 MA 251: Stats & Sports project

**Topics:**  Anything in any sport

**Content:** Creativity

Visualization

In and out of sample testing

Confidence intervals/regression/hypothesis testing/inference/measurement error

**Evaluation of presentation (10 points each)**

Creativity, Visualization, Depth of content, Accuracy of content, Presentation (volume, presence, eye contact, clarity)

**Evaluation of Markdown paper: formal paper (10 points each)**

Clarity, Presentation (neatness, spelling grammar), Feedback from presentation

**Hints:**

1. Focus (i) on asking a really interesting question and (ii) on picking right tool for your question, as opposed to just applying several different tools to random questions.
2. Be precise
3. Make sure you have data to answer your question

**Deadlines:**

**11/13: Project proposal, done in Markdown, printed and handed in (10 points)**

**11/20: Revised proposal, done in Markdown, printed and handed in (10 points)**

**11/27: Data, in .csv format (emailed, 10 points)**

**12/11: Presentation (50 points)**

**12/17: Final paper due, 5:00 to my mailbox in Harder Hall (30 points)**

**Proposal**

1. What question in what sport are you trying to answer
2. What techniques (fitting the criterion above) will you use
3. What newer techniques will you learn
4. What is your outcome variable (if applicable), and what are your explanatory variables. Also, state if these are numerical or categorical
5. Where does your data come from? If you are collecting it, identify how
6. What is your sample size?

**Revised proposal**

* Respond to feedback on initial proposal

**Data**

* Provide data in .csv format
* Simple and appropriate column labels
* Careful with capitalization/missing data

**Technical report (Markdown)**

General:

* Use hyperlinks for references (nothing formal needed except names).
* In general, show most code

Introduction ~ 1 pages:

* Why is your problem important? What other work has been done?
* Set-up what you are going to do. Why did you choose (method) that looks at relationship between (variable 1) and (variable 2)
* Describe data. Where does it come from? What are variable types/units?

Visualization & Descriptive statistics ~ 2-3 pages.

* Clear and concise summary of your data.
* Do not just show code and output. Describe it using words

Inference/predictions ~ 2 pages

- Clear and concise summary of your findings

- Statistical vs. practical significance

Conclusion ~ 1 pages

* Practical findings
* Who do your results generalize to? Other sports?
* Weaknesses in your design?
* Limitations in who was sampled or in how you collected the data
* Areas for future work