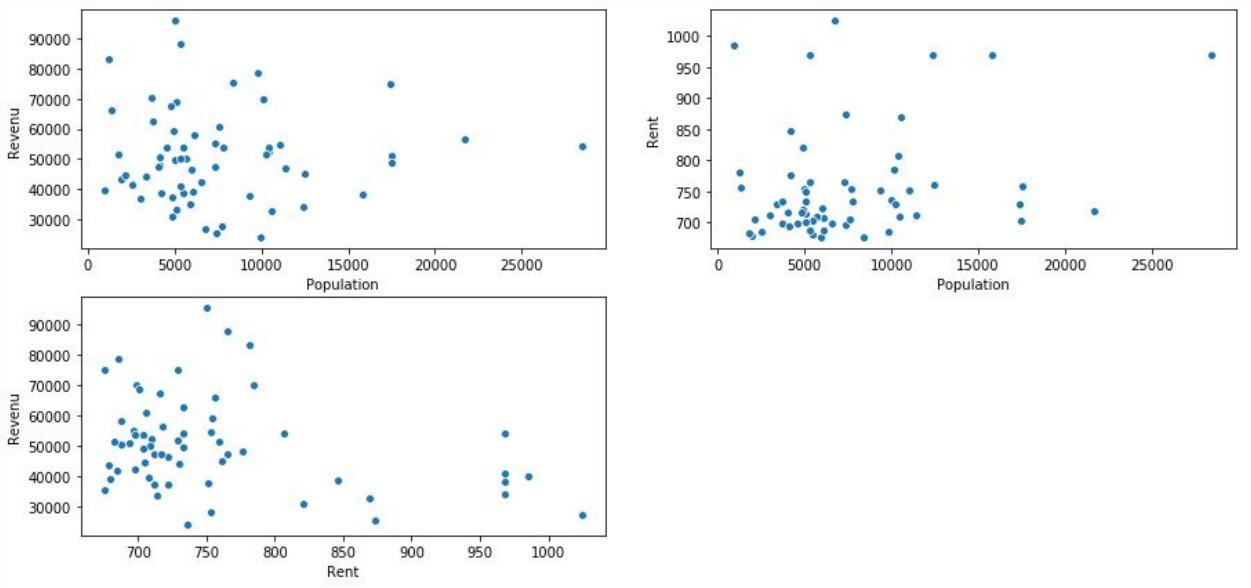
We filtered our dataset and used the k-means algorithm to get the clusters of each city.

Exploratoring & Modeling I



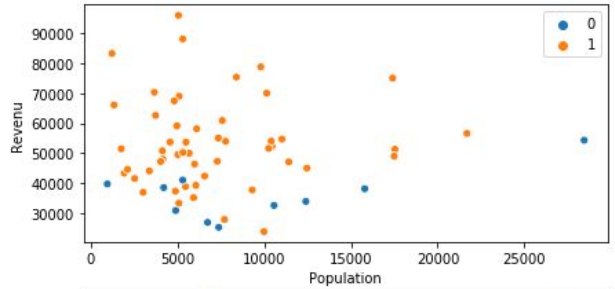
K-means algorithms helped us cluster the cities with higher revenue on alcohol consumption.
We filter our data

