**Player table 🡪**

**# create the table**

create table player( player\_name varchar(255), team\_id bigint, player\_id bigint, season bigint,id int unsigned, primary key(id));

alter table player modify id int unsigned not null auto\_increment;

alter table player add foreign key(team\_id) references team(team\_id);

**# load values into table from csv file**

load data local infile '/Users/christopheryang/Desktop/College/2ndYear/Summer2022/CS564/Project/finalproject/CS564/csv\_data/players.csv' into table player fields terminated by ',' lines terminated by'\n' ignore 1 lines;

**Team table 🡪**

**#create the table**

create table team( team\_id bigint, min\_year bigint, max\_year bigint, abbreviation varchar(255), nickname varchar(255), year\_founded bigint, city varchar(255), arena\_name varchar(255), arena\_capacity bigint, owner varchar(255), general\_manager varchar(255), head\_coach varchar(255), d\_league\_affiliation varchar(255), primary key(team\_id));

**#load values into table from csv**

load data local infile '/Users/christopheryang/Desktop/College/2ndYear/Summer2022/CS564/Project/finalproject/CS564/csv\_data/teams.csv' into table team fields terminated by ',' lines terminated by'\n' ignore 1 lines;

**Ranking table 🡪**

**#Create the table**

create table ranking( team\_id bigint, league\_id bigint, season\_id bigint, standings\_date date, conference varchar(255), team\_name varchar(255), games bigint, win bigint, loss bigint, win\_pct decimal(4,3), home\_record varchar(255), road\_record varchar(255), return\_to\_play varchar(255), id int unsigned, primary key(id));

alter table ranking modify id int unsigned not null auto\_increment;

**#load values into table from csv**

load data local infile '/Users/christopheryang/Desktop/College/2ndYear/Summer2022/CS564/Project/finalproject/CS564/csv\_data\_original/rankingunique.csv' into table ranking fields terminated by ',' lines terminated by'\n' ignore 1 lines;

alter table ranking delete column win\_pct, return\_to\_play;

alter table ranking add foreign key(team\_id) references team(team\_id);

**Game table🡪**

**#create the table**

create table game( game\_date date, game\_id bigint, game\_status varchar(255), home\_team\_id bigint, visitor\_team\_id bigint, season bigint, team\_id\_home bigint, pts\_home bigint, fg\_pct\_home decimal(4,3), ft\_pct\_home decimal(4,3), fg3\_pct\_home decimal(4,3), ast\_home bigint, reb\_home bigint, team\_id\_away bigint, pts\_away bigint, fg\_pct\_away decimal(4,3), ft\_pct\_away decimal(4,3), fg3\_pct\_away decimal(4,3), ast\_away bigint, reb\_away bigint, home\_team\_wins bigint, primary key(game\_id));

**#load values into the table**

load data local infile '/Users/christopheryang/Desktop/College/2ndYear/Summer2022/CS564/Project/finalproject/CS564/csv\_data\_original/games\_unique.csv' into table game fields terminated by ',' lines terminated by'\n' ignore 1 lines;

alter table game delete column team\_id\_home, team\_id\_away;

**Game details table🡪**

**#create the table**

create table game\_detail( game\_id bigint, team\_id bigint, abbreviation varchar(255), city varchar(255), player\_id bigint, player\_name varchar(255), nickname varchar(255), start\_position varchar(255), comment varchar(255), minutes varchar(255),fgm decimal(3,1), fga decimal(3,1), fg\_pct decimal(4,3), fg3m decimal(3,1), fg3a decimal(3,1), fg3\_pct decimal(4,3), ftm decimal(3,1), fta decimal(3,1), ft\_pct decimal(4,3), o\_reb decimal(3,1), d\_reb decimal(3,1), reb decimal(3,1), ast decimal(3,1), stl decimal(3,1), blk decimal(3,1), turnovers decimal(3,1), pf decimal(3,1), pts decimal(3,1), plus\_minus decimal(3,1), primary key(game\_id, player\_id, team\_id));

alter table game\_detail add foreign key(game\_id) references game(game\_id);

