FootballHome

Presented by: Zihan Wang

Simon Fraser University

Aug 1, 2019

Brief Overview of this project practice:

For this project, I designed a database called "FootballHome" for football fans to help them look up relative football information, including matches, football teams, team players, awards, etc. Basically, I designed seven main entities sets and four relationships in tables:

Entity sets:

- FootballTeam (TeamName, Country, City, Founded)

- HeadCoach (Name, DateOfBirth, CoachingYears, Gender)

- Award (TeamName, Competitions, Titles)

- Matches (GameDate, HomeTeam, AwayTeam, Location)

- Referees (Name, DateofBirth, Gender, Country)

- TeamPlayers (Name, DateOfBirth, Gender, Height, Position, Nationality)

- RetiredPlayers (Name, RetiredDate)

Relationships:

- BelongTo (PlayerName, TeamName) - connects TeamPlayers and FootballTeam

- Plays (TeamName, GameDate) – connects FootballTeam and Matches

- Manage (headCoachName, TeamName) – connects FootballTeam and HeadCoach

- Preside (GameDate, RefereeName) – connects Matches and Referees

Please note that I do not set the relation table for the relationship of weak entity sets and subclasses, just to avoid redundancy. The detailed design (graph) is attached later.

E/R Diagram:
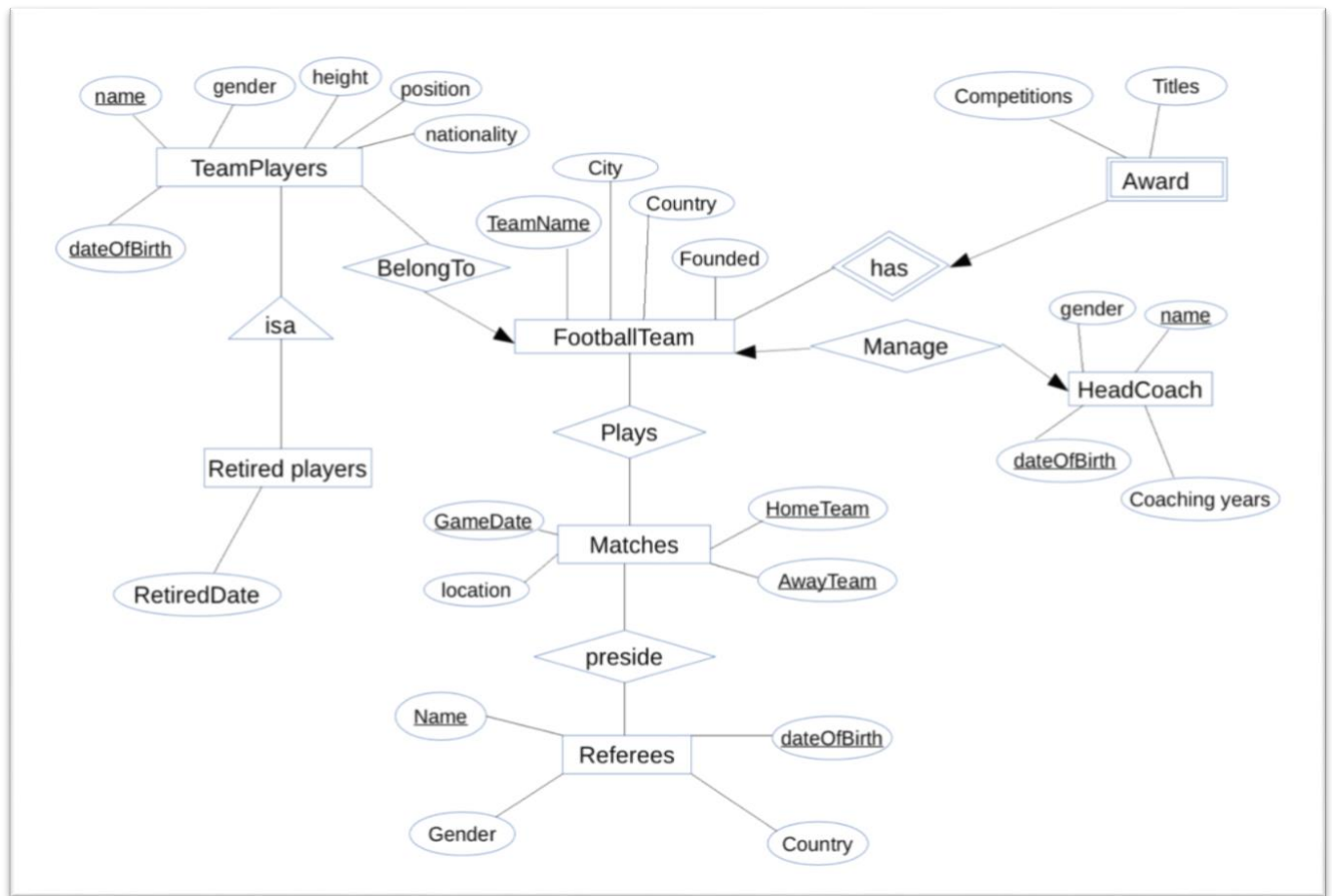
Anomalies Analysis:

I did a comprehensive review of my design and supposed that there did not occur many serious anomalies. Well, just to have a quick review, there are three main anomalies that may happen:

- Update Anomaly: one or more instances of duplicated data is updated, but not all.

- Insertion Anomaly: certain attributes cannot be inserted into the database without the presence of other attributes.

- Deletion Anomaly: certain attributes are lost because of the deletion of other attributes.

Here is a detailed analysis of each table:

- FootballTeam (<u>TeamName</u>, Country, City, Founded)

  A sample tuple of FootballTeam is like: ('Inter Milan', 'Italy', 'Milan', '1908-03-09'). The primary key of the table FootballTeam is TeamName. To simplify we will use "T, Co, Ci, F" to represent these four attributes.

  If we know the value of TeamName, we can obtain its origin Country, City and when Founded. By this, we can say that the Country, City and Founded are functionally depended (FD) on TeamName. Using the splitting rule, it could be shown like: "T -> Co, T -> Ci and T -> F". Notice that this is also the minimal basis for FootballTeam table.

  If we are concerned on the attributes closure, some combinations of Co, Ci and F with T also make sense as FD. For example, "T -> Co, Ci", "T -> Ci, T" or "T, Co -> Co, Ci, F" are acceptable FDs as well. As long as we know the value of a TeamName (it is unique: no 2 footabllTeams can have the same name and do match), we could know its origin country, city and founded date. When we update a football team, that team will be updated and will not influence the others. When we insert a new team, it does not make much sense to just input Country, City and Founded Date if we do not provide the TeamName. When doing deletion, some information may get lost with the deletion of TeamName. But they are not that important as we are concerned the TeamName best. Hence, there is no anomalies in FootballTeam table.

- HeadCoach (<u>Name, DateOfBirth</u>, CoachingYears, Gender)

Similay with the analysis before, for HeadCoach table, we use "N, D, C, G" to represent these four attributes.

FDs: as the primary key is N and D, the direct FDs are: "N,D -> C, G", which means CoachingYears and Gender are functionally depended on the Name and DateOfbirth. Some other acceptable FDs are like "N, D -> C", "N, D -> G" and "N, D, C -> G".

There are no FD Anomalies because: for update anomaly, it is almost impossible to have two coaches with the same name and date of birth, which means that every coach is independent and will not influence the others. The same works for deletion and insertion. We do not care about the information of CoachingYears and Gender as far as we know which coach it is. Hence, there is no serious anomalies in HeadCoach.

- Award (TeamName, Competitions, Titles)

  The award table is a subclass of FootballTeam, which shows how many awards that the football team has. We will use "Te, C Ti" to represent these three attributes.

  FDs: As the primary key is Te, the direct FDs are: "Te -> C, Ti", which means that competitions and titles are functionally depended by the teamname. Other optional Fds are like: "Te ->C" and "Te -> Ti".

  This is a simple table and there are no obvious anomalies. For update action, basically we just update the titles of that competitions if a team wins an award again. If it is a new award, we just create its competition title and set the title number to be 1. There does not

have two same competitions occur. For deletion / insertion, it may cause some anomaly if we are really concerned about one competition, let's say World Cup. However, the awards are belong to the team and together with which team wins it, the information could make sense. Hence there are no anomalies.

- Matches (GameDate, HomeTeam, AwayTeam, Location)

We use "G, H, A, L" to represent the table. The primary key is "G, H, A" together.

