

Lauren Gray
Cousera
IBM Data Professional Capstone Project
Feb 2020

Introduction/Business Problem

A client is looking to open a coffee shop in Pittsburgh, PA. There are many different neighborhoods in Pittsburgh, some are filled with hospitals, and some home to colleges and universities. Where would be the best place to open a coffee shop? Are there areas with too many coffee shops? Are there areas with no coffee shops? Are there any similar business trending in the area? Foursquare data can be used to explore answers to these questions.

Data

Dataframe in pandas was created by using an API call to Foursquare to retrieve data on venues in Pittsburgh. Geodata was obtained through [PGH Open Data](#).

Method

A call was made to Foursquare for 'Coffee.' Data was entered into a Pandas Dataframe. The results were mapped into a Folium Stamen Toner Map in blue circles. Geo data points of the 90 Pittsburgh neighborhoods were then also downloaded into a dataframe and mapped with Folium with green points. After visually exploring the placement of both sets of points on the map, I made the decision to do a second call to Foursquare with a query for 'hospital.' This data was then mapped with Folium in Yellow/orange color.

Results

The results show certain areas of Pittsburgh where there are neighborhoods with multiple hospitals located but lacking in coffee shops.

Discussion

Pittsburgh is full of hospitals all which contain many employees who could be potential customers. Lawrenceville and Oakland each have 8+ hospitals in the area with one or less coffee shops. Oakland is also home to many universities as well.

Conclusion

The best areas to look into opening a coffee shop would be areas close to hospitals where there are minimal other coffee shops and high foot traffic due to large employers(hospitals). Ideal neighborhood locations based on the data would be Lawrenceville, Oakland, and South Shore.