Moving to a New Location

Philip Kinney

Introduction

<u>The Problem</u>

Moving to a different state is difficult, especially with children. Many parents are hesitant to move because of the uncertainties of how it would disrupt their children's current lifestyle. What if there was a way to make the transition easier? While it is never easy to say, "Good-bye," to old friends and try to make new ones, what if there was an easy way to initiate your social life in a new location?

<u>The Solution</u>

I created a two tier computer program to assist families with gaining a general understanding of an area prior to moving. The first tier analyzes the county that they are thinking of moving to. The outputs reveal the top venues in the area broken down by borough. The second tier analyzes the major factors that make up the family's lifestyle in proximity to where they think they might want to live. The outputs reveal what is available within a 10 mile (16 km) radius from their desired location. The purpose is to see if where they think they might want to live would have similar venues as their current location, presumably making their social integration easier.

<u>Example</u>

In my example (shown below), I chose to analyze Chester County, Pennsylvania as my first tier analysis. I then narrowed my analysis in the second tier to what a family might be looking for if they were to consider moving to Malvern, Pennsylvania, a location within Chester County.

Data

Wikipedia.com: The names of the boroughs within Chester County were found on Wikipedia.com. This is a list of locations that will be used to gain an understanding of what is available within the different parts of the county.

Google: The coordinates of the boroughs were found using Google's search function and Google Maps. This list of coordinate data will be used to help search for venues and plot the data on a map.
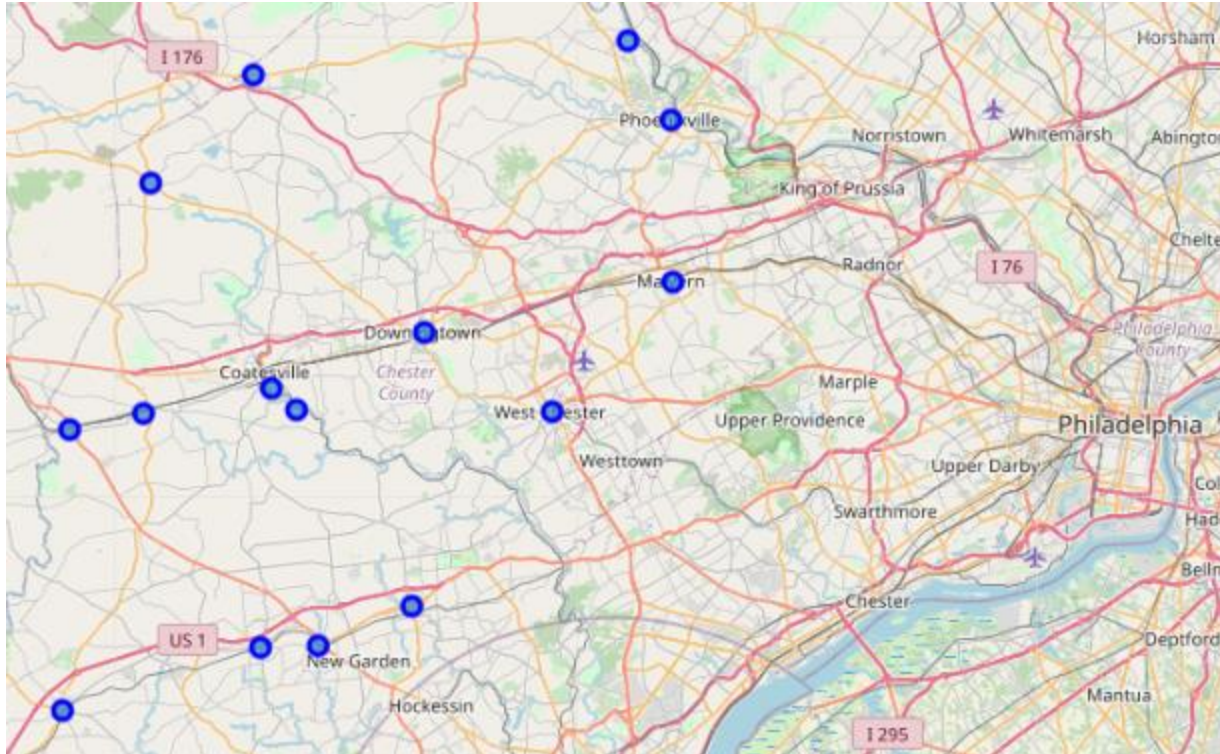
Venues: The venue data was all found using the Foursquare API. This list of data is what is available to do in the area, based on a specific radius (in meters) defined by the user.

