The Battle of Neighborhoods

Determining which major city is best to open Pizza Delivery service.

Determining which major city is best to open a Pizza Delivery service

- A small business owner wants to open up a pizza delivery service in a busy multicultural city.
- We will look at the Foursquare API data to determine which city would be most beneficial to open up this business.

Data acquisition

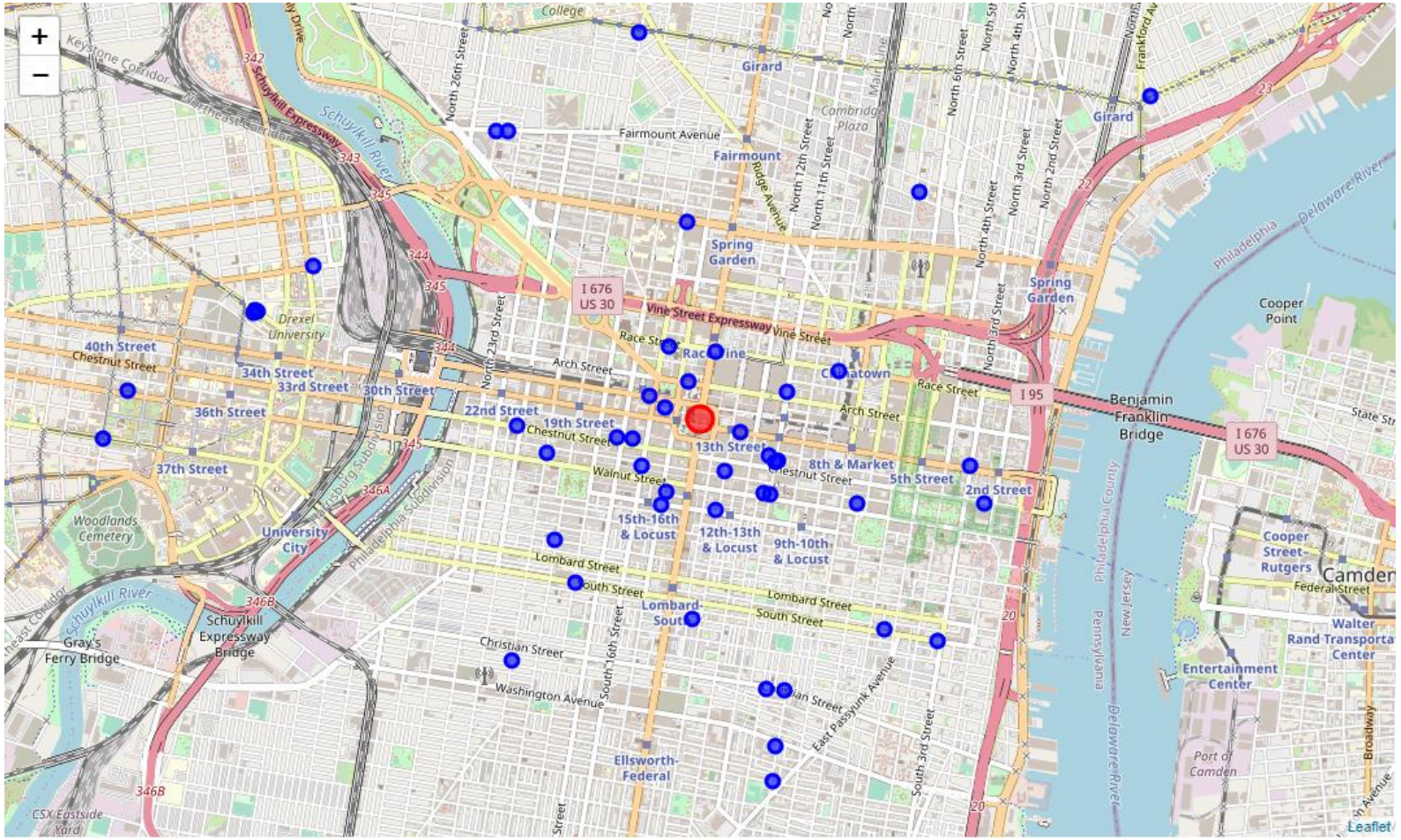
- Using the Foursquare API, we will only search pizza restaurants within a 5 mile radius of each cities center.
- We will find the most dense clusters within this area and use that centroid as our desired location.

