About the Dataset

The Dataset has 1 table and 18 columns

Table: pharma

Columns:

Column	Type
Distributor	varchar(255)
Customer_Name	text
City	varchar(255)
Country	varchar(20)
Latitude	double
Longitude	double
Channel	varchar(20)
Sub_channel	varchar(30)
Product_Name	varchar(255)
Product_Class	varchar(255)
Quantity	double
Price	int
Sales	double
Month	varchar(20)
Year	int
Name_of_Sales_Rep	varchar(30)
Manager	varchar(30)
Sales_Team	varchar(30)

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

```
Select * from Pharma;
```

OUTPUT

Distributor	Customer_Name	City	Country	Latitude	Longitude	Channel	Sub_channel	Product_Name	Product_Class
Gottlieb-Cruickshank	Zieme-Doyle and Kunze	Lublin	Poland	51.2333	22.5667	Hospital	Private	Topipizole	Mood Stabilizers
Gottlieb-Cruickshank	Feest PLC	Swiecie	Poland	53.4167	18.4333	Pharmacy	Retail	Choriotrisin	Antibiotics
Gottlieb-Cruickshank	Medhurst-Beer Pharmaceutical Limited	Rybnik	Poland	50.0833	18.5	Pharmacy	Institution	Acantaine	Antibiotics
Gottlieb-Cruickshank	Barton Ltd Pharma Plc	Czeladz	Poland	50.3333	19.0833	Hospital	Private	Lioletine Refliruvax	Analgesics
Gottlieb-Cruickshank	Keeling LLC Pharmacy	Olsztyn	Poland	53.78	20.4942	Pharmacy	Retail	Oxymotroban Fexoformin	Analgesics
Gottlieb-Cruickshank	Runte-Marquardt Pharmaceutical Ltd	Olecko	Poland	54.0333	22.5	Hospital	Private	Pazofenac	Mood Stabilizers
Gottlieb-Cruickshank	Blick-Pacocha and Schowalter	Inowroclaw	Poland	52.7958	18.2611	Pharmacy	Retail	Symbitrim	Analgesics
Gottlieb-Cruickshank	Leuschke PLC Pharmacy	Ciechanów	Poland	52.8817	20.6106	Pharmacy	Retail	Morphizolid Tianalin	Mood Stabilizers

This basic SQL function extracts all the results from the table.

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

```
Select distinct Country from Pharma;
```

OUTPUT

	Country
▶	Poland
	Germany

The query retrieves the **distinct countries** from the database. We have only **2 countries** in this table which is **Poland** & **Germany**.

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

```
Select Customer_Name,Sub_channel from pharma
where Sub_channel = "Retail";
```

OUTPUT

Customer_Name	Sub_channel
Feest PLC	Retail
Keeling LLC Pharmacy	Retail
Blick- Pacocha and Schowalter	Retail
Leuschke PLC Pharmacy	Retail
McClure- Zemlak and Dibbert Pharma Plc	Retail
Lindgren-Simonis Pharm	Retail
Will and Sons Pharma Plc	Retail
Jakubowski Inc Pharmaceutical Limited	Retail
Nader-Gaylord Pharmaceutical Ltd	Retail
Emard-O'Connell Pharmacy	Retail
Feest PLC	Retail
McCullough LLC Pharma Plc	Retail
Parisian- Hagenes and Skiles Pharma Plc	Retail

The query retrieves the **Customer Names** where **Sub-Channels Retail.**

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

```sql
Select count(*) as Count_Of_product_class, Product_Class
from pharma
Where Product_Class = "Antibiotics"
group by Product_Class;
```

OUTPUT

	Count_Of_product_class	Product_Class
▶	36979	Antibiotics

The output shows the **Total Quantity sold** for **Antibiotics.**

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

```
Select distinct Month from pharma;
```

