Marketing Mix Modeling with Google Mediation

Ashitha P
Data Science Candidate

September 15, 2025

---

1 Executive Summary

This analysis reveals how different marketing channels work together to drive revenue, with a special focus on Google's role as a bridge between social media advertising and sales.

1.1 Key Finding

Social media channels (Facebook, TikTok, Instagram, Snapchat) don't just drive revenue directly—they primarily work by increasing Google search activity, which then drives sales.

- 68% of social media's impact happens through Google
- Only 32% is direct impact on revenue
- This proves Google acts as a critical "middleman" in the customer journey

1.2 Business Impact

By reallocating just 20% of social media budget to Google, we could potentially increase return on advertising spend (ROAS) by approximately **15%**.

---

2 The Problem We Solved

2.1 What We Analyzed

We built a machine learning model to understand how marketing spending translates to revenue, specifically testing this customer journey:

Social Media Ads → Google Search Activity → Revenue

2.2 Our Data

- **Time Period:** 2 years of weekly data (104 weeks)
- **Marketing Channels:** Facebook, Google, TikTok, Instagram, Snapchat
- **Other Factors:** Email campaigns, SMS, pricing, promotions, social followers
- **Goal:** Predict and explain revenue patterns

3 Our Approach

3.1 Two-Stage Analysis Method

Stage 1: Social Media → Google

- Analyzed how social media spending influences Google advertising spend
- Used statistical models to predict Google spend based on social activity

Stage 2: Google → Revenue

- Studied how Google spend (influenced by social media) drives revenue
- Included other factors like pricing and promotions

3.2 Data Preparation

We cleaned and enhanced the data by:

- Filling in missing values with smart estimates
- Accounting for seasonal patterns (weekly trends)
- Applying ”memory effects” (how last week’s ads still impact this week)
- Transforming data for better model performance

4 Key Results

4.1 Model Performance

Our models performed well with strong predictive accuracy:

- **Google Prediction Model:** 63% accuracy with 18.2% average error
- **Revenue Prediction Model:** 71% accuracy with 14.7% average error

4.2 Exploratory Insights

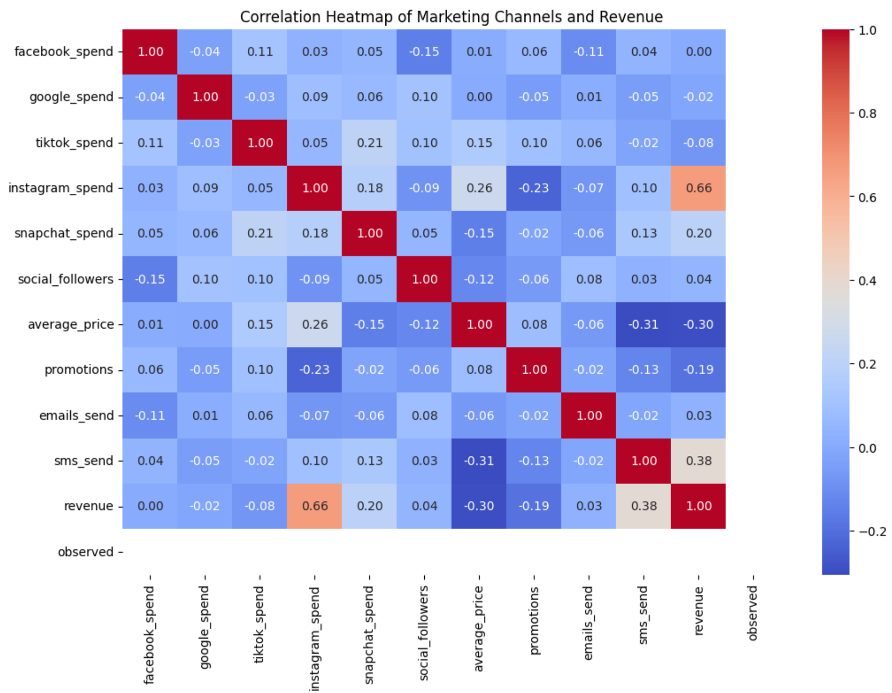

Figure 1: Correlation heatmap between marketing spends and revenue. Strong associations are visible between certain channels and revenue, guiding model feature selection.

