CHAPTER 3

How Web Analytics Works

INTRODUCTION

A basic understanding of how web analytics tools work is important for understanding their limitations and for interpreting the data you get.

There are two basic ways you can learn about user activity on websites: log files and page tagging. Log files keep track of what pages load upon each web page request and allow you to take a deep dive into a rich data set, but they can be challenging to use and deploy. Page tagging is less accurate than log files, but there are usually fewer obstacles to getting started with the process and when using page tagging tools. Within the category of page tagging, you will find tools that allow you to analyze how users move from page to page and tools for analyzing what they actually do on each page. This book mainly draws on Google Analytics for examples, which falls into the category of page tagging.

LOG FILE ANALYSIS

One of the two approaches to web analytics is log file analysis using tools like AWStats and Sawmill. It is worth understanding how this approach differs from page tagging, and while log file analysis is outside the scope of this book, the concepts covered in this book transfer well. Web servers keep records of transactions. Every time they get a request, whether it is from a browser or a search engine crawler, it gets recorded in the log and doesn't require that the user have JavaScript enabled, which is a weakness of page tagging tools. On the other hand, if the user accesses a cached version of a page rather than getting it from the server, the log doesn't record the transaction. The result is that the data are imperfect in different ways.

Why aren't we analyzing log files? Partly, it's because page tagging tools are well marketed and actually quite easy to use, relatively speaking. The other reason is that it may be far more challenging and expensive to install a log file analysis tool than to set up a page tagging tool. Log files are huge, and it can be challenging for IT departments to share them. Also, they may periodically

CONTENTS

Practical Web Analytics for User Experience. DOI: http://dx.doi.org/10.1016/B978-0-12-404619-1.00003-4

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" dir="ltr" lang="en-US">

    <title>Page Title</title>

<script type="text/javascript">

  var _gaq = _gaq || [];
  _gaq.push(['_setAccount', 'UA-XXXXXX-1']);
  _gaq.push(['_trackPageview']);

  (function() {
    var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
    ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') + '.google-analytics.com/ga.js';
    var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
  })();

