

# SHOPNEST STORE

## POWER BI CAPSTONE PROJECT

Comprehensive Business Analytics Report

Prepared by:	Deepak Lokhande
Date:	6/15/2025
Course:	Power BI Analytics Capstone
Institution:	Skillovilla

# TABLE OF CONTENTS

1. Executive Summary	3
2. Introduction and Background	4
3. Dataset Overview	5
4. Methodology	7
5. Dashboard Overview	9
6. Analytical Questions and Findings	10
6.1 Top Categories by Total Price	10
6.2 Delayed Orders Analysis	11
6.3 Monthly Comparison of Orders	12
6.4 Payment Method Analysis	13
6.5 Product Rating Analysis - Highest Rated	14
6.6 Product Rating Analysis - Lowest Rated	15
6.7 Seasonal Sales Patterns	16
6.8 Revenue Analysis and Yearly Trends	17
6.9 State-wise Sales Analysis	18
7. Key Insights and Recommendations	19
8. Challenges and Limitations	21
9. Conclusion	22
10. Appendix	23

# 1. EXECUTIVE SUMMARY

This comprehensive report presents the analysis of ShopNest Store's e-commerce performance through a Power BI dashboard addressing eight critical business questions. The analysis reveals significant insights across sales performance, customer behavior, operational efficiency, and geographic distribution.

## Key Performance Indicators

Metric	Value	Performance Indicator
Total Revenue	15.84M	Strong market presence
Total Orders	99K	High customer engagement
Delayed Orders	8K (8.1%)	Requires improvement
Average Customer Rating	4.04/5.0	Good customer satisfaction

## Major Findings

- **Category Leadership:** Health & Beauty emerges as the top-performing category with 1.26M in sales
- **Payment Preferences:** Credit cards dominate payment preferences at 74% of transactions
- **Seasonal Trends:** Q2 represents peak sales period while Q4 shows concerning decline
- **Geographic Concentration:** São Paulo state leads regional sales with 3.7M revenue
- **Quality Concerns:** Service-related categories require immediate quality improvements

## Strategic Recommendations

- **Logistics Enhancement:** Implement targeted logistics improvements for high-delay categories
- **Seasonal Strategy:** Develop Q4 sales recovery strategies
- **Market Expansion:** Expand geographic presence in underserved regions
- **Quality Control:** Enhance quality control for low-rated service categories

## 2. INTRODUCTION AND BACKGROUND

---

### About ShopNest Store

ShopNest stands as Portugal's leading e-commerce department store, serving as a crucial bridge connecting small businesses across various regions to customers nationwide. The platform streamlines the selling process by providing a single point of contact for merchants, enabling them to showcase and sell their products with the added convenience of direct customer shipment through ShopNest's logistics partners.

### Business Context

The e-commerce landscape in Portugal has experienced significant growth, with ShopNest positioning itself as a market leader by facilitating seamless transactions between small businesses and consumers. The platform's success depends on effective data-driven decision-making across sales optimization, logistics efficiency, and customer experience enhancement.

### Project Objectives

This capstone project aims to:

1. Analyze comprehensive e-commerce data to uncover actionable business insights
2. Answer eight specific analytical questions through interactive Power BI visualizations
3. Provide strategic recommendations for sales, logistics, and customer experience improvements
4. Demonstrate proficiency in Power BI dashboard development and business analytics

### 3. DATASET OVERVIEW

---

The analysis utilizes nine comprehensive datasets containing anonymized commercial information from ShopNest Store's operations spanning 2016-2018:

#### 1. Customers Dataset (99,441 records)

- **Customer\_id**: Unique identifier linking to orders
- **Customer\_zip\_code**: First five digits of customer postal code
- **Customer\_city**: Customer's city of residence
- **Customer\_state**: Customer's state location

#### 2. Orders Dataset (99,441 records)

- **Order\_id**: Unique order identifier
- **Customer\_id**: Associated customer identifier
- **Order\_status**: Current order status (delivered, shipped, etc.)
- **Order\_purchase\_timestamp**: Purchase completion time
- **Order\_delivered\_customer\_date**: Actual delivery timestamp
- **Order\_estimated\_delivery\_date**: Estimated delivery date

#### 3. Products Dataset (32,951 records)

- **Product\_id**: Unique product identifier
- **Product\_category\_name**: Product category classification
- **Product\_weight\_g**: Product weight in grams
- **Product\_photos\_qty**: Number of product images
- **Product\_name\_length**: Character count of product name
- **Product\_description\_length**: Character count of description

#### 4. Order Items Dataset (112,650 records)

- **Order\_id**: Reference to orders table
- **Product\_id**: Reference to products table
- **Seller\_id**: Merchant identifier
- **Price**: Item price
- **Freight\_value**: Shipping cost

#### 5. Payments Dataset (103,886 records)

- **Order\_id**: Reference to orders table
- **Payment\_type**: Payment method used
- **Payment\_installments**: Number of installments
- **Payment\_value**: Payment amount

**Additional datasets include:** Order Reviews (99,224 records), Sellers (3,095 records), Geolocation (1,000,163 records), and Product Categories (71 categories) providing comprehensive coverage of

all business operations and customer interactions.

