**Gameday Atmosphere—Feeding the masses**

John Draper

**Introduction/Business Problem**

Football Saturdays on American college campuses are always crazy, exciting events with great potential for restauranteurs to satisfy. While there are many similarities in a variety of 'College towns', there may be some opportunities to bring variety to the gameday experience. Potentially hundreds of thousands of fans converge on these towns for the opportunity to see their team ride to victory, and we are here to make sure they are well fed.

Our company has recognized an opportunity at THE Ohio State University and desires to find what the Buckeye faithful may be missing to supplement their gameday experience. This project will explore the 120 Division I NCAA stadiums and the surrounding area to characterize the gameday environment. We will then cluster the stadium environments by size of the stadium as well as the venues in the vicinity with special focus on bars and restaurants. Once the cities/stadiums are clustered, we will more closely explore the 'cluster' for Ohio State and find potential opportunities for growth rather that be in variety (what do similar teams have that we don't?) or rating (are the current local establishments rated low providing opportunity to acquire business?).

Overall, this project will examine the collegiate gameday environments in total to find if the similarities are more due to geographic location, conference affiliation, city size, or prestige of the football team. This analysis provides a deeper dive into understanding the underlying ecosystem of businesses surrounding college campuses.

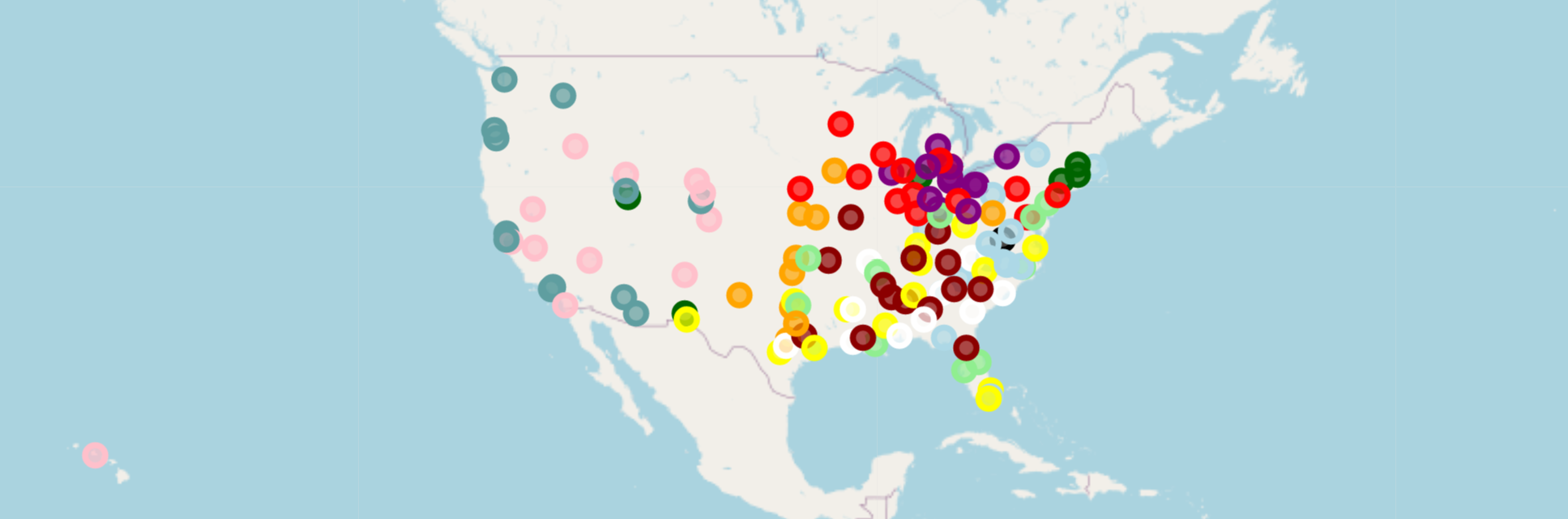
**Data**

The data for this project comes from Wikipedia1 containing the location, construction year, and capacity of all 120 FBS NCAA football stadiums. While additional information concerning the team and type of turf is provided, that information will not be used to cluster. However, conference affiliation will be used in examining the clusters after fitting the model to see if conference is related to the gameday enviroment. The population of each of the college towns will be collected via OpenDataSoft API. This data will also be used to create an interactive map to explore the distribution of collegiate environments in the United States.

