POWER BI SHOWCASE PROJECT DOCUMENTATION

NFL DATA 2002 – 2022 SEASON

ABBEY TOWSE
BUS 464

Table of Contents

Data Selection:

For my showcase project, I decided to use NFL data from the 2002 season through the 2022 season. Looking for data to use for my showcase project was challenging because I do not have access to a database; for context, Power BI projects typically pull from a company database. For my project, I needed to find multiple data tables that have the same primary key; in other words, multiple data tables recording data on the same subject. This proved to be hard to find from open-source data because most data is from individual studies using anonymous participants.

I ended up utilizing NFL data for my project because every season tracks a variety of data about the same teams; this provides me with a primary key for each NFL team per season. To create multiple tables, I gathered data about team statistics, draft orders, coaches, conferences, playoffs, and superbowls. Another reason why I chose to utilize NFL data is that I have worked with NFL data in other courses I have taken, so I am familiar with what statistics exist and what to use them for.

The reason I am using data starting from 2002 is that 2002 is when the NFL became the 32 teams and two conferences that are known today. Before 2002, there were different amounts of teams and conferences, which would make trying to compare and analyze data difficult.

Load & Prepare Data:

Coaches Table & NFL Coaches Table

Variables (both tables, coach table only, nfl_coaches table only)

1. Coach ID – the unique identifier for each coach
1. Coach – the name of the head coach that season
2. Team ID – the unique identifier for each team each season
3. Team Name – the name of the team
4. Year – the season
5. Years of Experience – the years of experience the head coach had during this season
6. To – the year the coach started head coaching
7. From – the year the coach finished head coaching

Power Query Transformations

1. Loaded table of coaches and their team for each season (21 tables)
    a. Excel files
2. Promoted the first row to the header for each table (21 tables)
3. Created a custom column to label the season for each table (21 tables)
    a. Example: “2002”, “2003”, …, “2021”, “2022”

4. Appended the 21 tables together to make 1 table with all coaches and all seasons
    a. coaches table
5. Removed columns from table that will not be used in analysis (coaches table)
    a. Games coached that regular season
    b. Wins that regular season
    c. Losses that season
    d. Ties that regular season
    e. Score for each regular season game
    f. Games coached that playoffs
    g. Wins that playoffs
    h. Losses that playoffs
    i. Ties that playoffs
    j. Score for each playoff game
    k. Remarks (Superbowl champion, AFC champion, etc.)
6. Changed team names in the Tm column from acronyms to spelled out team names (coaches table)
    a. To match the other data sources
    b. To make the data in the table easier to understand
    c. Example: “ARI” to “Arizona Cardinals”
7. Loaded table of all NFL coaches and their experience (1 table)
    a. TXT file
    b. nfl_coaches table
8. Removed columns that will not be used in analysis (nfl_coaches table)
    a. Rank
    b. Total Games Coaches
    c. Total Games Won
    d. Total Games Lost
    e. Total Games Tied
    f. Win-Loss Percentage
    g. Average Rank
    h. Best Rank
    i. Total Championships
    j. Total Superbowls
    k. Total Conference Titles
9. Filtered rows by the “To” column (last season coached) to only keep valued greater than or equals to 2002 (nfl_coaches table)
10. Removed extra “+” symbol next to some coaches’ names (nfl_coaches table)
    a. To make sure the coach’s name will be matched during merging
11. Merge the nfl_coaches table with the coaches table
12. Create a calculated column to get the years of experience of the head coach (coaches table)
    a. “Years of Experience” = “Year” – “From”

        i. Year – the current season
        ii. From – the year the coach started being a head coach
13. Removed columns from table that are no longer needed to create calculated columns (coaches table)
    a. To
    b. From
14. Renamed column (coaches table)
    a. "Tm" to "Team Name"
        i. To match the column name of other data sources
        ii. To remove acronyms to make the table easier to understand
15. Create an index column in the nfl_coaches table
16. Renamed column (nfl_coaches table)
    a. Index to "Coach ID"
        i. To make the column name accurately describe the data it is storing
17. Merge the "Coach ID" column from nfl_coaches
18. Reordered column to make "Coach ID" the first column in the table
    a. Having the primary key first improves table readability
19. Merged the "Team Name" and "Year" columns to create a "Team ID" column
20. Reordered column to make "Team ID" the second column in the table
    a. Having the ID columns at the beginning improves table readability

Conferences Table

Variables:

2. Franchise ID – the unique identifier for each team franchise
3. Team Name – the name of the team
4. Conference – the conference the team belongs to
5. Division – the division the team belongs to

Power Query Transformations

1. Loaded table of team names and conference the team belongs to (1 table)
    a. Excel file
2. Added Index Column from 1-32
3. Renamed column
    a. "Index" to "Franchise ID"
        i. To make the column name accurately describe the data it is storing
4. Reordered column to make "Franchise ID" the first column in the table
    a. Having the primary key first improves table readability
5. Manually entered data into Power BI to create a divisions table
    a. Team Name – the name of the team

    b. Division – the division the team belongs to
6. Merge the “Division” column from divisions table
7. Renamed the “Division” column to remove prefix
    a. “divisions.Division” to “Division”

Draft Order Table

Variables:

1. Team ID – the unique identifier for each team each year
2. Draft Order – the order the team had in the draft that year

Power Query Transformations

1. Loaded table of draft order for every season (1 table)
    a. CSV file
2. Remove column from the table that will not be used in analysis
    a. Index
3. Renamed columns
    a. “Tm” to “Team Name”
        i. To match the column name of other data sources
        ii. To remove acronyms to make the table easier to understand
    b. “draft_order” to “Previous Season Rank”
        i. Improve ability to clearly understand the data
        ii. Best practice to remove underscores
4. Duplicated “Team Name” and “Year” columns
    a. Created “Team Name – Copy” column
    b. Created “Year – Copy” column
5. Merged “Team Name – Copy” and “Year – Copy” together to make “Team ID” column
    a. Example: “Buffalo Bills_2021”
6. Reordered column to make “Team ID” the first column in the table
    a. Having the primary key first improves table readability
7. Remove columns holding repetitive data
    a. “Team Name”
    b. “Year”
8. Added a “Draft Order” column
    a. Previous Season Rank – (32 + 1)
9. Change type of “Draft Order” column to Whole Number

