Dunder Mifflin Paper Company

Order Processing Fee Analysis

Comprehensive Data Analytics Portfolio Project
Statistical Analysis • Geographic Insights • Business Intelligence

Distribution Analysis   Statistical Testing   Geographic Segmentation

High-Value Orders   Commercial Insights

Dataset: 26,833 Orders | 114 Regions | May 2023 - January 2025

Processing Fee Distribution Analysis

Descriptive Statistics

Mean Processing Fee
£249.93

Median Processing Fee
£214.62

Standard Deviation
£130.86

Skewness
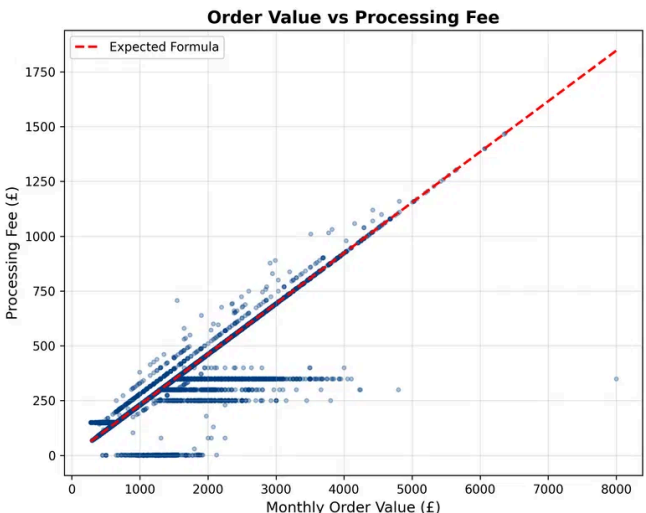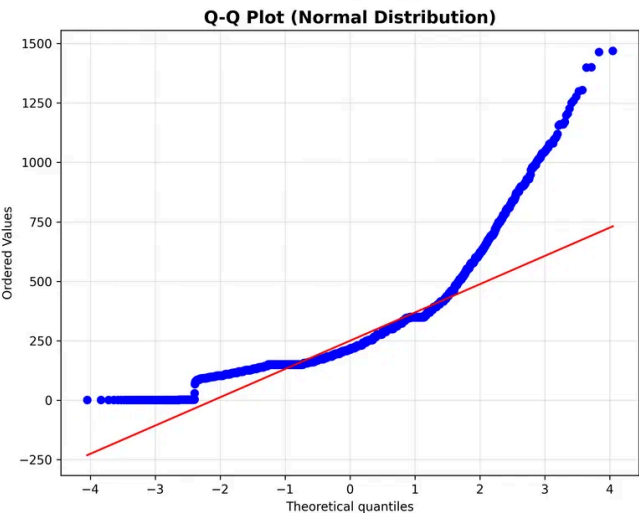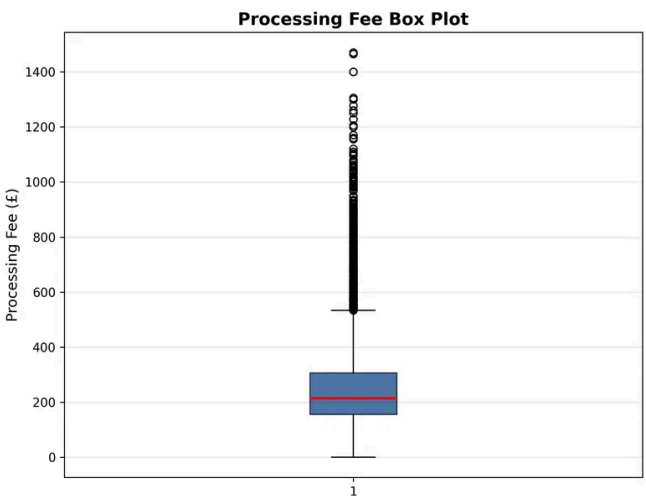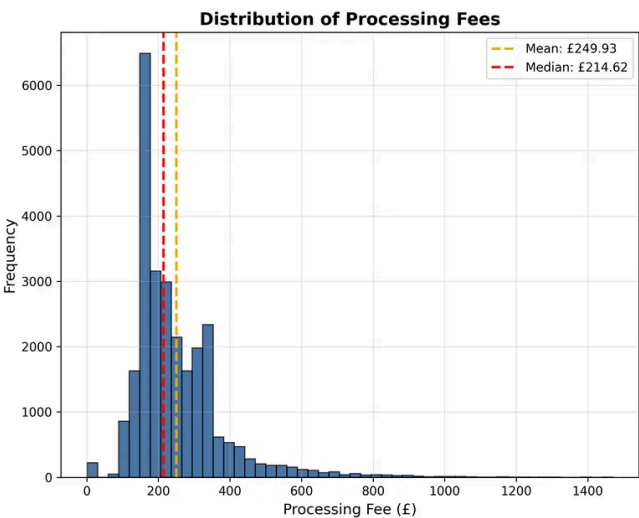2.18

