

Online Superstore: Sales and Performance Dashboard

Contents

Executive Summary.....	1
Project Introduction	2
Objectives	3
Dataset Description	4
Dataset Constraints	5
Methodology.....	6
Dashboard Overview.....	10
Recommendations	12
Future Work	13

Executive Summary

This report analyses a synthetic dataset reflecting transactions from an online superstore between 2019 and 2020, focusing on sales and delivery performance across various customer segments and product categories.

- **Dataset Overview:** The dataset includes details on orders and returns, capturing customer information, product categories, shipping methods, and sales figures, providing a comprehensive view of e-commerce operations.
- **Key Objectives:** The analysis aims to identify trends in sales and delivery performance, assess customer behaviour, and evaluate delivery compliance across different shipping methods.
- **Methodology:** Data was processed using Microsoft Power BI, employing DAX functions to create calculated columns and performance metrics, facilitating insights into operational efficiency.
- **Dashboard Features:** The Power BI dashboard presents key performance indicators (KPIs) such as total sales, average sale value, and delivery obligations met, allowing stakeholders to filter insights by year, product category, and shipping type.
- **Performance Insights:** The report highlights a 60.74% compliance rate for delivery obligations, with notable discrepancies in performance across different shipping methods, particularly First Class deliveries.

Recommendations for Improvement: Suggestions include investigating fulfilment delays in First Class shipping, enhancing visibility into order processing timelines, and prioritising improvements in high-sales regions to boost customer satisfaction.

Project Introduction

This report presents an analysis of a synthetic dataset simulating transactions from an online superstore between 2019 and 2020. The dataset captures a range of transaction details across different U.S. states, product categories, and customer segments, specifically Consumers, Corporates, and Home Offices. It has been designed for educational and analytical purposes to reflect realistic sales and delivery operations in a growing e-commerce business.

While the data does not represent actual business performance, it has been structured to mirror key operational scenarios. This includes order volumes over time, customer buying behaviour, and the impact of delivery logistics on customer satisfaction. A particular focus of the analysis is on delivery performance, examining whether items shipped using delivery classes (Same Day, First Class, Second Class, Standard Class) were processed on time, and in turn received within the expected timeframes.

The analysis was conducted using Microsoft Power BI, with extensive use of DAX functions to create aggregations, calculated columns, and delivery metrics. The report explores sales trends over time, regional performance, customer type behaviour, and delivery compliance. It aims to deliver insights that could support strategic decisions in order fulfilment, customer segmentation, and sales operations in a rapidly scaling online retail environment.

Objectives

The primary aim of this analysis is to explore key sales and delivery performance metrics within the Online Superstore dataset, with the goal of uncovering trends, inefficiencies, and actionable insights. Rather than testing predefined hypotheses, the focus is on understanding customer behaviour, delivery reliability, and sales dynamics over time to support operational improvements and strategic decision-making. In particular, this exploration includes:

- Identifying instances of delivery non-compliance by shipping type (e.g., Same Day, First Class, Standard) to evaluate where fulfilment processes may be falling short.
- Assessing sales performance over time, segmented by customer type (Consumer, Corporate, and Home Office).
- Exploring order volumes and trends across different states and time periods.
- Analysing the distribution of units sold across product categories.
- Calculating average sale values to understand purchasing behaviour by segment and over time.
- Highlighting peak sales periods to inform inventory and staffing decisions.

Dataset Description

The dataset consists of two primary tables: Orders and Returns. The Orders table captures detailed information on transactions made through an online superstore between 2019 and 2020. It includes customer details, product categories, shipping methods, sales figures, and order timelines. The *Returns* table provides information about returned orders, enabling an assessment of return rates and patterns. Together, these tables offer a comprehensive view of customer activity, sales performance, and delivery operations.

