Building a Market Value Research Tool for Secondhand Women's Fashion Resellers

Student: Kristina Frazier
Technical Sponsor: Dawoon Jung
DATA-793 Section 3 Data Science Capstone

Presentation Outline

Problem Statement

Technical Workflow

Application Demo

Summary

Learnings

Questions & Discussion

Resellers Lack Efficient, Data-Driven Price Guidance

Support smarter pricing: Provide guidance for sellers of secondhand women's fashion through data-driven price suggestions

Automate "comps" research: Replace manual browsing with a scalable tool that models prices using thousands of listing attributes

Enable market exploration: Help users discover trends in item types, materials, brands, and condition-based price differences

---

SOLUTION: A user-friendly web app that provides smart price recommendations and flexible tools for exploring the resale market.

Final Application Built in 7 Technical Phases

1: Data Collection

Select features;
Create API call to collect data

2: Exploratory Data Analysis (EDA)

Explore relationships between features
price target variable

3: Feature Engineering

Normalize variables

4: Modeling

Train and compare performance of OLS,
PCR, PLS, Ridge & Lasso models

5: Web App Development

Draft: Price Prediction & Insights
Final: Daily Data Refresh

6: Evaluation & Deployment

Deployment & Sharing
Beta Testing
Collect and address feedback

7: Closeout

Public Repo
User Guide

Price Data was Cleaned and Modeled for Prediction

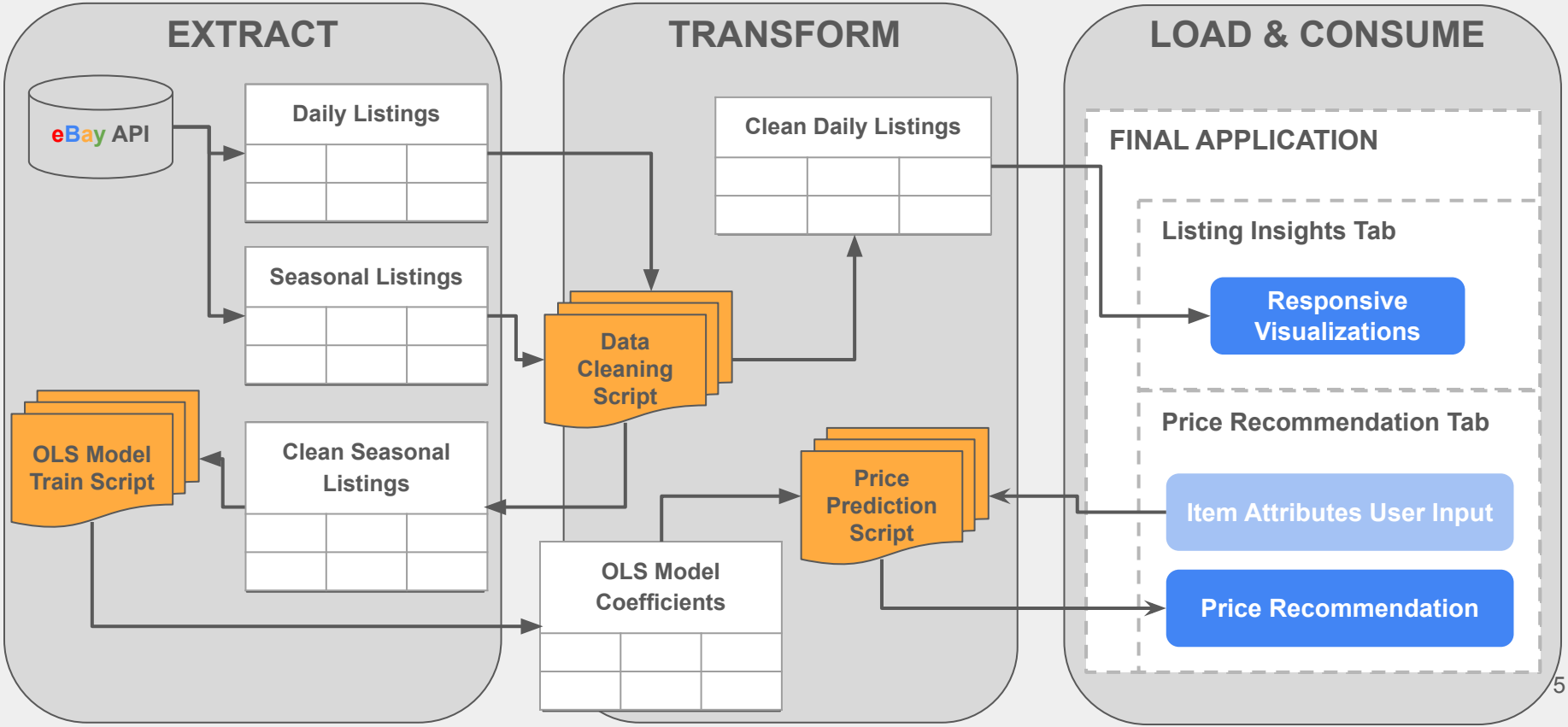
About the Data

- Collected daily in February & March 2025 (via eBay developer account)
- Clothing and shoes, not accessories
- 50,220 listings
- 9 categorical & 2 quantitative variables
- After normalizing data, 600+ dummy variables
- A “time-on-the-market” variable is not directly accessible; but seller-initiated price changes are captured

Data Cleaning and Modeling

- Data cleaning: write-in normalization; consolidating categories
- Log price for target variable
- Specified OLS & dimension reduction model types: Ridge, Lasso, PCR, and PLS models
- Performance measured with 30%/70% test train split comparing mean square error and comparing prediction interval variance
- OLS selected for interpretability and dynamic confidence intervals

Automated Data Pulls Drive Interactive Features



… So, Let's See the App!

App Deployed Publically Here: **https://reseller-market-research-tool.onrender.com/**

Use Case A: Item-Specific Price Guidance

"I have a new item I want to sell, and I know XXX about the item. How much I list the item for sale?"

Use Case B: Exploratory Research

"I'm considering selling {*a specific type of item OR items in XX material OR items in this price range, etc.* }. What does the current reselling market look like for items like this?

A Data-Driven Tool for Solo Reseller Entrepreneurs

Gathering data for comparable listings for competitive prices is time-consuming.

This “Market Value Research Tool” uses statistical modeling and descriptive stats to deliver smarter, faster listing data.

The tool enables both targeted item-specific price guidance and open-ended exploration of the reselling market as a whole.

This tool turns hours of manual market research into instant, data-driven pricing power — helping resellers price smarter, sell faster, and earn more

Many Opportunities for Technical Data Science Skills & Program Management

Content Learnings: APIs, virtual environments, model selection within a business context, reseller market data availability, competitive pricing motivations

Program Management Learnings: Time management; task prioritization; communicating progress; soliciting targeted feedback

Client Engagement Learnings: Anticipating user needs & motivation, user experience design

THANK YOU!

QUESTIONS?

A Special Thank You To:

Technical Sponsor, Prof. Dawoon Jung
Capstone Professor, Prof. Richard Ressler
DATA-793-003 classmates

Backup Slides

Price Data was Cleaned and Modeled for Prediction

Phases 3: Feature Engineering

- Cleaned and normalized listing data from eBay API & web scraping
- Addressed write-in variation using dictionaries, fuzzy matching, and RegEx
- Selected price-relevant predictors via correlation analysis and literature review

Phase 4: Model Development

- Specified OLS, Ridge, Lasso, PCR, and PLS models (log-price target)
- Chose OLS for interpretability and dynamic confidence intervals
- Final OLS model mitigates heteroscedasticity using robust covariance estimators