- Initially we were going to work with 5 major cities, but decided to trim it down to only using 2 major cities – both being located in Pennsylvania.
- **Philadelphia** and **Pittsburgh, PA**

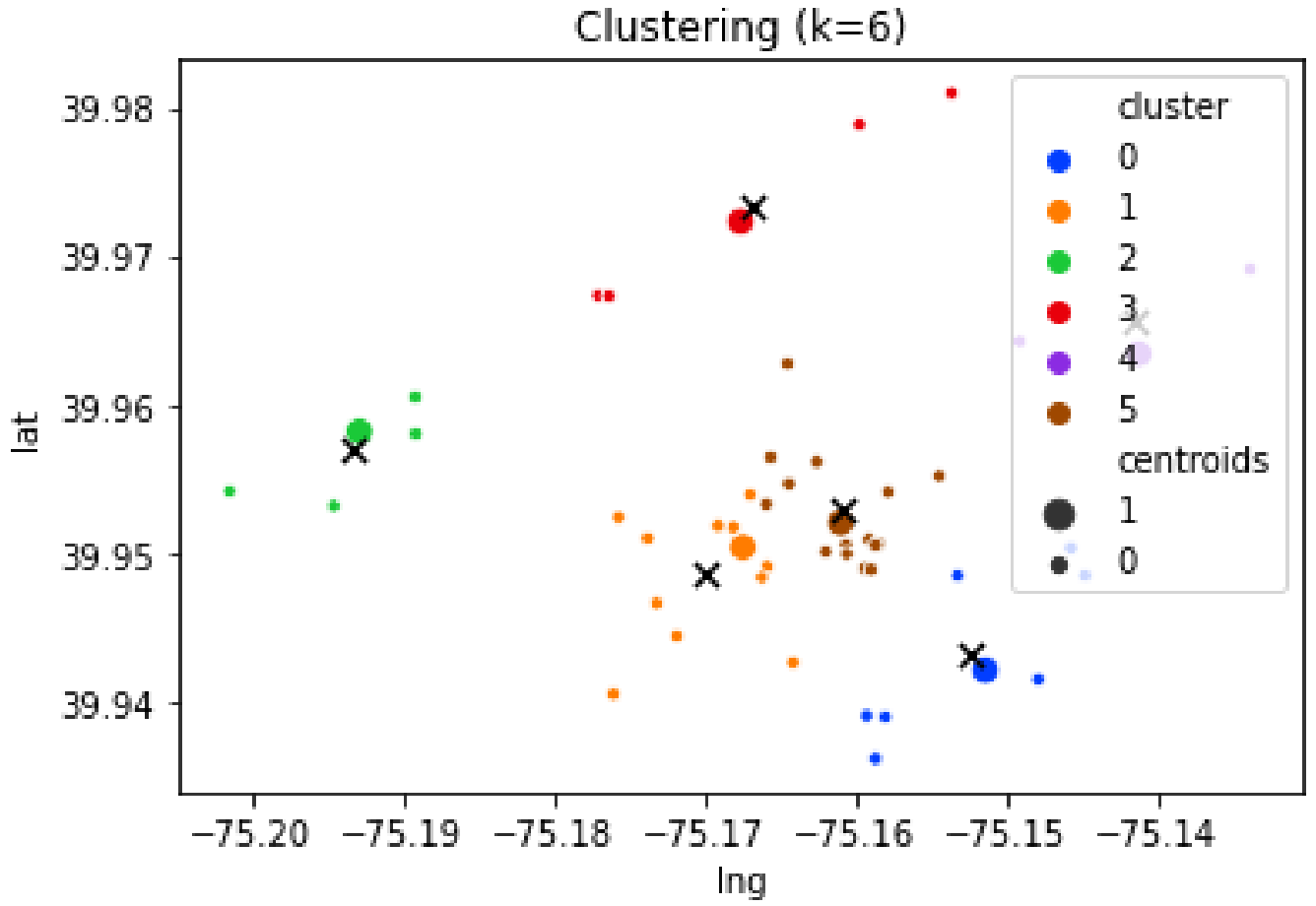
Query our first city, Philadelphia – and clean the data:

name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress
Jake's Pizza	Pizza Place	201 N Broad St	at Race St	39.956123	-75.162536	[{'label': 'display', 'lat': 39.95612338396027...	387	19107	US	Philadelphia	PA	United States	[201 N Broad St (at Race St), Philadelphia, PA...
Rex Pizza	Pizza Place	1526 Race St	NaN	39.956430	-75.165591	[{'label': 'display', 'lat': 39.95642960191507...	448	19102	US	Philadelphia	PA	United States	[1526 Race St, Philadelphia, PA 19102, United ...
Joe's Pizza	Pizza Place	122 S 16th St	at Sansom St	39.950372	-75.167422	[{'label': 'display', 'lat': 39.95037232183254...	423	19102	US	Philadelphia	PA	United States	[122 S 16th St (at Sansom St), Philadelphia, P...
Mix Brick Oven Pizza	Pizza Place	2101 Chestnut St	at 21st St	39.952399	-75.175665	[{'label': 'display', 'lat': 39.95239907148592...	1036	19103	US	Philadelphia	PA	United States	[2101 Chestnut St (at 21st St), Philadelphia, ...
Zio's Pizza	Pizza Place	111 S 13th St	13th & Sansom	39.950071	-75.161920	[{'label': 'display', 'lat': 39.95007113634469...	325	19107	US	Philadelphia	PA	United States	[111 S 13th St (13th & Sansom), Philadelphia, ...

Now plot the points on a folium map:



After doing a k-Means clustering algorithm on the data, we can visualize the separate clusters:



Further analysis:

After analysis the data further to see which cluster had the most restaurants in it:

- Cluster #5 shows that it contains the most restaurants with 16.