# table needs player\_id, season, team\_id or it needs its own id

**#load values into the table**

load data local infile '/Users/christopheryang/Desktop/College/2ndYear/Summer2022/CS564/Project/finalproject/CS564/csv\_data\_original/games\_details\_unique.csv' into table game\_detail fields terminated by ',' OPTIONALLY ENCLOSED BY '"' lines terminated by'\n' ignore 1 lines;

alter table game\_detail add column season;

alter table game\_detail add foreign key(season) references game(season); // not sure if this one works, find working version

alter table game\_detail add foreign key(game\_id) references game(game\_id);

alter table game\_detail add foreign key(team\_id) references team(team\_id);

alter table game\_detail add foreign key(player\_id) references player(player\_id);

**Got rid of duplicate values by:**

1. **Changing long decimals to round up or down 0.000…0001 to 0, 0.33299999 to .333 etc**
   1. **Use shift + command + l and command + f to find and apply changes to all occurrences**
2. **Sorting newly updated files into a new file and moving header row back to top of file**
   1. **Wc -l to check wc, sort -u ‘filename’ > ‘newfilename’, check wc and move header line back to top**
3. **Create tables without primary or foreign keys, load data, add primary keys, find duplicates in file if they exist**
   1. **Create table ‘name’ (…);**
   2. **Load data local infile…**
   3. **Alter table ‘name’ add primary keys(…);**
   4. **Find and remove duplicates if they exist**

**1610612742,0,22020,2020-12-29,West,Dallas,3,1,2,0.3329999999999999,0-0,1-2,**

**1610612763,0,22020,2020-12-29,West,Memphis,3,1,2,0.3329999999999999,0-2,1-0,**

**1610612760,0,22020,2020-12-29,West,Oklahoma City,3,1,2,0.3329999999999999,0-2,1-0,**

**1610612742,0,22020,2020-12-29,West,Dallas,3,1,2,0.333,0-0,1-2,**

**1610612763,0,22020,2020-12-29,West,Memphis,3,1,2,0.333,0-2,1-0,**

**1610612760,0,22020,2020-12-29,West,Oklahoma City,3,1,2,0.333,0-2,1-0,**

**:1,$s/0.3329999999999999/0.333/**

**2021-01-01,22000070,Final,1610612765,1610612738,2020,1610612765,96.0,0.409,0.65,0.324,18.0,48.0,1610612738,93.0,0.47,0.5,0.258,20.0,45.0,1**

**2021-01-01,22000076,Final,1610612743,1610612756,2020,1610612743,103.0,0.453,0.813,0.414,28.0,38.0,1610612756,106.0,0.487,0.905,0.371,24.0,37.0,0**

**2021-01-01,22000076,Final,1610612743,1610612756,2020,1610612743,103.0,0.459,0.813,0.414,28.0,38.0,1610612756,106.0,0.487,0.905,0.371,24.0,37.0,0**

**22000070,1610612765,DET,Detroit,1626143,Jahlil Okafor,,,DNP - Injury/Illness,,,,,,,,,,,,,,,,,,,,**

**22000070,1610612765,DET,Detroit,203924,Jerami Grant,,F,,37:26,9.0,24.0,0.375,2.0,9.0,0.222,4.0,5.0,0.8,1.0,3.0,4.0,0.0,1.0,0.0,1.0,2.0,24.0,6.0**

**22000070,1610612738,BOS,Boston,1627759,Jaylen Brown,,G,,37:34,12.0,20.0,0.6,1.0,7.0,0.143,0.0,0.0,0.0,2.0,7.0,9.0,2.0,3.0,2.0,4.0,3.0,25.0,-9.0**

**22000070,1610612738,BOS,Boston,1629684,Grant Williams,,,,19:23,4.0,6.0,0.667,0.0,1.0,0.0,0.0,0.0,0.0,2.0,2.0,4.0,0.0,0.0,0.0,0.0,4.0,8.0,-3.0**

**22000070,1610612765,DET,Detroit,1630165,Killian Hayes,,G,,23:24,0.0,5.0,0.0,0.0,2.0,0.0,0.0,0.0,0.0,0.0,1.0,1.0,6.0,0.0,0.0,2.0,2.0,0.0,-2.0**

**22000070,1610612738,BOS,Boston,1630202,Payton Pritchard,,,,16:00,0.0,3.0,0.0,0.0,3.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,2.0,2.0,1.0,2.0,0.0,0.0,1.0**

**22000070,1610612765,DET,Detroit,1630191,Isaiah Stewart,,,,18:24,3.0,4.0,0.75,0.0,0.0,0.0,2.0,3.0,0.667,2.0,2.0,4.0,0.0,0.0,0.0,1.0,3.0,8.0,-5.0**