

# 1. To Analyze the Most Visited Cities in the US

The objective of this analysis is to find out

1. Where the cities are located
2. What are the main venue categories from these cities
3. The relationship of the population and GDP with the number of visitors
4. Observations and Recommendations

## 2. Data to be used

Top 10 visited cities in the US is based on the information from the web site below. Data is captured in csv format and saved in IBM Watson asset storage.

<https://www.worldatlas.com/articles/the-most-visited-cities-in-the-us.html>

Rank	City	Number Of Visitors (Millions)
1	New York, New York	59.7
2	Chicago, Illinois	54.1
3	Atlanta, Georgia	51
4	Anaheim/Orange County, California	48.2
5	Orlando, Florida	48.0
6	Los Angeles, California	47.3
7	Las Vegas, Nevada	42.9
8	Philadelphia, Pennsylvania	42
9	San Diego, California	34.9
10	San Francisco, California	25

Nominal GDP information is imported directly from Wikipedia at the url below. A portion of the table from Wikipedia is screen captured as below.

[https://en.wikipedia.org/wiki/List\\_of\\_U.S.\\_metropolitan\\_areas\\_by\\_GDP](https://en.wikipedia.org/wiki/List_of_U.S._metropolitan_areas_by_GDP)

Nominal GDP for the top 50 metropolitan statistical areas (Millions of dollars)<sup>[2]</sup> [\[ edit \]](#)

2017 Rank ↕	Metropolitan area ↕	2017 ↕	2016 ↕	2015 ↕	2014 ↕	2013 ↕	2012 ↕
1	<a href="#">New York-Newark-Jersey City, NY-NJ-PA (Metropolitan Statistical Area)</a>	1,717,712	1,662,671	1,618,366	1,542,763	1,477,043	1,439,233
2	<a href="#">Los Angeles-Long Beach-Anaheim, CA (Metropolitan Statistical Area)</a>	1,043,735	996,432	967,101	901,980	852,034	820,863
3	<a href="#">Chicago-Naperville-Elgin, IL-IN-WI (Metropolitan Statistical Area)</a>	679,699	657,589	639,033	608,805	585,948	578,016

Population and city coordinates information are obtained from the website below. Data is captured in csv format and saved in IBM Watson asset storage. A portion of the table from the website is screen captured as below.

<https://public.opendatasoft.com/explore/dataset/1000-largest-us-cities-by-population-with-geographic-coordinates/table/?sort=-rank>

