Caitlin O'Grady
Emy Klepacki
Silvia Deyarmond

**PART 5 – Report (100 points)**

**1. The goals for your project including what APIs/websites you planned to work with and what data you planned to gather (10 points):**

- Our goal for this project was to learn about and utilize API data from different sports APIs. These included basketball (NBA), hockey (NHL), and baseball (MLB) data sets. Within each API, we hoped to create tables based on the available data, process the data, and finally create visual representations of the data. For the NBA, we planned to gather data player statistics per team and season, standings per team and season, games per team and season. We planned to gather this data from the RapidAPI NBA. For the NHL, we planned to gather team wins, team losses, goals per game, goals against per game, etc. For the MLB, we planned to gather data on team wins and losses. We planned to gather this data from https://gitlab.com/dword4/nhlapi.

**2. The goals that were achieved including what APIs/websites you actually worked with and what data you did gather (10 points)**

- We did work with the NBA RapidAPI. However, we shifted our focus from individual player statistics and rather focused on team statistics during one season. We gathered the statistics from the 2020 season on NBA teams points, points in the paint, rebounds, assists, personal fouls, steals, turnovers, and blocks. For NHL data, we continued to use the initial API above. However, we shifted our focus from team statistics to 2022 draft prospect statistics since the team statistics were limited. So, we gathered player pick overall, team ID, player ID, player name, and team name for each prospect. Due to issues

we switched from the MLB API to an NFL one, and calculated 2022 draft prospects with that, as well.


**3. The problems that you faced (10 points)**

- We faced problems when accessing the API data. Some APIs were outdated, so the sites exited with error. Some datasets were too minimal to create large tables (with at least 100 unique rows), so we had to do searching to find the right APIs. It was very hard figuring out how to only insert less than 25 rows of data per execution of the file. This was hard because it was a challenge figuring out where to pick back up/how to keep the spot and continue adding rows where it was left off without changing anything about the file. It was also a challenge because the NBA api had a limit of about 9 requests per minute, so it was hard to figure out how to time the file so that it would not overload the requests and therefore fail.

**4. The calculations from the data in the database (i.e. a screen shot) (10 points)**

```
Team Counts for 2022 NHL Draft:
Montréal Canadiens: 6
New Jersey Devils: 2
Arizona Coyotes: 7
Seattle Kraken: 8
Philadelphia Flyers: 2
Columbus Blue Jackets: 4
Chicago Blackhawks: 8
Detroit Red Wings: 3
Buffalo Sabres: 5
Anaheim Ducks: 4
Winnipeg Jets: 5
Vancouver Canucks: 2
Nashville Predators: 3
Dallas Stars: 3
Minnesota Wild: 5
Washington Capitals: 4
Pittsburgh Penguins: 1
St. Louis Blues: 3
San Jose Sharks: 4
Tampa Bay Lightning: 2
Edmonton Oilers: 1
Toronto Maple Leafs: 2
Vegas Golden Knights: 2
Los Angeles Kings: 1
Boston Bruins: 1
Calgary Flames: 1
```

```
Players Drafted Per Team in the 2022 NFL Draft:
New York: 18
Los Angeles: 16
Baltimore: 11
Green Bay: 11
Chicago: 11
Kansas City: 10
Minnesota: 10
New England: 10
Houston: 9
Tennessee: 9
Seattle: 9
Denver: 9
Cleveland: 9
San Francisco: 9
Dallas: 9
Detroit: 8
Buffalo: 8
Washington: 8
Atlanta: 8
Arizona: 8
Indianapolis: 8
Tampa Bay: 8
Pittsburgh: 7
Jacksonville: 7
Carolina: 6
Cincinnati: 6
Las Vegas: 6
Philadelphia: 5
New Orleans: 5
Miami: 4
```