Methodology: Tier 1 Analysis

<u>Exploratory Data Analysis</u>

Chester County is located west of Philadelphia, within the eastern section of the state. It has # of boroughs scattered throughout the county.

Based on the coordinates of the boroughs, using a search radius of 1600 meters per borough, there are 402 venues throughout the county within 121 unique venue categories.

Out[9]:

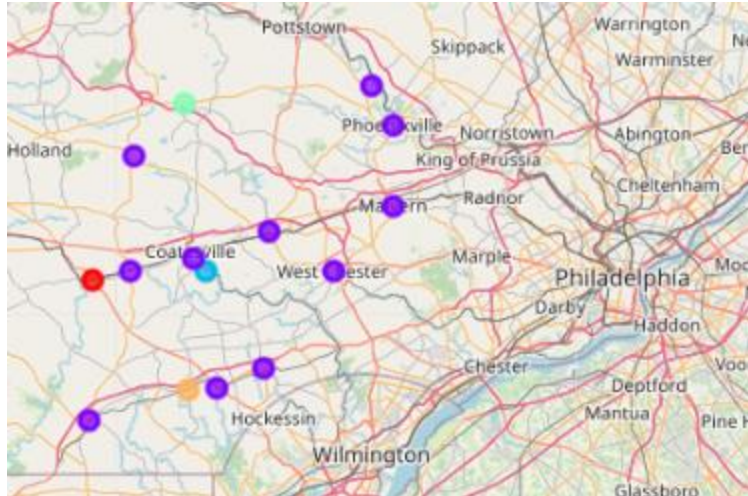|	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Atglen	39.9493	-75.9736	Reaching Higher LLC	39.956712	-75.980964	Construction & Landscaping
1	Atglen	39.9493	-75.9736	Schiffer Book Farm	39.948898	-75.957474	Bookstore
2	Avondale	39.8234	-75.7833	Va La Vineyards	39.817111	-75.772314	Vineyard
3	Avondale	39.8234	-75.7833	Glen Willow Orchards	39.835500	-75.785355	Farmers Market
4	Avondale	39.8234	-75.7833	Domino's Pizza	39.823428	-75.781126	Pizza Place
5	Avondale	39.8234	-75.7833	Dunkin'	39.823372	-75.780938	Donut Shop
6	Avondale	39.8234	-75.7833	Walgreens	39.834767	-75.793396	Pharmacy
7	Avondale	39.8234	-75.7833	Outfield Landscaping	39.813218	-75.775551	Construction & Landscaping
8	Avondale	39.8234	-75.7833	Jd Mushrooms	39.812199	-75.776009	Food
9	Avondale	39.8234	-75.7833	Walgreens	39.817253	-75.768410	Pharmacy
10	Avondale	39.8234	-75.7833	Gigante International Food Market	39.816902	-75.768210	Grocery Store
11	Downingtown	40.0065	-75.7033	Kerr Park	40.008275	-75.704929	Playground

```
print(venues.shape)
venues.groupby('Neighborhood').count()
print('There are {} unique categories.'.format(len(venues['Venue Category'].unique())))

(402, 7)
There are 121 unique categories.
```

Machine Learning

K-means clustering is a machine learning algorithm that clusters a collection of data points based on certain similarities. The outcome is to locate the center, or centroid, of the data to help further understand the similarities and differences between the data point. K-means clustering was used to further analyze the data within Chester County.

Coursera Capstone Final Report

Below are the results per cluster:

Cluster Label 0:

	Latitude	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	39.9493	Atglen	39.9493	-75.9736	0	Construction & Landscaping	Bookstore	Whisky Bar	Coffee Shop	Concert Hall	Convenience Store	Cosmetics Shop	Dance Studio	Deli / Bodega	Dessert Shop

Custer Label 1:

	Latitude	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	39.8234	Avondale	39.8234	-75.7833	1	Pharmacy	Construction & Landscaping	Farmers Market	Food	Donut Shop	Pizza Place	Grocery Store	Vineyard	College Auditorium	Concert Hall
2	40.0065	Downingtown	40.0065	-75.7033	1	American Restaurant	Brewery	Italian Restaurant	Gym / Fitness Center	Gas Station	Convenience Store	Coffee Shop	Pizza Place	Bakery	Café
4	40.0943	Honey Brook	40.0943	-75.9113	1	Gas Station	Pharmacy	Restaurant	Italian Restaurant	Grocery Store	Chinese Restaurant	Toy / Game Store	Pizza Place	Brewery	Elementary School
5	39.8468	Kennett Square	39.8468	-75.7116	1	Italian Restaurant	Pizza Place	Gym / Fitness Center	Breakfast Spot	Mexican Restaurant	Baseball Field	Gastropub	Discount Store	Bakery	Paper / Office Supplies Store
6	40.0362	Malvern	40.0362	-75.5138	1	American Restaurant	Sandwich Place	Pizza Place	Big Box Store	Hotel	Italian Restaurant	Coffee Shop	Playground	Photography Studio	Pharmacy
8	39.7854	Oxford	39.7854	-75.9788	1	Fast Food Restaurant	Discount Store	Big Box Store	Liquor Store	Pet Store	Pharmacy	Dance Studio	Pub	Business Service	Italian Restaurant
9	39.9594	Parkesburg	39.9594	-75.9168	1	Pharmacy	Gym / Fitness Center	Discount Store	Light Rail Station	Italian Restaurant	Ice Cream Shop	Gas Station	Fruit & Vegetable Store	Video Store	Football Stadium

Cluster Label 2:

	Latitude	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	39.9611	Modena	39.9611	-75.7996	2	Farm	Park	Thai Restaurant	Elementary School	College Auditorium	Concert Hall	Construction & Landscaping	Convenience Store	Cosmetics Shop	Dance Studio

Cluster Label 3:

	Latitude	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	40.1568	Elverson	40.1568	-75.8327	3	Park	Italian Restaurant	Construction & Landscaping	Train Station	Elementary School	College Auditorium	Concert Hall	Convenience Store	Cosmetics Shop	Dance Studio

Cluster Label 4:

	Latitude	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
14	39.8221	West Grove	39.8221	-75.8274	4	Park	Baseball Field	American Restaurant	Automotive Shop	Deli / Bodega	Bank	Pizza Place	Furniture / Home Store	Construction & Landscaping	Convenience Store

Methodology: Tier 2 Analysis

The search criteria is based on a hypothetical family who would like to replicate their lifestyle without having to drive more than 10 miles in any direction from where they live. I created 11 different venue categories and search for them using the foursquare API as Malvern, PA as my starting point. I then scrubbed the data and put all of the results into their own "master dataframe" for further analysis. I then grouped the venues by neighborhood to determine which area near Malvern, PA has the most venue categories that fit my hypothetical family's lifestyle. Below are the examples of the data I gathered.

The eleven categories are: Church, Aldi, Golf Course, Library, Shopping Mall, Gas Station, School, Museum, Park, Zoo, and Thai Restaurants

A portion of the combined dataframe:

Out[114]:

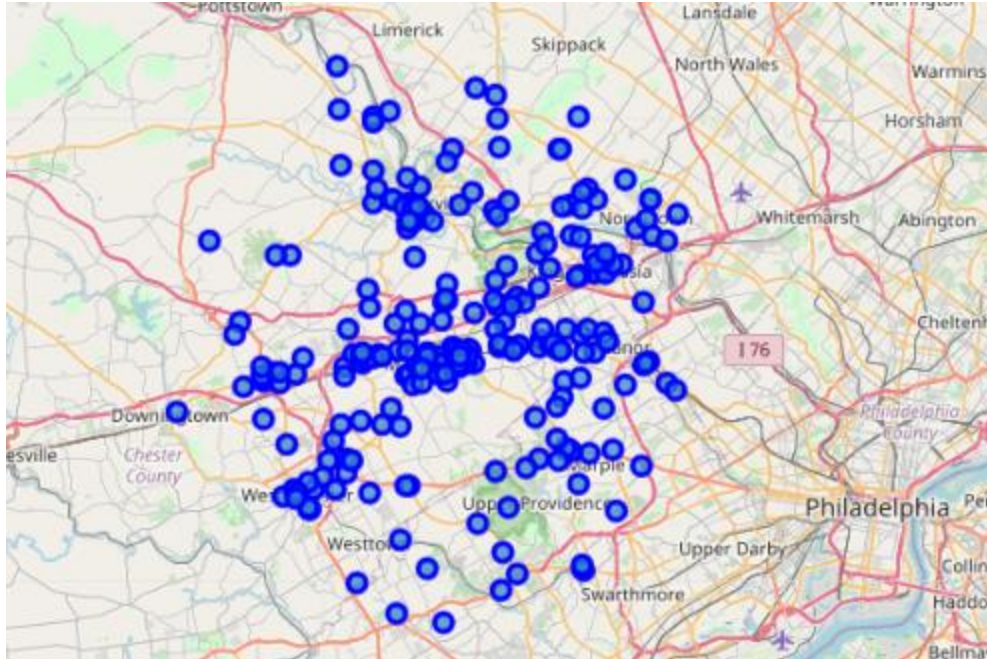|	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	postalCode
0	The Church of Saint Monica	Church	63 Main Ave	US	Berwyn	United States	at First Ave	5623	[63 Main Ave (at First Ave), Berwyn, PA 19312,...	[{'label': 'display', 'lat': 40.04608705725427...	40.046087	-75.440359	19312
1	Church of the Good Samaritan	Church	Lancaster Ave	US	Paoli	United States	NaN	1807	[Lancaster Ave, Paoli, PA 19301, United States]	[{'label': 'display', 'lat': 40.04091559619378...	40.040916	-75.490552	19301
2	St. Isaac Jogues Catholic Church	Church	50 W Walker Rd	US	Wayne	United States	NaN	6026	[50 W Walker Rd, Wayne, PA 19087, United States]	[{'label': 'display', 'lat': 40.07388665956120...	40.073887	-75.441086	19087
3	First Church Of Christ, Scientist	Church	NaN	US	Phoenixville	United States	NaN	1459	[Phoenixville, PA, United States]	[{'label': 'display', 'lat': 40.03923316394582...	40.039233	-75.508792	NaN
4	The Methodist Church	Church	Landowner Road	US	Malvern	United States	NaN	1571	[Landowner Road, Malvern, PA 19355, United Sta...	[{'label': 'display', 'lat': 40.038544, 'lng':...	40.038544	-75.510767	19355

[115]: penn.shape

Out[115]: (253, 15)

Map of the various venues:

Coursera Capstone Final Report

Summary description of the data:

```
penn.describe(include='all')
```

21]:

	Name	Categories	Neighborhood	Latitude	Longitude
count	253	253	243	253.000000	253.000000
unique	242	11	40	NaN	NaN
top	BP	Library	Malvern	NaN	NaN
freq	5	45	45	NaN	NaN
mean	NaN	NaN	NaN	40.051864	-75.480732
std	NaN	NaN	NaN	0.064029	0.083378
min	NaN	NaN	NaN	39.885068	-75.698385
25%	NaN	NaN	NaN	40.024321	-75.535052
50%	NaN	NaN	NaN	40.044193	-75.488419
75%	NaN	NaN	NaN	40.099600	-75.406838
max	NaN	NaN	NaN	40.210795	-75.317151

Analysis showing that the neighborhood of Berwyn has 10 matching venues across 6 of the 11 unique categories.

```
                               Name      Categories Neighborhood  \
1        The Church of Saint Monica          Church       Berwyn
2                 Easttown Library         Library       Berwyn
3                      CHS Library         Library       Berwyn
4    Tredyffrin-Easttown Middle School       School       Berwyn
5         Hillside Elementary School         School       Berwyn
6              Conestoga High School         School       Berwyn
7          Thai  Orchid Restaurant  Thai Restaurant       Berwyn
8                   Teegarden Park            Park       Berwyn
9        The Church Of Saint Monica          Church       Berwyn
10       Valley Fair shopping Center  Shopping Mall       Berwyn

     Latitude  Longitude
1   40.046087 -75.440359
2   40.044111 -75.441526
3   40.048180 -75.452540
4   40.047807 -75.448091
5   40.056866 -75.456100
6   40.046999 -75.453060
7   40.047695 -75.437849
8   40.061483 -75.449172
9   40.046040 -75.440033
10  40.070389 -75.442851
There are 6 unique categories.
```

Results

Tier 1 analysis

The results of the Tier 1 analysis show the various venues of Chester County, Pennsylvania. Based on the analysis, there are a wide diversity of venues within the county, giving way to plenty of activities to acclimate a new family into the area. Further analysis provided by the k-means machine learning programming demonstrate that most of the boroughs within Chester County are similar. This explains that our family has a wide selection within the county to choose where to live.

