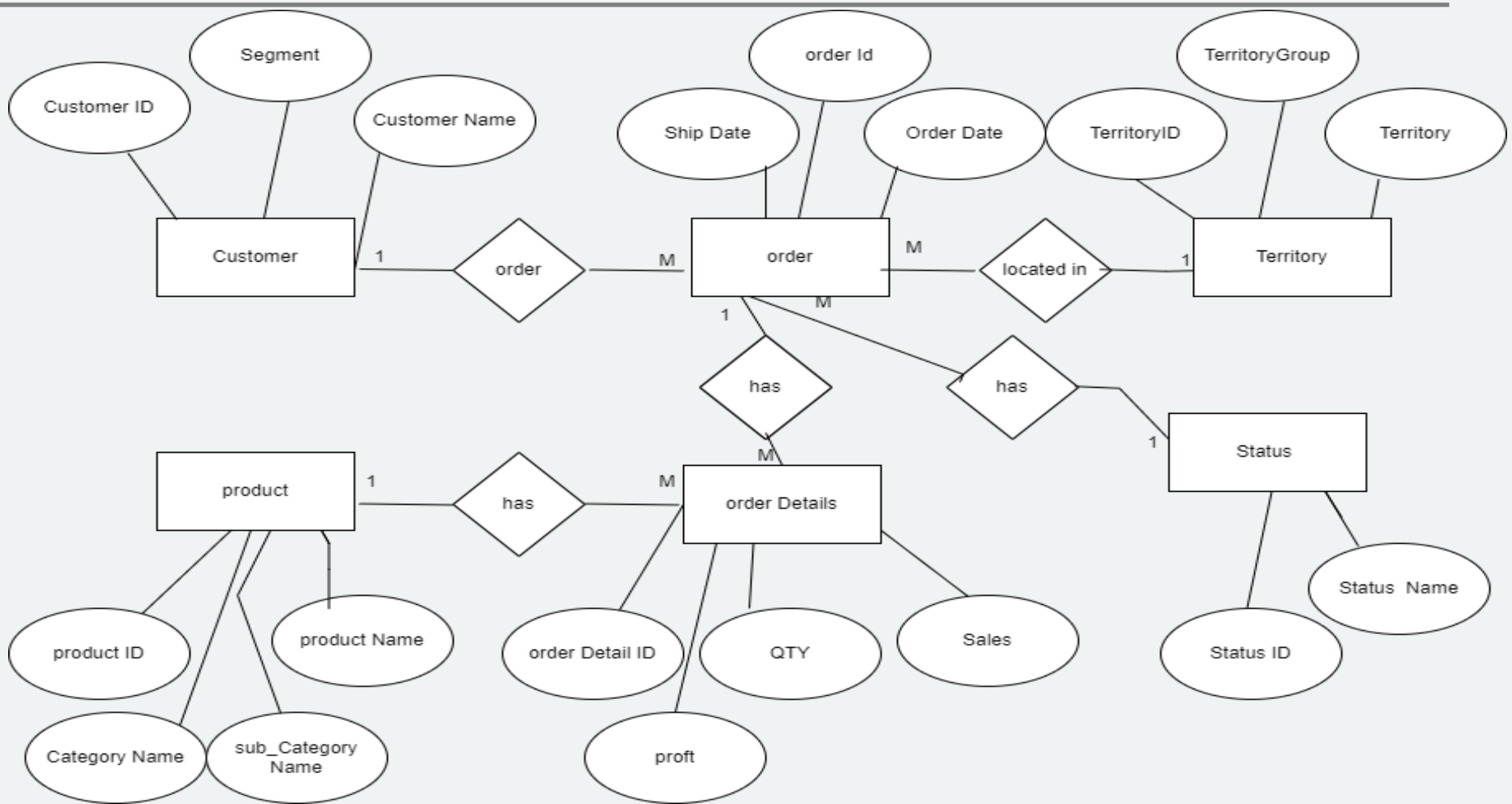
Superstore Dataset

Name: Nada Abdellatef Shaker
E-mail:nadaabdellatef22@gmail.com
phone:01027532602

Entity-Relationship Diagram (ERD) for a Superstore database



Entities and Attributes:

1.Customer

1. Customer ID
2. Customer Name
3. Segment

2.Order

1. Order ID
2. Order Date
3. Ship Date
4. Territory ID

3.Territory

1. Territory ID
2. Territory Name
3. Territory Group

Entities and Attributes:

4. Product

Product ID

Product Name

Category Name

Sub-Category Name

5. Order Details

Order Detail ID

Quantity (QTY)

