BasketballIndex

title: “NBA Role Evolution Across Eras”

description: Exploring how NBA player scoring styles and roles have changed over the decades using R visualization, modeling, and Shiny apps. categories: - Regression - Clustering - Shiny

## Module Overview

This SCORE module explores how NBA player roles and scoring styles have evolved over time — from the physical post play of the 1990s to today’s pace-and-space, three-point-heavy game. Using real NBA data from four key seasons across four decades (1995–96, 2004–05, 2014–15, and 2023–24), students will analyze trends in player statistics, use statistical modeling, apply clustering techniques, and build an interactive Shiny app to draw meaningful insights about the modern game.

## Learning Objectives

By the end of this module, students will be able to:

* Use dplyr and ggplot2 to summarize and visualize player stats across eras
* Fit and interpret a multiple linear regression model for scoring
* Perform k-means clustering to categorize player roles
* Build an interactive Shiny app for data exploration
* Reflect on how data analysis supports claims about changes in basketball strategy

## Sports Content Question

**How have NBA player roles and scoring styles evolved across the decades, and how can we use data and visual tools to understand that shift?**

## Data

The dataset contains per-game NBA player statistics from the 1996, 2005, 2015, and 2024 seasons. It includes points, assists, rebounds, 3-pointers, and other key metrics. An additional column labels each observation by “Era,” and clustering results are included.

Variable Descriptions

| Variable | Description |
| --- | --- |
| Player | Player name |
| Pos | Position (G, F, C) |
| PTS | Points per game |
| AST | Assists per game |
| TRB | Rebounds per game |
| 3P | 3-pointers made per game |
| 3PA | 3-pointers attempted |
| FG% | Field goal percentage |
| Era | Era label (1990s, 2000s, 2010s, 2020s) |
| Cluster | K-means cluster assignment |

**Data Source**

All statistics were collected from <https://www.basketball-reference.com> and are publicly available for educational use. Raw CSVs were merged, cleaned, and labeled in nba\_all.csv.

## Module Materials

* <nba_all.csv>: Merged dataset
* <worksheet.qmd>: Student worksheet for the module
* <worksheet_key.qmd>: Instructor answer key
* <app.R>: Full interactive Shiny app for exploring player

## Methods

* Students use dplyr and ggplot2 to create boxplots and scatterplots comparing scoring and efficiency across eras.
* A linear regression model is built to predict scoring based on other variables.
* K-means clustering groups players by performance style.
* A Shiny app brings everything together for interactive exploration. :::

## Conclusion

Students will develop insight into how basketball has changed and how to use statistical tools to communicate those changes clearly. This module encourages critical thinking about trends and empowers students to tell stories with sports data.