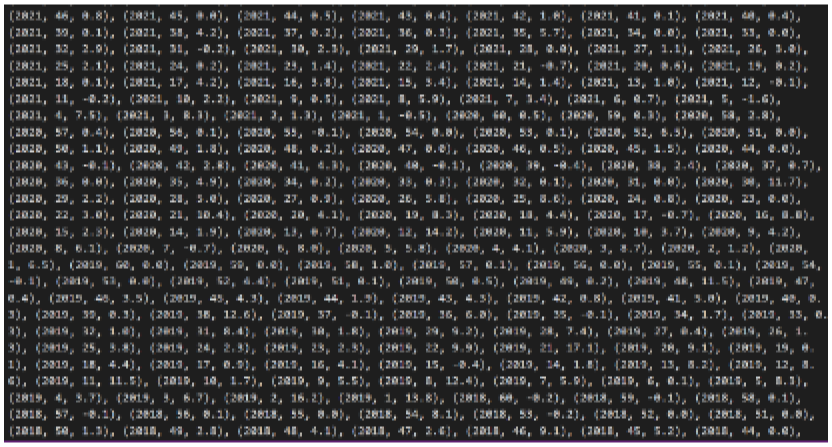
Poole Party: Casey Bressler, Samir Patel, and Thomas Chen
December 12, 2022
SI 206: Data-Oriented Programming
Final Project Report
Github link: https://github.com/dccaseyb/SI_206_Final_Project

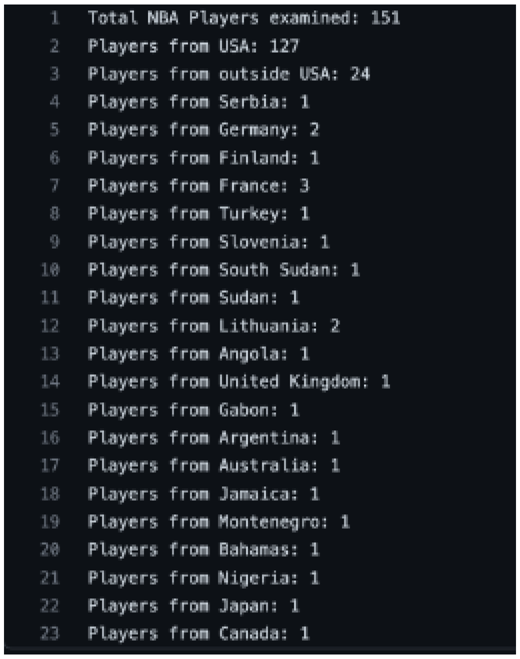
1. Our primary goals for this project were to examine the value of different NBA players with regard to their draft position and their home country. Our hypothesis was that the players taken with higher draft positions have, on average, accumulated higher value than lower draft positions and high win-share values. Additionally, we assumed that the United States would be the country with the most basketball players, but we were unsure of which other countries have also produced large amounts of NBA talent, so another goal was to determine which countries produce the most NBA-caliber players.
2. In order to examine these, we used the set of players drafted over the past five full seasons (2017-2021) to be our subset of the NBA population. Overall, we successfully achieved both of our goals for this project. First, regarding the value of players by their draft position, we did in fact find that players drafted higher in the draft tended to create more value. We also proved that the overwhelming majority (~84%, or 127 / 151) of players entering the NBA since 2017 were from the United States. The most common foreign nations for incoming NBA players were Germany (1.32%), France (~2%), and Lithuania (1.32%).
3. The largest problem that we faced when working on our project was the lack of reliable APIs that included the necessary data for this project. When initially researching our project, we found ESPN.com's APIs and believed they would certainly be functional for the purposes of this project. Unfortunately, after attempting to gather data from their various APIs, we found that as of August 31, ESPN no longer grants public access keys for their API, meaning that we had to find two new, separate APIs that included the data

we needed. Furthermore, some of the APIs we did find were not free; some were based on a subscription model while others were based on the “Freemium” model, with no upfront costs up to a certain amount of requests. Additionally, even the APIs we ended up using had missing information and data for certain players in certain contexts. For example, the API-NBA was missing start-year and career-length, as well as several other characteristics for several players. This forced us to only be able to examine the API player objects that had all sufficient data provided.

