

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: Hussain Salim |
| Course Date: |
|  |

**Table of contents**

[Day 1: Task 1 3](#_Toc193099469)

[Day 1: Task 2 4](#_Toc193099470)

[Day 3: Task 1 4](#_Toc193099471)

[Day 4: Task 1: SQL Practical 6](#_Toc193099472)

[Day 4: Task 2: Written (Optional) 14](#_Toc193099473)

[**Course Notes** 18](#_Toc193099474)

[**Additional Information** 19](#_Toc193099475)

# 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 used to distinguish each record within a database table. |
| How does this differ from a secondary key? | A primary key uniquely identifies a row or record in a relation whereas a secondary key the combination of fields used in order to create a unique identifier. |
| How are primary and foreign keys related? | The primary key and foreign keys are related as they provide a link between tables. A foreign key in one table references a primary in another table, which allows a connection to be created, allowing related data to be stored in different tables. |
| Provide a real-world example of a one-to-one relationship | One person can only have One passport |
| Provide a real-world example of a one-to-many relationship | One student can take many lessons to complete their GCSE’S |
| Provide a real-world example of a many-to-many relationship | Each Book in a library can have many authors |

# 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? | A relational database organizes data into structured tables with predefined relationships, while a non-relational database stores data in flexible formats like documents, graphs, or key-value pairs without fixed schemas. |
| What type of data would benefit off the non-relational model?  Why? | Non-relational databases excel with unstructured or semi-structured data, such as social media posts, IoT sensor data, or multimedia content, because they offer flexibility, scalability, and the ability to handle diverse data formats efficiently. |

# 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 when a table is joined with itself, often used for hierarchical or relational data, such as finding an employee's manager in an employee table containing both employee and manager IDs. |
| Right join | A **right join** retrieves all rows from the right table and matches them with rows from the left table, often used for scenarios like identifying products with suppliers even if some suppliers have no associated products. |
| Full join | A **full join** combines rows from both tables, including unmatched rows from either side by filling in nulls, and is often used for scenarios like merging customer data with transaction records to include all customers and transactions, even if some don't overlap. |
| Inner join | An **inner join** retrieves only the rows with matching values in both tables, often used for scenarios like linking customer orders with their corresponding details to show only completed transactions. |
| Cross join | A **cross join** produces a Cartesian product by combining every row from the first table with every row from the second table, often used for scenarios like creating a schedule by pairing all employees with all available shifts. |
| Left join | A **left join** retrieves all rows from the left table and matches them with rows from the right table, filling in nulls for unmatched rows, often used for scenarios like displaying all customers and their orders, including customers who haven't placed any orders. |

# Day 4: Task 1: SQL Practical

In your groups, work together to answer the below questions. It may be of benefit if one of you shares your screen with the group and as a team answer / take screen shots from there.

**Setting up the database:**

1. **Download world\_db(1)** [**here**](https://justit831-my.sharepoint.com/:u:/g/personal/danpe_justit_co_uk/Ef6vAaaYVi5FhHqKGxqnn60B9g2khoYekEIO3Y7J00UcJQ?e=pv9NNE)
2. **Follow each step to create your database** [**here**](https://justit831-my.sharepoint.com/:b:/g/personal/danpe_justit_co_uk/EdeCKl2Sas1Hl7u9amDy0fIB9jGVCKxSR0u2-lFOvS5rXw?e=xKv1U7)

**For each question I would like to see both the syntax used and the output.**

1. **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.

|  |
| --- |
| SELECT COUNTRY.Code AS CountryCode, COUNTRY.Name AS CountryName, COUNT(CITY.Name) AS Cities  FROM COUNTRY  JOIN CITY ON COUNTRY.Code = CITY.CountryCode  WHERE COUNTRY.Name LIKE 'United States%'  GROUP BY COUNTRY.Code, COUNTRY.Name; |

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.

|  |
| --- |
| SELECT Name, MAX(LIFEEXPECTANCY) AS LIFE\_EXPECTANCY  FROM COUNTRY  GROUP BY NAME  ORDER BY LIFE\_EXPECTANCY DESC |

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.

|  |
| --- |
| SELECT COUNTRY.Name AS CountryName, CITY.Name AS Cities  FROM COUNTRY  JOIN CITY ON COUNTRY.Code = CITY.CountryCode  WHERE CITY.Name LIKE '%NEW%' |

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.

|  |
| --- |
| SELECT NAME, POPULATION FROM CITY ORDER BY POPULATION DESC LIMIT 10 |

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.

|  |
| --- |
| SELECT NAME, POPULATION  FROM CITY  WHERE POPULATION >= '2000000'  ORDER BY POPULATION DESC |

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.

|  |
| --- |
| SELECT NAME  FROM CITY  WHERE NAME LIKE 'BE%' |

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.

|  |
| --- |
| SELECT NAME, POPULATION  FROM CITY  WHERE POPULATION BETWEEN '500000' AND '1000000'  ORDER BY POPULATION DESC |

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.

|  |
| --- |
| SELECT NAME FROM CITY ORDER BY NAME ASC |

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.

|  |
| --- |
| SELECT NAME, POPULATION  FROM CITY  ORDER BY POPULATION DESC  LIMIT 1 |

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.

|  |
| --- |
| SELECT DISTINCT(NAME), COUNT(NAME) AS CITY\_COUNT\_NAME FROM CITY GROUP BY NAME ORDER BY NAME |

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.

|  |
| --- |
| SELECT NAME, POPULATION FROM CITY ORDER BY POPULATION ASC LIMIT 1 |

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.

|  |
| --- |
| SELECT NAME, POPULATION  FROM COUNTRY  ORDER BY POPULATION DESC  LIMIT 1 |

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.

|  |
| --- |
| SELECT CAPITAL, NAME FROM COUNTRY WHERE NAME LIKE 'SPAIN' |

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.

|  |
| --- |
| SELECT COUNTRY.CONTINENT, CITY.NAME  FROM COUNTRY  JOIN CITY ON COUNTRY.NAME = CITY.NAME  WHERE COUNTRY.CONTINENT LIKE 'EUROPE' |

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.

|  |
| --- |
| SELECT AVG(POPULATION) AS POPULATION, NAME FROM COUNTRY GROUP BY NAME |

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.

|  |
| --- |
| SELECT DISTINCT COUNTRY.NAME AS COUNTRY, COUNTRY.POPULATION, CITY.NAME, COUNTRY.CAPITAL  FROM COUNTRY  JOIN CITY ON COUNTRY.CAPITAL = CITY.ID |

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.

|  |
| --- |
| SELECT NAME, POPULATION, SURFACEAREA, (POPULATION / SURFACEAREA) AS POPULATION\_DENSITY FROM COUNTRY ORDER BY POPULATION\_DENSITY ASC |

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.

|  |
| --- |
| SELECT NAME, POPULATION FROM CITY ORDER BY POPULATION DESC LIMIT 10 OFFSET 30; |

# 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:

|  |
| --- |
| **Additional 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.**