MODULE NAME:	MODULE CODE:
DATABASE (INTRODUCTION)	DATA6211
DATABASE (INTRODUCTION)	DATA6211d
DATABASE (INTRODUCTION)	DATA6211p

ASSESSMENT TYPE: ASSIGNMENT 1 (PAPER ONLY)

TOTAL MARK ALLOCATION: 100 MARKS

TOTAL HOURS: 10 HOURS

By submitting this assignment, you acknowledge that you have read and understood all the rules as per the terms in the registration contract, in particular the assignment and assessment rules in The IIE Assessment Strategy and Policy (IIE009), the intellectual integrity and plagiarism rules in the Intellectual Integrity Policy (IIE023), as well as any rules and regulations published in the student portal.

INSTRUCTIONS:

1. ***No material may be copied from original sources, even if referenced correctly, unless it is a direct quote indicated with quotation marks. No more than 10% of the assignment may consist of direct quotes.***
2. ***Any assignment with a similarity index of more than 25% will be scrutinised for plagiarism. Please ensure you attach an originality report to your assignment if required.***
3. ***Make a copy of your assignment before handing it in.***
4. *Assignments must be typed unless otherwise specified.*
5. *All work must be adequately and correctly referenced.*
6. *Begin each section on a new page.*
7. *Follow all instructions on the assignment cover sheet.*
8. *This is an individual assignment.*

Referencing Rubric

Providing evidence based on valid and referenced academic sources is a fundamental educational principle and the cornerstone of high-quality academic work. Hence, The IIE considers it essential to develop the referencing skills of our students in our commitment to achieve high academic standards. Part of achieving these high standards is referencing in a way that is consistent, technically correct and congruent. This is not plagiarism, which is handled differently.

Poor quality formatting in your referencing will result in a penalty **of a maximum of ten percent being deducted from the percentage awarded**, according to the following guidelines. Please note, however, that **evidence of plagiarism in the form of copied or uncited work (not referenced), absent reference lists, or exceptionally poor referencing, may result in action being taken in accordance with The IIE's Intellectual Integrity Policy (0023)**.

Markers are required to provide feedback to students by indicating **(circling/underlining) the information that best describes the student's work.**

Minor technical referencing errors: 5% deduction from the overall percentage – the student's work contains **five or more errors** listed in the minor errors column in the table below.

Major technical referencing errors: 10% deduction from the overall percentage – the student's work contains **five or more errors** listed in the major errors column in the table below**.**

If both minor and major errors are indicated, then 10% only (and not 5% or 15%) is deducted from the overall percentage. The examples provided below are not exhaustive but are provided to illustrate the error.

Required: **Technically correct referencing style**	**Minor errors in technical correctness of referencing style** **Deduct 5% from percentage awarded**	**Major errors in technical correctness of referencing style** **Deduct 10% from percentage awarded**
Consistency • The same referencing format has been used for all in-text references and in the bibliography/reference list.	Minor inconsistencies. • The referencing style is generally consistent, but there are one or two changes in the format of in-text referencing and/or in the bibliography. • For example, page numbers for direct quotes (in-text) have been provided for one source, but not in another instance. Two book chapters (bibliography) have been referenced in the bibliography in two different formats.	Major inconsistencies. • Poor and inconsistent referencing style used in-text and/or in the bibliography/ reference list. • Multiple formats for the same type of referencing have been used. • For example, the format for direct quotes (in-text) and/or book chapters (bibliography/ reference list) is different across multiple instances.
Technical correctness • Referencing format is technically correct throughout the submission. • Position of the reference: a reference is directly associated with every concept or idea. • For example, quotation marks, page numbers, years, etc. are applied correctly, sources in the bibliography/reference list are correctly presented.	**Generally, technically correct with some minor errors.** • The correct referencing format has been consistently used, but there are one or two errors. • Concepts and ideas are typically referenced, but a reference is missing from one small section of the work. • Position of the references: references are only given at the beginning or end of every paragraph. • For example, the student has incorrectly presented direct quotes (in-text) and/or book chapters (bibliography/reference list).	**Technically incorrect.** • The referencing format is incorrect. • Concepts and ideas are typically referenced, but a reference is missing from small sections of the work. • Position of the references: references are only given at the beginning or end of large sections of work. • For example, incorrect author information is provided, no year of publication is provided, quotation marks and/or page numbers for direct quotes missing, page numbers are provided for paraphrased material, the incorrect punctuation is used (in-text); the bibliography/reference list is not in alphabetical order, the incorrect format for a book chapter/journal article is used, information is missing e.g. no place of publication had been provided (bibliography); repeated sources on the reference list.
Congruence between in-text referencing and bibliography/ reference list • All sources are accurately reflected and are all accurately included in the bibliography/ reference list.	**Generally, congruence between the in-text referencing and the bibliography/ reference list with one or two errors.** • There is largely a match between the sources presented in-text and the bibliography. • For example, a source appears in the text, but not in the bibliography/ reference list or vice versa.	**A lack of congruence between the in-text referencing and the bibliography.** • No relationship/several incongruencies between the in-text referencing and the bibliography/reference list. • For example, sources are included in-text, but not in the bibliography and vice versa, a link, rather than the actual reference is provided in the bibliography.
In summary: the recording of references is accurate and complete.	In summary, at least **80%** of the sources are correctly reflected and included in a reference list.	In summary, at least **60%** of the sources are incorrectly reflected and/or not included in reference list.

