

Summary

Sales Growth

Product Performance

Product Details

Marketing Campaigns

Forecasting Trend

Customer Insights

Churn Prediction



Family Supermarket is a mid-sized retail enterprise that operates across multiple product categories, serving customers through both online and offline channels. Despite its established presence, the leadership team has expressed concern over declining repeat purchases and uneven sales performance. They suspect that these challenges are closely tied to data quality issues and a lack of comprehensive visibility into customer behavior, which limits their ability to design effective loyalty programs and optimize marketing strategies.

To investigate these concerns, Family Supermarket has provided five interconnected datasets derived from its operational systems: customers, products, sales, marketing, and reviews. The customer dataset captures demographic and transactional details, enabling segmentation and analysis of retention patterns. The product dataset outlines attributes such as category, pricing, and availability, which are essential for assortment planning and demand forecasting. Sales data reflects purchase trends across both channels, offering insights into seasonality, regional performance, and repeat buying behavior. Marketing records highlight campaign reach, expenditure, and conversion rates, allowing evaluation of promotional effectiveness. Finally, the review dataset provides sentiment and feedback, which can be leveraged to enhance product quality and customer experience.

By integrating and analyzing these datasets, Family Supermarket can uncover hidden patterns, identify inconsistencies, and build a unified view of customer journeys. This holistic approach will empower leadership to make data-driven decisions, strengthen customer relationships, and stabilize sales performance in an increasingly competitive retail environment.



