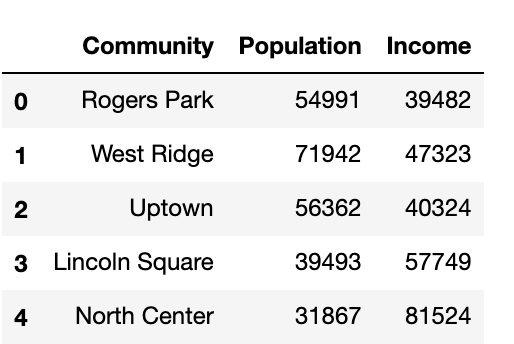
Explore Art & Entertainment in Chicago City

# Introduction

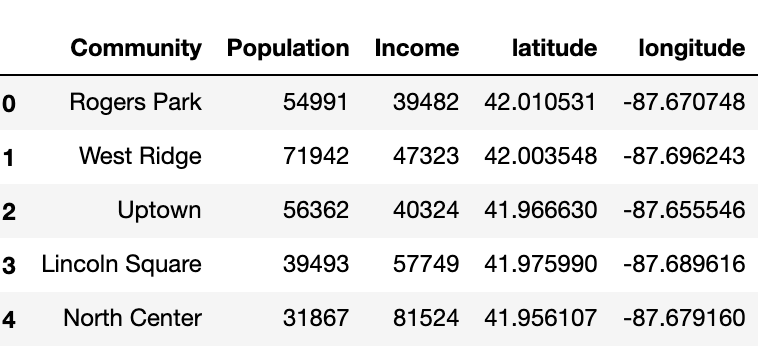
Chicago is one of the populous cities in the Mid-west of the United State. It is famous about the architectures, arts, music and academic environment. Every year people go to Chicago for different reasons. As an art and Music lover, I want to explore the art and entertainment of different Community in Chicago. As we know, foursquare API can provide us venues by category, combining with our geo data, I will find more insight about how to have fun in the Windy City.

# Data

1. use opencage.geocoder to get lat, lng
2. **Chicago-community-areas.csv** : community names with population and incomes.

