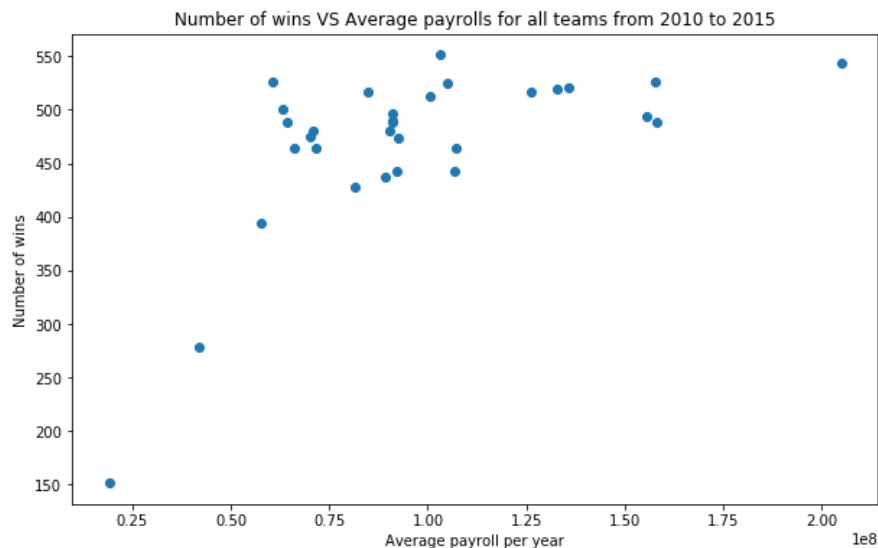# Summary for lab 2

For this lab, I analyzed the dataset of MLB baseball datasets, compute the relevant statistics, and created some visualizations. And I explored some of the relationships in the datasets, which are shown in this summary.

1. The relationship for a team's winning record and its payroll.

Based on this scatter graph, I think that the total wins are not quite related to the average payroll. Some teams with relatively small average payroll (3,000,000 - 5,000,000) also have a high number of wins which is larger than 500. But some teams with this payroll have a smaller number of payrolls (400 - 450). So there is not an obvious relationship between them. However, what is interesting here is that the teams with a very low payroll (2,000,000 is considered a very low payroll among all teams) have a very small number of wins, and the teams with a higher average payroll (a payroll that is larger than 6,000,000 is considered high here) tend to have a higher number of wins (larger than 500). So we can conclude that a team with a relatively higher average payroll usually will not have a very small number of wins, but there is not a necessary relationship between the average payroll and the number of wins.

