

SQL for Data Analysis

BASIC SELECT + FILTERING + SORTING

1. List all customers from USA sorted by their City.

Code:

```
select * from customer where Country = 'USA' order by City;
```

	Id	FirstName	LastName	City	Country	Phone
▶	65	Paula	Wilson	Albuquerque	USA	(505) 555-5939
	55	Rene	Phillips	Anchorage	USA	(907) 555-7584
	71	Jose	Pavarotti	Boise	USA	(208) 555-8097
	78	Liu	Wong	Butte	USA	(406) 555-5834
	36	Yoshi	Latimer	Elgin	USA	(503) 555-6874
	32	Howard	Snyder	Eugene	USA	(503) 555-7555
	82	Helvetius	Nagy	Kirkland	USA	(206) 555-8257
	75	Art	Braunschweiger	Lander	USA	(307) 555-4680
	48	Fran	Wilson	Portland	USA	(503) 555-9573
	77	Liz	Nixon	Portland	USA	(503) 555-3612
	45	Jaime	Yorres	San Francisco	USA	(415) 555-5938
	89	Karl	Jablonski	Seattle	USA	(206) 555-4112
	43	John	Steel	Walla Walla	USA	(509) 555-7969
▲	NULL	NULL	NULL	NULL	NULL	NULL

2.Retrieve the details of all orders placed in 2014-01-05.

code:

```
select * from orders where date_format(orderdate,'%Y-%m-%d') = '2014-01-05';
```

	Id	OrderDate	OrderNumber	CustomerId	TotalAmount
▶	566	2014-01-05 12:00:00	542943	67	648.00
	567	2014-01-05 12:00:00	542944	84	2070.00
	568	2014-01-05 12:00:00	542945	71	40.00
•	NULL	NULL	NULL	NULL	NULL

3.List all customers, they haven't placed any orders.

Code:

```
select concat(c.firstname,' ',c.lastname) as customer_name,o.id from customer  
c left join orders o on c.id=o.customerid where id is null;
```

	customer_name	id
▶	Marie Bertrand	NULL
	Diego Roel	NULL

4.Show the top 10 most recent orders with customer names and total amount.

Code:

```
select concat(firstname,' ',lastname) as customer_name,totalamount, orderdate  
from customer c join orders o on c.id=o.customerid order by orderdate desc  
limit 10;
```

	customer_name	totalamount	orderdate
►	Jytte Petersen	244.30	2014-05-06 12:00:00
	Michael Holz	586.00	2014-05-06 12:00:00
	Laurence Lebihan	1057.00	2014-05-06 12:00:00
	Paula Wilson	1374.60	2014-05-06 12:00:00
	Renate Messner	1873.50	2014-05-05 12:00:00
	Carlos González	510.00	2014-05-05 12:00:00
	Roland Mendel	5218.00	2014-05-05 12:00:00
	Guillermo Fernández	300.00	2014-05-05 12:00:00
	Sven Ottlieb	86.85	2014-05-04 12:00:00
	Lúcia Carvalho	2384.80	2014-05-04 12:00:00

5.Find all discontinued products.

Code:

```
select * from product where IsDiscontinued is not null and  
IsDiscontinued > 0;
```

	Id	ProductName	SupplierId	UnitPrice	Package	IsDiscontinued
►	5	Chef Anton's Gumbo Mix	2	21.35	36 boxes	1
	9	Mishi Kobe Niku	4	97.00	18 - 500 g pkgs.	1
	17	Alice Mutton	7	39.00	20 - 1 kg tins	1
	24	Guaraná Fantástica	10	4.50	12 - 355 ml cans	1
	28	Rössle Sauerkraut	12	45.60	25 - 825 g cans	1
	29	Thüringer Rostbratwurst	12	123.79	50 bags x 30 sausqs.	1
	42	Singaporean Hokkien Fried Mee	20	14.00	32 - 1 kg pkgs.	1
	53	Perth Pasties	24	32.80	48 pieces	1
*	NULL	NULL	NULL	NULL	NULL	NULL

AGGREGATIONS + GROUP BY

6. Calculate the total sales (sum of TotalAmount) for each country.

Code:

```
select country, sum(totalamount) as total_amount from customer c join  
orders o on c.id=o.CustomerId group by 1;
```

	country	total_amount
►	Germany	244640.63
	Mexico	24073.45
	UK	60616.51
	Sweden	59523.70
	France	85498.76
	Spain	19431.89
	Canada	55334.10
	Argentina	8119.10
	Switzerland	32919.50
	Brazil	114968.48
	Austria	139496.63
	Italy	16705.15
	Portugal	12468.65
	USA	263566.98
	Venezuela	60814.89
	Ireland	57317.39
	Belgium	35134.98
	Norway	5735.15
	Denmark	34782.25
	Finland	19778.45
	Poland	3531.95

