



Edun Joshua
Mbuotidem Awak
Makinde Kayode



**TACKLING POST-HARVEST LOSS
WITH TECHNOLOGY**



Post-Harvest Losses Problem: Reducing Post-Harvest Losses to Build Youth-Led Agri-Businesses

Investigating Strategies to Mitigate PHL

Edun Joshua

Mbuotidem Awak

Makinde Kayode

Organizers: Data Community Africa, AgriConnect, Dicey Tech, REDtech
Africa

Academic report

Lagos, November 2025

Post-Harvest Losses Problem: Reducing Post-Harvest Losses to Build Youth-Led Agri-Businesses

Copyright © 2025 - Edun Joshua, Mbuotidem Awak, Makinde Kayode; Data Community Africa.

This dissertation is original work, written solely for this purpose, and all the authors whose studies and publications contributed to it have been duly cited. Partial reproduction is allowed with acknowledgment of the author and reference to the year, institution - *Data Community Africa* - and public defense date.

Acknowledgements

Our deepest thanks go to the **Data Community Africa** community. We attribute significant professional growth to the foundation and resources they consistently provide.

We are also immensely grateful to **REDTech** and DiceyTech **DiceyTech** for sponsoring valuable hackathons like this. They empower data enthusiasts by offering crucial opportunities to challenge themselves and showcase their abilities.

Abstract

In Nigeria, millions of tons of food go to waste every year; not because of poor harvests, but because they never make it from farms to markets. To tackle these losses, we present Okra — a data-driven digital marketplace and logistics platform built to seamlessly connect farmers, buyers, and logistics providers in real time. Its stand-out feature is an AI-powered tool for predicting freshness and quantity of produce using image recognition, and helping buyers make faster, more informed decisions. Okra also uses data such as rainfall forecasts and harvest schedules to predict high-risk post-harvest loss periods, enabling early interventions like dispatching logistics or alerting farmers.

Keywords: Post Harvest Loss, Digital Marketplace, AI.

Contents

<i>List of Figures</i>	xi
<i>List of Tables</i>	xiv
1 Introduction	1
1.1 Business Performance Overview	1
1.2 Channel Distribution	1
2 Channel Performance	3
2.1 Introduction	3
2.2 Key Findings	3
2.2.1 The Profitability Paradox	3
2.2.2 Operational Metrics: Volume vs. Value	4
2.2.3 Geographic and Personnel Performance	4
2.3 Business Insights and Recommendations	4
2.4 Limitations	6
3 Customer Segmentation	7
3.1 Introduction	7
3.2 Methodology	7
3.3 Key Findings	8
3.3.1 Segment Distribution	8
3.3.2 The 161-Day Repeat Purchase Cycle	8
3.3.3 New vs. Repeat Customer Value	9
3.3.4 Geographic Concentration of Top Customers	9
3.4 Business Insights and Recommendations	9
3.5 Limitations	10
4 Product Optimization	11
4.1 Introduction	11
4.2 Methodology	11
4.3 Key Findings	11
4.3.1 Hero vs. Laggard Products	11
4.3.2 Market Basket and Bundling Opportunities	12
4.3.3 Channel-Specific Product Performance	12

4.3.4	The Corrosive Impact of Aggressive Discounting	13
4.4	Business Insights and Recommendations	13
4.5	Limitations	14
5	Conclusion and Recommendations	15
5.1	Summary of Findings	15
5.2	High-Impact Business Recommendations	16

Appendices

A	SQL Views	25
A.1	Theme 1: Channel Performance	25
A.2	Theme 2: Customer Segmentation	25
A.3	Theme 3: Product Optimization	25

List of Figures

1.1 Revenue by channel	2
2.1 Regional revenue distribution across territories (all years)	4
2.2 Trend of revenue growth by territory	5
3.1 Customer count distribution across RFM segments.	8
3.2 Average order value comparison: new versus repeat customers by month. .	9
4.1 Revenue vs. Profit Margin for Top 10 Products.	12
4.2 Category revenue mix by channel.	13
4.3 The impact of discounting on profitability across product categories.	13

List of Tables

1.1	Key Performance Indicators (KPIs)	1
1.2	Total Revenue by Sales Channel	2
2.1	Yearly Revenue and Profit Margin Comparison	3
2.2	Top 5 Salespeople by Revenue (Reseller Channel)	5
3.1	Customer Segmentation Distribution	8
4.1	Top 5 Frequently Purchased Product Pairs	12

1

Introduction

AdventureWorks operates as a global manufacturing entity, distributing products through a dual-channel strategy comprising an online direct-to-consumer interface and a robust reseller network. This report showcases findings from analysis of the business from 3 lenses: channel performance, customer behavior, and product trends. The objective is to identify actionable insights of business relevance, backed by quantitative evidence.

Tip

AdventureWorks has generated a total revenue of \$109.85M with a net profit of \$9.37M. The business relies heavily on its Reseller channel, which accounts for approximately 73% of total revenue.

1.1 Business Performance Overview

Initial exploratory data analysis of the dataset reveals a business with a substantial market footprint. The aggregate financial and operational metrics indicate a mature operation with a significant customer base. As detailed in [Table 1.1](#), the company has processed sales amounting to nearly \$110 million, serving over 19,000 customers with a high Average Order Value (AOV).

Table 1.1: Key Performance Indicators (KPIs)

Metric	Value
Total Revenue	\$109.85 M
Total Profit	\$9.37 M
Total Customers	19.12 K
Average Order Value (AOV)	\$3.49 K

1.2 Channel Distribution

The revenue stream of the business is bifurcated into Reseller and Online channels. [Figure 1.1](#) illustrates the volume of the Reseller channel, contributing the majority of

gross revenue.

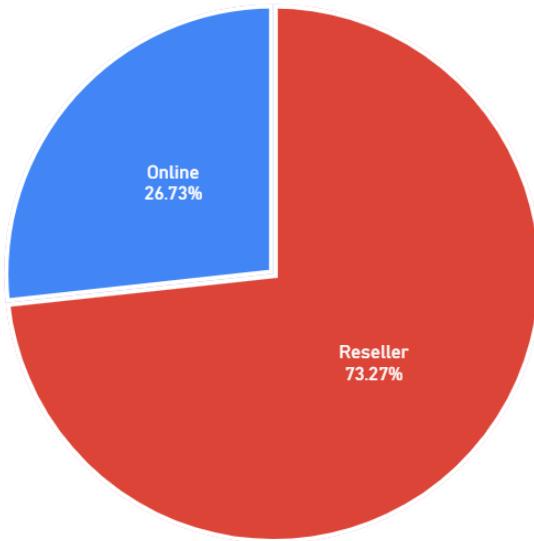


Figure 1.1: Revenue by channel

As presented in [Table 1.2](#), the Reseller channel generated approximately \$80.5 million, while the Online channel contributed \$29.4 million. This distribution suggests a B2B-centric business model where bulk purchasing or wholesale agreements has the greatest impact on financial performance.

