

# Data Technician

**Name: LAIBA SHAUKAT**

**Course Date: 12 MAY 2025**

**Table of contents**

Day 1: Task 1 ..... 3

Day 1: Task 2 ..... 7

Day 3: Task 1 ..... 8

Day 4: Task 1: Written .....15

Day 4: Task 2: SQL Practical .....21

Course Notes .....39

Additional Information.....40



## Day 1: Task 1

Please research and complete the below questions relating to key concepts of databases.

<b>What is a primary key?</b>	A primary key is a column (or set of columns) that uniquely identifies each record in a table. It cannot be NULL and must be unique.
<b>How does this differ from a secondary key?</b>	<p>Secondary Key (Alternate Key):</p> <ul style="list-style-type: none"><li>• A secondary key is any candidate key that is not chosen to be the primary key.</li><li>• It can also uniquely identify records, like the primary key, but it's not the main identifier.</li><li>• It is often used for querying or searching, not necessarily for enforcing record uniqueness (though it can be).</li><li>• It can be declared with a UNIQUE constraint.</li><li>• Example: If email is unique for each employee, it could be a secondary key.</li></ul> <pre>CREATE TABLE Employees (     employee_id INT PRIMARY KEY,     name VARCHAR(100),     email VARCHAR(100) UNIQUE );</pre>
<b>How are primary and foreign keys related?</b>	<p>Primary and foreign keys are closely related because they work together to create relationships between tables in a database.</p> <p>A primary key uniquely identifies each record in a table. A foreign key, on the other hand, is a field (or fields) in one table that refers to the primary key in another table. This creates a link between the two tables and ensures referential integrity—meaning, the foreign key must match an existing value in the related primary key column or be NULL (if allowed).</p> <p>For example, if you have a Customers table with customer_id as the primary key, and an Orders table where each order needs to know which customer it belongs to, the Orders table would</p>



	<p>include customer_id as a foreign key. This setup ensures that every order is connected to a valid customer.</p> <p>So in summary: a foreign key references a primary key in another table to establish a relationship between records in different tables.</p>
Provide a real-world example of a one-to-one relationship	<p>Here's a SQL example showing a one-to-one relationship between Persons and Passports:</p> <pre>-- Create the Persons table CREATE TABLE Persons (   person_id INT PRIMARY KEY,   name VARCHAR(100),   date_of_birth DATE );</pre> <p>-- Create the Passports table with a one-to-one relationship to Persons</p> <pre>CREATE TABLE Passports (   passport_id INT PRIMARY KEY,   person_id INT UNIQUE, -- Ensures one passport per person   issue_date DATE,   expiry_date DATE,   FOREIGN KEY (person_id) REFERENCES Persons(person_id) );</pre> <p><b><u>Key points:</u></b></p> <ul style="list-style-type: none"> <li>• person_id is the primary key in the Persons table.</li> <li>• person_id in the Passports table is a foreign key that also has a UNIQUE constraint, enforcing that each person can be linked to only one passport.</li> <li>• This structure ensures a strict one-to-one relationship.</li> </ul>
Provide a real-world example of a one-to-many relationship	<p>A real-world example of a one-to-many relationship is the relationship between a customer and their orders.</p> <p><b><u>Example:</u></b></p> <ul style="list-style-type: none"> <li>• One customer can place many orders.</li> <li>• But each order is placed by only one customer.</li> </ul>



	<pre>-- Create the Customers table CREATE TABLE Customers ( customer_id INT PRIMARY KEY, name VARCHAR(100), email VARCHAR(100) );</pre> <pre>-- Create the Orders table with a one-to-many relationship to Customers</pre> <pre>CREATE TABLE Orders ( order_id INT PRIMARY KEY, customer_id INT, -- Foreign key referencing Customers order_date DATE, total_amount DECIMAL(10, 2), FOREIGN KEY (customer_id) REFERENCES Customers(customer_id) );</pre> <p><b><u>Explanation:</u></b></p> <ul style="list-style-type: none"> <li>• customer_id in Orders is a foreign key pointing to the Customers table.</li> <li>• This setup allows each customer to have multiple orders, but each order is tied to only one customer—establishing a one-to-many relationship.</li> </ul>
<p><b>Provide a real-world example of a many-to-many relationship</b></p>	<p>A real-world example of a many-to-many relationship is the relationship between students and courses. Example:</p> <ul style="list-style-type: none"> <li>• One student can enroll in many courses.</li> <li>• One course can have many students enrolled in it.</li> </ul> <p>To represent this in a database, we use a junction (bridge) table to manage the relationship. Tables:</p> <p>Students Table: student_id (Primary Key)</p> <ul style="list-style-type: none"> <li>• name Courses Table: course_id (Primary Key)</li> <li>• course_name Enrolments Table (Junction table): student_id (Foreign Key) • course_id (Foreign Key)</li> <li>• enrollment_date Together, the student_id and course_id in Enrolments form a composite primary key.</li> </ul> <pre>-- Students table CREATE TABLE Students ( student_id INT PRIMARY KEY, name VARCHAR(100) );</pre>

```
-- Courses table CREATE TABLE Courses ( course_id INT PRIMARY KEY, course_name VARCHAR(100) );
```

```
-- Enrolments junction table to handle many-to-many relationship  
CREATE TABLE Enrolments ( student_id INT, course_id INT,  
enrollment_date DATE, PRIMARY KEY (student_id, course_id),  
FOREIGN KEY (student_id) REFERENCES Students(student_id),  
FOREIGN KEY (course_id) REFERENCES Courses(course_id) );  
Explanation:
```

- Each student can appear in the Enrolments table multiple times for different courses.
- Each course can appear multiple times for different students.
- This structure effectively captures the many-to-many relationship.

## Day 1: Task 2

Please research and complete the below questions relating to key concepts of databases.

<b>What is the difference between a relational and non-relational database?</b>	<p>A relational database organizes data into structured tables (rows and columns) where relationships between data are defined using keys—typically primary and foreign keys. These databases follow a strict schema and use SQL (Structured Query Language) for querying and managing data. They are ideal for applications where data integrity, complex queries, and transactions are important—such as banking systems, HR platforms, and traditional business applications. Examples include MySQL, PostgreSQL, Oracle, and SQL Server.</p> <p>On the other hand, a non-relational database (often called NoSQL) stores data in a more flexible format—such as documents, key-value pairs, graphs, or wide-column stores—without requiring a fixed schema. This allows for fast scalability and flexibility in handling unstructured or semi-structured data. Non-relational databases are commonly used in big data, real-time analytics, IoT, and applications with rapidly changing data. Examples include MongoDB (document-based), Redis (key-value), Cassandra (wide-column), and Neo4j (graph-based).</p> <p>In short, relational databases are structured, schema-based, and best for consistent, transactional data, while non-relational databases are more flexible and suited for dynamic or large-scale data with varying structures.</p>
<b>What type of data would benefit off the non-relational model?</b>  <b>Why?</b>	<p>Non-relational databases are ideal for handling data that is unstructured, semi-structured, or rapidly changing, because they offer flexibility, scalability, and performance that traditional relational models may struggle with. Here are some types of data that benefit most from a non-relational model and why:</p> <p><b>1. <u>Unstructured or Semi-Structured Data (e.g., JSON, XML, multimedia)</u></b></p> <p>Why: Non-relational databases like MongoDB allow storage of data in flexible formats such as JSON-like documents, which can have varying fields and nested structures. This is perfect for storing user profiles, product catalogs, or social media posts where not every record needs to have the same fields.</p>



## **2. Big Data and Real-Time Analytics**

Why: Wide-column stores like Cassandra or key-value stores like Redis are optimized for speed and horizontal scalability. They are great for handling massive volumes of data across distributed systems—like sensor data from IoT devices or real-time user interactions on websites.

