```
use amazon;
select * from  amazon_sales;



create table amazon_sale (
invoice_id int primary key,
branch varchar(30),
city varchar(30),
customer_type varchar(30) ,
gender varchar(10) ,
product_line varchar(100) ,
unit_price decimal ,
quantity int ,
VAT float ,
total decimal ,
dates date ,
times time ,
payment_method varchar(30) ,
cogs decimal,
gross_margin_percentage float,
gross_income decimal ,
rating float
);


select * from amazon_sale;


alter table amazon_sale
add column timeofday varchar(20);

set SQL_SAFE_UPDATES=0;

update amazon_sale set timeofday =
CASE
   WHEN times BETWEEN '05:00:00' AND '11:59:59' THEN 'Morning'
   WHEN times BETWEEN '12:00:00' AND '16:59:59' THEN 'Afternoon'
   ELSE 'Evening'
END;

ALTER TABLE amazon_sale
ADD COLUMN dayname VARCHAR(10);
```

```
UPDATE amazon_sale
SET dayname = DATE_FORMAT(dates, '%a');

ALTER TABLE amazon_sale
ADD COLUMN monthname VARCHAR(10);

UPDATE amazon_sale
SET monthname = DATE_FORMAT(dates, '%b');

ALTER TABLE amazon_sale
DROP COLUMN monthname;


-- 1. What is the count of distinct cities in the dataset?

select count(distinct city) from amazon_sale;

-- 2. For each branch, what is the corresponding city?

select branch, city from amazon_sale group by branch, city;

-- 3. What is the count of distinct product lines in the dataset?
select count(distinct product_line) from amazon_sale;

-- Sales and Revenue
-- 4. Which payment method occurs most frequently?



select payment_method, countmethod from(
select payment_method, count(payment_method)
as countmethod from amazon_sale group by payment_method)
as counts order by countmethod desc limit 2;


 -- 5. Which product line has the highest sales?

select product_line, sum(unit_price* quantity) as totalsales from amazon_sale
group by product_line order by totalsales desc limit 3;

--  6. How much revenue is generated each month?

select * from amazon_sale;
```

```
select monthname, sum(unit_price* quantity) as totalrevenue from amazon_sale
group by monthname order by monthname;

-- 7. In which month did the cost of goods sold (COGS) reach its peak?

select monthname, sum(cogs) as totalcogs from amazon_sale
group by monthname;


-- 8. Which product line generated the highest revenue?

select product_line, sum(unit_price* quantity) as totalrevenue from amazon_sale
group by product_line order by totalrevenue desc;


-- 9. In which city was the highest revenue recorded?

select city, sum(unit_price* quantity) as totalrevenue from amazon_sale
group by city order by totalrevenue desc;

--  10. Which product line incurred the highest VAT?
select product_line, sum(VAT) from amazon_sale GROUP BY product_line;


-- 11. For each product line, add a column indicating “Good” if its sales are above average,
-- otherwise “Bad.


select product_line, sales ,
case
when sales>avgsales  then "good" else "bad"
end as preferance
from(
select product_line, sum(unit_price* quantity) as sales,
 avg (sum(unit_price* quantity)) over () as avgsales
from amazon_sale group by product_line) as t;


-- 12. Identify the branch that exceeded the average number of products sold



select branch, sumofquanti from(
```

```
select branch , sum(quantity) as sumofquanti,
avg (sum(quantity)) over() as avgq from amazon_sale group by branch) as t
where sumofquanti>avgq;

-- 13. Which product line is most frequently associated with each gender?
SELECT gender, product_line, freq, rnk
FROM (
   SELECT gender,
        product_line,
        COUNT(*) AS freq,
        RANK() OVER (PARTITION BY gender ORDER BY COUNT(*) DESC) AS rnk
   FROM amazon_sale
   GROUP BY gender, product_line
) t  where rnk=1;

select gender, product_line , count(*) as freq from amazon_sale
group by gender, product_line;


-- 14. Calculate the average rating for each product line.
select product_line, avg(rating) from amazon_sale group by product_line;


-- 15. Identify the customer type contributing the highest revenue.
select * from amazon_sale;

select customer_type, sum(unit_price* quantity) as totalrevenue from amazon_sale
group by customer_type order by totalrevenue desc;


-- 16. Determine the city with the highest VAT percentage.
select city, VAT as totalrevenue from amazon_sale order by totalrevenue desc limit 1;

-- 17. Identify the customer type with the highest VAT payments.
select customer_type, VAT  from amazon_sale order by VAT desc limit 1;

-- 18. What is the count of distinct customer types?
select count(distinct customer_type) from amazon_sale;

-- 19. What is the count of distinct payment methods?
select count(distinct payment_method) from amazon_sale;

-- 20. Which customer type occurs most frequently?
```

```
select customer_type, count(*) as cn from amazon_sale group by customer_type order by cn
desc limit 1;

-- 21. Identify the customer type with the highest purchase frequency.
SELECT customer_type,
    COUNT(*) AS purchase_frequency
FROM amazon_sale
GROUP BY customer_type;

-- 22. Determine the predominant gender among customers.
select gender , count(*) as cn from amazon_sale group by gender
order by cn desc;

-- 23. Examine the distribution of genders within each branch. Time-Based Insights

SELECT branch,
    gender,
    COUNT(*) AS count
FROM amazon_sale
GROUP BY branch, gender
ORDER BY branch, gender;


-- 24. Count the sales occurrences for each time of day on every weekday.

select dayname, timeofday, count(*) as salescount from amazon_sale
group by timeofday, dayname having dayname not in ("Sat", "Sun");

-- 25. Identify the time of day when customers provide the most ratings.

SELECT timeofday,
    MAX(rating) AS highest_rating
FROM amazon_sale
GROUP BY timeofday;


-- 26. Determine the time of day with the highest customer ratings for each branch.select

select branch, timeofday, highest_rating from
(
SELECT branch, timeofday,  MAX(rating) AS highest_rating,
RANK() OVER (PARTITION BY branch order by max(rating) desc) AS rnk
FROM amazon_sale
GROUP BY timeofday ,branch
```

```
)
as t
where rnk=1;
use amazon;
select * from amazon_sale;

-- 27. Identify the day of the week with the highest average ratings.
-- 28. Determine the day of the week with the highest average ratings for each branch

select branch, dayname , ratings from (
select branch, dayname, avg(rating) as ratings,
rank() over(partition by branch order by avg(rating) desc) as rnk
 from amazon_sale group by branch, dayname) as t
 where rnk=1;

select sum(unit_price*quantity) from amazon_sale;
```