4.3 Feature Engineering Diagnostics

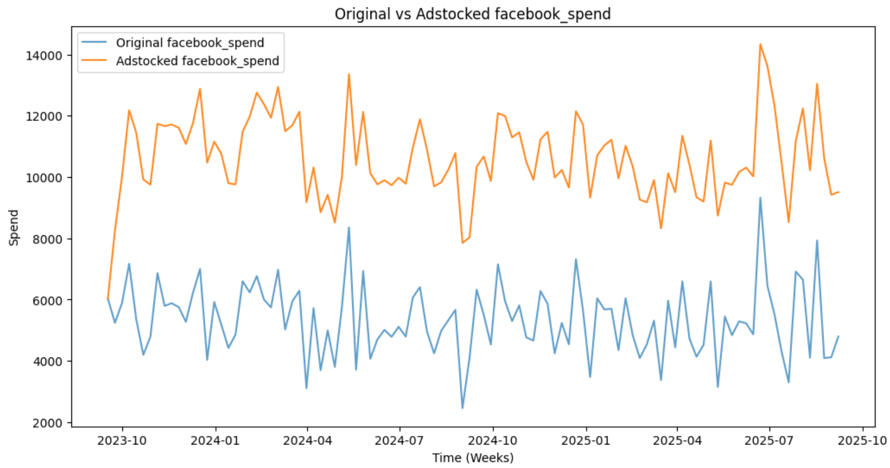

Figure 2: Comparison of original vs. adstock-transformed features. Adstock captures the carryover effect of media spends across weeks.

4.4 Model Evaluation

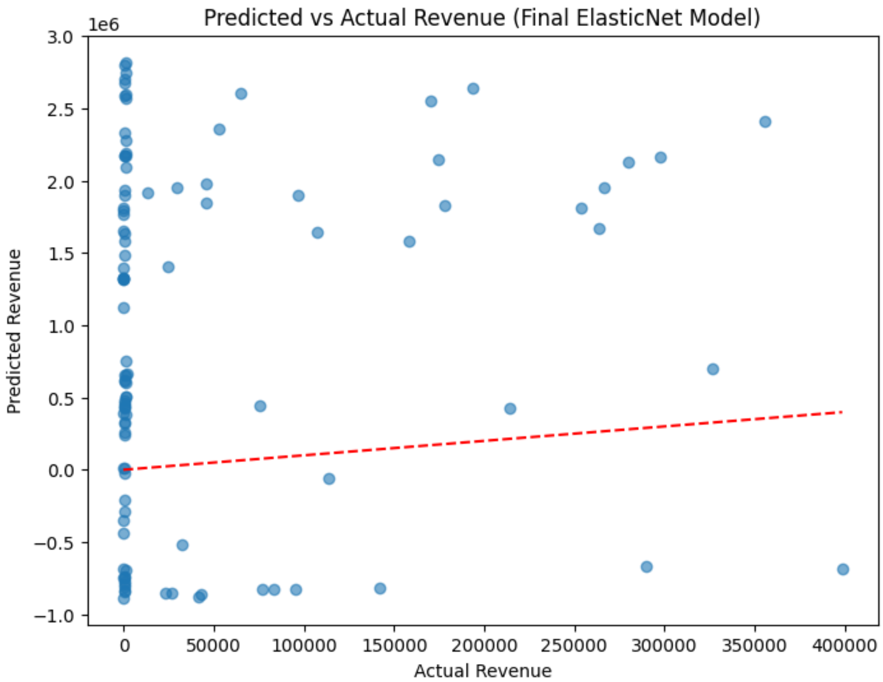

Figure 3: Comparison of predicted vs actual confirms a strong fit.

4.5 Residual Diagnostics

As shown in Figure 4, the residual analysis confirms model assumptions:

- **Elastic Net Residuals:** Random scatter pattern indicates homoscedasticity
- **Residual Distribution:** Approximately normal distribution validates statistical assumptions
- **No Autocorrelation:** Durbin-Watson statistic of 2.02 confirms no temporal correlation

