**Coursera Capstone – “The Battle of the Neighborhoods”**

*Comparative Analysis using location data via Foursquare*

# Problem Statement

Loudoun County is seated roughly 25 miles due west of Washington DC. It is the fastest growing county as well as the richest county per capita household income in the United States at $130,000. It is also where 80% of all internet traffic in the world flows. It seems like every week a new data center is being added to Loudoun County. By 2020, Apple, and Amazon will have second headquarters here in the area. It is the third most populous county in the state of Virginia and is a bedroom community for business moguls, diplomats, and politicians. It is also known for its rich history and farmlands which have been slowly converting into vineyards and breweries that draw crowds from all over the National Capital Area.

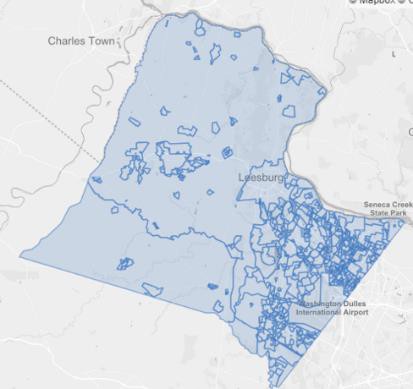
The objective of this Capstone project is to use Foursquare location data and regional clustering of

venue information to determine what might be the “best” place in Loudoun County to open a restaurant. What makes this so challenging is that there is no shortage of restaurants offering every type of the worlds culinary cuisine. From Mexican to Chinese, and Russian to Middle Eastern, the offerings are seemingly endless, thus competition, is high. Also, Loudoun County commands huge rent costs due to the overwhelmingly expensive cost of land and development. Therefore, the objective is to find a location that is in a reasonable rent area and a population dense area, with high disposable income.

# Data Overview

I will utilize data from Wikipedia regarding Loudoun County as well as Foursquare for this analysis. I will need to scrape the data, clean the data, and run analysis on the data. First I will have to determine the most likely neighborhoods based on mean rent prices and population density, as well as the disposable income. I will then consider the total number of venues and additional criteria to make a final determination. In addition, I am using data directly from the county office of planning and land management including Loudoun\_Zoning.csv, Loudoun\_population.csv, and Loudoun\_income.csv, seen below on the next page.



In addition I used the counties LOGIS mapping system and Folium to create choropleth maps of income and population as well as rent cost distributions to help in the identification of where to put the restaurant. As well as JSON files to identify county and city

boundaries. In identifying and creating this map I realized that it was actually easier and more cogent to use the 8 districts as my first layer of data visualization as I peeled back the layers on the county data. I did this because the County itself is organized into 8 major subdivisions as opposed to using city, town, or ordinance boundaries. This keeps the datum more clean and easier to follow.

esburg town proper.