Student Name: Joe Kirkup Student Number: 29026253

INDIVIDUAL ASSIGNMENT
(Weight: 40% OF THE OVERALL MODULE GRADE)

55-500998 DATABASE SYSTEMS FOR SOFTWARE APPLICATIONS
Module Leader: Kostas Domdouzis
Academic Year 2020/21

Submission Type: ELECTRONIC (through BLACKBOARD)

•**TASK 1 – Normalize the following table using the three Rules of Normalization and showing each of the three stages of their implementation. (Weight: 12.5% of the Assignment)**

Sheffield High School 'The Explorer'

Phone (0114) 225-6783
153 Ashby Rd
Sheffield
S5 2TU

Students Record

(Unique) StudentName	StudentYear	Exam	Date	Grade	Teacher	TeacherID
John	1	Physics	17/09/2016	70%	Dr Jones	12
George	1	Art	11/05/2016	65%	Dr Jones	12
Helen	2	Music	14/02/2017	55%	Dr Johnson	15
Sarah	3	Maths	15/07/2017	54%	Dr Vans	16
Paul	3	French	17/09/2016	85%	Dr Pascal	17

0NF:
StudentName, StudentYear, Exam, Date, Grade, Teacher, TeacherID

Student Name: Joe Kirkup Student Number: 29026253

1NF:
StudentName*, Exam, Date, Grade, TeacherID*
StudentName, StudentYear
TeacherID, Teacher

2NF:
StudentName*, ExamID*, Grade, TeacherID*
ExamID, Exam, Date
StudentName, StudentYear
TeacherID, Teacher

3NF:
StudentName*, AttemptID*
StudentName, StudentYear
AttemptID, ExamID, Date, Grade
ExamID, Exam, TeacherID
TeacherID, Teacher

•TASK 2 – Normalize the following table using the three Rules of Normalization and showing each of the three stages of their implementation. (Weight: 12.5% of the Assignment)

Employer Code	Student No	Employer Name	Employer Specialty	Student Name	Appointment Date/Time	Session Code	Time Alloc.
123	1414	Stevens	Artificial Intelligence	Johnson	12-10-2017 12:00pm	D	50
123	1513	Stevens	Artificial Intelligence	Patel	14-10-2017 13:00pm	B	60
123	1567	Stevens	Artificial Intelligence	Jamal	15-10-2017 13:30pm	F	20
123	2010	Stevens	Artificial Intelligence	Hope	15-10-2017 16:00pm	A	60
234	1414	Vidal	Graphics	Johnson	17-10-2017 08:00am	D	60

234	1567	Vidal	Graphics	Jamal	15-10-2017 16:00pm	F	10
234	1785	Vidal	Graphics	Michaels	17-10-2017 08:30am	A	10
365	1863	Matthews	Web Design	Wong	20-01-2017 15:00pm	A	30
365	1975	Matthews	Web Design	Holmes	21-02-2017 15:00pm	C	30
456	1414	Parsons	Algorithms	Jones	22-02-2017 16:00pm	D	10
456	1513	Parsons	Algorithms	Patel	06-06-2017 13:00pm	B	60
456	1634	Parsons	Algorithms	Peterson	15-10-2017 16:00pm	C	60
456	2011	Parsons	Algorithms	Siddiqi	17-10-2017 09:00am	A	30
456	2160	Parsons	Algorithms	King	20-12-2017 10:00am	A	30

0NF:
EmployerCode, StudentNo, EmployerName, EmployerSpecialty, StudentName, AppointmentDT, SessionCode, TimeAlloc

1NF:
<u>EmployerCode*</u>, StudentNo*, AppointmentDate, AppointmentTime, SessionCode, TimeAlloc
<u>EmployerCode</u>, EmployerName, EmployerSpecialty
<u>StudentNo</u>, StudentName

2NF:
<u>EmployerCode*, AppointmentID*</u>, StudentNo*
<u>AppointmentID</u>, AppointmentDate, AppointmentTime, SessionCode, TimeAlloc
<u>EmployerCode</u>, EmployerName, EmployerSpecialty
<u>StudentNo</u>, StudentName

3NF (no change):
<u>EmployerCode*, AppointmentID*</u>, StudentNo*
<u>AppointmentID</u>, AppointmentDate, AppointmentTime, SessionCode, TimeAlloc
<u>EmployerCode</u>, EmployerName, EmployerSpecialty

Student Name: Joe Kirkup   Student Number: 29026253

StudentNo, StudentName

•TASK 3 - Produce an Entity-Relationship Diagram (ERD) for the entire system presented in the following scenario.
The ERD should be usable by the Royal Astronomical Society in order to understand better how MARS missions operate and to extract conclusions on how to improve these missions further.

