

First we work to understand the structure and context of the data. We do this with exploratory data analysis. The purpose is to prime our intuition and hone the reporting development. Though we deploy quantitative testing procedures this is not a statistically rigorous stage of the process. Second, based on what we learn we build the performance dashboard. These visualizations are included at the end of this document but the raw data and Tableau source files are also included with this submission. Lastly, we outline key findings below.

Exploratory Data Analysis (EDA)

1. Understand the features
 - a. In Table 1 below, features shown in blue are added to the existing data set

Feature	Note
Date Code	Unique ID for week/year transaction to build time series
Marketing Channel Category	
Marketing Channel Type	
Marketing Channel Name	
Marketing Channel Category CODE	Assign integer 1-2
Marketing Channel Type CODE	Assign integer 1-4
Marketing Channel Name CODE	Assign integer 1-25
Week Number	
Year ID	
New Customers	
Existing Customers	Subtract new customers from total customers
Total Customers	
New/Existing Ratio	New customers divided by existing customers
Order Count	
Demand Sales	
Marketing Spend	
Order Multiple	Order count divided by total customers
Cost Per Click	Marketing spend divided by total customers
Cost Per Order	Marketing Spend divided by order count
Net Profit Margin	Sales less spend divided by sales

Table 1

2. Correlations

- a. Quick check to see if features are linearly related and possibly repetitive. The marketing channel categorical features are excluded. We normalize numerical features by z-score and compute the correlation matrix in Figure 1. Features are boxed informally as n related, cost related, and efficiency related.
- b. **Findings.** Dark blues and reds show the stronger linear relationships. There is repetition in the n related features, they are also highly correlated to spend and sales. Margin is obviously strongly inversely related to costs. For the next test we drop the repetitive features highlighted gray.

	New Customers	Existing Customers	Total Customers	New/Existing	Order Count	Marketing Spend	Cost Per Click	Cost Per Order	Order Multiple	Demand Sales	Net Profit Margin
New Customers	1										
Existing Customers	0.89	1									
Total Customers	0.97	0.98	1								
New/Existing	0.02	-0.14	-0.07	1							
Order Count	0.96	0.98	1.00	-0.07	1						
Marketing Spend	0.85	0.67	0.77	0.07	0.76	1					
Cost Per Click	-0.05	-0.10	-0.08	0.10	-0.08	0.03	1				
Cost Per Order	-0.05	-0.10	-0.08	0.10	-0.08	0.03	1.00	1			
Order Multiple	0.03	0.05	0.04	-0.01	0.04	0.00	-0.01	-0.03	1		
Demand Sales	0.94	0.98	0.99	-0.09	0.99	0.72	-0.09	-0.09	0.05	1	
Net Profit Margin	0.06	0.11	0.09	-0.08	0.09	0.00	-0.92	-0.92	0.09	0.10	1

Figure 1

3. Principal Component Analysis (PCA)
 - a. Here we conduct an unsupervised learning test. We have eight features. We alter their basis to n-orthogonal space and evaluate the distribution of variance.
 - b. Figure 2 below shows an elbow plot of the variance explained by each principle component. Note the kink at PC3 and its similarity to PC4.
 - c. **Findings.** Note the factor loadings in Table 2 below. All eight features indicate material influence across the first three components, as highlighted in blue. If a feature showed limited influence (a low factor score) in one of the first three components I would drop it but as it stands, we will keep these eight features for the next test.

