MARKETING ANALYTICS BUSINESS CASE

Presented By : Subeena k.k

INTRODUCTION TO BUSINESS PROBLEM

ShopEasy, an online retail business, is facing reduced customer engagement and conversion rates despite launching several new online marketing campaigns. They are reaching out to you to help conduct a detailed analysis and identify areas for improvement in their marketing strategies.

.

Key Points:

- **Reduced Customer Engagement:** The number of customer interactions and engagement with the site and marketing content has declined.
- **Decreased Conversion Rates:** Fewer site visitors are converting into paying customers.
- **High Marketing Expenses:** Significant investments in marketing campaigns are not yielding expected returns.
- **Need for Customer Feedback Analysis:** Understanding customer opinions about products and services is crucial for improving engagement and conversions.

KEY PERFORMANCE INDICATORS (KPIS)

- **Conversion Rate:** Percentage of website visitors who make a purchase.
- **Customer Engagement Rate:** Level of interaction with marketing content (clicks, likes, comments).
- **Average Order Value (AOV):** Average amount spent by a customer per transaction.
- **Customer Feedback Score:** Average rating from customer reviews.

GOALS

1. **Increase Conversion Rates:**
   - **Goal:** Identify factors impacting the conversion rate and provide recommendations to improve it.
   - **Insight:** Highlight key stages where visitors drop off and suggest improvements to optimize the conversion funnel.

2. **Enhance Customer Engagement:**
   - **Goal:** Determine which types of content drive the highest engagement.
   - **Insight:** Analyze interaction levels with different types of marketing content to inform better content strategies.

3. **Improve Customer Feedback Scores:**
   - **Goal:** Understand common themes in customer reviews and provide actionable insights.
   - **Insight:** Identify recurring positive and negative feedback to guide product and service improvements.

DATA OVERVIEW

This project is built on a relational dataset designed to uncover insights into customer behavior, engagement, product performance, and regional impact. The following six tables form the foundation of this analysis:

- **customers:** Stores individual customer profiles, including demographic attributes such as name, age, gender, email, and geographic identifiers. It's the core table for segmentation and personalization efforts.

- **customer_journey:** Chronicles the customer's path through the marketing funnel using data like visit dates, funnel stages (e.g. Awareness, Consideration), actions taken, and duration. This helps map customer progression and campaign effectiveness.

- **customer_reviews:** Contains customer feedback in the form of numeric ratings and review text. By tying reviews to products and dates, this table enables sentiment analysis and product satisfaction assessment.

- **engagement_data:** Captures digital interaction metrics (such as clicks, views, and likes) across campaigns and products. It’s critical for understanding content performance and user engagement trends.

- **products:** Lists each product’s ID, name, category, and price. This table connects product attributes with engagement, purchase behavior, and customer feedback.

- **geography:** Defines regions or locations associated with customers or campaigns. It's essential for analyzing geographic patterns in engagement, preferences, and sales trends.

STEP 1: DATA PREPARATION AND SQL-BASED CLEANING

In the initial phase of the project, I focused on preparing and cleaning the data using SQL queries to ensure high-quality and analysis-ready records. This step involved the following tasks:

- **Product Categorization:** Products were classified into price-based segments (Low, Medium, High) using conditional logic. This helps simplify product-level analysis and supports customer segmentation based on price preference.

- **Geographic Enrichment:** A LEFT JOIN was performed between the customers and geography tables to attach regional data (city and country) to each customer profile. This integration enables geographic trend analysis and regional campaign insights.

- **Review Data Exploration and Cleaning:** The customer_reviews table was explored, and whitespace inconsistencies in the ReviewText column were cleaned using REPLACE to ensure consistency and improve results in sentiment analysis.

- **Normalization of Engagement Data:** Content type entries were cleaned for casing and spelling inconsistencies (e.g., converting "Socialmedia" to "Social Media"). Also, complex view-click pairs were split into separate Views and Clicks fields, and engagement dates were formatted for clarity. Newsletter-related records were filtered out to focus on active engagement data.

- **Duplicate Detection using CTEs:** A Common Table Expression (CTE) was used to identify duplicate customer journey records based on key fields. This helped ensure accuracy in tracking customer interactions and campaign stages.

- **Journey Data Standardization:** Duplicate entries were removed, marketing funnel Stage values were standardized to uppercase, and missing Duration values were imputed using the average duration for each VisitDate.

STEP 2: SENTIMENT ANALYSIS USING PYTHON AND VADER

In the second phase of the project, I conducted sentiment analysis on customer review text to enrich the data with emotional context. This step was implemented using Python and the VADER (Valence Aware Dictionary for Sentiment Reasoning) sentiment analysis tool from the NLTK package. The key actions in this stage included:

- **Data Retrieval from SQL Server:** Used SQLAlchemy to connect to the SQL Server and extract reviews from the customer_reviews table into a Pandas DataFrame for processing.

