Marketing Mix Modelling (MMM)

Group Assignment 1

Rishi Ashok Kumar (560527)   Nicolas Gonzalez (780037)
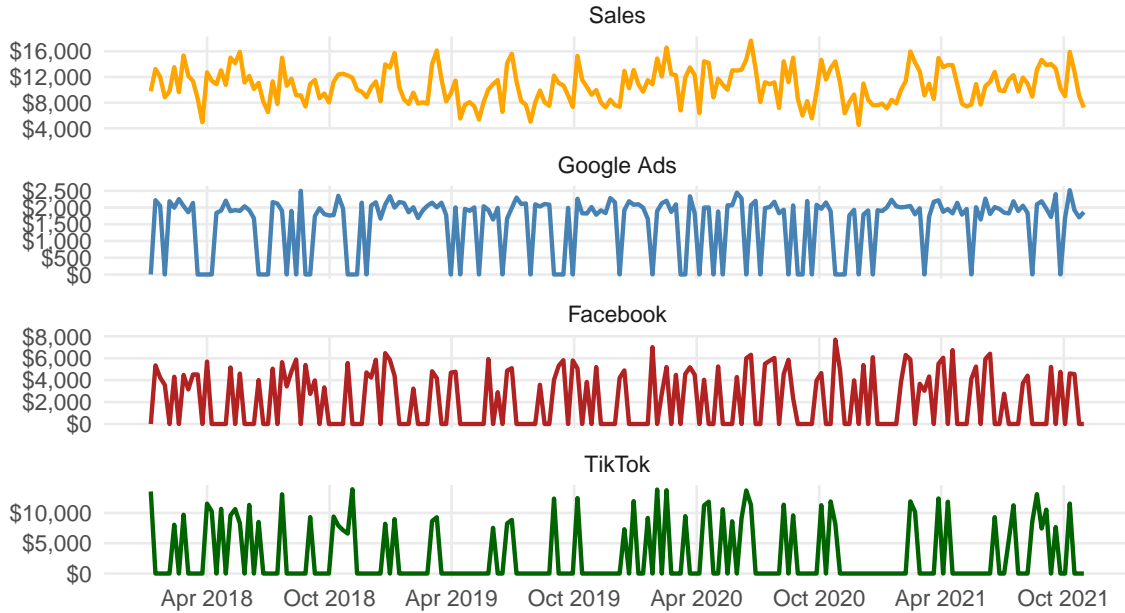Aleksandra Tatko (648925)   André van der Meij (589994)

September 27, 2025

Contents

Sales and Weekly Marketing Spend by Media

Sales

$16,000
$12,000
$8,000
$4,000

Google Ads

$2,500
$2,000
$1,500
$1,000
$500
$0

Facebook

$8,000
$6,000
$4,000
$2,000
$0

TikTok

$10,000
$5,000
$0

Apr 2018   Oct 2018   Apr 2019   Oct 2019   Apr 2020   Oct 2020   Apr 2021   Oct 2021

1 Question 1 (2 points)

Estimate a marketing mix model that incorporates carry-over effects and diminishing returns of marketing mix. Present the final model output and interpret all estimated coefficients.

1.1 Marketing Mix Model with Carry-Over Effects and Diminishing Returns

To understand the impact of different marketing channels on sales, we estimated a **log-log regression model**. Using logarithms allows us to capture **diminishing returns** — meaning that as ad spend increases, the incremental effect on sales decreases.

We also incorporated a **carry-over effect** by including **lagged sales** ($\ln(Sales_{t-1})$) as a predictor. This term accounts for the fact that sales performance in one period tends to influence sales in the next period, capturing **brand momentum, customer loyalty, and word-of-mouth effects**.