Sales

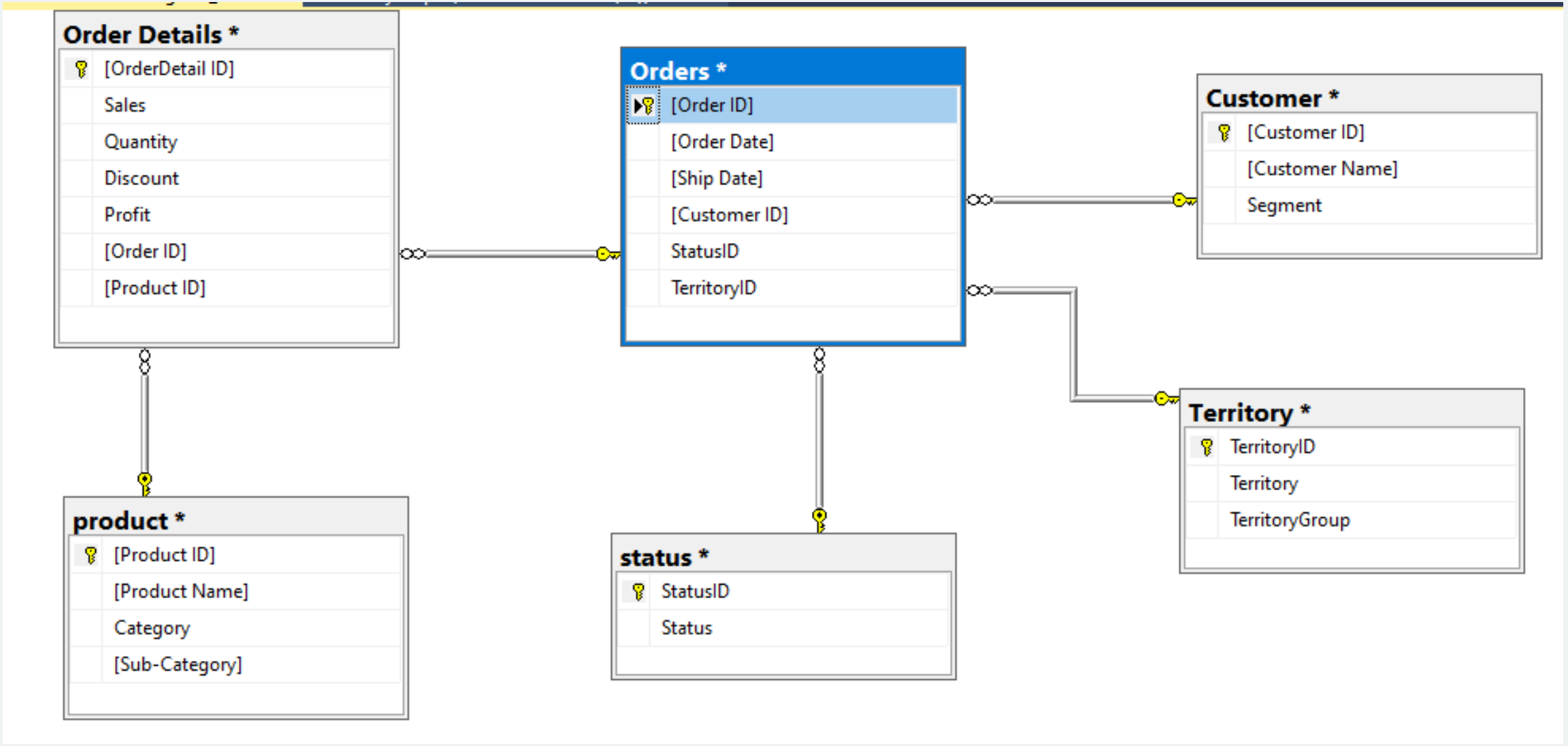
Profit

6. Status

Status ID

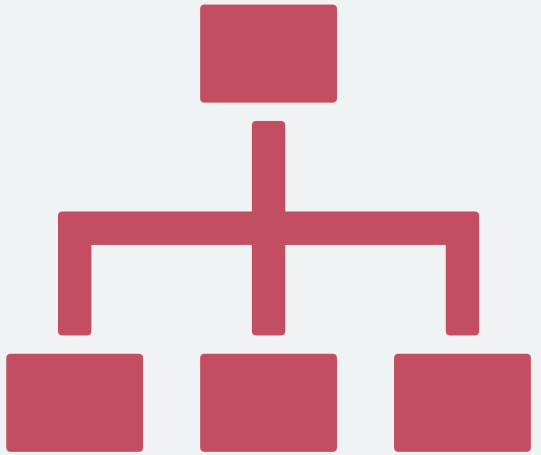
Status Name

Database Diagram



Relationships

- A **Customer** can place multiple **Orders**.
- An **Order** can contain multiple **Order Details**.
- Each **Order** is associated with a **Territory**.
- Each **Order Detail** includes a specific **Product**.
- **Order Details** has a **Status**.



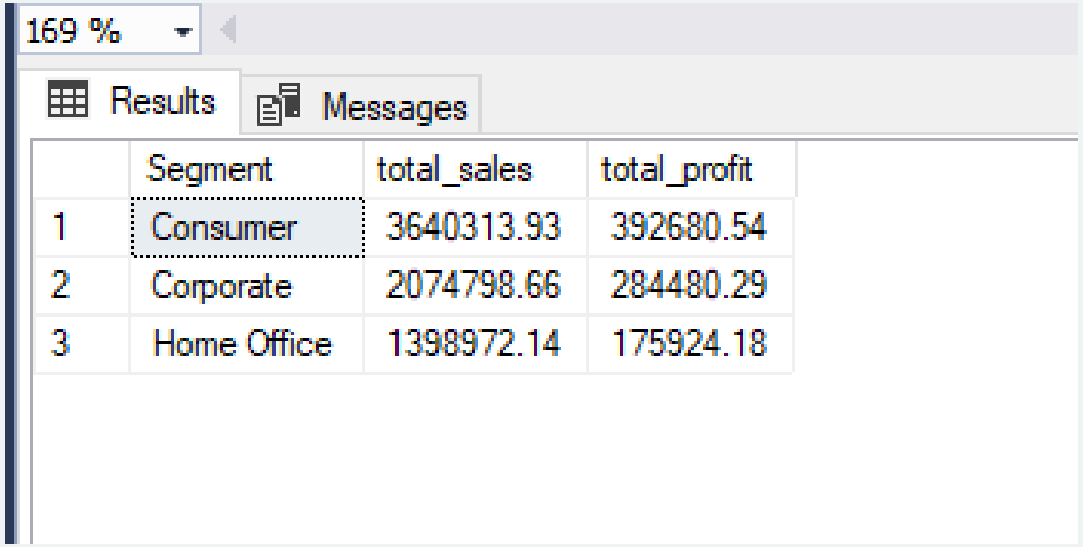
Queries

1. Which customer segment contributes the most to overall sales?

```sql
--1. Which customer segment contributes the most to overall sales?

SELECT Segment,ROUND (SUM(d.Sales), 2) AS total_sales,
ROUND (SUM(d.Profit), 2) AS total_profit
FROM [Order Details] d, Orders o,Customer c
where d.[Order ID]=o.[Order ID] and o.[Customer ID]=c.[Customer ID]
GROUP BY Segment
ORDER BY total_sales DESC, total_profit DESC
```

1. Result



	Segment	total_sales	total_profit
1	Consumer	3640313.93	392680.54
2	Corporate	2074798.66	284480.29
3	Home Office	1398972.14	175924.18

2. Which sub-category has the highest average profit margin?

```sql
--2. Which sub-category has the highest average profit margin?

SELECT TOP 1
[Sub-Category],
ROUND((SUM(d.Profit)/SUM(d.Sales))*100,2) AS Percentage_Profit_Margin
FROM product p inner join [Order Details] d on p.[Product ID]=d.[Product ID]
GROUP BY  [Sub-Category]
ORDER BY Percentage_Profit_Margin DESC
```

Results | Messages

	Sub-Category	Percentage_Profit_Margin
1	Labels	44.41

3. How does the sales volume vary across different territories?

```
SELECT  t.Territory  , SUM(Sales) as Total_Sales, SUM(Profit) as Total_Profits
FROM Territory t inner join Orders o on
t.TerritoryID= o.TerritoryID inner join [Order Details]
on [Order Details].[Order ID]=o.[Order ID]
GROUP BY t.Territory
ORDER BY Total_Profits ASC
```

