



MLB and Betting

Nick Brooks and Travis McKenzie

Data Pipeline



Our Data

- **MLB Score data**
 - Accessed from Tank01 MLB Live In-Game Real Time Statistics API
 - Variables of interest
 - Teams (home and away)
 - Final score
 - Game date
- **Betting Spreads data**
 - Accessed from The Odds API
 - Variables of interest
 - Teams (home and away)
 - Spreads
 - Prices
 - Bookmakers
- **Data Collection**
 - Three weeks from April 2 - April 23

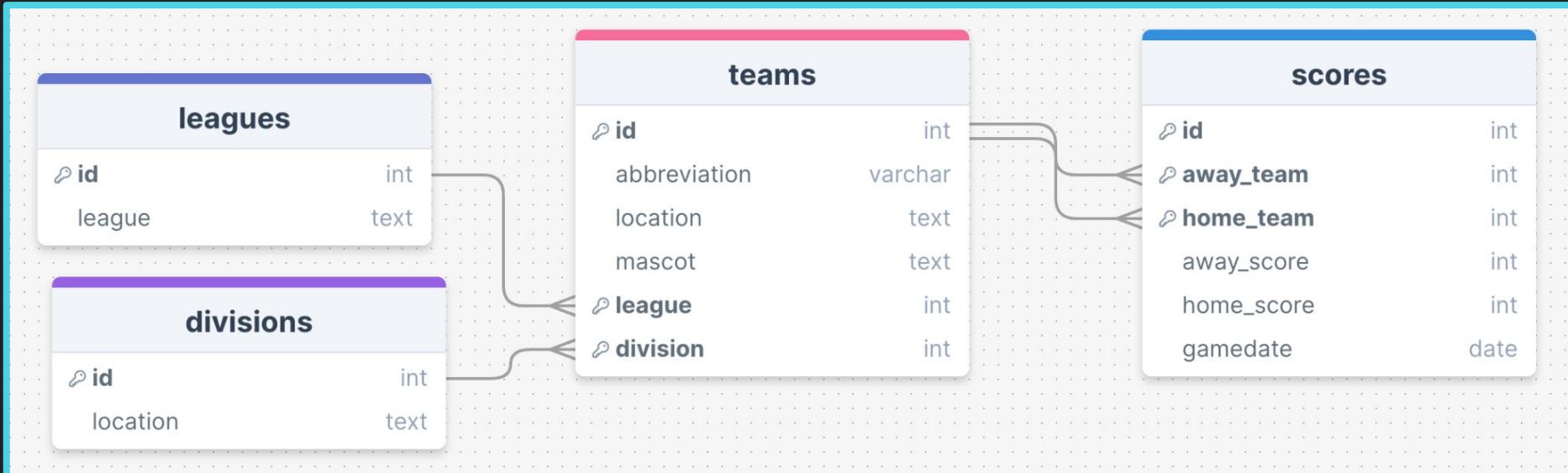
Sample Response - MLB Score Data

Key Variables

- Teams (home and away)
- Runs
- Date

```
gameStatus: "Completed"
lastUpdated: 2023073119235844
currentBatter: ""
bases: {} 0 keys
gameDate: "20230410"
awayResult: "W"
currentCount: ""
homeResult: "L"
away: "CHW"
▼ lineScore: {} 2 keys
  ▼ away: {} 5 keys
    H: "8"
    R: "4"
    team: "CHW"
    ▶ scoresByInning: {} 9 keys
      E: "3"
  ▼ home: {} 5 keys
    H: "4"
    R: "3"
    team: "MIN"
    ▶ scoresByInning: {} 9 keys
      E: "0"
    currentOuts: ""
    currentPitcher: ""
    currentInning: "Final"
    gameID: "20230410_CHW@MIN"
    home: "MIN"
```

