Paper SAS5661-2016

Bringing Google Analytics, Facebook, and Twitter Data to SAS® Visual Analytics

I-Kong Fu, Mark Chaves, Andrew Fagan, Jim Hazen, SAS Institute Inc.

ABSTRACT

Your marketing team would like to pull data from its different marketing activities into one report. What happens in Vegas might stay in Vegas, but what happens in your data does not have to stay there, locked in different tools or static spreadsheets. This paper demonstrates how to easily bring data from Google Analytics, Facebook, and Twitter into SAS® Visual Analytics in order to create interactive explorations and reports on these data along with your other data for better overall understanding of your marketing effectiveness.

INTRODUCTION

In today's digital business landscape, we often celebrate the "measurability" of digital data and its potential impact on business decisions and customer experiences. Based on research conducted by Forrester Research, businesses continue to struggle to derive actionable insights across digital data including data from web interactions and social media.

Marketing periodicals have cited Forrester's surveys which reveal that 42% of business leaders state that their top organizational challenge is uncovering actionable insights from social data¹. What's more, 34% of these same organizations are also having difficulty in sharing insights across departments within their enterprise².

In order to grow their top and bottom line, today's marketing organizations must deliver on their promise of meaningful customer experiences through efficient communications across many different touchpoints. Fortunately, a lot of these channels are digital and generate a lot of useful data. In fact, it should be no surprise that "customer experience" and "big data" have become perennial "trends" that show up on yearly "Marketing Trends" news articles³. Experts promote the importance of unearthing insights and then finding a way to communicate these learnings with management and operations.

As part of your marketing efforts, chances are your team is tracking information in separate software tools for social media, web analytics, offline sales leads, and so on. You and your colleagues might want to track most of this information at a higher level in one place so that you can quantify relationships across these data as well as ensure that insights gleaned are able to be easily shared with management.

Most of these digital tools have application programming interfaces (APIs) or support export of delimited files. As a result, you can bring much of the information together into SAS Visual Analytics for high-level exploration and reporting. More specifically, SAS Visual Analytics currently supports collection of data from Google Analytics, Facebook, and Twitter via built-in connections. Support for these web and social platforms is just the beginning as SAS will continue to add new data collectors as standard capabilities within SAS Visual Analytics.

SCENARIO

Let's use a fictitious scenario to illustrate how you might use these capabilities. In this scenario you work as an analyst for a major league baseball team and your marketing team is asking you to analyze what is driving traffic to the site in anticipation of Opening Day. You have been monitoring various digital marketing data sources that the team wishes to pull together in one dashboard:

Twitter: You have been looking at "macro" trending topics on social media such as Twitter
and are seeing more activity with hashtags and terms "#MLB", "#OpeningDay" and "Opening
Day". These are conversations the fans are having mostly independent of your or others'
campaigns. You can filter within these data by your team.

- Facebook: From a "macro" view, you also drill down to "owned" social properties, specifically, your team's Facebook page and those of your closest rivals. With the Facebook data, you can look at conversations and filter by type of posting (post, comment, or reply).
- Google Analytics: you have been tracking campaigns and traffic from different campaigns
 and channels. Based partly on the excitement you've been seeing in your various social
 media, your team wants to run an "Opening Day" campaign with various promotions around
 the home game. You discuss what aspects to promote: new ballpark, players, improved
 hitting. These ideas might be informed by the social media analysis. You also are looking at
 sales of players' jerseys and merchandise given the recent tweets and comments.

AVAILABLE DIGITAL DATA SOURCES

Like your team, many organizations are tracking web interactions (site visits, page views, sessions ...) and social media activity (Tweets, Retweets, Posts, Comments, Likes...) in separate applications and systems including the built-in systems from the particular social media channel. Starting with SAS Visual Analytics 7.2, there is now out-of-the-box connectivity to Google Analytics and Facebook, in addition to Twitter. These sources can be found in the "Other" category on the Import Data panel. See Figure 1 below.

Other

<u>Facebook</u> <u>Google Analytics</u> Twitter

Figure 1. Access these data sources in the Other category when importing data

Let's take a look at each one of these data sources and built-in access from SAS Visual Analytics.

