TheLook e-commerce Analysis

Query: ***Click here***

Table of contents

01 Business Understanding

Business Background & Core Problems

TheLook is an eCommerce clothing site company. Currently, the company is in the optimization mode caused by the potential crisis in 2023. The management has decided to cut off resources in some categories with the lowest growth in the past 1 year. On another side, they want to continue the analysis by understanding the retention behaviors of the users and how to increase the retention rate.

Objective

Growth Analysis

Analyze which product category has the lowest growth rate

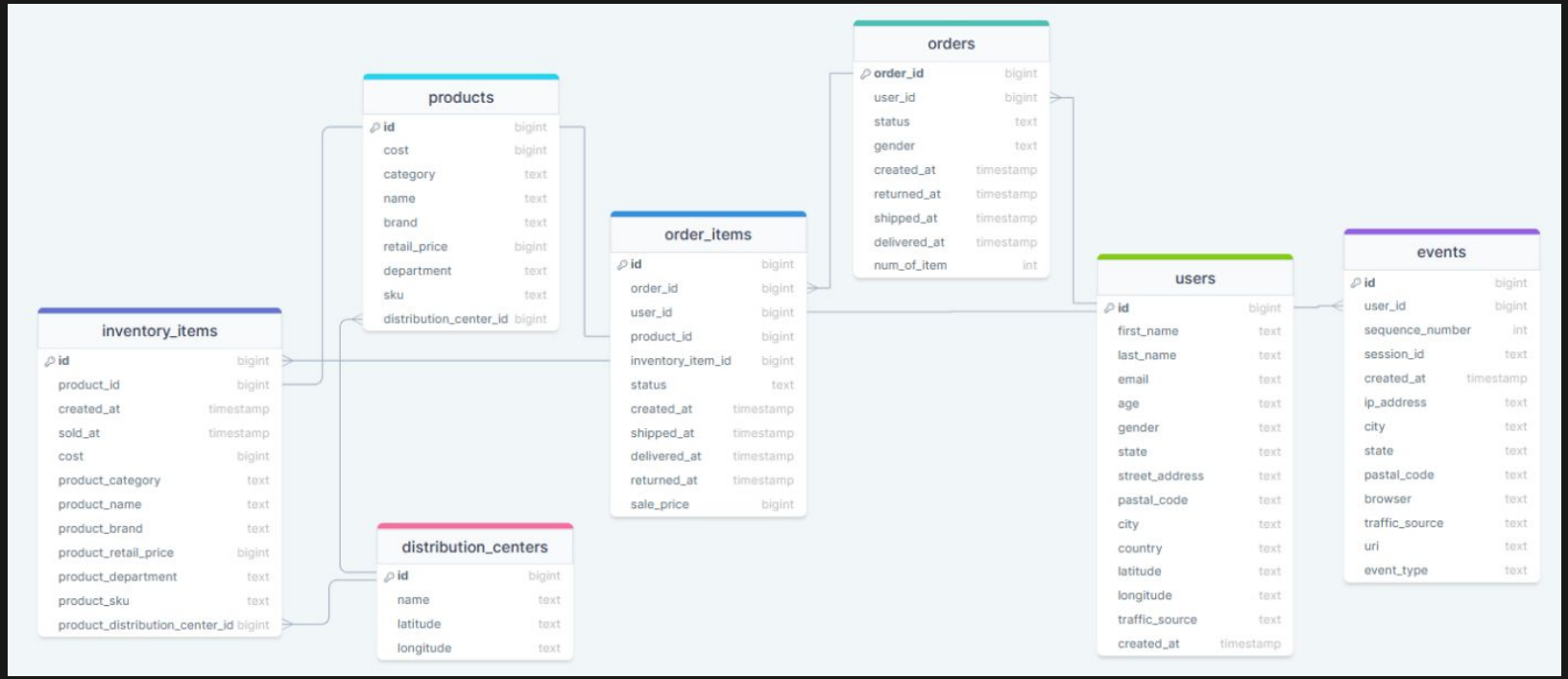
User Retention Analysis

Analyze the user retention based on their order behavior.

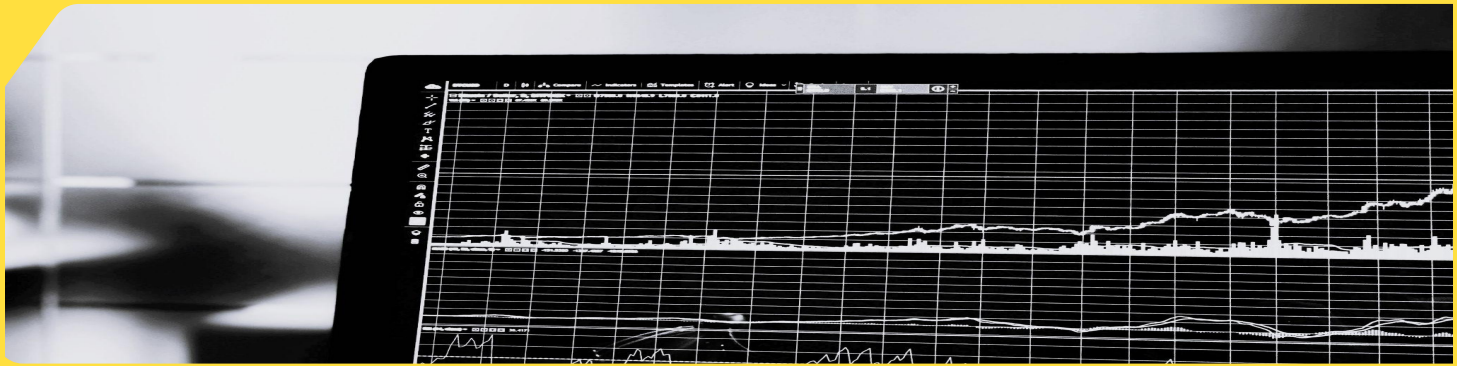
Give Recommendation

Based on those analyses, propose some solutions to answer the business problems

Dataset Entity Relationship Diagram (ERD)



