

CurveBall Consultants

**Project_0507_01
December 4, 2023**

Last Edited: Dec 3, 2023

***Pravah Malunjkar, Gautam Bhatia,
Rajat Marathe, Tanmay Sakharkar***

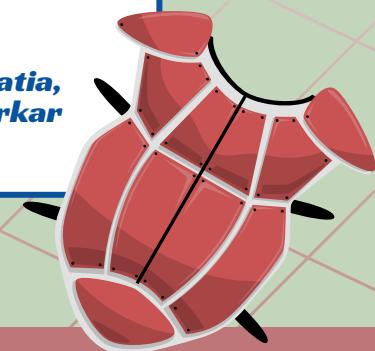


Table of Contents

***1. About
Our Company***

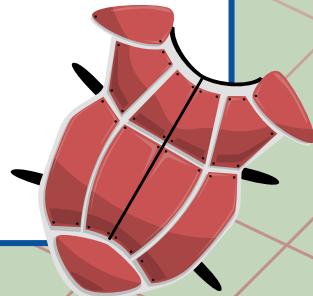
***2. Conceptual
Database Design***

***3. Logical
Database Design***

***4. Physical
Database Design***

5. Use Cases

About Our Company



Our Team



*Pravah
Malunjkar*



*Gautam
Bhatia*



*Rajat
Marathe*



*Tanmay
Sakharkar*

Mission Statement

*“To elevate the performance of
Maryland Softball using Analytics.”*

Our Objectives



Procure

Procure Maryland Softball game data from 1999 to 2022.



Build

Build a database for analyzing softball statistics.



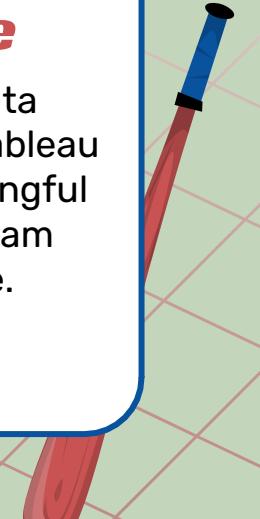
Store

Setting up SQL queries to analyze dataset



Analyze

Leveraging data visualization in Tableau to provide meaningful insights into team performance.



Business Description

Management

Specialization in comprehensive sports management and event organization services

Coordination

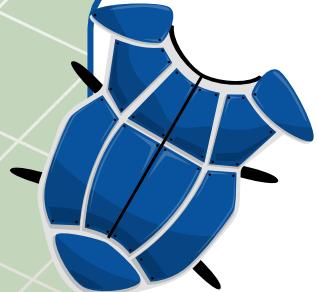
Expertise in streamlining the coordination of teams, matches, and locations

Communication

Emphasis on clear and efficient communication throughout the process

Delivery

Focus on delivering results in sports management and event organization



Our Users and Sources



Users

Our primary users include coaching staff, players, and management personnel associated with Maryland Softball. By leveraging analytics, we aim to empower these stakeholders with insights that go beyond traditional methods, fostering informed decision-making and strategic planning.



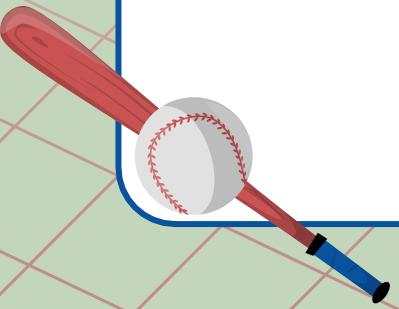
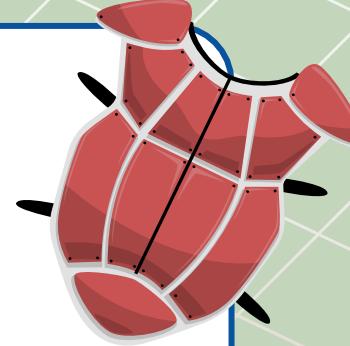
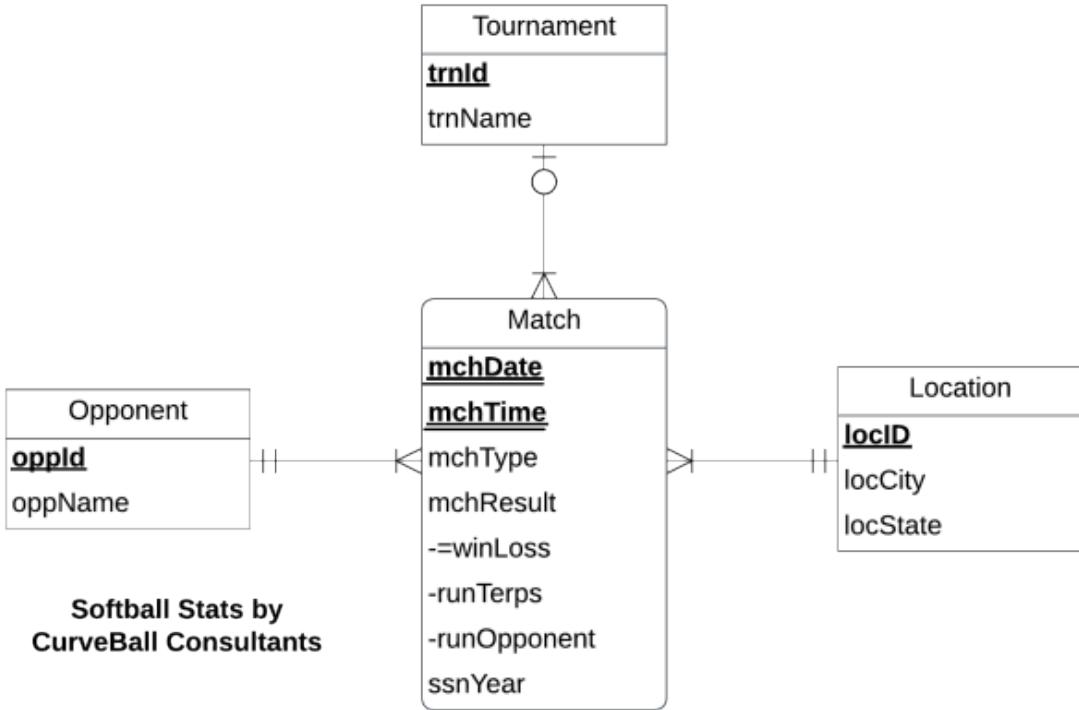
Data Source

