Finding the Next NBA City

Introduction

Background

- Professional sports is a multi-billion dollar industry
  - Major leagues include NFL, NBA, NHL and MLB
- Continually looking to expand to new markets
- The NBA is next likely candidate for expansion
  - Has seen revenue and ratings increase in recent years
  - Only 30 NBA teams vs 32 for NFL and NHL

Problem Statement

- Help the NBA find a new city for expansion

Collecting City Data

- City data collected via web scraping and the FourSquare city data API

Data	Source
US City population Data	https://en.m.wikipedia.org/wiki/List_of_United_States_cities_by_population
US TV Market Data	https://en.wikipedia.org/wiki/List_of_United_States_television_markets
US Per-Capita Income Data	https://en.wikipedia.org/wiki/List_of_United_States_metropolitan_areas_by_per_capita_income
List of current NBA cities	https://www.basketball-reference.com/teams/
City Venue/Interests Data	FourSquare API

Clean City Data

- Wikipedia city data cleaned and merged
  - Converted numerical strings and objects to floats
  - Extracted city TV Market and per-capita income from regional data
  - Merged tables
  - Interpolated for missing values based on population
- Added NBA column to distinguish if city has team
  - 1 = YES, 0 = NO
- Carried forward relevant features
  - City, State, Population, TV Market Size, Per-Capita-Income, NBA City

Add/Calculate Location Information

- Python GeoPy module used to...

  - Get latitude and longitude for each city

  - Calculate distance from each city to nearest NBA city

- Filter down data set by...

  - Removing non-NBA cities within 150km of current NBA cities

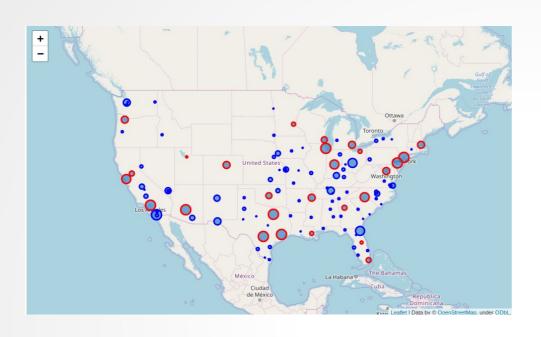 - Taking 120 most populous cities from remaining set

    – 27 NBA cities

    – 93 non-NBA cities

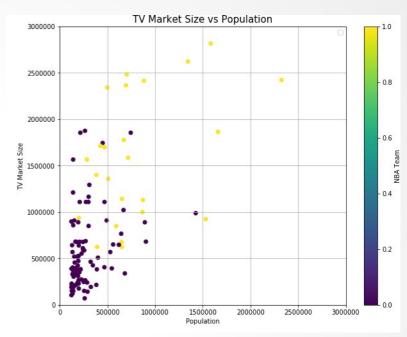| Rank	City	State	Population	Density	PCI	TV	NBA	Long	Lat	MinDistance
1	New York	New York	8398748.0	10933.0	24581.0	7100300.0	1	-74.006015	40.712728	0.000000
2	Los Angeles	California	3990456.0	3276.0	21170.0	5276600.0	1	-118.242767	34.053691	0.000000
3	Chicago	Illinois	2705994.0	4600.0	21435.5	3251370.0	1	-87.624421	41.875562	0.000000
4	Houston	Texas	2325502.0	1395.0	21701.0	2423360.0	1	-95.367697	29.758938	0.000000
5	Phoenix	Arizona	1660272.0	1200.0	21907.0	1864420.0	1	-112.077346	33.448587	0.000000
...	...	...	...	...	...	...	...	...	...	...
222	Fargo	North Dakota	124844.0	950.0	19910.0	228740.0	0	-96.789821	46.877229	345.669399
224	Columbia	Missouri	123180.0	717.0	20902.0	389590.0	0	-92.333737	38.951883	468.274367
225	Abilene	Texas	122999.0	442.0	17176.0	104440.0	0	-99.733301	32.446674	278.020276
226	Wilmington	North Carolina	122607.0	880.0	22100.0	190390.0	0	-77.944711	34.225728	287.757839
227	Hartford	Connecticut	122587.0	2735.0	34310.0	897870.0	0	-72.690855	41.764582	150.452019

Adding FourSquare Data

- "Explore" function used from FourSquare API
- Function called 9 times for each city
  - Once at city's Long/Lat
  - 8 times at permutations of Long/Lat +/-0.2
  - Gave more venue data (limited to 100 venues per call)
- One-hot encoded venue categories
- Venue categories averaged based on city
- Venue information appended to city data set
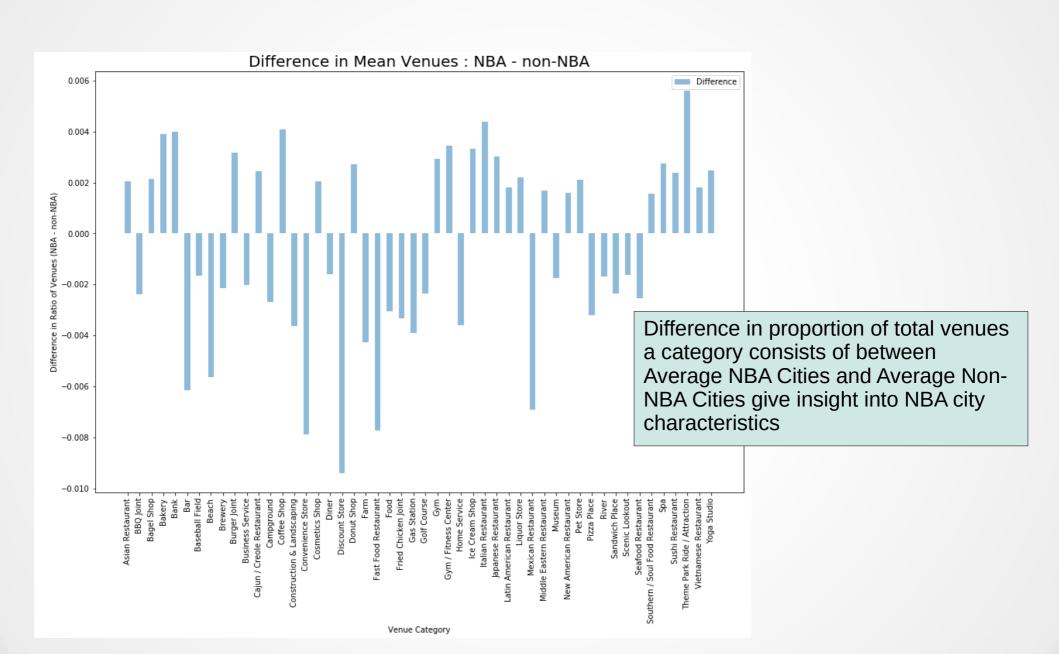
City Data Exploratory Analysis

