

Analysis of Current Market Competition for Conceptual Pizzeria/Bar in the Greater Pittsburgh, PA Region

1. Introduction

1.1. Background & Business Case

The market for pizza restaurants within the midsize city of Pittsburgh, Pennsylvania and its suburbs is significantly expansive. The city, in fact, has the second most pizzerias per capita within the United States at 9.9 per 10,000 people [1]. Coupled with the fact that the city contains the largest amount of bars per capita in the entire country at 11.8 per 10,000 people [1], this overall, high market interest is appealing to any prospective entrepreneur.

In a market which on the surface appears to be significantly saturated, however, it is crucial to delve deeper into what contributes to successful pizzerias/bars and what especially sets them apart from others found in more densely populated cities. As location ultimately plays a heavy factor in any successful business, the focus of this report is to provide prospective owners of a conceptual pizzeria/bar restaurant with a deliverable of a detailed city map, highlighting locations around Pittsburgh where the greatest competition currently resides. By focusing specifically on the city suburbs with the largest number of venues, the areas containing the heaviest concentration of both pizzerias and bars will be clearly identified. Finally, a detailed listing of each of the (25) major neighborhoods in scope and their respective (10) most common venues will be tabulated to provide a more complete picture of the landscape—thus allowing prospective owners to capitalize on the current market, geographically (while weighing the amount of risk to take with respect to competition).

[1] Clark, Jayne. "Where the Bars Are: Top U.S. Cities" *USA Today*, Gannett Satellite Information Network, 6-Dec-13, <www.usatoday.com/story/dispatches/2013/12/06/top-bar-and-pizza-cities/3882089/>

2. Data Acquisition

2.1. Data Sources

The first set of data utilized in this report consists of a comprehensive listing of Pittsburgh postal codes within Allegheny County (the county within which the city resides). This information can be found, split into 4 separate tables, on the following website: [http://www.ciclt.net/sn/clt/capitolimpact/gw_ziplist.aspx?FIPS=42003].

The second set of retrievable data, detailed further in **Section 2.2**, comes from the open data API for latitudinal and longitudinal coordinates by postal code. This API can be found at the following url, where {zip} = each desired postal code: [<https://public.opendatasoft.com/api/records/1.0/search/?dataset=us-zip-code-latitude-and-longitude&q=zip%3D{zip}>].

The final, key datasets required for compiling the competitor summary come from utilizing the Foursquare API—through which all relevant Pittsburgh venues by neighborhood have been pulled. A more detailed description with regards to how the Foursquare API input/output is optimized is found in **Section 2.2**.

2.2. Data Cleaning

Iterating over the full list of Allegheny County postal codes, the latitudinal/longitudinal data API was called for each individual zip code, thereby mapping to its respective coordinate. The open API provides necessary precision for each neighborhood to be accurately plotted onto a resultant map using Folium (i.e. latitude coordinates to 6 decimals and longitude coordinates to 5 decimals).

Next, the exhaustive list of city suburbs pulled out via the postal codes' website was cleaned by dropping rows containing:

- (i.) duplicated latitudinal/longitudinal coordinates
- (ii.) spelling variants referring to the same zip code and coordinates—i.e. Mount Washington vs. Mt. Washington
- (iii.) any bulk mail centers the open API returned with null coordinates

3. Methodology

Following cleaning of the dataset, the list of city suburbs—referred to in Foursquare as “Neighborhoods”—were successfully filtered to contain all (100+) legitimate areas within the region. It was, however, important to consolidate this down further to only the top (25) neighborhoods containing the greatest number of venue entries within the Foursquare database. More specifically, such neighborhoods were defined as those around Pittsburgh, containing (7) or more existing venues; this allowed for statistically significant clustering via k-means to occur. Otherwise, rural areas in Allegheny County, many with only 1 or 2 venues, inevitably pull in null venues when a one-hot coding approach is employed to filter out the make-up of top venues per region. For example, the Foursquare venue category “ATM” is automatically pulled in as the first null value in each neighborhood due to its initial positioning alphabetically. This is not representative and would skew overall venue data to a large cluster of rural areas inaccurately displaying “ATM”, among other various categories, as one of the top venues. Thus, the aforementioned (25) neighborhoods containing the most venues (as of March 2021)—opposed to neighborhoods based on largest population or other arbitrarily set metrics—were chosen as most representative.

