

Woods Procise

Professor Bow

Database Design

14 December 2025

## Databook Part 3: The Database

### Big Ten Football Conference

#### **I. Database Description**

Commonly, databases are designed to perform a wide array of tasks centered around data. They are created to serve the primary purpose of storing, structuring, and organizing oftentimes large and complex datasets. If databases successfully perform these operations, it enables the end users to access the readily available stored data, and retrieve the data to ultimately analyze, manipulate, and quite frankly, use the data for whatever purpose they would like.

This particular project will aim to create a user centered Big Ten Conference football database. The goal of this database will be to serve the purpose of storing, structuring, and organizing data for each football team in the Big Ten Conference. Specifically, the context of this Big Ten football database stores and collects data from the 2021 college football season up until the current 2025 college football season. Due to the everchanging college football landscape in recent years due in large part to conference realignment and the dramatic rise of name, image, and likeliness (NIL), this database seeks to store valuable data for Big Ten football teams ranging from historical information on each team by tracking records, final rankings, and bowl game appearances from past seasons, current up to date player and team statistics, attendance by game and season, as well as in depth financial data on each teams investment in their football program through various expenditures, and the total annual athletic revenue generated by each football program for their respective universities and the Big Ten Conference as a whole.

The data stored in this Big Ten Conference Football is extremely valuable and useful for a multitude of reasons and to a wide audience. End Users include the Big Ten Conference leadership and administration, Big Ten University athletic departments and administrators, Big Ten football teams, mostly players and coaches, as well as football analysts and college football fans. All these likely Big Ten football database users are all seeking various information specific to meet their interests and needs, that being a strength of this database as it stores a vast amount of information that is useful and of interest to many end users.

Overall, this Big Ten football database primary purposes and uses include providing historical insights on the evolution of the conference and its teams, statistical and predictive analysis on teams, coaches, and players, as well as financial tracking for universities, athletic departments, and the Big Ten Conference. Often, the purpose of this database and what exact information matters to which end users aligns with each other.

Specifically, historical insights are likely of high interest for fans who are interested in learning more about their teams in addition to football analysts and game commentators who have a strong desire and need to know historical information about the conference, teams, and players. In turn, the statistical and predictive analysis that this database enables through the tracking of player and team statistics for each season is of high value to most players and coaches as they use it to gain insights on their strengths and weaknesses so that they can formulate successful gameplans and improve in the future. All the player statistical data stored in this database could also be of interest to NFL scouts to gain data driven insights on a player's performance, and development during their time in college. Aside from those key uses, this database is imperative to Big Ten Conference and its universities as their main goal is to drive up revenue and fan engagement. From this database, they gain insights on fan attendance on a game to game and yearly basis to evaluate the state of each football program and the conference. This database is extremely useful for conference and university leadership as well through its financial tracking of spending allocations and revenue flow

providing thorough financial evaluation. Overall, the fourteen tables in this database, CONFERENCE, TEAM, STADIUM, COACH, PLAYER, GAME, GAMEPARTICIPATION, OFFENSIVEPLAYERSTATS, DEFENSIVEPLAYERSTATS, SEASONSTATS, SCHEDULE, TEAMSPENDING, TEAMREVENUE, and CONFERENCE REVENUE create relationships that store and update critical information that is highly beneficial to the Big Ten Conference, its football teams, players, coaches, and administrations as well as college football analysts, fans, and any end users in helping them accomplishing their goals.

## II. Data Dictionary

ENTITY	ATTRIBUTE NAME	CONTENTS	DATA TYPE	FORMAT	RANGE	REQUIRED	PK or FK	Reference
CONFERENCE	CONF_ID	Conference ID Code	INT(2)	99	0-50	YES	PK	
	CONF_NAME	Conference Name	VARCHAR (5)	XXXXXX		YES		
	CONF_FOUNDED_YEAR	Conference Founded Year	DATE	yyyy	1892 - 2013	YES		
	CONF_HQ_CITY	Conference Headquarters City	VARCHAR (25)	Xxxxxxx		YES		
	CONF_HQ_STATE	Conference Headquarters State	CHAR(2)	XX		YES		
	CONF_COMMISSIONER_NAME	Conference Commissioner Name	VARCHAR (30)	Xxxxxxx		YES		

TEAM	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	PK	
	CONF_ID	Conference ID Code	INT(2)	99	0-50	YES	FK	CONFERENCE
	TEAM_NAME	Team Name	VARCHAR (20)	Xxxxxxx		YES		
	TEAM_MASCOT	Mascot Name	VARCHAR (25)	Xxxxxxx		YES		

