Hillstrom's Catalog Attribution (the Catalog PhD)
Kevin Hillstrom, President, MineThatData
January 25, 2013

As discussed all over the internet, attribution is a vibrant topic in marketing. In catalog marketing, attribution is probably more valuable than elsewhere, simply because of the significant cost of putting catalogs in the mail.

Let me walk you through an example of how my clients approach this topic.

If clients have existing mail/holdout tests, I use the tests in my analysis. In less than half of the cases, this data exists. Let's pretend you have results from a test, results that look something like this:

Quarterly Catalog Mail / Holdout Test Results

	Mail	Holdout	Increment	Organic %	Catalog Driven
Mail Demand	$0.30	$0.03	$0.27	10.0%	90.0%
Phone Demand	$2.00	$0.30	$1.70	15.0%	85.0%
Pure Website Demand	$5.00	$3.75	$1.25	75.0%	25.0%
Email Demand	$3.00	$2.75	$0.25	91.7%	8.3%
Search Demand	$2.50	$1.38	$1.12	55.2%	44.8%
Affiliates / CSEs	$1.50	$0.73	$0.77	48.7%	51.3%
All Other Online Marketing	$1.50	$0.77	$0.73	51.3%	48.7%
iOS / Android / Mobile	$1.50	$1.45	$0.05	96.7%	3.3%
Retail Stores	$8.00	$7.00	$1.00	87.5%	12.5%
Totals	$25.30	$18.16	$7.14	71.8%	28.2%

Most of the attribution rules you'll apply to your business are contained in this table.

Orders that come through the mail (customer sends order form with check, in envelope) and via the call center are almost entirely generated because of catalog marketing. In this example, 90% of orders from the mail, and 85% of orders from the call center are truly catalog driven. We know this, because in the holdout group, customers basically stopped buying in these channels when catalogs were discontinued.

This becomes the attribution rule for these channels.
- 90% of mail demand is attributed to catalogs.
- 85% of phone demand is attributed to catalogs.

Look at pure website demand. Here, we see a different story, don't we? When we stop mailing catalogs, 75% of the demand still happens, whereas 25% of the demand is caused by catalog marketing. Email marketing, in this example, is similar. When

catalogs were discontinued, customers continued to spend 91.7% of their email marketing demand within that channel.
- 25% of pure online demand is attributed to catalogs.
- 8% of email demand is attributed to catalogs.

Online marketing has a different relationship. Here, catalogs cause customers to go online and search for the best price, discount codes, free shipping codes, and to comparison shop.
- 45% of search demand is attributed to catalogs.
- 51% of affiliate / CSE demand is attributed to catalogs.
- 49% of all other online marketing (retargeting, banners) is attributed to catalogs.

Not surprisingly, there isn't a connection between catalogs and mobile devices.
- 3% of mobile device demand is attributed to catalogs.

Retail demand is marginally influenced by catalogs. This is a common outcome.
- 13% of retail demand is attributed to catalogs in this case.

These rules are critical, when determining whether a customer should be mailed a catalog. Let's look at two customers, both equally valuable to this company.

Here is Customer #1:

Catalog Value Analysis			
	Spend	Attributed %	Catalog $
Mail Demand	$0.30	90.0%	$0.27
Phone Demand	$3.00	85.0%	$2.55
Pure Website Demand	$0.50	25.0%	$0.13
Email Demand	$0.00	8.3%	$0.00
Search Demand	$0.25	44.8%	$0.11
Affiliates / CSEs	$0.25	51.3%	$0.13
All Other Online Marketing	$0.25	48.7%	$0.12
iOS / Android / Mobile	$0.05	3.3%	$0.00
Retail Stores	$0.05	12.5%	$0.01
Totals	$4.65	71.3%	$3.31

Using our attribution rules, we can estimate that this customer, expected to spend $4.65 during the catalog window, will spend $3.31 because the catalog was mailed ... 71.3% of the demand this customer will generate is caused by catalog marketing. We can run a profit and loss statement on $3.31.
- Profit = $3.31 * 0.40 - $0.75 = $0.57.

Based on our attribution rules, this customer should be mailed a catalog.

