

Big Digital Data, Analytic Visualization, and the Opportunity of Digital Intelligence

Suneel Grover, SAS Institute Inc., Washington D.C.

ABSTRACT

Digital data has manifested into a classic BIG DATA challenge for marketers who want to push past the retroactive analysis limitations of traditional web analytics. The current ground swell of digital device adoption and variety of digital interactions grows larger year after year. The opportunity for "digital intelligence" has arrived, as traditional web analytic techniques were not designed for the breadth of channels, devices, and pace that fuels consumer experiences. In parallel, today's landscape for data visualization, advanced analytics, and our ability to process very large amounts of multi-channel information is changing. The democratization of analytics for the masses is upon us, and marketers have the opportunity to take advantage of descriptive, predictive, and (most importantly) prescriptive data-driven insights.

This presentation describes how organizations can use SAS products, specifically SAS® Visual Analytics and SAS® Adaptive Customer Experience, to overcome the limitations of web analytics, and support data-driven integrated marketing objectives.

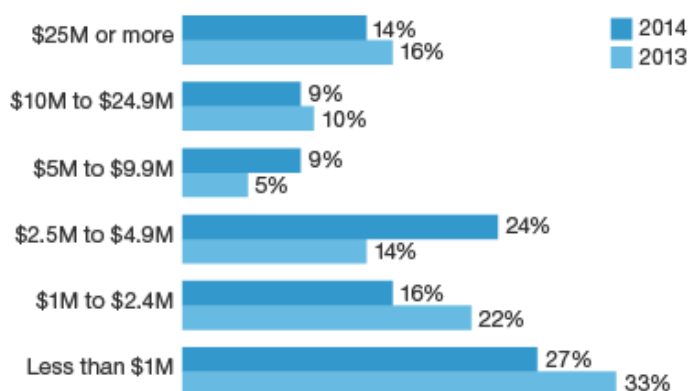
INTRODUCTION

Digital-savvy consumers today have high expectations when it comes to how brands interact with them. For example, they expect that the companies they engage with and buy from understand their preferences – regardless of the interaction channel. They expect them to provide a nonintrusive, yet personalized experience that adds value every step of the way.

According to Martin Gill at Forrester Research, digital business will become a reality in 2014. The question at the boardroom table is not "Should we do digital?" That battle has been won. The key question now is "How?" As senior executives wrestle with the dynamics of adapting their businesses to thrive in the age of the customer, digital business transformation will be firmly on the eBusiness agenda for 2014. ¹

"How much will your firm spend on its consumer-facing online presence in 2013 and in 2014 in US\$?"

(Please estimate as best you can [through the remainder of the year for 2013]; include development work as well as operating costs.)



Base: 128 eBusiness and channel strategy professionals
(percentages may not total 100 because of rounding)

Although a few high-spending firms are cutting back their eBusiness budgets, the vast majority of eBusiness professionals expect to have a higher budget in 2014 compared with 2013.

Source: Q4 2013 Global eBusiness And Channel Strategy Professional Online Survey

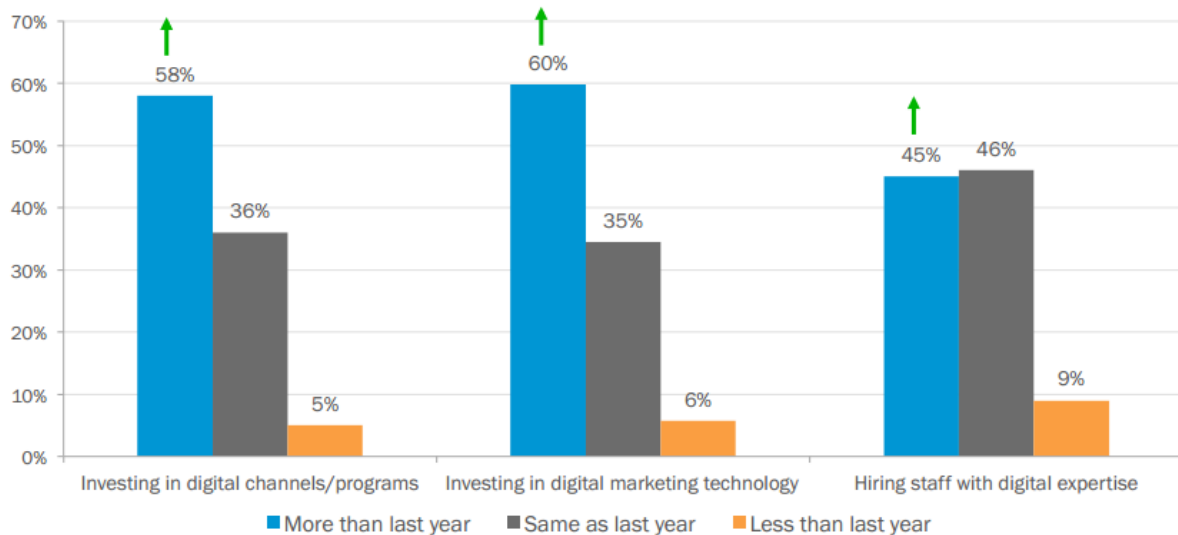
110449

Source: Forrester Research, Inc.

Figure 1. Predictions 2014—The Year of Digital Business

Research by Adobe in its September 2013 study “Digital Distress: What Keeps Marketers Up at Night?” validates this trend with the following insights:

1. Sixty percent of marketers expect their companies will invest more in digital marketing technology this year.
2. Only one in three marketers surveyed think that their companies are highly proficient in digital marketing.
3. Ultimately, 61% of all marketers think that, for most companies, digital marketing approaches are a constant cycle of trial and error.²



Q18. Please review the list of activities below and select how much more or less of each activity your company is doing compared to last year. Base: TOTAL (n=1017)

ADOBE | DIGITAL DISTRESS: What Keeps Marketers Up at Night?

Figure 2. DIGITAL DISTRESS—What Keeps Marketers Up at Night?

