Tyler Marsh

IST 659

Project Deliverable 2

**Table of Contents**

Project description/summary ……………………………………………………………………………………………….Page #2

Stakeholders .………………………………………………………………………………………………………………………..Page #2

Glossary of Entities and Attributes ………………………………………………………………………………………..Pages #3-5

Data Questions ……………………………………………………………………………………………………………………..Page #5-9

Models (ERD) ………………………………………………………………………………………………………………………..Page #10-11

Raw Data samples …………………………………………………………………………………………………………………Page#12-13

User Interface………………………………………………………………………………………………………………………..Page#14

Reflection and Summary…………………………………………………………………………………………………………Page#15

**Project Description**

There are many different factors that determine if an NFL team will win on any given Sunday. People most commonly think about how different player matchups such as star defenders being able to slow down star offensive players will decide who wins a game and by how much. Obviously other factors such as individual game plans and coaching play a factor as well. These are all well noted and blogged topics when it comes to sports betting and predictions. There are some factors that are not documented or talked about nearly as much within the sports fan and betting communities. Some fans believe that teams traveling a great distance are at a disadvantage such a team from California flying to Pennsylvania to play a game and vice versa. Other narratives that are not as widely documented include whether or not rivals play each other closer compared to other teams, whether or not Vegas’s betting lines may already be adjusting the spread and Over/Under for changing conditions such as playing in a dome versus in a cold climate or traveling over a significant longitude and or latitude. This project aims to investigate how these factors may influence how teams perform and how they may fluence what the proposed spreads and Over/Unders are.

Stake Holders

1. This research will help bloggers write their narratives.
2. It will help online forums that enjoy betting on NFL games consider the results or different variables when making a bet.
3. This project will entertain people who read the blogs that this information could appear on.
4. It could help broadcasters talk about these differences by providing an analysis rather than anecdotal evidence.

**Glossary of Entities and Attributes**

Entities & Their Attributes

Event- an event is a game sanctioned by the NFL

Attributes:

1. StadiumID -this is a foreign key that represents which stadium an event will be held at.
2. ScheduleDate -this is the date that the game is played
3. ScheduleWeek – this is the week of NFL games that the event is held on. They ascend in order so

week 1 is the first official game of the NFL season.

1. TeamHome – represents the team playing at their home stadium during the event.
2. TeamAway – represents the team playing at an opposing teams home stadium during the event.
3. SeasonYear – represents the year in which the event is held.

Stadium

Attributes:

1. Name – represents the stadium’s name.
2. Elevation – represents how far above or below sea level the stadium is in feet.
3. Latitude – represents the degrees of latitude the stadium’s location possesses.
4. Longitude – represents the degrees of longitude the stadium’s location possesses.
5. StadiumWeatherType – represents the climate of a stadium. The acceptable values are dome,

which is an indoor stadium with an internal temperature of 72 degrees, cold, moderate

and warm.

Teams

Attributes

1. TeamName – represents the names of the 32 NFL teams.
2. DivisionName – this is a foreign key that represents the division the team is in.
3. StadiumID – this is a foreign key that represents the stadium that a team plays their home

games in.

Division

Attributes

1. DivisionName – contains the names of the divisions in the NFL. There are listed below.
2. AFCNORTH – represents a division in the AFC (American Football Conference) that contains the following teams: Cleveland Browns, Pittsburgh Steelers, Baltimore Ravens, and Cincinnati Bengals.
3. AFCEAST– represents a division in the AFC (American Football Conference) that contains the

following teams: New England Patriots, Miami Dolphins, Buffalo Bills, and New York Jets.

1. AFCSOUTH - represents a division in the AFC (American Football Conference) that contains the

following teams: Indianapolis Colts, Tennessee Titans, Houston Texans, and Jacksonville Jaguars.

1. AFCWEST - represents a division in the AFC (American Football Conference) that contains the

following teams: Denver Broncos, Los Angeles Chargers, Oakland Raiders, and Kansas City Chiefs.

1. NFCNORTH - represents a division in the NFC (National Football Conference) that contains the

following teams: Green Bay Packers, Detroit Lions, Minnesota Vikings, and Chicago Bears.

1. NFCEAST – represents a division in the NFC (National Football Conference) that contains the

following teams: Philadelphia Eagles, Dallas Cowboys, New York Giants, and Washington Redskins.