## **3. Content Management Systems**

Why: Content like blog posts, articles, and media files often vary in structure. Document databases let you store diverse content types without redesigning your schema each time the structure changes.

## **4. E-commerce Product Data**

Why: Products often have unique attributes (e.g., shoes have sizes, books have authors). Non-relational models allow storing products with different fields in the same collection, making it easy to manage diverse inventories without complex joins or schema changes.

## **5. Social Networks or Graph Data**

Why: Graph databases like Neo4j excel at managing interconnected data—such as friendships, follows, or recommendations—where relationships between entities are as important as the data itself.

### **Summary:**

Non-relational databases are best when flexibility, scalability, and speed are more important than strict consistency or complex joins. They thrive in environments where data doesn't fit neatly into rows and columns, and where adapting quickly to changing data structures is essential.

## 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 type of SQL join where a table is joined to itself. This is useful when you need to compare rows within the same table — in other words, you're matching rows in a table with other rows in the same table based on a related column.





Since a table cannot refer to itself directly in a join, we use table aliases to differentiate the instances of the table.

### When to Use a SELF JOIN

You would use a SELF JOIN in situations like:

- Hierarchical relationships: such as employees and managers stored in the same table.
- Finding duplicates or comparing rows.
- Relating products to other products in the same category.

### Example Scenario: Employee and Manager

Each employee has a ManagerID that refers to another employee's EmployeeID.

To find out which employees report to which managers, you could use a SELF JOIN like this:

```
SELECT
    E1.Name AS Employee,
    E2.Name AS Manager
FROM
    Employees E1
JOIN
    Employees E2
ON
    E1.ManagerID = E2.EmployeeID;
```

### Summary

- SELF JOIN joins a table with itself using table aliases.
- It's used for comparing rows in the same table, often in hierarchical or relational structures.
- Useful in organizational charts, product relationships, and self-referencing data.



## Right join

A RIGHT JOIN (also known as RIGHT OUTER JOIN) returns all the rows from the right table, and the matching rows from the left table. If there is no match, NULL values are returned for the columns from the left table. This is useful when you want to retrieve all records from the second (right) table, regardless of whether they have related records in the first (left) table.

```
SELECT columns
```

```
FROM left_table
```

```
RIGHT JOIN right_table
```

```
ON left_table.column = right_table.column;
```

### **When to Use RIGHT JOIN:**

- To show all entries from the right table even if they have no matching entry in the left.

### **For example:**

- All products with or without sales
- All customers regardless of whether they've made purchases
- All classes offered, even if no students are enrolled.

### **Summary:**

- RIGHT JOIN keeps all rows from the right table.
- Fills in NULLs where there is no match in the left table.
- Best for cases where the right table is your "main" dataset and you want to see its full content.



## Full join

A FULL JOIN (also called FULL OUTER JOIN) returns all records from both tables, matching rows where possible. If there is no match, the result will contain NULLs for the missing side.

It's essentially a combination of a LEFT JOIN and a RIGHT JOIN:

- All records from the left table
- All records from the right table
- Matching records show combined data
- Non-matching records show NULL for the missing side.

SELECT columns

FROM table1

FULL JOIN table2

ON table1.column = table2.column;

### When to Use FULL JOIN:

- When you want to combine two tables and see everything from both, regardless of matches
- Great for reconciling data or comparing two datasets
- Example: All employees and all projects, even if not assigned
- All invoices and all payments, even if unmatched

### Summary:

- FULL JOIN = LEFT JOIN + RIGHT JOIN
- Includes all rows from both tables
- Uses NULLs where there's no match
- Useful for comprehensive reports, audits, or data checks.

## Inner join

An INNER JOIN is the most used type of join in SQL. It returns only the rows that have matching values in both tables. If there is no match between the rows, those rows are excluded from the result.

SELECT columns

FROM table1

INNER JOIN table2

ON table1.column = table2.column;

### When to Use INNER JOIN:



- When you only want rows that exist in both tables based on a condition.

- Examples:

- List of products sold
- Employees assigned to a project
- Customers who placed orders
- Students enrolled in classes

#### **Visual Summary:**

Imagine two overlapping circles (like a Venn diagram). The INNER JOIN returns only the middle part — the overlap — where data from both tables matches.

#### **Summary:**

- INNER JOIN returns only matching rows from both tables.
- It's the default type of join if you just write JOIN.
- Best for filtering data to only include related or relevant records.

### **Cross join**

A CROSS JOIN (also called a Cartesian Join) returns the Cartesian product of two tables — meaning it combines every row from the first table with every row from the second table.

\*If Table A has  $n$  rows and Table B has  $m$  rows, the result of a CROSS JOIN will have  $n \times m$  rows.

SELECT \*

FROM table1

CROSS JOIN table2;

#### **When to Use CROSS JOIN:**

- When you need every combination of values between two datasets
- Generating:
- Price lists for combinations (e.g., size  $\times$  color)
- Schedules (e.g., classes  $\times$  time slots)
- Test cases (e.g., software configs  $\times$  user roles)
- Often used with filters or WHERE clauses to reduce the number of combinations afterward

#### **Important Notes:**



- No ON clause is used in a CROSS JOIN — unless it's filtered with a WHERE.
- CROSS JOINS can produce very large result sets if both tables are big — so use with caution.
- Often used for reference data, not large transactional data.

**Summary:**

- CROSS JOIN = all combinations of rows from both tables
- Good for generating test data or possible configurations
- Can create large results if not managed properly

**Left join**

A LEFT JOIN (also known as a LEFT OUTER JOIN) returns all records from the left table, and the matched records from the right table. If there is no match, NULLs are returned for columns from the right table.

SELECT columns

FROM left\_table

LEFT JOIN right\_table

ON left\_table.column = right\_table.column;

**When to Use LEFT JOIN:**

- When you want to see all rows from the left table, even if there's no match in the right table.
- Customers and orders (including customers with no orders)
- Products and sales (including unsold products)
- Students and classes (including unenrolled students)

**Summary:**

- LEFT JOIN = all rows from the left table, and matching rows from the right table



- Returns NULL where no match exists
- Useful for identifying missing data or producing complete lists



## 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:**
  - a. What kind of data will the database need to store?
  - b. Who will be the users of the database, and what will they need to accomplish?
2. **Designing the Database Schema:**
  - a. How would you structure the database tables to efficiently store inventory, sales, and customer information?
  - b. What relationships between tables are necessary (e.g., how sales relate to inventory and customers)?
3. **Implementing the Database:**
  - a. What SQL commands would you use to create the database and its tables?
  - b. Provide examples of SQL statements for creating tables and defining relationships between them.
4. **Populating the Database:**
  - a. How would you input initial data into the database? Give examples of SQL INSERT statements.
5. **Maintaining the Database:**
  - a. What measures would you take to ensure the database remains accurate and up to date?
  - b. 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

### Setting Up a Database for a Small Retail Business

Creating a database for a small retail business involves careful planning, design, and implementation to ensure the system supports daily operations such as managing inventory, processing sales, and handling customer information, including a loyalty program. Here's how I would approach this project.

#### Understanding the Business Requirements

First, it's essential to identify the data the business needs to store. This includes:

- Inventory data: product name, category, quantity in stock, supplier, price, and restock level.
- Sales data: date of sale, items sold, quantity, total amount, and payment method.
- Customer data: name, contact details, loyalty card number, and accumulated loyalty points.

Users of the database will include store staff (to update inventory and record sales), managers (to analyse sales trends and stock levels), and possibly customers (for loyalty program tracking via a simple interface). Each user group will need specific access tailored to their tasks.

#### Designing the Database Schema

To support this functionality, the database will include the following main tables:

- Products: stores information about each product.
- Customers: holds customer details and loyalty information.
- Sales: records each sales transaction.
- Sales\_Items: records which products are sold in each transaction.
- Inventory: manages current stock levels (this could be merged with Products or separate depending on design choice).

