Effectiveness of Advertising Activities on Sales

Group 21

Members:
Aishwarya Prashant Kamat
Hanish Singla
Haridhakshini (Harisha) Subramoniapillai Ajeetha
Meng-Wei (Vivian) Wu

Table of Contents

Reallocation of Adv Spend required on Newspaper & Portal

In today's age, firms typically engage in quantifying the contribution of their advertising activities and redistributing resources across marketing activities in real time[1]. Given the vast advertising spend, it is reasonable for firms to be interested in understanding the impact of these activities on sales to optimize the spend for maximum revenue. Our client, a cosmetics firm, also attempts to determine the effectiveness of advertising on sales of a 4-year old product using the data on advertising spend over the last 42 months.

Using the data, a regression model has been developed to understand the variables impacting sales. The model accounts for diminishing returns of regressor variables and impact that previous sales will have on predicted sales (regressand var). The model suggests that for Offline adv. spend - Catalogs sent to all customers including existing, winback and new; and for Online adv. spend - only Newsletter and Portals significantly impact sales. Based on this model, we can infer that Newspaper, Portals and Win back Customers catalogs have the highest effectiveness and elasticity values. Therefore, the firm can strategize the spend across these three categories for maximum revenue. One caveat, this model treats advertising touch points in isolation, which is not entirely accurate as in today's market, customer purchase decisions are an outcome of an interaction between various marketing activities. We also found that transformations like log and model without intercept significantly increase the explained variation percentage so if these models do not violate any business principle, switching to these models is recommended.

Introduction

Cosmetic firms must determine which advertising activities are most effective, not only to increase profits but also improve the fame of their products. With the detailed data that parse product sales and advertising metrics, analytics can help establish the correlation between Sales and Adv. Spend variables. Allocating the right model, measuring outcomes, and validating strategies can help firms make better budget allocation decisions. In the analysis below, multiple adv. spend variables are accounted to quantify the precise combination of the most effective advertising activities. This will give insights to the company about how an advertising activity or a set of specific marketing channels will help drive the most sales.

Problem Formulation

In marketing, knowing precisely how all parts of a marketing campaign collectively drive sales and what happens when you adjust them [1] is essential, and today feasible as well. Consumers are exposed to an expanding, fragmented array of marketing touch points across media and sales channels. Similarly the stated cosmetics firm is using both Online and Offline media channels to target customers. The flowchart helps us understand the relationships among all the variables Appendix I (Figure.1).

The data is structured over 42 months, we have a Sales column and three aggregate ad spend columns - Total Advertising Spend; two columns it's composed of - Total Offline spend and Total Online spend. We have a total of 8 adv. activities, individually recorded under 10 columns - Catalogs, Mailings under Offline spend; and Banner, Search, Social Media, Newsletter, Retargeting and Portals ads spend under Online spend. The ad spend for Catalogs is further distributed across 3 types of Customers - Existing, Winback (customers who have not shopped in 6 months) and New.

Data Description

From the raw data, we have 13 independent variables which are shown in the flowchart above (Figure.1) and Sales as a dependent variable is the sales of items in units of month.

First, we do not take into consideration the aggregate columns - Total Ad Spend, Total Online Spend and Total Offline Spend as our regressor variables since they will be impacted by change in other variables. Second, we exclude Social Media and Banner fields as they have almost no values, which makes them less useful at helping us draw conclusions. From the boxplot in Appendix I (Figure.2), we observe that Sales has many outliers, while the rest of the variables do not. We calculate the descriptive statistics on every variable of interest to identify the distribution of this specific variable and check if we have sufficient data to run the linear regression model.

We notice that the count for the variables ExistCustomer, NewCustomer, Newsletter, Portals and Winback are the maximum. The maximum ad spend on existing customers is $567.60 and on new customers is $1131.57. Ad spend on portals has the least range and variation. Ad spend on a new customer and an existing customer has a negative correlation with Sales, while that on newsletter, portals and win back are positively correlated with Sales in Appendix I (Table.1).

