

# **Applied Data Science Capstone Project -The Battle of Neighborhoods**

## **New Restaurant Site Evaluation Recommendations**

Michael Hall  
mjhall@ucsd.edu  
December, 2018

## DATA

Data will be gathered from multiple sources to evaluate the communities. Data on the competing restaurants and nearby venues will be obtained from Foursquare. The community data will be obtained from San Diego Association of Governments (SANDAG). SANDAG is made up the 18 cities and unincorporated county governments. SANDAG develops annual demographic estimates and long range forecasts in addition to maintaining information from the U.S. Census Bureau. SANDAG will be providing the 2010 US Census data summarized for each community. SANDAG will also be providing the ESRI shapefile data that describes each of the communities. The ogr2ogr web client (OGRE - <https://ogre.adc4gis.com/> ) will be used to convert the ESRI shapefile data to JSON format.

Title:	Community Venues
Description:	Venue name, location, category and other firmographic information
Use:	Identify competitors and venues within a community
Format:	JSON
Source:	<a href="https://www.foursquare.com">https://www.foursquare.com</a>

Title:	Community Demographics
Description:	2010 US Census demographics aggregated and summarized by community
Use:	Profile of the population and other demographics of the community
Format:	Excel
Source:	<a href="http://datasurfer.sandag.org/">http://datasurfer.sandag.org/</a>

Title:	Community Location
Description:	Boundaries and centroids of the communities
Use:	Plotting the data on maps
Format:	ESRI Shapefile
Source:	<a href="http://rdw.sandag.org/">http://rdw.sandag.org/</a>