Operation Blue Print

Navigating the Waves of E-Commerce Success

Group 7

Members

Ankit Malhotra - 862468488

Nandini Ramakrishnan - 862465172

Pious Khemka - 862467916

Vishesh Shukla - 862464003

Abstract:

In response to the pressing challenge of identifying the optimal investment channel, "Blue" undertakes a comprehensive analysis of its e-commerce performance in the US. The report's overarching goal is to uncover opportunities for optimization and strategic growth, facilitating informed decision-making for channel investments. The methodology employed embraces a multifaceted approach, commencing with Exploratory Data Analysis (EDA) to discern patterns in the data. Subsequently, the analysis extends to profitability assessments at the product, customer, and regional levels, providing granular insights into performance nuances. Store profitability is evaluated to unravel geographical impacts, and a nuanced examination of marketing strategies explores channel profitability concerning marketing costs. The study delves into the influence of seasonal effects on sales, employing time series analysis to capture temporal trends. Additionally, optimization strategies are crafted for each channel, ensuring alignment with profitability objectives, while regression analysis identifies key factors influencing sales dynamics. This holistic methodology aims to furnish "Blue" with actionable insights, navigating the brand toward sustained success in the competitive US e-commerce market.

Keywords:

Keywords: E-commerce Optimization, Channel Investment Strategy, Profitability Analysis, Exploratory Data Analysis

1. Introduction / Background

In the ever-evolving landscape of e-commerce, the brand "Blue," with its roots embedded in India's textile heritage, has been a steadfast emblem of timeless elegance since its establishment in 2000. Beyond the realm of mere clothing retail, "Blue" has strategically positioned itself as a celebration of enduring fashion and a medium for individual expression. This report embarks on a strategic exploration of "Blue's" venture into the competitive realm of e-commerce in the United States.

Company Background:

"Blue's" journey is more than a narrative of retail; it is a narrative of tradition meeting innovation, of a brand that transcends trends to embody a philosophy of sustained style. Situated at the crossroads of tradition and modernity, "Blue" has become synonymous with a celebration of personal style, offering products that go beyond the ephemeral nature of fashion trends.

Mission and Values:

Guided by a mission to provide a platform for customers to express their unique identities through fashion, "Blue" stands apart in valuing the enduring nature of its offerings. The brand's commitment to self-expression and individuality extends beyond the commonplace, steering clear of fleeting trends in favor of a deeper resonance with enduring style.

Project Report Structure:

Structured to elucidate the background of the company, the daily life within its operational framework, and the overarching mission and values, this project report unfolds systematically. As we delve into the methodology section, the report intricately unveils the strategic approach

employed to address the challenge at hand—strategically investing in the most profitable e-commerce channel in the competitive US market. Each section is meticulously designed to offer insights and guide "Blue" towards optimized profitability and sustained strategic growth in the dynamic landscape of US e-commerce.

2. **Project Objective:**

This research endeavor is underpinned by the imperative objective of guiding strategic investment decisions for "Blue" within the context of the competitive US e-commerce milieu. The absence of discernible data-backed insights necessitates an in-depth exploration and analysis of factors influencing channel profitability. Employing a rigorous and comprehensive methodology, incorporating exploratory data analysis, profitability assessments at granular levels (product, customer, and regional), store profitability evaluations, marketing impact analyses, assessments of seasonal effects, time series analyses, and regression analyses, this report aspires to furnish "Blue" with empirically substantiated insights. The overarching ambition is to endow the company with the requisite information to make judicious and informed decisions pertaining to channel investments, thereby optimizing operational efficacy, and ensuring enduring growth within the dynamic contours of the US e-commerce landscape.

3. Data / Problem Analytics

3.1 Methodology

The methodology employed in this report integrates principles and techniques from the MGT 219 class, focusing on comprehensive data collection, analysis, and strategic decision-making.

Data Collection and Preparation:

The comprehensive data collection process for this analysis involved the acquisition and curation of five distinct data files. However, for the purpose of this report, three data files were selected based on their relevance to the analytical methodologies employed.

Channel Selection Rationale (File 1):

This file delves into the rationale behind the selection of specific channels, employing profit and sales comparison techniques. The information gleaned from this file informs the subsequent methodology of Channel-Wise Comparison, strategically aligning channel selection with profit maximization objectives.

Transaction Data (File 2):

This dataset, comprising columns such as Date, Season, Customer, Style, Size, PCS, Rate, and Total Amount, serves as the primary source for time series analysis. With a temporal span of 1.5 years, this file enables a nuanced exploration of trends and patterns over time.