Table 1 and Table 2 shown below is a description of the columns contained within each table:

Table 1 'Orders' Table Column Descriptions

Column Name	Description
Row ID	Unique index value for each transaction.
Order ID	Unique identifier for each customer order.
Order Date	Date when the order was placed.
Delivery Date	Date when the order was delivered to the customer.
Ship Mode	Shipping method used (e.g., Same Day, First Class, Second Class, Standard).
Customer ID	Unique identifier for each customer.
Customer Name	Full name of the customer who placed the order.
Segment	Type of customer: Consumer, Corporate, or Home Office.
Country	Country of the customer (all orders are within the United States).
City	City where the order was delivered.
State	U.S. state where the order was delivered.
Region	Geographic region within the U.S. (East, West, Central, or South).
Product ID	Unique identifier for the product.
Category	Broad classification of the product (Furniture, Office Supplies, Technology).
Sub-Category	More specific classification within each product category.
Product Name	Full descriptive name of the product.
Sales	Total dollar amount generated from the sale.
Quantity	Number of units of the product ordered.
Profit	Profit earned from the transaction.

Table 2 'Returns' Table Column Descriptions

Column Name	Description
Returned	Indicates if the product was returned ("Yes").
Return Date	Date the return was processed.
Return ID	Unique identifier for the returned order.

Dataset Constraints

Several limitations should be considered when interpreting the findings presented in this report. These constraints, stemming primarily from the structure and synthetic nature of the dataset, are outlined below:

- **Returns Table Limitations:** The Returns table was ultimately excluded from the dashboard, as none of the Return ID values corresponded to any Order ID in the Orders table. As a result, returns could only be evaluated at an aggregated level to assess the overall percentage of returned orders. Furthermore, the table lacked detail on which specific product types or customer segments were more prone to returns, limiting the ability to perform deeper return pattern analysis.
- **Missing Order Processing Date:** The dataset does not include an "Order Processed Date" field, which restricts the ability to assess the time taken between when an order was placed and when it was actually processed or fulfilled. This would have been valuable in understanding internal processing delays and their impact on delivery performance.
- **No Adjustment for Public Holidays:** Delivery performance metrics are based purely on calendar day differences, without accounting for weekends or public holidays. This may slightly skew the accuracy of delivery compliance calculations, as operational constraints on non-working days are not reflected in the data.

These constraints should be kept in mind when interpreting the insights presented, as they may impact the completeness or accuracy of certain conclusions.

Methodology

This section outlines the data preparation and processing steps undertaken to facilitate subsequent analysis. The aim was to ensure data quality, create necessary calculated fields, and define relevant performance measures to support the development of the dashboard.

Data Preparation

The dataset was imported into Power BI from an Excel source file. Initial data transformation was performed using Power Query, where headers were promoted, and each column was assigned an appropriate data type. Text strings were converted to text format, numeric fields were treated as whole numbers or decimals, and all date fields were formatted accordingly. Units were applied to numeric values where relevant to improve clarity and usability during analysis.

Data Modelling and Calculated Columns

Within the Power BI data model, new calculated columns were created in the Orders table to derive insights around delivery compliance:

1. **Shipping Time (Days):** Calculated as the difference between the Delivery Date and the Order Date to determine the number of days taken to deliver each order

```
Shipping Time (Days) = Orders[Delivery Date] - Orders[Order Date]
```

2. **Shipping Met?:** A conditional column using the SWITCH() and TRUE() functions to check whether the delivery met the expected timeline based on the shipping method selected.

```
Shipping Met? =  
SWITCH(  
    TRUE(),  
    Orders[Ship Mode] = "Same Day" && Orders[Shipping Time (Days)] <=  
0, "Met",  
    Orders[Ship Mode] = "First Class" && Orders[Shipping Time (Days)]  
<= 1, "Met",  
    Orders[Ship Mode] = "Second Class" && Orders[Shipping Time  
(Days)] <= 3, "Met",  
    Orders[Ship Mode] = "Standard Class" && Orders[Shipping Time  
(Days)] <= 5, "Met",  
    "Not Met"  
)
```

Shipping Compliance Thresholds

Shipping compliance was defined as follows, where t is the shipping time in days:

- **Same Day** $\rightarrow t = 0$
- **First Class** $\rightarrow t \leq 1$
- **Second Class** $\rightarrow t \leq 3$
- **Standard Class** $\rightarrow t \leq 5$

Any order exceeding these thresholds for its respective ship mode was marked as “Not Met.” This classification was used throughout the analysis to assess performance against customer delivery expectations.

