Identifying the Blitz Before the Snap

Andrea Wright

SLM 638

09 May 2023

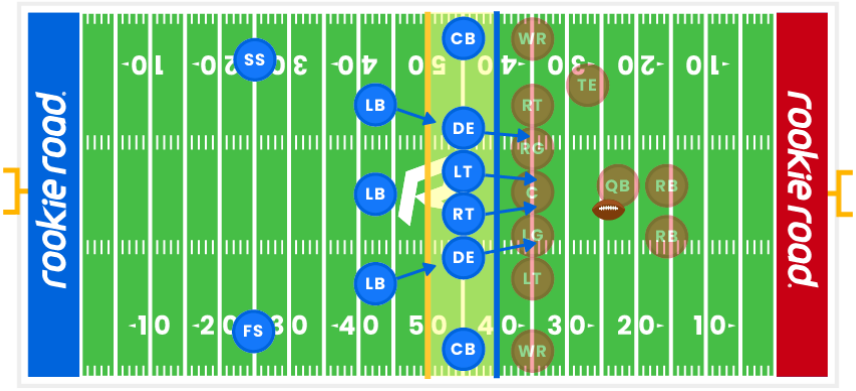
Abstract

The quarterback blitz is a risky but potentially worthwhile defensive maneuver in American football. This research seeks to evaluate pre-snap characteristics such as distance, location, time, and personnel of NFL teams on pass plays for their likelihood to indicate a quarterback blitz. A random forest model reported with 77% accuracy that zone coverage, increased number of defenders on the line of scrimmage and Cover-0 defenses were most associated with a blitz. This research is limited in that it reports what is “common knowledge” regarding a blitz and lacks depth in presenting new or novel information. Further research may present novel information if applied to a particular team or utilized GPS data more thoroughly.

I. Introduction

In American football, the defensive blitz is a defensive tactic used to disrupt the offense by sending more than a usual number of defensive players into the offensive backfield. A typical play involves only the linemen, defensive ends and tackles, encroaching into the offensive backzone defined by the line of scrimmage on the offensive side as pictured in Figure 1.

Figure 1
Illustration of a defensive blitz play



Football Blitz

Note. In this blitz, the two extra players are the outside linebackers, LB. In total, six defensive players are entering the offensive backzone, or the area behind the blue line of scrimmage.
Image Source: Rookie Road

With a blitz, linebackers or defensive backs may rush the offensive backzone A blitz may be used to sack a quarterback or force them to make rash decisions when throwing the ball. Although a risky tactic, the blitz can be an effective strategy in high-pressure situations.

II. Literature Review

If machine learning methods are being used by NFL staff to predict or evaluate opponents for a blitz, those are trade secrets. Some research exists to suggest that data can be used to predict play-calling. A master's thesis from the Massachusetts Institute of Technology applied machine learning models to play-by-play data to predict a run versus a pass. Research compared logistic regression, a neural network,

decision tree, and random forest to evaluate run versus pass plays based on pre-snap characteristics. The best model had a test accuracy of 80% (Goyal, 2019). Although not directly related to this research, this thesis suggests that play-by-play data can be effectively used to predict action in American football. Additionally, the NFL's Big Data Bowl (the source of this dataset) encourages participants to use play-by-play data to evaluate gameplay. A large amount of research exists that evaluates the results of games based on broad, contextual variables such as weather, indoor versus outdoor stadium, coaching changes, etc. Nwachukwu and Uzoma implemented a linear regression model and k-nearest neighbor model predicted NFL game results based on these factors. This research was, however, fairly superficial in that it evaluated using vague performance indicators such as team record that provide little insight into actual performance (Nwachukwu & Uzoma, 2015).

III. Hypothesis Development

This research seeks to evaluate defensive strategy in American football. More specifically, this research seeks to answer the question: Is it possible to determine or anticipate a blitz based on pre-snap characteristics?

IV. Method

a. Data

Data was collected from the 2022 Big Data Bowl hosted by Kaggle. This dataset included pass plays from 8 weeks of the regular 2021 season and included GPS data provided by the NFL and play-by-play data provided by Pro Football Focus.

b. Analytical Procedures

A blitz is defined as five or more defensive players rushing into the offensive backfield. Prior to data analysis, defining the blitz by evaluating GPS data was required. First, the offensive backfield was defined as any point behind the line of scrimmage. For each play, the number of defensive players crossing behind this point was counted. If that number was greater than or equal to five, the play was denoted as a blitz. This procedure was followed for every play in the dataset.