Fact Table

Variables

1. Team ID – the unique identifier for each team each year
6. Franchise ID – the unique identifier for each team franchise
2. Superbowl ID – the unique identifier for each super bowl game
1. Quarterback ID – the unique identifier for each quarterback name
3. Coach ID - the unique identifier for each coach
4. Year – the season the IDs belong to

Power Query Transformations

1. Duplicated the playoffs table to create a base table for the fact table to be built off of (1 table)
    a. To get the "Team ID" column
    b. To get the "Team Name" column
    c. To get the "Year" column
2. Merged the conference table to get the "Franchise ID" column
3. Merged the superbowl table to get the "Superbowl ID" column
4. Merged the quarterbacks table to get the "Quarterbacks ID" column
5. Merged the coaches table to get the "Coach ID" column
6. Remove columns holding repetitive data
    a. "Team Name"

Playoffs Table

Variables

1. Team ID – the unique identifier for each team each year
2. Seed – the seed the team had in the playoffs
    a. 0 if the team did not make the playoffs
3. Why – reason why the team made the playoffs

Power Query Transformations

1. Loaded table of playoffs seeds for each conference for each season (42 tables)
    a. Excel files
2. Created a custom column to label the season for each table (42 tables)
    a. Example: "2002", "2003", …, "2021", "2022"
3. Appended the 42 tables together to make 1 table with all playoff seeds for all conferences for all seasons
4. Renamed columns
    a. "Tm" to "Team Name"
        i. To match the column name of other data sources
        ii. To remove acronyms to make the table easier to understand

5. Removed columns from table that are already included in other tables to avoid repetitive data
    a. Wins that season
    b. Losses that season
    c. Ties that season
    d. Win-Loss percentage that season
    e. Simple Rating System that season
    f. Reason
6. Changed team names in the "Team Name" column to their most recent name
    a. To ensure all teams have the same number of data points in the model
    b. Example: "St. Louis Rams" to "Los Angeles Rams"
7. Got rid of null values in the table
    a. If the team did not go to the playoffs that year, fill the "Seed" column with 0
    b. If the team did not go to the playoffs that year, fill the "Why" column with "Did Not Make Playoffs"
8. Duplicated "Team Name" and "Year" columns
    a. Created "Team Name – Copy" column
    b. Created "Year – Copy" column
9. Merged "Team Name – Copy" and "Year – Copy" together to make "Team ID" column
    a. Example: "Baltimore Ravens_2010"
10. Reordered column to make "Team ID" the first column in the table
    a. Having the primary key first improves table readability
11. Remove columns holding repetitive data
    a. "Team Name"
    b. "Year"

Quarterbacks Table & Quarterbacks List Table

Variables (both tables, quarterbacks_list table only, quarterbacks table only)

2. Quarterback ID – the unique identifier for each quarterback name
3. Rank – the ranking of a quarterback based on total passing yards that season
4. Quarterback – the name of a team's best ranked quarterback (usually starting quarterback)
5. Team Name – the name of the team
6. Completions – the total number of completed passes the quarterback had that season
7. Attempts – the total number of pass attempts the quarterback had that season
8. Yards – the total number of passing yards the quarterback had that season
9. Interceptions – the total number of interceptions the quarterback had that season

10. Touch Downs – the total number of touch downs the quarterback had that season
11. Sacks – the total number of sacks the quarterback had that season
12. Year – the season

Power Query Transformations

1. Loaded table of quarterback statistics and rank for each season (21 tables)
    a. Excel files
2. Find the best quarterback for each team and keep those rows (21 tables, 32 rows each)
    a. Keep the first 32 rows
    b. If there are not 32 unique teams in the top 32 rows, keep more rows and then filter out duplicate teams until each team has only 1 quarterback listed
3. Created a custom column to label the season for each table (21 tables)
    a. Example: “2002”, “2003”, …, “2021”, “2022”
4. Appended the 21 tables together to make 1 table with all best quarterback statistics and rankings for all seasons (quarterbacks table)
5. Removed columns from table that will be not used in analysis (quarterbacks table)
    a. Games
    b. Games Started
    c. Quarter Back Record
    d. Completion Percentage
    e. Touch Down Percentage
    f. Interception Percentage
    g. First Downs
    h. Long gain
    i. Average Gain per Attempt
    j. Adjust Yards per Attempt
    k. Yards per Completion
    l. Yards per Game
    m. Rate
    n. Quarter Back Overall Performance
    o. Yards lost due to Sacks
    p. Sack Percentage
    q. Net Yards Gain per Attempt
    r. Adjusted Net Yards per Attempt
    s. Comebacks lead by Quarterback
    t. Game Winning Drives lead by Quarterback
6. Removed extra “+” or “*” symbol next to some quarterback names (quarterbacks table)
    a. To make sure the quarterback’s name is consistent throughout the table

7. Changed team names in the Tm column from acronyms to spelled out team names (quarterbacks table)
    a. To match the other data sources
    b. To make the data in the table easier to understand
    c. Example: “PHI” to “Philadelphia Eagles”
8. Renamed columns (quarterbacks table)
    a. “R” to “Rank”
        i. To remove acronyms to make the table easier to understand
    b. “Player” to “Quarterback”
        i. To make the table clearer that it is only looking at quarterbacks
    c. “Tm” to “Team Name”
        i. To match the column name of other data sources
        ii. To remove acronyms to make the table easier to understand
    d. “C” to “Completions”
        i. To remove acronyms to make the table easier to understand
    e. “A” to “Attempts”
        i. To remove acronyms to make the table easier to understand
    f. “Yds” to “Yards”
        i. To remove acronyms to make the table easier to understand
    g. “TD” to “Touch Downs”
        i. To remove acronyms to make the table easier to understand
    h. “Int” to “Interceptions”
        i. To remove acronyms to make the table easier to understand
    i. “Sk” to “Sacks”
        i. To remove acronyms to make the table easier to understand
9. Duplicated “quarterbacks” table to make “quarterbacks_list” table
10. Removed all columns except for “Quarterback” (quarterbacks_list table)
11. Removed duplicate values from “Quarterback” column (quarterbacks_list table)
12. Added Index Column (quarterbacks_list table)
13. Renamed column (quarterbacks_list table)
    a. “Index” to “Quarterback ID”
        i. To make the column name accurately describe the data it is storing
14. Reordered column to make “Quarterback ID” the first column in the table (quarterbacks_list table)
    a. Having the primary key first improves table readability
15. Merge the “Quarterback ID” column from quarterbacks_list
16. Reordered column to make “Quarterback ID” the first column in the table (quarterbacks table)
    a. Having the primary key first improves table readability
17. Merged the “Team Name” and “Year” columns to create a “Team ID” column
18. Remove columns holding repetitive data

a. “Team Name”
b. “Year”

