



Capstone Project: Global Fashion Retail Analytics

Analytics Journey Unveiled

Presented by Akshay Solanki, Data Analyst focused
on delivering actionable insights

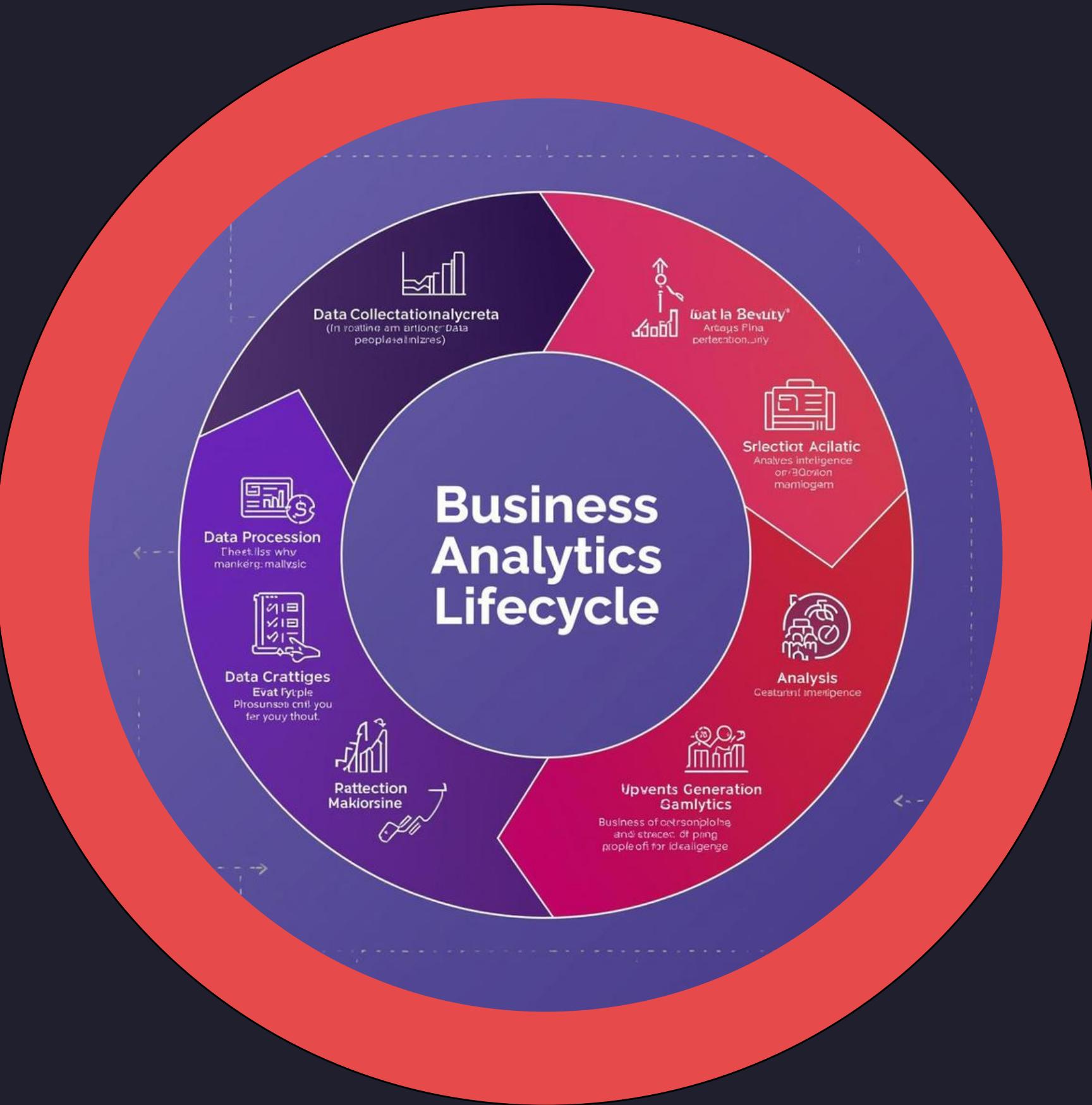


Global Fashion Retail Analytics



In today's competitive market, **fashion retail analytics** is essential for driving growth and efficiency. By leveraging SQL, Power BI, and machine learning, businesses can transform raw data into actionable insights. This approach enables strategic decision-making, optimizing inventory management, and enhancing customer experiences. Understanding these analytics tools is crucial for professionals aiming to excel in the fast-paced retail environment.

Data to Insights

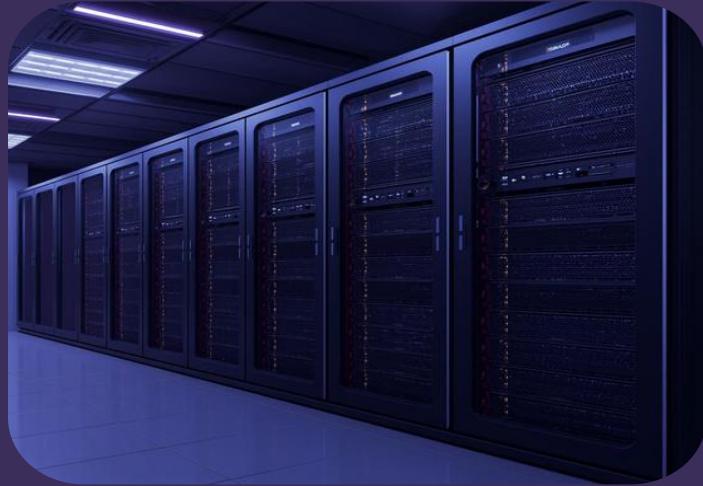


This section explores the **end-to-end analytics journey**, focusing on how SQL, Power BI, and Machine Learning converge to transform raw data into actionable insights. By leveraging advanced techniques, businesses can enhance decision-making, optimize operations, and drive strategic growth. Understanding each phase in this lifecycle is crucial for effectively utilizing analytics to gain a competitive edge in the global fashion retail industry.

Part 1: Advanced SQL Analysis – Retail Sales

🎯 Objective

Use advanced SQL to extract business-critical insights from Global Fashion Retail's operations, enabling leadership to make informed strategic decisions..



Data Extraction

Efficiently gathers data from various sources.

Data Transformation

Prepares data for analysis and reporting.

Insights Generation

Delivers actionable intelligence for decision-making.