Summary

shopping

Date

All

Channel

All

Segment

All

Location

All

- Total Orders
- Total Revenue
- Gross Profit

Customer Count

96K

Orders Per Customer

1.07

Churn Rate

80

Avg Review Score

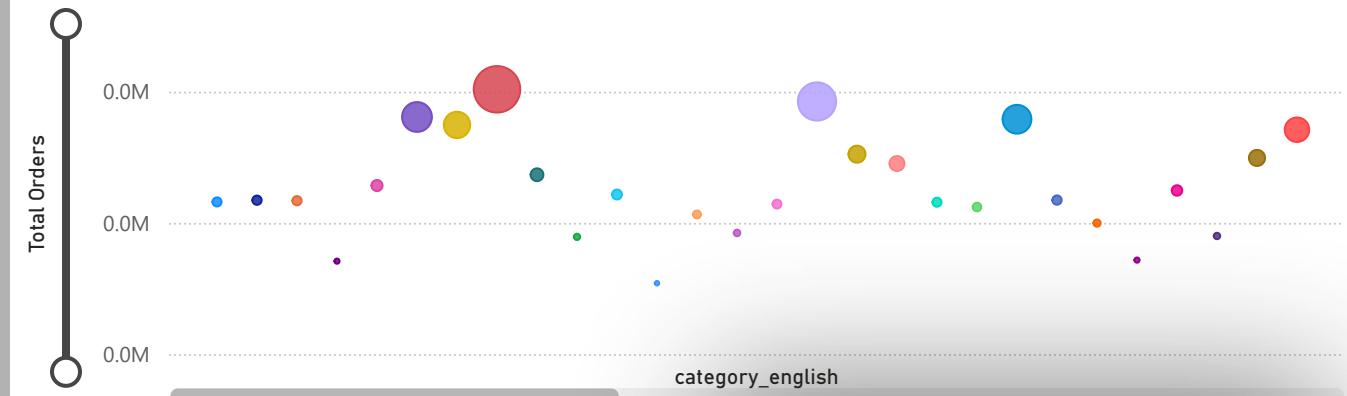
3.67

Avg Order Value

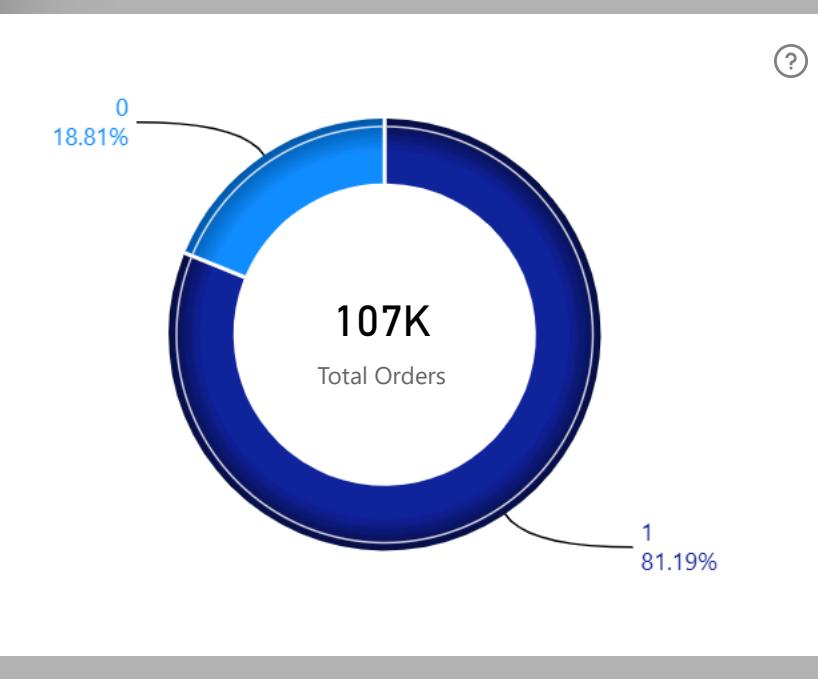
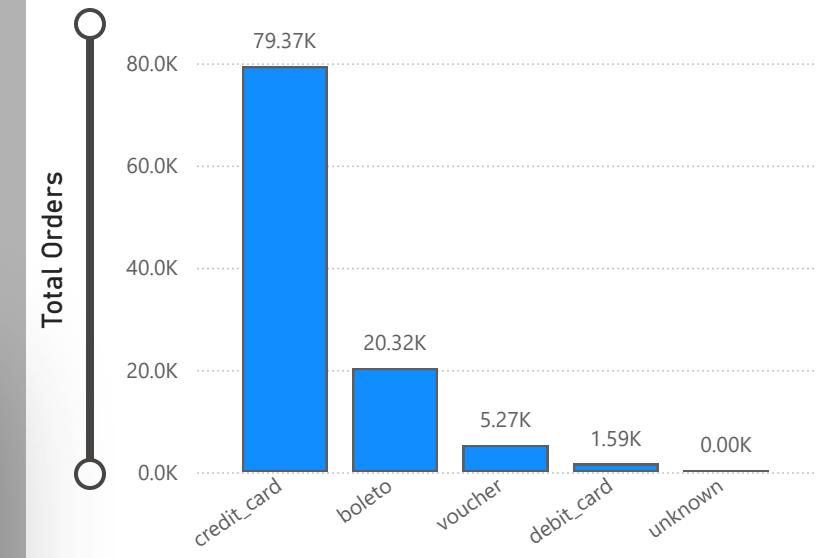
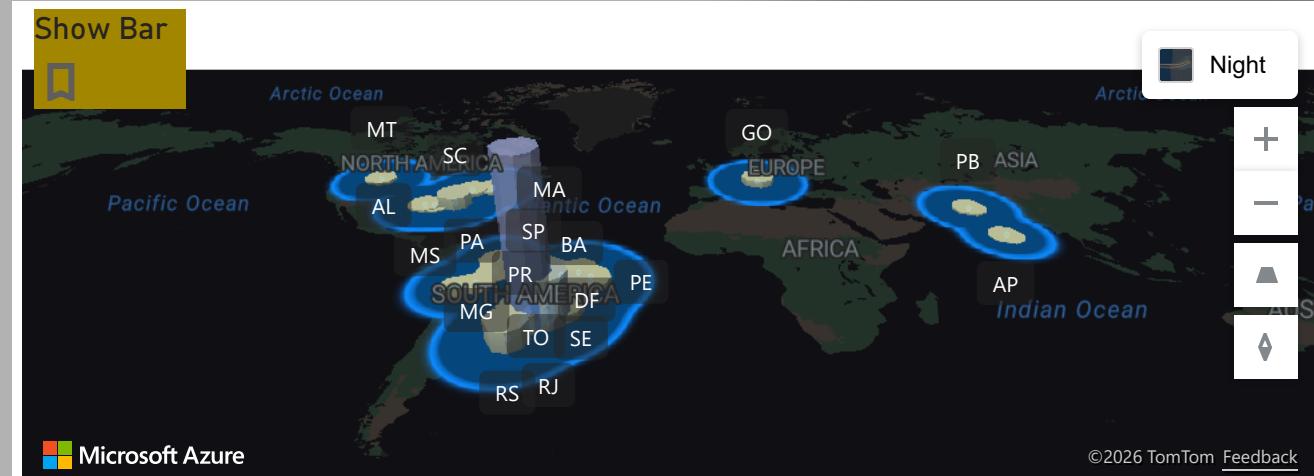
122.21

Top Product Revenue

962.9K



year	category	Top Product Revenue	Total Orders
2016	fashion_male_clothing	35.86	1
2016	fashion_shoes	40.95	1
2017	diapers_and_hygiene	49.85	1
2016	fashio_female_clothing	69.63	1
2016	unknown	95.92	2
Total		9,62,884.77	106549





Sales Growth

shopping

Date

All

Channel

All

Segment

All

Location

All

Avg Review Score

3.67

Forecasted Revenue

3.33M

Product Revenue

13.02M

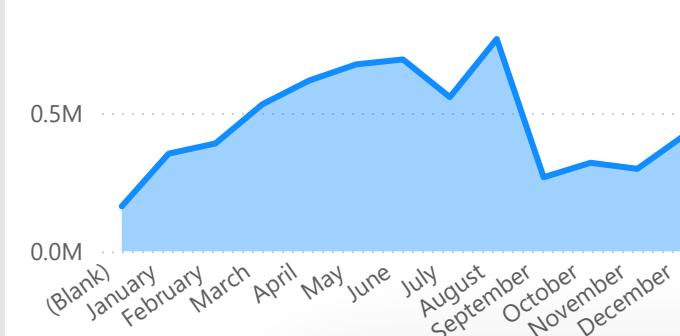
Top Product Revenue

963K

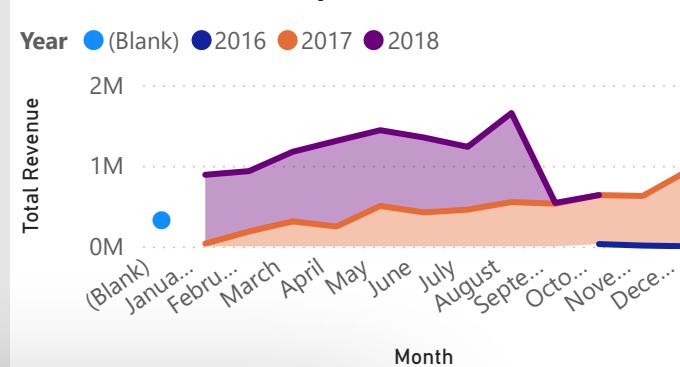
Avg Order Value

122.21

Gross Profit



Total Revenue over years



Date

Date	Actual Revenue	Forecasted Revenue
01 September 2016	259.11	0.00
01 October 2016	44,324.06	0.00
01 December 2016	19.62	0.00
01 January 2017	1,08,209.06	0.00
01 February 2017	2,26,652.36	0.00
01 March 2017	3,46,090.36	0.00
01 April 2017	3,32,747.52	0.00
01 May 2017	4,86,222.63	0.00
01 June 2017	4,16,141.19	0.00
01 July 2017	5,19,239.57	0.00
01 August 2017	5,51,995.29	0.00
01 September 2017	5,72,301.62	0.00
01 October 2017	6,14,368.47	0.00
01 November 2017	9,64,676.84	0.00
01 December 2017	7,39,710.78	0.00
01 January 2018	9,77,300.06	0.00
Total	1,30,21,390.90	33,26,676.61

