Anthony Chirinos, Abrar Mostofa, Ethan Papa

Intro/Question:

Does fan attendance rates (percentage of home stadium capacity/game) reflect manager turnover from 2008-2018? We expect that seasons in which a manager is fired will have lower average attendance. We also were curious about how average attendance correlated with win rates per team, which we addressed on the second slide. We expect that as win rates increase, average attendance will also increase. Manager firings also tend to indicate team performances and we expect to see teams with longer tenured coaches often having more success which we'll see based off of win totals.

Data Sources:

The ESPN dataset provided us with average game attendance for each team per year. We had to aggregate the average percent attendance by year and by team, and place them onto one spreadsheet so we could make figures from it. We also had to convert some of the raw average attendances to percent by factoring in stadium capacity.

The baseball reference data set told us about what managers were fired and what years. We used our previous data set and highlighted years in which manager firing was indicated to show differences.

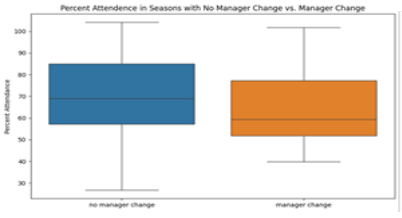
This team reference source provided us with the win totals per team from 2018 to 2008. This allowed us to compare those win percentages with average attendance.

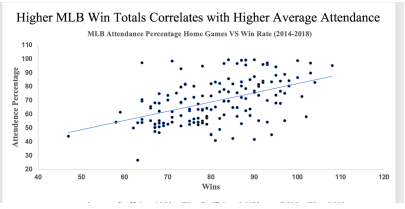
Methodology: (What graphs/models we're building)

We built a boxplot comparing the attendance percentage in seasons that had a manager change–either in season or after the season–with the attendance percentage vs. when the manager was retained by the team. Some limitations are factors that could be contributing to fan attendance, like win percentage or geographical location. We built a simple linear regression for this graph, measuring the coefficients with no manager change, coefficient with a manager change, and the p value for the manager change in order to see the contrast between the situations.

For the second graph, we built a scatter plot comparing win totals and attendance percentage per team by year. We used a simple linear regression model to find the intercept coefficient, the wins coefficient (slope) and the p values.

Results and Analysis:

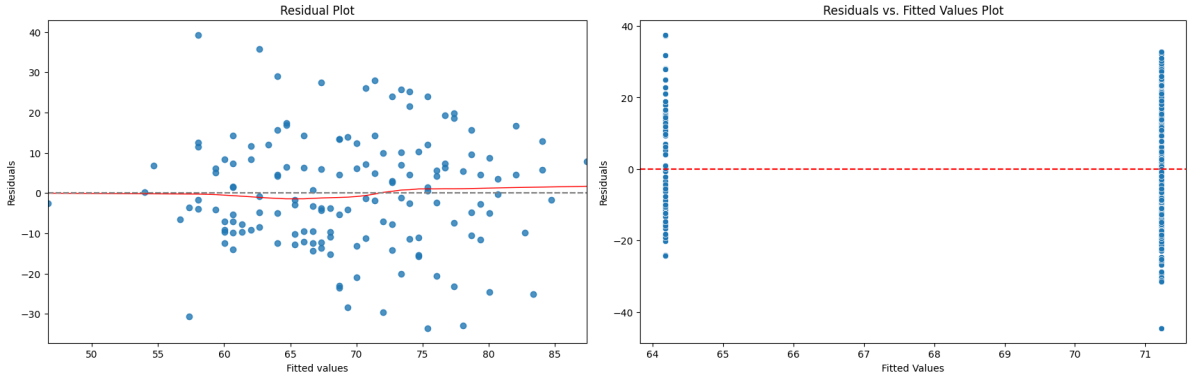




We found that attendance percentage was significantly higher in seasons with no manager change than in seasons with a manager change. For the first graph, we found that the constant coefficient was 71.2. This means that the average percent attendance in MLB seasons with no manager change was 71.2 percent. We also found that the manager change coefficient was -7.05, meaning that in seasons where the manager was fired, the average attendance was 7.05 percent lower than in seasons with no manager firing. This manager change's p value was 0.002, meaning that these results are statistically significant. Ultimately, there is a significant decrease in fan attendance when a manager was fired compared to when a manager was not fired.

We also found that higher win totals are correlated with higher attendance. For interpret coefficient, we got a value of 15.3, meaning that with wins of zero, the average attendance would be 15.3 percent. For wins coefficient, we had a value of 0.66, meaning that for every one increase in wins, average attendance increased by 0.66 percent. The p value for wins was 0.000, meaning the odds of getting values that extreme by random chance are 0.000. This means that those values are statistically significant, and average attendance increases as wins increase.

We also created residual plots for both graphs, included below. Both of these residual plots have means very close to zero, show relative homoskedasticity, and seem to have a random scatter. Because of this, we can say that our models are relatively valid.



Conclusion:

Ultimately, we found that percent attendance was significantly higher in seasons with no manager change than in seasons with a manager change. The data reflected that this variable caused a significant decrease in fan attendance across all teams, through all years. The second graph also shows that higher win totals meant higher win totals.

References:

Anthony Chirinos, Abrar Mostofa, Ethan Papa

https://pitt-my.sharepoint.com/:x:/r/personal/etp39_pitt_ed3u/Documents/Copy%20of%20finalr.xlsx?d=wb3c5a019536646a5b3bae21de7e8bbf4&csf=1&web=1&e=WcA1Pr MLB Attendance and Manager Turnover

BaseballReference.com , 2018 MLB Attendance - Major League Baseball - ESPN

https://www.teamrankings.com/mlb/stat/win-pct-all-games

Manager Turnover: Baseball Reference.com, ESPN team transaction logs that show manager changes

Fan Attendance: BaseballReference.com, https://www.espn.com/mlb/attendance/_/year/2018 (If page not found: search MLB 2018 attendance through ESPN on search browser) . Page should look like this