

*Final Four Square* Madness

Coursera Applied Data Science Capstone

Thomas Matich



A recent research paper I wrote examined, among various other trends and data points, an inherent East Coast bias in the Final Four host cities for the Men's NCAA Basketball Tournament held in March of every year.

Since the modern era of the tournament in 1985, when the field was expanded to 64 teams, there have only been 4 occurrences where the Final Four was hosted in a city that was in the Pacific or Mountain Time Zone.

In fact, it was quite surprising to learn that the last time a Final Four was hosted in the state of California was in San Diego in 1975.



The aim of my project is to compare and contrast Los Angeles (Staples Center) and San Francisco (Chase Center) as host cities/arena venues for a Final Four with the Four Square data that would include hotels, restaurants and bars.

The audience/stakeholders are the NCAA, tournament selection committee, LA/SF business owners/policymakers and college basketball fans/alumni.



## Getting started:

Because I was comparing two neighborhoods (Downtown LA, Mission Bay), I simply created two dataframes with the geocoordinates in Python.

```
data = [['Staples Center', 'Downtown Los Angeles', 34.043056, -118.267222]]  
sc = pd.DataFrame(data, columns = ['Venue', 'Neighborhood', 'Latitude', 'Longitude'])  
sc
```

	Venue	Neighborhood	Latitude	Longitude
0	Staples Center	Downtown Los Angeles	34.043056	-118.267222

```
data1 = [['Chase Center', 'Mission Bay', 37.768056, -122.3875]]  
cs = pd.DataFrame(data1, columns = ['Venue', 'Neighborhood', 'Latitude', 'Longitude'])  
cs
```

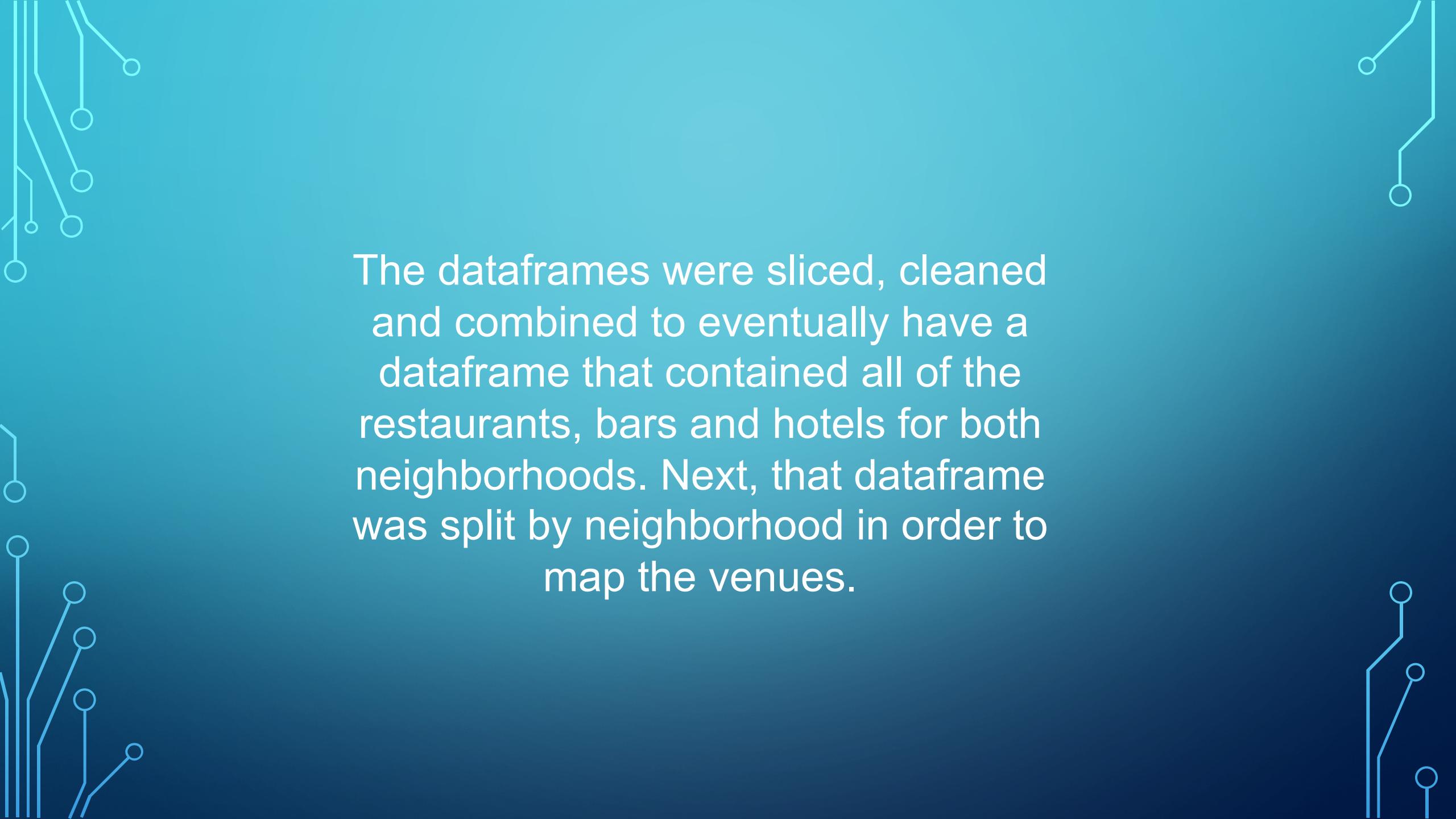
	Venue	Neighborhood	Latitude	Longitude
0	Chase Center	Mission Bay	37.768056	-122.3875