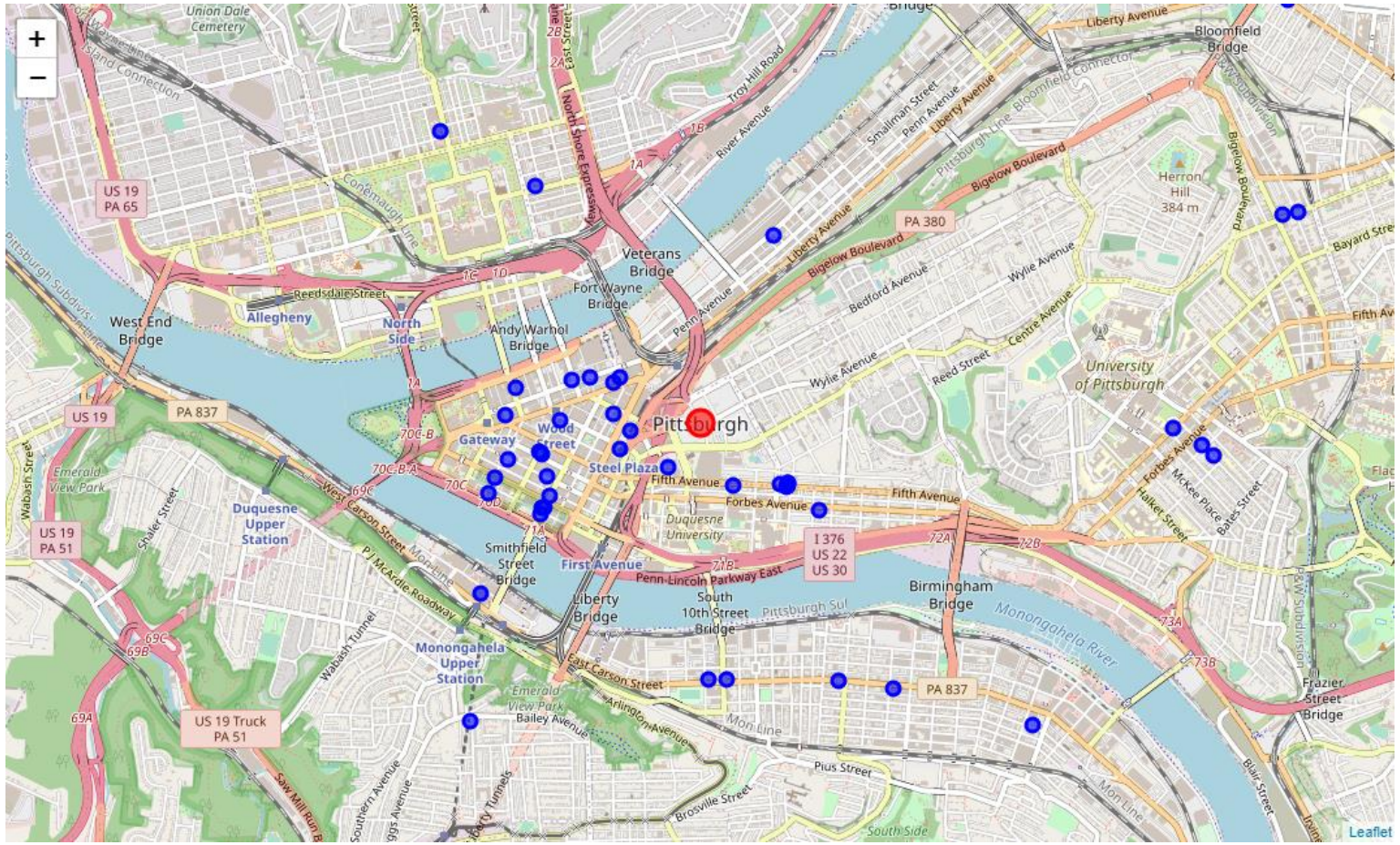
```
In [62]: phil_df['cluster'].value_counts().max(level=0)

Out[62]: 5    16
         1    12
         0     8
         2     6
         3     5
         4     3
         Name: cluster, dtype: int64
```

