Algorithmic Marketing Attribution and Conversion Journey Analysis Using SAS® Customer Intelligence 360

Malcolm Lightbody, Principal Product Manager, SAS Institute Inc.
Suneel Grover, Advisory Solutions Architect, SAS Institute Inc. and Professorial Lecturer,
The George Washington University (Department of Decision Science - School of Business)

ABSTRACT

Everyone has a marketing attribution[1] problem, and all attribution measurement methods are wrong. You hear that all the time. Like all urban myths, they are founded on some truth. Most organizations believe they can do better on attribution. They all perceive gaps in missing touchpoint data, multiple identities across devices, arbitrary decisions on weightings for rules, and uncertainty on what actions to drive from results.

Broadly speaking, the holy grail of media measurement is to analyze the impact and business value of all company-generated marketing interactions across the complex customer journey. Our goal is to take a transparent approach in discussing and demonstrating how SAS® has built data-driven marketing technology to help progress past typical attribution methods and make the business case for customer journey optimization.

INTRODUCTION

Being SAS, an analytic approach is naturally advocated to address the operational and process-related obstacles that are commonly experienced by customers. SAS wants to treat the obstacles as two sides of the same coin. The output of attribution analytics informs marketers about what touchpoints and sequence of activities drive conversion. This leads marketers to make strategic decisions about future investment levels, as well as more tactical decisions about what activities to run.

In an ideal world, the results of subsequent actions are fed back into the attribution model to increase not only its explanatory power, but also its predictive abilities, as shown below.

How We Think About Attribution
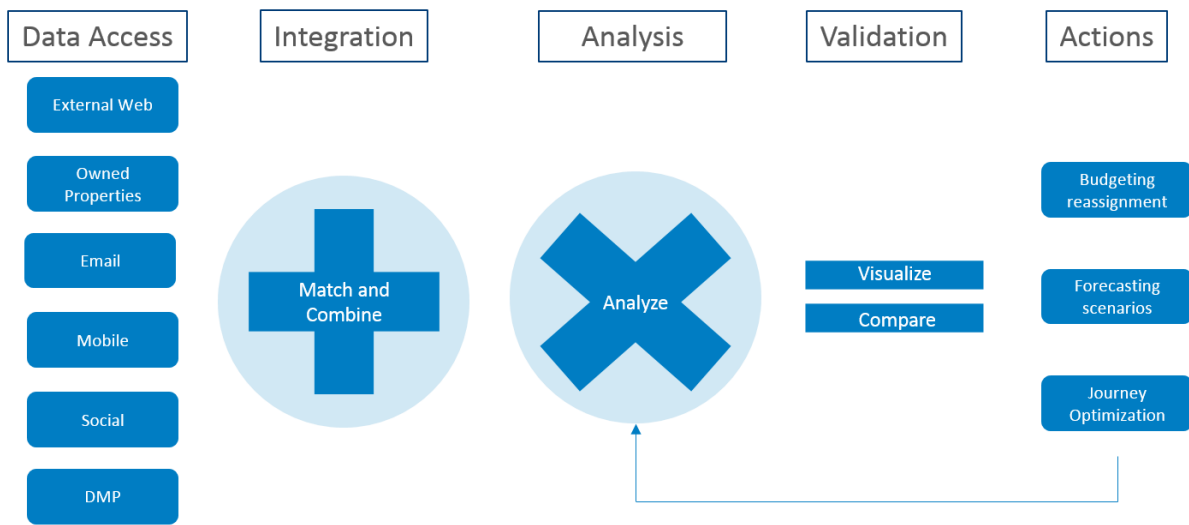
Figure 1. How we think about attribution

Figure 1 shows the main parts of an attribution project. The actual analysis is just part of the process, with upstream and downstream dependencies. But, this doesn't always happen as it should. Consider a standard attribution report. Let's for the moment ignore what technique was used to generate the result and place ourselves in the shoes of the marketer, who is trying to figure out what to do next.
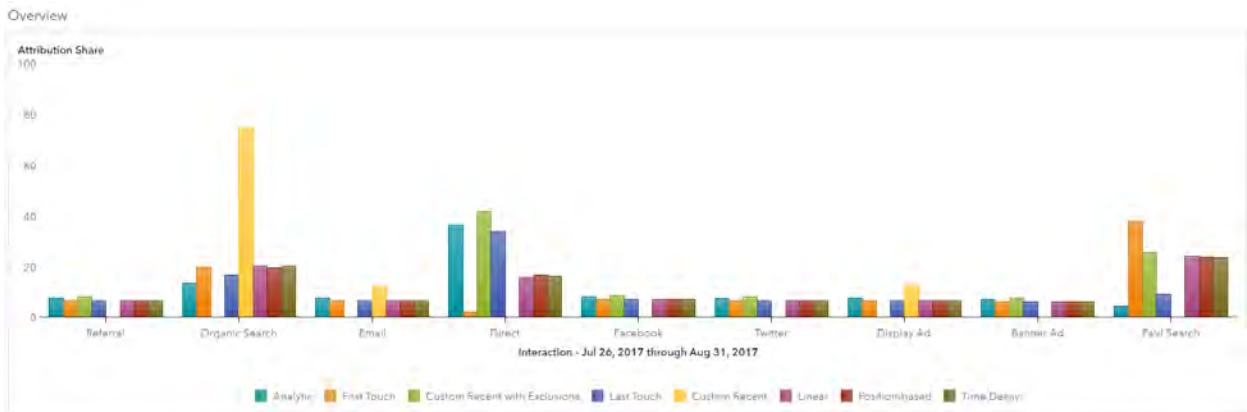
Figure 2. Attribution report

In the graph above, you see the results of an attribution analysis based on a variety of measurement methods.