Further research by James McCormick from Forrester Research observes that, given the significant new investment and revenue at stake for digital-centric business objectives, analytics is a critical function to support the development, validation, and optimization of digital channels. And web analytics is the logical hub of interactive channel analysis, based on the historical centrality of the website. Unfortunately, most firms’ traditional approaches to web analytics are vestiges of a simpler time, when understanding traffic sources and on-site user behavior were of paramount importance. If customer intelligence professionals do not update their approach to digital analytics, they will continue to be plagued by

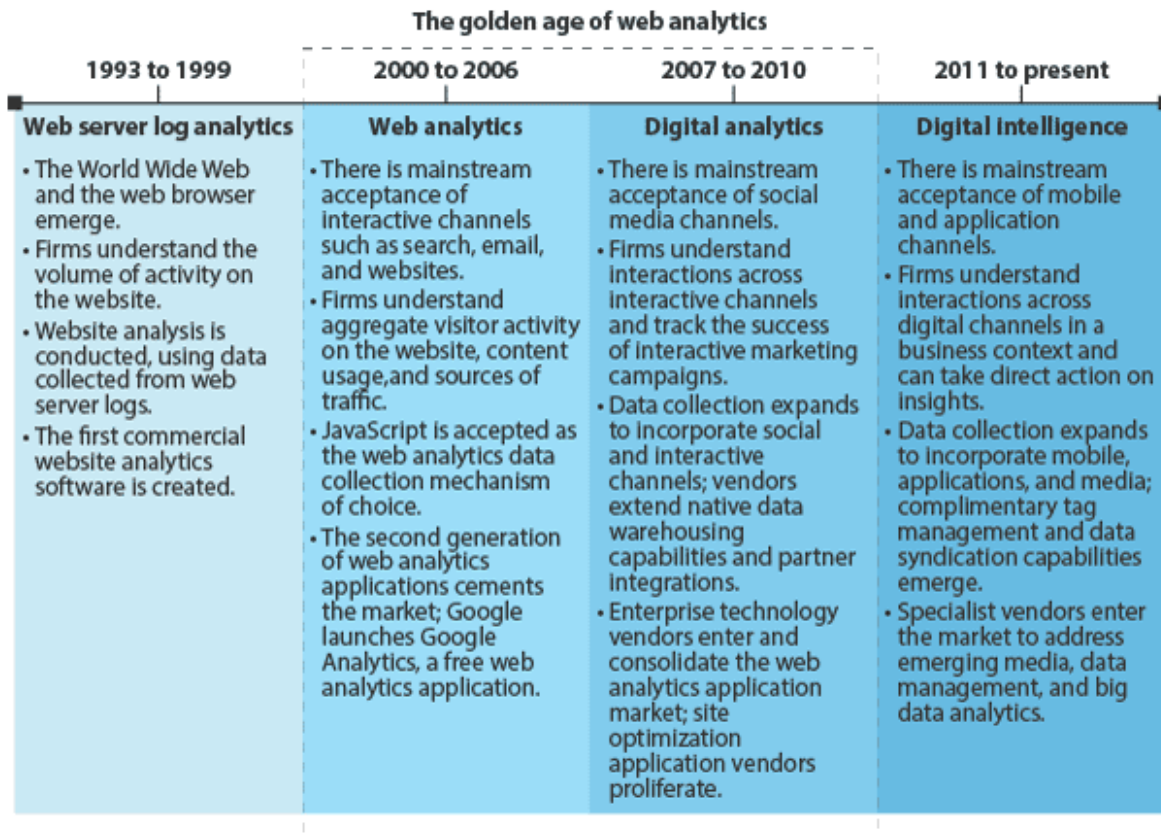
- Poor customer experiences
- Irrelevant business reporting
- Siloed customer insights.

To deliver comprehensive customer insights, firms seek to merge digital data with offline channels. Web and customer analytics teams are attempting to work together, but their projects struggle to get off the ground due to a clash of approaches and culture from

1. Data types — structured versus unstructured, known data versus anonymous
2. Skills — data scientist versus web geek
3. Analysis — advanced analytics versus “good enough” analytics
4. Time-to-delivery — best possible reporting versus satisfactory instantaneous reporting.

As marketing’s remit evolves to accommodate a growing catalog of intertwined touch points, so too must marketing analytics evolve. To close the gap between traditional web analytics and comprehensive analytics for digital marketing, organizations must adopt an updated approach. Forrester Research calls this “digital intelligence,” defined as

“The capture, management, and analysis of customer data to deliver a holistic view of the digital customer experience that drives the measurement, optimization, and execution of digital customer interactions.”³



61276

Source: Forrester Research, Inc.

Figure 3. The Journey to Digital Intelligence

A HISTORY OF FRAGMENTED DATA AND ANALYTICS TOOLS

Marketers have long had the challenge of stitching together different types of data and data sources to achieve on the vision of integrated marketing. This includes

- Internal first-party CRM and transactional data
- Digital analytics data from web analytics and advertising offerings like Google Analytics, DoubleClick, Adobe Omniture, or Right Media
- Third-party data from data management platforms like Nielsen, BlueKai, or X+1.

Unfortunately, it has been difficult to take advantage of such data sources because they have been stuck in silos. Marketers have struggled to bring together available online, offline, and third-party data in a way that is logical, efficient, and easy to use.

Until recently, many marketers depended primarily on various third-party web analytic tools designed to aggregate digital data to create reports that described what happened in the past. Obtaining an omni-channel, integrated view was extremely difficult across the fragmented digital ecosphere. As a result, it was practically impossible to get a data-centric, comprehensive view of the customer that could feed integrated marketing analytics, or more specifically, prescriptive recommendations for marketing processes.

While data-driven marketers and analysts have used powerful advanced analytics for many years to perform sophisticated analyses – such as regression, decision trees, or clustering – they have been limited to using offline data, primarily due to restrictions on access rights to online data from third-party technology vendors. On the flip side, web and digital analytics tools primarily aggregate and report on historical information and do not enable predictive analysis.

For the most part, web analytic tools are designed with the visualization of data as the primary driver for users, because data visualization enables a faster, deeper understanding of the insights and trends hidden within data in a more consumable manner. Their ease of use and visual appeal have helped marketers get a better understanding of the important trends and insights within data. However, data visualization largely has been very descriptive in nature – that is, primarily about reporting, business intelligence, and descriptive statistics.

INDIVIDUAL CUSTOMER-LEVEL DIGITAL DATA IS A VALUABLE COMMODITY

The challenge facing the marketing industry today is progressing beyond the multi-channel analytic limitations of aggregated data collection methods used by traditional web analytics. The opportunity is to have a digital data collection framework that enables both business intelligence and predictive analytics. This methodology requires organizations to collect data about

- Web, social, and mobile app sessions (that is, multi-domain visits)
- Accurately stitched-together digital visits to one anonymous visitor profile
- After authentication, accurately stitched together digital visits across devices, browsers, and multiple domains to one customer profile.

The vision of “digital intelligence” proclaims marketers focus on understanding the “who,” “what,” “where,” “when,” and most importantly, the “why” of digital experiences. However, this requires rethinking how you collect digital behavioral data, considering larger downstream business applications and use cases. For example, digital marketers love to perform website pathing analysis, which describes visitor page sequencing. What if marketers want to perform predictive analysis (like decision trees and regression) to identify what behaviors in a visitor’s digital journey are deemed most important in correlating with business conversion events? Other predictive digital marketing use cases include

- analytically forecasting website and mobile app visitation by traffic source, and identifying which ad-centric channels have the largest effect in predicting overall traffic
- predicting online and offline behavioral drivers of digital conversions using analytically driven segmentation techniques, and improving outbound and inbound targeting rules for future marketing communication and personalization efforts.

