SQL Warriors: Johnson L, Long N, Sebastian D, Kimberly V, Alexis R, Eli S

https://lucid.app/lucidchart/53c77ae5-d54b-4428-a3dc-1d2b5528171b/edit?viewport_loc=-2738%2C-667%2C3253%2C1536%2C0_0&invitationId=inv_e475cf67-9035-400f-a8d8-a18eea6503c0

MILESTONE 1

The job position we chose is a **Business Intelligence Analyst** position with the Arizona Cardinals Football team. Through our research, we learned that the Arizona Cardinals embrace data analytics to enhance their business processes, running from analyzing fan engagement to optimizing ticket sales. They are an organization in the NFL that uses business intelligence tools to drive data-driven decisions for a competitive advantage.

The reason we chose this position is that it involves data analytics and impactful decision-making within the sports industry. This role integrates SQL, Power BI, and Python, making it an ideal position that fits our college backgrounds, career goals, and interests. With this job position, we are able to fulfill our career goal of developing hands-on experience with applying technology to real-world scenarios. We also get to develop our interest in collaborating with other individuals, such as stakeholders, consultants, vendors, and individuals from the finance department. This interest of ours gives us the opportunity to enhance our interpersonal and leadership skills while contributing to our company's efficiency and future. Each member of our team brings a variety of skills and experiences that align with the job requirements. We've taken relevant courses such as database management systems (SQL) and introduction to business programming (Python). In this course, we learned the basics of SQL and its variety of statements. For example, we learned how to create a Normalized ERD with defined attributes,

entities, primary keys, and foreign keys. We also learned about the Normalization Process; organizing data and reducing redundancy. Creating an ERD gave us a blueprint to ensure our database was detailed and accurate.

Our shared interest in data extraction and analysis aligns closely with the team's goal of making data-driven decisions. By developing and optimizing SQL queries, we ensured efficient access to accurate information. This supported reporting, trend analysis, and strategic planning efforts across the team, helping to drive better business outcomes.

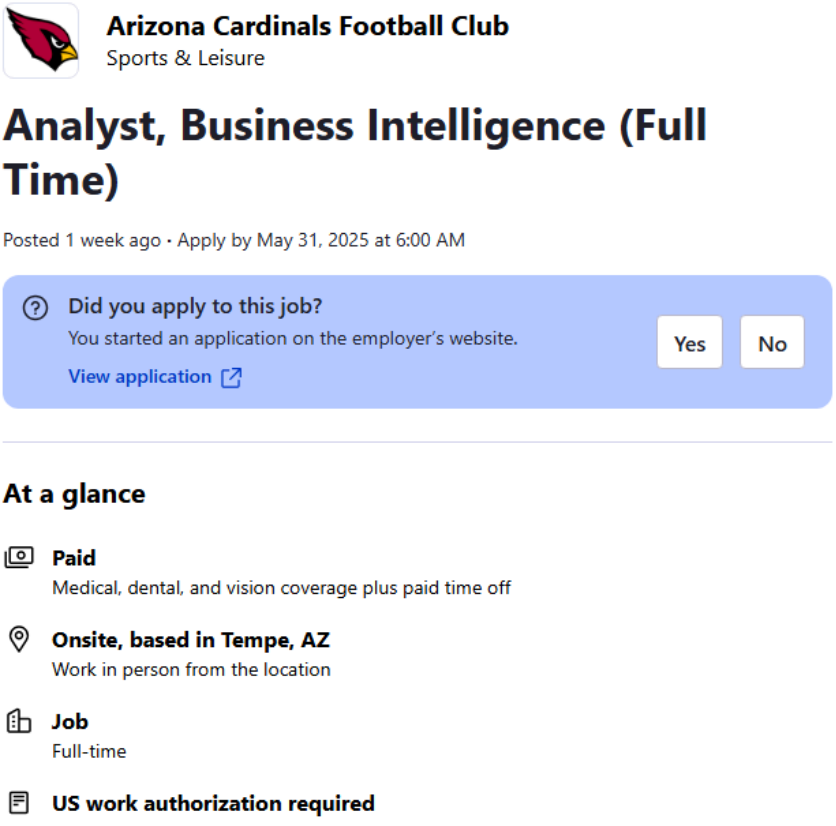
A member of the team has experience in building Power BI Dashboards to visualize key metrics and support data-driven decision-making. This is where he used DAX for specific calculations to create different types of bar charts, line graphs, and pie charts. Also, this member has a Certified SQL Developer certification, demonstrating proficiency in SQL fundamentals, including data querying, table joins, subqueries, and data manipulation.