RULES-BASED AND ALGORITHMIC MEASUREMENT METHODS

Before answering the question of which method should you focus on, let's do a quick review of assessment techniques.

LAST-TOUCH AND FIRST-TOUCH ATTRIBUTION

This type of attribution allocates 100% of the credit to either the last-touch or first-touch of the customer journey. This approach has genuine weaknesses. It ignores all other interactions with your brand across a
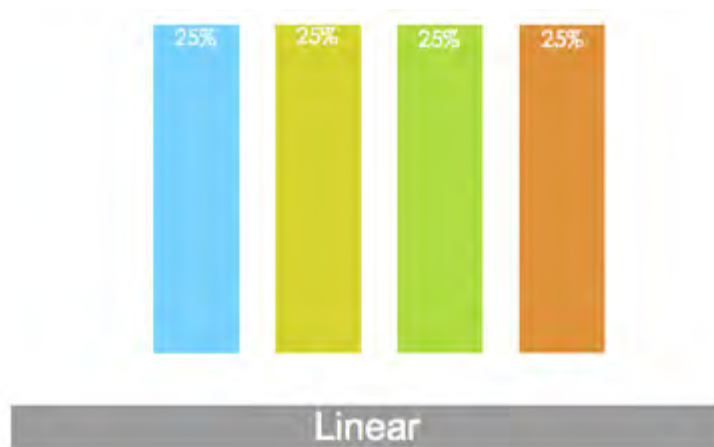
multi-touch journey.
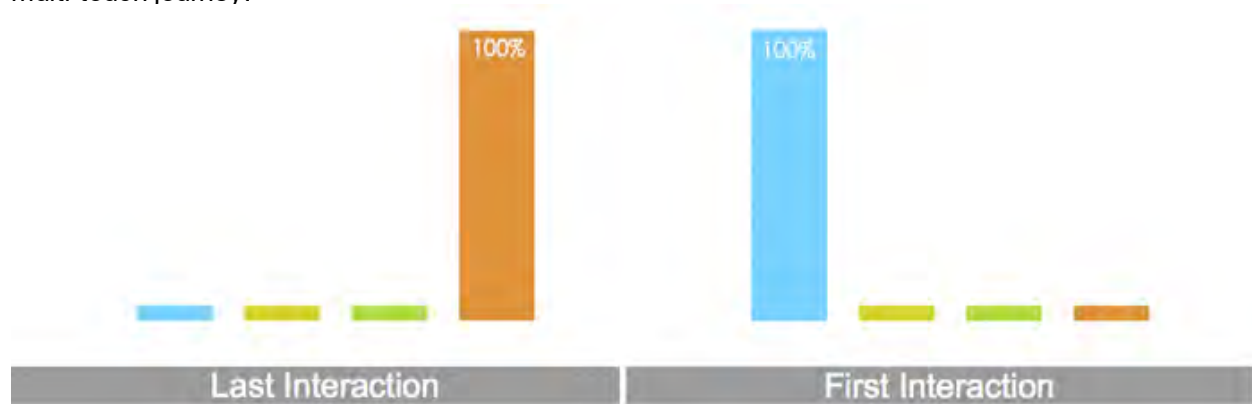
Figure 3. Last-touch and first-touch attribution

LINEAR ATTRIBUTION

Linear attribution arbitrarily allocates an equal credit weight to every interaction along the customer journey. Although slightly better than the last-touch and first-touch approach, linear attribution under-credits and over-credits specific interactions.

Figure 4. Linear attribution

TIME-DECAY AND POSITION-BASED ATTRIBUTION

Time-decay attribution arbitrarily biases the channel weighting based on the recency of the channel touches across the customer journey. If you support the concept of recency within RFM[2] analyses, there is some merit to this approach. Position-based attribution places more weight on the first touch and last touch, while providing less value to the interactions in between.
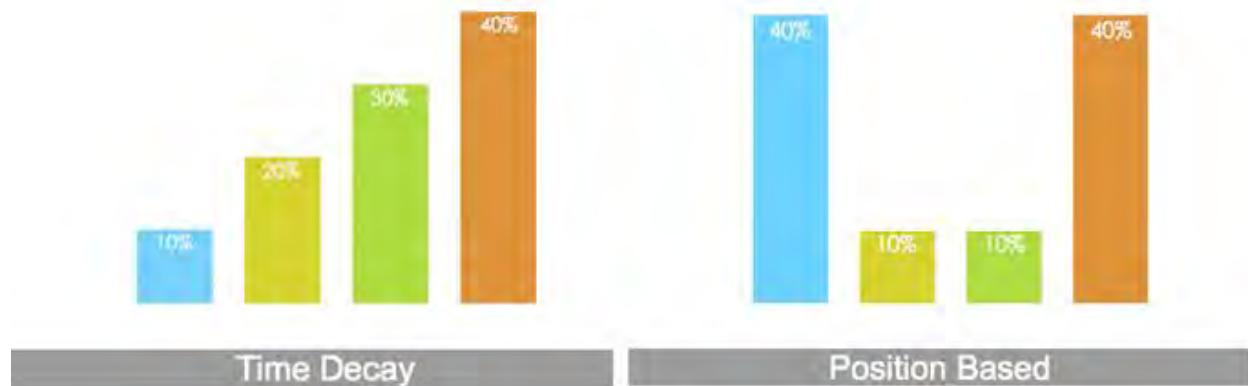
Figure 5. Time-decay and position-based attribution

