PART 4: Data Wrangling with SQL
A Real-World Case Study

Billing

InvoiceNumber	BillingDate	MaterialCode	CustomerCode	SellingPrice	Quantity	SalesValue
1183454841	January 1, 2020	23002342	15027792	21.5	80	1,722
1183455139	January 1, 2020	23002343	15204987	40	2	80
1183455416	January 2, 2020	23033619	15026117	38	202	7,692
1183455427	January 2, 2020	23033619	15103959	66	4	265
1183455572	January 2, 2020	23002344	15027191	17	48	822
1183455582	January 2, 2020	23002344	15028928	17.5	4	70
1183455755	January 2, 2020	23002342	15027419	20	140	2,826
1183456039	January 2, 2020	23033619	15031290	66	4	265

Label

- Primary Key
- Foreign Key
- Attribute

Product

MaterialCode	MaterialDescription	SupplierID	SupplierName	ProductGroup
23002342	CIMLUP 100MG 10S	100753	FedEx	CIMLUP
23002343	CIMLUP 200MG 10S	100753	FedEx	CIMLUP
23002344	CIMLUP 200MG 20S	100753	FedEx	CIMLUP
23002415	ASSERI 100ML	100797	UPS	ASSERI
23033619	NILIRB 50MG/ML 10	100797	UPS	NILIRB

Customer

CustomerCode	CustomerName	CustomerRegion	CustomerChannel
15025979	BWLLY	West	Email
15026117	CLLDY	South	Social Networks
15026118	CDEVY	West	Social Networks
15026126	CDGLY	West	Social Networks

SalesMan

CustomerRegion	SalesManID	SalesManName
West	T101	Bob
South	T102	John
East	T101	Bob
North	T103	Kelly
Central	T104	Sarah

PART 4: Data Wrangling with SQL
A Real-World Case Study

As a Data Analyst in a Sales Excellence team, your manager would like you to **evaluate sales performance** and **identify growth opportunities** for your products.

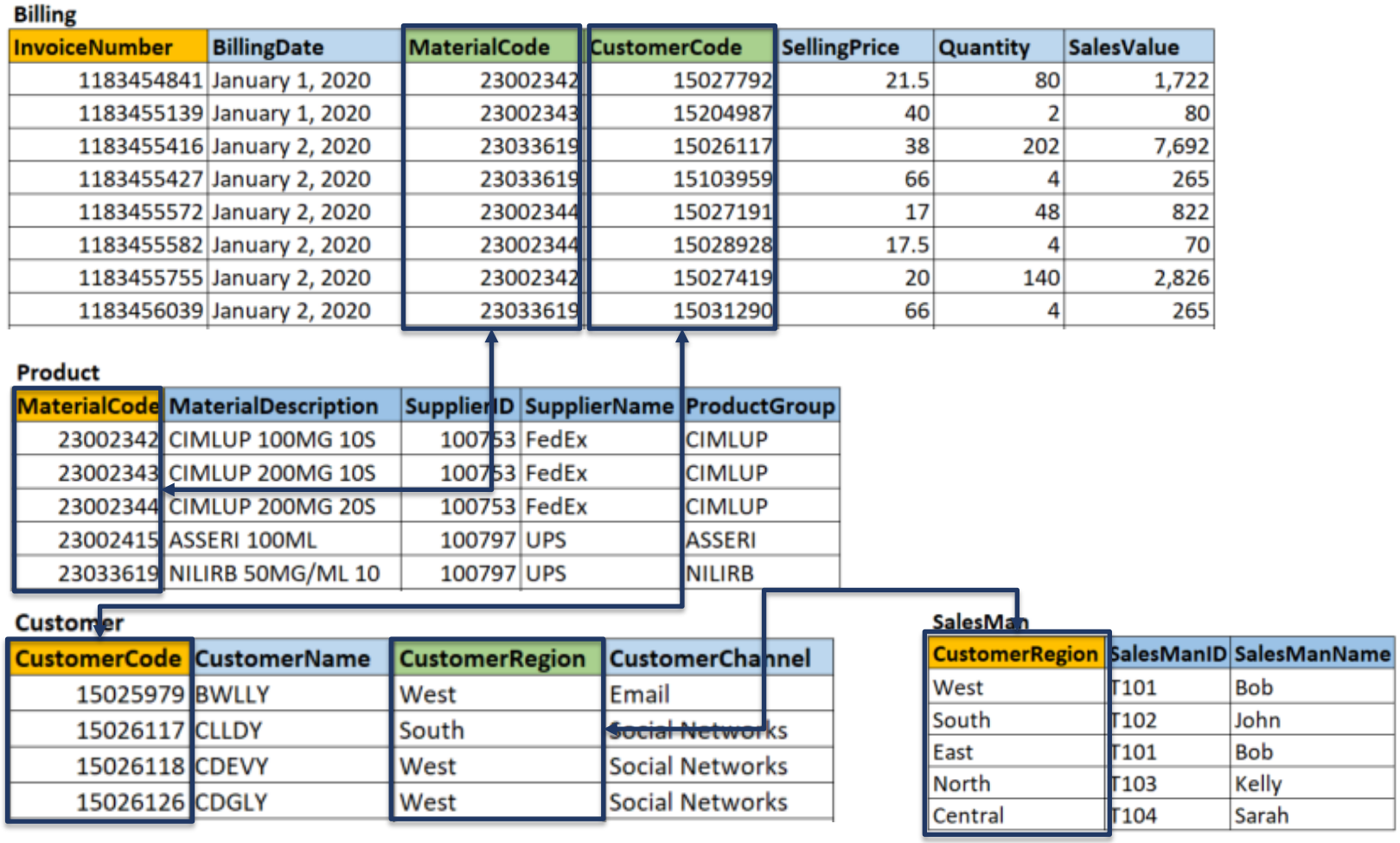
PART 4: Data Wrangling with SQL
A Real-World Case Study