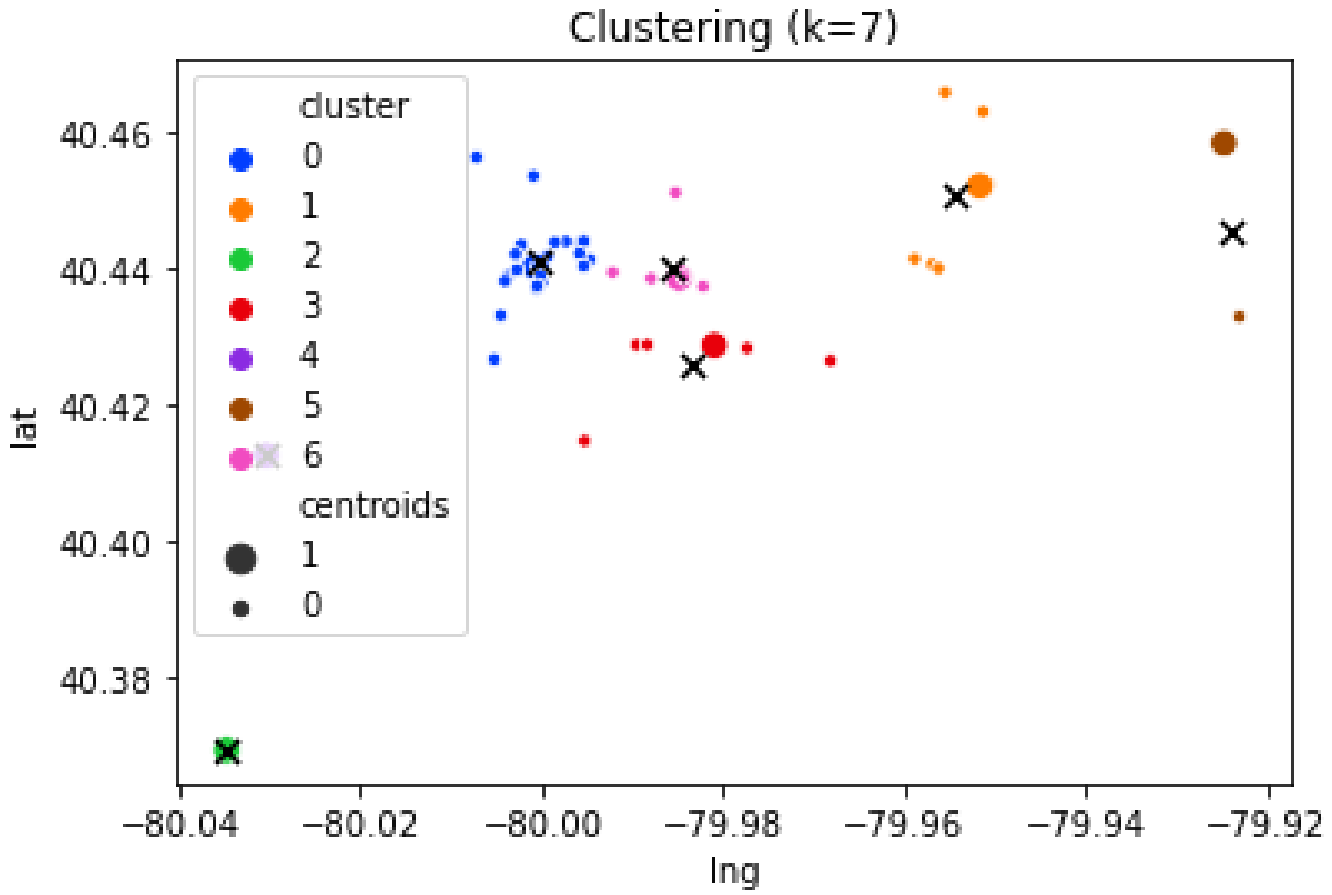
Query our second city, Pittsburgh – and clean the data:

name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	neighborhood
Pizza Parma	Pizza Place	963 Liberty Ave	40.443770	-79.995820	[{'label': 'display', 'lat': 40.44377, 'lng': ...	537	15222	US	Pittsburgh	PA	United States	[963 Liberty Ave, Pittsburgh, PA 15222, United...	NaN
Pizza Milano	Pizza Place	1304 5th Ave	40.438582	-79.987962	[{'label': 'display', 'lat': 40.4385821185373,...	390	15219	US	Pittsburgh	PA	United States	[1304 5th Ave, Pittsburgh, PA 15219, United St...	NaN
Genoa Pizza and Bar	Pizza Place	111 Market St	40.439001	-80.003671	[{'label': 'display', 'lat': 40.43900068922535...	1189	15222	US	Pittsburgh	PA	United States	[111 Market St, Pittsburgh, PA 15222, United S...	NaN
Domino's Pizza	Pizza Place	300 6th Ave Uppr 100	40.441832	-79.999361	[{'label': 'display', 'lat': 40.44183246264836...	785	15222	US	Pittsburgh	PA	United States	[300 6th Ave Uppr 100, Pittsburgh, PA 15222, U...	NaN
Napoli's Pizza	Pizza Place	1525 5th Ave	40.438725	-79.984352	[{'label': 'display', 'lat': 40.43872451782226...	587	15219	US	Pittsburgh	PA	United States	[1525 5th Ave, Pittsburgh, PA 15219, United St...	NaN

Now plot the points on a folium map:



After doing a k-Means clustering algorithm on the data, we can visualize the separate clusters:



Further analysis:

After analysis the data further to see which cluster had the most restaurants in it:

- Cluster #0 shows that it contains the most restaurants with 24.

