1. Module number	SET09702
2. Module title	Database Technology
3. Module leader	Brian Davison
4. Tutor with responsibility for this Assessment Student's first point of contact	Brian Davison
5. Assessment	SQL - Haulage
6. Weighting	60%
7. Size and/or time limits for assessment	Section 1 _Five_ (of 15) problems to be attempted. 1-5 attract a maximum of 5 points each 6-10 attract a maximum of 8 points each 11-15 attract a maximum of 12 points each Section 2 – up to 20 points Section 3 – up to 20 points
8. Deadline of submission	Hand in to Moodle by 1500 on Friday 23rd August 2019
9. Arrangements for submission	Answers to all sections in a single document. The section 1 answers should include SQL queries and the output generated. SQL code should be text formatted using a fixed-width font. Output may be in the form of a screen shot.
10. Assessment Regulations All assessments are subject to the University Regulations.	
11. The requirements for the assessment	
12. Special instructions	
13. Return of work	You will receive feedback via Moodle.
14. Assessment criteria	You will be assessed on the accuracy and the quality of your code.

Scenario

A haulage company manages its information using a relational database whose ER diagram is shown at the end of this document. The database provided contains information from a typical six-month period of operation. To fully understand the database structure, you will need to know some details about the way the company operates, and some of the terminology used:

- A single trip may last between 2 and 10 days
- Trips are grouped by their start dates. A trip starting in May for example is therefore deemed a May trip even if it ends in June.
- During a trip the driver visits a number of customer sites to pick up items of cargo which are then delivered to other customers
- Items may be picked up from a customer early in the trip, and other items delivered back to the same customer later on
- Routing is complex and is handled separately. This means you do not need to consider the relative location of customers on any given trip or the order in which they are visited
- Some items are fragile and must therefore have their condition checked and signed off by both pickup and delivery customers
- Some items are hazardous and may only be transported by drivers with appropriate qualifications
- *Manifest* is the term used for the items of cargo in transit
- Each manifest item is identified by a barcode which is used for checking and billing
- The *kerb weight* of the vehicle is its unladen weight
- *GVW* stands for *gross vehicle weight*. This is the maximum allowable laden weight for the vehicle

Section One - SELECT Statements (60 points)

The haulage database is available in MySQL format in Moodle as is the ER diagram in both png and drawio format.

Choose any **five** questions to answer. Questions 1 – 5 are worth 5 marks each, 6 – 10 are worth 8 marks, and 11 – 15 are worth 12 marks.

Each solution is a single SQL statement. Hard-coded values should be avoided except when the value is included in the question. The target answer has been provided for each question. The output from your SQL statement should reproduce it exactly including formatting and ordering.

1. **Trip 73440.** How many items were transported during trip 73440?

```
+-------+
| Items |
+-------+
|    19 |
+-------+
```

2. **Singleton.** Find the trip in which only a single item was transported.

```
+---------+
| trip_id |
+---------+
|   73738 |
+---------+
```

3. **Gavin Brandon.** Which company did Gavin Brandon deliver to between the 24th and 25th of April?

```
+--------------+
| company_name |
+--------------+
| Runnel Ltd.  |
+--------------+
```

4. **Long haul.** Which driver was responsible for the longest trip?

```
+------------+-----------+------+
| first_name | last_name | days |
+------------+-----------+------+
| Philip     | Slaight   |   11 |
+------------+-----------+------+
```

5. **Peak district.** Find the town where we do the most business – ie the one where the largest number of items are picked up and delivered.

```
+-----------+-------+
| town      | Items |
+-----------+-------+
| Gateshead |   328 |
+-----------+-------+
```

6. **Least used.** Find the five trucks that are least used during the six months covered by the data. Order by the number of trips they were used on.

```
+--------+--------------+--------------+-------+
| make   | model        | registration | trips |
+--------+--------------+--------------+-------+
| Scania | P94 4x2      | WY51OLV      |    17 |
| DAF    | FTGCF85.460V | PY12 ZYA     |    18 |
| Scania | R270 6x2     | BD08AOG      |    18 |
| DAF    | FTGCF85.460V | PY61 RNV     |    18 |
| DAF    | FTGCF85.460E | PY58 UHF     |    18 |
+--------+--------------+--------------+-------+
```

7. **Customer satisfaction**. Each month the company emails the five customers with the highest number of pickups (not manifest items) to check they are happy with the service. List the top five customers for June.

```
+-----------+----------------------+---------+
| reference | company_name         | Pickups |
+-----------+----------------------+---------+
|        99 | Temerarious & Co     |       9 |
|         3 | Trochiline Services  |       9 |
|       283 | Contemper Retail     |       9 |
|       264 | Byssiferous Industrial |     9 |
|         7 | Noumenalism Plc      |       8 |
+-----------+----------------------+---------+
```

8. **Gently does it.** Which drivers have never transported anything fragile? (NB Results are abbreviated)

```
+------------+-------------+
| first_name | last_name   |
+------------+-------------+
| Barry      | Thayre      |
| Henry      | Cobelli     |
| Rudyard    | Basillon    |
| Gareth     | Cruickshank |
| ….         | ….          |
| Edgar      | Strank      |
| Nadir      | Millbank    |
| Durant     | Dankersley  |
| Albert     | Phillimore  |
+------------+-------------+
```

9. **Travelling light**. Usually, the sequence of pickups and deliveries has to be carefully managed so as not to exceed the vehicle's capacity. However, if the total weight of manifest items for the whole trip does not exceed the limit, these checks can be skipped. How many trips can proceed without checking?

```
+----------+
| count(*) |
+----------+
|      341 |
+----------+
```

