R Workshop: Marketing Mix Modelling (MMM)

Dr Ashutosh Singh

March 2024

Contents

Content with * is optional.

You will need the following packages for this session:

- "dplyr"
- "readxl"

Marketing Mix Modelling, also known as market response modelling, uses historical marketing and sales data in statistical models to measure sales "bang" for marketing "buck". MMM is causal modelling in which researchers attempt to explain or predict market share or sales volume from marketing inputs while controlling for other sales drivers.

1 Marketing Effectiveness and Resource Allocation

Leverage *historical data* to quantify the effectiveness of marketing actions.

Load the data

```r
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
spending.data <- read_xls("advertising spending 1.xls")
```

Inspecting the data

```r
str(spending.data)
```

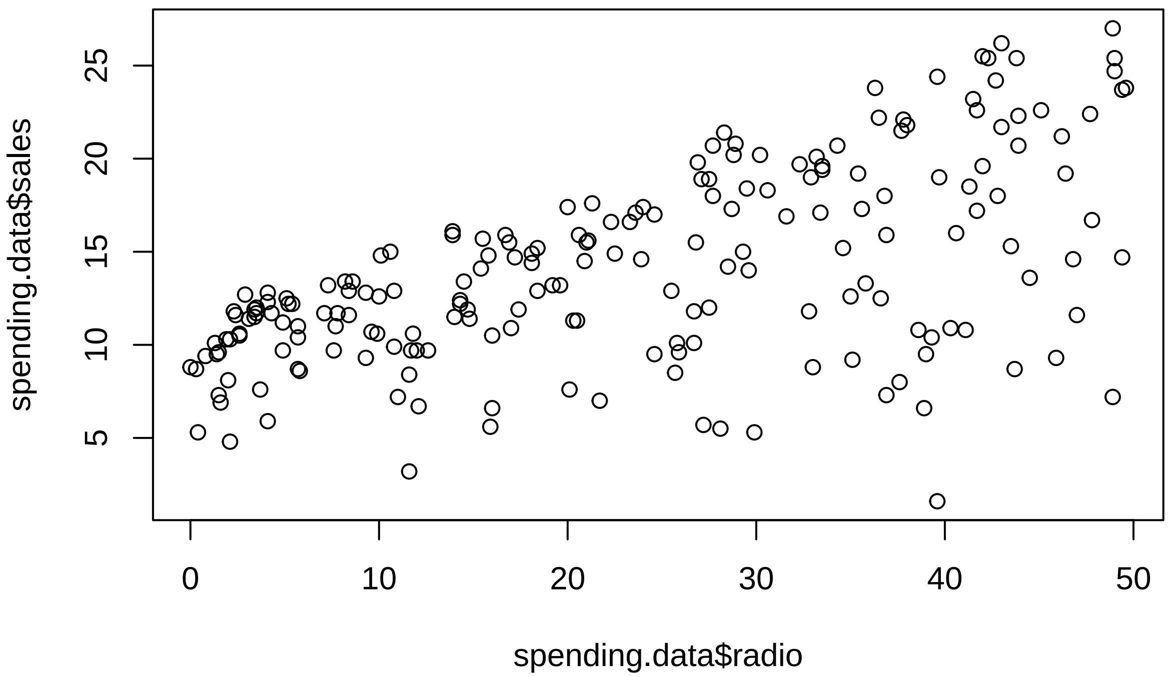
```
## tibble [200 x 5] (S3: tbl_df/tbl/data.frame)
##  $ week     : num [1:200] 1 2 3 4 5 6 7 8 9 10 ...
##  $ tv       : num [1:200] 230.1 44.5 17.2 151.5 180.8 ...
##  $ radio    : num [1:200] 37.8 39.3 45.9 41.3 10.8 48.9 32.8 19.6 2.1 2.6 ...
##  $ newspaper: num [1:200] 69.2 45.1 69.3 58.5 58.4 75 23.5 11.6 1 21.2 ...
##  $ sales    : num [1:200] 22.1 10.4 9.3 18.5 12.9 7.2 11.8 13.2 4.8 10.6 ...
```

Data contains weekly sales and tv, radio, and newspaper advertising (in thousands of pounds)