To succeed, this means getting data originating from the web or from mobile apps out of their traditional silos. This data, if collected and prepared appropriately, can be merged with your company’s first-party (or company-owned) customer data, and then streamed into your analytics, visualization, and marketing automation systems.

SOLUTION FOR COLLECTING, NORMALIZING, AND MAKING DIGITAL BEHAVIORAL DATA ACCESSIBLE WITH BUSINESS CONTEXT

According to January 2014 research performed by Celebris and Teradata ⁴, the biggest challenges facing digital marketing, data collection, and storage are as shown in the figure below.

What are the biggest data challenges facing your marketing team?

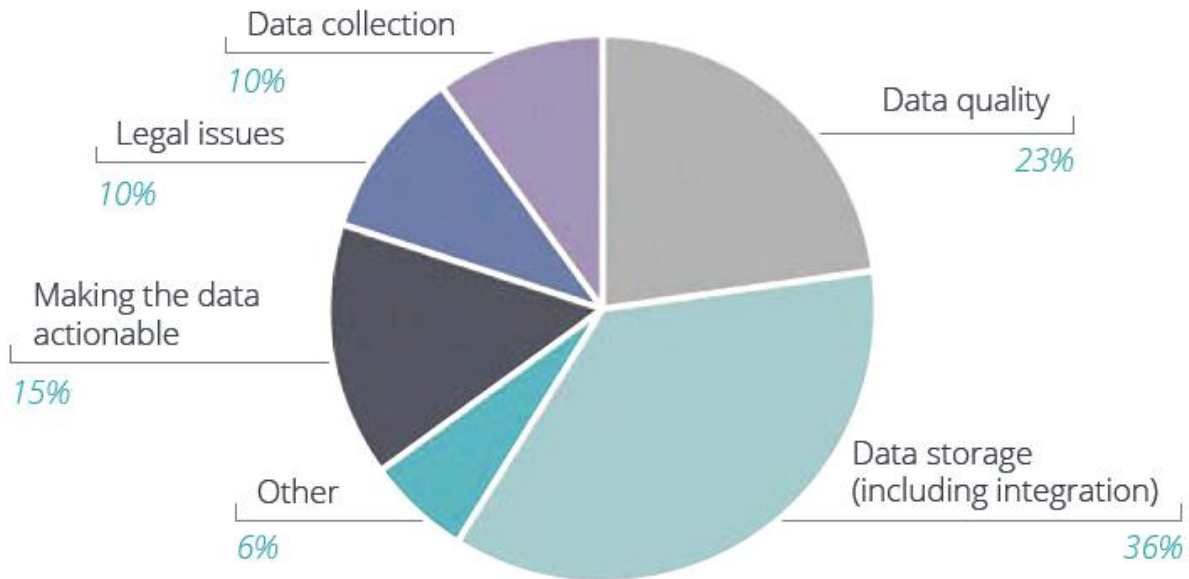


Figure 4. Biggest Obstacles Facing Digital Data Collection

What types of customer data do you collect?

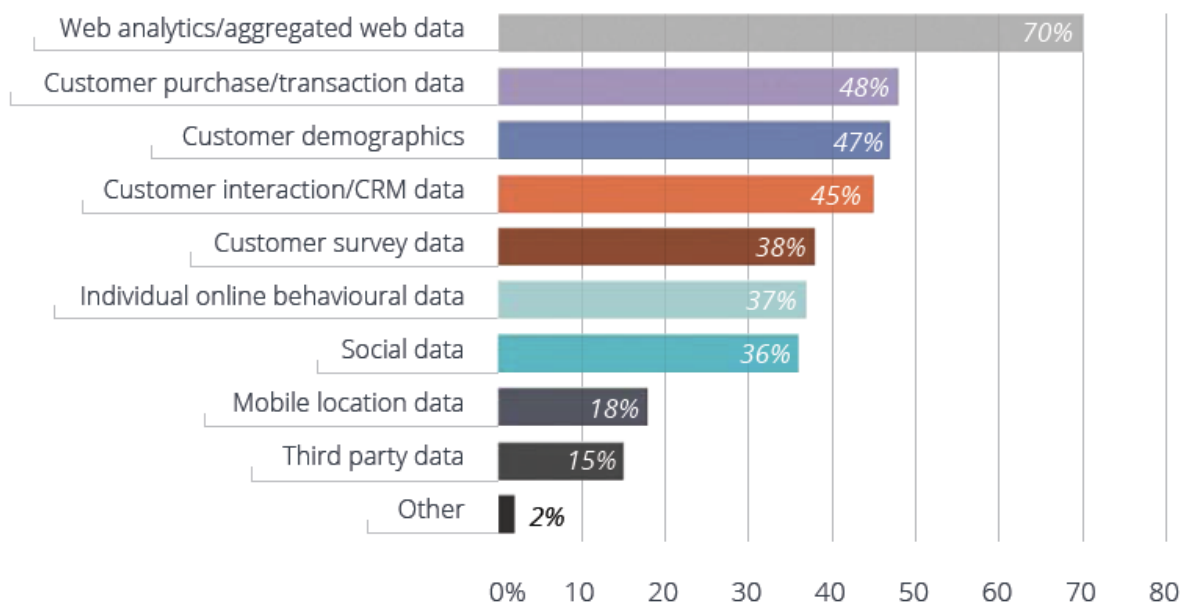


Figure 5. What Types of Customer Digital Data Do You Collect?

In the past few years, there has been a notable upswing in new and incremental abilities to process very large amounts of information. Data repositories – from Hadoop environments to traditional relational databases like SAP, Teradata, and Oracle – are getting bigger, stronger, and faster. Now that it's possible to handle very large amounts of information, we can approach digital data differently. As Figure 5 depicts, collecting web analytic data streams and preparing it for multi-channel integration is the biggest opportunity to focus on. Regretfully, aggregated web data is not the level of detail that can be utilized in predictive analysis, and cannot be de-aggregated.

SAS Adaptive Customer Experience is a digital marketing support solution that is designed to collect and feed big digital interaction data in the appropriate structure and context into an organization's Discovery, EDW, Visualization, Advanced Analytic, Marketing Automation, and Real-Time Decision systems. This solution provides the following capabilities:

THE COLLECTION OF DIGITAL INTERACTION BEHAVIOR VIA A SINGLE LINE INSERT

The single line insert (that is, Client Side Adaptor) records **all** online behavior down to the millisecond against the session, visitor, and the customer, providing an unprecedented level of insight into how your customers behave on digital properties – single and multi-domain – for websites, browser-based mobile apps, and social media branded pages. For iOS and Android native mobile applications, there is an API for the insert.

This approach has two key benefits as compared to traditional web analytic data collection.

