**# Set Up:**

**CREATE DATABASE SUPERSTORE;**

**USE SUPERSTORE;**

**show tables;**

**drop database superstore;**

**drop table s1;**

**SET GLOBAL local\_infile = 1;**

**CREATE TABLE s1 (**

**Row\_ID INT,**

**Order\_ID VARCHAR(20),**

**Order\_Date DATE,**

**Order\_Month TINYINT,**

**Order\_Month\_Name VARCHAR(15),**

**Order\_Year SMALLINT,**

**Ship\_Date DATE,**

**Ship\_Month TINYINT,**

**Ship\_Month\_Name VARCHAR(15),**

**Ship\_Year SMALLINT,**

**Ship\_Mode VARCHAR(50),**

**Customer\_ID VARCHAR(20),**

**Customer\_Name VARCHAR(100),**

**Segment VARCHAR(50),**

**Country VARCHAR(50),**

**City VARCHAR(50),**

**State VARCHAR(50),**

**Postal\_Code VARCHAR(10),**

**Region VARCHAR(50),**

**Product\_ID VARCHAR(20),**

**Category VARCHAR(50),**

**Sub\_Category VARCHAR(50),**

**Product\_Name VARCHAR(200),**

**Sales DECIMAL(10,2),**

**Quantity INT,**

**Discount DECIMAL(4,2),**

**Profit DECIMAL(10,2)**

**);**

**LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/superstore.csv'**

**INTO TABLE s1**

**FIELDS TERMINATED BY ','**

