

## GAME EMAIL MARKETING CASE

Pooja Bera(ETU20201756)

**1.What is the effect of opening the email on Logged Money Expenditure After? (Hint: Run a regression WITHOUT control variables and report the results.)**

After running the regression, we can see the effect of opening the email on Logged Money expenditure increases by 30% if there is 1 unit increase in opening the email. The p value is 0.045 which is less than the 0.05 significant value. Adjusted R square is 0.002 which states the proposition of the variability of the logged money over the opening of emails.

### Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.045 <sup>a</sup>	.002	.002	2.258762327

a. Predictors: (Constant), Open

b. Dependent Variable: Logged.MoneyA

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.807	.054		14.971	.000
	Open	.309	.154	.045	2.006	.045

a. Dependent Variable: Logged.MoneyA

**2. What control variables would you use and why? What is the effect of opening the email on Logged Money Expenditure After when you control for the relevant control variables? (Hint: Run a regression WITH control variables and report the results.)**

I have used the forward stepwise model to detect the following accuracy of 32.5%. The control variables I choose is Logged.MoneyBPD, ActiveDaysB ,Logged.ActivityA, and Lag. I choose these variables as they were significant and less coefficients of standard error.The R ,R square and Adjusted R square increased after adding the control variable which means that the variables added gave a better result. Hence ,makes it easier for us to predict the “Logged Money Expenditure After” in a more significant way.

With the help of unstandardized coefficients of B and std. error ,Standardised coefficients beta, t and significant value we derive some results. The model predicts that adding these control variables we see an increase /decrease depending on the variable reflecting in the “Logged Money Expenditure After”. For every one unit increase in Logged.Money BPD ,Logged ActivityA and lag we see it increases our “Logged Money Expenditure After” by 1.1%,21% and 1.8% respectively.Since, ActiveDaysB has standard error 0 and decreases the effect on the result by 0.1% we can consider it as it has significant value 0.017 less than 0.05

## Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.424 <sup>a</sup>	.180	.178	2.049133881

## Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.096	.189		-5.794	.000
	Logged.MoneyBPD	.011	.001	.324	15.884	.000
	ActiveDaysB	-.001	.000	-.052	-2.382	.017
	Logged.ActivityA	.210	.019	.353	10.965	.000
	Lag	.018	.004	.130	4.201	.000

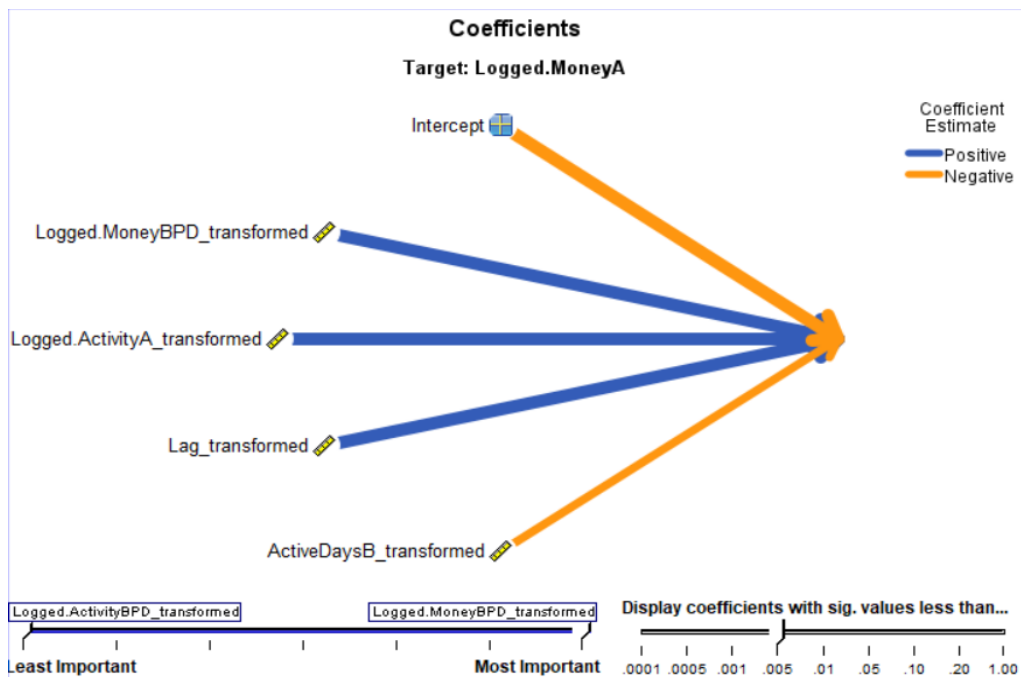
## ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1837.435	4	459.359	109.398	.000 <sup>b</sup>
	Residual	8376.905	1995	4.199		
	Total	10214.340	1999			