Finally, the stadium/location data will be joined with detailed information from foursquare to explore venues near the stadia as well as discover rating information. Foursquare provides information about the venues in the surrounding areas of the stadium. The focus will be on venue types related to entertainment on gameday (restaurants and bars, primarily). The crowd-source app also contains rating information that will be utilized to determine the quality of said restaurants in looking for a niche to fill.

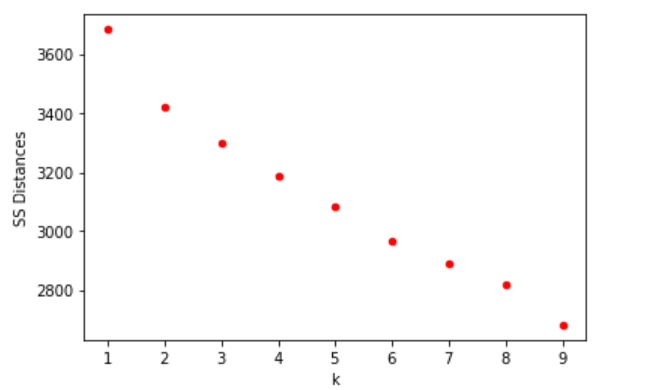
**Methodology**

We began with an exploratory analysis looking at the geolocation of the stadia around the country. As the map below shows, the stadia are primarily located in the eastern and central US. There are certainly colleges in the Western portion of the country, but most of the north central/non-coastal west is bereft of Division I colleges. This is likely due to terrain, population sparsity, and fewer large cities. It should be noted that the data scraped from Wikipedia needed to be cleaned to remove footnotes, commas, etc.

****

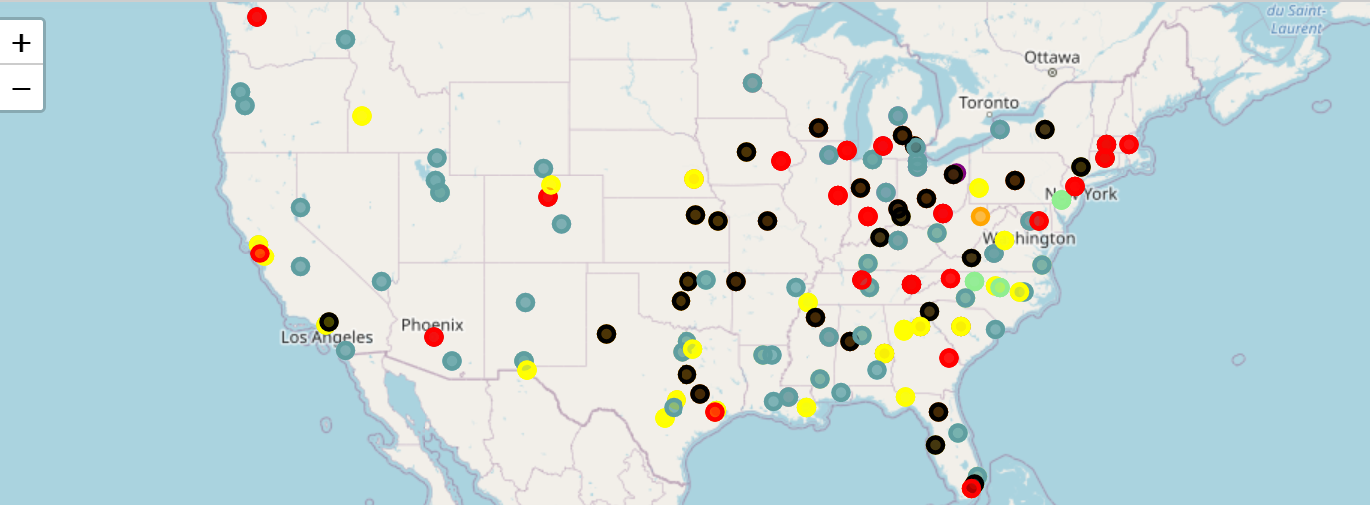
Next, the data of nearby venues was accessed via the Foursquare API. The information was collected similar to the previous project in which the most common venues in the vicinity of the stadium were discovered. Every business (up to 100 per stadium) within a radius of 1500 meters (~1 mile) was included, and the percentage of each type of business was calculated. Only businesses that are more common restaurants and bars are included. In addition, city population, the year the stadium was built, and stadium capacity were appended to provide additional information about the stadium/town. Determining 'food entities' was somewhat painstaking as the venues were combed through to determine their true category. Obviously, 'restaurants' and 'bars' were easily found, but 'BBQ Joint', 'Food Truck', etc. were more nuanced.

