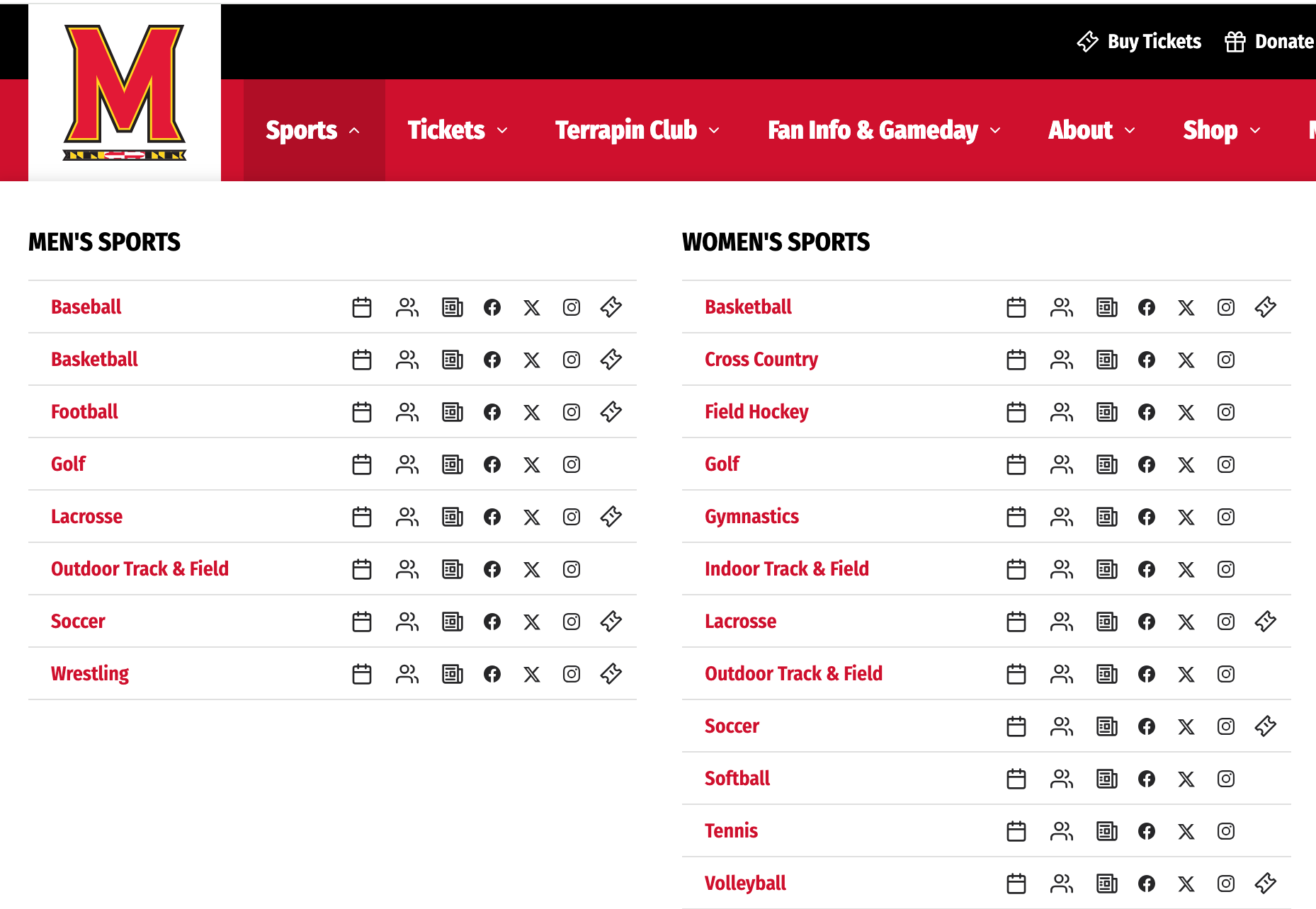
BUDT703: Database Management Systems

**Step 1: Project Proposal**

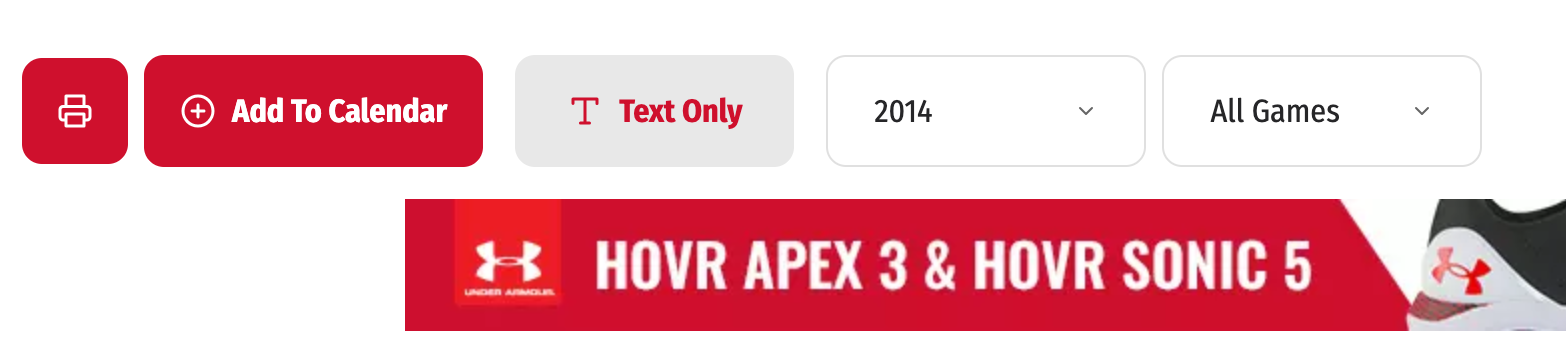
Collected data from University of Maryland’s sports website having information about the schedule of matches which were played from the year 2000 to 2023.

Data Source:

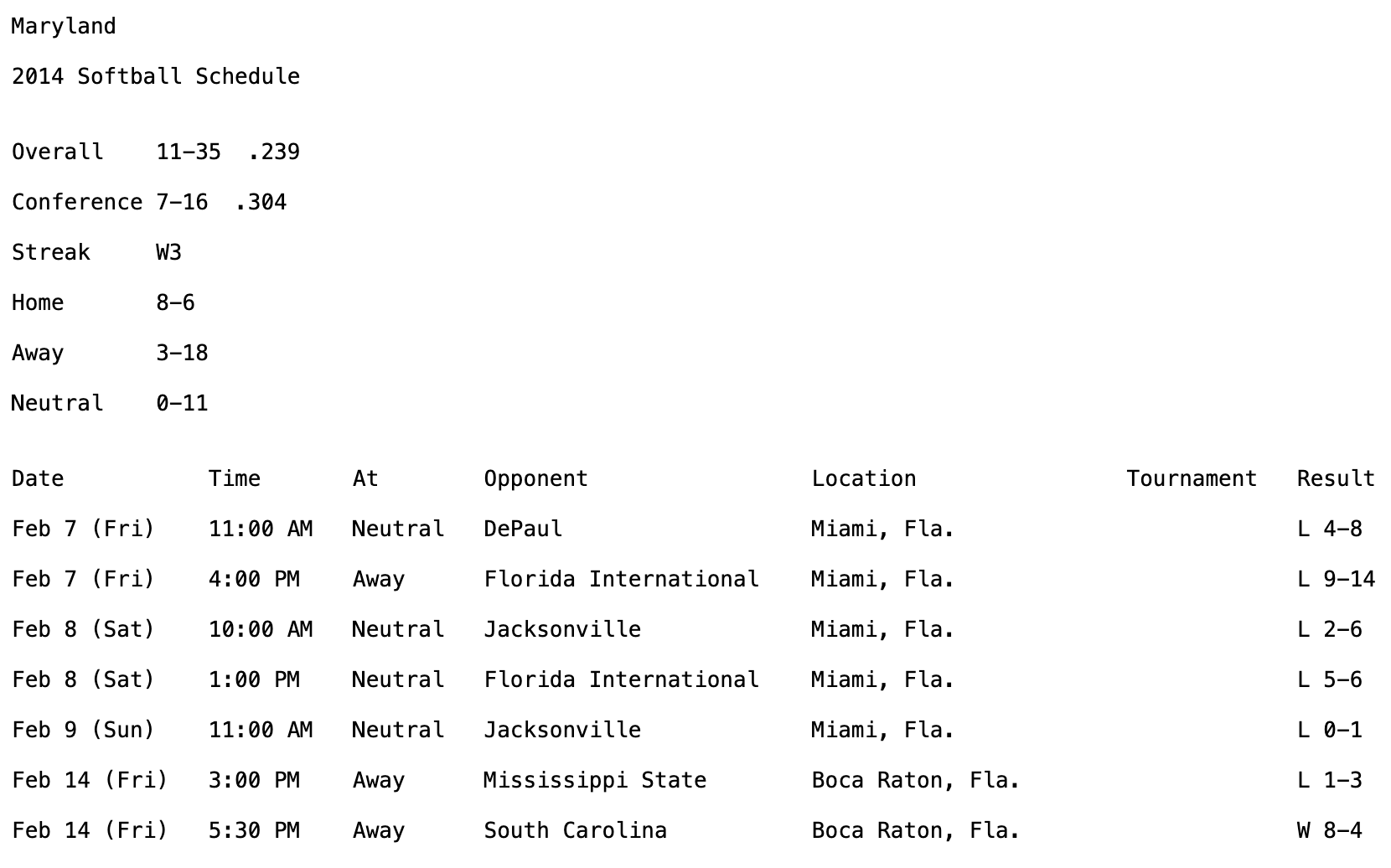
<https://umterps.com/sports/softball/schedule>



Selected the year for which we wanted to analyze the data-



Selected text only in order to view the data in table format-



We considered all the important rows such as date of match, location of match, opponent details and the final match result. Analyzed the data on MS Word and MS excel and then formed the mission statement and mission objectives accordingly.

We identified Entities, Relations, Primary and Foreign keys using which we made our ER diagram and Relational schema.

**MISSION STATEMENT**

As consultants, our mission is to serve as the catalyst for the University of Maryland women’s softball team's transformation, to equip the team with actionable recommendations, enabling them to elevate performance standards and establish a legacy of sporting achievement.