OUTPUT

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

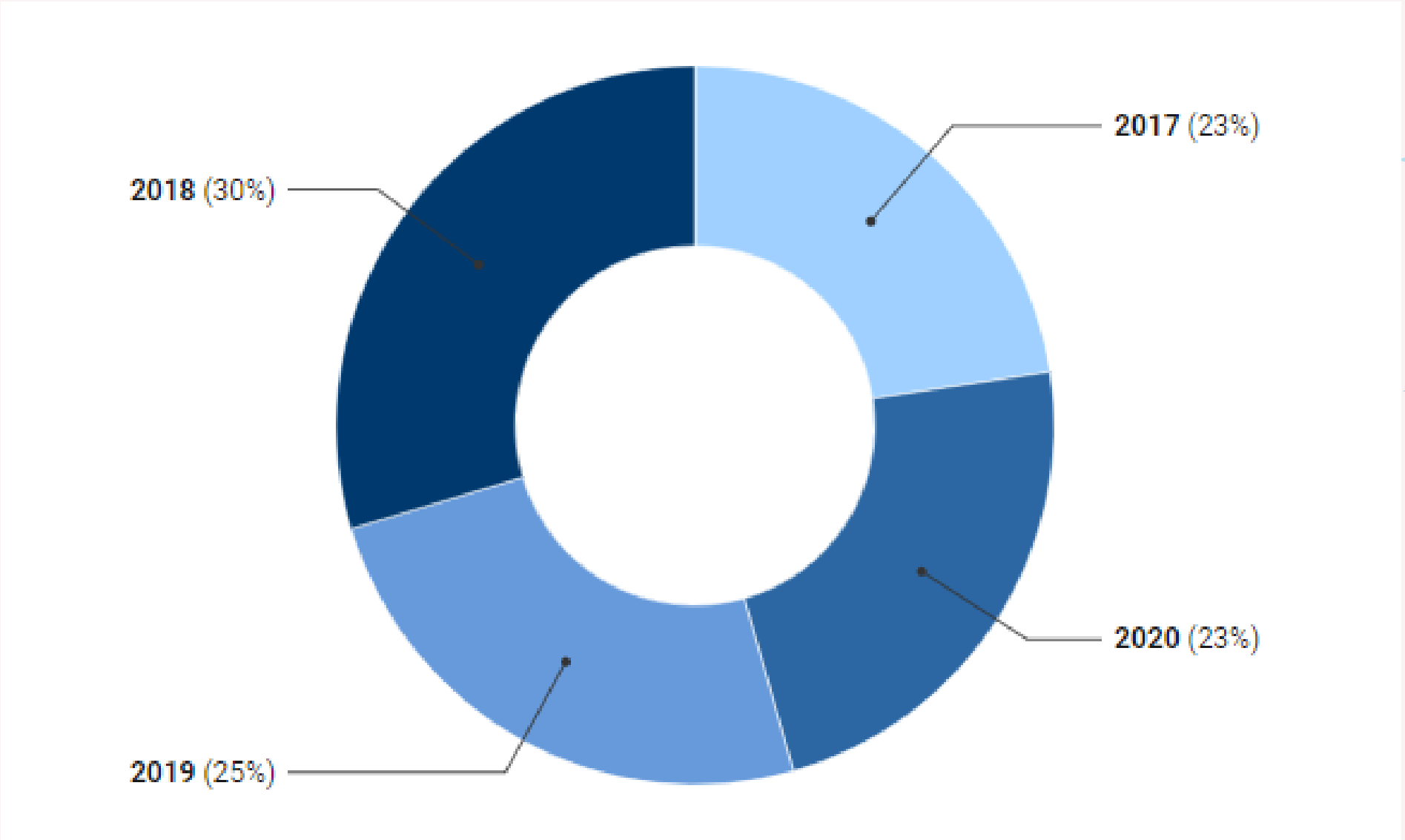The output shows all the **distinct months** present in the table.

