

## Detect and Remove Outliers

The screenshot shows the Microsoft SQL Server Enterprise Manager interface. The left pane displays the 'Northwind' database structure. The central pane shows a SQL query in the 'SQLQuery1.sql' file. The query uses a Common Table Expression (CTE) to rank rows by CustomerID and then filters for the first row (rn = 1) for each customer, which is a common technique to identify outliers or specific data points.

```
WITH cte1 as (  
    SELECT CustomerID, ShipName, ShipAddress, ShipPostalCode, ShipCountry,  
    ROW_NUMBER() OVER(partition by CustomerID order by CustomerID) as rn  
    from Orders  
)  
  
SELECT * FROM cte1 WHERE rn = 1
```

The bottom pane shows the 'Results' tab with a table of 5 rows. The first row is highlighted.

	CustomerID	ShipName	ShipAddress	ShipPostalCode	ShipCountry
1	ALFKI	Alfreds Futterkiste	Obere Str. 57	12209	Germany
2	ANATR	Ana Trujillo Emparedados y helados	Avda. de la Constitución 2222	05021	Mexico
3	ANTON	Antonio Moreno Taqueria	Mataderos 2312	05023	Mexico
4	AROUT	Around the Horn	Brook Farm Stratford St. Mary	CO7 6JX	UK
5	BERGS	Berglunds snabbköp	Berguvsvägen 8	S-958 22	Sweden

## Clean Data With Case Statements

The screenshot displays the Microsoft SQL Server Enterprise Manager interface. The left pane shows the 'Server Enterprise Explorer' tree with the 'Northwind' database selected. The right pane shows a query window with the following SQL code:

```
select Freight,  
CASE WHEN Freight between 0 and 50 THEN 'Low Charge'  
      WHEN freight between 50 and 200 THEN 'Medium Charge'  
      WHEN freight > 200 THEN 'High Charge'  
      Else 'No charge'  
END as Charge  
from Orders
```

Below the query window, the 'Results' tab is active, showing a table with 7 rows of data. The columns are 'Freight' and 'Charge'.

	Freight	Charge
1	32.38	Low Charge
2	11.61	Low Charge
3	65.83	Medium Charge
4	41.34	Low Charge
5	51.30	Medium Charge
6	58.17	Medium Charge
7	22.98	Low Charge

## Dealing with NULLs

The screenshot displays the SQL Server Enterprise Manager interface. The Object Explorer on the left shows the 'Northwind' database selected. The main query window, titled 'SQLQuery3.sql - EC...Administrator (66))', contains the following SQL query:

```
SELECT CustomerID, ContactName, City,  
ISNULL(Region, 'No Region') AS Region  
FROM Customers
```

Below the query window, the 'Results' tab is active, showing a table with 7 rows and 5 columns. The columns are CustomerID, ContactName, City, and Region. The 'Region' column contains the value 'No Region' for all rows, which is the result of the ISNULL function used in the query.

	CustomerID	ContactName	City	Region
1	ALFKI	Maria Anders	Berlin	No Region
2	ANATR	Ana Trujillo	México D.F.	No Region
3	ANTON	Antonio Moreno	México D.F.	No Region
4	AROUT	Thomas Hardy	London	No Region
5	BERGS	Christina Berglund	Luleå	No Region
6	BLAUS	Hanna Moos	Mannheim	No Region
7	BONAP	Frédérique Citeaux	Strasbourg	No Region

Northwind | Execute

SQLQuery4.sql - EC...Administrator (65))" | SQLQuery3.sql - EC...Administrator (66))"

```
SELECT QuantityPerUnit,  
CASE WHEN UnitsOnOrder >50 and UnitsInStock <20 then 'Restock Now'  
      WHEN UnitsOnOrder between 30 and 50 and UnitsInStock<20 THEN 'Restock Next Week'  
      WHEN UnitsOnOrder between 30 and 50 and UnitsInStock<50 THEN 'Restock Next Month'  
      WHEN UnitsOnOrder <30 and UnitsInStock<50 THEN 'Restock in 6 months'  
      ELSE 'Ask Manager'  
END as WhenToRestock  
from Products
```

100 %

Results | Messages

	QuantityPerUnit	WhenToRestock
1	10 boxes x 20 bags	Restock in 6 months
2	24 - 12 oz bottles	Restock Next Week
3	12 - 550 ml bottles	Restock Now
4	48 - 6 oz jars	Ask Manager
	36 boxes	Restock in 6 months
	12 - 8 oz jars	Ask Manager
	12 - 1 lb pkgs	Restock in 6 months

File Edit View Query Project Tools Window Help



Northwind

Execute

Obj... x

SQLQuery6.sql - EC...Administrator (62))

SQLQuery5.sql - EC...Administrator (67))\* x

Connect

EC2

```
with blank_fax as (  
  SELECT CustomerID, ContactName, City, Region,  
  ISNULL(Fax, 'No Fax') as Fax  
  from Customers  
)  
  
select * from blank_fax where fax = 'No Fax'
```

100 %

Results

Messages

	CustomerID	ContactName	City	Region	Fax
1	ANTON	Antonio Moreno	México D.F.	NULL	No Fax
2	BSBEV	Victoria Ashworth	London	NULL	No Fax
3	CHOPS	Yang Wang	Bern	NULL	No Fax
4	COMMI	Pedro Afonso	Sao Paulo	SP	No Fax
5	FAMIA	Aria Cruz	Sao Paulo	SP	No Fax
6	FOLKO	Maria Larsson	Bräcke	NULL	No Fax