

Frisbee Data Dynamics: Boston Glory

Team Members: Audrey Sellers, Mohamad Ali Saadeddine, Priyanka Chaudhari

Project Management and Workflow

Background Research:

Ultimate Frisbee (UF) is a sport that is played all over the world, but we will be focusing on the Ultimate Frisbee Association (UFA) throughout this project. During an ultimate game, each team has seven players on the field with an objective of getting the frisbee to the opposing team's end zone.¹ There are several strategies that involve offensive plays and defensive zones that teams implement to win games.¹ In general, the key to winning games is to score goals with the best lineup possible. The positions of ultimate are Handlers, Cutters, Poppers, Wings, the Cup, Short Deeps, and Middle Flats.² Our Business Advisor has informed us that they know position is highly correlated with scoring goals. It has also been seen that teams who win have significantly fewer backward passes,³ which aligns with our project's overall objective: to increase pass yardage and reduce players' turnover rates.

Along with goal scoring, it is imperative for ultimate teams to have a cohesive offensive lineup that executes assists and passes. There are different ways to decide which players should start and receive the most playing time. In fact, a research study conducted on a national UF team concluded that players exert more energy when playing higher-ranked opponents.⁴ This indicates a need for higher offensive performance, more effort, and calculated lineup formations during games against lower-ranked teams. Similarly, through feature selection, It has been found that ultimate teams are more likely to win when players who have more complete passes and assists.⁵ The same project also found that when players who have a high drop rate play, the team is less likely to win due to an increased turnover rate.⁵

¹ "Ultimate 101." 2018. WatchUFA. January 29, 2018. <https://watchufa.com/league/ultimate101>.

² "Ultimate Frisbee Positions - Which One Is Best for You?" n.d. Theuap.com. Accessed July 13, 2024. <https://www.theuap.com/ultimate-frisbee-positions-which-one-is-best-for-you>.

³ Lam, Hilary, Otto Kolbinger, Martin Lames, and Tiago Guedes Russomanno. 2021. "State Transition Modeling in Ultimate Frisbee: Adaptation of a Promising Method for Performance Analysis in Invasion Sports." *Frontiers in Psychology* 12: 664511. <https://doi.org/10.3389/fpsyg.2021.664511>.

⁴ Castillo, Daniel, Javier Raya-González, Aaron T. Scanlan, Marta Domínguez-Díez, and María C. Madueno. 2023. "Influence of Opponent Ranking on the Physical Demands Encountered during Ultimate Frisbee Match-Play." *Sports Biomechanics* 22 (7): 822–33. <https://doi.org/10.1080/14763141.2020.1766101>.

⁵ Go, Caitlin. n.d. "Exploring Predictors of Team Success in Ultimate Frisbee: An Analysis of Game Statistics for Stanford Women's Ultimate."

Problem Statement:

Currently, Boston Glory does not sufficiently utilize the vast amount of data available to make strategic decisions. The primary objective is to enhance the performance of the Boston Glory frisbee team by leveraging data analytics to provide actionable insights. The focus is on identifying key areas of offensive and defensive strengths and weaknesses and suggesting on-field strategies that can lead to wins and championships within the UFA league.

Scope:

This project will encompass data collection and preprocessing, exploratory data analysis, feature engineering, and machine learning model development. The data will be sourced from the UFA API and processed through a Google Cloud infrastructure, including Google Big Query for data storage and analysis. The goal is to analyze various performance metrics, such as points per game, yards received, and defensive weaknesses, to provide comprehensive insights for improving the team's strategies.

Expected Approach:**Data Collection and Preprocessing:**

Utilize the UFA API to collect relevant data from the year 2023-2024. Implement an ETL pipeline using Google Cloud Functions to automate data ingestion into Google Big Query. Perform data cleaning and preprocessing to ensure data quality like:

- [i] Dropped irrelevant columns that did not seem essential for our analysis and subject matter.
- [ii] Data Transformation: Converted categorical features to appropriate data types for analysis and modeling, including extracting players names and playerid's from the list form that they were initially in. This involved writing a function that could parse through the JSON list of strings and give us an individual set of values.

Exploratory Data Analysis (EDA):

Conduct EDA to understand the data distribution and identify key performance metrics. Visualize data to uncover trends and patterns like

- [i] Players with more than 5 PPG for Glory
- [ii] Comparison of Average points: Boston Glory V/S Other Teams
- [iii] Defensive Weakness Distribution by Team (Excluding Outliers)
- [iv] Playing Time and Performance Metrics: Boston Glory
- [v] Correlation Matrix of Playing Time and Performance Metrics: Boston Glory
- [vi] Correlation Matrix of Glory Key Statistics

Initial Data Analysis:

The initial data analysis revealed that Boston Glory scores fewer points per game (7.16 PPG) compared to the league average (7.43 PPG), indicating a need for offensive improvement. Key performance metrics for Boston Glory include yards received, yards per reception, seconds played, and yards thrown. A correlation matrix of these metrics highlighted significant relationships between them, providing a foundation for further analysis.

