Predicting NBA All-Team Selections

By: Kaley Wong

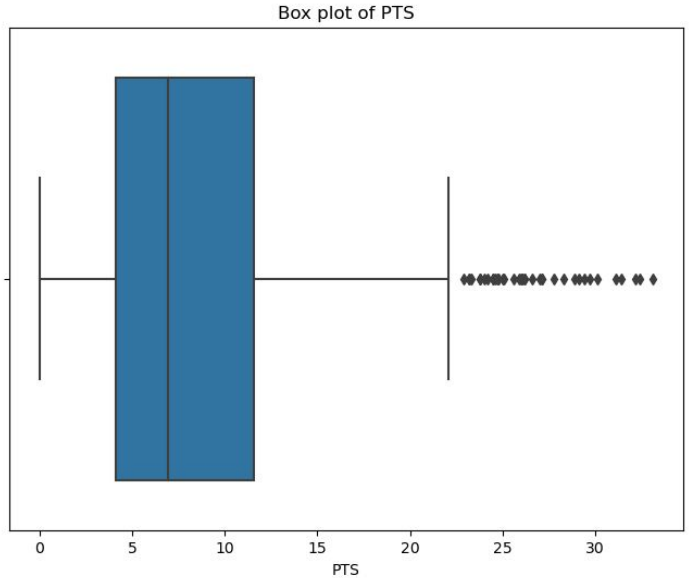
Introduction

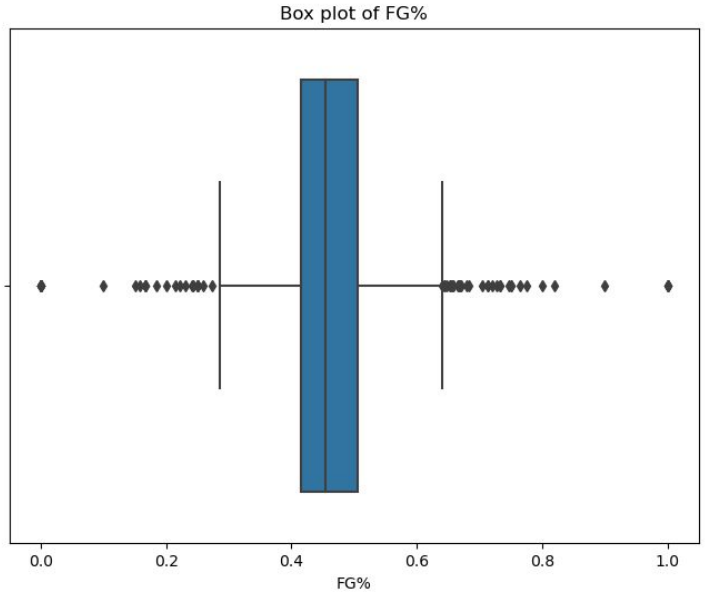
- All-NBA teams honor the league’s top players for each season
- By looking at a player's stats, we can help determine who will be chosen for these teams

Problem Statement

- Aim of this project is to develop predictive models to accurately forecast which NBA players are likely to be selected for the All-NBA Teams
- This is significance for team management, player recognition, and strategic decision-making in basketball

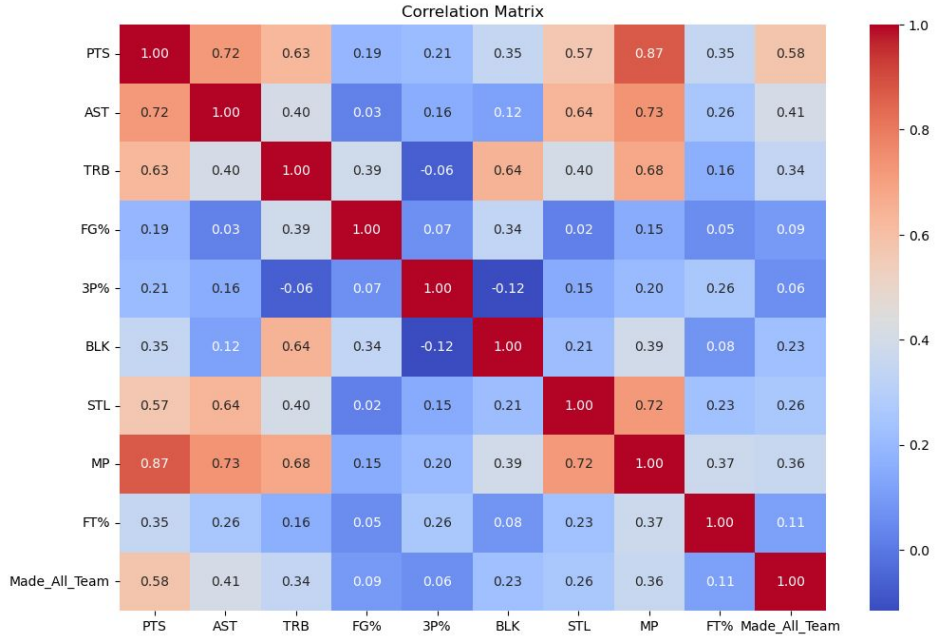
Exploratory Data Analysis





- Both graphs consist of higher value outliers, NBA player who has outstanding stats may be considered an outlier