Table 1.2: Total Revenue by Sales Channel

Channel	Total Revenue (\$)	Share of Total (%)
Reseller	80,487,704.18	73.27%
Online	29,358,677.22	26.73%
Total	109,846,381.40	100.00%

2

Channel Performance Analysis

2.1 Introduction

This chapter analyzes the performance of AdventureWorks' two primary sales channels: the Online (Direct-to-Consumer) and Reseller (Business-to-Business) channels. The objective is to evaluate their financial health, operational differences, and strategic value to the business. While the Reseller channel drives the majority of sales volume, the Online channel serves as the more stable and profitable revenue stream.

2.2 Key Findings

2.2.1 The Profitability Paradox

A core finding from the 2011–2014 data is the inverse relationship between revenue volume and profitability across channels. The Reseller channel, despite generating the bulk of revenue, consistently operates at a loss, whereas the Online channel maintains a high profit margin.

As detailed in [Table 2.1](#), the Reseller channel's revenue peaked at \$32.89 million in 2013, yet it incurred a loss of \$0.94 million (-2.85% margin). In contrast, the Online channel generated \$10.73 million in revenue that same year with a 40% profit margin. This suggests that aggressive volume-based discounts for resellers are severely eroding profitability.

Table 2.1: Yearly Revenue and Profit Margin Comparison

Year	Channel	Total Revenue (\$)	Total Profit (\$)	Margin (%)
2011	Online	3.86 M	1.54 M	39.91%
	Reseller	8.78 M	0.08 M	0.97%
2012	Online	6.39 M	2.38 M	37.28%
	Reseller	27.13 M	(1.43 M)	-5.29%
2013	Online	10.73 M	4.29 M	40.00%
	Reseller	32.89 M	(0.03 M)	-0.24%

*2014 data represents a partial fiscal year.

2.2.2 Operational Metrics: Volume vs. Value

The operational metrics of each channel highlight their distinct roles:

- **Reseller Channel:** Functions as a bulk-wholesale operation with a high Average Order Value (AOV) of **\$21,147.58** and an average of **16 items per order**.
- **Online Channel:** Operates as a high-frequency retail model with a lower AOV of **\$1,061.45** and an average of **2.18 items per order**.

This confirms that the channels serve fundamentally different customer needs and require distinct logistical and sales strategies.

2.2.3 Geographic and Personnel Performance

The North American market, particularly the **Southwest** territory (**\$7.19M** in 2013 revenue), remains the company's stronghold. However, international markets like the **United Kingdom** (1.23% YoY growth) and **France** (1.61% YoY growth) are emerging as key expansion targets.

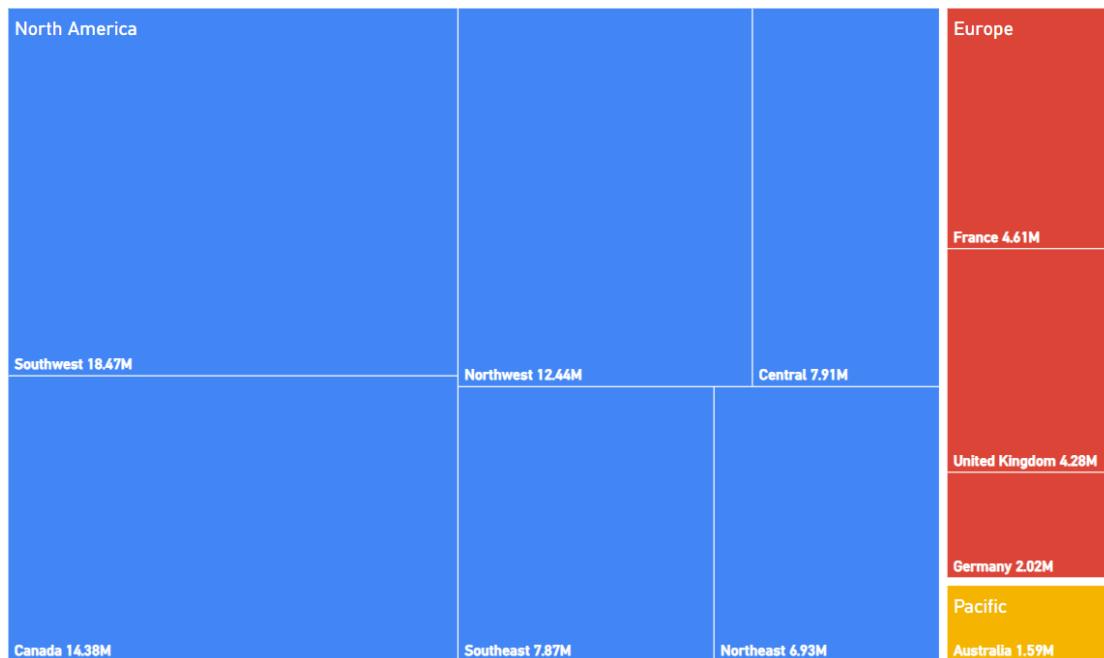


Figure 2.1: Regional revenue distribution across territories (all years).

Furthermore, analysis of the salesforce reveals a potential misalignment of incentives. As shown in **Table 2.2**, the top five revenue-generating salespeople in the Reseller channel all produced negative profit margins. For instance, Michael Blythe, the top performer with \$9.29M in revenue, generated a loss of \$281k.

2.3 Business Insights and Recommendations

The analysis points to a critical strategic tension: the channel that delivers the most revenue is also the one destroying the most value. The profitability paradox is not

