

SQL Project Solution

Create a Database Schema and Table Relationships for a Logistic Company's Data

Table of Contents

ABSTRACT	2
PROBLEM DESCRIPTION	3
SCOPE	4
TABLE DEFINITIONS	5
LOADING DATA	9
DATA PREPROCESSING (Working with dates)	13
EXPLORATORY DATA ANALYSIS (EDA)	15
CONCLUSION	20

ABSTRACT

Logistics is the support function of an organization and it means having the right object, at the right place, in the right time. Logistics deals with various kinds of methods to control the flow of resources from one place to another. One of the major and the most important factors that is costing is being dealt with utmost attention. The project is being designed keeping in attention the details of the various requirements of logistics such as keeping records of the goods i.e. their details and the kind of content that is stored in the shipment which is to be delivered.

A Relational Database Management System (RDBMS) is similar to DBMS. The difference is that in RDBMS, the entities and values in tables are related to one another. Also the tables are related to each other. Thus, it is called “Relational”.

PROBLEM DESCRIPTION

The logistics company provides services in both the international and domestic sectors. The logistics management takes into consideration every facility that has an impact on cost. It plays an important role in making the product conform to customer requirements. Also it involves efficient integration of suppliers, manufacturers, Import & export and other activities at many levels, from the strategic level through the tactical to the operational level.

Customers can send different types of shipping contents. Payment is to be paid at the same time the product is delivered to the client. Delivery boy and centre head can update the status of the shipment. Create a database schema and table relationships that can be used in any technology.

SCOPE

It is of critical importance to the organization how it delivers products & services to the customer, whether the product is tangible or intangible. Effective and efficient physical movement of the tangible product will speak of intangible services associated with the product and the organization which is delivering it.

In case of intangible products, the delivery of tangibles at the right place & right time will speak about its quality. On the macro level infrastructure such as various modes of transport, transportation equipment, storage facilities, connectivity and information processing are contributing to a large extent in the physical movement of goods produced in manufacturing, mining and agriculture Sectors.

This speed and reliability in distribution of products and services contribute to a great extent in the growth of a country's domestic and international trade.

TABLE DEFINITIONS

1) Employee_Details Table:

This table contains the information of the employees.

Column Name	Data Type	Description
Emp_ID	INT (5)	Employee ID (Primary Key)
Emp_NAME	VARCHAR (30)	Name of the employee
Emp-BRANCH	VARCHAR (15)	Branch name
Emp_DESIGNATION	VARCHAR (40)	Designation of the employee
Emp_ADDR	VARCHAR (100)	Address of the employee
Emp_CONT_NO	VARCHAR (10)	Contact Number of the employee

2) Membership Table:

This table contains the membership details of the customer or client.

Column Name	Data Type	Description
M_ID	INT	Membership ID associated with the client (Primary Key)
START_DATE	TEXT	Start date of the membership
END_DATE	TEXT	End date of the membership

3) Customer Table:

This table contains the information of the customers or clients.

Column Name	Data Type	Description
Cust_ID	INT (4)	Client ID (Primary Key)
Cust-NAME	VARCHAR (30)	Name of the client
Cust-EMAIL_ID	VARCHAR (50)	Email of the client
Cust_CONT_NO	VARCHAR (10)	Contact Number of the client
Cust_ADDR	VARCHAR (100)	Address of the client
Cust_TYPE	VARCHAR (30)	Type of client (Wholesale, Retail, Internal Goods)
Membership_M_ID	INT	Membership ID (Foreign Key)

4) Payment_Details Table:

This table contains the payment details.

Column Name	Data Type	Description
PAYMENT_ID	VARCHAR (40)	Payment Unique ID (Primary Key)
AMOUNT	INT	Price to be paid by the client
PAYMENT_STATUS	VARCHAR (10)	Payment status (Paid / Not Paid)
PAYMENT_DATE	TEXT	Date when payment is made by the client
PAYMENT_MODE	VARCHAR (25)	Mode of payment (COD / Card Payment)
Shipment_SH_ID	VARCHAR (6)	Shipment ID (Foreign Key)
Shipment_Client_C_ID	INT (4)	Client ID (Foreign Key)

5) Shipment_Details Table:

This table contains the shipment details.

