Using Machine Learning to Analyze NFL Penalty Data

Leo DiPerna, Eric Uehling

https://en.wikipedia.org/wiki/National_Football_League

Introduction

- Studying effects of different penalties on the outcomes of drives
- Studying which referee crews tend to call which penalties more often than others
- Finding insights that fans, teams, and analysts will find interesting and valuable
- Steps:
  - Data collection - Pro Football Reference
  - Data cleaning
  - Data analysis

PRO FOOTBALL REFERENCE

https://www.pro-football-reference.com/

Related Works

[1] Blaikie, Andrew D., et al. "NFL & NCAA Football Prediction Using Artificial Neural Networks." Proceedings of the Midstates Conference for Undergraduate Research in Computer Science and Mathematics, Denison University, Granville, OH. 2011.

[2] Anyama, Oscar Uzoma, and Chinwe Peace Igiri. "An Application of Linear Regression & Artificial Neural Network Model in the NFL Result Prediction." (2015).

[3] Snyder, Kevin, and Michael Lopez. "Consistency, Accuracy, and Fairness: A Study of Discretionary Penalties in the NFL." Journal of Quantitative Analysis in Sports 11.4 (2015): 219-230.

[4] Craig, Curtis, et al. "A Relationship Between Temperature and Aggression in NFL Football Penalties." Journal of Sport and Health Science, vol. 5, no. 2, 21 June 2016, pp. 205–210, https://doi.org/10.1016/j.jshs.2015.01.001.

[5] McDaniel, Zachary. NFL Penalty Analysis, Referee Influence and Penalty Trends Over Time. Diss. University of Iowa, 2021.

Data Scraping

San Francisco 49ers	Kansas City Chiefs
22	25
14-6	15-6
« Prev Game	« Prev Game
Coach: Kyle Shanahan	**Coach**: Andy Reid

Sunday Feb 11, 2024
Start Time: 6:30pm
Stadium: Allegiant Stadium
Attendance: 61,629
Time of Game: 4:06
Logos via Sports Logos.net / About logos

	1	2	3	4	OT	Final
San Francisco 49ers	0	10	0	9	3	22
Kansas City Chiefs	0	3	10	6	6	25

```
game_id,team_id,num,quarter,time,los,plays,length,net_yds,result
2009_1_TEN_PIT,PIT,1,1,15:00,PIT 42,3,1:44,2,Punt
2009_1_TEN_PIT,PIT,2,1,11:24,TEN 43,5,3:04,2,Punt
2009_1_TEN_PIT,PIT,3,1,6:44,PIT 27,3,1:55,-6,Punt
2009_1_TEN_PIT,PIT,4,1,1:38,PIT 21,3,1:32,3,Interception
2009_1_TEN_PIT,PIT,5,2,13:04,PIT 5,9,5:50,33,Punt
2009_1_TEN_PIT,PIT,6,2,2:14,PIT 21,5,0:52,79,Touchdown
2009_1_TEN_PIT,PIT,7,2,0:48,PIT 27,6,0:48,18,Interception
2009_1_TEN_PIT,PIT,8,3,11:57,PIT 46,3,2:04,9,Punt
```

```
Penalty,Date,Opp,Player,Pos,Week,Ref Crew,Quarter,Time,Down,Dist,Declined,Offsetting,Yardage,Home,Phase,Team,Year
Offensive Pass Interference,09/13/2009,San Francisco,L.Fitzgerald,WR,1,Don Carey,1,09:43,2,9,No,No,10,Yes,Off,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,M.Gandy,T,1,Don Carey,1,09:13,2,19,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,A.Wilson,SS,1,Don Carey,1,07:23,2,2,No,No,5,Yes,Def,arizona-cardinals,2009
Low Block,09/13/2009,San Francisco,R.Johnson,FS,1,Don Carey,1,05:34,0,0,No,No,8,Yes,ST,arizona-cardinals,2009
Unnecessary Roughness,09/13/2009,San Francisco,A.Wilson,SS,1,Don Carey,1,02:55,1,10,No,No,15,Yes,Def,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,A.Becht,TE,1,Don Carey,1,00:38,1,10,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,D.Dockett,DT,1,Don Carey,2,12:14,1,10,No,No,5,Yes,Def,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,A.Highsmith,ILB,1,Don Carey,2,05:36,4,2,No,No,5,Yes,ST,arizona-cardinals,2009
Delay of Game,09/13/2009,San Francisco,K.Warner,QB,1,Don Carey,3,01:32,2,4,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,D.Dockett,DT,1,Don Carey,4,07:39,2,7,No,No,4,Yes,Def,arizona-cardinals,2009
```

