Final Project Synthesis Paper by Masyn Grisel

Analytics plays a huge role in professional sports today. Teams use data more than ever to make smarter decisions and gain an edge over the competition. For my project, I’m building a performance prediction tool that baseball organizations can use to get a better idea of how their team ,or their opponent, might perform in an upcoming game. It’s meant to provide helpful insights by looking at both historical stats and current context.

The model will consider a few key factors that are known to influence team performance: recent performance trends, starting pitcher stats, strength of the opposing team’s pitcher, whether the game is at home or away, starting time of the game(day vs night), and the ballpark they’re playing in. These can all be used in baseball analysis and should give a realistic picture of what to expect going into a game.

This project ties directly into what I’ve learned in my program. It uses skills from data analytics, data management, software systems, and cybersecurity/networking. The prototype will be simple and focused, but the bigger idea behind it could be useful for teams in scouting and game planning. It’s a good example of how different areas of technology and analysis can come together to solve a real-world problem.

The following section outlines my understanding of the four curriculum areas – Data Analytics, Data Management, Systems Development, and Networking/Cybersecurity- and explains how each plays a role in the development and purpose of this project.

**Data Analytics**: To me, data analytics is the process of examining and interpreting data to uncover patterns, trends, and insights that lead to more informed decisions. In this project, I’ll use data analytics to explore which factors—like offensive performance, pitcher effectiveness, and ballpark conditions—most strongly impact team outcomes. By analyzing historical and contextual game data, I’ll look for meaningful patterns across matchups, locations, and rest days. These findings will support simple predictive models, such as linear regression, to estimate how a team might perform under given conditions. The goal is to turn these patterns into actionable insights that help teams make better decisions in preparation, scouting, and strategy.

**Data Management:** Data management, in my view, is about organizing, securing, and ensuring the reliability of data for easy access and analysis. Effective data management is key to making informed decisions and running accurate analyses. In this project, I’ll gather and store historical and contextual game data—like team metrics, pitcher stats, and ballpark conditions—in a relational database. This structure will allow me to link data points (team, pitcher, game) and retrieve them easily for analysis. I’ll also perform data cleaning and transformation tasks to correct errors, handle missing values, and standardize stats. Ensuring security through measures like user authentication and access control will protect sensitive data, while the organized structure will make it easy to use when building predictive models.

**Systems Development:** Systems development involves building, testing, and maintaining software that addresses a specific problem. For this project, it’s essential to develop a functional tool for predicting performance. I’ll create a web-based application using Django to integrate the front-end and back-end. The app will include user authentication, allowing users to view real-time performance data for teams and pitchers. The backend will handle data, logic, and user requests, using Django’s tools to connect to the database and display predictions. The user interface will be clean and intuitive, enabling easy input of teams and games, and displaying predictions clearly. Lastly, I’ll conduct unit and integration tests using Django’s testing framework to ensure the system is robust and error-free.

**Networking and Cybersecurity:** Networking and cybersecurity involve protecting systems to prevent unauthorized access and ensure smooth operations. For this project, I’ll ensure secure access to external data, while maintaining efficiency. The system will include a basic front-end interface, where access control will ensure only authorized users can interact with sensitive data, especially if a login or admin panel is added. If hosting the prototype online, I’ll ensure secure data transmission with HTTPS and properly manage user sessions to prevent unauthorized access.

The following section highlights how the four curriculum areas work together and support each other in my project.

The four key areas, Data Analytics, Data Management, Systems Development, and Networking/Cybersecurity, each bring something essential to the table when building my performance prediction tool.

Data Analytics is at the heart of the project. It helps me identify important patterns and trends in the game data, like which factors matter most when predicting how a team will perform. By using data analytics, I can turn raw data into actionable insights for the teams.

Data Management makes sure all that data gathered in data analysis is well-organized, accurate, and easy to access. If the data isn’t structured properly or is full of errors, it would be impossible to build reliable predictions. Using databases like MySQL or PostgreSQL, I can ensure the data is clean and easy to retrieve when it’s time to run analysis and generate insights.

Systems Development ties it all together. I’ll be using Django to build the application itself, which will let users interact with the system and get the predictions. The system will be a functioning tool that uses the first two curriculum areas and builds a user interface to allow for real predictions that will inform better decision making. This step helps turn all the data into an easy to use tool.

Finally, Networking and Cybersecurity is crucial to keep everything secure. I will uses secure protocols like HTTPS and add user authentication to ensure that only authorized users can access and interact with sensitive data.

Together, these areas form the backbone of the performance prediction tool, making sure it’s not just functional and insightful, but also secure and user-friendly. It’s a blend of technology and strategy, all working together to create a valuable resource for baseball teams.

The following section describes my previous projects and how they are connected to this project.

* **Django (Online Shop and Blog Projects):** Throughout my Django projects, I built web applications that integrated multiple functions, such as user authentication, data storage, and front-end display. These projects allowed me to practice systems development and data management using Django's architecture, where I handled the storage and retrieval of data, ensuring that the systems were secure, well-organized, and accessible. This experience directly ties to my current project by utilizing Django's capabilities for building a web-based tool that makes real-time predictions using performance data.
* **Excel (Data Analysis and Visualization):** I have used Excel extensively for data analysis, including building dashboards and performing statistical analysis. This hands-on experience with pivot tables, filters, and charts gave me a deep understanding of how to analyze and visualize data for actionable insights. In my current project, this knowledge helps me understand how to structure and clean data before performing analysis, as well as how to interpret performance trends to build predictive models.
* **RapidMiner (Logistic Regression for Predictive Modeling):** In a previous course, I applied logistic regression techniques using RapidMiner to make data-driven predictions, such as analyzing the likelihood of specific outcomes based on historical data. This experience with predictive modeling in RapidMiner has been instrumental in shaping my approach to the current project, where I plan to apply similar data analysis techniques to predict team performance based on a variety of factors.

In conclusion, this project demonstrates the integration of various skills I’ve developed throughout my coursework and hands-on projects. By applying data analytics, data management, systems development, and networking/cybersecurity, I’ve been able to create a robust tool for predicting baseball team performance, drawing on the strengths of each curriculum area. My experiences, combined with the lessons learned from my prior projects, have equipped me with the tools needed to build a solution that supports data-driven decision-making in a real-world scenario. Through this project, I’ve gained a deeper understanding of how technology can be leveraged to inform strategies, optimize performance, and ultimately enhance decision-making in the world of sports analytics.