TheLook eCommerce

RevoU FSDA Week 4 & 5 | SQL Assignment

by **Riadhi Nur Fajrina**

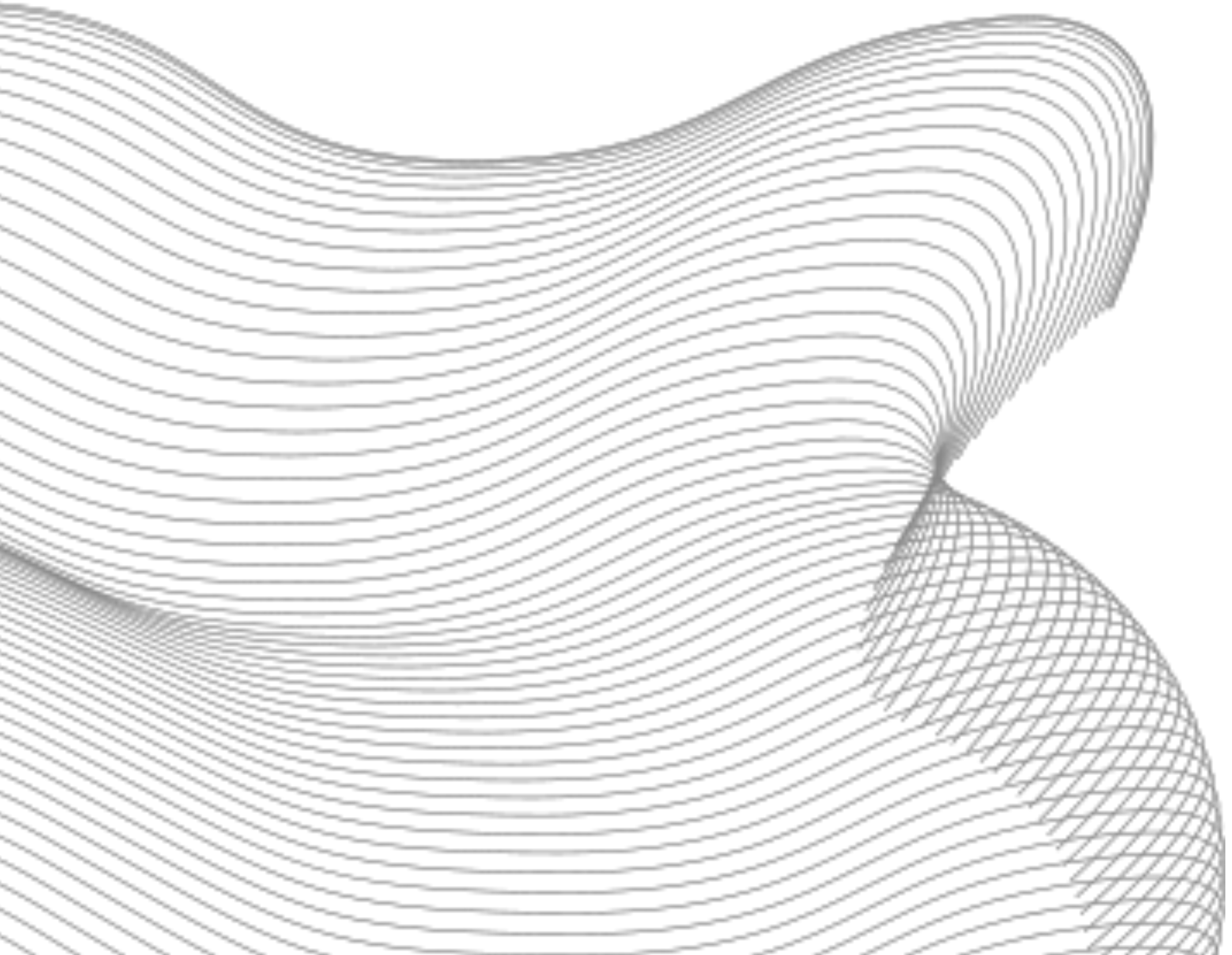
Table of Contents

● **01** Overview

● **02** Growth Analysis

● **03** User Retention Analysis

● **04** Conclusion



01
OVERVIEW

Business Overview