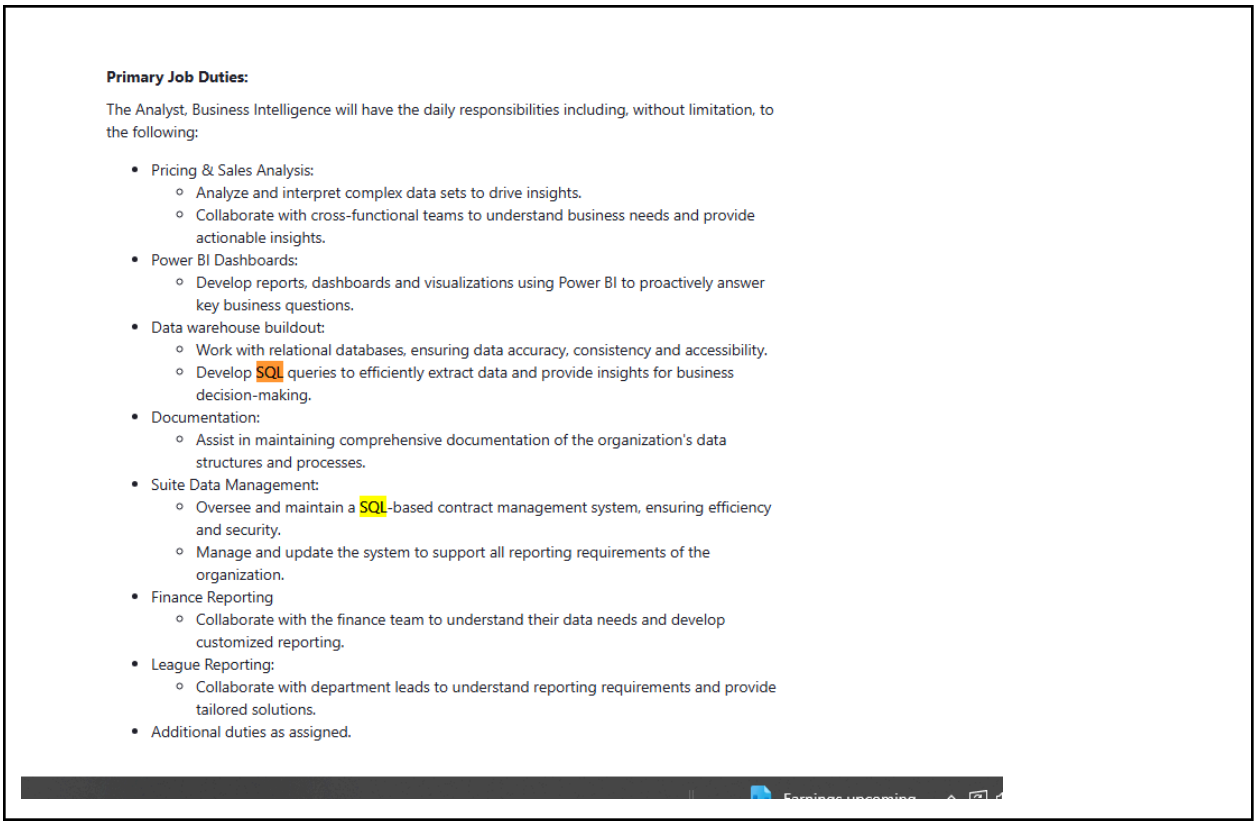
A member of the team has experience dealing with the Data warehouse buildout and Documentation, where for the Data warehouse buildout, he tracked and organized inventory, then recorded data in Excel. Using SQL to then sort the information and create a table for different categories of clothing, which would also fit the category documentation by creating the data tables.

A member of the team has experience with managing competing priorities, coordinating across diverse teams, and ensuring timely delivery of project deliverables while maintaining professionalism and clear communication throughout the project lifecycle.

