



Database Analysis and Design INF60009

Learning Summary Report

Name: PRANAV BALYAN

Student ID: 104680354

Part 1 -

Database Analysis and Design, the most interesting unit of Semester – 1 for me. It started with **Microsoft access**, a database tool and we have learned this for first three weeks.

In the first week, the unit content was about different types of data like structured data, unstructured data, and semi-structured data. Also, the different types of Relational Database Management Systems.

A database is a collection of tables, and we can store that data in Microsoft Access and perform operations like Create, read, update, and delete.

There are some rules which we got to learn in the initial weeks of Microsoft Access teaching – such as a role of Primary Key, foreign key, and constraints.

MOVIE	
Field Name	Data Type
MOVIEID	Number
TITLE	Short Text
RELYEAR	Number
RUNTIME	Number
RATINGCODE	Short Text
COLOUR_CODE	Short Text
TMDB_SCORE	Number
TMDB_VOTES	Number
IMDB_ID	Short Text

Talking of a table structure in Microsoft access, a table has a rows and columns.

The design view of a table shows us about the fields in a table and its datatype. This is simple for me to learn, and I started enjoying from the first lecture.

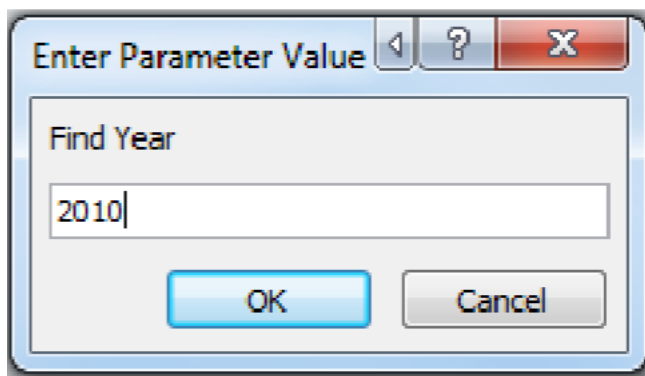
MOVIE								
MOVIEID	TITLE	RELYEAR	RUNTIME	RATINGCODE	COLOUR_CODE	TMDB_SCORE	TMDB_VOTES	IMDB_ID
22	Pirates of the Caribbean: The Curse of the Black Pearl	2003	143 M		C	6.9	2114	tt0325980
24	Kill Bill: Vol. 1	2003	111 MA		C	6.9	1021	tt0266697
70	Million Dollar Baby	2004	132 M		C	7	371	tt0405159
77	Memento	2000	113 MA		C	7.4	788	tt0209144
78	Blade Runner	1982	117 MA		C	7.5	986	tt0083658
95	Armageddon	1998	151 M		C	6.2	760	tt0120591
114	Pretty Woman	1990	119 MA		C	6.4	303	tt0100405
118	Charlie and the Chocolate Factory	2005	115 PG		C	6	426	tt0367594
120	The Lord of the Rings: The Fellowship of the Ring	2001	178 M		C	7.4	3515	tt0120737
121	The Lord of the Rings: The Two Towers	2002	179 M		C	7.4	2877	tt0167261

The above is the datasheet view of a table in Microsoft access. The datasheet view give us a table view of the data in our table.

The columns of a Microsoft Access are known by term field and rows are known by the term record. There is also an option for filtering the records in Microsoft access which is like that in Microsoft Excel. It helps us to view the filtered data.

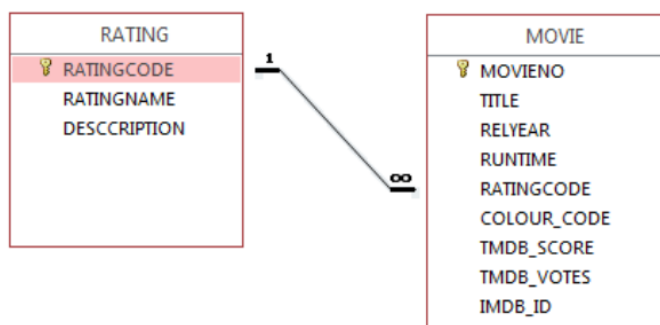
Moving to Access queries, A query can give us an answer for a simple question, like for MOVIE Database, what can be the Rating code of movies released in the year 2015.

I have also learned about AND, OR operators and their importance in a database. The text wildcards are also having significant importance and are to perform a pattern matching. The aggregation functions (count, sum, min, max and average) are simple to use in an. Access query.



Another interesting part for me was queries with parameters. A parameter is a temporary storage location in access database that is given a name. The above parameter will give us data only from the year 2010. We can also combine two queries with parameter.

Creating a new table from scratch was very interesting for me, first we must name the fields and choosing a suitable datatype for the same. After this, moving forward to relationships in Microsoft Access, Relationships helps in connecting the tables in a database and they are of different types like One to Many and Many to one.