</script>

</head>

<body>
```

FIGURE 3.1

An example of Google Analytics' tracking code. This small amount of code is what makes it work. Google Analytics will generate the appropriate code for your own website.

purge them, making it impossible to get much historical data. Installing the appropriate software and configuring databases appropriately may simply be beyond the IT capabilities of an organization. With a page tagging tool, you add a few lines of JavaScript to the pages of your website, which you may be able to do through the content management system that powers your website rather than manually changing every page. Based on your level of expertise or the amount of resources available to you, there's a good chance that it will simply be easier to get started using a page tagging tool. On the other hand, some kinds of highly interactive websites, such as those that make heavy use of AJAX, may be much harder to track with a page tagging tool than with log files—they would require extensive additional tagging of page elements, whereas log files, by their nature, simply track every request to the server.

PAGE TAGGING

The page tagging method works by adding a small piece of JavaScript code to all of the pages on your website that you want to track (Figure 3.1). When a user's browser loads a page, it opens up the HTML file and starts interpreting the code. When it hits the JavaScript code, it will send the following data to a database:

- What page just got loaded
- When it was loaded

- Where the user just came from (in terms of search engine or a link from another website)
- IP address
- Technical details like the user's browser, operating system, screen resolution, and colors

It also puts a cookie on the user's computer that will let the analytics tool know whether it's already recorded data from that user. From this small amount of data, web analytics tools are able to provide a rich set of information about user activity. By combining data about when pages were viewed with anonymous data on who viewed them, these tools can map out the sequence of pages that users viewed during their visit.

With page tagging solutions, there are two main categories: tools for tracking activity as users move from page to page, and tools for tracking activity on pages. Google Analytics falls mainly in the "move from page to page" category, as does Webtrends and Omniture (although they have capabilities for measuring on-page behavior).

GOOGLE ANALYTICS ASYNCHRONOUS TRACKING CODE

If you start looking at the source code on other people's websites, trying to find Google Analytics tracking code, sometimes you'll see it at the bottom of the page before the </body> tag and sometimes at the top of the page before the </head> tag. What you're seeing is technology in transition. Back in the "old" days of 2008, a page waiting to load a script tag would hold up the rest of the page load, so industry best practice was to put it at the bottom of your HTML. Having the code at the bottom does no harm to the end user, but lowers accuracy—when tracking code is placed at the bottom of a page, any time the user stops loading the page before it reaches the bottom, that page isn't tracked. There were many requests for development of a tag that could be at the top of the page and would also not slow down page loads. A couple of years ago, Google Analytics did just that, introducing a version of their tracking code (asynchronous code) that runs in the background while the page continues to load. That means it can go at the top of the page. This new code makes Google Analytics more accurate (sometimes much more accurate, depending on the website), but not everyone is using it yet, which is why you will still sometimes find tracking codes at the bottom of pages. If you're working on a website with the old version of the code or with the new code at the bottom of the page (which is surprisingly common), send the word out about the new code and try to get it changed and/or moved up to the top.

Cookies

A cookie is a piece of data that a website stores on users' computers. They are used to keep track of things like user preferences, what the user was doing on

