University of Pennsylvania
WUDAC - Datathon Spring 2025

Technical Report

Saaransh Pandey, Aakash Agarwal

1 Executive Summary

1.1 Objective

Objective We set out to predict which browsing sessions will lead to a repeat purchase on Amazon—and to test whether that same predictive framework holds up on other major e-commerce sites.

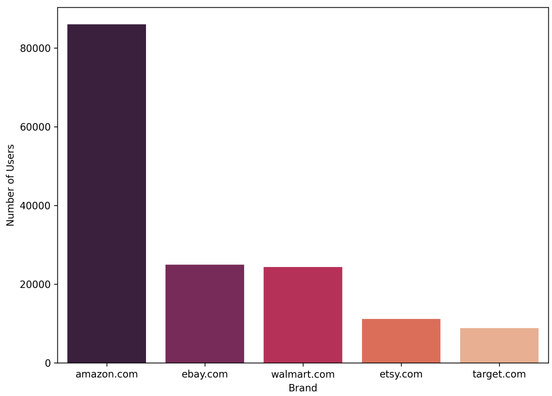

Figure 1: Distribution of users on different domains

1.2 Data & Approach

Using Comscore session (2020–2023) and transaction data, we computed Recency (days since last purchase), Frequency (number of sessions), and Monetary (basket spend) metrics at both the session and aggregated user level to succinctly capture purchase behavior. These RFM metrics are crucial because they distill complex browsing and purchase histories into clear indicators of customer value and intent, directly informing our segmentation and targeting strategies. We then applied k-means clustering to segment Amazon users into four distinct personas:

- Dormant (Cluster 0): Haven't bought in a long time; light site use; mid-sized orders.
- Occasional Buyers (Cluster 1): Recent but infrequent purchasers; brief visits; average spend.
- Frequent Browsers (Cluster 2): High visit frequency but skim pages; average basket size.
- Engaged Spenders (Cluster 3): Deep, lengthy sessions; large baskets; loyal recent buyers.

Building on those segments, we trained an XGBoost classifier to flag sessions likely to yield a repeat purchase.

1.3 Key Results

- Amazon Test Set:
  - Accuracy: 85%
  - Repeat-purchase (class 1) precision: 37%
  - Repeat-purchase recall: 59%
  - F_1 score: 0.45
- Cross-Domain Generalization (eBay, Walmart, Etsy, Target):
  - Accuracy: 78%
  - Repeat-purchase precision: 25%
  - Repeat-purchase recall: 60%
  - F_1 score: 0.35

Despite domain shifts, recall remains stable—demonstrating that our model reliably identifies high-value sessions across platforms.

1.4 ROI Analysis

Focusing on the top 10% of sessions by predicted repeat-purchase probability:

- Incremental repeat orders: $\simeq 1,486$
- Marketing cost: $0.50 per targeted session (assumption)
- ROI: 2.6x (net revenue ÷ cost)

1.5 Exploratory Insights

- Domain Concentration: Amazon accounts for $\simeq 85K$ of $\simeq 150K$ unique users and $\simeq 650K$ of $\simeq 930K$ total sessions—justifying our Amazon-centric modeling focus.
- Pages Viewed: Conversion rate climbs from $\simeq$ 7.5% (2–3 pages) to $\simeq$ 32% (51+ pages).
- Session Duration: Converted sessions skew longer—users who stay past 10 minutes convert at materially higher rates.
- Time of Day: Conversion peaks in late afternoon (2–6PM UTC), dipping in early morning hours.
- Referrer Type: "Marketplace" referrals (e.g., Amazon's own app or site) show the highest conversion ($\simeq$ 19%), versus $\simeq$ 14% for direct and $\simeq$ 8% for search or social.

1.6 Implications

- Targeting Strategy: Concentrate retargeting spend on the top 10% of sessions—especially those in the "Engaged Spenders" and "Frequent Browsers" clusters during peak hours (2–6 PM).
- Channel Focus: Prioritize marketplace referrals and personalize offers for users with deep browsing patterns (higher pages viewed).

1.7 Next Steps

- Enrich Features: Incorporate external signals (e.g., email-open rates, ad impressions).
- A/B Testing: Run causal experiments on personalized offers for high-probability sessions.
- Broader Roll-out: Extend the pipeline to additional domains (beyond the initial secondary domains) and refine cluster definitions over time.

2 Introduction

2.1 Background and Motivation

E-commerce has reshaped the retail landscape: in 2023, over 20% of global retail sales occurred online, generating vast streams of click-and-purchase data. Every browsing session—from page views and time on site to referral source and demographic context—offers a glimpse into consumer intent. Among the many metrics that drive sustainable growth, repeat purchases stand out; retaining an existing customer costs a fraction of acquiring a new one, and loyal shoppers contribute disproportionately to lifetime value.

2.2 Thesis Problem

In this work, we ask:

> *Can we accurately predict which browsing sessions on Amazon will lead to a repeat purchase, and does that same predictive model—without any re-tuning-generalize to other major e-commerce platforms?*

Predicting repeat purchases is valuable because returning customers drive a significant portion of revenue. Accurately flagging high-value sessions enables timely, personalized interventions that improve conversion and retention. If such models generalize across platforms, they offer scalable solutions for optimizing customer engagement and operational strategy.

