An MLB Bettor's Guide to Profitability

Building a Data Pipeline for MLB Games and Betting Odds

By Alex Weirth

Betting & Odds Domain Knowledge

American Odds Format

American odds are different than traditional odds (8:1 format). Positive odds signify a less likely event (greater return) and negative odds signify a more likely event (lesser return)

Moneyline Bets (ML)

A moneyline bet is a straightforward wager on which team will win a game, with odds reflecting the relative strength of the teams. You simply pick the winner, and if they win, you win your bet.

TEAM	Odds	Wager	To Win
Seahawks	(+250)	$1	$2.50
Rams	(-300)	$3	$1

Data Source & Ingestion

Scraping MLB games data and Bookmaker Odds

Railway

Our Data Sources

The Odds API

1. MLB Preseason Odds (x2)
2. MLB Regular Season Odds(x2)

4 endpoints total for **US1** & **US2** betting markets for **h2h** matchups for **mlb_preseason/mlb_baseball**

The MLB's Official API

1 API endpoint containing a **schedule** parameter that was a live scoreboard of the daily games
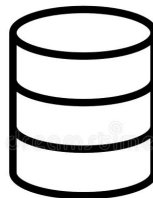
Ingestion

API Request - 1:00 am PST for US markets 1 & 2

API Request - 11:55 pm PST for Game Results (final states)
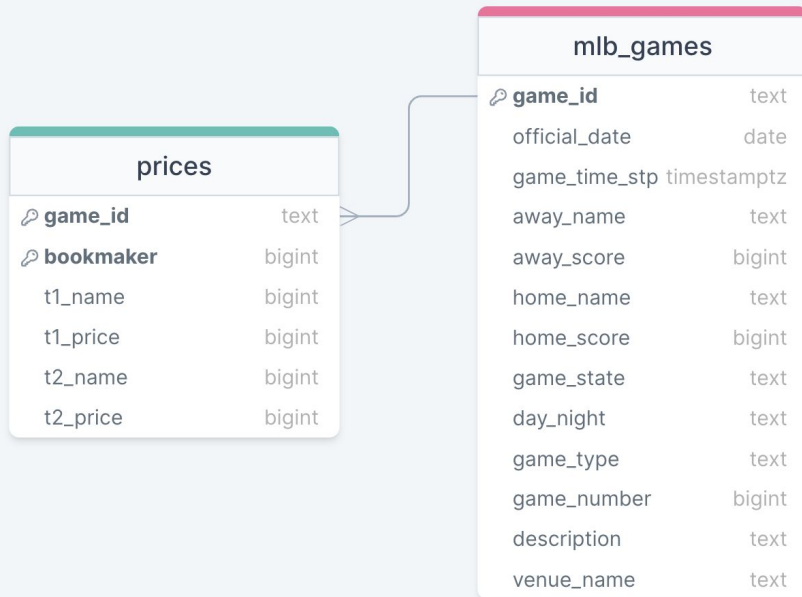
Modeling & Transformations

Cleaning raw JSON data, and creating a database in 3NF

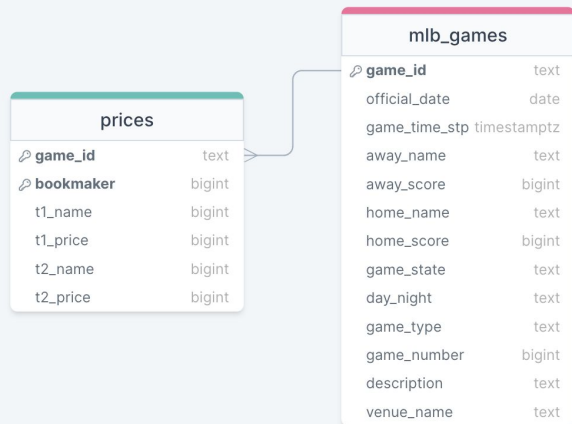# Transformation

Our final database updated & cleaned daily @ 2am:

- Scoreboard of daily MLB game Moneyline prices
- Scoreboard of MLB game results (from the previous day)

prices

🔑 game_id	text
🔑 bookmaker	bigint
t1_name	bigint
t1_price	bigint
t2_name	bigint
t2_price	bigint

mlb_games

🔑 game_id	text
official_date	date
game_time_stp	timestamptz
away_name	text
away_score	bigint
home_name	text
home_score	bigint
game_state	text
day_night	text
game_type	text
game_number	bigint
description	text
venue_name	text

Wrangling Tables

We struggled to get Power BI to wrangle the data in the structure we needed

- It was easier for us to do the necessary data manipulation in SQL.
- This led us to derive a few smaller support tables for our PowerBI report.
- The winnings variable calculates return on a $10 bet on every game.