Overall Feedback about the consistency, technical correctness and congruence between in-text referencing and bibliography:

.....

.....

Background

In the year 3306, Mars[1] is a green and blue paradise seen from orbit. With the last terraforming completed more than a 100 years ago now, Mars is one of the most beautiful planets in inhabited space. Tourists from almost all the star systems visit Mars every year to see the monuments to the original settlers, and to swim in the pristine oceans.

In New-Johannesburg, Tshepo is contemplating exactly that beauty this morning. Or rather, how to preserve that beauty. He has been reading ancient Earth history again lately, especially about the climate crisis. And he is worried that Mars might repeat Earth's mistakes if they do not consciously decide to do things differently.

Tshepo is the Group CEO of ACME Technologies, one of the leading groups of high-tech manufacturing companies on Mars. They design and manufacture virtual reality headsets, full body virtual reality suits, communicators, movie projectors and many other consumer gadgets. He has approached you to advise him about how to collect, store and process data for the group's new carbon accounting project code-named **ACME Green** – probably the first ever on Mars!

You should answer all the questions within the context of this scenario. Apply your knowledge – do not just present the textbook facts!

[1] If you want to read more about Mars, visit these websites:

1. http://www.astronomy.com/news/2018/09/scientists-draw-up-plan-to-colonize-mars;
2. https://www.spacex.com/mars;
3. https://www.imdb.com/title/tt4939064/

Question 1 **(Marks: 10)**

ACME Green is a huge, strategic project across all of the companies in the ACME Technologies group. Tshepo knows that technology will not be the only aspect he will need to manage well for the project to be successful – he also needs the **right people**.

List **five types of users** that would be using the **operational carbon accounting database**, and explain what the **responsibility** of each type of user will be in ACME Technologies.

Question 2 **(Marks: 25)**

Tshepo is a visionary leader with great ideas for his group of companies. But his background is in **electronic engineering**, not software development. He has decided that he needs to understand most of the **jargon** around databases and datamodels before he will feel comfortable with participating in the process to design the **carbon accounting database**.

Answer his questions below, providing **examples** that are **relevant** to **ACME Technologies** wherever possible.

Q.2.1	Tshepo wants to know what the **difference** is between **one-to-many**, **many-to-many** and **one-to-one** relationships. For each of the below **relationships**, provide an **explanation** of what it is and a **relevant example**.		
	Q.2.1.1	One-to-many relationship.	(4)
	Q.2.1.2	Many-to-many relationship.	(4)
	Q.2.1.3	One-to-one relationship.	(4)
Q.2.2	Two interns were overheard arguing over which **notation** to use for **entity relationship diagrams** – Crow's Foot, Chen or Unified Modelling Language (UML). It would be useful to see the three types of diagrams next to each other for comparison. Draw the below relationship in each of the three notations. *"A space station orbits around one and only one planet.* *A planet may have none, one or more space stations in orbit around it."*		
	Q.2.2.1	Crow's Foot.	(3)
	Q.2.2.2	Chen.	(3)
	Q.2.2.3	Unified Modelling Language (UML).	(3)