7. Show total number of orders placed by each customer.

Code:

```
select concat(firstname, ' ', lastname) as customer_name, count(customerid) as  
total_orders from customer c join orders o on c.id=o.CustomerId group by 1  
order by 2 ;
```

	customer_name	total_orders
►	Francisco Chang	1
	John Steel	2
	Manuel Pereira	2
	Yoshi Tannamuri	3
	Martín Sommer	3
	Carine Schmitt	3
	Elizabeth Brown	3
	Simon Crowther	3
	Liu Wong	3
	Helvetius Nagy	3
	Daniel Tonini	4
	Jaime Yorres	4
	Dominique Perrier	4
	Liz Nixon	4
	Janine Labrune	4
	Ana Trujillo	4
	Alejandra Camino	5
	Pedro Afonso	5
	Sergio Gutiérrez	5
	Paul Henriot	5
	Yoshi Latimer	5
	Eduardo Saave...	5
	Martine Rancé	5
	Yvonne Moncada	5
	Isabel de Castro	5
	Alexander Feuer	5

	customer_name	total_orders
	Maria Anders	6
	Paolo Accorti	6
	Sven Ottlieb	6
	Anabela Domin...	6
	Guillermo Ferná...	6
	Karin Josephs	6
	Jonas Bergulfsen	6
	Patricio Simpson	6
	Matti Karttunen	7
	Aria Cruz	7
	Hanna Moos	7
	Antonio Moreno	7
	Zbyszek Piestrz...	7
	Jytte Petersen	7
	Catherine Dewey	7
	Fran Wilson	8
	Yang Wang	8
	Ann Devon	8
	Lino Rodriguez	8
	Bernardo Batista	9
	Art Braunschwei...	9
	André Fonseca	9
	Hari Kumar	9
	Paula Parente	9
	Miguel Angel Pa...	10
	Victoria Ashworth	10

	customer_name	total_orders
	Helen Bennett	10
	Rita Müller	10
	Michael Holz	10
	Rene Phillips	10
	Mary Saveley	10
	Giovanni Rovelli	10
	Georg Pipps	10
	José Pedro Freyre	10
	Henriette Pfalz...	10
	Janete Limeira	11
	Palle Ibsen	11
	Howard Snyder	11
	Frédérique Citea...	11
	Maurizio Moroni	12
	Felipe Izquierdo	12
	Pascale Cartrain	12
	Thomas Hardy	13
	Jean Fresnière	13
	Lúcia Carvalho	13
	Karl Jablonski	14
	Mario Pontes	14
	Annette Roulet	14
	Elizabeth Lincoln	14
	Carlos González	14
	Philip Cramer	14

	Pirkko Koskitalo	15
	Renate Messner	15
	Peter Franken	15
	Laurence Lebihan	17
	Christina Berglund	18
	Paula Wilson	18
	Carlos Hernández	18
	Maria Larsson	19
	Patricia McKenna	19
	Horst Kloss	28
	Roland Mendel	30
	Jose Pavarotti	31

8.Find the top 5 customers based on the total purchase amount.

Code:

```
select concat(firstname,' ',lastname) as customer_name, sum(totalamount) as  
total_purchase_amount from customer c join orders o on  
c.id=o.customeridgroup by 1 order by 2 desc limit 5;
```

	customer_name	total_purchase_amount
►	Horst Kloss	117483.39
	Jose Pavarotti	115673.39
	Roland Mendel	113236.68
	Patricia McKenna	57317.39
	Paula Wilson	52245.90

9.Display monthly total sales from the orders table.