Column Name	Data Type	Description
SD_ID	VARCHAR (6)	Shipment ID (Primary Key)
SD_CONTENT	VARCHAR (40)	Type of shipping
SD_DOMAIN	VARCHAR (15)	Shipment Domain (International / Domestic)
SD_TYPE	VARCHAR (15)	Service Type (Express / Regular)
SD_WEIGHT	VARCHAR (10)	Shipment Weight
SD-CHARGES	INT (10)	Shipment Charges
SD-ADDR	VARCHAR (100)	Source Address
DS_ADDR	VARCHAR (100)	Destination Address
Customer_Cust_ID	INT (4)	Client ID (Foreign Key)

6)Status table:

This table contains the details about the delivery status.

Column Name	Data Type	Description
CURRENT_ST	VARCHAR (15)	Current status of the
SENT_DATE	TEXT	Date when shipment was sent
DELIVERY_DATE	TEXT	Date when the product was/will be delivered
SH_ID	VARCHAR (6)	Shipment ID (Primary Key)

6) Employee Manages Shipment Table:

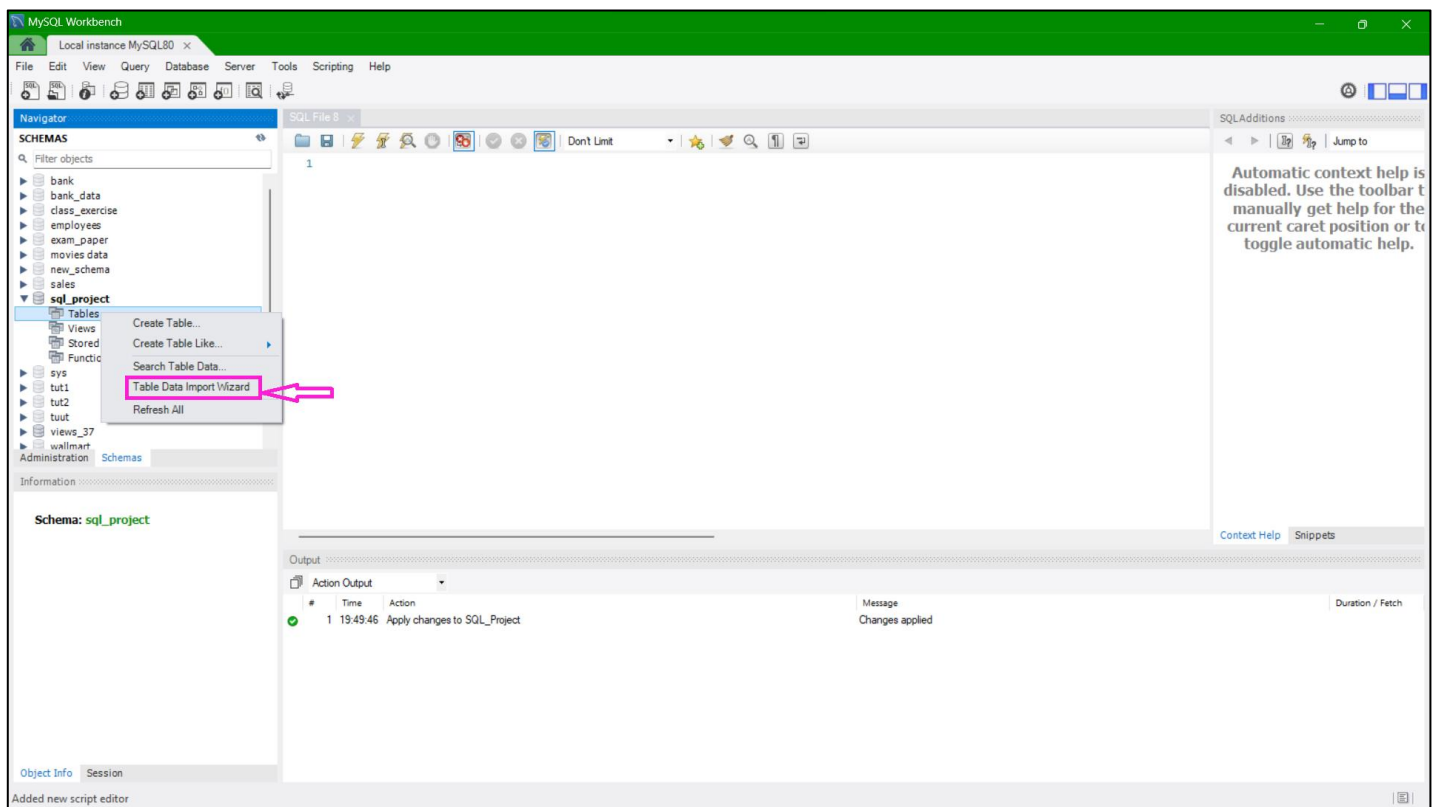
This is a relationship table between the employee and the shipment table.

