

SUPERSTORE & NORTHWIND DATA ANALYSIS PROJECT (SQL + EXCEL)

PRESENTED BY MARYAM RIAZ



Introduction

This project portfolio demonstrates SQL skills from beginner to advanced level using the Superstore dataset and Northwind database. The objective is to extract business insights, create visualizations in Excel, and present findings in a structured way.

Dataset 1: Superstore Sales Analysis

- Source: Kaggle Superstore Dataset
- ~10,000 sales records (Orders, Customers, Products, Regions).
- Columns: Order ID, Order Date, Ship Date, Customer, Segment, Region, Category, Sub-Category, Sales, Profit, Discount, Quantity.



SQL Queries

QUERY 1

What are the total sales and profit by region?

```
select region , sum(sales) as 'Total Sales' ,  
sum(profit) as 'Total Profit'  
from superstore  
group by region;
```

Insight:

West region generates the highest sales and profit, while the South lags in both. This suggests stronger customer demand in the West.

Row Labels	Sum of Total Sales	Sum of Total Profit
Central	497800.8728	40128.903
East	672194.054	90672.0127
South	388983.585	46035.689
West	713471.3445	106021.1495
Grand Total	2272449.856	282857.7542

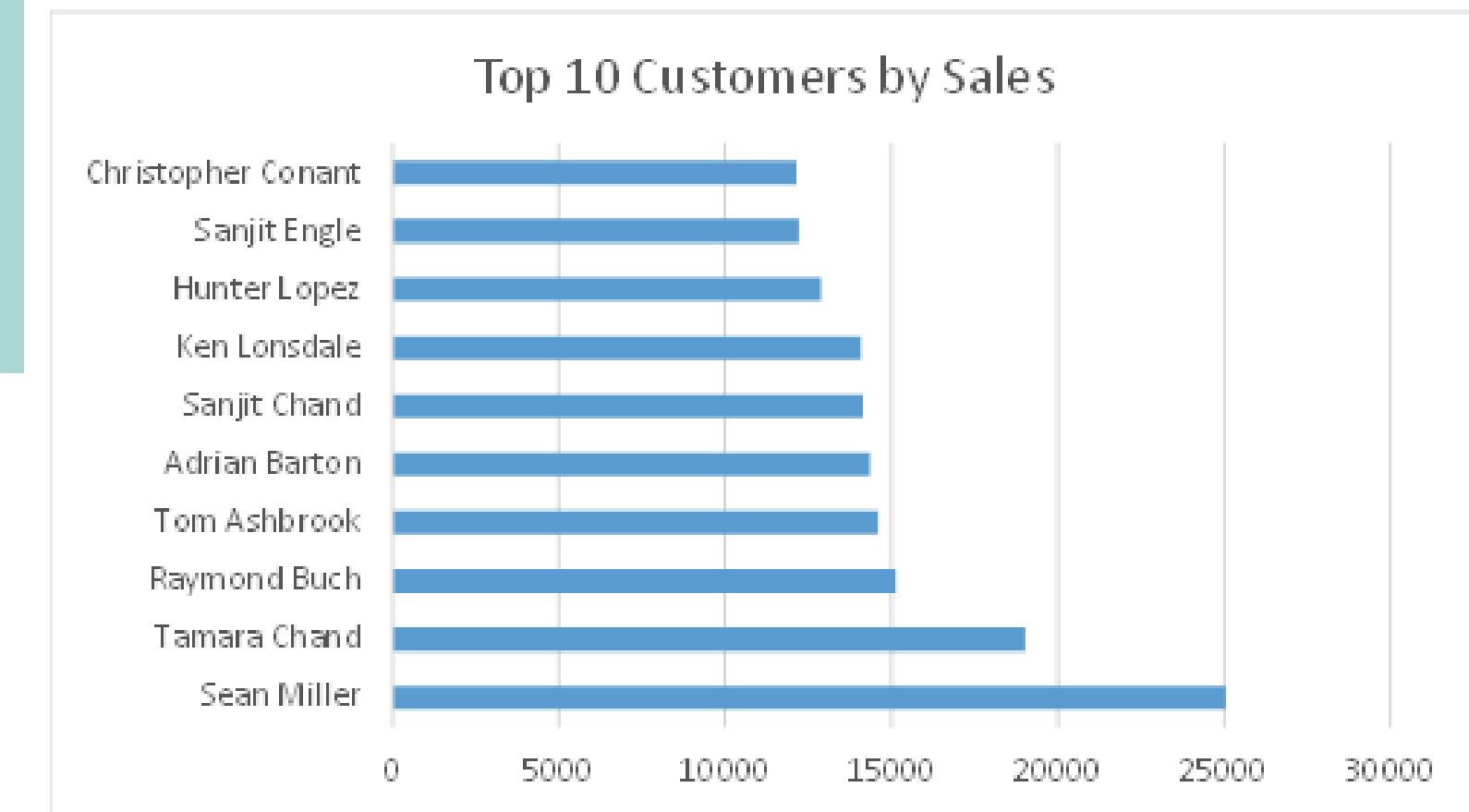


QUERY 2

Who are the top 10 customers by sales?

```
select `customer name`, sum(sales) as 'Total Sales'  
from superstore  
group by `customer name`  
order by `Total Sales` DESC  
limit 10 ;
```

customer name	Total Sales
Sean Miller	25043.05
Tamara Chand	19017.848
Raymond Buch	15117.339
Tom Ashbrook	14595.62
Adrian Barton	14355.611
Sanjit Chand	14142.334
Ken Lonsdale	14071.917
Hunter Lopez	12873.298
Sanjit Engle	12209.438
Christopher Conant	12129.072



Insight:

A few key customers drive most of the revenue, showing high customer concentration.

