

Introduction to NextPlay

Predicting NBA Success with Data



Siegfried Nacion
Data Scientist, Brooklyn College
SiegfriedAnthonyNacion@gmail.com

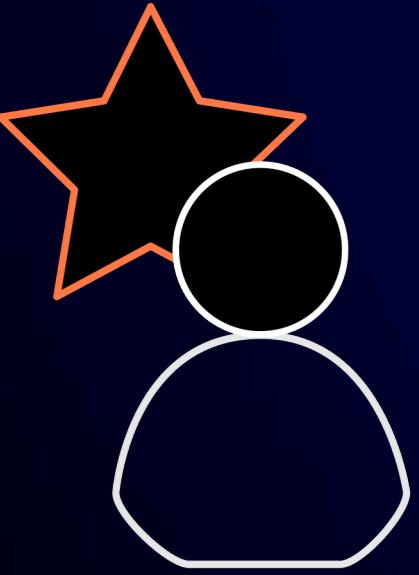
Agenda

- Project Introduction
- Customer Identification & Questions (PR-FAQ Process)
- Workflow Outline
- Scraping
- EDA Findings
- Machine Learning Breakdown
- Product Demo & Reflection

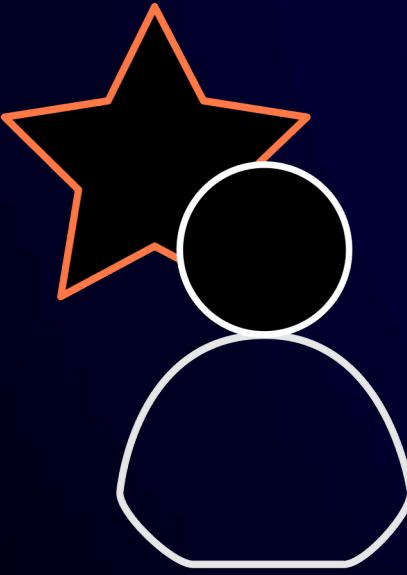
NextPlay's Mission: To have
an accurate outlook on the
upcoming generation of the
NBA

Customer Questions

1. Who is the customer?
2. What is the customer's problem?
3. What are/is the customer benefit[s]?



Who are my
customers?



- NBA front offices
- ESPN analysts
- Sportsbook companies and players

What is the customer problem or opportunity?

Today NBA front offices, ESPN Analysts, Sportsbook Players

have to actively scout NBA prospects in college

when they want to instantly and accurately see how well a prospect is going to play when they get drafted



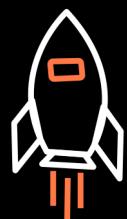
What is/are the customer benefit[s]?

Customer will predict a college player's first four years in the NBA

Fictitious Customer Testimonial

“As the Knicks’ President, a majority of our draft picks end up not performing as well as we’d like them to,” said Leon Rose. “Our front offices have been mainly using a mix of analytical and scouting methods to decide who we draft on a yearly basis. With NextPlay, I can see a college player’s initial career trajectory in a click of a button. This way, our front office is able to more efficiently focus their time on how to draft accordingly.”

Project Diagram



—

Scraping

Data
Source:



Scraping
module:

