

University Honors Program

Honors Thesis Preliminary Proposal

Name: Caroline Thomas

CSUID: 830913567

E-mail: thomasc@rams.colostate.edu

Phone: 3036888879

Honors Track: Track 1 2015

Intended Graduation: Fall of 2020

Current Semester: Fall of 2019

399 Instructor: Anne Merline

Thesis Adviser: Bailey Fosdick

Adviser's Position: Assistant Professor

Adviser's E-mail: bailey.fosdick@colostate.edu

1. Interest: Explain why you are interested in this topic.

This project is a fusion of two of my passions-- that is, my love for sports and my love for statistics/data analytics. I am a Statistics major so my interest in data analytics kind of speaks for itself but I am always blown away by the power of data and the tools that have been created to process it. Data analytics opens so many doors and can answer so many questions which is awesome! Statistics also involves computer programming which is something that I have grown to enjoy over the past couple years. My love for sports stems from a childhood of playing almost every sport. Softball, though, was the one sport that really stuck with me, resulting in 12 years of playing competitive softball. Since retiring from the game after my senior high school season, I have been looking for an opportunity to find my way back to softball, even from the sidelines. This past summer, I had a sports statistics internship working with the CSU Women's Volleyball team in conjunction with the Statistics Department, which I absolutely loved. I already knew that I loved sports and statistics but my work over the summer really showed me that sports analytics is something that I am really passionate about and would like to go into as a career in the future. Near the end of the summer, we reached out to the CSU Softball coaches to ask if they would be interested in working on some similar project with us, going forward and the head coach, Jen Fisher, was very interested!

2. Describe your thesis: What are the creative activities, research questions, and/or design features that are foundational to your thesis?

The goal of my thesis is to develop a method to evaluate how well CSU softball players execute their given game plan. That is, not whether or not they pitched a strike or got a hit, but rather how well were they able to execute their coach's instructions i.e pitch the pitches that were called, swing at pitches in the zone where they perform the best, and lay off of pitches that they are not as good at connecting with. The coaches are interested in having players who are coachable and able to do their job so they would like a way to "grade" how well each player is doing this. In doing this, these "grades" would also serve as motivation for the players to show them what they can improve upon and how well they did in the eyes of the coaches, even if their game statistics (batting average, ERA, etc.) don't seem as good. For the scope of this project, I will focus on the offensive side of this, that is

how well our batters are executing their game plan when at bat. According to CSU Head Coach, Jen Fisher, each situation (count) that players encounter when at bat can be “won” or “loss” and I will base player “grades” based on their ability to “win” a pitch—that does not necessarily correlate with whether the pitch was a ball or strike, but depending on the situation, did they do what Coach Fisher wanted them to do.

3. Describe your methods and approaches: How will you discover answers to research questions, create artistry, and/or develop design projects?

My first step will be completing a literature review of current methods for evaluating player performance in softball and baseball. Softball analytics is a largely unexplored field so I am not expecting great results from this. Then, I will use play-by-play data from last year’s (or perhaps the last several years’) home games (from Right View Pro software) as well as personalized player information from the head CSU softball coach, Jen Fisher, to develop my method using previous seasons as a test case. If I am able to achieve my goal using that data, then I will design a streamlined method to use during future seasons that will be able to have a quick turnaround (game-by-game results) for more immediate player feedback after a game. My methods will largely involve computing and statistical methods that will be performed using the program R. I will refine my methods to be put into an R package that will be publishable to be used by coaches in the future.

4. Outcome: Describe the creative activity and the expected research result, final product, and/or artistic expression that is the core of your thesis.

My primary final product will be an R package that I will publish online along with a written vignette that details the functions of my package. This will be similar to other vignettes that have been written for R packages with the purpose of explaining the purpose of the package and how to use the functions within it. This package will contain functions that will serve to clean the data, format the data correctly, and of course, perform the player evaluations. Additionally, it may be useful to the user to have some visualizations to go along with the data so I will write functions to do basic data visualizations as well. This R package should be ready and easy-to-use for coaches who would like to use my evaluation system in the future. I will also write a reflection paper that contains the results of my literature review, a description of my goals and methodologies, and an explanation of my final evaluation system.

HONR399 Instructor's Comments: Thank you for this preliminary proposal. You can now forward this on to your potential thesis adviser. Please contact me if you have any questions.

Student Signature: Caroline Thomas

Date: 12/2/2019 11:16:39 AM

Thesis Adviser Signature: Bailey Fosdick

Date: 12/2/2019 1:49:17 PM