Business Intelligent Solution for Days for Clothing

A Final Project For

Mis 584: Business Intelligence, Fall 2024

Authors:
Rucha Wange, Bijal Patel, Aman Sahu, & Vikrant Prayag

Advisor:
Professor Nima Kordzadeh

Date:
9 December 2024

Table of Content

I. Executive Summary

Days for Clothing, a major street style fashion business situated in Dharavi, Mumbai, India, has grown in popularity among young people, influencers, and stylists in the film industry, and is known for its unique looks. As consumer demand has increased and market competition has heated up, the brand acknowledges the importance of well-managed data for making educated decisions. Their current data management practices are ineffective and inefficient. While Excel is their major business intelligence tool, they also employ ithinklogistics, an ERP system that is coupled with their website's payment processor, Shopify. The Days for Clothing Project has created three prototypes to improve data support and visualization, resulting in better decision-making skills:

1. Sales Dashboard	2. Customer Insights and Activity Dashboard	3. Predictive Analytics

The Sales Dashboard will offer a complete, real-time snapshot of sales performance across many product categories. It will highlight developing patterns, track progress toward sales targets, and identify opportunities for improvement. This dashboard will allow the management team to quickly modify tactics and make educated decisions to increase income.

This Customer Insights and Activity dashboard will provide detailed insights into customer behavior and preferences by analyzing data from different touchpoints to identify trends in

purchase and interaction. It will be a valuable tool for customizing marketing campaigns and improving consumer experiences, eventually increasing customer loyalty and retention.

Finally, prediction models estimate future trends and client purchase behaviors with great accuracy. It will give essential insight into market demand of denim, aid in inventory management, and proactively handle possible interruptions. This proactive strategy will help the organization remain competitive and responsive to changing market conditions.

II. Introduction

2.1 Background on Days for Clothing

Days for Clothing (DFC), proudly "Made in Dharavi," India, is a premium streetwear company known for its design-first philosophy and dedication to quality. DFC was founded in 2017 by two college friends with the goal of making high-end fashion more accessible and affordable to a larger audience while maintaining workmanship. Because of its unique style and odd designs, the company experienced little client involvement in its early years. However, through constant brand-building efforts and inventive marketing methods, DFC effectively expanded its consumer base and acquired industry reputation. What began as a small warehouse operation has evolved into a thriving e-commerce business, serving over 10,000 customers globally. Its flagship store stays true to its roots while embracing a contemporary streetwear identity. By blending authenticity, originality, and innovation, DFC continues to redefine India’s streetwear scene, with aspirations of becoming a global denim powerhouse.

2.2 Current BI Practices

DFC relies primarily on MS Excel for data management and analysis, which limits its ability to perform complex analytics or scale efficiently. The absence of a centralized data warehouse makes it challenging to integrate and analyze data from multiple sources. Although its ERP system, iThinkLogistics, supports operational tasks like tracking orders, delivery and return, it lacks advanced data analytics features. With no dedicated BI team, DFC’s ten-member staff collectively manage marketing, sales, design, and inventory, reducing the focus on data-driven decision-making. Implementing specialized BI tools like Tableau and fostering a data-centric culture will be essential for the brand's continued growth and efficiency.

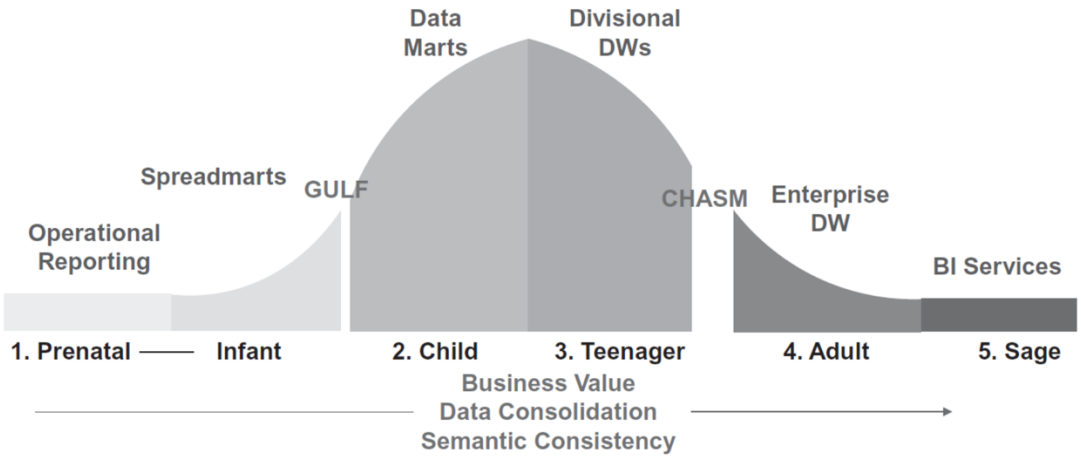


Fig 2.2 BI Framework

DFC is in between prenatal and infant stage of the Business Intelligence (BI) maturity model, indicating the organization has minimal BI capabilities.

Dependency on Basic Tools: DFC stores and analyzes data only using Microsoft Excel. Despite its versatility, Excel lacks the advanced analytical capabilities, automation, and scalability required for strong business intelligence processes.

Manual Procedures: There is a greater chance of mistakes, inefficiencies, and inconsistent data because the data management and analysis procedures are probably manual.

Limited Analytical Depth: Excel's basic analysis limits insights to basic descriptive statistics and patterns, making it difficult to perform predictive or prescriptive analytics.

Data management that is fragmented: Excel is not made to handle big datasets or integrate with other data sources. This restriction may result in fragmented data, which makes detailed analysis difficult.

2.3 Related Case Studies

This section will showcase how Zara and Urban Outfitters effectively utilize Business Intelligence (BI) tools and solutions to enhance business growth and operational efficiency, outlining how these practices could be incorporated by DFC to achieve similar outcomes.

2.3.1. Zara

Data Integration and Real-Time Analysis: To improve decision-making, Zara combines data from multiple sources, such as mobile apps, e-commerce platforms, and point-of-sale (POS) systems, to provide a consistent picture of sales and inventories.

Data Chat: Zara's Data Chat system enables staff members to ask basic questions regarding inventories and sales. This program forecasts future trends, aggregates data from social media, internet sales, and retail locations, and provides real-time updates on product performance. By matching with preferences, these insights help managers make better decisions more quickly, collaborate better, and increase customer satisfaction.

Zara uses predictive analytics to eliminate overproduction, prevent stockouts, and optimize production, which lowers waste and increases efficiency.

2.3.2. Urban Outfitters

Real-Time Reporting: Urban Outfitters offers almost real-time updates on key performance indicators (KPIs) like inventory, sales conversions, and top-selling products through the use of Qlik Sense. This enables retail employees quickly arrive at well-informed decisions.

AI-Powered Merchandise Planning: By introducing AI capabilities for demand forecasting, replenishment, and merchandise finance planning, the collaboration with o9 Solutions improved inventory alignment with consumer expectations and trends.

Localized Demand Insights: Urban Outfitters uses CB4's AI tools to find products that have unsatisfied local demand, which enables retailers to modify their merchandising plans and increase sales.

2.3.3. Takeaways for DFC

Using BI solutions like AI-powered planning, Data Chat platforms, real-time reporting, and predictive analytics, DFC can:

1. Improve decision-making with simple tools and seamless data integration.
2. Improve client satisfaction by attending to regional needs.
3. Reduce expenses and boost profitability by streamlining inventory control and operational effectiveness.
4. These tactics show how corporate processes can be transformed by BI tools and approaches, guaranteeing long-term growth and competitiveness.

III. Proposed BI Solution

3.1 Overview of Business Intelligence

Business intelligence (BI) is the technology-driven process of collecting, integrating, analyzing, and displaying data to help businesses make better decisions. For a growing firm like DFC, business intelligence is more than a luxury; it is a need. This allows them to:

- Gain insight into how their products are performing in various markets.
- Identify the most profitable consumer segments.
- Optimizing marketing efforts and resource allocation.
- Make accurate estimates regarding future trends.

