ANALYZING SALES DATA TO ENHANCE BUSINESS PERFORMANCE: A CASE STUDY ON AMAZON

Akshay Shaju x22152563
Caroline Vincent x22153926
Karthika Nair x22105522

1. ABSTRACT

This report presents a detailed analysis of sales data for Amazon, one of the largest e-commerce companies globally. The objective of this study is to explore the relationship between sales data, customer behaviour, and business performance to identify opportunities for improvement.

The research utilizes a multi-step approach, starting with data cleaning and exploratory data analysis using Python to gain initial insights. The cleaned data is then uploaded into a PostgreSQL database, enabling advanced data manipulation and analysis. Additionally, a Customer Relationship Management (CRM) system, specifically Salesforce, is implemented to further analyse customer interactions and behaviours. Finally, the findings are visualized using Power BI to provide meaningful insights and facilitate data-driven decision-making.

The analysis of the sales data aims to uncover valuable patterns, trends, and correlations. By understanding customer preferences, purchasing patterns, and market dynamics, Amazon can enhance its business operations and improve customer satisfaction. The report concludes with key findings from the analysis and offers recommendations on leveraging these insights. The goal is to optimize business strategies, enhance customer experiences, and drive sustainable growth for Amazon.

2. INTRODUCTION

The modern business landscape is characterized by the availability of vast amounts of data generated through various channels. Organizations are increasingly leveraging this data to gain insights into customer behaviour, improve operational efficiency, and drive business growth. In this report, we present a comprehensive analysis of sales data for Amazon, one of the world's leading e-commerce companies. By employing a combination of exploratory data analysis, database management, CRM implementation, and data visualization techniques, we aim to uncover valuable insights that can contribute to a deeper understanding of customer behaviour and drive improvements in business performance. Amazon, with its extensive product catalogue and global customer base, serves as an ideal case study for examining the potential of data analysis in optimizing business strategies.

To address this research question, we will follow a structured approach. Initially, we will utilize the Python programming language to clean the dataset, remove duplicates, and perform basic exploratory data analysis. Subsequently, we will establish a database in PostgreSQL, create tables, and upload the sales data. Through SQL queries, we will manipulate the data, enabling us to gain further insights and perform advanced analysis. Additionally, we will create an

Entity-Relationship (ER) diagram and a data dictionary to provide a comprehensive overview of the database structure.

As part of our analysis, we will also explore the implementation of a Customer Relationship Management (CRM) system using Salesforce. By uploading the sales data to Salesforce and leveraging its CRM functionalities, we can enhance customer interactions, improve customer segmentation, and tailor marketing efforts accordingly. To effectively communicate our findings, we will employ data visualization techniques, primarily utilizing Power BI. Through interactive visualizations and insightful dashboards, we will present the key findings from our analysis, enabling stakeholders to make data-driven decisions and optimize business strategies.

In conclusion, this report aims to demonstrate the significance of analysing sales data in improving business performance. By harnessing the power of exploratory data analysis, database management, CRM implementation, and data visualization, we aspire to uncover valuable insights into customer behaviour, identify trends, and support decision-making processes for Amazon. The subsequent sections of this report will delve into the methodology, data analysis, findings, and recommendations derived from our comprehensive study.

3. OBJECTIVE

The objective of this project is to analyse Amazon's sales data and customer insights to inform strategic decision-making and enhance business performance. The primary goal of this project is to extract meaningful insights from the dataset, identify key trends, and develop actionable recommendations for Amazon. By leveraging advanced data analysis techniques, including data cleaning, exploratory data analysis, SQL querying, and CRM implementation, the report aims to provide a comprehensive understanding of sales patterns, customer behaviour, and market dynamics. The insights derived from this analysis will assist Amazon in optimizing marketing campaigns, improving inventory management, and implementing customer-centric strategies to drive growth and foster long-term customer loyalty.

4. BUSINESS BACKGROUND

Amazon is one of the largest e-commerce players in India and is currently second only to Flipkart in terms of market share. However, Amazon has been steadily gaining ground on its competitor, and as of 2021, the two companies are almost neck and neck.

Fig1. American Technology Giant

The business idea is of an e-commerce platform that offers a wide range of products to customers worldwide with key aspects like product selection, customer experience, pricing and discounts, personalisation and customer reviews and ratings.

Amazon provides an extensive selection of products across various categories, including electronics, apparel, home goods, books, toys, and more. The product range includes items sourced directly from Amazon as well as those sold by third-party sellers through the Amazon marketplace. It operates a vast network of fulfilment centres worldwide to store, pack, and ship products efficiently and offers multiple delivery options, including standard shipping, expedited shipping, and Amazon Prime's fast and free shipping service. enables third-party sellers to sell their products through its platform, creating a marketplace that expands the selection available to customers.

Third-party sellers can utilize Amazon's infrastructure, logistics, and customer service capabilities to reach a broader customer base. It incorporates customer reviews and ratings on product listings, allowing customers to share their experiences and help others make informed purchase decisions. Reviews contribute to building trust and credibility within the Amazon retail ecosystem.

According to the dataset that we have chosen, the sales channel in viewpoint here is majorly www.amazon.in and the overall sales performance of Amazon has generated a revenue of Rs. 81,703,282 thus making it the most successful and profitable channel when compared to non-Amazon sales channel. The sales trend shows that the dataset is spread over the year 2022.

The business involves different shipping services for the orders which is tracked by the shipping service level variable levels. Once the products are shipped, the status of the couriers are tracked and recorded under the Courier Status column. Any delays or cancellations made corresponding to an order is recorded under the same variable.

Shipping performance: Assess the "Courier Status" column to track the status of shipments and identify any issues or delays in the shipping process. The fields SKU and ASIN help in analysing and calculating the quantity and revenue generated by each product where each product group is categorized, and then stored in the category variable.

Analysis of the "Ship City," "Ship State," "Ship Postal Code," help in visualising the geographical distribution of sales. The tracking of such details has deemed to be successful for conducting the best analysis through visualisations of the data.

4.1.SIX SIGMA

Six Sigma is a methodology that aims to improve business processes by reducing defects and variations in the output of those processes. In this project amazon sales have also noted the potential benefits of incorporating Six Sigma methodologies into its operations. One of the key principles of Six Sigma is the concept of "sigma levels," which represent the number of defects per million opportunities (DPMO) in each process. The higher the sigma level, the lower the number of defects in the process. The DMAIC methodology of Six Sigma can be applied to improve Amazon sales by following the steps below:

Define: In this step, the problem statement and the project goals are defined. For example, the problem could be low sales revenue or high return rates. The goal could be to increase sales revenue or reduce return rates.

Measure: In this step, data is collected and analysed to establish a baseline performance. The key performance indicators (KPIs) are identified, and the current performance is measured

against the target performance. For example, the KPIs for Amazon sales could be sales revenue, order fulfilment rate, and return rate.

Analyse: In this step, the root cause of the problem is identified using statistical tools and techniques. The data is analysed to determine the factors that contribute to the problem. For example, the analysis could reveal that the high return rate is due to incorrect sizing information on the product page.

Improve: In this step, solutions are developed and tested to address the root cause of the problem. For example, the product page could be updated with accurate sizing information. The solution is then implemented on a small scale to determine its effectiveness.

