# Capstone Project

The Battle of Neighborhoods

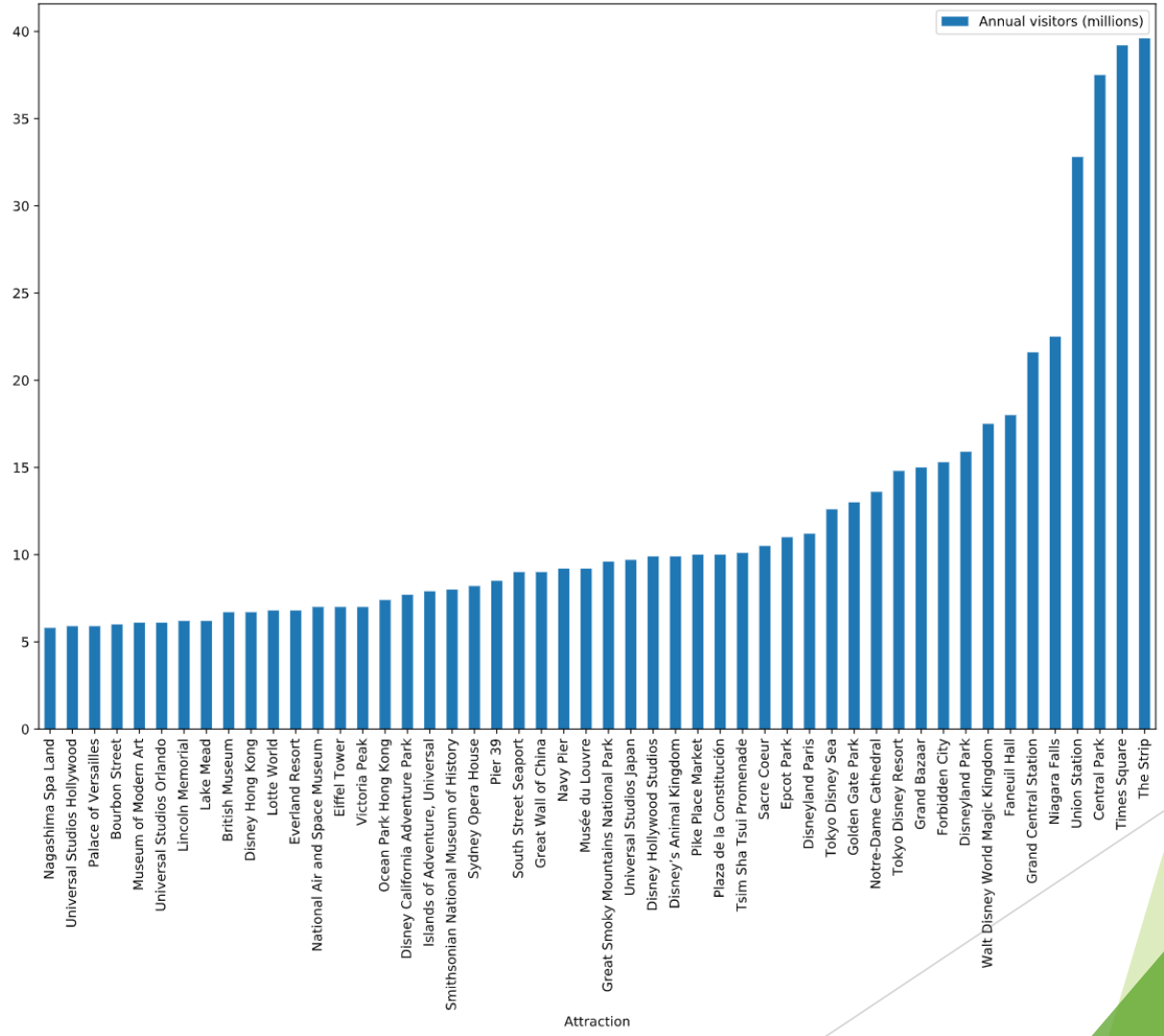
15 May 2020

# Intro

- Capstone Project for the IBM Data Science Professional Certificate
- Goal: Use Foursquare API, other data and machine learning algorithms on real world datasets
- Analysis of similiarity of top 50 most visited tourist attractions based on their recommended nearby venues

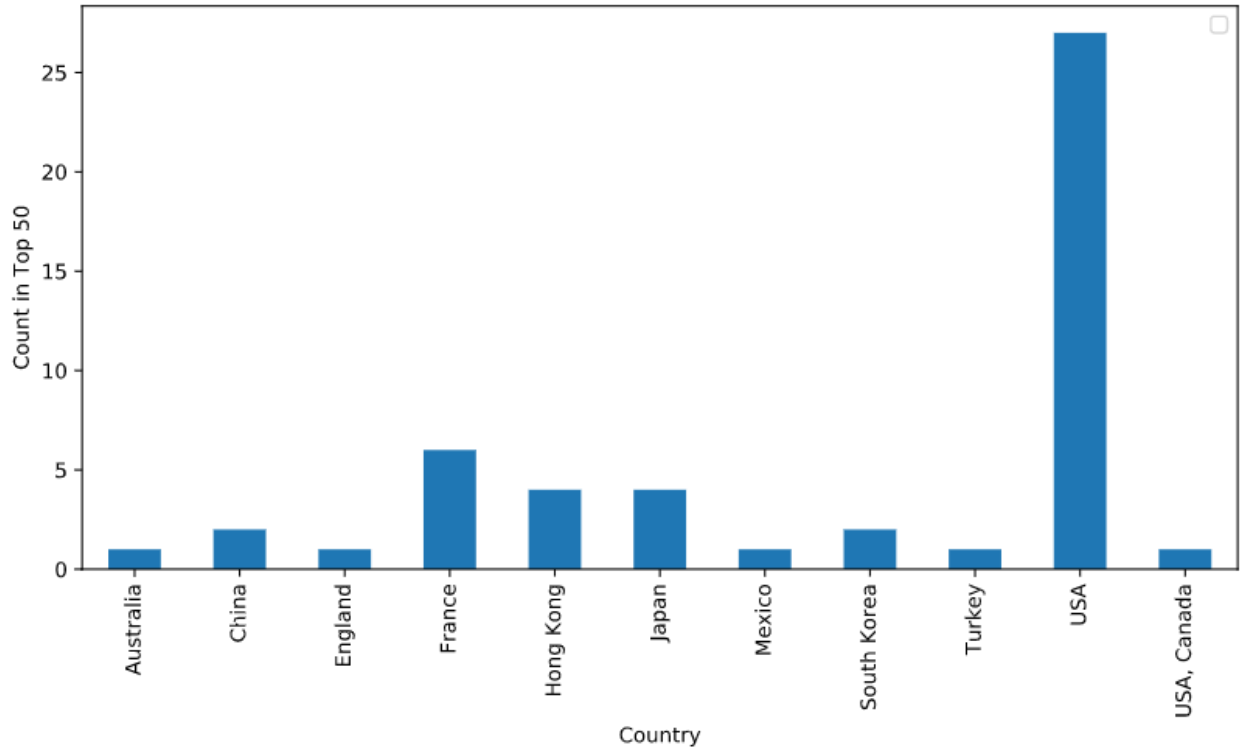
# Top 50 Most Visited Touris Attracions

- Data from blog Love Home Swap from 2015



# Top 50 Most Visited Touris Attracions

- Attractions are mainly in the US



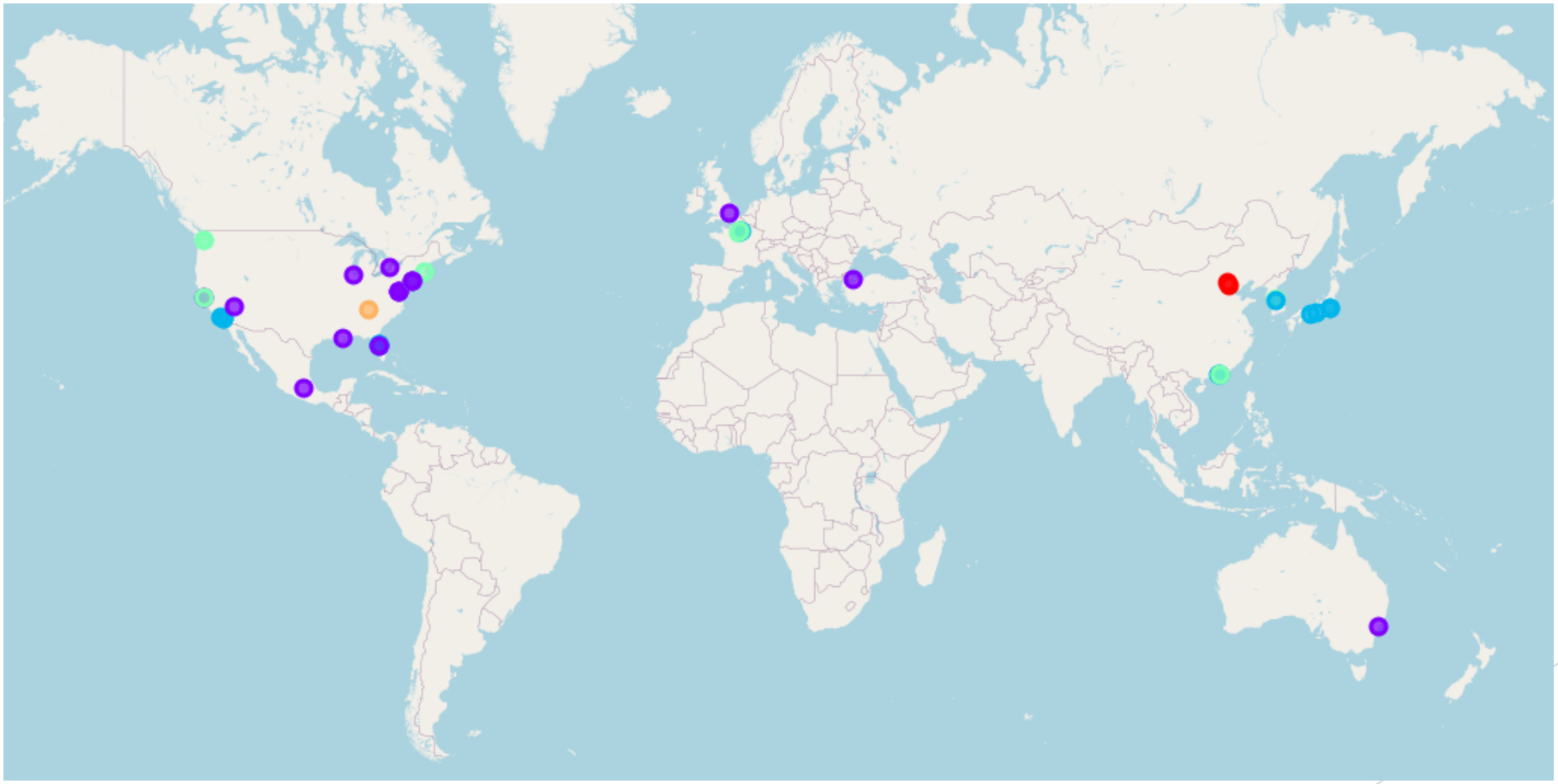
# Top 50 Most Visited Touris Attracions

# Using Forsquare API

- Collection of nearby recommended venues for each attraction
- Group by venue category (e.g. restaurant)
- Find top 10 recommended venue categories for each attraction
- Results in table as depicted below

| Attraction Name | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|---|---|---|---|---|
| Bourbon Street | Restaurant | Jazz Club | Bar | Hotel |
| British Museum | Restaurant | Theater | Hotel | Cocktail Bar |
| Central Park | Exhibit | Park | Restaurant | Garden |
| Disney California Adventure Park | Theme Park Ride / Attraction | Theme Park | Hotel | Gift Shop |
| Disney Hong Kong | Theme Park Ride / Attraction | Gift Shop | Restaurant | Hotel |

# Using k-Means Clustering to Group Attractions into 5 Clusters



# Result: Clusters

- **Cluster A:** Theme Parks such Universal Studios with mainly recommended attractions from the parks, restaurants and gift shops.
- **Cluster B:** Large downtown attractions such as the Eiffel Tower with recommendations for mainly nearby restaurants, hotels and coffee shops.
- **Cluster C:** Outdoors attractions such as Great Smoky Mountains NP.
- **Cluster D:** Chinese attractions (Great Wall, Forbidden Cities).
- **Cluster E:** Various other attractions such as Pier 39 in SF or Museums surrounded by restaurants and a wide variety of other venues.

# Discussion

- Results can be used for tourism executives to compare other, non top 50 attractions to draw conclusions on how to improve attractiveness of tourist locations by adding specific venues (e.g. more restaurants) nearby
- Business owners can find locations with a lack of their respective business (e.g. restaurants)