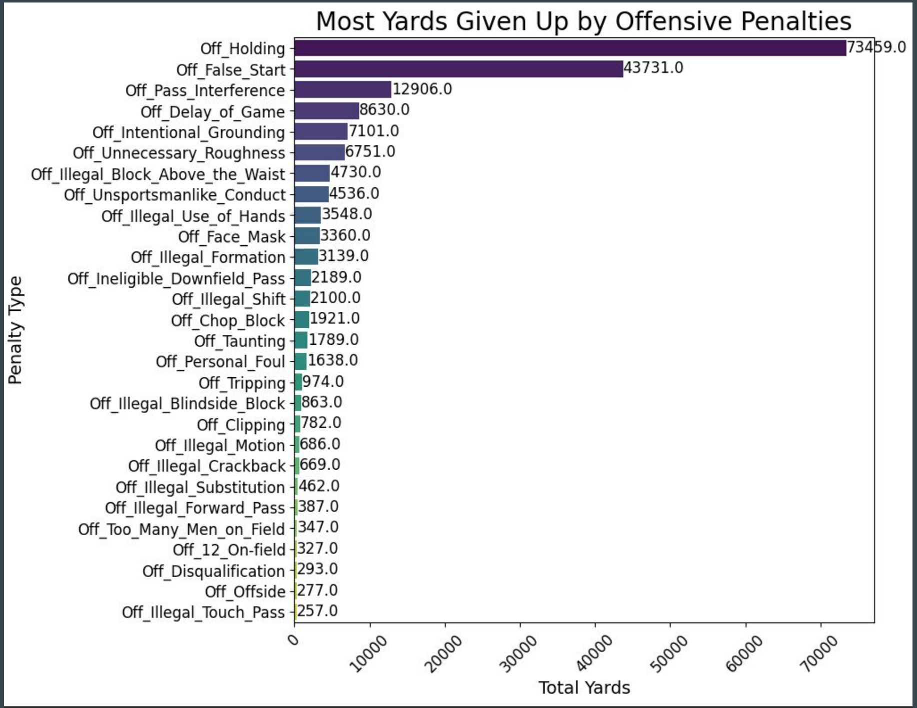
Data Cleaning

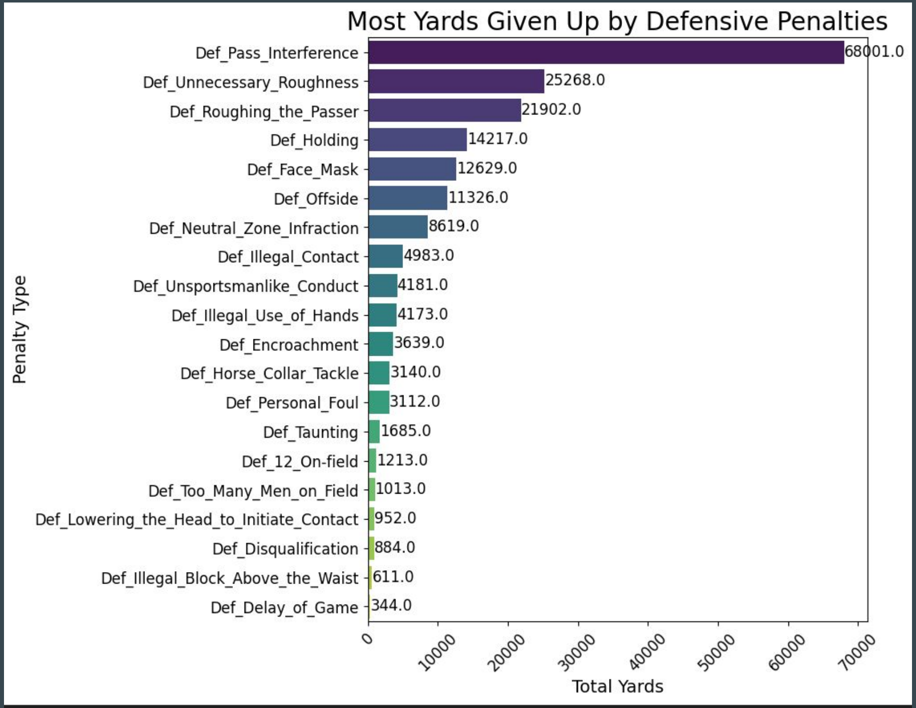
```
Penalty,Date,Opp,Player,Pos,Week,Ref Crew,Quarter,Time,Down,Dist,Declined,Offsetting,Yardage,Home,Phase,Team,Year
Offensive Pass Interference,09/13/2009,San Francisco,L.Fitzgerald,WR,1,Don Carey,1,09:43,2,9,No,No,10,Yes,Off,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,M.Gandy,T,1,Don Carey,1,09:13,2,19,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,A.Wilson,SS,1,Don Carey,1,07:23,2,2,No,No,5,Yes,Def,arizona-cardinals,2009
Low Block,09/13/2009,San Francisco,R.Johnson,FS,1,Don Carey,1,05:34,0,0,No,No,8,Yes,ST,arizona-cardinals,2009
Unnecessary Roughness,09/13/2009,San Francisco,A.Wilson,SS,1,Don Carey,1,02:55,1,10,No,No,15,Yes,Def,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,A.Becht,TE,1,Don Carey,1,00:38,1,10,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,D.Dockett,DT,1,Don Carey,2,12:14,1,10,No,No,5,Yes,Def,arizona-cardinals,2009
False Start,09/13/2009,San Francisco,A.Highsmith,ILB,1,Don Carey,2,05:36,4,2,No,No,5,Yes,ST,arizona-cardinals,2009
Delay of Game,09/13/2009,San Francisco,K.Warner,QB,1,Don Carey,3,01:32,2,4,No,No,5,Yes,Off,arizona-cardinals,2009
Defensive Offside,09/13/2009,San Francisco,D.Dockett,DT,1,Don Carey,4,07:39,2,7,No,No,4,Yes,Def,arizona-cardinals,2009
```

```
