

2025

# Cookie Media Marketing Analytics Report

CONVERSION, ENGAGEMENT, REVIEW IMPROVEMENT  
FRANCIS AUNG THU

## Contents

1.	Project Overview .....	1
2.	Dataset Summary.....	1
3.	Methodology.....	1
3.1	Data Understanding .....	1
3.2	Data Cleaning, Preparation, and Feature Engineering .....	2
4.	Business Goals and Analytical Focus.....	2
5.	Data Analysis using SQL (Business Transactions) .....	3
5.1	Customer Behaviour.....	3
5.2	Product Performance & Sales Insight .....	3
5.3	Conversion.....	4
5.4	Campaign & Engagement Efficiency .....	5
5.5	Customer Review .....	6
6.	Bringing Everything Into Power BI .....	6
7.	Actionable Recommendations .....	9

## 1. Project Overview

Despite high marketing expenditures, CookieMedia is currently facing a **triple threat** to its growth: declining customer engagement, stagnant conversion rates, and a lack of alignment between marketing content and consumer feedback. Significant investments in digital campaigns are failing to deliver expected returns, indicating an urgent need for data-driven strategy optimization.

## 2. Dataset Summary

The dataset for the Cookie Media project consists of five interconnected tables (customers, geography, engagement data, customer reviews, customer journey) representing the end-to-end customer journey, from initial marketing engagement to post-purchase feedback. This relational structure allows for a 360-degree analysis of conversion friction and product performance.

## 3. Methodology

### 3.1 Data Understanding

The project began by exploring each table in the database to understand the information available for analysis. Column names, data types, and each table's purpose were reviewed to build a clear picture of how customer details, geographic information, engagement behaviour, reviews, and journey stages were structured. Key fields such as CustomerID, ProductID, and GeographyID were identified early

on, as they would serve as the primary connectors between tables during the analysis.

### 3.2 Data Cleaning, Preparation, and Feature Engineering

Before jumping into analysis, the dataset was reviewed and prepared to ensure consistency and reliability.

- Missing values were checked across key fields such as age, gender, review text, and engagement metrics.
- Date fields and categorical variables were standardised to maintain consistency during joins and aggregations.
- Customer review text was cleaned in SQL (lowercasing, removing noise, and preparing tokens) to support sentiment analysis.
- Feature engineering: Several additional calculated fields were created to support behavioural analysis, segmentation, and performance measurement. These included purchase-interval metrics (average, minimum, and maximum days between purchases), product-level metrics such as purchase counts, contribution percentages, cumulative percentages, Pareto categories, and ABC classifications. Customer-level indicators such as churn risk and customer segment labels were also generated to enable more targeted analysis and clearer insights.
- SQL views and staging tables were created to streamline repeated joins, simplify calculations, and organise intermediate outputs.
- Customer review texts were processed in Python using a sentiment intensity analyzer and classified into five categories: very positive, positive, neutral, negative, and very negative. In addition, an intent analysis was performed using keyword patterns to group reviews into categories such as return/refund intent, inquiry, purchase intent, complaint, and usage intent. The resulting labels were incorporated back into SQL for deeper insights.

## 4. Business Goals and Analytical Focus

The analysis was structured around three key business goals:

1. Increase conversion rates

**Goal:** Identify factors influencing conversion rates and provide recommendations to improve them.

**Finding:** “Hero Products” were identified at the country level, enabling targeted strategies to boost conversion rates in specific markets

2. Enhance customer engagement

**Goal:** Determine which types of content drives the highest engagement

**Finding:** Campaigns and content types generating the strongest engagement were identified, highlighting where marketing efforts are most effective

### 3. Improve customer ratings

**Goal:** Understand common themes in customer reviews and provide actionable insights.

**Finding:** Customer sentiment was analysed at the product level, revealing areas for product and service improvement.

## 5. Data Analysis using SQL (Business Transactions)

### 5.1 Customer Behaviour

Q1: Customer Gender Distribution

	Gender	PersonCount
1	Female	54
2	Male	46

Q2: Which countries and regions have the most customers?

