-- Create table

CREATE TABLE IF NOT EXISTS sales(

invoice\_id VARCHAR(30) NOT NULL PRIMARY KEY,

branch VARCHAR(5) NOT NULL,

city VARCHAR(30) NOT NULL,

customer\_type VARCHAR(30) NOT NULL,

gender VARCHAR(30) NOT NULL,

product\_line VARCHAR(100) NOT NULL,

unit\_price DECIMAL(10,2) NOT NULL,

quantity INT NOT NULL,

tax\_pct FLOAT(6,4) NOT NULL,

total DECIMAL(12, 4) NOT NULL,

date DATETIME NOT NULL,

time TIME NOT NULL,

payment VARCHAR(15) NOT NULL,

cogs DECIMAL(10,2) NOT NULL,

gross\_margin\_pct FLOAT(11,9),

gross\_income DECIMAL(12, 4),

rating FLOAT(2, 1)

);

-----------------------------------------------------------------------------------------------------

------------------------------------- feature engineering -----------------------------------------

select \* from sales;

----------------------------- adding new columns --------------------------

----- adding new time column labelled as morning,afternoon , evening ------------

select time,time\_of\_day from sales;

alter table sales

add column time\_of\_day varchar(10);

update sales

set time\_of\_day = case

when hour(time\_sold)>=00 and hour(time\_sold)<12 then 'Morning '

when hour(time\_sold) >=12 and hour(time\_sold) <16 then 'Afternoon'

when hour(time\_sold) >=16 then 'Evening'

end;

select time\_sold , time\_of\_day from sales;

select \* from sales;

---- dayname column -----------------------

alter table sales

add column name\_of\_the\_day varchar(10);

UPDATE sales

SET name\_of\_the\_day = DAYNAME(date);

select date , name\_of\_the\_day from sales;

-------------------------- month name column ----------------

alter table sales

add column name\_of\_the\_month varchar(10);

UPDATE sales

SET name\_of\_the\_month = monthname(date);

select date , name\_of\_the\_month from sales;

select \* from sales;

## Business Questions To Answer

### Generic Question

-------------- 1. How many unique cities does the data have? ------------------

select distinct(city) from sales;

select count(distinct(city)) from sales;

-------------- 2. In which city is each branch? ---------------------

select city , branch

from sales

group by city,branch

order by city;

----- ### Product related questions ------------

-- 1. How many unique product lines does the data have?--

select distinct(product\_line) from sales;

select count(distinct(product\_line)) from sales;

-- 2. What is the most common payment method? --

select \* from

(

select payment , count(payment) as num\_of\_times\_trans ,

dense\_rank () over( order by count(payment) desc) as ranking

from sales

group by payment

) as s

where ranking=1;

-- 3. What is the most selling product line?---

select product\_line,counting from

(

select product\_line , count(product\_line) as counting ,

dense\_rank () over( order by count(product\_line) desc) as ranking

from sales

group by product\_line

) as s

where ranking=1;

-- 4. What is the total revenue by month? --

select unit\_price,quantity,cogs,tax\_pct,total,gross\_income,gross\_margin\_pct from sales;

select date from sales;

select name\_of\_the\_month,sum(total)

from sales

group by name\_of\_the\_month

order by name\_of\_the\_month;

-- 5. What month had the largest COGS? ------

select name\_of\_the\_month , sum(cogs) as sum\_of\_cogs

from sales

group by name\_of\_the\_month

order by sum\_of\_cogs desc;

-- 6. What product line had the largest revenue?

select product\_line,sum(total) as tot\_rev

from sales

group by product\_line

order by tot\_rev desc;

-- 7. What is the city with the largest revenue? --

select city,sum(total) as tot\_rev

from sales

group by city

order by tot\_rev desc;

-- 8. What product line had the largest VAT collection? ----

select product\_line, sum(tax\_pct) as tot\_tax

from sales

group by product\_line

order by tot\_tax desc;

-- 9. Fetch each product line and add a column to those product line

----- showing "Good", "Bad". Good if its greater than average sales

select product\_line, (select avg(total) from sales) as tot\_avg, avg(total) as avg\_of\_pro\_line,

case

when avg(total) >( select avg(total) from sales) then 'GOOD'

else 'BAD'

end as good\_bad\_category

from sales

group by product\_line;

