STREAMLINING INTERNATIONAL SHIPPING LOGISTICS

[1]R. Abu Azam, [2] Dr. S. Mangayarkarasi
[1]M.Sc. Student, [2]Professor,
Department of Advanced Computing & Analytics,
Vels Institute of Science, Technology &Advanced Studies. Chennai.
India.

Abstract- International shipping logistics involve handling large quantities of shipment data, shipment schedules, and transportation processes across various regions. The current project aims to develop a data-centric solution to optimize logistics processes using the DataCo Supply Chain dataset and database tools. The developed system is based on Python and SQLite database tools to effectively manage and analyze shipment data. The key features and functionalities of the developed system include goods tracking, shipment time analysis, cost optimization, and detection of delayed shipment alert. The system analyzes actual and scheduled shipment times to minimize shipment delays and optimize shipment efficiency.

Keywords - 3PL (Third-party Logistics) digitalization and end-to-end visibility.

I.INTRODUCTION

International shipping logistics plays a key role in the trade process by facilitating the transportation of goods from different countries and markets. As the e-commerce business and global trade are growing exponentially, it has become essential for the logistics company to manage the logistics process effectively and efficiently by tracking the shipment process, managing the delivery schedule, and controlling the different modes of transportation and operational costs. However, the traditional logistics management process faces many challenges such as delivery delay, route management, and increased logistic costs. International shipping logistics plays a crucial role in global trade and supply chain management [2].

To overcome such challenges in the logistics management process, data-driven approaches and database solutions can be implemented to analyze the shipment data and improve the decision-making process effectively. The objective of this project is to design a logistics analysis system for the “Streamlining International Shipping Logistics with Database Solutions” process by using the DataCo Supply Chain dataset. The dataset provides information about the shipment orders, customer locations, modes of transportation, and profitability of the orders. The dataset is processed and analyzed using Python and different data analysis libraries to obtain meaningful data from the dataset. Furthermore, the system incorporates a database management system based on SQLite, which enables efficient management of logistics information.

Through data analysis and visualization, the system helps in tracking goods from different regions, evaluating efficiency in terms of time spent during shipping, recognizing risks associated with deliveries, and evaluating market performance. Moreover, a notification alert system has been integrated to identify shipments that have a high risk of delayed delivery. Data analytics techniques are widely used to improve logistics efficiency and making.

II. LITERATURE SURVEY

International shipping logistics has attracted significant attention in recent times due to its importance in global trade and managing the supply chain. In earlier systems, there were no significant efficiencies, and operations were carried out manually. The introduction of technology in managing data through database management systems has been effective in managing large volumes of data related to international shipping logistics. Several studied have explored the role of data analytics in logistics optimization [1], [4]. Recent studies have also focused on the role of data analytics in improving international shipping logistics. The data analytics approach enables organizations to analyze data related to the transportation of goods and improve the efficiency of operations. Supply chain management strategies have been discussing extensively in previous work [2], [3]. It also enables organizations to track the movement of goods through the introduction of systems for tracking shipments. The recent studies also focused on improving the efficiency of international shipping operations by comparing actual and scheduled operations. Optimization techniques are used to improve transportation efficiency [5], [6].

The system also enables organizations to reduce the risks associated with late deliveries by sending early notifications to the logistics manager. The introduction of database systems and data analytics techniques in managing international shipping operations has been effective in improving efficiency and reducing delays.

III. PROJECT DESCRIPTION

The project **"Streamlining International Shipping Logistics"** is developed to improve the efficiency and management of international shipping operations using data analytics and database solutions. In modern logistics systems, handling shipment information, tracking goods, monitoring delivery schedules, and managing transportation costs are challenging tasks due to the large volume of logistics data generated daily. Traditional logistics systems often face problems such as delayed deliveries, inefficient shipment tracking, poor visibility of goods movement, and increased operational costs. To overcome these challenges, the proposed system uses a data-driven approach for analysing and managing logistics operations effectively. The system is developed using Python programming and SQLite database technology to process and store shipment-related information efficiently. The logistics dataset contains information related to customer locations, shipment status, delivery schedules, shipping modes, and profitability details.

