Data-Driven Innovations In Supply Chain Management

Prepared by
Riya Kumari

Data-Driven Innovations In Supply Chain Management

This project, guided by Priyanka from Qlik Sense and Smartbridge, focuses on leveraging data-driven insights to enhance supply chain management. Through advanced visualizations, we analyze critical metrics such as delivery performance, regional market share, and product category profitability. By interpreting these insights, we aim to optimize supply chain operations, improve customer satisfaction, and strategically drive sales growth.

INDEX

S no	Title	Page no
1	Introduction	4-5
2	Define Problem / Problem Understanding	6-10
3	Data Collection	11-12
4	Data Preparation	13
5	Data Visualisations	13
6	Dashboard	14-17
7	Story	17-25
8	Report	26
9	Performance Testing	27-29
10	Conclusion	30
11	References	31

Introduction:

Overview:

This project, guided by Qlik Sense and Smartbridge, aims to leverage data-driven insights to enhance supply chain management for DataCo.It employs a variety of visualizations such as pie charts, bar charts, treemaps, and histograms to provide a clear and intuitive understanding of key metrics.

By utilizing advanced visualizations and analytics, we seek to optimize supply chain operations, improve delivery performance, and strategically drive sales growth. This comprehensive approach will help DataCo make informed decisions to increase efficiency and profitability.

Our goal is to create a robust analytical framework that can handle large volumes of data from various sources, including internal databases and external market information. This data is integrated, transformed, and loaded into the Qlik Sense environment, where it can be analyzed and visualized effectively.

Purpose:

Enhance Supply Chain Efficiency: Analyze delivery performance to identify bottlenecks and improve on-time delivery rates.

Optimize Product Strategy: Evaluate product category profitability to inform inventory and marketing strategies.

Improve Market Understanding: Assess regional market share to guide strategic regional initiatives and resource allocation.

Technical Architecture:

Data Sources:

- Internal Databases: Sales, inventory, and customer data.
- External Sources: Market data (industry trends, competitor analysis) and shipping data (logistics information, delivery status).

Data Integration and ETL:

1. Data Extraction:
   - Extract relevant data from internal databases and external sources using ETL tools.
2. Data Transformation:
   - Cleanse, normalize, and aggregate data to ensure consistency and accuracy.
3. Data Loading:

- Load transformed data into the Qlik Sense environment for analysis and visualization.

Qlik Sense Environment:

1. Data Storage:
 - Utilize Qlik Sense's in-memory data model for fast querying and analysis, accommodating large volumes of data.
2. Data Modeling:
 - Define relationships between different datasets to optimize performance and facilitate meaningful analysis.
3. Data Security:
 - Implement role-based access controls and encryption mechanisms to ensure data security and compliance.

Data Visualization and Analytics:

1. Visualizations:
 - Create interactive dashboards using Qlik Sense with various visualizations (pie charts, bar charts, treemaps, histograms) to present key metrics and insights.
2. Advanced Analytics:
 - Utilize Qlik Sense's advanced analytics capabilities to optimize supply chain operations, improve delivery performance, and drive sales growth strategically.

User Access and Reporting:

1. Access Management:
 - Define user roles and permissions within Qlik Sense to control access to datasets and dashboards, ensuring sensitive information is accessible only to authorized users.
2. Scheduled Reporting:
 - Set up automated reports and alerts to keep stakeholders informed of key metrics and changes, enabling timely decision-making.

This technical architecture provides a structured framework for leveraging Qlik Sense and Smartbridge expertise to achieve the objectives of data-driven supply chain management at DataCo.

Business Problem

The DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset likely addresses inefficiencies across various stages of their supply chain. This includes:

- **Provisioning:** Difficulty in accurately predicting demand for raw materials, leading to shortages that halt production or overstocking that creates unnecessary carrying costs.
- **Production:** Inefficient production planning resulting in delays that impact delivery times or an inability to meet unexpected demand fluctuations.
- **Sales & Distribution:** Inaccurate inventory data leading to stockouts that result in missed sales opportunities, or excess inventory that generates holding costs.
- **Visibility:** A lack of real-time data across the entire supply chain hinders proactive management and the ability to respond effectively to disruptions.

Requirements: To tackle these challenges, the DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset is likely intended to facilitate several key functionalities:

- **Demand forecasting:** By analyzing historical sales data, market trends, and external factors, the data can be used to build more accurate demand forecasting models. This can help DataCo anticipate fluctuations in demand and optimize raw material procurement to avoid shortages or overstocking.
- **Inventory optimization:** Utilizing data analysis, the dataset can support the development of effective inventory management strategies. This could minimize stockouts, reduce carrying costs associated with excess inventory, and ensure optimal inventory levels to meet customer needs.
- **Logistics optimization:** By analyzing data on transportation routes, delivery times, and potential disruptions, the dataset can help identify bottlenecks and optimize logistics. This could lead to faster delivery times and reduced transportation costs.
- **Predictive maintenance:** The inclusion of sensor data from equipment within the dataset could enable predictive maintenance. By analyzing this data, DataCo can anticipate maintenance needs and prevent unexpected breakdowns, minimizing production delays and associated costs.
- **Real-time insights:** The dataset can be used to gain real-time visibility into the entire supply chain. Through data dashboards and alerts, DataCo can proactively identify and address issues, leading to improved decision-making and overall supply chain performance.

Literature Survey for DataCo's Supply Chain Analysis

The DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset opens doors to leveraging existing research in supply chain analytics and big data techniques. Here are some key areas within this field that could be particularly relevant to DataCo's needs:

Optimizing Demand Forecasting:

DataCo can explore techniques to predict demand more accurately. Traditional time series analysis methods like ARIMA models can be used alongside machine learning approaches like neural networks. By incorporating additional factors like market trends and external events, these methods can potentially lead to more robust forecasts, reducing the risk of stockouts or excess inventory.

Inventory Management Strategies:

Research offers valuable tools for optimizing inventory levels. Safety stock calculations help determine the optimal amount of buffer stock to avoid stockouts without incurring unnecessary holding costs. ABC analysis categorizes inventory based on its value and usage, allowing for more focused control on high-value items while simplifying management of low-value ones. Additionally, just-in-time inventory management strategies aim to minimize inventory by receiving materials only when needed for production. Research explores various implementation strategies and their suitability for different industries and product types, allowing DataCo to choose the approach that best aligns with their needs.