Model Development, Estimation, and Result

To estimate the sales based on these variables, we would like to talk about three important factors impacting model estimation:

1. Diminishing returns: As we know advertising provides diminishing returns on sales, therefore we use a square root transformation on all advertising amount variables to capture this effect.
2. Carry over Effect: Sales in a week say X is not only dependent upon the advertising in that week, but also has a carry over effect from previous weeks. For e.g. I came across the product I bought today through an advertisement in the previous week. So, to capture this effect, we introduce a lambda (Carry over effect) times one period of lagged sales.
3. Principle of Parsimony: If a particular variable doesn't add much value (criteria based on R-squared, Adjusted R-squared, AIC, BIC values) to the overall significance of the model, then we choose the reduced model (with fewer variables). Based on this, we removed Mailings, Retargeting and Search variables. All the comparisons have been added in the Appendix - IV

So, our marketing mix model allocation $Y_{it} \sim f(X_{1t}, X_{2t}, \dots X_{nt})$ can be estimated by:

$$Y_t = \beta_0 + \lambda Y_{t-1} + \beta_1 Z_{1t} + \beta_2 Z_{2t} + \beta_3 Z_{3t} + \beta_4 Z_{4t} + \beta_5 Z_{5t} + \epsilon_t$$

where Z_{it} is $\sqrt{Advertising\ amount}$ on i^{th} channel, Y_{t-1} is the lagged sales value, and ϵ_t is residuals.

Our focal model only considers five variables corresponding to Advertising amount on Existing Customer Catalogs, New Customer Catalogs, Winback Catalogs, Newsletter, Portals

The above model yields the following results and beta coefficient values for each channel (Table.2):

R squared value	*Adjusted R squared Value*	*AIC value*	*BIC value*
38.64	27.82	659.02	672.73

Table.2 - The result of the final focal model

Elasticity is calculated by the equation below:

$$\eta = \frac{\triangle Y_t\%}{\triangle X_{it}\%} = \frac{\beta_i}{2} \times \frac{\sqrt{X_i}}{(1-\lambda)Y}$$

where β_i denotes the correlation coefficients of each variable in the focal model, Y_i is the mean of dependent variables, and λ measures the carryover effect of the past sales result. The elasticity of all the variables in the final model is displayed below (Table.3).

	Intercept	ExistCust	NewCust	WinbackCust	Newsletter	Portals
Beta Estimates	2076.497	-23.923	-27.319	54.329	164.717	806.982
P-Val	0.060	0.149	0.058	0.035	0.210	0.009
Lambda	0.156	0.156	0.156	0.156	0.156	0.156
Mean of Y	4809.254	4809.254	4809.254	4809.254	4809.254	4809.254
Mean of X	0.000	567.600	272.870	83.424	20.734	5.246
Elasticity	0.000	-0.070	-0.056	0.061	0.092	0.228

Table.3 - The elasticity of each variable

Some of the key points we can infer from the above results are:

- P-values for only Portals, Winback Customer Catalogs, New Customer Catalogs and Intercept are statistically significant at 10% significance level.
- We see Portals and Newsletters have the highest beta coefficients implying they are the most effective ad channels while the beta coefficients for Existing Customer catalogs and New Customer catalogs are negative. We also see maximum elasticity in Newsletter, Portals as well as Win back Customers catalogs.

Recommendation and Managerial Implications

From the focal model, we learnt that we should reduce the investments in Catalogs of existing customers and new customers as their elasticities are negative, while we should keep investing in Newsletter, Portals and Catalogs of winback customers as they have positive and maximum elasticities.

Based on the 39% R-squared value, we can say that the current model does not help explain a lot of variation in Sales owing to budget allocation among these channels. Therefore, we have explored some other options for explaining this variation with other transformations (Log) and tweaks (with/without intercept, lagged variables as well as interaction). Based on the results in Appendix II we inferred that model without intercept improved the model significantly whereas there isnt much diff. in log vs square root models.

We also considered the synergy effect among each pair of variables and found that there are several advertising synergy effects. The result is shown in Appendix III (Table.7 and Figure.4). Based on the model, we can conclude that advertising synergy exists and may affect the sales amount and that $Z_{1t}*Z_{5t}$ and $Z_{4t}*Z_{7t}$ are statistically significant at the 10% significance level. Therefore, we should recalculate the elasticities in this model and reallocate the budget as some variables may affect the others.

Conclusion

After performing this model with the existing data, it suggests a causal relationship between all of the marketing activities affecting part of the sales in each month with significant impact from newsletter and portals. Additionally, we reach two main conclusions mentioned below:

1. For limitations, we found that it is possible to build a model with a relatively low R-squared but with a higher AIC, so we should determine which models to use based on the business perspectives rather than the mathematical values only. In addition, we could collect the data for a longer period of time (instead of 42 months). More the data, the more accurate the model. Additionally, we don’t have any data for banners and social media channels currently, so their significance was not accounted for in this analysis.
2. For the extensions of the current model, we should also consider several factors related to the market environment in the cosmetics industry, the economy around the globe, and the likelihood and speed with which the competition will be able to copy your product [3]. All these factors could affect the model so much and lead to completely different results, which may end up having a different marketing strategy.