#### Relationships:

- One customer can have many sales.
- Each sale can involve many products.
- Each product can be sold in many sales.





- Sales\_Items acts as a junction table between Sales and Products.

### Implementing the Database

Using SQL, the database and its tables can be created as follows:

```
CREATE DATABASE RetailDB;  
USE RetailDB;
```

```
CREATE TABLE Customers (  
    CustomerID INT PRIMARY KEY AUTO_INCREMENT,  
    Name VARCHAR(100),  
    Email VARCHAR(100),  
    Phone VARCHAR(15),  
    LoyaltyPoints INT DEFAULT 0  
);
```

```
CREATE TABLE Products (  
    ProductID INT PRIMARY KEY AUTO_INCREMENT,  
    Name VARCHAR(100),  
    Category VARCHAR(50),  
    Price DECIMAL(10,2),  
    Stock INT,  
    RestockLevel INT  
);
```

```
CREATE TABLE Sales (  
    SaleID INT PRIMARY KEY AUTO_INCREMENT,  
    CustomerID INT,  
    SaleDate DATETIME DEFAULT CURRENT_TIMESTAMP,  
    TotalAmount DECIMAL(10,2),  
    FOREIGN KEY (CustomerID) REFERENCES  
Customers(CustomerID)  
);
```

```
CREATE TABLE Sales_Items (  
    SaleItemID INT PRIMARY KEY AUTO_INCREMENT,  
    SaleID INT,  
    ProductID INT,
```



```
Quantity INT,  
FOREIGN KEY (SaleID) REFERENCES Sales(SaleID),  
FOREIGN KEY (ProductID) REFERENCES Products(ProductID)  
);
```

### **Populating the Database**

To input initial data, we use INSERT statements:

```
INSERT INTO Products (Name, Category, Price, Stock,  
RestockLevel)  
VALUES ('Milk', 'Dairy', 1.50, 50, 10);
```

```
INSERT INTO Customers (Name, Email, Phone)  
VALUES ('Jane Doe', 'jane@example.com', '123-456-7890');
```

```
INSERT INTO Sales (CustomerID, TotalAmount)  
VALUES (1, 3.00);
```

```
INSERT INTO Sales_Items (SaleID, ProductID, Quantity)  
VALUES (1, 1, 2);
```

### **Maintaining the Database**

To keep the database accurate and up to date, I would:

- Implement constraints (e.g., CHECK constraints on quantity and price).
- Use triggers to update stock levels automatically after a sale.
- Schedule regular backups using tools like mysqldump.
- Protect data using user roles and permissions, ensuring only authorized users can make changes.

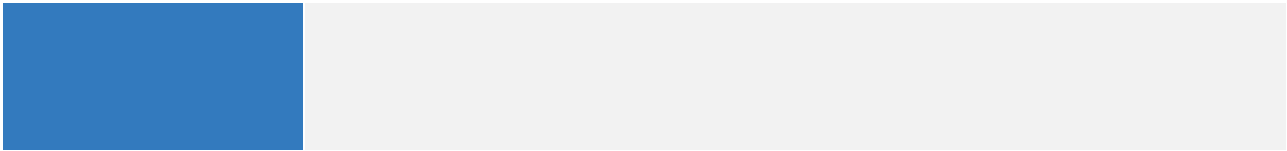
For example:

```
CREATE USER 'store_clerk'@'localhost' IDENTIFIED BY  
'password';  
GRANT SELECT, INSERT, UPDATE ON RetailDB.* TO  
'store_clerk'@'localhost';
```

In conclusion, designing and building a database for a small retail business requires a structured approach that includes understanding the business needs, creating a well-thought-out schema, implementing the system using SQL, and ensuring

long-term maintenance and security. With this approach, the business can efficiently manage its operations and improve customer service.





## Day 4: Task 2: SQL Practical

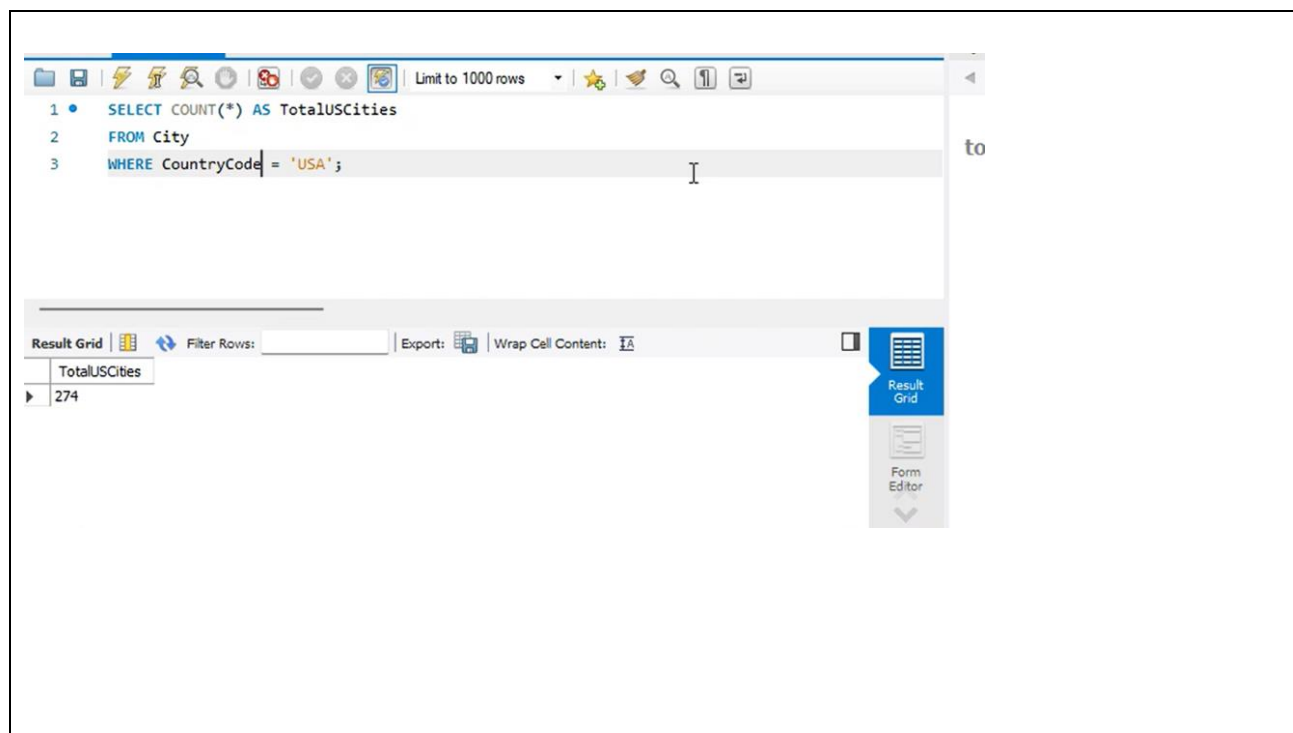
In your groups, work together to answer the below questions. It may be of benefit if one of you shares your screen with the group and as a team answer / take screen shots from there.

### Setting up the database:

1. Download world\_db(1) [here](#)
2. Follow each step to create your database [here](#)

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.



