A Geo-Analysis

of Charlotte, North Carolina

**Peer-graded Assignment:**

**Capstone Project - The Battle of Neighborhoods**

**(Week 1)**

**This document will be updated next week for**

**Week 2 of the assignment.**

Author: Amit Arya  
Publish Date: March 31, 2020

Contents

[1. Background & Problem Statement 2](#_Toc36643147)

[2. Data Description 3](#_Toc36643148)

[3. Methodology and Exploratory Data Analysis 4](#_Toc36643149)

[4. Results and Discussion 5](#_Toc36643150)

[5. Conclusion 6](#_Toc36643151)

A Geo-Analysis

of Charlotte, North Carolina

Author: Amit Arya  
Publish Date: March 31, 2020

# Background & Problem Statement

I reside in the Charlotte, North Carolina. After living here for 10 years, I moved to Pittsburgh, Pennsylvania for 5 years and have recently moved back to Charlotte about 2 years back.

Charlotte was always a vibrant metropolitan area supported by a family oriented suburban cities. Having come back after 5 years I see that the city has growth exponentially with many families preferring to move to warmer climate with good economic prospects. Charlotte fits this description.

The growth boom is wonderful to watch. While looking for a new home upon relocating back to Charlotte, we looked at the various suburban cities in the Charlotte Metropolitan Area for a home and we found that while you can buy a house in any nearby city, the amenities were not always the best in the cities with newer neighborhoods.

So my goal is to analyze which cities surrounding the Charlotte Metropolitan and solve for 2 problem from 2 different perspective:

1. **Home Seekers**: Assist people in evaluating various surrounding cities to enable them to find the location which has the amenities that they care about
2. **Business Owners:** Assist businesses in evaluating under served market to analyze what business opportunity may exist in a particular location.

# Data Description

The basic data that we need is a list of cities in the Charlotte Metropolitan Area. I performed a quick search on google and found several sites listed below:

1. [Will not be used] <https://www.gps-latitude-longitude.com/> This site requires location of each city to be fed in one at a time. This is very time consuming and therefore not very efficient.
2. [Will not be used] <https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/export/> This is an excellent source for city information and this dataset gave me a list of over 1,000 cities in North Carolina. However, the data was not very rich and I had no way to know which cities were around the Charlotte Metropolitan Area. Hence, I do not intend to use this dataset.
3. [Will not be used] <https://en.wikipedia.org/wiki/Charlotte%2C_North_Carolina> This link has some very rich information but did not have the structured data that I am looking for. Therefore, I will not be using this data source.
4. [**Will be used**] <https://simplemaps.com/data/us-cities> Simple map offers a download of most large cities in the world and I was able to download the list of US cities. I was further able to filter out just the cities in North Carolina. This consists of 736 cities. Since our target is cities surrounding the Charlotte Metropolitan Area, which fall in Mecklenburg county, I am going to further filter on the cities within the Mecklenburg county domain.

A copy of the city level data in North Carolina is available in my github repository. Link is provided below.

<https://raw.githubusercontent.com/amitarya76/Coursera_Capstone/master/NC_CITY_DATA.csv>

The data contains the following fields that I intend to uuse.:

|  |  |  |  |
| --- | --- | --- | --- |
| # | Field Name | Data Value Example | Data Type |
| 1 | city | Trenton | City Name |
| 2 | city\_ascii | Trenton | City Name |
| 3 | state\_id | NC | State Abbreviation |
| 4 | state\_name | North Carolina | State Name |
| 5 | county\_fips | 37103 | FIPS Code |
| 6 | county\_name | Jones | County Name |
| 7 | county\_fips\_all | 37103 | FIPS Code |
| 8 | lat | 35.064 | Latitude Data |
| 9 | lng | -77.3553 | Longitude Data |
| 10 | population | 285 | Population in '000 |
| 11 | density | 484 | Population Density |
| 12 | timezone | America/New\_York | Time Zone |
| 13 | zips | 28585 | Zip Code |

City Name allows for identification of cities. From the above dataset, I intend to use the county as the filter to get the cities in the “MECKLENBURG” county and combine that dataset with Foursquare API data to get venues and places in each city. The population and density data will inform our analysis on high, moderate and low growth areas.

I will use the above dataset from simple maps and use the Foursquare API data to look at venues and locations in the various cities surrounding the Charlotte Metropolitan Area to solve for the 2 problems outlined in the prior section.

# Methodology and Exploratory Data Analysis

To be completed in Week 2 per assignment instructions.

Step 1. Data Import

Step 2. Data Cleaning

Step 3. Data Merging

Step 4. Exploratory Analysis

# Results and Discussion

To be completed in Week 2 per assignment instructions.

# Conclusion

To be completed in Week 2 per assignment instructions.