Final Report: Predicting SCU Player Potential & NBA Draft Prospects

Jonah Engelmann, Raul Sharma, Luo Long, Arnay Konijeti

1. Introduction

Scouting and player development are crucial for identifying talent and maximizing player potential in basketball. Players who succeed in college often demonstrate consistent efficiency in scoring, defense, and overall performance. Additionally, players who get drafted into the NBA are said to possess standout athletic traits, such as elite shooting ability, vertical jump, and wingspan.

SCU has had only a few players make it to the NBA. This study develops a data-driven approach to analyze SCU player potential. It will analyze SCU men's players' performance and athletic metrics from two different angles. It will first examine if the players have the potential to be "good" college players by statistically comparing them to a manually created player with subjectively good stats. It will then predict if past NBA combine players (a player essentially trying out for the NBA) will get drafted, and then use those metrics to predict if SCU's players will be drafted.

2. Data Description

2.1 Dataset

3 datasets included:

- SCU roster performance metrics (PPG, FG%, BLK/G, STL/G, etc.)
- Physical metrics (height, weight, wingspan, vertical jump, speed, etc.)
    - SCU players
    - NBA draft combine players

2.2 Data Collection & Feature Engineering

- SCU Data was collected through manual inputs and calculations.
    - Additional stats like blocks per game and steals per game were computed.
    - **Feature Engineering**:
        - **Wingspan-to-height ratio** to analyze defensive impact.
        - **Blocks per game = total blocks / games played**.

    - **Steals per game = total steals / games played**.
- NBA combine player data was taken from [this public repository on NBA website]

2.3 Data Processing

- SCU Data:
  - Missing values were handled.
  - Measurements were standardized to ensure consistency.
- NBA combine data:
  - Converted string numeric metrics (such as height represented as 6’5”) to integers
  - Categorized all players as guard of forwards
  - Created boolean stat for each player is_drafted (indicating if they were drafted)
  - Dropped players who didn’t have data for 5+ metrics
  - Created 2 more clone tables with the target balanced by SNOTE

3. Experimental Design (Methodology)

3.1 Analysis Approach

- **Identifying "Good" College Players**
  - Criteria: Players who score efficiently and contribute across multiple categories.
  - Key Metrics:
    - PPG (Points Per Game)
    - FG% (Shooting Efficiency)
    - R/G (Rebounds Per Game)
    - STL/G + BLK/G (Defensive Impact)
  - **Analysis:**
    - Scatter plots for performance trends.
    - Higher FG% & scoring = potential key player.
    - High STL/G, BLK/G = elite defensive contributors.
- **Identifying NBA Draft Potential**
  - Criteria: Players with exceptional athletic traits aligning with NBA profiles.
  - Key Metrics:
    - Vertical Jump (explosiveness)
    - Wingspan (defensive impact)
    - Weight (physicality)
  - **Analysis:**
    - Predict if past NBA draft candidates got drafted.
    - Use the computed metrics to see if the model predicts our players will get drafted.

3.2 Models

a) **Who are our current good players and is anyone up next?**

i) **Modified Logistic Regression Model**
   1) Uses key features: PPG, FG%, R/G, STL/G, BLK/G, Wingspan-to-Height Ratio, Vertical Jump, and 3/4 Court Sprint Speed.
   2) Globally estimates the probability of a player becoming a star based on these metrics.

ii) **Decision Tree Model**
   1) Trained using historical SCU players' data, incorporating:
      (a) PPG, FG%, STL, BLK
      (b) Height, Wingspan, Vertical Jump, and Sprint Speed
   2) Players are categorized into quartiles based on PPG, as offensive dominance is a key factor.
   3) A Decision Tree classifier (max_depth=6, min_samples_split=3) is used for classification.
   4) The model defines thresholds that differentiate **Elite Players, Role Players, and Developmental Prospects**.

