Data Preparation

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

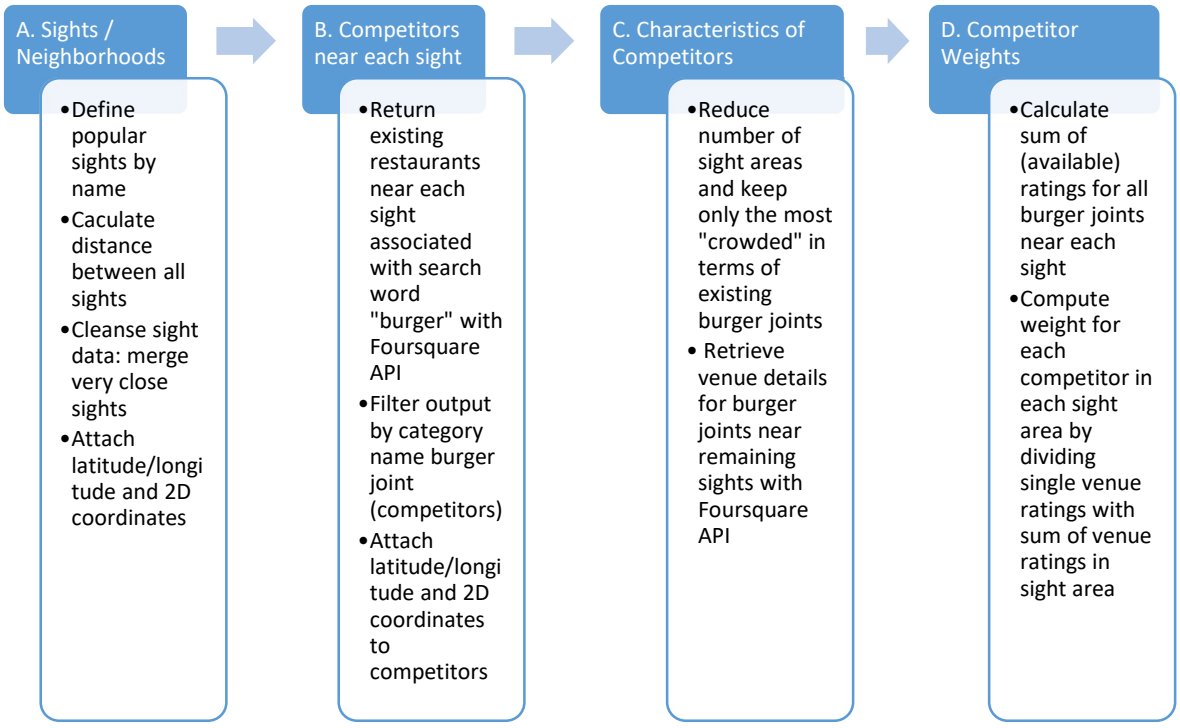
In particular, we compiled the following raw data:

- List of 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 as 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	

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

For this business case, we did assume 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

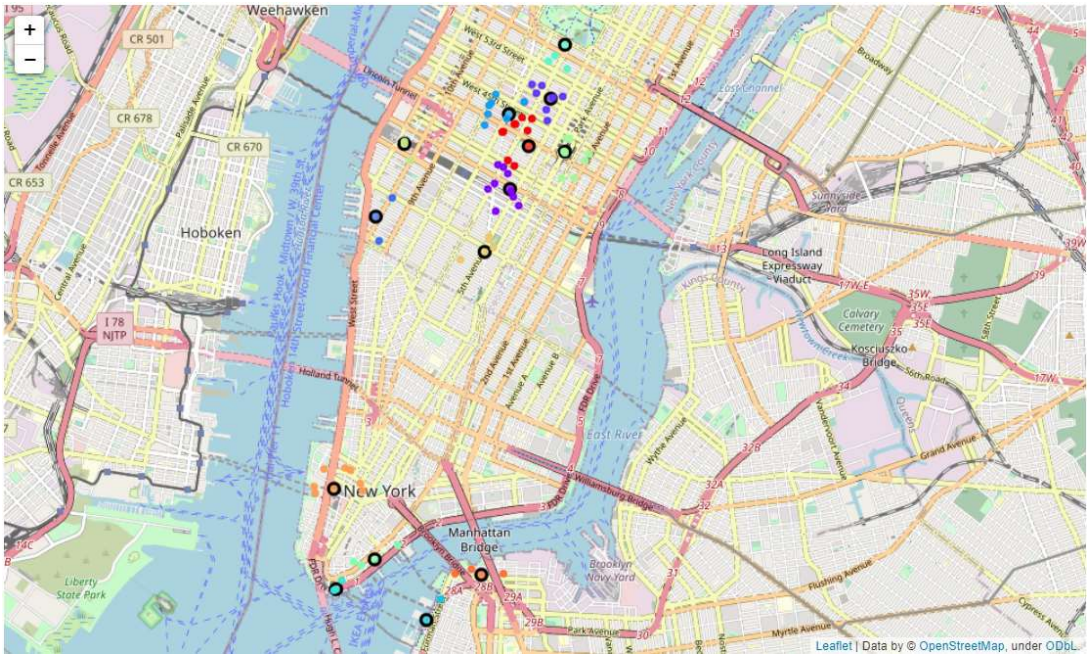
For each sight area we used the explore function of the Foursquare API with the query "burger" such as to retrieve at most 50 restaurants within 400 meters of the sight that are tagged with the word "burger" in the Foursquare database.

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 (markers with dark edge) & 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 farer away from higher rated existing

competitors than from lower rated competitor keeping “all other thing equal” and vice versa for the worst location.