visualIQ® Revealing Marketing Truth™

Untangling Attribution's Web of Confusion: A Primer for Marketers

"To regard as arising from a particular cause or source" – that's how attribution is defined in the dictionary. Yet, after more than a decade since the first offering appeared in the marketplace, advertisers, analysts and vendors can't seem to agree on a common definition for the purposes of marketing.

Opposing Perspectives

Some of the confusion may be due to the complexity of the subject matter or, if you're a conspiracy theorist, the result of deep rooted agendas set by those involved in media buying/planning with millions of dollars at stake. Whatever the cause, the industry is more confused today than ever before. One doesn't need to look further than the recent publications released by Gartner and Forrester Research to find basic examples of competing perspectives that leave the industry confused.

One doesn't need to look further than the recent publications released by Gartner and Forrester Research to find basic examples of competing perspectives that leave the industry confused.

Let's first look at Gartner's perspective as defined within their Market Guide for Attribution and Marketing Mix Modeling, where they recognize multi-touch attribution (MTA) and marketing mix modeling (MMM) as completely separate – but potentially integrated – marketing endeavors. Importantly, in distinguishing between technology providers and consulting firms, Gartner successfully uncovers that the bulk of marketing mix modeling today is purely a consulting, service-based exercise. Vendors claiming to have a viable software product for MMM are, in fact, just publishing the results of their consulting efforts into a technology interface no faster than monthly – hardly the speed required by marketers to gain any substantive benefit, but such is the nature of the media strategies that are measured by MMM: broadcast, TV, OOH and print as examples.

Gartner's Perspective

Gartner identifies that MTA is focused on addressable media at the user level, but largely isolates that user-level media to digital only. They suggest that "MTA is most useful when a marketer has user-level data for both the media and the goal. For example, when most of the media is digital and the sale happens in a digital channel or offline channel, that eventually can be directly and decisively tracked to a user and her/his activity. On the other hand, MMM is most useful when user-level information is not readily available for the media and/or the sale, such as retailers and car companies whose media is both digital and offline (e.g. TV and print) and whose sales happen offline"[1]. Gartner further states that, "Increasingly, marketing leaders use both methods – either in parallel or unified within the same platform – to expand the scope and precision of their measurement."[1].

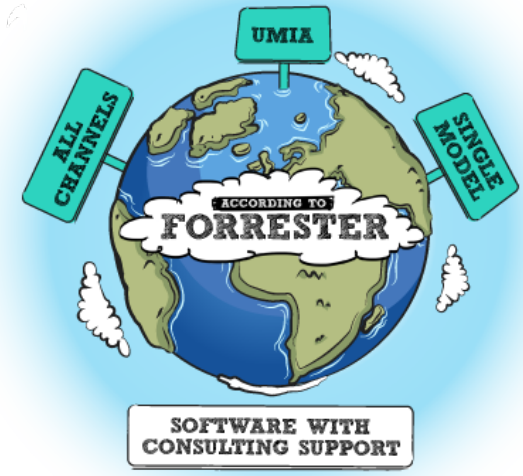
[1] Market Guide for Attribution and Marketing Mix Modeling, Gartner, September 30, 2016

Copyright © 2016 Visual IQ, Inc.

Forrester's Perspective

In contrast, the Forrester Wave™: Marketing Measurement And Optimization Solutions, Q4 2016 no longer covers attribution by name – as it has in the past – and instead focuses on the evaluation of "Marketing Measurement and Optimization" vendors attempting to create a new category called unified marketing impact analytics (UMIA). The criteria used for evaluating such vendors dramatically changed from previous reports on the topic of attribution due to Forrester's fundamental belief that vendors should have the ability to provide MTA and MMM using a single modeling approach. According to Forrester, "We have also seen that top-down techniques like marketing mix modeling and bottom-up approaches such as digital attribution share similar mathematical and statistical methodologies, encouraging measurement firms to test ways to combine them into a single technique Forrester calls unified marketing impact analytics (UMIA)"[2]. Forrester's view is based on the assumption that a single model can be used to effectively measure and optimize both addressable and non-addressable channels while providing a 'people-based' data architecture.