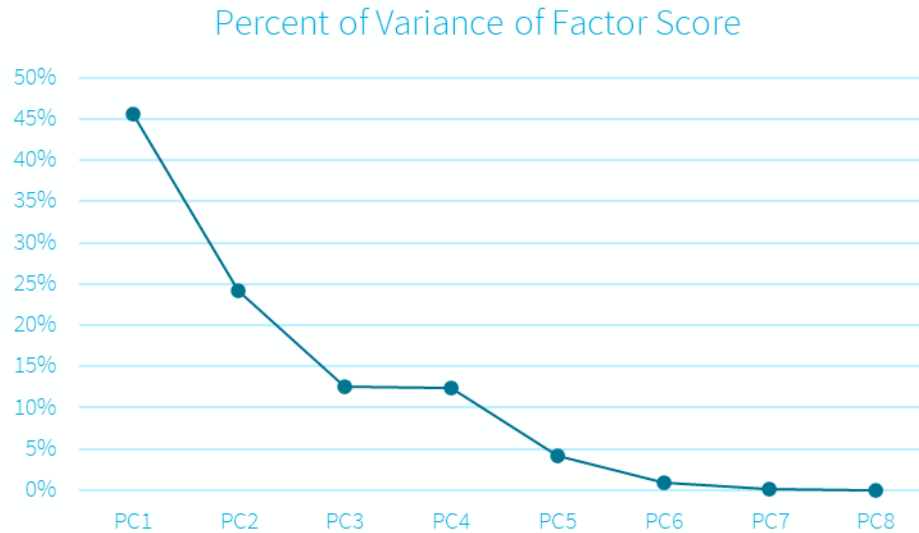


Figure 2

Eigenvectors		PC1	PC2	PC3
Total Customers		0.5177	-0.0432	-0.0143
New/Existng		-0.0384	-0.1404	0.8058
Order Count		0.5172	-0.0418	-0.0185
Marketing Spend		0.4344	-0.1282	0.1493
Cost Per Order		-0.0737	-0.6906	-0.1357
Order Multiple		0.0297	0.0849	-0.5419
Demand Sales		0.5109	-0.0322	-0.0406
Net Profit Margin		0.0832	0.6893	0.1187

Table 2

4. Cluster Analysis

- Here we deploy the k-means algorithm for another unsupervised learning test. The goal is to partition clusters with similar metrics across eight features and compute intra-cluster averages.
- We decide on the k-number of clusters based on Figure 3 below. The elbow plot shows a kink around three or four clusters. The metadata in cluster four shows a low intra-cluster n (< 30) so we proceed with three clusters.
- Figure 3 below summarize intra-cluster data across eight features. All figures are expressed in standard deviation from the feature mean, again because the data was standardized by z-score.
- Findings.** N related features obviously influence sales. Total customers and order count not only imply more sales, but they do so linearly, like correlations indicated. Marketing spend as it relates to total customers, order counts, and demand sales is somewhat mixed, which is quite interesting, and we will develop this shortly. More existing customers imply more sales and better margins. Cost per order of course affects margins

inversely and it appears to do so linearly. Order multiple, which had low correlations and the weakest PCA effect, seems to be somewhat irrelevant.

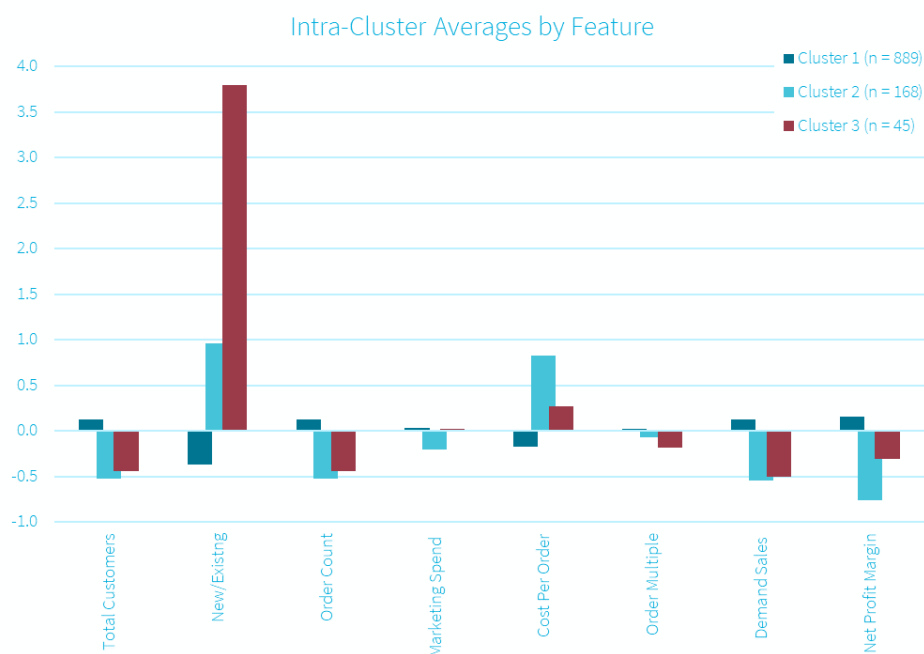


Figure 3

5. Regressions

- a. Lastly, we conduct supervised learning tests with some quick regressions. Here we try to determine key factors (dependent variable) as it relates to sales (independent variable). We use the original data set only (exclude the added ones used in previous tests) including the category codes for the three marketing channels. This last choice to include categorical data with numerical data is slightly hand-wavey, however, we are still just doing EDA.
- b. But here we need to make a key distinction. We know that in general more marketing spend leads to more sales, and this is shown below in Figure 4 (here there are large z-score outliers so we instead standardize by percentile rank). But not all line items in the original data have a marketing spend: all marketing spend entries are a paid marketing channel category, but not all paid marketing channel category entries include a marketing spend.
- c. Table 3 below shows linear and lasso regressions for two data sets: data with and without marketing spend. As such we drop the marketing channel category feature in the former (because they are all the same) and of course the marketing spend feature in the latter; dropped features are highlighted gray. Linear regression is typical least squares regression. Lasso regression is more complex but has the effect of simplifying the model by setting unimportant features to zero. There is a plethora of pitfalls associated with

both models, but again, this is EDA and we simply want to see if the model generalizes well to untrained data and if so the manner in which it weights the features.