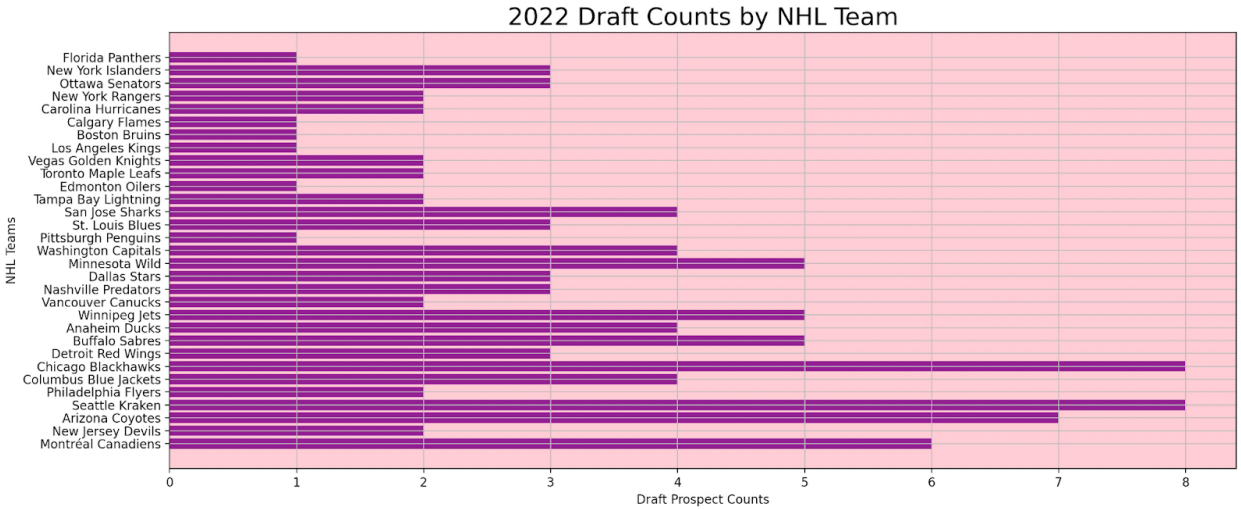
1.

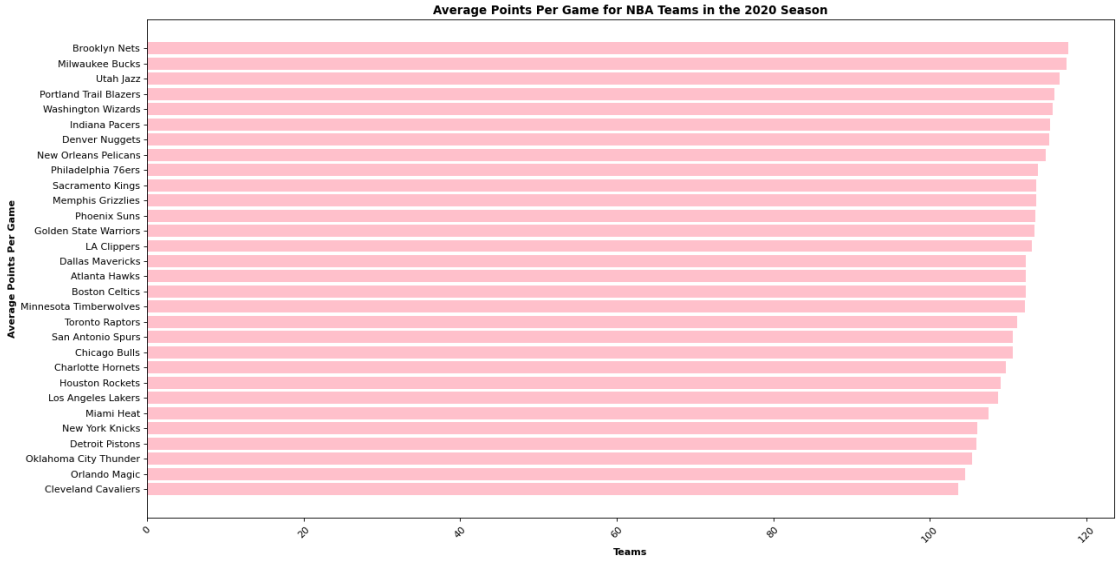
```
NBA Teams Average Points Per Game for the 2020 Season

The Atlanta Hawks had 10550 points in 94 games in the 2020 season, averaging 112.23 points per game.
-------------------------
The Boston Celtics had 8976 points in 80 games in the 2020 season, averaging 112.2 points per game.
-------------------------
The Brooklyn Nets had 10119 points in 86 games in the 2020 season, averaging 117.66 points per game.
-------------------------
The Charlotte Hornets had 8447 points in 77 games in the 2020 season, averaging 109.7 points per game.
-------------------------
The Chicago Bulls had 8406 points in 76 games in the 2020 season, averaging 110.61 points per game.
-------------------------
The Cleveland Cavaliers had 7875 points in 76 games in the 2020 season, averaging 103.62 points per game.
-------------------------
The Dallas Mavericks had 9205 points in 82 games in the 2020 season, averaging 112.26 points per game.
-------------------------
The Denver Nuggets had 9792 points in 85 games in the 2020 season, averaging 115.2 points per game.
-------------------------
The Detroit Pistons had 8052 points in 76 games in the 2020 season, averaging 105.95 points per game.
-------------------------
The Golden State Warriors had 8732 points in 77 games in the 2020 season, averaging 113.4 points per game.
-------------------------
The Houston Rockets had 8289 points in 76 games in the 2020 season, averaging 109.07 points per game.
-------------------------
The Indiana Pacers had 8878 points in 77 games in the 2020 season, averaging 115.3 points per game.
-------------------------
The LA Clippers had 10621 points in 94 games in the 2020 season, averaging 112.99 points per game.
-------------------------
The Los Angeles Lakers had 9019 points in 83 games in the 2020 season, averaging 108.66 points per game.
-------------------------
The Memphis Grizzlies had 9423 points in 83 games in the 2020 season, averaging 113.53 points per game.
-------------------------
The Miami Heat had 8382 points in 78 games in the 2020 season, averaging 107.46 points per game.
-------------------------
The Milwaukee Bucks had 11513 points in 98 games in the 2020 season, averaging 117.48 points per game.
-------------------------
```

