Beyond the Net: Analyzing Critical Performance Drivers for Orlando City Men’s Soccer

Abigail Wright, Georgina Wu,
Oma Persaud, and Shivani Ananth

Motivation

- Understanding the predictors that significantly impact the game outcome: Win, Loss, or Draw

- Creating this foundation of understanding can be beneficial to fans, team members, coaching staff, as well as many others.

- As of August 2024, 37% of adults 18 years and older considered themselves soccer fans in the US according for Forbes (Carosella, *Forbes: American Soccer Fans*).

Dataset Details

Data Description:

- The dataset originally contains
  - 26 predictive variables
  - 34 observations

Data Collection Process:

- This data was independently compiled by our team utilizing various sources concerning team member statistics, rankings, and specific game day data.

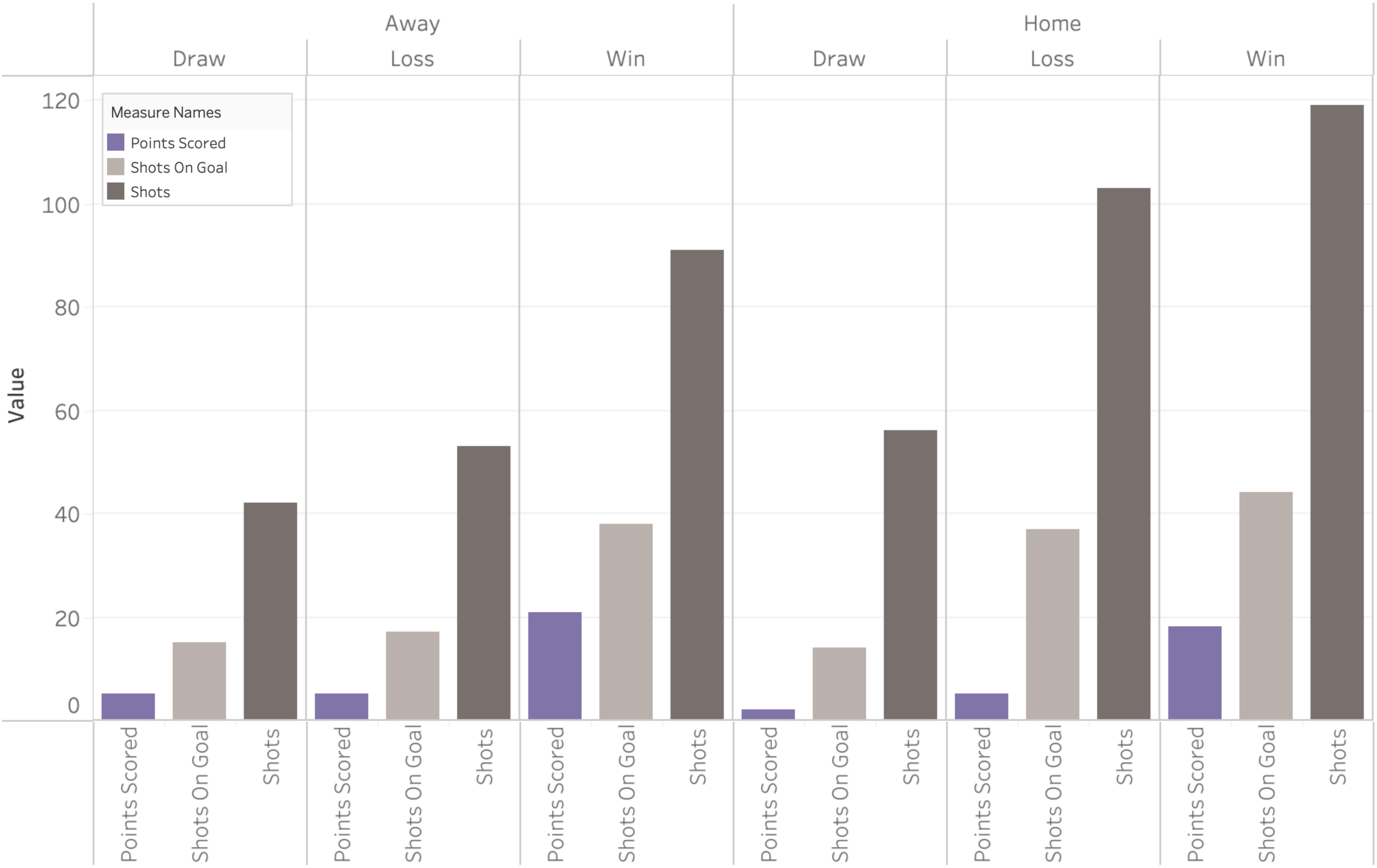
Main Sources:
  - Orlando City Team Stats
  - ESPN

Data Description

Dependent Variable:

Final Score

- Draw
- Win
- Loss



Data Description

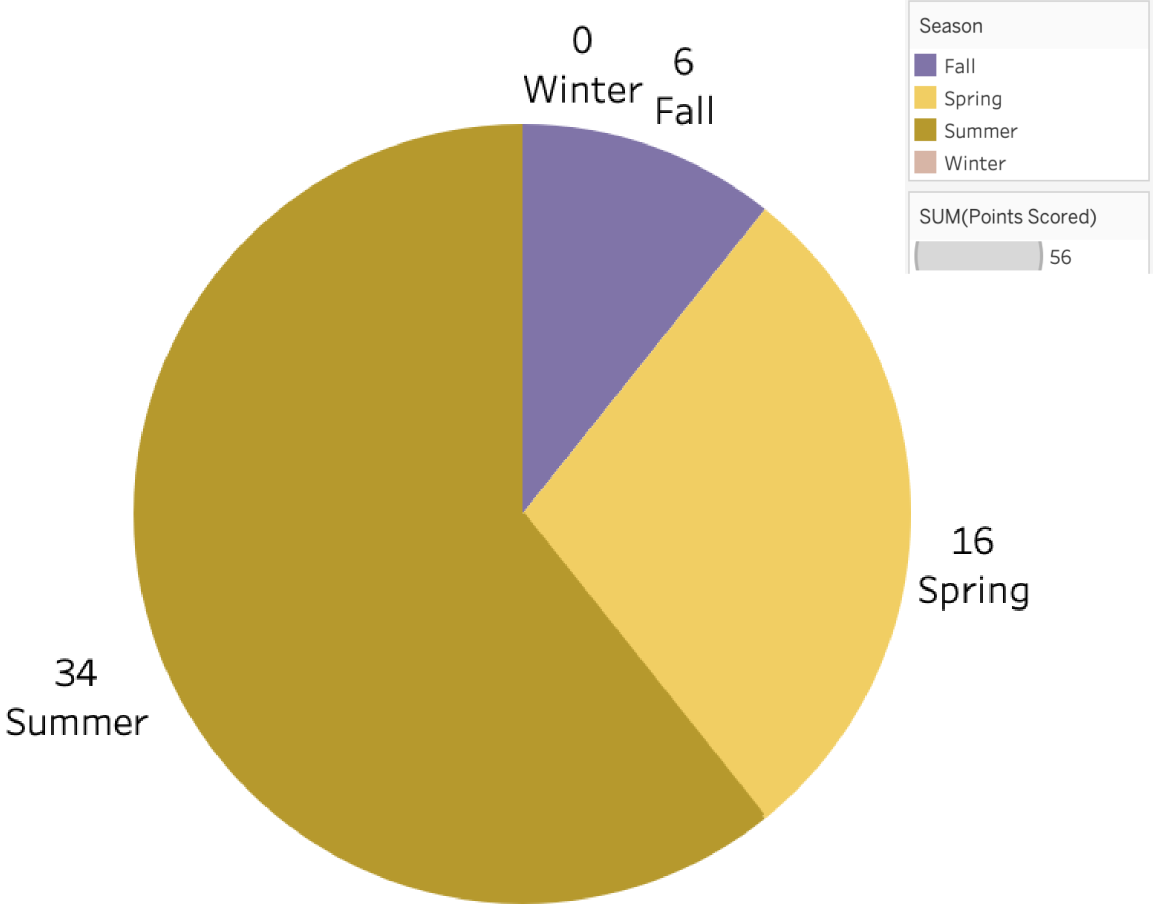
Independent Variables:

- Home or Away
- Possession Percentage
- Shots
- Shots on Goal
- Blocked Shots
- Total Passes
- Passing Accuracy Percentage
- Corners
- Total Corners
- Offsides
- Ariel Duals Won
- Goalkeeper Saves
- Clearances
- Fouls
- Yellow Cards
- Red Cards
- Top 3 Scoring Team members
- Top 3 Assisting Team Members
- Opposing Team
- Season
- Points Scored

Generalizations

Game Day: Original date of game, categorized by Orlando FL seasons.

- **Winter**: December to February
- **Spring**: March to May
- **Summer**: June to August
- **Fall**: September to November



Generalizations

Opposing Teams:

US Teams:

- Atlanta
- Austin
- Charlotte
- Chicago
- Cincinnati
- Columbus
- D.C
- Dallas
- Kanas City
- LAFCA
- Miami
- Minnesota
- Nashville
- New England
- New York
- New York City
- Philidelphia
- San Jose

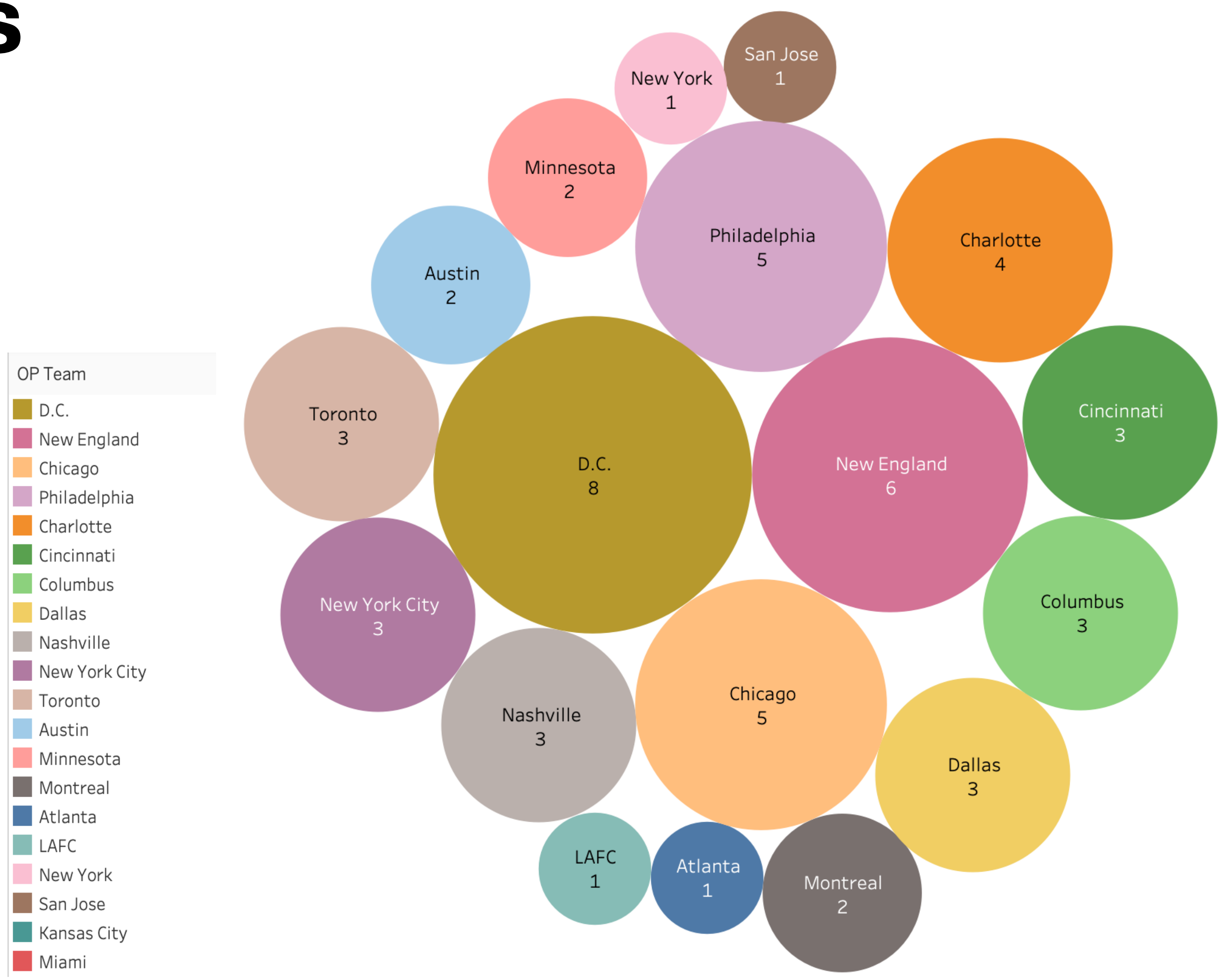
Non-US Teams:

- Montreal
- Toronto

Generalizations

Opposing Teams:

- Data concerning opposing team prior to generalization and points scored by Orlando City throughout the Regular 2024 Season



Top 3 Scoring Players

Duncan McGuire
#13
Forward
Orlando City
Senior
Games Played: 27
Minutes Played: 1599

Facundo Torres
#10
Forward
Orlando City
Senior
Games Played: 32
Minutes Played: 2642
Transferred December 2024

Ramiro Enrique
#7
Forward
Orlando City
Senior
Games Played: 20
Minutes Played: 1082

Top 3 Assists

Martín Ojeda
#10
Midfielder
Orlando City
Senior
Games Played: 34
Minutes Played: 1914

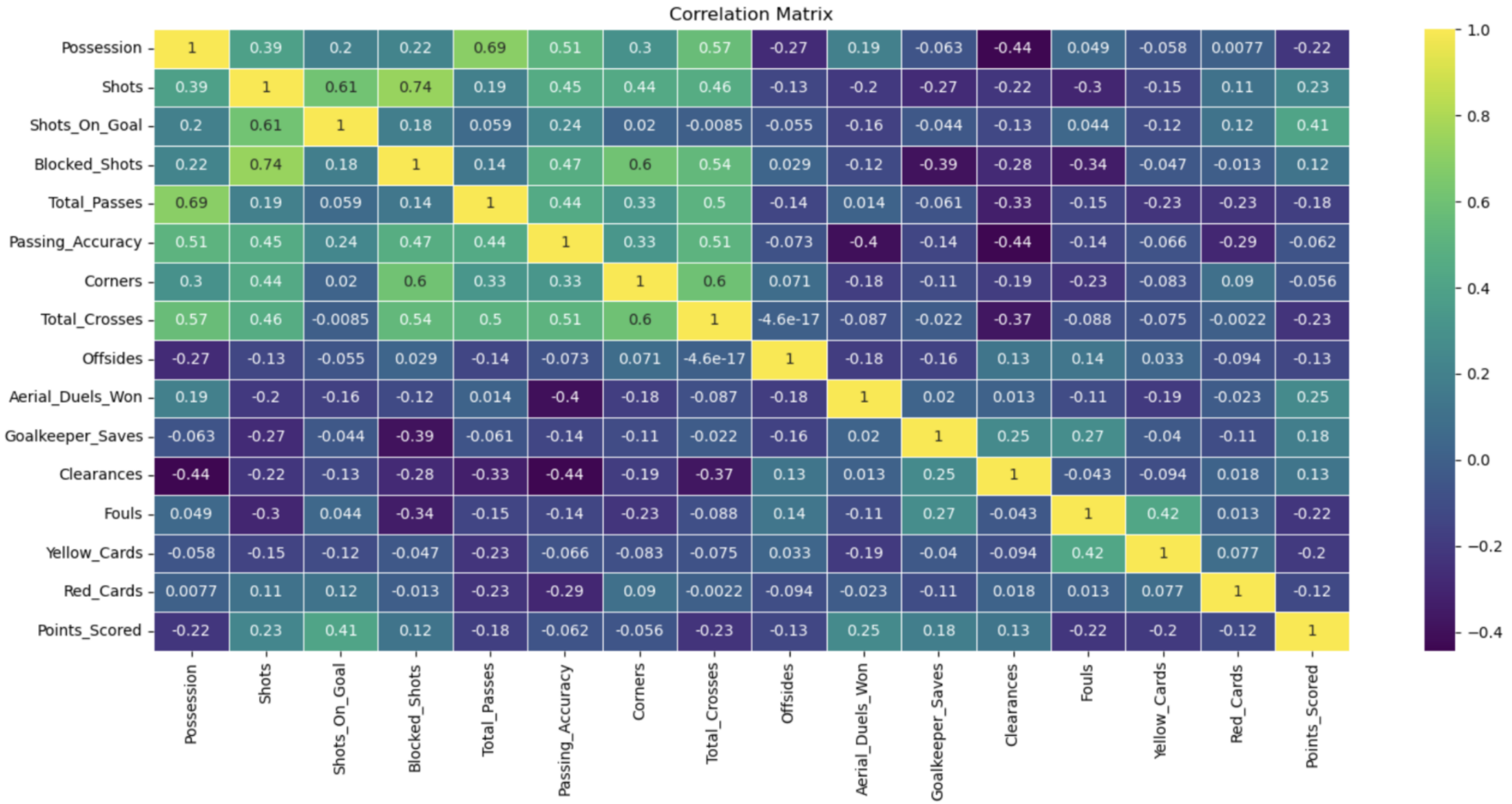
Nicolás Lodeiro
#14
Midfielder
Orlando City
Senior
Games Played: 34
Minutes Played: 1697

Iván Angulo
#77
Forward
Orlando City
Senior
Games Played: 34
Minutes Played: 2772

Exploratory Data Analysis

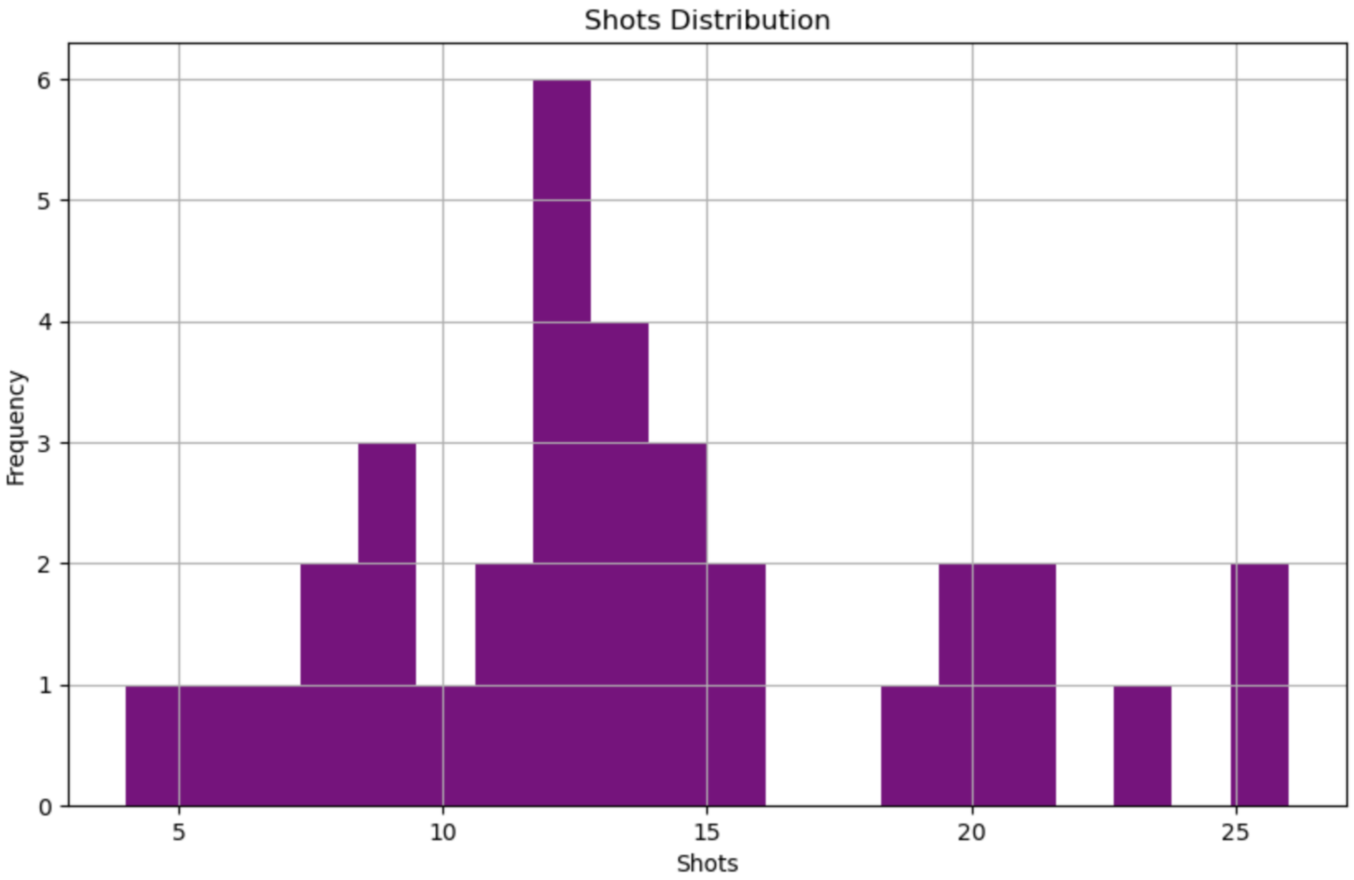
Correlation Matrix

- This matrix is visualizing all the numerical predictors:
    - Possession
    - Shots
    - Shots On Goal
    - Blocked Shots
    - Total Passes
    - Corners
    - Total Crosses
    - Offsides
    - Ariel Duals Won
    - Goalkeeper Saves
    - Clearances
    - Fouls
    - Yellow Cards
    - Red Cards
    - Points Scored



Exploratory Data Analysis
Shots Distribution

- The Orlando City players take mainly between 12-13 shots per game.