the website the last time he or she visited, and to keep track of whether a user is currently logged on to a website or not.

Web analytics tools use cookies to track whether a user has visited a website before and how the user originally got to the website (e.g., through a search engine). As the user moves from page to page, the cookie lets analytics know that these pageviews are all part of the same visit. If the user does not load a page on the website for a certain period of time, the visit ends.

Accuracy

The data you get from page tagging web analytics tools are not 100% accurate. The tools do, however, provide you with data that are good enough to work with, because they are internally consistent and the amount of error compared to the real amount of traffic is consistent.

The problem is that there is no perfect way to collect data about users' actions. Because page tagging tools rely on JavaScript, if a user's browser has JavaScript disabled, the tool won't gather any data. If users have installed a browser plug-in to prevent analytics tools from tracking their actions or activated privacy mode, then you won't be able to gather any data about them. Also, sometimes the Internet or the analytics tracking script just doesn't work properly and data don't get collected.

What makes web analytics useful is that the data are consistently accurate to within 10% of actual traffic numbers. Consistency lets you measure trends over time and make actionable decisions based on data.

We are left, then, with just a minor and somewhat philosophical point: the data in your web analytics tool are not a definitive statement of what really happened, but rather just what you can measure about what really happened. If you read that 10,467 people visited your homepage in a single week, that's simply how many visitors the tool recorded. Another web analytics tool (you can run more than one on the same website) would give you a slightly different number, a log file analysis would produce a slightly different number, and if you had perfect knowledge, you would know that the actual number was slightly different. If accuracy is critical, such as counting how many sales you received or how many people registered for an account, you should go to a more appropriate tool (like a CRM tool) as your source for "reality."

SMARTPHONES

You will be able to gather analytics data from users with smartphones such as Apple's iPhone. Smartphones have JavaScript-enabled browsers that can run JavaScript. However, you will probably not be able to get data about how people with feature phones (i.e., phones

that don't have JavaScript-enabled browsers) use your website. Depending on your users, lack of JavaScript capability could be a trivial concern or a huge problem—the United States has relatively few feature phones accessing websites, whereas developing countries see a great deal of usage.

Accounts and Profiles

In Google Analytics, each website has its own analytics account and each account has one or more profiles (Figure 3.2). Other web analytics tools use different models for organizing data, and if you are using one of them, you can safely move on to the next section.

The *account* is the repository of data. Each account has its own unique identifier (e.g., Pure Visibility's Google Analytics account is UA-461660) that ensures the tracking code sends data to the correct place.

A *profile* is the way the data are actually displayed to you. A profile may just present you with all the data in the account, but at the very least you will probably be working with a profile that filters out the IP address of the place you work so you don't skew the data. You can set up profiles to transform the data further, such as isolating parts of a larger website or rewriting URLs to cluster similar pages for easier analysis. Chapter 15 goes into greater detail on how profiles work.

This book focuses on Google Analytics for examples, but it is not the only web analytics tool that organizes data into profiles. This is a common and useful feature for organizing data into meaningful chunks.

