

LEARNING SUMMARY REPORT

INF10002–Database Analysis & Design

NGUYEN NAM TUNG (103181157)

Self-Assessment Details

The following checklists provide an overview of my self-assessment for this unit.

	Pass (P)	Credit (C)	Distinction (D)	High Distinction (HD)
Self-Assessment (please tick)				X

Self-assessment Statement

	Included (please tick)
Learning Summary Report _ Part 1 Completed	X
Test 1 and Test 2 are Completed (at least 10/20)	X
At least 4 Tasks are Completed	X

Minimum Pass Checklist

	Included (please tick)
Learning Summary Report _ Part 1 Completed	X
Test 1 and Test 2 are Completed (at least 10/20)	X
At least 8 Tasks Completed	X

Minimum Credit Checklist

	Included (please tick)
Learning Summary Parts 1 and 2	X
All 5 Pass Tasks Completed	X
All 5 Credit Tasks Completed	X
D task is at Satisfactory Level	X
Test 1 and Test 2 are Complete (at least 10/20 on first attempt)	X
Test 3 is Complete	X

Minimum Distinction Checklist

	Included (please tick)
Learning Summary Parts 1, 2 & 3 Completed	X

All 5 Pass Tasks Completed	X
All 5 Credit Tasks Completed	X
D task is at Satisfactory Level	X
HD task is at Satisfactory Level	X
Test 1 and Test 2 are Complete (at least 10/20 on first attempt)	X
Test 3 is Complete	X

Minimum High Distinction Checklist

Comments

This portfolio contains all of the work that I have completed for INF10002 Database Analysis and Design to a High Distinction level. In this portfolio, I will demonstrate my learning journey in this subject and present some justifications for why I deserve an HD grade. Firstly, I have finished all five pass and credit tasks. Besides that, I passed tests 1 and 2 on my first attempt with respective scores of 15/20 and 18/20. In addition, I completed test 3 with a grade higher than the minimum required for Distinction and High Distinction tasks. Even though I did not receive the grades for the Distinction and Good Distinction works, I believe that I have made a substantial contribution and put significant effort in order to earn a good grade.

Declaration

I declare that this portfolio is my individual work. I have not copied from any other student's work or from any other source except where due acknowledgment is made explicitly in the text, nor has any part of this submission been written for me by another person.

Signature: Nguyen Nam Tung

If the following section does not contain reflection on your learning, **you will not get more than 50** for this unit, regardless of other submissions and achievements.

PART 1 – All students to Complete

Introduction

This report summarises what I learnt in INF10002 Database Analysis and Design. It includes a self-assessment against the criteria described in the unit outline, a justification of the pieces included, details of the coverage of the unit's intended learning outcomes, and a reflection on my learning.

Coverage of the Intended Learning Outcomes

This section outlines how the pieces I have included demonstrate the depth of my understanding in relation to each of the unit's intended learning outcomes.

Reflections on Access and PowerBI

ILO 1:

Define and explain fundamental data and database concepts including tables, relations, keys, queries, transactions, and structured, semi-structured, unstructured data

Access and PowerBI were the primary tools used in the first several weeks. I was introduced to some important concepts, including Primary Key – a unique identifier of the tables, query – a command that may be used to extract desired data from the database, and foreign key – a key that is used to establish the associations between tables. There are additional definitions, but I believe the ones listed above are the most essential for understanding a database.

ILO 2:

Create, store, retrieve, exploit, and visualise data using modern database tools, functions, and techniques

In this section, I will use past portfolio tasks to illustrate my work and understanding.

To begin analysing the information, I needed to import the data using the import text wizard.

You can specify information about each of the fields you are importing. Select fields in the area below. You can then modify field information in the 'Field Options' area.

