



# Understanding Major League Baseball Attendance

March 25<sup>th</sup>, 2023

## Who's on First

Aakriti Aneja, Ram G S, Justin Mason, Divyansh Tripathi  
University of Minnesota – Carlson School of Management

# Baseball - America's Favorite Pastime is facing more competition than ever



As a business we want to deliver the *best possible experience* to the fans

Knowing what affects attendance and expected attendance helps us take informed business decisions leading to best possible fan experience



Our Solution is a one-stop-shop to take *informed business decisions*

## Our Solution

Interpretability

What affects Attendance?

&

Predictive

Expected Attendance

To help *manage finances, drive  
profitability & fan experience*



# Data Sources

## Primary Source

Game-Logs were used as the primary data source from which we team attributes (such as win record, rank, attendance), player attributes, and calendar features.

## External Sources:

- **Weather data:** *meteostat* API - The data provided by organizations like NOAA and DWD
- **Holidays:** Holidays API – Federal Holidays list in the US
- **Demographic & Economic indicators:**
  - Population - US Bureau of Labor Statistics
  - Consumer Price Index (CPI) - US Bureau of Labor Statistics
  - Unemployment rate - US Bureau of Labor Statistics
- **MLB team information:**
  - Division – ESPN
  - Payroll / Salary – The Baseball Cube

# Top Factors influencing Attendance for Minnesota Twins

Our Solution

Interpretability

Predictive



Day of Week  
(Mon vs Sat)

Holiday Games

Month of the year  
(Apr vs Aug)

