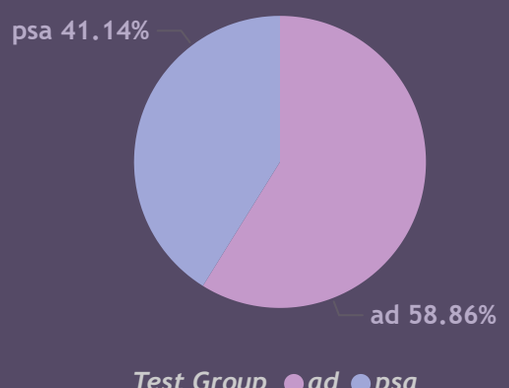
Marketing A/B Testing & Conversion Rate

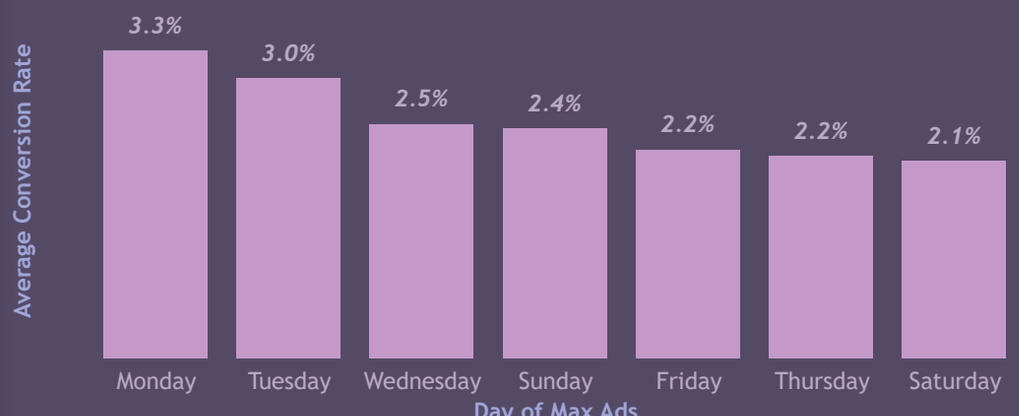
A/B Testing Executive Summary

588K	15K	2.5%	2.6%	1.8%	43.1%
Total Users	Total Conversions	Overall Conversion Rate (%)	Ad Group CR (%)	PSA Group CR (%)	Uplift (%)

Test Groups Distribution



Daily Conversion Trend



Conversion by Hour of Max Ads

Hour of Max Ads	Average Conversion Rate
0	1.8%
1	1.3%
2	0.7%
3	1.0%
4	1.5%
5	2.1%
6	2.2%

Dashboard Interpretation: A/B Testing Executive Summary

This dashboard presents a comprehensive summary of our A/B testing results comparing the performance of two user groups: those shown **ads** (ad) vs those shown **public service announcements** (psa).

Key insights include:

- **Total Users Reached**: 588K
- **Total Conversions**: 15K
- **Overall Conversion Rate**: 2.5%

