

# Summary

## Checkpoint 8

|  |          |
|--|----------|
| <b>Overview.....</b>                               | <b>1</b> |
| Group Work Option.....                             | 1        |
| <b>Assignment Objectives.....</b>                  | <b>1</b> |
| <b>What to Include in Your Checkpoint.....</b>     | <b>2</b> |
| <b>Visual and Writing Standards.....</b>           | <b>2</b> |
| <b>Submission Details.....</b>                     | <b>3</b> |
| <b>Deliverables.....</b>                           | <b>3</b> |
| <b>Grading &amp; Rubric (10 points total).....</b> | <b>3</b> |
| <b>Additional Notes.....</b>                       | <b>3</b> |

### Overview

Checkpoints 6 and 7 produced analysis and an improved result.

Checkpoint 8 turns that work into a short, clear executive summary that a decision maker can read in two minutes. Focus on the problem, the evidence, the takeaway, and the action.

### Group Work Option

You may work in groups of up to 3 students, but each student must maintain their own GitHub repository. Collaboration on topic and data collection is encouraged, but each student must submit an individual memo and ensure their repository contains all deliverables.

### Assignment Objectives

By completing this checkpoint, you will:

1. Communicate the question, the approach, and the result in plain language
2. Show one key exhibit that supports the takeaway
3. Make a clear recommendation tied to the result and to stakeholder goals
4. State risks and limits and how you would handle them
5. Outline next steps that keep the project moving

## DATA 6560 - Sports Analytics - Checkpoint 8 Fall 2025 What to Include in Your Checkpoint

|                             |  |   |
|-----------------------------|--|---|
|                             |  |   |
| Title Block and Stakeholder | Project title, team members, date, primary stakeholder or audience                           | Who is this for and why should they care right now?   |
| Problem and Decision        | One short paragraph that states the question and the decision that will use this result      | What decision are we informing?<br>What would a good decision look like in this context?                                    |
| Data in One Line            | Name the cleaned file from CP4, row definition, time span, and the scope of the sample       | What does one row represent?<br>What time period did you cover?<br>Are there any key filters the reader should know?        |
| Approach in One Line        | Name the method from CP6 or CP7 in plain language  | What did you compare or predict and why was that the right method here?   |
| Key Result                  | The headline number in sport or business units, plus a short sentence on why it matters      | How big is the effect or accuracy in terms a coach or manager would understand?<br>Is it large enough to change a decision? |
| One Key Exhibit             | A single figure or table that backs the headline. Include a one line takeaway under it       | Does this graphic directly support the claim above without extra clutter?<br>Are axes and units labeled and readable?       |
| Recommendation              | A clear, specific action tied to the result  | What should the stakeholder do, starting when, and how will we measure success?   |
| Risks & Limits              | The top 1 to 2 caveats that might change the call and how you would monitor or mitigate them | What could make this result fragile or misleading? What simple safeguard can we put in place?                               |
| Next Steps                  | Two to three concrete steps for the next week of work  | What will you refine, test, or deliver next to increase confidence or impact?   |

### Visual and Writing Standards

- Keep it to the point. Two pages maximum for the memo, not counting an optional appendix
- Use one primary figure. If you include more, move extras to an appendix
- Label titles, axes, and units. Put a one line takeaway under the figure
- Write in plain language. Prefer short sentences. Use numbers that people can feel, such as points, yards, percentage points, tickets, or dollars

- If you report accuracy or error, also report a simple baseline for context

## DATA 6560 - Sports Analytics - Checkpoint 8 Fall 2025 Submission Details

|               |   |
|---------------|---|
|               |   |
| Document Type | Google Doc in professional memo style, export to PDF  |
| Submission    | Share the Google Doc link on Classroom and upload the PDF to your repo in /reports. Save the figure in /figures   |
| File Naming   | <p>LastName_FirstName_CP8.pdf (example: Doe_Jane_CP8.pdf). Include your name(s), project title, and the checkpoint heading at the top of the memo.</p> <p>EACH STUDENT MUST UPLOAD TO INDIVIDUAL REPO</p> |
| Format Tips   | Include the single key exhibit file and keep any supporting tables or notebooks in the repo folders you used earlier. Optional appendix for extra figures if needed.                                      |

