Project: Querying a Large Relational Database

Submitted by: VENKATESH J M

Institution : Intellipaat

Problem Statement:

How to get details about customers by querying a database?

Introduction to SQL (Structured Query Language)

In the digital age, data has become the lifeblood of organizations and businesses worldwide. The ability to effectively store, retrieve, manipulate, and analyze data is critical for making informed decisions and driving success. This is where SQL, or Structured Query Language, emerges as an indispensable tool.

SQL is the universal language of data. It is the means by which we communicate with relational database management systems (RDBMS), the software that underpins much of the data infrastructure across industries. Whether you're managing a small business's customer records, analyzing massive datasets in a multinational corporation, or even building web applications, SQL is the foundation upon which you can build efficient, organized, and scalable data solutions.

At its core, SQL enables you to perform a multitude of tasks:

1. **Data Retrieval:** You can extract specific information from vast databases with precision, making it possible to answer questions like, "Which products sold the most last month?" or "Who are our top-performing customers?"
2. **Data Manipulation:** SQL allows you to add, update, and delete records, ensuring that your data remains accurate and up-to-date.
3. **Data Definition:** You can create and modify the structure of databases, defining tables, relationships, and constraints to maintain data integrity.
4. **Data Analysis:** SQL's power shines when it comes to aggregating data, performing calculations, and generating reports. It's a fundamental tool for data scientists and analysts.
5. **Automation:** You can automate repetitive tasks by creating stored procedures and triggers, which execute SQL commands when certain conditions are met.
6. **Security:** SQL provides robust security mechanisms, ensuring that sensitive data is protected from unauthorized access.

Table Basics and Data Types:

Tables: In a relational database, data is stored in tables, which are organized into rows and columns. Each row represents a record, and each column represents an attribute of that record.

Data Types: Data types define the type of data that can be stored in a column. Common data types include:

INTEGER: Used for whole numbers.

VARCHAR(size): Variable-length character strings.

DATE: Stores date values.

BOOLEAN: Represents true or false values.

DECIMAL(precision, scale): Used for decimal numbers with specified precision and scale.

BLOB: Binary Large Object, for storing binary data.

Various SQL Operators:

Comparison Operators:

- ‘ = ‘ (Equal)
- ‘< >’or ‘ != ‘ (Not Equal)
- < (Less Than)
- > (Greater Than)
- <= (Less Than or Equal To)
- >= (Greater Than or Equal To)

Logical Operators:

- **AND:** Combines multiple conditions, requiring all conditions to be true.
- **OR:** Combines multiple conditions, requiring at least one condition to be true.
- **NOT:** Negates a condition, making a true condition false and vice versa.

Arithmetic Operators:

+ (Addition)

- (Subtraction)

* (Multiplication)

/ (Division)

% (Modulo)

Concatenation Operator:

|| (Double Pipe): Used to concatenate strings.

IN Operator:

Allows you to specify a list of values to check against in a WHERE clause.

LIKE Operator:

Used for pattern matching in text fields, often with wildcard characters % (matches any sequence of characters) and _ (matches any single character).

Various SQL Functions:

Aggregate Functions:

SUM(): Calculates the sum of values in a column.

AVG(): Computes the average of values in a column.

COUNT(): Counts the number of rows.

MIN(): Finds the minimum value in a column.

MAX(): Finds the maximum value in a column.

String Functions:

CONCAT(): Concatenates strings.

SUBSTRING(): Extracts a substring from a string.

UPPER() and LOWER(): Convert text to uppercase or lowercase.

LENGTH() or LEN(): Returns the length of a string.

Date and Time Functions:

NOW(): Returns the current date and time.

DATE() and TIME(): Extract the date or time portion from a datetime.

DATEDIFF(): Calculates the difference between two dates.

Math Functions:

ABS(): Returns the absolute value of a number.

ROUND(): Rounds a number to a specified number of decimal places.

CEIL() and FLOOR(): Round up or down to the nearest integer.

Tasks To Be Performed:

1. Download the AdventureWorks database from the following location and restore it in your server:

Location:
https://github.com/Microsoft/sql-server-samples/releases/tag/adventureworks
File Name: AdventureWorks2012.bak

AdventureWorks is a sample database shipped with SQL Server and it can be downloaded from the GitHub site. AdventureWorks has replaced Northwind and Pubs sample databases that were available in SQL Server in 2005. Microsoft keeps updating the sample database as it releases new versions.