1. The single line insert captures all interactions. This significantly reduces tagging management process challenges, and improves analytic potential downstream.
2. The single line insert does not need to be updated when new content is added to the website or app. Therefore, the issues with outdated tags, broken tags, and missing tags ceases to exist.

ADDING BUSINESS CONTEXT TO SEMI-STRUCTURED DIGITAL DATA

Normalization is the conversion of raw event digital data into useable data with business context. Through normalization, data is classified into different types of digital activities, each with its own attributes. With SAS Adaptive Customer Experience, this data is organized and appropriately structured to be fed into an analytical data mart.

Skilled data analysts understand that performing data normalization for multi-channel integration is no small task. Even the seemingly simple aspect of defining what a web session is has hidden complexities. For example, how do you define sessions that span multiple data-processing windows or web domains? Regardless of how data normalization is done, you need to define and set up a huge number of business rules to drive the classification process. Let's illustrate by discussing two examples.

- Let's assume that you want to understand origination in detail. For example, did a customer find your business via an external referrer, an organic or paid search link, a display ad, or some other method? Because of the nature of websites, browsers, and browser versions, the same origination (for example, a specific Google organic search) might be represented many different ways in your data. Data governance and business rules are necessary to capture all of the different permutations of this digital pathing, as well as unify them into a coherent customer search record with specific attributes.
- Maybe you want to identify which pathing activities are resulting in an unusually high number of abandoned sessions. You know that there is an online form that most customers do not take the time to fill out – but you do not understand why. To identify and address this, you need to know where most people disengage at completing the form. Looking at raw data events alone would indicate form activity levels – and you could have multiple records per form. A visitor could also try to fill in the form twice in the same session – something you would need to use business rules to discern.

DATA ROLL UPS: TYING DIGITAL DATA TO THE CUSTOMER

Traditional web analytic solutions and their respective data collection methodologies can populate a data mart with aggregated data for web business intelligence. This gives you aggregated information about your website, such as

- Information about goals, transactions, and other types of activities
- Shopping basket information
- General details about visitor origination.

SAS Adaptive Customer Experience allows organizations to normalize this intelligence to individual customer records in the appropriate structure and format to utilize in data mining, predictive analytics, forecasting, and optimization. Without this, digital marketers are severely limited in terms of the level of detail to which they can

perform analysis. The normalization process requires adapting data management approaches and governance to address the unique nuances that come attached with digital behavior streams.

DATA MANAGEMENT: ENSURING COMPLETENESS AND QUALITY

SAS Adaptive Customer Experience utilizes SAS Data Management technologies to address the identification, matching, and consolidation of digital data. Together, these capabilities provide the appropriate context and level of accuracy for one-to-one customer analytics, marketing, and relationship management.

For example, given the many channels and devices that customers can use to interact with your organization, identity management is a complex problem for Marketing and IT Departments to tackle. How do you know if a given set of anonymous website visits from one IP address involves the same or different human being? Sometimes it is not possible to be sure, but sophisticated matching that uses retrospective processing can maximize effectiveness and accuracy.

SAS Adaptive Customer Experience provides capabilities for synchronizing data, eliminating duplicates, and tying data from current customer web or mobile app visits to data from past sessions to allow a complete view of customer behavior over time. Effective data management also involves ensuring data quality; for example, it is crucial that customer records do not become corrupted through false matching data. Good data quality improves the performance potential of marketing analytics and visualization.

DIGITAL DATA MODEL AND DATA OWNERSHIP: BREAKING DOWN THE WEB ANALYTIC SILO

SAS Adaptive Customer Experience provides an open, configurable, and fully documented ETL data model that consists of 56 structured tables for session, visitor, and customer level views. The information is stored in a digital data mart that can be accessed by any SQL query or SAS tool for querying, reporting, and analysis – such as SAS Visual Analytics or SAS Enterprise Miner.

The data model and organizational ownership has three key benefits as compared to traditional web analytic data collection.

1. The organization owns the digital data stream and can exploit it in any way they desire. There are no limitations as compared to a web analytic vendor owning and limiting access to the underlying data.
2. The data model is configurable, and can adapt to an organization's unique set of digital properties, objectives, and goals.
3. The data model significantly reduces the time analysts spend accessing and preparing data for downstream analytics and marketing. The model enables an analyst to join customer-level digital data with any other source of internal or third-party data, and efficiently progress in their workflow.

