

Business Analytics with Qlik

- K V BHARATH KUMAR

Data-Driven Innovations In Supply Chain

Management With Qlik Insights.

1. Define Problem / Problem Understanding:-

Supply chain management is a critical aspect of modern businesses, involving the coordination of various processes and activities to ensure the efficient flow of goods and services from suppliers to customers. However, traditional supply chain management approaches often rely on manual processes, historical data, and intuition, leading to inefficiencies, delays, and suboptimal decision-making. The business problem addressed in this project is the need for data-driven innovations that can streamline supply chain operations, enhance decision-making, and improve overall supply chain performance.

1.1. Specify The Business Problem:-

This project aims to revolutionise supply chain management through data-driven insights using Qlik. Leveraging advanced analytics, it seeks to optimize logistics, forecasting, and inventory management, enhancing operational efficiency and responsiveness.

This transformative project endeavours to reshape the landscape of supply chain management by harnessing the power of Qlik's data-driven insights. Employing cutting-edge analytics, it strives to revolutionise key facets such as logistics, forecasting, and inventory management, with the overarching goal of elevating operational efficiency and responsiveness to new heights.

1.2. Business Requirements:-

1. Data Integration Strategy: Develop a robust data integration framework to aggregate and centralize data from diverse supply chain sources, ensuring seamless data flow and consistency across the system.

2. Visualization Capabilities: Utilize Qlik's advanced visualization tools to create intuitive,

dynamic dashboards that provide stakeholders with clear and actionable insights into the entire supply chain ecosystem.

3. Advanced Analytics: Leverage Qlik's powerful analytics features to analyze historical logistics data, uncover patterns, and optimize transportation routes for enhanced efficiency and reduced costs.

4. Real-Time Tracking and Monitoring: Implement real-time tracking and monitoring solutions to enhance visibility into the movement of goods. This will help reduce lead times and minimize transportation costs by enabling timely interventions and adjustments.

5. Proactive Decision-Making: Utilize real-time analytics to facilitate quick and informed decision-making in response to unforeseen events or changes in demand, ensuring a proactive and responsive supply chain that can adapt to dynamic market conditions.

1.3.Literature Survey:-

A literature survey on the project theme of revolutionizing supply chain management through data-driven insights and advanced analytics reveals a growing body of research and scholarly articles focused on similar endeavours. Studies underscore the increasing recognition of the pivotal role that data analytics plays in transforming traditional supply chain processes. Research highlights the effectiveness of leveraging advanced analytics tools, such as Qlik, to enhance visibility and decision-making in supply chain operations. The study emphasises the positive impact on logistics optimization, forecasting accuracy, and inventory management efficiency.

Moreover, it delves into the broader landscape of data-driven supply chain transformations, exploring diverse analytical techniques and technologies. The findings showcase successful implementations, demonstrating notable improvements in operational efficiency and responsiveness across various industry sectors. In addition, it examines the challenges and opportunities associated with the adoption of data-driven insights in supply chain contexts. The literature emphasises the need for organizations to develop robust data governance frameworks and cultivate a data-driven culture to unlock the potential benefits fully.

1.4.Social Or Business Impact:-

1.4.1.Social Impact Analysis:-

1.Environmental Sustainability:

Optimized Routes: Reduces fuel consumption and emissions.

Reduced Waste: Improves inventory management to minimize overproduction.

2.Job Creation and Skill Development:

Employment Opportunities: Implementing advanced supply chain solutions can create jobs in data analysis, IT support, and supply chain management.

Training: Enhances employee skills and career prospects.

3 Enhanced Product Availability:

Better Delivery: Ensures timely availability of essential goods.

4.Economic Stability:

Supporting Local Businesses: Efficient supply chains can support local suppliers and manufacturers

Analysis:-

1.Cost Reduction:

Operational Efficiency: Lowers transportation and inventory costs.

Reduced Lead Times: Improves supply chain efficiency.

2.Increased Revenue:Customer Satisfaction: Enhances service levels and repeat business.

Market Responsiveness: Quickly adapts to market changes.

3.Risk Mitigation:

Proactive Management: Identifies and addresses potential disruptions.

Data-Driven Decisions: Reduces errors and enhances planning.

4.Competitive Advantage:

Innovative Solutions: Differentiates from competitors.

Agility: Quickly adapts to changes in demand.

5.Operational Visibility:

Comprehensive Insights: Provides a holistic view of the supply chain.

Performance Monitoring: Identifies areas for continuous improvement

2.Data Collection & Extraction From Database:-

Collecting data means gathering information that is needed for analysis. This involves measuring and recording details about the topics or variables of interest in a systematic way. Proper data

collection allows researchers to answer their questions, test ideas, evaluate results, and discover new insights from the data.

2.1.Downloading The Dataset:-

The dataset for this project was downloaded from Kaggle (kaggle.com), a website that provides many publicly available datasets. Specifically, the "DataCo Smart Supply Chain for Big Data Analysis" dataset, available at

"<https://www.kaggle.com/datasets/shashwatwork/dataco-smartsupply-chain-for-big-data-analysis/data>, was used".

In simple terms, data was first collected from various sources and compiled into a dataset. This dataset was then made publicly available on Kaggle at the link provided. For this project, the relevant dataset was downloaded from the specified link on Kaggle so it could be extracted and analysed to understand supply chain processes better.

The specific dataset used is the "DataCo Smart Supply Chain for Big Data Analysis" which can be accessed from Kaggle.com.