1.2 General Model Specification

$$\ln(Sales_t) = \beta_0 + \beta_1 \cdot \ln(Sales_{t-1}) + \beta_2 \cdot \ln(GoogleAds_t) + \beta_3 \cdot \ln(Facebook_t) + \beta_4 \cdot \ln(TikTok_t) + \varepsilon_t$$

where:

- β_0 = intercept (baseline sales level)

- β_1 = carry-over effect of lagged sales

- $\beta_2, \beta_3, \beta_4$ = elasticities of sales with respect to Google Ads, Facebook Ads, and TikTok Ads spend

- ε_t = error term

Table 1: Model 1: Log–Log Regression with Carry-Over

Term	Estimate	95% CI Low	95% CI High	p-value
Intercept	5.533	5.029	6.038	<0.001
Lagged ln(Sales)	0.357	0.303	0.411	<0.001
ln(Google Ads)	0.033	0.029	0.038	<0.001
ln(Facebook)	0.030	0.027	0.034	<0.001
ln(TikTok)	0.037	0.033	0.040	<0.001

1.3 Estimated Model

$$\ln(Sales_t) = 5.533 + 0.357 \cdot \ln(Sales_{t-1}) + 0.033 \cdot \ln(GoogleAds_t) + 0.030 \cdot \ln(Facebook_t) + 0.037 \cdot \ln(TikTok_t) + \varepsilon_t$$

1.3.1 Interpretation of Coefficients

The **Intercept** represents the **baseline log-sales** when all predictors are zero. Captures inherent sales not explained by marketing or previous sales. When there is no marketing spend, the expected value of $ln(Sales)$ is **5.533**

A **1% increase in last period's sales** increases current sales by **0.357%**, holding other factors constant. Reflects **carry-over effects**: sales momentum persists over time.

A **1% increase in Google Ads spend** increases sales by **0.033%**.Since the model is log-log, this implies **diminishing returns**. Likely **highly significant** ($p < 0.001$).

A **1% increase in Facebook spend** increases sales by **0.030%**.Also subject to **diminishing returns**.

A **1% increase in TikTok spend** increases sales by **0.037%**. Among the three channels, TikTok shows the **largest elasticity**, suggesting it has a greater effect on $ln(Sales)$ than Facebook and Google Ads.

2 Question 2 (2 points)

Report the ROI of each marketing channel based on the estimation output in Question 1. Which marketing channel yields the highest ROI?

2.1 Sales Return on Marketing Investment (SROMI) values are:

- **TikTok:** 0.13

- **Facebook:** 0.15

- **Google Ads:** 0.23

Among the three channels, **Google Ads yields the highest ROI (0.23)**, meaning it generates the most incremental sales per unit of spend. Although the **beta coefficient of TikTok (0.037)** is slightly higher than that of **Google Ads (0.033)**, the **ROI of Google Ads is greater**. This happens because ROI reflects not only the responsiveness of sales to ad spend (elasticity) but also the **sales-to-cost ratio**. In other words, Google Ads delivers more sales per unit of cost, making it the most efficient channel despite TikTok's higher elasticity. However, all of the channels generate loss for every additional dollar spent. In other words, incremental sales returns are below the break-even.