**MISSION OBJECTIVES:**

● OBJECTIVE 1: Analyze the top 10 teams where Maryland has lost against the maximum number of times

● OBJECTIVE 2: Analyze the top 10 away locations where Maryland has won the most number of times

● OBJECTIVE 3: Analyze the top 10 matches where Maryland has won with a shutout

● OBJECTIVE 4: Analyze the top 10 of teams against which Maryland has won the maximum number of times

● OBJECTIVE 5: Analyze the top 10 teams against whom Maryland has won by the highest margin

**RELATIONAL SCHEMAS**

Entities, Attributes and Primary Keys:

Match (**mchId** ,mchAt, mchScrUmd, mchScrOpp,locId)

Date (**datMch** , datDay)

Location (**locId** , locCty ,locState)

Opponent (**oppId** , oppName)

Play (***mchId*** , ***datMch*** , ***oppId***)

**Host: Binary Relationship**

1 location hosts 1 or more matches

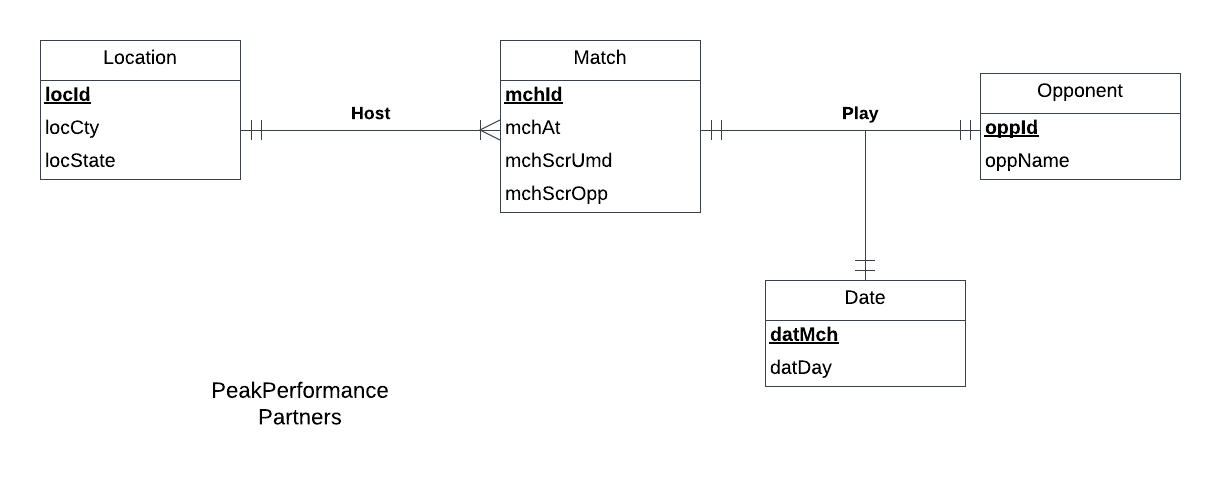
1 match to 1 location

**Play: Ternary Relationship**

1 match and 1 opponent to 1 date

1 match and 1 date to 1 opponent

1 date and 1 opponent to 1 match

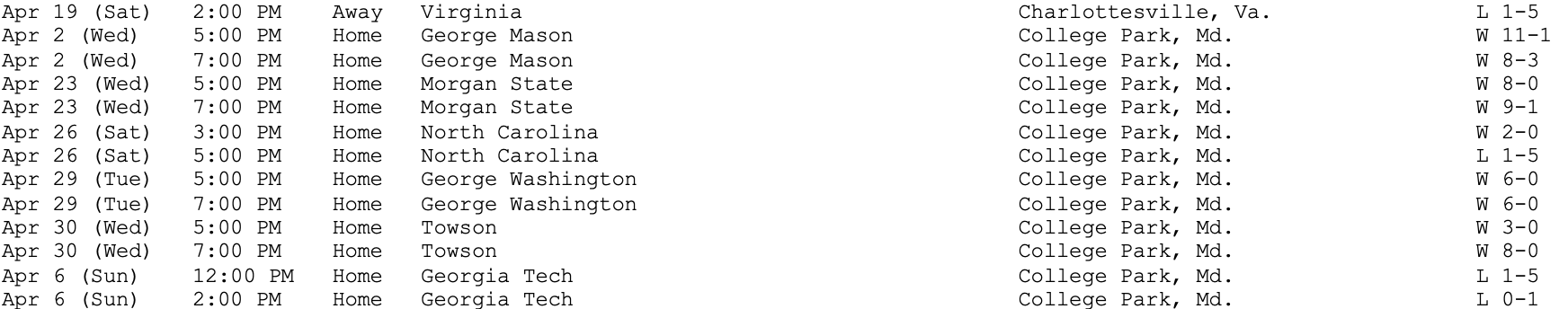


Business rules based on the objectives:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Relation | Foreign Key | Base Relation | Primary Key | Business rule | Constraint:  ON DELETE | Business rule | Constraint:  ON UPDATE |
| Host | locId | Location | locId | R1 | CASCADE | R2 | CASCADE |
| Host | mchId | Match | mchId | R3 | CASCADE | R4 | CASCADE |
| Play | mchId | Match | mchId | R5 | CASCADE | R6 | CASCADE |
| Play | datMch | Date | datMch | R7 | CASCADE | R8 | CASCADE |
| Play | oppId | Opponent | oppId | R9 | CASCADE | R1 | CASCADE |

**Step 2: Data Cleaning**

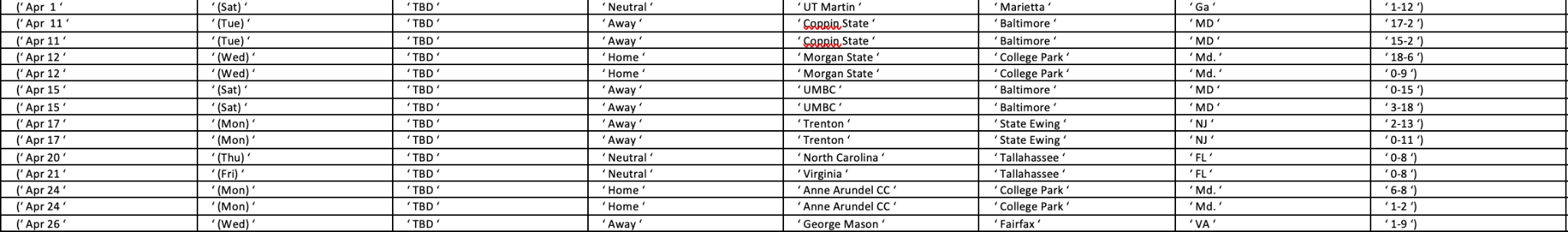
a) Copy pasted the entire columns in Microsoft word in order to remove spaces using find replace function.



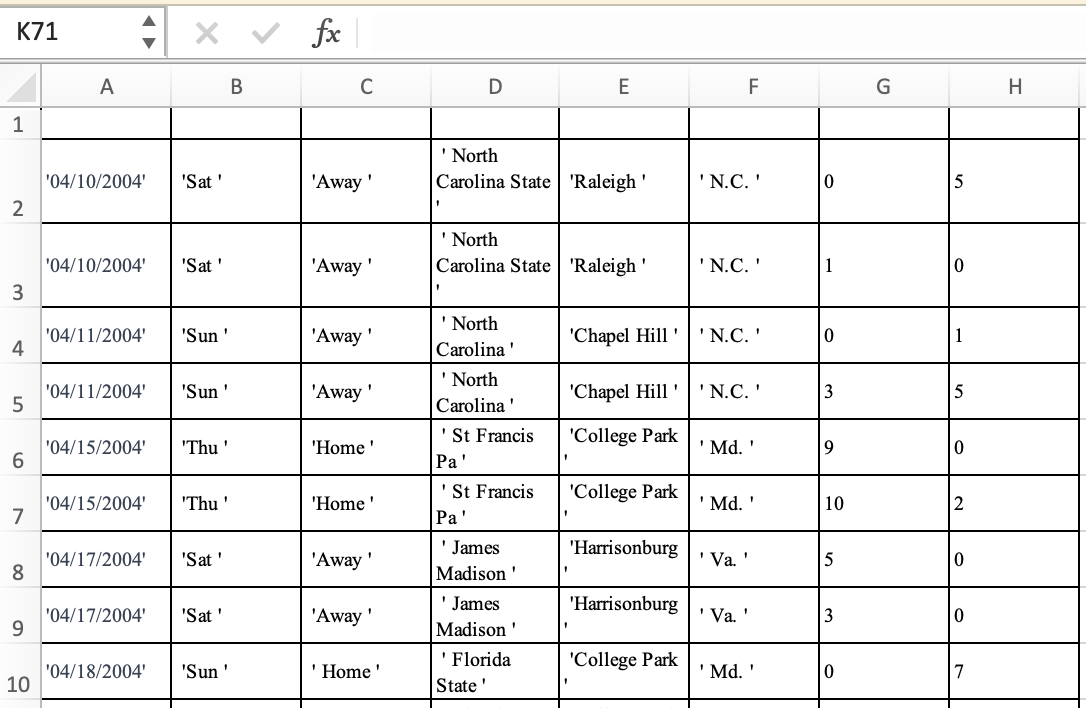
b) Arranged the data into tabular format using text to table function.

c) Removed the ‘Time result’ column and ‘L’ and ‘W’ from the tournament result.

d) Added quotes and brackets.

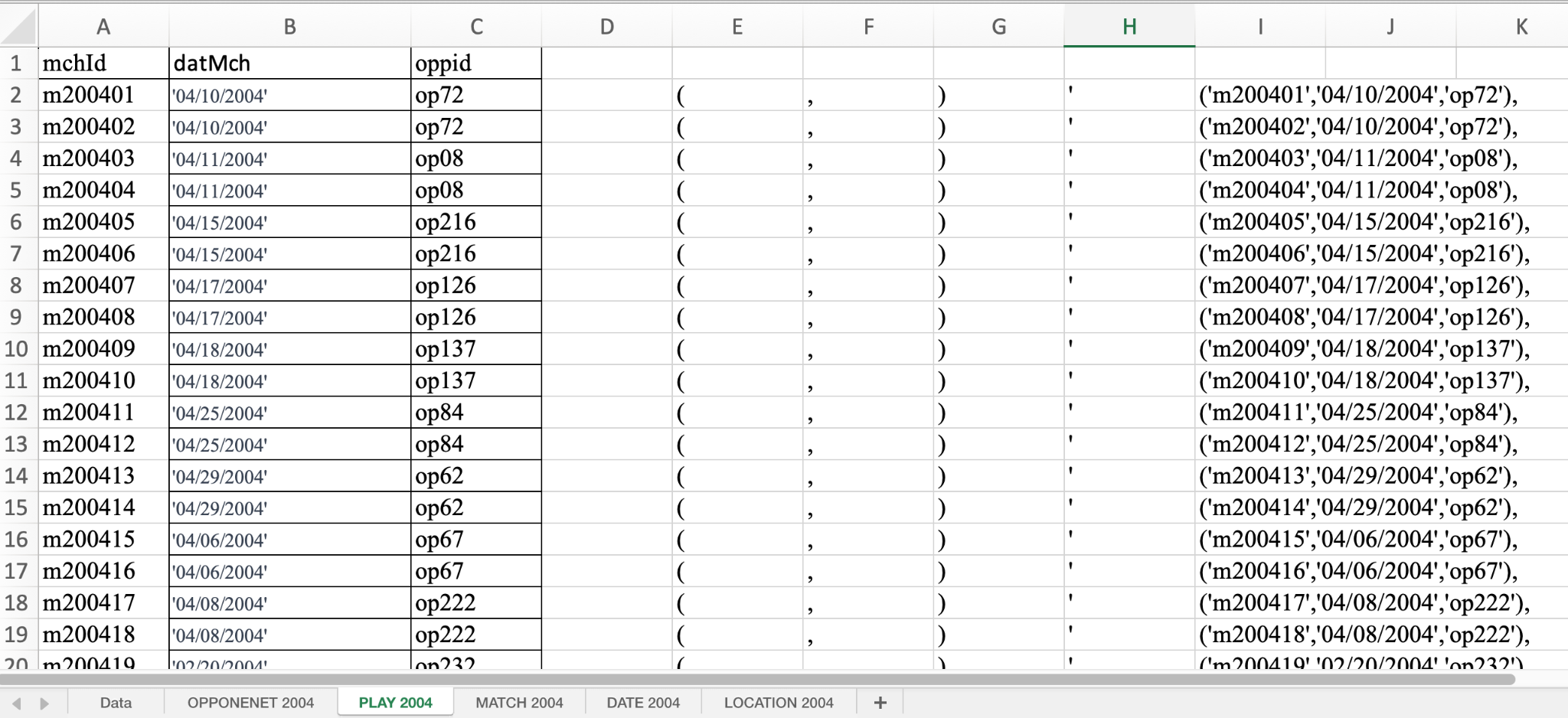


e) Pasted the data into Excel worksheet and removed null and empty cells.



