PHARMA DATA ANALYSIS

In our project, we learned how to leverage SQL to manage data effectively in databases. Mastering about 20 different types of SQL queries, we were able to extract specific information, perform calculations, and organize data neatly.

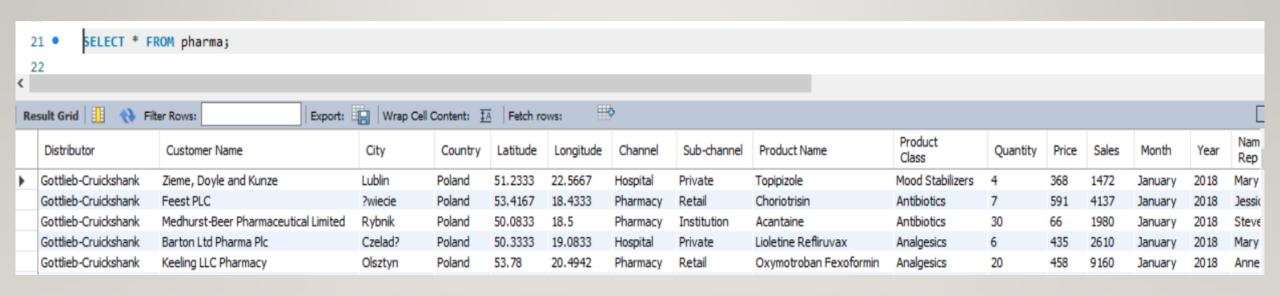
From simple tasks like finding a name in a list to more complex tasks like analyzing large amounts of data, SQL proved to be an invaluable tool.

This project underscored the importance of SQL in today's data-driven world, where efficiently handling large datasets is crucial for businesses.

Now, we're applying our SQL skills to analyze pharmaceutical sales data, enabling us to uncover valuable insights and drive informed decision-making for the pharmaceutical industry.

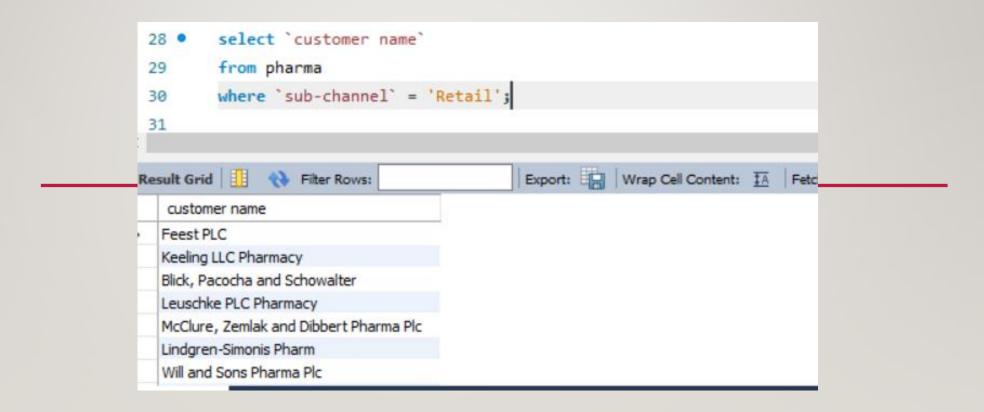


I. Retrieve all columns for all records in the dataset.



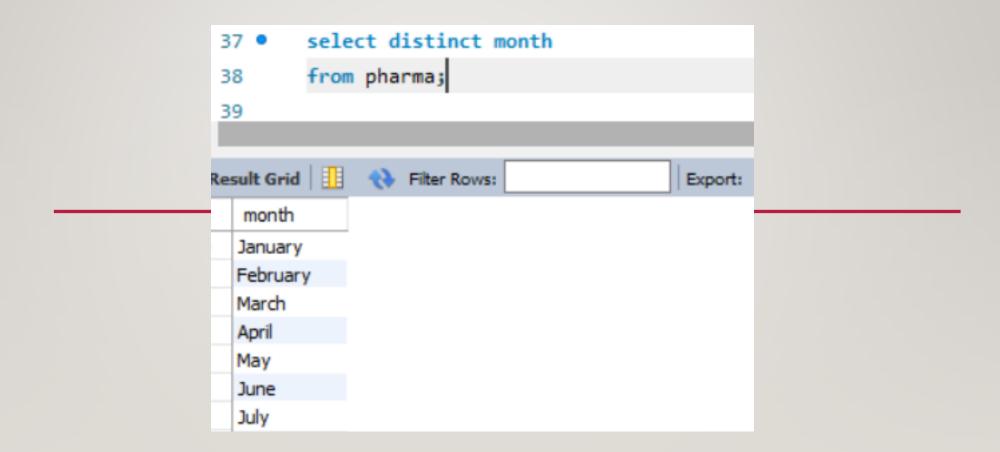
2. How many unique countries are represented in the dataset?