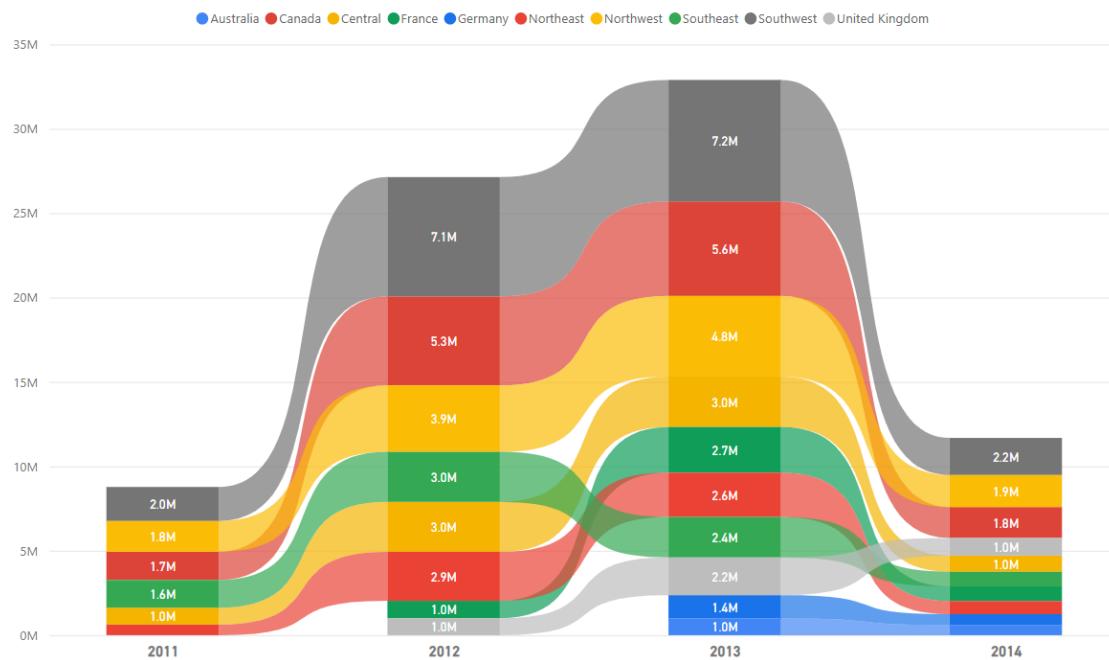


Figure 2.2: Trend of revenue growth by territory.

Table 2.2: Top 5 Salespeople by Revenue (Reseller Channel)

Salesperson	Total Revenue	Total Profit	Profit Margin
Michael Blythe	\$9,293,903	(\$281,662)	-3.03%
Jae Pak	\$8,503,338	(\$142,034)	-1.67%
Tsvi Reiter	\$7,171,012	(\$147,095)	-2.05%
Shu Ito	\$6,427,005	(\$396,323)	-6.17%
Amy Alberts	\$732,759	(\$24,279)	-3.31%

just a data anomaly but a symptom of flawed strategic priorities. The fact that top salespeople are also top loss-generators suggests that current incentive structures are counterproductive.

Key recommendations include:

- Restructure Reseller Incentives:** Shift sales commissions and targets from being based on gross revenue to net profit. This will disincentivize deep, unprofitable discounting to close deals.
- Implement Dynamic Pricing Tiers:** For the Reseller channel, introduce a tiered pricing structure where discounts are tied to volume thresholds that protect profit margins. For example, a 10% discount might apply to orders over \$50k, while a 15% discount requires orders over \$100k, with floors to prevent losses.
- Invest in Online Channel Growth:** Given its high profitability, the Online channel represents the most promising area for sustainable growth. Allocate marketing budget to campaigns aimed at increasing online traffic, conversion rates, and customer lifetime value.
- Focus International Expansion on Profitable Channels:** As the business expands into markets like the UK and France, prioritize establishing and growing the Online channel first to build a profitable foundation before scaling up reseller operations.

2.4 Limitations

- The analysis is based on data up to mid-2014, so more recent trends are not captured.
- The dataset does not include all operational costs (e.g., marketing spend, logistics overhead per channel), so the true net profitability of each channel may differ from the calculated margins.
- The root cause of the Reseller channel's unprofitability (e.g., specific products, customer segments, or regions) requires deeper, more granular analysis.

3

Customer Segmentation Analysis

3.1 Introduction

Following the analysis of channel performance, this chapter delves into the customer base to identify distinct segments with similar purchasing behaviors. By grouping customers, AdventureWorks can tailor marketing strategies, optimize resource allocation, and enhance customer lifetime value (CLV). The analysis focuses exclusively on the Online channel to understand direct-to-consumer behavior.

3.2 Methodology

The primary method employed was **RFM Analysis (Recency, Frequency, Monetary)**, a rule-based segmentation technique chosen for its interpretability and direct applicability to marketing actions.

Data Preparation: The analysis used sales transaction data from the Online channel, joined with customer demographic and geographic information. Key data sources included aggregated transactional data, geographic distributions, and repeat purchase metrics. Preprocessing involved:

1. **Filtering:** Restricting the dataset to Online channel sales.
2. **Aggregation:** Grouping transactions by CustomerID to calculate RFM metrics.
3. **Feature Engineering:** Deriving Recency (days since last purchase), Frequency (total number of orders), and Monetary (total spend).

Scoring Logic: Customers were scored on a scale of 1 to 4 for each RFM dimension using the NTILE(4) window function:

- **Recency (R):** Higher scores for more recent purchases.

- **Frequency (F):** Higher scores for more frequent purchases.
- **Monetary (M):** Higher scores for higher total spend.

An aggregate RFM_Score ($R + F + M$) was used to classify customers into five segments: *Champions*, *Loyal Customers*, *Potential Loyalists*, *At-Risk Customers*, and *Lost Customers*.

3.3 Key Findings

3.3.1 Segment Distribution

The analysis of 18,484 online customers revealed a highly concentrated value structure. As shown in [Table 3.1](#), the **Champions** (5.5% of customers) and **Loyal Customers** (33.8%) represent the high-value core of the business. Conversely, the **At-Risk** and **Lost** segments make up over 20% of the customer base, representing a significant pool of disengaged accounts.

Table 3.1: Customer Segmentation Distribution

Segment	Customer Count	% of Base	Business Implication
Champions	1,007	5.5%	High-value core; maximize retention.
Loyal Customers	6,257	33.8%	Stable revenue base; nurture loyalty.
Potential Loyalists	7,319	39.6%	Largest segment; prime growth target.
At-Risk Customers	2,907	15.7%	Require win-back; high recovery value.
Lost Customers	994	5.4%	Low-touch or reactivation pilots only.
Total	18,484	100%	—

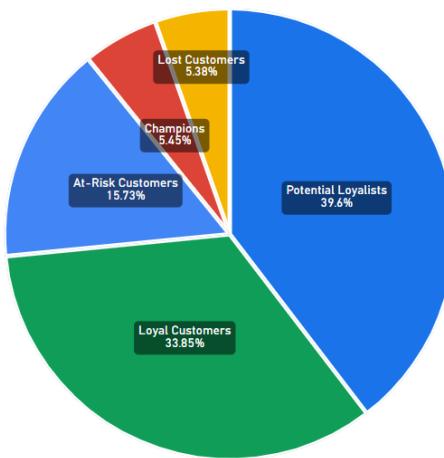


Figure 3.1: Customer count distribution across RFM segments.