Column Name	Data Type	Description
Employee_E_ID	INT (5)	Employee ID (Foreign Key)
Shipment_SH_ID	VARCHAR (6)	Shipment ID (Foreign Key)
Status_SH_ID	VARCHAR (6)	Shipment_ID from status table (Foreign Key)

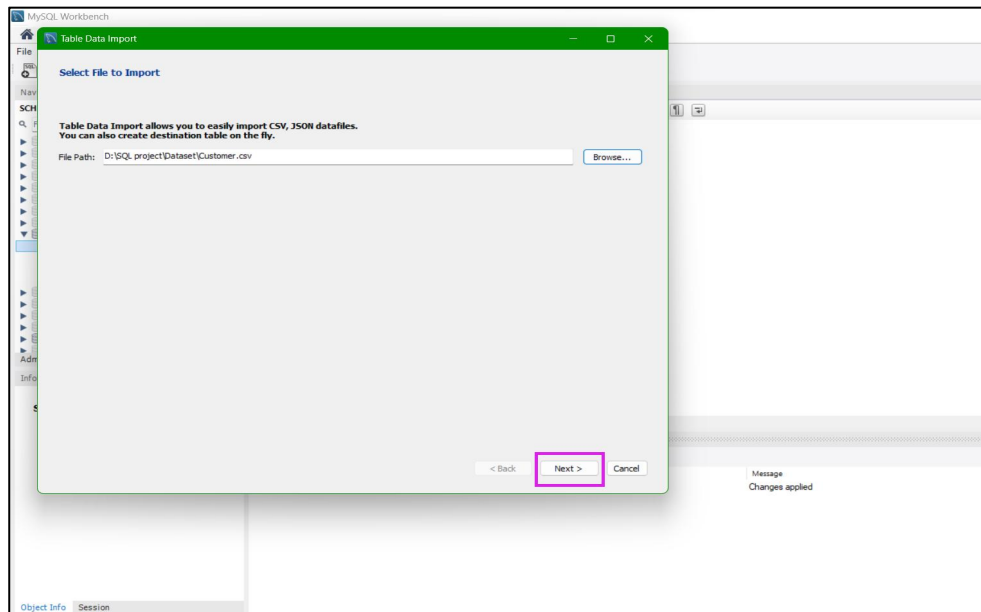
LOADING DATA

MySQL Workbench provides the facility to upload the csv file into the tables using the graphical user interface (GUI). Below are the steps for the same:

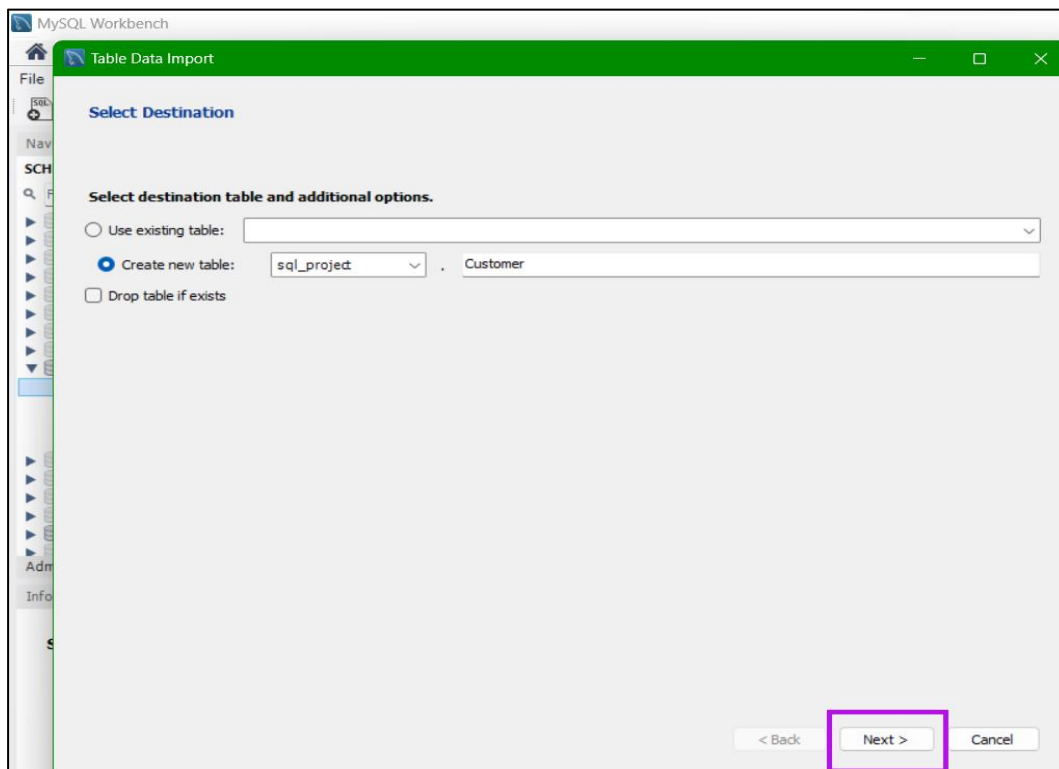
Step 1: Right-click on the table and click on “Table Data Import Wizard”.



Step 2: Browse and select the csv file and click “Next” button.



Step 3: Click “Next”



Step 4: Map the source and destination columns correctly and click “Next”.

Table Data Import

Configure Import Settings

Detected file format: csv

Encoding: utf-8

Columns:

Source Column	Dest Column
<input checked="" type="checkbox"/> E_ID	<input type="text" value="E_ID"/>
<input checked="" type="checkbox"/> E_NAME	<input type="text" value="E_name"/>
<input checked="" type="checkbox"/> E_DESIGNATION	<input type="text" value="E_designatio"/>
<input checked="" type="checkbox"/> E_ADDR	<input type="text" value="E_addr"/>
<input checked="" type="checkbox"/> E_BRANCH	<input type="text" value="E_branch"/>
<input checked="" type="checkbox"/> E_CONT_NO	<input type="text" value="E_Cont_no"/>

Source Columns are the columns from the csv file

Dest Columns are the columns from the employee table

E_ID	E_NAME	E_DESIGN...	E_ADDR	E_BRANCH	E_CONT_NO
582	Harriette	Market anal...	600 Block o...	TX	2754220306
396	Matthew	Chief financ...	EDDY ST /...	MA	4171197971
545	Geraldine	Transport...	800 Block o...	CT	8354987185
770	Brenda	Warehouse...	3300 Block...	UT	6348759218
991	Malie	Branch ma...	3200 Block...	OH	5096424869

< Back **Next >** Cancel

Step 5: Click “Next”

Table Data Import

Import Data

The following tasks will now be performed. Please monitor the execution.