# 1000 Largest US Cities By Population With Geographic Coordinates

Information

Table

Map

Analyze

Export

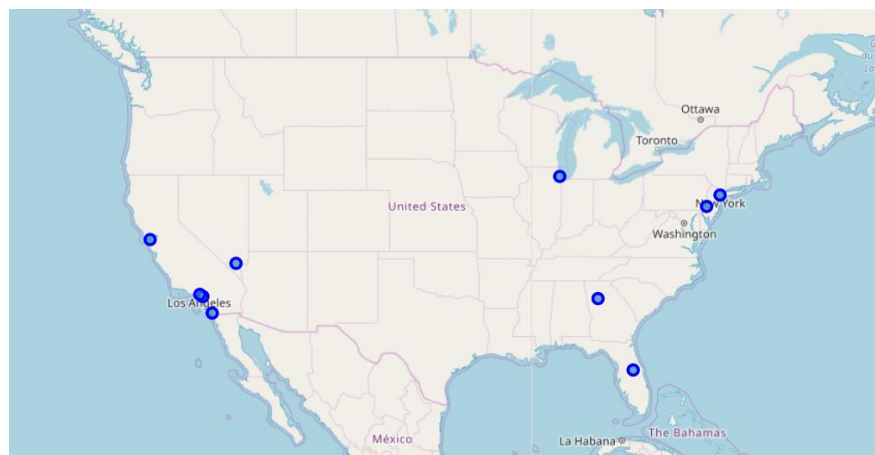
API

<

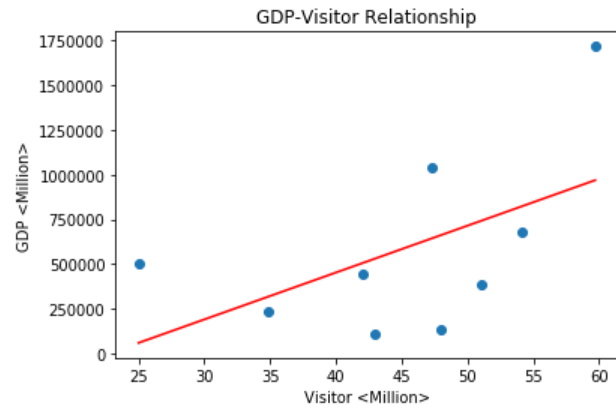
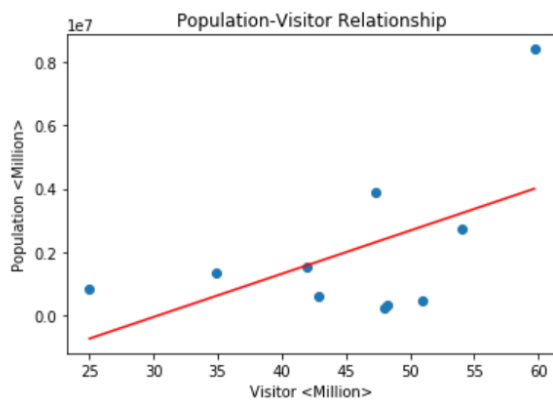
Venues information are obtained from Foursquare.

### 3. Results, Observations and Recommendations

The top 10 most visited cities are located at either West Coast or East Coast. We would expect the visitors are mostly either business travelers or tourists. It seems the central region is not the people's favor of choice for vacations or business travels.



By looking at the relationship between the city population and GDP with the number of visitors, we can see that visitors favor large cities and they are also one of the contributing factors to the GDP.

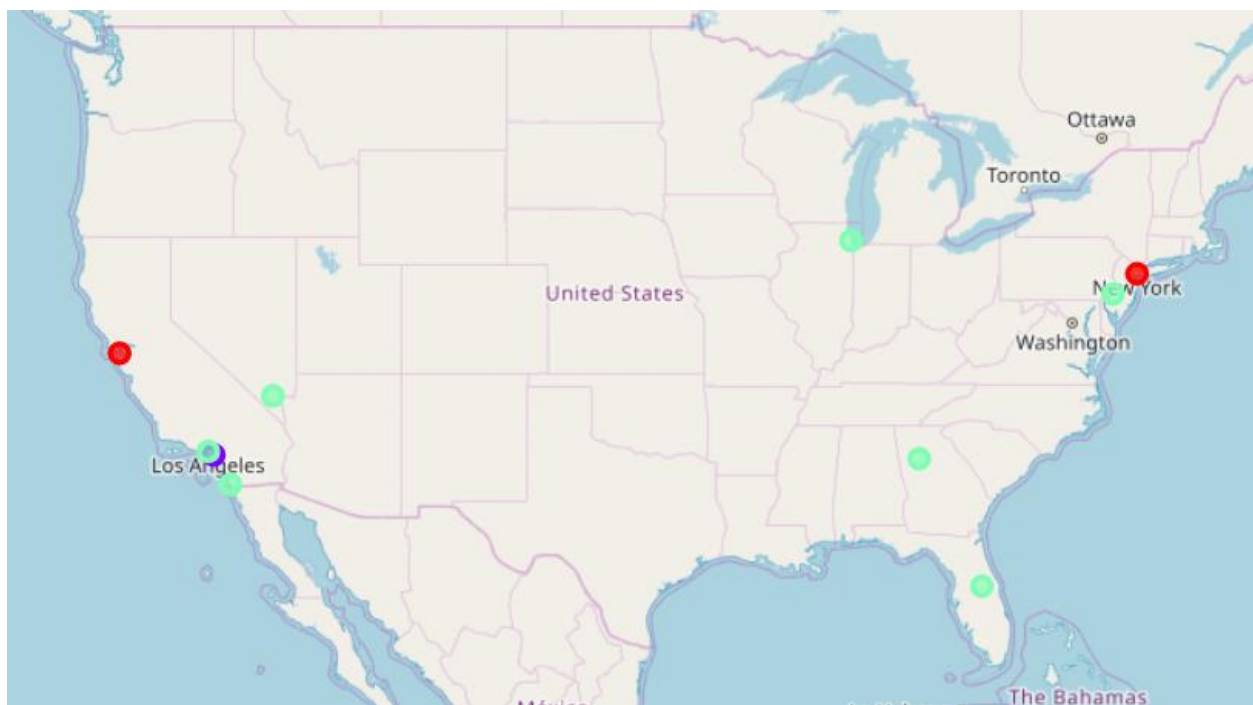


3 clusters are generated by K-means algorithm.

