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

SELECT COUNT(DISTINCT (city)) AS Distinct\_Cities

FROM amazon\_data\_odin\_capstoneproject;

SELECT DISTINCT (city) AS Distinct\_Cities

FROM amazon\_data\_odin\_capstoneproject;

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

SELECT DISTINCT Branch,City

FROM amazon\_data\_odin\_capstoneproject;

**-- 3.What is the count of distinct product lines in the dataset?**

SELECT COUNT(DISTINCT(product\_line)) AS Distinct\_productlines

FROM amazon\_data\_odin\_capstoneproject;

SELECT DISTINCT(product\_line) AS Distinct\_productlines

FROM amazon\_data\_odin\_capstoneproject;

**-- 4.Which payment method occurs most frequently?**

SELECT payment,COUNT(payment) AS Frequency\_of\_paymentmethod

FROM amazon\_data\_odin\_capstoneproject

GROUP BY payment

ORDER BY Frequency\_of\_paymentmethod desc;

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

SELECT product\_line,ROUND(SUM(total)) AS Sales

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY Sales DESC;

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

SELECT monthname,ROUND(SUM(total)) AS totalsales

FROM amazon\_data\_odin\_capstoneproject

GROUP BY monthname

ORDER BY totalsales DESC;

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

SELECT monthname,ROUND(SUM(cogs)) AS cost\_of\_goods\_sold

FROM amazon\_data\_odin\_capstoneproject

GROUP BY monthname

ORDER BY cost\_of\_goods\_sold DESC;

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

SELECT product\_line,ROUND(SUM(total),2) AS Revenue

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY revenue desc;

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

SELECT City,ROUND(SUM(total)) AS City\_with\_highest\_sales

FROM amazon\_data\_odin\_capstoneproject

GROUP BY City

ORDER BY City\_with\_highest\_sales desc;

**-- 10.Which product line incurred the highest Value Added Tax?**

SELECT product\_line,ROUND(SUM(Tax\_5percent)) AS Value\_added\_Tax

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY Value\_added\_Tax DESC;

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

ALTER table amazon\_data\_odin\_capstoneproject

ADD column Sales\_Type varchar(6);

SELECT product\_line,AVG(total) AS Avg\_sales

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY Avg\_sales DESC;

WITH AvgSales AS (

SELECT

product\_line,

AVG(total) AS Avg\_sales

FROM

amazon\_data\_odin\_capstoneproject

GROUP BY

product\_line

)

UPDATE amazon\_data\_odin\_capstoneproject

SET Sales\_Type = CASE

WHEN total > (SELECT Avg\_sales

FROM AvgSales

WHERE AvgSales.product\_line = amazon\_data\_odin\_capstoneproject.product\_line) THEN 'Good'

ELSE 'Bad'

END;

SELECT sales\_type,COUNT(Sales\_type) AS count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY sales\_type;

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

SELECT product\_line,SUM(quantity) Products\_sold

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY products\_sold desc;

SELECT product\_line,Avg(quantity) avgProducts\_sold

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY avgproducts\_sold desc;

SELECT product\_line,

SUM(quantity) AS total\_quantity,

COUNT(\*) AS row\_count,

AVG(quantity) AS average\_quantity

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line;

**-- 13.Which product line is most frequently associated with each gender?**

WITH GenderProductCount AS (

SELECT

gender,

product\_line,

COUNT(\*) AS product\_count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY gender, product\_line

)

SELECT

gender,

product\_line,

product\_count

FROM GenderProductCount;

/\*WHERE (gender, product\_count) IN (

SELECT

gender,

MAX(product\_count)

FROM GenderProductCount

GROUP BY gender

);\*/

**-- 14.Calculate the average rating for each product line.**

SELECT product\_line,ROUND(AVG(rating),2) as Avg\_rating

FROM amazon\_data\_odin\_capstoneproject

GROUP BY product\_line

ORDER BY Avg\_rating desc;

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

SELECT

dayname,

timeofday,

COUNT(\*) AS sales\_count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY dayname, timeofday

ORDER BY FIELD(dayname, 'Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday'),

FIELD(timeofday, 'Morning', 'Afternoon', 'Evening', 'Night');

**-- 16.Identify the customer type contributing the highest revenue.**

SELECT customer\_type,ROUND(SUM(total)) AS Revenue\_generating

FROM amazon\_data\_odin\_capstoneproject

GROUP BY customer\_type;

**-- 17.Determine the city with the highest VAT percentage.**

SELECT city,ROUND(SUM(tax\_5percent),2) AS value\_added\_tax\_percent

FROM amazon\_data\_odin\_capstoneproject

GROUP BY city

ORDER BY value\_added\_tax\_percent desc;

SELECT

city,

SUM(tax\_5percent) / SUM(total) \* 100 AS VAT\_percentage

FROM

amazon\_data\_odin\_capstoneproject

GROUP BY

city

ORDER BY

VAT\_percentage DESC;

**-- 18.Identify the customer type with the highest VAT payments.**

SELECT customer\_type,ROUND(SUM(tax\_5percent)) AS VAT\_payments

FROM amazon\_data\_odin\_capstoneproject

GROUP BY customer\_type;

**-- 19.What is the count of distinct customer types in the dataset?**

SELECT customer\_type,COUNT(customer\_type) AS Count\_of\_customer\_type

FROM amazon\_data\_odin\_capstoneproject

GROUP BY customer\_type;

**-- 20.What is the count of distinct payment methods in the dataset?**

SELECT Payment,COUNT(payment) AS Count\_of\_payment

FROM amazon\_data\_odin\_capstoneproject

GROUP BY payment;

**-- 21.Which customer type occurs most frequently?**

SELECT

customer\_type,

COUNT(\*) AS frequency

FROM

amazon\_data\_odin\_capstoneproject

GROUP BY

customer\_type

ORDER BY

frequency DESC;

**-- 22.Identify the customer type with the highest purchase frequency.**

SELECT

customer\_type,

ROUND(SUM(TOTAL),2) AS purchasing\_frequency

FROM

amazon\_data\_odin\_capstoneproject

GROUP BY

customer\_type

ORDER BY

purchasing\_frequency DESC;

**-- 23.Determine the predominant gender among customers.**

SELECT gender,count(gender) AS Gender\_count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY gender;

**-- 24.Examine the distribution of genders within each branch.**

SELECT

branch,gender,

COUNT(gender) AS Gender\_count

FROM

amazon\_data\_odin\_capstoneproject

GROUP BY

gender,branch;

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

SELECT timeofday,COUNT(rating) AS Ratings\_count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY timeofday

ORDER BY Ratings\_count desc;

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

SELECT branch,timeofday,COUNT(rating) AS Ratings\_count

FROM amazon\_data\_odin\_capstoneproject

GROUP BY timeofday,branch

ORDER BY Ratings\_count desc;

**-- 27.Identify the day of the week with the highest average ratings.**

SELECT dayname,ROUND(avg(rating),2) AS avg\_ratings

FROM amazon\_data\_odin\_capstoneproject

GROUP BY dayname

ORDER BY avg\_ratings DESC;

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

SELECT branch,dayname,ROUND(avg(rating),2) AS avg\_ratings

FROM amazon\_data\_odin\_capstoneproject

GROUP BY dayname,branch

ORDER BY avg\_ratings,branch DESC;