ONLINE RETAIL-PROJECT SQL

ARIANTY SIAHAAN

DOWNLOAD DATA HERE:

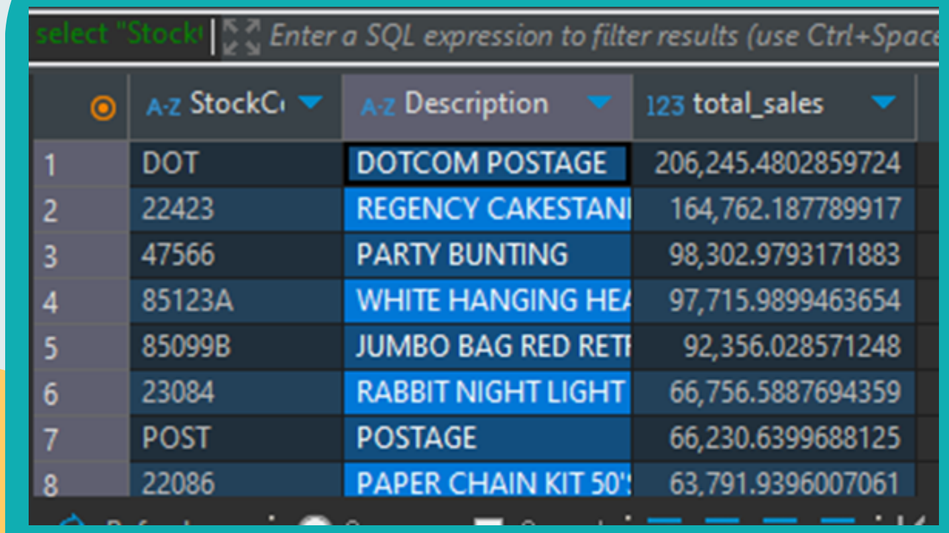
https://www.kaggle.com/datasets/thedevastator/online-retail-transaction-data/data

DATA ONLINE RETAIL KAGGLE

1. DISPLAYING TOTAL SALES PER PRODUCT

Question: What is the total sales amount for each product?

```sql
select
	"StockCode" ,
	"Description" ,
	sum("Quantity"*"UnitPrice") as total_sales
from
	online_retail
group by
	"StockCode" , "Description"
order by
	total_sales desc;
```



2. DISPLAYING THE TOP 5 CUSTOMERS BY TOTAL SPEND

Question: Which customers spent the most money on the online retail store?

```sql
select
	"CustomerID",
	sum("Quantity"*"UnitPrice") as total_spend
from
	online_retail or2
group by
	"CustomerID"
order by
	total_spend desc
limit 5;
```

	CustomerID	total_spend
1		1,447,682.1183872335
2	14646	279,489.0194430053
3	18102	256,438.4883605242
4	17450	187,482.1709718704
5	14911	132,572.6192782

3. DISPLAYING THE NUMBER OF ORDERS AND TOTAL SALES BY MONTH

Question: How many orders and total sales were generated each month?

```sql
select
	date_trunc('month', "InvoiceDate") as month,
	count(distinct "InvoiceNo") as number_of_orders,
	sum("Quantity"*"UnitPrice") as total_sales
from
	online_retail or2
group by
	month
order by
	month;
```

	month	number_of_orders	total_sales
1	2010-12-01 00:00:00.000	2,025	748,957.0175240934
2	2011-01-01 00:00:00.000	1,476	560,000.2572612464
3	2011-02-01 00:00:00.000	1,393	498,062.648327291
4	2011-03-01 00:00:00.000	1,983	683,267.0775260255
5	2011-04-01 00:00:00.000	1,744	493,207.1188206653
6	2011-05-01 00:00:00.000	2,162	723,333.5063029006
7	2011-06-01 00:00:00.000	2,012	691,123.1159918569
8	2011-07-01 00:00:00.000	1,927	681,300.1073726917
9	2011-08-01 00:00:00.000	1,737	682,680.5073843151

4. FIND THE MOST POPULAR PRODUCT

```sql
select
	"StockCode",
	"Description",
	sum("Quantity") as total_units_sold
from
	online_retail or2
group by
	"StockCode", "Description"
order by
	total_units_sold desc
limit 1;
```

Question: What is the most popular product based on the number of units sold?