**ENCLOSED BY '"'**

**LINES TERMINATED BY '\n'**

**IGNORE 1 LINES;**

**SHOW VARIABLES LIKE 'local\_infile';**

**SHOW VARIABLES LIKE 'secure\_file\_priv';**

**#MAIN ANALYSIS :**

**1. Total Sales and Profit by Year**

**SELECT**

**Order\_Year,**

**SUM(Sales) AS TOTAL\_SALES,**

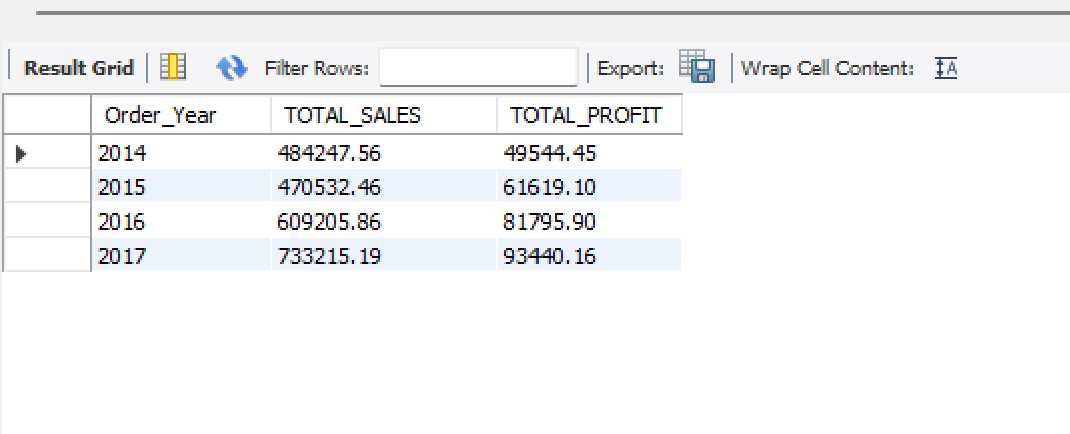
**SUM(Profit) AS TOTAL\_PROFIT**

**FROM**

**s1**

**GROUP BY Order\_Year**

**ORDER BY Order\_Year;**



**2. Monthly Sales Trend**

**SELECT**

**Order\_Month,**

**SUM(Sales) AS TOTAL\_SALES,**

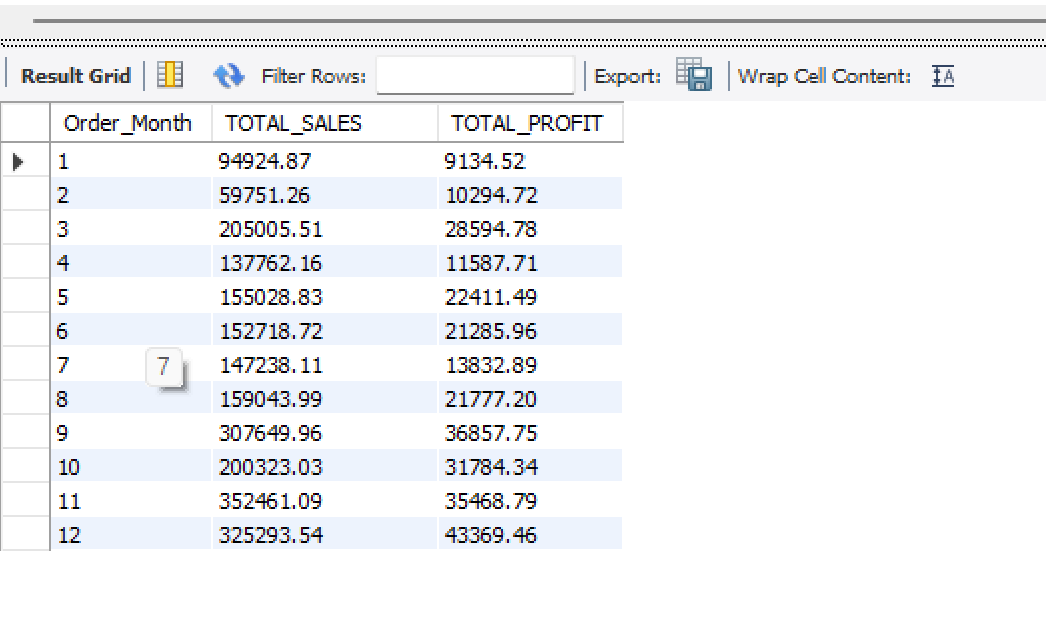
**SUM(Profit) AS TOTAL\_PROFIT**

**FROM**

**s1**

**GROUP BY Order\_Month**

**ORDER BY Order\_Month;**



**3.Profit Margin by Category**

**SELECT**

**Category,**

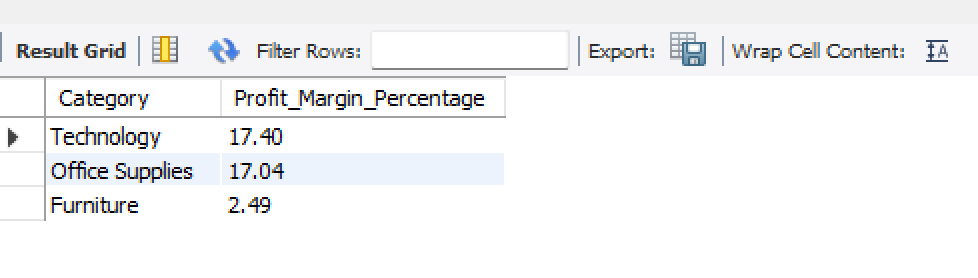
**ROUND(SUM(Profit) / SUM(Sales) \* 100, 2) AS Profit\_Margin\_Percentage**

**FROM**

**s1**

**GROUP BY Category**

**ORDER BY Profit\_Margin\_Percentage DESC;**



**4.Top 10 Products by Sales**

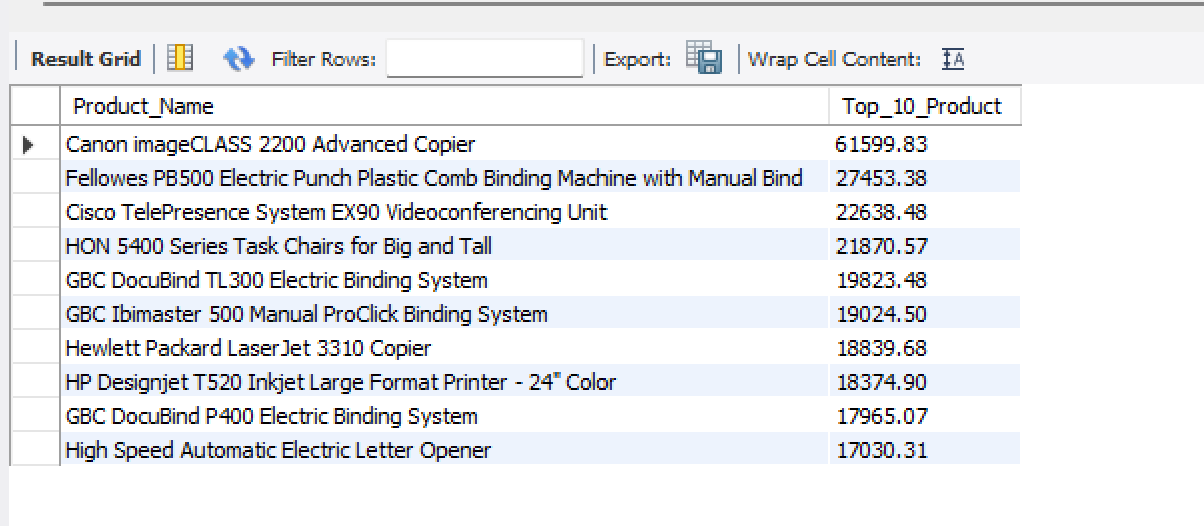
**Select Product\_Name,Sum(Sales) as Top\_10\_Product**