02 Growth Analysis



Growth Analysis: Parameters

Growth analyses will be conducted based on two parameters. We'll be comparing the annual total for both parameters and see how they grow in a year (from 2021 to 2022).

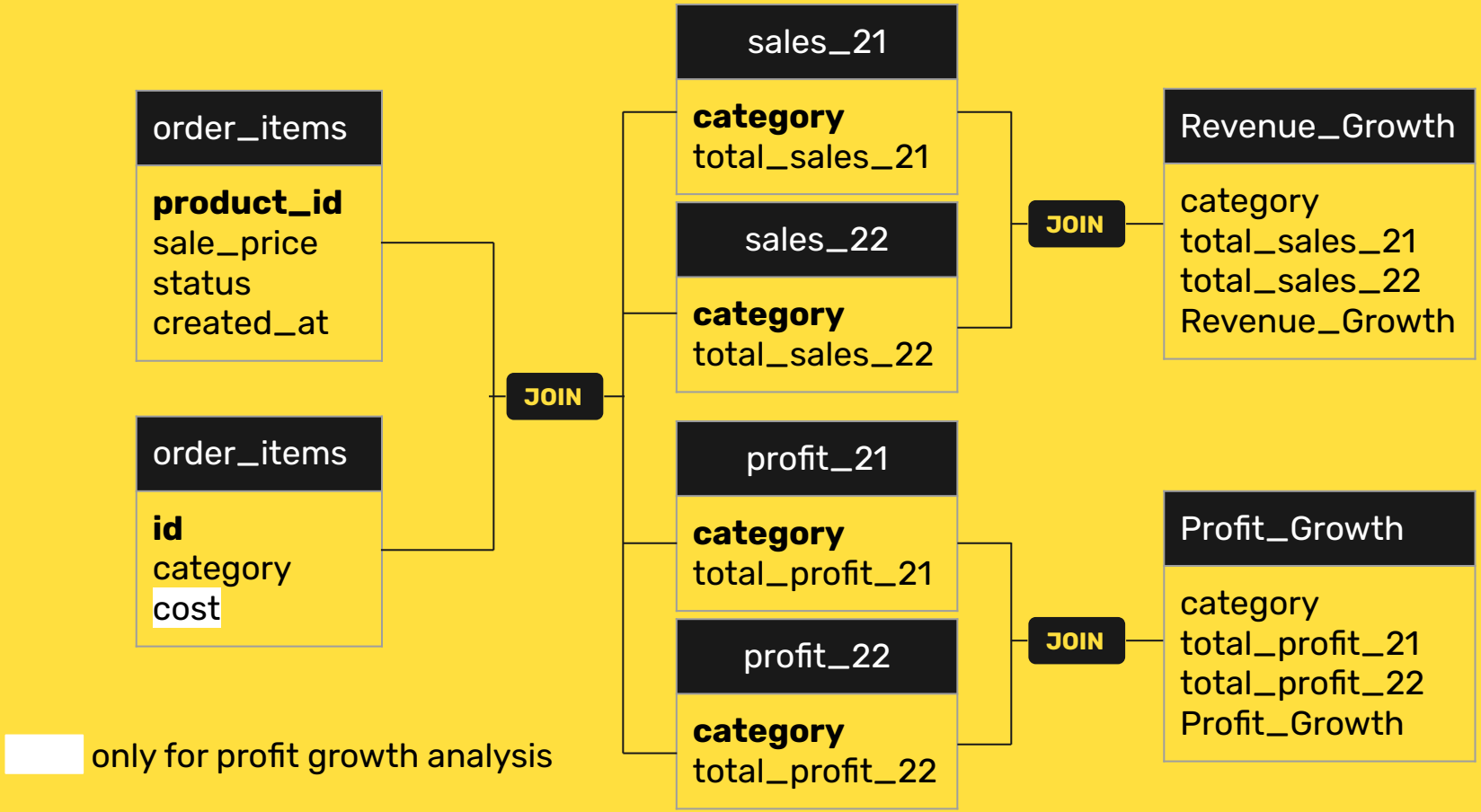
Revenue

The amount of money paid by customer for the products.
(sale price)

Profit

The actual amount that company gained by selling a product.
(sale price - cost)

