

E-Commerce Sales Analysis

Project Overview

This project aims to analyse e-commerce sales data to derive actionable insights that support strategic decision-making. The primary goals are to:

1. Identify sales trends and patterns.
2. Segment customers for targeted marketing.
3. Evaluate the performance of various product categories.
4. Assess the impact of discounts on revenue generation.
5. Optimize inventory and marketing strategies.

Key Outcomes:

- Identification of top-performing products and categories.
- Insights into the effectiveness of discounts and customer preferences for payment methods.
- Data-driven strategies for increasing revenue and enhancing customer engagement.

The process is divided into key stages:

1. Data Importing: Uploading the raw dataset into SQL for analysis.
2. Data Cleaning: Ensuring the dataset is consistent, accurate, and ready for analysis by handling missing values, standardizing formats, and removing duplicates.
3. SQL-Based Analysis: Using queries to extract insights on product performance, sales trends, customer segmentation, and more.

Dataset Overview

The dataset contains transaction-level data with the following critical fields:

- **Product_ID:** A unique identifier for each product.
- **Category:** Product category (e.g., Clothing, Electronics).
- **Price (Rs):** The original price of the product.
- **Discount (%):** Discount applied to the product.
- **Final_Price (Rs):** The price after applying the discount.
- **Purchase_Date:** The date of the transaction.
- **Payment_Method:** Payment method used (e.g., Credit Card, UPI).

- **User_ID:** A unique identifier for each customer.

Dataset Statistics

- **Number of Rows:** 10,000 (example size)
- **Number of Columns:** 8

Challenges Identified in the Dataset

1. **Null Values:** Missing entries in essential fields such as Category and Final_Price.
2. **Inconsistent Formats:** Varying formats for columns like Purchase_Date and Discount (%).
3. **Outliers:** Unrealistic values such as negative prices or discounts exceeding 100%.
4. **Duplicate Records:** Transactions with identical User_ID and Product_ID combinations.

Data Importing:

The e-commerce dataset is present in CSV format and imported into a SQL database for analysis. The steps followed for the data import process are outlined below, ensuring the data was successfully transferred and structured for querying.

Steps for Data Importing

1. Prepare the CSV File
 - Verified that the dataset in CSV format contained all the necessary fields and was free of formatting issues, such as empty rows or extra columns.
 - Ensured that the column headers matched the schema of the target SQL table.
2. Create the SQL Table
 - Before importing the data, a table schema was created in SQL with appropriate data types for each column:
 - This ensured that the data from the CSV file could be mapped correctly to the SQL table.
3. Import Data Using SQL Server Management Studio (SSMS)
 - Opened SQL Server Management Studio (SSMS).
 - Navigated to the Database where the table was created (SalesAnalysis).
 - Right-clicked on the database and selected Tasks > Import Data to launch the Import Wizard.

4. Data Import Wizard Steps

- Data Source: Selected "Flat File Source" and provided the path to the CSV file.
- File Format: Specified the delimiter as a comma (,) since it was a CSV file.
- Preview: Verified the data preview to ensure all columns were detected correctly.
- Destination: Chose the target SQL database and the previously created table (ecommerce_dataset_updated).
- Mapping: Mapped the columns from the CSV file to the SQL table fields.
- Run the Import: Completed the process by executing the import task, ensuring all rows were successfully transferred.

5. Validation of Imported Data

- Queried the table to verify that the data was imported correctly

Data Cleaning

To ensure the dataset's reliability and accuracy, a comprehensive data cleaning process was carried out.

1. Handling Missing Values

- A SQL query was executed to identify rows with null values in critical columns like Product_ID, Category, Final_Price, etc.
- Rows with missing values in essential fields were excluded, as retaining them could lead to skewed analysis results.

2. Standardizing Data Types

Data type inconsistencies were resolved to ensure smooth query execution:

- **Discount (%)**: Converted to DECIMAL(5, 2) for precise representation of percentage values.
- **Price (Rs)** and **Final_Price (Rs)**: Converted to DECIMAL(10, 2) to handle monetary values accurately.
- **Purchase_Date**: Converted to the DATE format for consistent date-based queries.
- **Text-Based Columns**: Columns like User_ID, Product_ID, Category, and Payment_Method were standardized as VARCHAR for consistent string handling.