Handshake job listing(external link)

https://recruiting2.ultipro.com/ARI1001CARD/JobBoard/bfe37693-2c0a-46b6-af4e-f984b08a8e7f/OpportunityDetail?opportunityId=b9f73a55-9f9b-43a3-b287-6eac0908767a&sourceId=3c66dea0-27a6-4209-9110-74503e50c909&gh_src=Handshake&iisn=Handshake&iis=Handshake&src=Handshake&source=Handshake&source=Handshake&ref=Handshake&utm_medium=Handshake&referral=Handshake&utm_source=Handshake&__jvst=Handshake&__jvsd=Handshake&sourceDetails=Handshake&trid=Handshake&lever-source%5B%5D=Handshake&rb=Handshake&jobBoardSource=Handshake&channel=Handshake&rcid=Handshake





Primary Job Duties:

The Analyst, Business Intelligence will have the daily responsibilities including, without limitation, to the following:

- Pricing & Sales Analysis:
  - Analyze and interpret complex data sets to drive insights.
  - Collaborate with cross-functional teams to understand business needs and provide actionable insights.
- Power BI Dashboards:
  - Develop reports, dashboards and visualizations using Power BI to proactively answer key business questions.
- Data warehouse buildout:
  - Work with relational databases, ensuring data accuracy, consistency and accessibility.
  - Develop SQL queries to efficiently extract data and provide insights for business decision-making.
- Documentation:
  - Assist in maintaining comprehensive documentation of the organization's data structures and processes.
- Suite Data Management:
  - Oversee and maintain a SQL-based contract management system, ensuring efficiency and security.
  - Manage and update the system to support all reporting requirements of the organization.
- Finance Reporting
  - Collaborate with the finance team to understand their data needs and develop customized reporting.
- League Reporting:
  - Collaborate with department leads to understand reporting requirements and provide tailored solutions.
- Additional duties as assigned.

MILESTONE 2:

SQL Warriors: Johnson L, Long N, Sebastian D, Kimberly V, Alexis R, Eli S

Part 1: Database Background

Business Operation

We will create a database that tracks game-day operations and fan engagement for the Business Intelligence Analyst role with the Arizona Cardinals. This database will collect data from fan attendance, merchandise purchases, and concession sales during home games. This allows for a better understanding of what people are doing during game hours

Data Requirements

The Fan entity stores personal information and contact details for each fan. The Game entity captures game schedules, opponent information, and event-specific details like date, time, and stadium location.

