**High-Level Document (HLD)**

**Amazon Sales Data**

**Analysis**

Revision Number: 1.0

Deepraj Arya

**Project Overview**

The Amazon Sales Data Analysis project aims to provide valuable insights and decision support to Amazon's business operations by analysing historical sales data. This High-Level Document (HLD) outlines the project's objectives, methodologies, and key findings.

**Project Objective**

The primary objective of this project is to analyse Amazon's historical sales data to gain insights into sales trends, profitability, and key performance indicators. The project will answer specific questions and provide actionable insights, including:

1. **Sales Trends Analysis:** Understand the trends and patterns in Amazon's sales data, both at a global and regional level.
2. **Profitability Analysis:** Analyse the profitability of Amazon's products and sales channels to identify areas for optimization.
3. **Outlier Detection and Handling:** Identify and understand outliers in the data, which could be extreme successes or failures, and develop strategies to leverage or mitigate them.
4. **Feature Engineering:** Create new features, such as Profit Margin, Sales Growth Rate, and Cost-to-Revenue Ratio, to gain deeper insights into business performance.
5. **Month-Wise and Year-Wise Analysis:** Analyse sales data on a monthly and yearly basis to identify seasonality, growth patterns, and yearly trends.
6. **Visualizations:** Create data visualizations to present key findings effectively to stakeholders.

**Methodology**

The Amazon Sales Data Analysis project follows a structured methodology, consisting of the following steps:

1. **Data Preparation:** The project begins by loading and preparing the raw sales data. This includes data cleaning, data type conversion, and handling missing values.
2. **Exploratory Data Analysis (EDA):** The project conducts exploratory data analysis to gain initial insights into the data. This includes summary statistics, data distribution, and correlation analysis.
3. **Outlier Detection and Handling:** Outliers are detected and handled using the Interquartile Range (IQR) method. Extreme outliers are identified, and potential reasons are analysed.
4. **Feature Engineering:** New features, such as Profit Margin, Sales Growth Rate, and Cost-to-Revenue Ratio, are created to provide deeper insights into business performance.
5. **Month-Wise and Year-Wise Analysis:** The project conducts a detailed analysis of sales data on a monthly and yearly basis to identify seasonality, growth patterns, and yearly trends.
6. **Visualizations:** Data visualizations, including line plots, box plots, and pairplots, are created to effectively present key findings to stakeholders.

**Key Findings**

The Amazon Sales Data Analysis project has generated the following key findings:

1. **Sales Trends Analysis**:

There is no strong trend or pattern in year-wise sales data.

Monthly sales data analysis reveals sporadic fluctuations with occasional growth in specific months.

1. **Profitability Analysis:**

Profit margins vary significantly by product category, with some categories demonstrating higher profitability.

Extreme outliers in profitability can be attributed to promotional campaigns, large one-time orders, and seasonal trends.

1. **Outlier Detection and Handling:**

Outliers are detected and treated, allowing for a more accurate representation of typical business performance.

1. **Feature Engineering:**

Key metrics, including Profit Margin, Sales Growth Rate, and Cost-to-Revenue Ratio, have been created to gain deeper insights.

1. **Month-Wise and Year-Wise Analysis:**

Sales growth rate exhibits sporadic variations with certain months demonstrating significant growth.

Year-wise analysis shows varying degrees of growth, with some years experiencing substantial increases.

1. **Visualizations:**

Data visualizations have been employed to present findings effectively, aiding in decision-making processes.

**Next Steps**

The insights and findings from this analysis will guide Amazon's business strategies and decision-making processes. The next steps for this project include:

* **Strategy Replication:** Replicate successful strategies that led to high profitability in the future.
* **Detailed Review of Unprofitable Transactions:** Conduct a detailed review of unprofitable transactions to understand their causes and take corrective actions.
* **Monitoring Mechanisms**: Implement monitoring mechanisms to detect similar outliers in the future, allowing for timely action.
* **Optimization Strategies:** Optimize pricing strategies, marketing campaigns, and overall business performance based on the insights provided.
* **Regular Updates:** Continuously update and re-run the analysis to adapt to changing market conditions and business strategies.