Field Options

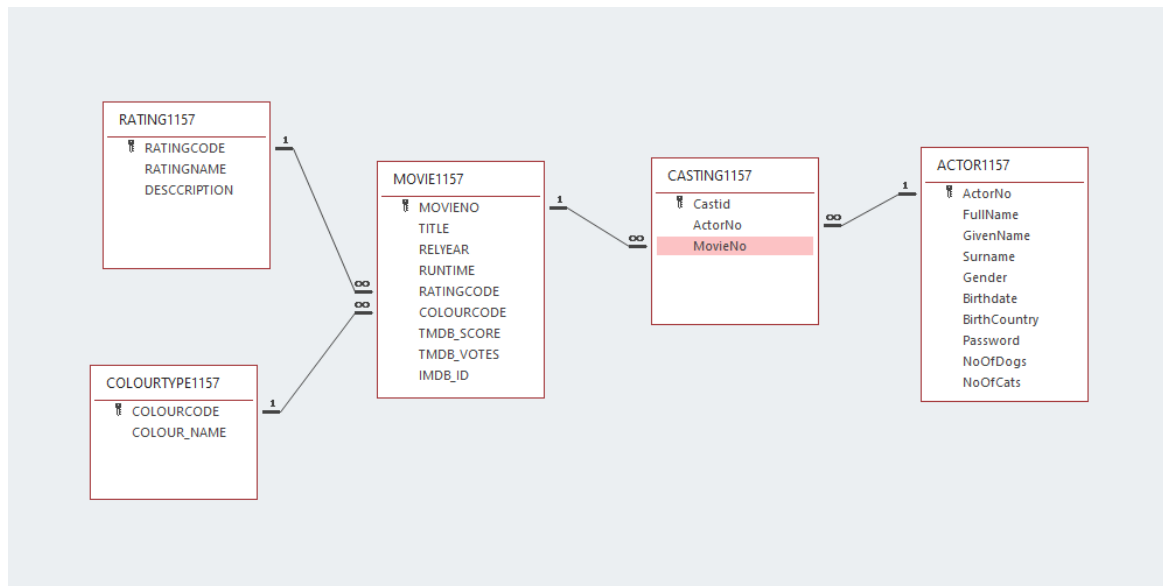
Field Name: Data Type:

Indexed: ☒ Yes (Duplicates OK) ☐ Do not import field (Skip)

StudentId	Gender	GivenName	Surname	DateOfBirth	Faculty
103181157	M	Tung	Nguyen Nam	6/03/2003	Faculty of Computer Science
286069	F	Madeline	Porteus	19/12/1999	Faculty of Business & Law
161176	M	Jayden	Cuthbert	6/12/1999	Faculty of Business & Law
651148	M	George	Southwell	24/11/1999	Faculty of Science Engineering and Te
238428	M	Alexander	Weingarth	10/11/1999	Faculty of Science Engineering and Te
395720	F	Isla	Wildman	7/10/1999	Faculty of Business & Law
268338	F	Lilian	Kessell	30/09/1999	Faculty of Health Arts and Design
603019	M	Thomas	Furnell	24/09/1999	Faculty of Business & Law
225802	M	Tyler	Giffen	17/09/1999	Faculty of Business & Law
843509	F	Abbey	Kerrigan	13/09/1999	Faculty of Business & Law
762199	F	Kiara	Ennor	17/08/1999	Faculty of Health Arts and Design
772698	F	Jorja	Savage	4/08/1999	Faculty of Science Engineering and Te
854578	F	Alannah	Fawkner	21/07/1999	Faculty of Business & Law
670465	M	Toby	Heidenreich	1/07/1999	Faculty of Science Engineering and Te

Advanced... Cancel < Back Next > Finish

Relationships between the tables must be created in order to retrieve data from multiple tables at the same time.



To get the information I needed, I could utilize Access' query design grid. For example, using the movie database, I might extract records about films released in 2009.

