Data-Driven Advice on choosing Location and Type of a New Restaurant in Berlin

Paul Seibel, March 2021

Introduction

This Project aims to show a technique to leverage geospatial data from the Foursquare API to solve a fictional Business Problem. This technique could be used anywhere without feeding it specific data and for any kind of venue.

Business Problem

Our client wants to open a new restaurant in Berlin. They want data-driven advice as to where and what kind of restaurant would be lucrative pick. This client already runs several restaurants in major European cities and has a business strategy in place which will dictate our basic methodology.

This strategy is relatively simple: Go to an area where you can be sure there will be lots of paying customers and offer something that is rare and unique in this place.

Methodology

Based on this strategy we state two presumptions:

- **The more restaurants in the area the better**
  even though there will be high competition in these areas, it is proven that a restaurant can be profitable there
- **The more unique the kind of restaurant the better**
  to deal with high competition there should be something to set our clients restaurant apart from the rest of the market. There is a constraint on that rule though:
  There should be **at least a single instance** of any suggested category so we have some proof that this kind of food can find buyers in this city but **the farer away the better** so the competition inside of the niche doesn't get to close.

So this determines the goal of our analysis: **Identify areas that have lots of restaurants and find out what kinds of restaurants would be unique there**.

Data

The specific locations of the areas to analyse will not come from an existing data since coordinates from existing units of Berlin (like Postal Code Areas, Administrative Districts, ...) are unevenly distributed over the city and have weird shapes. Instead I will calculate a matrix of points over a chosen area and analyse and compare the areas around them with each other.

The data about existing restaurants in these areas will come from the Foursquare API as required by the assignment. I will use the `/venues/search` endpoint and filter by "Food"-categories to get all restaurants in the area. I should end up with a grid of areas over the city (including Latitude, Longitude and an Area ID) and a list of venues for each of them (the venue data to work with will be the *category* from the Foursquare Database).

A big advantage of this approach is that it is not bound to any Berlin-specific data. Location and size of the analysed area are variable and would work anywhere (as long as Foursquare can deliver meaningful venue data for the chosen area).

Missing Data

For a business assessment our client has to include costs of running the place and at this point rent would be the most important factor. If we had some data on average rent price for commercial property per m² for each area, we could factor that in pretty easily. Unfortunately, I couldn't find data like that for free, if this was a real case and we had a budget we could buy this data.
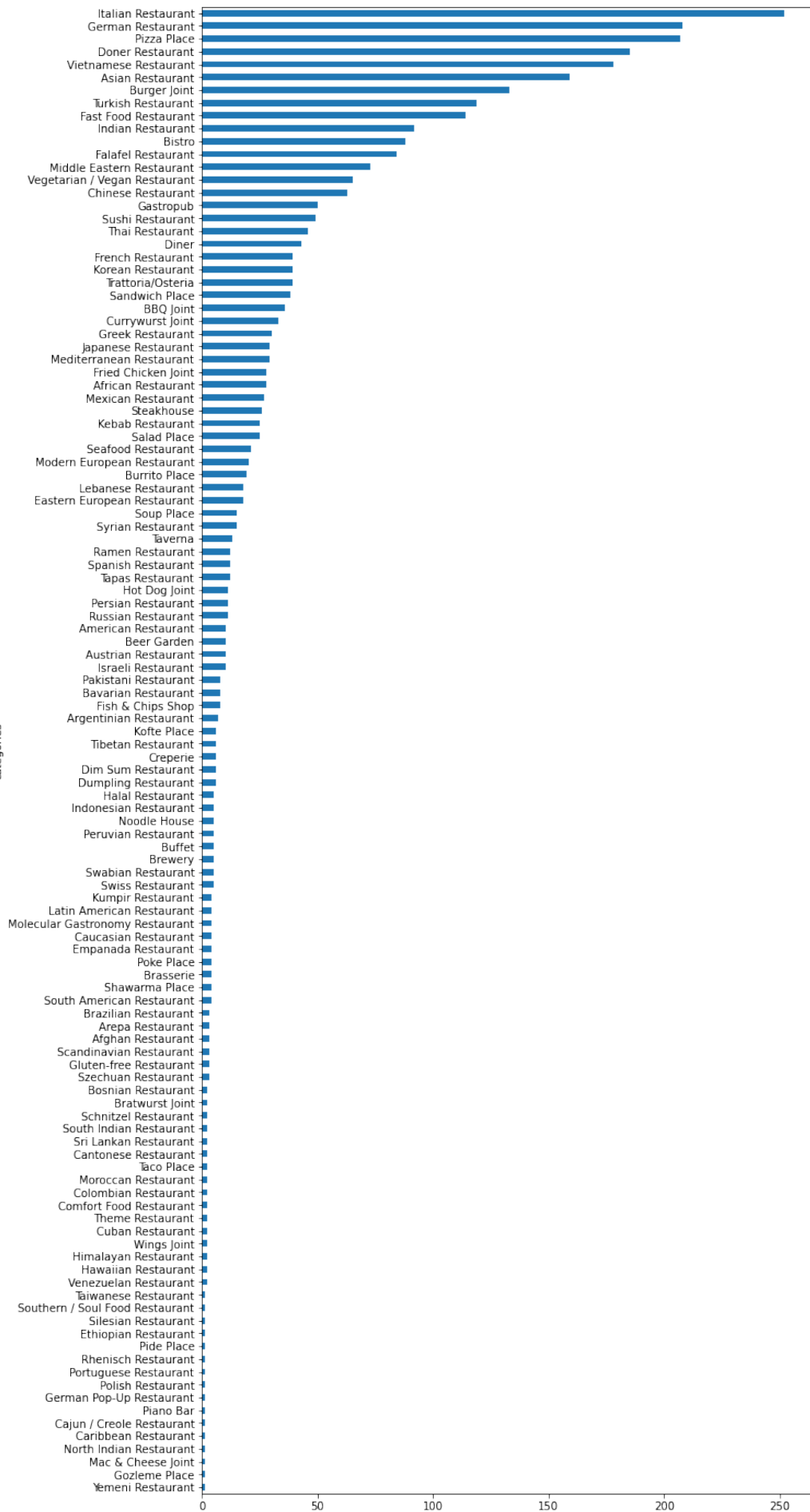
Analysis