Both dataframes were passed through separately with Four Square API calls to get neighborhood venues within an approximate 2 mile radius.

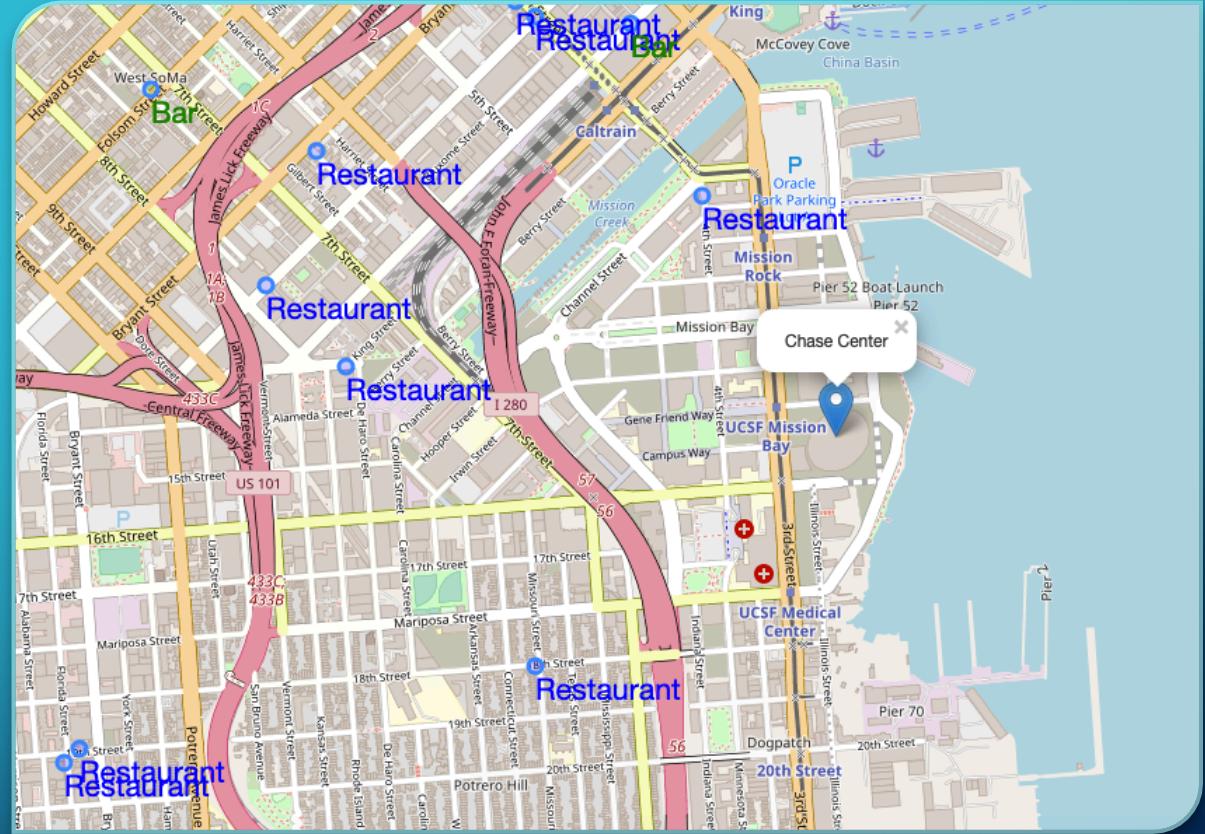
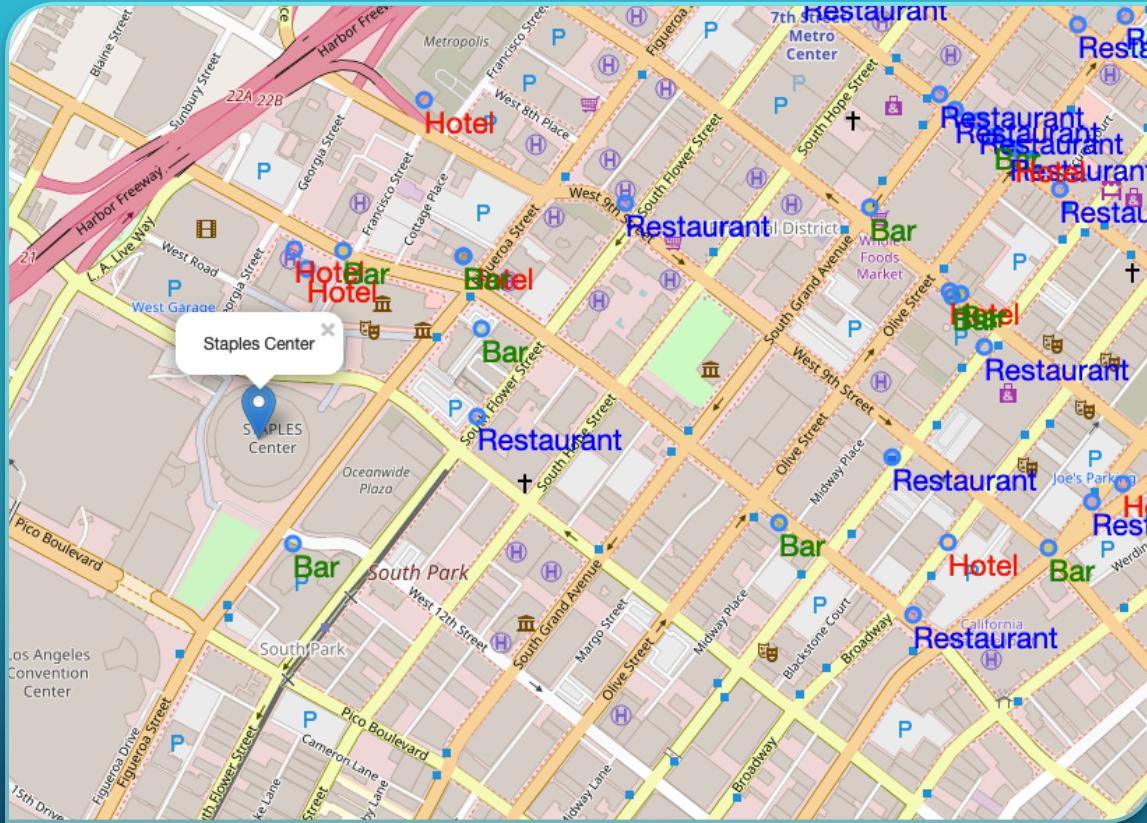
Eventually I was able to build a dataframe that combined the Top 10 Venues for each neighborhood. This was informative, but not good enough for mapping and comparing each neighborhood.

