COMP2714 Module 3 Case Study 1

Focus

The purpose of this case study is to help students develop practical experience in implementing a database system. You will learn how to identify and solve issues which commonly occur during database setup. Additionally, you will become familiar with SQL Data Definition Language (DDL) and its practical usage in a commercial environment.

Task

Using the correspondence on the pages below, complete the following tasks.

Section A – Error Identification / Resolution

In Peter's email to Elaine on the 2nd of May, he stated while they were exporting their finance database an error occurred causing some data corruption. Based on the information provided in this email and its attachments, complete the following two tasks.

A.1 – SQL File

Please read Correspondence 1, and attachments 1, 2 and 3 before attempting this question.

Identify three errors which occur in the SQL file. For each error:

1) Identify the line the error occurs on
2) State **briefly** the implications of this error (no more than 40 words)
3) Rewrite the line to fix the error

Note: Stating clear data corruption such as in the SQL for the "TimeLog" table will not be accepted.

Your answers must be provided in the format below:

Example Error	
Error	`hobby` varchar(1) NOT NULL,
Implications	An attribute of varchar with length 1 can only store a single letter.
Correction	`hobby` varchar(100) NOT NULL,

SEE NEXT PAGE FOR A.1 QUESTION BOX

SQL Error 1	
Error	
Implications	
Correction	

SQL Error 2	
Error	
Implications	
Correction	

SQL Error 3	
Error	
Implications	
Correction	

A.2 – CSV Files

Please read Correspondence 1, and attachments 1, 2 and 3 before attempting this question.

Based on the CSV files provided in Peter's first email, the INFS1200 tutors have identified three data entries which are either incorrect or cause runtime errors during the initial import. For each entry, you are required to write an SQL statement to remedy the error and also identify when the statement must be run. For each error you will be provided an SQL keyword which you must correctly use in your answer. An example has been provided below with the solution in red.

Note: You must use the SQL keyword provided for each given problem. Additionally, your SQL query must not cause any additional errors to occur when the CSV import is run.

Example Error	
Error	3039,Jennifer,Robinson,Developer
Explanation	This employee should have the role "Administration" not "Developer"
SQL Solution	UPDATE Employee SET role = "Administration" WHERE id = 3039
When to Run (Highlight Correct)	Before Employee CSV Import / After Employee CSV Import

SEE NEXT PAGE QUESTION BOXES

CSV Error 1	
Error	3035,Alex,MDonald,Administration
Explanation	“Mdonald” should be “McDonald”
SQL Solution	UPDATE
When to Run (Highlight Correct)	Before Employee CSV Import / After Employee CSV Import

CSV Error 2	
Error	2020 Marketing,Develop a marketing plan for 2020,43126.78,3009
Explanation	Tuple violates referential integrity constraint as no employee exists with ID 3009
SQL Solution	UPDATE
When to Run (Highlight Correct)	Before Project CSV Import / After Project CSV Import

CSV Error 3	
Error	Office Redesign,Redesign office to increase productivity,2000.50,3010
Explanation	Tuple violates referential integrity constraint as the Builder “Aaron Smith” with id 3010 is missing from the Employee table.
SQL Solution	INSERT
When to Run (Highlight Correct)	Before Project CSV Import / After Project CSV Import

Section B – SQL DDL (Table Setup)

Please read Correspondence 1, and attachments 1, 2 and 3 before attempting this section.

According to Peter’s email dated the 2nd of May, data corruption during a database export resulted in the SQL import queries for the TimeLog and ProjectExpenditureTimeline tables not being generated. Using the information provided in that email and its attachments, write the SQL queries to:

1) Generate the missing tables (without foreign keys)
2) Apply any needed foreign keys or missing constraints

Note: You must choose suitable data types when generating tables.

SEE NEXT PAGE QUESTION BOXES

YOUR SQL QUERIES TO GENERATE MISSING TABLES MUST FIT IN HERE

NOTE: NO FOREIGN KEY QUERIES SHOULD BE INCLUDED HERE

YOUR SQL QUERIES TO APPLY ANY NEEDED FOREIGN KEYS OR MISSING CONSTRAINTS MUST FIT IN HERE

Section C – SQL DDL (Table Modification)

Please read Correspondence 2 before attempting this section.

In Peter's email to Elaine dated the 5th of May, he presented several schema modifications for Dirt Road Driving payroll backend. For each schema modification, write SQL query(s) to apply the modification.

Schema Query 1	
Client Specification	Employee's middle names (if they have one) will also be stored along with their first name and last name.
SQL Query(s)	

Schema Query 2	
Client Specification	The employee's 'role' will be stored as a text datatype to allow for longer descriptions.
SQL Query(s)	

Schema Query 3	
Client Specification	The combination of first name and last name for each employee must be unique.
SQL Query(s)	

Schema Query 4 - A	
Client Specification	To increase productivity, each employee will be assigned to a "main project," which they are expected to spend the majority of their time on: - Add an attribute called "mainProject" to the employee's table which can store the main project name. Each employee already in the system should be assigned "Website Setup" as their main project.
SQL Query(s)	

Schema Query 4 - B	
Client Specification	To increase productivity, each employee will be assigned to a "main project," which they are expected to spend the majority of their time on: - This new attribute needs to be a foreign key referencing the Project.name attribute. Because each employee must have a valid main project this attribute should also not be null.
SQL Query(s)	

Section D – SQL DDL (Data Insertion/Deletion)

Please read Correspondence 2 before attempting this section.

In addition to schema modifications, Peter's second email dated 5th of May also requested for the database's data to be updated to reflect several developments within the company. For each of these requests, write SQL query(s) to make the needed changes in the needed tables. **Note**: at no point should your queries result in incorrect data being stored in the database, even if only temporarily.