d. **Findings:**

- i. Firstly, the scatter in Figure 4 shows not all marketing spend dollars are created equal. Some spends result in low sales and are thus expensive. While a notable cluster up and to the left suggest some marketing spends are better values.
- ii. For the regressions, the marketing spend models put the largest weight on order count, new customers, marketing spend, and the marketing channel name. Note that fewer new customers imply less sales, but also that less spend implies more sales. The former makes sense but the latter is in contradiction to the scatter below. The no marketing spend model puts the largest weight on total customers, order count, and new customers; all n related. Interesting that more new customers implies more sales in the linear model, which is opposite of the marketing spend linear model but to a much lesser extent.

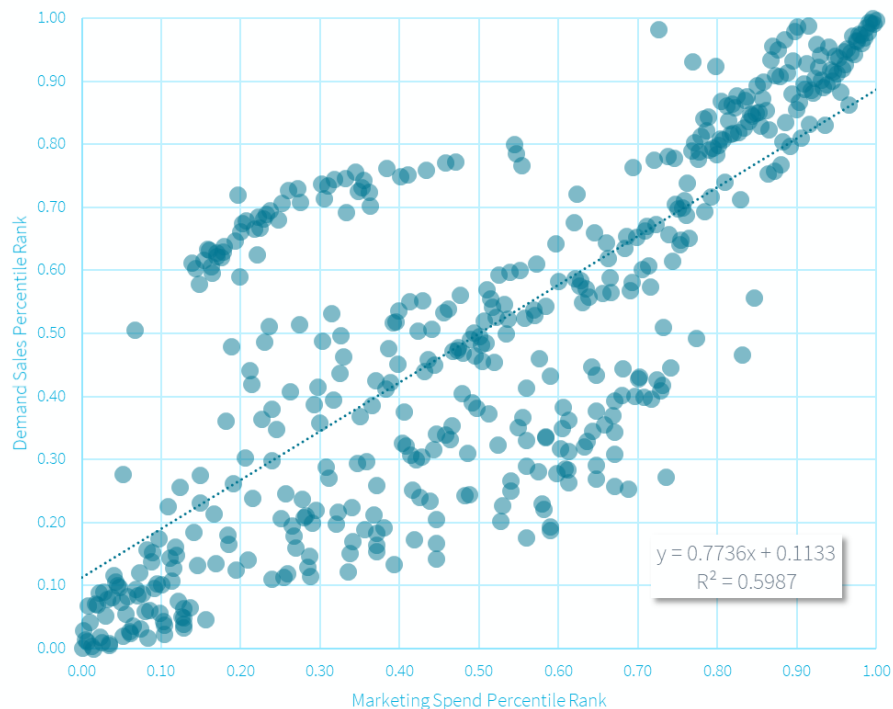


Figure 4

Marketing Spend				No Marketing Spend			
Linear Regression							
Train set MSE				0.0013			
Predicted MSE				0.0130			
intercept				0.0070			
Marketing Channel Type CODE				0.0269			
Marketing Channel Name CODE				-0.0301			
Week Number				0.0096			
New Customers				-0.2209			
Total Customers				0.0021			
Order Count				1.2986			
Marketing Spend				-0.0931			
Lasso Regression							
intercept				-0.4239			
Marketing Channel Type CODE				0.0000			
Marketing Channel Name CODE				-0.0740			
Week Number				0.0000			
New Customers				0.0000			
Total Customers				0.0000			
Order Count				0.0000			
Marketing Spend				0.0000			

