PSYLIQ

INTERNSHIP PROJECT

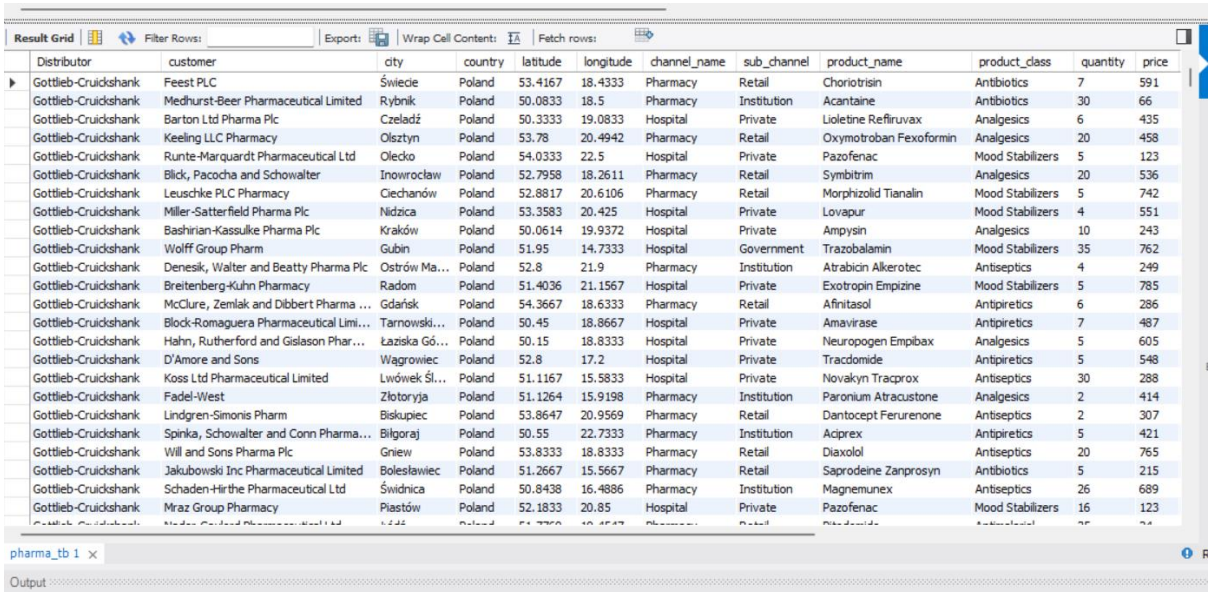
PHARMA DATA ANALYSIS USING SQL

BY

SHARUMATHI C

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

```
LOAD DATA INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\pharma.csv'

INTO TABLE pharma_tb

FIELDS TERMINATED BY ','

ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

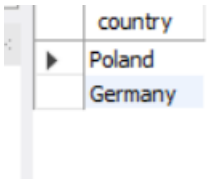
Distributor	customer	city	country	latitude	longitude	channel_name	sub_channel	product_name	product_class	quantity	price
Gottlieb-Cruickshank	Feest PLC	Świecie	Poland	53.4167	18.4333	Pharmacy	Retail	Choriotrisin	Antibiotics	7	591
Gottlieb-Cruickshank	Medhurst-Beer Pharmaceutical Limited	Rybnik	Poland	50.0833	18.5	Pharmacy	Institution	Acantaine	Antibiotics	30	66
Gottlieb-Cruickshank	Barton Ltd Pharma Plc	Czeladź	Poland	50.3333	19.0833	Hospital	Private	Lioletine Reflıruvax	Analgesics	6	435
Gottlieb-Cruickshank	Keeling LLC Pharmacy	Olsztyn	Poland	53.78	20.4942	Pharmacy	Retail	Oxymotroban Fexoformin	Analgesics	20	458
Gottlieb-Cruickshank	Runte-Marquardt Pharmaceutical Ltd	Olecko	Poland	54.0333	22.5	Hospital	Private	Pazofenac	Mood Stabilizers	5	123
Gottlieb-Cruickshank	Blick, Pacocha and Schowalter	Inowrocław	Poland	52.7958	18.2611	Pharmacy	Retail	Symbitrim	Analgesics	20	536
Gottlieb-Cruickshank	Leuschke PLC Pharmacy	Ciechanów	Poland	52.8817	20.6106	Pharmacy	Retail	Morphizolid Tianalin	Mood Stabilizers	5	742
Gottlieb-Cruickshank	Miller-Satterfield Pharma Plc	Nidzica	Poland	53.3583	20.425	Hospital	Private	Lovapur	Mood Stabilizers	4	551
Gottlieb-Cruickshank	Bashirian-Kassulke Pharma Plc	Kraków	Poland	50.0614	19.9372	Hospital	Private	Ampysin	Analgesics	10	243
Gottlieb-Cruickshank	Wolff Group Pharm	Gubin	Poland	51.95	14.7333	Hospital	Government	Trazobalamin	Mood Stabilizers	35	762
Gottlieb-Cruickshank	Denesik, Walter and Beatty Pharma Plc	Ostrów Ma…	Poland	52.8	21.9	Pharmacy	Institution	Atrabicin Alkerotec	Antiseptics	4	249