The Foursquare API was then used to display these (25) neighborhoods' top (10) most common venues to provide a comprehensive snapshot of the entire region's major markets. The last step for visualization and in order to provide the required customer competition overview, clustering the top venue findings was required. This was accomplished via k-means clustering of the one-hot coded overall venue means, which prominently displayed the specific pizzeria/bar grouping desired by a prospective owner. Note that neighborhoods included within this clustering on the Folium map, found in the results in **Section 4**, are designated by red markers (with the other clusters uniquely colored).

The k-value chosen for clustering was k=6. Implementing six unique clusters provided optimal sub-classification, uniquely identifying areas with a top pizzeria + bar split, which is what was sought after from the Folium-generated map deliverable. The red, pizza/bar grouping is labeled as 'Cluster 4'. A final summary read-out of 'Cluster 0' through 'Cluster 5' with their defining features is provided in **Section 4**.

4. Results

In **Figure 1**, below, the Folium-generated map displaying the complete list of Allegheny County suburbs by postal codes can be found highlighted with blue markers; this map served as initial basis.

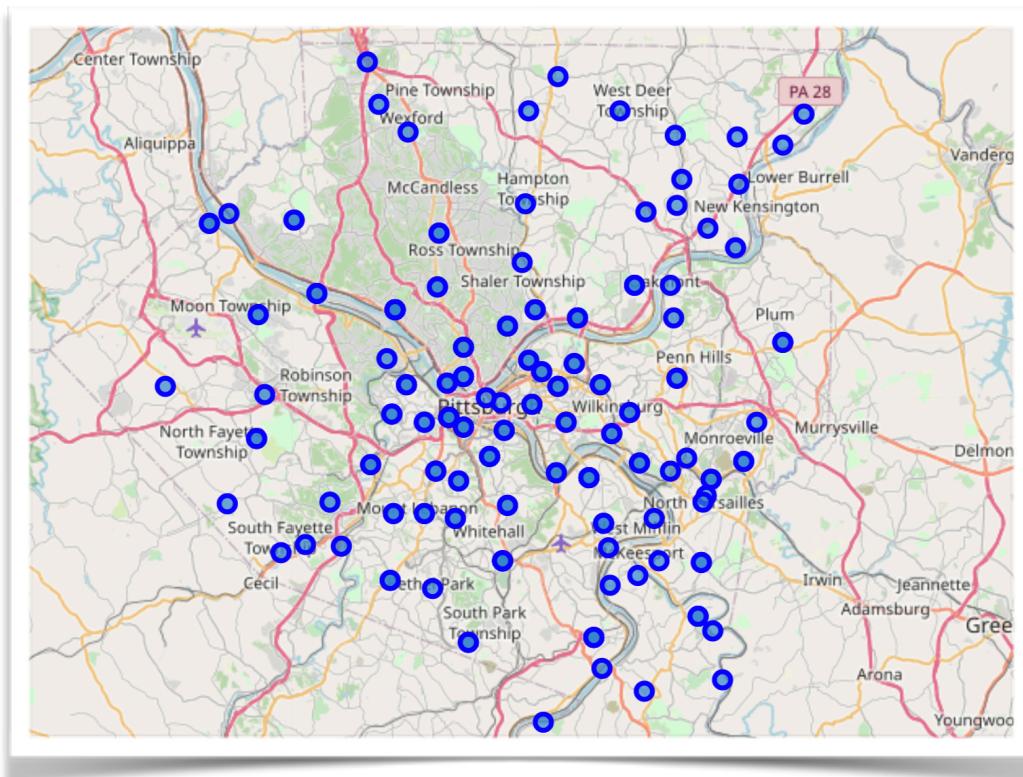


Figure 1 - Allegheny County Map: Overlay of Pittsburgh Suburbs by Postal Code

Next, **Figure 2**, found below, displays the Pittsburgh suburbs grouped into the earlier described (6) clusters. Area Cluster 4, with the greatest number of pizzeria and bar venues, is highlighted in red markers. **Tables 1-6** on the following pages provide further detail regarding each cluster.