19. Reordered column to make “Team ID” the second column in the table
    a. Having the ID columns at the beginning improves table readability

Standings Table

Variables

1. Team ID – the unique identifier for each team each year
2. Team Name – the name of the team
3. Win – the number of games won that season
4. Loss – the number of games lost that season
5. Tie – the number of games tied that season
6. Win-Loss Percentage – the percentage of games won versus lost
7. Points For – the total points the team has scored that season
8. Points Against – the total points scored against a team that season
9. Points Difference – “Points For” minus “Points Against”
10. Margin of Victory – the average points scored than their competitor that season
11. Strength of Schedule – the combined record of all teams in the schedule that season
12. Simple Rating System – the team rating that takes in account both “Point Difference” and “Strength of Schedule”
13. Offense Simple Rating System – the offense rating for a team that season
14. Defense Simple Rating System – the defense rating for a team that season
15. Year – the season

Power Query Transformations

1. Loaded table of team statistics for each conference for each season (42 tables)
    a. Excel files
2. Created a custom column to label the season for each table (42 tables)
    a. Example: “2002”, “2003”, …, “2021”, “2022”
3. Appended the 42 tables together to make 1 table with all team statistics for all conferences for all seasons
4. Renamed columns
    a. “Tm” to “Team Name”
        i. To match the column name of other data sources
        ii. To remove acronyms to make the table easier to understand
    b. “W” to “Win”
        i. To remove acronyms to make the table easier to understand
    c. “L” to “Loss”
        i. To remove acronyms to make the table easier to understand

    d. “T” to “Tie”
        i. To remove acronyms to make the table easier to understand
    e. “W-L%” to “Win-Loss Percentage”
        i. To remove acronyms to make the table easier to understand
    f. “PF” to “Points For”
        i. To remove acronyms to make the table easier to understand
    g. “PA” to “Points Against”
        i. To remove acronyms to make the table easier to understand
    h. “PD” to “Points Difference”
        i. To remove acronyms to make the table easier to understand
    i. “MoV” to “Margin of Victory”
        i. To remove acronyms to make the table easier to understand
    j. “SoS” to “Strength of Schedule”
        i. To remove acronyms to make the table easier to understand
    k. “SRS” to “Simple Rating System”
        i. To remove acronyms to make the table easier to understand
    l. “OSRS” to “Offense Simple Rating System”
        i. To remove acronyms to make the table easier to understand
    m. “DSRS” to “Defense Simple Rating System”
        i. To remove acronyms to make the table easier to understand
5. Changed team names in the “Team Name” column to their most recent name
    a. To ensure all teams have the same number of data points in the model
    b. Example: “Oakland Raiders” to “Las Vegas Raiders”
6. Removed extra “+” or “*” symbol next to some team names
    a. To make sure the team’s name will be matched when creating relationships
7. Got rid of null values in the table
    a. If a season/team did not have any ties that year, fill the “Tie” column with 0
8. Duplicated “Team Name” and “Year” columns
    a. Created “Team Name – Copy” column
    b. Created “Year – Copy” column
9. Merged “Team Name – Copy” and “Year – Copy” together to make “Team ID” column
    a. Example: “Philadelphia Eagles_2002”
10. Reordered column to make “Team ID” the first column in the table
    a. Having the primary key first improves table readability
11. Remove columns holding repetitive data
    a. “Team Name”
    b. “Year”

Superbowl Table

Variables

1. Superbowl ID – the unique identifier for each super bowl game
2. Winner (Team Name) – the team name that won the Superbowl that season
3. Loser (Team Name) – the team that lost the Superbowl that season
4. Winner (Team ID) – the unique identifier for the team that won the Superbowl that season
5. Loser (Team ID) – the unique identifier for the team that lost the Superbowl that season
6. State – the state the Superbowl took place in that season
7. Year – the season

Power Query Transformations

1. Loaded table of Superbowl history for each season (1 table)
    a. Excel file
2. Remove rows for Superbowls that happened before the 2002 season
3. Created a custom column to label the season for each row
    a. Example: “2002”, “2003”, …, “2021”, “2022”
4. Removed columns from table that will not be used in analysis
    a. Date
    b. Superbowl number
    c. Winner
    d. Points of Winner
    e. Loser
    f. Points of Loser
    g. Superbowl MVP
    h. Stadium
    i. City
5. Added Index Column from 57-37
    a. Went from 57 to 37 because those are the numbers of the Superbowls from 2002 to 2022
6. Renamed column
    a. “Index” to “Superbowl ID”
        i. To make the column name accurately describe the data it is storing
7. Reordered column to make “Superbowl ID” the first column in the table
    a. Having the primary key first improves table readability
8. Merge the Winner and Year Column to create a new Winner column in the format of Team ID
    a. Winner.1
9. Merge the Loser and Year Column to create a new Loser column in the format of Team ID

    a. Loser.1
10. Renamed column
    a. “Winner.1” to “Winner (Team ID)
        i. To make the column name accurately describe the data it is storing
    b. “Loser.1” to “Loser (Team ID)
        i. To make the column name accurately describe the data it is storing
    c. “Winner” to “Winner (Team Name)”
        i. To make the column name accurately describe the data it is storing
    d. “Loser” to “Loser (Team Name)”
        i. To make the column name accurately describe the data it is storing
11. Changed team names in the “Team Name” column to their most recent name
    a. To ensure all teams have the same number of data points in the model
    b. “Oakland Raiders” to “Las Vegas Raiders” in “Winner (Team Name)”
    c. “Oakland Raiders_2002” to “Las Vegas Raiders_2002” in “Winner (Team ID)”

Superbowl Results Table

Variables

1. Team ID – the unique identifier for each team each season
2. Result – whether or not the team won or lost the Superbowl
3. Year – the season

Power Query Transformations

1. Duplicated the Superbowl table
2. Removed columns containing data that is unnecessary for this table
    a. Superbowl ID
    b. Winner (Team Name)
    c. Loser (Team Name)
    d. State
3. Unpivoted the “Winner (Team ID)” and “Loser (Team ID)” columns creating “Value” and “Attribute” columns
    a. The “Value” column contains the Team ID
    b. The “Attribute” column contains either “Winner (Team ID)” or “Loser (Team ID)” for each row
4. Renamed columns
    a. “Value” to “Team ID”

        i. To make the column name accurately describe the data it is storing
    b. “Attribute” to “Result”
        i. To make the column name accurately describe the data it is storing
5. Replaced values in the “Result” column
    a. “Winner (Team ID” to “Winner”
        i. To make the data more concise
    b. “Loser (Team ID)” to “Loser”
        i. To make the data more concise
6. Reordered column to make “Team ID” the first column in the table
    a. Having the primary key first improves table readability

Data Model Design:

Table Descriptions

Coaches

The Coaches table contains information specific to what team each head coach coached for each year. Furthermore, this table includes the number of years of experience said coach had when coaching the corresponding team.