Sales Details and Region Data (File 3):

This dataset, featuring columns like Date, Status, ship-service-level, Category, Size, Size Category, Courier Status, Quality, Rate, Total Amount, Region, and B2B, is instrumental in regression analysis. It provides a comprehensive view of sales details across regions, facilitating a detailed examination of factors influencing sales dynamics. We have 3 months of sales data for this analysis.

Data Preparation Steps:

Channel-Wise Comparison:

Channel-Wise Comparison, informed by insights from Channel Selection Rationale (File 1), strategically aligns channel selection with profit maximization objectives. This process ensures that investments are directed toward channels with the highest potential for sales and profitability.

Exploratory Data Analysis (EDA):

EDA involves the computation of summary statistics on all available data. This process illuminates key characteristics, distributions, and outliers, providing a foundational understanding of the datasets.

Marketing Impact Analysis:

Marketing impact is assessed by analyzing profits from channels in relation to marketing costs. This analysis aims to identify the effectiveness of different marketing strategies and their contribution to overall profitability.

Seasonal Effect Assessment:

Leveraging seasonal information from the data, this analysis explores the impact of seasonal variations on sales and profitability. Insights derived guide strategic planning to navigate seasonal fluctuations.

Time Series Analysis:

Time series analysis is performed based on the transaction data (File 2), unraveling temporal trends and patterns within the sales and customer data over the 1.5-year period.

Multiple Regression Analysis:

Utilizing the sales details and region data (File 3), multiple regression analysis is conducted to identify and quantify the factors influencing sales dynamics. This method provides a nuanced understanding of the relationships between variables, guiding decision-making processes.

This meticulous data collection and preparation process lays the groundwork for a robust analytical framework, ensuring the reliability and relevance of the subsequent methodologies employed in this comprehensive report.

3.2 Data / Problem Analytics

Channel Comparison

Figure 1 Channel-wise sales comparison

The following depicts the side-ways stacked bar chart helping analyze which channel provides the opportunity of maximum sales the brand.

Exploratory Data Analysis (EDA):

On conducting Exploratory Data Analysis (EDA) for all channels sales' data from January, 2022 to July 2023, we study various trends in the data on the basis of seasons, timelines, size orders and customer type. Here the response variable is the total sales across all channels.