game_id,team_id,num,quarter,time,los,plays,length,net_yds,result
2009_1_TEN_PIT,PIT,1,1,15:00,PIT 42,3,1:44,2,Punt
2009_1_TEN_PIT,PIT,2,1,11:24,TEN 43,5,3:04,2,Punt
2009_1_TEN_PIT,PIT,3,1,6:44,PIT 27,3,1:55,-6,Punt
2009_1_TEN_PIT,PIT,4,1,1:38,PIT 21,3,1:32,3,Interception
2009_1_TEN_PIT,PIT,5,2,13:04,PIT 5,9,5:50,33,Punt
2009_1_TEN_PIT,PIT,6,2,2:14,PIT 21,5,0:52,79,Touchdown
2009_1_TEN_PIT,PIT,7,2,0:48,PIT 27,6,0:48,18,Interception
2009_1_TEN_PIT,PIT,8,3,11:57,PIT 46,3,2:04,9,Punt
```

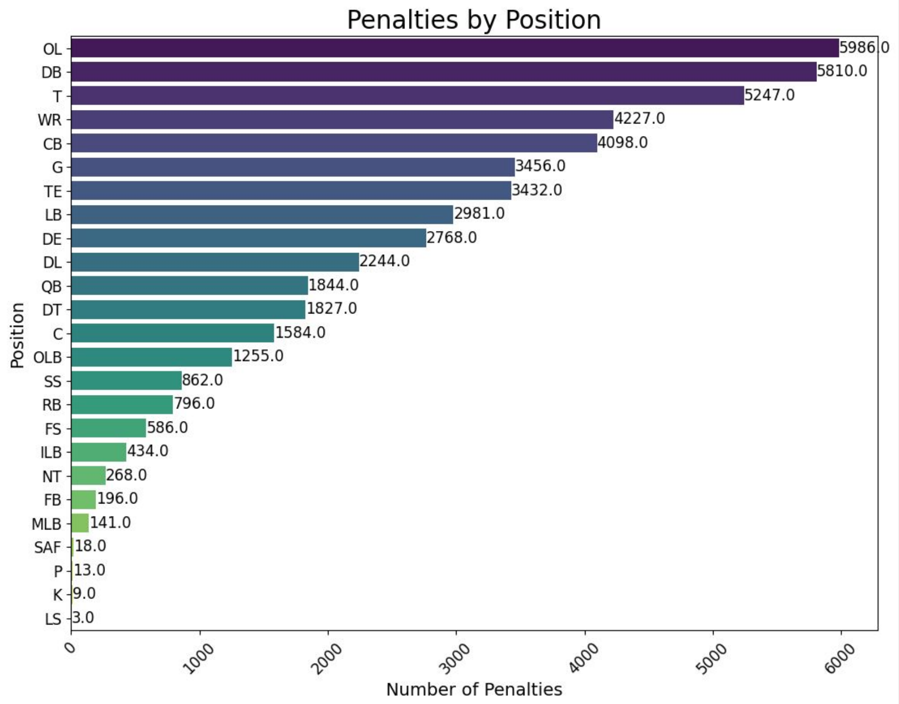
```
game_id,team_id,num,quarter,time,los,plays,length,net_yds,result,time_left,date,Def_Unnecessary_Roughness,Off_Illegal_Format
2009_1_TEN_PIT,PIT,1,1,15:00,58,3,1:44,2,Punt,01:00:00,2009-09-10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
2009_1_TEN_PIT,TEN,1,1,13:16,98,3,1:52,2,Punt,00:58:16,2009-09-10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
2009_1_TEN_PIT,PIT,2,1,11:24,43,5,3:04,2,Punt,00:56:24,2009-09-10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
2009_1_TEN_PIT,TEN,2,1,8:20,89,6,1:36,70,Missed FG,00:53:20,2009-09-10,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
2009_1_TEN_PIT,PIT,3,1,6:44,73,3,1:55,-6,Punt,00:51:44,2009-09-10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
2009_1_TEN_PIT,TEN,3,1,4:49,74,6,3:11,30,Interception,00:49:49,2009-09-10,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
```