* To ensure each team only had one head coach per season the head coach that completed the season with the team was selected.

Conference

The Conference table tracks what conference each franchise/team is a part of. This information includes “Franchise ID”, “Team Name”, and “Conference.

Draft_Order

The Draft_Order tables contains when a team got to pick in the draft for a season.

Note: The draft order in the table is pre trades meaning it represents how the team performed in the previous season.

Fact_Table

The Fact_Table table contains all the unique identifiers from all the tables in the model. These unique identifiers include “Team ID”, “Franchise ID”, “Coach ID”, “Quarterback ID”, and “Superbowl ID”.

NFL_Coaches

The NFL_Coaches table contains relevant information about everyone that has been a head NFL coach from 2002 to 2022. This information includes “Coach ID”, “Coach” (coach name), “From” (start year), and “To” (end year).

Playoffs

The Playoffs tables contains the seed a team had going into the playoffs for a particular season and a statement saying why they got the seed number that they did (Wild Card or Conference Champion).

Quarterbacks

The Quarterbacks table contains information specific to how a quarterback performed in a specific season for a specific team. This information includes “Attempts”, “Completions”, “Interceptions”, “Rank”, “Sacks”, “Touch Downs”, and “Yards.” This table also contains the age of the quarterback during the corresponding season.

Quarterback_List

The Quarterback_List table contains all starting* quarterbacks from 2002 to 2022.

* To ensure each team only had one quarterback per season the quarterback selected was a team’s highest rated quarterback for a given season.

Standings

The Standings table contains information specific to how a given team performed in a certain season. This information includes “Defense Simple Rating System”, “Loss”, “Margin of Victory”, “Offense Simple Rating System”, “Points Against”, “Points Difference”, “Points For”, “Simple Rating System”, “Strength of Schedule”, “Tie”, “Win”, and “Win-Loss Percentage.”

Superbowl

The Superbowl table contains information regarding the Superbowl that took place each season. This table records what team won the Superbowl ("Winner (Team Name)" / "Winner (Team ID)), what team went to the Superbowl but lost ("Loser (Team Name)" / "Loser (Team ID)), and what state the Superbowl was held in ("State").

Superbowl_Results

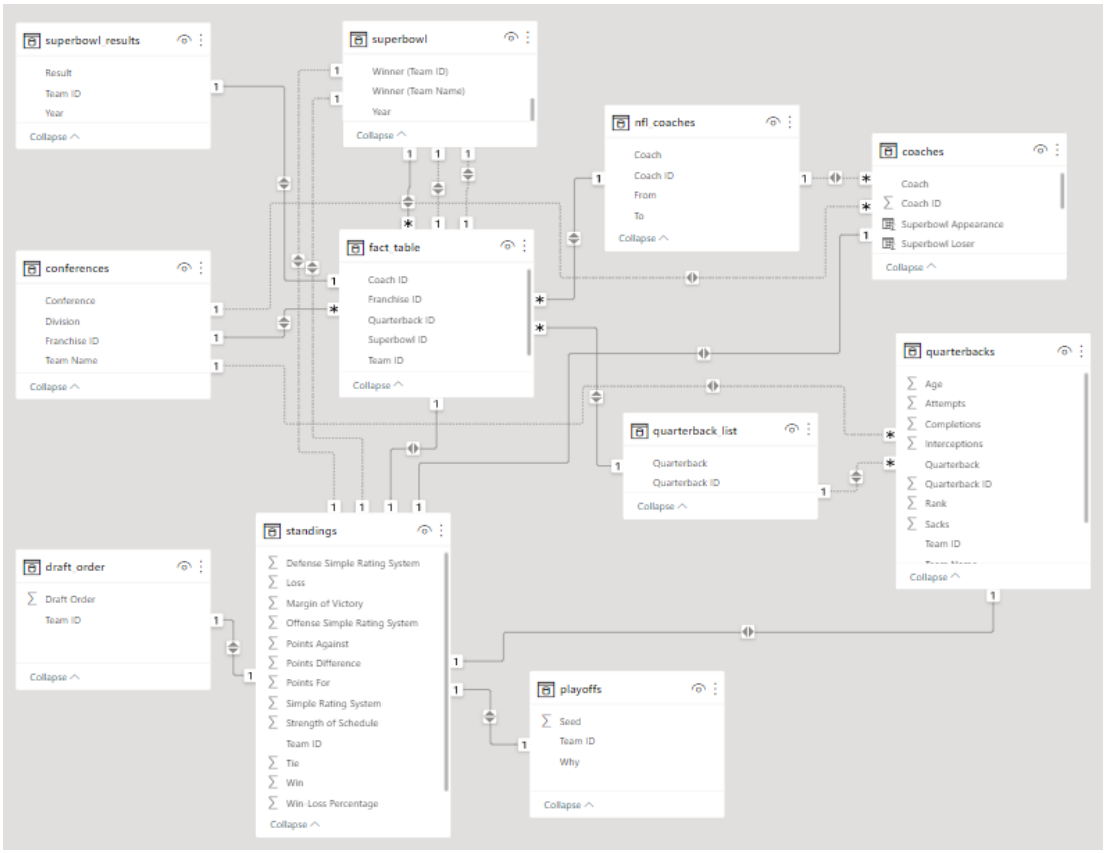
The Superbowl_Results table contains information regarding if a team won or lost a Superbowl game each year in which they participated in the Superbowl. This information includes "Team ID", "Result", and "Year".