Addressing this question requires (1) engineering session-level features that capture user engagement and purchase history, (2) segmenting customers into behaviorally meaningful clusters, and (3) building and evaluating a classifier that flags high-value sessions both on Amazon and on secondary domains (eBay, Walmart, Etsy, Target).

2.3 Data Overview

We leverage two proprietary Comscore datasets spanning 2020–2023:

- **sessions.csv** (∼13.8 M rows): session-level pageviews, duration, referral source, and household demographics (income, education, region).
- **transactions.csv** (∼2.3 M rows): item-level purchase records including product category, quantity, and basket spend.

These tables are joined on `machine_id` and `event_date`, yielding a unified view of browsing behavior and purchase outcomes.

2.4 Report Organization

The remainder of this report is structured as follows:

- **Section 3 – Thesis Problem:** formal statement of our predictive goal and evaluation criteria.
- **Section 4 – Data & Limitations:** detailed description of data sources, preprocessing steps, and known caveats.
- **Section 5 – Exploratory Data Analysis:** key patterns in RFM metrics, referral channels, and temporal trends.
- **Section 6 – Customer Segmentation:** k-means clustering methodology and cluster interpretations.
- **Section 7 – Predictive Modeling:** XGBoost implementation, performance on Amazon test data, and cross-domain generalization.
- **Section 8 – ROI Implications:** business impact assessment and targeting recommendations.
- **Section 9 – Conclusion & Future Work:** summary of findings and proposed next steps.

3 Thesis Problem

3.1 Predictive Objective

Our primary goal is to build a binary classifier that, for any given Amazon browsing session, predicts whether that session will result in a repeat purchase. Formally, let

$$y = \begin{cases} 1 & \text{if the session leads to a repeat purchase,} \\ 0 & \text{otherwise.} \end{cases}$$

We seek a model $f(x) \approx P(y = 1 \mid x)$, where x comprises session-level features (pages viewed, duration, referral type, etc) and user-history metrics (RFM).

3.2 Generalization Objective

Beyond Amazon, we evaluate whether f generalizes—without any re-tuning—to other major e-commerce domains (eBay, Walmart, Etsy, Target). This tests the robustness of our feature engineering and modeling pipeline under domain shift.

3.3 Evaluation Criteria

We assess performance using:

- **Area Under the ROC Curve (AUC):** measures rank ordering of positive vs. negative sessions.
- **Precision & Recall (class 1):** precision $= \frac{\text{TP}}{\text{TP+FP}}$, recall $= \frac{\text{TP}}{\text{TP+FN}}$.
- F_1 **Score:** harmonic mean of precision and recall.
- **Accuracy:** overall fraction of correct predictions.
- **Cross-Domain Drop:** difference in AUC between Amazon test set and secondary domains.

Since the dataset is imbalanced, accuracy alone is not a reliable metric. A successful solution achieves high AUC and balanced precision/recall on Amazon, with minimal degradation when applied to other platforms.

4 Data & Limitations

4.1 Preprocessing and Join Strategy

To combine browsing behavior with purchase outcomes, we joined `sessions.csv` and `transactions.csv` on:

$$\{\texttt{machine\_id}, \texttt{event\_date}\}$$

rather than on `site_session_id`, because:

- A large fraction of transaction rows had `site_session_id` = 0, preventing reliable session-level joins.
- Using `event_date` collapses all activity on the same day into a single "pseudo-session."

4.2 Limitations

- **Session granularity:** Merging on `event_date` may group multiple distinct sessions in a single day into one record. However, our analysis shows this occurs in very less number of cases, making the date-level join a pragmatic compromise.
- **Missing `site_session_id`:** Transactions with `site_session_id`=0 cannot be disambiguated; these are subsumed by the date-level join.
- **Demographic sparsity:** Some households have unknown or missing demographic fields (income, education). We treat "Unknown" as a separate category rather than impute.
- **Temporal alignment:** All sessions and transactions on the same UTC date are merged, potentially misaligning late-night browsing with early-morning purchases. Given the low incidence, we accept this minor temporal mismatch.

5 Exploratory Data Analysis

5.1 Demographic Distributions

We begin by examining household demographics for all users versus those who converted at least once (i.e., appear in `transactions.csv`).

- **Overall (Fig. 2):** Most users report having an associate degree, followed by some college education without a degree, and then a bachelor's degree—excluding the large proportion with unknown education levels (`hoh_most_education`). Household incomes are concentrated at the extremes, with many users earning below \$25K or above \$200K (`household_income`). Additionally, 2–3 person households are the most common, with the majority of users residing in the South (region 3) and identifying as Caucasian (race 1).
- **Successful (Fig. 3):** Users who converted show similar demographic trends, though they exhibit fewer unknown values in the education feature, suggesting more complete data among this group.