Data Extraction:-

The dataset was downloaded as a CSV file named 'DescriptionDataCoSupplyChain.csv'. The following steps were followed to extract and prepare the data for analysis:

1. *Download the Dataset: The dataset was downloaded from the Kaggle website.*
2. *Load the Data: The CSV file was loaded into a data analysis tool (such as Qlik) for further processing.*
3. *Initial Inspection: The dataset was inspected to understand its structure and content. This included checking the number of rows and columns, and looking at a few sample records to get a sense of the data*

2.2.Understanding The Dataset:-

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

1. *Type: Type Count*
2. *Days for shipping (real): Product shipment days*

3. Days for shipment (scheduled): product getting prepared for shipment
4. Benefit per item: profit earned per product
5. Sales per customer: No of products purchased by the customer
6. Delivery: Products delivery date.
7. Late_delivery_risk: percentage of late delivery risk
8. Category Id: product category ID
9. Category: product category
10. Customer City: Customer purchase city
11. Customer Country: Customer purchase country
12. Customer Email: Customer purchase Email
13. Customer Fname: Customer First name
14. Customer ID: Customer order ID
15. Customer Lname: Customer's last name
16. Customer Segment: Types of Customer\
17. Customer State: Customer order state
18. Customer Street: Customer address
19. Customer Zip code: Customer area code
20. Market: top 10 country Market
21. Order City: Customer purchase city
22. Order Country: Customer purchase country
23. Order Customer ID: Customer
24. order date (Date Orders): Customer order date
25. Order Item Product Price: product price
26. Order Item Profit Ratio: profit ratio
27. Order Item Quantity: No of orders placed
28. Sales: total no of sales
29. Order Item Total: total price of the order placed
30. Order Profit Per: product
31. Order Region: order placed region
32. Order State: order placed State
33. Order Status: order delivery status
34. Order Zip code: customer area code
35. Product Card ID: product number
36. Product Category Id: a product whose category belongs to

37. *Product: what product*
38. *Product Image: image of the product*
39. *Product Price: Price of the product.*

Data Quality:-

To ensure the reliability of the analysis, the data was assessed for quality. This involved:

40. *Checking for Missing Values: Identifying and handling any missing or null values in the dataset.*
41. *Ensuring Consistency: Verifying that data entries are consistent, for instance, ensuring that dates follow a uniform format.*
42. *Data Types: Confirming that each field has the appropriate data type (e.g., dates are in date format, numerical values are not stored as text).*

Key Insights from Data Exploration:-

From the initial data exploration, several key insights were identified:

43. *Sales and Profit Trends: Patterns in sales and profit margins across different product categories and regions.*
44. *Customer Segmentation: Understanding customer distribution across different segments and regions.*
45. *Order Patterns: Trends in order quantities and frequencies over time.*
46. *Shipping Performance: Analysis of shipping times and identification of any delays or bottlenecks.*

3. Data Preparation:-

Before data can be effectively visualised, it often needs to be preprocessed and transformed into a suitable format. This step, known as data preparation or data wrangling, is crucial for ensuring that the visualisations are accurate, insightful, and easy to interpret.

The process of preparing data for visualisation typically involves several steps:

1. **Data cleaning:** Removing or handling missing values, duplicates, errors, and inconsistencies in the data.
2. **Data transformation:** Converting data into a format that is suitable for visualisation, such as reshaping data from wide to long format, or vice versa.
3. **Data aggregation:** Grouping or summarising data based on specific variables or categories to create meaningful aggregates for visualisation.
4. **Feature engineering:** Creating new features or variables from existing ones that may

provide better insights or enable more effective visualisations.

5. **Data sub setting:** Selecting relevant subsets of the data based on specific criteria or filters to focus the visualisations on areas of interest.
6. **Data formatting:** Ensuring that data types (e.g., numerical, categorical, date/time) are correctly formatted for the chosen visualisation techniques.

By preparing the data properly, visualisations can accurately represent the underlying data, reveal patterns and trends more effectively, and support better decision-making based on the insights derived from the visualisations.

4.Data Visualization :-

Data visualisation is the process of creating graphical representations of data to help make complex datasets more accessible, intuitive, and easier to interpret. It involves transforming numbers, text, and raw data into visuals like charts, graphs, maps, and other visual elements. The goal is to leverage these visuals to represent information in a way that allows people to identify patterns quickly, trends, outliers, and key insights within the data. Effective data visualisations enable stakeholders to grasp crucial information at a glance, something that is often challenging with raw data alone. By presenting data visually, it becomes more comprehensible, even for those without extensive statistical or analytical expertise. Clear and compelling data visualisations act as a universal language for communicating data-driven insights to diverse audiences, enhancing communication and supporting better decisionmaking based on the insights derived from the visuals. Ultimately, data visualisation plays a crucial role in making complex datasets more accessible, understandable, and actionable, driving informed decision-making processes.

4.1.No Of Unique Visualisations:-

The number of unique visualisations that can be created with a given dataset. Some common types of visualisations that can be used to analyse the performance and efficiency include bar charts, line charts, heat maps, scatter plots, pie charts, Maps,KPI etc. These visualisations can be used to compare performance, track changes over time, show distribution, and relationships between variables, breakdown of revenue and customer demographics, workload, resource allocation.