Table Relationships

1. Coaches
    a. Many-to-One relationship to NFL_Coaches on "Coach ID" (Inactive)
    b. One-to-One relationship to Standings on "Team ID" (Active)
2. Conferences
    a. One-to-Many relationship to Coaches on "Team Name" (Inactive)
    b. One-to-Many relationship to Fact_Table on "Franchise ID" (Active)
    c. One-to-Many relationship to Quarterbacks on "Team Name" (Inactive)
3. Draft_Order
    a. One-to-One relationship with Standings on "Team ID" (Active)
4. Fact_Table
    a. Many-to-One relationship to Conferences on "Franchise ID" (Inactive)
    b. Many-to-One relationship to Superbowl on "Superbowl ID" (Active)
    c. One-to-One relationship to Superbowl on "Team ID"/"Winner (Team ID)" (Inactive)
    d. One-to-One relationship to Superbowl on "Team ID"/"Loser (Team ID)" (Inactive)
    e. Many-to-One relationship to NFL_Coaches on "Coach ID" (Active)
    f. Many-to-One relationship to Quarterbacks_List on "Quarterback ID" (Active)
    g. One-to-One relationship to Standings on "Team ID" (Active)
    h. One-to-One relationship to Superbowl_Results on "Team ID" (Active)
5. NFL_Coaches
    a. One-to-Many relationship to Coaches on "Coach ID" (Inactive)
    b. One-to-Many relationship to Fact_Table on "Coach ID" (Active)
6. Playoffs
    a. One-to-One relationship to Standings on "Team ID" (Active)
7. Quarterbacks

    a. Many-to-One relationship to Conference on “Team Name” (Inactive)
    b. Many-to-One relationship to Quarterbacks_List on “Quarterback ID” (Inactive)
    c. One-to-One relationship to Standings on “Team ID” (Active)
8. Quarterbacks_List
    a. One-to-Many relationship to Fact_Table on “Quarterback ID” (Active)
    b. One-to-Many relationship to Quarterbacks on “Quarterback ID” (Inactive)
9. Standings
    a. One-to-One relationship to Coaches on “Team ID” (Active)
    b. One-to-One relationship to Draft_Order on “Team ID” (Active)
    c. One-to-One relationship to Fact_Table on “Team ID” (Active)
    d. One-to-One relationship to Playoffs on “Team ID” (Active)
    e. One-to-One relationship to Quarterbacks on “Team ID” (Active)
    f. One-to-One relationship to Superbowl on “Team ID” (Inactive)
    g. One-to-One relationship to Superbowl on “Team ID” (Inactive)
10. Superbowl
    a. One-to-Many relationship to Fact_Table on “Superbowl ID” (Active)
    b. One-to-One relationship to Fact_Table on “Team ID”/“Winner (Team ID)” (Inactive)
    c. One-to-One relationship to Fact_Table on “Team ID”/“Loser (Team ID)” (Inactive)
    d. One-to-One relationship to Standings on “Team ID”/“Winner (Team ID)” “Team ID” (Inactive)
    e. One-to-One relationship to Standings on “Team ID”/“Loser (Team ID)” (Inactive)
11. Superbowl_Results
    a. One-to-One relationship with Fact_Table on “Team ID” (Active)

Model



Report Creation:

DAX Measures

Coaches Table

Superbowl Appearance = coaches[Superbowl Winner] + coaches[Superbowl Loser]

- A Boolean column (0 or 1) to show if a coached appeared in the Superbowl that season/team

Superbowl Loser = IF (coaches[Team ID] IN DISTINCT (superbowl[Loser (Team ID)]), 1, 0)

- A Boolean column (0 or 1) to show if a coach won or lost a Superbowl that season/team

Superbowl Winner = IF (coaches[Team ID] IN DISTINCT (superbowl[Winner (Team ID)]), 1, 0)

- A Boolean column (0 or 1) to show if a coach won or lost a Superbowl that season/team

Standings Table

```
Tie divided by 2 = INT(DIVIDE(SUM('standings'[Tie]), 2))
```

- A measure to use in the reports in order to not double the number of ties that happened
    - When a tie occurs it will be recorded under two teams because two teams are involved in a tied game; however, the number of ties total is half of that because there is one tie for every two teams.

Superbowl Table

```
Total hosted Superbowls since 2002 = COUNTROWS(FILTER(superbowl, EARLIER(superbowl[State]) = superbowl[State]))
```

- A column to show how many total times each state hosted the Superbowl since 2002

Report Pages

The Power BI has four report pages. Each of these report pages summarize different facets of data from the tables from the data model. Every table is included somewhere in my report; however, not every table column because of relevancy and space constraints. Slicers are synced across all tables so that the viewer can examine the data from different lenses.

Coaches

Tables/Columns used in Page

- Coaches Table
    - Coach
    - Coach ID
    - Superbowl Appearance (DAX measure)
    - Superbowl Winner (DAX measure)
    - Superbowl Loser (DAX measure)
    - Team ID
    - Team Name
    - Year
    - Years of Experience
- Fact Table
    - Year

Visuals used in Page

- Card
  - Average Years in the Head Coaching Industry
  - Number of Appearances in the Superbowl
  - Number of Distinct Coaches
  - Number of Losses in the Superbowl
  - Number of Wins in the Superbowl
- Clustered Column Chart
  - Coach's Years in Head Coaching Industry by Team ID
    - x axis: Team ID
    - y axis: Years in Head Coaching Industry
    - Tooltips: First Year
- Slicer
  - Coach (synced with Coach, Conferences, Quarterbacks, & Superbowl pages)
  - Team Name (synced with Coaches, Conferences, Quarterbacks, & Superbowl pages)
  - Season (synced with Coaches, Conferences, Quarterbacks, & Superbowl pages)

Conferences

Tables/Columns used in Page

- Coaches
  - Team Name
- Conferences
  - Conference
  - Division
- Draft Order Table
  - Draft Order
- Fact Table
  - Year
- Quarterbacks
  - Interceptions
  - Sacks
  - Touch Downs
  - Yards
- Standings
  - Simple Rating System
  - Tie divided by 2 (DAX measure)
- Superbowl

