



Terrapin Tactics

Project_0507_05

Sriram Ganesh



Jasvinder Singh

Suvedaranyan Alliarasi Murugaian

Taniya Agrawal

Our Client

- The UMD Softball Team, is a competitive collegiate softball team representing the University of Maryland.
- The team is composed of skilled athletes committed to excellence both on and off the field.
- The primary objective is to enhance the team's overall performance by leveraging data-driven insights.
- Terrapin Tactics aim to optimize strategies, identify strengths and weaknesses, and make informed decisions for more effective gameplay.





Mission Statement

We aim to help UMD softball team find their dream strategy! Using UMD's softball team data of past 20 years to identify key factors and discover insights to help improve their performance



Mission Objective

- **Venue Performance Evaluation:** Conduct a comprehensive comparison of the team's performance in various settings—home, away, and neutral venues—to inform strategic venue selection for future seasons and recommend modifications to enhance the home field advantage, thereby improving overall win rates.
- **Opponent Performance Analysis:** Examine historical game data to pinpoint the top 10 teams where UMD has excelled and the 10 where they have struggled the most. This analysis will be used to derive strategic insights aimed at bolstering team performance against a diverse range of opponents.
- **Seasonal Impact Review:** Investigate how different seasons and associated weather conditions affect team performance. This review will identify any seasonal patterns that could influence training schedules and match strategies.
- **Venue Win Percentage Breakdown:** Analyze UMD's win percentages at various venues, identifying the top and bottom 10 to uncover trends and inform future game planning, ensuring the team is positioned to capitalize on venues where they have historically performed well.



Agenda

Overview & Introduction

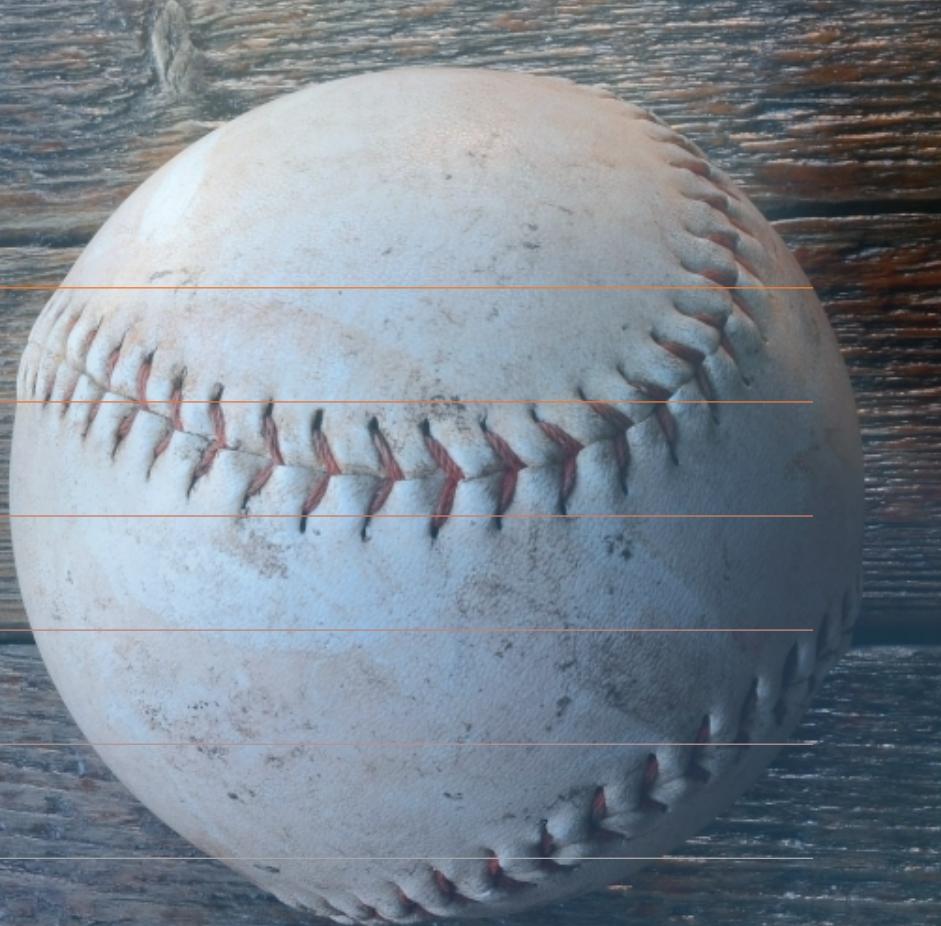
Conceptual Database Design

Data Definition Language

Data Manipulation Language

Use Case & Applications

Deep Dive





Project Flow



Data Collection

Collected data from
umterps.com



Data Preprocessing

Data & Time Formatting,
Redundancy Removal, Data
Cleaning, Primary key
assignment, appended
seasons attribute



Data Analysis

Performance analysis using
win % for different seasons,
venues and opponents



Data Visualisation

Visualizing use cases and
interpreting



Data Evaluation

Total rows: 1052 (2004-2023).
Comprehensive coverage from 2004 to 2023.

