|  |  |
| --- | --- |
|  | Abstract  SQL data analysis involves querying and examining data stored in relational databases using Structured Query Language (SQL). Analysts retrieve specific data subsets, perform transformations, and aggregate information using SQL functions. Filtering criteria are applied to focus on relevant data subsets for analysis. Various analytical tasks, such as trend analysis and anomaly detection, are conducted to extract insights. |

Jaya Sai NagaLakshmi Katrela

Table of Contents

Introduction to SQL

1. What is SQL
2. What SQL can do
3. Data Types
4. Relational Database, Constrains, Primary & Foreign key
5. Schema types, SQL Languages

Assignment-1

1. Tasks 1-6
2. Screen shorts

**EER Diagram**

EER diagram explanation

**Assignment-2**

1. Tasks 1-11
2. Screen shorts

**EER Diagram**

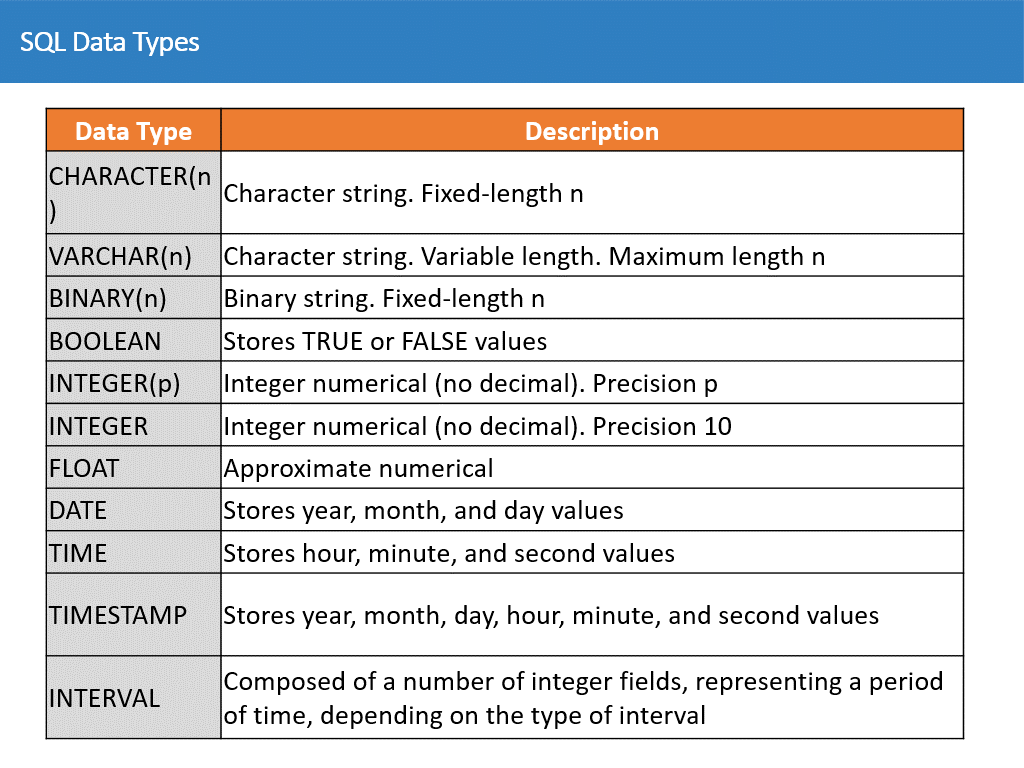
**Introduction to SQL**

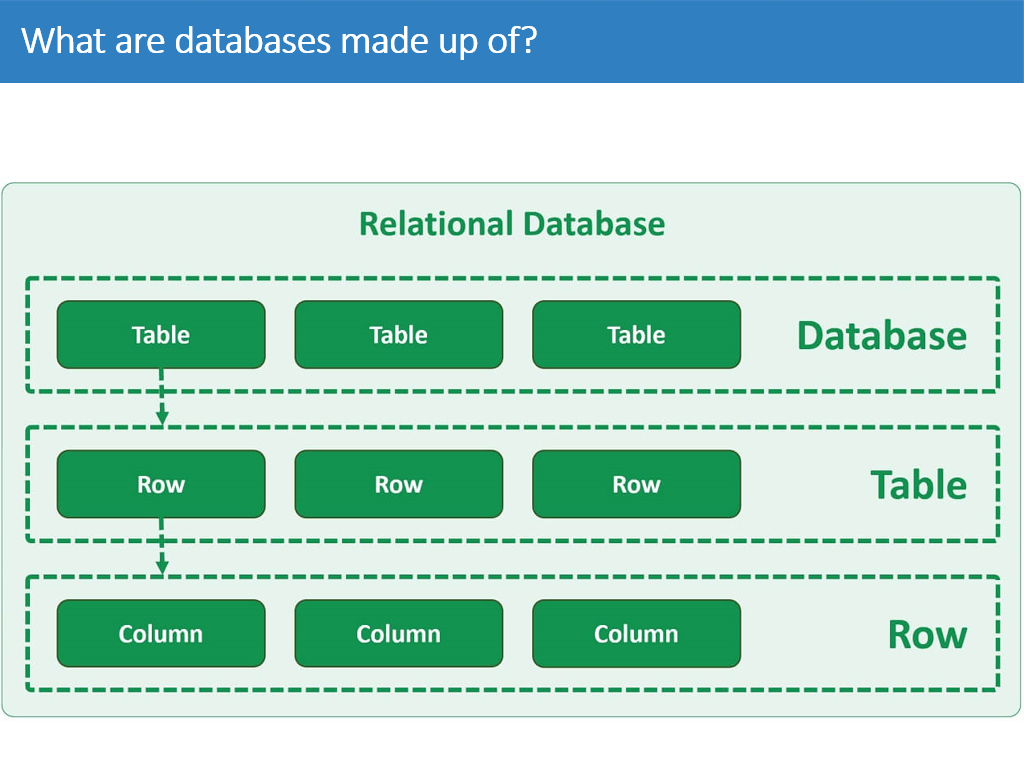
**What is SQL**

Structured Query Language, is a specialized programming language designed for managing and manipulating data stored in relational databases. It serves as the standard language for interacting with databases and is widely used across various industries and applications.

