

IBM Data Science Capstone
Robert Webster

BATTLE OF THE NEIGHBORHOODS

NEW YORK CITY
FY 2020

This report discusses the challenges of finding a good neighborhood to move to and start a business considering living expenses, availability of venues that are preferred, etc., and offers a solution by using modern Data Science Tools.

TABLE OF CONTENTS

Contents

Introduction	1
Data	Error! Bookmark not defined.
Methodology	Error! Bookmark not defined.
Results	Error! Bookmark not defined.
Discussion	Error! Bookmark not defined.
Conclusion	Error! Bookmark not defined.



Image courtesy of <https://www.goodfreephotos.com/united-states/new-york/new-york-city/face-of-the-statue-of-liberty-in-new-york.jpg.php>

BATTLE OF THE NEIGHBORHOODS

Introduction

BUSINESS PROBLEM

I want to start living in New York City (NYC), but I don't know what neighborhood to live in. I love to eat Chinese food so I want to live in a neighborhood that is small, but has a high density of Chinese food restaurants. Eventually, if the neighborhood is well-suited, I may consider opening my own restaurant – I want to open mine in a community of similar minded eaters. (If you want good Chinese food – go to China Town!)

TARGET AUDIENCE / STAKEHOLDERS

This analysis is a good project for data scientists seeking to explore available open-source tools. It is also a useful tool for someone who wants to move to NYC! The simple problem statement above is actually quite complex to resolve, but easily done with the correct tools and code

BACKGROUND

NYC is exciting and interesting and I like tall buildings. I enjoy the cultural diversity. I like to eat a lot of food, so I want there to be lots options to eat near-by. Specifically, I like to eat Chinese food.

Unfortunately (or fortunately – depending on your perspective) NYC is a gigantic city with many neighborhoods and a significant number of venues to choose from. This project would be nearly impossible to do by flipping through a phone book, or trying to search manually though google.

This project seeks to utilize the advantages of modern-day data science tools to solve the problem discussed above.