MIN Last Season  
League Rank

Unemployment  
rate

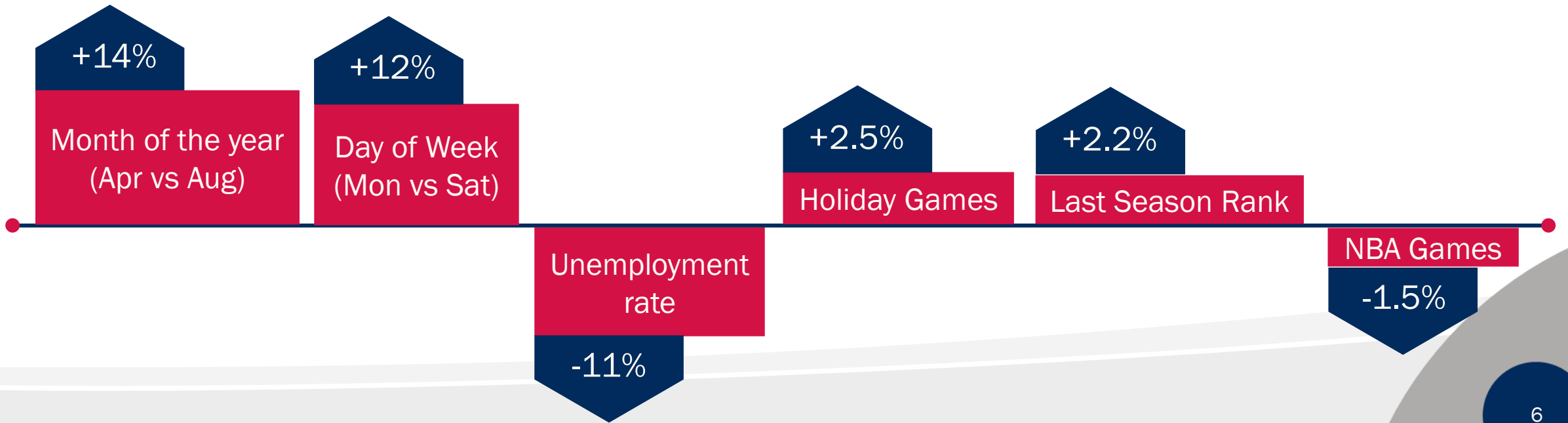
NBA Games

# Top Factors influencing Attendance for Minnesota Twins

Our Solution

Interpretability

Predictive



# Factors influencing attendance across Divisions

*\*For recent seasons*

## Our Solution

Interpretability

Predictive



## Geography

### AL Central



Highest attendance in