BeautifulSoup

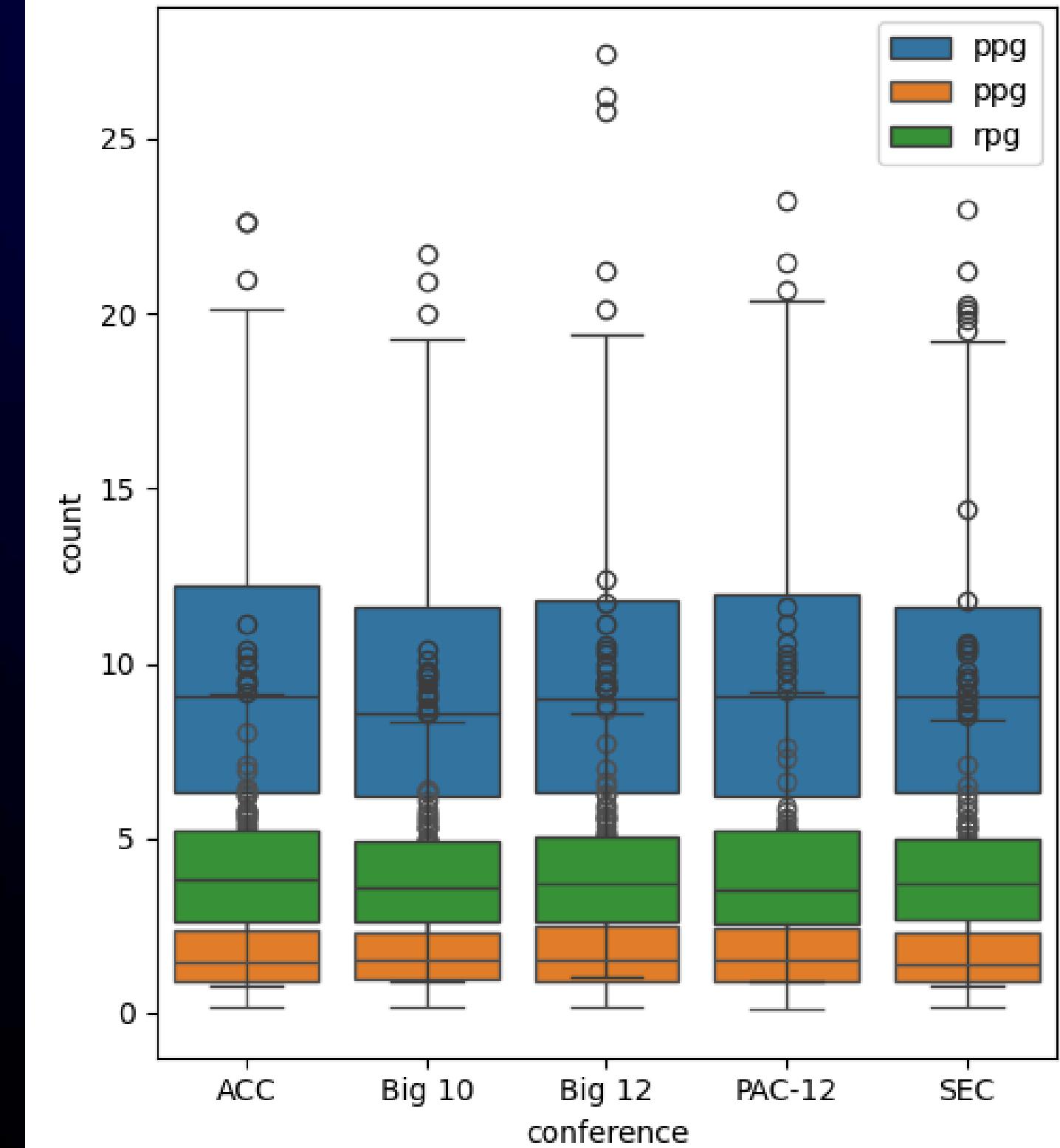
EDA Questions

- What are the average stats for players in the top 5 conferences?
- Are there differences in stats across conferences (e.g., average points, assists)?
- How many players transitioned to the NBA from each conference?

