SQL Assignment

1. **What is the total profit generated by each region, and how does it correlate with the total number of returns from each region?**

```
select
    region,sum(profit) as tota_profit,count(Returned) as return
from
    `Sample - Superstore_Orders` as o left join `Sample - Superstore_Returns` as r on o.`order id`=r.`order id`
group by
    region
```

	Region	tota_profit	return
1	Central	66.0585999999998	0
2	East	-1188.4077	7
3	South	-39.4633000000001	0
4	West	1059.9217	3

2. **Which customers have the highest average sales per order, and what is their return rate?**

```
select
    `customer name`,avg(sales) as avg_sales,(count(r.Returned)/count(o.`order id`))*1.0 as return_rate
from
    `Sample - Superstore_Orders` as o left join `Sample - Superstore_Returns` as r on o.`order id`=r.`order id`
group by
    `customer name`
order by
    avg_sales desc
```

	Customer Name	avg_sales	return_rate
1	Emily Burns	1044.63	0.0
2	Pete Kriz	665.88	0.0
3	Gene Hale	644.232	0.0
4	Brendan Sweed	640.496	0.0
5	Brosina Hoffman	530.614857142857	0.0
6	Janet Molinari	522.605	0.0
7	Claire Gute	496.95	0.0
8	Sean O'Donnell	489.97275	0.0
9	Tracy Blumstein	418.5605	0.0
10	Irene Maddox	407.976	0.0
11	Steve Nguyen	307.2383	0.0
12	Patrick O'Donnell	256.96	0.0
13	Joel Eaton	255.265	0.0
14	Christopher Schild	230.376	0.0
15	Paul Stevenson	213.115	0.0
16	Julie Creighton	200.984	0.0
17	Gary Mitchum	158.368	0.0
18	Linda Cazamias	147.168	0.0

	Customer Name	avg_sales	return_rate
19	Stewart Carmichael	112.45	0.0
20	Parhena Norris	96.53	0.0
21	Erin Smith	95.616	0.0
22	Ted Butterfield	93.5842857142857	1.0
23	Zuschuss Donatelli	81.5866666666667	0.0
24	Ruben Ausman	77.88	0.0
25	Elpida Rittenbach	77.88	0.0
26	Karen Daniels	75.88	0.0
27	Sandra Flanagan	71.372	0.0
28	Kunst Miller	66.5735	0.0
29	Rick Bensley	64.624	0.0
30	Alejandro Grove	55.5	0.0
31	Darren Powers	52.3875	0.0
32	Katherine Ducich	51.312	0.0
33	Eric Hoffmann	51.109	0.0
34	Ken Black	39.9	0.0
35	Harold Pawlan	35.677	0.0
36	Lena Hernandez	33.4	0.0

3. **What is the trend of monthly sales over the past three years, segmented by region and customer segment?**

```sql
with latest_year as(
select
    max(cast(substr(`order date`,-4)as int)) as max_year
from
    `Sample - Superstore_Orders`
),
filter_data as (
select
    substr(`order date`,-4) as year,substr(`order date`,4,3) as month,region,segment,sum(sales) as total_sales
from
    `Sample - Superstore_Orders`,latest_year
where cast(substr(`order date`,-4)as int) between latest_year.max_year - 2 and latest_year.max_year
group by
    year,month,region,segment
)
select *
from
    filter_data
order by
    year DESC;
```

	year	month	region	segment	total_sales
1	2017	Apr	South	Consumer	15.552
2	2017	Dec	Central	Corporate	38.196
3	2017	Jul	East	Consumer	71.372
4	2017	Jun	West	Consumer	51.312
5	2017	May	South	Consumer	301.96
6	2017	Nov	Central	Home Office	230.376
7	2017	Nov	East	Home Office	96.53
8	2017	Nov	West	Home Office	5.682
9	2017	Oct	Central	Consumer	26.15
10	2017	Oct	Central	Home Office	29.472
11	2017	Sep	Central	Consumer	19.05
12	2017	Sep	Central	Corporate	147.168
13	2017	Sep	South	Corporate	95.616
14	2016	Apr	Central	Home Office	158.368
15	2016	Aug	Central	Home Office	64.624
16	2016	Dec	Central	Corporate	1368.264
17	2016	Dec	East	Corporate	1045.21
18	2016	Dec	West	Consumer	407.976
19	2016	Jan	West	Consumer	102.218
20	2016	Jul	West	Corporate	77.88
21	2016	Jun	East	Consumer	721.89

