### Data-Driven Innovations In Supply Chain Management

#### 1 Introduction

In today's dynamic business landscape, leveraging data analytics has become imperative for organizations striving to gain a competitive edge. This project delves into the realm of business analytics, specifically focusing on visualizing data and compelling storytelling to extract meaningful insights. Utilizing the dataset named "DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS," our endeavor is to uncover actionable intelligence within the intricate web of supply chain operations. This introduction provides an overview of the project's scope, elucidates its purpose and objectives, and offers insights into the technical architecture underpinning this data-driven endeavor.

#### 1.1 Overview of the Project

At the core of this project lies the exploration and analysis of supply chain data to discern patterns, trends, and anomalies that can drive informed decision-making. Through the lens of business analytics, we aim to unearth hidden opportunities, mitigate risks, and optimize supply chain processes for enhanced efficiency and profitability. By harnessing the power of data visualization and storytelling techniques, we endeavor to transform raw data into actionable insights that empower stakeholders to navigate the complexities of modern supply chain management with confidence and clarity.

#### 1.2 Purpose and Objectives

The primary purpose of this project is to harness the potential of data analytics to revolutionize supply chain management practices. Our objectives include:

- Uncovering actionable insights from the DataCo SMART SUPPLY CHAIN dataset through comprehensive analysis.
- Visualizing key metrics and trends to facilitate informed decision-making and strategic planning.
- Crafting compelling narratives around the data to communicate insights effectively to stakeholders.
- Identifying opportunities for process optimization, cost reduction, and revenue enhancement within the supply chain ecosystem.
- Empowering organizations to leverage data-driven insights for competitive advantage and sustainable growth in today's dynamic marketplace.

#### 1.3 Technical Architecture Overview

The technical architecture of this project encompasses a robust framework for data acquisition, processing, analysis, and visualization. Leveraging cutting-edge tools and technologies in the field of business intelligence and analytics, our architecture is designed to seamlessly integrate disparate data sources, perform advanced analytics, and deliver actionable insights through intuitive visualizations and interactive dashboards. Key components of the technical architecture include data integration platforms, analytics engines, visualization tools, and cloud infrastructure to ensure scalability, flexibility, and reliability in handling large volumes of data for Big Data analysis.

### 2. Problem Understanding

#### 2.1 Business Problem Specification

In order to drive meaningful insights and improvements within the supply chain domain, it's essential to precisely articulate the business problem at hand. Within our project context, the business problem revolves around the lack of visibility, inefficiencies, and potential bottlenecks within the supply chain processes. This manifests in challenges such as inaccurate demand forecasting, suboptimal inventory management, and inefficient logistics operations. These issues ultimately lead to increased costs, missed opportunities, and diminished customer satisfaction. By delineating and understanding the specific pain points within the supply chain, we aim to tailor data-driven solutions that address these challenges head-on and drive tangible business outcomes.

#### 2.2 Business Requirements Analysis

A comprehensive analysis of business requirements is paramount to ensuring that our data-driven solutions align with organizational objectives and stakeholder expectations. Through extensive stakeholder engagement and requirements elicitation sessions, we have identified key functional and non-functional requirements that underpin our project. These requirements encompass diverse facets such as data accuracy, timeliness, scalability, usability, and security. Additionally, specific use cases and scenarios have been delineated to guide the development of actionable insights and analytics models. By prioritizing and addressing these requirements, we aim to deliver a solution that not only meets but exceeds the expectations of stakeholders, driving value across the supply chain ecosystem.

#### 2.3 Literature Survey on Supply Chain Management Innovations

A comprehensive literature survey has been conducted to explore emerging trends, innovations,

and best practices within the realm of supply chain management. This survey encompasses a wide array of scholarly articles, research papers, industry reports, and case studies from reputable sources. Key themes and topics explored include advancements in demand forecasting techniques, inventory optimization strategies, logistics and transportation management, supplier relationship management, and sustainability initiatives within the supply chain. By synthesizing insights from the literature survey, we aim to leverage proven methodologies and innovative approaches to inform our data analytics strategies and drive impactful outcomes within the supply chain domain.