	Neighborhood	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
0	Downtown Los Angeles	Bar	Hotel	Coffee Shop	French Restaurant	Sushi Restaurant	Theater	Sandwich Place	Roof Deck	Gym	Hotel Bar
1	Mission Bay	Coffee Shop	Brewery	Art Gallery	Café	Pizza Place	Gym	Vietnamese Restaurant	Spa	New American Restaurant	Baseball Stadium

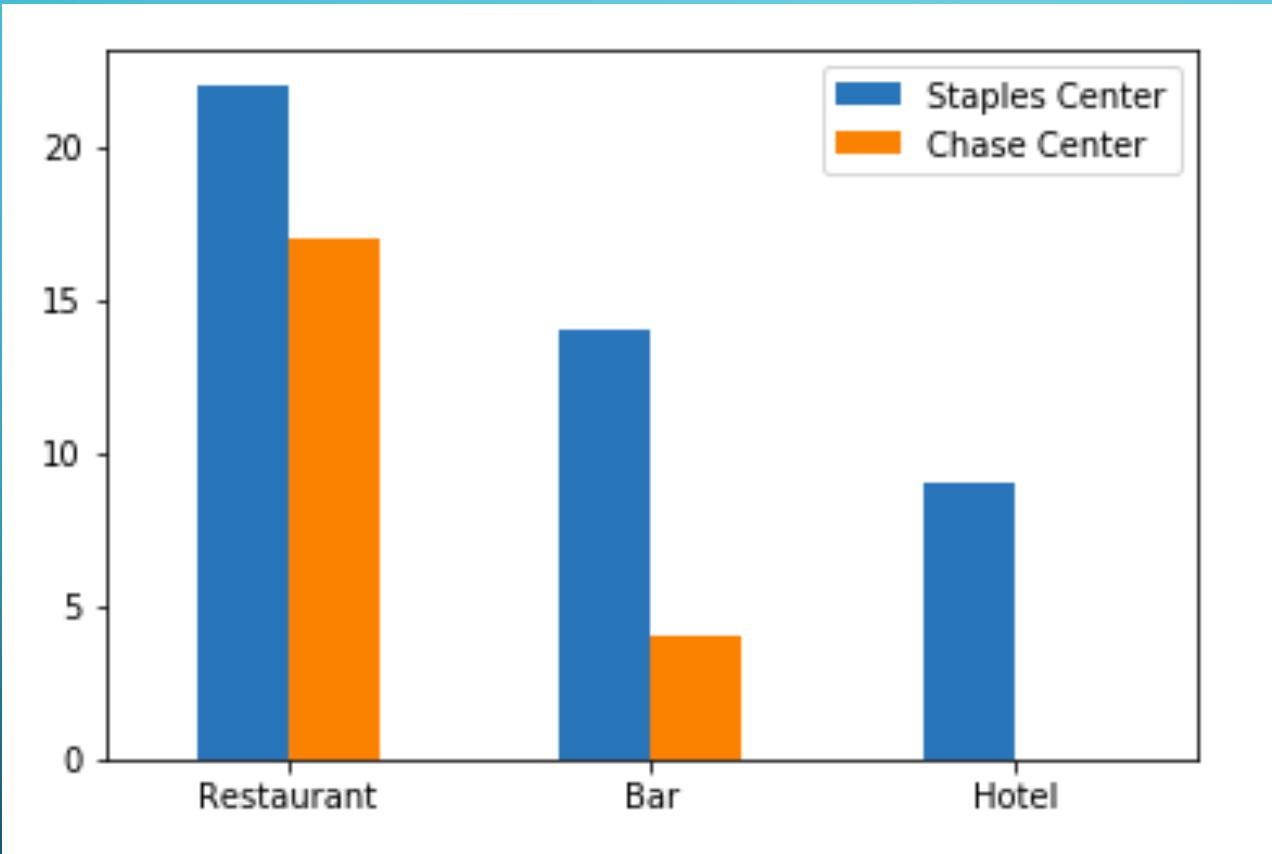


The dataframes were sliced, cleaned and combined to eventually have a dataframe that contained all of the restaurants, bars and hotels for both neighborhoods. Next, that dataframe was split by neighborhood in order to map the venues.

These two maps plot the bars, restaurants and hotels near Staples Center (left) and Chase Center (right).



Do you see any hotels near Chase Center? It appears Staples Center has more venues...



This bar graph clearly visualizes the difference in surrounding venues between Chase Center and Staples Center.



Given the clarity of the visualizations, it appears that Staples Center/Downtown LA would be a better host for a Final Four given the hotel accommodations and variety of restaurants and bars.

One important factor to consider is Chase Center is a brand new arena scheduled to open later in 2019 and the surrounding area could develop in the coming years.