Elijah McLaughlin

STA 3064

Project A

**Motivation:** My study is of interest because it can help determine the usefulness of SRS (Simple Rating System) in determining wins for NFL franchise wins. In this case, I am using the Houston Texans. Potential benefits could include fans of these franchises being either excited or apathetic for the upcoming seasons, sports bettors having better insights into placing bets for teams making the playoffs, and for NFL General Managers to set fair expectations for their teams.

**Data Description:** The data used is quantitative and numerical in nature. I collected the data from “Pro Football Reference”, which is a statistics warehouse full of NFL related statistics. The code is specifically pulled from: <https://www.pro-football-reference.com/teams/htx/> and saved as a .csv file. I removed the 2023 season as that is the current season and not needed as the it is still being played. The response variable is ‘W’, or wins, and the main predictor variable is ‘SRS’. Wins are the total amount of wins a franchise has given at the end of the season, and SRS is the Simple Rating System that Pro Football Reference created to show the rating of a given NFL franchise.

**Data Exploration:**

A screenshot of a computer

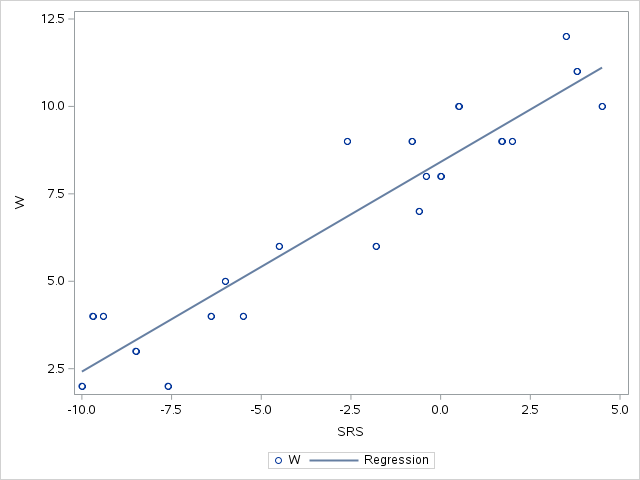
Description automatically generatedA graph with blue dots

Description automatically generatedA graph with blue dots

Description automatically generated

Other than importing the data correctly, no other data manipulation was needed. The scatterplot appears to show a linear relationship, and when adding in the LOESS line it appears to have a strong linear relationship.

**Model Fitting and Analysis:**

1.   
   Linear equation: W = 8.41207 + 0.59954\*SRS
2. R2=0.8723, and this high of an R2 value suggests that SRS has a strong correlation for ‘W’.
3. ‘Pr > F’ is <0.0001, meaning that there SRS is statistically significant. Model assumptions are met as it fits the residual plots.
4. I performed a square root transformation for 'W'. Results are similar, with R2 being 0.8605.
5. 95% confidence interval is [0.48940, 0.70968], and the point slope being approx. 0.59954. This shows a positive correlation between SRS and W, i.e. when SRS is higher W will be higher and vice versa.
6. 'Pr > F' is 0.0007, meaning that the results are statistically significant, and the model is effective.
7. 95% interval: [-1.96430, -0.14420]. This is a weak, negative correlation between SoS (Strength of Schedule) and W. Prediction interval: [5.41169, 7.90127]. This is predicting the Houston Texans to get between 5.4 and 7.9 wins this year.

**Conclusions:** Based off the analysis and findings for this report, SRS is a good indicator of wins for NFL franchises. There is strong and positive correlation between SRS (Simple Rating System) and Wins, meaning that the higher a team’s SRS there is a strong chance that the team will have more wins. However, when performing an analysis with SoS (Strength of Schedule), it revealed that SoS has a weak and negative correlation between SoS and wins. Just because a team’s SoS is high, does not mean that the team will have more wins. For NFL General Managers, this analysis can be used to help with determining draft position, trade acquisitions, and expected success for the season.