One more interesting thing to learn was Referential Integrity. There is an option in Microsoft Access to enforce referential integrity, which will display an error message, if we try to enter a Foreign Key value that does not match the Primary Key value of the referenced table.

Furthermore, studies about indexes and how it improves the performance of a database. There

are more functions like Update query, delete query and Crosstab query which also have their significant importance in a database.

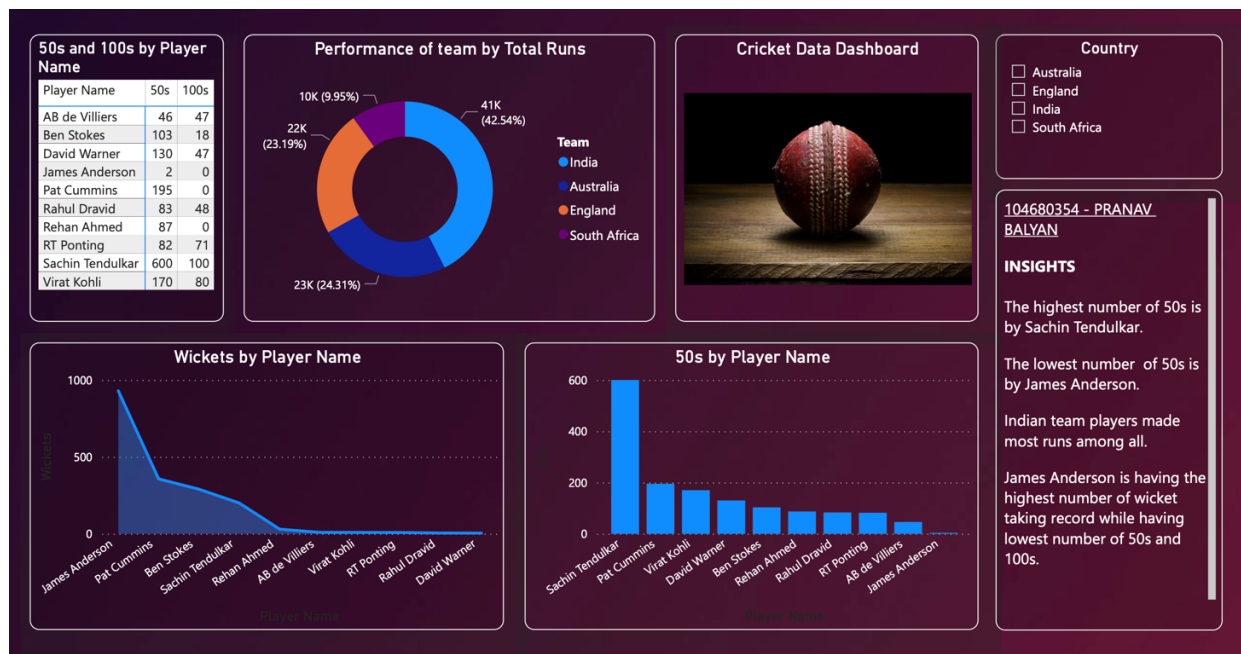
For updating data in a table, we first make a copy of that data because there is no option of UNDO in Microsoft Access and talking in aspect of working in an organization, it can be very harmful for the organization and could be very severe.

And the first three week learning with the Access concludes with the importing and exporting files and first task which was based on Microsoft Access and was smooth for me to answer.

Power-BI

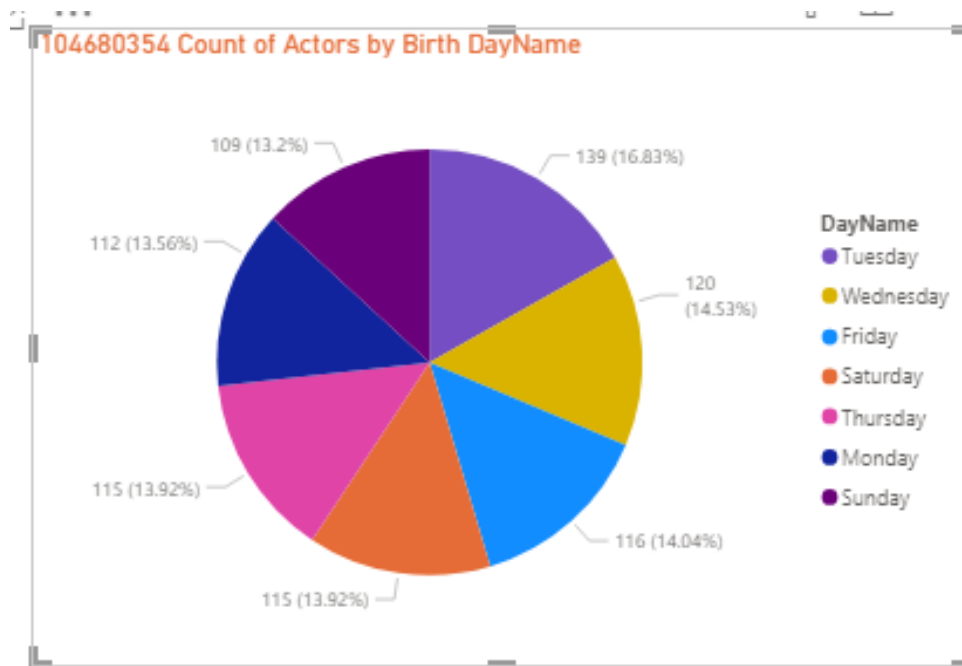
Starting with week four, the course began to concentrate on Microsoft Power BI, a powerful business intelligence application that greatly improved my comprehension of data gathering, integration, and analysis. Power BI presented a combination of challenges and inspiration by introducing innovative ways to show corporate data using advanced visualisations including charts, tables, and maps.