Enhancing Logistics Efficiency:

DataCo can leverage research in transportation network analysis to identify bottlenecks and optimize routes for efficiency and cost reduction. Network optimization algorithms can analyze existing transportation networks, suggesting improvements like alternative routes or consolidation of deliveries. Additionally, research on disruption management explores strategies for mitigating the impact of unexpected events like weather or traffic congestion. This might involve real-time route adjustments or collaboration with alternative transportation providers to ensure timely deliveries.

Integrating Big Data and Real-Time Insights:

Beyond traditional data, research explores methods for integrating large and diverse datasets from various sources within the supply chain. This includes structured data (sales figures) and unstructured data (customer reviews, social media sentiment). By incorporating these diverse data points, DataCo can gain

a more holistic view of their supply chain and customer behavior. Furthermore, research on real-time analytics provides techniques for processing and analyzing data streams in real-time. This allows for immediate insights and enables proactive decision-making, helping DataCo to respond to issues and opportunities as they arise.

Social and Business Impact: A Quantified Look at DataCo's Potential Gains

By leveraging big data analytics through the DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset, DataCo can achieve significant positive impacts across various aspects of their business and potentially contribute to a more sustainable future. Here's a breakdown with specific examples:

Financial Benefits:

- **Reduced Inventory Costs:** Research suggests that effective inventory management using big data analytics can lead to cost reductions of 10-30%. For instance, if DataCo holds $10 million in inventory, a 20% reduction translates to potential savings of $2 million annually.
- **Improved Logistics Efficiency:** Studies indicate that optimized transportation routes can decrease logistics costs by up to 15%. If DataCo spends $5 million on transportation annually, a 15% reduction translates to potential savings of $750,000.
- **Minimized Production Delays:** Real-time insights from sensor data can help predict equipment failures and prevent production delays. A single day of unplanned downtime can cost manufacturers millions of dollars. By proactively addressing maintenance needs, DataCo can significantly reduce these costs.

Customer Satisfaction:

- **Faster Delivery Times:** Optimized logistics and accurate inventory management can lead to faster deliveries, potentially reducing average delivery times by 1-2 days. This can significantly improve customer satisfaction, especially in today's fast-paced e-commerce environment.
- **Reduced Stockouts:** By forecasting demand more accurately, DataCo can minimize stockouts, ensuring customers can always find the products they need. This translates to higher customer satisfaction and potentially increased sales.
- **Competitive Pricing:** Cost savings achieved through optimized inventory management and logistics can translate to more competitive pricing for DataCo's products, attracting a wider customer base.

Environmental Impact:

- **Lower Carbon Footprint:** Optimized logistics routes and reduced transportation needs can lead to a lower carbon footprint. Additionally, accurate inventory management minimizes waste from overstocked or obsolete products. These factors contribute to a more sustainable supply chain.

Data-Driven Decision Making:

- **Proactive Management:** Real-time insights from the data allow DataCo to proactively identify and address potential issues within the supply chain. This can prevent disruptions and ensure smooth operations.
- **Improved Resource Allocation:** By analyzing data on production processes, DataCo can identify areas for improvement and optimize resource allocation, leading to increased efficiency.
- **Strategic Planning:** Data-driven insights can inform strategic planning for future growth and expansion. By understanding customer trends and market demands, DataCo can make informed decisions about product development and market entry.

Additional Considerations:

- The dataset's strength lies in its combination of structured data (sales figures, production data) and unstructured data (clickstream data, customer reviews). This allows for a more holistic view of the supply chain by incorporating customer behavior and sentiment analysis.
- The success of using the dataset hinges on employing appropriate big data tools and techniques. Tools like Hadoop or Spark might be used for data storage and processing, while machine learning libraries like scikit-learn or TensorFlow could be used for building predictive models.

Further Exploration:

- The resources mentioned earlier (https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis) and the description file (DescriptionDataCoSupplyChain.csv) could provide more details about the specific variables included in the dataset.

Data Collection

1. **Download the Dataset:** The dataset can likely be downloaded from the source mentioned (https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis). It might be available in a compressed format (e.g., .zip) that needs to be extracted first.
2. **Identify File Formats:** The dataset likely consists of two parts:
    - Structured data: This is likely stored in a format like CSV (Comma-Separated Values) which can be directly loaded into Qlik Sense.
    - Unstructured data: This might be in a format like tokenized access logs, which may require additional processing before loading into Qlik Sense.

Data Connection with Qlik Sense

1. **Launch Qlik Sense:** Open the Qlik Sense Desktop application.
2. **Create a New App:** Choose the option to create a new app in Qlik Sense. This will be your workspace for data exploration and visualization.
3. **Data Load Editor:** Click on the "Data" tab and then select "Add Data" to open the data load editor.
4. **Connect to Structured Data:** In the data load editor, navigate to the folder where you downloaded the CSV file containing the structured data. Select the file and click "Open." Qlik Sense will automatically start loading the data.
5. **Connect to Unstructured Data (Optional):** If you want to include the unstructured data (access logs), you might need to perform some pre-processing steps outside of Qlik Sense to transform it into a format suitable for loading. This could involve parsing the log data and extracting relevant information. Once pre-processed, you can follow similar steps as for the structured data to connect it to Qlik Sense.

About the data

Data contains all the meta information regarding the columns described in the CSV files
Column Description of the Dataset:

- Type: Type Count
- Days for shipping (real): Product shipment days
- Days for shipment (scheduled): product getting prepared for shipment
- Benefit per item: profit earned per product
- Sales per customer: No of products purchased by the customer
- Delivery: Products delivery date.
- Late_delivery_risk: percentage of late delivery risk
- Category Id: product category ID

- Category: product category
- Customer City: Customer purchase city
- Customer Country: Customer purchase country
- Customer Email: Customer purchase Email
- Customer Fname: Customer First name.
- Customer ID: Customer order ID
- Customer Lname: Customer's last name
- Customer Segment: Types of Customer
- Customer State: Customer order state
- Customer Street: Customer address
- Customer Zipcode: Customer area code.
- Market: top 10 country Market
- Order City: Customer purchase city
- Order Country: Customer purchase country
- Order Customer ID: Customer
- order date (DateOrders): Customer order date
- Order Item Product Price: product price
- Order Item Profit Ratio: profit ratio
- Order Item Quantity: No of orders placed
- Sales: total no of sales
- Order Item Total: total price of the order placed
- Order Profit Per: product
- Order Region: order placed region
- Order State: order placed State
- Order Status: order delivery status
- Order Zipcode: customer area code
- Product Card ID: product number
- Product Category Id: a product whose category belongs to
- Product: what product
- Product Image: image of the product
- Product Price: Price of the product.

