<u>CUSTOMER BEHAVIOR ANALYSIS</u>

<u>CONTENT</u>

- Introduction
- Problem Definition
- Project Structure
- Dataset
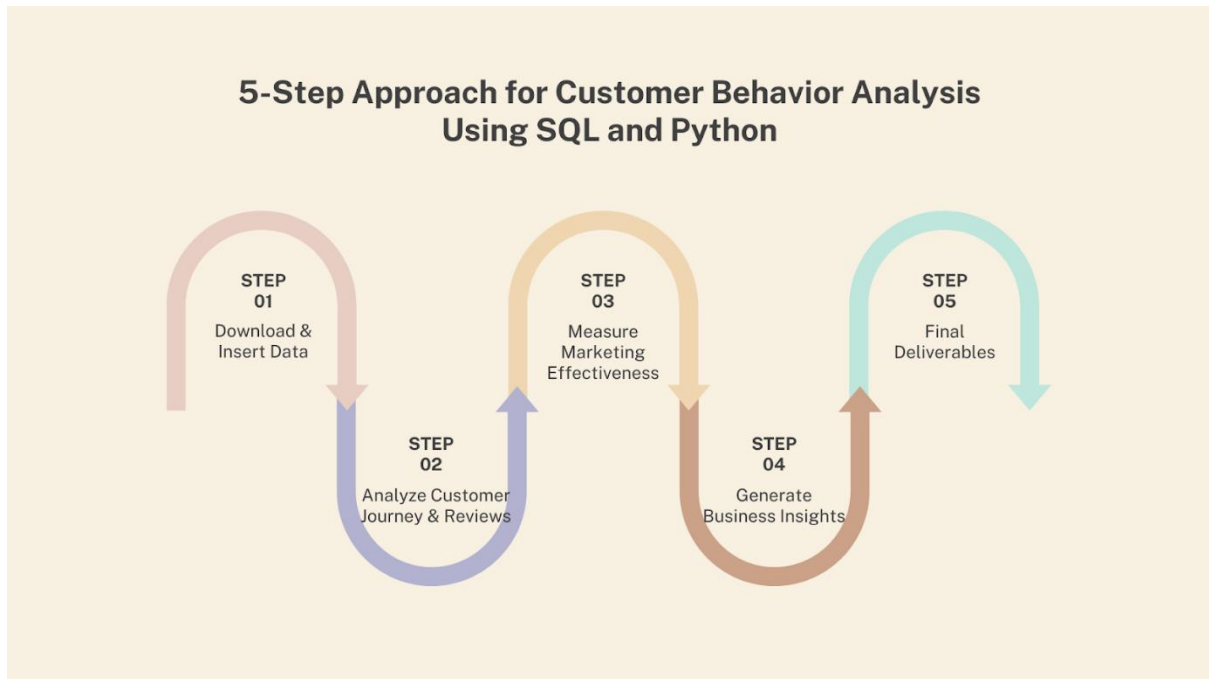- Languages
- Image
- Tables
- Chart

INTRODUCTION

Customer behavior analysis is the process of studying how customers interact with a business, including their purchasing patterns, preferences, and decision-making processes. The customer behaviour analysis will help to know about the customers' habits, such as how much they post on social media, when they look at your products and what drives them to buy, is all encompassed by the term "customer behaviour" It describes **how** customers shop, from how often they make a purchase to how they respond to your marketing campaigns. Understanding this behaviour will help you make predictions about what comes next and adapt your strategies for – from product to sales and marketing – to better reflect your customer's behaviour patterns. When you provide a customer experience that meets your customers' needs, you're more likely to drive sales and maximise your ROI.

This repository contains a comprehensive project to analyze the Customer behavior using machine learning techniques and various natural language processing techniques.

PROBLEM DEFINITION:

1. **Customer journey behavior** → Identify bottlenecks in the purchase process.

2. **Customer reviews & feedback** → Understand sentiment and satisfaction trends.

3. **Marketing effectiveness** → Measure the impact of engagement on conversion.

4. **Product & demographic analysis** → Identify high-performing products and customer segments.