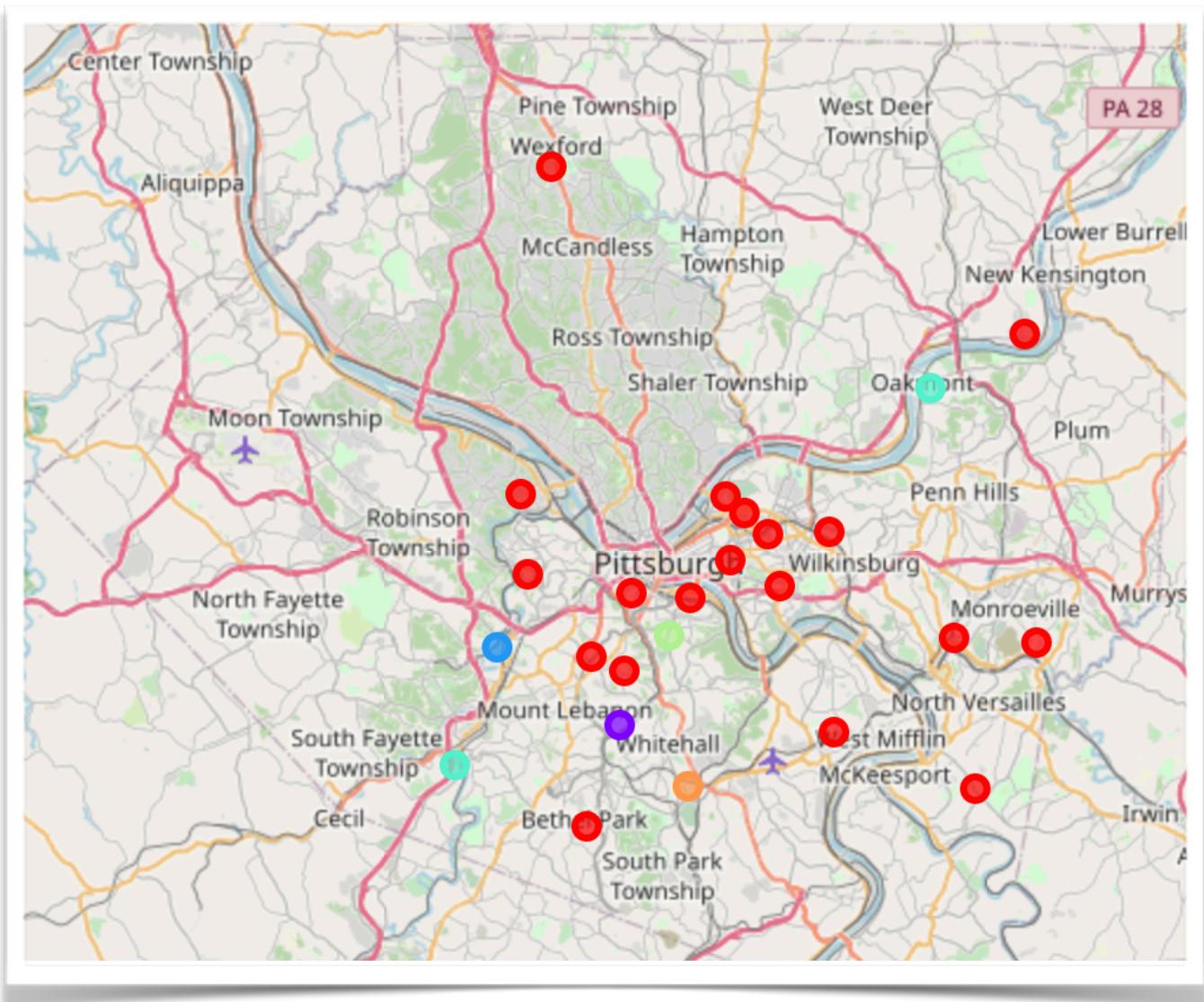


Figure 2 - Allegheny County Suburbs: K-Means Clustered by Top Foursquare Venues

Top venue listings, displayed by cluster, providing the scope of the existing market and marker colors are found below (continuing onto the following page), **Tables 1-6**:

City	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
Carnegie	Italian Restaurant	Library	Bar	Bank	Food & Drink Shop	Pizza Place	Miscellaneous Shop	Theater	Gift Shop	Other Great Outdoors

Table 1 - Cluster Label 0: Top Foursquare Venues (Marker Color: Dark Blue)

City	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
Bridgeville	Gas Station	American Restaurant	Sandwich Place	Chinese Restaurant	Coffee Shop	Mexican Restaurant	Performing Arts Venue	Pharmacy	Hotel	Post Office
Oakmont	Yoga Studio	Bakery	American Restaurant	Ice Cream Shop	Hobby Shop	Spa	Bank	Italian Restaurant	Video Game Store	Public Art

Table 2 - Cluster Label 1: Top Foursquare Venues (Gas, Yoga, Bakery, Marker Color: Cyan)

City	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
Mt Oliver	Fish & Chips Shop	Convenience Store	Fast Food Restaurant	American Restaurant	Fried Chicken Joint	Video Store	Hardware Store	Pizza Place	Performing Arts Venue	Park

Table 3 - Cluster Label 2: Top Foursquare Venues (Fish & Chips, Marker Color: Light Green)

City	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
Caste Village	Baseball Field	Chinese Restaurant	Bar	Bakery	Convenience Store	Grocery Store	Laundromat	Other Repair Shop	Other Great Outdoors	Outdoors & Recreation

Table 4 - Cluster Label 3: Top Foursquare Venues (Baseball Field, Marker Color: Orange)

City	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
Wexford	Fast Food Restaurant	Deli / Bodega	Bank	Moving Target	Bakery	Taco Place	Cosmetics Shop	ATM	Paper / Office Supplies Store	Performing Arts Venue
Bethel Park	Breakfast Spot	Sandwich Place	Bar	Gym	Pizza Place	Home Service	Mexican Restaurant	Beer Garden	Ice Cream Shop	Other Repair Shop
Pittsburgh	Bar	Pizza Place	Coffee Shop	Sandwich Place	American Restaurant	Moving Target	Italian Restaurant	Boat or Ferry	Tunnel	Bakery
McKeesport	Bar	Moving Target	Pizza Place	Boat or Ferry	Tunnel	Disc Golf	Optical Shop	Coffee Shop	Gym / Fitness Center	Shipping Store
McKees Rocks	Pizza Place	Breakfast Spot	Indie Movie Theater	Bakery	Pharmacy	Café	Spa	Bar	Bank	Plaza
Monroeville	Park	Pool	Construction & Landscaping	Bar	Bakery	Pub	Pizza Place	Flower Shop	Paper / Office Supplies Store	ATM
Springdale	Salon / Barbershop	Pizza Place	Discount Store	General Travel	Soup Place	Chinese Restaurant	Construction & Landscaping	Bar	Post Office	Pool
Turtle Creek	Pizza Place	Italian Restaurant	Bank	Convenience Store	Pharmacy	Smoke Shop	Garden Center	Discount Store	Astrologer	Library
Arsenal	Bar	Sandwich Place	Taco Place	Arts & Crafts Store	Karaoke Bar	Indie Movie Theater	Bus Station	Sports Bar	Coffee Shop	Gym
Carson	Bar	American Restaurant	Asian Restaurant	Sushi Restaurant	Thai Restaurant	Ice Cream Shop	Pub	Burger Joint	Dance Studio	Boutique
Crafton	Pizza Place	Pharmacy	Discount Store	Supermarket	Chinese Restaurant	Grocery Store	Fried Chicken Joint	Shoe Store	Bar	Bank
Homewood	Food	Bus Station	Café	Music Venue	Bar	Bank	Construction & Landscaping	Convenience Store	Gym / Fitness Center	Performing Arts Venue
Mount Washington	Baseball Field	Diner	Pharmacy	Sandwich Place	Bar	Bank	Bakery	Breakfast Spot	Scenic Lookout	Pizza Place
Oakland	Pizza Place	Coffee Shop	Chinese Restaurant	Bagel Shop	Burger Joint	Indian Restaurant	Cosmetics Shop	Museum	New American Restaurant	Dive Bar
South Hills	Pizza Place	Deli / Bodega	Movie Theater	Light Rail Station	Plaza	Park	Bar	Pet Store	Noodle House	Optical Shop
Squirrel Hill	Pizza Place	Coffee Shop	Dumpling Restaurant	Bagel Shop	Tea Room	Thai Restaurant	Toy / Game Store	Marijuana Dispensary	Sports Bar	Greek Restaurant
Bloomfield	Pizza Place	Coffee Shop	Grocery Store	Bar	Thai Restaurant	Art Gallery	Bookstore	Sandwich Place	New American Restaurant	Burger Joint
Brookline	Pizza Place	Italian Restaurant	Liquor Store	Spa	Grocery Store	Bike Shop	Convenience Store	Bakery	Lounge	Mediterranean Restaurant
Shadyside	Clothing Store	Coffee Shop	Asian Restaurant	Gym	Bakery	Japanese Restaurant	Salon / Barbershop	Bar	Gay Bar	Mexican Restaurant