FDs: The most important functional dependency is "G, H, A -> L", which means that the location of a match is functionally depended by the gamedate, hometeam and awayteam.

There are some kind of anomalies occurred in this table. For updating, everything works well as the GameDate and two against teams are uniquely defined together. However, for insertion and deletion, if we only care about the HomeTeam, AwayTeam and location, we cannot add a new match unless we have at least a clear GameDate for that. Also, there is some kind of redundancy because if two teams fight with each other, there will be two records stored in the database with the same GameDate and Location as well as two inverse HomeTeam and AwayTeam information. Actually, we just need one of them. No need of decomposition, but we need to simplify it a little bit. Hence, if the database sees two matches have the same GameDate and Location, just inverse HomeTeam and AwayTeam, it should just keep one and stop the insertion.

- Referees (Name, DateofBirth, Gender, Country)

The primay key is "Name and DateofBrith". We use "N, D, G, C" to represent the attributes. Referees preside the specific match. Unlike the table "HeadCoach", this is not one-to-one relationship with table Matches because for each match there could be many referees and for each referee, he/she could preside many matches. This is a many-to-many relationship.

The most obvious FD is "N, D -> G, C". Some other combinations like "N, D -> G", "N, D, G -> C" also work.

Like the table "HeadCoach", there is no anomaly for this table. Because we only care about the information of Gender and Country as far as we know the Referees' Name and DateOfBirth. Name and DateofBirth togerther as a key could almost identify a referee uniquely. It does not make much sense to leave the Gender and Country information separately. Hence, the table satisfies the requirements of avoiding insertion, deletion and update anomalies.

- TeamPlayers (<u>Name, DateOfBirth</u>, Gender, Height, Position, Nationality)
  The key is "Name and DateOfBirth". We use "N, D, G, H, P, Nation" to reprensent the whole attribute. Note that this table has a many-to-one relationship with FootballTeam because for each team there are many teamplayers.

  FDs: "N, D -> G, H, P, Nation". Using the splitting rule, it could be splitted into 4 different functional dependencies. Combine any "G, H, P, Nation" to the left side, the FDs make sense as well.

The corresponding information "gender, height, position, nationality" all make sense with the existence of a TeamPlayer. Hence, it gets rid of insertion, deletion and update anomalies.

- RetiredPlayers (Name, RetiredDate)

    The primary key is the Name of the player. Notice that RetiredPlayers is a subclass of TeamPlayers. They just have a special attribute "RetiredDate" to show that they are retired. We use "N, R" to represent the table.

    The most obvious FD is "N -> R", which means that the RetiredDate is functionally depended by the Name of players. As there are just two attributes, there are not many other important functional dependencies.

    Similarly, as this is a simple relation schema, it could avoid the anomalies. However, we notice a design mistake in this table. The schema should also contain the DateOfBirth attribute from TeamPlayers table. Hence, the improved schema should be (Name, DateOfBrith, RetiredDate).

- For relationship tables:

    As the relationship tables (BelongTo, Manage, Plays, Preside) just stores the keys of the tables that they connect with, they will not commit the insertion, deletion or update anomalies. All of their attributes form the primary key.

However, we notice a design mistake in some relationship tables. They should contain all keys of the connected tables. Our new design fixes that by adding all keys:

- o BelongTo (PlayerName, DateOfBirth, TeamName) - connects TeamPlayers and FootballTeam

- o Plays (TeamName, GameDate, HomeTeam, AwayTeam) – connects FootballTeam and Matches

- o Manage (headCoachName, DateOfBirth, TeamName) – connects FootballTeam and HeadCoach

- o Preside (GameDate, HomeTeam, AwayTeam, RefereeName) – connects Matches and Referees

Q2: SQL Schema

We changed some attributes for some tables. The new schemas are like:

Entity sets:

- FootballTeam (TeamName, Country, City, Founded)

- HeadCoach (Name, DateOfBirth, CoachingYears, Gender)

- Award (TeamName, Competitions, Titles)

- Matches (GameDate, HomeTeam, AwayTeam, Location)

- Referees (Name, DateofBirth, Gender, Country)

- TeamPlayers (Name, DateOfBirth, Gender, Height, Position, Nationality)

- RetiredPlayers (Name, DateOfBirth, RetiredDate)

Relationships:

- BelongTo (PlayerName, DateOfBirth, TeamName) - connects TeamPlayers and FootballTeam

- Plays (TeamName, GameDate, HomeTeam, AwayTeam) – connects FootballTeam and Matches

- Manage (headCoachName, DateOfBirth, TeamName) – connects FootballTeam and HeadCoach

- Preside (GameDate, HomeTeam, AwayTeam, RefereeName) – connects Matches and Referees

Q3: Populate Tables

```
CREATE TABLE FootballTeam(
    TeamName        varchar(255)  PRIMARY KEY,
    Country      text,
    City                text  DEFAULT NULL,
    Founded     date CHECK (Founded < date('now')) ,
    FOREIGN KEY (TeamName) REFERENCES Manage(TeamName),
    FOREIGN KEY (TeamName) REFERENCES BelongTo(TeamName),
    FOREIGN KEY (TeamName) REFERENCES Plays(TeamName)
);

CREATE TABLE HeadCoach(
    Name                varchar(255),
    DateOfBirth        date CHECK (DateOfBirth < date('now')),
    CoachingYears      INT,
    Gender              varchar(1) CONSTRAINT genderCheck CHECK (Gender IN ('M', 'F')),
    PRIMARY KEY     (Name, DateOfBirth),
    FOREIGN KEY (Name) REFERENCES Manage(headCoachName)
);

CREATE TABLE Award(
    TeamName          varchar(255) PRIMARY KEY,
    Competitions        TEXT,
    Titles              INT,
    FOREIGN KEY (TeamName) REFERENCES footballTeam(TeamName) ON DELETE CASCADE
);
```

```sql
CREATE TABLE Matches(
    GameDate        date CHECK (GameDate < date('now')),
    HomeTeam        varchar(255),
    AwayTeam        varchar(255),
    Location        TEXT,
    PRIMARY KEY    (GameDate, HomeTeam, AwayTeam),
    FOREIGN KEY (GameDate) REFERENCES Plays(GameDate),
    FOREIGN KEY (GameDate) REFERENCES Preside(GameDate)
);


CREATE TABLE Referees(
    Name            varchar(255),
    DateOfBirth     date  CHECK (DateOfBirth < date('now')),
    Gender          varchar(1)  CHECK (GENDER IN ('M', 'F')),
    Country         TEXT,
    PRIMARY KEY (Name, DateofBirth)
    FOREIGN KEY (Name) REFERENCES Preside(RefereeName)
);

CREATE TABLE TeamPlayers(
    Name            varchar(255),
    DateOfBirth     date  CHECK (DateOfBirth < date('now')),
    Gender          varchar(1)   CHECK (Gender IN ('M', 'F')),
    Height          FLOAT,
    Position        TEXT,
    Nationality     TEXT,
    PRIMARY KEY    (Name, DateOfBirth),
    FOREIGN KEY   (Name) REFERENCES BelongTo(PlayerName) ON DELETE CASCADE
);

CREATE TABLE RetiredPlayers(
    Name            varchar(255),
    DateofBirth     date  CHECK (DateOfBirth < date('now')),
    RetiredDate     date  CHECK (RetiredDate < date('now')),
    FOREIGN KEY  (Name, DateofBirth) REFERENCES TeamPlayers(name, DateOfBirth)
);

CREATE TABLE BelongTo(
    PlayerName    varchar(255),
    TeamName      varchar(255),
    DateofBirth    date  CHECK (DateOfBirth < date('now')),
    PRIMARY KEY (PlayerName, TeamName, DateofBirth)
);
```

```
CREATE TABLE Plays(
   TeamName           varchar(255),
   GameDate           date CHECK (GameDate < date('now')),
   HomeTeam           varchar(255),
   AwayTeam           varchar(255),
   PRIMARY KEY  (TeamName, GameDate, HomeTeam, AwayTeam)
);

CREATE TABLE Manage(
   headCoachName    varchar(255),
   TeamName           varchar(255),
   DateOfBirth        date CHECK (DateOfBirth < date('now')),
   PRIMARY KEY     (headCoachName, TeamName, DateOfBirth)
);

CREATE TABLE Preside(
   GameDate    date  CHECK (GameDate < date('now')),
   RefereeName varchar(255),
   HomeTeam           varchar(255),
   AwayTeam           varchar(255),
   PRIMARY KEY (GameDate, RefereeName, HomeTeam, AwayTeam)
);
```