**June & Aug**

12%

4%

*Colder states see highest attendance in warmer months*

## Weekday

### AL West

Attendance **NOT** affected by the **day of the week**

*Only division indifferent to week games (M-Th)*

## Covid (Residual)

### NL Central AL East

Still have not recovered their attendance to pre-Covid levels

*Some division are affected by Covid more than others*

## NBA Season

### AL Central

**-6.5%**

Drop in attendance during NBA season

*Only division affected by other prof sports*

# Factors affecting Baseball Attendance over the years

Our Solution

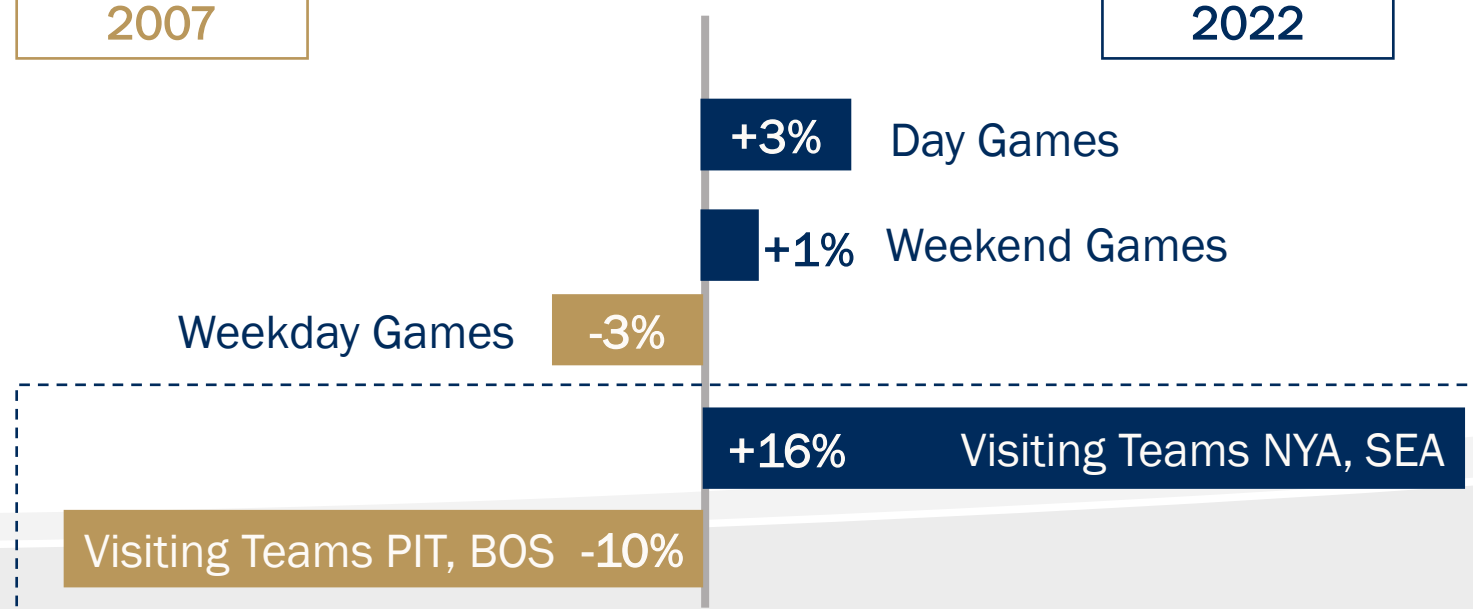
Interpretability

Predictive



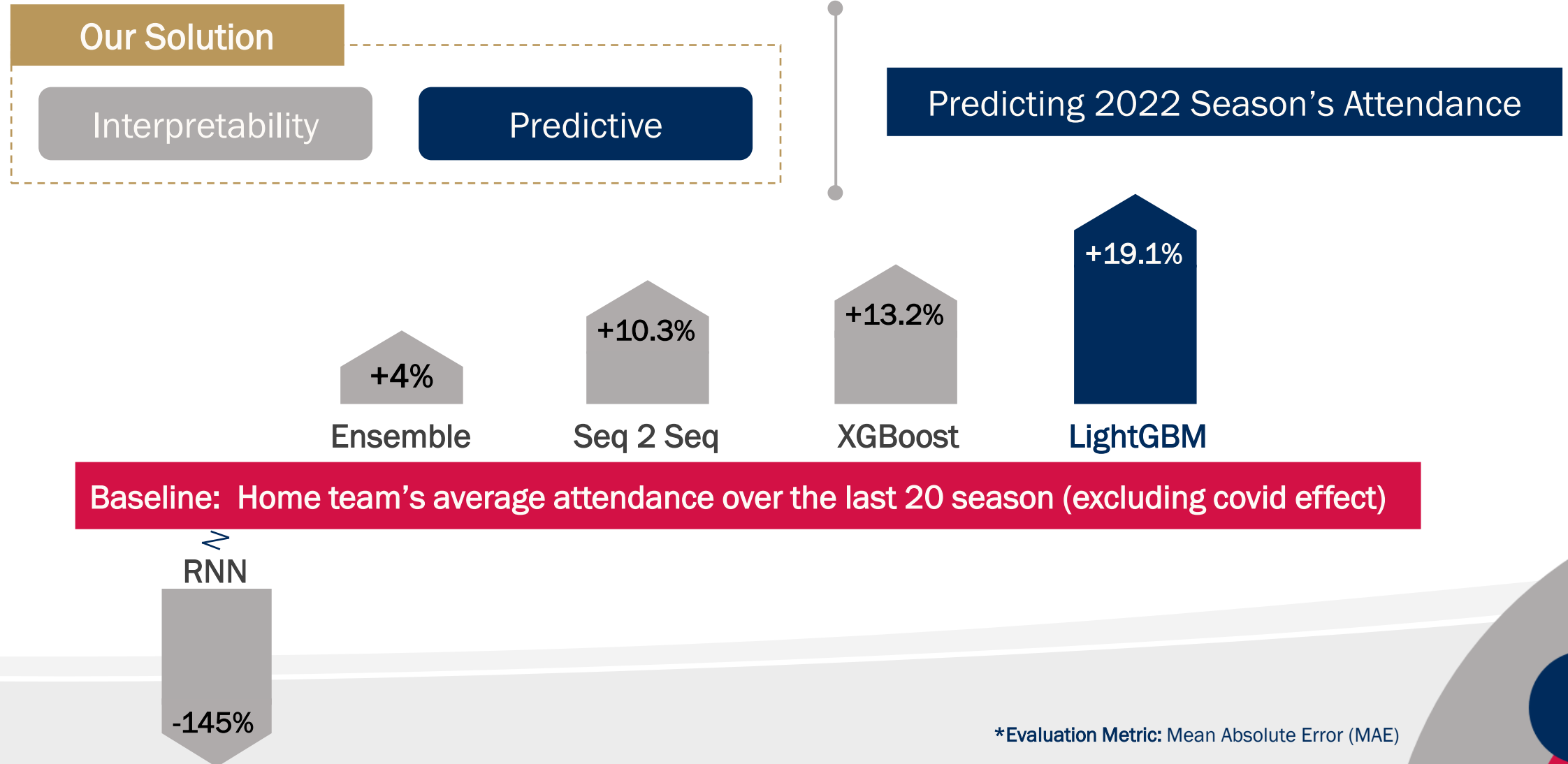
2007

2022

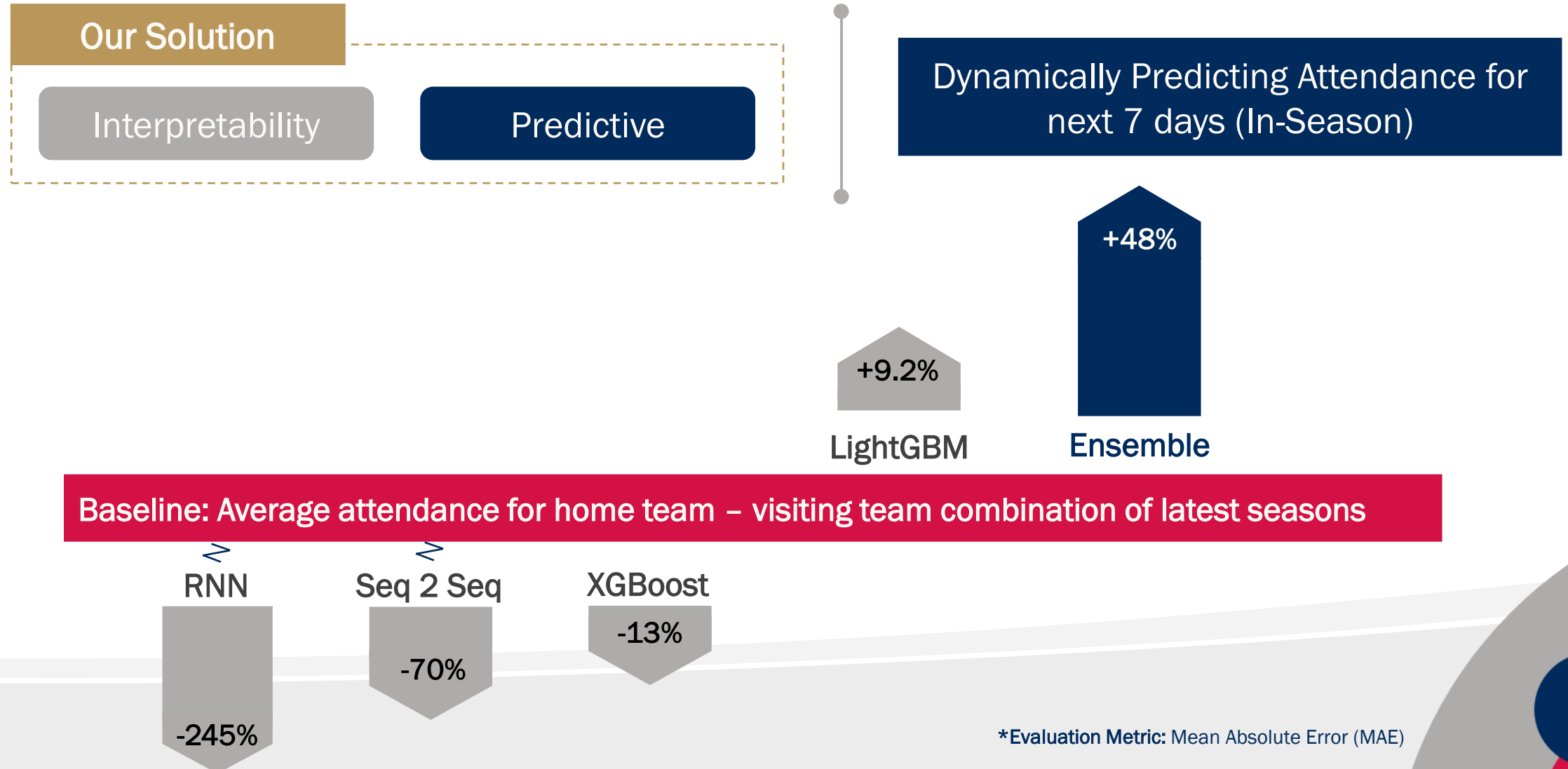