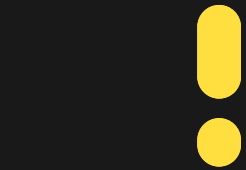
Growth Analysis: Query Process



SQL query: Revenue Growth

```sql
--What is the categories with the lowest revenue growth in the past 1 year?--
WITH sales_22 AS(
    SELECT category, SUM(sale_price) AS total_sales_22
    FROM `sql-project-376612.thelook_ecommerce.order_items` AS oi_22
    JOIN `sql-project-376612.thelook_ecommerce.products` AS pr_22
    ON pr_22.id=product_id
    WHERE status='Complete' AND EXTRACT(YEAR FROM created_at)=2022
    GROUP BY category
    ORDER BY total_sales_22),
  sales_21 AS (
    SELECT category, SUM(sale_price) AS total_sales_21
    FROM `sql-project-376612.thelook_ecommerce.order_items` AS oi_21
    JOIN `sql-project-376612.thelook_ecommerce.products` AS pr_21
    ON pr_21.id=product_id
    WHERE status='Complete' AND EXTRACT(YEAR FROM created_at)=2021
    GROUP BY category
    ORDER BY total_sales_21)

SELECT sales_22.category AS category,
  total_sales_21,
  total_sales_22,
  ROUND(((total_sales_22-total_sales_21)/total_sales_21*100),2) AS Revenue_Growth
FROM sales_22 JOIN sales_21
ON sales_22.category=sales_21.category
ORDER BY Revenue_Growth ASC;
```



Query Result: Revenue Growth

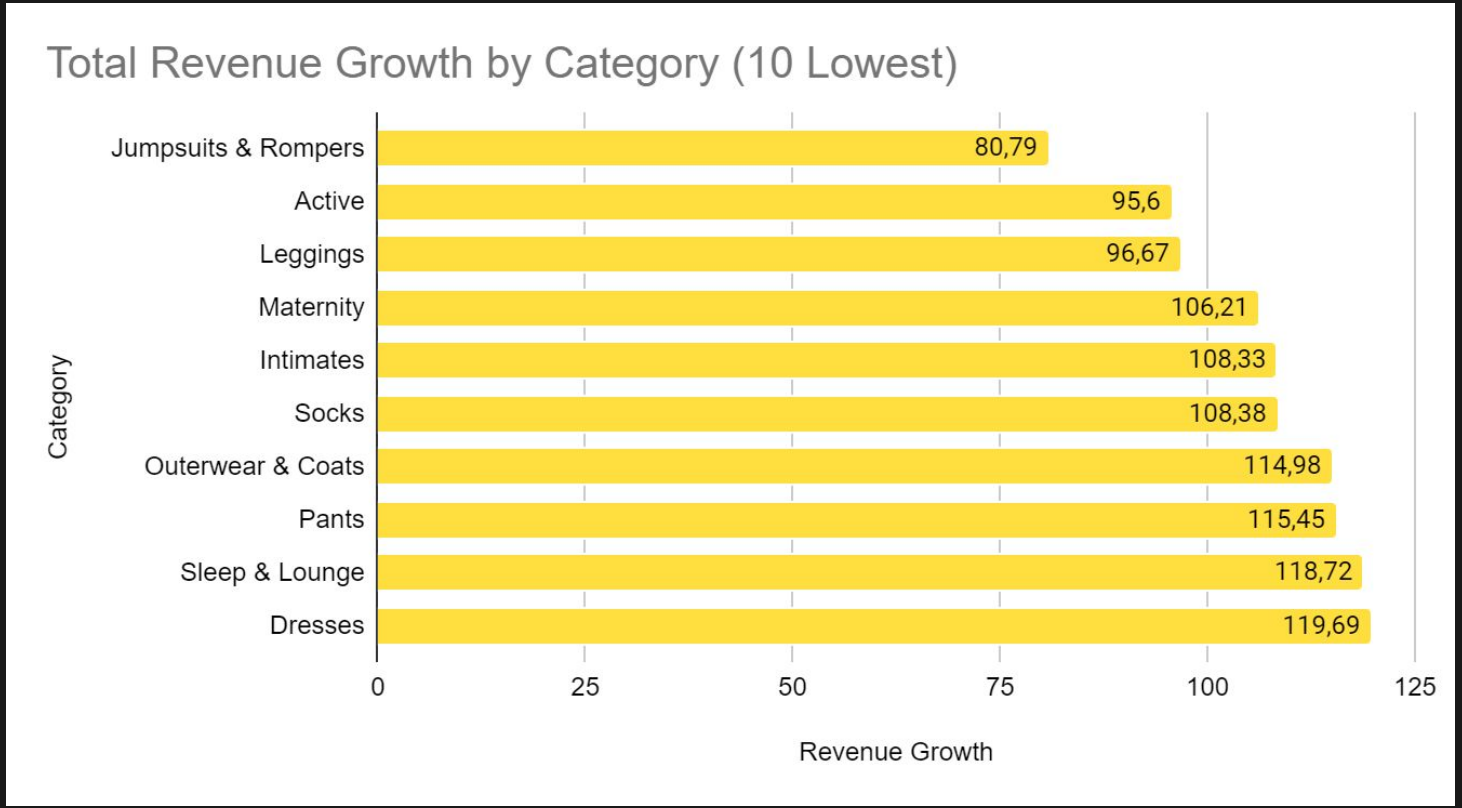
Schema

Field name	Type	Mode
category	STRING	NULLABLE
total_sales_21	FLOAT	NULLABLE
total_sales_22	FLOAT	NULLABLE
Revenue_Growth	FLOAT	NULLABLE