Table 3

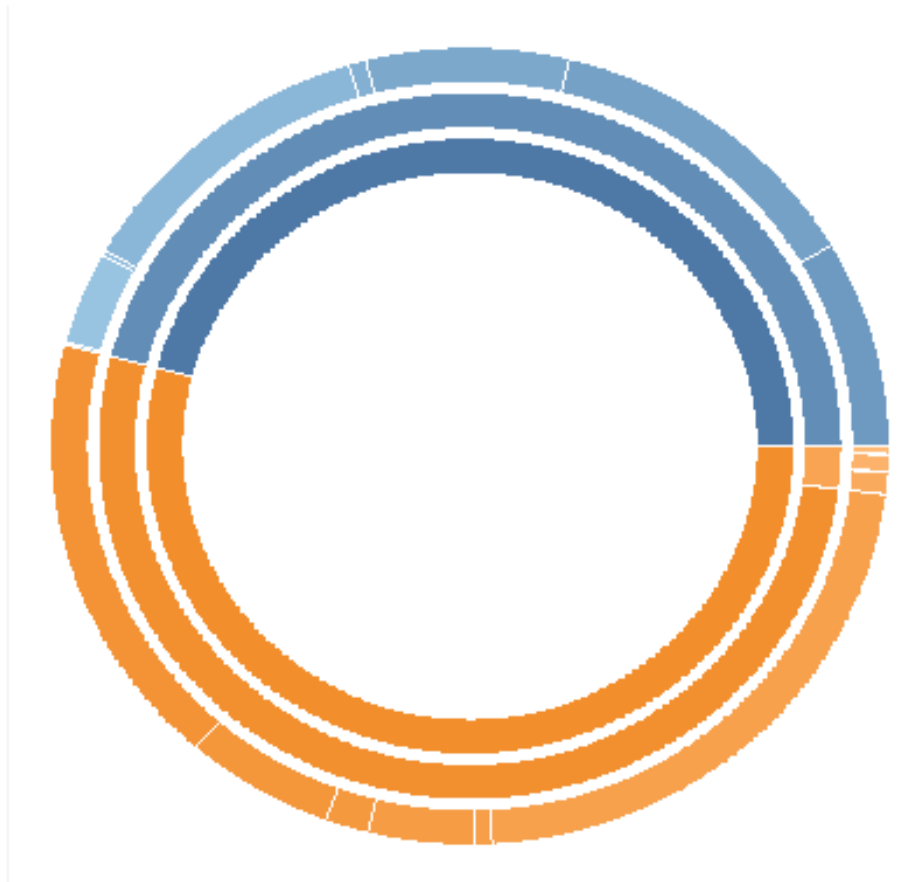
Build Tableau Dashboards

- Sources of sales (ecommerce_dash_1)
 - We show absolute dollar sales across the three marketing channels in a sunburst chart.
 - We also show these partitions in table form as a percentage of total sales.
- Sales trends (ecommerce_dash_2)
 - Here we show trailing twelve month sales figures, smoothed with a 12-week simple moving average, partitioned by sales quartiles, and also color coded to show paid or owned.
- Quality of sales (ecommerce_dash_3)
 - Here we partition across marketing spend. As we have shown these appear to be two different kinds of transaction. We then sort marketing channel name by percentage of sales. We then compare that number to four efficiency ratios. As informed by the above EDA these are orders per customer, sales per order, percentage existing customers, and profit margin.
 - Alongside we show percentile rank ratios in horizontal bar charts. The idea is we should be able to quickly identify inefficiencies with various sorting approaches.

Key Findings

1. Sales are somewhat evenly distributed across marketing category. There is minimal revenue coming from the paid brand type. Within the owned and legacy paid types there is high concentration at the channel name level. Four names account for 65% of sales: organic, direct, affiliate, and PLA.
2. For trends we see high seasonality across the top and into the third quartiles. These channels are also recently rolling over instead of ticking up as we enter what would be the start of the 2020 high sales season—possibly pandemic related. The middle quartiles appear to be more robust in the face of seasonality, though they account for far less revenue. These marketing channels are mostly paid and seem to be social media related marketing channels.
3. For quality of sales we bifurcate across marketing spend
 - a. For quality of sales that include a marketing spend note the spend per order and profit margin columns. When we sort by percentage of sales we would expect to see the bar charts descend in the same manner. The general aesthetic certainly suggests this but the branded paid search and display remarketing channels are the most efficient in terms of costs and margin, yet account for less sales. The sales per order bar chart also approximates this downwards shape; the most lucrative sales channels have the best dollar per order. The order per customer and percentage existing customer bar charts do not descend in magnitude. Here the top earners have more new customers and less orders per customer; whereas lower earners perhaps rely on existing customers and their tendency to submit more orders.
 - i. Key note here: Recall the above scatter in Figure 4 showing the best spend values up and to the left. When we isolate those transactions all of them come from the branded paid search channel. This is a top quartile earner and accounts for ~6% of the marketing spend cohort revenue. Note that the cost and margin bar charts identified this top-of-class efficiency.
 - b. For quality of sales that exclude a marketing spend we lose the cost and margin metrics. Just a general observation that these are mostly owned channels and are probably capex that we would see in the cash flows from investing section of the financial statements. Based on the aesthetic of the bar charts this cohort behaves like its peer. There are lower orders per customer in the higher earners. If we exclude the high group order figure, sales per order is higher in the highest earners. Reliance on existing customers favors the top earners to the same extent, but there are stronger existing customer percentages in lower earners.
4. In summary, costs and margins can be improved at the top of the marketing spend cohort but in general the descending nature of efficiencies alongside sales works well. The same could be said about sales per order in both marketing spend cohorts: the highest earners have the better sales per order figures. Also, for both marketing spend cohorts, the percentage of existing customers seems to help the higher earners (those with a high n) and the lower earners (also with a high n) but is more irrelevant in the middle earners. That is, the average percentage of existing customers across all channel names is ~60% and we see this in the high n earners. But it is above average in lower earners, as perhaps these channels with low n are more dependent on existing customers (and also lower sample size). Lastly, for orders per customer, there is limited variance across all