Exploratory Data Analysis

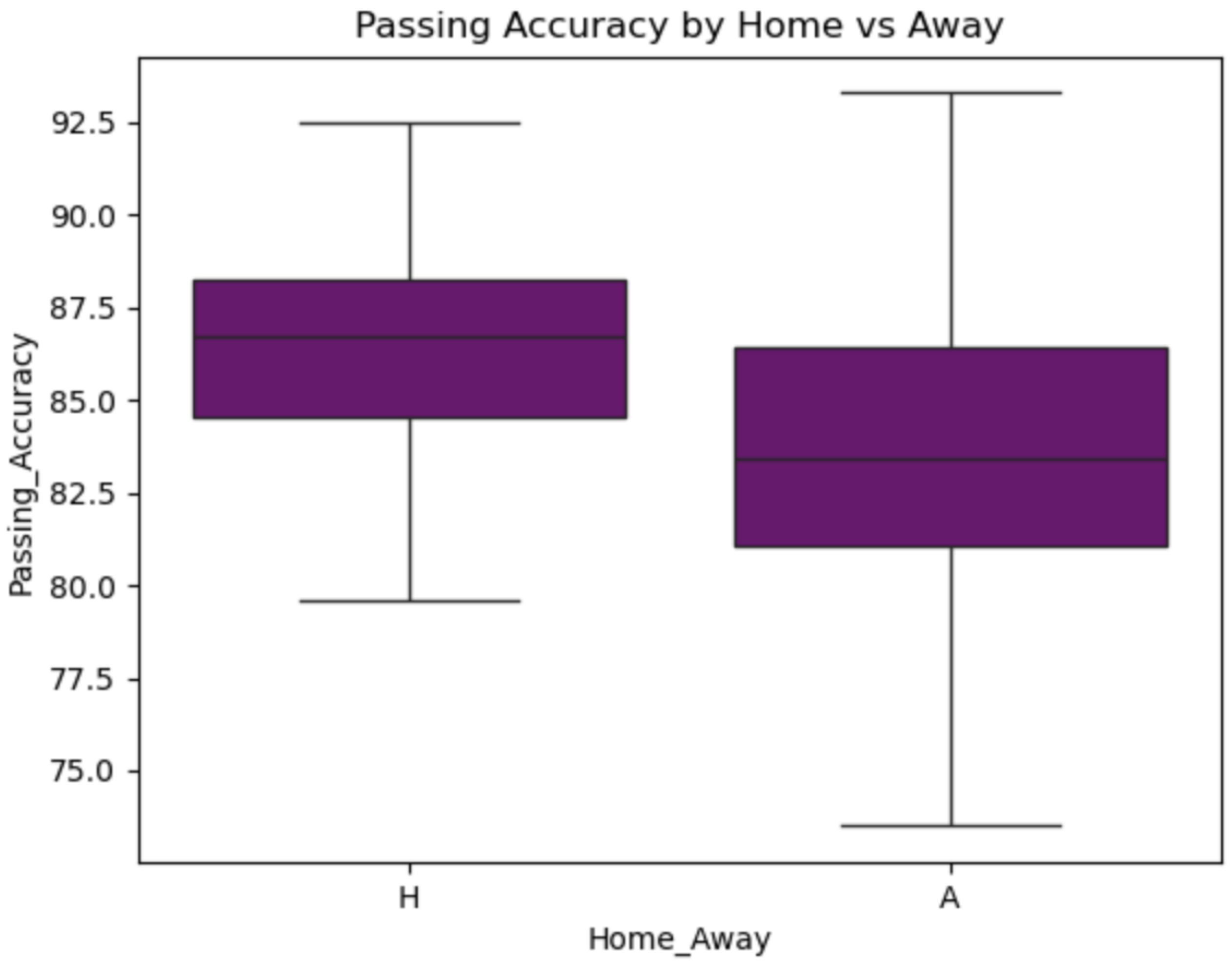
Shots Passing Accuracy

Home Game:

- Passing Accuracy Mean: 86%

Away Game:

- Passing Accuracy Mean: 83%

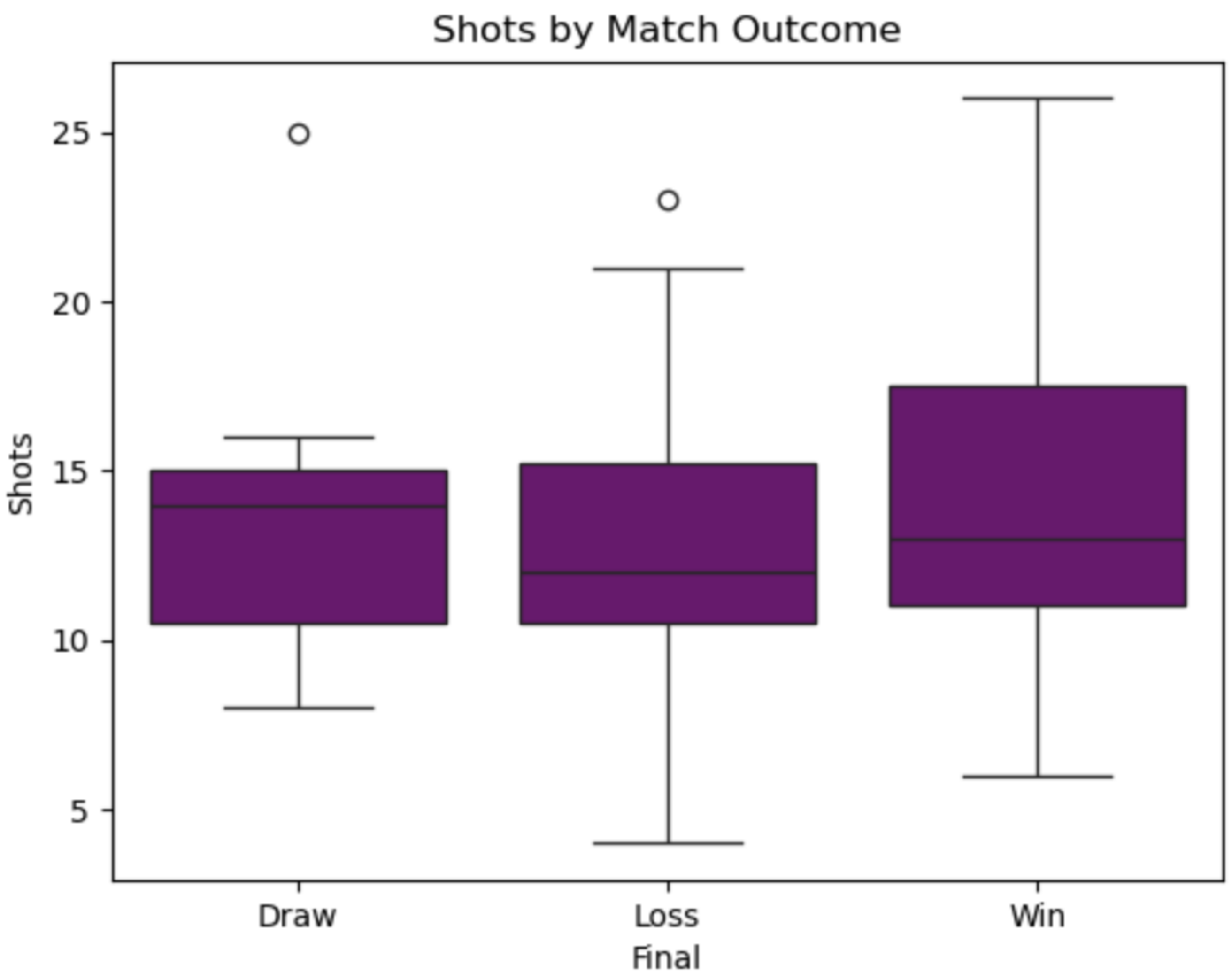


Exploratory Data Analysis

Shots by Match Outcome

From these box plots depicting game outcome vs shots some outliers are shown.

- Draw: 25 shots
- Loss: about 23 shots
- Win: None

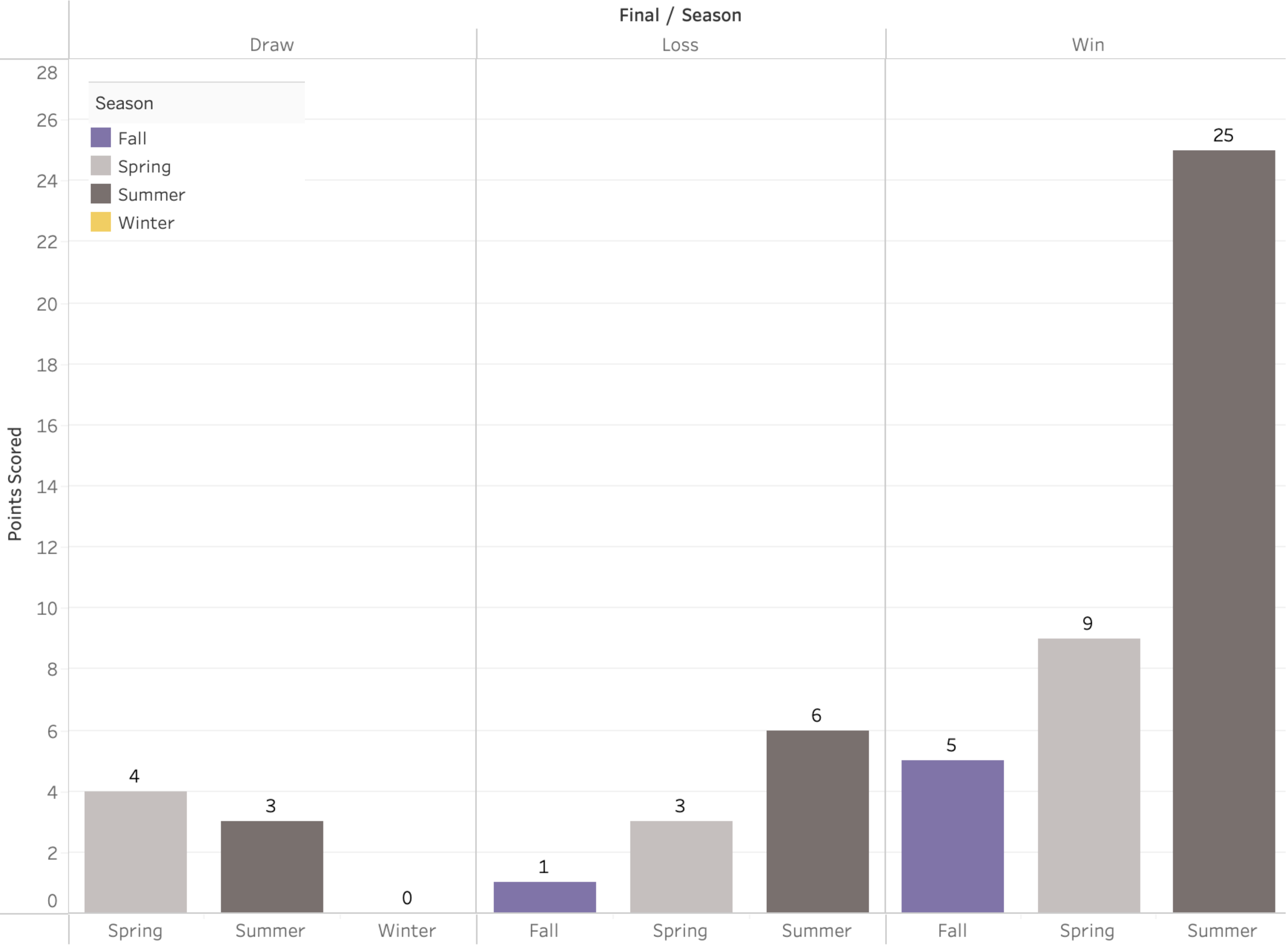


General Visualizations

Points Scored: Regular Season

Summer
16 games
- Draw: 3 points: 2 games
- Win: 25 points: 9 games
- Loss: 6 points: 5 games

Spring
14 Games
- Draw: 4 points: 4 games
- Win: 9 points: 4 games
- Loss: 3 points: 6 games