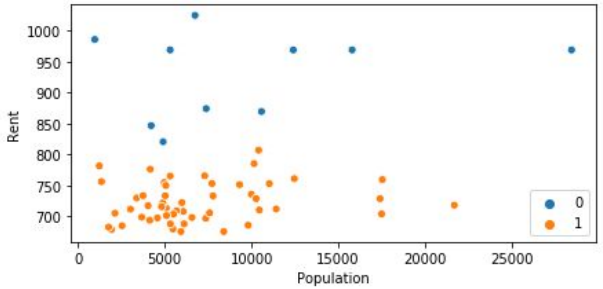
Exploratoring & Modeling II

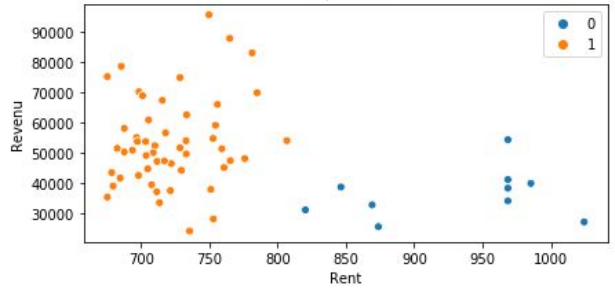


With our data we include the rent data to filter on cities with lower rent value

Exploratoring & Modeling II



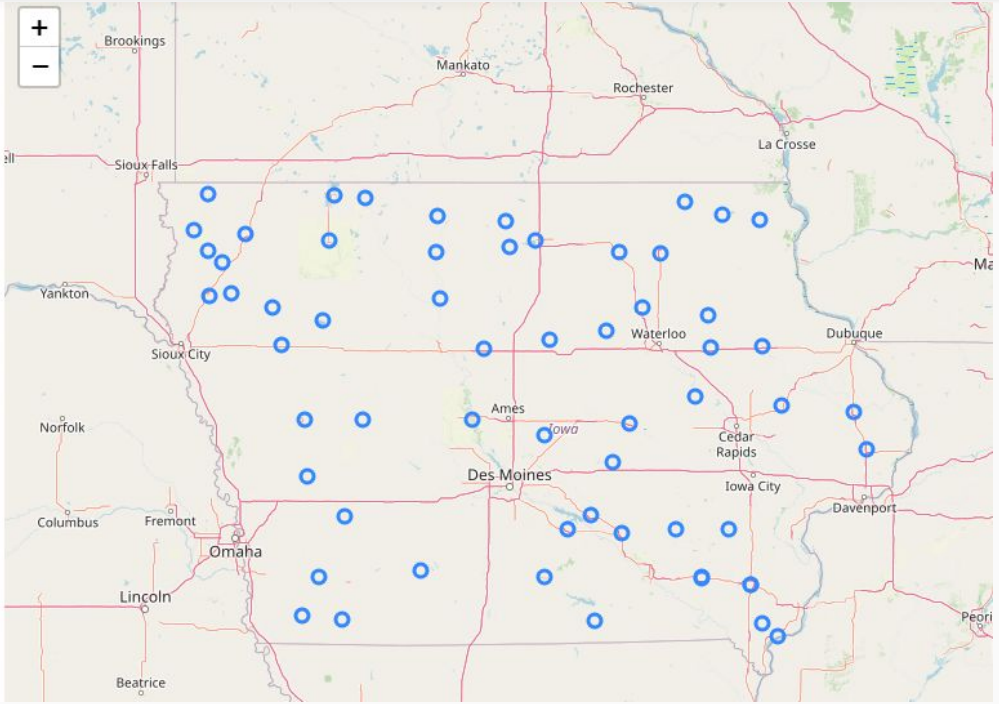




Now we choose the cluster who represent the lowest rent value.

Exploratoring & Modeling III

Using Geocoder we get location of each city and plot it using Folium package.



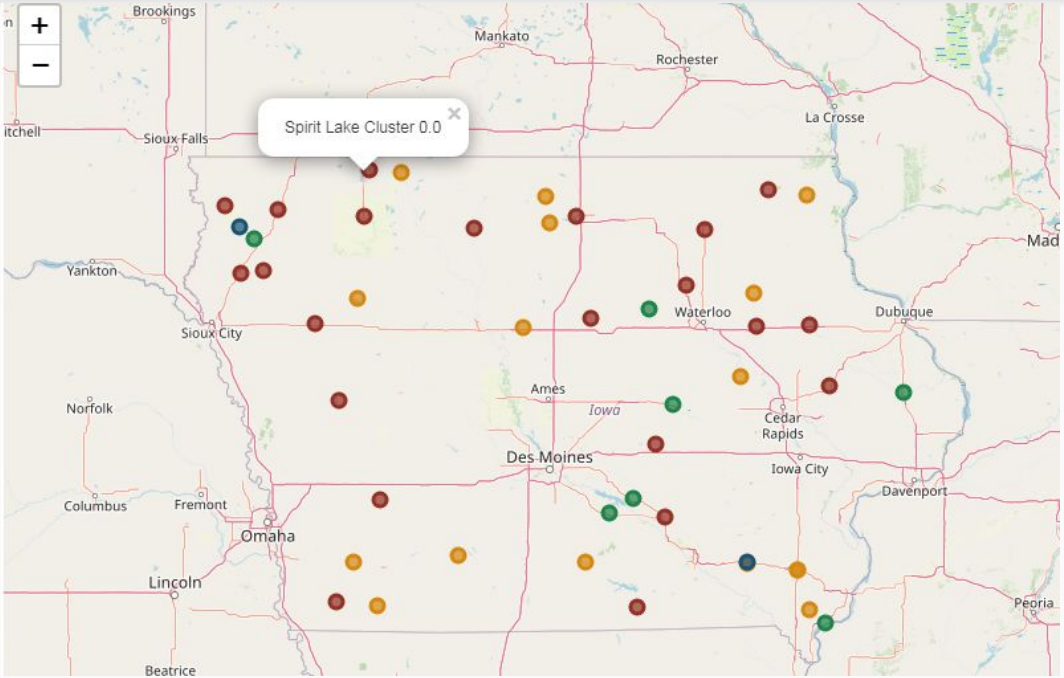
Exploratoring & Modeling III

Now we use the Folium API to get the list of the venues present in each city and their popularity