Results | Messages

	Territory	Total_Sales	Total_Profits
1	Germany	93181.6904	6847.4193
2	Australia	67061.628	12155.0596
3	United Kingdom	885521.0247	124135.8901
4	France	1044111.4311	129314.0956
5	Northwest	1588934.7785	197330.5155
6	Canada	3435274.1751	383302.0219

4.is there a significant difference in sales volume between different order statuses?

```
SELECT  s.Status  , SUM(Sales) as Total_Sales,SUM(Profit) as Total_Profits
FROM status s inner join Orders o
on s.StatusID= o.StatusID inner join [Order Details]
on [Order Details].[Order ID]=o.[Order ID]
GROUP BY s.Status
ORDER BY Total_Profits ASC
```

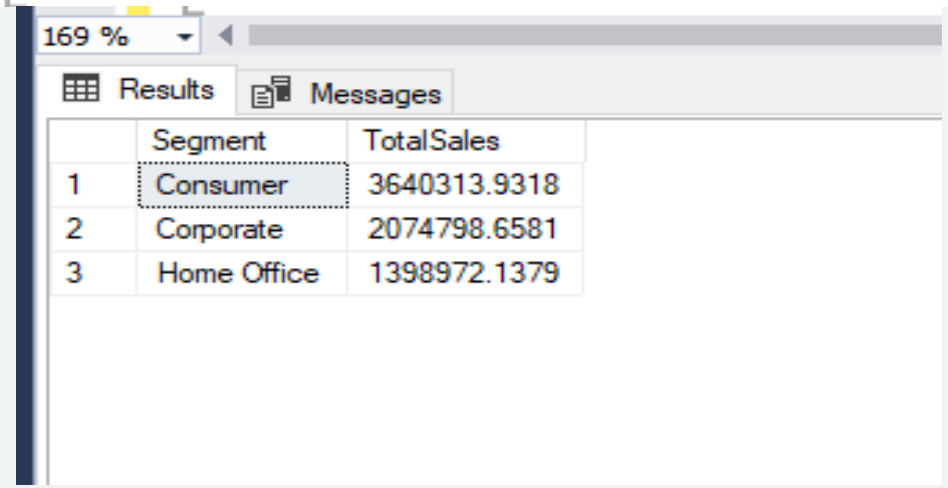
column Order ID(nvarchar, null)

Results | Messages

	Status	Total_Sales	Total_Profits
1	Backordered	397791.4598	44213.9364
2	Rejected	527462.736	68357.4153
3	Cancelled	899830.9335	82422.1268
4	In process	1582586.4361	157088.1067
5	Shipped	1744072.1179	196683.5552
6	Approved	1962341.0445	304319.8616

5. What factors influence sales more: the customer segment, the territory, or the product category? Provide a detailed analysis using a decomposition tree or another BI visualization.

```sql
--for Segment
SELECT c.Segment, SUM(od.Sales) AS TotalSales
FROM [Order Details] od
JOIN Orders o ON od.[Order ID] = o.[Order ID]
JOIN Customer c ON o.[Customer ID] = c.[Customer ID]
GROUP BY c.Segment
ORDER BY TotalSales DESC;
```



	Segment	TotalSales
1	Consumer	3640313.9318
2	Corporate	2074798.6581
3	Home Office	1398972.1379

```sql
--for territory
SELECT t.Territory, SUM(od.Sales) AS TotalSales
FROM [Order Details] od
JOIN Orders o ON od.[Order ID]  = o.[Order ID]
JOIN Territory t ON o.TerritoryID = t.TerritoryID
GROUP BY t.Territory
ORDER BY TotalSales DESC;
```

169 %

Results | Messages

	Territory	TotalSales
1	Canada	3435274.1751
2	Northwest	1588934.7785
3	France	1044111.4311
4	United Kingdom	885521.0247
5	Germany	93181.6904
6	Australia	67061.628

Product Category

```sql
----------Product Category
SELECT p.Category, SUM(od.Sales) AS TotalSales
FROM [Order Details] od
JOIN Product p ON od.[Product ID] = p.[Product ID]
GROUP BY p.Category
ORDER BY TotalSales DESC;
```

169 %

Results | Messages

	Category	TotalSales
1	Technology	836154.033
2	Furniture	741999.7953
3	Office Supplies	719047.032

6. Identify any seasonal trends in sales volume by analyzing the order and ship dates. How do these trends vary across different product categories?

```
SELECT
    p.Category AS product_category,
    MONTH(o.[Order Date])AS order_month,
    SUM(d.Sales) AS total_sales,
    COUNT(*) AS number_of_orders,
    MONTH(o.[Ship Date])AS  ship_month
    FROM [Order Details] d
JOIN product p ON d.[Product ID]= p.[Product ID]
JOIN Orders o ON o.[Order ID] = d.[Order ID]
GROUP BY p.Category,MONTH(o.[Order Date]) , MONTH(o.[Ship Date])
ORDER BY p.category, order_month;
```

Results | Messages

	product_category	order_month	total_sales	number_of_orders	ship_month
1	Furniture	1	78697.5928	204	1
2	Furniture	1	6016.67	30	2
3	Furniture	2	3239.848	14	3
4	Furniture	2	45292.2582	167	2
5	Furniture	3	109881.074	340	3
6	Furniture	3	17985.2818	42	4
7	Furniture	4	11947.2442	50	5
8	Furniture	4	91233.0385	302	4
9	Furniture	5	115189.1...	357	5

7. Determine the relationship between discount rates and profit margins. How do different discount levels impact overall profitability?

```
--Impact overall profitability?
SELECT
    d.Discount AS discount_rate,
    SUM(d.Profit / d.Sales) * 100 AS profit_margin,
    SUM(d.Profit) AS total_profit,
    COUNT(*) AS number_of_orders
FROM [Order Details] d
GROUP BY  d.Discount
ORDER BY d.Discount;
```

Results | Messages

	discount_rate	profit_margin	total_profit	number_of_orders
1	0	163209.00	320987.6032	4798
2	0.1	1464.04	9029.177	94
3	0.15	177.56	1418.9915	52
4	0.2	64670.00	90337.306	3657
5	0.3	-2620.47	-10369.2774	227
6	0.32	-470.46	-2391.1377	27
7	0.4	-4582.83	-23057.0504	206
8	0.45	-499.96	-2493.1111	11
9	0.5	-3624.00	-20506.4281	66

Query executed successfully.

