Google Analytics

INTELLIGENCE FOR INFORMATION PROFESSIONALS

by Beatriz Plaza

The launch of free web analytics tools by search engines can serve as a key online marketing tool for information managers. Analyzing the data revealed by Google Analytics (www.google.com/analytics) or Yahoo! Web Analytics (http://web.analytics.yahoo.com) helps webmasters fine-tune their sites more reliably. The result: increased website visibility and greater user satisfaction.

Web analytics have become websites' data warehouses, providing broad and nonstrategic analytics—probably too broad, in many cases. The search for and collection of relevant website information can be very time-consuming and costly for information professionals. However, the adoption of key metrics can contribute to reducing time and costs of finding relevant information about a website's performance.

Web analytics is not just about the number of visitors visiting a website; it is also about the quality of the traffic and what the visitors do after they enter the site. Does the website fulfill the needs of the organization it represents by meeting the needs of the target audience? How well does a website pursue its mission?

Analytics shows what type of visitors are on the site, where they came from, how they arrived at the website, what they did on the visit, and whether the organizations' objectives have been met. Web analytics serve as tools for information professionals to provide an understanding of the driving forces behind visitors' navigation through the website. Information on the number of clicks and pathways is useful, but interpretation of visitor behavior is a critical dimension if you intend to elevate web analytics into business intelligence.

LEARNING FROM WEB ANALYTICS

Nothing has transformed organizations in the past 10 years more than the internet and, to an even larger extent, Google. As the internet moves from being a means of technology to a channel of customer relationship and sales, information professionals, who have somewhat different objectives for their websites, should consider revolutionizing their website performance evaluation strategies. You can do this through new approaches that can assist practitioners in evaluating the visitors' preferences. With this knowledge, you can upgrade the sites' usefulness.

The aim of this article is to develop further the methodology, which I initiated, on the use of time series with Google Analytics' data and to supply some user-friendly metrics (that is, tips) for information practitioners. It is interesting to see how key performance indicators (KPIs) can help professionals, who can then make websites more effective.

For background on my research and other studies, please see the accompanying annotated bibliography.

RULES OF THE ROAD

When I performed time series analysis with Google Analytics, I noticed several phenomena that I have codified into three rules:

- Rule No. 1: Return visitors navigate deeper into the website and stay longer (that is, there is more time spent at the site and/or a greater number of pages viewed per visit).
- Rule No. 2: The lower the bounce rate (that is, the number of error visits), the longer the visit length (with regard to the time spent at the site and/or the number of pages viewed per visit).
- Rule No. 3: The lower the bounce rate (that is, the number of error visits), the greater the return visit rate.

I then tested these rules for simple cross-sectional data supplied by Google Analytics. Information professionals can take action based on measurable results using analytics to improve the effectiveness of the web content management and leverage information in their website. The aim of this work is to provide some KPIs using Google Analytics.

TESTING THE HYPOTHESES

Our tested website, Scholars on Bilbao (www.scholars-on-bilbao.info), exists to disseminate research and development information about the region. Launched in July 2006, it concentrates on "Cultural Policy as Development Policy" and provides academic works (abstract and link to PDF) that analyze the urban regeneration of the city of Bilbao, Spain, including the effect of the Guggenheim Museum Bilbao, cultural tourism, gentrification, uneven development, and creative industries.

In February 2007, the webmaster started to analyze web traffic using Google Analytics. From Feb. 4, 2007, to Jan. 30, 2010, Google Analytics registered 7,561 entries for those 1,092 days. Of those visits, 1,368 came directly to this site; referring sites sent 3,298 visits via 121 sources; and search engines sent 2,892 visits, mainly through Google. Reference site traffic is, by far, the main source of entries—almost 44% of the total incoming visits; that is, 3,298 entries through in-links, 1,820 of which enter from http://en.wikipedia.org.

How deep into the website do in-links visits navigate in comparison with other traffic sources? Are Wikipedia references more effective than other in-links? Which is the most effective traffic source? How deep do Google entries navigate? Which are the most effective keywords?

