1. Introduction

1.1 Background
When most Individuals relocate to a new city in search of a better job or a better life, they seldom think about the location of where to live in a city,generally the choice is left to immediate availability and convenience but Its not the best approach in selecting a neighbourhood to live in.Everyone chooses a location based on proximity to their workplace or they chose a place that is more suited to their budget but it's not easy to identify that without having prior knowledge regarding a place. To minimize the risk and errors we will be looking into proper research before making a decision on where to start searching for good housing, as making a decision like this can't be changed that easily and it involves a lot of work, time and money hence its important that the right decision is made in the first attempt in picking a house in the ideal neighbourhood based on one's requirement

1.2 Problem
The main issues that will be tackled in this project are the cost of housing and the safety of its neighbourhood,both these are determined using the housing sales data from Kaggle and NYC crime data from the cities database hosted online. The goal is to identify a borough and pick the best neighbourhood from that borough that has relatively low housing cost and is a safe place for a family, and implement k-means clustering to group them using the Foursquare data for venues in respective neighbourhoods.

1.3 Interests
Individuals who are trying to determine the cost and safety of the neighbourhoods in NY and the proximity of venues can utilize this report to make an informed decision before chosing a home in the city and it also provides information on the ideal cost of a neighbourhood for any business trying to open a new outlet in a particular neighbourhood.