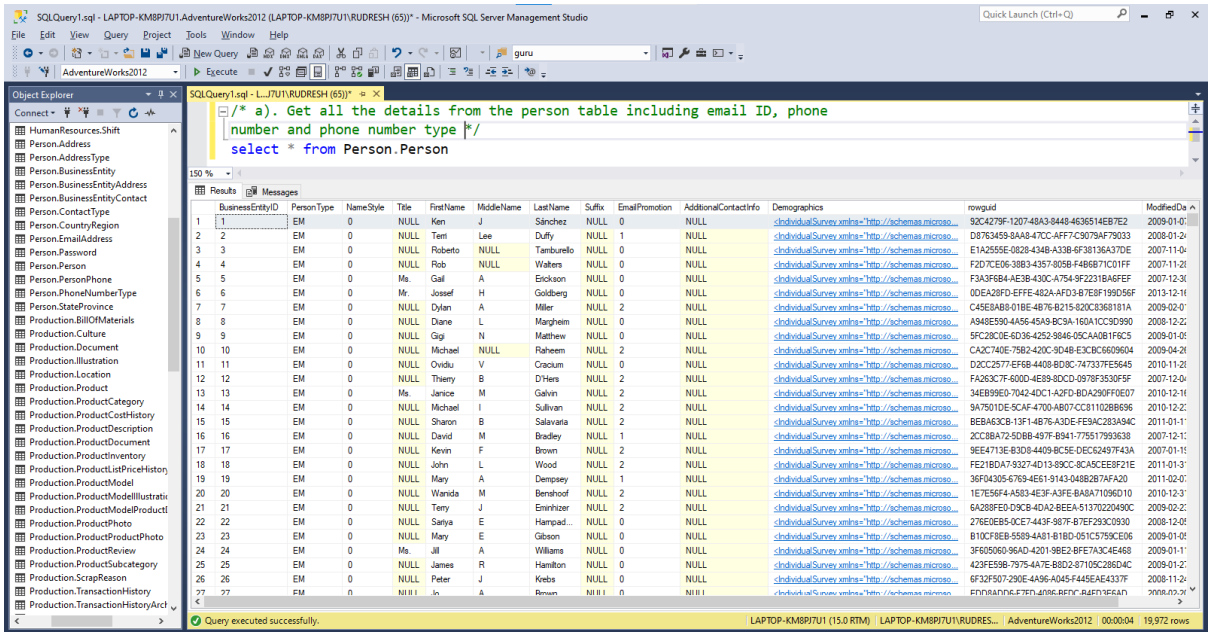
2. Restore Backup:

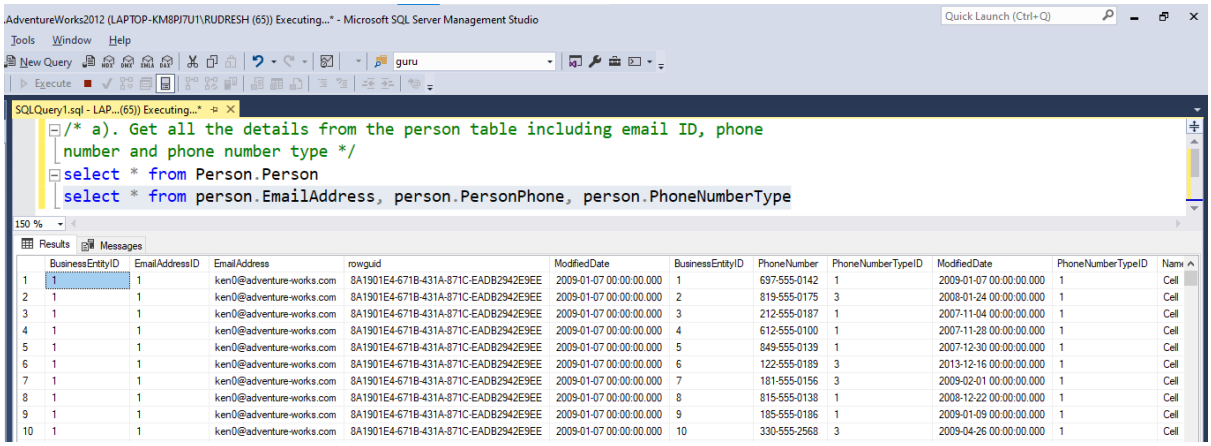
Follow the below steps to restore a backup of your database using SQL Server Management Studio:

a. Open SQL Server Management Studio and connect to the target SQL Server instance
b. Right-click on the Databases Node and select Restore Database
c. Select Device and click on the ellipsis (...)
d. In the dialog box, select Backup devices, click on Add, navigate to the database backup in the file system of the server, select the backup, and click on OK.
e. If needed, change the target location for the data and log files in the Files pane
Note: It is a best practice to place the data and log files on different drives.
f. Now, click on OK

