A Battle of Neighborhoods

Is the Pittsburgh Oakland student campus a food desert? If so, how can it be improved?

Recently, the term "Food Desert" has been to describe the phenomenon around some areas, including college campuses. People in these food deserts often assert that there is limited access to cheap grocery stores near campus, and this has measurable impacts on things like obesity, nutrition, and overall quality of life; as healthy options are limited only to those who can afford them. For students, already infamous for consuming cheap ramen, this sort of food desert can pose a serious problem.

1. Introduction and Discussion of the Business Objective and Problem

Modern students wish to have access to more food options. According to students, living close to campus often means having trouble accessing healthy, affordable food. If a student can commit some time to cooking, access to an affordable grocery store may significantly increase the quality of a student's life.

"If we did solve the issue of having no grocery store, there would be no need to get a meal plan for a lot of people because they're expensive for low quality food," student's opinion from (ref 1).

In this research I would like to:

1) Verify the student's assessment of situation, i.e., Oakland neighborhood of Pittsburgh is a food desert.
2) What choices student have in current situation.
3) Examine the benefits of opening a grocery store, both on students looking to buy produce and for any potential businesses looking to set up shop.

2. The Data Science Workflow.

HTTP requests would be made to this Foursquare API server using the zip codes of the Oakland city neighborhoods to pull location information (Latitude and Longitude).

The Foursquare API search feature would be enabled to collect the locations of grocery stores nearby the neighborhood. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 700. We would limit the types of stores in output to concentrate attention on particular store types.

In the same time unsupervised machine learning algorithm K-mean clustering would be applied to form the clusters of two categories of dormitories (Dormitories from Pittsburgh University and Carnegie-Mellon University). Next step: we can combine results into one dataset.

This dataset is spread over the large area and to make output more realistic we dividing the whole dataset on several residential subclusters and make calculation for center of each subcluster separately. Using these subclusters will determine if the stores we are interested in are in reasonable walking distances. Accessible walking distance is determined by reference (2).
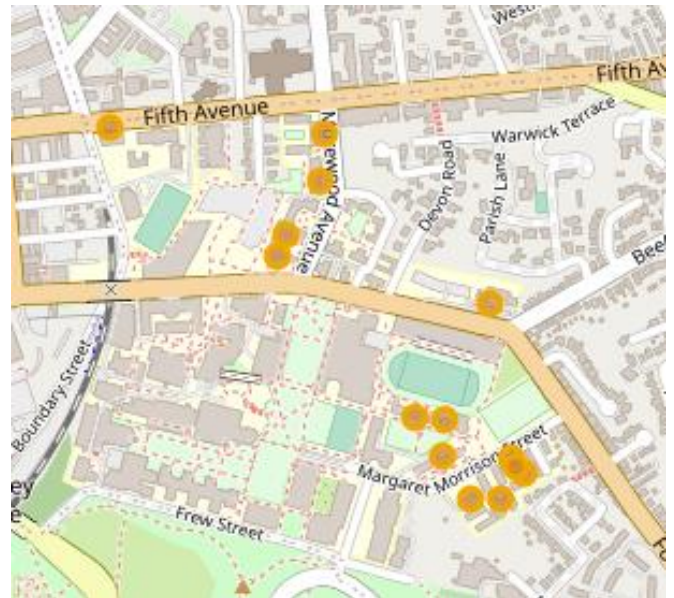
Considering the number of students in Pitt and Carnegie-Mellon and average student spending on vegetables per week (3) we can estimate total sales volume for this "new" store and determine if such an enterprise makes sense.

3. Maps of Residential Halls and Stores

It was found that colleges residents halls (from Pittsburgh University and Carnegie Mellon University) are really spread over the Oakland. Maps for dormitories blue circles (Pitt) and orange (Carnegie Mellon).
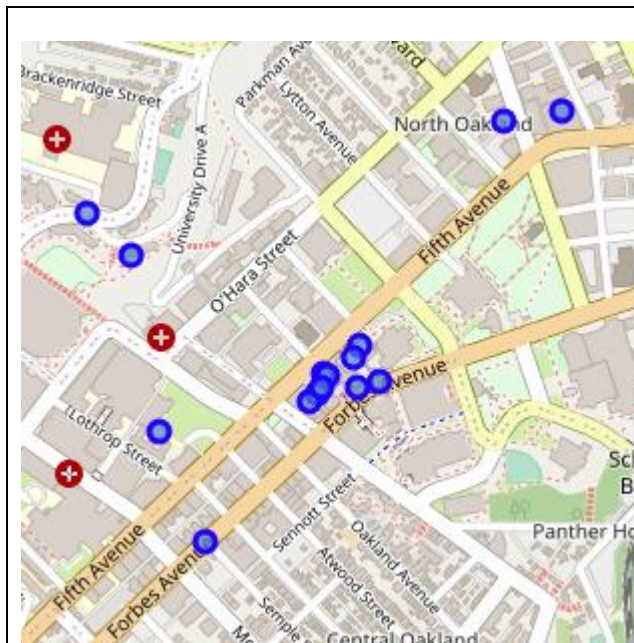