	year	month	region	segment	total_sales
20	2016	Jul	West	Corporate	77.88
21	2016	Jun	East	Consumer	721.89
22	2016	Jun	South	Consumer	75.88
23	2016	Jun	South	Corporate	224.9
24	2016	Jun	West	Corporate	14.62
25	2016	Mar	Central	Corporate	63.44
26	2016	Nov	South	Consumer	993.9
27	2016	Sep	Central	Corporate	77.88
28	2016	Sep	East	Consumer	4.616
29	2016	Sep	West	Corporate	100.164
30	2015	Apr	Central	Consumer	209.55
31	2015	Apr	Central	Home Office	213.115
32	2015	Apr	South	Consumer	1001.76
33	2015	Dec	Central	Home Office	1228.9532
34	2015	Jan	Central	Consumer	99.26
35	2015	Nov	Central	Home Office	71.354
36	2015	Nov	West	Consumer	266.294
37	2015	Oct	South	Consumer	979.9455
38	2015	Sep	East	Consumer	3329.434
39	2015	Sep	South	Corporate	200.984
40	2015	Sep	West	Consumer	1044.63

4. **Identify the top 10 products by sales in each region and determine their profit margins.**

```
with reg_pro_sales_all as(
select
    region,`product name`,sum(sales) as total_sales
from
    `Sample - Superstore_Orders`
group by
    region,`product name`
),
top_10_rank as (
select
    region,`product name`,total_sales,row_number() over(PARTITION by region order by total_sales desc) as rank
from
    reg_pro_sales_all
)

select
    region,`product name`,total_sales
from
    top_10_rank
where
    rank<=10
order by
    region,total_sales desc
```

	region	product name	total_sales
1	Central	GE 30524EE4	1097.544
2	Central	Stur-D-Stor Shelving, Vertical 5-…	665.88
3	Central	Atlantic Metals Mobile 3-Shelf …	532.3992
4	Central	Plantronics HL10 Handset Lifter	371.168
5	Central	Safco Industrial Wire Shelving	230.376
6	Central	Global Value Mid-Back Manager's …	213.115
7	Central	Global Fabric Manager's Chair, Dark …	212.058
8	Central	Gould Plastics 9-Pocket Panel Bin, …	211.96
9	Central	Electrix Architect's Clamp-On Swing …	190.92
10	Central	Eldon Portable Mobile Manager	158.368
11	East	Riverside Palais Royal Lawyers …	3083.43
12	East	AT&T CL83451 4-Handset Telephone	1029.95
13	East	Novimex Turbo Task Chair	319.41
14	East	Home/Office Personal File Carts	208.56
15	East	Howard Miller 13-3/4" Diameter …	124.2
16	East	9-3/4 Diameter Round Wall Clock	96.53
17	East	BOSTON Model 1800 Electric Pencil …	86.304
18	East	Imation 8gb Micro Traveldrive Usb …	75.0
19	East	Global Deluxe Stacking Chair, Gray	71.372
20	East	Plastic Binding Combs	48.48
21	South	Bretford CR4500 Series Slim …	957.5775