**What Can SQL do**

SQL allows users to perform a multitude of tasks, including retrieving specific data subsets from databases, inserting new records, updating existing data, and deleting unwanted information. Additionally, SQL enables the creation and modification of database structures, such as tables, indexes, views, and stored procedures. Its intuitive syntax and powerful capabilities make it an essential tool for database administrators, developers, data analysts, and anyone involved in data management and analysis tasks.





**Constraints** are limitations or rules placed on a field or column to ensure that data that is considered invalid is not entered.

**SQL PRIMARY KEY Constraint**

**•**The PRIMARY KEY constraint uniquely identifies each record in a table.

•Primary keys must contain UNIQUE values and cannot contain NULL values.

•A table can have only ONE primary key; and in the table, this primary key can consist of single or multiple columns (fields).

**SQL FOREIGN KEY Constraint​**

The FOREIGN KEY constraint is used to prevent actions that would destroy links between tables.​

A FOREIGN KEY is a field (or collection of fields) in one table, that refers to the PRIMARY KEY in another table.​

The table with the foreign key is called the child table, and the table with the primary key is called the referenced or parent table​.

**Schema Types**

**1.Conceptual Schema:** offer a big-picture view of what the system will contain, how it will be organized, and which business rules are involved.

**2.Star Schema:** Are less abstract, compared to conceptual schemas. They clearly define schema objects with information, such as table names, field names, entity relationships, and integrity constraints.

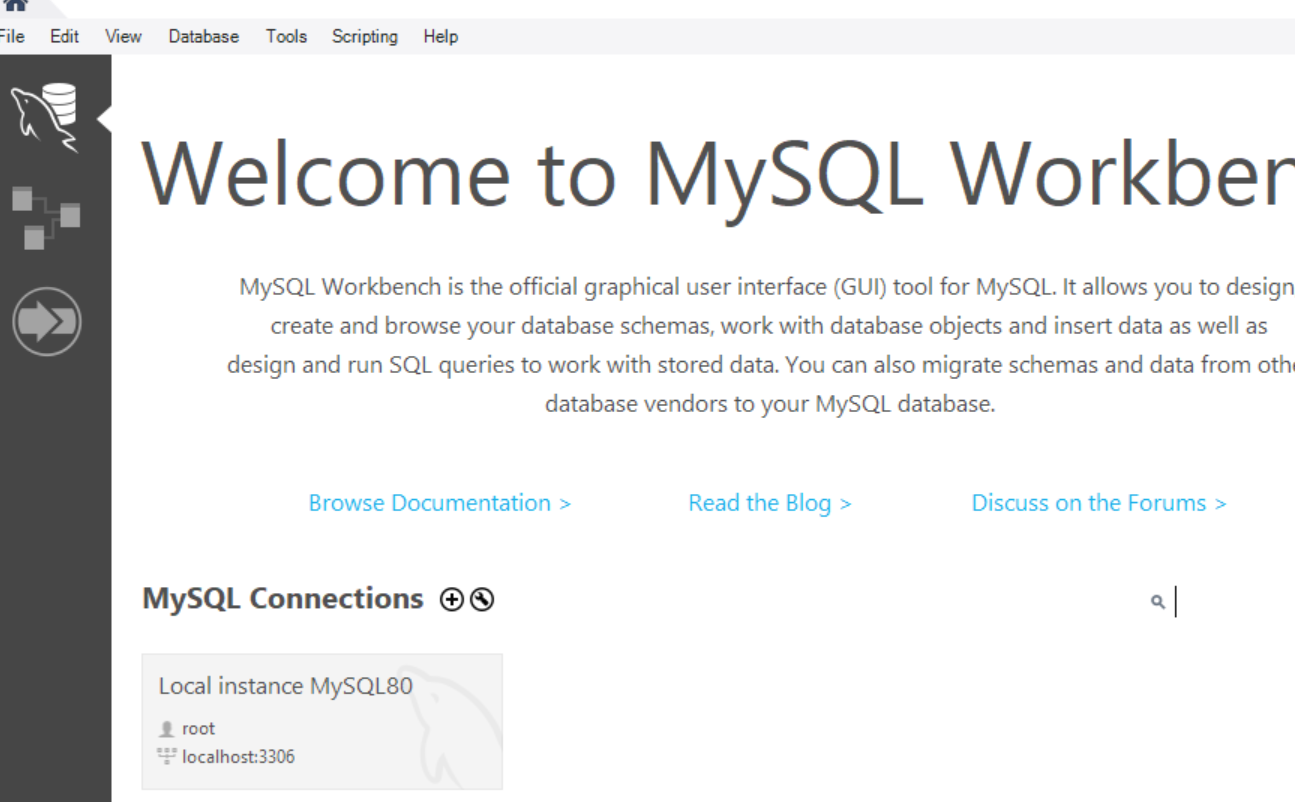
3.**Snowflake Schema:** In SQL, a snowflake schema is a database schema design where dimensional tables are normalized into multiple related tables, resembling the shape of a snowflake. It consists of a central fact table surrounded by multiple dimension tables, each of which may be further divided into sub-dimension tables.