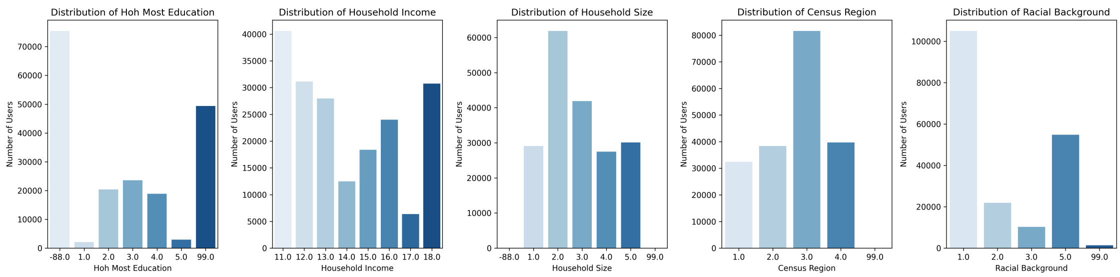

Figure 2: Overall User-Level Demographic Distribution

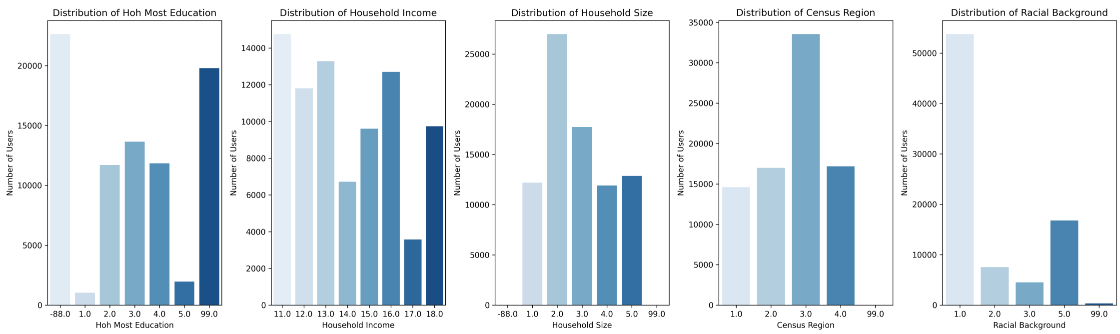

Figure 3: Demographics of Users with ≥ 1 Purchase

5.2 Session & Basket Characteristics

Key engagement and purchase-related features display notable skewness:

- **Basket Total (Fig. 4 (a)):** While the top 1% of sessions involve spending close to \$600, the vast majority of baskets are valued under \$50, indicating a heavy skew toward lower-value transactions.
- **Distinct Products (Fig. 4 (b)):** Most users tend to purchase only a few unique items per session, with over 80% of baskets containing between one and three distinct products.
- **Pages Viewed & Duration:** As illustrated in Figs. 5 (a) and 5 (b), sessions with higher page engagement and longer durations are more likely to lead to conversions, highlighting the importance of user involvement.

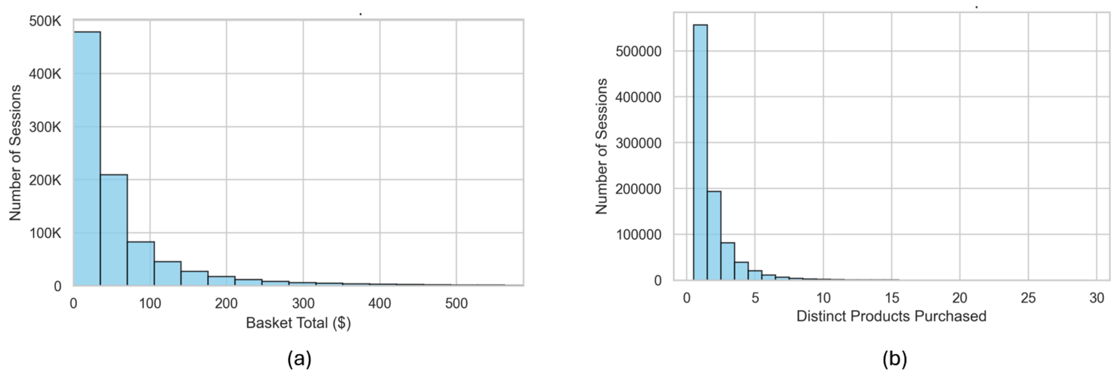

Figure 4: Distribution of (a) Basket total with the number of sessions (b) The number of distinct products in the basket

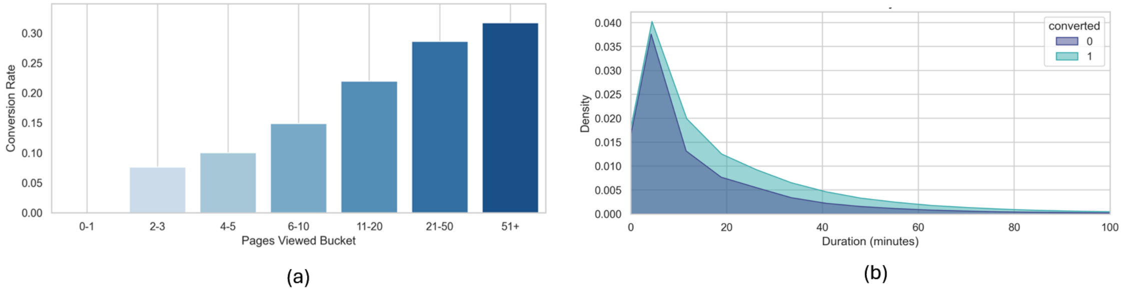

Figure 5: (a) Number of pages viewed vs Conversion Rate (b) Time Spent, in a session

5.3 Recency & Frequency Patterns

User history features provide insight into patterns of repeat engagement and purchasing behavior:

- **Days Since Last Session (Fig. 6 (a)):** The distribution peaks at zero, indicating a high frequency of same-day return visits, and shows additional spikes at yearly intervals, pointing to recurring annual shopping habits.
- **Days Since Last Purchase (Fig. 6 (b)):** A similar seasonal trend is observed, with the majority of purchases occurring within 30 days of the previous order, suggesting a tendency toward short-term repeat buying.
- **Historical Conversion Rate (Fig. 7):** Most users exhibit low prior conversion rates (under 10%), but a small segment demonstrates exceptionally high conversion rates (over 50%), signaling a valuable group of repeat buyers.