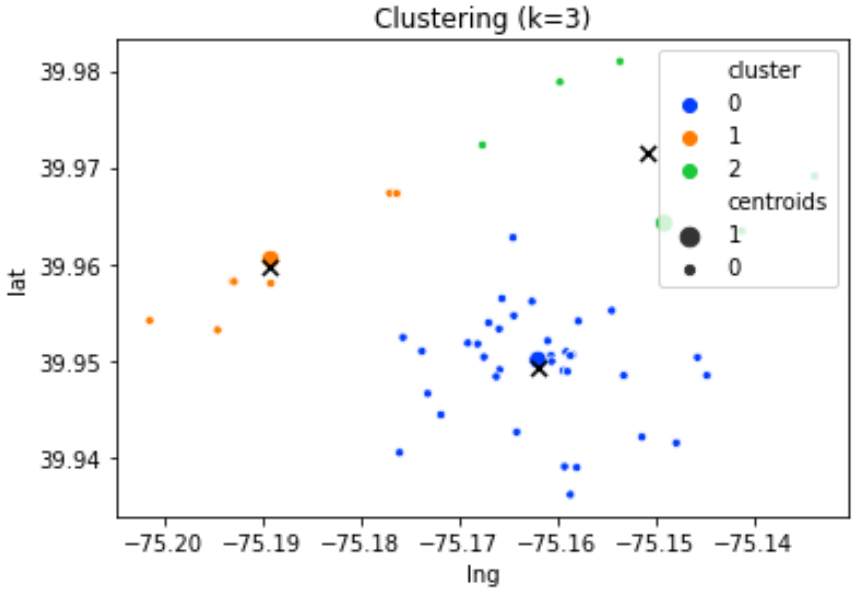
```
In [64]: pitt_df['cluster'].value_counts().max(level=0)
```

```
Out[64]: 0    24
         6     9
         1     7
         3     6
         5     2
         4     1
         2     1
         Name: cluster, dtype: int64
```

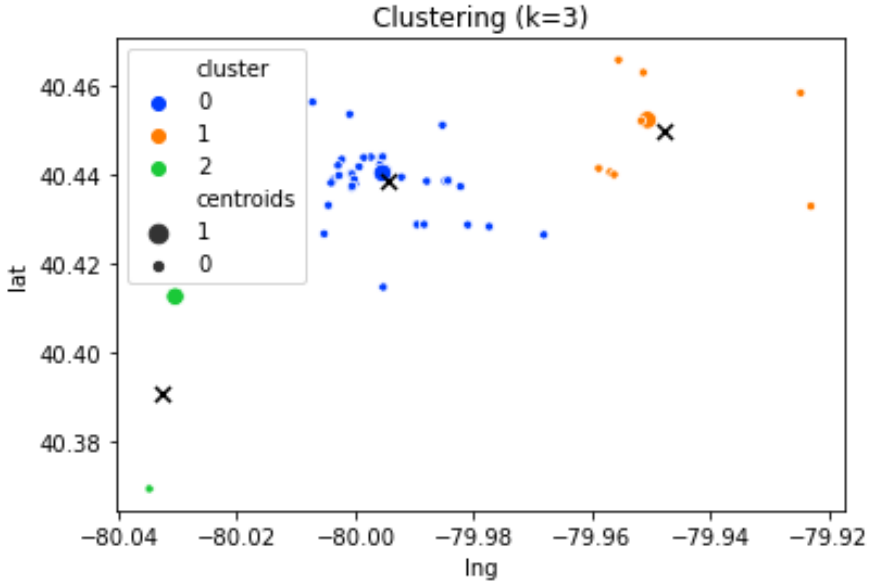
I think we can determine which city would be most beneficial to open our pizza delivery service in

- But just to be certain, let's change the **k** value to find a better value to service more customers.