[2] The Forrester Wave™: Marketing Measurement And Optimization Solutions, Q4 2016, Forrester Research, October 11, 2016



Copyright © 2016 Visual IQ, Inc.

Most of the industry defines the de-duplication of users across multiple devices (desktop, mobile etc.) as 'people-based'. This definition generally gets expanded to mean the ability to establish a unique user identifier (UID) that can be used to sync with multiple data sources. However, Forrester's definition involves the association of all data down to the customer level even though the exposure of a given advertiser's marketing and/or media is not independently tracked at the individual user level, nor can it be applied at scale. Further, unlike Gartner's view, Forrester's Wave report does not delineate between technology providers and services providers, further blurring the lines between what marketers can expect to solve via technology, and what needs require a heavy dose of consulting services.

If leading industry analysts, investing months of research into the topic of attribution, release contradicting reports on approaches and associated capabilities, it should come as no surprise that the larger marketplace is having difficulty understanding vendor capabilities and mapping them to their business problems to drive meaningful results.

In an effort to eliminate this confusion, Visual IQ has provided additional details below to cut through the hype, expose the pros and cons of the different attribution approaches available, and take one step closer to defining attribution for the industry.

Marketing Attribution Approaches

The objective of marketing attribution is to apply accurate credit to different marketing and/or advertising tactics. Such tactics generally involve the execution of paid, earned or owned media across both online and offline marketing channels and increasingly involves multiple devices. Some of these channels provide addressable touchpoints – which means vendors are able to track them at the user-level. Digital channels most frequently come to mind when discussing addressable marketing, but there are other offline channels that can also be tracked and recorded at a user level. Other channels contain non-addressable touchpoints, which means that they cannot be tied back to a specific user with complete certainty. For example, it is impossible to know when a specific user has been exposed to a magazine ad. Regardless of the touchpoint type, attribution involves the measurement of multiple marketing tactics in an effort to optimize those that are delivering the best results.

Forrester's 'one-model-to-solve-all' approach is not only unrealistic, but it also oversimplifies the reality of how attribution can be used by advertisers.

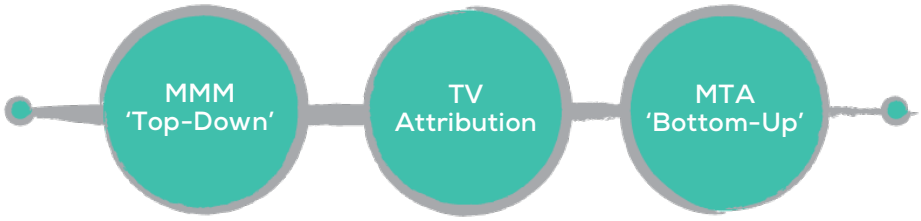
Levels of Granularity

It is important to understand that marketing and advertising tactics have different levels of granularity and that different attribution approaches have varying degrees of sophistication. For example, understanding yesterday's top performing display ad size for your business using in-store sales as your key performance indicator (KPI) is very different from attempting to understand the impact of each of your channels on a quarterly basis using total sales as your KPI. This is the reason why multiple attribution approaches exist and are utilized today.

The questions that these approaches answer tend to differ dramatically. As pointed out by Gartner, sophisticated advertisers use these approaches in combination to increase accuracy and maximize results.

Copyright © 2016 Visual IQ, Inc.

However, once advertisers truly understand the approaches in detail, very few believe Forrester's assumption that a single attribution model can meet all of their business objectives. Forrester's 'one-model-to-solve-all' approach is not only unrealistic, but it oversimplifies the reality of how attribution can be used by advertisers. Let's investigate:



1. **Marketing Mix Modeling (A.K.A 'Top-Down')** is actually a form of marketing attribution. Top-down attribution involves the collection of large amounts of summary level historical data in an effort to reallocate budgets across channels and sub-channels using various forms of logistical regression modeling. The data collected is often provided in numerous formats, which requires the vendors to either manually scrub and normalize the data, or require that it all be provided by the marketer in a common format. This approach also allows for the modeling of exogenous data – essentially understanding the impact of factors that fall outside of an advertiser's control, such as weather, economic conditions and approaches taken by competitors.