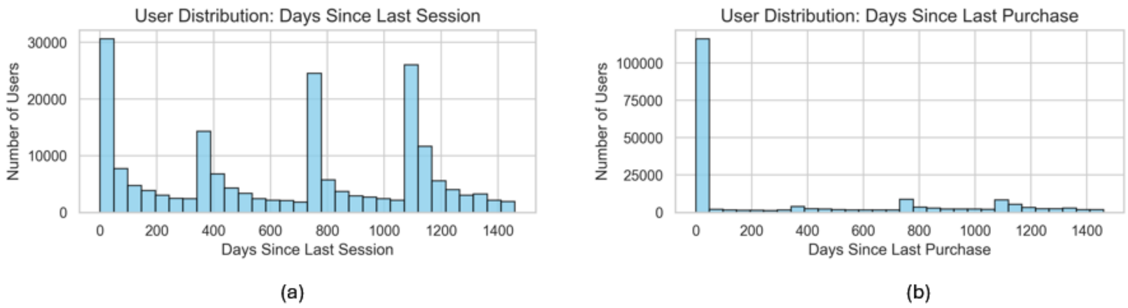

Figure 6: User Distribution: (a) Days Since Last Session (b) Days Since Last Purchase

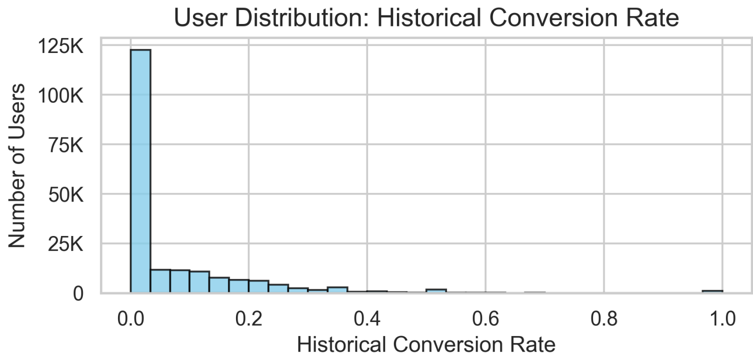

Figure 7: User Distribution: Historical Conversion Rate

5.4 Referral and Temporal Trends

Now we examine how different referral sources like, direct traffic, marketplace, search, social and temporal factors like time of day, day of the week influence conversion rates. By analyzing these dimensions, we aim to uncover actionable insights that can help optimize user acquisition strategies and improve overall marketing effectiveness.

- **Referrer Type (Fig. 8):** Marketplace and direct traffic yield the highest conversion rates, at approximately 19% and 14% respectively. In contrast, search, social, and other referral sources fall below the 10% mark.
- **Hour of Day (Fig. 9 (a)):** Conversion activity is highest during the afternoon hours (2–6 PM UTC), with a noticeable dip during the early morning.
- **Day of Week (Fig. 9 (b)):** Weekend traffic shows a modest increase in conversion rates compared to weekdays.

To summarize, conversion rates vary significantly based on both referral sources and timing. Marketplace and direct traffic tend to drive the most successful conversions, while other sources perform less effectively. In terms of timing, conversions are more likely to occur during afternoon hours, with weekends showing a slight edge over weekdays.

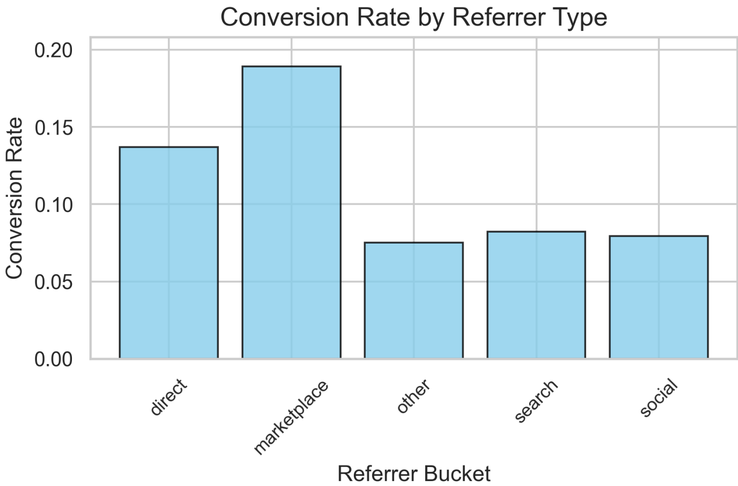

Figure 8: Conversion rates by referral source. Marketplace and direct channels outperform others.