3.2 The main Components of BI Solution

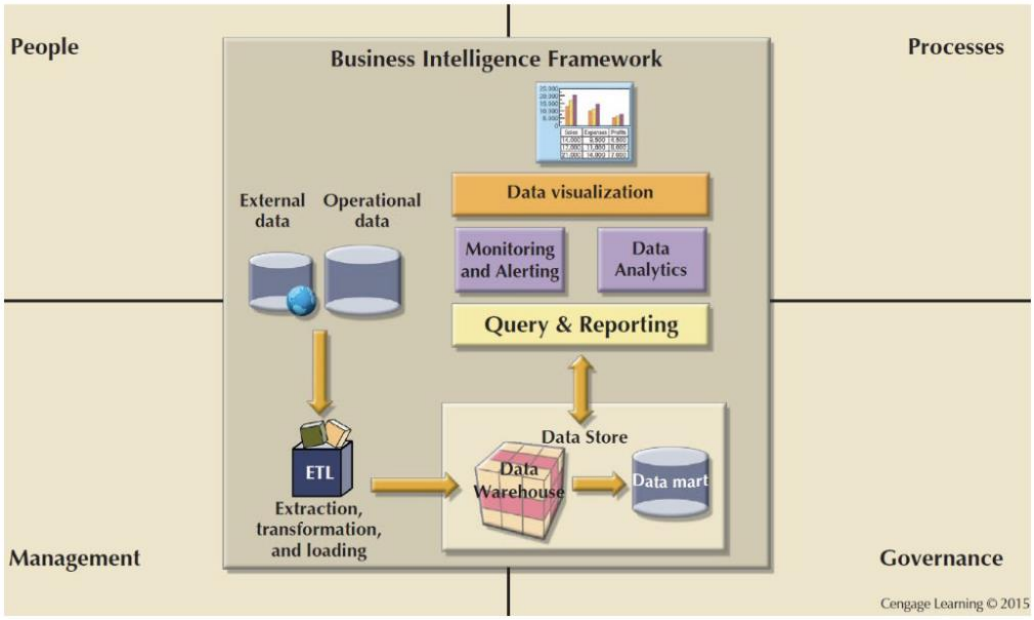


Fig 3.2 BI Framework

A Business Intelligence (BI) solution combines several essential elements that work together to transform raw data into meaningful insights. These components help businesses acquire, organize, and evaluate data, giving them the skills they need to make informed decisions. A BI system with a strong structure assures data consistency, accessibility, and relevance in support of strategic and operational objectives.

3.2.1 Data Sources

This project's data was mostly derived from internal and external sources, which provided crucial insights into DFC's operations and market performance.

Internal Data: The core dataset included DFC's sales data, marketing campaign performance data, and customer interaction data.
External data: External data includes metrics from social media platforms such as Instagram and Facebook, as well as website statistics from Google Analytics.

By combining different data sources, the BI solution delivers a complete picture of DFC's operational and market dynamics.

3.2.2 ETL Processes

Extract data from many sources, including Excel files, Shopify, and social media analytics platforms.

Transform: Clean and standardize data to ensure consistency across formats and metrics.
For example, unifying customer names and IDs across Shopify and social media platforms.

Load: Move the transformed data to a centralized data warehouse for quick access and analysis.

3.2.3 Data Wareshouse

The proposed data warehouse will be structured using a star schema, which will allow for effective data organization and retrieval. Products will include details such as denim styles, sizes, collections, and inventory levels; Customers will include demographic profiles, purchase histories, and engagement metrics; Time will track data over days, months, and years for trend analysis; and Sales will combine transactional data with profit margins and revenue metrics. This design enables smooth querying and serves as a solid platform for both operational insights and strategic decision-making.

3.2.4 Querying and Reporting Tools

Querying and reporting tools allow you to access and evaluate data stored in a data warehouse. For example, these technologies enable managers to measure weekly sales performance, assess the effectiveness of marketing efforts, and discover client segments driving the most revenue. These tools enable users to generate both predefined and ad hoc reports, allowing decision-makers to identify patterns, monitor KPIs, and efficiently address crucial business concerns.

3.2.5 Data Analytics

Data analytics is critical for transforming DFC's raw data into usable insights, which drive operational efficiency and strategic growth. Analytics assists in the identification of patterns such as top-performing denim styles, seasonal demand changes, and customer purchase behaviors by leveraging past sales and marketing data. Predictive analytics forecasts future patterns, such as

predicted sales volumes or customer preferences, helping DFC to plan inventory and improve marketing activities. For example, if analytics show an increase in demand for specific denim collections over the holiday season, the company might adjust its production and advertising to maximize income. Prescriptive analytics also makes actionable recommendations, such as reallocating inventory to high-demand areas or launching focused advertising for specific client segments. This comprehensive use of data analytics enables DFC to make more informed decisions, increase customer satisfaction, and boost its market position.

3.2.6 Data visualization

Data visualization transforms raw data into actionable insights through intuitive visuals, enabling DFC to interpret complex information and make data-driven decisions. Using Tableau, dynamic dashboards feature interactive tools like drill-downs, real-time updates, and customizable views. Examples include maps highlighting customer distribution by state, bar charts for product ratings (e.g., "Urban Luxe"), and donut charts showcasing revenue by sales channels. Line charts track weekly profit versus sales trends, while top customer charts identify key buyers for loyalty strategies. Tableau’s flexibility and best practices ensure clear, engaging visuals that support strategic growth and operational efficiency.

3.3 The Role and Value of Each Component

Each component of the proposed BI solution contributes significantly to the transformation of raw data into actionable insights for DFC. Internal and external data sources serve as the foundation for understanding consumer behavior, sales performance, and market trends. The ETL process guarantees that data is clean, consistent, and ready for analysis by extracting essential information, changing it into a standard format, and loading it into the data warehouse. The data warehouse, which is organized using a star schema, allows for efficient data storage and retrieval, thereby aiding both operational and strategic decision-making. Querying and reporting technologies provide quick access to insights via dynamic reports and dashboards, whereas data analytics uses complex approaches to forecast patterns and optimize decision-making. Finally, data visualization simplifies complex information by providing intuitive dashboards and visualizations, allowing DFC stakeholders to make well-informed timely, and strategic business choices. These components work together to form an integrated system that promotes the company's efficiency, growth, and competitive edge.

3.3.1 Proposed BI solution for DFC

DFC currently does not have a dedicated BI team or BI solution. All reporting and decision-making are primarily reliant on manual processes and basic technologies like Excel, limiting their ability to evaluate data efficiently. To bridge this gap, we created a customized BI system that gives actionable data for optimizing sales, understanding consumer behavior, and improving

marketing efforts. This solution includes three interactive dashboards designed to provide DFC with data-driven decision-making capabilities.

Sales Performance Dashboard

This dashboard analyzes DFC's sales data to provide a clear view of product and category performance. The key features include:

- A heatmap of monthly sales patterns for product categories such as denim, t-shirts, and cargo.
- A line chart depicting weekly profit and sales trends to help DFC track short-term performance and overall business success.
- The breakdown of product category contributions to overall sales enables the organization to optimize inventories and focus on high-performing items.

Customer Insights and Marketing Analytics Dashboard

This dashboard combines customer demographics, social media interaction, and sales data to provide a more detailed view of client preferences and marketing effectiveness. Features include:

- A geographic map of customer distribution by state, indicating places with good and weak performance.
- Customers are segmented based on their age, target demographic, and preferred sales methods.
- A bar chart ranks top customers based on total purchases, allowing DFC to create individualized loyalty programs.

Predictive Analytics Dashboard

This dashboard uses predictive analytics to help DFC plan for the future. This includes:

- Yearly sales growth trends for specific product categories, such as denim and cargo, helping forecast future demand.
- A heatmap of monthly sales trends can help discover seasonal patterns and influence production and inventory planning.
- Predictive insights into future sales performance enable DFC to proactively change marketing tactics and budget allocation.

3.3.2 Recommendation for Data Warehouse Structure

Given the lack of an existing BI infrastructure, we propose constructing a centralized data warehouse to store and analyze DFC's internal and external information. The warehouse is

intended to facilitate scalability and integration with data sources such as Shopify, social media networks, and Google Analytics. The proposed structure contains:

- Fact Table: A centralized database for sales figures like revenue, units sold, and profit margins.
- Dimension Tables:
    - Product Dimensions: Details on product categories, styles, and collections.
    - Customer Dimensions: Information on customer demographics, buying history, and engagement.
    - Time Dimension: Data is organized by day, month, and year to allow for analysis of trends and seasonal patterns.
    - Channel Dimension: Use data from in-store, online, and social media channels to assess performance by sales medium.

This star architecture simplifies data querying and analysis, making it accessible to a small team with no prior BI experience. DFC will be able to migrate from manual processes to a more efficient, data-driven approach once this framework is implemented. The dashboards built on this foundation will allow DFC to track performance, discover growth opportunities, and make educated decisions, paving the way for future scalability and success.