   The benefit to top-down attribution is that it can be used across both online and offline channels to incorporate an advertiser's entire marketing mix. In fact, as Forrester points out, there are a growing number of vendors that attempt to use the same modeling approach for both top-down and bottom-up (see below) using summary level data instead of actually tracking user exposure to marketing. The downside to top-down, and using it to address bottom-up, is the level of granularity, speed and consulting involved. Modeling occurs at only the channel/sub-channel level and on an infrequent basis (typically quarterly), and requires the involvement of data scientists (i.e. heavy services/consulting). This dramatically impacts an advertiser's ability to optimize its marketing, as all recommendations occur at a summary level, with no granular understanding of tactical marketing performance.

Modeling occurs at only the channel/sub-channel level and on an infrequent basis (typically quarterly) and requires the involvement of data scientists (i.e. heavy services/consulting).

2. **TV Attribution** starts by establishing a baseline for digital channel performance using bottom-up attribution (see below) without TV present in the marketplace, and then quantifies the incremental impact of TV impressions on digital responses using logistical regression modeling. Most vendors can provide fractionally attributed metrics by TV dimension including network, program, spot length, creative, geography, etc. Since most advertisers view TV as a top-of-funnel activity heavily focused on brand-related marketing, TV attribution is an effective way to understand how TV drives digital brand engagement KPI's like search queries and site visits, as well as final conversions. The modeling cadence for TV attribution can be relatively quick (weekly) and once implemented, uses an automated SaaS technology to eliminate manual modeling practices.

Copyright © 2016 Visual IQ, Inc.

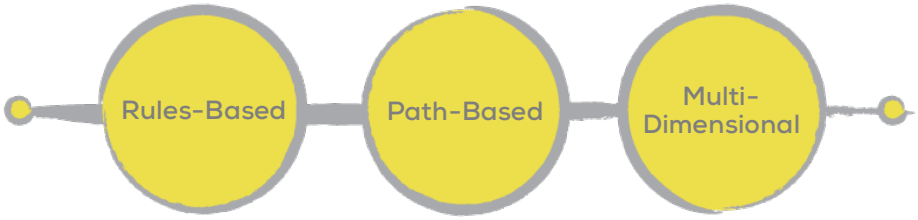
The pitfall with this approach occurs when some vendors attempt to link TV impressions to actual users based on their digital behavior after a TV spot is aired. If a user engages with a digital KPI like a search query, these vendors essentially assign a TV touchpoint within their marketing path. Unlike bottom-up attribution, which actually tracks a user traversing across marketing touchpoints, these vendors make statistical assumptions about a user's exposure to a TV impression, and then provide a path-based bottom-up attribution approach to produce optimization recommendations at the channel/sub-channel level, in a similar manner to top-down attribution. Unfortunately, the accuracy of these solutions dramatically decreases when attempting to optimize at more granular levels of media, which can negatively impact results (see 'path-based' below for more detail).

The challenge with bottom-up is that not all marketing tactics are addressable, which means that fractional credit is only allocated to the marketing tactics it has the ability to track.

3. **Multi-Touch Attribution (A.K.A 'Bottom-Up')** is an attribution approach that focuses on addressable channels and solves a problem that top-down cannot address. A unique identifier (UID) is used to track users across channels to understand the effectiveness of marketing tactics at extremely granular levels (keyword/placement). This happens at a much faster modeling cadence, with a full rebuild of the model occurring daily. Bottom-up attribution allows for multiple types of online and offline KPI's to be established by syncing unique user identifiers. Data is then automatically ingested and modeled using SaaS technology with optional levels of ongoing support post-implementation. Fractionally attributed data can be integrated directly into programmatic buying platforms to allow for near real-time optimization, and UID's can be synced with a host of different 1st and 3rd party data sources.

   The challenge with bottom-up is that not all marketing tactics are addressable, which means that fractional credit is only allocated to the marketing tactics it has the ability to track. In addition, unlike top-down which uses a common modeling approach across vendors (logistical regression) – the level of sophistication of bottom-up approaches can differ dramatically.

