Ecommerce

Streamline & Spotlight: Unraveling the Narrative of Supply Chain Excellence and Digital Footprints

Background:

DataCo Analysis Group, a leading consultancy in business analytics, specializes in providing comprehensive insights into various industry sectors. This case study, "Streamline & Spotlight," focuses on the intersection of supply chain operations and digital consumer behavior. With a detailed dataset encompassing supply chain metrics (DataCoSupplyChainDataset.csv) and digital access logs (TokenizedAccessLogs.csv), DataCo Analysis Group plays an instrumental role in helping businesses optimize their operations and understand digital engagement. The supply chain dataset offers insights into sales, logistics, and delivery performance, while the access logs shed light on consumer online interactions and preferences. In an era where efficient supply chain management and robust digital presence are crucial for business success, analyzing these datasets becomes imperative to uncover strategies for operational excellence and enhanced digital outreach.

Objective:

The objective of this case study is to conduct an in-depth analysis of the combined supply chain and digital access datasets provided by DataCo Analysis Group. Students will utilize advanced Excel techniques to dissect these datasets, revealing key patterns and insights. The primary tasks include data cleaning, integration, analysis, and the development of a comprehensive, interactive dashboard. This dashboard will serve as a critical tool in visualizing supply chain efficiencies, customer engagement metrics, and the synergy between them. The project aims to augment DataCo Analysis Group's consulting capabilities, offering actionable insights to clients for optimizing supply chain processes and improving digital marketing strategies. This analysis is also expected to contribute to a broader understanding of how supply chain and online consumer behavior are interlinked, thereby influencing strategic decisions in product management, marketing, and customer relationship management.

Data Source:

1. **Supply Chain Dataset:**

DataCoSupplyChainDataset.csv

This dataset is designed to analyze aspects of the supply chain such as shipping efficiency, customer demographics, sales performance, and product popularity.

FIELDS	DESCRIPTION
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 time
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 Asia, West Asia , West of USA , US Center , West Africa, Central Africa ,North Africa ,Western Europe ,Northern , Caribbean , South America ,East Africa ,Southern Europe , East of USA ,Canada ,Southern Africa , Central Asia , Europe , Central America, Eastern Europe , South of USA
Order State	: State of the region where the order is delivered
Order Status	: Order Status : COMPLETE , PENDING , CLOSED , PENDING_PAYMENT ,CANCELED , PROCESSING ,SUSPECTED_FRAUD ,ON_HOLD ,PAYMENT_REVIEW
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

2. **Access Logs Dataset**

TokenizedAccessLogs.csv

This dataset is for understanding user behavior on the website, the popularity of products, and the performance of different website sections.

1. **Product**: Name of the product.
2. **Category**: Category of the product.
3. **Date**: Date and time of the access.
4. **Month**: Month of the access.
5. **Hour**: Hour of the day when the access occurred.
6. **Department**: Department to which the product belongs.

7. **ip**: IP address of the user accessing the website.
8. **url**: Specific URL that was accessed on the website.

