

# **Final Four Square Madness**

Thomas Matich

Coursera Applied Data Science Capstone

## **I. Introduction**

The NCAA Men's Basketball Final Four is one of the largest, highest-grossing annual sporting events in the United States. According to a Forbes article in 2016, CBS paid over \$10.8 billion to secure the television rights from 2011 – 2024, an increase of 4,535% since 1986. The popularity of the event is interesting to reconcile with the fact that there is an inherent bias in the selection of Final Four host cities favoring the east coast. Since the tournament was expanded to its current format of 64 teams in 1985, there has been only 4 occurrences of a city in the Mountain or Pacific time zones hosting a Final Four. Ever more bewildering is that a Final Four has not been hosted in California since 1975.

The aim of this project was to propose a major city in California as a host venue for the Final Four. Two cities, Los Angeles and San Francisco, were identified as they both have large arena facilities. The Staples Center is located in Downtown Los Angeles and the Chase Center in San Francisco's Mission Bay. The intended audience and stakeholders would be the NCAA tournament selection committee, LA/SF business owners and policymakers/city officials, college basketball fans and alumni.

Having a culturally diverse and vibrant city like Los Angeles or San Francisco, both towns that have rich basketball cultures (Los Angeles Lakers, UCLA, Golden State Warriors) would present a refreshing, exciting atmosphere for the Final Four. It could also prompt the NCAA to consider other venues in the future to further expand the Final Four into an event that is more evenly distributed across the country over the years.

## **II. Data**

The data for this project consisted of a custom Python dataframe built with the geocoordinates of Staples Center and Chase Center. The geocoordinates of the arenas are publicly available from many sources. These two dataframes were then used to pass API calls to Foursquare to gather venues in each neighborhood within an approximate 2-mile radius.