To achieve our objectives, we've undertaken the ambitious task of procuring Maryland Softball game data spanning from 1999 to 2022. This comprehensive dataset serves as the backbone of our analytical endeavors, capturing the nuances of each match, player performance, and team dynamics. <https://umterps.com/sports/softball/schedule/>

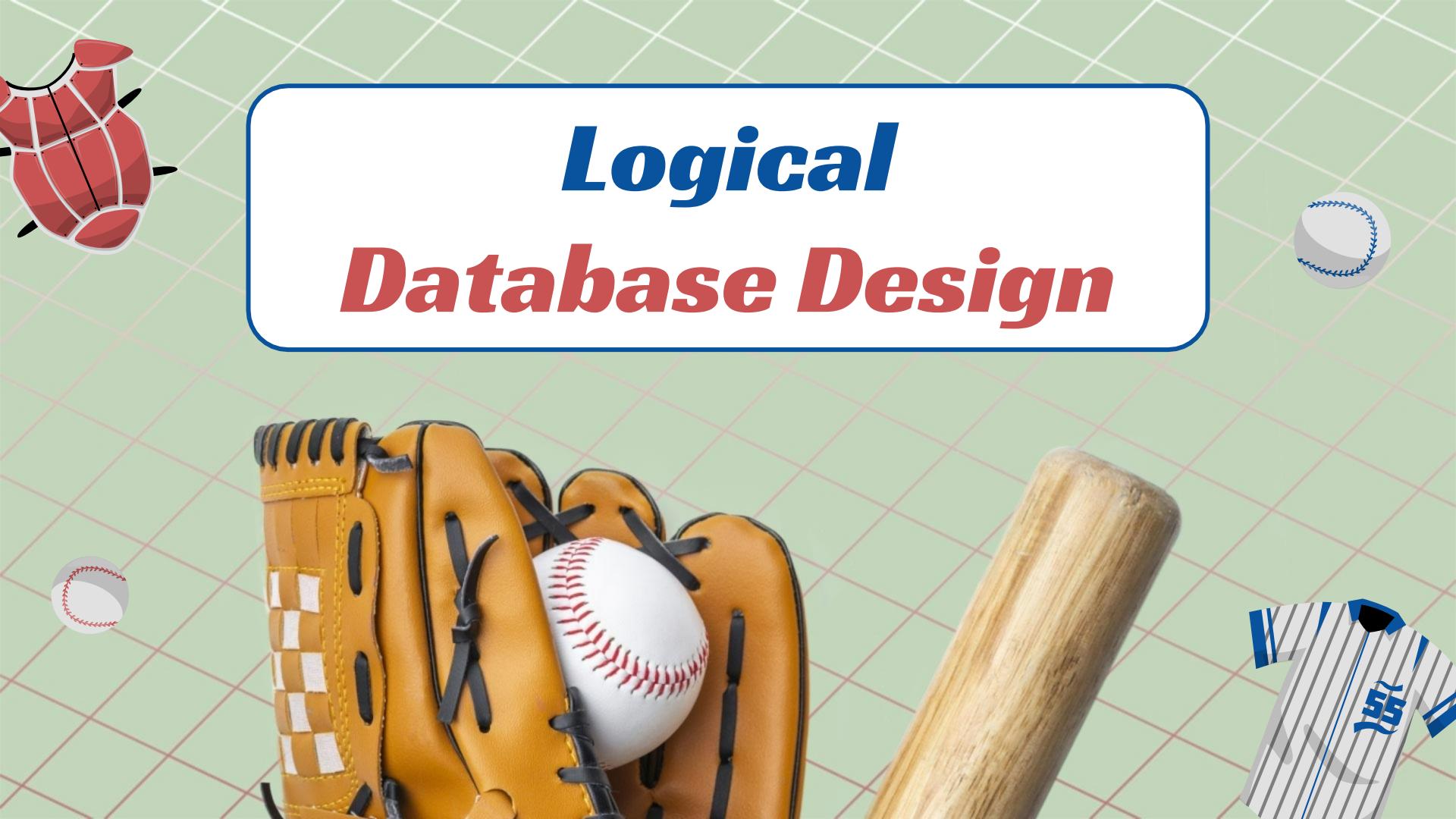
Conceptual Database Design



ER Diagram



Logical Database Design



Relational Schema

Relations:

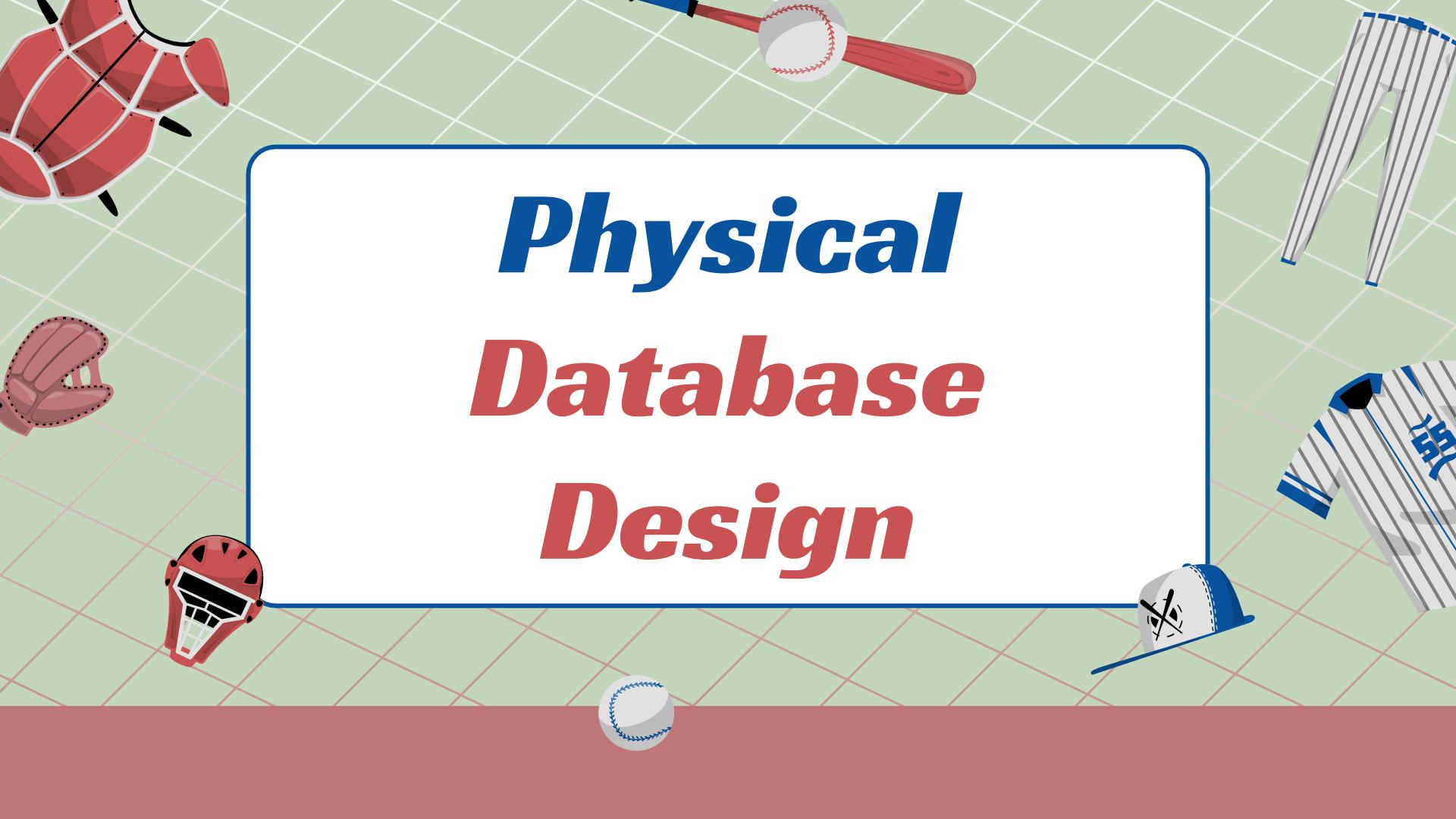
Opponent (oppId, oppName)

Tournament (trnId, trnName)

Location (locId, locName, locState)

Match (oppId, locId, trnId, mchDate, mchTime, mchType, runTerps, runOpponent, ssnYear)

Physical *Database* *Design*



CREATE Statement

Opponent Entity

```
CREATE TABLE [CurveBallConsultants.Opponent] (
    oppId CHAR (3) NOT NULL,
    oppName VARCHAR (40),
    CONSTRAINT pk_Opponent_oppId PRIMARY KEY
(oppId)
)
```

CREATE Statement

Location Entity

```
CREATE TABLE [CurveBallConsultants.Location] (
    locId CHAR(3) NOT NULL,
    locCity CHAR(50),
    locState VARCHAR (20),
    CONSTRAINT pk_Location_locId PRIMARY KEY
    (locId)
)
```

CREATE Statement

Tournament Entity

```
CREATE TABLE [CurveBallConsultants.Tournament] (
    trnId CHAR(3) NOT NULL,
    trnName CHAR (50),
    CONSTRAINT pk_Tournament_trnId PRIMARY KEY
    (trnId)
)
```

