CUSTOMER, PRODUCT AND SALES ANALYTICS USING SQL

OBJECTIVE OF THE PROJECT:

The project has the primary objective of conducting an in-depth analysis of the sales and customer data of a scale model car business based on a dataset that contains details of orders of customers from January of 2003 till May of 2005. This analysis is intended to provide the company with valuable insights into the overall sales performance, customer behaviour, and employee structure within the organization.
Through this project, we aim to identify the key drivers of the company's sales performance, as well as the most successful products and product lines.
Additionally, the project will investigate the employee structure within the company, including identifying top-performing sales employees.

ABOUT THE DATASET:

The data for the project will be sourced from several tables within the schema, including:

- ProductLines: This table provides a list of the different categories or product lines to which the scale model cars belong.

- Products: This table contains a list of all the scale model cars available for purchase, including their names, descriptions, prices and vendor.

- Offices: This table stores data about the sales offices of the business, including their locations and contact information.

- Employees: This table contains information about the employees of the business, including their names, contact information, and job titles. Here EmployeeNumber is the primary key. This table contains a foreign key ‘OfficeCode’ which connects the customer table to the employee table.

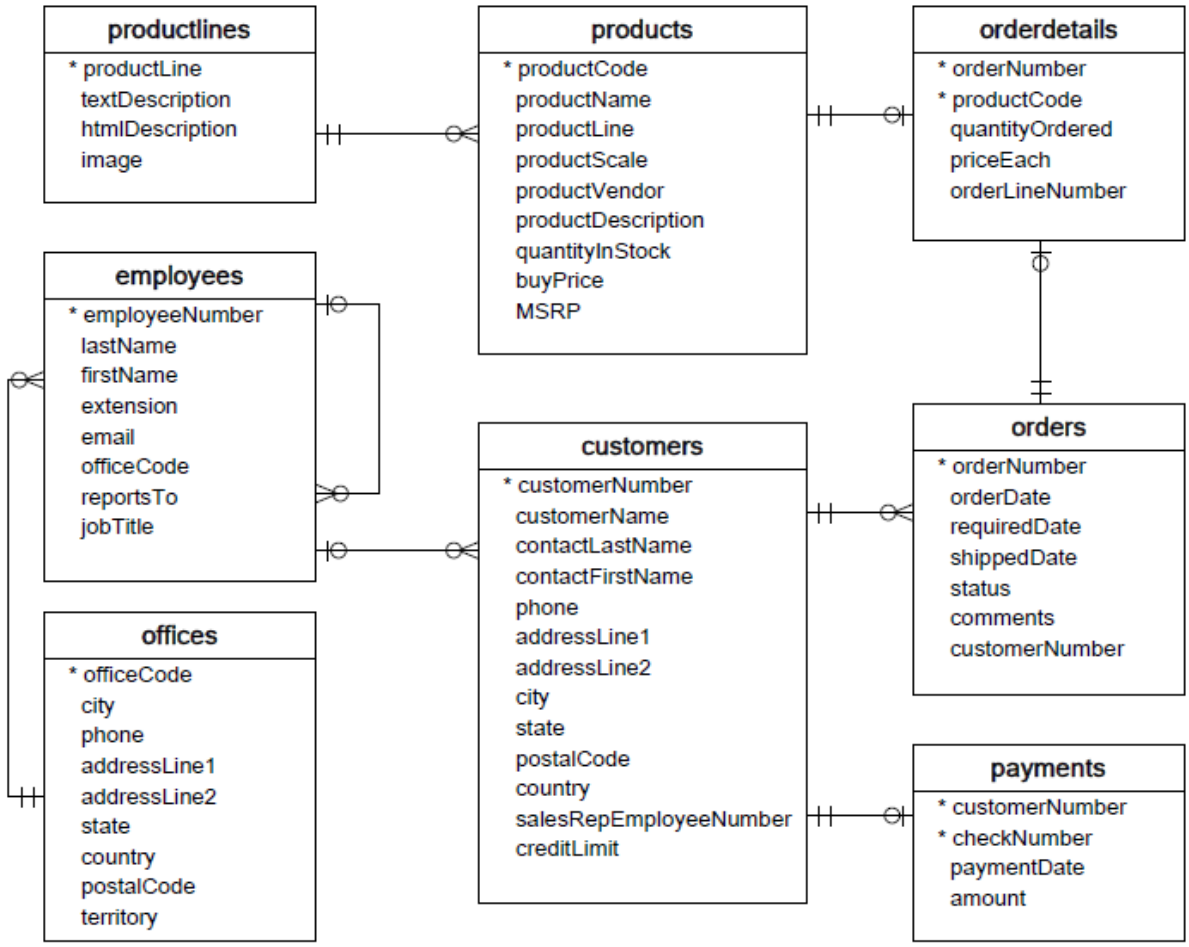
- Customers: This table stores information about the customers of the business, including their names, addresses, contact information and the employee that handles that customer. Here CustomerNumber is the primary key. This table contains a foreign key ‘SalesRepEmployeeNumber’ which connects the customer table to the employee table.

