Sports Analytics Syllabus (Fall 2022)

Instructor

Professor Mark Broadie
Office: E-mail:
mnb2@gsb.columbia.edu Phone:
212-854-4103

Course work and grading

Grades will be based on the following weights:

- Class participation (including Pollev): 10%
- Homework (individual) and guided project (group of 5 or less): 40%
- Concept check quizzes (individual): 15%
- Final project (group of 5 or less): 35%

The guided project is a group assignment that will give you practice for the final course project. You can choose one of six projects: draft analysis, expected goals in soccer, pulling the goalie in hockey (analysis with simulation or dynamic programming), clutch performance in baseball, and win totals betting markets.

Concept check quizzes: There will be daily concept check quizzes that are to be done individually (no discussion with other students).

The final project will be on a topic of your group's choice (subject to approval). The final project can be done in groups of five or less (no exceptions). Project presentations will be given in the last class session.

Prerequisites

The prerequisites for the course are the *Managerial Statistics* and *Business Analytics* core courses. Good Excel skills are required and we will use the Business Analytics Excel add-in for analyzing data. This means that **you need to have a laptop computer that is either a Windows machine or a Mac machine that can boot into Windows in order to run Windows Excel.**

Detailed knowledge of sports is not required, but familiarity with the rules of baseball, American football, basketball, hockey and golf will be assumed. For example, in baseball, you should understand the basic rules and what the terms bunting and base-stealing mean as described in `https://en.wikipedia.org/wiki/Baseball_rules`. For football, you should know about plays, downs, and scoring, as described in `https://en.wikipedia.org/wiki/American_football_rules`.

Core culture

Core culture (present and on time, prepared and participating) is expected of everyone in the course. It is important for learning the material and for us learning from each other—and is more fun for all of us.

Connection to the core

This course builds on the Business Analytics and Statistics core courses. The course will use the tools of statistical analysis, predictive analytics, optimization and simulation. Issues of risk and return from Corporate Finance will be used in the course, as will game theory concepts from Managerial Economics and Strategy.

Course description

Sports analytics refers to the use of data and quantitative methods to measure performance and make decisions to gain advantage in the competitive sports arena. This course builds on the Business Analytics core course and is designed to help students to develop and apply analytical skills that are useful in business, using sports as the application area. These skills include critical thinking, mathematical modeling, statistical analysis, predictive analytics, game theory, optimization and simulation. These skills will be applied to sports in this course, but are equally useful in many areas of business.

There will be three main topics in the course: (1) measuring and predicting player and team performance, (2) decision-making and strategy in sports, and (3) sports betting and fantasy sports. Typical questions addressed in sports analytics include: How to rank players or teams? How to predict future performance of players or teams? How much is a player on a team worth? How likely are extreme performances, i.e., streaks? Are there hot-hands in sports performances? Which decision is more likely to lead to a win (e.g., attempt a stolen base or not in baseball, punt or go for it on fourth down in football, dump and chase or not in hockey, pull the goalie or not in hockey)? How to form lineups in daily fantasy sports? Are betting markets efficient? How to manage money in sports betting?

The main sports discussed in the course will be baseball, football, basketball, hockey, and golf. Soccer, tennis, and other sports will be briefly discussed. Students are welcome to pursue any sport in more detail (e.g., cricket, rugby, auto racing, horse racing, Australian rules football, skiing, track and field, or even card games such as blackjack, poker, etc.) in a project.

Class sessions will involve a mixture of current events, lecture, discussion, and hands-on analysis with computers in class. Each session will typically address a question from a sport using an important analytical idea (e.g., mean reversion) together with a mathematical technique (e.g., regression). Because of the "laboratory" nature of part of the sessions, students should bring their laptops to each class.

Guest speakers

There will be two or three guest speakers during the course. Past speakers have been professionals working in sports analytics groups for professional teams or sports analytics companies.

Recommended textbook: *Mathletics*, 2009, Wayne Winston, Princeton University Press.

Additional recommended books

- *Scorecasting*, 2011, Moskowitz and Wertheim, Crown Archetype
- *The Success Equation*, 2012, Michael Mauboussin, Harvard Business Review Press
- *Every Shot Counts*, 2014, Mark Broadie, Gotham Books
- *Trading Bases*, 2013, Joe Peta, Dutton
- *Analyzing Baseball Data with R*, 2013, Chapman & Hall/CRC

Software

Assignments are designed to be done in Windows Excel with use of the Business Analytics Excel add-in. This add-in runs under Windows Excel but not Mac Excel (this is a limitation of Mac Excel which does not have the capability to run numerically intensive add-ins of this type). You are free to do the assignments in python (or any other programming language). A library of sports analytics routines in python will be provided.

Course outline

Sessions 1-4. Measuring and predicting the performance of players and teams

- Course overview
- Rating field goal kickers in the NFL
- Ratings teams and measuring the strength of schedule
- Predicting future performance of players and teams
- Assessing alternative models of performance
- Streaks, momentum, mean-reversion and hot hands in sports

Sessions 5-6. Decision making in sports

- Markov modeling of games: states, transition probabilities and state values
- Baseball: analysis of bunting, base-stealing and other strategies
- Football: analysis of run versus pass, punt or go-for-it
- Assessing the value of a play using run value added and win probability added
- Decision making with strategic interactions using game theory

Sessions 7-9. Sports betting and course wrap-up with golf analytics

- Overview: betting markets, odds, "prop" bets, setting betting lines
- Investigating the efficiency of betting markets
- Setting betting lines
- Money management and the Kelly criterion
- Analysis of prop bets
- Golf analytics: performance measurement, decision making and sports betting

Sessions 10. Project presentations