Expected Findings:

- [i] Identification of top-performing players and their contributions to the team's success.
- [ii] Insights into the team's defensive weaknesses, with Boston Glory averaging 0.21 goals scored against them versus a 0.39 league average.
- [iii] Recommendations for optimizing player lineups and in-game strategies to enhance overall team performance.

Feature Engineering:

Further going to develop features that capture critical aspects of gameplay, such as yardage, pass completion rates, and defensive stats.

Machine Learning Model Development:

- [i] We will be implementing supervised machine learning method (Linear Regression) for predicting Points per Game (Goals + Assists) using dependent variables such as 'Yard Received', 'Yard per Reception', 'Seconds Played', 'Yards Thrown
- [ii] Deliver multiple models and strategies to the Boston Glory front office and present the real time dashboards
- [iii] Provide actionable on-field strategy recommendations based on the model outcomes.
- [iv] Iterate and refine models based on feedback and additional data

Future Recommendations:

Deliver actionable insights and strategy recommendations to the Boston Glory front office using dashboards. Iterate and refine models based on feedback and additional data.

Risks/Limitations:

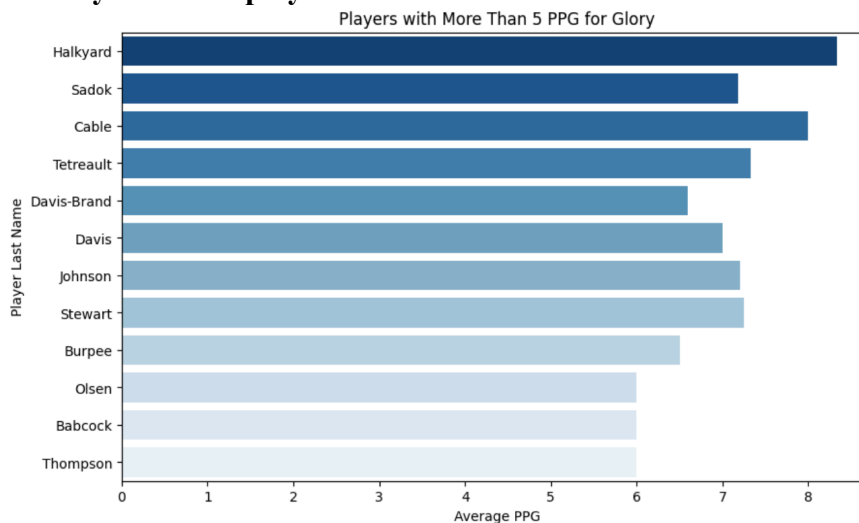
- [i] Missing values in certain columns can affect the accuracy of the analysis.
- [ii] Determining the most effective methodology for analysis is challenging.
- [iii] Limited historical data for some players and teams.
- [iv] Potential biases in the data sources used.
- [v] The need for continuous updates to models as new data becomes available.
- [vi] Balancing model complexity with interpretability for practical use by the coaching staff.
- [vii] Limitation of outlier data points such as players who do not feature.

References:

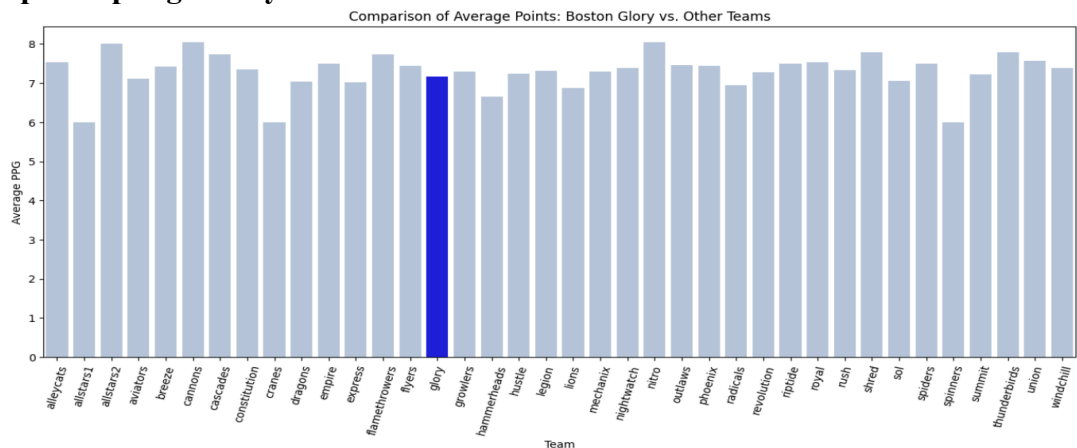
1. “Ultimate 101.” 2018. WatchUFA. January 29, 2018. <https://watchufa.com/league/ultimate101>.
2. “Ultimate Frisbee Positions - Which One Is Best for You?” n.d. Theuap.com. Accessed July 13, 2024. <https://www.theuap.com/ultimate-frisbee-positions-which-one-is-best-for-you>.
3. Lam, Hilary, Otto Kolbinger, Martin Lames, and Tiago Guedes Russomanno. 2021. “State Transition Modeling in Ultimate Frisbee: Adaptation of a Promising Method for Performance Analysis in Invasion Sports.” *Frontiers in Psychology* 12: 664511. <https://doi.org/10.3389/fpsyg.2021.664511>.
4. Castillo, Daniel, Javier Raya-González, Aaron T. Scanlan, Marta Domínguez-Díez, and María C. Madueno. 2023. “Influence of Opponent Ranking on the Physical Demands Encountered during Ultimate Frisbee Match-Play.” *Sports Biomechanics* 22 (7): 822–33. <https://doi.org/10.1080/14763141.2020.1766101>.
5. Go, Caitlin. n.d. “Exploring Predictors of Team Success in Ultimate Frisbee: An Analysis of Game Statistics for Stanford Women’s Ultimate.”