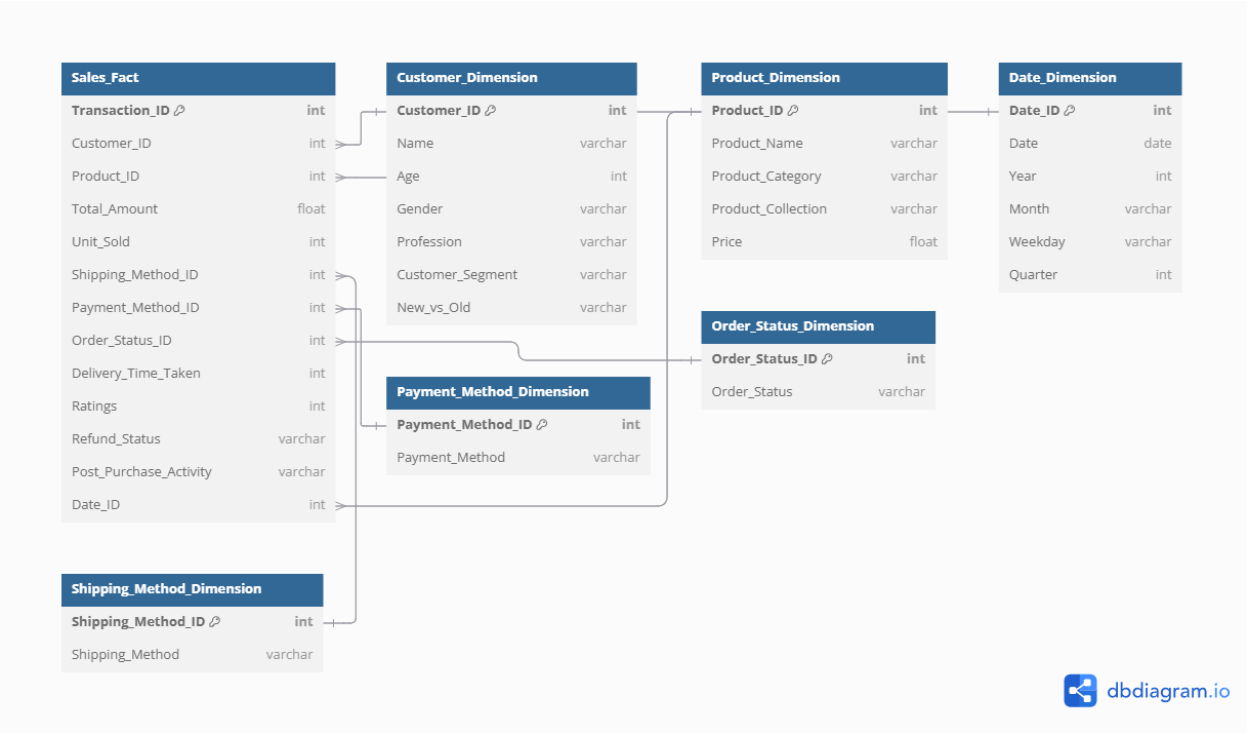


Fig 3.3.2 Star Schema of Sales Analytcis

The first schema depicts a Sales Fact Table associated with dimensions like Product, Customer, Date, Payment Method, Order Status, and Shipping Method. This structure centralizes critical sales data such as revenue, units sold, and ratings, allowing for extensive monitoring of sales trends, customer behavior, and delivery efficiency. It enables queries across several dimensions, allowing for more complete performance tracking.

For Customer Insights:

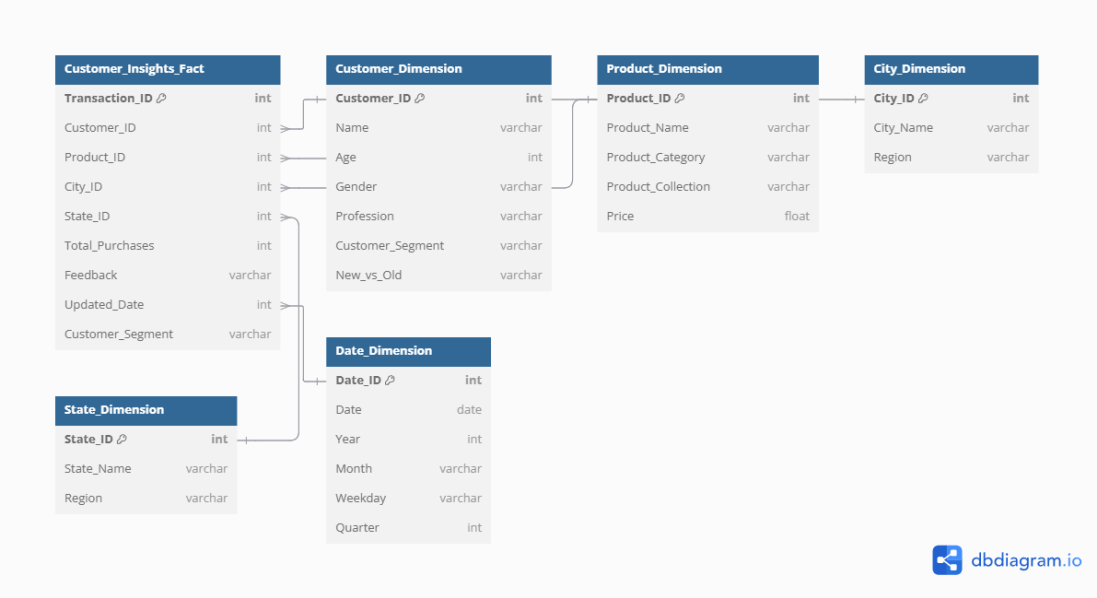


Fig 3.3.2 Star Schema for Customer Analytics

This schema focuses on Customer Insights, with a Fact Table that includes dimensions like Customer, Product, Date, City, and State. This schema focuses on customer-centric analysis by recording demographics, regional patterns, and purchase histories. It simplifies segmentation and feedback analysis, which are critical for targeted marketing and tailored consumer engagement.

3.3.3 Value of the Proposed Solution

The suggested data warehouse solution offers DFC a consolidated platform for integrating data from Shopify, social media, and Google Analytics, transforming fragmented procedures into simplified, data-driven approaches. It improves decision-making by providing precise insights into sales, customer behavior, and product performance, while also increasing operational efficiency through automated reporting and simplified queries. Customer-centric analysis enables targeted marketing and personalized engagement, resulting in increased retention and satisfaction. The scalable design facilitates future expansion by allowing DFC to optimize inventory, refine strategy, and acquire a competitive advantage in the market.

IV. Three Use Cases/Prototype

4.1 Sales Dashboard

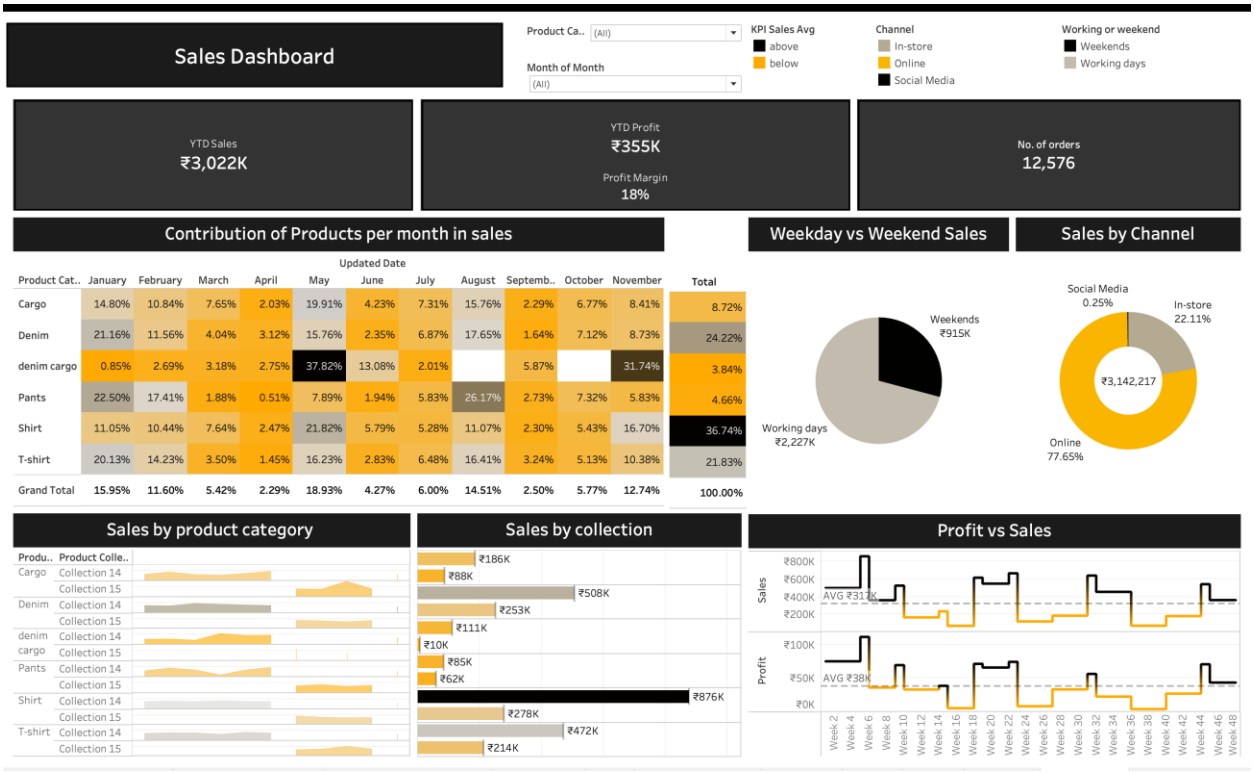
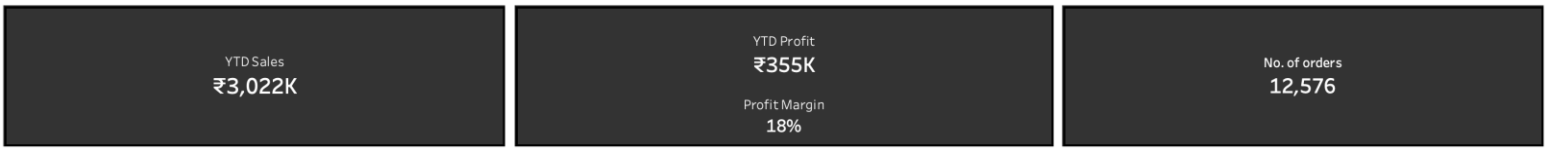


Fig 4.1 - Sales Dashboard

4.1.1 Overview

This comprehensive Sales Dashboard is an essential tool for DFC, offering a clear and actionable overview of important performance indicators such as total sales, profit margins, and order volume. DFC can detect trends and growth opportunities by examining product contributions, collection sales, and performance across channels such as online, in-store, and social media. Visualizations like weekday vs. working day sales and profit vs. sales comparisons allow for data-driven decision-making to manage inventory, better marketing strategies, and increase profitability. This dashboard is critical for streamlining processes, setting realistic goals, and creating long-term business success.