1. NFCSOUTH – represents a division in the NFC (National Football Conference) that contains the

following teams: Carolina Panthers, Tampa Bay Buccaneers, Atlanta Falcons, and New Orleans Saints.

1. NFCWEST – represents a division in the NFC (National Football Conference) that contains the

following teams: Los Angeles Rams, Seattle Seahawks, San Francisco 49ers, and Arizona Cardinals.

1. ConferenceName – a foreign key referring to what conference a division is in.

Conference

Attributes

1. Conference Name – name of the different conferences in the NFL. They are listed below:
2. AFC – stands for American Football Conference holds half the teams in the NFL.
3. NFC – stands for National Football Conference holds half the teams in the NFL.

Game Details

Attributes

1. HomeScore – represents the score of the home team in an event.
2. AwayScore- represents the score of the away team in an event.

the stadium the teams are playing in.

1. Temperature – represents the degrees in Fahrenheit of the stadium when the event is played.
2. WindMPH – represents the wind speed in miles per hour in the stadium.
3. Over/Under represents betting odds on how many total points will be scored in an event. If an

Over/Under is high experts from Vegas expect many points to be scored if it is low experts expect few points to be scored.

1. FavTeamSpread – represents betting odds on how many points a team must win by in order for

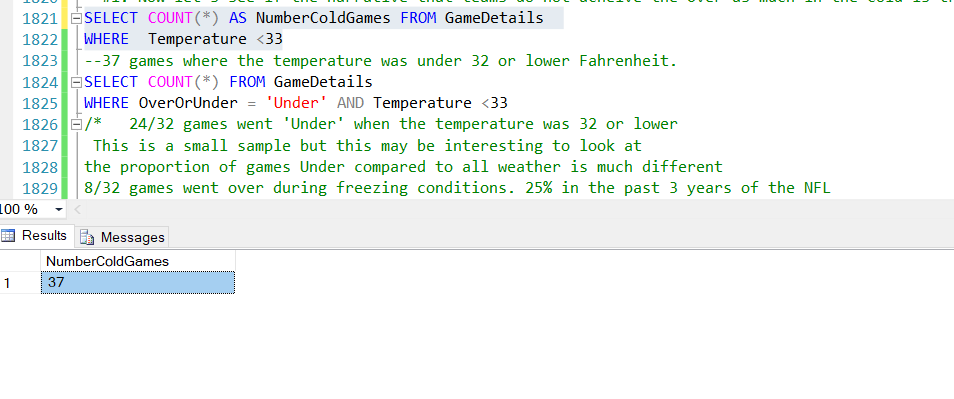
a bettor to win their wager. For example, if PIT is the favored team by -7 they would need to win by 8 or more points in order for a bettor to win their wager if they chose PIT.

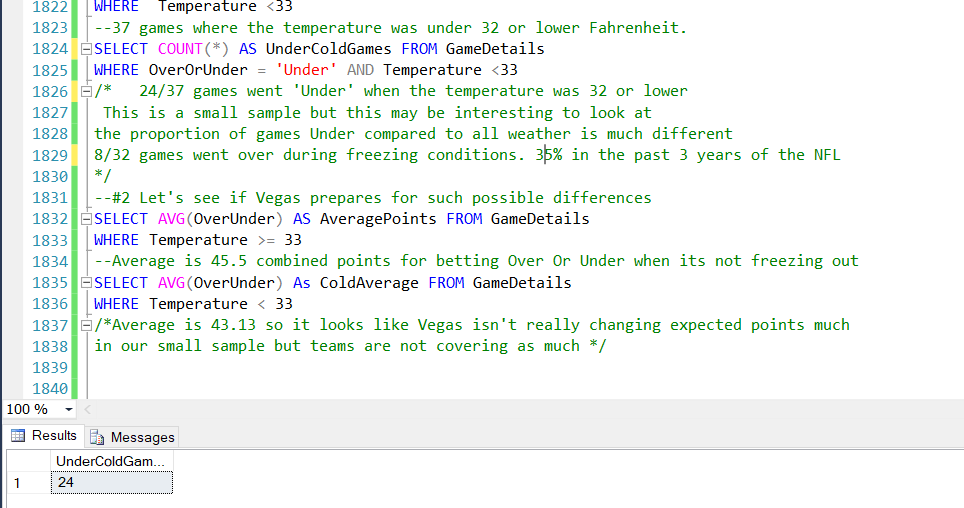
1. DidCover – represents whether the game covered the spread or not.
2. OverOrUnder – represents whether the point total was under or over the proposed betting line.