Data Query 1	
Client Specification	On the 3/5/2020 Annie Fang worked for 4 hours on the "Website Setup" project and this time has been approved.
SQL Query(s)	

Data Query 2	
Client Specification	Gabriele Cirulli was fired.
SQL Query(s)	

QUESTION BOXES CONTINUE ON NEXT PAGE

Data Query 3	
Client Specification	Due to the corrupt export, the ProjectExpenditureTimeline table needs to be populated. (Note: Because we only provided you with a snapshot of our data, we need a query which populates this tables **based of the TimeLog table and not just specific insert queries**.)
SQL Query(s)	

Data Query 4	
Client Specification	The Board of Directors approved a new $8000.00 project called “New Zealand expansion analysis” which is designed to look into the economic and logistical feasibility of setting up a branch of Dirt Road Driving in New Zealand. To lead this project, we have hired a new manager, “Charlie Johnson” and assigned him ID 2930. Charlie will also be assigned this project as his “main project.”
SQL Query(s)	

Correspondence

Correspondence 1:

From: peter@dirtroaddriving.com.au
To: infs1200@uq.edu.au
Date: 2/5/2020 03:51 PM
Subject: RE: Student Support for Industry Project

Hi Elaine,

I hope you are having a nice weekend! Please thank your student teams for their responses regarding the source of our problematic payroll system commands, it was really appreciated!

After consulting with our IT department, they have informed me that they would like to redeploy the current payroll database to a MySQL system. They feel this will prove more stable in the long run and provide greater opportunities for future expansion. Initially, we just wanted to just roll out a small portion of the database similar to what I sent in my last email. However, when our database administrator did an SQL export of the old database system, it appears the import file and data which was generated was possibly corrupted casing the import to fail. Due to the high use of the system, we cannot schedule any downtime within the next month to rerun the export. As such, we were wondering if your student team would be able to take a look at these files then identify any errors and write/rewrite the appropriate SQL statements to fix these errors?

I have attached three documents to this email to help with this:

1) The partial relational schema for the payroll system
2) The SQL import file
3) A zip file containing the CSV import files

Please let me know if your student teams require any further information. Also, please pass on our sincere appreciations to them for their help!

Kind regards,
Peter Thompson
Director of Innovation | Dirt Road Driving

ATTACHMENTS ON NEXT PAGE

Attachment 1: Partial Payroll Relational Schema

Employee [id, firstName, lastName, role]
Project [name, description, funding, projectLeader]
TimeLog [employeeID, projectName, date, hoursWorked, approved]
ProjectExpenditureTimeline [projectName, totalApprovedHours]

Project.projectLeader references Employee.id
TimeLog.employeeID references Employee.id
TimeLog.projectName references Project.name
ProjectExpenditureTimeline.projectName references Project.name

Attachment 2: SQL Import File

Click here to open the SQL attachment.
Note: If the hyperlink does not work, please manually open Payroll_Schema_Export.sql

Attachment 3: ZIP File

Note: If any of the hyperlinks do not work, please open the file manually. The csv files themselves are not corrupt and will open in a text editor however the information contained within them may be corrupt.

1) Employee.csv

Click here to open the Employee.csv attachment.

2) Project.csv

Click here to open the Project.csv attachment.

3) TimeLog.csv

Click here to open the TimeLog.csv attachment.

SEE NEXT PAGE FOR FURTHER CORRESPONDENCE

Correspondence 2:

From: peter@dirtroaddriving.com.au
To: infs1200@uq.edu.au
Date: 5/5/2020 11:32 AM
Subject: RE: Student Support for Industry Project

Hi Elaine,
Thank you for your call yesterday to confirm the INFS1200 student teams would be able to assist in solving these issues with importing our payroll system.

In addition to the fixing the corrupted queries and data I sent in my last email, I was also wondering if your student teams would also be able to assist in writing queries to make some minor adjustments to the payroll backend? I have divided these adjustments into two categories, schema queries and data queries. For each situation highlighted by these dot points, could your students please provide the necessary SQL queries to apply any needed changes in the backend?

Schema Queries:

1) Employee's middle names (if they have one) will also be stored along with their first name and last name.
2) The employee's 'role' will be stored as a text datatype to allow for longer descriptions.
3) The combination of first name and last name for each employee must be unique.
4) To increase productivity, each employee will be assigned to a "main project," which they are expected to spend the majority of their time on. In order to implement this, we need two tasks completed:
    A. Add an attribute called "mainProject" to the employee's table which can store the main project name. Each employee already in the system should be assigned "Website Setup" as their main project.
    B. This new attribute needs to be a foreign key referencing the Project.name attribute. Because each employee must have a valid main project, this attribute should also not be null.

Data Queries:

1) On the 3/5/2020, Annie Fang worked for 4 hours on the "Website Setup" project and this time has been approved.
2) Gabriele Cirulli was fired.
3) Due to the corrupt export, the ProjectExpenditureTimeline table needs to be populated. (Note: Because we only provided you with a snapshot of our data, we need a query which populates this table **based off the TimeLog table and not just specific insert queries**.)
4) The Board of Directors approved a new $8000.00 project called "New Zealand expansion analysis" which is designed to look into the economic and logistical feasibility of setting up a branch of Dirt Road Driving in New Zealand. To lead this project, we have hired a new manager, "Charlie Johnson", and assigned him ID 2930. Charlie will also be assigned this project as his "main project."

Sorry for these additional requests, we have been exceptional busy these last few days. Again, we'd like to thank you and your team for your ongoing support!

Kind regards,
Peter Thompson
Director of Innovation | Dirt Road Driving