Key Performance Indicators:

This section shows key business KPIs at a glance:

Year-to-date sales are ₹3,022K, providing a brief picture of total revenue generated.

The company's profitability is highlighted by its YTD profit of ₹355K and 18% margin.

The number of orders (12,576) indicates the entire transaction volume.

These short metrics provide stakeholders with a high-level overview of performance, allowing for quick assessments of revenue, profitability, and operational efficiency.

Sheet1- Contribution of products per month in Sales

Contribution of Products per month in sales

	Updated Date											
Product Cat..	January	February	March	April	May	June	July	August	Septemb..	October	November	Total
Cargo	14.80%	10.84%	7.65%	2.03%	19.91%	4.23%	7.31%	15.76%	2.29%	6.77%	8.41%	8.72%
Denim	21.16%	11.56%	4.04%	3.12%	15.76%	2.35%	6.87%	17.65%	1.64%	7.12%	8.73%	24.22%
denim cargo	0.85%	2.69%	3.18%	2.75%	37.82%	13.08%	2.01%		5.87%		31.74%	3.84%
Pants	22.50%	17.41%	1.88%	0.51%	7.89%	1.94%	5.83%	26.17%	2.73%	7.32%	5.83%	4.66%
Shirt	11.05%	10.44%	7.64%	2.47%	21.82%	5.79%	5.28%	11.07%	2.30%	5.43%	16.70%	36.74%
T-shirt	20.13%	14.23%	3.50%	1.45%	16.23%	2.83%	6.48%	16.41%	3.24%	5.13%	10.38%	21.83%
Grand Total	15.95%	11.60%	5.42%	2.29%	18.93%	4.27%	6.00%	14.51%	2.50%	5.77%	12.74%	100.00%

Fig 4.1.1 - Contribution of products per months in Sales

The heatmap visualizes the monthly contribution of each product category (e.g., Denim, Cargo, Shirts) to DFC's total sales. Each cell reflects a category's percentage share for a given month, with darker hues suggesting larger contributions and lighter shades indicating smaller contributions. Denim, for example, generally performs well, accounting for 24.22% of total sales each year, but Denim Cargo peaks in various months, such as May (37.82%).

Sheet2- Weekday vs Weekends Sales

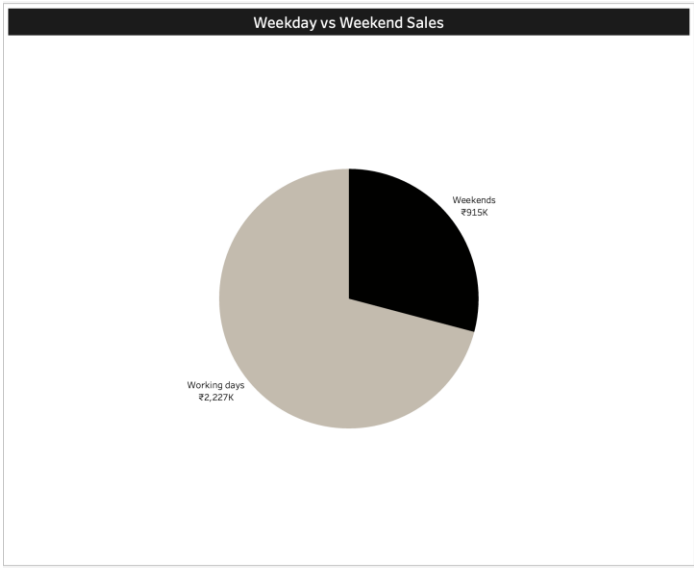


Fig 4.1.1 - Weekday vs Weekends sales.

The pie chart compares sales performance on weekdays and weekends, with ₹2,227K earned on weekdays and ₹915,000 on weekends. This shows that weekdays account for a much larger part of total revenues, indicating customer purchasing behavior tendencies.

This graphic assists DFC in determining the best periods for allocating resources, such as staffing or promotional initiatives, in order to optimize sales. The chart also suggests developing targeted marketing for weekends to enhance sales during low-performing periods, providing consistent income generation throughout the week.

Sheet 3 – Sales by Channel

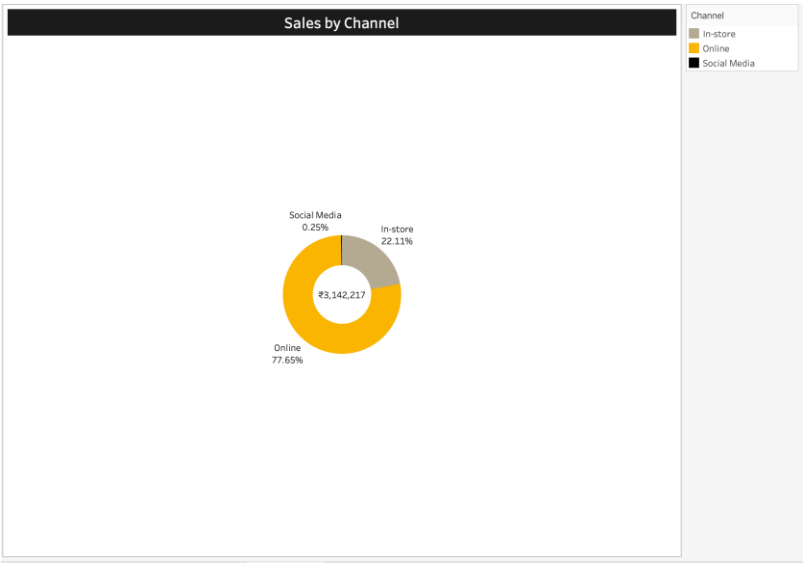


Fig 4.1.1 - Sales by channels

This donut chart breaks down DFC's sales by channel, indicating that online sales account for 77.65% of total revenue, followed by in-store sales at 22.11% and a minor contribution from social media at 0.25%. The depiction underlines the role of online platforms in driving income while also emphasizing future development opportunities for in-store and social media channels.

The graph also assists DFC in resource allocation by focusing efforts on improving online sales while also researching programs to increase in-store footfall and social media engagement. It also offers insights into adapting marketing campaigns and maximizing sales across all channels, allowing DFC to improve overall performance and customer reach.

Sheet4 – Sales by Product Categories and Collections

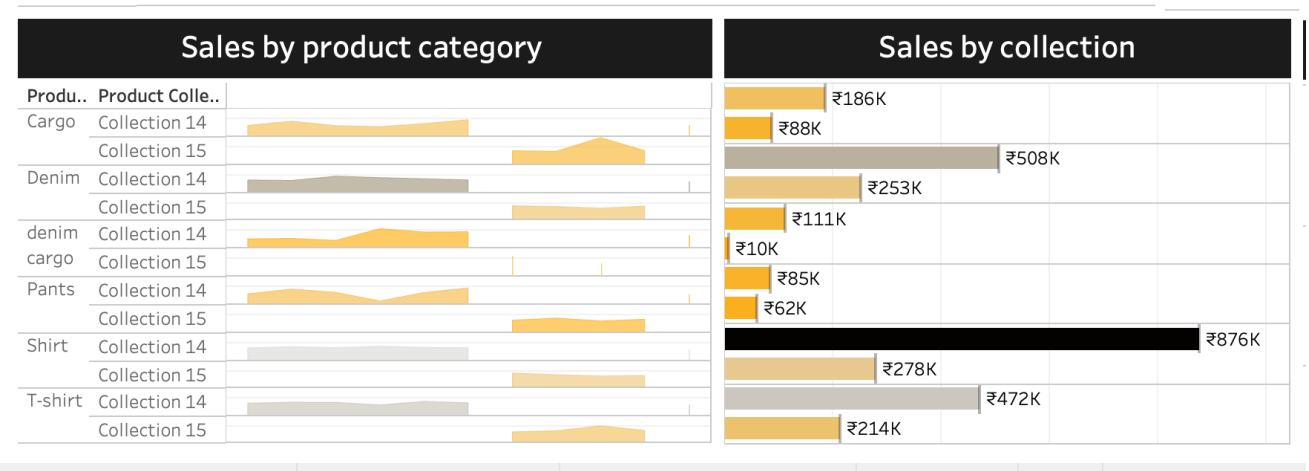


Fig 4.1.1 - Sales by product categories and collections.