### Deliverables

1. Executive Summary in PDF
2. One key exhibit saved in /figures

### Grading & Rubric (10 points total)

Your Checkpoint 8 will be evaluated on the following criteria:

|                       |   |   |
|-----------------------|---|---|
|                       |   |   |
| Clarity and Focus     | Problem, result, and action are easy to understand on a quick read. No extra fluff  | 3 |
| Evidence and Accuracy | Key number is clearly supported by the exhibit and matches the CP6 and CP7 work. Baseline context is provided when relevant | 3 |
| Actionability         | Recommendation is specific, time bound, and tied to stakeholder goals and metrics   | 2 |
| Risks and Limits      | Realistic caveats with a simple plan to monitor or mitigate   | 1 |
| Professional Polish   | Clean formatting, correct file names and locations, readable figure with a one line takeaway                                | 1 |

### Additional Notes

- Lead with the headline. Put the most important sentence in the first two lines
- Use the strongest chart you have. Remove anything that does not help the decision
- Convert stats into practical units when possible

- *If the recommendation depends on a threshold, name that threshold and why it was chosen* ●
- *If results are mixed, say what works and where it does not, then give the safest path forward*

DATA 6560 - Sports Analytics - Checkpoint 8 Fall 2025

To: Director of Basketball Strategy & Analytics

From: Connor Lewis

Analytics Team Date: December 15, 2025 Subject: Predictive Modeling of NBA Player Roles Using Efficiency Metrics - 2017-2018 Season

**Executive Summary** This memorandum has considered a data-driven approach toward player role analysis in the NBA, leveraging the dataset `NBA_2017_2018_cleaned.csv`. Using Multiple Linear Regression and a Gradient Boosted Classification model, we gauged which efficiency metrics are most related to whether a player performs at an Elite, Starter, or Role-player level. The Classifier had perfect accuracy on the test set, suggesting strong potential for informing roster decisions, contract valuations, and trade strategy.

**Key Insight:** The Gradient Boosted Classification model correctly classified all the players in the test sample with 100% accuracy, showing that efficiency metrics like PER, WS/48, and BPM will indeed support true separation between role tiers. This level of precision means that the model can serve as a dependable screening tool for determining undervalued contributors or overestimated assets ahead of personnel decisions.

**Supporting Evidence** A confusion matrix of model performance, which was summarized, had perfect diagonal alignment with zero misclassifications across Elite, Starter, and Role categories. This serves to confirm that the structure of the model captures the underlying pattern variation in player performance. A residuals-versus-fitted plot from the regression model highlighted where predictions are closely aligned with the observed tier scores and where individual deviations occur, offering transparency to model behavior.

**Implications for Decision-Makers:** This modeling framework will immediately support decisions related to rotation optimization, trade targeting, and contract negotiations. The team, by identifying players whose statistical profiles exceed their perceived market value, is able to pursue cost-efficient upgrades or avoid overpaying for inflated roles. The model flags players whose production does not justify their current usage or salary trajectory.

**Recommendation** Incorporate the classification model into the front office's weekly player evaluation workflow. Begin using it to identify possible trade targets and internal players whose roles should be adjusted. Success will be measured by whether model-identified players outperform or underperform expectations over the next 20–30 games.

**Risks & Limitations** The existing model is trained on a single season and relatively limited test sample, increasing the risk of overfitting. Player roles might also change due to coaching changes, injuries, or roster context. Expanding the dataset into multiple seasons will reduce these risks by monitoring predictive drift upon the arrival of new data.

**Next Steps**

Expand the data to encompass seasons 2018 through 2020 to increase generalizability.  
DATA 6560 - Sports Analytics - Checkpoint 8 Fall 2025

Validate this model on current-season data to verify real-time stability. Create a dashboard that shows coaches and executives how to explore player tier predictions interactively.