SQL Enhancements for Faster Query Execution

```
1  -- Convert raw Date (various formats) to a clean DATE column for reliable filtering & calculations
2  ALTER TABLE transactions
3  ADD COLUMN tx_date DATE GENERATED ALWAYS AS (
4      DATE(
5          COALESCE(
6              STR_TO_DATE(Date, '%Y-%m-%d %H:%i:%s'),
7              STR_TO_DATE(Date, '%d/%m/%Y %H:%i:%s'),
8              STR_TO_DATE(Date, '%m/%d/%Y %H:%i:%s'),
9              STR_TO_DATE(Date, '%Y-%m-%d'),
10             STR_TO_DATE(Date, '%d/%m/%Y'),
11             STR_TO_DATE(Date, '%m/%d/%Y')
12         )
13     )
14 ) STORED;
```

SQL Enhancements for Faster Query Execution

```
16  -- Index to speed up date-based queries (range, filtering, sorting)
17 • CREATE INDEX idx_tx_date ON transactions(tx_date);
18
19  -- Index discount date range for faster join on promotion periods
20 • CREATE INDEX idx_discount_dates ON discounts(start_date, end_date);
21
22  -- Customer + Date index to accelerate first purchase, history, and lifetime value queries
23 • CREATE INDEX idx_customer_id ON transactions(Customer_ID, tx_date);
24
25  -- Precompute Month (YYYY-MM) for easy reporting and grouping
26 • ALTER TABLE transactions
27   ADD COLUMN month_period CHAR(7) GENERATED ALWAYS AS (DATE_FORMAT(tx_date, '%Y-%m')) STORED;
28
29  -- Index for quick month-based summary reports
30 • CREATE INDEX idx_month_period ON transactions(month_period);
31
32  -- Store + Transaction Type index to speed up store-level sales/returns analysis
33 • CREATE INDEX idx_transactions_store_type ON transactions(Store_ID, Transaction_Type(50));
34
```

1. Customer Demographics Analysis

Retrieve the number of unique customers by country and city along with their average purchase value.

```
1  /*1. Customer Demographics Analysis
2   Retrieve the number of unique customers by country and city along with their average purchase value.*/
3  SELECT
4    c.Country,
5    c.City,
6    COUNT(DISTINCT t.Customer_ID) AS unique_customers,
7    ROUND(AVG(t.Invoice_Total), 2) AS avg_purchase_value
8  FROM
9    transactions t
10   JOIN
11     customers c ON t.Customer_ID = c.Customer_ID
12   GROUP BY c.Country , c.City
13   ORDER BY unique_customers DESC;
14
```

Result Grid | Filter Rows: Export: Wrap Cell Content:

	Country	City	unique_customers	avg_purchase_value
▶	中国	深圳	55941	732.33
	United States	New York	48565	95.11
	中国	北京	46232	739.37
	United States	Los Angeles	43913	95.14
	中国	上海	40976	735.1
	United States	Chicago	36521	96.26
	United States	Houston	33731	95.66

2. Sales vs Returns Trends

Calculate the percentage of returns compared to total transactions and track how it changes monthly.

```
1  /*2. Sales vs Returns Trends
2   Calculate the percentage of returns compared to total transactions and track how it changes monthly.*/
3   SELECT
4     month_period AS tx_month,
5     COUNT(*) AS total_lines,
6     SUM(CASE
7       WHEN Transaction_Type = 'Return' THEN 1
8       ELSE 0
9     END) AS return_lines,
10    ROUND(100 * SUM(CASE
11      WHEN Transaction_Type = 'Return' THEN 1
12      ELSE 0
13    END) / COUNT(*),2) AS pct_returns
14   FROM
15     transactions
16   GROUP BY month_period
17   ORDER BY month_period;|
```

Result Grid | Filter Rows: Export: Wrap Cell Content:

tx_month	total_lines	return_lines	pct_returns
2023-01	171759	9256	5.39
2023-02	119380	6501	5.45
2023-03	264005	14322	5.42
2023-04	187136	9986	5.34
2023-05	213135	11215	5.26

3. Top Performing Categories

Identify the top categories by total revenue and their share in overall sales.

```
1  /*3. Top Performing Categories
2   Identify the top categories by total revenue and their share in overall sales.*/
3  SELECT
4      p.Category, SUM(t.Line_Total) AS revenue
5  FROM
6      transactions t
7      JOIN
8      products p ON t.Product_ID = p.Product_ID
9  GROUP BY p.Category
10 ORDER BY revenue DESC;
11
```

The screenshot shows a database query results grid. At the top, there are navigation buttons for 'Result Grid' (highlighted), 'Filter Rows:', 'Export:' (with a CSV icon), and 'Wrap Cell Content:'. The results grid has two columns: 'Category' and 'revenue'. The data rows are:

	Category	revenue
▶	Feminine	340685345.8434937
	Masculine	326979394.612859
	Children	65080134.56652975

4. Payment Method Performance

Compare total sales volume and average basket value across payment methods.

```
1  /* 4. Payment Method Performance
2   Compare total sales volume and average basket value across payment methods.*/
3  • SELECT
4    Payment_Method,
5    COUNT(DISTINCT Invoice_ID) AS num_invoices,
6    SUM(Invoice_Total) AS total_sales,
7    ROUND(SUM(Invoice_Total) / COUNT(DISTINCT Invoice_ID),
8          2) AS avg_basket_value
9  FROM
10   transactions
11 GROUP BY Payment_Method
12 ORDER BY total_sales DESC;
13
```

Result Grid				
	Payment_Method	num_invoices	total_sales	avg_basket_value
▶	Credit Card	3633549	1251770773.7344487	344.5
	Cash	906855	310897563.13973475	342.83

5. Promotion Impact Analysis

Compare sales volume and revenue during discount periods versus non-discount periods.

```
1  /*#5. Promotion Impact Analysis
2   #Compare sales volume and revenue during discount periods versus non-discount periods.*/
3  •  SELECT
4    CASE
5      WHEN d.start_date IS NOT NULL THEN 'Promo'
6      ELSE 'Non-Promo'
7    END AS promo_status,
8    COUNT(DISTINCT t.Invoice_ID) AS num_invoices,
9    SUM(COALESCE(t.Invoice_Total, 0)) AS total_revenue,
10   SUM(COALESCE(t.Line_Total, 0)) AS total_line_revenue,
11   ROUND(AVG(COALESCE(t.Invoice_Total, 0)), 2) AS avg_invoice_value
12  FROM
13    transactions t
14    LEFT JOIN
15      discounts d ON t.tx_date BETWEEN d.start_date AND d.end_date
16  GROUP BY promo_status
17  ORDER BY promo_status DESC;
18
```

Result Grid Filter Rows: <input type="text"/> Export: Wrap Cell Content:					
	promo_status	num_invoices	total_revenue	total_line_revenue	avg_invoice_value
▶	Promo	1857086	2449226511.479398	1147563810.8861947	216.07
	Non-Promo	2685256	1058590401.7219491	496213959.0573331	279.06