1.1 Does radio advertising affect sales for an electronics brand?

Visualizing the data

```r
# scatter plot
plot(spending.data$radio, spending.data$sales)
```




```r
# Correlation between radio advertising and sales
cor(spending.data$radio, spending.data$sales)
```

```
## [1] 0.5762226
```

Simple linear regression

```r
regression <- lm(sales ~ radio, data = spending.data)
summary(regression)
```

```
##
## Call:
## lm(formula = sales ~ radio, data = spending.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.7305  -2.1324   0.7707   2.7775   8.1810
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.31164    0.56290  16.542   <2e-16 ***
## radio        0.20250    0.02041   9.921   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.275 on 198 degrees of freedom
## Multiple R-squared:  0.332,  Adjusted R-squared:  0.3287
## F-statistic: 98.42 on 1 and 198 DF,  p-value: < 2.2e-16
```

Question: How much will I sell if I spend 40,000 on radio ads? Answer: 17.31 = 9.31 +0.20 * 40

Let us practice #1!

How does the brand's TV advertising affect its sales? Run a simple linear regression with sales as the dependent variable and TV advertising spending as the only independent variable.

- How many units does the brand sell if it does not spend anything on TV advertising?
- If the brand increases TV advertising by £1,000, how much incremental sales will this generate?
- If the brand spends £150,000 on an advertising campaign, how many units can the brand expect to sell?

Solution

1.2 Accounting for multiple predictors: multiple linear regression

```
regression <- lm(sales ~ radio + tv, data=spending.data)
summary(regression)
```

```
##
## Call:
## lm(formula = sales ~ radio + tv, data = spending.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7977 -0.8752  0.2422  1.1708  2.8328
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.92110    0.29449   9.919   <2e-16 ***
## radio        0.18799    0.00804  23.382   <2e-16 ***
## tv           0.04575    0.00139  32.909   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.681 on 197 degrees of freedom
## Multiple R-squared:  0.8972, Adjusted R-squared:  0.8962
## F-statistic: 859.6 on 2 and 197 DF,  p-value: < 2.2e-16
```

Let us practice #2!

Run a multiple regression with sales as the dependent variable and TV, radio, and newspaper advertising spending as the independent variables.

What do you conclude about the relationship between newspaper advertising and sales?

Solutions

1.3 Allocating marketing budgets

1.3.1 Ratio of elasticities method

What is elasticity?

- % change in the response variable for a 1% change in the predictor variable
- Example: the % change in sales for a 1% change in advertising spending

** Let us start with an example:**

- Image that you have a £100,000 budget to spend on advertising
- A 1% increase in online advertising increases sales by 0.12%

- A 1% increase in offline advertising increases sales by 0.08%

** How to use the ratio of elasticity method? 1. Sum elasticity: 0.12 + 0.08 = 0.20. 2. Ratio of elasticity:

- Online: 0.12/0.20 = 60%
- Offline: 0.08/0.20 = 40%

3. Recommendation: Allocate 60% (£60,000) to online advertising and 40% (£40,000) to offline advertising.

1.3.2 How to obtain elasticities from a linear regression model?

- Advertising Elasticity = Advertising Estimate * (Baseline Advertising/Baseline Sales)
- Baseline Advertising = average radio advertising spending (*=23.26*)

```
mean(spending.data$radio)
```

```
## [1] 23.264
```

- Baseline Sales = average sales (*=14.02*)

```
mean(spending.data$sales)
```

```
## [1] 14.0225
```

- What is the elasticity of radio advertising? (**0.32**)

```
0.19 * (23.26/14.02)
```

```
## [1] 0.3152211
```

- A 1% increase in radio advertising results in a 0.32% increase in sales.

2 Marketing Mix Modelling

2.1 Modelling non-linear returns on investment

let us assume that an ad ran for an initial budget of £500,000 over 1 week, with the aim of improving the perception among UK consumers that Airbnb is an inclusive brand, but also to increase traffic to Airbnb's websites and bookings.

