

CSCI 5342 - Sports Data Mining (Fall 2024)

### Course Description

This is an advanced course in analyzing sports data for decision making. Identifying the metrics, types of analyses and making sense of sports-related data from a managerial business perspective. Use of industry tools to gather, learn, make predictions and visualize large sports data sets.

#### Class Time

This class is offered asynchronously online. While you control when you watch videos and work on assignments, be aware of course pacing and specific deadlines.

### **Instructor Information**

Dr. Robert P. Schumaker Professor, Computer Science Dept. rschumaker@uttyler.edu

#### Office Hours

DM through Slack (preferred), Zoom, email

If your inquiry is grade-related, please make a Zoom appointment

### **Textbook Information**

Analyzing Baseball Data with R (Marchi, Albert and Baumer)

ISBN: 978-0-81535-351-5

Sport Analytics (Fried and Mumcu)

ISBN: 978-1138667136

## Course Objective

This course is designed with the following goals:

- Identify a broad range of methods used in sports data acquisition, representation, analysis and reporting
- Demonstrate an understanding of statistics and their application to sport
- Develop an ability to recognize, formulate and analyze decision-making in sport
- Improve overall problem solving/analysis skills and critical thinking
- Conduct sports data acquisition, representation and prediction activities
- Assess current sports analytics trends and how they can apply to new areas

### Computer Account Access

Students will need a Patriot account and password for computer access. This information can be found at https://www.uttyler.edu/ccs

### Course Documents and Slides

This class will use Canvas for course documents, slides, quizzes and other class-related materials. Students are encouraged to check the website frequently during the course of the semester to keep up to date about course activity.

CSCI 5342 - Sports Data Mining (Fall 2024)

### Course Grading - Undergraduates

Course evaluation will be based on the following:

MLB Decision Analysis Report	50
NFL Decision Analysis Report	50
Total Points	100

## Course Grading - Graduates

Course evaluation will be based on the following:

MLB Decision Analysis Report	35
NFL Decision Analysis Report	35
New Topic Module (Slides and Code)	30
Total Points	100

### **Grading Scale**

- A 85.0 points or more
- B 70.0 to 84.999 points
- C 55.0 to 69.999 points
- D 40.0 to 54.999 points
- F 39.999 points or less

CSCI 5342 - Sports Data Mining (Fall 2024)

# Tentative Course Schedule and Assignments

Date	Concept	Readings
Aug 26-Sep 1	Introduction to Sports Data Mining	
	Introduction to Baseball	AB Ch 1-3
Sep 2-8	A Brief History of Baseball Statistics - Chadwick to StatCast	
	Batting Statistics	AB Ch 4
Sep 9-15	Modern Batting Metrics	
	States and Expected Values	AB Ch 5
Sep 16-22	Balls and Strikes	AB Ch 6
	Pitching Statistics	
Sep 23-29	Modern Pitching Metrics	AB Ch 7
	Fielding Statistics	
Sep 30-Oct 6	Park Statistics	AB Ch 11
	Career Trajectories	AB Ch 8
Oct 7-13	Simulation	AB Ch 9
	Streaky Performances	AB Ch 10
Oct 14-20	Baseball Postseason Discussion	
	Baseball Postseason Discussion	
Oct 21-27	Football - Metrics, State & Value Analysis and Decision Making	
	Football - Passing, Rushing and Conversions	
Oct 28-Nov 3	Football - EPA, Advanced Analytics and Visualization	
	Live Fantasy Football Draft	
Nov 4-10	Basketball - Metrics, Linear Weights and +/- Ratings	
	60 Minutes – Basketball Analytics (Video and Discussion)	
Nov 11-17	How Stats Won Football from Moneyball to FC Midtiylland	
	Research Talk – Predicting Wins and Spread in the Premier League	
Nov 18-24	Case Discussion – Managing a Youth Soccer Organization's Data	SA pg 27
	How Data Analytics is Changing Professional Hockey	
Nov 25-Dec 1	Thanksgiving Break (No Classes)	
	Thanksgiving Break (No Classes)	
Dec 2-6	Winning at Daily Fantasy Sports Using Analytics	
	Case Discussion – Catching Managerial Issues using Analytics	SA pg 84