Figure 2 Sales Timeline

Figure 3 Quarter Wise Sales Timeline

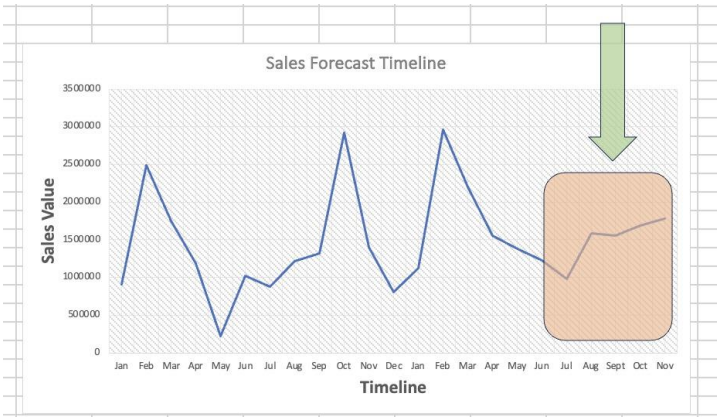

Figure 4 Sales Forecast Timeline

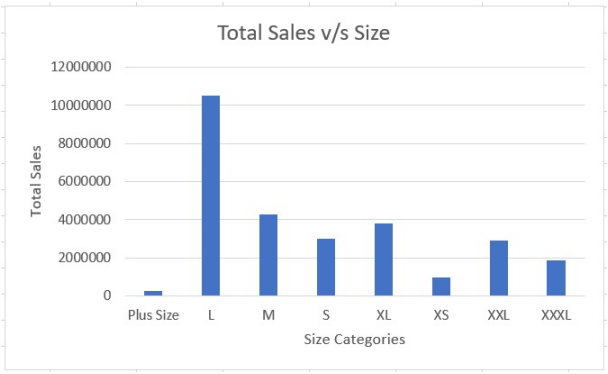

Figure 5 Total Sales v/s Size

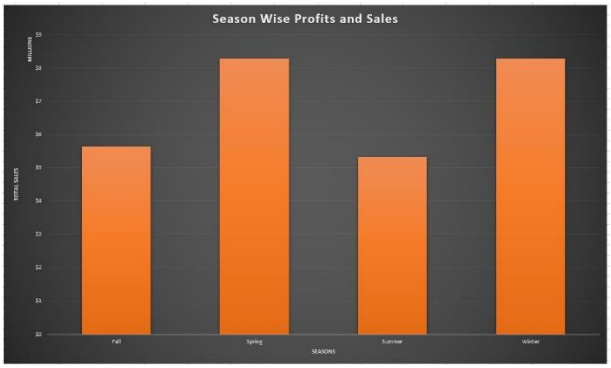

Figure 6 Season Sales

Figure 7 Timeline of Items Sold

Figure 8 Customer Word Cloud

On conducting Exploratory Data Analysis (EDA) specifically for the three months (April, May, June) in the Amazon dataset and with sales data of various other channels allows us to discern how independent variables (predictor variables) on the X-axis impact the dependent variables (response variable – sales amount), providing valuable insights into the underlying relationships.

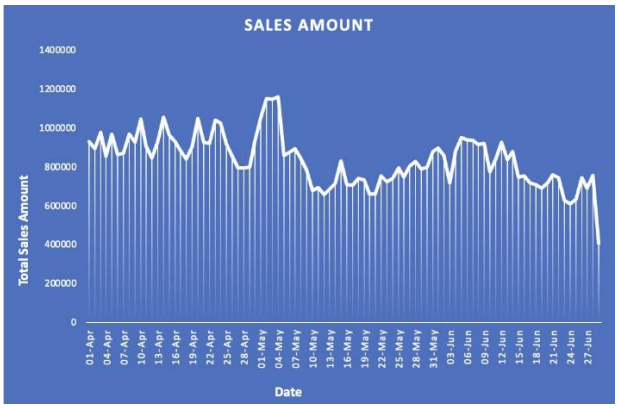

Figure 9 Sales Amount

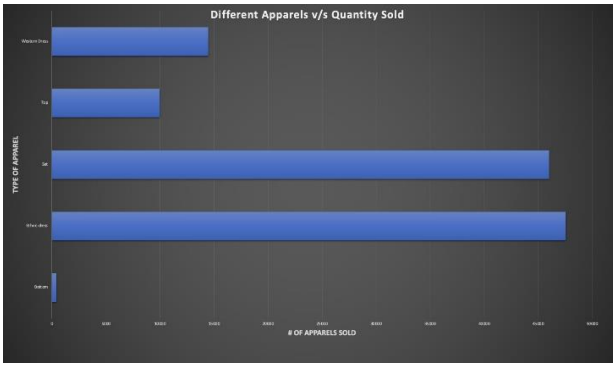

Figure 10Different Apparels v/s Quantity Sold

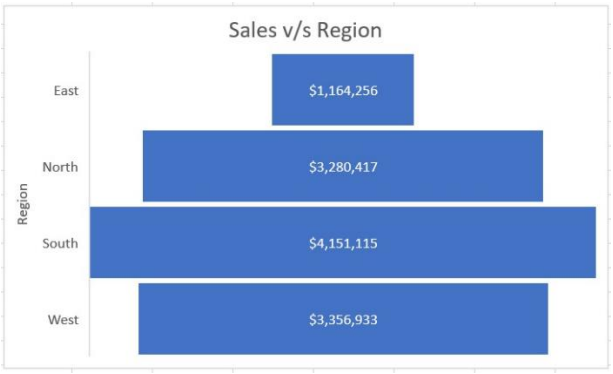

Figure 11 Sales v/s Region

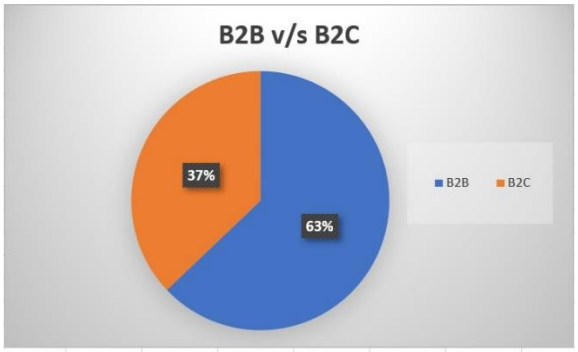

Figure 12 B2B v/s B2C

Regression Statistics	
Multiple R	0.9073079
R Square	0.8232077
Adjusted R Square	0.8231883
Standard Error	2092.3576
Observations	118430

