Shanmuka Siva Ganesh Palakula

ganeshkv1395@gmail.com

Pharma Data Solutions

1. Retrieve all columns for all records in the dataset.

Solution:

```
select * from pdata;
```

Distributor	customername	City	Country	Latitude	Longitude	Channel	subchannel	producti
Gottlieb-Cruickshank	Zieme, Doyle and Kunze	Lublin	Poland	51.2333	22.5667	Hospital	Private	Topipizole
Gottlieb-Cruickshank	Feest PLC	NA	Poland	53.4167	18.4333	Pharmacy	Retail	Choriotris
Gottlieb-Cruickshank	Medhurst-Beer Pharmaceutical Limited	Rybnik	Poland	50.0833	18.5	Pharmacy	Institution	Acantaine
Gottlieb-Cruickshank	Barton Ltd Pharma Plc	NA	Poland	50.3333	19.0833	Hospital	Private	Lioletine F
Gottlieb-Cruickshank	Keeling LLC Pharmacy	Olsztyn	Poland	53.78	20.4942	Pharmacy	Retail	Oxymotr
Gottlieb-Cruickshank	Runte-Marquardt Pharmaceutical Ltd	Olecko	Poland	54.0333	22.5	Hospital	Private	Pazofena
Gottlieb-Cruickshank	Blick, Pacocha and Schowalter	NA	Poland	52.7958	18.2611	Pharmacy	Retail	Symbitrin
Gottlieb-Cruickshank	Leuschke PLC Pharmacy	NA	Poland	52.8817	20.6106	Pharmacy	Retail	Morphizo
Gottlieb-Cruickshank	Miller-Satterfield Pharma Plc	Nidzica	Poland	53.3583	20.425	Hospital	Private	Lovapur
Gottlieb-Cruickshank	Bashirian-Kassulke Pharma Plc	NA	Poland	50.0614	19.9372	Hospital	Private	Ampysin
Gottlieb-Cruickshank	Wolff Group Pharm	Gubin	Poland	51.95	14.7333	Hospital	Government	Trazobala

2. How many unique countries are represented in the dataset?

Solution:

```
select distinct Country from pdata;
```

only two countries Poland and Germany.

Country
Poland
Germany

3. Select the names of all the customers on the 'Retail' channel.

Solution:

```
select * from pdata
where subchannel = 'Retail';
```

Distributor	customername	City	Country	Latitude	Longitude	Channel	subchannel
Gottlieb-Cruickshank	Feest PLC	NA	Poland	53.4167	18.4333	Pharmacy	Retail
Gottlieb-Cruickshank	Keeling LLC Pharmacy	Olsztyn	Poland	53.78	20.4942	Pharmacy	Retail
Gottlieb-Cruickshank	Blick, Pacocha and Schowalter	NA	Poland	52.7958	18.2611	Pharmacy	Retail
Gottlieb-Cruickshank	Leuschke PLC Pharmacy	NA	Poland	52.8817	20.6106	Pharmacy	Retail
Gottlieb-Cruickshank	McClure, Zemlak and Dibbert Pharma...	NA	Poland	54.3667	18.6333	Pharmacy	Retail
Gottlieb-Cruickshank	Lindgren-Simonis Pharm	Biskupiec	Poland	53.8647	20.9569	Pharmacy	Retail
Gottlieb-Cruickshank	Will and Sons Pharma Plc	Gniew	Poland	53.8333	18.8333	Pharmacy	Retail
Gottlieb-Cruickshank	Jakubowski Inc Pharmaceutical Limited	NA	Poland	51.2667	15.5667	Pharmacy	Retail
Gottlieb-Cruickshank	Nader-Gaylord Pharmaceutical Ltd	NA	Poland	51.7769	19.4547	Pharmacy	Retail
Gottlieb-Cruickshank	Emard-O'Connell Pharmacy	Zakopane	Poland	49.2994	19.9519	Pharmacy	Retail
Gottlieb-Cruickshank	Feest PLC	NA	Poland	53.4167	18.4333	Pharmacy	Retail
Gottlieb-Cruickshank	McCullough LLC Pharma Plc	NA	Poland	50.8	19.1167	Pharmacy	Retail
Gottlieb-Cruickshank	Parisian, Hagenes and Skiles Pharma Plc	Ciechoci...	Poland	52.8833	18.7833	Pharmacy	Retail

4. Find the total quantity sold for the 'Antibiotics' product class.

Solution:

```
select productclass, sum(Quantity) from pdata
where productclass = 'Antibiotics'
group by productclass;
```

productclass	sum(Quantity)
Antibiotics	1189013

5. List all the distinct months present in the dataset.

Solution:

```
select distinct Month from pdata;
```

Month
January
February
March
April
May
June
July
August
September
October
November
December