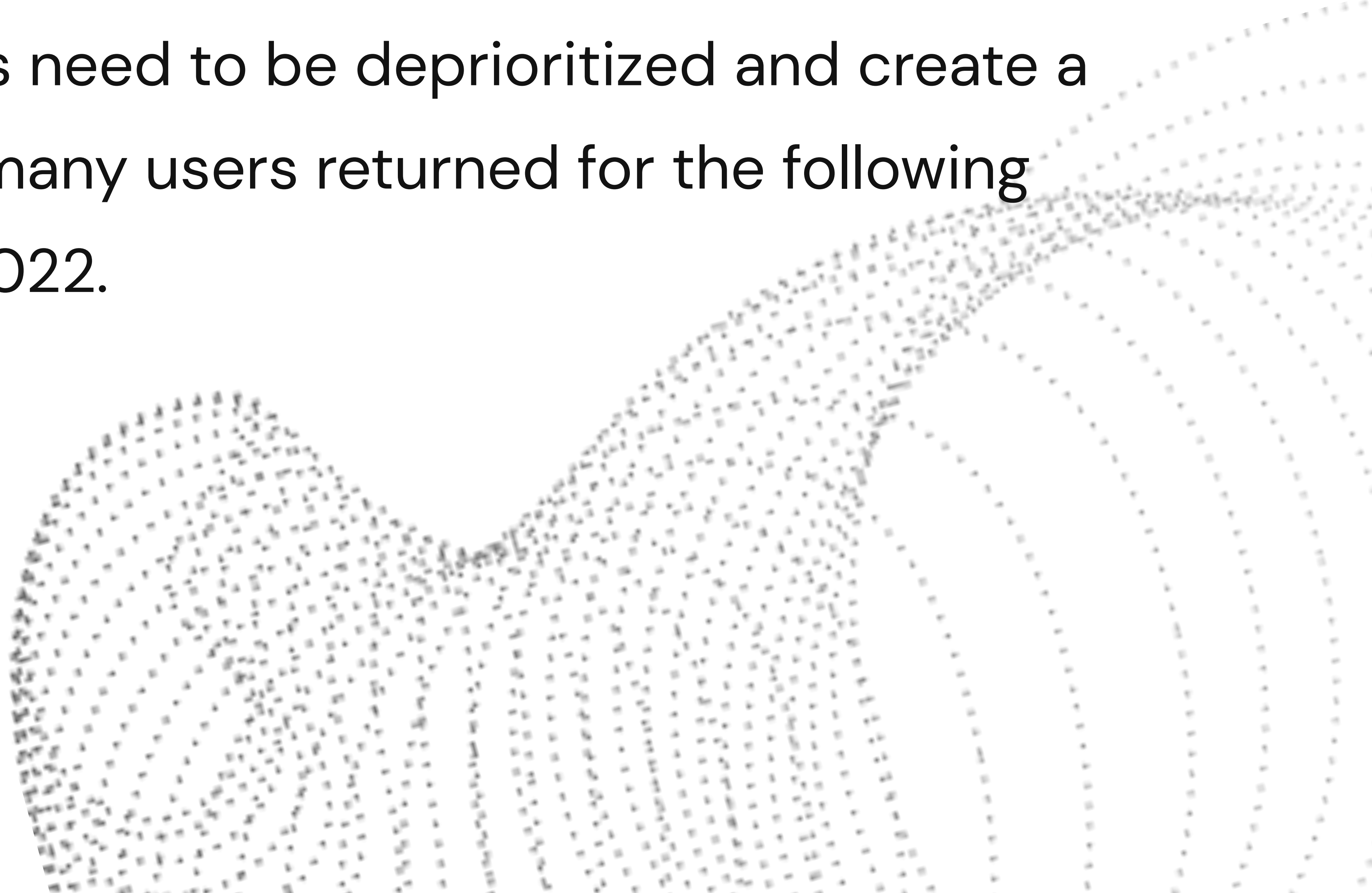
TheLook is a fictitious eCommerce clothing website. The company is currently in an optimization mode as a result of the potential crisis in 2023.

The management has decided to reduce resources in some categories that have seen the lowest growth in the last year.