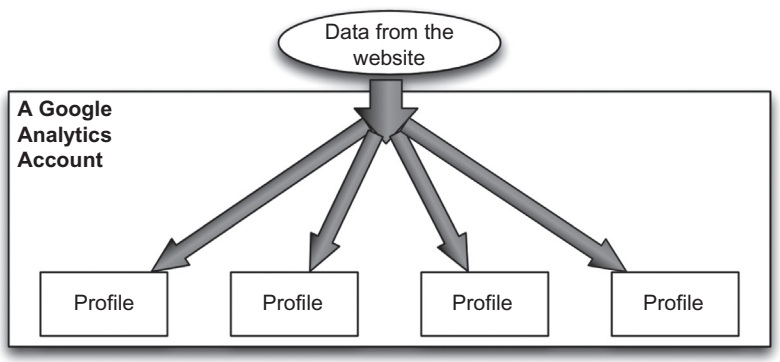


FIGURE 3.2

Data about website usage flow into a web analytics tool and then into one or more profiles. In Google Analytics, an account receives the data and contains one or more profiles.

INSTALLATION

This book doesn't go deep into the subject of setup and configuration—the concern is more with practical use of tools, and most readers will be in situations where other people are responsible for implementation. It may be helpful, though, to understand the basic steps of setting up web analytics. The general outline is that installation consists of three basic stages:

1. Create an account.
2. Set up profiles and filters.
3. Generate the tracking code to install on your website and install it.

These steps will vary from tool to tool; with Google Analytics, they are done without intervention from Google. Other tools will require interacting with the vendor for the first two steps.

In practice, steps 2 and 3 can be anything but simple and there may be multiple iterations of profile configuration. Setting up filters, which is the way you determine what data do and do not get captured, has the potential to be a tricky and time-consuming operation. It may involve things like finding IP addresses that your company uses so you can filter out your own visits. Generating tracking code entails putting together the JavaScript that has to go on every page or writing code to track interactive features on your website.

Click Analytics

There are also tools for analyzing user behavior on a page itself rather than movement between pages. They can capture where specifically users click (e.g., fields on a form or on a part of a page that look clickable but aren't), where they move the mouse, what keys they press, and how long users spend looking at the top of a page before scrolling down—basically, they let you make recordings of a user's session. You can view these data in aggregate—for example, by looking at a page of your website overlaid with a dot for everywhere that users clicked. Some tools even offer you the ability to play back individual user sessions as if you were standing over the user's shoulder.

These tools let you answer questions like these:

- In what order do users fill out fields in a form?
- Are there fields that take a long time for users to fill out?
- Where do users click on things that aren't clickable?
- On a clickable object, where exactly do users click?

Much like a page tagging analytics tool, click analytics tools involve inserting a snippet of JavaScript onto the pages you want to track.

Click analytics products, such as Tealeaf, CrazyEgg, and ClickTale, to name a few at the time of writing, are worth exploring if you want to gather more

evidence about user actions on pages. Chapter 12 will delve further into click analytics tools and analysis.

METRICS AND DIMENSIONS

Metrics and dimensions are key concepts in web analytics. Metrics are numeric measurements of various aspects of users' behavior, like how long they spend viewing a page or how many times users viewed a page.

Dimensions are the categories that user data may be grouped into, such as what browser they used or what keyword they searched for. Dimensions describe different attributes of users, their computers, how they got to your website, or even parts of your website that users visit, which you can use to divide your users into segments. When dimensions are paired with metrics, you can learn things like how many users use each kind of browser, how many views each page receives, and how many visits came from Texas versus New York last Thursday.

Let's take a look at a metric and a dimension to try to unpack this relationship. We will cover them in greater detail later, but a kind of metric is the pageview and a kind of dimension is the URL of a page. Every time a user loads a page in his or her browser, web analytics counts it as a pageview.