Control: In this step, the solutions are monitored and measured to ensure that they are effective in the long term. The performance is continuously monitored, and any deviations from the target are addressed.

By applying the DMAIC methodology, the organisation can identify and address the root cause of problems in sales data and continuously improve the performance.

Fig2. Six Sigma

4.2.MARKETING FUNNEL STRATEGY

Marketing funnel strategy is a concept used in marketing that describes the process through which potential customers move from being unaware of a brand to becoming loyal customers. The funnel consists of several stages, including awareness, interest, consideration, intent, and finally, purchase.

Awareness: At this stage, potential customers become aware of brand. In the context of Amazon sales, this includes visitors to our product pages or those who see our product in search results.

Interest: Once a potential customer becomes aware of the brand or product, they may begin to show interest by clicking on the product page or adding the product to their Wishlist.

Consideration: At this stage, the potential customer is actively considering the product and may be comparing it to similar products. This can be indicated by the customer spending a longer time on your product page or reading product reviews.

Intent: The potential customer has made the decision to purchase the product and is now showing intent to buy. This can be indicated by the customer adding the product to their cart or initiating the checkout process.

Evaluation: At this stage, the customer is evaluating the purchase decision by reviewing the details of the order, including shipping options and total cost.

Purchase: The customer completes the purchase and becomes a paying customer.

Overall, the marketing funnel strategy can help the organisation optimize their sales process and improve customer acquisition and retention. In Amazon sales, the strategy can be applied by using the data available in the columns of the dataset to identify areas for improvement and make data-driven decisions to improve the customer journey.

Fig3. Marketing Funnel Strategy

5. SYSTEM DESIGN

System design for this project involves planning and designing the entire system architecture, which includes the software tools, hardware, and infrastructure required to build and maintain the project. A high-level overview of the system design for the project is, In the first phase, the data is collected from various sources and stored in a raw format and used python for cleaning and data preparation. Once the data is cleaned and prepared, it needs to be transformed and loaded into the data warehouse. Postgres is used as the database management system in this project. Once the data is stored in the data warehouse, it can be utilized by the CRM system to improve customer interactions. This stage involves integrating the CRM with the data warehouse to extract customer-related information. Once the data is available in the CRM system, it can be used to generate reports, analyse customer behaviour, and gain insights into

business operations. Power BI is used for visualization in this project. The final stage involves maintaining the system and ensuring that it continues to work efficiently and effectively. This includes monitoring the performance of the system and making changes as necessary to ensure that it continues to meet the needs of the business.

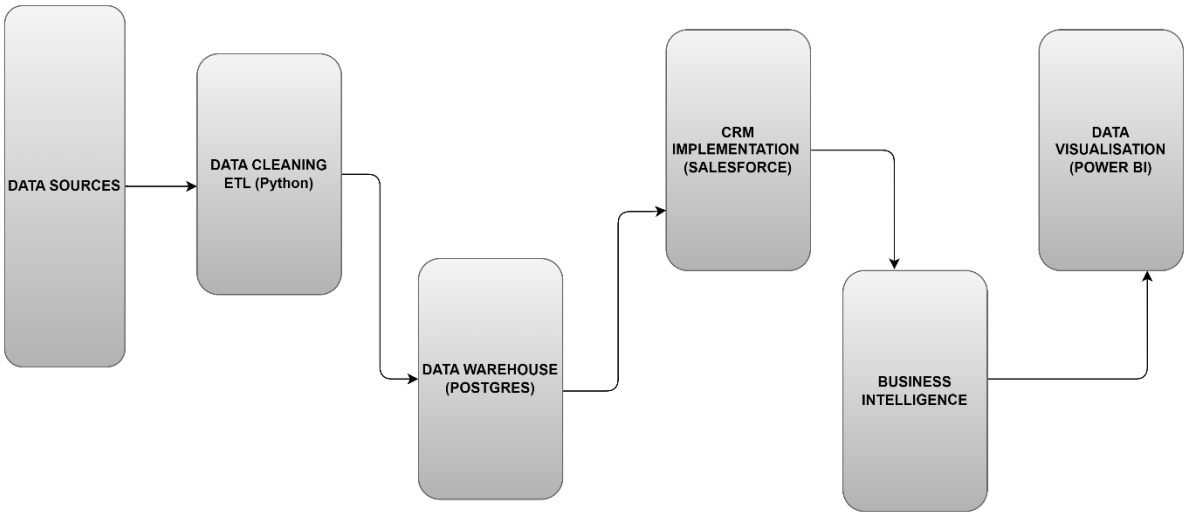


Fig4. Architecture Diagram

5.1. DATBASE DESIGN

Database design is the process of creating a database structure that is optimized for the efficient storage, retrieval, and management of data. The database design process involves identifying the entities and attributes that need to be stored and defining the relationships between these entities. This information is then used to create a schema or data model that defines the structure of the database.

Firstly, we need to design the schema for our data warehouse. This will involve identifying the entities and attributes that we need to store, and creating the necessary tables, views, and indexes to support our reporting needs. The schema design should be optimized for reporting, so we need to ensure that the tables are properly indexed, and that we use appropriate data types to minimize storage requirements and optimize query performance.

Next, we need to develop an ETL (Extract, Transform, Load) process to extract data from our source systems, transform it into the required format, and load it into our data warehouse. The ETL process in this project is done using python. The ETL process will typically involve the following steps:

Extract: Data is extracted from source systems such as operational databases, spreadsheets, or CSV files. The data may be extracted using APIs, database connectors, or custom scripts.

Transform: The extracted data is then transformed into the required format for the data warehouse. This may involve data cleaning, data validation, data enrichment, and data integration.

Load: The transformed data is loaded into the data warehouse, either directly into the target tables or into staging tables for further processing.

Once the data is loaded into our data warehouse, we can use it to build reports and visualizations using business intelligence tools such as Salesforce, Power BI.

Overall, the database design for this scenario involves creating a schema for our data warehouse, developing an ETL process to extract and transform data from source systems, and using business intelligence tools to analyse and report on the data. PostgreSQL is an excellent choice for this type of database design, as it provides a robust and scalable platform for storing and analysing large amounts of data.

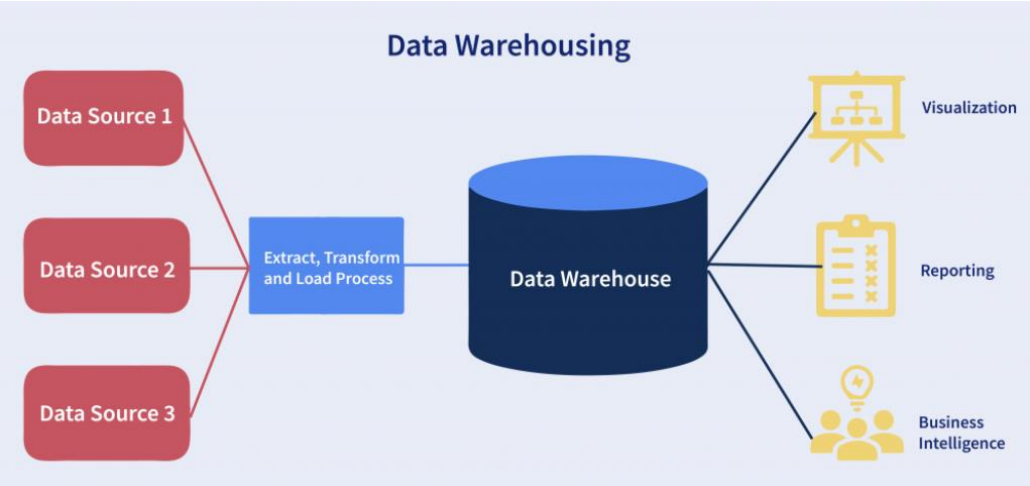


Fig5. Database Architecture

5.2. CUSTOMER RELATIONSHIP MANAGEMENT

Customer relationship management (CRM) is a strategy for managing and analysing customer interactions and data throughout the customer lifecycle, with the goal of improving customer retention, satisfaction, and ultimately driving sales growth. It involves collecting and analysing customer data to better understand their needs and preferences, and then using that information to develop targeted marketing and customer service campaigns. The goal of a CRM system is to improve customer retention and loyalty, streamline communication and sales processes, and increase revenue growth.