3.3.2 The 161-Day Repeat Purchase Cycle

A key behavioral insight is that the average time to a second purchase is **161 days** (approximately 5.3 months). This metric provides a clear window of opportunity for targeted marketing interventions to encourage repeat business and prevent customer churn.

3.3.3 New vs. Repeat Customer Value

Repeat customers are significantly more valuable than new customers, with Average Order Values (AOV) that are **8 to 20 times higher**. New customers typically have an AOV around \$1,600, while repeat customers spend \$6,000 or more per order. This highlights the strategic importance of retention over acquisition.

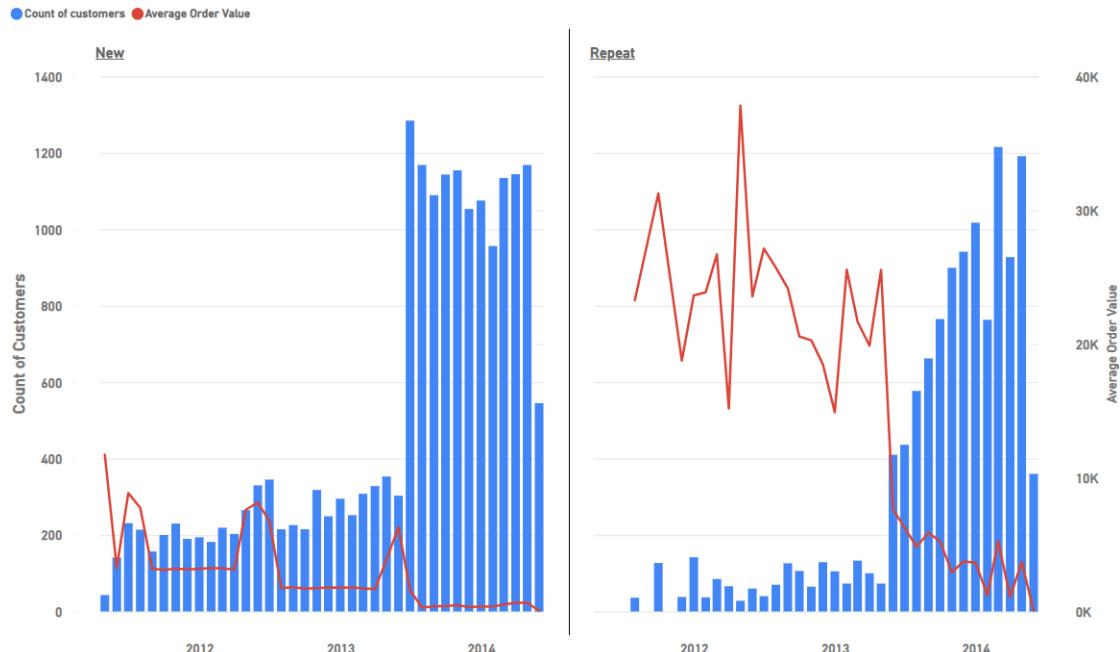


Figure 3.2: Average order value comparison: new versus repeat customers by month.

3.3.4 Geographic Concentration of Top Customers

The top 10% of customers are highly concentrated geographically. **England** (66 top customers) and the **Paris** region (33 top customers) are the most valuable international clusters. Australia also represents a significant base, with 52 top customers across Queensland and New South Wales.

3.4 Business Insights and Recommendations

The segmentation analysis provides a clear roadmap for data-driven marketing and retention strategies. The extreme concentration of value in a small customer segment, combined with the predictable 161-day repeat purchase cycle, offers powerful levers for growth.

- 1. Implement a Timed Nurture Campaign:** Launch an automated email and retargeting campaign that triggers between 90 and 120 days after a customer's first purchase. The goal is to re-engage customers during their consideration phase, just before the 161-day average repeat purchase window closes.

2. **Develop a "Champion" VIP Program:** Create an exclusive VIP program for the "Champion" segment. Benefits could include early access to new products, a dedicated customer service line, and invitations to special events. This focuses resources on retaining the most profitable customer group.
3. **Launch a "Potential to Loyal" Conversion Initiative:** Target the largest segment, "Potential Loyalists," with campaigns designed to increase their purchase frequency. Offer a small incentive (e.g., 15% off) for making a second purchase within 90 days to shorten the natural repeat cycle.
4. **Focus Geographic Marketing Efforts:** Allocate marketing and potentially logistical resources to the top geographic clusters: England, Paris, and Australia. This could include localized marketing content, targeted digital advertising, and exploring regional fulfillment options to improve service.

3.5 Limitations

- **Rank-Based Scoring:** The NTILE quartile system can mask variance within segments. For example, a customer with a \$50k lifetime spend might be in the same monetary quartile as one with a \$5k spend.
- **Data Recency:** Recency calculations are tied to the date the analysis was run and may not reflect current customer status.
- **Temporal Bias:** Newer customer cohorts have had less time to make repeat purchases, which may skew retention metrics downward compared to older cohorts.
- **Geographic Granularity:** The data is at the state/province level, which prevents more granular, city-level micro-targeting.

4

Product Optimization

4.1 Introduction

This chapter evaluates product performance to identify opportunities for portfolio optimization, improved profitability, and strategic cross-selling. The analysis focuses on four goals: identifying profitability extremes (hero vs. laggard products), analyzing cross-selling patterns, evaluating channel-specific trends, and assessing discount sensitivity.

4.2 Methodology

The analysis relied on sales and product data to assess profitability, bundling opportunities, and discount impacts. Methodologies included:

- **Profitability Analysis:** Products were ranked by revenue and profit margin to identify top and bottom performers.
- **Market Basket Analysis:** A self-join on sales data was used to find frequently co-purchased product pairs, revealing bundling opportunities.
- **Discount Impact Analysis:** Sales were segmented into discount buckets (e.g., 0-5%, >30%) to measure the effect of discounts on sales volume and profitability for different product categories.
- **Channel-Specific Analysis:** Product category performance was compared across the Online and Reseller channels to identify distinct sales patterns.

4.3 Key Findings

4.3.1 Hero vs. Laggard Products

The product portfolio exhibits significant performance variance. The *Mountain-200* series stands out as a "hero" product line, consistently delivering high revenue and strong profit margins (15-20%). Conversely, the *Road-250 Black* series, despite being

a top-10 revenue generator, is a "laggard" product that operates at a negative profit margin, effectively losing money on each sale.