Let's take a look at Customer #2. This customer, interestingly, is an online-centric customer.

Catalog Value Analysis			
	Spend	Attributed %	Catalog $
Mail Demand	$0.01	90.0%	$0.01
Phone Demand	$0.01	85.0%	$0.01
Pure Website Demand	$1.50	25.0%	$0.38
Email Demand	$1.50	8.3%	$0.12
Search Demand	$0.36	44.8%	$0.16
Affiliates / CSEs	$0.36	51.3%	$0.18
All Other Online Marketing	$0.36	48.7%	$0.18
iOS / Android / Mobile	$0.50	3.3%	$0.02
Retail Stores	$0.05	12.5%	$0.01
Totals	$4.65	22.8%	$1.06

This customer has the exact same future value ($4.65) as the first customer. However, this customer will only generate 23% of future demand because of catalog marketing. See, this customer utilizes marketing channels (email, online, mobile) that catalogs do not drive a lot of volume to.

As a result, when we run a simple profit and loss statement for this customer, we learn that the decision to mail a catalog to this customer is not a good one.
- $1.06 * 0.40 - $0.75 = ($0.33).

This is what catalog attribution is all about. For most, catalog attribution is a backward-looking analysis, generated by a complex set of statistical algorithms that are difficult to understand and highly variable due to the confidence intervals surrounding the parameter estimates.

In Hillstrom's Catalog PhD projects, we take a different approach to catalog attribution.

If test results exist, I incorporate the results into the algorithm. If test results do not exist, I use proprietary mathematical outcomes derived across twenty-five years, a career really, of business experience at Lands' End, Eddie Bauer, Nordstrom, and 90+ clients.

There are four key steps in an attribution / Catalog PhD project.

Step 1 = Predict the likelihood of a customer purchasing again next year. Let's assume, in this example, that the customer has a 40% probability of purchasing next year.

Step 2 = Predict how much a customer will spend if the customer purchases again next year. Let's assume, in this example, that the customer will spend $120 if the customer purchases again next year.

Step 3 = Predict the percentage of future demand that will be generated on an organic basis (outside of catalog marketing). Let's use the second customer as our example ... the customer will generate 23% of demand because of catalog marketing, 77% organically (independent of catalog marketing).

We can now estimate how much the customer will spend, annually, because of catalog marketing:
- Probability of Purchasing * Predicted Spend * (1 - Organic %).
- 40% * $120 * (1 - 0.77) = $11.04.

Step 4 = I use this information, in combination with the number of catalogs this customer was mailed last year, as a template for estimating the optimal number of catalogs to mail to a customer in the next year. Here's what the table looks like for a customer who actually received 12 catalogs last year.

Catalogs	Catalog $	All Other $	Total $	Total Profit
0	$0.00	$36.96	$36.96	$14.78
1	$3.19	$36.96	$40.15	$15.31
2	$4.51	$36.96	$41.47	$15.09
3	$5.52	$36.96	$42.48	$14.74
4	$6.37	$36.96	$43.33	$14.33
5	$7.13	$36.96	$44.09	$13.88
6	$7.81	$36.96	$44.77	$13.41
7	$8.43	$36.96	$45.39	$12.91
8	$9.01	$36.96	$45.97	$12.39
9	$9.56	$36.96	$46.52	$11.86
10	$10.08	$36.96	$47.04	$11.32
11	$10.57	$36.96	$47.53	$10.76
12	$11.04	$36.96	$48.00	$10.20

Last year, this customer segment generated $48.00 on 12 catalogs, yielding $10.20 profit.

This year, this customer segment should generate $40.15 on 1 catalog, yielding $15.31 profit.

I don't know about you, but I'd rather generate $15.31 profit instead of $10.20 profit.

And yes, we're going to take a demand hit in this case ... generating $40.15 instead of $48.00. But I now have 11 * $0.75 = $8.25 of incremental ad cost that I can choose to reinvest or pocket. Some catalogers reinvest this money in customer acquisition. Some catalogers reinvest this money in online marketing. Some catalogers invest this money

in new ventures. Finally, some catalogers pocket the money, and are much more profitable. The choice is yours.

On average, I generate $1,000,000 of incremental, annual profit for clients that average $100,000,000 in annual sales. That's not bad, is it?

