Swinburne University of Technology

Fundamentals of Data Management (2020 S2)

Doubtfire Submission

Test

Submitted By:
Yashkumar Pragneshkumar Mehta
102447560
2020/10/12 16:15

Tutor:
Sharon Stratsianis

October 12, 2020

COS20015 Fundamentals of Data Management

UNIT TEST- **Time** one (1) hour 3.00 pm to 4.00 pm (Total marks 100)

Please follow the instructions and upload your answers (screenshots are preferred) as a .pdf or Ms Word file into the Canvas. If you cannot access the software please type the answers. Your answers must be on the paper itself.

The submission link in the Canvas (Test) will be locked after 4.10 pm. Please upload your files before 4.10 pm on Monday 12th October 2020.

Student ID	102447560
Student Name	MEHTA YASHKUMAR
Class Time Eg.(Mon 8.30 – 10.30	Thursday 8.30 – 10.30

Answer All the questions. Make sure your answers are short and precise.

You can use (Q1.txt or Ed) files to check your answers for Q1.

1. The phone numbers collected from questionnaire is a mess. Design a regular expression to filter out those numbers that are stored in the standard format "+00-0-0000-0000" from the file called "Q1.txt" and redirect the results to the "cleaned.txt".

    Note: Only +61-3-9214-4980 and +61-3-9285-7706 are the valid results.   [10 Marks]

    grep -E '+[0-9]{2}-[0-9]-[0-9]{4}-[0-9]{4}' Q1.txt > cleaned.txt

2. Consider the following XML file (Q2.xml is given):

    ```
    <root>
     <students>
    ```

```
<element>
        <ID>100345</ID>
        <Nationality>USA</Nationality>
        <Program>ICT</Program>
        <age>23</age>
        <name>John</name>
</element>
<element>
        <ID>100876</ID>
        <Nationality>MALAYSIA</Nationality>
        <Program>CS</Program>
        <age>28</age>
        <name>Awang</name>
</element>
<element>
        <ID>100257</ID>
        <Nationality>AUSTRALIA</Nationality>
        <age>25</age>
        <name>Alex</name>
</element>
</students>
</root>
```

Write an XQUERY to display the information for all students who are not Malaysians or older than 25. You can use following website to check the answer: [10 Marks]

http://videlibri.sourceforge.net/cgi-bin/xidelcgi

```
for $y in root/students/element
where $y/Nationality != "MALAYSIA" or $y/age > 25
return ($y/ID, $y/Nationality, $y/Program, $y/age, $y/name)
```

3. Consider the following information to design a database for a catalogue of an online video service that offers access to popular TV series.

The following is the description of the application:
- Each series has a name and a description.
- Each series has many episodes. Each episode of a series has an episode number, the date it was first aired, and the length of the episode in minutes. The episode number uniquely identifies an episode with respect to the series, but two different series can have the same episode number.
- Viewers can comment on any episode. For each comment, we want to record its post date, content and the first name/last name of the viewer who commented.
- In this database:
  - The name of each series is unique.
  - A user can comment on one or more episodes of any series.
  - You may add extra columns if you need for primary keys.
  - You can model the user as a separate entity but registering the username with each comment is also acceptable.

Based on the above scenario, draw an ER or UML diagram. You can draw the diagram using https://app.diagrams.net/ or manually. Please mention if you make any assumptions. [15 Marks]

