TIPS ON WHERE TO LOCATE A BURGER JOINT IN NEW YORK CITY

Capstone Project- The Battle of Neighborhoods

By Cailin Gertner

Content

Introduction/Business Problem 2
Data Preparation 3
  Sights / Neighborhoods 4
  Competitors in each Sight Area 5
  Characteristics of Competitors 6
  Competitor Weights 7
Methodology 7
  Exploratory Data Analysis of Competition in each Sight Area 7
  Best and worst Location in each Sight Area 8
Results 14
  Analysis of different Sight Areas 14
  Worst location in each Sight Area 19
  Best location in each Sight Area 21
Discussion 22
  Selection of Sight and Business Strategy 22
  Worst and Best location in each Sight Area 24
Conclusion 24
References 24

Introduction/Business Problem

The success or failure of a new restaurant is largely determined by the appropriateness of its location, among other factors such as the price level and quality of food and service.

In this project, we will provide counsel on the advantages and disadvantages of specific locations for the branches of a new burger joint franchise in New York City.

Specifically, this report is targeted at stakeholders that have decided *already* to locate all branches of the new burger joint franchise within a walking distance to popular sightseeing attractions ("sights" hereafter) in the borough of Manhattan and identified tourists as main customer target group for the new business.

This seems to be a compelling strategy for a new burger joint franchise since locations within close reach of sightseeing attractions are naturally busy places and can thus generate many walk-in customers. In fact, restaurants located off the beaten path cannot rely on walk-in customers and thus need to work a lot harder towards attracting the same number of customers. Moreover, we expect many tourists near popular sights and primary targeting tourists as customers is admittedly a convincing strategy for a new burger joint franchise in New York City, due to:

- The sheer number of tourists in New York City; it being a very popular destination for global tourism
- Many tourists wanting an "all American" kind of burger eating experience when visiting New York City
- Tourists arguably being willing to pay a higher price for food than locals on average and hence there being more upside revenue potential for the business

In this project, the approach we take to assist the owners of the new burger joint franchise in their decision on where to locate the different branches is threefold:

1) We will first evaluate the overall attractiveness of the areas around the different sights based on a competitor analysis. Wherever appropriate, we will also recommend different business strategies based on the existing restaurant landscape in each sight area.
2) Secondly, within each sights area, we will offer advise with regards to a specific location where one should *not* open the branch of the new burger joint. This (least attractive) place will correspond to the location inside the sight area that minimizes the sum of weighted distances to all existing competitors (i.e. burger joints) in that area.
3) Thirdly, within each sights area, we will attempt to find the address which should represent the *best location* for the branch of the new burger joint. This location will correspond to the location inside the sight area that maximizes the sum of weighted distances to all existing competitors in that area.

All recommendations/results provided will thus be derived from the analysis of existing burger joints near each sight.

This is because we believe that knowledge of neighbouring restaurants is a major determinant for the selection of the right location and business strategy of a new restaurant. *Competitor analysis* represents a perfect use case as well to leverage modern location and user-generated content data.

Data Preparation

We used location- and user-generated content data provided by *Foursquare Labs Inc.* to analyze existing burger joints in each sight area.

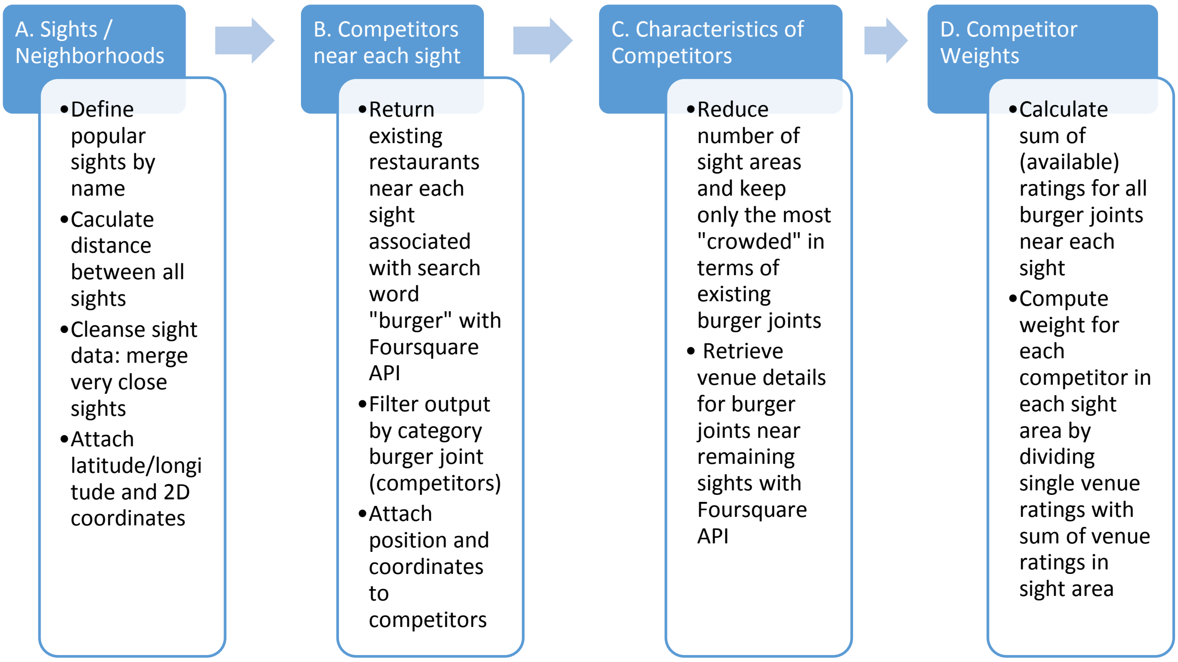
In particular, we compiled the following raw data:

- Names of popular sights in the borough of Manhattan in New York City
- Existing burger joints (i.e. competitors) near the different sights using the Foursquare API
- Characteristics of existing burger joints, i.e. price category, average user rating and distance from sight using Foursquare via premium calls to the API

From this raw data we derived the following data:

- Latitude/Longitude degrees for the different sights
- UTM coordinates in a 2D Cartesian coordinate system for all burger joints and sights
- Weights for burger joints near each sight, such that weights sum up to 1 for each sight area and burger joints with a higher average Foursquare user rating are allocated a higher weight

The key steps of the data preparation task are summarized in the following graphic:



Sights / Neighborhoods

We identified 17 sights popular among tourists visiting New York City:

Empire State Building	The Vessel
Top of the Rock	Flatiron
High Line Park	One World Trade Center
Times Square	One World Observatory
Brooklyn Bridge Park	DUMBO
Staten Island Ferry	The Edge
Central Park	Bryant Park
Wall Street	New York Public Library
Grand Central Terminal	

Source: https://lovingnewyork.de/sehenswuerdigkeiten/unsere-top-10-der-sehenswuerdigkeiten-in-new-york/

For this business case, we assumed that the owners of the new burger joint franchise wish to locate any branch within a comfortable walking distance from one of the above attractions.

We will assume that 400 meters is the maximum distance a potential customer would accept to walk, having completed his or her sightseeing activity, to reach a restaurant and take a meal. This defines 17 circular areas of same size around each sight, hereafter referred to as sight areas, that contain all potential interesting locations for the branches of the new burger joint given the initial stakeholder requirements specifications (StRS).