First Step is to get the venues and map their total number per area as a color code onto the map for a first idea of distribution:
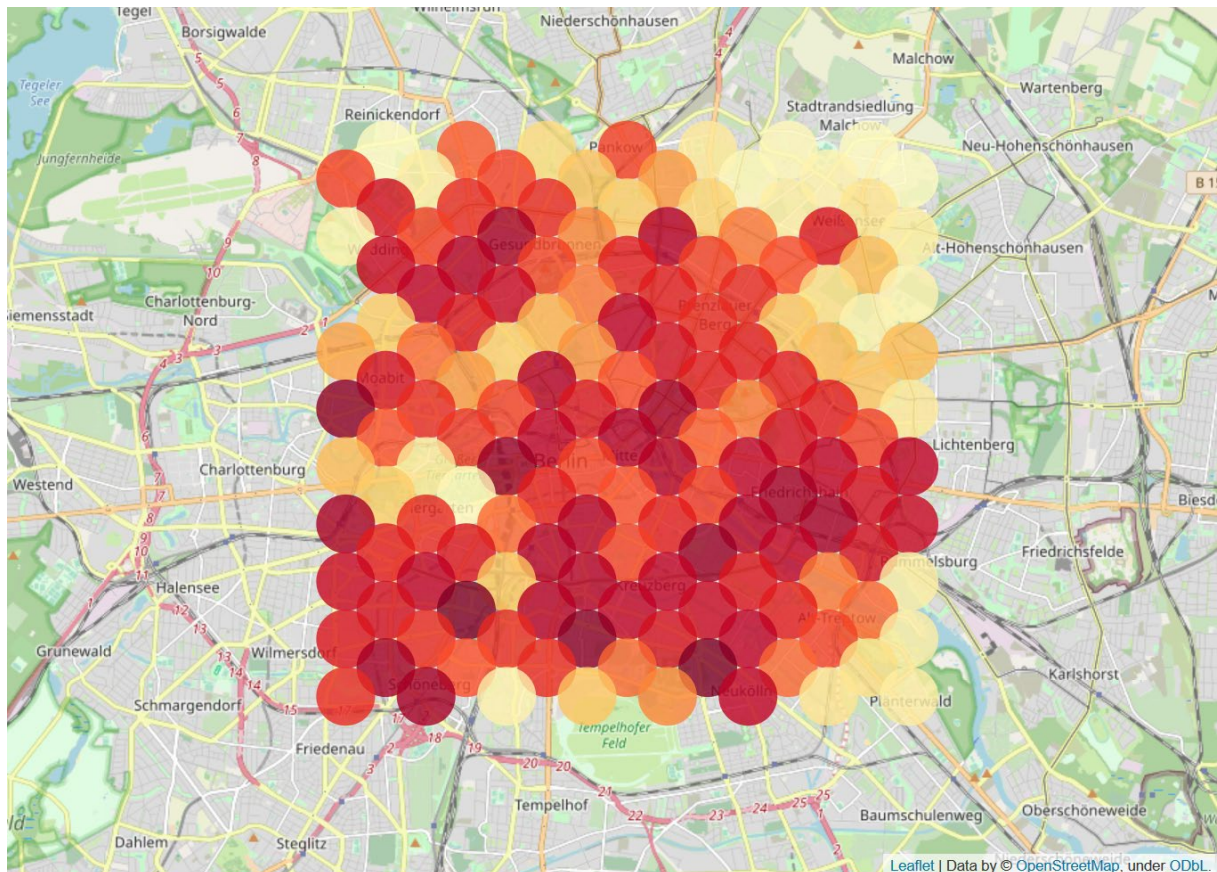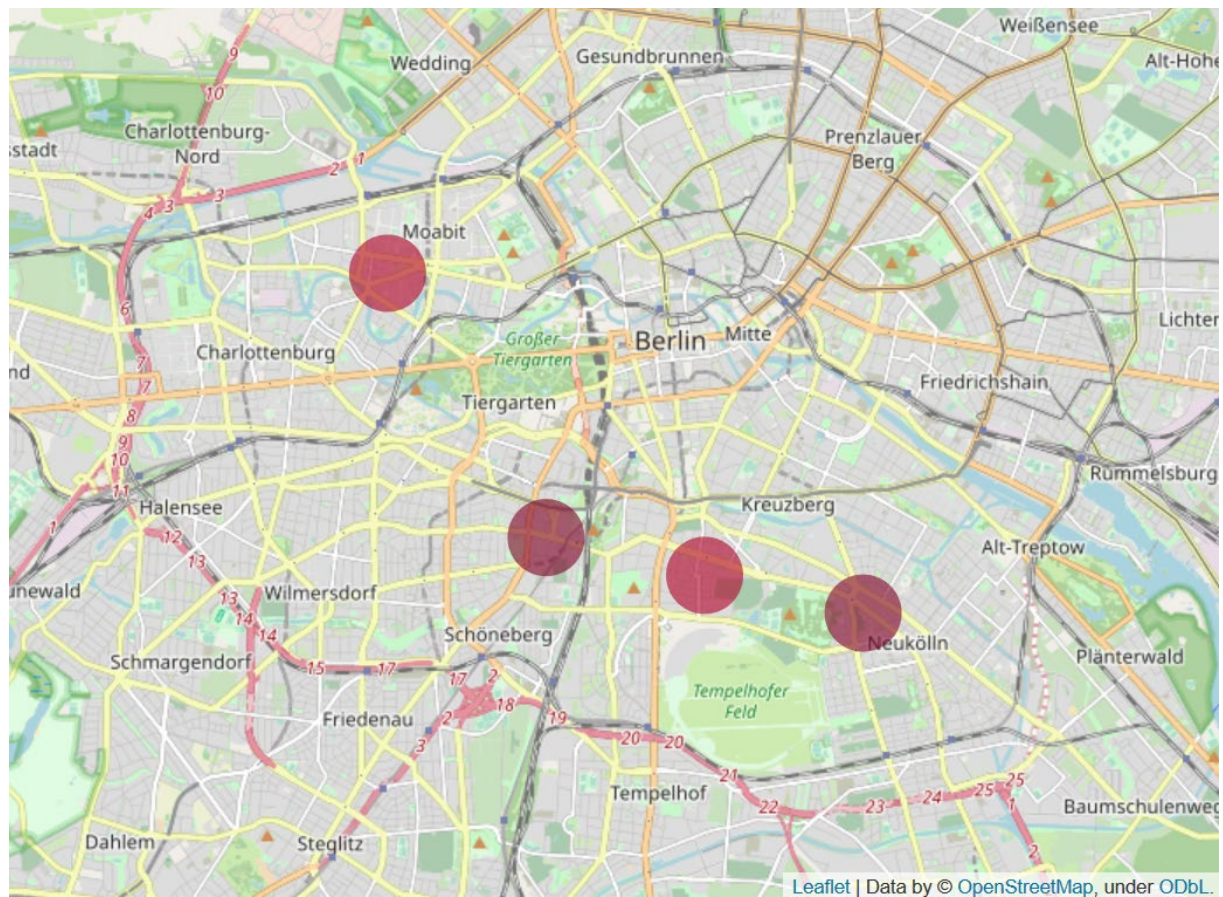