ALGORITHMIC ATTRIBUTION

In contrast, algorithmic attribution (sometimes referred to as "custom models") assigns data-driven conversion credit across all touchpoints preceding the conversion. It uses math typically associated with predictive analytics or machine learning to identify where credit is due. It analyzes both converting and non-converting consumer paths across all channels. Most important, it uses data to uncover the correlations and success factors within marketing efforts. Here is a video[3] summarizing a customer case study example to help demystify what this all means.
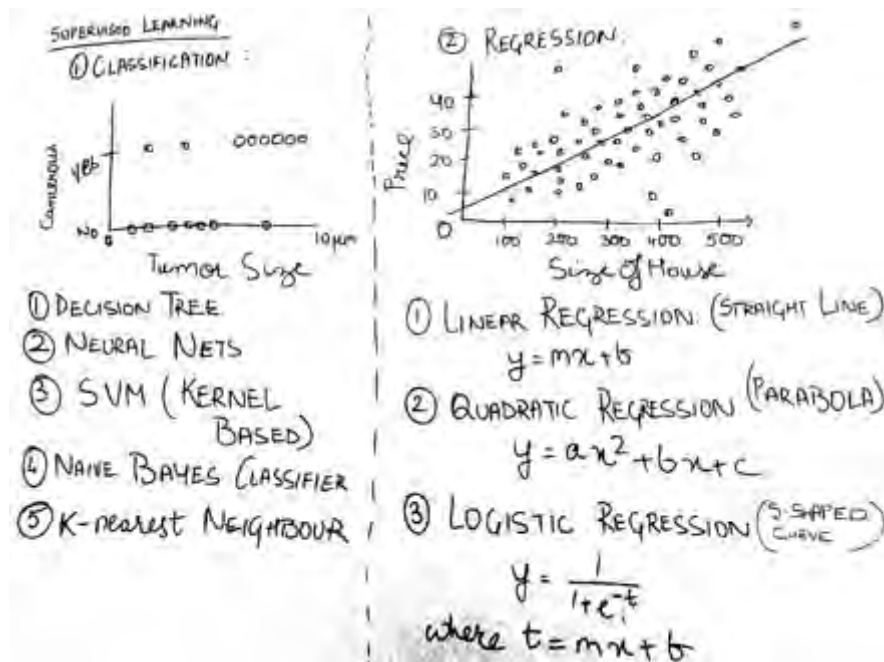

Figure 6. Algorithms

WHY DOESN'T EVERYONE USE ALGORITHMIC ATTRIBUTION?

Although many marketers recognize the value and importance of algorithmic attribution, adopting it hasn't been easy. There are several reasons.

- **Much-needed modernization**. The volume of data that you can collect is massive. This might overwhelm outdated data management and analytical platforms. This is especially true when you need to integrate multiple data sources. Organizations have a decision to make regarding modernization.

- **Scarcity of expertise**. Some believe the talent required to unlock the marketing value in data is scarce. However, there are countless universities offering undergraduate, graduate, and doctoral analytic and data science programs inside and outside the United States. Talent is flooding into the industry. The synergy between analysts and strategically minded marketers is the key to unlock this door.

- **Effective use of data**. Organizations are rethinking how they collect, analyze, and act on important data sources. Are you using all your crucial marketing data? How do you merge website and mobile app visitor data with email and display campaign data? If you accomplish all of this, how do you take prescriptive action between data, analytics, and your media delivery endpoints?

- **Getting business buy-in**. Algorithmic attribution is often perceived as a black box, which vested interest groups can use as a reason to maintain the status quo.

Returning to the question of which method should you ultimately focus on; the answer is, "It depends." An attribution report on its own cannot decide this. And, it doesn't even matter if the attribution report is generated using the most sophisticated algorithmic techniques. There are four things that the report won't tell you:

1. The elasticities of a single touchpoint.
2. The interdependencies between different touchpoints.
3. Cause and effect and timing dependencies.
4. Differences between separate groups of customers.

TOUCHPOINT ELASTICITY

The traditional approach to putting attribution results into action assumes a positive relationship between channel or touchpoint spending and conversion credit. This is the logic that leads a marketer to conclude that he should divert resources toward better performing touchpoints–double-spending on something big will make it much bigger because doubling-spending on something small will make it only a little less small.

Whenever you talk about elasticity—price, response, supply, and demand—you usually think of a curve that looks like the green line below. This comes from the experience of modeling many economic scenarios. What it is telling you is that you need to spend at least a certain amount to stimulate demand, but that there is also a point where the law of diminishing returns applies.
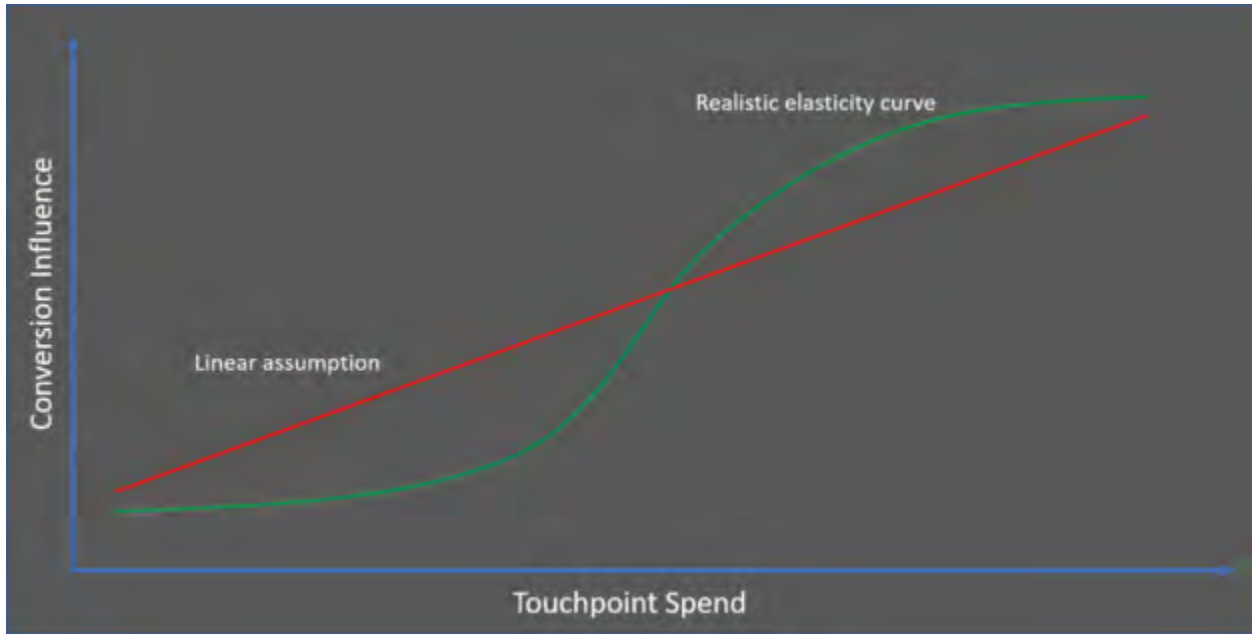
Figure 7. Economic scenarios—linear

