Coursera capstone Project

National Football League

Stadiums. Places to enjoy and business

Introduction

National Football League (NFL) is a professional American football league

The NFL was formed in 1920 as the American Professional Football Association (APFA) before renaming itself the National Football League for the 1922 season.

A total of 16.67 million fans attended NFL games during the 2019 regular season and each team welcomed an average total of more than 521 thousand spectators to their home games across the season. This meant that the league average attendance per regular season game stood at 66,151 for the 2019 season.

Business problem

Stadiums are a place that gathers millions of fans. Therefore, we have the main places for analysis and conclusion of two main problems

1. For fans

Stadiums segmenting will help many fans, traveling with their favorite team, to understand what type of stadium will be the next or in the nearest future. It will be easy to understand: it is easy to find a suitable place to stay or not, will be a lot of venues around or not etc.

2. For Business

We try to segment the stadium depending on the places that are nearby to identify areas that can be developed in a certain way (need more space for food, hotels or some others)

Data describe

We used data from network:

- Mainly from Wikipedia

- Home page of NFL (https://www.nfl.com/)

- Also we need a Forsquare API to make our main work

Data describe

We merged data from different resourses to receive final data to deal with.

Final data with stadiums looks like:

|	Division	Club	City	Stadium	Capacity	Coordinates	First season	Head coach	Latitude	Longitude
0	East	Buffalo Bills	Orchard Park, New York	New Era Field	71608	42.774°N 78.787°W	1960 (AFL), 1970 (NFL)	Sean McDermott	42.774	-78.787
1	East	Miami Dolphins	Miami Gardens, Florida	Hard Rock Stadium	64767	25.958°N 80.239°W	1966 (AFL), 1970 (NFL)	Brian Flores	25.958	-80.239
2	East	New England Patriots	Foxborough, Massachusetts	Gillette Stadium	65878	42.091°N 71.264°W	1960 (AFL), 1970 (NFL)	Bill Belichick	42.091	-71.264
3	East	New York Jets	East Rutherford, New Jersey	MetLife Stadium[C]	82500	40.814°N 74.074°W	1960 (AFL), 1970 (NFL)	Adam Gase	40.814	-74.074
4	North	Baltimore Ravens	Baltimore, Maryland	M&T Bank Stadium	71008	39.278°N 76.623°W	1996[D]	John Harbaugh	39.278	-76.623

Data describe

After, we merged stadiums info with data from Forsquare.

For, example, we used next items from Forsquare:

|	lat	lng	categories
0	42.773684383495166	-78.78722950890148	Football Stadium
1	42.776580316228205	-78.78879526962254	Other Great Outdoors
2	42.774717285786345	-78.78891698849554	Souvenir Shop
3	42.77079418868539	-78.78661834890876	Athletics & Sports
4	42.777837318401126	-78.78893909627917	Bar
5	42.77712200556276	-78.79191270072297	Hot Dog Joint
6	42.775482734543196	-78.79344287011612	Restaurant
7	42.77773361822128	-78.78984122129458	Bar
8	42.777868	-78.7911722	Coffee Shop
9	42.767301294177294	-78.79133099854224	Bar
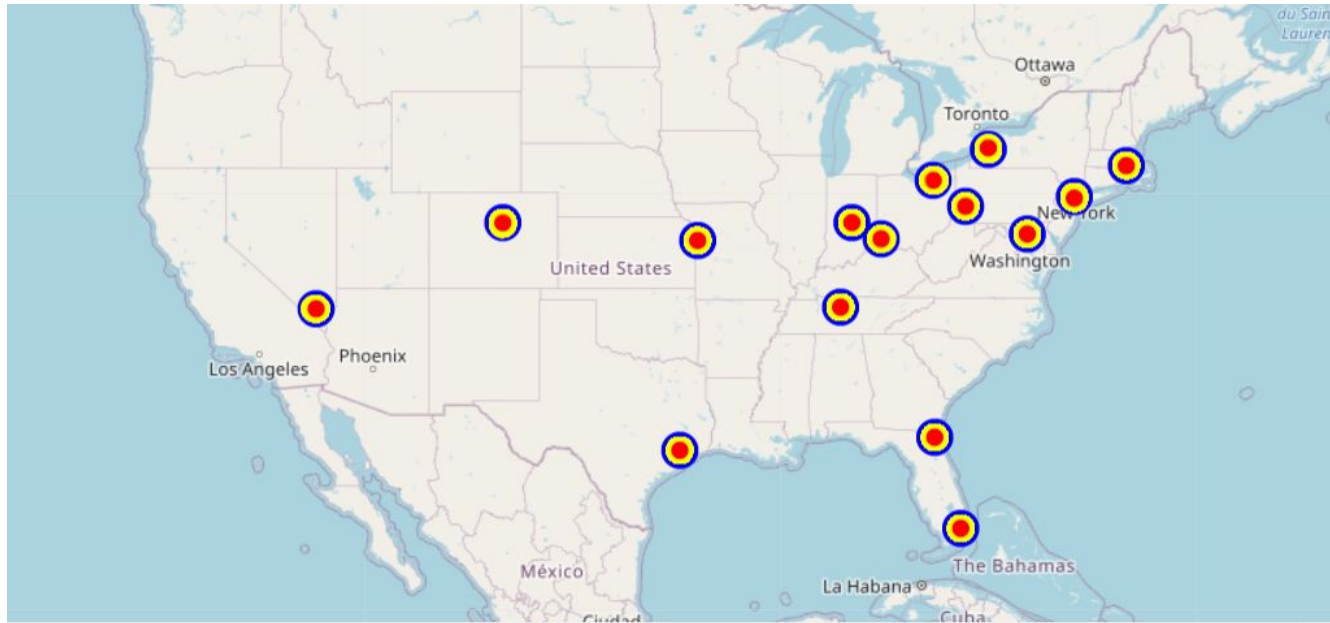
Methodology section

We can use map to better understand our data (for exemple for next stadiums):

Methodology section

We can use map to better understand our data (for exemple for next stadiums):

Indianapolis

Cleveland
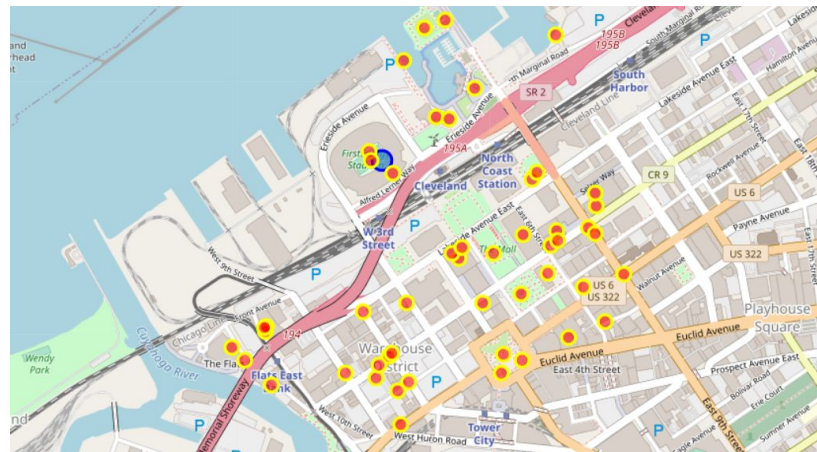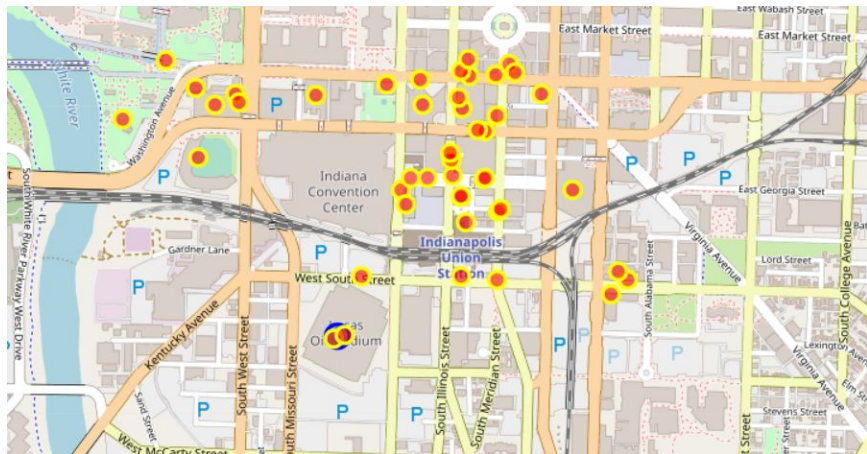
Methodology section

Finally, we used k-means to build our classification model.