- Because the Central Park spans a very large area as opposed to the other sightseeing spots, we decided to provide a more specific address to define the center for this sight area. We took the address of a point bordering the south of the Central Park area, since we think that a potential customer (i.e. mostly tourists) is more likely to enter the park from this angle.

- Having obtained the latitude/longitude degrees for all sights using geocoding, we noticed that some of the above sights lie very close together. Since we aim at defining mutually distinct, i.e. is non-overlapping sight areas of radius 400 meters for our analysis, we decided to treat those neighboring sights as one (paired) sight further in the analysis. This was the case for One World Trade Center and One World Observatory as well as Bryant Park and New York Public Library. The coordinates for the paired attractions were calculated by taking the average of the respective latitude/longitude group values.

Finally, we mapped the coordinates of the sights from latitude and longitude degrees to UTM coordinates in a 2D Cartesian coordinate system.

Competitors in each Sight Area

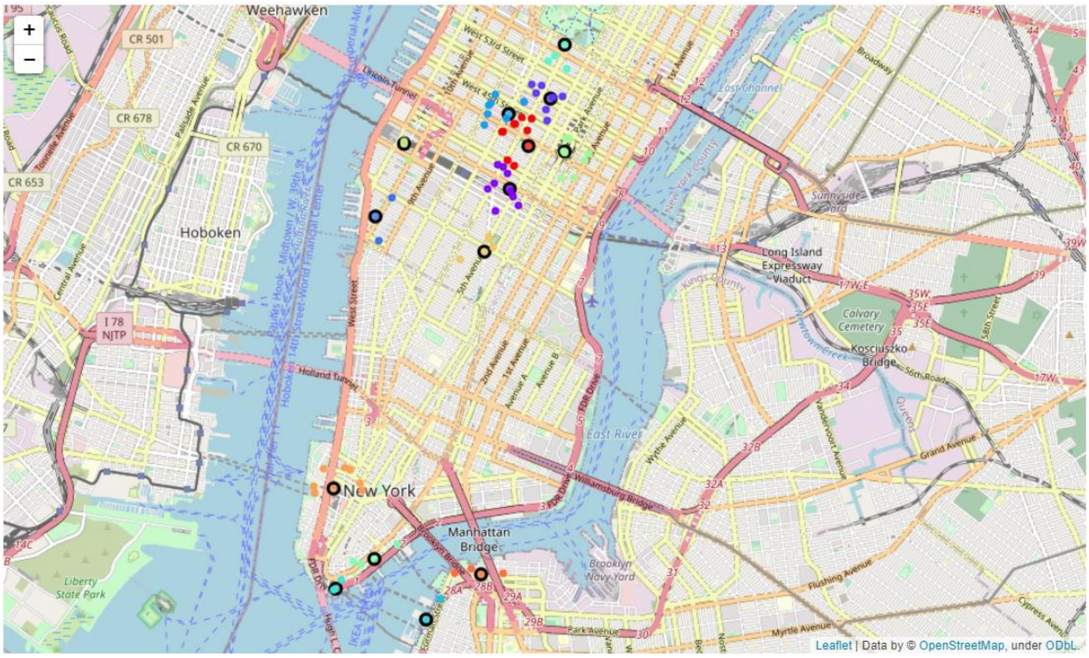
For each sight area we used the explore function of the Foursquare API with the query “burger” such as to retrieve at most 50 restaurants within 400 meters of the sight that are tagged with the word “burger” in the Foursquare database.

In fact, a tourist (i.e. potential customer) is likely to proceed the same way when looking for a restaurant serving burgers having completed his or her sightseeing activity. Thus, gathering the list of competitors of the new business near each sight in this manner is likely to increase the validity of our argument later.

An argument can be made as well that tourists are more likely in general to rely on outside recommendation systems/apps such as Foursquare to choose a place to eat since they usually have limited knowledge (based on personal experience or word-of mouth recommendations) of the local restaurant landscape. Thus, basing our analysis and findings on Foursquare data seemed to be particularly suitable given the fact that tourists were identified as the main target customer group for the new business.

We filtered the list of restaurants returned by the search query “burger” further and kept only restaurants near each sight categorized under the *primary* category name burger joints.

These restaurants represent the nearest established rivals a new burger joint franchise would need to face near each sight. We implicitly assume here that a potential customer wanting to eat a burger in a burger joint picks only a burger joint only from the list of items returned by the Foursquare search. Put differently, visiting another type of restaurant serving burgers (café, American restaurants, etc.) is thought to be *no* substitute for having a burger at a true burger franchise exclusively specialized in burgers.



1 Manhatten Map with sights (circle with dark border) & burger joints near each sight

Finally, we mapped the coordinates of the existing burger joints from latitude and longitude degrees to UTM coordinates in the 2D Cartesian coordinate system.

Characteristics of Competitors

At this point, we decided to keep only the top six sights according to the total number of burger joints nearby. This choice was motivated by two reasons:

- The sight areas with the highest number of existing burger joints potentially also correspond to the most attractive locations for the branches of the new burger joint. For instance, these sights could be more frequented by tourists or lie within a business district as opposed to the others. On the downside, this means we preselected the sights with the highest number of competitors.
- Receiving further details about the restaurants via the Foursquare API is only possible by issuing premium calls. Since we did subscribe to a Sandbox developer account, we were limited to 50 premium calls per day. Reducing the number of sights, thus allowed us to stay below that limit.

To analyze the competition near each of the six remaining sights we again made use of the Foursquare API and retrieved the venue details for each burger joint in each sight area using the respective venue id. Here below are the features we collected:

- Average user ratings of burger joints
- Price category of burger joints
- Distance of burger joints from sight defining respective sight area

These feature set will be used to evaluate the overall attractiveness of the different sight areas and recommend a specific strategy for the new business depending on the sight area.

Competitor Weights

For this project, we suppose that the worst location for a branch of the new burger joint franchise *inside* a sight area is given by the location that minimized the sum of weighted distances to all competitors in that area. Similarly, we suppose that the best location for a branch of the new burger joint franchise *inside* a sight area is given by the location that maximizes the sum of weighted distances to all competitors in that area.

To this end, we transformed the latitude/longitude degree coordinates to UTM coordinates in the 2D Cartesian coordinate system for all competitors and sights.

"Competitor" weights for each sight area were computed by dividing a burger joint's rating by the sum of ratings of all burger joints in that area. By construction, weights therefore sum up to 100% per sight area and competitors with a higher rating were allocated a higher weight in the location problems just described. This corresponds to the intuition that at the best location, the new branch should be located further away from higher rated existing competitors than from lower rated competitor keeping "all other thing equal" and vice versa for the worst location.

Methodology

Exploratory Data Analysis of Competition in each Sight Area

Competitors in each sight area were analyzed according to following aspects to evaluate the overall attractiveness of the different sight areas and recommend an appropriate strategy for a branch of the new burger joint franchise per sight area where applicable:

- Number of existing burger joints in sight area
- Average rating of existing burger joints in sight area
- Price categories of existing burger joints in sight area
- Average distance of existing burger joints in sight area from sight

To choose among sight areas, we considered the following:

- Locations with *a lot of* competition are generally less attractive, i.e. sight areas with a small number of existing burger joints are preferable
- Locations with *strong* competition are generally less attractive, i.e. sight areas with a higher average rating of competitors are to be avoided as well as sight areas with many cheap burger joints. This is based on the fact, that people are mainly sensitive to price and quality when choosing a restaurant.
- Sight areas with fewer competitors *near* the actual sight are more attractive, i.e. sight areas with a higher average distance of existing burger joints from the sight are more attractive. This is because, a potential customer currently at a sight and now looking for a burger joint nearby would prefer closer restaurants over further away restaurants, all things being equal.

