Understanding NBA Player Success: The Role of Physique, College, and Draft Background

Aditya Kushwaha   Kartik Gupta   Micaela DeLong

5/1/25

Introduction

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

FAIR Principles

1. Findable

- Metadata & Documentation: Ensure datasets are properly documented with clear descriptions ("NBA Player Statistics 2023" and "NBA Player Bios 2022-23").

- Constant Identifiers: Use stable URLs (Kaggle dataset DOI or NBA.com's official stats page).

- Searchable Keywords: Tag datasets with relevant terms ("NBA," "basketball analytics," "player performance").

2. Accessible

- Open Access: The Kaggle dataset is publicly available, while NBA.com's data may have usage restrictions (check terms of service).

- Authentication if Needed: If NBA.com requires API keys, document access procedures.

3. Interoperable

- Standardized Formats: Use CSV/JSON for compatibility.

- Common Variables: Merge datasets using player IDs/names for cross-analysis (e.g., linking stats to bios).

- Clear Units: Define metrics (height in cm/ft, 3P% as a decimal).

4. Reusable

- Licensing: Kaggle datasets often have CC licenses; NBA.com data may require attribution.

- Methodology Transparency: Document how performance metrics ("best players") are calculated.

CARE Principles

1. Collective Benefit

- Community Insights: Share findings to benefit teams, players, and fans (how draft position correlates with success).

- Avoid Harm: Ensure analysis doesn't reinforce biases (stereotyping players by physical attributes).

2. Authority to Control

- Player Consent: Acknowledge that player data is public but avoid misuse (invasive predictions about health).

- Respect NBA Policies: Follow league guidelines on data usage.

3. Responsibility

- Bias Mitigation: Check if conclusions favor certain demographics (does "best schools" overlook international players?).

- Transparency: Disclose limitations (stats don't capture leadership intangibles).

4. Ethics

- Privacy: Anonymization isn't needed for public stats, but avoid combining data to infer private details.

- Fair Representation: Ensure diverse player backgrounds (country, college) are analyzed equitably.

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).