Sprint Challenge 1

Josh Gaughan

September 7th, 2025

Question 1

Analytics have become especially important for today's sport business because of the prevalence of measured data and the insights and advantages that can be gained by utilizing them. The use of analytics is also important because teams are trying to find the best ways to protect their athletes, which are essentially their investments (Nibali 49). Data is also more easily obtainable because of technological progression which allows for non-invasive data collection (49). This gives teams even more incentive to use analytics as part of sport business. Nibali also gives a wonderful section containing "The driving forces of sport analytics". In it, she mentions the importance of not only using analytics properly, but also managing the data necessary to perform relevant analysis. The data and subsequent analysis includes wellness, GPS and performance data among several others which are used to monitor athlete health status (50). As mentioned in the case study of Michael Rennie regarding pro Australian Football, analytics can be imperative when planning training. GPS devices and the countless data they provide allow coaching staff to change training to better suit the condition of each athlete. This will help prevent improper, ineffective training (57). Analytics are also important for the prevention of injuries. According to the case study, there are about 150 missed games per club per season in Australian Football. The use of analytics allows for quicker, more effective rehabilitation of athletes (60). This means that the sports organizations most valuable investment, the athletes, are in playing condition as quickly as possible.

While analytics are helpful for athlete performance, they are also helpful for financial performance. In the mid 2010s, the Atlanta Hawks used analytics to examine ticket holder arrival times and traffic patterns to determine the optimal game time. Based on their findings, they decided to change their weekday home game start times to 8:00 pm from 7:30 pm. This produced a greater number of fans present at the time of tip-off (Bukstein 1-2). Analytics are also important for ticket sales and controlling the secondary market. Through ticket analytics the Bruins discovered the existence of extra ticket resellers in 2016. They cancelled ticket access to these resellers which allowed for greater ticket availability for individual fans (10). Analytics are also an important part of market research done by sports organizations. Market research allows them to increase revenue through branding and fan participation and satisfaction. The result is a more informed organization which is now equipped with insights that allow for the creation of a better fan experience. (17)

In the face of confirmation bias, the importance of analytics is highlighted when referring to sport business success. In a study done by David Romer, he analyzed fourth down occurrences and developed a model to predict success of a fourth down conversion versus kicking a field goal. He found that coaches departed from the decision that would produce the most success (McNerney). This could be blamed on cognitive bias. It is even more important to recognize the role analytics play because they could lead to better outcomes than coaches' intuition. Analytics are also important for providing a more objective view on phenomena and trends. The narrative fallacy, which is a "limited ability to look at sequences of facts without weaving an explanation into them…" (McNerney), partially explains why people will cherry-pick stats that they think are relevant to success or failure. Well sourced analytics are able to provide a more reliable justification for a team's performance.

Question 2

The CRISP-DM or Cross-Industry Standard Process for Data Mining is a standardized method for analytical problems to be addressed. Its purpose is to bring consistency into the field of analytics and provide a clearer path to success for those who work in the industry. It consists of 6 phases, which are typically followed in a particular order but also encourages individuals to be iterative and return to previous phases when necessary. Those phases include the business understanding, data understanding, data preparation, model generation, model evaluation and solution deployment. The business understanding phase includes familiarizing oneself with the goals a business has when beginning to tackle an analytics project. The data understanding phase consists of data collection and familiarization. One may choose to make initial observations of the patterns of the data. Data preparation includes the completion of the dataset that will be used in the following phases. One will also clean and organize the data, doing any transformations necessary for the models to run. The model generation phase is when statistical analysis is done using the data that was previously cleaned. Depending on the models being used, one might choose to revisit the data preparation phase to modify data to the models requirements. The model evaluation phase consists of understanding the results of the well-tuned models one has produced. Here, you will choose how the results will be used. The final phase, solution deployment, consists of gathering and presenting the results of all your work. Depending on the context, the presentation of these findings can be simple or not.

The descriptive, predictive and prescriptive frameworks are three semi-related perspectives of data analytics that can be used to answer different questions. Descriptive analytics is when someone tries to explain things that have already happened using data that already exists. For example: how many presidential votes were collected for each candidate in the 2024 presidential election in Marion county? The insights from the descriptive framework can be used to make better choices in the future. Predictive analytics on the other hand means someone is trying to

predict or forecast something that is expected to happen in the future by using past data and patterns. For example: How much will the total amount of gasoline purchased fluctuate if per gallon prices increase? Like the descriptive framework the insights gained from the predictive framework can also be used to make better choices for the future. However, they can also give an idea of what the results of those decisions may look like. Prescriptive analytics is an expansion of both the descriptive and predictive frameworks. It entails the use of past data to find the most efficient path to success. This framework provides a direction for someone to go into in order to maximize/minimize whatever it is they are seeking. For example: how many solar panels do we need to buy in order to save on energy costs while still getting the amount of energy we need? Results from the prescriptive framework are actionable and are meant to be used to make logical decisions.

