Team 4 Project Proposal MLB Stats

We would like to analyze different team statistics including:

1. Team Hitting Stats (Batting Avg, RBI, etc.)
2. Team Pitching Stats (Strikeouts, Saves, etc.)
3. Team Fielding Stats (Errors, Double Plays, etc.)

Supervised learning:

Logistic regression, or Neural Networks

* Compare individual stats against a determined threshold of success (bucket the number of wins <70, 70-90, 90+).

Unsupervised learning:

PCA, t-SNE, K-means Cluster Analysis

* Use a drop-down menu to choose a team. Then it tells you which teams are most like that team based on hitting/pitching/fielding stats. Also, overall which teams the chosen team is most like and had the most similar total stats.
* Using the grouping to analyze which stats have more influence on whether a team will be successful or not.