The first lessons taught us the basics of Power BI, starting with how to import data and set up relationships between datasets, which was similar to what we had learned in Microsoft Access. This basic information was very important because it made the change from learning about data structures in a simple database programme to working with large datasets in a business intelligence setting easy.



As part of the actual tasks, especially Task 2, students had to make a complicated dashboard that put what they had learned in class into practice. This job was especially helpful because it let me learn more about how Power BI can turn raw data into visualizations that look good and tell you something.

I eagerly sought out extra resources outside of the classroom to improve my learning experience. This meant doing things like following along with online lessons and using search engines like google to find answers to specific queries. These kinds of tasks not only helped me understand better, but they also showed me how important it is to learn on my own when using complicated software tools. There are a lot of businesses that can use the skills I learn in Power BI, such as marketing, finance, and healthcare. It is very helpful to be able to model, visualize, and understand data to make smart choices and stand out in the workplace.



The above pie charts is one of my answer from my Power-BI tasks.

There are a lot of businesses that can use the skills I learn in Power BI, such as marketing, finance, and healthcare. It is very helpful to be able to model, visualise, and understand data to make smart choices and stand out in the workplace.

My journey with Power BI has been very rewarding, full of obstacles and chances to learn and grow. As time goes on, I want to learn more about Power BI's more advanced features, like AI insights and the ability to combine big data, to get better at analysing data. I learned things that I can use in industry, and they will also have a big effect on my future work activities.

SQL

Week 5 was the start of our study of SQL (Structured Query Language), which is needed to work directly with databases. Before, we interacted with databases through the Microsoft Access frontend interface, which made it hard to see the SQL activities going on underneath. But that link was taken away this week, so we must use SQL tools to talk to the database management system (DBMS) directly. We learned how to use useful SQL commands, like the query: to get movie names and release years from 2015 onwards.

Reflecting on SQL, DML (Data Manipulation Language) commands like INSERT, UPDATE, and DELETE directly modify and manage data, while DDL (Data Definition Language) commands such as CREATE, ALTER, and DROP focus on structuring and defining database schemas. Each plays a crucial role in database interaction and management.

```
SELECT TITLE, RELYEAR FROM MOVIE WHERE RELYEAR >= 2015;
```

This activity showed how precise and direct SQL is when searching databases, which is important for managing and analysing data well.

We studied how to tell the difference between strong and weak entities in Entity Relationship Diagrams (ERDs) during weeks 6 and 7. Strong entities can be created on their own, and each instance has a primary key that makes it unique. Weak entities, on the other hand, get their definition from strong entities and use a partial key that can only be used with the main key of a strong entity to make a unique identification. We also learned about composite keys, which use more than one column to uniquely identify a table row, and how to write complicated SQL queries using "Group By" and "Having" clauses to do calculations like counts, sums, and averages. These skills are very important for making databases that are strong and properly reflect how data depends on each other in the real world.

In Lecture 8, we learned more advanced SQL processes that are needed to manage data structures and relationships that are very complicated. We learned how to model many-to-many (M:M) interactions, which are needed for things like classes and students, since each student can sign up for more than one class and classes can sign up for more than one student. To handle this level of complexity well, you need a join table. We also talked about SQL Outer Joins, which get records from one table that don't exactly match records in another table. This makes sure that no data is left out of the results. We also looked at Subqueries and SQL Unions. Subqueries let you filter data more dynamically, and SQL Unions combine the results of multiple SELECT lines into a single dataset. Lastly, we learned about SQL Views, which are fake tables based on SQL queries that make complex queries easier to use and improve security by limiting direct access to the database tables they are based on.

