The objective of this project was to build a predictive model for MLB baseball attendance based on twenty years of historical game log data, as well as provide actionable insight into underlying factors that influence home game attendance. Overarchingly, we aim to provide recommendations on operational, revenue, and financial strategy impacts for the Minnesota Twins.

Firstly, we have developed and compared the performance of several models following a variety of linear and non-linear approaches. Our final predictive model has shown a relatively high degree of accuracy in forecasting attendance for MLB games. Data exploration contributed to identifying factors such as seasonality, day of week, time of day, opponent, and previous game attendance to all have a significant impact on attendance levels.

Additionally, our analysis has provided insight and improvement into several areas of business operation that currently affect and can have a positive impact on stadium attendance and revenue. Moreover, we have collected external sources of stadium specific data to include factors such as seating capacity, city population, and club value and operating income.

We have begun mapping a set of actionable recommendations to help the Minnesota Twins optimize their operational and financial strategy, with the goal of maximizing attendance and revenue generated by both ticket, concession, and merchandise sales. These recommendations include utilizing dynamic ticket pricing, improving facility/labor processes and amenities, and leveraging targeted marketing through themed games, exclusive ticket bundles, and promotional giveaways.

There have been several technical risks throughout the project, such as feature engineering and the proper handling of existing and created variable, as well as the risk of an overfit model, that we have been managing and mitigating through a variety of measures. These measures have included close project management and regular communication within our group, as well as robust testing and validation of our predictive model.

To ensure a successful delivery of the project on time, we are currently focused on refining our recommendations and findings and reassessing our project in the scope of our sponsor’s expectations, to develop an implementation plan for our final insights and strategies. We are also conducting additional analysis to identify any potential areas of improvement for our predictive model to specifically account for progressive team performance; we will continue exploring other MLB organizations and gain understanding into what attendance and financial success looks like across the league. We believe that these next steps will help to ensure a successful delivery of the project and provide the Twins with the insight and recommendations they need to optimize their operational and financial strategy moving into the 2023 season.