One can simply count up all of the pageviews that a website receives—that is, every time every single page on a website was loaded in a user's browser. However, the pageviews metric becomes more useful when paired with the page dimension. You can count how many pageviews each individual page of your website received. You can count how many pageviews a group of pages received, such as the "About Us" section of your website. The page dimension gives you a way to divide all of the pageviews that web analytics measured into meaningful, useful segments.

Another possible pair is the average time on website metric and the mobile dimension. The mobile dimension only has two possible values—yes or no—based on whether or not the user visited the website using a mobile device like a phone or a tablet. Average time on website is the average amount of time that a group of users spent on your website. You can find out the average amount of time that all of your users spent on your website, but when you combine this metric with the mobile dimension, you can see the average amount of time that mobile users spent on your website and compare it to the average amount of time desktop computer users spent on your website. Again, the mobile dimension is just one possible way of segmenting metric data.

The term *metric* is widely used in web analytics tools, but *dimension* is Google's terminology. The concept of a dimension is used in other tools,

though (e.g., you have *elements* in Omniture), and overall you will find that the specific metrics and dimensions in this chapter exist in other analytics tools. They simply may vary in how easy they are to get to and what they are labeled.

The best way to understand these concepts better is to spend a bit of time seeing them in action. Later in this chapter, we will return to the concept of dimensions in the section "Interacting with Data in Google Analytics." For now, let's start with metrics, with an example of a report from Google Analytics (Figure 3.3). This is the "Visitors Overview" report, which gives you high-level data about the users coming to your website, in the sense that these values are sums and averages for all of the visitors to the website. The "Visitors Overview" report has a large graph that shows a specific, selectable metric over time, and then high-level metrics about the visitors to your website.

Visits

A visit is a single time that a person comes to a website, clicks around and views some pages, and then leaves. By itself, visits isn't a very useful metric, but much of this book involves slicing the total number of visitors into smaller segments according to how those visitors behave on the website, so the visits metric will come up repeatedly.

THE LIMITS OF TRACKING VISITS

Analytics tools can't get a perfect measurement of how long a visit lasts, because the tools can't actually tell when a user leaves the website they are tracking (see the "Average Visit Duration" section for more information). In Google Analytics, a visit ends when the user hasn't viewed any pages for 30 minutes (these 30 minutes are excluded from any calculations of how long the user spent on the website). If the user comes back to the same website later, it is a new visit—even if the user goes to the website, leaves it open in their browser for several hours without touching it, and then starts using the website again.

Unique Visitors (Metric)

If a user starts a new session on a website, analytics counts that as another visit. Unique visitors is a way of capturing how many individuals came to a website and whether they visit multiple times or just once.

A visitor is only unique within whatever timeframe you've selected—it doesn't reflect whether that user visited before or after the dates that you are analyzing. Analytics tools judge whether or not a visitor is unique by the presence of a cookie, so there's no way of knowing whether the same person is

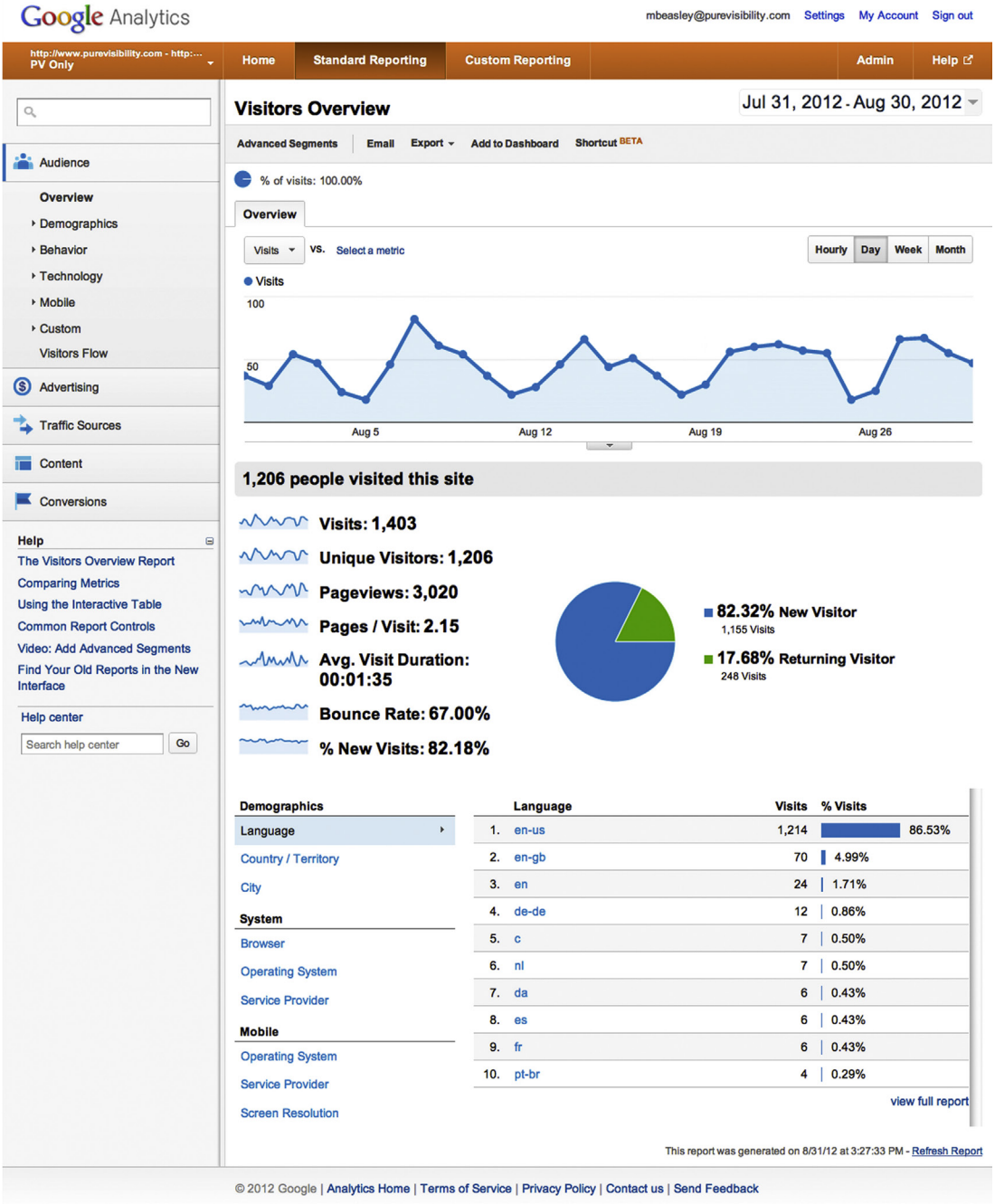


FIGURE 3.3

The "Visitors Overview" report, which shows high-level metrics for your entire website.

using the computer or if multiple people are. If a single person uses a desktop computer and then visits a website from his or her phone, he or she gets counted again as a unique visitor. This point is particularly important, as users increasingly access the same website through multiple devices, often as part of a single overarching task. People often focus on unique visitors because they want to learn how many individual people are coming to their website; however, its limitations make it better suited to using it in combination with other metrics rather than as a single, somewhat misleading, measurement of visitors to a website.

