

IBM Coursera Applied Data Science

Capstone

Tom Thomas

The Battle of the Neighborhoods

Introduction

The US is notorious for being one among the highest ranked countries in terms of obesity rates. The highly commercialized fast-food industry makes it difficult for the average American to stay healthy, especially in large cities, where advertisements are more prominent. This report explores how one could tackle the temptations of living in a large metropolitan city and remain healthy.

For this analysis, the aim is to focus on the two main factors at hand when it comes to an individual who may be attempting to stay healthy - fast-food restaurants and fitness centers. We will analyze the data to observe any patterns between these two types. Whether there are more fitness centers where there are fast-food restaurants or the opposite? And if there can be anything that can be done to reduce the desire to visit a fast-food chain restaurant.

Data

For a thorough analysis, we require the following data:

1. **List of Neighborhoods in a Metropolitan City** - We have chosen Los Angeles for this project, which is a large city with a lot of fast food options and fitness options allowing for a more concrete observation. We will collect this information from a project published by the LA times.

2. **Department of Public Health Data for LA County** - This dataset will be useful for evaluating correlation between obesity rates and the two main factors. This data will be sourced from the LA County Public Health Records.
3. **Geospatial Data using Geocoder Package** - We will use this package to capture the required Latitudes and Longitudes for LA County.
4. **GeoJSON Data for LA County** - This will be the data we use to create choropleth maps for Los Angeles.
5. **FourSquare Venue Data** - We will utilize the FourSquare API to collect venue information for the different neighborhoods.