- Orders: This table stores sales order data, including the date of the order, the customer who placed the order, order date and shipping date. Here OrderNumber is the primary key. This table contains a foreign key ‘CustomerNumber’ which connects the orders table to the customers table.

- OrderDetails: This table contains information about the line items included in each sales order, including the product purchased, the quantity, and the price. This table contains two foreign keys ‘OrderNumber’ and ‘ProductCode’ which connects the OrderDetails table to the orders table and products table.

- Payments: This table stores information about customer payments made against their account, including the payment date and the amount paid. This table contains a foreign key ‘CustomerNumber’ which connects the payments table to the customers table.

By utilizing data from these tables, we will be able to conduct a comprehensive analysis of the sales and customer data of the business, which will provide us with valuable insights into the company's performance, as well as the behaviour of its customers and employees.



QUESTIONS ANSWERED:

The questions that the project aims to answer include:

- How many orders were placed by each customer?
- Categorise customers into three categories: ‘churn’, ‘at risk’ and ‘Loyal or Retained’
- Which customer has the highest total purchase value?
- Which customer has never ordered?
- Which order has the max order value and which customer purchased it and which employee took the order?
- Which year was the best for sales?

- How much in %age, the sales had increased or decreased over the months in 2004?
- What is the median sales?
- Who are the top buyer from each state?
- Number of new customers over the years?
- Which product line has the highest total sales?
- What are the top most ordered products by quantity sold and total profit of each?
- How many products were sold in each quarter?
- Which products are most popular in each of the country?
- Who are the top-performing sales employees?

INSIGHTS GAINED FROM THE PROJECT:

The insights that the project aims to gain from the data include:

1) Identifying the most profitable product lines and products
2) Understanding the factors that impact sales performance
3) Identifying the most effective sales employees and strategies
4) Segmenting customers for targeted marketing and sales efforts
5) Identifying opportunities to improve payment and account management processes.

RANGE OF SQL FUNCTIONS USED:

Joins, case and when, sub-query, group by, order by, lag function, median, windows function, common table expression (CTE), date functions, concat, and with rollup

1. **How many orders were placed by each customer?**

```
SELECT C.CustomerNumber, C.customerName,
    COUNT(O.ORDERNumber) AS total_orders_made
FROM CUSTOMERS C
INNER JOIN ORDERS O
ON O.CustomerNumber = C.CustomerNumber
GROUP BY C.CustomerNumber
ORDER BY total_orders_made desc;
```

- EURO+SHOPPING CHANNEL HAS THE MOST NUMBER OF ORDERS

Result Grid Filter Rows: Search Export:

CustomerNumber	customerName	total_orders_made
141	Euro+ Shopping Channel	26
124	Mini Gifts Distributors Ltd.	17
353	Reims Collectables	5
323	Down Under Souveniers, Inc	5
114	Australian Collectors, Co.	5
148	Dragon Souveniers, Ltd.	5
145	Danish Wholesale Imports	5
496	Kelly's Gift Shop	4
131	Land of Toys Inc.	4
128	Blauer See Auto, Co.	4

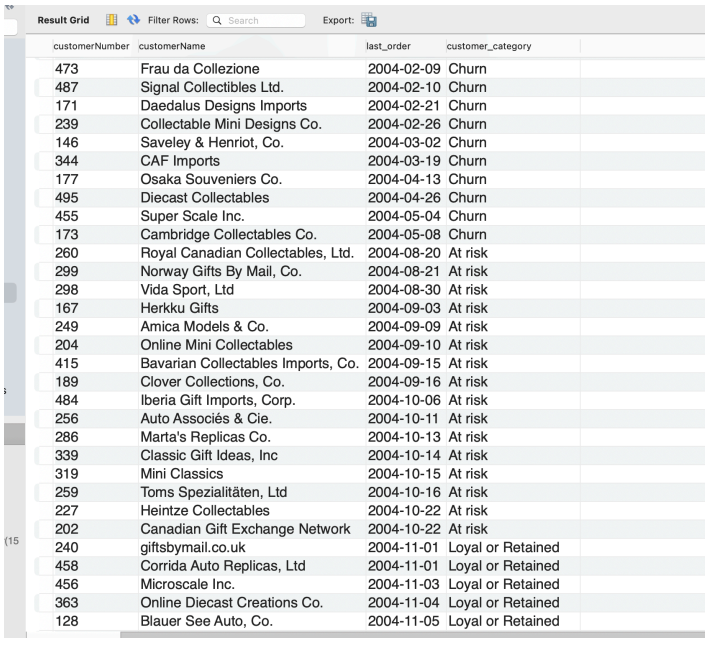
2. **Categorise customers into three categories: ‘churn’, ‘at risk’ and ‘Loyal or Retained’**

   Churn: Customers who have stopped doing business with the company. If a customer’s last purchase date is more than a year ago from May 2005 (i.e. before 2004-05-31) then they are categorised as ‘churn’.

   At risk: Customers who are showing signs that they may stop doing business with the company in the near future. If a customer’s last purchase date is more 7 months ago from May 2005 (i.e. between 2004-06-1 and 2004-10-31) then they are categorised as ‘at risk’.