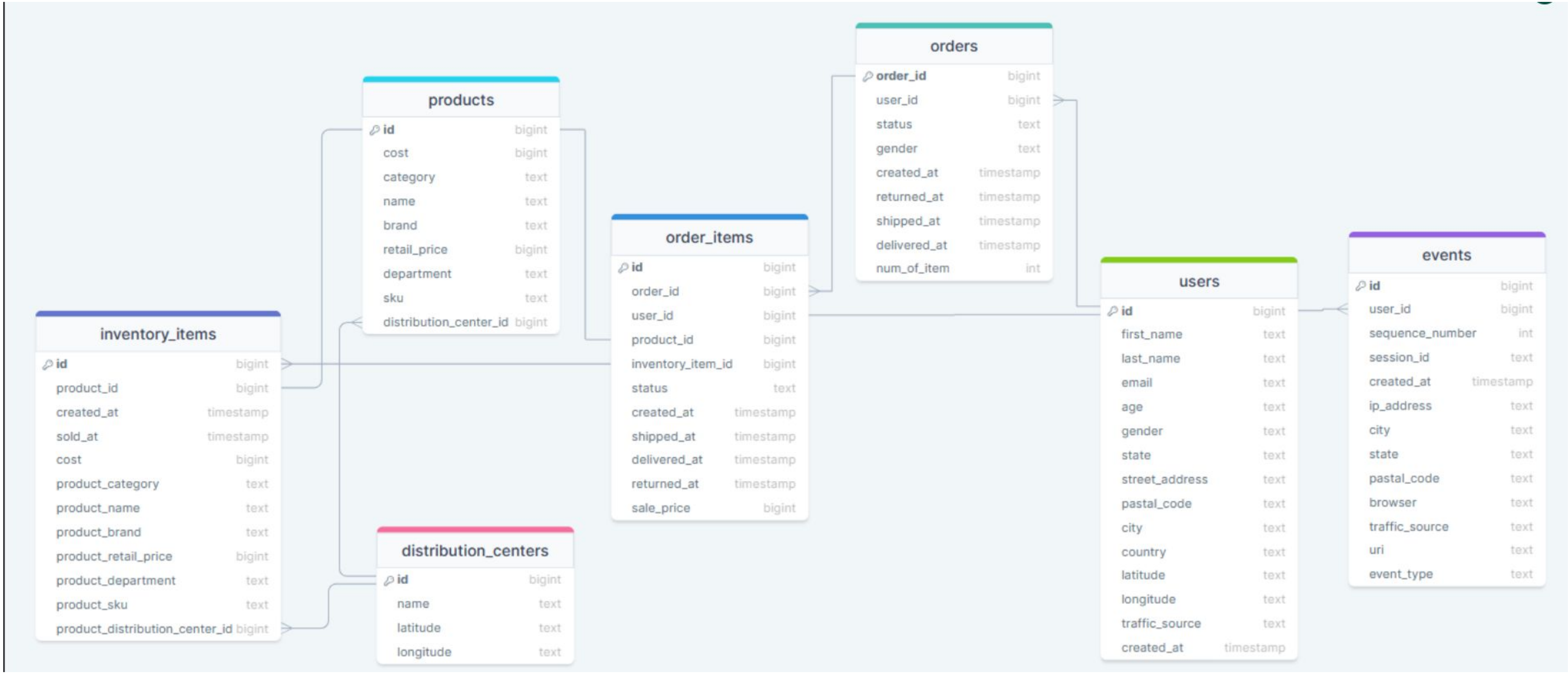
On the other hand, they want to continue the analysis by better understanding user retention behaviors and how to increase retention rates.

Objective

To make a recommendation by analyzing the lowest revenue and profit growth in the previous year and determine which categories need to be deprioritized and create a monthly retention cohort to calculate how many users returned for the following month in 2022.



ERD of the Dataset



orders	
order_id	bigint
user_id	bigint
status	text
gender	text
created_at	timestamp
returned_at	timestamp
shipped_at	timestamp
delivered_at	timestamp
num_of_item	int

products	
id	bigint
cost	bigint
category	text
name	text
brand	text
retail_price	bigint
department	text
sku	text
distribution_center_id	bigint

order_items	
id	bigint
order_id	bigint
user_id	bigint
product_id	bigint
inventory_item_id	bigint
status	text
created_at	timestamp
shipped_at	timestamp
delivered_at	timestamp
returned_at	timestamp
sale_price	bigint

events	
id	bigint
user_id	bigint
sequence_number	int
session_id	text
created_at	timestamp
ip_address	text
city	text
state	text
pastal_code	text
browser	text
traffic_source	text
uri	text
event_type	text

