COS 221 Practical 5

Untouchables

Team members:

Adam Osborne, Janco Spies, Selepo Sello

Paul Nhlapo, Lekgotla Motaung,

Tshegofatso Mapheto, Thapelo Thoka,

Task 1: Research

Chosen Sport: Athletics (100m sprint)

What is Athletics?
Athletics refers to a group of sporting events that involves competitive running, jumping, throwing, and walking. The most common athletics competitions are track and field, road running, cross country running, and racewalking. Even though in other parts of the world the term "athletics" refers to all sports in general, here we are talking specifically about the above-mentioned sporting codes as athletics. It is a sport which was founded in the early 9th century where it tested the physical endurance of competitors. At that time there were not as many sporting codes under the athletics label and it consisted mainly of footrace.

How does the sport work?
Athletics is a very competitive sport that is usually played by one individual participating against other individuals, but it is not only limited to this single-player format. There are other games that are played in teams like relay and cross country running which may be played as a team. The single-player games include jumping games (long jump and high jump) where the winner is determined by the highest jump in the high jump or by the longest jump in the long jump. The team games include:

- Relay is a sport consisting of a set of stages usually 4 where each stage is completed by a different member of the team. The runner finishing one stage is required to pass on a baton to the next runner while both are running in a marked exchange zone.

- Cross country running is a sport in which teams or individuals run a race on open-air courses over natural terrains such as dirt or grass. To win the race just like any other racing game the team or the individual has to come out first and reach the finish line before the other teams.

- Racewalking is a long-distance discipline where one's foot must appear to be in contact with the ground at all times. The rules of the game are simple: your foot must always be in contact with the ground until the race finishes and judges carefully assess whether the rule is followed throughout the race. To win you and your team just have to reach the finish line before other teams.

Tournament structure of the Athletics
The games of events are played in different seasons. The Athletics have: Racewalking Cross country running Track and field (Long jump, high jump, race, etc)

Categories
Here we have two females and males divisions which are also differentiated by age and abilities and disabilities of the participants.

Venues
Events such as long jump, high jump, relay and sprint are played at track and field stadiums. Then the other games like cross country racing are done outside on natural terrains.

Competitions
Competitions are usually held once every year nationally and internationally. Then there is a big main event that every athlete is looking forward to which is the Olympic Games. The Olympic Games happen every four years, alternating between summer and winter Olympics every two years in the four-year period.