High Risk

1,30,21,390.90

Total Revenue

33,26,676.61

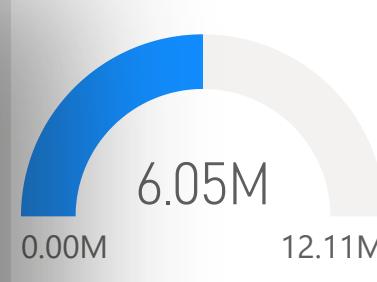
Forecasted Revenue

9,62,884.77

Top Product Revenue



Gross Profit



Total Orders over time

year ● 2016 ● 2017 ● 2018 ● Total Revenue





Product Performance

shopping

Date

All

Channel

All

Segment

All

Location

All

Avg Order Value

122.21

Avg Response Rate

0.33

Avg Review Score

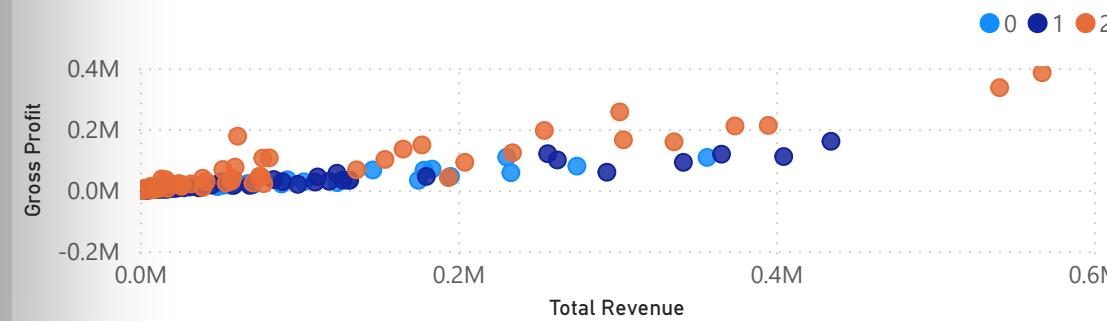
3.67

Product Revenue

13.02M

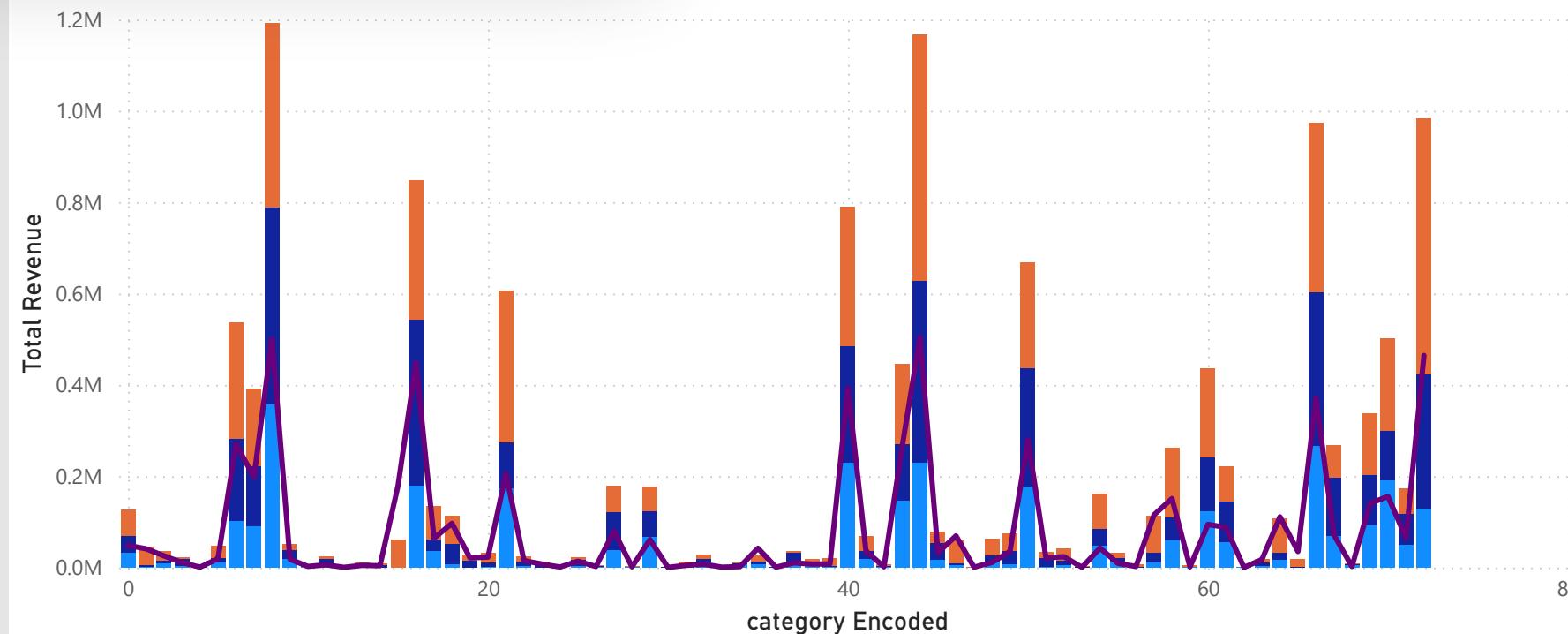


Category Revenue vs Profit by Segment

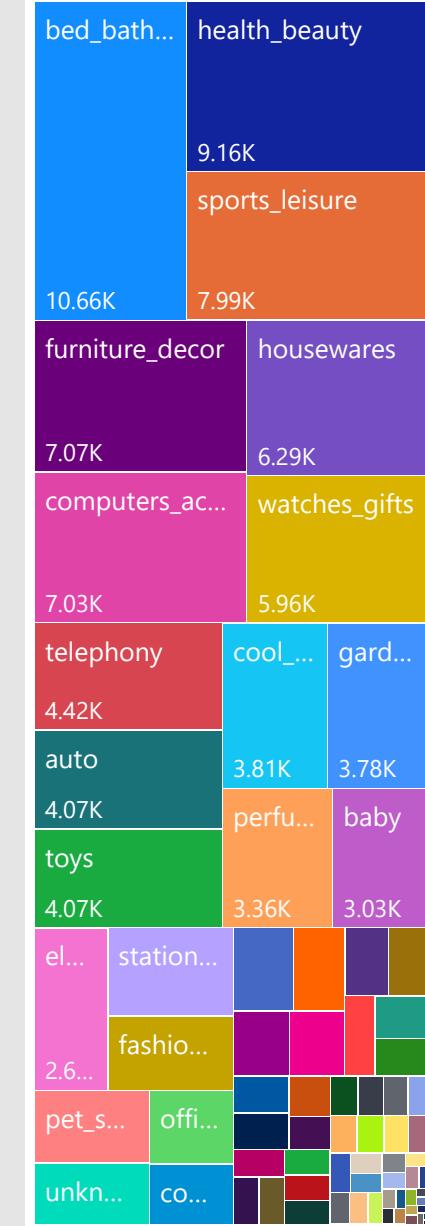


Revenue by Product

segment ● 0 ● 1 ● 2 ● Gross Profit



Orders by category





Product Details

shopping

Avg Order Value

122.21

Avg Response Rate

0.33

Avg Review Score

367

Product Revenue

13.02M

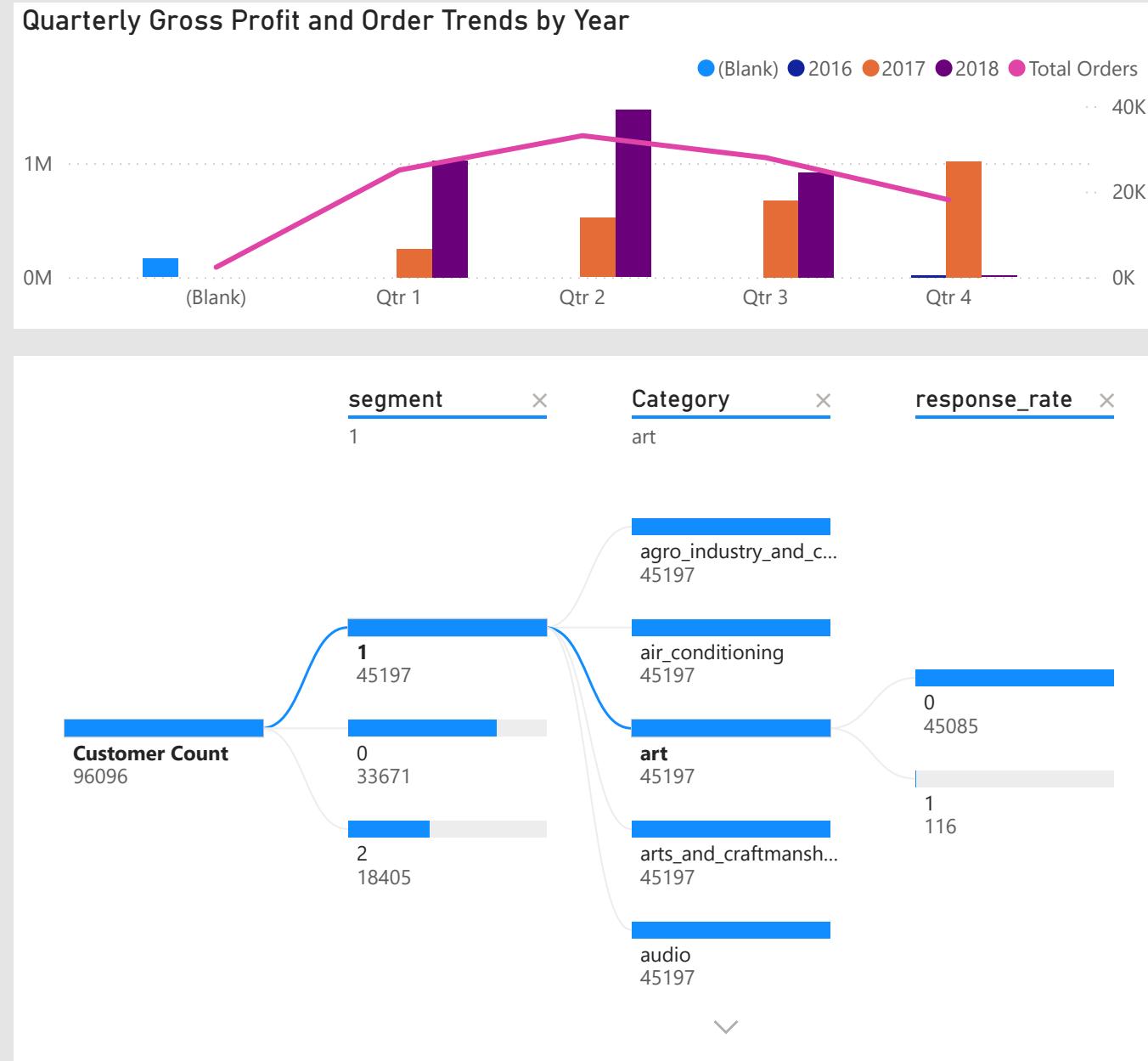
category Encoded	Total Orders	Gross Profit	Total Revenue
⊕ 8	10831	5,00,603.09	11,92,896.64