Google Analytics explicitly uses the label "unique visitors" to better differentiate it from "visits." You will also find "visitors" in other tools to express the same thing as "unique visitors."

UNIQUE VISITOR CAVEATS

Unique visitors falls under the category of "nice to know" but usually appears more useful than it actually is. Even though stakeholders may be very curious to learn what the real number of visitors is, knowing this number generally doesn't lead to any different decisions than the regular visits metric. More importantly, "unique" is relative to the time you're looking at. A visitor may be unique for the week you're looking at, but actually came to your website five times in the previous week. If you look at unique visitors one week at a time, that one user will get counted as one unique visit in both weeks. If you change your time range to cover both weeks, those unique visits don't get added together; rather, that user just shows up as one unique visit. You will get more value from investigating the behavior of new versus returning visitors, and perhaps, among those returning visitors, those that return frequently and those that do not (depending, of course, on who your users are).

Pageviews (Metric)

A pageview is a single time that a user went to a page. If the user goes to a page more than once during his or her visit, it's a separate pageview each time. In Figure 3.4, we can see that the user went to six pages total, repeating one of them. Each one counts as a separate pageview.

In Figure 3.4, if the person was the only person to use the website, all of the pages would have just one pageview except /widgets/, which would have two pageviews. The most obvious use of this metric is to get a relative measure of what pages people are visiting and which ones they're not. Comparing the pageviews of different pages will help you find out what pages are popular or unpopular with users. When you make design changes to your website, one of your goals may be to get more people to a particular page—that is, increase pageviews.

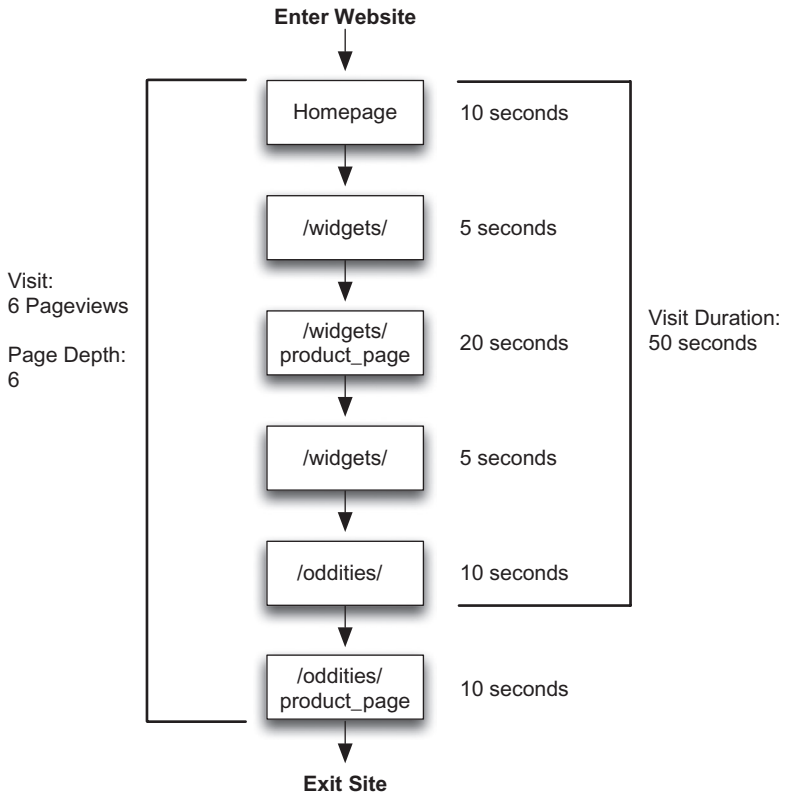


FIGURE 3.4

This diagram shows the relationship between pageviews, visits, and visit duration. A visit consists of one or more pageviews. The amount of time between each pageview is added together to calculate the visit duration.

Pages/Visit (Metric)

Pages/visit and page depth both capture the number of pages that users go to during their visit. If they visit the same page multiple times, each additional pageview is counted.

As with many other metrics, there is nothing intrinsically good about high or low numbers, but in a practical sense, high numbers are usually good because they usually reflect users taking the time to explore your website and/or do stuff. It is often useful to compare this metric across different segments of users or over time.

Average Visit Duration

Average visit duration is the amount of time that a user spends on your website. As previously discussed, any time that the user spends on the last page he

or she views isn't counted, which also means that users who bounce from the website have a visit duration of 0.

As we dive deeper into analytics data, it is common to compare two or more populations, usually as a proxy for how engaging or "sticky" a website is. When segmenting users in Google Analytics, you can do it according to visit duration, such as filtering out users whose visit duration was less than 45 seconds long, for example.

Bounce Rate (Metric)

The bounce rate of a page or a group of pages is the number of users who entered a website on a page and left without visiting any other pages, divided by the total number of users who entered the website on that page. It only counts users who start their visit on a page and excludes anyone who navigated from another page on the website. It's useful to keep this in mind, because the bounce rate may be calculated based on a very small number of visitors, if it is a page or group of pages that is deep within a website.