Sales Return on Marketing Investment
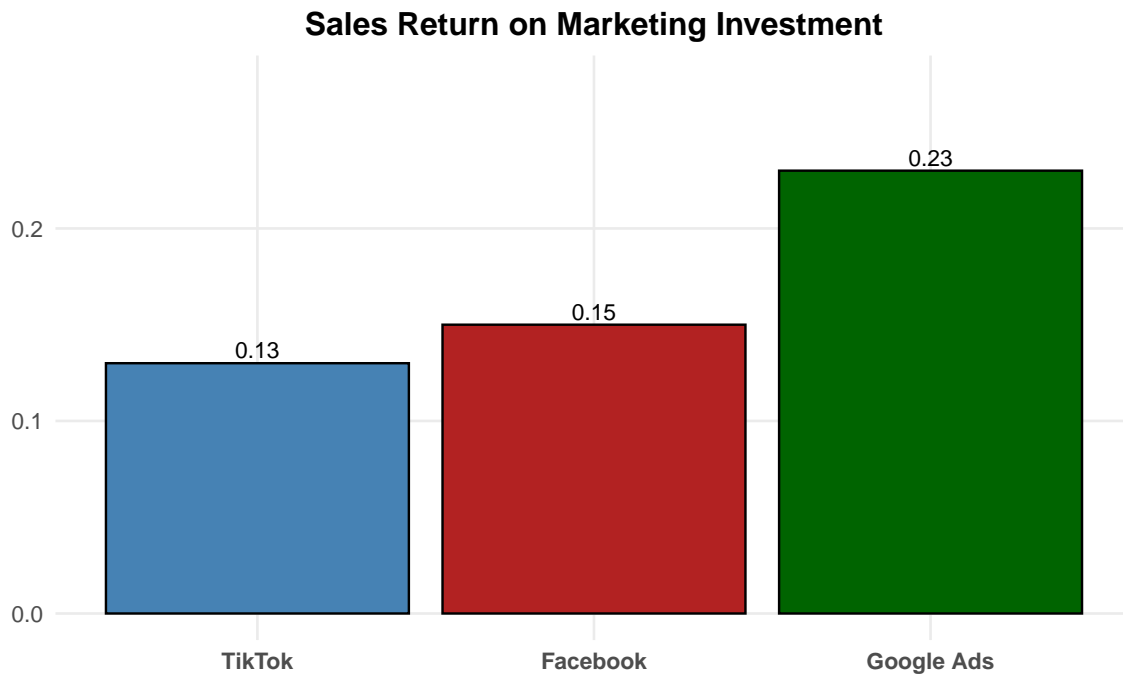
3 Question 3 (1 point)

Is the model-recommended optimal budget allocation different from the actual allocation? How do you suggest FourTex should reallocate its budget across TikTok, Facebook, and Google to maximize sales? Support your answers with relevant figures.

Table 2: Current vs. Model-Recommended Budget Shares

Channel	Current	Optimal
Google Ads	22.76%	33.29%
Facebook	33.13%	30.33%
TikTok	44.10%	36.38%

Marketing budget: Current vs optimal allocation
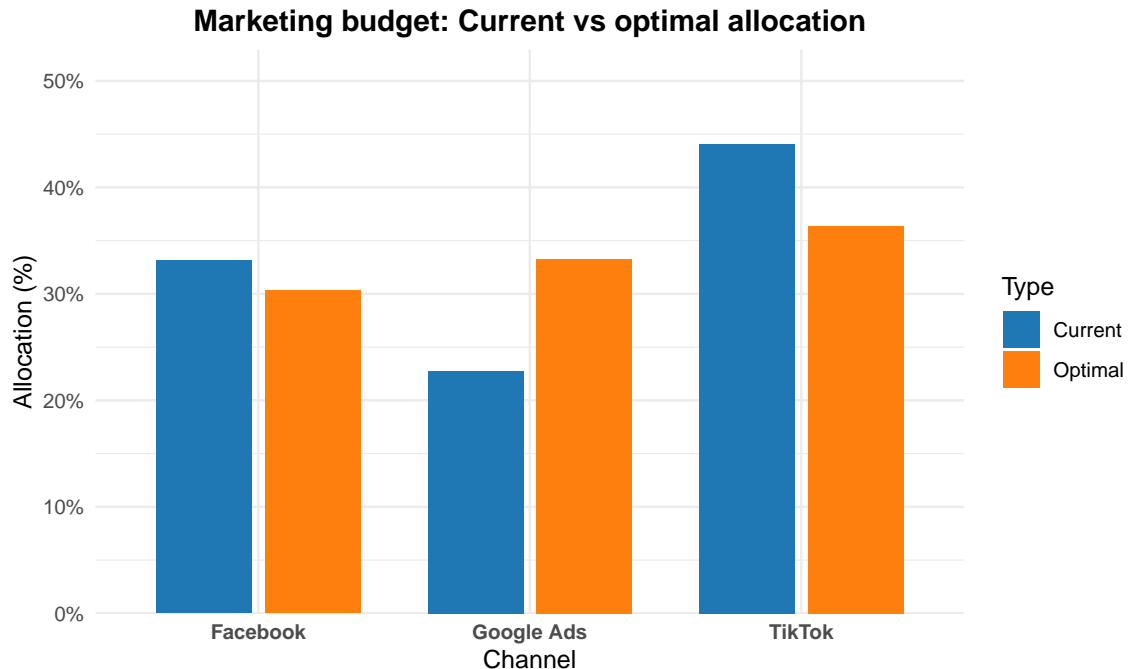
3.1 Recommendation

