



# *ticketmaster*

**Alvaro Aquino, Emmanuel Garcia, Frantzy Francols,  
Jonathan Pokorny, & Mickey Young**

---

# Agenda

- **Hypothesis**
- **Tactics**
- **Challenges & Obstacles**
- **Analysis & Findings**



# Hypothesis

**What costs more money: Music or Sports?**

**Are events more expensive in certain states?**

Using Ticketmaster API, we look to find the answer.



vs.



# Tools used:



Ticketmaster API



Libraries: Pandas, NumPy, Matplotlib, Requests, JSON, HVPlot.pandas



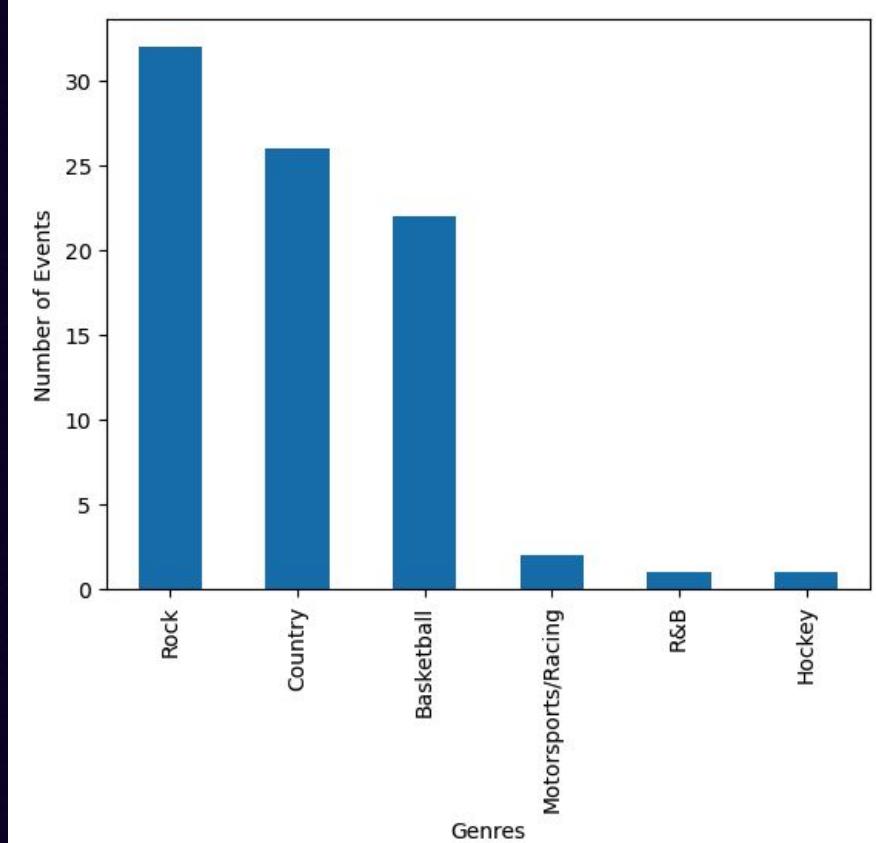
Export/Read CSV's



# SETUP

## OUR PROCESS:

- We analyzed ~100 events
- Both sports & music
- Categorized our data
- Created charts
- Extracted insights



