Motivation

While many talented athletes rise to stardom in the world of professional sports, there are few who leave such an enduring legacy as Tom Brady. Within the last several seasons, Brady's achievements have led to more discussion regarding the depth and longevity of his athletic abilities. The motivation for this analysis is to study Tom Brady's statistics as a quarterback in order to quantify several aspects of this discussion. The main objective is to determine Brady's biggest strengths as a player; a regression model can be utilized in this case to determine how different factors contribute to his overall passer rating. Another topic of interest is how his overall statistics compare over time to determine if he has maintained, increased, or decreased in ability throughout his 20-year long career. Analysis of coefficients in the regression model can be useful for this purpose.

Data Description

The data was collected from https://www.pro-football-reference.com/. The data includes Brady's regular season stats from every year in which he played 12 or more regular season games. After excluding the years in which he played less than 12 games, there are 19 observations in this data set. The response variable is Brady's touchdown-to-interception ratio, which is one way to quantify a quarterback's measure of success. A predictor variable used in this model is age.

Data Exploration

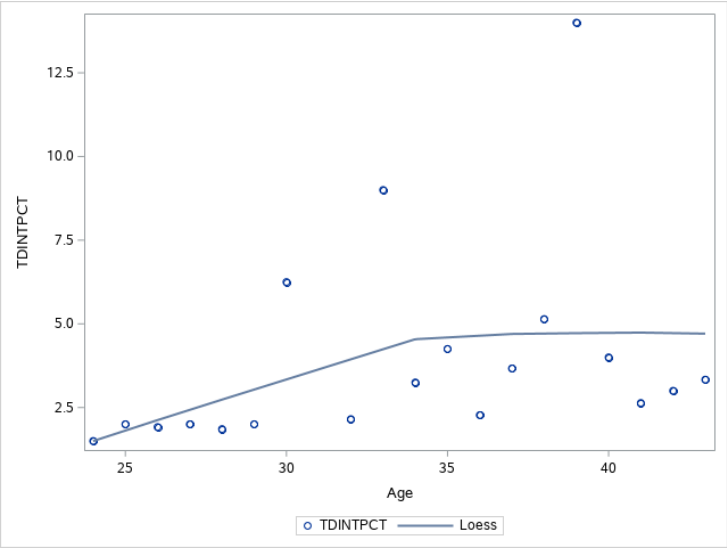
```
FILENAME CSV "/home/u52292162/Assignments3064/BradyStats3.csv" TERMSTR=CRLF;
proc import DATAFILE=CSV
                    OUT=BradyStats
                    DBMS=CSV
```

REPLACE;

RUN;

proc print data=BradyStats;

run;

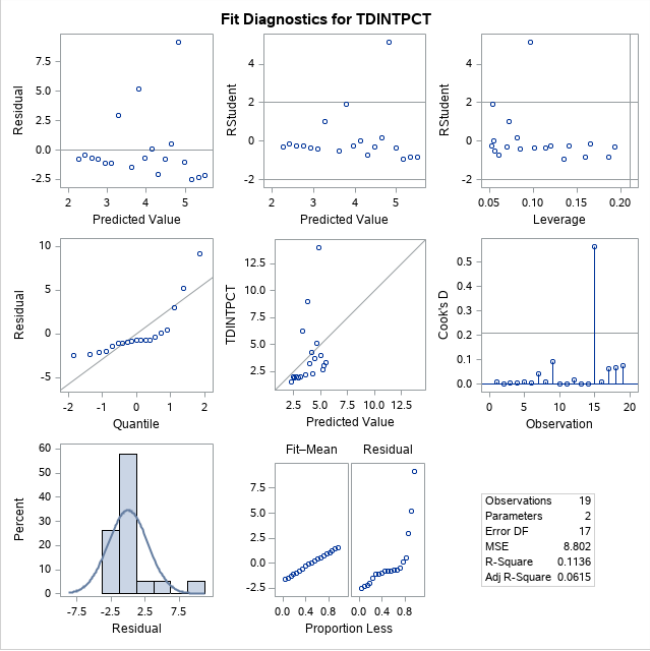


Statistics like total yards or pass attempts were removed due to lack of predictivity, as well as years in which Brady played fewer than 12 games. The table was saved as an excel worksheet, a column was added to divide the number of touchdowns by the number of interceptions in each column for the TD:INT response in each entry, exported as a csv file, and then imported into SAS. A blank line between the variable names and resulting data set was removed for formatting. The data does not seem to follow a linear pattern, as it oscillates higher and lower throughout time, although the linear regression line appears to capture the general median values throughout time.

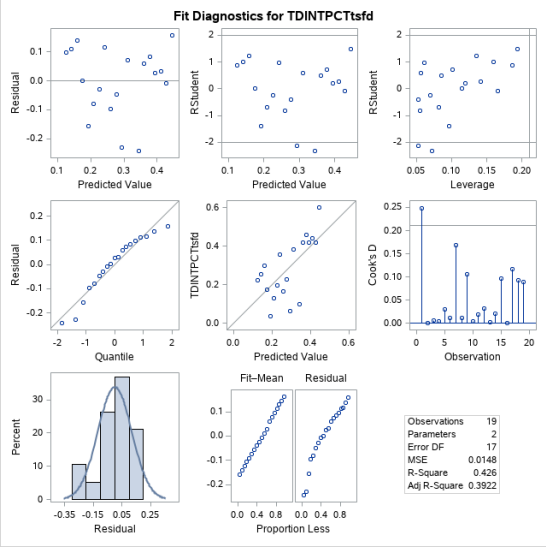
Model Fitting and Analysis

The equation of the simple linear regression model is $y = .17069x - 1.83372$. With a p-value of 0.1582, indicating this model is not significant, and an r-square of 0.1136, indicating

that only about 11% of the variance is explained by this model, it can be concluded this model can definitively be improved.



Analysis of residual plots confirm that assumptions for a linear model are not met, as residuals do not have constant variance, distribution of error is not centered at zero (heavily negative) and the quantile plot shows significant deviance from the pattern.



After utilizing a Box-Cox suggested transformation of lambda = -1.25, the model now indicates that there is now a strong predictive relationship between the predictor and response variables. Residual plots confirm that assumptions are now met for this model, as residual plots indicate random, constant variance with the distribution of error centered around zero. There is much less deviance from the pattern on the quantile plot.

The confidence interval at an alpha of 0.05 for the slope of this model is [-0.02682, -0.00683]. The predicted slope is -0.01683, indicating that Brady's touchdown-to-interception ratio seems to be decreasing only *very* slightly over time. An F-test on slope (Age) indicates a p-value of 0.0025, thus confirming that this model is statistically significant. At an age of 50, the model suggests Brady's TD:INT ratio would be 0.00765.

Conclusions

This model provides some statistical evidence suggesting that Tom Brady will, in fact, eventually get old. However, although the overall direction of his TD:INT ratio is negative, it seems to be decreasing extremely slowly over time. The model indicates a strong, but minimally negative, relationship between Brady's age and TD:INT ratio. Though many have argued over the past few years that statistics confirm Brady is no longer the quarterback he used to be, these statistics seem to suggest that he isn't headed towards retirement any time within the immediate future.