

CMU Softball Pitcher Efficiency

Scott Steinberg and Gautam Goel

1. Client Goals

Our clients are coaches Monica Harrison and Lindsay Mapes of the CMU Softball team. During each game, they call each pitch based on the current game situation. The goal of this project is to use pitch-by-pitch data from the 2019 and 2020 seasons to provide insight into pitcher performance during specific game scenarios. Doing so will allow for quicker in-game decisions, improving overall effectiveness and efficiency of pitching.

2. Softball Background

Pitch Type	Inside	Outside	Low	High	Curve
Change	Change Inside	Change Outside			
Curve		Curve Outside			
Drop	Drop Inside	Drop Outside			Drop Curve
Fastball	Fastball Inside	Fastball Outside	Fastball Low	Fastball High	
Rise	Rise Inside	Rise Outside			
Screw		Screw Outside			

This chart details the main pitch (rows) types and variations (columns) used by the softball team.

3. Data Collection

Below is a sample of the raw data we were provided. On the left are Pitching Evaluation Charts. They are created by pitchers postgame and show a summary of their pitch distributions from that game. On the right are Pitching Charts. They are created in real-time during each game. They track every pitch thrown to each batter and the subsequent outcomes.

4. Data Entry

In order to analyze the data, we first had to create a data product. Below is a layout of how we converted the Pitching Charts into a data frame.

Each block represents one at bat, with balls on top and strikes on bottom

Pitch #	Batter	Pitcher	Bats	Inning	Runs	Outs	Balls	Strikes	1st	2nd	3rd	Pitch	Result
1	Name	Name	R	1	0	0	0	0	0	0	0	Screw	Ball
2	Name	Name	R	1	0	0							
3	Name	Name	R	1	0	0							
4	Name	Name	R	1	0	0							

Pitch #	Batter	Pitcher	Bats	Inning	Runs	Outs	Balls	Strikes	1st	2nd	3rd	Pitch	Result
1	Name	Name	R	1	0	0	0	0	0	0	0	Screw	Ball
2	Name	Name	R	1	0	0	1	0	0	0	0	Curve	Ball
3	Name	Name	R	1	0	0							
4	Name	Name	R	1	0	0							

Pitch #	Batter	Pitcher	Bats	Inning	Runs	Outs	Balls	Strikes	1st	2nd	3rd	Pitch	Result
1	Name	Name	R	1	0	0	0	0	0	0	0	Screw	Ball
2	Name	Name	R	1	0	0	1	0	0	0	0	Curve	Ball
3	Name	Name	R	1	0	0	2	0	0	0	0	Screw	Strike
4	Name	Name	R	1	0	0							

Pitch #	Batter	Pitcher	Bats	Inning	Runs	Outs	Balls	Strikes	1st	2nd	3rd	Pitch	Result
1	Name	Name	R	1	0	0	0	0	0	0	0	Screw	Ball
2	Name	Name	R	1	0	0	1	0	0	0	0	Curve	Ball
3	Name	Name	R	1	0	0	2	0	0	0	0	Screw	Strike
4	Name	Name	R	1	0	0	2	1	0	0	0	Change	Ground Out

5. Data Product

Once we had a complete data frame. We developed an R Shiny App to allow for customizable visualizations of pitching data.

We had charts for Pitch Outcomes (above) and Pitch Types. Both charts have a variety of filters that could be applied as shown above.

Above example shows Pitch Type Distribution for all games against Brandeis in 2019 when the count was 3-2 (3 balls, 2 strikes)

6. Conclusions

We have now created a sustainable system by which the softball team can efficiently digitize Pitching Charts and immediately see visualizations by uploading data into R Shiny App. Doing so will allow for recognition and improvement of pitchers' weaknesses and better in game decisions.