Stat 6021: Guided Question Set 10

In this guided question set, we will use the "nfl.txt" data set that we used in the last module. As a reminder, the data are on NFL team performance from the 1976 season. The variables are:

- y: Games won (14-game season)
- x_1: Rushing yards (season)
- x_2: Passing yards (season)
- x_3: Punting average (yards/punt)
- x_4: Field goal percentage (FGs made/FGs attempted)
- x_5: Turnover differential (turnovers acquired minus turnovers lost)
- x_6: Penalty yards (season)
- x_7: Percent rushing (rushing plays/total plays)
- x_8: Opponents' rushing yards (season)
- x_9: Opponents' passing yards (season)

We will continue to regress the number of games won against three predictors: passing yards, x_2, percent rushing, x_7, and opponents' rushing yards in the season, x_8.

1. Create diagnostic plots for this regression. What are the plots telling us?
2. Generate partial regression plots for each of the predictors. Interpret what these plots are telling us.
3. Using externally studentized residuals, do we have any outliers? What teams are these?
4. Do we have any high leverage data points for this multiple linear regression? What teams are these?
5. Use $DFFITS_i$, $DFBETAS_{j,i}$, and Cook's distance to check for influential observations. What teams are influential?