Variables that referred to items occurring after the play ended were eliminated. These include items such as playResult (pass complete or incomplete), penalties, or play description (e.g. Joe Burrow pass complete short right to J'Marr Chase). Offensive and defensive personnel were provided. Offensive personnel was limited to include only the number of running backs, tight ends, or wide receivers. Defensive personnel included only defensive linemen, linebackers, or defensive backs. A final list of variables evaluated can be found in *IX. Appendix.*

c. Exploratory Data Analysis and Feature Importances

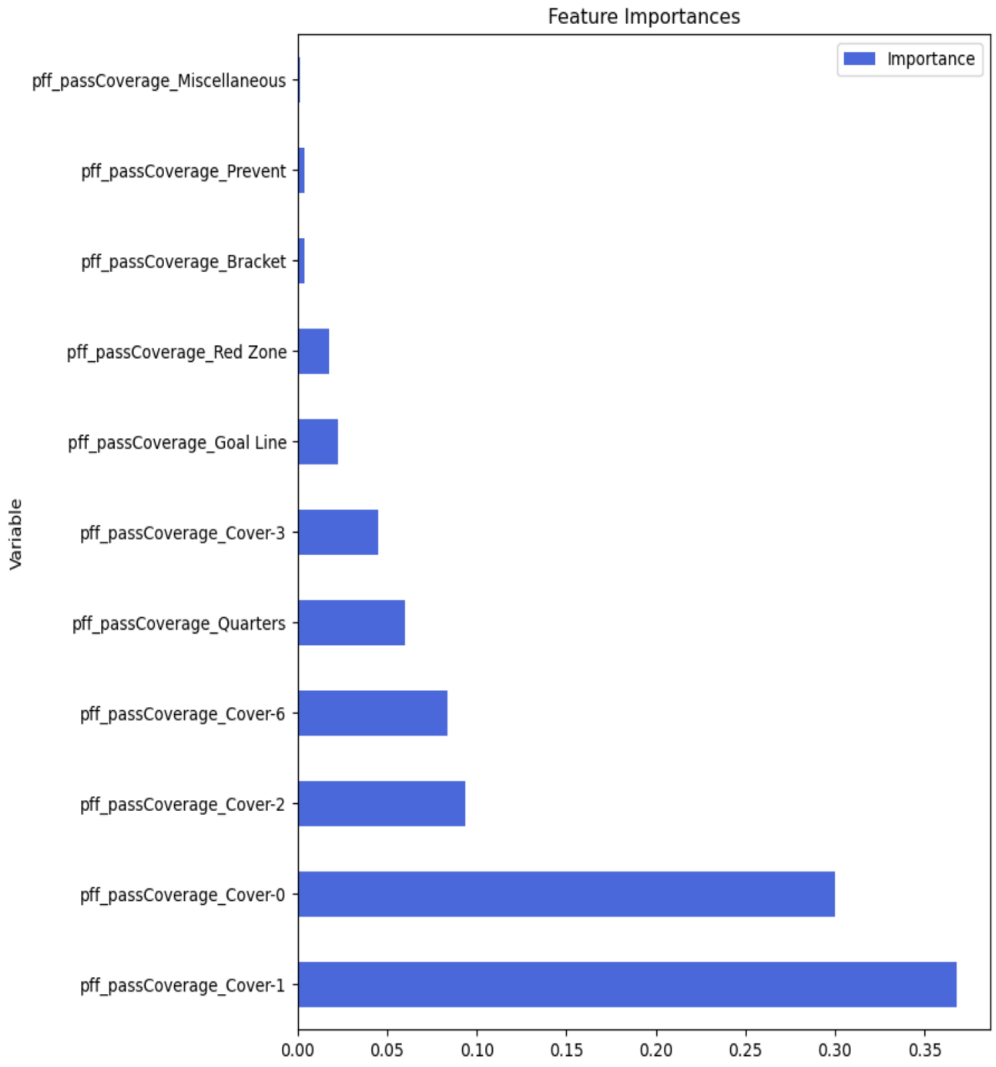
A significant amount of variables were both categorical and contained over ten degrees of freedom. An initial study was run without eliminating these degrees of freedom and resulted in algorithms that either did not converge or were largely overfit. As a result, pruning via feature importances was included to improve machine learning.

