# Final Report: Phase 1 Strategic Analysis

To: Rodrigues

From: Waqas Ahmed

Date: July 1, 2025

**Subject**: Phase 1 Analysis Complete: Definitive Findings & Path to Predictive Modeling

This report details the completion of our initial project phase. After a full day of intensive analysis, we have successfully built a clean analytical foundation and have uncovered several high-impact insights that answer our core business questions.

All code, scripts, and cleaned data files from this analysis are available for review in the project's GitHub repository, a link to which will be provided separately.

### I. Work Accomplished

The foundational work for this project is complete. This involved:

1. **Building a Clean Analytical Layer:** I have programmatically processed, cleaned, and reshaped all 10+ raw data sheets. This involved standardizing formats, correcting structural errors (especially in the complex TOPSHEET file), and creating a set of reliable cleaned\_\*.csv files that now form the basis for all analysis.
2. **Conducting In-Depth "Macro" Analysis:** I performed a series of analyses to build a high-level picture of business performance, including:
   * Return on Ad Spend (ROAS)
   * Overall Customer Acquisition Cost (CAC)
   * Monthly CAC and ROAS Trends
   * New vs. Existing Customer Composition
3. **Performing Rigorous Data Validation:** I manually cross-referenced our key findings with the source files and the Business Overview document. This process included a line-by-line manual recalculation of key data points to ensure our scripts were perfectly accurate. This validation not only confirmed our findings but also uncovered several data integrity issues in the source reporting, adding significant value beyond the initial scope.

### II. Definitive Findings

Our analysis provides clear, data-driven answers to the questions of "Where is growth coming from?" and "Who is driving the business?". I will be attaching the PNG charts and screenshots from our analytical dashboard for each of these findings.

#### Finding 1: Channel Profitability is a Tale of Two Extremes

Our final, validated analysis reveals a stark contrast in the efficiency of paid marketing channels.

|  |  |  |
| --- | --- | --- |
| **Channel** | **Corrected ROAS** | **Verdict** |
| Organic Search | **$46.60** | 🏆 Elite Performer |
| Email | **$38.74** | 🏆 Elite Performer |
| SMS | **$25.38** | ⭐ Strong Performer |
| Paid Search | **$5.18** | ✅ Profitable |
| Shopping | **$3.05** | ✅ Profitable |
| **Affiliates** | **$1.78** | ⚠️ **Problem Area** |
| **Paid Social** | **$0.73** | 🚨 **Critical Issue** |

* **Inefficient Channels:**
  + **Paid Social:** This is the most critical issue. The channel is chronically unprofitable with a **ROAS of $0.73** and a very high overall **CAC of $129.68**.
  + **Affiliates:** This channel is also unprofitable, with a lifetime **ROAS of $1.78** and an expensive **CAC of $111.77**.
* **Efficient Channels:**
  + **Paid Search:** This is a reliable and profitable channel with a healthy **ROAS of $5.18** and a much more efficient **CAC of $37.54**.
  + **Organic & Direct Channels:** Channels like Organic Search, Email, and SMS are hyper-efficient and represent the primary profit drivers of the business.

**Recommendation:** Immediately re-evaluate the strategy and budget for Paid Social and Affiliates. The data shows these channels, in their current state, are a significant drain on profitability.

#### Finding 2: The Business Operates on a "Leaky Bucket" Model

The analysis of our customer base shows two key trends:

1. **Stagnant Growth:** The total number of active customers is not growing consistently month-over-month.
2. **High Reliance on New Customers:** A very large portion of active customers each month are new.

This indicates that the business is constantly spending to acquire new customers who are not being effectively retained, leading to a "leaky bucket" where growth is difficult to sustain.

**Recommendation:** This finding underscores the critical need to shift focus from pure acquisition to customer retention. The next phase of our project, focusing on LTV and churn, will directly address this.

#### Note on Data Integrity

During our validation process, we discovered that the "New Customer" data for 2025 was incomplete in one of the source files, containing only four months of data (Jan-Apr) instead of the expected five. To ensure the accuracy of our final analysis, we made the strategic decision to cap all analyses at April 30, 2025, providing a fair, apples-to-apples comparison across all metrics. This rigorous approach ensures the findings presented here are as accurate and defensible as possible.

### III. Next Steps & Required Data

We have now extracted the maximum possible value from the existing summary-level data. The path forward requires moving from high-level reporting to customer-level predictive modeling.

To build the **Customer Lifetime Value (LTV)** and **Churn Prediction** models outlined in the project brief, I am now awaiting the final, more granular data files.

Action Required:

To proceed, I require the following datasets:

1. **A complete, row-level Orders table** (containing customer\_id, order\_id, product\_id, and order\_date).
2. **A complete Customers table** (containing customer\_id and first\_order\_date).

Upon receipt of this data, I will begin Phase 2 of our project. Thank you for your partnership. I am confident that the insights from the next phase will be even more impactful.

Google Drive Link: <https://drive.google.com/drive/folders/1Cm_-UInyt_WP5lfaYwjLsKvfwjjtrU-S?usp=sharing>