Gottlieb-Cruickshank	Breitenberg-Kuhn Pharmacy	Radom	Poland	51.4036	21.1567	Hospital	Private	Exotropin Empizine	Mood Stabilizers	5	785
Gottlieb-Cruickshank	McClure, Zemlak and Dibbert Pharma …	Gdańsk	Poland	54.3667	18.6333	Pharmacy	Retail	Afinitasol	Antipiretics	6	286
Gottlieb-Cruickshank	Block-Romaguera Pharmaceutical Limi…	Tarnowski…	Poland	50.45	18.8667	Hospital	Private	Amavirase	Antipiretics	7	487
Gottlieb-Cruickshank	Hahn, Rutherford and Gislason Phar…	Łaziska Gó…	Poland	50.15	18.8333	Hospital	Private	Neuropogen Empibax	Analgesics	5	605
Gottlieb-Cruickshank	D'Amore and Sons	Wągrowiec	Poland	52.8	17.2	Hospital	Private	Tracdomide	Antipiretics	5	548
Gottlieb-Cruickshank	Koss Ltd Pharmaceutical Limited	Lwówek Śl…	Poland	51.1167	15.5833	Hospital	Private	Novakyn Tracprox	Antiseptics	30	288
Gottlieb-Cruickshank	Fadel-West	Złotoryja	Poland	51.1264	15.9198	Pharmacy	Institution	Paronium Atracustone	Analgesics	2	414
Gottlieb-Cruickshank	Lindgren-Simonis Pharm	Biskupiec	Poland	53.8647	20.9569	Pharmacy	Retail	Dantocept Ferurenone	Antiseptics	2	307
Gottlieb-Cruickshank	Spinka, Schowalter and Conn Pharma…	Biłgoraj	Poland	50.55	22.7333	Pharmacy	Institution	Aciprex	Antipiretics	5	421
Gottlieb-Cruickshank	Will and Sons Pharma Plc	Gniew	Poland	53.8333	18.8333	Pharmacy	Retail	Diaxolol	Antiseptics	20	765
Gottlieb-Cruickshank	Jakubowski Inc Pharmaceutical Limited	Bolesławiec	Poland	51.2667	15.5667	Pharmacy	Retail	Saprodeine Zanprosyn	Antibiotics	5	215
Gottlieb-Cruickshank	Schaden-Hirthe Pharmaceutical Ltd	Świdnica	Poland	50.8438	16.4886	Pharmacy	Institution	Magnemunex	Antiseptics	26	689
Gottlieb-Cruickshank	Mraz Group Pharmacy	Piastów	Poland	52.1833	20.85	Hospital	Private	Pazofenac	Mood Stabilizers	16	123

pharma_tb 1 ×

Output

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

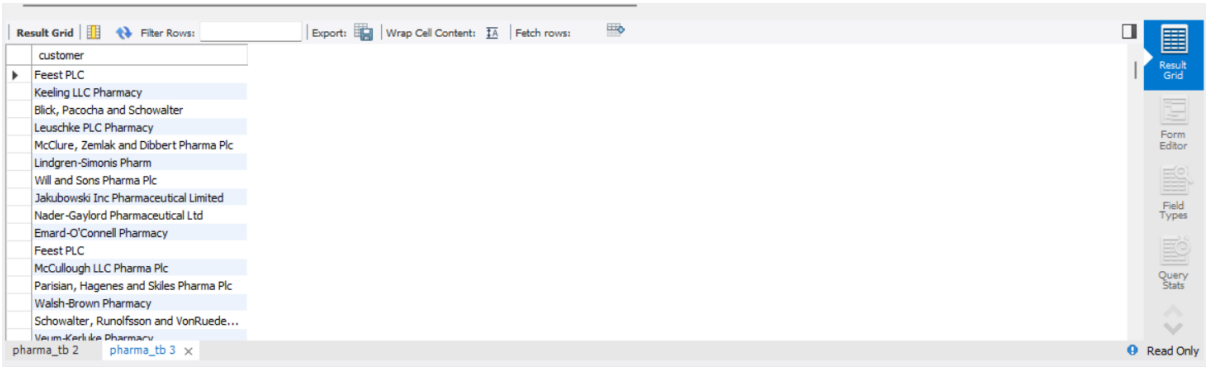
```
SELECT DISTINCT country from  pharma_tb;
```



country
Poland
Germany

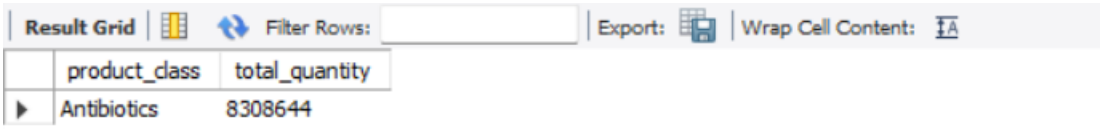
3. Select the names of all the customers on the 'Retail' channel.