**from s1**

**group by Product\_Name**

**order by Top\_10\_Product Desc**

**limit 10;**



**5. High Sales, Low Profit Products**

**SELECT**

**Product\_Name,**

**SUM(Sales) AS Total\_Sales,**

**SUM(Profit) AS Total\_Profit**

**FROM**

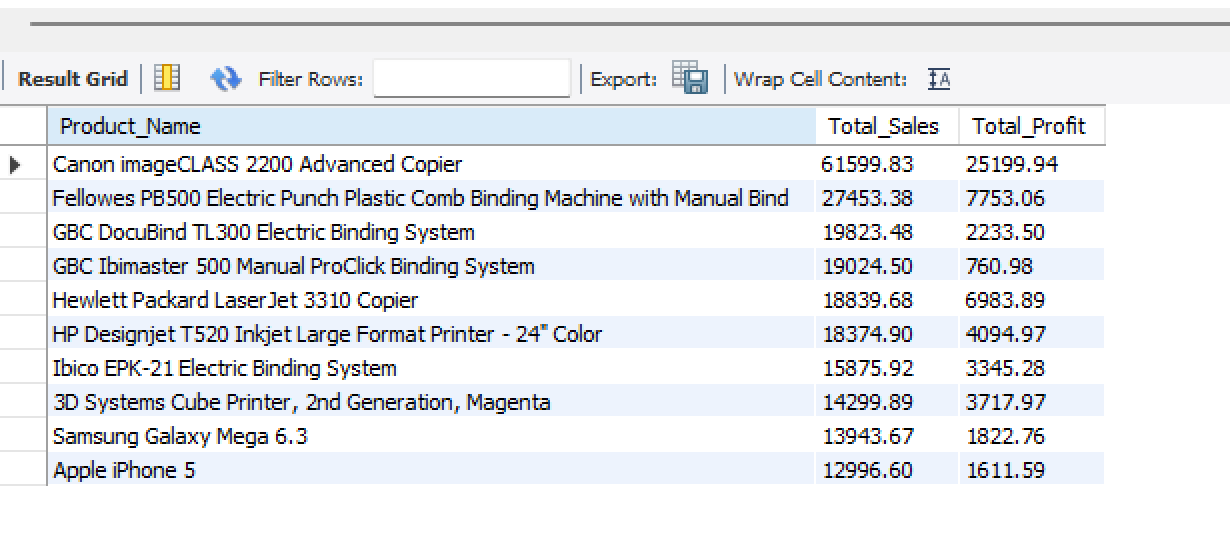
**s1**

**GROUP BY Product\_Name**

**HAVING SUM(Sales) > 5000 AND SUM(Profit) > 10**

**ORDER BY Total\_Sales DESC**

**LIMIT 10;**



**Geographical Insights:**

**6.Sales & Profit by Region and State**

**SELECT**

**Region,**

**State,**

**SUM(Sales) AS Total\_sales,**

**SUM(profit) AS Total\_Profit**

**FROM**

**s1**

**GROUP BY Region , State**

**ORDER BY Region , Total\_sales DESC;**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
| **7.Cities with Negative Profit**  **select City, Profit from s1 where Profit<0 order by City;** |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **Product & Category Analysis**  **8. Sales by Sub-Category**  **SELECT**  **Sub\_Category, SUM(Sales) AS Total\_Sales**  **FROM**  **s1**  **GROUP BY Sub\_Category**  **ORDER BY Total\_Sales DESC;** |  |  |  |
| **Shipping & Fulfillment:**  **9. Average Order-to-Ship Days**  **SELECT**  **ROUND(AVG(datediff(Ship\_Date , Order\_Date)),2) AS Average\_Order\_To\_Ship\_Days**  **FROM**  **s1;**    **10.Ship Mode Performance**  **SELECT**  **Ship\_Mode, COUNT(Row\_ID) AS COUNT\_Of\_Ship\_Mode**  **FROM**  **s1**  **GROUP BY Ship\_Mode**  **ORDER BY Count\_Of\_Ship\_Mode DESC;** |  |  |  |
|  |  |  |  |
| **Customer Behavior:**  **11. Top 10 Customers by Profit**  **SELECT**  **Customer\_Name,**  **COUNT(Customer\_ID) AS Count\_Of\_Customer,**  **SUM(Profit) AS Profitted\_Amount**  **FROM**  **s1**  **GROUP BY Customer\_Name**  **ORDER BY Profitted\_Amount DESC , Count\_Of\_Customer**  **LIMIT 10;** |  |  |  |