	region	product name	total_sales
20	East	Plastic Binding Combs	48.48
21	South	Bretford CR4500 Series Slim …	957.5775
22	South	High-Back Leather Manager's Chair	831.936
23	South	Hon Deluxe Fabric Upholstered …	731.94
24	South	Novimex Swivel Fabric Task Chair	301.96
25	South	Bush Somerset Collection Bookcase	261.96
26	South	1.7 Cubic Foot Compact "Cube" Office…	208.16
27	South	Jet-Pak Recycled Peel 'N' Seal Padde…	200.984
28	South	Tenex Traditional Chairmats for …	97.04
29	South	Advantus 10-Drawer Portable …	95.616
30	South	Snap-A-Way Black Print Carbonless …	75.88
31	West	Chromcraft Rectangular Conference …	1706.184
32	West	Hunt BOSTON Model 1606 High-Volume …	1113.024
33	West	Bretford CR4500 Series Slim …	1044.63
34	West	Konftel 250 Conference phone - …	911.424
35	West	Mitel 5320 IP Phone VoIP phone	907.152
36	West	Fellowes PB200 Plastic Comb Binding …	407.976
37	West	Cisco SPA 501G IP Phone	213.48
38	West	netTALK DUO VoIP Telephone Service	167.968
39	West	Xerox 1943	146.73
40	West	Belkin F5C206VTEL 6 Outlet Surge	114.9

5. Total Returns by Region and List of Customers Who Returned Orders

```
select
    o.region,o.`customer name`,count(r.Returned) as total_returns
from
    `Sample - Superstore_Orders` as o join `Sample - Superstore_Returns` as r
    on
        o.`order id`=r.`order id`
    group by
        region
```

Region	Customer Name	total_returns
East	Ted Butterfield	7
West	Jim Sink	3

6. **How does the average discount given affect the total sales and profit in each product category?**

```sql
select
    category,sum(profit) as total_profit,sum(sales) as total_sales,avg(discount)*100 as avg_discount
from
    `Sample - Superstore_Orders`
group by
        Category
    order by avg_discount
```

	Category	total_profit	total_sales	avg_discount
1	Technology	785.6685	5186.758	10.0
2	Office Supplies	763.7871	5233.994	15.1666666666667
3	Furniture	-1651.3463	11083.7797	20.2916666666667

7. **Which city has the highest return rate, and what are the common products returned from that city?**

```sql
with top_city as(
select o.city,
    CASE
            WHEN COUNT(DISTINCT o.`Order ID`) = 0 THEN 0
            ELSE COUNT(DISTINCT r.`Order ID`) * 1.0 / COUNT(DISTINCT o.`Order ID`) END AS return_rate
from
    `Sample - Superstore_Orders` as o left join `Sample - Superstore_Returns` as r
on
    o.`order id`=r.`order id`
group by o.city
order by return_rate DESC
limit 1)
select
    o.`city`,
    o.`product name`,
    COUNT(r.`Order ID`) AS return_count
from
    `Sample - Superstore_Orders` o
join
    `Sample - Superstore_Returns` r ON o.`Order ID` = r.`Order ID`
join
    top_city t ON o.`City` = t.`City`
group by
    o.`City`,
    o.`Product Name`
order by
    return_count DESC;
```

	City	Product Name	return_count
1	Troy	Array Parchment Paper, Assorted …	1
2	Troy	Home/Office Personal File Carts	1
3	Troy	Imation 8gb Micro Traveldrive Usb …	1
4	Troy	Novimex Turbo Task Chair	1
5	Troy	Plastic Binding Combs	1
6	Troy	Prang Dustless Chalk Sticks	1
7	Troy	Xerox 232	1

8. **What is the relationship between shipping mode and profit, considering different customer segments and regions?**

```
1  select
2      `ship mode`,`segment`,`region`,sum(profit) as total_profit from  `Sample - Superstore_Orders`
3  group by `ship mode`,`segment`,`region`
4  order by `ship mode`,`segment`,`region`
5
```