Enter Statements:

```

With SalesSummary AS
(Select state, postcode, SUM(salesthisyear) AS SumSalesThisYear,
SUM(saleslastyear) AS SumSalesLastYear,
SUM(salesthisyear + saleslastyear) AS "Total Sales for Last 2 Years"
FROM customer0354
WHERE rating IN (1, 2, 3)
GROUP BY state, postcode
ORDER by state ASC, Postcode DESC)

SELECT state, postcode, SumSalesThisYear, SumSalesLastYear
FROM SalesSummary
WHERE SumSalesThisYear = 0
AND SumSalesLastYear = 0;

```

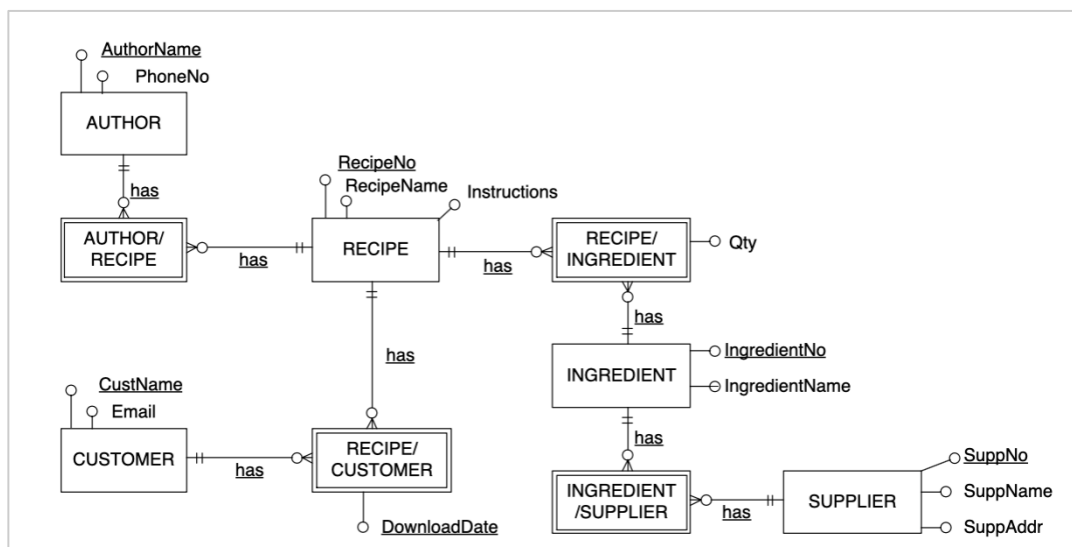
Execute Clear

With SalesSummary AS (Select state, postcode, SUM(salesthisyear) AS SumSalesThisYear, SUM(saleslastyear) AS SumSalesLastYear, SUM(salesthisyear + saleslastyear) AS "Total Sales for Last 2 Years" FROM customer0354 WHERE rating IN (1, 2, 3) GROUP BY state, postcode ORDER by state ASC, Postcode DESC) SELECT state, postcode, SumSalesThisYear, SumSalesLastYear FROM SalesSummary WHERE SumSalesThisYear = 0 AND SumSalesLastYear = 0

STATE	POSTCODE	SUMSALESTHISYEAR	SUMSALESLASTYEAR
New South Wales	2730	0	0
Tasmania	7172	0	0
Victoria	3715	0	0
Victoria	3525	0	0
Victoria	3233	0	0

OK. 5 rows selected.

We worked on making Entity Relationship Diagrams (ERDs) in week 9. ERDs are important tools in database design because they show clearly how different parts of a database are connected. The use of Natural Keys vs. Surrogate Keys in ER modelling was the subject of a major debate. Natural Keys, like Social Security numbers or email names, are better because they keep the data's integrity and make it easier to understand what the data is. We also looked at how feature entities change Foreign Key (FK) relationships in database tables, which is important for making sure that data is consistent and that references are correct.



This ERD depicts a recipe management system where authors create recipes that customers can download. Each recipe consists of various ingredients sourced from suppliers. The ERD tracks relationships such as which customer downloaded which recipe and the suppliers for each ingredient, highlighting the download dates and quantities used.

We talked quickly about database views this week. Views are virtual tables that are created by queries and make complex queries easier. They also make data more accessible and secure. The idea of Views was only quickly explained, but it is very important and will be looked at in more detail in later database units. As another useful skill, we learned how to use Microsoft Visio to make ERDs, which helped us better understand and explain database systems.

After thinking about what we've learned this week, it's clear how important it is to have a well-thought-out ERD and choose key types in a smart way when designing a database. These parts are necessary for making databases that work well, are stable, and can grow as needed. It's not just useful in school; the skills you learn are also necessary in any data-driven work setting. These basic database skills set the stage for further study and understanding of database management and design. They also get us ready for more difficult problems in future classes and in the workplace.