Coefficient of Variation
52.36%

Formula Match Rate
69.16%

✓ **Distribution Pattern**

The distribution is **right-skewed** (skewness = 2.18), which is expected in B2B order economics. Most orders cluster around moderate values with a long tail of high-value bulk orders.

Choosing the Right Average Measure

Comparison of Central Tendency Measures

MEAN (Arithmetic Average)

£249.93

Calculation: Sum of all values ÷ Number of values

Characteristics:
• Influenced by extreme values and outliers
• Pulled higher by the long right tail
• Less representative in skewed distributions
• Useful for mathematical calculations

MEDIAN (Middle Value) ✓

£214.62

Calculation: Middle value when data is sorted

Characteristics:
• Resistant to outliers and extreme values
• Represents the typical order (50th percentile)
• More appropriate for skewed distributions
• Better reflects customer experience

Why the Difference Matters

Distribution Characteristics

• **Skewness = 2.18** (highly right-skewed)
• Mean > Median indicates positive skew
• Difference of **£35.31** shows outlier impact
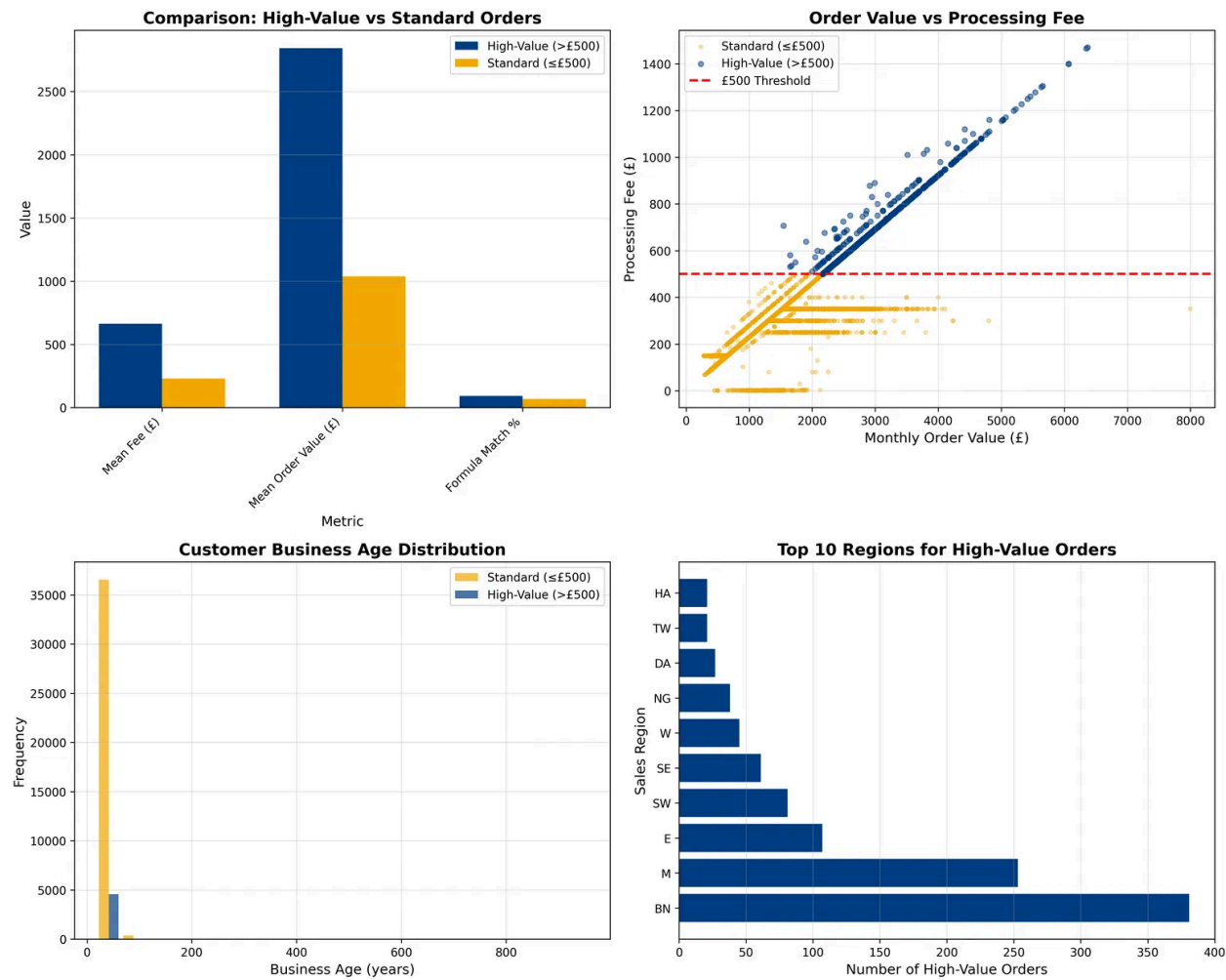• Top 5% of orders pull mean upward