   Loyal or Retained: Customers who continue to do business with the company and are satisfied with its products or services. Customers who have purchased in the last 7 months are loyal.

```
SELECT c.customerNumber,
c.customerName, max(o.orderdate) as last_order,
CASE
WHEN max(o.orderdate) <= '2004-05-31' THEN 'Churn'
WHEN max(o.orderdate) between '2004-06-01' and '2004-10-31' THEN 'At risk'
ELSE 'Loyal or Retained'
END AS customer_category
FROM CUSTOMERS C
INNER JOIN ORDERS O
ON O.CustomerNumber = C.CustomerNumber
GROUP BY C.CustomerNumber
ORDER BY last_order;
```



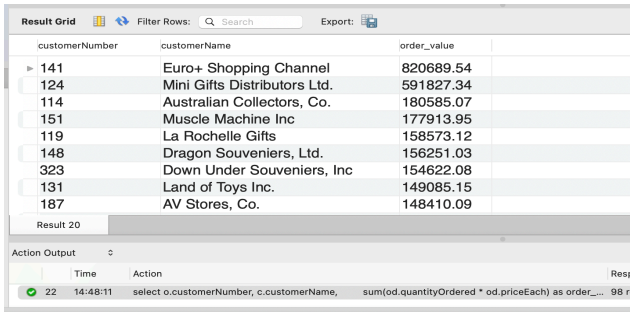
Result Grid Filter Rows: Search Export:

customerNumber	customerName	last_order	customer_category
473	Frau da Collezione	2004-02-09	Churn
487	Signal Collectibles Ltd.	2004-02-10	Churn
171	Daedalus Designs Imports	2004-02-21	Churn
239	Collectable Mini Designs Co.	2004-02-26	Churn
146	Saveley & Henriot, Co.	2004-03-02	Churn
344	CAF Imports	2004-03-19	Churn
177	Osaka Souveniers Co.	2004-04-13	Churn
495	Diecast Collectables	2004-04-26	Churn
455	Super Scale Inc.	2004-05-04	Churn
173	Cambridge Collectables Co.	2004-05-08	Churn
260	Royal Canadian Collectables, Ltd.	2004-08-20	At risk
299	Norway Gifts By Mail, Co.	2004-08-21	At risk
298	Vida Sport, Ltd	2004-08-30	At risk
167	Herkku Gifts	2004-09-03	At risk
249	Amica Models & Co.	2004-09-09	At risk
204	Online Mini Collectables	2004-09-10	At risk
415	Bavarian Collectables Imports, Co.	2004-09-15	At risk
189	Clover Collections, Co.	2004-09-16	At risk
484	Iberia Gift Imports, Corp.	2004-10-06	At risk
256	Auto Associés & Cie.	2004-10-11	At risk
286	Marta's Replicas Co.	2004-10-13	At risk
339	Classic Gift Ideas, Inc	2004-10-14	At risk
319	Mini Classics	2004-10-15	At risk
259	Toms Spezialitäten, Ltd	2004-10-16	At risk
227	Heintze Collectables	2004-10-22	At risk
202	Canadian Gift Exchange Network	2004-10-22	At risk
240	giftsbymail.co.uk	2004-11-01	Loyal or Retained
458	Corrida Auto Replicas, Ltd	2004-11-01	Loyal or Retained
456	Microscale Inc.	2004-11-03	Loyal or Retained
363	Online Diecast Creations Co.	2004-11-04	Loyal or Retained
128	Blauer See Auto, Co.	2004-11-05	Loyal or Retained

3. **Which customer has the highest total purchase value?**

```
select o.customerNumber, c.customerName,
    sum(od.quantityOrdered * od.priceEach) as order_value
from orderdetails od
inner join orders o
on od.ordernumber = o.ordernumber
inner join customers c
on o.customerNumber = c.customerNumber
group by o.customerNumber
order by order_value desc;
```

- EURO+SHOPPING CHANNEL HAS THE HIGHEST ORDER VALUE



customerNumber	customerName	order_value
141	Euro+ Shopping Channel	820689.54
124	Mini Gifts Distributors Ltd.	591827.34
114	Australian Collectors, Co.	180585.07
151	Muscle Machine Inc	177913.95
119	La Rochelle Gifts	158573.12
148	Dragon Souveniers, Ltd.	156251.03
323	Down Under Souveniers, Inc	154622.08
131	Land of Toys Inc.	149085.15
187	AV Stores, Co.	148410.09