Cluster 1 (Red) includes New York and San Francisco. It is likely that they are clustered together by venues that are favorable by business travelers.

Cluster 2 (Purple) includes Anaheim only. It is because of its Theme Park and Entertainment venues attracting tourists.

Cluster 3 (Green) includes Chicago, Atlanta, Orlando, Los Angeles, Las Vega, Philadelphia and San Diego. They are probably clustered together due of their venues good for both business travelers and tourist.



By looking at the top 5 venues, we observe that the categories including Theme Park, Attraction, Scenic Lookout, Hotel, Zoo, Casinos, Restaurants, Wine Bar and Coffee Shop are the common categories from these most visited cities. They are in fact the venues good for tourists and business travelers.

----Anaheim----			----Las Vegas----			----Orlando----		
	venue	freq		venue	freq		venue	freq
0	Theme Park Ride / Attraction	0.31	0	Sushi Restaurant	0.04	0	Bar	0.05
1	Theme Park	0.13	1	Steakhouse	0.04	1	American Restaurant	0.05
2	General Entertainment	0.04	2	American Restaurant	0.04	2	Café	0.04
3	Hotel	0.03	3	Casino	0.04	3	Wine Bar	0.03
4	Burger Joint	0.03	4	Theater	0.04	4	Theater	0.03
----Atlanta----			----Los Angeles----			----Philadelphia----		
	venue	freq		venue	freq		venue	freq
0	Trail	0.09	0	Coffee Shop	0.08	0	Coffee Shop	0.08
1	American Restaurant	0.05	1	Ice Cream Shop	0.05	1	Pizza Place	0.04
2	Park	0.05	2	Theater	0.05	2	Wine Bar	0.03
3	Brewery	0.04	3	Brewery	0.05	3	Steakhouse	0.03
4	Mediterranean Restaurant	0.03	4	Taco Place	0.04	4	Indian Restaurant	0.03
----Chicago----			----New York----			----San Diego----		
	venue	freq		venue	freq		venue	freq
0	Hotel	0.09	0	Park	0.14	0	Zoo Exhibit	0.09
1	Park	0.06	1	Bookstore	0.04	1	Park	0.09
2	New American Restaurant	0.05	2	Bakery	0.04	2	Hotel	0.05
3	Theater	0.05	3	Cycle Studio	0.03	3	Sushi Restaurant	0.04
4	Sandwich Place	0.03	4	Scenic Lookout	0.03	4	Seafood Restaurant	0.04
						----San Francisco----		
							venue	freq
						0	Yoga Studio	0.06
						1	Coffee Shop	0.05
						2	Grocery Store	0.05
						3	Park	0.05
						4	Ice Cream Shop	0.05

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Anaheim	Theme Park Ride / Attraction	Theme Park	General Entertainment	Gift Shop	Hotel
1	Atlanta	Trail	Park	American Restaurant	Brewery	Mexican Restaurant
2	Chicago	Hotel	Park	Theater	New American Restaurant	Grocery Store
3	Las Vegas	Casino	American Restaurant	Steakhouse	Sushi Restaurant	Theater
4	Los Angeles	Coffee Shop	Ice Cream Shop	Theater	Brewery	Taco Place
5	New York	Park	Bookstore	Bakery	Scenic Lookout	Yoga Studio
6	Orlando	American Restaurant	Bar	Café	Sushi Restaurant	Theater
7	Philadelphia	Coffee Shop	Pizza Place	Steakhouse	New American Restaurant	Wine Bar
8	San Diego	Zoo Exhibit	Park	Hotel	Seafood Restaurant	Sushi Restaurant
9	San Francisco	Yoga Studio	Ice Cream Shop	Coffee Shop	Grocery Store	Park

**Recommendation:**

As mentioned earlier, central region is not the people's favor of choice for vacations or business travels and its cities should build and promote these common venues in order to attract visitors and hence improve the GDP's.