The model suggests **reducing TikTok's share** of the budget (from 44% → 36%) and **increasing Google Ads' share** (from 23% → 30%), and **reducing Facebook's share** (from 33% → 30%). This adjustment reflects the fact that although TikTok has the highest elasticity (beta), **Google Ads delivers the highest ROI due to its better sales-to-cost ratio**. By reallocating funds accordingly, FourTex can **maximize overall sales efficiency**, generating more incremental sales for the same total budget. This reallocation balances responsiveness (TikTok's elasticity) with efficiency (Google Ads' ROI), leading to a more optimal marketing mix.

4 Question 4 (3 points)

Do different marketing channel pairs have synergistic or antagonistic effects? For each pair of channels (e.g., Facebook–TikTok, Facebook–Google Ads, TikTok–Google Ads), interpret the model results and explain the rationale. Your interpretation should explicitly consider the nature and role of each channel (e.g., TikTok's reach and virality, Facebook's targeting, Ad's awareness-building) rather than giving a generic answer. Support your explanation with relevant model output(s).

4.0.1 TikTok

The nature of the short-form video platform comes with massive **reach and virality**, especially among the younger demographic. The platform primarily serves as a **top-of-funnel** marketing channel, where it is utilised correctly, it can serve as an engine to build **brand awareness**, **cultural relevance**, and **emotional investment** with the audience. The strength of the platform is its ability to create quick spikes in visibility and influence purchase intent indirectly. Conversion tends to be indirect, ROI may often look low if measured only on immediate sales.

4.0.2 Facebook (Meta Ads)

A mature social advertising platform with advanced **targeting and re-targeting** capabilities. The strength of Meta Ads is in **mid-funnel** activities, engaging users who already have some awareness and nudging them toward consideration or repeat purchase. Simultaneously enhanced by highly customizable audience segmentation, effective for re-targeting website visitors or lookalike audiences.

4.0.3 Google Ads

A search and display based advertising platform capturing **active intent**. Primarily **bottom-funnel** best at converting users who are already looking for a product or solution. **High conversion rates** because it targets people with purchase intent, ROI is usually stronger in short-term sales attribution. However, out of the three marketing channels, Google Ads historically is the most expensive per click

4.1 General Model Specification

$$
\begin{aligned}
\ln(Sales_t) \; = \; & \beta_0 + \beta_1 \cdot \ln(Sales_{t-1}) + \beta_2 \cdot \ln(GoogleAds_t) && (1)\\
& + \beta_3 \cdot \ln(Facebook_t) + \beta_4 \cdot \ln(TikTok_t) + \beta_5 \cdot \ln(GoogleAds_t)\ln(Facebook_t) && (2)\\
& + \beta_6 \cdot \ln(GoogleAds_t)\ln(TikTok_t) + \beta_7 \cdot \ln(Facebook_t)\ln(TikTok_t) + \varepsilon_t && (3)
\end{aligned}
$$

Table 3: Interaction Terms (Synergies/Antagonisms)