	TEAM_CITY	Team Location City	VARCHAR (25)	Xxxxxxx		YES		
	TEAM_STATE	Team Location State	CHAR(2)	XX		YES		
	TEAM_COLORS	Team Colors	VARCHAR	Xxxxxxx		YES		
STADIUM	STADIUM_ID	Stadium ID Code	INT(3)	999	0-500	YES	PK	
	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	STADIUM_NAME	Stadium Name	VARCHAR (50)	Xxxxxxx		YES		
	STADIUM_CAPACITY	Stadium Full Capacity	INT(6)	999999	0-199999	YES		
	STADIUM_SURFACE_TYPE	Stadium Field Surface Type	VARCHAR (5)	Xxxxx	“Turf”, “Grass”	YES		
	STADIUM_CITY	Stadium City Location	VARCHAR (25)	Xxxxxxx		YES		

	STADIUM_STATE	Stadium State Location	CHAR(2)	XX		YES		
	STADIUM_ZIP	Stadium Zip Code	CHAR(5)	99999	0000099999	YES		
COACH	COACH_ID	Coach ID Code	INT(4)	999	0-9999	YES	PK	
	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	COACH_FNAME	Coach First Name	VARCHAR (30)			YES		
	COACH_LNAME	Coach Last Name	VARCHAR (30)			YES		
	COACH_POS_TITLE	Coach Position Title	VARCHAR (40)			YES		
	COACH_HIRE_DATE	Coach Hire Date	DATE	yyyy-mm-dd		YES		
	COACH_SALARY	Coach Salary	DECIMAL (10,2)	99,999,999.99	0.01 – 999999999.99	YES		
PLAYER	PLAYER_ID	Player ID Code	INT(7)	9999999	0-9999999	YES	PK	

	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	PLAYER_FNAME	Player First Name	VARCHAR (30)	Xxxxxxx		YES		
	PLAYER_LNAME	Player Last Name	VARCHAR (30)	Xxxxxxx		YES		

	PLAYER_POS	Player Position	VARCHAR (20)	Xxxxxxx		YES		
	PLAYER_JERSEY_NUM	Player Jersey Number	INT(2)	99	0-99	YES		
	PLAYER_CLASS_YEAR	Player Class Year	VARCHAR (20)	Xxxxxxx		YES		
	PLAYER_BIRTH_DATE	Player Birth Date	DATE	yyyy-mm-dd		YES		
	PLAYER_HOMETOWN_CITY	Player Hometown City	VARCHAR (25)	Xxxxxxx		YES		
	PLAYER_HOMETOWN_STATE	Player Hometown State	VARCHAR (2)	XX		YES		
	PLAYER_HOMETOWN_ZIP	Player Hometown Zip Code	CHAR(5)	99999	0000099999	YES		
GAME	GAME_ID	Game ID Code	INT(7)	9999999	0-9999999	YES	PK	
	SEASON_YEAR	Season Year	INT(4)	9999	2021 – 2025	YES	FK	SEASONSTATS
	GAME_WK_NUM	Game Week Number	INT(2)	99	0-16	YES		
	HOME_TEAM_ID	Home Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	AWAY_TEAM_ID	Away Team ID Code	INT(3)	999	0-500	YES	FK	TEAM

	STADIUM_ID	Stadium ID Code	INT(3)	999	0-500	YES	FK	STADIUM
	GAME_DATE	Game Date	DATE	yyyy-mm-dd		YES		
	GAME_KICKOFF_TIME	Game Kickoff Time	DATETIME	hh-mm		YES		
	GAME_ATTENDANCE	Total Attendance	INT(6)	999999	0-199999	YES		

	GAME_HOME_SCORE	Home Team Score	INT(2)	99	0-99	YES		
	GAME_AWAY_SCORE	Away Team Score	INT(2)	99	0-99	YES		
	GAME_TV_NETWORK	Game Broadcast Network	VARCHAR (20)	Xxxxxxx		YES		

GAMEPARTICIPATION	PLAYER_ID	Player ID Code	INT(7)	9999999	0-9999999	YES	PK/ FK	PLAYER
	GAME_ID	Game ID Code	INT(7)	9999999	0-9999999	YES	PK/ FK	GAME
	GP_TOT_MIN_PLAYED	Game Total Minutes Played	INT(2)	99	0-80	YES		
	GP_GAME_RATING	Game Viewership Rating	DECIMAL (3,1)	99.9	0-99.9	YES		

OFFENSIVEPLAYERSTATS	PLAYER_ID	Player ID Code	INT(7)	9999999	0-9999999	YES	PK, FK	PLAYER
----------------------	-----------	----------------	--------	---------	-----------	-----	-----------	--------