channels. The muddled aesthetic of this metric, plus what we said about this in EDA when we were calling it “order multiple” suggests it is perhaps void of insight.

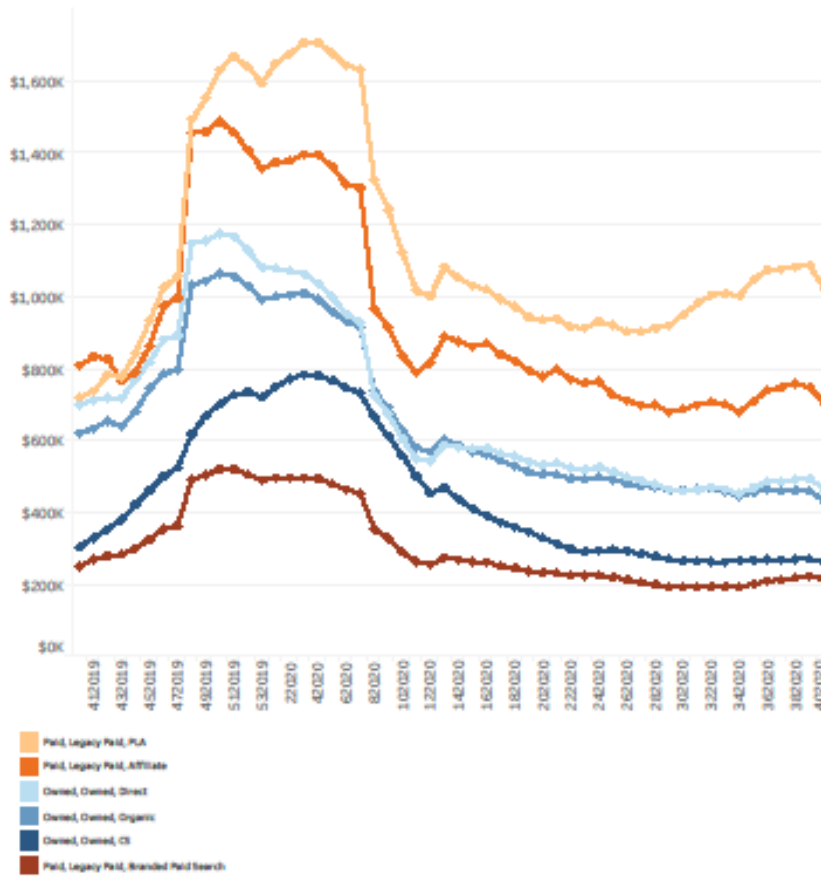


Percentage of Sales by Channel

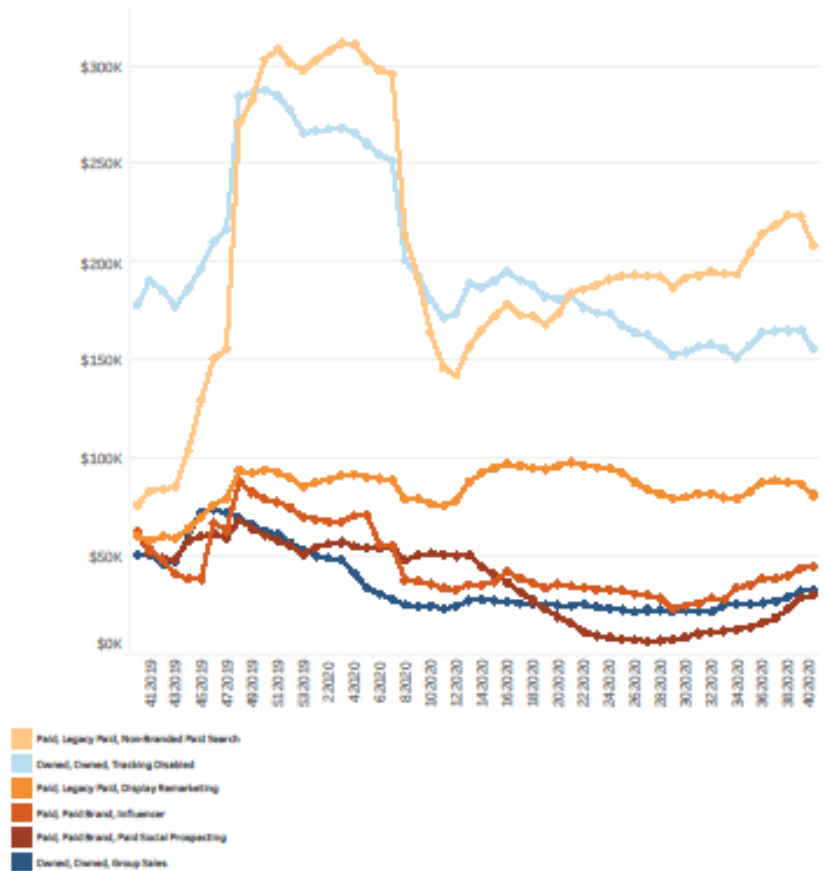
Owned		45.9%
Other Free	Shopping Search	0.0%
		0.0%
Owned	App	0.0%
	Vendor Marketing	0.1%
	Other Free	0.2%
	Group Sales	0.7%
	Tracking Disabled	3.7%
	Email	7.8%
	CS	8.4%
	Organic	12.2%
	Direct	12.8%