6. Customer Loyalty Tracking

Measure the proportion of new versus repeat customers each month along with their average spend.

```
1 /*6. Customer Loyalty Tracking
2 Measure the proportion of new versus repeat customers each month along with their average spend.*/
3 With ranked AS (
4     SELECT
5         Customer_ID,
6         Invoice_Total,
7         month_period,
8         MIN(tx_date) OVER (PARTITION BY Customer_ID) AS first_purchase_date
9     FROM transactions
10 )
11     SELECT
12         month_period,
13         CASE
14             WHEN month_period = DATE_FORMAT(first_purchase_date, '%Y-%m')
15                 THEN 'New'
16             ELSE 'Repeat'
17         END AS customer_type,
18         COUNT(DISTINCT Customer_ID) AS num_customers,
19         SUM(Invoice_Total) AS total_spend,
20         AVG(Invoice_Total) AS avg_spend_per_invoice
21     FROM ranked
22     GROUP BY month_period, customer_type
23     ORDER BY month_period, customer_type;
```

	month_period	customer_type	num_customers	total_spend	avg_spend_per_invoice
▶	2023-01	New	110456	29393439.400794506	171.13187315246657
	2023-02	New	68575	21914778.7212615	209.8553905203729
	2023-02	Repeat	9485	2512002.2095947266	168.0044281430395
	2023-03	New	134769	52939415.45378351	245.7874219603946
	2023-03	Repeat	29144	11874294.467366695	244.2365886578365
	2023-04	New	83311	34984542.5	271.49686088562606
	2023-04	Repeat	36317	15390286.48936367	264.0839851979078
	2023-05	New	85055	32164343.301448822	243.83921597968904
	2023-05	Repeat	50180	19702761.458032608	242.56418996186747
	2023-06	New	47897	21478472	297.59705152896515
	2023-06	Repeat	38465	17599974.04997921	293.0008332220019
	2023-07	New	49742	18685070.79996109	247.47127039575506
	2023-07	Repeat	46258	17695034.549646378	245.46776186616697
	2023-08	New	39920	16363937.398790836	273.35645388287986
	2023-08	Repeat	44196	18262958.251724243	266.4374972897256

7. Store Performance Ranking

Identify the top and bottom stores by revenue and return rates.

```
1 /*7. Store Performance Ranking
2 Identify the top and bottom stores by revenue and return rates.*/
3 • WITH store_metrics AS (
4     SELECT
5         s.Store_Name,
6         SUM(IF(t.Transaction_Type = 'Sale', t.Invoice_Total, 0)) AS total_revenue,
7         SUM(IF(t.Transaction_Type = 'Return', ABS(t.Invoice_Total), 0)) AS return_revenue,
8         COUNT(*) AS total_transactions,
9         SUM(IF(t.Transaction_Type = 'Return', 1, 0)) AS return_transactions,
10        ROUND(
11            SUM(IF(t.Transaction_Type = 'Return', 1, 0)) / COUNT(*) * 100, 2
12        ) AS return_rate_percent
13    FROM transactions t
14    JOIN stores s ON t.Store_ID = s.Store_ID
15    GROUP BY s.Store_Name
16),
17     ranked AS (
18         SELECT
19             Store_Name,
20             total_revenue,
21             return_revenue,
22             total_transactions,
23             return_transactions,
24             return_rate_percent,
25             ROW_NUMBER() OVER (ORDER BY total_revenue DESC) AS rank_top,
26             ROW_NUMBER() OVER (ORDER BY total_revenue ASC) AS rank_bottom
27     FROM store_metrics
28 )
```

```

29   SELECT
30     'Top' AS category,
31     Store_Name,
32     total_revenue,
33     return_revenue,
34     return_rate_percent,
35     rank_top AS rank_order
36   FROM ranked
37   WHERE rank_top <= 5
38
39 UNION ALL
40
41 SELECT
42   'Bottom' AS category,
43   Store_Name,
44   total_revenue,
45   return_revenue,
46   return_rate_percent,
47   rank_bottom AS rank_order
48   FROM ranked
49   WHERE rank_bottom <= 5
50   ORDER BY category, rank_order;

```

	category	Store_Name	total_revenue	return_revenue	return_rate_percent	rank_order
▶	Bottom	Store Glasgow	5117829.230078697	318819.60009384155	5.28	1
	Bottom	Store Liverpool	5732306.470118165	373706.39006757736	5.40	2
	Bottom	Store Braga	6110367.291038752	376210.62075042725	5.17	3
	Bottom	Store Valencia	6439399.218827248	408803.1993961334	5.39	4
	Bottom	Store Frankfurt am Main	6916802.560347319	430705.1000547409	5.26	5
	Top	Store 上海	300431731.3803978	18394811.96394348	5.22	1
	Top	Store 广州	293774875.9663429	18091074.676950455	5.23	2
	Top	Store 深圳	247006028.57431984	15734282.040592194	5.35	3
	Top	Store 北京	226342833.35508537	14726802.675647736	5.28	4
	Top	Store 重庆	158202086.87786293	9859128.769182205	5.32	5

8. Price Band Contribution

Segment products into price bands and analyze their contribution to total revenue and returns.

```
1  /*8. Price Band Contribution
2   Segment products into price bands and analyze their contribution to total revenue and returns.*/
3 • SELECT
4   CASE
5     WHEN t.Unit_Price < 50 THEN 'Low'
6     WHEN t.Unit_Price BETWEEN 50 AND 150 THEN 'Mid'
7     ELSE 'High'
8   END AS price_band,
9
10  SUM(CASE WHEN t.Transaction_Type = 'Sale' THEN t.Line_Total ELSE 0 END) AS total_revenue,
11
12  SUM(CASE WHEN t.Transaction_Type = 'Return' THEN ABS(t.Line_Total) ELSE 0 END) AS total_returns,
13
14  COUNT(*) AS total_transactions,
15
16  SUM(CASE WHEN t.Transaction_Type = 'Return' THEN 1 ELSE 0 END) AS return_transactions,
17
18  ROUND(
19    SUM(CASE WHEN t.Transaction_Type = 'Return' THEN 1 ELSE 0 END) /
20    COUNT(*) * 100, 2
21  ) AS return_rate_percent
22 FROM transactions t
23 GROUP BY price_band
24 ORDER BY total_revenue DESC;
```

	price_band	total_revenue	total_returns	total_transactions	return_transactions	return_rate_percent
	High	561441494.364975	31275577.501487732	1467221	77408	5.28
	Mid	125452513.89552498	7025771.970287323	1824245	96431	5.29
▶	Low	89166514.39432836	5014298.1601707935	3125361	165788	5.30