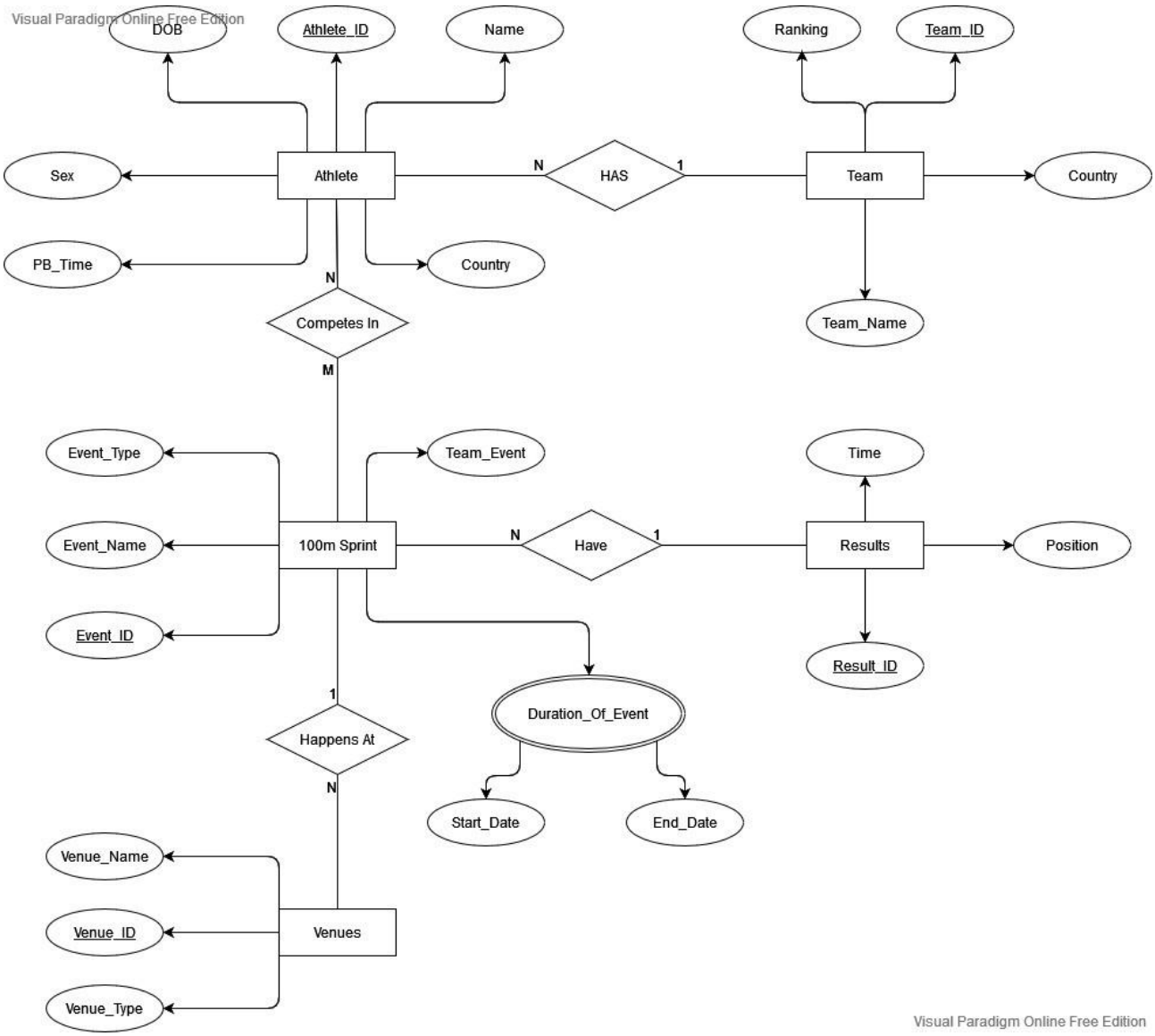
Task 2 (E)ER Diagram

Assumptions made:

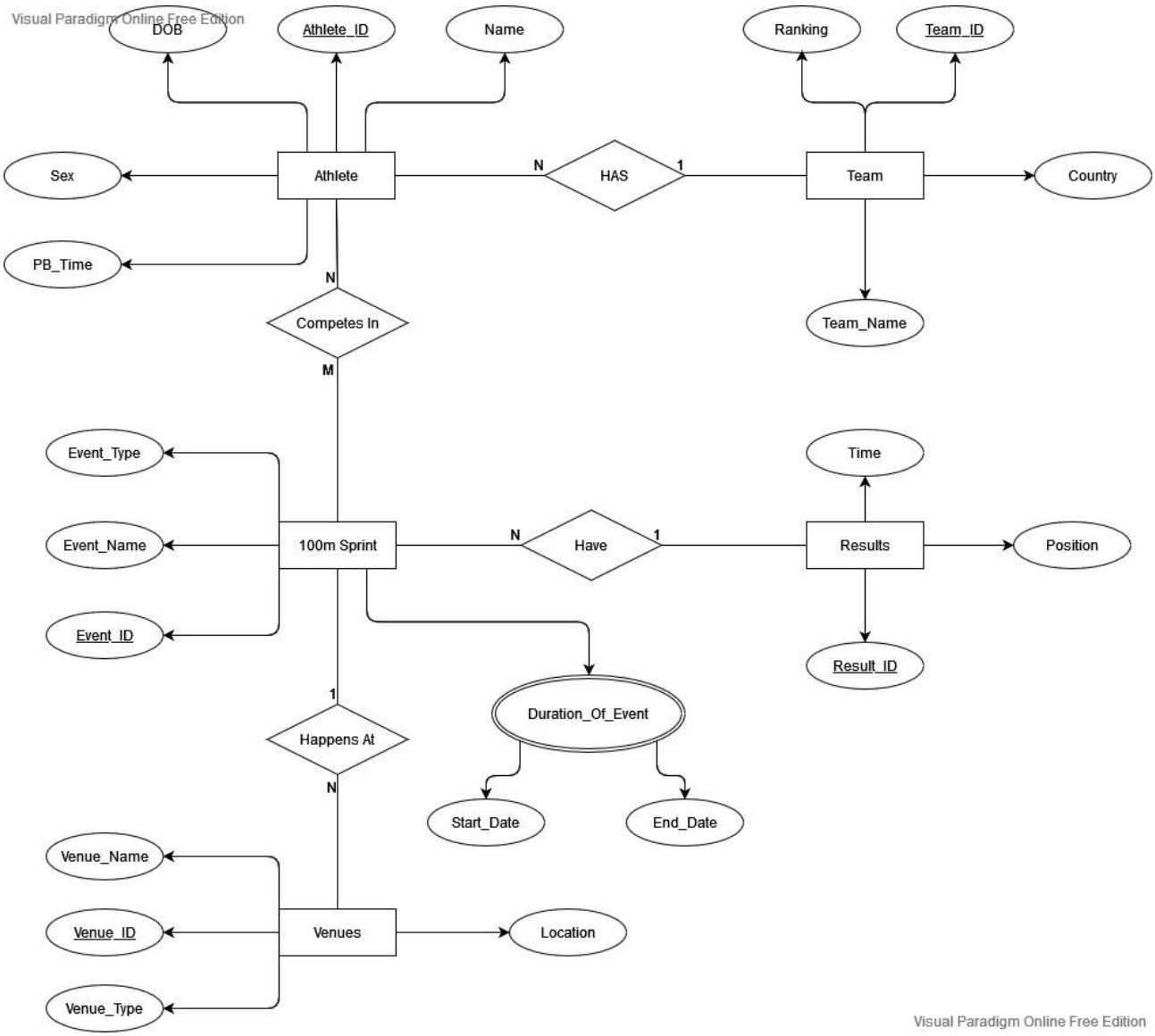
- All athletes must be part of exactly one team.

- All athletes who participate in an event finish the race and thus are awarded a time result

This is our first version of the EER diagram:



Wedecidedtoremovethecountryfieldfromtheathletetable,sincetheteamfieldalreadystoresthecountry oftheathlete.WealsoremovedtheTeam -Eventfieldfromthe100msprinttablesinceall100msprintevents areindividualevents.ThisisourfinalversionoftheEERdiagram:



Task 3 Relational mapping

The steps for the relational mapping are as follows:

Step 1: Mapping of Athlete, Team, 100m Sprint, Results and Venues entities(strong entity types):

TeamID attribute is added to Athlete and is a foreign key derived from TeamID in the Team entity The simple components of the composite attribute ”Duration Of Event” are included in the Duration Of Event entity. The primary key of ”Duration Of Event”, which is Event ID is selected as the primary key of the relation.

Step 2: (There are no weak entity types to map.)

Step 3:(There are no binary 1:1 relationships to map.)