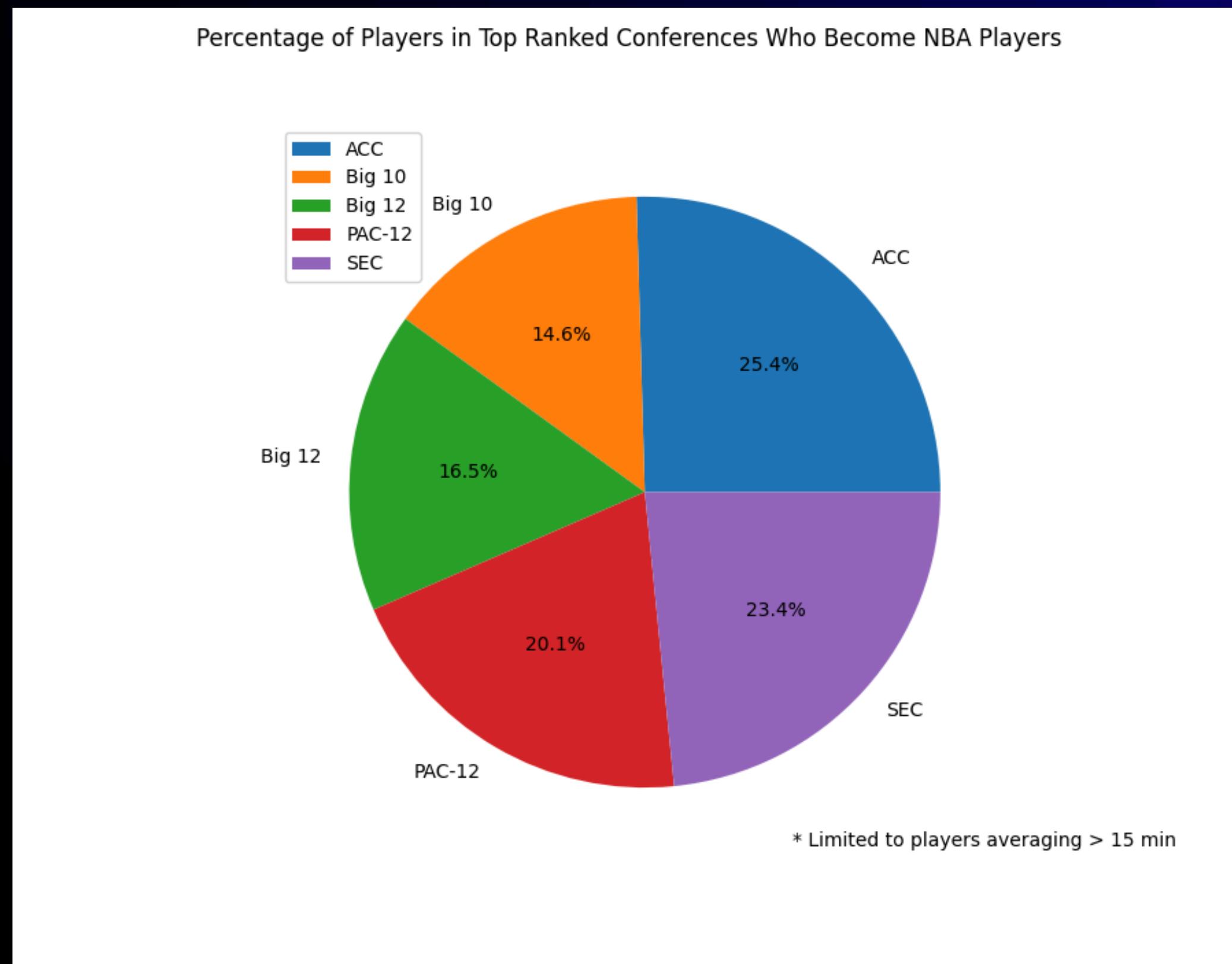
EDA Findings

	conference	pts	ast	trb
0	ACC	9.450163	1.789009	4.062359
1	Big 10	9.102105	1.797980	3.903300
2	Big 12	9.360966	1.848551	4.035944
3	PAC-12	9.328380	1.805636	3.973815
4	SEC	9.379231	1.736237	4.005633

Summary Statistics Across NCAA Conferences

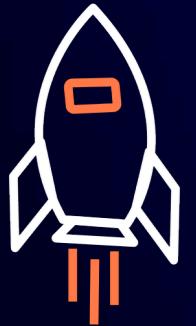


EDA Findings



EDA Findings

- No big difference in statistics between each conference
- All 5 conferences have a tendency to score over assist and rebound
- SEC, PAC-12, ACC favorites for NBA prospects