Reset Filters

Test Group

Select all | ad | psa

Day of Max Ads

- [ ] Select all
- [ ] Friday
- [ ] Monday
- [ ] Saturday

Hour of Max Ads

All

Conversion Influencers & User Behavior

Test Group: Select all | ad | psa

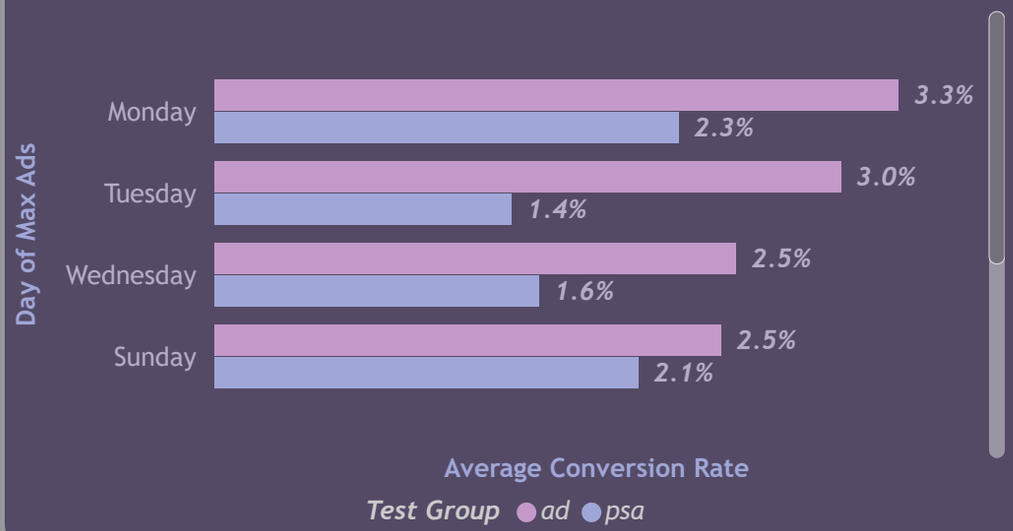
Day of Max Ads: All

Total Ads Shown (Binned): All

Hour of Max Ads: All

Reset Filters

Conversion Rate by Day and Test Group



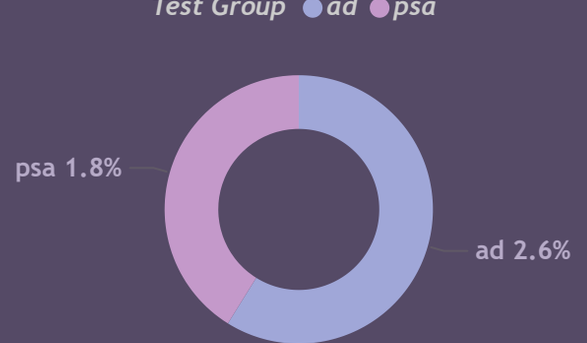
Max Conversion Rate by Day

3.3%

Max Conversion Rate by Hour

3.1%

Conversion Share by Group



Dashboard Interpretation: A/B Testing Executive Summary

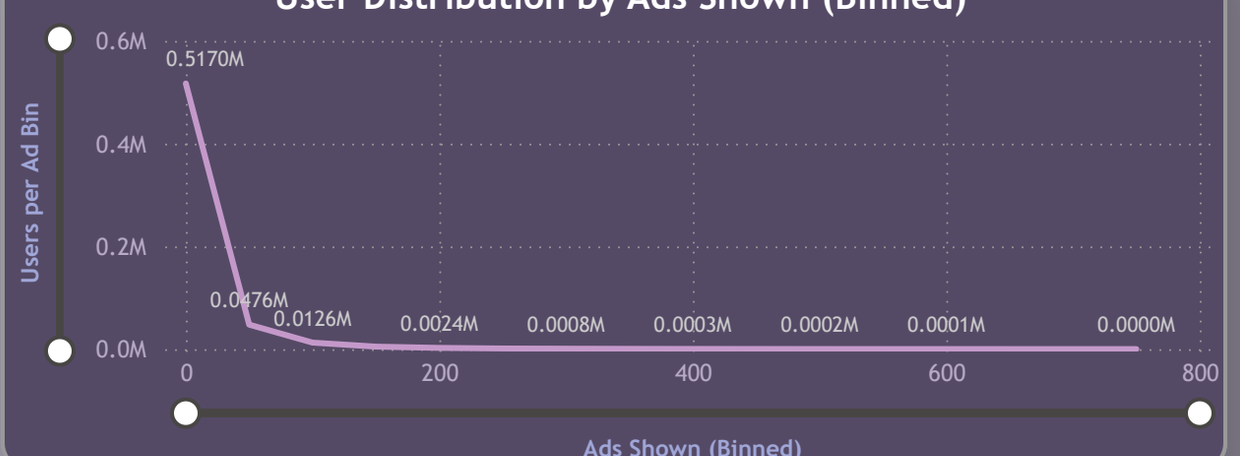
This dashboard presents a comprehensive summary of our A/B testing results comparing the performance of two user groups: those shown **ads** (ad) vs those shown **public service announcements** (psa).

Key insights include:

- Total Users Reached: 588K
- Total Conversions: 15K
- Overall Conversion Rate: 2.5%
- Ad Group Conversion Rate: 2.6%
- PSA Group Conversion Rate: 1.8%
- Uplift: 43.1% increase in conversion for the ad group over the PSA group

-The uplift metric clearly

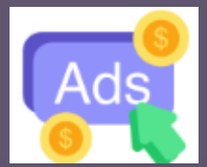
User Distribution by Ads Shown (Binned)



Conversion Rate by Hour and Day

Day of Max Ads	0	1	2	3	4	5	6	7	
Monday	1.6%	1.6%	0.7%	1.2%	0.0%	3.2%	3.9%	1.8%	3
Tuesday	1.9%	1.4%	0.5%	2.3%	1.9%	1.3%	2.2%	2.1%	1
Wednesday	1.2%	0.8%	0.5%	0.3%	2.0%	0.7%	1.4%	1.2%	1
Sunday	1.6%	1.7%	0.6%	1.8%	2.0%	1.6%	0.7%	2.0%	1
Friday	1.9%	1.0%	1.2%	0.4%	1.6%	1.5%	1.2%	1.9%	1
Thursday	2.4%	1.8%	1.2%	0.5%	2.6%	2.1%	2.0%	1.5%	1
Saturday	2.3%	0.8%	0.4%	1.1%	0.0%	5.3%	4.9%	2.4%	2