	SEASON_YEAR	Season Year	INT(4)	9999	2021 – 2025	YES	PK	
	OFF_GAMES_PLAYED	Offensive Player Games Played	INT(2)	99	0-16	YES		
	OFF_PASS_YARDS	Offensive Player Passing Yards	INT(4)	9,999	0-5,000	YES		
	OFF_RUSH_YARDS	Offensive Player Rush Yards	INT(4)	9,999	0-3,000	YES		
	OFF_RECEIVING_YARDS	Offensive Player Receiving Yards	INT(4)	9,999	0-3,000	YES		
	OFF_TOUCHDOWNS	Offensive Player Touchdowns	INT(2)	99	0-99	YES		
	OFF_INTERCEPTIONS_THROWN	Offensive Player Interceptions Thrown	INT(2)	99	0-50	YES		

	OFF_FUMBLES_LOST	Offensive Player Fumbles Lost	INT(2)	99	0-30	YES		
DEFENSIVEPLAYERSTATS	PLAYER_ID	Player ID Code	INT(7)	9999999	0-9999999	YES	PK/FK	PLAYER
	SEASON_YEAR	Season Year	INT(4)	9999	2021 – 2025	YES	PK	



	DEF_GAMES_PLAYED	Defensive Player Games Played	INT(2)	99	0-16	YES		
	DEF_INTERCEPTIONS	Defensive Player Interceptions	INT(2)	99	0-30	YES		
	DEF_TACKLES	Defensive Player Tackles	INT(3)	999	0-500	YES		
	DEF_SACKS	Defensive Player Sacks	INT(2)	99	0-30	YES		
	DEF_FORCED_FUMBLES	Defensive Player Forced Fumbles	INT(2)	99	0-30	YES		
	DEF_PASS_BREAKUPS	Defensive Player Pass Break Ups	INT(2)	99	0-50	YES		
SEASONSTATS	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	PK/FK	TEAM
	SEASON_YEAR	Season Year	INT(4)	9999	2021 – 2025	YES	PK	
	TEAM_WINS	Team Yearly Wins	INT(2)	99	0-16	YES		
	TEAM_LOSSES	Team Yearly Losses	INT(2)	99	0-16	YES		
	TEAM_AVG_POINTS_SCORED	Average Team Total Points Scored	DECIMAL (3,1)	99.9	0-99	YES		

	TEAM_AVG_POINTS_ALLOWED	Average Team Total Points Allowed	DECIMAL (3,1)	99.9	0-99	YES		
	TEAM_BOWL_GAME	Bowl Game Played In	VARCHAR (20)	Xxxxxxx		YES		
	TEAM_FINAL_RANK	Final Ranking	INT(2)	99	1-25	YES		

SCHEDULE	SCHEDULE_ID	Schedule ID Code	INT(3)	999	0-500	YES	PK	
	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	SCHEDULE_YEAR	Schedule Year	INT(4)	9999	2021 – 2025	YES		
	SCHEDULE_STRENGTH_RANK	Strength of Schedule	INT(3)	999	1-500	YES		
	SCHEDULE_BYE_WEEK	Bye Week Number	INT(2)	99	1-12	YES		
TEAMSPENDING	SPENDING_ID	Team Spending ID Code	INT(2)	99	0-18	YES	PK	
	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	SPENDING_FISCAL_YEAR	Fiscal Year	INT(4)	9999	1892-2026	YES		
	SPENDING_RECRUITING	Recruiting Expenses	INT(8)	99,999,999	0-19,999,999	YES		

	SPENDING_COACHING_STAFF	Coaching Staff Salary Pool	INT(8)	99,999,999	0-50,000,000	YES		
	SPENDING_FOOTBALL_FACILITIES	Football Facility Expenses	INT(9)	999,999,999	0-500,000,000	YES		
	SPENDING_TOTAL	Football Total Spending	INT(9)	999,999,999	0-500,000,000	YES		
TEAMREVENUE	TEAM_REV_ID	Team Revenue ID Code	INT(2)	99	0-18	YES	PK	
	TEAM_ID	Team ID Code	INT(3)	999	0-500	YES	FK	TEAM
	TEAM_REV_FISCAL_YEAR	Fiscal Year	INT(4)	9999	2021 – 2025	YES		
	TEAM_REV_SOURCE_TYPE	Team Revenue Source	VARCHAR (30)	Xxxxxxx		YES		
	TEAM_REV_TOT	Team Total Revenue	INT(8)	99,999,999	0-99,999,999	YES		
CONFERENCE REVENUE	CONF_REV_ID	Conference Total Revenue	INT(2)	99	0-50	YES	PK	
	CONF_ID	Conference ID Code	INT(2)	99	0-50	YES	FK	CONFERENCE
	CONF_REV_FISCAL_YEAR	Fiscal Year	INT(4)	9999	2021 – 2025	YES		