6. Calculate the total sales for each year.

Solution:

```
select Year, sum(Sales) from pdata
group by Year;
```

Year	sum(Sales)
2018	683989173.8
2017	2701480740.81559

7. Find the customer with the highest sales value.

Solution:

```
select customername, max(Sales) from pdata
group by customername
order by max(sales) DESC
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap C

customername	max(Sales)
Wiegand, Jast and Yost Pharmaceuti...	17225000

8. Get the names of all employees who are Sales Reps and are managed by 'James Goodwill'.

Solution:

```
select nameofsalesrep from pdata
where Manager = 'James Goodwill';
```

nameofsalesrep
Thompson Crawford
Erica Jones
Alan Ray
Erica Jones
Alan Ray
Alan Ray
Alan Ray
Alan Ray
Thompson Crawford
Erica Jones
Alan Ray
Thompson Crawford
Erica Jones

9. Retrieve the top 5 cities with the highest sales.

Solution:

```
select City, max(Sales) from pdata
group by City
order by max(Sales) DESC
LIMIT 5;
```

City	max(Sales)
Ludwigshafen	17225000
Filderstadt	14406000
Berlin	13734000
Melle	12080000
Wedel	10660000

10. Calculate the average price of products in each sub-channel.

Solution:

```
select subchannel, AVG(price) from pdata
group by subchannel;
```

subchannel	AVG(price)
Private	408.9635
Retail	412.5558
Institution	413.8806
Government	411.9139

11. Join the 'Employees' table with the 'Sales' table to get the name of the Sales Rep and the corresponding sales records.

Solution:

There is no unique column to join the table. If there is a unique column then we can establish the connection using `SELF JOIN` and we can retrieve the details of the sales rep and the corresponding sales records.

```
select distinct nameofsalesrep, SUM(sales) from pdata
group by nameofsalesrep;
```

This query retrieves the details of names of the sales rep and there corresponding sales records without using SELF JOIN.

nameofsalesrep	SUM(sales)
Mary Gerrard	284486324
Jessica Smith	266494277
Steve Pepple	237121455
Anne Wu	296178560.5089
Thompson Crawford	267212174
Sheila Stones	254180520
Stella Given	212575740
Morris Garcia	279108525
Erica Jones	254498319
Abigail Thompson	278862045.10669005
Daniel Gates	239601610
Alan Ray	268454950
Jimmy Grey	246695415

12. Retrieve all sales made by employees from ' Rendsburg ' in the year 2018.

Solution:

```
select nameofsalesrep, sales from pdata
where City = 'Rendsburg' and Year = 2018
group by nameofsalesrep;
```

13. Calculate the total sales for each product class, for each month, and order the results by year, month, and product class.

Solution:

```
SELECT distinct productclass, month, year, SUM(sales)
AS total_sales
FROM pdata
GROUP BY productclass, month, year
ORDER BY year, month, productclass;
```

productclass	month	year	total_sale
Analgesics	April	2017	32223716
Antibiotics	April	2017	40029226
Antimalarial	April	2017	17789675
Antipiretics	April	2017	22868812
Antiseptics	April	2017	42712211
Mood Stabilizers	April	2017	33176944
Analgesics	August	2017	49744520
Antibiotics	August	2017	32449096
Antimalarial	August	2017	25887712
Antipiretics	August	2017	39342305
Antiseptics	August	2017	45881555
Mood Stabilizers	August	2017	40529487
Analgesics	Dece	2017	64973444

14. Find the top 3 sales reps with the highest sales in 2019.

Solution:

```
select nameofsalesrep, sum(sales) from pdata
where year = 2018
group by nameofsalesrep
order by sum(sales) DESC
LIMIT 3;
```

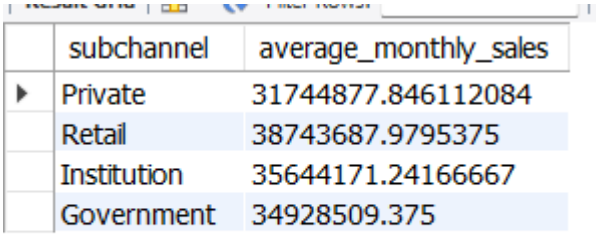
nameofsalesrep	sum(sales)
Jimmy Grey	57086313
Daniel Gates	56794687
Morris Garcia	56606142

15. Calculate the monthly total sales for each sub-channel, and then calculate the average monthly sales for each sub-channel over the years.

Solution:

```
select subchannel, avg(monthly_sales) as
average_monthly_sales
from (
select subchannel, month, year, sum(sales) as
monthly_sales
from pdata
```