The screenshot shows a SQL query editor interface. The query is as follows:

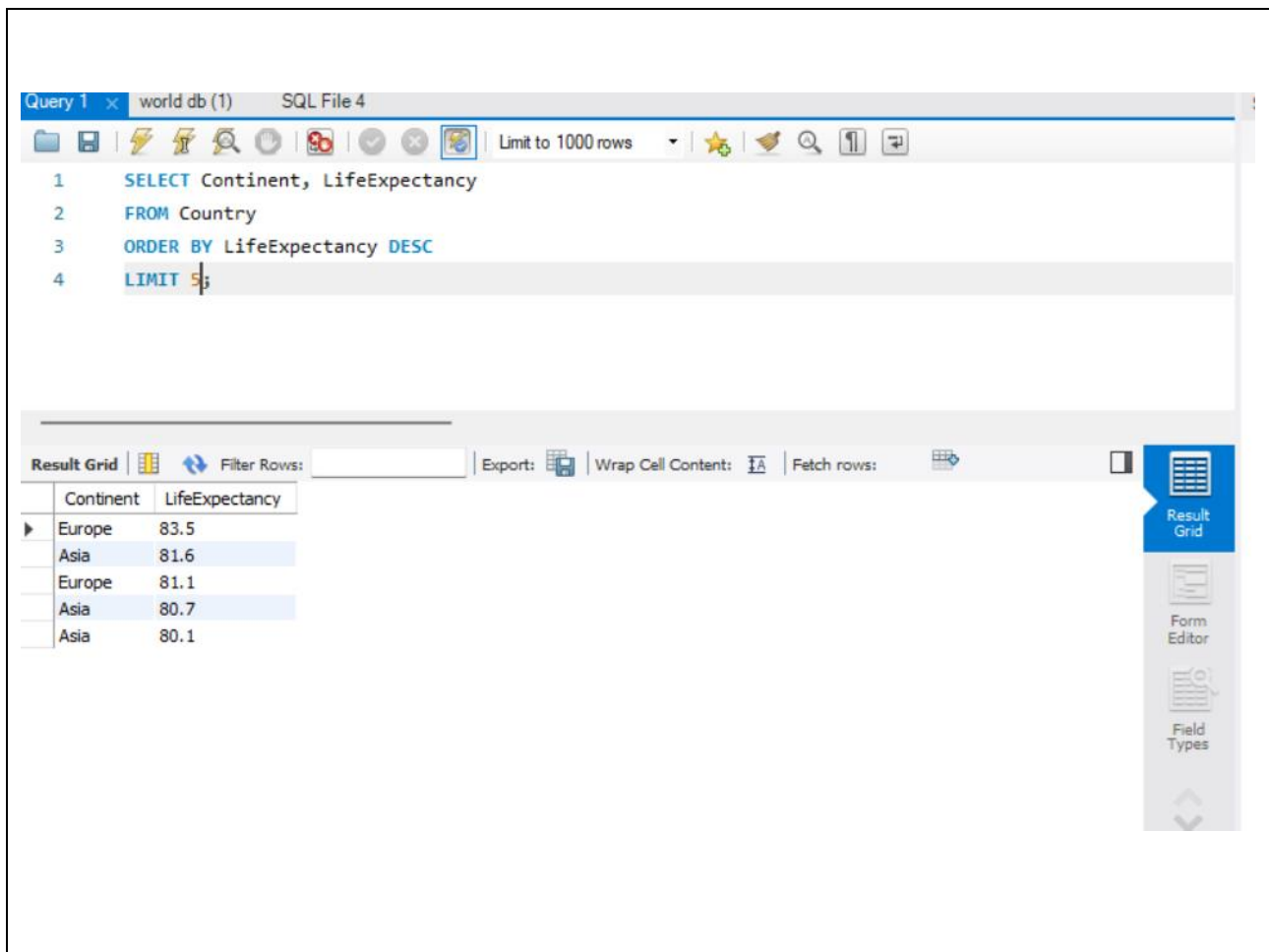
```
1 • SELECT COUNT(*) AS TotalUSCities
2 FROM City
3 WHERE CountryCode = 'USA';
```

The result grid at the bottom shows the output of the query:

TotalUSCities
274

The interface includes a toolbar at the top with various icons, a 'Limit to 1000 rows' dropdown, and a 'Result Grid' button on the right side.

2. **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.



Query 1 x world db (1) SQL File 4

Limit to 1000 rows

```

1 SELECT Continent, LifeExpectancy
2 FROM Country
3 ORDER BY LifeExpectancy DESC
4 LIMIT 5;

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows: |

	Continent	LifeExpectancy
▶	Europe	83.5
	Asia	81.6
	Europe	81.1
	Asia	80.7
	Asia	80.1

Result Grid  
Form Editor  
Field Types

3. **"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 • SELECT *
2 FROM city
3 WHERE name LIKE '%New%'
4

```

ID	Name	CountryCode	District	Population
137	Newcastle	AUS	New South Wales	270324
482	Newcastle upon Tyne	GBR	England	189150
502	Newport	GBR	Wales	139000
734	Newcastle	ZAF	KwaZulu-Natal	222993
936	Kowloon and New Kowloon	HKG	Kowloon and New Kowl	1987996
1106	New Bombay	IND	Maharashtra	307297
1109	New Delhi	IND	Delhi	301297
2857	Khanewal	PAK	Punjab	133000
3793	New York	USA	New York	8008278
3823	New Orleans	USA	Louisiana	484674
3855	Newark	USA	New Jersey	273546
3905	Newport News	USA	Virginia	180150

4. **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.

```

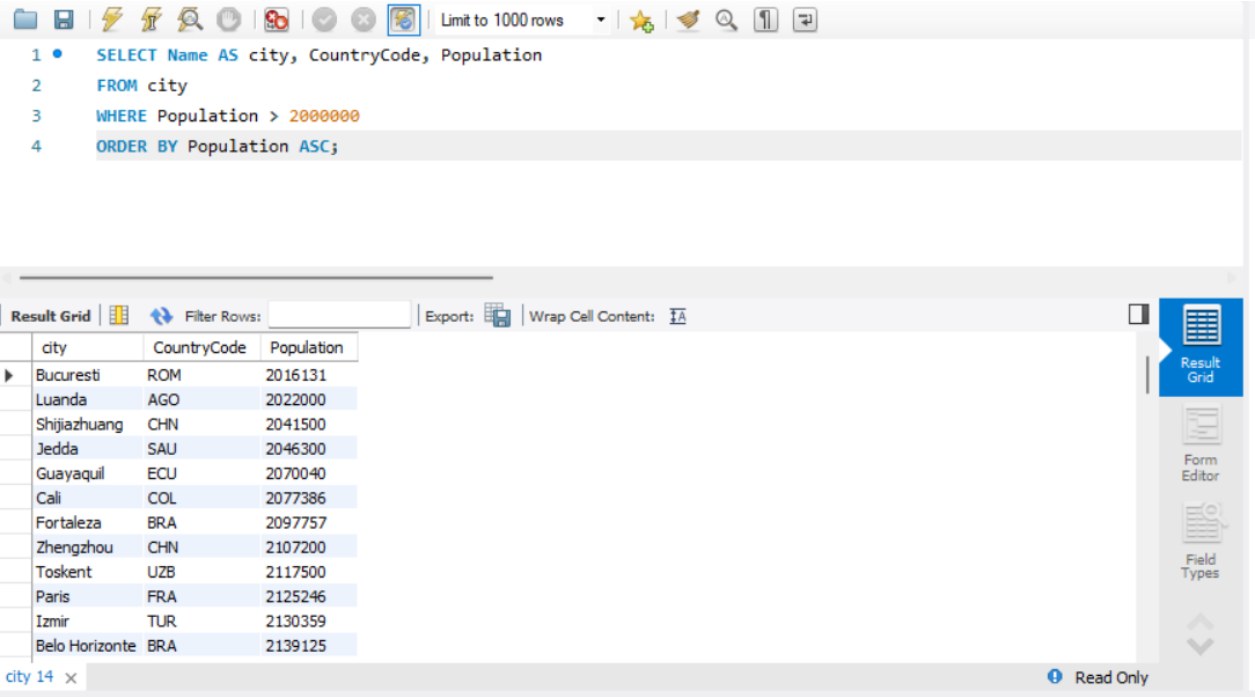
1 • SELECT *
2 FROM city
3 order by Population DESC
4 LIMIT 10;

```

ID	Name	CountryCode	District	Population
1024	Mumbai (Bombay)	IND	Maharashtra	10500000
2331	Seoul	KOR	Seoul	9981619
206	São Paulo	BRA	São Paulo	9968485
1890	Shanghai	CHN	Shanghai	9696300
939	Jakarta	IDN	Jakarta Raya	9604900
2822	Karachi	PAK	Sindh	9269265
3357	Istanbul	TUR	Istanbul	8787958
2515	Ciudad de México	MEX	Distrito Federal	8591309
3580	Moscow	RUS	Moscow (City)	8389200
3793	New York	USA	New York	8008278
*	NULL	NULL	NULL	NULL



5. **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.