Fig6. Customer Relationship Management

Salesforce is a cloud-based CRM platform that offers a wide range of features and functionalities to manage customer relationships. It can help businesses to store customer data, track customer interactions, manage sales pipelines, automate marketing campaigns, and provide customer service. In the context of Amazon sales dataset, Salesforce can be used to manage customer interactions across various sales channels, monitor sales performance, track order fulfilment, and provide customer support. Storing and organizing customer information such as contact details, order history, and preferences, Tracking sales performance across different channels and products, Monitoring order fulfilment and delivery status, Creating and managing marketing campaigns targeted towards specific customer segments, Providing customer support through various channels such as email, phone, and chat Generating reports and insights on customer behaviour and sales trends.

5.3.DATA VISUALISATION

About power BI, dashboard, and visualisation general

6. DATABASE IMPLEMENTATION USING POSTGRES

In this study, Database implementation using Postgres involves creating and setting up a database system using the Postgres database management system.

6.1.DATA DICTIONARY

Data dictionary is a collection of metadata that describes the structure and contents of a database or other information system. It provides a way to document the various components of a system, including tables, columns, indexes, relationships, data types, and other data elements. The main application of a data dictionary in a database is to ensure consistency and accuracy of data by providing a standardized definition and understanding of the data elements used in the system. It serves as a reference tool for developers, analysts, and end-users to understand the structure and meaning of the data.

A data dictionary can also facilitate data integration and sharing by providing a common language for different stakeholders and systems. It can help in data modelling and design by providing a clear understanding of the data elements and relationships. Additionally, it can improve data quality and governance by enforcing standards and policies related to data management. Overall, a data dictionary plays a crucial role in the effective and efficient management of data in a database or information system. For the Amazon sales data with the given column names, a data dictionary would define each column as follows:

Order_ID: The unique identifier for each order placed on Amazon,

Date: The date on which the order was placed,

Status: The status of the order (e.g., "processing", "shipped", "delivered", etc.)

Fulfilment: The method of order fulfilment (e.g., "Fulfilled by Amazon", "Merchant fulfilled")

Sales Channel: The channel through which the order was placed (e.g., "Amazon.com", "Amazon.in", etc.)

Ship service level: The shipping service level chosen by the customer (e.g., "standard", "two-day", "overnight", etc.)

Style: The style or design of the product ordered

SKU: The unique identifier for the product sold on Amazon.

Category: The product category to which the item belongs

Size: The size or dimensions of the product

ASIN: The unique identifier assigned by Amazon for each product sold on its platform

Courier Status: The status of the shipment (e.g., "out for delivery", "delivered", "returning to seller", etc.)

Qty: The quantity of the product ordered.

Currency: The currency in which the transaction was completed

Amount: The total amount paid for the order

Ship city: The city to which the order was shipped.

Ship state: The state or province to which the order was shipped.

Ship postal code: The postal code to which the order was shipped.

Ship country: The country to which the order was shipped.

B2B: Indicates whether the order was placed by a business or individual customer.

Fulfilled by: Indicates whether the order was fulfilled by Amazon or a third-party seller.

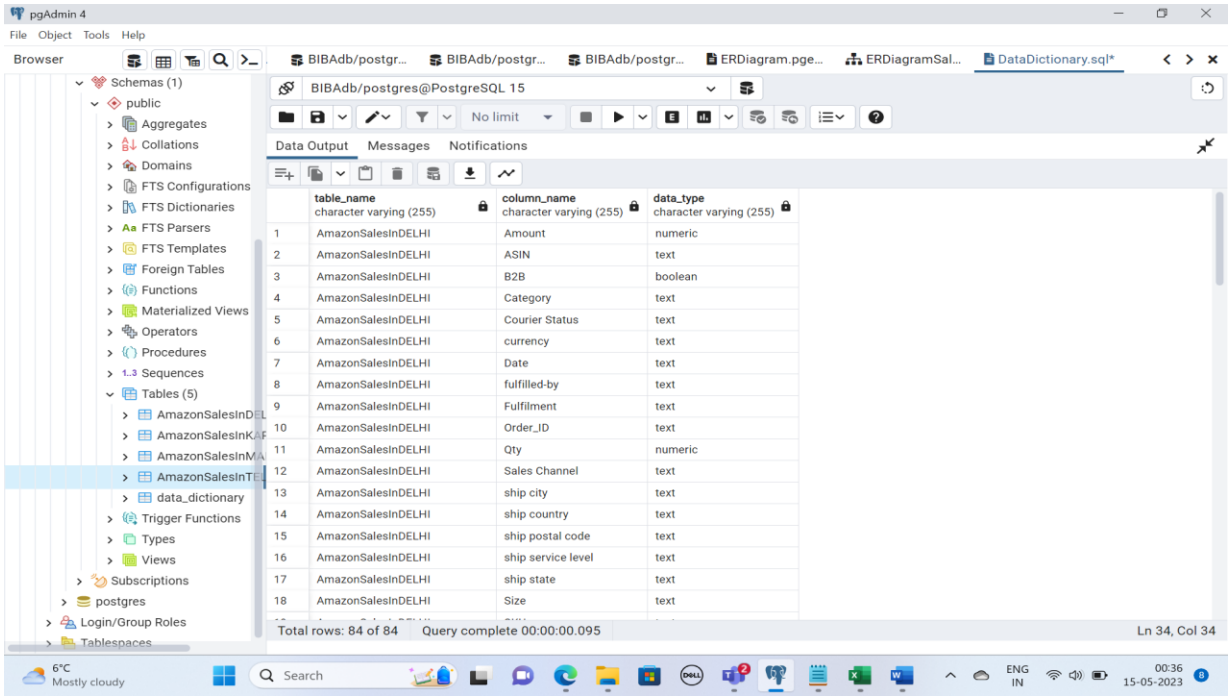


Fig7. Data Dictionary of Amazon sales

6.2.ENTITY-RELATIONSHIP DIAGRAM

An entity-relationship diagram (ERD) in PostgreSQL is a graphical representation of the data model that shows the entities, attributes, and relationships between them. It represents the logical structure of the data and provides a visual representation of the database schema.

The entity-relationship diagram (ERD) depicts the relationships between the four datasets - AmazonSalesInDELHI, AmazonSalesInTELANGANA, AmazonSalesInMAHARASHTRA, and AmazonSalesInKARNATAKA. The ERD represents the logical structure of the data and shows how the datasets are related to each other.