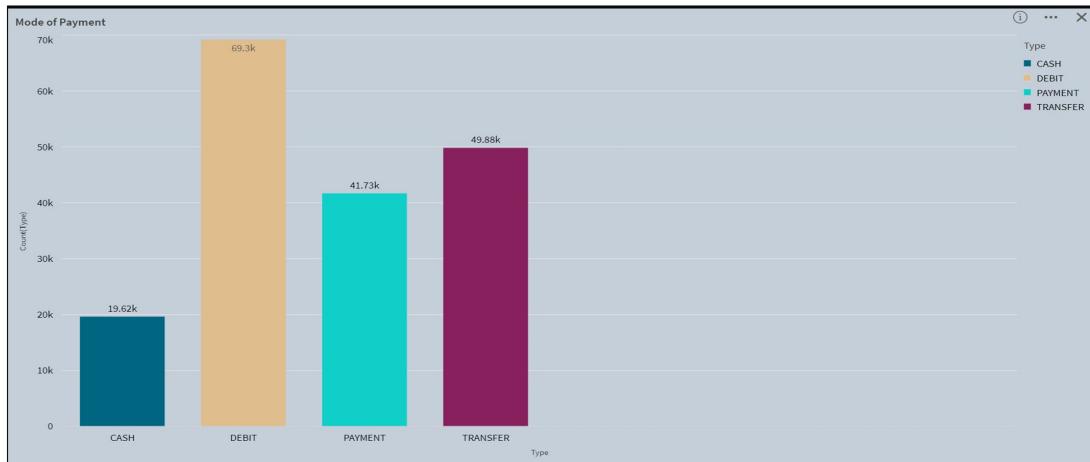
4.2.Visualisatons:-

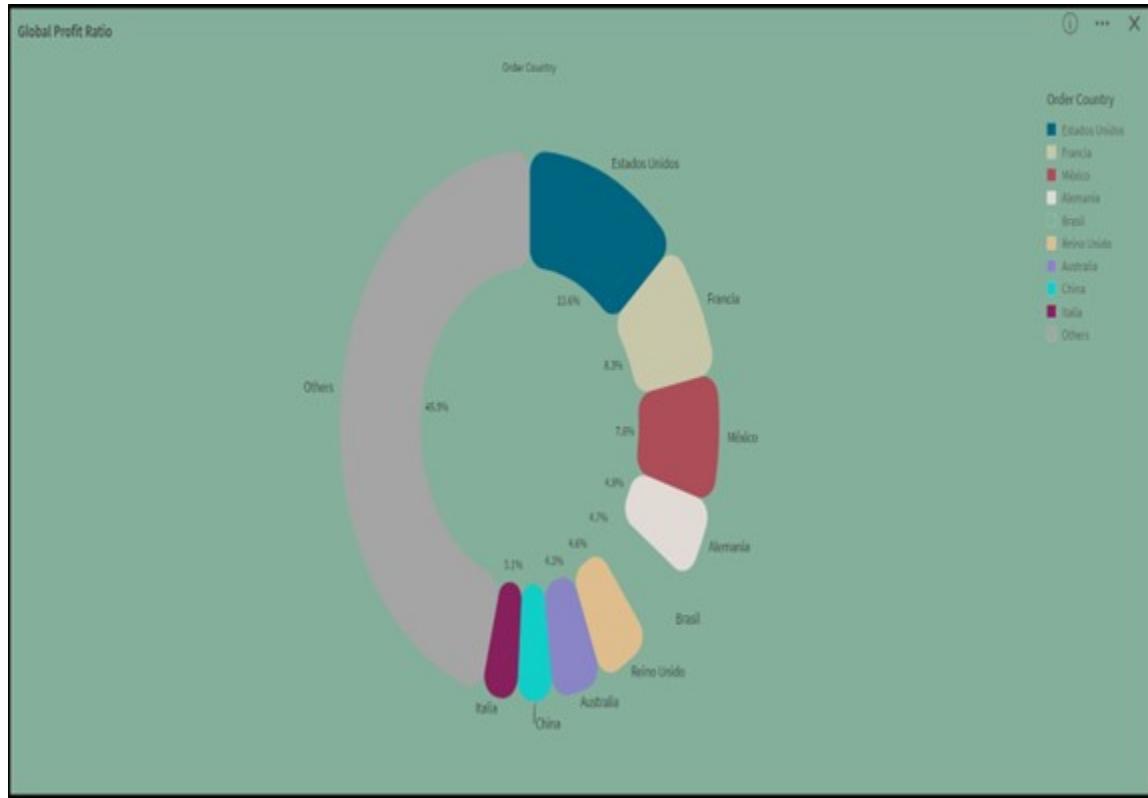
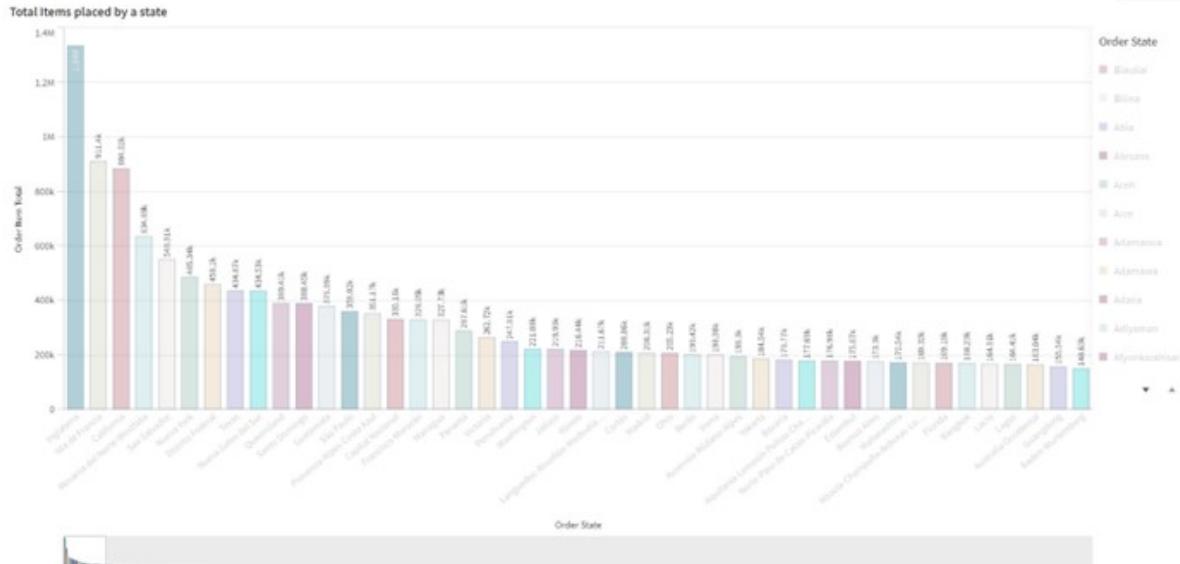
Below are the following unique visualisations that can be observed in this project for the chosen data set.