a. Dependent Variable: Logged.MoneyA

b. Predictors: (Constant), Lag, Logged.MoneyBPD, ActiveDaysB, Logged.ActivityA

With the F value =109.398 we see that null hypothesis is rejected and there is no explanatory power. The significant value is 0.00 which is a good indicator as well.

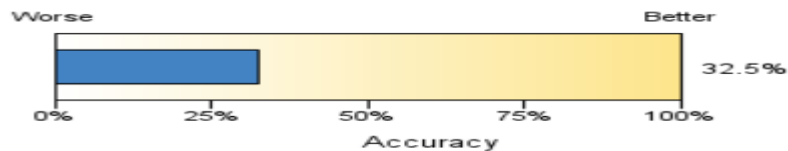


The least to most important can be seen above whose p value is less than 0.05.

### Model Summary

<b>Target</b>	Logged.MoneyA
<b>Automatic Data Preparation</b>	On
<b>Model Selection Method</b>	Forward Stepwise
<b>Information Criterion</b>	2,483.609

The information criterion is used to compare to models. Models with smaller information criterion values fit better.



A checkmark means the effect is in the model at this step using the forward stepwise.

	Step						
	1	2	3	4	5	6	7
<b>Information Criterion</b>	2,639.910	2,511.234	2,499.479	2,493.278	2,489.466	2,484.780	2,483.609
<b>Logged.MoneyBPD_transformed</b>	✓	✓	✓	✓	✓	✓	✓
<b>Logged.ActivityA_transformed</b>		✓	✓	✓	✓	✓	✓
<b>Lag_transformed</b>			✓	✓	✓	✓	✓
<b>Effect ActiveDaysB_transformed</b>				✓	✓	✓	✓
<b>ID_transformed</b>					✓	✓	✓
<b>Recency_transformed</b>						✓	✓
<b>Logged.ActivityBPD_transformed</b>							✓

### 3. What happens to the adjusted R2 after adding the control variables and why?

The adjusted R2 with the Control variables was 0.002 and adding the control variables was 0.178. The difference is huge 0.176 with respect to this given case as it effects our control variable which is associated with money and turnover of the company. Since we see an increase this means this results in a better model. The control variables are useful and make the “Logged Money Expenditure After” into a profit. As we know with every 1 unit of change in the Logged.Money BPD, Logged ActivityA and lag having fixed the other control variables fixed we get an increase of 1.1%, 21% and 1.8% respectively. Logged activity a gives the best results to the dependent variable.

### 4. Why is the treatment effect (i.e., the effect of opening the email on Logged Money Expenditure After) different with and without the control variables? (Hint: Think about what type of users are more likely to open the email and how controlling for relevant variables might alleviate this problem.)

The treatment effect with and without the control variables is that the users who spent money per day before the email, the amount of time user spent in the period after the email,

the number of days the user spent in the app before the email as well as email date and lag which is number of days between the email date and first day user logs in the app after the email . A user who satisfies these conditions holding the other values fixed gives a good turnover in the money expenditure. If we focus on the four control variables set, we can get loggedMoneyA.

**5. Basically, what the managers are doing is looking only at email receivers and comparing those who opened the email and those who did not. How do you evaluate this strategy? What better strategy would you recommend to the managers to investigate email effectiveness.**

Better strategy always depends on the need of the company .What they are sending and what is their expectation. We can investigate the email effectiveness by improving the quality of the email by personalization according to the customer as it will show a sense of connectivity rather than sending a generalised format to the whole globe. If we dig deeper into emailing data to seek what kind of email is more interesting like loyalty benefit , coupons ,newsletters and promotions. We can also use google analytics to monitor the bounce rate of the emails that means the email website link is suspicious and not reaching lot of users .