6. Calculate the total sales for each year.

```sql
Select year, CEIL(Sum(Sales)) as Total_sales
from pharma
group by Year
Order by year;
```

OUTPUT

	year	Total_sales
▶	2017	2734379001
	2018	3584302586
	2019	2995829371
	2020	2777435779

2017 (23%)
2018 (30%)
2019 (25%)
2020 (23%)

7. Find the customer with the highest sales value.

```sql
Select Customer_Name, MAX(Sales) as Highest_Sale
from pharma
group by Customer_Name
order by Highest_Sale DESC limit 1;
```

OUTPUT

Customer_Name	Highest_Sale
▶ Mraz-Kutch Pharma Plc	74205600

The query extracts the **customer name with highest sales**, which is **Mraz-Kutch Pharma Plc** and Highest sale is **74205600**

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

```sql
Select Name_of_Sales_Rep, Manager
from pharma
where Manager = 'James Goodwill' ;
```

OUTPUT

Name_of_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

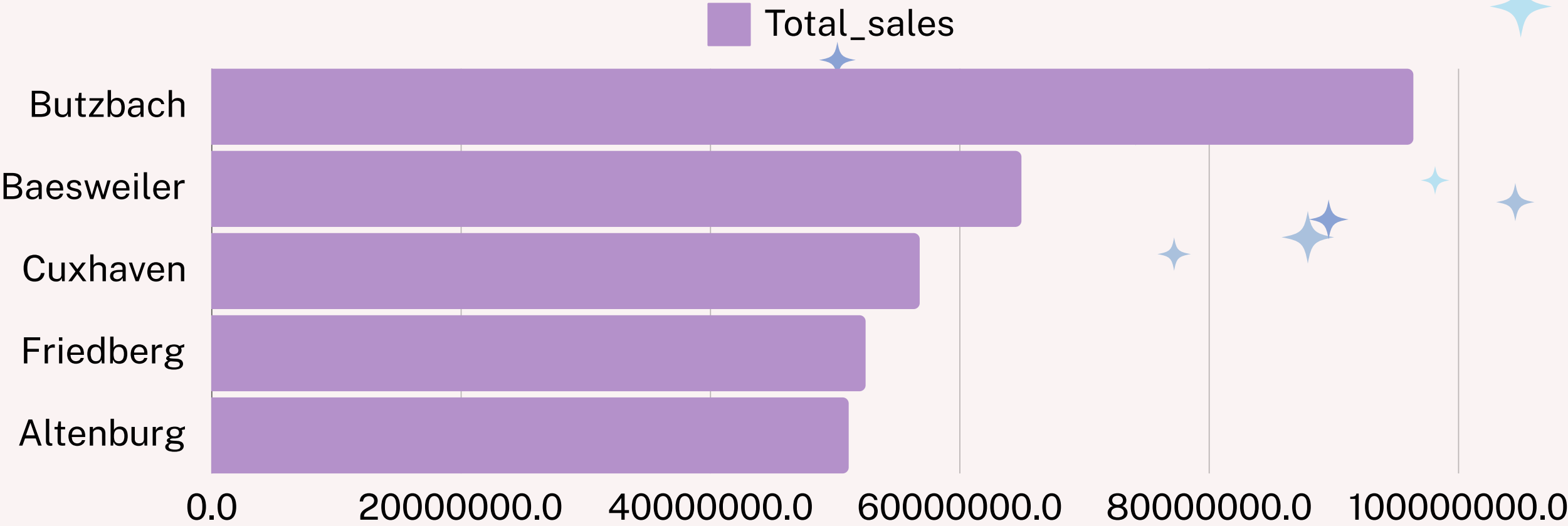The query retrieves the **all sales rep** name where **manager** is **James Goodwill.**

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

```sql
Select  City, Sum(Sales) as Total_sales
from Pharma
group by City
order by Total_sales DESC limit 5;
```

OUTPUT

	City	Total_sales
▶	Butzbach	96374052
	Baesweiler	64938713
	Cuxhaven	56779676
	Friedberg	52439674.6
	Altenburg	51084240

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

```sql
Select Sub_Channel, Round(AVG(Price),2) as Avg_Price
from pharma
group by Sub_Channel;
```

OUTPUT

Sub_Channel	Avg_Price
Private	410.72
Retail	412.81
Institution	411.95
Government	413.15

The query retrieves **average price** for all **sub channels.**

11. Get the name of the Sales Rep and the corresponding sales records.

```sql
Select Name_of_Sales_Rep, Count(Product_name) total_products,
Round(Sum(Quantity),2) total_quantities, Round(Sum(Sales),2) as total_sales
from pharma
group by Name_of_Sales_Rep;
```