MLB Score Data - Organized

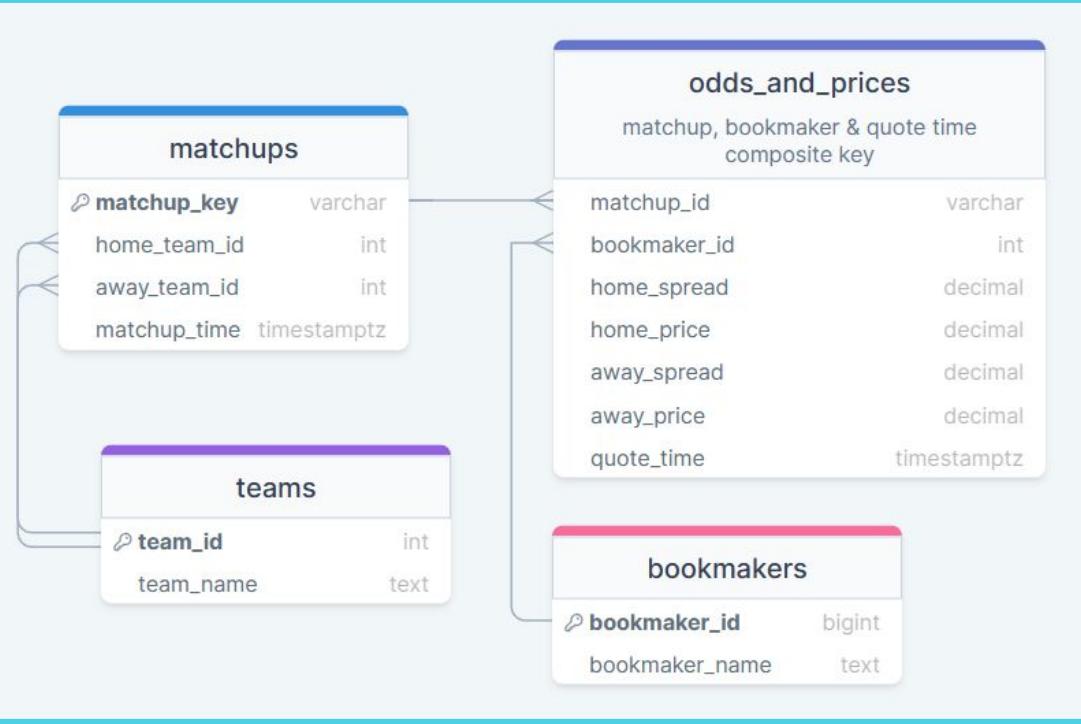


Raw JSON - Spread Betting Data

```
▼ object {7}
  id : 01042742cd7021ee2680705f0cad8c79
  away_team : New York Yankees
  home_team : Arizona Diamondbacks
  sport_key : baseball_mlb
  ▼ bookmakers [7]
    ▼ 0 {4}
      key : draftkings
      title : DraftKings
      ► markets [1]
      last_update : 2024-04-02T04:35:45Z
    ► 1 {4}
    ► 2 {4}
    ► 3 {4}
    ► 4 {4}
    ► 5 {4}
    ► 6 {4}
  sport_title : MLB
  commence_time : 2024-04-02T01:40:27Z
```

```
{"id":"01042742cd7021ee2680705f0cad8c79","away_team":"New York Yankees","home_team":"Arizona Diamondbacks","sport_key":"baseball_mlb","bookmakers":[{"key":"draftkings","title":"DraftKings","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":4.9},{"name":"New York Yankees","point":-2.5,"price":1.17}],"last_update":"2024-04-02T04:35:44Z"}],"last_update":"2024-04-02T04:35:45Z"}, {"key":"fanduel","title":"FanDuel","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":3.2}]}],"name":"New York Yankees","point":-2.5,"price":1.34}],"last_update":"2024-04-02T04:34:45Z"}],"last_update":"2024-04-02T04:34:45Z"}, {"key":"mybookieag","title":"MyBookie.ag","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":3}]}],"name":"New York Yankees","point":-2.5,"price":1.31}],"last_update":"2024-04-02T04:41:17Z"}],"last_update":"2024-04-02T04:41:17Z"}, {"key":"williamhill_us","title":"Caesars","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":3.4}]}],"name":"New York Yankees","point":-2.5,"price":1.31}],"last_update":"2024-04-02T04:34:44Z"}],"last_update":"2024-04-02T04:34:44Z"}, {"key":"unibet_us","title":"Unibet","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":6}]}],"name":"New York Yankees","point":-2.5,"price":1.1}],"last_update":"2024-04-02T04:37:04Z"}],"last_update":"2024-04-02T04:37:04Z"}, {"key":"betmgm","title":"BetMGM","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":4.6}]}],"name":"New York Yankees","point":-2.5,"price":1.18}],"last_update":"2024-04-02T04:39:12Z"}],"last_update":"2024-04-02T04:39:12Z"}, {"key":"wynnbet","title":"WynnBET","markets":[{"key":"spreads","outcomes":[{"name":"Arizona Diamondbacks","point":2.5,"price":3.39}]}],"name":"New York Yankees","point":-2.5,"price":1.29}],"last_update":"2024-04-02T04:34:45Z"}],"last_update":"2024-04-02T04:34:45Z"}], "sport_title":"MLB","commence_time":"2024-04-02T01:40:27Z"}]
```

MLB Spread Data - Organized



```
CREATE TABLE matchups (
    matchup_id VARCHAR REFERENCES matchups (matchup_key),
    bookmaker_id INT REFERENCES bookmakers (bookmaker_id),
    home_team_id INT REFERENCES teams (team_id),
    away_team_id INT REFERENCES teams (team_id),
    matchup_time TIMESTAMPTZ,
    CONSTRAINT matchup_key UNIQUE (matchup_id, bookmaker_id, home_team_id, away_team_id, matchup_time)
);

CREATE TABLE teams (
    team_id BIGINT PRIMARY KEY GENERATED ALWAYS AS IDENTITY,
    team_name TEXT UNIQUE
);

CREATE TABLE bookmakers (
    bookmaker_id SMALLINT PRIMARY KEY GENERATED ALWAYS AS IDENTITY,
    bookmaker_name TEXT UNIQUE
);

CREATE TABLE odds_and_prices (
    matchup_id VARCHAR REFERENCES matchups (matchup_key),
    bookmaker_id INT REFERENCES bookmakers (bookmaker_id),
    home_spread DECIMAL,
    home_price DECIMAL,
    away_spread DECIMAL,
    away_price DECIMAL,
    quote_time TIMESTAMPTZ,
    CONSTRAINT odds_key PRIMARY KEY (matchup_id, bookmaker_id, quote_time)
);
```



Data Service



/teams

- Returns all MLB teams, with their associated data
- Results are ordered in alphabetical order

<https://mlbbetapi-production.up.railway.app/teams>

```
[  
  {  
    "abbreviation": "ARI",  
    "location": "Arizona",  
    "mascot": "Diamondbacks",  
    "league": "National",  
    "division": "West"  
  },  
  {  
    "abbreviation": "ATL",  
    "location": "Atlanta",  
    "mascot": "Braves",  
    "league": "National",  
    "division": "East"  
  },  
  {  
    "abbreviation": "BAL",  
    "location": "Baltimore",  
    "mascot": "Orioles",  
    "league": "American",  
    "division": "East"  
  },  
  {  
    "abbreviation": "BOS",  
    "location": "Boston",  
    "mascot": "Red Sox",  
    "league": "American",  
    "division": "East"  
  },  
  {  
    "abbreviation": "CHC",  
    "location": "Chicago",  
    "mascot": "Cubs",  
    "league": "National",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "COL",  
    "location": "Colorado",  
    "mascot": "Rockies",  
    "league": "National",  
    "division": "West"  
  },  
  {  
    "abbreviation": "CWS",  
    "location": "Chicago",  
    "mascot": "White Sox",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "DET",  
    "location": "Detroit",  
    "mascot": "Tigers",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "HOU",  
    "location": "Houston",  
    "mascot": "Astros",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "KC",  
    "location": "Kansas City",  
    "mascot": "Royals",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "LAA",  
    "location": "Los Angeles",  
    "mascot": "Angels",  
    "league": "American",  
    "division": "West"  
  },  
  {  
    "abbreviation": "MIA",  
    "location": "Miami",  
    "mascot": "Marlins",  
    "league": "National",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "MIN",  
    "location": "Minnesota",  
    "mascot": "Twins",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "NYY",  
    "location": "New York",  
    "mascot": "Yankees",  
    "league": "American",  
    "division": "East"  
  },  
  {  
    "abbreviation": "NLO",  
    "location": "New York",  
    "mascot": "Mets",  
    "league": "National",  
    "division": "East"  
  },  
  {  
    "abbreviation": "OAK",  
    "location": "Oakland",  
    "mascot": "Athletics",  
    "league": "American",  
    "division": "West"  
  },  
  {  
    "abbreviation": "PHL",  
    "location": "Philadelphia",  
    "mascot": "Phillies",  
    "league": "National",  
    "division": "East"  
  },  
  {  
    "abbreviation": "PIT",  
    "location": "Pittsburgh",  
    "mascot": "Pirates",  
    "league": "National",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "SD",  
    "location": "San Diego",  
    "mascot": "Padres",  
    "league": "National",  
    "division": "West"  
  },  
  {  
    "abbreviation": "SEA",  
    "location": "Seattle",  
    "mascot": "Mariners",  
    "league": "American",  
    "division": "West"  
  },  
  {  
    "abbreviation": "STL",  
    "location": "St. Louis",  
    "mascot": "Cardinals",  
    "league": "National",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "TB",  
    "location": "Tampa Bay",  
    "mascot": "Rays",  
    "league": "American",  
    "division": "Central"  
  },  
  {  
    "abbreviation": "WAS",  
    "location": "Washington",  
    "mascot": "Nationals",  
    "league": "National",  
    "division": "East"  
  }]
```

