

**Data Technician**

|  |
| --- |
|  |

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

**Table of contents**

[Day 1: Task 1 3](#_Toc1614030192)

[Day 1: Task 2 3](#_Toc1905318678)

[Day 3: Task 1 4](#_Toc494476947)

[Day 4: Task 1: SQL Practical 5](#_Toc102053185)

[Day 4: Task 2: Written (Optional) 16](#_Toc1398854365)

[Course Notes 21](#_Toc1205192707)

[Additional Information 21](#_Toc1502713968)

# 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 field that is designed to uniquely identifies each row of a table. |
| How does this differ from a secondary key? | The secondary key is any other field that can be used to uniquely identify each row of a table but was not specifically designed to. This can be useful for filtering. |
| How are primary and foreign keys related? | A foreign key is a field in a table linked to the primary key of another table. Each value of the foreign key relates to a specific row in that other table, the same way the primary key relates to each specific row of this table. |
| Provide a real-world example of a one-to-one relationship | Every individual person will only have one National Insurance number. |
| Provide a real-world example of a one-to-many relationship | Each email address can only be owned by one person, but any given person could have several email addresses. Each person can only have one birth mother, but any given mother could have several children. |
| Provide a real-world example of a many-to-many relationship | Any given student could be enrolled in one or more classes, and any given class could have one or more students enrolled in 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? | A relational database specialises in data that can be structured in a rigid, organised way. This allows data sets to be related and linked to one another.  A non-relational database specialises in data that cannot easily be structured in a rigid, organised way, and doesn’t necessarily need to be related or linked to other data sets. |
| What type of data would benefit off the non-relational model?  Why? | Documents, images, audio/video, graphs, real-time data.  These types of data would benefit from the non-relational model as they wouldn’t have to conform to a rigid predefined schema, allowing more flexible and dynamic data storage and retrieval. |

# 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 join where a table is joined with itself, used for comparing rows within the same table. |
| Right join | A Right join returns all records from the right table, and the matching records from the left table. If there is no match, the left table columns return NULL. Ie, if you joined a list of all customers (right) with a list of orders (left), the join will return all customers regardless of whether they have placed an order. |
| Full join | A full join returns all records from both tables, returning NULL where there is no match. |
| Inner join | An Inner Join returns only the records that have matching values in both tables. |
| Cross join | A Cross Join returns the cartesian product of the two tables, ie every row in the first table is combined with every row in the second table, resulting in all possible combinations. |
| Left join | A Left Join returns all records from the left table, and the matching records from the right table. If there is no match, the right table columns return NULL. |

# 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 COUNT(name) FROM city WHERE CountryCode = “USA”;    274 |

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, LifeExpectancy FROM country ORDER BY LifeExpectancy DESC LIMIT 5;    Andorra |

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 \* FROM city WHERE LEFT(name, 3) = “New”;    The following: |

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 \* FROM city ORDER BY Population DESC LIMIT 10;    The following: |

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 \* FROM city WHERE Population > 2000000;    The following:    Several rows. |

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 \* FROM city WHERE LEFT(name,2) = “Be”;    The following:    Several rows. |

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 \* FROM city WHERE Population BETWEEN 500000 AND 1000000;    The following:    Several rows. |

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 \* FROM City ORDER BY Name ASC;    The Following:    Several rows. |

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 \* FROM City ORDER BY Population DESC LIMIT 1;    Mumbai (Bombay) |

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 name, count(Name) FROM city GROUP BY Name ORDER BY Name ASC;    The following:    (And DESC to show other counts) |

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 \* FROM city ORDER BY Population ASC LIMIT 1;    Adamstown |

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 \* FROM country ORDER BY Population DESC LIMIT 1;  China |

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 city.name AS Capital, country.name AS Country FROM city JOIN country ON city.ID=country.capital WHERE country.name = "Spain";  Madrid |

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 city.name AS CityName, country.name AS countryName FROM city JOIN country ON city.CountryCode = country.Code WHERE country.continent = "Europe";  The following:    Several rows. |

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(city.Population) AS AvgPopulation, country.name AS countryName FROM city JOIN country ON city.CountryCode = country.Code GROUP BY country.name ORDER BY country.name ASC;    The following: |

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 country.name AS Country, city.name AS Capital, city.population AS Population FROM city JOIN country ON city.ID=country.capital ORDER BY city.population DESC;  The following: |

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 AS PopulationDensity FROM country ORDER BY PopulationDensity ASC;  The following: |

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.

|  |
| --- |
| SELECT Name, Continent, GNP FROM country WHERE GNP > (SELECT AVG(GNP) FROM country) ORDER BY GNP DESC;    The following: |

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 \* FROM city ORDER BY Population LIMIT 10 OFFSET 30;    The following: |

# 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 | Day 4: Task 2: Written  1.a.  The data base will need to store data on products, buy prices, sell prices, stock quantity, sales quantity, customer info, customer purchase info,  1.b.  The database will be used by the owners of the shop to manage their supplies and maintain a profit.  2.ab.  •Master Inventory Table with productIDs, quantity, unit price, and category info  •Products table with information about the products, ProductID Primary Key  •Category table with information about the categories, CategoryID Primary Key  • Master Sales Table with orders, sales quantity, customerIDs  •Financials table with information about profit.  •Order Details for specific orders.  •Master Customers Table with information about Customers.  • Customer contact info.  • Customer purchase info.  3.a.  CREATE DATABASE ####  CREATE TABLE #### (field1, field2, etc);  4.a.  INSERT INTO #### (Field1, Field2, etc) VALUES (Field1Values, Field2Value, Etc);  5.a.  All sales info is entered directly into the database. Weekly visuals of data are produced to identify trends, outliers, and potential errors.  5.b.  SQL BACKUP DATABASE #### TO DISK = “####”; |
|  |  |

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