***Data Driven Innovation in Supply Chain Management with QLIK***

*by Shaikh Minhazuddin*

Introduction

1. Defining the Problem

In the ever-evolving landscape of global commerce, supply chain management plays a crucial role in ensuring the efficient flow of goods and services. Organizations are constantly seeking innovative ways to optimize their supply chain operations to stay competitive. This project, titled "Data Driven Innovation in Supply Chain Management with QLIK Insights," aims to leverage advanced data analytics to address key challenges within the supply chain and drive efficiency through actionable insights.

1. Problem Understanding

Specifying Business Problem:

Supply chains are inherently complex, involving multiple stakeholders, numerous processes, and vast amounts of data. Common issues faced in supply chain management include:

* Inventory Management: Balancing inventory levels to meet demand without overstocking or understocking.
* Order Fulfilment: Ensuring timely and accurate order processing to maintain customer satisfaction.
* Supplier Performance: Monitoring and improving supplier reliability and delivery times.
* Demand Forecasting: Accurately predicting future demand to optimize stock levels and reduce waste.

These challenges can lead to increased costs, reduced efficiency, and customer dissatisfaction if not addressed effectively.

1. Business Requirements

To tackle these problems, businesses need:

* Accurate Data Analysis: Tools and methodologies to analyse vast amounts of supply chain data accurately.
* Real-time Insights: Real-time visibility into supply chain operations to make informed decisions quickly.
* Predictive Analytics: Advanced analytics to forecast demand and manage inventory levels proactively.
* Performance Monitoring: Systems to track and improve supplier performance continuously.

1. Social and Business Impact

Effective supply chain management has far-reaching impacts:

* Business Impact: Improved supply chain efficiency leads to cost savings, better resource utilization, and enhanced customer satisfaction. This, in turn, drives profitability and competitive advantage.
* Social Impact: Efficient supply chains contribute to sustainability by reducing waste, optimizing resource use, and lowering carbon footprints. They also ensure the timely availability of essential goods, improving overall societal well-being.

Data Collection and Extraction

Data Source

The dataset for this project was sourced from Kaggle, a popular platform for data science and machine learning datasets. 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 Python or 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.

Data Fields

The dataset includes the following key fields:

* Type : Type of transaction made
* Days for shipping (real): Actual shipping days of the purchased product
* Days for shipment (scheduled): Days of scheduled delivery of the purchased product
* Benefit per order: Earnings per order placed
* Sales per customer: Total sales per customer made per customer
* Delivery Status: Delivery status of orders: Advance shipping , Late delivery , Shipping canceled , Shipping on tim...
* Late\_delivery\_risk: Categorical variable that indicates if sending is late (1), it is not late (0).
* Category Id: Product category code
* Category Name: Description of the product category
* Customer City: City where the customer made the purchase
* Customer Country: Country where the customer made the purchase
* Customer Email: Customer's email
* Customer Fname: Customer name
* Customer Id: Customer ID
* Customer Lname: Customer lastname
* Customer Password: Masked customer key
* Customer Segment: Types of Customers: Consumer , Corporate , Home Office
* Customer State: State to which the store where the purchase is registered belongs
* Customer Street: Street to which the store where the purchase is registered belongs
* Customer Zipcode: Customer Zipcode
* Department Id: Department code of store
* Department Name: Department name of store
* Latitude: Latitude corresponding to location of store
* Longitude: Longitude corresponding to location of store
* Market: Market to where the order is delivered : Africa , Europe , LATAM , Pacific Asia , USCA
* Order City: Destination city of the order
* Order Country: Destination country of the order
* Order Customer Id: Customer order code
* order date (DateOrders): Date on which the order is made
* Order Id: Order code
* Order Item Cardprod Id: Product code generated through the RFID reader
* Order Item Discount: Order item discount value
* Order Item Discount Rate: Order item discount percentage
* Order Item Id: Order item code
* Order Item Product Price: Price of products without discount
* Order Item Profit Ratio: Order Item Profit Ratio
* Order Item Quantity: Number of products per order
* Sales: Value in sales
* Order Item Total: Total amount per order
* Order Profit Per Order: Order Profit Per Order
* Order Region: Region of the world where the order is delivered : Southeast Asia ,South Asia ,Oceania ,Eastern ...
* Order State: State of the region where the order is delivered
* Order Status: Order Status : COMPLETE , PENDING , CLOSED , PENDING\_PAYMENT ,CANCELED , PROCESSING ,SUSPECTED\_FRAUD
* Product Card Id: Product code
* Product Category Id: Product category code
* Product Description: Product Description
* Product Image: Link of visit and purchase of the product
* Product Name: Product Name
* Product Price: Product Price
* Product Status: Status of the product stock :If it is 1 not available , 0 the product is available
* Shipping date (DateOrders): Exact date and time of shipment
* Shipping Mode: The following shipping modes are presented : Standard Class , First Class , Second Class , Same Day Delivery

Understanding the Data

Data Characteristics

The dataset consists of transactional data involving multiple variables related to products, orders, customers, and shipment details. Understanding the characteristics of this data is crucial for effective analysis and insights generation.

Data Quality

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

* Checking for Missing Values: Identifying and handling any missing or null values in the dataset.
* Ensuring Consistency: Verifying that data entries are consistent, for instance, ensuring that dates follow a uniform format.
* 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).