**SQL Languages**

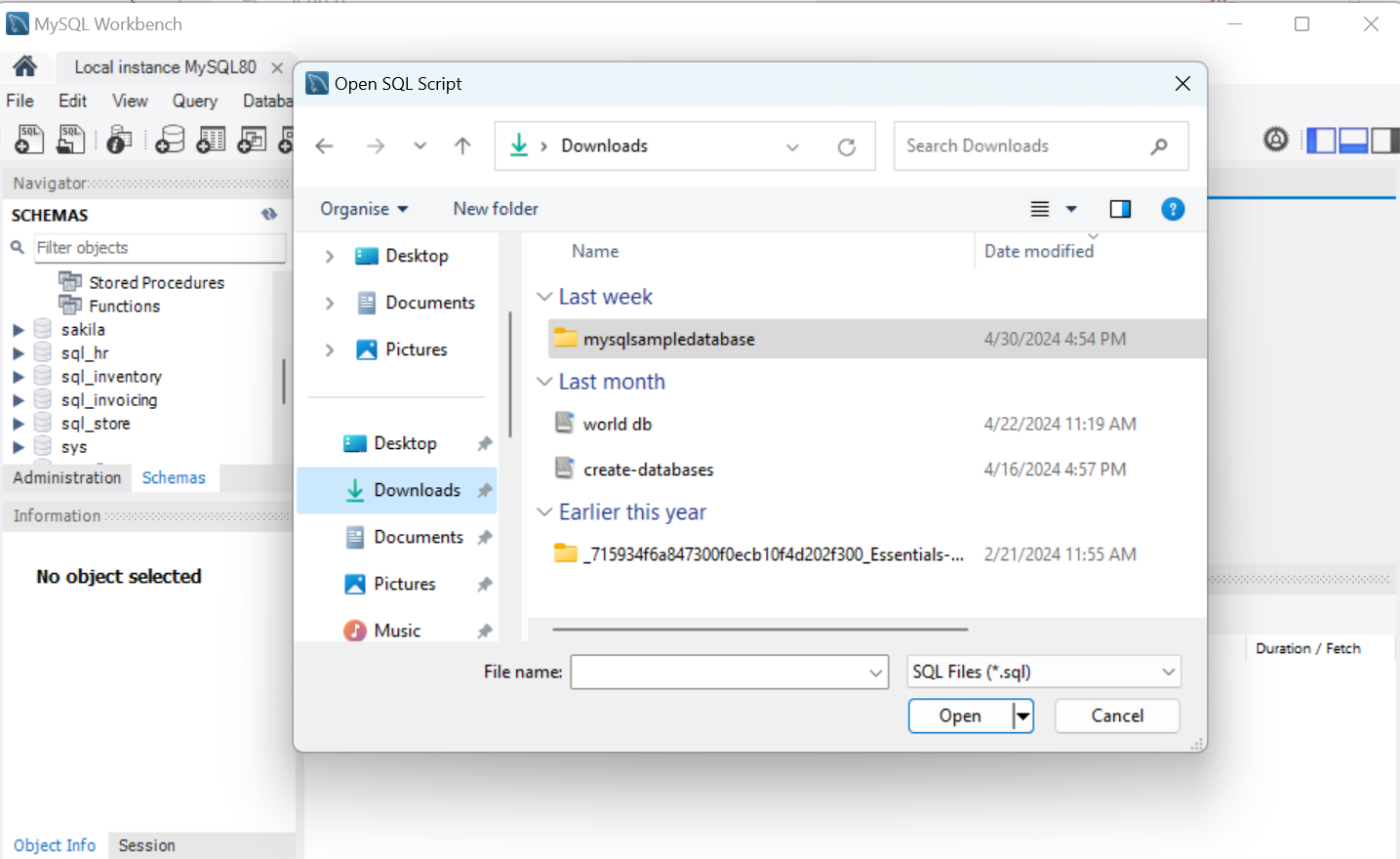
1. **DDL:** create, drop, alter, truncate, comment, rename.
2. **DML:** insert, update, delete, lock.
3. **TCL:** commit, rollback, save point, set transaction.
4. **DCL:** grant, revoke.

