



Data Management Project

NBA Data Warehouse Implementation - Spring 2025

Group Members:

- Najeh Alhalawani - 2223737
- Jamil Nassar - 2224424

Project Type:

Data Warehouse Implementation

Dataset Description:

We will build a Data Warehouse using NBA-related datasets, integrating information about games, game details, rankings, teams, players, player details, and head coaches. The datasets will be sourced from Kaggle, such as:

<https://www.kaggle.com/datasets/nathanlauga/nba-games?resource=download&select=games.csv>

<https://www.kaggle.com/datasets/justinas/nba-players-data>

https://www.kaggle.com/datasets/jacobbaruch/nba-head-coaches?select=NBA_head_coaches.csv

Project Overview:

This project focuses on integrating multiple NBA-related datasets into a unified Data Warehouse. We will preprocess data by dropping unwanted columns, aligning records to a common starting season, and handling missing values with appropriate imputation methods. The ETL process will be designed and executed to integrate these datasets seamlessly. We will construct the Entity-Relationship (ER) diagram, define the Dimensional Fact Model (DFM), and implement the structured schema in PostgreSQL. Finally, we will conduct OLAP Operations and meaningful ROLAP analysis on player performance, game outcomes, and team statistics over the seasons to derive valuable insights.