Tier 2 analysis

Based on our hypothetical search, Malvern returned with the highest results of 10 unique categories out of 11 that are located within a 10 mile radius. Three other neighborhoods returned 8 unique categories out of 11 within a 10 mile radius. They are Norristown, Phoenixville, and Westchester. Interestingly enough, the one category that was missing from Malvern, PA was the zoo, which was present within the Norristown search. Part of the Malvern data is shown below.

```
                                Name       Categories Neighborhood
76              Malvern Shopping Center    Shopping Mall      Malvern
77            Malvern Preparatory School          School      Malvern
78         Great valley school district office    School      Malvern
79                         PSGV library         Library      Malvern
80               Malvern Public Library         Library      Malvern
81                     The Phelps School          School      Malvern
82      Theodore S. A. Rubino memorial Park         Park      Malvern
83                   The Vanguard School          School      Malvern
84           Great Valley Shopping Center    Shopping Mall      Malvern
85               Wharton Esherick Museum          Museum      Malvern
86         General Wayne Elementary School         School      Malvern
87        St Patrick's Elementary School          School      Malvern
88                  The Methodist Church          Church      Malvern
89          K.D Markley Elementry School          School      Malvern
90       St. Patrick Roman Catholic Church        Church      Malvern
91                 Grant Griffiths Golf     Golf Course      Malvern
92             Chester Valley Golf Club     Golf Course      Malvern
93                      St. Pats Church          Church      Malvern
94       Great Valley Presbyterian Church        Church      Malvern
95                 Applebrook Golf Club     Golf Course      Malvern
96                   Royal Thai Orchid   Thai Restaurant      Malvern
97                 Malvern Memorial Park            Park      Malvern
98      Malvern Courtyard Shopping Center   Shopping Mall      Malvern
```

```
There are 10 unique categories.
```

Discussion

Based on my hypothetical scenario, the family looking to relocate to Chester County, PA would seem to fit in well in Malvern, which provided for 10 of their 11 unique venue categories. The lone category missing within a 10 miles radius of Malvern, PA was the zoo. Norristown, a nearby neighborhood to Malvern, which was tied for the second best fit, had a zoo within its 10 mile radius. I did a quick Google search of the distance between Malvern and zoo within Norristown and discovered that they are only 14 miles apart. If my hypothetical family were to move to the north east section of Malvern, they might have the zoo within their desired 10 mile radius and thus having all 11 of their preferred venues. This reinforces the possibility of Malvern, PA being the best fit for my hypothetical family to move, based on their current social preferences. The Norristown data is shown below.

```
                          Name      Categories Neighborhood    Latitude
140            Elmwood Park Zoo            Park   Norristown   40.133178
141        Norristown Farm Park            Park   Norristown   40.144084
142            Elmwood Park Zoo             Zoo   Norristown   40.133178
143       Jefferson Golf Course     Golf Course   Norristown   40.138588
144           Aldi Food Market    Grocery Store   Norristown   40.129382
145        Liberty Gas Station     Gas Station   Norristown   40.124704
146          Holy Savior Church          Church   Norristown   40.111720
147            CITGO Gas Station     Gas Station   Norristown   40.115620
148  Audubon Square Shopping Center  Shopping Mall   Norristown   40.128078
149          Sunoco Gas Station     Gas Station   Norristown   40.108330
150      Jeffersonville Golf Club     Golf Course   Norristown   40.136659
151     Norristown Public Library         Library   Norristown   40.121552

       Longitude
140   -75.337629
141   -75.356809
142   -75.337629
143   -75.385544
144   -75.398713
145   -75.317151
146   -75.336482
147   -75.347847
148   -75.404294
149   -75.324538
150   -75.389958
151   -75.340512
There are 8 unique categories.
```

Conclusion

It's not easy relocating a family to a new state. As a result, many people choose to stay where they are. I created a computer program that helps display and analyze various boroughs within a county and the venues located within. I was then able to analyze a particular borough to see if it would be a good fit for the family, based on their 11 venue category preferences. The end results showed that it would be possible for the new family to have a similar lifestyle in Malvern, Chester County, Pennsylvania without much driving.