Q.2.3	Somebody told Tshepo that data redundancy causes problems. And yet he has seen database designs that have **foreign keys** that look to him like **duplicate data**.		
	Q.2.3.1	Explain what **foreign keys** are.	(2)
	Q.2.3.2	Explain why foreign keys are **not redundant data**.	(2)

Question 3 **(Marks: 30)**

Tshepo has put a lot of thought into the data that he wants to store in the carbon accounting database. He has also spoken to the CEOs of a number of companies in the group to get their inputs. Between them, they have come up with a set of **business rules** to be used as the basis for the data model design.

Project ACME Green – Business Rules

1. A single database will be created that will track the carbon footprints[2] of all the companies within the ACME Technologies™ group.
2. All tables in the database must have surrogate primary keys.
3. Each company has a name and consists of one or more departments.
4. Each company has one employee that is the CEO of the company. That employee may only be the CEO of one company, and not all employees are CEOs.
5. Each department must have exactly one employee who is the Green Champion – the person responsible for any initiatives related to the carbon footprint of the department. That employee may only be the Green Champion for a single department, and not all employees are Green Champions.
6. The following data must be recorded for each employee: name and surname.
7. Some departments have products that they are responsible for. Each product must be allocated to exactly one department, and must have a description.
8. The main purpose of the database is to record carbon transactions. Each carbon transaction is accounted for by exactly one department. Optionally, the single product that the carbon transaction relates to can also be recorded.
9. The following data must be recorded for each carbon transaction: date and value.

[2] Read more about carbon footprint here:
1. http://www.takepart.com/flashcards/what-is-a-carbon-footprint/index.html;
2. https://science.howstuffworks.com/environmental/green-science/carbon-footprint.htm;
3. https://www.thecarbonreport.co.za/what-is-a-carbon-footprint/;

Draw an entity relationship diagram (**ERD**) in Unified Modelling Language (**UML**) notation according to these **business rules**. Marks will be awarded as follows:

Entities	5 marks
Primary and Foreign Keys	9 marks
Other Attributes	2 marks

Relationships with Names	6 marks
Multiplicities	6 marks
Correct UML Notation	2 marks