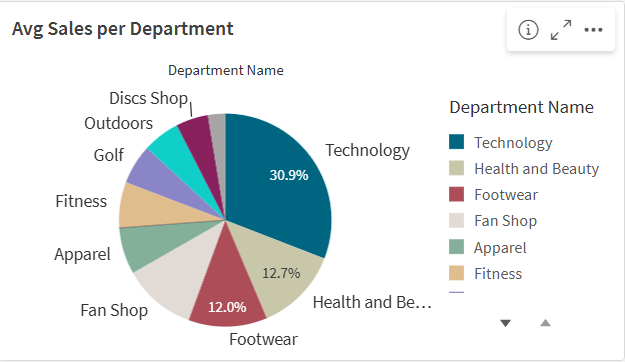
Key Insights from Data Exploration

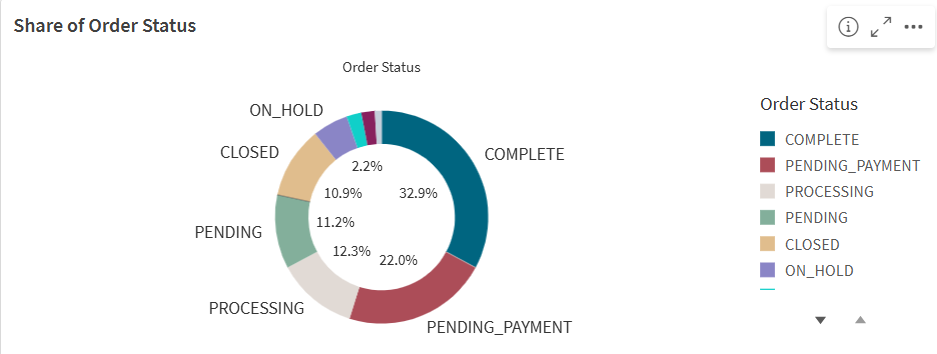
From the initial data exploration, several key insights were identified:

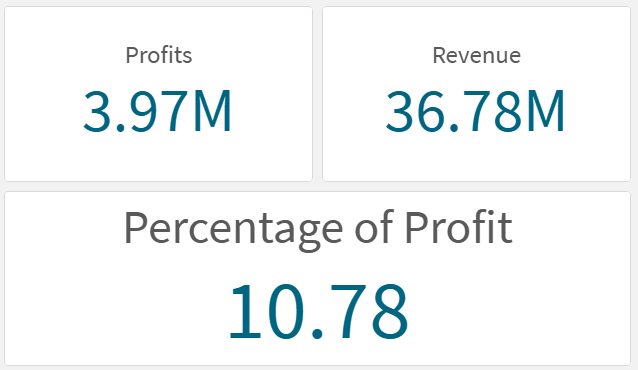
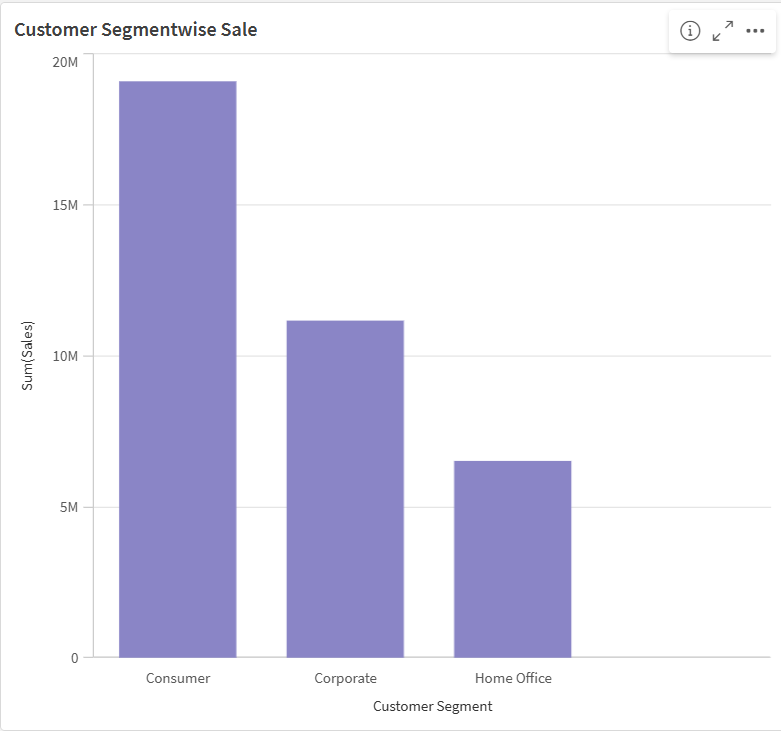
* Sales and Profit Trends: Patterns in sales and profit margins across different product categories and regions.
* Customer Segmentation: Understanding customer distribution across different segments and regions.
* Order Patterns: Trends in order quantities and frequencies over time.
* Shipping Performance: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

