Module Code: CS2DI17
Assignment report Title: Database Design and Implementation for Horseracing Media Centre
Student Number(s): 30008602, 30006477, 30020691
Date (when the work completed): 09/03/23
Actual hrs spent for the assignment: 6
Assignment evaluation (3 key points):

- *Understanding of database design using ER models*
- *Learning how to normalise ER models up to and including third normal form*
- *Using SQL to insert data and create queries to retrieve specific data needed.*

Table of Contents

Solution Design:

Task 1

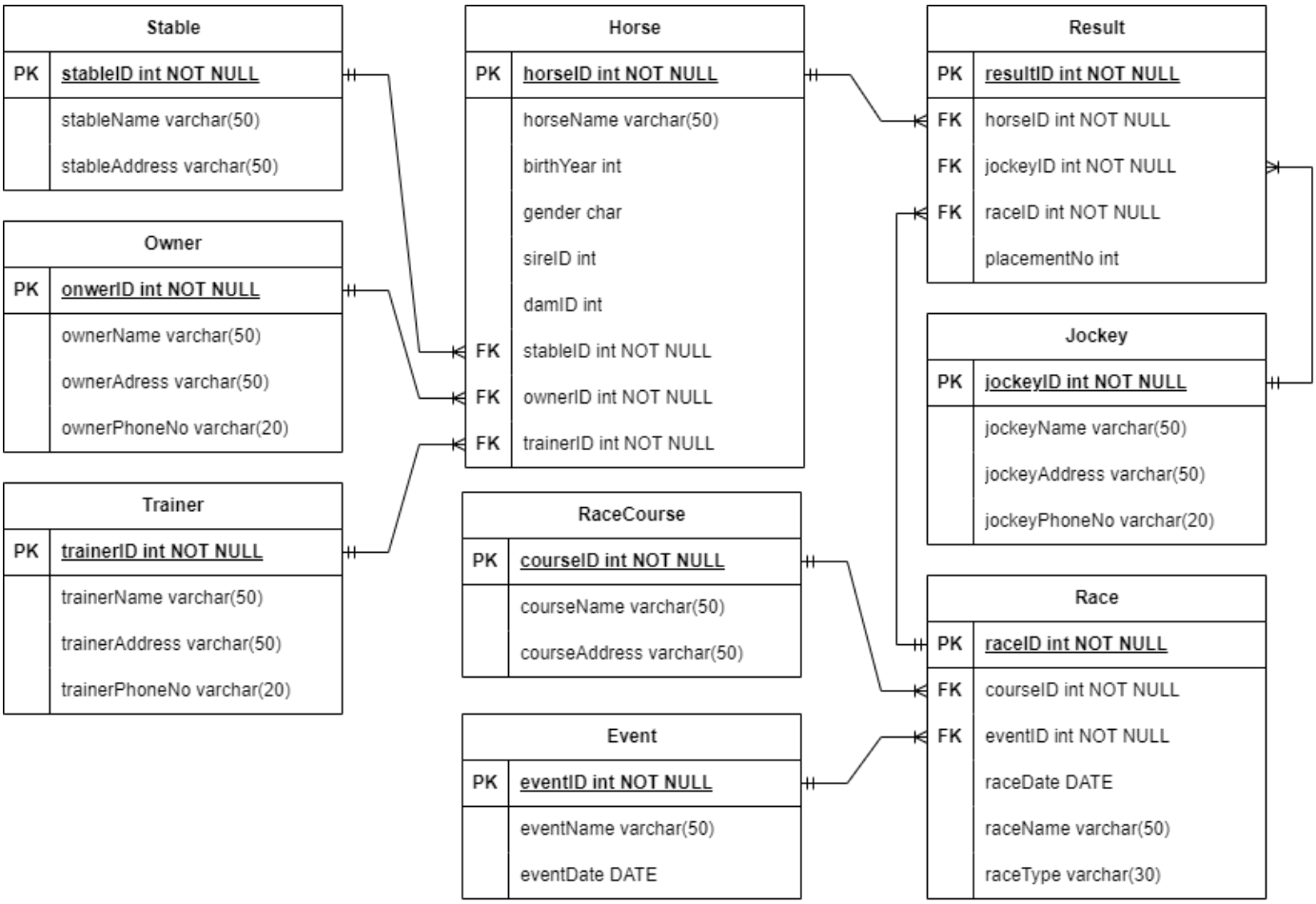

Figure 1 ER Model

Constraints:

- All primary keys cannot be NULL e.g. HorseID
- Horse table references stableID, ownerID and trainerID as foreign keys
- Race table references courseID and eventID as foreign keys
- Result table references horseID, jockeyID and raceID as foreign keys
- A horse must have one stable, owner and trainer
- A result must have one race, jockey, and horse
- A race must have one racecourse and event.

Assumptions:

- Sire and Dam horses are also racehorses
- A stable can have multiple horses
- An owner can have multiple horses
- A trainer can have multiple horses
- A course can have multiple races
- An event can have multiple races
- A race can have multiple results
- A horse can have multiple results
- A jockey can have multiple results

Task 2

Normalisation is the process in which redundant data is removed and the database consistency is improved. Large tables are split into smaller, more specialised ones to make it more logical. For our table to be normalised up to and including 3NF we went through each stage and adjusted the tabled based on the requirements.

