Sports Analytics Project

Fall 2023

A key part of sports analytics is preparation for applying sports analytics techniques to real sports problems. There are many benefits to this, including experience of carrying out an analytics project which will be transferable to other lines of work, and if you are interested in working in the field of sports analytics then many employers will look for prior experience of applying analytical techniques to sports analytics data.

The objective of this project is to carry out an analysis in sports that achieves one of the following objectives:

1. Answers or provides insight into an important sports question
2. Provides a benefit to a team/coach/player regarding some aspect of their sport
3. Predicts the outcomes of a sporting event

The deliverable of the project can take any of the following forms:

- A statistical analysis with documented results
- A visualization analysis that displays important metrics for a team
- A modelling analysis that uses a predictive model.

The steps involved in the project should include:

1. Identify a suitable question or problem of interest
2. Find a dataset that is suitable for answering the question/solving the problem
3. Create a project plan
4. Outline the dataset and describe the project
5. Carry out the analysis and produce a deliverable
6. Analyze the results of the project and extract actionable insights

Teams

You may work in teams of 1-5, it is up to you if you would like to work in a group or not. Grading will take group size into account in that more detailed work is expected from a larger group. If working in a group, please include a section in both the short report and the final report outlining the work done by each member of the team. We will discuss this in class. If working with team data that is sensitive, I appreciate you may wish to work alone.

Project Deliverables:

This project will account for 40% of the final grade for this course split amongst the following deliverables:

1. **Initial Proposal** outlining problem, dataset, and the project plan due on 11/15/2023 (10% of final grade)
2. **In-class presentation** delivered in the final week of class (11/28/2023 and 11/30/2023) presenting an executive summary of the project and your findings. (10% of final grade)
3. **Final detailed report** of project and process with code if appropriate, due on last day of class 12/11/2023. (35% of final grade)

Initial Proposal

Once you have decided upon a potential project, found a suitable dataset, and carried out an initial investigation of the data, please submit an initial proposal. This should be a short document describing the project you intend to carry out along with an overview of the dataset. This should include some initial summary visualizations. If you wish to discuss the project prior to submitting the initial proposal, please let me know and we can arrange a time to meet over zoom either during office hours or some other time.

Format:
Your problem outline report should be 1-3 pages in length. Please include code or graphs but note that these do not contribute to the page limit. Please include the following sections:

1. Motivation - What is the problem you are considering and why is this an interesting problem? Has this problem been attempted before?
2. Problem framing - What is the proposed solution to the problem?
3. Data overview - What is the dataset you are going to use to solve this problem? Describe the dataset and its characteristics.
4. Contribution (Does not count towards limit) - Which team members contributed to each area of the project.
5. R-code (Does not count towards limit) - R-script with code used for initial data analysis.

Please submit the short report via Canvas or email to mbarron2@nd.edu by 6.59pm on 11/8/2023. The report should be either a word document or a pdf file. The r-code should be sent as a separate r-script. The dataset used should also be submitted, the best method for this is likely to share the dataset via google drive.

The grading for the presentation will be based on the following factors:

1. Clarity of the proposal.
2. Accurate conveying of problem and solution.
3. Novelty of the problem.

In-class Presentation

The in-class presentations will be an executive summary of the problem, the solution, and findings. This should be treated as a high-level overview and does not need to go into extreme detail on the method used. In industry, especially sports, you will often present work to those who are less technical than you are and therefore you should attempt to describe the problem, briefly describe how it was solved and then discuss the insights gained.

The in-class presentations will be approximately 5-10 minutes in length.

The grading for the presentation will be based on the following factors:

4. Clarity of presentation.
5. Accurate conveying of problem and solution.
6. Insights drawn from the results.

Final Report

The final report should provide a comprehensive overview of the project and the findings that were generated from the project. This report should be readable and interpretable by someone who has no prior knowledge of your analysis.

The grading for the final report will be based on the following factors:

1. Clarity of presentation and explanation of project findings and results.
2. Technical quality of the work - Does the project make sense? Was the application of the methods reasonable?
3. Significance of the problem - Was this a real and interesting problem? Were the authors able to extract novel insights from the analysis?
4. Analysis of results and insights that can be drawn from the results.
5. Next steps - what would you do differently, what would be the process for improving the modelling results.
6. Peer grading (Teams only) – For those who work in teams part of the grade given to students will be determined by their team members. Each team member will be required to submit a grade for each of the members of their team.

Format

Your final report should be approximately 2-5 pages in length and not more than 6. Feel free to include additional graphs and figures that you believe provide supporting information, these will not count towards the page limit. The below is a rough guide to the format of the final report but feel free to modify it as you feel best suitable and feel free to discuss it with me in advance of final submission:

- Introduction (0.5-1 pages) - Outline the problem, its significance, and your motivation for pursuing the problem. Describe the inputs and outputs to the project in detail. An abstract paragraph outlining the entire project can also be included.

- Related work (0.5 pages) - What previous work has been done in this area and by whom? Remember to cite your sources. How is your approach different/an improvement on previous attempts to solve the problem?
- Methods (0.5 - 2 pages) – Describe the methods and approach used to solve the problem. Include the steps taken in your analysis and the reasons why.
- Discussion (1-2 pages) - Discuss the results/findings coming from the analysis. What were the insights gained from the analysis? What actions can be taken off the back of the project? Include suitable visualizations to support your conclusions.
- Conclusion and Future work (0.25-0.5 pages) - Provide a summary of the report and the key points from the analysis. What would be the next steps in this project given more time and resources?
- Contributions (Does not count towards limit) - What work was done by each of the team members on the project?
- Bibliography (Does not count towards limit) - List of sources used for the project.
- R-code (Does not count towards limit) - This should be provided separately as an r-script and should allow the findings and graphs in your project report to be recreated.

Your font size should not be smaller than 10pt for the final report. The final report can be submitted as a word document or pdf if you so wish. The r-code or otherwise used for the project should be submitted as a separate r-script.

Please submit your final report via Canvas or via email to mbarron2@nd.edu by 6.59pm 12/11/2023.

Citations:

For the project reports please include appropriate citations of sources which you make use of. These should be included in the report with references at appropriate points in the text where the information is used, and a bibliography should be included at the end of the report listing the sources used.

Some citation tools can be found at:
https://libguides.library.nd.edu/business-general/citing-sources

A guide to citing sources can be found here:
https://libguides.library.nd.edu/plagiarism