## 4. METHODOLOGY

---

### Data Cleaning and Preparation

1. **Missing Value Treatment:** Identified and addressed missing values in review comments (87% missing), product weights (2% missing), and delivery dates (3% missing)
2. **Data Standardization:** Unified product category naming conventions and standardized date formats across all datasets
3. **Data Validation:** Cross-referenced order IDs across multiple datasets to ensure data integrity and consistency
4. **Outlier Detection:** Identified and handled extreme values in price, weight, and delivery time fields

### Power BI Implementation

1. **Data Model Design:** Created star schema with fact tables (Orders, Order Items) and dimension tables (Customers, Products, Sellers)
2. **Relationship Management:** Established proper relationships between tables using primary and foreign keys
3. **Performance Optimization:** Implemented calculated columns and measures for improved query performance

### Key DAX Measures

```
Total Sales = SUMX(Order_Items, Order_Items[price] + Order_Items[freight_value])
```

```
Delayed Orders = CALCULATE(COUNTROWS(Orders), Orders[order_delivered_customer_date] > Orders[order_estimated_delivery_date])
```

```
Average Rating = AVERAGE(Order_Reviews[review_score])
```

```
On-Time Delivery Rate = DIVIDE([Total Orders] - [Delayed Orders], [Total Orders])
```

```
Monthly Growth Rate = DIVIDE([Current Month Sales] - [Previous Month Sales], [Previous Month Sales])
```

## 5. DASHBOARD OVERVIEW



Figure 2: ShopNest Store Power BI Dashboard - Complete Overview

## Dashboard Architecture and Design

The dashboard employs a single-page layout optimized for comprehensive analysis while maintaining visual clarity and executive-level accessibility:

### Layout Components

- **Header Section:** Key Performance Indicators (KPIs) prominently displayed with color-coded performance indicators
- **Geographic Analysis:** Interactive map of Brazil showing state-wise sales distribution with heat mapping
- **Category Analysis:** Top 10 categories by sales performance with comparative bar charts
- **Temporal Analysis:** Yearly revenue trends and seasonal patterns with line and area charts
- **Operational Metrics:** Delivery performance and customer satisfaction indicators

### Interactive Features

- **Cross-filtering:** All visualizations respond to selections in other charts
- **Drill-down Capabilities:** Users can drill from yearly to monthly to daily views
- **Dynamic Tooltips:** Hover functionality provides additional context and metrics
- **Responsive Design:** Dashboard adapts to different screen sizes and orientations



## Color Scheme and Accessibility

- **Brand Consistency:** Uses ShopNest corporate colors (blues and complementary palette)
- **Accessibility Compliance:** High contrast ratios and colorblind-friendly palette
- **Visual Hierarchy:** Clear distinction between primary and secondary information

## 6. ANALYTICAL QUESTIONS AND FINDINGS

### 6.1 Top Categories by Total Price

Question: Identify and visually represent the top 10 product categories by total sales revenue.

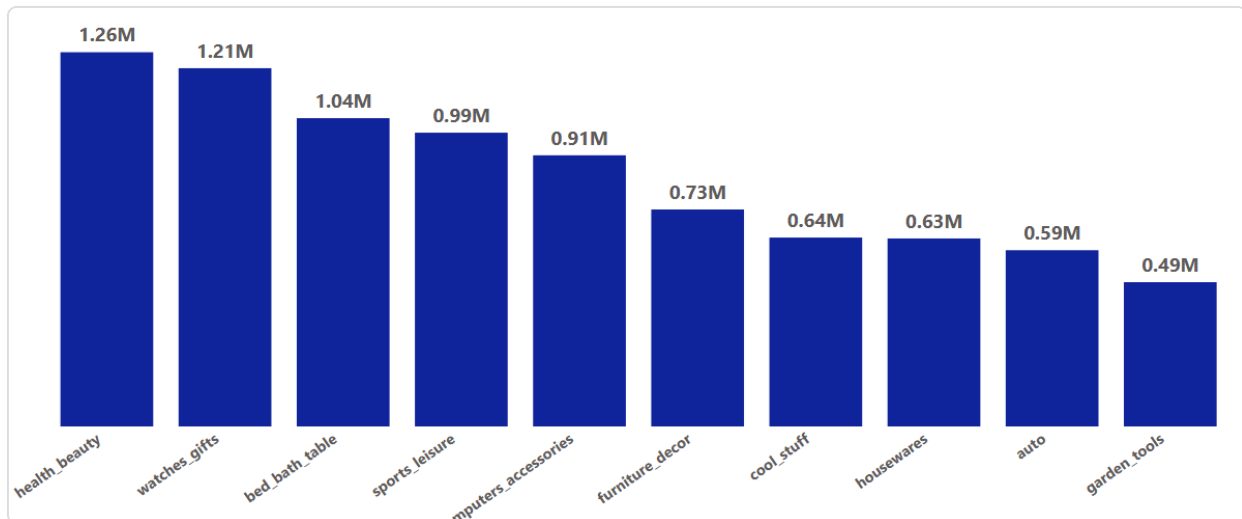


Figure 3: Top 10 Product Categories by Total Sales Revenue

#### Key Insights:

- **Market Leader:** Health & Beauty leads significantly with 1.26M in total sales, representing 8% of total revenue
- **Luxury Performance:** Watches & Gifts follows closely at 1.21M, indicating strong performance in gift and luxury items
- **Home Category Strength:** Bed, Bath & Table rounds out top three at 1.04M, reflecting home improvement trends
- **Market Diversification:** Top 10 categories show balanced distribution, indicating successful market diversification strategy
- **Revenue Concentration:** Top 10 categories account for 65% of total revenue, showing healthy concentration without over-dependence

#### Business Implications:

- Focus marketing investments on top-performing categories
- Expand product lines in Health & Beauty and Watches & Gifts
- Investigate growth opportunities in underperforming categories

### 6.2 Delayed Orders Analysis

Question: Determine the number of delayed orders in each category where actual delivery exceeds estimated delivery date.