Diagram
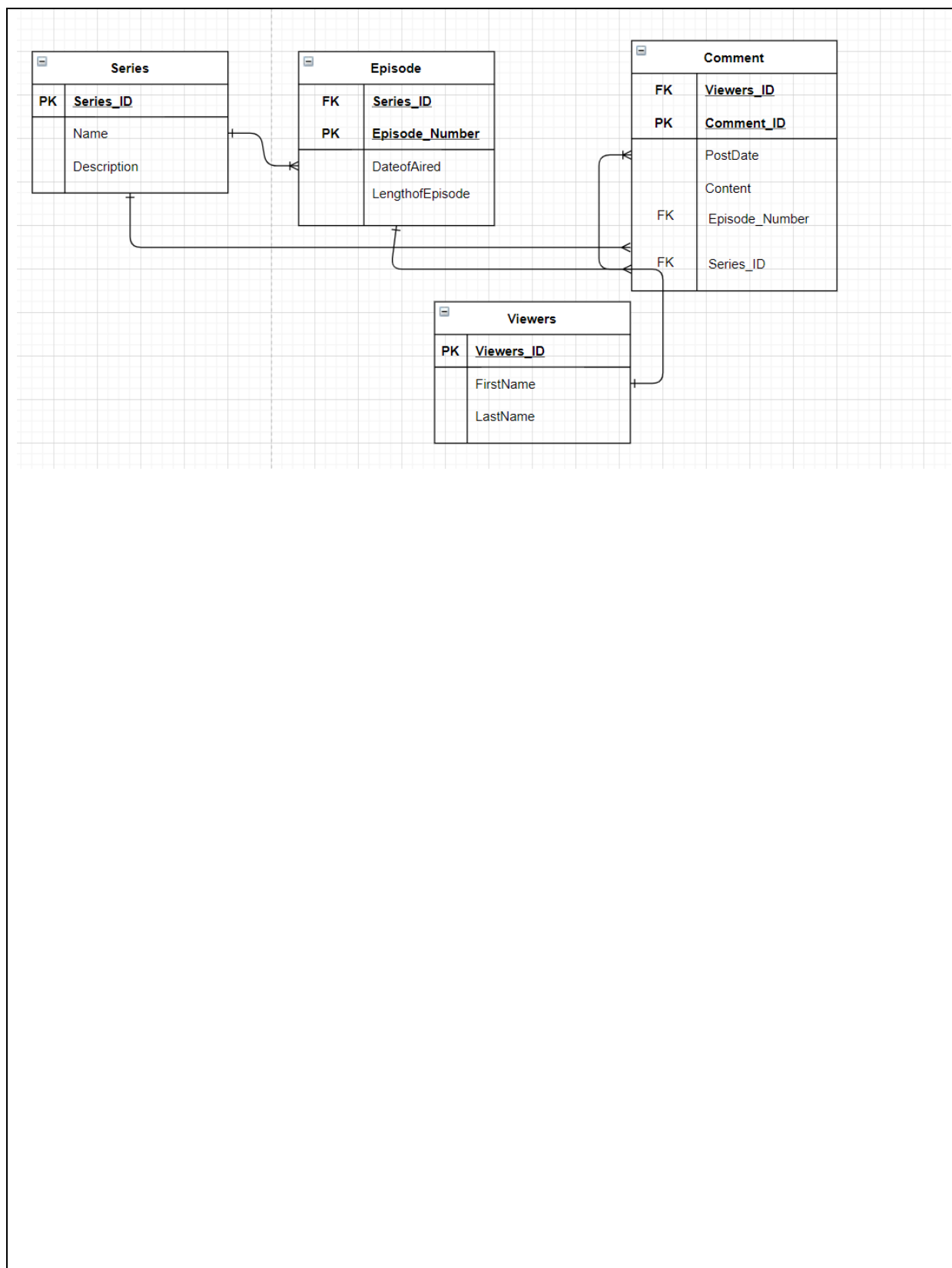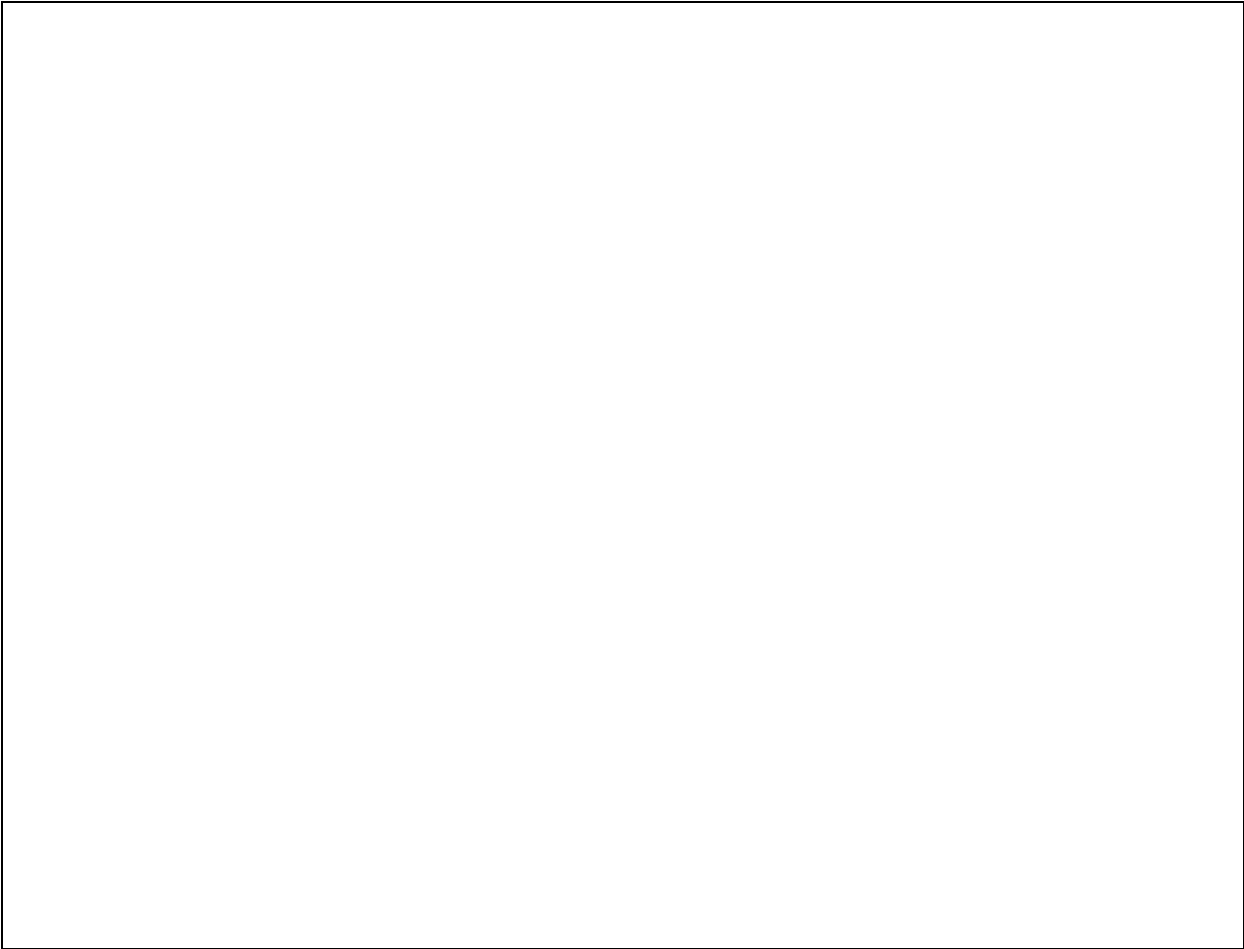
4. The following table stores details on employees and the projects they have worked on and the hours they spent on the project each week: [15 Marks]