References

[1] https://www.hindawi.com/journals/cin/2021/9863155/

[2] https://www.scribbr.com/statistics/akaike-information-criterion/

[3] https://hbr.org/1975/05/when-where-and-how-to-test-market

Appendix I

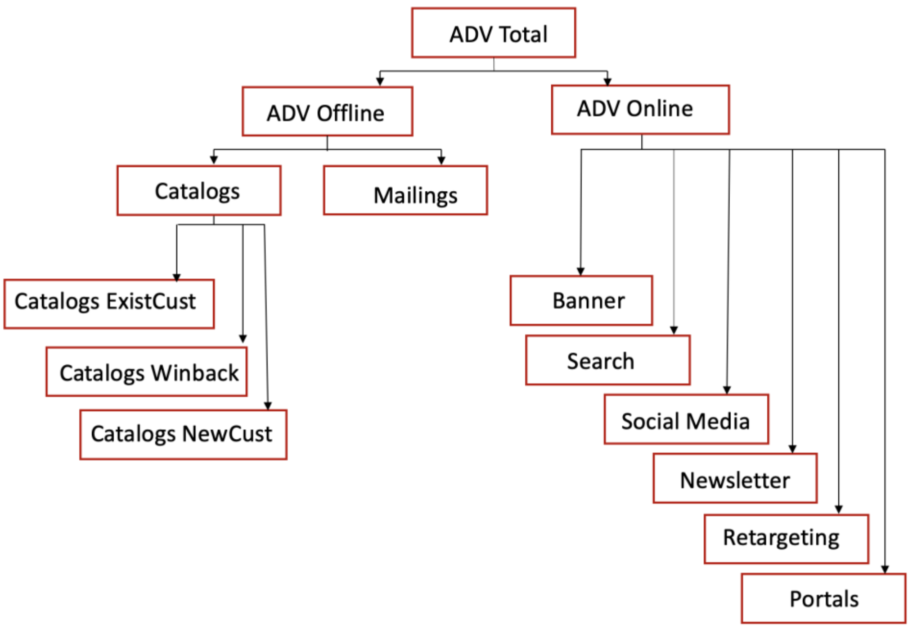

Figure.1 - The flowchart of all marketing activities

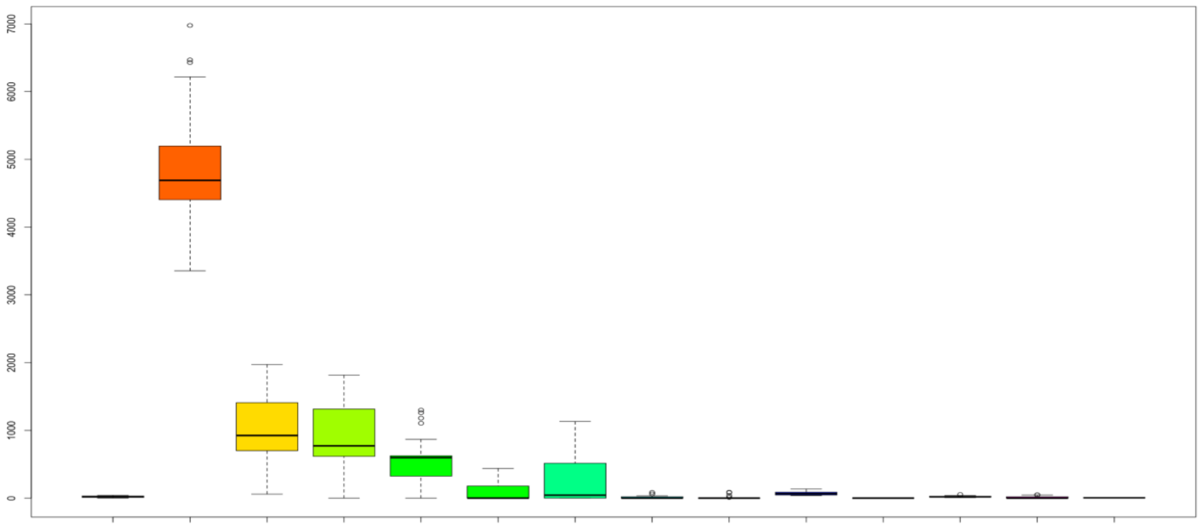

Figure.2 - The boxplots of all marketing activities

	ExistCustomer	NewCustomer	Newsletter	Portals	Winback
Count	39.00	21.00	42.00	42.00	16.00
Mean	567.60	272.87	20.73	5.25	83.42
Q1.25%	328.69	0.00	16.69	3.39	0.00
Median	597.96	43.63	19.78	4.71	0.00
Q3.75%	625.64	487.42	25.14	6.87	174.15
Min	0.00	0.00	7.06	2.54	0.00
Max	1298.69	1131.57	53.61	9.30	438.54
Range	1298.69	1131.57	46.55	6.76	438.54
Variance	87398.82	116458.86	72.70	3.83	16045.01
StdDev	295.63	341.26	8.53	1.96	126.67
CV_in_Percent	52.08	125.06	41.12	37.31	151.84
Skewness	0.44	0.99	1.34	0.43	1.23
Kurtosis	0.46	-0.10	3.48	-1.23	0.25
Correlation_with_Sales	-0.01	-0.04	0.06	0.49	0.03

Table.1 - Descriptive statistics on the most prominent variables.