Depending on the quality (i.e. ratings) and the price category of competitors in each sight area we can identify the optimal business strategy for a new burger joint per sight area.

An argument can be made, that for sight areas with many existing high-quality restaurants the new competitor should focus on offering lower priced food instead of increasing overall quality of food and service in order to stand a chance against the existing competition. This way the new restaurant could distinguish itself from existing competitors and specifically draw price sensitive customers.

In contrast, we think that for sight areas with many low-priced restaurants, it might be a good strategy for the new business to focus on quality of food and service instead of offering cheap burgers. This way, the new business can attain a higher user-based rating in the long run and beat competition by drawing customers more sensitive to quality than to price when deciding on a restaurant.

Best and worst Location in each Sight Area

Intuition

Within each sight area, i.e. among locations within a comfortable walking distance of 400 meters to the sight, which location should be avoided for the branch of the new burger joint franchise and which location should be optimally chosen?

To answer that question, we can start by taking the perspective of the customers targeted by the business owners, i.e. tourists having completed their sightseeing activity and now wandering off whilst looking for a burger joint nearby.

Among all burger joints returned by the Foursquare mobile app within a comfortable walking distance a tourist would typically prefer:

- Closer restaurants over further away restaurant
- Higher rated restaurants over lower rated restaurants

To maximize the chance that the potential customer will choose the branch of the new burger joint franchise, the branch's location should therefore be as far away as possible from other burger joints (whilst still being within a comfortable walking distance).

What is more, the new branch's location should be further away from higher rated burger joints in the sight area than from lower rated competitors since the potential customer would consider the new burger joint (which has no rating yet) a less attractive choice compared to higher rated equally distant competitors. This assumption is backed by the commonly observed behavior of people rather running the risk of "testing" an unrated restaurant if the alternatives nearby are not immediately compelling.

In contrast, we can think of the worst location for a new restaurant in each sight area as the location closest to all existing competitors in that area since the chances that the new restaurant will be selected are lower due to the number of close alternatives.

Moreover, the location for the new restaurant is worse, all other things being equal, the closer the new restaurant is located to higher rated competitors since the new restaurant would need to withstand stronger competition in that case.

Mathematical Modelling

We can rephrase these two types of location problems in mathematical terms, where they correspond to an optimization (maximization versus minimization) problem with inequality constraints.

In order to find the *optimal* location x_j for the new branch of the burger joint near sight j, we need to *maximize* the weighted sum of distances from the new branch to existing competitors $p_1, p_2, \ldots p_m$ (i.e. burger joints) near sight j. Therefore, the maximization problem looks as follows:

$$\underbrace{max}_{x_j=(x_{0,j},x_{1,j})} \quad F_j(x_j) = \sum_{i=1}^{m} w_i d(p_i, x_j)$$

s.t $d(c_{j,}x_j) \leq r$ i.e. $d(c_{j,}x_j) - r \leq 0$

where

$x_{j=}(x_{0,j}, x_{1,j})$ is the optimal location for the new branch in the 2D Cartesian coordinate system

$p_1, p_2, \ldots p_m$ denote the m restaurants/competitors in sight area j, the new branch of the burger joint x_j would need face

w_i denotes the weight for restaurant p_i calculated as:

$$w_i = \frac{rating_i}{\sum_{i=1}^{m} rating_i}$$

$d(p_i, x_j)$ stands for the Euclidean distance between the new branch x_j near sight j and an existing competitor p_i near sight j, defined as:

$$d(p_i, x_j) = \sqrt{(x_{j,0} - p_{i,0})^2 + (x_{j,1} - p_{i,1})^2}$$

$d(c_{j,}x_j)$ stands for the Euclidean distance between the new branch x_j and the sight j defined by its coordinates $c_{j=}(c_{0,j}, c_{1,j})$ in the two dimensional space

r denotes the radius defining the maximum tolerated distance of the location for the new branch with respect to sight j (400 meters)

In order to find the *worst* location x_j for the new branch of the burger joint near sight j, we need to *minimize* the weighted sum of distances from the new branch to existing competitors $p_1, p_2, \ldots p_m$ (i.e. burger joints) near sight j. The problem looks the same as above except that we minimize the function $F_j(x_j)$ for each sight j instead of maximizing $F_j(x_j)$.

Once we have solved the above problems by using different iterative solution techniques, we can transform the coordinates for the best/worst location of the new restaurant in each sight area j back to latitude and longitude degrees, visualize the best/worst locations on a map and finally generate an address name with reverse geocoding for each best/worst location to show to the stakeholders of the business case.

Worst Location in each Sight Area

In order to solve for the worst location for the new branch in each sight area we used the prominent *Weiszfeld* algorithm introduced in 1937 to solve the Fermat-Weber location problem.

The Fermat-Weber is one of the most famous problems in location theory and requires finding a point in $\mathbb{R}^N$ that minimizes the sum of weighted Euclidean distances to m given points. This is exactly the problem we did come up with, except for the fact that the Fermat-Weber does not take into account any inequality constraints.

Thus, formally speaking, the solution for the worst location provided by the *Weiszfeld* method does not necessarily need to meet our requirement that only locations within a circular area around sight j are considered. [1]

If $F_j(x_j)$ is convex the minimization problem is known to have a solution.

If the points for the competitors near location j are not all on the same line (i.e. they are not colinear) we know that the solution (i.e. worst location) is guaranteed to be unique as well. Looking at the scatter of competitors near each sightseeing attraction on the Folium map we have created, we can assert that this true.

The iterative *Weiszfeld* algorithm performs the following steps in order:

- We start with an initial guess for the worst location $x_{j,0} = (x_{0,j,0}, x_{1,j,0})$, the geometric center of all competitors near sight j

- For steps k=0 until convergence is reached, we update the coordinates for the worst location as follows:

[1] However, it is actually very likely that the solution provided by the *Weiszfeld* method complies with that requirement due to points/competitors being centered around the sight.

We update $x_{0,j,k+1} = \frac{\sum_{i=1}^{m} \frac{w_i}{d(p_i, x_{j,k})} * p_{i,0}}{\sum_{i=1}^{m} \frac{w_i}{d(p_i, x_{j,k})}}$

and $x_{1,j,k+1} = \frac{\sum_{i=1}^{m} \frac{w_i}{d(p_i, x_{j,k})} * p_{i,1}}{\sum_{i=1}^{m} \frac{w_i}{d(p_i, x_{j,k})}}$

If $d(F_j(x_{j,k+1}), F_j(x_{j,k})) \leq \mathcal{E}$ convergence is reached and $x_{j,k+1} = x_j^*$ where epsilon is a very small number (10 ^-7)

The *Weiszfeld* algorithm corresponds to a specification of the gradient descent method commonly applied in machine learning problems with a step size equal to:

$s = \frac{1}{\sum_{i=1}^{m} \frac{w_i}{d(p_i, x_{j,k})}}$

A disadvantage of the *Weiszfeld* algorithm is its rather poor rate of convergence towards the optimum. However, given the small number of points/competitors in each sight area, this is not noticeable.