# Predictive Component – Season's attendance



# Predictive Component – In-season attendance

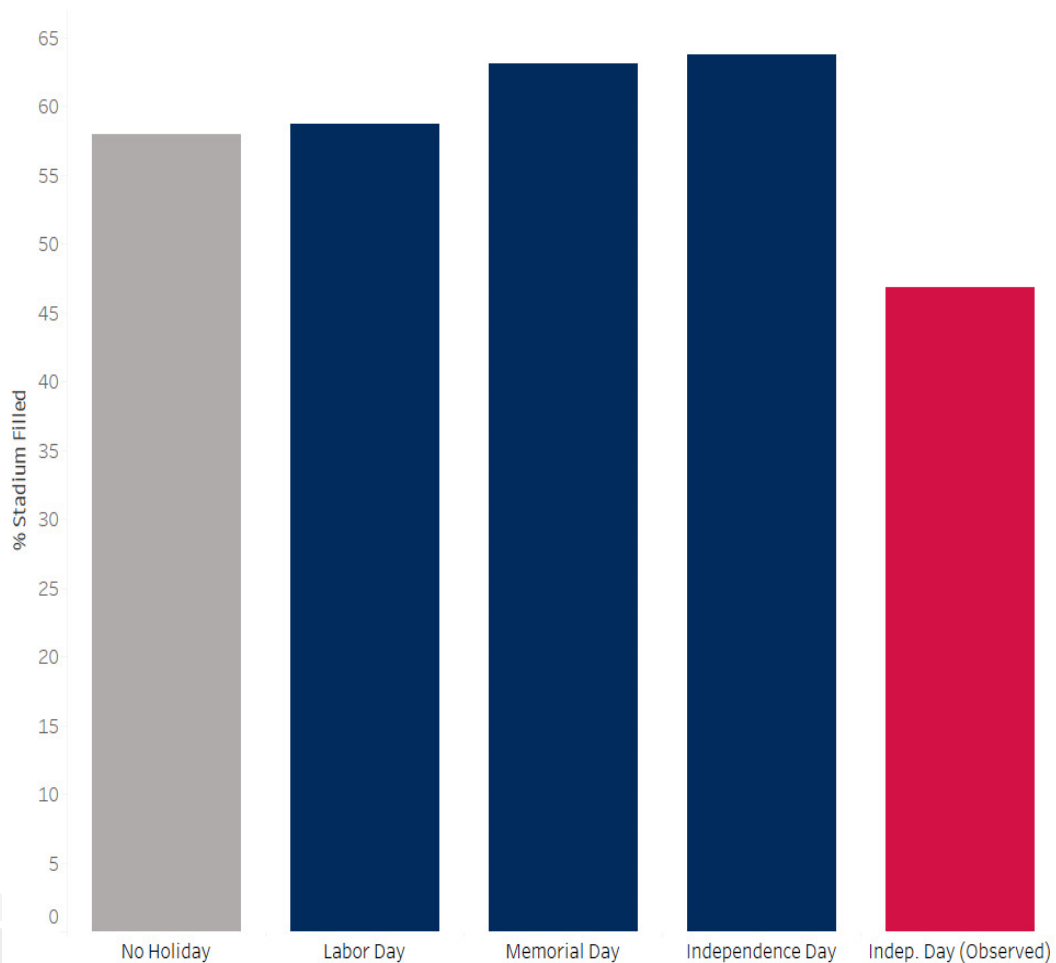


# Our solution can help organizers take informed decisions



# Holiday Games Drive High Attendance

Holiday Game Attendance



1% Labor Day      4% Memorial Day      6% Independence Day

However, **Observed Holidays** have negative effect (11.5%) on attendance – given the long weekend