The system performs various operations such as shipment tracking, shipping time analysis, cost optimization, and delayed shipment detection. Goods tracking functionality helps monitor the movement of products across different countries and regions. Shipping time analysis compares actual delivery time with scheduled delivery time to identify shipment delays and improve delivery performance. Cost analysis is used to study profit trends and identify high-performing markets for better business decisions.

The project also includes an alert and notification mechanism that identifies shipments with high risk of delayed delivery and provides early alerts to logistics managers. Data visualization techniques such as charts and graphs are used to represent logistics insights clearly and effectively. The integration of database management and data analytics helps improve supply chain visibility, reduce delivery delays, optimize logistics costs, and support efficient decision-making processes.

Overall, the proposed system provides an intelligent logistics management solution that enhances operational efficiency, improves shipment monitoring, and supports better international shipping management.

IV. TECHNICAL SPECIFICATIONS

The proposed system, "Streamlining International Shipping Logistics," is developed using Python programming language and SQLite database technology to efficiently manage and analyze logistics operations. Python is used because of its simplicity, flexibility, and powerful support for data analytics and database integration. Various Python libraries such as Pandas and NumPy are used for handling, preprocessing, and analyzing shipment data, while Matplotlib and Seaborn are utilized for graphical representation and visualization of logistics insights. The SQLite database is used for storing shipment information, customer details, delivery schedules, and profit-related data in a structured format.

The system performs operations such as goods tracking, shipping time analysis, cost optimization, delayed shipment detection, and alert notification generation. The project also supports data visualization techniques to represent shipment trends, delivery performance, and logistics analysis clearly through charts and graphs. The implementation is carried out using tools such as Jupiter Notebook, VS Code, or PyCharm in a Windows environment. Overall, the integration of data analytics techniques and database solutions helps improve logistics efficiency, shipment monitoring, delivery performance, and decision-making processes effectively.

V. PROPOSED METHODOLOGY

The proposed system uses a data-driven approach to improve international shipping logistics. This system uses database solutions and data analysis techniques. Firstly, the dataset from the DataCo Supply Chain is collected. This dataset contains information about shipment details, delivery status, shipping modes, locations of customers, and profitability of orders. Then, data preprocessing occurs. This includes cleaning the column names, handling missing data, and transforming the data. A new feature, 'shipping delays,' is introduced.

This feature calculates the difference between actual shipping time and scheduled shipping time. This helps in finding delays in shipping. Then, the data is stored in an SQLite database. This database helps in efficient data management. SQL queries are used to analyze key logistics metrics such as market-wise delivery risk and shipping modes. Moreover, data analysis and visualization are performed. This helps in understanding shipment patterns, efficiency in shipping, and costs. A goods tracking module is implemented. This helps in tracking the movement of goods from one region to another.

A shipping time analysis module is used to compare actual and scheduled time to reduce delays, and a cost analysis module is used to assess trends in profit. The proposed system applies data analytics techniques to analyze logistics data improve delivery performance [1], [4]. Cost optimization is achieved using data-driven approaches [5].

A notification alert system is also designed to identify shipments that are likely to be delivered late, thus helping the logistics manager take necessary actions to prevent delays. The entire methodology is based on database management and data analytics to track goods, reduce time, and lower costs, thus enhancing efficiency in logistics management.

