Coursera Capstone Project

(IBM Applied Data Science)

New Urgent Care Center in Boston, MA

Mrinal Shukla

November 2019

Introduction

The cost of ER care is increasing everyday and becoming unaffordable. As a result more and more Urgent Care Centers are coming up. A 2016 study in the Annals of Emergency Medicine found that ER treatment costs were about 10 times more (an average of about $2,200) than in an urgent care center (about $168) — even for patients with the same diagnosis. Also, the wait time is much less in Urgent Care centers than in ER rooms. The UCA's 2018 Benchmarking Report found that more than 70% of patients waited less than 20 minutes to see a provider at an urgent care center, and nearly 94% were seen in less than 30 minutes. Overall, 85% of urgent care centers patients are taken care of in under an hour. There has been a great increase in the number of Urgent Care Centers. The Urgent Care companies are looking for best possible venues to open more.

Business Problem

This study explores the neighborhoods of Boston to find the best possible locations for Urgent Care Centers. It uses Machine Learning Methodologies like Clustering and Segmentation to find the neighborhoods which would be best suited for opening Urgent Care Centers. This project aims to answer the business question: What locations are best suited for opening an Urgent Care Center in the city of Boston.

Target Audience

This project will be useful for Urgent Care companies who are looking to open new Urgent Care Centers in Boston. This project is time appropriate as the demand for more and more Urgent Care Centers is on the rise due to increasing ER care costs. According to Consumer Reports, the number of urgent care facilities increased from 6,400 in 2014 to 8,100 in 2018, with another 500 to 600 expected to open.

Data

To solve the problem we would need the following data:

- List of all the neighborhoods of Boston. This will define the scope of the project, which is limited to the city of Boston.
- Latitudes and Longitudes of all the neighborhoods of Boston. We will use this data to plot maps and get venue data.
- Venue data, particularly data related to Urgent Care Centers. We will use this data to perform segmentation and clustering.

Data Sources and Extraction Methods

The list of neighborhoods of Boston is obtained from Wikipedia (https://en.wikipedia.org/wiki/Neighborhoods_in_Boston) it has 22 neighborhoods. We get the data by scraping the webpage using the Requests and BeautifulSoup packages of Python.

We get the Longitude and Latitude of the neighborhoods by using the Geocoder package of Python. The venue data is retrieved using the FourSquare API. This returns various types of venues but we are only interested in Urgent Care Centers so we run a query for only these venues.

After retrieving the data we perform various Data Science methodologies like Data Cleaning, Data Wrangling, Visualization, Segmentation and Clustering to analyze the data and find the answer to our business problem: What is the best location to open a new Urgent Care Center in Boston. In the following sections we will elaborate on the Methodologies used and the conclusion from the analysis.