CONF_REV_SOUR CE_TYPE	Conference Revenue Source	VARCHAR (30)	Xxxxxxx		YES		
CONF_REV_TOT	Conference Total Revenue	INT(10)	9,999,999 ,999	0- 1,999,999,9 99	YES		

### III. Business Rules

#### CONFERENCE/TEAM

1. One CONFERENCE has at least one TEAM or has many
2. Each TEAM belongs to exactly one CONFERENCE
3. TEAM.CONFERENCE\_ID will not be null because every Big Ten TEAM is a part of the Big Ten CONFERENCE

#### CONFERENCE/CONFERENCEREVENUE

1. One CONFERENCE can have many CONFERENCEREVENUE records per fiscal year
2. Each CONFERENCEREVENUE record belongs to exactly one CONFERENCE

#### TEAM/PLAYER

1. Each TEAM has many PLAYERS
2. One PLAYER belongs to only one TEAM
3. PLAYER.TEAM\_ID cannot be null because every PLAYER is on a TEAM

#### TEAM/COACH

1. One TEAM is coached by one COACH or coached by many
2. Each COACH only coaches one TEAM
3. COACH.TEAM\_ID cannot be null because every TEAM has at least one COACH for each season

#### TEAM/STADIUM

1. One TEAM plays in exactly one home STADIUM
2. Each STADIUM is home to one TEAM
3. Neither TEAM nor STADIUM can be null to support one to one relationship

#### TEAM/SEASONSTATS

1. One TEAM can have many SEASONSTATS per year
2. Each SEASONSTATS record belong to only one team

#### TEAM/TEAMREVENUE

1. Each TEAM can have many TEAMREVENUE records per season
2. Each TEAMREVENUE record belongs to only one team

#### TEAM/TEAMSPENDING

1. Each TEAM can have many TEAMSPENDING records per season
2. Each TEAMSPENDING records belongs to only one team

#### TEAM/GAME

1. One TEAM plays in many GAMES as either the home TEAM or away TEAM
2. Each GAME must involve a home TEAM and an away TEAM

#### TEAM/SCHEDULE

1. Each TEAM plays one SCHEDULE per season
2. One SCHEDULE is assigned to one TEAM each year

#### PLAYER/GAME

1. One PLAYER can participate in one GAME or many
2. Each game can include one PLAYER or many
3. GAMEPARTICIPATION is bridge table
4. Neither GAME nor PLAYER are null

#### PLAYER/OFFENSIVEPLAYERSTATS

1. One PLAYER can record many OFFENSIVEPLAYERSTATS per season
2. Each OFFENSIVEPLAYERSTATS record belongs to one PLAYER

#### PLAYER/DEFENSIVEPLAYERSTATS

1. One PLAYER can record many OFFENSIVEPLAYERSTATS per season
2. Each OFFENSIVEPLAYERSTATS record belongs to one PLAYER

#### STADIUM/GAME

1. One STADIUM can host one GAME or can host many each year
2. Each GAME is played in one STADIUM

#### SEASONSTATS/GAME

1. One SEASONSTATS records stats from one GAME or many per year
2. Each GAME is recorded in one SEASONSTATS record

### IV. Entity Relationship Model (ERM) Components

Big Ten Conference Football Database Entity Relationship Model			
ENTITY	RELATIONSHIP	CONNECTIVITY	ENTITY
CONFERENCE	...consists of...	1:M	TEAM
CONFERENCE	...has...	1:M	CONFERENCEVENUE
TEAM	...has...	1:M	PLAYER
TEAM	...coached by...	1:M	COACH
TEAM	...plays in...	1:1	STADIUM
TEAM	...has...	1:M	SEASONSTATS
TEAM	...has...	1:M	TEAMREVENUE
TEAM	...has...	1:M	TEAMSPENDING
TEAM	...plays...	1:M	GAME
TEAM	...has...	1:1	SCHEDULE
PLAYER	...participates...	M:N	GAME
PLAYER	...records...	1:M	OFFENSIVEPLAYERSTATS
PLAYER	...records...	1:M	DEFENSIVEPLAYERSTATS
STADIUM	...hosts...	1:M	GAME
SEASONSTATS	...records ...	1:M	GAME

\*NOTE: GAMEPARTICIPATION is the bridge entity that enforces the M:N relationship between PLAYER and GAME

## V. Relational Schemas

CONFERENCE (**CONF\_ID**, CONF\_NAME, CONF\_FOUNDED\_YEAR,  
CONF\_HQ\_CITY, CONF\_HQ\_STATE, CONF\_COMMISSIONER\_NAME)

TEAM (**TEAM\_ID**, TEAM\_NAME, TEAM\_MASCOT, TEAM\_CITY, TEAM\_STATE,  
TEAM\_COLORS, CONF\_ID)