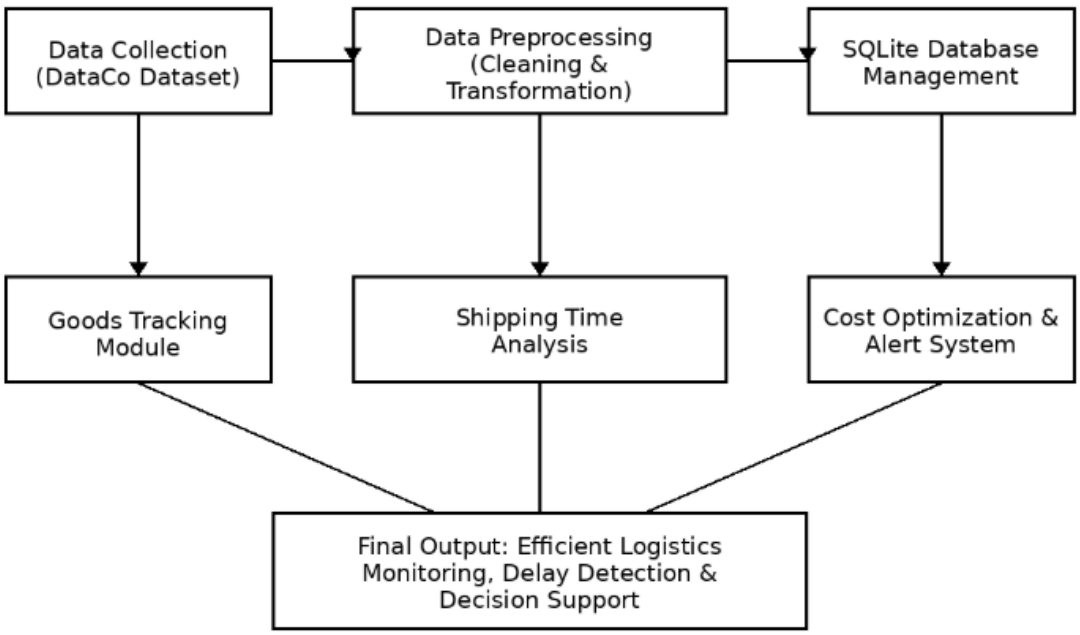

Figure 1: Various Methodologies of Proposed Architecture

VI. ARCHITECTURE DIAGRAM

1. Input Layer

The input layer is responsible for collecting logistics and shipment-related data from the DataCo Supply Chain dataset. The dataset contains information such as shipment details, customer information, delivery status, shipping modes, order locations, and profit details. This layer acts as the starting point of the logistics analysis system.

2. Data Preprocessing Layer

In this layer, the collected data is cleaned and transformed for analysis. Missing values, unwanted symbols, and inconsistent data are removed. Additional features such as shipping delay are generated by comparing actual shipping time with scheduled shipping time. The processed data is then prepared for storage and analysis.

3. Database Management Layer

The cleaned logistics data is stored in an SQLite database for efficient management and retrieval. SQL queries are used to access shipment records, analyse delivery risks, and manage logistics information effectively. This layer improves data organization and querying performance.

4. Analysis and Processing Layer

This layer performs logistics analysis using Python and data analytics techniques. Various modules such as goods tracking, shipping time analysis, cost optimization, and delayed shipment detection are executed in this stage. The system identifies shipment trends, delivery delays, and profitability insights for decision-making.

5. Visualization and Alert Layer

The analysed results are represented using charts, graphs, and reports through data visualization techniques. An alert and notification system is also implemented to identify delayed or high-risk shipments and provide early notifications to logistics managers.

6. Output Layer

The final output layer displays logistics insights, shipment tracking details, delay analysis, profit analysis, and decision-support information in a structured format. The output helps improve logistics efficiency, reduce delivery delays, optimize costs, and enhance supply chain management.

Figure. 2: Architecture Diagram

VII. INPUT DATA DESCRIPTION

The input data for this project is collected from the "DataCo Supply Chain" dataset, which provides detailed information regarding international shipping and logistics operations. The dataset contains various attributes such as transaction type, actual shipping days, scheduled shipping days, shipping modes, customer details, and product information to analyze shipment and delivery performance effectively.

A new attribute called "shipping delay" is generated by comparing actual shipping days with scheduled shipping days to identify delays in the shipment process. The dataset also includes product-related information such as category name and customer-related information such as customer country, which helps in analyzing shipment distribution across different regions.

Furthermore, location-based attributes including order city, order country, and order region are used to track the movement and distribution of goods between various locations. Financial attributes such as order item total and profit per order are utilized for cost and profitability analysis.

All these parameters collectively serve as the foundation for shipment tracking, delay detection, logistics analysis, and decision-making in the proposed international shipping logistics system.