⊕ 44	9175	5,04,115.17	11,66,177.46
⊕ 66	7944	3,71,594.35	9,73,092.98
⊕ 40	7075	3,92,680.35	7,88,803.11
⊕ 16	7072	4,47,873.68	8,46,889.25
⊕ 50	6261	2,79,371.70	6,66,887.22
⊕ 72	5922	4,64,273.34	9,82,833.65
⊕ 69	4398	1,39,836.83	3,38,271.69
⊕ 6	4101	2,70,173.81	5,35,762.89
⊕ 70	4027	1,55,776.39	5,01,290.29
⊕ 21	3788	2,05,624.27	6,05,861.97
⊕ 43	3770	2,62,320.51	4,45,148.31
⊕ 60	3353	94,319.58	4,35,458.79
⊕ 7	3106	1,97,889.01	3,91,872.70
⊕ 27	2623	79,574.60	1,78,686.62
⊕ 67	2419	71,416.73	2,67,128.82
⊕ 29	2049	61,044.83	1,77,938.99
⊕ 61	1832	87,662.41	2,21,124.20
⊕ 71	1479	64,572.77	1,72,574.40
⊕ 58	1349	1,50,964.77	2,62,874.06
⊕ 17	1112	64,100.76	1,35,000.55
⊕ 54	1090	42,492.92	1,61,793.02
⊕ 0	975	48,271.32	1,27,048.65
⊕ 18	801	96,633.31	1,12,294.49
⊕ 45	779	32,314.48	77,981.37
⊕ 57	650	1,15,097.11	1,12,462.55
⊕ 64	645	1,10,909.88	1,08,236.61
⊕ 9	540	18,103.08	50,102.16
Total	106549	60,53,681.40	1,30,21,390.90