    - Superbowl ID

Visuals used in Page

- Area Chart
    - Average of Interceptions by Year and Conference
        - x axis: Season
        - y axis: Average of Interceptions
        - Legend: Conferences
    - Average of Sacks by Year and Conference
        - x axis: Season
        - y axis: Average of Sacks
        - Legend: Conference
    - Average of Touch Downs by Year and Conference
        - x axis: Season
        - y axis: Average of Touch Downs
        - Legend: Conference
    - Average of Yards by Year and Conference
        - x axis: Season
        - y axis: Average of Yards
        - Legend: Conference
- Card
    - AFC Average Draft Order
        - Filters on this visual: Team Name (Baltimore Ravens, Buffalo Bills, Cincinnati Bengals, Cleveland Browns, Denver Broncos, Houston Texans, Indianapolis Colts, Jacksonville Jaguars, Kansas City Chiefs, Las Vegas Raiders, Los Angeles Chargers, New England Patriots, New York Jets, Pittsburgh Steelers, Tennessee Titans)
    - AFC Superbowl Wins
        - Filters on this visual: Team Name (Baltimore Ravens_2012, Denver Broncos_2015, Indianapolis Colts_2006, Kansas City Chiefs_2019, Kansas City Chiefs_2022, New England Patriots_2003, New England Patriots_2004, New England Patriots_2014, New England Patriots_2016, New England Patriots_2018, Pittsburgh Steelers_2005, Pittsburgh Steelers_2008)
    - AFC Ties
        - Filters on this visual: Team Name (Baltimore Ravens, Buffalo Bills, Cincinnati Bengals, Cleveland Browns, Denver Broncos, Houston Texans, Indianapolis Colts, Jacksonville Jaguars, Kansas City Chiefs, Las Vegas Raiders, Los Angeles Chargers, New England Patriots, New York Jets, Pittsburgh Steelers, Tennessee Titans)

    - NFC Average Draft Order
        - Filters on this visual: Team Name (Arizona Cardinals, Atlanta Falcons, Carolina Panthers, Chicago Bears, Dallas Cowboys, Detroit Lions, Green Bay Packers, Los Angeles Rams, Minnesota Vikings, New Orleans Saints, New York Giants, Philadelphia Eagles, San Francisco 49ers, Seattle Seahawks, Tampa Bay Buccaneers, Washington Commanders)
    - NFC Superbowl Wins
        - Filters on this visual: Team ID (Green Bay Packers_2010, Los Angeles Rams_2021, New Orleans Saints_2009, New York Giants_2007, New York Giants_2011, Philadelphia Eagles_2017, Seattle Seahawks_2013, Tampa Bay Buccaneers_2002, Tampa Bay Buccaneers_2020)
    - NFC Ties
        - Filters on this visual: Team Name (Arizona Cardinals, Atlanta Falcons, Carolina Panthers, Chicago Bears, Dallas Cowboys, Detroit Lions, Green Bay Packers, Los Angeles Rams, Minnesota Vikings, New Orleans Saints, New York Giants, Philadelphia Eagles, San Francisco 49ers, Seattle Seahawks, Tampa Bay Buccaneers, Washington Commanders)
- Clustered Column Chart
    - AFC Average of Simple Rating System by Division
        - x axis: Division
        - y axis: Average of Simple Rating System
        - Filters on this visual: Divisions (AFC East, AFC North, AFC South, AFC West)
    - NFC Average of Simple Rating System by Division
        - x axis: Division
        - y axis: Average of Simple Rating System
        - Filters on this visual: Divisions (NFC East, NFC North, NFC South, NFC West)
- Ribbon Chart
    - Average Margin of Victory by Year and Conference
        - x axis: Season
        - y axis: Average Margin of Victory
        - Legend: Conference
- Slicer
    - Team Name (synced with Coaches, Conferences, Quarterbacks, & Superbowl pages)
    - Season (synced with Coaches, Conferences, Quarterbacks, & Superbowl pages)

Special Features

- Buttons/Bookmarks
  - Yards
    - Action: Bookmark
      - Displays the “Average of Yards by Season and Conference” clustered column chart
  - Touch Downs
    - Action: Bookmark
      - Displays the “Average of Touch Downs by Season and Conference” clustered column chart
  - Interceptions
    - Action: Bookmark
      - Displays the “Average of Interceptions by Season and Conference” clustered column chart
  - Sacks
    - Action: Bookmark
      - Displays the “Average of Sacks by Season and Conference” clustered column chart

Quarterbacks

Tables/Columns used in Page

- Coaches Table
  - Team Name
- Fact Table
  - Year
- Quarterback List Table
  - Quarterback
- Quarterbacks Table
  - Age
  - Attempts
  - Completions
  - Interceptions
  - Quarterback
  - Rank
  - Sacks
  - Team Name
  - Touch Downs
  - Yards

Visuals used in Page

- Card
  - Total Yards Gained by Quarterback
  - Total Pass Attempts by Quarterback
  - Total Pass Completions by Quarterback
  - Total Interceptions by Quarterback
  - Total Sacks by Quarterback
  - Total Touch Downs by Quarterback
- Key Influencers (AI Visual)
  - What influences Rank to:
    - Analyze: Rank
    - Explain by: Sum of Interceptions, Sum of Touch Downs, Average of Age, Sum of Attempts, Sum of Completions, Sum of Sacks, Sum of Yards
- Line Chart
  - Quarterback Rank by Season
    - x axis: Season
    - y axis: Rank
    - Legend: Team Name
    - Tooltips: First Quarterback, Sum of Rank, Average of Age, Sum of Yards, Sum of Attempts, Sum of Completions, Sum of Interceptions, Sum of Sacks, Sum of Touch Downs
- Multi-row Card
  - Fields: Team Name, Quarterback, Rank, Age, Year
- Slicer
  - Quarterback (synced with Coaches, Conferences, Quarterbacks, & Superbowl)
  - Team Name (synced with Coaches, Conferences, Quarterbacks, & Superbowl)
  - Season (synced with Coaches, Conferences, Quarterbacks, & Superbowl)

Superbowl

Filters used on Page

- Team ID: Arizona Cardinals_2008, Atlanta Falcons_2016, Baltimore Ravens_2012, Carolina Panthers_2003, Carolina Panthers_2015, Chicago Bears_2006, Cincinnati Bengals_2021, Denver Broncos_2013, Denver Broncos_2015, Green Bay Packers_2010, Indianapolis Colts_2006, Indianapolis Colts_2009, Kansas City Chiefs_2019, Kansas City Chiefs_2020, Kansas City Chiefs_2022, Las Vegas Raiders_2002, Los Angeles Rams_2018, Los Angeles Rams_2021, New England Patriots_2003, New England Patriots_2004, New England Patriots_2007,

New England Patriots_2011, New England Patriots_2014, New England Patriots_2016, New England Patriots_2017, New England Patriots_2018, New Orleans Saints_2009, New York Giants_2007, New York Giants_2011, Philadelphia Eagles_2004, Philadelphia Eagles_2017, Philadelphia Eagles_2022, Pittsburgh Steelers_2005, Pittsburgh Steelers_2008, Pittsburgh Steelers_2010, San Francisco 49ers_2012, San Francisco 49ers_2019, Seattle Seahawks_2005, Seattle Seahawks_2013, Seattle Seahawks_2014, Tampa Bay Buccaneers_2002, Tampa Bay Buccanners_2020