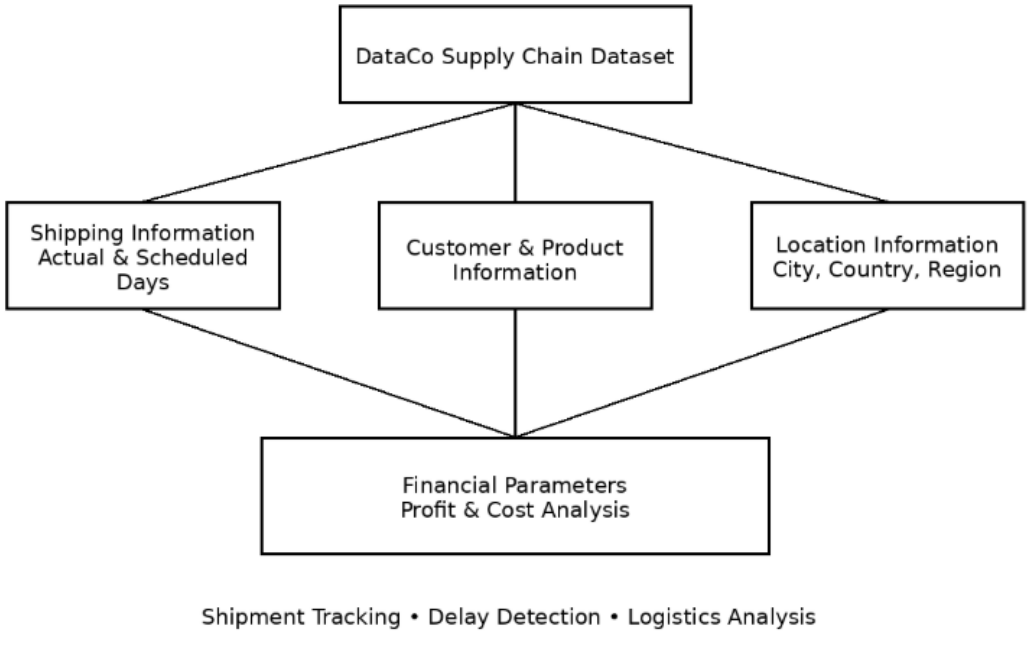

Figure 3: Input Data

VIII. PSEUDO CODE AND IMPLEMENTATION

#Import libraries

BEGIN

1. Import required libraries

Pandas, NumPy, Matplotlib, Seaborn, SQLite

2. Load Dataset

Read DataCoSupply.csv file into dataframe

3. Data Cleaning

- Remove spaces and special characters from column names
- Convert column names to lowercase

4. Feature Engineering

- Calculate shipping_delay = days_for_shipping_real - days_for_shipment_scheduled
- Convert late_delivery_risk to numeric format

5. Select Required Columns

- Extract logistics-related columns
- Remove missing values

6. Database Integration

- Create SQLite connection
- Store cleaned data into database table (shipping_data)

7. Perform SQL Queries

- Query 1: Calculate average late delivery risk by market
- Query 2: Analyze shipping mode performance (average delay

8. Goods Tracking Module

- Extract order_city, order_country, shipping_mode, delivery_status
- Display shipment tracking details

9. Shipping Time Analysis

- Calculate average real shipping time
- Calculate average scheduled shipping time
- Compare both values

10. Cost Analysis

- Group data by market
- Calculate average profit per order
- Identify high-performing markets

11. Notification Alert System

- Filter records where late_delivery_risk = 1
- Display number of risky shipments
- Show sample alerts

12. Data Visualization

- Plot shipment distribution by country
- Plot shipping time comparison
- Plot profit analysis
- Plot late delivery risk

13. Generate Insights

- Identify high-risk market
- Identify shipping mode with highest delay
- Display conclusions

END

IX. OUTPUT ANALYSIS

The output analysis evaluates the overall performance and efficiency of the proposed international shipping logistics system. The system processes logistics and shipment-related data to generate meaningful insights regarding shipment tracking, delivery performance, shipping delays, and cost optimization. The generated outputs help logistics managers and organizations make better operational decisions and improve supply chain efficiency.

