Dynamic Price Analysis of Aritzia E-commerce Products

1. Project Name and Team Members

Project Name: Dynamic Price Analysis of Aritzia E-commerce Products

Team Members:

Name	USC Email	USC ID
Mengxue Li	mengxue@usc.edu	6724826910

2. Short Description

This project analyzes temporal price changes in Aritzia's online product catalog to uncover discount patterns, cross-category price behaviors, and potential pricing cycles. Fashion e-commerce platforms frequently employ dynamic pricing mechanisms that remain unclear to consumers. Our goal is to provide data-driven insights into Aritzia's pricing behavior and empower consumers with better purchasing strategies.

Research Questions:

- Which product categories show the most frequent and significant discounts?
- How do product prices change over time?
- Are there identifiable patterns that can help consumers determine the best time to buy?

3. Data

Data Source:

Data was intended to be collected from the official Aritzia e-commerce website (https://www.aritzia.com) using Python's requests and BeautifulSoup libraries. However, due to website access restrictions, simulated data mimicking real Aritzia product data patterns was generated for this analysis.

Data Samples:

Metric	Value
Total Unique Products	236
Total Observations	2,360 (236 products × 10 days)
Collection Period	December 1-10, 2025
Categories Covered	5 (Outerwear, Dresses, Tops, Pants, Accessories)
Average Price	$148.63
Products on Sale	50.3%

4. Data Cleaning, Analysis & Visualization

Data Cleaning Process:

The raw data underwent several cleaning steps: (1) Removal of duplicate entries based on SKU, (2) Validation and standardization of price formats (rounding to 2 decimal places), (3) Recalculation of discount percentages for accuracy, (4) Addition of derived features including price tier (budget/mid-range/premium/luxury), discount tier (none/small/medium/large), and savings amount calculations.

Analysis Methods:

Three main analyses were conducted: (1) Category Discount Pattern Analysis examined which categories show the most frequent and significant discounts using aggregation and statistical measures. (2) Price Trend Analysis used time-series methods including linear regression to identify daily price change patterns. (3) Consumer Pattern Analysis identified actionable insights through correlation analysis and product-level tracking.

Key Findings:

Finding 1 - Category Discounts: Dresses had the highest discount frequency (61.1% of products on sale), while Tops showed the highest average discount rate (14.4%). Outerwear provided the largest absolute savings ($35.77 on average).

Finding 2 - Price Trends: A statistically significant decreasing trend in discount percentages was observed (slope: -0.145% per day, p-value: 0.001), suggesting discounts decreased over the collection period. Thursday showed the best average discounts.

Finding 3 - Consumer Patterns: 67 products maintained consistent discounts throughout the period. No strong correlation was found between original price and discount percentage (r=0.021), indicating discounts are not primarily driven by price point.

Visualizations Created:

Seven visualizations were generated: (1) Daily price trajectory line charts showing price and discount trends over time, (2) Category discount comparison bar charts, (3) Box plots of price and discount distributions by category, (4) Heatmap showing category-level discount patterns across days, (5) Scatter plot of original price vs. discount rate, (6) Correlation matrix heatmap, and (7) Discount tier distribution charts.

5. Changes from Original Proposal

Challenge Encountered: The original proposal planned to collect data directly from Aritzia's website using Python's requests and BeautifulSoup libraries through publicly available JSON endpoints. However, the website implemented access restrictions that prevented direct data collection.

Solution Implemented: To maintain the project's analytical rigor while working within constraints, realistic simulated data was generated that mimics actual e-commerce pricing patterns. The simulation includes: realistic price ranges for each category, dynamic price changes between days (15% daily change probability), appropriate discount distributions (10-50% discounts with weighted probabilities), and consistent product tracking across the 10-day collection period.

Impact on Analysis: While using simulated data, the analytical methods and visualization techniques remain exactly as proposed. The code infrastructure is designed to seamlessly handle real data when access becomes available.

6. Future Work

Given more time and resources, several improvements could be made:

1. Real Data Collection: Implement alternative data collection methods such as using Selenium for dynamic page rendering, or exploring official API partnerships with Aritzia.

2. Extended Collection Period: Collect data over 30-60 days to capture weekly and monthly patterns, including seasonal sales events (Black Friday, end-of-season sales).

3. Predictive Modeling: Develop machine learning models to predict future price changes and optimal purchase timing using historical patterns.

4. Competitive Analysis: Expand the analysis to compare Aritzia's pricing strategies with similar fashion retailers (Zara, H&M;, Reformation) to provide broader market context.

5. Consumer Alert System: Build a notification system that alerts consumers when products reach their lowest historical prices or when new discounts are applied.