Customer Review Sentiment & Insights Analysis – Full Project Report

1. Introduction

This project analyzes customer reviews from an e-commerce dataset using AWS cloud services. It automates sentiment analysis, key phrase extraction, text summarization, data processing, and dashboard visualization using Amazon S3, Lambda, Comprehend, DynamoDB, SageMaker, and QuickSight.

2. Objectives

- Automate ingestion of customer reviews
- Analyze sentiment using Amazon Comprehend
- Extract key phrases for insights
- Summarize all reviews using a SageMaker NLP model
- Store processed results in DynamoDB
- Build an interactive business intelligence dashboard in QuickSight

3. Architecture Diagram (Description)

The pipeline follows this structure:
1. Amazon S3 stores raw CSV review files.
2. S3 triggers a Lambda function for ETL and analysis.
3. Amazon Comprehend performs sentiment and key phrase extraction.
4. DynamoDB stores processed results.
5. SageMaker JumpStart summarization model produces review summaries.
6. Another Lambda stores generated summaries in DynamoDB.
7. QuickSight visualizes insights with KPIs and charts.

4. Dataset Description

Dataset: Women's Clothing E-Commerce Reviews.
Contains product reviews, ratings, department information, and free-text feedback.
Over 23,000 rows with fields such as Review Text, Rating, Division, Category, and Positive Feedback Count.

5. AWS Services Used

1. Amazon S3 – Dataset storage
2. AWS Lambda – ETL, batch processing, Comprehend calls
3. Amazon Comprehend – Sentiment analysis & key phrase extraction
4. DynamoDB – NoSQL storage for processed results
5. Amazon SageMaker – Text summarization using NLP model
6. Amazon QuickSight – Dashboard visualizations

6. Implementation Steps

Step-by-step implementation:
1. Create S3 bucket and upload dataset.
2. Create two Lambda functions: ingestion & summarization.
3. Configure IAM roles with correct permissions.
4. Use Comprehend inside Lambda to extract sentiment & key phrases.
5. Store structured output in DynamoDB.
6. Deploy SageMaker JumpStart model for summarization.
7. Generate summary and save to DynamoDB.
8. Export DynamoDB table to S3.
9. Connect QuickSight to S3 and build dashboard.

7. Visualizations in QuickSight

Dashboard includes:
- KPI: Total Reviews
- KPI: Average Rating
- KPI: % Positive Sentiment
- Donut Chart: Sentiment Distribution
- Bar Chart: Rating Distribution
- Heatmap: Product Category vs Sentiment
- Word Cloud: Key Phrases
- Table: ReviewText + Sentiment + Rating

8. Key Insights

Analysis reveals:
- Majority of customers express positive sentiment.
- Frequent issues include size, fitting, and delivery.
- Certain product categories have higher negative sentiment.
- Summarized text highlights themes like comfort, material quality, and value for money.

9. Conclusion

This end-to-end AWS pipeline organizes raw customer reviews into actionable insights. Automation reduces manual workload and helps businesses understand customer sentiment, improve quality, and make better decisions.