

Pharma Data Analysis

Task-2

**Aim:** To analyse the given dataset ‘Pharma\_data\_analysis.xlsx’ and perform the following queries in MS SQL.

**Database Name:** Pharma\_data$

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

**ANS:**

SELECT \* FROm Pharma\_data$;

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

**ANS:**

SELECT DISTINCT(country) FROM Pharma\_data$;

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

**ANS:**

SELECT [Customer Name] FROM Pharma\_data$ Where [Sub-channel]='Retail';

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

**ANS:**

SELECT COUNT(Quantity) AS [total quantity] FROM Pharma\_data$ WHERE [Product Class]='Antibiotics' ;

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

**ANS:**

SELECT DISTINCT(Month) FROM Pharma\_data$;

**6. Calculate the total sales for each year.**

**ANS:**

SELECT SUM(Sales),[Year] FROM Pharma\_data$ GROUP BY [Year];

**7. Find the customer with the highest sales value.**

**ANS:**

SELECT [Customer Name] FROM Pharma\_data$ GROUP BY[Customer Name] ORDER BY MAX(Sales) DESC;

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

**ANS:**

SELECT DISTINCT([Name of Sales Rep]) FROM Pharma\_data$ WHERE Manager='James Goodwill';

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

**ANS:**

SELECT TOP 5 City,MAX(Sales) FROM Pharma\_data$ GROUP BY City ;

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

**ANS:**

SELECT AVG(price),[Sub-channel] 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.**

**ANS:**

SELECT E.[Emp Name], p.Sales FROM general\_data$ AS e JOIN Pharma\_data$ AS p ON E.[Emp Name] = p.[Name of Sales Rep];

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

**ANS:**

SELECT [Name of Sales Rep],SUM(Sales) FROM Pharma\_data$ WHERE City='Rendsburg' AND Year=2018 GROUP BY [Name of Sales Rep];

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

**ANS:**

SELECT [Product Class],Month,Year, SUM(Sales)As total\_sales\_for\_each\_productclass FROM Pharma\_data$ GROUP BY

[Product Class],Month,Year ORDER BY Year,Month,[Product Class];

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

**ANS:**

SELECT TOP 3 Sales FROM Pharma\_data$ Where Year=2019**;**

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

**ANS:**

WITH MONTHLYSALES AS(

SELECT [Sub-channel],Month,Year,SUM(sales) AS TOTAl\_MONTHLY\_SALES FROM Pharma\_data$ GROUP BY [Sub-channel],Month,Year

)

SELECT [Sub-channel],AVG(TOTAl\_MONTHLY\_SALES) AS Sales\_average FROM MONTHLYSALES GROUP BY [Sub-channel];

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

**ANS:**

SELECT [Product Class],Sum(sales),AVG(Price),Sum(Quantity)

FROM Pharma\_data$ GROUP BY [Product Class];

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

**ANS:**

WITH top5customers AS(

SELECT

[Customer NAME],year AS saleyear,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 TOP 5 [Customer Name],saleyear,Totalsales FROM top5customers

where SalesRank <=5 ORDER BY saleyear,Totalsales DESC;

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

**ANS:**

SELECT AVG(Sales) AS SalesAVG,Year,Country FROM Pharma\_data$ GROUP BY Year,Country ORDER BY Year;

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

**ANS:**

WITH lowersales AS(

SELECT year AS SaleYear,month AS Sale\_month,SUM(sales) AS Totalsales,

RANK()OVER(PARTITION BY year ORDER BY SUM(sales) ASC)AS salesrank FROM Pharma\_data$ GROUP By Year,Month

)

SELECT Totalsales,Saleyear AS year,Sale\_month AS month FROM lowersales ORDER BY salesrank ;

**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.**

**ANS:**

With Totolsalesincountry AS(

SELECT [Sub-channel] ,SUM(Sales) AS Totalsales,Country FROM Pharma\_data$ GROUP BY [Sub-channel],Country),RankedCountries AS(

SELECT country,[Sub-channel],Totalsales,ROW\_NUMBER() OVER(PARTITION BY [Sub-Channel] ORDER BY Totalsales DESC)AS CountryRank FROM Totolsalesincountry

)

SELECT country,[Sub-channel],Totalsales FROM RankedCountries;