E-Commerce Pricing Optimisation Challenge: A Data Science Case Study

Background and Business Context

You are joining the Data Science & AI team at Temple & Webster, a rapidly growing e-commerce marketplace that sells a wide range of products across multiple categories. Temple & Webster has experienced significant growth over the past two years, expanding its product catalog from 20,000 to over 120,000 SKUs. However, the company faces increasing competition and needs to optimize its pricing strategy to maximize revenue while maintaining market competitiveness.

Currently, Temple & Webster uses a relatively simple pricing approach:

- Cost-plus pricing for most products
- Occasional manual adjustments based on competitor pricing
- Seasonal discounts applied uniformly across product categories
- Weekly promotions decided by merchandising teams with limited data support

This approach has several limitations:

- Prices aren't optimised for elasticity across different product categories
- Limited ability to respond to market changes in real-time
- No systematic way to test pricing strategies
- Inefficient inventory management due to suboptimal pricing decisions

The company's CEO has made pricing optimization a strategic priority and believes that implementing a data-driven, dynamic pricing system could increase revenues by 15-20% within the first year of implementation.

Your Challenge

As a senior data scientist candidate, your task is to develop a pricing optimization model that will help Temple & Webster maximize revenue across its product catalog. You will use historical sales data to understand price elasticity and develop a systematic approach to pricing that accounts for various factors influencing consumer purchasing decisions.

Dataset Description

For this case study, you will work with a simulated e-commerce dataset that mimics Temple & Webster's real-world data patterns. The dataset contains two years of transaction data for a sample of 500 products across 5 main categories.

Data Files

1. **products.csv**: Contains information about each product
    - product_id: Unique identifier for each product
    - category_id: Category identifier (1-5)
    - subcategory_id: Subcategory identifier
    - brand_id: Brand identifier
    - supplier_id: Supplier identifier
    - base_cost: Base cost of the product
    - quality_score: Product quality rating (1-10)
    - avg_competitor_price: Average price offered by competitors
    - is_seasonal: Flag indicating if product has seasonal demand (0/1)
    - launch_date: Date when product was added to catalog
2. **transactions.csv**: Contains historical sales data
    - transaction_id: Unique identifier for each transaction
    - product_id: Product identifier
    - timestamp: Date and time of purchase
    - price: Price at which product was sold
    - quantity: Number of units sold
    - user_id: Customer identifier
    - is_promotion: Flag indicating if product was on promotion (0/1)
    - promotion_type: Type of promotion if applicable
    - platform: ecommerce sales platform (web, mobile app, etc.)

3. **inventory.csv**: Contains historical inventory data
    - product_id: Product identifier
    - date: Date of inventory record
    - stock_level: Available stock
    - days_in_stock: Number of days product has been in stock
    - restock_date: Date when product was last restocked
    - restock_quantity: Quantity added during last restock

This dataset structure is inspired by real-world e-commerce data assets and allows for comprehensive pricing optimization analysis similar to those performed at major retailers.

Problem Statement

Temple & Webster needs to implement a data-driven pricing strategy that maximizes revenue while considering various business constraints. Your task is to develop a pricing optimization model that can:

1. Analyze historical data to understand price elasticity across different product categories and customer segments
2. Identify optimal price points for a selection of products to maximize revenue
3. Create a framework for dynamic pricing that adapts to changing market conditions
4. Design a system that can scale to handle the entire product catalog
5. Develop a simulation to test how your pricing recommendations would perform under different scenarios

Specific Requirements

1. **Price Elasticity Analysis**: Calculate price elasticity for different product categories and identify which products are most/least sensitive to price changes.
2. **Pricing Optimization Model**: Develop a model that recommends optimal price points for products to maximize revenue. The model should account for:
    - Product characteristics (category, quality, etc.)
    - Seasonality and temporal patterns

    - Competitor pricing
    - Inventory levels and constraints
    - Customer behavior patterns
3. **Implementation Strategy**: Provide a practical implementation plan for deploying your pricing optimization system in a production environment.
4. **Performance Evaluation**: Define appropriate metrics to evaluate the performance of your pricing model and show performance where relevant from backtesting. Additionally, demonstrate how you would measure its success in a real-world scenario.

Evaluation Criteria

Your solution will be evaluated based on:

1. **Technical approach**: The appropriateness and sophistication of analytical techniques, algorithms, and models used.
2. **Business understanding**: Demonstrating awareness of e-commerce business dynamics and how pricing influences customer behavior.
3. **Code quality**: Clean, well-documented, and efficient implementation that would be sustainable in a production environment.
4. **Innovation**: Creative approaches to solving the pricing optimization problem.
5. **Communication**: Clarity in explaining complex technical concepts and recommendations to both technical and non-technical stakeholders.
6. **Impact assessment**: Ability to quantify the potential business impact of your pricing recommendations.

Deliverables

Please provide the following deliverables:

1. **Code**: All code used for data exploration, model development, and evaluation. The code should be well-documented and reproducible.
2. **Technical report**: A comprehensive document (5-8 pages) explaining your approach, methodology, findings, and recommendations.

3. **Executive summary**: A 1-page summary of key findings and business recommendations for non-technical stakeholders.
4. **Presentation**: A 25-minute presentation (10-15 slides) explaining your approach and findings, followed by a 20-minute Q&A session.

You may use any tools or programming languages you prefer, but Python is recommended. Feel free to leverage common machine learning libraries and frameworks (e.g., scikit-learn, TensorFlow, PyTorch, etc).

Timeline

- Submission deadline: 5-7 days after completion
- Presentation date: 1-5 days after submission
- Share your complete case study to Marissa - Talent Acquisition Specialist via email marissa.ioannou@templeandwebster.com.au
- We recommend you to view our investor centre for insights on our business model & imagery for your presentation

Guidance on Approach

While we want to see your unique approach to this problem, here are some suggested steps:

1. **Exploratory data analysis**: Understand patterns in the dataset, identify relationships between variables, and generate insights that inform your modeling approach.
2. **Feature engineering**: Create relevant features that capture important aspects of product characteristics, market conditions, and customer behavior.
3. **Model development**: Develop models to predict demand at different price points and optimize for revenue. Consider various approaches such as:
4. **Simulation and validation**: Test your pricing recommendations through simulations and validate their potential impact on revenue.
5. **Implementation strategy**: Outline how your solution would be implemented in a production environment, including any API endpoints, data pipelines, or monitoring systems.