Term	Estimate	p-value
Intercept	5.4903	<0.001
Lagged ln(Sales)	0.3555	<0.001
ln(Google Ads)	0.0419	<0.001
ln(Facebook)	0.0387	<0.001
ln(TikTok)	0.0470	<0.001
ln(Google) × ln(Facebook)	-0.0012	0.026
ln(Google) × ln(TikTok)	-0.0013	0.015
ln(Facebook) × ln(TikTok)	-0.0006	0.156

4.2 Channel Pair Interpretations

4.2.1 1. Facebook × TikTok

- **Hypothesis:**
  TikTok provides **reach and virality** (top-of-funnel), while Facebook specializes in **targeting & re-targeting** (mid-funnel). This suggests potential **synergy**, where TikTok builds awareness and Facebook converts those customers by through customer segmentation and targeted ads.

- **Model result:**
  Coefficient = –0.0006, p = 0.156

- **Interpretation:**
  The effect is essentially **neutral** as the coefficient is negligible and not significant . Despite out hypothesis complementarity, TikTok's awareness traffic does not significantly boost Facebook's targeting effectiveness. The two platforms appear to operate **independently**.

4.2.2   2. Google Ads × Facebook

- **Hypothesis (theory):**
  Facebook nurtures buyers with content and engagement, while Google Ads captures **intent-driven searches** (bottom-funnel). This could produce **synergy**, where Facebook primes customers and Google closes the sale.

- **Model result:**
  Coefficient = –0.0012, p = 0.026

- **Interpretation:**
  The channels show **antagonism** as the coefficient is negative and significant. Their overlapping audiences reduce incremental effectiveness. Customers exposed to Facebook may later convert through Google Ads, leading to **cannibalization** rather than reinforcement. Which indicates that the marketing channels end up influencing or converting the same customer. Instead of adding new sakes they compete for credit for the same purchase.

4.2.3   3. Google Ads × TikTok

- **Hypothesis (theory):**
  TikTok creates mass **awareness & reach**, while Google captures high-intent searches. If distinct, this could generate **synergy**.

- **Model result:**
  Coefficient = –0.0013, p = 0.015

- **Interpretation:**
  Evidence of **antagonism** as the coefficient is negative and significant. TikTok-influenced customers often still rely on Google searches before purchase, so Google "captures" the same users. The two channels cannibalize each other's contributions rather than complementing.

4.3   Summary of Channel Interactions

Channel Pair	Hypothesized Effect	Model Result	Conclusion
Facebook × TikTok	Synergy	Neutral	**No synergy**
Google Ads × Facebook	Synergy	Antagonism	**Cannibalization**
Google Ads × TikTok	Synergy	Antagonism	**Cannibalization**

Overall Insight:
- No synergies were detected in the data.
- Significant **antagonism** occurs when Google Ads runs in parallel with either TikTok or Facebook.
- **Facebook–TikTok** appears neutral, so they can be combined without major loss of efficiency.