# Tactics & Methods

## TICKETMASTER API



Using JSON, requests, lists, loops, and DataFrames to collect data from the Ticketmaster API on upcoming events in the US

```
import pandas as pd
import requests
import json

sample = "https://app.ticketmaster.com/discovery/v2/events.json?countryCode=US&size=100&apikey=kEYjjJ2wWX9IwsLu8zG4E"
response = requests.get(sample).json()

event_list = []
dates = []
urls = []
price_min = []
prices_max = []
types = []
genres = []
sub_genres = []
states = []
venues = []
lats = []
longs = []
cities = []
avg_prices = []

counter = 0
for event in response["_embedded"]["events"]:
    try:
        print(event["name"])
        print("State:", event["_embedded"]["venues"][0]["state"]["name"])
        print("Venue:", event["_embedded"]["venues"][0][0]["name"])
        print("Event Type:", event["classification"][0]["segment"][0]["name"])
        print("Genre:", event["classification"][0][0]["genre"][0]["name"])
        print("Date:", event["dates"]["start"]["localdate"])
        print("URL:", event["url"])
        print("Minimum Ticket Price:", event["priceRanges"][0][0]["min"])
        print("Maximum Ticket Price:", event["priceRanges"][0][0]["max"])
        print("-----")
        event_list.append(event["name"])
        types.append(event["classification"][0][0]["segment"][0]["name"])
        genres.append(event["classification"][0][0]["genre"][0]["name"])
        sub_genres.append(event["classification"][0][0][0]["subGenre"][0]["name"])
        dates.append(event["dates"]["start"]["localDate"])
        urls.append(event["url"])
        states.append(event["_embedded"]["venues"][0][0]["state"]["name"])
        cities.append(event["_embedded"]["venues"][0][0][0]["city"]["name"])
        lats.append(event["_embedded"]["venues"][0][0]["location"]["latitude"])
        longs.append(event["_embedded"]["venues"][0][0][0]["location"]["longitude"])
    except:
        pass
    counter += 1
    if counter == 10:
        break
```

	Event Name	State	City	Latitude	Longitude	Venues	Event Type	Genre	Sub-Genre	Dates
0	Milwaukee Bucks v. Phoenix Suns	Wisconsin	Milwaukee	43.04228	-87.916896	Fiserv Forum	Sports	Basketball	NBA	2023-02-26
1	Atlanta Hawks vs. Phoenix Suns	Georgia	Atlanta	33.757796	-84.394569	State Farm Arena	Sports	Basketball	NBA	2023-02-09
2	Phoenix Suns vs. Sacramento Kings	Arizona	Phoenix	33.445899	-112.071313	Footprint Center	Sports	Basketball	NBA	2023-02-14
3	Indiana Pacers vs. Phoenix Suns	Indiana	Indianapolis	39.764064	-86.155507	Gainbridge Fieldhouse	Sports	Basketball	NBA	2023-02-10
4	Golden State Warriors vs. Phoenix Suns	California	San Francisco	37.76797	-122.387464	Chase Center	Sports	Basketball	NBA	2023-03-13
...	...	...	...	...	...	...	...	...	...	...
89	Shania Twain: Queen Of Me Tour	Ohio	Columbus	40.0080079	-83.0249081	Schottenstein Center	Music	Country	Country	2023-10-27
90	Shania Twain: Queen Of Me Tour	Texas	San Antonio	29.4330386	-98.4337987	AT&T Center	Music	Country	Country	2023-10-12
91	Monster Jam	California	Anaheim	33.800308	-117.88278	Angel Stadium of Anaheim	Sports	Motorsports/Racing	Motorsports/Racing	2023-02-18
92	Eagles	California	Lincoln	38.84026	-121.31368	The Venue at Thunder Valley Casino Resort	Music	Rock	Pop	2023-02-17
93	Shania Twain: Queen Of Me Tour	Tennessee	Knoxville	35.9517627	-83.925366	Thompson-Boling Arena	Music	Country	Country	2023-10-16



# Challenges and Obstacles

JUST A FEW....