3. Identifying Outliers

Outliers were flagged and addressed to improve data quality:

- Unrealistic discounts (e.g., greater than 100% or less than 0%) were identified and removed.
- Negative prices were flagged as invalid and excluded from the dataset.

4. Text Standardization

- Leading and trailing spaces in text-based columns (User_ID, Category, etc.) were removed.
- All values in Category and Payment_Method columns were converted to uppercase to ensure uniformity.


5. Removing Duplicate Records

- Transactions with the same User_ID and Product_ID were identified using a query.
- Duplicate records were excluded to avoid inflating results in the analysis.

Data Analysis

1. Monthly Sales by Category

- **Objective:** Analyse sales trends across categories to identify seasonality and stable demand.

Results		 Messages											
	Category	Jan_Sales	Feb_Sales	Mar_Sales	Apr_Sales	May_Sales	June_Sales	July_Sales	Aug_Sales	Sept_Sales	Oct_Sales	Nov_Sales	Dec_Sales
1	BOOKS	9368.09	8863.00	7473.42	10025.06	9983.96	6884.37	11141.00	10684.64	10729.26	14675.79	8811.36	2509.40
2	SPORTS	9018.21	9655.73	12263.01	9935.30	11610.72	6253.65	9596.55	10839.88	9049.90	9389.70	6436.37	4469.77
3	HOME & KITCHEN	10478.32	9061.16	8257.55	10437.77	10194.24	6362.79	11689.54	12332.53	11535.99	8261.84	7666.27	4050.08
4	BEAUTY	7934.45	10989.92	12214.45	10309.36	9804.06	6894.70	9060.36	8097.54	9013.24	10713.98	6283.14	2899.90
5	TOYS	10961.10	7832.08	8854.47	9773.90	10924.74	9253.38	9642.48	11805.28	10615.58	8009.08	6158.91	3458.69
6	ELECTRONICS	7758.47	10758.94	7090.57	7126.97	10184.83	9491.25	11423.45	8685.08	7660.34	9070.69	7833.31	3378.33
7	CLOTHING	9565.11	12728.05	10311.26	12628.23	7236.53	7912.02	11705.04	7944.02	9365.32	12158.32	9644.81	4116.13

- **Insights:**
 - **Festive and Seasonal Peaks:** Categories like Toys, Clothing, and Books show strong seasonality, peaking during holidays or key events.
 - **Consistent Performers:** Home & Kitchen maintains steady demand throughout the year, providing reliable revenue.
 - **Opportunities for Growth:** Categories like Electronics and Toys show potential for better performance during festive seasons with optimized strategies.
- **Recommendations:**
 - Focus on festive campaigns for categories like Toys and Books to capitalize on seasonal spikes.
 - Maintain a stable inventory of Home & Kitchen items to meet consistent demand.
 - Plan promotional campaigns around major product launches for Electronics to drive sales.

2. Trend Analysis with Discounts and Revenue

- **Objective:** Evaluate how discounts impact revenue trends across months and seasons.

	Year	Month	Avg_Discount	Total_Revenue
1	2024	1	17.779503	65083.75
2	2024	2	18.130990	69888.88
3	2024	3	18.792569	66464.73
4	2024	4	20.301724	70236.59
5	2024	5	21.235465	69939.08
6	2024	6	17.784090	53052.16
7	2024	7	18.623595	74258.42
8	2024	8	18.535502	70388.97
9	2024	9	18.389057	67969.63
10	2024	10	17.821637	72279.40
11	2024	11	18.818897	52834.17
12	2024	12	20.039370	24882.30

- **Insights:**
 - Seasonal Discounts: High discounts during April and December align with revenue spikes, indicating their effectiveness in driving sales.
 - Moderate Discounts: Months like July achieve significant revenue despite offering moderate discounts, showcasing the potential for profitability.
- **Recommendations:**
 - Use A/B testing to determine the optimal discount rates that maximize revenue without eroding profit margins.
 - Schedule larger discounts during festive periods to stimulate demand and maintain momentum.
 - Evaluate customer responsiveness to moderate discounts in off-peak months for consistent revenue flow.