APPLYING THE RULES

To provide a preliminary answer to these questions, I applied the already stated three rules (and key metrics) in order to measure simple cross-sectional data as follows:

1. The first step was to collect all the data (see Table 1)—the number of visits for each traffic source, session length (that is, the time spent on the site and/or the number of
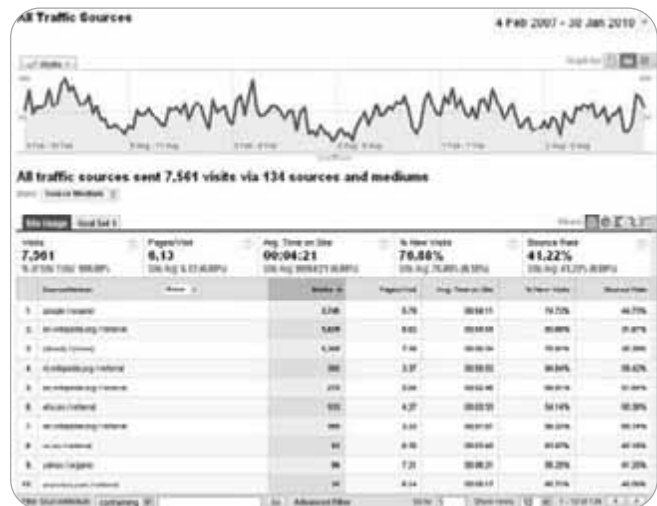
Google Analytics overview for Scholars on Bilbao

pages viewed per visit), the bounce rate, and the return visits rate. These indicators correspond to average values for the survey period.

2. Next, the traffic sources were sorted according to traffic volume; a ranking from the highest to the lowest traffic volume source was established (see Table 2). Next, the top 10 were selected (see marked traffic sources in Table 2).

3. Then, the top traffic sources with the highest return rate were selected (Table 3) to answer these key questions: Which are the traffic sources that generate traffic and also produce a high return rate? Which are the traffic sources that produce entries and achieve a return rate above average? In general, return visits travel deeper into the website and bounce less. It is for this reason that I chose the "maximizing return rate" criteria. However, the website owner can choose either to minimize "bounce rate" or to maximize session length, with regard to the time spent at the site and/or the number of pages viewed per visit. For this particular website, the most effective traffic sources are the keywords "scholar," "Bilbao," "museum," and "Guggenheim" for search engines, direct traffic, and Google (see Table 3).

4. Next, a scatter plot was created for the return rate against the number of pages viewed per visit for all the main traffic sources (Figure 2). This shows a positive relationship between return rates and the number of pages viewed per visit for the traffic sources (Rule No. 1).

5. The next stage was to scatter plot bounce rates against the number of pages viewed per visit (see Rule No. 2 in Figure 3). The aim here is to identify the qualified low-bounce traffic sources. The keyword "scholar" in search engines is by far the traffic source that qualifies with the lowest bounce rate. Then, well behind, direct traffic and the keyword "Bilbao" in search engines perform relatively well in terms of qualified low-bounce rate traffic.

6. The next stage was to scatter plot bounce rate against return rate for all the traffic sources (see Figure 4). Here,
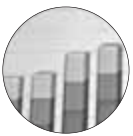
Table 1: Traffic Sources for Scholars on Bilbao (average values from Feb. 4, 2007, to Jan. 30, 2010)		Visits	Pages per Visit	Bounce Rate	Return Visits Rate
	Total	7.561	6.13	0.41	0.23
Traffic Sources	Direct Traffic	1.368	7.48	0.35	0.29
	Referring Sites	3.298	5.84	0.41	0.18
	Search Engines	2.892	5.81	0.45	0.26
Top 10 Referring Sites by Traffic	en.wikipedia.org / referral	1.820	6.62	0.32	0.19
	nl.wikipedia.org / referral	392	3.37	0.58	0.05
	es.wikipedia.org / referral	275	5.04	0.52	0.13
	ehu.es / referral	133	4.27	0.50	0.46
	de.wikipedia.org / referral	109	3.33	0.65	0.04
	uv.es / referral	93	6.35	0.45	0.16
	answers.com / referral	35	6.14	0.40	0.14
	plataformaurbana.cl / referral	32	3.78	0.44	0.16
	no.wikipedia.org / referral	21	1.43	0.90	0.05
	elearningeuropa.info / referral	19	3.47	0.47	0.63
Search Engines	Google	2.741	5.79	0.45	0.25
	Yahoo!	80	7.21	0.41	0.44
Top 15 Keywords by Traffic	All keywords	2.892	5.81	0.45	0.26
	Bilbao	1.485	7.36	0.36	0.34
	Urban	816	5.51	0.48	0.20
	Regeneration	583	5.86	0.42	0.22
	Guggenheim	457	6.32	0.42	0.32
	Scholar	324	10.61	0.16	0.50
	Museum	309	5.81	0.43	0.33
	Cultural	296	4.98	0.57	0.17
	City	179	4.55	0.59	0.12
	Culture	160	6.38	0.44	0.21
	Brand	146	4.93	0.62	0.11
	Effect	123	5.26	0.41	0.37
	Plaza	111	5.59	0.46	0.45
	Image	94	4.31	0.60	0.16
	European	94	4.87	0.62	0.14
	Tourism	92	4.34	0.52	0.09
	Design	91	3.57	0.57	0.10