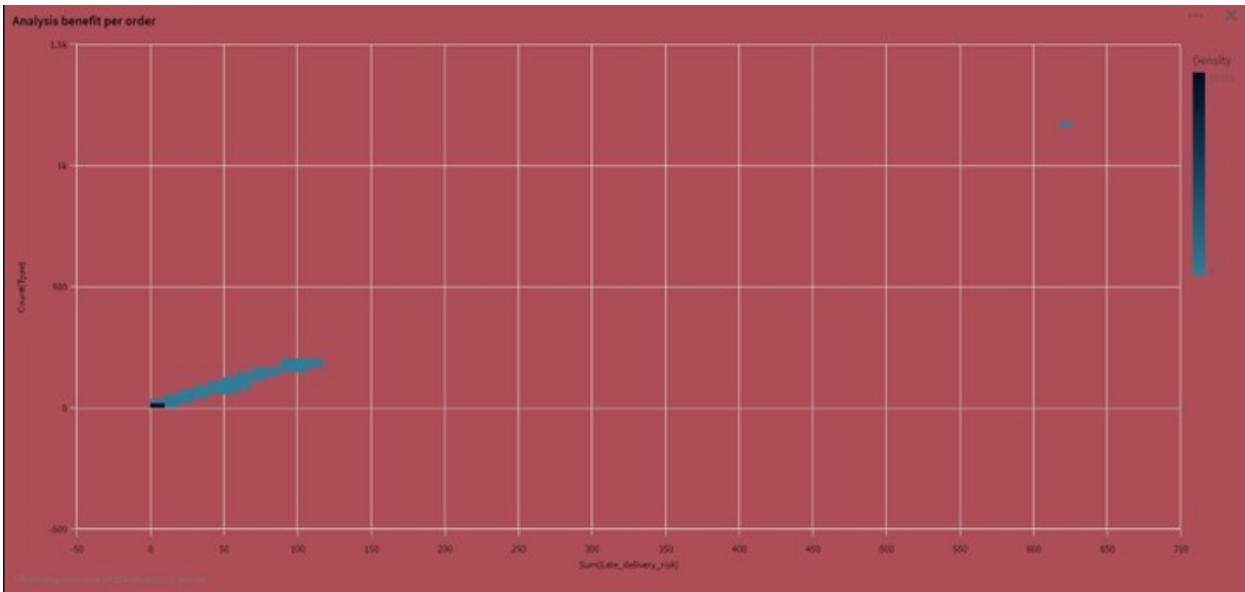
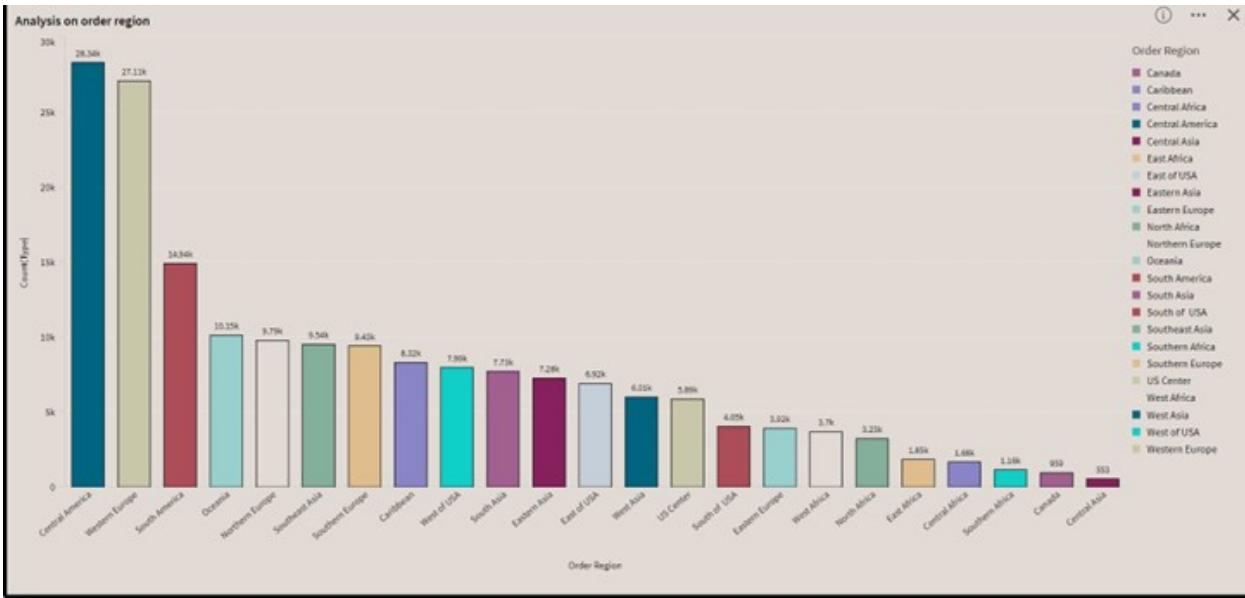
Sales
36.78M