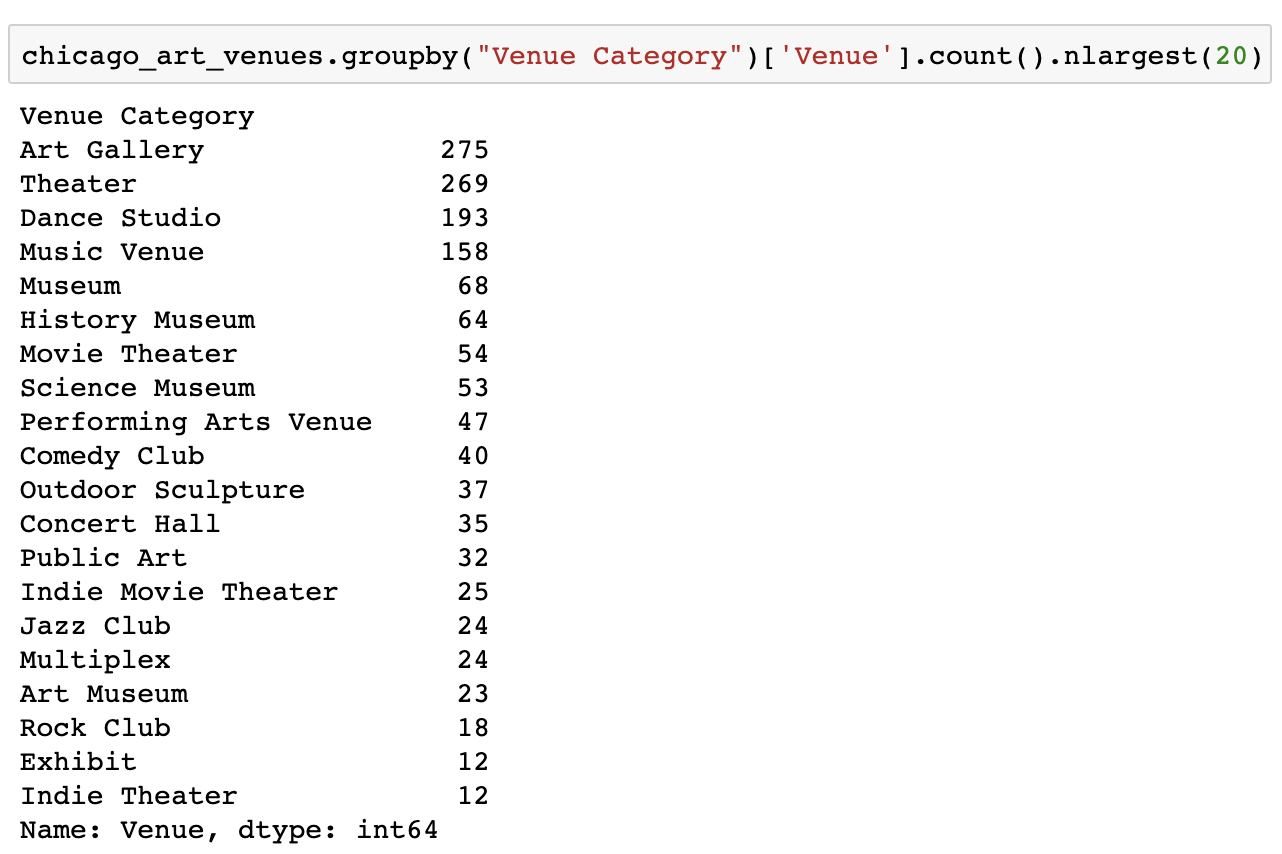


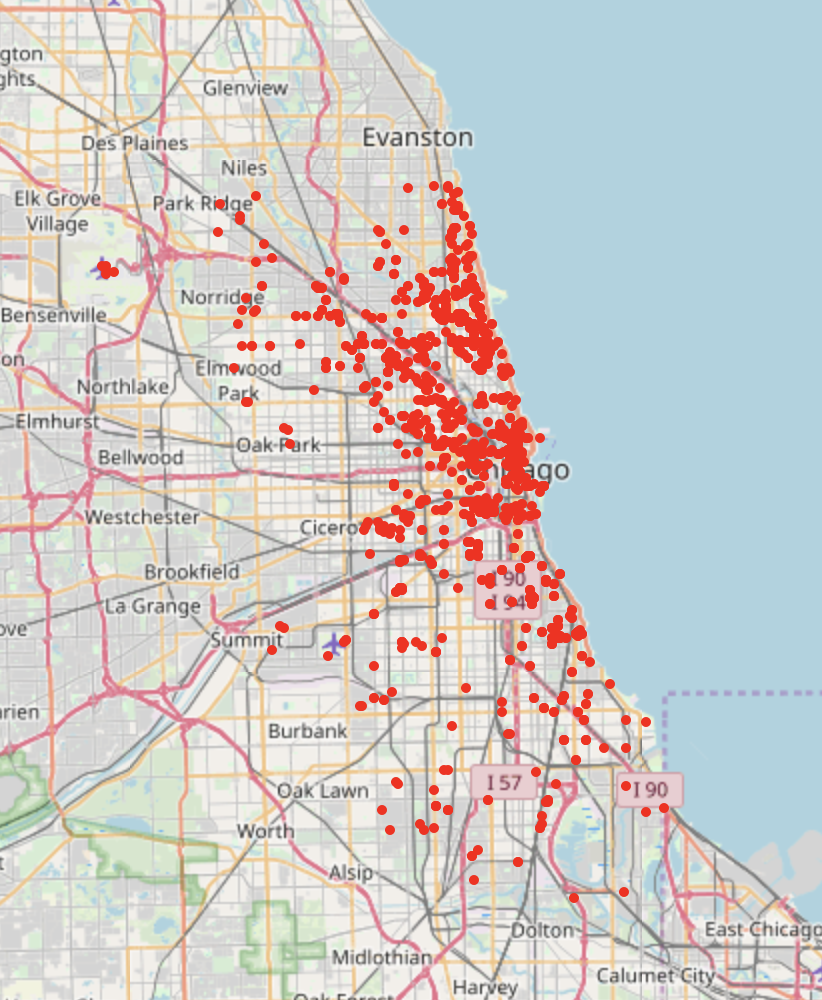
1. **chicago\_community.csv :** combined 1 and 2



1. **chicago\_art\_venues.csv** is the cleaned data contains all the art and entertainment venues in Chicago from foursquare

