Tourism Management 32 (2011) 477–481

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Google Analytics for measuring website performance

Beatriz Plaza*

Faculty of Economics, University of the Basque Country, Avda. Lehendakari Agirre 83, 48015 Bilbao, Spain

ARTICLE INFO

Article history:
Received 26 February 2010
Accepted 22 March 2010

Keywords:
Web analytics
Google Analytics
Time series
Online marketing
Ecommerce
eTourism

ABSTRACT

Performance measurement of tourism websites is becoming a critical issue for effective online marketing. The aim of this article is to analyse the effectiveness of entries (visit behaviour and length of sessions) depending on their traffic source: direct visit, in-link entries (for instance, en.wikipedia.org), and search engine visits (for example, Google). For this purpose, time series analysis of Google Analytics data is made use of. This method could be interesting for any tourism website optimizer.

1. Introduction

The Internet has revolutionized the business operations of the entire travel and tourism industry value chains. As the Internet has become the channel of customer relationships and sales, the performance measurement of tourism websites is becoming a strategic issue, critical for online marketing. Web analytics is on the increase.

Hundreds of thousands of tourism web owners worldwide have a web analyser program available to them. The web analyser provides plain and simple statistics concerning the website (number of visitors, the average number of page views per visitor, average page duration, most requested pages, domain classes and referrers). This article presents an experiment done with the information that Google Analytics offers for a R&D resource devoted to cultural tourism, about the number of visits on a website and the traffic source, which includes organic results in search engines, links from referral web pages or direct access. In other words, it investigates the differences between sessions started by direct connection by typing the site name, through a link on another site, or from a search engine, with regard to return visit behaviour and length of sessions. The importance of this paper is not the particular website http://www.scholars-on-bilbao.info, but the new methodology tested to arrive at these results. The case study must be presented only as a way to explain the new methodology, a method that could be interesting for a wider audience.

How deep do visitors navigate into the website? What is their internal performance depending on their traffic source? Are search engine visits more effective than referring site entries? This paper addresses these questions by time series analysis of Google Analytics data, in order to compare the performance of visits depending on their source: direct visits, referring site visits and search engine visits.

Why use Google Analytics? Firstly, and most importantly for the purpose of this study, it is used because Google Analytics provides time series data. Moreover, it is also employed because Google Analytics is a free service offered by Google that generates detailed statistics about the visits to a website, and which is a user friendly application with the guarantee of Google technology. This tracking application, external to the website, records traffic by inserting a small piece of HTML code into every page of the website. Google Analytics tells the web owner how visitors found the site and how they interact with it (see *Dashboard* in Fig. 1). Users will be able to compare the behaviour of visitors who were referred from search engines, from referring sites and emails, and direct visits, and thus gain insight into how to improve the site's content and design (Plaza, 2009).

2. Literature review

Many studies have been made about tourism firms utilizing the Internet for relationship building, business transactions and

* Tel.: +34 946013641; fax: +34 946017087.
E-mail address: beatriz.plaza@ehu.es

0261-5177/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tourman.2010.03.015