Result 20

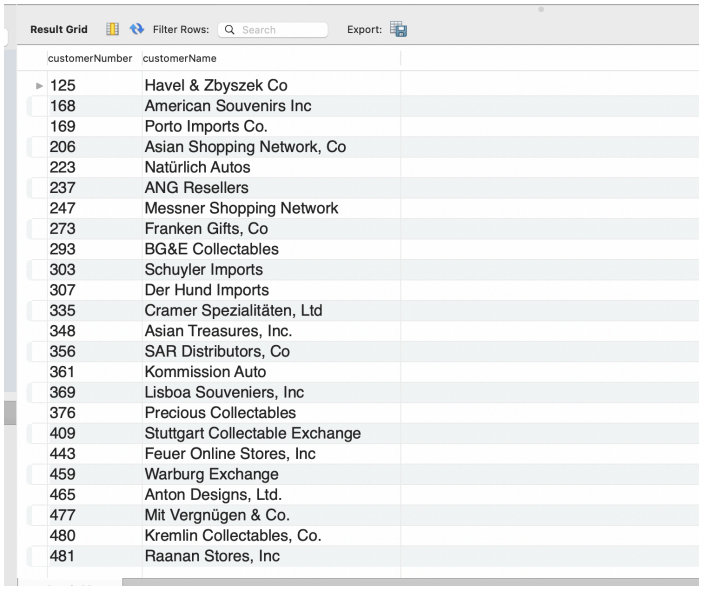
Action Output

	Time	Action	Resp
22	14:48:11	select o.customerNumber, c.customerName, sum(od.quantityOrdered * od.priceEach) as order_...	98 r

4. **Which customer has never ordered?**

```
select c.customerNumber, c.customerName
from customers c
left join orders o
on c.customerNumber = o.customerNumber
where o.customerNumber is null;
```

- There are 24 customers who have never ordered

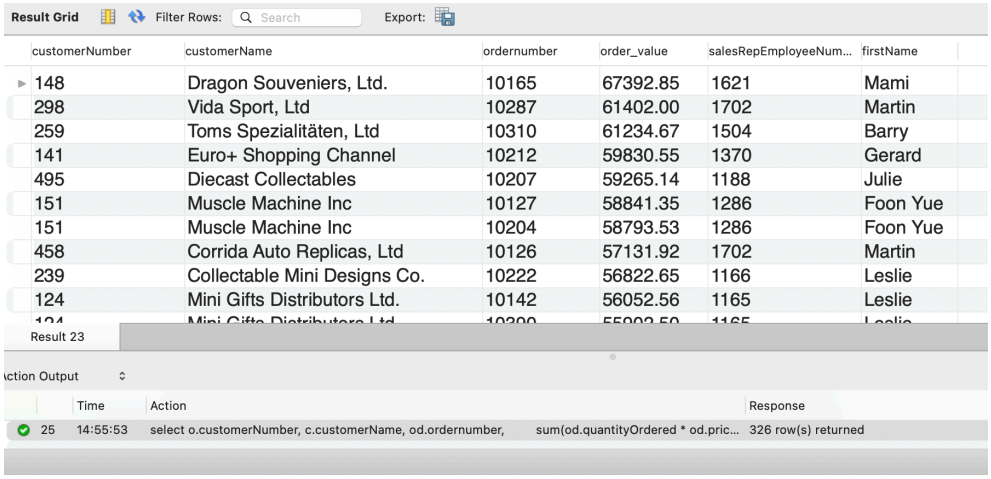


customerNumber	customerName
125	Havel & Zbyszek Co
168	American Souvenirs Inc
169	Porto Imports Co.
206	Asian Shopping Network, Co
223	Natürlich Autos
237	ANG Resellers
247	Messner Shopping Network
273	Franken Gifts, Co
293	BG&E Collectables
303	Schuyler Imports
307	Der Hund Imports
335	Cramer Spezialitäten, Ltd
348	Asian Treasures, Inc.
356	SAR Distributors, Co
361	Kommission Auto
369	Lisboa Souveniers, Inc
376	Precious Collectables
409	Stuttgart Collectable Exchange
443	Feuer Online Stores, Inc
459	Warburg Exchange
465	Anton Designs, Ltd.
477	Mit Vergnügen & Co.
480	Kremlin Collectables, Co.
481	Raanan Stores, Inc

5. **Which order has the max order value and which customer purchased it and which employee took the order?**

```
select o.customerNumber, c.customerName, od.ordernumber,
    sum(od.quantityOrdered * od.priceEach) as order_value,
    c.salesRepEmployeeNumber, e.firstName
from orderdetails od
inner join orders o
on od.ordernumber = o.ordernumber
inner join customers c
on o.customerNumber = c.customerNumber
inner join employees e
on c.salesRepEmployeeNumber = e.employeeNumber
group by od.ordernumber
order by order_value desc;
```

- ORDER NUMBER 10165, MADE BY DRAGON SOUVENIERS LTD HAS THE MAXIMUM ORDER VALUE



Result Grid Filter Rows: Search Export:

customerNumber	customerName	ordernumber	order_value	salesRepEmployeeNum...	firstName
148	Dragon Souveniers, Ltd.	10165	67392.85	1621	Mami
298	Vida Sport, Ltd	10287	61402.00	1702	Martin
259	Toms Spezialitäten, Ltd	10310	61234.67	1504	Barry
141	Euro+ Shopping Channel	10212	59830.55	1370	Gerard
495	Diecast Collectables	10207	59265.14	1188	Julie
151	Muscle Machine Inc	10127	58841.35	1286	Foon Yue
151	Muscle Machine Inc	10204	58793.53	1286	Foon Yue
458	Corrida Auto Replicas, Ltd	10126	57131.92	1702	Martin
239	Collectable Mini Designs Co.	10222	56822.65	1166	Leslie
124	Mini Gifts Distributors Ltd.	10142	56052.56	1165	Leslie