This data was then normalized for input into the kMeans clustering algorithm. Various values of k were explored to find the 'elbow' in variance explained. While a clear elbow wasn't observed (see below), a choice of 5 clusters was attempted to find potential groups in the data.

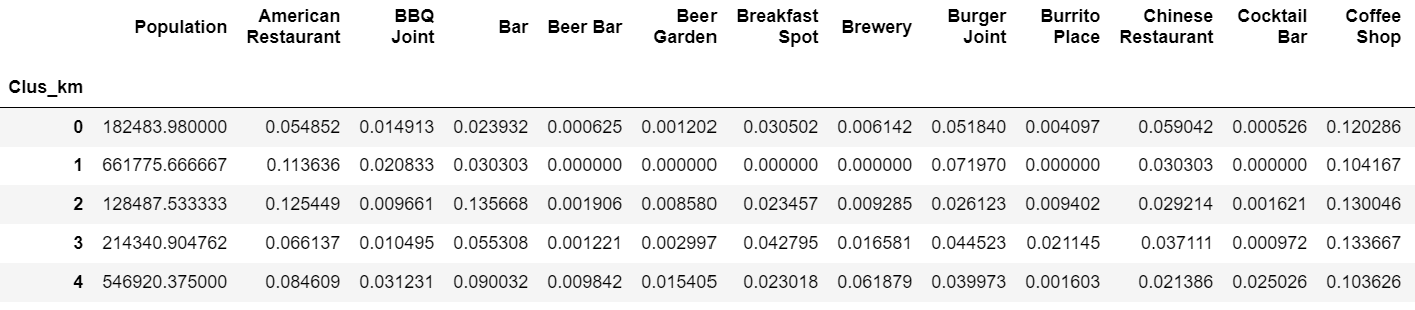


**Results**

We begin with a map of the stadiums color coded by cluster.



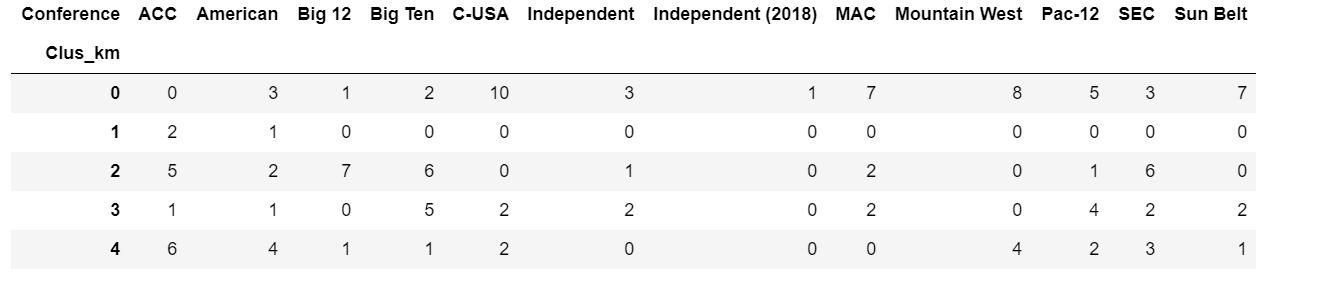
The means of the variables of interest across cluster showed the following:



****

****

Conference distribution across clusters was also examined.