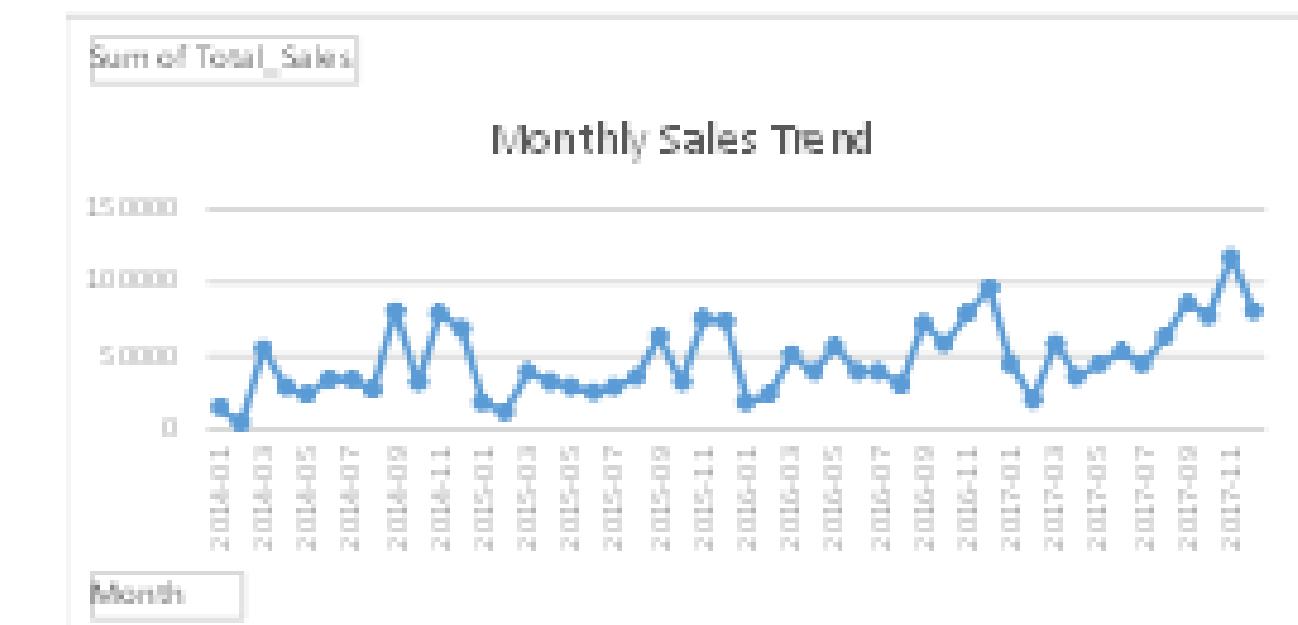
QUERY 3

What is the monthly sales trend over time?

```
SELECT DATE_FORMAT(STR_TO_DATE(`Order Date`, '%m/%d/%Y'), '%Y-%m')  
AS Month ,  
sum(sales) AS Total_Sales  
FROM superstore  
GROUP BY DATE_FORMAT(STR_TO_DATE(`Order Date`, '%m/%d/%Y'), '%Y-%m')  
ORDER BY Month;
```

Row Labels	Sum of Total_Sales	2015-11	74699.0335
2014-01	14161.349	2015-12	74478.4792
2014-02	4119.816	2016-01	18432.591
2014-03	55526.199	2016-02	22706.415
2014-04	28139.561	2016-03	50832.549
2014-05	23634.667	2016-04	38587.141
2014-06	34508.9956	2016-05	56457.32
2014-07	33500.873	2016-06	39628.192
2014-08	27603.5125	2016-07	39108.008