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



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

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



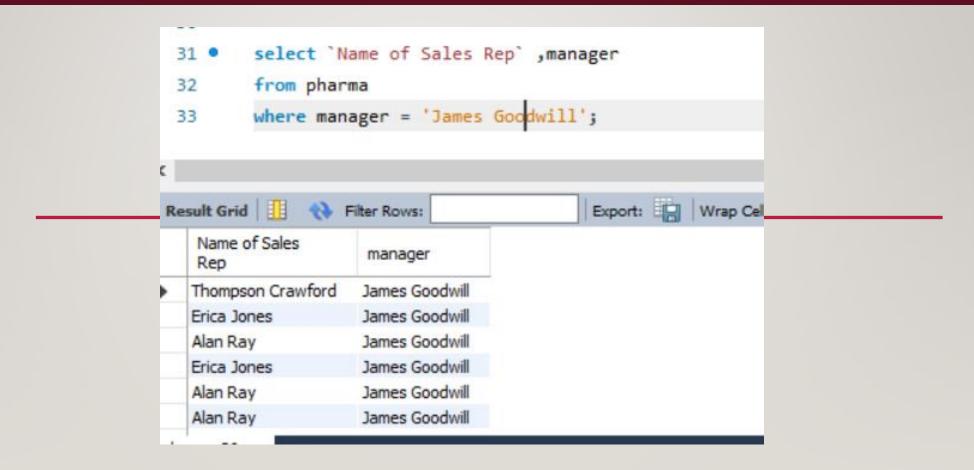
6. Calculate the total sales for each year.

```
select year, sum(sales) as total_sales
 40 •
         from pharma
 41
         group by year;
 42
 43
                                              Export: Wr.
Result Grid | | No. | Filter Rows:
          total_sales
   year
         3506897354
  2018
         2701480741
  2017
         2930937133
  2019
         2108573686
  2020
```

7. Find the customer with the highest sales value.

```
44 •
       select `customer name` , sum(sales) as total_sales
        from pharma
45
46
        group by 1
       order by 2 desc
47
        limit 1;
48
Export: Wrap Cell Content:
  customer name
                    total_sales
  Mraz-Kutch Pharma Plc 92441208
```

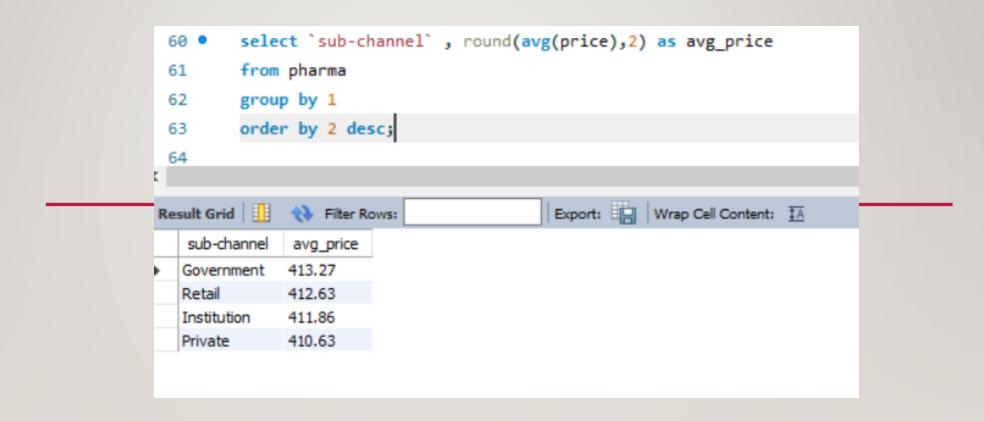
8. Get the names of all employees who are Sales Reps and are managed by 'James Goodwill'.