Sales By Product Category

This chart depicts monthly sales for each product category across collections. Cargo reaches its high in March with ₹20,947 in Collection 14, while Denim continuously outperforms Collection 15.

Sales by Collection.

The report shows that Collection 15 Shirts had the highest overall sales (₹876K), followed by Denim (₹253K).

Key insights for DFC include

- monitoring seasonal changes for inventory and production planning.
- Identifies the best-performing collections for resource allocation.
- Directs marketing efforts to boost underperforming segments.

Sheet 5 – Profit Vs Sales

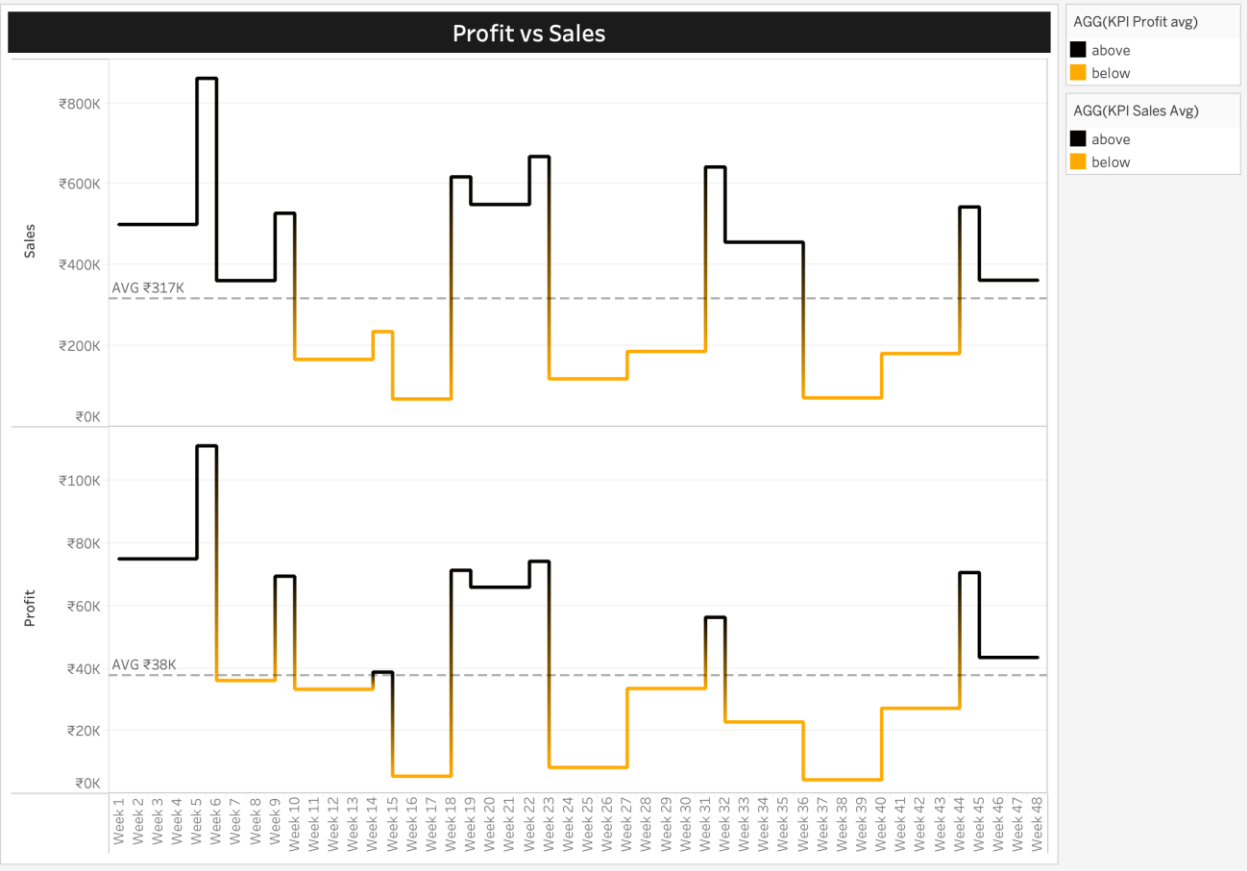


Fig 4.1.1 - Profit vs Sales

This line chart illustrates DFC's weekly profit and sales patterns. The figure compares performance to average benchmarks, with black lines representing weeks when sales or profit

exceeded the average and orange lines indicating below-average weeks.
Key insights:

- **Weekly Trends:** Tracks changes in sales and profit over 48 weeks, helping discover high and low-performing periods.
- **Above versus Below Average:** Identifies weeks with higher or lower sales and profit averages (₹317K vs. ₹38K, respectively).
- **Alignment of Sales and Profit:** Highlights correlations or discrepancies between sales and profitability, which aids in finding areas where strong sales may not transfer into increased profits.

Sales Dashboard

The entire sales dashboard is extremely valuable to DFC because it provides a comprehensive perspective of the company's performance in key areas. It allows tracking of sales, profit, and product contributions, which aids in the identification of high-performing categories such as denim and shirts. Insights regarding sales channels emphasize online sales' dominance (77.65%), which guides budget allocation and marketing activities. The dashboard also displays weekday versus weekend sales patterns, enabling for more effective staffing and promotional initiatives. Furthermore, profit vs. sales trends and collection-specific sales graphics assist DFC in monitoring financial health, identifying development prospects, and addressing underperforming regions. By combining these insights, the dashboard enables DFC to make data-driven decisions, optimize operations, and accomplish strategic growth objectives.

Interactive Filters

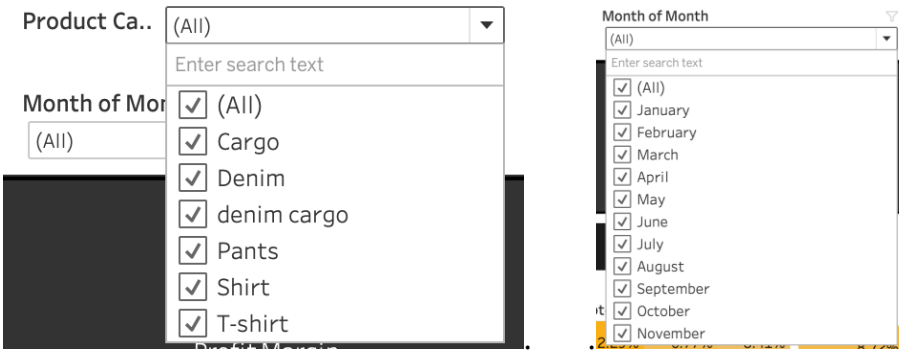


Fig 4.1.1 - Interactive Filters.

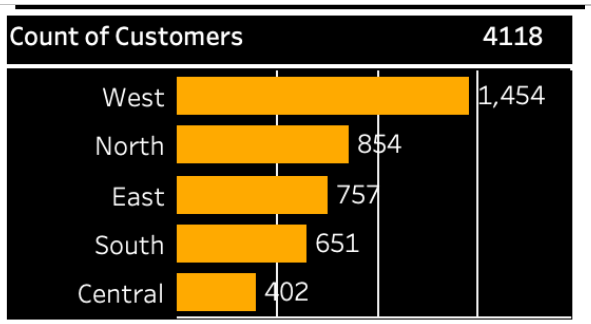
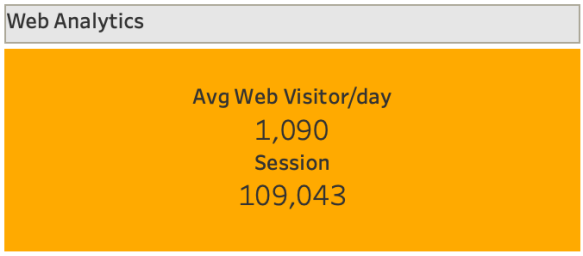
The interactive filter for product categories and months improves the dashboard's usability and performance.

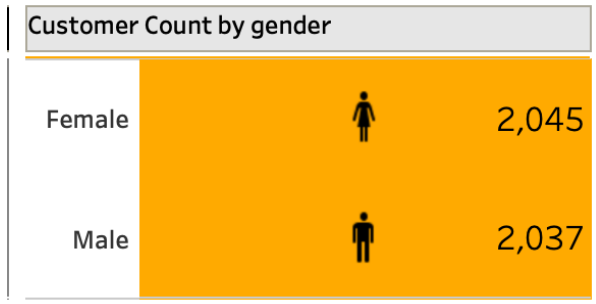
Product Category Filter: Users can restrict down data to specific categories like denim, shirts, and cargo. This aids in analyzing the success of particular product lines, identifying key contributors, and resolving underperforming areas.

Month Filter: Allows users to focus on specific months or compare performance over many time periods. This is especially valuable for monitoring seasonal trends, campaign results, and monthly sales surges.

These filters offer flexibility, allowing stakeholders to customize dashboard displays to meet their individual requirements, resulting in more precise and actionable insights. They ensure that the dashboard is dynamic, user-friendly, and highly adaptable for decision-making.