#### 3. Data Collection

#### 3.1 Collect the Dataset

The foundation of any successful data-driven project lies in the quality and comprehensiveness of the dataset. In this phase, our first objective is to procure a robust dataset that encapsulates the intricacies of supply chain operations. Leveraging the DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset, we gain access to a rich repository of structured and unstructured data spanning various facets of the supply chain, including sales, procurement, inventory, logistics, and customer interactions. Through rigorous data collection processes, we ensure that the dataset is representative, accurate, and up-to-date, laying the groundwork for meaningful analysis and insights generation.

#### 3.2 Connect Data with Qlik Sense

With the dataset in hand, the next crucial step is to seamlessly integrate it with the Qlik Sense platform for streamlined data analysis and visualization. Leveraging Qlik's robust connectivity options and data integration capabilities, we establish a secure and efficient connection between the dataset and the Qlik Sense environment. This involves configuring data connections, defining data models, and optimizing data loading processes to ensure optimal performance and usability. By connecting the dataset with Qlik Sense, we empower users to explore and interact with the data dynamically, unlocking actionable insights and driving informed decision-making across the supply chain ecosystem.

#### 4. Data Preparation

#### 4.1 Prepare the Data for Visualization

Effective data visualization hinges upon the quality and cleanliness of the underlying data. In this phase, we focus on preparing the dataset for visualization by employing a series of data preparation techniques. This includes data cleaning, transformation, normalization, and enrichment processes to ensure that the data is accurate, consistent, and conducive to visualization. Through automated data cleansing algorithms and manual validation checks, we address data anomalies, missing values, and inconsistencies, thereby enhancing the reliability

and integrity of the dataset. Additionally, we structure the data in a format that is optimized for visualization, incorporating relevant dimensions, measures, and hierarchies to facilitate meaningful analysis and exploration within the Qlik Sense environment.

#### **5 Data Visualizations**

#### 5.1 Visualizations

#### 1. Total Items placed by the Customer in the Country

This graph is a pie chart titled "Total Items Placed by the Customer in Country." It shows the distribution of items placed by customers in two different countries:

- 1. **EE. UU.** (United States)
- 2. Puerto Rico

#### **Explanation:**

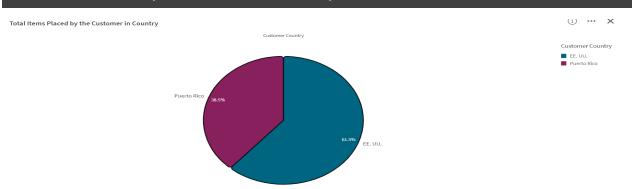
- **EE. UU.** (United States): Represented by the blue section of the pie chart, which makes up 61.5% of the total items placed by customers.
- **Puerto Rico**: Represented by the purple section of the pie chart, which makes up 38.5% of the total items placed by customers.

#### Interpretation:

- A larger proportion of items are placed by customers in the United States (61.5%) compared to Puerto Rico (38.5%).
- This distribution indicates that the majority of the activity or orders come from the United States.

The pie chart effectively visualizes the relative contributions of each country to the total number of items placed by customers.

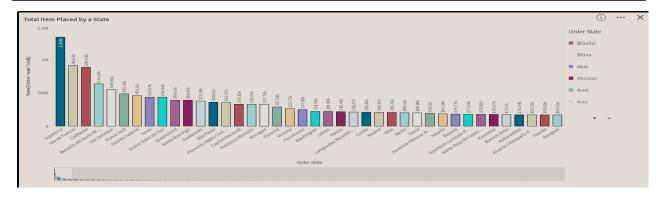
### Total Items Placed by the Customer in Country



#### 2. Total Items Placed by a State

This bar chart titled "Total Item Placed by a State" visualizes the total number of items ordered from various states or regions. The x-axis represents different states or regions, while the y-axis represents the sum of items placed in thousands. Each bar corresponds to a state or region, with the height indicating the total number of items ordered by customers from that area. England leads with 1.34 million items placed, followed by Isla d. with 911.4k items and California with 884.01k items. Other notable regions include Rena. with 634.69k items, San S. with 549.91k items, and Nuev. with 485.31k items. In the lower range, regions like Lagos have the least items placed at 164.91k, with others like Bangalore, Alsace, and Florida placing around 169k to 170k items. The chart highlights the geographical distribution of order volumes, with higher activity in regions like England and Isla d., and a gradual decline as you move towards regions like Lagos. This distribution is crucial for understanding order patterns, aiding in supply chain planning and resource allocation.