	Ship Mode	Segment	Region	total_profit
1	First Class	Consumer	East	102.2808
2	First Class	Consumer	South	35.6636
3	First Class	Consumer	West	17.9592
4	First Class	Corporate	Central	-1.54690000000013
5	First Class	Corporate	South	64.2384
6	First Class	Home Office	Central	-48.9549
7	Second Class	Consumer	Central	41.673
8	Second Class	Consumer	East	-1.0196
9	Second Class	Consumer	South	294.7112
10	Second Class	Consumer	West	51.8639
11	Second Class	Corporate	West	6.8714
12	Second Class	Home Office	Central	23.804
13	Second Class	Home Office	East	40.5426
14	Standard Class	Consumer	Central	96.4239
15	Standard Class	Consumer	East	-1635.1536
16	Standard Class	Consumer	South	-506.4456
17	Standard Class	Consumer	West	790.8092
18	Standard Class	Corporate	Central	59.8896
19	Standard Class	Corporate	East	304.9421
20	Standard Class	Corporate	South	72.3691
21	Standard Class	Corporate	West	196.206

9. **Identify the top 5 customers with the highest total sales, and analyze their purchasing patterns over time.**

```sql
with top_customers as(
select
    `customer id`,
    `customer name`,sum(sales) as total_sales
from
    `Sample - Superstore_Orders`
group by
    `customer name`
order by
    total_sales DESC
limit 5
)
select
    o.`customer id`,
    o.`customer name`,
    substr(`Order Date`, 4, 8) AS year_month,
    SUM(o.`Sales`) AS monthly_sales
from
    `Sample - Superstore_Orders` o
join
    top_customers t ON o.`customer id` = t.`customer id`
group by
    o.`customer id`,o.`customer name`,year_month
```

	Customer ID	Customer Name	year_month	monthly_sales
1	BH-11710	Brosina Hoffman	Jun-2014	3714.304
2	BS-11590	Brendan Sweed	Dec-2014	1280.992
3	GH-14485	Gene Hale	Dec-2016	1288.464
4	SN-20710	Steve Nguyen	Dec-2015	1228.9532
5	TB-21520	Tracy Blumstein	Sep-2015	3329.434
6	TB-21520	Tracy Blumstein	Sep-2017	19.05

10. How do sales and profits vary across different sub-categories within each product category?

```sql
SELECT
    `Category`,
    `Sub-Category`,
    SUM(Sales) AS total_sales,
    SUM(Profit) AS total_profit
FROM
    `Sample - Superstore_Orders`
GROUP BY
    `Category`, `Sub-Category`
ORDER BY
    `Category`, `Sub-Category`;
```

	Category	Sub-Category	total_sales	total_profit
1	Furniture	Bookcases	3877.7892	-1670.115
2	Furniture	Chairs	2771.781	131.2134
3	Furniture	Furnishings	725.818	-54.9878
4	Furniture	Tables	3708.3915	-57.4569000000002
5	Office Supplies	Appliances	530.09	5.08879999999999
6	Office Supplies	Art	1287.068	145.7102
7	Office Supplies	Binders	722.41	223.4431
8	Office Supplies	Envelopes	317.576	99.3241
9	Office Supplies	Fasteners	15.26	6.2566
10	Office Supplies	Labels	95.96	45.1628
11	Office Supplies	Paper	417.708	183.2701
12	Office Supplies	Storage	1847.922	55.5314
13	Technology	Accessories	245.52	52.7433
14	Technology	Phones	4941.238	732.9252

11. Determine the impact of discounts on sales and profits for orders that were shipped via 'First Class' and 'Standard Class'.

```
1   select `ship mode`,discount,sum(sales) as total_sales ,sum(profit) as total_profit
2   from
3       `Sample - Superstore_Orders`
4   where
5       `ship mode` in ('First Class','Standard Class')
6   group by
7       `ship mode`, discount
8   order by
9       `ship mode`, discount
```

	Ship Mode	Discount	total_sales	total_profit
1	First Class	0.0	651.42	206.7004
2	First Class	0.1	319.41	7.09800000000001
3	First Class	0.2	1454.952	111.564
4	First Class	0.6	200.628	-153.7878
5	First Class	0.8	1.248	-1.9344
6	Standard Class	0.0	3768.16	868.0827
7	Standard Class	0.2	7598.352	671.5992
8	Standard Class	0.3	425.173	-30.3695
9	Standard Class	0.32	532.3992	-46.9764
10	Standard Class	0.45	957.5775	-383.031
11	Standard Class	0.5	3083.43	-1665.0522
12	Standard Class	0.7	22.158	-16.5562
13	Standard Class	0.8	71.354	-127.674