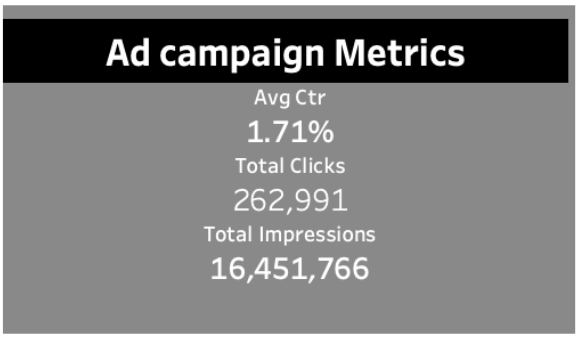
4.2. Customer Insights and Activity Dashboard

KPI's	Description
	Regional Customer Distribution: The Indian map splits the country into five regions—West, North, East, South, and Central—and shows client counts for each. This assists in identifying strong and weak market sectors.
	Web Analytics: Provides data on average daily web visitors (1,090) and total sessions (109,043) to help understand online traffic and engagement.
	Average Sales per Customer by Region: Displays average revenue per customer by region (e.g., ₹5.43K in the East), indicating high-value markets.



Customer Demographics:

Shows the gender breakdown of clients (2,045 females and 2,037 males), which aids in focused marketing tactics.



Ad Campaign Metrics:

Provides a summary of campaign performance using important metrics such as average CTR (1.71%), total clicks, and impressions. Clicking "Ad Campaign Metrics" opens

Sheet 1- High and Low Performing States

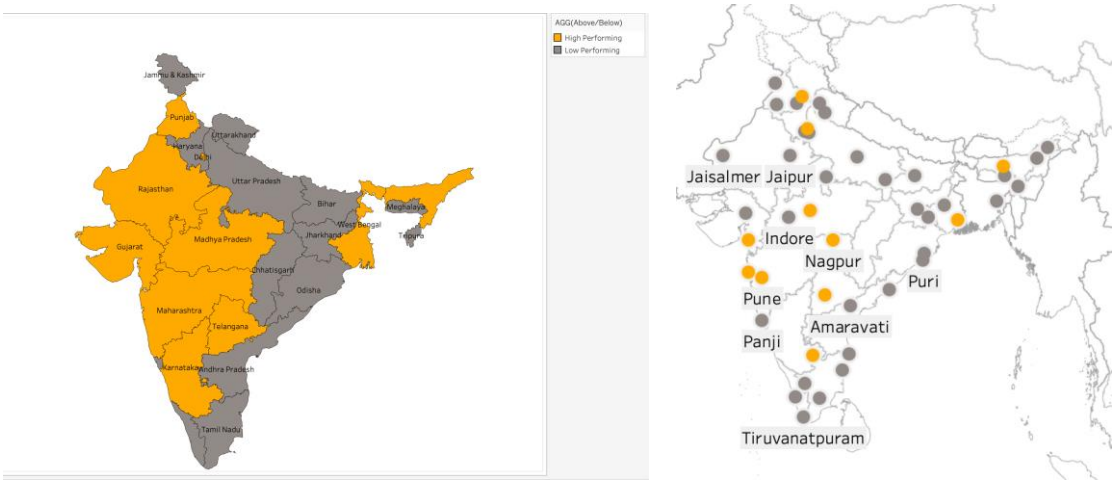
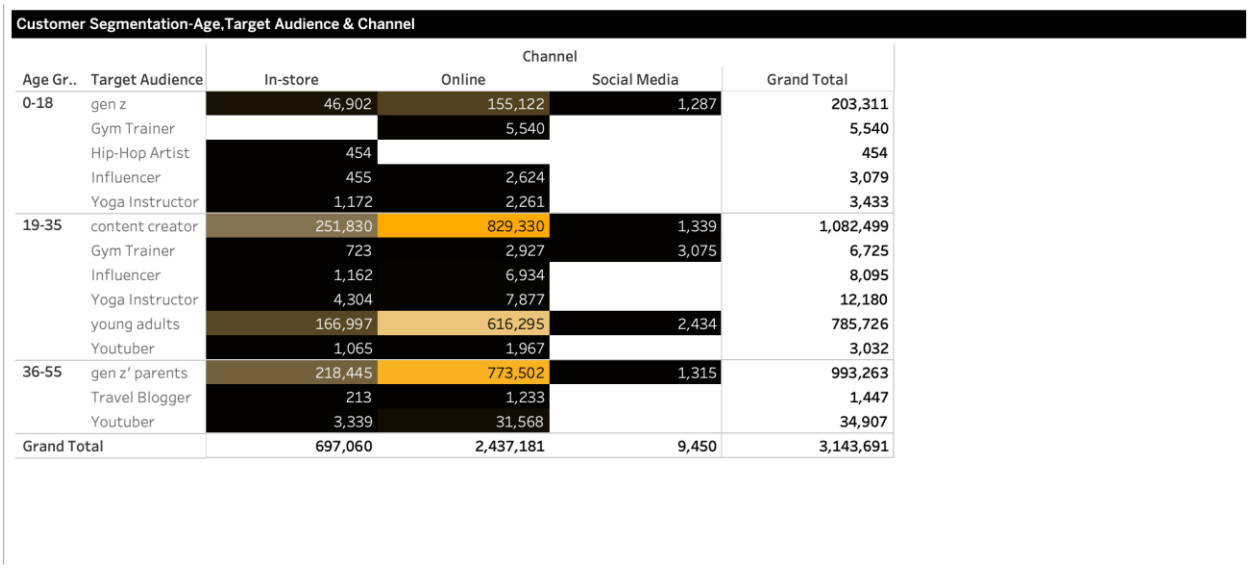


Fig 4.2 - High and Low performing states and cities.

The map visualizations show DFC's consumer dispersion. The state-level view displays high- and low-performing states based on customer count, providing a regional perspective. By toggling the switch button to "Customer count/City," the view changes to a more detailed city-level distribution, allowing you to pinpoint customer density within certain cities. This toggle

functionality allows users to smoothly transition between broader state insights and particular city-level data, assisting with strategic decisions for regional marketing and resource allocation.

Sheet 2 – Customer Segmentation: Age, target Audience and Channel



Customer Segmentation-Age,Target Audience & Channel

		Channel			
Age Gr..	Target Audience	In-store	Online	Social Media	Grand Total
0-18	gen z	46,902	155,122	1,287	203,311
	Gym Trainer		5,540		5,540
	Hip-Hop Artist	454			454
	Influencer	455	2,624		3,079
	Yoga Instructor	1,172	2,261		3,433
19-35	content creator	251,830	829,330	1,339	1,082,499
	Gym Trainer	723	2,927	3,075	6,725
	Influencer	1,162	6,934		8,095
	Yoga Instructor	4,304	7,877		12,180
	young adults	166,997	616,295	2,434	785,726
	Youtuber	1,065	1,967		3,032
36-55	gen z' parents	218,445	773,502	1,315	993,263
	Travel Blogger	213	1,233		1,447
	Youtuber	3,339	31,568		34,907
Grand Total		697,060	2,437,181	9,450	3,143,691

Fig 4.2 Customer Segmentation

This chart segments customers into age groups, target audiences, and sales channels (in-store, internet, and social media). It demonstrates that the 19-35 age group is the most involved, particularly online, with content creators and young adults making major contributions. The 0-18 demographic prefers online platforms, whereas clients aged 36-55 are more active both in-store and online. These analytics enable DFC to adjust marketing tactics, target high-performing personas, and discover growth possibilities in both digital and in-store channels.

Sheet 3- Top 10 Customers / Total Purchases and Frequency

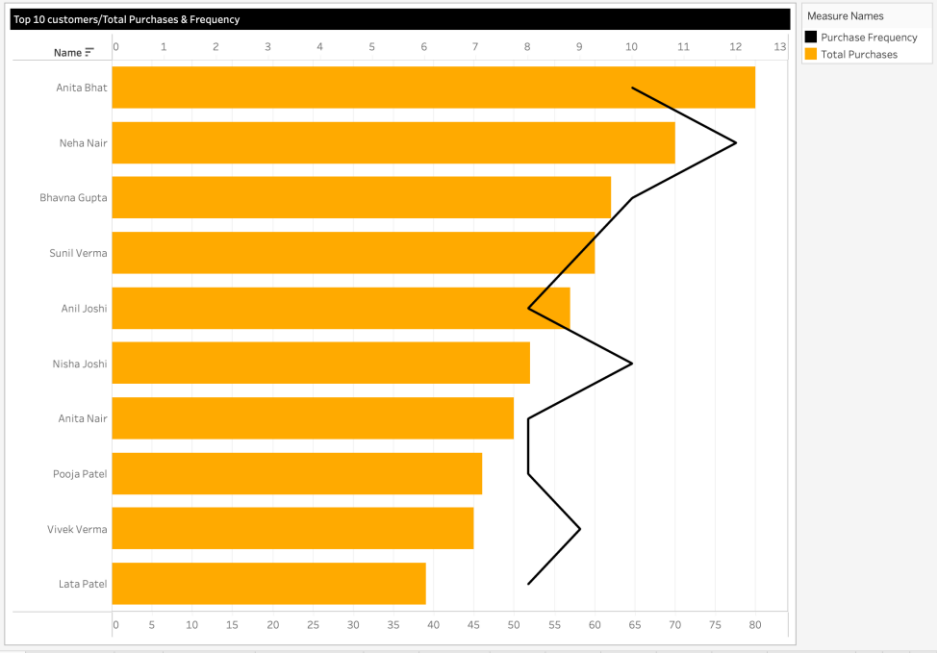


Fig 4.2 Top 10 customers/total purchases and frequency.