(Weight: 35% of the Assignment)

Mars is the fourth planet from the Sun and the second-smallest planet in the Solar System, after Mercury.

Several rovers have been dispatched to Mars:

- *Mars 2, Prop-M* rover, 1971, [redacted] landing failed taking Prop-M with it. The *Mars 2* and *3* spacecraft from the USSR had identical 4.5 kg *Prop-M* rovers. They were to move on skis while connected to the landers with cables.
- *Mars 3, Prop-M* rover, 1971, lost [redacted] 20 seconds after landing.
- *Sojourner* rover, Mars Pathfinder, landed successfully [redacted] July 4, 1997. Communications [redacted] lost [redacted] September 27, 1997.
- *Beagle 2, Planetary Under-surface Tool*, lost with Beagle 2 on deployment [redacted] in 2003. A compressed spring mechanism was designed to allow movement across the surface at a rate of 1 cm per 5 seconds and to burrow into the ground and collect a subsurface sample in a cavity in its tip.
- *Spirit* (MER-A), Mars Exploration Rover, launched on June 10, 2003 at 13:58:47 EDT and landed successfully on January 4, 2004. Nearly 6 years after the original mission limit, Spirit had covered a total distance of 7.73 km (4.80 mi) but its wheels became trapped in sand. [redacted] The last communication received from the rover was on March 22, 2010, [redacted]
- *Opportunity* (MER-B), Mars Exploration Rover, launched on July 7, 2003 at 23:18:15 EDT [redacted] landed successfully [redacted] January 25, 2004. [redacted] total distance of 40.25 km (25.01 mi). [redacted] still operational [redacted] as of September 10, 2020.
- *Curiosity*, Mars Science Laboratory (MSL), [redacted] November 26, 2011 at 10:02 EST and landed in the Aeolis Palus plain near Aeolis Mons (informally "Mount Sharp") in Gale Crater [redacted] August 6, 2012, 05:31 UTC. [redacted] still operational as of September 10, 2020.

Curiosity [redacted] launched [redacted] November 26, 2011, at 15:02 UTC [redacted] landed in Gale Crater on

Student Name: Joe Kirkup Student Number: 29026253

Mars on August 6, 2012, 05:17 UTC. [redacted] mass of 899 kg (1,982 lb) [redacted] The rover is 2.9 m (9.5 ft) long by 2.7 m (8.9 ft) wide by 2.2 m (7.2 ft) in height. *Curiosity* is powered by a radioisotope thermoelectric generator (RTG), like the successful Viking 1 and Viking 2 Mars landers in 1976. Radioisotope power systems (RPSs) are generators that produce electricity from the decay of radioactive isotopes, such as plutonium-238, which is a non-fissile isotope of plutonium. Heat given off by the decay of this isotope is converted into electric voltage by thermocouples, providing constant power during all seasons and through the day and night.

The temperatures at the landing site can vary from −127 to 40 °C (−197 to 104 °F); therefore, the thermal system of Curiosity will warm the rover for most of the Martian year. The thermal system will do so in several ways: passively, through the dissipation to internal components; by electrical heaters strategically placed on key components; and by using the rover heat rejection system (HRS). It uses fluid pumped through 60 m (200 ft) of tubing in the rover body so that sensitive components are kept at optimal temperatures.

Curiosity uses two identical on-board rover computers, called Rover Computer Element (RCE) which contain radiation hardened memory to tolerate the extreme radiation from space and to safeguard against power-off cycles. The computers run the VxWorks real-time operating system (RTOS). Each computer's memory includes 256 kB of EEPROM, 256 MB of DRAM, and 2 GB of flash memory.

In order to communicate with Earth, Curiosity is equipped with several telecommunication means – an X-band Transmitter/Receiver that can communicate directly with Earth, and a UHF Electra-Lite software-defined radio for communicating with Mars orbiters.

Communication with orbiters is expected to be the main path for data return to Earth, since the orbiters have both more power and larger antennas than *Curiosity*, thus allowing for faster transmission speeds. The rover has two UHF radios, the signals of which the 2001 Mars Odyssey orbiter is capable of relaying back to Earth. An average of 14 minutes, 6 seconds will be required for signals to travel between Earth and Mars. *Curiosity* can communicate with Earth directly at speeds up-to 32 kbit/s, but the bulk of the data transfer should be relayed through the Mars Reconnaissance Orbiter and Odyssey orbiter. Communication from and to *Curiosity* relies on internationally agreed space data communications protocols as defined by the Consultative Committee for Space Data Systems.