16	Corporate	2015	Q3	200.984	62.8075
17	Corporate	2016	Q1	31.72	13.9888
18	Corporate	2016	Q2	79.84	23.7032666666667
19	Corporate	2016	Q3	51.1848	8.90016
20	Corporate	2016	Q4	402.245666666667	50.2001333333333
21	Corporate	2017	Q3	121.392	13.059
22	Corporate	2017	Q4	12.732	-1.6784
23	Home Office	2015	Q2	213.115	-15.2225
24	Home Office	2015	Q4	216.717866666667	-18.771
25	Home Office	2016	Q2	158.368	13.8572
26	Home Office	2016	Q3	64.624	22.6184
27	Home Office	2017	Q4	90.515	-0.563375000000001

12. What is the average order value and profit per customer segment for each quarter of the year?

```sql
SELECT segment,
    substr(`Order Date`, 8, 4) AS year,
    CASE
        WHEN substr(`Order Date`, 4, 3) in ('Jan','Feb','Mar') THEN 'Q1'
        WHEN substr(`Order Date`, 4, 3) in ('Apr','May','Jun') THEN 'Q2'
        WHEN substr(`Order Date`, 4, 3) in ('Jul','Aug','Sep') THEN 'Q3'
        WHEN substr(`Order Date`, 4, 3) in ('Oct','Nov','Dec') THEN 'Q4' ELSE '00' END AS quarter,
    avg(Sales) AS avg_order_value, avg(profit) as avg_profit
from
    `Sample - Superstore_Orders`
group by segment,year,quarter
ORDER BY
    `Segment`, year, quarter;
```

	Segment	year	quarter	avg_order_value	avg_profit
1	Consumer	2014	Q2	471.2255	38.8448375
2	Consumer	2014	Q3	81.5866666666667	8.62579999999999
3	Consumer	2014	Q4	186.686	3.53922
4	Consumer	2015	Q1	33.0866666666667	12.9688666666667
5	Consumer	2015	Q2	173.044285714286	-8.28875714285715
6	Consumer	2015	Q3	546.758	-175.9592125
7	Consumer	2015	Q4	207.706583333333	-45.6177166666667
8	Consumer	2016	Q1	51.109	7.99825
9	Consumer	2016	Q2	79.777	14.89984
10	Consumer	2016	Q3	4.616	1.731
11	Consumer	2016	Q4	467.292	131.3626
12	Consumer	2017	Q2	122.941333333333	18.8726666666667
13	Consumer	2017	Q3	45.211	3.8717
14	Consumer	2017	Q4	13.075	4.8767
15	Corporate	2014	Q4	640.496	87.1452
16	Corporate	2015	Q3	200.984	62.8075
17	Corporate	2016	Q1	31.72	13.9888
18	Corporate	2016	Q2	79.84	23.7032666666667
19	Corporate	2016	Q3	51.1848	8.90016
20	Corporate	2016	Q4	402.245666666667	50.2001333333333
21	Corporate	2017	Q3	121.392	13.059
22	Corporate	2017	Q4	12.732	-1.6784
23	Home Office	2015	Q2	213.115	-15.2225
24	Home Office	2015	Q4	216.717866666667	-18.771
25	Home Office	2016	Q2	158.368	13.8572
26	Home Office	2016	Q3	64.624	22.6184
27	Home Office	2017	Q4	90.515	-0.563375000000001