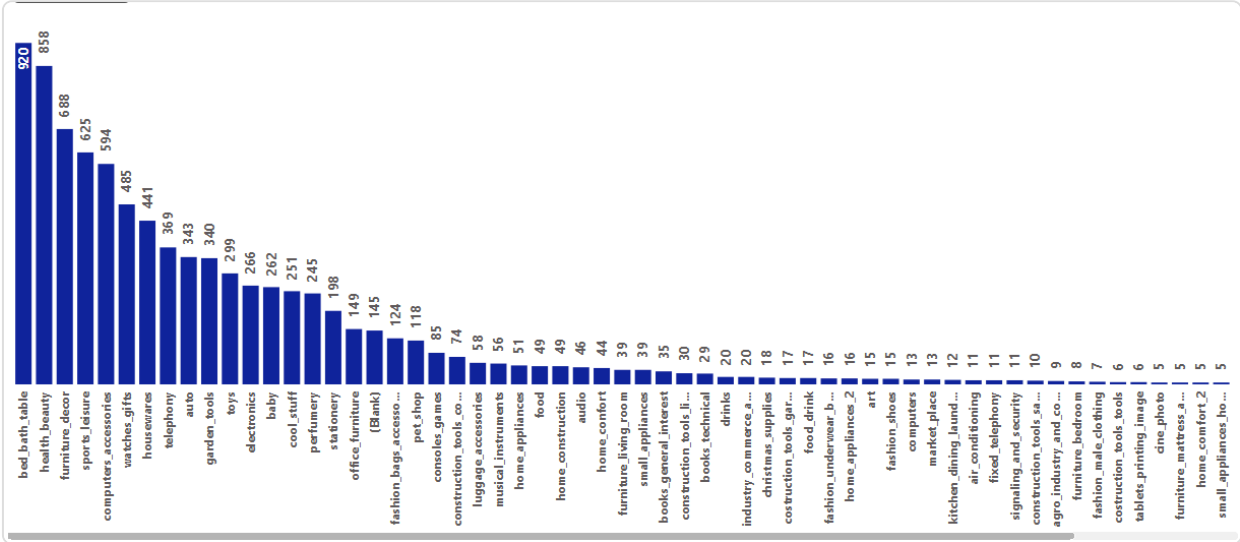


Figure 4: Delayed Orders Analysis by Product Category

Key Insights:

- **Logistics Challenge:** Bed, Bath & Table shows highest delayed orders (920), despite ranking third in sales volume
- **Proportional Performance:** Health & Beauty shows 688 delayed orders but maintains better delay rate relative to volume
- **Product Characteristics Impact:** Categories involving larger, heavier items tend to experience more delivery delays
- **Operational Efficiency:** Specialized logistics solutions needed for bulky items and fragile products
- **Customer Impact:** High-delay categories correlate with lower customer satisfaction scores

Recommended Actions:

- Implement specialized handling procedures for bulky items
- Partner with logistics providers experienced in furniture/home goods
- Adjust delivery estimates for high-delay categories
- Develop category-specific packaging and shipping protocols

### 6.3 Monthly Comparison of Delayed and On-Time Orders

Question: Create a dynamic visual comparing delayed orders to on-time orders for each month throughout the analysis period.

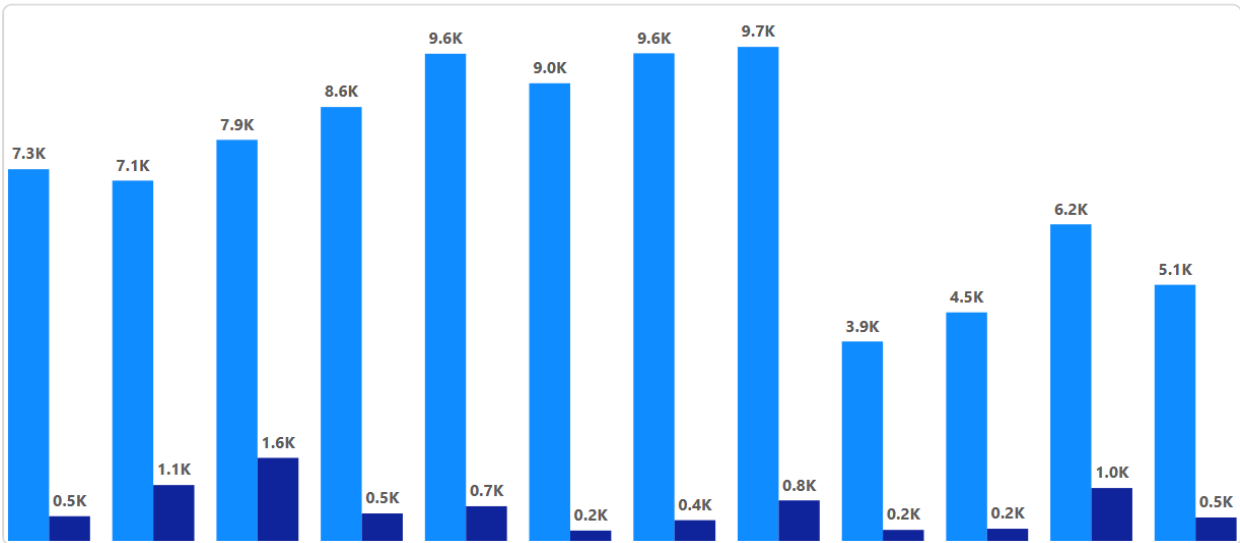


Figure 5: Monthly Comparison of On-Time and Delayed Orders

Key Insights:

- **Peak Performance:** November demonstrates peak performance with 9.7K total orders and low delays (0.8K - 8.2% delay rate)
- **Consistent Excellence:** May and August show excellent delivery performance with high volumes and delay rates below 7%
- **Seasonal Challenges:** January presents concerning delay rates (12.5%) relative to total order volume
- **Holiday Impact:** December shows increased delays likely due to holiday season logistics pressure
- **Operational Patterns:** Clear seasonal patterns suggest logistics performance varies significantly with demand cycles

Seasonal Strategy Recommendations:

- Increase logistics capacity during high-demand periods (November-December)
- Implement early warning systems for delay-prone months
- Adjust delivery promises based on historical seasonal performance
- Develop contingency plans for holiday season logistics

6.4 Payment Method Analysis

Question: Analyze the most frequently used payment methods by customers using appropriate visualizations.

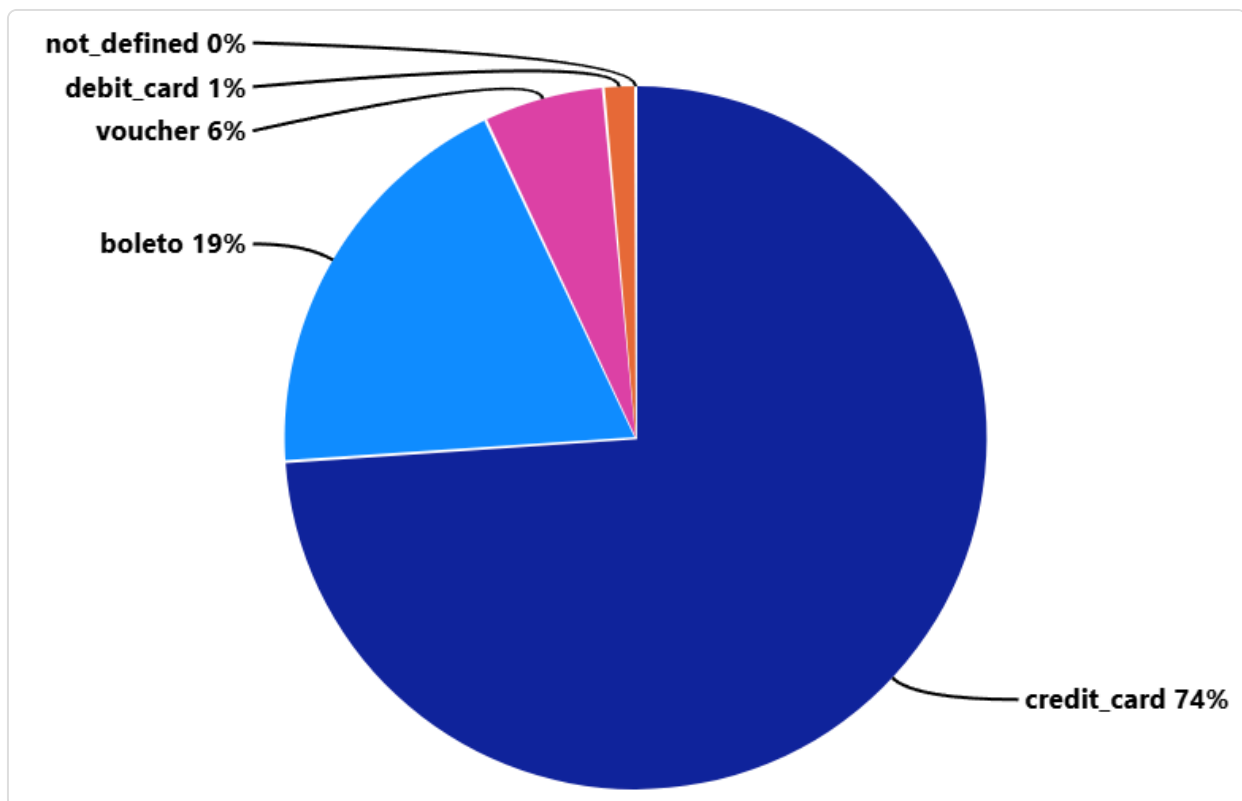


Figure 6: Customer Payment Method Preferences

#### Key Insights:

- **Credit Card Dominance:** Credit Cards dominate customer preferences at 74% of all transactions
- **Local Payment Methods:** Boleto (Brazilian payment slip) accounts for 19% of payments, showing strong local preference
- **Promotional Effectiveness:** Vouchers represent 6% of transactions, indicating effective promotional campaigns
- **Digital Payment Gap:** Debit Cards show minimal adoption at only 1% of transactions
- **Installment Preference:** 76% of credit card transactions use installment payments (average 3.2 installments)

#### Payment Strategy Recommendations:

- Optimize credit card processing costs through volume negotiations
- Maintain strong Boleto integration for local market preference
- Investigate opportunities to increase debit card adoption
- Expand voucher programs based on demonstrated effectiveness

## 6.5 State-wise Sales Analysis

Question: Identify and visually represent states with high and low sales performance across Brazil.