	City	City Latitude	City Longitude	id	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Algona	43.069966	-94.233019	4d66ed1f58155481542bde55	The Perky Parrot After Dark	43.069039	-94.236002	Cocktail Bar
1	Algona	43.069966	-94.233019	4bf5b5469abec9b69d8124e8	Billie Jo's Bar & Grill	43.068514	-94.237650	Bar
2	Algona	43.069966	-94.233019	4bac324df964a52082ea3ae3	Pep's	43.068985	-94.237120	Bar
3	Algona	43.069966	-94.233019	52685ff5498e9cba961df12f	Locker Room Bar & Grill	43.068753	-94.234676	Sports Bar
4	Anamosa	42.108337	-91.285159	4c018e80b58376b0145e443c	Tucker's Tavern	42.108258	-91.284374	Bar

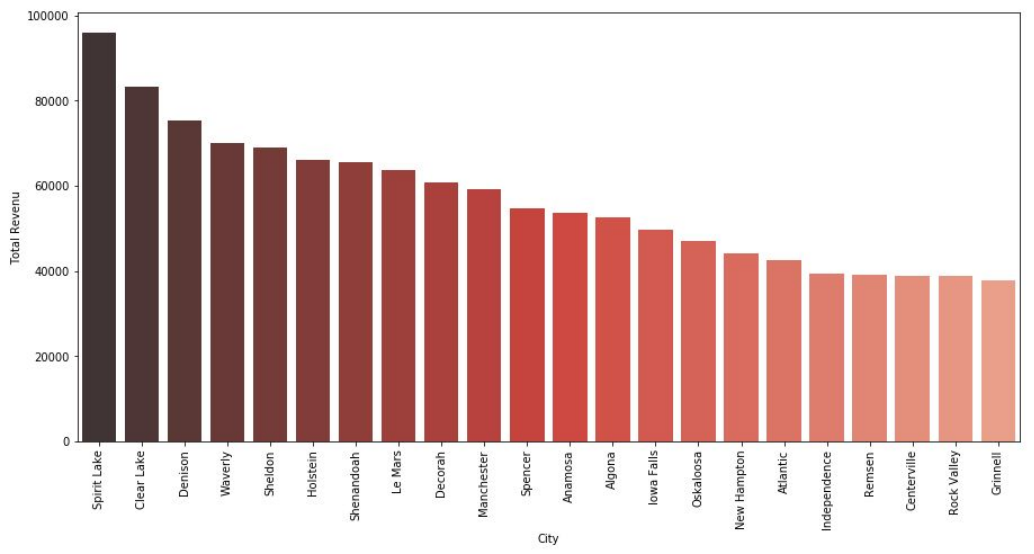
Exploratoring & Modeling III

Now we use the Folium API to get the list of the venues present in each city and use k-means algorithm to classify the cities based on their category venue



Exploratoring & Modeling III

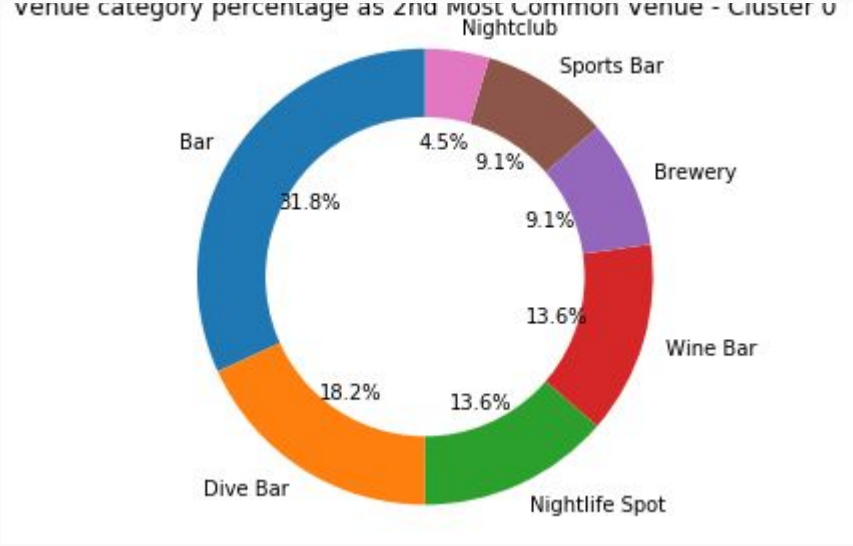
We found that the first cluster knows a diversity in the most common venues categories controlled mostly with regular bars.