/scores/{page}

- Returns 50 results per page
- Results are ordered by the game date

<https://mlbbetapi-production.up.railway.app/scores/0>

```
[{"away_team": "Chicago Cubs", "home_team": "Texas Rangers", "away_score": 9, "home_score": 5, "gamedate": "2024-03-31"}, {"away_team": "Boston Red Sox", "home_team": "Seattle Mariners", "away_score": 5, "home_score": 1, "gamedate": "2024-03-31"}, {"away_team": "St. Louis Cardinals", "home_team": "Los Angeles Dodgers", "away_score": 4, "home_score": 5, "gamedate": "2024-03-31"}, {"away_team": "Colorado Rockies", "home_team": "Arizona Diamondbacks", "away_score": 1, "home_score": 5, "gamedate": "2024-03-31"}, {"away_team": "San Francisco Giants", "home_team": "San Diego Padres", "away_score": 4, "home_score": 13, "gamedate": "2024-03-31"}]
```

/teamscores/{team}

- Returns all score results for a specified team
- Results are ordered by the game date

<https://mlbbetapi-production.up.railway.app/teamscores/sea>

```
[  
  {  
    "away_team": "Boston Red Sox",  
    "home_team": "Seattle Mariners",  
    "away_score": 5,  
    "home_score": 1,  
    "gamedate": "2024-03-31"  
  },  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "away_score": 4,  
    "home_score": 5,  
    "gamedate": "2024-04-01"  
  },  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "away_score": 5,  
    "home_score": 2,  
    "gamedate": "2024-04-02"  
  },  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "away_score": 8,  
    "home_score": 0,  
    "gamedate": "2024-04-03"  
  },  
  {  
    "away_team": "Seattle Mariners",  
    "home_team": "Milwaukee Brewers",  
    "away_score": 5,  
    "home_score": 6,  
    "gamedate": "2024-04-05"  
  },  
]
```

/lines/{page}

- Returns betting lines of MLB games, with the teams, game date, spread, prices, and bookmakers
- Results are ordered in chronological order

<https://mlbbetapi-production.up.railway.app/lines/0>

```
[{"away_team": "New York Yankees", "home_team": "Arizona Diamondbacks", "game_start_time": "2024-04-02T01:40:27", "bookmaker": "FanDuel", "home_spread": 2.5, "home_price": 3.2, "away_spread": -2.5, "away_price": 1.34}, {"away_team": "New York Yankees", "home_team": "Arizona Diamondbacks", "game_start_time": "2024-04-02T01:40:27", "bookmaker": "MyBookie.ag", "home_spread": 2.5, "home_price": 3, "away_spread": -2.5, "away_price": 1.31}, {"away_team": "New York Yankees", "home_team": "Arizona Diamondbacks", "game_start_time": "2024-04-02T01:40:27", "bookmaker": "DraftKings", "home_spread": 2.5, "home_price": 4.9, "away_spread": -2.5, "away_price": 1.17}, {"away_team": "New York Yankees", "home_team": "Arizona Diamondbacks", "game_start_time": "2024-04-02T01:40:27", "bookmaker": "Unibet", "home_spread": 2.5, "home_price": 6, "away_spread": -2.5, "away_price": 1.1} ]
```

/teamlines/{team}

- Returns betting lines of MLB games for a specified team, with the teams, game date, spread, prices, and bookmakers
- Results are ordered in chronological order

<https://mlbbetapi-production.up.railway.app/teamlines/sea>

```
[  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "game_start_time": "2024-04-03T01:40:00",  
    "bookmaker": "Caesars",  
    "home_spread": 1.5,  
    "home_price": 1.45,  
    "away_spread": -1.5,  
    "away_price": 2.8  
  },  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "game_start_time": "2024-04-03T01:40:00",  
    "bookmaker": "BetMGM",  
    "home_spread": -1.5,  
    "home_price": 2.65,  
    "away_spread": 1.5,  
    "away_price": 1.5  
  },  
  {  
    "away_team": "Cleveland Guardians",  
    "home_team": "Seattle Mariners",  
    "game_start_time": "2024-04-03T01:40:00",  
    "bookmaker": "BetOnline.ag",  
    "home_spread": 1.5,  
    "home_price": 1.47,  
    "away_spread": -1.5,  
    "away_price": 2.82  
  },  
]
```

Analysis



Betting Definitions

gamedate	mkt	home_team	home_score	away_team	away_score	home_spread	away_spread
<date>	<chr>	<chr>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
1 2024-04-02	BetMGM	Arizona Diamondbacks	7	New York Yankees	0	2.5	-2.5