		45.9%
Paid		54.1%
Legacy Paid	Paid Social Retention	0.7%
	Display Remarketing	1.7%
	Non-Branded Paid Search	4.1%
	Branded Paid Search	5.8%
	Affiliate	17.7%
	PLA	22.2%
		52.2%
Paid Brand	YouTube	0.0%
	Facebook PLA	0.0%
	Podcast	0.0%
	Paid Social MDF	0.0%
	Pinterest	0.0%
	Paid Social Owned Brands	0.0%
	Social Media	0.3%
	Paid Social Prospecting	0.7%
	Influencer	0.9%
		1.9%

Owned
Paid

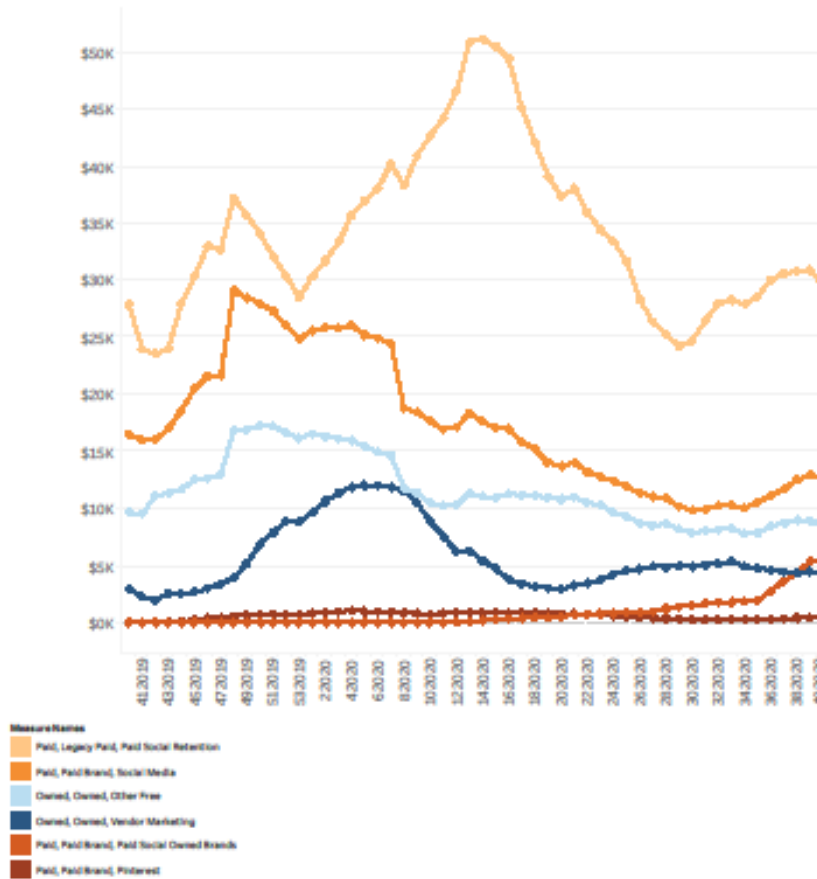
Top Quartile TTM Sales by Channel Name (12-Week SMA)



