

Predicting the Concentration of Starbucks Locations using Socioeconomic Data

Prepared by:

Trevor Stull

February 2020

Coursera Capstone Project

Table of Contents

Table of Contents	2
List of Figures.....	2
List of Tables	2
List of Appendices	2
1.0 Introduction/Background	3
1.1 Background	3
1.2 Problem Statement	3
1.3 Interest	3
2.0 Data Acquisition and Cleaning	3
2.1 Data Sources.....	3

List of Figures

No table of figures entries found.

List of Tables

No table of figures entries found.

List of Appendices

Appendix A	[Insert Title]
Appendix B	[Insert Title]
Appendix C	[Insert Title]

1.0 Introduction/Background

1.1 Background

The Starbucks Coffee Company was founded in 1971 with one coffee shop in Seattle, Washington. Since those humble beginnings, the company has grown to an international company with coffee shops open across the world. Like with any company, knowing what customer the product is trying to reach is important to Starbucks' success and growth. It is also paramount for a company to maximize its resources and ensure the company is efficiently using them. Keeping this in mind, it would be worthwhile to predict the number of Starbucks stores needed in an area based on its population's socioeconomic makeup. This would allow the company to use its resources more efficiently by either adding store locations in areas predicted to need more stores than current store locations or eliminate store locations in areas that are predicted not to need the current number of stores.

1.2 Problem Statement

Data that might contribute to the concentration of Starbucks in an area might be the area's population, average age, median household income, or race/ethnicity. This project will aim to try and quantify these relations and develop an expression to predict the number of Starbucks in an area.

1.3 Interest

This information could be used by the Starbucks Coffee Company to determine the optimum number of Starbucks stores needed in a location or could be leveraged to determine the typical Starbucks customer in America.

2.0 Data Acquisition and Cleaning

2.1 Data Sources

Foursquare data will be leveraged to gather Starbucks location information throughout Dallas and Tarrant counties in North Texas. Socioeconomic factors such as population counts, average age, race/ethnicity, and median household income will be downloaded from US Census data from American Fact Finder and Integrated Public Use Microdata Series (IPUMS) National Historical Geographic Information System (NHGIS) service.