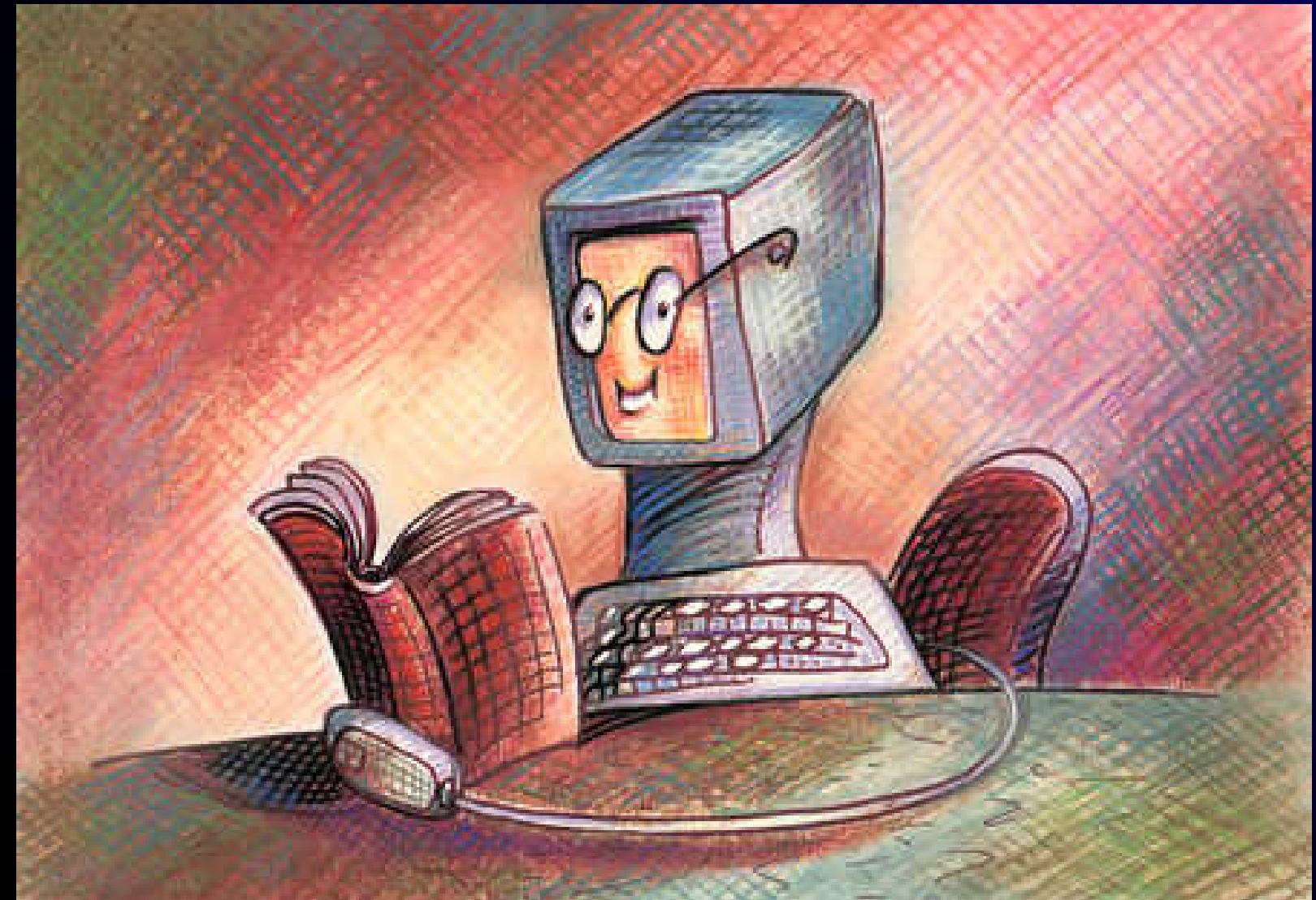


Model Creation

What is Machine Learning?

Per Google:

“The use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.”



Teaching a machine how to think using examples instead of rules.

How NextPlay Uses Machine Learning

- Input: College basketball stats like points, assists, rebounds, free throws, steals, blocks, etc.
- Output:
 - 1) Predicted NBA stats for the first 4 years, including points, assists, and rebounds
 - 2) Impact score that ranks players into tiers (Elite, High Impact, Medium Impact , Low Impact).

NBA Statistic Prediction Tool

- Features: Statistics in college
- Target Variables: Predicted statistics in the NBA
- RandomForestRegressor:
 - Handles non-linear relationships and complex interactions between features.
 - Robust to outliers and overfitting
- MultiOutputRegressor:
 - Allows predicting multiple NBA stats (e.g., points, rebounds, assists) simultaneously.

