**BATTLE OF THE NEIGHBORHOODS**

***Schools govern Relocation to New Place- Myth or Reality: Case Study of Charlotte Area, NC***

**Introduction**

*The Battle of the Neighborhoods* is a capstone project for Coursera and IBM’s *Data Science Specialization*. This project encompasses the vision of the Data Scientist to portray the tools learned during 9 courses on *data analysis* for the certification. One of the requirements is to use Foursquare location to explore and solve the neighborhood analysis using clustering method.

**Background**

*“Schools govern Relocation to New Place- Myth or Reality: Case Study of Charlotte Area, NC”* is a project taken to understand how the schools affect the relocation to a new place/ neighborhood. Charlotte area is known for its Banking jobs and the increase of financial institutions making Charlotte and its surrounding areas, their financial headquarter.

The goal of this project is to find a neighborhood for a person relocating with his family to Charlotte, NC. A family man requires an easily accessible neighborhood with some good schools. Certainty and uncertainty of renting or owning a house revolves around a good neighborhood. Estate agents try to their potential to sell houses that are most profitable to them. Personally, having a little insight about the area and its neighborhoods can help the family plan their relocation smoothly.

This study is an attempt to find such a neighborhood for a family with school going children.

**Data Requirement**

Best approach will be to follow a *Top-Down Approach* for getting the best results. The identified data for such a study is:

* Affluent Neighborhoods in Charlotte Area,
* Real Estate Values of Charlotte Area,
* List of Schools,
* School Ranking,
* Accessibility of neighborhoods to nearby venues,
* Crime Rate, if possible

**Methodology**

As per the earlier studies, it has been established that real estate is dependent on its surroundings and proximity to important landmarks. To establish the same correlation, following methodology with be followed:

*Data Collection*- Since primary data collection is out of scope of this project, secondary data and tertiary data will be used for analysis.

*Data Wrangling*- To get an insight of the data, data cleaning is one of the most important tasks. As per the requirement, data will be filtered from the available sources. As of now, data has been acquired from websites like: http://www2.cms.k12.nc.us/cmsdepartments/accountability/Pages/PerformanceDashboardNew.aspx

(for List of Schools and School Performance Grades); https://www.zipdatamaps.com/mecklenburg-nc-county-zipcodes (for Zip code names); https://nces.ed.gov/ccd/schoolsearch/school\_list.asp?Search=1&DistrictID=3702970 (for School codes); https://www.niche.com/k12/search/best-public-elementary-schools/d/charlotte-mecklenburg-schools-nc/ (for School and Neighborhood Rankings); https://www.zillow.com/charlotte-nc-28205/home-values/

(for real estate values); https://opendata.arcgis.com/datasets/87376bdb0cb3490cbda39935626f6604\_0.geojson?where=STATE%20like%20'%25NC%25'%20AND%20CITY%20like%20'%25Huntersville%25'(for maps); https://simplemaps.com/data/us-cities (for US cities data).

*Data at a Glance- Exploratory Data Analysis*- From the available data, it seems that instead of top down approach, bottom’s up approach will be followed to check the correlation of the data. Data Visualization will play an important aspect for the inspection and at a glance overview of the available data.

*Data Analysis- Statistical Analysis*- Any statistical inference will be of great help for finding the desired results. Data will be plotted on the map and Foursquare venues will be explored. This will help to make selection easy with quick identification of the neighborhoods and the schools. K-clustering will be used to analyze the clusters and find the best neighborhood for relocation.