Curiosity is equipped with six 50 cm (20 in) diameter wheels in a rocker-bogie suspension. The suspension system also serves as landing gear for the vehicle, unlike its smaller predecessors. Each wheel has cleats and is independently actuated and geared, providing for climbing in soft sand and scrambling over rocks. Each front and rear wheel can be independently steered, allowing the vehicle to turn in place as well as execute arcing turns. Each wheel has a pattern that helps it maintain traction but also leaves patterned tracks in the sandy surface of Mars.

Curiosity has 17 cameras: HazCams (8), NavCams (4), MastCams (2), MAHLI (1), MARDI (1), and ChemCam (1). Each MastCam includes the Medium Angle Camera (MAC) which has a 34 mm (1.3 in) focal length, a 15° field of view, and can yield 22 cm/pixel (8.7 in/pixel) scale at 1 km (0.62 mi). The other camera in the MastCam is the Narrow Angle Camera (NAC), which has a 100 mm (3.9 in) focal length, a 5.1° field of view, and can yield 7.4 cm/pixel (2.9 in/pixel) scale at 1 km (0.62 mi). A pair of MastCams were developed which include zoom lenses, but these were not included in the rover because of the time required to test the new hardware and the looming November 2011 launch date. Each MastCam has eight gigabytes of flash memory, which is capable of storing over 5,500 raw images, and can apply real time lossless data compression.

ChemCam is actually two different instruments combined as one: a laser-induced breakdown spectroscopy (LIBS) and a Remote Micro Imager (RMI) telescope. The purpose of the LIBS instrument is to provide elemental compositions of rock and soil, while the RMI will give ChemCam scientists high-resolution images of the sampling areas of the rocks and soil that LIBS targets. ChemCam has the ability to record up to 6,144 different wavelengths of ultraviolet, visible, and infrared light.

MAHLI is a camera on the rover's robotic arm and acquires microscopic images of rock and soil. MAHLI can take true-colour images at 1600×1200 pixels with a resolution as high as 14.5 micro-meters per pixel. MAHLI has an 18.3 to 21.3 mm (0.72 to 0.84 in) focal length and a 33.8–38.5° field of view. MAHLI has both white and ultraviolet LED illumination for imaging in darkness or fluorescence imaging. MAHLI also has mechanical focusing in a range from infinite to millimetre distances.

Curiosity stores the images generated by MastCams, the ChemCam and MAHLI in three different databases. Field-Programmable Gate Arrays (FPGAs) are used by Curiosity in order to categorize these images based on their significance and store them in the respective database. The significance of an image is defined by image processing algorithms that identify specific features of each image (e.g. density, contrast) and produce a decision about them. For example, if the image processing algorithms identify the significance of an image to be above 70%, then they store the image in Database A. If the significance of the image is between 45% to 69%, it is stored in Database B while for significance below 45%, then the image is stored in Database C. It is possible that an image is elevated to Database A from Database B or even from Database C depending on the conditions of the mission of Curiosity on the surface of Mars or inversely, an image can be downgraded to Database C from Database A.

The most significant pictures are emitted by FPGAs through the UHF Electra-Lite software-defined radio to the Mars orbiters and then back to earth. All the images though are used for obstacle track identification on the Mars surface and coordinates definition for the rover. These coordinates are sent directly to Earth through Curiosity's X band transmitter so that NASA knows exactly the route followed by the rover.

Database on missions:
0NF:
<u>RoverName</u>, LaunchDate, LandingSuccess, LandingDate, Failure, FailureDate, Transport, TotalDistance, Weight

3NF (no change past 1NF):
<u>RoverName</u>, LaunchID*, Failure, FailureDate, Transport, TotalDistance, Weight
<u>LaunchID</u>, LaunchDate, LandingSuccess, LandingDate