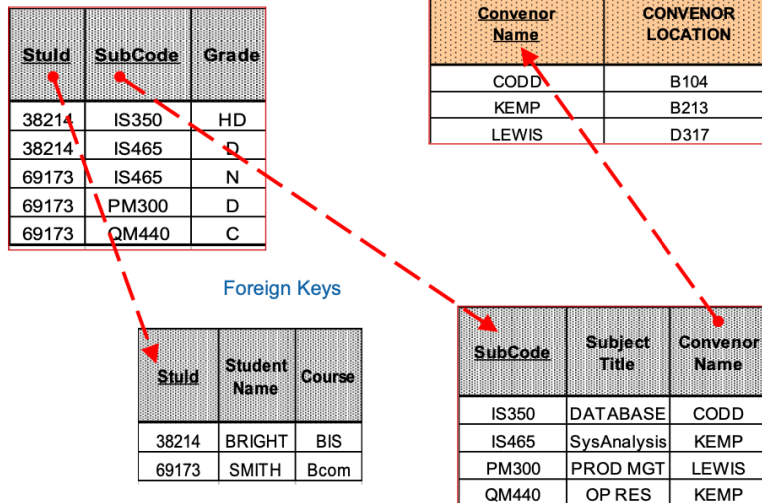
Normalisation and JSON

We learned about Normalisation and JSON files in Document Databases in Week 10 of our Database Analysis and Design class. This helped us understand database management systems better.

Normalisation is a thorough way to improve the structure of a database by getting rid of unnecessary dependencies and duplicate data. The process includes breaking up a single large table into several smaller tables that don't repeat information. It starts by getting rid of repeating groups to get to the First Normal Form (1NF). Then it moves on to the Second Normal Form (2NF) by getting rid of partial key dependencies. Finally, it gets to the Third Normal Form (3NF) by getting rid of non-key dependencies. These steps help stop common problems that happen when databases are used, like mistakes when adding, deleting, or updating data. This makes the database more reliable and efficient.

During the lesson, an example of an unnormalized table with student data was shown. The structure was messed up by things like repeating groups and dependencies. By using normalisation, these problems were fixed in a planned way. The data was organised into several tables that are linked by main and foreign keys. This not only makes the data display clearer, but it also makes it easier to maintain and query.

All relations are now in 3NF



Doc Databases, on the other hand, keep data in JSON (JavaScript Object Notation) files. Document databases don't need a set schema and store data in rows and columns like relational databases do. Instead, each document can store data in a flexible and nested JSON format. This adaptability is especially helpful for programmes that work with different kinds of data that change quickly and don't need structured schema limits.

We looked at how Document Databases keep data, where information is kept in documents instead of rows and properties instead of columns. This method provides a more flexible and expandable way to handle data, making it good for modern apps like discussion boards and social media sites where data structures change and adapt over time.

People who attended the session learned how normalisation and JSON-based Document Databases can help with different types of data handling. Normalisation tries to make relational databases more efficient and consistent, while Document Databases are more flexible and scalable. This shows that database designers can use a variety of approaches based on their needs.

Thinking about these things has really helped me see database design in a new way. It has also made me realise how important it is to pick the right database structure based on the needs and type of data that needs to be handled. It also taught me how to better evaluate and use data management techniques in future projects, whether they use older relational databases or newer non-relational systems.

Transactions and ACID, ETL, Data Warehousing

Anyone who wants to become a database worker or analyst will benefit greatly from the lecture in Week 11 on Normalisation, Database Transactions, and Data Warehousing. By organising data in a way that reduces duplication and maximises integrity, normalisation not only makes database designs more efficient but also makes apps run faster and be able to handle more users. This skill is necessary to build strong databases that can handle complicated queries more quickly and reliably, which is very important in fields that depend on technology.

Product

ProdID	Brand	Description	PricePerOne	QtyInStock
G43546	Gucci	Leather mid-heel pump	1050.00	12

Action

ActionID	ActionDateTime	Action	ProdID	ProdQty	ProdCost
1008	21/01/2021	Purchase	G43546	2	2100.00
1026	23/01/2021	Return	G43546	-1	1050.00

- Write the set of SQL statements to complete each of the transactions in the Action table and update Product table. Make sure you consult lecture slides and lecture recording explaining Transactions concepts.

FOR ActionID - 1008

Update Product Table –

UPDATE PRODUCT

SET QtyInStock = QtyInStock -2

WHERE ProdID = 'G43546'

Inserting in Action Table –

INSERT INTO ACTION (ActionID, ActionDateTime, Action, ProdID, ProdQty, ProdCost)

VALUES (1008, '21/02/2021', Purchase, 'G43546', 2, 2100);

FOR ActionID – 1026

Update Product Table –

UPDATE PRODUCT

SET QtyInStock = QtyInStock + 1

WHERE ProdID = 'G43546'

Inserting in Action Table –

INSERT INTO ACTION (ActionID, ActionDateTime, Action, ProdID, ProdQty, ProdCost)

VALUES (1026, '21/02/2021', Purchase, 'G43546', -1, 1050);

These SQL queries manage inventory by updating product quantities and recording actions for purchases, adjusting stock levels for product ID 'G43546' in the database.