2014-09	81495.8068	2016-08	31014.4023
2014-10	31384.941	2016-09	71848.2509
2014-11	78297.2407	2016-10	58120.423
2014-12	69379.8365	2016-11	78454.9018
2015-01	18085.1156	2016-12	96075.067
2015-02	11924.272	2017-01	43860.286
2015-03	38621.292	2017-02	20262.3224
2015-04	32640.4825	2017-03	58739.2238
2015-05	29325.9705	2017-04	36020.1731
2015-06	24659.684	2017-05	44095.3522
2015-07	28524.521	2017-06	52242.5617
2015-08	36380.9282	2017-07	44490.448

2017-09	86487.305
2017-10	77542.4832
2017-11	117383.381
2017-12	81227.6158
Grand Total	2272449.856



Insight:

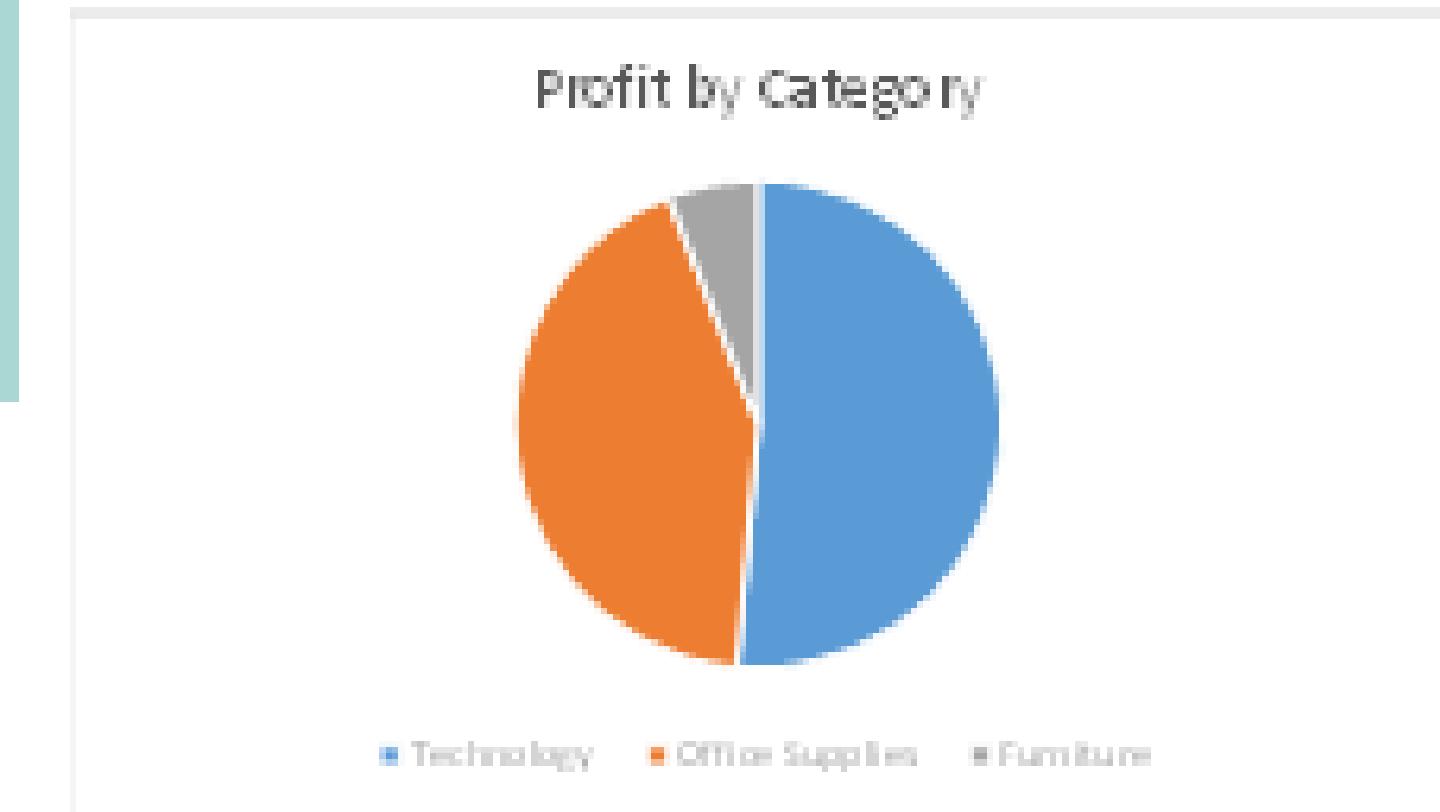
Sales fluctuate month to month, with visible peaks in certain periods.