As a Data Analyst in a Sales Excellence team, your manager would like you to **evaluate sales performance** and **identify growth opportunities for your products**.

List down some of the key business questions that you would like to address.

1. Who are the top 5 customers in sales value for the past 6 months?
2. Which region has the largest sales growth in Q2 vs Q1?
3. Which customer channel is the largest channel? And what percentage it contributes to the total sales?
4. How many customers are new customers, who only started to purchase in most recent 1 month? List down the new customer names.
5. Show each Salesman's sales size for each product group and its growth rate between Q2 and Q1.

PART 4: Data Wrangling with SQL
A Real-World Case Study



Billing

InvoiceNumber	BillingDate	MaterialCode	CustomerCode	SellingPrice	Quantity	SalesValue
1183454841	January 1, 2020	23002342	15027792	21.5	80	1,722
1183455139	January 1, 2020	23002343	15204987	40	2	80
1183455416	January 2, 2020	23033619	15026117	38	202	7,692
1183455427	January 2, 2020	23033619	15103959	66	4	265
1183455572	January 2, 2020	23002344	15027191	17	48	822
1183455582	January 2, 2020	23002344	15028928	17.5	4	70
1183455755	January 2, 2020	23002342	15027419	20	140	2,826
1183456039	January 2, 2020	23033619	15031290	66	4	265

Product

MaterialCode	MaterialDescription	SupplierID	SupplierName	ProductGroup
23002342	CIMLUP 100MG 10S	100753	FedEx	CIMLUP
23002343	CIMLUP 200MG 10S	100753	FedEx	CIMLUP
23002344	CIMLUP 200MG 20S	100753	FedEx	CIMLUP
23002415	ASSERI 100ML	100797	UPS	ASSERI
23033619	NILIRB 50MG/ML 10	100797	UPS	NILIRB

Customer

CustomerCode	CustomerName	CustomerRegion	CustomerChannel
15025979	BWLLY	West	Email
15026117	CLLDY	South	Social Networks
15026118	CDEVY	West	Social Networks
15026126	CDGLY	West	Social Networks

SalesMan

CustomerRegion	SalesManID	SalesManName
West	T101	Bob
South	T102	John
East	T101	Bob
North	T103	Kelly
Central	T104	Sarah