OUTPUT

Name_of_Sales_Rep	total_products	total_quantities	total_sales
Mary Gerrard	19403	2159034.76	896371338.92
Jessica Smith	19623	2215327.16	905365331
Steve Pepple	19620	2295304.43	896917146.57
Anne Wu	19563	2281091.1	939393009.18
Thompson Crawford	19622	2178452.31	901037756.17
Sheila Stones	19655	2437304.12	980758388.24
Stella Given	19539	2158895.52	907751560.42
Morris Garcia	19461	2297423.81	918488474.5
Erica Jones	19471	2188352	894711806
Abigail Thompson	19571	2348109.05	999602125.87
Daniel Gates	19439	2371922.21	980465449.19
Alan Ray	19612	2123160.6	867166970.2
Jimmy Grey	19503	2339281.74	1003917379.94

The Query retrieves **Total number of products, Total quantities** and **Total Sales** generated **by each Sales Rep.**

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

```
Select Name_of_Sales_Rep, Sales, City, year
from pharma where city = "Rendsburg" and Year = 2018;
```

OUTPUT

Name_of_Sales_Rep	Sales	City	year
Jessica Smith	4128	Rendsburg	2018
Daniel Gates	595	Rendsburg	2018
Alan Ray	2690	Rendsburg	2018
Sheila Stones	124	Rendsburg	2018
Thompson Crawford	70380	Rendsburg	2018
Jimmy Grey	1350	Rendsburg	2018
Anne Wu	2120	Rendsburg	2018
Erica Jones	3664	Rendsburg	2018
Morris Garcia	4870	Rendsburg	2018
Stella Given	22900	Rendsburg	2018
Morris Garcia	912	Rendsburg	2018
Morris Garcia	37100	Rendsburg	2018
Morris Garcia	63360	Rendsburg	2018

This query shows the **Name of sales rep**, and their **sales** from year **2018** and from **Rendsburg** city .

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

```sql
Select Product_class, Month, year, Sum(sales) as total_sales
from pharma
group by  Product_class, Month, year
order by month, year, Product_class;
```

OUTPUT

Product_class	Month	year	total_sales
Analgesics	April	2017	32342274
Antibiotics	April	2017	40080462
Antimalarial	April	2017	19380719
Antipiretics	April	2017	23072216
Antiseptics	April	2017	43187735
Mood Stabilizers	April	2017	33596290
Analgesics	April	2018	41293802
Antibiotics	April	2018	53940546
Antimalarial	April	2018	31306902
Antipiretics	April	2018	37849709
Antiseptics	April	2018	51459139.8
Mood Stabilizers	April	2018	36098409
Analgesics	April	2019	75916169

This query shows the **Product Class,** and their **total sales** by each **Year** and **Month**
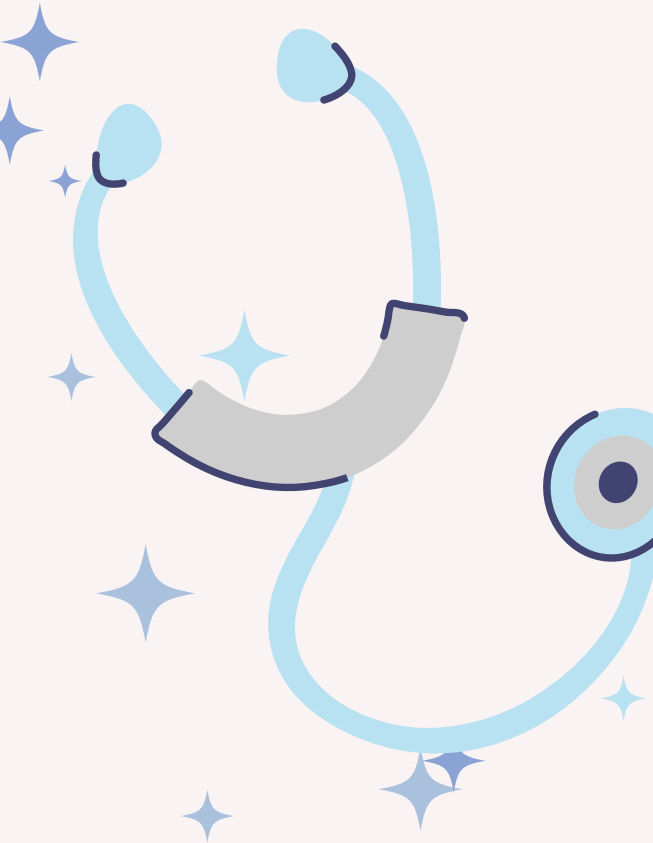
14. Find the top 3 sales reps with the highest sales in 2019.

```sql
Select  Name_of_Sales_Rep, Year, Sum(sales) as total_sales
from pharma
where Year = 2019
group by  Name_of_Sales_Rep, Year
order by total_sales DESC limit 3;
```