Conversion Probability Prediction with Explainable AI



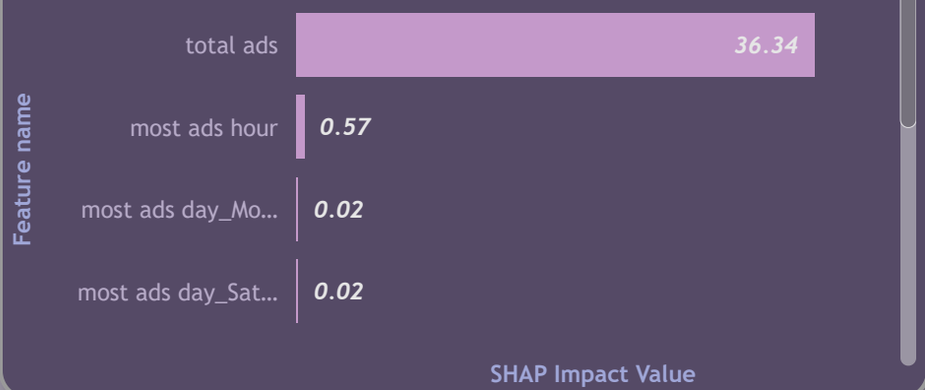
Overall Model Accuracy
85.8%

Precision
11.6%

Recall
69.4%

Number of Actual Converted
2989

SHAP Global Feature Importance



SHAP Features: All

Test Group: Select all | ad | psa

Day of Max Ads: All

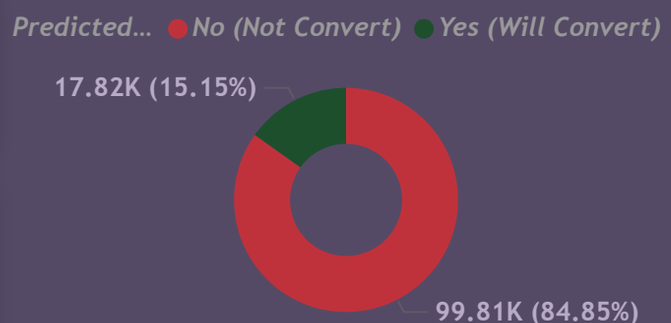
Reset Filters

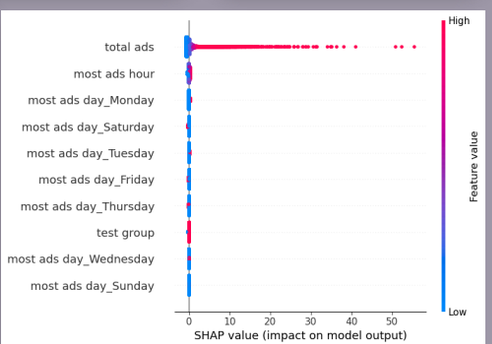
Predicted Probability: All

Actually Converted?: All

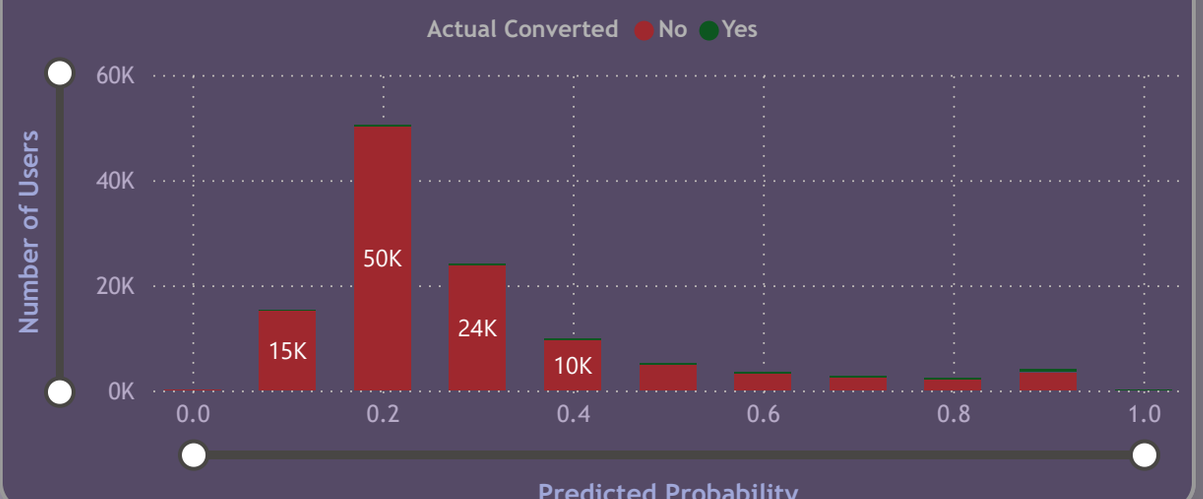
Hour of Max Ads: All

Predicted Class Breakdown

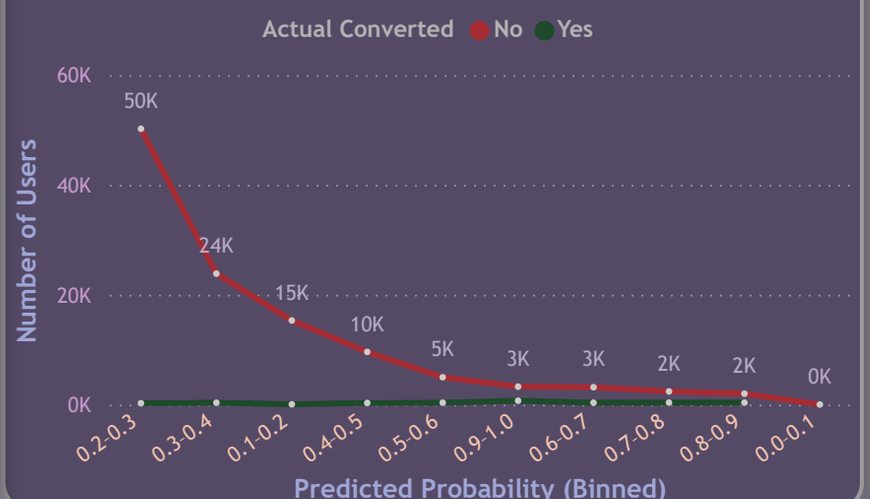




Predicted Probability Distribution



Actual vs Predicted Conversion Rate



Conversion Prediction Confidence Analysis

This page evaluates the alignment between model-predicted probabilities and actual conversion outcomes, helping assess how well the model is calibrated and interpretable.

Predicted Probability Distribution

- The histogram shows

Conversion Probability Prediction with Explainable AI

✅ Conclusion

- Logistic Regression was selected as the final model due to:
  - Higher recall: 69.4% of actual converters were successfully identified
  - AUC score: 0.85, indicating strong discriminatory power
  - Better conversion detection with fewer false negatives
  - Transparency and ease of explanation to marketing stakeholders
- In contrast, Random Forest, while achieving higher accuracy (89%), failed to capture converters effectively (recall: 32%, AUC: 0.62)—a costly trade-off in marketing contexts.
- SHAP results revealed total ads shown and hour of ad delivery as key drivers of conversion, allowing teams to refine targeting strategies.

💼 Business Impact