Preview

category	total_sales_21	total_sales_22	Revenue_Growth
Jumpsuits & Rompers	2160.29999...	3905.64001...	80.79
Active	27755.6900...	54291.3900...	95.6
Leggings	5392.90998...	10606.1199...	96.67
Maternity	15622.3999...	32215.2499...	106.21
Intimates	25945.4700...	54051.4900...	108.33
Socks	6710.26997...	13983.0500...	108.38
Outerwear & Coats	80561.2999...	173189.649...	114.98
Pants	23860.9300...	51408.2402...	115.45
Sleep & Lounge	31142.0700...	68112.4001...	118.72
Dresses	27359.3901...	60105.9601...	119.69

Revenue Growth



SQL query: Profit Growth

```sql
--What is the categories with the lowest profit growth in the past 1 year?--
WITH profit_22 AS(
    SELECT category, SUM(sale_price-cost) total_profit_22
    FROM `sql-project-376612.thelook_ecommerce.order_items` AS oi_22
    JOIN `sql-project-376612.thelook_ecommerce.products` AS pr_22
    ON pr_22.id=product_id
    WHERE status='Complete' AND EXTRACT(YEAR FROM created_at)=2022
    GROUP BY category),
  profit_21 AS(
    SELECT category, SUM(sale_price-cost) total_profit_21
    FROM `sql-project-376612.thelook_ecommerce.order_items` AS oi_21
    JOIN `sql-project-376612.thelook_ecommerce.products` AS pr_21
    ON pr_21.id=product_id
    WHERE status='Complete' AND EXTRACT(YEAR FROM created_at)=2021
    GROUP BY category)

SELECT profit_22.category AS category,
  total_profit_21,
  total_profit_22,
  ROUND(((total_profit_22-total_profit_21)/total_profit_21*100),2) AS Profit_Growth
FROM profit_22 JOIN profit_21
ON profit_22.category=profit_21.category
ORDER BY Profit_Growth ASC;
```

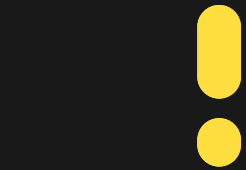
Query Result: Profit Growth

Schema

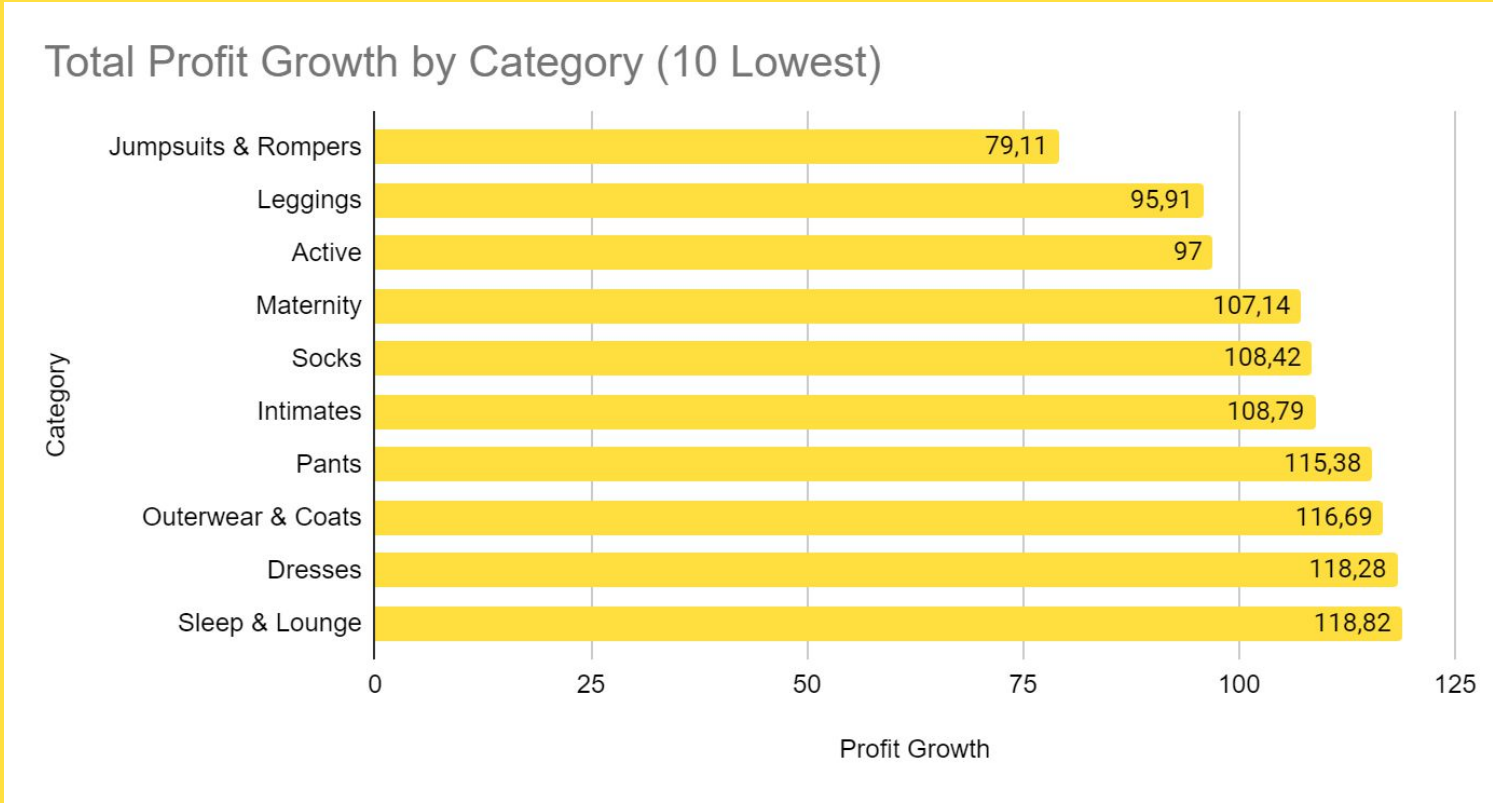
Field name	Type	Mode
category	STRING	NULLABLE
total_profit_21	FLOAT	NULLABLE
total_profit_22	FLOAT	NULLABLE
Profit_Growth	FLOAT	NULLABLE