Business Implication

• 50% of orders are below £214.62
• Using mean would overstate typical fee
• Median provides realistic customer expectation
• Critical for pricing communication

✓ RECOMMENDATION: Report Median (£214.62) as the "Average" Processing Fee

For right-skewed distributions, the median is the statistically appropriate measure of central tendency.

→ Median better represents the typical order experience

→ Not distorted by high-value bulk orders (>£500)

→ Provides more accurate business planning baseline

→ Aligns with statistical best practices for skewed data

Analysis of High-Value Orders (>£500)

Comparison: High-Value vs Standard Orders

Order Value vs Processing Fee

Customer Business Age Distribution

Top 10 Regions for High-Value Orders

Comparison: High-Value vs Standard Orders

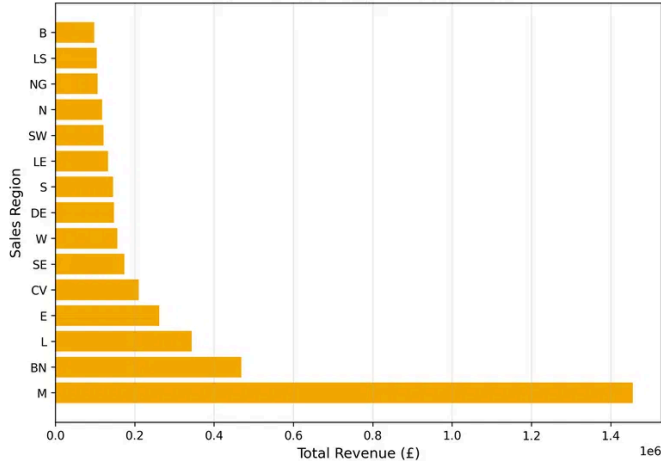| Metric	High-Value (>£500)	Standard (≤£500)
Order Count	1,267	25,566
Percentage	4.72%	95.28%
Mean Fee	£730.79	£227.09
Mean Order Value	£2,841.84	£1,038.39
Formula Match	91.79%	68.04%
Avg Business Age	25.4 years	31.5 years