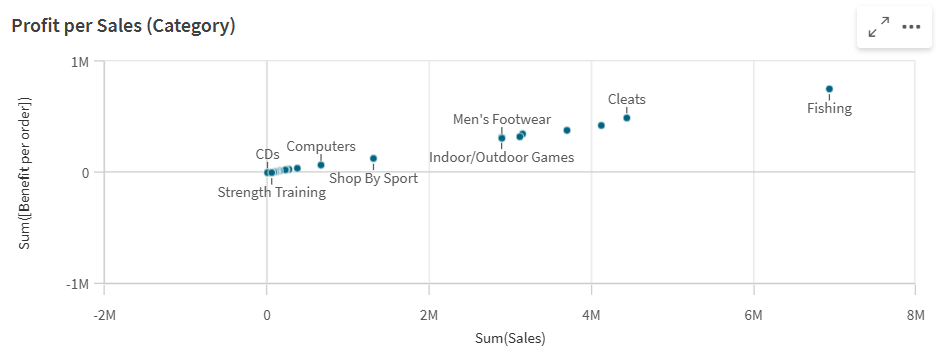
Data Visualization

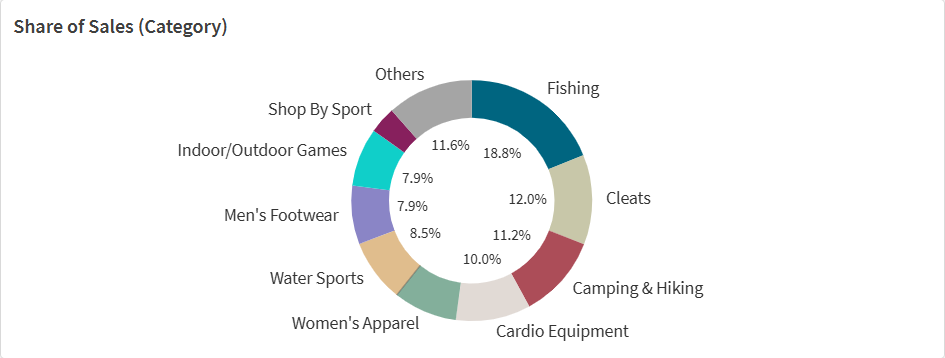
Below are the some of the visualization from the project:





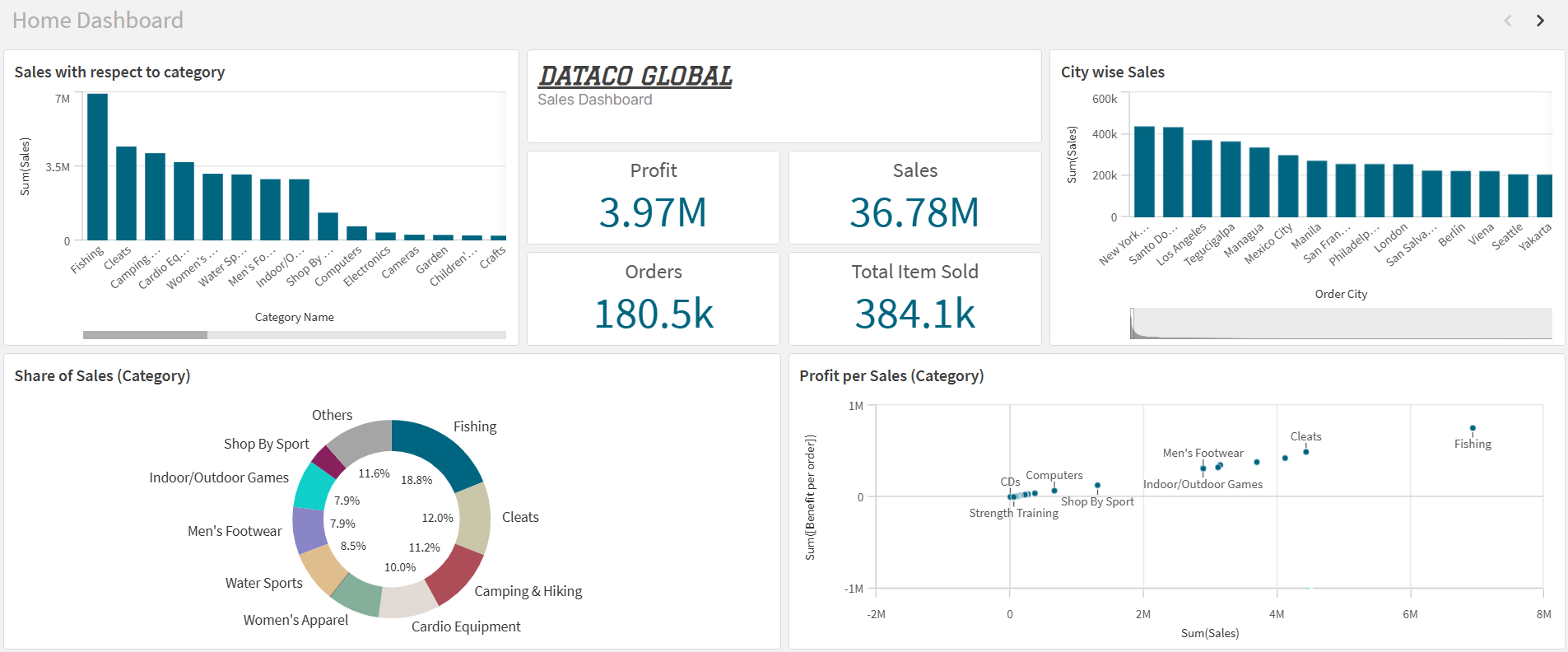
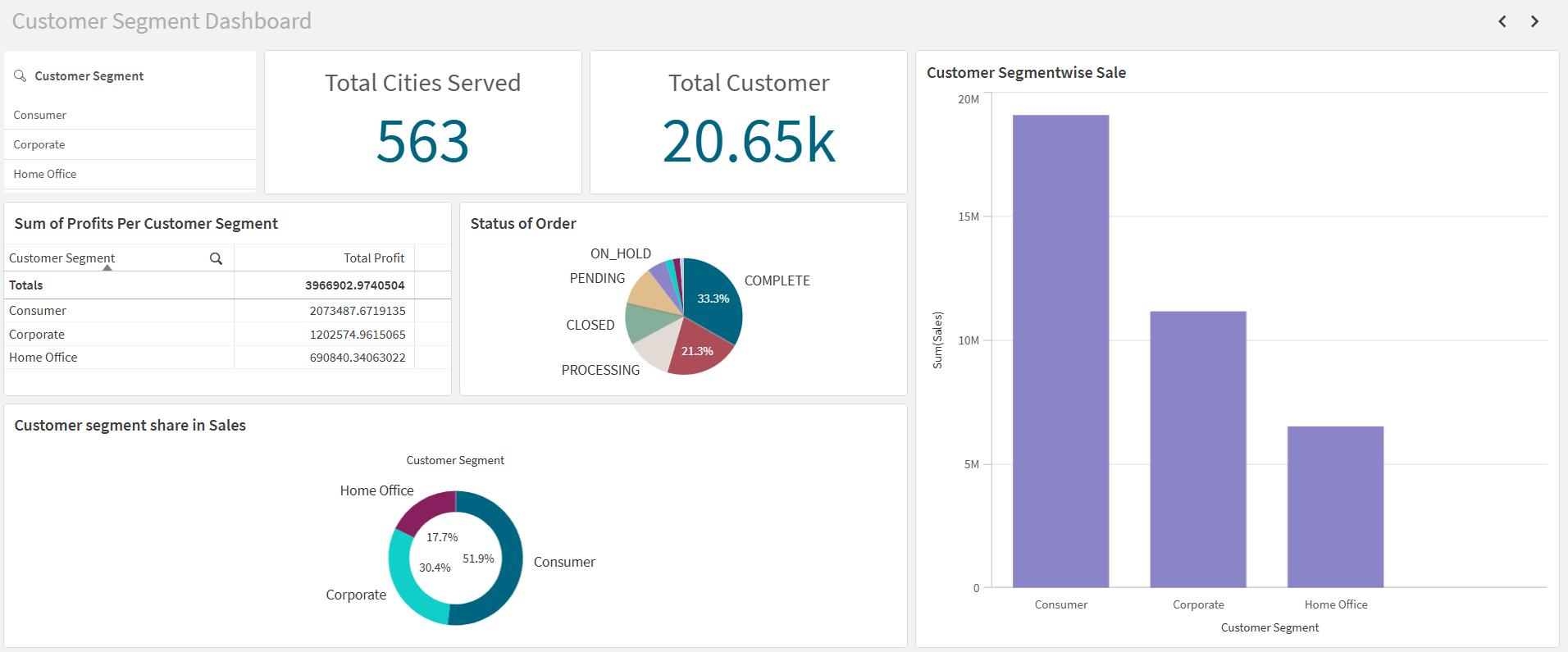


