Med Frezghi and Ajaeya Masti
Professor Barbara Ericson
SI 206 Final Project
Group Name: Bikini Bottom Boys
13 December 2021

Project Goals

Our original goals were to utilize the balldontlie API to retrieve NBA teams' total points and wins from the 2019 season. We also used the national census' population estimate API to get the population per city. Our intention was to compare the two databases and see if the number of fans in a city affects how well a professional basketball team performs in that city. To do this, we would calculate the number of fans per NBA team and the average number of points per amount of fans in the city.

Achieved Goals

We were able to successfully use the balldontlie API to pull game data from the 2018 season. Although we weren't able to utilize the population estimate API, we were able to parse a website's team population estimate data for their metro population. It wasn't exactly our original goal, but it did allow us to compare the points each team scored in 2018 to their city's size and see if there is any correlation between the two.

Problems Faced

We weren't able to use the population estimates API from the census because we weren't able to retrieve the API key. The NBA data from the balldontlie API was also difficult to traverse, since there was a lot of specification necessary in order to pull the correct game scores from the right year we wanted.

Calculation Files

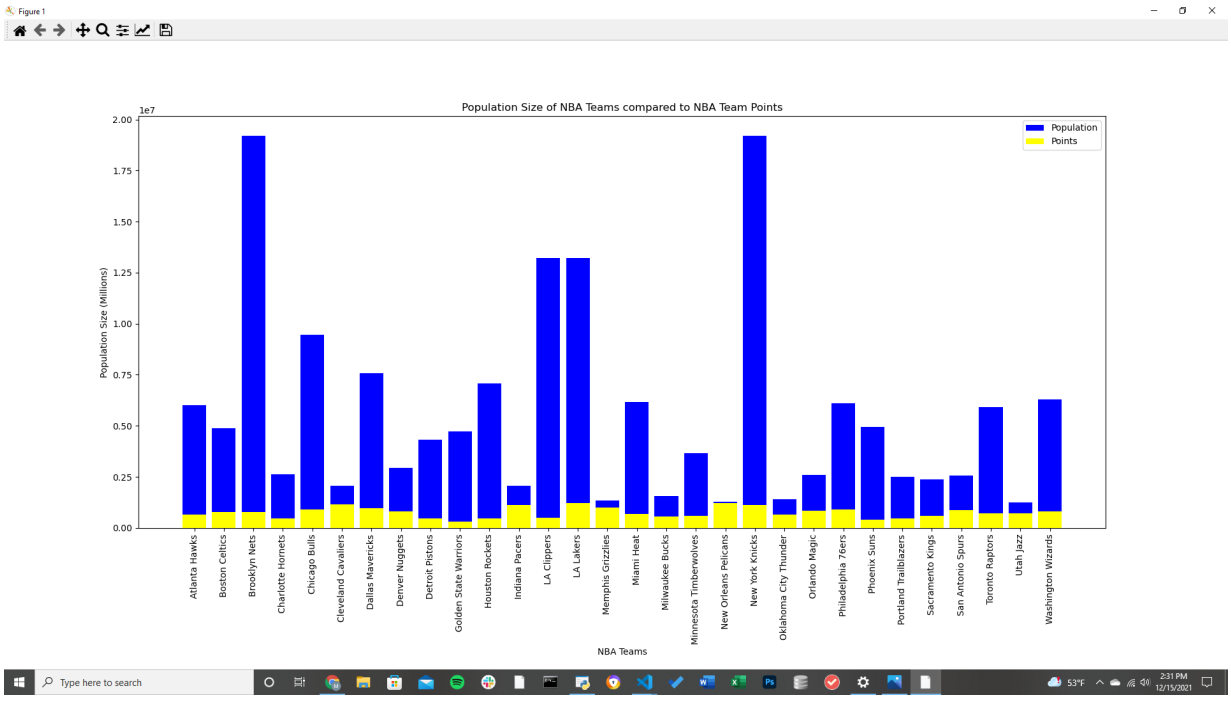
Mean points per team in regards to population size.
This data is stored in a dictionary in the find_mean_difference function.