Philadelphia



Pittsburgh



Which cluster had more?

```
In [70]: phil_df['cluster'].value_counts().max(level=0)
```

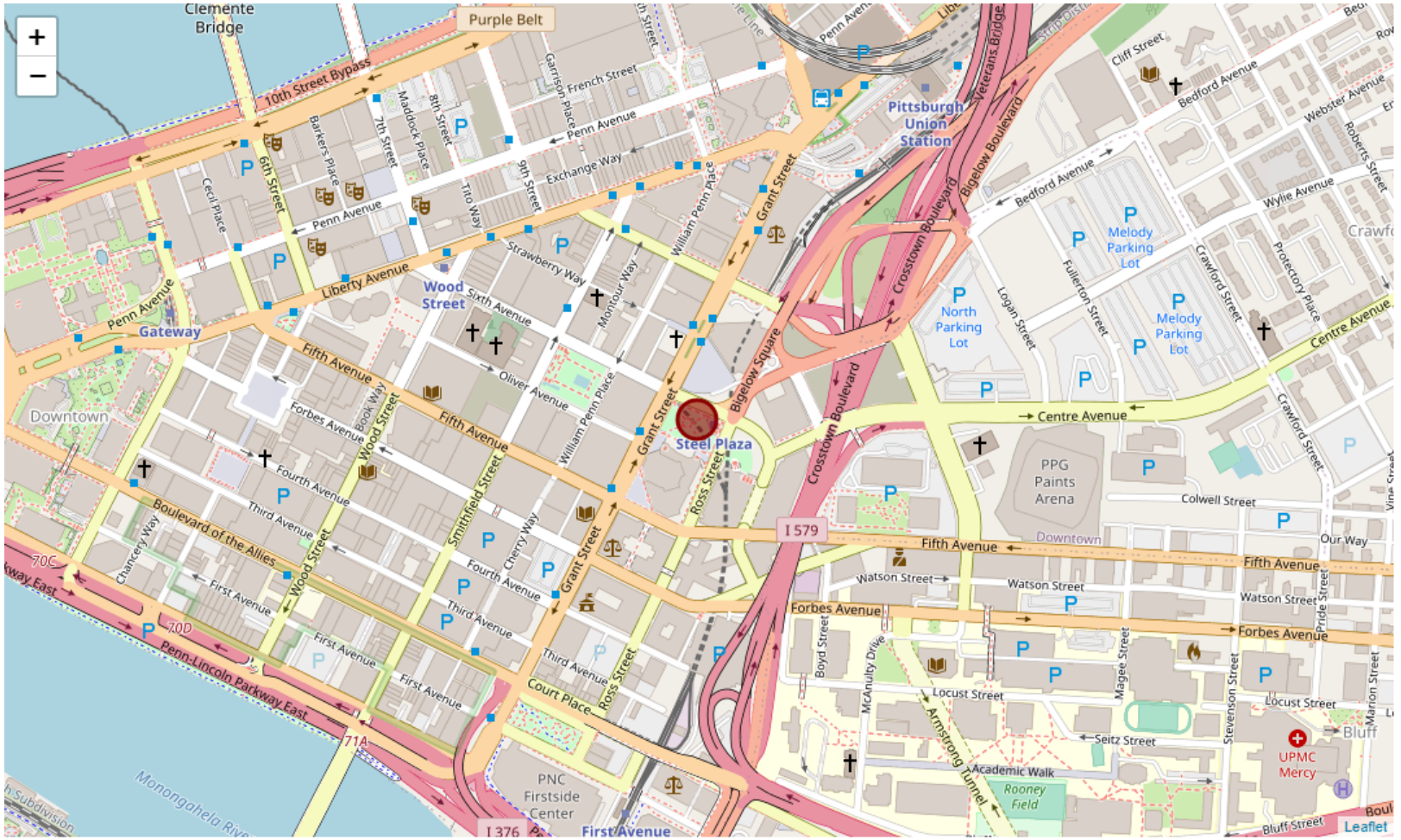
```
Out[70]: 0    36
         1     8
         2     6
         Name: cluster, dtype: int64
```

```
In [67]: pitt_df['cluster'].value_counts().max(level=0)
```

```
Out[67]: 0    39
         1     9
         2     2
         Name: cluster, dtype: int64
```

Our Pittsburgh cluster had the most restaurants in it, with 39.

We would want to open our pizza delivery service as close to these coordinates as possible:



(40.440394, -79.995415)