OUTPUT

Name_of_Sales_Rep	Year	total_sales
Jimmy Grey	2019	314995074.94
Sheila Stones	2019	279118750.24
Daniel Gates	2019	249718747.19

This query shows the **Name of the Sales Rep** and their **total sales** in year **2019** which were **top 3.**

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

```sql
SELECT Sub_channel,
       YEAR,
       Month,
       SUM(Sales) AS Total_Sales,
       Round(AVG(Sales),2) AS Avg_Sales
FROM pharma
GROUP BY Sub_channel, YEAR, Month;
```

OUTPUT

Sub_channel	YEAR	Month	Total_Sales	Avg_Sales
Government	2017	April	46409590	43131.59
Government	2017	August	63241033	46914.71
Government	2017	December	63455907	61369.35
Government	2017	February	57220319	48864.49
Government	2017	January	38914778	36233.5
Government	2017	July	71116089	48843.47
Government	2017	June	85616946	65406.38
Government	2017	March	63679935	49097.87
Government	2017	May	53129802	38140.56
Government	2017	November	59974317	43428.18
Government	2017	October	45432266	36000.21
Government	2017	September	61474367	49061.75
Government	2018	April	72582193	42495.43

This SQL query shows the **total sales** and **average sales** per **month** for **each sub-channel** in the pharmaceutical dataset, **grouped by sub-channel, year, and month.**

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

```sql
Select Product_class, Round(AVG(price),2) as Avg_Price,
Round(Sum(Quantity),2) as total_quantity, Round(Sum(sales),2) as total_sales
from Pharma
group by Product_class;
```

OUTPUT

	Product_class	Avg_Price	total_quantity	total_sales
▶	Mood Stabilizers	400.49	5291909.14	2099699926.64
	Antibiotics	419.67	4272215.85	1801884712.54
	Analgesics	432.57	5711335.78	2435020926.29
	Antiseptics	412.40	5616306.72	2288548713.65
	Antipiretics	469.05	4170238.06	1939003231.18
	Antimalarial	337.67	4331653.26	1527789225.9

This SQL query calculates the **average price, total quantity,** and **total sales** for **each product class** in the pharmaceutical dataset. It groups the data by product class and rounds the average price, total quantity, and total sales to two decimal places.

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

```sql
Select Customer_name, YEAR, Highest_sales, Highest_yearly_sales_ranked from
(
Select Customer_name, YEAR, Sum(sales) as Highest_sales,
dense_rank() over (partition by Year order by Sum(sales) DESC) as Highest_yearly_sales_ranked
from pharma
group by Customer_name, YEAR
) as subquery
where Highest_yearly_sales_ranked between 1 and 5;
```

OUTPUT

Customer_name	YEAR	Highest_sales	Highest_yearly_sales_ranked
Wiegand- Jast and Yost Pharmaceutical Ltd	2017	20949794	1
Raynor-Graham	2017	20787932	2
Fadel-West Pharmaceutical Ltd	2017	19381932	3
Leannon-West Pharmaceutical Limited	2017	16715449	4
Kuphal- Herzog and Purdy	2017	16711749	5
Barrows- Zboncak and Reichert Pharm	2018	22867841	1
Watsica- Larson and Labadie Pharmaceutical Ltd	2018	20978197	2
Zemlak Group Pharm	2018	20691357	3
Senger-Kirlin Pharmaceutical Ltd	2018	20055524	4
McDermott Inc Pharmacy	2018	19909167	5
Mraz-Kutch Pharma Plc	2019	76528560	1
Zemlak-Witting	2019	36676029	2
Streich PLC	2019	31558442	3

This SQL query shows the **top 5 customers** with the **highest yearly sales in each year**, along with their corresponding sales amounts and ranking based on their yearly sales. It filters the results to include only the customers ranked between 1 and 5 in terms of yearly sales.