Impact Classification

Low Impact

AJ Green
Zabian Dowdell
CJ Wilcox
Yuta Watanabe

Medium Impact

Bojan Bogdanović
Yogi Ferrell
Aaron Brooks
Wayne Ellington

High Impact

Aaron Gordon
Zach Lavine
Al Jefferson
Willie Cauley-Stein

Elite

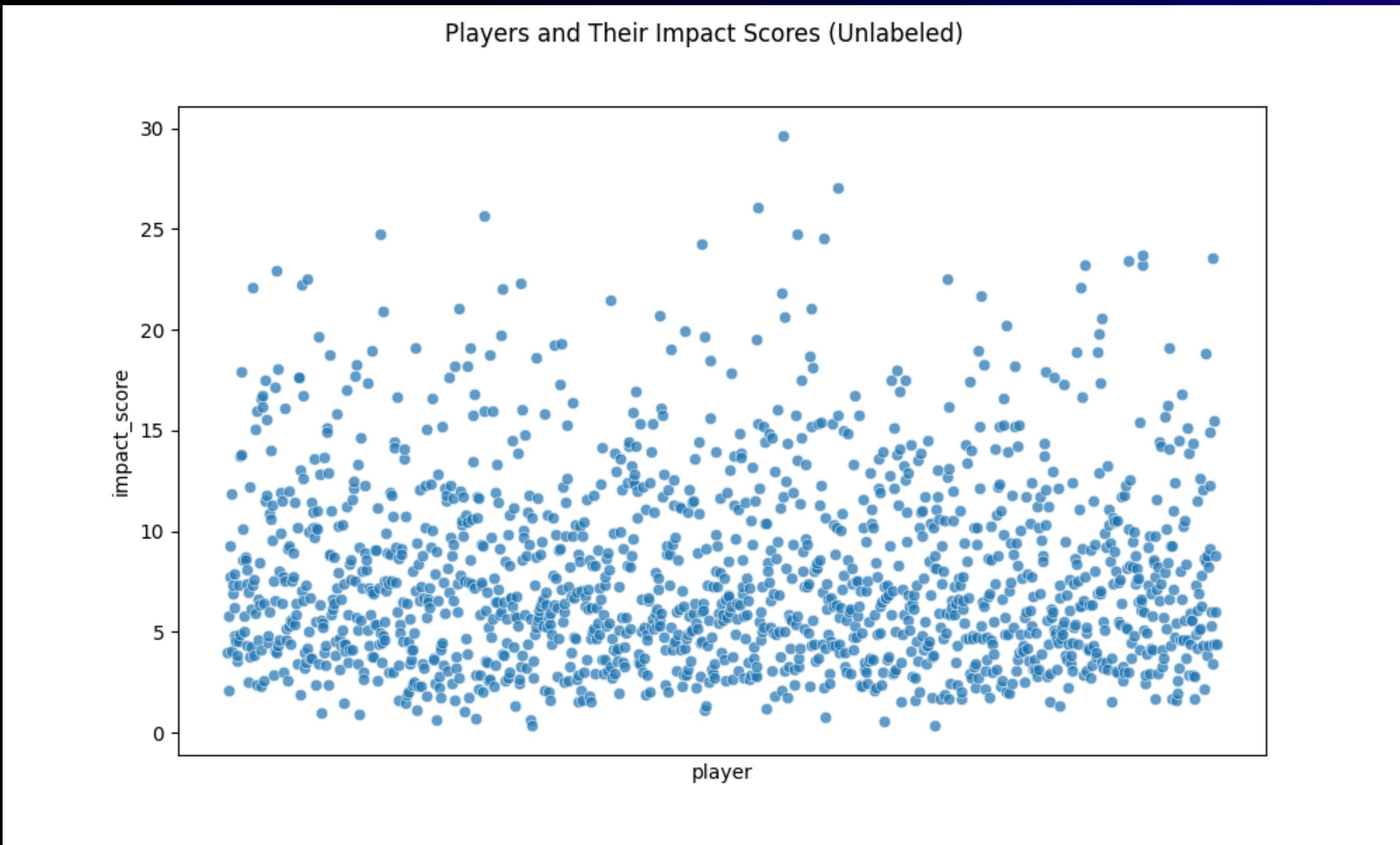
LeBron James
Donovan Mitchell
Kyrie Irving
Carmelo Anthony