8.Analyze the effect of order status on delivery time. Is there a significant difference in delivery times for

--different order statuses?

```sql
SELECT
    s.Status AS order_status,
    AVG(DATEDIFF(day, o.[Order Date] , o.[Ship Date])) AS average_delivery_time
    ,COUNT(*) AS number_of_orders
FROM Orders o inner join status s on o.StatusID=s.StatusID
WHERE o.[Ship Date] IS NOT NULL  -- Exclude orders that haven't been shipped yet
GROUP BY s.Status
ORDER BY average_delivery_time;
```

Results | Messages

	order_status	average_delivery_time	number_of_orders
1	Approved	3	2722
2	Backordered	3	497
3	In process	3	2576
4	Rejected	3	510
5	Shipped	3	2471
6	Cancelled	4	1218

9.which product sub-categories have shown the most growth in sales over the past years? Provide a yearover-year analysis

```
select p.[Sub-Category], sum(d.Sales) as total_sales,
YEAR(o.[Ship Date]) as ship_years from Product p inner join [Order Details] d
on p.[Product ID] = d.[Product ID]
inner join Orders o on o.[Order ID] = d.[Order ID]
group by p.[Sub-Category], YEAR(o.[Ship Date])
order by YEAR(o.[Ship Date])
```

Results | Messages

	Sub-Category	total_sales	ship_years
1	Labels	6744.406	2016
2	Bookcases	48226.1884	2016
3	Tables	136211.0345	2016
4	Art	15899.524	2016
5	Supplies	52147.596	2016
6	Chairs	226972.794	2016
7	Fasteners	2178.772	2016
8	Copiers	23509.54	2016
9	Machines	305690.284	2016

10. Develop a predictive model to forecast sales for the next quarter based on historical data. Consider factors such as product category, customer segment, and territory

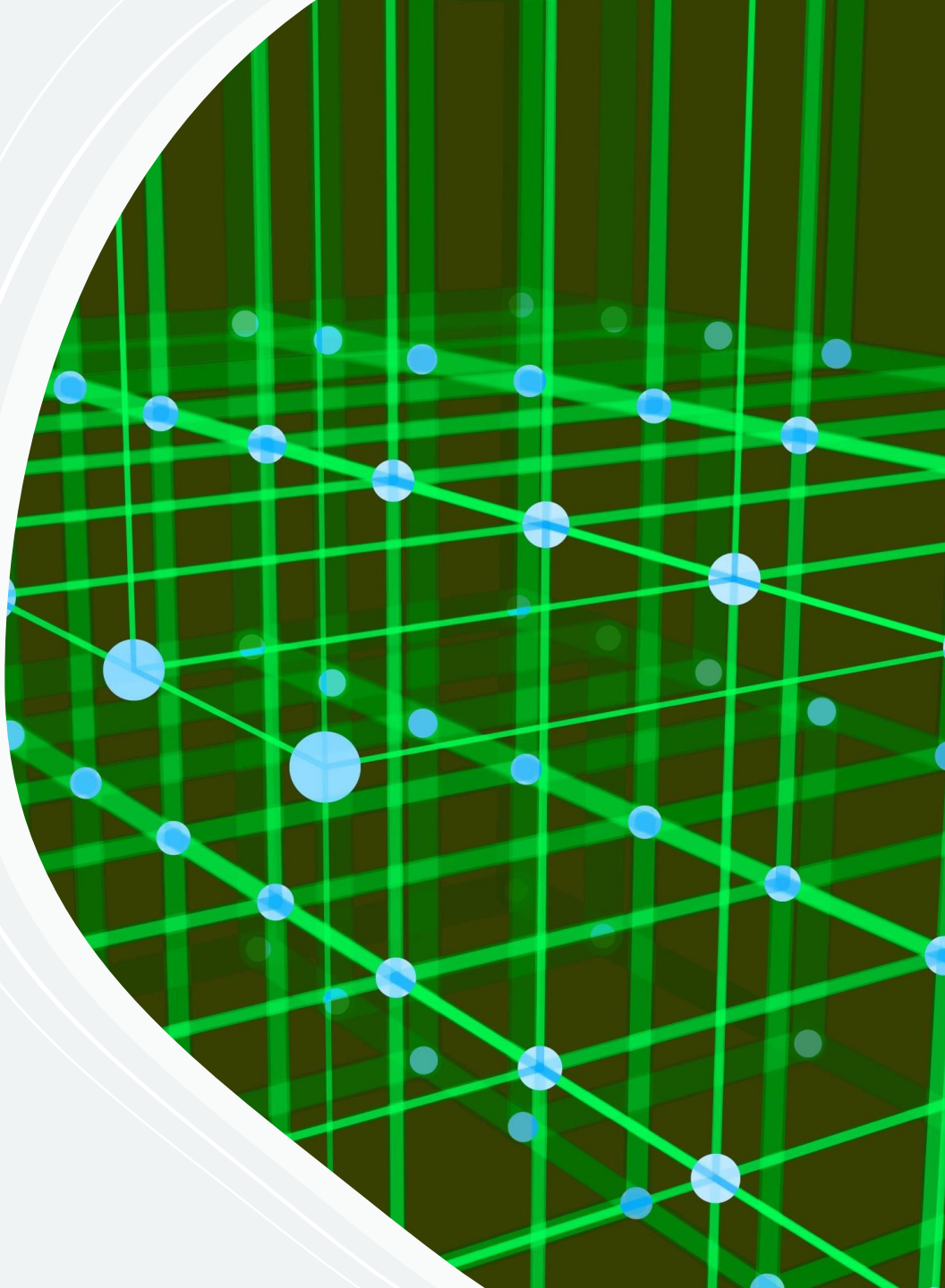
```sql
WITH sales_per_quarter AS (
    SELECT
        o.[Order Date],
        CASE
            WHEN MONTH(o.[Order Date]) IN (1, 2, 3) THEN 'Q1'
            WHEN MONTH(o.[Order Date]) IN (4, 5, 6) THEN 'Q2'
            WHEN MONTH(o.[Order Date]) IN (7, 8, 9) THEN 'Q3'
            ELSE 'Q4'
        END AS sales_quarter,
        d.Sales,
        d.Profit
    FROM
        Orders o
    INNER JOIN
        [Order Details] d
    ON
        o.[Order ID] = d.[Order ID]
)
```