prices	
game_id	text
bookmaker	bigint
t1_name	bigint
t1_price	bigint
t2_name	bigint
t2_price	bigint

mlb_games	
game_id	text
official_date	date
game_time_stp	timestamptz
away_name	text
away_score	bigint
home_name	text
home_score	bigint
game_state	text
day_night	text
game_type	text
game_number	bigint
description	text
venue_name	text

prices_split	
game_id	text
team_name	text
bookmaker	text
price	bigint
commence_time_pst	timestamptz

mlb_games_split	
game_id	text
team_name	text
official_date	date
is_winner	boolean
game_type	text

profits	
game_id	bigint
team_name	text
is_winner	bigint
game_type	bigint
bookmaker	bigint
price	bigint
winnings	bigint

Our Questions of Interest

In baseball, a team's profitability from consistent betting reflects its performance compared to sportsbooks' predictions (that use extensive ML modeling), showcasing how well it meets or exceeds expectations.

1. **For the 2024 MLB Season, which teams have been most and least profitable?**

2. **What sportsbooks offer the best odds if we wanted to bet on a game for a given day?**

*** Spring training games, while scraped, were not included in the analysis

Findings

What teams have been most and least profitable?

1. The Royals, Orioles, and Mets have been most profitable across all sportsbooks

National League East							
Team	W	L	Pct	GB	Home	Away	L10
Braves	17	6	.739	-	9-3	8-3	9-1
Phillies	15	10	.600	3.0	10-6	5-4	7-3
Mets	13	11	.542	4.5	6-6	7-5	7-3
Nationals	10	13	.435	7.0	4-7	6-6	5-5
Marlins	6	20	.231	12.5	2-11	4-9	3-7

2. The **White Sox**, Twins, and Astros have been the least profitable across all sportsbooks

American League Central							
Team	W	L	Pct	GB	Home	Away	L10
Guardians	17	7	.708	-	7-4	10-3	8-2
Royals	15	10	.600	2.5	11-5	4-5	5-5
Tigers	14	11	.560	3.5	4-7	10-4	5-5
Twins	10	13	.435	6.5	5-6	5-7	4-6
White Sox	3	21	.125	14.0	2-9	1-12	1-9

Findings

What sportsbooks offer the best odds if we wanted to bet on a game for a given day?

1. Odds change daily and for each game, however we now have a useful tool to find what sportsbooks will give us the best return on our money.

2. We did notice that on average, larger sportsbooks make more conservative and less favorable odds (likely due to their number of users), and smaller sportsbooks sometimes have more attractive/favorable odds for bettors (possibly trying to attract customers).

Moneyline Prices by Sportsbook		
Today's prices shown for selected team		
bookmaker	price	commence_time_pst
SuperBook	-190	4/25/2024
MyBookie.ag	-192	4/25/2024
LowVig.ag	-193	4/25/2024
BetOnline.ag	-195	4/25/2024
WynnBET	-195	4/25/2024
Bovada	-198	4/25/2024
BetMGM	-200	4/25/2024
BetUS	-200	4/25/2024
PointsBet (US)	-200	4/25/2024
Wind Creek	-200	4/25/2024
DraftKings	-205	4/25/2024
Caesars	-210	4/25/2024
ESPN BET	-210	4/25/2024
FanDuel	-210	4/25/2024
Hard Rock Bet	-210	4/25/2024
SI Sportsbook	-213	4/25/2024
betPARX	-215	4/25/2024
BetRivers	-215	4/25/2024
Fliff	-215	4/25/2024
Unibet	-215	4/25/2024

Limitations & Future Directions

Hurdles: API missingness, accessing JSON data, wrangling in power BI, connecting Power BI to data, and publishing Power BI.

Future direction: With the Railway trial plan, making constant API calls, and often running of the Docker cleaner container was not possible to the fullest extent: **creating a more frequent CRON schedule would provide more recently updated resources for where someone should place their bets.**

Future direction: API did not contain complete data for all sportsbooks, and some sportsbooks do not provide bets for all games. Find either more complete API's or other endpoints to supplement missingness.

Future direction: look to collect odds on other bets (overs/under, spreads, player props) to create a very useful tool for sports bettors and provide more metrics on profitability of teams.