```
SELECT customer

FROM pharma_tb

WHERE sub_channel = 'Retail';
```



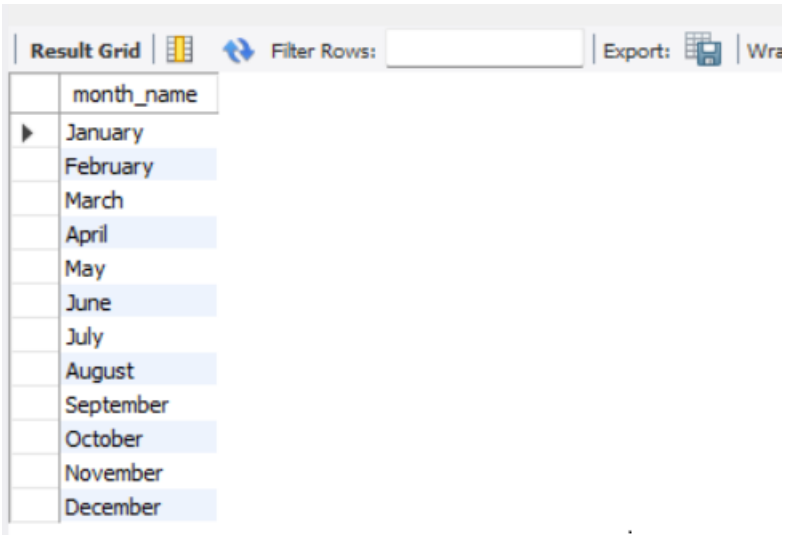
4. Find the total quantity sold for the ' Antibiotics' product class.

```
SELECT product_class ,sum(quantity) total_quantity

FROM pharma_tb

group by product_class

having product_class='Antibiotics';
```



5. List all the distinct months present in the dataset

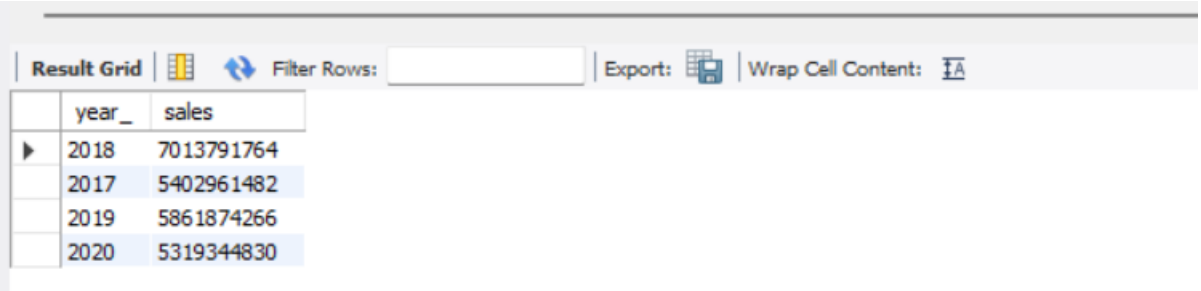
```
select distinct month_name from pharma_tb;
```



6. Calculate the total sales for each year

select year_,sum(sales) as sales

from pharma_tb group by year_;



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

year_	sales
2018	7013791764
2017	5402961482
2019	5861874266
2020	5319344830

7. Find the customer with the highest sales value.

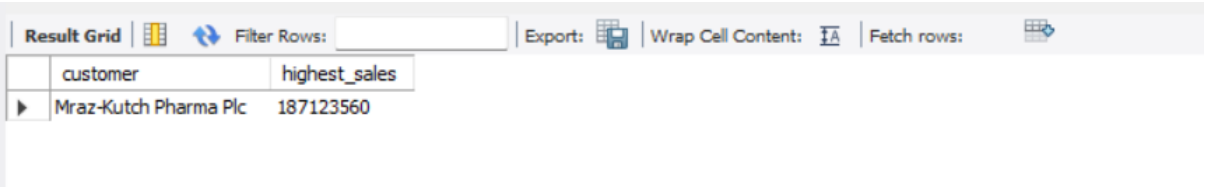
SELECT customer, SUM(sales) AS highest_sales

FROM pharma_tb

GROUP BY customer

ORDER BY highest_sales DESC

LIMIT 1;

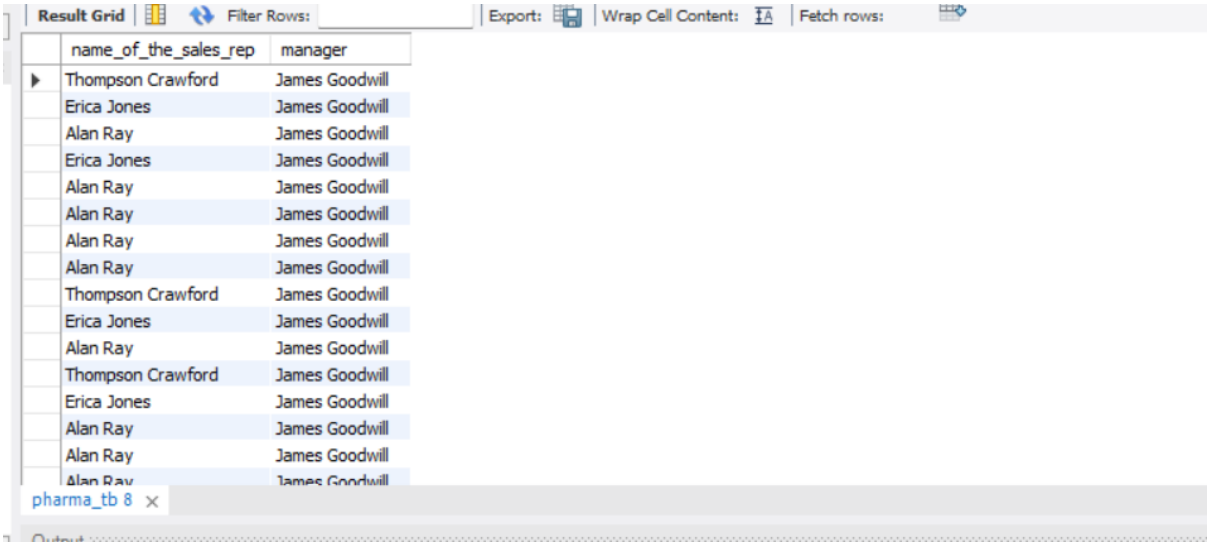


Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

customer	highest_sales
Mraz-Kutch Pharma Plc	187123560

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

select name_of_the_sales_rep,manager from pharma_tb where manager='James Goodwill';



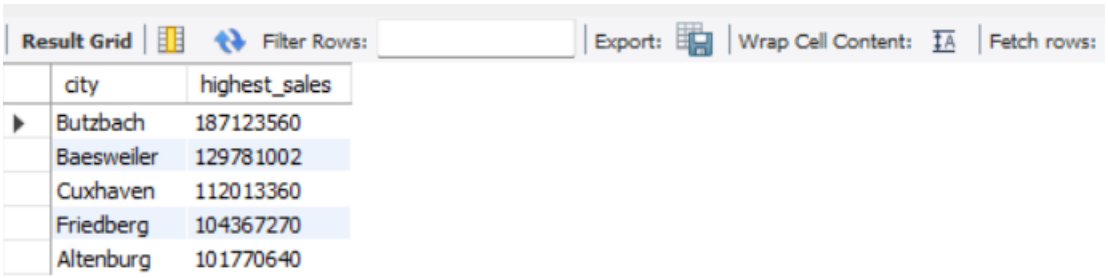
Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

name_of_the_sales_rep	manager
Thompson Crawford	James Goodwill
Erica Jones	James Goodwill
Alan Ray	James Goodwill
Erica Jones	James Goodwill
Alan Ray	James Goodwill
Alan Ray	James Goodwill
Alan Ray	James Goodwill
Alan Ray	James Goodwill
Thompson Crawford	James Goodwill
Erica Jones	James Goodwill
Alan Ray	James Goodwill
Thompson Crawford	James Goodwill
Erica Jones	James Goodwill
Alan Ray	James Goodwill
Alan Ray	James Goodwill
Alan Ray	James Goodwill

pharma_tb 8

9. Retrieve the top 5 cities with the highest sales