Exploratoring & Modeling III

This cluster is characterized with popularity for the nightlife spots, average rent value of 700$.

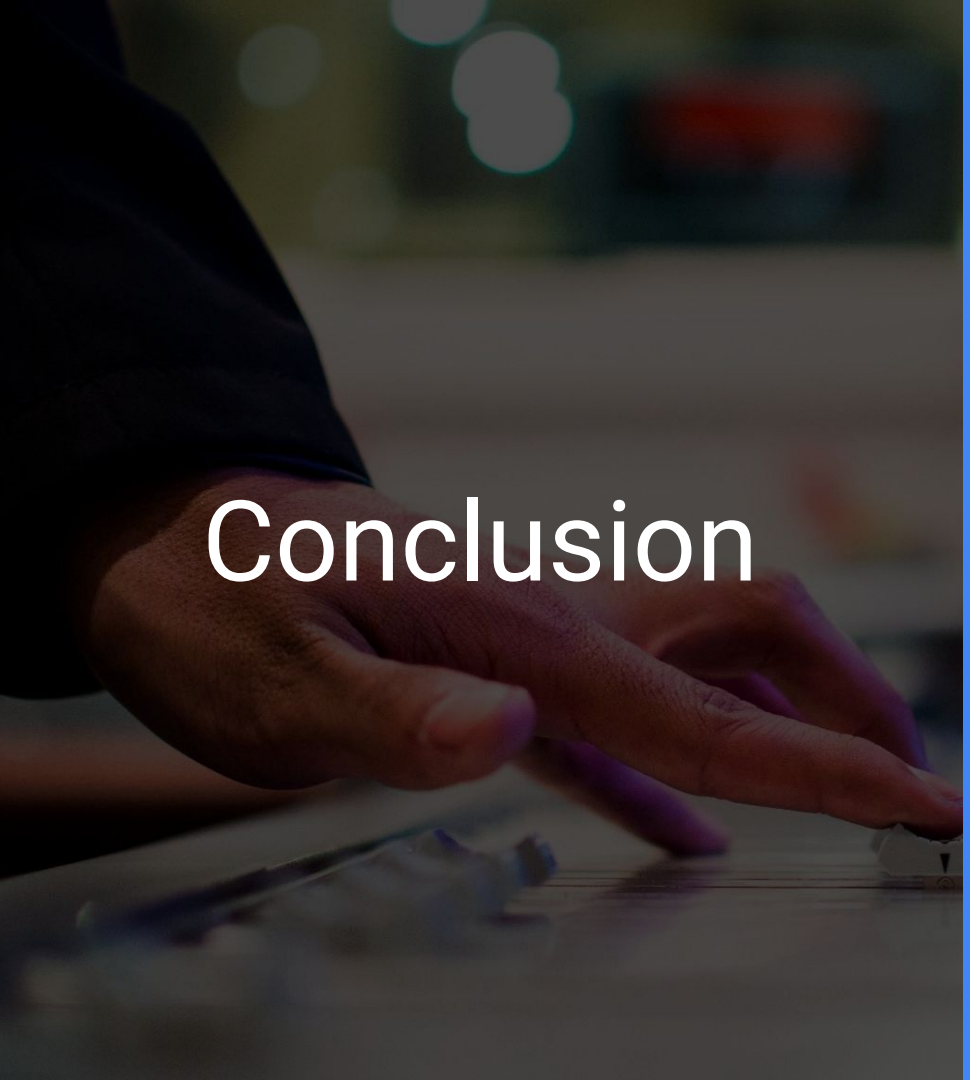
It is the best one to fit our demands.



Conclusion

Spirit Lake is the top city with the highest revenu with 95849$ estimated Population and a rent average of 750$ that could affect approximately 9% of average annual revenue.

Pubs are known as the 3rd Most Common Venue in the city. This city knows big popularity with the diversity of venues categories
Spirit Lake is the best place to open '***Maclaren's Pub***'.



Conclusion

Spirit Lake is the top city with the highest revenu with 95849$ estimated Population and a rent average of 750$ that could affect approximately 9% of average annual revenue.

Pubs are known as the 3rd Most Common Venue in the city. This city knows big popularity with the diversity of venues categories

Spirit Lake is the best place to open ‘Maclaren's Pub’.