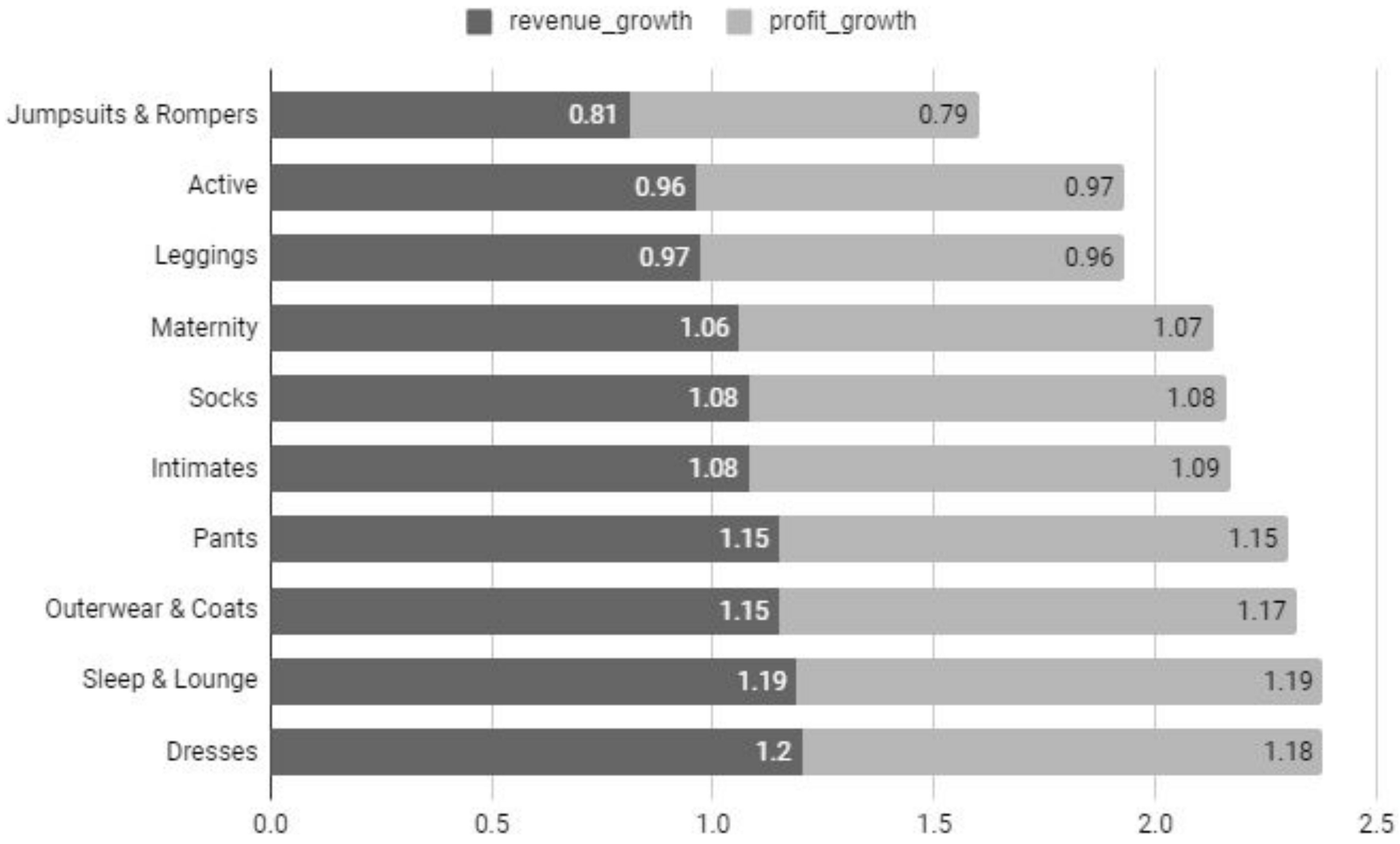
users	
id	bigint
first_name	text
last_name	text
email	text
age	text
gender	text
state	text
street_address	text
pastal_code	text
city	text
country	text
latitude	text
longitude	text
traffic_source	text
created_at	timestamp

inventory_items	
id	bigint
product_id	bigint
created_at	timestamp
sold_at	timestamp
cost	bigint
product_category	text
product_name	text
product_brand	text
product_retail_price	bigint
product_department	text
product_sku	text
product_distribution_center_id	bigint