**Data Questions**

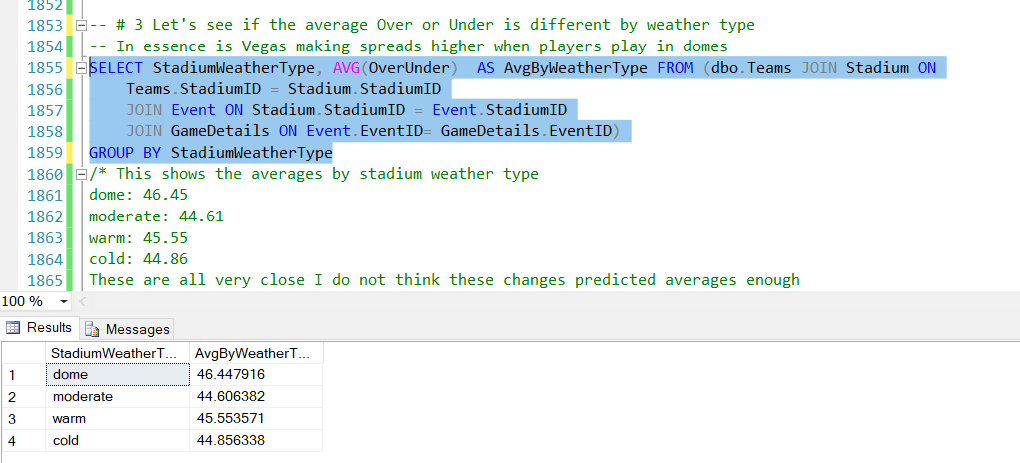
These are the fundamentals questions and concepts this project aims to answer and explore.

1. Is the narrative NFL games played in the cold have fewer points scored and thus fewer games that achieve the Over in the OverOrUnder category?
2. Does it seem like Vegas is putting out lower OverOrUnder totals for games that are played in cold weather?