The variables possessionTeam and defensiveTeam were evaluated for feature importances. Because the research question seeks to evaluate all NFL teams, this was conducted as an exploratory data analysis rather than as a method of pruning.

Feature importances evaluated using the amount of "spread" between two variables to determine the cutoff threshold. For example in Figure 2, the difference between `pff_passCoverage_Cover-2` and `pff_passCoverage_Cover-0` is displayed. `pff_passCoverage_Cover-2` has a feature importance of 0.10, which is less than 50% of the next highest value, `pff_passCoverage_Cover-0` at 0.29. This method was utilized for all remaining feature importance evaluations.

Figure 2

Feature Importances for categorical variable `pff_passCoverage`



Offensive and defensive personnel were evaluated and found that just two offensive and five defensive formations could be best associated with a blitz.

Table 1
Variables modified by feature importance evaluation

Variable	Original DOF	Modified DOF
personnelO	21	2
personnelD	27	5
pff_passCoverage	10	2

Correlation was evaluated prior to algorithm selection. A threshold of 75% was set. The variables `numBackfield` and `isBlitz` were eliminated. This set of variables suggests that the number of players in the backfield is closely associated with a blitz. `numBackfield` was a variable created when evaluating whether a blitz occurred, so this is a logical correlation. `numBackfield` is dropped. Next, the variables `pff_passCoverage_Cover-1` and `pff_passCoverageType_Zone` were beyond the threshold. Because Zone coverage is a more generalized variable, `pff_passCoverage_Cover-1` is dropped.

Data was scaled and split into 70% training and 30% test datasets prior to algorithm implementation.

d. Algorithm Selection

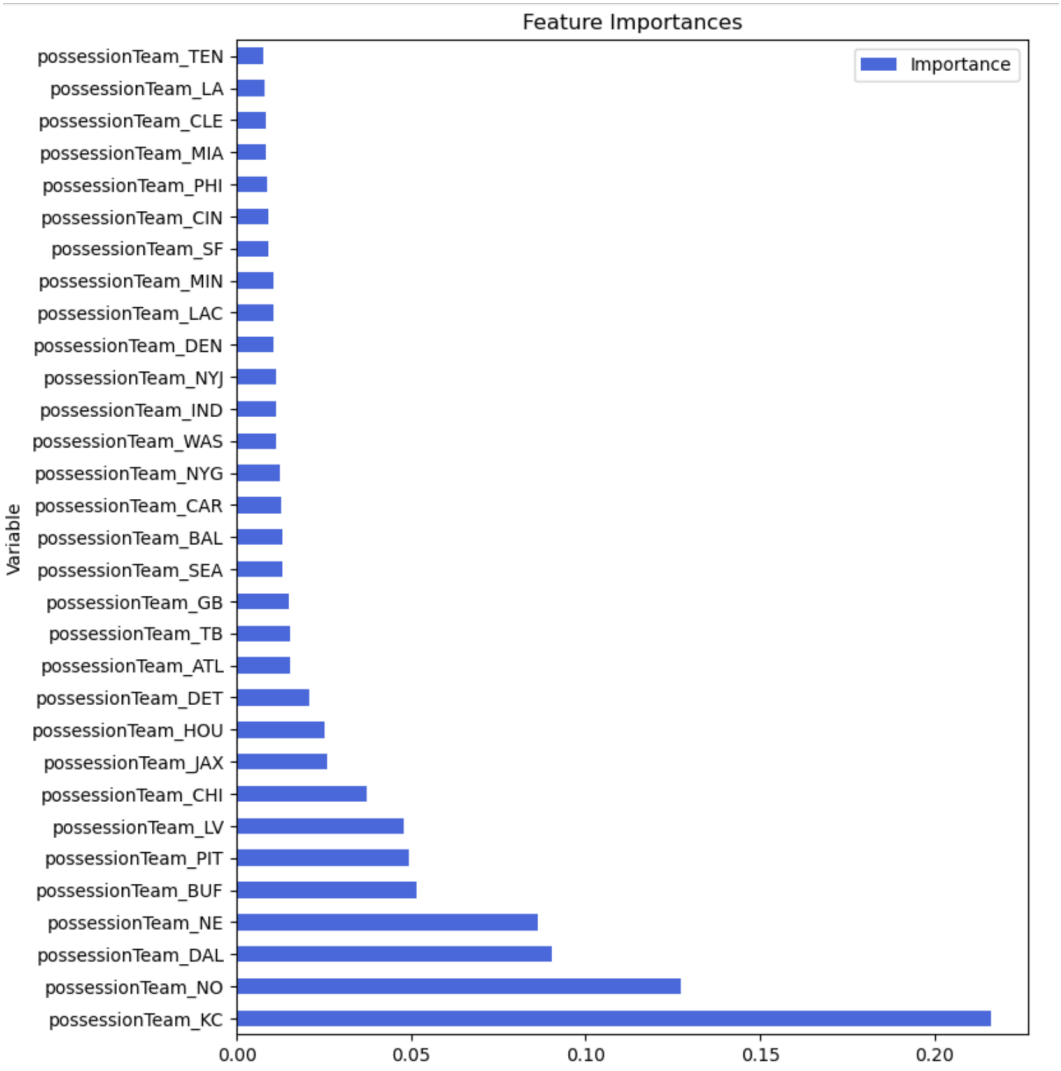
A variety of classification algorithms were evaluated for this research. Logistic regression was evaluated using both Lasso and Ridge regularization. A bagging ensemble and random forest were also evaluated using sklearn's GridSearchCV to find the best hyperparameters. Each algorithm was evaluated for accuracy, precision, AUC and top 10% lift.