cURL vs. JSON

---



Service Disruptions

---



Inconsistent Data

---



Free Events API

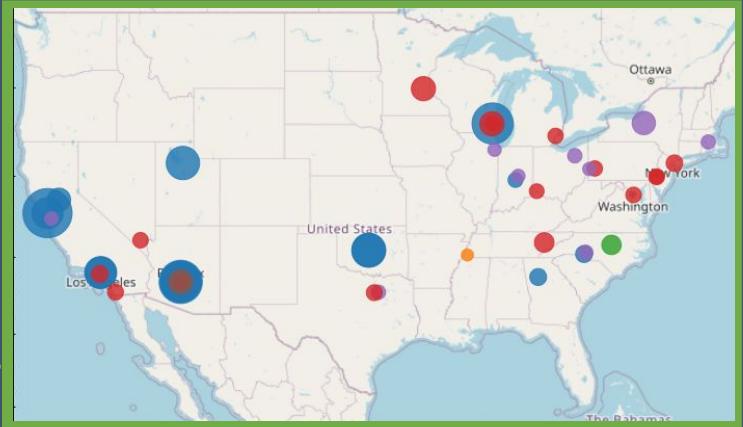
---

# HVPLOTS | Event Locations

## VISUALIZING OUR HYPOTHESIS

Using hvplot.pandas to visually show Average Ticket Prices for Events and determine which are the most expensive to attend

```
# %%capture --no-display  
  
map_plot = events_df.hvplot.points(  
    "Longitude",  
    "Latitude",  
    geo = True,  
    size = "Average Price $",  
    scale = 1,  
    color = "Genre",  
    alpha = 0.8,  
    tiles = "OSM",  
    frame_width = 700,  
    frame_height = 400,  
    hover_cols = ["Event Name", "Venues"]  
)  
  
map_plot
```



### Findings:

Most Expensive Event to Attend:  
???????

Most Expensive Genre of Music:  
???????

