

# Project Synopsis

## Project Title:

Analysis of Customer Reviews and Orders to Enhance Product Strategies Using Big Data Technologies

## Group Number:

Group 1

## Project Overview:

The primary objective of this project is to analyze the Amazon Customer Reviews Dataset to extract actionable insights, identify trends, and create compelling visualizations to better understand customer behavior. These insights aim to enhance product recommendations and business strategies.

The dataset comprises millions of customer reviews from diverse product categories available on Amazon's website. Each review includes contextual details, ratings, and other relevant information. The analysis emphasizes identifying popular products and categories based on the volume of reviews and average ratings. These findings are presented through interactive dashboards and reports created in Power BI to ensure effective communication with stakeholders.

## Technologies and Tools Used:

- Data Storage and Management: Amazon S3
- Data Processing and Analysis: Apache Spark, PySpark, Python (Pandas, NumPy), SQL, AWS Glue, AWS Crawler, Amazon Athena
- Visualization: Power BI
- Project Management and Documentation: Jupyter Notebooks, GitHub

## **Project Synopsis**

### **Conclusion:**

This project demonstrates the effective use of big data technologies and machine learning techniques to derive valuable insights from the Amazon Customer Reviews Dataset. By analyzing customer reviews and behavior, businesses can make data-driven decisions to refine product offerings, improve customer satisfaction, and drive overall business growth.