Step4:Mapping1:Nrelationships

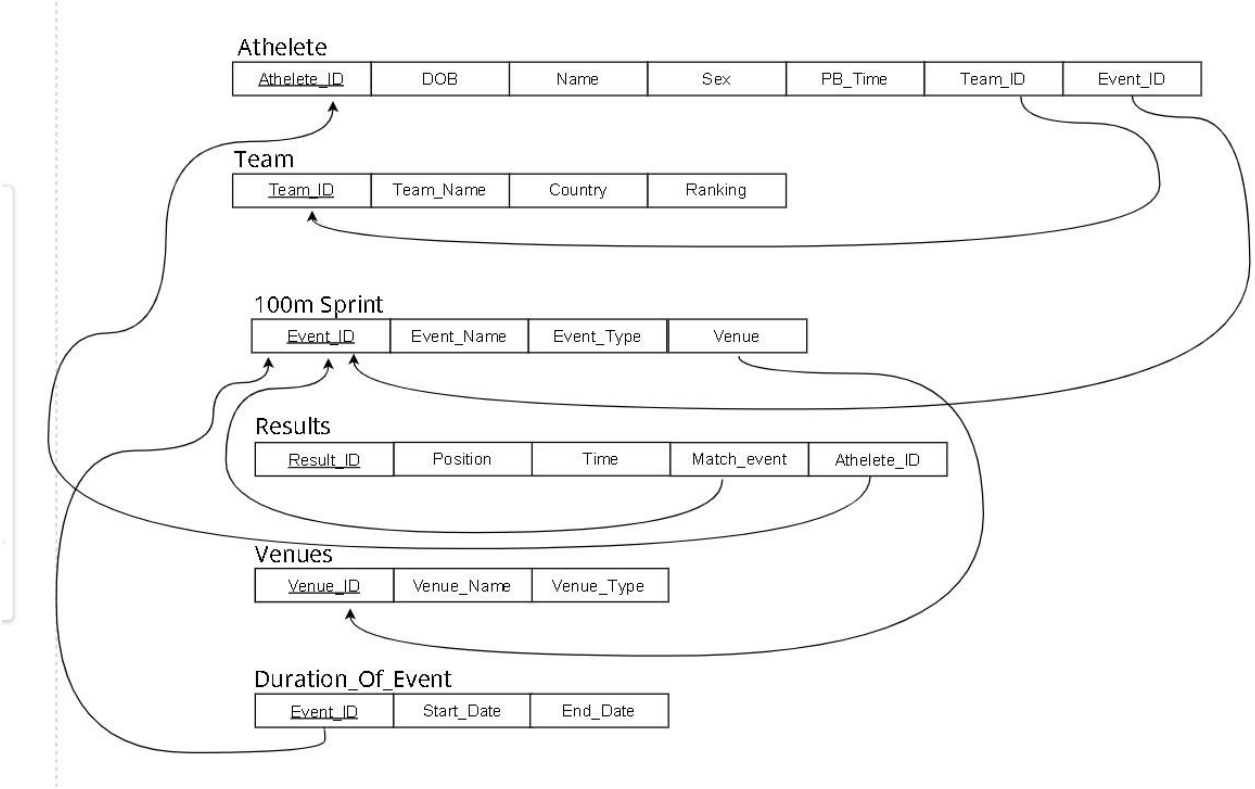
A ”Venue” attribute is added to 100m sprint and it is a foreign key to the VenueID field in the Venues entity for the relationship between ”100m Sprint” and ”Venues”.This relates to the mapping of the ”Happens At” relationship. A ”Match event” attribute is added to the Results entity, this is a foreign key to Event ID which is the primary key of the 100m Sprint entity.This relates to the mapping of the ”Have” relationship. A ”Team ID” attribute is added to the Athlete entity, this is a foreign key to the Team ID field in the Team entity. This relates to the mapping of the ”HAS” relationship. A ”Event ID” attribute is added to the Athlete entity, this is a foreign key to the ”Event _ID” field in the 100m Sprint entity. This relates to the mapping of the ”Competes in” relationship.

Step 5:(There are no binary M:N relationships to map.)