```
SELECT city, SUM(sales) AS highest_sales
FROM pharma_tb
GROUP BY city
ORDER BY highest_sales DESC
LIMIT 5;
```

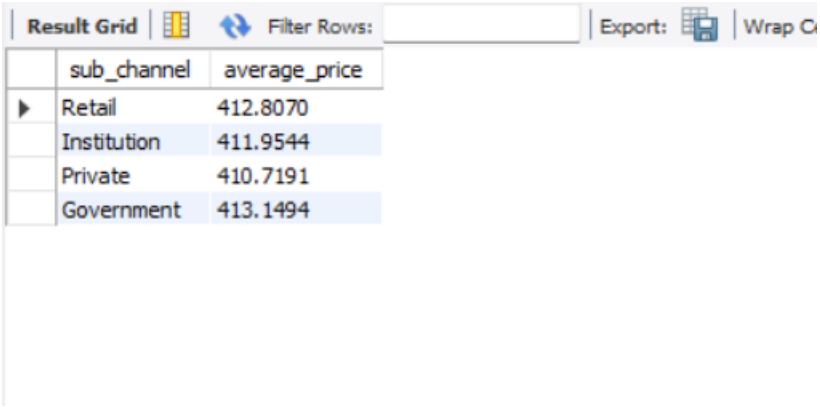


Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

city	highest_sales
Butzbach	187123560
Baesweiler	129781002
Cuxhaven	112013360
Friedberg	104367270
Altenburg	101770640

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