	Market_Region	Country	PersonCount	PercentageTotal	Country_Rank
1	Southern Europe	Spain	18	18	1
2	Southern Europe	Italy	12	12	2
3	DACH	Germany	11	11	3
4	Western Europe	UK	10	10	4
5	DACH	Austria	10	10	4
6	Benelux	Belgium	9	9	5
7	Benelux	Netherlands	9	9	5
8	Other Europe	Sweden	8	8	6
9	DACH	Switzerland	8	8	6
10	Western Europe	France	5	5	7

### 5.2 Product Performance & Sales Insight

Q3: Top 3 countries by purchase count

	Market_Region	Country	PurchaseCount
1	Southern Europe	Spain	34
2	DACH	Germany	26
3	Western Europe	UK	24

Customers take around 10 months from first interaction to first purchase—which is quite long.

Q4: What is the most purchased product?

	ProductName	↑↓	PurchaseCount	↑↓
1	Hockey Stick		15	
2	Ice Skates		13	
3	Ski Boots		13	
4	Baseball Glove		13	
5	Climbing Rope		12	
6	Cycling Helmet		12	
7	Surfboard		12	
8	Fitness Tracker		11	
9	Golf Clubs		10	
10	Kayak		10	

### 5.3 Conversion

Q5: What is the average Days\_to\_First\_Purchase?

	Average_Days_to_First_Purchase	↑↓
1	307	

Q6: Loyal & Quick Customer. What was their most common First\_Purchase\_Date product? Can we use that product as a 'Hero Product' to acquire more conversion users?

	ProductID	↑↓	ProductName	↑↓	PurchaseCount	↑↓	ContributionPct	↑↓	CumulativePct	↑↓	ParetoCategory	↑↓	ABC_Class	↑↓
1	15		Climbing Rope		4		13		12		Top 80%		A	
2	20		Boxing Gloves		4		13		25		Top 80%		A	
3	19		Hockey Stick		3		9		34		Top 80%		A	
4	3		Yoga Mat		3		9		43		Top 80%		A	
5	10		Golf Clubs		3		9		53		Top 80%		A	
6	1		Running Shoes		2		6		59		Top 80%		A	
7	4		Dumbbells		2		6		65		Top 80%		A	
8	14		Cycling Helmet		2		6		71		Top 80%		B	
9	9		Baseball Glove		2		6		78		Top 80%		B	
10	13		Swim Goggles		1		3		81		Bottom 20%		B	
11	17		Surfboard		1		3		84		Bottom 20%		B	
12	5		Soccer Ball		1		3		87		Bottom 20%		B	
13	6		Tennis Racket		1		3		90		Bottom 20%		B	
14	7		Basketball		1		3		93		Bottom 20%		C	

Top 80% and Bottom 20%, a business can optimize inventory. Prevent stockouts of high-velocity items, and marketing should focus on the underperforming products 'Bottom 20%'.

The top 6 products contribute 60% of the total sales. Those products can be hero products.

#### Q7: UK Top 5 selling products in cumulative sales percentage

	Country	ProductName	SalesCount	TotalCountrySales	SalesRank	CumulativeSalesPct
1	UK	Kayak	4	24	1	16
2	UK	Ice Skates	3	24	2	29
3	UK	Yoga Mat	2	24	3	37
4	UK	Surfboard	2	24	4	45
5	UK	Climbing Rope	2	24	5	54

#### 5.4 Campaign & Engagement Efficiency

##### Q8: Top 10 campaign IDs driving the highest quality traffic (conversion efficiency)?

	CampaignID	ContentType	PurchaseCount	Views	Clicks	LIKES	Purchase_per_100_views
1	17	socialmedia	11	10	0	0	110
2	11	newsletter	8	10	0	0	80
3	4	socialmedia	7	13	0	0	53
4	11	video	14	27	0	0	51
5	2	video	11	22	0	0	50
6	1	socialmedia	12	26	1	0	46
7	10	socialmedia	11	26	0	0	42
8	13	blog	10	24	1	0	41
9	17	blog	11	27	0	0	40
10	6	socialmedia	14	36	1	0	38

These campaigns reach a very high-intent audience, and targeting is very precise. Fewer views but very strong conversions. Highest-quality traffic sources.