**Assignment-1**

**Import and Execute SQL Script**

****

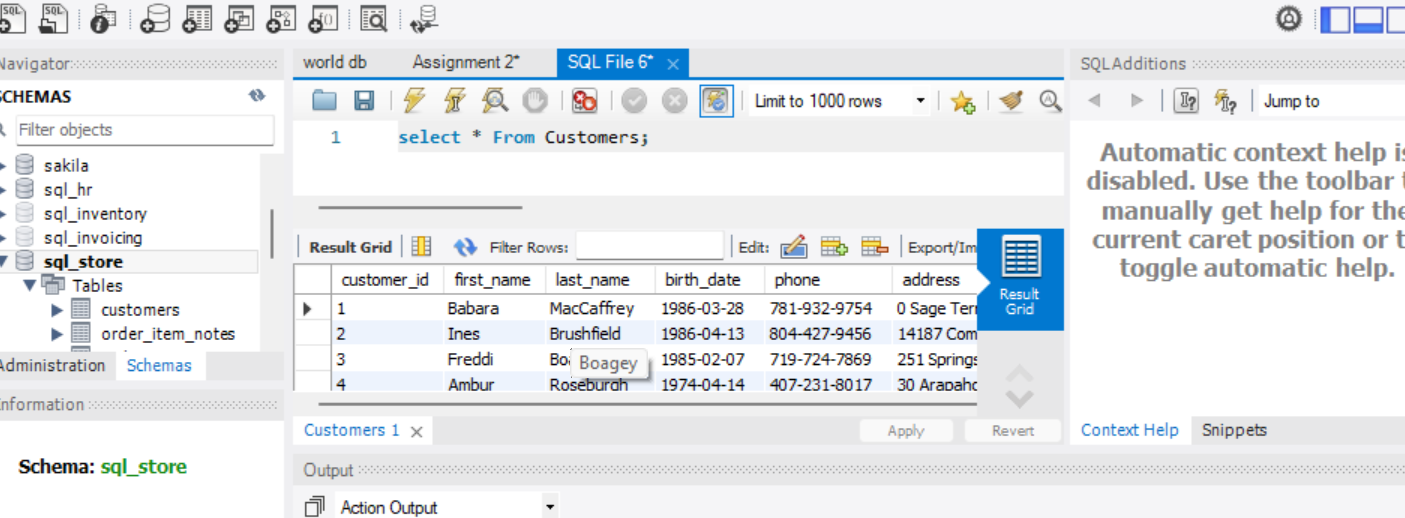
**Open the My SQL Workbench and login to the local instance.**

****

First download the files go to my SQL workbench navigate to files click on open SQL scripts navigate to the downloaded files and click open and run the script.

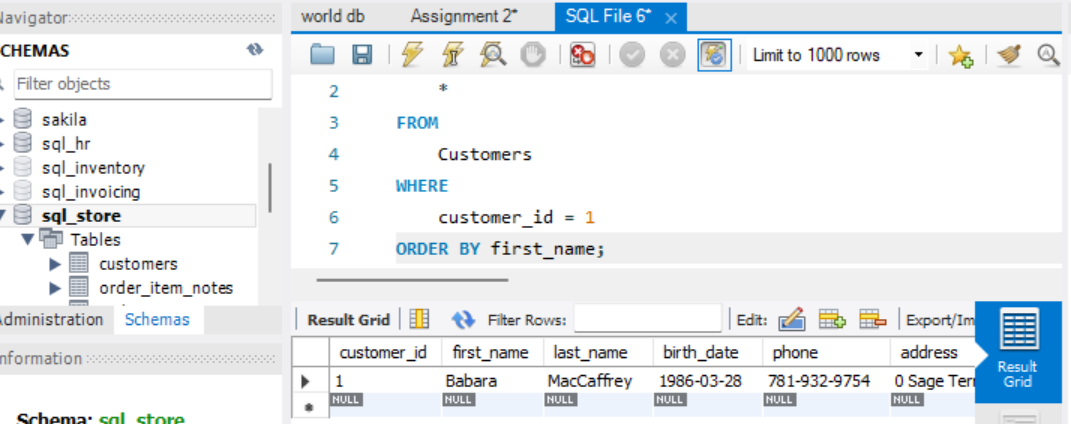
View the Customers Table using SELECT command.

**Query-1**



Once the schema is ready either double click on the schemas or use command “**use sql\_store**”.

Using SELECT \* From customers; command we can view all the data in customer table.



Here we are giving the condition using Where command to choose the data which satisfy the condition and to display it and order it based on the first\_name.

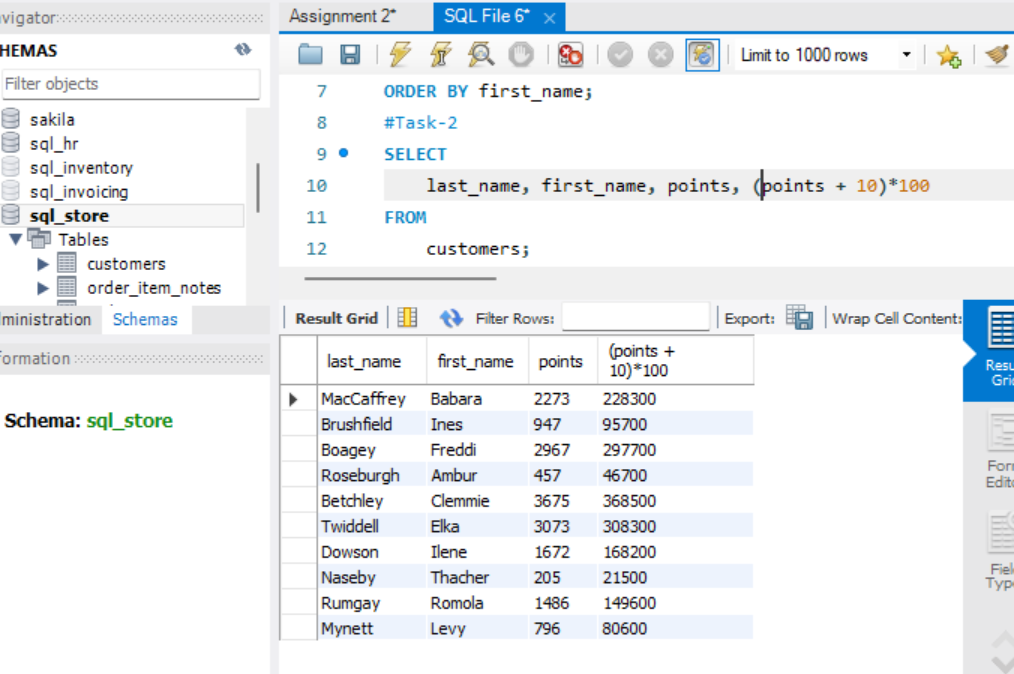
**Query-2**

**A screenshot of a computer

Description automatically generated**

Here we have displayed the selected columns and done the addition for points column and display it in separate column.

