

## First deliverable – Battle of the Neighborhoods

The first deliverable of the capstone project consists in 2 descriptions:

1. A description of the problem and a discussion of the background.
2. A description of the data and how it will be used to solve de problem.

### Description of the problem and a discussion of the background

As per my personal work experience as CRM analyst what I constantly work on doing is better understanding our customers based on self-gathered data (based mainly on surveys amongst other sources). It is because of this fact that I would like to do a demographic analysis in order to hopefully be able to replicate said analysis (or at least use as a guide) a more in depth analysis for the company I work for.

For this specific project – I will look into the City of Toronto (provided I have already gathered and cleaned the data in previous week's deliverables) in order to better understand Foursquare users in Toronto – be it Torontonians, Canadians (not from Toronto), foreigners living in Toronto and tourists. This analysis will relate neighborhoods with specific personality affinities.

This project will (besides help me for future projects regarding my own work) be useful for people who might want to understand better how people in Toronto behave and the needs they have in specific neighborhoods, seeing for example that people in very close neighborhoods go to Cafes whilst having no café in their actual neighborhood – This will help entrepreneurs start new businesses.

**Note:** *Keep in mind that, as I am completely new to Python and lack a whole bunch of practice and experience. For this project's idea I am going with the assumption that only places visited are enough to do a demographic analysis in terms of personality's affinities. I am therefore not taking into account ethnicity, age, gender, sexual orientation, etc., into the research.*

### Description of the data and how it will be used to solve de problem.

I personally don't think that the data needs much explaining provided that I will be using the same information from previous weeks in order to create my own project. It is none the less a bit different in terms of datasets created.

In a very general grasp I will be gathering the data, processing it based on my needs, clustering the data and then conclude with the analysis of the information gathered.

Keep in mind that there are some details or descriptions in the code as to “what I am doing” but the full description on or process will be shown in the final deliverable of the project alongside a presentation with the results/analysis.