Question 3

https://www.youtube.com/watch?v=V3OaXc0Heco&list=PLhbPeSFiFnAbTHGwepXMqLhD2RwTex0jQ&index=52

Above, I have linked the youtube page for the conference video I watched. It was at SSAC19 and the topic was: "Building Smarter Venues: How Gillette Stadium Uses Data to Drive the Bottom Line." Bronson Green, the director of strategic operations with Kraft Sports Management, begins the presentation by quickly addressing the first phase in the CRISP-DM, which is business understanding. Green talks about their partnership with KAGR as being important for making Gillette Stadium as efficient as possible using data-backed insights (Green 0:56). They understand the business related goals of making even more informed decisions when it comes to changes to Gillette stadium's processes. They also recognized that while they want to increase the scope of the property, they also want to be strategic in managing the new spaces so the experiences aren't lacking due to poor management (5:45).

The next phase covered was the data understanding phase. They wanted to "get our arms wrapped around the data" (6:25). In other words, they wanted to use KAGR to familiarize themselves with the data so they could begin using it. Green mentions the desire to visualize the information so their attention is focused on projects that are the most important to their business. He also mentions trying to understand the trends of workers in the stadium so they can make more informed decisions about who to hire and how long they should be working for.

The data preparation phase was also briefly discussed. Green talks about the data collection process and how their maintenance system was out of date. They went through the process of cleaning up the data that was poorly formatted or even missing (20:49-21:07). He also talks about a nine to twelve month process to get everything running which included the creation of dashboards, cleaning the data and producing visualizations (30:15).

While the model generation and model evaluation phases aren't technically covered in this presentation, Green does talk about analysis done on the data. He doesn't discuss a specific statistical model that was built and evaluated but rather talks about the various visualizations they produced from each dashboard.

Green discusses the solution deployment phase which includes deploying operational changes based on the data and its insights to see changes in efficiency. He also talks about multiple reports produced from the efficiency data in the dashboards. These reports are used to present the data and key areas of action for management staff at Gillette stadium. He also mentions how "this information can be used in contract negotiations" (23:17) to hold workers accountable so they work in a manner supported by the data. Additionally, as new data emerges, it can be used to make more informed decisions beyond worker insights . For example, Green talks about understanding money spent on utilities and how that can become more structured (24:37).

Question 4

In the context of player analytics, 5 potential analytics questions are:

1. What sort of contributions do individual players have to their team? What specific attributes do players have that make them valuable to a team's success?
2. In what situations do certain players succeed or fail? Do they tend to perform better in high stakes scenarios or against difficult opponents or in an away setting or not?
3. What is a player's tendency to get injured or be unfit to play? Are they on the older side or have prior injuries that reduce their reliability?
4. In what way is a player getting better or worse over time? Are they showing signs of improvement as their experience level and exposure increases?
5. What position does a player's attributes link up with the most. If they are used to playing a position (say in college or high school), does that automatically mean they should play that position?

Question 5

My ideal career goal is to work in the field of golf analytics. There are a variety of ways that professional golf is measured that provide data that can be analyzed. Specifically, raw shot data provides so much information that can be turned into individual player performance metrics. Additionally, environmental and course related data have significant influences on certain players. For example, course conditions (like green/fairway firmness and speed, course length and rough length) and weather conditions (like temperature, humidity and wind) will cater to certain players over others. All of this data can be combined together to determine a player's likelihood of success at a tournament, round or even on an individual hole.

Works Cited

Nibali, Maria. "The Data Game: Analyzing our way to Better Sport Performance", *Sport Analytics: A data-driven approach to sport business and management*, edited by Gil Fried, Ceyda Mumcu, 2016.

Bukstein, Scott. Evolution and Impact of Business Analytics in Sport. 2016

McNerney, Samuel. "Cognitive Biases in Sports: The Irrationality of Coaches, Commentators and Fans." Scientific American, 22 Sept. 2011, https://www.scientificamerican.com/blog/guest-blog/cognitive-biases-in-sports-the-irrationality-of-coaches-commentators-and-fans/

Green, Bronson. "Building Smarter Venues: How Gillette Stadium Uses Data to Drive the Bottom Line." YouTube, uploaded by 42 Analytics, 9 Mar. 2019, https://www.youtube.com/watch?v=V3OaXc0Heco&list=PLhbPeSFiFnAbTHGwepXMqLhD2RwTex0jQ&index=53