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	13	2.41395E+12	1.85688E+11	42414.378	0
Residual	118416	5.18421E+11	4377960.289		
Total	118429	2.93237E+12			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1309.969	46.46562947	-28.19221966	2.76E-174	-1401.041	-1218.897	-1401.041	-1218.897
Ship_level_stand	23.429728	13.98738116	1.675061805	0.0939247	-3.985316	50.844771	-3.985316	50.844771
Category_Set	-69.71988	17.34570004	-4.019432823	5.837E-05	-103.7172	-35.72258	-103.7172	-35.72258
Category_Ethnic	122.80683	17.45398499	7.036034181	1.989E-12	88.597303	157.01637	88.597303	157.01637
Size_categegory_Small	-19.40967	19.4924869	-0.995751312	0.319373	-57.61463	18.795293	-57.61463	18.795293
Size_category_Large	-7.671195	19.03355294	-0.403035362	0.6869229	-44.97665	29.634265	-44.97665	29.634265
Status_Shipped	-95.3404	31.92956277	-2.985960009	0.0028275	-157.9218	-32.75896	-157.9218	-32.75896
Status_Unshipped	-77.30488	41.44926501	-1.865048229	0.0621771	-158.5448	3.9350187	-158.5448	3.9350187
Qty	709.66908	4.84910346	146.3505761	0	700.16492	719.17325	700.16492	719.17325
Rate	2.0255153	0.029159277	69.46383826	0	1.9683636	2.082667	1.9683636	2.082667
Region_North	7.6156346	22.46189492	0.339046843	0.734575	-36.40932	51.64059	-36.40932	51.64059
Region_South	4.2979286	20.46005046	0.210064416	0.8336178	-35.80344	44.3993	-35.80344	44.3993
Region_West	-11.64929	21.75307618	-0.535523994	0.5922886	-54.28498	30.986387	-54.28498	30.986387
B2B	-464.2787	369.1222884	-1.257791029	0.2084698	-1187.752	259.19508	-1187.752	259.19508

Figure 13 Regression Analysis

4. Findings / Conclusions

Our comprehensive analysis of sales channels provides valuable insights into the performance of each platform, paving the way for strategic recommendations aimed at optimizing our marketing efforts. Following a detailed examination, we observe that Amazon Fashion emerges as the top-performing channel, generating the highest sales and profits, with Poshmark, House of India, and CBazaar following suit. In light of these findings, a decision analysis recommends reallocating marketing efforts, specifically suggesting a reduction in focus on CBazaar and a heightened emphasis on Amazon Fashion.

Exploratory Data Analysis - Sales Across the US:

Sales Trends: Over the last two quarters, a discernible drop in sales is noted, prompting further investigation into the factors contributing to this decline.

Size Dynamics: Across a spectrum of sizes, "L" consistently outperforms others, indicating a size preference that can inform inventory planning and marketing strategies.

Seasonal Patterns: Sales exhibit a seasonal trend, with Winter leading, followed by Spring, Fall, and Summer. This insight is crucial for targeted marketing campaigns aligned with seasonal demands.

Customer Profile: A word cloud analysis of customer names reveals a significant portion of orders originating from Boutiques, providing valuable insights into our customer base.

Exploratory Data Analysis - Sales on Amazon Fashion Across US:

Recent Sales Trends: A recent drop in sales on Amazon Fashion prompts further investigation to identify potential causes and rectify any underlying issues.

Size and Category Influence: Size "L" consistently dominates sales, reinforcing its significance. In terms of categories, Ethnic dresses outshine others, followed by Sets, Western wear, Tops, and Bottoms, guiding product placement and marketing strategies.

Regional Contributions: Sales data across regions indicates a notable contribution from Southern states, followed by the West, North, and East. This regional breakdown is invaluable for targeted marketing and inventory management.

Sales Composition - B2B vs. B2C: B2B sales constitute a substantial 63%, emphasizing the importance of fostering strong relationships with business partners.

Multiple Linear Regression (MLR) Insights:

Y=−1309.97+23.43×Ship_level_stand−69.72×Category_Set+122.81×Category_Ethnic−19.41×Size_category_Small−7.67×Size_category_Large−95.34×Status_Shipped−77.30×Status_Unshipped+709.67×Qty+2.03×Rate+7.62×Region_North+4.30×Region_South−11.65×Region_West−464.28×B2B

The MLR model, with an R-squared of 0.823, indicates a high level of predictability. The adjusted R-squared of 0.823 suggests that 82.3% of the variability in the dependent variable is explained by the independent variables, indicating a well-fitted model. The MLR equation further highlights key contributing factors, including shipping level, category types, size categories, order status, quantity, rate, regional factors, and B2B involvement.