Preview

category	total_profit_21	total_profit_22	Profit_Growth
Jumpsuits & Rompers	1006.24884...	1802.26509...	79.11
Leggings	2154.83639...	4221.60657...	95.91
Active	16073.3203...	31664.7270...	97.0
Maternity	8702.41303...	18025.7682...	107.14
Socks	2653.19413...	5529.80387...	108.42
Intimates	12138.5897...	25343.6781...	108.79
Pants	12885.7278...	27753.7556...	115.38
Outerwear & Coats	44449.2895...	96316.9253...	116.69
Dresses	15087.9395...	32933.9813...	118.28
Sleep & Lounge	15998.2584...	35007.6880...	118.82



Profit Growth



10 Lowest Revenue & Profit Growth

Revenue	Profit
Jumpsuits & Rompers	Jumpsuits & Rompers
Leggings	Active
Active	Leggings
Maternity	Maternity
Socks	Intimates
Intimates	Socks
Pants	Outerwear & Coats
Outerwear & Coats	Pants
Dresses	Sleep & Lounge
Sleep & Lounge	Dresses

From this comparison table, we can see that the 10 lowest growing product categories are almost similar and differ only slightly in rankings.

Jumpsuits & Rompers are the category with the lowest revenue and profit growth. For this reason, this category might be the best choice to deprioritize.

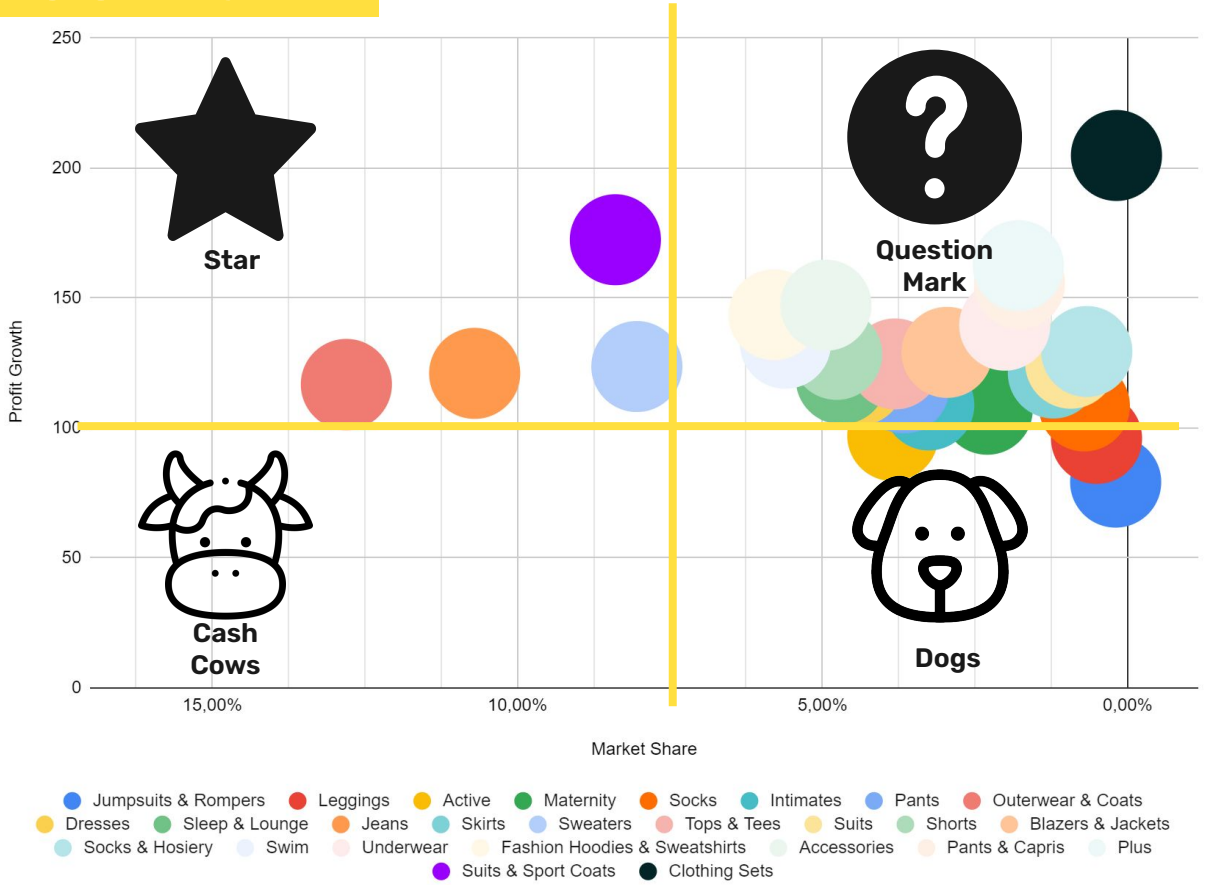
Further Analysis: BCG Matrix

To provide deeper insights into the development of our product category, we use the BCG Matrix to project our results. By using the BCG Matrix, we can analyze our product portfolio holistically and make strategic decisions regarding the placement of resources. This helps us to evaluate our products and determine which ones to invest, divest or retain.

