

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: |
| Course Date: |
|  |

**Table of contents**

[Day 1: Task 1 3](#_Toc203576404)

[Day 1: Task 2 4](#_Toc203576405)

[Day 2: Task 1 5](#_Toc203576406)

[Day 3: Task 1 7](#_Toc203576407)

[Day 3: Task 2 8](#_Toc203576408)

[Day 4: Task 1: SQL Practical 10](#_Toc203576409)

[Day 4: Task 2: Written (Optional) 14](#_Toc203576410)

[**Course Notes** 16](#_Toc203576411)

[**ditional Information** 16](#_Toc203576412)

# Day 1: Task 1

Please research and complete the below questions relating to key concepts of databases.

|  |  |
| --- | --- |
| What is a primary key? | A primary key is a unique identifier for each record in a database table. It ensures that every row in the table is distinct. |
| How does this differ from a secondary key? | A secondary key isn’t selected as the primary key but can still uniquely identify rows in a table |
| How are primary and foreign keys related? | They are related as a foreign key is just the primary key of another table that is being linked |
| Provide a real-world example of a one-to-one relationship | A common example of one-to-one relationships is an individual person and their fingerprints as they are unique to everyone |
| Provide a real-world example of a one-to-many relationship | An example would be a retailer to their products as they don’t just sell one thing at a time and would have to maintain records of all products being sold |
| Provide a real-world example of a many-to-many relationship | A common example would be the relationship between students and a university as it is possible for a student to study more than one course, and each course typically has multiple students associated with it |

# Day 1: Task 2

Please research and complete the below questions relating to key concepts of databases.

|  |  |
| --- | --- |
| What is the difference between a relational and non-relational database? | Relational databases store data in structured, tabular formats with rows and columns, where each table has a predefined schema and relationships between tables are defined using primary and foreign keys whereas non-relational databases use flexible, schema-less models like key-value, document, column-family, or graph structures, which are better suited for handling large volumes of unstructured or semi-structured data |
| What type of data would benefit off the non-relational model?  Why? | Social media data as it is a combination of different data types such as numeric values, text values and images which would not collate together in a relational modelled database and would require separate databases for each data type |

# Day 2: Task 1

You are a junior data analyst at **Northwind Traders**. Your manager has assigned you a series of real-world business tasks. For each scenario:

1. **Write an SQL query** to solve the request. Write an SQL query to solve the request.
2. Run the query in MySQL Workbench using the Northwind database.
3. Paste your query in the box below each question as evidence of completion.

**1.**📘 *Retrieve Full Customer Data*  
Your manager has asked you to export the full list of all customer details into a report.  
➡️ **Write a SQL query to retrieve all columns from the Customers table.**

|  |
| --- |
| **SELECT \***  **FROM Customers;** |

**2.**📘 *Customer Names and Cities for Marketing*  
A marketing analyst is targeting a campaign based on customer names and their cities.  
➡️ **Write a SQL query to retrieve only the CustomerName and City columns from the Customers table.**

|  |
| --- |
| **SELECT CustomerName, City**  **FROM Customers;** |

**3.** 📘 *Unique Cities for Delivery Network Expansion*  
The logistics team wants to know all the different cities where customers are located (no duplicates).  
➡️ **Write a SQL query to retrieve distinct values from the City column in the Customers table.**

|  |
| --- |
| **SELECT DISTINCT City**  **FROM Customers;** |

**4.**📘 *High-Value Products Report*  
The product manager wants to analyse products priced over £50.  
➡️ **Write a SQL query to retrieve all columns from the Products table where the Price is greater than 50.**

|  |
| --- |
| **SELECT \***  **FROM Products**  **WHERE Price > 50;** |

**5.**📘 *International Customers Targeting (USA & UK)*  
The marketing team is preparing campaigns for customers in the USA and UK.  
➡️ **Write a SQL query to retrieve all columns from the Customers table where the Country is either 'USA' or 'UK'.**

|  |
| --- |
| SELECT \*  FROM Customers  WHERE Country IN ('USA', 'UK'); |

**6.**📘 *Recent Orders Report*  
Your team needs to analyze the latest orders.  
➡️ **Write a SQL query to retrieve all columns from the Orders table, sorted by OrderDate in descending order.**

|  |
| --- |
| SELECT \*  FROM Orders  ORDER BY OrderDate DESC; |

**7.**📘 *Mid-Range Products Listing*  
You are preparing a product list for items priced between £20 and £50.  
➡️ **Write a SQL query to retrieve all columns from the Products table where the Price is between 20 and 50, ordered by descending Price.**

|  |
| --- |
| SELECT \*  FROM Products  WHERE Price BETWEEN 20 AND 50  ORDER BY Price DESC; |