----- 10. Which branch sold more products than average product sold? ------------

select branch , sum(quantity) as qt\_sold\_in\_this\_branch,

(select sum(quantity)/count(distinct branch) from sales) as avg\_qt\_sold\_at\_all\_branches

from sales

group by branch

having sum(quantity) >(select sum(quantity)/count(distinct branch) from sales);

----- 11. What is the most common product line liked by each gender?---------

select gender , product\_line , product\_purchased from

(

select gender , product\_line , sum(quantity) as product\_purchased,

dense\_rank() over( partition by gender order by sum(quantity) desc) as rnk

from sales

group by gender , product\_line

) as s

where rnk=1;

-------- 12. What is the average rating of each product line?-------

select product\_line , avg(rating) as avg\_rating

from sales

group by product\_line

order by avg\_rating desc;

### Sales

-------------- 1. Number of qt. and amount of sales made in each time of the day per weekday ------

select name\_of\_the\_day,time\_of\_day,sum(quantity),sum(total) ,

dense\_rank() over(partition by name\_of\_the\_day order by time\_of\_day desc)

from sales

group by name\_of\_the\_day,time\_of\_day;

-------------- 2. Which of the customers spend the most amount of money and

-------------- which customers purchase the most no. of goods quantity wise ? --------

select distinct(customer\_type)

from sales;

select customer\_type , sum(total) as tot\_amt

from sales

group by customer\_type

order by tot\_amt desc;

select customer\_type ,product\_line, sum(total) as tot\_amt

from sales

group by customer\_type,product\_line

order by customer\_type,tot\_amt ;

select customer\_type , sum(quantity) as tot\_cnt

from sales

group by customer\_type

order by tot\_cnt desc;

select customer\_type ,product\_line, sum(quantity) as tot\_cnt

from sales

group by customer\_type,product\_line

order by customer\_type,tot\_cnt ;

-- 3. Which city has the largest tax percent/ VAT (\*\*Value Added Tax\*\*)?----------

select city , sum(tax\_pct) as vat

from sales

group by city

order by vat desc;

-------- 4. Which customer type pays the most in VAT? -------

select customer\_type , sum(tax\_pct) as vat

from sales

group by customer\_type

order by vat desc;

### Customer

----------------------- 1. How many unique customer types does the data have?

select distinct(customer\_type)

from sales;

select count(distinct(customer\_type))

from sales;

----------- 2. How many unique payment methods does the data have?-

select distinct(payment)

from sales;

select count(distinct(payment))

from sales;

----------------- 3. What is the most common customer type? -----

select customer\_type , count(customer\_type) cnt

from sales

group by customer\_type

order by cnt desc ;

----------------- 4. Which customer type buys the most moneywise,quantitywise? -------

select customer\_type , sum(total) amt

from sales

group by customer\_type

order by amt desc;

select customer\_type , sum(quantity) qt

from sales

group by customer\_type

order by qt desc;

-------------------- 5. What is the gender of most of the customers?------

select customer\_type , gender , count(gender) cnt

from sales

group by customer\_type,gender

order by customer\_type asc ,gender desc;

----------- 6. What is the gender distribution per branch? ----------

select branch ,gender , count(gender) as count\_of\_individuals

from sales

group by branch , gender

order by branch , gender desc ;

-------------------- 7. Which time of the day do customers give most ratings?-------

select time\_of\_day , count(rating) as cnt\_of\_ratings

from sales

group by time\_of\_day

order by cnt\_of\_ratings desc ;

----------------- 8. Which time of the day do customers give most ratings per branch?

select branch , time\_of\_day , count(rating) as cnt\_of\_ratings

from sales

group by branch , time\_of\_day

order by branch , cnt\_of\_ratings desc ;

-------------------- 9. Which day of the week has the best avg ratings? ------

select name\_of\_the\_day , avg(rating) as avg\_ratings

from sales

group by name\_of\_the\_day

order by avg(rating) desc limit 1;

select name\_of\_the\_day,avg\_ratings from

(select name\_of\_the\_day , avg(rating) as avg\_ratings ,

dense\_rank() over(order by avg(rating) desc) rnk

from sales

group by name\_of\_the\_day) s

where rnk=1;

-------------------- 10. Which day of the week has the best average ratings per branch?

select branch,name\_of\_the\_day,avg\_ratings from

(select branch,name\_of\_the\_day , avg(rating) as avg\_ratings ,

dense\_rank() over(partition by branch order by avg(rating) desc) rnk

from sales

group by name\_of\_the\_day,branch) s

where rnk=1;