+ Expected Attendance (Predicted) =

Operational

Optimizing # of concession stands, staff required

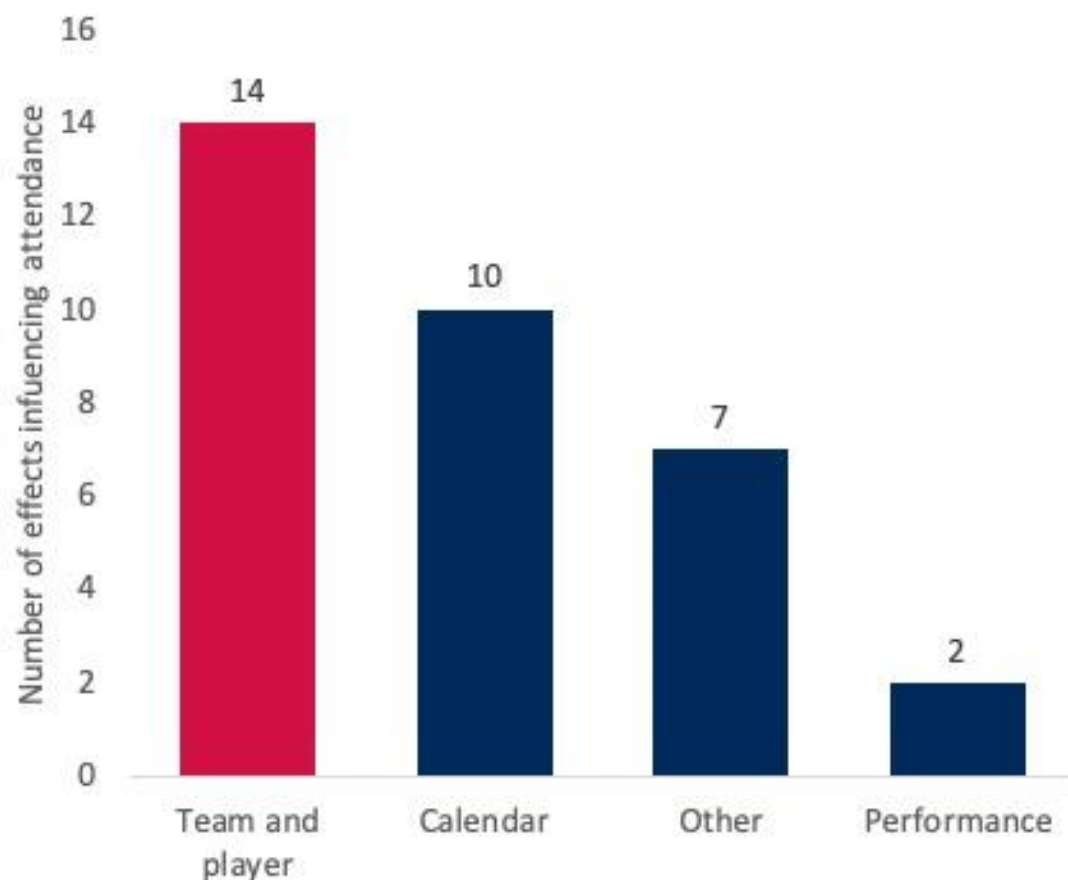
Sales and Revenue

Potential Special Event placement

Financial and Strategic

Marketing Budget optimization

# Teams and players are the largest significant factors driving attendance



## Top visiting teams



25%



18%



17%



Expected Attendance (Predicted)



Operational

Optimizing # of concession stands, staff required

Sales and Revenue

Special events to offset worse teams

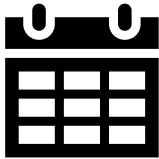
Financial and Strategic

Market teams and players

# Schedule Optimization – What can we control?

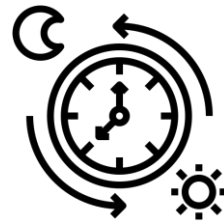
We can control and experiment with three dimensions

1



On which dates should the twins play?