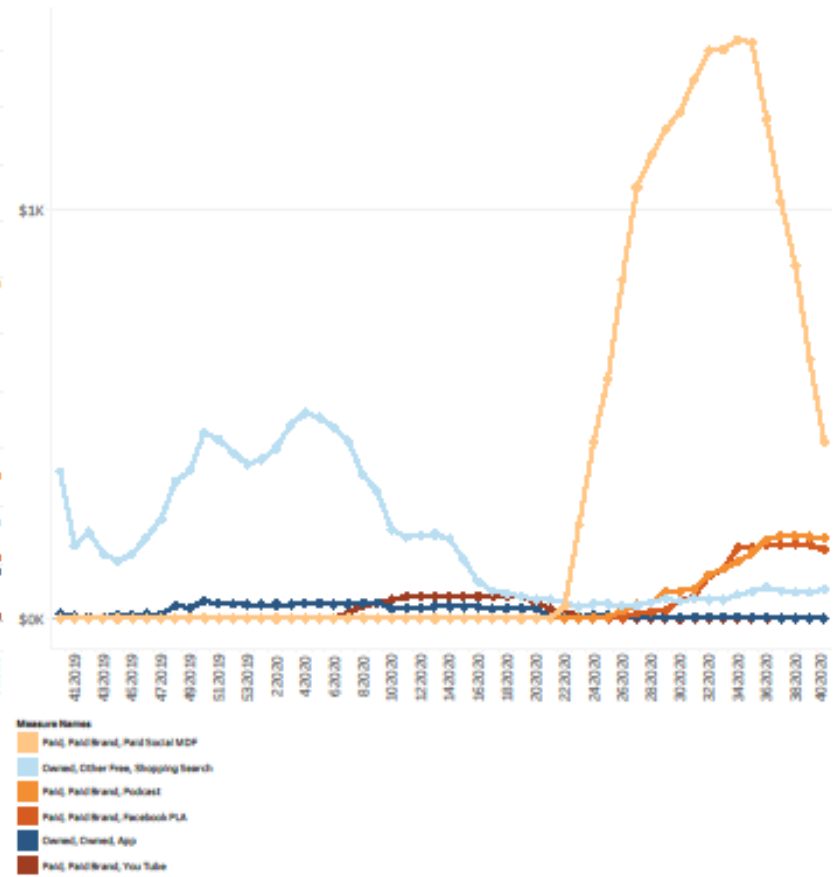
Third Quartile TTM Sales by Channel Name (12-Week SMA)



Second Quartile TTM Sales by Channel Name (12-Week SMA)



Bottom Quartile TTM Sales by Channel Name (12-Week SMA)



Quality of Sales

Mkt Spend	Marketing Channel Name	Marketing Channel Category	% Sales (by spend partition)	Order per Cust	Sales per Order	% Existing Customers	Marketing Spend per Order	Profit Margin			
Marketing Spend	PLA	Paid	22.26%	0.97	\$154	54%	\$33	78%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Affiliate	Paid	17.76%	0.93	\$178	61%	\$18	90%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Branded Paid Search	Paid	5.79%	0.97	\$183	62%	\$6	96%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Non-Branded Paid Search	Paid	4.14%	0.98	\$178	55%	\$39	78%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Display Remarketing	Paid	1.67%	0.96	\$173	79%	\$32	82%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Influencer	Paid	0.88%	0.96	\$116	36%	\$58	50%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Paid Social Prospecting	Paid	0.69%	0.99	\$102	19%	\$63	38%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Paid Social Retention	Paid	0.66%	0.98	\$113	89%	\$43	62%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Paid Social Owned Brands	Paid	0.02%	0.98	\$96	82%	\$91	6%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Pinterest	Paid	0.01%	0.99	\$130	48%	\$424	-227%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
No Marketing Spend	Podcast	Paid	0.00%	1.00	\$120	40%	\$718	-499%	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Direct	Owned	12.81%	0.94	\$179	63%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Organic	Owned	12.19%	0.96	\$174	60%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	CS	Owned	8.40%	0.94	\$193	82%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Email	Owned	7.84%	0.94	\$162	84%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Tracking Disabled	Owned	3.76%	0.94	\$159	64%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Group Sales	Owned	0.67%	0.89	\$1,999	68%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Social Media	Paid	0.33%	0.97	\$153	53%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Vendor Marketing	Owned	0.12%	0.98	\$176	32%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Paid Social Owned Brands	Paid	0.01%	0.98	\$87	91%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Paid Social MDF	Paid	0.01%	0.97	\$112	92%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Pinterest	Paid	0.01%	0.85	\$107	42%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Shopping Search	Owned	0.00%	0.98	\$150	44%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Other Free	Owned	0.00%	0.98	\$150	44%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Facebook PLA	Paid	0.00%	1.00	\$107	32%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	Podcast	Paid	0.00%	1.00	\$106	35%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	App	Owned	0.00%	1.00	\$99	78%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
	YouTube	Paid	0.00%	1.00	\$100	57%	Null	Null	<div><div></div></div>	<div><div></div></div>	<div><div></div></div>
									0.0	0.5	1.0
									Order per Cust Percentile Rank		
									0.0	0.5	1.0
									Sales per Order Percentile Rank		
									0.0	0.5	1.0
									% Existing Customers Percentile Rank		