General Visualizations

Points Scored: Regular Season

Winter

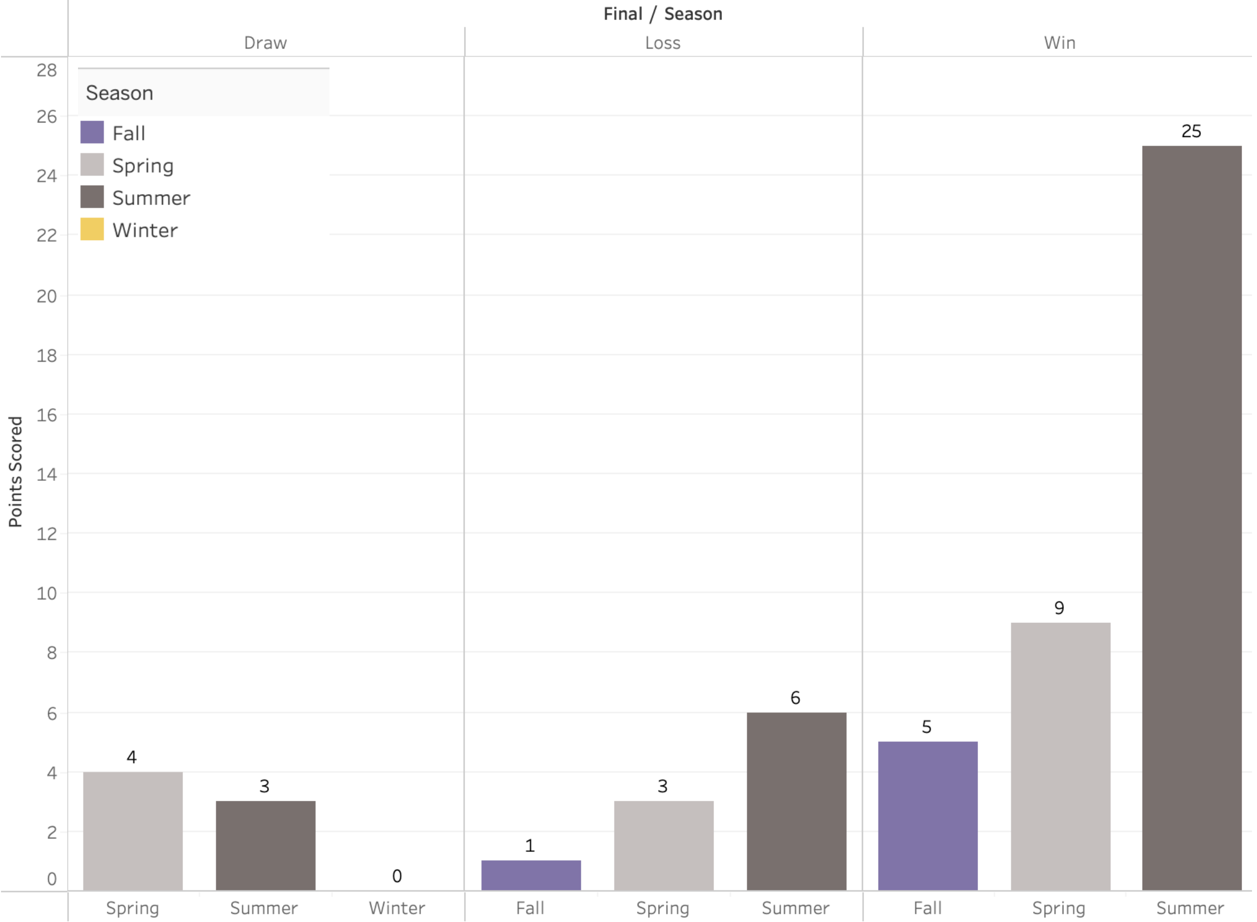
1 Game

- Draw: 0 points: 1 game
- Win: No games
- Loss: No games

Fall

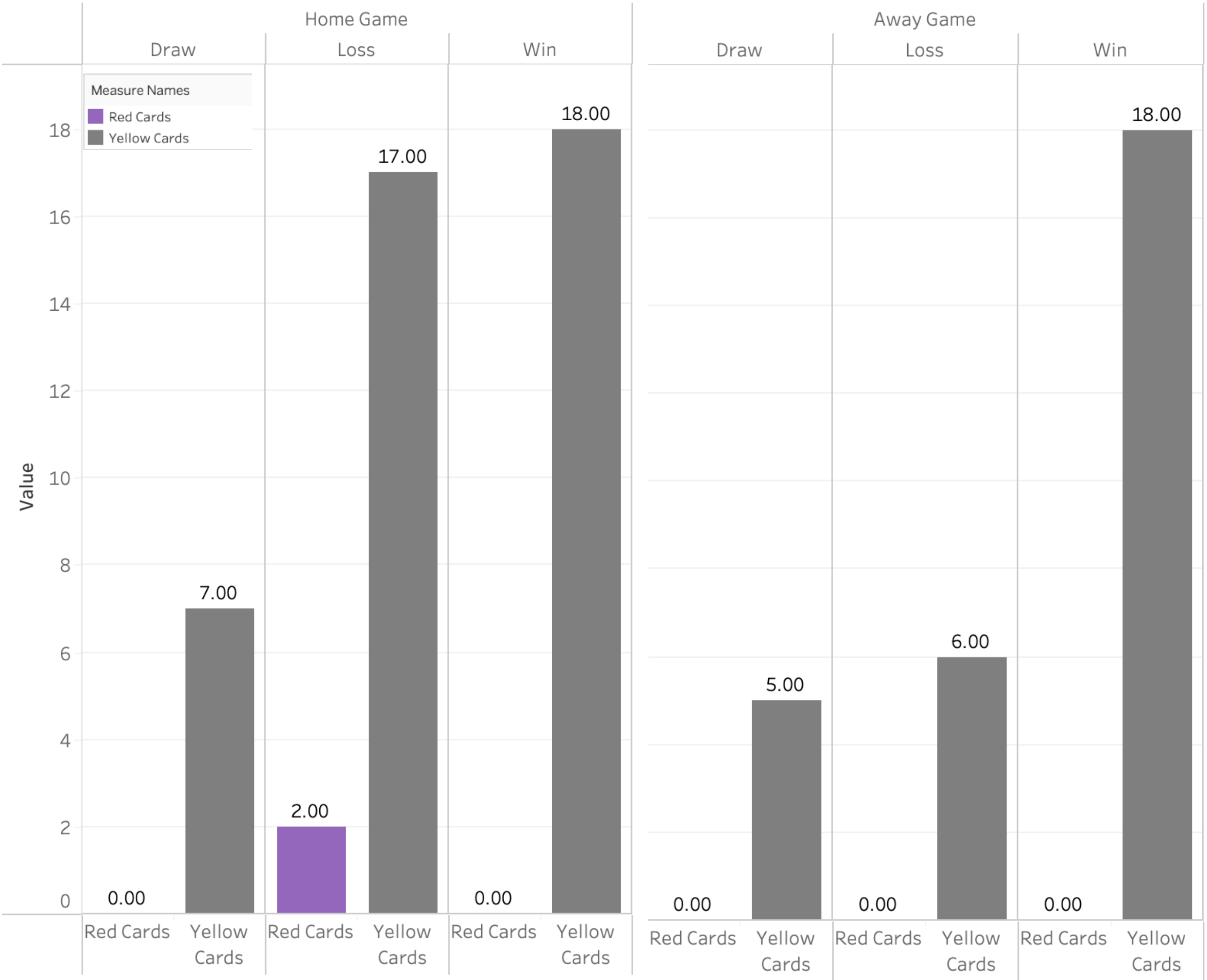
3 Games

- Draw: No games
- Win: 5 points: 2 games
- Loss: 1 point: 1 game



General Visualization

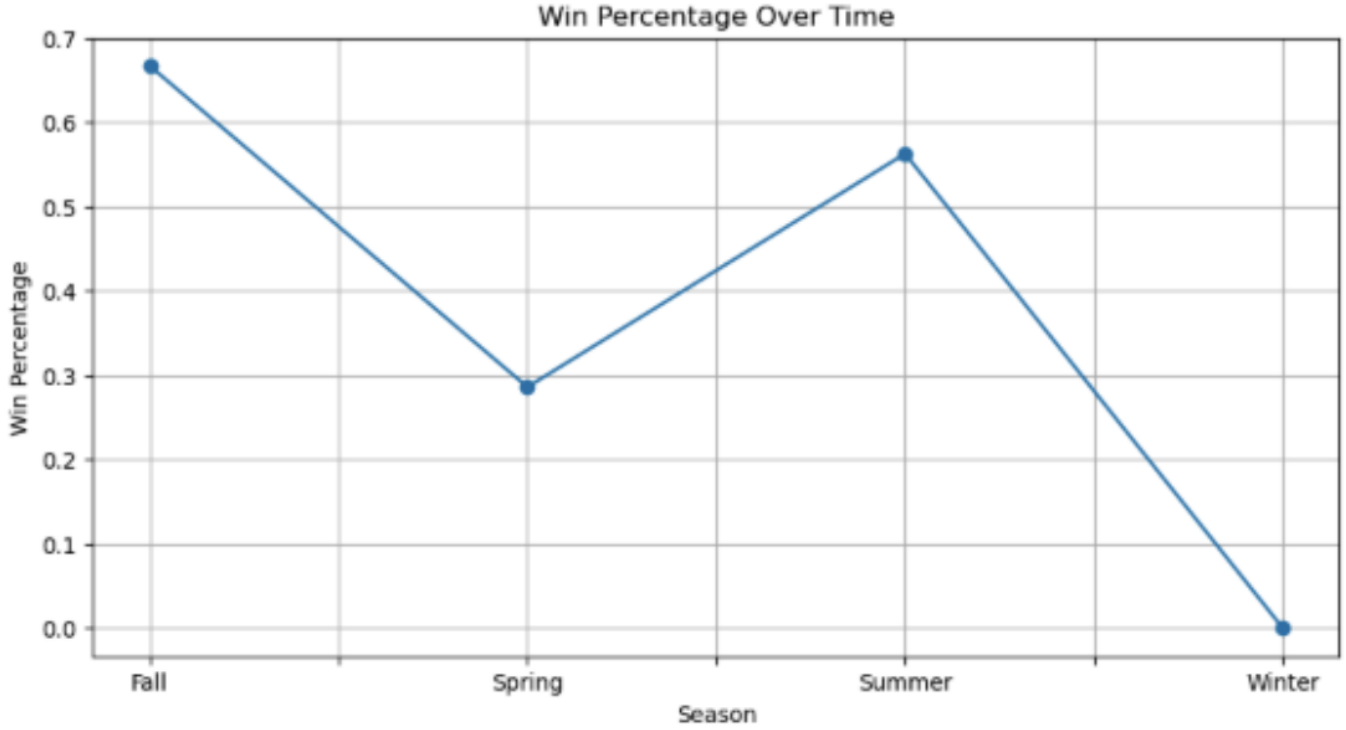
Data concerning the number of Yellow Cards and Red Cards given to Orlando City plays during matches that resulted in Draw, Win, or Loss.