9. Basket Size Insights

Calculate the average number of items and revenue per basket across different countries.

```
1 /*9. Basket Size Insights
2 Calculate the average number of items and revenue per basket across different countries.*/
3 WITH basket_level AS (
4     SELECT
5         t.Invoice_ID,
6         c.Country,
7         SUM(t.Quantity) AS total_items,
8         MAX(t.Invoice_Total) AS total_revenue -- take MAX once per invoice
9     FROM transactions t
10    JOIN customers c
11      ON t.Customer_ID = c.Customer_ID
12    GROUP BY t.Invoice_ID, c.Country
13 )
14     SELECT
15         Country,
16         COUNT(*) AS num_baskets, -- number of baskets per country
17         ROUND(AVG(total_items),2) AS avg_items_per_basket, -- average number of items per basket
18         ROUND(AVG(total_revenue),2) AS avg_revenue_per_basket -- average revenue per basket
19     FROM basket_level
20    GROUP BY Country
21    ORDER BY Country;
```

Country	num_baskets	avg_items_per_basket	avg_revenue_per_basket
Deutschland	530596	1.56	57.28
España	406241	1.55	57.75
France	459272	1.56	57.37
Portugal	410517	1.55	57.37
United Kingdom	443718	1.56	40.3
United States	1192732	1.55	63.61
中国	1097328	1.56	491.84

10. Monthly Growth Trends

Analyze month-over-month revenue growth overall and for the top three categories.

```
1  /*10. Monthly Growth Trends
2   Analyze month-over-month revenue growth overall and for the top three categories.*/
3   -- Step 1: Calculate overall monthly revenue
4   WITH monthly_revenue AS (
5     SELECT
6       month_period,
7       SUM(Invoice_Total) AS total_revenue
8     FROM transactions
9     GROUP BY month_period
10   ),
11
12   -- Step 2: Calculate monthly revenue per category
13   category_monthly AS (
14     SELECT
15       t.month_period,
16       p.Category,
17       SUM(t.Invoice_Total) AS category_revenue
18     FROM transactions t
19     JOIN products p ON t.Product_ID = p.Product_ID
20     GROUP BY t.month_period, p.Category
21   ),
22
23   -- Step 3: Rank categories per month by revenue
24   ranked_categories AS (
25     SELECT
26       *,
27       RANK() OVER (PARTITION BY month_period ORDER BY category_revenue DESC) AS rnk
28     FROM category_monthly
29   ),
30
31   -- Step 4: Keep top 3 categories per month and compute MoM growth for each
32   category_growth AS (
33     SELECT
34       month_period,
35       Category,
36       category_revenue,
37       ROUND(
38         (category_revenue - LAG(category_revenue) OVER (PARTITION BY Category ORDER BY month_period)) /
39         LAG(category_revenue) OVER (PARTITION BY Category ORDER BY month_period) * 100, 2
40       ) AS category_mom_growth
41     FROM ranked_categories
42     WHERE rnk <= 3
43   ),
```

```

45  -- Step 5: Compute overall MoM growth
46  overall_growth AS (
47      SELECT
48          month_period,
49          total_revenue,
50          ROUND(
51              (total_revenue - LAG(total_revenue) OVER (ORDER BY month_period)) /
52              LAG(total_revenue) OVER (ORDER BY month_period) * 100, 2
53          ) AS overall_mom_growth
54      FROM monthly_revenue
55  )
56
57  -- Step 6: Combine overall + category results
58  SELECT
59      o.month_period,
60      o.total_revenue,
61      o.overall_mom_growth,
62      c.Category,
63      c.category_revenue,
64      c.category_mom_growth
65  FROM overall_growth o
66  LEFT JOIN category_growth c
67  ON o.month_period = c.month_period
68  ORDER BY o.month_period, c.category_revenue DESC;

```

Result Grid | Filter Rows: _____ | Export: Wrap Cell Content:

	month_period	total_revenue	overall_mom_growth	Category	category_revenue	category_mom_growth
▶	2023-01	29393439.400794506	NULL	Feminine	13381403.869159698	NULL
	2023-01	29393439.400794506	NULL	Masculine	12426312.361250877	NULL
	2023-01	29393439.400794506	NULL	Children	3585723.17038393	NULL
	2023-02	24426780.930856228	-16.9	Feminine	11364591.890260696	-15.07
	2023-02	24426780.930856228	-16.9	Masculine	10833070.700621605	-12.82
	2023-02	24426780.930856228	-16.9	Children	2229118.3399739265	-37.83
	2023-03	64813709.92115021	165.34	Feminine	28872348.195737362	154.06
	2023-03	64813709.92115021	165.34	Masculine	26953989.4027009	148.81
	2023-03	64813709.92115021	165.34	Children	8987372.322711945	303.18
	2023-04	50374828.98936367	-22.28	Feminine	23840558.51971054	-17.43
	2023-04	50374828.98936367	-22.28	Masculine	22119399.38944149	-17.94
	2023-04	50374828.98936367	-22.28	Children	4414871.080211639	-50.88
	2023-05	51867104.75948143	2.96	Feminine	23180216.687903404	-2.77
	2023-05	51867104.75948143	2.96	Masculine	21969968.49178505	-0.68
	2023-05	51867104.75948143	2.96	Children	6716919.579792976	52.14
	2023-06	39078446.04997921	-24.66	Feminine	18668503.75001049	-19.46
	2023-06	39078446.04997921	-24.66	Masculine	16963942.86997509	-22.79
	2023-06	39078446.04997921	-24.66	Children	3445999.4299936295	-48.7

Data Visualization



Power BI offers **dynamic insights** through interactive data visualization tools, empowering businesses to make informed decisions. By leveraging advanced analytics and machine learning, users can easily identify trends and patterns. This approach enhances strategic thinking, enabling organizations to adapt to market changes effectively. With a user-friendly interface, Power BI transforms complex data into clear visuals, fostering collaboration among stakeholders for optimized outcomes.

Part 2: Interactive Dashboards – Power BI

🎯 Objective

Build a multi-page interactive dashboard to visualize sales trends, product performance, promotions, and customer behavior, enabling executives to drill down into specific metrics.



Data Extraction

Efficiently gathers data from various sources.



Data Transformation

Prepares data for analysis and reporting.



Insights Generation

Delivers actionable intelligence for decision-making.

