DATA 6560 - Sports Analytics - Checkpoint 1 Fall 2025 Proposal Memo

Checkpoint 1

This week marks the first major checkpoint of your semester-long Sports Analytics portfolio project.

In this Proposal Memo, you will officially propose your project topic, define the problem you plan to address, and outline your intended approach. This checkpoint ensures your project is feasible, well-defined, and grounded in real datam setting the foundation for the rest of the semester.

Overview

Each student (or team) will submit a professional proposal memo (~1 page) that introduces their project idea, identifies a dataset, and outlines a high-level plan for analysis.

Group Work Option

You may work in groups of up to 3 students, but each student must maintain their own GitHub repository for this project (we'll cover setup in a later checkpoint). You are encouraged to collaborate on the same topic, share data, and divide tasks.

Assignment Objectives

By completing this checkpoint, you will:

1. Select a feasible sports analytics topic and define a clear research question or problem.
2. Identify a relevant dataset or data source that can address your question.
3. Explain why your topic matters and who benefits from the insights.
4. Outline a preliminary analysis plan at a high level.
5. Practice writing in a clear, concise, professional memo format.

DATA 6560 - Sports Analytics - Checkpoint 1 Fall 2025 What to Include in Your Memo

Keep your memo around 1 page, single-spaced. Use professional tone and formatting suitable for a stakeholder audience.

Project Title & Topic Summary	Provide a working title and a 1–2 sentence overview of your topic or question.	What is your project about, in plain language?
Problem Statement & Relevance	Describe the problem or question you'll address and why it matters.	Who cares about this problem, and why? What is the real-world sports context or impact?
Intended Dataset (Source & Description)	Identify the dataset(s) you'll use (or plan to obtain). Explain what the data contains and why it fits your question.	What's your data source? What does it include (sport, season, variables)? How will you get it?
Initial Plan or Method	Outline how you might analyze the data at a high level.	What kind of analysis do you envision - statistical comparison, predictive modeling, simulation, visualization?
Next Steps	Describe your immediate plans after submitting the proposal.	What are your next 2-3 actions (obtain data, clean it, refine question, gather feedback)?

*Feel free to use headings or bullet points **within your memo** to organize these sections, but ensure the memo reads as a professional narrative (not just an outline). The tone should be formal yet concise, as if you are writing to a project supervisor or stakeholder interested in your analysis.*

Submission Details

Document Type	Google Doc (professional memo format)
Submission	Share the Google Doc link on Google Classroom
Format Tips	Include your name and project title at the top, as well as any group members

Grading & Rubric (10 points total)

Your Proposal Memo will be evaluated on the following criteria:
DATA 6560 - Sports Analytics - Checkpoint 1 Fall 2025

Clarity of Problem Statement	Clearly defines a focused, specific question or problem that	2

	is easy to understand.	
Feasibility of Topic & Data	Topic is realistic for the semester, and data is suitable or obtainable.	2
Relevance & Significance	Explains why the topic matters and identifies key stakeholders.	2
Structure & Professional Tone	Organized, well-formatted, and written professionally (as a memo).	2
Writing Quality	Clear, concise, grammatically correct, and polished.	2

Additional Notes

You do not need to submit your GitHub repo yet, just ensure you know how to create one for later checkpoints.

Choose a topic you care about; passion often leads to better analysis and presentation later on!!

Connor Project Title: Evaluating NBA Player Efficiency by Salary Tier Date

Subject: Proposal – NBA Player Efficiency by Salary Tier

Project Title & Topic Summary

NBA Player Efficiency by Salary Tier: Which Players Are Underpaid and Overpaid? This project examines NBA player efficiency by salary tier to determine which players are underpaid or overpaid according to performance metrics.

Problem Statement & Relevance

NBA teams have tight salary caps, and there is a great deal of strategy involved in the decision about who stays under contract and at what price. Moreover, very high salaries do not necessarily indicate equivalently strong performances. This project investigates whether teams are efficiently allocating salary by comparing PER or other player performance metrics to cost. The results will help front offices, agents, analysts, and fans by highlighting value players and spots in the roster that are inefficient.

Target Dataset: Source and Description

Source: NBA Player Stats and Salaries Dataset – Kaggle Description: This dataset contains player-level data from recent NBA seasons. It contains many fields that describe each player, including Player Name, Team, Position, Salary, PER (Player Efficiency Rating), Minutes Played, Points, Assists, Rebounds, and Age. The data is publicly available and well-suited for evaluating performance-to-pay relationships.

DATA 6560 - Sports Analytics - Checkpoint 1 Fall 2025 Initial Approach or Procedure

Cleaning: Remove duplicates, standardize player names, and handle missing values.

Feature Engineering: With each tier of salary (e.g., <$5M, $5–15M, >$15M), normalize performance metrics per 36 minutes.

Exploratory Analysis: Visualize PER distribution across salary tiers and detect outliers. Modeling: Use regression to predict PER based on salary and other features; flag players with high residuals. Visualization: Construct Tableau dashboards on efficiency vs salary, highlighting top value and overpaid players. Next Steps Download and examine the dataset from Kaggle. Clean and prepare the data for analysis.