2. The relationship of a team's batting average and its win-loss record.

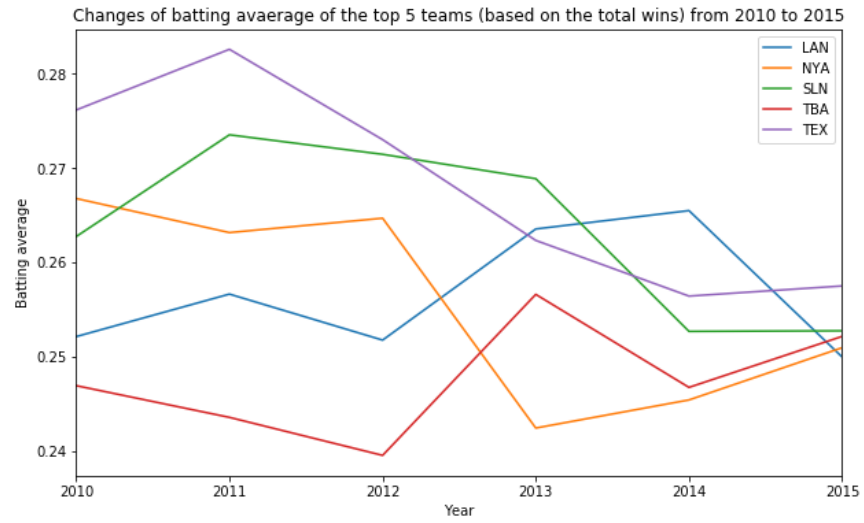Average batting VS Win loss for all teams from 2010 to 2015
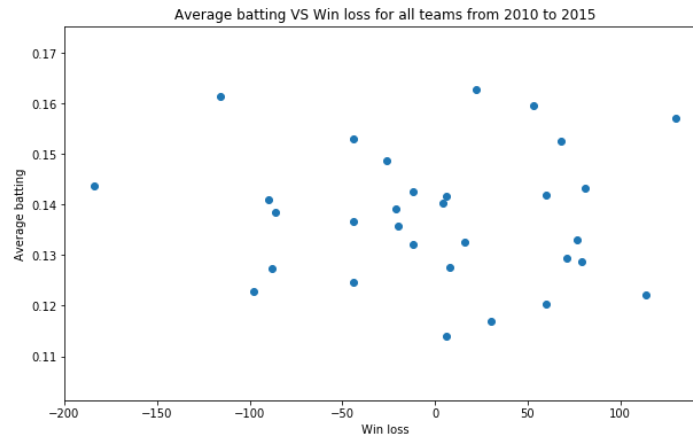
Based on this graph, I think there is not a relationship between the average batting and the win loss. We notice that for a certain value of win-loss, the average batting can be either low or high.

3. The relationship of a team's win-loss record and its pitching performance.

Wins-loss record VS pitching performance for all teams from 2010 to 2015
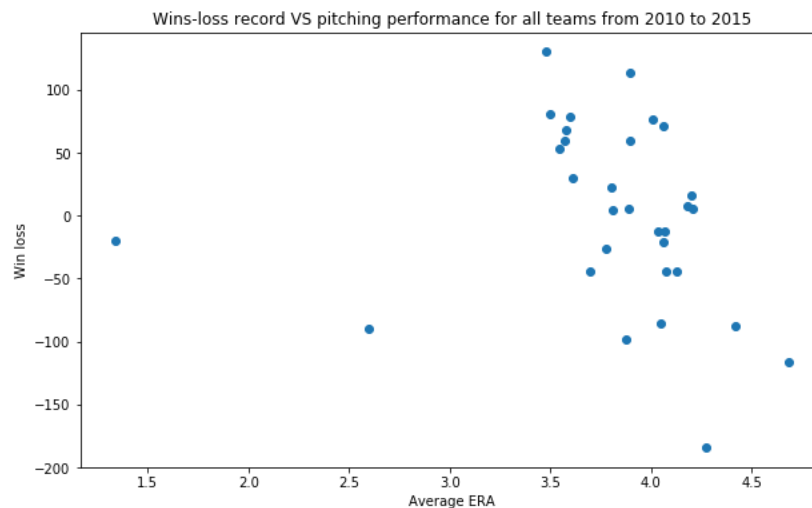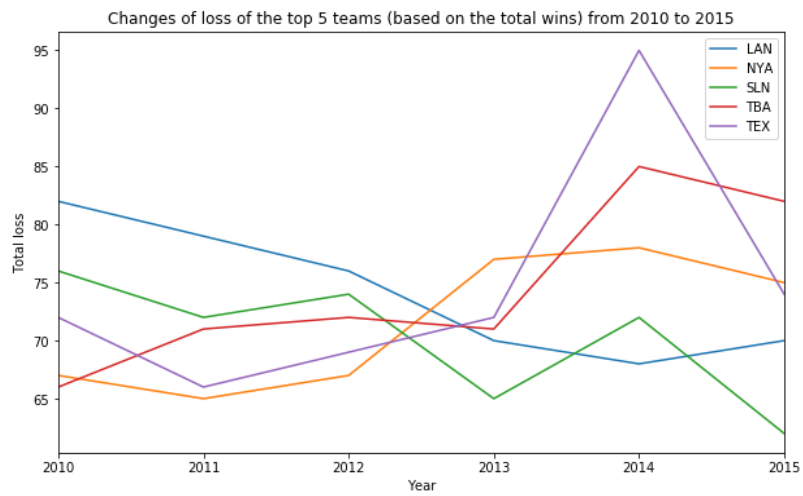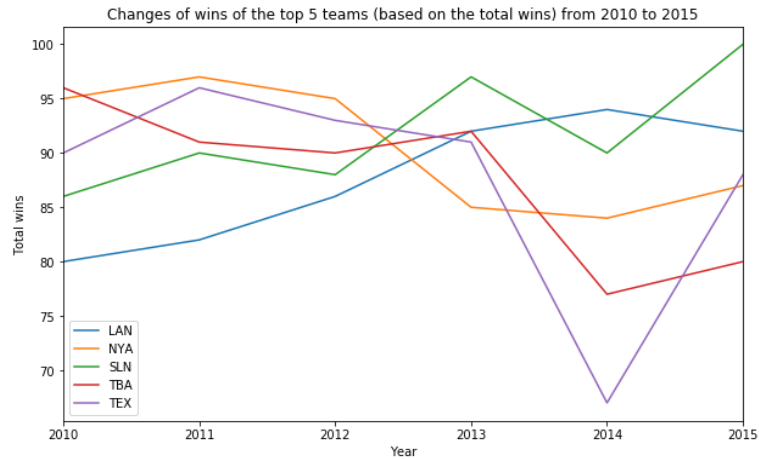