| Pitt University Residency Halls (central part)	Carnegie Mellon University Residency Halls

All (Pitt + Carnegie Mellon) residency halls have been combined into one set. To present data in more effective way it was decided to divide this combined data set (green points total 29) into clusters (K- neighbors) – blue stars. Several clusters numbers have been tried and, finally n_clusters=3 have been chosen.
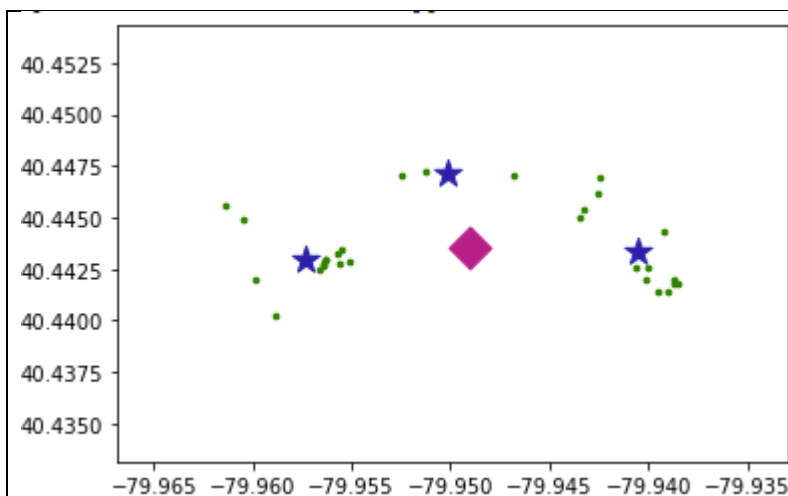
For next step: selecting all grocery store at particular radius center of whole set was selected Purple diamond.

Grocery stores around center of whole residential hall dataset have been searched using particular restriction on Foursquare code for Grocery Store "4bf58dd8d48988d118951735" to limit output ventures to reasonable number. In this case output includes all type of objects