Date

Channel

Segment

Location





Marketing Campaigns

shopping

Date

All

Channel

All

Segment

All

Location

All

Marketing Spend

2M

Segment Revenue

13.02M

Sum of conversions

164

Total Response Rate

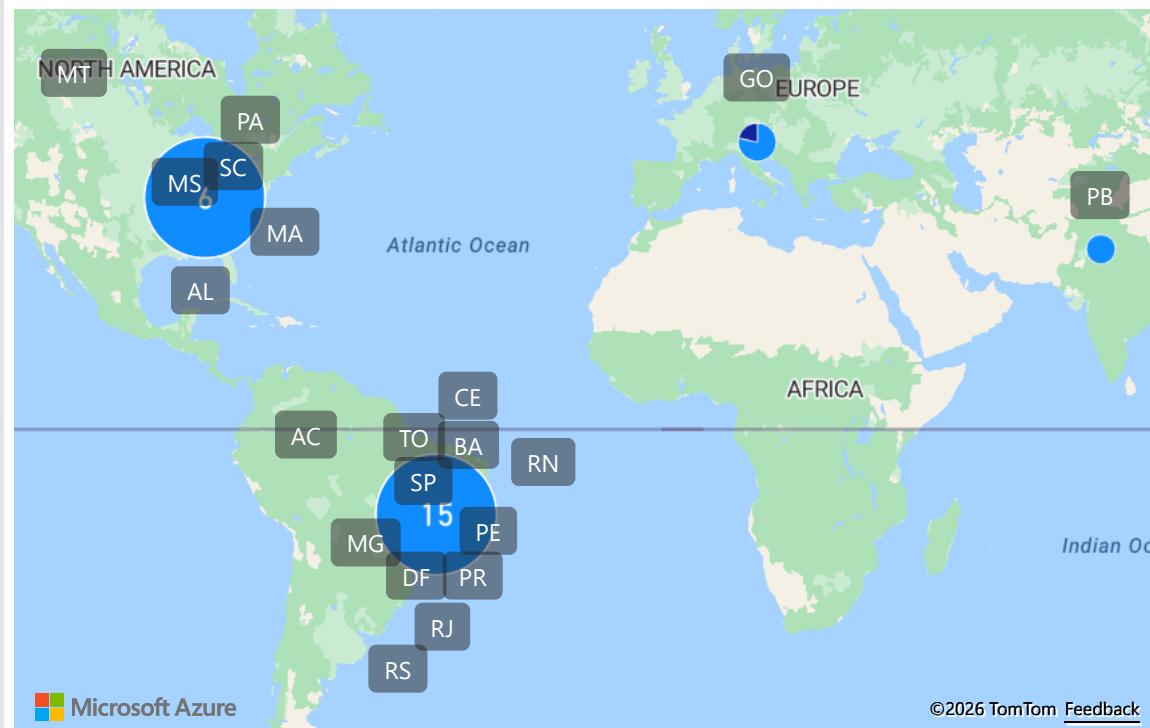
166.17

Avg Response Rate

0.33

Spend and Response_rate by state and conversions

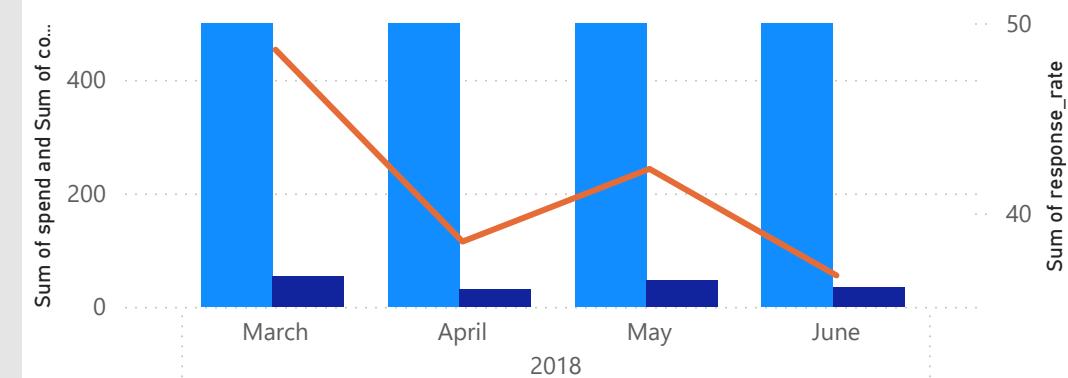
conversions ● 0 ● 1



start_date	campaign_id	spend_band	Marketing Spend	Avg Response Rate	city
21 May 2018	C1000	6-8K	7583	0.00	Sao Paulo
15 March 2018	C1001	4-6K	5888	0.50	Unai
04 March 2018	C1002	6-8K	7543	0.33	Icara
03 June 2018	C1003	0-2K	1755	0.25	Duque De Caxias
05 April 2018	C1004	0-2K	1484	0.40	Mutum
01 April 2018	C1005	2-4K	2217	0.33	Campo Grande
Total			2237467	0.33	

