PROJECT REPORT: AMAZON SHIPPING ANALYTICS USING TABLEAU

DATE

OCTOBER 24, 2024

PREPARED BY

SUBHASRI DATTA

23MCA20053
MASTER OF COMPUTER APPLICATIONS
CHANDIGARH UNIVERSITY

AKNOWLEDGEMENT

I would like to express my deepest gratitude to ***Ms. Amanjot Kaur***, my subject teacher, for her invaluable guidance, continuous support, and encouragement throughout the course of the Amazon Shopping Analysis project. Her insightful feedback and expertise have been crucial in shaping this work.

I am also deeply indebted to ***Dr. Abdullah***, the Head of Department, for his support and for providing a conducive academic environment that fostered the successful completion of this project.

A special thanks to ***Chandigarh University*** for providing the resources and platform necessary for academic growth and the completion of this research.

Lastly, I extend my appreciation to ***Google*** for the availability of its resources, tools, and data that significantly contributed to the research process.

Table of Contents

- **Introduction**
- **Data Overview**
- **Objectives of Analysis**
- **Data Preprocessing & Assumptions**
    - Data Cleaning
    - Data Filtering
    - Assumptions
    - Screenshot (Data Preprocessing and Filters)
- **Visualization and Analysis**
- **5.1 Shipping Status Overview**
    - Description
    - Screenshot (Total Orders & Shipping Status Summary)
    - Key Insights
- **5.2 Product Category Analysis**
    - Description
    - Screenshot (Product Category Analysis Bar Chart)
    - Key Insights
- **5.3 Daily Shipping Performance**
    - Description
    - Screenshot (Daily Shipping Performance Line Chart)
    - Key Insights
- **5.4 Regional Shipping Challenges**
    - Description
    - Screenshot (Map Visualization for Unshipped Orders)
    - Key Insights
- **Key Insights and Business Implications**
- **Conclusion and Recommendations**
    - Conclusion
    - Recommendations
- **Appendices**
    - Screenshots (All Tableau Visualizations)
    - Data Dictionary

Introduction

Title: Amazon Shipping Analytics (July2018-May 2020)

Objective: The primary goal of this project is to conduct an in-depth analysis of Amazon's shipping data from April 2020 to uncover patterns and insights that can inform better decision-making. The focus is on understanding the performance of the shipping process, specifically examining the rate of shipped versus unshipped orders, identifying product categories experiencing shipping issues, and analyzing regional shipping inefficiencies. The insights derived from this analysis aim to help Amazon optimize its logistics and enhance customer satisfaction.

Scope: The scope of the analysis includes:

- Total number of orders placed during April 2020.
- Breakdown of shipping status, focusing on shipped versus unshipped orders.
- Identification of product categories that experienced higher shipping failure rates or delays.
- Examination of geographical regions where shipping delays or inefficiencies occurred.
- Daily shipping trends to understand fluctuations in shipping performance throughout the month.

Challenges Addressed:

- Shipping delays due to logistical constraints.
- High rates of unshipped orders in specific categories or regions.
- Regional inefficiencies in the supply chain.

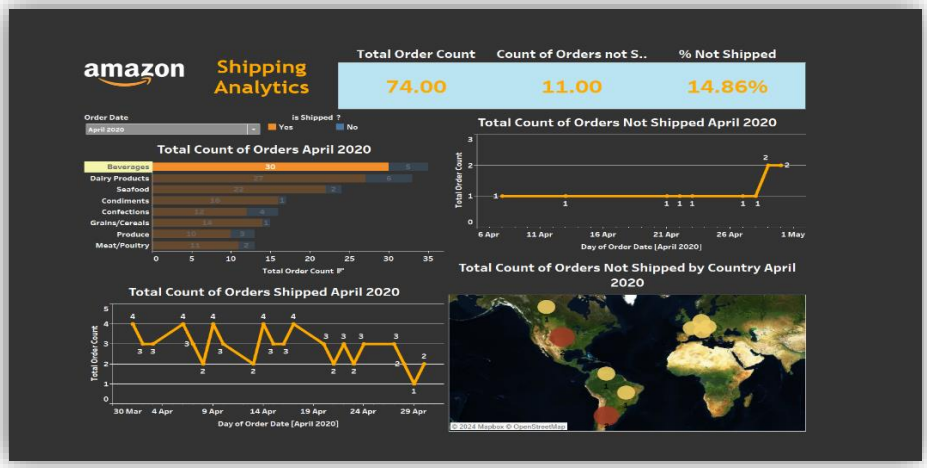
Tools Used: Tableau has been employed for data visualization to create interactive dashboards and charts that highlight key insights. The data analysis focuses on Amazon's shipping data for April 2020, enabling a detailed breakdown of trends and performance metrics. Various Tableau features such as pivot tables, filters, and calculated fields were used to manipulate and visualize the data.