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

```
select city, sum(sales) as total_sales
 54 •
55
        from pharma
        group by 1
56
        order by 2 desc
        limit 5;
58
59
        select `sub-channel` , round(avg(price),2) as avg
                                        Export: Wrap Cel
total_sales
   city
  Butzbach
            92441208
  Baesweiler
            62935628
  Cuxhaven
            54910769
  Friedberg
            48843679
  Emsdetten
            45368998
```

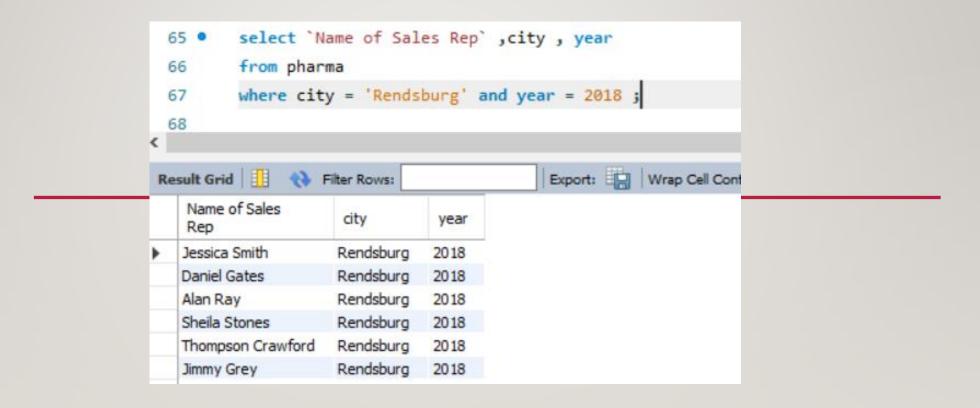
10. Calculate the average price of products in each sub-channel.



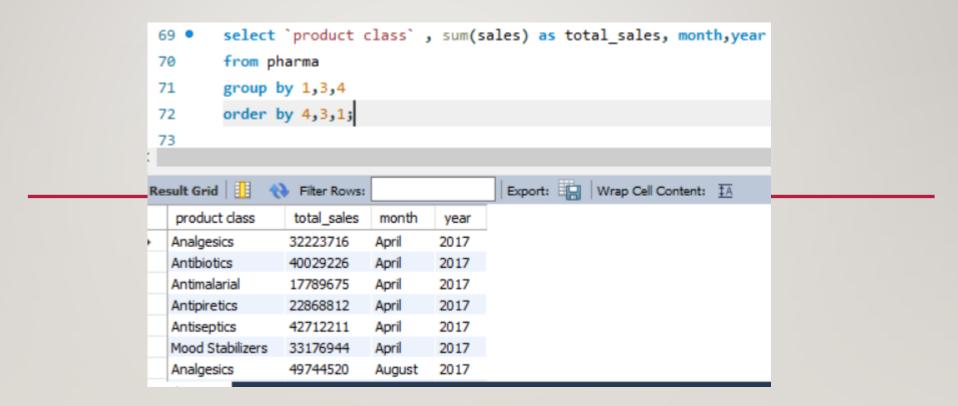
11. Join the 'Employees' table with the 'Sales' table to get the name of the Sales Rep and the corresponding sales records

3 €		ect p.name_of_sale from pharma_data p roup by p.name_of_		les)
Ħ	}: -		44 × 6 × 2 -44	
		NAME_OF_SALES_REP	SUM(P.SALES)	
>	1	Mary Gerrard	875270762.91093	
	2	Morris Garcia	901195482.5	
	3	Erica Jones	871372192	
	4	Abigail Thompson	981056993.864903	
	5	Jimmy Grey	985969993.944742	
	6	Jessica Smith	881698369.002429	
	7	Stella Given	888340902.41899	
	8	Steve Pepple	875449982.57143	
		Alan Ray	842637242.2	
100	10	Thompson Crawford	866964886.178331	
	11	Sheila Stones	958203898.244147	
100	12	Anne Wu	920168301.173581	
1	13	Daniel Gates "	950658635.185934	