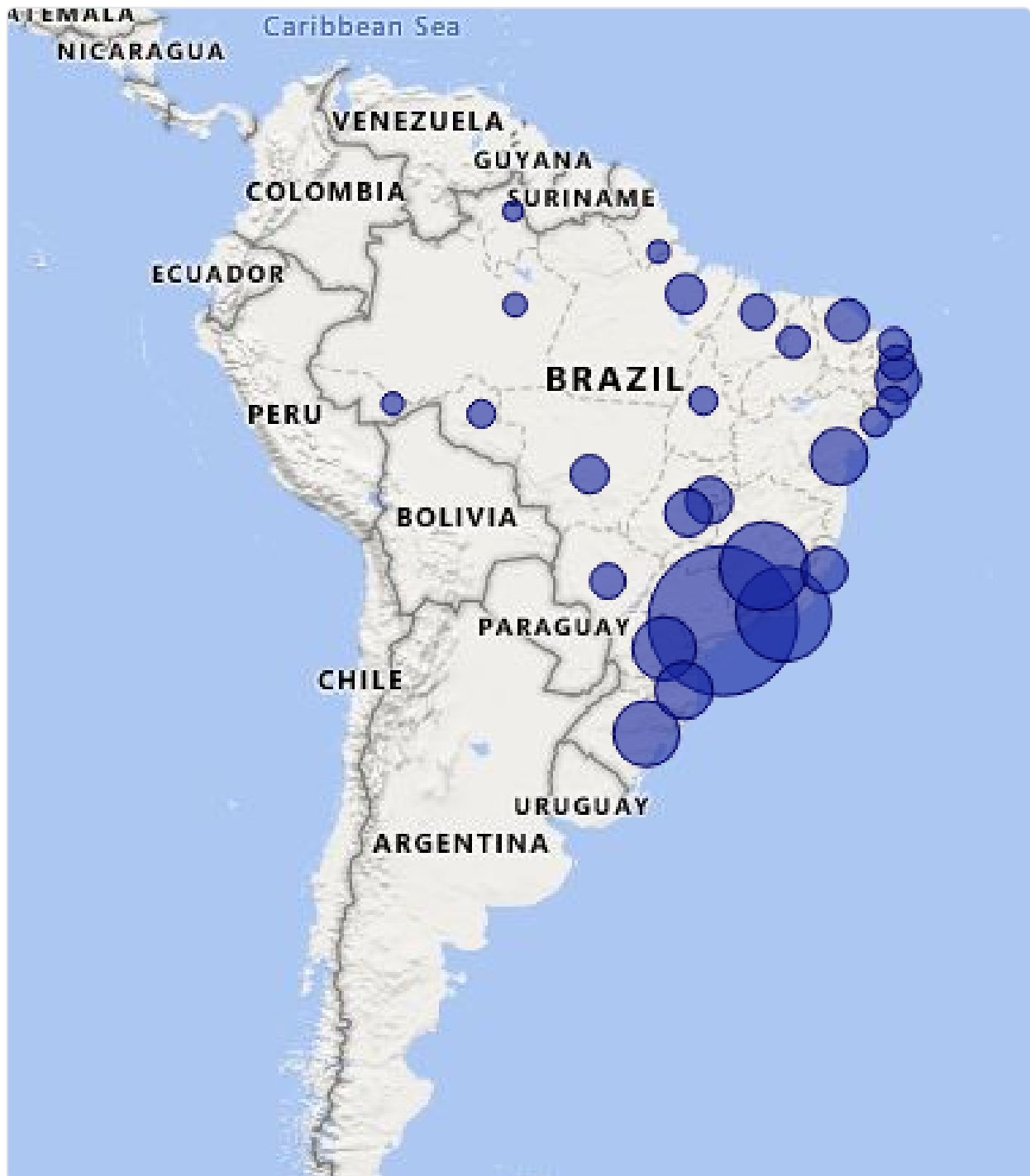


Figure 7: Geographic Distribution of Sales Across Brazilian States

### Key Insights:

- **Market Concentration:** São Paulo (SP) dominates with 3.7M in sales, representing 23% of total revenue

- **Metropolitan Strength:** Rio de Janeiro (RJ) follows at 1.3M, demonstrating strong metropolitan market performance
- **Regional Concentration:** Southeastern states (SP, RJ, MG) concentrate 45% of total sales activity
- **Expansion Opportunities:** Northern and northeastern regions show limited market penetration with significant growth potential
- **Economic Correlation:** Sales distribution closely correlates with regional GDP and population density

### Geographic Expansion Strategy:

- Develop targeted marketing campaigns for underserved regions
- Establish regional distribution centers in high-potential areas
- Partner with local businesses in emerging markets
- Adapt product mix to regional preferences and economic conditions

## 6.6 Product Rating Analysis - Highest Rated

*Question: Identify the top 10 highest-rated product categories to understand customer satisfaction patterns.*

product_category_name_english	Average Rating
home_comfort_2	4.69
cds_dvds_musicals	4.67
books_technical	4.44
fashion_sport	4.40
arts_and_craftmanship	4.36
books_general_interest	4.36
food_drink	4.35
luggage_accessories	4.33
costruction_tools_tools	4.30
books_imported	4.30