Label

- Primary Key
- Foreign Key
- Attribute

PART 4: Data Wrangling with SQL
A Real-World Case Study

1. Who are the **top 5 customers** in **sales value** for the **past 6 months**?

Billing

InvoiceNumber	BillingDate	MaterialCode	CustomerCode	SellingPrice	Quantity	SalesValue
1183454841	January 1, 2020	23002342	15027792	21.5	80	1,722
1183455139	January 1, 2020	23002343	15204987	40	2	80
1183455416	January 2, 2020	23033619	15026117	38	202	7,692
1183455427	January 2, 2020	23033619	15103959	66	4	265
1183455572	January 2, 2020	23002344	15027191	17	48	822
1183455582	January 2, 2020	23002344	15028928	17.5	4	70
1183455755	January 2, 2020	23002342	15027419	20	140	2,826
1183456039	January 2, 2020	23033619	15031290	66	4	265

Label
- Primary Key
- Foreign Key
- Attribute

Product

MaterialCode	MaterialDescription	SupplierID	SupplierName	ProductGroup
23002342	CIMLUP 100MG 10S	100753	FedEx	CIMLUP
23002343	CIMLUP 200MG 10S	100753	FedEx	CIMLUP
23002344	CIMLUP 200MG 20S	100753	FedEx	CIMLUP
23002415	ASSERI 100ML	100797	UPS	ASSERI
23033619	NILIRB 50MG/ML 10	100797	UPS	NILIRB

Customer

CustomerCode	CustomerName	CustomerRegion	CustomerChannel
15025979	BWLLY	West	Email
15026117	CLLDY	South	Social Networks
15026118	CDEVY	West	Social Networks
15026126	CDGLY	West	Social Networks

SalesMan

CustomerRegion	SalesManID	SalesManName
West	T101	Bob
South	T102	John
East	T101	Bob
North	T103	Kelly
Central	T104	Sarah

PART 4: Data Wrangling with SQL
A Real-World Case Study

1. Who are the top 5 **customers** in sales value for the past 6 months?

Analysis

- Tops 5 and Past 6 months are two constraints in the questions. We can use **ORDER BY, LIMIT** and **BETWEEN AND** statement.
- A customer name might have one or more customer codes. Therefore, we need to firstly **JOIN** the **Billing** table and **Customer** table, and then rank the customer names by sales values.

Solution

```
SELECT Customer.CustomerName,SUM(Billing.SalesValue) AS "Sales Value"
FROM Billing
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
WHERE BillingDate BETWEEN '2020-01-01' AND '2020-06-30'
GROUP BY Customer.CustomerName
ORDER BY SUM(Billing.SalesValue) DESC
LIMIT 5;
```

```
SELECT Customer.CustomerName,SUM(Billing.SalesValue) AS "Sales Value"
FROM Billing
LEFT JOIN Customer
ON  Billing.CustomerCode = Customer.CustomerCode
WHERE BillingDate BETWEEN
    (SELECT DATE_ADD(MAX(BillingDate), INTERVAL -6 MONTH) FROM Billing)
    AND (SELECT MAX(BillingDate) FROM Billing)
GROUP BY Customer.CustomerName
ORDER BY SUM(Billing.SalesValue) DESC
LIMIT 5;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

1. What are the top 5 **materials** in sales value for the past 6 months?

Analysis

- Tops 5 and Past 6 months are two constraints in the questions. We can use **ORDER BY, LIMIT** and **BETWEEN AND** statement.
- A material description might have one or more material codes. Therefore, we need to firstly **JOIN** the **Billing** table and **Product** table, and then rank the **material description** by sales values.

Solution

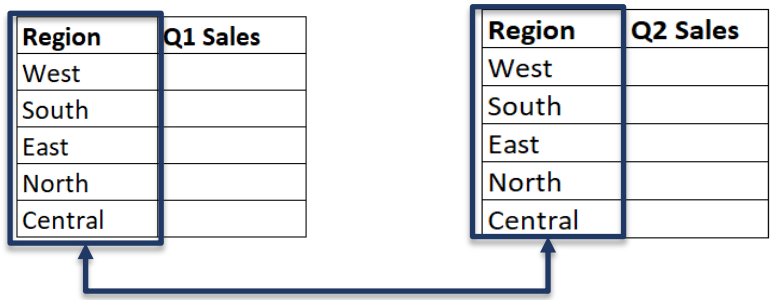
```
SELECT Product.MaterialDescription,SUM(Billing.SalesValue) AS "Sales
Value" FROM Billing
LEFT JOIN Product
ON Billing.MaterialCode = Product.MaterialCode
WHERE BillingDate BETWEEN '2020-01-01' AND '2020-06-30'
GROUP BY Product.MaterialCode
ORDER BY SUM(Billing.SalesValue) DESC
LIMIT 5;
```