Redundancy and Data Trimming:

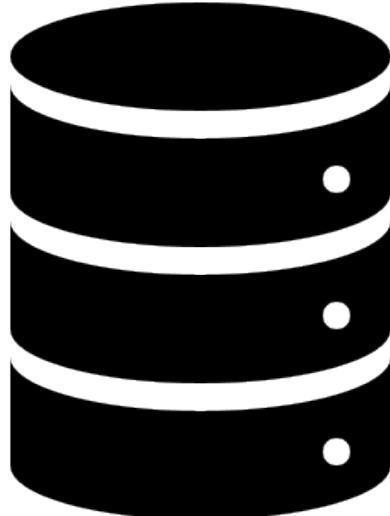
- Identified and eliminated redundant fields.
- Removed non-essential data for client relevance.

Impact on Data Quality:

- Streamlined dataset for improved clarity.
- Enhanced relevance and efficiency.

Next Steps:

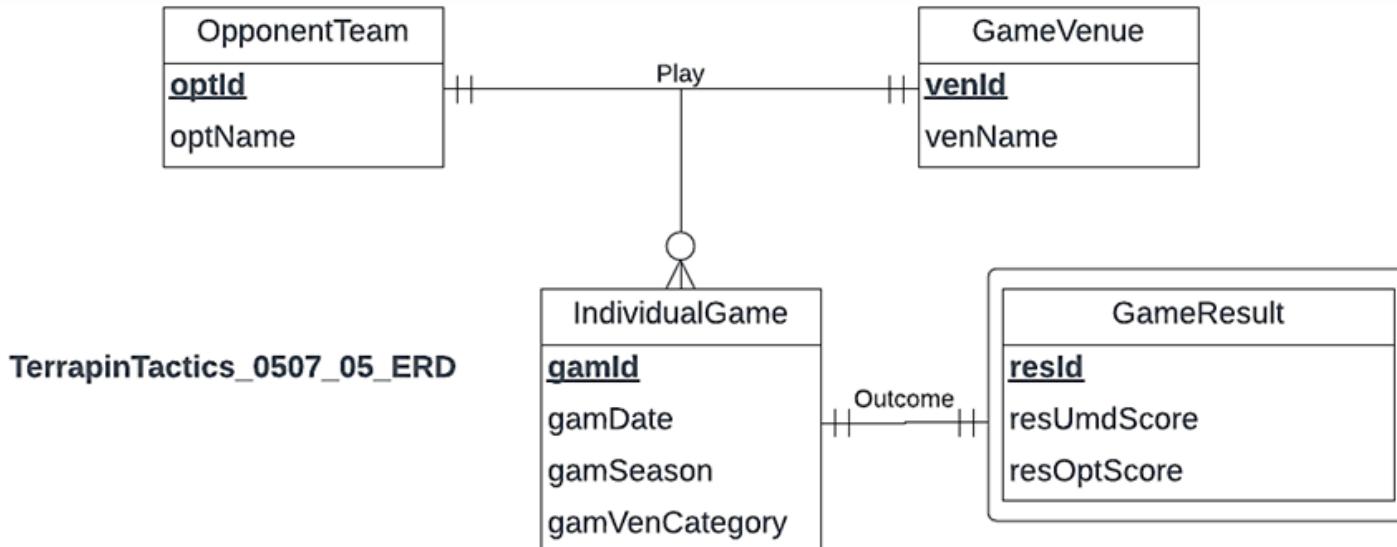
- Outline the subsequent steps in data processing.
- Invite questions for further clarification.



Database Design



Conceptual Database Design





Logical Database Design

OpponentTeam (optId, optName)

IndividualGame (gamId, gamDate, gamSeason, gamVenCategory)

GameVenue (venId, venName)

GameResult (resId, resUmdScore, resOptScore, *gamId*)

Play (gamId, optId, venId)



Referential Integrities

Relation	Foreign Key	Base Relation	Primary Key	Business Rule	Business Constraint: ON DELETE	Business Rule	Business Constraint: ON UPDATE
Outcome	gamId	IndividualGame	gamId	R1	NO ACTION	R1	NO ACTION
Outcome	resId	GameResult	resId	R1	NO ACTION	R1	NO ACTION
Play	venId	GameVenue	venId	R2	NO ACTION	R3	CASCADE
Play	gamId	IndividualGame	gamId	R4	NO ACTION	R4	NO ACTION
Play	optId	OpponentTeam	optId	R4	NO ACTION	R4	NO ACTION



[R1] When a game has been completed, the details of the IndividualGame and GameResult cannot be updated or deleted.

[R2] When a GameVenue details have been deleted, the corresponding details of the OpponentTeam and Individual Game cannot be deleted.

