Welcome to my **Pharma** Data Analysis project done as part of my Internship with PSYLIQ. This program helps unlock SQL's potential in understanding data.

Pharma Data Assessment Details

1. **Retrieve all columns for all records in the dataset.**
   SELECT * FROM Pharma_data;

2. **How many unique countries are represented in the dataset?**
   SELECT COUNT(DISTINCT Country) AS UniqueCountriesCount
   FROM Pharma_data;

3. **Select the names of all the customers on the 'Retail' channel.**
   SELECT Customer_Name
   FROM Pharma_data
   WHERE Channel = 'Retail';

4. **Find the total quantity sold for the ' Antibiotics' product class.**
   SELECT SUM(Quantity) AS TotalQuantity
   FROM Pharma_data
   WHERE Product_Class = 'Antibiotics';

5. **List all the distinct months present in the dataset.**
   SELECT DISTINCT Month
   FROM Pharma_data;

6. **Calculate the total sales for each year.**
   SELECT Year,SUM(Sales) AS TotalSales
   FROM Pharma_data
   GROUP BY Year;

7. **Find the customer with the highest sales value.**
   ```sql
   SELECT TOP 1 Customer_Name,MAX(Sales) AS HighestSales
   FROM Pharma_data
   GROUP BY Customer_Name
   ORDER BY HighestSales DESC;
   ```

8. **Get the names of all employees who are Sales Reps and are managed by 'James Goodwill'.**
   ```sql
   SELECT DISTINCT(a.[Name_of_Sales_Rep])
   FROM pharma_data a
   JOIN pharma_data m ON a.Manager = m.[Name_of_Sales_Rep]
   WHERE m.Manager = 'James Goodwill'
    AND a.[Sales_Team] = 'Sales Rep';
   ```

9. **Retrieve the top 5 cities with the highest sales.**
   ```sql
   SELECT TOP 5 City,SUM(Sales) AS HighestSales
    FROM Pharma_data
    GROUP BY City
    ORDER BY HighestSales DESC;
   ```

10. **Calculate the average price of products in each sub-channel.**
    ```sql
    SELECT Sub_channel,AVG(Price) AS AveragePrice
     FROM Pharma_data
     GROUP BY Sub_channel;
    ```

11. **Join the 'Employees' table with the 'Sales' table to get the name of the Sales Rep and the corresponding sales records.**
    ```sql
    --SELECT e.Employee_Name, p.*
    --FROM Employees AS e
    --JOIN pharma_data AS p ON e.Name_of_Sales_Rep = p.Name_of_Sales_Rep;
    ```

12. **Retrieve all sales made by employees from ' Rendsburg ' in the year 2018.**
    ```sql
    SELECT *
    FROM Pharma_data
    WHERE City = 'Rendsburg' AND Year = '2018';
    ```

13. **Calculate the total sales for each product class, for each month, and order the results by year, month, and product class.**
    ```sql
    SELECT Year,Month,Product_Class
    FROM Pharma_data
    GROUP BY Year,Month,Product_Class
    ORDER BY Year,Month,Product_Class;
    ```

14. Find the top 3 sales reps with the highest sales in 2019.
```sql
SELECT TOP 3 Name_of_Sales_Rep,SUM(Sales) AS HighestSales
FROM Pharma_data
WHERE Year = 2019
GROUP BY Name_of_Sales_Rep
ORDER BY HighestSales;
```

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
WITH MonthlyTotalSales AS (
    SELECT
        Year AS SalesYear,
        Month AS SalesMonth,
        Sub_channel,
        SUM(Sales) AS MonthlySales
    FROM Pharma_data
    GROUP BY Year,Month,Sub_channel
),
AverageMonthlySales AS (
    SELECT
        Sub_channel,
        AVG(MonthlySales) AS AvgMonthlySales
    FROM MonthlyTotalSales
    GROUP BY Sub_channel
)
SELECT
    Sub_channel,
    AVG(AvgMonthlySales) AS AverageMonthlySales
FROM AverageMonthlySales
GROUP BY Sub_channel;
```

16. Create a summary report that includes the total sales, average price, and total quantity sold for each product class.
```sql
SELECT
    Product_Class,
    SUM(Sales) AS TotalSales,
    AVG(Price) AS AveragePrice,
    SUM(Quantity) AS TotalQuantity
FROM Pharma_data
GROUP BY Product_Class;
```

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

