Superstore Data Analysis Using MySQL

FINAL PROJECT BY RISHABH SHARMA

Introduction

- **Dataset: Sample Superstore (Sales, Segments, Products, Regions etc)**
- **Tool: MySQL Workbench / MySQL Command Line**
- **Objective: Extract meaningful insights using SQL queries**

Project Objective

- **Perform data analysis on Superstore dataset**
- **Understand sales, profit, shipping, and customer patterns**
- **Use SQL queries to find business insights**

Methodology

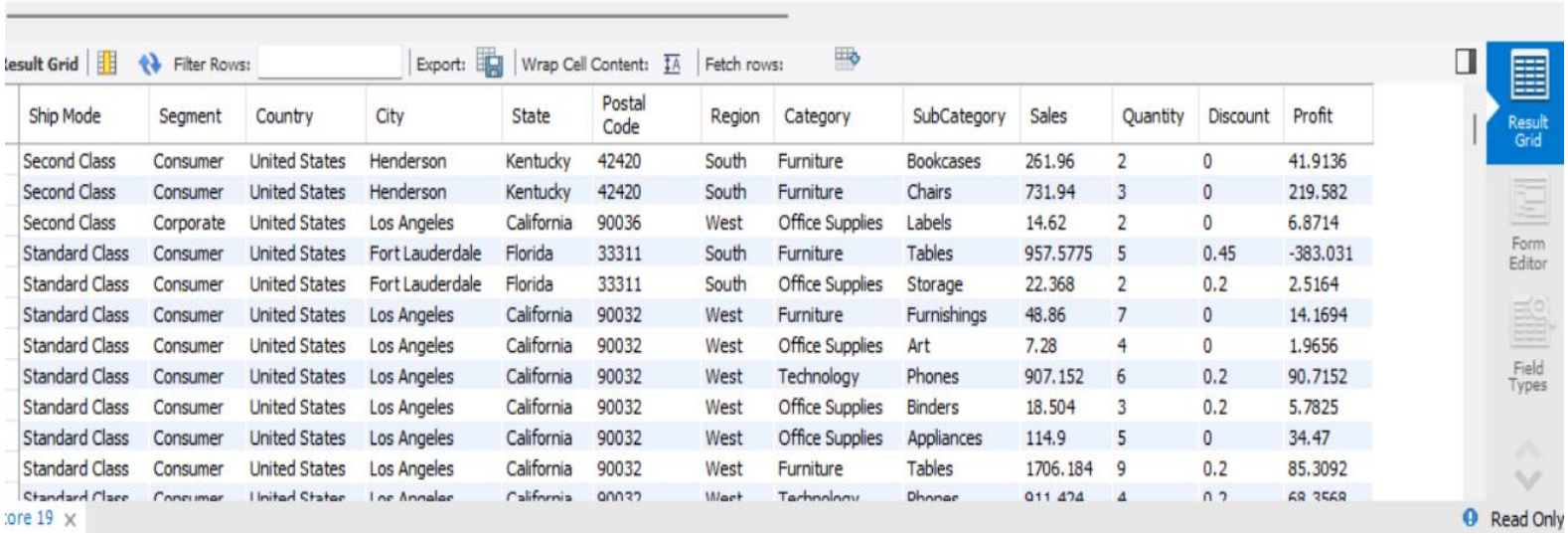
- **Step 1: Import CSV into MySQL**
- **Step 2: Create database & table schema**
- **Step 3: Write 25 SQL queries (aggregate, window, analytical)**
- **Step 4: Analyze results for insights**

Queries Implemented

- **Profitability analysis (products, Sales, regions)**
- **Customer lifetime value & top customers**
- **Shipping delays & order performance**
- **High-discount loss orders**
- **Regional and segment contribution**