4. File

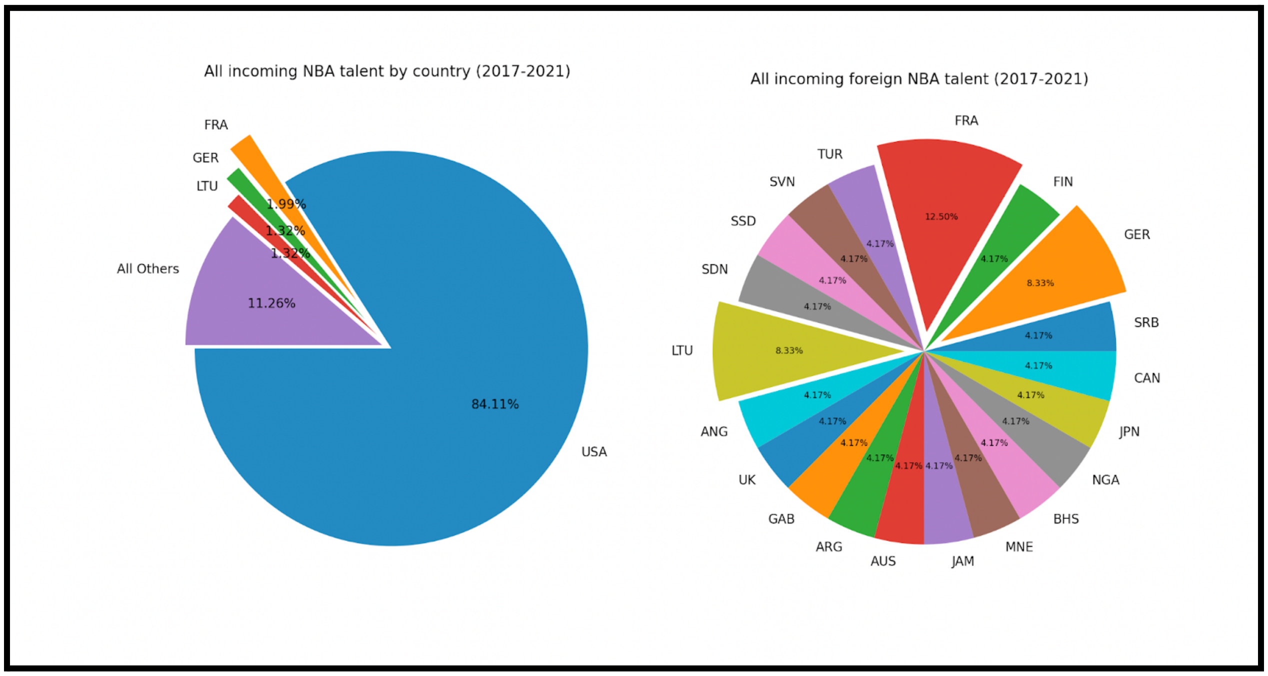


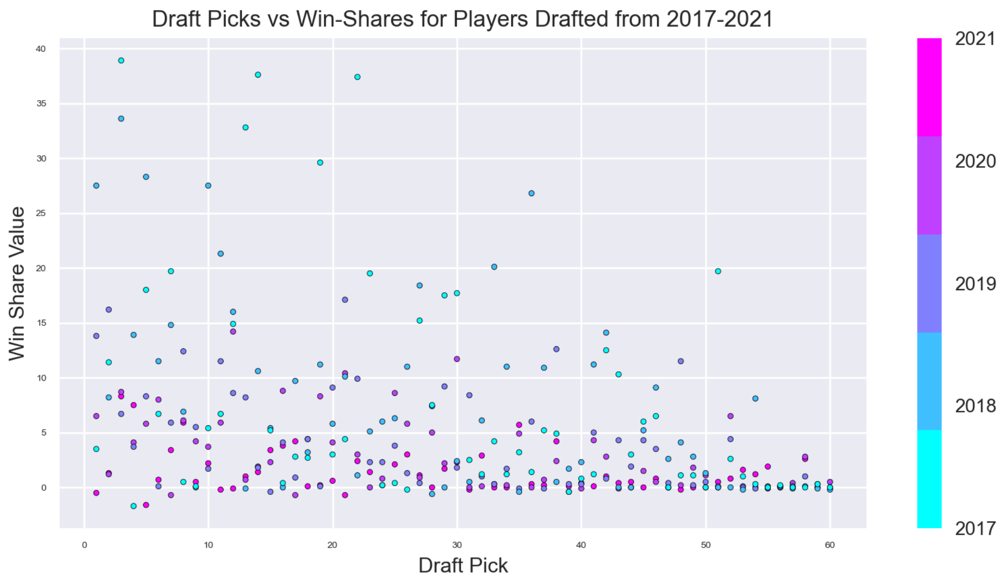
```
(2021, 46, 0.8), (2021, 45, 0.0), (2021, 44, 0.5), (2021, 43, 0.4), (2021, 42, 1.0), (2021, 41, 0.1), (2021, 40, 0.4),
(2021, 39, 0.1), (2021, 38, 4.2), (2021, 37, 0.2), (2021, 36, 0.3), (2021, 35, 5.7), (2021, 34, 0.0), (2021, 33, 0.0),
(2021, 32, 2.9), (2021, 31, -0.2), (2021, 30, 2.3), (2021, 29, 1.7), (2021, 28, 0.0), (2021, 27, 1.1), (2021, 26, 3.0),
(2021, 25, 2.1), (2021, 24, 0.2), (2021, 23, 1.4), (2021, 22, 2.4), (2021, 21, -0.7), (2021, 20, 0.6), (2021, 19, 0.2),
(2021, 18, 0.1), (2021, 17, 4.2), (2021, 16, 3.8), (2021, 15, 3.4), (2021, 14, 1.4), (2021, 13, 1.0), (2021, 12, -0.1),
(2021, 11, -0.2), (2021, 10, 2.2), (2021, 9, 0.5), (2021, 8, 5.9), (2021, 7, 3.4), (2021, 6, 0.7), (2021, 5, -1.6),
(2021, 4, 7.5), (2021, 3, 8.3), (2021, 2, 1.3), (2021, 1, -0.5), (2020, 60, 0.5), (2020, 59, 0.3), (2020, 58, 2.8),
(2020, 57, 0.4), (2020, 56, 0.1), (2020, 55, -0.1), (2020, 54, 0.0), (2020, 53, 0.1), (2020, 52, 6.5), (2020, 51, 0.0),
(2020, 50, 1.1), (2020, 49, 1.8), (2020, 48, 0.2), (2020, 47, 0.0), (2020, 46, 0.5), (2020, 45, 1.5), (2020, 44, 0.0),
(2020, 43, -0.1), (2020, 42, 2.8), (2020, 41, 4.3), (2020, 40, -0.1), (2020, 39, -0.4), (2020, 38, 2.4), (2020, 37, 0.7),
(2020, 36, 0.0), (2020, 35, 4.9), (2020, 34, 0.2), (2020, 33, 0.3), (2020, 32, 0.1), (2020, 31, 0.0), (2020, 30, 11.7),
(2020, 29, 2.2), (2020, 28, 5.0), (2020, 27, 0.9), (2020, 26, 5.8), (2020, 25, 8.6), (2020, 24, 0.8), (2020, 23, 0.0),
(2020, 22, 3.0), (2020, 21, 10.4), (2020, 20, 4.1), (2020, 19, 8.3), (2020, 18, 4.4), (2020, 17, -0.7), (2020, 16, 8.8),
(2020, 15, 2.3), (2020, 14, 1.9), (2020, 13, 0.7), (2020, 12, 14.2), (2020, 11, 5.9), (2020, 10, 3.7), (2020, 9, 4.2),
(2020, 8, 6.1), (2020, 7, -0.7), (2020, 6, 8.0), (2020, 5, 5.8), (2020, 4, 4.1), (2020, 3, 8.7), (2020, 2, 1.2), (2020,