☐ Prepare Import

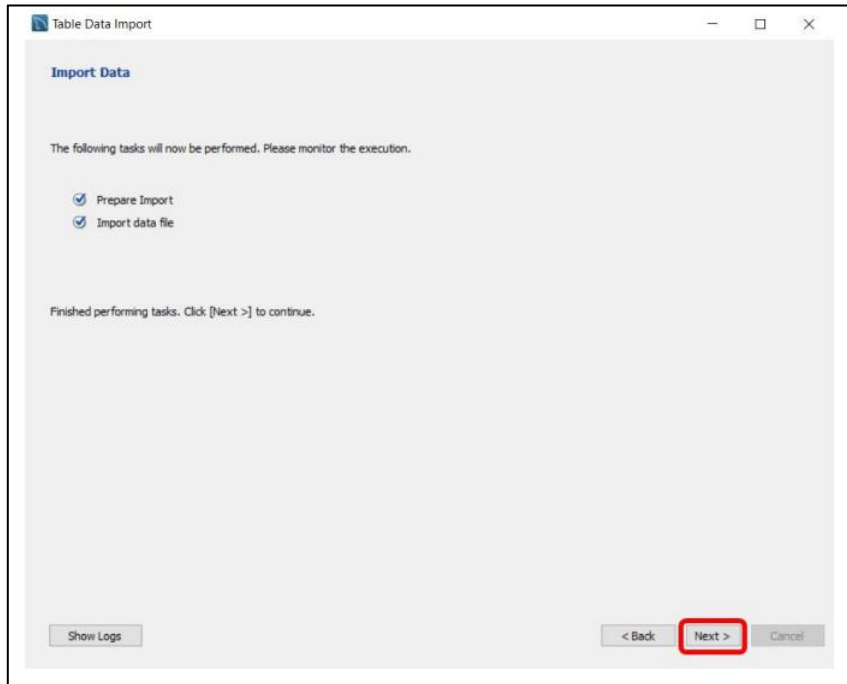
☒ Import data file

Click [Next >] to execute.

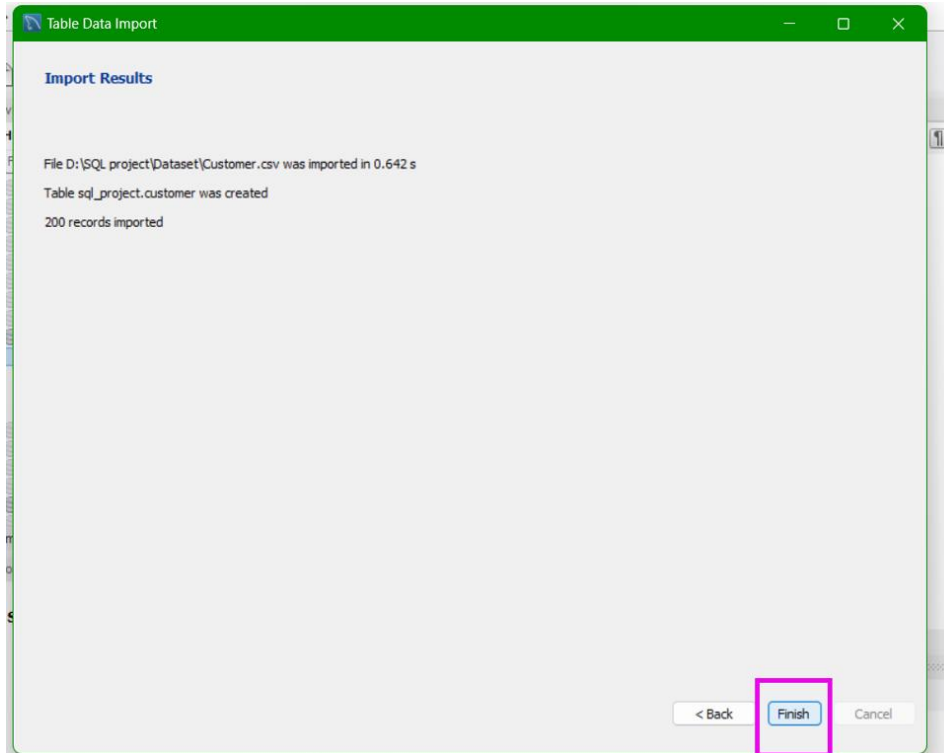
Show Logs

< Back **Next >** Cancel

Step 6: Click “Next”



Step 7: Data imported successfully. Click “Finish”



DATA PREPROCESSING (Working with dates)

It is not necessary that the dates obtained from csv files will always be in the same format. They can be separated by '/' (slash) or by '-'. Also dates can be in any format like 'dd-mm-yy', 'dd-mm-yyyy', 'yyyy-mm-dd', etc and many more.