Different Bottom-Up Approaches



- **Rules-Based**
  This form of bottom-up attribution uses subjective rules to allocate credit to marketing touchpoints. Generally, these rules focus on the user's journey across marketing touchpoints, leading to a desired KPI and include first click, last click, even weighting, starter/player/closer, and U-shape (i.e. allocating credit to both the first and last touchpoints). In all cases, marketing tactics are measured based on the approach chosen by the marketer instead of the utilization of machine learning. This type of attribution tends to be observational and directional in nature due to the lack of sophistication, allowing for only basic forms of marketing optimization.

Copyright © 2016 Visual IQ, Inc.

- **Path-Based**
  Path-based attribution algorithmically compares user journeys across media touchpoints to isolate the fractional influence that a given touchpoint has on a specific KPI. While it is more sophisticated than rules based, in order to work effectively, the path-based methodology requires users to experience similar touchpoint paths so they can allocate fractional credit to any single touchpoint. This approach tends to be more effective when attempting to measure the impact of broader marketing channels or sub-channels, similar to top-down. This is because finding enough statistically significant user paths to compare at the marketing channel level (display, search, email, affiliate, etc.) is much easier than finding and comparing users who engaged with the same types of granular media dimensions (i.e. a specific email creative, search keyword, display size, social tactic, etc.). Fractional credit also tends to be applied only to the touchpoints involved in a user's path to conversion, instead of understanding the effectiveness of all marketing touchpoints, independent of the conversion path.

Fractional credit also tends to be applied only to the touchpoints involved in a user's path to conversion, instead of understanding the effectiveness of all marketing touchpoints, independent of the conversion path.

Accuracy Issues Arise

Accuracy issues arise when path-based vendors attempt to allocate fractional credit to granular media dimensions, since this approach prohibits modeling at those levels. For example, due to the inability to find common user paths to compare at granular levels (i.e. data scarcity), user paths tend to be modeled at the channel level only. Using one of many modeling approaches to overcome data scarcity (logistical regression, bayesian, game theory, ensemble, etc.), path-based attribution vendors cascade the influence of a given channel down to all of the media dimensions within it. This means that all campaigns, publishers, placements, creatives, sizes, etc., within the display channel receive the same attribution weight even though there could be major performance differences between these media dimensions. Publishers all perform differently (e.g. ESPN vs. AOL vs. Yahoo) and placements across each publisher derive different results. The problem with a path-based methodology is that advertisers have no idea what has actually been modeled versus what has been inaccurately assumed.