|  |  |  |  |
| **12.Segment-wise Profitability**  **SELECT**  **Segment,**  **Sum(Sales) AS Total\_Sales,**  **SUM(Profit) AS Total\_Profit**  **FROM**  **s1**  **GROUP BY Segment**  **ORDER BY Total\_Profit DESC;**    **Discount and Loss Analysis:**  **13.Profit vs. Discount**  **SELECT**  **Product\_Name,Discount,**  **SUM(Sales) AS total\_sales,**  **SUM(Profit) AS total\_profit,**  **ROUND(SUM(Profit) / SUM(Sales)\*100, 2) AS Best\_Profit\_Percentage**  **FROM**  **s1**  **GROUP BY Product\_Name,Discount**  **HAVING Best\_Profit\_Percentage > 0**  **ORDER BY Product\_Name,Best\_Profit\_Percentage DESC;** |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **14.Orders with Losses**  **SELECT**  **Product\_Name,Discount,**  **Sales,**  **Profit**  **FROM**  **s1**  **where profit<0**  **Order by Profit Asc;** |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Time & Trends**

**15.Sales by Weekday**

**SELECT**

**Weekday,**

**ROUND(SUM(Sales), 2) AS Total\_Sales**

**FROM**

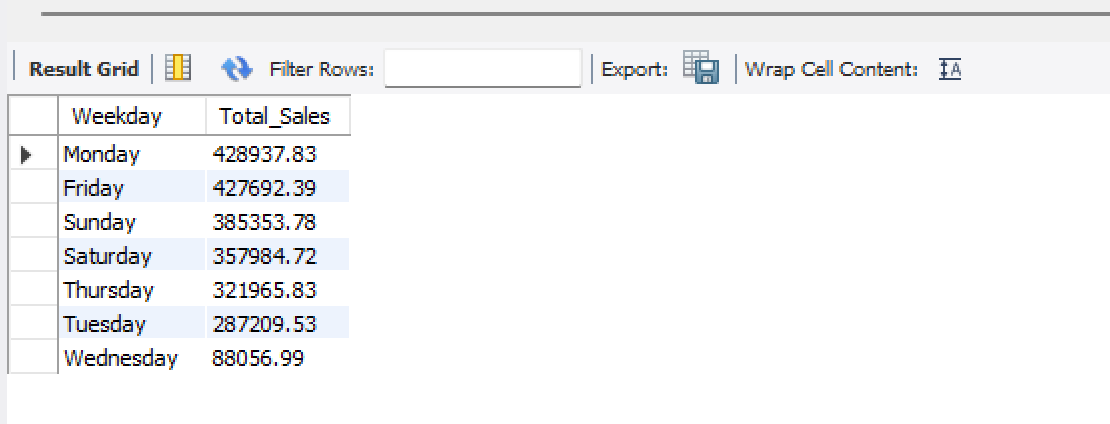
**s1**

**GROUP BY**

**Weekday**

**ORDER BY Total\_Sales Desc,**

**FIELD(Weekday, 'Monday','Tuesday','Wednesday','Thursday','Friday','Saturday','Sunday');**



16. Which discount ranges give us the best profit?

SELECT

CASE

WHEN Discount = 0 THEN '0%'

WHEN Discount > 0 AND Discount <= 0.1 THEN '0-10%'

WHEN Discount > 0.1 AND Discount <= 0.2 THEN '11-20%'

WHEN Discount > 0.2 AND Discount <= 0.3 THEN '21-30%'

WHEN Discount > 0.3 AND Discount <= 0.4 THEN '31-40%'