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



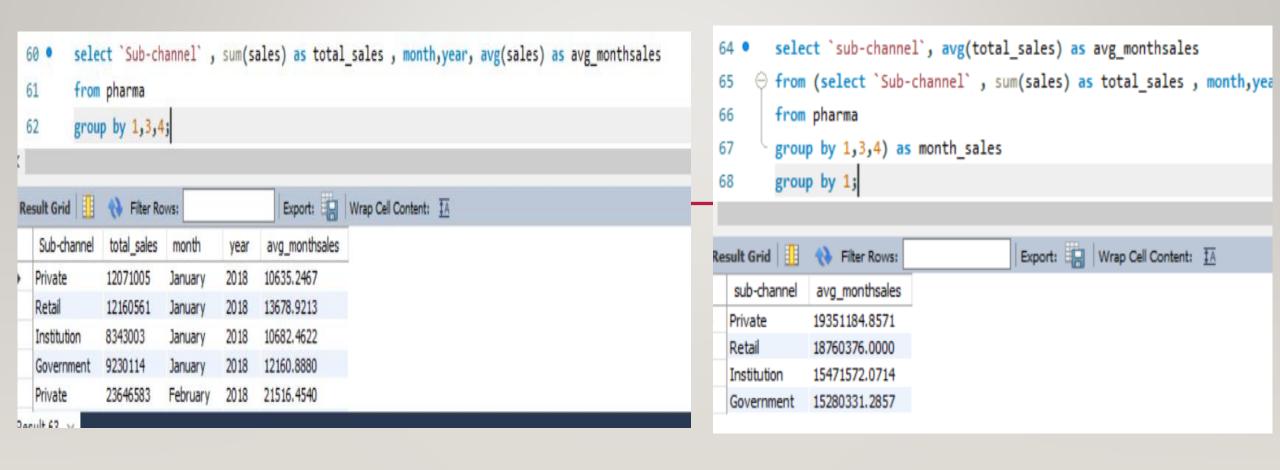
13. Calculate the total sales for each product class, for each month, and order the results by year, month, and product class.



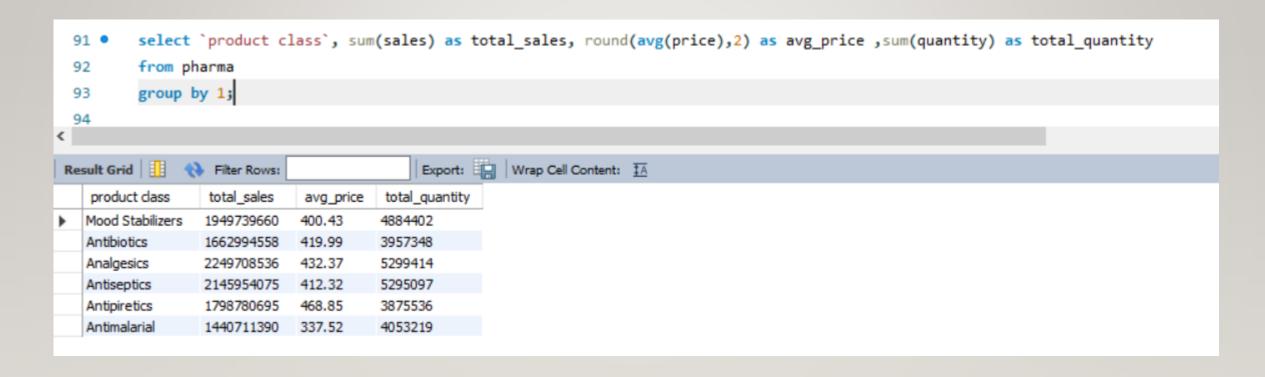
14. Find the top 3 sales reps with the highest sales in 2019.

```
select `Name of Sales Rep`, sum(sales) as total_sales
74 •
75
        from pharma
        where year = 2019
        group by 1
77
78
         order by 2 desc
        limit 3;
79
 80
Result Grid
                                            Export: Wrap Cell Content: 1A
              Filter Rows:
   Name of Sales
                   total_sales
   Rep
  Jimmy Grey
                   310551051
  Sheila Stones
                   266924378
  Daniel Gates
                   245363929
```

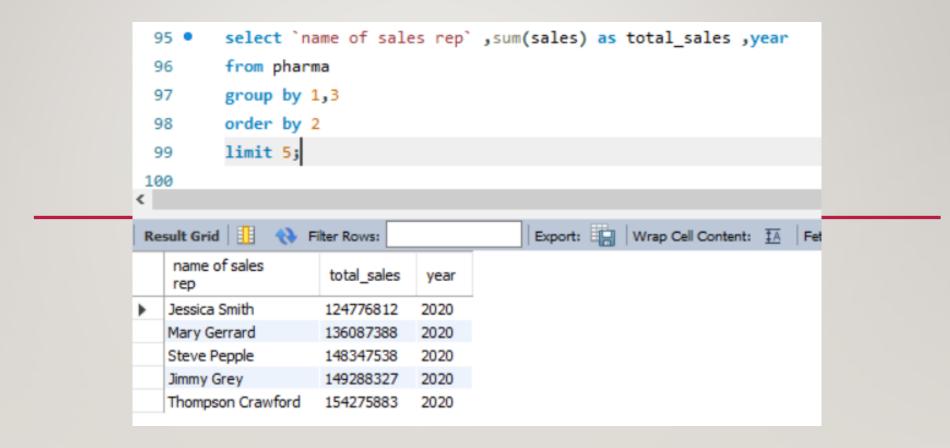
15. Calculate the monthly total sales for each sub-channel, and then calculate the average monthly sales for each sub-channel over the years.



