

**Data Technician**

|  |
| --- |
|  |

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

**Table of contents**

[Day 1: Task 1 3](#_Toc106466110)

[Day 1: Task 2 3](#_Toc1673105856)

[Day 3: Task 1 4](#_Toc1682646377)

[Day 4: Task 1: Written 6](#_Toc135095808)

[Day 4: Task 2: SQL Practical 9](#_Toc1228298484)

[Course Notes 19](#_Toc631049525)

[Additional Information 19](#_Toc1612679949)

# 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 table. It cannot be empty and must be unique. |
| How does this differ from a secondary key? | A secondary key (or alternate key) can also identify records but is not the main key. It’s optional and not always unique. |
| How are primary and foreign keys related? | A foreign key in one table refers to a primary key in another table. This creates a link between the two tables. |
| Provide a real-world example of a one-to-one relationship | 1. Countries and their capitals — each country has one capital, and each capital belongs to one country. 2. A person and their passport — each person has one passport, and each passport belongs to one person. 3. A car and its registration certificate — each car has one certificate, and each certificate belongs to one car. 4. A student and their student ID card — each student has one card, and each card belongs to one student. |
| Provide a real-world example of a one-to-many relationship | 1. A parent and their children — one parent can have many children. 2. A teacher and students — one teacher can teach many students. 3. A company and its employees — one company can have many employees. 4. A library and its books — one library can have many books. |
| Provide a real-world example of a many-to-many relationship | 1. Actors and films — actors can act in many films, and each film can have many actors. 2. Authors and books — one author can write many books, and one book can be written by many authors. 3. Students and courses — students can take many courses, and each course can have many students. 4. Many pilots can fly many airplanes. 5. Doctors and patients — one doctor can treat many patients, and each patient can see many doctors. |

# 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? | 1. Data Model: Non-relational databases use various data models, such as document, key-value, graph, or wide-column, while relational databases use the relational model. 2. Schema Flexibility: Non-relational databases often have more flexible schemas, allowing for changes without requiring significant restructuring. 3. Performance: Non-relational databases can be optimised for specific workloads, such as real-time analytics or high-throughput operations. 4. Consistency: Some non-relational databases may prioritise availability and performance over strong consistency, meaning data may not be immediately consistent across all nodes.   **Simple rule:** Pick **relational** for strict structure, strong transactions, and complex joins. Pick **non-relational** when you need flexible schema, very high scale, or data that is naturally documents/graphs/keys. |
| What type of data would benefit off the non-relational model?  Why? | * **Changing / semi-structured data** (user profiles, app settings) — you can add/remove fields without migrations. * **Nested / hierarchical data** (orders with items, comments with replies) — one document holds the whole object → fewer joins, faster reads. * **High-volume events** (logs, clickstream, IoT, metrics) — very fast writes and horizontal scaling. * **Graph relationships** (social networks, recommendations) — graph DBs do quick multi-hop queries. * **Cache / sessions** — key-value stores are ultra-fast. * **Full-text search** (docs, logs) — search engines give ranking and filters out of the box. |

# 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 | I designed all the missing slides in the same style and used the same tables throughout to better understand the differences between the joins. |
| Right join |  |
| Full join |  |
| Inner join |  |
| Cross join |  |
| Left join |  |

# Day 4: Task 1: Written

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 | **1. Understanding the Business Requirements** The small corner shop sells groceries and household products. The database needs to store:   * **Products** (name, category, price, stock quantity) * **Sales** (date, products sold, quantity, price) * **Customers** (name, contact information, loyalty program details) * **Loyalty points or discounts** for customers.   The **users** of the database are shop employees and the owner.   * Employees will use the system to search for products, process sales, and register new customers. * The owner will use it to manage stock and analyse sales.  **2. Designing the Database Schema** The database will include these main tables:   * Products — to store product information * Customers — to store customer and loyalty program details * Sales — to store each sale (date, customer, etc.) * SaleItems — to store details of each product sold during a sale   **Relationships**:   * Sales is linked to Customers by CustomerID * SaleItems is linked to Sales by SaleID and to Products by ProductID  **3. Implementing the Database (SQL)**      **4. Populating the Database**  **5. Maintaining the Database** To keep the database accurate and secure:   * update stock levels after each sale; * make regular backups; * use user roles to control access (e.g., cashiers vs. admin); * duplicate all information; * protect against hacker attacks * use foreign keys and constraints to keep data consistent. * develop a user-friendly interface to ensure proper operation of the database. * offer constantly updated charts by sales, customer, and product to ensure rapid response. |
|  |  |

# Day 4: Task 2: 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.

|  |
| --- |
|  |

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.

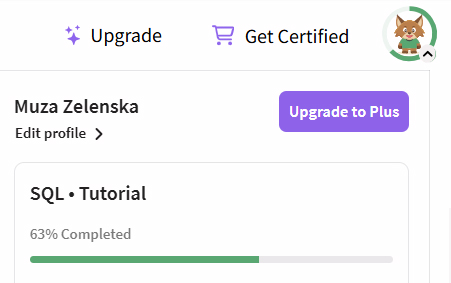
|  |
| --- |
| Tables that were used |

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

A good resource for practicing and improving your knowledge is this site.

<https://www.w3profile.com/Muza>

[W3Schools Online Web Tutorials](https://www.w3schools.com/)



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