Data Overview

Data Source: The dataset used for this analysis is either a simulated dataset or publicly available data that closely mirrors Amazon's actual transaction and shipping data for April 2020. This dataset is designed to provide realistic insights into Amazon's shipping performance during this period.

Key Columns:

- **Order Date**: The date when each order was placed, which helps in tracking daily shipping trends and understanding peak order periods.
- **Product Category**: The category of the product being ordered (e.g., Beverages, Dairy Products, Electronics), allowing for the identification of categories with high or low shipping performance.
- **Shipping Status**: This binary column indicates whether the order was shipped successfully or not. Entries are marked as "Yes" for shipped orders and "No" for unshipped ones, which helps in calculating the overall shipping success rate.
- **Country/Region**: The geographical location of the customer placing the order. This column is critical for analyzing regional shipping inefficiencies and understanding logistical challenges across different areas.



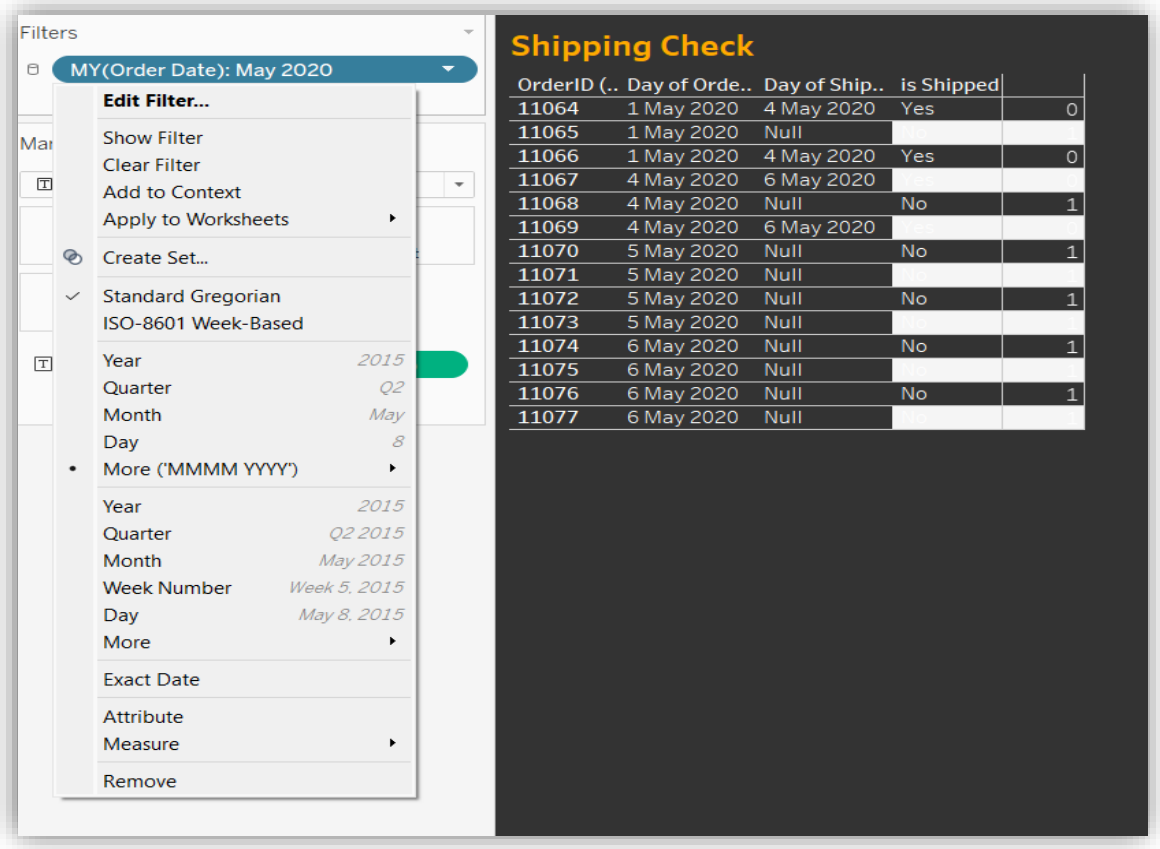
Objectives of Analysis

1. **Assess Shipping Efficiency**:
    - The analysis aims to evaluate the overall efficiency of Amazon's shipping process by comparing the total number of orders that were shipped against those that remained unshipped. This comparison will provide insights into the company's ability to fulfill orders during April 2020, helping to identify any bottlenecks in the logistics chain.
    -
2. **Identify Product Categories with Shipping Challenges**:
    - One key focus is to examine the top-selling product categories (e.g., Beverages, Dairy Products, Electronics) and assess their shipping performance. This involves identifying categories where a significant number of orders faced shipping delays or failures, potentially highlighting supply chain or inventory management issues specific to certain product types.
    -
3. **Track Daily Shipping Performance**:
    - The analysis will break down shipping data by day to identify trends in shipping performance over the course of April 2020. By tracking daily shipping metrics, the project aims to uncover specific days when shipping failures were higher, which could be linked to external factors such as holidays, peak demand periods, or logistical constraints
    -
4. **Map Regional Shipping Challenges**:
    - Another critical objective is to map shipping performance across different geographical regions. This involves identifying regions or countries with higher percentages of unshipped orders, which could indicate issues with regional fulfillment centers, shipping routes, or local restrictions. Understanding these regional disparities can help Amazon optimize its delivery operations in underperforming areas.

Data Preprocessing & Assumptions

1. **Data Cleaning**:
    - Ensuring the dataset is clean and ready for analysis is a crucial step. All missing values, especially in critical columns like "Order Date" and "Shipping Status," were handled by either removing incomplete records or imputing values where appropriate.
    - Special attention was given to date formatting. Order dates were standardized to a consistent format (e.g., DD/MM/YYYY), ensuring accurate filtering and analysis of data within the April 2020 timeframe.
2. **Data Filtering**:
    - To maintain the accuracy and relevance of the analysis, only orders placed during April 2020 were considered. This filtering step ensured that no outliers, such as orders from other months, or incomplete or test transactions, were included in the final analysis.
    - The dataset was further filtered to remove orders with missing or invalid shipping statuses, focusing solely on valid shipped and unshipped data points for clear insight into performance.
3. **Assumptions**:

    - For the purpose of this analysis, it is assumed that any order marked with a shipping status of “No” indicates that the order was either delayed or canceled. While the dataset doesn’t specify the exact reason for an unshipped order, this assumption helps categorize and quantify shipping failures in the analysis.
    - It is also assumed that the dataset is representative of Amazon’s actual operations for April 2020, even though the data may be simulated or publicly available.
    - The dataset, whether simulated or publicly available, is assumed to be representative of Amazon's actual shipping operations for April 2020, enabling accurate insights despite potential differences in the source of the data.



Visualization and Analysis

Shipping Status Overview

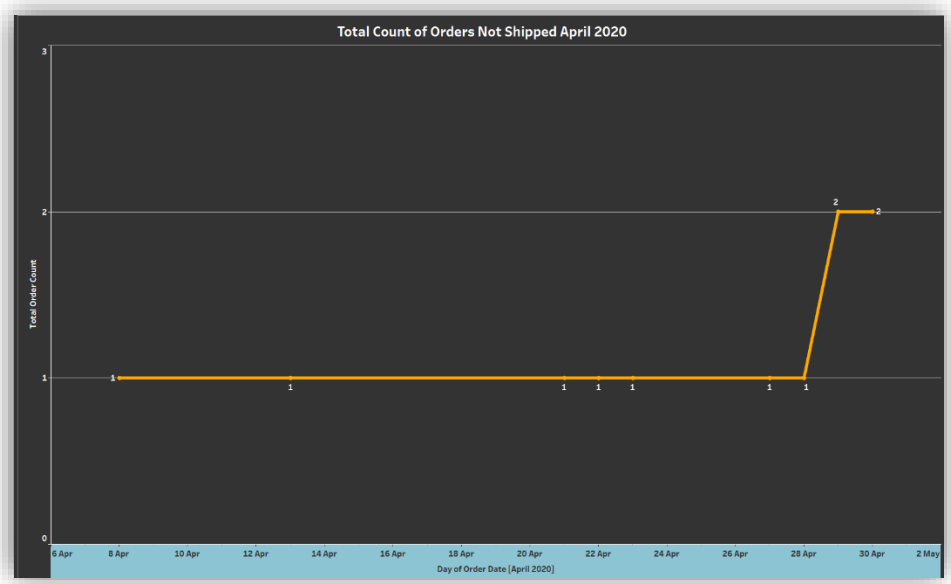
Description:
This visualization provides an overview of Amazon's shipping performance in April 2020. It summarizes the total number of orders, the number of unshipped orders, and the percentage of unshipped orders. This high-level insight gives a quick snapshot of shipping efficiency.