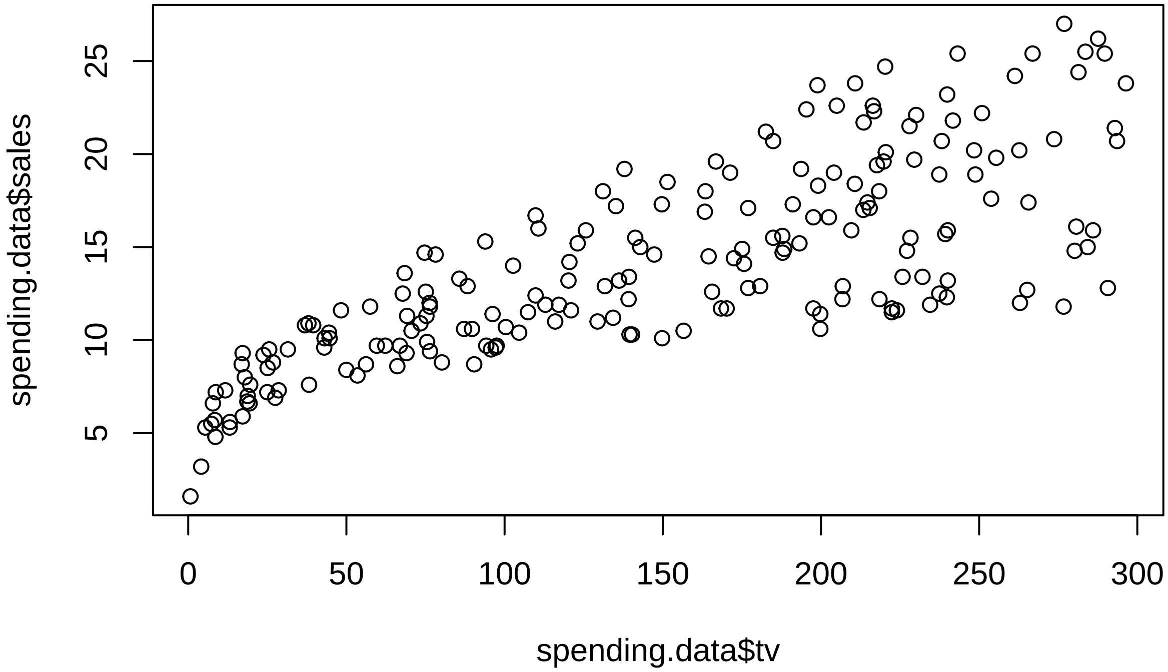
As a result of the campaign, Airbnb sees an increase in bookings of 1%.

A year later, Airbnb decides to run a comparable ad, again for 1 week, but to double the spending to £1,000,000, in the hope that bookings will now increase by 2%.

Do you think this will happen? Why (not)?

** Is the relation between TV and advertising linear?"

```
plot(spending.data$tv, spending.data$sales)
```




```r
summary(spending.data$sales)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.60   10.38   12.90   14.02   17.40   27.00
```

```r
summary(spending.data$tv)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.70   74.38  149.75  147.04  218.82  296.40
```

```r
summary(spending.data$radio)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   9.975  22.900  23.264  36.525  49.600
```

```r
summary(spending.data$newspaper)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.30   12.75   25.75   30.55   45.10  114.00
```

```r
regression <- lm(log(sales) ~ log(radio+0.01) + log(tv) + log(newspaper),
                 data=spending.data)
summary(regression)
```

```
##
## Call:
## lm(formula = log(sales) ~ log(radio + 0.01) + log(tv) + log(newspaper),
##     data = spending.data)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45346 -0.08881 -0.01746  0.06781  0.78863
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.479773   0.052031   9.221   <2e-16 ***
## log(radio + 0.01) 0.144177   0.007865  18.333   <2e-16 ***
## log(tv)           0.349297   0.008857  39.437   <2e-16 ***
## log(newspaper)    0.017488   0.009306   1.879   0.0617 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1254 on 196 degrees of freedom
## Multiple R-squared:  0.9098, Adjusted R-squared:  0.9084
## F-statistic: 658.7 on 3 and 196 DF,  p-value: < 2.2e-16
```

We have added 0.01 to radio, as the min of radio is 0, which returns invalid result of log (radio).

Which model is better, linear or log-log?

- Plot the relationship between variables
- Compare the fit of a linear and log-log model

How to interpret coefficients?

- A 1% increase in radio advertising results in a 0.14% increase in sales
- A 1% increase in TV advertising results in a 0.35% increase in sales

Coefficients are elasticities!