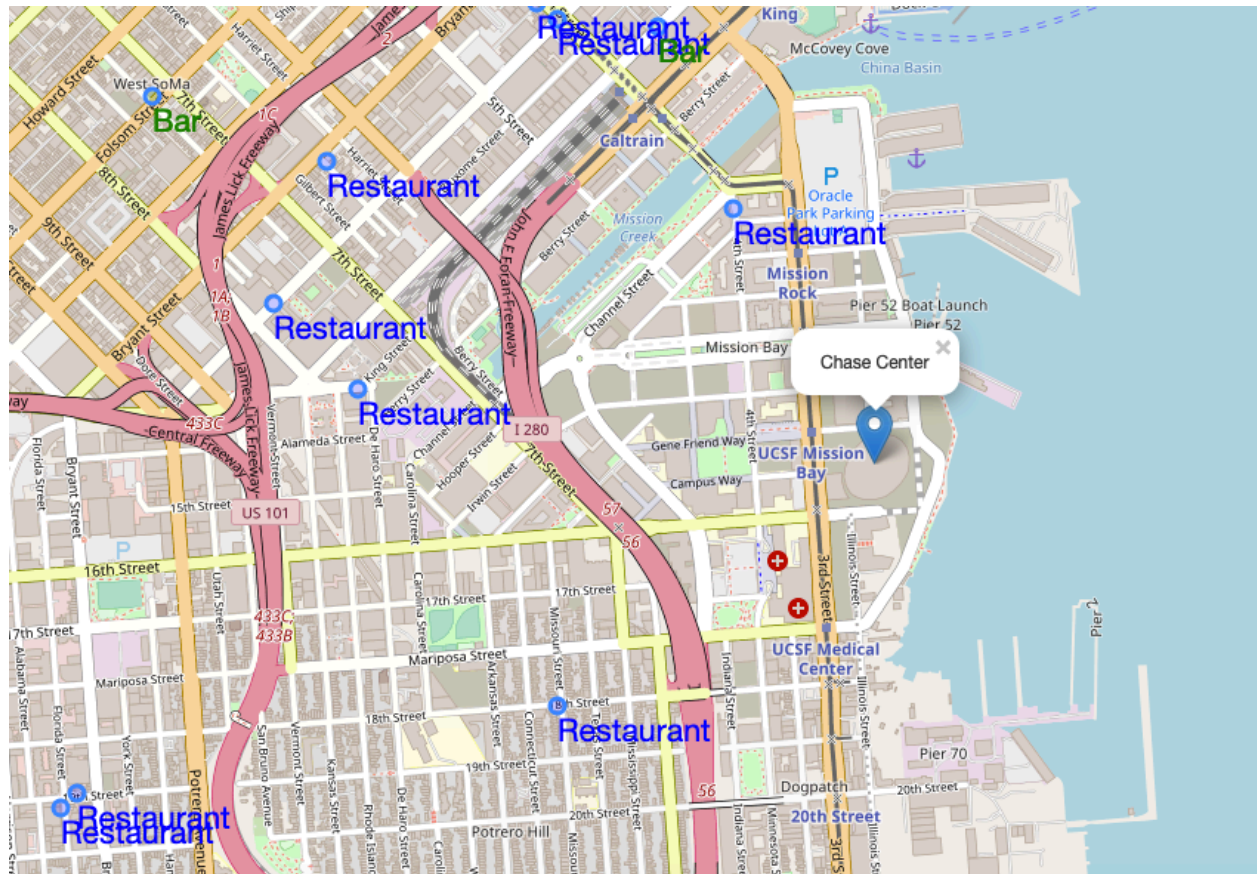
## **III. Methodology**

The end-goal of the data analysis was to be able to plot restaurants, bars and hotels in Downtown Los Angeles and Mission Bay in Python. The thought process was that the arena with more of the aforementioned surrounding venues would be a better host for the Final Four. Once venues were captured from Four Square, they were loaded into a dataframe for the two neighborhoods. This process involved loops, one hot encoding and function methods. The data was then sliced, cleaned and combined to where there was one dataframe that contained all of the restaurants,

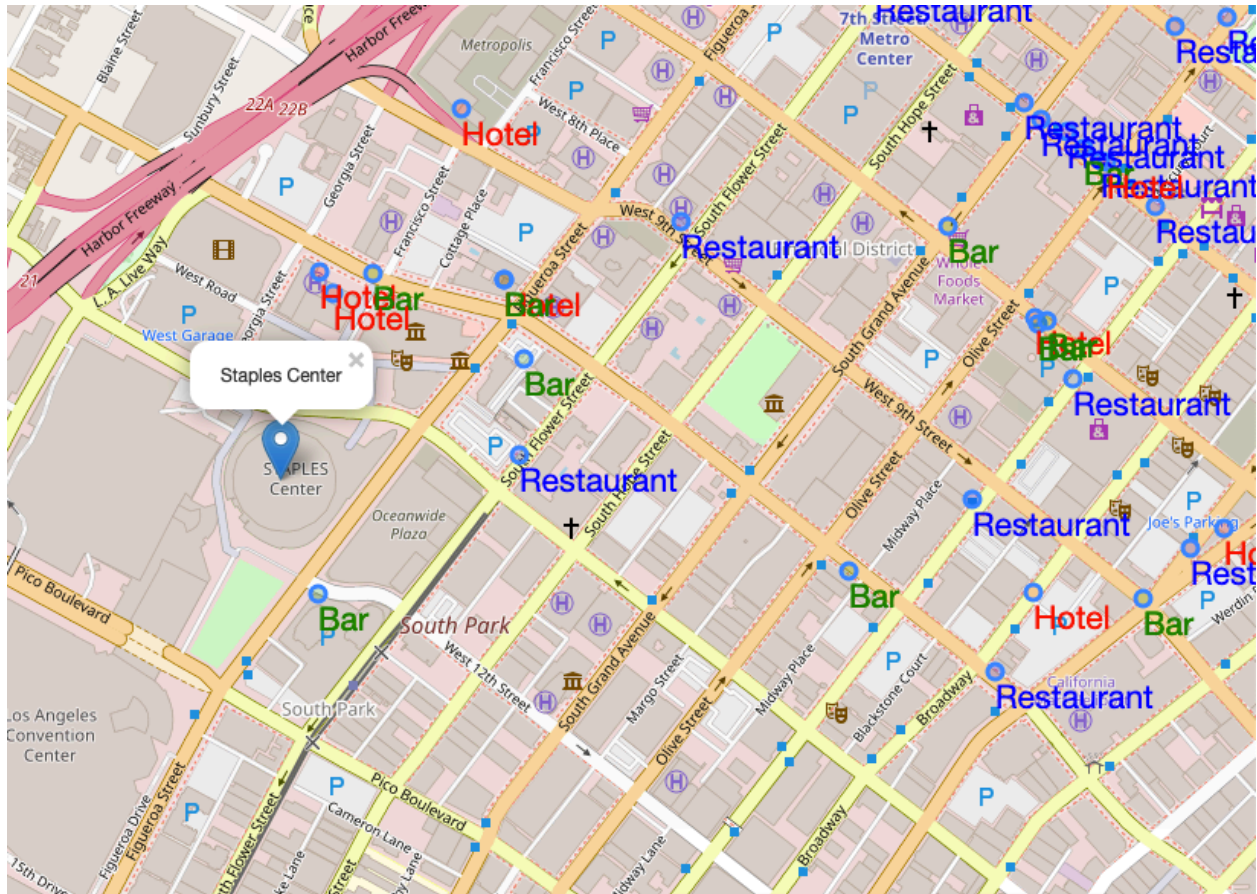
hotels and bars for both Downtown Los Angeles and Mission Bay. This dataframe was then split into the two separate neighborhoods in order to map the venues surrounds each arena.