QUERY 4

Which product categories are most profitable?

```
select category ,sum(profit) as 'Total_Profitable'  
from superstore  
group by category  
order by Total_Profitable DESC;
```

category	Total_Profitable
Technology	145387.0966
Office Supplies	120489.8864
Furniture	16980.7712



Insight:

Technology is the most profitable category, while others lag behind.

QUERY 5

*Which orders resulted in losses
(negative profit)?*

```
select `order Id`, `Product Name`, sales, profit
from superstore
where profit < 0
group by `order Id`, `Product Name`, sales, profit;
```



Insight:

Some products are sold at a loss due to discounts or costs.

order Id	Product Name	sales	profit
US/2015-108986	Bretford CR4500 Series SI	957.5775	-383.031
US/2015-118983	Holmes Replacement Filter	68.81	-123.858
US/2015-118983	Storex DuraTech Recycle	2.544	-3.816
US/2017-156909	Global Deluxe Stacking Cr	71.371	-1.0196
US/2015-150630	Riverside Pabis Royal Law	3083.43	-1665.0522
US/2015-150630	Avery Recycled Flexi-View	9.618	-7.0532
CA/2015-117580	Electrix Architect's Clamp	190.92	-147.963
CA/2015-117415	Atlantic Metals Mobile 34	532.3992	-46.9764
CA/2015-117415	Global Fabric Manager's C	212.058	-15.147
US/2015-164175	Global Value Mid-Back Mi	213.115	-15.2225
US/2015-134026	High Back Leather Manag	831.936	-114.3912
US/2015-134026	Safco Industrial Wire Shel	72.784	-18.196
US/2017-118038	Economy Binders	1.248	-1.9344
US/2014-147606	Eldon Expressions Desk A	19.3	-14.475
US/2017-119662	Safco Industrial Wire Shel	230.376	-48.9549
US/2017-109484	Flexible Leather Look Ch	5.682	-3.788
CA/2015-158568	SanDisk Ultra 64 GB Micro	95.976	-10.7973
CA/2015-158568	Avery Hidden Tab Divider	1.788	-3.0396
US/2015-156867	Logitech K350 2.4Ghz Wir	238.896	-26.8758
US/2015-156867	Deflect-o DuraMat Lighw	102.36	-3.8385
US/2015-136476	GBC DocuBind 300 Electri	157.794	-115.7156
US/2014-152030	Global Deluxe High Back I	600.558	-8.5794
US/2014-134614	Bevits 44 x 96 Conference	617.7	-407.682
US/2017-107272	Avery Durable Slant Ring I	2.388	-1.8308
US/2016-125969	Global Task Chair, Black	81.424	-9.1602
US/2017-164147	Anker 36W 4-Port USB W	59.97	-11.994
CA/2014-139892	Lexmark MX611dhe Mon	8158.952	-1369.992
CA/2014-139892	Space Solutions HD Indust	275.928	-58.6347
CA/2014-139892	SAFCO Arco Folding Chair	1740.06	-24.858
CA/2014-139892	Kendington 7 Outlet Mast	177.98	-453.849
US/2014-100853	Kensington 7 Outlet Mast	52.448	-131.12
US/2017-152366	Acco 7 Outlet Masterploc	97.264	-243.16
US/2015-101511	Padded Folding Chairs, Bl	396.802	-1133.72
US/2015-101511	Acme Rosewood Handle I	15.88	-3.7715
CA/2014-133690	BoxOffice By Design Rect	218.75	-161.875
US/2017-116701	Eureka Sanitaire Commode	66.284	-178.9668
CA/2015-146262	O'Sullivan 2 Door Banister	452.45	-244.323