In tables PAYMENT and MEMBERSHIP, the date is in the format "%Y-%m-%d". In the STATUS table the date is of the format "%m/%d/%Y".

There can be some dates that are entered erroneously like 02/31/1999. There are only 28 or 29 days in the month of February. But the date '31' can be wrongly entered.

Steps to perform while dealing with the dates after importing the data from csv files:

1. Look for erroneous dates

There can be dates where the month is greater than 12. For example: Find the erroneous date from the column 'DELIVERY_DATE' in the 'STATUS' table where the month is greater than 12.

```
SELECT DELIVERY_DATE FROM STATUS
WHERE CAST(substring_index(DELIVERY_DATE, '/', 1) AS UNSIGNED) >
12;
```

Search for the records where the month is February but the date is erroneously entered as 30 and 31.

```
SELECT * FROM STATUS
WHERE CAST(substring_index(DELIVERY_DATE, '/', 1) AS UNSIGNED) = 2
AND
CAST(substring_index(substring_index(DELIVERY_DATE, '/', 2), '/',
-1) AS UNSIGNED) > 29;
SELECT * FROM STATUS
WHERE CAST(substring_index(SENT_DATE, '/', 1) AS UNSIGNED) = 2
AND
```

```
CAST(substring_index(substring_index(SENT_DATE, '/', 2), '/', -1)
AS UNSIGNED) > 29;
SELECT * FROM PAYMENT_DETAILS
WHERE CAST(substring_index(substring_index(PAYMENT_DATE, '-', 2),
'- ', -1) AS UNSIGNED) = 2
AND
CAST(substring_index(PAYMENT_DATE, '-', -1) AS UNSIGNED) > 29;
```

2. Convert the string in the date format

```
UPDATE PAYMENT_DETAILS
SET Payment_Date = STR_TO_DATE(Payment_Date, '%Y-%m- %d');
UPDATE STATUS
SET Delivery_Date = STR_TO_DATE(Delivery_Date, '%m/%d/%Y'),
Sent_Date = STR_TO_DATE(Sent_Date, '%m/%d/%Y');
UPDATE MEMBERSHIP
SET Start_Date = STR_TO_DATE(Start_Date, '%Y-%m-%d'),
End_Date = STR_TO_DATE(End_Date, '%Y-%m-%d');
```

2. Change the data type of the column to DATE

```
ALTER TABLE PAYMENT_DETAILS
MODIFY COLUMN Payment_Date Date;
ALTER TABLE STATUS
MODIFY COLUMN Delivery_Date Date,
MODIFY COLUMN Sent_Date Date ;
ALTER TABLE MEMBERSHIP
MODIFY COLUMN Start_Date Date,
MODIFY COLUMN End_Date Date ;
```

EXPLORATORY DATA ANALYSIS (EDA)

1) Extract all the employees whose name starts with A and ends with A.

```
select E_NAME from employee_details  
where E_NAME Like 'A%A';
```

2) Find all the common names from Employee_Details names and Customer names.

```
SELECT DISTINCT(E_name) FROM Employee_Details WHERE  
E_name IN (SELECT C_name FROM Customer AS cus);
```

3) Create a view 'PaymentNotDone' of those customers who have not paid the amount.

```
CREATE VIEW PaymentNotDone AS  
SELECT * FROM sql_project  
WHERE PAYMENT_STATUS = 'NOT PAID';
```

-- Selecting all the observations of the newly created view 'PaymentNotDone'

```
SELECT * FROM PaymentNotDone;
```