Source: Google Analytics for www.scholars-on-bilbao.info

there is a negative relationship between bounce rate and return rate (Rule No. 3). The keyword "scholar" is by far the traffic source that qualifies both the highest return rate and the lowest bounce rate. Then, well behind, direct traffic and the keyword "Bilbao" in search engines also perform relatively well in terms of higher return rates and qualified low bounce rate traffic.

EFFECTIVE KEYWORDS FOR SEO

In summary, the keywords "scholar," "Bilbao," "museum," and "Guggenheim"; direct entries; and Google are responsible for driving quality traffic to the site, whereas the keywords "regeneration" and "urban" in search engines and the referrals from http://en.wikipedia.org and http://nl.wikipedia.org underperform (Table 3). This is an extremely cost-effective method of analysis. While not 100% statistically robust, it can be useful to optimize your site for search engines.

Search engine optimization (SEO) is the ongoing process of designing the site to increase the number of visits, improve the quality of the visitors, and raise the ranking of the site by search engines. The ultimate goal of SEO is to drive qualified traffic without including pay-per-click (PPC) or paid search.

The key performance indicators uncovered by web analytics can help information professionals use tested keywords to achieve their SEO aims. It can also change the way they approach paid search, leading them to focus less on expensive keyword bidding (Google AdWords) and look instead at the long tail, protecting themselves against rising PPC rates caused by the growing demand for paid search.

| Table 2: Traffic Sources for Scholars on Bilbao Sorted According to Traffic (average values from Feb. 4, 2007, to Jan. 30, 2010)	Visits	Pages per Visit	Bounce Rate	Return Visits Rate
Total	7.561	6.13	0.41	0.23
Refering Sites	3.298	5.84	0.41	0.18
Search Engines	2.892	5.81	0.45	0.26
Google	2.741	5.79	0.45	0.25
en.wikipedia.org/referral	1.820	6.62	0.32	0.19
Bilbao/keyword	1.485	7.36	0.36	0.34
Direct Traffic	1.368	7.48	0.35	0.29
Urban/keyword	816	5.51	0.48	0.20
Regeneration/keyword	583	5.86	0.42	0.22
Guggenheim/keyword	457	6.32	0.42	0.32
nl.wikipedia.org/referral	392	3.37	0.58	0.05
Scholar/keyword	324	10.61	0.16	0.50
Museum/keyword	309	5.81	0.43	0.33
Cultural	296	4.98	0.57	0.17
es.wikipedia.org/referral	275	5.04	0.52	0.13
City	179	4.55	0.59	0.12
Culture	160	6.38	0.44	0.21
Brand	146	4.93	0.62	0.11
ehu.es/referral	133	4.27	0.50	0.46
Effect	123	5.26	0.41	0.37
Plaza	111	5.59	0.46	0.45
de.wikipedia.org/referral	109	3.33	0.65	0.04
Image	94	4.31	0.60	0.16
European	94	4.87	0.62	0.14
uv.es/referral	93	6.35	0.45	0.16
Tourism	92	4.34	0.52	0.09
Design	91	3.57	0.57	0.10
Yahoo	80	7.21	0.41	0.44
answers.com/referral	35	6.14	0.40	0.14
plataformaurbana.cl/referral	32	3.78	0.44	0.16
no.wikipedia.org/referral	21	1.43	0.90	0.05
elearningeuropa.info/referral	19	3.47	0.47	0.63