Dataset 2: Northwind Database

- Microsoft sample relational database.
- Tables: Customers, Employees, Orders, Order Details, Products, Suppliers, Shippers.



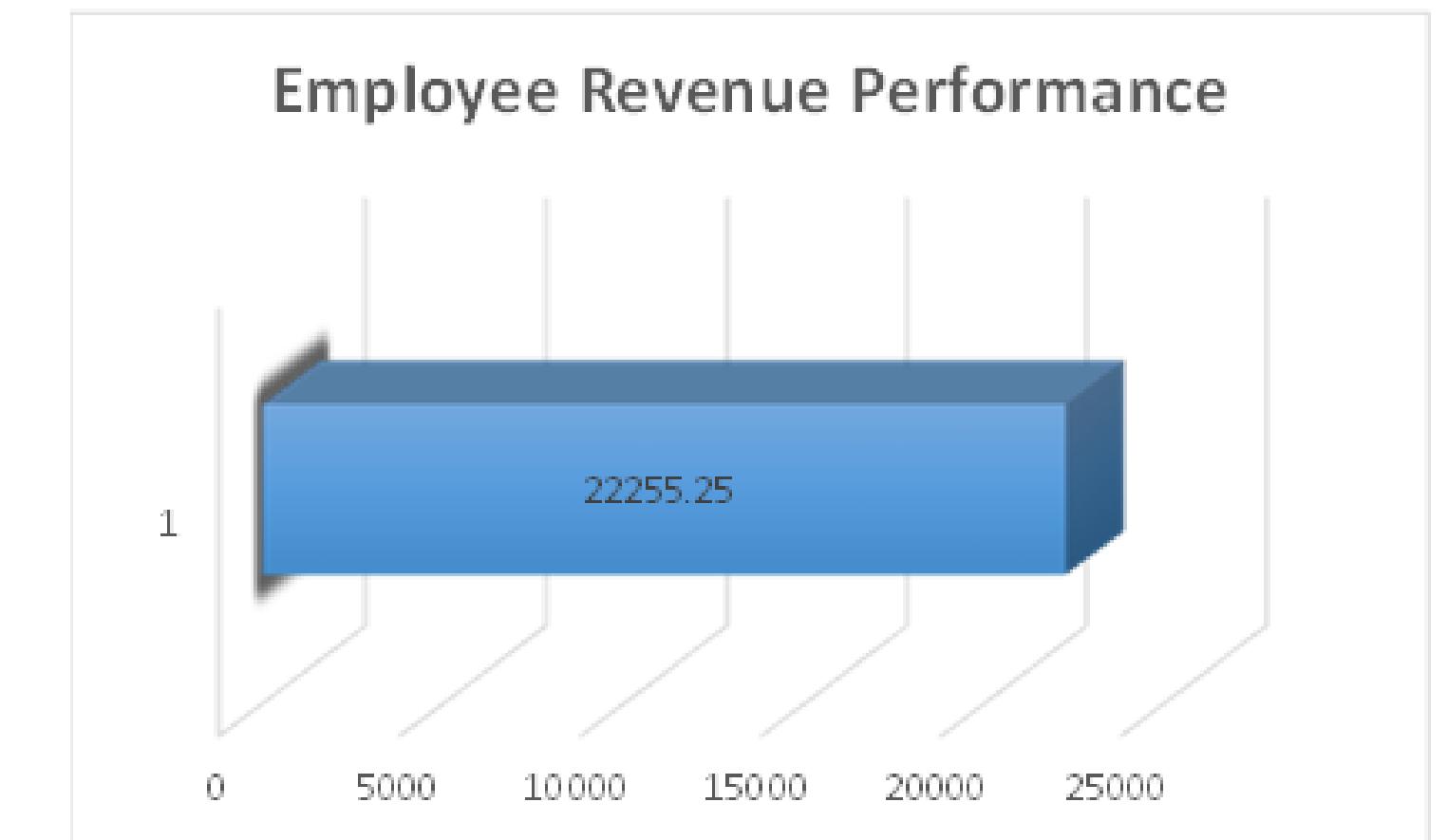
SQL Queries

QUERY 1

Which employee generated the highest revenue in the last year?

```
select e.id, e.first_name, e.last_name,  
sum(od.unit_price * od.quantity * (1 - od.discount)) as 'Highest_Revenue'  
from orders o  
