World Database

Mohamed Mohamed

Table of Contents

[Creating a database 2](#_Toc171432572)

[Understanding the Business Requirements 2](#_Toc171432573)

[Designing the Database Schema 3](#_Toc171432574)

[Implementing the Database 4](#_Toc171432575)

[Populating the Database 5](#_Toc171432576)

[Maintaining the Database 6](#_Toc171432577)

[SQL Processing 7](#_Toc171432578)

[1. Count Cities in USA 7](#_Toc171432579)

[2. Country with Highest Life Expectancy 7](#_Toc171432580)

[*3.* New Year Promotion: Featuring Cities with 'New’ 8](#_Toc171432581)

[*4.* Display Columns with Limit (First 10 Rows) 9](#_Toc171432582)

[*5.* Cities with Population Larger than 2,000,000: 10](#_Toc171432583)

[*6.* Cities Beginning with 'Be' Prefix 11](#_Toc171432584)

[*7.* Cities with Population Between 500,000-1,000,000: 12](#_Toc171432585)

[*8.* Display Cities Sorted by Name in Ascending Order: 13](#_Toc171432586)

[9. Most Populated City: 15](#_Toc171432587)

[10. City Name Frequency Analysis: Supporting Geography Education 16](#_Toc171432588)

[11. City with the Lowest Population 17](#_Toc171432589)

[12. Country with Largest Population 17](#_Toc171432590)

[*13.* Capital of Spain 18](#_Toc171432591)

[14. Country with Highest Life Expectancy 18](#_Toc171432592)

[15. Cities in Europe 19](#_Toc171432593)

[16. Average Population by Country 20](#_Toc171432594)

[17. Capital Cities Population Comparison 21](#_Toc171432595)

[18. Countries with Low Population Density 22](#_Toc171432596)

[19. Cities with High GDP per Capita 23](#_Toc171432597)

[20. Display Columns with Limit (Rows 31-40) 24](#_Toc171432598)

[Summary 25](#_Toc171432599)

# Creating a database

**Scope:**

*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.*

## Understanding the Business Requirements

To streamline operations for this business though a database system, it is important to identify and structure the key data the database will need to store. Considering the scope of managing inventory, sales, and customer information, as well as incorporating a loyalty program, the following types of data will be necessary:

* Inventory Data
* Sales Data
* Customer Data
* Supplier Data
* Promotion Data

Each user group will have different needs to access the database. The following are users of their databases and have different needs:

* Store owner/management team

The store owner or management team will need to access sales report, they will need ability to view and manage inventory data in real time, access customer data and loyalty program this will provide insight on how effective the promotions and discounts are.

* Cashiers

The task of cashiers will need to process sales transactions, update inventory and enrol customers in the loyalty program. As a result of these tasks, they will need to access the database to search for product information and availability, look at customer information and loyalty points.

* Inventory managers

The task of inventory managers will be to monitor and update stock levels. Receive and record any new stock shipments. With these responsibilities they will require the database to use real-time inventory tracking tools, alerts for anything that is low in stock and easy input interface to record and track new stock requests.

* Customer Service Reps

Customer service reps will handle customer inquiries and complaints, along with cashiers handle refunds and exchanges, and provide information on loyalty program and customer account status. With that, they will need the database to access detailed customer reports, like personal details, purchasing history and loyalty programme. Tools for exchange and refunds and enrolling customers into the loyalty program.

* Marketing personnels (for promotion)

Depending on the size of the business, the marketing personnels could be a large team or a small group, regardless their tasks would involve to design and create promotional campaign, track the effectiveness of it, analyse the data that comes with it and manage loyalty program incentives. They will need to access the database to look at sales during promotion periods. Observe purchasing behaviour data and customer demographic. They will need tools to set up and manage promotion within the system. Also need to have reporting tools to measure the success of marketing initiatives.