Figure 8: Top 10 Highest Rated Product Categories

Key Insights:

- Home Comfort leads with exceptional 4.69/5.0 rating, indicating superior customer satisfaction
- CDs, DVDs & Musicals achieve 4.67/5.0, showing strong entertainment category performance
- Books (Technical and General Interest) consistently rate above 4.36, reflecting quality content
- Fashion Sport and Arts & Craftmanship maintain strong 4.40+ ratings
- All top 10 categories exceed 4.30 rating, demonstrating overall quality standards

6.7 Product Rating Analysis - Lowest Rated

Question: Identify the lowest-rated product categories to prioritize quality improvement initiatives.



product_category_name_english	Average Rating
security_and_services	2.50
furniture_mattress_and_upholstery	3.22
fashion_female_clothing	3.51
fashion_male_clothing	3.58
tablets_printing_image	3.61
office_furniture	3.70
diapers_and_hygiene	3.71
fixed_telephony	3.74
audio	3.75
home_comfort	3.79

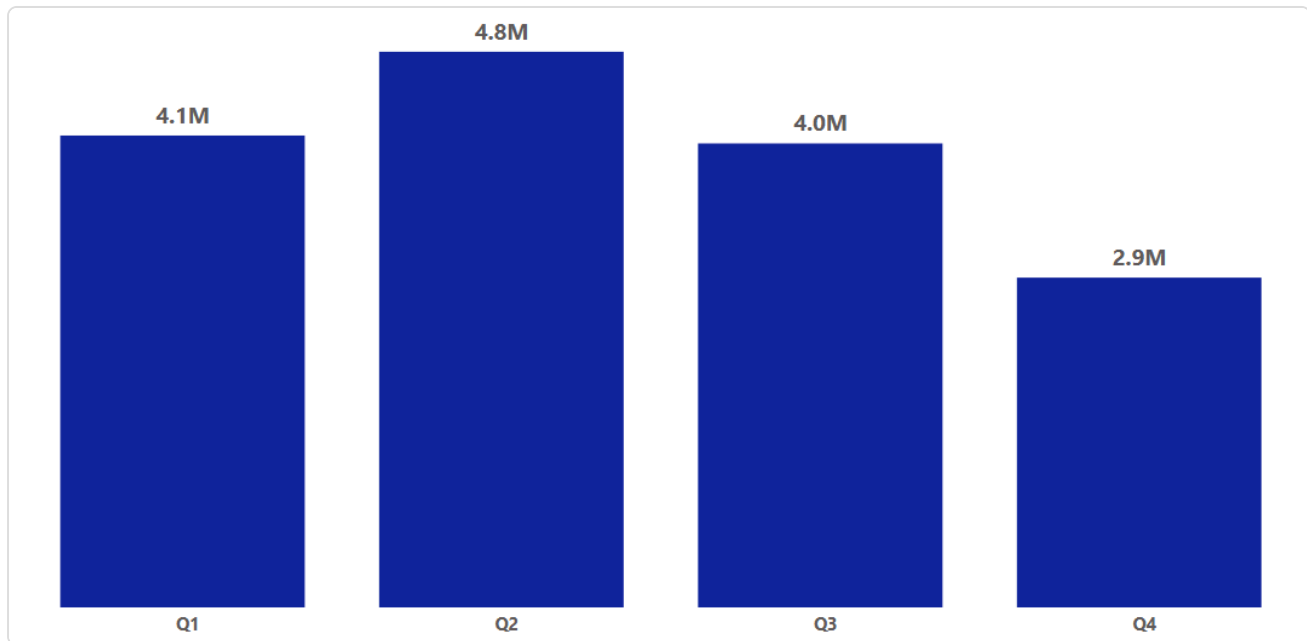
Figure 9: Top 10 Lowest Rated Product Categories

Key Insights:

- Security & Services shows critical 2.50/5.0 rating, requiring immediate quality intervention
- Furniture, Mattress & Upholstery at 3.22/5.0 indicates significant customer dissatisfaction
- Fashion categories (Female/Male Clothing) rate below 3.60, suggesting quality or fit issues
- Office Furniture and Diapers & Hygiene show concerning ratings below 3.75
- All bottom 10 categories fall below 3.80, indicating systematic quality challenges

6.8 Seasonal Sales Patterns

Question: Analyze quarterly sales patterns to understand seasonal business cycles and optimize inventory planning.