Next we provide the client with an overview all over Berlin. The following bar chart shows which kinds of restaurants are most frequent and on the opposite which are relatively unique.
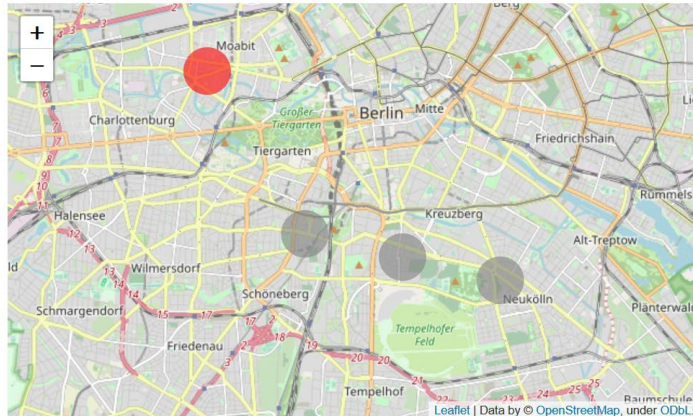
Since only the top tier areas come into question we filter out the lower 95% and up with following map of areas for further analysis:

Discussion of the results of each point
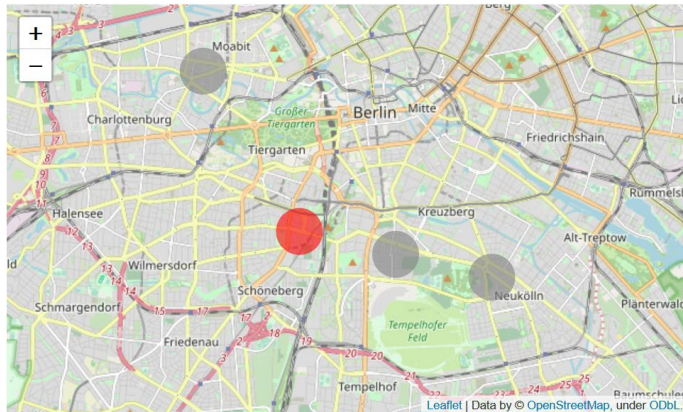
Spot 67: Moabit

Overview for Spot 67

|	Category	Shortest Distance	Avg. Distance	Amount
0	Cajun / Creole Restaurant	9.812888	9.812888	1
1	Venezuelan Restaurant	9.531717	9.586921	2
2	South American Restaurant	8.971761	9.428084	4
3	Taco Place	8.730618	8.730618	2
4	Portuguese Restaurant	8.092994	8.092994	1
5	Caribbean Restaurant	8.005869	8.005869	1
6	Silesian Restaurant	7.781005	7.781005	1
7	Mac & Cheese Joint	7.134290	7.134290	1
8	Moroccan Restaurant	7.131226	7.131226	2
9	Cantonese Restaurant	7.071257	7.705468	2
10	Kumpir Restaurant	7.050807	8.279400	4
11	Taiwanese Restaurant	6.940798	6.940798	1
12	Ethiopian Restaurant	6.135919	6.135919	1
13	Himalayan Restaurant	6.106335	6.106335	2
14	Colombian Restaurant	5.760098	7.158930	2
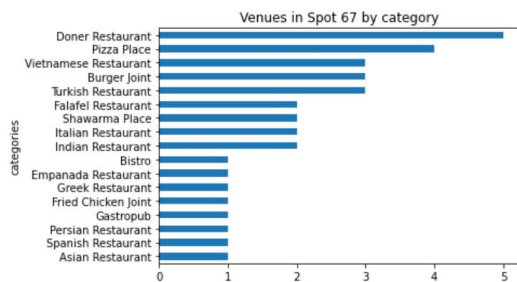
Dominated by Italian, Turkish/Arabic and Vietnamese Restaurants. But from our table we can see that this place doesn't have **Southern or Latino American restaurants** anywhere near it. So this could be combination to explore further.

