PROBLEM

As a Business Intelligence Analyst at the Superstore, you are asked to do some analysis supports from the internal BI team and from the other teams such as Marketing, Business Development, Sales, etc.

Here are some of the most urgent analysis requests from other teams. As a member of the BI Analyst who works most efficiently, you are asked to do assists to the following problems!

superstore_customer

Store the customer's data.

	customer_id [PK] character	customer_name character varying	segment character vary	country character vary	city characte	state characte	postal_code character vary	region character
1	AA-10315	Alex Avila	Consumer	United Stat...	Minne...	Minne...	55407	Central
2	AA-10375	Allen Armold	Consumer	United Stat...	Mesa	Arizona	85204	West
3	AA-10480	Andrew Allen	Consumer	United Stat...	Conc...	North ...	28027	South
4	AA-10645	Anna Andreadi	Consumer	United Stat...	Chester	Penns...	19013	East
5	AB-10015	Aaron Bergman	Consumer	United Stat...	Seattle	Washi...	98103	West
6	AB-10060	Adam Bellavan...	Home Office	United Stat...	New Y...	New Y...	10009	East
7	AB-10105	Adrian Barton	Consumer	United Stat...	Phoe...	Arizona	85023	West
8	AB-10150	Aimee Bixby	Consumer	United Stat...	Long ...	New Y...	11561	East
9	AB-10165	Alan Barnes	Consumer	United Stat...	Los A...	Califor...	90036	West
10	AB-10255	Alejandro Balle...	Home Office	United Stat...	Lorain	Ohio	44052	East
11	AB-10600	Ann Blume	Corporate	United Stat...	Tucson	Arizona	85705	West
12	AC-10420	Alyssa Crouse	Corporate	United Stat...	San Fr...	Califor...	94122	West
13	AC-10450	Amy Cox	Consumer	United Stat...	Minne...	Minne...	55407	Central
14	AC-10615	Ann Chong	Corporate	United Stat...	New Y...	New Y...	10009	East
15	AC-10660	Anna Chung	Consumer	United Stat...	Hunts...	Texas	77340	Central
16	AD-10180	Alan Dominguez	Home Office	United Stat...	Houst...	Texas	77041	Central
17	AF-10870	Art Ferguson	Consumer	United Stat...	Colleg...	Texas	77840	Central
18	AF-10885	Art Foster	Consumer	United Stat...	Louis...	Kentu...	40214	South
19	AG-10270	Alejandro Grove	Consumer	United Stat...	West ...	Utah	84084	West

superstore_product

Store the data of the product.

	product_id [PK] character varyir	product_name character varying (255)	category character va	sub_category character varyi
1	FUR-BO-10000112	Bush Birmingham Collection Boo...	Furniture	Bookcases
2	FUR-BO-10000330	Sauder Camden County Barrister ...	Furniture	Bookcases
3	FUR-BO-10000362	Sauder Inglewood Library Bookca...	Furniture	Bookcases
4	FUR-BO-10000468	O'Sullivan 2-Shelf Heavy-Duty Bo...	Furniture	Bookcases
5	FUR-BO-10000711	Hon Metal Bookcases, Gray	Furniture	Bookcases
6	FUR-BO-10000780	O'Sullivan Plantations 2-Door Libr...	Furniture	Bookcases
7	FUR-BO-10001337	O'Sullivan Living Dimensions 2-Sh...	Furniture	Bookcases
8	FUR-BO-10001519	O'Sullivan 3-Shelf Heavy-Duty Bo...	Furniture	Bookcases
9	FUR-BO-10001567	Bush Westfield Collection Bookca...	Furniture	Bookcases
10	FUR-BO-10001601	Sauder Mission Library with Door...	Furniture	Bookcases
11	FUR-BO-10001608	Hon Metal Bookcases, Black	Furniture	Bookcases
12	FUR-BO-10001619	O'Sullivan Cherrywood Estates Tr...	Furniture	Bookcases
13	FUR-BO-10001798	Bush Somerset Collection Bookc...	Furniture	Bookcases
14	FUR-BO-10001811	Atlantic Metals Mobile 5-Shelf Bo...	Furniture	Bookcases
15	FUR-BO-10001918	Sauder Forest Hills Library with D...	Furniture	Bookcases
16	FUR-BO-10001972	O'Sullivan 4-Shelf Bookcase in Od...	Furniture	Bookcases
17	FUR-BO-10002202	Atlantic Metals Mobile 2-Shelf Bo...	Furniture	Bookcases
18	FUR-BO-10002206	Bush Saratoga Collection 5-Shelf ...	Furniture	Bookcases
19	FUR-BO-10002213	DMI Eclipse Executive Suite Book...	Furniture	Bookcases

superstore_order

Store the data of the customer's order.