The ERD includes four entities, one for each dataset, and shows the attributes that describe each entity. The attributes include fields such as Order_ID, Status, Fulfilment, ship service level, and Date, among others. These attributes define the characteristics of the data stored in the datasets.

The ERD also shows the relationships between the entities. In this case, the entities are related through the Order_ID field, which is the primary key of AmazonSalesInDELHI dataset. This allows for a one-to-many relationship between the entities, with each sale record in the datasets related to one location (either Karnataka, Telangana, or Maharashtra).To represent this relationship, the ERD uses lines with arrowheads pointing to the location entity, indicating the direction of the relationship. This allows for a clear understanding of the relationship between the datasets and how they are connected.

Overall, the entity-relationship diagram provides a visual representation of the relationship between the datasets and helps in understanding the data structure. It serves as a useful reference for database administrators and developers who need to work with the data and make changes to the database.

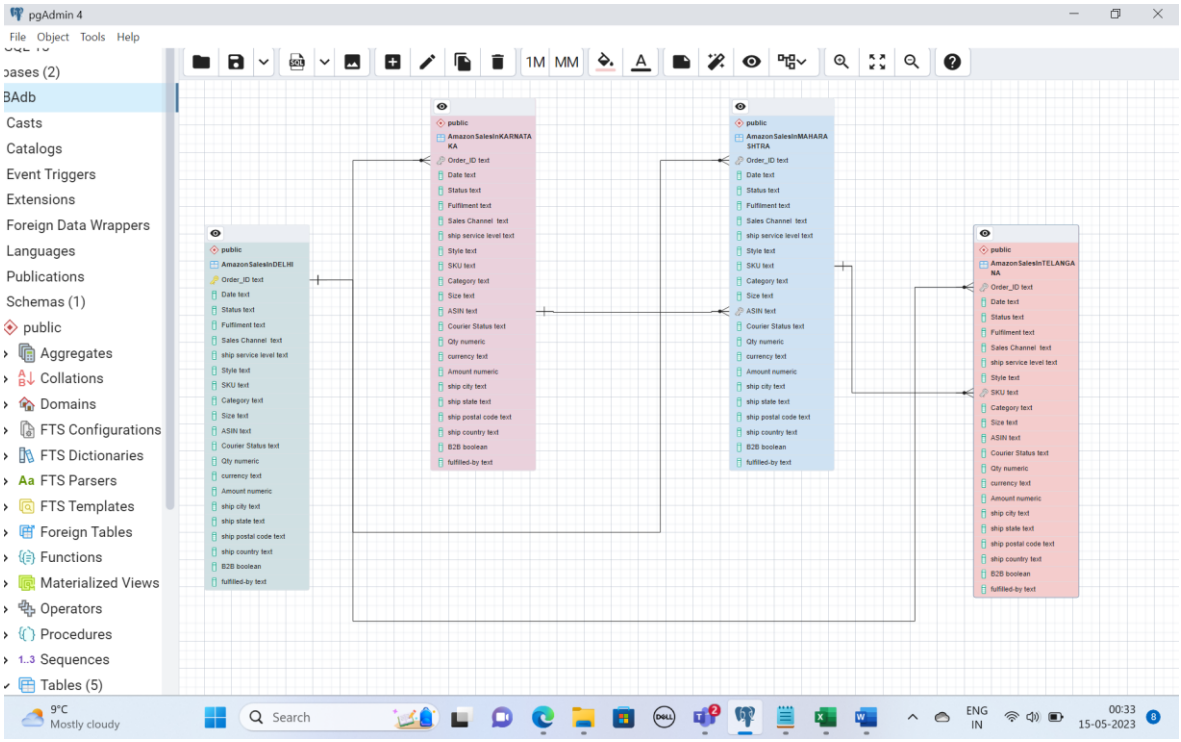


Fig8. ER Diagram

7. CRM IMPLEMENTATION USING SALESFORCE

In this study, we use Salesforce platform for the implementation of CRM. Salesforce is a cloud-based CRM platform that provides businesses with a range of tools for managing customer relationships. Salesforce can be customized to meet the specific needs of different businesses, and it can be used to manage customer data, track sales and marketing activities, automate processes, and provide analytics and insights.