PROJECT STRUCTURE:

5-Step Approach for Customer Behavior Analysis Using SQL and Python

STEP 01 - Download & Insert Data
STEP 02 - Analyze Customer Journey & Reviews
STEP 03 - Measure Marketing Effectiveness
STEP 04 - Generate Business Insights
STEP 05 - Final Deliverables

DATA SET:

1. customer_journey.csv - customer journey through different stages
2. customer_review.csv – detailed customer reviews on different product
3. customer.csv – customers detail
4. engagement_data.csv – factors which influencing customer engagement
5. geography.csv- customer and product based on location
6. product.csv – product description

LANGUAGES:

- Python
- SQL

Letter from the Marketing Manager (Jane Doe)

Subject: Request for Data Analysis to Improve Marketing Strategy

Hi Data Analyst,

I hope this email finds you well. I'm the Marketing Manager at ShopEasy, and we've been facing some serious challenges with our marketing campaigns lately. Despite increasing our investment, we've seen:
- A decline in customer engagement and site interactions
- A drop in conversion rates
- Rising marketing expenses with low ROI

We need a data-driven strategy to understand what's working and what's not. Your role is to analyze our customer journey data, product reviews, and conversion patterns to identify opportunities for improvement.

We believe the key questions to answer are:
- What factors are influencing customer engagement?
- At what stage are customers dropping off in their journey?
- How do customer reviews impact purchasing behavior?
- Which products, locations, and customer segments are performing well?

We have provided access to our databases containing customer behavior records, reviews, and transactions. Your insights will be critical in optimizing our marketing efforts and driving better business outcomes.

Looking forward to your analysis.

Best regards,

Jane Doe
Marketing Manager, ShopEasy

TABLES:

TABLE1: DURATION PER STAGE:

Stage	Average Duration
Checkout	45.05
Home page	160.53
Product Page	182.76

- **Checkout** (45.05s): Likely the fastest stage - users complete this quickly because they are not ready to purchase. They just check out whether the product is available.

- **Home page** (160.54s): **High engaged** stage. Users likely browse featured content, promotions, or navigation. Compare with bounce rates—if high, the page might need better UX.

- **Product Page** (182.77s): Longest duration - Users deeply evaluate products (reading descriptions, reviews, etc.) The duration Indicates interest or could signal confusing layouts or slow loading times. Check if this correlates with add-to-cart rates.

TABLE2: IDENTIFY THE DROP OFF POINT

Journey ID	Customer ID	Stage	Visit Date
1	64	Checkout	2024-06-10
2	94	Checkout	2025-07-09
4	33	Checkout	2025-05-28
11	23	Checkout	2023-08-22
21	58	Checkout	2024-03-07
24	77	Checkout	2025-06-25
32	67	Checkout	2023-09-04

37	38	Checkout	2024-06-18
47	30	Checkout	2024-02-18
52	15	Checkout	2025-08-07
54	9	Checkout	2024-11-19
55	1	Checkout	2024-03-17
71	40	Checkout	2023-07-13
93	43	Checkout	2023-08-22

Critical Drop-off Point: Checkout stage has an extremely high drop-off. 13 out of 15 customers who reached checkout abandoned the process. Only 2 customers completed purchases. To reduce the drop off,

- **Simplify Checkout:** Reduce form fields, enable guest checkout.

- **Transparent Pricing:** Show shipping costs earlier (e.g., on cart page).

- **Trust Signals:** Add SSL badges, money-back guarantees.

- **Exit-Intent Offers:** Discounts or free shipping to retain abandoning users.

TABLE3: OVERALL CUSTOMER RATING BASED ON PRODUCT:

Product ID	Product Name	Total Reviews	Average Rating	Minimum Rating	Maximum Rating	Price
8	Football Helmet	3	5.0000	5	5	327.36
19	Hockey Stick	5	4.4000	4	5	225.12
18	Running Shoes	3	4.0000	3	5	262.32
15	Soccer Ball	6	4.0000	2	5	410.17
1	Ski Boots	4	4.0000	3	5	472.32
11	Climbing Rope	6	4.0000	3	5	26.21
20	Volleyball	2	4.0000	4	4	196.68
5	Boxing Gloves	3	4.0000	4	4	44.75

17	Surfboard	9	3.8889	2	5	81.59
6	Tennis Racket	7	3.8571	1	5	173.83
3	Yoga Mat	7	3.7143	1	5	37.56
13	Swim Goggles	9	3.6667	3	4	259.4
2	Fitness Tracker	7	3.5714	2	5	223.75
10	Golf Clubs	4	3.5000	2	5	340.2
9	Baseball Glove	5	3.4000	3	4	41.26
16	Kayak	10	3.4000	2	4	275.43
4	Dumbbells	5	3.0000	1	5	145.97
12	Ice Skates	2	3.0000	3	3	36.07
7	Basketball	3	2.6667	2	3	42.8

Key Observations

Metric	Findings
Highest-Rated Products	Football Helmet (5.0), Hockey Stick (4.4), Running Shoes (4.0)
Lowest-Rated Products	Basketball (2.67), Dumbbells (3.0), Ice Skates (3.0)
Most Reviewed	Kayak (10), Surfboard (9), Swim Goggles (9) – Indicates high customer interest.
Least Reviewed	Volleyball (2), Ice Skates (2), Football Helmet (3) – May need more exposure.