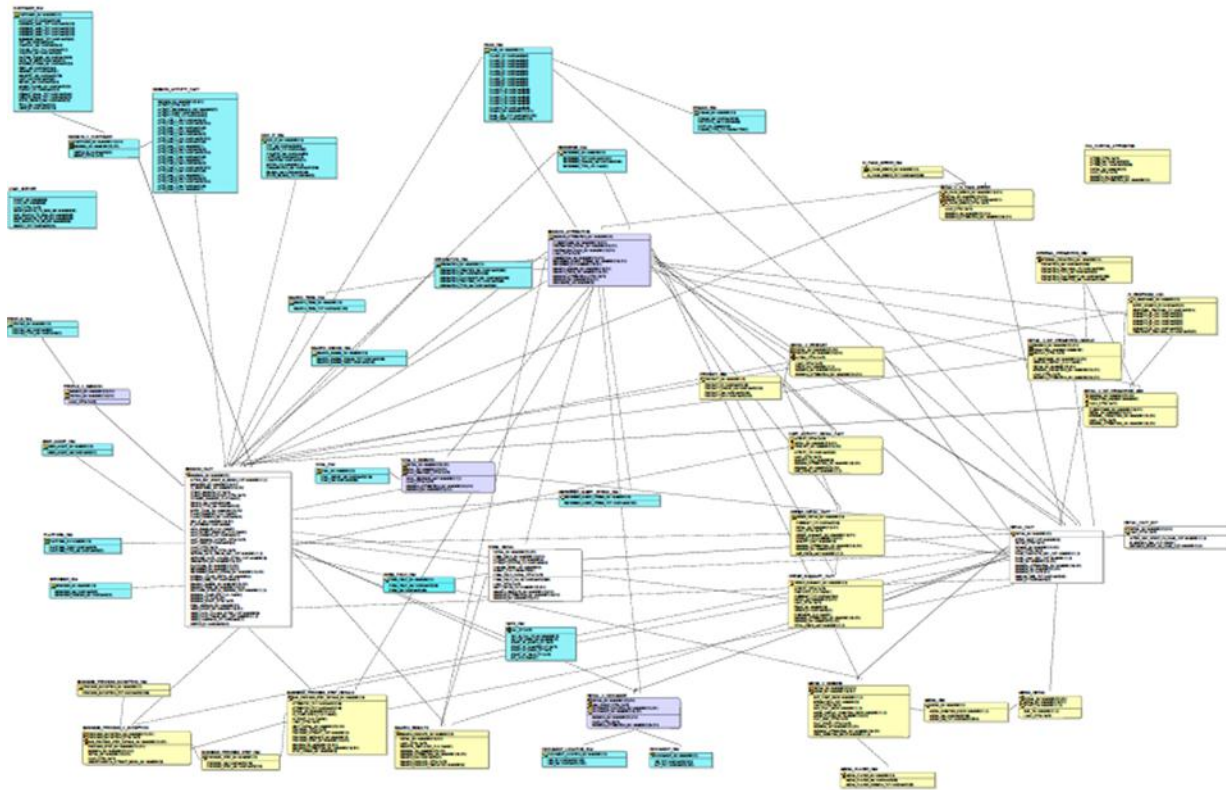


Figure 6. SAS Adaptive Customer Experience ETL Data Model

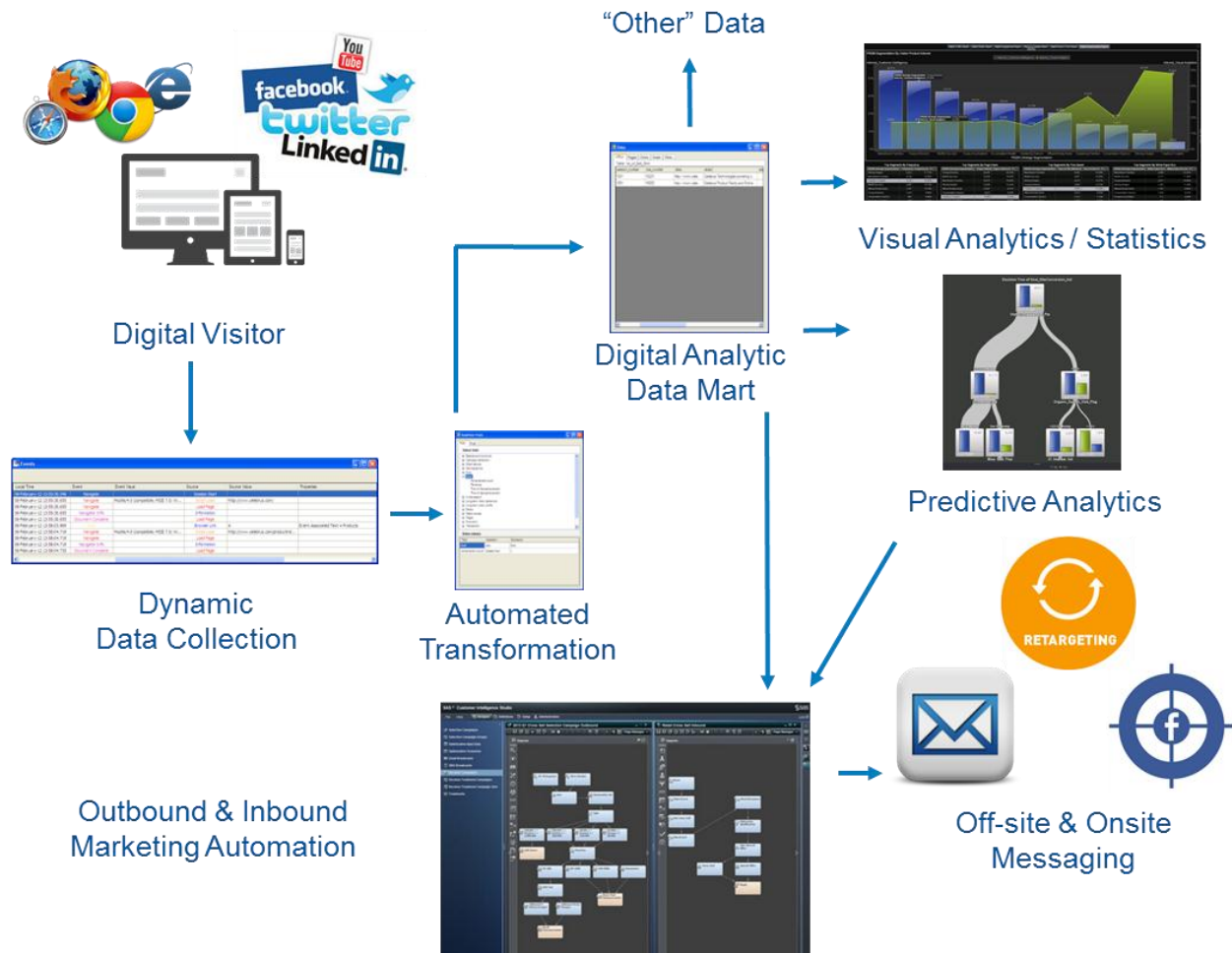


Figure 7. Visualizing SAS Adaptive Customer Experience

SAS VISUAL ANALYTICS ENABLES PRESCRIPTIVE ANALYSIS FOR DIGITAL AND INTEGRATED MARKETING

Predictive analytics and exploratory data mining thrive on detailed data. When we can bring together very granular digital data streams that highlight consumer behavior and feed that into visual predictive models, we can improve our approaches to segmentation, personalization, ad targeting, and customer experience management.

Today we have an attractive opportunity to watch predictive analytic and visualization technology mesh together through SAS Visual Analytics. The biggest value of this data-agnostic, advanced visualization platform is that it allows users to see predictive insights that could never before be seen using traditional web analytic tools. As the famous mathematician John W. Tukey said in his 1977 book *Exploratory Data Analysis*, “The greatest value of a picture is when it forces us to notice what we never expected to see.”

SAS VISUAL ANALYTICS AND PREDICTIVE MARKETING – DEMO #1: VISUAL ADVANCED FORECASTING

To illustrate how advanced visual analytics can help organizations improve their approach toward digital intelligence, let’s go through an example of analyzing digital visitation to SAS.com (collected by SAS Adaptive Customer Experience, which utilizes SAS Visual Analytics as the front-end user interface of the solution offering), both from a historical and predictive perspective.

HISTORICAL VIEW

Suppose that a manager asks, “What did our web traffic look like over the last few months?” We can get the answer in just a couple of clicks by assigning the Visitor Date and Visit Identifier elements to the visual.



Figure 8. Historical Traffic Visitation Pattern

PREDICTIVE VIEW

Now suppose the department manager asks, “What’s going to happen to web traffic in the next two weeks?” In one click, we can show a forecast of expected site traffic of any duration – no coding required.



Figure 9. Web Traffic Forecast

What's more, SAS Visual Analytics uses champion-challenger forecasting. The tool uses six different forecasting algorithms to fit the data in near-real time, and the algorithm that is most statistically accurate is selected for the visualization. In other words, you get the most accurate result, no matter what your quantitative skill level is.

IMPROVING THE FORECAST

Users can improve how this model predicts future website traffic by providing more information from which it can learn. In Figures 8 and 9, the visualization only represented visitors by date. Now we will add more data elements describing the originating visitor traffic sources – paid search, organic search, social, blog, affiliate, and direct visitors who came to SAS.com without the stimulus of an advertisement.



