2025 DATA CHALLENGE

LAUNCH YOUR CAREER

The SMT Data Challenge is an advanced data competition that invites students to analyze real-world player-tracking data. Projects are open-ended and data-driven, rather than strictly quantitative, and they emphasize process, relevance, creativity, and communications. With its uniquely-available player-tracking data, the Data Challenge has become a rich recruiting ground for MLB teams; of the students who have submitted projects, ***more than 20% have been hired by professional teams or sports companies.***

2025 DATA CHALLENGE PROJECT TOPIC

"Good" baseball is all about the success or failure of decision-making. *Did what happened on the field go according to plan?* This year, we want you to use tracking data to quantitatively infer and/or understand player or team "intent". For example:

- Determining who should field a ball and assigning responsibility for misplays
- Evaluating when a fielder should make a play or not
- Analyzing whether a fielder or baserunner moves as intended
- Investigating what an advance scout can deduce about a team's strategy

MLB teams have asked to see:

- Effective use of data visualization—compelling graphics that advance storytelling
- Proficiency with messy data—effective cleaning, analysis, and justifications for inclusion/exclusion

| REGISTRATION *Deadline April 30*

Students may participate individually or in teams of up to 4 students. Individuals who want help building a team can identify themselves when registering.
Data will be made available on or after May 5.

| SUBMISSION CONTENT *Deadline Aug. 1*

- Narrative Paper (PDF max: 2000 words)
- Analysis Code
- Results (CSV)

CLICK OR SCAN TO REGISTER

March 14 - April 30

| JUDGING

Judges from academia, industry, journalism, and sports will evaluate submissions based on creativity, relevance, methodology, storytelling and communication to technical and non-technical audiences.

| FINALISTS *Announced Oct. 6*

• Three finalist teams will be invited to a Showcase Weekend at SMT headquarters in Durham, NC to present to a panel of judges. *(Mid-Nov, date TBD)*

| OTHER DATES

Optional Analysis Demo, Q/A

- *May 8: Introduction*
- *May 22: Analysis*
- *June 5: Data Visualization*
- *July 3: Technical Writing*

Virtual Office Hours Available May 6 - July 31

Must be 18 years or older and enrolled in both Spring and Fall 2025 semesters to enter.

Note: *Unlike previous years, students are not split into undergraduate and graduate divisions, but all compete in the same group.*

For questions, contact us at datachallenge@smt.com