The screenshot displays a database query interface. At the top, a toolbar includes icons for file operations, query execution, and a 'Limit to 1000 rows' dropdown. Below the toolbar, a SQL query is entered in a text area:

```
1 • SELECT Name AS city, CountryCode, Population
2 FROM city
3 WHERE Population > 2000000
4 ORDER BY Population ASC;
```

Below the query editor, the 'Result Grid' tab is active, showing a table with the following data:

city	CountryCode	Population
Bucuresti	ROM	2016131
Luanda	AGO	2022000
Shijiazhuang	CHN	2041500
Jedda	SAU	2046300
Guayaquil	ECU	2070040
Cali	COL	2077386
Fortaleza	BRA	2097757
Zhengzhou	CHN	2107200
Toskent	UZB	2117500
Paris	FRA	2125246
Izmir	TUR	2130359
Belo Horizonte	BRA	2139125

On the right side of the interface, there is a sidebar with buttons for 'Result Grid', 'Form Editor', and 'Field Types'. At the bottom of the window, a tab labeled 'city 14' is visible, and a 'Read Only' status indicator is present.

6. **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.



Query 1 x world db (1) SQL File 4

Limit to 1000 rows

```

1 • SELECT Name AS City
2 FROM city
3 WHERE Name LIKE 'Be%' ;

```

Result Grid

City
Béjaia
Béchar
Benguela
Berazategui
Belize City
Belmopan
Belo Horizonte
Belém
Belford Roxo
Betim
Bento Gonçalves
Belfast

city 18 x Read Only

7. **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.

Query 1 x world db (1) SQL File 4

Limit to 1000 rows

```

1 • SELECT *
2 FROM city
3 WHERE Population BETWEEN 500000 AND 1000000
4 ORDER BY Population ASC;

```

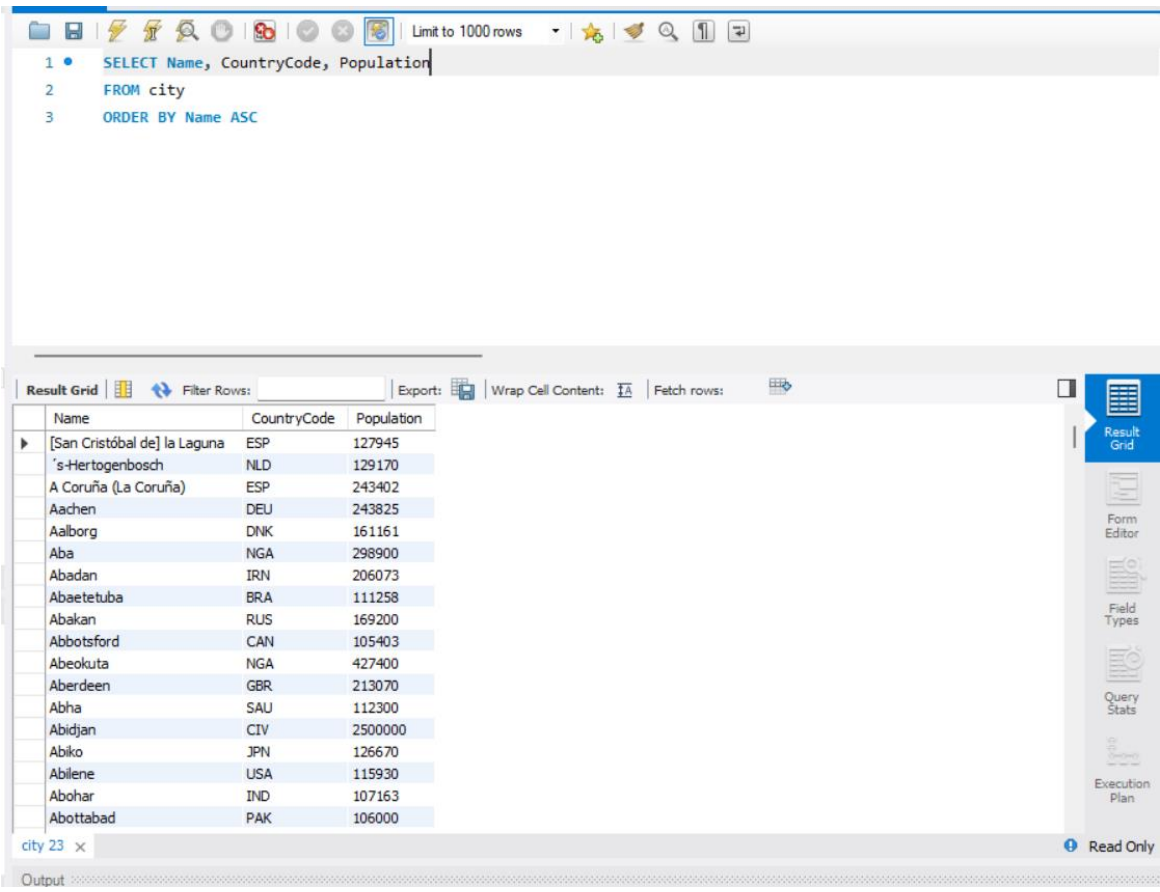
Result Grid

ID	Name	CountryCode	District	Population
2297	Pointe-Noire	COG	Kouilou	500000
3612	Tjumen	RUS	Tjumen	503400
1780	Sanaa	YEM	Sanaa	503600
1073	Chandigarh	IND	Chandigarh	504094
2490	Salé	MAR	Rabat-Salé-Zammour-Z	504420
771	Pasig	PHL	National Capital Reg	505058
1072	Gorakhpur	IND	Uttar Pradesh	505566
3611	Tula	RUS	Tula	506100
3821	Oklahoma City	USA	Oklahoma	506132
3252	Hims	SYR	Hims	507404
3434	Mykolajiv	UKR	Mykolajiv	508000
2807	Oslo	NOR	Oslo	508726

city 20 x Apply Revert



8. **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.