Tables/Columns used in Page

- Coaches Table
  - Years of Experience
- Conferences Table
  - Conference
  - Division
- Draft Order Table
  - Draft Order
- Fact Table
  - Team ID
  - Year
- Playoffs Table
  - Seed
- Quarterbacks Table
  - Age
- Standings Table
  - Simple Rating System
  - Strength of Schedule
  - Win-Loss Percentage
- Superbowl Table
  - Loser (Team Name)
  - Superbowl ID
  - State (set column to location)
  - Total hosted Superbowls since 2002 (DAX Measure)
  - Winner (Team Name)
- Superbowl Results Table
  - Result
  - Year

Visuals used in Page

- Card
  - Average Age of Starting Quarterback
  - Average Order in the Draft Pick (Before Trades)
  - Average Playoff Seed
  - Coach's Average Years in the Head Coach Industry
- Filled Map
  - State and Total hosted Superbowls since 2002
    - Location: State
    - Legend: Total hosted Superbowls since 2002
    - Tooltips: First Year, Last Year
- Pie Chart
  - Distribution of Superbowl Wins by Team Name since 2002
    - Legend: Winner (Team Name)
    - Values: Count of Superbowl ID
    - Tooltips: First Year, Last Year
  - Distribution of Superbowl Losses by Team Name since 2002
    - Legend: Loser (Team Name)
    - Values: Count of Superbowl ID
    - Tooltips: First Year, Last Year
- Table
  - Columns: Team ID, Average of Simple Rating System, Average of Strength of Schedule, Average of Win-Loss Percentage, Year
- Slicers
  - Conference (synched with Coaches, Conferences, Quarterbacks, & Superbowl page)
  - Division (synced with Coaches, Conferences, Quarterbacks, & Superbowl page)
  - Result
  - Season (synced with Coaches, Conferences, Quarterbacks, & Superbowl page)

Color Scheme

For my report I wanted to make the visuals lines/slices/columns match the color scheme for the team it is representing that statistics for. I wanted to do this to give the report a customized theme that would speak to sports fans. I also customized visuals that split stats by conference with the NFC/AFC colors, so the representations are intuitive.

NFL/AFC/NFC HEX Colors

- #013369: Dark Midnight Blue
- #D50A0: Rosso Corsa

Team HEX Colors

- Arizona Cardinals - #97233F: Red
- Atlanta Falcons - #A71930: Red
- Baltimore Ravens - #241773: Purple
- Buffalo Bills - #0038D: Blue
- Carolina Panthers - #0085CA: Carolina Blue
- Chicago Bears - #0B162A: Dark Navy
- Cincinnati Bengals - #FB4F14: Orange
- Cleveland Browns - #311D00: Brown
- Dallas Cowboys - #003594: Royal Blue
- Denver Broncos - #FB4F14: Broncos Orange
- Detroit Lions - #0076B6: Honolulu Blue
- Green Bay Packers - #203731: Dark Green
- Houston Texans - #03202F: Deep Steel Blue
- Indianapolis Colts - #A2AAAD: Gray
- Jacksonville Jaguars - #006778: Teal
- Kansas City Chiefs - #E31837: Red
- Las Vegas Raiders - #000000: Raiders Black
- Los Angeles Chargers - #0080C6: Powder Blue
- Los Angeles Rams - #003594: Blue
- Miami Dolphins - #008E97: Aqua
- Minnesota Vikings - #4F2683: Purple
- New England Patriots - #002244: Nautical Blue
- New Orleans Saints - #D3BC8D: Old Gold
- New York Giants - #0B2265: Dark Blue
- New York Jets - #125740: Gotham Green
- Philadelphia Eagles - #004c54: Midnight Green
- Pittsburgh Steelers - #FFB612: Gold
- San Francisco 49ers - #AA0000: 49ers Red
- Seattle Seahawks - #69BE28: Action Green
- Tampa Bay Buccaneers - #D50A0A: Red
- Tennessee Titans - #0C2340: Titans Navy
- Washington Commanders - #5A1414: Burgundy

Mobile Layout

One of the features of Power BI is that it is a mobile application accessible. Reports created on a PC can be rearranged in a mobile layout allowing people to access the data in a report anywhere. Despite my report not being used by people other than myself, and I do not need to access it from my mobile device, I wanted to develop the mobile layout to get experience using as many features of Power BI as possible.

For the design of my mobile layout, I decided to think of my report visuals as groups. Mobile device screens limit the amount of content a user can view at once, so I wanted to ensure that related visuals are neighboring in my mobile layout. Additionally, for navigation ease, I placed all my slicers at the top of the screen. I placed recurrent slicers like “Season” and “Team Name” in the same location on each mobile report page layout to allow the users to navigate the mobile report using muscle memory. Furthermore, I placed all slicers at the top because having all the slicers together prevents the user from searching the page for the slicer they want to use; it also allows the user to easily determine what slicers are available on each page.

I would also like to note, although not required, I did include every visual from my report in my mobile layouts. I chose to do this so that potential mobile report users are not confused searching for a visual from the PC report on the mobile report but are unable to find it.

Perform Data Analysis:

The way I created and designed my report is ideal for team-to-team comparisons. Many of my visuals/graphs are overwhelming to look at if not examining a specific subset of teams/coaches/quarterbacks/years. With that said, for my dashboard, I aim to showcase what my report is designed for, ie team-to-team comparisons.

Before starting to create my dashboard, I explored my data thoroughly. My goal was to find two teams that are reasonable to compare from different divisions. By reasonable to compare I mean two teams that a sports fan in real life would be interested in examining the statistical difference between. Most rivalries in football are within the same conference and/or division, so I decided to look at cities/states that have multiple NFL teams. I examined teams like the Los Angeles Chargers and Los Angeles Rams, New York Jets and New York Giants, and the Tampa Bay Buccaneers, Miami Dolphins, and Jacksonville Jaguars; however, I did not use these team comparisons for my dashboard because only one team in each of these groups have gone to the Superbowl since 2002, which would make my Superbowl visuals from my report unusable in my dashboard.