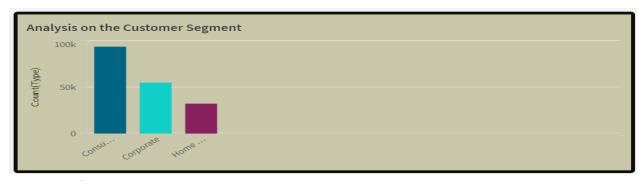
### Total Items Placed by a State



#### 3. Analysis on Customer Segment

This bar chart titled "Analysis on the Customer Segment" visualizes the distribution of orders across different customer segments: Consumer, Corporate, and Home Office. The y-axis represents the count of orders. The Consumer segment, represented by the blue bar, has the highest number of orders, nearing 100,000. The Corporate segment, depicted by the cyan bar, has a moderate number of orders, around 60,000. The Home Office segment, shown by the purple bar, has the lowest order volume, approximately 40,000. This analysis indicates that the Consumer segment is the largest in terms of order volume, followed by the Corporate segment, with the Home Office segment generating the least orders. These insights are valuable for tailoring marketing strategies, optimizing inventory, and enhancing customer service to meet the specific needs of each segment.

### Analysis on Customer Segement



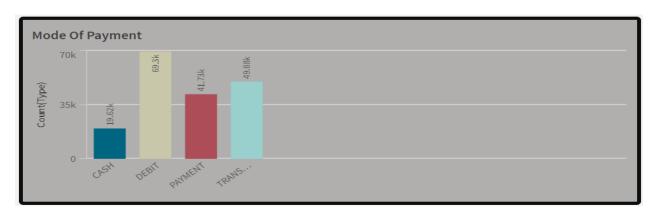
#### 4. Mode of Payment

The bar chart displays the distribution of different modes of payment used, categorized by the count of transactions for each type. The four modes of payment shown are CASH, DEBIT, PAYMENT, and TRANS..., with the y-axis representing the count of transactions in thousands.

- 1. **CASH**: This mode has the lowest count of transactions, amounting to 19.62k.
- 2. **DEBIT**: This is the most frequently used mode of payment, with 69.3k transactions.
- 3. **PAYMENT**: This mode has a moderate count, with 41.73k transactions.
- 4. **TRANS...**: The exact name of this mode is truncated, but it has a significant number of transactions, totaling 49.88k.

The chart clearly indicates that DEBIT is the most popular mode of payment among users, while CASH is the least used. The other two categories, PAYMENT and TRANS..., fall in between, with TRANS... having a higher count than PAYMENT.

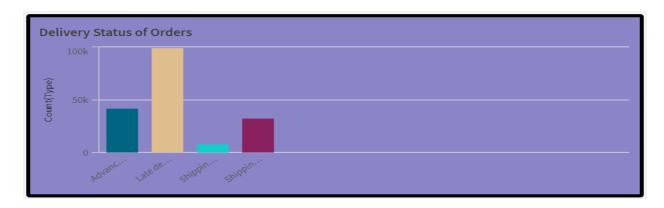
### Mode Of Payment



#### 5. Delivery status of orders

The bottom-right chart categorizes orders based on their delivery status: Advance, Late delivery, Shipping, and Shipped. "Late delivery" stands out with the highest count of nearly 100k transactions. "Advance" deliveries account for a moderate number of transactions, around 50k. "Shipped" orders are fewer, with about 25k transactions, while "Shipping" orders have the lowest count. This highlights that late deliveries are a significant issue, with the majority of orders falling into this category.

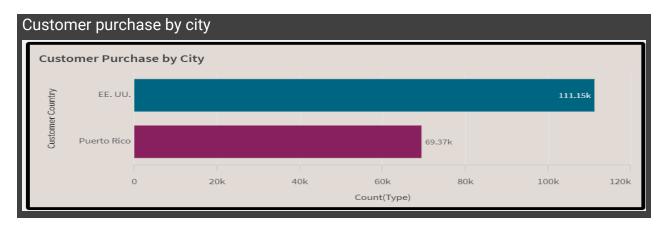
#### Delivery status of orders



#### 6. Customer purchase by city

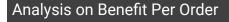
The bottom-left chart compares customer purchases between two locations: the United States (EE. UU.) and Puerto Rico. The United States has a significantly higher transaction count at 111.15k, compared to Puerto Rico's 69.37k. This indicates that customers in the United States

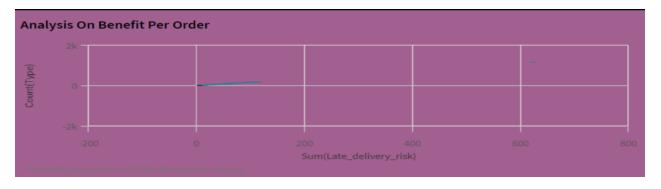
are more active in making purchases than those in Puerto Rico.