Best Location in each Sight Area

To solve for the best location for the new branch in each sight area we attempted to adapt Newton's iterative method to our problem. The main challenge here was to incorporate the inequality constraint correctly. In order to account for the inequality constraint ($d(c_j, x_j) \leq r$) we defined a so-called Lagrangian function corresponding to our maximization problem.

The Lagrangian function includes a cost term that penalizes choices for x_j that are not conform to the inequality constraint. Instead of directly maximizing the objective function $F_j(x_j)$ directly, we therefore attempted to maximize the Lagrangian function instead:

$$\underbrace{max}_{x_j=(x_{0,j}, x_{1,j}), \lambda} \quad L_j(x_j, \lambda) = \sum_{i=1}^{m} w_i d(p_i, x_j) - \lambda(d(c_j, x_j) - r)$$

where λ denotes the Lagrangian multiplier or slack variable.

For values of x_j that are further away from the location c_j of sight j the value of the Lagrangian function will decrease (even more so the higher λ is chosen)

If the point $x_j^* = (x_{0,j}^*, x_{1,j}^*)$ and the slack variable λ^* are a solution to our problem, we know that the following Karush-Kuhn-Tucker (KKT) conditions must be satisfied:

$$(1) \frac{dL_j}{dx_{o,j}} (x_j^*, \lambda^*) = 0$$

(2) $\frac{dL_j}{dx_{1,j}}\left(x_j^*, \lambda^*\right) = 0$

(3) $\lambda^*\left(d\left(c_j, x_j^*\right) - r\right) = 0$ (complementary slackness condition)

(4) $\lambda^* \geq 0$

(5) $d\left(c_j, x_j^*\right) - r \leq 0$

These are necessary but not sufficient conditions for a local maximum. Nevertheless, if we have identified a potential solution for the constrained maximization problem, we can verify that these KKT conditions hold.

If the inequality constraint for the optimal location of the new restaurant is not binding, i.e. if the distance from the optimal location for the new branch to sight j is anyway strictly less than r, λ^* needs to be zero in order to satisfy the complementary slackness condition (3). This is straightforward, if you consider the fact that, we did not need to add a cost term to our objective $F_j(x_j)$ to begin with in this particular case.

We apply the following iterative primal dual Inter point Newton-method based algorithm in order to minimize $-L_j\left(x_j, \lambda\right)$ which is equivalent to maximizing $L_j\left(x_j, \lambda\right)$:

(0) We start with an initial guess $\left(x_{j,0}, \lambda_0\right)$ for the location of the new branch and slack variable of the cost term at the optimum.

(1) For steps k=0 until the overall stopping criterion is satisfied we update the coordinates of the location for the new branch $x_{j,\mathrm{k}} = (x_{0,j,k}, x_{1,j,k})$ and the slack variable λ_{k} as follows:

$$X_{j,k+1} = X_{j,k} - G\left(X_{j,k}\right) * H_{\left(X_{j,k}\right)}{}^{-1}$$

Where

${X_{j,k}}^T = (x_{0,j,k}, x_{1,j,k}, \lambda_{\mathrm{k}})$ denotes the parameter vector at step k

$G\left(X_{j,k}\right)$ is the gradient vector evaluated at step k with the intermediate parameter values ${X_{j,k}}^T$. It is given by the first-order derivatives of our function $-L_j$ with respect to all parameters evaluated at step k:

$$G\left(X_{j,k}\right)^T = \left(-\frac{dL_j}{dx_{0,j}}\left(X_{j,k}\right), -\frac{dL_j}{dx_{1,j}}\left(X_{j,k}\right), -\frac{dL_j}{d\lambda}\left(X_{j,k}\right)\right)$$

$H_{(X_{j,k})}{}^{-1}$ denotes the inverted symmetric Hessian matrix of second-order partial derivatives of $-L_j(x_j, \lambda)$ with respect to the parameters evaluated at step k.

We repeat this until the number of iterations exceeds the maximum number of iterations or the following stopping criterion is satisfied:
$G(X_{j,k})^T * G(X_{j,k}) < \alpha$

This corresponds to the classic Newton method.

(2) check if $0.5 < \frac{\lambda_{k+1}}{\lambda_k} \leq \frac{1}{0.5}$

If the left side is valuated, reset μ_{k+1}=0.5
If the left side is valuated, reset μ_{k+1}=2

If number of iteration is greater than the maximum number of iterations or $\frac{L_j(x_{j,k+1},\lambda_{k+1}) - L_j(x_{j,k},\lambda_k)}{1+abs(L_j(x_{j,k},\lambda_k))} < \varepsilon$ exit the loop, otherwise go back to (1)

The gradient G at step k can be evaluated by computing the following first-order derivatives:

$$-\frac{dL_j}{dx_{o,j}}(x_{j,k}, \lambda_k) = -\sum_{i=1}^{m} \frac{w_i * (x_{0,j,k} - p_{i,0})}{d(p_i, x_{j,k})} + \lambda * \frac{(x_{0,j,k} - c_{0,j})}{d(c_j, x_{j,k}))}$$

$$-\frac{dL_j}{dx_{1,j}}(x_{j,k}, \lambda_k) = -\sum_{i=1}^{m} \frac{w_i * (x_{1,j,k} - p_{i,1})}{d(p_i, x_{j,k})} + \lambda * \frac{(x_{1,j,k} - c_{1,j})}{d(c_j, x_{j,k}))}$$

$$-\frac{dL_j}{d\lambda}(x_{j,k}, \lambda_k) = -(r - d(c_j, x_{j,k}))$$

The Hessian matrix H at step k can be evaluated by computing the following second-order derivatives:

$$-\frac{dL_j}{dx_{o,j}dx_{o,j}}(x_{j,k}, \lambda_k) = -\sum_{i=1}^{m} \frac{w_i * (1 - \left(\frac{x_{0,j,k} - p_{i,0}}{d(p_i, x_{j,k})}\right)^2)}{d(p_i, x_{j,k})} + \lambda_k * \frac{1 - \left(\frac{x_{0,j,k} - c_{j,0}}{d(c_j, x_{j,k})}\right)^2}{d(c_j, x_{j,k})}$$

$$-\frac{dL_j}{dx_{1,j}dx_{1,j}}(x_{j,k}, \lambda_k) = -\sum_{i=1}^{m} \frac{w_i * (1 - \left(\frac{x_{1,j,k} - p_{i,1}}{d(p_i, x_{j,k})}\right)^2)}{d(p_i, x_{j,k})} + \lambda_k * \frac{1 - \left(\frac{x_{1,j,k} - c_{j,1}}{d(c_j, x_{j,k})}\right)^2}{d(c_j, x_{j,k})}$$

$$-\frac{dL_j}{dx_{o,j}dx_{1,j}}(x_{j,k}, \lambda_k) = -\sum_{i=1}^{m} \frac{-w_i * \frac{(x_{0,j,k} - p_{i,0})(x_{1,j,k} - p_{i,1})}{d(p_i, x_{j,k})^2}}{d(p_i, x_{j,k})} - \lambda_k \frac{\frac{(x_{0,j,k} - c_{j,0}) * (x_{1,j,k} - c_{j,1})}{d(c_j, x_{j,k})^2}}{d(c_j, x_{j,k})}$$

$$-\frac{dL_j}{dx_{o,j}\, d\lambda}(x_{j,k}, \lambda_k) = \frac{x_{0,j,k} - c_{0,j}}{d(c_j, x_{j,k})}$$

$$-\frac{dL_j}{dx_{1,j}\,d\lambda}\left(x_{j,k},\lambda_k\right) = \frac{x_{1,j,k} - c_{1,j}}{d(c_j, x_{j,k})}$$

