CMSE MidTerm Project Report: Baseball Hitting metrics EDA

Sangeetha Suresh

Goal:

Statistical analysis of games can determine a strong correlation between wins and the strategy used. Being largely a numbers game, baseball can easily produce such insights, which can help payers strategize their play style.
The project aimed to determine relationships between hittings metrics and wins for games over the last ten years. Moreover, the relationships between wins and winning tendency, play style, health, and covid were also analyzed and explored.

What do you learn from the app?

1. There exists a clear relationship between wins and, Runs created and slugging percentage.
2. The divisional winners tend to have an equal number of defensive and offensive players.
3. The world series winners are primarily defensive, but surprisingly, the offensive payers have better hitting metrics. This shows that It's better to focus on stopping runs than scoring.
4. Covid greatly impacted the players, and all teams seemed to have underperformed. This could mean the health or audience effect impacted the players.

Visualizations:

Bar graphs, Regression plots, Scatter plots, and stacked graphs were used to represent the relationships between hitting metrics and wins.
For the visualizations, libraries like Matplotlib and Altair were used.
The graphs used gray and orange to differentiate and highlight the subject.

Preprocessing.

The data did not have nan values, but it did not have the hitting metrics explicitly, so they had to be determined explicitly.