# Individual Stats on Game Outcome

Collin Travasos

# Hypothesis

Can we use a single players stat line to predict the outcome of the game?

Can we use a single statistic from a player to predict the outcome of the game?

Can coaches use this information to adjust their gameplan?

Can gamblers use this information to take advantage of sports betting agencies or vice versa?

# Data

- Single game box score stats
  - Team Stats
    - Score
    - Starting Lineup
    - Team Totals
  - Player Stats
    - PTS, REB, AST, etc
    - Minutes
- 2015-16 to 2023-24

# Data Manipulation

- Combine Player & Team Datasets
  - using game id
- Create New Columns
  - win/loss
  - team - game - player
- Remove Columns
  - duplicates
  - misc
- Adjust Names

- Remove Blowouts
  - games decided by 25 or more
- Remove Bench Players
  - starters only
  - played more than 12 minutes
- Only All-Stars?
  - too limited applications
  - only small increase in accuracy

# What’s missing

- Defense
- Defense
- Defense

# Models

# Random Forest Models - All Players

```
Confusion Matrix and Statistics

          Reference
Prediction Loss  Win
      Loss 6485 5529
      Win  4769 5779

               Accuracy : 0.5436
                 95% CI : (0.537, 0.5501)
    No Information Rate : 0.5012
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.0873

 Mcnemar's Test P-Value : 7.468e-14

            Sensitivity : 0.5762
            Specificity : 0.5111
         Pos Pred Value : 0.5398
         Neg Pred Value : 0.5479
             Prevalence : 0.4988
         Detection Rate : 0.2874
   Detection Prevalence : 0.5325
      Balanced Accuracy : 0.5436

       'Positive' Class : Loss
```

```
Confusion Matrix and Statistics

          Reference
Prediction Loss  Win
      Loss 7306 6307
      Win  3948 5001

               Accuracy : 0.5455
                 95% CI : (0.5389, 0.552)
    No Information Rate : 0.5012
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.0914

 Mcnemar's Test P-Value : < 2.2e-16

            Sensitivity : 0.6492
            Specificity : 0.4423
         Pos Pred Value : 0.5367
         Neg Pred Value : 0.5588
             Prevalence : 0.4988
         Detection Rate : 0.3238
   Detection Prevalence : 0.6034
      Balanced Accuracy : 0.5457

       'Positive' Class : Loss
```

# Other Models - Lebron James

Logistic Regression - Gradient Boosting - K Nearest Neighbor

```
Confusion Matrix and Statistics

          Reference
Prediction Loss Win
      Loss   15  13
      Win    36  71

               Accuracy : 0.637
                 95% CI : (0.5499, 0.718)
    No Information Rate : 0.6222
    P-Value [Acc > NIR] : 0.397677

                  Kappa : 0.1529

 Mcnemar's Test P-Value : 0.001673

            Sensitivity : 0.2941
            Specificity : 0.8452
         Pos Pred Value : 0.5357
         Neg Pred Value : 0.6636
             Prevalence : 0.3778
         Detection Rate : 0.1111
   Detection Prevalence : 0.2074
      Balanced Accuracy : 0.5697

       'Positive' Class : Loss
```

```
Confusion Matrix and Statistics

          Reference
Prediction Loss Win
      Loss   16  14
      Win    35  70

               Accuracy : 0.637
                 95% CI : (0.5499, 0.718)
    No Information Rate : 0.6222
    P-Value [Acc > NIR] : 0.397677

                  Kappa : 0.16

 Mcnemar's Test P-Value : 0.004275

            Sensitivity : 0.3137
            Specificity : 0.8333
         Pos Pred Value : 0.5333
         Neg Pred Value : 0.6667
             Prevalence : 0.3778
         Detection Rate : 0.1185
   Detection Prevalence : 0.2222
      Balanced Accuracy : 0.5735

       'Positive' Class : Loss
```

```
Confusion Matrix and Statistics

          Reference
Prediction Loss Win
      Loss   21  25
      Win    30  59

               Accuracy : 0.5926
                 95% CI : (0.5047, 0.6763)
    No Information Rate : 0.6222
    P-Value [Acc > NIR] : 0.7885

                  Kappa : 0.1164

 Mcnemar's Test P-Value : 0.5896

            Sensitivity : 0.4118
            Specificity : 0.7024
         Pos Pred Value : 0.4565
         Neg Pred Value : 0.6629
             Prevalence : 0.3778
         Detection Rate : 0.1556
   Detection Prevalence : 0.3407
      Balanced Accuracy : 0.5571

       'Positive' Class : Loss
```