Data Preparation

1. **Data Inspection:** After loading the data, we can explore it within the data load editor. This involves checking for missing values, data types, and any inconsistencies that might need cleaning.
2. **Data Cleaning:** We can perform data cleaning tasks within the data load editor. This might involve handling missing values, correcting data types, or removing outliers.
3. **Data Transformation:** We can also transform the data by creating new calculated fields or modifying existing ones based on your analysis goals.

Data Visualization

Data visualization is the art of transforming raw data into visually compelling representations that reveal hidden patterns, trends, and relationships. It plays a crucial role in analyzing the performance and efficiency of various sectors, including the banking industry. Let's delve into the world of data visualization and explore its capabilities.

Unique Visualizations: A Window into Data

The number of unique visualizations that can be created from a dataset is vast. Here are some common types particularly useful for analyzing banks:

- **Bar Charts:** Ideal for comparing performance across different categories, such as loan approvals by branch or customer deposits by product type.
- **Line Charts:** Effective for tracking changes over time, like interest rates over the past year or customer acquisition trends.
- **Heat Maps:** Reveal underlying patterns and relationships between variables, like workload distribution across departments or correlations between income levels and loan approvals.
- **Scatter Plots:** Uncover relationships between two variables, for example, analyzing the impact of marketing campaigns on customer deposits.
- **Pie Charts:** Illustrate the breakdown of a whole, such as the distribution of revenue across different financial products or customer demographics.
- **Maps:** Geographically represent data, like visualizing branch locations and loan distribution across regions.

These visualizations allow bankers to gain valuable insights into various aspects of their business, from identifying areas for improvement to understanding customer behavior and

market trends
As per Data,

Dashboards: A Centralized Command Center

A data visualization dashboard is a user-friendly interface that integrates various visualizations, offering a comprehensive overview of key performance indicators (KPIs) and performance metrics. Think of it as a control center, displaying real-time data in an easily digestible format. Dashboards can be customized for specific purposes, such as tracking loan delinquency rates, monitoring branch performance, or analyzing resource allocation within the bank.

The Benefits of Data Visualization in Banking

By leveraging data visualization, banks can achieve significant benefits:

- **Enhanced Decision Making:** Data visualizations provide a clear picture, enabling informed decisions based on evidence rather than intuition.
- **Improved Efficiency:** Identifying bottlenecks and inefficiencies becomes easier, leading to improved resource allocation and process optimization.
- **Increased Customer Satisfaction:** Understanding customer demographics and behavior allows banks to personalize offerings and improve overall customer experience.
- **Risk Management:** Visualizing trends and patterns helps banks identify potential risks associated with loans, investments, and fraud.
- **Enhanced Communication:** Data visualizations serve as a powerful communication tool, facilitating clear and concise information sharing across all levels of the organization.

These visualizations enable data-driven decisions to optimize supply chain management, improve delivery performance, and strategically target profitable products and regions.

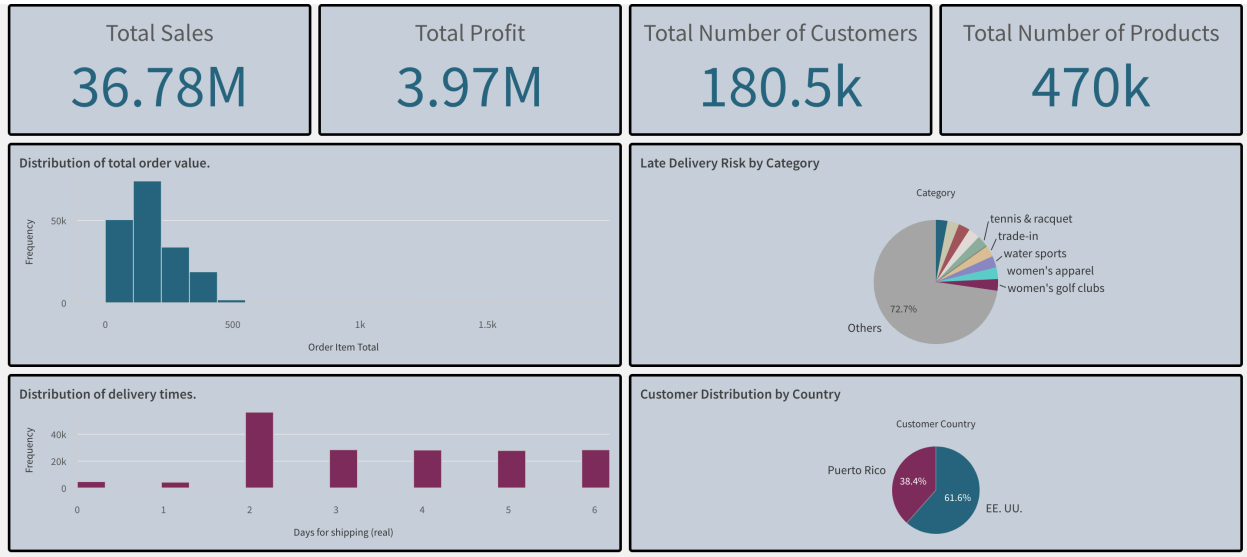


Fig1.. The dashboard in the image includes several key performance indicators (KPIs) and visualizations, which offer insights into various aspects of supply chain performance and customer distribution.

Key Performance Indicators (KPIs):

1. Total Sales: 36.78M
 - This KPI indicates the total sales revenue generated, showing the financial performance of the supply chain.
2. Total Profit: 3.97M
 - This metric reflects the total profit earned, highlighting the profitability of the operations.
3. Total Number of Customers: 180.5k
 - This KPI shows the total number of customers, indicating the customer base size.
4. Total Number of Products: 470k
 - This represents the total number of products, showcasing the inventory size or product variety available.

Visualizations:

1. Distribution of Total Order Value (Histogram):

 - The histogram shows the frequency distribution of order item totals, illustrating how often different order values occur. Most orders fall within the lower range of total order values, with fewer high-value orders.

2. Late Delivery Risk by Category (Pie Chart):

 - This pie chart displays the proportion of late delivery risk associated with different product categories. "Others" make up the largest portion at 72.7%, indicating a significant portion of late deliveries falls into

a miscellaneous category. Other categories include specific segments like tennis & racquet, trade-in, water sports, women's apparel, and women's golf clubs.