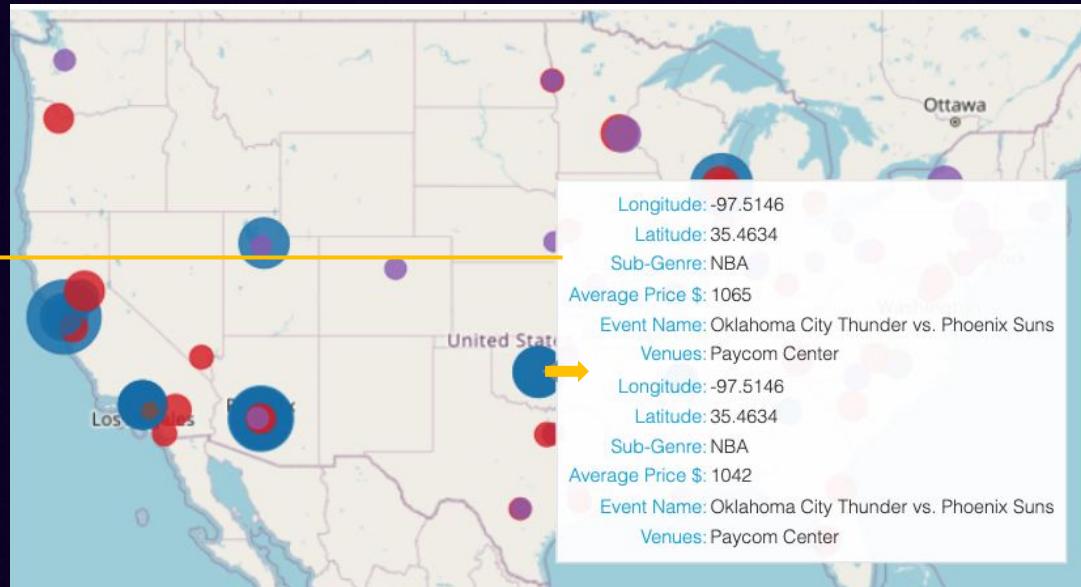
# HVPLOTS | Features

INSIGHTS FROM THE INTERACTIVE



## Hover Column

EVENT INFORMATION



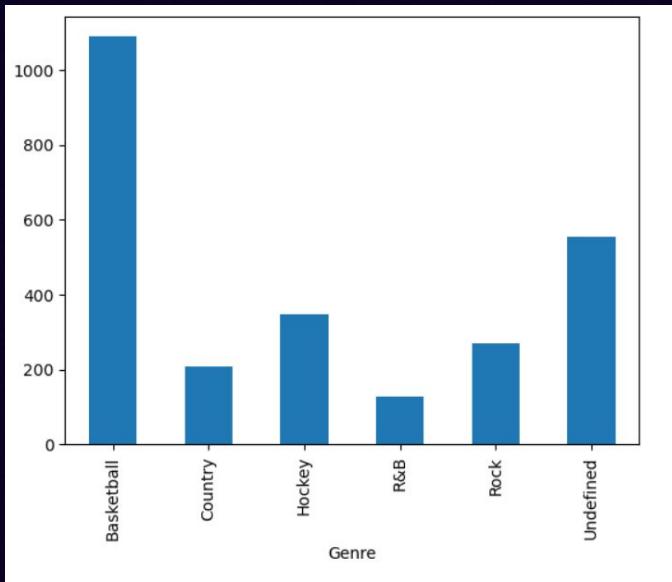
ticketmaster



# Removing “undefined” & re-ordering columns

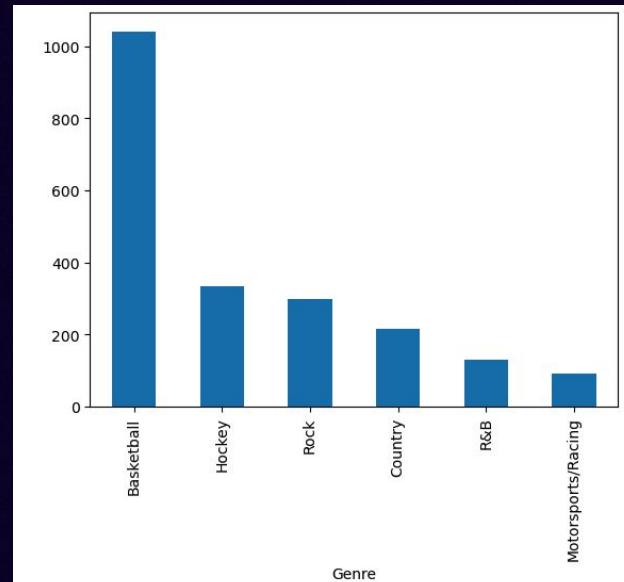
BEFORE

Average Ticket Price (\$)



AFTER

Average Ticket Price (\$)



DataFrame

EXCLUDING “UNDEFINED”

```
# Filtering out undefined
events_df = raw_df[raw_df["Sub-Genre"] != "Undefined"]
```

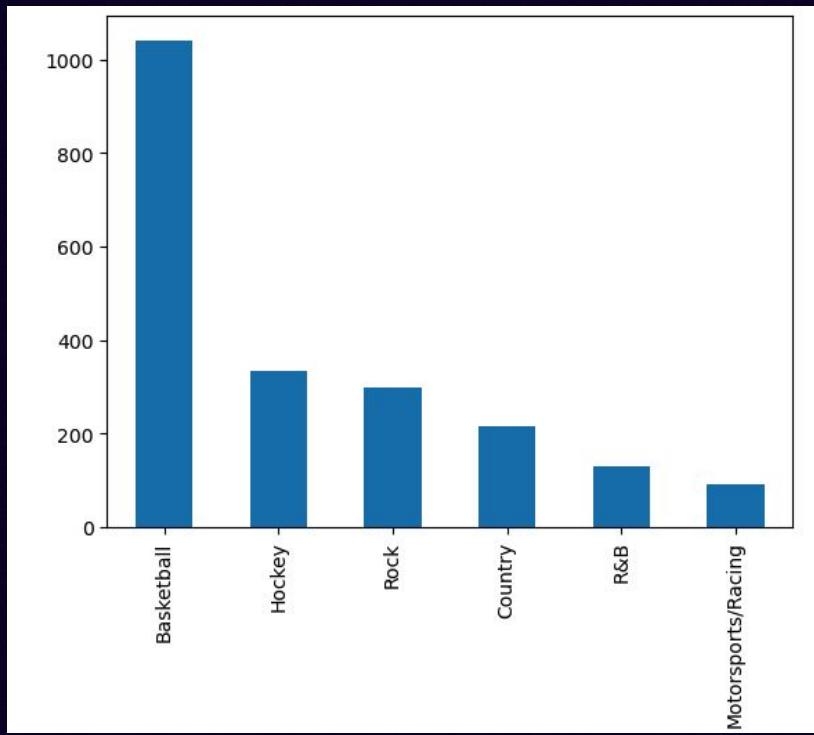
Bar Plot

RE-ORDERING COLUMNS

```
avgprice_genre_df = events_df.groupby(["Genre"])["Average Price $"].mean()
new_avgprice_genre_df = avgprice_genre_df[["Basketball", "Hockey", "Rock", "Country", "R&B", "Motorsports/Racing"]]
new_avgprice_genre_df.plot(kind='bar');
```