Summary

Linear

- Equation: $Y = \beta_0 + \beta_1 X$
- Interpretation: One unit change in X leads to $beta_1$ unit change in Y
- When to use? Linear relation between X and Y

Log-Log

- Equation: $log(Y) = \beta_0 + \beta_1 log(X)$
- Interpretation: One percent change in X leads to β_1 percent change in Y
- When to use? A non-linear relation between X and Y

Which one to use?

- Plot the data to learn about the relation between X and Y.
- Estimate both models and identify the best fitting model

2.2 Modelling media synergy

- The combined use of marketing mix instruments can create **synergy effects**
- **Synergy effect** arises when the joint impact of multiple media exceeds the total of their individual parts.

Is TV advertising more effective in the presence of radio advertising?

```r
regression <- lm(log(sales) ~ log(radio+0.01) + log(tv) + log(newspaper)
                 + log(radio+0.01)*log(tv), data=spending.data)

summary(regression)
```

```
##
## Call:
## lm(formula = log(sales) ~ log(radio + 0.01) + log(tv) + log(newspaper) +
##     log(radio + 0.01) * log(tv), data = spending.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29117 -0.06889 -0.02084  0.05787  0.74453
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                1.347498   0.101811  13.235  < 2e-16 ***
## log(radio + 0.01)         -0.152661   0.032200  -4.741 4.10e-06 ***
## log(tv)                    0.153799   0.022032   6.981 4.49e-11 ***
## log(newspaper)             0.019947   0.007741   2.577   0.0107 *
## log(radio + 0.01):log(tv)  0.066311   0.007043   9.415  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1043 on 195 degrees of freedom
## Multiple R-squared:  0.938,  Adjusted R-squared:  0.9367
## F-statistic:   737 on 4 and 195 DF,  p-value: < 2.2e-16
```

Use centered variables

```r
center <- function(x) {
          scale(x, scale = F)
} # scale = F, means only center not standardize
```

```r
spending.data <- spending.data %>%
  mutate(radio_log_centered = center(log(radio+0.01)),
         tv_log_centered = center(log(tv)),
         newspaper_log_centered = center(log(newspaper)))

regression <- lm(log(sales) ~ radio_log_centered +
                   tv_log_centered + newspaper_log_centered +
                   radio_log_centered*tv_log_centered,
                 data=spending.data)
summary(regression)
```

```
##
## Call:
## lm(formula = log(sales) ~ radio_log_centered + tv_log_centered +
##     newspaper_log_centered + radio_log_centered * tv_log_centered,
##     data = spending.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29117 -0.06889 -0.02084  0.05787  0.74453
##
```

```
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      2.561748   0.007376 347.306   <2e-16 ***
## radio_log_centered               0.157146   0.006681  23.521   <2e-16 ***
## tv_log_centered                  0.337076   0.007476  45.086   <2e-16 ***
## newspaper_log_centered           0.019947   0.007741   2.577   0.0107 *
## radio_log_centered:tv_log_centered 0.066311   0.007043   9.415   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1043 on 195 degrees of freedom
## Multiple R-squared:  0.938,  Adjusted R-squared:  0.9367
## F-statistic:   737 on 4 and 195 DF,  p-value: < 2.2e-16
```

Let us practice #3! Run a log-log model with sales as the dependent variable and TV, radio, and newspaper advertising spending as the independent variables. In addition to allowing for synergy effects between radio and tv advertising, test for synergy effects between radio and newspaper advertising.

Do you find any evidence of a synergy effect between radio and newspaper advertising?

Solution:

2.3 Modelling carryover effects

So far, we assumed that advertising in a given time period would only affect sales in that same time period. In reality, **consumers' responses to advertising can be delayed**. Not accounting for carryover effects can cause advertising elasticities to be under-valued.

Advertising adstock measures the effect of advertising beyond the current time period.

The theory behind adstock is that marketing exposures build awareness in consumers' minds. That awareness does not disappear right after the consumers see the ad but rather remains in their memory. Memory decays over the weeks and hence the decaying portion of adstock.

$$Adstock_t = Advertising_t + \lambda Adstock_{t-1}$$

In each time period, you are assumed to retain a fraction (λ) of your previous advertising stock.

For instance, if λ equals 0.3, then adstock from one time period ago still has a 30% effect in the current time period.

