# Evaluating Cities in Fairfax County of Virginia, USA Jichong Wu | June, 2020

#### 1. INTRODUCTION

#### 1.1 The Background

The Fairfax County is a county of the Virginia State of the United States located in the suburban of the capital city Washington, D.C. The county is one of the well-developed city clusters that has a promising future to continue to flourish, in terms of population, demographics, income level, location, and employment opportunities.

The Fairfax County is the 3<sup>rd</sup> highest-income county in US, accounts for 13% of Virginia's population and 20% of the DC-VA-MD Metropolitan Area population, and is home to seven Fortune 500 companies. Also, Amazon's second headquarter offices has been selected to locate right next to the Fairfax County.

#### 1.2 The Problem

The county has a total area of 406 square miles (1,050 km<sup>2</sup>), including 58 cities or 80 zip codes. With the rapid development and also as a resident living in the area myself, it is helpful to know which cities have the higher potential or in other words more attractive to live in.

This project is designed to evaluate all the 58 cities in Fairfax County and identify those "good" cities to live in. Coding will use data in areas of population, housing market, income level, crime, and school quality, Foursquare location data will also be used to explore and segment these cities.

#### 1.3 Target Audience

The result of this project might be of interest to the current and future residents of the Fairfax County, as well as to property investors.

#### 2. DATA COLLECTION

#### 2.1 Data Source

City values are typically determined by the following factors: real estate market trends, demographic trends, income level and job access, crime and safety, school quality, and nearby features. The following datasets were collected accordingly (Table 1):

1. Fairfax County city list and zip code list. This is the foundation of this project and other indicator data will be organized by this city and/or zip code list. Official lists gained from the Fairfax County Government was selected for this data needs.

- 2. **Real estate market**. Property values in each city is a good indicator of the value and potential of the city. The popular real estate website Zillow has its own house value index by zip code. Higher index indicates higher city value.
- **3. Demographics**. Population is a good indicator for a city. The better a city is, the more residents it has. Official 2019 population data from the Fairfax County Government was selected, the higher population the more positive for a city assessment. **Data Example:** screenshot of the data source website and raw data of Fairfax County
  - **Data Example:** screenshot of the data source website and raw data of Fairfax County city population by zip code
- **4. Income level.** Good cities tend to have more high-income residents. The per capita income from 2014-2018 by zip code/city of Fairfax County from the popular income statistics source American Community Survey was selected for this assessment indicator. The higher per capita income, the more positive for a city.
- **5. Crime and safety.** Crime rate is another common indicator for city evaluation. The Weekly Crime Data from the Fairfax County Police Department will be used, and it is a real time database that's always updated for the most recent 7 days. The more total crime number for a selected city, the lower value of this city.
  - **Data Example:** screenshot of the crime dataset with dates and by zip code
- 6. **School quality.** School quality is another common and important indicator. School scores of the public elementary schools of Fairfax County for 2019 from the popular school rating website SchoolDigger was selected for this indicator. School scores will be aggregated by zip code/city, the higher score the better city.
- 7. Nearby features. Location is everything for a city. It makes a city more valuable and attractive if, for example, it's close to shopping centers, entertainment businesses, food and beverage amenities, city centers, recreational centers, sought-after tourist areas, and office buildings. Foursquare venue category data will be used to explore and evaluate the nearby features for a given city, cities with similar characteristics may be segmented and grouped in this analysis.