**8.**📘 *Local Marketing in the US (Portland & Kirkland)*  
You’re focusing on customers from Portland and Kirkland in the USA.  
➡️ **Write a SQL query to retrieve all columns from the Customers table where the Country is 'USA' and City is either 'Portland' or 'Kirkland', ordered by ascending CustomerName.**

|  |
| --- |
| SELECT \*  FROM Customers  WHERE Country = 'USA'  AND City IN ('Portland', 'Kirkland')  ORDER BY CustomerName ASC; |

**9.** 📘 *Customers from the UK or London*  
Prepare a list for promotional emails targeting customers from the UK or those based in London.  
➡️ **Write a SQL query to retrieve all columns from the Customers table where the Country is 'UK' or City is 'London', ordered by descending CustomerName.**

|  |
| --- |
| SELECT \*  FROM Customers  WHERE Country = 'UK'  OR City = 'London'  ORDER BY CustomerName DESC; |

**10.**📘 *Product Inventory for Selected Categories*  
The inventory team needs products from Category 1 or 2, sorted alphabetically.  
➡️ **Write a SQL query to retrieve all columns from the Products table where the CategoryID is 1 or 2, ordered by ascending ProductName.**

**SELECT \***

**FROM Products**

**WHERE CategoryID IN (1, 2)**

**ORDER BY ProductName ASC;**

# Day 3: Task 1

Please research the below ‘JOIN’ types, explain what they are and provide an example of the types of data it would be used on.

|  |  |
| --- | --- |
| Self-join | A self-join is a regular SQL join where a table is joined with itself. This is useful when you want to compare rows within the same table for example Employees and their managers |
| Right join | A Right Join returns all rows from the right table, and the matching rows from the left table. If there's no match, you'll still get the right table's row, with NULL values for the left table’s columns. |
| Full join | A full join returns all rows from both tables, matching rows where available, and NULL where there's no match. |
| Inner join | An inner join returns only the rows with matching values in both tables based on a specified condition. |
| Cross join | A cross join returns the every row from the first table is combined with every row from the second table. |
| Left join | A left join returns: all rows from the left table, and the matching rows from the right table,  if there's no match, the result will include NULL values for columns from the right table. |

# Day 3: Task 2

You are a junior data analyst at **Northwind Traders**. Your manager has assigned you a series of real-world business tasks. For each scenario:

1. **Write an SQL query** to solve the request. Write an SQL query to solve the request.
2. Run the query in MySQL Workbench using the Northwind database.
3. Paste your query in the box below each question as evidence of completion.

**1.**📘 *Linking Products to Suppliers*  
Management wants to know which supplier provides each product in the inventory.  
➡️ **Write a SQL query to find the supplier of each product.**

|  |
| --- |
| SELECT products.productname AS productname,  suppliers.suppliername AS suppliername  FROM products  JOIN suppliers ON products.supplierid = suppliers.supplierid; |

**2.** 📘 *Classifying Products by Category*  
The category manager is reviewing how products are organized.  
➡️ **Write a SQL query to find the category of each product.**

|  |
| --- |
| SELECT  products.productname AS productname,  categories.categoryname AS categoryname  FROM products  JOIN categories ON products.categoryid = categories.categoryid; |

**3.**📘 *Category-Specific Product Report: Meat/Poultry*  
The food department wants to view all items in the *Meat/Poultry* category.  
➡️ **Write a SQL query to retrieve all products belonging to the Meat/Poultry category.**

|  |
| --- |
| SELECT  products.productname AS productname,  categories.categoryname AS categoryname  FROM products  JOIN categories ON products.categoryid = categories.categoryid  WHERE categories.categoryname = 'Meat/Poultry'; |

**4.**📘 *Complete Order Overview*  
The business team wants to see a detailed order list with customer and employee information.  
➡️ **Write a SQL query to retrieve the Order ID, Order Date, Customer Name, and Employee Name for all orders.**

|  |
| --- |
| SELECT  orders.orderid AS orderid,  orders.orderdate AS orderdate,  customers.customername AS customername,  employees.firstname AS firstname,  employees.lastname AS lastname  FROM orders  JOIN customers ON orders.customerid = customers.customerid  JOIN employees ON orders.employeeid = employees.employeeid; |

**5.**📘 *Supply Chain Overview Report*  
Your manager wants to see the product name, its category, and the name of its supplier all in one report.  
➡️ **Write a SQL query to retrieve the Product Name, Category Name, and Supplier Name for all products.**

|  |
| --- |
| SELECT  products.productname,  categories.categoryname,  suppliers.suppliername  FROM products  JOIN categories ON products.categoryid = categories.categoryid  JOIN suppliers ON products.supplierid = suppliers.supplierid; |