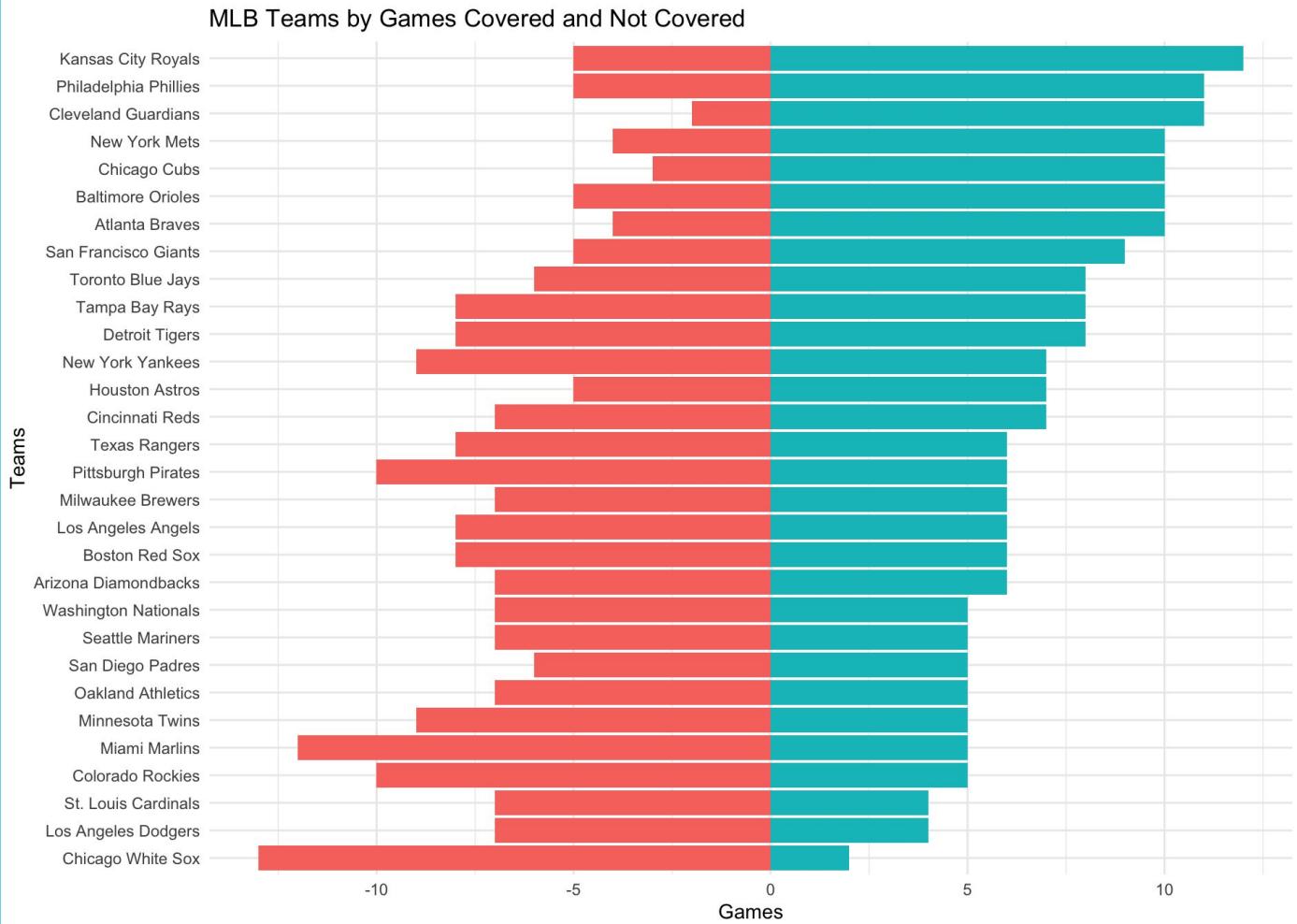
- **Spread:** number of runs that bookmakers predict is the margin of victory/defeat
 - Negative spread means that team is favored
 - Yankees' spread is -2.5, so they are favored to win by 2.5 runs
- **Cover:** covering the spread is winning by more than the predicted margin of victory as a favorite OR losing by less than the predicted margin of defeat as an underdog
 - For the Yankees to cover, they would have to win by more than 2.5 runs
 - For the Diamondbacks to cover, they would have to lose by less than 2.5 runs or win
- In our example, the **Diamondbacks covered** and the **Yankees did not cover**

Questions of Interest

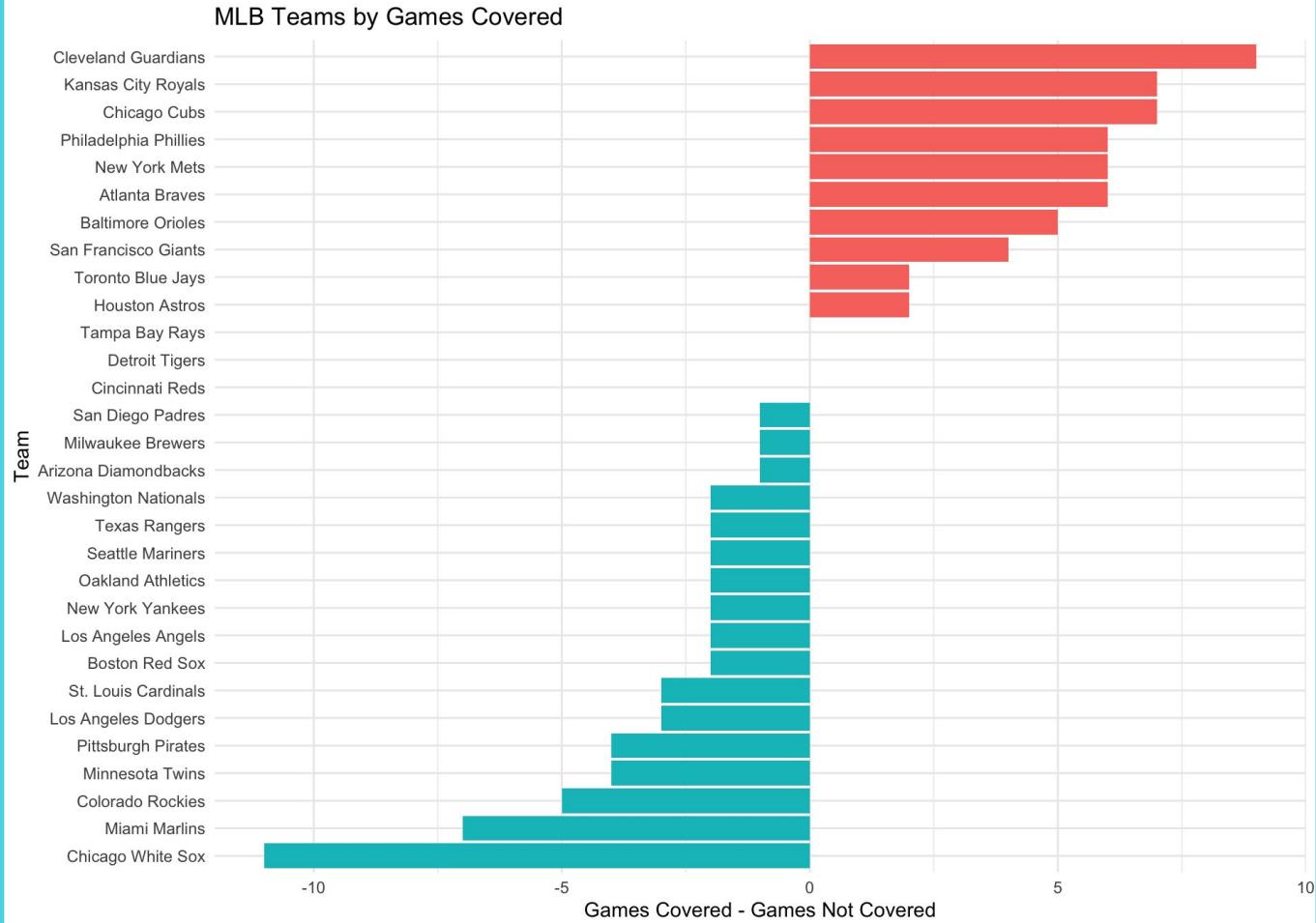
We are **not** sports bettors, so we wanted to ask questions that, should we decide to bet, would help us make better (bettor) decisions

- How are the teams performing against the spread for our window of observation?
- Which teams cover the spread most often?
- Factors for covering
 - Do favorites or underdogs cover more often?
 - Does the away team cover more often?
- What types of variation exist (if any) across bookmakers?
- How are the Mariner's (our mutual favorite baseball team) performing when considering the spread?