The BCG Matrix categorizes our products into four different categories based on their market share and growth rate.

Market Share = (profit of each category/total profit)

BCG Matrix


250
200
150
100
50
0
Profit Growth
15,00%
10,00%
5,00%
0,00%
Market Share
Star
Question Mark
Cash Cows
Dogs
Jumpsuits & Rompers
Leggings
Active
Maternity
Socks
Intimates
Pants
Outerwear & Coats
Dresses
Sleep & Lounge
Jeans
Skirts
Sweaters
Tops & Tees
Suits
Shorts
Blazers & Jackets
Socks & Hosiery
Swim
Underwear
Fashion Hoodies & Sweatshirts
Accessories
Pants & Capris
Plus
Suits & Sport Coats
Clothing Sets

Insights and Recommendation

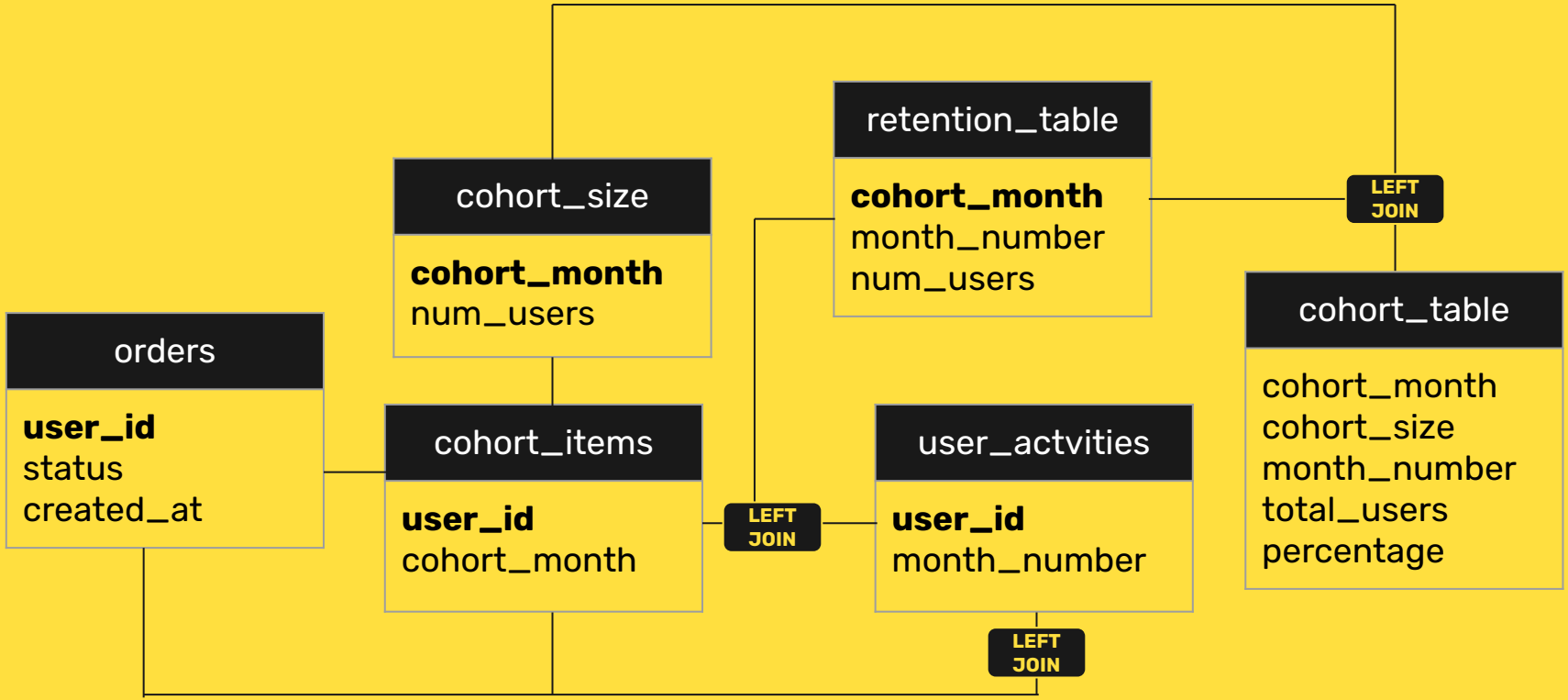
Deprioritize	Highly Potential
From the BCG Matrix, we can see that the categories Jumpsuits & Rompers, Active, and Leggings fall into the 'dog' area. For this reason, we can lower the priority for this category.	Then the Outerwear & Coats, Jeans, and Sweater categories are very close to the cash cows area. We can maximize this by making it the best source of income.

03 User Retention Analysis

User Retention: Cohort Analysis

Cohort analysis will be conducted to gather insights about user retention based on their purchasing behaviour. We will analyze users who have **successfully completed their orders for the first time, in which the order was made in 2022**.