**6.**📘 *Yearly Order Summary – 1996*  
The team is auditing customer orders made in 1996.  
➡️ **Write a SQL query to create a report for all the orders of 1996 and their customers.**

|  |
| --- |
| SELECT  orders.orderid AS orderid,  orders.orderdate AS orderdate,  customers.customername AS customername  FROM orders  JOIN customers ON orders.customerid = customers.customerid  WHERE YEAR(orders.orderdate) = 1996; |

**7.**📘 *Product Count by Category*  
The product team wants to know how many products exist under each category.  
➡️ **Write a SQL query to retrieve all categories along with the number of products in each category.**

|  |
| --- |
| SELECT  categories.categoryname AS categoryname,  COUNT(products.productid) AS productcount  FROM categories  LEFT JOIN products ON categories.categoryid = products.categoryid  GROUP BY categories.categoryname; |

**8.**📘 *Sales Volume Breakdown*  
The sales department wants to analyze how much of each product was ordered and at what price.  
➡️ **Write a SQL query to retrieve all products with their prices and the quantity ordered for each product.**

|  |
| --- |
| SELECT products.productname AS productname, products.price AS price, SUM(order\_details.quantity) AS quantityordered FROM products LEFT JOIN order\_details ON products.productid = order\_details.productid GROUP BY products.productname, products.price; |

# Day 4: Task 1: SQL Practical

You are working as a **junior data analyst** for an international research organization analyzing countries, cities, and languages from around the world. Your job is to write and run SQL queries that provide useful global insights.

For each task:

1. **Write the appropriate SQL query** to meet the scenario's objective.
2. **Run your query** using **MySQL Workbench** and the **world** database.
3. **Take a screenshot** that clearly shows both your SQL code and the result table.
4. **Paste the screenshot in the box provided for each question as evidence of completion.**

**Setting up the database:**

1. **Download world\_db** [**here**](https://b2wcompletetraining057-my.sharepoint.com/:u:/g/personal/yusufs_justit_co_uk/EaLslMes0ZOMw_TMvFQqZykBuiDh836OaSlR6LXXn9m-Ig?e=6Rrusg)
2. **Follow each step to create your database** [**here**](https://b2wcompletetraining057-my.sharepoint.com/:v:/g/personal/yusufs_justit_co_uk/ETXToXTS31xDrvO8lE1MZfcBh1cy5vTqMVxftJ3tF5wwYA?e=Bj1RgY)
3. **Count Cities in USA:** *Scenario:* You've been tasked with conducting a demographic analysis of cities in the United States. Your first step is to determine the total number of cities within the country to provide a baseline for further analysis.

|  |
| --- |
|  |

1. **Country with Highest Life Expectancy:** *Scenario:* As part of a global health initiative, you've been assigned to identify the country with the highest life expectancy. This information will be crucial for prioritising healthcare resources and interventions.

|  |
| --- |
|  |

1. **"New Year Promotion: Featuring Cities with 'New :** *Scenario:* In anticipation of the upcoming New Year, your travel agency is gearing up for a special promotion featuring cities with names including the word 'New'. You're tasked with swiftly compiling a list of all cities from around the world. This curated selection will be essential in creating promotional materials and enticing travellers with exciting destinations to kick off the New Year in style.

|  |
| --- |
|  |

1. **Display Columns with Limit (First 10 Rows):** *Scenario:* You're tasked with providing a brief overview of the most populous cities in the world. To keep the report concise, you're instructed to list only the first 10 cities by population from the database.

|  |
| --- |
|  |

1. **Cities with Population Larger than 2,000,000:** *Scenario:* A real estate developer is interested in cities with substantial population sizes for potential investment opportunities. You're tasked with identifying cities from the database with populations exceeding 2 million to focus their research efforts.

|  |
| --- |
|  |

1. **Cities Beginning with 'Be' Prefix:** *Scenario:* A travel blogger is planning a series of articles featuring cities with unique names. You're tasked with compiling a list of cities from the database that start with the prefix 'Be' to assist in the blogger's content creation process.

|  |
| --- |
|  |

1. **Cities with Population Between 500,000-1,000,000:** *Scenario:* An urban planning committee needs to identify mid-sized cities suitable for infrastructure development projects. You're tasked with identifying cities with populations ranging between 500,000 and 1 million to inform their decision-making process.

|  |
| --- |
|  |

1. **Display Cities Sorted by Name in Ascending Order:** *Scenario:* A geography teacher is preparing a lesson on alphabetical order using city names. You're tasked with providing a sorted list of cities from the database in ascending order by name to support the lesson plan.

|  |
| --- |
|  |

1. **Most Populated City:** *Scenario:* A real estate investment firm is interested in cities with significant population densities for potential development projects. You're tasked with identifying the most populated city from the database to guide their investment decisions and strategic planning.