*Source: Google Analytics for www.scholars-on-bilbao.info*

Table 3: Traffic Sources for Scholars on Bilbao Sorted According to Return Rate (average values from Feb. 4, 2007, to Jan. 30, 2010)	Visits	Pages per Visit	Bounce Rate	Return Visits Rate
elearningeuropa.info referral	19	3.47	0.47	0.63
Scholar/keyword	324	10.61	0.16	0.50
ehu.es referral	133	4.27	0.50	0.46
Plaza	111	5.59	0.46	0.45
Yahoo	80	7.21	0.41	0.44
Effect	123	5.26	0.41	0.37
Bilbao/keyword	1.485	7.36	0.36	0.34
Museum/keyword	309	5.81	0.43	0.33
Guggenheim/keyword	457	6.32	0.42	0.32
Direct Traffic	1.368	7.48	0.35	0.29
Search Engines	2.892	5.81	0.45	0.26
All keywords	2.892	5.81	0.45	0.26
Google	2.741	5.79	0.45	0.25
Average		**6.13**	**0.41**	**0.23**
Regeneration/keyword	583	5.86	0.42	0.22
Culture	160	6.38	0.44	0.21
Urban/keyword	816	5.51	0.48	0.20
en.wikipedia.org/referral	1.820	6.62	0.32	0.19
Refering Sites	3.298	5.84	0.41	0.18
Cultural	296	4.98	0.57	0.17
uv.es/referral	93	6.35	0.45	0.16
Image	94	4.31	0.60	0.16
plataformaurbana.cl/referral	32	3.78	0.44	0.16
answers.com/referral	35	6.14	0.40	0.14
European	94	4.87	0.62	0.14
es.wikipedia.org/referral	275	5.04	0.52	0.13
City	179	4.55	0.59	0.12
Brand	146	4.93	0.62	0.11
Design	91	3.57	0.57	0.10
Tourism	92	4.34	0.52	0.09
nl.wikipedia.org/referral	392	3.37	0.58	0.05
no.wikipedia.org/referral	21	1.43	0.90	0.05
de.wikipedia.org referral	109	3.33	0.65	0.04

Source: Google Analytics for www.scholars-on-bilbao.info
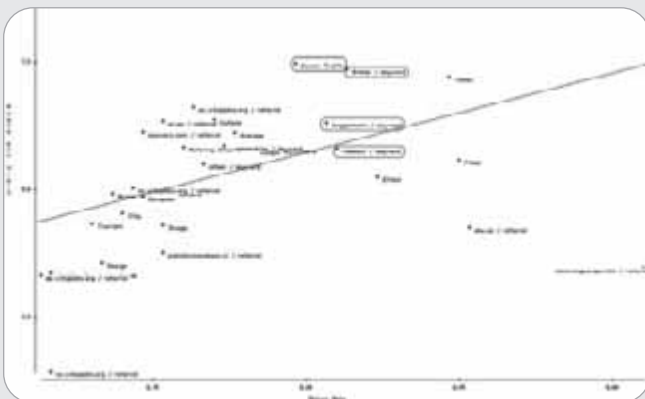
Rule No. 1: Return visits navigate deeper into the website and stay longer.
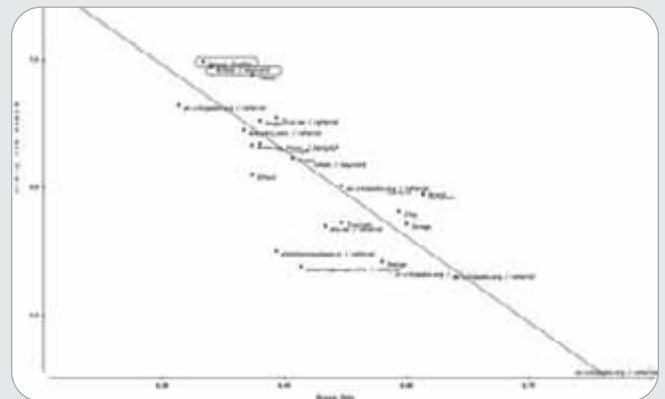
Rule No. 2: The lower the bounce rate, the greater the visit duration

Suggested Readings

Betty, Paul. "Assessing Homegrown Library Collections: Using Google Analytics to Track Use of Screencasts and Flash-Based Learning Objects." *Journal of Electronic Resources Librarianship*. Vol. 21, No. 1 (2009): pp. 75–92.

Explores the use of Google Analytics to track usage statistics for interactive Shockwave Flash (.swf) files, the common file output for screencast and Flash projects.