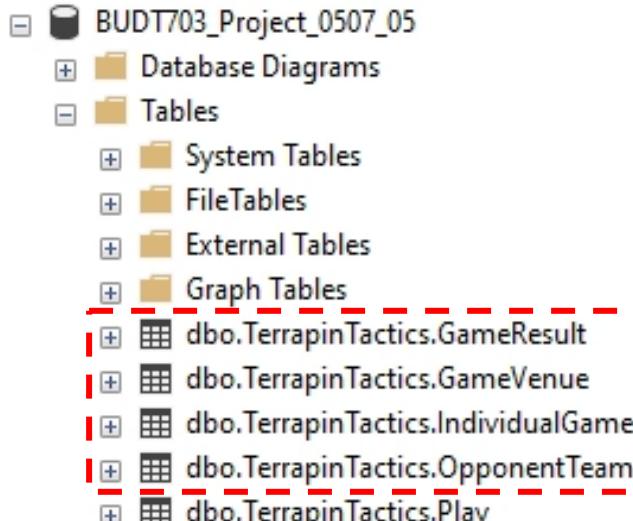
[R3] When a GameVenue details have been updated, it should cascade to update the corresponding details of the OpponentTeam and Individual Game.

[R4] When a game has been played, the corresponding details of the OpponentTeam and Individual Game cannot be updated or deleted.

Business Rules



Physical Database Design



```
CREATE TABLE [TerrapinTactics.IndividualGame] (
    gamId CHAR (5) NOT NULL,
    gamDate DATE,
    gamSeason VARCHAR (10),
    gamVenCategory VARCHAR (10),
    CONSTRAINT pk_IndividualGame_gamId PRIMARY KEY (gamId)
)
```

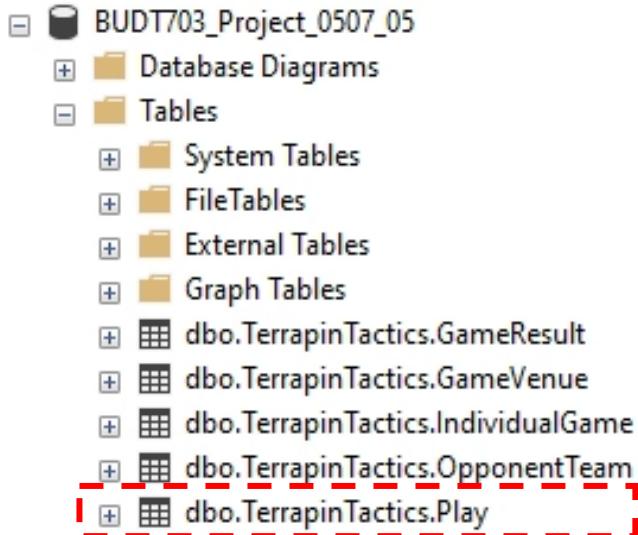
```
CREATE TABLE [TerrapinTactics.GameResult] (
    resId CHAR (5) NOT NULL,
    resUmdScore INT,
    resOptScore INT,
    gamId CHAR (5),
    CONSTRAINT pk_GameResult_resId PRIMARY KEY (resId),
    CONSTRAINT fk_GameResult_gamId FOREIGN KEY (gamId)
        REFERENCES [TerrapinTactics.IndividualGame] (gamId)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION
)
```

```
CREATE TABLE [TerrapinTactics.GameVenue] (
    venId CHAR (5) NOT NULL,
    venName VARCHAR (50),
    CONSTRAINT pk_GameVenue_venId PRIMARY KEY (venId)
)
```

```
CREATE TABLE [TerrapinTactics.OpponentTeam] (
    optId CHAR (5) NOT NULL,
    optName VARCHAR (50),
    CONSTRAINT pk_OpponentTeam_optId PRIMARY KEY (optId),
)
```



Physical Database Design



```
CREATE TABLE [TerrapinTactics.Play] (
    gamId CHAR (5),
    optId CHAR (5),
    venId CHAR (5),
    CONSTRAINT pk_Play_gamId_optId_venId PRIMARY KEY (gamId,optId,venId),
    CONSTRAINT fk_Play_gamId FOREIGN KEY (gamId)
        REFERENCES [TerrapinTactics.IndividualGame] (gamId)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION,
    CONSTRAINT fk_Play_optId FOREIGN KEY (optId)
        REFERENCES [TerrapinTactics.OpponentTeam] (optId)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION,
    CONSTRAINT fk_Play_venId FOREIGN KEY (venId)
        REFERENCES [TerrapinTactics.GameVenue] (venId)
        ON DELETE NO ACTION
        ON UPDATE CASCADE
)
```

Use Cases & Applicatio n





Key Points

- In case of home venue, the visiting team bats first.
- In cases of neutral venues, the team with the higher ranking assumes the role of the home team, and the other team bats first.
- Softball maintains standardized dimensions for the diamond field and pitching arc, but the outfield fence dimension and wall height may exhibit variability across different venues.
- Pitchers typically experience favorable conditions in winter, as the conditions provide a natural edge in control and movement. Due to reduced elasticity during winters the ball becomes harder to hit and often leads to more groundouts and pop-ups.
- Hitters experience advantageous circumstances in spring, as the warmer temperature and softer fields leads to predictable ball and allows power hitters to push the ball farther with higher launch angles.