The goods tracking output provides information about the movement of shipments across different countries and regions. It helps in monitoring delivery status, identifying shipment locations, and improving supply chain visibility. The shipping time analysis compares actual shipping time with scheduled shipping time to identify delayed deliveries and measure shipping efficiency.

The cost and profit analysis output helps in evaluating the profitability of different markets and shipment categories. By analysing logistics costs and profits, the system supports better financial planning and cost optimization. The notification alert system identifies high-risk shipments and generates alerts for delayed deliveries, enabling logistics managers to take corrective actions at the right time.

Various visualization techniques such as charts, graphs, and comparative reports are used to represent logistics insights clearly and effectively. These visual outputs help users easily understand shipment trends, delivery performance, delay patterns, and market-wise profitability. Overall, the output analysis demonstrates that the proposed system improves logistics monitoring, reduces shipment delays, enhances decision-making, and increases operational efficiency in international shipping logistics.

X. RESULTS AND DISCUSSIONS

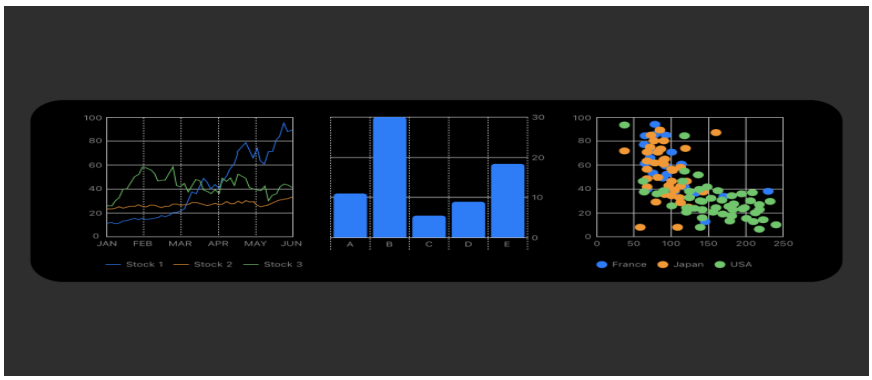

Figure 4: Output Image

Figure 4 The above figure represents the visual output analysis of the proposed international shipping logistics system. The graphs are used to analyze shipment performance, logistics trends, and delivery efficiency. The first graph illustrates the variation and performance trends over different time periods. The second graph represents comparative analysis between different shipment categories and operational values. The third graph displays the relationship between multiple shipment parameters across different countries and regions.

XI. CONCLUSION

The project, “Streamlining International Shipping Logistics with Database Solutions,” was a success in demonstrating the application and benefits of data analytics and database systems in enhancing international shipping logistics. This was made possible by utilizing the data available from the DataCo Supply Chain dataset, which was then used to analyze and monitor the movement of goods from one region to another. The application of data preprocessing and feature engineering, such as computing the shipping delay, helped in identifying the problems in the shipping process. The utilization of an SQLite database helped in efficiently managing and querying the data for efficient analysis. The project demonstrates that integrating data analytics with database systems improves logistics efficiency and decision-making.

XII. REFERENCES

[1] A. K. Gupta and R. Singh, "Data Analytics in Supply Chain Management: A Review," International Journal of Logistics Systems and Management, vol. 35, no. 2, pp. 123–135, 2020.

[2] M. Christopher, Logistics and Supply Chain Management, 5th ed. London, U.K.: Pearson Education, 2016.

[3] S. Chopra and P. Meindl, Supply Chain Management: Strategy, Planning, and Operation, 7th ed. New Delhi, India: Pearson, 2019.

[4] J. Wang, Y. Zhang, and L. Liu, "Big Data Analytics for Intelligent Logistics Management," IEEE Access, vol. 6, pp. 27562–27574, 2018.

[5] R. K. Jain and S. Bansal, "Optimization of Transportation and Logistics Using Data Mining Techniques," International Journal of Computer Applications, vol. 182, no. 10, pp. 25–30, 2019.