STADIUM (**STADIUM\_ID**, STADIUM\_NAME, STADIUM\_CAPACITY,  
STADIUM\_SURFACE\_TYPE, STADIUM\_CITY, STADIUM\_STATE, STADIUM\_ZIP,  
TEAM\_ID)

COACH (**COACH\_ID**, COACH\_FNAME, COACH\_LNAME, COACH\_POS\_TITLE,  
COACH\_HIRE\_DATE, COACH\_SALARY, TEAM\_ID)

PLAYER (**PLAYER\_ID**, PLAYER\_FNAME, PLAYER\_LNAME, PLAYER\_POS,  
PLAYER\_JERSEY\_NUM, PLAYER\_CLASS\_YEAR, PLAYER\_BIRTH\_DATE,  
PLAYER\_HOMETOWN\_CITY, PLAYER\_HOMETOWN\_STATE,  
PLAYER\_HOMETOWN\_ZIP, TEAM\_ID)

GAME (**GAME\_ID**, SEASON\_YEAR, GAME\_WK\_NUM, HOME\_TEAM\_ID,  
AWAY\_TEAM\_ID, STADIUM\_ID, GAME\_DATE, GAME\_ATTENDANCE,  
GAME\_HOME\_SCORE, GAME\_AWAY\_SCORE, GAME\_TV\_NETWORK)

GAMEPARTICIPATION (**PLAYER\_ID**, **GAME\_ID**, GP\_TOT\_MIN\_PLAYED,  
GP\_GAME\_RATING)

OFFENSIVEPLAYERSTATS (**PLAYER\_ID**, **SEASON\_YEAR**, OFF\_GAMES\_PLAYED,  
OFF\_PASS\_YARDS, OFF\_RUSH\_YARDS, OFF\_RECEIVING\_YARDS,  
OFF\_TOUCHDOWNS\_OFF\_INTERCEPTIONS\_THROWN, OFF\_FUMBLES\_LOST)

DEFENSIVEPLAYERSTATS (**PLAYER\_ID**, **SEASON\_YEAR**, DEF\_GAMES\_PLAYED,  
DEF\_INTERCEPTIONS, DEF\_TACKLES, DEF\_SACKS, DEF\_FORCED\_FUMBLES,  
DEF\_PASS\_BREAK\_UPS)

SEASONSTATS (**TEAM\_ID**, **SEASON\_YEAR**, TEAM\_WINS, TEAM\_LOSSES,  
TEAM\_AVG\_POINTS\_SCORED, TEAM\_AVG\_POINTS\_ALLOWED,  
TEAM\_BOWL\_GAME, TEAM\_FINAL\_RANK)

SCHEDULE (**SCHEDULE\_ID**, TEAM\_ID, SCHEDULE\_STRENGTH\_RANK,  
SCHEDULE\_YEAR, SCHEDULE\_BYE\_WEEK)

TEAMSPENDING (**SPENDING\_ID**, SPENDING\_FISCAL\_YEAR,  
SPENDING\_RECRUITING, SPENDING\_COACHING\_STAFF,

SPENDING\_FOOTBALL\_FACILITIES, SPENDING\_TOTAL, TEAM\_ID)

TEAMREVENUE (TEAM\_REV\_ID, TEAM\_REV\_FISCAL\_YEAR,  
TEAM\_REV\_SOURCE\_TYPE, TEAM\_REV\_TOT, TEAM\_ID)

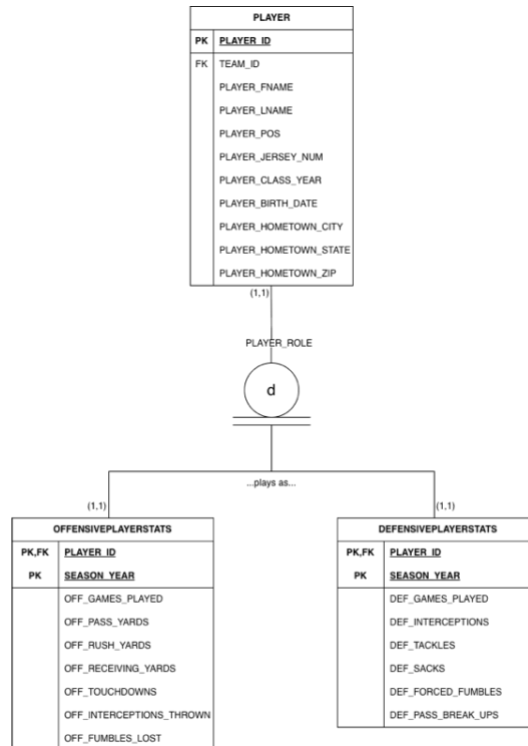
CONFERENCEVENUE (CONF\_REV\_ID, CONF\_REV\_FISCAL\_YEAR,  
CONF\_REV\_SOURCE\_TYPE, CONF\_REV\_TOT, CONF\_ID)

## VI. Entity Relationship Diagram (Crows Foot notation)





## VII. Subtype/Supertype diagram



## VIII. Normalized Table (1NF – 3NF)

### Pre Normalization (UNF)

This unnormalized table is a combination of the COACH and TEAM entities including all their attributes before any normalization has taken place. There are many problems with this unnormalized table as the attributes are clustered together, there are repeating values, and all values are not atomic.