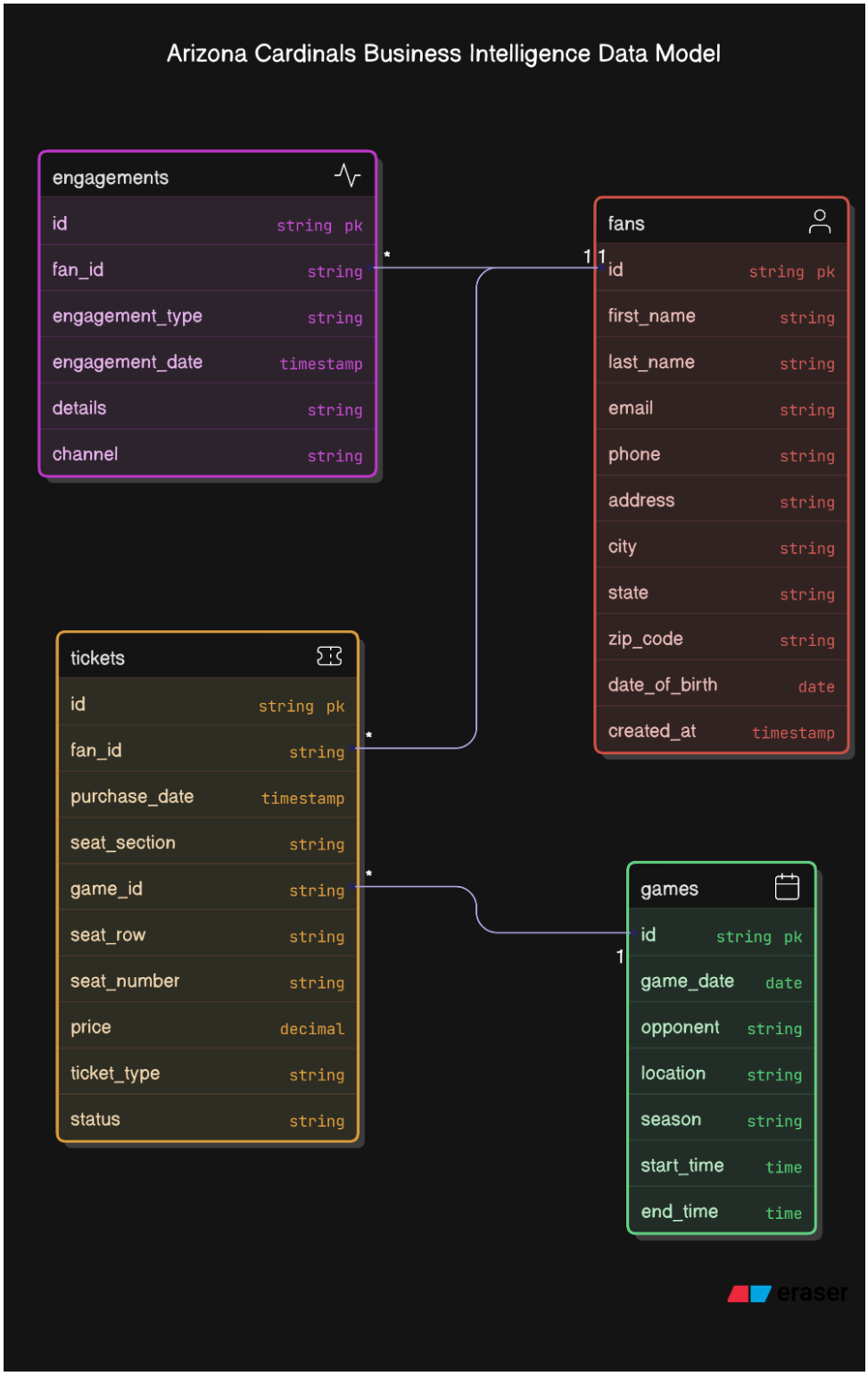
- The Ticket entity links fans to specific games, recording ticket purchase history, seat location, ticket type, and purchase date. Lastly,

- The Engagement entity tracks sales, loyalty program participation, newsletter signups, fan merchandise transactions, food, and beverages

Operational Context

The database supports the Arizona Cardinals' business operation as it collects the data of what is going on with their customers. With the collection of data, the Cardinals can make better decisions as they would know what the people want more.

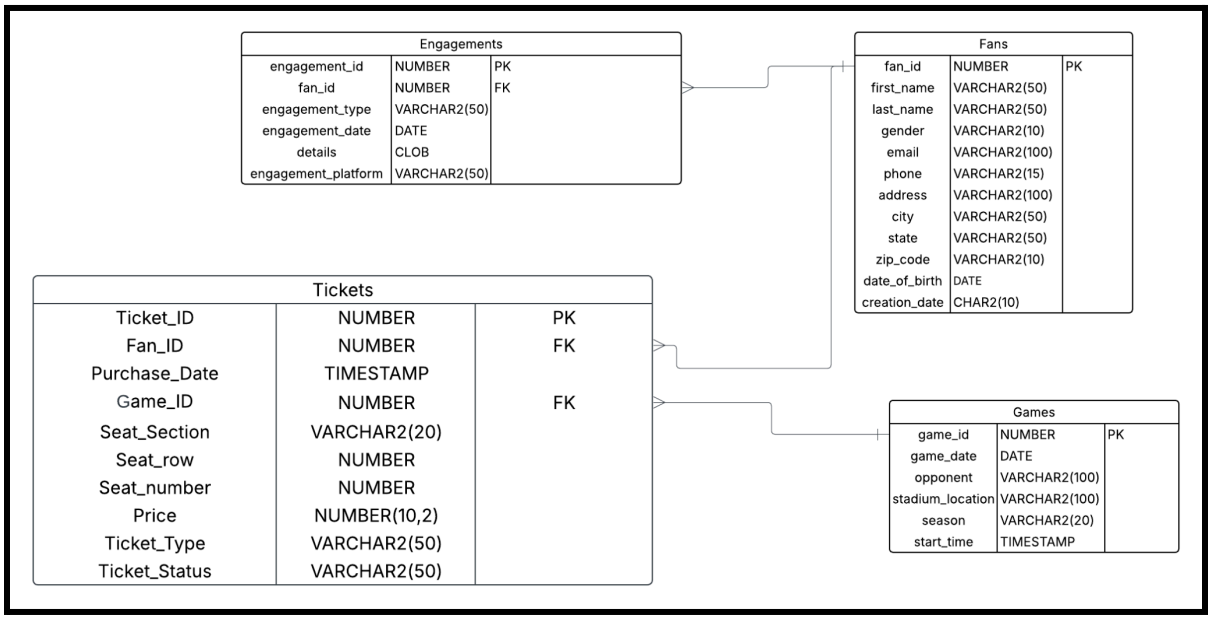
Part 2: Draft ERD: Include the draft ERD generated by DiagramGPT.