13. **Analyze the sales performance of products that have been returned at least once, compared to those that have never been returned.**

```sql
WITH ReturnStatus AS (
    SELECT
        `Product Name`,
        CASE WHEN SUM(CASE WHEN `Returned` = 'Yes' THEN 1 ELSE 0 END) > 0 THEN 'Returned' ELSE 'Not Returned' END AS return_status
    FROM
        `Sample - Superstore_Orders` as s  left join `Sample - Superstore_Returns` as r  on s.`order id`=r.`order id`
    GROUP BY
        `Product Name`),
SalesPerformance AS (
    SELECT
        r.return_status,
        SUM(s.`Sales`) AS total_sales,
        COUNT(DISTINCT s.`Order ID`) AS order_count,
        AVG(s.`Sales`) AS avg_sales_per_order
    FROM
        `Sample - Superstore_Orders` s JOIN ReturnStatus r
    ON
        s.`Product Name` = r.`Product Name`
    GROUP BY
        r.return_status
)
SELECT
    return_status,
    total_sales,
    order_count,
    avg_sales_per_order
FROM
    SalesPerformance
ORDER BY
    return_status;
```

	return_status	total_sales	order_count	avg_sales_per_order
1	Not Returned	20557.1097	47	233.603519318182
2	Returned	947.422	4	78.9518333333333

14. How do sales and profits differ between orders that have been returned and those that have not, across different regions and customer segments?

```sql
WITH OrderReturnStatus AS (
    SELECT
        o.`Order ID`,
        o.`Region`,
        o.`Segment`,
        o.`Sales`,
        o.`Profit`,
        CASE
            WHEN r.`Returned` = 'Yes' THEN 'Returned'
            ELSE 'Not Returned'
        END AS return_status
    FROM
        `Sample - Superstore_Orders` o
    LEFT JOIN
        `Sample - Superstore_Returns` r
    ON
        o.`Order ID` = r.`Order ID`
),
AggregatedMetrics AS (
    SELECT
        `Region`,
        `Segment`,
        `return_status`,
        SUM(`Sales`) AS total_sales,
        SUM(`Profit`) AS total_profit,
        COUNT(DISTINCT `Order ID`) AS order_count
    FROM
        OrderReturnStatus
    GROUP BY
        `Region`,
        `Segment`,
    GROUP BY
        `Region`,
        `Segment`,
        `return_status`
)
SELECT
    `Region`,
    `Segment`,
    `return_status`,
    total_sales,
    total_profit,
    order_count,
    total_sales * 1.0 / order_count AS avg_sales_per_order,
    total_profit * 1.0 / order_count AS avg_profit_per_order
FROM
    AggregatedMetrics
ORDER BY
    `Region`,
    `Segment`,
    `return_status`;
```

	Region	Segment	return_status	total_sales	total_profit	order_count	avg_sales_per_order	avg_profit_per_order
1	Central	Consumer	Not Returned	1251.15	138.0969	7	178.735714285714	19.7281285714286
2	Central	Corporate	Not Returned	1694.948	58.3426999999999	6	282.491333333333	9.72378333333331
3	Central	Home Office	Not Returned	1996.2622	-130.381	7	285.180314285714	-18.6258571428572
4	East	Consumer	Not Returned	3472.222	-1636.1732	4	868.0555	-409.0433
5	East	Consumer	Returned	655.09	102.2808	1	655.09	102.2808
6	East	Corporate	Not Returned	1045.21	304.9421	1	1045.21	304.9421
7	East	Home Office	Not Returned	96.53	40.5426	1	96.53	40.5426
8	South	Consumer	Not Returned	3368.9975	-176.0708	6	561.499583333333	-29.3451333333334
9	South	Corporate	Not Returned	521.5	136.6075	3	173.833333333333	45.5358333333333
10	West	Consumer	Not Returned	5923.284	860.6323	9	658.142666666667	95.6258111111111
11	West	Corporate	Not Returned	1373.492	185.0558	3	457.830666666667	61.6852666666667
12	West	Corporate	Returned	100.164	18.0216	1	100.164	18.0216
13	West	Home Office	Not Returned	5.682	-3.788	1	5.682	-3.788