{'New York Knicks': 19215424.266666666, 'Brooklyn Nets': 19215424.266666666, 'LA Lakers': 13214041.266666668, 'LA Clippers': 13214041.266666668, 'Chicago Bulls': 9457781.266666668, 'Philadelphia 76ers': 6101676.266666667, 'Dallas Mavericks': 7572378.266666667, 'Toronto Raptors': 5927282.266666667, 'Golden State Warriors': 4731045.266666667, 'Atlanta Hawks': 6019606.266666667, 'Houston Rockets': 7065383.266666667, 'Washington Wizards': 6279729.266666667, 'Boston Celtics': 4872261.266666667, 'Phoenix
Suns': 4947445.266666667, 'Minnesota Timberwolves': 3654150.2666666666, 'Detroit Pistons': 4318871.266666667, 'Denver Nuggets': 2931657.2666666666, 'Orlando Magic': 2607389.2666666666, 'Miami Heat': 6165730.266666667, 'Cleveland Cavaliers': 2047691.2666666666, 'Sacramento Kings': 2362972.2666666666, 'Portland Trailblazers': 2491654.2666666666, 'Charlotte Hornets': 2636125.2666666666, 'Indiana Pacers': 2073779.2666666666, 'Utah Jazz': 1231938.2666666666, 'San Antonio Spurs': 2550202.2666666666, 'Milwaukee Bucks': 1574421.2666666666, 'Oklahoma City Thunder': 1408192.2666666666, 'New Orleans Pelicans': 1269772.2666666666, 'Memphis Grizzlies': 1345287.2666666666}

```
abs(points_mean - pop_data[key])
```

The absolute value of the mean of total points scored in the 2018 NBA season - the population size of each respective team city.