Field:	MOVIEID	TITLE	RATINGCODE	RELYEAR	RUNTIME
Table:	MOVIE1157	MOVIE1157	MOVIE1157	MOVIE1157	MOVIE1157
Sort:	Descending				
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:				2009	
or:					

This was the output:

MOVIEID	TITLE	RATINGCODE	RELYEAR	RUNTIME
23082	The Invention of Lying	M	2009	100
19995	Avatar	M	2009	162
19908	Zombieland	M	2009	88
18785	The Hangover	MA	2009	100
16614	Adventureland	M	2009	107
13448	Angels & Demons	M	2009	138
10528	Sherlock Holmes	M	2009	128
8952	I Love You Phillip Morris	MA	2009	98
7980	The Lovely Bones	M	2009	136
767	Harry Potter and the Half-Blood Prince	PG	2009	153
534	Terminator Salvation	M	2009	115

I was also introduced to parameters, which is a way to fetch the records you want based on user input. In this case, I may use this design grid to get data from consumers whose age is younger than the entry age.

Customer Id	Given Name	Family Name	DOB	The Age of the Customer
CUST1157	CUST1157	CUST1157	CUST1157	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				<=[Age you entered]

Furthermore, total functions including as GROUP BY, COUNT, and SUM can be utilized to retrieve data from a database. For example, if I want to know the number of actors by gender, I may create a query design grid that looks like this:

Field:	Gender	Gender
Table:	ACTOR1157	ACTOR1157
Total:	Group By	Count
Sort:		
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:		
or:		

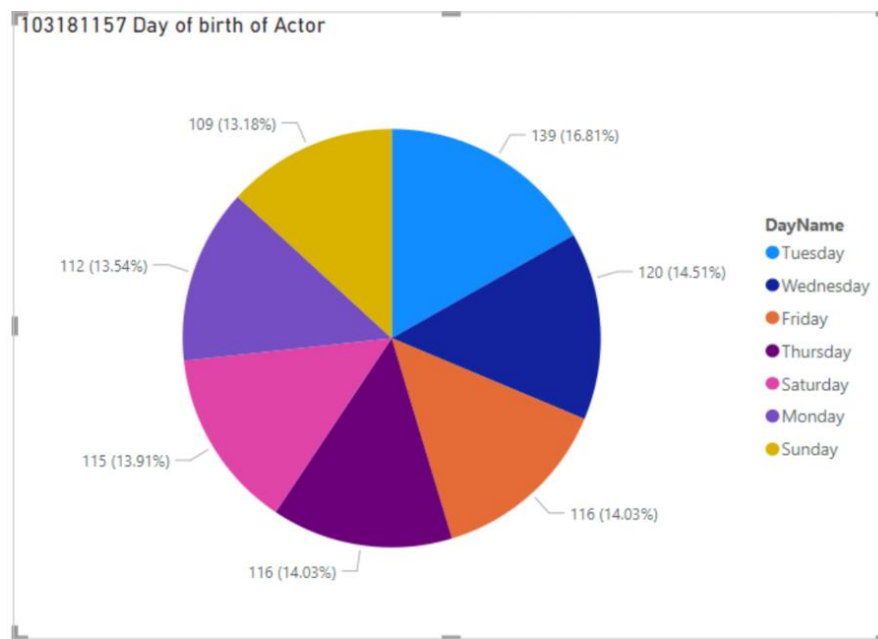
This is the output:

T1PI_1157	
Gender	CountOfGer
	224
M	602

You can obtain records by using a query within a query. This approach must be used in my Distinction task

Field:	Person Id	Name	Gender	Minutes
Table:	People	People	People	Results
Sort:				
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:				
or:				<[SELECT Avg([Results].[Minutes]) FROM [Results] WHERE [Results].[Event] = 2]