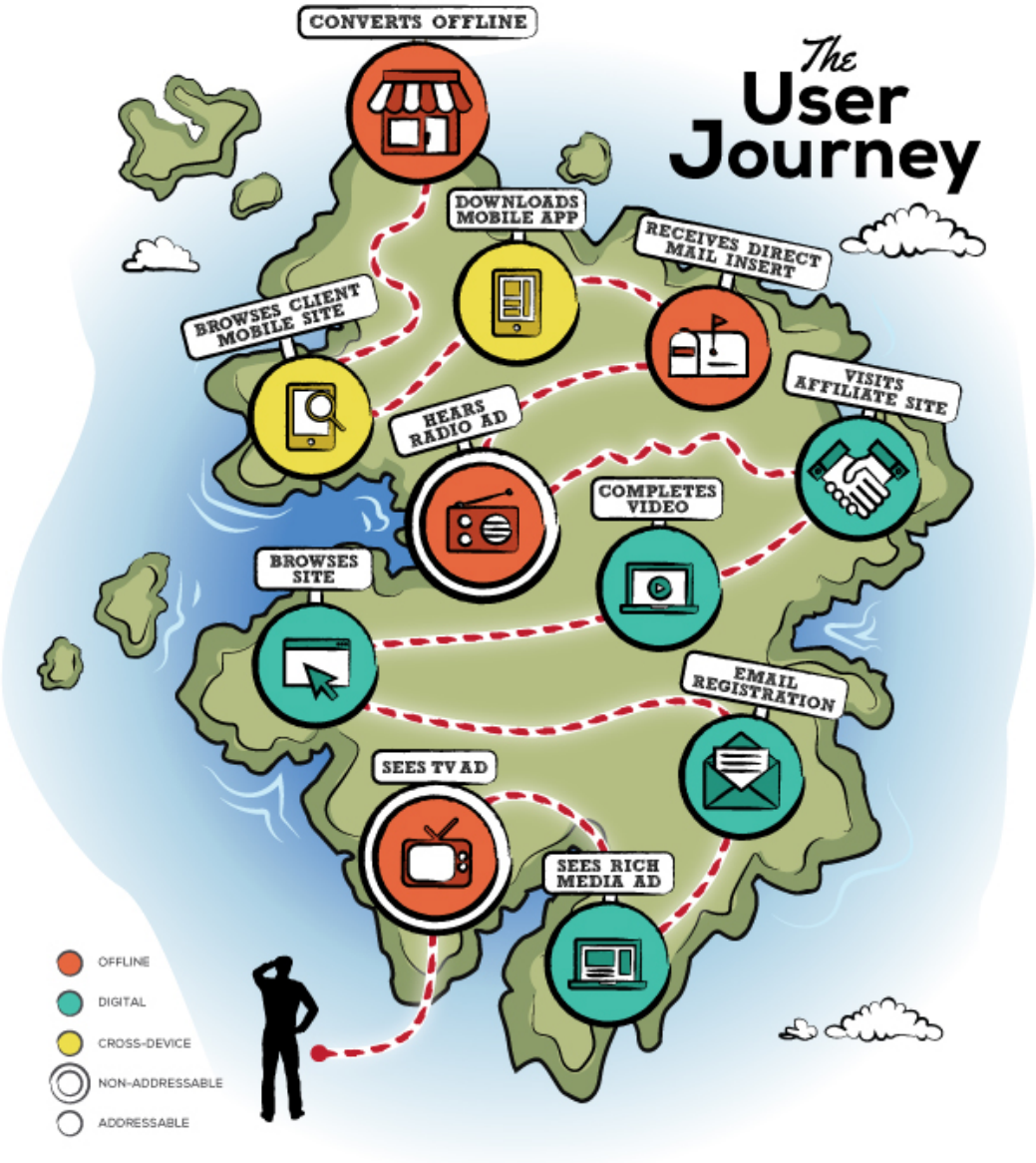
Unfortunately, advertisers cannot control a user's path or the order of media exposure prior to completing the desired KPI, which renders the value of the media dimensions inaccurate when attempting to optimize in isolation.

Further, the fractional credit being assigned to each of these media dimensions is dependent on the presence of other channels in the user's path. Using this approach when attempting to optimize a single media dimension, like a specific search keyword, advertisers would have to force consumers to see the other media touchpoints in a given path in order for the value assigned to the search keyword to be accurate. Unfortunately, advertisers cannot control a user's path or the order of media exposure prior to completing the desired KPI, which renders the value of the media dimensions inaccurate when attempting to optimize in isolation.

As a result, path-based attribution vendors are rarely able to use cross channel insights to provide accurate intra-channel optimization recommendations. Scenario planning and marketing recommendations tend to focus on shifting budgets at the channel/sub-channel levels only.

Copyright © 2016 Visual IQ, Inc.

- **Multi-Dimensional**
  Multi-dimensional attribution is similar to path-based attribution in that it establishes a single UID and tracks a user's path across addressable marketing touchpoints that lead to the intended KPI, either online or offline. However, instead of comparing user journeys to algorithmically assign fractional credit to marketing touchpoints, it models all the media dimensions of each touchpoint independently, and in combination, across all addressable channels (paid, earned and owned) to understand how exposure to each dimension changes consumer behavior. Media dimensions like campaign, publisher, placement, creative, size, etc., are common across most advertisers, while dimensions such as brand, geography, mobile device type, marketing tactic, etc., tend to be custom to specific advertisers based on their business model. There are multiple benefits to this approach.