Measures and KPIs

To support detailed analysis and visualisation, several DAX measures were created to quantify key performance indicators:

% Shipping Obligations Met

```
% Shipping Obligations Met =  
COALESCE(  
    DIVIDE([Shipping Obligation Met], COUNTROWS(Orders), 0),  
    0  
)
```

This measure calculates the percentage of total orders where the delivery met the company's stated shipping timeline. It divides the number of "on-time" orders by the total number of orders. COALESCE ensures the result defaults to 0 if there are no orders.

Average Order Value

```
Average Order Value = [Total Sales Amount]/[Total Orders]
```

Indicates the average revenue generated per order by dividing the total sales amount by the total number of orders. This helps assess customer spending per transaction.

Average Quantity per Order

```
Average Quantity per Order =  
DIVIDE(  
    SUM(Orders[Quantity]),  
    DISTINCTCOUNT(Orders[Order ID]),  
    0  
)
```

Calculates the average number of units sold per order, providing insight into purchasing behaviour, e.g. whether customers typically buy in bulk or make smaller purchases.

Average Shipping Time (Days)

Average Shipping Time (Days) = AVERAGE(Orders[Shipping Time (Days)])

Computes the mean number of days it took for orders to be delivered. Useful for evaluating delivery speed across different shipping methods and time periods.

Profit

Profit = SUM(Orders[Profit])

Sums the profit generated from all orders in the dataset. This can be compared against sales to assess profitability by segment or product category.

Shipping Obligation Met

```
Shipping Obligation Met =  
CALCULATE(  
    COUNTRROWS(Orders),  
    FILTER(  
        Orders,  
        (Orders[Ship Mode] = "Same Day" && Orders[Shipping Time (days)]  
        <= 0) ||  
        (Orders[Ship Mode] = "First Class" && Orders[Shipping Time  
(days)] <= 1) ||  
        (Orders[Ship Mode] = "Second Class" && Orders[Shipping Time  
(days)] <= 3) ||  
        (Orders[Ship Mode] = "Standard Class" && Orders[Shipping Time  
(days)] <= 5)  
    )  
)
```

Counts the number of orders that were delivered within the expected timeframe based on their shipping method. Forms the basis of the delivery compliance analysis.

Shipping Obligation Not Met

```
Shipping Obligation Not Met =  
CALCULATE(  
    COUNTRROWS(Orders),  
    FILTER(  
        Orders,  
        (Orders[Ship Mode] = "Same Day" && Orders[Shipping Time (days)]  
        > 0) ||  
        (Orders[Ship Mode] = "First Class" && Orders[Shipping Time  
(days)] > 1) ||  
        (Orders[Ship Mode] = "Second Class" && Orders[Shipping Time  
(days)] > 3) ||  
        (Orders[Ship Mode] = "Standard Class" && Orders[Shipping Time  
(days)] > 5)  
    )  
)
```

Identifies the total number of orders that did not meet the expected delivery timeframe. This measure supports identifying issues with delivery performance.

Total Customers

Total Customers = `DISTINCTCOUNT(Orders[Customer Name])`

Calculates the number of unique customers who made purchases. Useful for segmenting customer behaviour and assessing market reach.

Total Orders

Total Orders = `DISTINCTCOUNT(Orders[Order ID])`

Returns the total number of individual orders placed, forming the denominator for several other KPIs.

Total Quantity Sold

Total Quantity Sold = `SUM(Orders[Quantity])`

Measures the total number of product units sold across all orders. Helps assess demand and product volume.

Total Returns

Total Returns = `DISTINCTCOUNT>Returns[Return ID])`

Counts the total number of recorded product returns. While returns could not be matched directly to specific orders, this aggregate count offers a general indication of return activity.

Total Sales Amount

Total Sales Amount = `SUM(Orders[Sales])`

Sums the revenue generated from all transactions, forming a key metric for overall sales performance.