Super Store Table

```
1 •  create database SuperstoreDB;
2 •  use SuperstoreDB;
3 •  select * from store;
4
5
6
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Ship Mode	Segment	Country	City	State	Postal Code	Region	Category	SubCategory	Sales	Quantity	Discount	Profit
Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Bookcases	261.96	2	0	41.9136
Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Chairs	731.94	3	0	219.582
Second Class	Corporate	United States	Los Angeles	California	90036	West	Office Supplies	Labels	14.62	2	0	6.8714
Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Furniture	Tables	957.5775	5	0.45	-383.031
Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Office Supplies	Storage	22.368	2	0.2	2.5164
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Furniture	Furnishings	48.86	7	0	14.1694
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Art	7.28	4	0	1.9656
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Technology	Phones	907.152	6	0.2	90.7152
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Binders	18.504	3	0.2	5.7825
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Office Supplies	Appliances	114.9	5	0	34.47
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Furniture	Tables	1706.184	9	0.2	85.3092
Standard Class	Consumer	United States	Los Angeles	California	90032	West	Technology	Phones	911.424	4	0.2	68.3568

Result Grid | Form Editor | Field Types | Read Only

1. Top 10 most profitable products

```
4
5   /*1. Top 10 most profitable products*/
6   SELECT SubCategory Product_Name, SUM(Profit) AS Total_Profit
7   FROM store
8   GROUP BY SubCategory
9   ORDER BY Total_Profit DESC
10  LIMIT 10;
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Product_Name	Total_Profit
Copiers	55617.82490000001
Phones	44515.7306
Accessories	41936.63569999993
Paper	34053.56929999997
Binders	30221.763299999995
Chairs	26590.166300000026
Storage	21278.826399999998
Appliances	18138.005399999995
Furnishings	13059.143599999983
Envelopes	6964.176700000003

Result Grid
Form Editor
Field Types

2. Products causing the highest losses

```
11
12    /*2. Products causing the highest losses*/
13 •  SELECT SubCategory As Product_Name, SUM(Profit) AS Total_Profit
14    FROM store
15    GROUP BY SubCategory
16    HAVING SUM(Profit) < 0
17    ORDER BY Total_Profit ASC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Product_Name	Total_Profit
Tables	-17725.481100000008
Bookcases	-3472.5559999999978
Supplies	-1189.0994999999984

3. Total Sales And Profit

```
/* Total Sales And Profit*/
SELECT SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
FROM store;
```

sult Grid | Filter Rows: | Export: | Wrap Cell Content:

Total_Sales	Total_Profit
2297200.860299955	286397.0217000013

4. region wise Sales

```
32      /*3. region wise Sales*/
33 •    SELECT Region, SUM(Sales) AS Total_Sales
34      FROM store
35      GROUP BY Region
36      ORDER BY Total_Sales DESC;
```

Result Grid | Filter Rows: | Export:

Region	Total_Sales
West	725457.8245000006
East	678781.2399999979
Central	501239.8908000005
South	391721.9050000003

5. Region Wise Total Profit

```
42      /* Region Wise Total Profit */
43 •    SELECT Region, SUM(Profit) AS Total_Profit
44      FROM store
45      GROUP BY Region
46      ORDER BY Total_Profit DESC;
```

Result Grid | Filter Rows: | Export: | W

Region	Total_Profit
West	108418.44890000018
East	91522.78000000026
South	46749.430300000065
Central	39706.36249999998

6. Sales and profit by category

```
48      /*Sales and profit by category*/
49      SELECT Category, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
50      FROM store
51      GROUP BY Category;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Category	Total_Sales	Total_Profit
Furniture	741999.7952999998	18451.2728
Office Supplies	719047.0320000029	122490.80080000011
Technology	836154.0329999966	145454.9480999999

7. Sales and profit by sub-category

```
52      /*Sales and profit by sub-category*/
53 •    SELECT SubCategory, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
54      FROM store
55      GROUP BY SubCategory
56      ORDER BY Total_Sales DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

SubCategory	Total_Sales	Total_Profit
Phones	330007.0540000001	44515.7306
Chairs	328449.10300000076	26590.166300000026
Storage	223843.60800000012	21278.826399999998
Tables	206965.5320000001	-17725.481100000008
Binders	203412.7330000001	30221.763299999995
Machines	189238.63099999996	3384.7569
Accessories	167380.3180000001	41936.63569999993
Copiers	149528.02999999994	55617.82490000001
Bookcases	114879.99629999997	-3472.5559999999978
Appliances	107532.161	18138.005399999995
Furnishings	91705.16400000005	13059.143599999983
Paper	78479.20600000002	34053.56929999997
Supplies	46673.538000000015	-1189.0994999999984
Art	27118.791999999954	6527.786999999998
Envelopes	16476.402	6964.176700000003

8. Top 10 states by sales