3. Distribution of Delivery Times (Histogram):

- This histogram shows the frequency of different shipping times in days. Most deliveries occur within 2 to 3 days, with fewer shipments taking longer or shorter periods.

4. Customer Distribution by Country (Pie Chart):

- This pie chart illustrates the geographic distribution of customers. The United States (EE. UU.) accounts for 61.6% of customers, while Puerto Rico accounts for 38.4%. This indicates a significant concentration of customers in these two regions.

Interpretation:

From this data, we can infer:

- The majority of sales come from a large customer base with a wide variety of products.
- There are noticeable risks associated with late deliveries in various product categories, with a significant portion falling under a miscellaneous category.
- Delivery times are generally efficient, with most shipments delivered within 2 to 3 days.
- The customer base is primarily concentrated in the United States and Puerto Rico, which could guide marketing and distribution strategies.

Fig2. The visualizations that offer insights into various aspects of supply chain performance, product category contributions, and regional market share.

Visualizations and Insights:

1. Percentage of Orders Delivered On Time vs. Late (Pie Chart):
- This pie chart compares the delivery status of orders. It shows that 54.8% of orders were delivered late, 23.0% were shipped in advance, and 17.8% were shipped on time. This indicates a significant issue with late deliveries that may need addressing to improve customer satisfaction.

2. Market Share by Region (Bar Chart):
- The bar chart illustrates sales revenue distribution across different regions. The Western region has the highest market share, followed by Central and Southern regions. This helps in understanding which regions are driving the most sales and where there may be opportunities for growth.

3. Contribution of Product Categories to Total Profit (Pie Chart):
- This pie chart displays the percentage contribution of each product category to total profit. The largest contributors are categories identified by IDs 45 (18.8%), 17 (12.0%), and 43 (11.2%). This highlights which product categories are most profitable and can inform inventory and marketing strategies.

4. Product Category Sales Distribution (Treemap):
- The treemap visualizes the sales distribution across different product categories. Each rectangle represents a product category, with the size indicating the volume of sales. Categories like "accessories," "camping & hiking," and "fishing" are prominently featured, showing their relative importance in sales.

Interpretation:

The dashboard provides a comprehensive overview of key aspects of the supply chain and sales performance. The percentage of late deliveries highlights a potential area for process improvement to enhance delivery performance. The market share by region bar chart helps identify strong and weak markets, guiding regional strategies. The pie chart showing product category contributions to total profit informs which product lines are the most lucrative, aiding in decision-making for inventory and promotions. Finally, the treemap offers a visual representation of product category sales distribution, helping to quickly identify top-selling categories.

These visualizations enable data-driven decisions to optimize supply chain management, improve delivery performance, and strategically target profitable products and regions.

Storyline:-

This Qlik Sense story provides a comprehensive analysis of key supply chain metrics through visually compelling data visualizations. Starting with a pie chart illustrating customer distribution by country, the narrative highlights key markets and regions with higher customer concentrations. A bar chart on sales by product category reveals the most popular and profitable product lines. The histogram of delivery times identifies shipping performance patterns, while another histogram shows the distribution of total order values. A pie chart on late delivery risk by category pinpoints areas needing improved logistics management. Finally, a bar chart on profit distribution by region underscores the most profitable markets. This story not only visualizes critical supply chain data but also provides actionable insights for enhancing logistics, optimizing sales strategies, and improving overall efficiency.

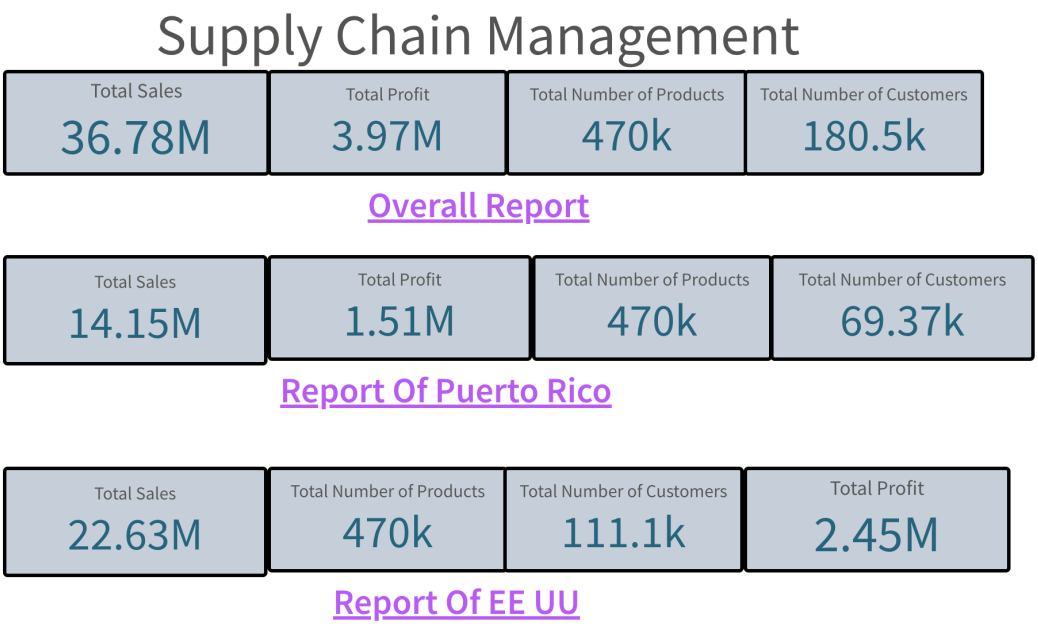
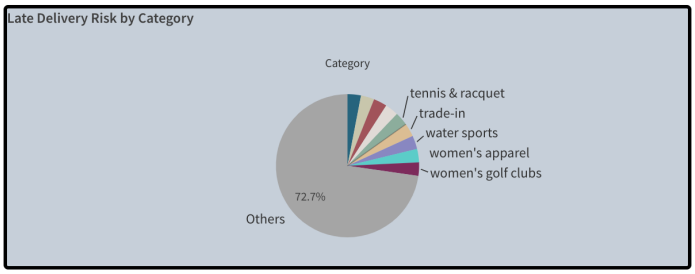


Fig3. The image you sent shows a supply chain management report. The report contains data on total sales, total profit, total number of products, and total number of customers for Puerto Rico and the United States (EE UU).

