

# Gym Planner

Coursera Applied Data Science Capstone  
John Lunn 4/13/20

# Introduction

In 2018, there were 62.5 million gym members in the United States. From 2000 to 2017, gym membership rates have experienced steady growth over the past several years. In 2017, there were 38,477 gyms in the U.S. The types of gyms that people attend evolve based on fashion as well as research around the benefits of different exercises as well as the detriments. The most popular recent Gym types are Crossfit, Orange Theory, Stationary Cycles, Pilates, and Yoga has got increasingly more popular. The Bay Area and California are considered some of the fittest areas in the US.

# The Business Problem

Owning a Gym franchise is a good way for people who are passionate about fitness to combine their hobby with their employment. However, the Bay Area in California has many competitive gyms and which of the current fitness trends is suitable for which area. Can we use publicly available census data to predict where and what type of gym to open?

A prospective Gym franchise owner would be interested in this research in order to help make the difficult decision on where to open a gym. Franchise brands like “Orange Theory” or “SoulCycle” could also be interested in the outcomes of this study.

# The Data

We will use data for the cities in the Bay Area as so:

Geo data to Describe the Bay Area

- [https://en.wikipedia.org/wiki/List\\_of\\_cities\\_and\\_towns\\_in\\_the\\_San\\_Francisco\\_Bay\\_Area](https://en.wikipedia.org/wiki/List_of_cities_and_towns_in_the_San_Francisco_Bay_Area)
- <https://public.opendatasoft.com/explore/?sort=modified>

Census Data:

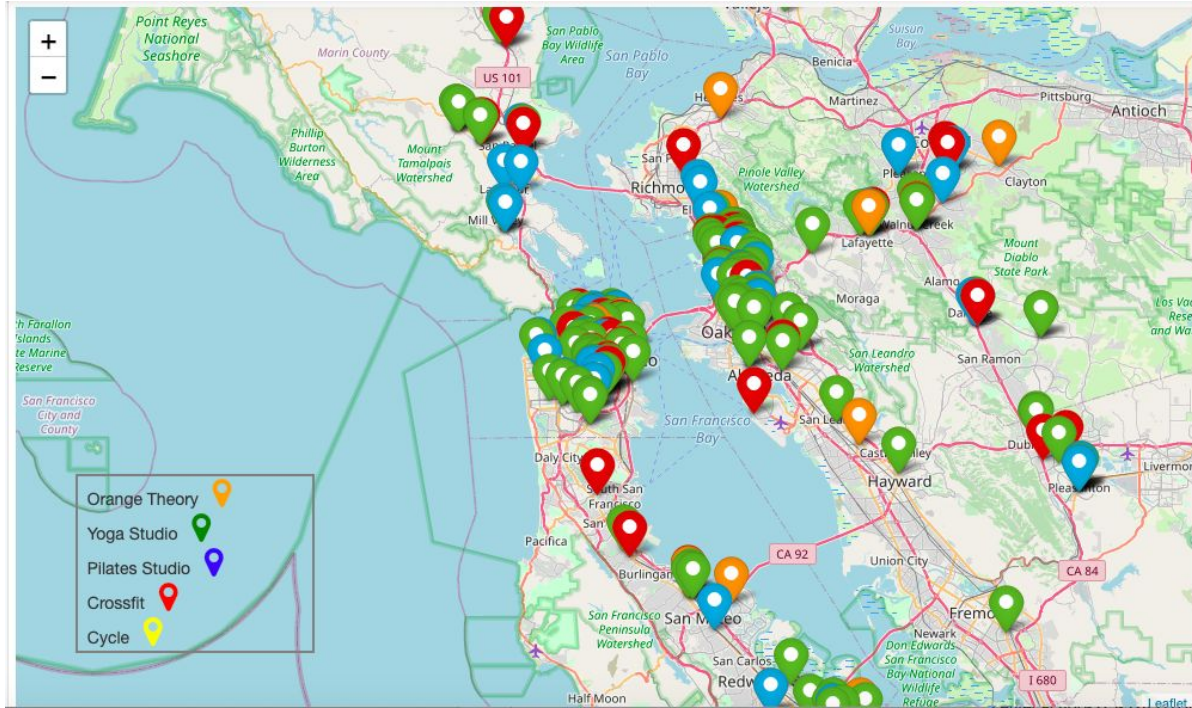
<https://censusreporter.org/>

Gym Data:

- Foursquare to get information about existing locations of brands and Gyms

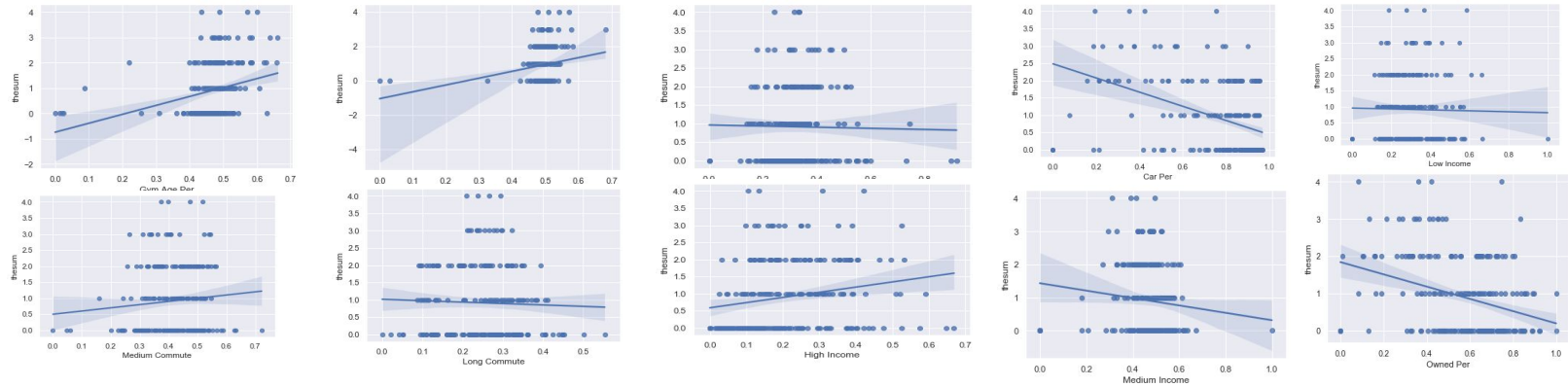
<https://censusreporter.org/>

# Gym Distribution in the Bay Area



The Distribution of our 4 chosen Categories of Gym are shown on this Map.

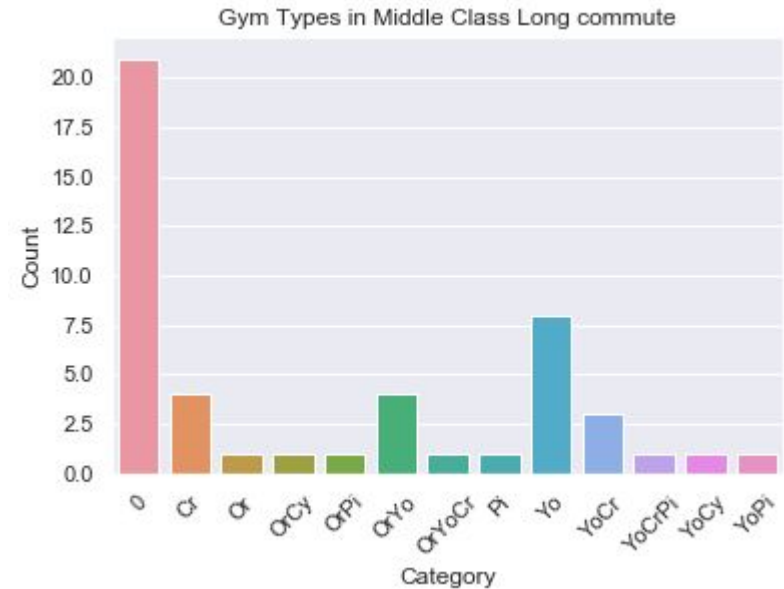
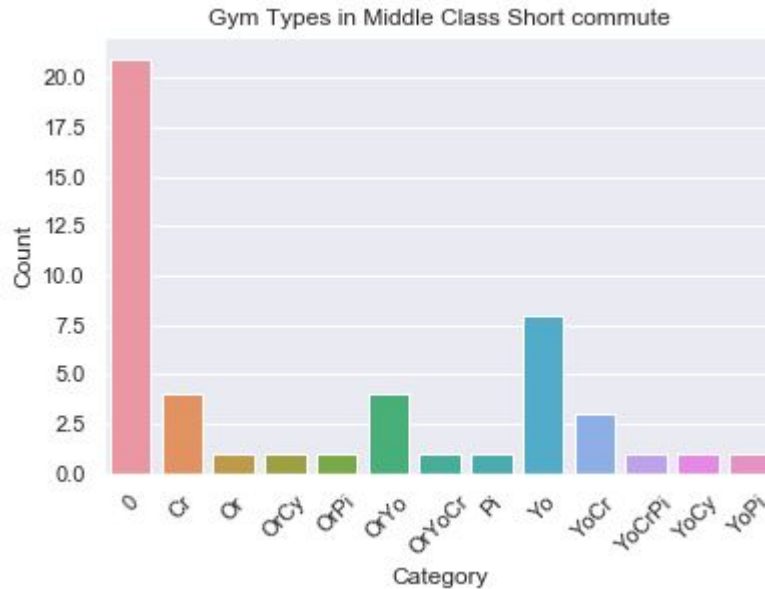
# Feature Selection