Part 3: ERD Improvements(List at least three specific improvements, with brief reasoning for each)

1) We want to make the ID attribute for each table more specific so that it can be differentiated in our ERD(ex, Engagement_ID, Fan_ID, Ticket_ID, Game_ID). We also want to rename some specific table attributes listed below because we want our ERD to have more entity clarity.
    - Rename the channel entity(engagement table) to engagement_platform, so it is less confusing
    - Rename the channel created_at(fans table) to creation_Date
    - Rename status(tickets table) to ticket_status
    - Rename location(games table) to stadium_location
2) We are removing the end_time(game table) attribute because it is unnecessary, and this data is harder to retrieve because there is no set time for a football game to end.
3) We added data types for each attribute to make the database easier to build and use. Choosing data types like string, date, timestamp, and decimal helps store information correctly and keeps the database organized for future SQL work.
4) We also decided to add the attribute gender to the fans table to have better customer insight. This addition will support more effective organizational marketing, personalized promotions, and improved event planning based on fan demographics.

PART 4:Final ERD: Insert the PNG or JPEG image of the finalized ERD created in LucidChart.



Example Data Tables: Include at least three rows of example data per entity.

FAN

Fan_ID	First_name	Last_name	Gender	Email	Phone
001	Johnson	Luong	Male	jluong1020@gmail.com	(669) 274-5819
002	Leann	Nguyen	Female	Leannnguyen05@gmail.com	(480) 265 7081
0003	Sebastian	Dang	Male	dangsebastian212@gmail.com	(602) 895 1253

FAN continued

Address	City	State	DOB	creation_date
123 Blossom Hill Rd	San Jose	California	06/05/1999	01/04/2020
473 Trevor Street	Tucson	Arizona	07/14/1974	08/09/2010
8371 Holmes Ave	Mesa	Arizona	08/31/1987	01/21/2014

Engagements

engagement_Id	Fan_Id	engagement_type	engagement_date	details	engagement_platform
1	003	Survey Participation	02/18/2017	Fan Satisfaction based on overall experience	Mobile App
2	001	Contest Entry	09/08/2022	Entered to win an autographed football from	Website

				Hall of Fame player	
3	002	Merchandise Purchase	11/14/2015	Bought a Jersey and hat	Stadium's Team Store

Tickets:

ticket_id	Fan_ID	Purchase_Date	Game_ID	Price	Ticket_Type	Ticket_Status	Seat_Section	Seat_row	Seat_number
3001	001	2024-06-25	103	125.00	Standard	Confirmed	Section 451	7	1
3002	002	2023-05-25	101	225.00	VIP	Cancelled	Section 134	4	19
3003	003	2024-04-22	102	125.00	Standard	Confirmed	Section 125	1	5

GAMES:

Game_ID	Game_date	Opponent	Stadium_location	Season	Start Time
103	2024-09-07	Las Vegas Raiders	Statefarm	2024-2025	1:25 PM
101	2023-11-18	Dallas Cowboys	AT&T Stadium	2023-2024	5:20 PM
102	2024-10-06	San Francisco 49ers	Levis Stadium	2024-2025	1:05 PM

PART 6: Final review on the ERD along with the example data(Be prepared to use example scenarios to explain all your ERD definitions.)

