**NYC data Property Analysis -Final capstone Project**

**Introduction / Business Problem**

New York City (NYC) is one of the most populated cities in USA and second most populous city in North America. Its is comprised of 5 boroughs namely Brooklyn, Manhattan, Staten Island, Bronx and Queens. The 5 Boroughs cover overall land area of about 784 Km square with a population of about roughly 8,398,748 in year 2018. As a Neighbor of New York City, I have chosen this location for my capstone project.

New York City is culturally very diverse with population density of 159 people per square Km and described as financial, and media capital of the world and it’s a center for commerce entertainment, research, technology, education, politics, tourism, art, fashion, and sports.

NYC is one of the largest metropolitan cities with over 20 million people and home to headquarters of United Nations. There are more than 100 neighborhoods divided among 5 boroughs with Manhattan titled the most expensive real estate Markets. New York city is most powerful city economically and financially, it is also home to largest stock exchanges the NASDAQ and New York Stock Exchange.

As we can see from the statistics NYC is a very diverse and financial capital, we can derive many ideas and problems like: if I am looking to open a restaurant or business, I would like to explore neighborhoods /areas with low real estate property values? If some is looking for office / house to rent which area should they prefer and why?

With help of foursquare location data and raw data (NYC property data) and other tools we explore further to cluster based on borough information and venue data obtained using foursquare and come up with a solution to some of the problems mentioned above.

**Reference:**

<https://en.wikipedia.org/wiki/New_York_City>

Note: some of the statistical information (population info) has been taken from the link above.

**Data Section**

1.<https://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page>

2. Foursquare location data will also be used.

Data consists of rolling property sales data for all 5 boroughs and information about taxes, type of property, neighborhood, date of sale, square footage info etc. The has been corresponds to 12-month period (year 2018). The source had sales data recorded per each Borough. I have consolidated the data in one data source.

Description:

The data contain property sales data across all 5 boroughs

Manhattan (1), Brooklyn (3), Staten Island (5), Bronx (2), Queens (4)

Neighborhood info: Name of the Neighborhood where the property dwells

Building Class Category: 01 ONE FAMILY DWELLINGS, 21 OFFICE BUILDINGs, COMMERICAL CONDOS etc. (There are about 44 Categories)

Tax class at Present: There 3 to 4 diff tax classes applied based on building class category

BOROUGH

NEIGHBORHOOD

BUILDING CLASS CATEGORY

TAX CLASS AT PRESENT

BLOCK

LOT

EASE-MENT

BUILDING CLASS AT PRESENT

ADDRESS

APARTMENT NUMBER

ZIP CODE

RESIDENTIAL UNITS

COMMERCIAL UNITS

TOTAL UNITS

LAND SQUARE FEET

GROSS SQUARE FEET

YEAR BUILT

TAX CLASS AT TIME OF SALE

BUILDING CLASS AT TIME OF SALE

SALE PRICE

SALE DATE