

**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? | It is a unique value column or group of columns to recognise records. Uniqueness of column so it can be key data to be used in data relations. |
| How does this differ from a secondary key? | Primary key should have its uniqueness and index but the secondary key does not necessarily need an index. Null value is not allowed for the primary key but it is allowed for the secondary key. A table can have only 1 primary key but multiple secondary keys. |
| How are primary and foreign keys related? | A primary key is unique and cannot have a null value. Foreign key guarantee integrity. The foreign key and primary key are connected through a reference relationship. |
| Provide a real-world example of a one-to-one relationship | I would like to bring a driving license since only 1 driver's license can be issued to one person. It is a unique example. Make a table of drivers and licenses and connect them with the driver ID attribute. |
| Provide a real-world example of a one-to-many relationship | One-to-many is the most used relationship in the real world. It could be found in most of the business data. The customer table and Orders table can be the one. The customer entity will contain the customer id as the primary key, customer name email, phone number etc. The order entity will have the order id as its primary key and the customer id as a foreign key to be related to the customer table. This results in customers having multiple orders. |
| Provide a real-world example of a many-to-many relationship | Can be exampled with students and courses. A student can take multiple courses and a course can be taken. Setting up 3 entities Students, Courses and Enrollments. Let student id and course id to be primary key for each entity and enrolment to have both keys as foreign keys. The enrollment entity will do mid-table to prevent potential data overlapping. |

# 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 has a table and clear schema. Each row has a unique primary key and relations to other tables. It has a scale-up extension that powers up the server with extra RAM or CPU. It uses SQL for queries and has a structured data type. Non-relational data has various data types such as documents, key-value, graphs, videos etc. It has independent data storage and uses scale-out upgrades which means having more servers to extend. It is a Non-SQL query that uses API, JSON and key-value. |
| What type of data would benefit off the non-relational model?  Why? | A relational database is good to be used where data integrity is very important and needs clear relations between tables. It could be used for bank or financial databases, as well as with CRM systems such as.  Non-relation databases is good to be used for Big databases such as social media, IoT or live chat programs like WhatsApp. |

# 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 | It is a regular join where a table is joined with itself. This is used when you want to compare rows within the same table, treating it as if it were 2 separate tables. |
| Right join | It returns all records from the right table and the matched records from the left table. Null values are returned for columns from the left table. |
| Full join | It returns all records when there is a match in either the left or right table.  When combining customer data and order data to view all 2 data together, full join is used. |
| Inner join | It returns only the records where there is a match in both tables. It is good to use when finding customers who placed orders. |
| Cross join | It returns the Cartesian product of the two tables, meaning every row from the first table is combined with every row from the second table. |
| Left join | It returns all records from the left table and the matched records from the right table. No match will be Null.  Good to use listing all customers and their orders including customers who placed no orders. |

# 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 | Made company called fastandfresh that sells fresh groceries online. For data I thought essential are in the ER diagram below.    The users going to use this data is store manager and staff. Manager will need sales data for marketing predictions and staff will need data for inventory to manage stocks.  Database schema is designed to be one to many, connecting customer Information table and Inventory to Sales table using customerID and productID as foreign key. In this schema relation, stock will be adjusted upon sales and will be able to track customer’s purchase behaviour.    Used above SQL to create database.  I would back up data daily base and weekly base for back up plan. Will set daily data to be backed up for a week then automatically deleted after weekly data is backed up. Weekly data base will be stored in archive storage for back up plan. Also to remain data up to date, would set data refresh script to be activated every 1 hour. |

# 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)**
2. **Follow each step to create your database**

**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 CountryCode, count(\*) Total\_country from City WHERE countrycode="USA"; |

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 MaxLp from country group by name order by MaxLp 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 Name FROM city WHERE 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 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%" order by name asc; |

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; |

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, max(population) as MaxP from city group by name order by MaxP 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 name, count(\*)as occurrence from city group by Name order by Name asc; |

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;  SELECT Name, population from country order by population desc LIMIT 10; |

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 \* from city where countryCode='ESP' and District='Madrid'; |

1. **Country with Highest Life Expectancy:** *Scenario:* A healthcare foundation is conducting research on global health indicators. You're tasked with identifying the country with the highest life expectancy from the database to inform their efforts in improving healthcare systems and policies.

|  |
| --- |
| SELECT Name, LifeExpectancy from country order by LifeExpectancy desc LIMIT 1; |

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 Name, continent from Country WHERE Continent = '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 region, avg(population) as AvgP from country group by region order by AvgP desc; |

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 ID, Name, Population from city WHERE ID IN (SELECT capital from Country) order by Name asc; |

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.

|  |
| --- |
| **Cannot understand the questions at all** |

1. **Cities with High GDP per Capital:** *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 (SELECT Name From City WHERE City.ID = Country.capital) AS CapitalName, Name as 'Nation', GNP FROM Country WHERE GNP >= (SELECT AVG(GNP) from Country) AND Capital IN (SELECT ID FROM city) order by GNP desc; |

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 LIMIT 10 OFFSET 30; |

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