- Johnson Luong is an Arizona Cardinals Fan and they created their account on 09/08/2022 to enter themselves to win an autographed football from Hall of Fame player Larry Fitzgerald. He purchased a ticket on June 25th, 2024, to watch the Arizona Cardinals play against the Las Vegas Raiders. This game occurred on August 7th, 2024, at State Farm Stadium at 1:25 PM. The tickets he purchased were for section 451, row 7, seat number
- Leann Nguyen is an Arizona Cardinals fan, and she created her account on November 14th, 2015, when she purchased a jersey and a hat. She purchased a ticket on May 25th, 2023, to watch the Arizona Cardinals play the Cowboys. This game occurred on November 18th, 2023, at 5:20 pm, but Leaann could not make it to the game, so she cancelled ahead of time. She was supposed to sit in Section 451, Row 7, Seat 1.
- Sebastian Dang is an Arizona Cardinals fan, and he created his account on February 19th, 2017, when he filled out a Fan Satisfaction survey on the mobile app. He purchased a ticket to watch the Arizona Cardinals play against the San Francisco 49ers on October 6th, 2024, at 1:05 pm. His seats were located in Section 125, row 1, seat 5.

MILESTONE 3:

Step 1: Implement Your Database in Oracle LiveSQL

```
DROP TABLE Engagements CASCADE CONSTRAINTS;
DROP TABLE Tickets CASCADE CONSTRAINTS;
DROP TABLE Games CASCADE CONSTRAINTS;
DROP TABLE Fans CASCADE CONSTRAINTS;


CREATE TABLE Fans (
   fan_id         NUMBER PRIMARY KEY,
   first_name     VARCHAR2(50),
   last_name      VARCHAR2(50),
   gender         VARCHAR2(10),
```

```
    email           VARCHAR2(100),
    phone           VARCHAR2(15),
    address         VARCHAR2(100),
    city            VARCHAR2(50),
    state           VARCHAR2(50),

    zip_code        VARCHAR2(10),
    date_of_birth   DATE,
    creation_date   TIMESTAMP
);
-- Games Table
CREATE TABLE Games (
    game_id           NUMBER PRIMARY KEY,
    game_date         DATE,
    opponent          VARCHAR2(100),
    stadium_location  VARCHAR2(100),
    season            VARCHAR2(20),
    start_time        TIMESTAMP  -- Changed from TIME to TIMESTAMP
);

-- Engagements Table
CREATE TABLE Engagements (
    engagement_id       NUMBER PRIMARY KEY,
    fan_id              NUMBER,
    engagement_type     VARCHAR2(50),
    engagement_date     TIMESTAMP,
    details             CLOB,
    engagement_platform VARCHAR2(50),
    CONSTRAINT fk_engagement_fan FOREIGN KEY (fan_id) REFERENCES Fans(fan_id)
);

-- Tickets Table
CREATE TABLE Tickets (
    ticket_id      NUMBER PRIMARY KEY,
    fan_id         NUMBER,
    purchase_date  TIMESTAMP,
    game_id        NUMBER,
    seat_section   VARCHAR2(20),
    seat_row       NUMBER,
    seat_number    NUMBER,
    price          NUMBER(10, 2),  -- Changed from DECIMAL to NUMBER
    ticket_type    VARCHAR2(50),
    ticket_status  VARCHAR2(50),
    CONSTRAINT fk_ticket_fan FOREIGN KEY (fan_id) REFERENCES Fans(fan_id),
```

```
    CONSTRAINT fk_ticket_game FOREIGN KEY (game_id) REFERENCES Games(game_id)
);
```

--For Fans

```
INSERT INTO Fans (fan_id, first_name, last_name, gender, email, phone, address, city, state, date_of_birth, creation_date)
VALUES (1, 'Johnson', 'Luong', 'Male', 'jluong1020@gmail.com', 6692745819, '123 Blossom Hill Rd', 'San Jose', 'California', TO_DATE('06/05/1999', 'MM/DD/YYYY'), TO_TIMESTAMP('01/04/2020', 'MM/DD/YYYY'));

INSERT INTO Fans (fan_id, first_name, last_name, gender, email, phone, address, city, state, date_of_birth, creation_date)
VALUES (2, 'Leann', 'Nguyen', 'Female', 'Leannnguyen05@gmail.com', 4802657081, '473 Trevor Street', 'Tucson', 'Arizona', TO_DATE('07/14/1974', 'MM/DD/YYYY'), TO_TIMESTAMP('08/09/2010', 'MM/DD/YYYY'));

INSERT INTO Fans (fan_id, first_name, last_name, gender, email, phone, address, city, state, date_of_birth, creation_date)
VALUES (3, 'Sebastian', 'Dang', 'Male', 'dangsebastian212@gmail.com', 6028951253, '8371 Holmes Ave', 'Mesa', 'Arizona', TO_DATE('08/31/1987', 'MM/DD/YYYY'), TO_TIMESTAMP('01/21/2014', 'MM/DD/YYYY'));
```

