

Course Title:	Fundamentals of Data Engineering
Course Number:	COE 848
Semester/Year (e.g.F2016)	W2024

Instructor:	Dr. Faezeh Ensan
--------------------	------------------

<i>Assignment/Lab Number:</i>	Lab 1
<i>Assignment/Lab Title:</i>	Setup and Project Specification

<i>Submission Date:</i>	January 25, 2024
<i>Due Date:</i>	January 25, 2024

Student LAST Name	Student FIRST Name	Student Number	Section	Signature*
Saini	Harsanjam	501055402	04	Harsanjam

*By signing above you attest that you have contributed to this written lab report and confirm that all work you have contributed to this lab report is your own work. Any suspicion of copying or plagiarism in this work will result in an investigation of Academic Misconduct and may result in a "0" on the work, an "F" in the course, or possibly more severe penalties, as well as a Disciplinary Notice on your academic record under the Student Code of Academic Conduct, which can be found online at: <http://www.ryerson.ca/senate/current/pol60.pdf>

Project Specification

For this term's project, I am excited to delve into the world of basketball by developing a comprehensive database and analysis system focused on the NBA (National Basketball Association). The primary goal of this project is to create an application that caters to basketball enthusiasts like me, providing a unified platform to organize, analyze, and explore various facets of the NBA, from player statistics to team performances.

My choice of this topic is deeply rooted in my enduring passion for basketball. Having followed the NBA closely for several years, I recognize the league's significance and the wealth of data it generates. This project allows me to merge my technical skills with my love for the sport, creating an application that not only showcases my capabilities but also provides a valuable resource for NBA fans.

The NBA Explorer project aims to capture a wide range of data, including player information, team details, game statistics, and more. This database will serve as a centralized repository of information about NBA players actively participating in the current season, as well as details about the teams competing in the league. The system will allow users to not only track ongoing seasons but also explore historical data, making it a valuable resource for NBA fans, analysts, and researchers.

The features of the NBA Explorer database include comprehensive player information, detailing aspects such as name, age, height, weight, position, college, and various performance metrics like points per game, assists per game, and field goal percentage. Similarly, team information encompasses key details such as team name, coach, founded year, win-loss record, conference, and division.

Database Structure:

Entities:

- Player
- Team
- Game

Attributes:

- **Player:** PlayerID (Primary Key), Name, Age, Height, Weight, Position, College, Points per game, Assists per game, Rebounds per game, Field Goal Percentage, Free Throw Percentage, Three-Point Percentage
 - PlayerID serves as the primary key, uniquely identifying each player.

- **Team:** TeamID (Primary Key), Team Name, Coach, Founded Year, Win-Loss Record, Conference, Division
 - TeamID serves as the primary key, ensuring the uniqueness of each team
- **Game:** GameID (Primary Key) is the primary key, uniquely identifying each game.
 - Date, Venue, Attendance, Final Scores, and Key Moments provide comprehensive information about each game.

Relationships:

Player plays for Team (**Many-to-One**):

- A player can play for one team, but a team can have many players.

Team has many Players (**One-to-Many**):

- A team has multiple players, but each player is associated with only one team.

Game involves Teams (**Many-to-Many**):

- Multiple teams participate in a game, and a team can play in multiple games. Similarly, a game involves multiple teams.

This project is not merely a technical endeavor but a fusion of my expertise and passion. By creating a comprehensive NBA database, I aim to contribute to the basketball community and further enhance my skills in database management and analysis.

To provide users with a powerful analytical tool, a list of queries has been crafted to extract meaningful insights from the database. Questions range from specific player performance in a given game to team statistics for the entire season. These queries offer a spectrum of information, including scoring details, shooting efficiency, team standings, and individual player contributions.

List of Queries:

1. How many points did a specific player score in a particular game?
2. What is the average assists per game for a specific player in the current NBA season?
3. How many rebounds did a specific player grab in a particular game?
4. What is the field goal percentage of a particular player in the current NBA season?
5. How many wins and losses does a specific team have in the current NBA season?
6. What is the average points per game for a specific team in the current NBA season?
7. Who is the top scorer in the league for the current season?
8. Which team has the best free-throw percentage in the current NBA season?
9. What is the three-point percentage of a specific player in the current NBA season?
10. How many assists did a specific player accumulate throughout the entire NBA season?
11. What is the win-loss record for a particular team against its top rival?
12. What is the average age of NBA players in the current season?
13. Which team has the highest average rebounds per game in the current NBA season?