TEAM_ID	TEAM_NAME	TEAM_MASCOT	TEAM_CITY	TEAM_STATE	TEAM_COLORS	COACH_ID	COACH_NAME	COACH_POS_TITLE	COACH_HIRED_DATE	COACH_SALARY
1	Indiana	Hoosiers	Bloomington	Indiana	Cream/Crimson	101	Curt Cignetti	Head Coach	2023-11-30	11,200,000
1	Indiana	Hoosiers	Bloomington	Indiana	Cream/Crimson	102	Briant Haines	DC	2023-11-30	2,300,000
2	Ohio St	Buckeyes	Columbus	Ohio	Scarlet/Gray	201	Ryan Day	Head Coach	2019-01-02	12,000,000
2	Ohio St	Buckeyes	Columbus	Ohio	Scarlet/Gray	202	Matt Patricia	DC	2024-03-01	2,000,000

## First Normal Form Conversion (1NF)

This 1NF table has two unique identifiers in TEAM\_ID and COACH\_ID. Both attributes act as composite primary keys as they uniquely identify the other attributes in the table. TEAM\_ID uniquely identifies TEAM\_NAME, TEAM\_MASCOT, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, and TEAM\_SECONDARY\_COLOR. COACH\_ID uniquely identifies COACH\_FNAME, COACH\_LNAME, COACH\_POS\_TITLE, COACH\_HIRE\_DATE, and COACH\_SALARY.

In the 1NF table below, repeating groups and multi-valued attributes are nonexistent. Team colors and coach name were split into two columns each rather than having all the data stored in a single column. This means that each column in the table successfully has atomic values as a coach's first name and last name cannot be further divided. There is still a partial dependency existing that will be shown in the dependency diagram below.

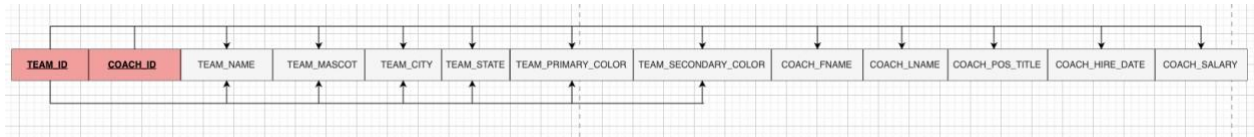
TEAM_ID	TEAM_NAME	TEAM_MASCOT	TEAM_CITY	TEAM_STATE	TEAM_PRIMARY_COLOR	TEAM_SECONDARY_COLOR	COACH_ID	COACH_FNAME	COACH_LNAME	COACH_POS_TITLE	COACH_HIRE_DATE	COACH_SALARY
1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson	101	Curt	Cignetti	Head Coach	2023-11-30	11,200,000
1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson	102	Bryant	Haines	DC	2023-11-30	2,300,000
2	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray	201	Ryan	Day	Head Coach	2019-01-02	12,000,000
2	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray	202	Matt	Patricia	DC	2024-03-01	2,000,000

## 1NF Dependency Diagram

(TEAM\_ID, COACH\_ID, TEAM\_NAME, TEAM\_MASCOT, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, TEAM\_SECONDARY\_COLOR, COACH\_FNAME, COACH\_LNAME, COACH\_POS\_TITLE, COACH\_HIRE\_DATE, COACH\_SALARY)

Partial Dependency:

(TEAM\_ID  $\twoheadrightarrow$  TEAM\_NAME, TEAM\_MASCOT, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, TEAM\_SECONDARY\_COLOR)



## SECOND NORMAL FORM CONVERSION (2NF)

The second normal form requires the dependencies to be investigated. Specifically, it demands for the removal of partial dependencies, which analyzes the attributes that are only dependent on a portion of the primary key. Here, in the above table, there is an evident partial dependency, including the reliance of TEAM\_NAME, TEAM\_MASCOT, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, and TEAM\_SECONDARY\_COLOR on the primary key TEAM\_ID.