WORKLOAD

EmpID	EmpName	ProjID	ProjName	HoursPerWeek
E01	Smith	P02	Database	10
E02	Smith	P01	Web Portal	5
E03	Robinson	P02	Database	20

    a. Explain why the above table is not in 2nd normal form. Then normalise it to 2nd normal form.

    b. Once the table is in 2nd normal form, test if the result is in 3rd normal form and explain why you think it **is** or **is not** in 3NF.

A : above table is not in 2nd normal form as there partial dependencies exist in it. EmpName only depends on EmpID, and ProjName is only dependent on ProjID.

EmpID	EmpName
E01	Smith
E02	Smith
E03	Robinson

ProjID	ProjName
P02	Database
P01	Web Portal
P02	Database

| EmpID	ProjID	HoursPerWeek
E01	P02	10
E02	P01	5
E03	P02	20

B: Above table in in 2nd normal form as there is no partial dependencies exists. Above tables are also in third normal form as there are no transitive dependencies exist because there are no columns exist that are dependent on primary key but not need primary key to uniquely identify them.
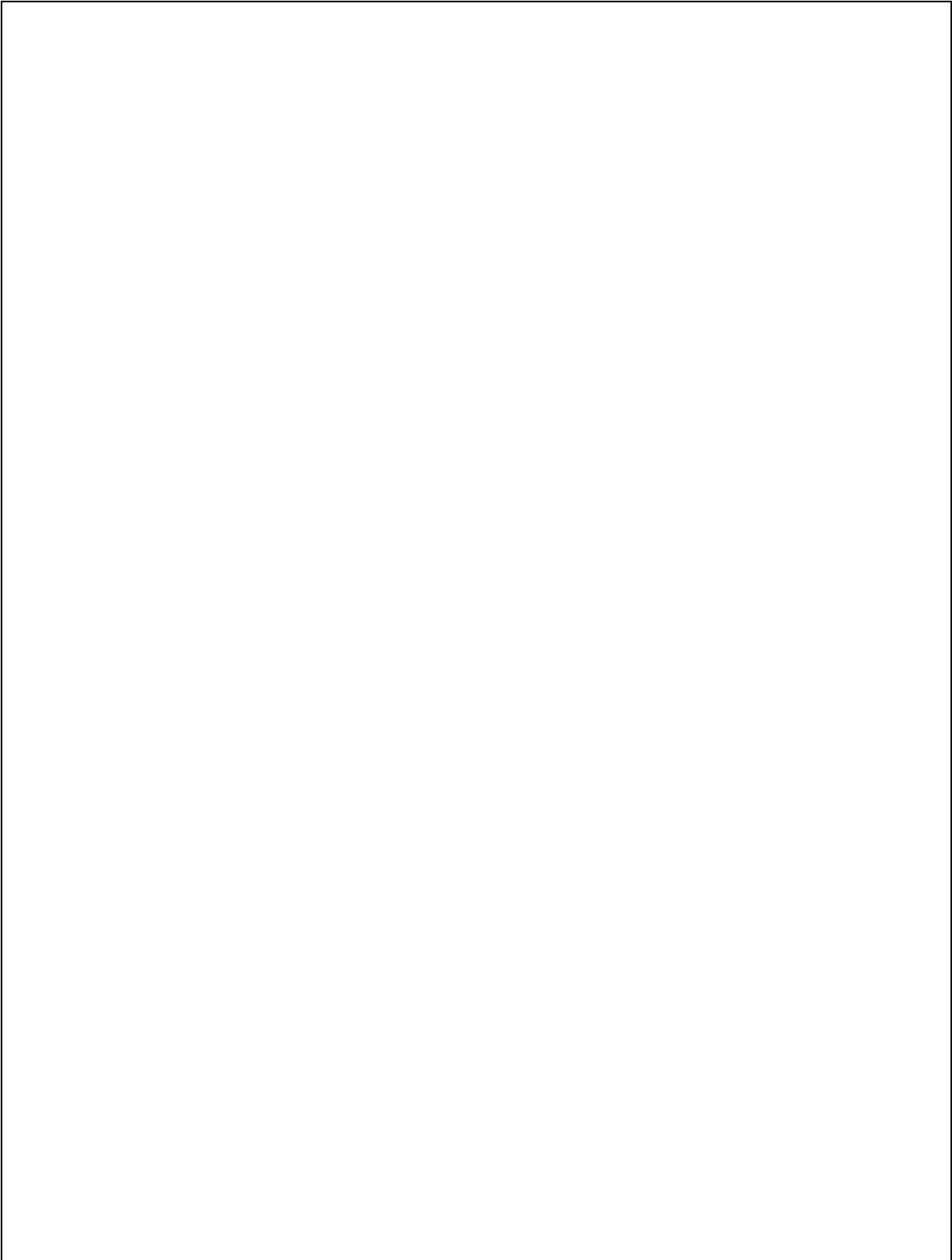
5. Consider the following ERD
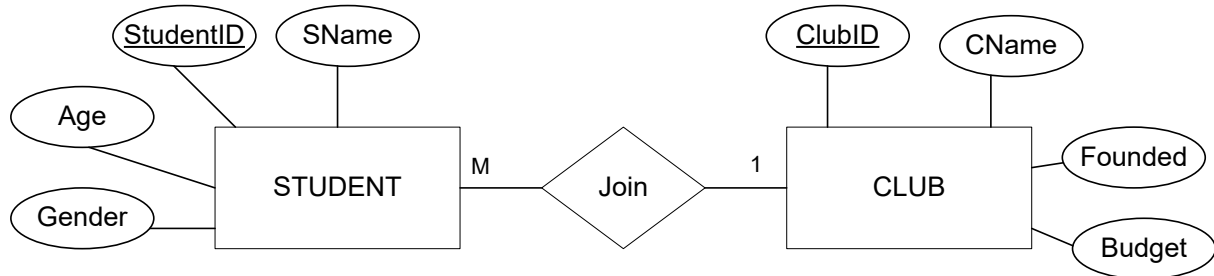
5.1 Write the SQL DDL statements to create the tables shown in the ERD above. Include primary and foreign keys. (The tables in 5.2 give you hints regarding the datatypes)