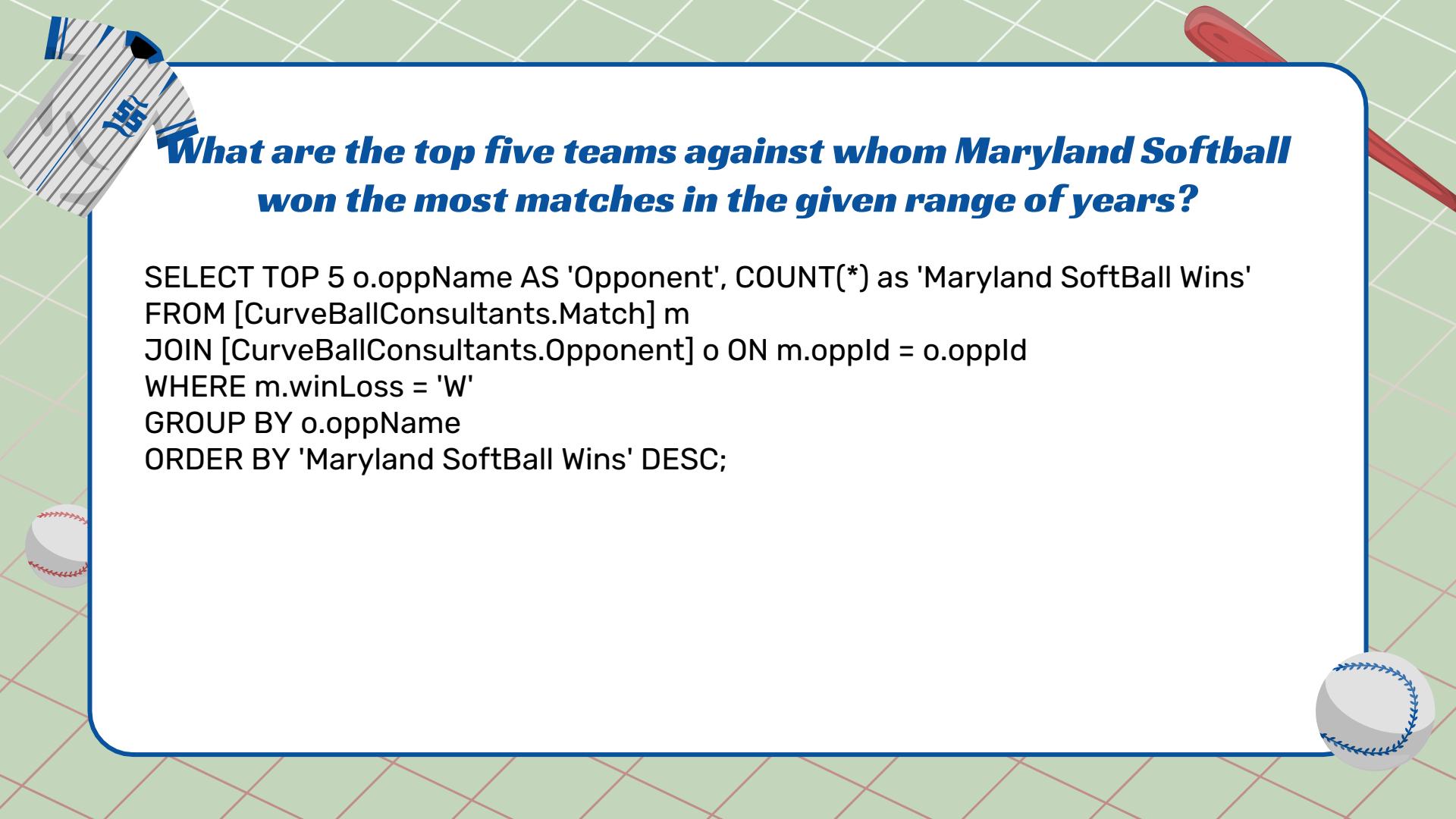
CREATE Statement

Match (associative entity)

```
CREATE TABLE [CurveBallConsultants.Match] (
    oppId CHAR(3) NOT NULL,
    locId CHAR(3) NOT NULL,
    trnId CHAR(3),
    mchDate DATE NOT NULL,
    mchTime TIME NOT NULL,
    mchType VARCHAR (7),
    runTerps INT,
    runOpponent INT,
    winLoss AS (
        CASE
            WHEN runTerps - runOpponent > 0 THEN 'W'
            WHEN runTerps - runOpponent < 0 THEN 'L'
            ELSE 'D'
        END
    ),
);
```

```
    ssnYear CHAR(4),
    CONSTRAINT
    pk_Match_oppId_locId_mchDate_mchTime PRIMARY KEY (oppId,
    locId, mchDate, mchTime),
    CONSTRAINT fk_Match_oppId FOREIGN KEY (oppId)
    REFERENCES
    [CurveballConsultants.Opponent] (oppId)
    ON DELETE NO ACTION ON UPDATE
    CASCADE,
    CONSTRAINT fk_Match_locId FOREIGN KEY (locId)
    REFERENCES
    [CurveballConsultants.Location] (locId)
    ON DELETE NO ACTION ON UPDATE
    CASCADE,
    CONSTRAINT fk_Match_trnId FOREIGN KEY (trnId)
    REFERENCES
    [CurveBallConsultants.Tournament] (trnId)
    ON DELETE NO ACTION ON UPDATE
    CASCADE
);
```

Use Cases



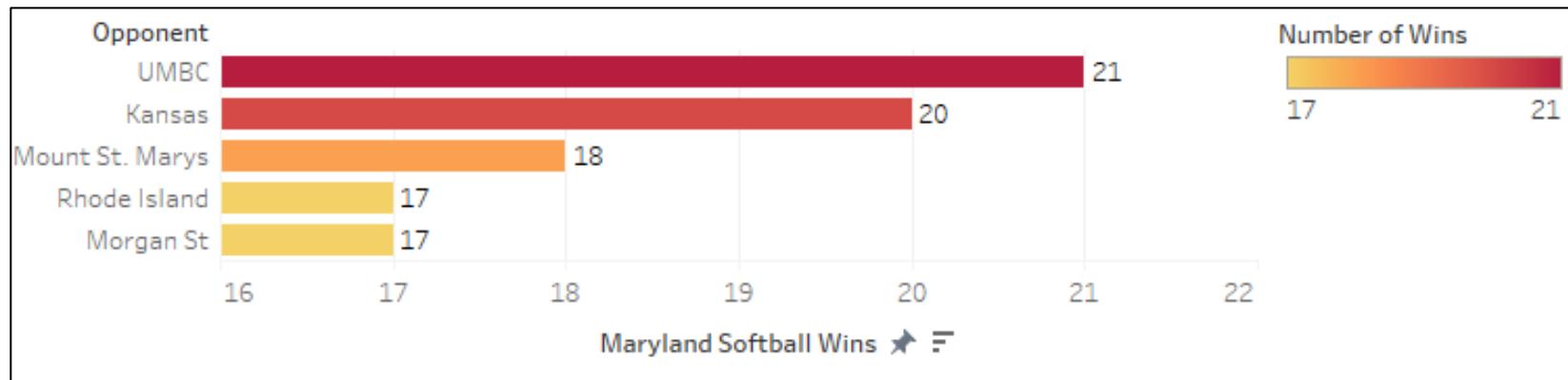
What are the top five teams against whom Maryland Softball won the most matches in the given range of years?

```
SELECT TOP 5 o.oppName AS 'Opponent', COUNT(*) as 'Maryland SoftBall Wins'  
FROM [CurveBallConsultants.Match] m  
JOIN [CurveBallConsultants.Opponent] o ON m.oppId = o.oppId  
WHERE m.winLoss = 'W'  
GROUP BY o.oppName  
ORDER BY 'Maryland SoftBall Wins' DESC;
```

What are the top five teams against whom Maryland Softball won the most matches in the given range of years?

