Fritz Kussow, Matt Damsma, Timothy Gilmore, Alonso Carrasco

# The Indiana Pacers 2025 Play-off Run: Discovering Trends to Predict Future Game Outcomes

# Introduction

- The Goal: Predict the outcome (Win/Loss) of Indiana Pacers games based on the *composition* of the teams playing, rather than just team-level averages.
- Box score data from each game
- Base on 2024-25 Indiana Pacers season

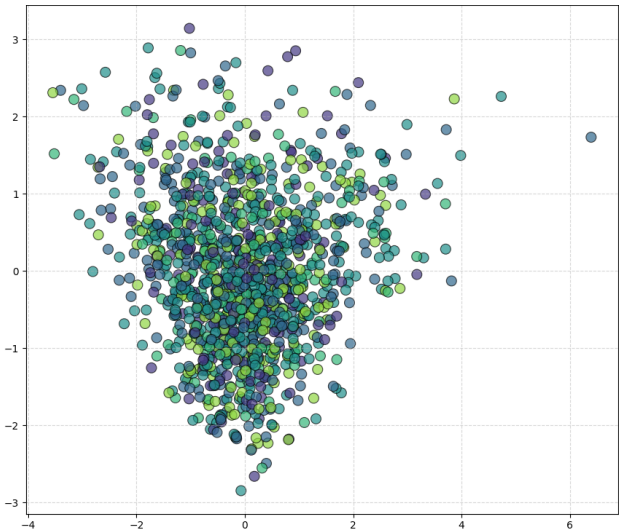
# Dataset

- The dataset came from the nba_api
- Data includes each individual players stats
- We standardized data to be per 36min
- Preprocessing Steps:
  - Filtering: Excluded players with < 10 minutes played to remove "garbage time" noise.
  - Normalization: Converted all key stats (Points, Rebounds, Assists, etc.) to "Per 36 Minutes" to compare starters and bench players on an equal footing.
- Dataset:
  - Scope: All Pacers games from the 2024-2025 Season (Regular Season + Playoffs).
  - Volume: 105 games total, ~2,890 player-game records.

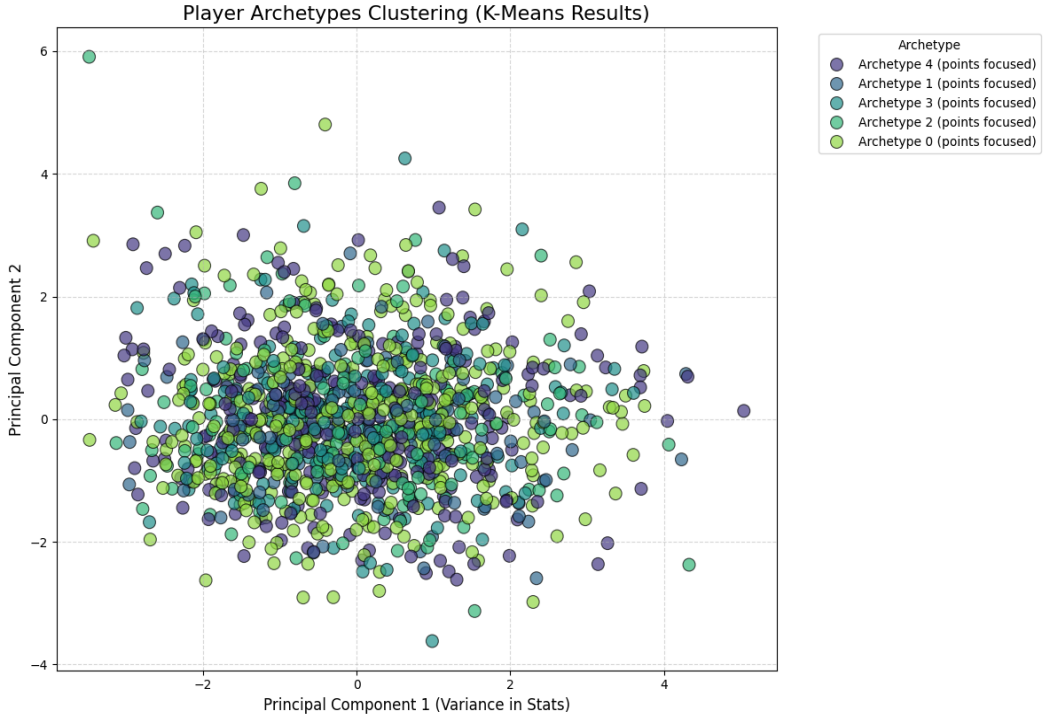
# Feature Engineering: Player Archetypes Labels

- We tried several different combinations of stats in order to categorize the players into five archetypes
- Using these stats, we would be able to see what clusters form and categorize players that way

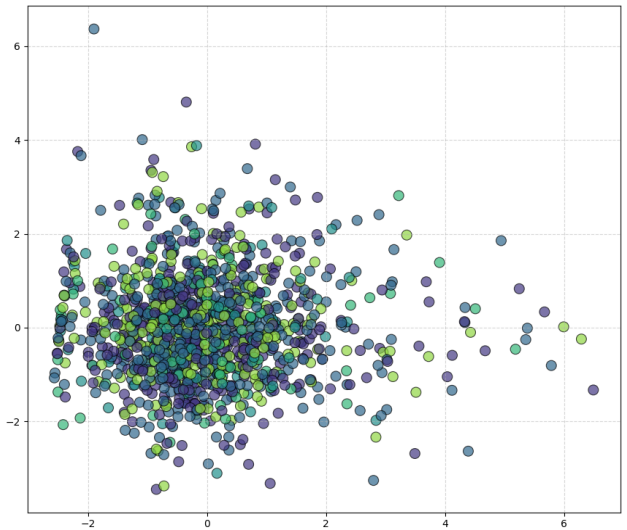
Pts, reb, ast, blk, stl



Selected: pts, reb, ast, stl, blk, to, 3pm, fga




Pts, reb, ast, blk, stl, 3P%, fg%



# Testing

- Used 80% of the games for training, 20% for testing
- K-NN Classification Accuracy = 0.62
- Naive Bayes accuracy = 0.57
- Logistic Regression accuracy = 0.52

K-NN Classification

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.00 | 0.00 | 0.00 | 6 |
| 1 | 0.68 | 0.87 | 0.76 | 15 |
| accuracy | | | 0.62 | 21 |
| macro avg | 0.34 | 0.43 | 0.38 | 21 |
| weighted avg | 0.49 | 0.62 | 0.55 | 21 |

Naïve Bayes

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.20 | 0.17 | 0.18 | 6 |
| 1 | 0.69 | 0.73 | 0.71 | 15 |
| accuracy | | | 0.57 | 21 |
| macro avg | 0.44 | 0.45 | 0.45 | 21 |
| weighted avg | 0.55 | 0.57 | 0.56 | 21 |

Logistic Regression

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.17 | 0.17 | 0.17 | 6 |
| 1 | 0.67 | 0.67 | 0.67 | 15 |
| accuracy | | | 0.52 | 21 |
| macro avg | 0.42 | 0.42 | 0.42 | 21 |
| weighted avg | 0.52 | 0.52 | 0.52 | 21 |