This bar chart shows the top ten customers based on total purchases (yellow bars) and buy frequency (black line). It identifies major contributors such as Anita Bhat and Neha Nair, who have substantial purchase volumes and make regular transactions. The dual visualization enables DFC to identify loyal clients and analyze their purchasing habits. This knowledge enables DFC to create tailored loyalty programs, optimize customer retention methods, and focus marketing efforts on high-value customers

4.3. Predictive Analytics

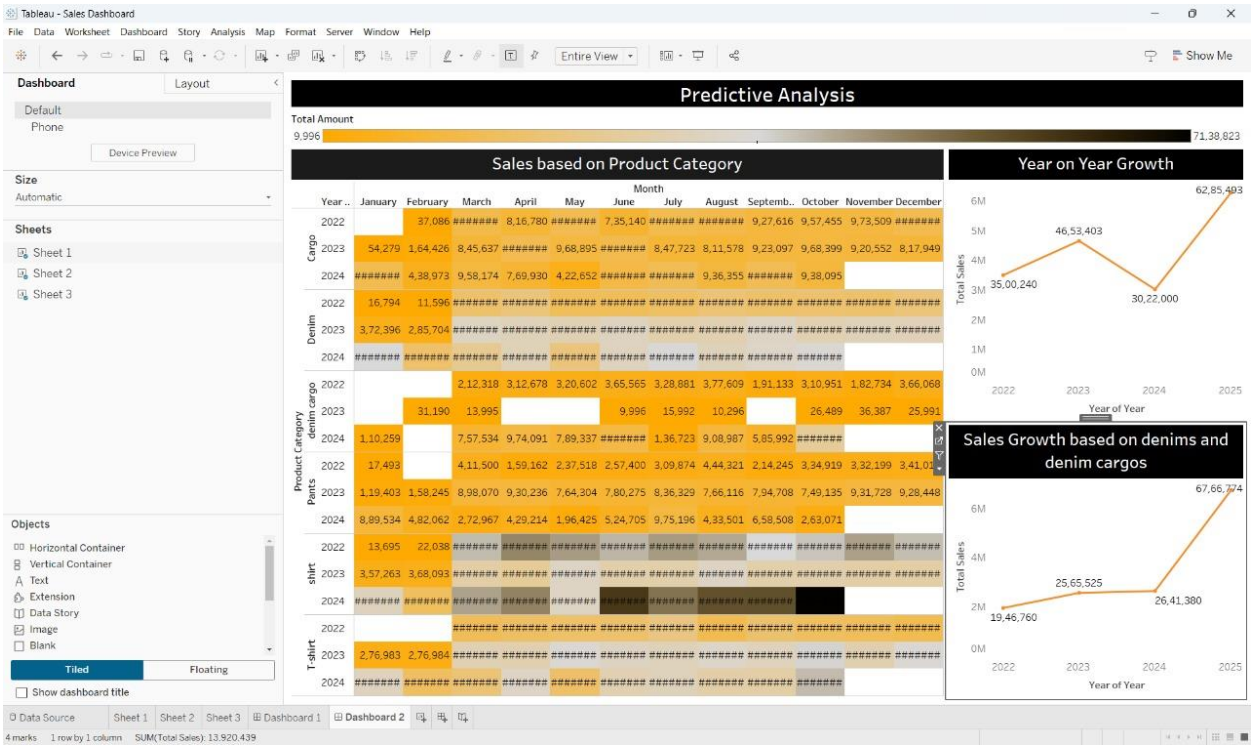


Fig 4.3- Predictive analytics dashboard.

```
mse_test = mean_squared_error(y_test, pred_test)
mse_test
```

```
12087554.455055008
```

```
r2_test = r2_score(y_test, pred_test)
r2_test
```

```
0.7976142243070137
```

Fig 4.3 MSE and R-Sqaured values from python environment.

In this analysis we decided to do predictive analytics to know what the future sales will be and what the sales of denim will be in 2025. We decided to do this analysis as knowing the prediction will help them achieve their vision.

4.2.1. Below are the steps used to do the predictive analysis

Step 1: Importing Python Libraries

We've imported essential libraries such as pandas, numpy, matplotlib, and sklearn to streamline our work with the dataset. These tools are crucial for various tasks: pandas and numpy will help us manipulate and clean the data efficiently, while matplotlib can be used to create visualizations, if needed, to better understand the data. Additionally, we've brought in sklearn, a powerful library for machine learning, which will assist us in building and fine-tuning predictive models later in the process.

Step 2: Basic Data Cleaning

In this step, we started by examining the dataset to understand its structure and content. Commands such as df.head(), df.shape, and df.info() were utilized to explore the first few rows of the data, check the overall dimensions, and review the data types and column information. This helped us identify any potential issues with the dataset, such as missing or irrelevant data.

Next, we addressed null values to ensure the dataset was clean and ready for analysis. Missing values were handled appropriately, either by filling them with relevant values or by dropping rows/columns where necessary. Additionally, we identified and removed unnecessary columns that did not contribute to our analysis, reducing noise and improving the overall clarity of the dataset.

Step 3: Data Transformation

In this step, we focused on transforming the data to make it suitable for analysis and modeling. Specifically, columns with a data type of object, often representing categorical data, were converted into numerical values.

To achieve this, we used Label Encoder, a method from the sklearn library that assigns a unique integer to each category in a column. This transformation ensures that the categorical data is in a numerical format, which is essential for machine learning algorithms to process and interpret the data effectively. This step prepares the dataset for further analysis and predictive modeling.

Step 4: Train-Test Split

To prepare the data for modeling, we used the train_test_split function from the sklearn library to divide the dataset into training and testing sets. The data was split in an 80:20 ratio, where 80% of the data was allocated for training and 20% was reserved for testing.

The training set is used to train our machine learning model, enabling it to learn patterns and relationships within the data. The testing set, consisting of unseen data, is used to evaluate the model's performance. By comparing the model's predictions on the test set with the actual ground truth values, we can calculate metrics like accuracy and gain insights into how well the model generalizes to new data. This split ensures a robust evaluation of the model's predictive capabilities.

Step 5: Model Selection and Training

Given that our target variable is continuous, we chose Linear Regression as our predictive model. This algorithm is well-suited for understanding the relationship between input features and a continuous output variable.

Once the model was selected, we trained it using the training dataset (80% of the data). After training, we evaluated the model's performance on the test dataset (20% of the data). By comparing the model's predictions on the test data with the actual values, we calculated its accuracy, which came out to be 78%.

This indicates that the model is reasonably effective in capturing the patterns in the data, though there may still be room for improvement through further optimization or by exploring alternative models.

Step 6: Final Projections on 2025 Data

After evaluating the model's performance and confirming its accuracy, we used it to make predictions on the 2025 data. The trained Linear Regression model, having learned from the historical patterns in the dataset, was applied to this new data to generate projections for the target variable.

These projections provide insights and estimates for 2025, enabling data-driven decision-making and strategic planning based on the model's outputs. This step demonstrates the practical application of the model to real-world scenarios, showcasing its utility beyond evaluation.

Note: The discrepancy between the predictive analytics graph in the PowerPoint presentation and the sales dashboard figures is due to the use of different datasets at the time of analysis. The dataset for predictive analytics contained inconsistencies and additional data not included in the final dataset used for the sales dashboard. Moreover, the dataset was unclean, with some customer IDs mistakenly included in the sales numbers, which contributed to the mismatch.

V. Implementation

5.1 Implementation Process

The successful implementation of the Sales and Customer Insights & Activity Dashboard for Days for Clothing involves addressing managerial, technical, and ethical implications in a structured and detailed manner. The dashboard is an important tool for the retail store manager and the founder and the CEO, providing actionable insights to improve operations and drive strategic decisions. Here, we discuss the BI implementation process using Kotter's eight-step model for organizational transformation, outline technical challenges, and also touch on ethical considerations.