Figure 10. Adding Underlying Factors to Improve the Accuracy of the Forecast

By adding these segments to the forecast model's consideration, we can see that the confidence interval (that is, best- and worst-case scenarios) of the prediction gets much tighter, showcasing accuracy improvement in the model's prediction compared with the earlier iteration. In addition, significance testing identifies which segments have an impact on the prediction.

Below the forecast line graph in Figure 10, SAS Visual Analytics provides the ability to perform scenario analysis, sometimes referred to as *simulation* or *what-if analysis*. The model identified that two segments – organic search and paid search – had a significant effect on the forecast. Users can now simulate inflated and deflated effects of these independent variables.

As a digital marketer – and more specifically, a digital advertiser or media planner – you have a limited amount of control over organic search traffic. You have more control over paid search, which is an ad-centric channel. What if we increased our paid search advertising budget by varying amounts? What effect would that have on overall site traffic? That is actually very easy to answer.

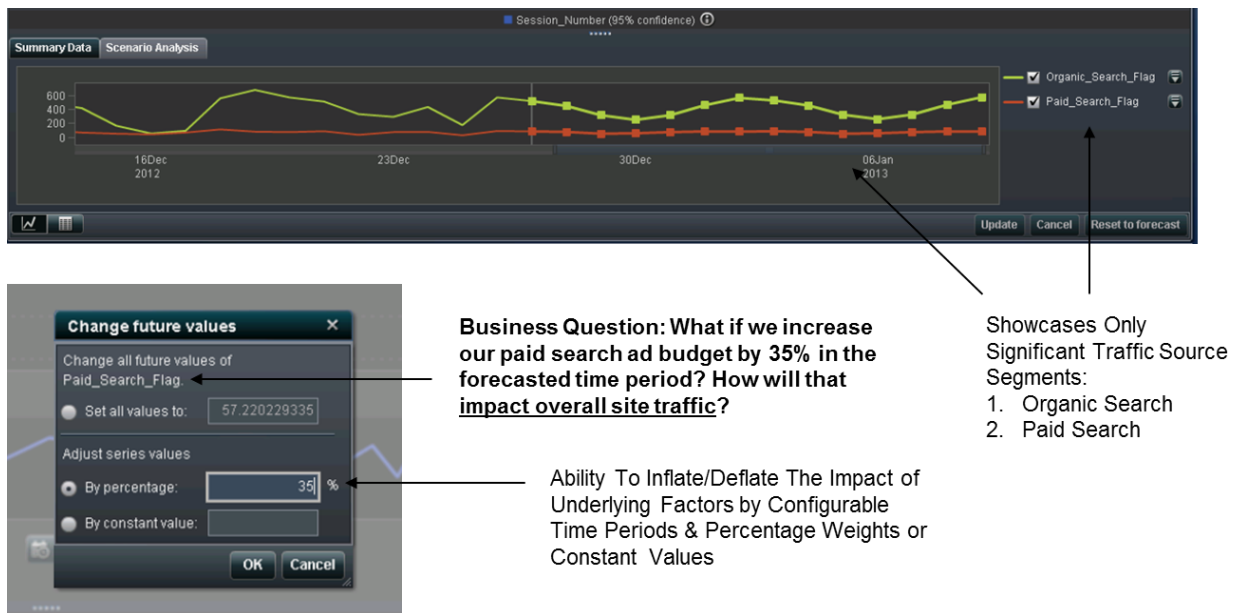


Figure 11. Inputs for Digital Advertising Simulation

In Figure 11, we have simulated a 35-percent increase in paid search advertising. Let's see how this change will affect the traffic pattern forecast for the entire website. With today's ever-changing ad budgets and short time windows, having the ability to simulate increases or reductions in ad spending in different marketing channels can be very valuable.

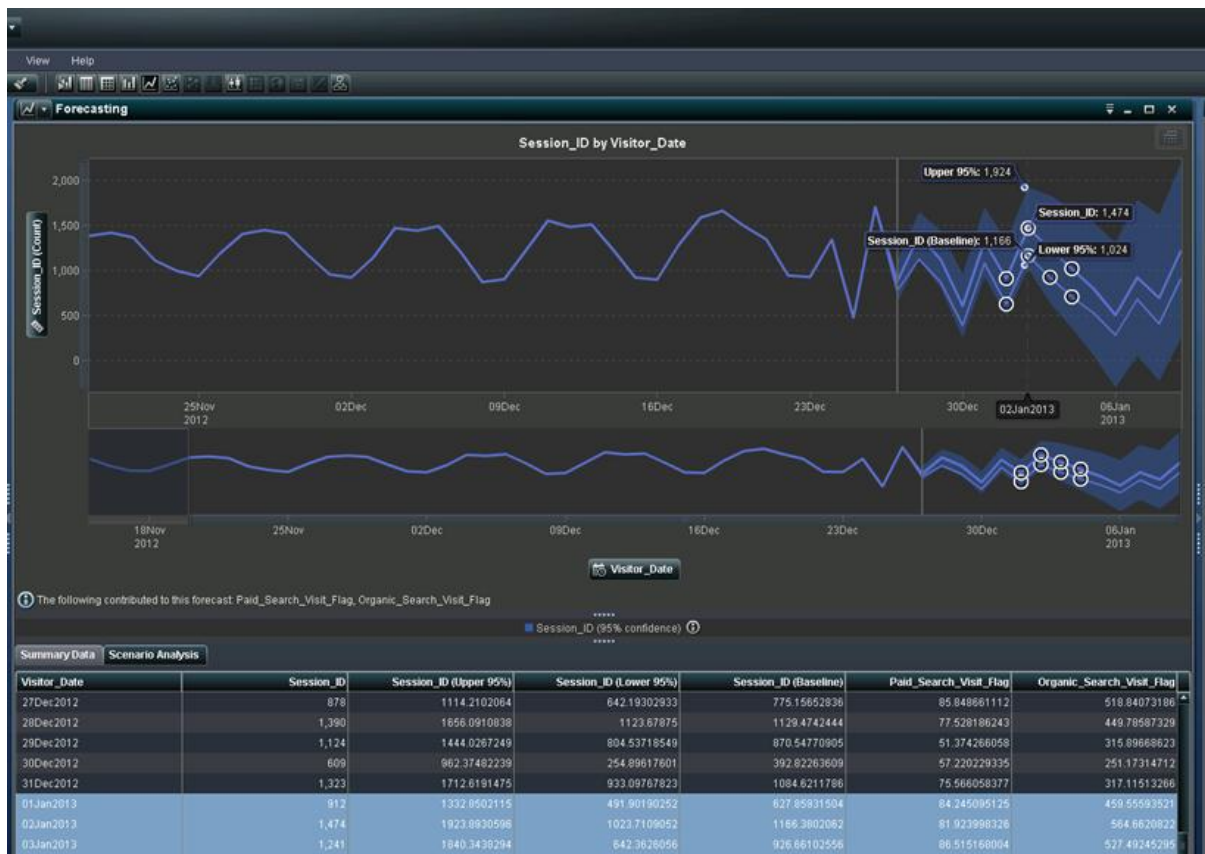


Figure 12. Forecast Simulation Visualized

Now we have two numbers representing website traffic for January 2 in Figure 12. The baseline was the original prediction – 1,085. If we increase paid search by 35 percent, we can expect 1,323 visitors to the site. That means that a 35-percent increase in ad spending on paid search is predicted to produce a 22-percent increase in overall traffic over the next two weeks.