Fashion Retail Trends & Growth Tracker Sales & Returns Overview

732.74M

Total Revenue

43.32M

Total Returns

4.47%

Return Rate % by Order

1.39

Average Basket Size

4.54M

Total Transactions (Net)

Color

BEIGE

BLACK

BLUE

Transaction Type

Return

Sale

Year, YearMonth

2023

2024

2025

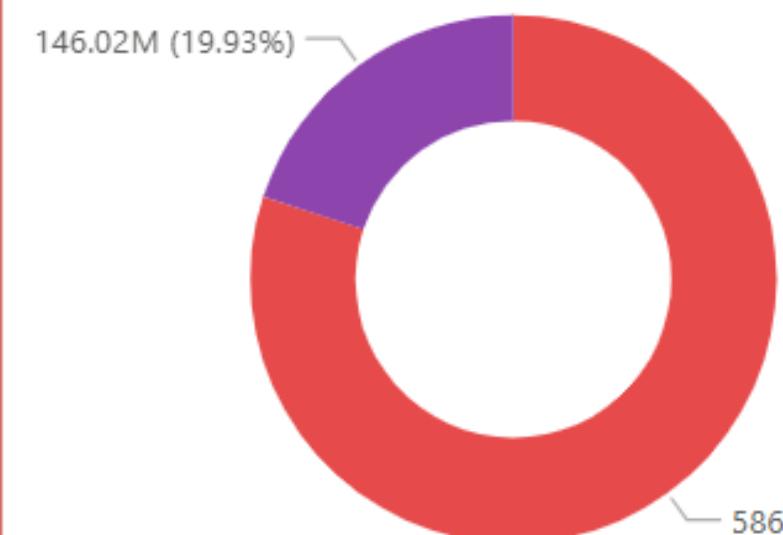
Gender

D

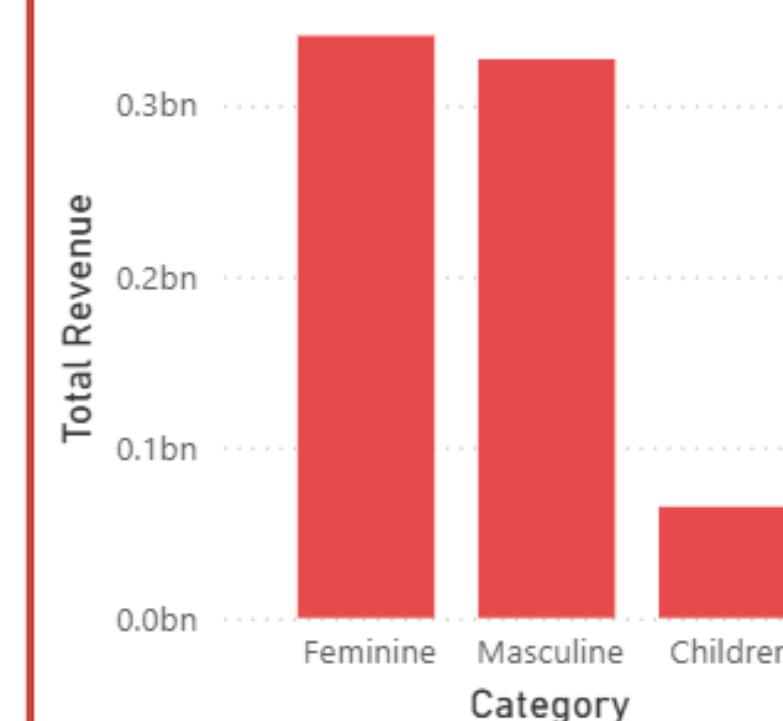
F

M

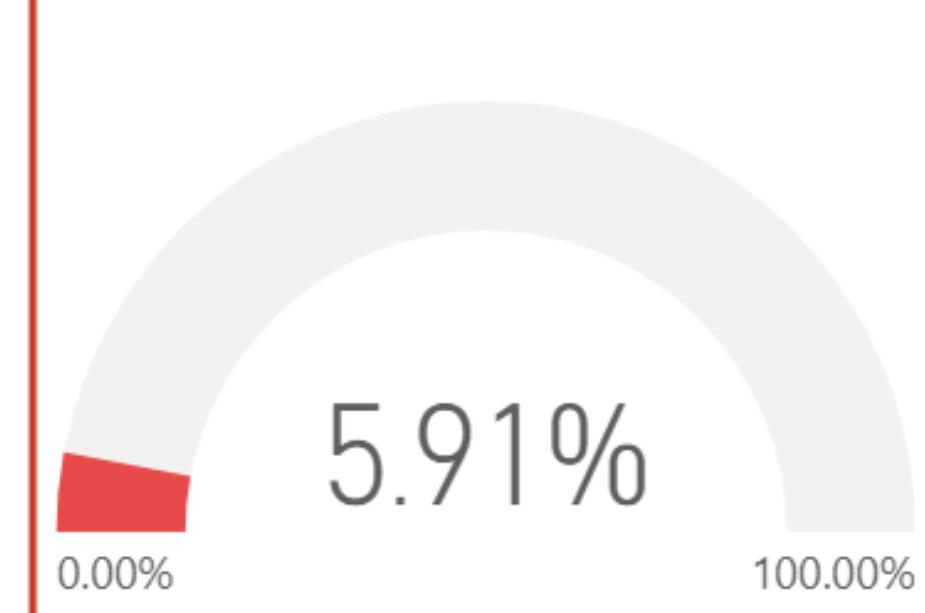
Total Revenue by Payment Method



Total Revenue by Category



Return Rate % by revenue



Total Revenue by YearMonth



Geographic Insights

Category & Product...

Promotion Effectiveness

Customer Behavior

Sales & Returns Overview

Key Insights –

- ₹732.74M total revenue with ₹43.32M total returns ($\approx 4.47\%$ by order).
- Feminine and Masculine categories dominate sales, while Children's contributes the least.
- Credit cards account for 80% of total revenue; cash only 20%, indicating strong digital adoption.
- Average basket size is 1.39 items per transaction across 4.54M total transactions.
- Return rate by revenue: 5.91%, suggesting moderate product returns and potential sizing or satisfaction issues.
- Revenue trend shows fluctuations with clear seasonal peaks, implying high activity during promotional or festive periods.

Suggestions

- Investigate product return reasons, especially in top-selling categories (Feminine & Masculine).
- Use return analytics to optimize product descriptions, sizing guides, and quality checks.
- Encourage cashless payments further through credit/debit card discounts or rewards.
- Monitor basket size trends—bundle complementary products to raise average order size.
- Align sales campaigns with peak months to maximize seasonal demand.

Fashion Retail Trends & Growth Tracker Geographic Insights

1.64M

Customers by Region

2.43%

Revenue MoM %

732.74M

Total Revenue

Year, YearMonth

2023

2024

Payment Meth...

Cash

Credit Card

Country, City

ä,ä½

Deutschland

Customer Type

New

Non Buyer

Year	Total Revenue	Revenue LY	Revenue Growth %
2023	30,49,55,981.03		0.00
2024	38,29,07,128.76	30,49,55,981.03	0.26
2025	4,48,81,765.20	6,96,86,448.94	-0.36

Average Spend per Region by City



⚠ This visual type is being retired soon....