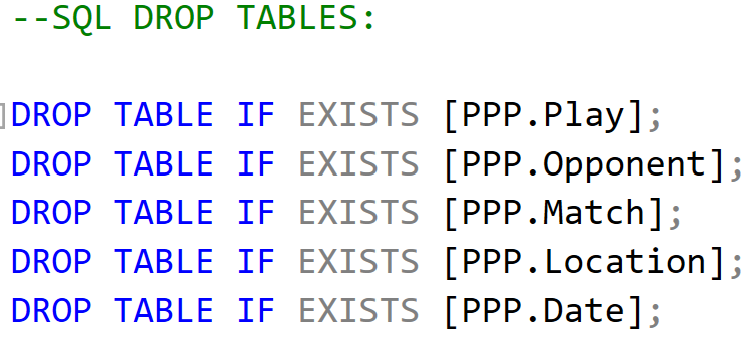
f) Created primary keys for Match as mchId, date as datMch, Location as locId and for Opponent as oppId.

g) Arranged the data as per the entities and created tables.

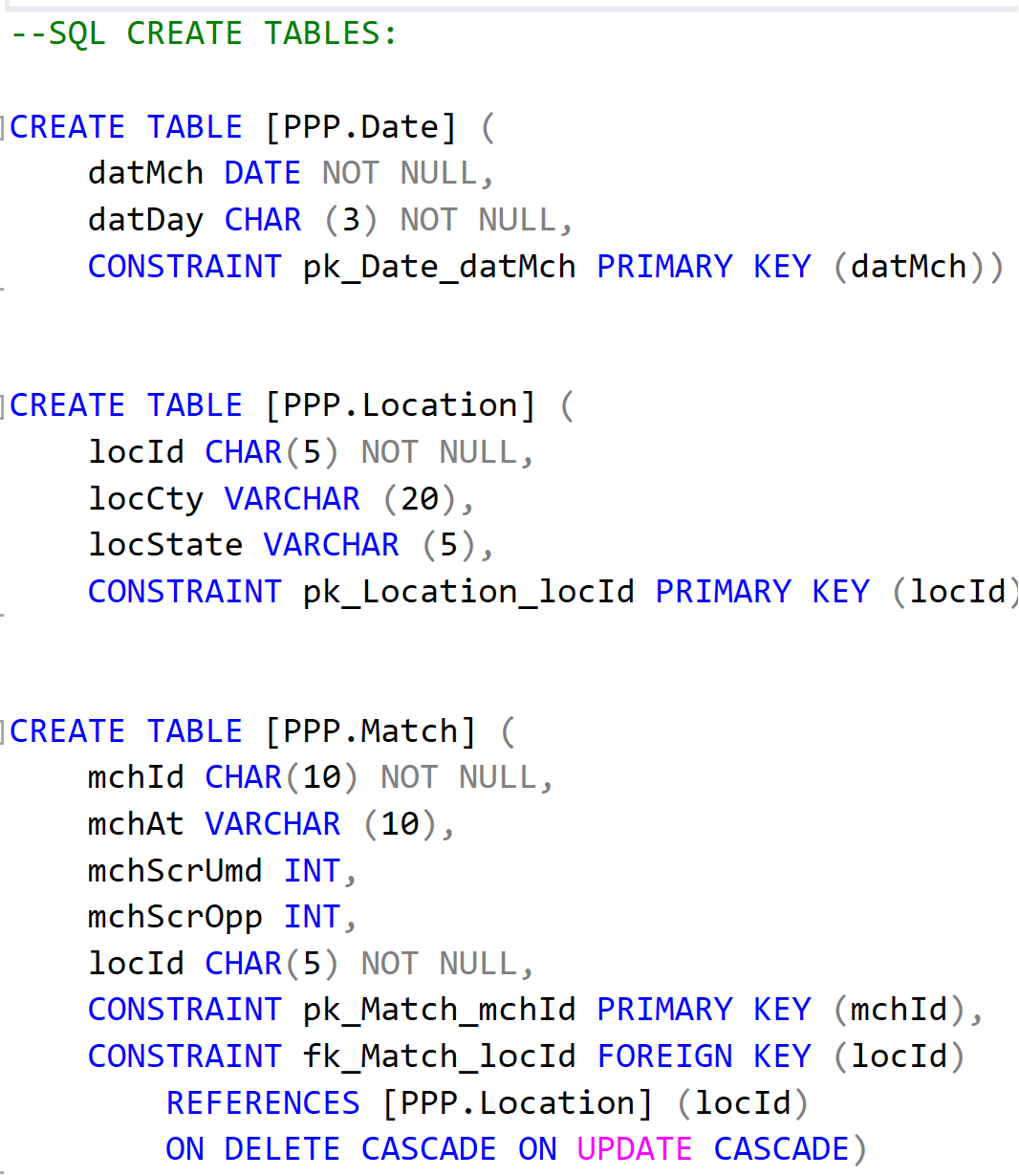


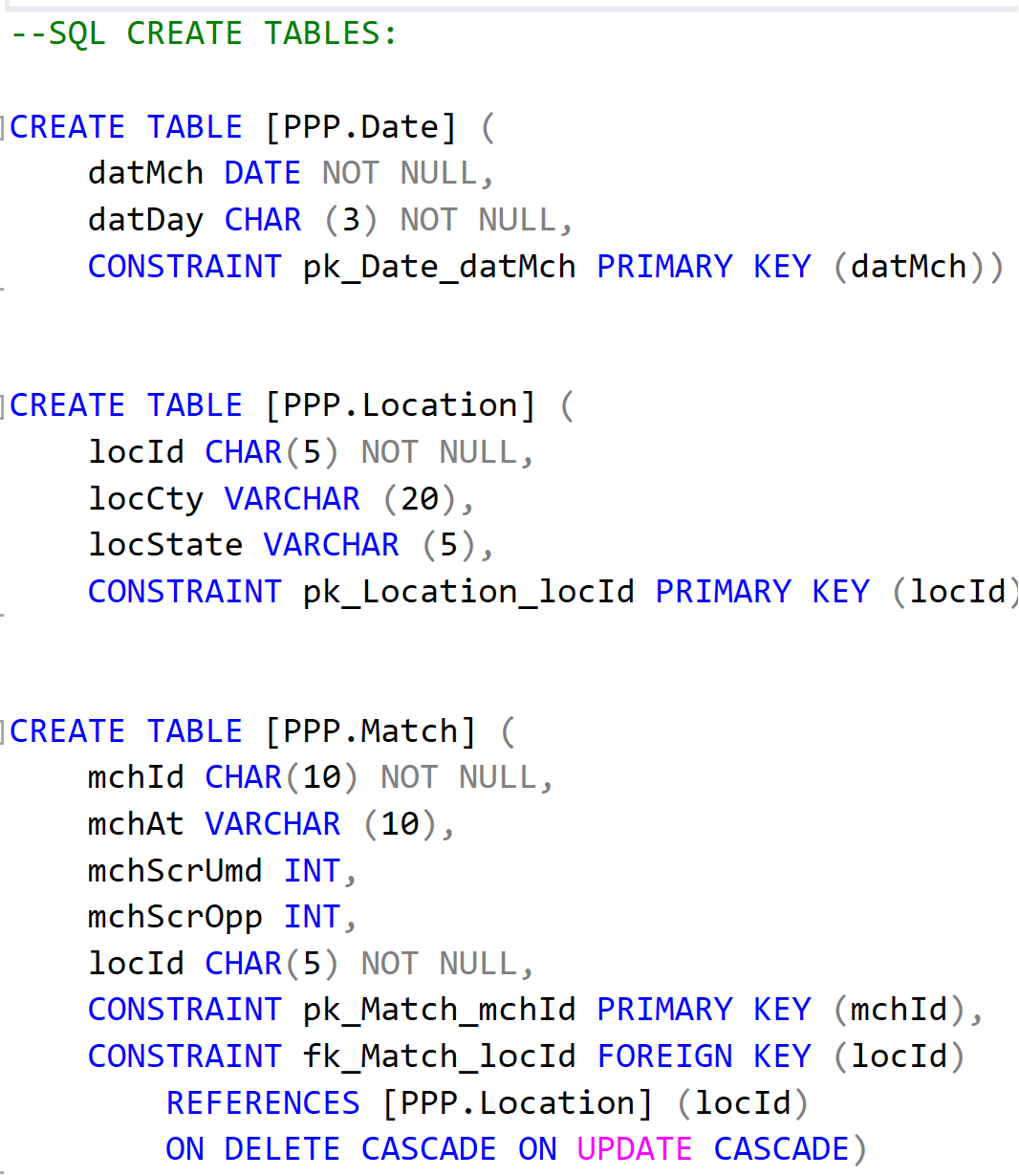
**Step 3: Creating tables in SQL**

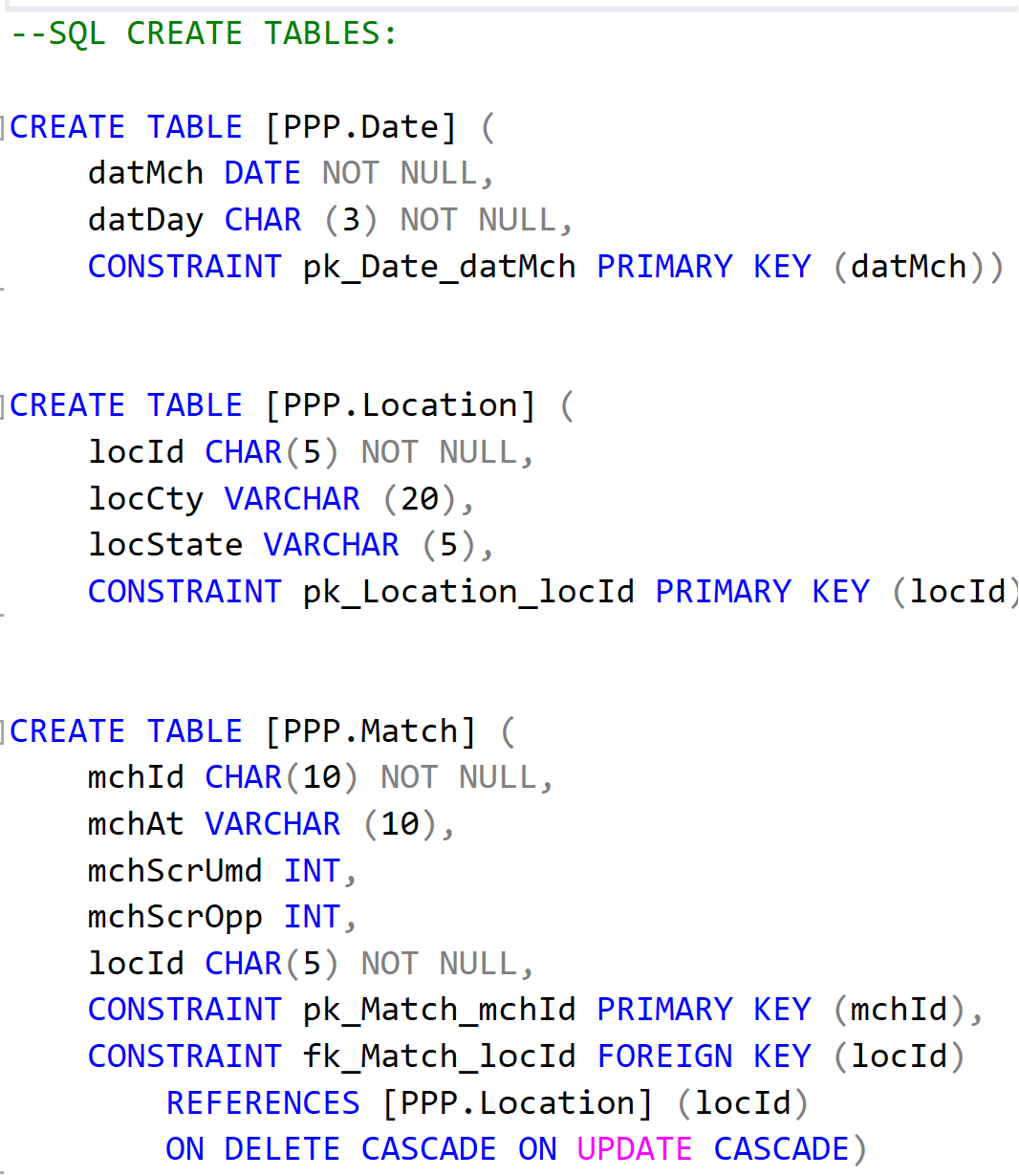
DROP TABLE:

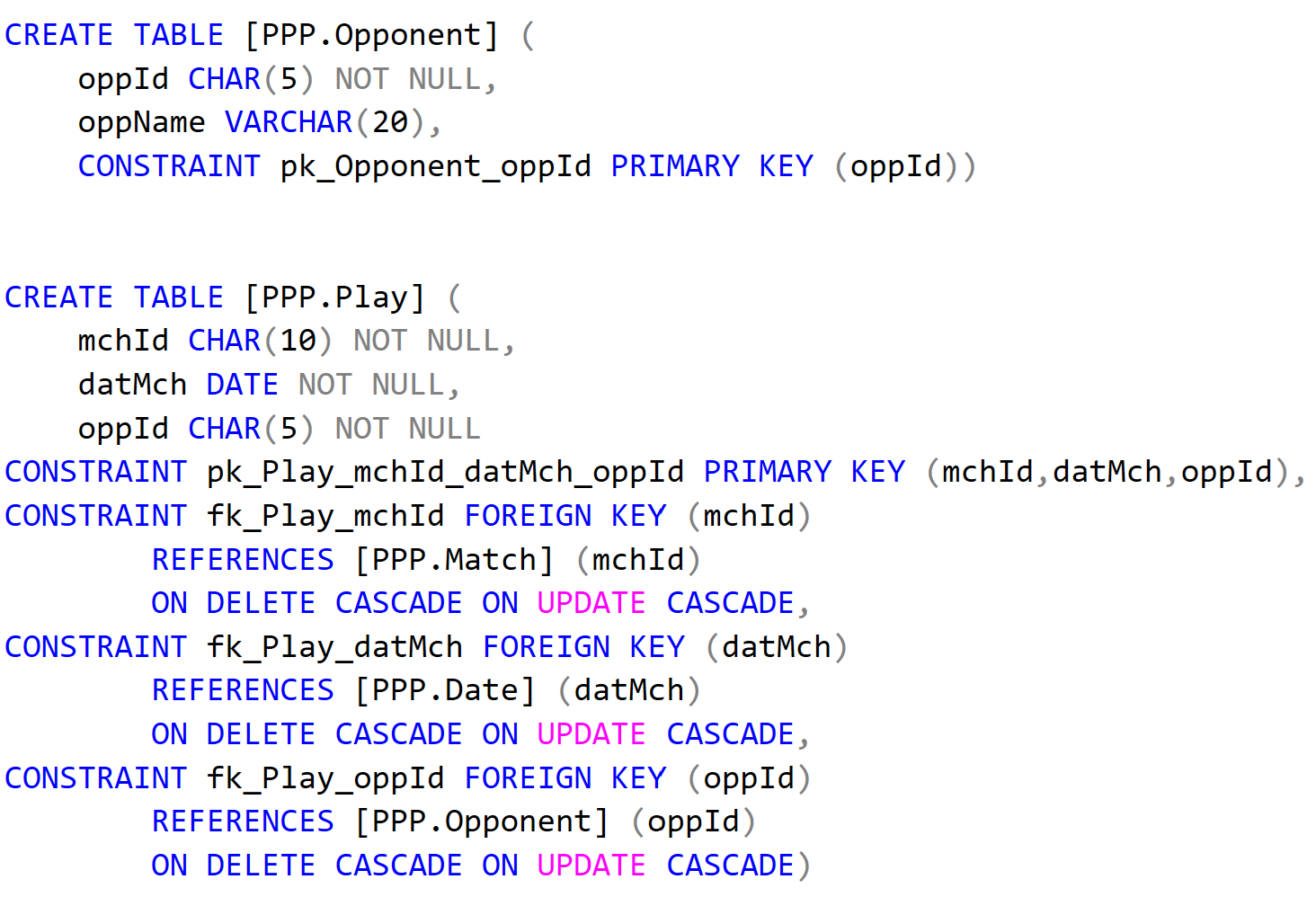


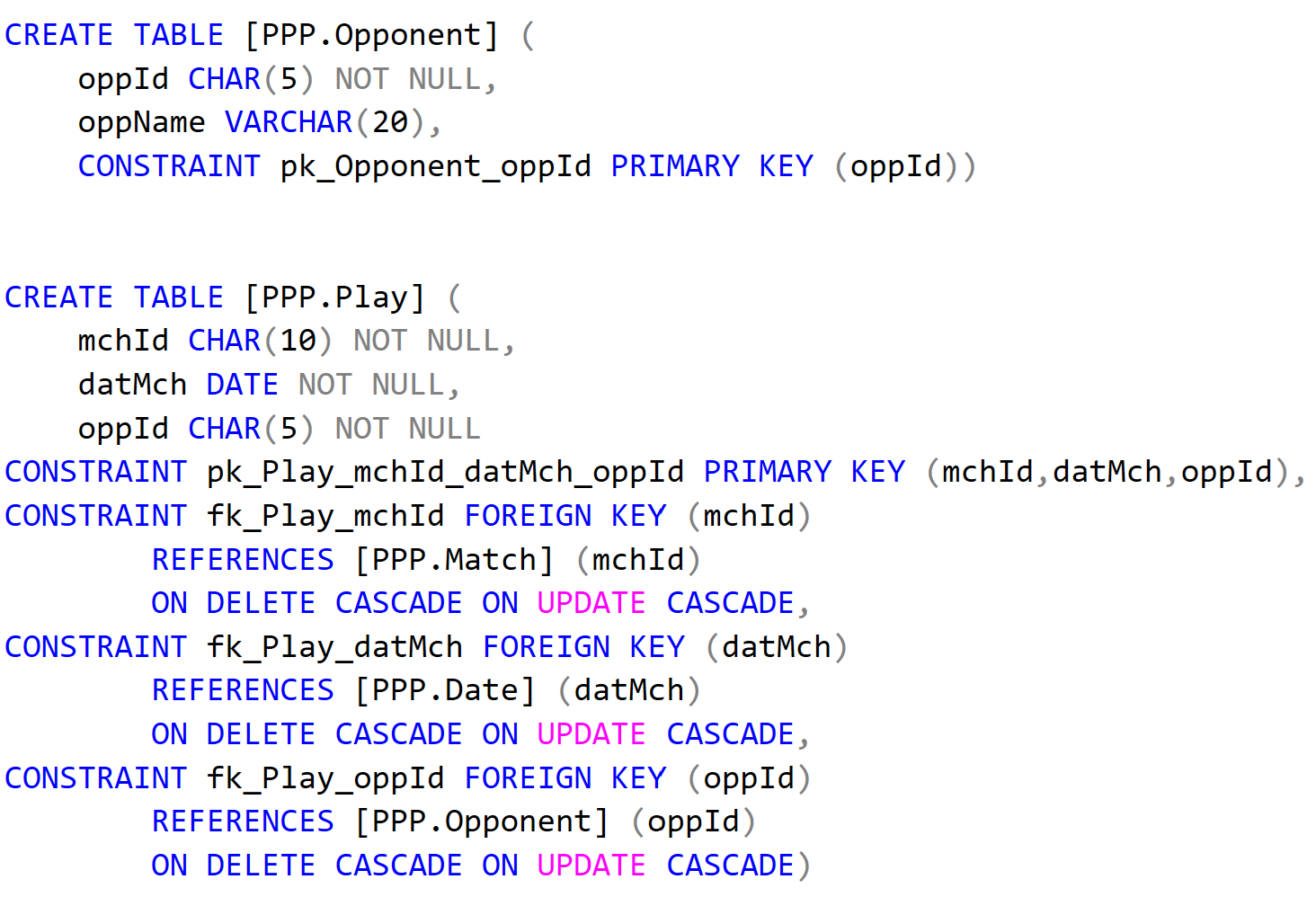
CREATE TABLE:



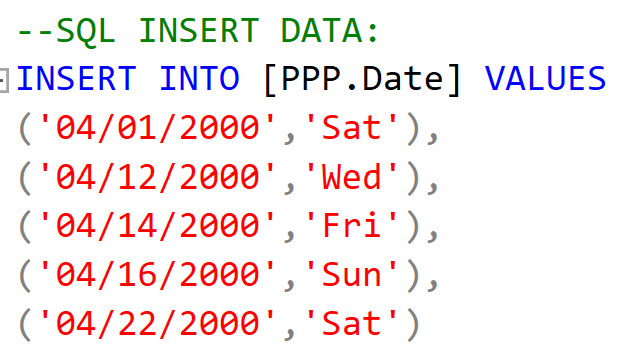


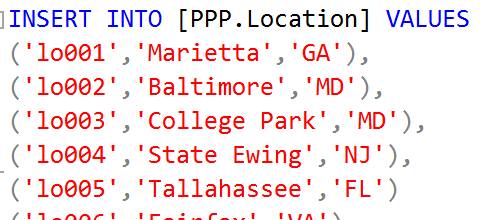


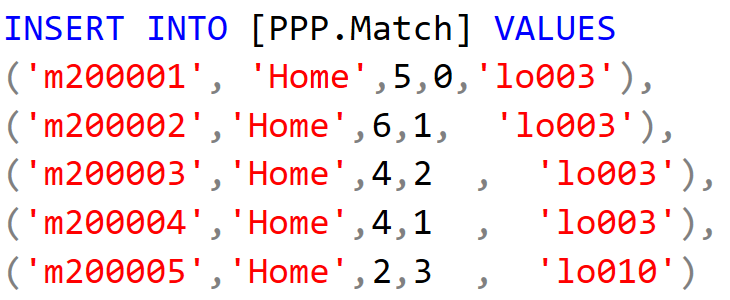


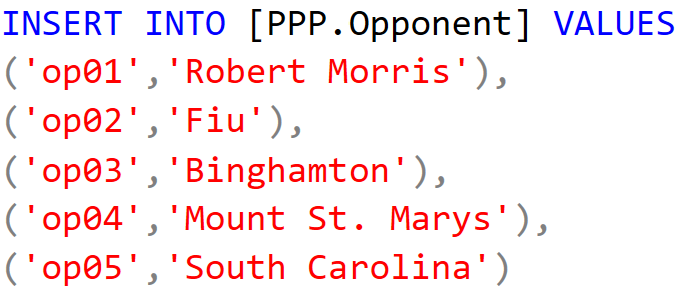


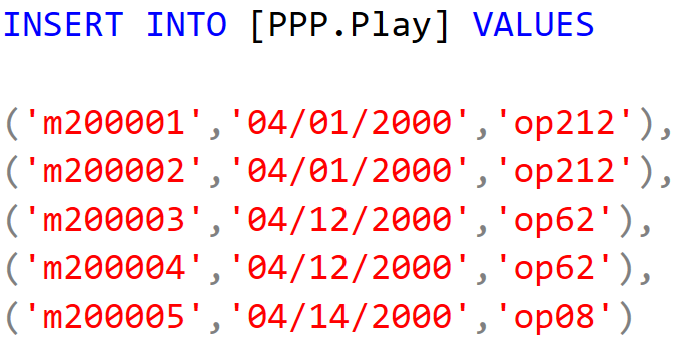
INSERT VALUES:





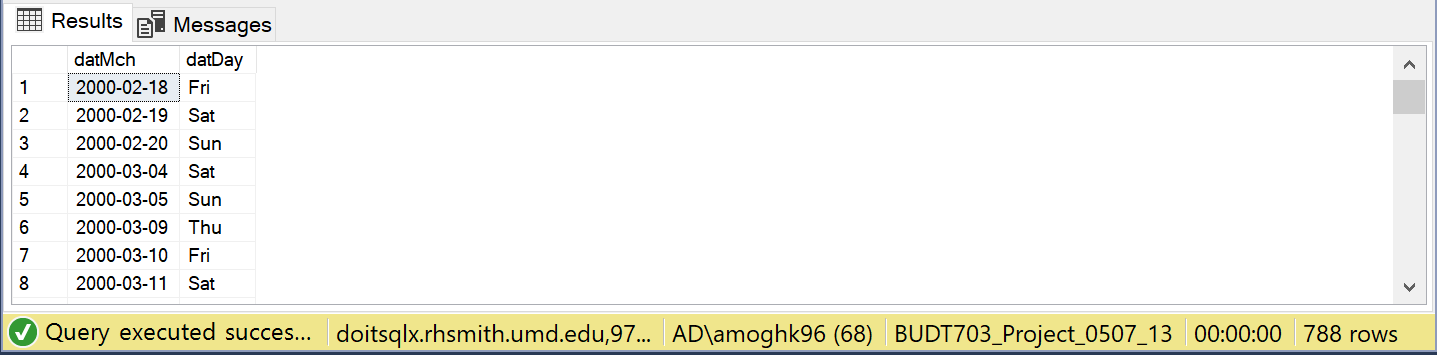




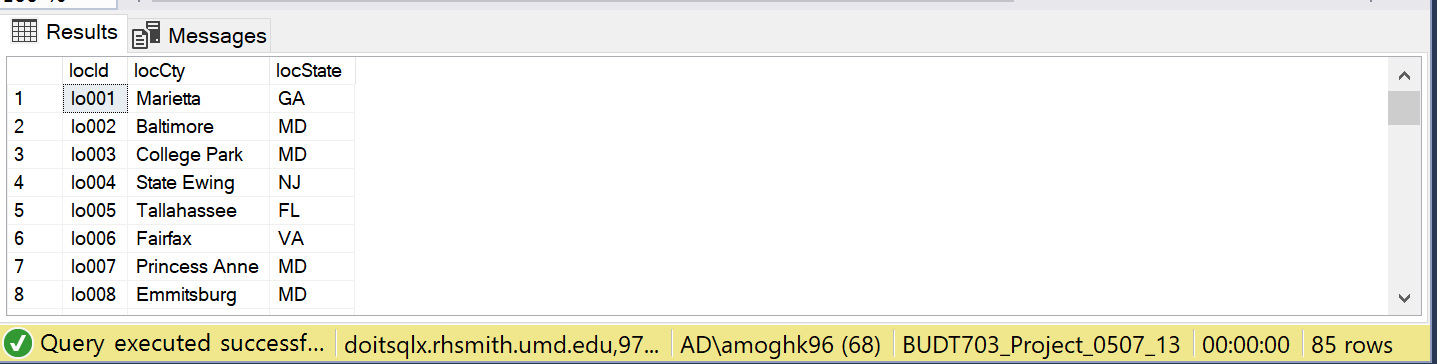


TABLES:

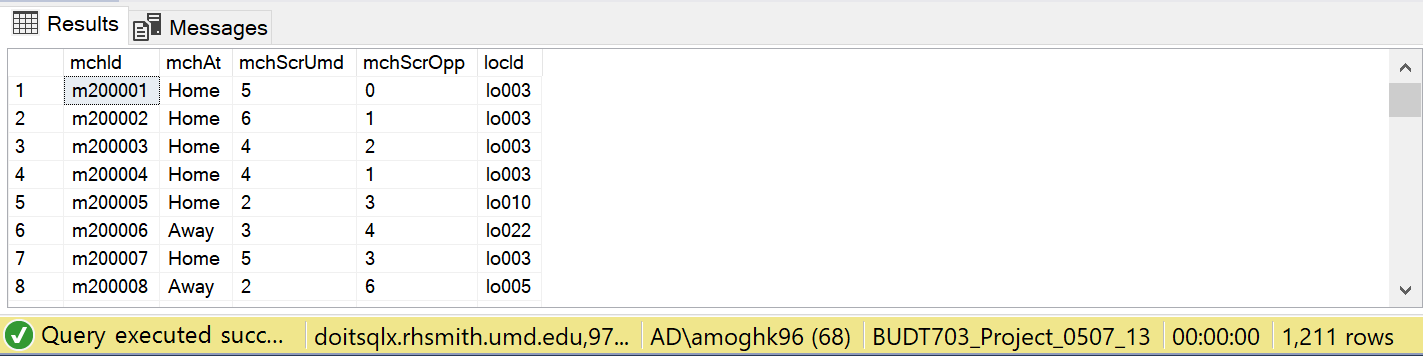
DATE



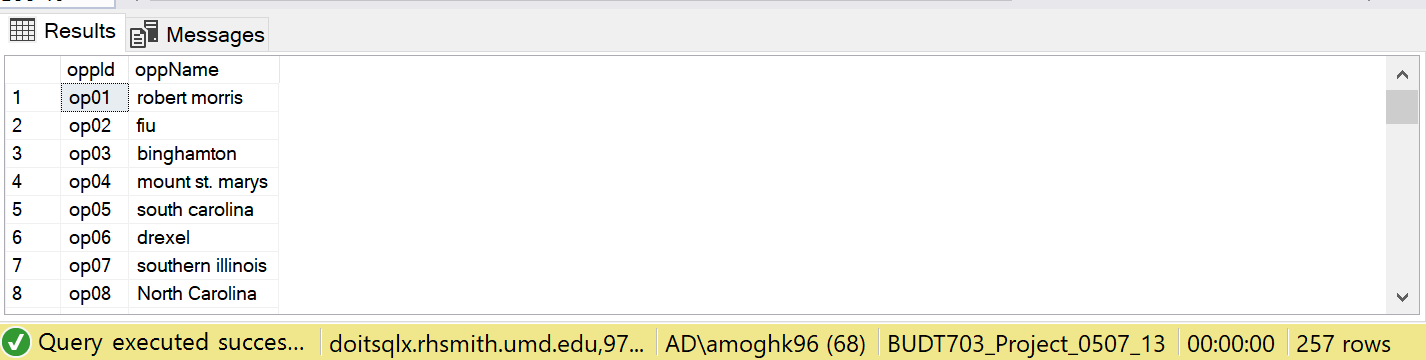
LOCATION



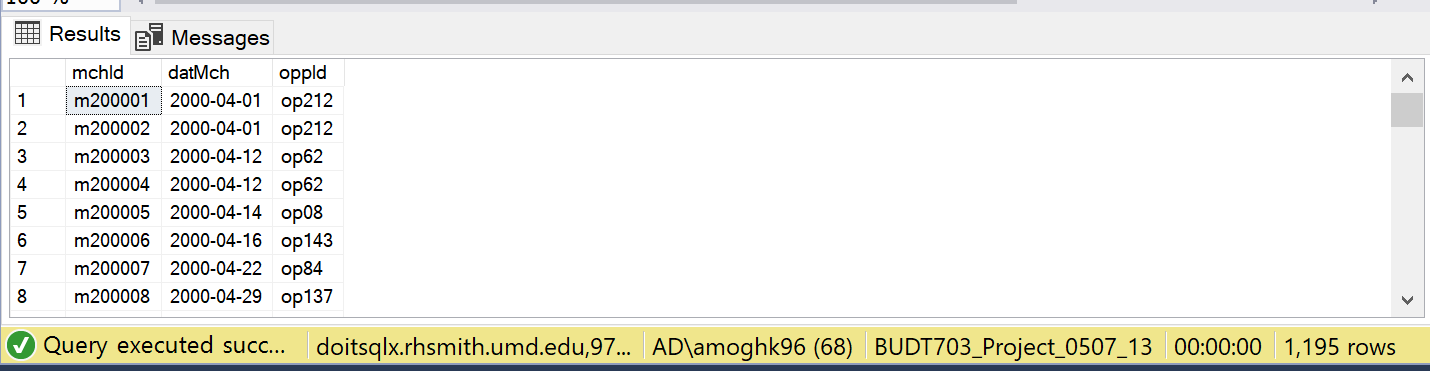
MATCH



OPPONENT



PLAY



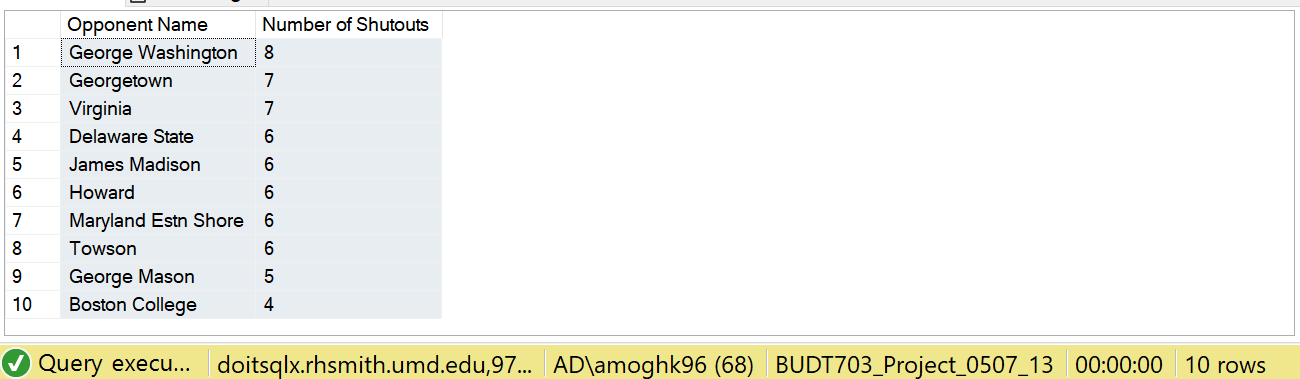
In SQL only 1000 rows can be inserted into one single insert command, due to which we used 2 insert commands for inputting values in Match and Play table.

**Step 4: Executing queries in SQL as per the mission objectives**

Query 1

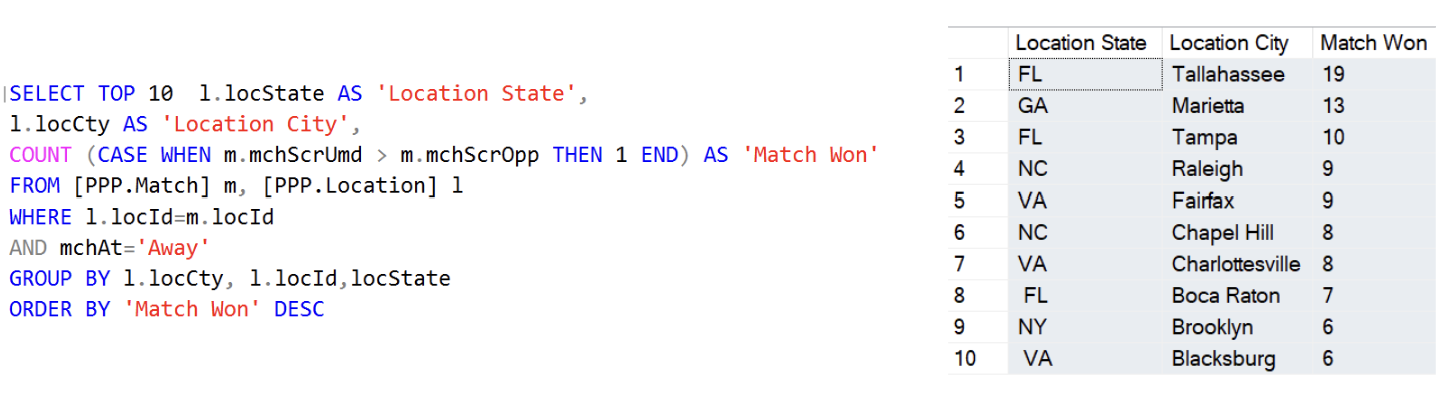
**What are the details of the top 10 teams which Maryland has won against keeping a Shutout?**