# Average Price of All Events

Average Ticket Price (\$)



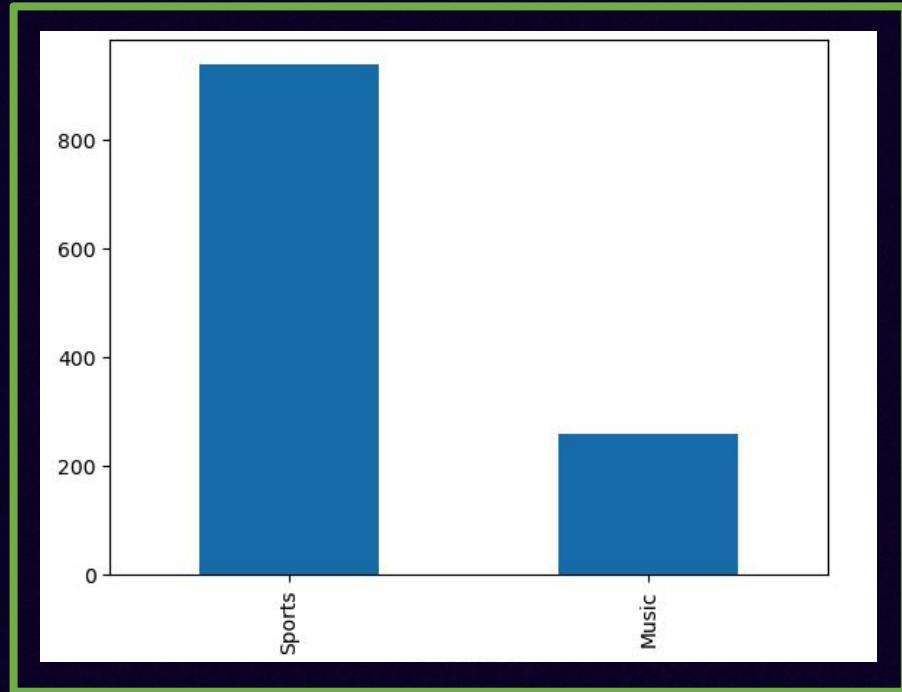
Sub-Genre

# Music vs. Sports

```
1 avgprice_type_df.plot(kind='bar');
```

The average price of sports tickets was nearly 3x as much as music tickets

Average Ticket Price (\$)



Event Type

# Music vs. Sports

```
# %%capture --no-display

map_plot = events_df.hvplot.points(
    "Longitude",
    "Latitude",
    geo = True,
    size = "Average Price $",
    scale = 1,
    color = "Event Type",
    alpha = 0.8,
    tiles = "OSM",
    frame_width = 700,
    frame_height = 400,
    hover_cols = ["Event Name", "Venues"]
)

map_plot
```