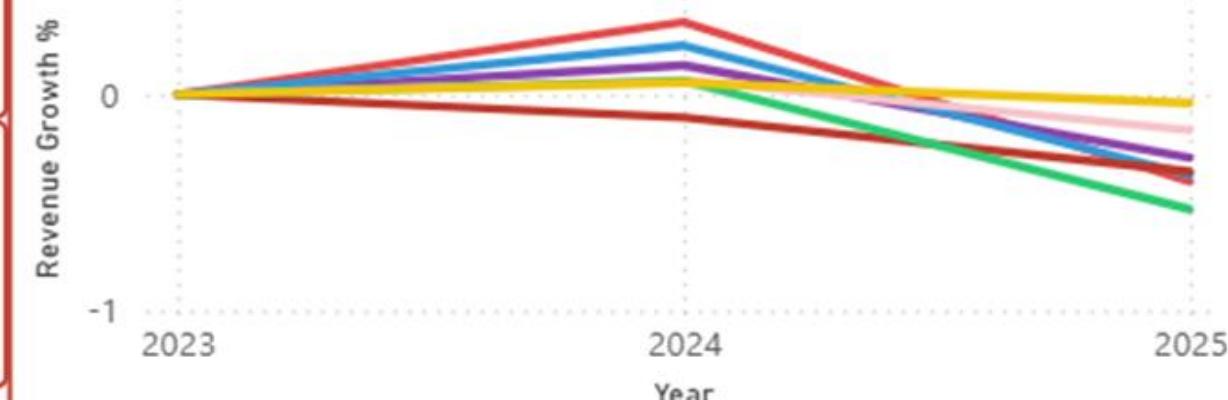
Customers by Region by Country and City

City ● USA ● UK ● Spain ● France



Revenue Growth % by Year and Country

Country ● USA ● Deutschland ● Spain ● France ● Portugal



Revenue MoM % by YearMonth

Revenue MoM %



Sales & Returns Overview

Category & Product Insights

Promotion Effectiveness

Customer Behavior

Geographic Insights

Key Insights –

₹732.7M total revenue with 1.64M customers; +2.43% MoM growth.

2024 saw +26% YoY growth, while 2025 shows a slight dip (-0.36%).

Europe, especially Germany, drives the majority of sales.

Revenue shows strong seasonal spikes aligned with promotions.

Suggestions

Focus marketing and inventory efforts on peak months and European markets.

Investigate 2025's early decline — potential product mix or campaign timing issue.

Use predictive analytics to forecast demand seasonality.

Category & Product Insights

Key Insights –

₹340.7M revenue across 18K products from 4.34M transactions.

Feminine category leads with 46% share, followed by Masculine (45%) and Children (9%).

Average return rate is 5.6%, consistent across all categories.

Monthly sales patterns show stable performance with periodic spikes.

Suggestions

Prioritize Feminine and Masculine categories for new product launches.

Target Children's category for growth—expand assortment or run targeted campaigns.

Maintain current return control strategies to sustain customer satisfaction.

Fashion Retail Trends & Growth Tracker Promotion Effectiveness

732.74M

Total Revenue

26.10M

Total Revenue Promo

432.69K

Lift in Sales (Amount)

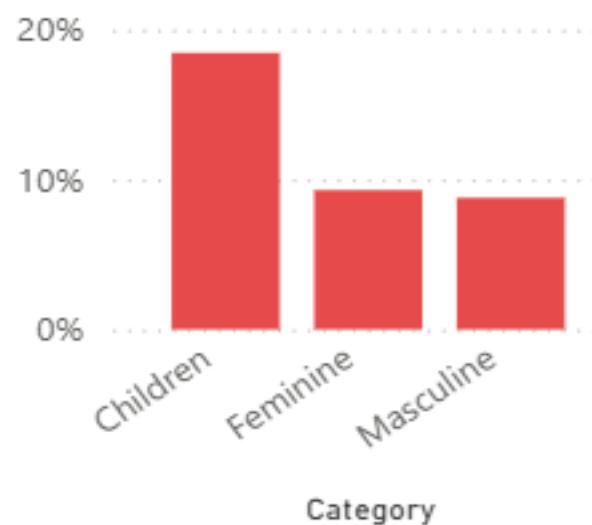
3.56%

% Sales Under Promo

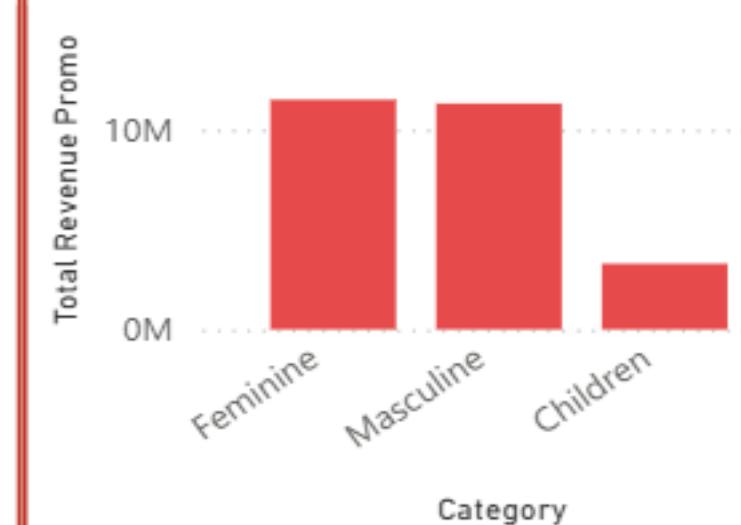
9.89%

Lift in Sales %

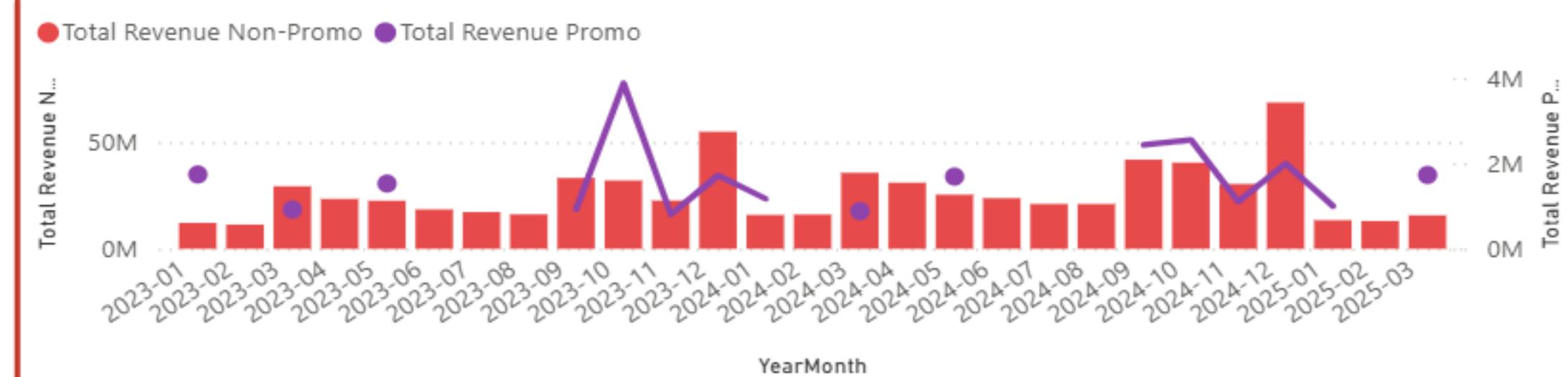
Lift in Sales % by Category



Total Revenue Promo by Category



Total Revenue Non-Promo and Total Revenue Promo by YearMonth



Date

All

DayOfWeek	01-01-2023	15-03-2023	01-05-2023	01-09-2023	01-10-2023	27-11-2023	20-12-2023	01-01-2024
Monday								
Tuesday								
Wednesday								
Friday								
Saturday								
Sunday								

Category, Sub Category

All

Customer Type

New

Non Buyer

Sales & Returns Overview

Geographic Insights

Category & Product Insights

Customer Behavior

Promotion Effectiveness

Key Insights –

₹26.1M of total revenue (3.56%) came from promotions, adding ₹432.7K in incremental sales.