```
select sub_channel,avg(price) as average_price from pharma_tb
group by sub_channel;
```



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

sub_channel	average_price
Retail	412.8070
Institution	411.9544
Private	410.7191
Government	413.1494

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

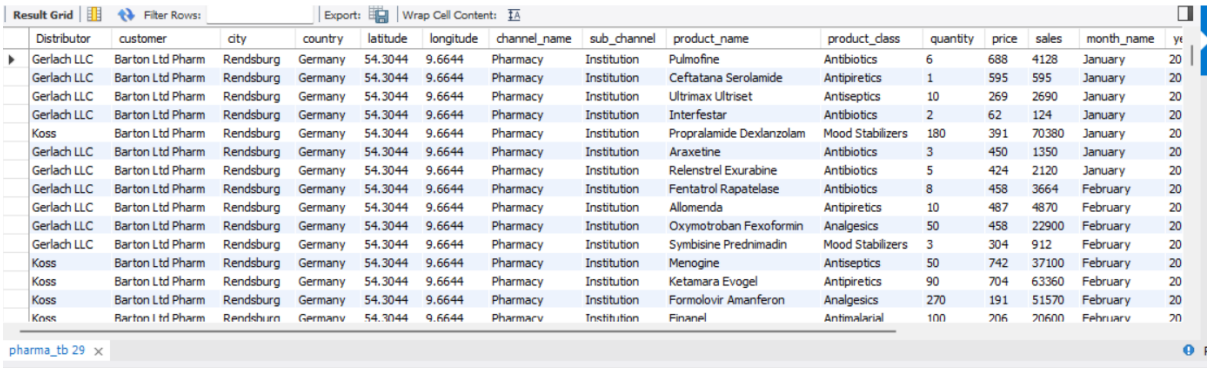
select name_of_the_sales_rep,sum(sales) from pharma_tb group by name_of_the_sales_rep;

Result Grid | Filter Rows:

name_of_the_sales_rep	sum(sales)
Jessica Smith	1763396738
Steve Pepple	1750899966
Mary Gerrard	1750538582
Anne Wu	1840336604
Thompson Crawford	1733929772
Sheila Stones	1916407796
Stella Given	1776681806
Morris Garcia	1802390966
Erica Jones	1742744384
Abigail Thompson	1962113986
Daniel Gates	1901317270
Alan Ray	1685274484
Jimmy Grey	1971939988

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

select*from pharma_tb where city='Rendsburg' and year_=2018;

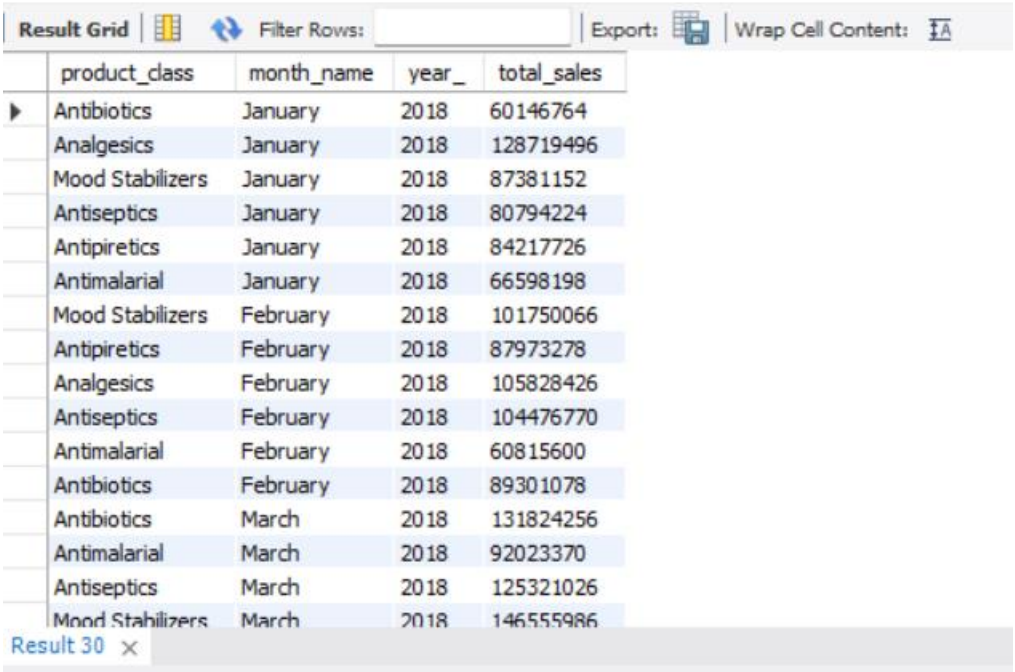


Distributor	customer	city	country	latitude	longitude	channel_name	sub_channel	product_name	product_class	quantity	price	sales	month_name	ye
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Pulmofine	Antibiotics	6	688	4128	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Ceftatana Serolamide	Antipiretics	1	595	595	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Ultrimax Ultriset	Antiseptics	10	269	2690	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Interfestar	Antibiotics	2	62	124	January	20
Koss	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Propralamide Dexlanzolam	Mood Stabilizers	180	391	70380	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Araxetine	Antibiotics	3	450	1350	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Relenstrel Exurabine	Antibiotics	5	424	2120	January	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Fentatrol Rapatelase	Antibiotics	8	458	3664	February	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Allomenda	Antipiretics	10	487	4870	February	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Oxymotroban Fexoformin	Analgesics	50	458	22900	February	20
Gerlach LLC	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Symbisine Prednimadin	Mood Stabilizers	3	304	912	February	20
Koss	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Menogine	Antiseptics	50	742	37100	February	20
Koss	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Ketamara Evogel	Antipiretics	90	704	63360	February	20
Koss	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Formolovir Amanferon	Analgesics	270	191	51570	February	20
Koss	Barton Ltd Pharm	Rendsburg	Germany	54.3044	9.6644	Pharmacy	Institution	Finanel	Antimalarial	100	206	20600	February	20

pharma_tb 29

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

```
SELECT product_class, month_name, year_, SUM(sales) as total_sales
FROM pharma_tb
GROUP BY product_class, month_name, year_
LIMIT 50000;
```



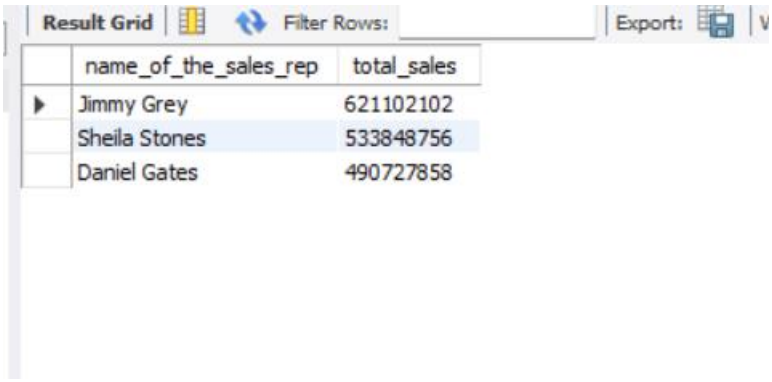
product_class	month_name	year_	total_sales
Antibiotics	January	2018	60146764
Analgesics	January	2018	128719496
Mood Stabilizers	January	2018	87381152
Antiseptics	January	2018	80794224
Antipiretics	January	2018	84217726
Antimalarial	January	2018	66598198
Mood Stabilizers	February	2018	101750066
Antipiretics	February	2018	87973278
Analgesics	February	2018	105828426
Antiseptics	February	2018	104476770
Antimalarial	February	2018	60815600
Antibiotics	February	2018	89301078
Antibiotics	March	2018	131824256
Antimalarial	March	2018	92023370
Antiseptics	March	2018	125321026
Mood Stabilizers	March	2018	146555986

Result 30

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

