**CDACL-006 Market Analysis**

use project\_orders;

SELECT \* from aisles;

select \* from departments;

select \* from order\_products\_train;

select \* from orders;

select \* from products;

**-- 1 What are the top 10 aisles with the highest number of products?**

SELECT

aisle,

COUNT(product\_id) AS product\_count

FROM

aisles a

JOIN

products p ON a.aisle\_id = p.aisle\_id

GROUP BY

a.aisle

ORDER BY

product\_count DESC

LIMIT 10;

**-- 2 How many unique departments are there in the dataset?**

SELECT

COUNT(DISTINCT department\_id) AS unique\_departments

FROM

departments;

**-- 3 What is the distribution of products across departments?**

SELECT

d.department,

COUNT(p.product\_id) AS product\_count

FROM

products p

JOIN

departments d ON p.department\_id = d.department\_id

GROUP BY

d.department

ORDER BY

product\_count DESC;

**-- 4. What are the top 10 products with the highest reorder rates?**

SELECT

p.product\_name,

COUNT(CASE WHEN opt.reordered = 1 THEN 1 END) / COUNT(opt.product\_id) AS reorder\_rate

FROM

order\_products\_train opt

JOIN

products p ON opt.product\_id = p.product\_id

GROUP BY

p.product\_id, p.product\_name

ORDER BY

reorder\_rate DESC

LIMIT 10;

**-- 5 How many unique users have placed orders in the dataset?**

SELECT

COUNT(DISTINCT user\_id) AS unique\_user\_count

FROM

orders;

**-- 6. What is the average number of days between orders for each user?**

SELECT user\_id,

ROUND(AVG(days\_since\_prior\_order)) AS avg\_days\_between\_orders

FROM orders

WHERE days\_since\_prior\_order IS NOT NULL

GROUP BY user\_id;

**-- 7 What are the peak hours of order placement during the day?**

SELECT

order\_hour\_of\_day,

COUNT(order\_id) AS order\_count

FROM

orders

GROUP BY

order\_hour\_of\_day

ORDER BY

order\_count DESC;

**-- 8 How does order volume vary by day of the week?**

SELECT

order\_dow,

COUNT(order\_id) AS order\_count

FROM

orders

GROUP BY

order\_dow

ORDER BY

order\_dow;

**-- 9 What are the top 10 most ordered products?**

SELECT

p.product\_name,

COUNT(opt.product\_id) AS order\_count

FROM

order\_products\_train opt

JOIN

products p ON opt.product\_id = p.product\_id

GROUP BY

p.product\_name

ORDER BY

order\_count DESC

LIMIT 10;

**-- 10 How many users have placed orders in each department?**

SELECT

d.department\_id,

COUNT(DISTINCT o.user\_id) AS num\_users

FROM

orders o

JOIN

order\_products\_train opt ON o.order\_id = opt.order\_id

JOIN

products p ON opt.product\_id = p.product\_id

JOIN

departments d ON p.department\_id = d.department\_id

GROUP BY

d.department\_id

ORDER BY

num\_users DESC;

**-- 11 What is the average number of products per order?**

SELECT

AVG(product\_count) AS avg\_products\_per\_order

FROM (

SELECT

o.order\_id,

COUNT(opt.product\_id) AS product\_count

FROM

orders o

JOIN

order\_products\_train opt ON o.order\_id = opt.order\_id

GROUP BY

o.order\_id

) AS product\_counts;

**-- 12 What are the most reordered products in each department?**

SELECT

d.department\_id,

p.product\_name,

COUNT(opt.product\_id) AS reorder\_count

FROM

order\_products\_train opt

JOIN

products p ON opt.product\_id = p.product\_id

JOIN

departments d ON p.department\_id = d.department\_id

WHERE

opt.reordered = 1

GROUP BY

d.department\_id, p.product\_name

ORDER BY

d.department\_id, reorder\_count DESC;

**-- 13 How many products have been reordered more than once?**

SELECT

COUNT(DISTINCT opt.product\_id) AS products\_reordered\_more\_than\_once

FROM

order\_products\_train opt

WHERE

opt.reordered = 1

GROUP BY

opt.product\_id

HAVING

COUNT(opt.product\_id) > 1;

**-- 14 What is the average number of products added to the cart per order?**

SELECT

AVG(product\_count) AS avg\_products\_per\_order

FROM (

SELECT

o.order\_id,

COUNT(opt.product\_id) AS product\_count

FROM

orders o

JOIN

order\_products\_train opt ON o.order\_id = opt.order\_id

GROUP BY

o.order\_id

) AS product\_counts;

**-- 15 How does the number of orders vary by hour of the day?**

SELECT

o.order\_hour\_of\_day,

COUNT(o.order\_id) AS order\_count

FROM

orders o

GROUP BY

o.order\_hour\_of\_day

ORDER BY

o.order\_hour\_of\_day;