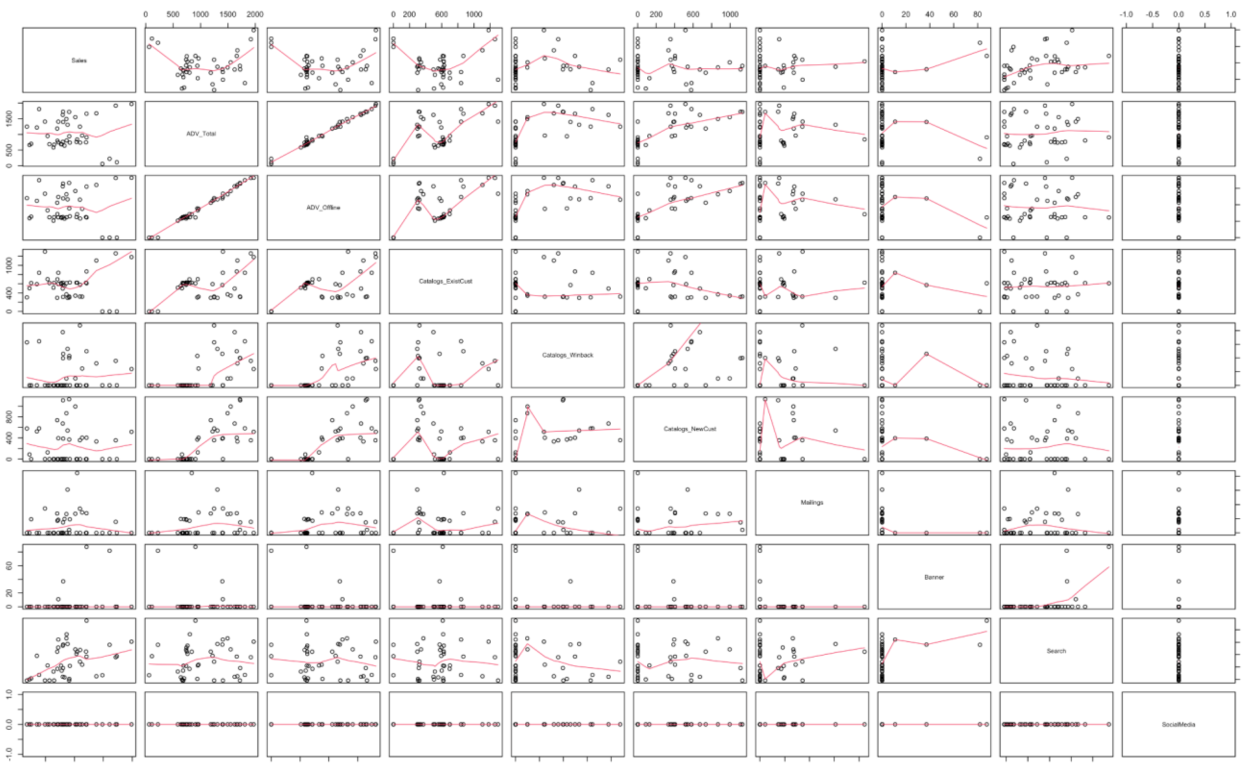

Figure.3 - The scatter plots among all variables

	Search	Banner	Mailings	Retargeting
Count	42.00	4.00	17.00	17.00
Mean	69.83	5.18	11.42	10.85
Q1.25%	45.38	0.00	0.00	0.00
Median	66.11	0.00	0.00	0.00
Q3.75%	88.19	0.00	19.24	18.56
Min	38.17	0.00	0.00	0.00
Max	134.87	87.61	84.47	49.30
Range	96.71	87.61	84.47	49.30
Variance	636.24	359.89	335.48	259.01
StdDev	25.22	18.97	18.32	16.09
CV_in_Percent	36.12	366.27	160.33	148.37
Skewness	0.38	3.62	2.02	1.27
Kurtosis	-0.81	12.06	4.59	0.31
Correlation_with_Sales	0.40	0.25	0.02	0.18

Table.4 - Descriptive statistics for unchosen variables

Appendix II

With the same independent variables in the final focal model, we use the natural log of the same advertising variables and include/exclude the intercept/lagged variable to see what the difference is. The result is shown below:

1. Use square root for diminishing returns (The focal model)

	With intercept and lagged variable	With Intercept but without lagged variable	Without Intercept but with lagged variable	Without Intercept and lagged variable
p-value	0.007335	0.002505	<2.2e-16	<2.2e-16
R-Squared	0.3876	0.3881	0.9842	0.9778
Adj. R-Squared	0.2795	0.3032	0.9814	0.9748
AIC	658.9511	674.6298	657.2686	684.5423
BIC	672.6596	686.7935	669.2636	694.9683

Table.5 - The statistics of the models (transform variables by square root)

2. Use natural log for diminishing returns

	With intercept and lagged variable	With Intercept but without lagged variable	Without Intercept but with lagged variable	Without Intercept and lagged variable
p-value	0.007335	0.002505	<2.2e-16	<2.2e-16
R-Squared	0.3876	0.3881	0.9842	0.9778
Adj. R-Squared	0.2795	0.3032	0.9814	0.9748
AIC	658.9511	674.6298	657.2686	684.5423
BIC	672.6596	686.7935	669.2636	694.9683