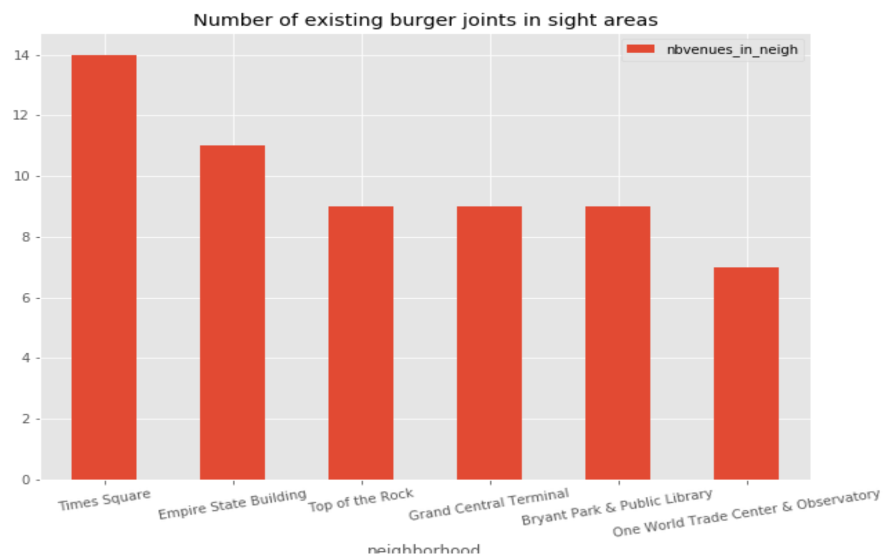
$$-\frac{dL_j}{d\lambda d\lambda}\left(x_{j,k},\lambda_k\right) = 0$$

Results

Analysis of different Sight Areas

Number of Competitors

The number of existing burger joints is highest near Times Square (14) and Empire State Building (11) and lowest near One World Trade Center & Observatory (7). For the other 3 sight areas we found exactly 9 competitors.



2 Number of Competitors near each Attraction

Competitor Ratings

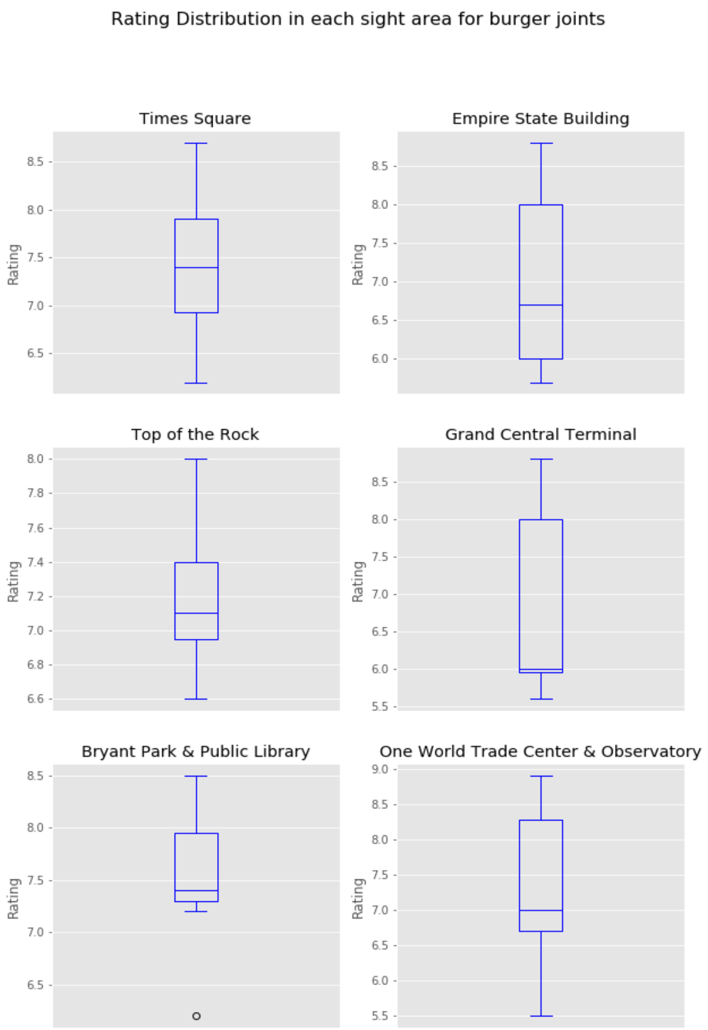
We computed the average of ratings of existing burger joints for each sight area:

	neighborhood	avg_venue_rating
5	Bryant Park & Public Library	7.528571
2	Times Square	7.375000
4	One World Trade Center & Observatory	7.300000
1	Top of the Rock	7.200000
0	Empire State Building	7.042857
3	Grand Central Terminal	6.942857

3 Average Ratings of Competitors near each Attraction

We found that average ratings per sight area are very similar (approx. 7 to 7.5), although existing burger joints were rated highest near Bryant Park & Public Library and Times Square and lowest near Empire State Building and Grand Central Terminal.

We analyzed the distribution of competitor ratings near each sight further by plotting the box plot of venue ratings for each sight:



4 Rating Distribution for Competitors near each sight

These box plots allow us to assess the distribution of the competitor ratings in each sight area according to the following five main dimensions (Minimum, First quartile, Median, Third quartile, Maximum)

The distribution of restaurant ratings seems to be nicely symmetric for burger joints near Times square but of course this is also the sight area with highest number restaurants inside.

The remaining distributions seem to be right skewed. Therefore, median values lie below the computed average ratings meaning that 50% of the burger joints in the sight area are rated below average.

The Bryant Park & Public Library area has the highest average rating of burger joints, however, there exists one negative outlier among burger joints:

```
venue.name                           HB Burger
venue.id                  49dceeecf964a520c05f1fe3
neighborhood        Bryant Park & Public Library
venue.rating                                   6.2
```

Distances of Competitors from Sightseeing Attractions

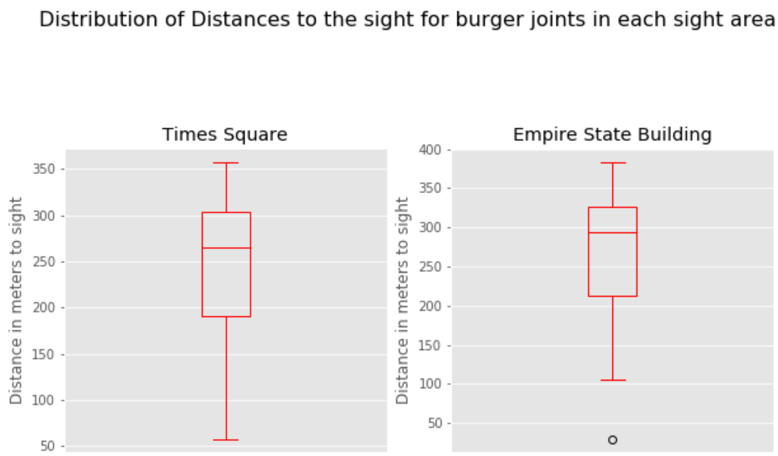
We computed the average of distances of existing burger joints in each sight area from the respective sight:

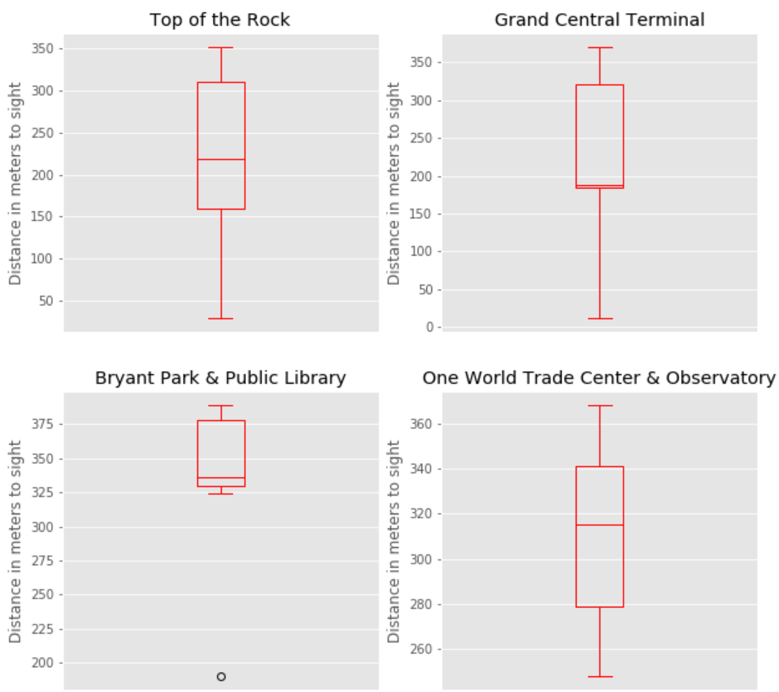
	neigh.id	neighborhood	avg_venue_dist
5	14	Bryant Park & Public Library	335.888889
4	11	One World Trade Center & Observatory	310.142857
0	0	Empire State Building	254.181818
2	3	Times Square	245.785714
1	1	Top of the Rock	226.666667
3	8	Grand Central Terminal	213.111111

5 Average Distance to Sight of Competitors near each Attraction

We found that burger joints are located nearest to the sight in the area around Grand Central Terminal, whereas the average distance to the attraction is larger for burger joints near Bryant Park & Public Library and On World Trade Center & Observatory.

We analyzed the distribution of competitor distances in each sight area from the sight further by plotting the box plot of competitor distances to the sight per sigh area:





6 Distribution of Distances to Sight for Competitors near each sight

Relative to other restaurants near the respective attraction we can see that near Empire State Building and Bryant Park there exist two restaurants that are located particularly close to the sight:

```
venue.name                       Smashburger
venue.id                534330a7498e89554e17da87
neighborhood               Empire State Building
venue.distance                                30
venue.rating                                 7.5
```

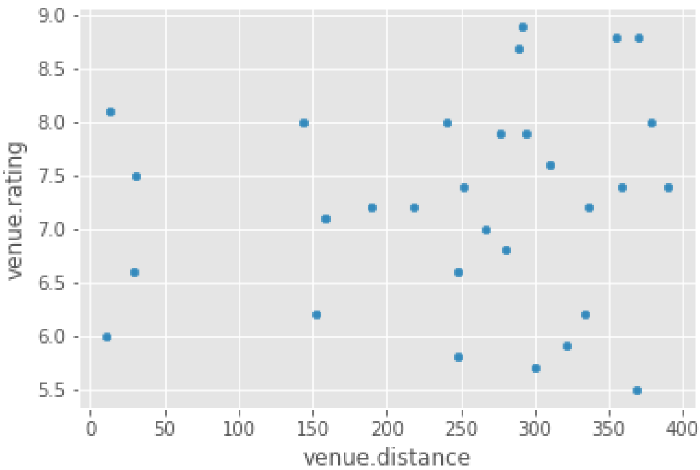
```
venue.name                           STK Midtown
venue.id                    4eb2ec5077c846fcfeb2d366
neighborhood            Bryant Park & Public Library
venue.distance                                   190
venue.rating                                     7.2
```

We know that the Grand Central Terminal area has the lowest average rating of burger joints whereas the area around near Bryant Park & Public Library has the highest average rating. Interestingly, now we notice that the average distance of restaurants to the attraction is smallest near Grand Central Terminal and highest near Bryant Park.

Can we say that the quality of burger joints increases the further away they are located from the attraction?

Intuitively, this could be the case since, as mentioned in the introduction, restaurants located off the beaten path need to work harder towards luring customers for instance by offering better quality.

