SEPTEMBER 2, 2018

# THE BATTLE OF NEIGHBORHOODS REPORT

## COURSERA – CAPSTONE PROJECT

TIMBAKE DANG

# TABLE OF CONTENTS

**BUSINESS PROBLEM ..... 3**
**Executive Summary ..... 3**
**DATA - STATISTICS & REPORTS ..... 4**
**Top 20 Venues Recommended By Category in city, neighborhoods ..... 4**
**Top 5 venues have the most interesting (likes) by category in city, neighborhoods ..... 4**
**Top 5 venues have the highest rating by category in city, neighborhoods ..... 4**
**Top 20 venues have the most visits, users by category in city, neighborhoods ..... 4**
**Top 10 categories have the most visits, users by city, neighborhoods ..... 5**

# BUSINESS PROBLEM

## EXECUTIVE SUMMARY

There are many existing venues in neighborhoods in New York as well as Toronto city. How can we know which venue by category and city have the most interesting from people as well as have the most number of users checked-in? Then how can we know the competition of the venues of the same category in a city or between neighborhoods in a city bases on the statistic number of users, visits, checked-ins, likes, rating?

Through by answering this questions, we can have insight about the competition between venues of the same category in a city or in a neighborhood, then we can find out the reasons could make these venues to be attractive to the people.

Finally, after we know reasons and have the accurate statistic data about the existing competitors, the behavior of users, the user culture, the demands are not served or are not served well, the needs in future, we will develop strategies for changes in the market, sustainable growth as well as expansion business to new market.

# DATA - STATISTICS & REPORTS

The data used in these statistics and reports was resources retrieved from:

- Toronto data: Wikipedia (Toronto postal code) combined with neighborhoods, venues data by using Foursquare APIs.

- New York data: included neighborhoods, venues data gathered from https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json

Before using Foursquare APIs to get neighborhoods, venues data, I used Geocoder library to get coordinates (latitude, longitude) for each neighborhood, instead of Google Geoencoding APIs.

Then I combined 2 data frames Toronto and New York together in order to perform comparison, statistic and create reports.

## TOP 20 VENUES RECOMMENDED BY CATEGORY IN CITY, NEIGHBORHOODS

## TOP 5 VENUES HAVE THE MOST INTERESTING (LIKES) BY CATEGORY IN CITY, NEIGHBORHOODS

## TOP 5 VENUES HAVE THE HIGHEST RATING BY CATEGORY IN CITY, NEIGHBORHOODS

## TOP 20 VENUES HAVE THE MOST VISITS, USERS BY CATEGORY IN CITY, NEIGHBORHOODS

# TOP 10 CATEGORIES HAVE THE MOST VISITS, USERS BY CITY, NEIGHBORHOODS