We have multiplied the (poins+10)\*100

****

**Task-1**

**A screenshot of a computer

Description automatically generated**

**Here we have changed the (points+10)\*100 column name as discount\_factor using alias.**

**Task-2**

**A screenshot of a computer

Description automatically generated**

We have selected name, unit\_price, and increase the value of unit\_price by 10% and named it as new\_price using alias.

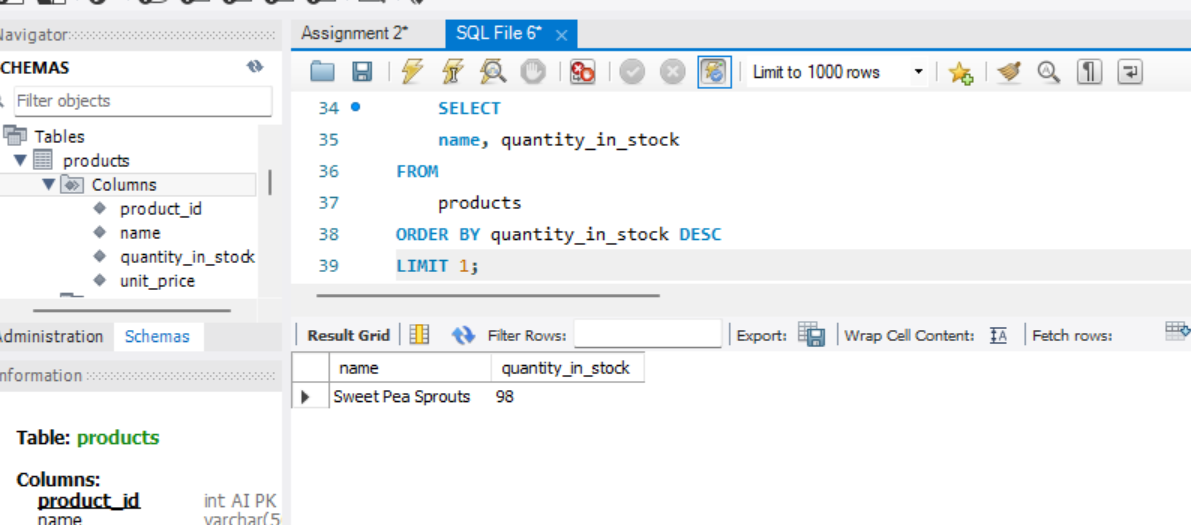
**Task-3**

**A screenshot of a computer

Description automatically generated**

We have displayed the customers whose date of birth is >1991-01-01 using where command.

**Task-4**

****

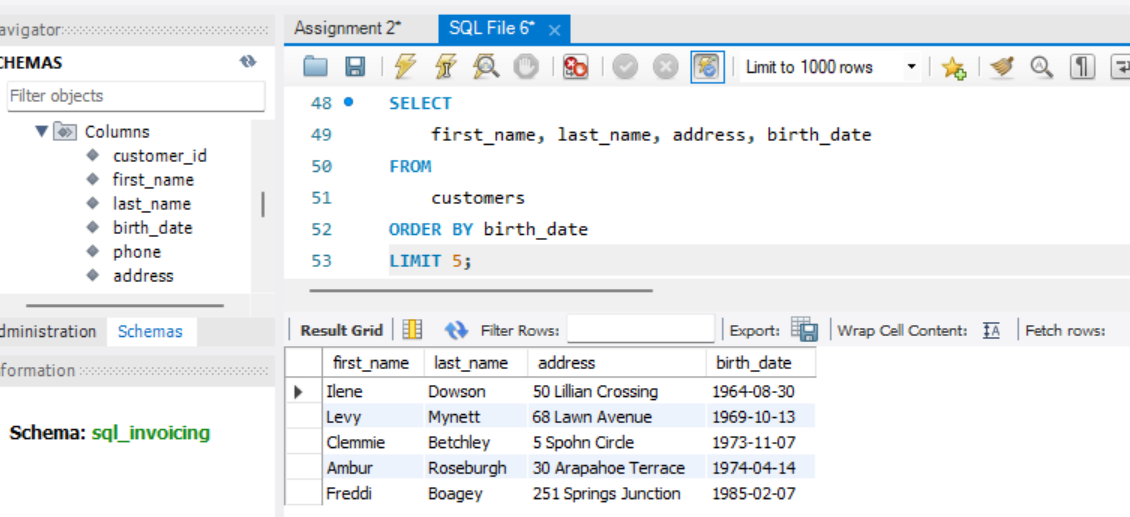
Change the schema to sql\_inventory from products table display the name of the product which has highest amount of stock and display only the top 1 product name.

**Task-5**

****

Here we have listed the name of the product which has highest price using order by and limit command.