Case 1: Year on year win percentage by

```
DROP VIEW IF EXISTS SeasonalPerformanceTrends
GO
CREATE VIEW SeasonalPerformanceTrends AS
SELECT YEAR(i.gamDate) AS Year, gamSeason,
       COUNT(*) AS TotalGames,
       SUM(CASE WHEN resUmdScore > resOptScore THEN 1 ELSE 0 END) AS Wins,
       SUM(CASE WHEN resUmdScore < resOptScore THEN 1 ELSE 0 END) AS Losses,
       ROUND(
           CASE
               WHEN SUM(CASE WHEN resUmdScore < resOptScore THEN 1 ELSE 0 END) = 0 THEN NULL
               ELSE CAST(SUM(CASE WHEN resUmdScore > resOptScore THEN 1 ELSE 0 END) AS FLOAT) / SUM(CASE WHEN resUmdScore < resOptScore THEN 1 ELSE 0 END)
           END, 2) AS WinLossRatio
FROM [TerrapinTactics.IndividualGame] i
JOIN [TerrapinTactics.GameResult] ON i.gameId = [TerrapinTactics.GameResult].gameId
GROUP BY YEAR(gamDate), gamSeason
GO
SELECT * FROM SeasonalPerformanceTrends
ORDER BY Year
```

	Year	gamSeason	TotalGames	Wins	Losses	WinPercentage
1	2004	Spring	50	28	22	56
2	2004	Winter	10	5	5	50
3	2005	Spring	40	23	17	57.5
4	2005	Winter	5	1	4	20
5	2006	Spring	48	26	22	54.17
6	2006	Winter	11	6	5	54.55
7	2007	Spring	48	28	20	58.33
8	2007	Winter	11	5	6	45.45

	gamSeason	WinPercentage
1	Winter	47.18
2	Spring	50.62



Visualization: Year on year win percentage by season





Case 2: Year on Year win percentage by

Venue

```
DROP VIEW IF EXISTS HomeAwayPerformancePerYear
GO
CREATE VIEW HomeAwayPerformancePerYear AS
SELECT
    YEAR(i.gamDate) AS Year,
    i.gamVenCategory AS 'Venue Category',
    COUNT(*) AS TotalGames,
    SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS Wins,
    SUM(CASE WHEN g.resUmdScore < g.resOptScore THEN 1 ELSE 0 END) AS Losses,
    ROUND(
        CAST(SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS FLOAT) / COUNT(*) * 100, 2
    ) AS WinPercentage
FROM [TerrapinTactics.IndividualGame] i
JOIN [TerrapinTactics.GameResult] g ON i.resId = g.resId
GROUP BY YEAR(i.gamDate), i.gamVenCategory;
GO

SELECT * FROM HomeAwayPerformancePerYear
ORDER BY Year, 'Venue Category';
```