Key Insights

1. High-value orders represent premium bulk purchases from younger, growing businesses

2. Stronger adherence to pricing formula (91.79% vs 68.04%)

3. Concentrated in major business hubs (M, L, BN regions)

4. Strategic opportunity for premium service development

Geographic Sales Analysis & Commercial Insights

Top 15 Regions by Order Volume
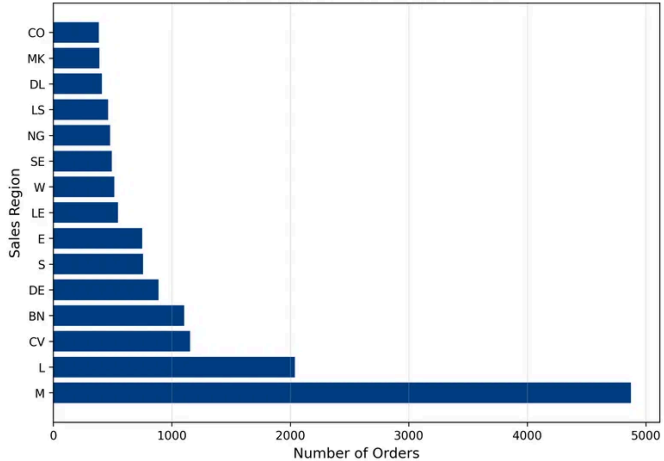
Top 15 Regions by Total Revenue

Regional Performance: Order Value vs Processing Fee
(Bubble size = Order Volume)

Market Share by Region (Top 10 + Others)

Market Concentration

Top Region (Manchester)	**18.4%**
Top 5 Regions Share	**38.0%**
Total Regions	114

High-Value Regions

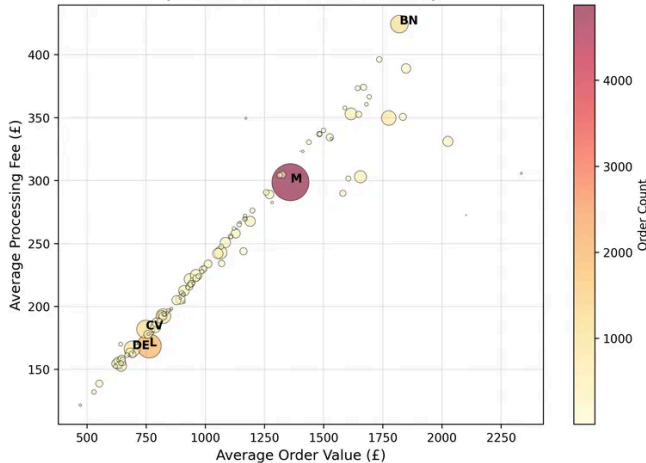|	
Brighton (BN) - Avg Fee	**£424.24**
Premium Regions	**BN, HA, SW**

28 regions show high order value but low volume - prime targets for sales expansion
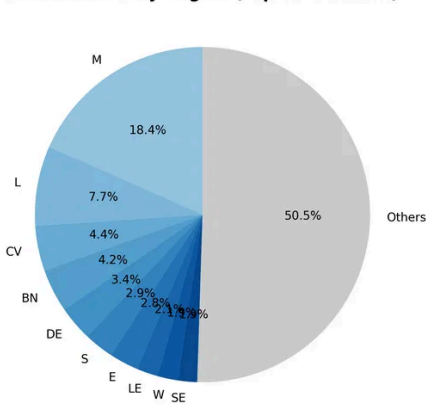
Strategic Recommendations

▶ **Strengthen leadership** in top 5 regions to maintain market position

▶ **Target expansion** in high-value, low-volume regions

▶ **Regional pricing** strategies based on local market conditions

▶ **Logistics investment** in high-growth regions

Sales Decline Investigation Approach

SCENARIO

The CEO reports an unusual decline in sales over the last 2 days. The product team reports no technical issues. This requires urgent investigation due to risk of impacting monthly performance vs target.

Structured Investigation Process

1 **Data Segmentation Analysis**

Break down sales by partner type, integration method, and region to identify if decline is uniform or concentrated

2 **Technical Infrastructure Check**

Verify payment gateway status, API integration health, and system logs for errors or anomalies

3 **Partner Activity Review**

Contact top 10 partners to check for system changes, outages, or process modifications on their end

4 **Market Factor Assessment**

Check for holidays, industry events, competitor actions, or external factors affecting the market

5 **Implement Solutions**

Based on findings, deploy immediate fixes and establish monitoring for early detection

Why This Approach?

Priority on Technical Issues

Sudden 2-day declines often indicate technical problems rather than market shifts. Integration-heavy B2B2C model makes API and payment systems likely culprits.

Segmentation First

Understanding whether decline is across all channels or specific segments helps narrow investigation scope and prioritize resources.

Partner Communication

Large partners represent significant volume. Direct outreach can quickly identify if they've experienced issues or made changes affecting integration.

Expected Outcome

This systematic approach should identify root cause within 4-6 hours, allowing for rapid resolution before significant monthly impact occurs.