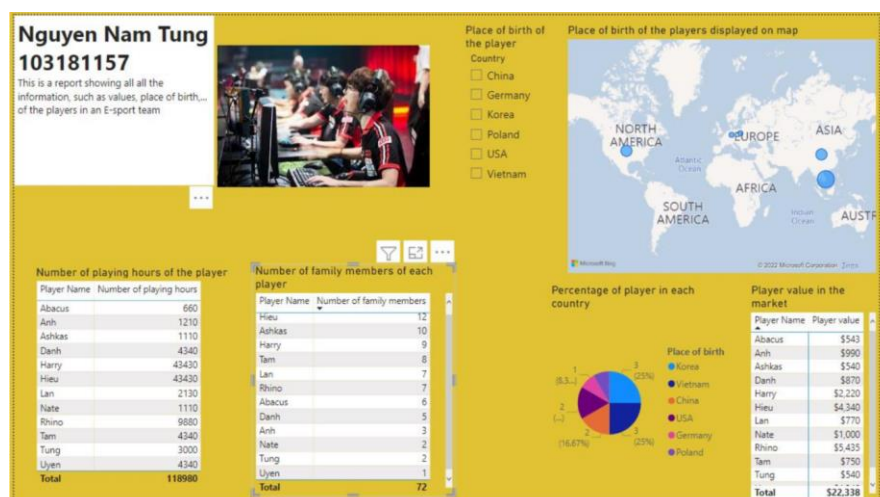
Finally, I was able to visualize the data that I recorded in many formats such as table, pie chart, map, ... This is a sample of the visualizations that I have done.



ILO 3:

Understand and apply conventional data modelling techniques to solve practical database design problems....

I had a chance to apply Power Bi and Access to visualize a practical database design. In task 2C, I could create my own visualizations of an E-sport team. This is how it look likes:



Reflections on SQL and ERDs

ILO 1:

Define and explain fundamental data and database concepts including tables, relations, keys, queries, transactions, and structured, semi-structured, unstructured data

In the following weeks, I was introduced to concepts like ERD, SQL, and Relational Schema, as well as definitions like entity - a real-world object, attribute - a property of an entity, weak entity, and strong entity, which aid in the building of an ERD diagram. SQL Junior was offered to me as a platform for writing query language, and additionally I used draw.io to create ERD diagrams.

ILO 2:

Create, store, retrieve, exploit, and visualise data using modern database tools, functions, and techniques

To completely grasp the principles, I will use my past work to demonstrate what I have learned.

First, I needed to create a table to store the data.

Enter Statements:

```
DROP TABLE CAR1157;
CREATE TABLE CAR1157 (REGNO VARCHAR(100), MAKE VARCHAR (100), MODEL
VARCHAR(100), SALEDATE DATE, ODOMETERREADING NUMBER,
PRIMARY KEY (REGNO))
```

Execute Clear

Insert statements are used to put data in a table:

Enter Statements:

```
INSERT INTO CAR1157 VALUES ('1AMT7U', 'ASTON MARTIN', 'VANTAGE',
TO_DATE('12/03/2022', 'dd/mm/yyyy'),'5300');
INSERT INTO CAR1157 VALUES ('103181157', 'NGUYEN NAM TUNG', 'GAMER',
TO_DATE('06/03/2022', 'dd/mm/yyyy'),'10000');
INSERT INTO CAR1157 VALUES ('1AMTFGR', 'MIN MAX', 'HAI',
TO_DATE('12/06/2022', 'dd/mm/yyyy'),'77300');
```

Execute Clear

This is how the outcome appears:

SELECT * FROM CAR1157 ORDER BY REGNO

REGNO	MAKE	MODEL	SALEDATE	ODOMETERREADING
103181157	NGUYEN NAM TUNG	GAMER	06-MAR-22	10000
1AMT7U	ASTON MARTIN	VANTAGE	12-MAR-22	5300
1AMTFGR	MIN MAX	HAI	12-JUN-22	77300
FWFEFE	TOYOTA	MINS	16-JUL-99	7676