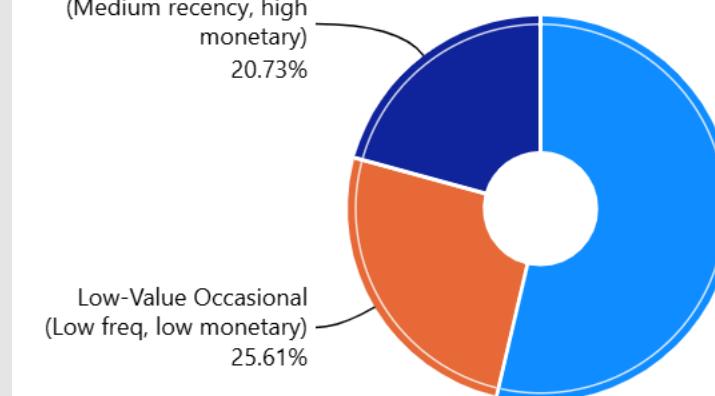
Marketing Performance

● Sum of spend ● Sum of conversions ● Sum of response_rate



Conversions by channel

High-Value Loyalists
(Medium recency, high monetary)
20.73%



At-Risk Customers (High recency, low engagement)
53.66%

Low-Value Occasional
(Low freq, low monetary)
25.61%



Demand Forecasting

shopping

Date

All

Channel

All

Segment

All

Location

All

▼

Seasonal Growth

Forecasted

Forecasted Orders

Forecasted Revenue

Total Revenue

13.02M

Avg Order Value

122.21

Total Orders

107K

Churn Rate

80.41

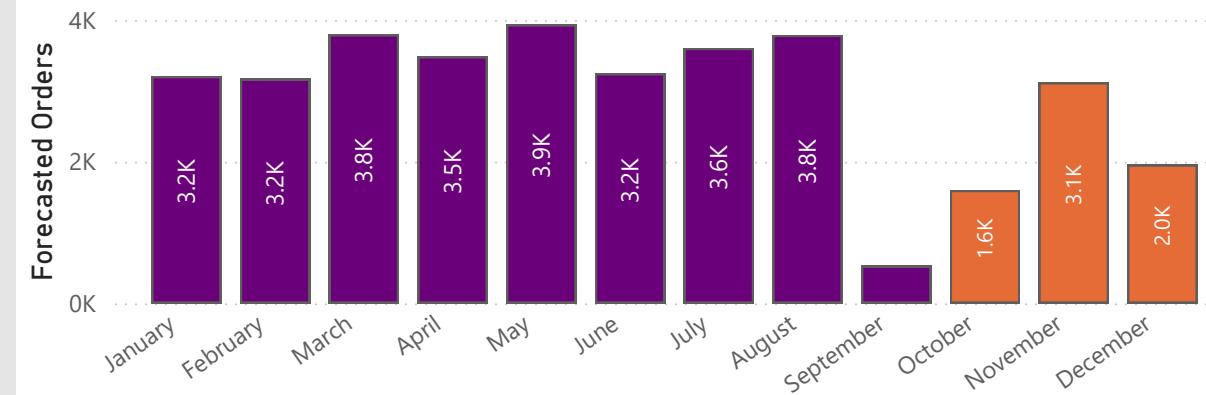
What is happening? Forecasted demand shows a significant projected spike in **November (3.1K orders)** and **December (2.2K orders)**, following a period of relatively stable order volume through the summer.

Why is it happening? Strong historical seasonal trends drive Q4 performance. Geographically, demand is concentrated in specific clusters across **North America** and **Brazil**, which are the primary engines for this projected growth.

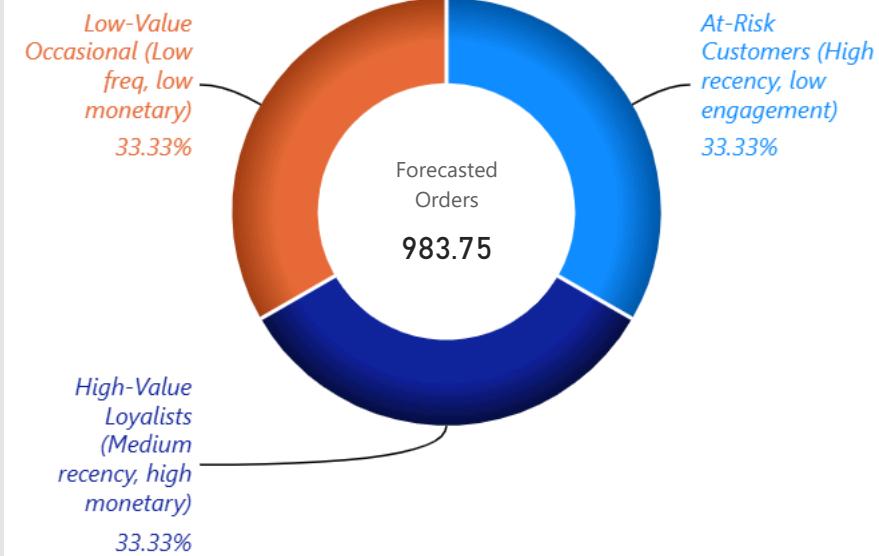
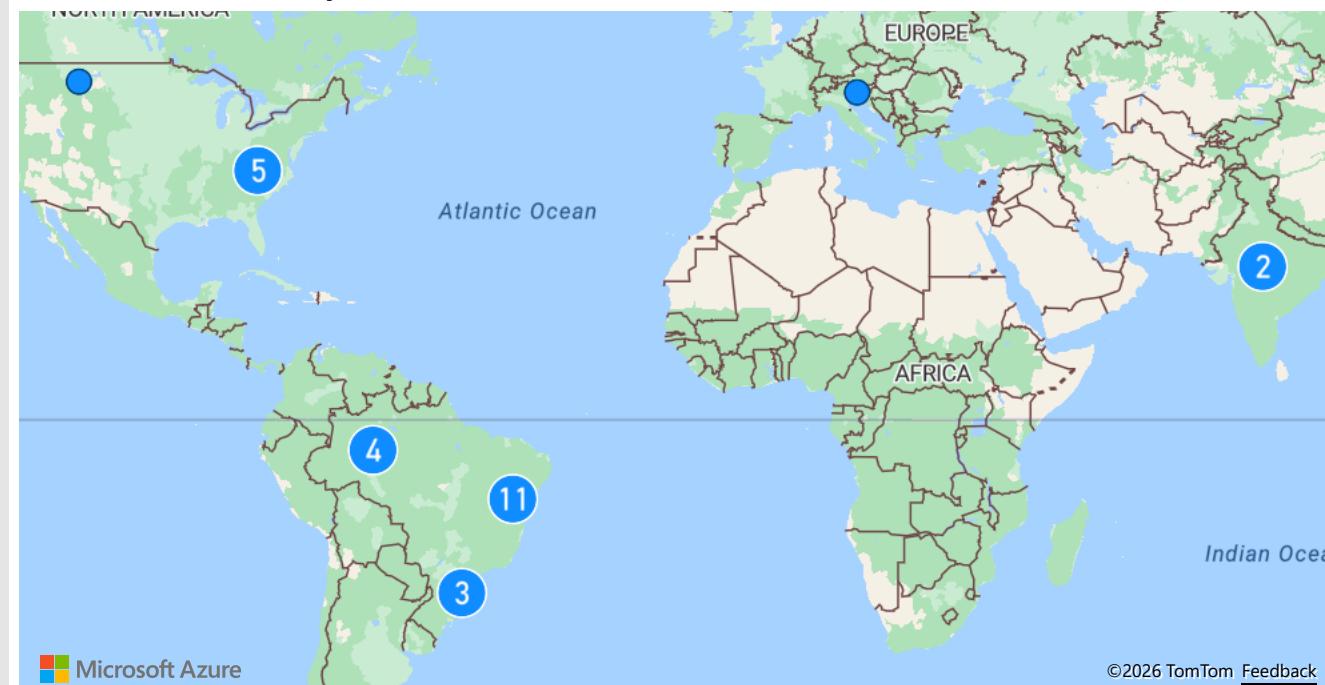
What should we do next?