Fig9. CRM Tool

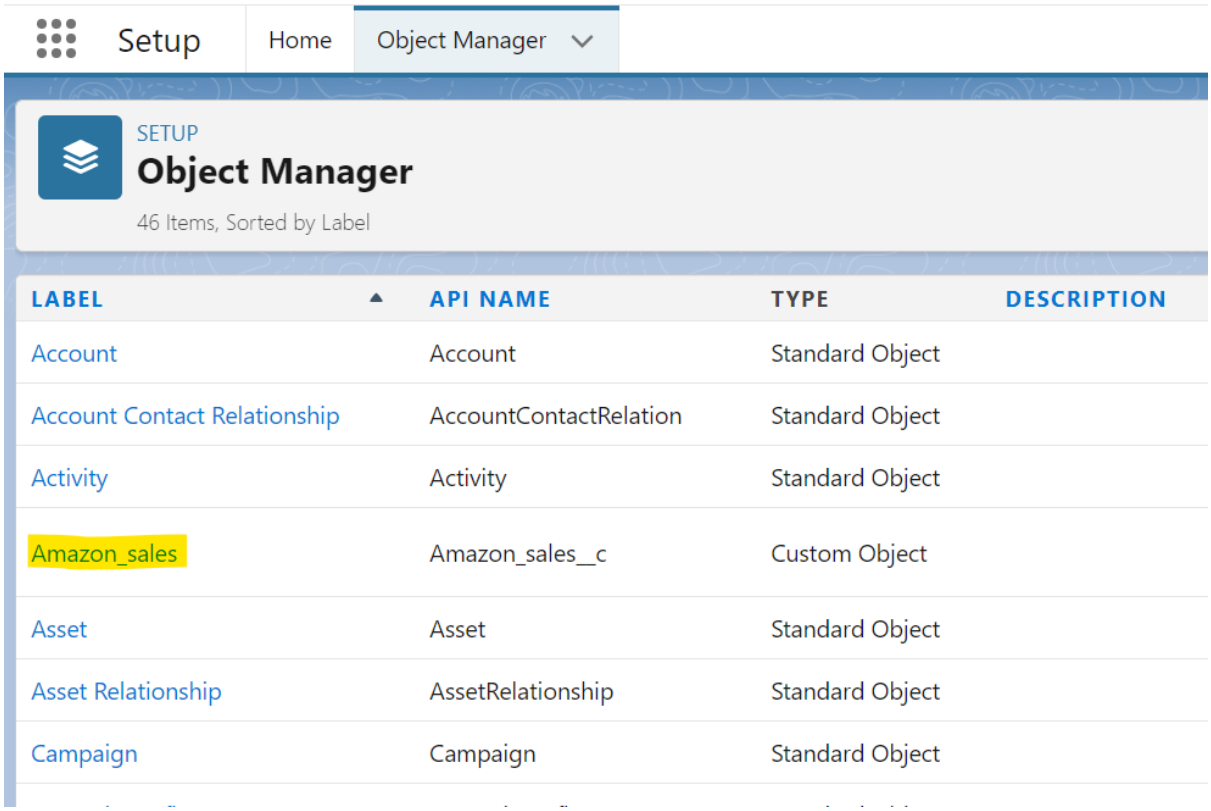


Fig10. Object Creation

DATA IMPORTED

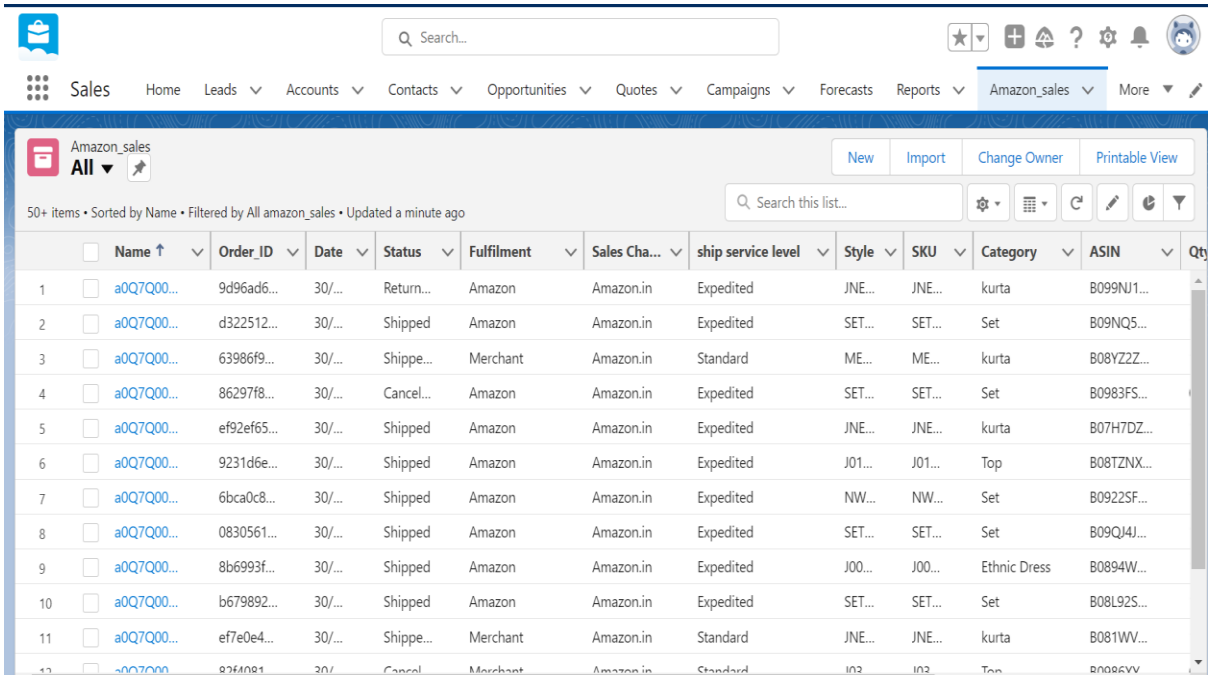


Fig11. Inserted data

7.1.EMAIL TRIGGERING

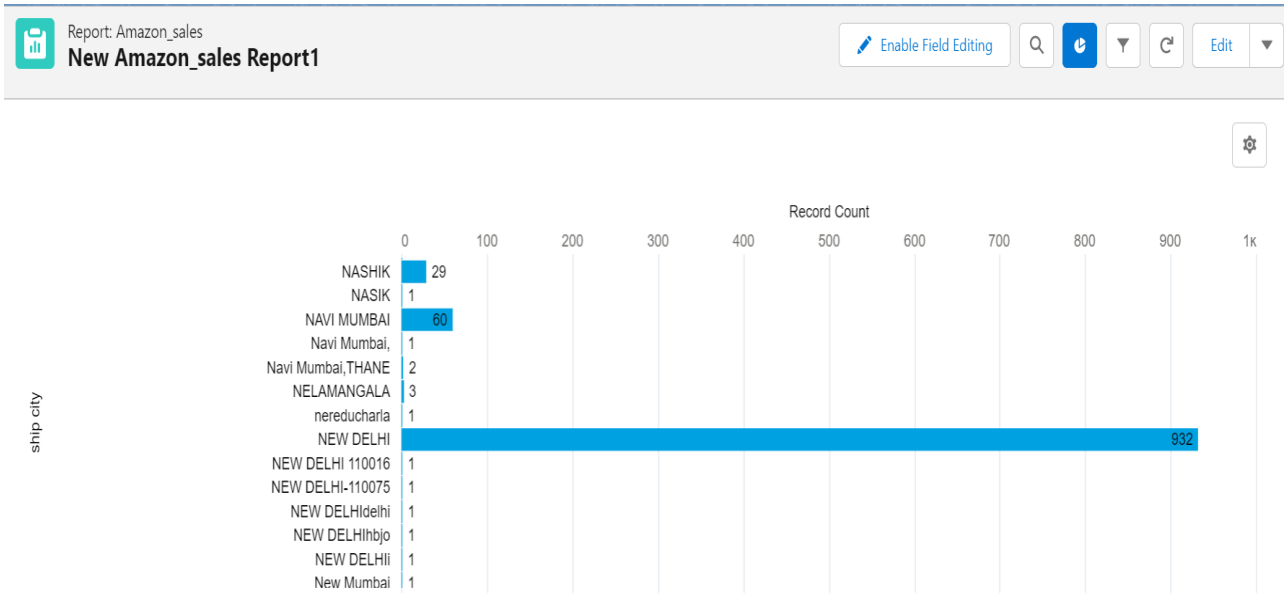
When an order is marked as "shipped" in the system, an email can be triggered to inform the customer that their order is on its way. This email can include the order details, tracking information, and an estimated delivery date. This type of triggered email can help keep customers informed and engaged with the order process, and can also help reduce the workload on customer service teams by providing customers with the information they need in a timely and automated manner. Overall, using an email triggering action like this can help improve the customer experience and increase customer satisfaction.

7.2.REPORTS

By city, which city in india

Report: Amazon_sales
New Amazon_sales Report1

ship city	Amazon_sales: Name	Category	Qty	Courier Status
ADDAKAL (2)	a0Q7Q000008wANX	Set	0	Shipped
	a0Q7Q000008wANY	kurta	1	Unshipped
Subtotal				
Adilabad (2)	a0Q7Q000008wAPB	Set	1	Shipped
	a0Q7Q000008wAW2	kurta	1	Shipped
Subtotal				
AHMADNAGAR (1)	a0Q7Q000008wA97	kurta	1	Shipped
Subtotal				
AKOLA (3)	a0Q7Q000008wA3O	kurta	0	Cancelled
	a0Q7Q000008wA6B	kurta	1	Shipped
	a0Q7Q000008wACM	kurta	1	Shipped