##### Purchase per 100 views by content type

	ContentType	Purchase_per_100_views
1	newsletter	0.379512308391
2	video	0.378938042360
3	blog	0.354978937576
4	socialmedia	0.353802903096

Newsletter and video content deliver the highest conversion efficiency, generating approximately 0.38 purchases per 100 views. Blog and social media content convert slightly lower at around 0.35 purchases per 100 views. This suggests that newsletter and video formats attract more purchase-ready users.

Q9: Which ContentType (e.g., Blog vs. Video) leads to the most ViewsClicksCombined for our highest-priced items (Price > £350)?"

	ContentType	TotalViews	TotalClicks	TotalViewsClicksCombined
1	blog	528309	103574	631883
2	newsletter	373972	74610	448582
3	socialmedia	481141	99491	580632
4	video	381638	70425	452063

A blog is the most effective content type for the highest-priced items.

## 5.5 Customer Review

Q10: any correlation between customer sentiment in reviews count and their average days between purchases?

	Sentiment_Label	Avg_Days_Between_Purchases	ReviewCount
1	Very Negative	245	4
2	Negative	276	189
3	Neutral	332	246
4	Very Positive	301	223
5	Positive	270	102

Very negative and negative are highly engaged. They buy often, so they notice flaws more quickly. Neutral and very positive have the highest reviewcount. This indicates a healthy, active customer base, but the high Neutral count suggests a need for more "surprise and delight" marketing to move them toward Positive.

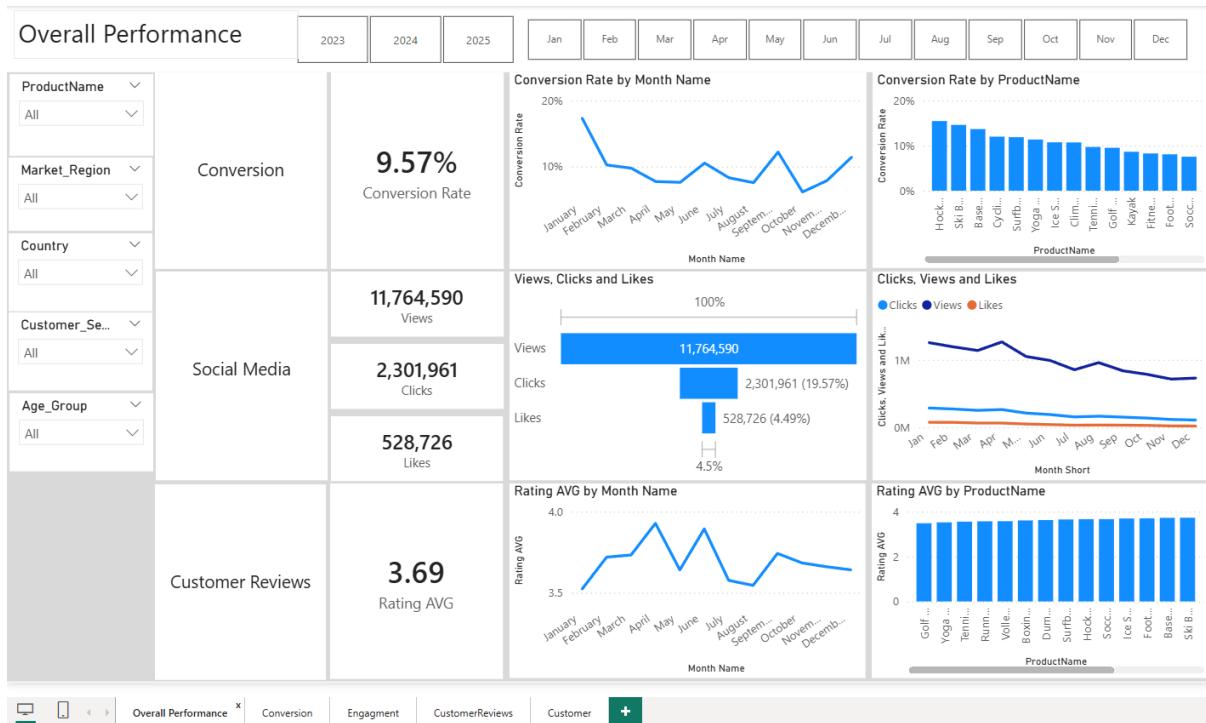
## 6. Bringing Everything Into Power BI

Once the SQL analysis was complete, the cleaned views and aggregated tables were imported into Power BI.