TWITTER

With more than 320 million users, Twitter is an important micro-blogging platform. There are many social media listening platforms that collect Twitter data on their own. In fact, Twitter recently introduced their own Twitter Analytics (analytics.twitter.com) to provide data on key metrics including the "Reach" of your tweets.

These data are available for export. However, with SAS Visual Analytics you can not only blend Twitter data with other data sources, you can also take your analysis further by using built-in text analytics to process hashtags and others key terms. Furthermore, advanced linguistic processing will also help you to automatically categorize the text into topics and perform sentiment analysis.

Connection to Twitter requires you to log on to Twitter and authorize the SAS Visual Analytics application for use with Twitter. The connection is very similar to Google Analytics' connection which will be described in more detail in a section below. Queries are defined by searching for a specific set of hashtags or other search terms. Maximum tweet results are set by Twitter's Public API. Additionally, you may choose not to import retweets.

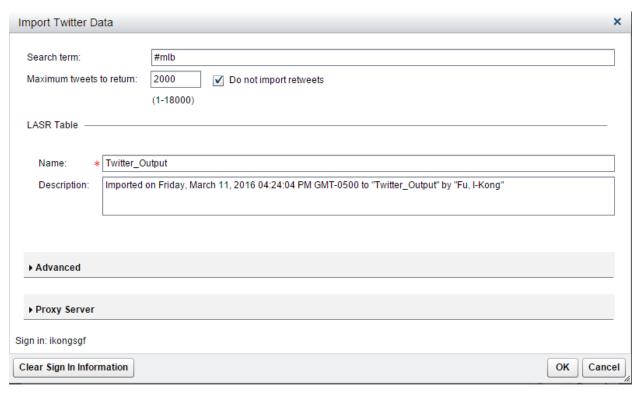


Figure 2. Retrieving data from Twitter with a search term

After we retrieve data based on a search term, we want to categorize the topics in the tweets and analyze the sentiment. Right click on the body variable and turn it into a Document Collection.

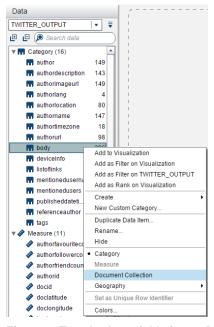


Figure 3. Turn body variable into a Document Collection for text analytics

We will be prompted to specify a unique row identifier. Choose docid.

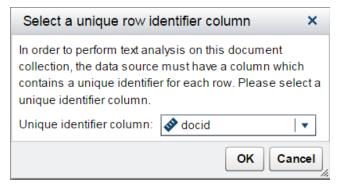


Figure 4. Use docid for the row identifier

Using the auto chart feature in SAS Visual Analytics Explorer, we will get topics with a word cloud for each topic. However we aren't just interested in seeing words by frequency. We would like to have an idea of the categories or groupings within this topic and how fans feel when posting their tweets. SAS Visual Analytics includes some built-in capabilities from SAS text analytics technology. We see the automatic categorization of our topic in a drop down at the top left. We can then right click to turn on sentiment analysis for additional analysis. Clicking on a term will show the sentiment score for a document or tweet in this case along with emoticon style traffic highlighting in a panel below.

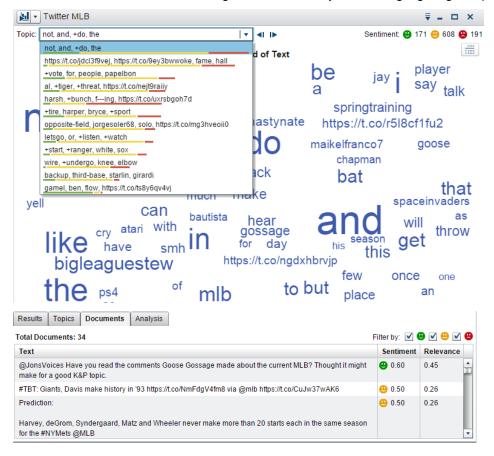


Figure 5. Sentiment analysis

It appears that fans are very excited about the upcoming seasons. Next let's take a look at our Facebook page for social media analysis of content that is much more specific to our organization.