- Inventory Prep:** Scale up stock levels specifically for the North American and Brazilian regions by October to meet the 3K+ November order forecast.
- Resource Allocation:** Shift marketing spend toward the **High-Value Loyalist** and **At-Risk** segments (who each account for 33.33% of forecasted orders) to maximize the year-end surge.

Year ● 2016 ● 2017 ● 2018 ● 2019



Forecasted Orders by state





Customer Insights

shopping

Date

All

Channel

All

Segment

All

Location

All

Total Revenue

13.02M

Avg Order Value

122.21

Total Orders

107K

Customer Count

96K

Forecasted Orders

983.75

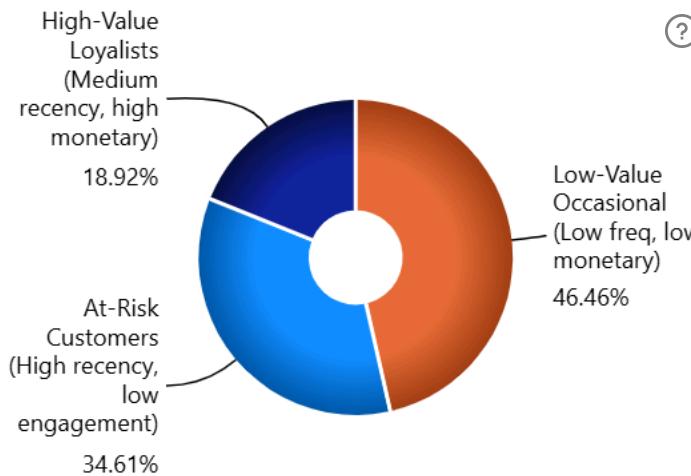
Churn Rate

80.41

Avg Review Score

3.67

Customer Count by profile



Segment Imbalance

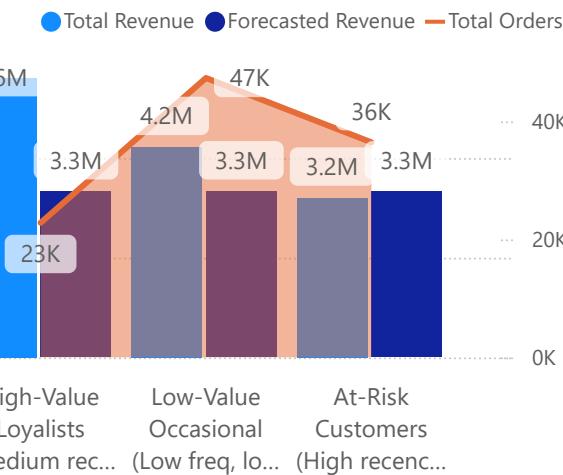
• **What is happening?** Our base is heavily skewed toward **Low-Value Occasional** shoppers (46.46%), while the **High-Value Loyalists** who drive the bulk of our revenue (\$5.6M) make up only 18.92% of the total population.

• **Why is it happening?** A significant portion of our base (34.61%) has fallen into the **At-Risk** category due to high recency. The Retention Snapshot shows a sharp drop in customer activity after the November peak, indicating a failure to retain holiday shoppers.

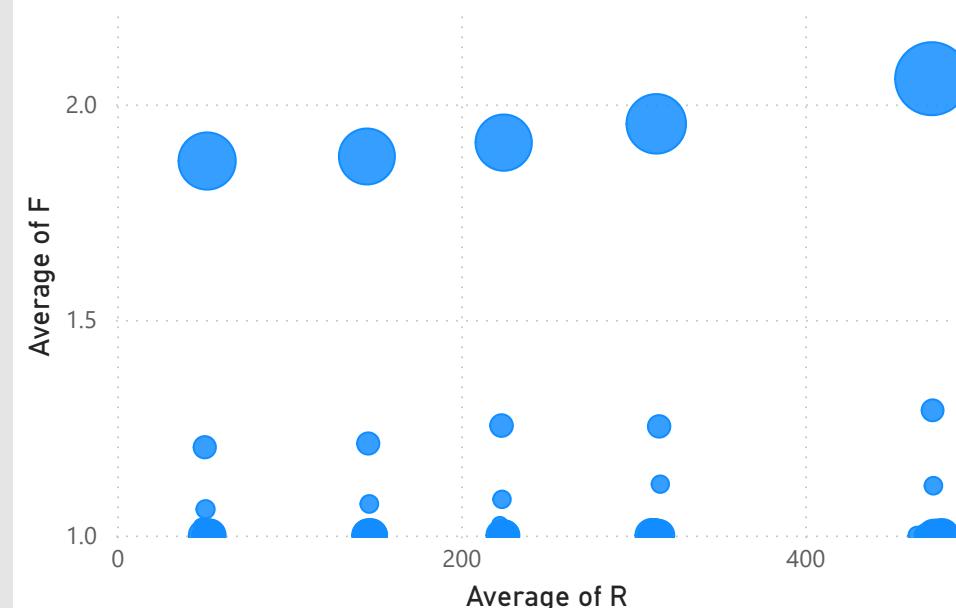
• What should we do next?