[15 Marks]

Use XAMPP or MySQL Workbench or Ed.

_____

_____

CREATE TABLE club (ClubID varchar(3) NOT NULL, CName varchar(25) NOT NULL, Founded int (4) NOT NULL, Budget int(5) NOT NULL, PRIMARY KEY (ClubID));

_____

CREATE TABLE student (SName varchar(25) NOT NULL, StudentID int(10) NOT NULL, Age int(3) NOT NULL, Gender ENUM('Male', 'Female'), ClubID varchar(3) NOT NULL, PRIMARY KEY (StudentID), FOREIGN KEY (ClubID) REFERENCES club(ClubID));

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

5.2 Consider the following data. (It is not necessary to run the code and upload the output) [10 Marks]

STUDENT

StudentID	SName	Gender	Age	ClubID
3234	Alfred Smith	Male	20	BSK
2244	McJohnson Robert	Male	22	
2389	Jessica Low	Female	20	JPA
4211	Roland Devingo	Male	24	
4383	Jane Usa Khan	Female	21	BKY
4450	Elaine Fong	Female	20	JPA

CLUB

ClubID	CName	Founded	Budget
BKY	Bakery Club	2010	2546
PDC	Photomedia and Design	2005	1345
JPA	Japanese Anime	2009	3453
BSK	Basketball	2011	6744

Answer the following:

a. Write an SQL statement to display student name and club name in ascending order of the student name. Include all clubs even if they do not have any members.

SELECT student.SName, club.CName from student right join club
_____

_____

_____

_____

_____

_____

_____

b. What is the value returned by this statement?

```
SELECT COUNT( CLUBID )
FROM STUDENT
```

4

5.3 If the database administrator in the University has turned off auto commit, consider the following:

Student 4211 has joined the Photomedia and Design club. When the system admin wanted to update the database, she wrote the following statement: [15 Marks]

```
UPDATE STUDENT
SET CLUBID = "PDC"
```

Answer the following:

a. How does the above **SQL statement** produce false information in the Student table?

It will set clubid for every student to PDC

b. How do you undo the effect of the SQL update statement?
ROLLBACK;

c. Re-write the statement to produce the correct output.
UPDATE STUDENT SET CLUBID = "PDC" WHERE StudentID = '4211'

d. Since **autocommit is off**, how do you make changes permanent in the Student table?

COMMIT;

5.4 Write the SQL statements that would DROP both tables. (**Note: Order is significant**).

[10 Marks]

DROP TABLE student;

DROP TABLE club;

---End of Paper---