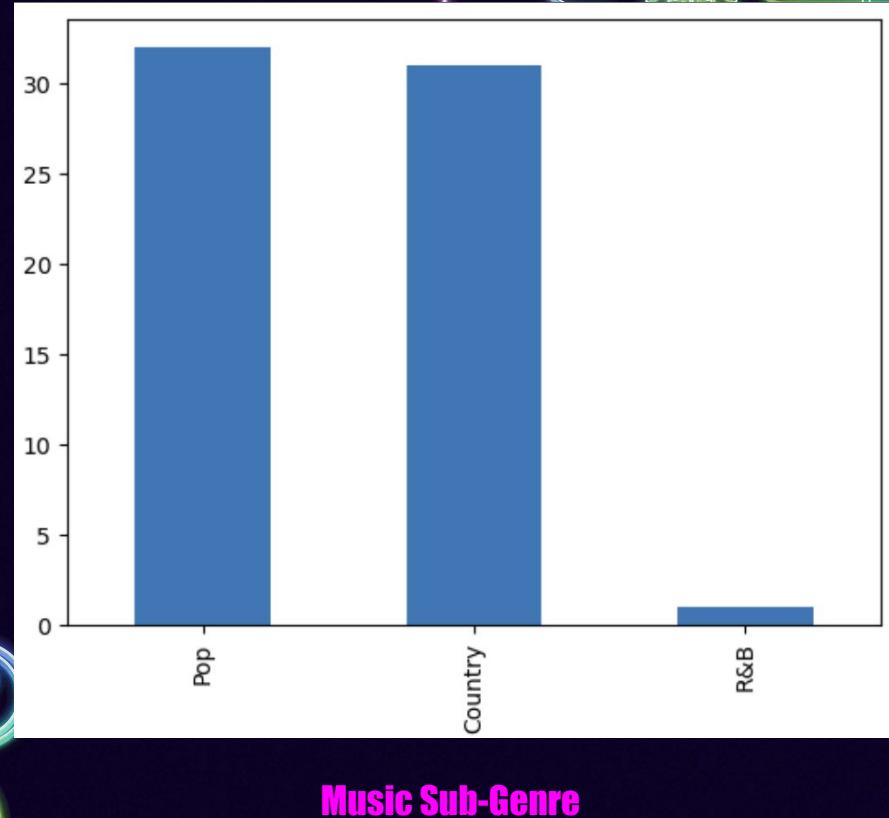


- Sports
- Music

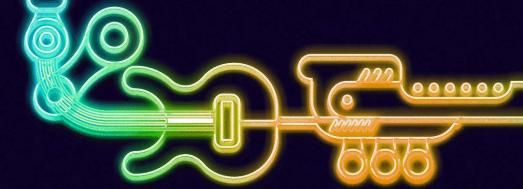
ticketmaster

# Number of Events Analyzed by Genre

```
music_counts = music_df[ "Sub-Genre" ].value_counts()  
music_counts.plot(kind="bar")  
plt.xlabel("Music Genre")  
plt.xticks(rotation=90)  
plt.ylabel("Number of Events")  
plt.show()
```



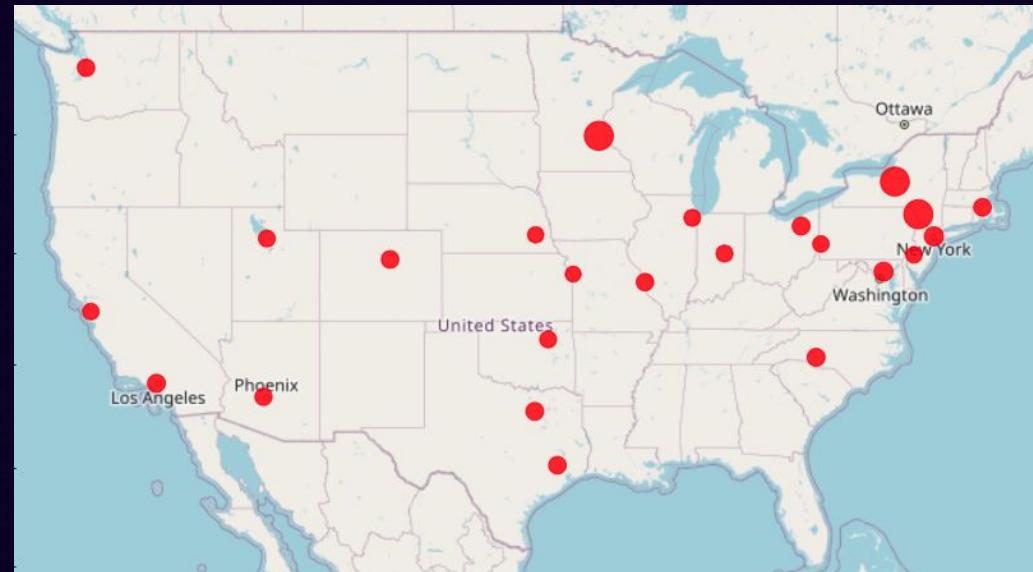
# Is country music more popular in the northeast or South?



