**Applied Data Science Capstone Project – The Restaurant Battle of Neighborhood in Philadelphia**

**Problem:** English can be a difficult language for immigrants coming to United States. The county Philadelphia in PA provides English speaking classes for the immigrants. Immigrants would be interested to taste different types of food in Philadelphia and it’s surrounding counties. Immigrants who prefer to eat their preferred cuisine, finding the right spot to eat in Philadelphia can be difficult and frustrating.

This project provides information regarding most dense/frequent cuisines in the following counties in and around Philadelphia: Philadelphia, Berks, Bucks, Chester, Cumberland, Delaware, Lancaster, Montgomery in PA. Camden, Salem counties in NJ. Thus the goal of this exercise gives in which county will you find large number or even concentration of which type of restaurants. Where to eat Chinese food? Where to eat Italian food?

**Data Preparation:** Following data sources will be needed to extract/generate the required information.

1) Philadelphia ESL class locations csv dataset is publicly available at <https://www.opendataphilly.org>. It contains Postal code, Latitude, Longitude of Philadelphia county.

2) Neighborhood, City, Postal code, Latitude and Longitude information of Philadelphia and surrounding 10 counties information is available from URL, <https://www.opendataphilly.org>. By passing the latitude and longitude into Foursquare API, we can explore venue information like Neighborhood Latitude, Neighborhood Longitude, Venue, Venue Latitude, Venue Longitude, Venue Category. I will use this data about restaurants in the neighborhoods around Philadelphia to show the density of restaurants, type of cuisine. After the data collection we can run k-means clustering to cluster the cuisine type by county/Neighborhood.