```
SELECT name_of_the_sales_rep, SUM(sales) AS total_sales
FROM pharma_tb
WHERE YEAR_ = 2019
GROUP BY name_of_the_sales_rep
ORDER BY total_sales DESC
```

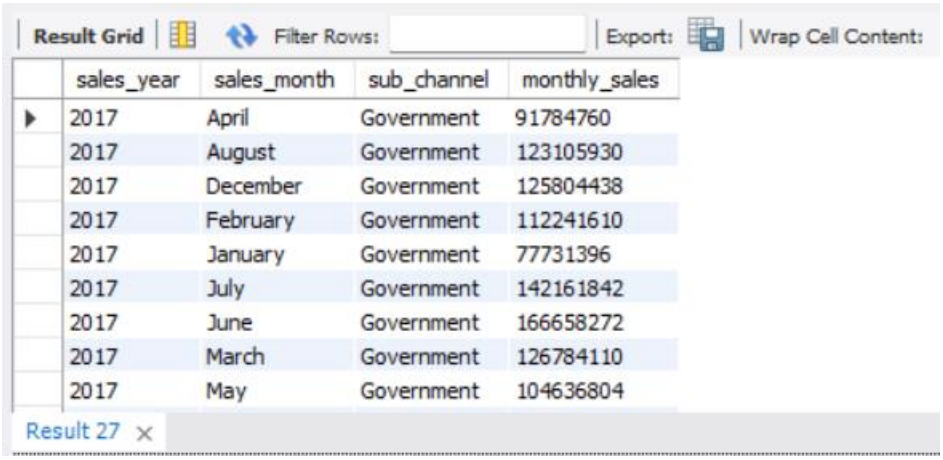
LIMIT 3;



name_of_the_sales_rep	total_sales
Jimmy Grey	621102102
Sheila Stones	533848756
Daniel Gates	490727858

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

```
SELECT
   YEAR_ AS sales_year,
   month_name  AS sales_month,
   sub_channel,
   SUM(sales) AS monthly_sales
FROM pharma_tb
GROUP BY sales_year, sales_month, sub_channel
ORDER BY sub_channel, sales_year, sales_month;
```



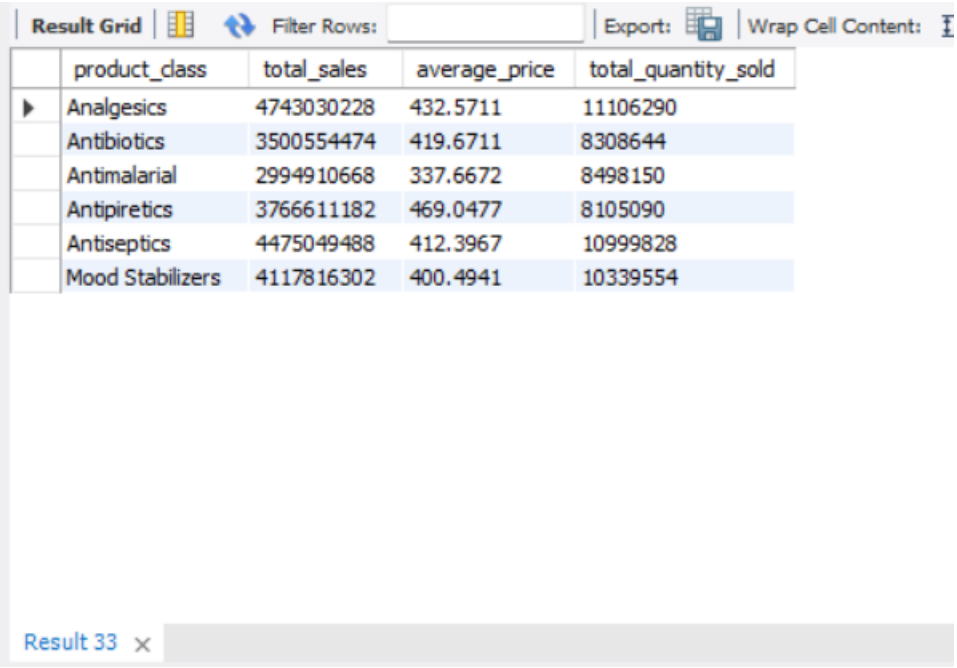
sales_year	sales_month	sub_channel	monthly_sales
2017	April	Government	91784760
2017	August	Government	123105930
2017	December	Government	125804438
2017	February	Government	112241610
2017	January	Government	77731396
2017	July	Government	142161842
2017	June	Government	166658272
2017	March	Government	126784110
2017	May	Government	104636804

Result 27

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