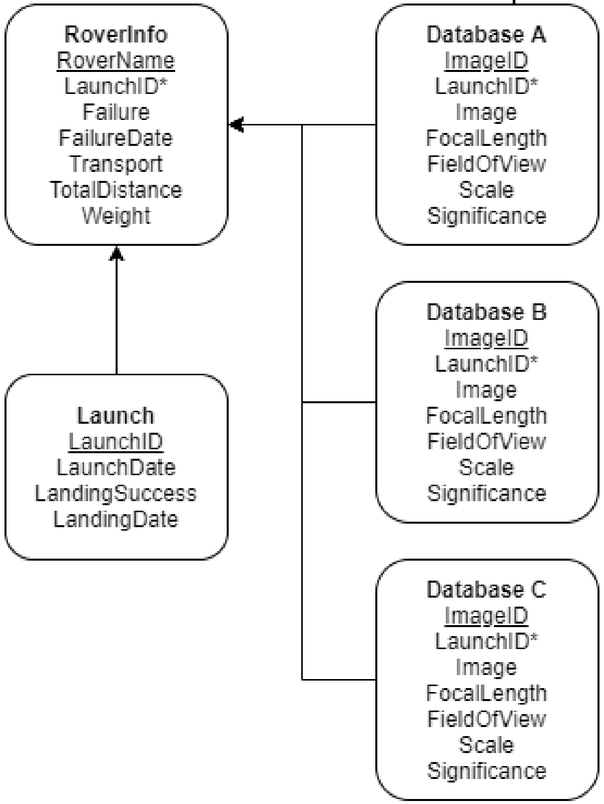
Student Name: Joe Kirkup Student Number: 29026253

Databases on images from Curiosity:
Database A
<u>ImageID</u>, Image, FocalLength, FieldOfView, Scale, Significance (S>=70%)

Database B
<u>ImageID</u>, Image, FocalLength, FieldOfView, Scale, Significance (45<=S>70%)

Database B
<u>ImageID</u>, Image, FocalLength, FieldOfView, Scale, Significance (S<45%)

Shown below is the ERD I have developed based on this extraction of data fields and normalisation.



Student Name: Joe Kirkup Student Number: 29026253

•TASK 4

Identify characteristics of your ER diagram that make the database efficient.

When designing this ERD I first ensured that all significant data outlined in the scenario would be represented in the database, even if in a simplified form as this opens many analytical doors. Next, I laid out the entities and their fields, and normalised the planned database to 3NF to ensure integrity of data and a great balance between storage and speed of operation, thinking of the infamous time-space trade-off.

I believe that the simplistic design I have presented represents all significant information brought forth as easily accessible and queryable data in a well-formed database system. While not all explicit details are included, I believe this is a strength as it allows for the database to serve its true purpose – displaying patterns in *data*, not lengthy write-ups about technical details.

[*Max Number of Words: 400*]

[Weight: 15%]

•TASK 5

For each of the following questions, provide one PL/SQL procedure, one function or one cursor that will check the following:

5A) How many images are stored in Database B based on their significance?

```
SELECT COUNT(ImageID)
FROM DATABASE_B
GROUP BY ImageID
HAVING Significance >= 45 AND Significance < 70
```

[Weight: 8%]

5B) How many images are elevated from Database C to Database A?

```
SELECT COUNT(ImageID)
FROM DATABASE_C
```

```
GROUP BY ImageID
HAVING Significance >= 70
```

[Weight: 9%]

5C) How many coordinates collected by the images of Databases B & C are sent directly to Earth through Curiosity's X band transmitter?

```
SELECT COUNT(ImageID)
FROM DATABASE_A A
INNER JOIN DATABASE_B B
        ON A.Significance = B.Significance
INNER JOIN DATABASE_C C
        ON B.Significance = C.Significance
GROUP BY ImageID
HAVING Significance >= 70
```

[Weight: 8%]

LEARNING OUTCOMES

LO Ref	Learning Outcome
1	Design databases for non-complex scenarios using appropriate notations and theories, including underlying set notations.
2	Implement, manipulate and query these databases using standard approaches.
3	Identify and discuss issues relating to databases, such as query optimisation, data integrity, security, reliability, data protection and curation.

MARKING OF THE INDIVIDUAL ASSIGNMENT

TASKS 1 & 2 Marking

Tasks 1 & 2 are Normalization exercises that show the level of comprehension by the students of the Normalization Rules. The distribution of the marks for each of Tasks 1 and 2 is the following:

The marker should consider even small things that the student does when he/she normalizes the table and should also consider the overall logic.