distribution_centers	
id	bigint
name	text
latitude	text
longitude	text

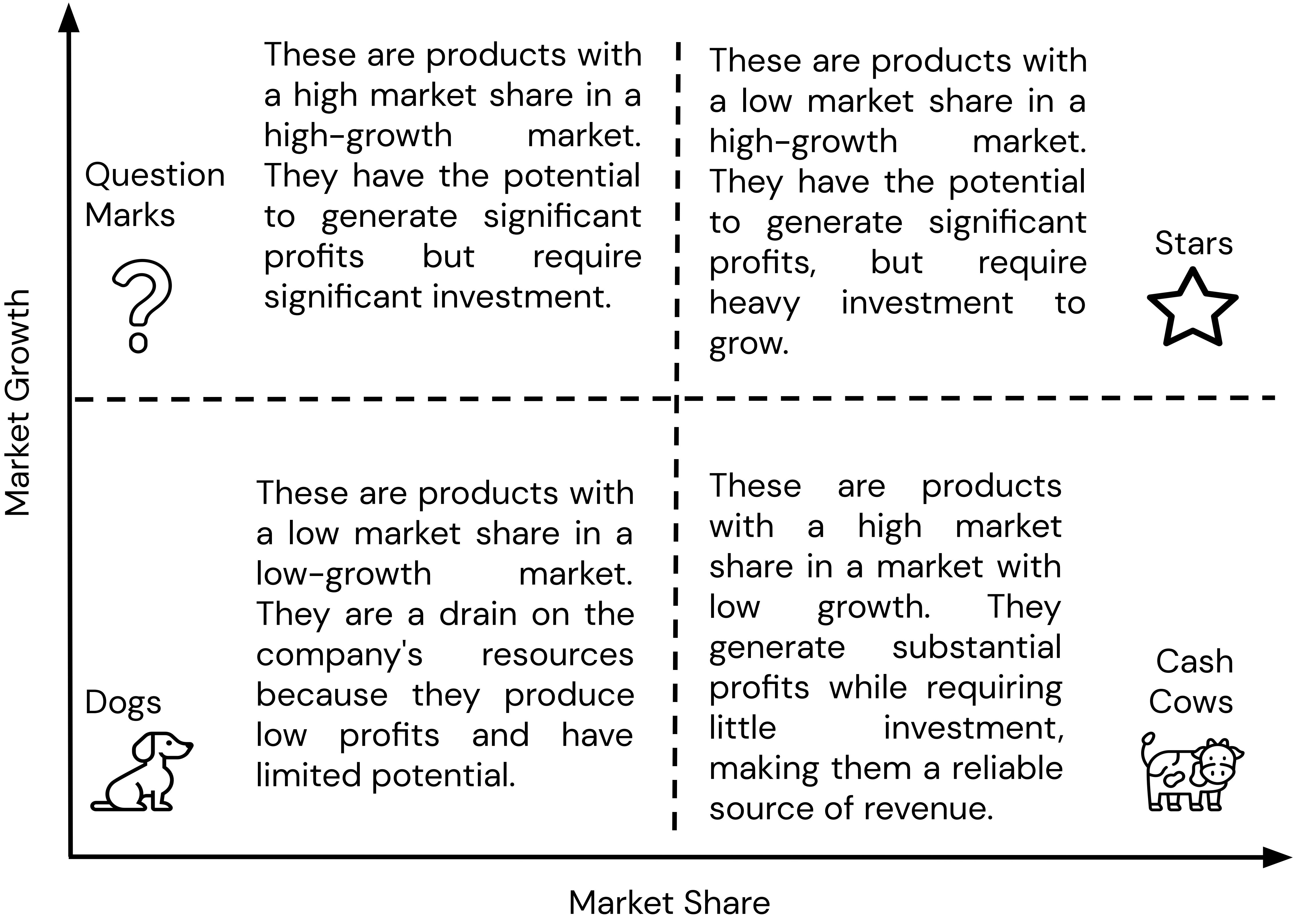
02
GROWTH ANALYSIS

Lowest Growth in the Past Year



When compared to other categories, **Jumpsuits and Rompers** had the lowest revenue and profit growth rates in the previous year, followed by **Active** and **Leggings.**

BCG Matrix



Question Marks

These are products with a high market share in a high-growth market. They have the potential to generate significant profits but require significant investment.

These are products with a low market share in a high-growth market. They have the potential to generate significant profits, but require heavy investment to grow.