- NBA cities (RED)

- NBA candidates (BLUE)

- Size of marker proportional to size of city

- Some candidate cities fall in similar population and TV market range as NBA cities

- Population alone isn't predictor of NBA city success
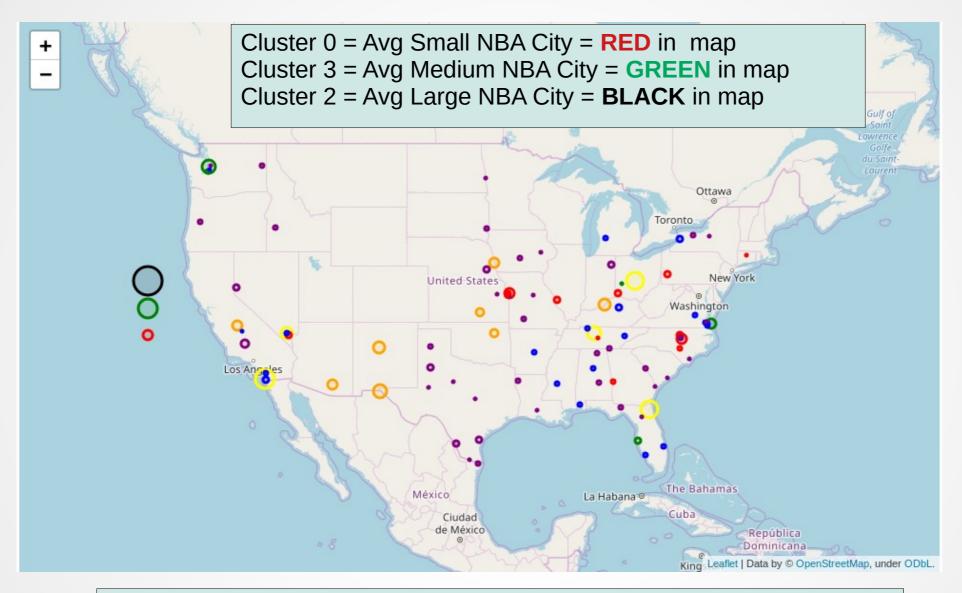
FourSquare Exploratory Analysis

Difference in Mean Venues : NBA - non-NBA

Difference in proportion of total venues a category consists of between Average NBA Cities and Average Non-NBA Cities give insight into NBA city characteristics

Modeling – KMeans Cluster

Cluster Number	Number of Cities	NBA City?
0	14	Avg. Small
1	20	--
2	1	Avg. Large
3	6	Avg. Medium
4	5	--
5	9	--
6	41	--

- Data prepared for modeling
  - Dropped NBA cities
  - Created 3 average NBA "cities"
    - Large = 10 largest cities
    - Medium = 10 middle cities
    - Small = 7 smallest cities

- KMeans Clustering used to group data
  - # of Clusters = 7
  - Initializer = k-means++

Non-NBA cities clustered against 3 categories of an "Average" NBA city

Results

Cluster 0 = Avg Small NBA City = **RED** in map
Cluster 3 = Avg Medium NBA City = **GREEN** in map
Cluster 2 = Avg Large NBA City = **BLACK** in map

Cluster 3 with 5 non-NBA cities best candidate for future NBA city based on data in study

Discussion

- Closer inspection of 5 cities in cluster 3 reveals Seattle, WA recommended selection

  - Population and TV market largest among candidates

  - Tacoma, WA also on this list, potentially supports a Seattle franchise

Cluster 3 – Detailed Results

	City	State	Population	TV	PCI	Lat	Long
3	Seattle	Washington	744,955	1,854,810	39,322	48	-122
14	Virginia Beach	Virginia	450,189	1,744,733	20,328	37	-76
30	St. Petersburg	Florida	265,098	1,875,420	21,784	28	-83
42	Tacoma	Washington	216,279	1,854,810	39,322	47	-122
78	Dayton	Ohio	140,640	1,565,890	21,598	40	-84
94	Med	Med	648,833	1,428,650	28,424	40	-95

Study recommends Seattle, WA as next NBA city location

Closing

- Study looked at candidate NBA cities based on city and locale data

- Recommendation of study is **Seattle, WA** as next location

- Additional information would need to be explored before final decision, such as…

  - Is the city willing to fund or build a stadium?

  - Is there an area within the city to build a new stadium or otherwise house the team?

  - Is there ample public transit and/or parking nearby to get people to the stadium?