Decision Tree Examples

- **Player A - Predicted as "Elite"**
  - **Stats:**
    - Wingspan: 85 inches ✅ (Long reach, great for defense & finishing)
    - FG%: 52% ✅ (Highly efficient scorer)
    - PTS/G: 18.5 ✅ (High offensive contribution)
    - Vertical Jump: 37 inches ✅ (Strong athleticism, good for rebounds & dunks)
    - Sprint Speed: 3.1 sec ✅ (Fast mover on the court)
  - **Why "Elite"?**
    - The model prioritizes **high PTS/G and FG%** as indicators of offensive dominance.
    - Athletic ability (Vertical Jump, Sprint Speed) further strengthens the classification.
    - Elite players in past data had similar attributes—long wingspan, efficient shooting, and strong vertical leap.
- **Player B - Predicted as "Role Player"**
  - **Stats:**
    - Wingspan: 78 inches ❌ (Average, not a major advantage)
    - FG%: 44% ✅ (Decent, but not elite)
    - PTS/G: 8.2 ❌ (Not a primary scorer)
    - STL: 2.5 ✅ (Strong defensive presence)
    - BLK: 1.8 ✅ (Solid rim protection, good help defender)
  - **Why "Role Player"?**
    - The model identified **defensive strengths (high steals & blocks).**

- Lower scoring but solid efficiency (FG%) suggests a **specialist role rather than a star.**
- Past players in this category were effective in team roles but not primary scorers.

b) Are any of our players NBA ready or have NBA potential?

i) **Logistic Regression Model** for NBA combine players

- Features:
  - HAND LENGTH (inches), HAND WIDTH (inches), HEIGHT W/O SHOES, STANDING REACH, WEIGHT (LBS), WINGSPAN, Lane Agility Time seconds), Three Quarter Sprint (seconds), Standing Vertical Leap (inches), Max Vertical Leap (inches)Target: is_drafted
- Target:
  - Is_drafted
- Predicts if a player will drafted from these metrics
- 2 models were made, one for forwards one for guards

ii) **Logistic Regression Model** for NBA combine players for SCU's metrics

- Features:
  - Height w/o Shoes, Standing Reach, Weight (lbs), Wingspan, 3/4 Court Sprint, Vertical Jump, Run Jump
- Target:
  - Is_drafted
- Predicts if a player will drafted from these reduced metrics
- Theta/Scalars from the result will be used to predict if SCU players will get drafted

4. Implementation (Analysis & Visualizations)

4.1 Tools Used

- **Python** (primary programming language)
- **Pandas & NumPy** (data manipulation)
- **Scikit-learn** (model training and evaluation)
- **Matplotlib & Seaborn** (data visualization)
- **Dash** (interactive dashboard development)
- **Plotly** (visualization components including histograms, comparison plots, and the ability to hover over points to see player details)
- **Logistic regression model** & Decision tree classification model
- **Correlation Analysis** - Applied to identify which physical metrics (wingspan, hand measurements, vertical leap, etc.) most strongly predict draft success for different position groups.

4.2 Key Visualizations

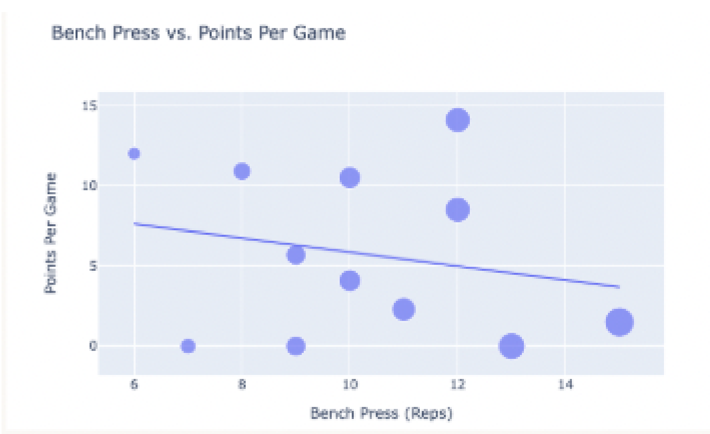
- **Scatter plot**: Height vs. PPG (Do taller players score more?)
- **Bubble Chart**: Vertical Jump vs. Rebounds (Do explosive players rebound better?)
- **Histogram**: PTS/G distribution (Who are the most efficient scorers?)
- **Wingspan-to-Height Ratio vs. Rebounds** (Do longer players grab more rebounds?)
- **Bench Press vs. Strength Impact** (Does upper body strength matter?)

