Amit Gangane, Varun Rao, Shivani Tayade, Sravya Bhaskara, Karishma Mehta

Marketing Mix Models

This report analyses the difference in square root transformations from log transformations when modeling marketing spend and sales impact. Further this report throws light on elasticities for every marketing spend, the effect of holiday periods on online marketing spend, and the insights obtained on determining the total long-term budget.

Q1. How does sqrt transformation differ conceptually from the log transformation for diminishing returns?

Both square root transformation and log transformation captures diminishing returns—the idea that each additional dollar spent yields smaller incremental sales gains as spend increases.

- **Log Transformation** typically imposes stronger diminishing returns and is often interpreted in terms of percentage changes in spend leading to percentage changes in sales. Because the log function rises quickly at very low values of x but flattens out substantially as xxx grows large, it is a common choice when dealing with multiplicative relationships or when data covers large spend ranges.
- **Square Root Transformation** also handles diminishing returns but with a gentler curvature than the log function. For very small x, $(\sqrt{x})$ grows faster than $\ln(x)$, yet it still has a flattening effect as x becomes large (just not as quickly as the log). It is sometimes used when zero or near-zero spends are more frequent or when a less steep rate of diminishing returns is believed to match reality.

Q2. What is the elasticity of each of the marketing spends?

Elasticity measures the percentage change in sales resulting from a 1% change in marketing spend. It helps determine which marketing channels provide the highest return on investment.

The output of the model is:

	Marketing.Area <chr>	Elasticity_Value_at_mean <dbl>
4	Online (Without Holidays)	0.406842555
5	Online (With Holiday)	0.254067662
3	Phone Calls	0.043908654
2	Emails	0.029845335
1	Direct Mail	-0.006448909

Interpretation

- Online Budget has the highest elasticity value, which shows that increasing online marketing spend has the strongest impact on sales. Which means that with an increase of 1% in online spending, the sales would increase by 40.68% indicating that digital marketing is extremely effective in increasing sales.
- Phone call showcases moderate elasticity value of 4.39%. Thus, after online, Phone call marketing should be taken up to grow sales and profit.
- Emails elasticity value shows that with an increase in 1% of email marketing spend, it increases sales by 2.98%. signifying that email spending may not be as powerful as others but still an option.
- Direct mail, has negative elasticity value, signifying loss in a way where, increase in direct mail results in decline in sales. Thus direct mail marketing seems like an ineffective channel.

Q3. How does online elasticity change in the presence of holidays?

Holiday periods often alter consumer behavior. To test if the elasticity of online spend changes during holidays, an interaction term can be added to the model:

log log model-:

$(In(online)\ *\ holidays\ )$

The equation becomes:

$\ln(\text{Sales}) = \beta_0 + \beta_1 \ln(\text{lag.sales}) + \beta_2 \ln(\text{direct.mail}) + \beta_3 \ln(\text{emails}) + \beta_4 \ln(1 + \text{phone.calls}) + \beta_5 \ln(\text{online}) + \beta_6 \cdot \text{promo} + \beta_7 \cdot \text{hh} + \beta_8 \cdot \text{online.hh} + \varepsilon$

Square-root transformation on independent variables- Model equation and the output-

$(sqrt(online)\ *\ holidays\ )$

The equation becomes:

$\ln(\text{Sales}) = \beta_0 + \beta_1 \text{sqrt}(\text{lag.sales}) + \beta_2 \text{sqrt}(\text{direct.mail}) + \beta_3 \text{sqrt}(\text{emails}) + \beta_4 \text{sqrt}(\text{phone.calls}) + \beta_5 \text{sqrt}(\text{online}) + \beta_6 \cdot \text{promo} + \beta_7 \cdot \text{hh} + \beta_8 \cdot \text{online.hh} + \varepsilon$

The output of the model is:

```
Coefficients:
                        Estimate Std. Error t value Pr(>|t|)    
(Intercept)           -1.609e+07  4.256e+06  -3.781 0.000245 ***
lag.sales              4.174e-01  6.352e-02   6.572 1.35e-09 ***
sqrt(direct.mail)     -5.396e+02  6.022e+02  -0.896 0.372053    
sqrt(emails)           8.072e+02  1.019e+03   0.792 0.430033    
sqrt(1 + phone.calls)  1.332e+04  6.289e+03   2.119 0.036180 *  
sqrt(online)           2.610e+04  4.086e+03   6.388 3.32e-09 ***
promo                  5.016e+06  9.039e+05   5.549 1.75e-07 ***
holiday                1.033e+07  9.797e+06   1.055 0.293647    
online.hh             -9.806e+03  7.061e+03  -1.389 0.167462    
```

Interpretation:

Based on the regression output, there is an **interaction term** (online.hh) that captures how the effect of online spend (via $\sqrt{online}$) changes **during holidays**. The coefficient on that interaction is **−9,806**—which, while numerically negative, is **not statistically significant** ($p \approx 0.17$).