|  |
| --- |
|  |

1. **City Name Frequency Analysis: Supporting Geography Education** *Scenario*: In a geography class, students are learning about the distribution of city names around the world. The teacher, in preparation for a lesson on city name frequencies, wants to provide students with a list of unique city names sorted alphabetically, along with their respective counts of occurrences in the database. You're tasked with this sorted list to support the geography teacher.

|  |
| --- |
|  |

1. **City with the Lowest Population:** *Scenario:* A census bureau is conducting an analysis of urban population distribution. You're tasked with identifying the city with the lowest population from the database to provide a comprehensive overview of demographic trends.

|  |
| --- |
|  |

1. **Country with Largest Population:** *Scenario:* A global economic research institute requires data on countries with the largest populations for a comprehensive analysis. You're tasked with identifying the country with the highest population from the database to provide valuable insights into demographic trends.

|  |
| --- |
|  |

1. **Capital of Spain:** *Scenario:* A travel agency is organising tours across Europe and needs accurate information on capital cities. You're tasked with identifying the capital of Spain from the database to ensure itinerary accuracy and provide travellers with essential destination information.

|  |
| --- |
|  |

1. **Cities in Europe:** *Scenario:* A European cultural exchange program is seeking to connect students with cities across the continent. You're tasked with compiling a list of cities located in Europe from the database to facilitate program planning and student engagement.

|  |
| --- |
|  |

1. **Average Population by Country:** *Scenario:* A demographic research team is conducting a comparative analysis of population distributions across countries. You're tasked with calculating the average population for each country from the database to provide valuable insights into global population trends.

|  |
| --- |
|  |

1. **Capital Cities Population Comparison:** *Scenario:* A statistical analysis firm is examining population distributions between capital cities worldwide. You're tasked with comparing the populations of capital cities from different countries to identify trends and patterns in urban demographics.

|  |
| --- |
|  |

1. **Countries with Low Population Density:** *Scenario:* An agricultural research institute is studying countries with low population densities for potential agricultural development projects. You're tasked with identifying countries with sparse populations from the database to support the institute's research efforts.

|  |
| --- |
|  |

1. **Cities with High GDP per Capita:** *Scenario:* An economic consulting firm is analysing cities with high GDP per capita for investment opportunities. You're tasked with identifying cities with above-average GDP per capita from the database to assist the firm in identifying potential investment destinations.

|  |
| --- |
|  |

1. **Display Columns with Limit (Rows 31-40):** *Scenario:* A market research firm requires detailed information on cities beyond the top rankings for a comprehensive analysis. You're tasked with providing data on cities ranked between 31st and 40th by population to ensure a thorough understanding of urban demographics.

|  |
| --- |
|  |

# Day 4: Task 2: Written (Optional)

In your groups, discuss and complete the below activity. You can either nominate one writer or split the elements between you. Everyone however must have the completed work below:

*Imagine you have been hired by a small retail business that wants to streamline its operations by creating a new database system. This database will be used to manage inventory, sales, and customer information. The business is a small corner shop that sells a range of groceries and domestic products. It might help to picture your local convenience store and think of what they sell. They also have a loyalty program, which you will need to consider when deciding what tables to create.*

*Write a 500-word essay explaining the steps you would take to set up and create this database. Your essay should cover the following points:*

1. ***Understanding the Business Requirements****:*
   1. *What kind of data will the database need to store?*
   2. *Who will be the users of the database, and what will they need to accomplish?*
2. ***Designing the Database Schema****:*
   1. *How would you structure the database tables to efficiently store inventory, sales, and customer information?*
   2. *What relationships between tables are necessary (e.g., how sales relate to inventory and customers)?*
3. ***Implementing the Database****:*
   1. *What SQL commands would you use to create the database and its tables?*
   2. *Provide examples of SQL statements for creating tables and defining relationships between them.*
4. ***Populating the Database****:*
   1. *How would you input initial data into the database? Give examples of SQL INSERT statements.*
5. ***Maintaining the Database****:*
   1. *What measures would you take to ensure the database remains accurate and up to date?*
   2. *How would you handle backups and data security?*

*Your essay should include specific examples of SQL commands and explain why each step is necessary for creating a functional and efficient database for the retail business.*

|  |  |
| --- | --- |
| Please write your 500-word essay here |  |

|  |
| --- |
| **Course Notes** |

It is recommended to take notes from the course, use the space below to do so, or use the revision guide shared with the class:

|  |
| --- |
| **ditional Information** |

We have included a range of additional links to further resources and information that you may find useful, these can be found within your revision guide.

**END OF WORKBOOK**

**Please check through your work thoroughly before submitting and update the table of contents if required.**

**Please send your completed work booklet to your trainer.**