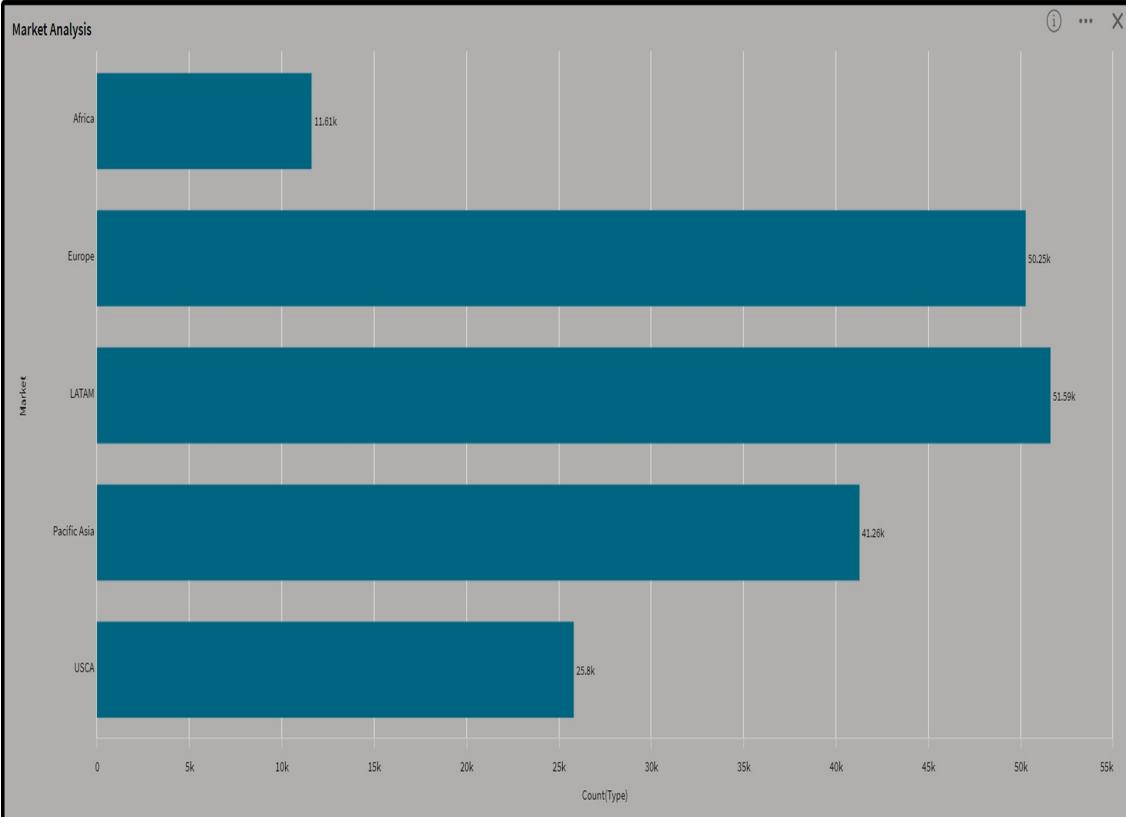
Profit Margin
3.97M

Delivery Risk
98.98k

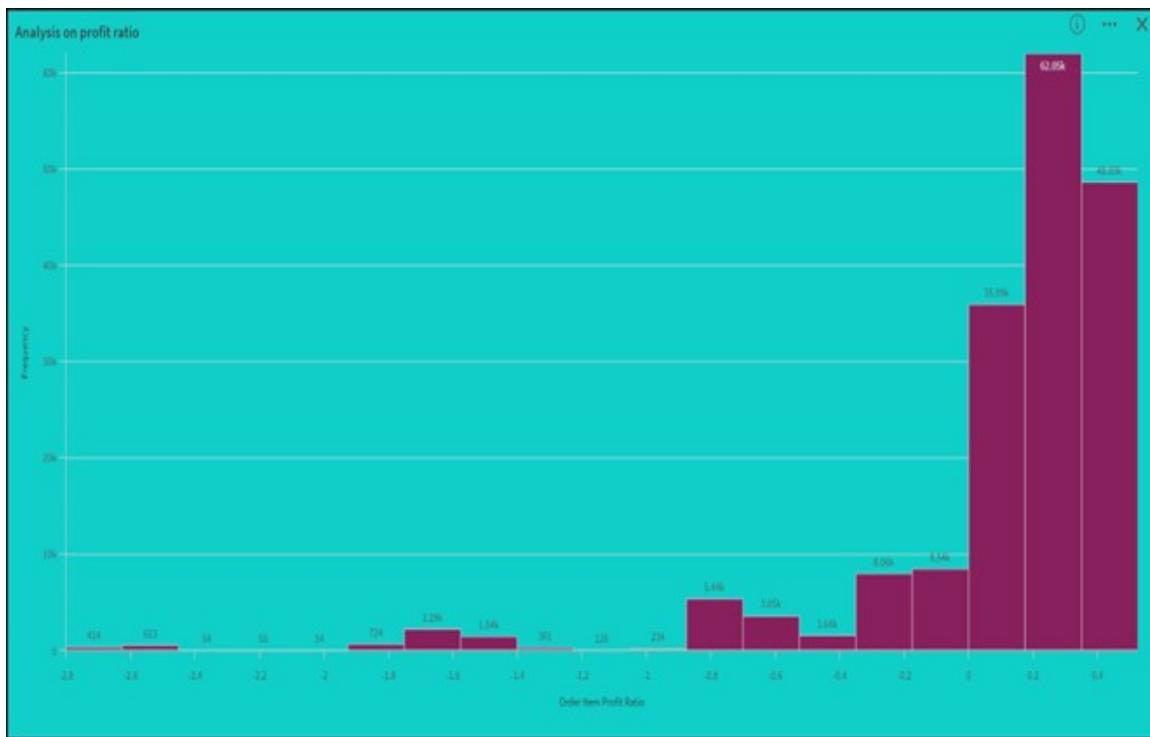


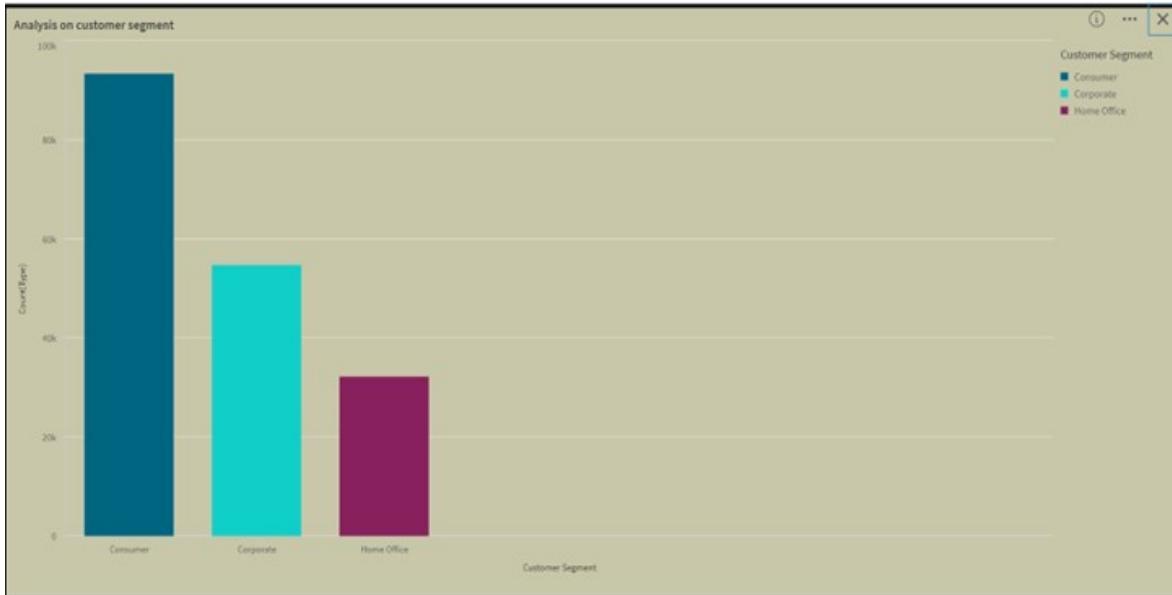






Customer Id	Customer City
	i ↗ ...
order date (DateOrders)	
Order Region	
Product Category Id	
Product Price	





5.Dashboard:-

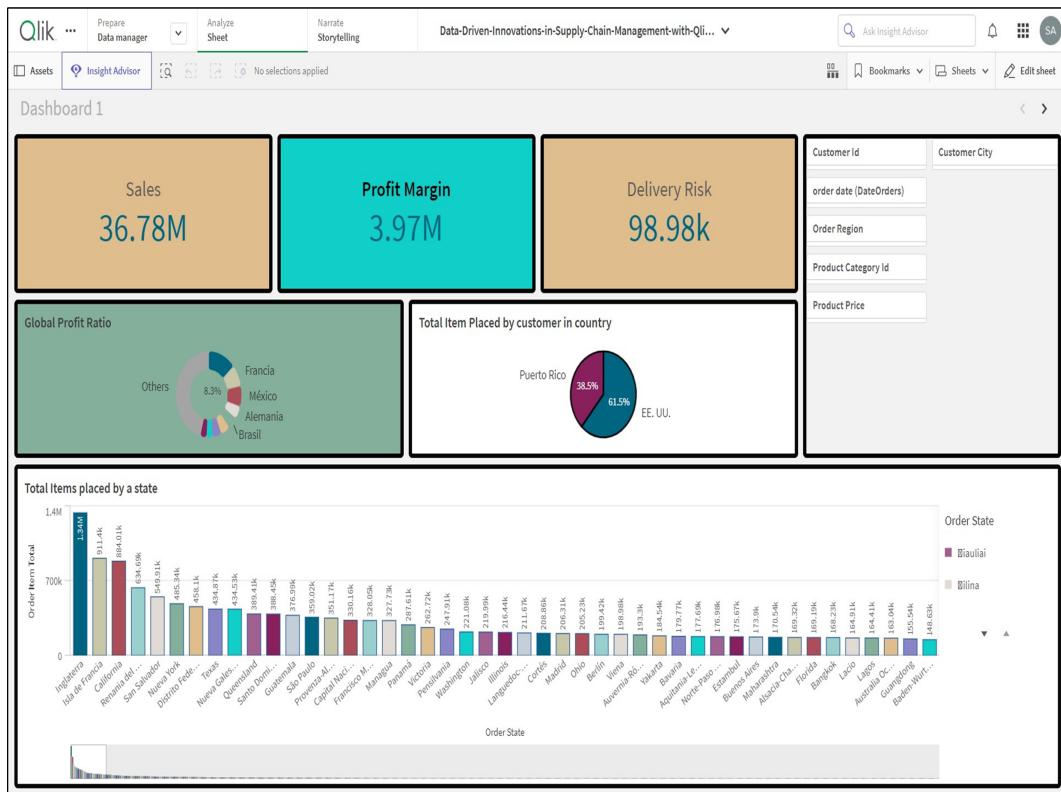
A dashboard is a graphical user interface that presents critical information and data in an organized, visually appealing, and easy-to-comprehend format. It serves as a centralised hub or control panel that consolidates and displays multiple data sources, key performance indicators(KPIs), and other essential metrics through interactive charts, graphs, tables, and other data visualisation elements.