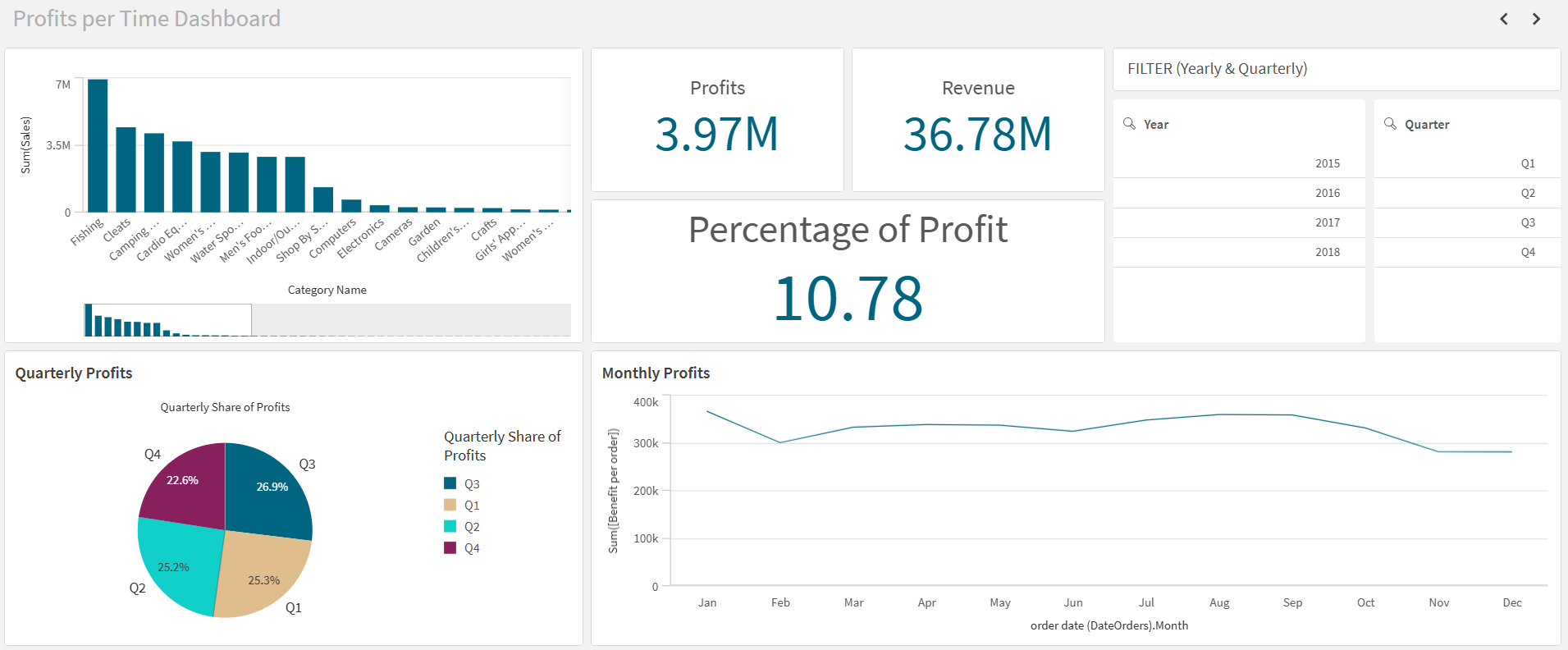




Dashboard Creation

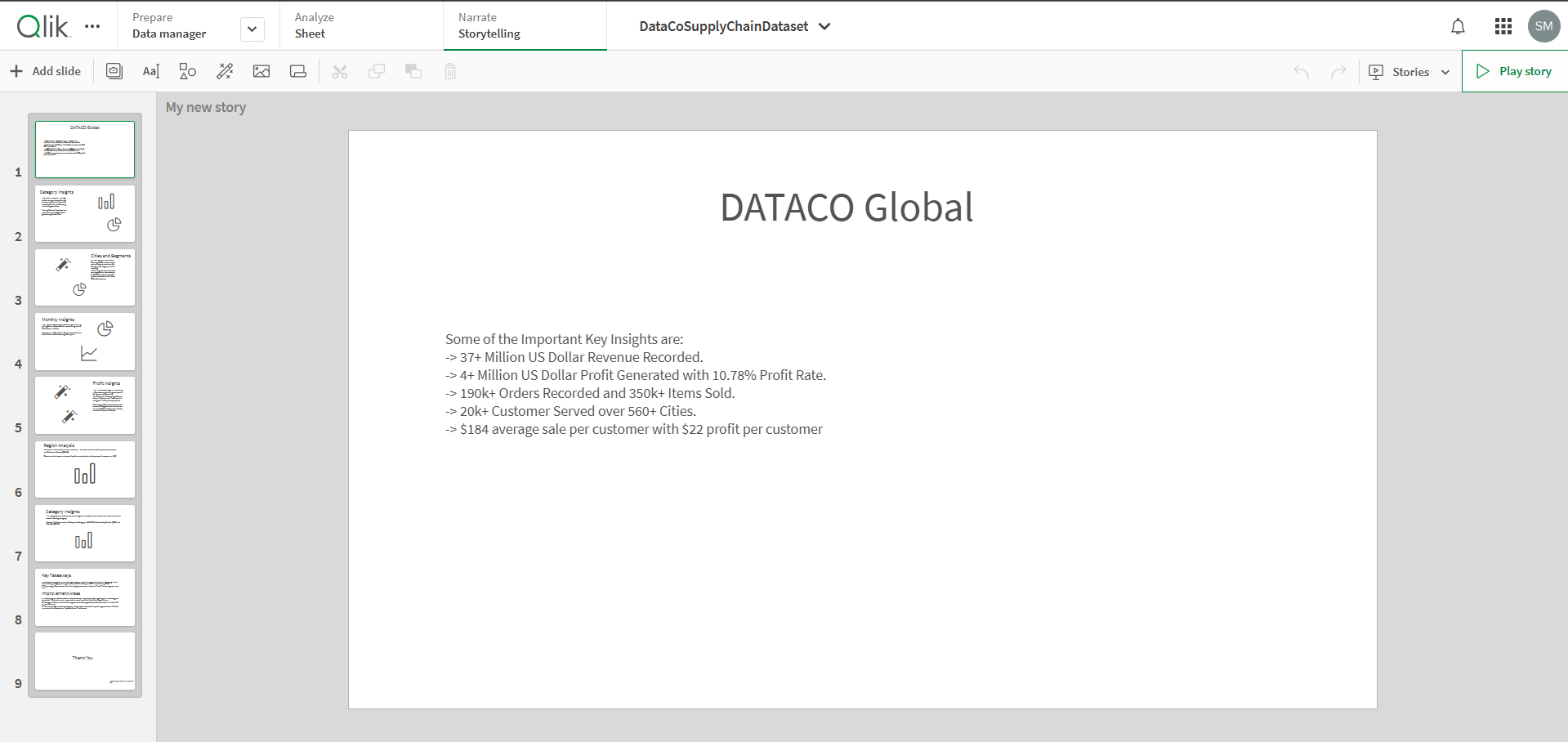
Here are some of the Dashboard snapshots from the projects:

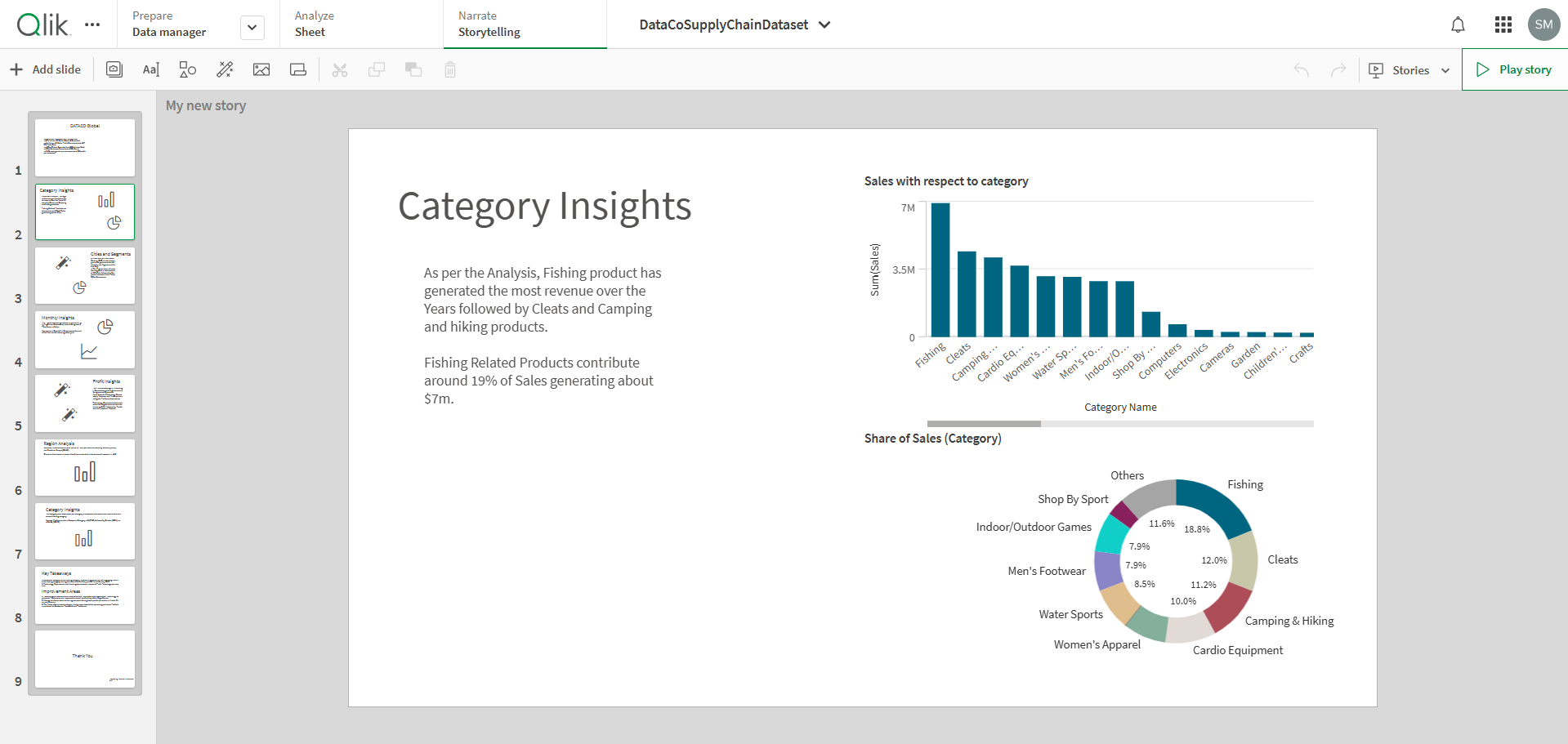


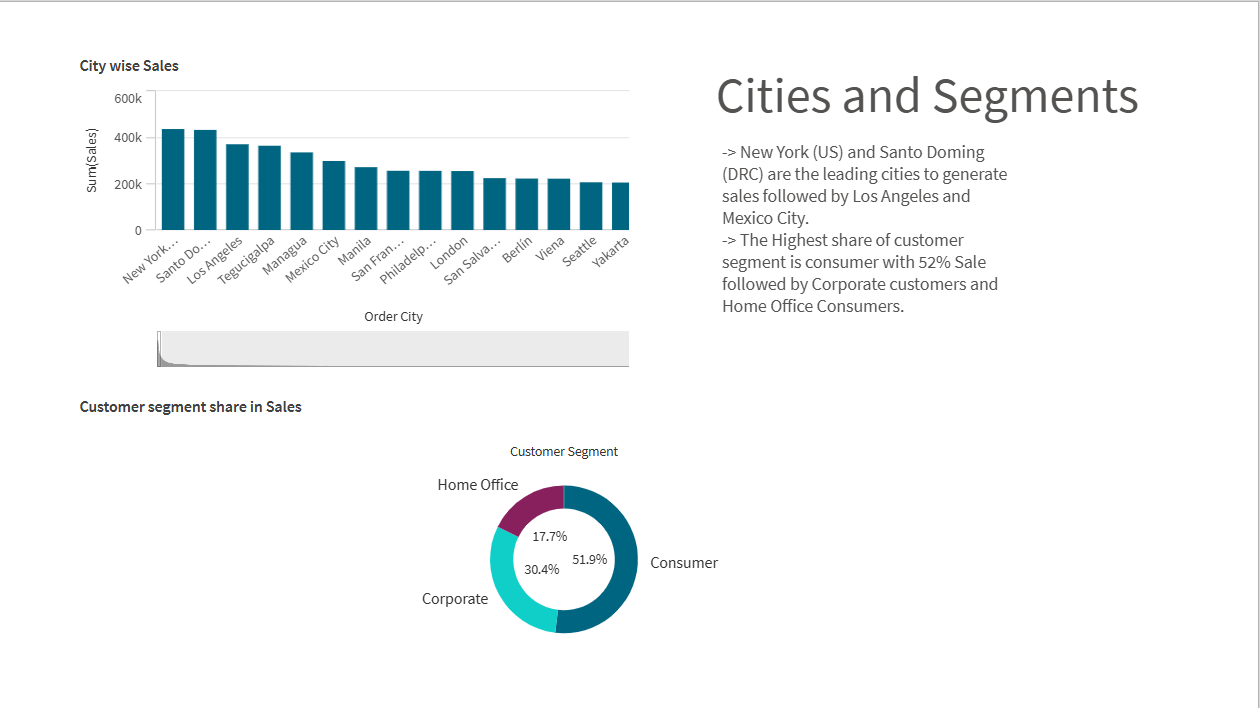


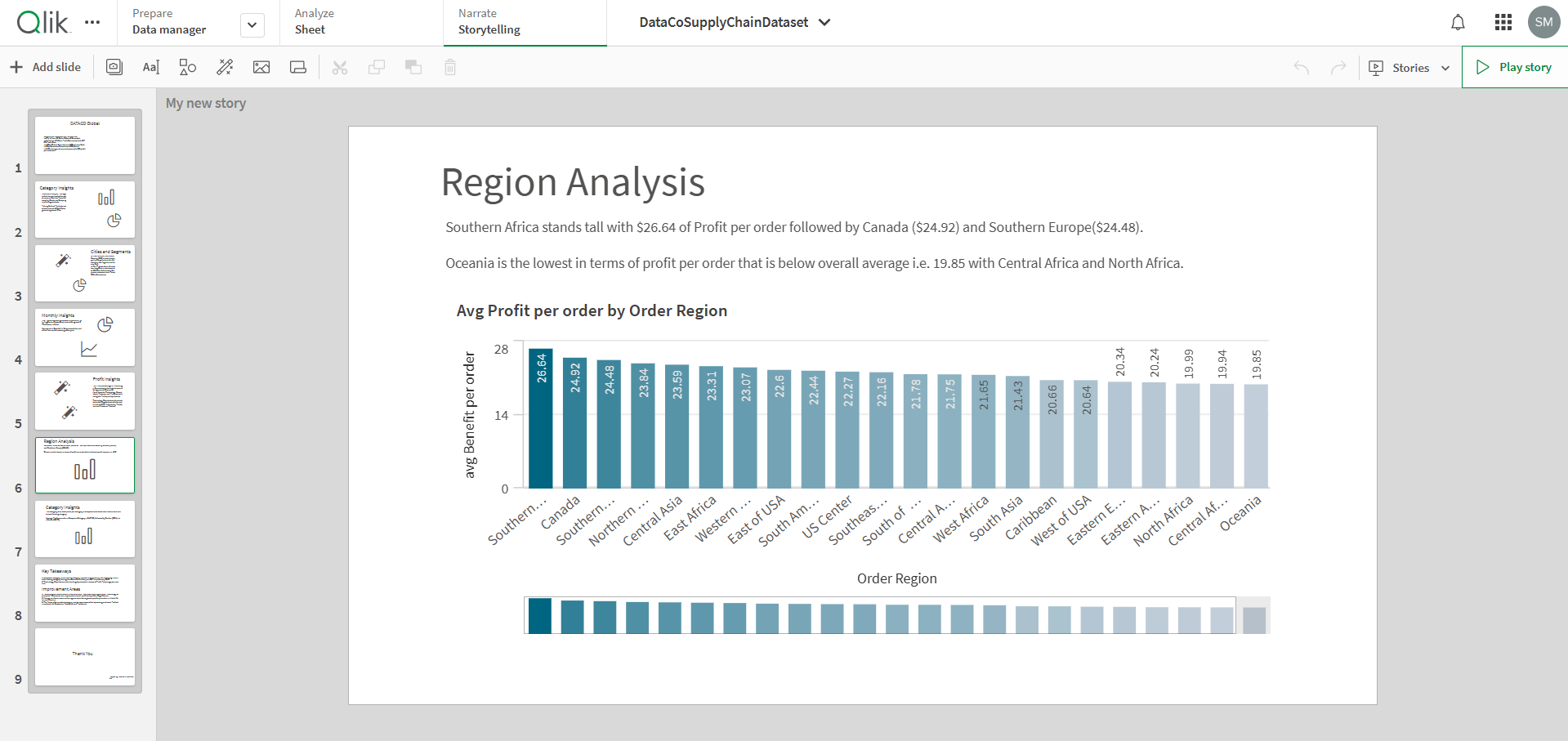
StoryTelling

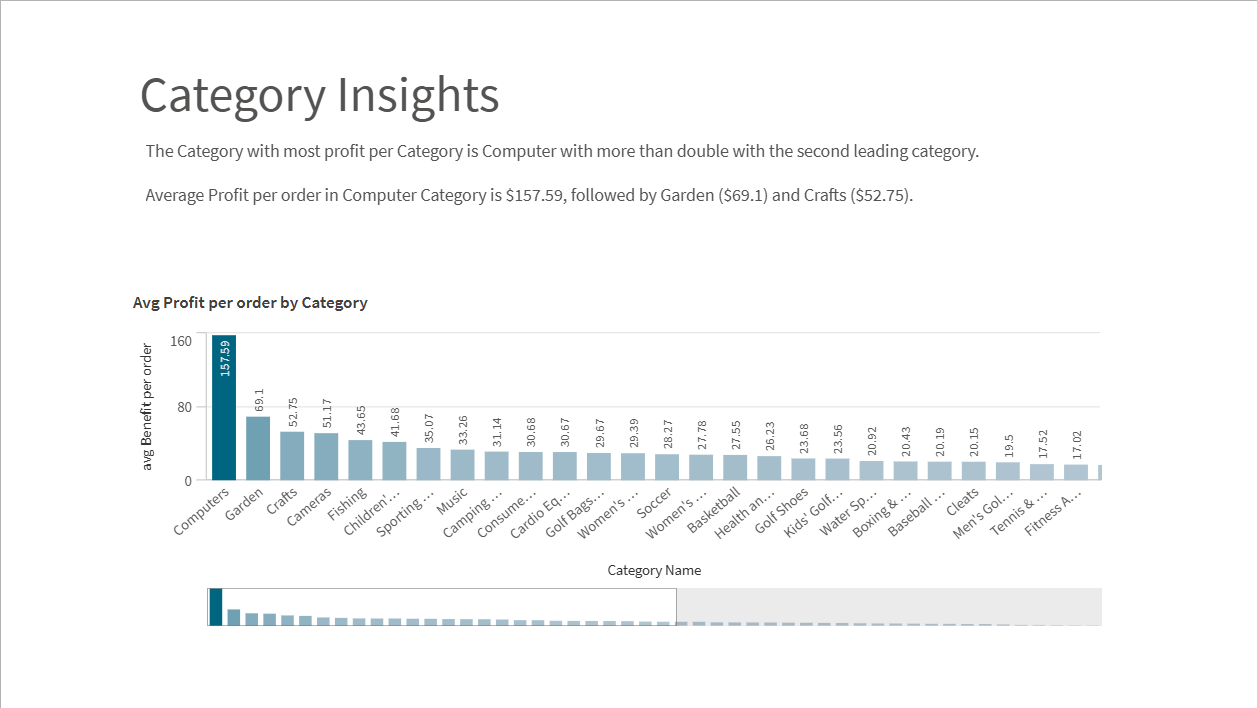
Here are some of the snapshots of the Story Telling from Qlik Sense.











Performance Testing

Dataset Overview

* Total Unique Columns:53
* Total Unique Rows: 180,497

To ensure that the data analytics process is efficient and scalable, performance testing was conducted. This involved evaluating how the dataset performs under various conditions and identifying any potential bottlenecks.

Number of Columns and Rows

The dataset consists of 53 unique columns and 180,497 unique rows. The large size of the dataset necessitates robust performance testing to handle data processing, querying, and visualization efficiently.

Key Insights from the Projects

1. Computer Category:

* The Computer category stands out as the leading category in terms of profitability.