**Task-6**

****

Here we have selected the oldest customers first name, last name, address, date of birth.

**Creating a EER Diagram**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A computer screen shot of a computer

Description automatically generated

Here we have created the EER diagram for sql\_store schema which contains table (customers, orders, products, order\_items, shippers, order\_statuses).

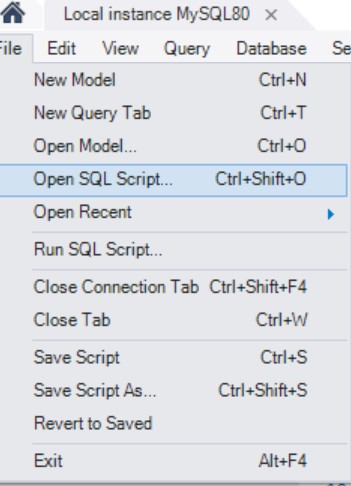
Each table tis linked to orders table to with primary key and foreign key, order items is linked to products to check weather the product is available or not and how many are in stock etc.

Once the order is place there will be a set of data under the order like order id, status, customer details.

From the order table to view the customer details based on customers\_id as primary key from the customer table and customers\_id as foreign key in orders table we can retrieve the customer details like wise we can retrieve the details of the order items, order status, shipping\_id etc.

**Assignment-2**

Import & execute worlddb file to my SQL workbench.



A screenshot of a computer

Description automatically generated

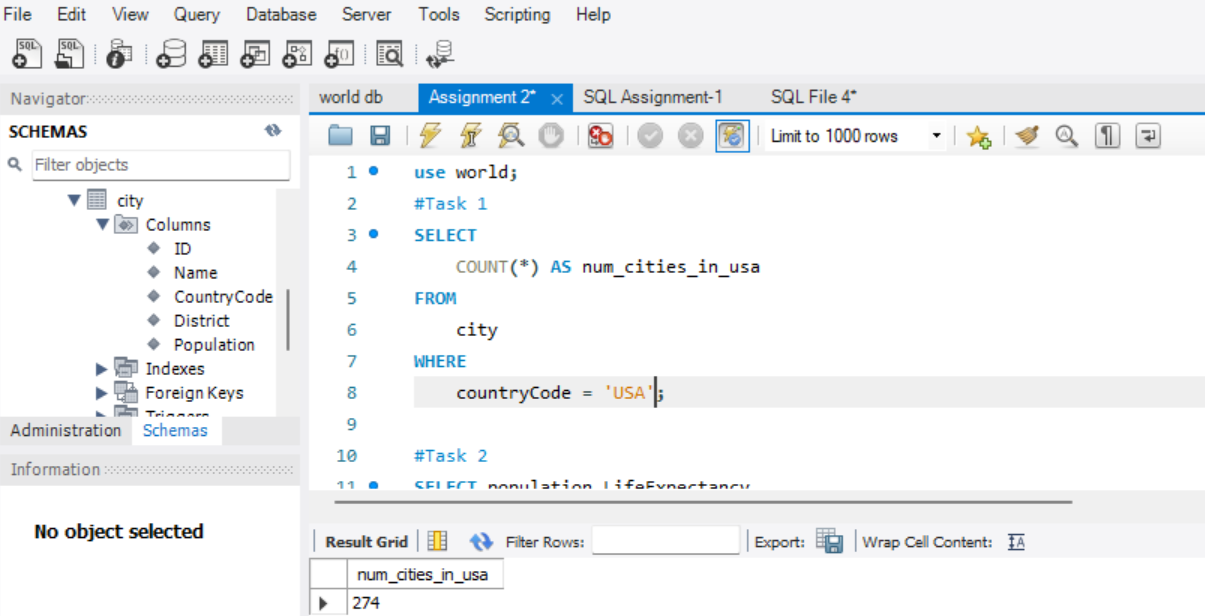
Click on file and select open SQL script the code will appear in the query tab and run the script. Once the script was executed correctly it will appear as shown in below image.

A screenshot of a computer

Description automatically generated

**Task-1**

**Count the countries in USA using count function**

****

Here we have used the count function to count the number of cities using where given the condition to count only for the country USA.

**Task-2**

**To find the population and life expectancy for people in Argentina**

A screenshot of a computer

Description automatically generated

We have selected the population and life expectancy column from country table to see the Argentina’s population and life expectancy.

**Task-3**

**Find out the highest life expectancy of country**

**A screenshot of a computer

Description automatically generated**

Select the name, life expectancy columns using order by for life expectancy in descending order and limit the result to top 1 country.

**Task-4**

**Select 25 cities around the world starting with letter F**

**A screenshot of a computer

Description automatically generated**

We have selected names from city table by using where condition to select only the city that start with letter ‘F’