Copyright © 2016 Visual IQ, Inc.

Multi-Dimensional Benefits

- Because of the exponentially larger data set, data scarcity tactics are not used. The fractionally attributed metrics are based on the actual modeling of each marketing dimension.
- Multi-dimensional modeling allows marketers to optimize media at any level of granularity. So rather than shifting budgets at the channel or sub-channel level, marketers can optimize each keyword, placement, creative, etc., using accurate metrics.
- Models can be rebuilt daily and applied to each new day's marketing performance data, instead of using a legacy model, which is common amongst path-based attribution vendors.
- Audience and cross-device data can be used to sync with the UID to understand the performance of media dimensions across different audience segments and devices.
- Fractional metrics can be easily validated across all dimensions using proven model validation techniques, or by testing dimension-level recommendations in market.

Multi-dimensional attribution is widely regarded by advertisers as the most sophisticated marketing attribution approach available today.

Multi-Dimensional Challenges

- Unlike more common statistical modeling approaches such as logistical regression or bayesian, multi-dimensional modeling is unfamiliar to most marketers and analysts, some of whom have difficulty articulating the methodology to stakeholders within their organizations. If stakeholder approval is required to execute marketing recommendations or there is organizational turnover, this lack of understanding can sometimes lead to optimization stagnation and a significant increase in opportunity cost for the brand.
- Similar to other bottom-up attribution approaches, multi-dimensional modeling is restricted to addressable channels only and is best utilized in combination with TV attribution and/or top-down attribution to ensure that offline marketing tactics being implemented by the advertiser are receiving accurate attribution credit.

Confusion

Unfortunately, marketing attribution is a complex topic that is difficult for not only advertisers, but even the analyst community to easily understand, due to the overall pace at which marketing technology evolves, as well as an unlimited supply of vendor hype.

The fact is that there are many instances where multiple vendors are being used simultaneously by advertisers because of the significant discrepancy in vendor capabilities, with each vendor providing a different level of value based on their heritage and the business objectives they are attempting to solve. For example, according to one advertiser, one of the vendors that achieved 'leader' status in the The Forrester Wave™:

Copyright © 2016 Visual IQ, Inc.

Marketing Measurement And Optimization Solutions is being used by its organization to understand campaign performance by audience demographic. Yet, that vendor was not even covered by Gartner's Market Guide for Attribution and Marketing Mix Modeling. To complicate things further, that same advertiser has combined a second and third vendor listed in the Forrester Wave to complete its business requirements.

Using a single modeling approach such as logistical regression across bottom-up, TV attribution and top-down limits accurate measurement to the channel/sub-channel levels and prohibits dimension-level optimization.

Recommendations

When selecting a marketing attribution vendor, it is paramount that advertisers understand proven vendor capabilities in detail, discount the future capabilities promised by vendors, and take larger marketing trends into account.

Relying solely on the competing perspectives published by Gartner and Forrester Research moves advertisers into dangerous territory where single-vendor solutions may not deliver what they promise and 'vendor hype' is, once again, used to describe the state of attribution. Let's review some of the key points advertisers should remember and explore:

- Unified attribution methodologies can be highly effective so long as the granular marketing accuracy isn't sacrificed. Using a single modeling approach such as logistical regression across bottom-up, TV attribution and top-down limits accurate measurement to the channel/sub-channel levels and prohibits dimension-level optimization.

- The market is slowly becoming more addressable, allowing for bottom-up to expand its attribution coverage. This is due to budgets shifting to digital and mobile, with less budget being allocated year-over-year to channels such as radio, print, outdoor, etc.