Result 23

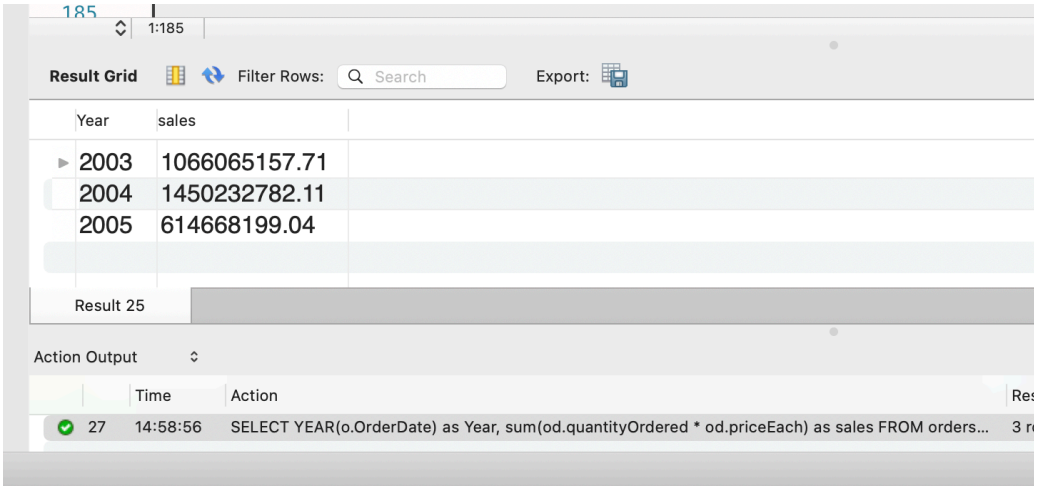
Action Output

	Time	Action	Response
25	14:55:53	select o.customerNumber, c.customerName, od.ordernumber, sum(od.quantityOrdered * od.pric...	326 row(s) returned

6. **Which year was the best for sales?**

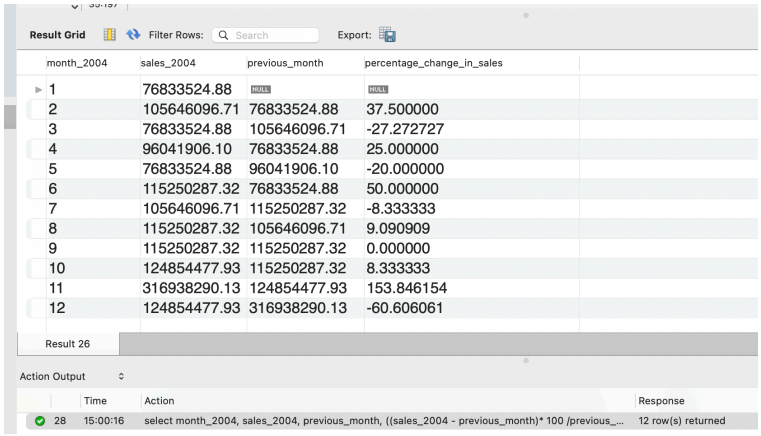
```
SELECT YEAR(o.OrderDate) as Year, sum(od.quantityOrdered * od.priceEach) as sales
FROM orders o
inner join orderdetails od
group by Year
ORDER by Year;
```

- Sales is highest for the year 2004



7. **How much in %age, the sales had increased or decreased over the months in 2004?**

```
SELECT month_2004, sales_2004, previous_month,
((sales_2004 - previous_month)* 100 /previous_month) as percentage_change_in_sales
from
(select month(o.OrderDate) as month_2004,
sum(od.quantityOrdered * od.priceEach) as sales_2004,
lag(sum(od.quantityOrdered * od.priceEach)) over(order by month(o.OrderDate)) as
previous_month
FROM orders o
inner join orderdetails od
where year(o.OrderDate) = 2004
group by month_2004) as temptable;
```



8. **What is the median sales?**

```
SET @rowindex := -1;

SELECT
  round(AVG(s.SALES), 2) as Median
FROM
   (SELECT @rowindex:=@rowindex + 1 AS rowindex,
```

```
        SUM(od.quantityOrdered * od.priceEach) AS SALES
    FROM Orderdetails od
    ORDER BY SALES) AS s
WHERE
s.rowindex IN (FLOOR(@rowindex / 2), CEIL(@rowindex / 2));
```

9. Who are the top buyer from each state? Solved using windows function?

```
with cte as (
select c.customerNumber ,c.country,
    sum(od.quantityOrdered * od.priceEach) as sales,
    dense_rank() over (partition by c.country order by sum(od.quantityOrdered *
od.priceEach) desc) as top
from customers c
inner join  orders o
on  c.customerNumber = o.customerNumber
inner join orderdetails od
on o.orderNumber = od.orderNumber
GROUP BY c.country, c.customerNumber)
select customerNumber, country, sales, top
from cte
having top = 1
order by sales desc;
```