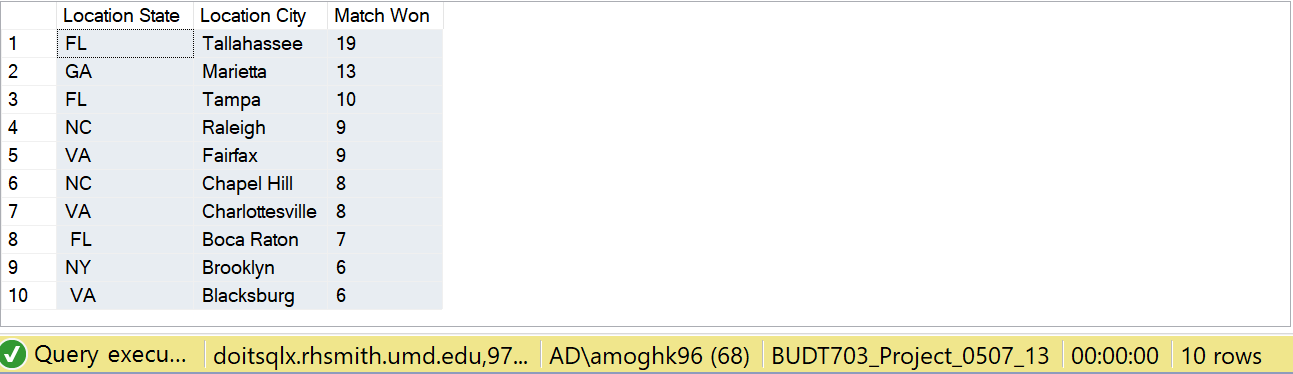




Query 2

**What are the details of the top 10 away locations where Maryland has won the most number of times?**

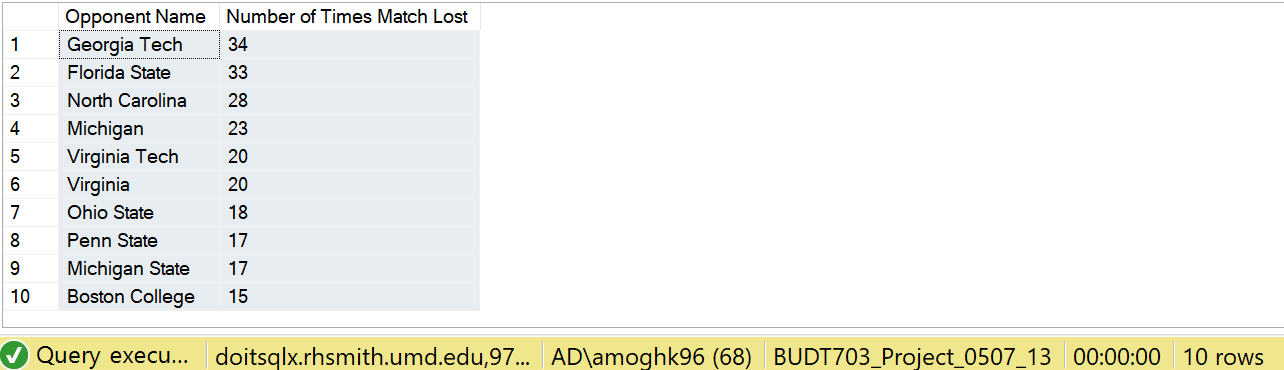


****

**Query 3**

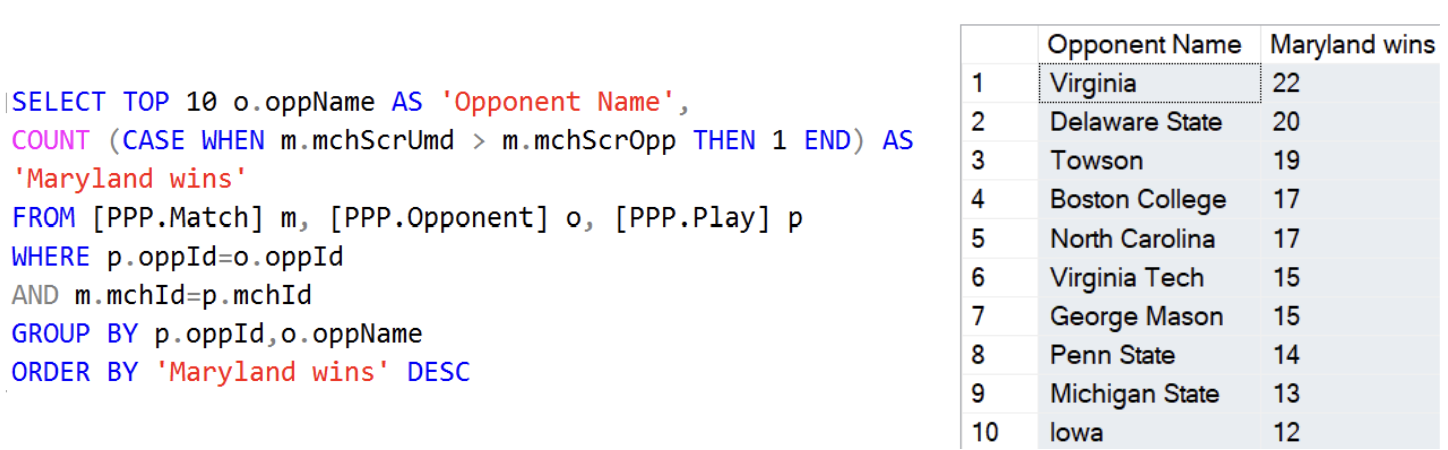
**What are the details of the top 10 teams that Maryland has lost against the maximum number of times?**

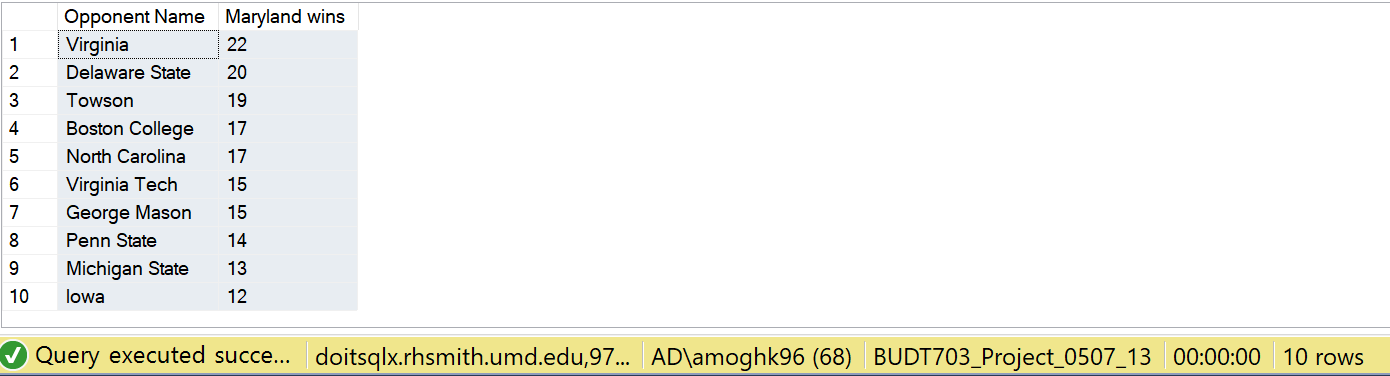
****

****

**Query 4**

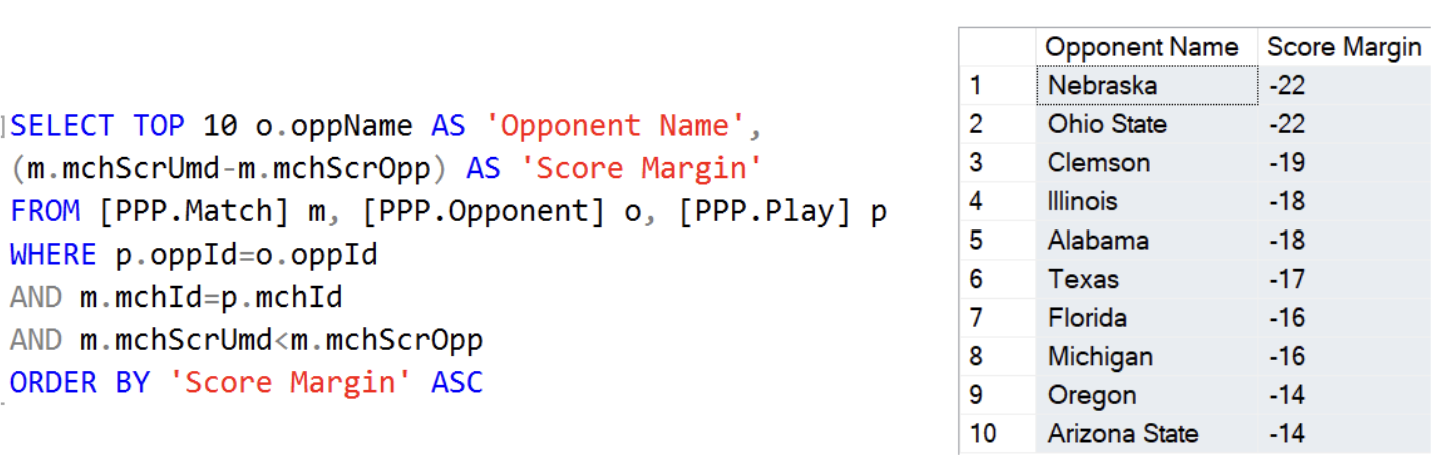
**What are the details of the top 10 teams against which Maryland has won the maximum number of times?**

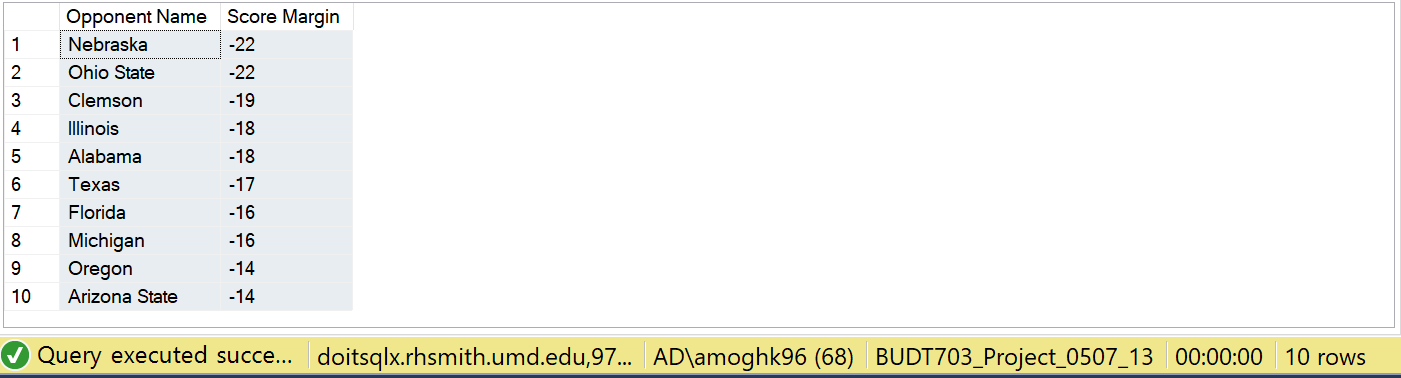
****

****

**Query 5**

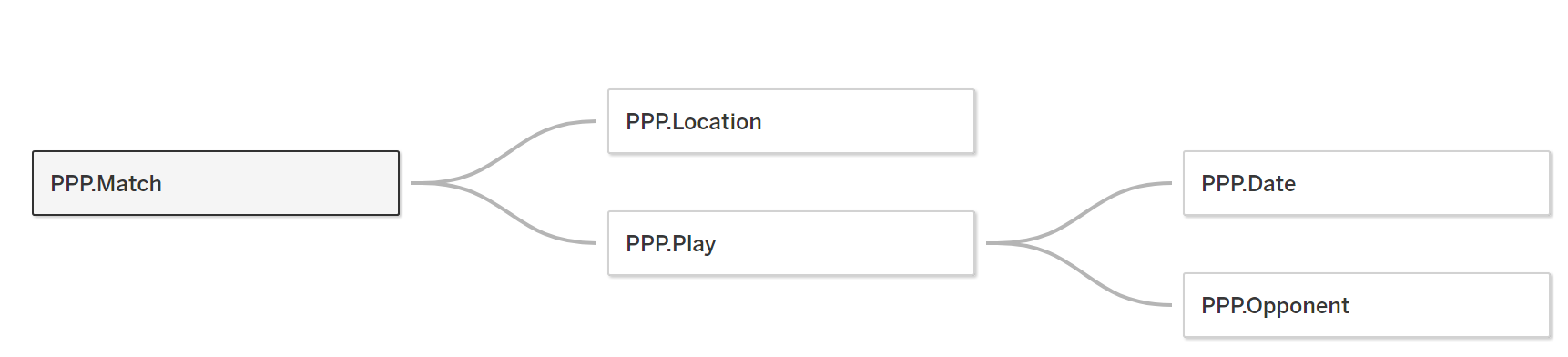
**What are the details of the top 10 teams against which Maryland has lost by the highest margin?**