FACEBOOK

With over 1 billion users worldwide, Facebook is a premier social network platform. As with Twitter, there are many social media listening platforms that collect Facebook data on their own and Facebook also has their own "Insights API" which provides basic metrics, which we can import via the exported csv file.

However we will also explore sentiment analysis in the text data. Because we are exploring text on public Facebook pages or "fan pages", when we access the Facebook data source we do not need to log in using a specific account. We will simply enter a page name and SAS Visual Analytics will communicate with Facebook to retrieve text. Facebook has text that is more structured than the Twitter data with different levels of text: posts, comments, and replies to the comments. We can explore data from a public page such as our own or other pages. It may be helpful to analyze competitors' pages, for example. We can also get some basic metrics from this data. See figure 6 below.

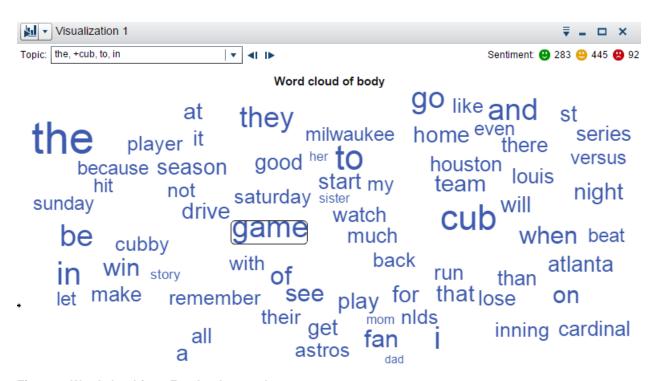


Figure 6. Word cloud from Facebooks text data

With this data, you can also look at metrics on Likes and Shares and filter the word cloud topics by the posting type (posts, comments, replies to comments). You can then save these explorations for your team to analyze later.

With a lot of interesting conversation happening in social media about the season and your team, you want to take a look at traffic to your site, and especially when driven from related campaigns and posts on social channels. With discussions about particular star players, your team is also wondering about ecommerce sales of fan merchandise as the season is about to start. Let's take a look at Google Analytics next.

GOOGLE ANALYTICS

Many organizations are tracking web analytics information with Google Analytics. Even when there is a paid web analytics solution in their organization, teams often choose to use Google Analytics as a secondary source. Here is how you can establish access to your Google Analytics account:

When clicking the Google Analytics link you will be prompted to obtain an access code from Google that you will need to paste back into SAS Visual Analytics.

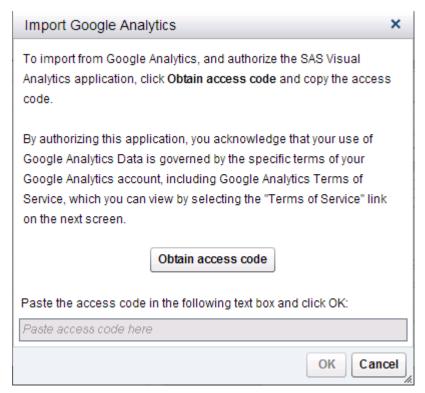


Figure 7. Obtain access code from Google Analytics

Clicking the button will take you to a Google page. After signing in to your Google account, you will be prompted to allow SAS Visual Analytics to access your Google Analytics data.

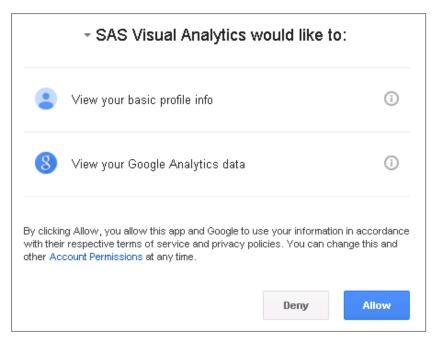


Figure 8. Give permission to access your data

SAS Visual Analytics needs permission in order to retrieve the information for you. Click "Allow" and copy the access code returned by Google into the box above and then you can begin accessing your Google Analytics data. The data are in predefined groupings of dimensions and metrics. The Google Analytics API enables you to retrieve groupings of up to seven (7) dimensions and ten (10) metrics in one request.