Stars

Dogs

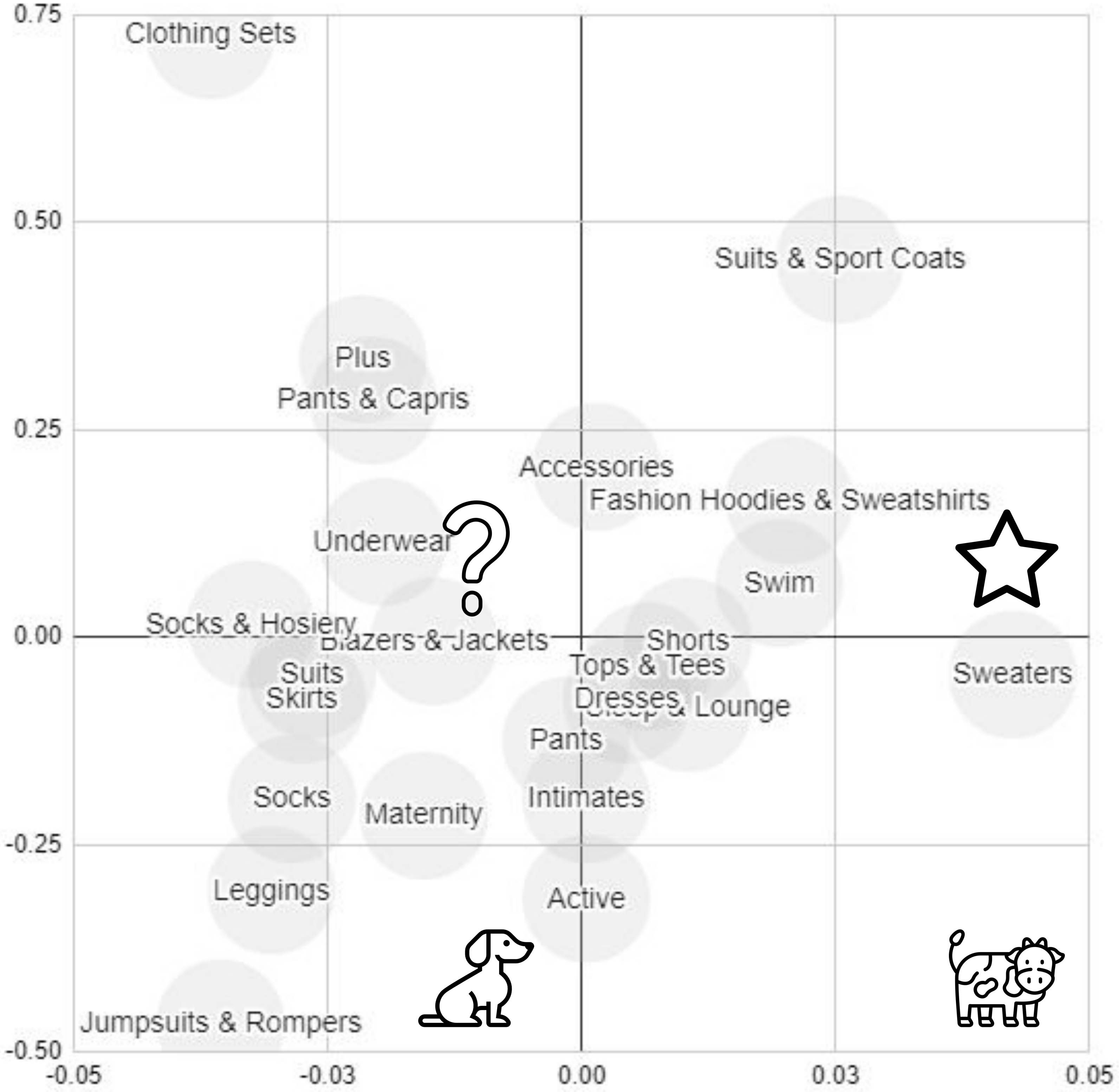
These are products with a low market share in a low-growth market. They are a drain on the company's resources because they produce low profits and have limited potential.

These are products with a high market share in a market with low growth. They generate substantial profits while requiring little investment, making them a reliable source of revenue.

Cash Cows

To better understand the growth of the category, we projected the results into the BCG Matrix. We use the BCG Matrix to analyze our product portfolio and choose where to allocate strategic resources and prioritize investments.

BCG Matrix & Insights



We can deprioritize **Jumpsuits & Rompers and Leggings** that fall in the bottom of the *dog* quadrant based on the BCG Matrix. It could mean reducing marketing spend or reallocating resources to focus on higher potential categories such as **Suits & Sport Coats** or **Fashion Hoodies & Sweatshirts** in *star* quadrant.