The screenshot shows a database query tool interface. The top section contains a SQL query editor with the following text:

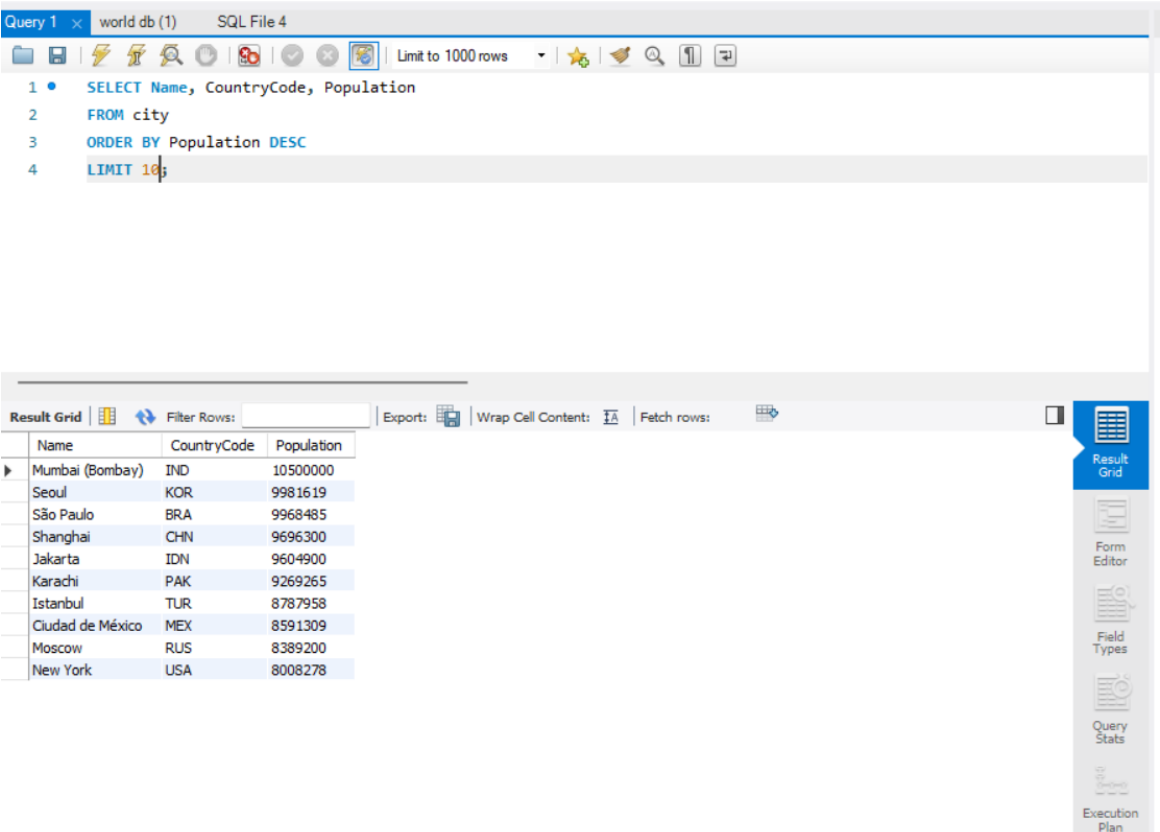
```
1 • SELECT Name, CountryCode, Population
2 FROM city
3 ORDER BY Name ASC
```

Below the query editor is a "Result Grid" displaying the results of the query. The grid has three columns: Name, CountryCode, and Population. The results are sorted alphabetically by city name. The first few rows are:

Name	CountryCode	Population
[San Cristóbal de] la Laguna	ESP	127945
's-Hertogenbosch	NLD	129170
A Coruña (La Coruña)	ESP	243402
Aachen	DEU	243825
Aalborg	DNK	161161
Aba	NGA	298900
Abadan	IRN	206073
Abaetetuba	BRA	111258
Abakan	RUS	169200
Abbotsford	CAN	105403
Abeokuta	NGA	427400
Aberdeen	GBR	213070
Abha	SAU	112300
Abidjan	CIV	2500000
Abiko	JPN	126670
Ablene	USA	115930
Abohar	IND	107163
Abottabad	PAK	106000

The interface also includes a toolbar with various icons for query execution, a "Filter Rows" field, and an "Export" button. On the right side, there is a sidebar with options for "Result Grid", "Form Editor", "Field Types", "Query Stats", and "Execution Plan". The bottom of the window shows a tab labeled "city 23" and an "Output" section.

9. **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.



The screenshot shows a SQL query editor window titled "Query 1" with a tab for "world db (1)" and "SQL File 4". The query is as follows:

```
1 • SELECT Name, CountryCode, Population
2 FROM city
3 ORDER BY Population DESC
4 LIMIT 10;
```

Below the query editor, the "Result Grid" displays the top 10 most populated cities. The grid has columns for Name, CountryCode, and Population. The results are as follows:

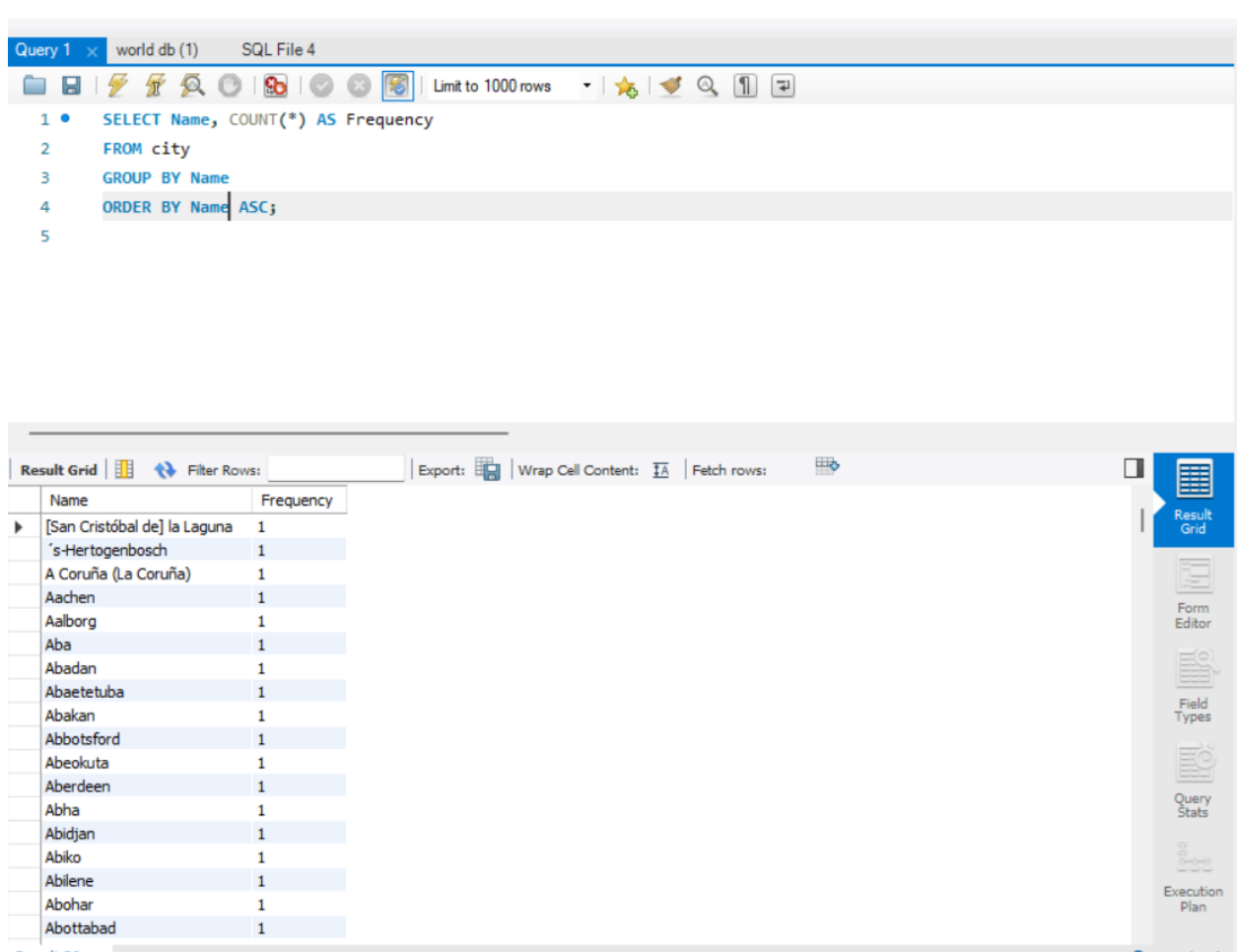
Name	CountryCode	Population
Mumbai (Bombay)	IND	10500000
Seoul	KOR	9981619
São Paulo	BRA	9968485
Shanghai	CHN	9696300
Jakarta	IDN	9604900
Karachi	PAK	9269265
Istanbul	TUR	8787958
Ciudad de México	MEX	8591309
Moscow	RUS	8389200
New York	USA	8008278

The interface includes a toolbar with icons for file operations, a "Limit to 1000 rows" dropdown, and a "Filter Rows" input field. The right sidebar contains buttons for "Result Grid", "Form Editor", "Field Types", "Query Stats", and "Execution Plan".