General Visualizations

Win Percentage Over Seasons

- Highest win percentage in Fall (67%)
- Sharp drop in Spring (29%)
- Go back up in the Summer (55%)
- Dramatic decline to 0% in the Winter



General Visualizations

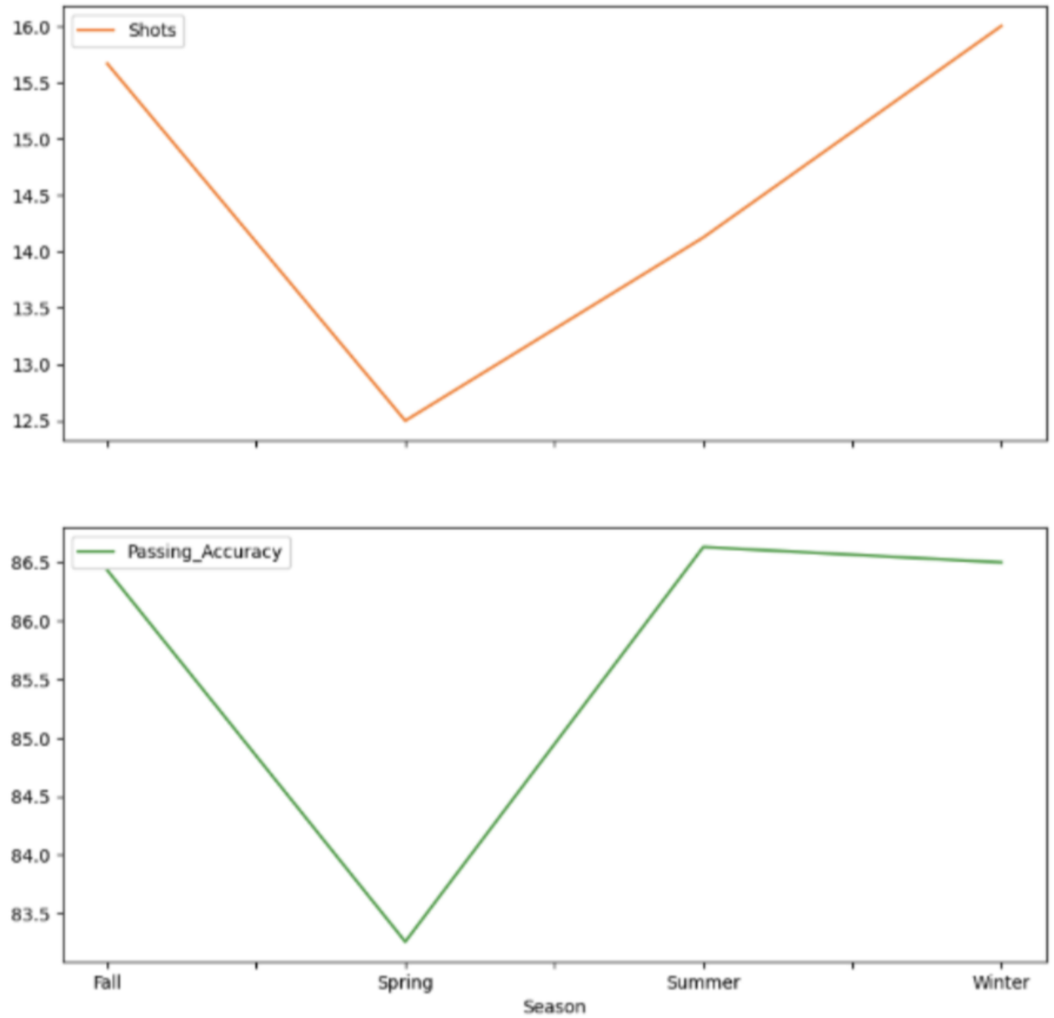
Shots and Accuracy Over Seasons

Shots Taken Per Season

- High in the Fall (16)
- Lowest in Spring (12.5)
- Steady increase for Summer and Winter

Passing Accuracy

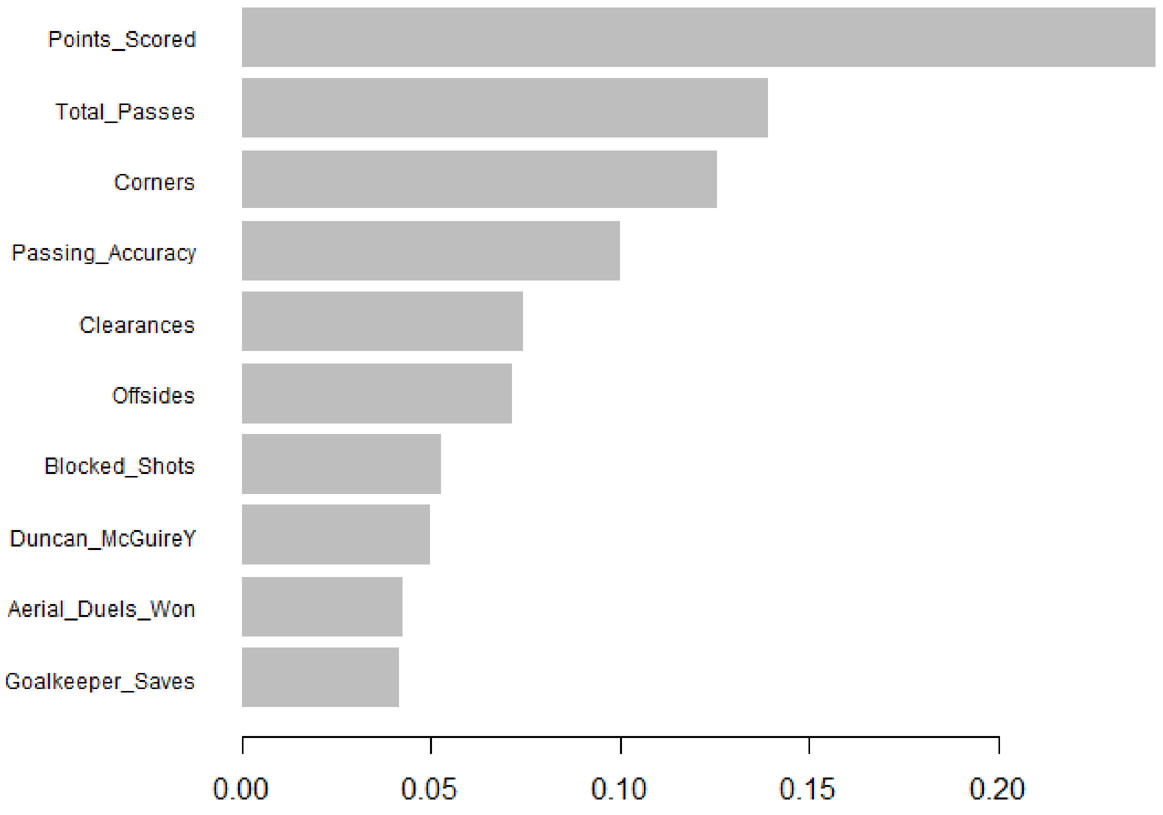
- High in the Fall (86.5%)
- Declined in Spring (83.3%)
- Spiked to highest in the Summer (86.7%) and was steady in Winter (86.5%)



Methodology

Full Model Construction

XGBOOST MODEL WITH ALL VARIABLES



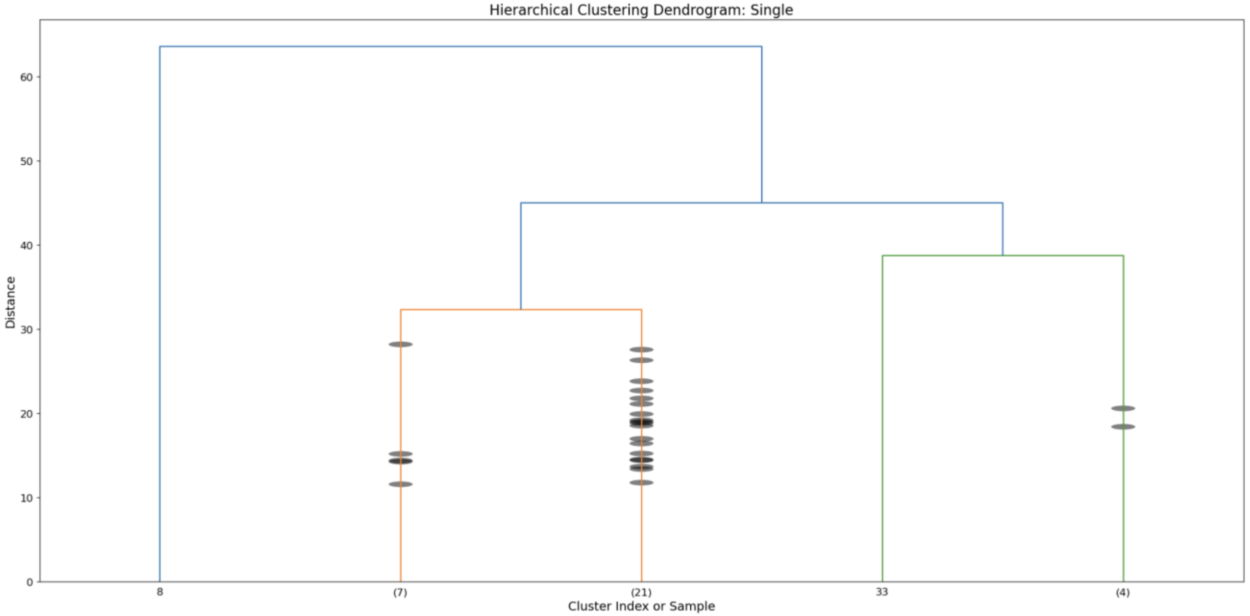
Feature: The variable in order of importance
Gain: How much better does the model get every time it uses this feature?
Cover: How many data points are impacted when I split on this feature?
Frequency: How often does this feature show up in trees?
Test Accuracy: 0.571

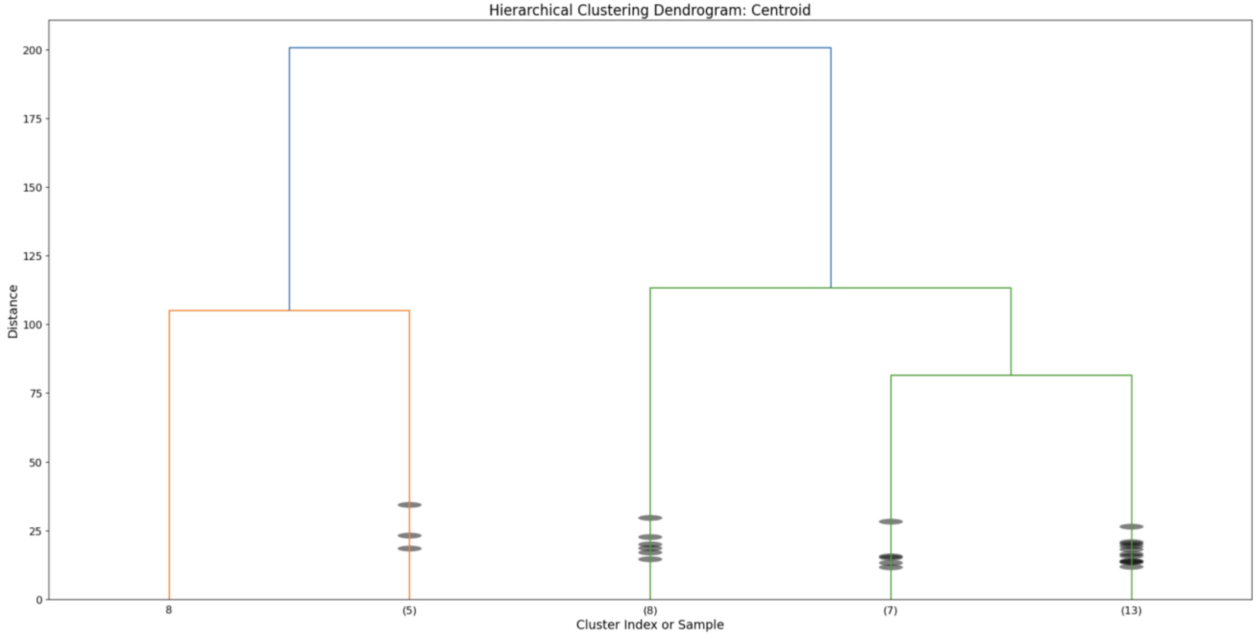
	Feature	Gain	Cover	Frequency
1:	Points_Scored	0.241834073	0.22037953	0.13793103
2:	Total_Passes	0.139245869	0.13216897	0.08620690
3:	Corners	0.125644222	0.16280789	0.13793103
4:	Passing_Accuracy	0.100289760	0.08191641	0.13793103
5:	Clearances	0.074804550	0.07269149	0.06896552
6:	Offsides	0.071679720	0.04908535	0.08620690
7:	Blocked_Shots	0.052986131	0.03352655	0.05172414
8:	Duncan_McGuireY	0.049933113	0.11091328	0.08620690
9:	Aerial_Duels_Won	0.042675919	0.06094731	0.06896552
10:	Goalkeeper_Saves	0.041963628	0.01220484	0.03448276
11:	Possession	0.033124387	0.03247722	0.05172414
12:	Yellow_Cards	0.020742116	0.01641519	0.03448276
13:	Shots_On_Goal	0.005076512	0.01446598	0.01724138

Hierarchical Clustering

Linkage Method: Single vs. Centroid

- Notably different clustering patterns between Single and Centroid linkage patterns.
- Sample 8 here depicts the farthest point from the others in the Single Method. Thus, links last to the other clusters.
- Dots represent data points within the clusters.
- (21) has the most clusters in the Single Method, whereas the points are spaced out between (5), (8), (7), and (13) in the Centroid Method.