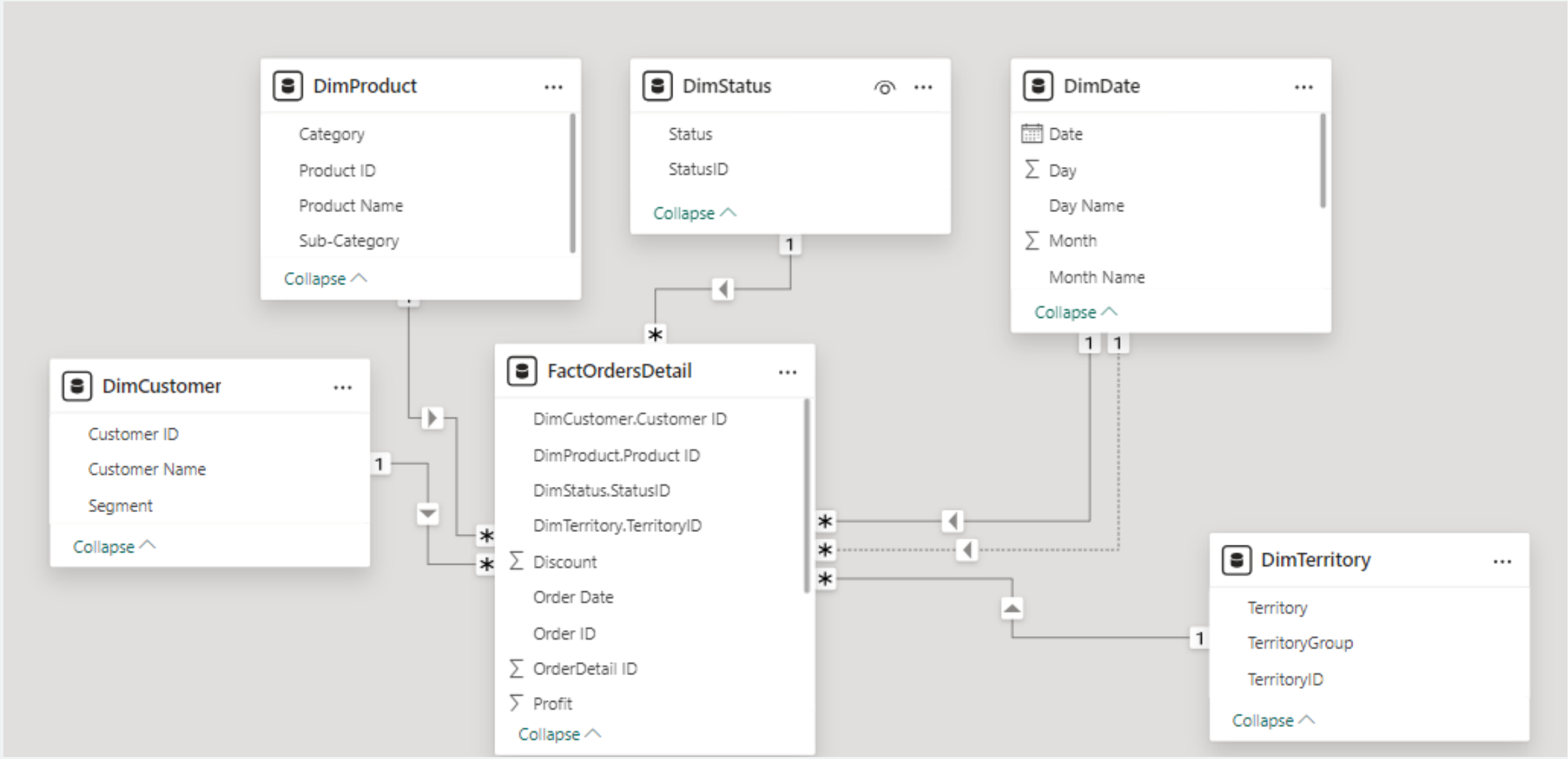
```sql
SELECT
    YEAR([Order Date]) AS year,
    sales_quarter,
    ROUND(SUM(Sales), 2) AS total_sales,
    ROUND(SUM(Profit), 2) AS total_profit
FROM
    sales_per_quarter
GROUP BY
    YEAR([Order Date]),
    sales_quarter
ORDER BY
    total_sales DESC,
    total_profit DESC,
    year,
    sales_quarter;
```

Results | Messages

	year	sales_quarter	total_sales	total_profit
1	2019	Q4	786319.48	81515.52
2	2019	Q3	726574.06	100548.15
3	2018	Q4	690650.37	113579.23
4	2017	Q4	531092.71	68664.92
5	2016	Q4	516038.14	63085.71
6	2017	Q3	503020.89	55241.28
7	2016	Q3	486652.67	35149.09
8	2018	Q3	428901.77	56237.28
9	2018	Q2	400087.43	45945.80
10	2019	Q1	371368.48	68670.91
11	2019	Q2	364620.69	33061.91