Appendix:

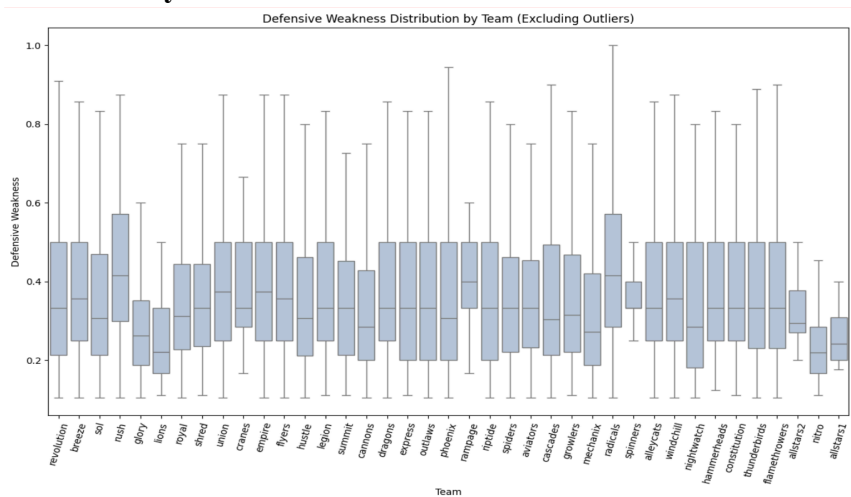
A. Top Boston Glory offensive players.



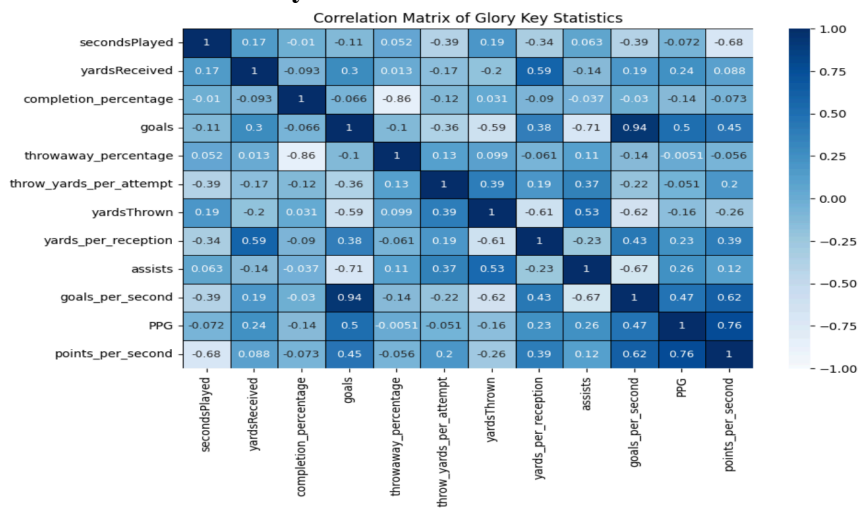
B. Average points per game by team in the UFA.



C. Defensive weakness by team in the UFA



D. Correlation matrix of UFA key attributes.



Contribution:

Name	Coding	Theory
Audrey Sellers	BigQuery SQL and Exploratory Data Analysis	Analysis in Google Colab, Poster, Powerpoint Presentation, Report
Mohamad Ali Saadeddine	BigQuery SQL	Poster, Powerpoint Presentation, Report
Priyanka Chaudhari (Project Manager)	ETL Pipeline(data collection and processing) and BigQuery SQL	Poster, Powerpoint Presentation, Report