Stores considered in radii 3.0 km from center of the whole residency set.
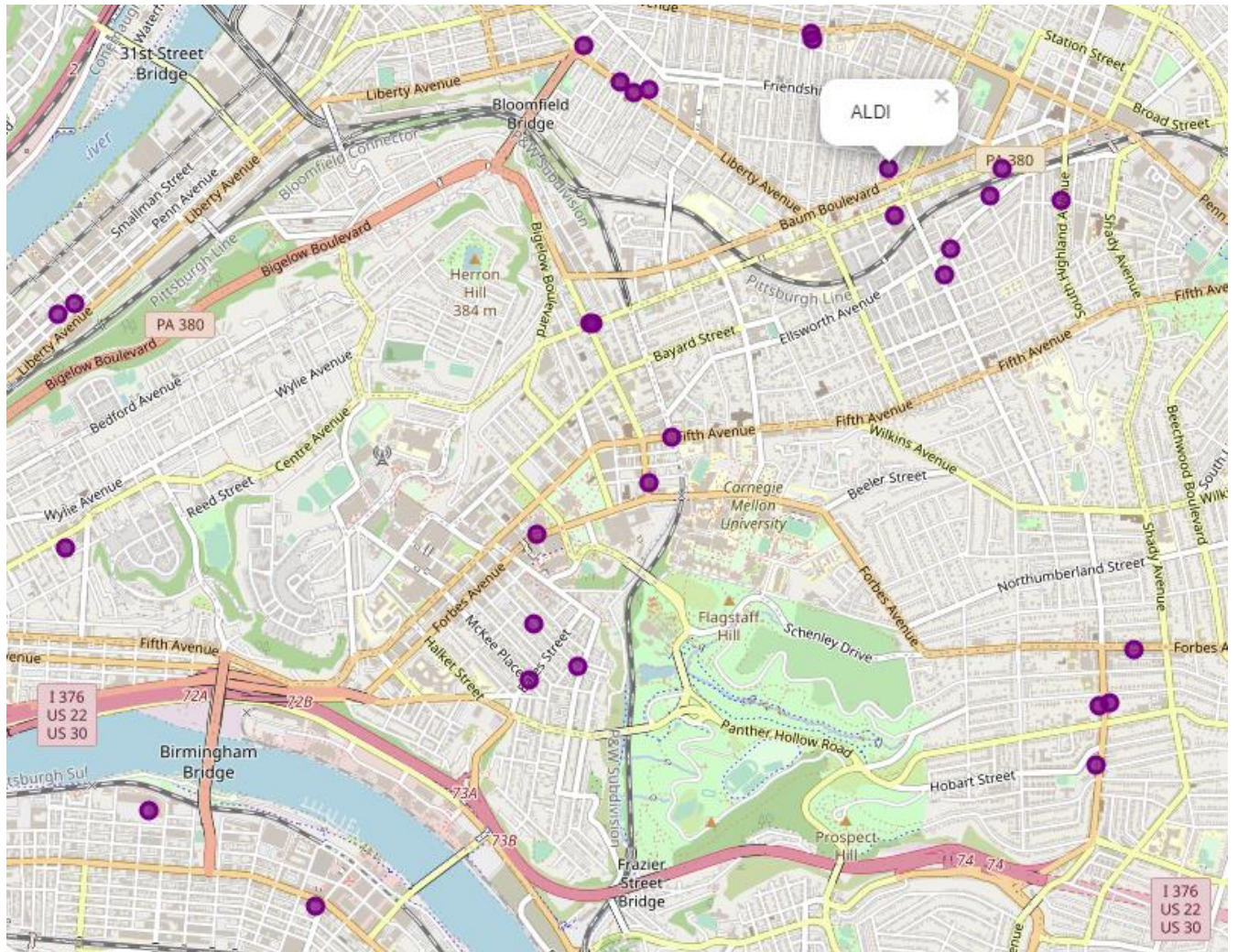
Distances are shown in a table for each store to the center to residential subcluster (Distance1, Distance2 and Distance3 correspondingly. Next step is to create a function which calculates distance between stores and subclusters centers (see Jupyter notebook)

4. Results

From first glance students have some choices (at least for 2 out of 3 clasters).
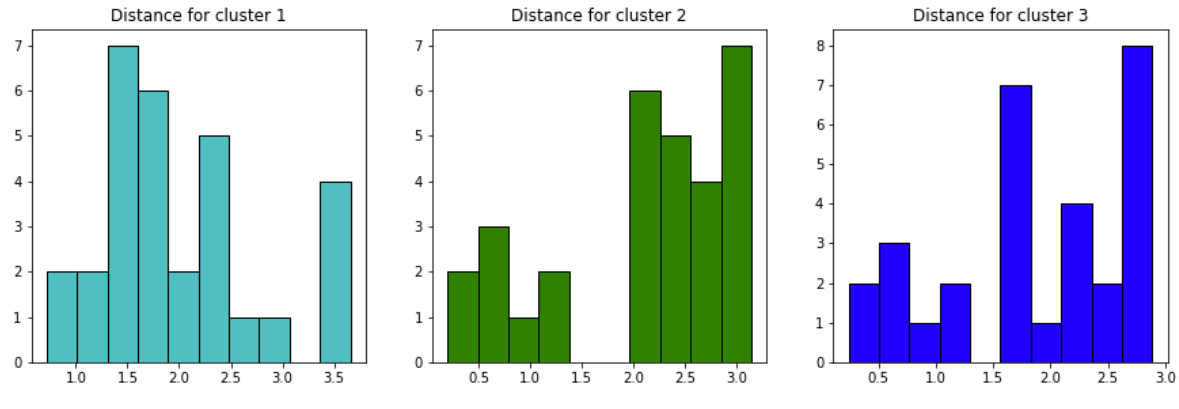
However, if we analyze what kind of stores are close to residential clusters then picture is not so bright.