```
group by subchannel, month, year
) as subchannel_monthly_sales
group by subchannel;
```

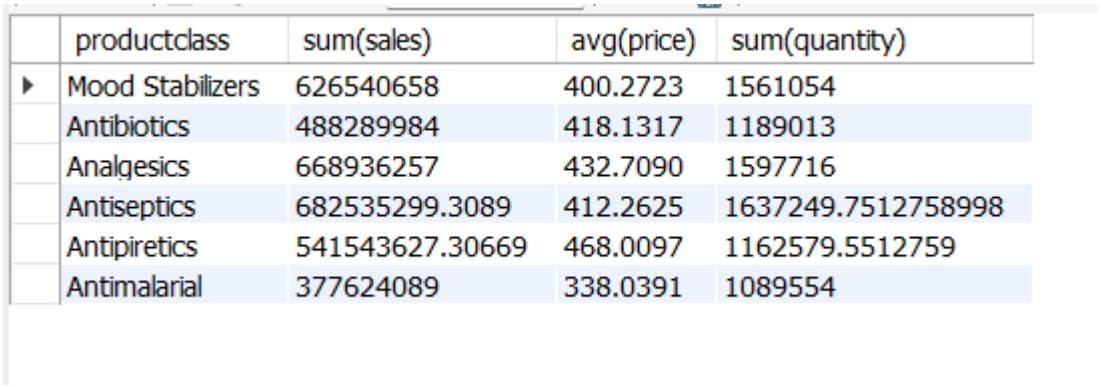


subchannel	average_monthly_sales
Private	31744877.846112084
Retail	38743687.9795375
Institution	35644171.24166667
Government	34928509.375

16. Create a summary report that includes the total sales, average price, and total quantity sold for each product class.

Solution:

```
select productclass, sum(sales), avg(price),
sum(quantity) from pdata
group by productclass;
```



productclass	sum(sales)	avg(price)	sum(quantity)
Mood Stabilizers	626540658	400.2723	1561054
Antibiotics	488289984	418.1317	1189013
Analgesics	668936257	432.7090	1597716
Antiseptics	682535299.3089	412.2625	1637249.7512758998
Antipiretics	541543627.30669	468.0097	1162579.5512759
Antimalarial	377624089	338.0391	1089554

Summary Report:

This query selects the productclass and calculates total sales, average price and total quantity

- Total sales: the sum(sales) adds up the sales for each productclass
- Average Price: the avg(price) computes the average of the price for each productclass.
- Total quantity: the sum(quantity) adds up the quantity for each productclass.

17. Find the top 5 customers with the highest sales for each year.

Solution:

```
select year, nameofsalesrep, total_sales
from (
select year, nameofsalesrep, sum(sales) as total_sales
from pdata
group by year, nameofsalesrep
) as yearly_sales
where (year, total_sales) in
(select year, max(total_sales) as max_sales from
(
select year, nameofsalesrep, sum(sales) as total_sales
from pdata
group by year, nameofsalesrep
) as yearly_sales_max
group by year )
order by year;
```

year	nameofsalesrep	total_sales
2017	Anne Wu	240168631.5089
2018	Jimmy Grey	57086313

18. Calculate the year-over-year growth in sales for each country.

Solution:

```
select
    current.year as year,
    current.country as country,
    current.sales as current_year_sales,
```

```
    previous.sales as previous_year_sales,
    ((current.sales - previous.sales) /
previous.sales) * 100 as sales_growth_percentage
from (
select country, year, sum(sales) as sales from pdata
group by country, year
) as current
left join
(select country, year, sum(sales) as sales from pdata
group by country, year) as previous on current.country
= previous.country and current.year = previous.year +
1
order by current.country, current.year;
```

year	country	current_year_sales	previous_year_sales	sales_growth_percentage
2017	Germany	2701480740.81559	NULL	NULL
2018	Germany	3109372	2701480740.81559	-99.88490119685024
2018	Poland	680879801.8	NULL	NULL

19. List the months with the lowest sales for each year.

Solution:

```
select year, month, min(sum_sales) as min_total_sales
from (select year, month, sum(sales) as sum_sales from
pdata
group by year, month) as total_sales
group by year, month
order by min_total_sales;
```

year	month	min_total_sa
2018	April	36361345.8
2018	December	44208503
2018	January	44914055
2018	July	48643178
2018	May	50167387
2018	March	55415538
2018	June	58733162
2018	February	62348889
2018	October	62967497
2018	August	66857565
2018	September	75542574
2018	November	77829480
2017	January	151872184

20. Calculate the total sales for each sub-channel in each country, and then find the country with the highest total sales for each sub-channel.

Solution:

```
select subchannel, country, max(sum_sales)
from (
select subchannel, country, sum(sales) as sum_sales
from pdata
group by subchannel, country
) as total_sales
group by subchannel, country
order by max(sum_sales);
```

subchannel	country	max(sum_sales)
Government	Poland	137326867
Institution	Poland	155610090.8
Retail	Poland	180809570
Private	Poland	207133274
Private	Germany	554743794.30669
Institution	Germany	699850019
Government	Germany	700957358
Retail	Germany	749038941.5088999