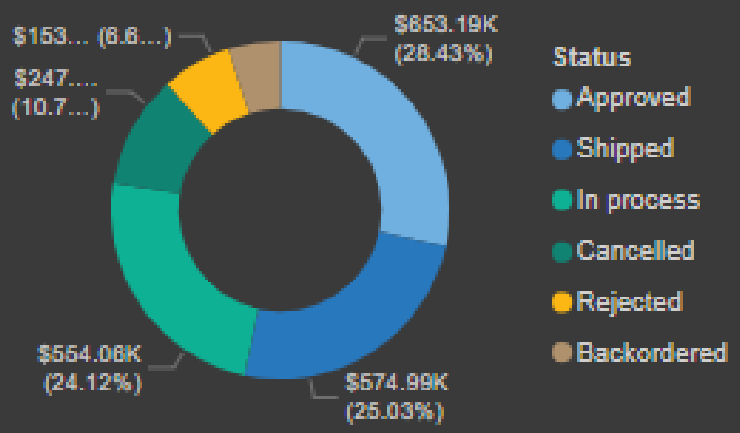
Dimensional Model



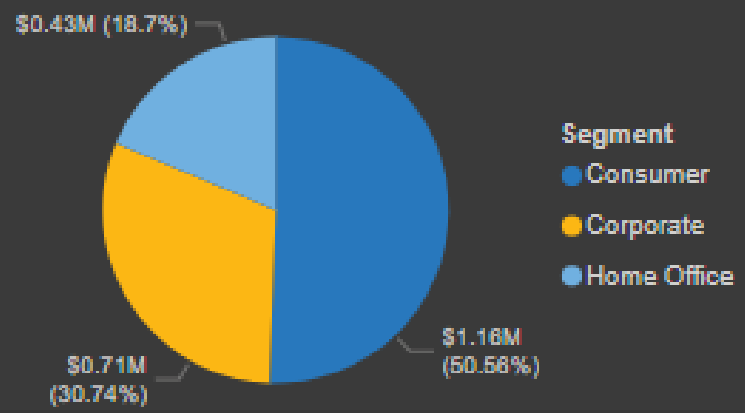

DimProduct
Category
Product ID
Product Name
Sub-Category
Collapse
DimStatus
Status
StatusID
Collapse
DimDate
Date
Day
Day Name
Month
Month Name
Collapse
DimCustomer
Customer ID
Customer Name
Segment
Collapse
FactOrdersDetail
DimCustomer.Customer ID
DimProduct.Product ID
DimStatus.StatusID
DimTerritory.TerritoryID
Discount
Order Date
Order ID
OrderDetail ID
Profit
Collapse
DimTerritory
Territory
TerritoryGroup
TerritoryID
Collapse
1
*