join order_details od on o.id = od.order_id  
join employees e on e.id = o.employee_id  
where YEAR(o.order_date) = 2006  
group by e.id, e.first_name, e.last_name  
order by Highest_Revenue DESC  
limit 1;
```

<u>id</u>	<u>first_name</u>	<u>last_name</u>	<u>Highest_Revenue</u>
1	Nancy	Freehafer	22255.25



Insight:

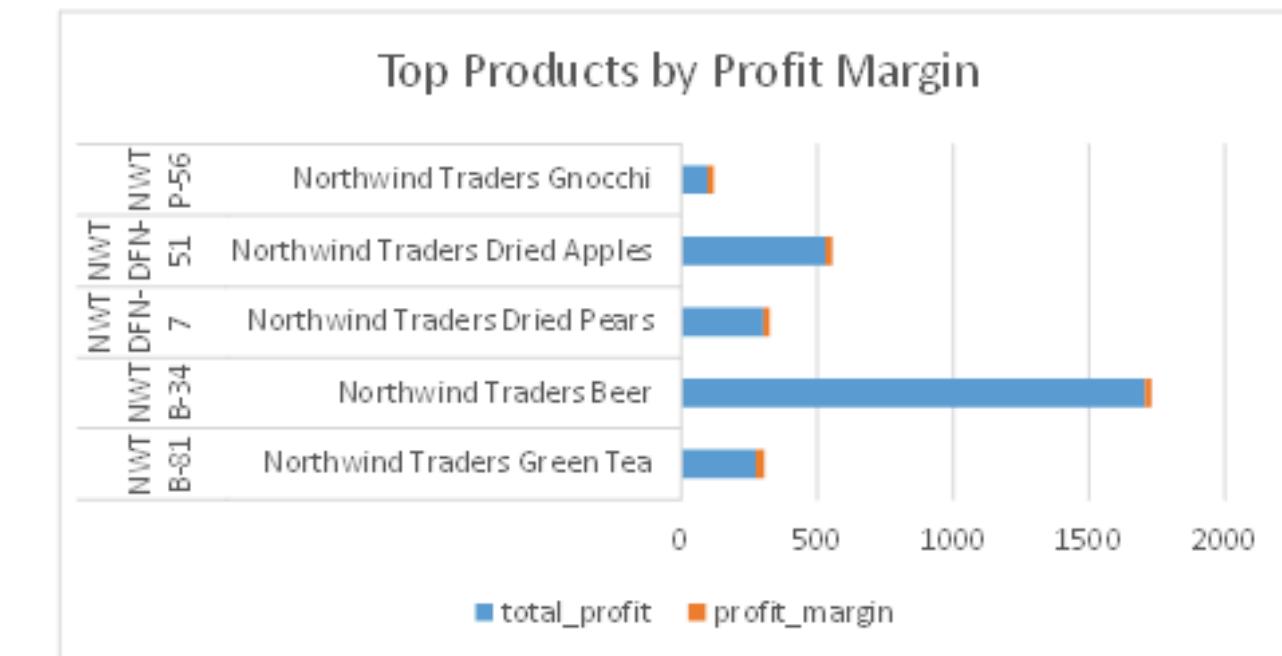
One employee generated the most revenue in 2006, highlighting top performance.