Exploratory Data Analysis



- Which performance metrics had the highest correlations with Made All Team?
  - 0.58: Points per Game
  - 0.41: Assist per Game
  - 0.36: Minutes Played

Model Selection

Methods Used:

- Logistic Regression
- Random Forest Classifier
- Gradient Boosting Machine

Features Used:

- Points per game (PTS)
- Assists per game (AST)
- Rebounds per game (TRB)
- Field goal percentage (FG%)
- Three-point field goal percentage (3P%)
- Blocks per game (BLK)
- Steals per game (STL)
- Minutes played per game (MP)
- Free throw percentage (FT%)

Logistic Regression

```
Accuracy: 0.9629629629629629
Classification Report:
              precision    recall  f1-score   support

         0.0       0.97      0.99      0.98       104
         1.0       0.50      0.25      0.33         4

    accuracy                           0.96       108
   macro avg       0.74      0.62      0.66       108
weighted avg       0.95      0.96      0.96       108

Confusion Matrix:
 [[103   1]
 [  3   1]]
ROC AUC Score: 0.6201923076923077
```

- 6/108 Incorrect Predictions
- Performs well in Accuracy and ROC AUC Score
- Lower Precision and Recall

Random Forest Classifier

```
Accuracy: 0.9814814814814815
Classification Report:
              precision    recall  f1-score   support

         0.0       0.99      0.99      0.99       104
         1.0       0.75      0.75      0.75         4

    accuracy                           0.98       108
   macro avg       0.87      0.87      0.87       108
weighted avg       0.98      0.98      0.98       108

Confusion Matrix:
 [[103   1]
 [  1   3]]
ROC AUC Score: 0.8701923076923078
```

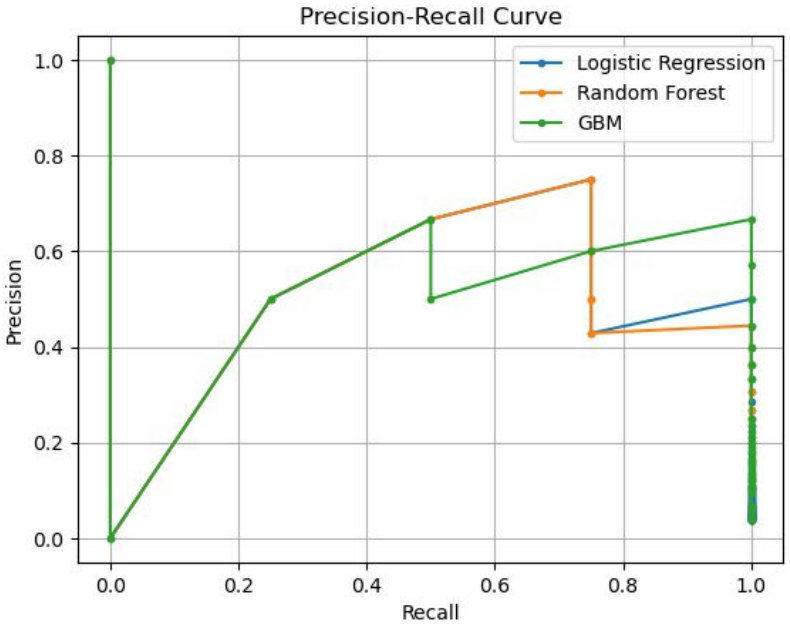
- 8/108 Incorrect Predictions
- Performs well in Accuracy and ROC AUC Score
- Has limitations in Predicting Class 1

Gradient Boosting Machine

```
Accuracy: 0.9722222222222222
Classification Report:
              precision    recall  f1-score   support

         0.0       0.99      0.98      0.99       104
         1.0       0.60      0.75      0.67         4

    accuracy                           0.97       108
   macro avg       0.80      0.87      0.83       108
weighted avg       0.98      0.97      0.97       108

Confusion Matrix:
 [[102   2]
 [  1   3]]
ROC AUC Score: 0.8653846153846153
```

- 9/108 Incorrect Predictions
- Performs well in Accuracy and ROC AUC Score
- Lower in Precision, Recall and F1-Score

Model Selection



- Random Forest performs the best to predicting NBA All-Team selection
  - Highest accuracy, precision, recall, F1-Score, and ROC AUC score

Conclusion

- Uncovered key performance metrics such as points per game and field goal percentage as influential factors in predicting NBA All-Team placements
- Recognized the importance of acknowledging model limitations and need to explore additional factors to make better predictions for NBA All-Team selections