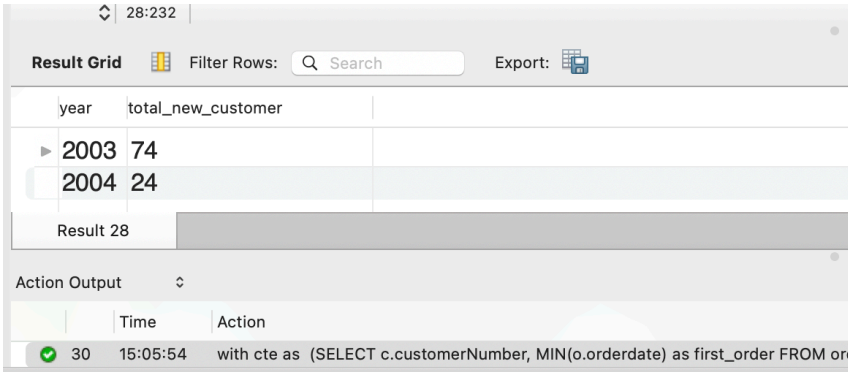
Result Grid | Filter Rows: Search | Export:

customerNumber	country	sales	top
141	Spain	820689.54	1
124	USA	591827.34	1
114	Australia	180585.07	1
119	France	158573.12	1
148	Singapore	156251.03	1
323	New Zealand	154622.08	1
187	UK	148410.09	1
382	Austria	137480.07	1
145	Denmark	129085.12	1
278	Italy	127529.69	1
448	Sweden	120943.53	1
298	Switzerland	108777.92	1
398	Japan	105548.73	1
121	Norway	104224.79	1
334	Finland	103896.74	1
259	Germany	89223.14	1
385	Philippines	87468.30	1
314	Belgium	70851.58	1
202	Canada	70122.19	1
189	Ireland	49898.27	1
211	Hong Kong	45480.79	1

10. Number of new customers over the years?

```
with cte as
(SELECT c.customerNumber, MIN(o.orderdate) as first_order
FROM orders o
inner join customers c
on c.customerNumber = o.customerNumber
group by 1)
select year(first_order) as year, count(customerNumber) as total_new_customer
```

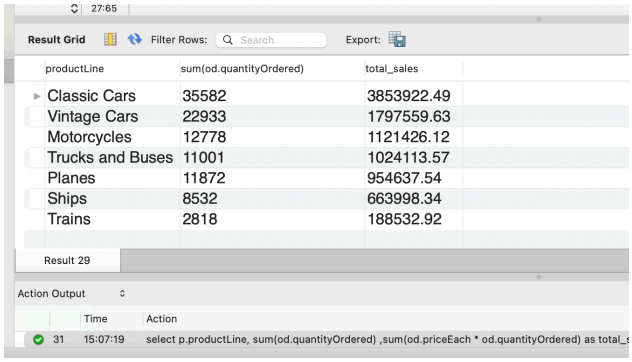
from cte
group by year(first_order);



11. Which product line has the highest and lowest total sales?

SELECT p.productLine, sum(od.quantityOrdered) ,sum(od.priceEach * od.quantityOrdered) as total_sales
from orderdetails od
inner join products p
on p.productCode = od.productCode
group by p.productLine
order by total_sales desc;

- classic cars have the most sales and trains have the least sales

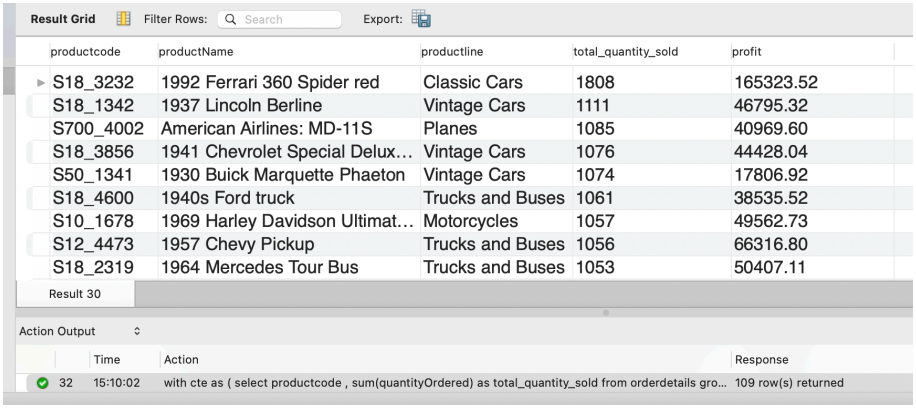


12. What are the top most ordered products by quantity sold and what is the total profit of each?

with cte as (
select productcode , sum(quantityOrdered) as total_quantity_sold
from orderdetails
group by productcode)
select p.productcode, p.productName, p.productline ,total_quantity_sold ,
total_quantity_sold * (MSRP - buyPrice) as profit
from products p
join cte
on p.productcode = cte.productcode

order by total_quantity_sold desc;

- 1992 Ferrari 360 spider red is the most ordered product and has made the maximum profit