In order to select records from many tables, INNER JOIN can be used

Enter Statements:

```
SELECT '103181157' AS STUID, MOVIE1157.TITLE, MOVIE1157.RELYEAR,
MOVIE1157.RATINGCODE, RATING1157.SHORTDESC, COLOURTYPE1157.COLOUR
NAME
FROM MOVIE1157
INNER JOIN RATING1157
ON MOVIE1157.RATINGCODE = RATING1157.RATINGCODE
INNER JOIN COLOURTYPE1157
ON MOVIE1157.COLOURCODE = COLOURTYPE1157.COLOURCODE
WHERE MOVIE1157.RATINGCODE IN ('M','MA') AND MOVIE1157.RUNTIME > 170
ORDER BY MOVIE1157.MOVIE1157
```

Execute Clear

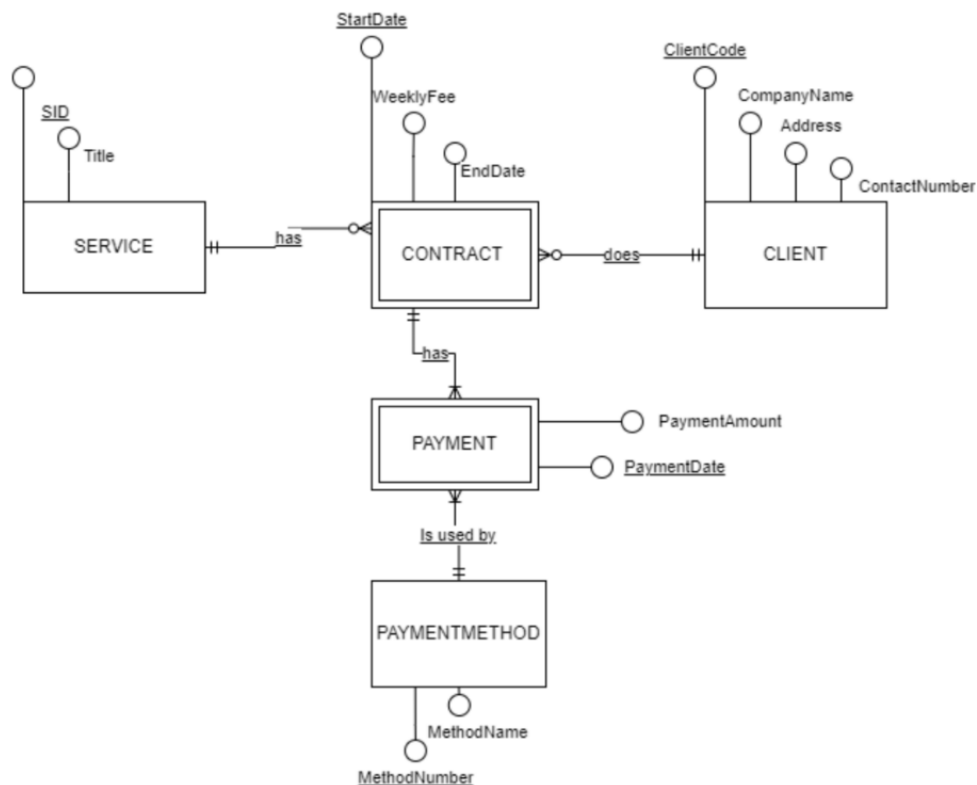
SELECT '103181157' AS STUID, MOVIE1157.TITLE, MOVIE1157.RELYEAR, MOVIE1157.RATINGCODE, RATING1157.SHORTDESC, COLOURTYPE1157.COLOURCODE = COLOURTYPE1157.COLOURCODE WHERE MOVIE1157.RATINGCODE IN ('M','MA') AND MOVIE1157.RUNTIME > 170 ORDER BY MOVIE1157.MOVIE1157