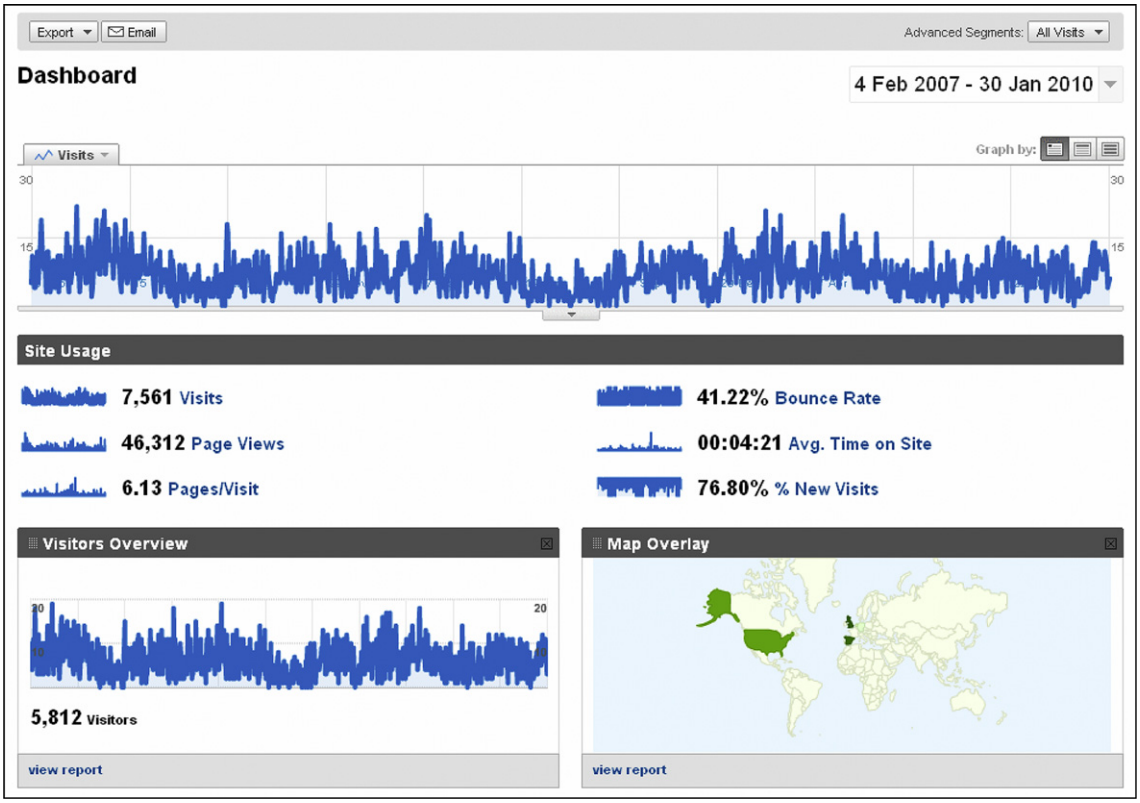

Fig. 1. Google Analytics dashboard overview for www.scholars-on-bilbao.info (daily data, 4 Feb 2007–30 Jan 2010). Source: Google Analytics for www.scholars-on-bilbao.info.

information search (for a literature review on information technologies in tourism see Buhalis & Law, 2008; Law, Leung, & Buhalis, 2009; Law, Qi, & Buhalis, 2010).

In addition, several scientific articles have analyzed the use of Google Analytics and evaluated its usefulness as a web analytics tool. (Hasan, Morris, & Probets, 2009; Plaza, 2009; Rodriguez-Burrel 2009). Some statistical matters with regard to the use of Google Analytics data in combination with time series methodology were developed and fine-tuned by Plaza (2009).

3. Methodology

3.1. Website profile

In July 2006 a non-profit organization (based in Gernika, Basque Country-Spain) launched http://www.scholars-on-bilbao.info (Art4pax Foundation, 2008) in order to improve the dissemination of R&D results in the field of 'Cultural Tourism' scientific production, through the exchange of research work on