On the other hand, marketing analysts tend to assume that the red linear model fits, which isn't the case (except as an approximation over a small region). And, what you can't tell from the results of an attribution analysis is where exactly current investment levels lie on the green line. Ideally, you would want to be somewhere in the region of point B in the graph below because you know that the impact of any increased spending would be greatest. Knowing that you are at point A tells you that you need to invest significantly more to see a benefit, whereas any additional spending is wasted at point C (and might even be detrimental, though this isn't shown in Figure 8).
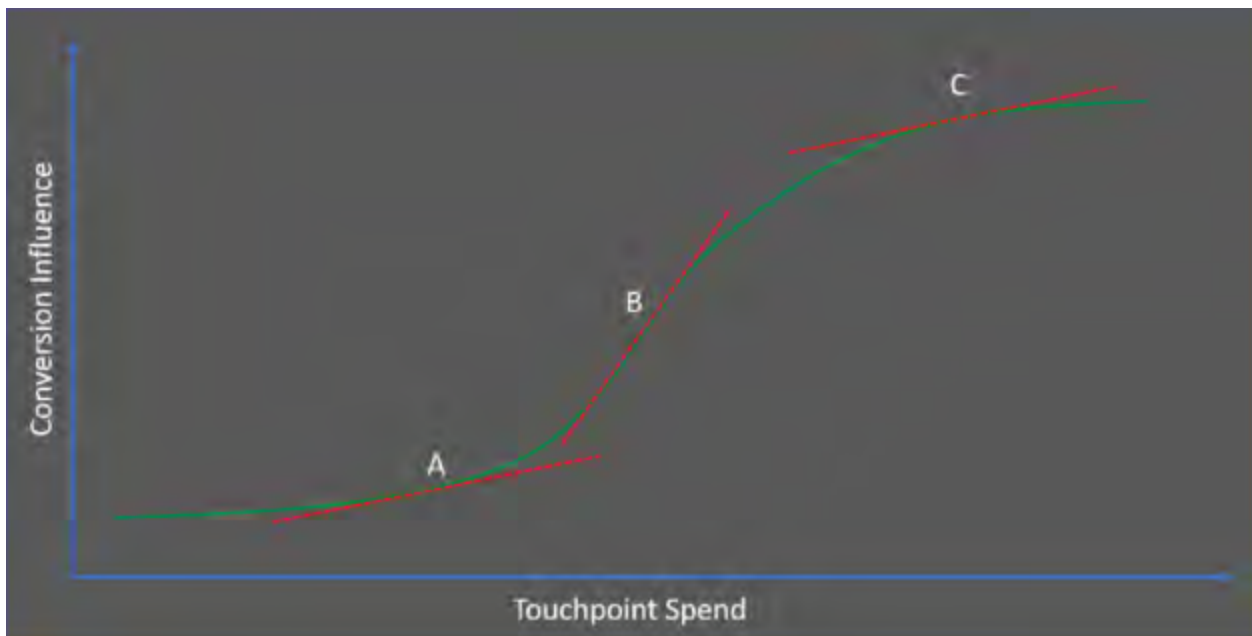
Figure 8. Economic scenarios–points A, B, and C

Successful attribution analysis currently requires an iterative approach to determining the best actions and discovering the shape of the response curve.

TOUCHPOINT INTERACTIONS

With a limited investment pot, there is always going to be competition between different marketing activities and touchpoints. Thus, the more you spend on one channel, the less you can spend on another. The influence that one channel has on another is what is being discussed here.

Visualize a scenario where a user receives a viewable impression of a display media ad, and subsequently searches for the advertised product—perhaps because she remembers seeing the ad. In another scenario, the user receives a promotional email, and then later that day clicks on a social media ad for that specific product. In both cases, there is a question as to how much the first activity influenced the second and how to accurately measure this.

CAUSE AND EFFECT

Does the sequence of ad media impressions matter? Does one act as a warm-up for the other and, if so, what is the optimal timing between the two? This is the area where attribution and customer journeys meet. What are the common sequences of activity that drive conversion? How can these be influenced? And, how does this compare with how the organization thinks that a customer interacts with them? Keep in mind that this holds true only for touchpoints that can be measured.

CUSTOMER DEPENDENCIES

Can you realistically apply a single attribution model when you have different groups of customers (or segments)? Assume that homogeneity simplifies the analysis (but might come at a cost of missed opportunities). Plus, a single user does not always present himself or herself with a consistent way. Users might have different objectives on subsequent visits or even multiple objectives in one visit.