# Lebron James RF Predictions

```
Confusion Matrix and Statistics

          Reference
Prediction  0  1
         0 15 11
         1 36 73

               Accuracy : 0.6519
                 95% CI : (0.5651, 0.7317)
    No Information Rate : 0.6222
    P-Value [Acc > NIR] : 0.2687482

                  Kappa : 0.1806

 Mcnemar's Test P-Value : 0.0004639

            Sensitivity : 0.2941
            Specificity : 0.8690
         Pos Pred Value : 0.5769
         Neg Pred Value : 0.6697
             Prevalence : 0.3778
         Detection Rate : 0.1111
   Detection Prevalence : 0.1926
      Balanced Accuracy : 0.5816

       'Positive' Class : 0
```

# Lebron James

"We can predict with 65.19 % accuracy that the win probability for the team when LeBron James has an excellent game is: 68.8 %, a good game is: 69.6 %, an average game is: 42 %, a bad game is: 52 %, and a terrible game is: 42.2 %"

# Stephen Curry

"We can predict with 75.78 % accuracy that the win probability for the team when Stephen Curry has an excellent game is: 76.2 %, a good game is: 86.6 %, an average game is: 66.2 %, a bad game is: 65 %, and a terrible game is: 66.2 %"

# Klay Thompson

"We can predict with 74.36 % accuracy that the win probability for the team when Klay Thompson has an excellent game is: 75.2 %, a good game is: 85.2 %, an average game is: 74.4 %, a bad game is: 63.4 %, and a terrible game is: 26.4 %"

# Draymond Green

"We can predict with 64.44 % accuracy that the win probability for the team when Draymond Green has an excellent game is: 62.4 %, a good game is: 95.6 %, an average game is: 79.4 %, a bad game is: 37.6 %, and a terrible game is: 31.2 %"

# As Requested:

# Josh Hart

"We can predict with 52.27 % accuracy that the win probability for the team when Josh Hart has an excellent game is: 39.8 %, a good game is: 57.2 %, an average game is: 30.2 %, a bad game is: 22.2 %, and a terrible game is: 67.2 %"

# TJ McConnell

"We can predict with 31.58 % accuracy that the win probability for the team when T.J. McConnell has an excellent game is: 29.6 %, a good game is: 18 %, an average game is: 47.6 %, a bad game is: 48.2 %, and a terrible game is: 25.6 %"

# Grayson Allen

"We can predict with 58 % accuracy that the win probability for the team when Grayson Allen has an excellent game is: 58.2 %, a good game is: 65.4 %, an average game is: 72.4 %, a bad game is: 70 %, and a terrible game is: 53.8 %"

# Wemby

"We can predict with 84.62 % accuracy that the win probability for the team when Victor Wembanyama has an excellent game is: 46.8 %, a good game is: 54.8 %, an average game is: 7.6 %, a bad game is: 7 %, and a terrible game is: 13.8 %"

# Luka Doncic

"We can predict with 53.93 % accuracy that the win probability for the team when Luka Dončić has an excellent game is: 46.4 %, a good game is: 78 %, an average game is: 59.4 %, a bad game is: 24.8 %, and a terrible game is: 28.4 %"

# John Wall

"We can predict with 52.63 % accuracy that the win probability for the team when John Wall has an excellent game is: 43.2 %, a good game is: 43.2 %, an average game is: 43 %, a bad game is: 34.2 %, and a terrible game is: 41.2 %"

# Mo Bamba

"We can predict with 66.67 % accuracy that the win probability for the team when Mo Bamba has an excellent game is: 28.8 %, a good game is: 7.8 %, an average game is: 20.6 %, a bad game is: 55.8 %, and a terrible game is: 41 %"

# Tyler Herro

"We can predict with 60 % accuracy that the win probability for the team when Tyler Herro has an excellent game is: 57.8 %, a good game is: 76.4 %, an average game is: 56 %, a bad game is: 37 %, and a terrible game is: 26.8 %"

# Joe Ingles

"We can predict with 58.62 % accuracy that the win probability for the team when Joe Ingles has an excellent game is: 55.6 %, a good game is: 76.2 %, an average game is: 72.6 %, a bad game is: 55.6 %, and a terrible game is: 12.2 %"

# Pat Connaughton

"We can predict with 53.85 % accuracy that the win probability for the team when Pat Connaughton has an excellent game is: 47 %, a good game is: 59.2 %, an average game is: 61.8 %, a bad game is: 43 %, and a terrible game is: 35.6 %"

# Conclusions

- Playing within designated role is important
  - Stars can take over and win a game
  - Stars doing too much can be detrimental
  - Stars need to show up
  - Role players are needed
  - Some players should focus more on handshakes and cheering
- Basketball is a team sport

# Next Steps

- Add The Opposition
  - team and individual
- Alter Specific Stats
  - maintain averages but increase points, shooting percentage, etc.
- Include Teammates
  - “big 3” success

# Applications

## Coaching

- Pre Game Planning
  - player target stats
- In Game Adjustments
  - increase / decrease efforts

## Betting

- Live Betting
  - player pace
- Parlays
  - winning / losing metrics