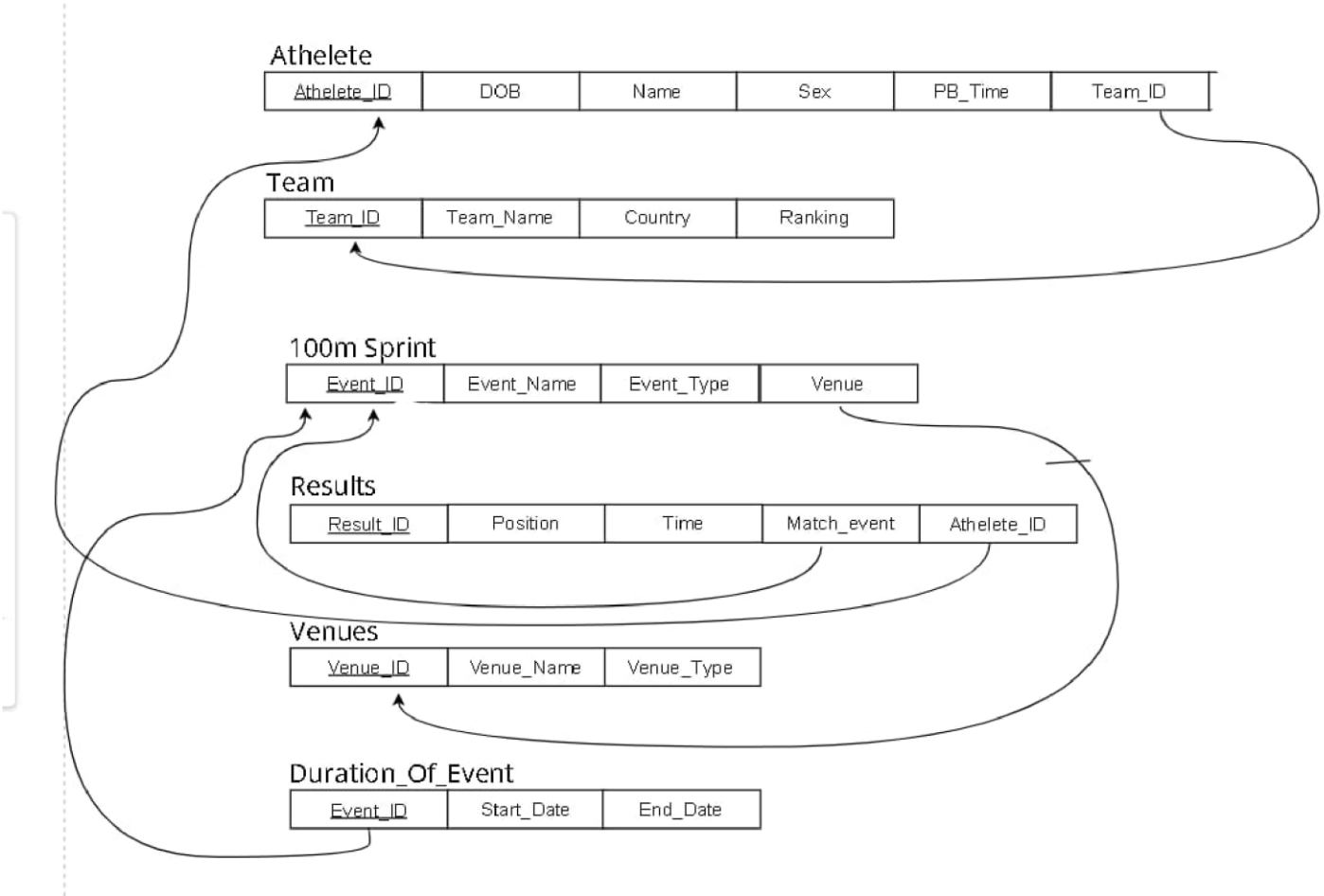
Step 6: Mapping of multivalued attributes.

The relation ”Duration Of Event” has been added in order to map the ”Duration Of Event” multivalued attribute in the 100m Sprint entity.