Cohort Analysis: Query Process



SQL query: Cohort Analysis

```sql
--Cohort Analysis of users who first purchased in 2022--
WITH
  --bucketing users into cohort, based on the time of their first sucsessful purchase--
  cohort_items AS (
    SELECT
      user_id,
      MIN(DATE(DATE_TRUNC(created_at,MONTH))) AS cohort_month --finding first order date,
converted to be grouped into month
    FROM `sql-project-376612.thelook_ecommerce.orders`
    WHERE status='Complete' --//Filter only completed order
    GROUP BY user_id),

  --Finding "Cohort Size" (how many users made their first purchase in each month)--
  cohort_size AS (
    SELECT cohort_month, COUNT(user_id) AS num_users
    FROM cohort_items
    GROUP BY cohort_month),
```

SQL query: Cohort Analysis

```
--Check for user order activity--
  user_activities AS(
    SELECT
      orders.user_id AS user_id,
      DATE_DIFF(DATE(DATE_TRUNC(created_at,MONTH)),cohort.cohort_month,MONTH) AS
month_number
    FROM `sql-project-376612.thelook_ecommerce.orders` orders
      LEFT JOIN cohort_items AS cohort
      ON orders.user_id = cohort.user_id
    WHERE
      EXTRACT(year FROM cohort.cohort_month)=2022 --//Filter user with first purchase on
2022
      AND EXTRACT(YEAR FROM created_at)=2022 --//Filter user who ordered in 2022
    GROUP BY 1, 2),
```

SQL query: Cohort Analysis

```
--Show user retention table (by Number of Users)--
 retention_table as (
   SELECT
     cohort_month,
     month_number,
     COUNT(user_activities.user_id) AS num_users
   FROM user_activities
     LEFT JOIN cohort_items
     ON user_activities.user_id = cohort_items.user_id
   GROUP BY cohort_month, month_number)
```

SQL query: Cohort Analysis

```
--Show final cohort analysis (cohort_month, size, month_number, percentage)--
SELECT
  retention_table.cohort_month AS cohort_month,
  cohort_size.num_users AS cohort_size,
  month_number,
  retention_table.num_users AS total_users,
  CAST(retention_table.num_users AS decimal)/cohort_size.num_users AS percentage
FROM retention_table
  LEFT JOIN cohort_size
  ON retention_table.cohort_month = cohort_size.cohort_month
WHERE retention_table.cohort_month IS NOT NULL AND month_number>=0
ORDER BY cohort_month, month_number;
```

Query Result: Cohort Analysis

Schema

Field name	Type	Mode
cohort_month	DATE	NULLABLE
cohort_size	INTEGER	NULLABLE
month_number	INTEGER	NULLABLE
total_users	INTEGER	NULLABLE
percentage	NUMERIC	NULLABLE

Preview

cohort_month	cohort_size	month_number	total_users	percentage
2022-08-01	1208	0	1208	1
2022-08-01	1208	1	84	0.069536424
2022-08-01	1208	2	96	0.079470199
2022-08-01	1208	3	105	0.08692053
2022-08-01	1208	4	96	0.079470199
2022-09-01	1329	0	1329	1
2022-09-01	1329	1	108	0.081264108
2022-09-01	1329	2	110	0.082768999
2022-09-01	1329	3	112	0.08427389
2022-01-01	779	0	779	1

Cohort Chart: Number of User

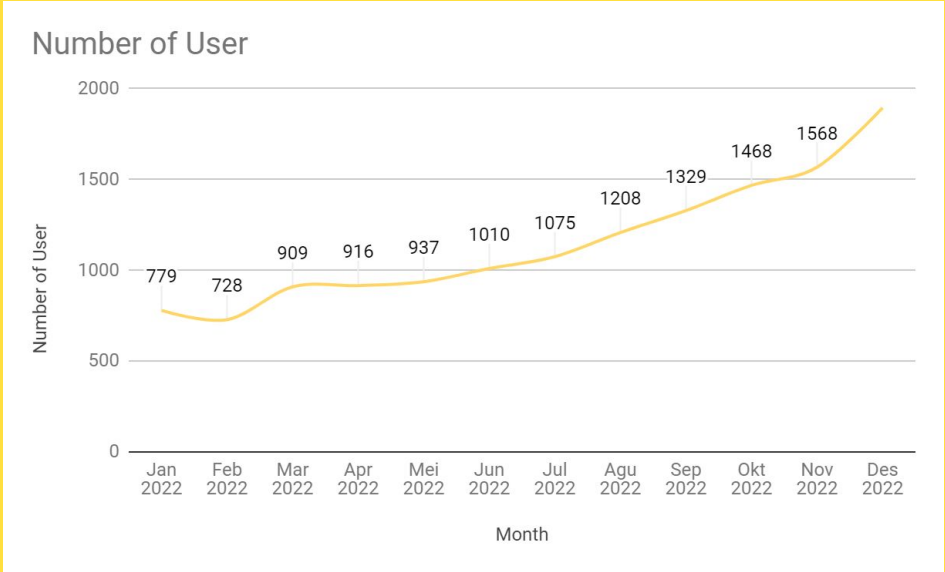
Month	0	1	2	3	4	5	6	7	8	9	10	11
Jan 2022	779	25	37	30	29	35	34	23	32	38	28	36
Feb 2022	728	41	24	33	34	42	36	36	43	40	33	
Mar 2022	909	48	47	39	41	58	43	52	56	53		
Apr 2022	916	46	49	57	60	52	48	52	52			
Mei 2022	937	47	59	53	53	57	69	62				
Jun 2022	1010	65	56	53	60	52	60					
Jul 2022	1075	84	73	75	69	72						
Agu 2022	1208	84	96	105	96							
Sep 2022	1329	108	110	112								
Okt 2022	1468	144	142									
Nov 2022	1568	200										
Des 2022	1895											