Uncovering those differences informs the customer journey analysis. The idea that all this information informs strategic planning and day-to-day operational marketing activities is gaining traction with many SAS customers.

THE SAS® CUSTOMER INTELLIGENCE 360 APPROACH

SAS is delivering on its vision of artificial intelligence-driven[4] customer journey optimization. The first step is attribution and journey path analysis. This step uses historical and current information to build representations of what the future might look like. It is an exciting topic, and one that requires technical know-how as well as market acceptance of the perceived loss of control over individual decisions. It will not happen overnight, so SAS is pacing itself and releasing a set of capabilities that build toward that future state.

The first area to consider is attribution. SAS recognizes that attribution is a mature market and that there are incumbent vendors with years of experience. However, organizations still believe they have an attribution problem. There are numerous reasons for this. (Here is an example[5].) The reality is many organizations have customers or prospects who often fly under the radar across the interaction landscape. SAS believes the way forward is to build an evolving, comprehensive picture in the visible areas of consumer data and expand this view as new touchpoint data becomes available.

The SAS approach to attribution is to use first-party marketing data to find the best method to apportion credit to the different touchpoints, while considering interaction and timing effects. SAS uses an approach that includes rules-based attribution, but the focus is on a SAS algorithmic method, which is referred to as "analytic attribution."

ANALYTIC ATTRIBUTION

SAS has developed an analytic approach that is channel-independent. The data can come from any touchpoint, provided it fits into a transactional format. SAS Customer Intelligence 360 considers traffic originations as well as internal promotions when calculating the share of conversion credit. Data processing is automatic and uses a transactional stream representing, for each user ID, a time-stamped sequence of interactions across the different touchpoints as shown below:

Figure 9. Analytic attribution—data processing

The first part of the analysis is to identify the distinct journeys contained in the data. From these journeys, SAS determines which lead to conversion and (equally important), which do not.

SAS uses a machine learning technique to compile the journeys, looking for the strength of associations in the data that link different touchpoints together. This is not dependent on there being a conversion, unlike rules-based methods that use a conversion as the endpoint. This is important because you need to be able to compare touchpoint roles in cases where a journey converts to those where there is no conversion.

Visitor ID	Timestamp	Event
1	03/01/15	Display
1	03/20/15	Display
1	03/21/15	Email
1	03/22/15	$
2	04/10/15	Email
2	04/23/15	Search
2	04/24/15	$
3	04/25/15	Email
4	04/30/15	Display
...		

Visitor ID	Journeys
1	Display
1	Display, Email, $
2	Email
2	Search, $
3	Email
4	Display
...	

Figure 10. Analytic attribution—compiling journeys

For the analytically inclined, these journeys form the analytic base table (one row per journey) used as the input data view for the attribution analysis. The individual touchpoints become the predictors, with values for each row based on the number of times that touchpoint was seen in the visitor journey. What you are trying to predict is the final column—a "1" indicates the journey converted and a "0" means it didn't. The resulting machine learning model is used to identify the contributions that different touchpoints made to overall conversion. SAS uses an iterative method to calculate the incremental lift from touchpoints to determine attribution weights.

Journey ID	Journey	Other Channels Credit	Display Credit	Email Credit	Search Credit
1	Display, Email, Search, Display	0.20	0.30	0.10	0.40
2	Display	0.35	0.65		
3	Display, Display	0.75	0.25		
4	Search	0.71			0.29
5	Search, Email	0.15		0.60	0.25
...					
n	Email, Search	0.450		0.30	0.25
	Total channel conversion credit:	10,000	13,000	7,000	20,000
	Channel attribution:	20.00%	26.00%	14.00%	40.00%

Figure 11. Analytic attribution—determining incremental lift

All identified journeys are used to calculate the final conversion credit. In the chart above, you can see that display advertising is credited with driving 26% of the conversion.

JOURNEY ANALYTICS AND TIMING

What Figure 11 doesn't show is any indication of the sequence or timing of interactions. The SAS approach to including sequencing is to further refine the input data. Instead of simply having the sequence "search then email" as a journey, SAS differentiates based on the time between the two items. The result is many more predictor columns, each one relating not just to a touchpoint, but also to the time delay.

| Journey	Probability Conversion	Incremental Lift	Other Channels Credit	Display Credit	Email Credit	Search Credit
Baseline (due to other channels e.g. print, radio, TV, etc.)	0.010	0.000	0.010			
Display_5	0.020	0.010		0.010		
Display_5, Email_2	0.025	0.005			0.005	
Display_5, Email_2, Search_1	0.045	0.020				0.020
Display_5, Email_2, Search_1, Display_1	0.050	0.005		0.005		
Total channel lift credit:			0.010	0.015	0.005	0.020
Channel conversion credit:			0.200	0.300	0.100	0.400

Figure 12. Analytic attribution—timing

TOUCHPOINT INTERACTION

Another effect in the attribution model is the interaction between the different touchpoints. The principle of how SAS tackles this is the same as it is for time-based attribution. SAS creates new predictors as inputs to the model based on finding relevant combinations of channels. SAS then uses a statistical modeling technique to detect the presence and validity of interactions rather than simply binning and counting. Finally, SAS combines variables in the analytic base table that include the time dependency, as well as identifying the main interaction effects.