QUERY 2

What products have the highest profit margin?

```
SELECT
    p.product_code,
    p.product_name,
    SUM((od.unit_price - p.standard_cost) * od.quantity) AS total_profit,
    AVG((od.unit_price - p.standard_cost) / od.unit_price * 100) AS profit_margin
FROM products p
JOIN order_details od
    ON od.product_id = p.id
GROUP BY p.product_code, p.product_name
ORDER BY profit_margin DESC
LIMIT 5;
```

product_code	product_name	total_profit	profit_margin
NWTB-81	Northwind Traders Green Tea	272.25	33.11036789
NWTB-34	Northwind Traders Beer	1704.5	25
NWTDFN-7	Northwind Traders Dried Pears	300	25
NWTDFN-51	Northwind Traders Dried Apples	530	25
NWTP-56	Northwind Traders Gnocchi	95	25



Insight:

A few products have the highest profit margins, making them the most profitable to sell.

QUERY 3

Which customers haven't ordered in the last 6 months?

```
select c.first_name
from customers c
where c.id not in (
    select distinct o.customer_id
    from orders o
    where o.order_date >= DATE_SUB(CURDATE(), INTERVAL 6 MONTH));
```

Row Labels
Alexander
Amritansh
Andre
Anna
Antonio
Bernard
Carlos
Catherine
Christina
Daniel
Elizabeth
first_name
Francisco
George
Helena
Jean Philippe
John
Jonas
Karen
Ludana
Martin
Michael
Ming Yang
Peter
Roland
Run
Soo Jung
Sven
Thomas

Insight:

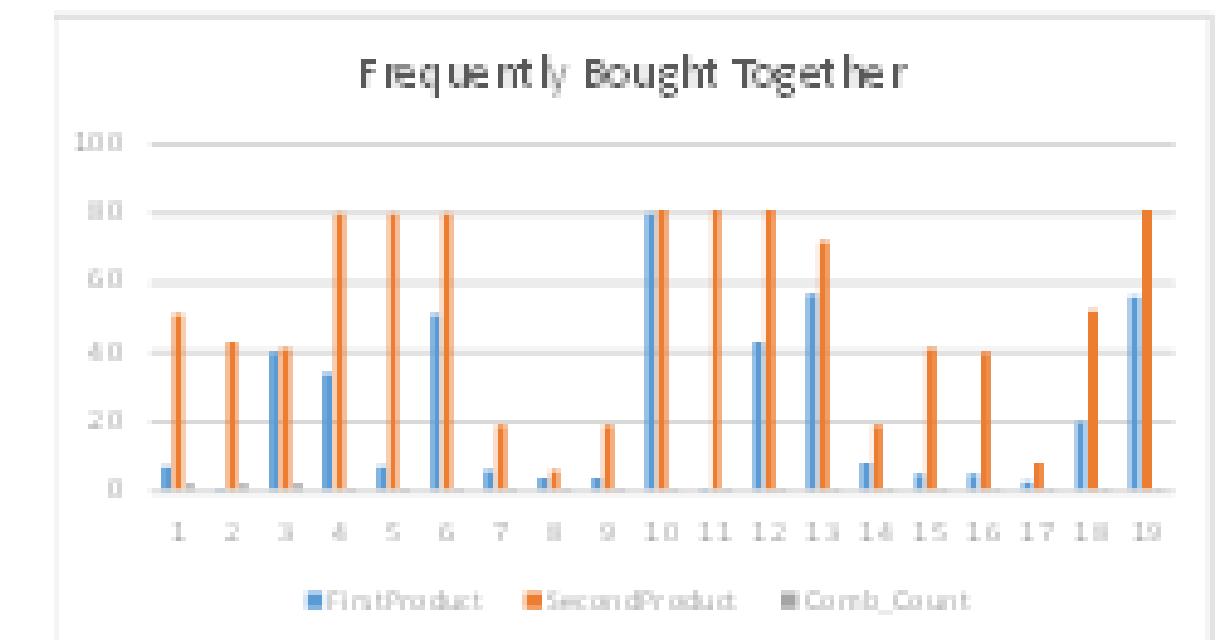
Several customers have not placed any orders in the last six months, showing retention gaps.