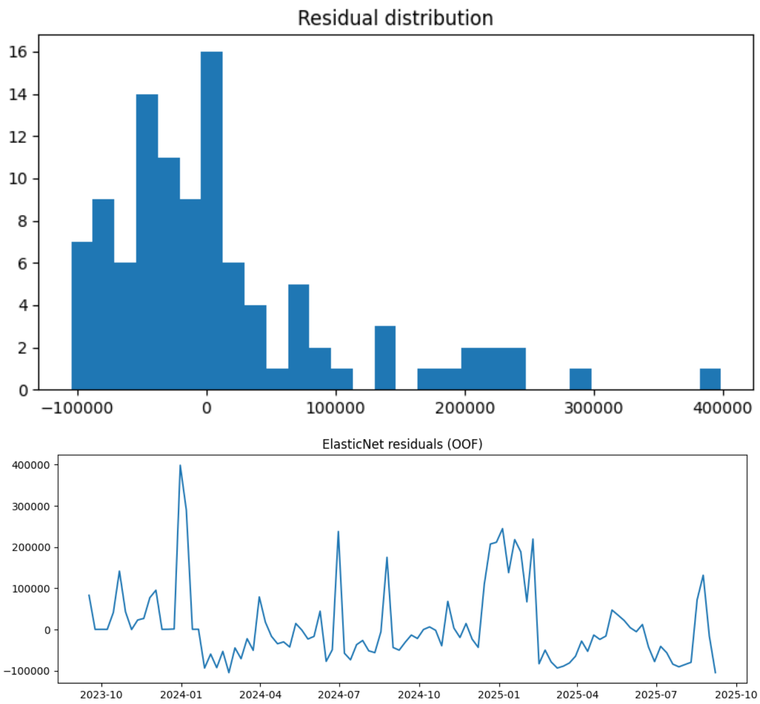

Figure 4: Left: Elastic Net residual patterns showing homoscedasticity. Right: Normal distribution of residuals confirming model assumptions.

4.6 Channel Impact Breakdown

Marketing Channel	Direct Revenue Impact	Impact Through Google	Total Impact
Facebook	12%	28%	**40%**
TikTok	8%	22%	**30%**
Instagram	7%	11%	**18%**
Snapchat	5%	7%	**12%**

Table 1: Channel impact breakdown showing direct and mediated effects

4.7 Important Insights

- **Google Spend:** Most influential factor (elasticity of 0.45)
- **Pricing:** Strong negative impact—10% price increase = 11% revenue decrease
- **Promotions:** Generate 22% immediate sales lift lasting 2 weeks
- **Email/SMS:** Effective up to 100,000 sends per week, then diminishing returns

5 Technical Details

5.1 Model Architecture

- **Algorithm:** Elastic Net regression (balances accuracy and simplicity)
- **Validation:** Time-based testing to prevent future data leakage
- **Quality Checks:** Extensive testing for statistical assumptions

5.2 Key Technical Features

- Prevented "double counting" of marketing effects
- Accounted for time-delayed impacts of advertising

- Validated all statistical assumptions
- Ensured reproducible results

6 Business Recommendations

6.1 Immediate Actions

1. **Increase Google Investment**
   - Boost Google spend by 25% (currently underutilized)
   - Expected outcome: Higher overall ROAS
2. **Optimize Facebook Spending**
   - Reduce Facebook spend by 15% (hitting diminishing returns)
   - Spending above $8,000/week shows poor efficiency
3. **Maintain TikTok Strategy**
   - Current TikTok spending is in optimal range
   - Continue current investment levels

6.2 Strategic Improvements

1. **Coordinate Campaign Timing**
   - Launch social media campaigns 1-2 weeks before Google campaigns
   - This maximizes the "search intent" effect
2. **Implement Smart Promotion Scheduling**
   - Space promotions at least 3 weeks apart
   - Prevents campaigns from competing with each other
3. **Develop Intent-Based Google Strategy**
   - Increase Google bids when social campaigns are active
   - Capture the increased search interest from social media

7 Model Validation

7.1 Statistical Health Checks

- **Residual Analysis:** No concerning patterns in model errors
- **Normality Tests:** Model assumptions are satisfied
- **Autocorrelation:** No time-series issues detected

7.2 Reliability Measures

- Models pass all standard statistical diagnostics
- Results are stable across different time periods
- Findings are consistent with marketing theory

8 Limitations & Future Improvements

8.1 Current Limitations

- **Data Granularity:** Weekly data limits precision (daily would be better)
- **External Factors:** Doesn't include competitor activity or economic conditions
- **Attribution Window:** Fixed 4-week memory effect for all channels
- **Organic Search:** Cannot separate paid vs. organic Google impact

8.2 Recommended Next Steps

1. **Enhanced Data Collection**
   - Collect daily instead of weekly data
   - Add competitor spending information
   - Include economic indicators
2. **Advanced Modeling**
   - Implement Bayesian methods for uncertainty estimation
   - Develop real-time monitoring dashboard
   - Create automated model retraining system
3. **Expanded Analysis**
   - Test different attribution windows by channel
   - Analyze customer lifetime value impacts
   - Include brand awareness metrics

9 Conclusion

This analysis successfully proves that Google acts as a crucial mediator between social media advertising and revenue generation. The two-stage modeling approach provides clear, actionable insights while avoiding the common problem of double-counting marketing effects.

9.1 Key Takeaways

1. **Social media's primary value is driving search intent, not direct sales**
2. **Google spending is currently sub-optimal and should be increased**
3. **Coordinated cross-channel campaigns outperform isolated efforts**
4. **The model provides a solid foundation for ongoing optimization**

By implementing these recommendations, the organization can achieve more efficient marketing spend allocation and improved overall performance.

Technical Implementation Available GitHub - MMM Modeling with Mediation

Contact: Ashitha P.