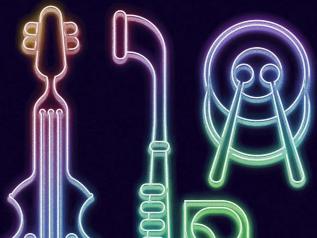
Filtered number of music events by Country and created a hvPlot to show events by location



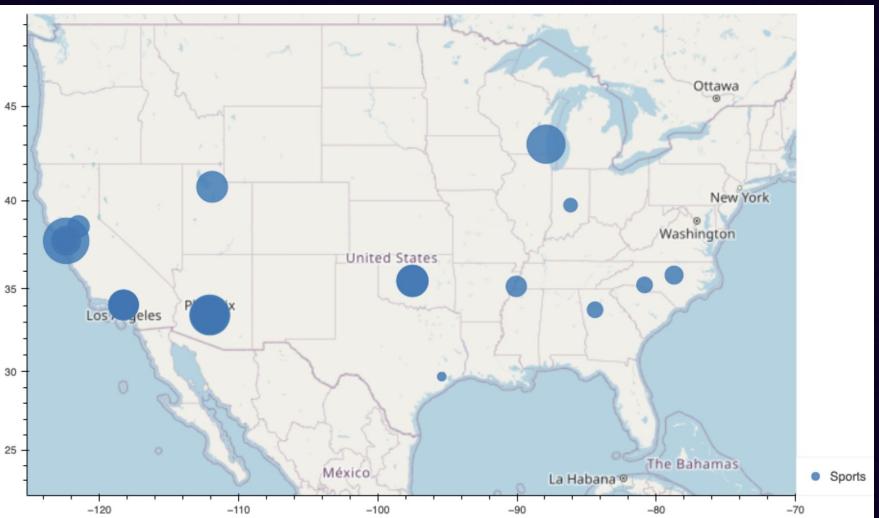
Data show that it's actually most popular in New York



```
country_filter = events_df[  
    (events_df["Event Type"] == "Music") &  
    (events_df["Sub-Genre"] != "Undefined") &  
    (events_df["Genre"] == "Country")]  
country_df = pd.DataFrame(country_filter)  
country_df
```



# Highest Average Ticket Price for Sports By Location



## DataFrame

### FILTERED FOR SPORTS

```
sports_filter = events_df[events_df["Event Type"] == "Sports"]
sports_df = pd.DataFrame(sports_filter)
sports_df
```