I include a variant of a five-year forecast in every project I perform. I do this, because it is important for Executives to understand the levers available to grow a business. These days, I create the five-year forecast using the Judy, Jennifer, and Jasmine personas.

- Judy = 60 year old, rural catalog shopping maven.
- Jennifer = 44 year old online shopping expert.
- Jasmine = 28 year old social, mobile, local customer.

Catalogs fully resonate with Judy, influence Jennifer, and have minimal impact over Jasmine. By moving the personas into segment form, I can forecast where your business is headed. If your business is Judy-centric, you're going to be a cataloger into the foreseeable future. If your business is Jennifer-centric, you have significant online marketing opportunities. If your business is Jasmine-centric, listen to all of the pundits, because their social / mobile / local advice is directly relevant to your business.

A forecast might look something like this (see the next page):

Five Year Forecast Model

			Beginning Inventory	End of Year 0	End of Year 1	End of Year 2	End of Year 3	End of Year 4	End of Year 5
0 to 12 Month Buyers	1x LY	Judy	27,953	25,532	24,202	23,617	23,306	23,153	23,073
		Jennifer	13,847	14,011	13,966	13,955	13,940	13,931	13,926
		Jasmine	44,619	48,382	49,069	49,690	49,952	50,091	50,158
	2x+ LY	Judy	15,667	13,852	12,797	12,255	11,975	11,833	11,760
		Jennifer	13,882	14,059	13,981	13,927	13,886	13,862	13,848
		Jasmine	25,080	28,075	29,335	29,993	30,311	30,469	30,546
13-24 Month Buyers		Judy	31,990	26,085	23,786	22,487	21,888	21,571	21,413
		Jennifer	12,204	11,399	11,540	11,493	11,472	11,453	11,441
		Jasmine	44,745	41,878	45,598	46,524	47,191	47,485	47,638
New/Reactivated Buyers	1x LY	Judy		13,128	13,128	13,128	13,128	13,128	13,128
		Jennifer		6,965	6,965	6,965	6,965	6,965	6,965
		Jasmine		29,687	29,687	29,687	29,687	29,687	29,687
	2x+ LY	Judy		3,389	3,389	3,389	3,389	3,389	3,389
		Jennifer		3,415	3,415	3,415	3,415	3,415	3,415
		Jasmine		8,760	8,760	8,760	8,760	8,760	8,760
12 Month Buyer Totals		Buyers	141,048	143,911	143,349	143,436	143,370	143,340	143,311
		1x LY	86,419	87,925	87,236	87,262	87,198	87,176	87,157
		2x+ LY	54,629	55,986	56,113	56,175	56,172	56,164	56,154
		Newbies		65,344	65,344	65,344	65,344	65,344	65,344
		Judy	43,620	39,384	36,999	35,872	35,282	34,986	34,833
		Jennifer	27,729	28,070	27,947	27,881	27,826	27,794	27,774
		Jasmine	69,699	76,457	78,403	79,683	80,263	80,561	80,704
Total Annual Demand				$32,543,696	$32,578,581	$32,587,589	$32,572,619	$32,560,749	$32,551,249

This is a business, for instance, that is Jasmine-centric, and is not forecast to grow fast. There will be many opportunities to trim circulation among the Jennifer/Jasmine-centric audience, and to reallocate those dollars along the activities that interest Jasmine.

All attribution / Catalog PhD projects include a five-year forecast, so that the Executive can get a good view of where the business is headed, and can understand why newer marketing activities do/do-not work.

Equations and Grades

A series of equations are generated from an Attribution / Catalog PhD project. Clients either implement the equations in-house, have their database provider implement the equations for them, or ask me to score their customer file on a quarterly basis.

Most of my clients ask me to modify the equations into grades. This is done to make it easier to implement my findings. The grades look something like this:

A = Mail as many catalogs as is humanly possible. Create new in-home dates!
B = Mail the maximum number of catalogs in your existing strategy.
C = Mail customer 3-9 times per year (average = 6).
D = Mail customer 1-3 times per year.
F = Mail customer 0-1 times per year.