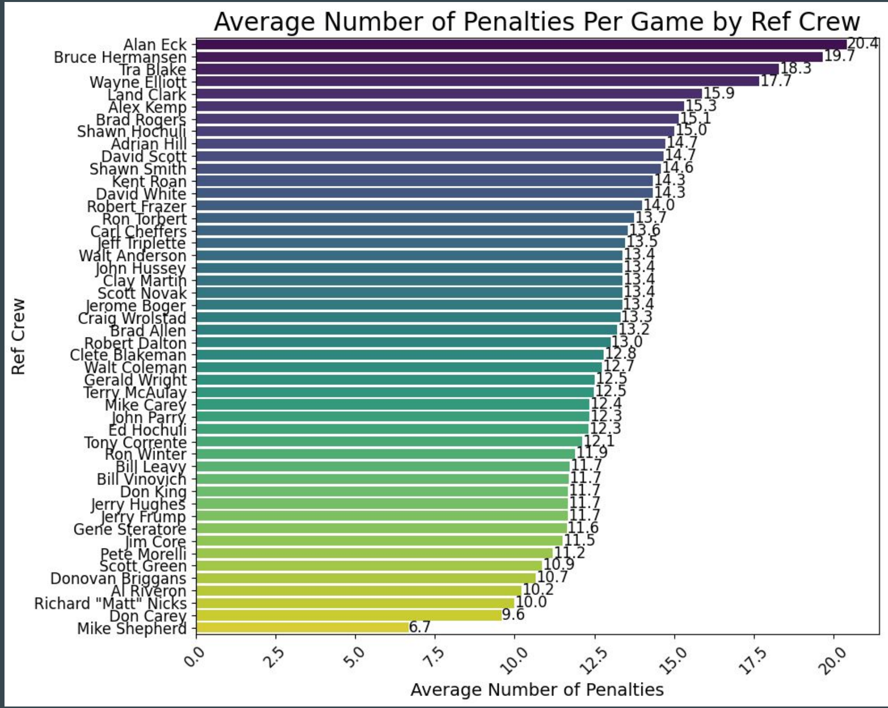
Exploratory Data Analysis



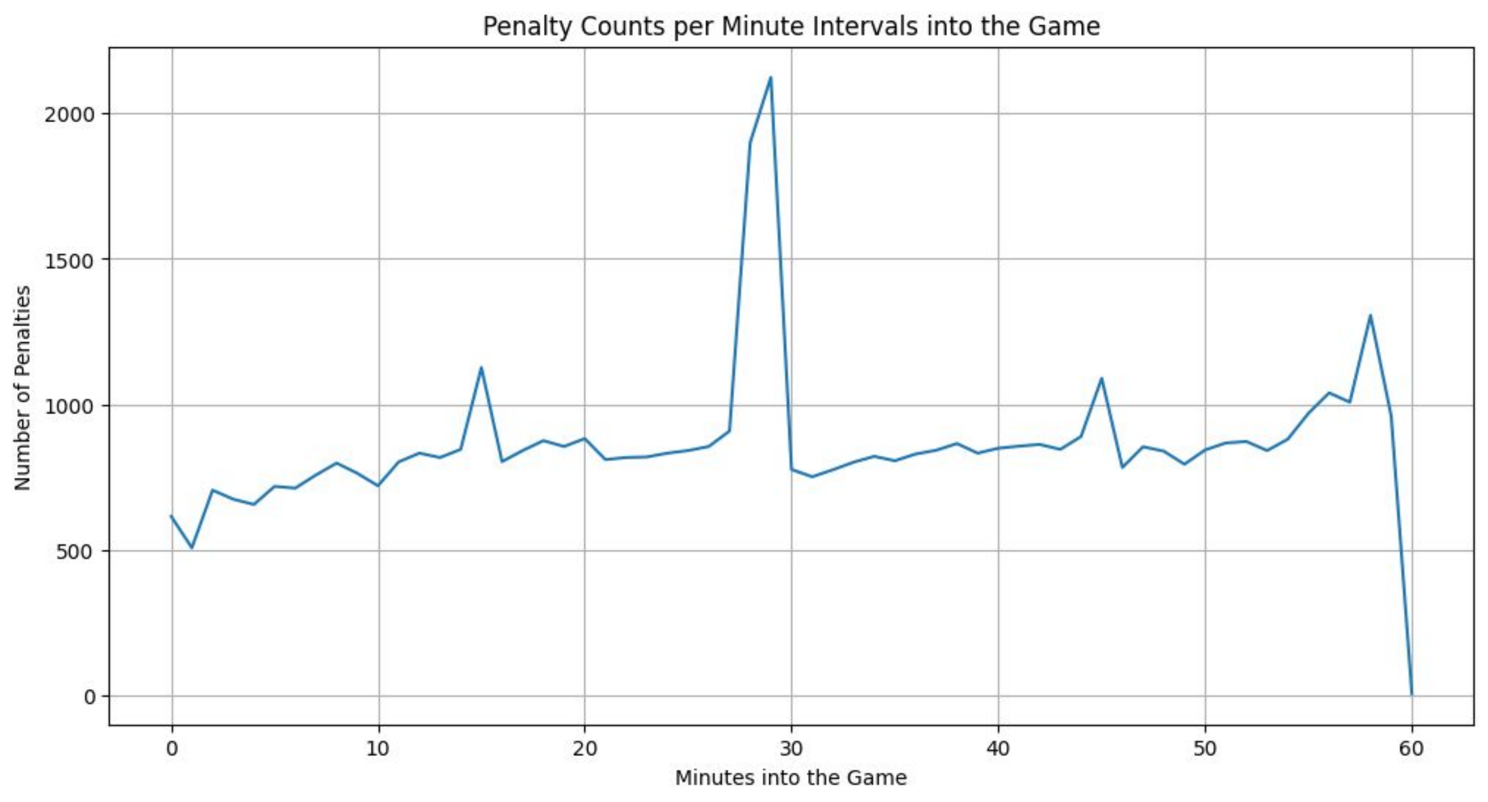


Exploratory Data Analysis, Pt. 2





Exploratory Data Analysis, Pt. 3



Drive's Model - Predicting Drive Outcomes

```python
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.metrics import mean_squared_error, r2_score

# Initialize the Gradient Boosting Regressor
gbm_regressor = GradientBoostingRegressor(n_estimators=100, learning_rate=0.1, max_depth=3, random_state=42)

# Train the regressor
gbm_regressor.fit(X_train_reg, y_train_reg)

# Predicting the test set results
y_pred_reg = gbm_regressor.predict(X_test_reg)

# Evaluate the regression model
mse = mean_squared_error(y_test_reg, y_pred_reg)
r2 = r2_score(y_test_reg, y_pred_reg)
print("Mean Squared Error:", mse)
print("R^2 Score:", r2)
```

✓ 2.5s

```
Mean Squared Error: 6.88506327387569
R^2 Score: 0.13576086162277534
```

```python
def predict_points(total_off_pen, total_def_pen, total_off_pen_yard
    input_features = pd.DataFrame([{
        'total_off_pen': total_off_pen,
        'total_def_pen': total_def_pen,
        'total_off_pen_yards': total_off_pen_yards,
        'total_def_pen_yards': total_def_pen_yards,
        'los': los,
        'time_left_seconds': time_left_seconds
    }])
    predicted_points = gbm_regressor.predict(input_features)
    return predicted_points[0]

# Example usage of the prediction function for regression
points = predict_points(0, 3, 0, 30, 25, 800)
print(f"Predicted Points: {points}")
```