#### IV: Results

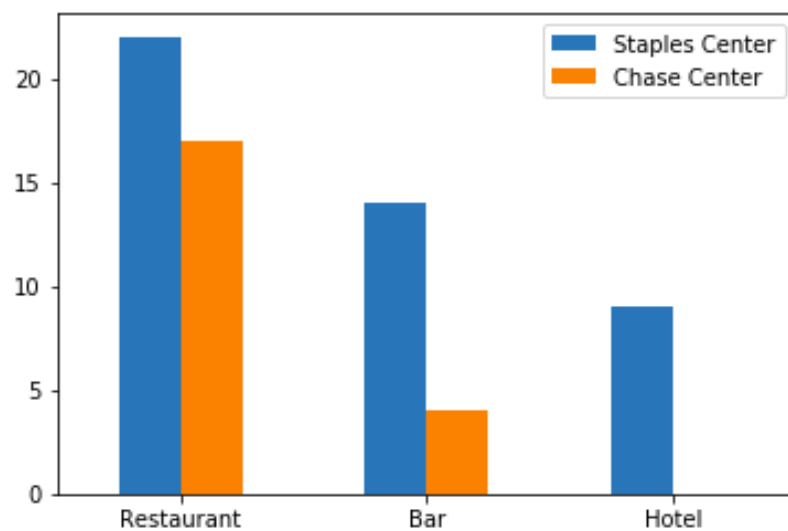
It became apparently fairly early in the data analysis stage that the neighborhood of Mission Bay was sparse on hotels. But further investigation was need and the mapping would provide helpful to visualize the dispersion of venues related to each area.



The figure above represents the venues near Chase Center. There are no hotels in the immediate vicinity and the majority of restaurants and bars are located across water or highway. The figure on the next page is the map of venues near Staples Center, where there are hotels, bars and restaurants located much closer to the area.



While these maps are helpful visualization, simple descriptive statistics of the total number of these venues could further validate the findings. Therefore, a simple bar chart was produced to highlight the disparity of venues between Staples Center and Chase Center in the next figure.



With the Staples Center being surrounded by 9 hotels as opposed to Chase Center having none nearby was a pretty glaring statistic. Staples Center's Downtown Los Angeles neighborhood also had more restaurants and bars than Chase Center's Mission Bay. To confirm Foursquare's finding of "zero" hotels in Mission Bay, I consulted Google Maps, and found that the closet hotel was 1 mile away. There are several hotels located with 0.25-0.5 miles of Staples Center, however.

## **V: Discussion**

Based on the findings of venues near Staples Center and Chase Center, it is easy to recommend Staples Center/Downtown Los Angeles as a host for the Final Four as they are better positioned with closer hotel accommodations and have more restaurants and bars. With its positioning on Mission Bay, Chase Center seems to be more isolated from other parts of San Francisco, therefore the area would need to be further developed.

## **VI: Conclusion**

Foursquare is a great resource of location data and venues, but it is not all encompassing. It is also easy to be fooled by maps and user generated reviews and databases, therefore without actually visiting the surrounding neighborhoods of Staples Center and Mission Center, it would be difficult to recommend one venue based on data alone. For selecting a Final Four city, one should also compare these potential sites against past host cities to look for common attributes.