Impact Classification Formula

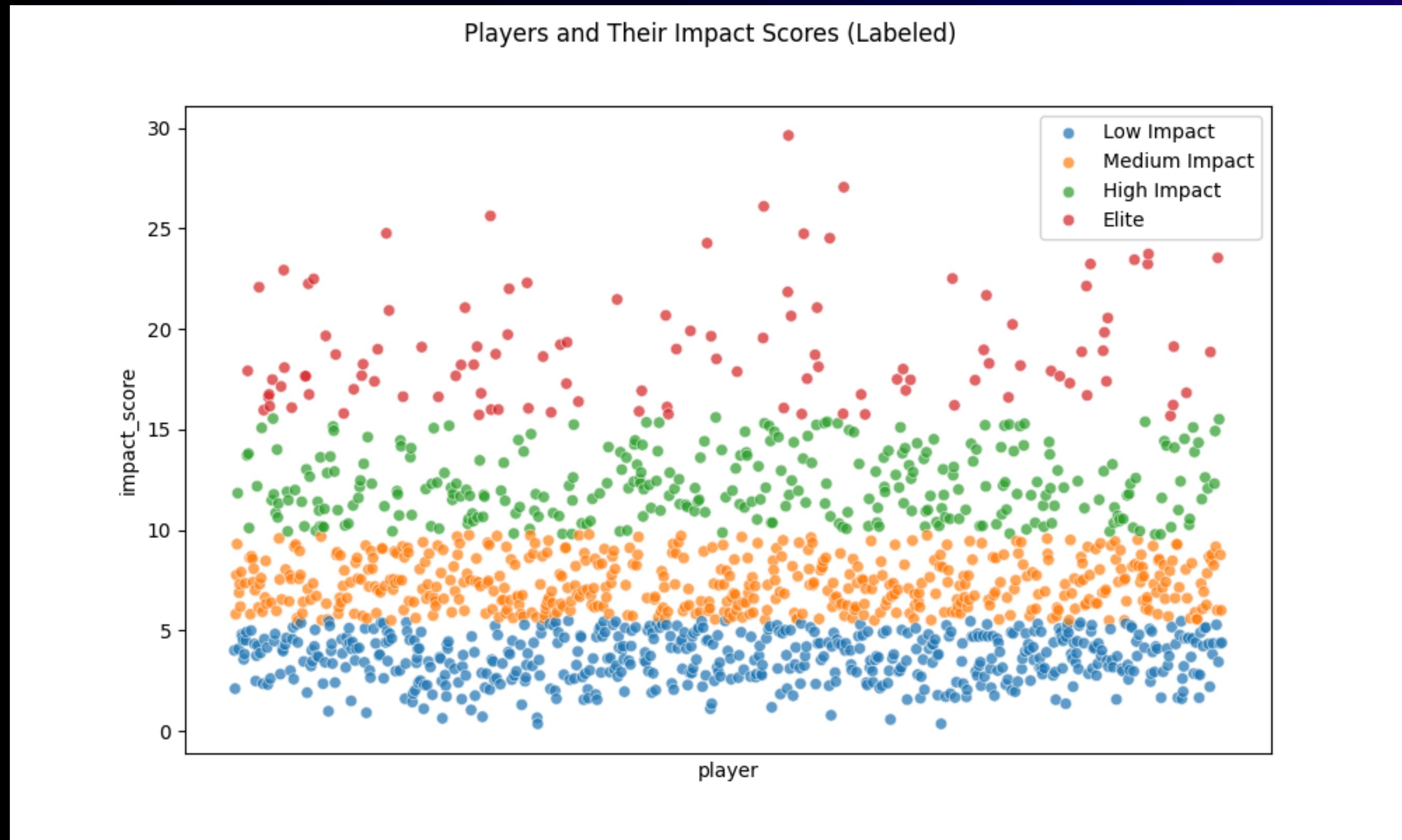
Impact Score =

$$\text{PPG} + \text{APG} + \text{DRB} + (0.5 \times \text{ORB}) + \text{FG} + \text{FT} + \text{STL} + (0.5 \times \text{BLK}) - \text{TOV} - \text{FGA} - \text{FTA}$$