STUID	TITLE	RELYEAR	RATINGCODE	SHORTDESC	COLOURNAME
103181157	The Lord of the Rings: The Fellowship of the Ring	2001	M	MATURE (15+)	Colour Movie
103181157	The Lord of the Rings: The Two Towers	2002	M	MATURE (15+)	Colour Movie
103181157	The Lord of the Rings: The Return of the King	2003	M	MATURE (15+)	Colour Movie
103181157	Schindlers List	1993	M	MATURE (15+)	Black & White Movie
103181157	Titanic	1997	M	MATURE (15+)	Colour Movie
103181157	Pearl Harbor	2001	M	MATURE (15+)	Colour Movie
103181157	Hamlet	1996	M	MATURE (15+)	Colour Movie
103181157	Cloud Atlas	2012	MA	MATURE ACCOMPANIED (15+)	Colour Movie
103181157	The Irishman	2019	MA	MATURE ACCOMPANIED (15+)	Colour Movie

ILO 3:

Understand and apply conventional data modelling techniques to solve practical database design problems....

I have the chance to create an ERD diagram for a practical process. For example, the Distinction task requires me to construct an ERD diagram for a contract with payments between a customer and a company's services.



Normalisations, Transactions, JSON

In the last weeks, we practiced additional ERD diagram and relational schema exercises and were introduced to normalization, which aims to produce three normal forms: 1NF, 2NF, 3NF. For instance, this is the original table for tasks 5P and 5C:

CustId	Name	Phone	CarRego	MakeModel	StartDate	ReturnDate
125	John Coles	0401112233	1AU8HK 1LM3AB	Mazda 3 Hyundai i30	31/08/2020 14/11/2020	7/09/2020 21/11/2020
278	Erin Trump	0466121455	1AU8HK 1KA2CA 1CZ8JK 1AU8HK	Mazda 3 Toyota Camry Mazda 3 Mazda 3	12/09/2020 1/10/2020 10/11/2020 26/11/2020	19/09/2020 8/10/2020 12/11/2020 1/12/2020
721	Emma Knox	0423544117	1LM3AB	Hyundai i30	10/09/2020	13/09/2020

This is after presented in 1NF Normalization

CustID	Name	Phone	CarRego	MakeModel	StartDate	ReturnDate
125	John Coles	0401112233	1AU8HK	Mazda 3	31/08/2020	7/09/2020
125	John Coles	0401112233	1LM3AB	Hyundai i30	14/11/2020	21/11/2020
278	Erin Trump	0466121455	1AU8HK	Mazda 3	12/09/2020	19/09/2020
278	Erin Trump	0466121455	1KA2CA	Toyota	1/10/2020	8/10/2020
278	Erin Trump	0466121455	1CZ8JK	Camry	10/11/2020	12/11/2020
278	Erin Trump	0466121455	1AU8HK	Mazda 3	26/11/2020	1/12/2020
721	Emma Knox	0423544117	1LM3AB	Hyundai i30	10/09/2020	13/09/2020

This is after presented in 3NF Normalization

CustID	Name	Phone
125	John Coles	0401112233
278	Erin Trump	0466121455
721	Emma Knox	0423544117

CarRego	MakeModel
1AU8HK	Mazda 3
1KA2CA	Toyota
1CZ8JK	Camry
1LM3AB	Hyundai i30

CustID	CarRego	StartDate	ReturnDate
125	1AU8HK	31/08/2020	7/09/2020
125	1LM3AB	14/11/2020	21/11/2020
278	1AU8HK	12/09/2020	19/09/2020
278	1KA2CA	1/10/2020	8/10/2020
278	1CZ8JK	10/11/2020	12/11/2020
278	1AU8HK	26/11/2020	1/12/2020
721	1LM3AB	10/09/2020	13/09/2020

I was also introduced to transaction, JSON, and DocumentDBs, as well as explained how a sale or refund activity can go wrong. In the future, as a data analyst, I will have to take these factors into consideration.