Part 1: Excel Data Analysis: Manipulation, Formulas and Functions

1. **Missing Data Handling:** Identify and address missing data in the datasets. Are there any patterns in the missing data that can be noted?
2. **Date Analysis**: Calculate the average shipping delay (difference between 'Days for shipping (real)' and 'Days for shipment (scheduled)') for each product category.
3. **Profit Margin Calculation**: Create a formula to calculate the profit margin for each order and categorize them as 'High', 'Medium', or 'Low'.
4. **Customer Geographic Analysis**: Use COUNTIF and other functions to analyze the distribution of customers across different cities and countries.
5. **Sales Trend Analysis**: Analyze monthly sales trends over the years and identify peak sales months using date functions.
6. **Top 5 products:** Identify the top 5 products with the highest sales.
7. **Advanced Filtering for High-Risk Deliveries**: Use advanced filtering to identify orders with a high risk of late delivery and high sales value.
8. **Product Popularity Index**: Develop an index to rate product popularity based on sales volume and frequency.
9. **Dynamic Data Range for Sales Analysis**: Create dynamic named ranges for different product categories and use them to calculate total sales for each category.
10. **Order Processing Efficiency**: Calculate the average processing time for orders (from order date to shipping date) and identify the department with the best performance.
11. **Customer Loyalty Assessment**: Assess customer loyalty by calculating the average number of orders per customer.
12. **Delivery Status Breakdown by Market**: Use a pivot table to analyze the breakdown of delivery status (e.g., on time, late) by market regions.
13. **Peak Traffic Time Analysis**: Determine the peak hours of website traffic and which products are most viewed during these times.
14. **Category Popularity by Month**: Use pivot tables to analyze which product categories are most popular in each month.
15. **User Engagement Analysis**: Calculate the average number of product views per IP address to assess user engagement.
16. **Departmental Traffic Analysis**: Analyze which department's products are most frequently accessed using COUNTIF and SUMIF functions.
17. **Most Visited URLs**: Identify and rank the most frequently visited URLs in the dataset.
18. **Product Interest vs. Sales Performance Analysis**: Compare the frequency of product views from the Access Logs dataset with the sales data for the same products in the Supply Chain dataset. Determine if higher online views correlate with higher sales.
19. **Customer Geographic Interest Analysis**: Match the customer cities from the Supply Chain dataset with IP addresses from the Access Logs dataset (approximate analysis due to the nature of IP geolocation). Analyze which cities show the most online engagement compared to actual sales.

20. **Time Series Analysis of Product Interest and Sales**: Compare the monthly trends in product views from the Access Logs dataset with the monthly sales trends of those products in the Supply Chain dataset. Identify any lag or lead relationship between interest and sales.
21. **Complex Profit Margin Analysis with Conditional Logic**: Develop an advanced formula to calculate the profit margin for each order, factoring in different variables such as product category, shipping mode, and delivery status. For instance, adjust the profit margin calculation based on whether the delivery was late or on time, and whether the shipping mode was standard or express. (This would involve nested IF statements or a combination of IF with other functions like VLOOKUP.)
22. **Predictive Sales Analysis Using Regression**: Use Excel's data analysis toolpak to perform a regression analysis. Predict future sales for different product categories based on historical data, considering variables such as market region, customer segment, and seasonality. This will involve setting up the regression model, interpreting the coefficients, and understanding the statistical significance of the results.
23. **Dynamic Inventory Optimization Model**: Create a model that calculates optimal inventory levels for different products based on sales velocity, lead time for shipping (real vs. scheduled), and buffer stock requirements. This model should dynamically adjust to changes in sales data and shipping performance over time. Use Excel functions like AVERAGE, STDEV, and various logical functions to build this model. Additionally, incorporate Excel's scenario manager to see how changes in sales trends or shipping performance affect inventory recommendations.

(Note: Show Visualizations wherever possible in Part 1)

Part 2: Building an Excel Dashboard

Leverage Excel's data visualization and interactive tools to create an insightful and dynamic dashboard. This dashboard should provide a clear and interactive overview of key performance indicators (KPIs), trends, and insights in the supply chain and web access areas. The goal is to develop a tool that facilitates data-driven decision-making and offers a comprehensive understanding of the underlying business processes.

Key Elements to Include:

1. **Interactive Filters:**
   - Include options to filter data by time periods (e.g., month, year), product categories, departments, and customer locations.
2. **Supply Chain Metrics:**
   - Total Sales Over Time: A line or area chart displaying trends in sales.
   - Top Performing Products: A bar or column chart showing the products with the highest sales.
   - Order Delivery Status: A pie chart or donut chart showing the proportion of on-time vs. late deliveries.

3. **Access Logs Analysis:**
   - Product View Trends: A line chart showing the number of product views over time.
   - Most Viewed Products: A bar chart displaying the most frequently viewed products.
   - User Activity Heatmap: A heatmap showing product views by hour and day.
4. **Comparative Analysis:**
   - A scatter plot or bubble chart comparing sales vs. number of views for products.
   - A bar chart comparing the product categories in both datasets.
5. **Data Tables:**
   - Include summary tables showing key metrics like total sales, average delivery time, total views, and unique visitors.
6. **Dashboard Usability:**
   - Ensure that the dashboard is user-friendly, with clear labels and legends.
   - Implement slicers or dropdown menus for dynamic interactivity.
7. **Advanced Visualization:**
   - Incorporate advanced charts like sparklines or conditional formatting to highlight trends or anomalies.