ON-SITE VIEWABLE IMPRESSIONS

Adding on-site viewable impressions (promotions, testing, and other forms of personalization) to the attribution mix enables the marketer to understand which of these influence conversions. Being able to tease out interactions between internal and external touchpoints gives the marketer a better picture of how to tune the overall on-site experience for a business advantage. For example, visualizing conversion paths that combine external and internal touchpoints leads to insight that helps optimize journeys to maximize conversion goal achievement.
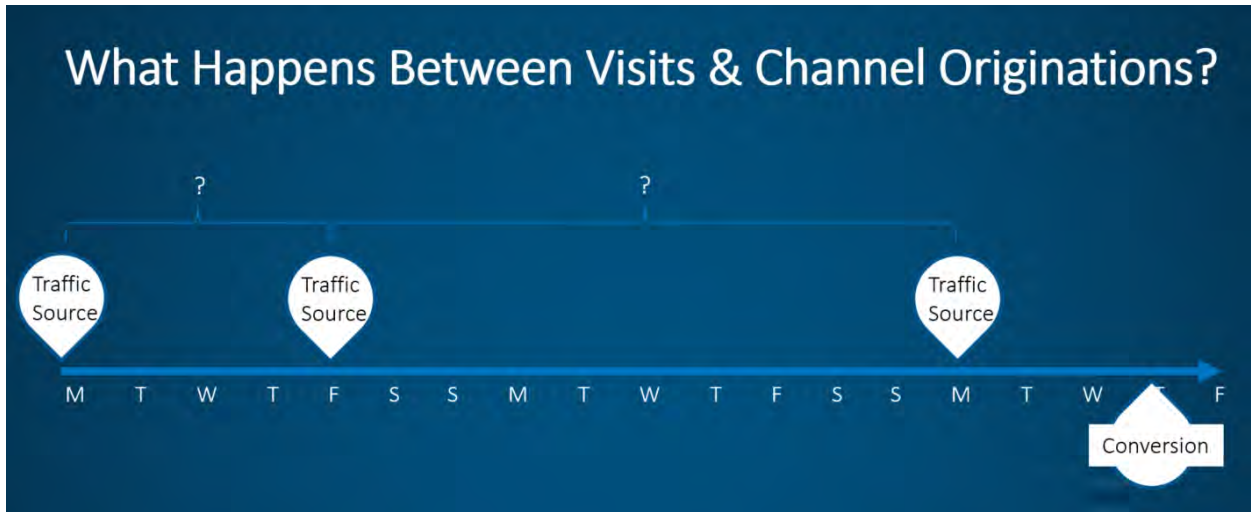
Figure 13. What happens between visits and channel originations?

A challenge is the volume of on-site viewable impressions. A typical website consists of many pages that have many content spots (or areas available for personalization). Each of these spots can contain multiple candidate messages and creatives. Website visitors do things that obfuscate their true intentions. Not deliberately, of course, but as part of the normal visit process. For example, refreshing pages, going from one page back to another in a short space of time, and so on. A spot set up for personalization on a given page might fire many times in a short space of time. Each time this creates an impression record.

The SAS approach to this challenge is two-fold. First, enable the marketer to define which on-site interactions are candidates for attribution analytics. Even that can still lead to multiple interactions within short time windows. This leads to the use of machine learning techniques to determine which sequence lengths are significant and which can be collapsed together. Like the descriptions about time-based attribution and touchpoint interaction detection, SAS looks for the strength of association between individual occurrences of on-site viewable impressions and the overall attribution goal. The result is more manageable path lengths and a higher signal-to-noise ratio in the attribution model's input data.


Figure 14. On-site viewable impressions

PUTTING IT ALL TOGETHER

This paper has talked about what you should do and how you should go about it. Up to now, it's unclear how the story ends.

Marketers have expectations about what they see as output from attribution analysis, and SAS Customer Intelligence 360 offers easy-to-understand reporting views for each defined channel. Analysts can compare the results from any rules-based method with algorithmic attribution. Iterating different courses of action, based on comparisons, can help build confidence in analytic methods. At the very least, it can help choose the best rules-based method.
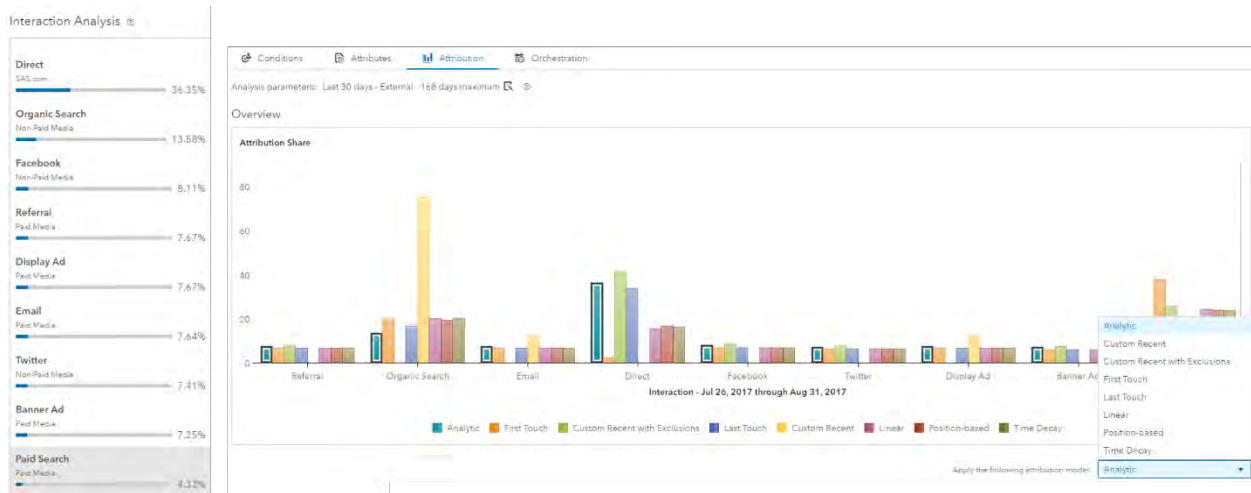

Figure 15. Attribution—reporting view

INTERACTION JOURNEY ANALYSIS

The exciting news is how SAS is using attribution as a first step into journey analytics and, ultimately, into customer journey optimization. Being able to credit paid search marketing with 4.3% of conversion revenue is useful to know, but this leads to more questions.