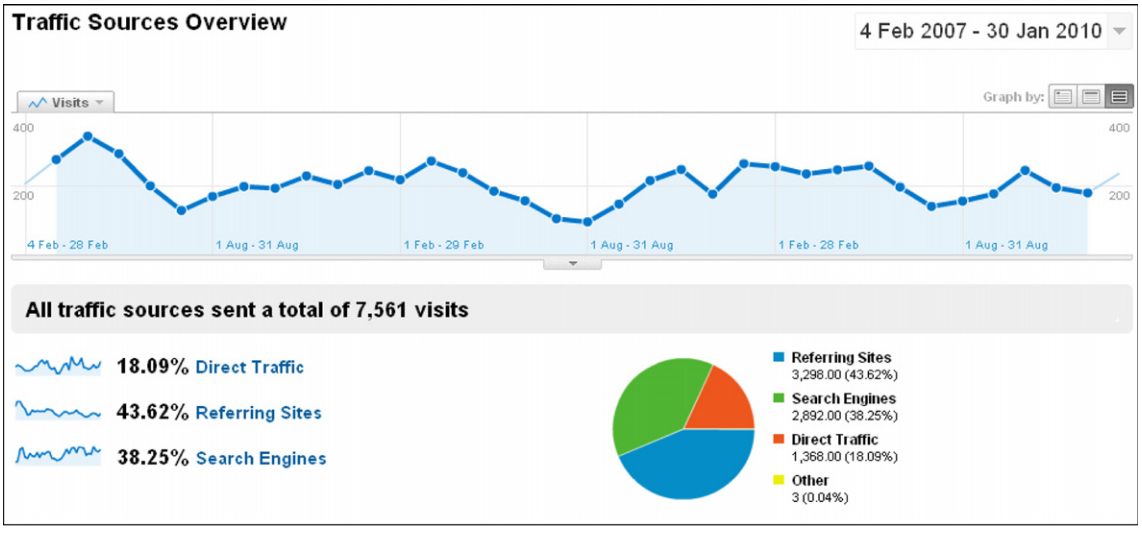

Fig. 2. Google Analytics traffic sources overview for www.scholars-on-bilbao.info (monthly data, 4 Feb 2007–30 Jan 2010). Source: Google Analytics for www.scholars-on-bilbao.info.

Top Traffic Sources

Sources	Visits	% visits
google (organic)	2,741	36.25%
en.wikipedia.org (referral)	1,820	24.07%
(direct) ((none))	1,368	18.09%
nl.wikipedia.org (referral)	392	5.18%
es.wikipedia.org (referral)	275	3.64%

view full report

Fig. 3. Top Traffic Sources for www.scholars-on-bilbao.info (4 Feb 2007–30 Jan 2010). Source: Google Analytics for www.scholars-on-bilbao.info.

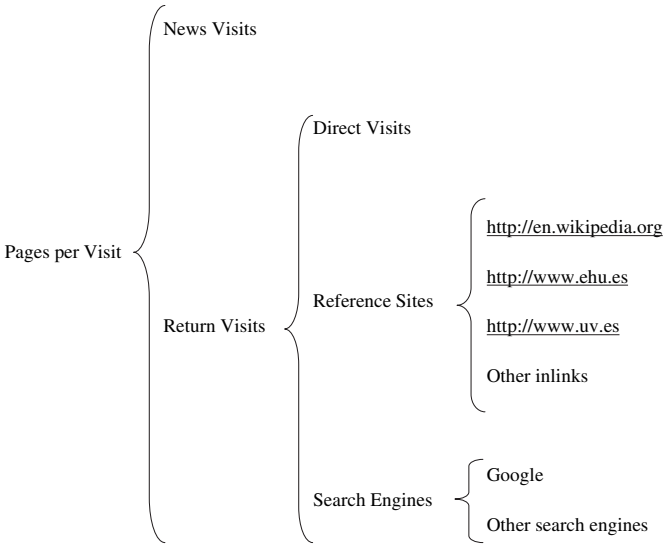

Fig. 4. Sequence of the performed regressions.

the Guggenheim Museum Bilbao case. This locally based website encompasses academic works that analyse the 'Guggenheim Effect' (cultural tourism, the Guggenheim Museum Bilbao and dilemmas, creative industries and artists). Each article includes the abstract and a web-link to its pdf/word file. Due to the fact that each one is displayed in a single page, the number of pages per visit tells whether the visitors are attracted by the content or not (visit length).

Table 1
Regression for pages per visit (weekly data, 4 Feb 2007–30 Jan 2010).