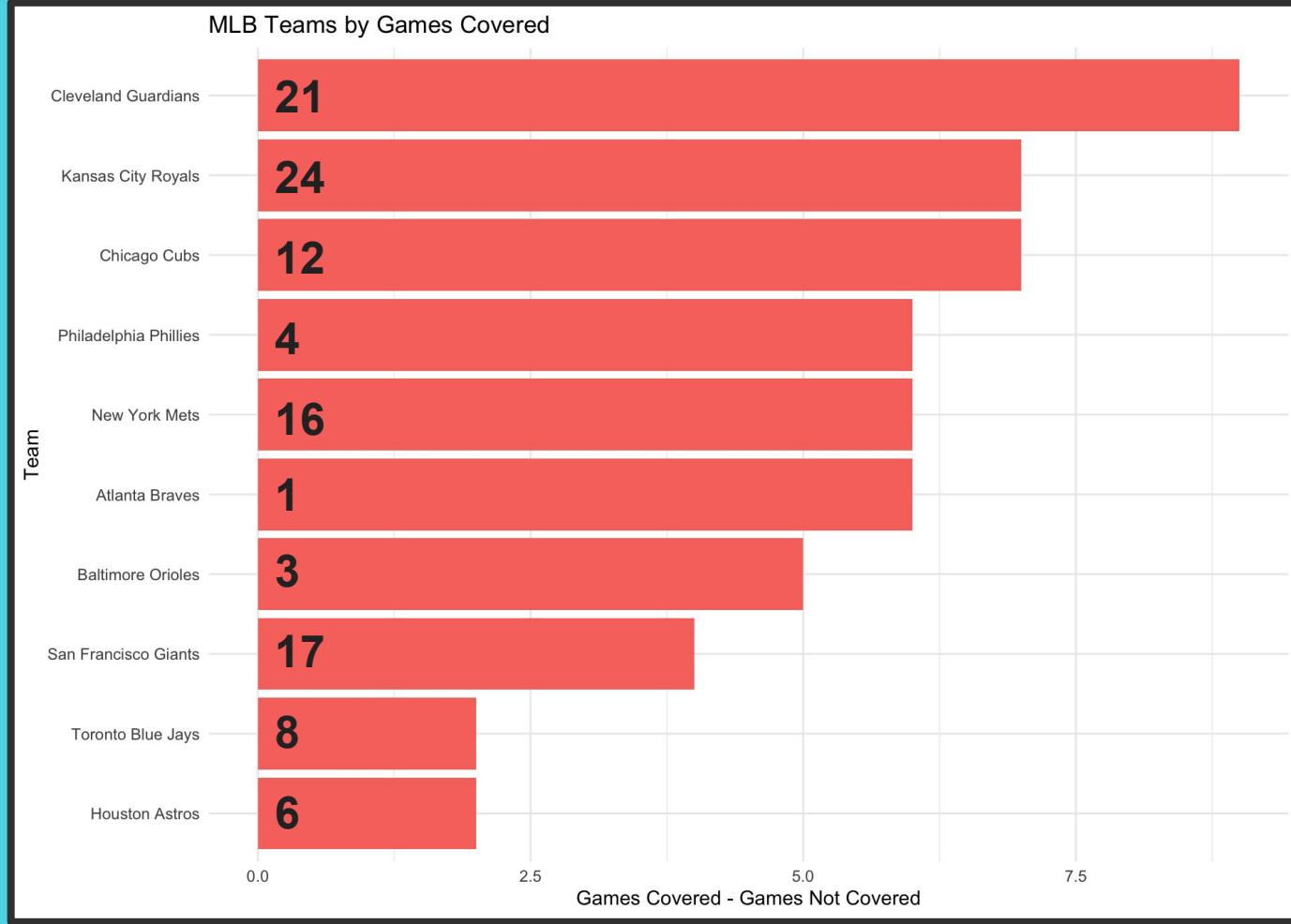
How are teams performing against the spread?



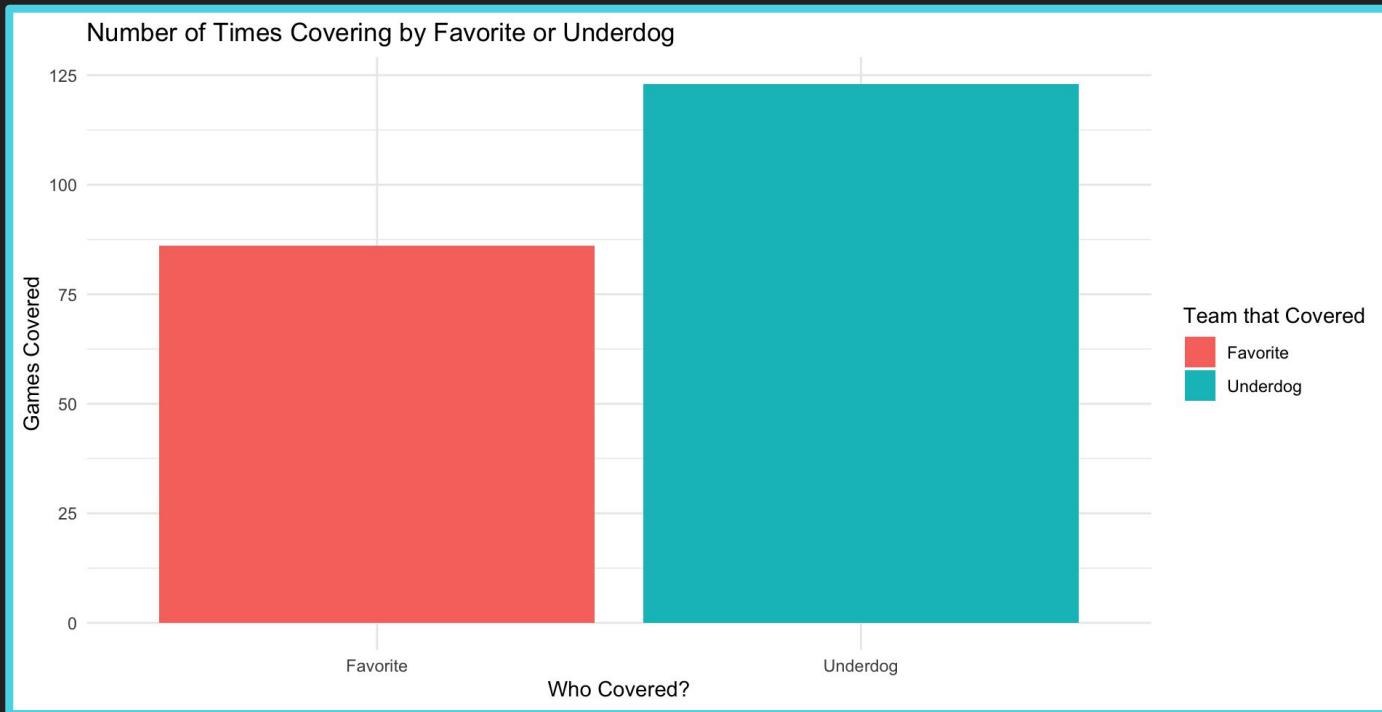
Which teams cover the spread the most?



