

# **Applied Data Science Capstone**

## **The Battle of Neighborhoods**

### **Week 1**

*by: Josemar Tadeu Migowski da S.Carvalho  
October 2019*

## Introduction

The purpose of this project is to identify the commercial neighborhood of residential areas in the city of Brasilia (DF) - Brazil, in order to support decision making for investors interested in starting commercial activities or services to meet the current and future demands of each region.

The work began by obtaining shapefiles from the Federal District state containing the geometry of the neighborhoods in order to allow specifying the target neighborhood of the analyzes to be formulated. These files were obtained from the IBGE (Brazilian Institute of Geography and Statistics) website at the URL <https://portaldemapas.ibge.gov.br/portal.php#mapa201483>.

## Problem:

Investors are always looking for information on the most appropriate types of businesses and places to start a business and minimize the risks involved. For this reason, we decided to present a study that seeks the appropriate information for decision making regarding the type of business and the most appropriate location for its implementation in the Federal District - Brazil region.

This project aims to gather and cross-check information obtained from various sources to facilitate decision making for investors interested in investing in the Federal District - Brazil region.

## Methodology:

### Data Sources:

For the development of this work, several open data sources were consulted in order to identify the pertinent information for the project in question. Among which are:

1. CODEPLAN - Planalto Central Development Company website, where it provides a report containing data on the Urban Density of each Administrative Region of the Federal District. Available at: [http://www.codeplan.df.gov.br/wp-content/uploads/2018/02/TD\\_22\\_Densidades\\_Urbanas\\_nas\\_Regi%C3%B5es\\_Administrativas\\_DF.pdf](http://www.codeplan.df.gov.br/wp-content/uploads/2018/02/TD_22_Densidades_Urbanas_nas_Regi%C3%B5es_Administrativas_DF.pdf);
2. Georeferencing data making it possible to present the results of the work visually, facilitating their comprehension. ShapeFile available at the link: <http://dados.gov.br/dataset/dados-georreferenciados-siturb-segeth/resource/c5898ba4-d1a4-49fb-9075-60f6dda74097>;
3. FourSquare API - The search for information on the types of business existing in the administrative regions of DF was only possible thanks to the availability of the FourSquare API. Through this API, it became possible to perform business-related searches within 3000 meters from the central point of each administrative region. It should be noted that as the use of this API was made for testing purposes only, the results obtained are limited.