15. Identify the correlation between order quantity and profit margin for each product sub-category

```sql
with Correlation_formula as (
    Select
        `sub-category`,
        count(*) as n,
        sum(`quantity`) as total_quantity,
        sum((profit / sales) * 100) AS profit_margin,
        sum(`quantity` * (profit / sales) * 100) AS quantity_mul_profit_margin,
        sum(`quantity` * `quantity`) AS quantity_square,
        sum(((profit / sales) * 100) * ((profit / sales) * 100)) AS profit_margin_square
    from
        `Sample - Superstore_Orders`
    group by
        `sub-category`
)
select
    `sub-category`,
    total_quantity,
    profit_margin,
    (n * quantity_mul_profit_margin - total_quantity * profit_margin) /
    sqrt((n * quantity_square - total_quantity * total_quantity) * (n * profit_margin_square - profit_margin * profit_margin)) AS correlat
from
    Correlation_formula
order by
    correlation_quantity_profit_margin DESC;
```

	sub-category	total_quantity	profit_margin	correlation_quantity_profit_margin
1	Storage	39.0	76.75	0.134377835054025
2	Binders	54.0	-62.3333333333335	0.0420317480584363
3	Furnishings	41.0	-20.7500000000001	-0.0235597341571733
4	Appliances	24.0	-68.0	-0.0478435401716933
5	Tables	17.0	-12.0	-0.0907841299003206
6	Phones	41.0	164.5	-0.127409976499844
7	Paper	31.0	435.75	-0.479142498646424
8	Labels	10.0	142.0	-0.5
9	Accessories	13.0	156.0	-0.594088525786005
10	Chairs	29.0	32.7579365079365	-0.622656528397848
11	Art	38.0	236.75	-0.735790645906699
12	Envelopes	18.0	96.25	-0.96076892283053
13	Bookcases	12.0	-46.8235294117647	-0.98606199573392

16. How does the frequency of orders placed by each customer segment vary across different months of the year?

```
select
    segment,substr(`order date`,-4) as year,substr(`order date` ,4,3) as month,count(`order id`) as total_orders
from
    `Sample - Superstore_Orders`
group by
    segment,year,month
order by segment,year,month
```

	Segment	year	month	total_orders
1	Consumer	2014	Aug	3
2	Consumer	2014	Jun	7
3	Consumer	2014	May	1
4	Consumer	2014	Nov	2
5	Consumer	2014	Oct	3
6	Consumer	2015	Apr	7
7	Consumer	2015	Jan	3
8	Consumer	2015	Nov	4
9	Consumer	2015	Oct	2
10	Consumer	2015	Sep	8
11	Consumer	2016	Dec	1
12	Consumer	2016	Jan	2
13	Consumer	2016	Jun	10
14	Consumer	2016	Nov	2
15	Consumer	2016	Sep	1
16	Consumer	2017	Apr	1
17	Consumer	2017	Jul	1
18	Consumer	2017	Jun	1
19	Consumer	2017	May	1
20	Consumer	2017	Oct	2
21	Consumer	2017	Sep	1
22	Corporate	2014	Dec	2
23	Corporate	2015	Sep	1
24	Corporate	2016	Dec	6
25	Corporate	2016	Jul	1
26	Corporate	2016	Jun	3
27	Corporate	2016	Mar	2
28	Corporate	2016	Sep	4
29	Corporate	2017	Dec	3
30	Corporate	2017	Sep	2
31	Home Office	2015	Apr	1
32	Home Office	2015	Dec	4
33	Home Office	2015	Nov	2
34	Home Office	2016	Apr	1
35	Home Office	2016	Aug	1
36	Home Office	2017	Nov	3
37	Home Office	2017	Oct	1