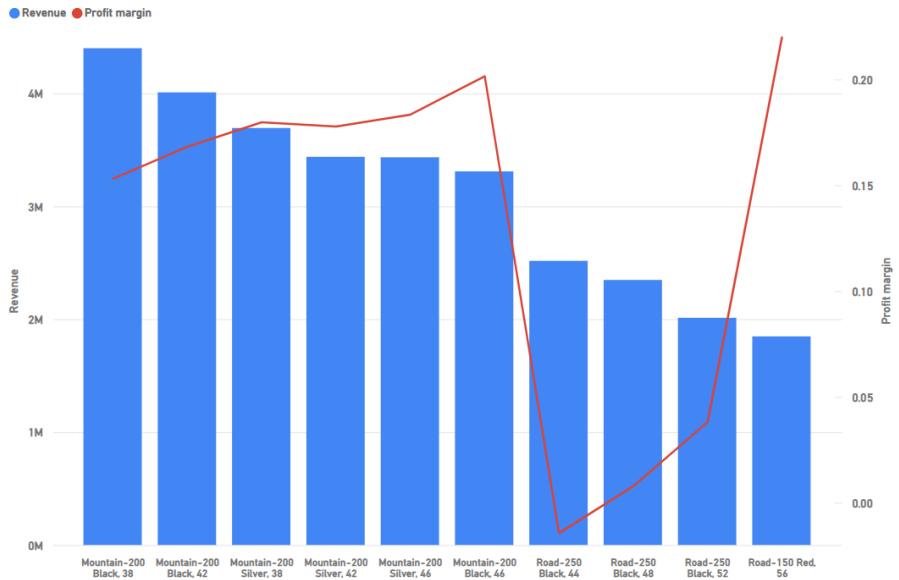


Figure 4.1: Revenue vs. Profit Margin for Top 10 Products.

4.3.2 Market Basket and Bundling Opportunities

Market basket analysis revealed strong cross-selling patterns that suggest clear bundling opportunities. Key patterns include:

- **Component Pairing:** High-frequency pairing of components like frames and forks, common in reseller assembly orders.
- **Accessory Add-ons:** Frequent co-purchase of accessories, such as helmets and patch kits, particularly in the Online channel.
- **Upsell Paths:** A significant number of customers purchasing both *Mountain-200* and *Mountain-300* series bikes, indicating a clear upsell opportunity.

Table 4.1: Top 5 Frequently Purchased Product Pairs

Product A	Product B	Frequency	Insight
LL Road Frame - Red, 48	Road-650 Black, 58	225	High-volume component pairing.
Road-650 Black, 58	Road-650 Black, 60	319	Bulk reseller stocking.
Sport-100 Helmet, Red	Patch Kit/8 Patches	132	Strong accessory add-on.
Mountain-200 Silver, 38	Mountain-300 Black, 40	193	Cross-model buyer overlap.
Hitch Rack - 4-Bike	Touring-3000 Yellow, 54	71	High-value accessory attachment.

4.3.3 Channel-Specific Product Performance

Product categories perform very differently depending on the sales channel. **Components** are sold exclusively through the Reseller channel, generating \$11.8M in revenue with zero online presence. In contrast, **Accessories** are most popular online, with 55% of their revenue coming from that channel. **Bikes** remain heavily reliant on the Reseller channel, which accounts for 70% of bike revenue.

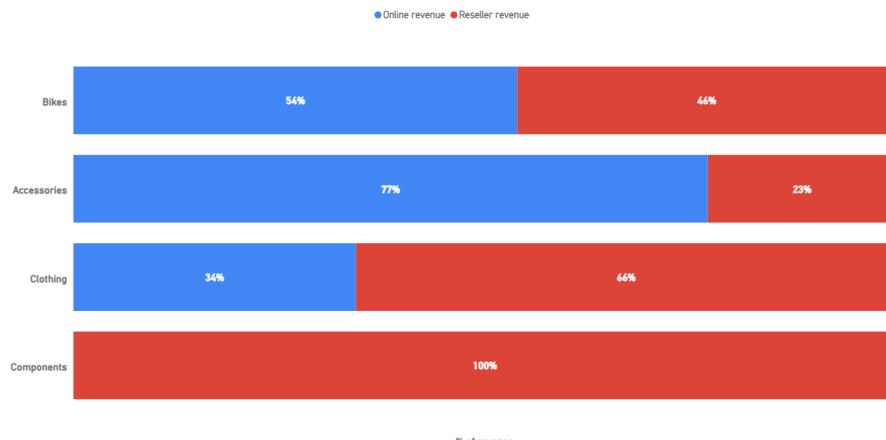


Figure 4.2: Category revenue mix by channel.

4.3.4 The Corrosive Impact of Aggressive Discounting

Discounting has a dramatically different impact across product categories. For Road Bikes, minimal discounts (0-5%) yield a healthy profit of \$62 per unit. However, deep discounts (>30%) on Mountain Bikes lead to a loss of **\$846 per unit**. This demonstrates that undisciplined discounting, likely used to drive volume in the Reseller channel, is a primary driver of unprofitability.



Figure 4.3: The impact of discounting on profitability across product categories.

4.4 Business Insights and Recommendations

The product portfolio is unbalanced, with a handful of hero products subsidizing a long tail of underperformers and outright loss-makers. Channel-specific product strategies and disciplined pricing are essential for improving overall profitability.

- 1. Conduct a profitability audit on the Road-250 series:** The *Road-250 Black* series is a major revenue driver but a net loss for the company. An immediate review of its cost of goods sold (COGS), pricing, and promotional strategy is required. The business must decide whether to adjust pricing to make it profitable or discontinue it.

2. **Implement an improved discounting policy:** Introduce a company-wide policy that caps discounts, particularly on high-value items like Mountain Bikes and Road Bikes. A 15% discount ceiling should be enforced, with any exceptions requiring senior management approval. This would immediately address the significant losses from deep discounting.
3. **Create strategic product bundles:** Capitalize on the market basket analysis by creating official product bundles. For example, a "Reseller Build Kit" could package popular frame and component pairs, while an "Online Adventure Bundle" could combine a bike with high-margin accessories like helmets and lights.
4. **Develop channel-specific product strategies:** Acknowledge and formalize the de facto channel specializations. Position the Reseller channel as the exclusive home for components and complex bike builds. Position the Online channel as the primary destination for accessories, clothing, and direct-to-consumer bike sales with a focus on higher-margin configurations.

4.5 Limitations

- **Cost Data Granularity:** The analysis is based on standard cost data, which may not capture all manufacturing and supply chain costs. The root cause of high COGS in laggard products requires a more detailed financial investigation.
- **Temporal Blind Spots:** The aggregate analysis does not account for product life-cycle stage (e.g., launch vs. decline) or seasonality, which could influence profitability.
- **Lack of Competitive Context:** Without market pricing data, it is difficult to determine whether losses on certain products are due to competitive pressures or internal inefficiencies.