- **Guards Correlation Analysis -** Identifies which physical attributes most strongly predict draft status for guards, both positively and negatively.
- **Forwards Correlation Analysis -** Determines the most significant physical measurements that correlate with draft success specifically for forwards.
- **Position Group Comparison Dashboard** - Visualizes the key differences between guards and forwards across all metrics, highlighting how draft success factors vary by position.

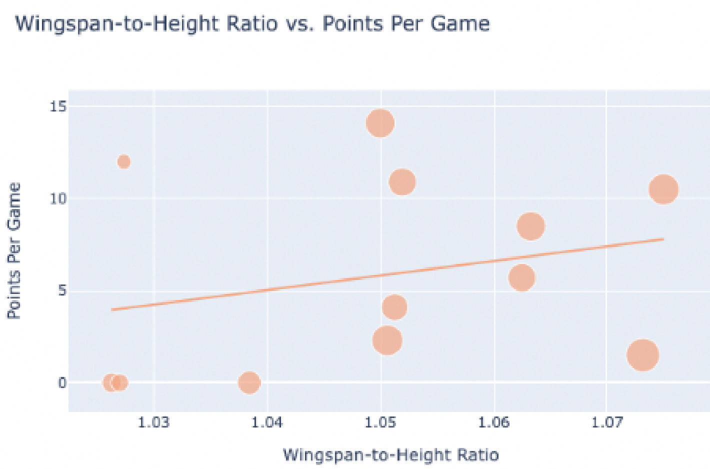
4.3 Findings

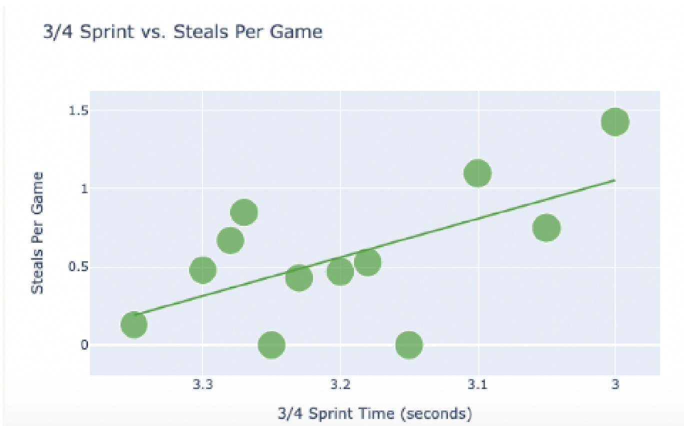
- Players with **high vertical jumps** tend to get more rebounds & blocks.
- **Wingspan ratio strongly correlates with points per game, rebounds per game, and blocks per game**.
- **FG% is a better predictor of success than raw athleticism**.
- **Speed** is a good predictor of how well a player can defend and their **steals per game**
- Some metrics such as bench press are not relevant

Visuals:

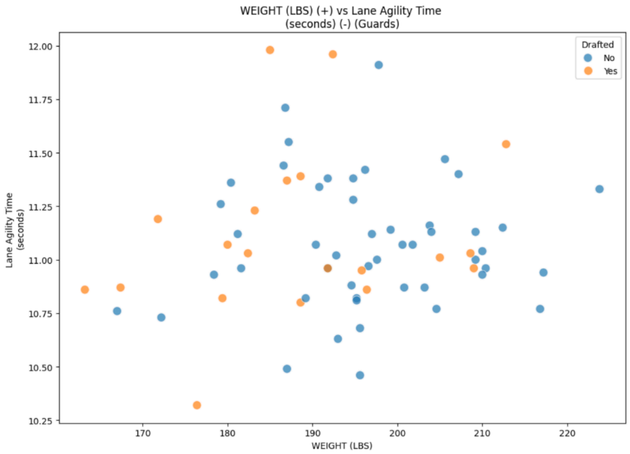


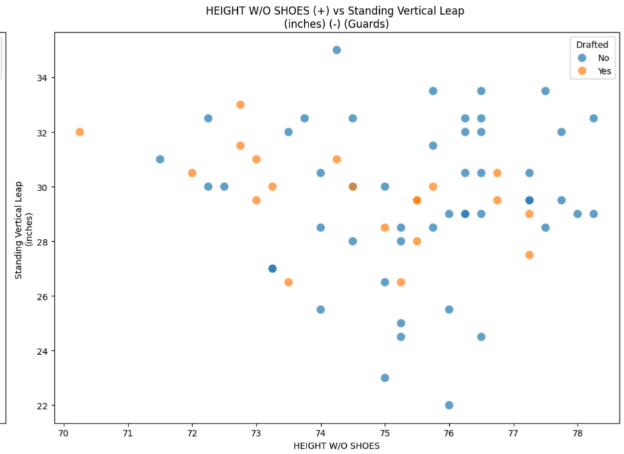

Scoring Performance
Shooting Efficiency
Points Distribution
Physical Metrics vs. Performance
Scoring Performance
This chart shows how many points each player scores per game, with color coding to indicate scoring efficiency (Field Goal %).
Scoring Performance (PPG vs Efficiency)
Points Per Game
0
5
10

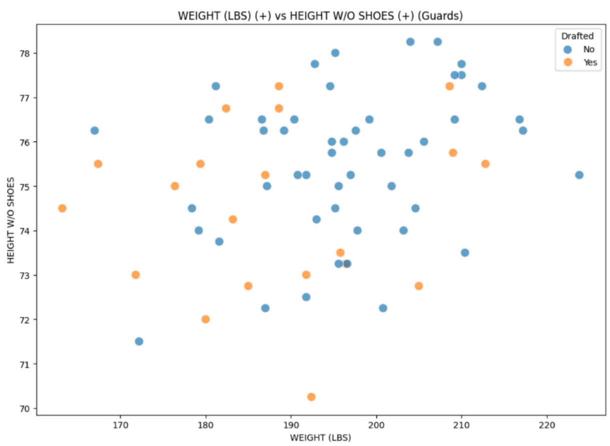

Wingspan-to-Height Ratio vs. Points Per Game
Points Per Game
0
5
10
15
1.03
1.04
1.05
1.06
1.07
Wingspan-to-Height Ratio

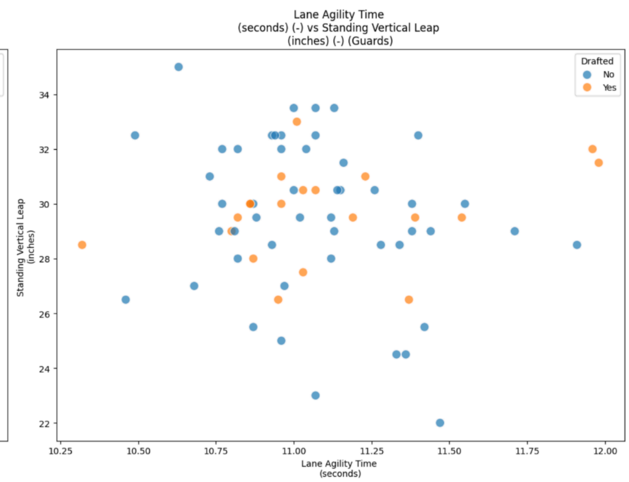

3/4 Sprint vs. Steals Per Game
Steals Per Game
0
0.5
1
1.5
3.3
3.2
3.1
3
3/4 Sprint Time (seconds)

Guards:

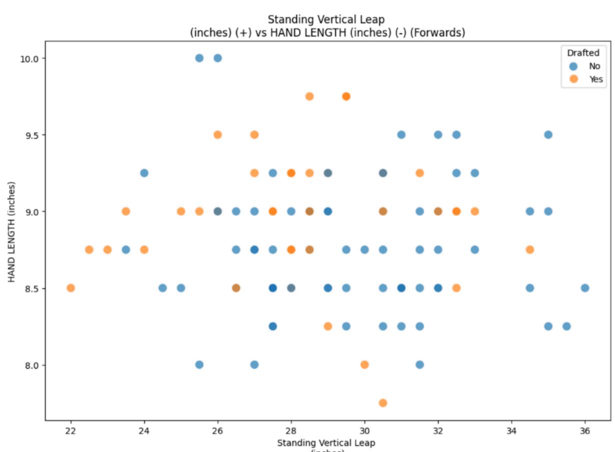

WEIGHT (LBS) (+) vs Lane Agility Time
(seconds) (-) (Guards)
Lane Agility Time
(seconds)
WEIGHT (LBS)
Drafted
No
Yes