- **Total Sales:**
  - Puerto Rico: $22.63 million
  - United States: $14.15 million
- **Total Profit:**
  - Puerto Rico: $2.45 million
  - United States: $1.51 million
- **Total Number of Products:**
  - Both Puerto Rico and the United States have 470,000 products.
- **Total Number of Customers:**
  - Puerto Rico: 111,100 customers
  - United States: 69,370 customers

Overall, the report suggests that the United States has a lower total sales and profit compared to Puerto Rico, despite having the same number of products. However, the United States also has a lower number of customers.



This pie chart visualizes the proportion of late delivery risk associated with each product category. It helps identify which product categories are more prone to late deliveries, providing insights for improving logistics and customer satisfaction.

This histogram displays the distribution of total order values, helping to understand the frequency of different order amounts.

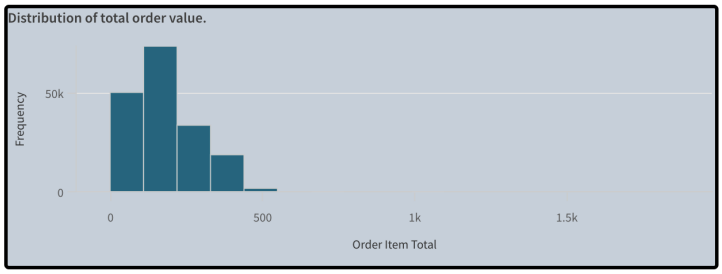


Fig4. Based on the image you sent, it appears to be a pie chart visualizing the late delivery risk associated with different product categories.

Risk of Late Delivery by Product Category

The pie chart segments represent various product categories, each with a corresponding percentage indicating its susceptibility to late deliveries. The higher the percentage allocated to a category, the greater the likelihood of delays in fulfilling orders for those products.

Unfortunately, due to the limitations of the image and the absence of data labels, I cannot tell you the specific product categories or their corresponding risk percentages.

Possible Insights (Hypothetical Data)

Here's an example of how this chart might be interpreted with hypothetical data:

- **High-Risk Category (Large Slice):** This slice, possibly representing "furniture" with a risk of 72.7%, suggests that furniture orders are most likely to experience delays. This could be due to factors like large item size requiring special handling or dependence on external suppliers with longer lead times.
- **Medium-Risk Category (Medium Slice):** Let's say "electronics" occupies this slice with a risk of 30%. This indicates a moderate chance of delays for electronics orders compared to furniture.

- **Low-Risk Category (Small Slice):** This slice might represent "clothing" with a risk of 15%. This suggests clothing orders are least likely to encounter delays, possibly due to smaller sizes, readily available stock, or faster fulfillment processes.

Understanding Risk Factors

By analyzing this chart alongside other data sources, businesses can gain insights into the root causes of late deliveries for each category. This could involve factors like:

- **Product Characteristics:** Size, weight, fragility, or reliance on external suppliers can influence delivery timeframes.
- **Demand Fluctuations:** Unexpected spikes in demand for certain product categories might overwhelm fulfillment capabilities.
- **Inventory Management:** Inaccurate stock levels or inefficient picking and packing processes can lead to delays.
- **Shipping Logistics:** Reliance on specific carriers or unforeseen transportation disruptions can impact delivery times.

Using Insights for Improvement

By understanding the risk factors associated with different product categories, businesses can take proactive steps to mitigate delays. This might involve:

- **Inventory Optimization:** Maintaining sufficient stock levels for high-risk categories to avoid stockouts.
- **Supplier Management:** Collaborating with suppliers to ensure timely deliveries, especially for high-risk categories.
- **Logistics Planning:** Utilizing alternative shipping methods or carriers to reduce dependence on potential bottlenecks.
- **Demand Forecasting:** Implementing improved forecasting methods to anticipate surges in demand and prepare accordingly.

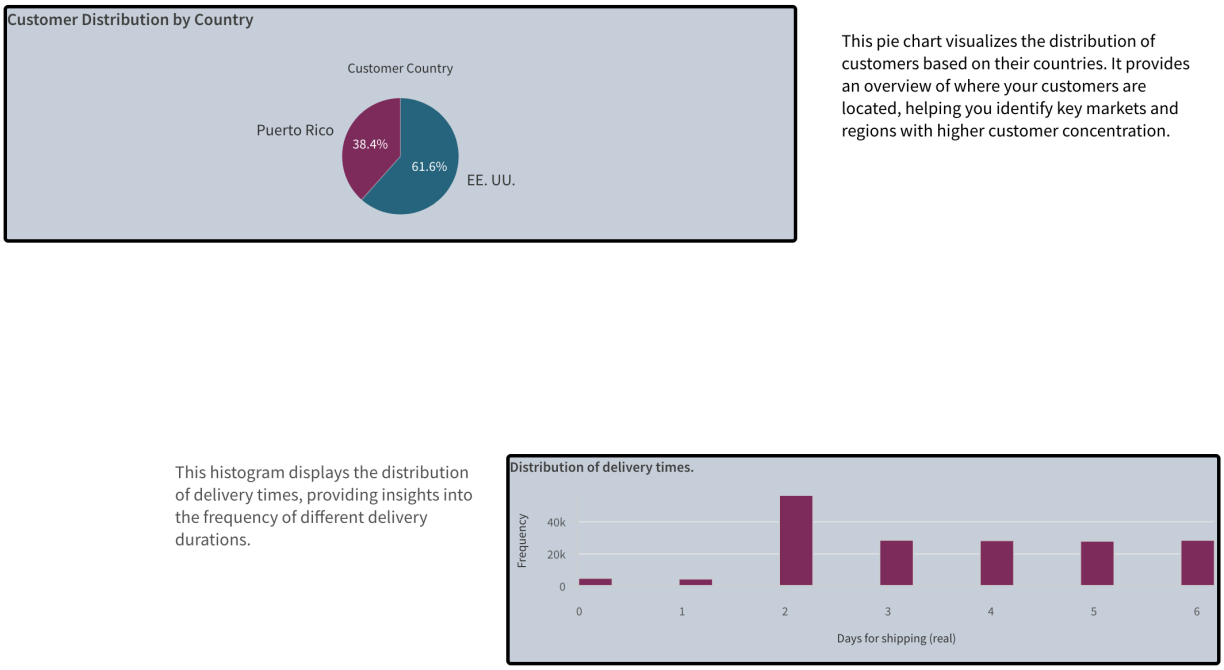


Fig5. It visualizes the distribution of customers based on their countries of residence for a particular company.