5

Conclusion and Recommendations

5.1 Summary of Findings

This report has analyzed the AdventureWorks business through three critical lenses: channel performance, customer segmentation, and product optimization. The analysis reveals a company with a strong market presence but facing significant strategic challenges that impact its overall profitability.

The **Channel Performance** analysis highlighted a stark "profitability paradox": the Reseller channel generates 73% of total revenue but operates at a consistent loss, while the smaller Online channel is highly profitable, boasting margins of around 40%. This suggests that the company's growth engine is simultaneously its biggest drain on value.

The **Customer Segmentation** of the Online channel identified a highly concentrated customer base where a small percentage of "Champion" and "Loyal" customers drive a disproportionate share of revenue. Furthermore, the analysis pinpointed a clear, actionable **161-day average repeat purchase cycle**, providing a timeline for targeted marketing interventions.

The **Product Optimization** analysis revealed significant imbalances in the product portfolio. "Hero" products like the Mountain-200 series are highly profitable, but their gains are offset by "laggard" products like the Road-250 Black series, which generates substantial revenue at a net loss. The analysis also showed that undisciplined discounting is a major cause of unprofitability, particularly for high-value items like Mountain Bikes.

5.2 High-Impact Business Recommendations

Based on the synthesized findings, five high-impact recommendations are proposed to address the core strategic challenges and unlock sustainable, profitable growth for AdventureWorks.

1. Address the Profitability Paradox by Overhauling Reseller Strategy:

Insight: The Reseller channel is a revenue-driving but loss-making machine, likely due to misaligned incentives and aggressive discounting.

Recommendation: Immediately restructure the Reseller channel's strategy. Shift sales team incentives from gross revenue to **net profit**. Implement a mandatory tiered discount structure that protects margins, and consider establishing a minimum order value to ensure that the logistical complexity of reseller orders is justified by their profitability.

2. Focus Retention Efforts on High-Value Segments and Geographies:

Insight: A small fraction of online customers (5.5% "Champions") and a few key geographic clusters (England, Paris, Australia) deliver a large portion of the business's value.

Recommendation: Launch a "**Champion**" VIP program to nurture and retain these top-tier customers. Simultaneously, concentrate digital marketing spend and potentially logistical resources in the identified high-value geographic regions to maximize ROI and strengthen market presence where it matters most.

3. Leverage the 161-Day Repeat Purchase Cycle for Proactive Marketing:

Insight: Customers take, on average, 161 days to make a second purchase, creating a predictable window for engagement.

Recommendation: Implement an **automated marketing campaign** that triggers between day 90 and day 120 post-purchase. This campaign should use personalized content and modest incentives to re-engage customers during their natural consideration phase, aiming to shorten the repeat purchase cycle and increase customer lifetime value.

4. Aggressively Optimize the Product Portfolio for Profitability:

Insight: The product portfolio is unbalanced, with loss-making "laggards" canceling out the gains from profitable "heroes."

Recommendation: Conduct an urgent **profitability audit of the top 20 revenue-generating products**, starting with the Road-250 Black series. Based on the findings, either adjust prices to ensure profitability or make data-driven decisions to discontinue unprofitable products. Simultaneously, promote and invest in the "hero" products that are proven to be profitable.

5. Formalize Channel-Specific Product Strategies and Bundles:

Insight: Product categories have clear affinities for specific channels (e.g., Components for Resellers, Accessories for Online), and certain products are frequently bought together.

Recommendation: Create and market **strategic product bundles** tailored to each channel. For the Reseller channel, offer "Reseller Build Kits" that package frequently paired components. For the Online channel, create "Adventure Bundles" that pair high-margin accessories with bikes. This approach will increase Average Order Value (AOV) while respecting the natural purchasing behavior of each channel's customer base.

Appendices

A

SQL Views

This appendix contains the SQL code for the views created to support the analysis in the main body of the report.

A.1 Theme 1: Channel Performance

Listing 1: View for Channel Revenue Comparison

```
1 CREATE OR ALTER VIEW Analytics.vChannelPerformance AS
2 SELECT
3     YEAR(OrderDate) AS OrderYear,
4     Channel,
5     SUM(LineTotal) AS TotalRevenue,
6     SUM(LineProfit) AS TotalProfit,
7     CASE WHEN SUM(LineTotal) = 0 THEN NULL
8         ELSE SUM(LineProfit) * 1.0 / SUM(LineTotal) END AS ProfitMargin
9     FROM Analytics.Fact_Sales
10    GROUP BY YEAR(OrderDate), Channel;
```

A.2 Theme 2: Customer Segmentation

A.3 Theme 3: Product Optimization

Listing 2: View for Monthly Sales Trends

```

1 CREATE OR ALTER VIEW Analytics.v_ChannelMonthly AS
2 SELECT
3     YEAR(OrderDate) AS OrderYear,
4     MONTH(OrderDate) AS OrderMonth,
5     DATEFROMPARTS(YEAR(OrderDate), MONTH(OrderDate), 1) AS MonthStart,
6     Channel,
7     SUM(LineTotal) AS MonthlyRevenue,
8     COUNT(DISTINCT SalesOrderID) AS NumberOfOrders
9 FROM Analytics.Fact_Sales
10 GROUP BY YEAR(OrderDate), MONTH(OrderDate), DATEFROMPARTS(YEAR(OrderDate),
11   ↳ MONTH(OrderDate), 1), Channel;

```

Listing 3: View for Geospatial Performance

```

1 CREATE OR ALTER VIEW Analytics.v_ResellerTerritoryYoY AS
2 WITH TerritorySales AS (
3     SELECT
4         st.Name AS TerritoryName,
5         st.[Group] AS TerritoryGroup,
6         YEAR(fs.OrderDate) AS OrderYear,
7         SUM(fs.LineTotal) AS TotalRevenue
8     FROM Analytics.Fact_Sales fs
9     JOIN Sales.SalesTerritory st ON fs.TerritoryID = st.TerritoryID
10    WHERE fs.Channel = 'Reseller'
11    GROUP BY st.Name, st.[Group], YEAR(fs.OrderDate)
12 ),
13 YoY AS (
14     SELECT
15         TerritoryName,
16         TerritoryGroup,
17         OrderYear,
18         TotalRevenue,
19         LAG(TotalRevenue) OVER (PARTITION BY TerritoryName ORDER BY OrderYear) AS
20           ↳ PrevRevenue
21     FROM TerritorySales
22 )
23 SELECT
24     TerritoryName,
25     TerritoryGroup,
26     OrderYear,
27     TotalRevenue,
28     PrevRevenue,
29     CASE
30         WHEN PrevRevenue IS NULL OR PrevRevenue = 0 THEN NULL
31         ELSE (TotalRevenue - PrevRevenue) * 1.0 / PrevRevenue
32     END AS YoYGrowthPct
33 FROM YoY
34 ORDER BY TotalRevenue DESC, OrderYear;