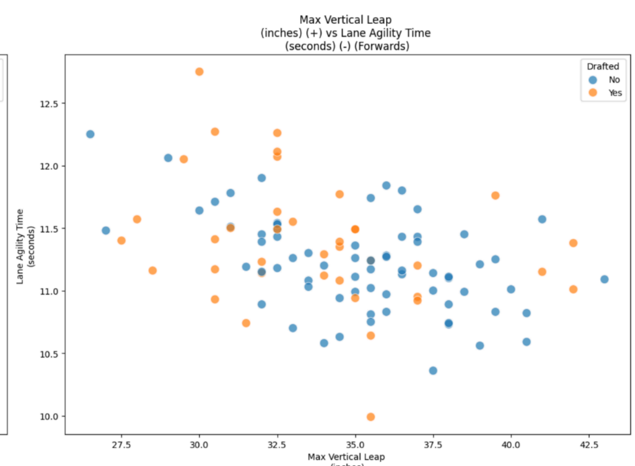

HEIGHT W/O SHOES (+) vs Standing Vertical Leap
(inches) (-) (Guards)
Standing Vertical Leap
(inches)
HEIGHT W/O SHOES
Drafted
No
Yes

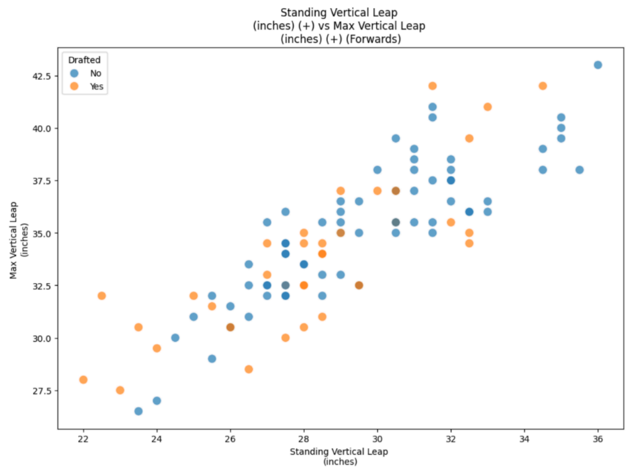

WEIGHT (LBS) (+) vs HEIGHT W/O SHOES (+) (Guards)
HEIGHT W/O SHOES
WEIGHT (LBS)
Drafted
No
Yes


Lane Agility Time
(seconds) (-) vs Standing Vertical Leap
(inches) (-) (Guards)
Standing Vertical Leap
(inches)
Lane Agility Time
(seconds)
Drafted
No
Yes

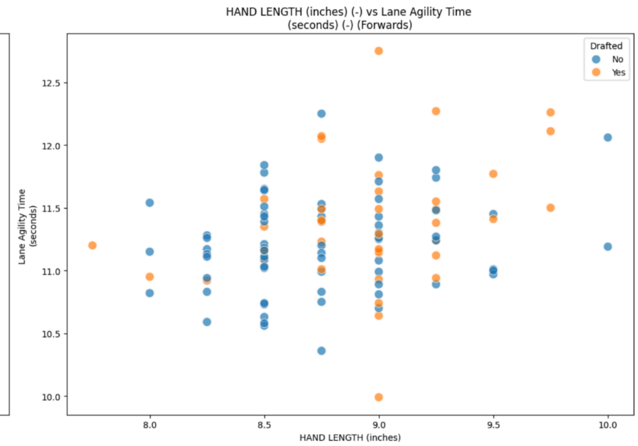
Forwards:


Standing Vertical Leap
(inches) (+) vs HAND LENGTH (inches) (-) (Forwards)
HAND LENGTH (inches)
Standing Vertical Leap
(inches)
Drafted
No
Yes