Student Name: Joe Kirkup Student Number: 29026253

	0-19%	20-33%	34-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Tasks 1 & 2 (25% of the Assignment Grade) **Normalization Marking Grid**	Very little or no understanding at all.	Many important attributes missing. Many incorrect relationships as expressed by using keys.	Some important attributes appearing. Some correct dependencies both in terms of logic (that corresponds to the requirements of the given problem). Some correct use of keys.	Most important attributes present. Dependencies mostly correct and corresponding at a substantial degree to the logic of the problem. Most PKs and FKs indicated.	Some minor discrepancies in keys and dependencies.	All attributes and relationships shown correctly and supported by all correct PKs and FKs.	A 'perfect' Normalization solution with exactly the correct attributes; primary and foreign keys will be indicated in an unambiguous, easily read way.	All the requirements set in the range (80-89) satisfied plus provision of alternative solutions with appropriate explanation for the provision of these solutions. This range of grades shows that the student examined the tasks in a more-in-depth manner and provided more work than what the assignment was asking for.

TASK 3 Marking

	0-19%	20-39%	40-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Task 3 (35% of the Assignment Grade)	Very little or no understanding at all.	Many important entities missing or	Some important entities appearing. Some correct relationships	Most important entities present. Sensible and	Even more important entities present. Even more sensible and	All entities and relationships shown correctly and	A 'perfect' ERD with exactly the correct attributes;	All the requirements set in the range

Entity-Relationship Diagram Marking Grid		poorly described in terms of attributes. Many incorrect relationships. No annotations in relationships at all.	both in terms of logic (that corresponds to the requirements of the given scenario) and Optionality & Cardinality. Some correct use of keys. Some annotations.	adequate attributes shown. Relationships (Cardinality & Optionality) mostly correct and corresponding at a substantial degree to the logic of the scenario. Most PKs and FKs indicated.	adequate attributes shown. Even more correct relationships shown and corresponding even more to the logic of the scenario. Even more correct PKs and FKs. Possibly, some minor discrepancies between keys and relationships.	supported by all correct PKs and FKs. All annotations present in relationships. All the aspects and logic of the scenario presented in the ER diagram.	primary and foreign keys will be indicated in an unambiguous, easily read way; relationships (Cardinality and Optionality) correct. Perfect depiction of all the aspects and logic of the scenario. Thorough and clear annotations in every relationship.	(80-89) satisfied plus provision of alternative solutions with appropriate explanation for the provision of these solutions. This range of grades shows that the student examined the task in a more-in-depth manner and provided more work than what the assignment was asking for.

TASK 4 Marking

	0-19%	**20-39%**	**40-49%**	**50-59%**	**60-69%**	**70-79%**	**80-89%**	**90-100%**

Student Name: Joe Kirkup Student Number: 29026253

Task 4 (15% of the Assignment Grade) **Identify features from your developed ER Model that are very important for the efficiency of your database.**	Missing or demonstrates little or no understanding	Some attempt, but badly flawed	A basic identification of features demonstrating understanding, or a more comprehensive solution which may be notionally or semantically flawed.	Demonstrates a clear understanding. May be some errors.	A set of features demonstrating very clear understanding . May be few errors.	A correct solution which closely models the problem domain.	A 'perfect' solution with all the necessary features..	All the requirements set in the range (80-89) satisfied plus provision of alternative solutions with appropriate explanation for the provision of these solutions. This range of grades shows that the student examined the task in a more-in-depth manner and provided more work than what the assignment was asking for.

TASK 5 Marking

Student Name: Joe Kirkup Student Number: 29026253

Task 5 (25% of the Assignment Grade) **Distribution of Grades between Sub-Tasks:** **Task 5A (8%)** **Task 5B (9%)** **Task 5C (8%)** **[*The comments exactly on the right of this box correspond to each sub-task of Task 3*]**	Very little or no understanding at all.	A very basic SQL Script with many errors in the syntax of its statements.	A basic SQL Script in which the syntax is semantically flawed.	A more comprehensive SQL Script which may be semantically flawed.	Demonstrates clear understanding. May be some errors in the syntax.	A comprehensive solution, which closely models the problem domain.	A complete solution which models the problem domain.	A complete solution which models exactly the problem domain plus alternative solutions.

INSTRUCTIONS

For marks equal or greater to 90% on each task, you will need to provide a perfect solution for the task plus any other alternative solution for it.

For the realisation of Task 3, you can use any ER Drawing software program. I personally use the Gliffy editor (https://www.gliffy.com/) which offers a free trial period during which you can complete the assignment. There is also the Flowchart Maker in this address: https://www.draw.io/. You are free of course to use any software you would like to. You can even draw the ERD on paper and scan it (as the assignment requires an electronic submission). In this case though, you need to make sure that your diagram is really clear.