17. What are the top 5 most profitable products in each region, and how frequently are they returned?

```
with tol_pro as(
select
    o.region,o.`product name`,sum(o.profit)as total_profit,count(CASE when r.Returned = 'Yes' then 1 END) as return_count
from
    `Sample - Superstore_Orders` as o
left join
    `Sample - Superstore_Returns` as r
on o.`order id`=r.`order id`
group by region,`product name`),
rank_product as(
select
    region,`product name`,total_profit,return_count,
    rank() over(PARTITION by region order by total_profit desc) as rank
from tol_pro)
select region,`product name`,total_profit,return_count from rank_product where rank<=5
```

	region	product name	total_profit	return_count
1	Central	GE 30524EE4	123.4737	0
2	Central	Plantronics HL10 Handset Lifter	41.7564	0
3	Central	#10-4 1/8" x 9 1/2" Premium Diagonal…	35.415	0
4	Central	Avery 485	35.3346	0
5	Central	Avery Personal Creations Heavyweight…	22.6184	0
6	East	AT&T CL83451 4-Handset Telephone	298.6855	0
7	East	Home/Office Personal File Carts	52.14	1
8	East	9-3/4 Diameter Round Wall Clock	40.5426	0
9	East	Plastic Binding Combs	16.362	1
10	East	Xerox 232	15.552	1
11	South	Hon Deluxe Fabric Upholstered …	219.582	0
12	South	Jet-Pak Recycled Peel 'N' Seal Padde…	62.8075	0
13	South	1.7 Cubic Foot Compact "Cube" Office…	56.2032	0
14	South	Bush Somerset Collection Bookcase	41.9136	0
15	South	Snap-A-Way Black Print Carbonless …	35.6636	0
16	West	Bretford CR4500 Series Slim …	240.2649	0
17	West	Fellowes PB200 Plastic Comb Binding …	132.5922	0
18	West	Hunt BOSTON Model 1606 High-Volume …	111.3024	0
19	West	Mitel 5320 IP Phone VoIP phone	90.7152	0
20	West	Chromcraft Rectangular Conference …	85.3091999999998	0

18. Determine the average shipping time for orders in each region, and analyze its impact on customer satisfaction based on return rates.

SQL 1*

```
    a.Avg_Shipping_Time,
    r.Return_Rate
FROM
    Avg_Shipping_Time a
JOIN
    Return_Rates r ON a."Region" = r."Region"
ORDER BY
    a."Region";
```

	Region	Avg_Shipping_Time	Return_Rate
1	Central	NULL	0.0
2	East	NULL	0.333333333333333
3	South	NULL	0.0
4	West	NULL	0.1

19. How do the sales of technology products compare to furniture products in terms of total sales, profit, and return rates?

```
select
    category,sum(sales) as total_sales,sum(profit) as total_profit,count(r.Returned)
from
    `Sample - Superstore_Orders` as o join `Sample - Superstore_Returns` as r on o.`order id`=r.`order id`
where
    category in ('Furniture','Technology')
group by category
```

	Category	total_sales	total_profit	count(r.Returned)
1	Furniture	319.41	7.09800000000001	1
2	Technology	103.584	11.5782	2

20. What is the total sales and profit for orders placed by corporate customers in the 'West' region, and how does it compare to consumer customers in the same region?

```sql
select
    segment,region,sum(sales) as total_sales,sum(profit) as profit
from
    `Sample - Superstore_Orders`
where
    segment in ('Corporate','Consumer') and region=('West')
group by segment,region
```

	Segment	Region	total_sales	profit
1	Consumer	West	5923.284	860.6323
2	Corporate	West	1473.656	203.0774

21. Identify the key factors that influence the likelihood of an order being returned, considering customer segment, region, product category, and shipping mode.

SQL 1*

```sql
        "Ship Mode",
        AVG(Returned) AS Return_Rate
    FROM
        Return_Analysis
    GROUP BY
        "Customer Segment",
        "Region",
        "Product Category",
        "Ship Mode"
)
```

	Customer Segment	Region	Product Category	Ship Mode	Return_Rate
1	Customer Segment	Region	Product Category	First Class	0.368421052631579
2	Customer Segment	Region	Product Category	Standard Class	0.0491803278688525
3	Customer Segment	Region	Product Category	Second Class	0.0