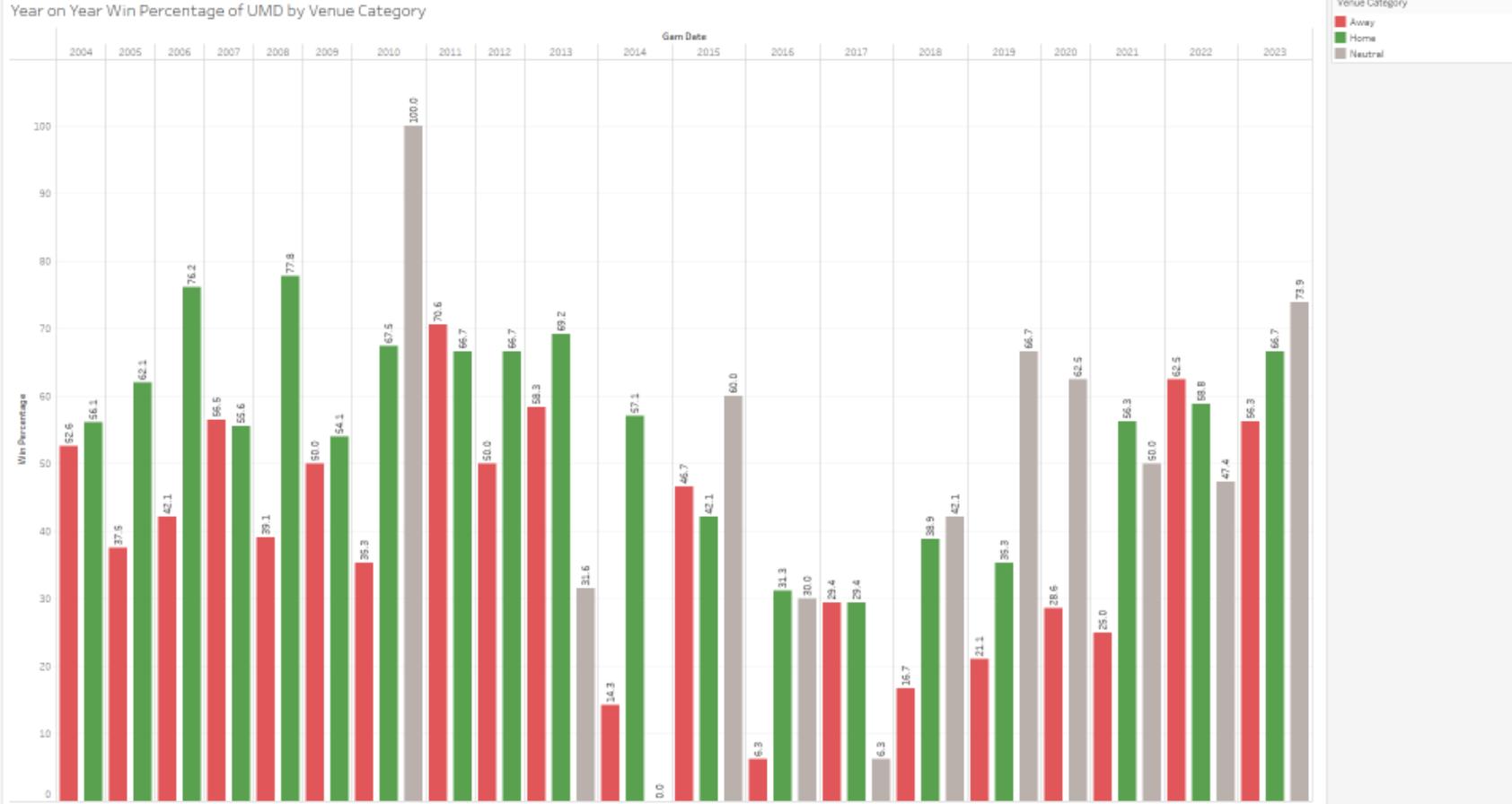
Results					
	Year	Venue Category	TotalGames	Wins	Losses
1	2004	Away	19	10	9
2	2004	Home	41	23	18
3	2005	Away	16	6	10
4	2005	Home	29	18	11
5	2006	Away	38	16	22
6	2006	Home	21	16	5
7	2007	Away	23	13	10
8	2007	Home	36	20	16

	Venue Category	TotalGames	Wins	Losses	WinPercentage
1	Away	364	146	218	40.11
2	Home	496	292	204	58.87
3	Neutral	192	86	105	44.79



Visualization: Year on Year win percentage by

Venue



Case 3: Top 10 win % against opponents against whom UMD has played 10 or more games

```
DROP VIEW IF EXISTS TopTenOpponentWinPercentages
GO
CREATE VIEW TopTenOpponentWinPercentages AS
WITH WinPercentages AS (
    SELECT
        o.optName AS Opponent, COUNT(*) AS TotalGames,
        SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS Wins,
        SUM(CASE WHEN g.resUmdScore < g.resOptScore THEN 1 ELSE 0 END) AS Losses,
        CAST(SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS FLOAT) / COUNT(*) * 100 AS WinPercentage
    FROM [TerrapinTactics.IndividualGame] i
    JOIN [TerrapinTactics.GameResult] g ON i.gamId = g.gamId
    JOIN [TerrapinTactics.Play] p ON i.gamId = p.gamId
    JOIN [TerrapinTactics.OpponentTeam] o ON p.optId = o.optId
    GROUP BY o.optName
)
SELECT TOP 10
    Opponent, TotalGames, Wins, Losses, ROUND(WinPercentage, 2) AS WinPercentage
FROM WinPercentages
WHERE TotalGames >= 10
ORDER BY WinPercentage DESC;
GO
SELECT * FROM TopTenOpponentWinPercentages;
```