I had to decide between two groups of team comparisons for my dashboard: Pittsburgh Steelers/Philadelphia Eagles and Indianapolis Colts/Chicago Bears. I ended up going with the Indianapolis Colts/Chicago Bears for a couple of reasons. The first reason is that the teams are rivals because Indianapolis and Chicago are in close proximity. The second

reason, which set them apart from the Pittsburgh Steelers/Philadelphia Eagles, is that the Bears and the Colts played each other in the 2006 season Superbowl. Not only have both teams gone to the Superbowl, rendering my Superbowl data usable, but I can directly compare how the teams compared in 2006 when they both made it to the Superbowl.

Dashboard Creations:

Web Layout:

All visuals I pinned from my report to my dashboard were first filtered with a slicer to only include the Indianapolis Colts and Chicago Bears allowing me to directly compare the two teams.

I pinned the following visuals from my report to my dashboard:

- Report: Distribution of Superbowl Wins by Team Name since 2002
  - Dashboard: Distribution of Colts & Bears Superbowl Wins since 2002
- Report: Distribution of Superbowl Losses by Team Name since 2002
  - Dashboard: Distribution of Colts & Bears Superbowl Losses since 2002
- Report: Average of Touch Downs by Season
  - Dashboard: Average of Touch Downs by Season by Quarterback
- Report: Average of Yards by Season
  - Dashboard: Average of Yards by Season by Quarterback
- Report: Average of Interceptions by Season
  - Dashboard: Average of Interceptions by Season by Quarterback
- Report: Average of Sacks by Season
  - Dashboard: Average of Sacks by Season by Quarterback
- Report: Coach's Years in Head Coach Industry by Team ID
  - Dashboard: Coach's Years in Head Coach Industry by Team ID
- Report: Quarterback Rank by Season
  - Dashboard: Quarterback Rank by Season
- Report: Average of Margin of Victory by Season and Conference
  - Dashboard: Average of Margin of Victory by Season
- Report: Average of Simple Rating System, Strength of Schedule, and Win-Loss Percentage by Year and Team ID
  - Dashboard: Average of Simple Rating System, Strength of Schedule, and Win-Loss Percentage by Year and Team ID
- Report: Average Age of Starting Quarterback Teams with Superbowl Appearances
  - Dashboard: Bears Quarterback Age 2006 Season
- Report: Average Order in the Draft Pick (Before Trades) Teams with Superbowl Appearances
  - Dashboard: Bears Draft Order 2006 Season (Before Trades)
- Report: Average Playoff Seed Teams with Superbowl Appearances

    - Dashboard: Bears Playoff Seed 2006 Season
- Report: Coach's Average Years in the Head Coaching Industry Teams with Superbowl Appearances
    - Dashboard: Bears Coach's Experience 2006 Season
- Report: Average Age of Starting Quarterback Teams with Superbowl Appearances
    - Dashboard: Colts Quarterback Age 2006 Season
- Report: Average Order in the Draft Pick (Before Trades) Teams with Superbowl Appearances
    - Dashboard: Colts Draft Order 2006 Season
- Report: Average Playoff Seed Teams with Superbowl Appearances
    - Dashboard: Colts Playoff Seed 2006 Season
- Report: Coach's Average Years in the Head Coaching Industry Teams with Superbowl Appearances
    - Dashboard: Colts Coach's Experience 2006 Season

Mobile Layout

One feature of Power BI is that it is a mobile application accessible. Dashboards created on a PC can be rearranged in a mobile layout allowing people to access the story told in the dashboard anywhere. Despite my dashboard not being used by people other than myself, and I do not need to access it from my mobile device, I wanted to develop the mobile layout because I want to get experience using as many features of Power BI as possible.

For the design of my mobile layout, I decided to think of my dashboard visuals as groups. Mobile device screens limit the amount of content a user can view at once, so I wanted to ensure that visuals that are directly related to each other/desirable to compare are neighboring in my mobile layout. This is particularly important for my dashboard because the story it is telling is a comparison between the Chicago Bears and Indianapolis Colts.

I would also like to note, although not required, I did include every visual in my dashboard in my mobile layouts. The reason I chose to do this is so potential mobile dashboard users are not confused searching for a visual from the PC dashboard on the mobile dashboard but are unable to find it.

Appendix:

Hex Code Data

NFL Colors

https://www.schemecolor.com/national-football-league-nfl-logo-colors.php

- Got the HEX codes that correspond with the NFL/AFC/NFC logo

Team Colors

https://teamcolorcodes.com/nfl-team-color-codes/

- Got the HEX codes that correspond with team colors

Table Data

Coaches Table & NFL Coaches Table

https://www.pro-football-reference.com/years/2002/coaches.htm

- Downloaded each file from site as html file and had to convert to excel spreadsheet to use in Power BI (21 files – 1 file for each season)

https://www.pro-football-reference.com/coaches/

- Downloaded file from site as a txt file (1 file)

Conference Table

Created excel file myself using my knowledge of which teams are in each conference and division (1 file)

Draft Order Table

https://www.prosportstransactions.com/football/DraftTrades/Years/2002.htm

- Used python code to scrape this website and download data as a CSV file (1 file)
  - Link to python code: https://colab.research.google.com/drive/1tvyR31-xjkMhrFO47eC6jE4PV6ibXn2w?usp=sharing
- The reason I scraped this website instead of downloading an excel/csv table is because this website did not have the option to download on the page

- I did not scrape the other tables I got from websites because I wanted to demonstrate my ability to use Power Query Editor to append tables, keep columns, etc.
- In the data I scraped I did most of the data cleaning in my python code and did not use Power Query Editor, which is not the point of my senior project.
    - I would like to note that I recognize cleaning data before importing it in Power BI creates better performance.

Playoffs Table

https://www.pro-football-reference.com/years/2002/playoffs.htm

- Downloaded each file from site as html file and had to convert to excel spreadsheet to use in Power BI (42 files – 1 file for each conference each season)

Quarterbacks Table & Quarterbacks List Table

https://www.pro-football-reference.com/years/2002/passing.htm

- Downloaded each file from site as html file and had to convert to excel spreadsheet to use in Power BI (21 files – 1 file for each season)

Standings Table

https://www.pro-football-reference.com/years/2002/

- Downloaded each file from site as html file and had to convert to excel spreadsheet to use in Power BI (42 files – 1 file for each conference each season)

Superbowl Table & Superbowl Results Table

https://www.pro-football-reference.com/super-bowl/

- Downloaded file from site as html file and had to convert to excel spreadsheet to use in Power BI (1 file)