WHEN Discount > 0.4 AND Discount <= 0.5 THEN '41-50%'

ELSE '>50%'

END AS Discount\_Range,

COUNT(Order\_ID) AS Total\_Orders,

SUM(Sales) AS Total\_Sales,

SUM(Profit) AS Total\_Profit,

ROUND(SUM(Profit)/SUM(Sales)\*100, 2) AS Profit\_Margin\_Percentage,

ROW\_NUMBER() OVER (

ORDER BY ROUND(SUM(Profit)/SUM(Sales)\*100, 2) DESC

) AS Rank\_Number

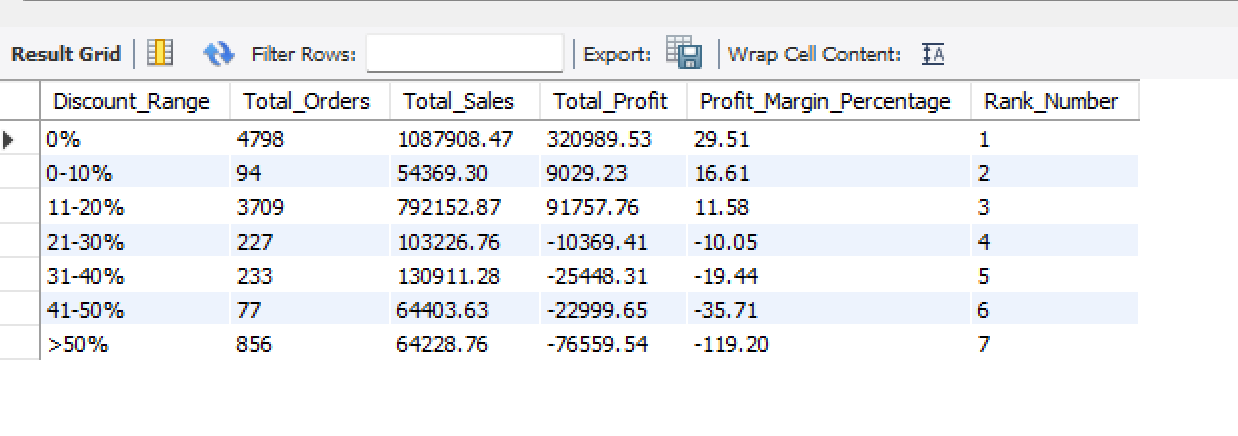
FROM

s1

GROUP BY

Discount\_Range

ORDER BY

Rank\_Number;

**17. Is there a point where giving more discount leads to a loss?**

**SELECT**

**Discount,**

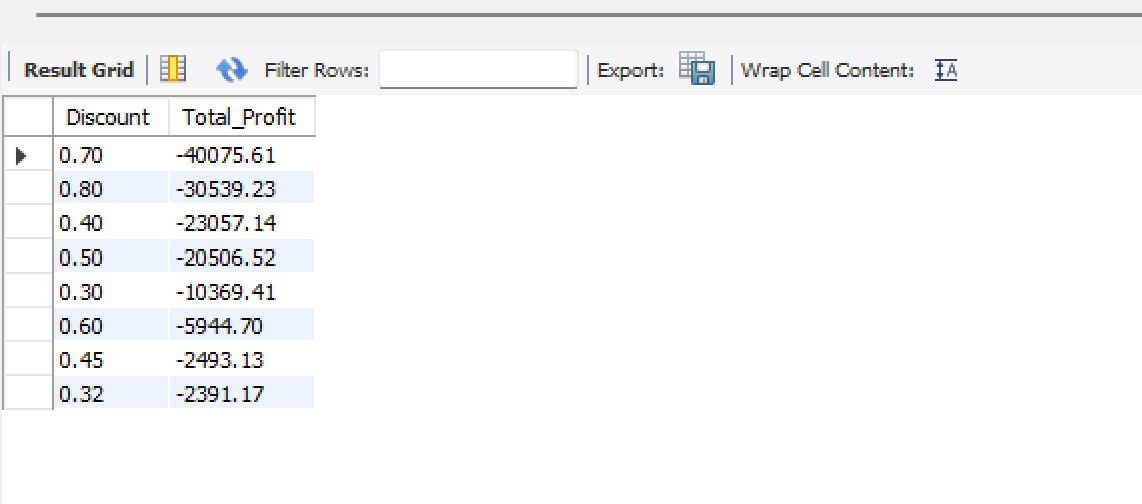
**ROUND(SUM(Profit), 2) AS Total\_Profit**