Recommendations

1. **Promote High-Rated Products**:

   o Feature Football Helmet, Hockey Stick, and Running Shoes in campaigns.

   o Use testimonials like *"Rated 5/5 by customers!"*.

2. **Improve Low-Rated Products**:

   o **Basketball**: Investigate quality issues – consider supplier switch.

   o **Dumbbells/Ice Skates**: Add sizing guides or quality assurances.

3. **Encourage More Reviews**:

   o Offer discounts for reviews on low-count items (e.g., Volleyball).

TABLE6: BEST PERFORMING PRODUCT PER REGION:

Country	City	Product ID	Purchase Count
Austria	Vienna	1.0	3
Belgium	Brussels	2.0	2
France	Paris	4.0	1
Germany	Berlin	2.0	1
Italy	Rome	10.0	3
Netherlands	Amsterdam	15.0	2
Spain	Madrid	8.0	3
Sweden	Stockholm	12.0	1
Switzerland	Zurich	18.0	3
UK	London	19.0	2

Patterns & Trends

1. **Regional Preferences**:

   o **Winter Sports**: Ski Boots (Austria), Ice Skates (Sweden).

   o **Team Sports**: Football Helmet (Spain), Soccer Ball (Netherlands).

   o **Fitness Tech**: Fitness Tracker (Belgium, Germany).

   o **Lifestyle/Leisure**: Golf Clubs (Italy), Running Shoes (Switzerland).

2. **Purchase Count Discrepancies**:

   o **High Performers**: Ski Boots (3), Golf Clubs (3), Football Helmet (3), Running Shoes (3).

   o **Low Performers**: Dumbbells (1), Ice Skates (1) – may need targeted promotions.

Recommendations

- **Austria/Switzerland**: Bundle Ski Boots with winter gear (e.g., gloves, goggles).

- **Spain/Netherlands**: Partner with local football clubs to promote helmets/balls.

- **Italy**: Market Golf Clubs as luxury items (e.g., "VIP Golf Packages").

Boost Underperforming Regions

- **France (Dumbbells)**: Run fitness challenges with Parisian gyms.

- **Germany (Fitness Tracker)**: Collaborate with tech influencers in Berlin.

- **Sweden (Ice Skates)**: Launch winter-themed campaigns ahead of colder months.

Inventory & Marketing Alignment

- Stock more **Ski Boots** in Austria, **Running Shoes** in Switzerland.

- Test discounts on **Ice Skates** in Sweden if demand is seasonal.

Letter from the Customer Experience Manager (John Smith)

Subject: Request for Customer Feedback Analysis

Hi Data Analyst,

I'm the Customer Experience Manager at ShopEasy, and I need your help in analyzing our customer reviews and feedback. We've noticed a decline in customer satisfaction and engagement, which is negatively affecting our overall conversion rates.
Our main concerns include:
- Customers leaving negative reviews about certain products.
- Low repeat purchase rates, indicating dissatisfaction.
- A lack of structured insights into customer opinions.

We have collected customer reviews, ratings, and purchase history. Your role is to:
1. Analyze customer sentiment from the reviews.
2. Identify key complaints and improvement areas.
3. Find patterns between negative reviews and product performance.
4. Recommend strategies to enhance customer satisfaction.
5.

We hope your expertise in SQL and Python will provide actionable insights to improve customer experience and retention.