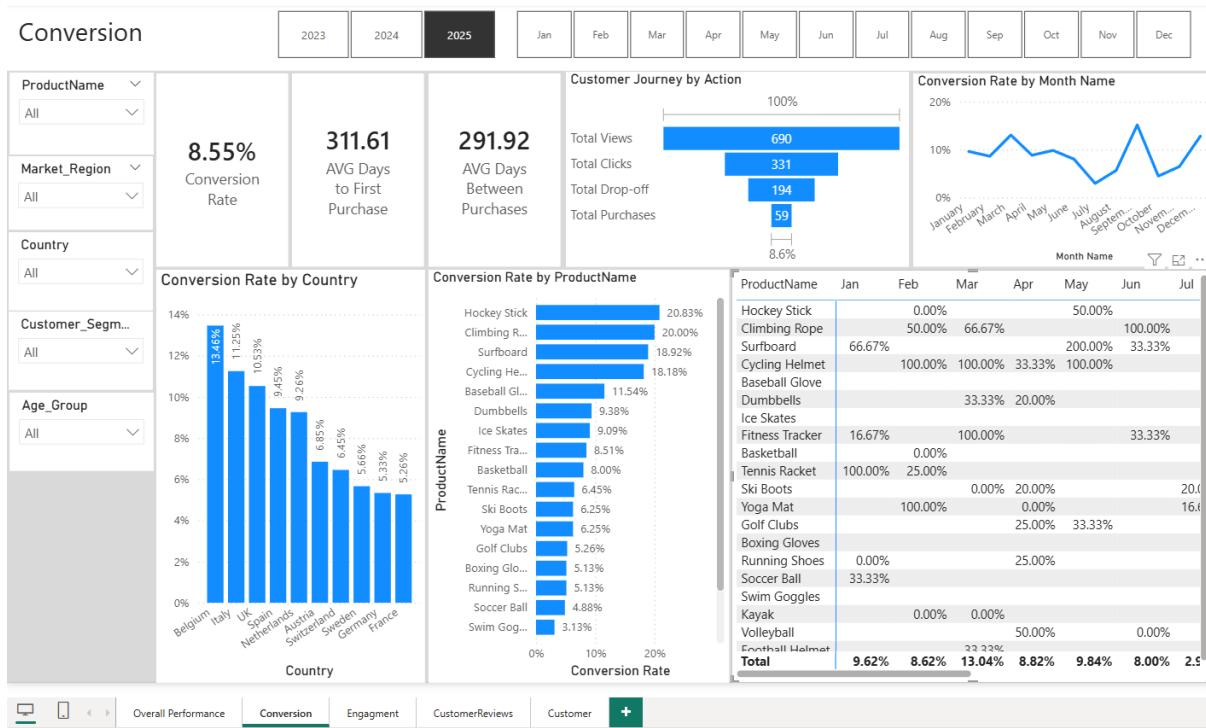
Dashboards were developed to visualise:

- Overall Performance
- Customer segments
- Social Media performance
- Customer Review
- Customer

Power BI served as the presentation layer, bringing together all SQL and Python outputs into an interactive, easy-to-navigate format for stakeholders.