Dashboard Overview

The interactive Power BI dashboard provides a high-level summary of sales performance, customer behaviour, and delivery fulfilment across the U.S. It is designed to allow stakeholders to monitor key metrics and filter insights by year, product category, and shipping type. The dashboard is composed of four KPI cards, several slicers, and a series of visualisations that support trend analysis and operational insights. Figure 1 shown below is a snapshot of the Sales and Delivery Performance Dashboard:

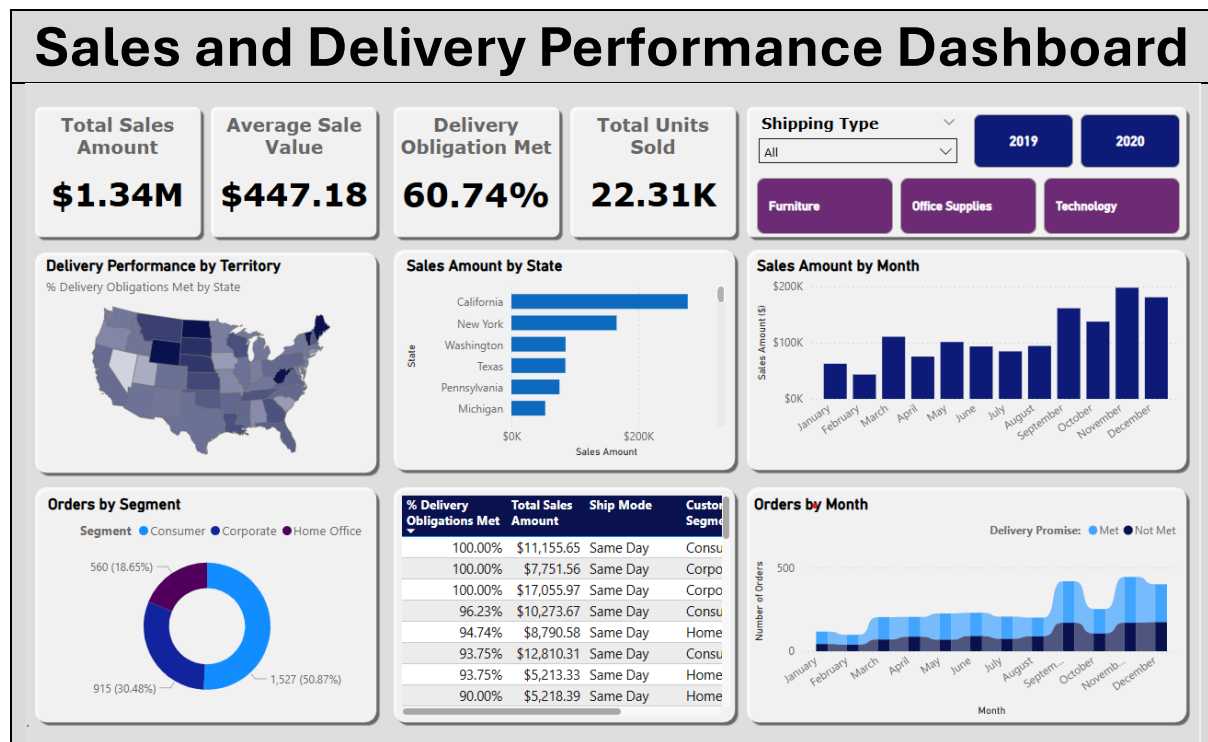


Figure 1 Snapshot of Sales and Delivery Performance Dashboard

KPI Cards

1. Total Sales Amount

Displays the total revenue generated from all recorded transactions. In the current snapshot, the dataset has produced **\$1.34M** in sales.

2. Average Sale Value

Shows the average amount spent per order, calculated as total sales divided by total orders. This KPI provides insight into customer purchasing patterns, currently averaging **\$447.18** per order.

3. Delivery Obligation Met

Represents the percentage of orders delivered within the promised shipping timeframes, based on the logic applied per shipping mode. Currently, **60.74%** of orders met their delivery obligations, offering a key indicator of shipping reliability.

4. Total Units Sold

Summarises the number of product units sold across all orders, with **22.31K** units recorded in this snapshot.

Interactive Filters (Slicers)

1. Shipping Type (Dropdown)

Allows users to filter the entire dashboard by shipping mode, such as Same Day, First Class, Second Class, or Standard Class. This is useful for isolating delivery performance or sales trends by service level.

2. Year Filter (Button Slicers)

Enables filtering the data by transaction year, specifically 2019 and 2020. Users can switch between these periods to view how KPIs and performance indicators vary over time.