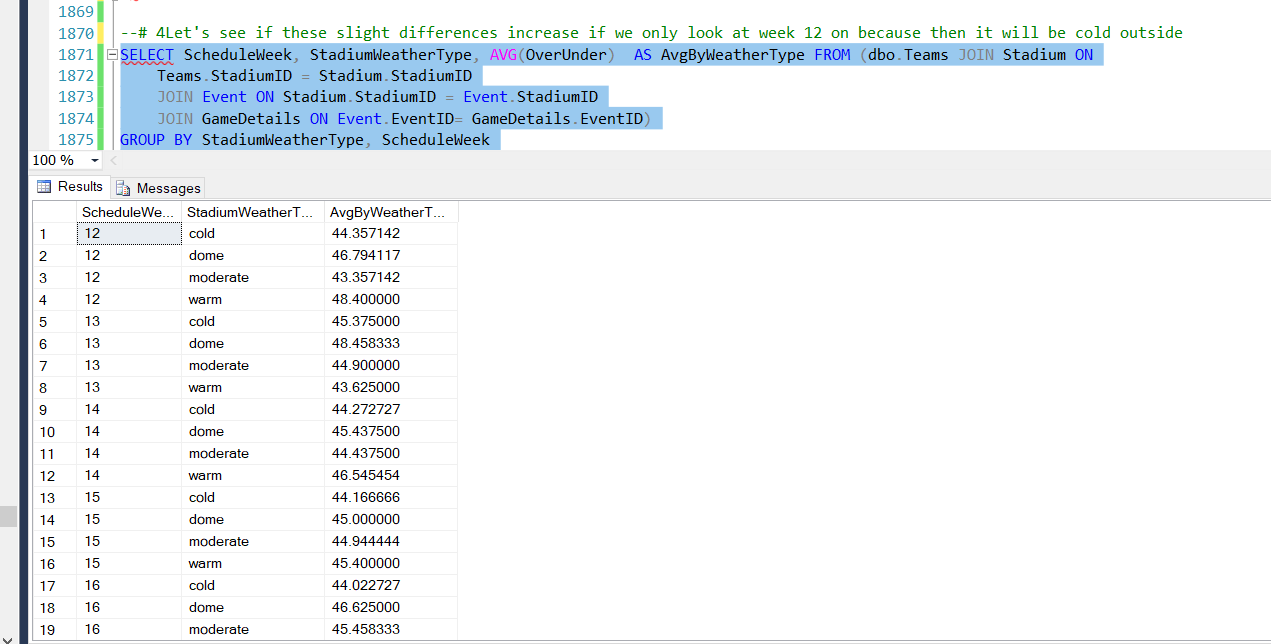




1. Are the OverOrUnder lines different by weather type of the stadiums?



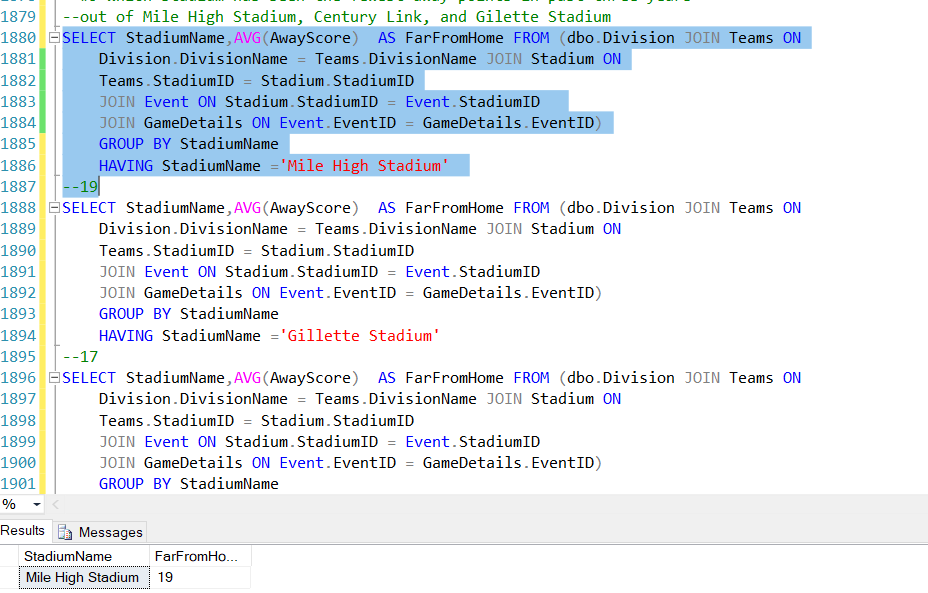
1. Are these trends different if we look at the OverOrUnder lines later in the year when it actually gets cold?