****

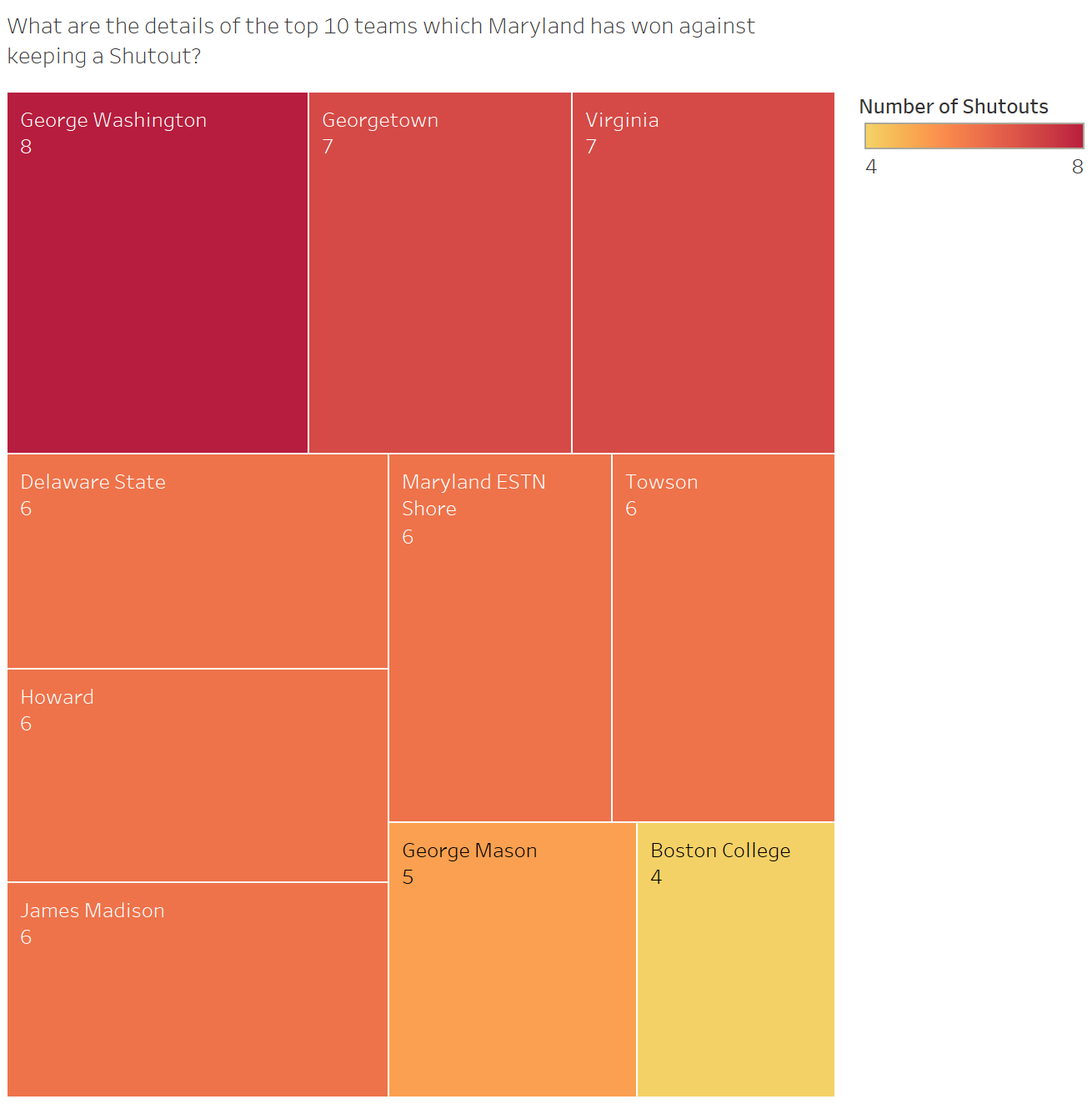
****

**Step 5: Data Visualization using Tableau**

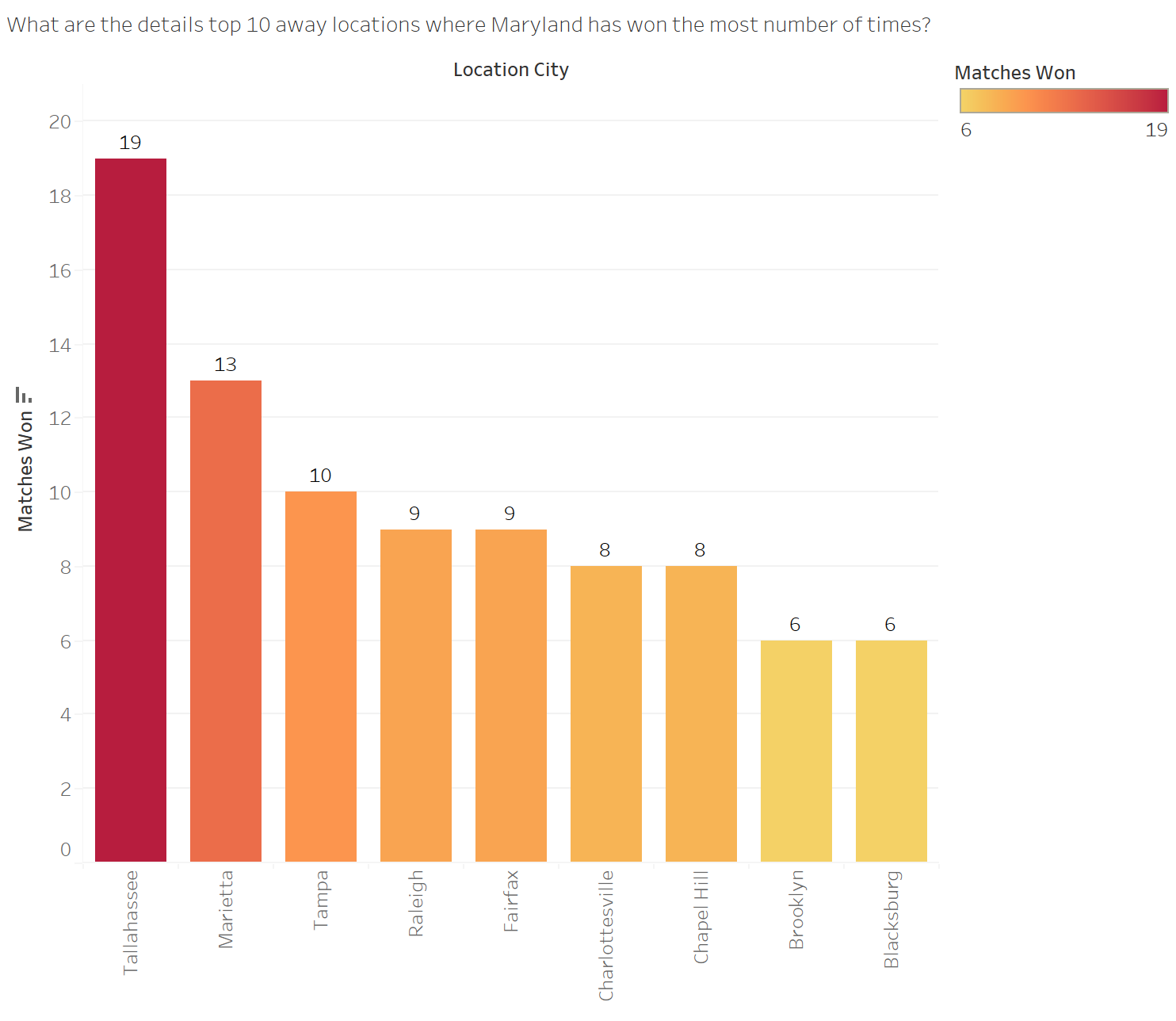
**Created connections between all 4 entities.**

****

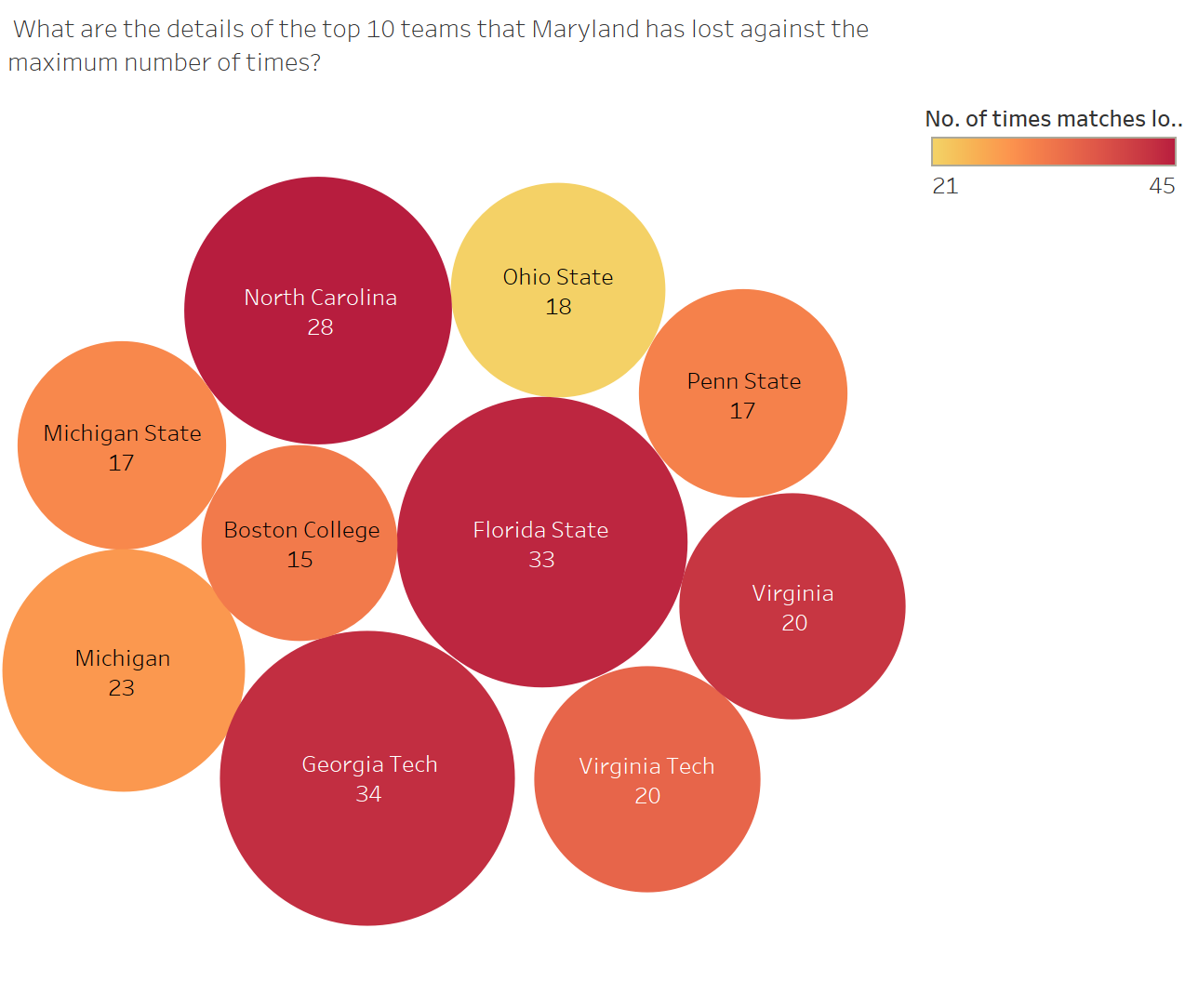
We created a calculated field for calculating Shutout and visualized it using Tree map as seen below.