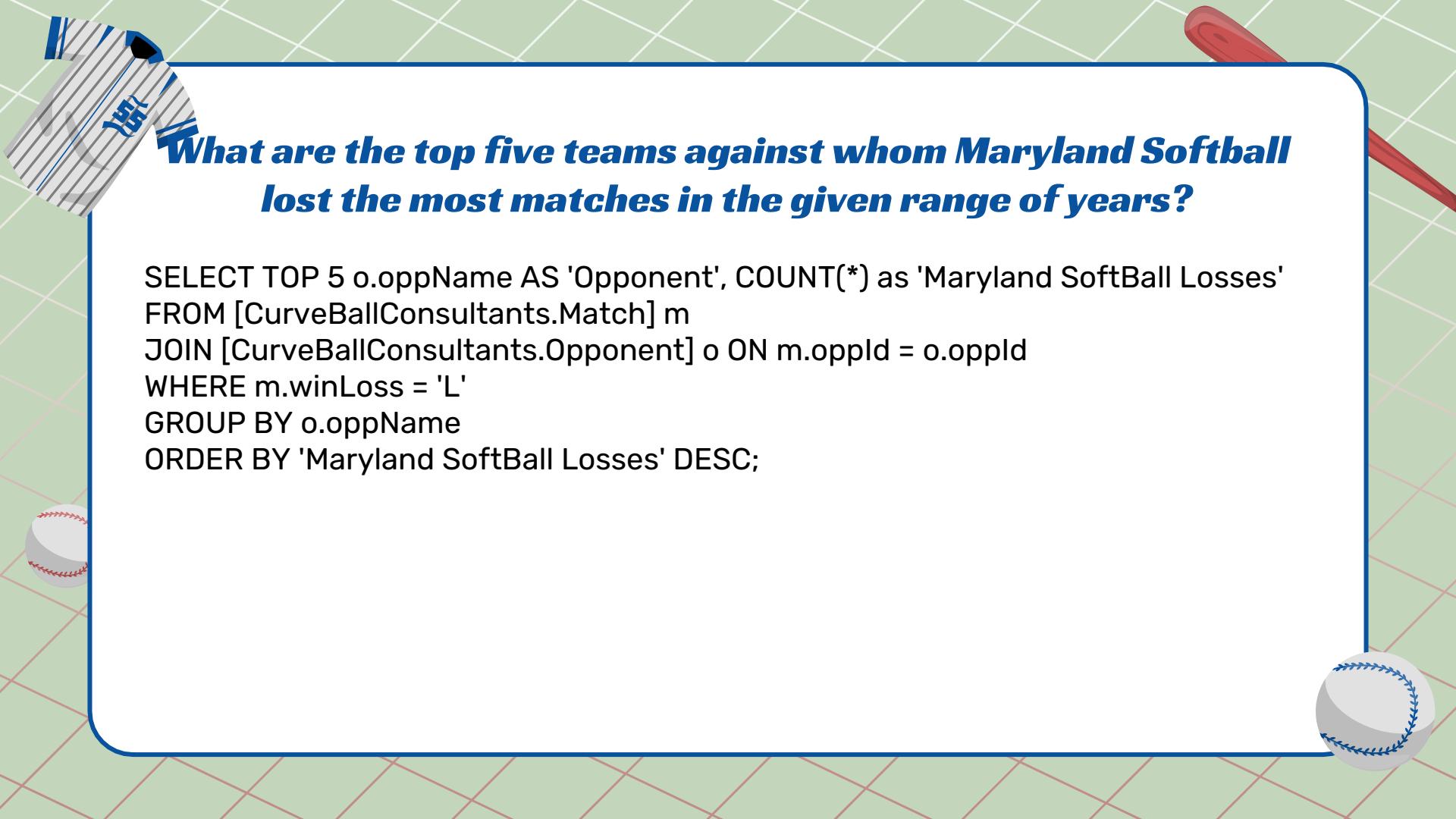
	Opponent	Maryland SoftBall Wins
1	UMBC	21
2	Kansas	20
3	Mount St. Marys	18
4	Morgan St	17
5	Rhode Island	17

What are the top five teams against whom Maryland Softball won the most matches in the given range of years?



Interpretation

- The bar graph indicates that for the given range of years (1999 - 2022), UMBC, Kansas, Mount St. Mary's, Rhode Island, and Morgan St. are the teams that the Maryland Softball Team has won the most games against with 21 wins, 20 wins, 18 wins, 17 wins and 17 wins respectively.



What are the top five teams against whom Maryland Softball lost the most matches in the given range of years?

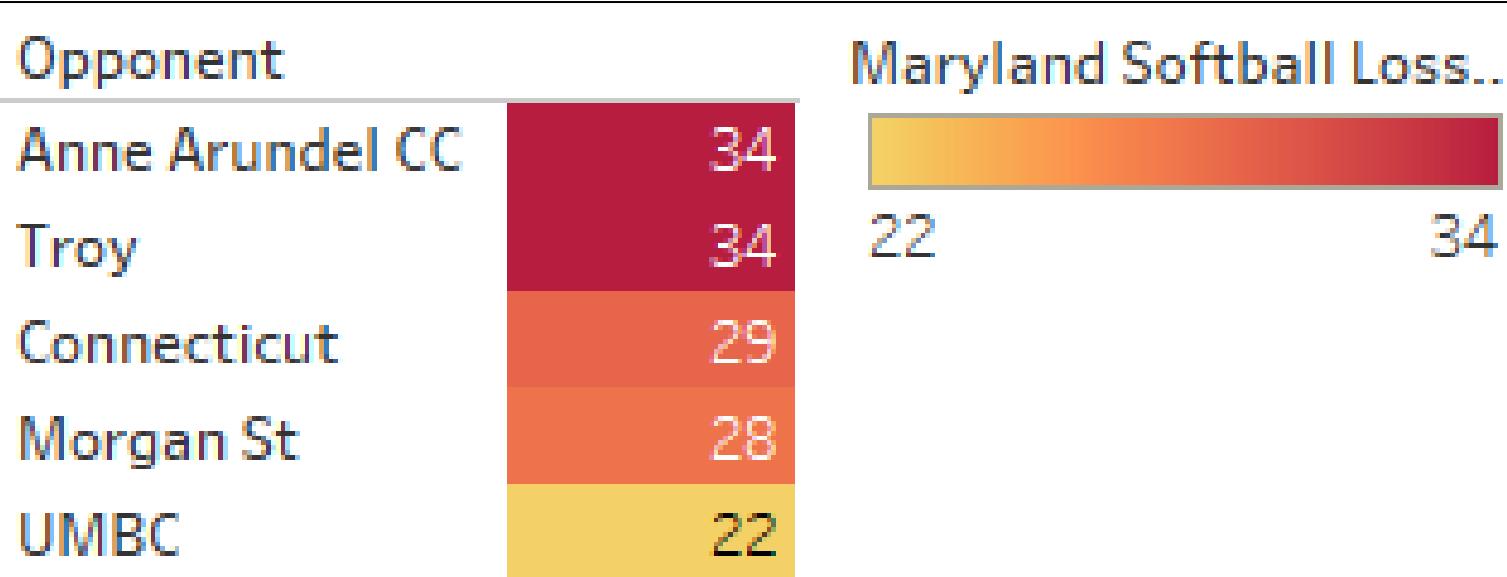
```
SELECT TOP 5 o.oppName AS 'Opponent', COUNT(*) as 'Maryland SoftBall Losses'  
FROM [CurveBallConsultants.Match] m  
JOIN [CurveBallConsultants.Opponent] o ON m.oppId = o.oppId  
WHERE m.winLoss = 'L'  
GROUP BY o.oppName  
ORDER BY 'Maryland SoftBall Losses' DESC;
```



What are the top five teams against whom Maryland Softball lost the most matches in the given range of years?