Understanding database transactions and their ACID properties helps you handle and protect database operations well, making sure that data stays consistent and reliable even when multiple people are trying to access it at the same time or the system fails. This knowledge is especially useful in fields like banking and e-commerce, where the security of transactions is very important.

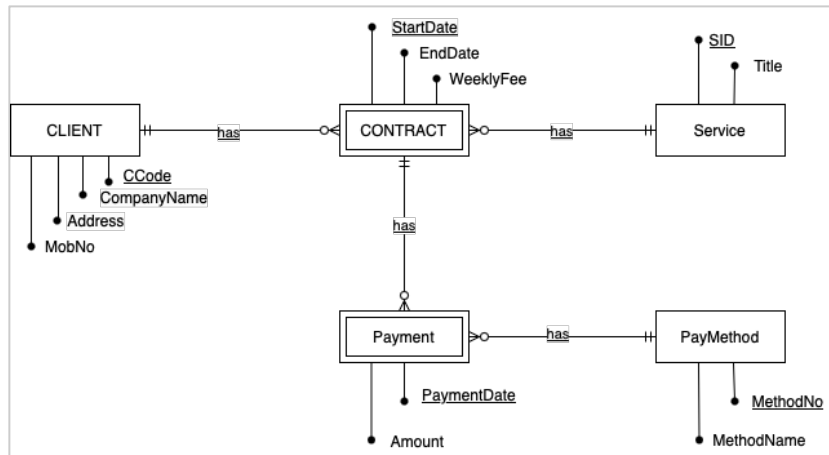
Also, knowing about Data Warehousing and ETL gives you the tools to get insights from big, different sets of data, which makes it easier to make strategic decisions. This skill is necessary for jobs that involve business intelligence and analytics, where you need to turn raw data into information that can be used to make the business better and give you a competitive edge.

Overall, these skills are not only useful in school, but they are also in high demand on the job market. They will set you up for a great career in a wide range of fields, from IT and finance to strategic business roles.

Part 2 [for D & HD students]

After completing the Distinction Task, I've gained important knowledge and skills that have helped me understand and use database design and management better.

ERDs and Schemas



It was both hard and fun to make Entity-Relationship Diagrams (ERDs) and schemas that went with them for two complicated situations with many entities and relationships. This activity strengthened my ability to organise and create databases visually in a way that closely matches how systems work in the real world. Because I had to stay away from surrogate keys, I had to think more deeply about the real identifiers in a system. This helped me learn more about data security and relational design.

Building the Database

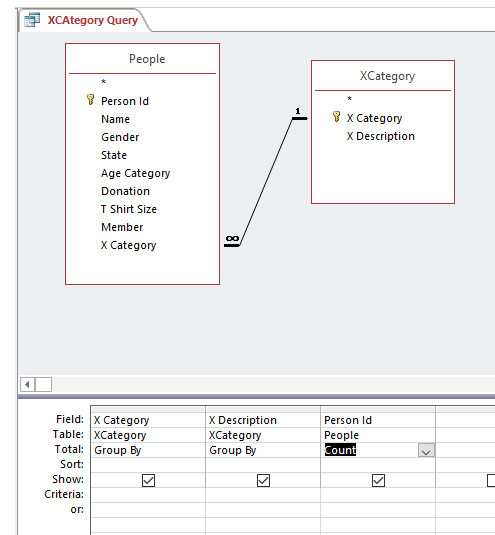
Putting my theoretical knowledge to use in the real world by building a database for Outdoor Events Inc. (OEI) using data they gave me was a great experience. Setting up tables, defining relationships, and making sure the database layout was optimised for fast queries were all part of this. Working with real data during the importing process from CSV files was especially helpful because it simulated real-life data management situations.

Queries and Insert Data

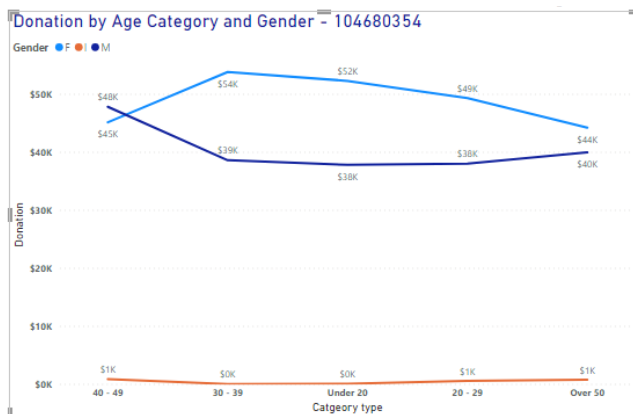
Using SQL skills by writing queries to get specific data from the database was a great way to practice. It was important for this part of the job to have a deep understanding of how data is linked and how to quickly get it without slowing down the database.

