

# Median Income and Venue Analysis of New York City

Zachary Avant

June 3, 2021

## 1. Introduction

New York City is famously competitive, and any new business has to be aware not only of what businesses are already in place, but also the population density and median income of a neighborhood's residents. By clustering New York zip codes based on nearby venues, median income, and density, we can find regions that share distinct characteristics. This information can be useful not only for people looking to start a business, but to residents or immigrants looking to move, or tourists trying to make the most of their visit.

## 2. Data

Data has been gathered from Citizens' Commission for Children of New York's [median income dataset](#), as well as BetaNYC's [neighborhood dataset](#). These were joined along zip codes, and together provide population densities, median income, and other useful features. Missing numeric data was replaced with the mean value of its column. Zip codes were utilized to find geospatial data, which was put into [Foursquare](#)'s developer API to find neighboring venues.