Principal Component Analysis

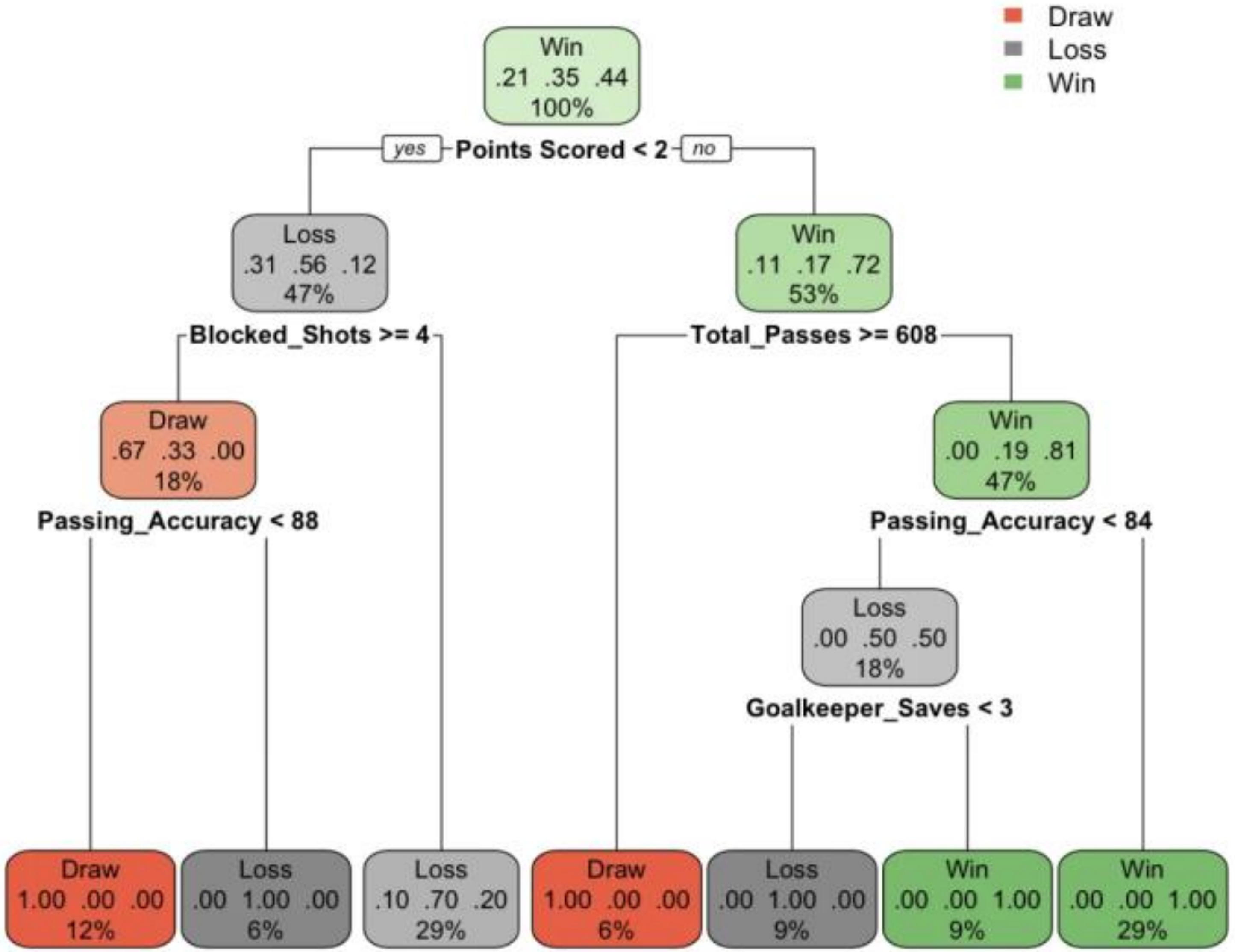
- **Top 3 Influential Predictors:**
  - Total_Crosses
  - Possession
  - Total_Passes

- **Top 5 Influential Predictors:**
  - Clearances
  - Passing_Accuracy
  - Total_Crosses
  - Possession
  - Total_Passes

- **Top 7 Influential Predictors:**
  - Corners
  - Shots
  - Clearances
  - Passing_Accuracy
  - Total_Crosses
  - Possession
  - Total_Passes

Decision Tree

Utilizing RStudio



Methodology

Significant Variables:

From investigating multiple approaches to best find the most significant variables, the following show the variables most likely to predict a win, loss, or draw for Orlando City Soccer

- Points Scored
- Passing Accuracy
- Corners
- Shots
- Clearances
- Total Crosses
- Possession
- Total Passes
- Goal Keeper Saves
- Blocked Shots
- Possession

Key Players Seen in Analysis to be relevant in predictions:

Duncan McGuire
#13
Forward
Orlando City
Senior
Games Played: 27
Minutes Played: 1599

Joined Orlando City Soccer in December 2022

Methodology

Insignificant Variables:

- Home or Away
- Shots on Goal
- Offsides
- Aerial Duels Won
- Fouls
- Yellow Cards
- Red Cards
- Team
- Season
- Facundo Torres
- Ramiro Enrique
- Martin Ojeda
- Nicolas Lodeiro
- Ivan Angulo
- OP Team
- Game Day

Player with more time and games, but not deemed as significant:

Facundo Torres
#10
Forward
Orlando City
Senior
Games Played: 32
Minutes Played: 2642
Transferred December 2024

Most likely related to when he got traded into the team despite his stats being higher than Duncan's. Left after the 2024 season ended.

Classification Models in Python

- We also evaluated the data on four classifications in Python using 'sklearn':
    - Logistic Regression
    - Naïve Bayes
    - Support Vector Machine (SVM)
    - Random Forest

Full Models

```
--- Naive Bayes Model with All Predictors ---
Accuracy on the test set: 0.545

Classification Report:
              precision    recall  f1-score   support

        Draw       0.25      0.33      0.29         3
        Loss       0.33      0.33      0.33         3
         Win       1.00      0.80      0.89         5

    accuracy                           0.55        11
   macro avg       0.53      0.49      0.50        11
weighted avg       0.61      0.55      0.57        11
```

```
--- Random Forest Model with All Predictors ---
Accuracy on the test set: 0.455

Classification Report:
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.29      0.67      0.40         3
         Win       0.75      0.60      0.67         5

    accuracy                           0.45        11
   macro avg       0.35      0.42      0.36        11
weighted avg       0.42      0.45      0.41        11
```

```
--- Logistic Regression Model with All Predictors ---
Accuracy on the test set: 0.545
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.43      1.00      0.60         3
         Win       1.00      0.60      0.75         5

    accuracy                           0.55        11
   macro avg       0.48      0.53      0.45        11
weighted avg       0.57      0.55      0.50        11
```

```
--- Support Vector Machine (SVM) Model with All Predictors ---
Accuracy on the test set: 0.545

Classification Report:
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.40      0.67      0.50         3
         Win       0.67      0.80      0.73         5

    accuracy                           0.55        11
   macro avg       0.36      0.49      0.41        11
weighted avg       0.41      0.55      0.47        11
```

Reduced Model: 7 Features

```
--- Evaluation with Top 7 Predictors ---

--- Evaluating Logistic Regression with Reduced Predicto
Accuracy with 7 predictors: 0.364
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.33      0.67      0.44         3
         Win       0.50      0.40      0.44         5

    accuracy                           0.36        11
   macro avg       0.28      0.36      0.30        11
weighted avg       0.32      0.36      0.32        11


--- Evaluating Naive Bayes with Reduced Predictors ---
Accuracy with 7 predictors: 0.545
              precision    recall  f1-score   support

        Draw       0.50      0.33      0.40         3
        Loss       0.50      0.33      0.40         3
         Win       0.57      0.80      0.67         5

    accuracy                           0.55        11
   macro avg       0.52      0.49      0.49        11
weighted avg       0.53      0.55      0.52        11
```

```
--- Evaluating SVM with Reduced Predictors ---
Accuracy with 7 predictors: 0.455
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.25      0.33      0.29         3
         Win       0.57      0.80      0.67         5

    accuracy                           0.45        11
   macro avg       0.27      0.38      0.32        11
weighted avg       0.33      0.45      0.38        11


--- Evaluating Random Forest with Reduced Predictors ---
Accuracy with 7 predictors: 0.455
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.33      0.67      0.44         3
         Win       0.60      0.60      0.60         5

    accuracy                           0.45        11
   macro avg       0.31      0.42      0.35        11
weighted avg       0.36      0.45      0.39        11
```

Reduced Model: 5 Features

```
--- Evaluation with Top 5 Predictors ---

--- Evaluating Logistic Regression with Reduced Predictors ---
Accuracy with 5 predictors: 0.364
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.29      0.67      0.40         3
         Win       0.50      0.40      0.44         5

    accuracy                           0.36        11
   macro avg       0.26      0.36      0.28        11
weighted avg       0.31      0.36      0.31        11


--- Evaluating Naive Bayes with Reduced Predictors ---
Accuracy with 5 predictors: 0.636
              precision    recall  f1-score   support

        Draw       1.00      0.33      0.50         3
        Loss       0.50      0.67      0.57         3
         Win       0.67      0.80      0.73         5

    accuracy                           0.64        11
   macro avg       0.72      0.60      0.60        11
weighted avg       0.71      0.64      0.62        11
```

```
--- Evaluating SVM with Reduced Predictors ---
Accuracy with 5 predictors: 0.455
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.25      0.33      0.29         3
         Win       0.57      0.80      0.67         5

    accuracy                           0.45        11
   macro avg       0.27      0.38      0.32        11
weighted avg       0.33      0.45      0.38        11


--- Evaluating Random Forest with Reduced Predictors ---
Accuracy with 5 predictors: 0.455
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.43      1.00      0.60         3
         Win       0.50      0.40      0.44         5

    accuracy                           0.45        11
   macro avg       0.31      0.47      0.35        11
weighted avg       0.34      0.45      0.37        11
```

Reduced Model: 3 Features

```
--- Evaluation with Top 3 Predictors ---

--- Evaluating Logistic Regression with Reduced Predictors ---
Accuracy with 3 predictors: 0.455
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.33      0.67      0.44         3
         Win       0.60      0.60      0.60         5

    accuracy                           0.45        11
   macro avg       0.31      0.42      0.35        11
weighted avg       0.36      0.45      0.39        11


--- Evaluating Naive Bayes with Reduced Predictors ---
Accuracy with 3 predictors: 0.636
              precision    recall  f1-score   support

        Draw       1.00      0.33      0.50         3
        Loss       0.50      0.67      0.57         3
         Win       0.67      0.80      0.73         5

    accuracy                           0.64        11
   macro avg       0.72      0.60      0.60        11
weighted avg       0.71      0.64      0.62        11
```

```
--- Evaluating SVM with Reduced Predictors ---
Accuracy with 3 predictors: 0.545
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.40      0.67      0.50         3
         Win       0.67      0.80      0.73         5

    accuracy                           0.55        11
   macro avg       0.36      0.49      0.41        11
weighted avg       0.41      0.55      0.47        11


--- Evaluating Random Forest with Reduced Predictors ---
Accuracy with 3 predictors: 0.364
              precision    recall  f1-score   support

        Draw       0.00      0.00      0.00         3
        Loss       0.17      0.33      0.22         3
         Win       0.60      0.60      0.60         5

    accuracy                           0.36        11
   macro avg       0.26      0.31      0.27        11
weighted avg       0.32      0.36      0.33        11
```