In conclusion, our findings provide a comprehensive understanding of sales trends across channels, guiding strategic decisions to optimize marketing efforts. The recommendations to shift focus from CBazaar to Amazon Fashion align with data-driven insights. The nuanced exploration of size, seasonality, customer profiles, and regional dynamics equips us with actionable intelligence for targeted marketing, inventory management, and overall business optimization. The robustness of our MLR model, as evidenced by the high R-squared and adjusted R-squared values, instills confidence in its predictive capabilities, further supporting strategic decision-making for sustainable business growth.

5. Managerial Implications

1. To optimize channel investments, Brand Blue is recommended to strategically allocate resources and prioritize the Amazon channel, acknowledging its superior profitability

compared to the other three channels. This targeted investment is poised to maximize returns and bolster the overall financial performance of the brand.

2. In a forward-looking approach, Brand Blue should consider the possibility of discontinuing Channel 4 in the future. This recommendation stems from the channel's high operational costs and relatively low sales contribution. By evaluating and potentially phasing out this less lucrative channel, the company can streamline its operations for enhanced efficiency
3. To enhance the partnership with Amazon, Brand Blue should actively pursue discussions to negotiate a reduction in commission rates. Given the substantial volume of sales on this channel, a decreased commission could lead to improved profit margins, fostering a mutually beneficial and cost-effective collaboration.
4. Recognizing the significance of Quarter 1, Brand Blue should strategically align marketing efforts, inventory management, and promotional activities across all channels during this period. As Quarter 1 consistently demonstrates the highest sales, optimizing strategies during this timeframe is crucial for setting a positive trajectory for the entire fiscal year.

6. Idea Sharing

During this project, our team has undergone a substantial learning curve, with one of the pivotal lessons centering around the significance of data cleaning. At the heart of our endeavors, we discovered that the data we initially encountered was far from pristine. Unraveling the intricacies of data cleaning became not just a task but a critical skill that lay at the foundation of our project's success.

A significant revelation that surfaced during our exploration was the inherent messiness of raw data. As we delved into the datasets, we encountered inconsistencies, missing values, and

anomalies that necessitated meticulous cleaning. This hands-on experience underscored the importance of data quality, accuracy, and reliability in any analytical undertaking.

Moreover, our journey led us to a profound appreciation for the complexity of formulas in tools such as Excel. One notable challenge arose when categorizing all U.S. cities into four distinct regions. The conventional approach, involving lengthy and intricate formulas utilizing functions like CHOOSE and IF, highlighted the limitations of manual efforts. Recognizing the need for efficiency and precision, we turned to AI tools, which proved instrumental in automating and streamlining this categorization process.

The decision to leverage AI not only alleviated the burden of manual sorting but also illuminated the power of technology in enhancing the analytical workflow. Embracing AI tools not only expedited our processes but also opened new avenues for exploration, reinforcing the notion that technology can be a valuable ally in the realm of data analysis.

Beyond the technical aspects, this project reinforced the importance of adaptability and the readiness to seek innovative solutions. Our initial struggles with manual data manipulation prompted us to reassess our approach, leading to the incorporation of advanced tools that significantly augmented our capabilities.

In essence, this project served as a multifaceted learning experience. From the intricacies of data cleaning to the realization of the limitations of traditional methods, we emerged not only with newfound technical skills but also with a broader perspective on the dynamic landscape of data analytics. As we move forward, these lessons will undoubtedly shape our future projects, emphasizing the continuous evolution and adaptation required in the ever-changing field of data analysis.

7. References

Data Source: https://www.kaggle.com/code/taufikhidayatzaza/e-commerce-sales-analysis

8. Appendix

Time	Contents
October 4 – October 25, 2023	Project Proposal Design and Presentation Preparation
October 25, 2023	Project Proposal Presentation
November 1, 2023	Data Cleaning
November 8, 2023	Data Pre-processing and Reconciliation
November 15, 2023	Forming Models (Optimization and Regression)
November 22, 2023	Forming Models, Decision Analysis, and Report Generation
November 29, 2023	Final Project Paper Submission and Final Project Presentation