Store Name	Distance1	Distance2	Distance3
Las Palmas 2	1.327975	0.471583	0.988822
Kohli's Indian Groceries	0.715185	0.773415	0.257046
Whole Foods Market	1.974436	2.986873	2.229191
Seoul Mart	0.716726	0.956715	0.237151
ALDI	1.762701	2.565146	1.811791
Bombay Food Market	1.356590	1.105445	0.559568
Forbes Street Market	1.225021	0.198352	0.619682
Donatelli's Italian Food Center	2.201166	2.208588	1.647775
ALDI	2.347614	2.824773	2.129996
Oakland Mini Market	1.231662	0.750943	1.118326
ALDI	2.927489	1.984998	2.683571
Groceria Italiana	2.188782	2.242953	1.664177
Market District Supermarket	1.556009	2.440281	1.682335
S & D Polish Deli	3.657603	2.340364	2.760410
Murray Avenue Kosher	1.702014	3.003147	2.621548
Tokyo Japanese Store	1.833952	2.865285	2.108694

Store Name	Distance1	Distance2	Distance3
Merante Groceria	1.467014	0.724521	1.247984
Sultan Bey	1.364247	1.096482	0.559960
WFH Oriental Market | Wing Fat Hong	3.603276	2.301148	2.701576
Centre Ave Shop N Save	3.490205	2.067533	2.722389
ALDI	1.397531	2.478500	1.729673
Fuel On	1.518976	2.566369	1.812981
Green Market	2.274671	2.238652	1.698694
Bottom Dollar	2.324191	2.808702	2.111720
Giant Eagle Supermarket	1.744074	3.053141	2.664249
Panda Supermarket	1.741149	3.107215	2.657395
Giant Eagle Supermarket	3.375393	2.129016	2.887022
Engel's Market	1.857324	3.085169	2.765936
Shur Save	2.497859	2.364941	1.876770
R. Market	2.015972	3.142916	2.392073

Code defined in Foursquare API includes results not just for grocery stores but also for restaurants (bakeries) which sell food as secondary business. They can provide some outstanding ethnic

food, but they are significantly more expensive than ordinary grocery stores. If one does not care about the price tag, a lot of food is still available near campus, but for a college student who is expected to be buried in debt it is not the most viable choice(Forbes Street Market is around 25% more expensive than average grocery store)[4]. Expect to waste half a day waiting for a bus (both directions) or else prepare to walk for upwards of an hour total with a backpack to inexpensive stores which are little bit far. For most students, especially those with jobs outside of studying, it is waste of precious time. However, this situation can be also considered an opportunity for any discount store to open at new location at somewhere between cluster centers.

5. Limits and Suggestions for Future Research.

More sophisticated data processing can be applied to process html file from website of each store to understand how expensive this particular store is for analyzing. However it requires access to big data tools and resources to analyze prices for thousands of products.

We do not consider real walking distances (e.g. tracks from google map), but straight lines between dorms and stores. In more precise approach it should be automated generation of walking distances in (for example Google maps).

6. Conclusion

For students originating from families with a moderate income, the university campus and surrounding area do not provide easy or cheap food options. However, considering the total number of students ~34,000 (Pitt) and ~ 13,000 (Carnegie-Mellon) and assuming that half of them will use a new hypothetical store in the neighborhood (we can make opening such store economically attractive. Middle size grocery store with `8,000-14,000 sq. feet of retail area and gross sales volume of 600$ per sf/per year which around average ($11-12/per sf per week) for grocery food sale industry, would be profitable. For example, Aldi, Trader Joe, Save-A-Lot, or other stores in same price range can be a good fit for this location. But we should consider its dependency from regulations which are out of scope of this work.

References.

1) (Publication in Pittsburgh Newspaper, August 29, 2018)https://www.pghcitypaper.com/pittsburgh/food-deserts-oakland-and-downtown-make-healthy-eating-a-challenge-for-pittsburgh-students/Content?oid=10300727

2) An Examination of College Students' Produce Consumption and Purchasing Behavior: A Case Study in California Walker, Bernadette Marie; McGarry Wolf, Marianne; Schroeter, Christiane Journal of Food Distribution Research, 40, 1 2009-03 Page range: 180-186

3) Walking Distance by Trip Purpose and Population Subgroups. Yong Yang, PhD and Ana V. Diez-Roux, PhD, MD Am J Prev. Med. 2012 Jul; 43(1): 11–19.

4) Forbes Street Market: Is It Worth Shopping There? By Erin Lalli Pitt Contributor • Lifestyle November 13, 2018
https://www.hercampus.com/school/pitt/forbes-street-market-it-worth-shopping-there

5) 27 Retail Grocery Industry Statistics, Trends & Analysis. Oct 6, 2018 by Brandon Gaille.
https://brandongaille.com/27-retail-grocery-industry-statistics-trends-analysis/