Queries:

```
INSERT INTO footballTeam VALUES
('Inter Milan', 'Italy', 'Milan', '1908-03-09'),
('Chelsea','England','London', '1905-03-10'),
('Arsenal','England','London','1886-00-00'),
('Manchester City','England','Manchester','1880-01-01'),
('AC Milan','Italy','Milan','1899-12-16'),
('Real Madrid','Spain','Madrid','1902-03-06'),
('Barcelona','Spain','Baecelona','1899-11-29'),
('Bayern Munchen','Germany','Munchen','1900-02-27'),
('Borussia Dortmund','Germany','Dortmund','1909-12-19'),
('Paris Saint-Germain','France','Paris','1970-08-12');

INSERT INTO Manage Values
 ('Antonio Conte', 'Inter Milan'),
('Niko Kovac','Bayern Munchen'),
('Maurizio Sarri','Chelsea'),
('Unai Emery Etxegoien','Arsenal'),
('Josep Guardiola i Sala','Manchester City'),
('Gennaro Ivan Gattuso','AC Milan'),
('Zinedine Zidane','Real Madrid'),
('Ernesto Valverde Tejedor','Barcelona'),
('Alston Atwood','Borussia Dortmund'),
('Thomas Tuchel','Paris Saint-Germain');
```

```sql
INSERT INTO HeadCoach VALUES
('Antonio Conte','1957-11-02',2,'M'),
('Niko Kovač','1971-10-15',2,'M'),
('Maurizio Sarri','1959-01-10',2,'M'),
('Unai Emery Etxegoien', '1971-11-03', 1, 'M'),
('Josep Guardiola i Sala', '1971-01-08', 3, 'M'),
('Gennaro Ivan Gattuso', '1978-01-09', 1, 'M'),
('Thomas Tuchel','1973-08-29',2,'M'),
('Marco Giampaolo','1967-08-02',1,'M'),
('Zinedine Zidane','1972-06-23',2,'M'),
('Ernesto Valverde Tejedor','1964-02-09',2,'M'),
('Alston Atwood','1972-01-12',3,'M');

INSERT INTO Plays VALUES
 ('Inter Milan', '2019-06-02'), ('AC Milan', '2019-06-02'),
('Bayern Munchen','2018-04-01'),('Borussia Dortmund','2018-04-01'),
('Borussia Dortmund','2017-12-21'),('Bayern Munchen','2017-12-21'),
('Bayern Munchen','2017-04-09'),('Borussia Dortmund','2017-04-09'),
('Bayern Munchen','2018-11-09'),('Borussia Dortmund','2018-11-09'),
('Borussia Dortmund','2017-11-05'),('Bayern Munchen','2017-11-05'),
('Baecelona','2016-04-03'),('Real Madrid','2016-04-03'),
('Baecelona','2015-03-25'),('Real Madrid','2015-03-25'),
('Baecelona','2013-10-26'),('Real Madrid','2013-10-26'),
('Real Madrid','2015-11-22'),('Baecelona','2015-11-22'),
('Real Madrid','2014-10-25'),('Baecelona','2014-10-25');

INSERT INTO Matches VALUES
('2019-06-02', 'Inter Milan', 'AC Milan', 'Milan'),
('2018-04-01','Bayern Munchen','Borussia Dortmund','Munchen'),
('2017-12-21','Bayern Munchen','Borussia Dortmund','Munchen'),
('2017-04-09','Bayern Munchen','Borussia Dortmund','Munchen'),
('2018-11-09','Borussia Dortmund','Bayern Munchen','Dortmund'),
('2017-11-05','Borussia Dortmund','Bayern Munchen','Dortmund'),
('2016-04-03','Barcelona','Real Madrid','Barcelona'),
('2015-03-25','Barcelona','Real Madrid','Barcelona'),
('2013-10-26','Barcelona','Real Madrid','Barcelona'),
('2015-11-22','Real Madrid','Baecelona','Madrid'),
('2014-10-25','Real Madrid','Baecelona','Madrid');

INSERT INTO preside VALUES
('2019-06-02', 'Ivy Lavish'),
('2018-04-01','Charles Zhao'),
('2017-12-21','Andy Lucus'),
('2017-04-09','Steven Kolin'),
('2018-11-09','Ivy Lavish'),
('2017-11-05','Aiden Liam'),
('2016-04-03','Nathan Jayden'),
('2014-10-25','Nathan Jayden'),
('2015-11-22','Brandon Kulun'),
```