Table 5 - Cluster Label 4: Top Foursquare Venues (Bar/Pizzeria, Marker Color: Red)

City	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
Castle Shannon	Dessert Shop	Light Rail Station	Coffee Shop	Train Station	Bakery	Playground	Bar	Print Shop	Pub	Noodle House

Table 6 - Cluster Label 5: Top Foursquare Venues (Dessert Shop, Marker Color: Purple)

5. Discussion

Analyzing top venue data gleaned from the Foursquare tables and interpreting the resultant clustered map, the following details emerged: areas nearest the city center in downtown Pittsburgh, and generally to the east and south of the city, showed the greatest concentration of current competition for the conceptual pizzeria/bar. Perhaps more interestingly, there is a sizable area north of the city where fewer overall venues were shown to exist, including a lesser amount of pizza restaurants and bars within the county. This area's population is by no means rural and is on par with many of the other suburbs south of the city, indicating that an unmet overall market may be present. Additional information essential to prospective owners—other top venues, both intra- and inter-Cluster 4, were shown to contain the following: Chinese/American restaurants, coffee shops/cafés, delis, diners, bakeries, yoga studios, gas stations, baseball fields, and parks.

6. Conclusion

Upon analyzing the market through a thoroughly refined review of venues provided by Foursquare data, the ideal location for prospective owners seeking to capitalize on the relatively most-underserved area for pizzerias/bars is the North Hills region of Pittsburgh—spanning from the confluence of the three rivers downtown, northward to the city limits of Wexford.

Alternatively, if prospective owners are driven by interest in investing within the area containing the largest concentration of existing competition, a location recommendation of the following is supported from statistical review: the approximate 10 mile radius to the south, east, and west of downtown Pittsburgh—with epicenter being just southeast of the city center.

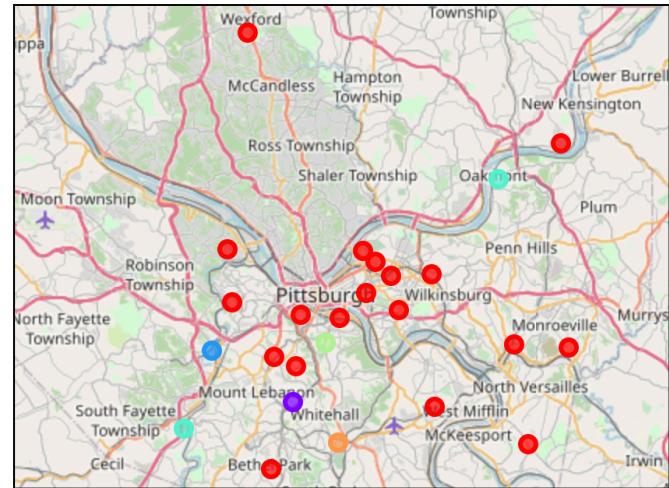


Figure 3: Pittsburgh Map - Snapshot Summary