Group One

Alvaro Aquino, Emmanuel Garcia, Frantzy Francois,

Jonathan Pokorny, and Mickey Young

GitHub Link: https://github.com/mshawn12/group1-ticketmaster

Ticketmaster

Hypothesis:

- What events cost more money to attend?

- Are events more expensive in certain states?

Project Description & Outline

In order to test the Group's hypothesis, the team leveraged the Ticketmaster API along with a number of tools and tactics such as Python, Pandas, NumPy, Matplotlib, HVPlots, and various export methods. Using loops, lists, and DataFrames, the team pulled everything from the Event Name, Date, Location (City, State, Longitude, Latitude), Venue, Event Type, Genre, Sub-Genre, Minimum Ticket Price ($), and Maximum Ticket Price ($). Once the core data was collected, the team used a series of filters and visualizations to better understand the story the data was telling. Over (15) DataFrames, (7) map plots, and (12) bar charts were developed to generate the team's findings and analyses.

Datasets to Be Used

Ticketmaster API

Goals

1. To implement what we learned in class to attack our first ever data science project.

Research Questions to Answer

- Music vs Sports
  - Number of Events Analyzed by Genre
    - Is country music more popular in the northeast or south

  - Number of Events by Sport
- Highest Average Price Across the Country by Sport and Music Genre

Rough Breakdown of tasks

1. By becoming familiar with the documentation we learned how to use the Ticketmaster API and identify what data can be extracted using the free version
2. Ensure all the proper libraries are imported to perform task
3. Leverage cURL method to request event data
4. Generate core DataFrames to house all event data
5. Share core Jupyter Notebook file with team in order to do individual testing
6. Create loops and lists in order to extract and store data
7. Clean data
8. Establish relevant data filters and additional DataFrames
9. Discuss findings as a team to determine what hypothesis we we want to address
10. Create Visualizations to tell story
11. Reformat visualizations
12. Export visualizations
13. Construct Powerpoint to present findings
14. Create presentation script & conduct dry runs and timing
15. Record live presentation
16. Create GitHub repository

Conclusion

1. Basketball had the highest average price per ticket in sports, and rock had the highest average price for music.

		Min Price $	Max Price $	Average Price $
Genre	**State**			
Basketball	**Arizona**	114.625000	3764.000000	1996.625000
	Wisconsin	77.000000	3310.000000	1732.000000
	Oklahoma	27.500000	2052.000000	1053.500000
	Utah	28.000000	2000.000000	1028.000000

2. An interesting observation was that Wisconsin and Arizona were the most expensive locations for ticket sales.

	Min Price $	Max Price $	Average Price $
State			
Arizona	114.625000	3764.000000	1996.625000
Wisconsin	77.000000	3310.000000	1732.000000

3. Hockey tickets were the most expensive in North Carolina.

4. There are many types of factors that can influence the data sets, such as time of year. California and Arizona are much more pleasant destinations in the winter, as opposed to the Northeast. Also, music artists announce more concerts in the spring and summer, likely because many sports run in the fall and winter.

5. We didn't expect to see Rock ticket prices being the most expensive.

Genre	Average Price $
Basketball	1327.488636
Hockey	533.812500
Rock	299.479688
Country	204.929839
Motorsports/Racing	94.250000

6. We thought Country music would be more popular in the South, but it was actually more popular in New York
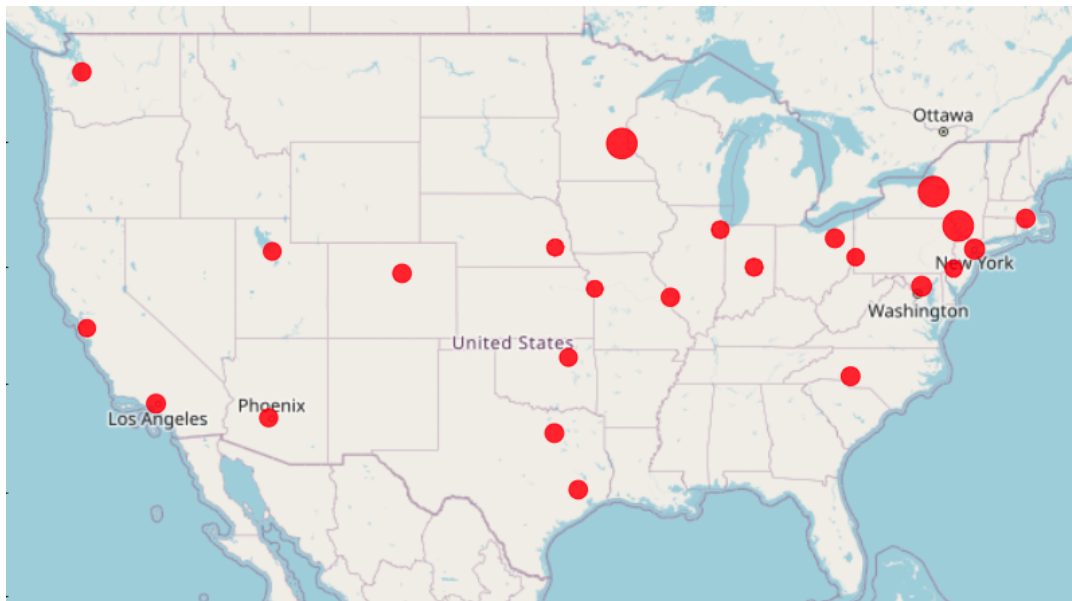
7. Wisconsin and Arizona were the overall most expensive venues for events, with the Southwest having the highest number of events.

8. Out of the ~100 events analyzed, the cheapest tickets were $15 USD for Monster Jam (a Motorsports Racing event in Minnesota) and the most expensive tickets were $4,999 USD for Phoenix Suns vs. Oklahoma City Thunder (an NBA game in Arizona)

9. The Mean Average Price for Tickets was about $500.93 USD

	Min Price $	Max Price $	Average Price $
count	89.000000	89.000000	89.000000
mean	75.641685	850.577865	500.930618
std	51.724047	1054.300660	547.789023
min	15.000000	110.000000	80.000000
25%	40.000000	255.950000	188.925000
50%	55.000000	399.500000	249.925000
75%	109.000000	925.950000	530.000000
max	299.950000	4999.000000	2648.500000

Future Plans

For this project, we used a free API key from Ticketmaster, which was useful, but certainly did not provide unlimited requests and types of data. If we were to expand on this project in the future, we might consider spending money on a more premium API key from ticketmaster and collecting much more data. We would also look at secondary markets for scalped ticket data, and consider API keys for sites like StubHub. Lastly, we would also consider web-scraping as a last resort to gather any data not obtained through Ticketmaster and StubHub. For example, we could look at places like eBay or Craigslist.