```
SELECT Product. MaterialDescription,SUM(Billing.SalesValue) AS "Sales Value"
FROM Billing
LEFT JOIN Product
ON  Billing. MaterialCode = Product. MaterialCode
WHERE BillingDate BETWEEN
    (SELECT DATE_ADD(MAX(BillingDate), INTERVAL -6 MONTH) FROM Billing)
    AND (SELECT MAX(BillingDate) FROM Billing)
GROUP BY Product.MaterialCode
ORDER BY SUM(Billing.SalesValue) DESC
LIMIT 5;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

2. Which region has the largest sales growth in Q2 vs Q1?

First, we need to get the sales in Q2 and Q1 by region. GROUP BY and BETWEEN AND can be used. We will have two tables.



Region	Q1 Sales
West	
South	
East	
North	
Central	

Region	Q2 Sales
West	
South	
East	
North	
Central	

Solution

```
SELECT Customer.CustomerRegion,SUM(Billing.SalesValue) AS "SalesValue" FROM Billing
    LEFT JOIN Customer
    ON  Billing.CustomerCode = Customer.CustomerCode
    WHERE BillingDate BETWEEN '2020-04-01' AND '2020-06-30'
    GROUP BY Customer.CustomerRegion
```

```
SELECT Customer.CustomerRegion,SUM(Billing.SalesValue) AS "SalesValue" FROM Billing
    LEFT JOIN Customer
    ON  Billing.CustomerCode = Customer.CustomerCode
    WHERE BillingDate BETWEEN '2020-01-01' AND '2020-03-31'
    GROUP BY Customer.CustomerRegion
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

2. Which region has the largest sales growth in Q2 vs Q1?

Second, we can join the table by region. And calculate the growth rate.

Region	Q1 Sales
West	
South	
East	
North	
Central	

Region	Q2 Sales
West	
South	
East	
North	
Central	

Region	Q1 Sales	Q2 Sales	Growth
West			
South			
East			
North			
Central			

Solution

```
SELECT q2.CustomerRegion, q2.SalesValue AS "q2_Sales",q1.SalesValue AS
"q1_Sales", (q2.SalesValue/q1.SalesValue - 1) AS "Q2vsQ1_Growth"
FROM Q2_Sales_Table AS q2
LEFT JOIN Q1_Sales_Table AS q1
ON q2.CustomerRegion = q1.CustomerRegion
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

2. Which region has the largest sales growth in Q2 vs Q1?

Analysis

- The growth rate is defined by (Sales in Q2 – Sales in Q1)/Sales in Q1. First, we need to get the sales in Q2 and Q1 by region. GROUP BY and BETWEEN AND can be used. We will have two tables.

Region	Q1 Sales
West	
South	
East	
North	
Central	

Region	Q2 Sales
West	
South	
East	
North	
Central	

- Second, we can join the table by region. And calculate the growth rate.

Region	Q1 Sales	Q2 Sales	Growth
West			
South			
East			
North			
Central			

- Then ORDER BY Growth for the second step result.