3. Perform the following with help of the above database:

a. Get all the details from the person table including email ID, phone number and phone number type





b. Get the details of the sales header order made in May 2011

```
--b.    Get the details of the sales header order made in May 2011
select * from Sales.SalesOrderHeader
WHERE month(OrderDate) = 5 AND year(OrderDate) = 2011
```

150 %

Results | Messages

	SalesOrderID	RevisionNumber	OrderDate	DueDate	ShipDate	Status	OnlineOrderFlag	SalesOrderNumber	PurchaseOrderNumber	AccountNumber	CustomerID	SalesPersonID	TerritoryID	BillT
1	43659	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43659	PO522145787	10-4020-000676	29825	279	5	985
2	43660	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43660	PO18850127500	10-4020-000117	29672	279	5	921
3	43661	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43661	PO18473189620	10-4020-000442	29734	282	6	517
4	43662	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43662	PO18444174044	10-4020-000227	29994	282	6	482
5	43663	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43663	PO18009186470	10-4020-000510	29565	276	4	107
6	43664	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43664	PO16617121983	10-4020-000397	29898	280	1	876
7	43665	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43665	PO16588191572	10-4020-000146	29580	283	1	849
8	43666	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43666	PO16008173883	10-4020-000511	30052	276	4	107
9	43667	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43667	PO15428132599	10-4020-000646	29974	277	3	629
10	43668	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43668	PO14732180295	10-4020-000514	29614	282	6	529
11	43669	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43669	PO14123169936	10-4020-000578	29747	283	1	895
12	43670	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43670	PO14384116310	10-4020-000504	29566	275	3	810
13	43671	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43671	PO13978119376	10-4020-000200	29890	283	1	855
14	43672	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43672	PO13862153537	10-4020-000119	30067	282	6	464
15	43673	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43673	PO13775141242	10-4020-000618	29844	275	2	821
16	43674	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43674	PO12760141756	10-4020-000083	29596	282	6	458
17	43675	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43675	PO12412186464	10-4020-000670	29827	277	3	631
18	43676	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43676	PO11861165059	10-4020-000017	29811	275	5	755
19	43677	8	2011-05-31 00:00:00.000	2011-06-12 00:00:00.000	2011-06-07 00:00:00.000	5	0	SO43677	PO11049174786	10-4020-000679	29824	278	6	556

Query executed successfully. | LAPTOP-KM8PJ7U1 (15.0 RTM) | LAPTOP-KM8PJ7U1\RUDRES... | AdventureWorks2012 | 00:00:00 | 43 rows

c. Get the details of the sales details order made in the month of May 2011

```
--c. Get the details of the sales details order made in the month of May 2011
select * from Sales.SalesOrderDetail
WHERE year(ModifiedDate) = 2011
```