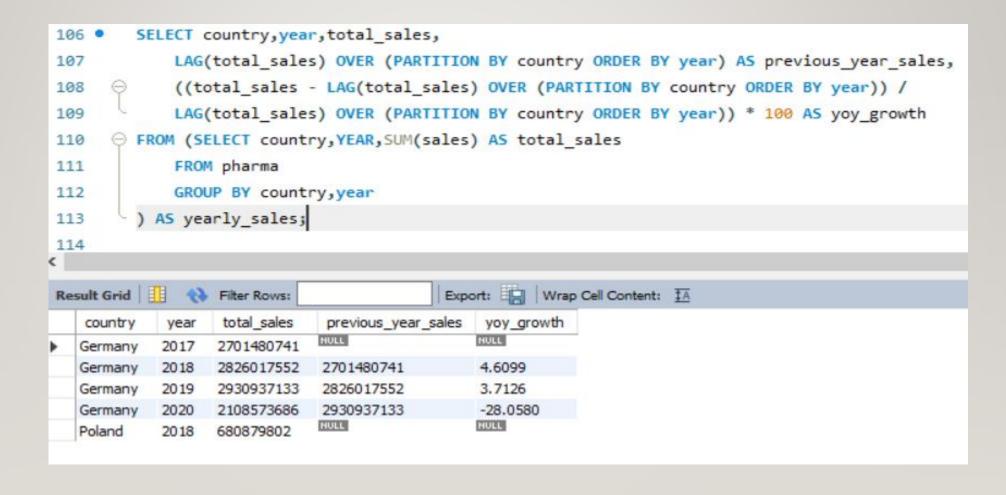
16. Create a summary report that includes the total sales, average price, and total quantity sold for each product class.



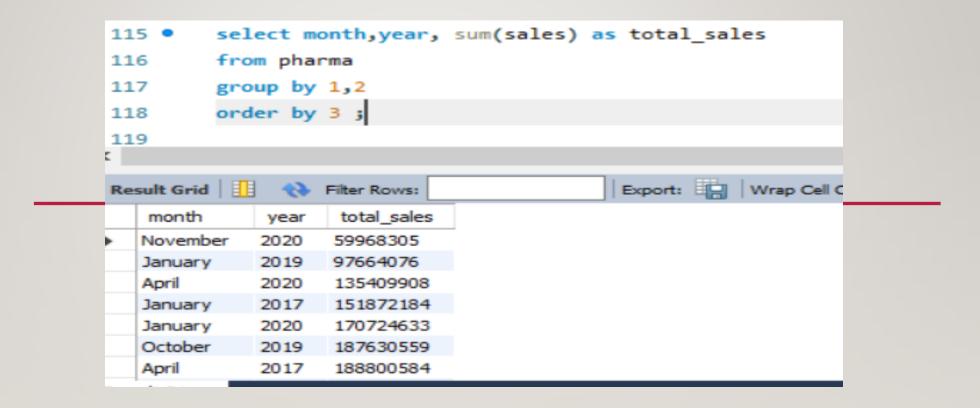
17. Find the top 5 customers with the highest sales for each year.



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



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



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

```
135 • ⊝ with subchannelsales as (select country, `sub-channel`, sum(sales) as total sales
        from pharma
136
137
        group by 1,2),
        maxsubchannel as (select `sub-channel`, max(total_sales) as max_sales
138
        from subchannelsales
139
        group by 1)
140
         select sc.country,sc.`sub-channel`,sc.total_sales from subchannelsales sc
141
         join maxsubchannel msc on sc.`sub-channel` = msc.`sub-channel` and sc.total_sales = msc.max_sales;
142
143
Result Grid Filter Rows:
                                       Export: Wrap Cell Content: TA
            sub-channel
                        total sales
   country
            Government
                       2763800068
                       2217743137
            Private
                       2607947184
            Institution
            Retail
                       2977518723
```