To make our model meet first normal form (1NF) we made sure each table had a primary key so that each table had a unique data point for reference. This is set to not be null as it is required for each row in that table. Primary keys can also be used in other tables using foreign keys. Figure 1 shows each table with a primary key e.g., “horseID”, “resultID”, “stableID” etc. We also need to make sure that each table has no repeating data and that each cell has one value. This applies for our model as no column has the same name and also will only allow one input field. E.g., trainerName only allows for one name and does not state multiple trainers.

To make our model second normal form (2NF), we made sure it doesn’t have any partial dependencies, which means that each non-key attribute is dependent on the primary key. For example, in the Horse table, the “horseID” is the primary key, and the other attributes are dependent on the primary key, this is the same for all the tables making them 2NF.

Our model is also in third normal form (3NF). This is because it does not have any transitive dependencies, meaning that each non-key attribute is not dependant on another non-key attribute. For example, in the Horse table, the “sireID” and “damID” are both dependant on the “horseID”, but not each other, this is again the same for each table making it 3NF.

Solution Implementation:

Task 3

We used MySQL as a workbench for inserting and testing queries to retrieve sets of data.

Creating the table:

```
CREATE TABLE `Horse` (
    `horseID` INT  NOT NULL ,
    `name` text  NOT NULL ,
    `birthYear` YEAR  NOT NULL ,
    `gender` CHAR  NOT NULL ,
    `sireID` INT  NOT NULL ,
    `damID` INT  NOT NULL ,
    `stableID` INT  NOT NULL ,
    `ownerID` INT  NOT NULL ,
    `trainerID` INT  NOT NULL ,
    PRIMARY KEY (
        `horseID`
    )
);

CREATE TABLE `Owner` (
    `OwnerID` INT  NOT NULL ,
    `name` text  NOT NULL ,
    `address` text  NOT NULL ,
    `phoneNo` text  NOT NULL ,
    PRIMARY KEY (
        `OwnerID`
    )
);
```

Figure 2 SQL statements for creating tables

Populating the table:

```
INSERT INTO Trainer (trainerID, name, address, phoneNo)
VALUES
    (1, 'Mike Smith', '111 Oak St', '447488810589'),
    (2, 'Sara Jones', '222 Pine Ave', '447488810587'),
    (3, 'Tom Williams', '333 Cedar Rd', '447458196483');

INSERT INTO Racecourse (courseID, name, location)
VALUES
    (1, 'Churchill Downs', 'Louisville'),
    (2, 'Santa Anita Park', 'Arcadia'),
    (3, 'Saratoga Race Course', 'Saratoga Springs');
```

Figure 3 SQL statements for populating tables

All table data:

horseID	name	birthYear	gender	sireID	damID	stableID	ownerID	trainerID
1	Thunderbolt	2018	M	15	13	2	1	2
2	Firecracker	2019	F	15	14	3	3	3
3	Lightning	2016	M	11	12	2	1	2
4	Sonic	2017	M	11	12	1	3	1
5	Blaze	2015	F	10	11	2	5	1
6	Phoenix	2020	F	9	8	2	4	3
7	Aurora	2014	F	15	14	2	1	2
8	Rocket	2013	M	16	17	3	3	2
9	Comet	2012	M	15	11	2	3	3
10	Meteor	2011	M	19	21	2	4	1
11	Rainbow	2010	F	23	14	2	4	1
12	Moonlight	2009	F	25	27	1	5	2
13	Starlight	2008	F	24	21	2	1	1
14	Sunshine	2007	F	23	27	3	3	3
15	Storm	2006	M	5	30	29	2	3

Figure 4 Horse table

OwnerID	name	address	phoneNo
1	John Smith	123 Main St	447893932101
2	Mary Johnson	456 High St	447700185211
3	Bob Brown	789 Maple Ave	447700184666
4	Alice Green	111 Oak St	447700184665
5	David Davis	222 Pine St	447488811031

Figure 5 Owner table

jockeyID	name	address	phoneNo
1	David Lee	444 Elm St	447700150913
2	Lynn Bridgeman	555 Maple Ave	447893932118
3	Robert Garcia	666 Cedar Rd	447893932116
4	Emily Brown	777 Walnut St	447893932114
5	Michael Davis	888 Oak Ave	447893932113
6	Amy Johnson	999 Pine Ave	447893932110
7	Martin Jones	1011 Cedar Rd	447893932107
8	Julia Smith	1213 Walnut St	447893932108
9	Ahmed Bedair	1415 Elm St	447893932101
10	Lauren Lee	1617 Maple Ave	447893932104

Figure 6 Jockey table

trainerID	name	address	phoneNo
1	Mike Smith	111 Oak St	447488810589
2	Sara Jones	222 Pine Ave	447488810587
3	Tom Williams	333 Cedar Rd	447458196483