Spot 121: Schöneberg

Overview for Spot 121

|	Category	Shortest Distance	Avg. Distance	Amount
0	Cajun / Creole Restaurant	8.864151	8.864151	1
1	South American Restaurant	7.247364	7.493090	4
2	Gozleme Place	7.064765	7.064765	1
3	Caribbean Restaurant	6.920378	6.920378	1
4	Venezuelan Restaurant	6.860319	7.107173	2
5	Pide Place	6.849867	6.849867	1
6	Yemeni Restaurant	6.505200	6.505200	1
7	Taco Place	6.317247	6.317247	2
8	Himalayan Restaurant	6.201434	6.201434	2
9	Mac & Cheese Joint	5.772596	5.772596	1
10	Caucasian Restaurant	5.609926	6.202171	4
11	Halal Restaurant	5.595043	6.696821	5
12	Portuguese Restaurant	5.510409	5.510409	1
13	Gluten-free Restaurant	5.423435	5.650594	3
14	Bosnian Restaurant	5.308295	6.019995	2
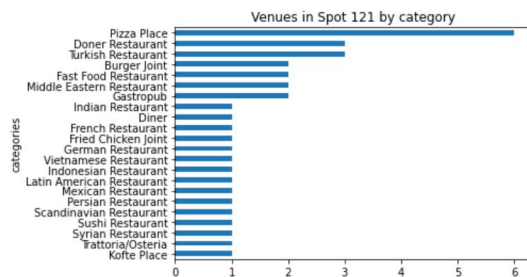
Many Pizza places but also Turkish and Burgers. It's not to far from Moabit and we have similar results with the **South American Restaurants** being far of. There are some other intersting candiates on the list like **Gluten-free or Bosnian**.

Spot 130: Kreuzberg

Overview for Spot 130
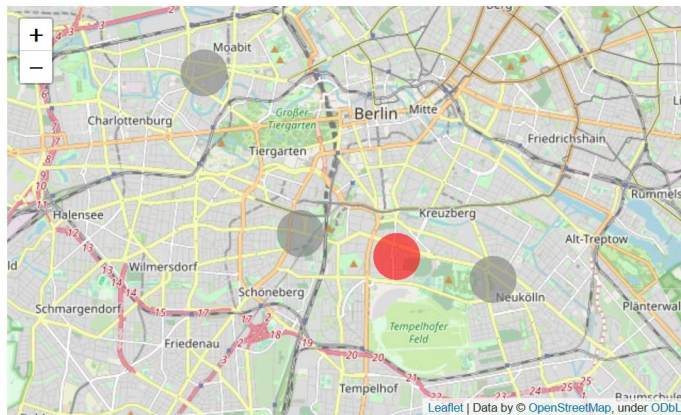
|	Category	Shortest Distance	Avg. Distance	Amount
0	Gozleme Place	8.225912	8.225912	1
1	Yemeni Restaurant	7.500651	7.500651	1
2	Cajun / Creole Restaurant	7.445228	7.445228	1
3	Pide Place	7.283960	7.283960	1
4	Halal Restaurant	6.202677	7.227516	5
5	Himalayan Restaurant	5.638975	5.638975	2
6	Caribbean Restaurant	5.597346	5.597346	1
7	South American Restaurant	5.567511	5.705910	4
8	Wings Joint	5.129000	6.301985	2
9	Hawaiian Restaurant	5.108834	5.108834	2
10	Venezuelan Restaurant	4.773746	5.119032	2
11	Polish Restaurant	4.617111	4.617111	1
12	Mac & Cheese Joint	4.450763	4.450763	1
13	Taco Place	4.384596	4.384596	2
14	Afghan Restaurant	4.298873	6.002631	3
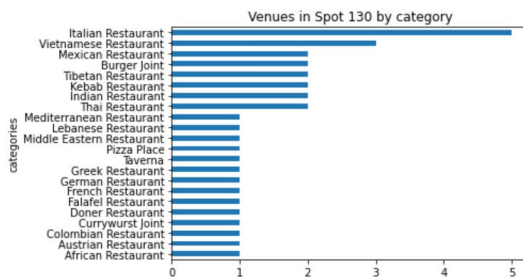
Many Italian Restaurants but also lots of Burgers and Vietnamese. There are also lot's of diversity from Mexican over Greek to African, you can get it here. Still our result poses some interesting possibilites for something unique. Except for the obligatory Döner/Kebab places there are surprisingly few Turkish/Middle-Eastern places here. The next **Gözleme place** is over 8km away! Restaurents specializing in **Wings** or **Hawaiian** would also be unique for this area.