4) Find the frequency (in percentage) of each of the class of the payment mode

```
SET @total_count = 0;
SELECT COUNT(*) INTO @total_count FROM Pyament_Deatils;
SELECT PAYMENT_MODE, ROUND((COUNT(PAYMENT_MODE) /
@total_count) * 100,2) AS Percentage_Contribution
FROM Payment_Details
GROUP BY PAYMENT_MODE;
```

5) What is the highest total payable amount ?

```
SELECT MAX(Amount) FROM sql_project;
```

6) Extract the customer id and the customer name of the customers who were or will be the member of the branch for more than 10 years

```
SELECT C_ID, C_NAME, START_DATE, END_DATE,
ROUND(DATEDIFF(END_DATE, START_DATE)/365,0)
    AS Membership_Years FROM sql_project
HAVING Membership_Years > 10;
```

7) Who got the product delivered on the next day the product was sent?

```
SELECT * FROM sql_project
    HAVING DELIVERY_DATE-SENT_DATE = 1;
SELECT * FROM sql_project
    HAVING DATEDIFF(DELIVERY_DATE, SENT_DATE)=1;
```


8) Which shipping content had the highest total amount (Top 5).

```
SELECT SH_CONTENT, SUM(AMOUNT) AS Content_Wise_Amount
FROM sql_project
GROUP BY (SH_CONTENT)
ORDER BY Content_Wise_Amount DESC
LIMIT 5;
```

9) Which product categories from shipment content are transferred more?

```
SELECT SH_CONTENT, COUNT(SH_CONTENT)
      AS Content_Wise_Count
FROM sql_project
GROUP BY (SH_CONTENT)
ORDER BY Content_Wise_Count DESC
LIMIT 5;
```

10) Create a new view 'TXLogistics' where employee branch is Texas.

```
CREATE VIEW TXLogistics AS
      SELECT * FROM sql_project
            WHERE E_BRANCH = 'TX';

SELECT * FROM TXLogistics;
```

11) Texas(TX) branch is giving 5% discount on total payable amount. Create a new column 'New_Price' for payable price after applying discount.

```
ALTER VIEW TXLogistics
    AS SELECT *, AMOUNT - ((AMOUNT * 5)/100) AS New_Price
    FROM logistics_Emp
    WHERE E_BRANCH = 'TX';
SELECT * FROM TXLogistics;
```

12) Drop the view TXLogistics

```
DROP VIEW TXLogistics;
```

13) The employee branch in New York (NY) is shutdown temporarily. Thus, the the branch needs to be replaced to New Jersey (NJ).

```
SELECT * FROM sql_project WHERE E_BRANCH = 'NY';
```

```
UPDATE sql_project
    SET E_BRANCH = 'NJ'
WHERE E_BRANCH = 'NY';
```

```
SELECT * FROM sql_project;
```

14) Finding the unique designations of the employees.

```
SELECT DISTINCT(E_DESIGNATION) FROM Employee_Details;
```

15) Rename the column SER_TYPE to SERVICE_TYPE.

```
ALTER TABLE sql_project  
CHANGE SER_TYPE SERVICE_TYPE VARCHAR (15);
```

16) Which service type is preferred more?

```
SELECT SERVICE_TYPE, COUNT(SERVICE_TYPE)  
      AS Frequency  
FROM sql_project  
GROUP BY SERVICE_TYPE  
ORDER BY Frequency DESC;
```

17) Find the shipment id and shipment content where the weight is greater than the average weight.

```
SELECT SH_ID, SH_CONTENT, SH_WEIGHT FROM  
Shipment_Details  
WHERE SH_WEIGHT > (SELECT AVG(SH_WEIGHT) FROM  
Shipment_Details);
```

CONCLUSION

The rise in the demand of transportation of shipment from one place to another and due to development of better transportation facilities all around the globe, logistics has taken a vital position in business processes all around the world. The increase in the amount of transfer of contents has also contributed to the development of logistics this rapidly.

Due to all these factors it became necessary to keep track of all whereabouts of the shipments. Logistics system not only helps us to keep track of them but also provides us with better solutions and helps us to get maximum utilization of the available resources. Keeping track of any shipment and knowing its current status becomes easy.

Thank You