We plotted a scatter plot of all burger joint's ratings and their respective distances to the sight and modelled the relationship between ratings and distances by a linear regression:




```
                            OLS Regression Results
==============================================================================
Dep. Variable:           venue.rating   R-squared:                       0.017
Model:                            OLS   Adj. R-squared:                 -0.018
Method:                 Least Squares   F-statistic:                    0.4739
Date:                Mon, 29 Jun 2020   Prob (F-statistic):              0.497
Time:                        21:33:00   Log-Likelihood:                -40.976
No. Observations:                  30   AIC:                             85.95
Df Residuals:                      28   BIC:                             88.75
Df Model:                           1
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------
const              6.9684      0.442     15.758      0.000       6.063       7.874
venue.distance     0.0011      0.002      0.688      0.497      -0.002       0.004
==============================================================================
Omnibus:                        1.218   Durbin-Watson:                   1.732
Prob(Omnibus):                  0.544   Jarque-Bera (JB):                0.968
Skew:                          -0.172   Prob(JB):                        0.616
Kurtosis:                       2.190   Cond. No.                         670.
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

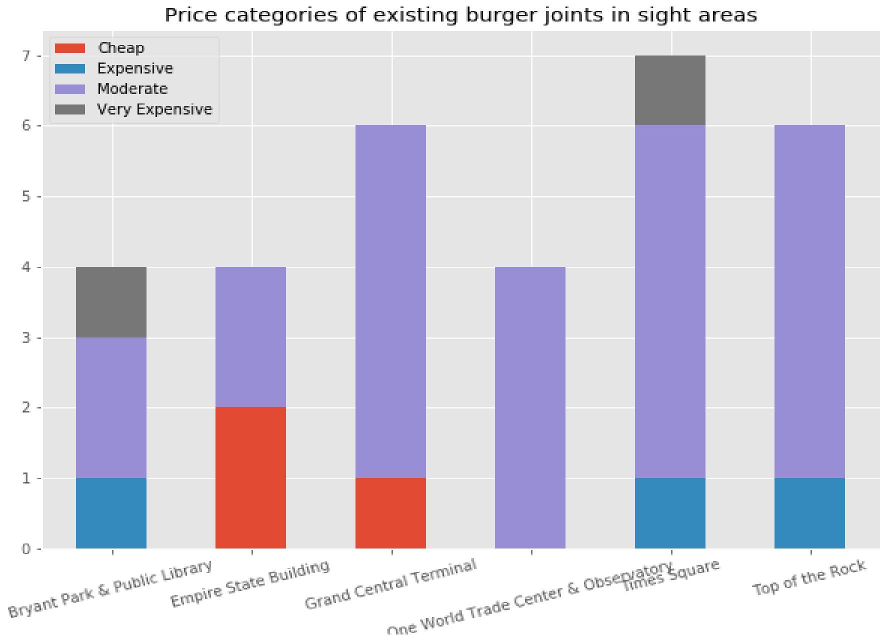
7. Relationship between Distance to Sight and Ratings of Competitors

We cannot reject the null hypothesis that the slope coefficient of the independent variable distance to sight is null in the regression according to the t-statistic and p-value.

Therefore, we cannot confirm that there is a significant positive relationship between distance to attractions and ratings of restaurants from this small sample.

Price Range of Competitors

We have plotted the number of venues in each price category:



8. Distribution of Price Categories of Competitors near each Attraction

From this graph, we concluded the following:

- Cheap burger joints are mainly located near Empire State Building and Grand Central Terminal. Near Empire State Building cheap restaurants account for an even larger proportion of all burger joints.
- The only attractions with very expensive burger joints in the surrounding area are Bryant Park & Public Library and Times Square. Expensive and very expensive burger joints account for a larger proportion of all burger joints in the area around Bryant Park & Public Library than Time Square.

Given that the average rating of burger joints was the highest near Bryant Park and Times Square, that may hint towards a positive correlation between quality (i.e. ratings) and price.

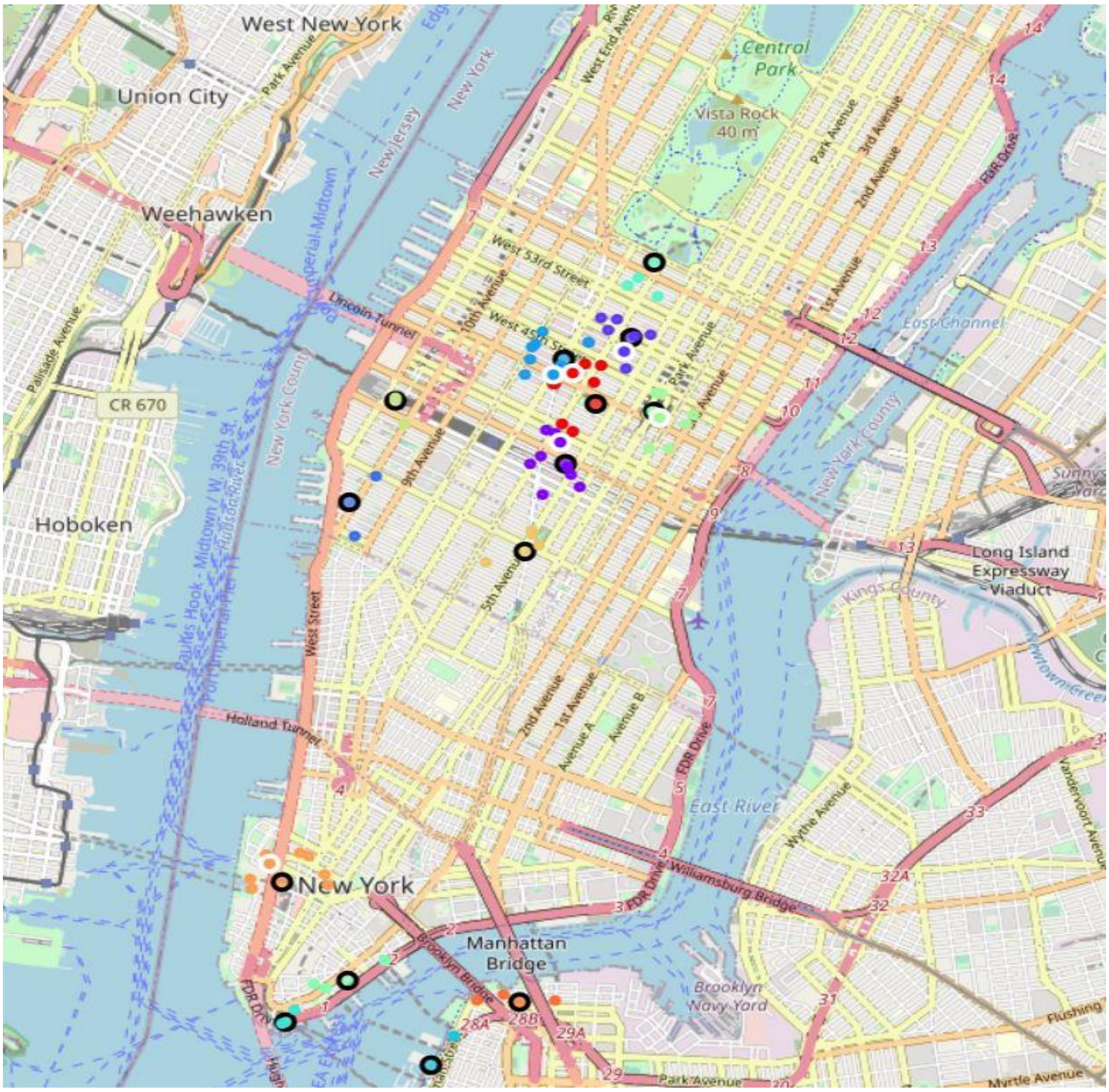
Worst location in each Sight Area

As described in the methodology section we were able to return the following addresses for the worst location within each sight area for a new burger joint:

neigh.id	neighborhood	Weiszfeld.address
3	Times Square	42nd Street–Times Square–Port Authority Bus Te...
0	Empire State Building	51, West 35th Street, Midtown South, Manhattan...
1	Top of the Rock	38, West 48th Street, Midtown, Manhattan, Manh...
8	Grand Central Terminal	Grand Hyatt New York, 109, Park Avenue, Murray...
14	Bryant Park & Public Library	Heartland Brewery, 127, West 43rd Street, Time...
11	One World Trade Center & Observatory	Goldman Sachs Tower, 200, West Street, Battery...

9. Addresses of Worst Location near each Attraction for a Branch of the new Burger Joint

These addresses correspond to the location that would have the smallest overall weighted distance to all existing competitors near the attraction. We can visualize the results also on the map of Manhattan:



10. Worst location near each Attraction (circles with white border)

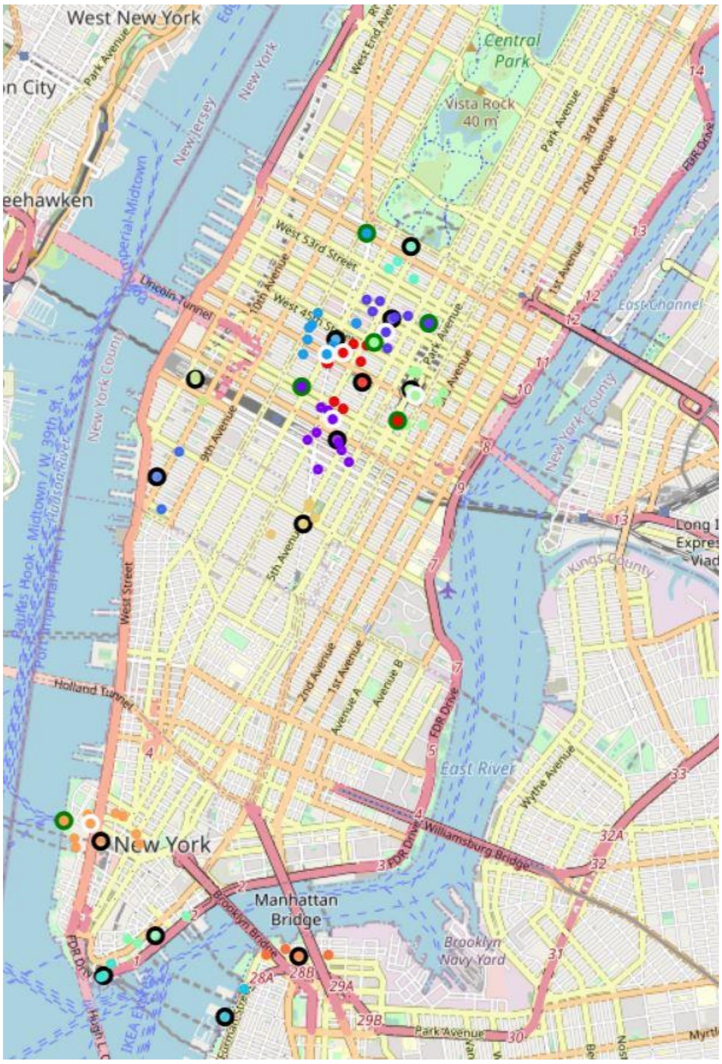
Best location in each Sight Area

As described in the methodology section we were able to return the following addresses for the best location within each sight area for a new burger joint:

neigh.id	neighborhood	Newton.address
3	Times Square	3, Columbus Circle, Manhattan, Manhattan Commu...
0	Empire State Building	200, West 37th Street, Garment District, Manha...
1	Top of the Rock	Look Building, 488, Madison Avenue, Midtown Ea...
8	Grand Central Terminal	59, West 46th Street, Midtown, Manhattan, Manh...
14	Bryant Park & Public Library	100, East 39th Street, Murray Hill, Manhattan,...
11	One World Trade Center & Observatory	Brookfield Place / Battery Park City Ferry Ter...

11. Addresses of Best Location near each Attraction for a Branch of the new Burger Joint

These addresses correspond to the location that would have the largest overall weighted distance to all existing competitors near the attraction.



12. Best location near each Attraction (circles with green border)

Discussion

Selection of Sight and Business Strategy

Based on the results of our analysis of competition near each of the six sight we compiled the following table with an overview of the analyzed features:

	# of Competitors	*Avg. Rating*	*Avg. Venue Distance*	*Price Category*	*Total (yellow omitted)*
Empire State Building	+	-	0	- -	***-2***
Top of the Rock	0	0	-	+	***0***
Times Square	++	+	0	++	***-1***
Grand Central Terminal	0	- -	- -	-	***-1***
One World Trade Center & Observatory	-	0	+	0	***2***
Bryant Park & Public Library	0	++	++	+++	***3***

13 Ranking of Attraction / Neighborhood by analyzed features