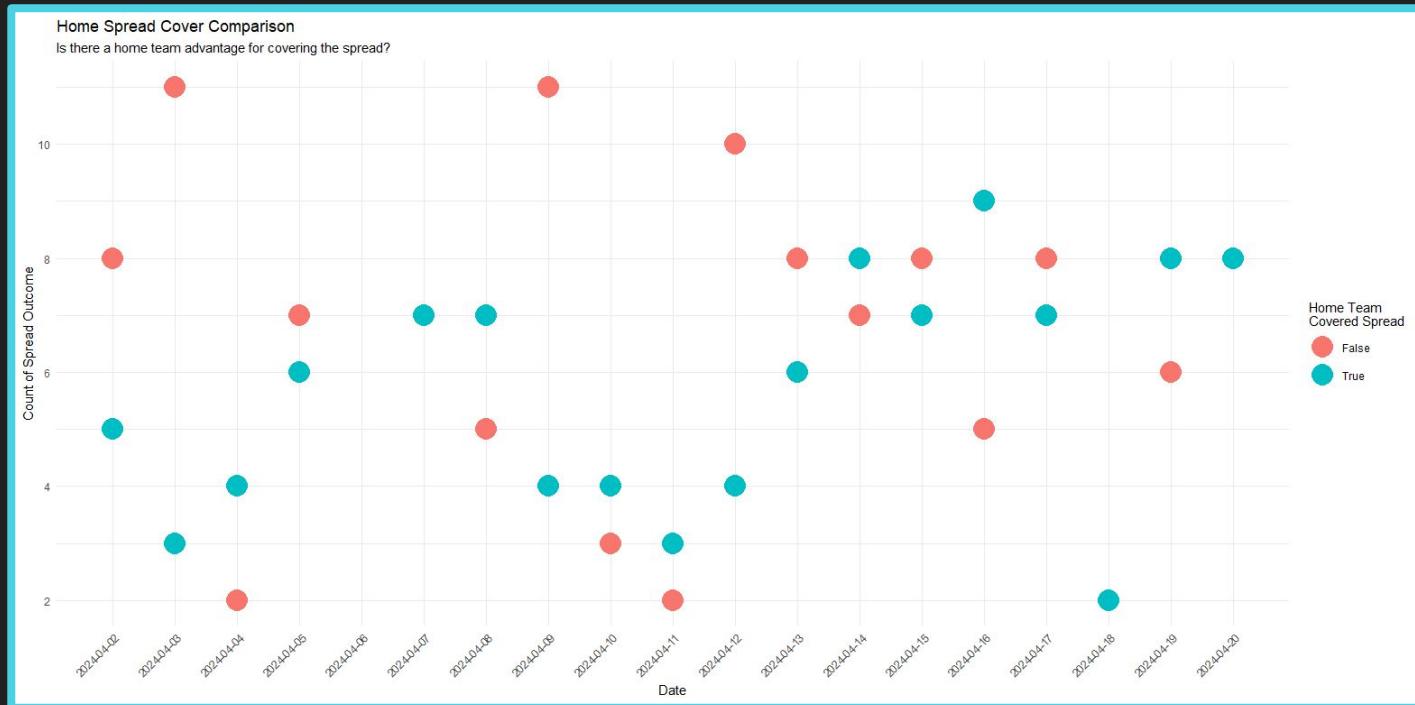
Comparing
spread
covering
teams with
their
preseason
ranking on
MLB.com



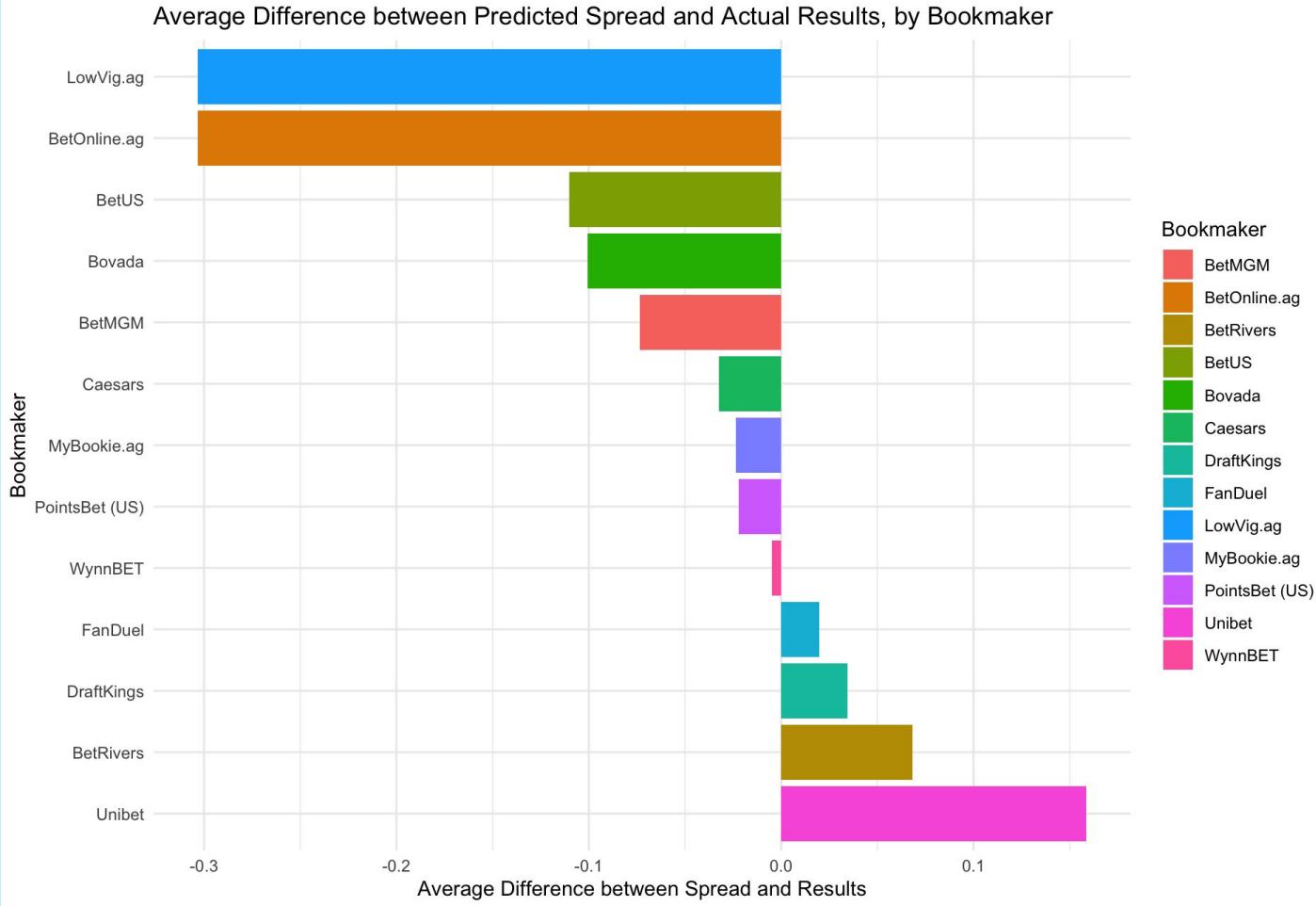
Do favorites
or
underdogs
cover more
often?