Bhatnagar, Alka. "Web Analytics for Business Intelligence: Beyond Hits and Sessions." *ONLINE*. Vol. 33, No. 6 (2009): pp. 32–35.

Explains the use of web analytics as they apply to library websites and how libraries must find new metrics to determine success.

Fang, Wei. "Using Google Analytics for Improving Library Website Content and Design: A Case Study." *Library Philosophy and Practice 2007* (June), LPP Special Issue: Libraries and Google (2007).

Presents a case study of how a law library used Google Analytics to redesign its website.

Hasan, Layla; Morris, Anne; Probets, Steve. "Using Google Analytics to Evaluate the Usability of E-Commerce Sites." Lecture Notes in Computer Science. Vol. 5619 (2009): pp. 697–706.

Suggests specific web metrics that are useful for quickly identifying potential usability problems of ecommerce websites.

Plaza, Beatriz. "Monitoring Web Traffic Source Effectiveness with Google Analytics. An Experiment with Time Series." Aslib Proceedings. Vol. 61, No. 5 (2009): pp. 474–482.

Explores some statistical matters with regard to the use of Google Analytics data in combination with time series methodology.

Plaza, Beatriz. "Google Analytics for Measuring Website Performance." *Tourism Management*. (2010). doi:10.1016/j.tourman.2010.03.015.

Analyzes the effectiveness of entries (visit behavior and length of sessions) depending on their traffic source, using time series analysis; sets up Google Analytics eMetrics for information professionals, using basic cross-sectional data.

Rodriguez-Burrel, Jorge. "Google Analytics: Buena, bonito y gratis" ("Google Analytics: Good and Nice and Free"). *Profesional de la Información*. Vol. 18, No. 1 (2009): pp. 67–71.

Details the use of Google Analytics for online marketing campaigns (article is in Spanish; abstract is in English).

Web Analytics Association. The Web Analytics Association; www.webanalyticsassociation.org.

International professional association for those involved in the measurement, collection, analysis, and reporting of internet data for the purposes of understanding and optimizing web usage.

This way, information professionals can even allocate their budget to niche terms and combinations, allowing them to reach well-defined micromarkets for niche products and services. Google Analytics Intelligence will help you understand the long tail, which can become vitally critical for executives at information businesses and for libraries.

IMPLICATIONS FOR WEBSITES

Which are the most effective traffic sources for your particular website? The traffic sources that perform most effectively are those that 1) drive a higher traffic volume, 2) have the highest return rate, 3) have the largest visit length, and 4) have the lowest bounce rate. The results obtained with cross-sectional data are quite consistent with those results I obtained through time series analysis.

These rules can be easily applied to basic cross-sectional data supplied by Google Analytics and can be helpful, especially for information experts, although a statistically robust study of web analytics requires a more thorough analysis of the time series for different traffic sources.

The importance of this research is not to identify KPI for just one particular Bilbao website. It lies in the methodology tested to arrive at these results. The agenda for future research requires the repetition of the experiment with different websites, to delimit more accurately the effectiveness of different traffic sources and to compare these results with other case studies. The agenda for future research also calls for an analysis of which traffic sources should perform better, and why.
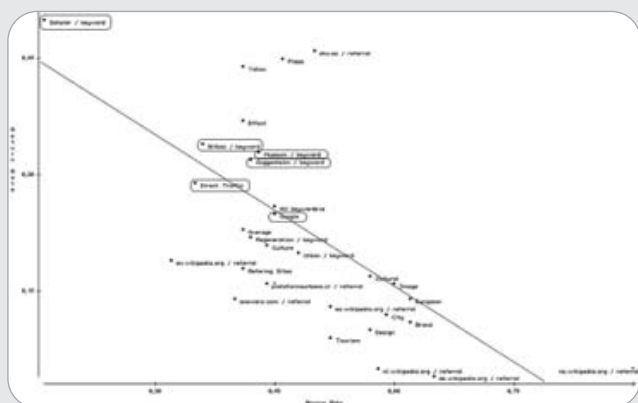
Rule No. 3: The lower the bounce rate, the greater the return visit rate

Beatriz Plaza (beatriz.plaza@ehu.es) is on the Faculty of Economics, University of the Basque Country, Bilbao, Spain.
Comments? Email the editor (marydee@xmission.com).