You should know that many of my projects result in a housefile circulation decrease of between 15% and 35%, in total, on an annual basis.

Profitability Metrics

To complete the project accurately, I need to know two pieces of information.
- Flow-Through Percentage: The percentage of demand that flows-through to profit, after accounting for returns, cost of goods sold, and pick/pack/ship expense. Do not include advertising cost in this metric. This number is usually somewhere between 35% and 55% of demand.
- Average Catalog Cost: This can be delivered as a total variable cost (i.e. $0.55) or as a cost per thousand pages circulated ($7.00).

Project Cost

Project costs are straightforward, for a four week attribution / Catalog PhD project. They are based on the size of a business. Half of the project fee is due prior to starting a project. Half of the project fee is due within fifteen days of project completion. Project fees are non-negotiable.
- Annual Sales of $1 to $9,999,999 = $10,000.
- Annual Sales of $10,000,000 to $29,999,999 = $20,000.
- Annual Sales of $30,000,000 to $59,999,999 = $27,000.
- Annual Sales of $60,000,000 to $99,999,999 = $35,000.
- Annual Sales of $100,000,000 to $999,999,999 = $45,000.
- Annual Sales of $1,000,000,000 or more = $55,000.

File Requirements:

A project requires just a handful of files, with very simple file layouts. The files are typically sent to my FTP server. The files must be sent in .csv format.

File #1 (.csv format) = One Row per Item Purchased (at least 36 months … only include applicable fields).

Customer Number:	A unique identifier that tells me which customer purchased the item.
Order Date:	The date (20091224) an item was purchased by the customer.
Order Number:	The order number associated with the order.
Item Number:	The unique code associated with the item purchased.
Merchandise Division:	A high-level identifier (furniture, lighting, bathware, etc.) for each item.
Merchandise Brand:	If applicable, the brand represented by the item (Ugg Boots, for instance).
Quantity:	The number of items purchased.
Price:	The price of the item.
Demand:	Quantity * Price.
Discounts:	Any % off or free shipping etc. applied to this item in the purchase.
Physical Channel:	Phone, Online, Store, iPhone, Android, Tablet, App : The physical channel the item was purchased in.
Advertising Channel:	The ad channel that drove the order (e-mail, catalog, affiliates, paid search, natural search, etc.). This can be generated internally, or can come from the co-op/vendor who does matchbacks.
Store Number:	The number of the store a customer purchased from, if applicable.
Catalog ID:	The ID from the catalog the customer purchased from, if applicable.
Gross Margin:	The margin associated with the item purchased.
Tender Type:	How payment was made (credit card, cash, check, etc.).

File #2 (.csv format) = One Row per Customer (only include the fields that are applicable or easy to get to).

Customer Number:	A unique identifier for each customer.
First Purchase Date:	The date a customer first purchased merchandise.
First Purchase Channel:	The physical channel where a first purchase happened (phone, online, store).
First Ad Channel:	The ad channel where a first purchase happened (e-mail, catalog, affiliates, search, etc.)
Distance From A Store:	In miles (15, 75, 294), only if applicable.
Zip Code:	Zip code where the customer lives.
Demographic Info:	Anything you wish to append can be added.
E-Mail Subscriber Date:	Date customer opted-in to receive e-mail campaigns.
E-Mail Opt-Out Date:	Date customer opted-out of e-mail campaigns (or became undeliverable).
Twitter ID:	Id from Twitter … or any other appended social media overlay data.

File #3 (.csv format) = One Row for Each Catalog Mailed, Past Twelve Months

Customer Number:	A unique identifier for each customer.
Mail Date:	The date a customer was mailed a catalog.
Key Code:	The code printed on the back of every catalog, segment info.
Catalog Identifier:	The identifier that tells what the catalog is (spring, holiday).
Catalog Cost:	Cost to mail the catalog (variable) to that customer.

File #4 (.csv format) = One Row for Every Website Visit, Past Two Years (optional)

Customer Number:	A unique identifier for each customer.
Visit Date:	The date a customer first purchased merchandise.
Referring URL:	The site the customer visited prior to this website visit, summarized, or the entire URL.
Depth Indicator:	Depth customer got into site (home page, shopping cart, checkout, purchase).