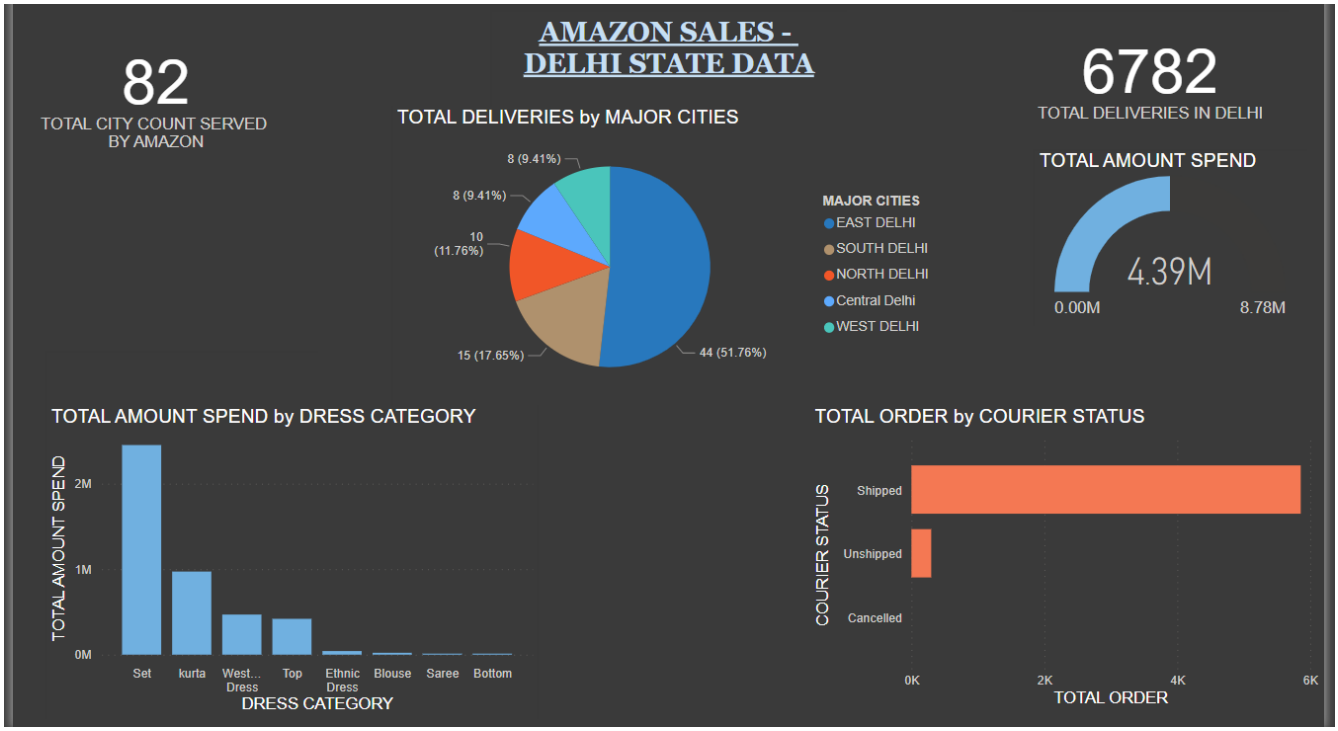
NEW DELHI HAS the highest count with 932

8.DATA VISUALISATION USING POWER BI

Power BI is a powerful business intelligence tool that can greatly enhance our project analysis and reporting capabilities. With its intuitive interface and rich set of features, Power BI allows us to connect to multiple data sources, transform and cleanse the data, and create insightful visualizations and reports. Power BI also enables us to create dynamic dashboards, where we can consolidate multiple visualizations onto a single screen. This allows for easy monitoring and comparison of different metrics and KPIs related to our project goals. Additionally, Power BI supports interactive filtering and slicing, which enables us to drill down into specific subsets of the data and perform ad-hoc analysis. Overall, Power BI empowers us to present our findings in a visually appealing and interactive manner, making it easier for stakeholders to understand and derive insights from the data. By leveraging the capabilities of Power BI, we can effectively communicate the outcomes of our analysis and provide actionable recommendations based on the project objectives.

In our project, we will utilize Power BI to create interactive dashboards and visualize the datasets related to Amazon's retail sales and orders in India. Specifically, we will focus on Delhi, Karnataka, Maharashtra, and Telangana regions. By leveraging Power BI's visualization capabilities, we aim to gain insights into sales performance, order trends, and regional variations. These visual representations will support informed decision-making and help optimize sales strategies in the selected regions.

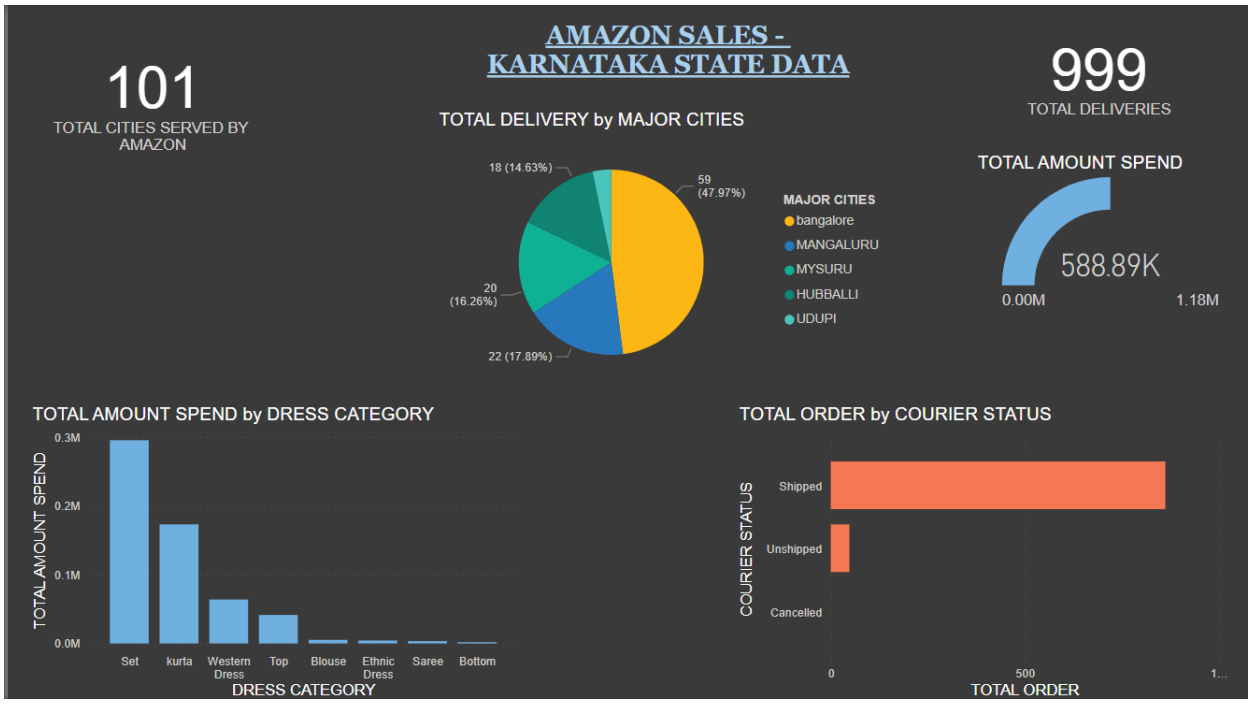
8.1 DELHI AMAZON SALES



In this dashboard, we can see that the total number of cities served by Amazon in Delhi is 82.