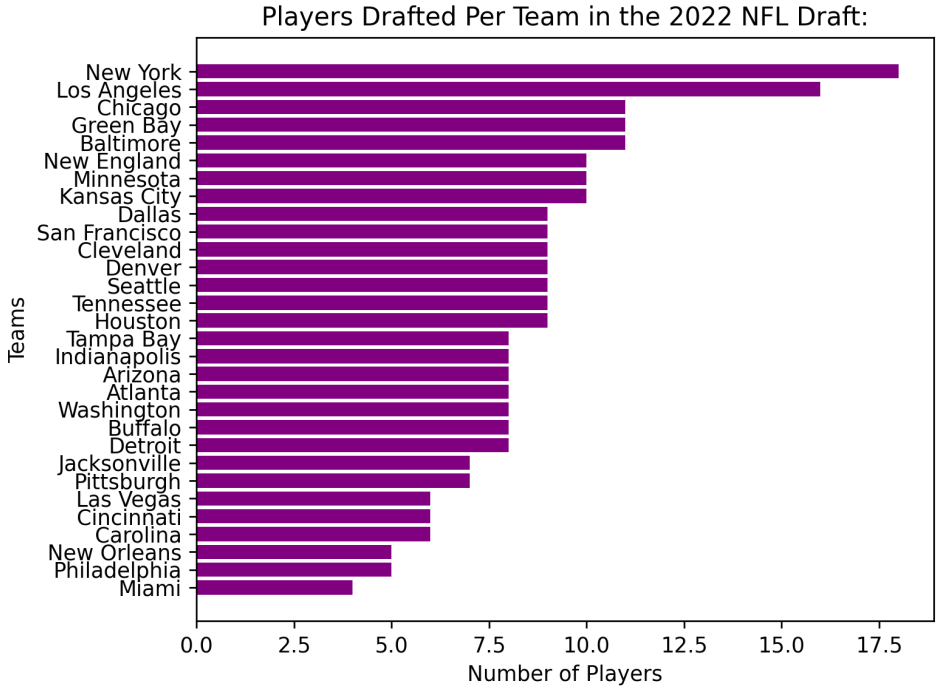
2.

```
The Minnesota Timberwolves had 8411 points in 75 games in the 2020 season, averaging 112.15 points per game.
-------------------------
The New Orleans Pelicans had 8492 points in 74 games in the 2020 season, averaging 114.76 points per game.
-------------------------
The New York Knicks had 8591 points in 81 games in the 2020 season, averaging 106.06 points per game.
-------------------------
The Oklahoma City Thunder had 7797 points in 74 games in the 2020 season, averaging 105.36 points per game.
-------------------------
The Orlando Magic had 7946 points in 76 games in the 2020 season, averaging 104.55 points per game.
-------------------------
The Philadelphia 76ers had 9677 points in 85 games in the 2020 season, averaging 113.85 points per game.
-------------------------
The Phoenix Suns had 11115 points in 98 games in the 2020 season, averaging 113.42 points per game.
-------------------------
The Portland Trail Blazers had 9501 points in 82 games in the 2020 season, averaging 115.87 points per game.
-------------------------
The Sacramento Kings had 8631 points in 76 games in the 2020 season, averaging 113.57 points per game.
-------------------------
The San Antonio Spurs had 8406 points in 76 games in the 2020 season, averaging 110.61 points per game.
-------------------------
The Toronto Raptors had 8339 points in 75 games in the 2020 season, averaging 111.19 points per game.
-------------------------
The Utah Jazz had 10023 points in 86 games in the 2020 season, averaging 116.55 points per game.
-------------------------
The Washington Wizards had 9489 points in 82 games in the 2020 season, averaging 115.72 points per game.
-------------------------
The Brooklyn Nets had the highest average points per game at 117.66 points per game
AND
The Cleveland Cavaliers had the lowest average points per game at 103.62 points per game
```

