Connor Mignone
Daron Greenblatt
Dong Chen
Ty Painter

Case Study #2

We were quick to discover that our interests aligned in the realm of sports. What better sport to kick off a deep dive data analysis in than baseball - the game that fueled an arm's race of advanced analysis in professional sports. Our project proposal involves pitch-level data regarding every pitch thrown in a Major League Baseball (MLB) game from 2015-2018. Major League Baseball (MLB) is a professional baseball organization, the oldest of the four major professional sports leagues in the United States and Canada. The dataset includes 40 columns with each row giving a description of the speed, location, and movement of a pitch. The aim of our analysis is to be insightful for both pitchers and hitters, along with their respective coaches. Our objective is to understand pitchers' pitch selection and what makes pitches most effective under certain conditions such as pitcher/batter matchup, base runners, and score. If possible, we would like to predict the likelihood of which pitches will be made in certain situations to prepare a batter. On the flipside, we envision our analysis would be utilized by pitchers to enhance their chances of successfully minimizing runs scored. The results of our analysis would be of benefit to any MLB organization when developing scouting reports to evaluate a pitcher on their team, or to scout an opposing pitcher to help provide insight to their own hitters.
https://www.kaggle.com/pschale/mlb-pitch-data-20152018?select=pitches.csv