Quality of Sales

Mkt Spend	Marketing Channel Name	Marketing Channel Category	% Sales (by spend partition)	Order per Cust	Sales per Order	% Existing Customers	Marketing Spend per Order	Profit Margin						
Marketing Spend	PLA	Paid	22.26%	0.97	\$154	54%	\$33	78%	<div><div></div></div>	<div><div></div></div>				
	Affiliate	Paid	17.76%	0.93	\$178	61%	\$18	90%	<div><div></div></div>	<div><div></div></div>				
	Branded Paid Search	Paid	5.79%	0.97	\$183	62%	\$6	96%	<div><div></div></div>	<div><div></div></div>				
	Non-Branded Paid Search	Paid	4.14%	0.98	\$178	55%	\$39	78%	<div><div></div></div>	<div><div></div></div>				
	Display Remarketing	Paid	1.67%	0.96	\$173	79%	\$32	82%	<div><div></div></div>	<div><div></div></div>				
	Influencer	Paid	0.88%	0.96	\$116	36%	\$58	50%	<div><div></div></div>	<div><div></div></div>				
	Paid Social Prospecting	Paid	0.69%	0.99	\$102	19%	\$63	38%	<div><div></div></div>	<div><div></div></div>				
	Paid Social Retention	Paid	0.66%	0.98	\$113	89%	\$43	62%	<div><div></div></div>	<div><div></div></div>				
	Paid Social Owned Brands	Paid	0.02%	0.98	\$96	82%	\$91	6%	<div><div></div></div>	<div><div></div></div>				
	Pinterest	Paid	0.01%	0.99	\$130	48%	\$424	-227%	<div><div></div></div>	<div><div></div></div>				
Podcast	Paid	0.00%	1.00	\$120	40%	\$718	-499%	<div><div></div></div>	<div><div></div></div>					
No Marketing Spend	Direct	Owned	12.81%	0.94	\$179	63%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Organic	Owned	12.19%	0.96	\$174	60%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	CS	Owned	8.40%	0.94	\$193	82%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Email	Owned	7.84%	0.94	\$162	84%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Tracking Disabled	Owned	3.76%	0.94	\$159	64%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Group Sales	Owned	0.67%	0.89	\$1,999	68%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Social Media	Paid	0.33%	0.97	\$153	53%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Vendor Marketing	Owned	0.12%	0.98	\$176	32%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Paid Social Owned Brands	Paid	0.01%	0.98	\$87	91%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Paid Social MDF	Paid	0.01%	0.97	\$112	92%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Pinterest	Paid	0.01%	0.85	\$107	42%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Shopping Search	Owned	0.00%	0.98	\$150	44%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Other Free	Owned	0.00%	0.98	\$150	44%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Facebook PLA	Paid	0.00%	1.00	\$107	32%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	Podcast	Paid	0.00%	1.00	\$106	35%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	App	Owned	0.00%	1.00	\$99	78%	Null	Null	<div><div></div></div>	<div><div></div></div>				
	YouTube	Paid	0.00%	1.00	\$100	57%	Null	Null	<div><div></div></div>	<div><div></div></div>				
										0.0	0.5	1.0	0.0	0.5
									Marketing Spend per Order Percentile Rank			Profit Margin Percentile Rank		