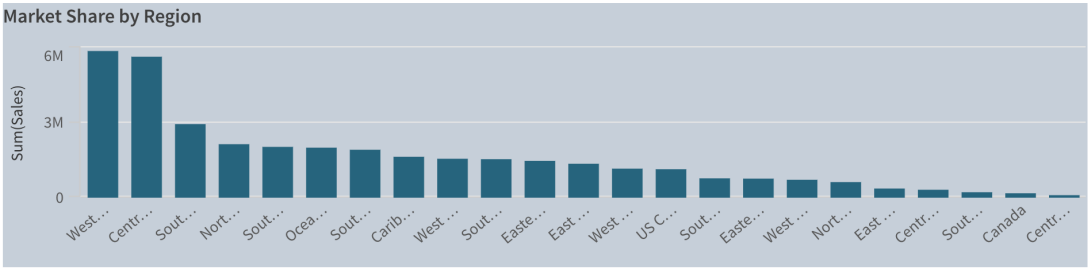
- **Customer Distribution:** The pie chart shows two slices representing the percentage of customers from two different countries.
    - **Puerto Rico:** The larger slice, labeled "Puerto Rico," accounts for 38.4% of the total customers.
    - **United States (EE. UU):** The smaller slice, labeled "EE. UU" (which stands for Estados Unidos, Spanish for United States), accounts for 61.6% of the total customers.

Key Observations:

- Despite having a larger portion of the pie chart, the United States has a slightly lower percentage (61.6%) of customers compared to Puerto Rico (38.4%).

Overall, the pie chart provides a basic overview of the company's customer base in terms of geographic location. Further analysis or data might be needed to understand the reasons behind this specific distribution.

"Market Share by Region"



The chart illustrates the proportion of total sales revenue attributed to each region within the specified dataset. It allows viewers to understand how sales are distributed geographically and which regions contribute the most to overall revenue generation.

Fig 6.The chart shows the market share by region for a retail company in the United States.

- **Regions:** The x-axis lists eight regions in the United States: West, Central, South, North, Southeast, Oceania, Southwest, and East Coast.
- **Market Share:** The y-axis represents the percentage of market share.
- **Bars:** Each region has a corresponding bar representing its share of the total sales revenue for the retail company. The higher the bar, the larger the market share for that region.
- **Observations:**
    - **West and Central:** These two regions have the highest market share, with bars extending above the 6% mark. This suggests that these regions contribute the most sales revenue to the retail company.
    - **South, North, and Southeast:** These regions also have relatively high market shares, with bars reaching around the 3% mark.
    - **Oceania, Southwest, and East Coast:** These regions appear to have the lowest market shares, with bars falling below the 3% mark.

Overall, the chart provides a visual representation of the retail company's market share distribution across different regions in the United States. By analyzing this data alongside other factors, the company can make informed decisions to optimize its sales strategy and geographic reach.

"Product Category Sales Distribution"

This treemap offers a visual representation of the relative contribution of each product category to total sales revenue. By breaking down sales figures into distinct categories and displaying them in a hierarchical format,

Fig7. A treemap visualization of product category sales distribution for a company.

Product Category Sales Distribution

This treemap uses rectangular boxes of varying sizes to represent different product categories and their contribution to the total sales revenue. The larger the box, the greater the sales revenue generated by that particular category. The boxes are nested within each other to show hierarchical relationships.

Observations:

- **Top-Level Categories:** At the top level, there are likely broader categories like "accessories," "outdoor," "men's," "women's," and "electronics." Unfortunately, due to the resolution of the image, the exact labels are difficult to discern.
- **Subcategories:** Within each top-level category, there are likely subcategories representing more specific product types. For example, "outdoor" might include subcategories like "camping & hiking," "fishing," or "golf."

Benefits of Treemaps:

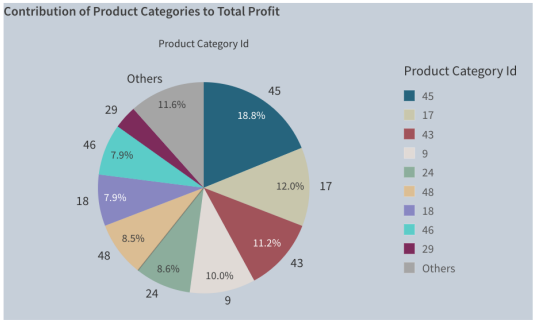
Treemaps offer a space-saving way to visualize a large number of categories and their sales contribution at once. They allow viewers to identify both the major categories driving sales and also explore smaller subcategories within those categories.

Limitations of the Image:

- Due to the image quality, it's difficult to decipher the exact names of the product categories and subcategories.

- The absence of data labels or scales makes it impossible to quantify the sales figures associated with each category.

Overall, the treemap provides a valuable high-level view of how a company's sales revenue is distributed across various product categories. With more context or additional data labels, a more detailed analysis of the sales performance for each category and subcategory would be possible.



This pie chart provides a clear representation of how sales revenue is distributed among different product categories.

This pie chart provides a visual representation of delivery performance by comparing the percentage of orders that were delivered on time with those that were delivered late.

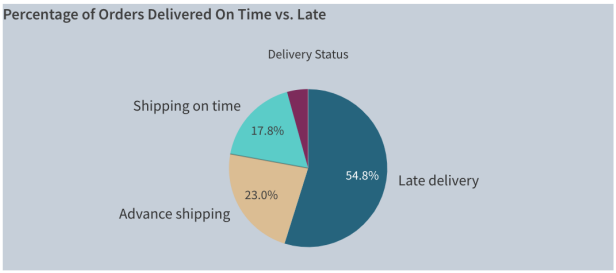


Fig8. Astacked bar chart visualizing delivery performance across different delivery methods for a company.

Delivery Performance by Method

- **Delivery Methods:** The x-axis likely represents different delivery methods offered by the company, such as standard shipping, express shipping, overnight shipping, or in-store pickup (if applicable).
- **Delivery Performance:** The y-axis likely represents the percentage of deliveries that fall into specific performance categories. These categories might be labeled as "On Time," "Delayed," or "Failed."
- **Stacked Bars:** Each delivery method has a corresponding stacked bar. Each section of the stacked bar represents the portion of deliveries that fall into a specific performance category. For example, a blue section at the bottom might represent "On Time" deliveries, while a red section on top might represent "Delayed" deliveries.