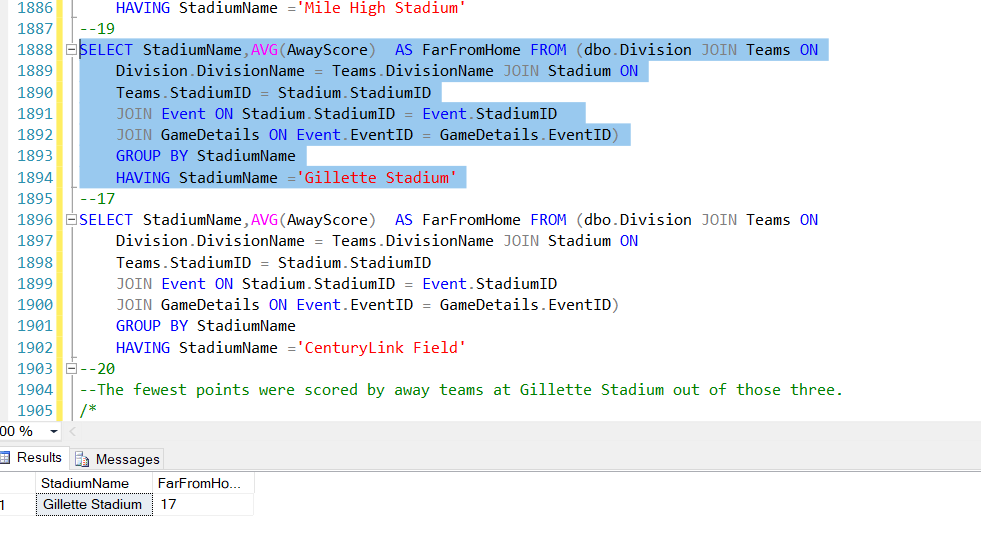


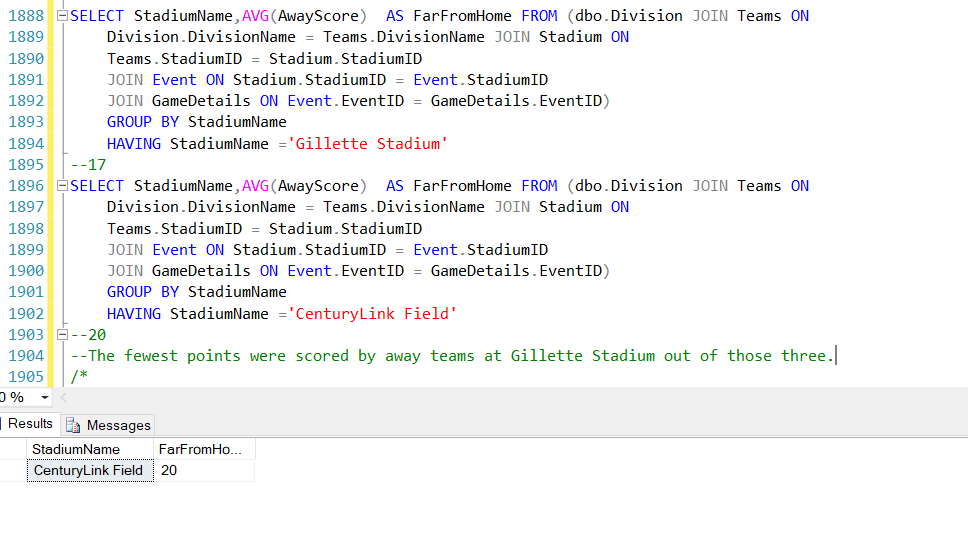
It doesn’t appear to be the case that there is a significant difference after Week 12 in the NFL. We can see that sometimes there is a different a couple of points, but nothing too major.