**Task-5**

**Display the 10 ID, Name, Population columns from city table**

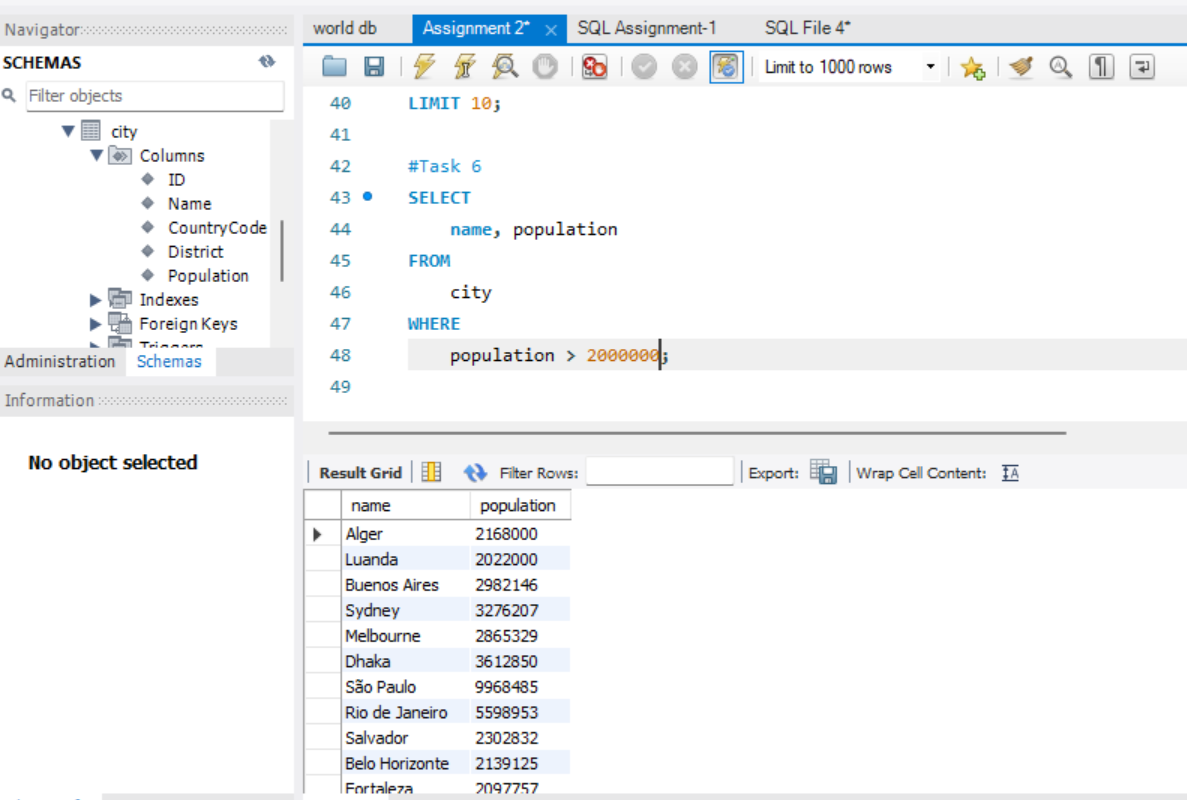
**A screenshot of a computer

Description automatically generated**

Using select statement selected id, name, population columns from city table and limit to first 10.

**Task-6**

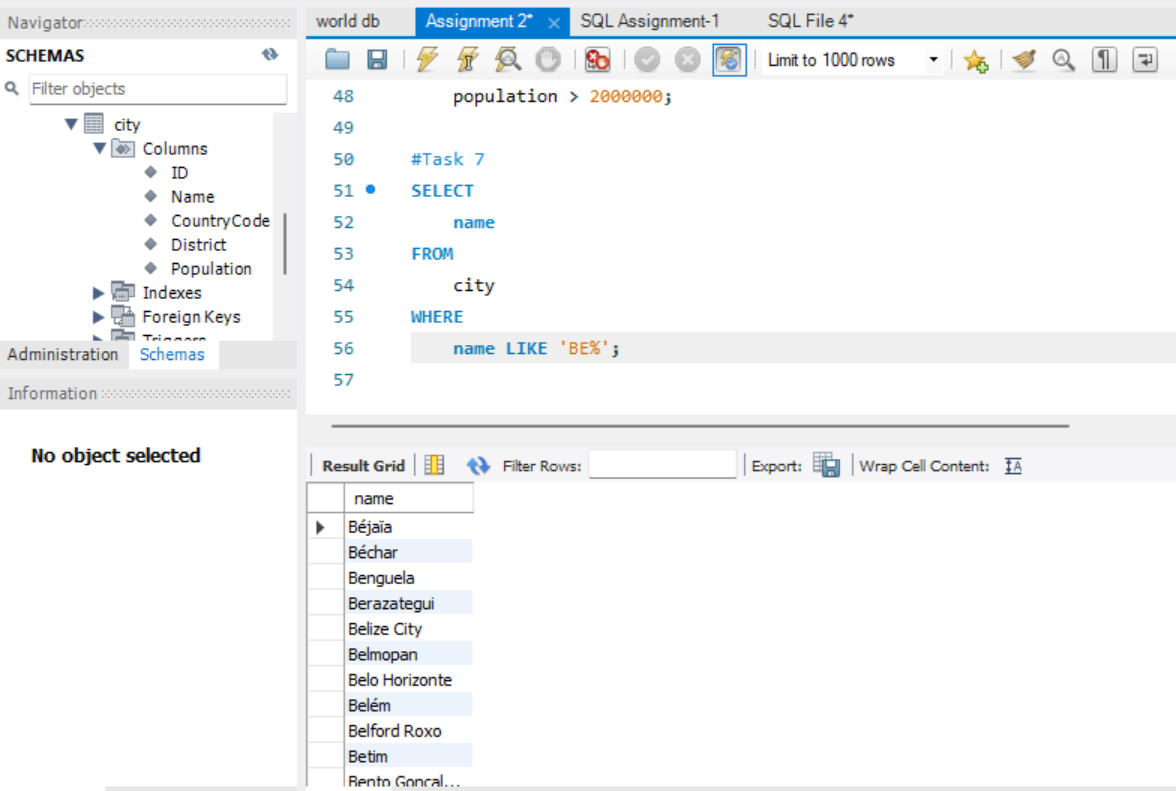
**Find the cities whose population is greater than 2000000 from city table**

****

Select the name, population from city table using where condition to select only the cities whose population is greater than 2000000.