Table.6 - The statistics of the models (transform variables by natural log)

After comparing all these above models, we found that as we make a transformation on the variables, no matter square root or natural log, the results are the same. However, if we remove the intercepts or lagged variables, we will get completely different results. From the tables above, we can see that the R-squared increases largely if we remove the intercept, which means that the model is more statistically significant but does not mean that the data points are fitted better in the model. We should instead consider this issue on the business side, not only the mathematical side.

Appendix III

We start by adding the first pair Z_{1t}*Z_{2t} , running the model and comparing it to the original focal model. The AIC decreases largely with an increase in R-squared, which means that this new model works better. We continue to add on different pairs of variables and leave the synergy effect if the adjusted model has a larger R-squared but smaller AIC. Finally, the final focal model with the synergy effects of Z_{1t}*Z_{2t} , Z_{1t}*Z_{5t} , Z_{1t}*Z_{7t} , Z_{2t}*Z_{4t} , and Z_{4t}*Z_{7t} has the largest R-squared and the smallest AIC.

	The final focal model	With Z1 * Z2	With Z1 * Z2 and Z1 * Z4	With Z1 * Z2 and Z1 * Z5	With Z1 * Z2 and Z1 * Z5 and Z1 * Z7	With Z1 * Z2 and Z1 * Z5 and Z1 * Z7 and Z2 * Z4	With Z1 * Z2 and Z1 * Z5 and Z1 * Z7 and Z2 * Z4 and Z4 * Z7
p-value	0.007335	5.20E-05	0.0001433	6.64E-05	3.26E-05	3.56E-05	1.31E-05
R-Squared	0.3876	0.5898	0.5901	0.6117	0.6569	0.6798	0.7257
Adj. R-Squared	0.2795	0.5028	0.4876	0.5146	0.5573	0.5731	0.6217
AIC	658.9511	644.5202	646.4891	644.27	641.194	640.3591	636.0167
BIC	672.6596	659.9423	663.6249	661.4057	660.433	660.922	658.2931