1, 6.5), (2019, 60, 0.0), (2019, 59, 0.0), (2019, 58, 1.0), (2019, 57, 0.1), (2019, 56, 0.0), (2019, 55, 0.1), (2019, 54,
-0.1), (2019, 53, 0.0), (2019, 52, 4.4), (2019, 51, 0.1), (2019, 50, 0.5), (2019, 49, 0.2), (2019, 48, 11.5), (2019, 47,
0.4), (2019, 46, 3.5), (2019, 45, 4.3), (2019, 44, 1.9), (2019, 43, 4.3), (2019, 42, 0.8), (2019, 41, 3.0), (2019, 40, 0.
3), (2019, 39, 0.3), (2019, 38, 12.6), (2019, 37, -0.1), (2019, 36, 6.0), (2019, 35, -0.1), (2019, 34, 1.7), (2019, 33, 0.
3), (2019, 32, 1.0), (2019, 31, 8.4), (2019, 30, 1.8), (2019, 29, 9.2), (2019, 28, 7.4), (2019, 27, 0.4), (2019, 26, 1.
3), (2019, 25, 3.8), (2019, 24, 2.3), (2019, 23, 2.3), (2019, 22, 9.9), (2019, 21, 17.1), (2019, 20, 9.1), (2019, 19, 0.
1), (2019, 18, 4.4), (2019, 17, 0.9), (2019, 16, 4.1), (2019, 15, -0.4), (2019, 14, 1.8), (2019, 13, 8.2), (2019, 12, 8.
6), (2019, 11, 11.5), (2019, 10, 1.7), (2019, 9, 5.5), (2019, 8, 12.4), (2019, 7, 5.0), (2019, 6, 0.1), (2019, 5, 8.3),
(2019, 4, 3.7), (2019, 3, 6.7), (2019, 2, 16.2), (2019, 1, 13.8), (2018, 60, -0.2), (2018, 59, -0.1), (2018, 58, 0.1),
(2018, 57, -0.1), (2018, 56, 0.1), (2018, 55, 0.0), (2018, 54, 8.1), (2018, 53, -0.2), (2018, 52, 0.0), (2018, 51, 0.0),
(2018, 50, 1.3), (2018, 49, 2.8), (2018, 48, 4.1), (2018, 47, 2.6), (2018, 46, 9.1), (2018, 45, 5.2), (2018, 44, 0.0),
```



```
Total NBA Players examined: 151
Players from USA: 127
Players from outside USA: 24
Players from Serbia: 1
Players from Germany: 2
Players from Finland: 1
Players from France: 3
Players from Turkey: 1
Players from Slovenia: 1
Players from South Sudan: 1
Players from Sudan: 1
Players from Lithuania: 2
Players from Angola: 1
Players from United Kingdom: 1
Players from Gabon: 1
Players from Argentina: 1
Players from Australia: 1
Players from Jamaica: 1
Players from Montenegro: 1
Players from Bahamas: 1
Players from Nigeria: 1
Players from Japan: 1
Players from Canada: 1
```