** How to do this in R?**

```
adstock <- function(x, rate){
  return(as.numeric(stats::filter(x=x, filter=rate, method="recursive")))
}
# filter() function from stats package applies linear filtering to a univariate time series.
```

- "rate = 0.1" sets λ to 0.1 (i.e. 10%)
- You can empirically test multiple values of λ.

```
spending.data <- spending.data %>%
  mutate(tv_adstock = adstock(tv,0.1),
         newspaper_adstock = adstock(newspaper, 0.1),
         radio_adstock = adstock(radio, 0.1))
```

```
regression <- lm(log(sales) ~ log(radio+0.01) + log(tv_adstock) +
                   log(newspaper), data=spending.data)
summary(regression)
```

```
##
## Call:
## lm(formula = log(sales) ~ log(radio + 0.01) + log(tv_adstock) +
##     log(newspaper), data = spending.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.96605 -0.08394 -0.00081  0.07966  0.81860
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.09547    0.08418  -1.134   0.2581
## log(radio + 0.01)  0.13177    0.01008  13.074   <2e-16 ***
## log(tv_adstock)    0.45582    0.01542  29.567   <2e-16 ***
## log(newspaper)     0.02219    0.01190   1.865   0.0637 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1604 on 196 degrees of freedom
## Multiple R-squared:  0.8523, Adjusted R-squared:  0.8501
## F-statistic: 377.1 on 3 and 196 DF,  p-value: < 2.2e-16
```

Let us practice #4!

Run a log-log model with sales as the dependent variables and TV advertising stock, radio advertising stock, and newspaper advertising stock as independent variables.

Start by setting λ to 0.1, and then adjust it to 0.3 and 0.5.

Do you find any evidence for the carryover effects of advertising?

2.4 Predictive accuracy *

```
# Total number of rows in the data frame
n <- nrow(spending.data)

# Number of rows for the training set (80% of the dataset)
n_train <- round(0.80 * n)

# Training data
spending.data.train <- subset(spending.data, week <= n_train)

# Holdout data
spending.data.holdout <- subset(spending.data, week > n_train)

# Estimation on training data
regression <- lm(log(sales) ~ log(radio_adstock) + log(tv_adstock) +
                   log(newspaper_adstock), data=spending.data.train)
summary(regression)
```

```
##
## Call:
## lm(formula = log(sales) ~ log(radio_adstock) + log(tv_adstock) +
##     log(newspaper_adstock), data = spending.data.train)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.95544 -0.06206 -0.00140  0.07359  0.60782
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -0.37264    0.09377  -3.974 0.000108 ***
## log(radio_adstock)       0.19215    0.01473  13.048  < 2e-16 ***
## log(tv_adstock)          0.47022    0.01566  30.031  < 2e-16 ***
## log(newspaper_adstock)   0.02057    0.01545   1.332 0.184852
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1483 on 156 degrees of freedom
## Multiple R-squared:  0.8812, Adjusted R-squared:  0.8789
## F-statistic: 385.6 on 3 and 156 DF,  p-value: < 2.2e-16
```

```r
# Predict sales on holdout data
spending.data.holdout$predicted_sales_log <-
  predict(object =regression, newdata = spending.data.holdout)

# Convert predicted log sales to actual sales
spending.data.holdout$predicted_sales <-
  exp(spending.data.holdout$predicted_sales_log)

# Quantify predictive accuracy: Mean Average Percentage Error (MAPE)
mape <- mean(abs((spending.data.holdout$sales
                  -spending.data.holdout$predicted_sales)
                 /spending.data.holdout$sales))
mape # Reflects the average percentage error in a given week
```

```
## [1] 0.1010304
```

```r
# Plot actual versus predicted sales
plot(spending.data.holdout$week, spending.data.holdout$sales,
     type="l", col="blue") # Plot actual sales

lines(spending.data.holdout$week,spending.data.holdout$predicted_sales,
      type = "l", col = "red") # Add predicted sales

legend("topleft", legend=c("Actual sales", "Predicted sales"),
       col=c("blue", "red"), lty = 1:2, cex=0.6)
```

Actual sales
Predicted sales
spending.data.holdout$sales
spending.data.holdout$week
10
15
20
25
160
170
180
190
200