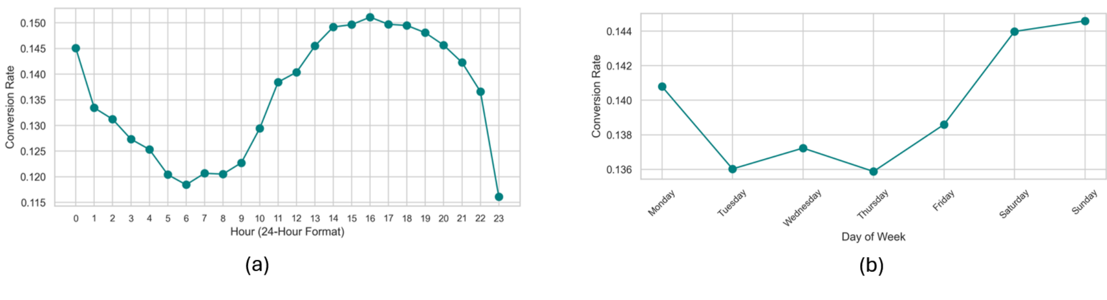

Figure 9: Temporal conversion trends by (a) hour of day and (b) day of week.

6 Customer Segmentation

6.1 Feature Construction

To profile each Amazon user, we aggregated sessions and transactions into five behavioral metrics plus demographics:

- **Frequency (`freq_sessions`):** number of unique session dates per user.
- **Recency (`avg_recency`):** average days between successive purchases. Lower is better
- **Monetary (`avg_basket`):** mean basket total per user (avg spend).
- **Engagement:** mean pages viewed (`avg_pages`) and mean session duration in minutes (`avg_duration`) per user.
- **Demographics:** head-of-household education, income, household size, census region, and racial background.

In code, we first grouped `sessions_amz` by `machine_id` to compute frequency, pages, and duration. We then merged daily baskets from `transactions_amz` to derive monetary and recency, and joined the earliest demographic record. This yielded a user-level table of RFM, engagement, and demographic features.

6.2 Determining k via Elbow and Silhouette

We standardized the 5 behavioral features and ran k-means for $k = 2$–10. Figure 10 (a) the inertia curve exhibits a clear "elbow" at $k = 4$,while the average silhouette scores in Figure 10 (b) also peak at this value. Both metrics consistently suggest that four clusters offer the most meaningful segmentation.

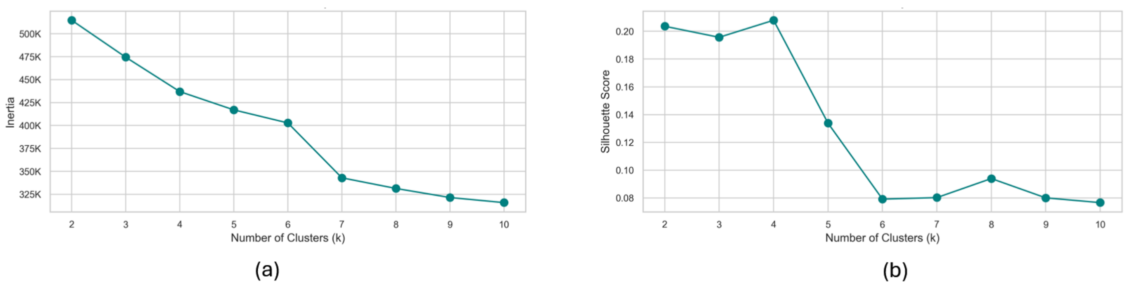

Figure 10: (a) Different value of interia for different k values in ELBOW method (b) Different silhouette score for different k values in Silhoutte method

6.3 Cluster Profiles

With $k = 4$, we analyzed the distribution of the five behavioral features across clusters (Fig. 11) to interpret and label each group:

- **Cluster 0 – Dormant:**
  - Characterized by the highest recency (median $\simeq$ 300 days) and the lowest frequency of visits.
  - Moderate levels of basket size (median $\simeq$ \$50), page views ($\simeq$ 8), and session duration ($\simeq$ 8 min).

  These users engage infrequently but tend to place average-sized orders when they return, suggesting low but not negligible value.

- **Cluster 1 – Occasional Buyers:**
  - Moderate recency (median $\simeq$ 30 days) with relatively low frequency.
  - Engagement levels and basket sizes closely mirror those of Cluster 0.

  These are infrequent yet consistent purchasers who exhibit straightforward browsing behavior and moderate order values.

- **Cluster 2 – Frequent Browsers:**
  - Most recent activity (median recency < 20 days) and the highest frequency of sessions.
  - Average engagement (pages $\simeq$ 9, duration $\simeq$ 9 min) and basket value (median $\simeq$ \$50).

  This group represents high-frequency visitors who regularly browse and occasionally purchase, contributing consistently to site traffic.

- **Cluster 3 – Engaged Spenders:**
  - Low recency and moderate frequency.
  - Highest levels of engagement (pages $\simeq$ 23, duration $\simeq$ 22 min) and basket value (median $\simeq$ \$60).

  These users are highly engaged, tend to explore thoroughly, and consistently place the largest orders—marking them as a valuable segment for targeted strategies.