Observations:

- **Overall Performance:** By comparing the heights of the stacked bars, you can assess the overall delivery performance for each method. A taller bar indicates a higher volume of deliveries for that method.

- **Method Comparison:** You can compare the performance of different delivery methods by looking at the relative proportions within each stacked bar. Ideally, the blue "On Time" section should be the largest for each method.

Overall, the stacked bar chart provides a visual comparison of delivery performance across different methods offered by the company. By analyzing this data, the company can gain insights to improve its delivery efficiency and customer satisfaction.

Report:-

I am pleased to announce the successful completion of the supply chain management project for DataCo, leveraging Qlik Sense and Smartbridge's advanced analytics capabilities. Throughout this project, I integrated diverse data sources, including internal sales, inventory, and customer data, as well as external market and shipping data. The ETL (Extract, Transform, Load) process was meticulously executed to ensure data consistency, accuracy, and readiness for analysis.

In the Qlik Sense environment, I developed comprehensive data models, ensuring robust relationships between datasets and optimizing performance for fast querying and real-time analysis. I implemented data security measures to protect sensitive information, restricting access to authorized users only.

I created interactive dashboards featuring a variety of visualizations—pie charts, bar charts, treemaps, and histograms—that provide clear insights into key metrics such as sales distribution, market share, and delivery performance. These visualizations allow users to apply data filters for detailed, granular analysis and better decision-making.

I established key performance indicators (KPIs) to monitor critical metrics like total sales, profit, customer numbers, and delivery efficiency. Automated reports and alerts keep stakeholders informed of important changes and trends.

Performance testing confirmed the system's ability to handle large datasets efficiently, ensuring responsive data filters and optimized calculation fields. This project has successfully delivered a powerful analytical toolset that enhances DataCo's supply chain operations, drives sales growth, and boosts overall profitability.

Overall , this report,utilizing Qlik Sense and Smartbridge's advanced data-driven capabilities, focuses on enhancing DataCo's supply chain management. By integrating diverse data sources, implementing robust ETL processes, and employing secure and efficient data modeling, we aim to provide comprehensive insights into supply chain operations. The interactive dashboards and detailed visualizations, including pie charts, bar charts, treemaps, and histograms, facilitate real-time analysis and KPI monitoring, allowing for informed decision-making. User access is meticulously managed to ensure data security, while scheduled reporting keeps stakeholders updated on key metrics. Performance testing ensures the system handles large datasets efficiently, with responsive data filters and optimized calculation fields. This holistic approach aims to improve delivery performance, optimize operations, and drive sales growth, ultimately boosting DataCo's efficiency and profitability.

Performance testing:-

Performance testing in supply chain management systems encompasses several key aspects to ensure optimal system responsiveness and efficiency:

1. Amount of Data Rendered to DB:

This metric evaluates the volume of data loaded into the database and its impact on query execution time and system resources. By analyzing the amount of data fetched from the database during operations such as data loading, filtering, and visualization rendering, stakeholders can assess the scalability and performance of the system. Optimizing data loading processes, implementing efficient data retrieval techniques such as indexing, and partitioning large datasets can help minimize latency and enhance system responsiveness.

1. Prepare your data in a structured format suitable for analysis.
2. Log in to Qlik Sense using your credentials.
3. Navigate to the data loading interface within Qlik Sense.
4. Choose the appropriate data source from which to upload data.
5. Select the data tables or files you want to upload into Qlik.
6. Specify any data transformations or mappings required during the loading process.
7. Perform data modeling tasks such as creating associations between tables and defining key fields.
8. Save your app in Qlik Sense and ensure it is published for authorized users.
9. Test the app to ensure data accuracy and functionality.
10. Set up scheduled data refreshes if your data is regularly updated.
11. Collaborate with stakeholders and share insights derived from the data using Qlik's collaboration features.

2. Utilization of Data Filters:

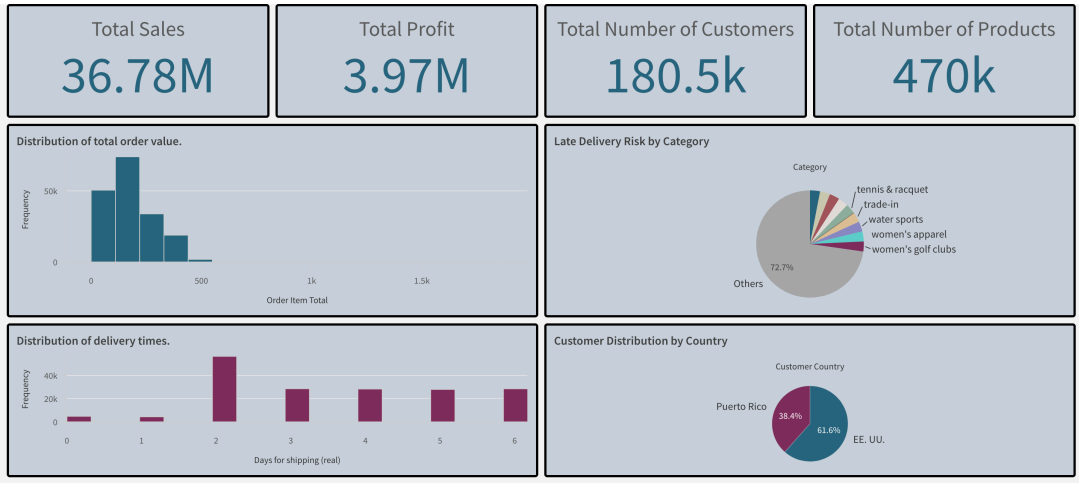
Data filters play a crucial role in refining search results and reducing data retrieval times by selecting relevant data subsets based on user-defined criteria. Performance testing assesses the efficiency of data filters in streamlining data access and improving user experience. By analyzing the utilization of data filters across various system functionalities, stakeholders can identify bottlenecks and optimize filter configurations to enhance query performance and responsiveness.

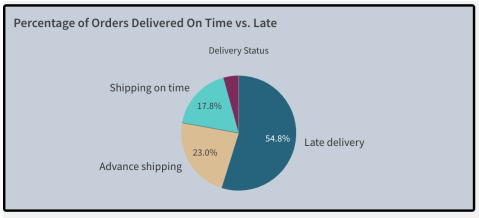
1. Open your Qlik Sense app and navigate to the desired sheet or dashboard.
2. Identify the fields available for filtering.
3. Access the filter panel and select the field you want to filter.
4. Choose the specific values or range to filter by.