## Designing the Database Schema

A schema is the blueprint of the database, similar to a floorplan of an apartment. This will provide information on what is found in each table and how the relationships are formed between them. This will allow processing queries on software like SQL. As mentioned before the database will consist of the following data:

* Inventory Data: A table that contains information about the product, the primary key will be (product\_id). Information like product name, category, supplier information, product price and selling price.
* Sales Data: This will be a table consisting of information about transaction details and information. The primary key will be (transaction\_id) and will contain collects that includes date and time of purchase, quantity sold and total price.
* Customer Data: This will be a table that will have the primary key as (customer\_id) and consist of personal and sensitive information of customers like name, address and possibly bank details.
* Supplier Data: This will be a table consisting information about the supplier and supplier\_id will be the primary key and have information about their name and address and contact details.
* Promotion Data: This will store information about past and present promotions and promotion\_id will be the primary key.

The relationships between the tables will be as follows:

* Product from the inventory data will be found in the sales data
* Sales data information will be found in the customer data
* Sales data information will be associated with supplier data and promotion data.

Primary keys will be in different tables as foreign keys allowing relationships to form, to help streamline queries when analysing the database.

## Implementing the Database

SQL commands provide a foundational structure for the database, this setup allows for efficient data storage, retrieval, and management to support the business's operational needs.

To create a database the following commands are to be used:

CREATE DATABASE ‘Name\_Of\_Database’

In the example it will look like the following:



To create the tables that is within the SQL command to use will be:

CREATE DATABASE ‘Name\_Of\_Database’

USE ‘‘Name\_Of\_Database’

CREATE TABLE ‘Title\_Of\_Table’ (

Example:





## Populating the Database

To populate the database, which refers to completing it with columns and values in each of them, a number of things needs to be considered.

For example, the name of the column, the data type and the values that needs to be inserted.

The SQL command for a 4 columns in a table named Students will look like this

CREATE TABLE ‘Student Data’ (

Student\_id, **INTEGER**

First\_name, **TEXT**

Names of columns

Last\_name, **TEXT**

Age, **INTEGER**

Date\_enrolled**, DATE**);

The blue highlighted words represents the data types for each column.

|  |  |
| --- | --- |
| Data Type | Description |
| INTEGER | The integer datatypes can store whole integer values like the count of a number or an age. |
| TEXT | The text based datatypes can store strings and text in all sorts of locales. |
| DATE | Data type to store date stamp to keep track of events that take place. The same is applied for time also. |

Entering this command, a table with the column headings will be created, which looks like this:

**Student Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Student\_id | First\_name | Last\_name | Age | Date\_enrolled |

To populate this database, values needed to be added to it. The following SQL command can be used:

For the first record:

INSERT INTO ‘Student Data’

VALUES (1, ‘Mohamed’, ‘Dahir’, ‘16’, ‘05/07/2024’)

(2, ‘John’, ‘Douglas’, ‘15’, ‘05/07/2024’);

When the SQL command is run, the following table will be updated as follows

**Student Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Student\_id | First\_name | Last\_name | Age | Date\_enrolled |
| 1 | Mohamed | Dahir | 16 | 05/07/2024 |
| 2 | John | Douglas | 15 | 05/07/2024 |

## Maintaining the Database

To ensure the database remains accurate and up-to-date. Some key measures needs to be implemented to ensure the integrity of the database is upheld.

Using primary and foreign keys to ensure relationships between tables are maintained. Apply constraints to them also, to ensure that the data entry is valid for example “NOT NULL” and to avoid repeated entry, the record must be unique and using primary and foreign keys ensures that. To keep the data consistent, use appropriate data types for example in the first\_name column use VARCHAR (100) as opposed to TEXT or CHAR to ensure spaces aren’t used to fill the gaps.

Schedule regular backups to prevent loss and quick recovery in events of data being lost.

Security measures need to be implement to ensure GDPR policy is followed and to avoid breach and limit access to authorised personnel only. So security measures will involve user access controls, data encryption and audit logs.