**-- 16 What is the distribution of order sizes (number of products per order)?**

SELECT

product\_count,

COUNT(\*) AS order\_count

FROM (

SELECT

o.order\_id,

COUNT(opt.product\_id) AS product\_count

FROM

order\_products\_train opt

JOIN

orders o ON o.order\_id = opt.order\_id

GROUP BY

o.order\_id

) AS order\_sizes

GROUP BY

product\_count

ORDER BY

product\_count;

**-- 17 What is the average reorder rate for products in each aisle?**

SELECT

a.aisle,

AVG(CASE WHEN opt.reordered = 1 THEN 1 ELSE 0 END) AS avg\_reorder\_rate

FROM

order\_products\_train opt

JOIN

products p ON opt.product\_id = p.product\_id

JOIN

aisles a ON p.aisle\_id = a.aisle\_id

GROUP BY

a.aisle

ORDER BY

avg\_reorder\_rate DESC;

**-- 18 How does the average order size vary by day of the week?**

select

order\_dow,

COUNT(order\_id) as total\_orders,

FLOOR(COUNT(order\_id) / count(distinct order\_id)) as avg\_order\_size

from

orders

group by

order\_dow

order by

order\_dow;

**-- 19 What are the top 10 users with the highest number of orders?**

SELECT

o.user\_id,

COUNT(o.order\_id) AS order\_count

FROM

orders o

GROUP BY

o.user\_id

ORDER BY

order\_count DESC

LIMIT 10;

**-- 20 How many products belong to each aisle and department?**

SELECT

a.aisle,

d.department,

COUNT(p.product\_id) AS product\_count

FROM

products p

JOIN

aisles a ON p.aisle\_id = a.aisle\_id

JOIN

departments d ON p.department\_id = d.department\_id

GROUP BY

a.aisle, d.department

ORDER BY

d.department, a.aisle;

PURCHASE BEHAVIOUR ANALYSIS :-

Customer Purchasing Behaviour:

-- Average Order Frequency by Day of the Week

SELECT

order\_dow,

COUNT(order\_id) AS num\_orders,

AVG(order\_number) AS avg\_order\_number

FROM

orders

GROUP BY

order\_dow

ORDER BY

order\_dow;

-- Average Order Size

SELECT

order\_id,

COUNT(product\_id) AS order\_size

FROM

order\_products\_train

GROUP BY

order\_id;

-- Reordered Products

SELECT

p.product\_name,

COUNT(\*) AS reorder\_count

FROM

order\_products\_train op

JOIN

products p ON op.product\_id = p.product\_id

WHERE

op.reordered = 1

GROUP BY

p.product\_name

ORDER BY

reorder\_count DESC;

Product Performance Analysis

-- Top-Selling Products

SELECT

p.product\_name,

COUNT(op.product\_id) AS total\_sales

FROM

order\_products\_train op

JOIN

products p ON op.product\_id = p.product\_id

GROUP BY

p.product\_name

ORDER BY

total\_sales DESC

LIMIT 10;

-- Department Performance

SELECT

d.department,

COUNT(op.product\_id) AS total\_sales

FROM

order\_products\_train op

JOIN

products p ON op.product\_id = p.product\_id

JOIN

departments d ON p.department\_id = d.department\_id

GROUP BY

d.department

ORDER BY

total\_sales DESC;

-- Aisle Performance

SELECT

a.aisle,

COUNT(op.product\_id) AS total\_sales

FROM

order\_products\_train op

JOIN

products p ON op.product\_id = p.product\_id

JOIN

aisles a ON p.aisle\_id = a.aisle\_id

GROUP BY

a.aisle

ORDER BY

total\_sales DESC;

Customer Behavior Segmentation

-- Frequency of Purchases

SELECT

user\_id,

COUNT(order\_id) AS total\_orders

FROM

orders

GROUP BY

user\_id

HAVING

total\_orders > 5;

-- Average Spend per Customer

SELECT

user\_id,

SUM(op.order\_amount) AS total\_spent

FROM

orders o

JOIN

order\_products\_train op ON o.order\_id = op.order\_id

GROUP BY

user\_id

ORDER BY

total\_spent DESC;

Order Time Analysis

-- Average Order Size by Hour of Day

SELECT

order\_hour\_of\_day,

AVG(order\_size) AS avg\_order\_size

FROM (

SELECT

o.order\_id,

COUNT(op.product\_id) AS order\_size,

o.order\_hour\_of\_day

FROM

order\_products\_train op

JOIN

orders o ON op.order\_id = o.order\_id

GROUP BY

o.order\_id, o.order\_hour\_of\_day

) AS order\_data

GROUP BY

order\_hour\_of\_day

ORDER BY

order\_hour\_of\_day;