**Relationship between United States Metropolitan GDP and Population Venue Preference**

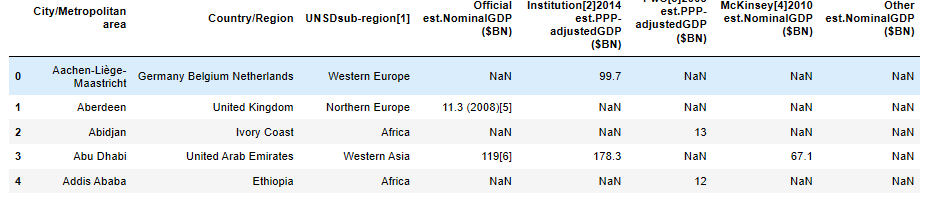
**Fanqin Zhou 10/11/2019**

**1. Introduction**

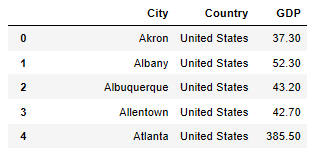
Money can't buy happiness, but would money change the definition of happiness? Wealth affects human behavior and preference, as learned from Psychology 101 in college. Studies shown that money alters how you value your time and effort. In fact, average GDP is a measurement of the wealthiness of a given city. Does wealth have anything to do with people’s value? How about ethics? Does lack of wealth or GDP of a city effect addiction? Alcohol consumption?

**2. Data acquisition and cleaning**

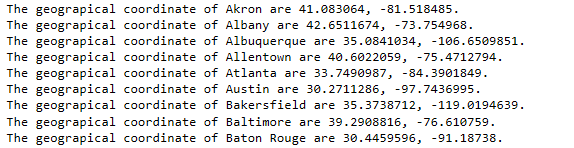
From Wikipedia (https://en.wikipedia.org/wiki/List\_of\_cities\_by\_GDP). We obtained data for metropolitan GDP around the worl. In this study, we are only focusing on the United States.

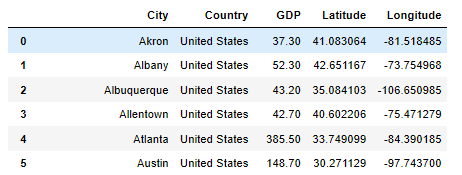


We will be focusing on GDP only. We will drop the irrelevant columns and clean up the data. Worcester does not have a GDP value in this table, thus we will ignore this city.



We will be using foursquare location data to obtain geographic coordinates for each city:





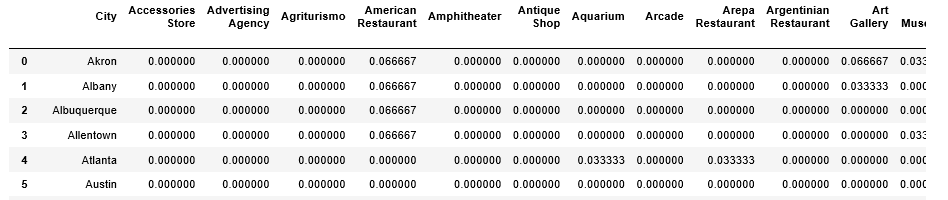
Now, we have obtained our desired table for this study.

**3. Popular Venue Generation**

Using foursquare location data, we can get venue info around a certain latitude and longitude, with a radius of 500, we will find the top 30 most popular venues in a city:



We analyze each city on their most popular venues:

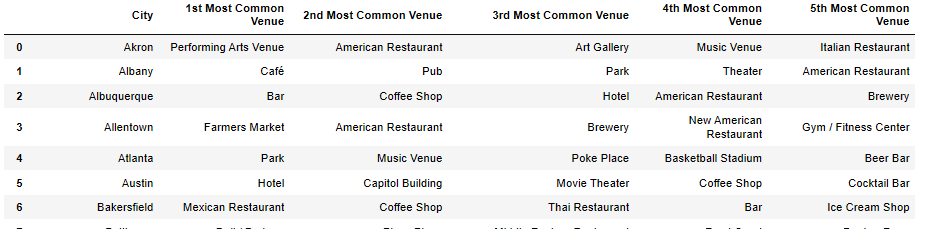


From this, we generate the top 10 venues for the city:



**3.1 Categorization**

Notice we really have a lot of too detailed venue type, such as: “fast food restaurant”, “American Restaurant”, “New American restaurant”, etc. It is necessary to group them into more inclusive categories.



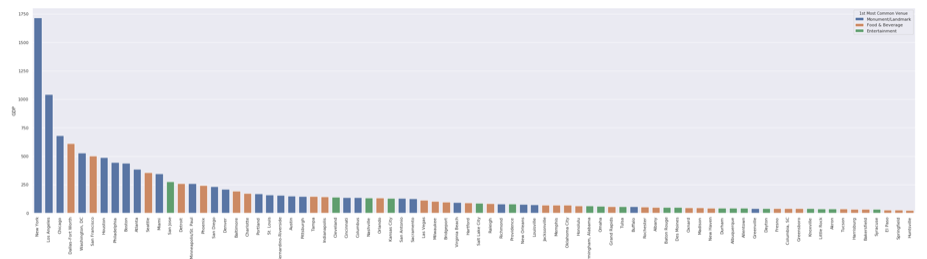
In this study, we are mostly interested in the GDP effect on 3 big categories: “Monument/Landmark”, “Food & Beverage” and “Entertainment”.



After transformation, our table looks like this:

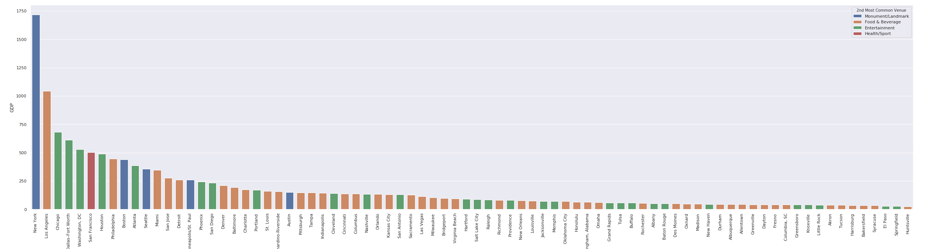


**4. Finding**

By looking at cities from higher GDP to lower GDP, we see that "Food & Beverage" spread evenly throughout the GPD span, "Monument/Landmark" seems to be most popular among high GPD cities, "Entertainment" seems to be most popular among lower GPD cities. 

*In graph, blue indicates “Monument/landmark”, orange denotes “Food & Beverage”, green denotes “Entertainment”.*

let's see of the same holds true for 2nd popular venues



*In graph, blue indicates “Monument/landmark”, orange denotes “Food & Beverage”, green denotes “Entertainment”.*

while Monument/Landmark lost popularity, same rule still holds true. However, Entertainment seems to be distributed evenly amongst cities.

**5. Conclusion**

From data, we can conclude that “Monument/Landmark” is the most popular among high GDP cities while “Entertainment” is the most popular among lower GDP cities. While this does not indicate any casual relationship, we can see that higher GDP does have an effect on people’s preference. It is also interesting to see that GDP does not influence “Food & Beverage” a lot.