**Batch Time Analysis of Transactional Data**

**Introduction:**

* Lenodo, a multinational e-commerce giant, aims to maximize its transactional data potential through AWS Big Data tools.
* As a data analytics consultant, the focus is on revealing insights into best-selling items and their popular countries, emphasizing the significance of data-driven decisions in business strategy.

**Objective**:

* The primary objective of this project is to utilize AWS Big Data tools to perform data engineering, uncover transactional patterns, and provide actionable insights for Lenodo.
* By transforming raw transactional data into a more efficient and query-friendly format, the company aims to enhance its understanding of product performance, brands, categories, and segments.
* These insights will empower different teams within Lenodo to make informed decisions, leading to more effective digital marketing campaigns and overall business optimization.

**Project Execution:**

**1. S3 Bucket Creation and CSV Upload**:

* The project kicks off with the creation of a unique S3 bucket to store the CSV file containing nightly exported transactional data.
* It is imperative that the CSV file is in UTF-8 format for compatibility and seamless processing within the AWS environment.
* This step ensures a centralized and easily accessible data repository.

**2. Crawler Creation for CSV Data:**

* A crucial step in data engineering is the creation of a crawler to traverse the CSV data stored in the S3 bucket.
* The crawler generates a metadata catalogue, offering a structured view of the available data. This metadata catalogue becomes the foundation for subsequent processing and analysis.

**3. Glue Job for Data Transformation:**

* To enhance the efficiency of data warehouse queries, a Glue job is implemented to transform the raw CSV data into the Parquet format.
* Parquet's columnar storage and compression capabilities make it an optimal choice for analytical queries. The transformation process is designed to streamline data access and retrieval.

**4. Crawler for Parquet Data:**

* Following the transformation, another crawler is deployed to navigate the Parquet data files.
* This crawler generates a metadata catalogue specific to the Parquet format, enabling seamless integration with Athena for query execution.

**5**. **Athena Query for Insights:**

* The final step involves querying the data using Athena to gain insights into the best-selling items and the countries where these items are most popular.
* Leveraging the power of SQL queries, the analysis focuses on extracting valuable information that will guide marketing, product development, sales, and procurement strategies.

**OUTPUT:**

BestSellingProduct – DOTCOM POSTAGE

CountryWithMostSales – United Kingdom

**Conclusion:**

In conclusion, the Batch Time Analysis of Transactional Data project exemplifies the strategic use of AWS Big Data tools to unlock the hidden potential within Lenodo's transactional data. By establishing an efficient data processing pipeline, the project empowers Lenodo's teams with actionable insights, fostering data-driven decision-making across various business functions. This initiative not only enhances the overall understanding of product performance but also positions Lenodo for more targeted and effective digital marketing campaigns in the dynamic e-commerce landscape. Through the integration of AWS Big Data services, Lenodo is well-positioned to navigate the complexities of modern data analytics, driving innovation and growth in the highly competitive e-commerce industry.