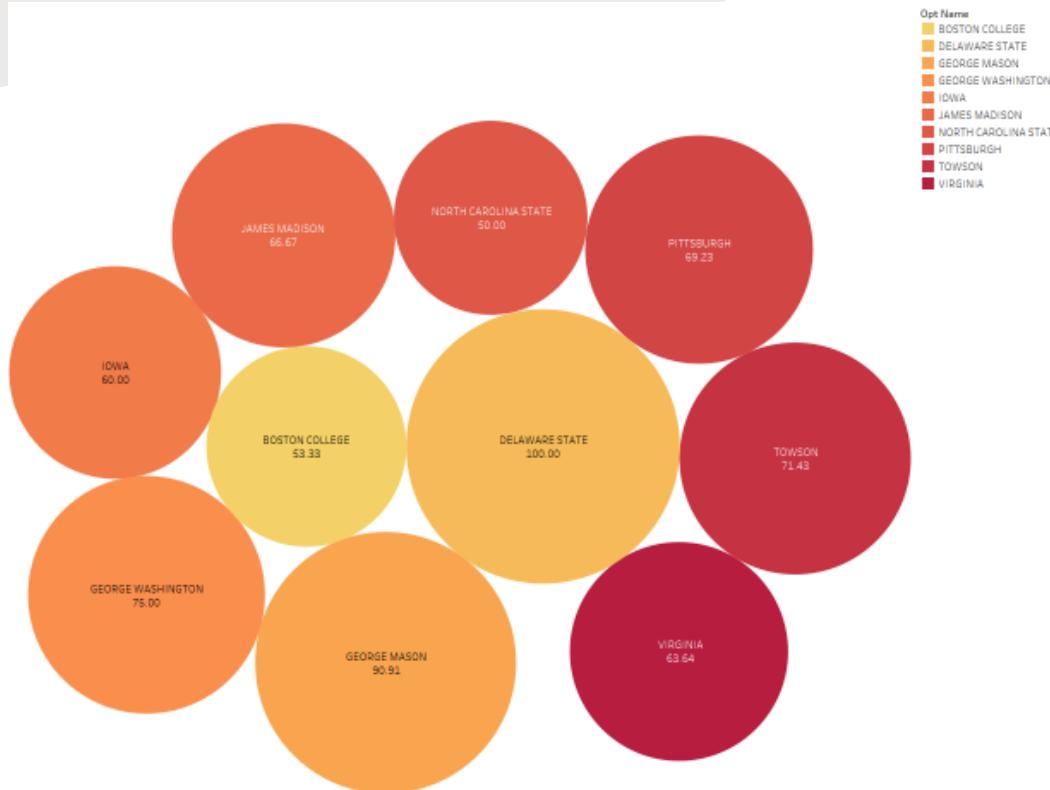
Results Messages

	Opponent	TotalGames	Wins	Losses	WinPercentage
1	DELAWARE STATE	20	20	0	100
2	GEORGE MASON	11	10	1	90.91
3	GEORGE WASHINGTON	12	9	3	75
4	TOWSON	14	10	4	71.43
5	PITTSBURGH	13	9	4	69.23
6	JAMES MADISON	12	8	4	66.67
7	VIRGINIA	33	21	12	63.64
8	IOWA	20	12	8	60

lmd.edu,97... DESKTOP-L068QVT\sri... BUDT703_Project_0507_05 | 00:00:00 | 10 rows



Visualization: Top 10 win % against opponents against whom UMD has played 10 or more games





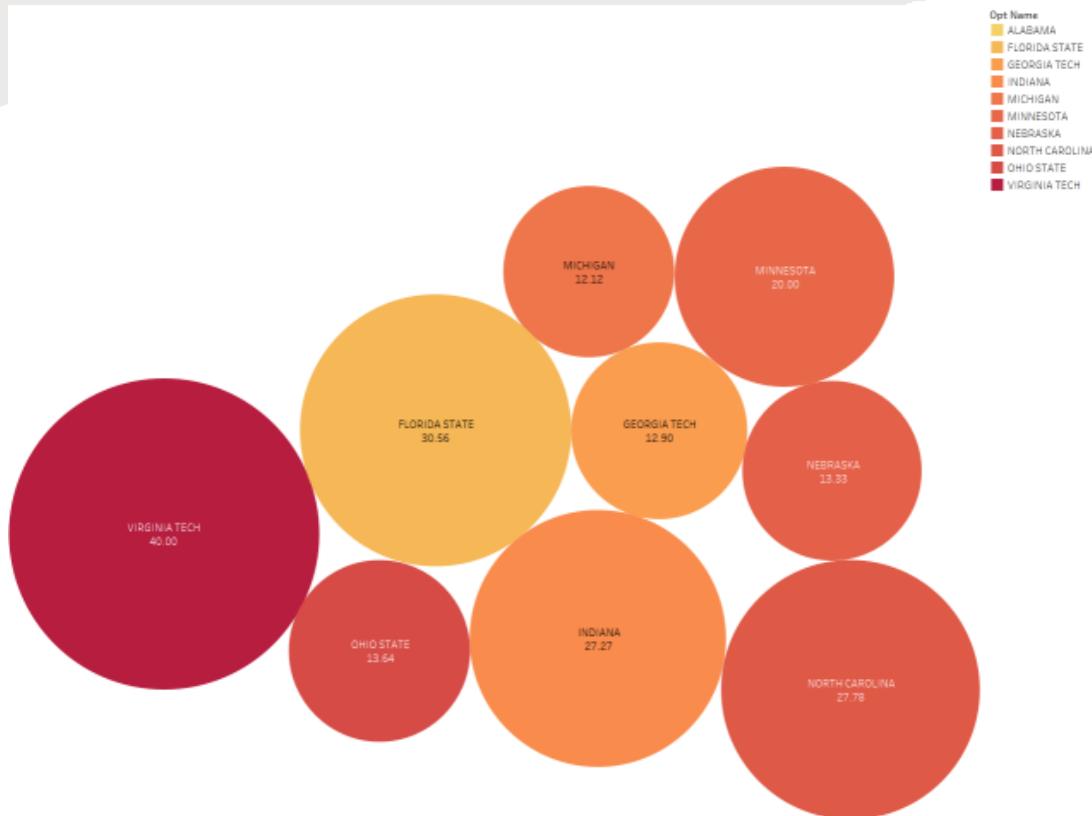
Case 4: Worst 10 win % against opponents against whom UMD has played 5 or more games

```
DROP VIEW IF EXISTS BottomTenOpponentWinPercentages
GO
CREATE VIEW BottomTenOpponentWinPercentages AS
WITH WinPercentages AS (
    SELECT
        o.optName AS Opponent, COUNT(*) AS TotalGames, SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS Wins,
        SUM(CASE WHEN g.resUmdScore < g.resOptScore THEN 1 ELSE 0 END) AS Losses,
        CAST(SUM(CASE WHEN g.resUmdScore > g.resOptScore THEN 1 ELSE 0 END) AS FLOAT) / COUNT(*) * 100 AS WinPercentage
    FROM [TerrapinTactics.IndividualGame] i
    JOIN [TerrapinTactics.GameResult] g ON i.gamId = g.gamId
    JOIN [TerrapinTactics.Play] p ON i.gamId = p.gamId
    JOIN [TerrapinTactics.OpponentTeam] o ON p.optId = o.optId
    GROUP BY o.optName
)
SELECT TOP 10
    Opponent, TotalGames, Wins, Losses, ROUND(WinPercentage, 2) AS WinPercentage
FROM WinPercentages
WHERE TotalGames >= 10
ORDER BY WinPercentage ASC; -- Ordering by ascending win percentage to get the bottom 10
GO
SELECT * FROM BottomTenOpponentWinPercentages;
```

