

## CUSTOMER ANALYSIS: ASSIGNMENT 1

The objective of the assignment is to find what drives sales and which platform of advertising should we spend our money on. First, running a multiple regression will give a better understanding of the relationship between the variables.

Multiple Regression with sales as dependent variable

<i>Regression Statistics</i>	
Multiple R	0.752045947
R Square	0.565573107
Adjusted R Square	0.407599691
Standard Error	112.0340306
Observations	31

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	8	359496.4867	44937.06083	3.580179012	0.008328622
Residual	22	276135.7282	12551.62401		
Total	30	635632.2148			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
<b>Intercept</b>	148.80	114.17	1.30	0.21	-87.96	385.57	-87.96	385.57
<b>Offline - Print</b>	-0.33	0.74	-0.44	0.66	-1.85	1.20	-1.85	1.20
<b>Offline - TV</b>	1.66	0.64	2.58	0.02	0.33	3.00	0.33	3.00
<b>Online - Ad Network</b>	0.43	0.77	0.56	0.58	-1.17	2.04	-1.17	2.04
<b>Online - Google SEO</b>	1.46	0.71	2.04	0.05	-0.02	2.94	-0.02	2.94
<b>Online - Paid Search</b>	1.28	0.78	1.65	0.11	-0.33	2.89	-0.33	2.89
<b>Online - Facebook</b>	2.86	0.90	3.16	0.00	0.98	4.73	0.98	4.73
<b>Online - Twitter</b>	-0.07	0.77	-0.09	0.93	-1.68	1.54	-1.68	1.54
<b>Online - Instagram</b>	0.29	0.71	0.41	0.68	-1.18	1.77	-1.18	1.77

- The F-test significance is less than 0.05 indicating that the regression model is significant at 95% confidence level.
- The R<sup>2</sup> reveals that 56.6% of the variability in Sales is explained by the independent variables.
- The regression equation for this model is **Sales = 148.8 + 0.32 Print + 1.66 TV + 0.43 Ad Network + 1.46 Google SEO + 1.28 Paid Search + 2.85 Facebook + 0.07 Twitter + 0.3 Instagram**
- Only the TV ads among the offline ads and Facebook among the online ads are significant at 95%.
- At 90% confidence level, the Google SEO ads are significant as well.

The following plots show the trend of sales for each additional dollar spent towards a particular advertising campaign.

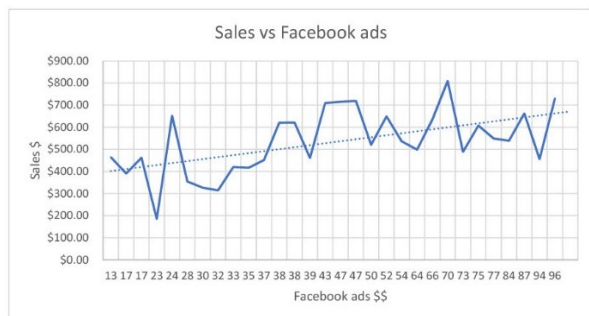


Figure 1

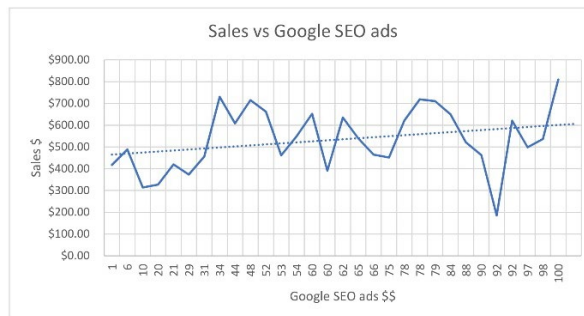


Figure 2

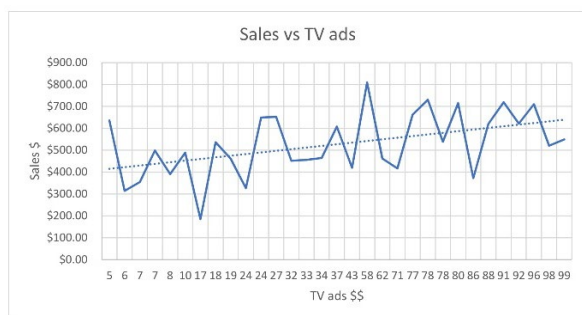


Figure 3

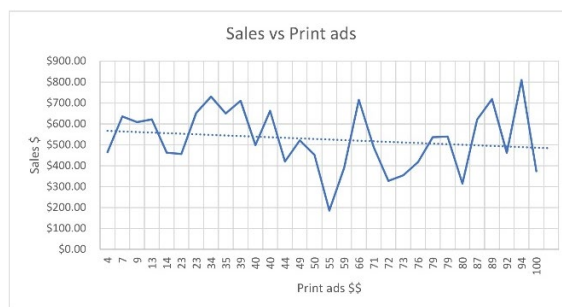


Figure 4

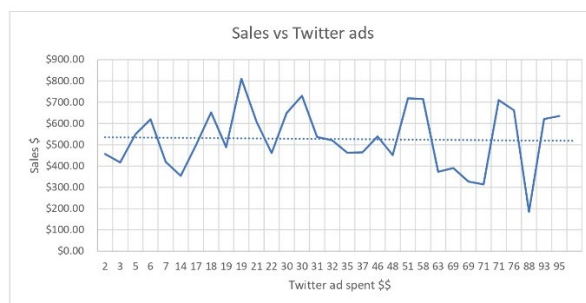


Figure 5

- From figures 1,2 and 3 it is evident that the Facebook ads, TV ads and the Google SEO ads have an increasing trend line indicating that with the increase in the money spent on these ads, the sales are increasing. The p-value based inferences in the previous page are also supportive of the same.
- On the contrary, with increase in money spent on Print ads, the sales are decreasing. Though it might not indicate that print ads are causing bad publicity, we can conclude that it is not significantly influencing sales. It would be advisable to not spend on print ads.

Some additional inferences:

- Data shows that TV, Facebook and Google SEO are the most successful marketing vehicles for the product. Given that Print ads and relatively newer social media platforms such as Instagram/Twitter are not yielding sales as much, it wouldn't be too far fetched to say that the product's target audience are either Gen Y or early Gen Z (80s-90s).
- The Twitter ads have a slightly decreasing trend line (*Figure 5*) which can be analyzed further with a larger dataset. Data spread over a few months would help arrive at a conclusion on whether advertising on this platform is beneficial.