Table.7 - The result of models with synergy effects

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 7557.3991  1512.0513   4.998 2.55e-05 ***
SqM1        -161.7776    53.0486  -3.050  0.00486 ** 
SqM2         -18.1899    87.7882  -0.207  0.83730    
SqM7        -307.1741   123.2171  -2.493  0.01863 *  
SqM4          86.1499   120.9831   0.712  0.48210    
SqM5        -104.8732   586.4601  -0.179  0.85932    
Stm1          -0.2335     0.1730  -1.350  0.18760    
SqM1:SqM2      2.4856     2.5426   0.978  0.33637    
SqM1:SqM7      3.1638     3.7886   0.835  0.41050    
SqM1:SqM5     37.5008    22.8267   1.643  0.11121    
SqM2:SqM4    -11.1182    16.3209  -0.681  0.50113    
SqM7:SqM4     58.8920    26.5345   2.219  0.03444 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 489.1 on 29 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.7262,	Adjusted R-squared:  0.6224 
F-statistic: 6.994 on 11 and 29 DF,  p-value: 1.279e-05
```

Figure.4 - *The summary of the final model with Synergy Effect*

Appendix IV

Comparison with model including all the advertising channel variables.

Reduced Model: Only includes 5 advertising variables (Existing Customer Catalogs, New Customer Catalogs, Winback Catalogs, Newsletter, Portals)

$$Y_t = \beta_0 + \lambda Y_{t-1} + \beta_1 Z_{1t} + \beta_2 Z_{2t} + \beta_3 Z_{3t} + \beta_4 Z_{4t} + \beta_5 Z_{5t} + \epsilon_t$$

Full Model: Includes all the advertising channel variables (Except Social Media and Banner)

$$Y_t = \beta_0 + \lambda Y_{t-1} + \beta_1 Z_{1t} + \beta_2 Z_{2t} + \beta_3 Z_{3t} + \beta_4 Z_{4t} + \beta_5 Z_{5t} + \beta_6 Z_{6t} + \beta_7 Z_{7t} + \beta_8 Z_{8t} + \beta_9 Z_{9t} + \epsilon_t$$

Comparison:

Figure.5 - Reduced Model Results

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2076.4968  1066.0810   1.948  0.05974 .
SqM1         -23.9232    16.2031  -1.476  0.14902
SqM2         -27.3190    13.9015  -1.965  0.05761 .
SqM7          54.3291    24.7737   2.193  0.03524 *
SqM4         164.7168   128.9270   1.278  0.21005
SqM5         806.9819   291.6213   2.767  0.00908 **
Stm1           0.1563     0.1833   0.853  0.39957
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 676.2 on 34 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.3864,    Adjusted R-squared:  0.2782
F-statistic: 3.569 on 6 and 34 DF,  p-value: 0.00753
```

Figure.6 - Full Model Results

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 1560.11816 1305.48737   1.195   0.2414
SqM1         -25.42388   17.45955  -1.456   0.1557
SqM2         -25.29200   15.65407  -1.616   0.1166
SqM3         147.32006  222.56314   0.662   0.5131
SqM4         125.69160  145.20000   0.866   0.3936
SqM5         866.98152  631.84484   1.372   0.1802
SqM6         -17.72565   46.26717  -0.383   0.7043
SqM7          49.71340   28.51097   1.744   0.0915 .
SqM8         -86.53916   95.53703  -0.906   0.3723
SqM9         -13.94797   62.94786  -0.222   0.8261
Stm1           0.07296    0.21486   0.340   0.7365
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 707.9 on 30 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.4065,    Adjusted R-squared:  0.2087
F-statistic: 2.055 on 10 and 30 DF,  p-value: 0.06231
```

Summary parameters:

#	R-Squared	Adjusted R-Squared	Overall P-Value	AIC	BIC
Reduced	38.64	27.82	.007	659	672
Full	40.65	20.87	.06	665	686

Table.8 - Comparisons of the parameters between the reduced model and the full model

Comparison using ANOVA:

```
Analysis of Variance Table

Model 1: multdata$Sales ~ SqM1 + SqM2 + SqM7 + SqM4 + SqM5 + Stm1
Model 2: multdata$Sales ~ SqM1 + SqM2 + SqM3 + SqM4 + SqM5 + SqM6 + SqM7 +
    SqM8 + SqM9 + Stm1
  Res.Df      RSS Df Sum of Sq      F Pr(>F)
1     34 15545155
2     30 15035695  4    509460 0.2541 0.9049
```

Figure.7 - ANOVA table of the reduced model and the full model

All these indicate that the full model does not significantly increase the model therefore by principle of parsimony we go forward with the reduced model.