**Table 1: Data Collection Summary** 

#	Data	Source	Method
1	Fairfax County city and zip code list	Fairfax County Government, Open Geospatial Data https://opendata.arcgis.com/datasets/3111976184004077 b836f535f31eadf1_21.csv	Import from a downloadable csv file from the source
2	Property Value	Zillow, home value index (by zip code and by city), current <a href="https://www.zillow.com/fairfax-county-va/home-values/">https://www.zillow.com/fairfax-county-va/home-values/</a>	Import from a downloadable csv file from the source
3	Population	Fairfax County Government, Economic, Demographic and Statistical Research (EDSR) unit, demographics data 2019 <a href="https://www.fairfaxcounty.gov/gisapps/iplsmap/IPLSExport.ashx?Type=Z&amp;Year=2019&amp;Param=ALL">https://www.fairfaxcounty.gov/gisapps/iplsmap/IPLSExport.ashx?Type=Z&amp;Year=2019&amp;Param=ALL</a>	Import from a downloadable csv file from the source
4	Income	American Community Survey, per capita income, 2014-2018 http://www.livehealthyfairfax.org/indicators/index/view?indicatorId=15&localeTypeId=3&periodId=1581	Import from a downloadable csv file from the source

5	Crime	Fairfax County Police Department, Weekly Crime Data, current <a href="https://www.fairfaxcounty.gov/apps/pfsu/api/file/crimereportsfromsp">https://www.fairfaxcounty.gov/apps/pfsu/api/file/crimereportsfromsp</a>	Import from a downloadable csv file from the source
6	School quality	SchoolDigger, Fairfax County elementary public school scores, 2019 <a href="https://www.schooldigger.com/qo/VA/zip/22033/search.aspx">https://www.schooldigger.com/qo/VA/zip/22033/search.aspx</a>	Make an excel spreadsheet
7	Nearby features	Foursquare	Venue category feature (most common venues)

# 2.2 Data Cleaning

The Fairfax city and zip code list was downloaded from the government website. Cleaning work included rename the column titles and drop irrelevant columns to get data frame Fairfax\_city\_list.

The house value data can't be scraped from the website but it was downloadable. Original data separates house values by zip code and by city and they are different subsets. Cleaning work focused on obtaining a complete list of house values including all zip codes and cities the source has. Also, a dollar sign and thousand separator were added to the values for easy reading. Then the two lists were merged with the *Fairfax\_city\_list* data frame, named *df\_fairfax\_masterlist*. Following datasets will follow cities and zip codes from this master list.

The population data had no issues and data frame <u>df\_pop\_clean</u> was created.

The income data required grouping by zip code and calculating the average values after grouping. Data frame <u>df\_income\_clean</u> was created.

The real-time (constantly being updated) crime data has a non-integer zip code value being as point zero. Decimal place was needed to be removed, then crime data was grouped by zip code and counted the total. Data frame *df\_crime\_clean* was created.

The school scores were calculated by mean value after grouping by zip code. Data frame *df\_school* was created.

Then, merging these data frames (data frame <u>df\_merged\_final</u> was created) and using geopy to get the geo-coordinates for each zip code.

Using Foursquare API, 100 venues within 15,000 meters for each zip code were obtained (data frame *df\_fairfax\_venue\_category*), then the top 5 most common venues for each zip code were able to obtain (data frame *fairfax\_top\_venues*). Adding these top 5 venues back to the main data frame *df\_merged\_final* to get *df\_final*.

#### 2.3 Data Feature Selection

For each zip code, population, income (average per capita), house value, total crime number, school rating scores, and the top 5 venues were the feature that will be used for evaluating each city.

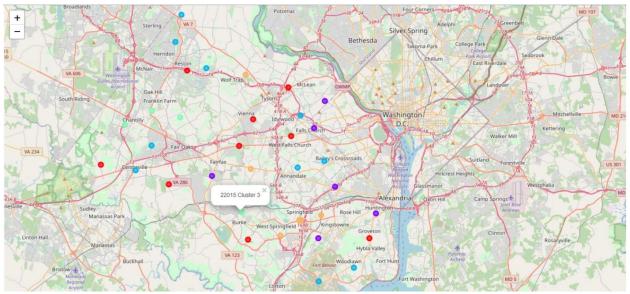
Data frame <u>df\_nor</u> was then created after applying data normalization for column population, income, house value, total crime number, and school rating scores for the calculation.