- **Total Order Count**: 74
- **Unshipped Orders**: 11 (14.86% unshipped)

Key Insights:

- Around 15% of all orders were not shipped, signaling potential bottlenecks in the shipping process.

- Identifying the root cause behind these unshipped orders could help in mitigating delays or cancellations.



Product Category Analysis

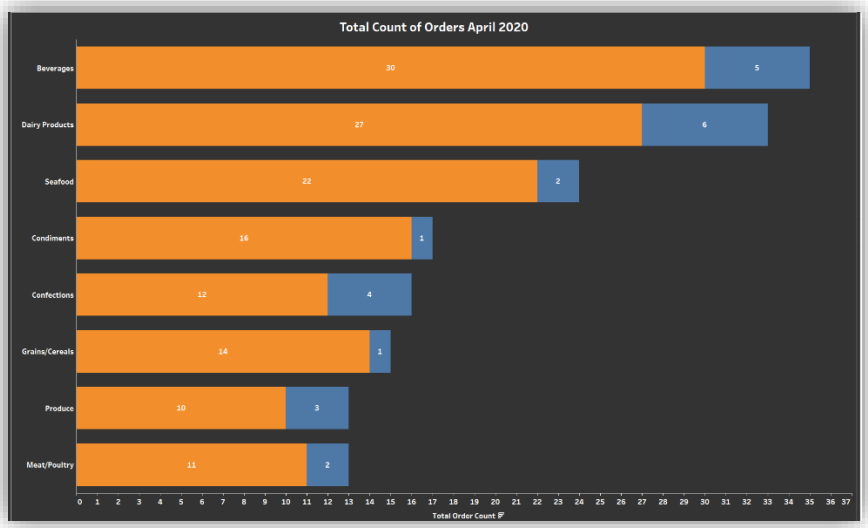
Description:
This bar chart visualizes the total number of orders for each product category, broken down by whether they were shipped or unshipped. The goal is to identify which product categories experience the most shipping challenges.

- **Top Categories**: Beverages (30 orders), Dairy Products (27 orders), Seafood (22 orders), with significant unshipped orders in each category.
- **Categories with Unshipped Orders**: Beverages and Dairy Products see higher counts of unshipped orders, contributing to overall shipping inefficiency.

Key Insights:

- Beverages, Dairy Products, and Seafood are the top three categories by order count, but they also face notable shipping issues.

- Beverages alone contributed 5 unshipped orders, making it the category with the highest shipping failures, indicating possible supply chain or inventory issues for this category.



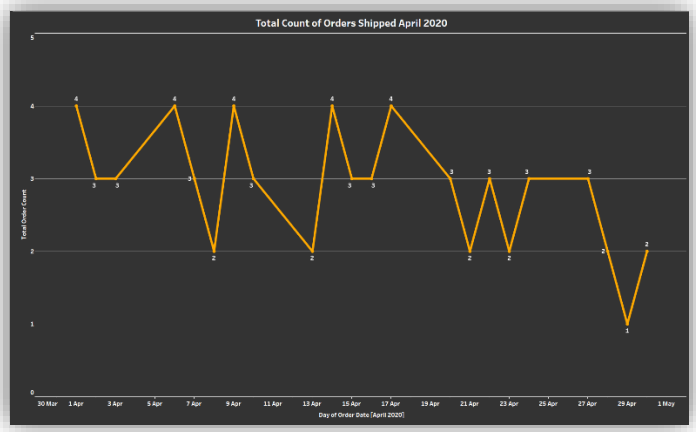
Daily Shipping Performance

Description:
A line chart visualizes the daily performance of shipped and unshipped orders, providing insight into how shipping fluctuated throughout April 2020. This helps in understanding whether shipping performance was consistent or if certain days faced more challenges.

- **Shipped Orders Trend**: The daily count of shipped orders fluctuates between 2 and 4, showing variability in shipping efficiency.
- **Unshipped Orders Trend**: Unshipped orders occur sporadically, with a noticeable increase towards the end of the month.