150 %

Results | Messages

	SalesOrderID	SalesOrderDetailID	CarrierTrackingNumber	OrderQty	ProductID	SpecialOfferID	UnitPrice	UnitPriceDiscount	LineTotal	rowguid	ModifiedDate
1	43659	1	4911-403C-98	1	776	1	2024.994	0.00	2024.994000	B207C96D-D9E6-402B-8470-2CC176C42283	2011-05-31 00:00:00.000
2	43659	2	4911-403C-98	3	777	1	2024.994	0.00	6074.982000	7ABB600D-1E77-41BE-9FE5-B9142CFC08FA	2011-05-31 00:00:00.000
3	43659	3	4911-403C-98	1	778	1	2024.994	0.00	2024.994000	475CF8C6-49F6-486E-B0AD-AFC6A50CDD2F	2011-05-31 00:00:00.000
4	43659	4	4911-403C-98	1	771	1	2039.994	0.00	2039.994000	04C4DE91-5815-45D6-8670-F462719FBCE3	2011-05-31 00:00:00.000
5	43659	5	4911-403C-98	1	772	1	2039.994	0.00	2039.994000	5A74C7D2-E641-438E-A7AC-37BF23280301	2011-05-31 00:00:00.000
6	43659	6	4911-403C-98	2	773	1	2039.994	0.00	4079.988000	CE472532-A4C0-45BA-816E-EEFD3FD848B3	2011-05-31 00:00:00.000
7	43659	7	4911-403C-98	1	774	1	2039.994	0.00	2039.994000	80667840-F962-4EE3-96E0-AECA108E0D4F	2011-05-31 00:00:00.000
8	43659	8	4911-403C-98	3	714	1	28.8404	0.00	86.521200	E9D54907-E7B7-4969-80D9-76BA69F8A836	2011-05-31 00:00:00.000
9	43659	9	4911-403C-98	1	716	1	28.8404	0.00	28.840400	AA542630-BDCD-4CE5-89A0-C1BF82747725	2011-05-31 00:00:00.000
10	43659	10	4911-403C-98	6	709	1	5.70	0.00	34.200000	AC769034-3C2F-495C-A5A7-3B71CDB25D4E	2011-05-31 00:00:00.000
11	43659	11	4911-403C-98	2	712	1	5.1865	0.00	10.373000	06A66921-6B9F-4199-A912-DDAFD383472B	2011-05-31 00:00:00.000
12	43659	12	4911-403C-98	4	711	1	20.1865	0.00	80.746000	0E371EE3-253E-4BB0-B813-83CF4224F972	2011-05-31 00:00:00.000
13	43660	13	6431-4D57-83	1	762	1	419.4589	0.00	419.458900	419A1302-AC7A-4044-97B2-66D9D14CD02E	2011-05-31 00:00:00.000
14	43660	14	6431-4D57-83	1	758	1	874.794	0.00	874.794000	5D0B2B03-1D4C-4C34-9696-C14C58E7301C	2011-05-31 00:00:00.000
15	43661	15	4E0A-4F89-AE	1	745	1	809.76	0.00	809.760000	EDE1759E-6733-4C7B-A43F-DC6F48002D8A	2011-05-31 00:00:00.000
16	43661	16	4E0A-4F89-AE	1	743	1	714.7043	0.00	714.704300	FE10BF09-D477-485B-9541-27AE8053A6D4	2011-05-31 00:00:00.000
17	43661	17	4E0A-4F89-AE	2	747	1	714.7043	0.00	1429.408600	B136852E-24C9-4006-8048-B14AEFE6C337	2011-05-31 00:00:00.000
18	43661	18	4E0A-4F89-AE	4	712	1	5.1865	0.00	20.746000	F0F410B6-BC93-4D77-B70D-E4E2465D98EE	2011-05-31 00:00:00.000
19	43661	19	4E0A-4F89-AE	4	715	1	28.8404	0.00	115.361600	AE4CDA7D-9B6E-4ABB-A391-E9E9DDB10125	2011-05-31 00:00:00.000
20	43661	20	4E0A-4F89-AE	2	742	1	722.5949	0.00	1445.189800	054A9D3B-A178-4656-B2B0-9B9BDDF1C345	2011-05-31 00:00:00.000

Query executed successfully. | LAPTOP-KM8PJ7U1 (15.0 RTM) | LAPTOP-KM8PJ7U1\RUDRES... | AdventureWorks2012 | 00:00:00 | 5,716 rows

d. Get the total sales made in May 2011

```
--d.    Get the total sales made in May 2011
select * from Sales.SalesOrderDetail

SELECT SalesOrderID, SUM(orderqty) AS Total_sales
FROM Sales.SalesOrderDetail
WHERE YEAR(modifieddate) = 2011
GROUP BY SalesOrderID

select sum(orderqty) as total_sales from Sales.SalesOrderDetail
where year(modifieddate) = 2011
```

150 %

Results | Messages

	total_sales
1	12888

e. Get the total sales made in the year 2011 by month order by increasing sales

```
--e.    Get the total sales made in the year 2011 by month order by increasing sales
select * from Sales.SalesOrderDetail
select
    datepart(month, modifieddate) AS Month,
    sum(orderqty) AS Total_Sales
from
    Sales.SalesOrderDetail
where
    year(modifieddate) = 2011
GROUP BY
    DATEPART(MONTH, modifieddate)
ORDER BY
    Total_Sales;
```

150 %

Results | Messages | Client Statistics

	Month	Total_Sales
1	6	141
2	9	157
3	11	230
4	5	825
5	12	1040
6	7	2209
7	8	2904
8	10	5382