**Task-7**

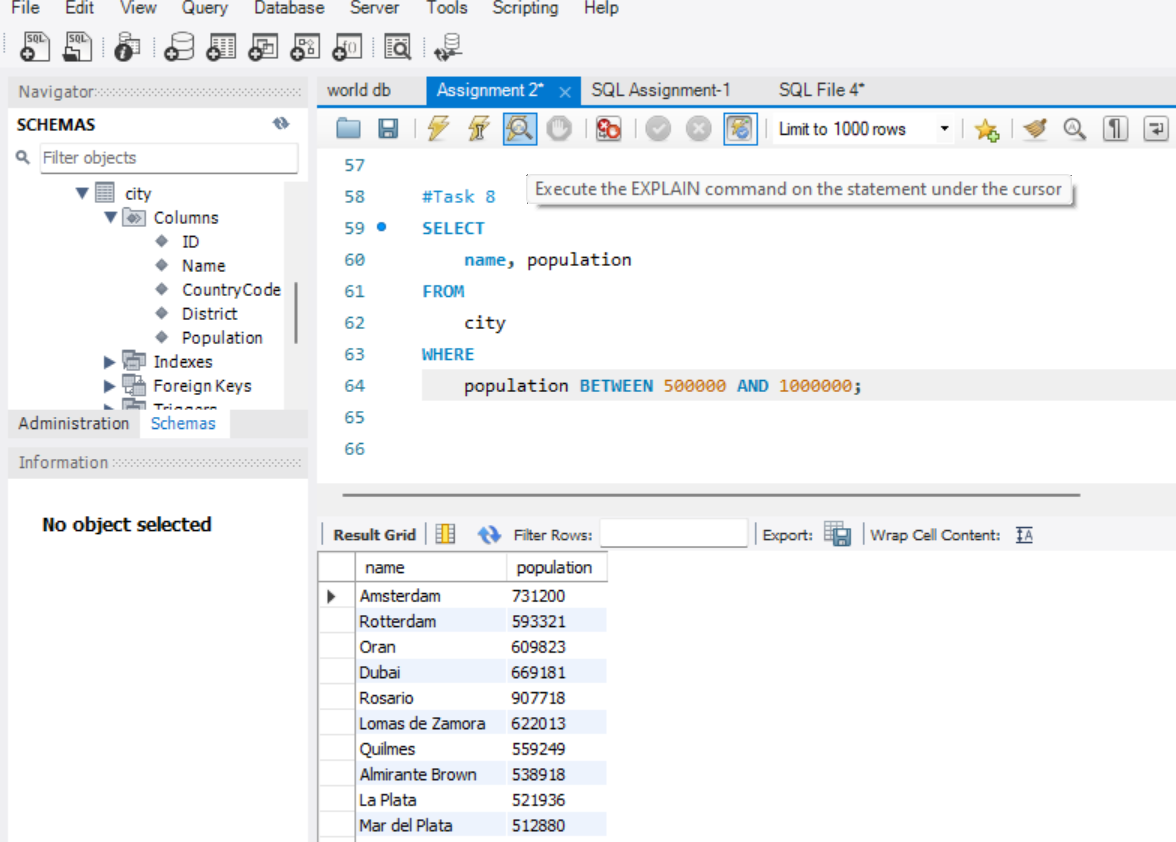
**Find the cities from city table city name that starts with BE**

****

Select the name from city table where the name starts with BE

**Task-8**

**Find the cities whose population is between 500000-1000000**

****

Selecting the name and population columns from city table whose population is between 500000-1000000 using where condition.

**Task-9**

**Find the cities with lowest population**

**A screenshot of a computer

Description automatically generated**

Select the name, population and order it by population from lowest to highest.

**Task-10**

**Display all cities sort by names in ascending order**

**A screenshot of a computer

Description automatically generated**

Selecting the name and displaying it in ascending order.

**Task-11**

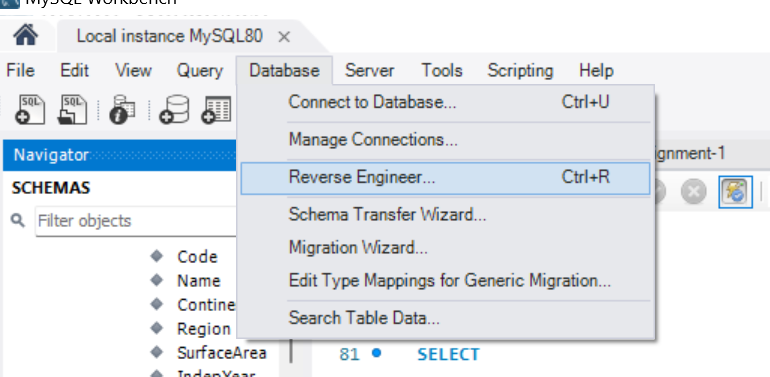
**Find the country with largest population**

**A screenshot of a computer

Description automatically generated**

Select name, population order by population from highest to lowest and limit to top 10 countries.

**Creating EER Diagram**

****

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer screen

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**