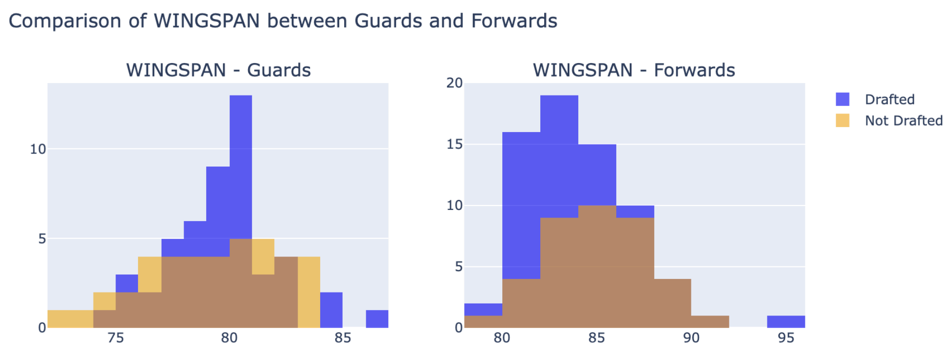

Max Vertical Leap
(inches) (+) vs Lane Agility Time
(seconds) (-) (Forwards)
Lane Agility Time
(seconds)
Max Vertical Leap
(inches)
Drafted
No
Yes



Standing Vertical Leap
(inches) (+) vs Max Vertical Leap
(inches) (+) (Forwards)
Drafted
No
Yes
Max Vertical Leap
(inches)
Standing Vertical Leap
(inches)

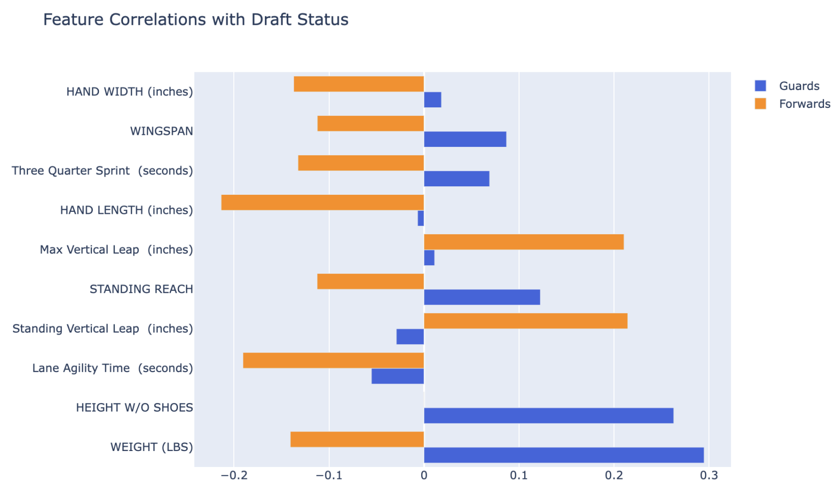


HAND LENGTH (inches) (-) vs Lane Agility Time
(seconds) (-) (Forwards)
Drafted
No
Yes
Lane Agility Time
(seconds)
HAND LENGTH (inches)



Comparison of WINGSPAN between Guards and Forwards
WINGSPAN - Guards
WINGSPAN - Forwards
Drafted
Not Drafted

WINGSPAN vs HEIGHT W/O SHOES - Guards and Forwards
Guards - Not Drafted
Guards - Drafted
Forwards - Not Drafted
Forwards - Drafted
HEIGHT W/O SHOES
WINGSPAN



What the visualizations tell us:

- For guards, size metrics are the strongest draft predictors, with weight and height showing the highest positive correlations, suggesting NBA teams prioritize larger guards who can defend multiple positions and finish through contact.
- For forwards, athletic performance metrics matter most, with vertical leap measurements showing the strongest positive correlations while hand length and agility time are negatively correlated, indicating explosive jumping ability is more valued than size metrics at these positions.
- There's a striking positional contrast in how metrics predict draft success, with many measurements (like weight and height) that positively correlate for guards actually showing negative correlations for forwards, revealing position-specific evaluation criteria in NBA talent assessment.

5. Results (Findings & Insights)

5.1 SCU Player Potential

- Players with **high FG% and scoring numbers** are already contributing significantly.
- **Strong defenders** have high STL/G + BLK/G values.
- Some players show **steady improvement over time**, indicating strong future potential.
- **Prediction**: Tyeree Bryan was projected as SCU's next star player due to his physical and performance improvements.