Reflection

The most important things I learnt:

The most significant thing I've learnt in this unit is the mindset for maintaining, assessing, and visualizing a database utilizing many tools like Power BI, Access, SQL, and ERD... This is considered as the most significant to me since it allows me to study any database and serves as a foundation for future study of more difficult database-related topics.

I found the following topics particularly challenging:

I believe that the ERD diagram is the most difficult topic in this course. Although I have previously taken a course in data analytics, this is still a very new concept to me. Occasionally, I must deal with a narrative containing numerous entities and attributes, and it can take me many hours to complete a single ERD diagram. This must be something I need to work on in the future.

I found the following topics particularly interesting:

Access and Power BI are the two topics that I find most satisfying. I am able to design my own database tables and use them to generate beautiful visualisations, such as a map or chart. It was a very enjoyable studying experience.

I feel I learnt these topics, concepts, and/or tools really well:

I believe that I have learned queries really well. As a student of Computer Science, I have a lot of programming skills that allow me to easily do tasks that require writing SQL queries. In addition, I have learnt a lot about Query language from external resources, which also contribute to the development of my SQL skills.

I still need to work on the following areas:

As previously stated, the most challenging aspect of this unit for me is working with complex ERD diagrams including numerous entities and attributes at the same time. This is something I need to practice and develop in the long run.

I explored the following topics beyond the unit tasks:

I am currently studying a major in Data Science, thus doing research on this topic is crucial. I've studied Data Analysis and Data Visualization in Python, one of the most popular programming languages, utilizing external libraries like Panda, NumPy, and Matplotlib. I have also read about a

database security concept known as SQL Injection, which is crucial to keep in mind while developing a website that uses a database to manage information. This summer, I intend to learn more about Machine Learning, which is one of the skills a data scientist must have.

The things that helped me most were:

The tutorial task is one of the things that aids my learning the most. I can practice the questions extensively, thereby enhancing my performance in the tests. In addition, my tutor is extremely helpful. He provides me with numerous helpful suggestions for enhancing the quality of my portfolio tasks, as well as recommendations for upgrading my High Distinction assignment.

This unit will help me in the future:

The unit provides a lot of background about database analysis. This will be a great foundation for my future profession as a data scientist.

If I did this unit again, I would do the following things differently:

I would try to apply database principles not just to the portfolio tasks, but also to other projects in order to improve my understanding. I can also examine additional sources such as websites and YouTube to expand my database knowledge. Finally, I may redo the tutorials and portfolio tasks multiple times to recall and comprehend the underlying logical concepts behind them.

PART 2 – Distinction Students to Complete

Reflect on your learning and development of the Distinction Task. You can use the reflection points from Part 1 or you may want to organize your discussion in some other way.

In your reflections include the following points

- Performance in Test 3
- Elaborate on aspects that you found challenging/inspiring/interesting or different to your expectations. Explain why?
- Discuss the approach that you used to solve problems and how the concepts learnt in this unit helped.
- Include discussion of how ideas/techniques/principles from this unit can be used in your further learning inside and outside university

I believe exam 3 is the most difficult aspect of this unit. I completed tests 1 and 2 without difficulty, and I anticipated that test 3 would also be straightforward. Even though I had spent a great amount of time studying for the exam, I had numerous problems completing it under time pressures. I had to cope with a lengthy narrative followed by a complex ERD diagram, which, as I've indicated, caused me the most confusion. I barely completed the last question one minute before the test's end time. I never imagined that I could pass this exam with a score of 3 out of 5. For the Distinction task, I found it much more challenging than the pass and credit tasks, but I spent a great deal of time and effort on it, so I am optimistic that I will receive a good grade, even if the grade has not yet been released. In the Distinction task, I must apply all of the unit's concepts, such as power BI, Access, and query, to a higher level. Due to the increased length of the narrative, it also took me more time to examine the entities and attributes. The ability to visualize data with Power BI is the most interesting aspect of this unit. I can create my own diagrams and play around with enormous amounts of data.