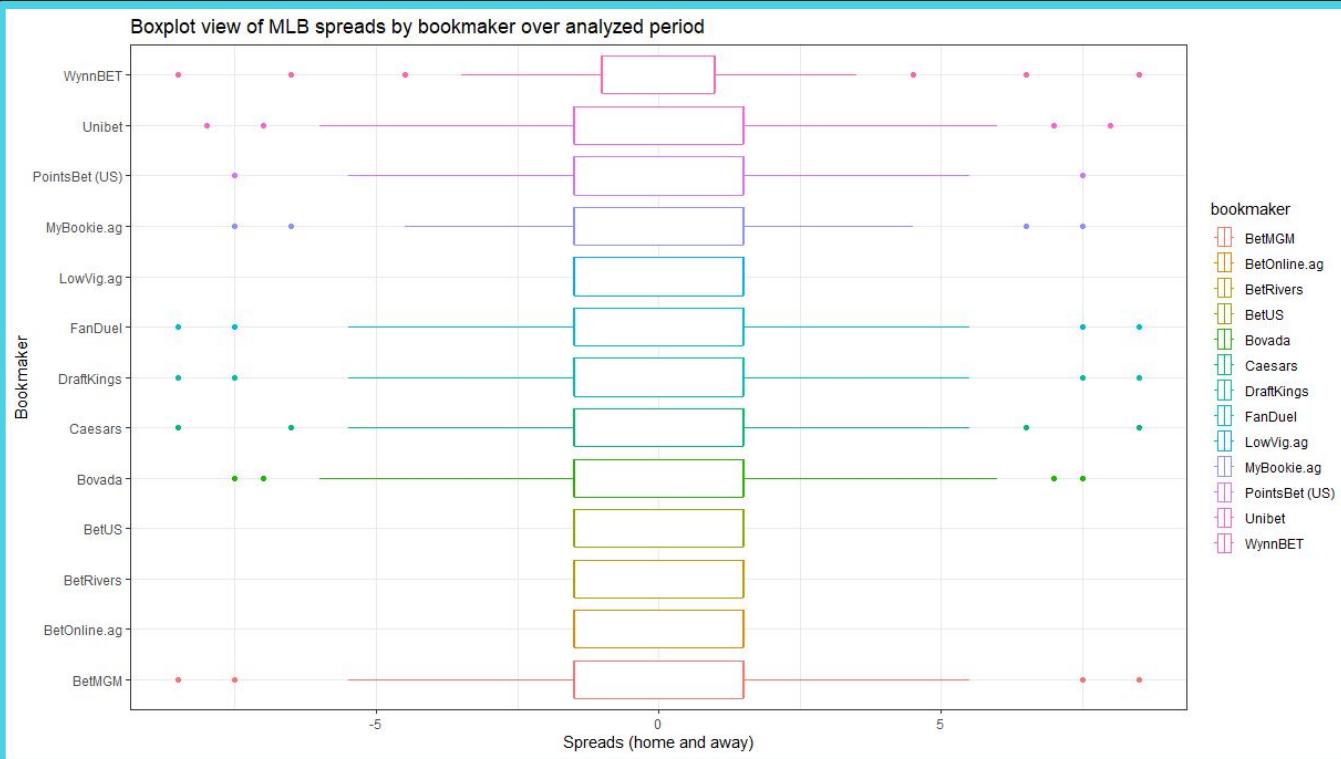
Does the home team have a spread cover advantage?



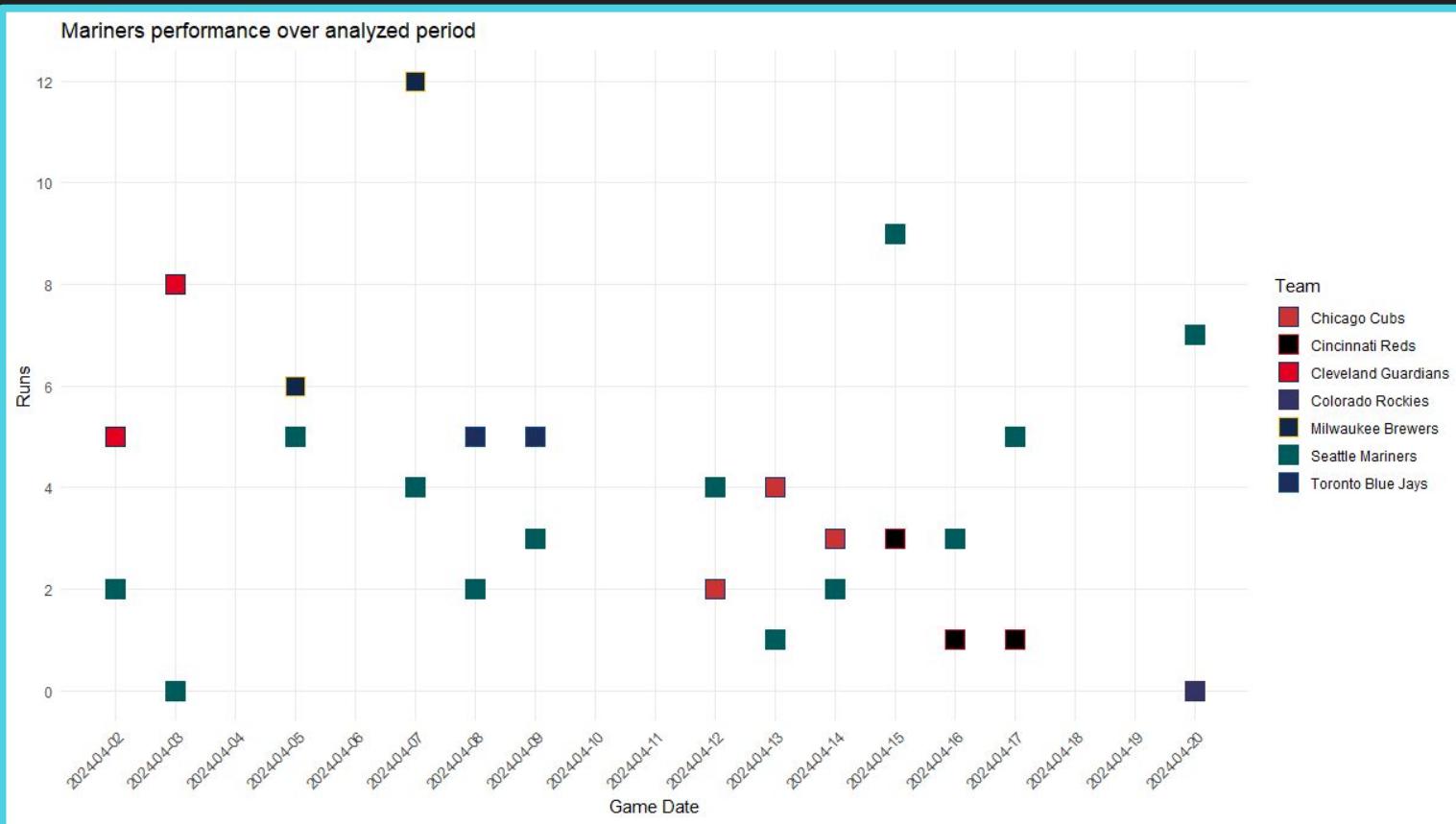
How accurate are the spreads?