5.2 Managerial Challenges

Building the Sales Dashboard

Building the sales dashboard involves overcoming organizational challenges and fostering team communication to gain support and ensure alignment with business needs. Using Kotter's eight-step model, the process unfolds as follows:

1. **Establish a Sense of Urgency:**
   DFC currently lacks a centralized data platform and designated personnel for data maintenance. Reliance on MS Excel for data visualization results in fragmented data, inefficient processes, and time-consuming workflows. This hampers a comprehensive understanding of the brand's operations, limiting informed decision-making. Addressing these challenges requires implementing a Business Intelligence (BI) solution. A BI system would enable the CEO, founders, and store managers to centralize and streamline data analysis, enhancing decision-making, boosting efficiency, and driving better outcomes.
2. **Building a Powerful Coalition:**
   Assemble a cross-functional team comprising stakeholders and team members. Hire a IT specialist, or data analyst. This diverse team will integrate technical expertise with business insights, ensuring the project meets operational needs and strategic objectives for successful implementation and make data-driven decisions.
3. **Develop a Vision & Strategy:**
   The vision focuses on creating an E-commerce with a centralized real-time system and BI support-interactive dashboard that consolidates retail data, enhances data accuracy, supports demand forecasting, and enables customer analysis for better decision-making. Communicate this vision clearly, emphasizing how it replaces manual processes with a streamlined, data-driven approach.
4. **Communicate the Vision:**
   Describe the whole plan and the reason behind this shift through meetings, workshops,

and training sessions to introduce stakeholders and team members to the dashboard's capabilities and benefits. These sessions should highlight practical use cases, such as tracking product performance, monitoring customer behavior, and analyzing sales trends.

5. **Empower Others to Act on the Vision:**
   Conduct training sessions to empower store managers to use the dashboard effectively. Customizable views and interactive graphs allow them to adapt the tool to their specific needs without requiring technical expertise.
6. **Generate Short-Term Wins:**
   Reward staff for BI-driven training. Implement a phased rollout, starting with daily sales tracking and inventory management. Store managers will quickly experience time savings and more accurate reporting, boosting their confidence in the system and fostering better decision-making.
7. **Consolidate Gains and Drive More Change:**
   Expand the BI system by integrating additional datasets, such as social media analytics, and automating repetitive tasks to increase functionality. Establish feedback loops to refine features based on user input.
8. **Institutionalize New Approaches:**
   Integrate the Dashboards into daily operations, supporting strategic planning and continuous improvement. Regular updates and data-driven decision making should ensure the system evolves with the business' changing needs.

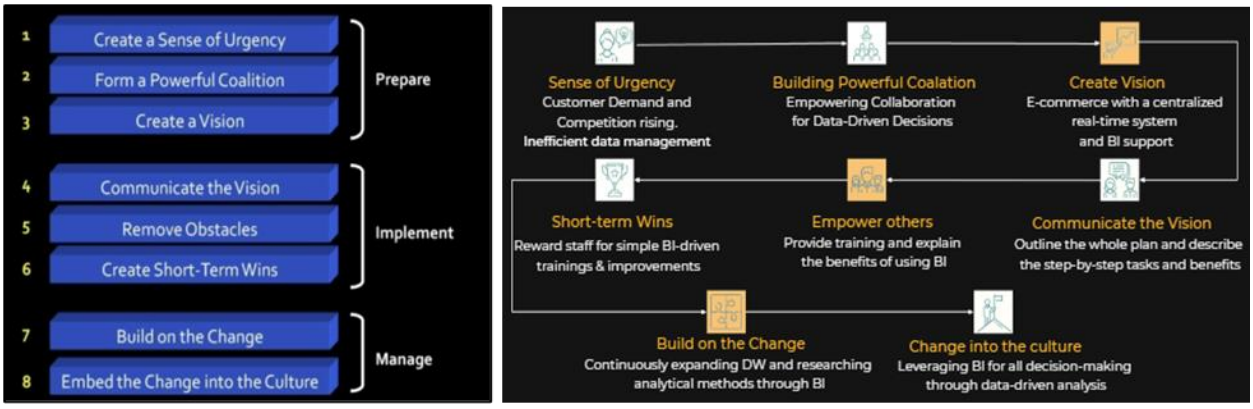


Figure 5.2. Kotter's 8-Step Model showcasing organizational transformation and implementation utilizing Business Intelligence solutions.

5.3 Technical Challenges

5.3.1 Data Quality

One of the major technical challenges encountered during the implementation of the dashboards was ensuring data quality. Retail data frequently contains missing or incomplete fields, such as purchase dates, customer demographics, time stamps, or other essential details required for in-depth analysis. The absence of such critical information can compromise the accuracy of insights

and hinder data-driven decision-making. Moreover, inconsistent data formats from various systems further complicated the integration process. Another significant issue was the occurrence of duplicate records, where multiple entries existed for a single transaction. If left unresolved, this could skew key metrics like total sales or customer counts.

5.3.2 Our Approach for Data Management

To overcome these difficulties, a strong data management approach was required. The first stage was to build up automatic data cleaning routines to discover and repair errors like missing values and duplicate entries before loading the data into the warehouse. These procedures were strengthened by validation tests at each level of the ETL pipeline to ensure data consistency and correctness. In addition, a detailed data governance structure was built, outlining stakeholder roles for ensuring data quality. This involved clearly defining ownership of data integrity and developing rules for addressing disputes. These precautions guaranteed that the dashboard was constructed on dependable and accurate data.

5.4 Ethical Implications

5.4.1 Privacy

Analyzing DFC's sales data presents notable privacy concerns, particularly when sensitive customer information is involved. Sales records often contain personal details such as purchase history, payment methods, contact numbers, and demographic data. Mishandling this information or allowing unauthorized access could breach customer trust and cause significant reputational harm. Additionally, failing to comply with data privacy regulations could expose the organization to legal penalties, including hefty fines. Therefore, safeguarding customer data through secure management practices is essential for the successful implementation of the dashboard.

In contrast to sales transaction data, the customer feedback collection process at DFC inherently addresses privacy concerns, as data privacy terms are explicitly mentioned on the feedback form. We recommend extending this transparency to sales data and other data collection efforts where feasible. Clearly communicating data usage policies can help alleviate customer concerns about data tracking for personalization and remarketing purposes.

5.4.2 Our Approach for Ethical BI

To address data privacy issues, our BI deployment approach must incorporate proactive efforts to protect consumer information. Data anonymization is a critical procedure during data preparation. By removing personally identifying information (PII) from datasets like as names, contact information, and payment details, our study can focus on aggregated insights rather than individual consumer behaviors. This technique reduces the danger of data breaches while preserving vital business information.

Additionally, openness is critical to our data management approach. All key stakeholders, including store managers and analysts, must be fully aware of the extent, purpose, and constraints of data collection and analysis. Clearly defined policies guarantee that data is only utilized for its intended purpose, ensuring responsibility throughout the firm.

Regular compliance audits should be carried out to ensure that data protection laws and ethical standards are followed, therefore reducing legal risks and creating confidence. By adhering to these principles, we want to strike a balance between data-driven decision-making and client privacy, resulting in a safe and transparent business intelligence environment.

VI. Conclusion

6.1 Summary of Work

As highlighted in the previous section about DFC's existing operational difficulties and illustrating how a Business Intelligence (BI) solution can drive growth through data-driven decision-making, the company has opted to use Tableau as its primary BI platform and hire a dedicated data analyst. This strategic approach intends to improve data analytic skills and streamline company procedures.

The installation entails creating a Customer Insights & Activity Dashboard, a Sales Dashboard, and a Predictive Model. These technologies give critical insights into consumer behavior, sales success, and market trends. DFC's owners, shop managers, and team members may use this data to make educated decisions about staff scheduling during peak hours, inventory management, replenishing popular goods, and creating efficient sales campaigns.

By embracing these sophisticated BI tools, DFC will be better positioned to understand its clients, improve the shopping experience, and boost its market presence as India's leading street-style clothes brand. In the long term, this data-driven strategy might help it achieve its goal of becoming a leading denim brand in the competitive fashion business.

6.2 Recommendations for the Days for Clothing

To make the most of its new BI tools, we advise DFC to implement the following strategic measures:

1. Extend Data Utilization: Build upon the existing data architecture by incorporating data from other sources, like social media involvement, online buying habits, and consumer loyalty programs. Deeper insight into company expansion will be made possible by the more thorough understanding of consumer preferences and purchasing trends that will result from this.

2. Enhance Customer Engagement: Use dashboard analytics to create customized marketing campaigns, new product releases, and in-store activities. To sustain steady foot traffic and increase sales, target slower sales periods with promotional offers or interactive customer experiences.

3. Empower the Team: Provide continual training for store managers, team leaders, and marketing personnel to analyze and act on dashboard findings. This will allow them to make confident data-driven decisions, resulting in increased operational efficiency and customer happiness.

4. Adopt a Data-Driven Mindset: Continuously analyze and adapt business strategies based on current data patterns. Continuously analyze sales performance, inventory levels, and consumer feedback to fine-tune marketing activities, improve product assortments, and remain ahead of market trends.

5. Ensure Data Privacy and Security: follow best practices for secure data management. To gain and maintain customer trust, implement explicit data governance policies, verify legal compliance, and update security processes on a regular basis.

These steps will help Days for Clothing make the most of its new BI tools, ensuring it stays competitive and continues to thrive in India's dynamic retail environment.

VII. References

Inditex Group Annual Report 2023. https://annualreport2023.inditex.com/en

Urban Outfitters Press Releases. https://investor.urbn.com/financial-news-and-events/press-releases

Kotter, J. P. (1996). Leading Change. Harvard Business Review Press.

Tableau Software. "Dashboard Creation Best Practices." https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-home.htm

European Union. "General Data Protection Regulation (GDPR) Overview." https://gdpr-info.eu/

Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. Wiley.