V. Results

a. Exploratory Data Analysis

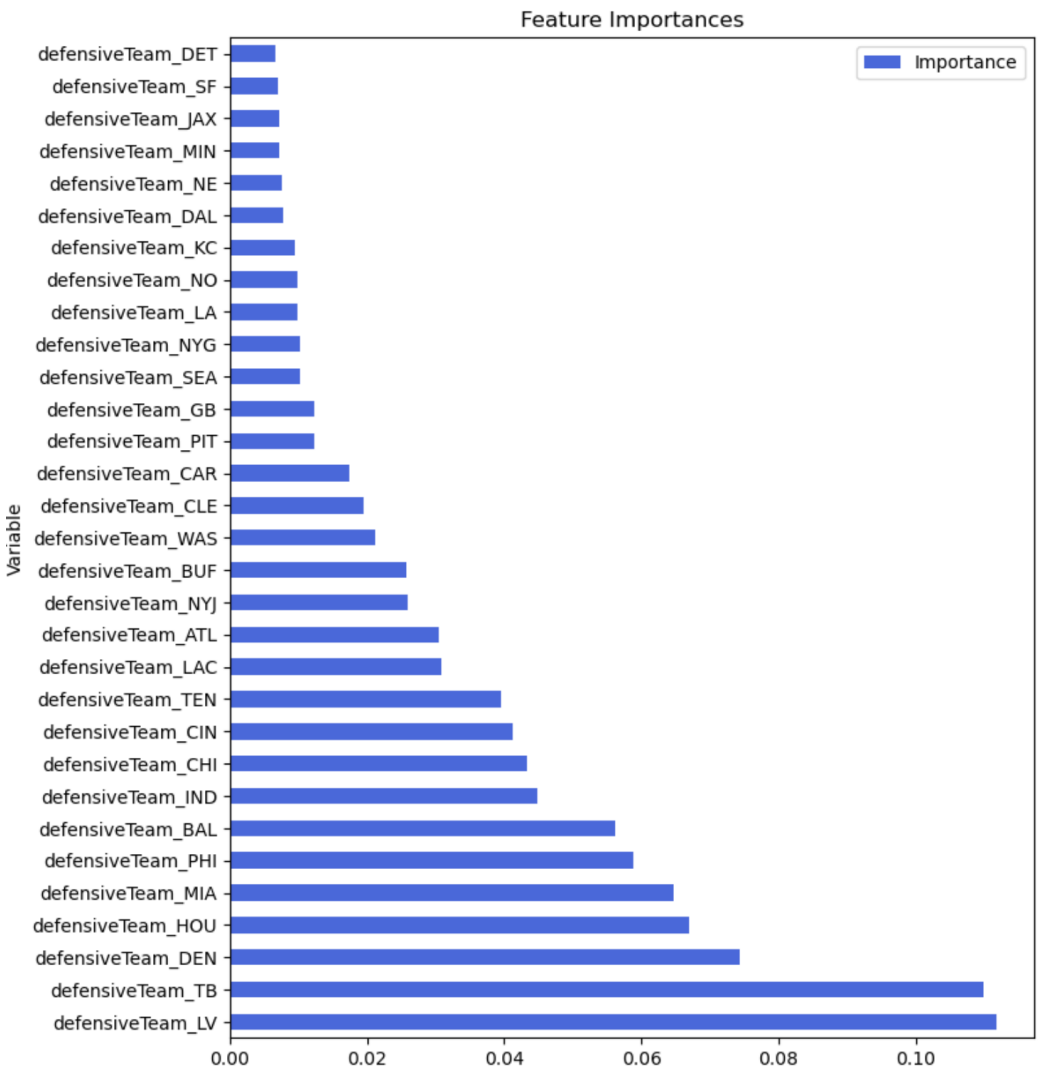
Evaluating feature importances with respect to the offensive and defensive team exhibited unique trends regarding the blitz. For the offensive teams, the Kansas City Chiefs ranked significantly higher than all remaining NFL teams as pictured in Figure 2. This indicates that of all NFL teams, the Kansas City Chiefs are most likely to have a blitz imposed on them. The next highest was the New Orleans Saints, followed by the Dallas Cowboys and New England Patriots, who had similar feature importances.