DashBoards

Total Sales by Status



Total Sales by Segment



$2.3M
Total Sales

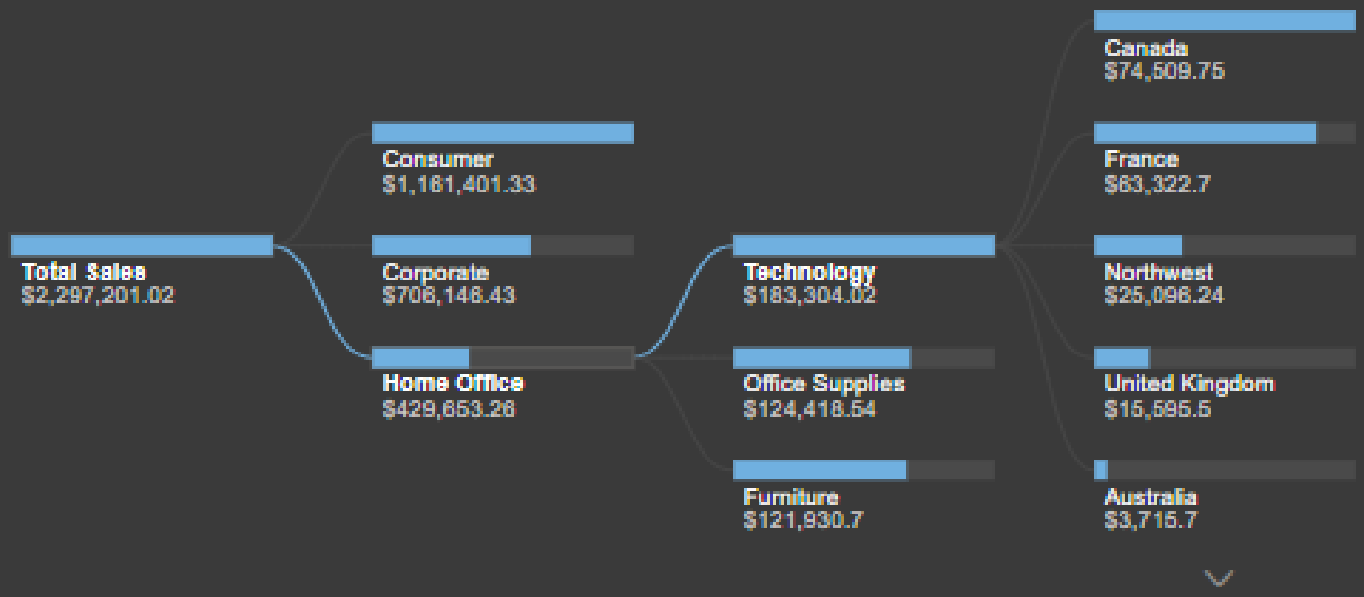
$28...
Sum of Profit

12.47
Average Profit Margin

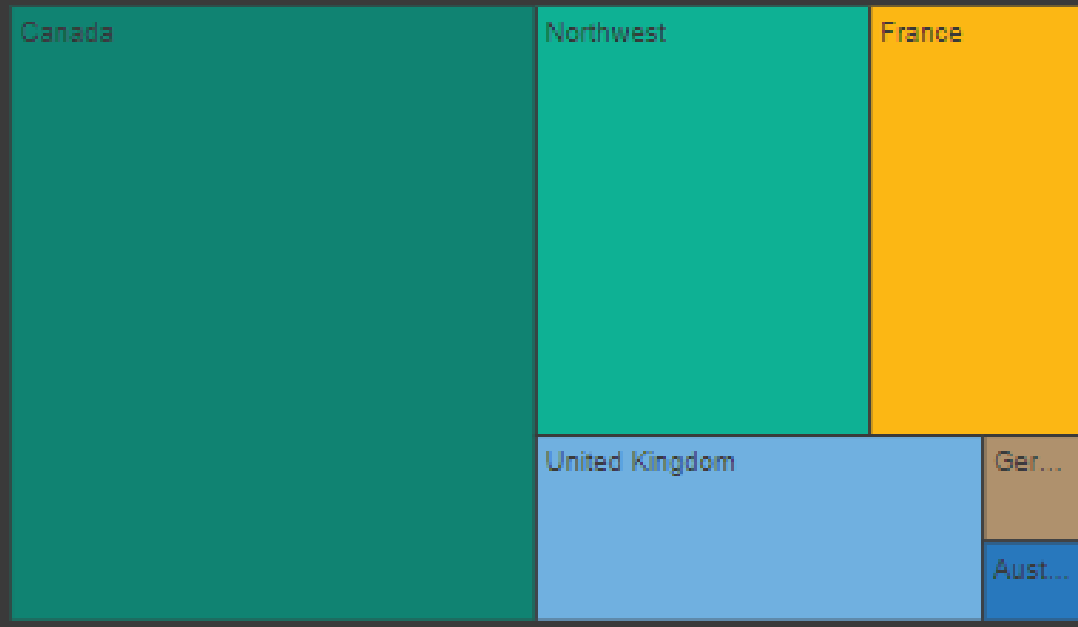
Segment ×
Home Office

Category ×
Technology

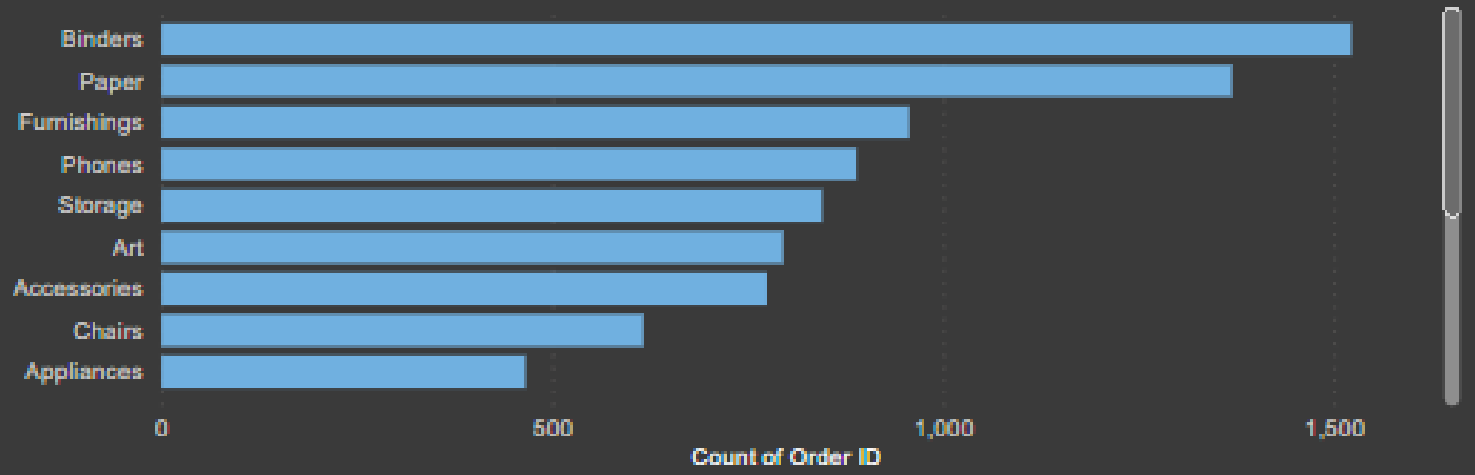
Territory ×



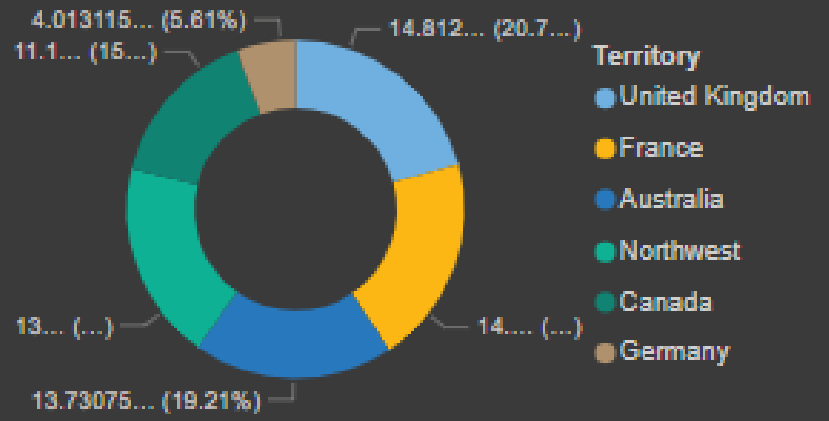
Total Sales by Territory



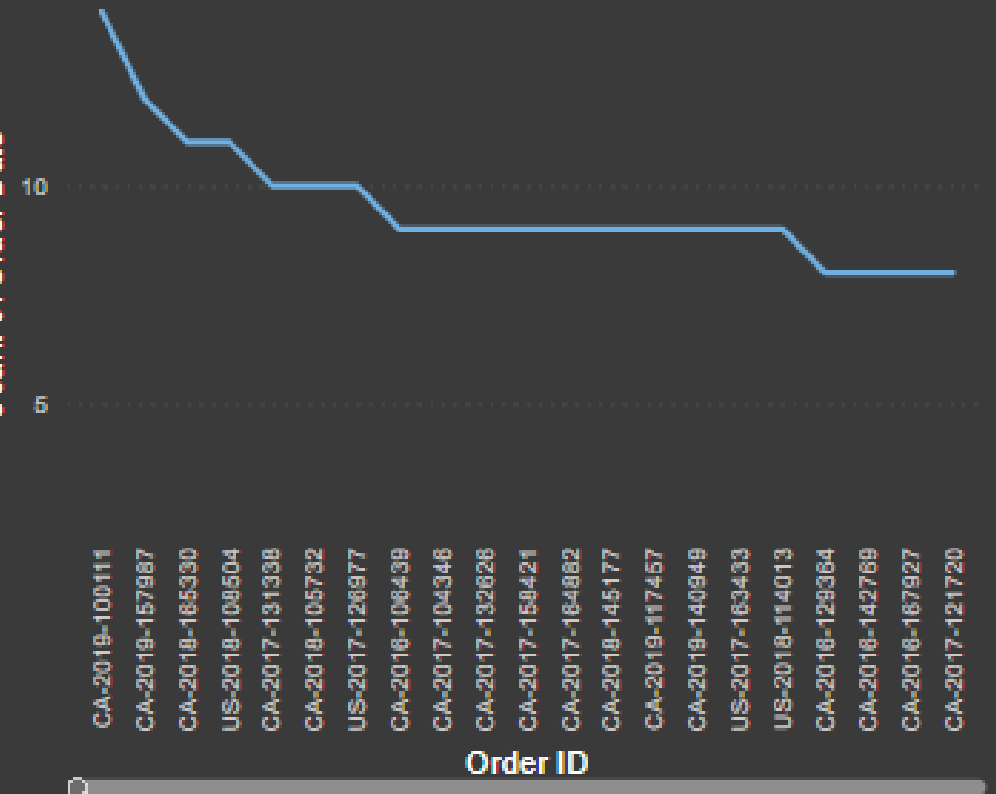
ount of Order ID by Sub-Category



Average Profit Margin by Territory

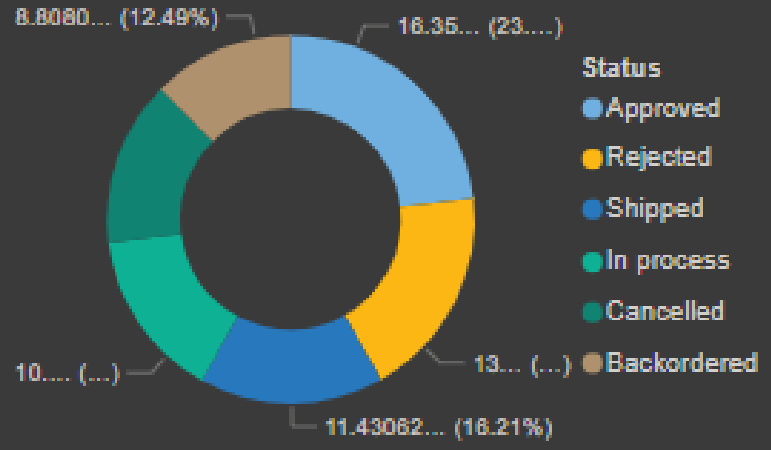