Variable	Coefficient	Standard error	*t*-Statistic	Probability
Constant	4.95	0.44	11.04	0.000
New Visits	0.001	0.013	0.09	0.935
Return Visits	0.069	0.022	3.068	0.002
$N = 156$				
$R^2 = 0.44$	*F*-statistic = 11.66			Prob(*F*-statistic) = 0.00
Breusch–Godfrey Serial Correlation LM Test: *F*-statistic 0.72				Probability 0.48
White Heteroskedasticity Test: *F*-statistic 1.22				Probability 0.27
Jaque–Bera 4.28				Probability 0.11
Augmented Dickey–Fuller Unit Root Tests for Variables:				
ADF Test Statistic for 'Pages per Visit': −5.8				5% Critical Value −2.88
ADF Test Statistic for 'New Visits': −3.2				5% Critical Value −2.88
ADF Test Statistic for 'Return Visits': −3.20				5% Critical Value −2.88

3.2. Methodology and hypothesis testing

Google Analytics allows users to export report data in MS Excel format, which when transformed can be analyzed with

Table 2
Regression for return visits (weekly data, 4 Feb 2007–30 Jan 2010).

Variable	Coefficient	Standard error	t-Statistic	Probability
Constant	−5.49	1.26	−4.33	0.000
Direct Visits	0.43	0.06	6.80	0.000
Referring Sites Visits	0.27	0.03	6.82	0.000
Search Engine Visits	0.37	0.04	7.66	0.000
AR(1)	0.32	0.07	4.12	0.000
$N = 155$				
$R^2 = 0.69$	*F*-statistic = 68			Prob(*F*-statistic) = 0.00
Breusch–Godfrey Serial Correlation LM Test: *F*-statistic 0.92				Probability 0.39
White Heteroskedasticity Test: *F*-statistic 1.11				Probability 0.35
Jaque–Bera 0.50				Probability 0.77
Augmented Dickey–Fuller Unit Root Tests for Variables:				
ADF Test Statistic for 'Return Visits': −3.20				5% Critical Value −2.88
ADF Test Statistic for 'Direct Visits': −3.84				5% Critical Value −2.88
ADF Test Statistic for 'Referring Sites Visits': −2.99				5% Critical Value −2.88
ADF Test Statistic for 'Search Engine Visits': −3.52				5% Critical Value −2.88

Table 3
Regression for Return Visits (weekly data, 4 Feb 2007 to 30 Jan 2010)

Variable	Coefficient	Standard error	*t*-Statistic	Probability
Constant	−5.50	1.28	−4.27	0.000
Direct Visits	0.42	0.06	6.46	0.000
en.wikipedia.org	0.23	0.06	3.38	0.000
ehu.es	0.38	0.27	1.41	0.160
uv.es	0.12	0.26	0.48	0.630
Other in-links	0.33	0.08	3.99	0.000
Google	0.39	0.05	7.48	0.000
Other Search Engines	0.21	0.24	0.86	0.387
AR(1)	0.32	0.08	4.04	0.000
$N = 155$				
$R^2 = 0.69$	*F*-statistic = 37.43			Prob(*F*-statistic) = 0.00
Breusch–Godfrey Serial Correlation LM Test: *F*-statistic 0.94				Probability 0.39
White Heteroskedasticity Test: *F*-statistic 1.33				Probability 0.18
Jaque–Bera 0.65				Probability 0.72
Augmented Dickey–Fuller Unit Root Tests for Variables:				
ADF Test Statistic for 'Return Visits': −3.20				5% Critical Value −2.88
ADF Test Statistic for 'Direct Visits: −3.84				5% Critical Value −2.88
ADF Test Statistic for 'en.wikipedia.org': −3.15				5% Critical Value −2.88
ADF Test Statistic for 'ehu.es': −4.10				5% Critical Value −2.88
ADF Test Statistic for 'uv.es: −8.64				5% Critical Value −2.88
ADF Test Statistic for 'Other inlinks': −2.46				5% Critical Value −2.88
ADF Test Statistic for 'Google': −3.38				5% Critical Value −2.88
ADF Test Statistic for 'Other Search Engine Visits': −4.33				5% Critical Value −2.88

