NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

College of Computing and Data Science

Web Scraping for E-Commerce Retailers

Applying Web Scraping Techniques to Optimize Pricing and Channel Strategies

Completed By: Lim Yi   **Supervisor:** Assistant Professor Zhao Jun

Abstract:

This project demonstrates the effectiveness of web scraping to empower e-commerce retailers with actionable market insights. Using Python-based tools such as Selenium and Beautiful Soup, data was collected from five major marketplaces – Amazon, ASOS, JD Sports, Farfetch, and Footlocker. The data provides comprehensive insights into price distributions, competitor strategies, and consumer sentiment. Targeted at smaller emerging e-commerce retailers, the solution offers scalability, resource efficiency, and enhanced depth of analysis compared to existing tools, enabling optimal pricing and channel strategies to compete effectively in dynamic e-commerce environments.

Motivation:

Smaller e-commerce retailers lack access to timely and detailed market insights. Existing tools lack depth and scalability, limiting their utility. Web scraping provides the ability to efficiently collect and analyse large volumes of market data, empowering smaller retailers to thrive in competitive marketplaces.

Methodology:

- Identification of target e-commerce marketplaces
- Inspection of webpage structure
- Extraction of product data using Python tools
- Processing and cleaning of scraped data
- Visualization of insights using Tableau
- Analysis of marketplace dynamics

Results:

The project successfully scraped and analysed over 53,000 products from five major e-commerce marketplaces. The visualized insights revealed key trends, including price distributions, competitor pricing strategies, and consumer sentiment. These results provide smaller retailers with actionable data to optimize their pricing and channel strategies, enabling them to compete effectively in dynamic markets.
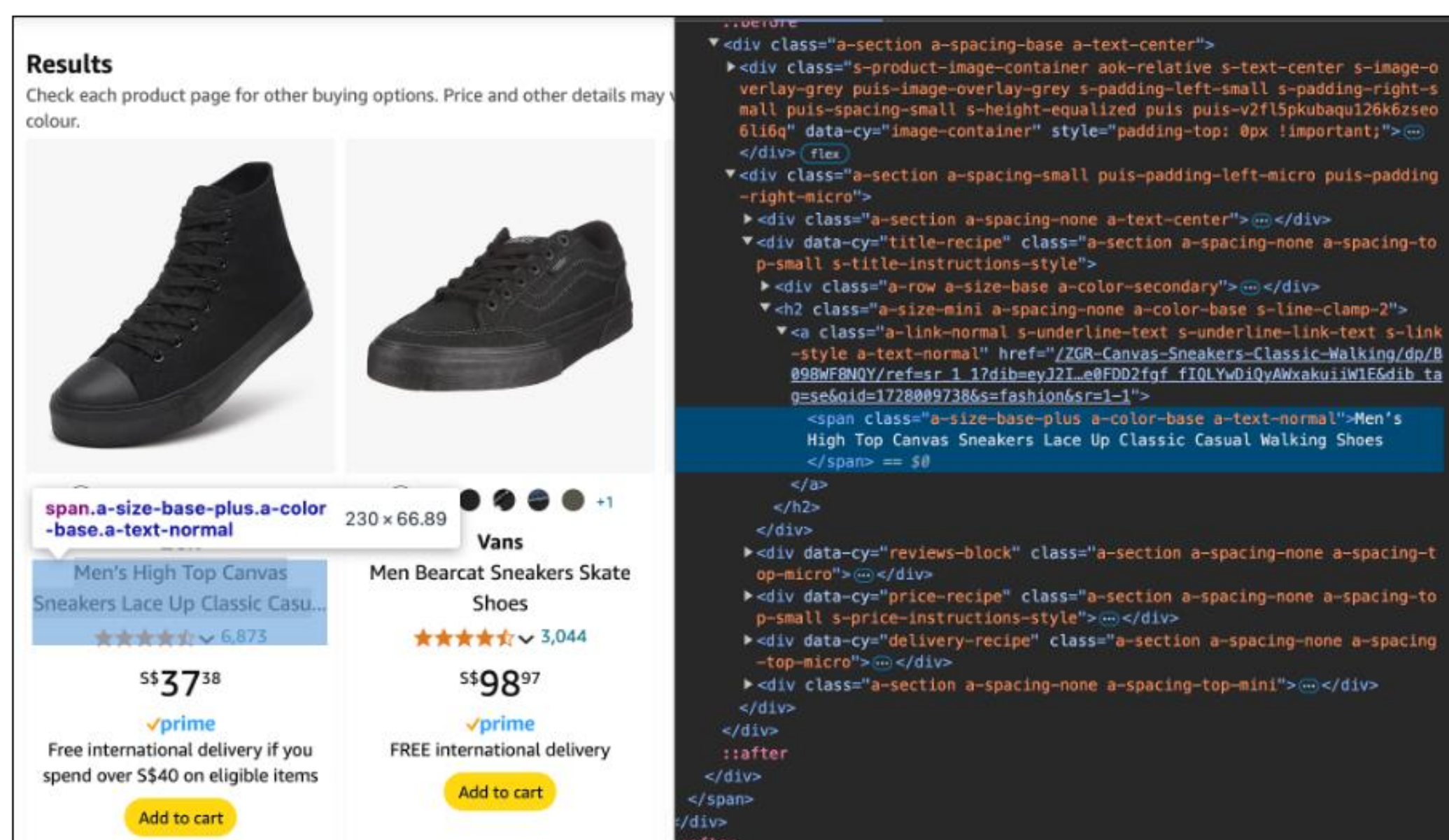
Analysis of webpage HTML code to develop automated scraping script
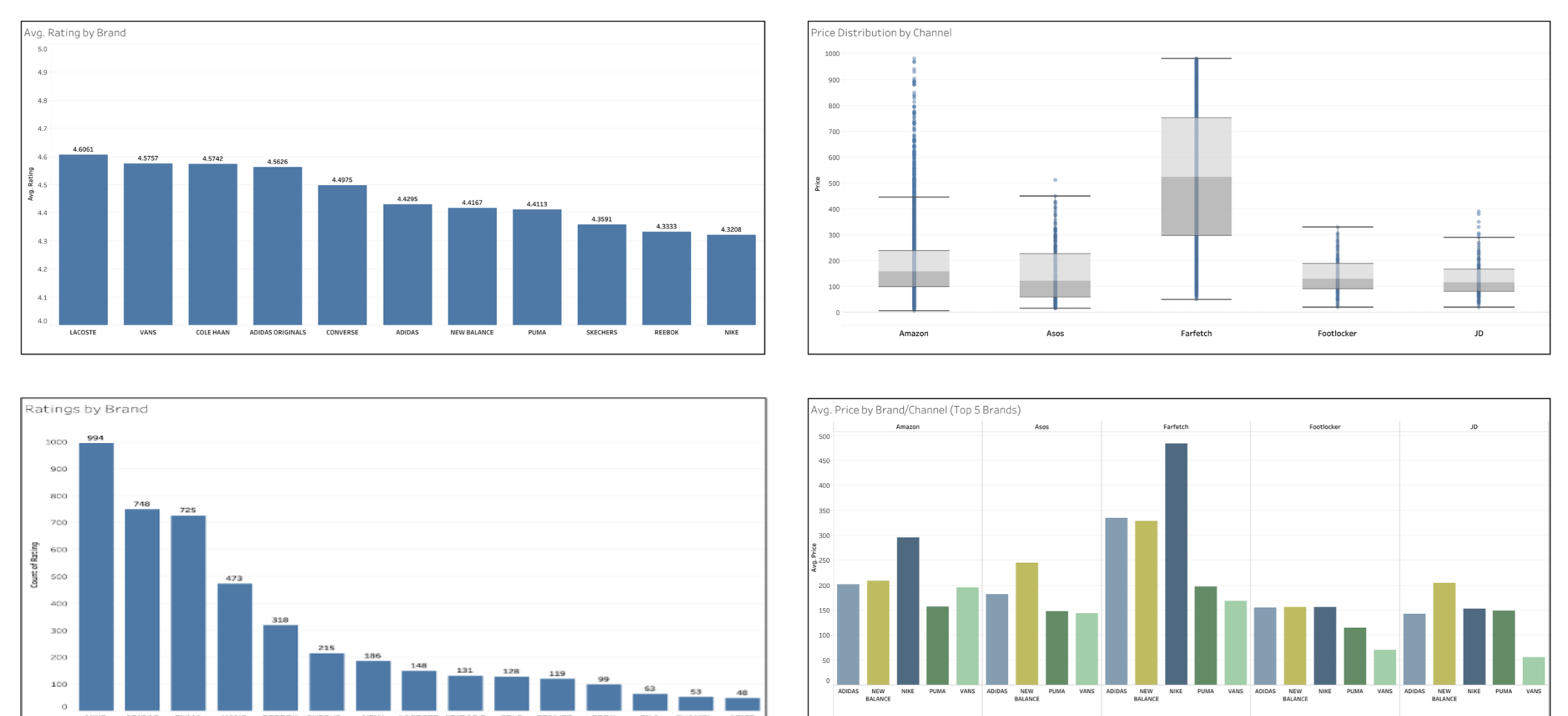
Visualization of scraped data to generate actionable insights