Solution

```
SELECT q2.CustomerRegion, q2.SalesValue AS "q2_Sales",q1.SalesValue AS "q1_Sales",
(q2.SalesValue/q1.SalesValue - 1) AS "Q2vsQ1_Growth"  FROM
    (SELECT Customer.CustomerRegion,SUM(Billing.SalesValue) AS "SalesValue" FROM Billing
    LEFT JOIN Customer
    ON  Billing.CustomerCode = Customer.CustomerCode
    WHERE BillingDate BETWEEN '2020-04-01' AND '2020-06-30'
    GROUP BY Customer.CustomerRegion
    ) AS q2
LEFT JOIN
 (
    SELECT Customer.CustomerRegion,SUM(Billing.SalesValue) AS "SalesValue" FROM Billing
    LEFT JOIN Customer
    ON  Billing.CustomerCode = Customer.CustomerCode
    WHERE BillingDate BETWEEN '2020-01-01' AND '2020-03-31'
    GROUP BY Customer.CustomerRegion
 ) AS q1
 ON q2.CustomerRegion = q1.CustomerRegion
 ORDER BY (q2.SalesValue/q1.SalesValue - 1) DESC;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

3. Which channel is the largest channel? What % it contributes to the total sales?

- First, we can use GROUP BY and ORDER BY to get the each channel sales. And the total sales value.

CustomerChannel	SalesValue
Email	
Social Networks	
Phone	
Search Engine	
Door Visit	
Reviews	
Directory	

SalesValue

Solution

```
SELECT Customer.CustomerChannel,SUM(Billing.SalesValue) AS "SalesValue"
FROM Billing
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
GROUP BY Customer.CustomerChannel
ORDER BY SUM(Billing.SalesValue) DESC
```

```
SELECT SUM(Billing.SalesValue) AS "SalesValue" FROM Billing
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

3. Which channel is the largest channel? What % it contributes to the total sales?

- Secondly, we can join the two tables by a dummy key to generate a table with two columns: individual_channel_sales, and total_sales.

CustomerChannel	SalesValue	DummyKey
Email		1
Social Networks		1
Phone		1
Search Engine		1
Door Visit		1
Reviews		1
Directory		1

DummyKey	SalesValue
1	

CustomerChannel	individual_sales	DummyKey	total_sales
Email		1	
Social Networks		1	
Phone		1	
Search Engine		1	
Door Visit		1	
Reviews		1	
Directory		1	

Solution

```
SELECT Customer.CustomerChannel,SUM(Billing.SalesValue) AS "SalesValue",
'1' AS "DummyKey" FROM Billing
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
GROUP BY Customer.CustomerChannel
ORDER BY SUM(Billing.SalesValue) DESC;


SELECT SUM(Billing.SalesValue) AS "TotalSales",
'1' AS "DummyKey" FROM Billing;


SELECT lt.CustomerChannel, lt.SalesValue, rt.SalesValue AS "TotalSales"
FROM
Channel_Sales_Table AS lt
LEFT JOIN
Total_Sales_Table AS rt
ON lt.DummyKey = rt.DummyKey;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

3. Which channel is the largest channel? What % it contributes to the total sales?

- Final step, we can get the contribution by dividing individual channel sales with total sales. And then rank the contribution.

CustomerChannel	individual_sales	DummyKey	total_sales	contribution
Email		1		
Social Networks		1		
Phone		1		
Search Engine		1		
Door Visit		1		
Reviews		1		
Directory		1		

Solution

```
SELECT lt.CustomerChannel, lt.SalesValue, rt.SalesValue AS
"TotalSales",lt.SalesValue/rt.SalesValue AS "Contribution"
FROM
 (
 SELECT Customer.CustomerChannel,SUM(Billing.SalesValue) AS "SalesValue",
 '1'   AS "DummyKey" FROM Billing
 LEFT JOIN Customer
 ON Billing.CustomerCode = Customer.CustomerCode
 GROUP BY Customer.CustomerChannel
 ORDER BY SUM(Billing.SalesValue) DESC
 ) AS lt
LEFT JOIN
 (
 SELECT SUM(Billing.SalesValue) AS "SalesValue", '1' AS "DummyKey"
 FROM Billing
 ) AS rt
ON lt.DummyKey = rt.DummyKey
ORDER BY lt.SalesValue/rt.SalesValue DESC;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

4. New customer list who started to purchase in most recent 1 month

- First, we need to get the list of customers who purchased in recent 1 month. We can use WHERE condition to filter based on billing date.

- Second, we should also get the list of customers who made purchase in past months (more than 1 month ago).