('2013-10-26','Anthony Tylor'),
('2016-04-03','Aiden Liam');

INSERT INTO Referees VALUES
('DuoAnziyi', '1991-02-14', 'M', 'China'),
('Ivy Lavish', '1995-07-01', 'F', 'England'),
('Charles Zhao','1993-09-28','M','America'),
('George He','1994-05-06','M','Canada'),
('Andy Lucus','1990-09-21','M','France'),
('Steven Kolin','1989-10-23','M','Italy'),
('David locke','1991-12-11','M','France'),
('Evyleen Dove','1990-03-24','F','Italy'),
('Anthony Tylor','1996-12-21','M','France'),
('Brandon Kulun','1995-11-30','M','England'),
('Aiden Liam','1992-07-04','M','Brazil'),
('Nathan Jayden','1990-05-07','M','England');

INSERT INTO BelongTo VALUES
('Robert Lewandowski','Bayern Munchen'),
('Manuel Neuer','Bayern Munchen'),
('Sven Ulreich','Bayern Munchen'),
('Thomas Muller','Bayern Munchen'),
('Niklas Sule','Bayern Munchen'),
('Jerome Agyenim Boateng','Bayern Munchen'),
('David Alaba','Bayern Munchen'),
('Joshua Kimmich','Bayern Munchen'),
('Thiago','Bayern Munchen'),
('Javi Martinez','Bayern Munchen'),
('Leon Goretzka','Bayern Munchen'),
('Corentin Tolisso','Bayern Munchen');

INSERT INTO TeamPlayers VALUES
('Robert Lewandowski','1988-08-21','M',1.85,'Striker','Poland'),
('Manuel Neuer','1986-03-27','M',1.93,'GoalKeeper','West Germany'),
('Sven Ulreich','1988-08-03','M',1.92,'GoalKeeper','West Germany'),
('Thomas Muller','1989-09-13','M',1.86,'Forward','West Germany'),
('Niklas Sule','1995-09-03','M',1.95,'Centre back','Germany'),
('Jerome Agyenim Boateng','1988-09-03','M',1.92,'Defender','Germany'),
('David Alaba','1992-06-24','M',1.80,'Defender','Austria'),
('Joshua Kimmich','1995-02-08','M',1.76,'Central midfielder','Germany'),
('Thiago','1991-04-11','M',1.74,'Central midfielder','Italy'),
('Javi Martinez','1988-09-02','M',1.90,'Certre back','Spain'),
('Leon Goretzka','1995-02-06','M',1.89,'Central midfielder','Germany'),
('Corentin Tolisso','1994-08-03','M',1.80,'midfielder','France');

INSERT INTO Award VALUES
('Bayern Munichen','2009-04-18 CXE',1),
('Barcelona','2017-01-31 CXE',2),
('Paris Saint-Germain','2005-11-01 CXE',1),

('Real Madrid','2012-06-30 CXE',1),
('Manchester City','2014-02-28 TAE',3),
('Arsneal','2015-05-31 CNM',2),
('AC Milan','2000-10-11 CNM',1),
('Inter Milan','2017-10-12 CNM',3),
('Chelsea','2003-06-01 TAE',5),
('Liverpool','2003-01-12 ZIYI',1),
('Bayern Munchen','1998-01-01 SSR',2);

INSERT INTO RetiredPlayers VALUES
('GeorgeHe', '1978-02-14', '1998-02-14'),
('ZIYI ANNE', '1968-02-14', '2016-02-28'),
('Bobby Bart', '1978-02-14', '2008-04-28'),
('Nancy Lee', '1996-02-28', '2019-06-31'),
('Nils Burinben', '1986-02-28', '2013-02-28'),
('Charlse Munchen', '1978-02-14', '1996-02-28'),
('Garry Yapa', '1978-02-14', '1996-02-28'),
('Jiandooz Hey', '1978-03-14', '1996-02-28'),
('Doris Wu', '1978-02-11', '2005-02-28'),
('Perry Hansen', '1968-05-24', '1996-02-28'),
('Bruce Bell', '1998-02-14', '1996-01-31');