Sports Analytics in Practice – Syllabus

Stern School of Business, New York University
Professors Daniel Altman and Philip Z. Maymin
Six-week seminar – Spring 2017, Thursdays at 6 pm

Introduction

Sports analytics is a fast-growing field that is moving far beyond the innovations in recruiting and tactics chronicled in Michael Lewis’s Moneyball. Teams in basketball, football, ice hockey, soccer, and now eSports have followed baseball in building analytics departments to support myriad aspects of decision-making on and off the field. This seminar aims to explore recent trends in sports analytics from a practical point of view, offering students the skills and ideas to create analytics of potential value to professional sports enterprises.

This year, the course will follow the narrative of a sports executive trying to recruit and trade players, determine coaching strategies, and maximize the probability of a championship in the next five years.

Registration requirements

The following requirements are non-negotiable:

- Completion of or concurrent enrolment in the Stern “Data Boot Camp” course, or prior coursework and permission of this course’s instructors
- Competence in a computer language and/or software package capable of statistical analysis: R, Python, SAS, Stata, Matlab, etc.
- An understanding of basic probability, statistics, and prediction, including linear regression, properties of the normal distribution, and types of errors

Students with questions about these requirements should contact one of the professors.

Lecturers

Daniel Altman is Adjunct Associate Professor of Economics at the NYU Stern School of Business. He is the founder of North Yard Analytics LLC, a sports data consulting firm that serves sports teams and leagues around the world, particularly in soccer. His analysis has been featured in The New Yorker, The
Economist, the Financial Times, The Telegraph, Bloomberg Sports StatsInsights, Quartz, and other media. He is also senior editor for economics at Foreign Policy magazine and the author of four books, including the international bestseller *Outrageous Fortunes: The Twelve Surprising Trends That Will Reshape the Global Economy*. He has a B.A., M.A., and Ph.D. in Economics from Harvard University.

**Philip Z. Maymin** Dr. Philip Z. Maymin is Associate Professor of Analytics and Finance at the University of Bridgeport Trefz School of Business. He is also the founding managing editor of Algorithmic Finance and the co-founder and co-editor-in-chief of the Journal of Sports Analytics. He has also been an analytics consultant with several NBA teams and is the Chief Analytics Officer for Vantage Sports. He has been a portfolio manager at Long-Term Capital Management, Ellington Management Group, and his own hedge fund, Maymin Capital Management. He has also been a policy scholar for a free market think tank, a Justice of the Peace, a Congressional candidate, an Assistant Professor of Finance and Risk Engineering at the NYU School of Engineering, and a columnist for American Banker, the Fairfield County Weekly and LewRockwell.com. He has a Ph.D. in Finance from the University of Chicago and a B.A. in Computer Science and M.S. in Applied Math from Harvard University.

**Assignments and grading**

Students will receive a large data set from a popular sport to form the basis of an analytics project of their choosing. Students may also choose to use their own data with approval of the instructors. They will also be encouraged to work in groups combining different skills (such as scraping, data wrangling, statistical analysis, programming, data visualization, communication, and the like). Each group will give a brief presentation in the final weeks of the course to receive feedback from instructors and classmates. The project will be completed within one week after the end of the course and will constitute 70% of each student’s grade. The remaining 30% will come from class participation as evaluated by the professors and teaching fellow. Attendance is not the same as participation.

**Lectures**

**Week 1: Professor Maymin**

Understanding sports data

By standard definitions, big data can mean large quantities of data, or a large variety of irregular data types, or a real-time high-frequency feed of data. In sports contexts, big data currently tends to be of much smaller breadth and
scope than big data applications in other contexts, but typically too large and too irregular to fit into a standard spreadsheet. Modern sports analytics now needs to include approaches and systems for processing and analyzing optical tracking, social networks, and other comparatively large new datasets. What data should our executive use to make decisions, and how should these data be handled?


**Week 2: Professor Altman**
Creating performance metrics for players and teams

Sports executives and coaches want to measure the performance of individual players and entire teams; the resulting metrics can guide recruiting, tactics, and financial planning. But many metrics offer little or no information that can contribute meaningfully to decisions. To avoid this problem, metrics must have properties of aggregation, consistency, predictiveness, incentive compatibility, and more. These properties apply regardless of the sport in question, though in practice analysts must choose among a variety of suboptimal options. What metrics should our executive use to evaluate players and gauge the benefits of specific tactics during games?

- **Priority Reading**: Altman, Daniel (2013), “The ten traits of ideal soccer metrics.” Bloomberg Sports StatsInsights (September 18)


Week 3: Professor Altman
Forecasting and prediction

Forecasting results is important not just for bettors but also for teams, both in deciding to whether to strengthen their rosters and planning their future finances. Many forecasting models depend on power scores, but others use ratings inspired by chess rankings or simulations. Incorporating Bayesian aspects into the prediction of results is also an important innovation. A final component is the application of constraints or assumptions based on the historical behavior of teams and leagues. How will our sports executive know the probability of a championship in the next five seasons?


Week 4: Professor Maymin  
Machine learning in sports

The original statistical revolution in sports was largely an inspired application of linear regression to available data. The next wave will come from leveraging the advances in machine learning techniques, allowing a computer to learn on its own from past data. Some of the hottest techniques include random forests, support vector machines, nearest neighbors, and deep learning. We will go over the basics of these approaches, and look at applications to drafts and player projections from past performance. How can machine learning help our executive to identify useful players, tactics, and overall approaches to winning?


Week 5: Professor Maymin  
Sports data visualization

Standard bar charts, histograms, and time series plots are often extremely unhelpful in sports contexts where the goal is usually not to present a single answer but to help refine an interesting question. Better sports data visualization does not necessarily have to be more complicated—for example, basketball shot charts are simple and effective—but it does have to do a better job at quickly providing the needed information. Human eyesight is fine-grained and attuned to pattern and complexity recognition and comparison: a good visualization is usually far more important and useful than a great quantitative model. How will our executive present analytical results in order to convince ownership to adopt new priorities?
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7042478&isnumber=7042473

• **Priority Reading:** Owens, Sean Gabriel; Jankun-Kelly, T.J. (2013), "Visualizations for Exploration of American Football Season and Play Data," presented at the 1st Workshop on Sports Data Visualization at IEEE VIS in Atlanta.  

http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7042533&isnumber=7042473

http://ssrn.com/abstract=2200560

Week 6: Professor Altman

Profit and loss

Sports analysts frequently assume that the teams they are modeling have points, wins, positions in league standings, or trophies as their objectives. But sports teams are businesses, and sometimes even public companies. As a result, the ultimate objective of many decisions may be long-term profitability. Transforming game-based metrics into cash-based metrics is therefore an essential part of the analyst’s toolkit. We will consider how to evaluate players as capital assets and assess their transfer or trade values. Will our executive’s strategy for winning be profitable or not?

• **Priority Reading:** Altman, Daniel (2015), “Player contributions to cash flow,” presentation.  