I don't think there is an obvious relationship between the average ERA and the number of wins. We can notice that the teams with similar average ERA (when ERA is in the range of 3.5 to 4) have different number of win-loss. But what is interesting is that when the ERA is greater than 4.3, the win-loss are all very poor.
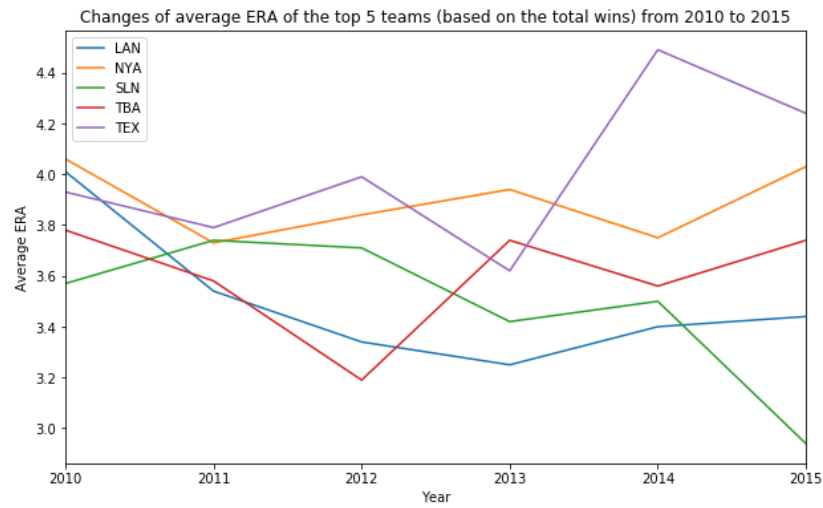
4. For the top 5 teams (based on winning record) that have the most total wins between 2010 and 2015:
    1) How are their Batting Averages changed from 2010-2015

Changes of batting avaerage of the top 5 teams (based on the total wins) from 2010 to 2015

2) How are their Wins/Losses changed from 2010-2015

Changes of wins of the top 5 teams (based on the total wins) from 2010 to 2015

Changes of loss of the top 5 teams (based on the total wins) from 2010 to 2015

3) How are their average ERAs changed from 2010-2015

Changes of average ERA of the top 5 teams (based on the total wins) from 2010 to 2015

4) How are their annual payrolls changed from 2010-2015

Changes of average annual payroll of the top 5 teams (based on the total wins) from 2010 to 2015