Solution

```
SELECT Customer.CustomerName,SUM(Billing.SalesValue) AS "SalesValue"
FROM Billing
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
WHERE Billing.BillingDate BETWEEN '2020-06-01' AND '2020-06-30'
GROUP BY Customer.CustomerName
ORDER BY SUM(Billing.SalesValue) DESC
```

```
SELECT DISTINCT Customer.CustomerName FROM Billing
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
WHERE Billing.BillingDate <='2020-05-31'
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

4. New customer list who started to purchase in most recent 1 month

- Thirdly, we can exclude the customers who purchased (more than 1 month ago) from the list which purchased in recent 1 month. NOT IN operators can help to exclude results.

Solution

```
SELECT t.CustomerName AS "NewCustomer", t.SalesValue FROM (
  SELECT Customer.CustomerName,SUM(Billing.SalesValue) AS "SalesValue"
  FROM Billing
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  WHERE Billing.BillingDate BETWEEN '2020-06-01' AND '2020-06-30'
  GROUP BY Customer.CustomerName
  ORDER BY SUM(Billing.SalesValue) DESC ) AS t
WHERE
CustomerName NOT IN
(
  SELECT DISTINCT Customer.CustomerName FROM Billing
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  WHERE Billing.BillingDate <='2020-05-31'
)
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

4. Lost customer list who purchased before (> 3 months ago), but stopped purchasing in recent 3 months

- We can just modify the WHERE condition and change the filters.

Solution

```
SELECT t.CustomerName AS “LostCustomer", t.SalesValue FROM (
  SELECT Customer.CustomerName,SUM(Billing.SalesValue) AS "SalesValue"
  FROM Billing
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  WHERE Billing.BillingDate <= '2020-03-31'
  GROUP BY Customer.CustomerName
  ORDER BY SUM(Billing.SalesValue) DESC ) AS t
WHERE
CustomerName NOT IN
(
  SELECT DISTINCT Customer.CustomerName FROM Billing
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  WHERE Billing.BillingDate >='2020-04-01'
)
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

5. Show each Salesman's sales size for each product group and its growth rate between Q2 and Q1.

Billing Table

InvoiceNumber	BillingDate	MaterialCode	CustomerCode	SellingPrice	Quantity	SalesValue
1183454841	January 1, 2020	23002342	15027792	21.5	80	1,722
1183455139	January 1, 2020	23002343	15204987	40	2	80
1183455416	January 2, 2020	23033619	15026117	38	202	7,692
1183455427	January 2, 2020	23033619	15103959	66	4	265
1183455572	January 2, 2020	23002344	15027191	17	48	822
1183455582	January 2, 2020	23002344	15028928	17.5	4	70
1183455755	January 2, 2020	23002342	15027419	20	140	2,826
1183456039	January 2, 2020	23033619	15031290	66	4	265

Product

MaterialCode	MaterialDescription	SupplierID	SupplierName	ProductGroup
23002342	CIMLUP 100MG 10S	100753	FedEx	CIMLUP
23002343	CIMLUP 200MG 10S	100753	FedEx	CIMLUP
23002344	CIMLUP 200MG 20S	100753	FedEx	CIMLUP
23002415	ASSERI 100ML	100797	UPS	ASSERI
23033619	NILIRB 50MG/ML 10	100797	UPS	NILIRB

Customer

CustomerCode	CustomerName	CustomerRegion	CustomerChannel
15025979	BWLLY	West	Email
15026117	CLLDY	South	Social Networks
15026118	CDEVY	West	Social Networks
15026126	CDGLY	West	Social Networks

SalesMan

CustomerRegion	SalesManID	SalesManName
West	T101	Bob
South	T102	John
East	T101	Bob
North	T103	Kelly
Central	T104	Sarah

Result Table

SalesMan	ProductGroup	Growth Rate
Kelly	DUOXIB	0.37
John	DUOXIB	0.19
Sarah	ASSERI	0
Bob	ASSERI	-0.04
Bob	NILIRB	-0.05

PART 4: Data Wrangling with SQL
A Real-World Case Study

5. Show each Salesman's sales size for each product group and its growth rate between Q2 and Q1.

- Firstly, we need to join the three tables to link the product group and salesman into a single table. It's easier for us to do data wrangling after building the table relationship.
- Billing table is the base table, and then we left join with Product, Customer tables on Material Code and Customer Code. From Customer Region, left join again with SalesMan table on Customer Region.

Solution