****

**Discussion**

Based on the map, we find the following:

* The Blue cluster (cluster = 0) appears spread evenly throughout the country.
* The Green cluster (cluster = 1) is the smallest cluster with only a few cities in North Carolina and Philadelphia.
* The Black cluster (cluster = 2) appears primarily in the eastern United States (Midwest) and in a wide size range of cities. Ohio State lies within this cluster.
* The Red cluster (cluster = 3) is primarily in the middle of the country (again more Midwestern) but in larger cities cities.
* The Yellow cluster (cluster = 5) is primarily in the southern US. Note: Hawaii was in this cluster but cut off for legibility. Below, you can find a summary of the clusters and a brief description of each cluster.

We will now describe the clusters focusing on the nuances present.

**Cluster 0: Small School, Small City, Americana**

Cluster 0 contains many of the midmajor schools spread throughout the country. Examples are Northern Illinois, New Mexico State, and Toledo. Larger Power 5 schools also enter this cluster (Mississippi State, Virginia, Washington State), but these schools are usually in smaller, more rural towns. The food tastes generally run simple with Fast-Food and Pizza dominating the menu.

**Cluster 1: Nothing nearby (somewhat confusing)**

Cluster 1 is the smallest cluster containing only NC State, Wake Forest, and Temple. Interestingly enough, there are very few venues listed on foursquare of the restaurant variety. This is likely do to a stadium in a more rural area (I’ve been to NC State and the stadium is hidden). It is somewhat confusing to find Temple in this cluster in downtown Philadelphia.

**Cluster 2: Hitting Bars in Middle America with Big Football**

Cluster 2 contains some many of the power 5 schools that have large stadiums (i.e., Ohio State,Penn State, Oklahoma State, Arkansas, Alabama, Texas A&M). These schools have large capacity stadiums (usually due to strong teams) but the average city population is a little smaller than expected. The selection of restaurants is eclectic, but the item that stands out is bars and American cuisine. Cluster 2 has the highest number of bars near the stadium. This isn’t terribly surprising if you’ve attended a high level football game in the Midwest. Ohio State is in this cluster so we will examine it in more detail later.

**Cluster 3: Mid-level Football in a large city (but less booze)**

Cluster 3 contains many of the lower tier teams from Power 5 schools (Boston College, Arizona State, Indiana, etc.). It also contains some group of 5 schools in larger cities (FIU, Rice, etc.). The taste are spread evenly across the board, but the emphasis on bars found in cluster 2 isn’t present.

**Cluster 4: Big City with more modern, hipster tastes**

Cluster 4 also runs the gamut of tastes with no real surprises, but more modern, hip venues seem to arise more often (Gastropubs, Seafood, fewer bars). The cities are very large (Atlanta, Austin, LA, Miami, etc.) and traditional spots for a more youthful demographic. Outside of cluster 0 with the small schools, cluster 4 shows the largest presence in the Western US which fits this mold.

**Conclusion: So what should I do at OSU?**

When comparing OSU to the average of its cluster, a few things become readily apparent.

1. Columbus is saturated with bars near the stadium (23.5% of venues compared to the average of 13.6%)
2. There is a lack of American restaurants (0% vs. 12.5% for the average of the cluster).
3. Pizza is also better represented in Columbus than the cluster (17.6% vs. 11.4%)

It appears that trying to enter an already heavily saturated bar scene may be a mistake unless a niche can be exploited (pubs, and beer bars were also more represented at OSU than the cluster). Pizza is also well represented. The void appears to be in more traditional American cuisine whether that be a restaurant, food truck, or fast food restaurant—although burger joints are slightly higher than the average. Food trucks with traditional cuisine may be the sweet-spot to fill a need. Chicken finger fast food restaurants are pretty popular now so that might scratch the itch of the Buckeye fan who might have tied a few on after the big win.