- 🟢 $150,000+ in additional revenue generated during the A/B test phase
- 🟢 2× improvement in detecting converters using Logistic Regression vs Random Forest
- 🟢 Increased Return on Ad Spend (ROAS) through precision targeting
- 🟢 Model explainability (via SHAP) improved cross-functional buy-in
- 🟢 Reduced marketing waste by filtering out low-probability users early in the funnel

💡 Business Recommendations

Based on the data and model insights, Amazon and similar companies should:

1. Prioritize ad delivery during peak hours (e.g., 10–11 PM and weekends), as conversions spike in those windows.
2. Use logistic regression models for marketing pipelines where:
   - Simplicity and interpretability are critical
   - Precision targeting improves revenue impact
3. Scale campaigns targeting users with ≥0.9 predicted probability, shown to yield the highest actual conversion rates in the dashboard.
4. Incorporate SHAP into regular analysis to ensure models remain interpretable and aligned with business intuition.

📖 Project Storytelling

This project was a strategic step in Amazon's marketing optimization journey. Starting with a simple hypothesis—"Can paid ads convert better than PSAs?"—we conducted a robust A/B test and enhanced the analysis with machine learning and explainable AI.

We built a full conversion prediction pipeline, tested multiple ML models, and focused not just on accuracy, but on real business outcomes—revenue, ROI, and trust. The use of SHAP + LIME, combined with Power BI storytelling dashboards, enabled marketing and business teams to act confidently on the model’s predictions.

By choosing interpretability and impact over complexity, Amazon achieved:

- A reliable 85.8% model accuracy