**FROM s1**

**GROUP BY Discount**

**HAVING Total\_Profit < 0**

**ORDER BY Total\_Profit;**



**18. What’s the optimal discount level?**

**SELECT**

**Discount,**

**ROUND(SUM(Sales), 2) AS Total\_Sales,**

**ROUND(SUM(Profit), 2) AS Total\_Profit,**

**ROUND(SUM(Profit)/SUM(Sales) \* 100, 2) AS Profit\_Margin\_Percent**

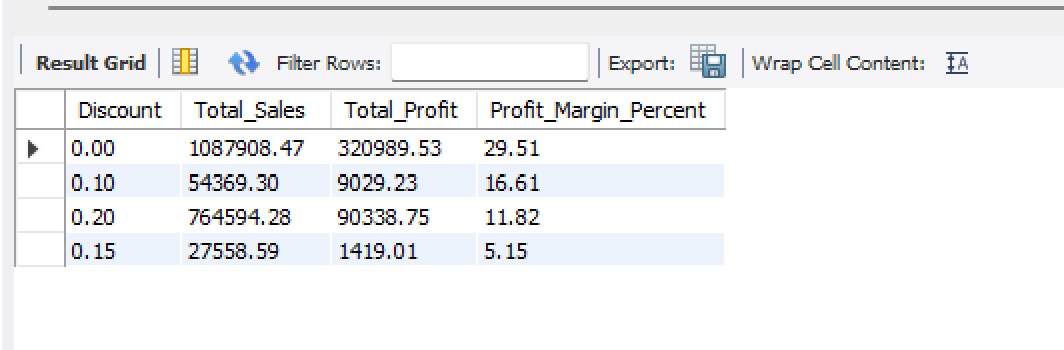
**FROM s1**

**GROUP BY Discount**

**having Profit\_Margin\_Percent>0**

**ORDER BY Profit\_Margin\_Percent DESC**

**LIMIT 5;**



**19.How many orders used each discount level?**

**SELECT**

**Discount,**

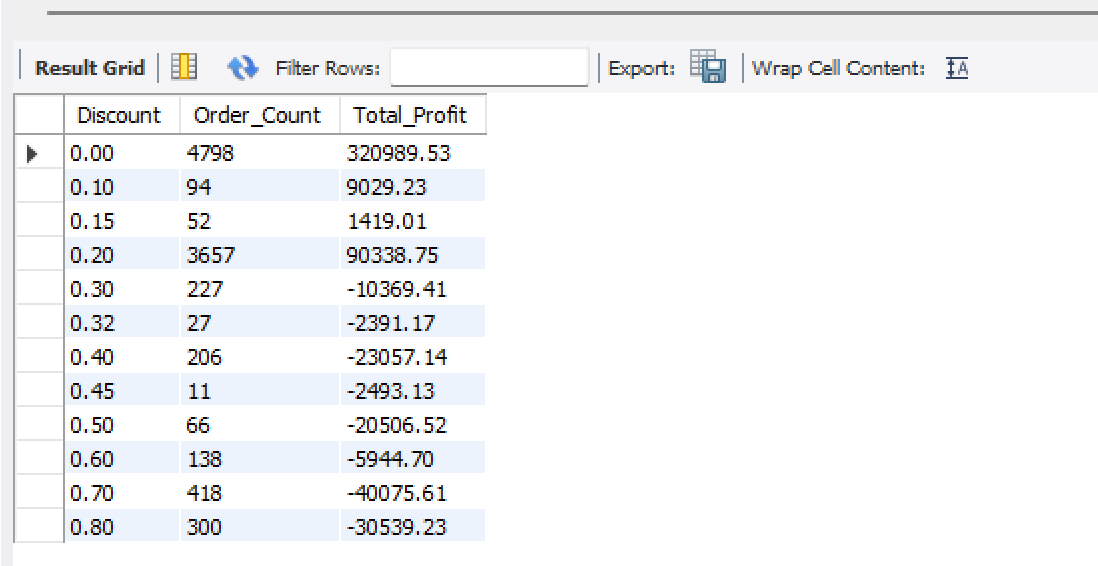
**COUNT(\*) AS Order\_Count,**

**Sum(Profit) as Total\_Profit**

**FROM s1**

**GROUP BY Discount**

**ORDER BY Discount;**



**THANK YOU**