Data Frame *df\_rank* was the final results after ranking and comparison.

## 3. DATA ANALYSIS

## 3.1 City Segmentation & Clustering

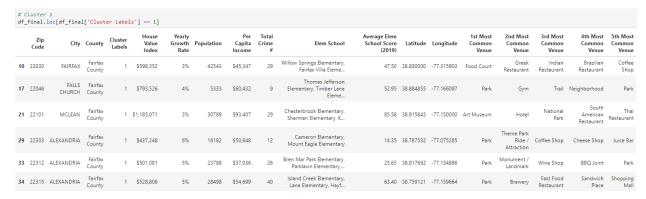
Five segments were identified based on the most common venues for each zip code.



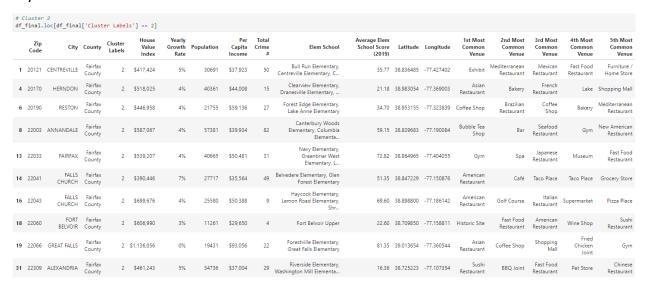
Cluster0 is featured with a convenient store-cluster.



Cluster1 has a higher house value with many parks.



# Cluster2 is restaurant-clusters with many American and Asian food, and they are all closer to city centers.



# Cities in Cluster3 locate in suburban area and most of them are relatively away from city centers.

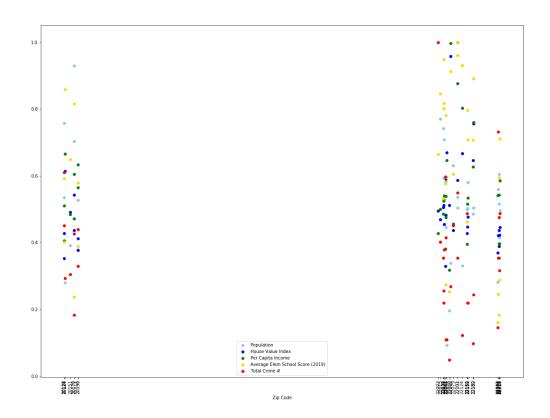


Cities in Cluster4 are restaurant-cluster area and are away from city centers.

	# Cluster 4 df_final.loc[df_final['Cluster Labels'] == 4]																	
	Zip Code	City	County	Cluster Labels	House Value Index	Yearly Growth Rate	Population	Per Capita Income	Total Crime #	Elem School	Average Elem School Score (2019)	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
12	22032	FAIRFAX	Fairfax County	4	\$628,088	4%	29317	\$49,660	18	Wakefield Forest Elementary, Oak View Elementa	71.43	38.837883	-77.258316	Diner	Coffee Shop	Fast Food Restaurant	Sushi Restaurant	Grocery Store
20	22079	LORTON	Fairfax County	4	\$517,581	4%	36203	\$42,645	37	Laurel Hill Elementary, Gunston Elementary, Lo	53.87	38.779636	-77.184847	Mediterranean Restaurant	Toy / Game Store	Italian Restaurant	Warehouse Store	Brewery
24	22150	SPRINGFIELD	Fairfax County	4	\$506,861	4%	28626	\$36,946	40	Key Center, Springfield Estates Elementary, Fo	41.25	38.779030	-77.187414	Coffee Shop	Brewery	Italian Restaurant	Gym / Fitness Center	Toy / Game Store
25	22152	SPRINGFIELD	Fairfax County	4	\$529,866	5%	28867	\$48,209	18	Keene Mill Elementary, West Springfield Elemen	70.98	38.762084	-77.215550	Bubble Tea Shop	Thai Restaurant	BBQ Joint	Salad Place	Grocery Store
28	22182	VIENNA	Fairfax County	4	\$896,903	2%	28952	\$71,029	20	Wolftrap Elementary, Colvin Run Elementary, Ki	79.32	38.945437	-77.266294	Restaurant	Portuguese Restaurant	Salad Place	BBQ Joint	Gym
32	22310	ALEXANDRIA	Fairfax County	4	\$517,466	5%	29621	\$50,766	39	Clermont Elementary, Bush Hill Elementary, Fra	52.98	38.785138	-77.155274	Waterfront	Afghan Restaurant	Theater	Park	Mediterranean Restaurant