Figure 2
Evaluation of feature importances for offensive team



The same procedure was followed for defensive teams. Conversely to the analysis on offensive teams, feature importances for defense indicates that these teams are those most associated with a blitz. In Figure 3, the Tampa Bay Buccaneers and Las Vegas Raiders are nearly equal in feature importance and rank higher than the next following team. This indicates that these two teams' defenses enact the most blitzes compared to all other NFL teams.

Figure 3
Evaluation of feature importances for defensive team



b. Logistic Regression: Lasso and Ridge

A previous attempt at logistic regression had resulted in a nearly perfect model, an indicator of overfitting. After pruning and scaling the data, both Lasso and Ridge regression were implemented to address this issue. Lasso showed 75.43% accuracy with an AUC score of 0.76. Ridge showed exactly the same values. Confusion matrices for both regression models differed by single observations.

c. Bagging Ensemble

A bagging ensemble was selected in an effort to also reduce overfitting. The results of this model showed 75.29% accuracy and the lowest recall of all models, 19.00%. AUC was also the lowest of all models, 0.69.

d. Random Forest

Random forest methods can also be used in datasets that risk overfitting. A random forest provides simplicity and accuracy in classification problems. This model proved to be the best of all models, with 76.66% accuracy and 0.77 AUC. Additionally, lift was 2.50, the highest of all models.

e. Final Results

Comparing accuracy, recall, AUC, and lift of all models, Random Forest was ultimately the best model for this dataset. A full table of the details for each model can be found in Table 2.

Table 2
Comparing all models

Model	Accuracy	Recall	AUC	Lift Top 10%
Logistic Regression, Lasso	75.43%	77.74%	0.76	2.42
Logistic Regression, Ridge	75.43%	77.74%	0.76	2.42
Bagging Ensemble	75.29%	19.00%	0.69	2.28
Random Forest	76.66%	26.00%	0.77	2.50

When comparing all metrics, accuracy and AUC were deemed the most important. Between accuracy and recall, accuracy was more important due to its ability to evaluate the model as a whole, rather than the number of times the model could detect a particular category. AUC is similar to accuracy in that it evaluates the ability of the model to distinguish positive and negative outcomes. The random forest model was ultimately selected.

The random forest model suggests that zone coverage is the most significant feature when predicting a blitz. However, this is followed by the number of defenders in the box (close to the line of scrimmage) and Cover-0. These variables are nearly equal in feature importance, as seen in Table 3, and this strategically makes sense. Cover-0 is a defensive formation where no players exist in the defensive backfield. Instead, these players line up close to the line of scrimmage (or in the box) - nearer to the offensive back zone, hypothetically to induce a blitz.

Table 3
Feature importances for Random Forest model

Feature	Importance
pff_passCoverageType_Zone	0.162068
defendersInBox	0.068083
pff_passCoverage_Cover-0	0.063794
yardsToGo	0.054857
secondsRemaining	0.048669
yardlineNumber	0.046129
absoluteYardlineNumber	0.045141
down	0.040297

The remaining importance variables are time and situation dependent. It appears as though position on the field (`yardsToGo`, `yardlineNumber`, and `absoluteYardlineNumber`), time remaining in the quarter (`secondsRemaining`) and down are all associated with a blitz.