Visualizations



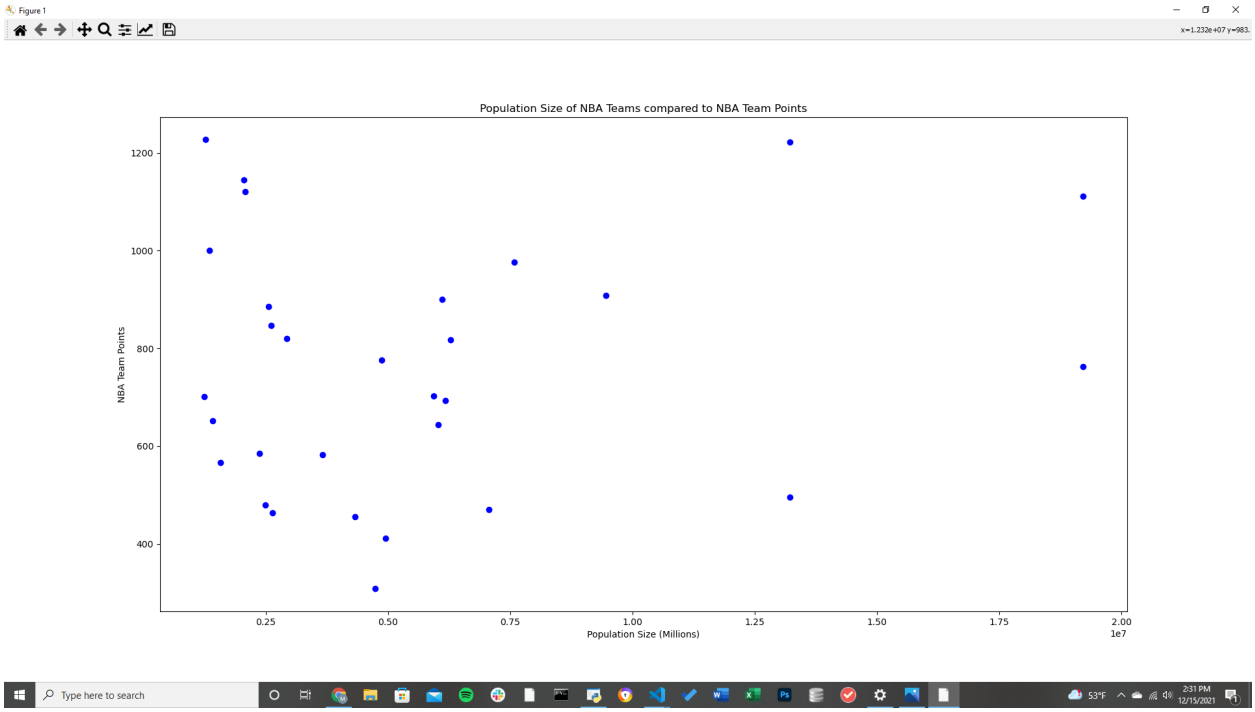


Figure 1
x=1.232e+07 y=983.
Population Size of NBA Teams compared to NBA Team Points
NBA Team Points
Population Size (Millions)
1e7
1200
1000
800
600
400
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Type here to search
53°F
2:31 PM
12/15/2021