```

Listing 4: View for Average Order Metrics

```

1 CREATE OR ALTER VIEW Analytics.v_ChannelOrderMetrics AS
2 WITH OrderMetrics AS (
3     SELECT
4         SalesOrderID,
5         Channel,
6         SUM(LineTotal) AS OrderTotalValue,
7         COUNT(SalesOrderDetailID) AS ItemsPerOrder
8     FROM Analytics.Fact_Sales
9     GROUP BY SalesOrderID, Channel
10    )
11    SELECT
12        Channel,
13        AVG(OrderTotalValue) AS AverageOrderValue,
14        AVG(CAST(ItemsPerOrder AS FLOAT)) AS AverageItemsPerOrder,
15        COUNT(*) AS TotalOrders
16    FROM OrderMetrics
17    GROUP BY Channel;

```

Listing 5: View for Top Salespeople

```

1 CREATE OR ALTER VIEW Analytics.v_TopResellerSalespeople AS
2 WITH SalesPersonPerformance AS (
3     SELECT
4         fs.SalesPersonID,
5         COALESCE(p.FirstName + ' ' + p.LastName, 'Unknown') AS SalesPersonName,
6         SUM(fs.LineTotal) AS TotalRevenue,
7         SUM(fs.LineProfit) AS TotalProfit
8     FROM Analytics.Fact_Sales fs
9     LEFT JOIN Person.Person p ON fs.SalesPersonID = p.BusinessEntityID
10    WHERE fs.Channel = 'Reseller' AND fs.SalesPersonID IS NOT NULL
11    GROUP BY fs.SalesPersonID, p.FirstName, p.LastName
12    )
13    SELECT TOP (5)
14        SalesPersonName,
15        TotalRevenue,
16        TotalProfit,
17        CASE WHEN TotalRevenue = 0 THEN NULL ELSE TotalProfit * 1.0 / TotalRevenue END AS
18            ↳ ProfitMargin
19    FROM SalesPersonPerformance
20    ORDER BY TotalRevenue DESC;

```

Listing 6: View for RFM Analysis

```

1 CREATE OR ALTER VIEW Analytics.v_RFMs_Online AS
2 WITH RFM_Base AS (
3     SELECT
4         c.CustomerID,
5             COALESCE(p.FirstName + ' ' + p.LastName, 'Company/Unknown') AS CustomerName,
6             MAX(fs.OrderDate) AS LastOrderDate,
7             DATEDIFF(day, MAX(fs.OrderDate), GETDATE()) AS Recency,
8             COUNT(DISTINCT fs.SalesOrderID) AS Frequency,
9             SUM(fs.LineTotal) AS Monetary
10    FROM Analytics.Fact_Sales fs
11   JOIN Sales.Customer c ON fs.CustomerID = c.CustomerID
12  LEFT JOIN Person.Person p ON c.PersonID = p.BusinessEntityID
13 WHERE fs.Channel = 'Online'
14 GROUP BY c.CustomerID, p.FirstName, p.LastName
15 ),
16 RFM_Scores AS (
17     SELECT
18         CustomerID,
19         CustomerName,
20         Recency,
21         Frequency,
22         Monetary,
23         NTILE(4) OVER (ORDER BY Recency ASC) AS R_Score,      -- smaller recency = better
24         NTILE(4) OVER (ORDER BY Frequency DESC) AS F_Score,    -- larger frequency = better
25         NTILE(4) OVER (ORDER BY Monetary DESC) AS M_Score     -- larger monetary = better
26    FROM RFM_Base
27 )
28 SELECT
29     CustomerID,
30     CustomerName,
31     Recency,
32     Frequency,
33     Monetary,
34     R_Score, F_Score, M_Score,
35     (R_Score + F_Score + M_Score) AS RFM_Score,
36     CASE
37         WHEN (R_Score + F_Score + M_Score) >= 11 THEN 'Champions'
38         WHEN (R_Score + F_Score + M_Score) >= 9 THEN 'Loyal Customers'
39         WHEN (R_Score + F_Score + M_Score) >= 6 THEN 'Potential Loyalists'
40         WHEN (R_Score + F_Score + M_Score) >= 4 THEN 'At-Risk Customers'
41         ELSE 'Lost Customers'
42     END AS CustomerSegment
43 FROM RFM_Scores
44 ORDER BY RFM_Score DESC;

```

Listing 7: View for New vs. Repeat Customers

```

1 CREATE OR ALTER VIEW Analytics.v_MonthlyNewRepeat AS
2 WITH CustomerFirstOrder AS (
3     SELECT CustomerID, MIN(OrderDate) AS FirstOrderDate
4     FROM Analytics.Fact_Sales
5     GROUP BY CustomerID
6 ),
7 MonthlyCustomerActivity AS (
8     SELECT
9         fs.SalesOrderID,
10        fs.CustomerID,
11        fs.OrderDate,
12        fs.LineTotal,
13        CASE WHEN CAST(fs.OrderDate AS DATE) = CAST(cfo.FirstOrderDate AS DATE) THEN 'New'
14             ELSE 'Repeat' END AS CustomerType
15     FROM Analytics.Fact_Sales fs
16     JOIN CustomerFirstOrder cfo ON fs.CustomerID = cfo.CustomerID
17 )
18     SELECT
19         FORMAT(OrderDate, 'yyyy-MM') AS OrderMonth,
20         CustomerType,
21         COUNT(DISTINCT CustomerID) AS NumberOfCustomers,
22         SUM(LineTotal) * 1.0 / NULLIF(COUNT(DISTINCT SalesOrderID),0) AS AverageOrderValue
23     FROM MonthlyCustomerActivity
24     GROUP BY FORMAT(OrderDate, 'yyyy-MM'), CustomerType
25     ORDER BY OrderMonth, CustomerType;

```

Listing 8: View for Time to Repeat Purchase

```

1 CREATE OR ALTER VIEW Analytics.vAvgDaysToSecondPurchase AS
2 WITH Purchases AS (
3     SELECT
4         CustomerID,
5         OrderDate,
6         ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY OrderDate) AS rn,
7         LEAD(OrderDate) OVER (PARTITION BY CustomerID ORDER BY OrderDate) AS NextOrderDate
8     FROM Analytics.Fact_Sales
9     WHERE Channel = 'Online'
10    )
11    SELECT
12        AVG(CAST(DATEDIFF(day, OrderDate, NextOrderDate) AS FLOAT)) AS AvgDaysToRepeatPurchase
13    FROM Purchases
14    WHERE rn = 1 AND NextOrderDate IS NOT NULL;