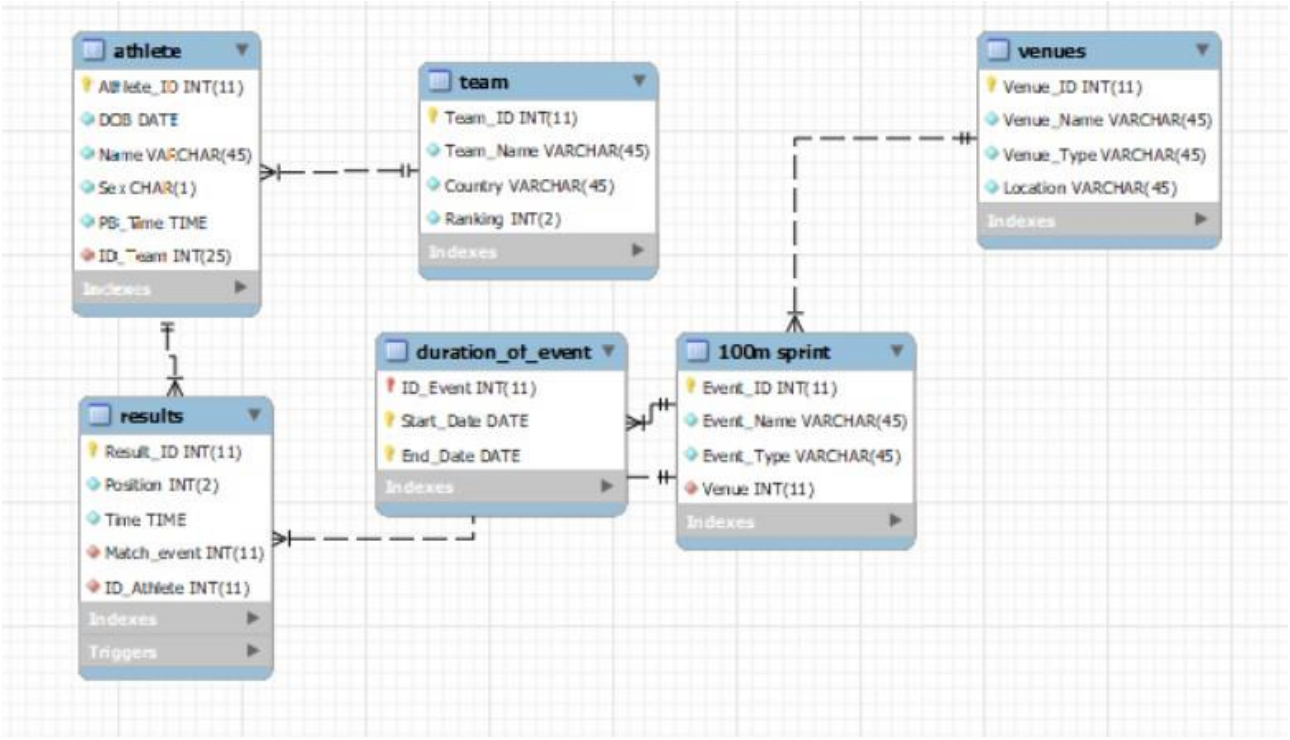
This is our first version of the relational model:



We decided to remove the Event ID field from the athlete table, since this would mean that each athlete can only participate in one event or would have duplicate entries for each event they participated in. This is our final version of the relational model:


Athelete
Athelete_ID
DOB
Name
Sex
PB_Time
Team_ID
Team
Team_ID
Team_Name
Country
Ranking
100m Sprint
Event_ID
Event_Name
Event_Type
Venue
Results
Result_ID
Position
Time
Match_event
Athelete_ID
Venues
Venue_ID
Venue_Name
Venue_Type
Duration_Of_Event
Event_ID
Start_Date
End_Date