#### 7. Analysis on benefit per order

The top-left scatter plot shows the relationship between the sum of late delivery risk and the count of orders. The data points are scattered across the plot, with a slight upward trend indicating that as the late delivery risk increases, the count of orders also shows a minor increase. This could imply that orders with higher risks of late delivery are relatively frequent, suggesting a need to address factors contributing to this risk to improve overall efficiency and customer satisfaction.





#### 8. Analysis of Profit Ratio

The top-right histogram illustrates the frequency distribution of the order item profit ratio. Most of the data is clustered around a profit ratio close to zero, with a notable peak at 0.5, showing around 62.05k transactions. Other significant frequencies are seen at profit ratios of -2.5, 0, and 0.5. This distribution suggests that while many transactions are marginally profitable or break-

even, there are distinct groups of highly profitable and unprofitable transactions. Identifying and understanding the characteristics of these groups can help in strategizing to improve profitability

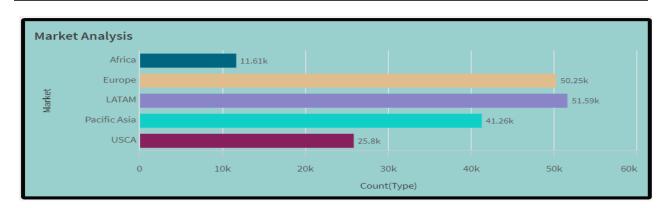
### Analysis of Profit Ratio



#### **Market Analysis**

The bottom-left bar chart breaks down the count of transactions across various markets: Africa, Europe, LATAM, Pacific Asia, and USCA. LATAM has the highest transaction count at 51.59k, followed closely by Europe with 50.25k transactions. USCA and Pacific Asia also show substantial activity with 25.8k and 41.26k transactions respectively, while Africa has the lowest at 11.61k. This chart highlights the geographical distribution of business activity, indicating that LATAM and Europe are the most active markets, which can inform marketing and operational strategies.

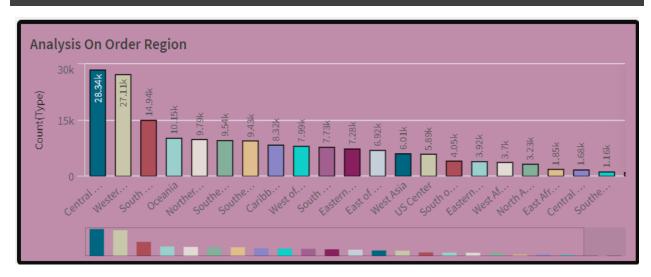
#### Market Analysis



### 10 Analysis on Order Region

The bottom-right bar chart provides a detailed breakdown of order counts by specific regions. South America, with 14.94k orders, is highlighted in the tooltip. Other notable regions include Western Europe, Eastern Europe, Pacific Asia, and North America, each with varying transaction counts. This detailed regional analysis helps in understanding the distribution of orders across different parts of the world, which can be crucial for logistics planning, regional marketing strategies, and resource allocation to meet regional demands effectively.