Spot 139: Neuköln

Overview for Spot 139
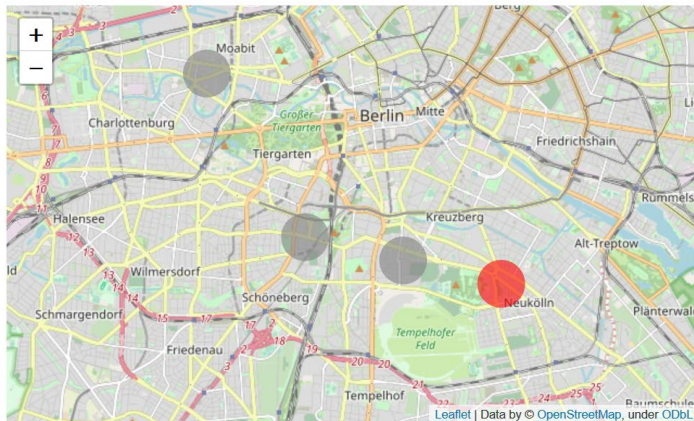
	Category	Shortest Distance	Avg. Distance	Amount
0	Gozleme Place	9.736320	9.736320	1
1	Yemeni Restaurant	8.920628	8.920628	1
2	Pide Place	8.280401	8.280401	1
3	Halal Restaurant	7.417417	8.305867	5
4	Hawaiian Restaurant	6.957212	6.957212	2
5	Wings Joint	6.907672	7.955858	2
6	Cajun / Creole Restaurant	6.454600	6.454600	1
7	Theme Restaurant	6.289800	6.601135	2
8	Bratwurst Joint	6.249989	6.249989	2
9	German Pop-Up Restaurant	6.044577	6.044577	1
10	Afghan Restaurant	5.902118	7.548614	3
11	Himalayan Restaurant	5.874779	5.874779	2
12	Szechuan Restaurant	5.732105	6.060388	3
13	Polish Restaurant	5.394472	5.394472	1
14	Southern / Soul Food Restaurant	5.259891	5.259891	1
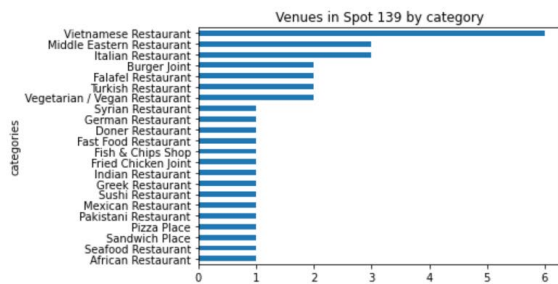
This area has many Vietnamese Resaturants but you have to travel quite far to find **Szechuan** kitchen. Also there again there a some middle easter/turkish options that would be unique here like **Gözleme, Pide or specifically Halal kitchen**

Conclusion

To actually pick a restaurant location and it takes more. But we saved our client from opening yet another Vietnamese Restaurant in an area where there is one on every corner already. We also were able to give them some ideas for concepts that a customer would have to ride all across town to find.

In a further step it would be interesting to factor rent prices into this analysis.