Key Insights:

- There is no clear trend of improvement in shipping performance throughout the month.
- Towards the end of April, the number of unshipped orders spikes, suggesting potential supply chain or logistical disruptions during that period.



Regional Shipping Challenges

Description:
This map chart highlights the geographical distribution of unshipped orders, allowing us to pinpoint regions that face more shipping issues. The focus is on identifying areas where logistical or infrastructure challenges may be hindering Amazon's ability to ship orders efficiently.

Key Insights:

- Significant geographical disparities in shipping performance are evident, with South America and Southeast Asia facing the highest number of shipping failures.
- These regions may experience logistical challenges, such as poor infrastructure or supply chain disruptions, contributing to the delays.

Key Insights and Business Implications

Unshipped Orders:

- **Insight**: 15% of total orders remain unshipped, indicating that a significant portion of customer demand is unmet.
- **Business Implications**: This level of unshipped orders can severely impact customer satisfaction, leading to potential frustration and decreased loyalty. Customers may choose competitors if they consistently face shipping delays. It's crucial for Amazon to address these issues promptly to enhance customer experience and retention rates.

Product Categories at Risk:

- **Insight**: High-demand categories such as Beverages and Dairy Products are experiencing notable shipping challenges, with a substantial number of unshipped orders.
- **Business Implications**: The prevalence of unshipped orders in these categories underscores the need for improved inventory management and supply chain processes. Implementing better forecasting tools, enhancing supplier relationships, and optimizing stock levels could help mitigate these issues and ensure that high-demand products are consistently available for shipping.

Daily Trends:

- **Insight**: There is a distinct spike in unshipped orders towards the end of April, suggesting possible capacity issues or stock shortages during this period.
- **Business Implications**: This trend indicates that Amazon may need to assess its operational capacity and stock management practices. Potential solutions could include scaling up warehouse resources during peak periods, enhancing real-time inventory.

Regional Bottlenecks:

- **Insight**: Certain regions, particularly in South America and Southeast Asia, are experiencing disproportionate shipping delays.
- **Business Implications**: This disparity highlights the necessity for region-specific shipping solutions and improvements. Investing in local distribution centers, enhancing partnerships with local logistics providers, and tailoring shipping strategies to address regional challenges can improve overall shipping performance and reduce delays. Additionally, understanding regional market dynamics can help Amazon better align its services to meet customer needs effectively.

Conclusion and Recommendations

Conclusion:
The analysis of Amazon's shipping performance for April 2020 reveals significant areas for improvement, particularly regarding high-demand product categories and specific regions facing logistical challenges. With 15% of total orders unshipped, it is evident that inefficiencies exist within the shipping process, affecting customer satisfaction and overall business performance. Addressing these challenges is crucial for enhancing operational efficiency and maintaining customer loyalty.

Recommendations:

Improve Inventory Management:

- **Focus on Top-Selling Categories**: It is essential to optimize stock levels for high-demand categories like Beverages and Dairy Products. This can be achieved by:
  - Implementing advanced demand forecasting tools that analyze historical sales data and seasonal trends to predict future demand accurately.
  - Collaborating closely with suppliers to ensure timely replenishment of stock and avoid shortages.

Regional Shipping Enhancements:

- **Address Bottlenecks**: To tackle shipping delays in underperforming regions:
  - Strengthen local supply chains by investing in regional distribution centers that can expedite order fulfillment and delivery.
  - Consider partnerships with reliable third-party logistics (3PL) providers to enhance shipping capabilities and leverage their local expertise and networks.
  - Analyze regional shipping data to identify specific challenges and tailor shipping strategies to meet local demands and logistics constraints.

Monitor Trends Regularly:

- **Proactive Approach**: Establish a framework for the continuous monitoring of shipping performance metrics, including daily order counts, shipping status, and category-specific analyses. This can be achieved by:
  - Utilizing dashboards and analytics tools to track key performance indicators (KPIs) in real-time.
  - Implementing alerts for anomalies in shipping performance, allowing the team to address potential issues before they escalate.
  - Conducting regular reviews of shipping performance data to identify trends and develop strategies to mitigate future risks.

These recommendations aim to create a more efficient shipping process that can adapt to changing market demands and enhance customer satisfaction. By focusing on inventory management, regional shipping enhancements, and proactive monitoring, Amazon can significantly improve its overall shipping performance and better meet customer expectations.