2



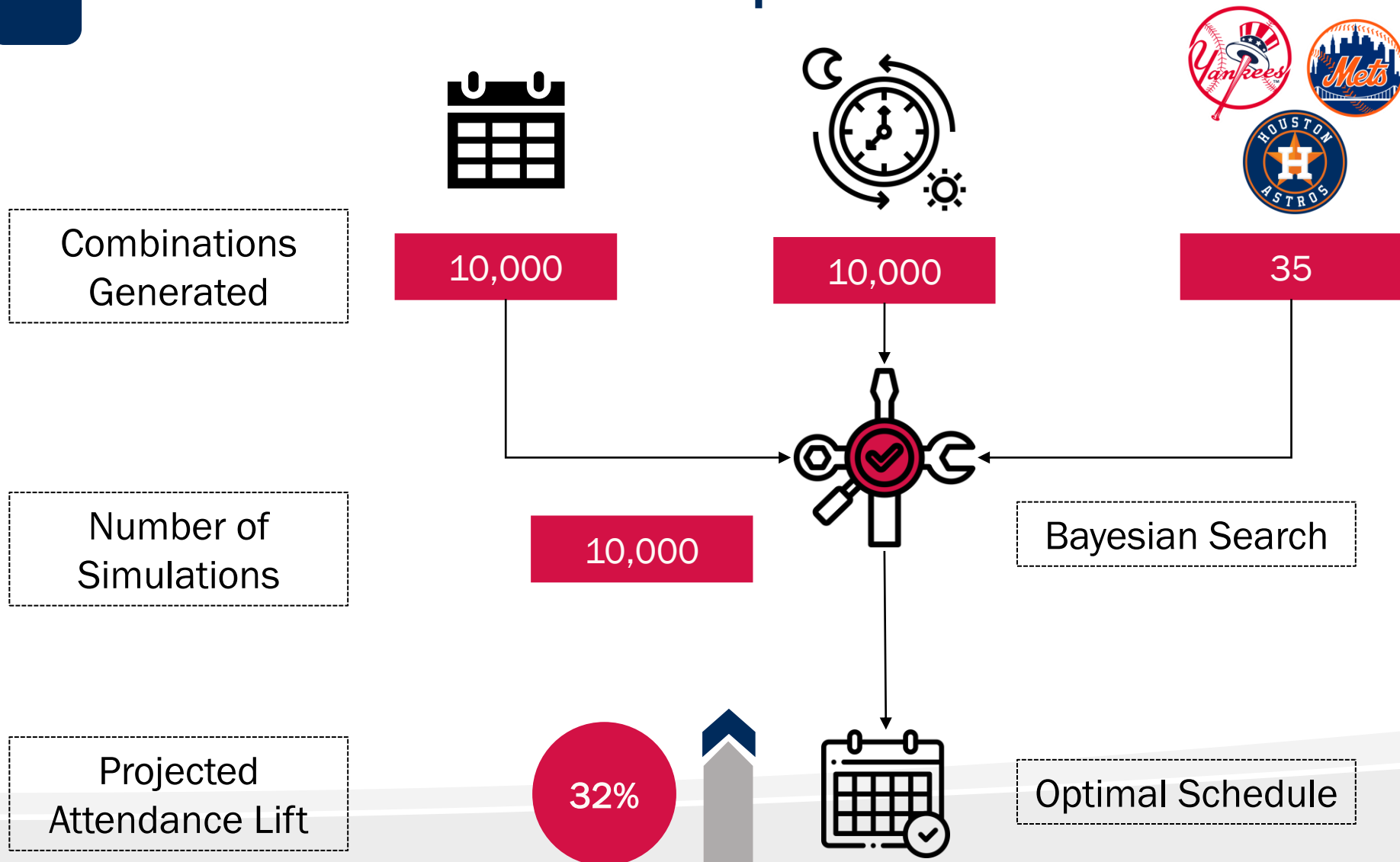
Should the Twins play during Day or Night?

3



When should the Twins play their different opponents?

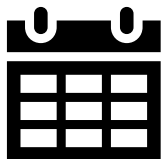
# Method for Schedule Optimization



# Why is the schedule optimal?



1



Prioritizes June to August, with **43%** of games scheduled during those months.

2



**54%** of weekend games are day-games, while only **24%** weekday are day games.

3



Only **40%** of big-ticket games are on weekends compared to **54%** of the other teams



# Why chose our solution?

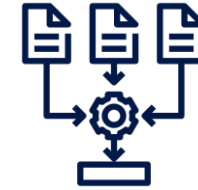
1



Incorporate **Interpretability** and **Predictive** models to help give Direction & Quantify business decisions

- Holidays drive high attendance for MN Twins
- Teams and players are the largest attendance driving factors

2



Predictive Models & Optimizer

- Use state-of-the-art methods
- Provide ~19% better season predictions and 48% better in-season predictions
- Create schedule which could boost attendance by 32%



# Special Thanks!



**Mochen Yang**

*Faculty Advisor*



**Lee Thomas**

*Communications Expert*



# Meet Our Team – Who's on First



**Aakriti Aneja**

*MSBA Class of 2023*  
[aakritianeja.umn@gmail.com](mailto:aakritianeja.umn@gmail.com)



**Ram G S**

*MSBA Class of 2023*  
[egsram@gmail.com](mailto:egsram@gmail.com)



**Justin Mason**

*MSBA Class of 2023*  
[justinkmason@gmail.com](mailto:justinkmason@gmail.com)



**Divyansh Tripathi**

*MSBA Class of 2023*  
[divyanshmani@gmail.com](mailto:divyanshmani@gmail.com)