```
59      /*Top 10 states by sales*/
60 •    SELECT State, SUM(Sales) AS Total_Sales
61      FROM store
62      GROUP BY State
63      ORDER BY Total_Sales DESC
64      LIMIT 10;
```

Result Grid | Filter Rows: | Export:

State	Total_Sales
California	457687.631500001
New York	310876.2709999998
Texas	170188.04580000002
Washington	138641.26999999993
Pennsylvania	116511.91400000003
Florida	89473.708
Illinois	80166.10099999986
Ohio	78258.13599999993
Michigan	76269.61400000002
Virginia	70636.71999999999

9. Bottom 10 states by profit

```
66      /*Bottom 10 states by profit */
67 •    SELECT State, SUM(Profit) AS Total_Profit
68      FROM store
69      GROUP BY State
70      ORDER BY Total_Profit ASC
71      LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wr

State	Total_Profit
Texas	-25729.3563
Ohio	-16971.376600000018
Pennsylvania	-15559.960300000013
Illinois	-12607.88699999998
North Carolina	-7490.912200000003
Colorado	-6527.857900000001
Tennessee	-5341.6936
Arizona	-3427.9246
Florida	-3399.3017
Oregon	-1190.4704999999992

10. Average discount per category

```
73    /* Average discount per category*/
74    SELECT Category, AVG(Discount) AS Avg_Discount
75    FROM store
76    GROUP BY Category;
77
```

Result Grid | Filter Rows: | Export: | Wrap Cel

Category	Avg_Discount
Furniture	0.17392267798208297
Office Supplies	0.15728509790906708
Technology	0.1323226854358401

11. Average discount per sub-category

```
78      /*Average discount per sub-category*/
79 •    SELECT SubCategory, AVG(Discount) AS Avg_Discount
80      FROM store
81      GROUP BY SubCategory
82      ORDER BY Avg_Discount DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell C

SubCategory	Avg_Discount
Binders	0.3722915298752447
Machines	0.30608695652173884
Tables	0.26128526645768047
Bookcases	0.211140350877193
Chairs	0.17017828200972449
Appliances	0.16652360515021464
Copiers	0.16176470588235295
Phones	0.15455568053993343
Furnishings	0.13834900731452457
Fasteners	0.08202764976958511
Envelopes	0.0803149606299211
Accessories	0.07845161290322622
Supplies	0.0768421052631578
Paper	0.07489051094890578
Art	0.07487437185929686

12. Sales contribution by region (%)

```
84      /*Sales contribution by region (%)*/
85 •    SELECT Region,
86             SUM(Sales) AS Total_Sales,
87             (SUM(Sales) * 100.0 / (SELECT SUM(Sales) FROM store)) AS Sales_Percentage
88      FROM store
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Region	Total_Sales	Sales_Percentage
South	391721.9050000003	17.052139922533875
West	725457.8245000006	31.58007804355752
Central	501239.8908000005	21.819593552413675
East	678781.2399999979	29.54818848149685

13. Profit margin by category

```
91      /*Profit margin by category*/
92      SELECT Category,
93             SUM(Profit) AS Total_Profit,
94             SUM(Sales) AS Total_Sales,
95             (SUM(Profit)/SUM(Sales))*100 AS Profit_Margin_Percentage
96      FROM store
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Category	Total_Profit	Total_Sales	Profit_Margin_Percentage
Furniture	18451.2728	741999.7952999998	2.486695133458887
Office Supplies	122490.80080000011	719047.0320000029	17.0351583900286
Technology	145454.9480999999	836154.0329999966	17.395712076891996

14. State with maximum loss

```
 99    /*State with maximum loss*/
100 •  SELECT State, SUM(Profit) AS Total_Profit
101    FROM store
102    GROUP BY State
103    ORDER BY Total_Profit ASC
104    LIMIT 1;
105
```

Result Grid | Filter Rows: | Export:

State	Total_Profit
Texas	-25729.3563

15. Segment-wise sales

```
106     /*Segment-wise sales*/
107 •   SELECT Segment, SUM(Sales) AS Total_Sales
108     FROM store
109     GROUP BY Segment;
```

Result Grid | Filter Rows: | Export:

Segment	Total_Sales
Consumer	1161401.3449999888
Corporate	706146.3668000001
Home Office	429653.1485000003

16. Segment-wise profit