A low bounce rate is generally a good thing. A high bounce rate may be fine if the purpose of a page is to show the user some information and then send him or her on his or her way without doing anything, but that situation may not come up very much. As with other metrics, "high" and "low" are relative to other pages on your website (although in my experience, bounce rates lower than 30% are quite low).

% New Visits (Metric)

The % new visits, for a given set of users, is the number of users who have never been to the website before divided by the total number of users. In Google Analytics, "new" means they haven't visited within the last two years. Unfortunately, this metric doesn't reflect people who are using a different computer or even a different browser or who clear their cookies.

The % new visits metric can be useful as a way to describe users who you are analyzing or as a way to measure success if you are attempting to get users to come back to your website after their initial visit.

VISITOR TYPE (DIMENSION)

A closely related dimension is visitor type. Users are classified as either "new" or "returning" based on whether or not their browser has a cookie from this website from the past two years.

Using These Metrics

In isolation, these metrics won't answer particularly interesting questions. They serve more as jumping-off points for deeper inquiry—it's when we dig into other reports and compare these metrics over time and segment them according to users' behavior, whether according to a single dimension or a combination of dimensions, and filtering by the values captured as metrics, that we are able to gain the deepest insights.

INTERACTING WITH DATA IN GOOGLE ANALYTICS

These metrics and others become more meaningful when we split them up according to different dimensions (and, in later chapters, combinations of dimensions)—that is, basic segmentation. To see segmentation in action, we turn to an example of a typical report in Google Analytics. If you click on Content in the side navigation, then Site Content, and then All Pages, you will reach the "All Pages" report. We will return to this report in Chapter 7, but for now, we look at this report to learn more about interacting with data within the tool (in later chapters we will export data and work with them). Interacting with tables of data is one of the most important and common activities you will engage in when working with web analytics, and this section will give you a sense of the power you have to manipulate data.

Again, we see a big chart at the top of the report that displays a metric over time (days, weeks, months, and, in some reports, hours). It can display any of the metrics that you see in the table in the lower half of Figure 3.5.

Below the large chart, there are averages or sums for all of the metrics in the report, for all of the data in the entire time period selected—not just the data visible in the table.

The table is where we encounter a dimension in use to divide up metrics. In this case, the dimension is page, with each row showing the various metrics' data for just that page. We can see in Figure 3.5 that the average time on page for the various pages ranges from 33 seconds (/about/jobs/) to 3 minutes and 52 seconds (/services/archives/seo-articles/delivery-methods/), but in this report, above the table, we can see that the average amount of time that users spend looking at pages for the entire website is 1 minute 24 seconds.

The rest of this chapter deals with interacting with the data in the table itself, which will be a common activity in Google Analytics because so much of the data are displayed in a tabular format. You will spend a great deal of time looking at data that are divided up across a single dimension (and, later in this chapter, two dimensions in the same table), and the main difference between many reports is what dimension it uses to segment the data.

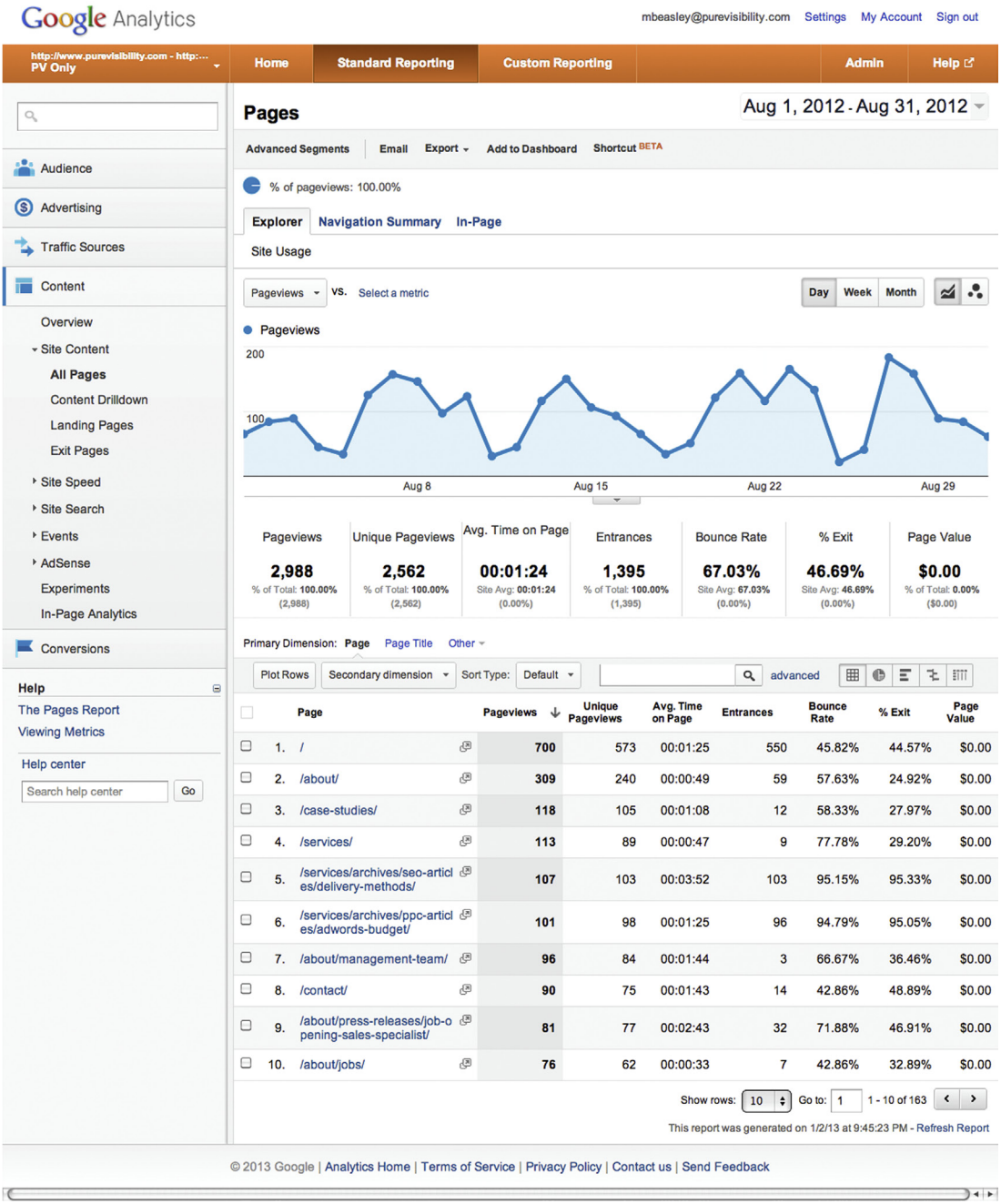