QUERY 4

Which product combinations are most frequently ordered together?

```
select od1.product_id as 'FirstProduct',
       od2.product_id as 'SecondProduct', count(*) as 'Comb_Count'
  from order_details od1
  join order_details od2 on od1.order_id = od2.order_id
  and od1.product_id < od2.product_id
 group by od1.product_id, od2.product_id
order by Comb_Count DESC;
```

FirstProduct	SecondProduct	Comb_Count
7	51	2
1	43	2
40	41	2
34	80	1
7	80	1
51	80	1
6	19	1
4	6	1
4	19	1
80	81	1
1	81	1
43	81	1
57	72	1
8	19	1
5	41	1
5	40	1
3	8	1
20	52	1
56	81	1



Insight:

Certain product combinations are frequently ordered together, suggesting cross-sell opportunities.

QUERY 5

Which suppliers provide the most products?

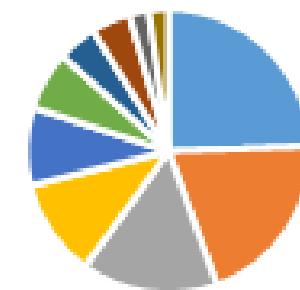
```
select s.first_name, count(p.id) as 'MostProductSells'  
from suppliers s  
join products p on s.id = p.supplier_ids  
group by s.first_name  
order by MostProductSells DESC;
```

Insight:

A few suppliers dominate product sourcing by providing the most items.

first_name	MostProductSells
Satomi	11
Elizabeth A.	9
Cornelia	7
Luis	5
Stuart	4
Bryn Paul	3
Naoki	2
Madeleine	2
Amaya	1
Mikael	1

Top Suppliers by Product Count



- Satomi ■ Elizabeth A. ■ Cornelia ■ Luis ■ Stuart
- Bryn Paul ■ Naoki ■ Madeleine ■ Amaya ■ Mikael

TOOLS & TECHNOLOGIES

MYSQL

To query and extract meaningful business insights.

GITHUB/ LINKEDIN

Platforms to showcase the project portfolio and reach recruiters.



CANVA

For presenting the queries, insights, and dashboards in a visually appealing way.

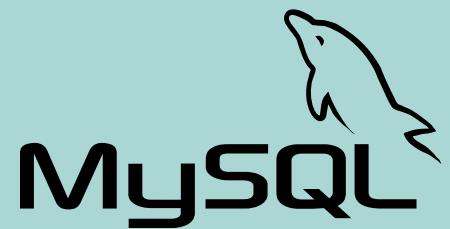
EXCEL

To design dashboards, create pivot tables, and build charts for visualization.

Key Insights & Recommendations

- High-value customers and top revenue products identified.
- Seasonal trends reveal sales peaks useful for planning.
- Certain products and discounts cause losses → review pricing strategy.
- Inactive customers highlight retention opportunities.
- Supplier and employee performance metrics guide sourcing and HR decisions.

Skills Demonstrated



**SQL (joins,
aggregation,
grouping,
filtering, date
functions)**



**Business
Analysis
(interpreting
sales, profit,
customers,
suppliers)**



**Data
Visualization
(Excel charts,
pivot tables,
dashboards)**

Conclusion

- This project demonstrates ability to analyze real-world datasets, apply SQL queries for business insights, and design dashboards for decision-making.
- Shows beginner → intermediate skills in SQL and Excel, with clear communication of insights.



**Thank you
very much!**

PRESENTED BY MARYAM RIAZ