Promotions lifted total sales by 9.9%, with the Children's category seeing the highest uplift (~18%).

Feminine and Masculine categories contribute most to promo revenue.

Sales spikes coincide with promo periods, especially on weekends.

Suggestions

Run targeted weekend promotions for maximum traction.

Increase promo investment in Children's products where lift is highest.

Conduct ROI analysis to identify underperforming campaigns and optimize spend.

Fashion Retail Trends & Growth Tracker Customer Behavior

1.64M

Total Customers

15.75%

% of New Customers

62.37%

% of Repeat Custom...

21.88%

% of Non Buyer

168.94

AOV (Net)

1.46

Average Items per Order

Customer Type

- New
- Non Buyer
- Repeat

Country, City

- USA
- Deutschland

Store ID

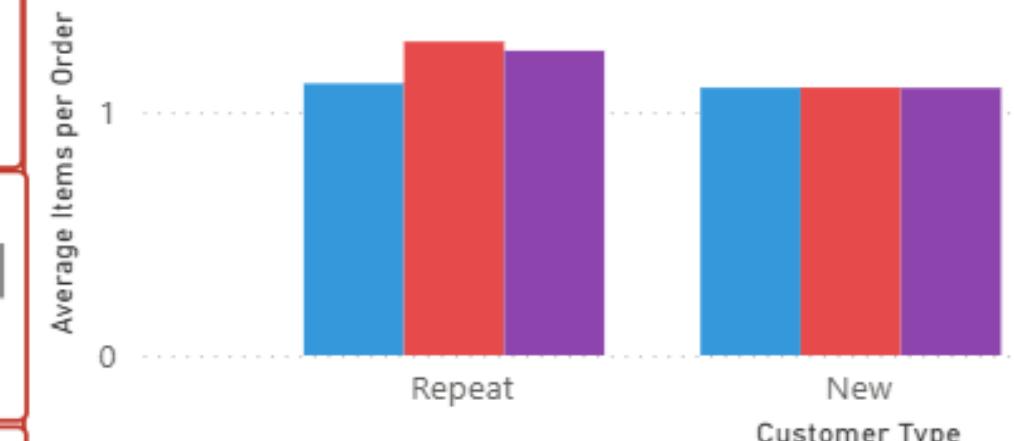
All

YearMonth

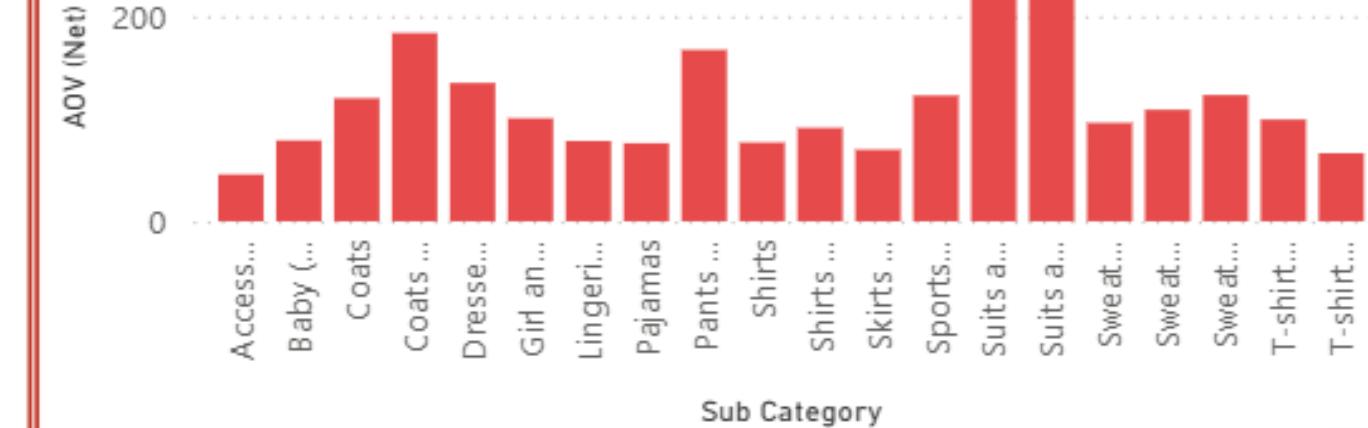
All

Average Items per Order by Customer Type and Category

Category ● Children ● Feminine ● Masculine

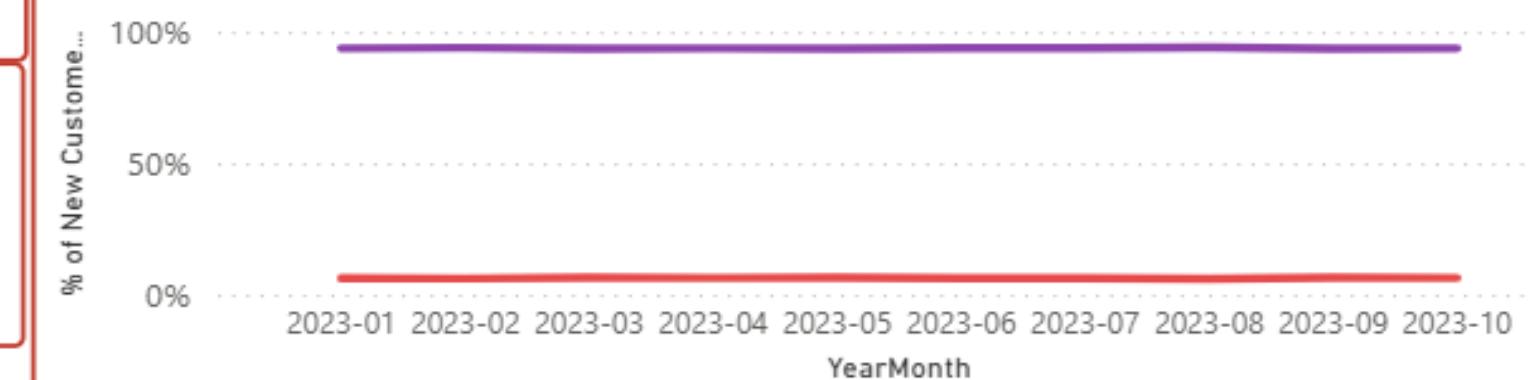


AOV (Net) by Sub Category

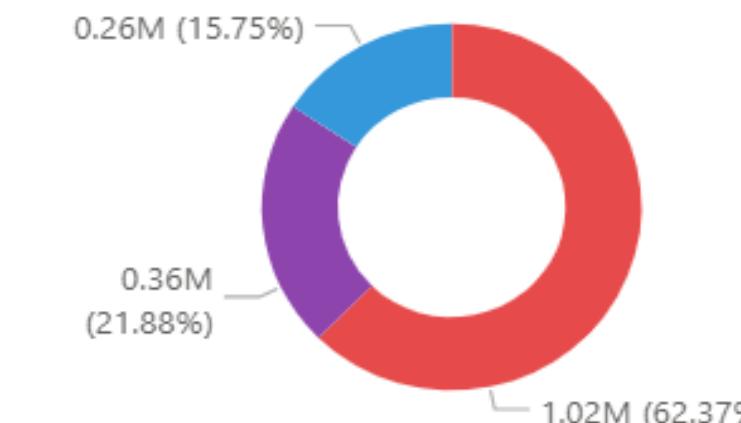


% of New Customers, % of Repeat Customers and % of Non Buyer by YearMonth

● % of New Customers ● % of Repeat Customers ● % of Non Buyer



Total Customers by Customer Type



Customer Type
● Repeat
● Non Buyer
● New

Sales & Returns
Overview

Geographic Insights

Category & Product
Insights

Promotion Effectiveness

Customer Behavior

Key Insights –

1.64M Total Customers:

62.37% Repeat, 15.75% New, and 21.88% Non-Buyers.

Repeat buyers show stronger engagement with higher AOV (₹168.9) and better basket size (1.46 items/order).

Feminine & Children categories show consistent preference among repeat customers.

Customer mix is stable month-over-month, indicating good retention, but new acquisition remains moderate.

Suggestions

Strengthen Retention with loyalty benefits, early access, or exclusive offers for repeat buyers.

Boost New Customer Acquisition via targeted promotions, first-time discounts, or influencer campaigns.

Re-engage Non-Buyers (21.8%) through personalized nudges, abandoned cart reminders, or welcome deals.

Predictive Analytics



Machine Learning for Predictive Analytics involves using algorithms to analyze data, identify patterns, and make predictions about future trends. By leveraging historical data, businesses can gain valuable insights that inform decision-making, enhance customer experiences, and ultimately drive strategic growth. This innovative approach empowers organizations to stay ahead of the competition and adapt to changing market dynamics efficiently.

Part 3: Machine Learning – Predictive Modeling

Objective

Use machine learning to predict returns and forecast demand, helping optimize operations and inventory.



Data Extraction

Efficiently gathers data from various sources.



Data Transformation

Prepares data for analysis and reporting.



Insights Generation

Delivers actionable intelligence for decision-making.

Tasks & Models

1. Return Prediction Model

Goal: Predict whether a transaction will result in a return

Models: Logistic Regression, Random Forest

Evaluation: Precision, recall, F1-score, ROC-AUC

2. Monthly Demand Forecasting

Goal: Forecast monthly units sold by category and region

Models: SARIMA, ETS, or Gradient Boosting models

Evaluation: MAPE, RMSE against baseline models

3. (Optional) Market Basket Analysis