1. Which of these three Stadiums has seen the fewest away points in the past three years?

* Mile High Stadium
* Gillette Stadium
* CenturyLink Field



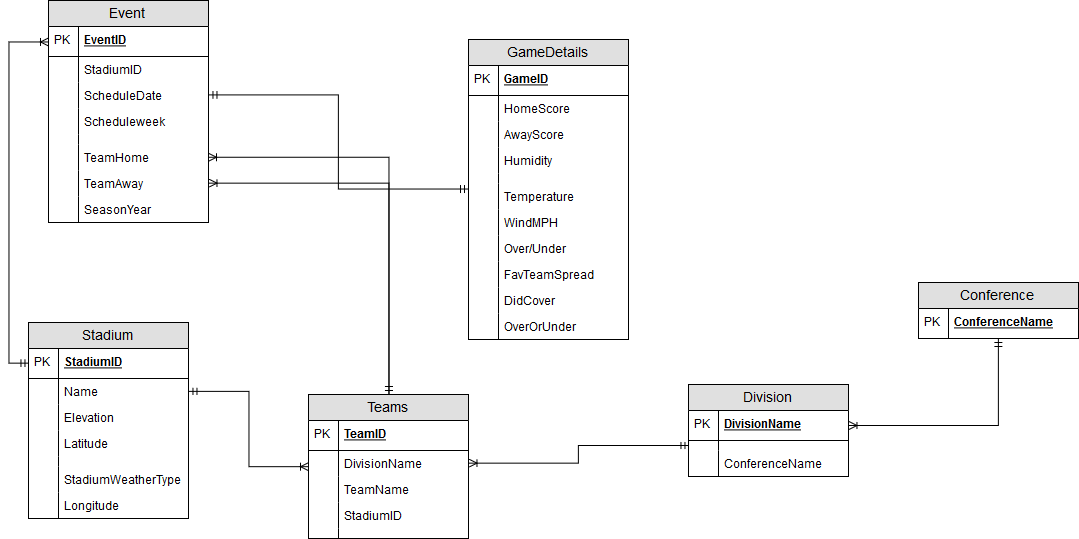




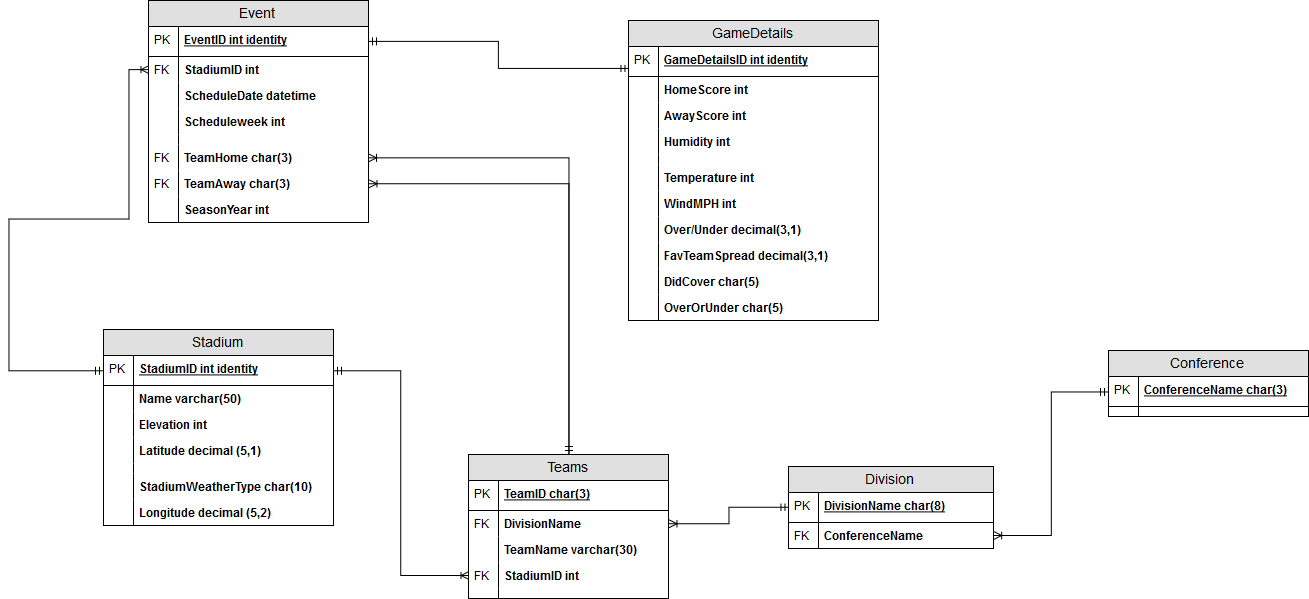
It appears that Gillette Stadium has seen the fewest points scored by opposing teams.