5 Appendix

```r
## ----include=FALSE-------------------------------------------------
# install pacman if you don't have it yet
if (!require("pacman")) install.packages("pacman")

# load all packages at once
pacman::p_load(
  tidyverse,
  scales,
  ggplot2,
  dplyr,
  tinytex,
  kableExtra,
  broom
)

# Global knitr options for clean output
knitr::opts_chunk$set(
  echo = FALSE,
  message = FALSE,
  warning = FALSE,
  fig.align = "center",
  fig.width = 7,      # ~ full text width
  fig.height = 4.2,   # comfortable aspect
  out.width = "92%"
)

# A consistent, readable ggplot theme
theme_mm <- theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold"),
    plot.subtitle = element_text(margin = margin(b = 8)),
    plot.caption = element_text(size = 9, color = "grey30"),
    axis.title.y = element_text(margin = margin(r = 6)),
    axis.title.x = element_text(margin = margin(t = 6)),
    panel.grid.minor = element_blank(),
    legend.position = "none",
    strip.text = element_text(face = "bold")
  )
theme_set(theme_mm)



## ----include=FALSE-------------------------------------------------
# Read data
```

```r
stopifnot(file.exists("marketing_mix_assignment.csv"))
data <- readr::read_csv("marketing_mix_assignment.csv", show_col_types = FALSE)

# Standardize names
data <- data %>%
  rename(
    google_ads = `Google Ads`,
    facebook   = Facebook,
    tiktok     = TikTok,
    sales      = Sales
  )

# Parse Date ("mm/dd/yyyy" like "09/10/2025")
data <- data %>% mutate(Date = as.Date(Date, format = "%m/%d/%Y"))

# Quick quality checks
missing_counts <- colSums(is.na(dplyr::select(data, google_ads, facebook, tiktok, sales)))
neg_counts <- sapply(dplyr::select(data, google_ads, facebook, tiktok, sales), function(x) sum

tibble::tibble(
  Metric = c("Missing values", "Negative values"),
  Google_Ads = c(missing_counts[["google_ads"]], neg_counts[["google_ads"]]),
  Facebook   = c(missing_counts[["facebook"]],   neg_counts[["facebook"]]),
  TikTok     = c(missing_counts[["tiktok"]],     neg_counts[["tiktok"]]),
  Sales      = c(missing_counts[["sales"]],      neg_counts[["sales"]])
) |>
  kable(caption = "Data Quality Checks") |>
  kable_styling(full_width = FALSE, position = "center")



## ------------------------------------------------------------------------------
# Desired facet order (top → bottom)
order_levels <- c("Sales", "Google Ads", "Facebook", "TikTok")

long <- data %>%
  select(Date, Sales = sales, `Google Ads` = google_ads,
         Facebook = facebook, TikTok = tiktok) %>%
  pivot_longer(-Date, names_to = "series", values_to = "value") %>%
  mutate(
    series = factor(series, levels = order_levels)  # puts Sales first
  )

ggplot(long, aes(Date, value, color = series)) +
  geom_line(linewidth = 0.9) +
  facet_wrap(~ series, ncol = 1, scales = "free_y") +
  scale_color_manual(values = c(
```

```r
    "Sales"     = "orange",
    "Google Ads" = "steelblue",
    "Facebook"  = "firebrick",
    "TikTok"    = "darkgreen"
  )) +
  labs(title = "Sales and Weekly Marketing Spend by Media", x = NULL, y = NULL) +
  scale_x_date(date_breaks = "6 months", date_labels = "%b %Y") +
  scale_y_continuous(labels = label_dollar(prefix = "$")) +
  theme_minimal(base_size = 12) +
  theme(panel.grid.minor = element_blank(),
        legend.position = "none")




## ---------------------------------------------------------------------------
# Build time series + log transforms
google_ads <- data$google_ads
facebook   <- data$facebook
tiktok     <- data$tiktok
sales      <- data$sales

ln_google_ads <- log(google_ads + 1)
ln_facebook   <- log(facebook   + 1)