The total number of deliveries in Delhi in the provided specific period is 6782 and people have spent almost 4.39 million on orders from Amazon. The given bar chart shows that most of them have spent money on set clothes followed by Kurtis. Almost 90 percent of the orders have been shipped in Delhi and most orders have been delivered in East Delhi followed by South Delhi

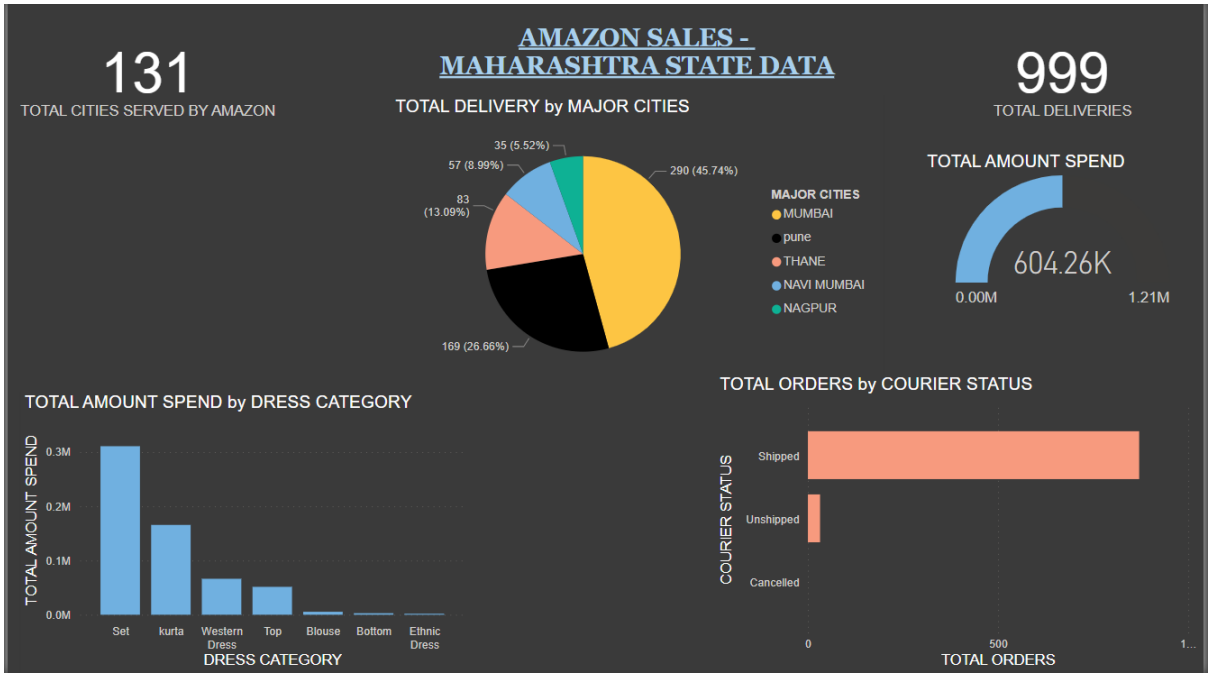
8.2. KARNATAKA AMAZON SALES



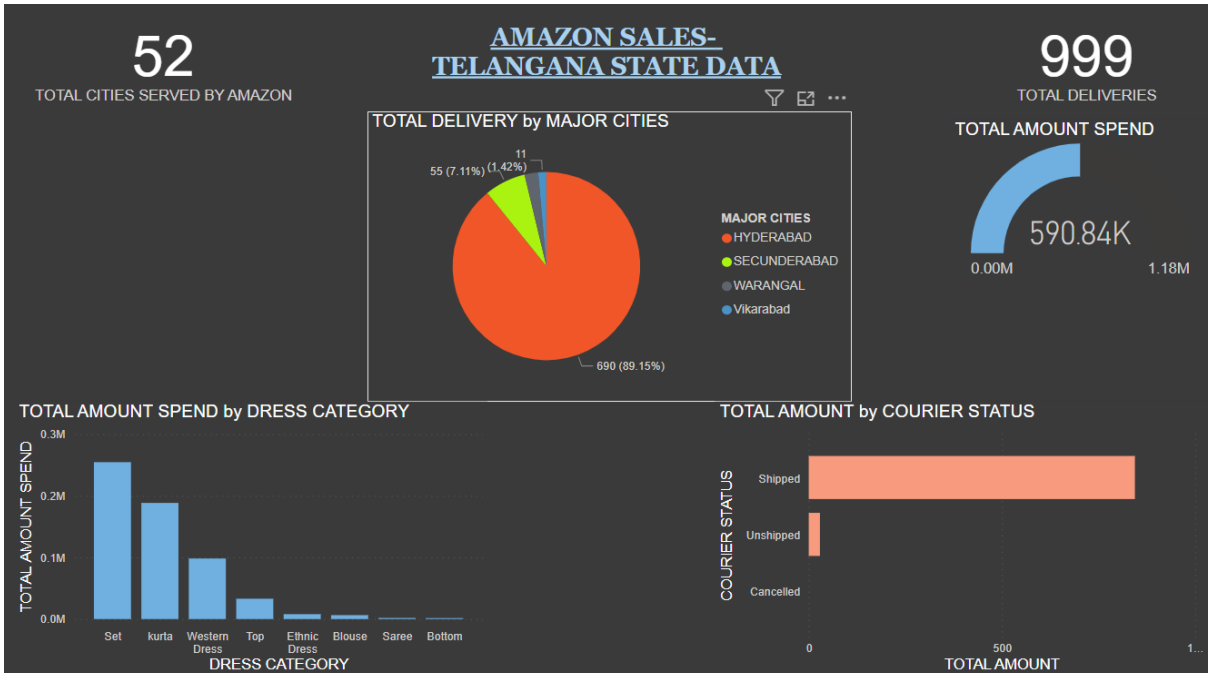
In this dashboard, we can see that the total number of cities served by Amazon in Karnataka is 101. The total number of deliveries in Delhi in the provided specific period is 999 and people have spent almost 5.88 lakh money on orders from Amazon. The given bar chart shows that most of them have spent money on set clothes for almost 2 lakh followed by Kurtis. Only 10 percent of the orders were unshipped, and most orders have been delivered in Bengaluru that accounts almost half of the total orders followed by Mangalore and Mysore.

8.3. MAHARASHTRA AMAZON SALES

In the below dashboard, we can see that the total number of cities served by Amazon in Maharashtra is 131, this is because it is one of the largest state in India. The total number of deliveries in Maharashtra in the provided specific period is 999 and people have spent almost 6.33 lakhs money on orders from Amazon. The given bar chart shows that most of them have spent money on set clothes above 2 lakhs followed by Kurtis. This trend is seen in the same manner all over the country. Just as all other states, only 10 percent of the orders were unshipped, and remaining orders were successfully shipped. Most orders have been delivered in Mumbai that accounts for more than half of the total orders followed by Pune and Thane.