This partial dependency violates 2NF requirements as these attributes solely depend on TEAM\_ID, and not on COACH\_ID. To fix this, is to further normalize the table by splitting it into two different entities, COACH and TEAM. The TEAM table will include only the direct TEAM\_ID dependent attributes, and the COACH table will include the attributes directly dependent on COACH\_ID along with TEAM\_ID, as it will need to reference the TEAM entity after this step of the normalization process.

**Table: TEAM**

TEAM_ID	TEAM_NAME	TEAM_MASCOT	TEAM_CITY	TEAM_STATE	TEAM_PRIMARY_COLOR	TEAM_SECONDARY_COLOR
1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson
1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson
2	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray
2	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray

**Table: COACH**

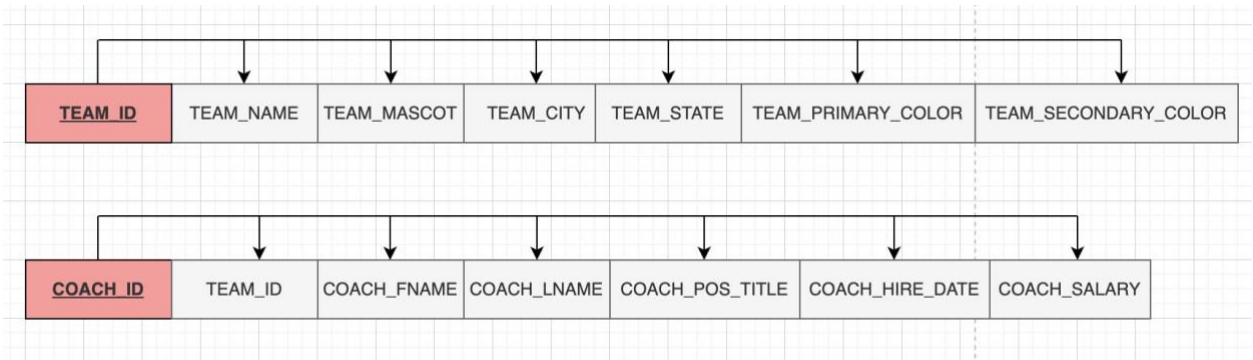
COACH_ID	TEAM_ID	COACH_FNAME	COACH_LNAME	COACH_POS_TITLE	COACH_HIRE_DATE	COACH_SALARY
101	1	Curt	Cignetti	Head Coach	2023-11-30	11,200,000
102	1	Bryant	Haines	DC	2023-11-30	2,300,000
201	2	Ryan	Day	Head Coach	2019-01-02	12,000,000

202	2	Matt	Patricia	DC	2024-03-01	2,000,000
-----	---	------	----------	----	------------	-----------

### 2NF Dependency Diagrams

TEAM (TEAM\_ID, TEAM\_NAME, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, TEAM\_SECONDARY\_COLOR)

COACH (COACH\_ID, TEAM\_ID, COACH\_FNAME, COACH\_LNAME, COACH\_POS\_TITLE, COACH\_HIRE\_DATE, COACH\_SALARY)



### Third Normal Form Conversion (3NF)

For 3NF, the goal is to further normalize a table by eliminating any transitive dependencies. Here, there are no transitive dependencies, and the partial dependency has already been removed in the second normal conversion. Because of this, there is no need to further normalize these tables. The design of these tables can be improved though, specifically the TEAM table, as a foreign key can be added, CONF\_ID, which connects TEAM to the CONFERENCE entity. Aside from this addition, the tables are in 3NF and have no repeating groups, the attribute values are atomic, and there are no partial or transitive dependencies.

**Table: TEAM**

TEAM_ID	CONF_ID	TEAM_NAME	TEAM_MASCOT	TEAM_CITY	TEAM_STATE	TEAM_PRIMARY_COLOR	TEAM_SECONDARY_COLOR
1	1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson
1	1	Indiana	Hoosiers	Bloomington	IN	Cream	Crimson
2	1	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray

2	1	Ohio St	Buckeyes	Columbus	OH	Scarlet	Gray
---	---	---------	----------	----------	----	---------	------