```sql
WITH RankedSales AS (
    SELECT Customer_Name,Year AS SalesYear,SUM(Sales) AS TotalSales,
    ROW_NUMBER() OVER(PARTITION BY Year ORDER BY SUM(Sales) DESC) AS
SalesRank
    FROM Pharma_data
    GROUP BY Customer_Name,Year
)
SELECT Customer_Name,SalesYear,TotalSales
FROM RankedSales
WHERE SalesRank <=5;
```

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

```sql
WITH SalesByYear AS (
    SELECT Year As SalesYear,Country,SUM(Sales) AS TotalSales
    FROM Pharma_data
    GROUP BY Year,Country
)
SELECT Country,SalesYear,TotalSales,
LAG(TotalSales) OVER(PARTITION BY Country ORDER BY SalesYear) AS
PreviousYearSales,
CASE
    WHEN LAG(TotalSales) OVER(PARTITION BY Country ORDER BY SalesYear) IS
NULL THEN NULL
    ELSE (TotalSales -LAG(TotalSales) OVER(PARTITION BY Country ORDER BY
SalesYear)) / LAG(TotalSales) OVER (PARTITION BY Country ORDER BY SalesYear)
*100
END AS YearOverYearGrowth
FROM SalesByYear
ORDER BY Country,SalesYear;
```

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

```sql
SELECT Year, Month, MIN(Sales) AS LowestSales
FROM Pharma_data
GROUP BY Year, Month
ORDER BY 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.

WITH SubChannelSales AS (
    SELECT
        Country,Sub_channel,SUM(Sales) AS TotalSales,
        RANK() OVER(PARTITION BY Sub_Channel ORDER BY SUM(Sales) DESC) AS
CountryRank
    FROM Pharma_data
    GROUP BY Country,Sub_Channel
)
SELECT Country,Sub_Channel,TotalSales,CountryRank
FROM SubChannelSales
WHERE CountryRank = 1
ORDER BY Sub_Channel;

Results From SQL Server

	Name_of_Sales_Rep	HighestSales
1	Thompson Crawford	189365359.178331
2	Mary Gerrard	194240491.91093
3	Anne Wu	201172379.664681

	Sub_channel	AverageMonthlySales
1	Retail	69647852.0914831
2	Institution	59900317.464489
3	Government	63713338.4988745
4	Private	52550734.4908911

	Product_Class	TotalSales	AveragePrice	TotalQuantity
1	Antipiretics	1883305591.17649	469	4052544.0572775
2	Antimalarial	1497455333.90892	337	4249075.24967097
3	Antibiotics	1750277236.54387	419	4154321.8570175
4	Mood Stabili...	2058909622.63676	400	5169781.14213925
5	Analgesics	2371515114.28388	432	5553143.78359865
6	Antiseptics	2237524743.6455	412	5499912.71284735

	Customer_Name	SalesYear	TotalSales
1	Parker, Green and Emmerich Pharma Plc	2020	51565996
2	Torphy, Pfeffer and Jakubowski Pharm	2020	27598295
3	Stehr-Champlin Pharmacy	2020	26698534
4	Runolfsson, Swaniawski and Jaskolski ...	2020	25132379
5	Goldner-Tillman Pharm	2020	24981248
6	Mraz-Kutch Pharma Plc	2019	76494324
7	Zemlak-Witting	2019	36611325
8	Streich PLC	2019	31116982

	Country	SalesYear	TotalSales	PreviousYearSales	YearOverYearGrowth
1	Germany	2017	2701480740.81559	NULL	NULL
2	Germany	2018	2826017551.8	2701480740.81559	4.60994628252697
3	Germany	2019	2930937132.77983	2826017551.8	3.71263019626326
4	Germany	2020	2659672415	2930937132.77983	-9.25522129922137
5	Poland	2018	680879801.8	NULL	NULL

	Year	Month	LowestSales
1	2017	October	-301600
2	2017	June	-984000
3	2017	April	-702900

	Year	Month	LowestSales
1	2017	October	-301600
2	2017	June	-984000
3	2017	April	-702900
4	2017	July	-176400
5	2017	Dece...	-133250
6	2017	January	-29325
7	2017	Febru...	-502400
8	2017	March	-345600

	Country	Sub_Channel	TotalSales	CountryRank
1	Germany	Government	2920913380.94598	1
2	Germany	Institution	2719605147.49547	1
3	Germany	Private	2315301981.56277	1
4	Germany	Retail	3162287330.39119	1