5. Visualizations





30
25
20
15
10

The above chart shows the distribution of players by draft position, showing that the players in our data set were extremely balanced between positions.

6. Code Running Instructions
7. Documentation For Each Function

Country.py

- def getResponse(cur, conn):
    - Inputs: cursor & connection objects
    - Purpose: get requests, create json_file, add appropriate items to database table
    - Output: None
- def addToTable(cur, conn, id, s, f, l, c):
    - Inputs: cursor & connection objects, player_id, player_start_year, player_firstname, player_lastname, player_country
    - Purpose: add appropriate items to table in database
    - Output: None
- def setUpDatabase(db_name):
    - Inputs: database name string
    - Purpose: creates database as well as cursor & connection objects
    - Output: None
- def draw_graph():
    - Inputs: cursor & connection objects
    - Purpose: draw visualization for incoming NBA players (2017-2021) by country
    - Output: 2 Pie Charts attached above

Winshares.py

- def scrapedata(years):
    - Input: list of years (NBA draft years)
    - Purpose: Adds appropriate items to database
    - Output: None
- def removeNull(row):
    - Input: List of strings
    - Purpose: Replaces empty values (collected from data) and replaces with a 0
    - Output: None
- def finaldata(draftyear, draftpick, winshares):
    - Input: database draft year, draft pick, and win shares lists
    - Purpose: Combines lists into a database
    - Output: None
- def drawgraph(draftpick, winshare, draftyear):
    - Input: database draft year, draft pick, and win shares lists
    - Purpose: Creates a scatterplot of win shares vs draft pick with the draft year as a third color variable
    - Output: Scatterplot

8. Resources We Used

Date	Issue Description	Resource Location	Result
12/5/22	The ESPN API no longer offers public access keys	GitHub	We found an API to use for the project
12/7/22	Issues accessing the data in the NBA	GitHub and The API's Slack	We were finally able to access the

	API we were using	Channel	data, we needed specific custom headers within the request
12/8/22	The NBA API would crash when taking multiple requests in one run of the code	API’s Slack Channel	We implemented a sleeper to prevent the API from crashing