ount of Order Date by Order ID

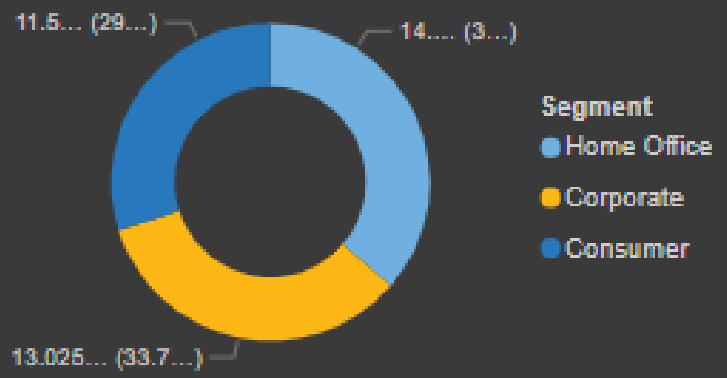


Sub-Category	Average Profit Margin
Labels	44.42
Paper	43.39
Envelopes	42.27
Copiers	37.20
Fasteners	31.40
Accessories	25.05
Art	24.07
Appliances	16.87
Binders	14.86
Furnishings	14.24
Phones	13.49
Storage	9.51
Chairs	8.10
Machines	1.79
Supplies	-2.55
Bookcases	-3.02
Tables	-8.56
Total	**12.47**

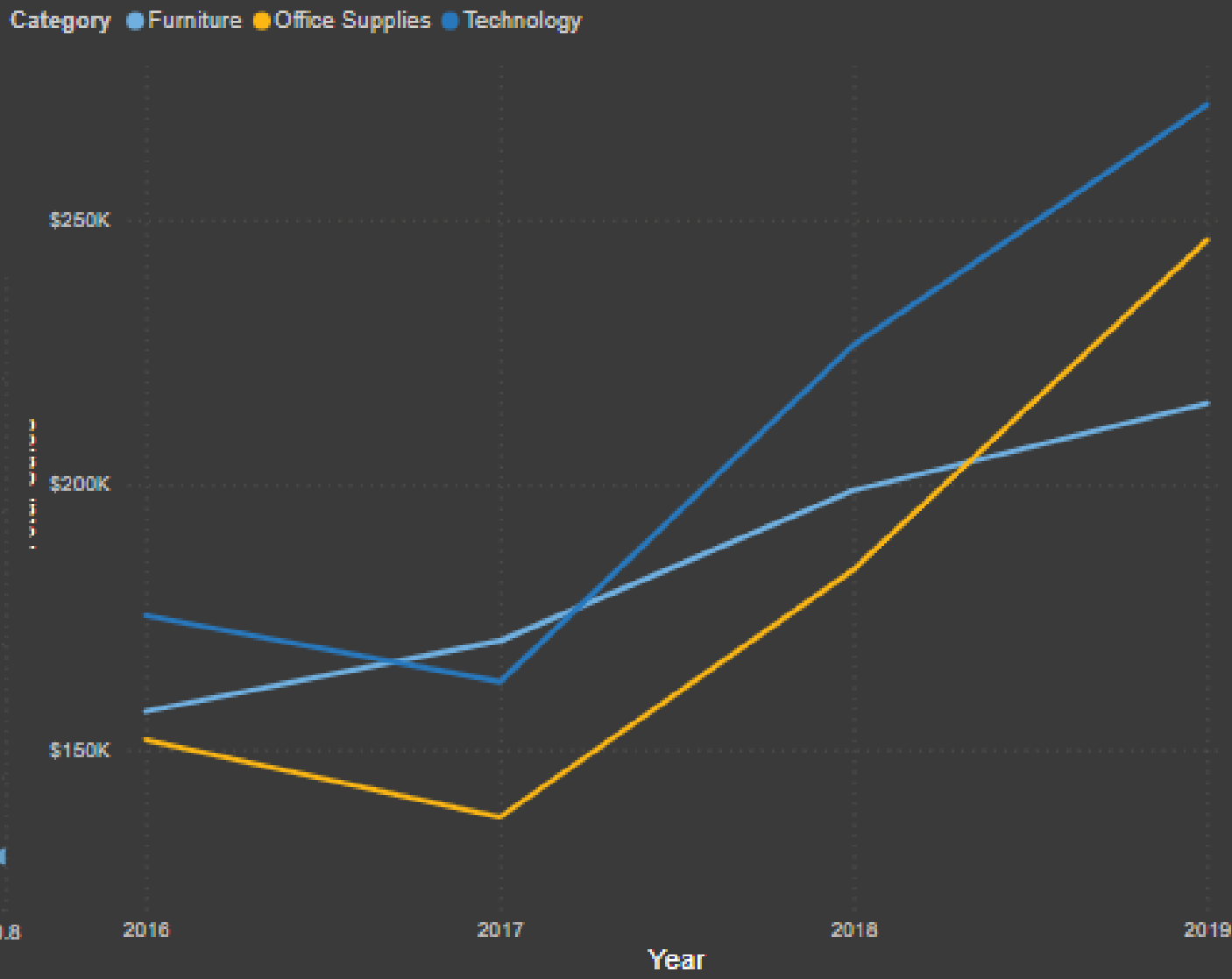
Average Profit Margin by Status



Average Profit Margin by Segment



Total Sales by Year and Category



Average Profit Margin by Discount