03

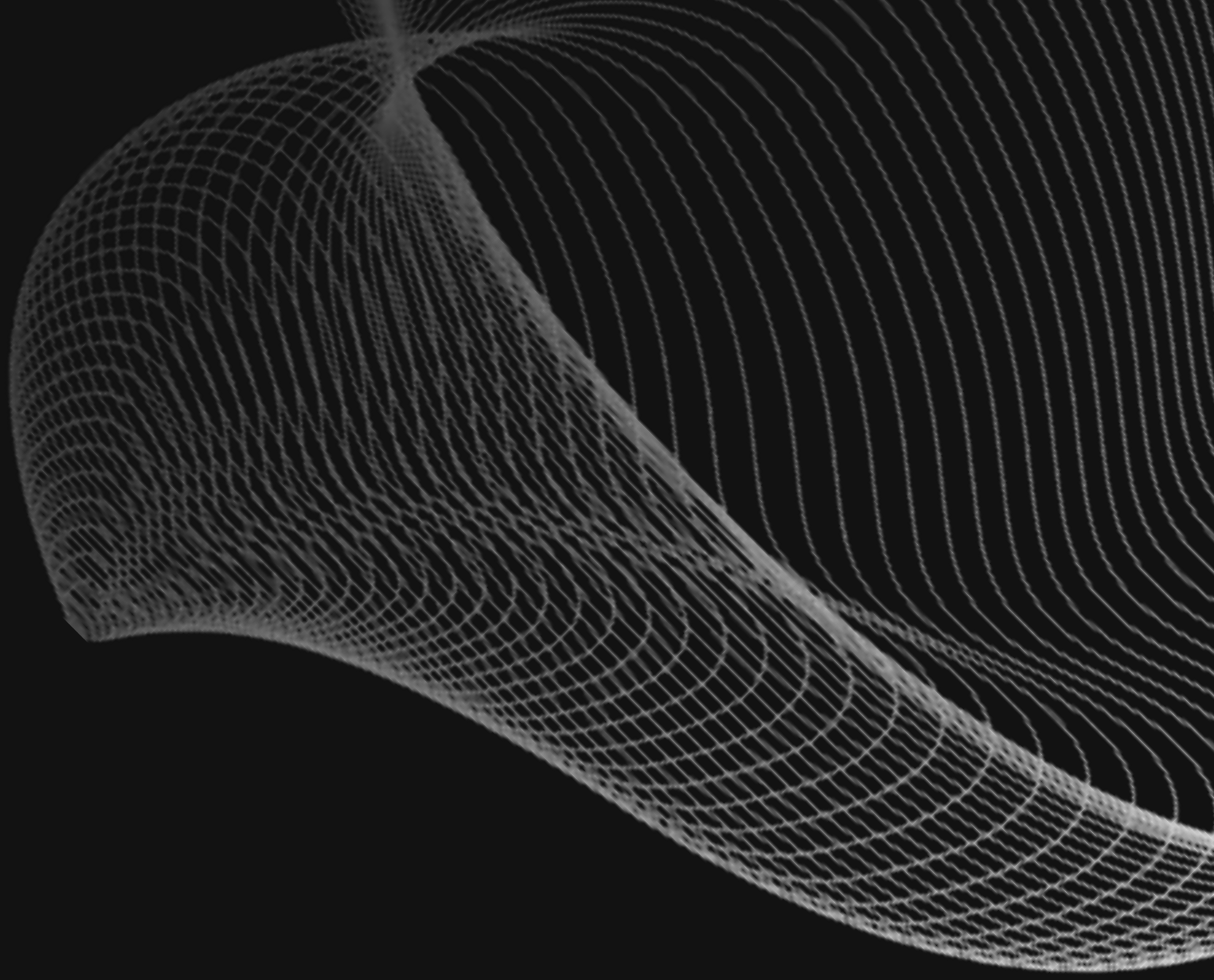
USER RETENTION ANALYSIS

Percentage of Users Who Returned After Their First Purchase

Month	Cohort Size	0	1	2	3	4	5	6	7	8	9	10	11	12
Jan 2022	2236	100.00%	5.77%	5.50%	4.34%	4.38%	4.20%	4.56%	4.34%	4.79%	4.47%	3.89%	4.61%	4.11%
Feb 2022	2169	100.00%	7.47%	4.20%	4.66%	5.26%	5.30%	4.56%	5.07%	5.39%	4.66%	4.56%	4.15%	
Mar 2022	2512	100.00%	6.93%	5.25%	5.10%	5.21%	4.58%	4.90%	5.25%	5.41%	5.65%	4.42%		
Apr 2022	2636	100.00%	6.94%	5.42%	5.61%	6.49%	5.88%	4.74%	5.54%	5.54%	5.12%			
May 2022	2735	100.00%	7.28%	6.25%	6.84%	5.81%	6.00%	6.76%	5.41%	5.59%				
Jun 2022	2788	100.00%	8.82%	5.92%	6.17%	6.38%	5.38%	6.35%	4.95%					
Jul 2022	3041	100.00%	8.39%	7.23%	6.97%	7.10%	7.27%	6.15%						
Aug 2022	3177	100.00%	9.73%	7.81%	8.22%	7.81%	6.26%							
Sep 2022	3468	100.00%	10.81%	9.08%	8.59%	8.13%								
Oct 2022	3720	100.00%	12.10%	10.22%	9.01%									
Nov 2022	4118	100.00%	14.84%	11.53%										
Dec 2022	4672	100.00%	18.66%											

Insights from Cohort Analysis

- According to the cohort analysis, the total number of users has been increased. The largest cohort was in March 2022, with 2,512 users, and the smallest was in February 2022, with 2,169 users.
- The percentage of users who return tends to decline over time. For example, in the January 2022 cohort, 5.77% of users returned in the first month, but only 4.11% returned by the twelfth month. This pattern is consistent across all cohorts.
- It is possible that the increase in users seen from August to December due to seasonal trends as the fall and winter seasons transition.



04
CONCLUSION

Conclusion

- We can deprioritizing J**umpsuits & Rompers** and **Leggings** category by reducing marketing spend or reallocating resources to focus on higher potential categories such as **Suits & Sport Coats** or **Fashion Hoodies & Sweatshirts** category.
- We may create promotional campaign for existing and old users to pique their interest and come back, such as cashbacks or discounts with a certain minimum spend.