Results Messages

	Opponent	TotalGames	Wins	Losses	WinPercentage
1	ALABAMA	10	0	10	0
2	MICHIGAN	33	4	29	12.12
3	GEORGIA TECH	31	4	27	12.9
4	NEBRASKA	15	2	13	13.33
5	OHIO STATE	22	3	19	13.64
6	MINNESOTA	15	3	12	20
7	INDIANA	22	6	16	27.27
8	NORTH CAROLINA	36	10	26	27.78

Visualization: Worst 10 win % against opponents against whom UMD has played 5 or more games





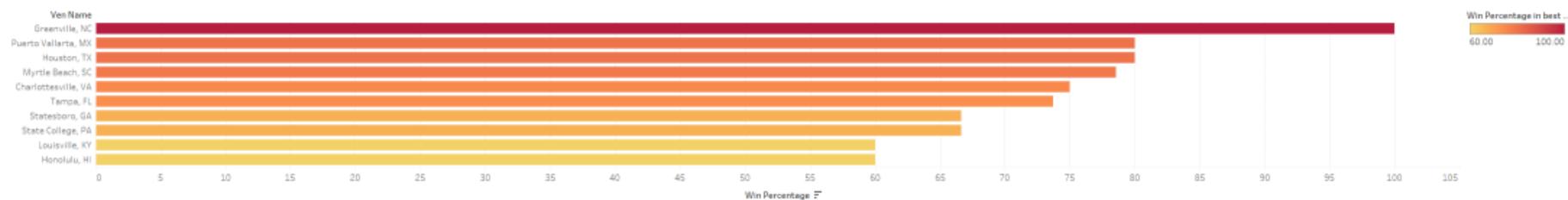
Case 5:

```
DROP VIEW IF EXISTS TopVenuesByWinPercentage
GO
CREATE VIEW TopVenuesByWinPercentage AS
WITH VenueWins AS (
    SELECT
        gv.venId, gv.venName, COUNT(CASE WHEN gr.resUmdScore > gr.resOptScore THEN 1 END) AS Wins,
        COUNT(*) AS TotalGames
    FROM [TerrapinTactics.GameVenue] gv
    JOIN [TerrapinTactics.Play] p ON gv.venId = p.venId
    JOIN [TerrapinTactics.GameResult] gr ON p.gamId = gr.gamId
    GROUP BY gv.venId, gv.venName
    HAVING COUNT(*) >= 5 -- Ensures only venues where UMD has played 5 or more games are considered
)
SELECT TOP 10
    venName, Wins, TotalGames, ROUND((CAST(Wins AS FLOAT) / TotalGames) * 100, 2) AS WinPercentage
FROM VenueWins
ORDER BY WinPercentage DESC, venName;
GO
SELECT * FROM TopVenuesByWinPercentage;
```

	venName	Wins	TotalGames	WinPercentage
1	Greenville, NC	5	5	100
2	Houston, TX	4	5	80
3	Puerto Vallarta, MX	4	5	80
4	Myrtle Beach, SC	11	14	78.57
5	Charlottesville, VA	9	12	75
6	Tampa, FL	14	19	73.68
7	State College, PA	4	6	66.67
8	Statesboro, GA	6	9	66.67
9	Honolulu, HI	3	5	60
10	Louisville, KY	3	5	60