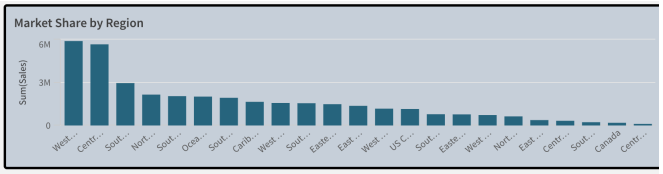
5. Analyze the data with applied filters to understand the subset of information.
6. Combine multiple filters to refine your analysis further.
7. Clear filters when necessary to view the entire dataset.

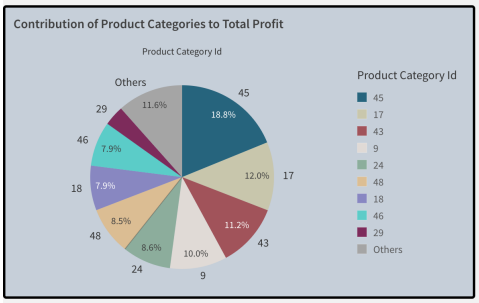
3. Number of Visualizations/Graphs:

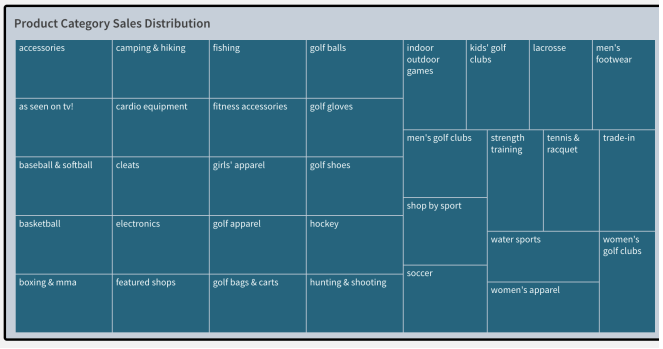
Visualizations and graphs play a crucial role in presenting data insights and facilitating decision-making in supply chain management systems. However, an excessive number or complexity of visualizations can impact dashboard loading times and user interaction. Performance testing evaluates the performance impact of the number and complexity of visualizations on system responsiveness. By analyzing dashboard loading times, rendering performance, and resource utilization, stakeholders can optimize visualization design and layout to maintain responsiveness and usability without compromising data presentation quality.











Based on the visuals in Qlik Sense:

1. Treemap Visualization: Product Category Sales Distribution

- The treemap offers a hierarchical view of sales distribution across product categories.
- Larger rectangles indicate categories with higher sales contributions, providing insights into the most profitable product categories.
- Categories with smaller rectangles may represent areas for potential growth or optimization.

2. Contribution of Product Categories to Total Sales (Pie Chart)

- This pie chart illustrates the percentage contribution of each product category to total sales.
- It highlights the relative importance of each category in driving overall sales revenue.
- Identifies key product categories that significantly impact total sales, aiding in strategic decision-making and resource allocation.

3. Percentage of Orders Delivered On Time vs. Late (Pie Chart)

- The pie chart compares the percentage of orders delivered on time versus those delivered late.
- It provides insights into delivery performance, with a higher percentage of on-time deliveries indicating efficient operations.
- A higher percentage of late deliveries may signal areas for improvement in the supply chain or logistics processes.

4. Market Share by Region (Bar Chart)

- This bar chart illustrates the market share of sales revenue across different regions.
- Identifies regions contributing the most to total sales, guiding geographic expansion or targeted marketing efforts.
- Regional sales performance can be analyzed to identify growth opportunities or areas requiring additional focus.

5. Customer Distribution by Country (Pie Chart)

- The pie chart shows the distribution of customers based on their countries.
- It highlights geographic concentration of customers and potential opportunities for market expansion or localization efforts.
- Variances in customer distribution across countries may indicate differing market dynamics or customer preferences.

6. Histograms of Delivery Time and Order Item Total

- The histograms provide insights into the distribution of delivery times and total order values, respectively.
- They help identify patterns and outliers in delivery performance and order values.
- Analysis of delivery time distributions can inform optimization strategies for logistics and processes.

Conclusion:-

The visuals in Qlik Sense provide comprehensive insights into various facets of supply chain management. The treemap and pie chart visuals elucidate the sales distribution across product categories, shedding light on the relative importance of each category in driving overall revenue.

These visualizations go beyond mere numbers and tables, transforming data into clear and actionable visuals. Treemaps and pie charts, for instance, shed light on product category sales distribution. By seeing the relative importance of each category, stakeholders can identify the key drivers of overall revenue and tailor strategies to maximize their contribution.

Delivery performance is another crucial aspect of supply chain management. Qlik Sense utilizes pie charts to offer a clear assessment in this area. By comparing on-time deliveries against late deliveries, businesses can pinpoint areas for improvement in logistics and fulfillment processes. This not only optimizes operations but also ensures customer satisfaction, a critical factor for success.

Understanding where your customers are located is vital for targeted marketing and reaching new markets. Qlik Sense empowers stakeholders with this knowledge through bar charts detailing market share by region. These charts reveal the key regions contributing significantly to sales revenue, allowing for the development of targeted marketing strategies and informed decisions about geographic expansion.

Data visualization in Qlik Sense extends beyond sales and delivery. Pie charts depicting customer distribution by country provide valuable insights into geographic customer concentration. This information plays a vital role in market segmentation and localization efforts. By understanding customer demographics, businesses can ensure their products and services resonate with specific segments, fostering stronger customer relationships.

Finally, histograms play a crucial role in optimizing logistics and fulfillment processes. Qlik Sense utilizes these visualizations to depict distribution patterns and outliers in delivery time and order item totals. By analyzing these patterns, stakeholders can develop optimization strategies, leading to improved logistics efficiency and streamlined fulfillment processes.

In conclusion, Qlik Sense data visualization empowers stakeholders within the supply chain with a holistic view of their operations. By transforming complex data into clear and actionable insights, these visualizations enable informed decision-making, streamlined operations, and ultimately, a high-performing supply chain.

References:-

1)https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis/data

2)https://medium.com/@dsyafz/supply-chain-analysis-dataco-global-259f404078ae

3) https://jovian.com/fajarrafli2/ecommerce-supply-chain-data-analysis

4) https://www.tutorialspoint.com/qlikview/index.htm