Dashboards leverage real-time monitoring capabilities, providing up-to-date visibility into data as it changes, allowing users to track performance and monitor metrics in real-time. They are specifically designed for targeted purposes or use cases, tailoring the displayed information to the needs of different users, teams, or industries. Dashboards are widely used across various sectors,

such as business, finance, manufacturing, healthcare, and more, as they enable stakeholders to grasp critical insights quickly, identify patterns and trends, and make informed, data-driven decisions by presenting complex data in a visually appealing and digestible format.

5.1. Responsive And Design Of Dashboard:-

Here are some of the following project Dashboard looks like :



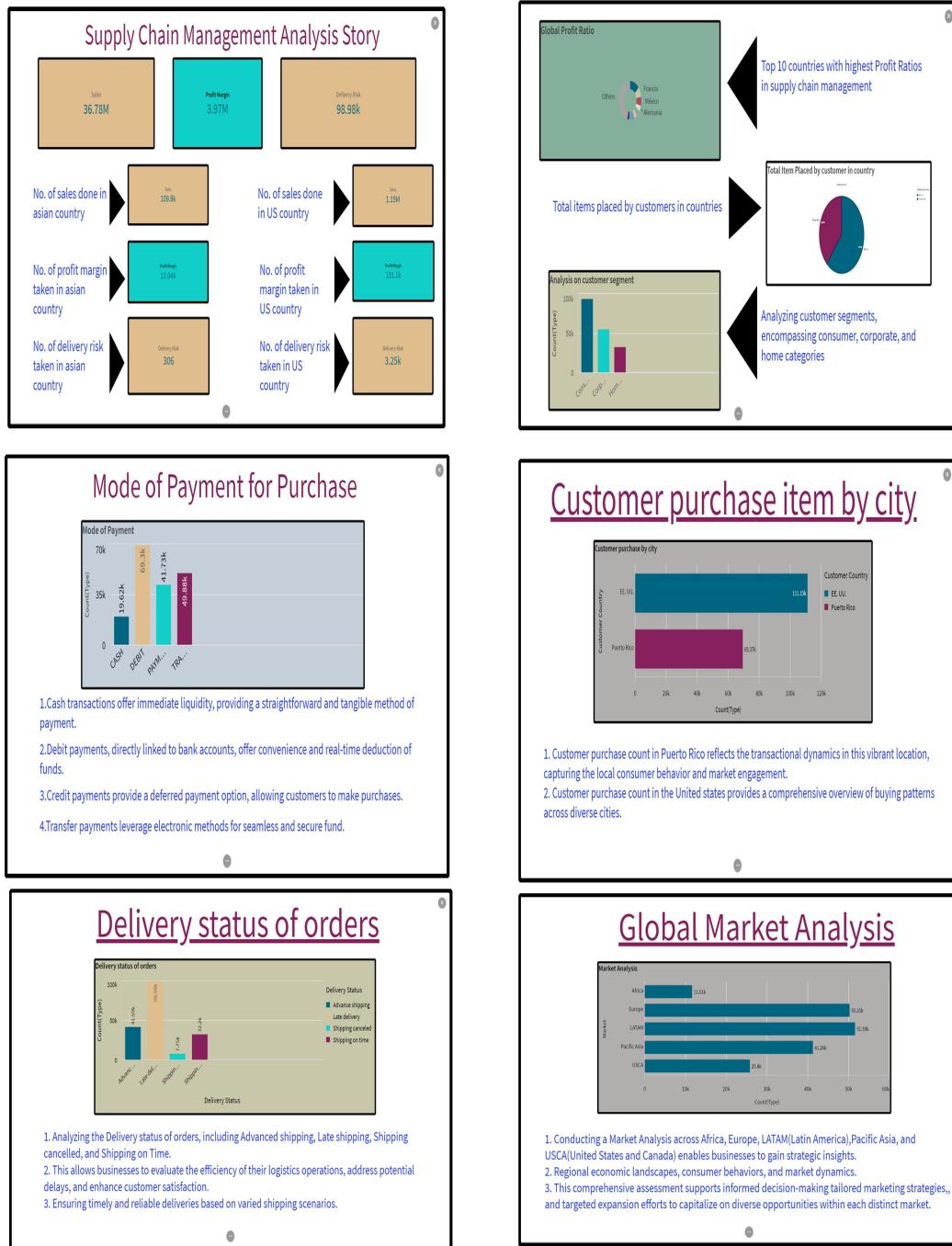


6.Story:-

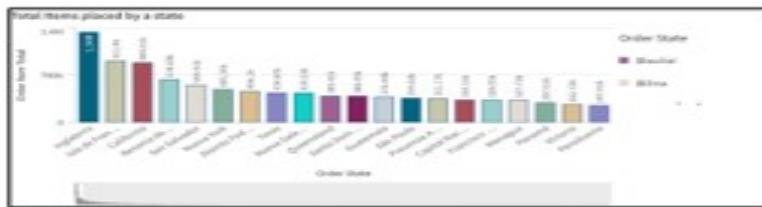
A data story is a narrative approach to presenting data and analysis in a way that is engaging, logical, and easy to comprehend. It involves structuring the information as a story, with a clear introduction that provides context and sets the stage, a body that systematically presents the data, analysis, and key findings through visualisations and explanations, and a conclusion that summarises the main insights and highlights their implications. Data stories aim to transform complex data into a cohesive and compelling narrative, guiding the audience through the data in a storytelling format. This approach leverages the power of storytelling to make data more relatable, memorable, and impactful. Data stories can be conveyed through various mediums, such as reports, presentations, interactive dashboards, or videos, making the information more accessible and understandable to a diverse audience.