	StockCode	Description	total_units_so
1	84077	WORLD WAR 2 GLIDEF	53,847

5. FIND THE CUSTOMER WITH THE HIGHEST PURCHASE IN A SINGLE TRANSACTION

Question: Which customer made the highest purchase in a single transaction?

```
select
	"CustomerID",
	"InvoiceNo",
	sum("Quantity" * "UnitPrice") as total_transaction_value
from
	online_retail or2
group by
	"CustomerID", "InvoiceNo"
order by
	total_transaction_value	desc
limit 1;
```

A-Z CustomerID	A-Z InvoiceNo	123 total_transacti
16446	581483	168,469.5938205719

6. DISPLAY THE NUMBER OF TRANSACTIONS FOR EACH PRODUCT

Question: How many transactions were made for each product?

```sql
select
	"StockCode",
	"Description",
	count(distinct "InvoiceNo") as number_of_transactions
from
	online_retail or2
group by
	"StockCode", "Description"
order by
	number_of_transactions desc;
```

	StockCode	Description	number_of_tr
1	85123A	WHITE HANGING HEA	2,235
2	22423	REGENCY CAKESTAND	2,169
3	85099B	JUMBO BAG RED RETR	2,135
4	47566	PARTY BUNTING	1,706
5	20725	LUNCH BAG RED RETR	1,607
6	84879	ASSORTED COLOUR B	1,467
7	22720	SET OF 3 CAKE TINS PA	1,458
8	21212	PACK OF 72 RETROSP(	1,334

7. FIND THE TOTAL NUMBER OF PRODUCTS SOLD BY EACH CUSTOMER

Question: How many different products did each customer purchase?

```sql
select
	"CustomerID",
	count(distinct "StockCode") as products_purchased
from
	online_retail or2
group by
	"CustomerID"
order by
	products_purchased desc;
```

	CustomerID	products_purchased
1		3,810
2	14911	1,794
3	12748	1,769
4	17841	1,331
5	14096	1,121
6	14298	884
7	14606	832
8	14769	718

8. SHOW THE FOLLOWING TABLE: INVOICEDATE, CUSTOMER ID, UNIT PRICE, GROSS PROFIT PERCENTAGE, NET SALES, NET PROFIT

a. Add a Discount Percentage column with the following conditions:

- From April 1, 2010, to December 31, 2010, set the discount percentage to 0.3.
- From January 1, 2011, to December 31, 2011, set the discount percentage to 0.35.
- For all other dates, set the discount percentage to 0.4.

b. Create a column for InvoiceDate, Unit Price, and Gross Profit Percentage:

- Categorize the data based on quintiles for the Unit Price, with the following quintile percentages for gross profit: 0.1, 0.15, 0.2, and 0.25.

c. Add a variable for Net Sales using the formula:

- Net Sales = Unit Price * (1 - Discount Percentage).

d. Add a variable for Net Profit using the formula:
Net Profit = Unit Price * (1 - Discount Percentage) * Gross Profit.

- Sort the data in descending order based on InvoiceDate, from the most recent to the oldest.

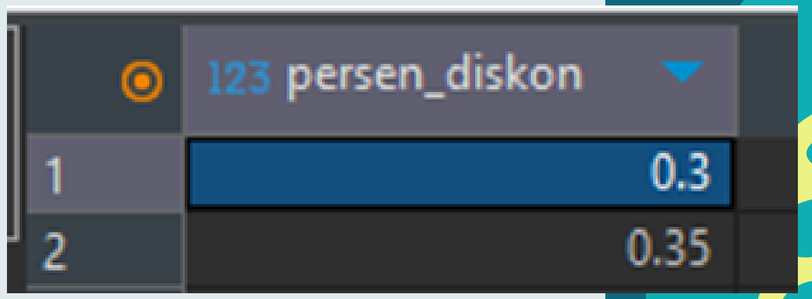
8A. ADD A DISCOUNT PERCENTAGE COLUMN WITH THE FOLLOWING CONDITIONS:

```sql
--menambahkan kolom persen_diskon
alter table online_retail
add column persen_diskon decimal (5,2);

--mengupdate isian kolom persen_diskon
update online_retail
set persen_diskon = case
WHEN "InvoiceDate" BETWEEN '2010-04-01' AND '2010-12-31' THEN 0.30
    WHEN "InvoiceDate" BETWEEN '2011-01-01' AND '2011-12-31' THEN 0.35
    ELSE 0.40
end;

--melihat kolom persen diskon dengan pivot
select "persen_diskon"
from online_retail or2
group by persen_diskon ;

--melihat tabel invoice date dan persen diskon
select "InvoiceDate", "persen_diskon"
from online_retail or2 ;
```



	123 persen_diskon
1	0.3
2	0.35

	InvoiceDate	123 persen_diskon
1	2010-12-01 09:41:00.000	0.3
2	2010-12-01 09:41:00.000	0.3
3	2010-12-01 09:41:00.000	0.3
4	2010-12-01 09:41:00.000	0.3
5	2010-12-01 09:41:00.000	0.3
6	2010-12-01 09:41:00.000	0.3
7	2010-12-01 09:41:00.000	0.3
8	2010-12-01 09:41:00.000	0.3
9	2010-12-01 09:41:00.000	0.3
10	2010-12-01 09:41:00.000	0.3
11	2010-12-01 09:41:00.000	0.3