Best regards,
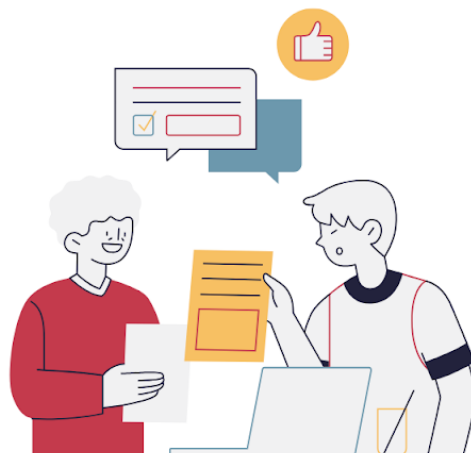John Smith
Customer Experience Manager, ShopEasy

TABLE4: KEY COMPLAINTS

Product ID	Product Name	Review Text	Rating
13	Swim Goggle	Average experience, nothing special.	3
17	Surfboard	Average experience, nothing special.	2
16	Kayak	Average experience, nothing special.	2
10	Golf Clubs	Disappointed with the performance.	2
2	Fitness Tracker	Product did not meet my expectations.	2
7	Basketball	The product is okay, but the instruction	2
15	Climbing Rope	The product arrived late.	2
3	Yoga Mat	Not worth the money.	1
4	Dumbbells	The product stopped working after a month.	1
6	Tennis Racket	I had a bad experience with this product.1.0	1

1. Product Quality Issues – Dumbbells, Tennis Racket and Basketball
2. Performance Dissatisfaction - Golf Clubs, Fitness Tracker, Yoga Mat
3. Delivery issue: Climbing Rope, Surfboard/Kayak

Recommendations:

- Upgrade materials; extend warranty. Strengthen frame; quality testing. Add illustrated manuals/videos.
- Offer demo units; highlight pro endorsements, Improve accuracy
- Switch to reliable shipping partners.
- Offer real-time tracking updates.
- Add unique features (e.g., anti-fog).

TABLE5: PATTERN BETWEEN NEGATIVE REVIEW AND PRODUCT PERFORMANCE

Rating	Product ID	Review Text	Clicks	View	Purchase
1	3	Not worth the money.	0	2	0
	4	The product stopped working after a month.	3	2	0
	6	I had a bad experience with this product.	1	2	0
2	2	Product did not meet my expectations.	2	3	1
	7	The product is okay, but the instructions were unclear.	0	0	0
	10	Disappointed with the performance.	2	5	0
	15	The product arrived late.	1	3	2
	16	Average experience, nothing special.	1	3	1
2.5	17	Average experience, nothing special.	0	2	2
3.0	13	Average experience, nothing special.	1	2	0

1. **Rating 1 = No Purchases**

   o Products with **1-star reviews consistently fail to convert** (0 purchases).

   o **Urgent Fix Needed**: Quality recalls or replacements.

2. **Rating 2 = Minimal Purchases**

   o Only **1-2 purchases** despite higher views (e.g., Fitness Tracker, Kayak).

   o **Risk of Returns**: Likely unsatisfied customers.

3. **Negative Reviews Kill Trust**

   o **Dumbbells**: 3 clicks but 0 buys – users research but abandon due to reviews.

   o **Golf Clubs**: 5 views but 0 buys – performance complaints deter sales.

4. **"Average Experience" = Low Engagement**

   o Products with generic negative feedback (**Swim Goggles, Surfboard**) get fewer clicks.

SENTIMENT ANALYSIS

TABLE7: COMMON ACTION LEADING TO SUCCESSFUL CONVERSION

Content Type	Total Likes	Total views & click	Purchase-count
Newsletter	1708	28755	11
Video	1552	19126	11
Social media	897	20423	7
Blog	290	22358	9

TABLE8: CONVERSION RATES BY CONTENT TYPE

Content Type	Purchase Count	Count	Conversion Rate (%)
Newsletter	11	17	64.705882
Video	11	31	35.483871
Blog	9	28	32.142857
Social media	7	24	29.166667

Newsletters Dominate

- Highest conversion rate (**64.7%**) despite moderate views.

- Suggests **highly targeted, intent-driven audience**.

Videos Drive Balanced Performance

- Second-highest conversions (11) with strong likes (1,552).

- **Engaging but less direct** than newsletters.

Blogs & Social Media Lag in Conversions

- Blogs get **most views (22,358)** but low conversions (32.1%).

- Social media has **decent likes (897)** but worst conversion rate (29.2%).

Recommendation:

1. **Double Down on Newsletters**: Allocate **40% of content budget** here (highest ROI).

2. **Optimize Videos for Sales**: Add **time-limited offers** in descriptions.

3. **Repurpose Blogs**: Turn top-performing articles into **video scripts**.

4. **Revamp Social Strategy**: Shift focus to **Reels/TikTok** for younger demographics.

CHARTS:

Conversion Rates by Content Type