Future direction: also expand to different sports!

https://the-odds-api.com/ (Reference Website for Odds API's)

API Endpoint for MLB Preseason US Region 1 Odds:

https://api.the-odds-api.com/v4/sports/mlb_preseason/odds?regions=us&oddsFormat=american&apiKey=

API Endpoint for MLB Preseason US Region 1 Odds:

https://api.the-odds-api.com/v4/sports/mlb_preseason/odds?regions=us2&oddsFormat=american&apiKey=

API Endpoint for MLB Preseason US Region 1 Odds:

https://api.the-odds-api.com/v4/sports/baseball_mlb/odds?regions=us&oddsFormat=american&apiKey=

API Endpoint for MLB Preseason US Region 1 Odds:

https://api.the-odds-api.com/v4/sports/baseball_mlb/odds?regions=us2&oddsFormat=american&apiKey=

https://statsapi.mlb.com/ (Reference Website for MLB game scoreboard)

API Endpoint for Daily Scoreboard:

https://statsapi.mlb.com/api/v1/schedule/games/?sportId=1

BEAT THE ODDS - Betting Navigator

Find the best price for your favorite team!

Choose your team from the list below or consult today's game schedule to locate a specific team. The Moneyline (ML) odds for the selected team's game are displayed in the adjacent table:

Team:

Minnesota Twins ▾

Today's Games

Thursday, April 25, 2024	Milwaukee Brewers @ Pittsburgh Pirates
Thursday, April 25, 2024	Boston Red Sox @ Cleveland Guardians
Thursday, April 25, 2024	Chicago White Sox @ Minnesota Twins
Thursday, April 25, 2024	Philadelphia Phillies @ Cincinnati Reds
Thursday, April 25, 2024	Toronto Blue Jays @ Kansas City Royals
Thursday, April 25, 2024	Houston Astros @ Chicago Cubs
Thursday, April 25, 2024	Seattle Mariners @ Texas Rangers
Thursday, April 25, 2024	San Diego Padres @ Colorado Rockies
Thursday, April 25, 2024	Los Angeles Dodgers @ Washington Nationals
Thursday, April 25, 2024	Oakland Athletics @ New York Yankees

Moneyline Prices by Sportsbook

Today's prices shown for selected team

bookmaker	price ↓	commence_time_pst
SuperBook	-190	4/25/2024
MyBookie.ag	-192	4/25/2024
LowVig.ag	-193	4/25/2024
BetOnline.ag	-195	4/25/2024
WynnBET	-195	4/25/2024
Bovada	-198	4/25/2024
BetMGM	-200	4/25/2024
BetUS	-200	4/25/2024
PointsBet (US)	-200	4/25/2024
Wind Creek	-200	4/25/2024
DraftKings	-205	4/25/2024
Caesars	-210	4/25/2024
ESPN BET	-210	4/25/2024
FanDuel	-210	4/25/2024
Hard Rock Bet	-210	4/25/2024
SI Sportsbook	-213	4/25/2024
betPARX	-215	4/25/2024
BetRivers	-215	4/25/2024
Fliff	-215	4/25/2024
Unibet	-215	4/25/2024

Profitability by Team – 2024 MLB Season
Net Earnings by Team on a $10 ML Bet Each Game

Bookmaker

- Select all
- BetMGM
- BetOnline.ag
- betPARX
- BetRivers
- BetUS
- Bovada
- Caesars
- DraftKings
- ESPN BET
- FanDuel
- Fliff
- Hard Rock Bet
- LowVig.ag
- MyBookie.ag
- PointsBet (US)
- SI Sportsbook
- SuperBook
- Tipico
- Unibet
- Wind Creek
- WynnBET

Game Type

- ◉ Regular Season
- ○ Spring Training

Team Name	Net Winnings ($)
Kansas City Royals	$1.6K
New York Mets	$0.9K
Baltimore Orioles	$0.9K
Cleveland Guardians	$0.8K
Milwaukee Brewers	$0.7K
Atlanta Braves	$0.6K
Cincinnati Reds	$0.6K
Chicago Cubs	$0.5K
Philadelphia Phillies	$0.4K
New York Yankees	$0.4K
Tampa Bay Rays	$0.3K
San Francisco Giants	$0.2K
Detroit Tigers	$0.2K
Boston Red Sox	($0.0K)
Oakland Athletics	($0.0K)
Washington Nationals	($0.1K)
Los Angeles Dodgers	($0.2K)
Toronto Blue Jays	($0.2K)
San Diego Padres	($0.2K)
Seattle Mariners	($0.4K)
St. Louis Cardinals	($0.5K)
Pittsburgh Pirates	($0.5K)
Arizona Diamondbacks	($0.5K)
Los Angeles Angels	($0.7K)
Texas Rangers	($0.7K)
Colorado Rockies	($0.9K)
Miami Marlins	($1.0K)
Minnesota Twins	($1.3K)
Houston Astros	($1.3K)
Chicago White Sox	($2.4K)