	order_id character	order_date date	ship_date date	ship_mode character varying	customer_id character var	product_id character va	sales numeric	quantity integer	discount numeric	profit numeric
1	CA-201...	2016-11-08	2016-11-11	Second Class	CG-12520	FUR-BO-...	261.96	2	0.0	41.9136
2	CA-201...	2016-11-08	2016-11-11	Second Class	CG-12520	FUR-CH-...	731.9399999999999	3	0.0	219.58199999999997
3	CA-201...	2016-06-12	2016-06-16	Second Class	DV-13045	OFF-LA-1...	14.62	2	0.0	6.8713999999999995
4	US-201...	2015-10-11	2015-10-18	Standard Class	SO-20335	FUR-TA-1...	957.5775	5	0.45	-383.03100000000006
5	US-201...	2015-10-11	2015-10-18	Standard Class	SO-20335	OFF-ST-1...	22.368000000000002	2	0.2	2.516399999999999
6	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	FUR-FU-1...	48.86	7	0.0	14.169399999999996
7	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	OFF-AR-...	7.28	4	0.0	1.9656000000000002
8	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	TEC-PH-...	907.152	6	0.2	90.71520000000004
9	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	OFF-BI-1...	18.504	3	0.2	5.7825
10	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	OFF-AP-...	114.9	5	0.0	34.46999999999999
11	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	FUR-TA-1...	1706.1840000000002	9	0.2	85.3091999999998
12	CA-201...	2014-06-09	2014-06-14	Standard Class	BH-11710	TEC-PH-...	911.424	4	0.2	68.35680000000002
13	CA-201...	2017-04-15	2017-04-20	Standard Class	AA-10480	OFF-PA-1...	15.552000000000003	3	0.2	5.4432
14	CA-201...	2016-12-05	2016-12-10	Standard Class	IM-15070	OFF-BI-1...	407.97600000000006	3	0.2	132.59219999999993
15	US-201...	2015-11-22	2015-11-26	Standard Class	HP-14815	OFF-AP-...	68.80999999999999	5	0.8	-123.858
16	US-201...	2015-11-22	2015-11-26	Standard Class	HP-14815	OFF-BI-1...	2.5439999999999996	3	0.8	-3.8160000000000016
17	CA-201...	2014-11-11	2014-11-18	Standard Class	PK-19075	OFF-ST-1...	665.88	6	0.0	13.317599999999999
18	CA-201...	2014-05-13	2014-05-15	Second Class	AG-10270	OFF-ST-1...	55.5	2	0.0	9.989999999999995
19	CA-201...	2014-08-27	2014-09-01	Second Class	ZD-21925	OFF-AR-...	8.56	2	0.0	2.4823999999999993

CASE 1

SAME DAY Ship Mode service is a service where the product ordered by the customer can be sent directly on the same day as the day of order. But in reality, not all customers who order receive the benefits of this service well. In other words, there are also some SAME DAY orders that are not delivered on the same day as the order day. The Operations Team would like to analyze this problem to be followed up. You are asked to display the number of SAME DAY orders which is experiencing delays in delivery.

SOLUTION

Query Editor | Query History

```
1  SELECT
2      COUNT(1) AS count_late_shipped
3  FROM
4      superstore_order
5  WHERE
6      ship_mode = 'Same Day' AND
7      order_date != ship_date
```