Task 4 Relational exclusion



```
CREATE DATABASE IF NOT EXISTS ‘uxxxxxxxx_untouchables‘ /*!40100 DEFAULT CHARACTER SET utf8mb4 */; USE
‘uxxxxxxxx_untouchables‘;
-- MariaDB dump 10.19 Distrib 10.7.3-MariaDB, for Win64 (AMD64) --
-- Host: wheatley.cs.up.ac.za Database: u21434159_untouchables
-- ------------------------------------------------------
-- Server version 10.3.31-MariaDB-0+deb10u1

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!40101 SET NAMES utf8 */;
/*!40103 SET @OLD_TIME_ZONE=@@TIME_ZONE */;
/*!40103 SET TIME_ZONE=’+00:00’ */;
/*!40014 SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0 */;
/*!40014 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0 */;
/*!40101 SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE=’NO_AUTO_VALUE_ON_ZERO’ */;
/*!40111 SET @OLD_SQL_NOTES=@@SQL_NOTES, SQL_NOTES=0 */;

--
-- Table structure for table ‘100m sprint‘
--

DROP TABLE IF EXISTS ‘100m sprint‘;
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */; CREATE TABLE
‘100m sprint‘ (
  ‘Event_ID‘ int(11) NOT NULL AUTO_INCREMENT,
  ‘Event_Name‘ varchar(45) NOT NULL,
  ‘Event_Type‘ varchar(45) NOT NULL,
  ‘Venue‘ int(11) NOT NULL,
  PRIMARY KEY (‘Event_ID‘),
  KEY ‘Venue_idx‘ (‘Venue‘),
  CONSTRAINT ‘Venue‘ FOREIGN KEY (‘Venue‘) REFERENCES ‘venues‘ (‘Venue_ID‘) ON DELETE NO ACTION ON UPDATE NO
      ACTION
) ENGINE=InnoDB AUTO_INCREMENT=9 DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */; --
```

```
-- Table structure for table ‘athlete‘
--

DROP TABLE IF EXISTS ‘athlete‘;
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */; CREATE TABLE
‘athlete‘ (
  ‘Athlete_ID‘ int(11) NOT NULL AUTO_INCREMENT,
  ‘DOB‘ date NOT NULL,
  ‘Name‘ varchar(45) NOT NULL,
  ‘Sex‘ char(1) NOT NULL,
  ‘PB_Time‘ time(2) NOT NULL DEFAULT ’00:00:00.00’,
  ‘ID_Team‘ int(25) NOT NULL,
  PRIMARY KEY (‘Athlete_ID‘),
  KEY ‘Event_ID_idx‘ (‘ID_Team‘),
  CONSTRAINT ‘Team_ID‘ FOREIGN KEY (‘ID_Team‘) REFERENCES ‘team‘ (‘Team_ID‘) ON DELETE NO ACTION ON UPDATE NO
      ACTION
) ENGINE=InnoDB AUTO_INCREMENT=56 DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Table structure for table ‘duration_of_event‘ --

DROP TABLE IF EXISTS ‘duration_of_event‘;
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */;
CREATE TABLE ‘duration_of_event‘ (
  ‘ID_Event‘ int(11) NOT NULL,
  ‘Start_Date‘ date NOT NULL,
  ‘End_Date‘ date NOT NULL,
  PRIMARY KEY (‘ID_Event‘,‘Start_Date‘,‘End_Date‘),
  KEY ‘Events_idx‘ (‘ID_Event‘),
  CONSTRAINT ‘Events‘ FOREIGN KEY (‘ID_Event‘) REFERENCES ‘100m sprint‘ (‘Event_ID‘) ON DELETE NO ACTION ON UPDATE NO
      ACTION
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Table structure for table ‘results‘
--

DROP TABLE IF EXISTS ‘results‘;
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */; CREATE TABLE
‘results‘ (
  ‘Result_ID‘ int(11) NOT NULL AUTO_INCREMENT,
  ‘Position‘ int(2) NOT NULL,
  ‘Time‘ time(2) NOT NULL,
  ‘Match_event‘ int(11) NOT NULL,
  ‘ID_Athlete‘ int(11) NOT NULL,
  PRIMARY KEY (‘Result_ID‘),
  KEY ‘Event ID_idx‘ (‘Match_event‘),
  KEY ‘athlete_idx‘ (‘ID_Athlete‘),
  CONSTRAINT ‘Event‘ FOREIGN KEY (‘Match_event‘) REFERENCES ‘100m sprint‘ (‘Event_ID‘) ON DELETE NO ACTION ON UPDATE
      NO ACTION,
  CONSTRAINT ‘athlete‘ FOREIGN KEY (‘ID_Athlete‘) REFERENCES ‘athlete‘ (‘Athlete_ID‘) ON DELETE NO ACTION ON UPDATE NO
      ACTION
```

```
) ENGINE=InnoDB AUTO_INCREMENT=65 DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Table structure for table ‘team‘
--

DROP TABLE IF EXISTS ‘team‘;
```

```
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */; CREATE TABLE
‘team‘ (
  ‘Team_ID‘ int(11) NOT NULL AUTO_INCREMENT,
  ‘Team_Name‘ varchar(45) NOT NULL,
  ‘Country‘ varchar(45) NOT NULL,
  ‘Ranking‘ int(2) NOT NULL, PRIMARY KEY
  (‘Team_ID‘)
) ENGINE=InnoDB AUTO_INCREMENT=21 DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Table structure for table ‘venues‘
--

DROP TABLE IF EXISTS ‘venues‘;
/*!40101 SET @saved_cs_client = @@character_set_client */;
/*!40101 SET character_set_client = utf8 */; CREATE TABLE
‘venues‘ (
  ‘Venue_ID‘ int(11) NOT NULL AUTO_INCREMENT,
  ‘Venue_Name‘ varchar(45) NOT NULL,
  ‘Venue_Type‘ varchar(45) NOT NULL,
  ‘Location‘ varchar(45) NOT NULL, PRIMARY KEY
  (‘Venue_ID‘)
) ENGINE=InnoDB AUTO_INCREMENT=5 DEFAULT CHARSET=utf8;
/*!40101 SET character_set_client = @saved_cs_client */;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
/*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
/*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;

-- Dump completed on 2022-06-07 15:01:58
```