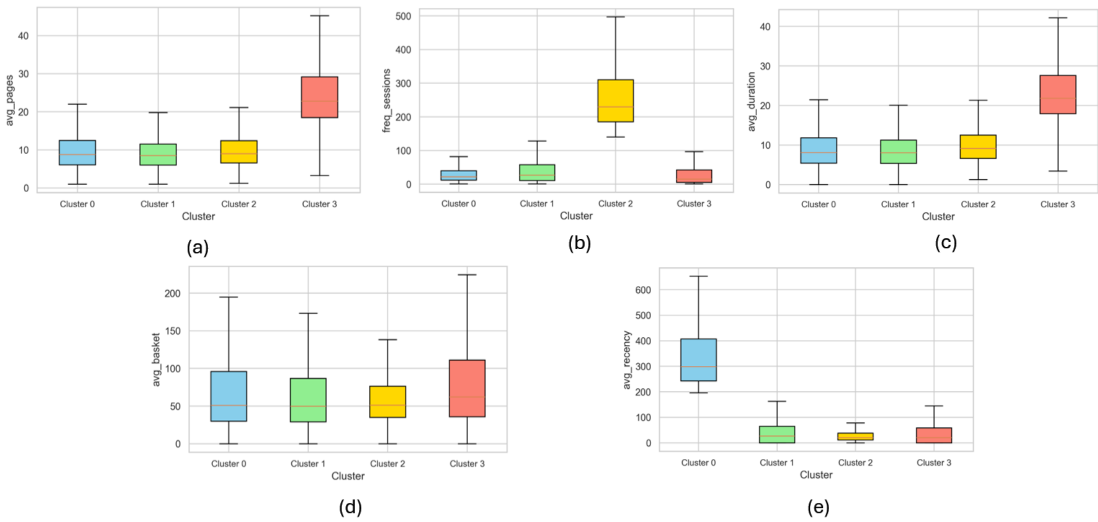

Figure 11: Distributions of behavioral features across clusters: (a) Page Views, (b) Frequency, (c) Session Duration, (d) Basket Value, (e) Recency

7 Predictive Modeling

7.1 Session-Level Feature Engineering

All modeling was performed at the session level (using `machine_id+event_date`), since many transaction records lacked valid session IDs. We derived the following features:

- **Categorical** (`cat_cols`): `weekday`, `hour`, `month`, `ref_bucket`, `hoh_most_education`, `household_income`, `census_region`, `household_size`, `hoh_oldest_age`, `children`.
- **Numerical** (`num_cols`): `pages_viewed`, `duration`, `days_since_last_purchase`.

We excluded `racial_background` to avoid introducing demographic bias. `days_since_last_purchase` was computed by joining transactions on date and taking the interval since the user's previous purchase.

7.2 Feature Selection

To better understand the drivers of conversion, we conducted a series of feature-level analyses evaluating variable importance, correlation strength, and statistical significance. This multi-faceted approach helps identify the most predictive features, uncover meaningful relationships with conversion outcomes, and eliminate less informative variables to improve model efficiency and interpretability.

- *Importance:* Random Forest feature importances (Figure 12 (a)) ranked `pages_viewed`, `duration`, and `days_since_last_purchase` as the top three predictors.
- *Correlation:* Pearson correlations (Figure 12(b) confirmed that `pages_viewed` and `duration` correlate positively with conversion, while `days_since_last_purchase` is negatively correlated.
- *Statistical Significance:* Chi-square and ANOVA tests (Figure 12 (c)) identified several high-p-value categorical levels (e.g. `weekday_Tuesday`, `census_region_99`) which were dropped later for a more parsimony model.

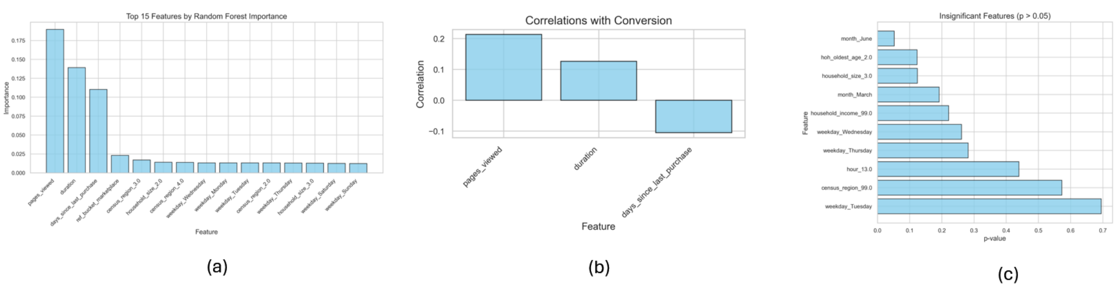

Figure 12: (a) Feature Importance based on Random Forest (b) Feature Corelation (c) Insignificant Features

7.3 Baseline Model Comparison

We benchmarked three classifiers on an Amazon test set (conversion rate ≈15%):

- **XGBoost** (`use_label_encoder=False, eval_metric=logloss`)
- **Logistic Regression** (`max_iter=1000`)
- **Random Forest** (100 trees)

As shown in Table 1, XGBoost outperformed the other models across all key metrics, achieving the highest ROC-AUC (0.87) and F_1 score (0.30). It also recorded a precision of 0.50 and a recall of 0.22 for the minority (positive) class. However, as illustrated by the confusion matrix in Figure 13, a substantial number of purchasers were still missed at the default classification threshold, reflecting the inherent difficulty of detecting rare positive outcomes in imbalanced datasets.