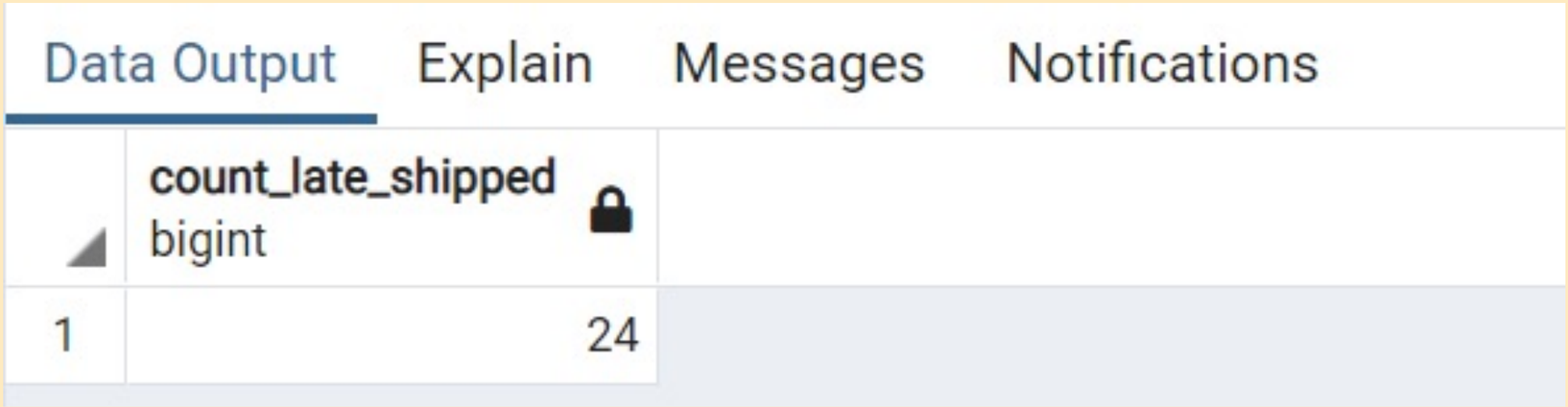
Counts the total rows in the table

Rename the column using AS

Table's name

FIlter the record where ship_mode is 'Same Day' and order_date is not the same as ship_date

Query Result



Data Output | Explain | Messages | Notifications

	count_late_shipped bigint
1	24

There are 24 SAME DAY orders which is experiencing delays in delivery.

CASE 2

The Business team would like to conduct further analysis of the company's profitability. This time, they wants to see the relationship between the amount of discount given and profitability received by the company. You are asked to display this relationship with shows the average profit for each discount level, where the discount level criteria are as follows:

- LOW if the discount is below 0.2 (excluding 0.2),
-MODERATE if the discount starts from 0.2 to below 0.4 (excluding 0.4)
- HIGH if the discount starts from 0.4 and above.

SOLUTION

Query Editor | Query History

```sql
SELECT
    CASE
        WHEN discount < 0.2 THEN 'LOW'
        WHEN discount >= 0.4 THEN 'HIGH'
        ELSE 'MODERATE'
    END level_discount,
    AVG(profit) AS average_profit
FROM
    superstore_order
GROUP BY 1
ORDER BY 2 DESC
```

Goes through conditions and returns a value when the condition is met

Aggregate function to returns the average value of profit column

Group the aggregation with the first column (level_discount)

Order the records with the second column (average_profit) in decending

Query Result

Data Output | Explain | Messages | Notifications

	level_discount text	average_profit numeric
1	LOW	67.03797971278316737
2	MODERATE	19.835564024546144 75
3	HIGH	-107.65201132572433 39

The higher the discount level, the lower the average profit

CASE 3

The Sales Team asked the Business Intelligence Analyst to analyze the performance of Category and Subcategory of products owned by the company. You are asked to returns the following metrics for each of the existing Category–Subcategory pairs:

- Average discount
- Average profit

Don't forget to display the full Category and Subcategory names instead of just display only the Product ID to make it easier for the Sales Team to understand the results of your analysis!

SOLUTION

Query Editor | Query History

```sql
SELECT
    p.category,
    p.sub_category,
    AVG(o.discount) AS average_discount,
    AVG(o.profit) AS average_profit
FROM
    superstore_order o
LEFT JOIN
    superstore_product p
    ON o.product_id = p.product_id
GROUP BY 1,2
ORDER BY 1,2
```

Use a prefix (p._) to indicate the reference table

Give the alternative name of the table

Join with superstore_product table

Match the key column between two tables

Query Result

Data Output

	category character varying (25)	sub_category character varying (25)	average_discount numeric	average_profit numeric
1	Furniture	Bookcases	0.21114035087719298246	-15.2305087719298380
2	Furniture	Chairs	0.17017828200972447326	43.0958935170178139
3	Furniture	Furnishings	0.13834900731452455590	13.64591807732497215
4	Furniture	Tables	0.26128526645768025078	-55.5657714733542497
5	Office Supplies	Appliances	0.16652360515021459227	38.92275836909870306
6	Office Supplies	Art	0.07487437185929648241	8.20073743718592866
7	Office Supplies	Binders	0.37229152987524622456	19.8435740643466793
8	Office Supplies	Envelopes	0.08031496062992125984	27.4180185039370051
9	Office Supplies	Fasteners	0.08202764976958525346	4.37565990783410111
10	Office Supplies	Labels	0.06861318681318681320	15.2369615384615381
11	Office Supplies	Paper	0.07489051094890510949	24.8566199270072976
12	Office Supplies	Storage	0.07470449172576832151	25.15227706855791304
13	Office Supplies	Supplies	0.07684210526315789474	-6.25841842105263644
14	Technology	Accessories	0.07845161290322580645	54.11178799999999635
15	Technology	Copiers	0.16176470588235294118	817.9091897058822724
16	Technology	Machines	0.30608695652173913043	29.4326686956520713
17	Technology	Phones	0.15455568053993250844	50.07393768278964770