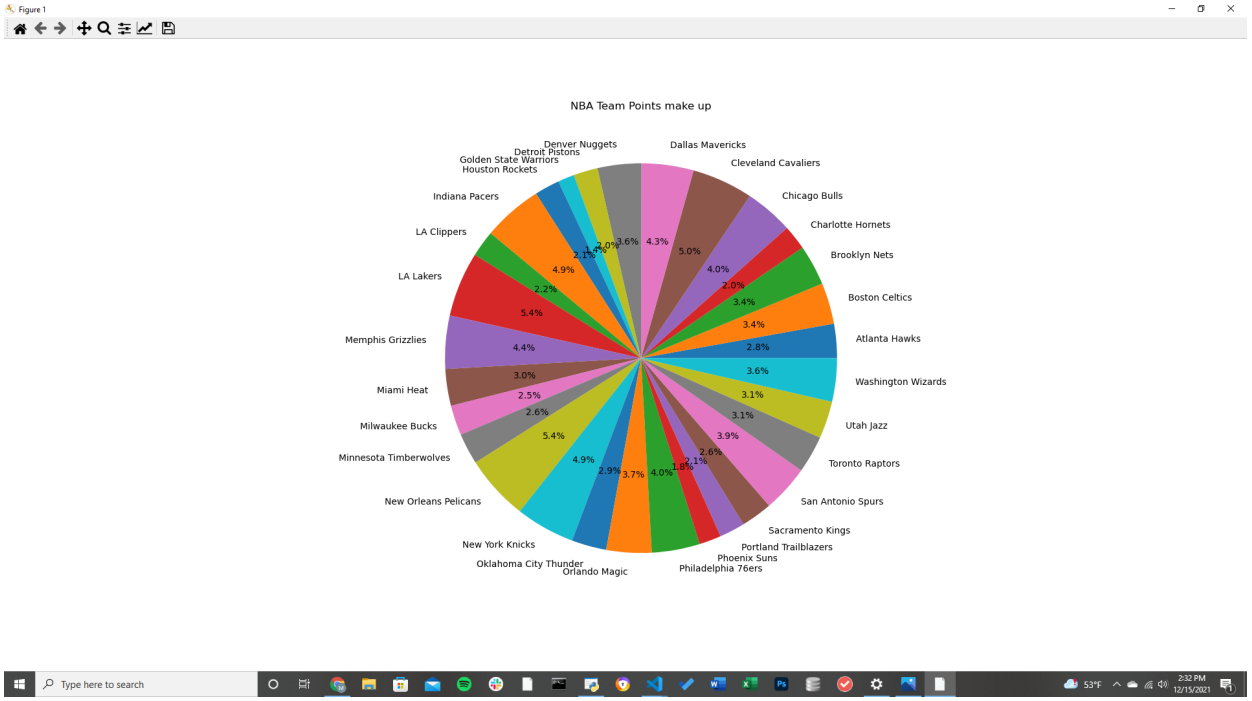
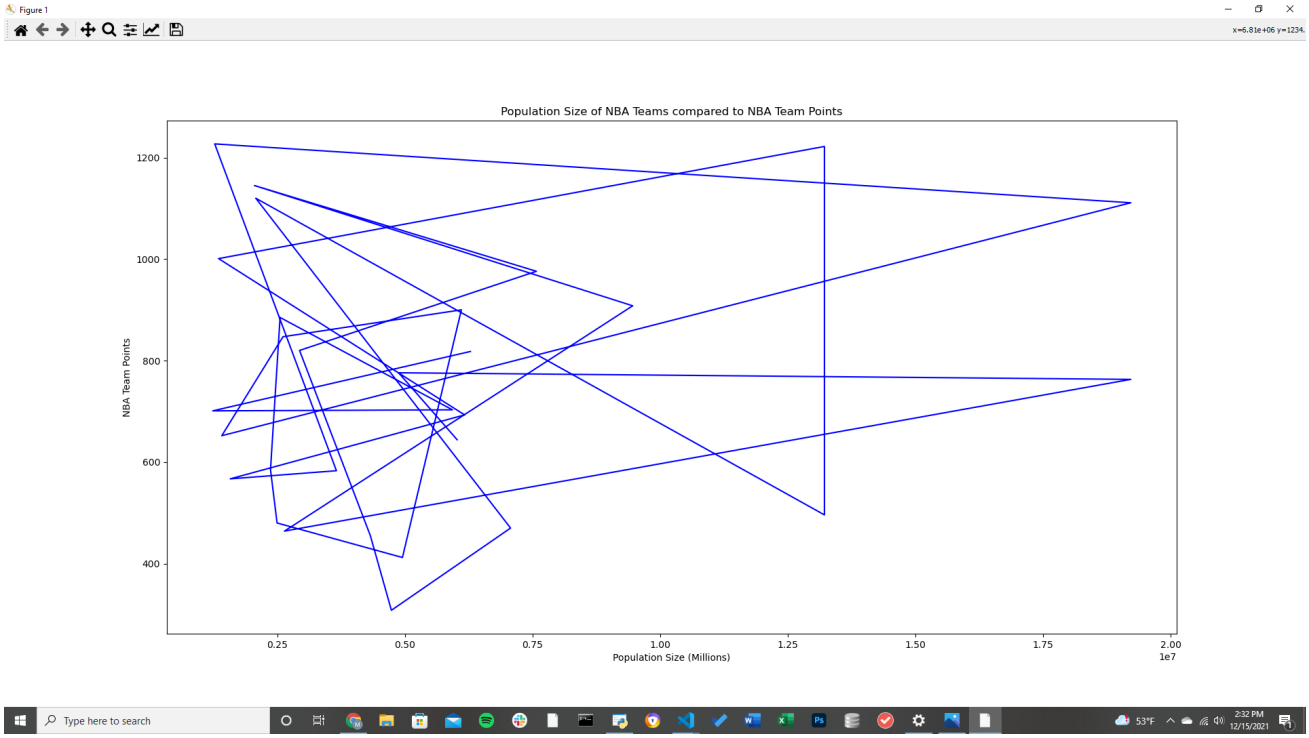