3. Top Products by Revenue

- **Objective:** Identify the top five products contributing the most to overall revenue.

	Product_ID	Total_Revenue
1	5a4c2797-7	496.82
2	2a8fff8b-4	495.02
3	475d4cf1-c	493.04
4	e88dc97f-5	492.41
5	67d7c91c-b	491.70

- **Insights:**
 - A few products significantly dominate revenue generation, indicating strong customer preferences or product appeal.

- High-performing products, such as those with IDs 5a4c2797-7 and 2a8fff8b-4, consistently outpace others in sales.
- **Recommendations:**
 - Prioritize inventory management and marketing support for top-performing products to avoid stockouts.
 - Analyse customer feedback and reviews to understand why these products outperform and apply the findings to improve other offerings.

4. Category with the Highest Revenue

- **Objective:** Determine which product category contributes the most to total revenue.

	Category	Total_Revenue
1	CLOTHING	115314.84

- **Insights:**
 - Clothing: Generates the highest revenue, even with lower discounts compared to other categories, indicating strong demand.
- **Recommendations:**
 - Expand the product range within the Clothing category to meet diverse customer preferences.
 - Allocate marketing resources and design campaigns specifically for this high-performing category.
 - Conduct customer surveys to further understand preferences and optimize offerings in Clothing.

5. Effectiveness of Discounts on Revenue

- **Objective:** Assess the relationship between discounts and revenue generation across categories.

	Category	Average_Discount	Total_Revenue
1	HOME & KITCHEN	19.608378	110328.08
2	SPORTS	19.326923	108518.79
3	ELECTRONICS	19.267068	100462.23
4	BOOKS	19.035580	111149.35
5	BEAUTY	18.475247	104215.10
6	TOYS	18.126195	107289.69
7	CLOTHING	17.919020	115314.84

- **Insights:**

- Home & Kitchen: Achieves significant revenue increases with higher discounts (19.6% average).
- Clothing: Performs well with minimal discounts (17.9% average), reflecting strong demand elasticity.
- **Recommendations:**
 - Offer higher discounts for price-sensitive categories like Home & Kitchen to maximize sales volume.
 - Maintain moderate discounts for categories with inherent demand, such as Clothing, to protect profit margins.
 - Experiment with dynamic pricing strategies for mid-performing categories like Electronics to find the optimal balance between discounts and revenue.

6. Preferred Payment Methods

- **Objective:** Identify customer preferences for payment methods and their contribution to revenue.

Results		Messages		
	Payment_Method	Number_of_Transactions	Total_Revenue	Average_Transaction_Value
1	CREDIT CARD	760	156018.02	205.286868
2	UPI	757	154684.81	204.339247
3	DEBIT CARD	731	153362.26	209.797893
4	NET BANKING	716	146804.40	205.034078
5	CASH ON DELIVERY	696	146408.59	210.357169

- **Insights:**
 - Credit Card and UPI: Lead in transaction frequency, indicating widespread adoption of digital payments.
 - Cash on Delivery (COD): While less frequent, COD has the highest average transaction value, highlighting its relevance for high-value purchases.
- **Recommendations:**
 - Promote digital payments through incentives like cashback or reward points to encourage usage.
 - Maintain and enhance logistics for COD to ensure customer satisfaction, particularly for regions or demographics preferring this method.
 - Educate customers on the benefits of digital payments to further increase their adoption.

7. Sales Spikes and Drops

- **Objective:** Identify specific time periods with significant sales increases or decreases.

	Purchase_Date	Daily_Sales
1	2024-01-01	950.91
2	2024-01-02	2536.21
3	2024-01-03	2336.35
4	2024-01-04	2035.24
5	2024-01-05	1800.38
6	2024-01-06	3465.40
7	2024-01-07	2692.41
8	2024-01-08	755.46
9	2024-01-09	1607.63
10	2024-01-10	2278.97
11	2024-01-11	2017.88
12	2024-01-13	1518.67
13	2024-01-14	3274.35
14	2024-01-15	1739.44
15	2024-01-16	2364.28
16	2024-01-17	1931.54
17	2024-01-18	2556.13
18	2024-01-19	1457.84
19	2024-01-20	1783.44

- **Insights:**

- Sales spikes are observed during holidays, promotions, and festive periods, such as mid-July (₹4,697.54 on 2024-07-18).
- Post-holiday periods exhibit a slump (e.g., ₹950.91 on 2024-01-01), reflecting reduced customer activity.