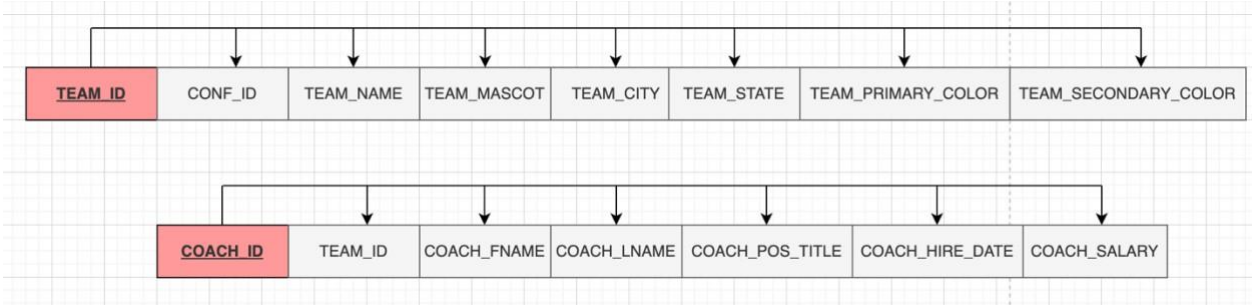
Table: COACH

COACH_ID	TEAM_ID	COACH_FNAME	COACH_LNAME	COACH_POS_TITLE	COACH_HIRE_DATE	COACH_SALARY
101	1	Curt	Cignetti	Head Coach	2023-11-30	11,200,000
102	1	Bryant	Haines	DC	2023-11-30	2,300,000
201	2	Ryan	Day	Head Coach	2019-01-02	12,000,000
202	2	Matt	Patricia	DC	2024-03-01	2,000,000

3NF Dependency Diagrams

TEAM (TEAM ID, CONF\_ID, TEAM\_NAME, TEAM\_CITY, TEAM\_STATE, TEAM\_PRIMARY\_COLOR, TEAM\_SECONDARY\_COLOR)

COACH (COACH ID, TEAM\_ID, COACH\_FNAME, COACH\_LNAME, COACH\_POS\_TITLE, COACH\_HIRE\_DATE, COACH\_SALARY)



IX. Big Ten Database Questions to Query

- 1. What revenue source had the highest total revenue for the 2024 CFB season for Indiana University?
- 2. How many games did each team win in the 2024 CFB season?
- 3. Who were the top 5 players with the most touchdowns in the 2024 season?

## SQL Queries mapping the above questions

### 1. Query that pulls from one table

```
SELECT
    TEAM_REV_SOURCE_TYPE,
    TEAM_REV_TOT
FROM
    TEAMREVENUE
WHERE
    TEAM_ID = 101
    AND TEAM_REV_FISCAL_YEAR = 2024
ORDER BY
    TEAM_REV_TOT DESC
LIMIT 1;
```

#### Query Results:

TEAM_REV_SOURCE_TYPE	TEAM_REV_TOT
Media Rights	84000000

### 2. Query that pulls from two tables

```
SELECT
    TEAM.TEAM_NAME,
    SEASONSTATS.TEAM_WINS
FROM
    TEAM
JOIN
    SEASONSTATS
ON
    TEAM.TEAM_ID = SEASONSTATS.TEAM_ID
WHERE
    SEASONSTATS.SEASON_YEAR = 2024
ORDER BY
    SEASONSTATS.TEAM_WINS DESC;
```

#### Query Results:

TEAM_NAME	TEAM_WINS	▼ 1
Michigan	13	
Ohio State	12	
Penn State	10	
Purdue	7	
Indiana	5	

### 3. Query that includes a subquery

```

SELECT
    PLAYER.PLAYER_FNAME,
    PLAYER.PLAYER_LNAME,
    OFFENSIVEPLAYERSTATS.OFF_TOUCHDOWNS
FROM
    PLAYER
JOIN
    OFFENSIVEPLAYERSTATS
ON
    PLAYER.PLAYER_ID = OFFENSIVEPLAYERSTATS.PLAYER_ID
WHERE
    OFFENSIVEPLAYERSTATS.SEASON_YEAR = 2024
AND
    OFFENSIVEPLAYERSTATS.OFF_TOUCHDOWNS >= (
        SELECT
            MIN(OFF_TOUCHDOWNS)
        FROM (
            SELECT
                OFF_TOUCHDOWNS
            FROM
                OFFENSIVEPLAYERSTATS
            WHERE
                SEASON_YEAR = 2024
            ORDER BY
                OFF_TOUCHDOWNS DESC
            LIMIT 5
        ) AS TOP5
    )
ORDER BY

```



OFFENSIVEPLAYERSTATS.OFF\_TOUCHDOWNS DESC LIMIT 5;

Query Results:

PLAYER_FNAME	PLAYER_LNAME	OFF_TOUCHDOWNS	▼ 1
J.J.	McCarthy	24	
Hudson	Card	19	
Tayven	Jackson	18	
TreVeyon	Henderson	14	
Jaylin	Lucas	6	

### Database Data Works Cited

Introducing Chatgpt | openai. (n.d.). <https://openai.com/index/chatgpt/>

Sports Reference. (n.d.). *2024 big ten conference coaches: College football at sports*. Sports-Reference.com. <https://www.sports-reference.com/cfb/conferences/big-ten/2024-coaches.html>