Customer Cohort Analysis & CLV Insights Report

**1. Introduction**

Customer behavior analysis is critical for understanding purchasing patterns and business growth. This report presents insights from **Cohort Analysis** and **Customer Lifetime Value (CLV)** calculations based on an e-commerce dataset.  
The goal of this study is to:

* Segment customers into cohorts based on their **first purchase month**.
* Analyse **retention trends** over time.
* Identify **high-value customers** using CLV.
* Provide **actionable recommendations** for improving customer engagement and revenue.

**2. Dataset Overview**

Before performing cohort analysis, we examined the dataset to understand its structure:

* The dataset contained **541909 rows** and **8 columns**, with details on invoices, customers, products, and purchases.
* Data cleaning steps:
  + **Removed duplicate records** to avoid overcounting.
  + **Handled missing values** in key columns.
  + **Filtered out cancelled orders and negative quantities** to ensure accuracy.

**3. Cohort Analysis Details**

We created cohorts based on **the customer’s first purchase month**, assigning each customer to a specific group.

**3.1 Number of Cohorts Created**

* We segmented the dataset into **13 cohorts**, where each cohort represents customers acquired in a specific month.

**3.2 Cohorts Created**

Customers were grouped into the following cohorts:

* **[Cohort 1: 12/2010]**
* **[Cohort 2: 01/2011]**
* **[Cohort 3: 02/2011]**
* **[Cohort 4: 03/2011]**
* **[Cohort 5: 04/2011]**
* **[Cohort 6: 05/2011]**
* **[Cohort 7: 06/2011]**
* **[Cohort 8: 07/2011]**
* **[Cohort 9: 08/2011]**
* **[Cohort 10: 09/2011]**
* **[Cohort 11: 10/2011]**
* **[Cohort 12: 11/2011]**
* **[Cohort 13: 12/2011]**

**3.3 Customer Retention Trends**

* The **highest retention** was observed in the **first month** of each cohort.
* A **steep decline** in retention occurred after the **second month**, suggesting the need for stronger engagement.
* Very few customers remained active after **six months**, indicating room for improvement in long-term retention.
* The **best cohort** in terms of retention was **[12/2010]**, with a retention rate of **38.2%** in the second month and **33.4%** in the third month.
* The **worst cohort** was **[11/2011]**, where retention dropped sharply to **13.4%** after the first month.

**4. Customer Lifetime Value (CLV) Analysis**

* CLV was calculated for each customer, identifying **high-value customers**.
* The **top 3 customers** with the highest CLV are:
  1. **Customer ID 14646:** $23290.76
  2. **Customer ID 18102:** $21369.88
  3. **Customer ID 14096:** $19040.31
* These customers contributed significantly to total revenue.

**5. Key Insights & Recommendations**

**5.1 Insights**

1. **Retention Drops After Month 1**: Most customers do not return after their first purchase.
2. **High-Value Customers Exist**: A small percentage of customers contribute significantly to revenue.
3. **Best and Worst Cohorts Identified**: Some cohorts exhibit stronger purchasing patterns than others.

**5.2 Recommendations**

1. **Improve Retention Strategies**
   * Implement **loyalty programs** to encourage repeat purchases.
   * Use **personalized email campaigns** targeting inactive customers.
2. **Focus on High CLV Customers**
   * Offer **exclusive deals and premium services** to top customers.
   * Provide **personalized discounts** based on purchase history.
3. **Enhance First-Month Engagement**
   * Send **follow-up reminders** and **discount offers** after the first purchase.
   * Improve customer onboarding to **increase repeat purchases**.

**6. Conclusion**

This analysis highlights the importance of **customer retention and high-value customer engagement**.  
By focusing on top cohorts and improving retention strategies, businesses can maximize revenue and build long-term customer loyalty.