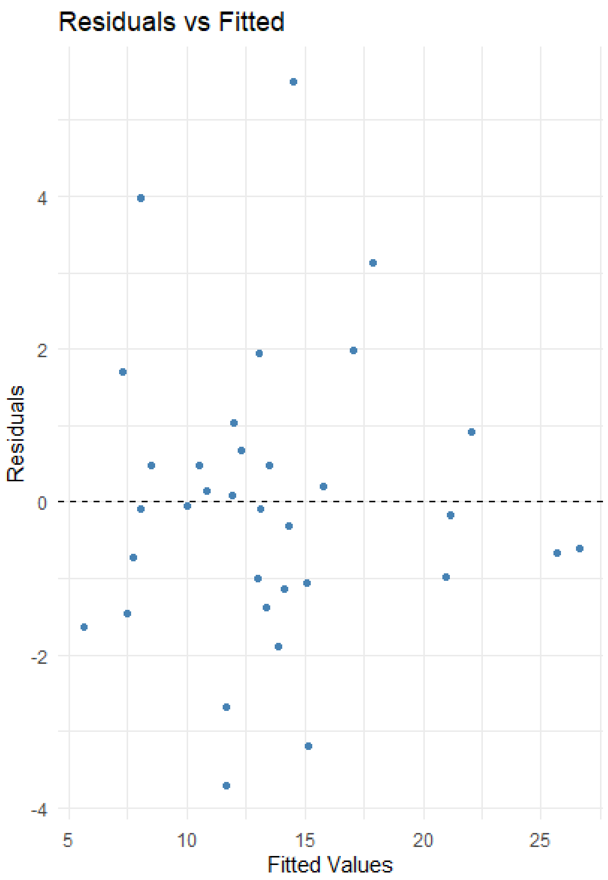
Model Diagnostics

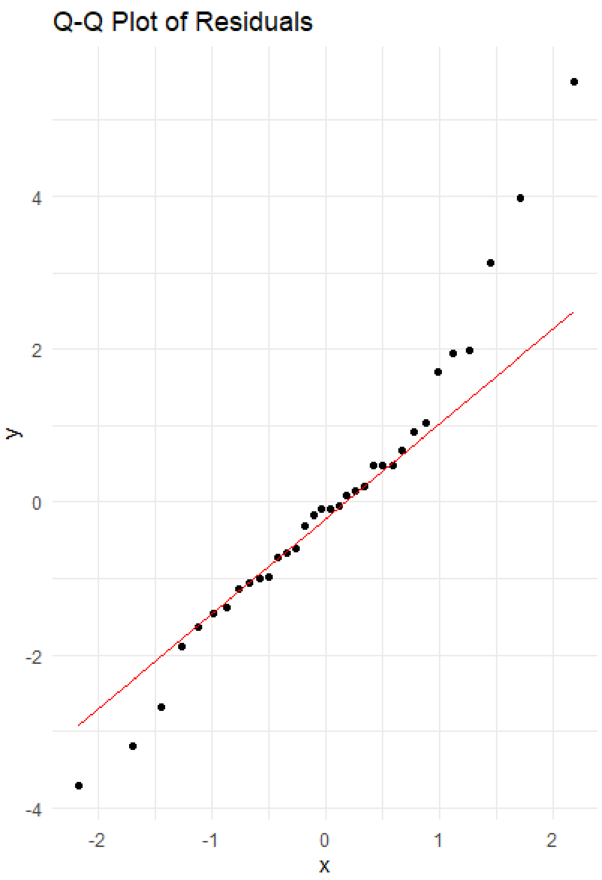
Classification Models Used:
- XGBoost
- Decision Trees

Other Methods Used:
- Hierarchical Clustering

Methods Used to evaluate performance and accuracy of models:
- Variable Importance
- Training and Testing Accuracy
- Confusion Matrix

Regression Model Diagnostics:
- Residual Plots Randomly Scattered
- Q-Q Plot normally distributed
- VIF (Variance Inflation Factor to check for multicollinearity





Could be open to a transformation for the final model if a regression model is used based on Q-Q plot

Model Diagnostics

```
Confusion Matrix and Statistics

          Reference
Prediction FALSE TRUE
     FALSE    32    0
     TRUE      0    2

               Accuracy : 1
                 95% CI : (0.8972, 1)
    No Information Rate : 0.9412
    P-Value [Acc > NIR] : 0.1273

                  Kappa : 1

 Mcnemar's Test P-Value : NA

            Sensitivity : 1.0000
            Specificity : 1.0000
         Pos Pred Value : 1.0000
         Neg Pred Value : 1.0000
             Prevalence : 0.9412
         Detection Rate : 0.9412
   Detection Prevalence : 0.9412
      Balanced Accuracy : 1.0000

       'Positive' Class : FALSE
```

Logistic Regression Model:

- Confusion matrix has 100% accuracy
- P-value of .1273 shows high performance, but dataset is small

Why is this no good?

- Likely to be overfitting
- May not generalize well to new data

```
> print(vif_values)
      Possession    Shots_On_Goal
        3.858655         1.259225
   Blocked_Shots     Total_Passes
        3.011039         2.399375
Passing_Accuracy          Corners
        2.962442         2.156599
   Total_Crosses  Aerial_Duels_Won
        2.824983         2.103487
 Goalkeeper_Saves      Clearances
        1.446675         1.508471
           Fouls
        1.504515
```

All VIF scores are under the value 5 which suggests there are no multicollinearity issues in the dataset

Outlier Detection and Removal

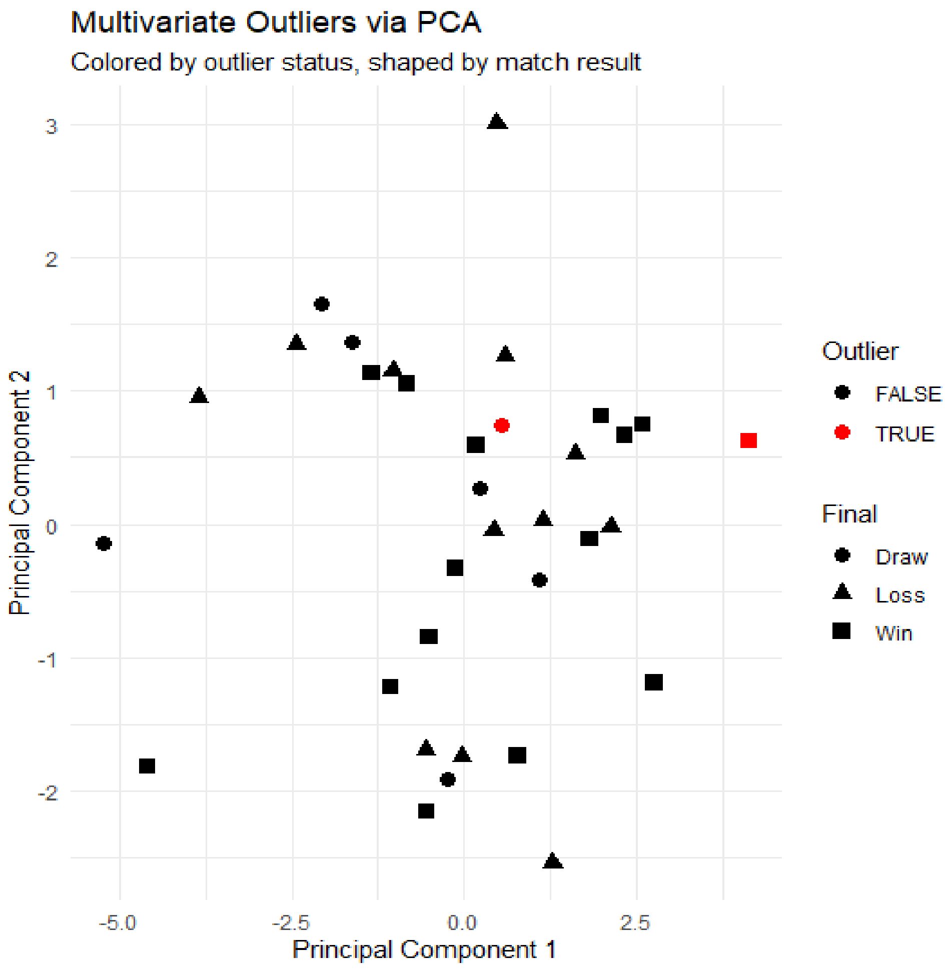
Due to smaller size of the dataset, the two outlier games were used

To find the outliers, we calculated a distance score based on all key match stats together.

By visualizing the outliers with PCA, we can compare the matches based on if it was a win, loss, or draw

Computed with Mahalanobis Distance:
measures distance relative to the centroid — a base or central point which can be thought of as an overall mean for multivariate data

```
Player Presence in Outlier Games:
> print(player_outliers)
# A tibble: 2 x 5
  Game_Day OP_Team  Facundo_Torres Duncan_McGuire Ramiro_Enrique
  <fct>    <fct>    <chr>          <chr>          <chr>
1 30-Apr   Montreal Y              Y              N
2 22-Jun   Chicago  Y              Y              N
```



Outlier Detection and Removal (continued)

- **Shots:** Unusually high in games at indices **17** and **21**
- **Blocked Shots:** Spike at game **21**
- **Total Passes:** Very high count in game **8**
- **Total Crosses:** High in games **16** and **17**
- **Offsides:** Surged in games **0**, **13**, and **29**
- **Goalkeeper Saves:** Exceptionally high in game **18**
- **Red Cards:** Rare events in games **2** and **9**

•**Why Keep the Outliers?**

•Reflect **real match events** (e.g., aggressive plays, defensive pressure, or red card scenarios)

•May indicate **tactical shifts** or **game-changing moments**

•Removing them could **hide important patterns**, especially with a **small dataset**