	Opponent	Maryland SoftBall Losses
1	Anne Arundel CC	34
2	Troy	34
3	Connecticut	29
4	Morgan St	28
5	UMBC	22

What are the top five teams against whom Maryland Softball lost the most matches in the given range of years?



Interpretation

- The table indicates that for the given range of years (1999 - 2022), Anne Arundel CC, Troy, Connecticut, Morgan St and UMBC have been the teams that Maryland Softball Team has lost the most matches against with 34 losses, 34 losses, 29 losses, 28 losses, 22 losses respectively.
- The previous bar graph and this table indicates that Maryland has played numerous matches against UMBC and Morgan St. and are behind on the head to head count by 21-22 and 17-28.

Which are the top 5 locations that have most matches in the given range of years?

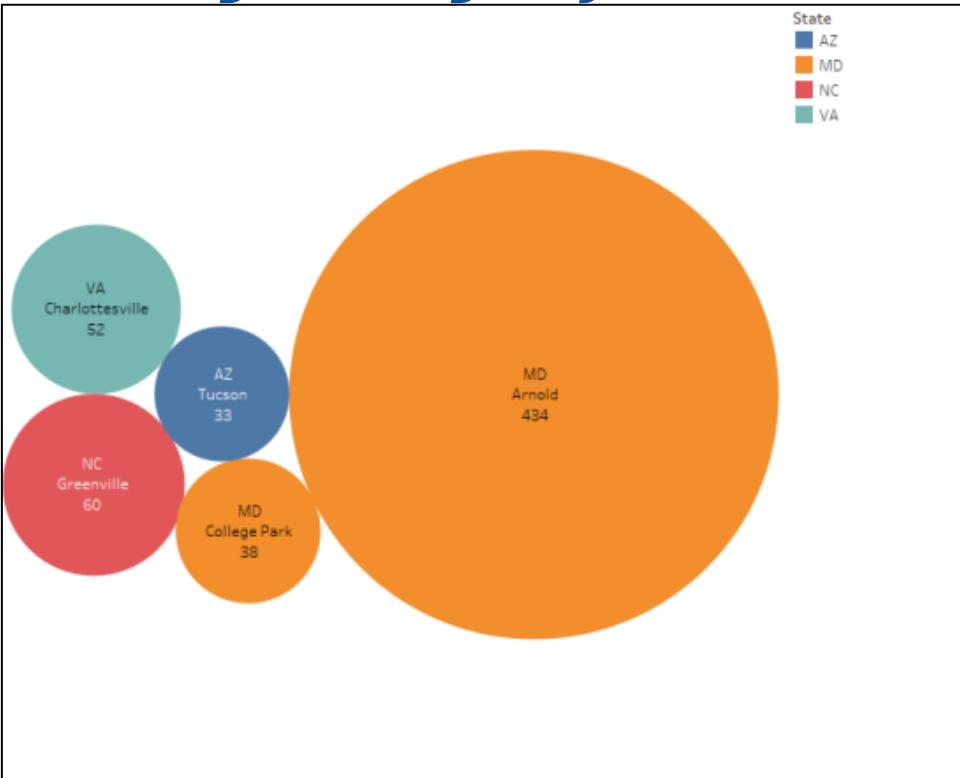
```
SELECT TOP 5 locCity AS 'City', locState AS 'State', COUNT(*) AS 'Matches Hosted'  
FROM [CurveBallConsultants.Match] m  
JOIN [CurveBallConsultants.Location] l ON m.locId = l.locId  
WHERE m.ssnYear BETWEEN '1999' AND '2022'  
GROUP BY l.locCity, l.locState  
ORDER BY COUNT(*) DESC;
```



Which are the top 5 locations that have most matches in the given range of years?

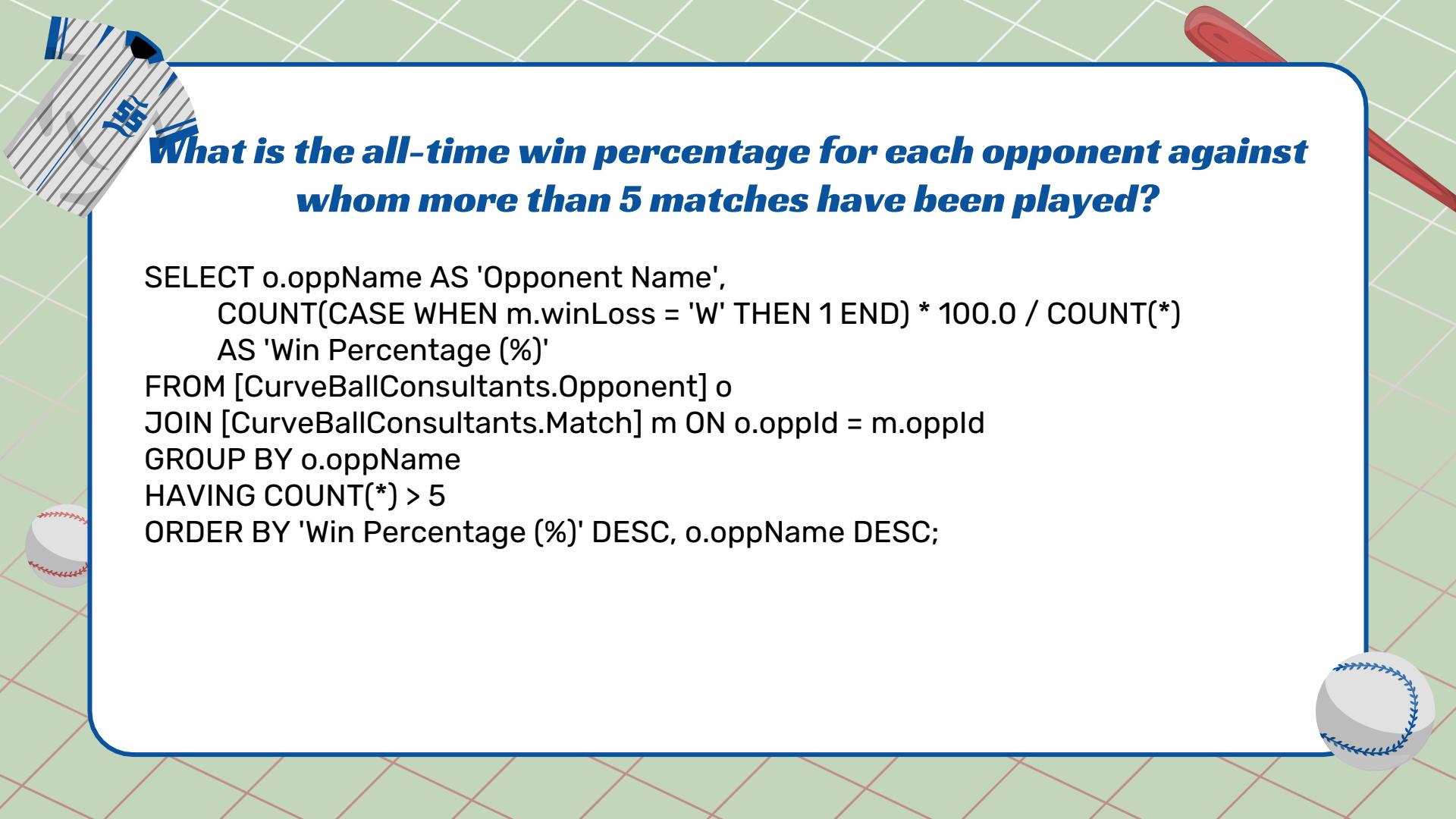
	City	State	Matches Hosted
1	Arnold	MD	434
2	Greenville	NC	60
3	Charlottesville	VA	52
4	College Park	MD	38
5	Tucson	AZ	33