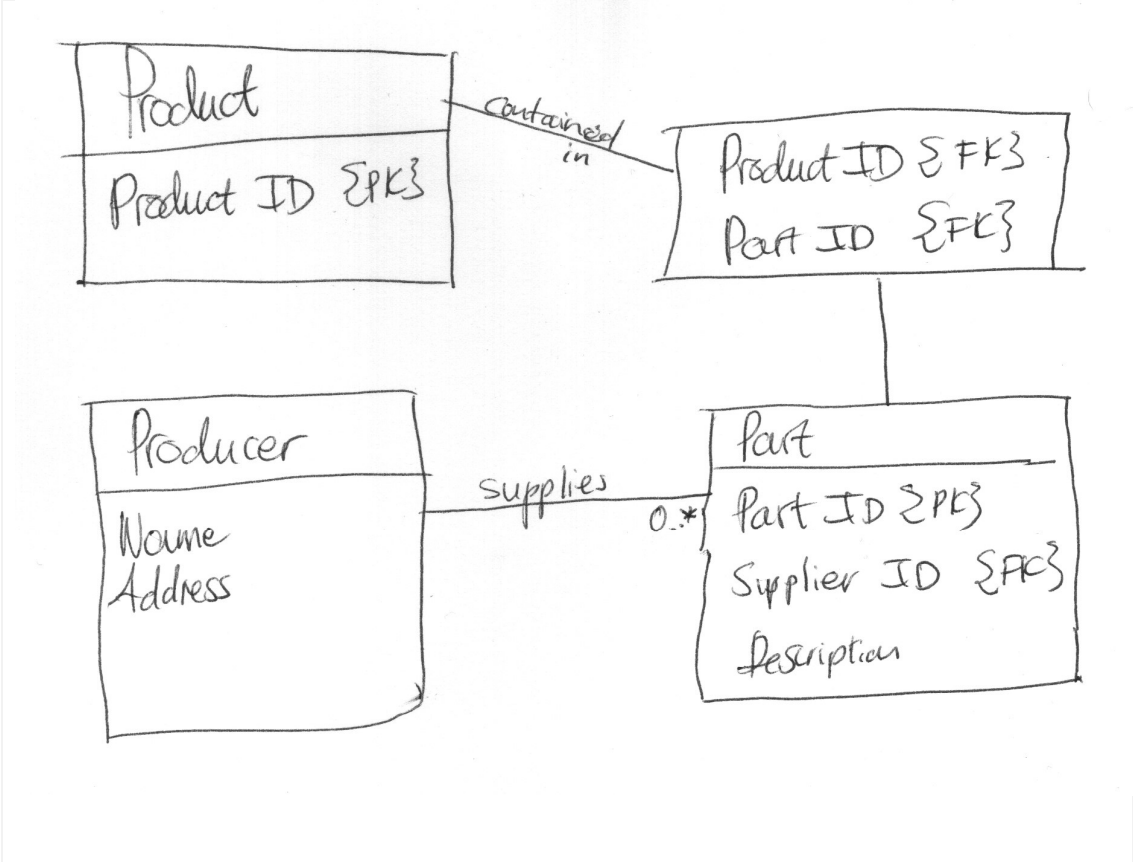
Question 4 **(Marks: 10)**

Now that Tshepo knows that you are an expert with ERDs, he wants you to review a UML ERD that an intern created for another project. Below is the **UML ERD** and the **business rules** used to create it. Which **10 changes** would you propose should be made to the **ERD** to improve its **quality**, so that it can be successfully **implemented** in a **relational database**?

You can make **annotations** on the diagram and/or **describe** the changes in text.

Project ACME Grey – Business Rules

1. All tables in the database must have surrogate primary keys.
2. A product consists of one or more parts. Not all parts captured in the system are part of a product yet, and a part can be used in many products.
3. Each part is supplied by at least one supplier, but multiple suppliers could potentially supply the same part.
4. The following information must be stored for each product: description.



Question 5 **(Marks: 25)**

While the datamodel design process was still underway, an enterprising accountant started putting together data in a **spreadsheet** already. Tshepo applauded his enthusiasm, but he is worried about the **integrity of the data**, despite the fact that the accountant says his data is in **first normal form** already.

Here is a **sample** of the spreadsheet data, in first normal form (**underlined** columns names indicate the **composite primary key**):

Company ID	Company Name	Department ID	Department Name	Product ID	Product Name	Product Type ID	Product Type Name	Date	Value
1	ACME VR	1	Production	1	VR1000	1	Headset	3306/01/23	521
1	ACME VR	1	Production	0	No Product			3306/01/24	-250
1	ACME VR	2	HR	0	No Product			3306/01/23	105
2	ACME Movies	3	Finance	0	No Product			3306/01/23	999
2	ACME Movies	4	Production	2	MC9	2	Camera	3306/01/23	50

Q.5.1 Explain how each of the following **data anomalies** could arise in this **sample data**:

Q.5.1.1 Update anomaly. (2)

	Q.5.1.2	Insertion anomaly.	(2)
	Q.5.1.3	Deletion anomaly.	(2)
Q.5.2	Draw a **dependency diagram** to show the **partial** and **transitive dependencies** in this data.		(6)
Q.5.3	What are the **two major differences** between the **accountant's data** and the **business rules** that were provided by Tshepo? For each difference, explain which option is **better**, and **why**.		(6)
Q.5.4	A software developer suggested to create (among others) a **Product** table for the second normal form of the data (a sample of the data is shown below). **Normalise** this table to **third normal form**, showing **all steps**. Show table(s) with the data.		(7)

Product ID	**Product Name**	**Product Type ID**	**Product Type Name**
1	VR1000	1	Headset
2	MC9	2	Camera
3	AR26	1	Headset

[TOTAL MARKS: 100]