--preview dataset

select*from sample s

limit 10;

•	123 row_id 🔻	A-z order_id 🔻	⊘ order_date ▼	⊘ ship_date ▼	A-z ship_mode 🔻	A-z customer_id	A-z customer_
1	1	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute
2	2	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute
3	3	CA-2016-138688	2016-06-12	2016-06-16	Second Class	DV-13045	Darrin Van Hu
4	4	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donne
5	5	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donne
6	6	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffn
7	7	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffn
8	8	CA-2014-115812	2014-06-09	2014-06-14	Standard Class	BH-11710	Brosina Hoffn
Q	q	CΔ-2014-115812	2014-06-09	2014-06-14	Standard Class	RH-11710	Rrosina Hoffn

as: 10 data teratas

- --Check Data Quality
- --Check Missing Values

select*from sample s

where row_id is null or order_id is null or customer_id is null or sales is null;--artinya tidak ada data yang null di variabel tsb

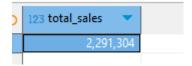
)	123 row_id	•	A-z order_id	•	Order_date	•	// ship_date	•	A-z ship_mode	•	A-z customer_id	•	A-z customer_nai
ı													
ı													
ı													
-													

as: kalo data kosongan ini itu ,tidak ada yang null yah.

- --exploration data analyst (EDA)
- --1.Berapa total sales revenueu?

select sum(sales) as total_sales

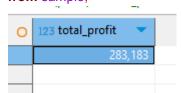
from sample;



--2. Berapa total profit yang didapatkan?

select sum(profit) as total_profit

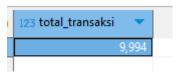
from sample;



--3. Berapa jumlah transaksi yang tercatat ?.

select count(order_id) as total_transaksi

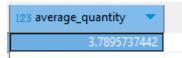
from sample;



--4. Berapa jumlah rata-rata quantity penjualan?

select avg(quantity) as average_quantity

from sample;



--5. Menampilkan produk yang paling banyak di order? select category, sum(quantity) as total_order_quantity from sample s group by category;



select sub_category, **sum** (quantity) **as** *total_order_quantity* **from** sample *s*

and the same

group by sub_category

order by total_order_quantity desc

limit 1;



--6.ranking penjualan berdasarkan customer id (top 5 sales)

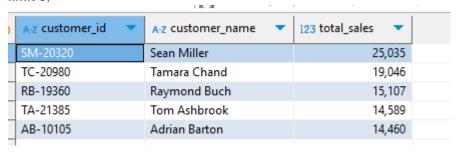
select customer_id, customer_name, **sum**(sales) **as** *total_sales*

from sample s

group by customer_id, customer_name

order by total_sales desc

limit 5;



--7. ranking customer berdasarkan salesnya dengan opartisi segmentnya:

with ranking_customer as (select customer_id, segment, sum(sales) as total_sales

from sample

select customer_id, segment, total_sales,

rank() over(partition by segment order by total_sales desc) as ranked

from ranking_customer; A-z customer_id A-z segment 123 total_sales 123 ranked RB-19360 15,107 1 Consumer 14,460 2 AB-10105 Consumer KL-16645 Consumer 3 14,157 SC-20095 Consumer 14,127 4 5 12,865 HL-15040 Consumer SE-20110 Consumer 12,198 6 CC-12370 Consumer 12,122 7 GT-14710 Consumer 11,803 8 9 BM-11140 Consumer 11,778 CV/ 20265 Cancumar 11 //51 10

1. Buatlah monthly sales trend.

```
SELECT
EXTRACT(YEAR FROM order_date) AS year,
EXTRACT(MONTH FROM order_date) AS month,
SUM(sales) AS total_sales
FROM
sample s
GROUP BY
EXTRACT(YEAR FROM order_date),
EXTRACT(MONTH FROM order_date)
ORDER BY
year,
month;
```

0	123 year 🔻	123 month	123 total_sales 🔻
1	2,014	1	14,188
2	2,014	2	4,492
3	2,014	3	55,602
4	2,014	4	28,215
5	2,014	5	23,578
6	2,014	6	34,519
7	2,014	7	33,859
8	2,014	8	27,820
9	2,014	9	81,615
10	2,014	10	31,357
11	2,014	11	78,434
12	2,014	12	69,384
13	2,015	1	18,145
14	2,015	2	11,909
15	2,015	3	38,642
16	2,015	4	34,098
17	2,015	5	30,046
18	2,015	6	24,717
19	2,015	7	28,682
20	2,015	8	36,810

2.Identifikasi customer yang telah melakukan setidaknya lima pembelian dan hitung nilai rata-rata pesanan customer.

```
SELECT customer_name, COUNT(*) AS TotalOrders, ROUND(AVG(sales), 2) AS

AverageOrderValue

FROM sample s

GROUP BY customer_name

HAVING COUNT(*) >= 5

ORDER BY TotalOrders, AverageOrderValue DESC;
```

•	A-Z customer_name 🔻	123 totalorders	123 averageordervalue
1	Kelly Collister	5	781.2
2	Robert Dilbeck	5	556.8
3	George Zrebassa	5	490.4
4	Bobby Elias	5	451.8
5	Justin Hirsh	5	450.8
6	Denny Joy	5	401.8
7	Pamela Coakley	5	366
8	Georgia Rosenberg	5	256.2
9	Liz Willingham	5	251
10	Elpida Rittenbach	5	248.6
11	Claire Gute	5	229.2
12	Alex Russell	5	210.4
13	Victor Preis	5	198.4
14	Alyssa Crouse	5	184.6
15	Mary O'Rourke	5	184
16	Sally Knutson	5	176
17	David Kendrick	5	159
		-	

3. Temukan 5 pelanggan teratas yang telah melakukan total penjualan tertinggi di setiap negara bagian, bersama dengan kategori produk yang paling banyak mereka beli. Gunakan CTE

```
WITH CustomerSales AS (
SELECT
customer_name,
state,
category,
SUM(sales) AS Total_Sales,
RANK() OVER (PARTITION BY state ORDER BY SUM(sales) DESC) AS Sales_Rank
FROM sample
GROUP BY customer_name, state, category
)
SELECT
customer_name,state,
category AS Predominant Product Category, Total_Sales,Sales_Rank
FROM CustomerSales
WHERE Sales_Rank <= 5
ORDER BY state, Total_Sales DESC;
```

•	A-z customer_name	A-z state 🔻	A-z predominant_product_category •	123 total_sales 🔻	123 sales_rank 🔻
1	Mark Cousins	Alabama	Technology	3,040	1
2	Natalie Webber	Alabama	Furniture	1,857	
3	Irene Maddox	Alabama	Furniture	1,819	
4	Alan Schoenberger	Alabama	Technology	1,394	
5	Karen Carlisle	Alabama	Technology	1,319	5
6	John Murray	Arizona	Technology	1,879	
7	Arthur Gainer	Arizona	Technology	1,502	2
8	Tanja Norvell	Arizona	Furniture	1,295	
9	Maris LaWare	Arizona	Furniture	1,273	4
10	Brendan Sweed	Arizona	Office Supplies	1.113	5