8B. CREATE A COLUMN FOR INVOICEDATE, UNIT PRICE, AND GROSS PROFIT PERCENTAGE

```
select or2."InvoiceDate", or2."UnitPrice",
case
	when OR2."UnitPrice" <=1.25 then 0.1
	when or2."UnitPrice" >1.25  and or2."UnitPrice" <=2.08 then 0.15
	when or2."UnitPrice" >2.08 and or2."UnitPrice" <=4.13 then 0.2
	else 0.25
end as persentase_gross_laba
from online_retail or2 ;
```

	InvoiceDate	UnitPrice	persentase_gross_laba
1	2011-11-03 13:24:00.000	18	0.25
2	2011-11-03 14:28:00.000	3.29	0.2
3	2011-11-03 15:01:00.000	1.45	0.15
4	2011-11-03 16:15:00.000	28	0.25
5	2011-11-03 16:30:00.000	5.75	0.25
6	2011-11-03 18:49:00.000	0.39	0.1
7	2011-11-04 11:35:00.000	0	0.1
8	2011-11-04 11:41:00.000	0.85	0.1
9	2011-11-04 12:45:00.000	0.85	0.1
10	2011-11-04 13:15:00.000	0.21	0.1
11	2011-11-04 13:26:00.000	2.1	0.2
12	2011-11-04 15:29:00.000	1.65	0.15

8C. ADD A VARIABLE FOR NET SALES USING THE FORMULA

```
select or2."InvoiceDate", or2."CustomerID", or2."UnitPrice",
case
	when OR2."UnitPrice" <=1.25 then 0.1
	when or2."UnitPrice" >1.25  and or2."UnitPrice" <=2.08 then 0.15
	when or2."UnitPrice" >2.08 and or2."UnitPrice" <=4.13 then 0.2
	else 0.25
end as persentase_gross_laba,
	(or2."UnitPrice"*(1- or2."persen_diskon")) as penjualan_bersih
	from
online_retail or2;
```

	InvoiceDate	CustomerID	UnitPrice	persentase_gross	penjualan_bersih
1	2011-09-07 15:07:00.000	12637	1.65	0.15	1.0724999845
2	2011-09-08 10:09:00.000		1.45	0.15	0.942500031
3	2011-09-08 12:42:00.000	13137	2.1	0.2	1.364999938
4	2011-09-08 13:38:00.000	12540	28	0.25	18.2
5	2011-09-08 15:27:00.000		4.13	0.25	2.6845000744
6	2011-09-09 10:51:00.000	12674	0.42	0.1	0.2729999915
7	2011-09-09 12:02:00.000		116.69	0.25	75.8485015869
8	2011-09-11 12:23:00.000	13324	0.42	0.1	0.2729999915
9	2011-09-13 11:39:00.000	13136	4.95	0.25	3.217499876
10	2011-09-13 12:32:00.000	14096	1.63	0.15	1.0594999969
11	2011-09-14 09:38:00.000	14911	0.39	0.1	0.2534999907
12	2011-09-14 14:19:00.000	12726	0.42	0.1	0.2729999915

8D. ADD A VARIABLE FOR NET PROFIT USING THE FORMULA

```
select or2."InvoiceDate", or2."CustomerID", or2."UnitPrice",
case
	when OR2."UnitPrice" <=1.25 then 0.1
	when or2."UnitPrice" >1.25  and or2."UnitPrice" <=2.08 then 0.15
	when or2."UnitPrice" >2.08 and or2."UnitPrice" <=4.13 then 0.2
	else 0.25
end as persentase_gross_laba,
	(or2."UnitPrice"*(1- or2."persen_diskon")) as penjualan_bersih,
	("UnitPrice" * (1- "persen_diskon")*
case
	when OR2."UnitPrice" <=1.25 then 0.1
	when or2."UnitPrice" >1.25  and or2."UnitPrice" <=2.08 then 0.15
	when or2."UnitPrice" >2.08 and or2."UnitPrice" <=4.13 then 0.2
	else 0.25
	end) as keuntungan_bersih
from
online_retail or2
order by "InvoiceDate" DESC;
```

	InvoiceDate	Cust	UnitPrice	persentase_gross	penjualan_bersih	keuntungan_ber
1	2011-12-09 12:50:00.000	12680	4.95	0.25	3.217499876	0.804374969
2	2011-12-09 12:50:00.000	12680	4.15	0.25	2.697500062	0.6743750155
3	2011-12-09 12:50:00.000	12680	2.1	0.2	1.364999938	0.2729999876
4	2011-12-09 12:50:00.000	12680	0.85	0.1	0.5525000155	0.0552500015
5	2011-12-09 12:50:00.000	12680	4.15	0.25	2.697500062	0.6743750155
6	2011-12-09 12:50:00.000	12680	1.95	0.15	1.267500031	0.1901250046
7	2011-12-09 12:50:00.000	12680	1.95	0.15	1.267500031	0.1901250046

THANK YOU

Linkedin:
Arianty Siahaan

Portfolio:
https://linktr.ee/AriantySiahaan