Table 1: Model Performance Metrics

Model	ROC AUC	Accuracy	Precision	Recall	F1
XGBoost	0.869272	0.897263	0.500616	0.215369	0.301171
Logistic Regression	0.795352	0.895298	0.436151	0.063497	0.110856
Random Forest	0.860523	0.896417	0.489726	0.183639	0.267115

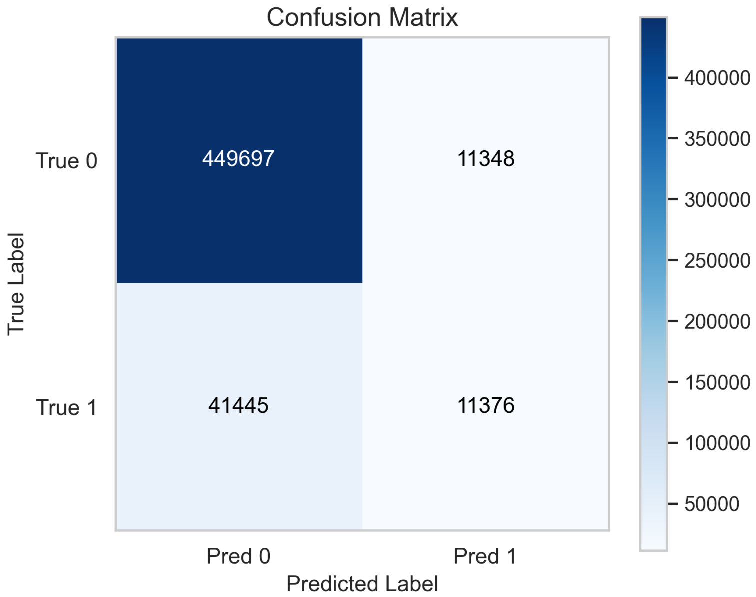

Figure 13: Confusion Matrix for baseline XGBOOST classifier

7.4 Threshold Optimization

To improve recall (critical for capturing high-value sessions), we tuned the decision threshold t to maximize F_1:

$$t^* = \arg \max_{t \in [0.1, 0.9]} F_1\big(y,\ \hat{p} > t\big).$$

The optimal threshold was $t^* \approx 0.73$. Under this threshold (Figure 14), XGBoost achieves:

- Precision: 0.37
- Recall: 0.59
- F_1: 0.45
- Accuracy: 0.85

This reflects approximately a $2.7\times$ improvement in recall, accompanied by a modest reduction in precision—an acceptable trade-off in marketing contexts, where false positives are less costly than missed opportunities to engage potential customers.

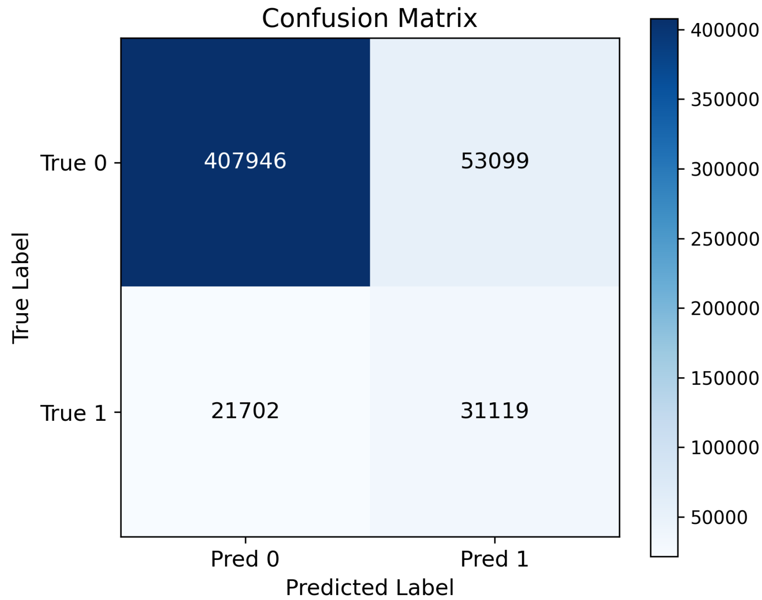

Figure 14: Confusion Matrix for threshold tuned XGBOOST classifier

7.5 Cross-Domain Generalization

Applying the tuned XGBoost model to sessions from eBay, Walmart, Etsy, and Target (conversion rate ≈10%) yields:

- Accuracy: 0.78
- Precision: 0.25
- Recall: 0.60
- F_1: 0.35

(Table 2). The recall remains consistent across domains, highlighting the model's robustness to domain shift and its effectiveness for cross-platform targeting.

Table 2: Classification Report for Non-Amazon, Personalized Model

Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (Non-Purchasers)	0.94	0.80	0.86	1,970,708
1 (Purchasers)	0.25	0.60	0.35	227,170
Accuracy			0.78	2,197,878
Macro Avg	0.60	0.70	0.61	2,197,878
Weighted Avg	0.87	0.78	0.81	2,197,878

7.6 Parsimony Analysis

Finally, we trained a "light" XGBoost using only the top three numerical features (`pages_viewed`, `duration`, `days_since_last_purchase`). On Amazon test data, this parsimony model achieves:

- Accuracy: 0.81
- Precision: 0.30
- Recall: 0.63

- F_1: 0.41

On other domains, it attains accuracy 0.79, precision 0.26, recall 0.58, F_1 0.36. This minimal model simplifies deployment with only a small trade-off in overall performance.

Business Implications. By tuning for higher recall, we can identify the majority of high-intent sessions (nearly 60%), enabling more effective retargeting campaigns. The parsimony model shows that just three behavioral signals suffice, making real-time scoring feasible at scale across multiple e-commerce platforms.