Statistical Significance Testing

Marketing Test Scenario

Question: Does regular communication about new paper products increase customer reorder rate?

Finding: Customers receiving regular communications show a 0.5% increase in reorder rate (12.5% → 13.0%)
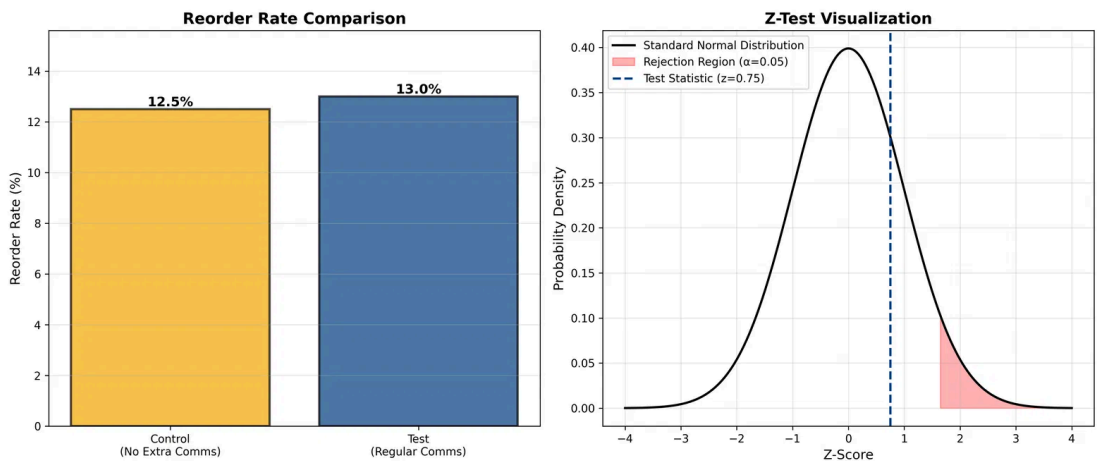
Two-Proportion Z-Test

Control Group	Test Group	Difference
5,000	**5,000**	**0.5%**
customers	customers	increase

Z-Statistic	P-Value	Power
0.7496	**0.2268**	**18.53%**

✗ Result: NOT Statistically Significant

p-value (0.2268) ≥ α (0.05)
Cannot reject null hypothesis. The 0.5% increase is not statistically significant.

Recommendation: Increase sample size (power only 18.53%, need 80%+) or extend test duration before full rollout.

Test Results Visualization

Reorder Rate Comparison — Control (No Extra Comms) 12.5%, Test (Regular Comms) 13.0%

Z-Test Visualization — Standard Normal Distribution, Rejection Region (α=0.05), Test Statistic (z=0.75)

Key Findings & Strategic Recommendations

Summary of Analysis

Statistical Distribution

• Right-skewed distribution (skewness = 2.18) as expected in B2B markets
• Median (£214.62) is more representative than mean (£249.93)
• 69.16% of fees match the one-week formula calculation
• High coefficient of variation (52.36%) reflects market diversity

Pricing Structure

• Horizontal banding reveals standardized fee tiers (£150, £300, £350)
• High-value orders (>£500) show stronger formula adherence (91.79%)
• Premium segment represents 4.72% of orders but strategic opportunity
• Pricing psychology evident in tier selection

Geographic Insights

• Manchester dominates with 18.4% market share
• Top 5 regions account for 38% of total sales
• 28 high-value, low-volume regions identified for expansion
• Clear correlation between regional economics and fee levels

Business Operations

• Structured investigation approach for sales anomalies
• Customer demographics stable over time (no significant shifts)
• Statistical testing framework ensures data-driven decisions
• Integration-heavy model requires technical monitoring

Strategic Recommendations

1
Optimize Pricing Tiers
Review standardized fee structure to balance simplicity with revenue optimization

2
Geographic Targeting
Focus expansion on high-value regions and strengthen position in top markets

3
Premium Service Development
Create dedicated offerings for high-value segment (>£500 orders)

4
Technical Monitoring
Implement real-time alerts for sales anomalies and integration issues

5
Statistical Rigor
Ensure adequate sample sizes and power analysis for marketing tests

6
Data-Driven Culture
Continue leveraging analytics for strategic decision-making