```

Listing 9: View for Customer Geographics

```

1 CREATE OR ALTER VIEW Analytics.vTopCustomerGeography AS
2
3 WITH CustSpend AS (
4     SELECT
5         fs.CustomerID,
6         SUM(fs.LineTotal) AS TotalSpent
7     FROM Analytics.Fact_Sales fs
8     WHERE fs.Channel = 'Online'
9     GROUP BY fs.CustomerID
10 ),
11 Ranked AS (
12     SELECT
13         cs.*,
14         NTILE(10) OVER (ORDER BY TotalSpent DESC) AS Decile
15     FROM CustSpend cs
16 ),
17 TopDecile AS (
18     SELECT CustomerID, TotalSpent FROM Ranked WHERE Decile = 1 -- top 10%
19 )
20 SELECT
21     st.Name AS Territory,
22     sp.Name AS State,
23     a.City,
24     COUNT(DISTINCT td.CustomerID) AS NumberOfTopCustomers
25 FROM TopDecile td
26 JOIN Sales.Customer c ON td.CustomerID = c.CustomerID
27 JOIN Person.BusinessEntityAddress bea ON c.PersonID = bea.BusinessEntityID
28 JOIN Person.Address a ON bea.AddressID = a.AddressID
29 JOIN Person.StateProvince sp ON a.StateProvinceID = sp.StateProvinceID
30 JOIN Sales.SalesTerritory st ON sp.TerritoryID = st.TerritoryID
31 GROUP BY st.Name, sp.Name, a.City
32 ORDER BY NumberOfTopCustomers DESC;

```

Listing 10: View for Top and Bottom Products

```
1 CREATE OR ALTER VIEW Analytics.vTopBottomProducts AS
2 WITH ProdAgg AS (
3     SELECT
4         p.ProductID,
5         p.ProductName,
6         p.CategoryName,
7         p.SubcategoryName,
8         SUM(fs.LineTotal) AS TotalRevenue,
9         SUM(fs.LineProfit) AS TotalProfit
10    FROM Analytics.Fact_Sales fs
11   JOIN Analytics.Dim_Product p ON fs.ProductID = p.ProductID
12  GROUP BY p.ProductID, p.ProductName, p.CategoryName, p.SubcategoryName
13 ),
14 Top10 AS (
15     SELECT TOP (10) 'Top 10' AS Category, ProductID, ProductName, CategoryName,
16             SubcategoryName, TotalRevenue, TotalProfit
17     FROM ProdAgg ORDER BY TotalRevenue DESC
18 ),
19 Bottom10 AS (
20     SELECT TOP (10) 'Bottom 10' AS Category, ProductID, ProductName, CategoryName,
21             SubcategoryName, TotalRevenue, TotalProfit
22     FROM ProdAgg ORDER BY TotalRevenue ASC
23 )
24     SELECT *, CASE WHEN TotalRevenue = 0 THEN NULL ELSE TotalProfit * 1.0 / TotalRevenue
25             END AS ProfitMargin FROM Top10
26 UNION ALL
27     SELECT *, CASE WHEN TotalRevenue = 0 THEN NULL ELSE TotalProfit * 1.0 / TotalRevenue
28             END AS ProfitMargin FROM Bottom10;
```

Listing 11: View for Market Basket Analysis

```

1 CREATE OR ALTER VIEW Analytics.vFrequentProductPairs AS
2     WITH OrderProducts AS (
3         SELECT
4             fs.SalesOrderID,
5             p.ProductName,
6             p.ProductID
7         FROM
8             Analytics.Fact_Sales fs
9         JOIN
10            Analytics.Dim_Product p ON fs.ProductID = p.ProductID
11    )
12     SELECT
13         a.ProductName AS ProductA,
14         b.ProductName AS ProductB,
15         COUNT(*) AS PairFrequency
16     FROM OrderProducts a
17     JOIN OrderProducts b ON a.SalesOrderID = b.SalesOrderID AND a.ProductID < b.ProductID
18     GROUP BY a.ProductName,
19             b.ProductName
20     ORDER BY PairFrequency DESC;

```

Listing 12: View for Category by Channel

```

1 CREATE OR ALTER VIEW Analytics.vCategoryChannelRevenue AS
2     SELECT
3         p.CategoryName,
4         SUM(CASE WHEN fs.Channel = 'Online' THEN fs.LineTotal ELSE 0 END) AS OnlineRevenue,
5         SUM(CASE WHEN fs.Channel = 'Reseller' THEN fs.LineTotal ELSE 0 END) AS
6             ResellerRevenue,
7         SUM(fs.LineTotal) AS TotalRevenue
8     FROM Analytics.Fact_Sales fs
9     JOIN Analytics.Dim_Product p ON fs.ProductID = p.ProductID
10    GROUP BY p.CategoryName
11    ORDER BY TotalRevenue DESC;

```

Listing 13: View for Discount Impact Analysis

```
1 CREATE OR ALTER VIEW Analytics.vDiscountImpact AS
2 WITH DiscountBuckets AS (
3     SELECT
4         p.CategoryName,
5         p.SubcategoryName,
6         fs.UnitPriceDiscount,
7         CASE
8             WHEN fs.UnitPriceDiscount BETWEEN 0 AND 0.05 THEN '0-5%'
9             WHEN fs.UnitPriceDiscount > 0.05 AND fs.UnitPriceDiscount <= 0.15 THEN '5-15%'
10            WHEN fs.UnitPriceDiscount > 0.15 AND fs.UnitPriceDiscount <= 0.30 THEN '15-30%'
11            ELSE '>30%'
12        END AS DiscountBucket,
13        fs.OrderQty,
14        fs.LineProfit
15    FROM Analytics.Fact_Sales fs
16    JOIN Analytics.Dim_Product p ON fs.ProductID = p.ProductID
17    WHERE p.SubcategoryName IN ('Mountain Bikes', 'Road Bikes', 'Jerseys', 'Helmets')
18 )
19 SELECT
20     CategoryName,
21     SubcategoryName,
22     DiscountBucket,
23     SUM(OrderQty) AS TotalQuantitySold,
24     SUM(LineProfit) AS TotalProfit,
25     CASE WHEN SUM(OrderQty)=0 THEN NULL ELSE SUM(LineProfit) * 1.0 / SUM(OrderQty) END AS
26     → ProfitPerUnit
27 FROM DiscountBuckets
28 GROUP BY CategoryName, SubcategoryName, DiscountBucket
29 ORDER BY SubcategoryName, DiscountBucket;
```