8 ROI Implications

To quantify business impact, we measure lift by comparing conversion rates in the top decile of model-scored sessions against the baseline conversion, then translate that lift into ROI assuming \$0.50 cost per targeted session.

8.1 Targeting Lift Analysis

Using our tuned XGBoost model on Amazon sessions, we rank each session by its predicted purchase probability. We then compare:

1. **Baseline conversion rate (CR)**:

$$\mathrm{CR}_{\mathrm{base}} \;=\; 10.28\%,$$

i.e. overall fraction of sessions that convert.

2. **Top 10% sessions CR**:

$$\mathrm{CR}_{\mathrm{top10\%}} \;=\; 13.17\%,$$

i.e. conversion among the highest-scoring decile.

3. **Lift**:

$$\mathrm{Lift} \;=\; \frac{\mathrm{CR}_{\mathrm{top10\%}}}{\mathrm{CR}_{\mathrm{base}}} \;\approx\; 1.28\times.$$

A random 10% sample converts at roughly the baseline rate ($\simeq$ 10.3%), so model-based targeting yields a 28% lift in conversion versus random outreach.

8.2 Incremental Orders and Revenue

Let N_{top} be the number of sessions in the top 10%. On our test set of 513,866 sessions,

$$N_{\mathrm{top}} \approx 51{,}387.$$

The model identifies

$$\Delta\mathrm{orders} \;=\; N_{\mathrm{top}} \times \left(\mathrm{CR}_{\mathrm{top10\%}} - \mathrm{CR}_{\mathrm{base}}\right) \;\approx\; 1{,}486$$

additional orders.

We compute average order value (AOV) by domain (Table 3):

Table 3: Average Order Value by Domain

Domain	AOV (\$)
amazon.com	62.21
ebay.com	65.33
etsy.com	25.43
target.com	79.64
walmart.com	97.51

For Amazon:

$$\text{Incremental revenue} = \Delta\mathrm{orders} \times \mathrm{AOV}_{\mathrm{AMZ}} \approx 1{,}486 \times \$62.21 \;=\; \$92{,}400.$$

8.3 Cost Assumptions and ROI

Assume a marketing outreach cost of \$0.50 per targeted session. Total cost for the top decile:

$$\text{Cost} = N_{\text{top}} \times \$0.50 \approx 51{,}387 \times 0.50 = \$25{,}694.$$

Net incremental profit:

$$\text{Profit} = 92{,}400 - 25{,}694 = \$66{,}706.$$

Return on investment (ROI):

$$\text{ROI} = \frac{\text{Profit}}{\text{Cost}} \approx 2.60 \quad \big(260\%\big).$$

Hence we get a return on investment of 260% assuming we deploy the ML model.

8.4 Recommendations

- **Focus outreach on the top decile** of predicted sessions to maximize incremental conversions.
- **Budgeting:** At \$0.50 per session, every dollar invested yields \$2.60 in net revenue—supporting a reallocation of marketing spend toward model-driven campaigns.
- **Sensitivity:** If outreach cost or AOV changes, ROI scales linearly; revisit this analysis as campaign economics evolve.

9 Conclusion & Future Work

9.1 Thesis Revisited

We set out to demonstrate that a combination of session-level behavioral metrics (pages viewed, session duration, recency) and basic demographics can reliably predict which e-commerce sessions will convert. Our analyses—from RFM exploratory insights to k-means segmentation to XGBoost modeling—validate this thesis and quantify its business impact.

9.2 Key Findings

- **User Segments:** Four distinct customer clusters emerged, ranging from *high-frequency, low-recency* "power shoppers" to *infrequent, long-recency* browsers, each with unique RFM and demographic profiles.
- **Predictive Performance:** An XGBoost classifier achieved ROC-AUC $\simeq$ 0.87 on Amazon sessions, outperforming logistic regression and random forest baselines. Precision–recall trade-offs were tuned to balance campaign costs versus incremental orders.
- **Lift & ROI:** Targeting the top decile of predicted sessions yields a 1.28x lift in conversion and—under a \$0.50 outreach cost assumption—delivers a 260% ROI, corresponding to $\simeq$ 1,500 additional orders and \$66K net profit on our test sample.
- **Cross-Domain Generalization:** Parsimonious models using only pages viewed, duration, and days since last purchase maintained 0.79 accuracy on non-Amazon domains (eBay, Target, Walmart), confirming transferability.

9.3 Future Work

1. **Feature Enrichment:**
   - Incorporate product-category affinities (e.g. "electronics vs. home goods") and dynamic pricing/promotions.
   - Leverage clickstream path patterns (e.g. sequence of page types) via sequence models.
2. **Online Experimentation:**
   - A/B test model-driven outreach versus generic campaigns to validate lift and ROI in production.
   - Test different messaging or incentives by customer segment.
3. **Fairness & Ethics:**
   - Audit models for unintended biases across demographic groups.
   - Incorporate fairness constraints if needed.
4. **Real-Time Scoring:**
   - Deploy a streaming inference pipeline to score sessions as they occur.
   - Integrate with real-time personalization engines.
5. **Longitudinal Monitoring:**
   - Track model performance drift over time as user behavior and market conditions evolve.
   - Schedule periodic retraining and recalibration.

By combining rigorous EDA, robust segmentation, and predictive modeling with clear ROI calculations, this project lays the groundwork for data-driven targeting strategies that can be scaled across platforms and continuously optimized.