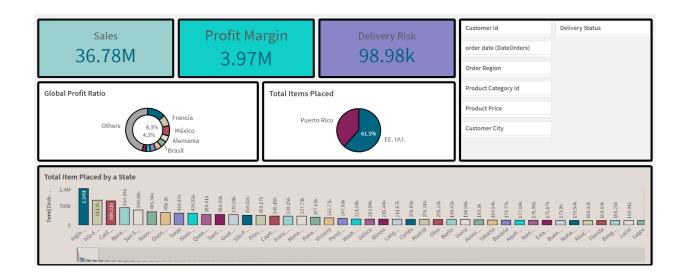
### Analysis on Order Region



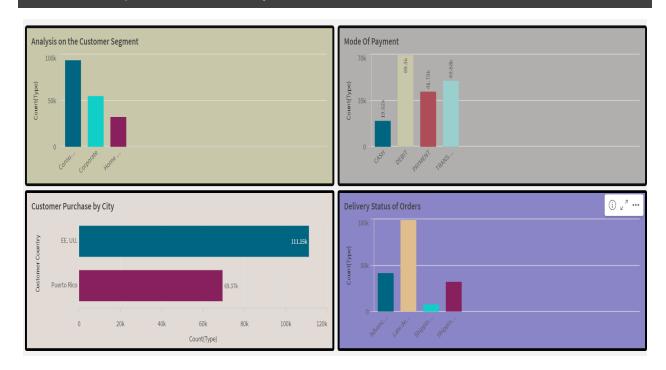
### 6 Dashboards

#### Dashboard 1

#### Global Sales and Performance Dashboard

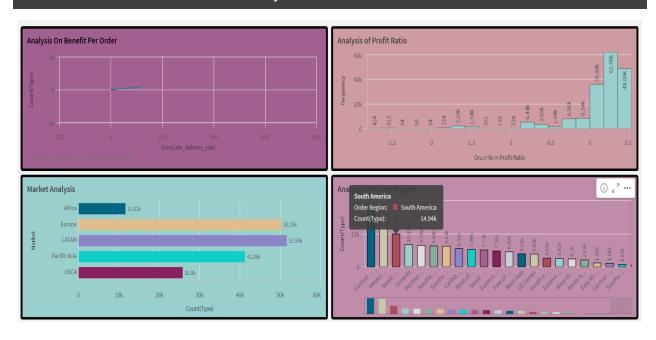


Dashboard 2
Customer Insights and Order Analysis Dashboard



### **Dashboard 3**

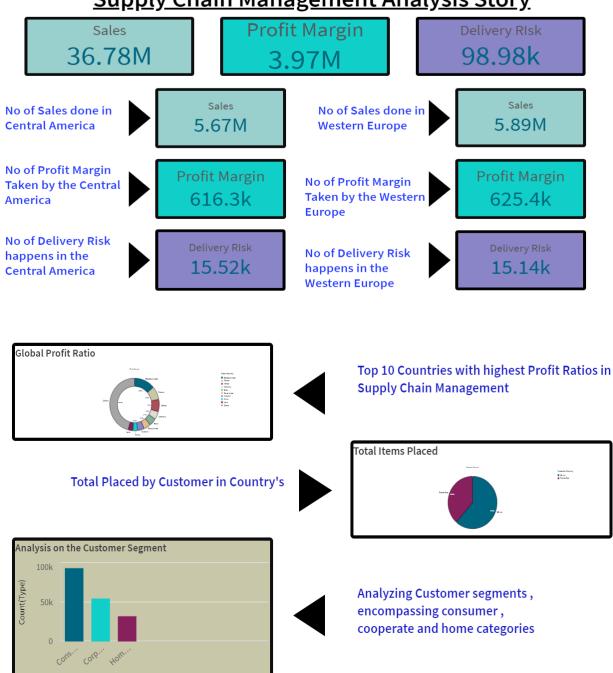
### Order Performance and Market Analysis Dashboard



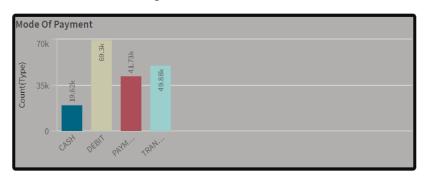
### 7. Report

### 7.1 Report Creation

# **Supply Chain Management Analysis Story**

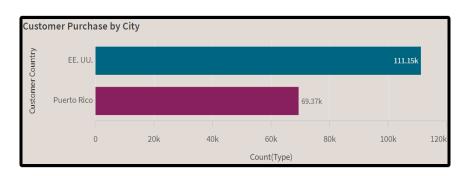


## **Mode of Payment for Purchase**



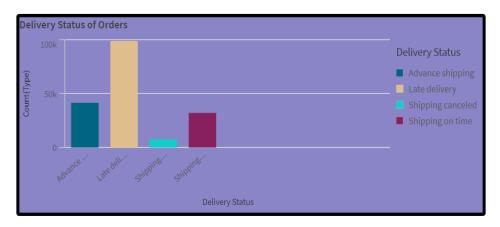
- Cash transactions offer immediate liquidity, providing a straightforward and tangible method of payment.
- Debit payments, directly linked to bank accounts, offer convenience and real-time deduction of funds
- Credit payments provide a deferred payment option, allowing customers to make purchases
- Transfer payments leverage electronic methods for seamless and secure fund