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

```sql
SELECT Country, YEAR, SUM(Sales) AS Total_Sales,
    LAG(SUM(Sales)) OVER (PARTITION BY Country ORDER BY YEAR) AS Previous_Year_Sales,
    CASE WHEN LAG(SUM(Sales)) OVER (PARTITION BY Country ORDER BY YEAR ROWS BETWEEN 1 PRECEDING AND 1 PRECEDING) IS NOT NULL THEN
    (SUM(Sales) - LAG(SUM(Sales)) OVER (PARTITION BY Country ORDER BY YEAR ROWS BETWEEN 1 PRECEDING AND 1 PRECEDING)) / LAG(SUM(Sales))
    OVER (PARTITION BY Country ORDER BY YEAR ROWS BETWEEN 1 PRECEDING AND 1 PRECEDING) * 100
    ELSE NULL
    END AS Year_Over_Year_Growth
FROM pharma
GROUP BY Country, YEAR
ORDER BY Country, YEAR;
```

OUTPUT

Country	YEAR	Total_Sales	Previous_Year_Sales	Year_Over_Year_Growth
Germany	2017	2734379000.8199997	NULL	NULL
Germany	2018	2903061325.8	2734379000.8199997	6.168944573133979
Germany	2019	2995829370.78	2903061325.8	3.19552481222338
Germany	2020	2777435779	2995829370.78	-7.289920911721978
Poland	2018	681241259.8	NULL	NULL

This SQL query shows the **total sales for each country in each year**, along with the sales from the previous year and the **year-over-year growth rate**. It calculates the year-over-year growth rate as the percentage change in sales from the previous year to the current year for each country.

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

```sql
Select Year, Month, Lowest_Sales
from
(
Select Year, Month, Sum(sales) as Lowest_sales,
Dense_rank() over (partition by YEAR order by Sum(sales) ASC) as Ranked_as_per_lowest_sale
from pharma
group by Year, Month
) as Subquery
Where Ranked_as_per_lowest_sale = 1
order by Year, Month ;
```

OUTPUT

	Year	Month	Lowest_Sales
▶	2017	January	152062486
	2018	December	218797059
	2019	January	99897488
	2020	April	144991672

This SQL query retrieves the **year, month, and lowest sales for each month across all years.** It calculates the lowest sales for each month and assigns rankings based on these lowest sales amounts. Then, it selects the records where the lowest sales are ranked as 1, indicating the lowest sales amount for each month, and sorts the results by year and month.
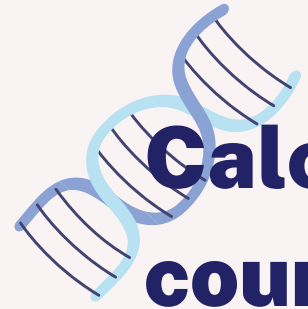
Calculate the total sales for each sub-channel in each country, and then find the country with the highest total sales for each sub-channel.

```sql
WITH cte AS (
    SELECT Sub_Channel,Country,SUM(Sales) AS Total_Sales,
    RANK() OVER (PARTITION BY Sub_Channel ORDER BY SUM(Sales) DESC) AS Sales_Rank
    FROM pharma
    GROUP BY Sub_Channel, Country
)
SELECT Sub_Channel, Country, Total_Sales
FROM cte
WHERE Sales_Rank = 1;
```

OUTPUT

Sub_Channel	Country	Total_Sales
Government	Germany	2994131420.93
Institution	Germany	2798082735.51
Private	Germany	2380073631.5599995
Retail	Germany	3238417688.4

This SQL query selects the **top-selling country within each sub-channel** of the pharmaceutical sales data. It first calculates the **total sales** for each **country** within **each sub-channel** and assigns a rank based on the sales amount. Then, it selects the records where the sales rank is equal to 1, indicating the country with the highest sales within each sub-channel.

Thank you for your attention