✓ 0.0s

```
Predicted Points: 5.046858987798735
```

3 Def Penalties, 30 Yards, 800 seconds

0 Def Penalties, 0 Yards, 25 los

Drive's Model - Predicting Drive Outcomes, Pt. 2

```
# Initialize the Gradient Boosting Classifier
gbm_model = GradientBoostingClassifier(n_estimators=100, learning_rate=0.1, max_depth=3, random_state=42)

# Train the model
gbm_model.fit(X_train, y_train)

# Predicting the test set results
y_pred = gbm_model.predict(X_test)

# Evaluate the model by specifying all labels known to the LabelEncoder
classification_report_result = classification_report(y_test, y_pred, labels=np.arange(len(le.classes_)), t
print("Classification Report:\n", classification_report_result)

# Extracting and printing the feature importance
feature_importance = gbm_model.feature_importances_
features_df = pd.DataFrame({'Feature': features.columns, 'Importance': feature_importance}).sort_values(by
print("\nFeature Importance:\n", features_df)
```

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

  Field Goal       0.43      0.02      0.03      4034
   Touchdown       0.54      0.21      0.30      5857
        Zero       0.68      0.97      0.80     18001

    accuracy                           0.67     27892
   macro avg       0.55      0.40      0.38     27892
weighted avg       0.62      0.67      0.58     27892


Feature Importance:
               Feature  Importance
1       total_def_pen    0.429765
4                 los    0.242568
5   time_left_seconds    0.187329
3  total_def_pen_yards    0.080356
2  total_off_pen_yards    0.031086
0       total_off_pen    0.028895
```

Penalty's Model - Predicting Number

```python
# Define predictor variables
predictors = ['team_id', 'opp_id', 'year', 'week', 'ref_crew', 'home', 'postseason']

# Identify the top 5 most common penalty types
top_penalty_codes = df_grouped['penalty'].value_counts().nlargest(5).index.tolist()

# Store MSE and R-squared values for overall evaluation
overall_mse = []
overall_r2 = []

# Train and evaluate a model for each of the top 5 penalty types
for penalty_code in top_penalty_codes:
    # Filter data for the current penalty type
    df_penalty = df_grouped[df_grouped['penalty'] == penalty_code]

    # Model formula
    formula = 'count ~ ' + ' + '.join(predictors)

    # Fit the model with explicit alpha to avoid warnings
    model = glm(formula, data=df_penalty, family=NegativeBinomial(alpha=1.0)).fit()
    models[penalty_code] = model
```

```python
def predict_penalties(team_id, opp_id, year, week, ref_crew, home, postseason):
    # Encode input data
    input_data = {
        'team_id': label_encoders['team_id'].transform([team_id])[0],
        'opp_id': label_encoders['opp_id'].transform([opp_id])[0],
        'year': year,
        'week': week,
        'ref_crew': label_encoders['ref_crew'].transform([ref_crew])[0],
        'home': label_encoders['home'].transform([home])[0],
        'postseason': label_encoders['postseason'].transform([postseason])[0]
    }

    predictions = {}
    for penalty_code, model in models.items():
        features_df = pd.DataFrame([input_data])
        predicted_count = model.predict(features_df)[0]
        penalty_type = label_encoders['penalty'].inverse_transform([penalty_code])[0]
        predictions[penalty_type] = max(0, predicted_count)  # Ensure non-negative predictions

    return pd.DataFrame([predictions], index=['Predicted Count'])

# Example prediction
predict_penalties('DAL', 'SEA', 2023, 10, 'Bill Leavy', 'Yes', 'No')
```

✓ 0.0s

	Off_Holding	Off_False_Start	Def_Pass_Interference	Def_Holding	Def_Offside
Predicted Count	2.174033	1.920161	1.610521	1.691706	1.573771

Use Cases - What do the Models Uncover?

- Sports Betting
  - Prop Bets
  - Regular Bets
- Team Strategy
  - Deciding Schemes
  - Deciding Targets

Conclusion and Future Work

- We have found some interesting insights on the effect of penalties on games
- Some penalties tend to have a much bigger impact on games than others
- We intend to create some additional models:
    - The effect of penalties on team performance over the course of a game
    - The effect of penalties on the final standings of a season

Questions?