Based on how your organization manages budgets and decisions, you could explore different what-if scenarios. For example, you could determine if the impact of increasing paid search advertising by 25-percent or 45-percent would be worth the investment. This would be valuable information, indeed, for a decision-making leader within the organization.

SAS VISUAL ANALYTICS, DIGITAL SEGMENTATION, AND OUTBOUND MARKETING – DEMO #2: VISUAL DECISION TREES

To illustrate how visual decision trees can help organizations improve their predictive marketing approach toward analytically defined segmentation and data-driven campaign management, let's review a second example of analyzing digital visitation to SAS.com, and identify an attractive audience for a future marketing communication.

WHAT DRIVES CONVERSIONS ON SAS.COM?

Suppose that a manager asks, “What are the most important factors that differentiate visitors who convert and do not convert on our website?” This is the perfect hypothesis for a supervised predictive model.

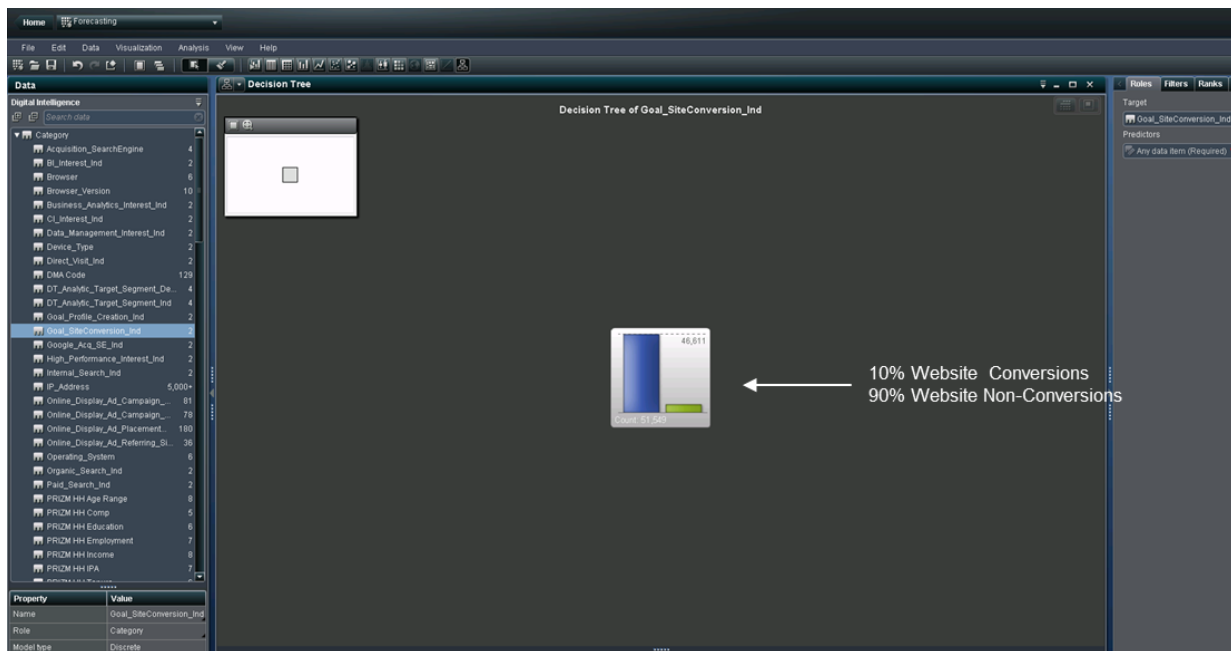


Figure 13. Assigning the Target Variable

After our target variable has been added, we observe 90% of visitors did not convert, and 10% of visitors did convert. Users can now visually select potential predictors to identify unique characteristics of higher-value and lower-value audience segments utilizing an advanced data mining algorithm that intensively iterates through thousands of potential scenarios before arriving at a statistically supported answer.

Figure 15 provides one example of how the decision tree analyzed the parent population of digital visitors, and delivered insights into important behaviors that help explain why this specific segment is attractive. Specifically, visitors to SAS.com who showcase high scores of engagement, and originated from an organic search before arriving, showcase higher likelihoods of converting as compared to their peers who displayed other behaviors. It is the interaction of these two characteristics together that drives this result. The next logical step is to take action on this insight.

