Title: Review Scrapping and Analysis using Flask, MongoDB, and AWS

1) Introduction:

The project aims to automate the process of scraping and analyzing product reviews from Flipkart, utilizing various technologies such as Flask, MongoDB, and AWS. By extracting valuable insights from customer feedback, businesses can make data-driven decisions to improve their products and enhance customer satisfaction.

2) Problems:

a) Manual review analysis: Analyzing a large number of reviews manually is a time-consuming and arduous task. The project addresses this challenge by automating the review collection and analysis process.

b) Data storage and management: Storing and managing a significant amount of review data requires an efficient and scalable database solution. MongoDB is employed to ensure seamless storage and retrieval of scraped reviews.

c) Scalability and accessibility: The project leverages the AWS cloud infrastructure to deploy the system, ensuring it can handle growing data volumes and remains easily accessible from any location.

3) Objective:

The primary objective is to streamline the review scraping process by developing a web scraping script that extracts product reviews from Flipkart. These reviews are then stored in a MongoDB database for efficient management. Additionally, a Flask application is created to create an API that enables easy retrieval and analysis of the scraped data.

4) Impact and Value:

a) Customer insights: By analyzing a large dataset of product reviews, businesses gain valuable insights into customer preferences, opinions, and pain points. This information can drive improvements in product design, marketing strategies, and customer support.

b) Product improvements: Identifying common themes and issues raised in reviews helps businesses understand areas where their products may need improvement. This allows them to make data-driven decisions to enhance product quality and address customer concerns effectively.

c) Decision-making: The availability of a comprehensive review dataset empowers businesses to make informed decisions, such as identifying top-performing products, assessing market trends, and monitoring customer satisfaction over time.

5) Workflow:

a) Web Scraping: Develop a robust web scraping script that retrieves product reviews from Flipkart. Extract relevant information such as review content, ratings, and timestamps.

b) Flask Application: Build a Flask application that acts as an interface between the database and users. This application creates an API that allows easy access to the stored review data.

c) MongoDB Integration: Establish a connection with a MongoDB database and store the scraped review data in a structured manner. Implement indexing and querying techniques for efficient data retrieval.

d) AWS Deployment: Deploy the project on AWS cloud infrastructure for improved scalability, reliability, and accessibility. Utilize AWS services such as EC2, S3, and Route 53 to ensure seamless operation.

6) Conclusion:

The project successfully automates the process of scraping product reviews from Flipkart, storing them in a MongoDB database, and providing access to the data through a Flask-based API deployed on AWS. This end-to-end solution enables businesses to gain valuable customer insights, improve their products, and make informed decisions. By leveraging the power of technologies like Flask, MongoDB, and AWS, the project offers a scalable and efficient solution for review scraping and analysis.