- **Recommendations:**

- Schedule marketing campaigns and flash sales around anticipated high-demand periods like holidays and weekends.
- Implement targeted discounts or free shipping offers during low-demand periods to boost sales.
- Analyse customer feedback during peak and low periods to fine-tune marketing strategies.

8. Discounts Driving Sales Across Categories

- **Objective:** Evaluate the impact of discounts on sales across different product categories.

	Category	Total_Revenue	Average_Discount	Discount_Impact
1	BOOKS	111149.35	19.035580	18766.98650000
2	SPORTS	108518.79	19.326923	17732.90050000
3	HOME & KITCHEN	110328.08	19.608378	18347.74250000
4	BEAUTY	104215.10	18.475247	16485.21500000
5	TOYS	107289.69	18.126195	16087.60900000
6	ELECTRONICS	100462.23	19.267068	16959.65650000
7	CLOTHING	115314.84	17.919020	17577.68050000

- **Insights:**

- Price-sensitive categories like Home & Kitchen benefit most from discounts.
- High-performing categories like Clothing generate strong revenue even with minimal discounts.

- **Recommendations:**
 - Focus discounts on price-sensitive categories to maximize sales volumes.
 - Use minimal discounting for categories with strong baseline demand, such as Clothing, to protect profit margins.

9. Category Performance Over Time

- **Objective:** Track how each category performs across months to identify revenue trends.

	Year	Month	Category	Monthly_Sales		Year	Month	Category	Monthly_Sales
1	2024	1	SPORTS	9018.21	43	2024	7	SPORTS	9596.55
2	2024	1	ELECTRONICS	7758.47	44	2024	7	ELECTRONICS	11423.45
3	2024	1	CLOTHING	9565.11	45	2024	7	HOME & KITCHEN	11689.54
4	2024	1	HOME & KITCHEN	10478.32	46	2024	7	TOYS	9642.48
5	2024	1	TOYS	10961.10	47	2024	7	BEAUTY	9060.36
6	2024	1	BEAUTY	7934.45	48	2024	7	BOOKS	11141.00
7	2024	1	BOOKS	9368.09	49	2024	7	CLOTHING	11705.04
8	2024	2	ELECTRONICS	10758.94	50	2024	8	HOME & KITCHEN	12332.53
9	2024	2	CLOTHING	12728.05	51	2024	8	BEAUTY	8097.54
10	2024	2	TOYS	7832.08	52	2024	8	CLOTHING	7944.02
11	2024	2	BEAUTY	10989.92	53	2024	8	SPORTS	10839.88
12	2024	2	HOME & KITCHEN	9061.16	54	2024	8	BOOKS	10684.64
13	2024	2	BOOKS	8863.00	55	2024	8	ELECTRONICS	8685.08
14	2024	2	SPORTS	9655.73	56	2024	8	TOYS	11805.28
15	2024	3	TOYS	8854.47	57	2024	9	BOOKS	10729.26
16	2024	3	CLOTHING	10311.26	58	2024	9	HOME & KITCHEN	11535.99
17	2024	3	BEAUTY	12214.45	59	2024	9	SPORTS	9049.90
18	2024	3	ELECTRONICS	7090.57	60	2024	9	ELECTRONICS	7660.34
19	2024	3	HOME & KITCHEN	8257.55	61	2024	9	CLOTHING	9365.32
20	2024	3	BOOKS	7473.42	62	2024	9	TOYS	10615.58
21	2024	3	SPORTS	12263.01	63	2024	9	BEAUTY	9013.24
22	2024	4	BEAUTY	10309.36	64	2024	10	BOOKS	14675.79
23	2024	4	TOYS	9773.90	65	2024	10	ELECTRONICS	9070.69
24	2024	4	HOME & KITCHEN	10437.77	66	2024	10	SPORTS	9389.70
25	2024	4	BOOKS	10025.06	67	2024	10	CLOTHING	12158.32
26	2024	4	SPORTS	9935.30	68	2024	10	BEAUTY	10713.98
27	2024	4	ELECTRONICS	7126.97	69	2024	10	TOYS	8009.08
28	2024	4	CLOTHING	12628.23	70	2024	10	HOME & KITCHEN	8261.84
29	2024	5	SPORTS	11610.72	71	2024	11	BOOKS	8811.36
30	2024	5	HOME & KITCHEN	10194.24	72	2024	11	CLOTHING	9644.81
31	2024	5	BOOKS	9983.96	73	2024	11	ELECTRONICS	7833.31
32	2024	5	BEAUTY	9804.06	74	2024	11	SPORTS	6436.37
33	2024	5	TOYS	10924.74	75	2024	11	BEAUTY	6283.14
34	2024	5	ELECTRONICS	10184.83	76	2024	11	HOME & KITCHEN	7666.27
35	2024	5	CLOTHING	7236.53	77	2024	11	TOYS	6158.91
36	2024	6	SPORTS	6253.65	78	2024	12	SPORTS	4469.77
37	2024	6	HOME & KITCHEN	6362.79	79	2024	12	ELECTRONICS	3378.33
38	2024	6	CLOTHING	7912.02	80	2024	12	BEAUTY	2899.90
39	2024	6	ELECTRONICS	9491.25	81	2024	12	CLOTHING	4116.13
40	2024	6	BOOKS	6884.37	82	2024	12	BOOKS	2509.40
41	2024	6	BEAUTY	6894.70	83	2024	12	TOYS	3458.69
42	2024	6	TOYS	9253.38	84	2024	12	HOME & KITCHEN	4050.08