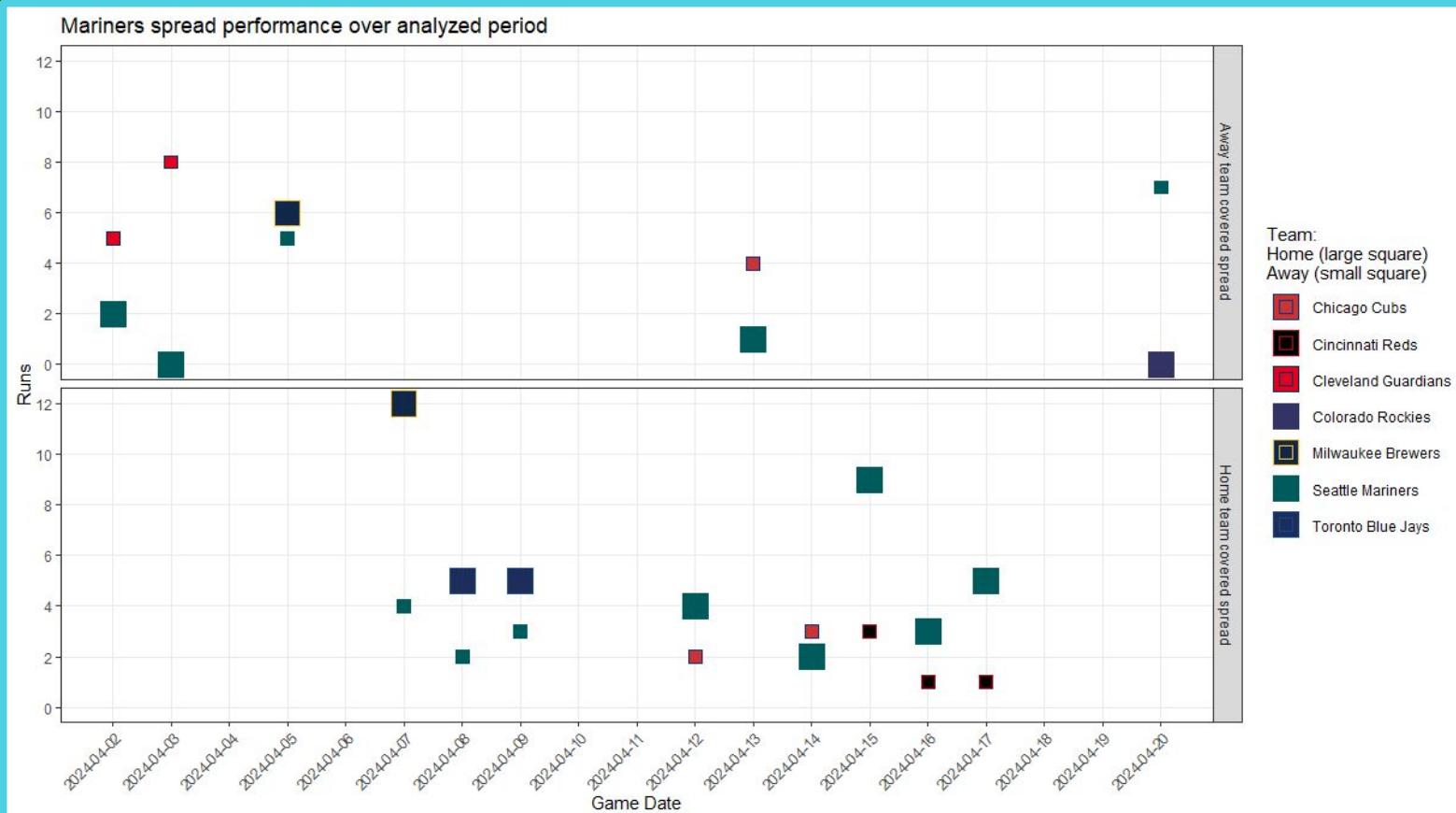
Spread comparison across bookmakers



How are the Mariners doing?



Mariners home or away while covering spread or not



Conclusion



Conclusions

- Teams that cover the spread most are mainly teams that are outperforming preseason expectations
- Underdogs are covering at a significantly higher rate than favorites
 - Likely due to being so early in the season, bookmakers are still trying to figure out how teams stack up
- Not seeing a significant advantage to home or away teams based on our limited window of games.
- WynnBet has the most accurate spreads on average
- Limited variation across bookmakers. Spreads are not particularly dynamic.
- Mariners were a bad bet to cover the spread for the first half of our data and a great bet to cover the spread for the second half.

Challenges

- **Doubleheaders**
 - MLB API did not distinguish between the games of a doubleheader, giving them both the same date and teams, but different scores
- **How many lines?**
 - Betting API would have a varying number of lines different bookmakers, from 5 to 13 lines per game
- **Data accuracy**
 - Some spreads and scores were incorrect when obtained from the API
- **Spread data normalization**
 - More complicated than expected
- **Learning about sports betting**
 - Neither of us are sports bettors, so we had to do research to learn about the prices and spreads

Looking Forward

- Expand to other sports
- Increase features and service options
 - Dynamic dashboard showing comparisons
- Collecting a larger sample size of games
- Continue to explore data and try to find more connections
- More resolution on the meaning of the price variable in the context of the API
- Exploring other variables and how they impact teams covering spreads
 - Game attendance
 - Weather
 - Traffic....

Questions?