Which are the top 5 locations that have most matches in the given range of years?



Interpretation

- Arnold is the location with the most matches played in the given range of years.
- Greenville, Charlottesville, College Park and Tucson follow Arnold as the locations with the most matches played.



What is the all-time win percentage for each opponent against whom more than 5 matches have been played?

```
SELECT o.oppName AS 'Opponent Name',
       COUNT(CASE WHEN m.winLoss = 'W' THEN 1 END) * 100.0 / COUNT(*)
             AS 'Win Percentage (%)'
  FROM [CurveBallConsultants.Opponent] o
 JOIN [CurveBallConsultants.Match] m ON o.oppId = m.oppId
 GROUP BY o.oppName
 HAVING COUNT(*) > 5
 ORDER BY 'Win Percentage (%)' DESC, o.oppName DESC;
```



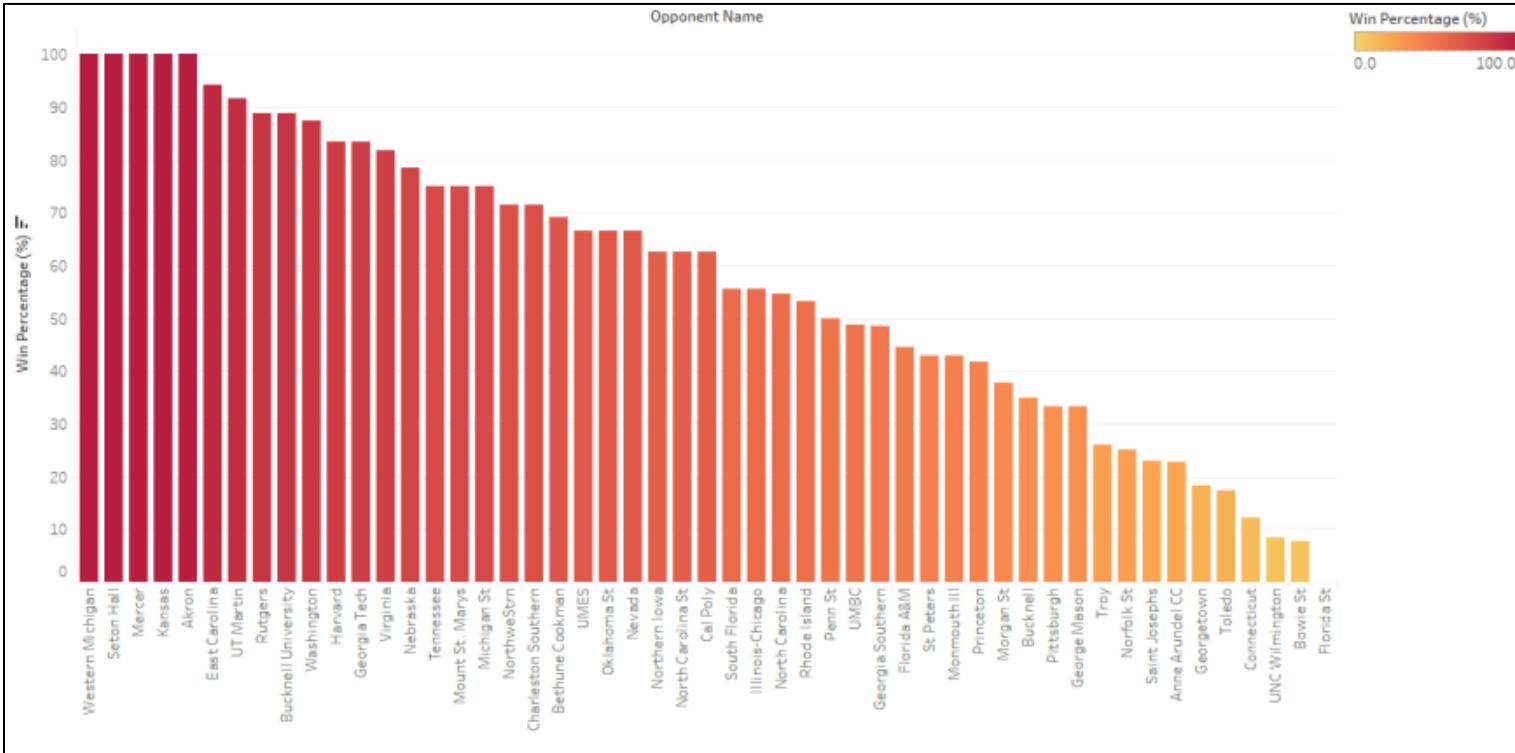
What is the all-time win percentage for each opponent against whom more than 5 matches have been played?