- **Insights:**
 - Toys see a clear peak during holidays, aligning with festive shopping trends.

- Electronics show a decline in December, potentially due to end-of-year promotions or customer spending shifts.
- Home & Kitchen exhibits steady performance throughout the year.
- **Recommendations:**
 - Schedule seasonal promotions for categories like Toys to maximize holiday revenue.
 - Address inventory or marketing gaps for categories like Electronics during low-performing months.
 - Focus on maintaining inventory for stable categories like Home & Kitchen to ensure consistent revenue.

10. Moving Averages of Product Sales

- **Objective: Identify trends in product sales using 7-day moving averages.**

	Product_ID	Purchase_Date	Daily_Sales	Moving_Avg_7_Days
1	003d1f09-c	2024-08-30	172.15	172.150000
2	004ae9eb-0	2024-09-14	61.27	61.270000
3	0063786d-0	2024-12-10	354.92	354.920000
4	006cd2ab-5	2024-01-05	414.27	414.270000
5	00b3c8d1-b	2024-02-18	432.76	432.760000
6	00bb902e-c	2024-09-22	147.52	147.520000
7	00cabebb-5	2024-07-07	229.44	229.440000
8	00d20e0d-6	2024-03-20	337.58	337.580000
9	00fcb9c0-8	2024-06-24	315.70	315.700000
10	012b2dda-1	2024-05-30	315.45	315.450000
11	015db72c-0	2024-10-05	89.32	89.320000
12	0163c4eb-b	2024-09-19	125.20	125.200000
13	0169a3ab-a	2024-06-06	430.96	430.960000
14	018f23ea-7	2024-07-23	217.87	217.870000
15	01b89dba-d	2024-05-26	84.24	84.240000
16	01c1488d-2	2024-07-24	430.38	430.380000
17	01d22c80-5	2024-08-07	281.72	281.720000
18	01f61e5d-d	2024-10-25	240.64	240.640000
19	021b5323-6	2024-01-17	30.63	30.630000
20	023a6d02-4	2024-11-16	159.75	159.750000
21	024bf585-e	2024-03-23	368.70	368.700000
22	025cebdc-b	2024-07-22	122.45	122.450000
23	027fc4d8-d	2024-01-17	146.75	146.750000

- **Insights:**
 - Periodic spikes indicate seasonal or promotional influences on specific products.
 - Steady declines in moving averages may signal reduced customer interest or increased competition.
- **Recommendations:**
 - Increase marketing support for products showing consistent upward trends.
 - Investigate and address the causes of declining sales for underperforming products.
 - Use moving averages to forecast demand and plan inventory replenishment.