Query Result: Cohort Analysis

Month	Cohort Size	0	1	2	3	4	5	6	7	8	9	10	11
Jan 2022	779	100,00%	3,21%	4,75%	3,85%	3,72%	4,49%	4,36%	2,95%	4,11%	4,88%	3,59%	4,62%
Feb 2022	728	100,00%	5,63%	3,30%	4,53%	4,67%	5,77%	4,95%	4,95%	5,91%	5,49%	4,53%	
Mar 2022	909	100,00%	5,28%	5,17%	4,29%	4,51%	6,38%	4,73%	5,72%	6,16%	5,83%		
Apr 2022	916	100,00%	5,02%	5,35%	6,22%	6,55%	5,68%	5,24%	5,68%	5,68%			
Mei 2022	937	100,00%	5,02%	6,30%	5,66%	5,66%	6,08%	7,36%	6,62%				
Jun 2022	1010	100,00%	6,44%	5,54%	5,25%	5,94%	5,15%	5,94%					
Jul 2022	1075	100,00%	7,81%	6,79%	6,98%	6,42%	6,70%						
Agu 2022	1208	100,00%	6,95%	7,95%	8,69%	7,95%							
Sep 2022	1329	100,00%	8,13%	8,28%	8,43%								
Okt 2022	1468	100,00%	9,81%	9,67%									
Nov 2022	1568	100,00%	12,76%										
Des 2022	1895	100,00%											

Insights and Recommendation

Month	Cohort Size
Jan 2022	779
Feb 2022	728
Mar 2022	909
Apr 2022	916
Mei 2022	937
Jun 2022	1010
Jul 2022	1075
Agu 2022	1208
Sep 2022	1329
Okt 2022	1468
Nov 2022	1568
Des 2022	1895



From the cohort size/number of users, it can be seen that the number of users who have successfully made continuous orders in 2022 **has generally continued to increase**. This may be due to an **effective user acquisition strategy, therefore we can maximize and improve this strategy**.

Insights and Recommendation

Month	0	1
Jan 2022	779	25
Feb 2022	728	41
Mar 2022	909	48
Apr 2022	916	46
Mei 2022	937	47
Jun 2022	1010	65
Jul 2022	1075	84
Agu 2022	1208	84
Sep 2022	1329	108
Okt 2022	1468	144
Nov 2022	1568	200
Des 2022	1895	

If you look at it, very few users place another order the following month after their first purchase. In January, the number of users who bought again in the following month **was only 3.21%, which is the month with the lowest repeat orders in 2022**.

In terms of increasing repeat orders for first time buyers, **we can do promotions such as giving coupons for first time buyers that can entice users to buy regularly**.

Insights and Recommendation

Month	0	1	2	3	4	5	6	7	8	9	10	11
Jan 2022	100,00%	3,21%	4,75%	3,85%	3,72%	4,49%	4,36%	2,95%	4,11%	4,88%	3,59%	4,62%
Feb 2022	100,00%	5,63%	3,30%	4,53%	4,67%	5,77%	4,95%	4,95%	5,91%	5,49%	4,53%	
Mar 2022	100,00%	5,28%	5,17%	4,29%	4,51%	6,38%	4,73%	5,72%	6,16%	5,83%		
Apr 2022	100,00%	5,02%	5,35%	6,22%	6,55%	5,68%	5,24%	5,68%	5,68%			
Mei 2022	100,00%	5,02%	6,30%	5,66%	5,66%	6,08%	7,36%	6,62%				
Jun 2022	100,00%	6,44%	5,54%	5,25%	5,94%	5,15%	5,94%					
Jul 2022	100,00%	7,81%	6,79%	6,98%	6,42%	6,70%						
Agu 2022	100,00%	6,95%	7,95%	8,69%	7,95%							
Sep 2022	100,00%	8,13%	8,28%	8,43%								
Okt 2022	100,00%	9,81%	9,67%									
Nov 2022	100,00%	12,76%										
Des 2022	100,00%											

If seen from **January to June**, visually in that month there **was a pretty bad decline in repeat orders**. However, from **July to December**, the decline in repeat orders **was not too bad**. Therefore, **we can focus more on new users from July to December so they can be more active in making transactions**.

04 Recommendation

Recommendation

Deprioritize	Prioritize	Repeat Order	Existing User
We can deprioritize the product categories for Jumpsuits & Rompers, Active, and Leggings.	We can invest further in the Outerwear & Coats, Jeans, and Sweaters categories.	To increase the number of repeat orders, we can carry out promotions such as giving coupons to first buyers in order to attract user transaction interest.	To increase the number of repeat orders, we can focus more on existing users by giving discounts or other promos to increase user transactions.

Thanks!

Contact me:

rafiqnaufal97@gmail.com
+62 812 809 05778
linkedin.com/in/rafiqnaufal