

Wrangle Report

The gather phase of the project was done in the following steps: I loaded the csv file 'pass-2016.csv' into a dataframe called df_original. I then made a copy called df for wrangling purposes.

The cleaning portion of the project was done in the following steps: I found that there was extra data in the 'Player' column that appeared to be some sort of id. I also found asterisks in some of the rows as well, with no discernable meaning. In both cases, I removed the data from the column. I found that there were multiple data points in the column 'QBrec'. I put each of these data points into their own column, the three being called 'QBwin', 'QBlose', and 'QBtie'. I found that there was data in the 'Pos' column that contained the same data in both lowercase and uppercase form. I converted the entire column to uppercase for consistency.

Finally, I put the cleaned dataframe into a csv file called 'pass-2016-master'.