Opening a Fitness Center Based on Regional Health Patterns

1. **Introduction/Business Problem**

The purpose of this report is to utilize freely available data to determine the optimal location for opening a new fitness center. In this case, the optimal location would be a highly unhealthy region of the United States where there are little to no fitness centers within a reasonable traveling distance. In 2015-2016, about 39.8% of adults in America were obese[1], while about 18.5% of American children ages 2-19 years old were obese[2]. About 9.4% of the US population has been diagnosed with diabetes[3]. Since this data was sampled and presented, these numbers have continued to rise every year. Opening a fitness center in a highly unhealthy location where the most people would benefit from it could help slowly reverse this trend. If successful, the process could be repeated for more unhealthy areas, as to benefit as many people as possible.

The stakeholders in this project consist of unhealthy people, which for the purposes of this report, will be Americans who are obese and/or have diabetes. This can also be extended to their families and friends, as they do not want to have any unhealthy loved ones. This could also possibly extend to medical offices and hospitals, as a decrease in unhealthy Americans would likely lead to a little less stress for them.

Based on this information, the question being asked here is: *Can we use available data to determine optimal locations for opening new fitness centers?*

2. **Data**

The data being used for this report comes from FourSquare.com, and the USDA ERS (United States Department of Agriculture Economic Research Service). The USDA ERS hosts a dataset[4] on their website called the Food Environment Atlas, which details data on demographics (such as age and race), access to nearby grocery stores and fast food restaurants, SNAP benefits, health/physical activity rates, income level, and much more. The main set of data from this dataset that will be used here is the data in the Health section; more specifically, information on physical activity rates, and adult obesity/diabetes rates from 2008 and 2013. Other data from the dataset will be used as well for comparison and analysis at times.

FourSquare.com is an online platform where users can search for different types of venues in their current location, such as different restaurants, shopping locations, and entertainment options. The FourSquare API will be leveraged in order to pull and visualize this data; specifically, data regarding nearby fitness centers will be utilized to help determine optimal locations for opening new fitness centers.

A. References

[1] Adult Obesity Facts, Centers for Disease Control and Prevention. August 13, 2018. Retrieved from https://www.cdc.gov/obesity/data/adult.html

[2] Childhood Obesity Facts, Centers for Disease Control and Prevention. August 13, 2018. Retrieved from https://www.cdc.gov/obesity/data/childhood.html

[3] National Diabetes Statistics Report, Centers for Disease Control and Prevention. February 24, 2018. Retrieved from https://www.cdc.gov/diabetes/data/statistics/statistics-report.html

[4] Data Access and Documentation Downloads, United States Department of Agriculture Economic Research Service. March 27, 2018. Retrieved from https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads/