Influence of Venue Types on International Tourism in United States Cities

Jake Peterson

January 24, 2020

1. Introduction

1.1 Background

International tourism into the United States generated $1.2 trillion and supported 7.6 million jobs in 2011, which grew 3.5% and outpaced the growth rate of the United States economy, 1.7%, as a whole.[I1] Tourism to the United States helps local economies, job production, and benefits conservation of many historical and environmental sites.

1.2 Problem

In 2011, 62 million international visitors came to the United States, however their destinations were largely disproportionate and condensed to nearly 60 cities in the United States.[I1] Considering there are nearly 20,000 cities in the United States, it is important to determine the factors that promote international travel to these 60 United States cities in order to disperse visitors to other areas in the United States.[I2] By attracting visitors to other United States destinations, the total number of international visitors will increase, and thus, will not only increase the amount of money tourism generates but will distribute the wealth more equally throughout the United States.

One of the largest influencers on international travel destination is the types of venues within each United States city. By determining common types of venues that popular travel destinations have, less popular travel destinations can implement various types of buildings and locations that will promote international tourism. Using international tourism data and venue location data by cities, the types of venues that are most popular among cities with high international tourism rates will be determined.

1.3 Interest

The types of venues that attract international tourism will be presented to state and federal legislators to provide insight on how strategic placement of various types of venues can increase international tourism rates among other United States cities. This information can help governing bodies support local economies, provide employment, and benefit conservation of habitats and historic sites.

2. Data

2.1 Data Sources

Inbound international travel estimates by United States cities for 2017 and 2018 were acquired from TravelTrade.gov.[D1] The location coordinates for the 61 most visited cities were found using Google Maps API geocoding. The coordinates were then used to obtain the 60 most popular venues within each city using Foursquare API.

2.2 Data Cleaning

The international travel estimates data was cleaned to remove unnecessary columns and to rename all columns appropriately. The estimated visitors for 2017 and 2018 were condensed into one column containing the average visitors per year. In addition, the city format provided by TravelTrade was Metropolitan Statistical Area

(MSA), so each city area was converted to the primary city in each MSA. For example, New York-White Plains-Wayne, NY-NJ was converted to New York, NY. This conversion was performed so Google's geolocator could work correctly. Once cleaned, the data was evaluated to ensure no outliers, missing data, or duplicate data was present.

Once the individual cities and their respective average international visitors per year were determined, the location coordinates for each city were found using Google's geolocator. The latitude and longitude values were evaluated and mapped using folium to verify they were correct.

Finally, the 60 most popular venues within a 30 km radius from the city center were acquired. Information on the venue's name, location, category type, and rank of popularity within the city were gathered. Gyms and grocery store venues were removed from all cities with the assumption that these types of venues had little influence on international travel, and were likely popular due to usage by local residents. The 30 most popular venues were then taken so each city had an equal number of venues under consideration. The venue data was evaluated to guarantee there were no duplicate venues by comparing longitude and latitude values for each venue.

3. Methodology

The focus of this project was on discovering how the types of venues within major United States cities impact international tourism.

The 61 most visited cities in the United States and visitation data were acquired from TravelTrade.gov. Using Google's geolocator, the location coordinates for each United States city was acquired. Foursquare identified the most popular venues within each city.

The strength of correlation and certainty of correlation between each venue type and the average number of international visitors for each United States city was determined. The venues that demonstrated a correlation with a minimum of a weak certainty ($p < 0.1$) were determined and classified based on the value of the correlation and p-value using provided correlation definitions from DM Stat-1.[13] United States cities were clustered (using k-means clustering) by the 20 most common venue types within each city. The United States cities were remapped to explore the distribution across the United States of each cluster. By evaluating the cluster distributions, the city location and climate types were assessed to determine if the location of each city has an influence on the common types of venues within each city. This information would be useful when considering the types of common venues based on city climate and location.

4. Results

4.1 Venue Type Influence on International Visitation

Analysis of the venue type's impact on international visitation demonstrated four venue types (Fountain, Tennis Stadium, Memorial Site, and Art Museum) that have a moderate positive correlation with strong certainty, nine venue types (Theme Park, Bridge, Public Art, Park, Performing Arts Venue, Canal, College Residence Hall, Observatory, and Building) that have a moderate positive correlation with moderate certainty, two venue types (Monument/Landmark and Art Gallery) that have a weak positive correlation with moderate certainty, one venue type (Shopping Mall) that has a weak positive correlation with weak certainty, and three venue types (Fast Food Restaurant, American Restaurant, and Warehouse Store) that have a weak negative correlation with weak certainty.