Not all combinations can go together. SAS Visual Analytics provides seven (7) predefined groupings at this time, and customizable user-defined groupings will be available in a future release. There are groupings for Acquisition, Behavior, Ecommerce, Geography, Mobile, Searches, and Technology. These combinations of dimensions and metrics are based on the most popular queries to the API according to Google. See Figure 2 below for the selection screen for the groupings, called "Study area" in the UI.

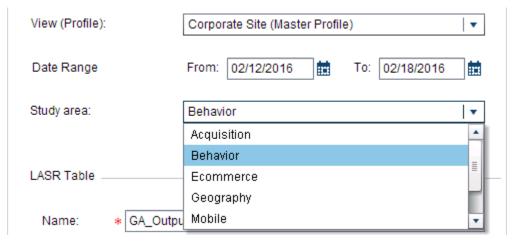


Figure 9. Study areas are groupings of dimensions and metrics

Figure 10 shows an example of the Acquisition area as well as other choices a user can make. This can include the date range which defaults to the previous 7 days and the account, property, and profile. In this example we will retrieve data for the master profile for our site.

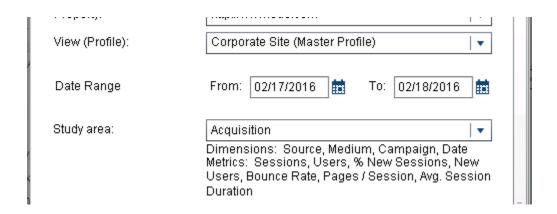


Figure 10. Acquisition area example

In the figure, we can see the User Interface (UI) names for our dimensions and metrics. When we retrieve the data, the variables in our table will show using the API names. Usually, that just means that there is a "ga:" in front of the UI name. For example, Source will show as "ga:source". However, users can change the names in SAS Visual Analytics if desired.

Once we have selected a grouping, SAS Visual Analytics will begin to retrieve data from Google Analytics. If you have a reasonably high traffic site and/or a large date range you will probably encounter a message that sampling is used as in Figure 4 below.

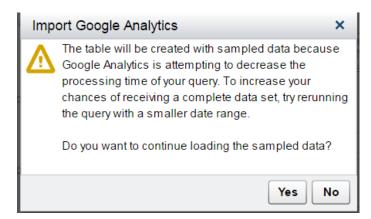


Figure 11. Sampled data

Your results will be representative of your actual metrics based on a sample of your data so that the Google API can return everyone's requests in a timely way while being fair to all users. Also, you might be retrieving a large amount of data and encounter this dialog:

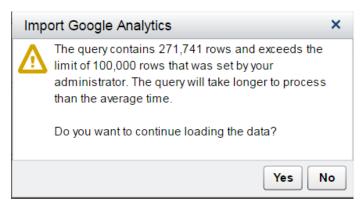


Figure 12. Large amount of data

We can see from above that the Acquisition grouping is about where visitors came from before arriving to your site. Let's look at a simple example using this grouping. We have various social campaigns happening across some of our different social channels. We have one called "Opening Day" and another called "Player Jerseys". We will create a hierarchy using the Source, Medium, and Campaign variables. From a bar chart, we can quickly see that for our Opening Day campaign we have more traffic (sessions) from Facebook and for our Player Jerseys campaign we have more sessions from our Twitter activity.

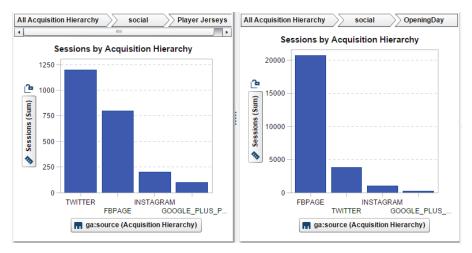


Figure 13. Exploring our traffic sources in a dynamic hierarchy

Let's examine Behavior next. With these dimensions and metrics, we can show basic metrics on pageviews and sessions. We will show our pageviews by date to see the trend over time.