## 5. The visualization that you created (i.e. screen shot or image file) (10 points)






**6. Instructions for running your code (10 points)**

- FOR NBA FILES: There is a time delay written into the execution of NBAdatacomp.py. It should take about 4 minutes to add about 22 rows across the NBATeams and the NBASeasonStats tables in the database, 11 rows each. It will have to be run about 10 times in order to add the complete number of rows. If there is a key error about response, it just means that the code was run too quickly and the API didn't give a response. Just give it about a minute or so and then continue to run the file. The NBAcalculations.py should work just fine, it will create a txt file with the data calculations done as well as create a matplotlib pop up of the graph I created.
- FOR NHL FILES: Run the NHL.py file to add to the database, do calculations, and create visualizations. If the DB table is too full, uncomment line 112 where it drops the table, and simply run the code 4 times to get to 100 values.
- FOR NFL FILES: Run NFL.py. The last line runs the function that adds 25 rows to the specified database 11 times for the year 2022, so this can be edited to change database, number of API calls, or year.

**7. Documentation for each function that you wrote. This includes describing the input and output for each function (20 points)**

Caitlin O'Grady
Emy Klepacki
Silvia Deyarmond

**NBA**

1. Teamtable
    - Input:
        - Start = the id number of the team you would like to start at when adding rows
        - Db = database
    - Output:
        - None
2. Getstats
    - Input:
        - Id = the id number of the team you want stats for
        - Db = database
    - Output:
        - None
3. Teaminfo
    - Input:
        - Idlist: list of team id numbers that you want info for
    - Output:
        - A list of tuples with the structure of (id,name,games in a season, total points)
4. Calc_pper
    - Input:
        - Pperlist = A list of tuples with the structure of (id,name,games in a season, total points)
        - F = file to write to
    - Output:
        - None
5. Vis
    - Input
        - Tavers = a list of tuples with the structure (name, points per game)
    - Output
        - none

**NHL**

1. setUpDatabase():
    - Input
        - None
    - Output

        - cur, conn
2. getAPI(url):
    - Input
        - url for DB
    - Output
        - file to access data
3. create_team_table(cur, conn):
    - Input
        - cur, conn
    - Output
        - None (creates table)
4. def add_draft_data(cur, conn, jfile):
    - Input
        - cur, conn, file from create_team_table
    - Output
        - None (adds data)
5. calculate_teams(cur, conn):
    - Input
        - cur, conn
    - Output
        - dictionary of team data
6. visualizeData(cur, conn, dct):
    - Input
        - cur, conn, dictionary from calculate_teams
    - Output
        - None (graph visualization)

**NFL**

1. load_data_25(db, year)
    - Input
        - File name for a database
        - Year of data to get
    - Output
        - None (commits to the given database)
2. load_data_full(db, num, year)
    - Input
        - File name for a database
        - Number of times to fetch data
        - Year of data to get

    - Output
        - Dict of players drafted per team from the database
        - Also writes a txt file with that data
3. graphdata(datadict, title)
    - Input
        - Dict of teams from load_data_full
        - Title for the graph
    - Output
        - None (draws a pyplot graph)

8. You must also clearly document all resources you used. The documentation should be of the following form (20 points)

**Date – Issue – Description – Location of Resource – Result (did it solve the issue?)**

| Date | Issue Description | Location of Resource | Result |
|---|---|---|---|
| | Issues figuring out how to structure matplotlib and create a graph | https://stackabuse.com/rotate-axis-labels-in-matplotlib/ | Produced a workable graph |
| | How to add a time delay into the code so as not to overwhelm the requests per minute of the API | https://realpython.com/python-sleep/ | Added a time delay and successfully avoided request overload |

| Date | Issue Description | Location of Resource | Result |
|---|---|---|---|
| 4/19/23 | How to make sure table data is not duplicated | https://www.geeksforgeeks.org/python-mysql-insert-record-if-not-exists-in-table/ | Ensured that player results were not repeated in database tables |
| 4/20/23 | How to customize graph on Matplotlib | https://matplotlib.org/stable/gallery/lines_bars_and_markers/bar | Successfully was able to change graph to make it personable |

| | | | |
|---|---|---|---|
| | | _label_demo.html | |