Further explorative analysis demonstrates that Parks are the most common type of venue among all cities with 20 cities having Park as the most common venue type, 10 cities having Park as the second most common venue type, and 4 cities having Park as the third most common venue type.

4.2 City Location Influence on Venue Type

Clustering cities by the popularity of venue types within each city and calculating the inertia of each kmeans value demonstrated the optimal number of clusters was three. After grouping the cities into three clusters and remapping them with cluster labels on a United States map, it is clear that the geographical location of the city plays an important role in the types of common venues in each city. Cluster 0 contains cities that have common venue types of Park, Scenic Lookout, Trails, and Plaza and cities in this cluster have an average of 2,991,000 visitors per year. Cluster 1 cities have common venue types of Beach, Surf Spot, and Hotel and cities in this cluster have an average of 1,360,000 visitors per year. Cluster 2 contains cities with common venue types of Park, Warehouse Store, Supermarkets, and Fast Food Restaurants and cities in this cluster have an average of 420,000 visitors per year.

5. Discussion

International tourism is an important aspect of the United States economy and understanding the factors that influence tourism rates is important. The type of venues that are located within United States cities play a significant role on the rate of international tourism. It was determined that fountains, tennis stadiums, memorial sites, and art museums have the greatest positive correlation to international tourism rates with strong certainty. In which case, it would be in governing bodies best interest to build these types of venues in cities throughout the United States. In addition, fast food restaurants, american restaurants, and warehouse stores have a negative correlation to international tourism rates, and thus, should not be implemented into city building plans.

It was visualized that the geographical location of cities has a major impact on the type of venues that are commonly present. Cities close to the coastline tend to have more

coastline specific venues such as beaches, surf spots, and resort hotels. Cities near mountain ranges and forests tend to have venues such as parks, science lookouts, and trails. Clustered cities that were scattered across the United States more commonly had venues such as parks, warehouse stores, supermarkets, and fast food restaurants. Using this information, legislators can also consider the geographical location of each city to consider which venue types may be more successful. In these cases, if a city is located near the coastline it would be most beneficial to create public beaches and surf spots. Inland cities would benefit most by building parks, art museums, and fountains while reducing the number of fast food restaurants and warehouse stores.

Further evaluation and consideration of the data indicates some potential flaws with various venue types. There were only four tennis stadiums total, and one was from New York, NY, which has the highest international tourism rate among all cities. In this case, the perception of tennis stadiums impact on international tourism rates may be relatively skewed. However, if available, future research can utilize Foursquare "checkincounts" to create a more accurate weight based on the popularity of each venue type. A more accurate correlation between each venue type and international tourism rate could then be obtained. Also, venue data specific to international tourism would help isolate the types of venues that are more important to promoting international tourism. Foursquare offers only data on the popularity of various venues, and since it is likely that international travelers don't visit all locations that local residents frequent, such as gyms and grocery stores, more specific data would provide a stronger correlation between venue type and international tourism rate. In addition, memorial sites have an interesting influence on international tourism rates. It is likely that the importance and connection of the memorial site to international tourists plays a significant role on how popular each site would be in attracting tourists. Also, many cities do not have memorial sites simply because there is no historical moment that is significant enough to build a memorial site within the respective city. In these cases, it may be important to consider a historical lens, and not the type of venues, when determining the popularity of international tourism within United States cities. For example, when comparing the 9/11 Memorial Pool in New York, NY to other memorial sites a historical lens may be more important than simply looking at the venue type when determining a memorial sites impact on international tourism. Future research can also further this research by determining areas that are likely to be visited by tourists in one trip. This would be important to consider because if two locations are closely located, than their popular venues would also likely have an impact on each individual cities international travel rates.

6. Conclusion

International tourism is an important aspect of the United States economy and understanding the factors that influence tourism can benefit United States economies, job

production, and conservation of many historical and environmental sites. Legislators should focus on building fountains, parks, art museums, theme parks, and bridges while reducing the number of fast food restaurants, american restaurants, and warehouse stores in United States cities to promote international tourism while also carefully considering the geographical aspects of each city.

Works Cited

Data Sources:

1. **https://travel.trade.gov/outreachpages/inbound.general_information.inbound_overview.asp**

Information Sources:

1. **https://travel.trade.gov/pdf/national-travel-and-tourism-strategy.pdf**
2. **http://worldpopulationreview.com/us-cities/how-many-cities-are-in-the-us/**
3. **http://www.dmstat1.com/res/TheCorrelationCoefficientDefined.html**