US Total Media Ad Spending Growth, by Media, 2015-2020
% change

	2015	2016	2017	2018	2019	2020
Digital	**20.4%**	**20.5%**	**14.9%**	**12.5%**	**11.0%**	**9.5%**
—Mobile	65.5%	45.0%	25.0%	20.0%	13.0%	11.5%
Out-of-home	**4.1%**	**2.6%**	**2.0%**	**1.5%**	**1.0%**	**1.0%**
TV*	**0.5%**	**3.5%**	**2.0%**	**2.5%**	**2.0%**	**2.5%**
Radio**	**-2.6%**	**-1.0%**	**0.2%**	**0.3%**	**0.2%**	**0.2%**
Print	**-7.5%**	**-6.1%**	**-2.5%**	**-0.8%**	**-0.3%**	**0.4%**
—Magazines***	-7.0%	-4.0%	-1.0%	0.5%	0.4%	0.4%
—Newspapers***	-8.0%	-8.0%	-4.0%	-2.0%	-1.0%	0.5%
Directories***	**-7.0%**	**-6.8%**	**-4.0%**	**-3.0%**	**-2.0%**	**-2.0%**
Total	**4.4%**	**6.9%**	**5.9%**	**5.8%**	**5.3%**	**5.1%**

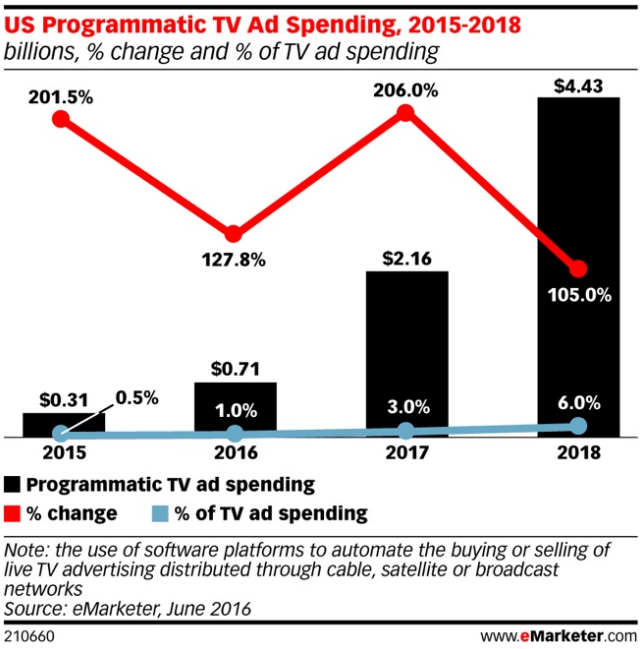
*Note: *excludes digital; **excludes off-air radio and digital; ***print only; excludes digital*
Source: eMarketer, Sep 2016

215849 www.eMarketer.com

- Similar to website analytics, bottom-up and TV attribution are now being offered as self-service approaches using a SaaS-based platform, augmented by the vendor's implementation and ongoing consulting services. This means that advertisers can become certified on vendor capabilities and develop their own attribution practice without involving data scientists or heavy vendor consulting services.

Copyright © 2016 Visual IQ, Inc.

- The digital and mobile marketing worlds are increasingly becoming programmatic with non-addressable channels such as TV also moving in the same direction, allowing for measurement of TV as a true addressable channel. The drive to programmatic creates an opportunity to distribute fractionally attributed metrics directly into programmatic platforms on a daily basis for tactical optimization at dimension-levels.



Sophisticated advertisers use attribution to optimize their marketing tactics at varying degrees of granularity to drive results.

- Today, major user-level data sources like Facebook are allowing a subset of media-agnostic attribution vendors to measure and optimize impressions, and gain access to its predominately deterministic cross-device data via partnerships with its Atlas subsidiary. Some vendors will leverage a UID to sync directly with Atlas for bottom-up attribution, while others will leverage Facebook's summary level data to provide a top-down attribution approach. It is important to understand the nuances of both approaches.

- There is no longer discord between understanding audience performance separately from cross channel marketing performance. Sophisticated attribution vendors now have the ability to sync with 1st and 3rd party audience sources (i.e. DMP's, CRM, etc.) to allow advertisers to understand the fractional performance of dimension-level marketing by audience segment.

- Finally, advertisers should not think of attribution purely as a measurement problem. Sophisticated advertisers use attribution to optimize their marketing tactics at varying degrees of granularity to drive results. This requires that attribution vendors provide a modeling cadence, RTB integrations and robust scenario planning capabilities that support that objective.

Copyright © 2016 Visual IQ, Inc.