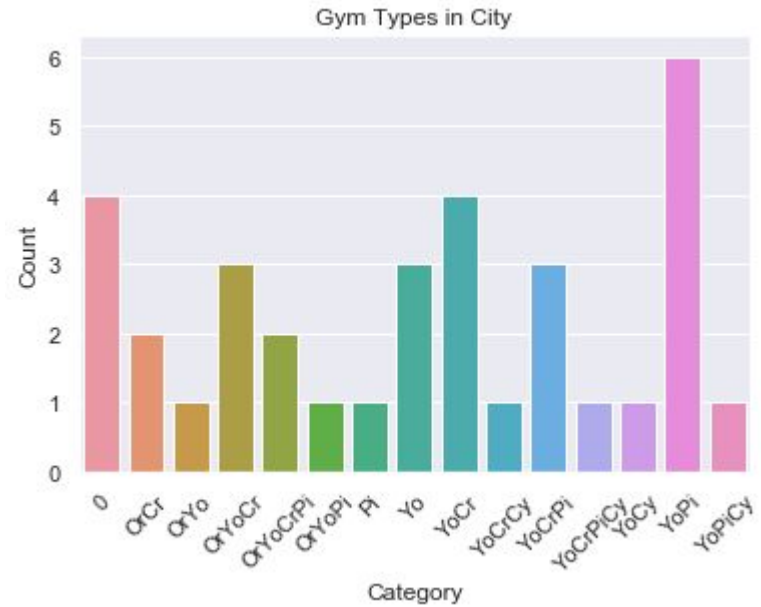
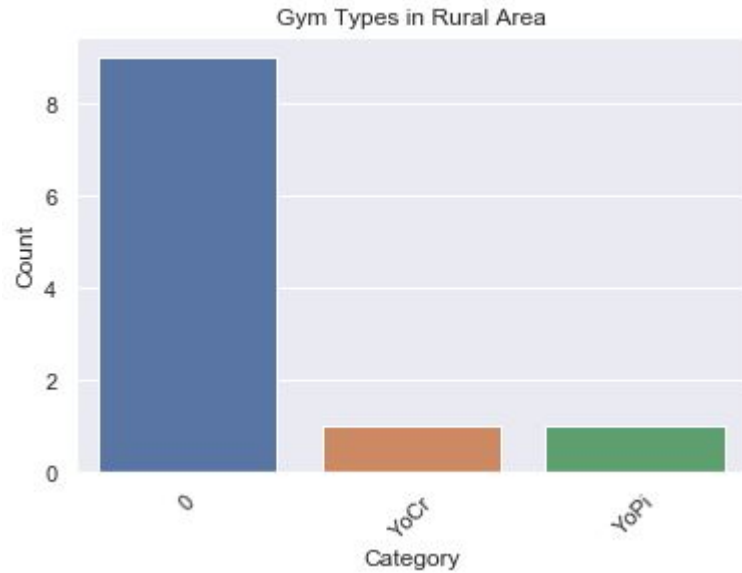
Features from the Census report were selected based on their influence on Gym presence. These included features such as % of people of Gym Age, Mean commute, Income Range and Own vs Rent Ratio

# Data Clustering

K Mean Clustering divided the Data into 6 Clusters and their Gym Distribution was examined