Result Grid Filter Rows: Search Export:

productcode	productName	productline	total_quantity_sold	profit
S18_3232	1992 Ferrari 360 Spider red	Classic Cars	1808	165323.52
S18_1342	1937 Lincoln Berline	Vintage Cars	1111	46795.32
S700_4002	American Airlines: MD-11S	Planes	1085	40969.60
S18_3856	1941 Chevrolet Special Delux...	Vintage Cars	1076	44428.04
S50_1341	1930 Buick Marquette Phaeton	Vintage Cars	1074	17806.92
S18_4600	1940s Ford truck	Trucks and Buses	1061	38535.52
S10_1678	1969 Harley Davidson Ultimat...	Motorcycles	1057	49562.73
S12_4473	1957 Chevy Pickup	Trucks and Buses	1056	66316.80
S18_2319	1964 Mercedes Tour Bus	Trucks and Buses	1053	50407.11

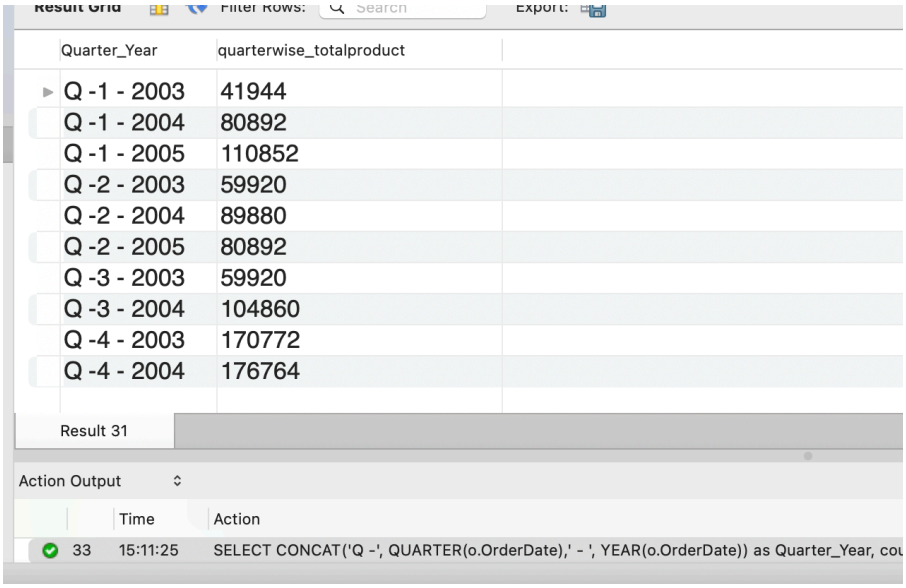
Result 30

Action Output

	Time	Action	Response
32	15:10:02	with cte as (select productcode , sum(quantityOrdered) as total_quantity_sold from orderdetails gro...	109 row(s) returned

13. **How many products were sold in each quarter?**

```
SELECT CONCAT('Q -', QUARTER(o.OrderDate),' - ', YEAR(o.OrderDate)) as Quarter_Year,
count(productCode) as quarterwise_totalproduct
FROM orders o
inner join orderdetails od
group by Quarter_Year
ORDER by Quarter_Year;
```



Result Grid Filter Rows: Search Export:

Quarter_Year	quarterwise_totalproduct
Q -1 - 2003	41944
Q -1 - 2004	80892
Q -1 - 2005	110852
Q -2 - 2003	59920
Q -2 - 2004	89880
Q -2 - 2005	80892
Q -3 - 2003	59920
Q -3 - 2004	104860
Q -4 - 2003	170772
Q -4 - 2004	176764

Result 31

Action Output

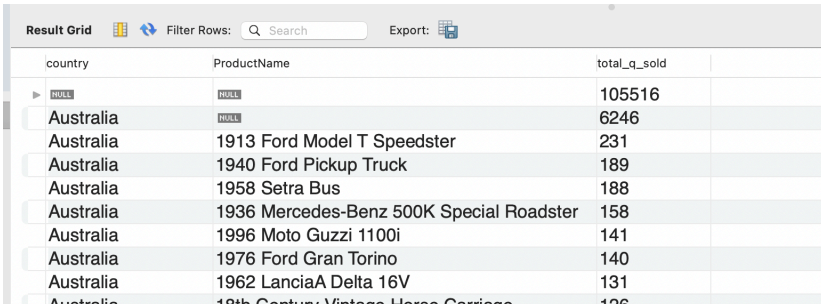
	Time	Action
33	15:11:25	SELECT CONCAT('Q -', QUARTER(o.OrderDate),' - ', YEAR(o.OrderDate)) as Quarter_Year, cou

14. Which products are most popular in which country?

```
SELECT c.country, p.ProductName,SUM(od.quantityOrdered) AS total_q_sold
FROM Orders o
INNER JOIN OrderDetails od
ON o.OrderNumber = od.OrderNumber
```