```
111     /*Segment-wise profit*/
112 •   SELECT Segment, SUM(Profit) AS Total_Profit
113     FROM store
114     GROUP BY Segment;
```

Result Grid | Filter Rows: | Export:

Segment	Total_Profit
Consumer	134119.20919999972
Corporate	91979.13400000021
Home Office	60298.678500000075

17. Orders where discount > 0.3 OR profit < 0

```
116    /*Orders where discount > 0.3 OR profit < 0*/
117 •  SELECT *
118    FROM store
119    WHERE Discount > 0.64 OR Profit < 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Ship Mode	Segment	Country	City	State	Postal Code	Region	Category	SubCategory	Sales	Quantity	Discount	Profit
Standard Class	Consumer	United States	Philadelphia	Pennsylvania	19134	East	Office Supplies	Storage	36.336	3	0.2	-7.2672
Standard Class	Consumer	United States	Philadelphia	Pennsylvania	19134	East	Office Supplies	Supplies	666.248	1	0.2	-149.9058
Second Class	Consumer	United States	Chicago	Illinois	60653	Central	Office Supplies	Binders	42.616	7	0.8	-68.1856
Standard Class	Consumer	United States	Springfield	Ohio	45503	East	Office Supplies	Storage	646.776	9	0.2	-145.5246
Standard Class	Consumer	United States	Fort Worth	Texas	76106	Central	Furniture	Furnishings	66.112	4	0.6	-84.2928
Second Class	Corporate	United States	Chicago	Illinois	60653	Central	Office Supplies	Binders	11.364	3	0.8	-17.046
Second Class	Corporate	United States	Chicago	Illinois	60653	Central	Office Supplies	Supplies	8.72	5	0.2	-1.744
Standard Class	Consumer	United States	Los Angeles	California	90045	West	Furniture	Tables	1335.68	4	0.2	-217.048
Standard Class	Consumer	United States	Chicago	Illinois	60610	Central	Furniture	Furnishings	8.792	1	0.6	-5.7148
Standard Class	Consumer	United States	Denver	Colorado	80219	West	Office Supplies	Binders	6.783	7	0.7	-4.7481
Standard Class	Consumer	United States	Salem	Oregon	97301	West	Technology	Phones	84.784	2	0.2	-20.1362
Standard Class	Consumer	United States	Salem	Oregon	97301	West	Office Supplies	Binders	16.821	3	0.7	-12.8961
Standard Class	Consumer	United States	Philadelphia	Pennsylvania	19120	East	Technology	Phones	143.982	3	0.4	-28.7964
Standard Class	Consumer	United States	Philadelphia	Pennsylvania	19120	East	Technology	Phones	494.376	4	0.4	-115.3544

18. Cities with sales above 50,000 AND profit margin below 5%

```
4   /*Cities with sales above 50,000 AND profit margin below 5%*/
5   SELECT City, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit,
6          (SUM(Profit)/SUM(Sales))*100 AS Profit_Margin
7   FROM store
8   GROUP BY City
9   HAVING SUM(Sales) > 50000 AND (SUM(Profit)/SUM(Sales))*100 < 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

City	Total_Sales	Total_Profit	Profit_Margin
Philadelphia	109077.01300000008	-13837.767400000012	-12.686236099992948
Houston	64504.76039999994	-10153.548499999997	-15.740773916586793

19. Most profitable category in each segment

```
121     /*Most profitable category in each segment*/
122 •   SELECT Segment, Category, SUM(Profit) AS Total_Profit
123     FROM store
124     GROUP BY Segment, Category
125     ORDER BY Segment, Total_Profit DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Conter

Segment	Category	Total_Profit
Consumer	Technology	70797.80960000002
Consumer	Office Supplies	56330.321000000076
Consumer	Furniture	6991.078599999996
Corporate	Technology	44166.99799999998
Corporate	Office Supplies	40227.32019999998
Corporate	Furniture	7584.8158
Home Office	Technology	30490.14049999998
Home Office	Office Supplies	25933.15959999994
Home Office	Furniture	3875.3783999999982

20. Highest discount given in any state

```
127     /*Highest discount given in any state*/
128 •   SELECT State, MAX(Discount) AS Max_Discount
129     FROM store
130     GROUP BY State
131     ORDER BY Max_Discount DESC
132     LIMIT 1;
```

Result Grid | Filter Rows: | Export:

State	Max_Discount
Texas	0.8

21. List states whose profit is higher than the average state profit

```sql
/*List states whose profit is higher than the average state profit*/
SELECT State, SUM(Profit) AS Total_Profit
FROM store
GROUP BY State
HAVING SUM(Profit) > (
    SELECT AVG(TotalProfit)
    FROM (
        SELECT State, SUM(Profit) AS TotalProfit
        FROM store
        GROUP BY State
    ) AS Sub
);
```

Result Grid | Filter Rows:

State	Total_Profit
Kentucky	11199.696600000005
California	76381.38710000017
Washington	33402.651699999995
Wisconsin	8401.800399999998
Minnesota	10823.1874
Michigan	24463.187599999994
Delaware	9977.374800000001
Indiana	18382.936300000005
New York	74038.54860000005
Virginia	18597.9504
Missouri	6436.210499999999
New Jersey	9772.9138
Massachus...	6785.501600000005
Georgia	16250.043300000003
Rhode Island	7285.629300000001

22. Sub-categories where average quantity ordered > average quantity across dataset

```
25    /*Sub-categories where average quantity ordered > average quantity across dataset*/
26    SELECT SubCategory, AVG(Quantity) AS Avg_Quantity
27    FROM store
28    GROUP BY SubCategory
29    HAVING AVG(Quantity) > (
30        SELECT AVG(Quantity) FROM store
31    );
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

SubCategory	Avg_Quantity
Bookcases	3.8070
Chairs	3.8185
Labels	3.8462
Tables	3.8903
Binders	3.9225
Accessories	3.8400
Fasteners	4.2120
Machines	3.8261

23. States where total discount given is above the average total discount across all states

```sql
/*States where total discount given is above the average total discount across all states*/
SELECT State, SUM(Discount) AS Total_Discount
FROM store
GROUP BY State
HAVING SUM(Discount) > (
    SELECT AVG(StateDiscount)
    FROM (
        SELECT State, SUM(Discount) AS StateDiscount
        FROM store
        GROUP BY State
    ) AS Sub
);
```

Result Grid | Filter Rows:

State	Total_Discount
California	145.6000000000004
Florida	114.65000000000074
North Carolina	70.60000000000029
Washington	32.39999999999992
Texas	364.6399999999986
Pennsylvania	192.89999999999955
Illinois	191.89999999999947
New York	62.40000000000031
Arizona	68.00000000000026
Tennessee	53.300000000000125
Oregon	35.79999999999998
Colorado	57.600000000000186
Ohio	152.4000000000001

24. Find postal codes where the total quantity sold is above the average quantity of all postal codes

```sql
SELECT PostalCode, SUM(Quantity) AS Total_Quantity
FROM store
GROUP BY PostalCode
HAVING SUM(Quantity) > (
    SELECT AVG(PostQty)
    FROM (
        SELECT PostalCode, SUM(Quantity) AS PostQty
        FROM store
        GROUP BY PostalCode
    ) AS Sub
);
```

Result Grid | Filter Rows:

PostalCode	Total_Quantity
42420	185
90036	468
33311	65
90032	438
98103	539
76106	102
94109	515
19140	564
90049	580
77095	381
77041	443
10024	840
90004	460
60610	389
22153	198

25.Analyzing Shipping Modes with Above-Average Sales using Nested Subquery

```sql
/*Analyzing Shipping Modes with Above-Average Sales using Nested Subquery*/
SELECT ShipMode, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
FROM store
GROUP BY ShipMode
HAVING SUM(Sales) > (
    SELECT AVG(ModeSales)
    FROM (
        SELECT ShipMode, SUM(Sales) AS ModeSales
        FROM store
        GROUP BY ShipMode
    ) AS Sub
)
ORDER BY Total_Sales DESC;
```

Result Grid | Filter Rows: | Export:

ShipMode	Total_Sales	Total_Profit
Standard Class	1358215.7429999826	164088.7874999997

Key Insights

- **Some products generate losses despite high sales (discount impact)**
- **Technology category drives most profit**
- **Certain states have poor profit margins**
- **Corporate segment contributes maximum revenue**
- **Faster shipping = higher customer satisfaction**

Conclusion

- **SQL is powerful for analyzing real-world business datasets**
- **Superstore data reveals key insights on profitability, sales & customers**
- **Data-driven decision making improves business growth**

THANK YOU
SO MUCH