# Exploratory Data Analysis

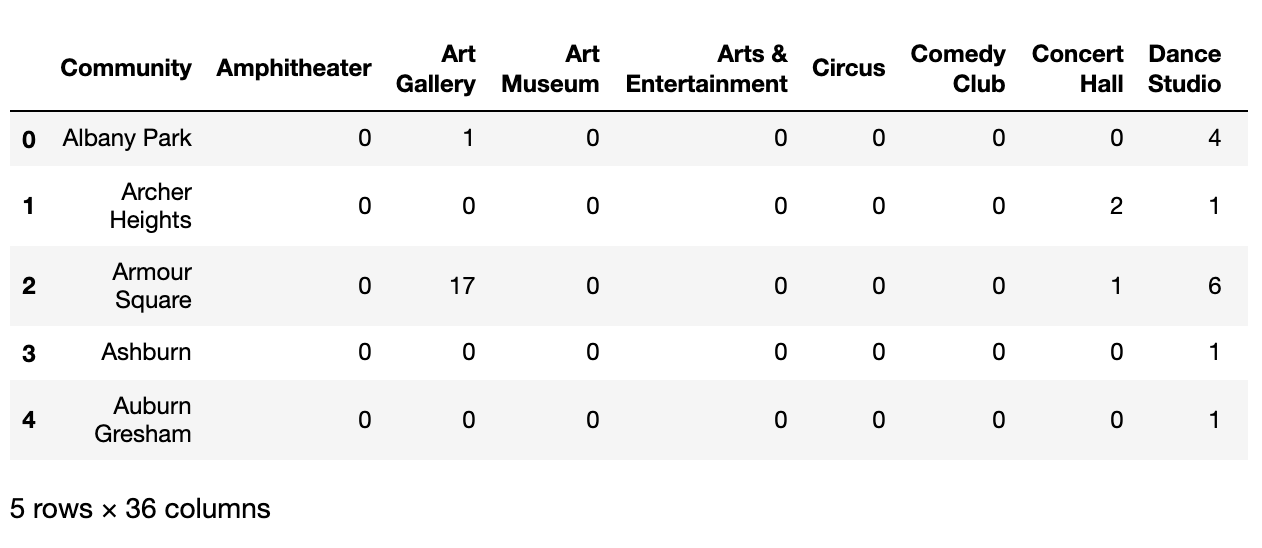
We print out the top 10 art and entertainment categories



Chicago Art and Entertainment Venues Map, there are 35 art and entertainment categories in the whole city with the total amount of 1540 venues. We can see that Chicago is a very art friendly city with a lot of art galleries and theaters. There are also many dance studio and music venues which means that people love dancing and music very much. The amount of museums (adding up museum and history museum, as well as science museums) are also greater than a hundreds, these implies that it must be a culture and historical center of the mid-west area.



# Clusters

* One-hot Encoding for the venue categories.
* GroupBy on the Community column and obtain the mean of the one-hot encoded venue categories.
*  Transpose the data-frame at step 2 and arrange in descending order.

Because K-Means algorithm is one of the most common cluster methods of unsupervised learning. I tried Elbow Method and Silhouette Method For Optimal k and the k-cluster = 4.

# 

# 

# 

# Discussion

# 

Cluster 0 (Red) : A lot of Dance Studio, Art Gallery, theatre, music venue.

Cluster 1 (Purple) : Mainly West Part, with Art Gallery as most common venues.

Cluster 2 (Blue): Up North, mostly theater and music venues.

Cluster 3 (Green): South Part, Top venues are all different kind of museum.

From the cluster result, we can see Cluster 0 is around the city except downtown, indicating that most community of the city has very good art environment that allows you to music, dancing and art. Communities from Cluster 1 are mostly located on the west and north. From my personal experience, these communities are mostly young and talented people that came to the city for their dream career. That makes much sense that these place usually have very affordable rent price and always art friendly so that galleries and music venues are the biggest part for entertainment there. The most common venues in Cluster 2 is theatre. As we know theater are always for a bit elder people who has everything settled and like high class elegancy. Since Cluster 2 contains downtown to up north by the lake, those communities are mostly expensive and classic. Cluster 3 in on the south-side and mostly different kind of museum. We know south-side is where university of Chicago at, a world-famous university for both art and science. All those museums are main surrounding the university, providing the best historical and scientific learning resources for the communities.