	Opponent Name	Win Percentage (%)
1	Western Michigan	100.00
2	Seton Hall	100.00
3	Mercer	100.00
4	Kansas	100.00
5	Akron	100.00
6	East Carolina	94.12
7	UT Martin	91.67
8	Rutgers	88.89*
9	Bucknell University	88.89
10	Washington	87.50
11	Harvard	83.33
12	Georgia Tech	83.33
13	Virginia	81.82
14	Nebraska	78.57
15	Tennessee	75.00
16	Mount St. Marys	75.00
17	Michigan St	75.00

	Opponent Name	Win Percentage (%)
18	Northwestern	71.43
19	Charleston Southern	71.43
20	Bethune Cookman	69.23
21	UMES	66.67
22	Oklahoma St	66.67
23	Nevada	66.67
24	Northern Iowa	62.50
25	North Carolina St	62.50
26	Cal Poly	62.50
27	South Florida	55.56
28	Illinois-Chicago	55.56
29	North Carolina	54.55
30	Rhode Island	53.13
31	Penn St	50.00
32	UMBC	48.84
33	Georgia Southern	48.48
34	Florida A&M	44.44

	Opponent Name	Win Percentage (%)
35	St Peters	42.86
36	Monmouth III	42.86
37	Princeton	41.67
38	Morgan St	37.78
39	Bucknell	35.00
40	Pittsburgh	33.33
41	George Mason	33.33
42	Troy	26.09
43	Norfolk St	25.00
44	Saint Josephs	23.08
45	Anne Arundel CC	22.73
46	Georgetown	18.18
47	Toledo	17.39
48	Connecticut	12.12
49	UNC Wilmington	8.33
50	Bowie St	7.69
51	Florida St	0.00

What is the all-time win percentage for each opponent against whom more than 5 matches have been played?



Interpretation

- Strongest opponents:
 - UNC Wilmington, Bowie State, and Florida State seem to be the strongest opponents for Maryland Softball, as they have the lowest win percentages against them.
 - Maryland Softball has a win rate of less than 25% against these teams.
- Weakest opponents:
 - Western Michigan, Seton Hall, and Mercer are the weakest opponents for Maryland Softball, as they have a 100% win rate against them.
 - Additionally, Maryland has a win rate of over 75% against several other opponents, including Washington, Harvard, and UMES.
- Balanced competition:
 - Against several opponents, such as Penn State, UMBC, and Illinois-Chicago, Maryland Softball has a win rate around 50%, suggesting a more balanced competition.

What is Maryland's win percentage at home across all years?

```
SELECT m.ssnYear AS 'Season',
       COUNT(CASE WHEN m.winLoss = 'W' THEN 1 END) * 100.0 / COUNT(*)
             AS 'Win Percentage (%)'
  FROM [CurveBallConsultants.Match] m
 WHERE m.mchType = 'Home'
 GROUP BY m.ssnYear;
```

What is Maryland's win percentage at home across all years?

	Season	Win Percentage (%)
1	1999	76.00
2	2000	78.57
3	2001	80.00
4	2002	66.67
5	2003	66.67
6	2004	56.10
7	2005	62.07
8	2006	76.19
9	2007	55.56
10	2008	77.78
11	2009	54.05
12	2010	67.50
13	2011	66.67

	Season	Win Percentage (%)
12	2010	67.50
13	2011	66.67
14	2012	66.67
15	2013	69.23
16	2014	57.14
17	2015	42.11
18	2016	31.25
19	2017	29.41
20	2018	38.89
21	2019	35.29
22	2021	56.25
23	2022	58.82

What is the win percentage at home across all years?



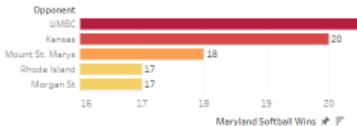
Interpretation

- The Terps had a strong start to the 21st century, winning at least 70% of their home games each year from 1999 to 2001.
- From 2002 to 2005, the Terps' win percentage at home fluctuated between 60% and 70%.
- From 2006 to 2009, the Terps' had a very inconsistent performance as they won 76% of their home games in 2006, which dropped to 56% in 2006. They improved their performance and won 77% of their home games in 2008 but had a poor performance again in 2009 by winning only 55% of their home games.
- From 2010 to 2013, the Terps' win percentage at home remained relatively stable, between 67% and 69%.
- From 2014 to 2017, the Terps' home game performances kept dropping as their win percentage declined from 57% to 29%.
- The Terps' had a relatively better home game performance in 2018 and 2019 as there win percentage increased to 38% and 35% respectively.
- The Terps' had a more convincing home game performance in 2021 and 2022 as their win percentage increased to 56% and 59% respectively.

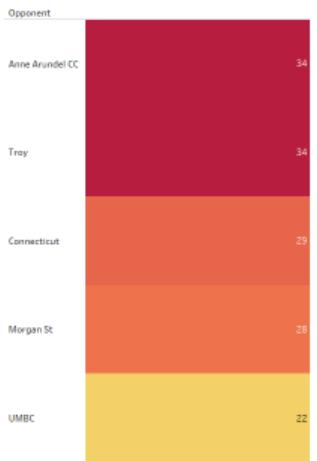
It is important to note that the win percentage at home is just one measure of a team's success. The Terps have still had some successful seasons in recent years, even though their win percentage at home has declined. For example, the Terps won the Big Ten Conference tournament in 2019 and 2022.



What are the top five teams against whom Maryland Softball won the most matches in the given range of years?



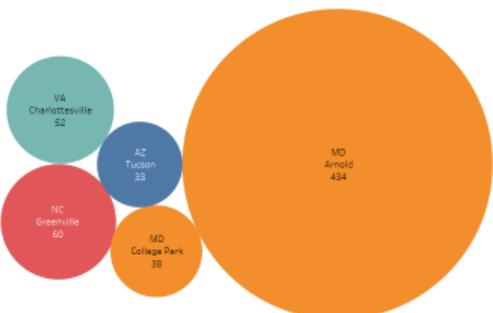
What are the top five teams against whom Maryland Softball lost the most matches in the given range of years?



What is the win percentage at home across all years?



Which are the top 5 locations that have most matches in the given range of years?



**Thank
You!**

Go Terps!