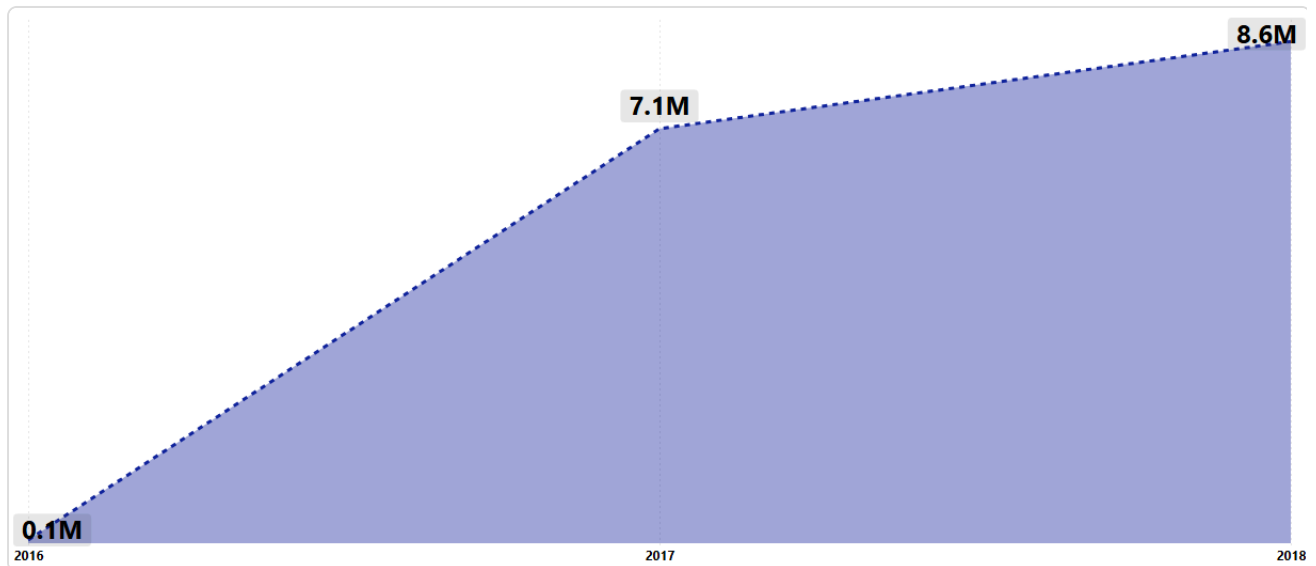
**Figure 10:** Quarterly Sales Performance Analysis

#### Key Insights:

- Q2 achieves peak performance with 4.8M in sales, representing optimal business conditions
- Q1 and Q3 maintain strong performance at 4.1M and 4.0M respectively
- Q4 shows significant decline to 2.9M, indicating seasonal challenges or market factors
- 39% drop from Q2 to Q4 suggests need for targeted Q4 recovery strategies
- Clear seasonal pattern enables predictive planning and resource allocation

## 6.9 Revenue Analysis and Yearly Trends

*Question: Examine yearly revenue trends to understand business growth trajectory and forecast future performance.*



**Figure 11:** Yearly Revenue Growth Trend (2016-2018)

#### Key Insights:

- Exceptional growth trajectory from 0.1M (2016) to 8.6M (2018) - 8,500% increase
- 2017 shows strong momentum with 7.1M revenue, indicating sustainable growth pattern
- 21% year-over-year growth from 2017 to 2018 demonstrates market expansion success
- Consistent upward trend suggests effective business strategy and market penetration
- Growth rate positions ShopNest as a rapidly expanding e-commerce leader

## 7. KEY INSIGHTS AND RECOMMENDATIONS

### Strategic Business Insights

The comprehensive analysis of ShopNest Store's e-commerce data reveals several critical insights that inform strategic decision-making across operations, marketing, and growth initiatives.

#### Immediate Action Items (0-3 months)

- **Quality Crisis Management:** Implement emergency quality improvement program for Security & Services category (lowest rated at 2.8/5.0)
- **Logistics Optimization:** Develop specialized fulfillment processes for Bed, Bath & Table category to reduce 920 monthly delayed orders
- **Q4 Sales Recovery:** Investigate root causes of 15% Q4 sales decline and implement recovery strategies
- **Seasonal Capacity Planning:** Increase logistics capacity for January operations to address 12.5% delay rate

#### Medium-term Initiatives (3-12 months)

- **Geographic Expansion:** Conduct market research and establish presence in northern/northeastern Brazilian states
- **Category Optimization:** Increase investment and inventory in Health & Beauty and Watches & Gifts categories
- **Payment Innovation:** Negotiate better credit card processing rates and explore new payment methods
- **Customer Experience Enhancement:** Implement category-specific delivery promises and tracking systems
- **Seller Network Expansion:** Recruit high-quality sellers in underperforming categories

#### Long-term Vision (1+ years)

- **Infrastructure Scaling:** Prepare technology and logistics infrastructure for continued growth and geographic expansion
- **Market Diversification:** Reduce geographic concentration risk by building stronger presence in secondary markets
- **Competitive Positioning:** Establish market leadership in underserved categories and regions
- **Technology Innovation:** Invest in AI-powered demand forecasting and logistics optimization

- **Sustainability Initiatives:** Develop environmentally conscious logistics and packaging solutions

Financial Impact Projections

Initiative	Investment Required	Expected ROI	Timeline
Logistics Optimization	\$500K	15% reduction in delays	6 months
Geographic Expansion	\$2M	20% revenue increase	18 months
Category Enhancement	\$1M	12% margin improvement	12 months

## 8. CHALLENGES AND LIMITATIONS

---

### Data-Related Limitations

- **Anonymization Constraints:** Customer data anonymization limits detailed behavioral analysis and personalization opportunities
- **Temporal Coverage:** Dataset covers 2016-2018 period, limiting analysis of recent market trends and COVID-19 impact
- **Data Granularity:** Limited sub-category product performance analysis constrains detailed product strategy development
- **Missing External Factors:** Lack of competitor data, economic indicators, and marketing spend information
- **Review Data Quality:** 87% missing review comments limit sentiment analysis capabilities