1. What conversion journeys included paid search?

2. Where in the journey did paid search touch the customer?

3. What is the best timing between a paid search interaction and others?

4. What types of customers included touches with paid search, and how do they differ from those who didn't?

The Interaction Journey Analysis feature answers these questions and more. For each channel, you can discover:

- The top conversion paths that include the channel.

- The relative contribution of that channel in any given conversion path.

- The relative position of that channel in the conversion path.
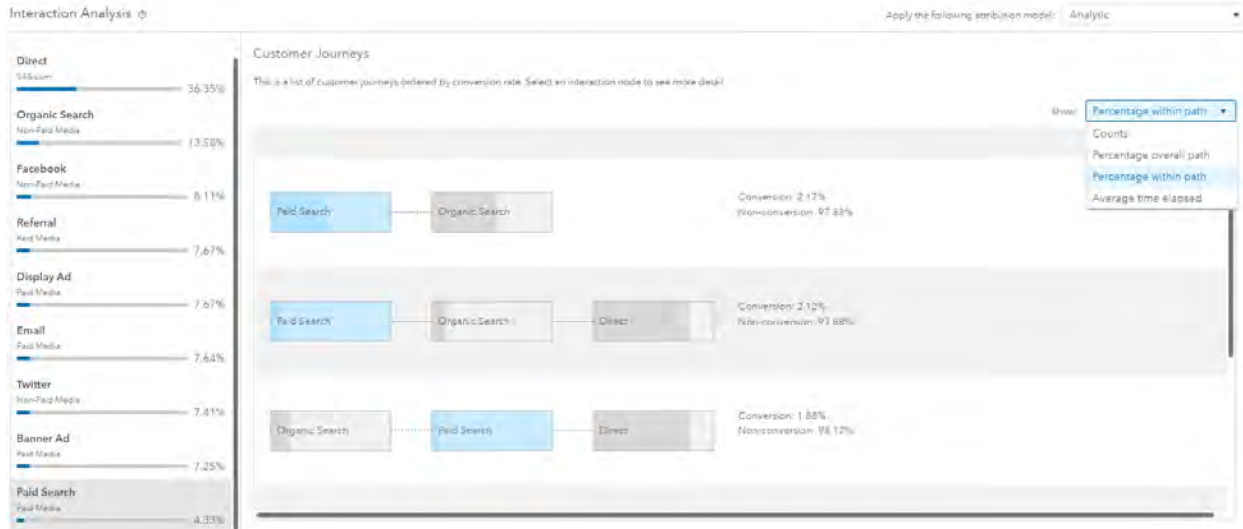
Figure 16. Interaction journey analysis

In Figure 16, selecting **Paid Search** in the left pane displays the top paths that include paid search (on the right). The shading within each channel in a path shows its relative importance for that specific path. And, its position in the conversion path lets you know whether it played more of an acquisition role or a closing role.

PATH TIMING

There is one last item worth mentioning, and that is the ability to visualize timing effects in conversion path analytics.
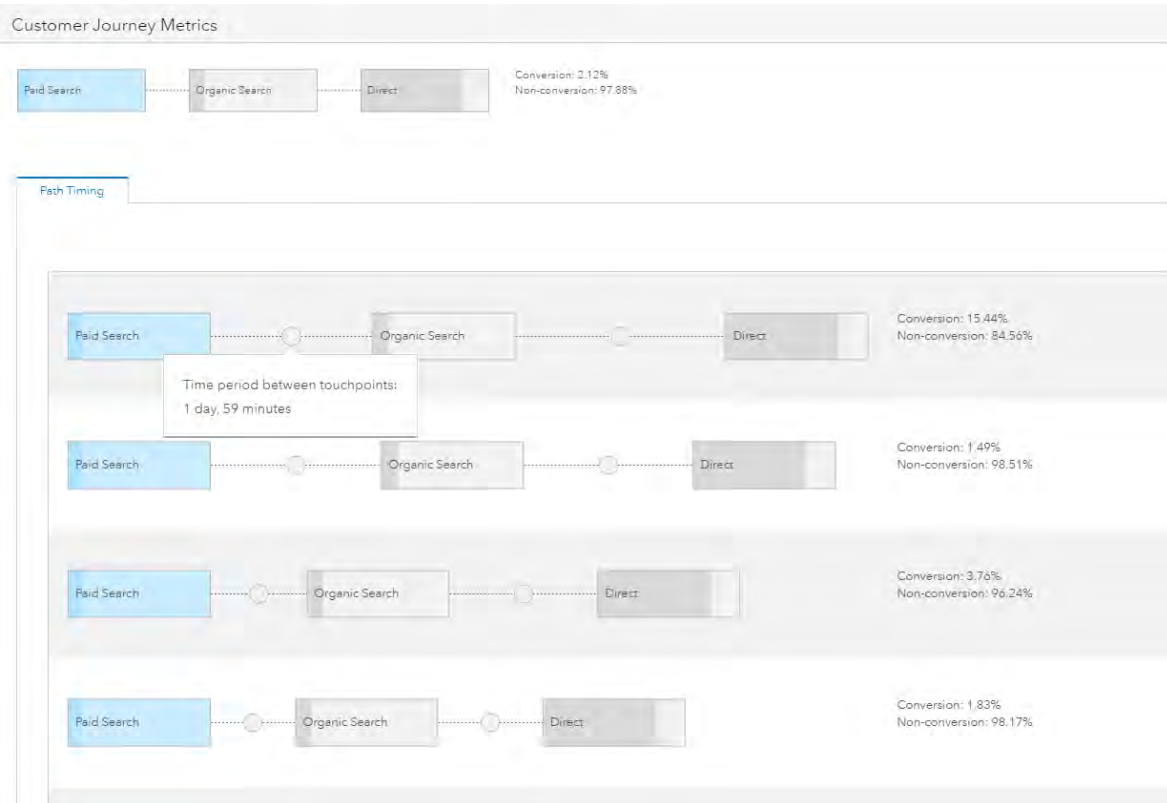
Figure 17. Conversion path analysis