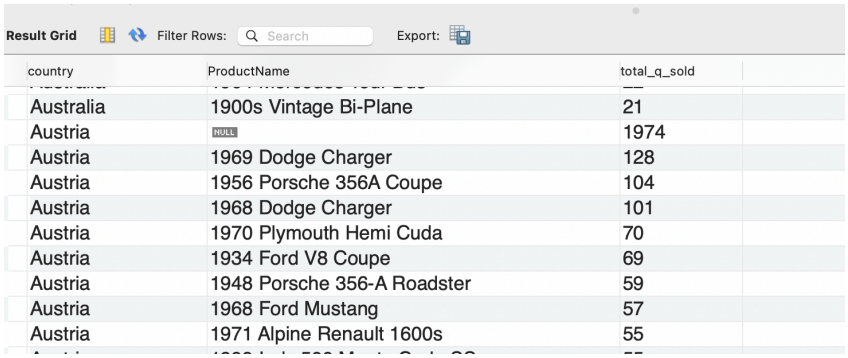
```
INNER JOIN Products p
ON od.ProductCode = p.ProductCode
INNER JOIN Customers c
ON o.CustomerNumber = c.CustomerNumber
GROUP BY
 c.country,
 p.ProductName with rollup
ORDER BY
 c.country,
 total_q_sold desc;
```

- WITH ROLLUP gives us the summarization of total number of products sold the those countries



Result Grid Filter Rows: Search Export:

country	ProductName	total_q_sold
NULL	NULL	105516
Australia	NULL	6246
Australia	1913 Ford Model T Speedster	231
Australia	1940 Ford Pickup Truck	189
Australia	1958 Setra Bus	188
Australia	1936 Mercedes-Benz 500K Special Roadster	158
Australia	1996 Moto Guzzi 1100i	141
Australia	1976 Ford Gran Torino	140
Australia	1962 LanciaA Delta 16V	131

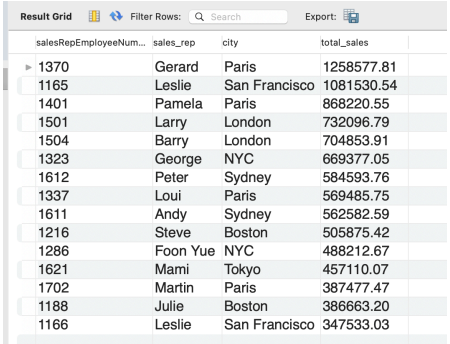


Result Grid Filter Rows: Search Export:

country	ProductName	total_q_sold
Australia	1900s Vintage Bi-Plane	21
Austria	NULL	1974
Austria	1969 Dodge Charger	128
Austria	1956 Porsche 356A Coupe	104
Austria	1968 Dodge Charger	101
Austria	1970 Plymouth Hemi Cuda	70
Austria	1934 Ford V8 Coupe	69
Austria	1948 Porsche 356-A Roadster	59
Austria	1968 Ford Mustang	57
Austria	1971 Alpine Renault 1600s	55

15. Who are the top-performing sales employees?

```
SELECT c.salesRepEmployeeNumber, e.firstName as sales_rep, offices.city,
sum(od.quantityOrdered * od.priceEach) as total_sales
from orderdetails od
inner join orders o
on od.ordernumber = o.ordernumber
inner join customers c
on o.customerNumber = c.customerNumber
inner join employees e
on c.salesRepEmployeeNumber = e.employeeNumber
inner join offices
on e.officecode = offices.officecode
group by c.salesRepEmployeeNumber
order by total_sales desc;
```



Result Grid Filter Rows: Search Export:

salesRepEmployeeNum...	sales_rep	city	total_sales
1370	Gerard	Paris	1258577.81
1165	Leslie	San Francisco	1081530.54
1401	Pamela	Paris	868220.55
1501	Larry	London	732096.79
1504	Barry	London	704853.91
1323	George	NYC	669377.05
1612	Peter	Sydney	584593.76
1337	Loui	Paris	569485.75
1611	Andy	Sydney	562582.59
1216	Steve	Boston	505875.42
1286	Foon Yue	NYC	488212.67
1621	Mami	Tokyo	457110.07
1702	Martin	Paris	387477.47
1188	Julie	Boston	386663.20
1166	Leslie	San Francisco	347533.03

CONCLUSION:

According to the data analysis done, it can be concluded that:

- Euro + shopping channel and mini gifts distributors Ltd are the most frequent customers of this business and they also have the highest order or purchase value. They are high-value customers and the business can focus their effort on providing them personalised services and incentives to keep them happy.
- There are many ‘churn’ and ‘at risk’ customers. The business should take actions to prevent further churn and retain the at-risk customers. They can conduct exit surveys, analyse customer feedback and complaints. They can come up with loyal or reward schemes to retain the at-risk customers.
- The order with maximum value was made by Dragon Souveniers, Ltd.
- It is seen from the analysis that sale was highest in 2004 and lowest in 2005. The reason for low sales in 2005 is that the data we have on sales of the year 2005 is of first 5 months.
- In the year 2004, the month of November had the highest sale, and saw the biggest change from the previous month.
- The number of new customers decreased over the years, with no new customer until May 2005.
- Classic cars are the most popular and profitable product line. As per the popularity of the product lines the business can evaluate the effectiveness of their marketing strategies for each of them. 1992 Ferrari 360 spider red is the most ordered product, this information can help the business allocate resources and invest in products that are performing well.
- Quarter 4 of the years 2003 and 2004 had high sales.
- Gerard is the top performing sales representative. The business should provide incentives or bonuses to top-performing sales representatives, this motivates other sales reps to work harder.