Task 6

The sample data for our database is real-world Olympic Games data sourced from the website https://olympics.com/ and was captured by hand in a php script. We chose to capture our data via this method since it allows the data to be loaded easily from the web application into the database every time it is launched, ensuring that the application always interacts with the correct version of the data.

Task 7

Query #1

The first query we will consider is one that calculates the average time of all athletes in a team and then sorts the teams by their average times.

```
SELECT t.Team_Name, Avg(r.Time) as AvgTime
FROM team as t
INNER JOIN athlete as a
ON t.Team_ID = a.ID_Team
INNER JOIN results as r
ON a.Athlete_ID = r.ID_Athlete
```

```
GROUP BY t.Team_Name
ORDER BY AvgTime;
```

This query is very slow since it performs two joins on entire tables and takes 0.0054 seconds to execute. We can optimize this query by limiting the amount of columns projected out before the joins so that the tables that are joined are smaller:

```
SELECT t.Team_Name, Avg(f1.Time) as AvgTime
FROM (
    SELECT Time, ID_Team
    FROM athlete as a INNER JOIN
    results as r ON
    a.Athlete_ID = r.ID_Athlete) as f1
INNER JOIN team as t ON f1.ID_Team = t.Team_ID
GROUP BY t.Team_Name
ORDER BY AvgTime;
```

Our new query only takes 0.0044 seconds to execute which means we achieved an 18% gain in speed by changing our query structure.

Query #2

The second query we will consider is one displays all athletes that have finished on the podium and their respective teams and personal best times.

```
SELECT a.Name, t.Team_Name, a.PB_Time
FROM athlete as a
INNER JOIN team as t
ON a.ID_Team = t.Team_ID
INNER JOIN results as r
ON a.Athlete_ID = r.ID_Athlete
WHERE r.Position < 4
ORDER BY r.Position;
```

This query is very slow since it performs two joins on entire tables and takes 0.0096 seconds to execute. We can optimize this query by limiting the amount of rows and columns projected out before the joins so that the overall table that have to be joined as smaller:

```
SELECT a.Name, t.Team_Name, a.PB_Time
FROM (SELECT Name, PB_Time, Athlete_ID, ID_Team FROM athlete) as a
INNER JOIN (SELECT Team_Name, Team_ID FROM team) as t
ON a.ID_Team = t.Team_ID
INNER JOIN (SELECT Position, ID_Athlete FROM results WHERE Position < 4) as r
ON a.Athlete_ID = r.ID_Athlete
ORDER BY r.Position;
```

Our new query only takes 0.0046 seconds to execute which means we achieved an 52% gain in speed by changing our query structure.

Query #3

The third query we will consider is the one that does self-join of two teams and then orders them by country name:

```
SELECT A.Team_Name AS Team_Name1, B.Team_Name AS Team_Name2, A.Country
FROM team A, team B
WHERE A.Team_ID <> B.Team_ID AND A.Country = B.Country
ORDER BY A.Country
```

This query is very fast to perform, it does a self join on columns that have the same country name this is because at the moment the database does not have two teams from the same country. The time taken for the query to execute is 0.0026 seconds.