Figure 7 Trainer table

courseID	name	location
1	Churchill Downs	Louisville
2	Santa Anita Park	Arcadia
3	Saratoga Race Course	Saratoga Springs

Figure 8 Course table

eventID	eventName	eventDate
1	American Oaks	2022-05-09
2	Arkansas Derby	2022-12-30
3	Florida Derby	2023-03-29

Figure 9 Event table

raceID	courseID	raceName	raceDate	raceType
1	1	Turf Classic Stakes	2022-05-09	flat race
2	1	Man o War Stakes	2022-05-10	endurance race
3	2	Gamely Stakes	2022-12-30	harness race
4	2	Shoemaker Mile	2023-01-02	flat race
5	3	Pegasus World Cup	2023-03-29	flat race
6	3	Beholder Mile Stakes	2023-03-30	flat race

Figure 10 Course table

resultID	horseID	jockeyID	raceID	eventID	placementNumber
1	1	1	1	1	1
2	2	2	1	1	2
3	3	3	1	1	3
4	4	4	1	1	4
5	5	5	1	1	5
6	6	6	1	1	6
7	7	7	1	1	7
8	8	8	1	1	8
9	9	9	1	1	9
10	10	10	1	1	10
11	6	2	2	1	3
12	7	1	2	1	6
13	8	3	2	1	1
14	9	5	2	1	7
15	10	4	2	1	2
16	11	6	2	1	9
17	12	8	2	1	4
18	13	7	2	1	5
19	14	9	2	1	8
20	15	10	2	1	10
21	4	3	3	2	1
22	6	6	3	2	9
23	7	1	3	2	3
24	3	7	3	2	8
25	8	2	3	2	4
26	12	9	3	2	5
27	10	4	3	2	7
28	9	5	3	2	10
29	11	8	3	2	2
30	14	10	3	2	6

Figure 11 Result table

	31	1	1	4	2	4
	32	2	9	4	2	8
	33	15	3	4	2	1
	34	12	4	4	2	6
	35	4	5	4	2	2
	36	6	7	4	2	7
	37	8	6	4	2	3
	38	14	10	4	2	10
	39	3	9	4	2	9
	40	5	8	4	2	5
	41	2	1	5	3	1
	42	1	2	5	3	3
	43	3	3	5	3	10
	44	4	4	5	3	6
	45	7	5	5	3	9
	46	5	6	5	3	5
	47	10	7	5	3	8
	48	8	8	5	3	7
	49	9	9	5	3	4
	50	15	10	5	3	2
	51	1	1	6	3	8
	52	4	2	6	3	7
	53	3	3	6	3	10
	54	10	4	6	3	2
	55	6	5	6	3	4
	56	11	6	6	3	9
	57	7	7	6	3	3
	58	5	8	6	3	5
	59	9	9	6	3	1
	60	13	10	6	3	6

Figure 12 Result table cont.

Task 4

Test 1- Who has ridden the winners this year?

```
SELECT Jockey.name AS JockeyName
FROM Result
JOIN Jockey ON Result.jockeyID = Jockey.jockeyID
JOIN Race ON Result.raceID = Race.raceID
WHERE YEAR(Race.raceDate) = YEAR(NOW()) AND
Result.placementNumber = 1;
```

JockeyName
Robert Garcia
David Lee
Ahmed Bedair

To retrieve the winners who won this year we need to select all the jockeys and filter it down to the ones that have a result including placementNumber = 1 and ones that took place this year using the raceDate. In addition, we use join statements to link the jockey to the result with their respected IDs as well as linking the race to the result.

Test 2 Produce a league table of winning trainers.

```
SELECT Trainer.name AS Trainer, COUNT(*) AS Wins
FROM Result
JOIN Horse ON Result.horseID = Horse.horseID
JOIN Trainer ON Horse.trainerID = Trainer.trainerID
GROUP BY Trainer.name
ORDER BY Wins DESC;
```

Trainer	Wins
Sara Jones	22
Tom Williams	19
Mike Smith	19

This SQL query retrieves data from tables named Result, Horse, and Trainer. It groups the trainers by name and counts the number of wins for each trainer by joining the tables on their respective IDs. Finally, it sorts the results in descending order of wins.

Test 3 List all the winners of the Florida Derby and Arkansas Derby in the same year.

```
SELECT Horse.name AS Horse, Event.eventName AS Event,
Result.placementNumber AS Placement, YEAR(racedate) AS YEAR
FROM Horse
JOIN Result ON Horse.horseID = Result.horseID
JOIN Race ON Result.raceID = Race.raceID
JOIN Event ON Result.eventID = Event.EventID
WHERE Event.eventName IN ('Florida Derby', 'Arkansas Derby')
AND Result.placementNumber = 1
AND YEAR(racedate) = 2023;
```

This query joins tables "Horse", "Result", "Race", and "Event" to select the names of horses, the name of the event, the placement number, and the year of the race for all races in the events 'Florida Derby' and 'Arkansas Derby' where the placement number is 1 and the year is 2023 to retrieve the winners of the ‘Florida Derby’ and ‘Arkansas Derby’ in the same year.