Predicting injuries in SU Men’s Basketball

Overview of the Analysis and Findings

Index

Introduction

Background:

- Importance of muscle imbalance monitoring in sports.
- Objective: Analyze muscle imbalance data to predict injury risk and inform preventive strategies.

Research Question:

- Can muscle imbalance metrics predict player injuries?
- Can data collected during player session predict muscle imbalances?

Data Overview

Dataset Overview:

- **Injury Data:** Player history and injury types.
- **Imbalance Data:** Key metrics like Hamstring to Quad Ratio, Quad Imbalance Percent, etc.
- **Session Data:** Player metrics collected during training sessions such as, distance ran, jump height.

Steps Taken:

- **Prepared data:** Merged data sets, dropped unnecessary columns, removed post injury imbalance data to not skew data.
- **Classification:** Used a classification model to find a relationship between muscle imbalances and injury risk.
- **Regression:** Used regression to find a relationship between player session data and muscle imbalances therefore tying all data sets together.
- **Plots:** Plotted relevant data.

EDA

- **Exploratory Data Analysis**
  - **Summary Statistics**
    - Generated mean, median, quartiles, min and max
    - Key metrics: Hamstring to Quad Ratio, Quad Imbalance Percent, etc.
  - **Visualizations**
    - Plot highlighting safe zones: areas representing ranges of muscle imbalances where injury risk decreases.
    - Highlights observed differences across groups.

Model Development

Approach

Explore relationships between muscle imbalance metrics and injury likelihood using various models

- Further explore the relationships between training metrics and muscle imbalances using regression
- Tied the 2 by outputting a ‘safe range’ for muscle imbalances.
- Explore relationships between muscle imbalance metrics and injury likelihood using Classifiers

Iterative Feature Selection

Tested different combinations of imbalance features to find the best predictors.

Model Evaluation

Performance Metrics:

- Mean Squared Error (MSE): Lower values indicate better model performance.
- R-squared: Higher values indicate better model fit.

Key Findings:

- Models involving Groin Imbalance Percent and Quad Imbalance Percent had lower MSE, suggesting these are significant predictors.
- Polynomial regression models with multiple features showed the best fit (highest R-squared).

Observations and Insights

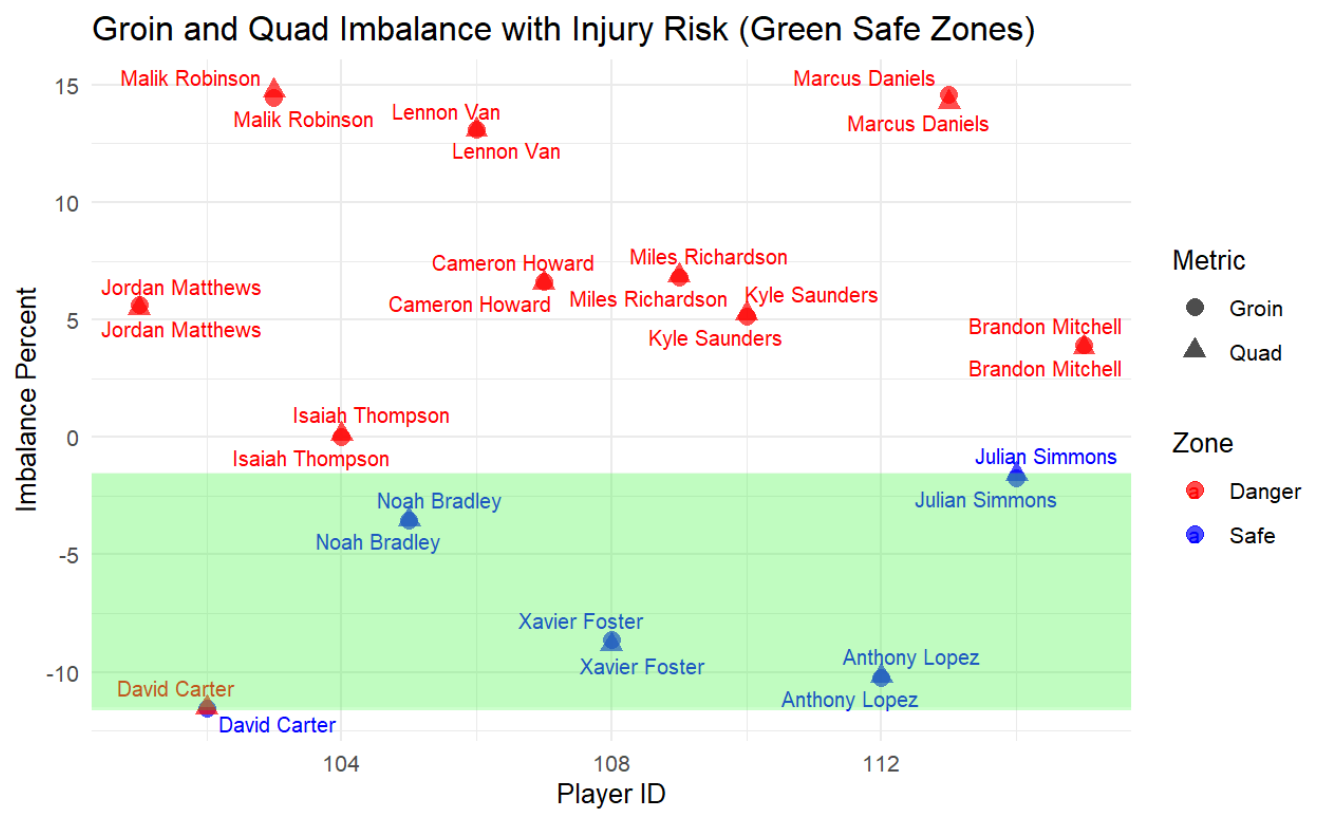
Key Observations:

- Muscle imbalance metrics are valuable predictors of injury risk.
- Groin and Quad imbalance metrics together stand out as the most significant predictors.
- Time played, speed, and max jump height seem to be particularly good predictors for imbalances.
- Models can guide targeted training and recovery programs to reduce injury likelihood.

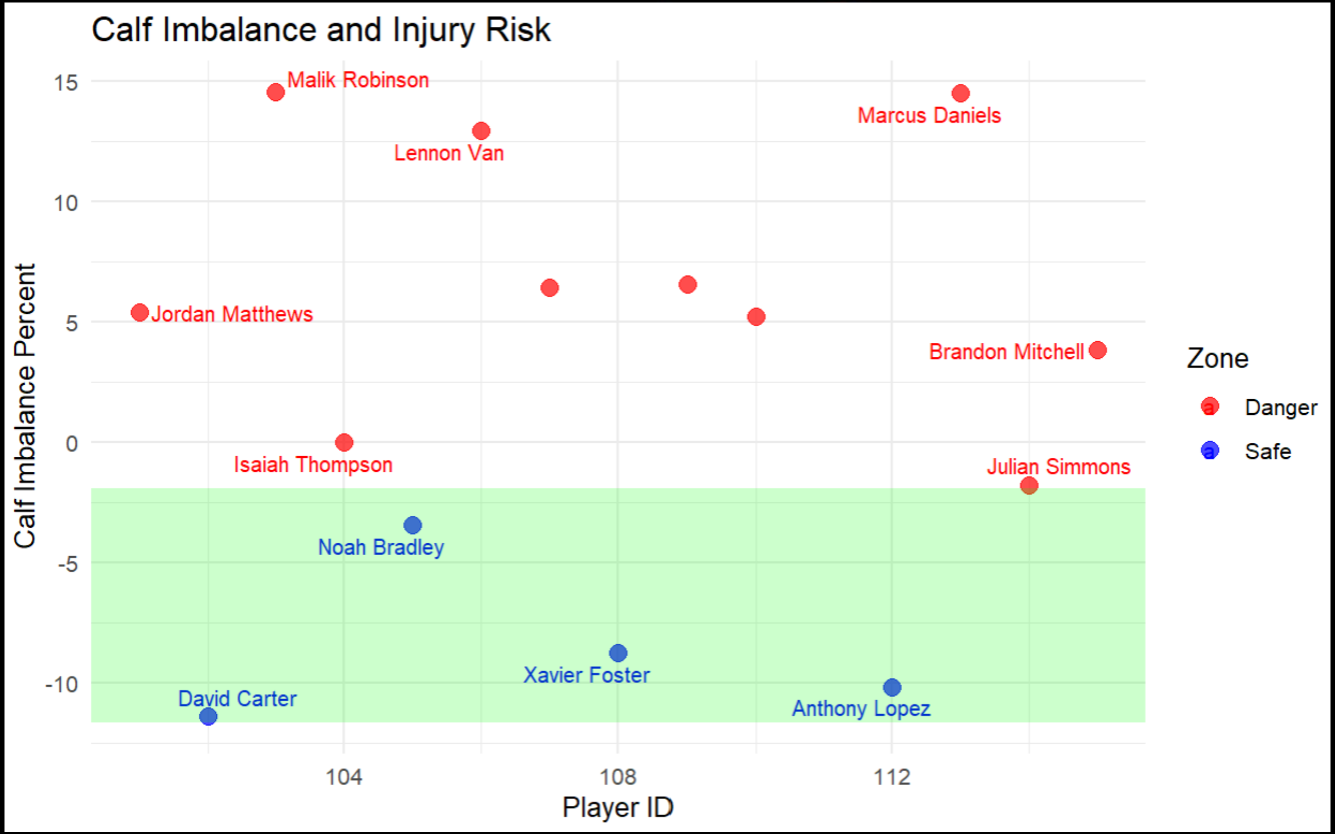
Potential Applications:

- Inform player-specific training plans.
- Develop injury prevention strategies based on high-risk imbalance thresholds.

Observations and Insights



Observations and Insights



Recommendation

Expand Data:

Collect more comprehensive data or generate synthetic samples.

Refine Models:

Focus on enhancing polynomial regression models.

Implement Findings:

Use imbalance thresholds to design preventive measures and improve player safety.

Possible Application

Store session data live, and use the model in an app to flag players after training sessions as possible injury risks.

Physical therapists can use the imbalances to tailor rehab plans to target specific imbalances.

Conclusion

Summary:

- Tracking muscle imbalances, especially between the groin and quadriceps, helps identify injury risks early. Addressing these imbalances can prevent injuries and promote safer movement.

Next Steps:

- Collaborate with teams to validate findings.
- Implement predictive models for injury prevention.