**Battle of the Neighborhoods – Gyms vs Crimes**

**Introduction**

**1.) Background**

Health should be looked at as a big asset to a person. Gyms encourage people to stay healthy and eat healthy in order to maintain an overall well-being. In Las Vegas gyms used to be focused on pushing a person to their limits. Boxing gyms used to be popular within mob communities and these gyms would encourage young boxers to be the best. Recently gyms have gone away from pushing the limits only to having social hour while focusing on health. Although many people can work out at home, having a community to push the limit is highly encouraged.

Looking at Southern Nevada health district data, it is found that around 25% of Las Vegas locals are obese. This is highly alarming as this means that people in Las Vegas are not only overweight but obese.

**2.) Business Problem**

Americans have become more obese over the year, having an abundance of food available wherever possible. Las Vegas is not an outlier, as this city has a fast food joint in every corner. There should be gyms in every corner as well. The problem with gyms is that they are expensive to maintain and not having many patrons can cause the owners to go negative on their accounts.

Most gyms in Las Vegas were built in the late 80s and early 90s and most of these gyms have criminal activities in these areas which prevent people from going to these gyms at these times. Although having more equipment would be necessary to have a properly operated gym, understanding crimes in the area is also encouraged as people can have their car stolen, or get robbed while going to the gym at 5 in the morning.

We will try to find out where the best place would be to run and operate a brand-new gym, a place where crime is minimal.

**Data Sources & Cleaning**

**1.) Data Sources**

The data sources that will be used for the project are from data.world and Foursquare. Crimes data was gathered from data.world, the name being Las Vegas Metropolitan calls for service. While the gym data was gathered from the Foursquare API.

**2.) Cleaning**

Data cleaning started by importing the datafile, which was imported into IBM cloud, from there the crimes data was available as a data frame. Within the data frame, there were missing values, these were dropped with axis = 0. The date was in string form so we changed that into datetime, this would be useful for linear analysis, yet in this data we limited to only 1000 crimes that were committed most recent, this information was gathered mostly in June of 2018. We then transformed the crime types into dummy variables this information will be used for clustering later on. We rename some fields in the data and also transform certain criteria such as “Location” into two columns and turn them into float64 type.

For the foursquare data we used a search query that brings only data from gyms. This information was then plotted.

**Methodology**

**1.) Data Analysis**

The data analysis that was used was a pie chart which explains the percentage of crimes that are most common Las Vegas. This data was gathered from the 1000 most recent crimes. We have the ability to see that “Other Disturbances” is the most common type of crime, which is around 38.4 of the data. We will later see that this data will be in its own cluster.

**A picture containing umbrella

Description automatically generated**

**A close up of a map

Description automatically generated**

This map shows us the crimes vs the gyms that are available, only 50 gyms are present in Foursquare, so they are mostly in the center of Las Vegas. We can see that the redder the clusters get the more amount of crimes are within this area. The gyms are surrounded by crimes in this case.

**2.) Machine Learning**

For the machine learning we decided to use KMeans Clustering, as this would allow us to view the data in clusters and have an understanding on how these crimes are related. A k-cluster of 7 was used to give us insights of the data. The most crowded cluster was cluster 2 with 384 values, this number seems familiar as it is the same amount in the pie chart where “Other Disturbances” take up majority. The least filled cluster is 6, where it shows the auto crimes. This is either “RECOVERED AUTO VEHICLE” or “AUTO BURGLARY”. These are the light blue spots in the map.

**A close up of a map

Description automatically generated**

**Results**

**1.) Discussion**

From the map above and the analysis we can see that the purple spots or “Other Disturbances” are most often found within the crimes data. This can give us an understanding of how we can approach this if we were to build a gym. Gyms that are moving towards indoor and outdoor activities such as CrossFit gyms should focus on staying further away from the red and orange areas, as these are full of assault/battery crimes.

**Conclusions**

Crimes happen everywhere, and there is never going to be a place where there are no wrongs being committed. Understanding where crimes happen more often can relieve the feeling of anxious and stress. Location is everything and having a populated area will always come with crimes, we just need to find areas that are not so heavy with crimes. The area around the east side of the map shows little crime activity and would be the best area to have and build a gym. The area on the north side of the map is not covered as that is considered a different city, City of North Las Vegas. My conclusion will be in the far east, or southwest of the map.