- **Sentiment Scoring with VADER:** Applied the VADER SentimentIntensityAnalyzer to compute compound sentiment scores for each review. These scores range from -1 (very negative) to +1 (very positive), providing a quantitative emotional measure.

- **Dual-Source Sentiment Categorization:** Developed a function to categorize reviews into labels like Positive, Negative, Mixed Positive, Mixed Negative, and Neutral based on both the compound sentiment score and the original customer rating (1–5 stars). This hybrid approach improved the reliability of sentiment tagging.

- **Sentiment Bucketing:** Grouped reviews into predefined sentiment score ranges (e.g., 0.5 to 1.0 for strongly positive) to facilitate analysis and visualization of overall sentiment trends.

- **Export of Enriched Data:** The resulting DataFrame—including sentiment scores, categories, and buckets—was exported to a CSV file named fact_customer_reviews_with_sentiment.csv for further use in dashboards or reporting.

STEP 3: DASHBOARD DEVELOPMENT AND INSIGHT GENERATION (POWER BI)

In the third phase, I developed a series of interactive dashboards using Power BI to translate raw data and analytical results into actionable marketing insights. These dashboards visualize customer behavior, campaign engagement, and product sentiment using data prepared in earlier steps.

1. Conversion Dashboard

Purpose: To analyze customer progression through the marketing funnel and identify top-performing products and time periods.

Key Measures:

- Conversion Rate
- Monthly Conversion Trend
- Conversion by Product & Stage

Insights:

- Overall conversion rate is 9.45%, peaking in January (16.46%) and dipping in September (7.33%).
- Hockey Stick had the highest product conversion rate (14.7%); Swim Goggles had the lowest (5.5%).
- The mid-funnel stages show high user drop-off, indicating potential improvement areas in campaign targeting or user experience.

2. Social Media Engagement Dashboard

Purpose: To evaluate content performance across views, clicks, and likes by content type and product.

Key Measures:

- Total Views, Clicks, and Likes
- Engagement Rate (Clicks/Views)
- Content Type & Product-wise Engagement

Insights:

- Engagement shows a clear decline from January to December, indicating potential campaign fatigue or need for content refresh.
- Blogs performed well in early months, while video content remained stable throughout the year.
- Basketball received the highest views (5,29,860), while Ice Skates received the lowest (1,04,966).

3. Customer Reviews Dashboard

Purpose: To assess customer satisfaction using sentiment analysis, ratings, and temporal feedback trends.

Key Measures:

- Average Rating
- Sentiment Category & Score Distribution
- Monthly Rating Trends

Insights:

- Average customer rating is 3.69, indicating moderately positive satisfaction.
- Reviews include a spread of positive, mixed, and negative sentiments, showing varied customer experiences.
- Sentiment and rating patterns fluctuate monthly, suggesting impact from seasonal campaigns or product releases.

4. Sentiment + Review Behavior Dashboard

Purpose: To explore sentiment trends by product and identify reviewers by behavior clusters (e.g., brand advocates vs. detractors).

Key Measures:

- Sentiment Score (VADER compound)
- Sentiment Category (Hybrid: score + rating)
- Sentiment Buckets

Insights:

- Sentiment is mostly positive, but mixed and negative experiences exist across several products.
- Certain reviewers leave consistently high or low ratings—valuable for identifying customer personas and follow-up strategies.
- Product filtering enabled deep dives into individual product performance from a customer experience lens.

CONCLUSION

This project offered a comprehensive analysis of customer behavior, engagement patterns, and sentiment through data-driven storytelling. By integrating structured customer and campaign data with textual review insights, we were able to paint a full picture of how products perform across different platforms, time periods, and audience segments. The dashboards and analytical models revealed seasonal trends, product popularity cycles, engagement fatigue, and mixed customer perceptions—all of which are crucial for strategic decision-making in marketing and product planning.

RECOMMENDATIONS

1. **Revise Campaign Timing & Content Strategy:** Engagement rates and conversions clearly peaked in early months (especially January), then declined. Future campaigns should focus on timely launches, leveraging seasonal interest, and maintaining freshness in content formats.
2. **Focus on High-Performing Products** Products like Hockey Stick and Running Shoes consistently outperformed others. These can be emphasized in bundled offers or featured campaigns, while underperforming products may benefit from repositioning or promotion.
3. **Improve Funnel Efficiency:** Significant drop-offs between initial views and purchases suggest an opportunity to optimize the customer journey—possibly by simplifying checkout experiences or offering mid-funnel incentives.
4. **Address Customer Pain Points:** Sentiment analysis and review patterns revealed mixed or negative experiences for certain products. These should be reviewed for quality, support issues, or mismatched customer expectations.
5. **Personalize Based on Demographics & Geography:** With access to customer region and demographic attributes, personalized targeting can enhance relevance and response rates—especially for regionally favored products.
6. **Monitor and Adapt Social Media Performance:** As clicks and likes declined over time, reevaluating social media content and channel strategies is essential. Test alternative formats (like short-form video or interactive posts) and consider A/B testing campaign timing.

THANK YOU