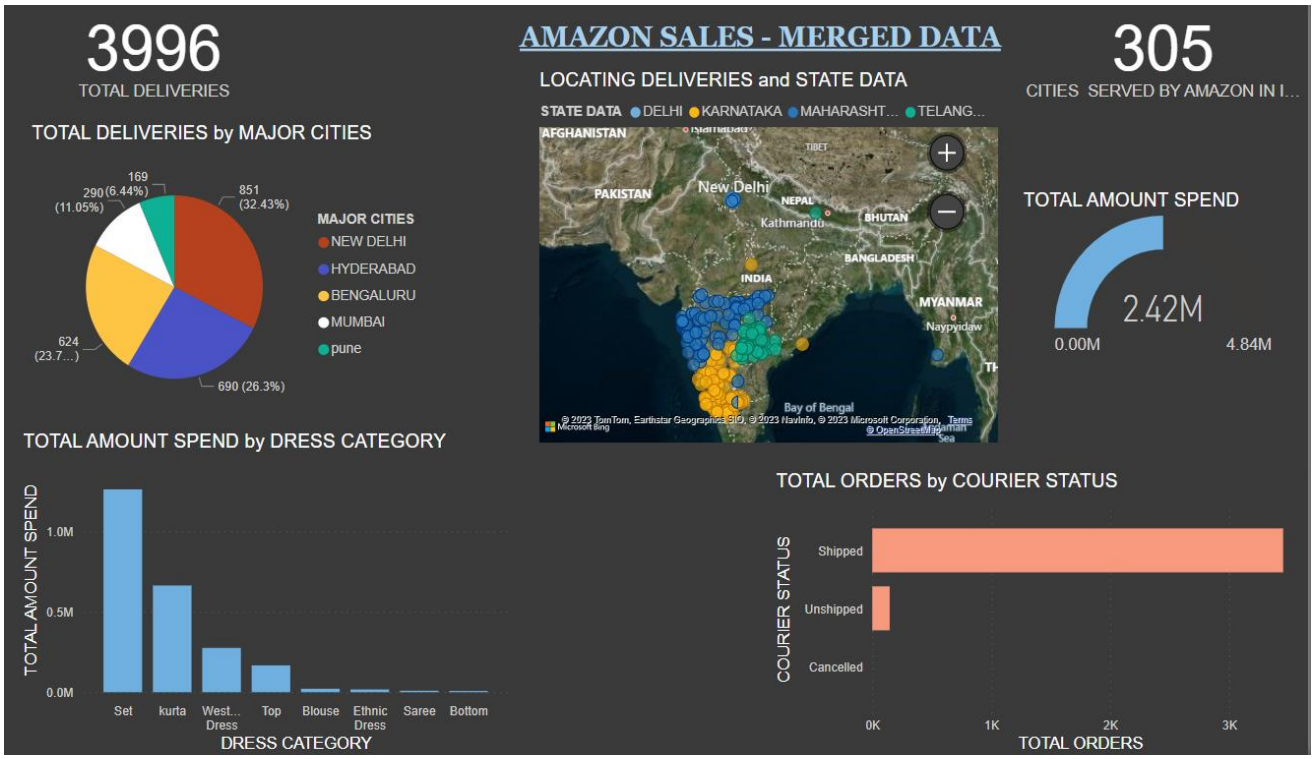
8.4. TELANGANA AMAZON SALES



In this dashboard, we can see that the total number of cities served by Amazon in Telangana is 52. The total number of deliveries in Telangana in the provided specific period is 999 and people have spent almost 5.90 lakhs money on orders from Amazon. The given bar chart shows that most of them have spent money on set clothes above 2 lakhs followed by Kurtis. Here we can see an increment in orders for the western dress. Just as all other states, only 10 percent of

the orders were unshipped, and remaining orders were successfully shipped. Most orders have been delivered in Hyderabad that accounts for most of the total orders followed by Secunderabad.

8.5. MERGED AMAZON SALES



The merged dataset dashboard in Power BI provides valuable insights into Amazon's orders and sales in India for a specific period, encompassing the states of Delhi, Karnataka, Maharashtra, and Telangana. This comprehensive dashboard consolidates data from these regions, allowing for a holistic understanding of customer behaviour and market trends.

The dashboard reveals that a total of 3,996 orders were placed during the specified period, indicating a substantial volume of transactions. These orders were served across 305 different cities, highlighting the widespread reach of Amazon's retail operations in India. Impressively, customers spent a total of 2.42 million on their orders, reflecting the significant economic impact of e-commerce in the country. Examining the order status, it is noteworthy that the majority of orders, exceeding 3,000, were successfully shipped to customers. This indicates a reliable and efficient delivery process. However, a smaller proportion of orders, less than 900, encountered issues such as cancellations or being unshipped, likely due to various circumstances such as stock availability or customer preferences.

Analysing the regional distribution of orders, it becomes apparent that New Delhi emerged as the leading city in terms of order volume, followed closely by Hyderabad and Bangalore. This highlights the importance of these metropolitan areas as key markets for Amazon's retail

business in India. Furthermore, the dashboard provides insights into customer preferences in terms of product categories. Among the various product types, Kurtis and set dresses witnessed the highest monetary transactions, indicating a strong demand for these fashion items. On the other hand, sarees recorded comparatively lower spending, suggesting a lesser preference for this category during the analysed period. But we can see a significant increase in the number of people ordering Western Dresses too.

Overall, the merged dataset dashboard in Power BI offers a comprehensive overview of Amazon's orders and sales in India, emphasizing key metrics such as order volume, regional distribution, and product category performance. These insights provide valuable information for strategic decision-making, enabling Amazon to tailor its marketing strategies, inventory management, and customer engagement initiatives to maximize sales and enhance customer satisfaction in the Indian market.

8. IMPROVED SOLUTIONS IN THE ORGANISATION

As a business analyst working in Amazon sales in India, there are several solutions and strategies that can be implemented to improve business performance, enhance sales, and address any identified flaws. These solutions are derived from the insights gained from the merged dataset dashboard and aim to optimize operations and capitalize on influential factors.

- Targeted Marketing Strategies: Based on the regional distribution of orders, it is evident that certain cities, such as New Delhi, Hyderabad, and Bangalore, exhibit higher order volumes. To capitalize on these key markets, Amazon can develop targeted marketing campaigns tailored to these regions. By focusing on localized promotions, partnerships with local influencers, and region-specific advertisements, the company can further penetrate these markets and increase customer engagement.
- Inventory Optimization: Analysing the product categories reveals that Kurtis and set dresses attract higher spending compared to sarees. To maximize sales and ensure optimal inventory levels, Amazon should consider adjusting its product assortment and stocking strategies. Increasing the variety and availability of popular items, while also monitoring demand trends, will allow for better inventory management and help meet customer preferences effectively.
- Streamlined Delivery Process: While the majority of orders were successfully shipped, the smaller proportion of unshipped or cancelled orders indicates room for improvement in the delivery process. Amazon can focus on enhancing logistics and supply chain management to minimize disruptions and improve overall delivery efficiency. This can involve implementing real-time tracking systems, optimizing warehouse operations, and establishing proactive communication channels with customers to address any potential issues promptly.
- Customer Feedback and Engagement: The insights gained from the merged dataset dashboard provide valuable information on customer behaviour and preferences. Amazon can leverage this data to enhance customer engagement and satisfaction. Implementing targeted feedback collection mechanisms, such as post-purchase surveys or product reviews, can help gather insights directly from customers. This information can then be used to refine products, tailor recommendations, and personalize the overall shopping experience, ultimately driving customer loyalty and repeat purchases.

- Market Expansion Opportunities: While the analysed dataset focuses on specific regions, the insights obtained can serve as a foundation for identifying potential market expansion opportunities. By exploring emerging cities and regions with promising growth potential, Amazon can strategically plan its expansion efforts. Conducting market research, analysing demographic and economic factors, and leveraging data-driven insights will enable the company to make informed decisions regarding market entry, infrastructure development, and customer acquisition strategies.

10. CONCLUSION