VII. Discussion

a. Hypothesis Evaluation

The original hypothesis sought to answer the question: Is it possible to determine or anticipate a blitz based on pre-snap characteristics? Results suggest that yes, with 76.66% accuracy via a random forest model, a blitz can be detected using pre-snap characteristics. Further, zone coverage, number of defenders in the box, Cover-0, and situational characteristics such as time remaining, field location, down and distance are most associated with a blitz.

b. Practical Implications

Results of this study can be used as supplementary information during pre-game preparation. This model would best be applied to all plays of a particular team, perhaps the opponent for that week. Teams differ greatly in coaching (and thus, play-calling) style, playing style, and talent. This research would be more specific and potentially more accurate if applied to a particular team.

Conversely, this research could be used as a self-evaluation - are there patterns to when we, as a team, choose to blitz? This point of view follows the same logic as applying this research to an opposing team - application to a single team would result in clearer, more definitive results as opposed to those that

could be identified by any armchair quarterback such as down and distance, location on field, and defenders on the line of scrimmage.

VIII. Limitations and Future Research

Feature importances of this research suggest that yes, specific pre-snap characteristics can indicate a blitz. However, to what extent are these characteristics novel and interesting? Characteristics relating to down and distance, time, and position are all obvious indicators of a blitz. Further research should more effectively utilize the GPS data. Current research utilizes the GPS data only to identify a blitz. However, players could be evaluated on an individual basis for "tells" or slight movements that may indicate a blitz.

Beyond this dataset, one interesting area of research could be 3-dimensional movement analysis. Technology exists for this and is frequently used in baseball. In a football context, this could be used to evaluate the tells or movements discussed earlier. This type of analysis is currently completed using film review. Availability of this type of data would remove the human element, but could also improve efficiency of coaching staff.

IX. Appendix

a. Sources

Goyal, G. (2019). *Leveraging Machine Learning to Predict Playcalling Tendencies in the NFL [Master's thesis, Massachusetts Institute of Technology]*.

Nwachukwu, E.O., & Uzoma, A.O. (2015). A Hybrid Prediction System for American NFL Results. *International Journal of Computer Applications Technology and Research, 4*, 42-47.

Rookie Road. (N/A). Football Blitz [Image]. https://www.rookieroad.com/football/plays/blitz/.

b. Variable List

This list refers to variables prior to dummy encoding. If a variable is dummy encoded, that is denoted in the 'Type' column.

Table 4
Variable list and descriptions

Variable Name	Description	Type	Example
`Quarter`	Quarter of the game	Integer	1
`Down`	Down of the series	Integer	3
`yardsToGo`	Yards to first down	Integer	2
`posessionTeam`	Offensive team, possesses ball	String → Dummy Variable	TB
`defensiveTeam`	Defensive team	String → Dummy Variable	DAL
`yardlineNumber`	Line of scrimmage location on field	Integer	33
`preSnapHomeScore`	Score of home team prior to ball snap of this play	Integer	0
`preSnapVisitorScore`	Score of away team prior to ball snap of this play	Integer	7
`absoluteYardlineNumber`	Distance from endzone for possession team	Float	43.0
`offenseFormation`	Formation of offensive team	String → Dummy Variable	SINGLEBACK
`personnelO`	Description of offensive personnel	String → Dummy variable	1 RB, 1 TE, 3 WR
`defendersInBox`	Number of defenders in close proximity to line of scrimmage	Float	6.0
`personnelD`	Description of defensive personnel	String → Dummy variable	4 DL, 2 LB, 5 DB
`dropBackType`	Dropback description of quarterback	String → Dummy variable	TRADITIONAL

pff_playAction	Binary variable denoting play action pass	Boolean	0
pff_passCoverageType	Description of defensive pass coverage	String → Dummy variable	Cover-1
numBackfield	Number of defensive players that move into the offensive backfield	Integer	4
isBlitz	Binary variable denoting if a blitz occurred	Boolean	0
weekNum	Week during which game occurred	Integer	2
secondsRemaining	Second remaining in the quarter	Integer	813