time series statistical programs. In this case, the software *EViews* is made use of. A data set with 7561 entries for 1092 days drawn from Google Analytics was employed to analyse the performance of the website from 4 February 2007 to 30 January 2010; enough for obtaining valid results. Of those visits 1368 came directly to this site, referring sites sent 3298 visits via 121 sources, and search engines sent a total of 2892 visits (Fig. 2), mainly through Google (Fig. 3). Reference site traffic is, by far, the main source of entries for www.scholars-on-bilbao.info: almost 44 per cent of the total incoming visits; that is, 3298 entries through in-links, 1820 of which enter from en.wikipedia.org (Fig. 3). But 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? How deep do Google entries navigate?

Several regressions are undertaken (see the sequence of regressions in Fig. 4). Weekly data is utilized (Plaza, 2009). Dickey–Fuller stationary tests are calculated for each variable and all show stationary (Tables 1–3). The Breusch–Godfrey Serial Correlation LM Test is used to check serial autocorrelation. The White Test is used to test heteroskedasticity, and the Jarque–Bera statistic to test normality. The presence of outliers is corrected through the use of dummies. The roots of the AR and MA processes are outside the unit circle. The regressions are well-adjusted. The fitted estimations are shown in Tables 1–3.

4. Results

Results from Table 1 show that the number of pages per entry grows by 0.06 out of every return visit, whereas the marginal effect of new visits is nil. That is to say that return visits are the main engine for nurturing session length for www.scholars-on-bilbao.info (see Fig. 5), and bounce less (Fig. 6). But, which type of traffic source nurtures these return visits?

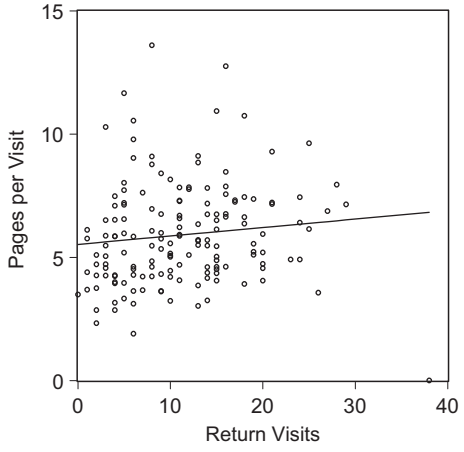

Fig. 5. Return visits navigate deeper into the website and stay longer (weekly data, 4 Feb 2007–30 Jan 2010). Source: Google Analytics data for www.scholars-on-bilbao.info.

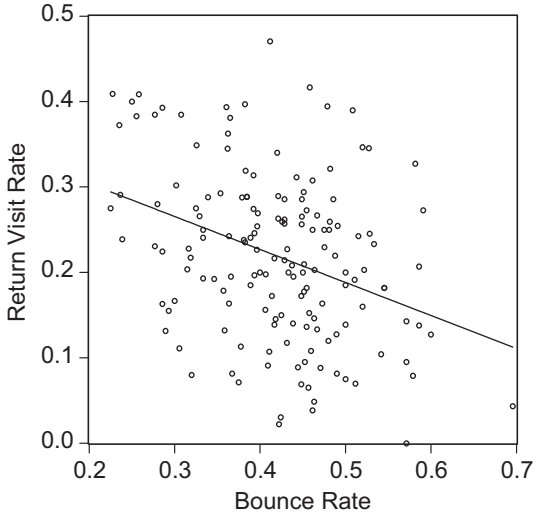

Fig. 6. The less the bounce rate, the more the return visit rate (weekly data, 4 Feb 2007–30 Jan 2010). Source: Google Analytics data for www.scholars-on-bilbao.info.