6.1.Design of story / Storytelling:-

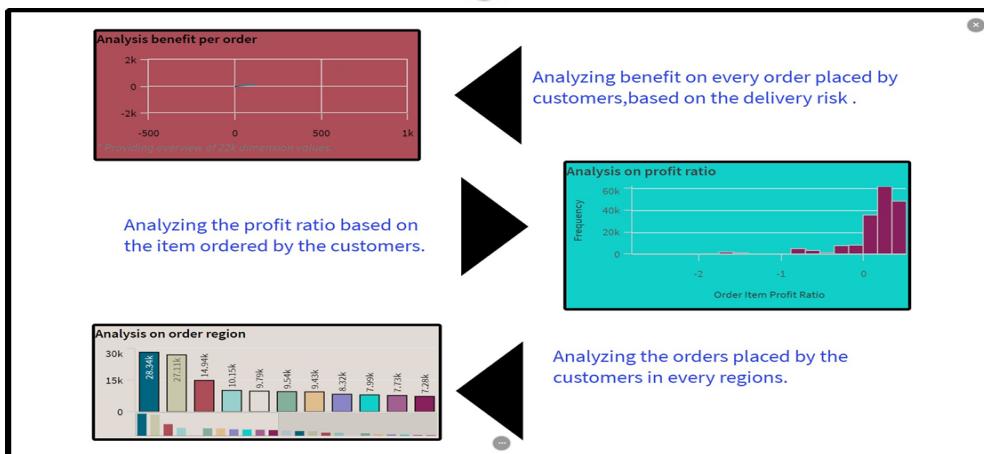
The following figures explain the story of the graphs:-



Total Items placed by a state



- Comparing the Total ordered Items placed by customers in a state and placing them in an order.



7.1.Amount of data loaded:-

Performance testing of the "Amount of Data Loaded" involves evaluating how efficiently and effectively a system, software application, or database handles large volumes of data during the loading process. This testing helps ensure that the system can manage the expected data volumes within acceptable performance criteria. Here are the key aspects and steps involved in performance testing for data loading:

1. Define Objectives and Metrics.
 2. Prepare the Environment like Test Environment Setup and Data Preparation.
 3. Plan the Test Scenarios and Execute Tests by Load Testing, Stress Testing and

Volume Testing.

4. Monitor and Measure the resource Utilization, Load Time, Throughput.

5. Reporting the Detailed Report with Visualizations.

The screenshot shows the Qlik Sense interface. On the left, there is a 'Search' bar and a 'Filter by table' dropdown set to 'All tables'. Below this is a list of various data fields and tables: Benefit per order, Category, Category Id, Category Name, Customer City, Customer Country, Customer Email, Customer Fname, Customer Id, Customer Lname, Customer Password, Customer Segment, Customer State, Customer Street, Customer Zipcode, Date, Days for shipment (r...), Days for shipping (r...), Delivery Status, Department, Department Id, Department Name, Hour, IP, Late_delivery_risk, Latitude, Longitude, Longitude_Latitude, Market, Month, Order City, Order Country, Order Customer Id, order date (Dat...), Order Id, Order Item Cardpro..., Order Item Discount, Order Item Discoun..., Order Item Id, Order Item Product..., Order Item Profit Ra..., Order Item Quantity, Order Item Total, Order Profit Per Order, Order Region, Order State, Order Status, Order Zipcode, Product, Product Card Id, Product Category Id, Product Description, Product Image, Product Name, Product Price, Product Status, Sales, Sales per customer, and url. In the center, there are two filter panes. The top pane contains 'shipping date (...)' with a calendar icon, 'Shipping Mode' with a gear icon, 'Type' with a grid icon, and 'url' with a magnifying glass icon. The bottom pane has a '+ Add data' button and a 'Q' icon. The overall layout is clean and organized, typical of a business intelligence tool like Qlik Sense.

7.2.Utilization Of Data Filters:-

Utilization of Data Filters in Qlik Analysis refers to the practice of applying criteria or conditions to restrict datasets within Qlik applications. These filters help users focus on specific subsets of data, enabling more targeted and meaningful analysis. By selecting relevant dimensions, values, or expressions, users can refine their data views, extract insights, and make informed decisions based on the filtered dataset. Using data filters in Qlik enhances data analysis by allowing users to narrow down datasets, focus on specific information, and gain more precise insights. Data filters in Qlik help users narrow down datasets, focus on specific information, and gain precise insights. Here's a concise guide/an overview of how to effectively utilize data filters in Qlik analysis:

1.Understanding Filters: Criteria applied to restrict data, removing noise for focused analysis.

2.Types of Filters: Selection, List Box, Advanced Search, and Set Analysis.

3.Applying Filters: Interactively clicking data points, using list boxes, date range filters, and custom set analysis.

4.Best Practices: Start broad and narrow down progressively. Use clear naming conventions. Combine filters wisely.

5.Examples: Sales analysis, customer segmentation, inventory management.

6.Benefits: Enhanced focus, improved performance, customized insights, interactive exploration.

7.Implementation: Design filter panes, use master items, document filter logic.

Utilizing data filters effectively enhances analysis, improves focus, and tailors insights to specific business needs in Qlik.