# Data Clustering



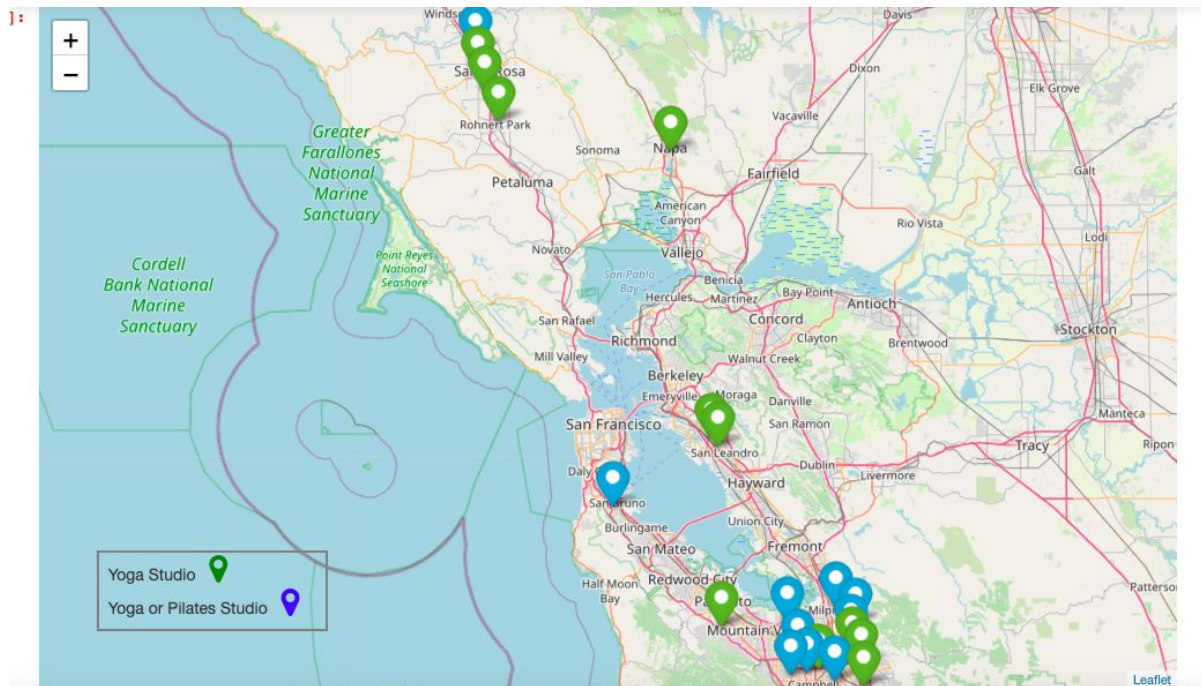


# Where should you put a Gym?

For a person deciding to set up a Gym in the bay area then the order of preference of where they should set up is as follows.

- 1) **The Middle-Class low commute:** the area has the most opportunity with the lowest risk, Over half already have gyms and this leaves half of the area as potential venues for gyms with the right demographics
- 2) **Middle-Class long commute:** There are plenty of no gym zips but the large % of empty zips suggests demand is not there possibly caused by workers working out at work, or at gyms near work, or not at all due to exhausting commute times
- 3) **Upper-Class areas:** If you were to set up a Gym in these areas you would be best off choosing to build a Yoga or Pilates Studio.
- 4) **City:** There is over-saturation in the city of Gyms and smaller Zip codes mean it's not hard to travel. The one exception is the upcoming area of Oakland by the Bay Bridge which has the right demographics but no Gym yet.
- 5) **Low Income:** The cost of trending Gyms is too high for these areas which is backed by the data.

# Helping decide what type of Gym in Cluster 2 Zip codes



The most accurate model still only provides an accuracy of 0.31428 but shows Yoga and Pilates are your best Choice.

# Conclusion

The results of this study can show is that using census data, some areas are more likely to succeed as others as locations for Gyms and that you are better off choosing middle-class non-Large City areas as the place to locate such a Gym. In so much as using the census data to build a model to help you chose a gym type, we have a very low accuracy model. However, the model has determined the most likely to succeed would be a Yoga or Pilates Gym. These gym types have the lowest starting capital investment in that very little equipment is required to start this type of gyms and all that really is needed is floor space. Therefore as a conclusion as a prospective gym owner in the Bay area you should pick a middle-class low commute area and set up a Yoga/Pilates studio.