Impact Classification



Impact Classification



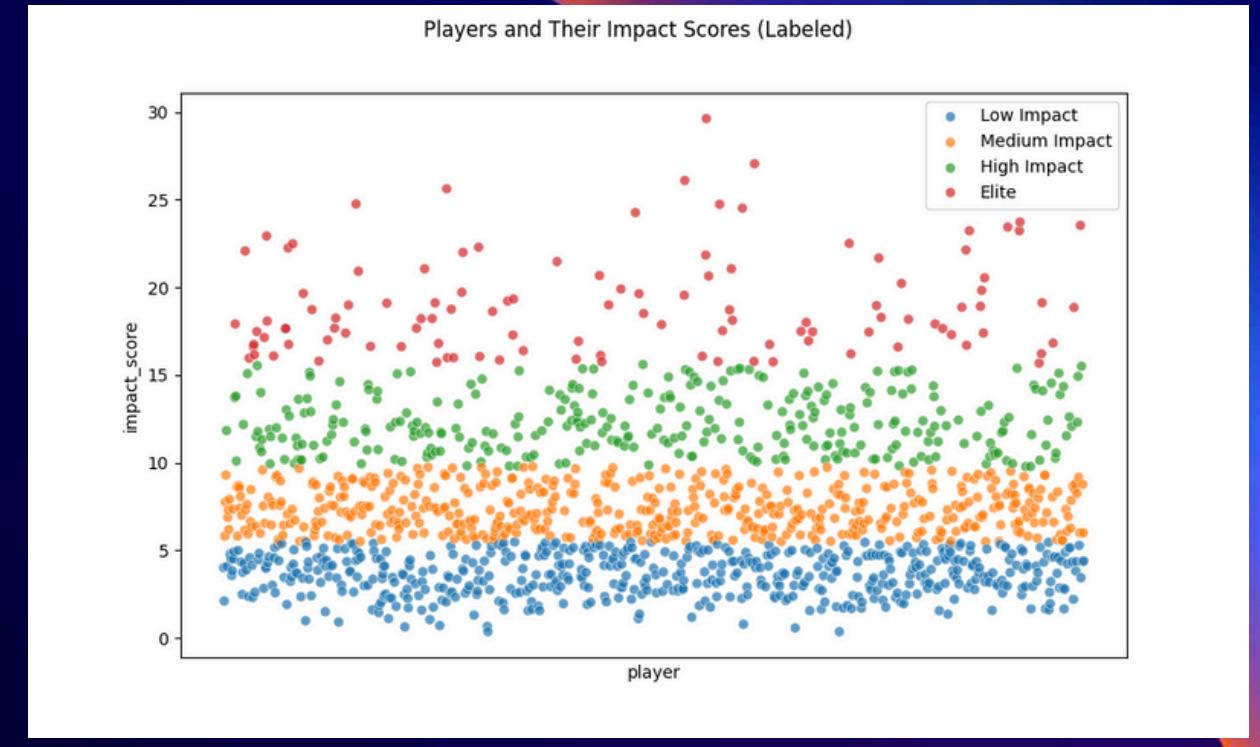
Impact Classification

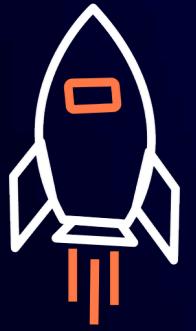
Elite: Impact score > 15.6

High Impact: $9.77 < \text{Impact Score} < 15.6$

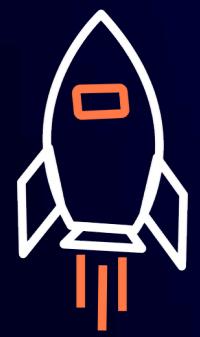
Medium Impact: $5.49 < \text{Impact Score} < 9.77$

Low Impact: Impact score < 5.49





NextPlay
Demo



Reflection

Introduction to NextPlay

Predicting NBA Success with Data



Siegfried Nacion
Data Scientist, Brooklyn College
SiegfriedAnthonyNacion@gmail.com