Data entry processes will need to established to ensure maintenance and updating the database. To reduce risk and errors, validation rules have to applied and also batch processing to import large amount of data.

There must be documentation on how to appropriately use and interact with the database and training programs to help continuously develop employees.

# SQL Processing

In this part of the assignment, I will take screenshots of the SQL command I have used from mySQL and take screenshots of the query made. Along with that, I will explain how the query could be potentially used in context.

1. Count Cities in USA

A close-up of a number of words

Description automatically generated

Query outcome



This information could be used as a baseline for further analysis.

1. Country with Highest Life Expectancy

A close-up of words

Description automatically generated

Query outcome



This information will be crucial for prioritizing healthcare resources and interventions.

1. New Year Promotion: Featuring Cities with 'New’

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.

A white background with blue text

Description automatically generated

Query outcome

A screenshot of a computer

Description automatically generated

Point of clarification in the SQL command, a gap was left after New, to avoid showing cities with the word new appearing in the middle of the world. If there was no gap the SQL command and query outcome would look like the following:

A white background with blue text

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Display Columns with Limit (First 10 Rows)

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.

A white background with blue text

Description automatically generated

Query outcome

A screenshot of a list of cities

Description automatically generated

1. Cities with Population Larger than 2,000,000:

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. Cities Beginning with 'Be' Prefix

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.

A close-up of a sign

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. Cities with Population Between 500,000-1,000,000:

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.

A close-up of a computer code

Description automatically generated

Snippet of query outcome:

A screenshot of a computer

Description automatically generated

1. Display Cities Sorted by Name in Ascending Order:

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

Point of clarification, the first two names on this list don’t comply with the idea that the list of cities are in alphabetical order. This is a result of an error in data entry and it can be solved in a number of ways.

The rows could be edited as follows:

A screenshot of a computer code

Description automatically generated

Which would cause the query when run to produce this outcome

A screenshot of a computer

Description automatically generated

1. Most Populated City:

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.

A white background with blue text

Description automatically generated

Snippet of query outcome



1. City Name Frequency Analysis: Supporting Geography Education

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's lesson.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. City with the Lowest Population

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A close-up of a white background

Description automatically generated

1. Country with Largest Population

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.

A white background with blue text

Description automatically generated

Snippet of query outcome



1. Capital of Spain

A travel agency is organizing 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.

A screenshot of a computer code

Description automatically generated

Snippet of query outcome



1. Country with Highest Life Expectancy

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.

A close-up of words

Description automatically generated

Snippet of query outcome



1. Cities in Europe

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. Average Population by Country

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A table with numbers and letters

Description automatically generated with medium confidence

1. Capital Cities Population Comparison

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. Countries with Low Population Density

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a data

Description automatically generated

1. Cities with High GDP per Capita

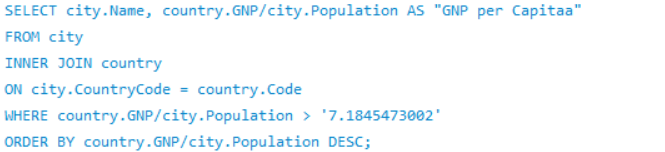
An economic consulting firm is analyzing cities with high GDP per capita for investment opportunities. You're tasked with identifying cities with above-average GDP per capita from the database to assist the firm in identifying potential investment destinations.

Firstly determine the average

A white background with blue text

Description automatically generated





Snippet of query outcome

A screenshot of a computer

Description automatically generated

1. Display Columns with Limit (Rows 31-40)

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.

A white background with blue text

Description automatically generated

Snippet of query outcome

A screenshot of a computer

Description automatically generated

# Summary

Using a dataset containing the names of cities along with their respective countries, population, Gross National Product (GNP), and surface area provides a comprehensive overview of urban demographics and economic indicators. This dataset can be invaluable for analysing urban growth patterns, economic performance, and spatial distribution of resources. By examining the population data, one can identify densely populated areas and potential markets.