A. Introduction

As someone who has studied, worked, and lived in Manhattan for the past twelve years, moving away due to COVID-19 has been a big turning point in my life. Alongside this change in location is the decision to shift my career towards data science, so it was only natural for me to focus my capstone project for IBM Data Science Professional Certificate on the spread of COVID-19 in Manhattan.

As one of the five boroughs of New York City, Manhattan is home to roughly 1.63 million people all living in just 23 square miles. This translates to a population density of 70,826 people per square mile, or 27,346 per square kilometer, which makes Manhattan or New York County the most densely populated county in the United States. Sadly, this statistic is also one of the key reasons why Manhattan was hit hard by COVID-19 in the first half of 2020.

Beyond just residents of Manhattan, during a typical workday, commuters would push these figures to 3.9 million people or a population density of over 170,000 people per square mile. Combine these office hotspots with tourist hotspots like Time Squares and Central Park and cultural hotspots like Lincoln Center and Broadway, it is pretty clear Manhattan is particularly vulnerable to infectious diseases such as COVID-19.

What is unclear or left unexplored is how different venues within smaller neighborhoods inside Manhattan contributed or impacted the spread of COVID-19. So in this project, we will utilize the Foursquare API to first breakdown the neighborhoods of Manhattan by their venue types and then cross reference this data with COVID-19 data from NYC Department of Health and Mental Hygiene.

Ultimately, this project will utilize the data science techniques I learned in the course to shed light on the transmission of COVID-19 in cities like Manhattan and hopefully pass on insight to help other cities to take the right measures to curb the effects of COVID-19 on our livelihood. Lastly, I hope to continue to develop my data science skills and expand on this project beyond the scope of this project in order to help alleviate the impact of future infectious diseases both in terms of disease prevention and city planning.