### SPORTS MAP PLOT

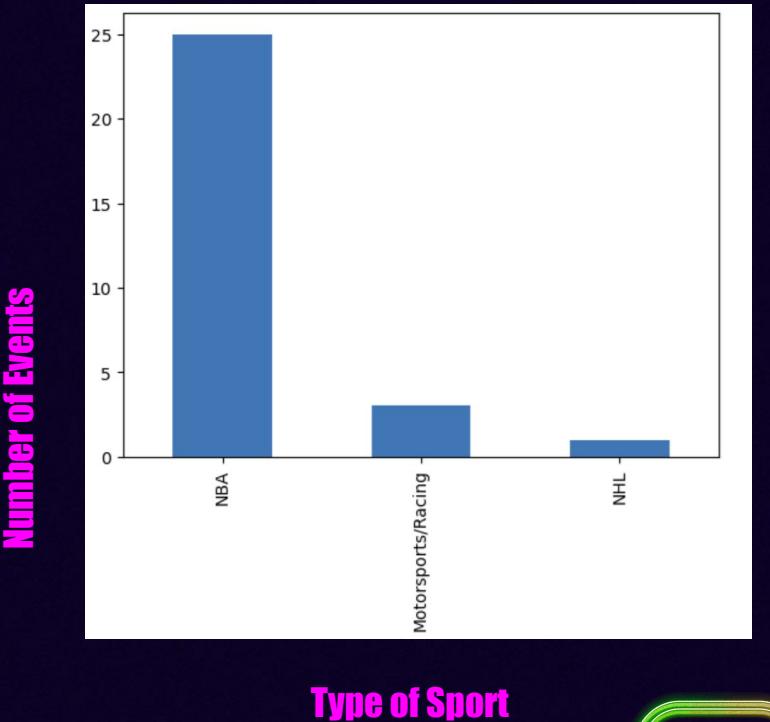
```
# %%capture --no-display

sports_plot = sports_df.hvplot.points(
    "Longitude",
    "Latitude",
    geo = True,
    size = "Average Price $",
    scale = 1,
    color = "Event Type",
    alpha = 0.8,
    tiles = "OSM",
    frame_width = 800,
    frame_height = 500,
    hover_cols = ["Event Name", "Venues", "Genre"]
)

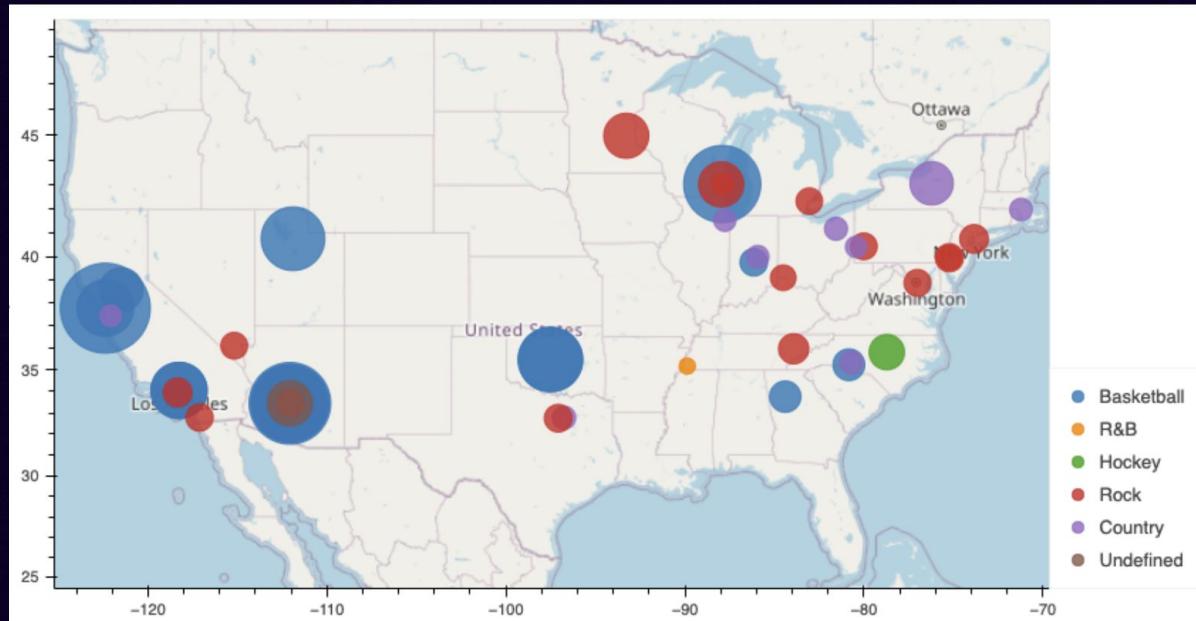
sports_plot
```

# Number of Events by Sport

- Basketball currently has the highest number of events
- Hockey regular season runs from October to April
- Basketball regular season runs from October to April
- Motorsports/Racing typically starts in February and ends in March



## **Highest Average Price Across the Country by Sport and Music Genre**



**Basketball had the overall highest average price per ticket in sports**



**Rock had the overall highest average price per ticket in music**



# Conclusion



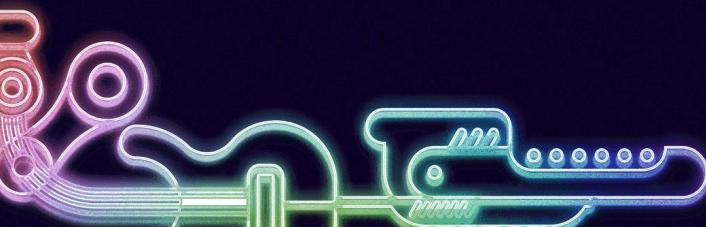
**We didn't expect to see Rock ticket prices being the most expensive**



**We thought Country music would be more popular in the South,  
but it was actually more popular in New York**



**Wisconsin and Arizona were the overall most expensive venues for  
events, with the Southwest having the highest number of events**





# Future Plans

If we needed to revisit the project, and needed more data, we might consider the following:



Getting a more premium API key from TicketMaster



Use API's from other websites, such as StubHub



Web-scrape other sites that do not have API's

# Thanks!

**Alvaro Aquino, Emmanuel Garcia, Frantzy Francols,  
Jonathan Pokorny, & Mickey Young**

**ticketmaster**