10. **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.



The screenshot shows a database query editor interface. The top pane displays an SQL query:

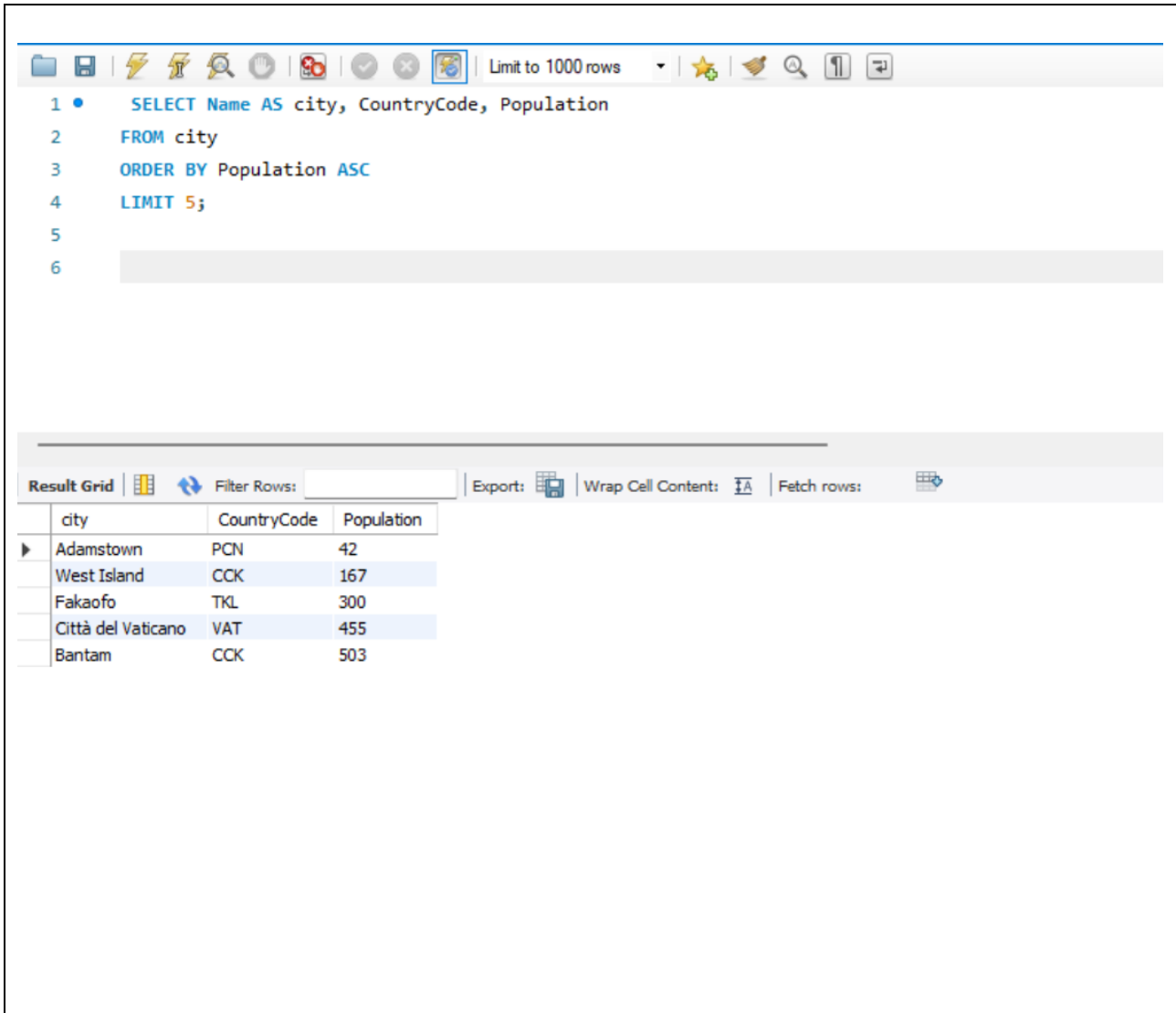
```
1 • SELECT Name, COUNT(*) AS Frequency
2   FROM city
3   GROUP BY Name
4   ORDER BY Name ASC;
5
```

The bottom pane shows the results of the query in a table format. The table has two columns: 'Name' and 'Frequency'. The results are sorted alphabetically by city name. The first few rows are:

Name	Frequency
[San Cristóbal de] la Laguna	1
's-Hertogenbosch	1
A Coruña (La Coruña)	1
Aachen	1
Aalborg	1
Aba	1
Abadan	1
Abaetetuba	1
Abakan	1
Abbotsford	1
Abeokuta	1
Aberdeen	1
Abha	1
Abidjan	1
Abiko	1
Abilene	1
Abohar	1
Abottabad	1

The interface includes a toolbar with various icons for file operations, a 'Limit to 1000 rows' dropdown, and a 'Result Grid' button on the right. The bottom status bar shows 'Result 29' and 'Read Only'.

11. **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.



The screenshot displays a database query interface. The top toolbar includes icons for file operations, a search icon, and a dropdown menu set to "Limit to 1000 rows". The SQL query editor contains the following code:

```
1 • SELECT Name AS city, CountryCode, Population
2 FROM city
3 ORDER BY Population ASC
4 LIMIT 5;
5
6
```

Below the query editor is the "Result Grid" section, which includes a "Filter Rows:" input field, an "Export:" button, a "Wrap Cell Content:" toggle, and a "Fetch rows:" button. The results are displayed in a table with three columns: city, CountryCode, and Population.

city	CountryCode	Population
Adamstown	PCN	42
West Island	CCK	167
Fakaofo	TKL	300
Città del Vaticano	VAT	455
Bantam	CCK	503

12. **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.

The screenshot shows a SQL query editor with the following code:

```
1 • SELECT Name AS countryName, Population
2 FROM country
3 ORDER BY Population DESC
4 LIMIT 5;
```

Below the query editor, the results are displayed in a table:

	countryName	Population
▶	China	1277558000
	India	1013662000
	United States	278357000
	Indonesia	212107000
	Brazil	170115000

	countryName	Population
▶	China	1277558000
	India	1013662000
	United States	278357000
	Indonesia	212107000
	Brazil	170115000

13. **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.

Altering file\* world db (1) SQL File 5\*

Limit to 1000 rows

```
1 • SELECT city.Name AS capital_city
2 FROM country
3 JOIN city ON country.Capital = city.ID
4 WHERE country.Name = 'Spain';
5
6
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	capital_city
▶	Madrid

14. **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.



[illegible]



Limit to 1000 rows

```

1 • SELECT country.Name AS country_name,
2     AVG(city.Population) AS avg_city_population
3 FROM city
4 JOIN country ON city.CountryCode = country.Code
5 GROUP BY country.Name
6 ORDER BY avg_city_population DESC;
7

