|  |  |
| --- | --- |
| **Name: Jal Bafana** | **Roll no: K005** |
| **Class: Btech. Cyber Security** | **Batch: K1** |
| **Date of Experiment: 14.01.2025** | **Date of Submission: 14.01.2025** |

**Aim: To Analyse the following in data warehousing and mining**

**(A)**

1. **Application implementing DW & DM.**
2. **Domain of the application.**
3. **Key concepts implemented.**
4. **Benefit to the company. (Outcome after implementation)**

**(B) Software implementing Data Warehouse**

**Analyse the following**

1. **Software name**
2. **Name of company**
3. **Key features of Software**
4. **Application Overview**

* **Walmart's Retail Analytics Platform**:
  + Walmart, as a global retail giant, employs Data Warehousing (DW) and Data Mining (DM) technologies extensively to analyze vast amounts of transaction and customer data.
  + Walmart's data warehouse aggregates transaction data from over 11,000 stores worldwide, allowing the company to derive insights that optimize everything from inventory management to pricing strategies.
  + **Data Mining** helps Walmart analyze this data, identify shopping trends, predict customer needs, and fine-tune its operations.

1. **Domain of the Application:**

* Walmart operates in the retail domain, with a focus on grocery and general merchandise sales. The company has implemented data warehousing (DW) and data mining (DM) techniques to improve operational efficiency, inventory management, and customer service across thousands of stores and e-commerce platforms worldwide.

1. **Key Concepts Implemented**
   * **Data Warehouse**: Walmart operates one of the largest data warehouses globally, collecting and processing enormous volumes of transactional data from thousands of stores. This centralized data storage system supports complex analysis and provides valuable insights that influence business operations.
   * **Big Data Analytics**: Walmart processes massive amounts of unstructured data (about 2.5 petabytes every hour), including customer interactions, sales transactions, and external market data. This helps the company optimize inventory levels, personalize promotions, and enhance its forecasting models.
   * **Predictive Analytics**: Walmart employs predictive analytics to forecast demand, optimize pricing, and improve supply chain management. By analyzing patterns in customer behavior and sales data, Walmart can predict future trends and tailor its inventory and marketing strategies accordingly.
2. **Benefits to the Company**
   * **Operational Efficiency**: Walmart's use of DW and DM enables real-time data insights that help streamline supply chain operations, optimize inventory, and reduce excess stock or stockouts. This improves logistics management and cost-efficiency across the company.
   * **Enhanced Customer Experience**: By analyzing customer behavior and shopping patterns, Walmart tailor’s product recommendations, promotional offers, and store layouts to better meet customer needs. This personalized experience increases customer satisfaction and loyalty.
   * **Data-Driven Decision Making**: The ability to process and analyze large volumes of data helps Walmart make data-driven decisions. This includes optimizing product assortments, setting competitive prices, and improving store operations.

**Software Name: Teradata**

**Company Name**: Teradata Corporation

**Key Features**:

* **Scalable Data Warehouse**: Teradata provides a high-performance, scalable data warehouse solution capable of handling massive amounts of structured and unstructured data. Walmart uses Teradata to store and manage petabytes of data from its global retail operations, which include point-of-sale transactions, inventory levels, and customer interactions.
* **Advanced Analytics**: Teradata integrates with advanced analytics tools, allowing Walmart to run complex queries and analyses on large datasets. This enables Walmart to gain deeper insights into customer behavior, inventory trends, and operational efficiencies.
* **Optimized for Large Datasets**: Teradata is specifically designed to handle large-scale data environments. Walmart benefits from its ability to process and store large datasets efficiently while maintaining high-speed performance, even when performing complex joins and aggregations.
* **Data Integration**: Teradata supports the integration of data from multiple sources, including transactional systems, CRM platforms, and third-party applications. This helps Walmart create a unified view of its operations and better understand customer preferences and behavior.

**Software Name: Amazon Redshift**

**Company Name**: Amazon Web Services (AWS)

**Key Features**:

* **Fully Managed Cloud Data Warehouse**: Amazon Redshift is a cloud-based data warehouse that removes the need for complex hardware infrastructure management. Walmart leverages Redshift to store and process vast amounts of data on AWS's secure and scalable cloud platform.
* **Scalable Architecture**: Redshift allows Walmart to scale its data warehousing needs easily, depending on the size of the data and the complexity of the queries. This flexibility ensures that Walmart can handle a growing volume of data from multiple sources without worrying about the limitations of on-premises infrastructure.
* **High-Performance Querying**: Redshift uses columnar storage and parallel query execution, which ensures that Walmart can run fast queries on large datasets. This enables Walmart to quickly analyze historical sales data, inventory levels, and customer behavior across different regions.
* **Seamless Integration with AWS Ecosystem**: Amazon Redshift integrates seamlessly with other AWS services, such as **AWS S3** (for data storage), **AWS Glue** (for data preparation), and **Amazon QuickSight** (for business intelligence). This makes it easy for Walmart to create a comprehensive and efficient data pipeline for analytics.
* **Cost-Effective**: Amazon Redshift offers a pay-as-you-go pricing model, which is beneficial for Walmart as it only pays for the storage and computing resources it uses, allowing for efficient cost management at scale.

**Observation:** By using **Teradata** and **Amazon Redshift**, Walmart has successfully implemented a robust data warehousing strategy. Teradata provides a powerful on-premise solution for handling large datasets and running complex queries, while Amazon Redshift offers flexibility and scalability in the cloud. Together, these solutions enable Walmart to process and analyze vast amounts of data efficiently, supporting the company's operational needs and customer-focused strategies.

**Conclusion:** Walmart's integration of **Teradata** and **Amazon Redshift** into its data warehousing strategy enables the company to make faster, more informed decisions across its global operations. These software solutions provide Walmart with the scalability, performance, and flexibility needed to handle its growing data demands while improving operational efficiency and customer satisfaction.

**Citations:**

1. <https://www.igi-global.com/chapter/business-data-warehouse/7798>
2. <https://www.projectpro.io/article/how-big-data-analysis-helped-increase-walmarts-sales-turnover/109>
3. <https://www.onehouse.ai/blog/enabling-walmarts-data-lakehouse-with-apache-hudi>
4. <https://dataforest.ai/blog/practical-data-warehousing-successful-cases>
5. <https://www.sas.com/en_in/insights/articles/analytics/how-walmart-makes-data-work-for-its-customers.html>
6. <https://www.researchgate.net/publication/297425872_Business_Data_Warehouse_The_Case_of_Wal-Mart>
7. <https://www.globalspec.com/reference/23313/203279/chapter-9-postimplementation-of-the-wal-mart-data-warehouse>