XCategory Query

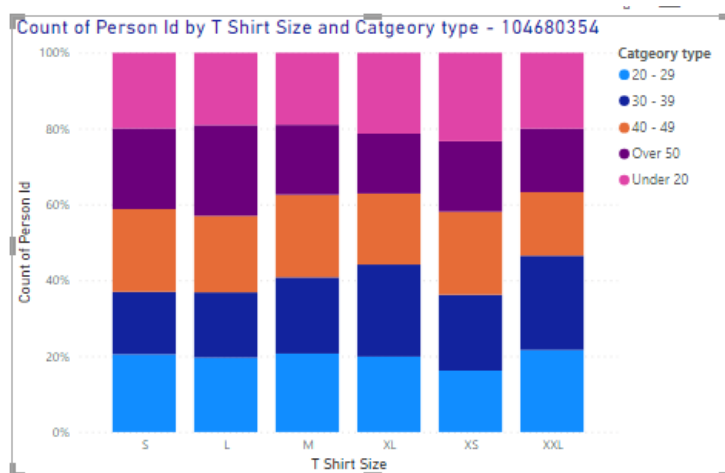
X Category	X Description	CountOfPerson
A	Sprinters	491
B	Rouleurs	436
C	Climbers	470
D	Puncheur	446



Visualisations



Using Power BI to make visualisations of OEI data was a useful experience that helped me connect managing raw data with business intelligence. This task made it clear how important it is to show data visually and how it can be used to help stakeholders make better decisions.



This Distinction Task really put my skills to the test in terms of how well I can plan, configure,

and use a computer system. It needed the use of ERD generation, SQL querying, and data visualisation all at the same time. These are all very important skills in data science and database management. The hands-on experience I got from this job has helped me get ready for more difficult and demanding database projects in the real world. The skills I learned on this project, especially with SQL and Power BI, are ones I expect to use a lot in my future job, maybe in positions that need strong data management and analysis. Also, the things I learned from this job will help me a lot as I learn more about advanced database architecture and business analytics.

Part 3 [for HD students]

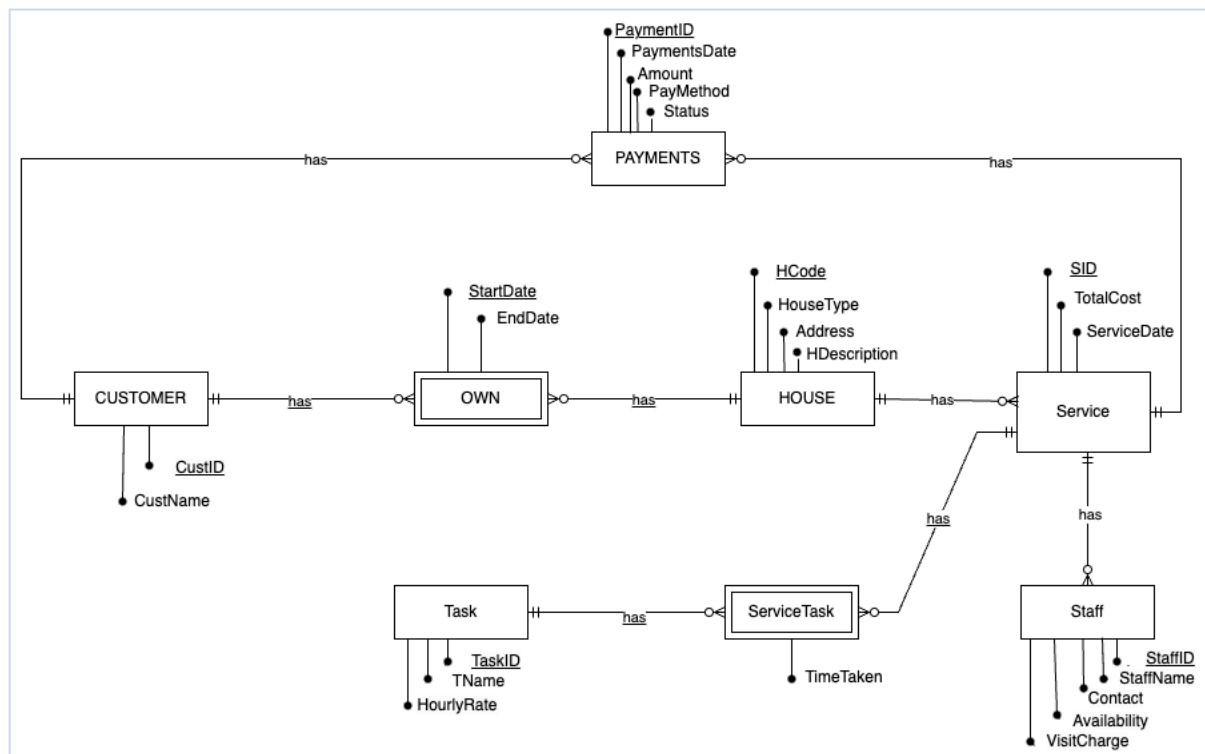
Finished the High Distinction Task, which gave me a deep experience in database design, analysis, and visualization. It ended with a big project that combined all the skills and information I learned in the course.