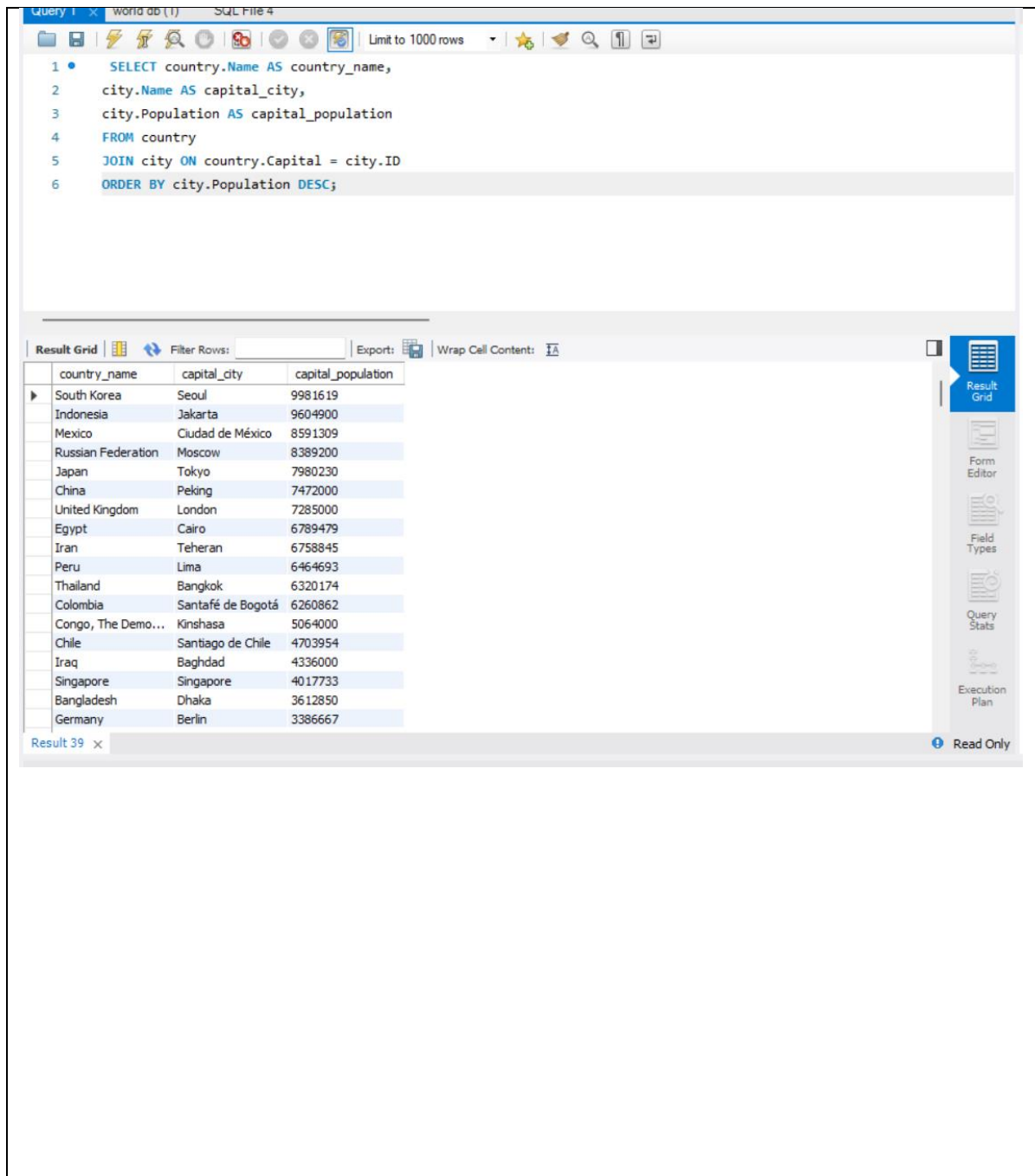
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

country_name	avg_city_population
Singapore	4017733.0000
Hong Kong	1650316.5000
Uruguay	1236000.0000
Guinea	1090610.0000
Uganda	890800.0000
Liberia	850000.0000
Sierra Leone	850000.0000
Mali	809552.0000
Australia	808119.0000
Mongolia	773700.0000
Congo	725000.0000
Libyan Arab J...	674251.7500
Lebanon	670000.0000
Thailand	662763.4167
Côte d'Ivoire	638227.4000
Azerbaijan	616000.0000
Iraq	595069.4000
Afghanistan	583025.0000

Result 38 x Read Only

16. **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.



The screenshot shows a SQL query editor with the following query:

```

1 • SELECT country.Name AS country_name,
2     city.Name AS capital_city,
3     city.Population AS capital_population
4 FROM country
5 JOIN city ON country.Capital = city.ID
6 ORDER BY city.Population DESC;

```

The result grid displays the following data:

country_name	capital_city	capital_population
South Korea	Seoul	9981619
Indonesia	Jakarta	9604900
Mexico	Ciudad de México	8591309
Russian Federation	Moscow	8389200
Japan	Tokyo	7980230
China	Peking	7472000
United Kingdom	London	7285000
Egypt	Cairo	6789479
Iran	Teheran	6758845
Peru	Lima	6464693
Thailand	Bangkok	6320174
Colombia	Santafé de Bogotá	6260862
Congo, The Demo...	Kinshasa	5064000
Chile	Santiago de Chile	4703954
Iraq	Baghdad	4336000
Singapore	Singapore	4017733
Bangladesh	Dhaka	3612850
Germany	Berlin	3386667

17. **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.



Query 1

world db (1)

SQL File 4

SQLAddit

Limit to 1000 rows

```

1 SELECT Name AS Country_Name, Population , SurfaceArea , (Population / SurfaceArea) AS PopulationDensity
2 FROM country
3 WHERE SurfaceArea > 0
4 ORDER BY PopulationDensity ASC;
5
6
7

```

Auto  
disab  
man  
curre  
tog

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Country_Name	Population	SurfaceArea	PopulationDensity
Antarctica	0	13120000.00	0.0000
French Southern territories	0	7780.00	0.0000
Bouvet Island	0	59.00	0.0000
Heard Island and McDonald Islands	0	359.00	0.0000
British Indian Ocean Territory	0	78.00	0.0000
South Georgia and the South Sandwich Islands	0	3903.00	0.0000
United States Minor Outlying Islands	0	16.00	0.0000
Greenland	56000	2166090.00	0.0259
Svalbard and Jan Mayen	3200	62422.00	0.0513
Falkland Islands	2000	12173.00	0.1643
Pitcairn	50	49.00	1.0204
Western Sahara	293000	266000.00	1.1015
Mongolia	2662000	1566500.00	1.6993
French Guiana	181000	90000.00	2.0111
Namibia	1726000	824292.00	2.0939
Australia	18886000	7741220.00	2.4397
Suriname	417000	163265.00	2.5541
Mauritania	2670000	1025520.00	2.6036

Result Grid

Form Editor

Field Types

Query Stats

Execution Plan

Result 41

Read Only

Context H

Output

18. **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  SELECT
2      ci.Name AS city_name,
3      co.Name AS country_name,
4      co.GNP AS country_gdp,
5      co.Population AS country_population,
6      (co.GNP / co.Population) AS gdp_per_capita
7  FROM
8      city ci
9  JOIN country co ON ci.CountryCode = co.Code
10 WHERE co.Population > 0 AND
11      (co.GNP / co.Population) > (
12          SELECT AVG(GNP / Population)
13          FROM country
14          WHERE Population > 0
15      )
16 ORDER BY
17     gdp_per_capita DESC;

```

city_name	country_name	country_gdp	country_population	gdp_per_capita
Luxembourg [Luxemburg/Lëtzebuerg]	Luxembourg	16321.00	435700	0.037459
Zürich	Switzerland	264478.00	7160400	0.036936
Geneve	Switzerland	264478.00	7160400	0.036936
Basel	Switzerland	264478.00	7160400	0.036936
Bern	Switzerland	264478.00	7160400	0.036936
Lausanne	Switzerland	264478.00	7160400	0.036936
Saint George	Bermuda	2328.00	65000	0.035815
Hamilton	Bermuda	2328.00	65000	0.035815
Bandar Seri Begawan	Brunei	11705.00	328000	0.035686
Schaan	Liechtenstein	1119.00	32300	0.034644

19. **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.



[illegible]

--



## 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:



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.**