Query Result

Order by average discount and give limitation to 5 to see the top 5 category and sub-category pairs that have the biggest average discount

```
12  ORDER BY 3 DESC
13  LIMIT 5
```

Result:

Data Output

	category character varying (25)	sub_category character varying (25)	average_discount numeric	average_profit numeric
1	Office Supplies	Binders	0.37229152987524622456	19.8435740643466793
2	Technology	Machines	0.30608695652173913043	29.4326686956520713
3	Furniture	Tables	0.26128526645768025078	-55.5657714733542497
4	Furniture	Bookcases	0.21114035087719298246	-15.2305087719298380
5	Furniture	Chairs	0.17017828200972447326	43.0958935170178139

The highest average discount, which is around 0.37 is owned by the BINDERS subcategory of the OFFICE SUPPLIES category

Query Result

Order by average profit and give limitation to 5 to see the top 5 category and sub-category pairs that have the biggest average profit

```
12  ORDER BY 4 DESC
13  LIMIT 5
```

Result:

Data Output

	category character varying (25)	sub_category character varying (25)	average_discount numeric	average_profit numeric
1	Technology	Copiers	0.16176470588235294118	817.9091897058822724
2	Technology	Accessories	0.07845161290322580645	54.11178799999999635
3	Technology	Phones	0.15455568053993250844	50.07393768278964770
4	Furniture	Chairs	0.17017828200972447326	43.0958935170178139
5	Office Supplies	Appliances	0.16652360515021459227	38.92275836909870306

The TECHNOLOGY category seems to dominate the highest average profit with the COPIERS, ACCESSORIES and PHONES subcategories

CASE 4

The Business Development team is considering further expansion in California, Texas and also Georgia. As material for their consideration, you are asked to display the performance of each of the Customer Segments in that states on 2016 only. The requested performance metrics are as follows:

- Total sales
- Average profit

SOLUTION

Query Editor | Query History

```sql
SELECT
    c.segment,
    SUM(o.sales) AS total_sales,
    AVG(o.profit) AS average_profit
FROM
    superstore_order o
LEFT JOIN
    superstore_customer c
    ON o.customer_id = c.customer_id
WHERE
    extract(year from o.order_date) = 2016 AND
    c.state IN('California', 'Texas', 'Georgia')
GROUP BY 1
ORDER BY 2 DESC
```

Extract the year value from order_date

Shorthand for multiple OR conditions

Query Result

Data Output | Explain | Messages | Notifications

	segment character varying (25)	total_sales numeric	average_profit numeric
1	Consumer	90982.3196000000012181	30.32965562913906877
2	Corporate	50951.9110000000016284	33.573490566037731 46
3	Home Office	34897.9529999999997967	34.66199395973154455

The HOME OFFICE segment has the highest average profit but the lowest total sales from other segments

CASE 5

The Business team is interested to see which region has the most number of customers who loves discounts. Therefore, the Business Team asks you as a Business Intelligence Analyst to display the number of people/customers who have an average discount above 0.4 for each existing region.

STEP 1

Query Editor | Query History

```
1  WITH
2  temp AS (
3      SELECT
4          customer_id,
5          AVG(discount)
6      FROM
7          superstore_order
8      GROUP BY 1
9      HAVING AVG(discount) > 0.4
10 )
```

Create a temporary table that can be used later

This temporary table (temp) is used to save the records of customer_id that have average discount above 0.4

Filter the result of aggregation

Data Output | Explain | Messages | Notifications

	customer_id [PK] character varying (25)	avg numeric
1	GH-14485	0.46666666666666666667
2	LH-17020	0.46666666666666666667
3	TS-21085	0.70000000000000000000
4	MG-18205	0.45000000000000000000
5	AG-10765	0.46000000000000000000
6	CD-11980	0.43333333333333333333
7	RH-19555	0.53333333333333333333
8	VS-21820	0.47500000000000000000
9	TP-21415	0.46200000000000000000

Query Result

There are 9 customers who have average discount above 0.4

STEP 2

Query Editor | Query History

```
12  SELECT
13      c.region,
14      COUNT(1)
15  FROM
16      temp
17  JOIN
18      superstore_customer c
19      ON temp.customer_id = c.customer_id
20  GROUP BY 1
```

After finding the customer_id and save it in the temp table, combine it with the customer table and count the number of occurrences of each regions

Query Result

Data Output | Explain | Messages | Notifications

	region character varying (10)	count bigint
1	South	2
2	West	3
3	East	2
4	Central	2

The WEST Region has the most customers who have average discount above 0.4