**Website**: <https://www.basketball-reference.com/>

Name and the URLs (at least 5) of the site you will be scraping data from.

1. LeBron James: <https://www.basketball-reference.com/players/j/jamesle01.html>
2. NBA 2024-25 Draft: <https://www.basketball-reference.com/leagues/NBA_2025.html>
3. Lakers 2024-25 Season Summary: <https://www.basketball-reference.com/teams/LAL/2025.html>
4. NBA 2024-25 Coaches: <https://www.basketball-reference.com/leagues/NBA_2025_coaches.html>
5. NBA 2024 Draft: <https://www.basketball-reference.com/draft/NBA_2024.html>

5 Data Science Questions (2 plot):

1. How has LeBron James’ scoring efficiency (TS%, PER) changed over the last 10 years?
2. Which teams in the eastern conference had the best win-lose percentage and how does it correlate with opponent points per game?
3. In the 2024-25 season, how does the Lakers' performance compare to their opponents? Judge stats like field goal percentage, total rebounds, and total points
4. Which coaches of the 2024-2025 season had the highest win ratio?
5. What impact do the Round 1 2024 draft members with the top 10 points-per-game have on their team wins? Assess points per game (PTS), assists (AST), and win shares (WS) among the top 10 points-per-game draft players.

Data Scraping Plan:

First, use requests and BeautifulSoup or pandas.read\_html() for fast scraping of tables.

Scrape tables with table IDs or captions.

Tables from pages:

**2024-2025 Season Summary:** Use "Per Game Stats" and "Advanced Stats" tables.

id="per\_game-team"

<table class="stats\_table sortable now\_sortable" id="per\_game-team" data-cols-to-freeze=",2">to-freeze="1,2">

<caption>Per Game Stats Table</caption>

id="advanced-team"

<table class="stats\_table sortable now\_sortable" id="advanced-team" data-cols-to-freeze=",2">

<caption>Advanced Stats Table</caption>

**LA Lakers**: Use "Roster" and "Advanced" tables.

id="roster"

<table class="sortable stats\_table now\_sortable" id="roster" data-cols-to-freeze=",2">

<caption>Roster Table</caption>

id="advanced"

<table class="stats\_table sortable soc now\_sortable" id="advanced" data-cols-to-freeze=",2" data-soc-sum-scope-type="player\_season" data-soc-sum-phase-type="reg" data-soc-sum-params="null" data-soc-sum-year="2024">

<caption>Advanced Table</caption>

**LeBron James:** Use "Per Game" and “Advanced” table under career stats.

id="per\_game\_stats"

<table class="stats\_table sortable row\_summable suppress\_headers soc now\_sortable" id="per\_game\_stats" data-cols-to-freeze="1,3" data-soc-sum-scope-type="player\_season" data-soc-sum-phase-type="reg" data-soc-sum-table-type="PlayerPerGame" data-soc-sum-entity-id="jamesle01" data-soc-sum-params="null">

<caption>Per Game Table</caption>

id="advanced"

<table class="stats\_table sortable row\_summable suppress\_headers soc now\_sortable" id="advanced" data-cols-to-freeze="1,3" data-soc-sum-scope-type="player\_season" data-soc-sum-phase-type="reg" data-soc-sum-table-type="Advanced" data-soc-sum-entity-id="jamesle01" data-soc-sum-params="null">

<caption>Advanced Table</caption>

**2024 Coaches**: Use "2024-2025 NBA Coaches" table with team win-loss (W-L) data.

id="NBA\_coaches"

<table class="sortable stats\_table now\_sortable" id="NBA\_coaches" data-cols-to-freeze="1,2">

<caption>2024-25 NBA Coaches Table</caption>

**2024 Draft**: Use "NBA Draft Table" for rookie stats.

id="stats"

<table class="sortable stats\_table now\_sortable" id="stats" data-cols-to-freeze="2,4">

<caption>60 Selections Table</caption>

**Question 4:** Which teams in the eastern conference had the best win-lose percentage and how does it correlate with opponent points per game? (plot)

<https://www.basketball-reference.com/leagues/NBA_2025.html>

<table class="suppress\_all sortable stats\_table now\_sortable sticky\_table eq1 re1 le1" id="confs\_standings\_E" data-cols-to-freeze=",1">

<caption>Conference Standings Table</caption>

<th aria-label="Win-Loss Percentage" data-stat="win\_loss\_pct" scope="col" class=" poptip right" data-tip="Win-Loss Percentage">W/L%</th>

<th aria-label="Opponent Points Per Game" data-stat="opp\_pts\_per\_g" scope="col" class=" poptip right" data-tip="Opponent Points Per Game">PA/G</th>

Plot: Create a scatter plot comparing win-loss percentage vs opponent points per game for each team in the eastern conference.

The X axis will be the opponent points per game.

The Y axis will be a win percentage.

Each team can be a different color point on the plot.

**Question 5**: How did LeBron James' 2023–2024 season compare to his career averages in key stats like points, assists, and minutes per game?

Lebron James Stats: <https://www.basketball-reference.com/players/j/jamesle01.html>

Extract the per-season stat table and compare it with career stats.

Scrape stats from the page:

Points per game (PPG)

Assists per game (APG)

Minutes per game (MPG)

Table:<caption>Per Game Table</caption>

2023-24 season table row:<tr id="per\_game\_stats.2024" data-row="20" class>

minutes:<td class="right " data-stat="mp\_per\_g" csk="35.2676056338">35.3</td>

assists:<td class="right " data-stat="ast\_per\_g" csk="8.2957746479">8.3</td>

points:<td class="right " data-stat="pts\_per\_g" csk="25.6619718310">25.7</td>

Table with career:<div class="stats\_pullout">

Points:<span class="poptip" data-tip="Points">…</span>

Season:<p>24.4</p>

Career:<p>27.0</p>

Assists: <span class="poptip" data-tip="Assists">…</span>

Season:<p>8.2</p>

Career:<p>7.4</p>

Plot: Create a bar chat with both career average and 2023–2024 season values side by side.

**Restrictions in robots.txt**

Robots.txt file: <https://www.basketball-reference.com/robots.txt>

Restrictions:

* AhrefsBot, GPTBot are completely blocked.
* Twitter is not blocked.

These pages are blocked:

* /basketball/
* /blazers/
* /dump/
* /fc/
* /my/
* /7103
* /play-index/\*.cgi?\*
* /play-index/plus/\*.cgi?\*
* \*/gamelog/
* \*/splits/
* \*/on-off/
* \*/lineups/
* \*/shooting/
* /req/
* /short/
* nocdn/

The [plagiarism.org](http://plagiarism.org/) robot and [www.slysearch.com](http://www.slysearch.com/) are also blocked.

These robots are blocked:

* SlySearch
* GroundControl
* Ground-Control
* Carmine
* Skynet
* The-Matrix
* Matrix
* HAL9000

Restriction: Crawl/request delay of 3 seconds must be followed.