Figure 1
NBA Team Points make up
Dallas Mavericks
Cleveland Cavaliers
Chicago Bulls
Charlotte Hornets
Brooklyn Nets
Boston Celtics
Atlanta Hawks
Washington Wizards
Utah Jazz
Toronto Raptors
San Antonio Spurs
Sacramento Kings
Portland Trailblazers
Phoenix Suns
Philadelphia 76ers
Orlando Magic
Oklahoma City Thunder
New York Knicks
New Orleans Pelicans
Minnesota Timberwolves
Milwaukee Bucks
Miami Heat
Memphis Grizzlies
LA Lakers
LA Clippers
Indiana Pacers
Houston Rockets
Golden State Warriors
Detroit Pistons
Denver Nuggets
4.3%
5.0%
4.0%
2.0%
3.4%
3.4%
2.8%
3.6%
3.1%
3.1%
3.9%
2.6%
2.1%
1.8%
4.0%
3.7%
2.9%
4.9%
5.4%
2.6%
2.5%
3.0%
4.4%
5.4%
2.2%
4.9%
2.1%
1.4%
2.0%
3.6%
Type here to search
53°F
2:32 PM
12/15/2021



Instructions for Running Code

The following resources are needed to run this project:

- Sqlite3
- Matplotlib
- Json
- Regex
- BeautifulSoup
- readindata.py

To run the code, all you have to do is run the readindata.py Python file. This will retrieve the data and create the visualizations.

Function Documentation

- get_nba_data
    - Retrieves team game data from the Balldontlie API
    - Creates the team database

    - Also creates the team_data values: home_team, full_name, and home_team_score
    - Returns list of dictionaries of each home team's score for the 2018 season.
- Read_data_from_db
    - Reads in data from team_database.db
    - Selects out team data from database
    - Creates new dictionary to store teams and points that they scored for the 2018 season
    - Returns team_dict dictionary
- Get_pop_data
    - Creates a beautiful soup object to parse data on the page https://hoop-social.com/nba-team-market-size-rankings/
    - Finds the table on the website
    - Find rows of each team and appends the population information to a list.
    - Uses regex to eliminate the numbers of the team names on the table
    - Creates a second table in the team_database.db called "Populations" to store the population data of every team.
    - Returns dictionary of team population sizes.
- Barchart_population_size
    - Takes input in the form of the population data per city.
    - Sort points scored by teams in alphabetical order
    - Finds the mean points scored per team
    - Plots the population data against each team's city and the number of points each team scored. Makes the first bar blue and the second yellow.
    - Labels the x-axis "NBA Teams" and the y-axis "Population Size"
    - Titles the bar chart "Population Size of NBA Teams compared to NBA Team Points"
- Scatter_points
    - Takes input in the form of the population data per city

    - Takes the points scored from each team and sorts them by their names (i.e. Atlanta Hawks before Boston Celtics)
    - Creates a scatter plot with Population size on the x-axis
    - NBA team points are plotted via the y-axis
    - Titles the plot “Population Size of NBA Teams compared to NBA Team Points”
    - Visualizes data into a scatter plot
- Pie_chart
    - Takes input in the form of the U.S. population data per city
    - Sorts scored points by team name, like scatter_points
    - Appoints sorted team scores into another list
    - Plots a pie chart based on the percentage of the points that are scored by each team
- Line_graph
    - Once again, takes input in the form of U.S. population size per city
    - Sorts point scored by each team by their team name.
    - Appoints sorted team scores into another list
    - Plots population data against the team scores, makes visualization the color blue.
    - Defines the x-axis as Population Size and the y-axis as Team Points.
    - Titles the line graph “Population Size of NBA Teams compared to NBA team points)

Resource Documentation

Date	Issue Description	Location of Resource	Result (did it solve the issue)
December 2, 2021	**How to use balldontlie API**	**https://www.balldontlie.io/#get-a-specific-team**	**yes**

December 2, 2021	**How to use population estimate API**	**https://www.census.gov/data/developers/data-sets/popest-popproj/popest.html**	**no**
December 5, 2021	**How to create databases with SQLite**	**https://www.sqlitetutorial.net/sqlite-python/create-tables/**	**yes**
December 10, 2021	**How to use matplotlib to create visualizations**	**https://matplotlib.org/stable/tutorials/introductory/usage.html#sphx-glr-tutorials-introductory-usage-py**	**yes**
December 12, 2021	**Parse team population data**	**https://hoop-social.com/nba-team-market-size-rankings/**	**yes**