Prediction CSV File:

Player_Predictions

Player	Good_Player_Label	Predicted_Probability	Predicted_Label	Height w/ Shoes	Wingspan	Vertical Jump	PTS/G	FG%
Kosy Akametu	0	0.42772335181944200	0	77.25	80.25	32.5	4.1	0.565
Carlos Stewart	1	0.6604076100965940	1	72.5	78.5	34.5	12.0	0.448
Jacob Ensminger	0	0.3157868031801310	0	81.0	78.0	27.0	2.3	0.279
Adama Bal	1	0.5719675080653320	1	78.5	82.0	30.0	14.1	0.424
Tyeree Bryan	0	0.6056123433033800	1	77.25	80.0	31.0	10.9	0.475
Brenton Knapper	0	0.42522556052992200	0	71.25	73.5	28.5	4.0	0.451
Elijah Mahi	1	0.5739031181890340	1	78.25	82.0	31.0	11.6	0.497
Johnny O'Neil	0	0.4722522788446850	0	82.75	82.5	26.0	6.7	0.398
Bukky Oboye	0	0.3615549398722490	0	85.0	88.0	32.0	0.8	0.4
Christoph Tilly	1	0.7388837075658270	1	84.5	111.0	24.0	13.2	0.576
Cam Tongue	0	0.47581819809424200	0	78.25	82.0	27.5	6.2	0.478
Allen Graves	0	0.2489771229062020	0	80.75	84.0	25.5		
Christian Hammond	0	0.22530679452283700	0	75.5	75.5	24.0		
Synthetic Star	1	0.5699282434951180	1			35.0	13.5	0.48
Synthetic Bench	0	0.36780965543099600	0			28.0	8.2	0.4

Decision Tree Result:

```
=== Predicted Player Levels ===
              Player Predicted_Level
0       Kosy Akametu           Bench
1     Carlos Stewart     Role Player
2    Jacob Ensminger           Bench
3          Adama Bal           Elite
4       Tyeree Bryan     Role Player
5    Brenton Knapper           Bench
6        Elijah Mahi     Role Player
7       Luke McEldon           Bench
8      Johnny O'Neil  Potential Star
9        Bukky Oboye           Bench
10   Christoph Tilly           Elite
11        Cam Tongue           Bench
12      Allen Graves           Bench
13 Christian Hammond           Bench
```

5.2 NBA Draft Potential

- Full report of Logistics Regression's ability to predict if a combine player will be drafted for guards, guards with SNOTE, forwards, and forwards with SNOTE:

```
Results for df_G:
Accuracy: 0.6190
Classification Report:
              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

--------------------------------------------------
Results for df_G_balanced:
Accuracy: 0.6333
Classification Report:
              precision    recall  f1-score   support

           0       0.64      0.60      0.62        15
           1       0.62      0.67      0.65        15

    accuracy                           0.63        30
   macro avg       0.63      0.63      0.63        30
weighted avg       0.63      0.63      0.63        30

--------------------------------------------------
Results for df_F:
Accuracy: 0.5625
Classification Report:
              precision    recall  f1-score   support

           0       0.20      0.09      0.12        11
           1       0.63      0.81      0.71        21

    accuracy                           0.56        32
   macro avg       0.41      0.45      0.42        32
weighted avg       0.48      0.56      0.51        32

--------------------------------------------------
Results for df_F_balanced:
Accuracy: 0.6098
Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.67      0.64        21
           1       0.61      0.55      0.58        20

    accuracy                           0.61        41
   macro avg       0.61      0.61      0.61        41
weighted avg       0.61      0.61      0.61        41

--------------------------------------------------
```

  - SNOTE models all performed better
  - Mediocre accuracies and F-1 scores
  - High recalls for models without SNOTE (likely due to imbalanced target)
- Prediction if current SCU roster will be drafted:

```
Player: Kosy Akametu, Position: Guard, Predicted Draft Status: Drafted
Player: Carlos Stewart, Position: Guard, Predicted Draft Status: Not Drafted
Player: Jacob Ensminger, Position: Guard, Predicted Draft Status: Drafted
Player: Adama Bal, Position: Guard, Predicted Draft Status: Drafted
Player: Brenton Knapper, Position: Guard, Predicted Draft Status: Not Drafted
Player: Elijah Mahi, Position: Guard, Predicted Draft Status: Drafted
Player: Christian Hammond, Position: Guard, Predicted Draft Status: Drafted
Player: Johnny O'Neil, Position: Forward, Predicted Draft Status: Drafted
Player: Bukky Oboye, Position: Forward, Predicted Draft Status: Drafted
Player: Christoph Tilly, Position: Forward, Predicted Draft Status: Not Drafted
Player: Cam Tongue, Position: Forward, Predicted Draft Status: Not Drafted
Player: Allen Graves, Position: Forward, Predicted Draft Status: Not Drafted
```

  - 7 of our 12 players predicted to be drafted
    - 5 of the 7 guards
    - 2 of the 5 forwards

6. Interpretation (Discussion & Takeaways)

6.1 Player potential Takeaways

- The best SCU players may become key college contributors, but **NBA potential is low**.
- More **skill-focused players (high FG%, shooting efficiency)** tend to outperform purely athletic players.
- **Adding Catapult wearable data (speed, agility, player movement)** could improve future analyses.