- Conversion Focus:** Target "Low-Value Occasional" customers with bundle offers to increase their purchase frequency.
- Retention Fix:** Deploy a re-engagement campaign in January to prevent the annual post-holiday "churn cliff".

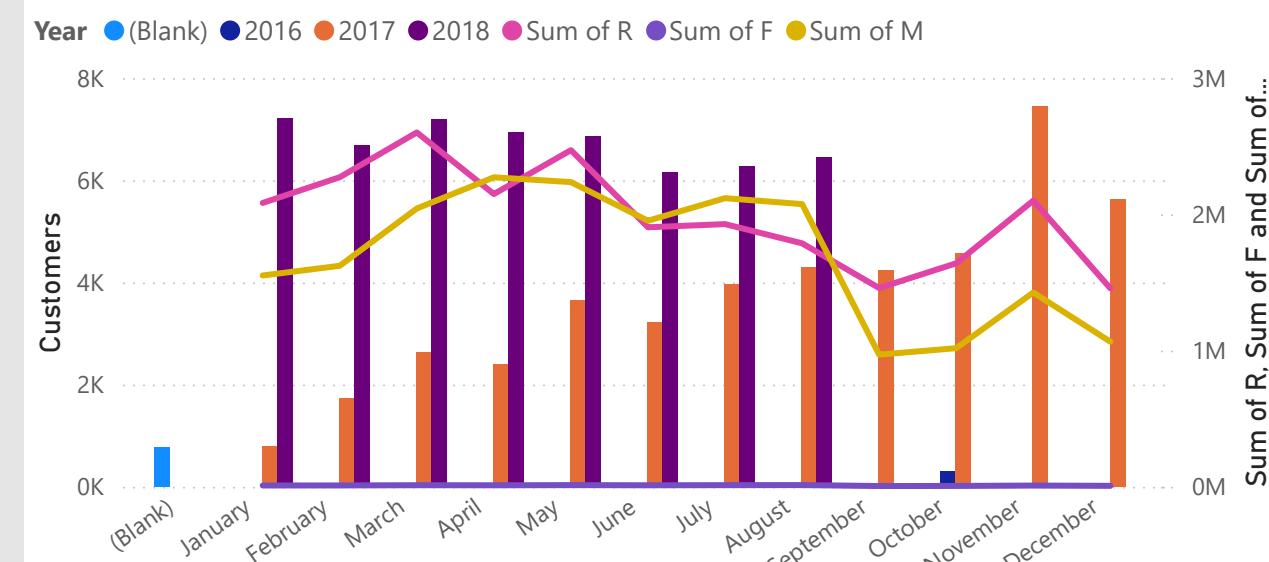
Revenue, Orders and Customer Metrics by Profile



RFM Metrics by RFM Score



Customer Retention Snapshot





Churn Prediction

shopping

Date

All

Channel

All

Segment

All

Location

All

Total Revenue

13.02M

Avg Order Value

122.21

Total Orders

107K

Customer Count

96K

Forecasted Orders

983.75

Churn Rate

80.41

Avg Review Score

3.67

Churn Risk Level

High Risk

What is happening? Churn is at a critical **80.4%**. Our largest customer group (40.4%) is currently "At-Risk," having not purchased in over **383 days** on average.

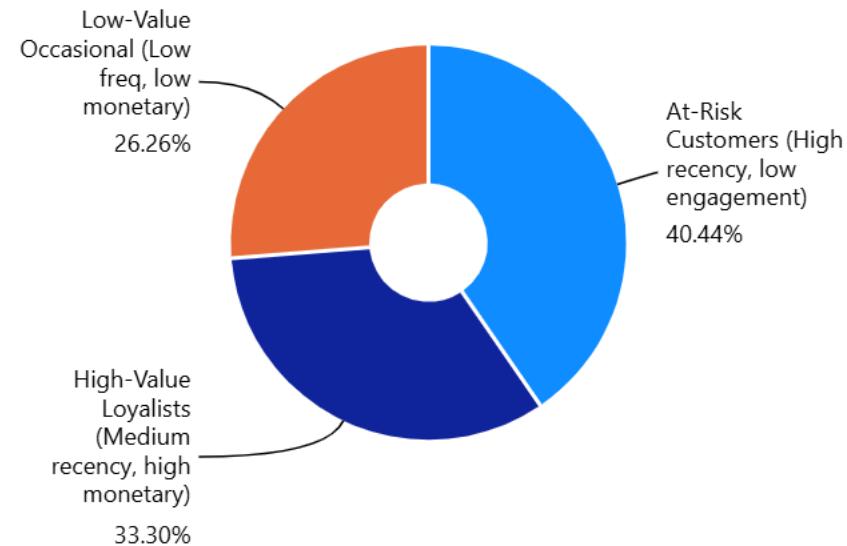
Why is it happening? Low engagement and stagnant review scores (**3.67/5**). We are losing "High-Value Loyalists" as they cross the 200-day inactivity mark without a re-engagement trigger.

What should we do next?

- Automate Win-Back:** Send high-value offers to "At-Risk" users immediately.
- Protect Margin:** Incentivize "Loyalists" (AOV \$245) to purchase before they hit the 250-day dormancy threshold.



Churn Rate by profile



profile	Avg Order Value	Average of days_since_last_order	Sum of response_rate	Sum of conversions	Sum of churn_flag	Average of review_score
At-Risk Customers (High recency, low engagement)	87.88	383.43	97.36	88	34614	3.78
High-Value Loyalists (Medium recency, high monetary)	245.83	239.60	30.12	34	15353	3.56
Low-Value Occasional (Low freq, low monetary)	89.09	135.97	38.68	42	29997	3.63
Total	122.21	241.53	166.17	164	79964	3.67