

IBM Data Science Capstone Project, The Battle of New York's Boroughs

1. Objective. New York's Boroughs analysis and comparison.

The objective of the project is to compare and establish differences, if they exist, between the five boroughs that form New York, the Bronx, Manhattan, Brooklyn, Queens, and Staten Island, focussing on the distribution and clustering of venues.

Background.

Urban development and the distribution of important city landmarks reveal the history of the city and its social, cultural, and economic development, and from its current composition it's possible to define differences within and between its geographical and political divisions. From its conception as a world capital, the unique composition of New York's boroughs presents interesting insights into multiculturalism and approach to urban development, and potential clues to its future and priorities.

This type of analysis may be helpful to private enterprises if they want to establish the basis for a more specific and complete analysis on the possibilities for new franchises, a new brand, or service. The public sector can also gain insights into the commercial, and cultural distribution of the city, and define long term projects of urban development in function of its priorities and budget constraints.

2. Data.

The data that will be used to analyse New York's boroughs comes from the Spatial Data Repository, that presents the geographical distribution of New York's neighborhoods as well as its political divisions. It will help to establish the base of the geographical distribution of New York, and of the data that will be built on top of it.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

The second data set to be used is the venues information provided by Foursquare, presenting the geographical description of the existent venues in New York City. With it, we will be able to define the

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
3	Wakefield	40.894705	-73.847201	Shell	40.894187	-73.845862	Gas Station
4	Wakefield	40.894705	-73.847201	Cooler Runnings Jamaican Restaurant Inc	40.898083	-73.850259	Caribbean Restaurant

characteristics of the neighborhoods and boroughs of the city according to the frequency and type of establishment that it presents

3. Data sources.

1. [2014 New York City Neighborhood Names](#). New York Spatial Data Repository.
2. [2010 New York City Boroughs](#). New York Spatial Data Repository.