Visualization:

Top 10 best performing venues for UMD





Case 6:

```
DROP VIEW IF EXISTS BottomVenuesByWinPercentage;
GO
CREATE VIEW BottomVenuesByWinPercentage AS
WITH VenueWins AS (
    SELECT gv.venId, gv.venName, COUNT(CASE WHEN gr.resUmdScore > gr.resOptScore THEN 1 END) AS Wins,
    COUNT(*) AS TotalGames
    FROM [TerrapinTactics.GameVenue] gv
    JOIN [TerrapinTactics.Play] p ON gv.venId = p.venId
    JOIN [TerrapinTactics.GameResult] gr ON p.gamId = gr.gamId
    GROUP BY
        gv.venId, gv.venName
    HAVING
        COUNT(*) >= 5
)
SELECT TOP 10
    venName, Wins, TotalGames, ROUND((CAST(Wins AS FLOAT) / TotalGames) * 100, 2) AS WinPercentage
FROM VenueWins
ORDER BY WinPercentage ASC, venName; -- Order by ascending win percentage
GO
SELECT * FROM BottomVenuesByWinPercentage;
```

Results Messages

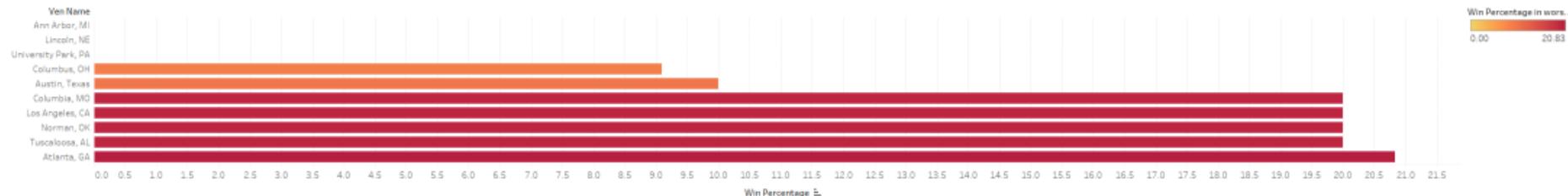
	venName	Wins	TotalGames	WinPercentage
1	Ann Arbor, MI	0	10	0
2	Lincoln, NE	0	7	0
3	University Park, PA	0	5	0
4	Columbus, OH	1	11	9.09
5	Austin, Texas	1	10	10
6	Columbia, MO	1	5	20
7	Lafayette, LA	1	5	20
8	Los Angeles, CA	1	5	20

umd.edu,97... | DESKTOP-L068QVT\sri... | BUDT703_Project_0507_05 | 00:00:00 | 10 rows



Visualization:

Top 10 worst performing venues for UMD



Year on Year Win Percentage of UMD by Season



Seasonal win perce..

6.88 91.67

Year on Year Win Percentage of UMD by Venue Category



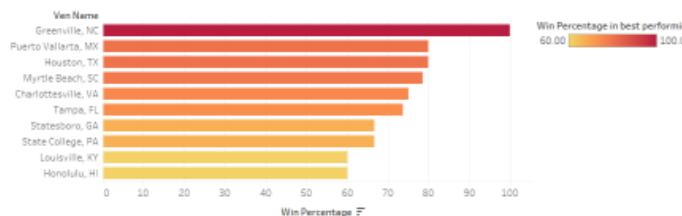
UMD Dominance: Top 10 Teams with the Highest Win Percentage



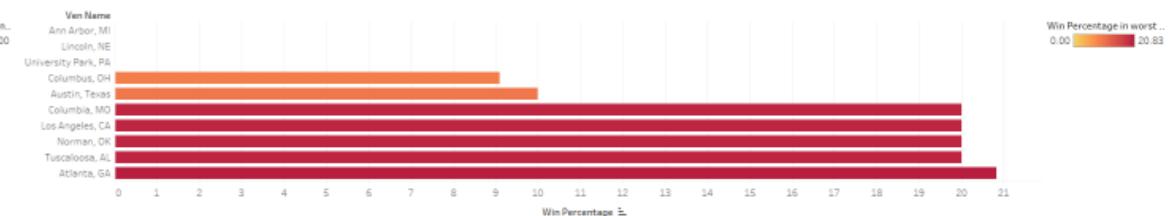
UMD's Toughest Opponents: Top 10 Teams with the Lowest Win Percentage



Top 10 best performing venues for UMD



Top 10 worst performing venues for UMD





Deep Dive

Opponent Analysis: Identify style of play of opponents and form a counter strategy

Seasonal Analysis: Identify reasons of why we do well in spring compared to winter and suggest ways to improve

Identify if there is an advantage to the home team and ways to nullify that vs different opponents

Average score at venue: Pick offensive formation on high scoring venues and defensive on low scoring venues.

Quantification of hitting and fielding ability of individual players.



Thank you