ln_tiktok     <- log(tiktok     + 1)
ln_sales      <- log(sales      + 1)

# One-period lag of ln(Sales)
Lag1_ln_sales <- dplyr::lag(ln_sales, 1)

# Drop first row with NA lag
model_df <- tibble(
  ln_sales = ln_sales,
  Lag1_ln_sales,
  ln_google_ads,
  ln_facebook,
  ln_tiktok
) %>% drop_na()

# Fit model
regression1 <- lm(
  ln_sales ~ Lag1_ln_sales + ln_google_ads + ln_facebook + ln_tiktok,
  data = model_df
)

# Tidy summary table
reg1_tidy <- broom::tidy(regression1, conf.int = TRUE) %>%
```

```r
  mutate(term = recode(term,
    "(Intercept)"   = "Intercept",
    "Lag1_ln_sales" = "Lagged ln(Sales)",
    "ln_google_ads" = "ln(Google Ads)",
    "ln_facebook"   = "ln(Facebook)",
    "ln_tiktok"     = "ln(TikTok)"
  ))

reg1_glance <- broom::glance(regression1)

reg1_tidy %>%
  mutate(
    estimate = round(estimate, 3),
    conf.low = round(conf.low, 3),
    conf.high= round(conf.high, 3),
    p.value  = scales::pvalue(p.value)
  ) %>%
  select(Term = term, Estimate = estimate, `95% CI Low` = conf.low, `95% CI High` = conf.high,
  kable(caption = "Model 1: Log-Log Regression with Carry-Over") %>%
  kable_styling(full_width = FALSE, position = "center")




## ----include=FALSE------------------------------------------------------
fitted_vals <- fitted(regression1)
actual_vals <- model_df$ln_sales

tibble(
  Index = seq_along(actual_vals),
  `Actual ln(Sales)` = actual_vals,
  `Fitted ln(Sales)` = fitted_vals
) %>%
  pivot_longer(-Index, names_to = "Series", values_to = "Value") %>%
  ggplot(aes(Index, Value, colour = Series, linetype = Series)) +
  geom_line(linewidth = 0.9) +
  labs(x = "Observation", y = "Log Sales") +
  theme(legend.position = "top")




## ----include=FALSE------------------------------------------------------
# Pull betas
beta_google_ads <- coef(regression1)[["ln_google_ads"]]
beta_facebook   <- coef(regression1)[["ln_facebook"]]
beta_tiktok     <- coef(regression1)[["ln_tiktok"]]
```

```r
# Baselines
avg_sales    <- mean(sales, na.rm = TRUE)
avg_google   <- mean(google_ads, na.rm = TRUE)
avg_facebook <- mean(facebook,   na.rm = TRUE)
avg_tiktok   <- mean(tiktok,     na.rm = TRUE)

# Unit effects (theta)
theta_google <- beta_google_ads * (avg_sales / avg_google)
theta_fb     <- beta_facebook   * (avg_sales / avg_facebook)
theta_ttk    <- beta_tiktok     * (avg_sales / avg_tiktok)

# Spend totals
google_spend   <- sum(google_ads, na.rm = TRUE)
facebook_spend <- sum(facebook,   na.rm = TRUE)
tiktok_spend   <- sum(tiktok,     na.rm = TRUE)

# Contributions
contrib_google <- theta_google * google_spend
contrib_fb     <- theta_fb     * facebook_spend
contrib_ttk    <- theta_ttk    * tiktok_spend

# ROI (Sales per unit spend)
roi_df <- tibble(
  Channel = c("Google Ads", "Facebook", "TikTok"),
  ROI     = c(contrib_google / google_spend,
              contrib_fb     / facebook_spend,
              contrib_ttk    / tiktok_spend)
) %>%
  mutate(ROI = round(ROI, 2))

roi_df %>%
  kable(caption = "Sales Return on Marketing Investment (SROMI) by Channel") %>%
  kable_styling(full_width = FALSE, position = "center")



## ---------------------------------------------------------------------------
roi_df %>%
  ggplot(aes(reorder(Channel, ROI), ROI)) +
  geom_col(color = "black", fill = c("steelblue", "firebrick", "darkgreen")[rank(roi_df$ROI)])
  geom_text(aes(label = ROI), vjust = -0.3, size = 3.6) +
  labs(
    x = NULL,
    y = NULL,  # remove y-axis label
    title = "Sales Return on Marketing Investment"  # put it on top
  ) +
  coord_cartesian(ylim = c(0, max(roi_df$ROI) * 1.2)) +
```

```r