Code:

```
select date_format(orderdate,'%Y-%m') as month,sum(totalamount) as  
total_amount from orders group by 1 ;
```

	month	total_amount
►	2012-07	30192.10
	2012-08	26609.40
	2012-09	27636.00
	2012-10	41203.60
	2012-11	49704.00
	2012-12	50953.40
	2013-01	66692.80
	2013-02	41207.20
	2013-03	39979.90
	2013-04	55699.39
	2013-05	56823.70
	2013-06	39088.00
	2013-07	55464.93
	2013-08	49981.69
	2013-09	59733.02
	2013-10	70328.50
	2013-11	45913.36
	2013-12	77476.26
	2014-01	100854.72
	2014-02	104561.95
	2014-03	109825.45
	2014-04	134630.56
	2014-05	19898.66

JOINS (INNER, LEFT, RIGHT)

10. List order details with product names and customer names.

Code:

```
select o.id, concat(c.firstname, ' ', c.lastname) as customer_name, p.productname,  
o.orderdate, o.ordernumber, o.customerid, o.totalamount from customer c join  
orders o on c.id=o.customerid join orderitem oi on o.id = oi.OrderId join  
product p on oi.ProductId = p.id;
```

	id	customer_name	productname	orderdate	ordernumber	customerid	totalamount
	429	Miguel Angel Pa...	Teatime Chocolate Biscuits	2013-09-22 12:00:00	542806	80	534.85
	429	Miguel Angel Pa...	Gula Malacca	2013-09-22 12:00:00	542806	80	534.85
	595	Miguel Angel Pa...	Queso Cabrales	2014-01-20 12:00:00	542972	80	975.00
	595	Miguel Angel Pa...	Ipoh Coffee	2014-01-20 12:00:00	542972	80	975.00
	595	Miguel Angel Pa...	Scottish Longbreads	2014-01-20 12:00:00	542972	80	975.00
	595	Miguel Angel Pa...	Outback Lager	2014-01-20 12:00:00	542972	80	975.00
	668	Miguel Angel Pa...	Alice Mutton	2014-02-27 12:00:00	543045	80	539.50
	668	Miguel Angel Pa...	Geitost	2014-02-27 12:00:00	543045	80	539.50
	668	Miguel Angel Pa...	Tourtière	2014-02-27 12:00:00	543045	80	539.50
	822	Miguel Angel Pa...	Chartreuse verte	2014-05-04 12:00:00	543199	80	360.00
	42	Victoria Ashworth	Aniseed Syrup	2012-08-26 12:00:00	542419	11	479.40
	42	Victoria Ashworth	Wimmers gute Semmelknödel	2012-08-26 12:00:00	542419	11	479.40
	224	Victoria Ashworth	Uncle Bob's Organic Dried Pears	2013-03-11 12:00:00	542601	11	1328.00
	224	Victoria Ashworth	Gnocchi di nonna Alice	2013-03-11 12:00:00	542601	11	1328.00
	237	Victoria Ashworth	Sir Rodney's Scones	2013-03-24 12:00:00	542614	11	386.20
	237	Victoria Ashworth	Boston Crab Meat	2013-03-24 12:00:00	542614	11	386.20

11.Show all products and their supplier names, even if some suppliers didn't supply anything (LEFT JOIN)

code:

```
select productname,s.companyname,s.contactname from product p left join supplier s  
on p.SupplierId = s.Id;
```

	productname	companyname	contactname
►	Chai	Exotic Liquids	Charlotte Cooper
	Chang	Exotic Liquids	Charlotte Cooper
	Aniseed Syrup	Exotic Liquids	Charlotte Cooper
	Chef Anton's Cajun Seasoning	New Orleans Cajun Delights	Shelley Burke
	Chef Anton's Gumbo Mix	New Orleans Cajun Delights	Shelley Burke
	Grandma's Boysenberry Spread	Grandma Kelly's Homestead	Regina Murphy
	Uncle Bob's Organic Dried Pears	Grandma Kelly's Homestead	Regina Murphy
	Northwoods Cranberry Sauce	Grandma Kelly's Homestead	Regina Murphy
	Mishi Kobe Niku	Tokyo Traders	Yoshi Nagase
	Ikura	Tokyo Traders	Yoshi Nagase
	Queso Cabrales	Cooperativa de Quesos 'La...	Antonio del Valle ...
	Queso Manchego La Pastora	Cooperativa de Quesos 'La...	Antonio del Valle ...
	Konbu	Mayumi's	Mayumi Ohno
	Tofu	Mayumi's	Mayumi Ohno
	Genen Shouyu	Mayumi's	Mayumi Ohno
	Pavlova	Pavlova, Ltd.	Ian Devling
	Alice Mutton	Pavlova, Ltd.	Ian Devling