```
SELECT Billing.BillingDate,Product.ProductGroup,
SalesMan.SalesManName,SUM(Billing.SalesValue) AS "SalesValue"
FROM Billing
LEFT JOIN Product
ON Billing.MaterialCode = Product. MaterialCode
LEFT JOIN Customer
ON Billing.CustomerCode = Customer.CustomerCode
LEFT JOIN SalesMan
ON Customer.CustomerRegion = SalesMan.CustomerRegion
GROUP BY Billing.BillingDate,Product.ProductGroup, SalesMan.SalesManName;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

5. Show each Salesman's sales size for each product group and its growth rate between Q2 and Q1.

- Second, we have known how to get the quarterly growth rate in (2). We can replace it with the table we just generated. However, the join key should include both **ProductGroup** and **SalesManName**.

SalesMan	ProductGroup	Q2 Sales
Kelly	DUOXIB	
John	DUOXIB	
Sarah	ASSERI	
Bob	ASSERI	
Bob	NILIRB	

SalesMan	ProductGroup	Q1 Sales
Kelly	DUOXIB	
John	DUOXIB	
Sarah	ASSERI	
Bob	ASSERI	
Bob	NILIRB	

```
SELECT q2.CustomerRegion, q2.SalesValue AS "q2_Sales",q1.SalesValue AS "q1_Sales",
(q2.SalesValue/q1.SalesValue - 1) AS "Q2vsQ1_Growth"  FROM
Q2_Sales_Table AS q2
LEFT JOIN
Q1_Sales_Table AS q1
ON q2.ProductGroup = q1.ProductGroup AND q2.SalesManName = q1.SalesManName
ORDER BY (q2.SalesValue/q1.SalesValue - 1) DESC;
```

Solution

```
SELECT q2.ProductGroup, q2.SalesManName,SUM(q2.SalesValue) AS "Q2_SalesValue", SUM(q1.SalesValue) AS
"Q1_SalesValue",SUM(q2.SalesValue)/SUM(q1.SalesValue) -1 AS "GrowthRate" FROM
(
  SELECT Product.ProductGroup, SalesMan.SalesManName,SUM(Billing.SalesValue) AS "SalesValue"
  FROM Billing
  LEFT JOIN Product
  ON Billing.MaterialCode = Product. MaterialCode
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  LEFT JOIN SalesMan
  ON Customer.CustomerRegion = SalesMan.CustomerRegion
  WHERE  Billing.BillingDate BETWEEN '2020-04-01' AND '2020-06-30'
  GROUP BY Product.ProductGroup, SalesMan.SalesManName
) AS q2
LEFT JOIN
(
  SELECT Product.ProductGroup, SalesMan.SalesManName,SUM(Billing.SalesValue) AS "SalesValue"
  FROM Billing
  LEFT JOIN Product
  ON Billing.MaterialCode = Product. MaterialCode
  LEFT JOIN Customer
  ON Billing.CustomerCode = Customer.CustomerCode
  LEFT JOIN SalesMan
  ON Customer.CustomerRegion = SalesMan.CustomerRegion
  WHERE  Billing.BillingDate BETWEEN '2020-01-01' AND '2020-03-31'
  GROUP BY Product.ProductGroup, SalesMan.SalesManName
) AS q1
ON
q2.ProductGroup = q1.ProductGroup AND q2.SalesManName = q1.SalesManName
GROUP BY q2.ProductGroup, q2.SalesManName
ORDER BY SUM(q2.SalesValue)/SUM(q1.SalesValue) -1 DESC;
```

PART 4: Data Wrangling with SQL
A Real-World Case Study

As a Data Analyst in a Sales Excellence team, your manager would like you to **evaluate sales performance** and **identify growth opportunities for your products**.

List down some of the key business questions that you would like to address.

1. Who are the top 5 customers in sales value for the past 6 months?
2. Which region has the largest sales growth in Q2 vs Q1?
3. Which customer channel is the largest channel? And what percentage it contributes to the total sales?
4. How many customers are new customers, who only started to purchase in most recent 1 month? List down the new customer names.
5. Show each Salesman's sales size for each product group and its growth rate between Q2 and Q1.