--For Engagements

```
INSERT INTO Engagements (engagement_id, fan_id, engagement_type, engagement_date, details, engagement_platform)
VALUES (1, 3, 'Survey Participation', TO_TIMESTAMP('02/18/2017', 'MM/DD/YYYY'), 'Fan Satisfaction based on overall experience', 'Mobile App');

INSERT INTO Engagements (engagement_id, fan_id, engagement_type, engagement_date, details, engagement_platform)
VALUES (2, 1, 'Contest Entry', TO_TIMESTAMP('09/08/2022', 'MM/DD/YYYY'), 'Entered to win an autographed football from Hall of Fame player', 'Website');

INSERT INTO Engagements (engagement_id, fan_id, engagement_type, engagement_date, details, engagement_platform)
VALUES (3, 2, 'Merchandise Purchase', TO_TIMESTAMP('11/14/2015', 'MM/DD/YYYY'), 'Bought a Jersey and hat', 'Stadium’s Team Store');
```

--For games

```
INSERT INTO Games (game_id, game_date, opponent, stadium_location, season, start_time)
```

```
VALUES (103, TO_DATE('2024-09-07', 'YYYY-MM-DD'), 'Las Vegas Raiders', 'Statefarm',
'2024-2025', TO_TIMESTAMP('2024-09-07 13:25:00', 'YYYY-MM-DD HH24:MI:SS'));

INSERT INTO Games (game_id, game_date, opponent, stadium_location, season, start_time)
VALUES (101, TO_DATE('2023-11-18', 'YYYY-MM-DD'), 'Dallas Cowboys', 'AT&T Stadium',
'2023-2024', TO_TIMESTAMP('2023-11-18 17:20:00', 'YYYY-MM-DD HH24:MI:SS'));

INSERT INTO Games (game_id, game_date, opponent, stadium_location, season, start_time)
VALUES (102, TO_DATE('2024-10-06', 'YYYY-MM-DD'), 'San Francisco 49ers', 'Levis Stadium',
'2024-2025', TO_TIMESTAMP('2024-10-06 13:05:00', 'YYYY-MM-DD HH24:MI:SS'));
```

--For tickets

```
INSERT INTO Tickets (ticket_id, fan_id, purchase_date, game_id, price, ticket_type,
ticket_status, seat_section, seat_row, seat_number)
VALUES (3001, 1, TO_TIMESTAMP('2024-06-25', 'YYYY-MM-DD'), 103, 125.00, 'Standard',
'Confirmed', 'Section 451', 7, 1);

INSERT INTO Tickets (ticket_id, fan_id, purchase_date, game_id, price, ticket_type,
ticket_status, seat_section, seat_row, seat_number)
VALUES (3002, 2, TO_TIMESTAMP('2023-05-25', 'YYYY-MM-DD'), 101, 225.00, 'VIP',
'Cancelled', 'Section 134', 4, 19);

INSERT INTO Tickets (ticket_id, fan_id, purchase_date, game_id, price, ticket_type,
ticket_status, seat_section, seat_row, seat_number)
VALUES (3003, 3, TO_TIMESTAMP('2024-04-22', 'YYYY-MM-DD'), 102, 125.00, 'Standard',
'Confirmed', 'Section 125', 1, 5);
```

Step 2: Design Meaningful Analysis Questions

1) Which games were able to generate the highest ticket revenue, and what is the number of fan engagements for that game?

This question helps us identify which games produce the highest ticket revenue and see if there is high fan engagement. Being able to compare revenue alongside fan engagement provides valuable insights for the Arizona Cardinals organization because they can be able to understand the factors of higher fan engagement. These factors can include things like marketing campaigns, giveaways, and opponents(rivals). These insights will help the Arizona Cardinals make better data-driven decisions for the future.