```
SELECT
```

```
    product_class,
    SUM(sales) AS total_sales,
    AVG(price) AS average_price,
    SUM(quantity) AS total_quantity_sold
FROM pharma_tb
GROUP BY product_class
ORDER BY product_class;
```



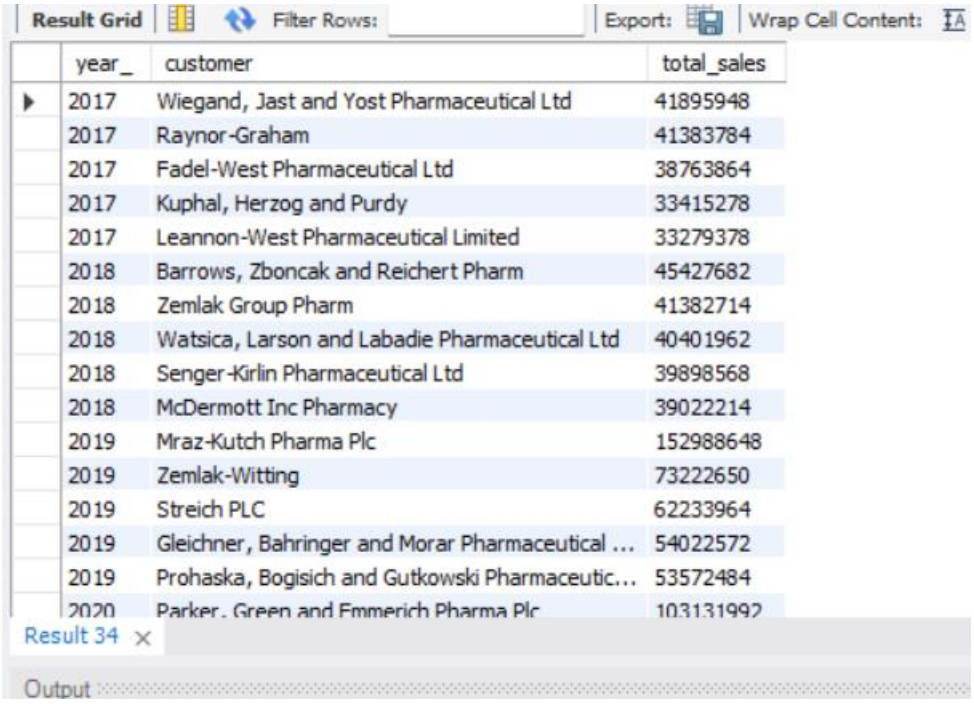
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

product_class	total_sales	average_price	total_quantity_sold
Analgesics	4743030228	432.5711	11106290
Antibiotics	3500554474	419.6711	8308644
Antimalarial	2994910668	337.6672	8498150
Antipiretics	3766611182	469.0477	8105090
Antiseptics	4475049488	412.3967	10999828
Mood Stabilizers	4117816302	400.4941	10339554

Result 33

17. Find the top 5 customers with the highest sales for each year.SELECT year_, customer, total_sales

```
FROM (
    SELECT year_, customer, SUM(sales) AS total_sales,
        ROW_NUMBER() OVER (PARTITION BY year_ ORDER BY SUM(sales) DESC) AS sales_rank
    FROM pharma_tb
    GROUP BY year_, customer
) ranked_sales
WHERE sales_rank <= 5;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

year_	customer	total_sales
2017	Wiegand, Jast and Yost Pharmaceutical Ltd	41895948
2017	Raynor-Graham	41383784
2017	Fadel-West Pharmaceutical Ltd	38763864
2017	Kuphal, Herzog and Purdy	33415278
2017	Leannon-West Pharmaceutical Limited	33279378
2018	Barrows, Zboncak and Reichert Pharm	45427682
2018	Zemlak Group Pharm	41382714
2018	Watsica, Larson and Labadie Pharmaceutical Ltd	40401962
2018	Senger-Kirlin Pharmaceutical Ltd	39898568
2018	McDermott Inc Pharmacy	39022214
2019	Mraz-Kutch Pharma Plc	152988648
2019	Zemlak-Witting	73222650
2019	Streich PLC	62233964
2019	Gleichner, Bahringer and Morar Pharmaceutical ...	54022572
2019	Prohaska, Bogisich and Gutkowski Pharmaceutic...	53572484
2020	Parker, Green and Emmerich Pharma Plc	103131992

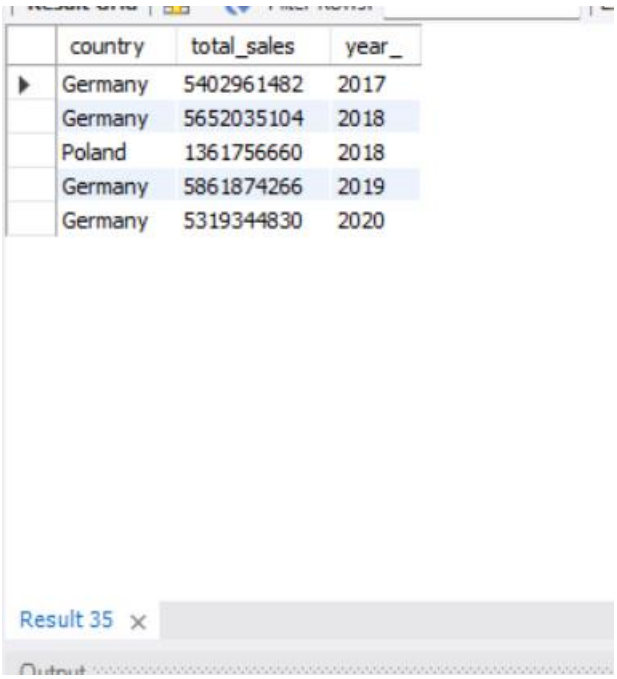
Result 34 ×

Output

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

select country,round(sum(sales))as total_sales,year_ from pharma_tb group by year_,country

order by year_;