As a Computer Science student majoring in Data Science, the concepts, and ideas I have acquired in this course will be extremely beneficial to my future job. My database knowledge and ability to produce, save, retrieve, and visualise data can be enhanced. Next semester, I will also study a course titled Introduction to Data Science, and this information will be very useful in helping me get a head start on the material.

PART 3 – High Distinction Students to Complete

Reflect on your performance on the High Distinction Task. You can use the reflection points from Part 1, or you may want to organize your discussion in some other way.

In your reflections include the following points

- Present areas that you have personally explored (or will explore) beyond the expectations of the unit, as well as indicate the areas that you plan to learn further on your own and why?
- Highlight ideas/techniques/principles that can be generalized and used in other areas or for further learning and your career.

The High Distinction task is one of the longest assignments that I must complete this semester. From my perspective, the HD assignment has the same level of complexity as the D task, but I must build my own story, input data, and create queries and visualizations from scratch. To complete the HD assignment, I had to evaluate the lecture slides and portfolio tasks to ensure I had covered all the topics. Creating ERD and Narrative is the most significant step because it serves as the foundation for your database and other tasks, such as data entry and visualisations. Although, at this point of writing the Learning Summary, I have not completed the HD assignment, I still believe I am eligible for an HD grade owing to my hard effort and focus on this work.

In addition to taking the course, I have completed a project that needed me to use Query Language. I was required to construct a website for a quiz in which users would answer questions and their scores would be recorded. It looks like this:

What is CoffeeScript?

- ☐ A coding language
- ☒ A type of cold-brew coffee
- ☐ An addon for a video game
- ☐ A simplified version of another coding language
- ☐ A more complicated version of another coding language

What language is CoffeeScript based on?

- ☒ HTML
- ☐ CSS
- ☐ JavaScript
- ☐ TypeScript
- ☐ None of the above

What year was CoffeeScript created?

Please Select ▼

Why was CoffeeScript created?

Write your answer here...

What was the biggest reason for the decline of CoffeeScript?

Write your answer here...

This is how the score will be stored in the database table of the administration page:

All Attempts

Sort by: [Student ID](#) | [Given Name](#) | [Family Name](#) | [Score](#)
 Student Attempt | [List of students who got 100% on their first attempt](#) | [List of students who got less than 50% on their second attempt](#)
 Search (Given Name):
 Search (Student ID):
[Search](#)
[Log out](#)

Attempt ID	Attempt Date	Attempt Time	Student ID	Given Name	Family Name	Attempt Number	Score	Change Score	Delete Attempt
355	2022-05-27	22:19:49	545454545	tbgbgb	bggbg	1	100	<input type="text"/> Save	Delete
356	2022-05-27	22:35:33	5656565	Thanh	Nam	1	100	<input type="text"/> Save	Delete
358	2022-05-27	22:56:21	102649968	Thanh	Tran	1	100	<input type="text"/> Save	Delete
361	2022-05-27	22:59:20	5656565	tbtdh	etefef	2	8	<input type="text"/> Save	Delete

This page allows you to search for records by name, modify the record's score, and delete an attempt. Basically, its functionality consists of performing queries such as DELETE, SELECT, and UPDATE to retrieve the desired records. Furthermore, I have studied Data Analysis and Visualization utilizing libraries like Panda, NumPy, and Matplotlib. To visualize data, you must write lines of code to generate a meaningful diagram and charts, which is significantly more difficult than using Power BI and Access. Finally, I have also explored some information about SQL Injection, which is a technique that can be used to hack websites and destroy database. This is a factor that I need to take a lot of consideration when creating my own data system in my future internship projects.

My future plans include expanding my knowledge of advanced SQL and Machine Learning, which will aid me on the path to becoming a data scientist. I also will study more Access and Power BI as they are the main modelling tools used by large technology organizations. Therefore, I believe mastery in these programmes is crucial and by gaining this subject's knowledge and experience, I would get an ample advantage in my future profession.