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold"), # center & bold
    axis.text.x = element_text(face = "bold")
  )




## -------------------------------------------------------------------------
# Actual spend distribution
actual_spend <- tibble(
  Channel = c("Google Ads", "Facebook", "TikTok"),
  Spend   = c(google_spend, facebook_spend, tiktok_spend)
) %>% mutate(Share = Spend / sum(Spend),
        Type  = "Current")

# Elasticity-based optimal share (proportional to betas)
beta_sum <- beta_google_ads + beta_facebook + beta_tiktok
opt_share <- tibble(
  Channel = c("Google Ads", "Facebook", "TikTok"),
  Share   = c(beta_google_ads, beta_facebook, beta_tiktok) / beta_sum,
  Type = "Optimal"
)

bind_rows(
  actual_spend %>% transmute(Channel, Type = "Current", Share),
  opt_share    %>% transmute(Channel, Type = "Optimal", Share)
) %>%
  mutate(SharePct = scales::percent(Share)) %>%
  select(-Share) %>%                            # drop numeric Share
  pivot_wider(id_cols = Channel,                # ensure 1 row per Channel
            names_from = Type, values_from = SharePct) %>%
  kable(caption = "Current vs. Model-Recommended Budget Shares") %>%
  kable_styling(full_width = FALSE, position = "center")




## -------------------------------------------------------------------------
plot_df <- bind_rows(actual_spend, opt_share) %>%
  mutate(
    Channel = factor(Channel, levels = c("Facebook", "Google Ads", "TikTok")),
    Type    = factor(Type, levels = c("Current", "Optimal"))
  )

ggplot(plot_df, aes(x = Channel, y = Share, fill = Type)) +
  geom_col(position = position_dodge(width = 0.8), width = 0.7) +
```

```r
    scale_y_continuous(labels = scales::percent_format(accuracy = 1),
                       expand = c(0, 0)) +
    scale_fill_manual(values = c("Current" = "#1f77b4", "Optimal" = "#ff7f0e")) +
    labs( title = "Marketing budget: Current vs optimal allocation", x = "Channel", y = "Allocat
    coord_cartesian(ylim = c(0, max(plot_df$Share) * 1.2)) +   # ~20% headroom
    theme_minimal() +
    theme(
      plot.title = element_text(hjust = 0.5, face = "bold"),    # centered & bold
      axis.text.x = element_text(face = "bold"),
      legend.position = "right"
    )




## --------------------------------------------------------------------------
regression2 <- lm(
  ln_sales ~ Lag1_ln_sales + ln_google_ads + ln_facebook + ln_tiktok +
    ln_google_ads:ln_facebook +
    ln_google_ads:ln_tiktok +
    ln_facebook:ln_tiktok,
  data = model_df
)

reg2_tidy <- broom::tidy(regression2) %>%
  mutate(term = recode(term,
    "(Intercept)"                       = "Intercept",
    "Lag1_ln_sales"                     = "Lagged ln(Sales)",
    "ln_google_ads"                     = "ln(Google Ads)",
    "ln_facebook"                       = "ln(Facebook)",
    "ln_tiktok"                         = "ln(TikTok)",
    "ln_google_ads:ln_facebook"         = "ln(Google) × ln(Facebook)",
    "ln_google_ads:ln_tiktok"           = "ln(Google) × ln(TikTok)",
    "ln_facebook:ln_tiktok"             = "ln(Facebook) × ln(TikTok)"
  ))

reg2_tidy %>%
  mutate(estimate = round(estimate, 4),
         p.value  = scales::pvalue(p.value)) %>%
  select(Term = term, Estimate = estimate, `p-value` = p.value) %>%
  kable(caption = "Interaction Terms (Synergies/Antagonisms)") %>%
  kable_styling(full_width = FALSE, position = "center")



## ----echo=FALSE, results='asis'-----------------------------------------
# Extract the R code from the current Rmd
```

```r
code_file <- knitr::purl("Assignment1.rmd", quiet = TRUE)
cat("```r\n")
cat(readLines(code_file), sep = "\n")
cat("\n```")
```