# 3.2 Scatter Plot Visualization & Comparison

Normalized column data population, income, house value, total crime number, and school rating scores were able to be compared on a same scale. Scatter plot was applied to visualize zip code strengths and weakness. Red color was used for crime # as it's easy to standout and it's different than other data values – the higher value the worse. The dots on the top part of the chart (excerpt the red dot) indicate good cities



# 3.3 Weighted Average Score & City Ranking

To calculate the weighted average scores, first the normalized crime data was converted to negative values since the higher the worse. Then all normalized data values of column data population, income, house value, total crime number, and school rating scores were added up. Finally ranking of each zip code was created by sort the total add-up score.

	Zip ode	Rank	City	County	Cluster Labels	House Value Index	Yearly Growth Rate	Population	Per Capita Income	Total Crime #	Elem School	Average Elem School Score (2019)	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
<b>21</b> 22	101	3.143407	MCLEAN	Fairfax County	1	1.000000	3%	0.536571	1.000000	-0.353659	Chesterbrook Elementary, Sherman Elementary, K	0.960494	38.915843	-77.150000	Art Museum	Hotel	National Park	South American Restaurant	Thai Restaurant
19 22	:066	2.938240	GREAT FALLS	Fairfax County	2	0.958640	0%	0.338631	0.996242	-0.268293	Forestville Elementary, Great Falls Elementary	0.913019	39.013654	-77.360544	Asian Restaurant	Coffee Shop	Shopping Mall	Fried Chicken Joint	Gym
<b>28</b> 22	182	2.668150	VIENNA	Fairfax County	4	0.756835	2%	0.504557	0.760425	-0.243902	Wolftrap Elementary, Colvin Run Elementary, Ki	0.890236	38.945437	-77.266294	Restaurant	Portuguese Restaurant	Salad Place	BBQ Joint	Gym
23 22	124	2.608919	OAKTON	Fairfax County	3	0.667175	2%	0.330824	0.802456	-0.121951	Oakton Elementary, Waples Mill Elementary	0.930415	38.896743	-77.362288	Gym	American Restaurant	Spa	Deli / Bodega	Pizza Place
5 20	171	2.467336	HERNDON	Fairfax County	3	0.543319	4%	0.929227	0.605458	-0.426829	Oak Hill Elementary, Floris Elementary, Fox Mi	0.816162	38.904368	-77.403626	Trail	Coffee Shop	Airport Lounge	Brewery	Fast Food Restaurant
22 22	102	2.419160	MCLEAN	Fairfax County	0	0.586869	4%	0.504017	0.877054	-0.548780	Spring Hill Elementary	1.000000	38.931275	-77.203997	Wine Shop	Vietnamese Restaurant	Butcher	Neighborhood	Park
<b>27</b> 22	180	2.368264	VIENNA	Fairfax County	0	0.646942	3%	0.485509	0.626752	-0.097561	Stenwood Elementary, Louise Archer Elementary,	0.706622	38.894999	-77.254909	Taco Place	Shopping Mall	Butcher	Supermarket	Thai Restaurant
11 22	031	2.324079	FAIRFAX	Fairfax County	0	0.511751	4%	0.593820	0.525111	-0.256098	Mosby Woods Elementary, Mantua Elementary, Fai	0.949495	38.864393	-77.275968	American Restaurant	Neighborhood	Seafood Restaurant	Soup Place	Thai Restaurant