12. List customers who placed orders.

Code:

```
select distinct concat(c.firstname, ' ', c.lastname) as customer_name from customer c  
inner join orders o on c.id=o.customerid;
```

	customer_name
▶	Paolo Accorti
	Pedro Afonso
	Maria Anders
	Miguel Angel Paolino
	Victoria Ashworth
	Bernardo Batista
	Helen Bennett
	Christina Berglund
	Jonas Bergulfsen
	Art Braunschweiger
	Elizabeth Brown
	Alejandra Camino
	Pascale Cartrain
	Lúcia Carvalho
	Francisco Chang
	Frédérique Citeaux
	Philip Cramer

SUBQUERIES

13. Find customers who ordered more than 3 times.

Code:

```
select concat(firstname, ' ', lastname) as customer_name from customer
where id in (select customerid from orders group by customerid having
count(*) > 3 );
```

	customer_name
►	Paolo Accorti
	Pedro Afonso
	Maria Anders
	Miguel Angel Paolino
	Victoria Ashworth
	Bernardo Batista
	Helen Bennett
	Christina Berglund
	Jonas Bergulfsen
	Art Braunschweiger
	Alejandra Camino
	Pascale Cartrain
	Lúcia Carvalho
	Frédérique Citeaux
	Philip Cramer
	Aria Cruz

14. List products that have been ordered more than 100 times total.

Code:

```
Select p.ProductName from product p join orderitem oi on p.Id = oi.ProductId group by p.Id, p.ProductName having SUM(oi.Quantity) > 100;
```

	ProductName
▶	Alice Mutton
	Aniseed Syrup
	Boston Crab Meat
	Camembert Pierrot
	Carnarvon Tigers
	Chai
	Chang
	Chartreuse verte
	Chef Anton's Cajun Seasoning
	Chef Anton's Gumbo Mix
	Chocolade
	Côte de Blaye
	Escargots de Bourgogne
	Filo Mix
	Flotemysost
	Geitost
	Genen Shouyu
	Gnocchi di nonna Alice
	Gorgonzola Telino
	Grandma's Boysenberry Spre...
	Gravad lax
	Guaraná Fantástica

15.Find the highest spending customer.

Code:

```
select concat(firstname,' ',lastname) as customer_name ,sum(totalamount) as  
total_amount_spent from customer c join orders o on c.id=o.CustomerIdgroup by 1  
order by 2 desc limit 1;
```

	customer_name	total_amount_spent
▶	Horst Kloss	117483.39

16.Show the number of distinct products ordered.

Code:

```
select count(distinct productid) as Distinct_Products_Ordered from orderitem;
```

	Distinct_Products_Ordered
▶	77

17.Retrieve all orders that had a total amount above the average order amount.

Code:

```
select * from orders where TotalAmount >(select avg(TotalAmount) from orders);
```

	Id	OrderDate	OrderNumber	CustomerId	TotalAmount
►	2	2012-07-05 12:00:00	542379	79	1863.40
	3	2012-07-08 12:00:00	542380	34	1813.00
	5	2012-07-09 12:00:00	542382	76	3730.00
	8	2012-07-12 12:00:00	542385	68	2490.50
	11	2012-07-17 12:00:00	542388	20	2018.60
	13	2012-07-19 12:00:00	542390	56	1746.20
	16	2012-07-23 12:00:00	542393	20	2464.80
	20	2012-07-29 12:00:00	542397	25	4031.00
	26	2012-08-05 12:00:00	542403	63	2142.40
	38	2012-08-20 12:00:00	542415	63	2179.20
	39	2012-08-21 12:00:00	542416	63	3016.00
	43	2012-08-27 12:00:00	542420	15	2169.00
	47	2012-08-30 12:00:00	542424	65	1887.60
	51	2012-09-05 12:00:00	542428	37	3127.00
	55	2012-09-10 12:00:00	542432	76	2708.80
	58	2012-09-13 12:00:00	542435	55	4157.00

AGGREGATE FUNCTIONS

18.What is the average quantity per order item?

Code:

```
select avg(quantity) as average_quantity from orderitem;
```

	average_quantity
▶	23.8130

19.What's the maximum unit price of any product?

Code:

```
select max(unitprice) as maximum_unit_price from orderitem;
```

	maximum_unit_price
▶	263.50

20.Find the total revenue generated per product.

Code:

```
select productid,sum(quantity*unitprice) as total_revenue from orderitem group by 1;
```

	productid	total_revenue
▶	1	14277.60
	2	18559.20
	3	3080.00
	4	9424.80
	5	5801.15
	6	7345.00
	7	22464.00
	8	13760.00
	9	8827.00
	10	22140.20
	11	13902.00
	12	12866.80
	13	5234.40
	14	8630.40
	15	1813.50
	16	18748.05