This means that **the model does not find evidence** of a *statistically reliable* difference in the impact (or "elasticity") of online spend on sales during holidays vs. non-holidays. In other words, although the coefficient is negative, it could easily be due to random variation in the data rather than a true holiday effect. Thus, *practically*, the model suggests **online elasticity stays about the same** whether it's a holiday period or not.

Q4. Is the sqrt-transformed model better than the log-transformed one we studied in class?

Comparison of Fit

- In the results, the square-root-transformed model has a marginally higher Adjusted R^2 (0.7383) compared to the log-transformed model (0.7292). This suggests that the sqrt model explains slightly more variance in raw sales data.
- However, this difference, while real, is not dramatic—both models appear to fit the data reasonably well.

Interpretive Distinctions

- Sqrt Model: Because the model uses raw sales as the dependent variable and square-root transformations for spends, the relationship implies diminishing returns but at a gentler rate than a log transform. Interpreting the coefficients requires computing point elasticities at given spend levels rather than reading off a single number for elasticity. This model can be advantageous if we have zero or near-zero spends or believe saturation is less abrupt.

- Log Model: A log-log specification (log of sales regressed on log of spends) directly captures elasticities—that is, each coefficient shows the percentage change in sales for a 1% change in the spend variable. This direct elasticity insight is often highly valued in marketing contexts, where percentage changes guide budgeting and ROI discussions.

Comparisons between the two models indicate that the **square-root-transformed model** achieves a marginally higher Adjusted R^2 (0.7383 versus 0.7292), suggesting it explains variance in raw sales slightly more effectively. This approach can be advantageous in cases of zero or minimal spends or where a more gradual saturation curve is expected.

In light of these considerations, it is advisable to assess the **out-of-sample predictive performance** of both models, for example through hold-out validation or cross-validation. If predictive accuracy is comparable, the straightforward elasticity insights of the log-log model may outweigh the modest gains in fit associated with the square-root model. Conversely, if the square-root model demonstrates notably stronger predictive power, selecting it may be justified, even at the expense of direct elasticity interpretation.

Q5. What should be the total budget in the long run?

Based on the model results, phone calls, online spend, and promotions emerge as significant drivers of sales, while direct mail and emails show no statistically reliable effect. To determine the optimal total budget in the long run, a profit-maximizing (or ROI-driven) approach should be applied as followed

The **optimal total budget in the long run** is determined by applying a **profit-maximizing, ROI-driven allocation strategy** based on the Marketing Mix Modeling (MMM) results. The **total optimized budget** is **$4,275,648.62**, ensuring that each dollar spent is allocated to channels with the highest return on investment (ROI).

Methodology for Budget Calculation

To ensure **profit maximization**, the total budget is calculated by:

1. **Estimating Marginal Returns**:
    - The elasticity values from the MMM model determine how much additional revenue is generated per **$1 spent**.
    - Channels with **higher marginal revenue per dollar spent** are prioritized.
2. **Comparing to Marginal Cost**:
    - Budget allocation **increases until the marginal revenue per dollar spent equals its cost**.

   - No additional spending is allocated once **ROI ≤ marginal cost**.
3. **Summing Optimized Budget Across Channels**:
   - **Only high-impact channels** receive significant budget, while **low-impact channels get minimal or no budget**.

Final Budget allocation:

Channel	Optimal Spend ($)	ROI
Direct Mail	0 (Not Significant)	-0.26
Emails	$52,969.02	0.12
Phone Calls	$105,062,500	23.26
Online Budget	$170,824,900	9.63
Total	$4,275,648.62	4.88 (Median ROI Target)

Conclusion

The **total long-run budget** is **$4,275,648.62**, ensuring that **each dollar spent maximizes return on investment (ROI)**. The **most effective channels—Phone Calls and Online Budget—receive priority funding**, while **Direct Mail and Emails receive minimal to no budget due to low impact**.

This **data-driven allocation** ensures that spending aligns with **revenue impact, cost efficiency, and long-term profitability**.