6.2 Training Recommendations

- **Players with low vertical jumps but good shooting** should focus on skill development over athleticism.
- **Elite defenders (high STL/G + BLK/G)** should refine on-ball defensive techniques.
- **Players with long wingspans** should develop shot-blocking and rebounding skills to maximize their strengths.

6.3 Player Draft Potential Takeaways

- NBA combine players' metrics are not great at predicting if they will be drafted!
  - As shown accuracies and F-1s in the 60s
- The models predicted 7 of our 12 players to get drafted even though none likely will
  - Shows the poor performance of the model
- It is slightly easier to predict if a guard will get drafted than a forward
  - As shown by the slightly higher accuracies for the guard models
- The models that weren't balanced by SNOTE predicted way more players to be drafted than undrafted
  - As shown by the higher recalls for the non-SNOTE models for the predicted target 1
- SCU's guards are likely more similar to NBA guards than SCU's forwards
  - As shown by 5 of 7 guards being predicted to be drafted compared to 2 of 5 forwards

7. References

1. The Effects of Functional Training on Physical Fitness and Skill-Related Performance Among Basketball Players: A Systematic Review

This systematic review examines how functional training (FT) impacts physical fitness and skill-related performance in basketball players. The findings indicate that FT significantly improves muscle strength, linear speed, cardiovascular endurance, flexibility, balance, and muscular endurance. However, the effects on power, change-of-direction speed, and

basketball-related performance were inconsistent. This insight suggests that while FT enhances certain physical attributes, its direct translation to on-court performance may vary.

Relevance: Understanding the specific benefits of FT can inform training regimens aimed at improving targeted physical attributes in SCU players, potentially enhancing their overall performance

2. Relationship Between Physical Fitness and Game-Related Statistics in Elite Professional Basketball Players: Regular Season vs. Playoffs

This research analyzes the correlations between players' anthropometric and physical characteristics with game-related statistics during regular seasons and playoffs. The study found few correlations during the regular season, and these correlations diminished further during playoff games, suggesting that physical attributes are not strong predictors of technical performance in high-stakes games.

Relevance: This insight emphasizes that while physical fitness is essential, technical skills, decision-making, and psychological factors may play more significant roles in a player's performance during critical game moments. This understanding can guide a more holistic approach to player development at SCU.

3. The Predictive Power of the NBA Draft Combine: A Statistical Analysis

This analysis evaluates the extent to which NBA Draft Combine metrics predict a player's future impact in the league. The study highlights that while certain metrics like wingspan-to-height ratio, vertical jump, and 3/4 court sprint speed are considered, their predictive power on actual performance varies.

Relevance: This analysis supports your hypothesis that specific physical factors—wingspan-to-height ratio, vertical jump, and sprint speed—are influential in assessing player potential. Incorporating these metrics into your player performance model aligns with findings that these attributes can impact a player's effectiveness on the court.

8. Conclusion

8.1 Summary

- SCU players can **improve their draft chances** by focusing on shooting efficiency & defensive impact.
- Players with **strong physical tools (wingspan, vertical jump)** should refine their skills to maximize their potential.
- Future work could **compare SCU data to past NBA draftees** for better predictions.
- Statistics measured at the NBA draft combine aren't good for predicting if a player will get drafted alone!
- SCU players' physical metrics aren't significantly different from players that go to NBA draft combine

8.2 Next Steps

- Expand the dataset with **Catapult wearable data**.
- Develop a **predictive model** using machine learning to classify future SCU prospects.
- Compare SCU player trajectories with **historical NBA draftees** to refine prediction accuracy
- Expand the dataset and models to include SCU's Women's Basketball Team
- Try other classification models on the combine database to predict if players will be drafted