Goal: Identify frequently bought-together products to improve cross-selling

Technique: Apriori or FP-Growth algorithm

Deliverable: Top product association rules with lift and confidence metrics



Technical Artifacts & Reproducibility

To ensure complete transparency the full codebase, feature set, and model evaluation results are available on GitHub.

Content: Contains the complete Python code for both the return prediction model and the sales forecasting model.

Evaluation: Includes detailed model summaries, feature importance visualizations, and cross-validation outputs.

Version Control: The repository will host all future model iterations and performance tracking.

Complete Codebase: <https://github.com/AkshaySolanki-DA/Global-Fashion-Retail-Analytics.git>

Return Prediction Model

Key Insights ✅ Both models show near-perfect performance with 100% accuracy.🎯 No class imbalance issue—the model predicts both classes (0 and 1) equally well.🔍 ROC-AUC ~1.0 indicates excellent separability between returned and non-returned transactions.⚠️ Warning: Such high scores may indicate overfitting—models might be memorizing training data.

Moderate Forecast Accuracy: MAPE values range between 19% (best) and 40% (worst). Indicates decent performance, but inconsistency across months. Forecast errors still significant in certain periods (above 30%).

⚠️ Improvement Needed High MAPE (>30%) suggests seasonal spikes or trends are not fully captured. External factors like promotions, holidays, or product launches may need to be included.

✅ Next Steps Tune SARIMA parameters or test Prophet / Machine Learning models. Add exogenous variables (price, discount, events) for better accuracy.

Monthly Sales Forecasting – Key Insights

① **Model Performance – Poor Accuracy** MAPE: 136% | RMSE: 224K → Forecast is highly inaccurate.

ETS model is unsuitable – it over-smooths and fails to follow real sales behavior.

② **Sales Pattern – Highly Volatile** Extreme spikes (up to 800K+ in Dec 2024) vs lows near 130K. Strong month-to-month swings → suggests impact of promotions, seasonality, or events.

③ **Forecast Failure** Model predicts a flat ~380K, missing the real drop to ~200K in Jan 2025. Severe overestimation in test period → model doesn't capture true dynamics.

④ **Recommendations (Action Plan)** Drop ETS Model – Not suitable for volatile retail data. 🔎 Investigate Spikes – Analyze promotions, launches, and macro events. 🤖 Adopt Advanced Models – Use ARIMA, Prophet, or ML models with external factors.

Market Basket Analysis – Key Insights (Top 15 Rules) Strongest Product Associations

Most frequent combinations involve Prod_11, Prod_19, and Prod_8. Shoppers often buy Prod_11 together with Prod_19 & Prod_8 (and vice versa).  Cross-Selling Opportunities
Bundles to consider:
Prod_11 + Prod_19 + Prod_8
Prod_16 + Prod_13 + Prod_11
Prod_1 + Prod_2 → Prod_9
Prod_10 + Prod_3 → Prod_4

 Metrics Interpretation Support (~0.0025): These combos occur in ~0.25% of all transactions (rare, but significant in volume). Confidence (~14–15%): If a shopper buys the antecedent pair, there's a ~15% chance they add the consequent.

Lift (>1.0): All rules show positive association — buying A increases probability of buying

B.  Business Actions Create combo offers/bundles around top pairs/trios. Use these associations for product recommendations on website/checkout. Cross-promote frequently linked products in-store & online.



Thank You