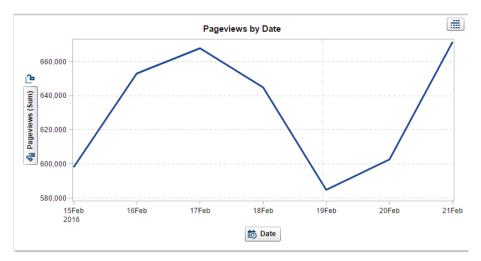


Figure 14. Pageview trend over time

Next let's examine ecommerce. In Figure 15 we have calculated average order value from our transactions and transaction amounts. In this case we have granular data that makes sense for a new computed column. In other cases such as Pageviews Per Session, we are retrieving totals that Google Analytics has already calculated. Therefore we have to be careful not to do calculations such as averages of averages.

Date	Transactions	Average Order Value (AOV)
Feb/15/2016	1162	108.59
Feb/16/2016	1154	130.22
Feb/17/2016	1242	159.33
Feb/18/2016	1222	109.24
Feb/19/2016	932	114.32
Feb/20/2016	1332	88.29
Feb/21/2016	1694	166.24

Figure 15. Table showing ecommerce data including a calculated column

What about custom groups of measures and dimensions? Suppose we want to look at transactions for traffic associated with a campaign. We are not able to join the data due to it returning at different levels of aggregation. To address this kind of question, SAS Visual Analytics will support this custom grouping along with this kind of filtering in the future. For now as a workaround we can use various free and commercial tools such as the Google Analytics Query Explorer tool to retrieve data and import a tabseparated file: https://ga-dev-tools.appspot.com/query-explorer/

Those with more extensive needs and experience using APIs could also use some custom programming with the Core Reporting API: https://developers.google.com/analytics/devguides/reporting/core/v3/

FUTURE DIRECTIONS

This paper has given you an introduction to bringing together social and web data into SAS Visual Analytics so that the team can start looking at data together in one location. Just the process of creating a shared dashboard can lead to more strategic insights and conversations about what is happening and the story shown in the data. Support for Twitter, Facebook, and Google Analytics is just the beginning. SAS will continue to expand data access options such as custom groups of dimensions and metrics for Google Analytics, and will build out new data connections to new digital data sources (like YouTube, others); we will also be enabling more operational delivery of these data streams so that "always on" data streams are available for production environments and not just analytical workstations.

CONCLUSION

For businesses to compete in increasingly complex markets, the ability to learn, process, and act has become a strategic imperative.

To fulfill this need, business managers and leaders will need to recognize that:

- Data becomes increasingly actionable and prescriptive when mashed together
- Business results and customer experiences will improve once everyone (across all touchpoints) are on the same page
- The importance of considering how to operationalize data collection, processing, and learning.

By tapping into digital data streams (including data from web interactions and social media) and blending these data with other corporate data assets, organizations will increasingly uncover market opportunities, orchestrate meaningful customer experiences, and drive efficiency and effectiveness.

REFERENCES

¹ Tode, Chantal. "Social listening is not ready for prime time: Forrester". Mobile Marketer magazine January 14, 2015. Available at http://www.mobilemarketer.com/cms/news/research/19541.html

RECOMMENDED READING

- Inbound Marketing: Attract, Engage, and Delight Customers Online by Brian Halligan and Dharmesh Shah
- Google Analytics Dimensions and Metrics Explorer: https://developers.google.com/analytics/devguides/reporting/core/dimsmets

CONTACT INFORMATION

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

I-Kong Fu SAS i-kong fu@sas (

i-kong.fu@sas.com

Mark Chaves SAS Mark.Chaves@sas.com

Andrew Fagan
SAS
Andrew.Fagan@sas.com

Jim Hazen SAS Jim.Hazen@sas.com

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.

² Kaneshige, Tom. "Why Social Listening Platforms Are Failing". CIO magazine January 27, 2015. Available at http://www.cio.com/article/2876295/social-media/why-social-listening-platforms-are-failing.html

³ Newman, Daniel. "The Top 10 Marketing Trends That Will Define 2016". Fortune magazine November 3, 2015. Available at http://www.forbes.com/sites/danielnewman/2015/11/03/the-top-10-marketing-trends-that-will-define-2016