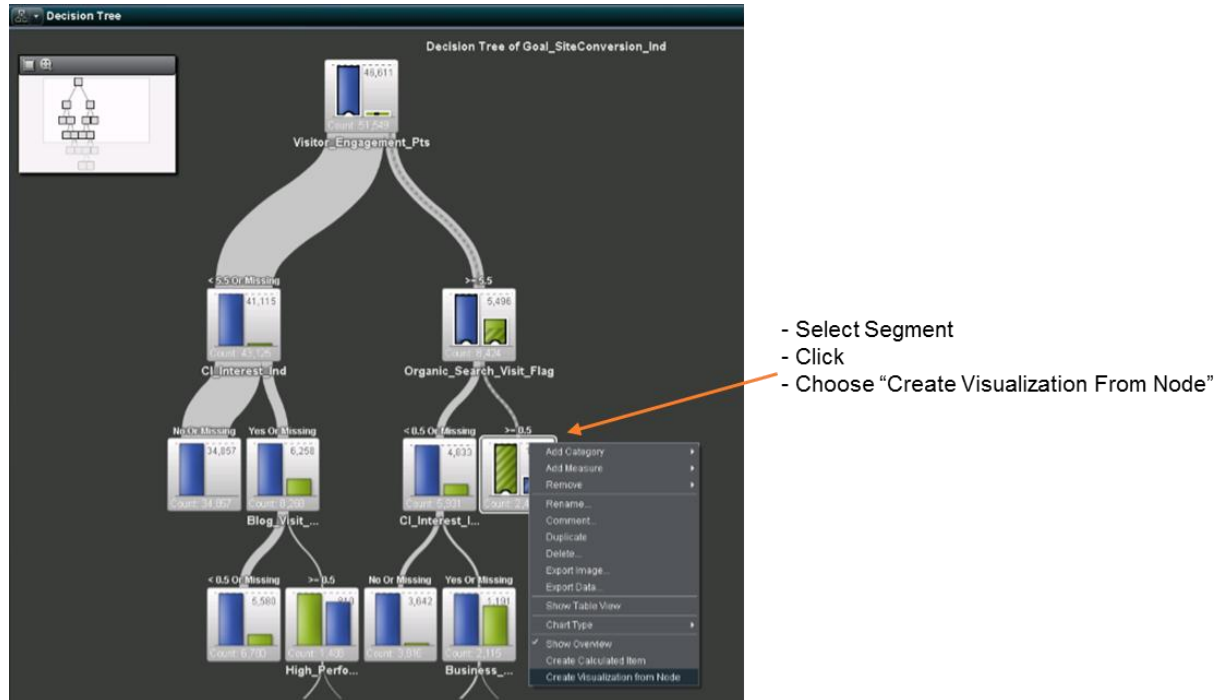


Figure 16. Visual Segment Selection

Users have the ability to select unique segments, and subset the audience to a new visualization. This is empowering as no programming or coding is required, and the intelligence of the algorithm to define this audience segment will carry over to the next step within the filter. In addition, marketers can easily filter out members of this segment who have achieved conversion, and focus on the look-a-like prospects who are showcasing conversion signals, but have not crossed the finish line.

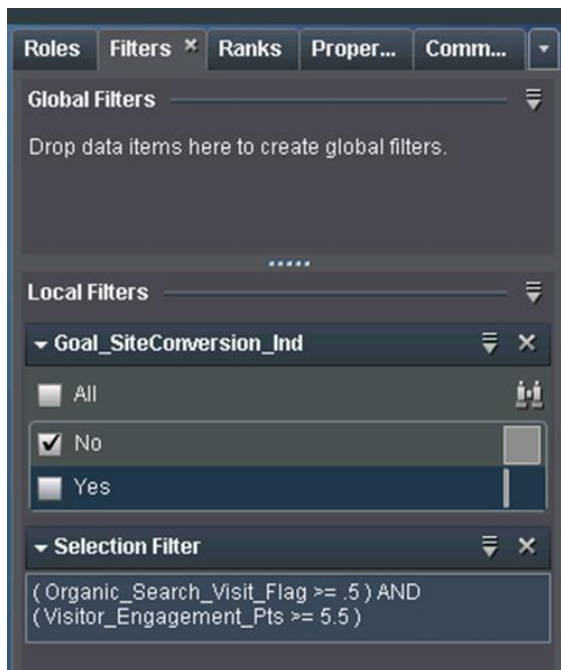


Figure 17. Prepopulated Filter

The last step for the user is to create a targeted audience list. After selecting the variables, this attractive segment can now be used in a marketing automation system, real-time decision platform, or communication optimization exercise.

Intelligen...							
sis View Help							
Visualization 1							
Customer_ID	City	State_Region	Country	Visual_Analytics_Interest_Ind	Data_Management_Interest_Ind	Business_Analytics_Interest_Ind	
10090457	Mumbai	Maharashtra	India	No	Yes	No	
10089443	Cary	North Carolina	United States	No	No	Yes	
10089126	Milton	Massachusetts	United States	Yes	No	No	
10088842	Brisbane	Queensland	Australia	No	Yes	No	
10087559	Stamford	Connecticut	United States	No	No	No	
10086818	Laurel	Maryland	United States	No	Yes	No	
10086272	Pittsburgh	Pennsylvania	United States	No	Yes	No	
10086189	Frederick	Maryland	United States	No	Yes	No	
10085741	Wichita	Kansas	United States	No	No	No	
10085471	Bowling Green	Ohio	United States	No	No	No	
10085470	Santiago	Region Metropol...	Chile	No	No	Yes	
10083995	Mill Valley	California	United States	No	No	No	
10083918	Izmir	Izmir	Turkey	No	Yes	Yes	
10082933	Lisbon	Lisboa	Portugal	No	No	Yes	
10082856	Victor	New York	United States	No	No	No	
10082295	Unknown	Unknown	United States	No	Yes	No	
10082142	Unknown	Unknown	Japan	No	No	Yes	
10080435	Zebulon	North Carolina	United States	No	No	No	
10080287	Winter Park	Florida	United States	No	No	No	
10080074	Shanghai	Zhejiang	China	No	Yes	No	
10078741	Unknown	Unknown	Greece	No	No	No	
10077344	Durham	North Carolina	United States	No	No	No	
10077240	Dallas	Texas	United States	No	No	No	
10076844	London	Barking and Dag...	United Kingdom	No	No	No	
10076781	Centreville	Virginia	United States	No	Yes	No	
10075398	Kirchberg	Ausser-Rhoden	Switzerland	No	No	Yes	
10075025	Auckland	Auckland	New Zealand	No	Yes	No	
10074974	Tula De Allende	Hidalgo	Mexico	No	No	No	
10071585	Mexico	Chiapas	Mexico	No	No	Yes	
10071360	Falls Church	Virginia	United States	No	No	No	

Figure 18. Audience Segment Table Available for Export

CONCLUSION

Consumers today have high expectations when it comes to how brands interact with them. Marketers understand this. They know that to meet these expectations, they need deeper insight into their customers across all channels – including web, social, and mobile. To meet this challenge, the two worlds of advanced analytic analysts and digital analysts must converge, and begin working together rather than in silos.

SAS Adaptive Customer Experience helps organizations achieve this by collecting detailed digital data and providing business context to connect the integrated marketing department to online customers. Predictive marketing capabilities accessible to data miners and business analysts through SAS Visual Analytics are designed to meet the market's rising demand of enabling big data analytics to exploit digital marketing insights.

To learn more about these solutions, please visit the following:

SAS Adaptive Customer Experience: http://www.sas.com/en_us/software/customer-intelligence/customer-experience.html

SAS Visual Analytics: http://www.sas.com/en_us/software/business-intelligence/visual-analytics.html

REFERENCES

- ¹ Gill, Martin. December 19, 2013. "Predictions 2014: The Year of Digital Business." *Forrester Research*.
- ² Adobe. September 23, 2013. "Digital Distress: What Keeps Marketers Up at Night?" *Adobe Systems Inc.*
- ³ McCormick, James. November 11, 2013. "Digital Intelligence Replaces Web Analytics." *Forrester Research*.
- ⁴ Gordon, Ruth / Hulls, Katharine. January 2014. "Digital Marketing Insights Report 2014." *Teradata/Celebris Technologies*.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:

Suneel Grover, M.B.A., M.S.
Sr. Solutions Architect - SAS Institute Inc.
Adjunct Lecturer – The George Washington University
Suneel.Grover@sas.com
<http://www.sas.com>
LinkedIn: <http://www.linkedin.com/in/suneelgrover>
Twitter: @suneelgrover

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.