

Hello Borromean Digital HR,

I hope you are doing great. Here are the answers to your questions regarding final project.

The final project will be conducted by a team, including Yunzhe Liu (UMID: 97654360) and Zhonghan Xie (UMID: 83866049).

1. The dataset we plan to use is an open dataset shared on Kaggle: *Football Data from Transfermarkt* (please see <https://www.kaggle.com/datasets/davidcariboo/player-scores?select=players.csv>). It is web-scraped from Transfermarkt (please see <https://www.transfermarkt.com/>), which is a website comprehensively collecting and documenting soccer-related data but it does not share any data files directly. The scraped dataset covers information about clubs, games, players, etc., but we will primarily focus on player transfers.
2. We will first normalize the dataset with concentration on transfer-related information. The key question of interest will include:
  - (a) Are attackers valued more than defenders in general?
  - (b) What's the temporal trend of transfer valuation in the last decade?
3. Here is the planned ERD for our final project:

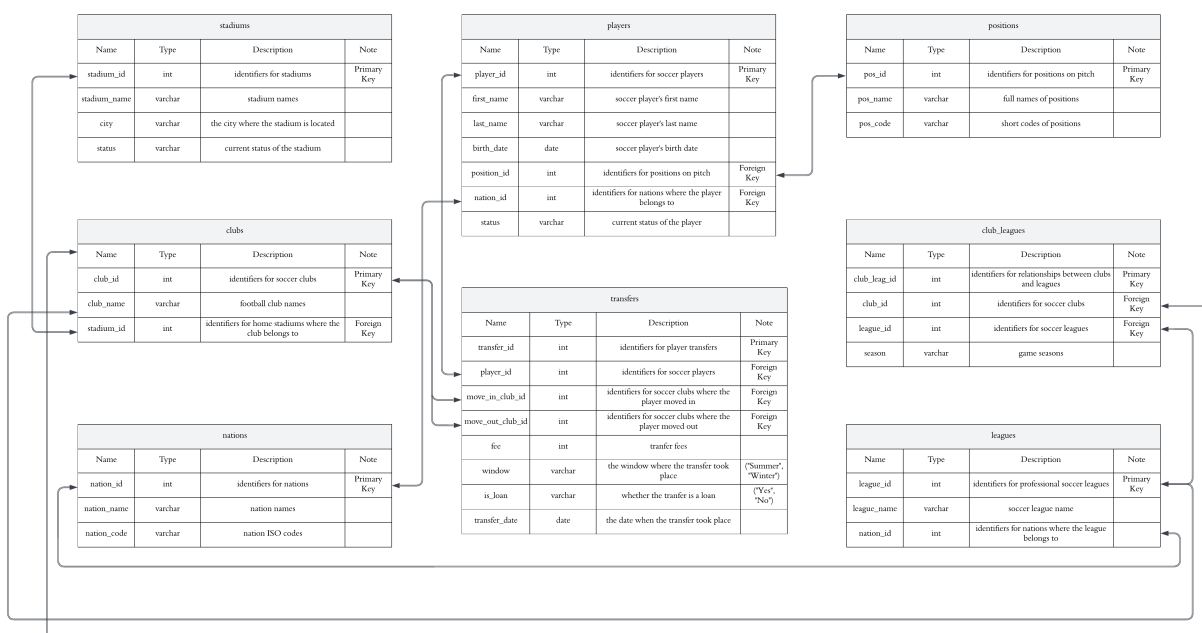


Figure 1: The ERD for the final project.

From,

Yunzhe Liu and Zhonghan Xie