* It also emerged as the top-performing category during the second half of 2017 and in January 2018, indicating its growing importance and potential for sustained high performance.

2. Technology Department:

* The Technology Department is the frontrunner in terms of profit percentage per revenue.
* This highlights the department's efficiency and effectiveness in converting revenue into profit, making it a crucial area for strategic focus.

Improvement Areas

1. Focusing on Technology Products:

* Analyzing recent trends in sales and profit data reveals that an increased focus on products from the Technology Department can significantly boost the overall profitability of the organization.
* By leveraging the strengths of this department, the organization can capitalize on high-margin products and drive growth.

2. Regional Focus:

* It's important to pay additional attention to regions that currently contribute lesser profits per order, specifically North Africa and Oceania.
* Targeted strategies in these regions could help improve their profitability and balance the overall profit distribution across different markets.

3. Product Category Focus:

* Shifting the focus towards emerging and more profitable product categories like Computers, Footwear, and Healthcare can further enhance the organization’s profitability.
* These categories show strong potential for growth and profitability, making them ideal candidates for increased investment and marketing efforts.

Conclusion

This project has demonstrated the power of data-driven decision-making in supply chain management. By leveraging advanced analytics tools like Qlik Insights, organizations can uncover valuable insights, streamline operations, and enhance profitability. The findings and recommendations from this project provide a strong foundation for ongoing improvements and strategic initiatives in supply chain management.