In Figure 17, there is a single conversion path (shown at the top). Below that, there is a breakdown into different sequences depending on the timing between components. This ability to discover natural path cadences is invaluable for the marketer looking to improve the efficiency of journeys.

NEXT STEPS

The initial attribution release of SAS Customer Intelligence 360 showcases analytic strengths and reinforces the SAS vision to be the leader in bringing advanced techniques to the digital marketer. It's analogous to self-driving automobiles. These techniques represent the first step in a journey that aims to link the rear-view mirror (think "traditional attribution") to forward sensors (think "predictive journey optimization") to provide advanced driver assistance capabilities.

Practical next steps welcome Journey Optimization capabilities. The aim is to apply self-learning to generate insight and understand what sequences of actions and creatives drive conversion over time. Subsequently, this knowledge can drive the display of appropriate creatives to visitors to achieve the business goal. The selection methodology is based on:

- What sequences of creatives drive conversion from different visitor states?
- Improvements in conversion rate over time.
- Improvements in average number creatives required to generate conversion over time.

The approach uses reinforcement learning[6] to determine, for each customer, the optimum sequence of messages to maximize a conversion goal. A simple analogy is to view this as multi-arm-bandit testing[7] on steroids.

Reinforcement learning is often observed in applications for robotics, gaming, and navigation. The algorithm discovers, through trial and error, which actions yield the greatest rewards. This type of learning has three primary components: the agent (the learner or decision maker), the environment (everything the agent interacts with), and actions (what the agent can do). The objective is for the agent to choose actions that maximize the expected reward over a given amount of time. The agent will reach the goal much faster by following a good policy. So, the goal in reinforcement learning is to learn the best policy.

In this description, SAS Customer Intelligence 360 is the agent, and by using a prediction model to compute the expected value of presenting different creatives based on visitor state information, it will select the creative with the highest expected value into the digital experience. Visitor state information can include the following:

- Creatives shown to the visitor
- Demographic information about the visitor
- On-site behavioral information

The prediction model is periodically retrained to retrieve and consider recent visitor interactions and associated responses (or non-responses).

CONCLUSION

At SAS, we are passionate about customer journey optimization. It is exciting to see artificial intelligence approaches such as reinforcement learning emerging in product development cycles. We[8] will continue to innovate and support the evolving needs of marketers by embracing machine learning and artificial intelligence, because at the end of the day, analytics is the heartbeat of modern organizations.

REFERENCES

1. SAS Institute Inc. 2017. "Marketing attribution: Giving credit where credit is due." Accessed February 15, 2018. https://www.sas.com/en_us/whitepapers/marketing-attribution-108515.html.

2. SAS Institute Inc. 2014. "Customer segmentation using RFM in SAS Enterprise Guide." Accessed February 16, 2018. https://blogs.sas.com/content/sasdummy/2014/01/20/rfm-in-sas-eg/.

3. SAS Institute Inc. 2017. "How did one hospitality organization measure its marketing effectiveness?" Accessed February 17, 2018. http://players.brightcove.net/1872491364001/default_default/index.html?videoId=5401125978001.

4. SAS Institute Inc. 2018. "Artificial Intelligence - What it is and why it matters." Accessed February 18, 2018. https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html.

5. SAS Institute Inc. 2015. "Making the case for algorithmic digital attribution." Accessed February 19, 2018. https://blogs.sas.com/content/customeranalytics/2015/12/08/making-case-algorithmic-digital-attribution/.

6. SAS Institute Inc. 2018. "Machine Learning – What It Is And Why It Matters." Accessed February 20, 2018. https://www.sas.com/en_us/insights/analytics/machine-learning.html.

7. SAS Institute Inc. 2016. "SAS® Analytics for Multi-armed Bandit Market Analysis." Accessed February 21, 2018. http://support.sas.com/resources/papers/proceedings16/SAS6581-2016.pdf.

8. SAS Institute Inc. 2017. "SAS® Customer Intelligence 360 Engage – A Marketer's Overview." Accessed February 22, 2018. https://www.youtube.com/watch?v=p4lwA4wC-fg.

RECOMMENDED READING

- *Learn SAS Customer Intelligence 360:* https://www.sas.com/en_us/learn/software/customer-intelligence-360.html

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Malcolm Lightbody
Principal Product Manager
SAS Institute, Inc.
Email: malcolm.lightbody@sas.com
Web: https://blogs.sas.com/content/author/malcomlightbody/

Suneel Grover, M.B.A., M.S.
Principal Solutions Architect / Professorial Lecturer
SAS Institute, Inc. / The George Washington University
Washington D.C.
Email: suneel.grover@sas.com
LinkedIn: https://www.linkedin.com/in/suneelgrover
Twitter: @suneelgrover
Web: https://blogs.sas.com/content/author/suneelgrover/