2) **What engagement platform is most popular amongst fans?** This question enables the Cardinals to identify which platforms are used actively by fans. This will help them identify what drives fan engagement the most and tailor their marketing strategies to increase to fan loyalty. They can do things like analyze demographics and try targeting a wider audience of fans.

This question helps the organization understand where and how fans are interacting. By analyzing platform preferences by location, the team can make informed decisions about where to invest in digital vs. in-person engagement and tailor outreach strategies based on regional behavior patterns.

Step 3: Write and Document SQL Queries

--Answer to Question 1: Which games were able to generate the highest ticket revenue, and what is the number of fan engagements for that game?

```
SELECT g.game_id,g.opponent,g.game_date, (So to answer this question, we need the game data such as ID, opponent, and date)
SUM(t.price) AS total_revenue,     –(To calculate the total revenue, we need to sum the ticket prices for each game)
    COUNT(e.engagement_id) AS total_engagements  -- (We also need to count engagements for the fans who bought tickets for that game)
FROM games g
JOIN tickets t ON g.game_id = t.game_id   (We connect each ticket ot the game it was purchased for since the ticket prices are tied to the record of tickets sold)
LEFT JOIN engagements e ON t.fan_id = e.fan_id  (Join engagements to fans who bought tickets to a game. We use left join specifically so that fans without any engagements are still accounted for, but it'll just show up as 0 engagements)
GROUP BY g.game_id, g.opponent, g.game_date  (Since we are aggregating sum and count, we need to group them by each unique game. This ensures that we get a row per game instead of per ticket)
ORDER BY total_revenue DESC;  (Since we're focusing on the games with the highest revenue, we need this data to pop up at the top)
```

--Answer to Question 2: What engagement platform is most popular amongst fans?

```
SELECT engagement_platform, COUNT(*) AS total_uses FROM engagements
 (Takes the platform they engaged on and then adds a count to it and labels it as total_uses)
WHERE engagement_platform IS NOT NULL      (This is where it would filter out the information that is NULL)
GROUP BY engagement_platform (Groups all the rows into a table by each platform)
ORDER BY total_uses DESC;(This is where it would show the Ranking of platforms by popularity)
```

SQL Worksheet

```
SELECT g.game_id,g.opponent,g.game_date,SUM(t.price) AS total_revenue,                    -- Total ticket revenue
    COUNT(e.engagement_id) AS total_engagements      -- Total fan engagements for ticket buyers
FROM games g
JOIN tickets t ON g.game_id = t.game_id
LEFT JOIN engagements e ON t.fan_id = e.fan_id
GROUP BY g.game_id, g.opponent, g.game_date
ORDER BY total_revenue DESC;
```

GAME_ID	OPPONENT	GAME_DATE	TOTAL_REVENUE	TOTAL_ENGAGEMENTS
101	Dallas Cowboys	18-NOV-23	225	1
102	San Francisco 49ers	06-OCT-24	125	1
103	Las Vegas Raiders	07-SEP-24	125	1

Download CSV

3 rows selected.

ENGAGEMENT_PLATFORM	TOTAL_USES
Mobile App	1
Website	1
Stadium’s Team Store	1

Step 4: Incorporate Feedback

We received feedback that our data types for the ERD did not match up with the example data. Our original data type for phone numbers was integer but this was incorrect because we formatted it as (XXX)-XXX-XXX. After realizing this, in our SQL code, we implemented it with the data type VarChar. Our original date format was MM/DD/YYYY, but we ended up formatting it as DD-MM-YYYY for better readability. We also made the mistake of keeping the zipcodes data type as an integer, it should be a varchar because it’s not used for calculations. We also realized that we need to be

inclusive of other countries' zip codes, so we changed the length from varchart length from 5 to 10.

When creating our tables, we violated a referential integrity constraint because we created the tickets table before the games table. This was an issue because the ticket table contained a foreign key (game_id) that references the game table. If we did try to insert tickets before the games are defined, it would essentially be like selling a ticket to an event that the system does not recognize. So, to maintain data integrity, we changed our code and created the games table first.