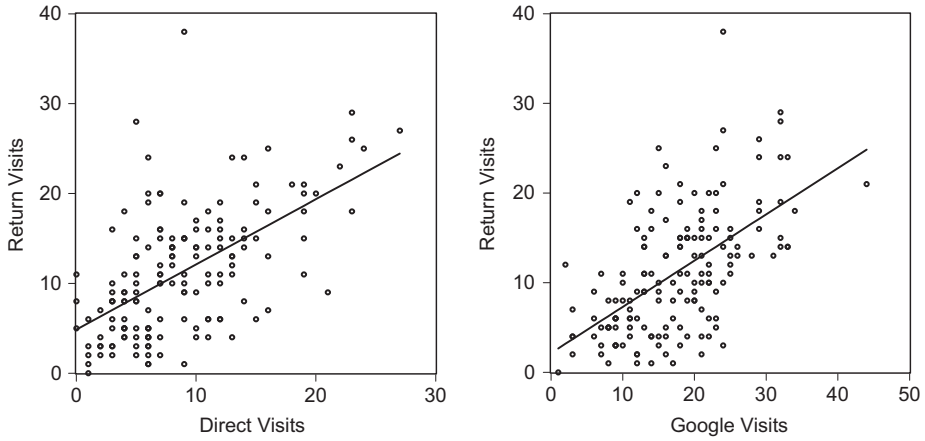

Fig. 7. Direct visits and Google entries return the most (weekly data, 4 Feb 2007–30 Jan 2010).

According to the reading of the results in Table 2, 0.43 out of every direct entry visit returns, 0.36 out of every search engine entry visits the site again, and only 0.24 out of every referee site visit returns. In other words, for our particular website, direct visits are the most effective ones, followed by search engine visits and only thirdly link-entries (Plaza, 2009).

According to Table 3, the effectiveness of the in-links from www.ehu.esandwww.uv.esis null, whereas 0.21 out of every http://en.wikipedia.org driven entry visits the site again, and 0.29 out of every 'Other In-links' visit returns. In other words, for our particular website, http://en.wikipedia.org driven entries are effective, showing an adequate return visit behaviour and length of sessions; although 'Other In-links' are shown to be even more effective with 0.33 return visits per entry.

Last, but not least, visits through Google are also shown to be effective, with 0.39 return visits per Google entry. The effectiveness of other search engines shows null for this particular website (Table 3).

In summary, for our particular website direct visits (Fig. 7) are the most effective ones, followed by Google entries (Fig. 7) and only thirdly http://en.wikipedia.org visits. Moreover, the performed time series analysis with Google Analytics shows 1) that return visits navigate deeper into the website and stay longer, and 2) that the less the bounce rate (error entries), the greater the visit duration (pages per visit and/or time at website). The importance of this article is not the particular website, but the methodology tested to arrive at these results. The agenda for future research requests the repetition of the experiment with different tourism websites, to delimit more accurately the effectiveness of different traffic sources and to compare these results with other case studies. Tourist firms have to revolutionize their web analytics strategy with effective methods that can assist practitioners to evaluate their website performance and subsequent online marketing effectiveness.

References

Art4pax Foundation. (2008). Scholars on Bilbao. http://www.scholars-on-bilbao.info.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the internet—the state of eTourism research. *Tourism Management, 29*(4), 609–623.

Hasan, L., Morris, A., & Probets, S. (2009). Using Google analytics to evaluate the usability of e-commerce sites. *Lecture Notes in Computer Science, 5619*, 697–706.

Law, R., Leung, R., & Buhalis, D. (2009). Information technology applications in hospitality and tourism: a review of publications from 2005 to 2007. *Journal of Travel & Tourism Marketing, 26*, 599–623.

Law, R., Qi, S., & Buhalis, D. (2010). Progress in tourism management: a review of website evaluation in tourism research. *Tourism Management, 31*(3), 297–313.

Plaza, B. (2009). Monitoring web traffic source effectiveness with Google Analytics. An experiment with time series. *Aslib Proceedings, 61*(5), 474–482.

Rodriguez-Burrel, J. (2009). Google Analytics: good and nice and free. *Profesional de la Información, 18*(1), 67–71.