•Will be considered in **interpretation and modeling**, not discarded

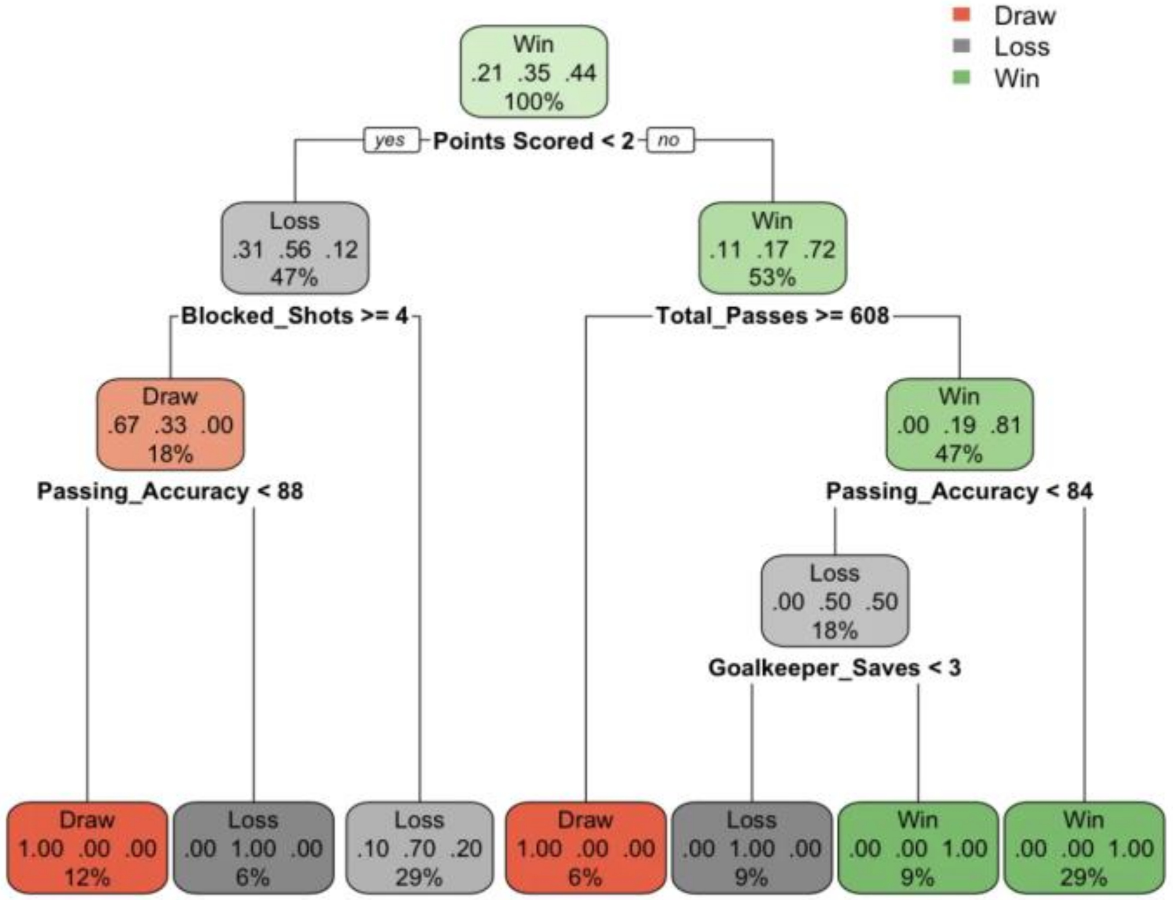
```
Outliers for Possession are located at indices: Index([], dtype='int64')
Outliers for Shots are located at indices: Index([17, 21], dtype='int64')
Outliers for Shots_On_Goal are located at indices: Index([], dtype='int64')
Outliers for Blocked_Shots are located at indices: Index([21], dtype='int64')
Outliers for Total_Passes are located at indices: Index([8], dtype='int64')
Outliers for Passing_Accuracy are located at indices: Index([], dtype='int64')
Outliers for Corners are located at indices: Index([], dtype='int64')
Outliers for Total_Crosses are located at indices: Index([16, 17], dtype='int64')
Outliers for Offsides are located at indices: Index([0, 13, 29], dtype='int64')
Outliers for Aerial_Duels_Won are located at indices: Index([], dtype='int64')
Outliers for Goalkeeper_Saves are located at indices: Index([18], dtype='int64')
Outliers for Clearances are located at indices: Index([], dtype='int64')
Outliers for Fouls are located at indices: Index([], dtype='int64')
Outliers for Yellow_Cards are located at indices: Index([], dtype='int64')
Outliers for Red_Cards are located at indices: Index([2, 9], dtype='int64')
Outliers for Points Scored are located at indices: Index([], dtype='int64')
Detected outliers in the dataset:
   Final Home_Away  Possession  Shots  Shots_On_Goal  Blocked_Shots  \
0   Draw         H        57.7     16              4              5
2   Loss         H        55.0     23              8              4
8   Draw         A        46.7      9              4              2
9   Loss         H        46.2      9              4              3
13  Loss         H        42.6     12              5              4
16  Loss         H        55.7     19              1              8
17  Draw         A        60.5     25              9              7
18   Win         H        42.4      7              5              0
21   Win         H        55.6     26             10             10
29  Loss         A        35.4     12              4              4

    Total_Passes  Passing_Accuracy  Corners  Total_Crosses  ...  \
0            579              86.5        9             25  ...
2            451              79.6        8              9  ...
8            746              80.9        8              6  ...
9            362              80.7        5             13  ...
13           435              88.5        9             12  ...
16           627              92.5        9             31  ...
17           654              93.3       11             28  ...
18           382              83.8        1              1  ...
21           561              91.3       12             15  ...
29           382              83.5        3              5  ...

    Facundo_Torres  Duncan_McGuire  Ramiro_Enrique  Martín_Ojeda  \
0                Y               Y               Y             Y
2                N               Y               Y             Y
8                Y               Y               N             Y
9                Y               Y               N             Y
13               Y               N               Y             Y
16               Y               Y               N             Y
17               Y               Y               N             Y
18               Y               Y               N             Y
21               Y               Y               Y             Y
29               Y               Y               Y             Y

    \tNicolás_Lodeiro  Iván_Angulo      OP_Team   Game_Day Points Scored    Win
0                   Y            Y     Montreal 2024-02-24             0  False
2                   Y            Y    Minnesota 2024-03-09             2  False
8                   Y            Y     Montreal 2024-04-30             2  False
9                   Y            Y  Cincinnati 2024-05-04             0  False
13                  Y            Y     Columbus 2024-05-25             0  False
16                  Y            Y         LAFC 2024-06-15             1  False
17                  Y            Y    Charlotte 2024-06-19             2  False
18                  Y            Y      Chicago 2024-06-22             4   True
21                  Y            Y         D.C. 2024-07-06             5   True
29                  Y            Y     Columbus 2024-09-21             3  False
--------------------------------------------------------------------------------
```

Final Model



Draw
Loss
Win
Win
.21 .35 .44
100%
yes
Points Scored < 2
no
Loss
.31 .56 .12
47%
Blocked_Shots >= 4
Win
.11 .17 .72
53%
Total_Passes >= 608
Draw
.67 .33 .00
18%
Passing_Accuracy < 88
Win
.00 .19 .81
47%
Passing_Accuracy < 84
Loss
.00 .50 .50
18%
Goalkeeper_Saves < 3
Draw
1.00 .00 .00
12%
Loss
.00 1.00 .00
6%
Loss
.10 .70 .20
29%
Draw
1.00 .00 .00
6%
Loss
.00 1.00 .00
9%
Win
.00 .00 1.00
9%
Win
.00 .00 1.00
29%

Results & Summary

- Overall Accuracy: 91.2%
- Kappa: 0.862
  - Almost perfect agreement between predictions and actual values
  - P-value is statistically significant at 1.08e-08

Statistics By Class:
- Draws and Wins are predicted with high precision and specificity
- Losses have perfect sensitivity, but a slightly lower precision
- Only 3 misclassifications

```
Confusion Matrix and Statistics

          Reference
Prediction Draw Loss Win
      Draw    6    0   0
      Loss    1   12   2
      Win     0    0  13

Overall Statistics

               Accuracy : 0.9118
                 95% CI : (0.7632, 0.9814)
    No Information Rate : 0.4412
    P-Value [Acc > NIR] : 1.082e-08

                  Kappa : 0.862

 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: Draw Class: Loss Class: Win
Sensitivity               0.8571      1.0000     0.8667
Specificity               1.0000      0.8636     1.0000
Pos Pred Value            1.0000      0.8000     1.0000
Neg Pred Value            0.9643      1.0000     0.9048
Prevalence                0.2059      0.3529     0.4412
Detection Rate            0.1765      0.3529     0.3824
Detection Prevalence      0.1765      0.4412     0.3824
Balanced Accuracy         0.9286      0.9318     0.9333
```

Variable Importance:
Top 5 Influential Predictors in Tree
- Passing Accuracy
- Total Passes
- Points Scored
- Corners
- Shots on Goal

```
Passing_Accuracy    Total_Passes   Points_Scored
        7.470833        4.811887        4.691993
           Shots      Possession          Season
        2.929167        2.417443        2.091248
        Offsides
        0.625000

         Corners   Shots_On_Goal Goalkeeper_Saves
        4.572222        4.052747         3.929167
 Aerial_Duels_Won           Fouls        Home_Away
        1.333333        1.333333         1.172998
                                     Blocked_Shots
                                          3.858333
                                       Clearances
                                          0.625000
```

Interpretations

- Based on the top 5 important variables, the most important factors in a game is:
- Ball Control, Aggressiveness and Attacking Pressure, and Offensive Strength are most influential in determining a win, loss or draw
- Other factors that influence a match show the balance of attack and defense in a game such as
  - Goal Keeper Saves
  - Blocked Shots
  - Possession

Study Limitations

- Though this study highlighted significant factors in match outcome limitations remain.
- Additional predictor possibilities:
  - Injury
    - During match: Orlando City or opposing team
    - Prior to match
    - Resulting from match
  - Complete team roster
    - All members playing in match
    - Duration member played in match
    - Members that started/ ended the match
    - Career length in general
    - Career length at Orlando City

Work Cited

- "A New Metric for Evaluating 1v1 Ability." *Stats Perform*, www.statsperform.com/resource/a-new-metric-for-evaluating-1v1-ability/.
- "Glossary of Soccer Terms." *Marcus Jewish Community Center of Atlanta*, www.atlantajcc.org/clientuploads/PDF/SF/YLC/SOCCER/Fall%20Soccer/GlossarySoccerTerms.pdf.
- "Opta Event Definitions." *Stats Perform*, www.statsperform.com/opta-event-definitions/.
- "Opta Football Stats Definitions." *The Analyst*, www.theanalyst.com/2024/07/opta-football-stats-definitions/.
- "Orlando City SC Player: Duncan McGuire." *Orlando City Soccer Club*, www.orlandocitysc.com/players/duncan-mcguire/.
- "Orlando City SC Player: Facundo Torres." *Orlando City Soccer Club*, www.orlandocitysc.com/players/facundo-torres/.
- "Orlando City SC Player: Iván Angulo." *Orlando City Soccer Club*, www.orlandocitysc.com/players/ivan-angulo/.
- "Orlando City SC Player: Martín Ojeda." *Orlando City Soccer Club*, www.orlandocitysc.com/players/martin-ojeda/.
- "Orlando City SC Player: Nicolás Lodeiro." *Orlando City Soccer Club*, www.orlandocitysc.com/players/nicolas-lodeiro/.
- "Orlando City SC Player: Ramiro Enrique." *Orlando City Soccer Club*, www.orlandocitysc.com/players/ramiro-enrique/.
- "Orlando City SC Schedule." *Orlando City Soccer Club*, www.orlandocitysc.com/schedule/.
- "Orlando City SC Stats." *Orlando City Soccer Club*, www.orlandocitysc.com/stats/.

Work Cited

- "Orlando City SC Team Statistics." *ESPN*, www.espn.com/soccer/team/stats/_/id/12011/league/USA.1/season/2024/view/scoring.
- "Seasons in Florida." *Florida Smart*, www.floridasmart.com/seasons-in-florida.
- "Soccer Statistics Manual." *National Collegiate Athletic Association*, fs.ncaa.org/Docs/stats/Stats_Manuals/Soccer/2009ez.pdf.
- "Soccer Terms." *US Soccer Players*, www.ussoccerplayers.com/soccer-terms.
- "The History of Soccer." *Bundesliga*, www.bundesliga.com/en/faq/all-you-need-to-know-about-soccer/the-history-of-soccer-10560.
- "U.S. Soccer History Timeline." *U.S. Soccer*, www.ussoccer.com/history/timeline.
- "Weather in Orlando." *TripSavvy*, www.tripsavvy.com/weather-in-orlando-2546241.
- McKinley, Eliot. "Reinventing the Passing Wheel: What Determines a Good Passer?" *American Soccer Analysis*, www.americansocceranalysis.com/home/2018/10/11/whatdeterminesagoodpasser.
- Scarosella, Vita. "The Arms Race for the Hearts and Minds of American Soccer Fans." *Forbes*, 6 Aug. 2024, https://www.forbes.com/sites/vitascarosella/2024/08/06/the-arms-race-for-the-hearts-and-minds-of-american-soccer-fans/.