SQL Syntax

```
WITH
 sales2022 AS(
  SELECT
    category
    ,EXTRACT(YEAR FROM ord.created_at) IN (2022) AS year2022
    ,SUM (sale_price) AS total_sales2022
    ,SUM (sale_price - cost) AS total_profit2022
   FROM `sql-project-376612.thelook_ecommerce.order_items` AS ord
    INNER JOIN `sql-project-376612.thelook_ecommerce.products` AS prod
    ON prod.id=product_id
   WHERE status='Complete'
    AND EXTRACT(YEAR FROM ord.created_at)=2022
   GROUP BY 1,2
 )

 ,sales2021 AS (
  SELECT
    category
    ,EXTRACT(YEAR FROM ord.created_at) IN (2021) AS year2021
    ,SUM (sale_price) AS total_sales2021
    ,SUM (sale_price - cost) AS total_profit2021
   FROM `sql-project-376612.thelook_ecommerce.order_items` AS ord
   INNER JOIN `sql-project-376612.thelook_ecommerce.products` AS prod
    ON prod.id=product_id
   WHERE status='Complete'
    AND EXTRACT(YEAR FROM ord.created_at)=2021
   GROUP BY 1,2
 )
```

```
,growth AS(
  SELECT
    sales2022.category AS category
    ,total_sales2021
    ,total_sales2022
    ,ROUND(((total_sales2022-total_sales2021)/total_sales2021),2) AS revenue_growth
    ,ROUND(total_sales2022/SUM(total_sales2022) OVER(PARTITION BY year2022),3) AS
revenue_market_share
    ,total_profit2021
    ,total_profit2022
    ,ROUND(((total_profit2022-total_profit2021)/total_profit2021),2) AS profit_growth
    ,ROUND(total_profit2022/SUM(total_profit2022) OVER(PARTITION BY year2022),3) AS
profit_market_share
   FROM sales2022
    INNER JOIN sales2021
    ON sales2022.category=sales2021.category
   ORDER BY revenue_growth, profit_growth
 )

SELECT
   *
  ,CASE
   WHEN revenue_market_share >= AVG(revenue_market_share) OVER() AND revenue_growth >=
AVG(revenue_growth) OVER()
    THEN 'Star'
   WHEN revenue_market_share >= AVG(revenue_market_share) OVER() AND revenue_growth <
AVG(revenue_growth) OVER()
    THEN 'Cash Cow'
   WHEN revenue_market_share < AVG(revenue_market_share) OVER() AND revenue_growth >=
AVG(revenue_growth) OVER()
    THEN 'Question Mark'
   ELSE 'Dog'
   END quadrant
 FROM growth
 ORDER BY revenue_growth , profit_growth
;
```

SQL Syntax

```sql
WITH c_order AS (
  SELECT
      user_id
      ,MIN(DATE(DATE_TRUNC(created_at, MONTH))) as coh_month
    FROM `sql-project-376612.thelook_ecommerce.order_items`
    GROUP BY 1
  )

  ,c_user AS (
    SELECT
          c_order.*
      FROM c_order
        LEFT JOIN `sql-project-376612.thelook_ecommerce.users` AS user
        ON c_order.user_id = user.id
      WHERE coh_month >= DATE(DATE_TRUNC(user.created_at, MONTH))
  )

  ,c_act AS (
    SELECT
        act.user_id AS user_id
        ,DATE_DIFF(
            DATE(DATE_TRUNC(created_at, MONTH)),
            coh.coh_month,
            MONTH
        ) AS month_num
    FROM `sql-project-376612.thelook_ecommerce.order_items` AS act
      LEFT JOIN c_user AS coh
      ON act.user_id = coh.user_id
      WHERE EXTRACT(YEAR FROM coh.coh_month) IN (2022)
      GROUP BY 1,2
  )
```

```sql
  ,c_size AS (
    SELECT
        coh_month
        ,COUNT(1) AS num_users
      FROM c_user
      GROUP BY 1
      ORDER BY 1
  )

 ,ret AS (
    SELECT
        c_user.coh_month
      ,c_act.month_num
        ,COUNT(1) AS num_users
      FROM c_act
        LEFT JOIN c_user
        ON c_act.user_id = c_user.user_id
      GROUP BY 1,2
  )

SELECT t1.coh_month AS Cohort_Month
    ,t2.num_users AS Cohort_Size
    ,t1.month_num AS Month_Num
    ,t1.num_users AS Total_Users
    ,t1.num_users/t2.num_users AS Percent
  FROM ret AS t1
    LEFT JOIN c_size AS t2
    ON t1.coh_month = t2.coh_month
  WHERE t1.month_num IS NOT NULL
  ORDER BY 1,3
;
```

Thank you!

Feel free to comment below, DMs, or reach me at riadhigodjay@gmail.com for any feedbacks or suggestions.