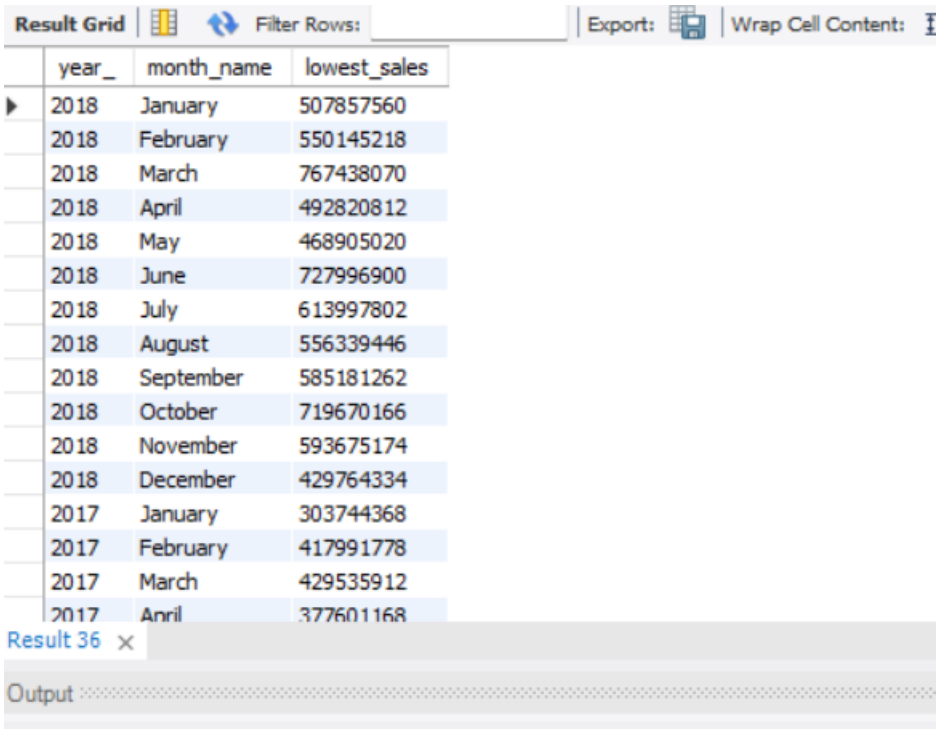
country	total_sales	year_
Germany	5402961482	2017
Germany	5652035104	2018
Poland	1361756660	2018
Germany	5861874266	2019
Germany	5319344830	2020

Result 35 ×

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

SELECT year_, month_name, MIN(total_sales) AS lowest_sales

```
FROM (
   SELECT YEAR_ AS year_, MONTH_NAME AS month_name, SUM(sales) AS total_sales
   FROM pharma_tb
   GROUP BY YEAR_, MONTH_NAME
) AS sales
GROUP BY year_, month_name
LIMIT 0, 50000;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

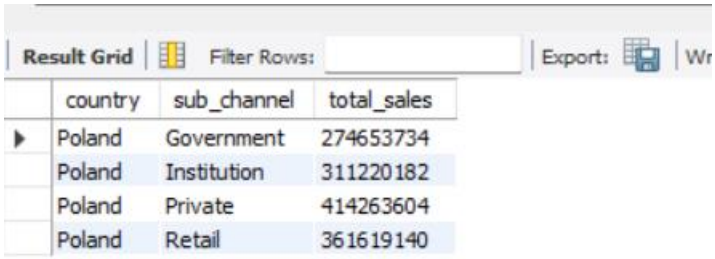
year_	month_name	lowest_sales
2018	January	507857560
2018	February	550145218
2018	March	767438070
2018	April	492820812
2018	May	468905020
2018	June	727996900
2018	July	613997802
2018	August	556339446
2018	September	585181262
2018	October	719670166
2018	November	593675174
2018	December	429764334
2017	January	303744368
2017	February	417991778
2017	March	429535912
2017	April	377601168

Result 36 | Output

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.

```
WITH tot_sales AS (
   SELECT
      RANK() OVER(PARTITION BY sub_channel ORDER BY SUM(sales)) AS rnk,
      SUM(sales) AS total_sales,
      country,
      sub_channel
```

```
    FROM pharma_tb
    GROUP BY country, sub_channel
)
SELECT
    tot_sales.country,
    tot_sales.sub_channel,
    tot_sales.total_sales
FROM tot_sales
WHERE tot_sales.rnk = 1;
```



Result Grid | Filter Rows: | Export: | Wr

country	sub_channel	total_sales
Poland	Government	274653734
Poland	Institution	311220182
Poland	Private	414263604
Poland	Retail	361619140

Result 37 ×

Output