

**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 the column or columns that contain values that uniquely identify each row in a table |
| How does this differ from a secondary key? | Primary keys are defined on tables objects only. In SQL, Table extension objects inherit the primary key of the table object they extend, so any key that define in a table extension object is considered a secondary key. |
| How are primary and foreign keys related? | Primary Keys serve as unique identifiers for each row in a database table. Foreign keys link data in one table to the data in another table. |
| Provide a real-world example of a one-to-one relationship | One to One relationship:  One Employee has one Contract.  One driver has one car. |
| Provide a real-world example of a one-to-many relationship | Parent child relationship  Teachers and students' relationship  Book authors and their books |
| Provide a real-world example of a many-to-many relationship | Student and courses.  Actor and films- An actor can act in multiple movies. |

# 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? | Non-relational databases, often referred to as NoSQL databases, store data in a non-tabular form, making them more flexible than traditional relational databases. Unlike relational databases that use tables with rows and columns, non-relational databases use various data models such as documents, key-value pairs, graphs, and wide-column stores. |
| What type of data would benefit off the non-relational model?  Why? | A [non-relational database](https://www.geeksforgeeks.org/non-relational-databases-and-their-types/) is a type of database that does not rely on the traditional tabular structure of rows and columns found in relational databases. Instead, it uses flexible data models such as key-value pairs, documents, graphs, and wide-column stores.  This flexibility allows non-relational databases to manage unstructured, semi-structured, and structured data efficiently. They were designed when data was expected to be partitioned across multiple machines to scale, in contrast to relational databases, which assumed the data would stay on a single machine |

# 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 | The self join, as its name implies, joins a table to itself. To use a self join, the table must contain a column (call it X) that acts as the primary key and a different column (call it Y) that stores values that can be matched up with the values in Column X. The values of Columns X and Y do not have to be the same for any given row, and the value in Column Y may even be null. |
| Right join | The RIGHT JOIN keyword returns all records from the right table (table2), and the matching records from the left table (table1). The result is 0 records from the left side, if there is no match |
| Full join | The FULL OUTER JOIN keyword returns all records when there is a match in left (table1) or right (table2) table records. |
| Inner join | The INNER JOIN keyword selects records that have matching values in both tables. |
| Cross join | the CROSS JOIN is a unique join operation that returns the Cartesian product of two or more tables. This means it matches each row from the left table with every row from the right table, resulting in a combination of all possible pairs of records. |
| Left join | The LEFT JOIN keyword returns all records from the left table (table1), and the matching records from the right table (table2). The result is 0 records from the right side, if there is no match. |

# 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 | **Database for Fast and Fresh Grocery Shop**  **1.The Fast and Fresh Grocery Shop need few tables to store data.**   1. Product Table (Inventory Table)- In Product Table it will have product details like Product\_id, Product\_name, Stock and Price. For example 1 , Milk(20 Units in stock), 1.45 each 2. Sales Table – Sales Table will need Sales details like ( Sales\_ID, Quantity and Sales\_Date) For example: 1, 2 Milk, 15th March 2025. 3. Customer Table- Customer Table needs customer details like Name, Email Address, Loyalty Points. For Example (Dan Healy , [DanH@hotmail.com](mailto:DanH@hotmail.com) , 100)   Usage of Database:   * The Shop Manager needs to track stock and sales trends and profit. * The Shop Staff needs to record transactions and check loyalty points for the customers. * The IT Support team will maintain the database and ensure it runs smoothly and securely store database frequently.   **2. Designing of Database -Schema chart**    **Implementing / Creating Database in SQL:**  CREATE DATABASE IF NOT EXISTS FastAndFresh;  Inventory\_Table  CREATE TABLE Inventory1 ( productID int(4),  ProductName varchar(1000) DEFAULT NULL,  UnitPrice decimal(10,2),  Stock int(4),  PRIMARY KEY (productID) )  ;  Customer Information Table  CREATE TABLE CustomerInfo  ( CustomerID int,  CustomerName varchar(1000) DEFAULT NULL,  Email varchar(200) NOT NULL,  LoyalytPoint int,  PRIMARY KEY (CustomerID) )  ;  Sales Table  CREATE TABLE SalesInfo  ( SalesID int,  CustomerID int,  productID int,  Quantity int not NULL,  TotalPrice int NOT NULL,  SalesDate Date,  PRIMARY KEY (SalesID),  Foreign key (productID) References inventory (productID),  Foreign Key (CustomerID) REFERENCES customerinfo (CustomerID))  ;  Inserting values in customer Table  insert into customerinfo (CustomerId,CustomerName,email,LoyalytPoint) values (1,'Dan Healy','DanH@hotmail.com',100);  insert into customerinfo (CustomerId,CustomerName,email,LoyalytPoint) values (2,'Indie Smat','IndiS@hotmail.com',50);  Insertind Values in Inventory Table  insert into Inventory (productID,ProductName,UnitPrice,Stock) values (1,'Milk',1.50,20);  insert into Inventory (productID,ProductName,UnitPrice,Stock) values (2,'Eggs',3.25,100);  Insert Values in Salesinfo  into salesinfo (SalesID,CustomerID,productID,Quantity,TotalPrice,SalesDate) values (1,1,1,2,4,'2025-01-04');  into salesinfo (SalesID,CustomerID,productID,Quantity,TotalPrice,SalesDate) values (2,2,2,2,3,'2025-01-04'); |

# 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\_Cities 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,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; |

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 MaxPP from city Group by Name order by MaxPP 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 Occurance 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, min(Population) as MinPP from city Group by Name order by MinPP 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; |

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.

|  |
| --- |
| Select Name, Min(Population) as MinP from Country group by Name Order By MinP 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.

|  |
| --- |
| Select country.Name as CountryName, City.name as CityName,GNP, Capital from Country  Inner Join City on Country.Capital = City.Id 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 , Max(population) as MAXPOP from City group by Name Order by MAXPOP limit 10 offset 40; |

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