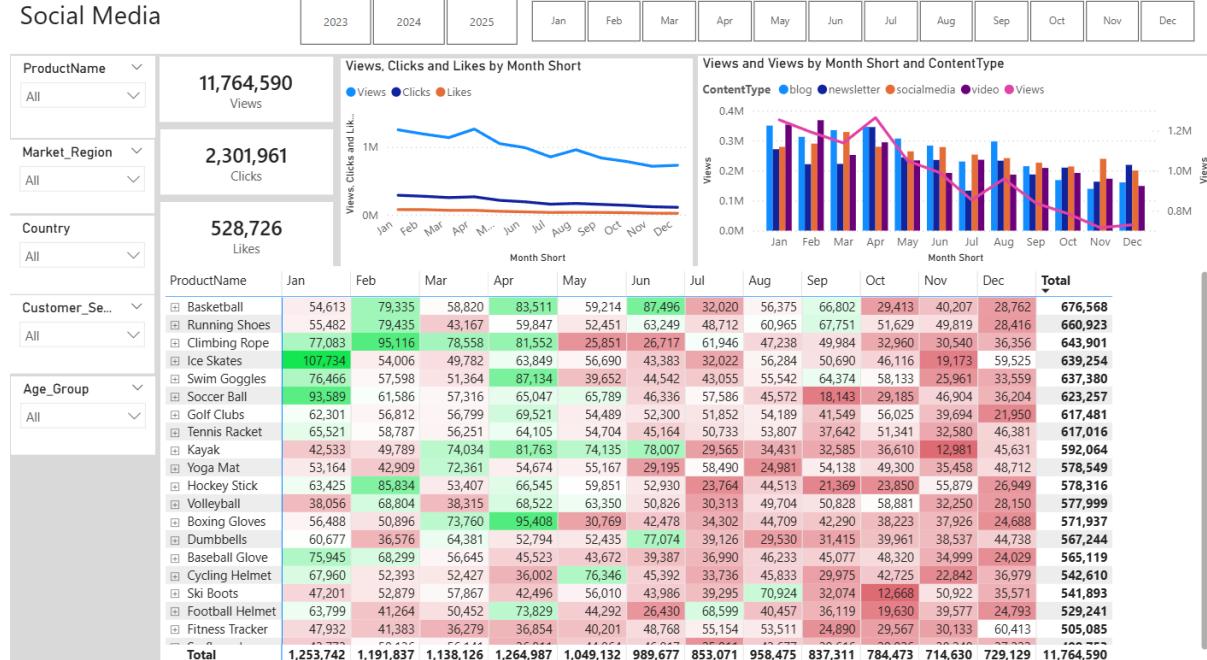


## Overall Performance Dashboard



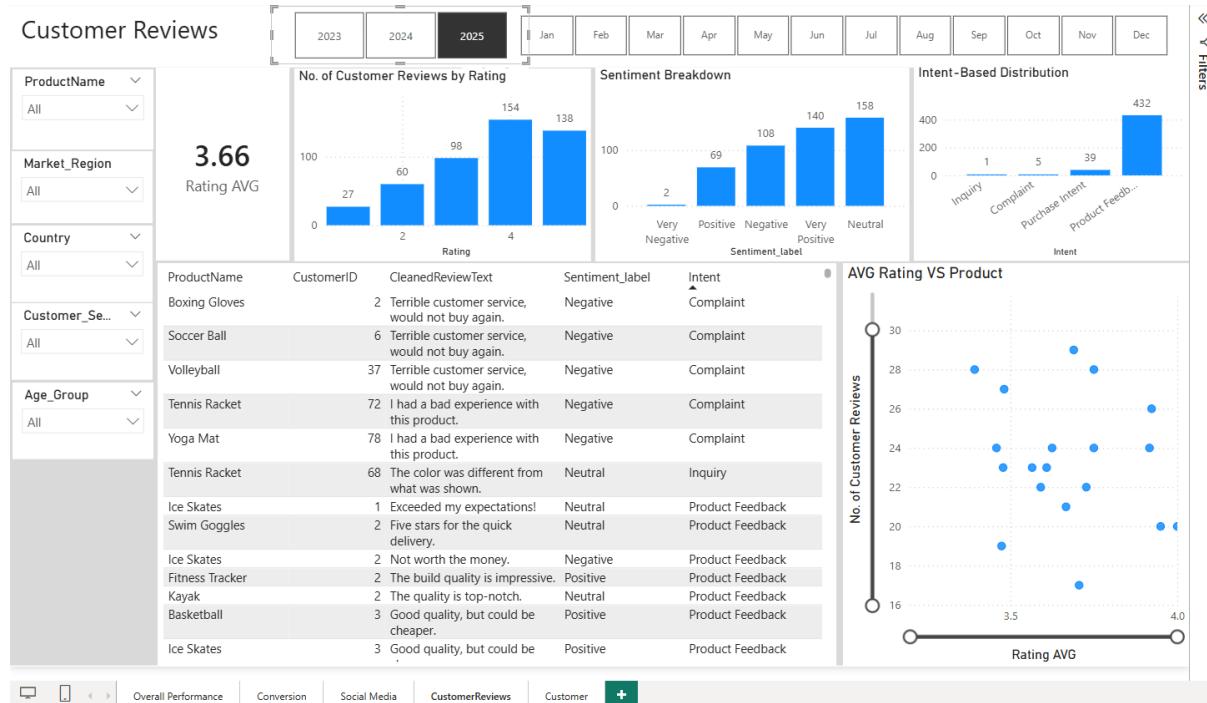
## Conversion Focus Dashboard

## Social Media

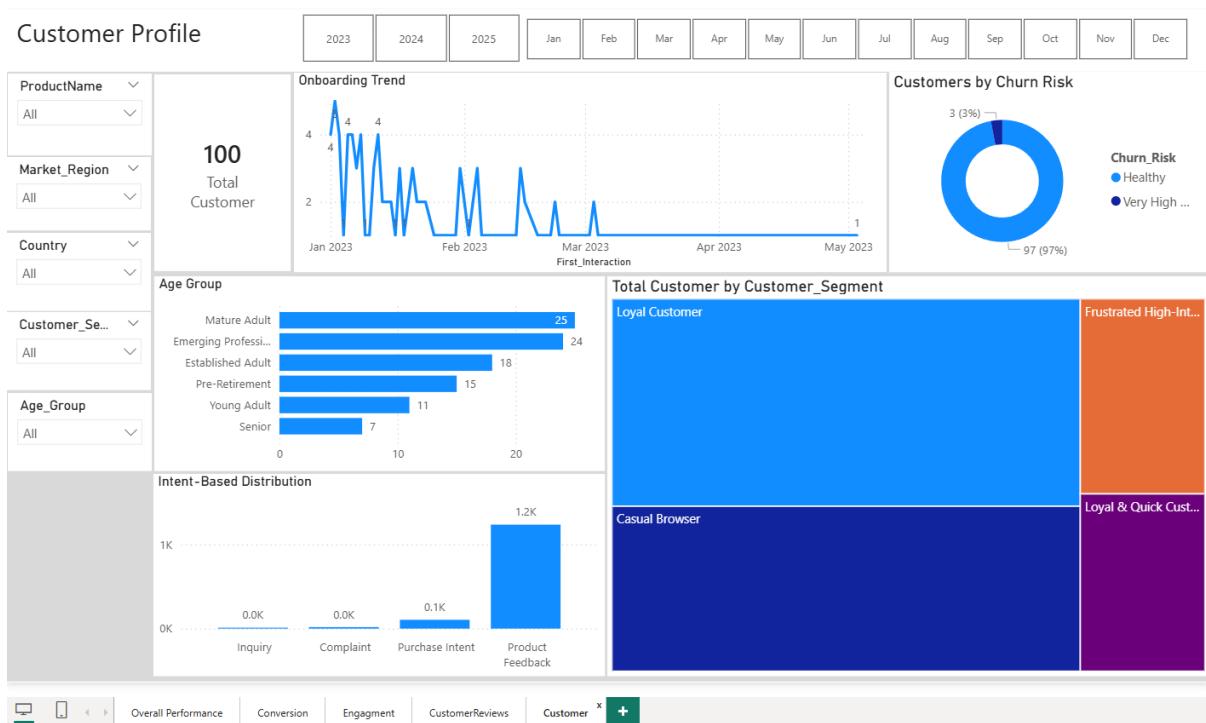


## Engagement Focus Dashboard

### Customer Reviews



## Customer Review Focus Dashboard



Customer Focus Dashboard

## 7. Actionable Recommendations

By bridging the gap between "what" customers are buying (Pareto) and "why" they are buying (Sentiment), this project provides the following roadmap:

1. Reduce Funnel Friction: Optimize the checkout experience for "Hero" products to recapture lost conversions.
2. Targeted Reinvestment: Shift marketing budget toward Newsletters and Video for general conversion, and Blogs for high-ticket items.
3. Reputation Management: Proactively address negative feedback on top-performing products to protect core revenue.
4. Inventory Optimization: Prioritize "Top 80%" items to prevent stockouts and focus marketing on the "Bottom 20%" to clear slow-moving stock.