### **Customer Purchase Item by City**



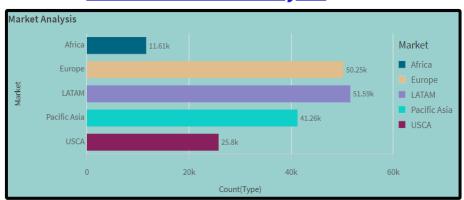
- Customer purchase count in Puerto Rico reflects the transactional dynamics in this vibrant location, capturing the local consumer behavior and market engagement.
- Customer purchase count in the United States provides a comprehensive overview of buying patterns across diverse cities.

## **Delivery Status of Order**



- Analyzing the delivery status of orders, including Advanced Shipping, Late Shipping, Shipping Canceled, and Shipping on Time.
- This allows businesses to evaluate the efficiency of their logistics operations, address potential delays, and enhance customer satisfaction.
- Ensuring timely and reliable deliveries based on varied shipping scenarios.

### **Global Market Analysis**



- Conducting a market analysis across Africa, Europe, LATAM (Latin America), Pacific Asia, and USCA (United States and Canada) enables businesses to gain strategic insights.
- Regional economic landscapes, consumer behaviors, and market dynamics.
- This comprehensive assessment supports informed decision-making, tailored marketing strategies, and targeted expansion efforts to capitalize on diverse opportunities within each distinct market.

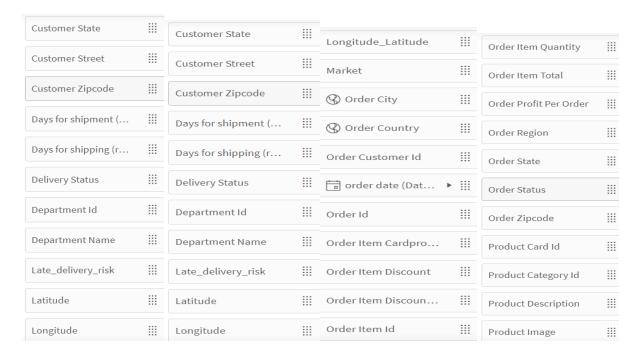
#### 8. Performance Testing

#### 8.1 Amount of Data Loaded

Performance testing for data loading evaluates how well a system handles large volumes of data. This ensures the system meets performance criteria for efficiency and effectiveness. The key steps involved are:

- 1. **Define Objectives and Metrics:** Establish performance goals such as maximum load times and acceptable throughput rates.
- 2. **Prepare the Environment:** Set up a test environment that mirrors production, including necessary hardware and realistic test datasets.
- 3. **Plan and Execute Tests:** Develop load, stress, and volume testing scenarios to assess system performance under various conditions.
- 4. **Monitor and Measure:** Track resource utilization, load times, and throughput during tests to identify performance bottlenecks.
- 5. **Reporting:** Compile a detailed report with visualizations to present findings, highlighting key metrics and recommendations for optimization.

By following these steps, performance testing ensures that the system can efficiently manage large data volumes, supporting effective data-driven decisions.



#### 8.2 Utilization of Data Filters

Utilization of data filters in Qlik Sense refers to applying criteria to restrict datasets, enabling targeted and meaningful analysis. These filters help users focus on specific subsets of data, extract precise insights, and make informed decisions. By setting various dimensions, values, or conditions, data filters enhance analysis by narrowing down information and emphasizing relevant data. Here's an overview of effectively utilizing data filters in Qlik analysis:

- 1. **Understanding Filters:** Apply criteria to restrict data, removing noise for a focused analysis.
- 2. **Types of Filters:** Use selection, list boxes, advanced search, and set analysis.
- 3. **Applying Filters:** Interactively click data points, use list boxes, data range filters, and custom set analysis.
- 4. **Best Practices:** Start broad and progressively narrow down, use clear naming conventions, and combine filters wisely.
- 5. **Examples:** Apply filters for sales analysis, customer segmentation, and inventory management.
- 6. **Benefits:** Enhance focus, improve performance, customize insights, and create interactive applications.
- 7. **Implementation:** Design filter panes, use master items, and document filter logic.

Effectively utilizing data filters enhances analysis, improves focus, and tailors insights to specific business needs.