For each category we interested (hotels, restaurants, bars, clubs, cafes), we calculate the score (score = number of each venue/ capacity of the stadium).

Based on the scores, we conducted a segmentation.

As a result we have 3 groups:
0 - stadium with low level of quantity of venues around
1 - stadium with medium level of quantity of venues around
2 - stadium with high level of quantity of venues around

Result section

Our resulting data with all stadiums.

6 stadiums with 0 level
18 stadiums with 1 level
6 stadiums with 2 level

Stadium	Bar	Club	Hotel	Restaurant	Coffee Shop	Labels
AT&T Stadium	0.000025	0.000000	0.000000	0.000025	0.000000	1
Allegiant Stadium	0.000046	0.000015	0.000031	0.000154	0.000015	2
Arrowhead Stadium	0.000026	0.000000	0.000000	0.000000	0.000000	1
Bank of America Stadium	0.000000	0.000000	0.000000	0.000026	0.000000	1
CenturyLink Field	0.000029	0.000000	0.000029	0.000217	0.000014	2
Empower Field at Mile High	0.000013	0.000000	0.000000	0.000026	0.000000	1
FedExField	0.000012	0.000000	0.000000	0.000037	0.000000	1
FirstEnergy Stadium	0.000000	0.000000	0.000029	0.000015	0.000000	1
Ford Field	0.000015	0.000000	0.000015	0.000108	0.000015	2
Gillette Stadium	0.000046	0.000000	0.000015	0.000121	0.000000	0
Hard Rock Stadium	0.000015	0.000015	0.000000	0.000015	0.000000	1
Heinz Field	0.000044	0.000000	0.000000	0.000044	0.000015	0
Lambeau Field	0.000012	0.000000	0.000012	0.000012	0.000000	1
Levi's Stadium	0.000015	0.000000	0.000000	0.000015	0.000015	1
Lincoln Financial Field	0.000043	0.000000	0.000000	0.000058	0.000014	0
Lucas Oil Stadium	0.000015	0.000015	0.000015	0.000030	0.000015	2
M&T Bank Stadium	0.000014	0.000000	0.000000	0.000000	0.000000	1
Mercedes-Benz Stadium	0.000014	0.000000	0.000000	0.000056	0.000000	1
Mercedes-Benz Superdome	0.000000	0.000000	0.000027	0.000027	0.000014	2
MetLife Stadium[C]	0.000048	0.000000	0.000000	0.000024	0.000000	0
NRG Stadium	0.000042	0.000000	0.000000	0.000069	0.000014	0
New Era Field	0.000014	0.000000	0.000000	0.000014	0.000000	1
Nissan Stadium	0.000014	0.000000	0.000014	0.000000	0.000000	1
Paul Brown Stadium	0.000031	0.000000	0.000000	0.000015	0.000015	0
Raymond James Stadium	0.000000	0.000000	0.000000	0.000030	0.000000	1
SoFi Stadium[F]	0.000000	0.000000	0.000000	0.000000	0.000000	1
Soldier Field	0.000016	0.000000	0.000000	0.000016	0.000000	1
State Farm Stadium	0.000000	0.000000	0.000000	0.000016	0.000016	1
TIAA Bank Field[E]	0.000015	0.000000	0.000000	0.000000	0.000000	1
U.S. Bank Stadium	0.000030	0.000000	0.000030	0.000090	0.000015	2

Conclusion section

We explored that 24 stadiums with 0 or 1 level and only 6 stadiums with 2 level.

So,
For fans:
most stadiums have very good conditions for fans.
this is not only the about the stadium but also the venues around.
everyone can find appropriate place before and after the game

For business:
a significant number of fans, as well as not full coverage of venues, provides an opportunity to develop specific areas and business activities during the games.