11. Customer Segmentation Based on Spending

- **Objective:** Segment customers into spending categories for targeted engagement.

	User_ID	Total_Spent	Spending_Segment
103	06c442e0	150.05	Low Spender
104	06d390ac	349.59	Low Spender
105	0741d501	71.08	Low Spender
106	074c931e	48.48	Low Spender
107	0750df4a	182.87	Low Spender
108	075fc126	93.94	Low Spender
109	076f21ca	42.85	Low Spender
110	077651c7	216.23	Low Spender
111	078124d1	314.67	Low Spender
112	078fcf7a	232.22	Low Spender
113	07bba05f	327.74	Low Spender
114	07d9828a	288.05	Low Spender
115	07da90ff	102.55	Low Spender
116	07e93a67	136.10	Low Spender
117	07fd6558	76.12	Low Spender
118	080f21f4	83.30	Low Spender
119	084b0e23	71.39	Low Spender
120	084d1679	155.33	Low Spender
121	08602a9c	84.00	Low Spender
122	0888b113	52.27	Low Spender
123	088e7c77	266.10	Low Spender
124	08a65189	315.45	Low Spender
125	08a6752d	24.64	Low Spender

- **Insights:**
 - High spenders (₹1,000+) constitute 20% of customers but contribute to 50% of total revenue.
 - Low spenders account for a significant portion of the customer base but have minimal revenue impact.
- **Recommendations:**
 - Develop loyalty programs and exclusive offers to retain high spenders and encourage repeat purchases.
 - Use targeted promotions, such as small discounts or free shipping, to convert low spenders into moderate or high spenders.

12. Top-Selling Products by Category

- **Objective:** Identify best-performing products within each category for targeted strategies.

	Category	Product_ID	Total_Revenue
1	BEAUTY	28714133-6	486.79
2	BEAUTY	ee59d876-b	479.19
3	BEAUTY	2bed8c90-3	473.42
4	BEAUTY	74ebdb33-e	471.43
5	BEAUTY	e907f4d3-9	471.25
6	BEAUTY	cf9bf8ad-9	470.33
7	BEAUTY	1a41a273-7	468.43
8	BEAUTY	5e3a3479-c	467.38
9	BEAUTY	9f0098a7-5	458.94
10	BEAUTY	e80edd02-2	456.78
11	BEAUTY	669c2558-4	454.48
12	BEAUTY	273017e8-2	453.66
13	BEAUTY	492eb03f-1	453.41
14	BEAUTY	f0ed22c9-3	452.74
15	BEAUTY	6556f3ed-5	450.56
16	BEAUTY	fa938c93-7	450.24
17	BEAUTY	3825f76f-9	449.80
18	BEAUTY	9ffd92ed-4	449.69
19	BEAUTY	df747575-7	448.96
20	BEAUTY	c30ec252-d	444.79
21	BEAUTY	aacdd2c4-c	440.91
22	BEAUTY	5e363698-2	436.53
23	BEAUTY	26fe9d8a-2	435.68
24	BEAUTY	97332847-9	433.41

- **Insights:**
 - Certain products consistently dominate sales within their categories, indicating strong customer preferences.
- **Recommendations:**
 - Highlight these top-performing products in marketing campaigns and promotional activities.
 - Ensure sufficient stock levels for these products to prevent lost sales opportunities.

Key Findings

- **Seasonality and Demand:** Categories like Toys and Books thrive during festive seasons, while Home & Kitchen offers consistent revenue opportunities year-round.
- **Top Categories:** Clothing generates the highest revenue; Books and Toys perform best during festive seasons.
- **Discount Optimization:** While discounts effectively boost sales, a tailored strategy is crucial for balancing profitability and demand.
- **Customer Engagement:** High spenders should be prioritized for loyalty programs, while targeted campaigns for low spenders can enhance overall revenue.
- **Payment Method Trends:** Digital payments dominate customer preferences, but Cash on Delivery retains relevance for high-value transactions.
- **Actionable Insights:** Ensuring adequate inventory, leveraging top-performing products, and aligning promotions with sales trends can significantly enhance business performance.