10. **Average number of trips**. What is the average duration of a trip in days for each month? Order the results by month.

```
+------------+-------+
| trip_month | trips |
+------------+-------+
| January    |   3.7 |
| February   |   3.3 |
| March      |   3.6 |
| April      |   3.5 |
| May        |   3.5 |
| June       |   3.5 |
| July       |   1.0 |
+------------+-------+
```

11. **Dangerous driving**. For all trips where hazardous good were transported, find the percentage of each category of item in the manifest. Sort in descending order of the percentage of hazardous items. (NB results are abbreviated)

```
+---------+------+------+------+
| trip_id | A    | B    | C    |
+---------+------+------+------+
|   73832 | 44%  | 0%   | 56%  |
|   73404 | 60%  | 0%   | 40%  |
|   73773 | 63%  | 0%   | 38%  |
|   73551 | 64%  | 0%   | 36%  |
|   73013 | 67%  | 0%   | 33%  |
|   ...   | ...  | ...  | ...  |
|   74059 | 96%  | 0%   | 4%   |
|   73049 | 96%  | 0%   | 4%   |
+---------+------+------+------+
```

12. **Unused trucks.** List the registration numbers of the trucks that were not in use between 1 and 5 April inclusive.

```
+--------------+
| registration |
+--------------+
| SDU 567M     |
| PY06 BYP     |
| PY61 RNU     |
| BD60BVF      |
| BD08AOF      |
+--------------+
```

13. **Bonus**. If a driver works more than 22 days in any one month, they are paid at a higher rate for the extra days. List the drivers who qualify for bonus payments for each month in the data and include the number of extra days worked. Drivers who are not eligible for a bonus should not be shown. Order by month and number of days descending.

```
+----------+----------------------+------+-----------+
| Month    | Name                 | Days | Bonus days |
+----------+----------------------+------+-----------+
| January  | Igor Woodruffe       | 24   | 2         |
| January  | Henry Cobelli        | 23   | 1         |
| February | Tristan Crumbie      | 23   | 1         |
| March    | Oscar Nutten         | 27   | 5         |
| March    | Daniel Miliffe       | 26   | 4         |
| March    | Eden Blackbrough     | 23   | 1         |
| March    | Henry Cobelli        | 23   | 1         |
| March    | Barry Thayre         | 23   | 1         |
| April    | Durant Kewzick       | 24   | 2         |
| April    | Leonardo Charlet     | 24   | 2         |
| May      | Solomon Alessandrucci | 24  | 2         |
| June     | Lee Rookledge        | 28   | 6         |
| June     | Durant Kewzick       | 26   | 4         |
+----------+----------------------+------+-----------+
```

14. **Peak week.** Find the busiest week based on the number of departures and returns. Show the start date assuming that a week starts on Monday and the number of departures and returns.

```
+------------+-----------+
| Week       | Movements |
+------------+-----------+
| 2012-04-02 |       105 |
+------------+-----------+
```

15. **Capacity factor**. 100% capacity is when every truck is in use every day. If some trucks are idle, the capacity factor is less than 100%. What is the total capacity factor for the company for the time period covered by the data?

```
+----------+
| capacity |
+----------+
| 43%      |
+----------+
```

Section 2 – Database Design (20 points)

You are required to extend the haulage system to include customer account managers.

The following facilities are required:

- Only drivers are currently recorded – a new staff role of *customer account manager* (CAM) is needed.
- Every customer must be assigned a CAM, and one CAM can serve several customers
- CAMs may change over time – it must be possible to find out who was the CAM at the time of any trip, even if that person has subsequently left the company
- It must be possible to identify a customer's current CAM

A correct solution for the problems above is worth up to 70%; to get marks over 70% assuming that everything else is correct you should also add the facility for CAMs to handle customers' queries:

- The details of any initial query must be recorded
- The CAM's initial response must be recorded (potentially at a later date than the original query)
- Both the customer and the CAM can add follow-up entries to the query history
- It must be possible to display the entire history of a query including all follow-up items in chronological order
- Queries must be marked as closed when no further action is required
- It must be possible to calculate a performance rating for CAMs based on the promptness of their response to a customer and the number of open queries for their customers

You are required to submit:

1. An update of the ER diagram to include tables/attributes to support these additional features
2. SQL statements to:
   a. alter the database to include any additional tables and attributes required
   b. insert sample data for at least three customers and their CAMs
   c. query the database to retrieve the information needed to satisfy the above requirements

Section 3 - YouHaul.com (20 points)

A global haulage management system is being considered. When it is fully operational we will have thousands of haulage contractors using our online service to manage their scheduling. The essential data stored is to be identical to the solution described here, but many independent contractors can use system simultaneously. The system should be able cope with up to 1 million contractors with working patterns similar to the example that you have. Potential customers will be able to post a pickup and/or delivery request to a regional queue where it can be serviced by any of the local contractors. A 'region' is a geographical area with roughly a 200 km radius.

Write a short report of around 1000 words (± 10%) outlining options for implementing this database. You should compare the suitability of a MySQL solution against a non-relational alternative.

Your report should include:

- An introduction explaining the problem.
- The advantages and disadvantages of a MySQL based solution.
- The advantages and disadvantages of a NoSQL solution
- The selection of a solution and a defence of your selection.
- A high-level description of your proposed solution – ie how it might be implemented
- Conclusions