16 22	043	2.247035	FALLS CHURCH	Fairfax County	2	0.590409	4%	0.445792	0.539446	-0.109756	Haycock Elementary, Lemon Road Elementary, Shr	0.781145	38.898800	-77.186142	American Restaurant	Golf Course	Italian Restaurant	Supermarket	Pizza Place
9 22	015	2.183607	BURKE	Fairfax County	3	0.469689	4%	0.770516	0.500723	-0.402439	Terra Centre Elementary, Cherry Run Elementary	0.845118	38.788896	-77.272504	Afghan Restaurant	Portuguese Restaurant	American Restaurant	Grocery Store	American Restaurant
<b>12</b> 22	:032	2.154742	FAIRFAX	Fairfax County	4	0.530000	4%	0.510918	0.531652	-0.219512	Wakefield Forest Elementary, Oak View Elementa	0.801684	38.837883	-77.258316	Diner	Coffee Shop	Fast Food Restaurant	Sushi Restaurant	Grocery Store
13 22	:033	2.143360	FAIRFAX	Fairfax County	2	0.455000	4%	0.708684	0.540441	-0.378049	Navy Elementary, Greenbriar West Elementary, L	0.817284	38.864965	-77.404055	Gym	Spa	Japanese Restaurant	Museum	Fast Food Restaurant
<b>2</b> 20	1124	2.125094	CLIFTON	Fairfax County	0	0.613926	3%	0.279918	0.665346	-0.292683	Union Mill Elementary	0.858586	38.820280	-77.378782	Coffee Shop	Italian Restaurant	Indian Restaurant	Pizza Place	Golf Course
<b>26</b> 22	153	2.081439	SPRINGFIELD	Fairfax County	0	0.477286	5%	0.580732	0.533954	-0.219512	Sangster Elementary, Newington Forest Elementa	0.708979	38.757142	-77.262515	Coffee Shop	Trail	Food Court	Pizza Place	Bakery
<b>25</b> 22	152	2.043432	SPRINGFIELD	Fairfax County	4	0.447118	5%	0.503076	0.516118	-0.219512	Keene Mill Elementary, West Springfield Elemen	0.796633	38.762084	-77.215550	Bubble Tea Shop	Thai Restaurant	BBQ Joint	Salad Place	Grocery Store
10 22	:030	1.911249	FAIRFAX	Fairfax County	1	0.504908	3%	0.741413	0.485478	-0.353659	Willow Springs Elementary, Fairfax Villa Eleme	0.533109	38.830000	-77.315903	Food Court	Greek Restaurant	Indian Restaurant	Brazilian Restaurant	Coffee Shop

# 4. CONCLUSION

According to the ranking, here's top 10 cities of Fairfax County, it is very much consistent with the general sense and reality:

#	Zip Code	City	Score
1	22101	MCLEAN	3.143407
2	22066	GREAT FALLS	2.93824
3	22182	VIENNA	2.66815
4	22124	OAKTON	2.608919
5	20171	HERNDON	2.467336
6	22102	MCLEAN	2.41916
7	22180	VIENNA	2.368264
8	22031	FAIRFAX	2.324079

9	22043	FALLS CHURCH	2.247035
10	22015	BURKE	2.183607

#### **5. FUTURE STUDIES**

To further improve this research, the indicators can be expanded and more dataset can be added to make the calculation system more robust. Also, user ratings for selected venues in each city/zipcode from Foursquare can be added. It's stronger to evaluate a city by comparing aggregated user ratings. Lastly, more sophisticated evaluation/rating calculation methods can be applied than the weighted average model. A regression model can also be built to predict city value and development outlook.