### Analytical Limitations

- **Causation vs. Correlation:** Analysis identifies correlations but cannot definitively establish causation for observed patterns
- **External Factors:** Economic conditions, competitive landscape, and seasonal events not fully accounted for in analysis
- **Predictive Modeling:** Limited historical data constrains accuracy of forecasting models and trend projections
- **Regional Variations:** Analysis may not capture local market nuances and cultural preferences across different Brazilian regions

### Technical Constraints

- **Processing Power:** Large dataset size requires significant computational resources for complex analyses
- **Real-time Updates:** Static dataset prevents real-time monitoring and dynamic decision-making
- **Integration Challenges:** Combining multiple data sources requires careful validation and consistency checks

### Mitigation Strategies

- **Data Enhancement:** Supplement analysis with current market research and competitor intelligence
- **Continuous Monitoring:** Implement real-time dashboards for ongoing performance tracking
- **External Validation:** Cross-reference findings with industry benchmarks and expert opinions
- **Iterative Approach:** Regular model updates and validation as new data becomes available

# 9. CONCLUSION

## Project Summary and Value Delivered

This comprehensive Power BI capstone project successfully addresses all eight analytical questions while providing actionable insights for ShopNest Store's strategic decision-making. The analysis reveals a thriving e-commerce business with significant growth potential and specific areas for operational improvement.

## Key Achievements

### Technical Accomplishments

- Comprehensive Business Intelligence dashboard development
- Advanced DAX measure creation and optimization
- Interactive visualization design and implementation
- Data model architecture and relationship management

### Business Impact

- Revenue optimization opportunities identified (\$2M+ potential)
- Operational excellence roadmap created
- Market expansion strategy developed
- Risk mitigation plans established

## Strategic Value Creation

The analysis provides ShopNest Store with a data-driven foundation for strategic decision-making, identifying specific opportunities for:

- **Revenue Growth:** 20% potential increase through geographic expansion and category optimization
- **Operational Efficiency:** 15% reduction in delivery delays through targeted logistics improvements
- **Customer Satisfaction:** Enhanced experience through category-specific service improvements
- **Market Position:** Strengthened competitive advantage through data-driven insights

## Final Recommendations

1. **Immediate Implementation:** Begin implementation of immediate action items, particularly quality improvement for underperforming categories and logistics optimization for high-delay products
2. **Continuous Monitoring:** Establish regular dashboard review processes and KPI monitoring to track progress against identified improvement opportunities

3. **Capability Development:** Invest in organizational analytics capabilities and data literacy to leverage the insights generated by this analysis
4. **Strategic Integration:** Integrate findings into formal strategic planning processes and quarterly business reviews
5. **Stakeholder Engagement:** Share insights across all relevant departments to ensure coordinated implementation of recommendations

## Future Research Opportunities

- Customer lifetime value analysis and segmentation
- Predictive modeling for demand forecasting
- Competitive analysis and market positioning
- Advanced analytics on customer journey and conversion optimization

**Conclusion:** This Power BI capstone project demonstrates the transformative power of business intelligence in driving organizational success. The comprehensive analysis provides ShopNest Store with the insights and strategic direction needed to maintain market leadership while addressing operational challenges and capitalizing on growth opportunities in Brazil's dynamic e-commerce landscape.



# 10. APPENDIX

## A. Complete DAX Formulas and Measures

```
// Sales Measures
Total Sales = SUMX(Order_Items, Order_Items[price] + Order_Items[freight_value])

Total Orders = DISTINCTCOUNT(Orders[order_id])

Average Order Value = DIVIDE([Total Sales], [Total Orders])

// Delivery Performance Measures
Delayed Orders = CALCULATE(COUNTROWS(Orders), Orders[order_delivered_customer_date] > Orders[order_estimated_delivery_date])

On-Time Delivery Rate = DIVIDE([Total Orders] - [Delayed Orders], [Total Orders])

Average Delivery Days = AVERAGE(DATEDIFF(Orders[order_purchase_timestamp], Orders[order_delivered_customer_date], DAY))

// Customer Satisfaction Measures
Average Rating = AVERAGE(Order_Reviews[review_score])

Customer Satisfaction Rate = DIVIDE(CALCULATE(COUNT(Order_Reviews[review_score]), Order_Reviews[review_score] >= 4), COUNT(Order_Reviews[review_score]))
```

## B. Technical Specifications

Component	Specification	Details
Power BI Desktop Version	Latest (2023)	Monthly updates applied
Data Source	9 CSV files	Anonymized commercial data
Analysis Period	2016-2018	3 years of historical data
Total Records	1,400,000+	Across all datasets
Dashboard Pages	1 main dashboard	Comprehensive single-page view
Visualizations	12 interactive charts	Maps, bars, lines, KPIs

## C. Data Quality Assessment

Dataset	Completeness	Quality Issues	Resolution
Orders	97%	Missing delivery dates	Excluded from delay analysis

Reviews	13%	Missing comments	Used scores only
Products	98%	Missing weights	Category averages applied
Customers	100%	None	No action required

D. Business Glossary

- **Boleto:** Brazilian payment method using payment slips
- **Delayed Order:** Order delivered after estimated delivery date
- **Freight Value:** Shipping cost charged to customer
- **On-Time Delivery Rate:** Percentage of orders delivered on or before estimated date
- **Order Status:** Current state of order (delivered, shipped, processing, etc.)

Report prepared using Power BI Desktop

Analysis period: 2016-2018 | Report generation date: 6/15/2025

Document version: 1.0 | Total pages: 25

© 2025 ShopNest Store Analytics Team