Part 1 - Business Story

CleanHome Australia's business story was mostly about switching from managing spreadsheets by hand to using a central database system. This story not only set the scene for the whole project, but it also helped me understand the problems that service-based companies face with running their businesses and keeping track of their data. The story was very important in shaping the database's design and features, making sure it met specific business needs like managing reservations, keeping customer information organised, and keeping track of payments.

Part 2: ERDs and Relational Schemas

It took a lot of accuracy and knowledge of how CleanHome Australia's business works to make the Entity-Relationship Diagram (ERD) and the schema that went with it. This part of the project helped me get better at seeing and organizing data in a way that accurately represents complicated connections in the real world and protects the integrity of the data. The difficulty of correctly displaying links without using substitute keys made me value natural keys even more and their part in keeping data consistent.



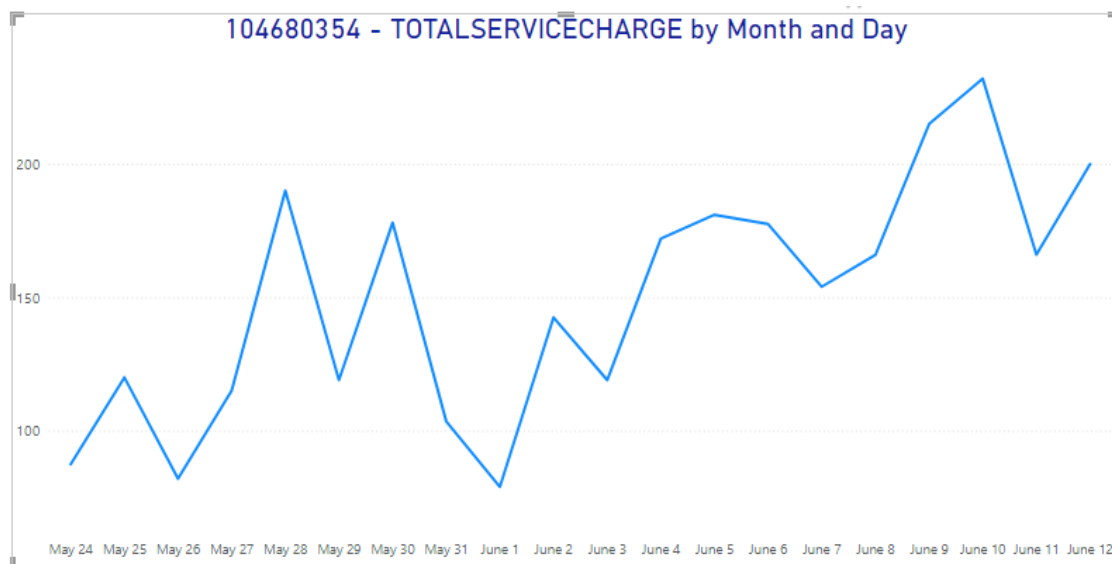
Part 3: The Database Solution

Setting up the database in MS Access and writing SQL scripts to create tables and change data showed how theoretical information could be used in real life. This part of the project was especially satisfying because it turned the ERD and structure into a database that could be used in the real world. Setting up main and foreign keys, defining tables, and making sure the right data types were used helped me understand database architecture better.

Part 4: Test Data and Queries

Adding test data to the database, such as actual scenarios and edge cases, made the system ready for thorough testing and simulating real operations. Creating complicated searches to get useful data from the database was an important part of showing that the database was good at giving us useful information that we could use to make business decisions.

Part 5: Visualizations



It took both imagination and technical know-how to make Power BI visualizations that did a good job of showing the data. This part of the task showed how important it is to present facts clearly and how that can affect business strategies. The visualizations that were made showed important business measures concisely, making it easy to get to important data insights.

Conclusion: This High Distinction Task has been a turning point in my learning, bringing together all the different parts of database creation and use. It pushed me to think of new and useful ways to use what I knew. The result is a project that not only meets school requirements but also prepares me for work in database management and data analysis. The experience has given me a lot more faith and skill in my ability to manage complicated database projects from the beginning to the end.