3. Product Category (Button Slicers)

Users can toggle between the three main product categories: Furniture, Office Supplies, and Technology. This supports focused analysis on performance by product line.

Visualisations

1. Delivery Performance by Territory (Shape Map)

A choropleth-style map shows the percentage of delivery obligations met by state. Darker regions indicate stronger shipping performance, providing a geographic overview of delivery reliability across the U.S.

2. Sales Amount by State (Bar Chart)

Highlights the top-performing states in terms of sales revenue. California, New York, and Washington lead in total sales, aiding in regional market analysis.

3. Sales Amount by Month (Column Chart)

Displays monthly sales totals to identify seasonal trends or peak sales periods. December shows the highest volume, suggesting strong holiday-driven performance.

4. Orders by Segment (Donut Chart)

Breaks down customer orders by market segment: Consumer, Corporate, and Home Office. Consumer customers make up the majority (**50.87%**) of orders, followed by Corporate and Home Office segments.

5. % Delivery Obligations Met (Table)

A detailed table lists combinations of shipping mode, customer segment, and product category, along with the associated percentage of on-time deliveries and total sales. This enables a granular view of delivery performance across different segments, service levels, and product types, helping identify where fulfilment targets are being met or missed.

6. Orders by Month (Ribbon Chart)

A ribbon chart visualises how order volumes fluctuate month-to-month, while also distinguishing between orders that met delivery expectations and those that didn't. This layered approach would potentially help track both operational and customer satisfaction metrics over time.

Recommendations

The analysis has highlighted several delivery performance patterns that warrant attention. While overall shipping obligation compliance stands at approximately **60.7%**, key disparities between shipping modes and regions suggest opportunities for operational improvement.

- **Investigate First Class Fulfilment Delays**

First Class deliveries exhibited the lowest compliance rate at just 26%, despite being the second-fastest and a premium shipping option. This contrasts sharply with Same Day deliveries, which achieved 95% compliance. This suggests a systemic issue in how First Class orders are handled, potentially due to fulfilment delays, bottlenecks in processing, or misalignment between logistics partners and expected service levels. A thorough process review is recommended to identify root causes and improve delivery reliability for this segment.

- **Improve Visibility into Order Processing Timelines**

The dataset lacks an "Order Processed Date," making it difficult to evaluate the full delivery lifecycle. Capturing when an order is processed, distinct from when it is placed would allow for clearer insight into lead times and help identify delays in internal operations versus shipping carriers.

- **Assess Low-Volume, High-Compliance States with Caution**

States like Wyoming, North Dakota, West Virginia, Maine, and Vermont show 100% compliance, but each had fewer than 5 orders. These results are not statistically significant and should not drive decision-making without larger sample sizes.

- **Prioritise Improvements in High-Sales Regions**

The Western region, which generated over \$440,000 in sales, had a moderate compliance rate of 60%. Notably, First Class deliveries in this region had only 31% compliance, while Same Day reached 89%. Focused improvements in this high-performing region, especially around First Class logistics, could yield substantial business value and improve customer retention.

- **Consider Customer Impact and Brand Perception**

The poor performance of premium delivery options may negatively impact customer satisfaction and long-term loyalty. Communicating realistic delivery expectations at

checkout or offering proactive compensation (e.g., free upgrades or discounts) could help mitigate dissatisfaction.

Future Work

To strengthen future iterations of this analysis and build on the insights uncovered, the following actions are recommended:

- **Incorporate Order Processing and Carrier Data**
Adding fields for when orders are processed and shipped by carriers will enable a complete view of the order-to-delivery journey.
- **Account for Non-Working Days and Holidays**
Delivery obligations do not currently adjust for weekends or public holidays. Enhancing the logic to accommodate these variables will yield more accurate assessments of performance.
- **Segment Analysis by Product Type and Profitability**
While sales volume is analysed, segmenting delivery compliance by **high-margin products** could help prioritise logistics improvements where they are most profitable.
- **Monitor Delivery Trends Over Time**
Expanding the analysis across more years, or on a rolling basis, can help assess if delivery performance is improving, worsening, or affected by seasonal trends.
- **Customer Feedback Integration**
If available, correlating delivery performance with customer satisfaction ratings or complaints can validate the operational impact on user experience.