****

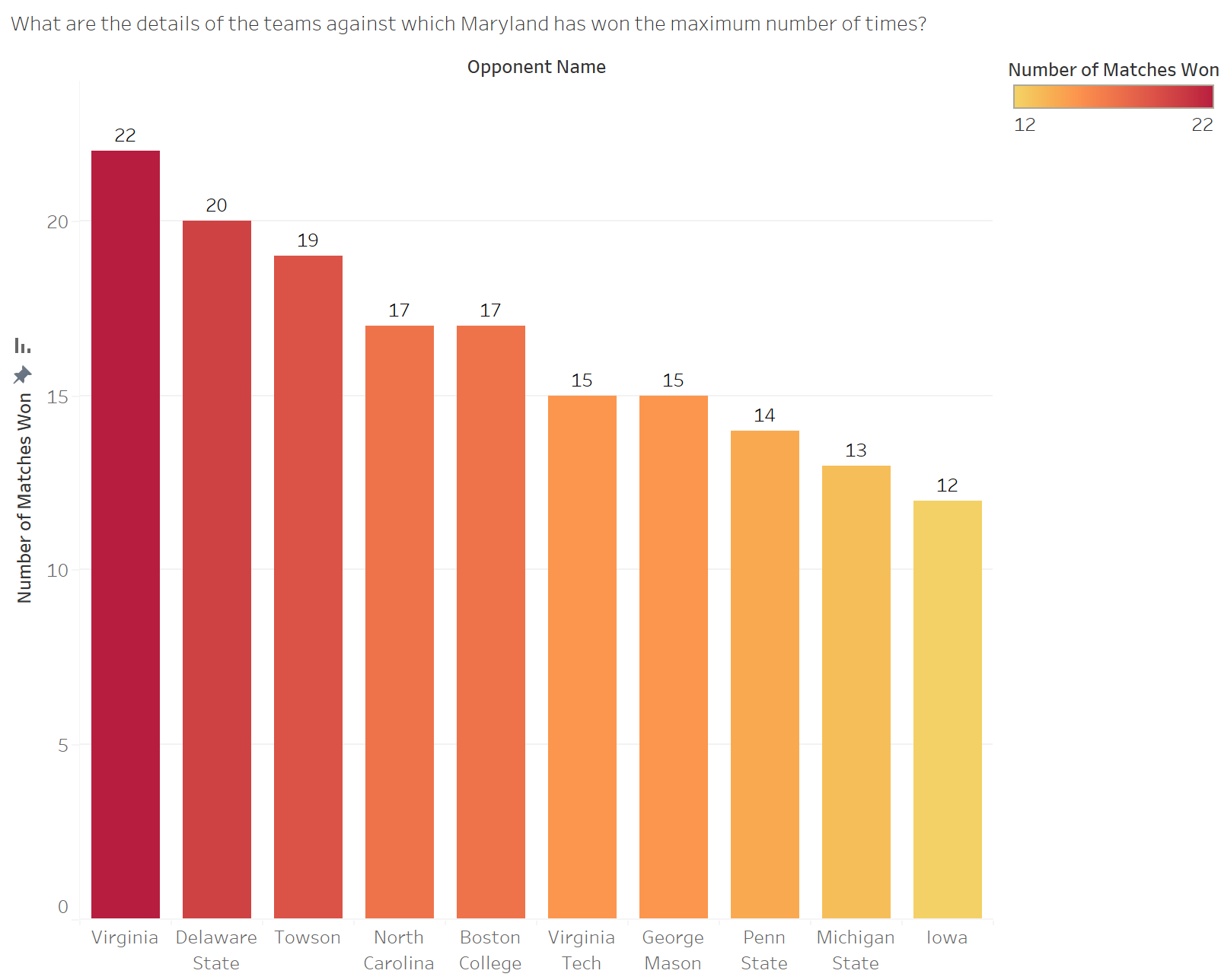
We created a calculated field for when Maryland’s score was greater than Opponent’s score. We also specified the location to be ‘away’ and visualized it using a Bar chart.

****

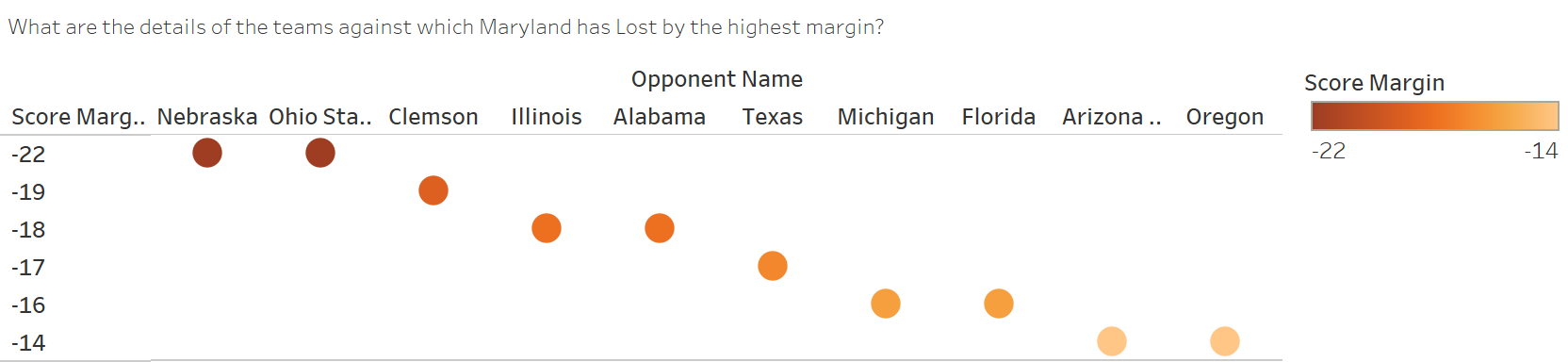
We created a calculated field for when Opponent’s score was greater than Maryland’s score and visualized its total count using a Bubble chart.



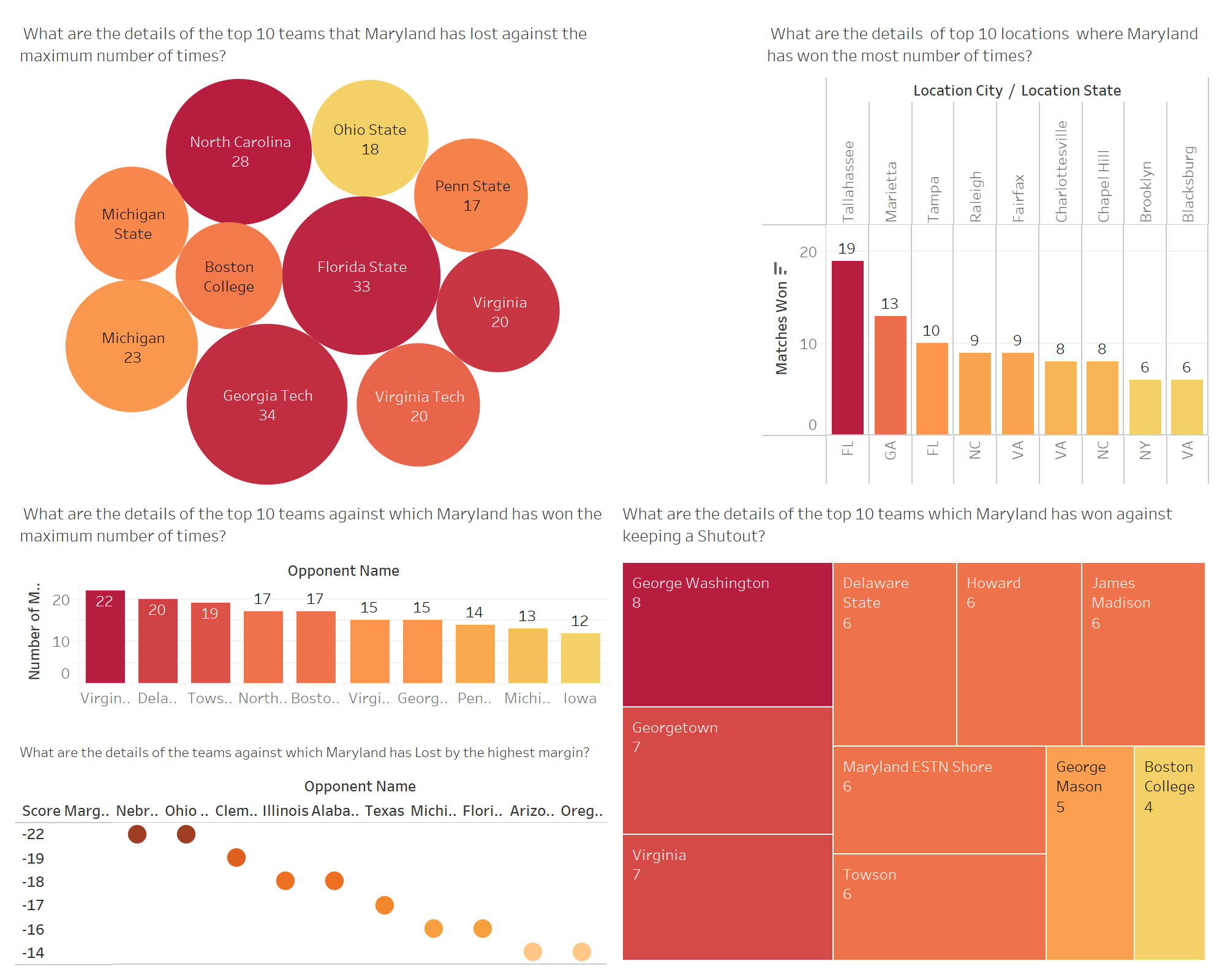
We created a calculated field for when Maryland’s score was greater than Opponent’s score and visualized their maximum wins using a Bar chart.



We created a calculated field for getting the Score Margin (Maryland Score-Opponent Score) and visualized it using a dot chart.



Final Tableau Dashboard-

****

**INSIGHTS:**

* Number of “Shutouts” indicate impressive defensive play from the team, as it reflects a strong performance in preventing the opposing team's hitters from reaching home plate. It signifies effective pitching, solid fielding, and teamwork to keep the opponent from scoring. Maryland’s team showed impressive defensive play against George Washington 8 times, Georgetown 7 times etc.
* In contrast, Maryland’s defeat with the highest margins against teams like Nebraska, Ohio State, Clemson etc. show where their pitching, fielding and overall defense was weak. They should focus on defensive aspects of their game against these opponents.
* Maryland has dominated the most against teams like Virginia, Delaware State, Towson etc. They should be confident when playing against these teams.
* The maximum number of defeats against teams like Georgia Tech, Florida State, North Carolina etc. shed light on whom Maryland should practice more, get more fan support and improve their overall performance.