- The plus and minus signs indicate the magnitude and direction of the features. However, since for instance a higher average rating of existing competitors would make a sight area *less* attractive, depending on the analyzed feature, plus and minus signs are colored in green or red to indicated a positive (green) effect on the attractiveness of the sight area or negative (red) effect on the attractiveness of the sight area.
- The value 0 stands for the fact that relative to all other sight areas the feature does neither seem to add to nor diminish the attractiveness of the sight area. In the "Total" column we aggregate the number of negative effects (red) and positive effects (green) to rank the sight areas according to their attractiveness.

Now we can provide the following priority list to the stakeholders as to which sight areas to target when setting up the branches of the new burger joint franchise in New York City/Manhattan:

1. **Bryant Park & Public Library**
2. **One World Trade Center & Observatory**
3. **Top of the Rock**
4. **Times Square and Grand Central Terminal**
5. **Empire State Building**

Sight areas are attractive for different reasons, implying that different business strategies are more promising than others for opening a branch of the new burger joint near the respective sight:

- Near **Bryant Park & Public Library** one can find *mainly higher priced and highly rated* burger joints. The average distance to both sights is also the largest. This combination of features has led to the area being at the top of our recommendation list.
  A suitable business strategy would be offering *lower priced* burgers *closer to the sight* targeting price sensitive costumers even if that should entail a lower quality standard (and thus future rating) since it would be difficult to compete in terms of quality with existing (highly rated) competitors.
- For **One World Trade Center & Observatory** there are *no* specific strategies lending themselves. The main advantage of the location is the *smaller number of existing burger joints* in the area and *slightly higher distance* of existing competitors to the sight
- **Grand Central Terminal** and **Times Square** are equally recommended (3rd priority) to stakeholders albeit for different reasons.
  - **Grand Central Terminal** has an *average nu*mber of burger joints which are the *lowest rated* compare to all other sight areas. However, these are cheap and located well (i.e. close to the Terminal) as well. Thus, a good strategy might be to set up a "high quality" burger joint where customers might enjoy lingering when waiting for their onward journey. This branch would be able to score a higher rating and thus be chosen by travellers sensitive to quality when inspecting the list of items returned by Foursquare.
  - **Times Square** has the highest number of competitors however these are on average more expensive (second most expensive compared to all sights). Thus, we recommend setting up a lower-priced burger joint here. Avg. distance to Times Square is average, therefore the new business could try to find a better location (closer to the main attraction) than existing competitors as well.
- **Empire State Building** was ranked last since existing burger joints are the cheapest here among all and the number of competitors is high too. However, avg. rating is less than elsewhere and distance to the sight average. Therefore, we recommend a quality-based strategy as for Grand Central Terminal.

Of course, we have left aside important features such as expected restaurant lease when forming these recommendations. It is likely that leases increase the closer the location is to the main attraction, thus counteracting a strategy based on price for example. This would be another interesting aspect to incorporate into the analysis.

Worst and Best location in each Sight Area

The implementation of the *Weiszfeld* method converged nicely.

Based on the literature, an implementation of primal dual interior-point method or, alternatively, barrier method should be used to handle an optimization problem with inequality constraints.

We found the best location in each sight area by implementing a primal dual interior-point Newton method. Instead of working with a maximum tolerated radius of 400 meters from the best location to the sight, we took a maximum radius of 502 meters for the inequality constraint. This is because, that was actually the highest Euclidean distance of over all existing burger joints from their allocated attraction in the 2D Cartesian coordinate system we mapped the sight and restaurant latitude/longitude coordinates to.

We tried to verify that the Karush-Kuhn-Tucker (KKT) conditions were satisfied at the point returned by the algorithm. However, these are necessary but not sufficient conditions for a local optimum and it may well be that we found a local maximum as opposed to a global maximum as well.

Thus, our results with respect to the best location of branches within each sight area should be treated with caution until further verification.

Conclusion

In this paper we provided valuable insights related to the attractiveness and appropriate business strategy for setting up the branches of a new burger joint franchise given the business requirement that branches should be located near to popular attractions.

We have also advised against and recommended a specific location near each sight based on the fact that it would be the closest or, respectively, furthest away from all competitors.

References

https://en.wikipedia.org/wiki/Interior-point_method

https://www.mathematik-informatik.uni-wuerzburg.de/fileadmin/10040700/paper/FermatWeberNewton.pdf

https://slideplayer.com/slide/7601919/