**Models**

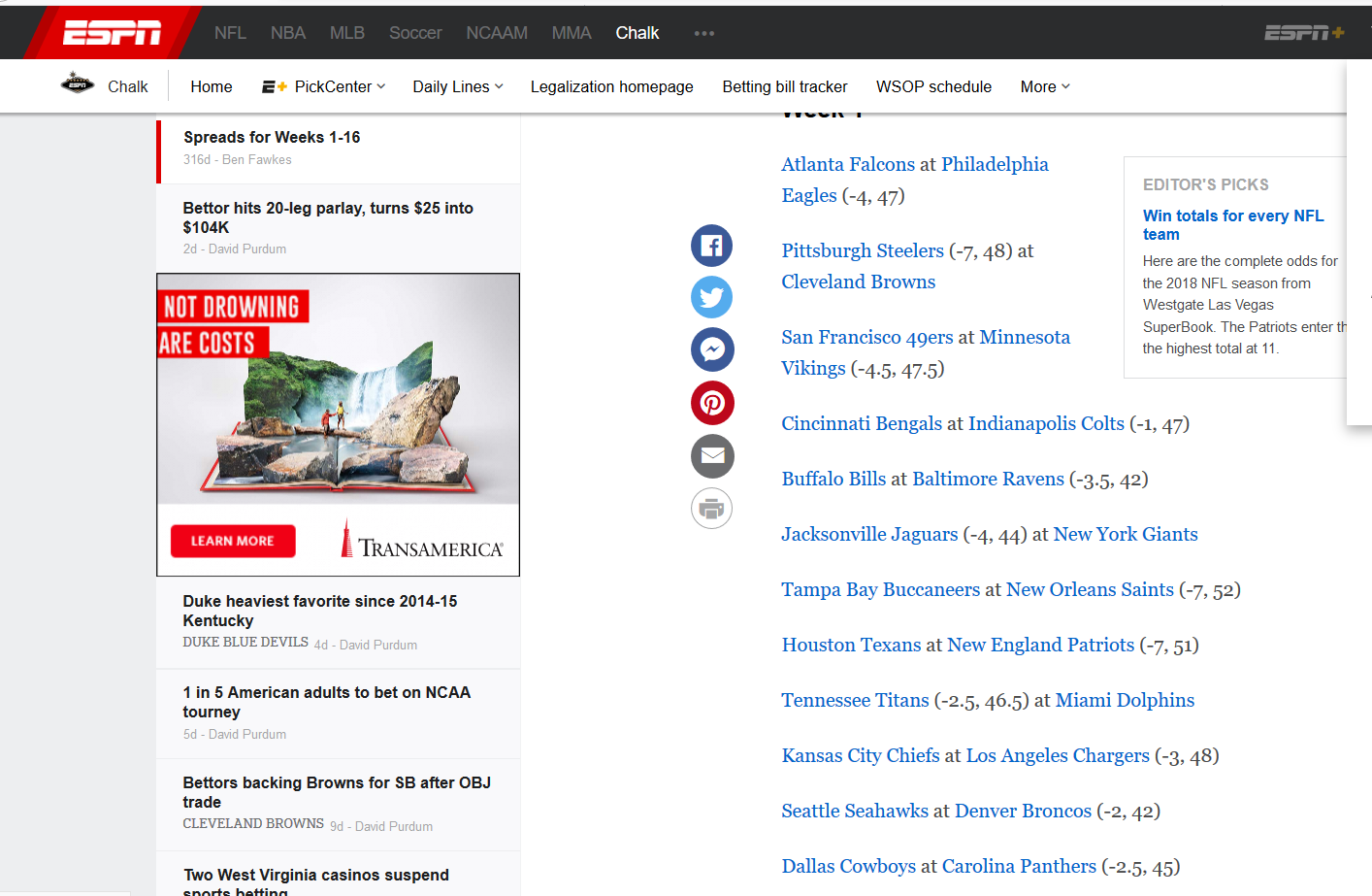
ERD



Logical Model

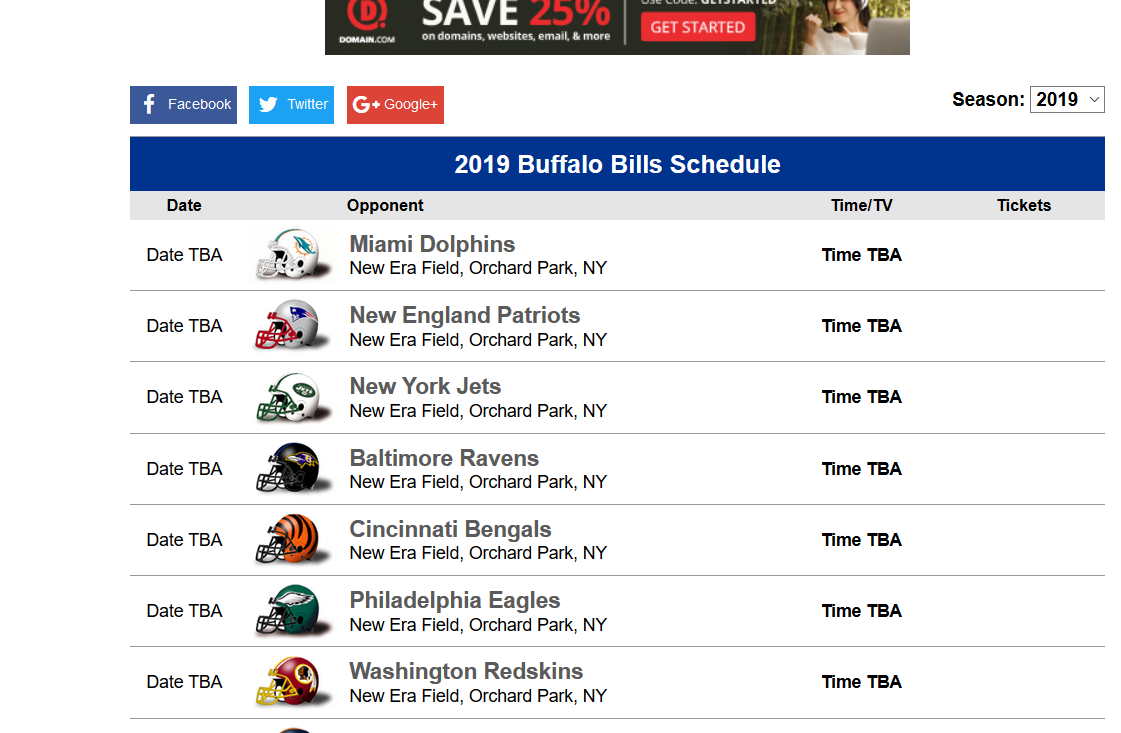


Raw Data Samples:

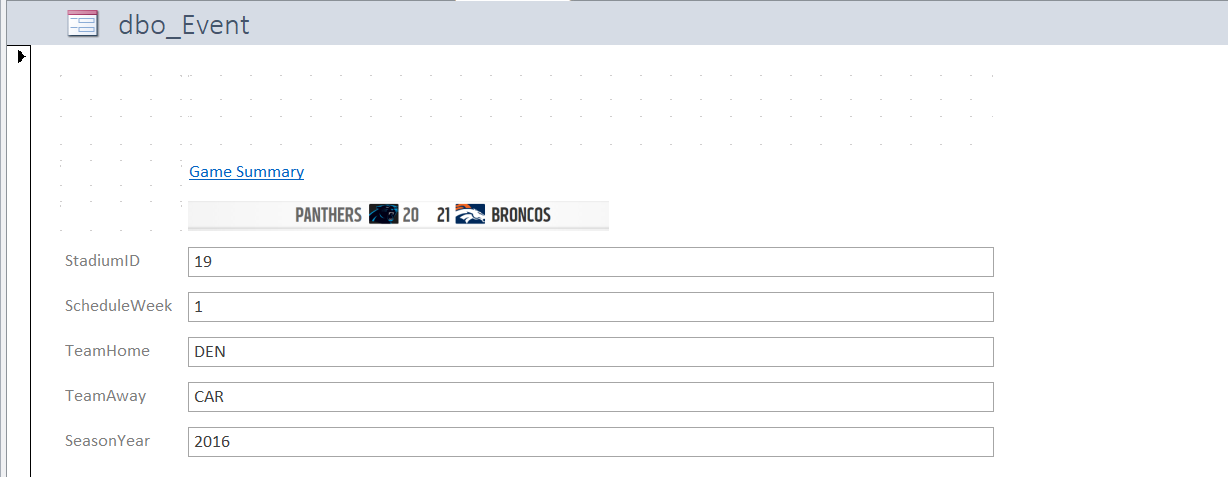




Above are examples of website where the raw data will come from. The first shows spreads and the second currently does not have spreads for games because at the time of making this it is far from football season. Vegas insider is a popular website to check out the spreads and odds for football games. Currently there are no 2019 completed NFL schedules. After the schedule is determined it will be released. For now, you can find teams upcoming opponents by looking at specific teams. Here is an example.



User Interface



To make the interface more user friendly, I made an example user interface in Microsoft Access. I attached a small picture showing the logos of the teams the played, the home team in bold, and the scored as well. Above there is a hyperlink called “Game Summary” that will take users to a summary of the game on the NFL’s website. Here they can get more information on the game if they so choose.

*Reflection & Summary*

*If I had to do this project again I would have chosen a topic with less data points and downloaded and configured MySQL on my laptop. This project took a long time due to time spent data munging and ensuring that each data point would match up with each other corresponding table. I have all data points from 2002 to present however, it isn’t exactly practical to use a million data points for a project. Downloading MySQL and configuring it to my computer would have allowed me any privileges within the database that connecting remotely to Syracuse Universities did not allow.*

*I would have allowed for much more time learning more advanced queries to run against the database. I still have the information, so I may practice on the current database that I installed.*

*Again, the project went fine overall. I learned a lot about database design and the idiosyncrasies of excel when concatenating text to create my inserts. Practicing the overall process has given me an appreciate for the process. My interest in databases has increased, and I would say that once you create your own database and run code against it the concepts become clearer.*

*Attached is the SQL file containing each line of code from the repeatable physical database design to the last data question answered.*