FIGURE 3.5

An example of the "All Pages" report, a typical report in Google Analytics from an interaction perspective.

Plot Rows

There is a checkbox on the far left side of every row of data. You can check one or more boxes and then click on the Plot Rows button to change the big chart so that instead of displaying a line for all data combined, it also shows lines for the rows you have selected. In Figure 3.5, the chart shows the sum of all pageviews every day. You could select just the homepage and the /about/ page (rows 1 and 2 of the table) and display how many pageviews those pages receive per day, as in Figure 3.6. It can be particularly helpful to do so to find out if a page follows the same overall trend as the rest of the website, or if it moves different, such as having a sudden surge of pageviews.

Secondary Dimension

By default, the table takes a single dimension, such as all the pages that users may view, and displays a row for each value. The Secondary Dimension button lets you add another dimension column to the table. When you add another dimension to a table you go to a finer level of granularity with your data because you are now segmenting it according to two dimensions. For example, instead of just looking at usage data for pages on your website, you could then divide the data for each page into whether the user was new or returning (Figure 3.7).

Sort Type

By default, tables are sorted according to some metric (often, visits or pageviews), typically in descending order. In the example of Figure 3.5, the data are sorted starting with the page with the most pageviews, the home page, and then the next most viewed page, and so on. If you are comparing two time ranges, it sorts according to the data from the more recent time range, as seen in Figure 3.8.

When comparing time ranges, you can choose two other sort types from the Sort Type dropdown menu. One is Absolute Change, which sorts the rows according to the size of the difference between the two time ranges. You can see in Figure 3.9 that whereas the homepage had the most pageviews (as shown in Figure 3.5), the page /about/press-releases/job-opening-sales-specialist/ had the greatest growth in pageviews when comparing August to May. Sorting tables by Absolute Change often fills the top rows of a table with data that are not very useful, as in the case of Figure 3.9, where the top 10 pages are rarely viewed compared to other pages in the website. This situation is where Advanced Search can be helpful—you can use it to filter out any pages with fewer than 100 pageviews, for example.

Another sorting feature that is occasionally useful is dropdown option Weighted Sort. Weighted Sort is only rarely available in Google Analytics and

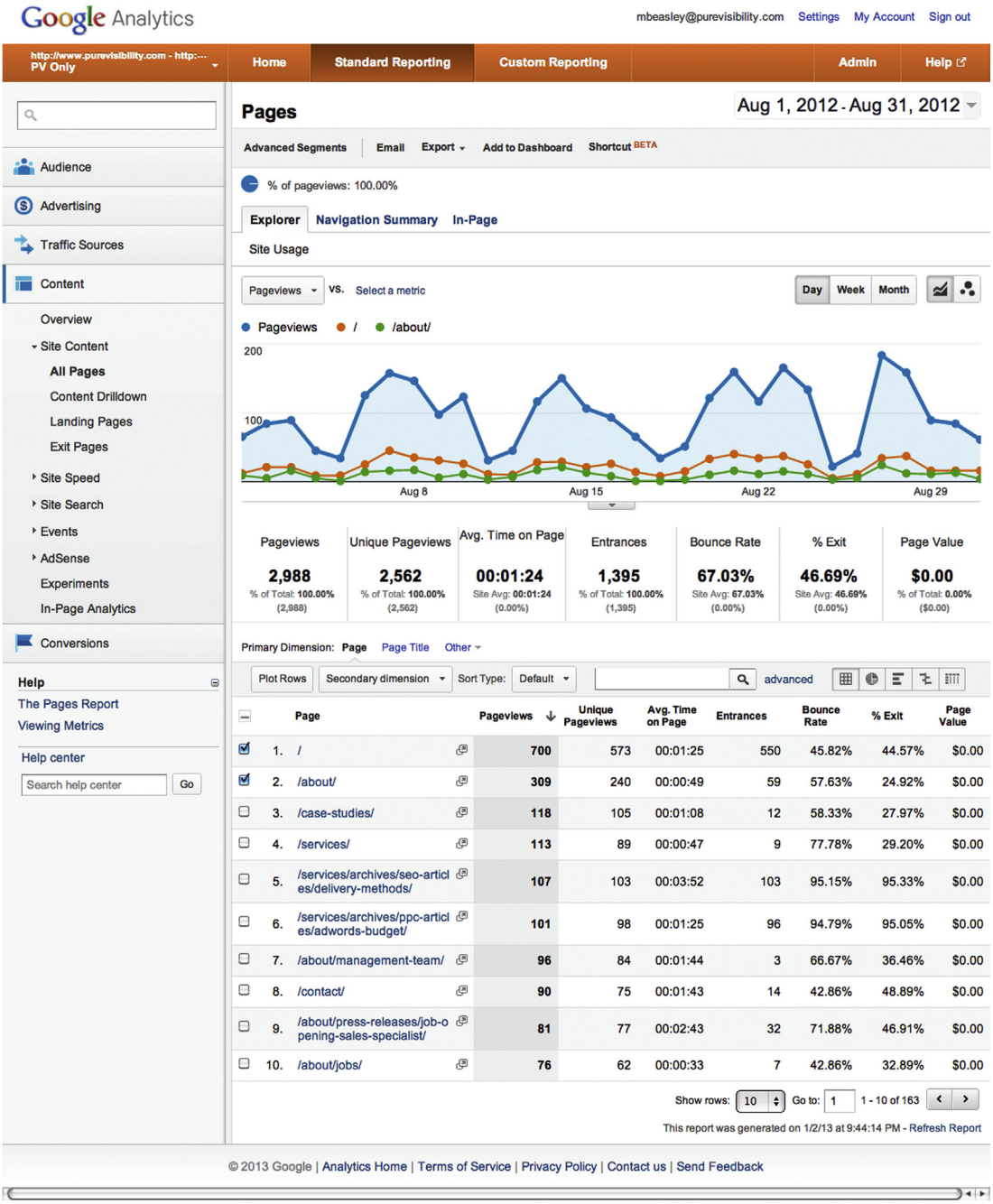


FIGURE 3.6

The report from Figure 3.5, but with only data from rows 1 and 2, the homepage and /about/ page, displayed in the graph. This was done by selecting them using the checkboxes on the left side of the table and clicking on the Plot Rows button.

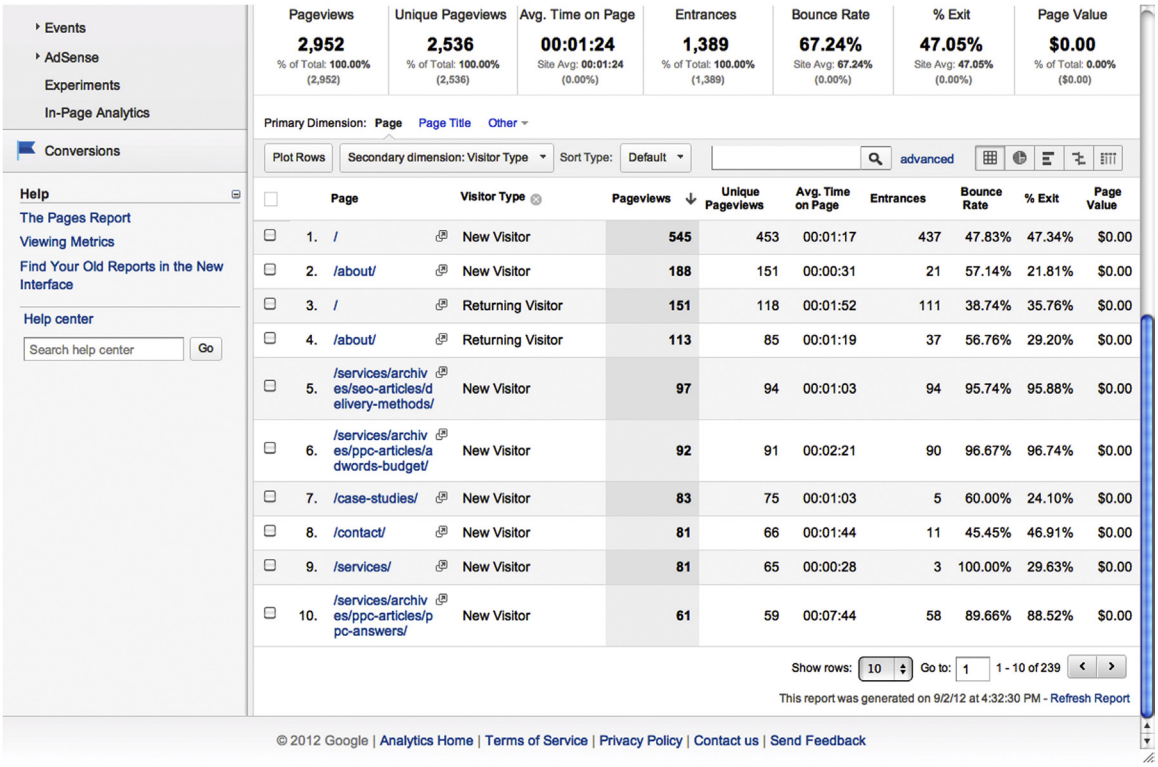


FIGURE 3.7

The data from Figure 3.5, but with a secondary dimension added to divide data about pages according to whether the users were new to the website or returning.

is a sorting algorithm that balances the degree of change in two values against the size of those values.

Search

There are two types of search available: simple and advanced. The simple search lets you enter a string of characters, and then it shows only the rows of data where that string appears. If you were to use simple search for the report in Figure 3.5, you could search for "services" and it would only show data for the pages that include "services" in the URL. Performing a search will also change the data above the table that shows sums and averages (Figure 3.10).

Advanced search is, unsurprisingly, more advanced. It lets you filter the table according to metrics and dimensions that aren't explicitly exposed in the table and combine multiple filters using AND and OR. This feature works much like advanced segments, which we will cover in Chapter 9.

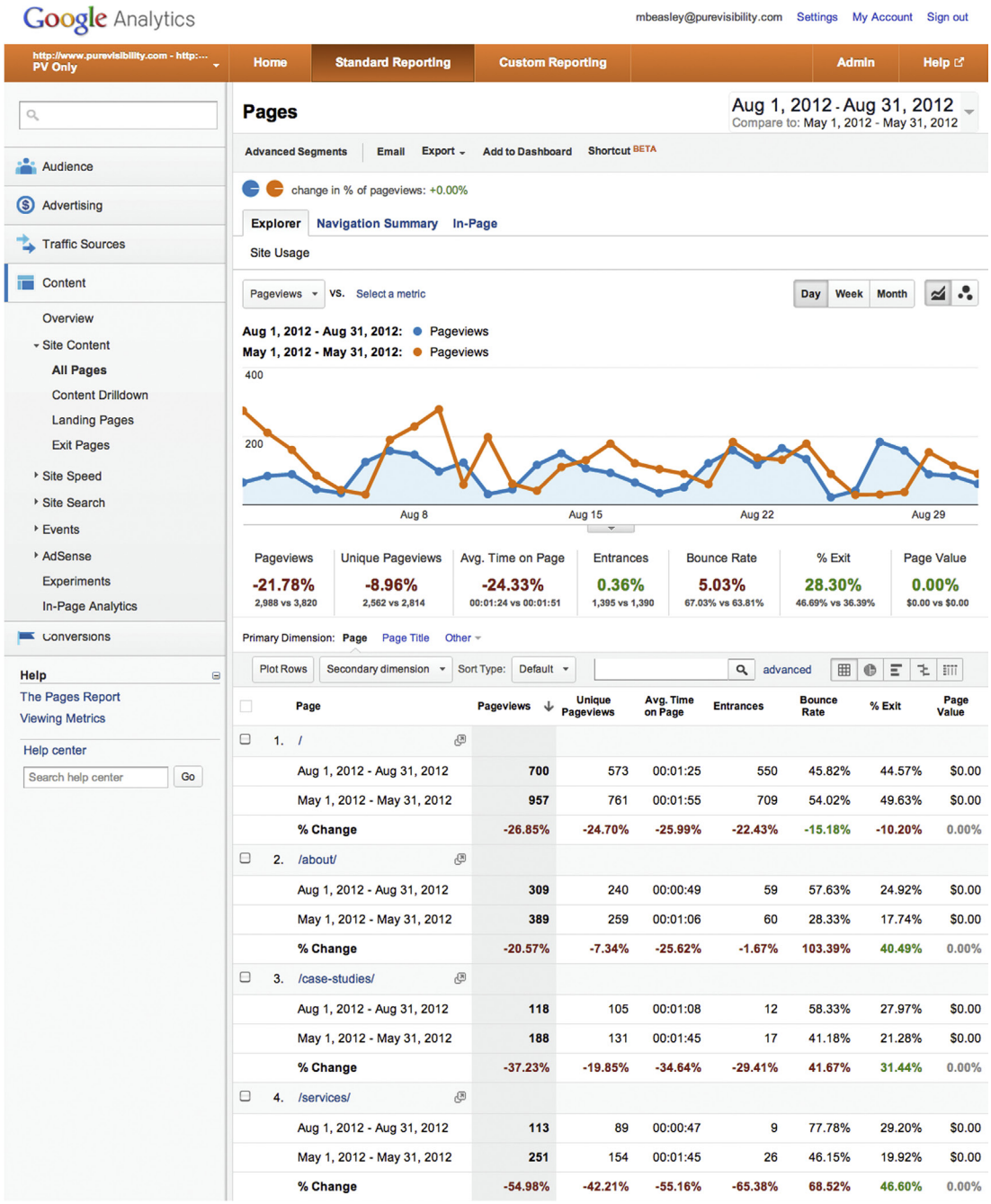
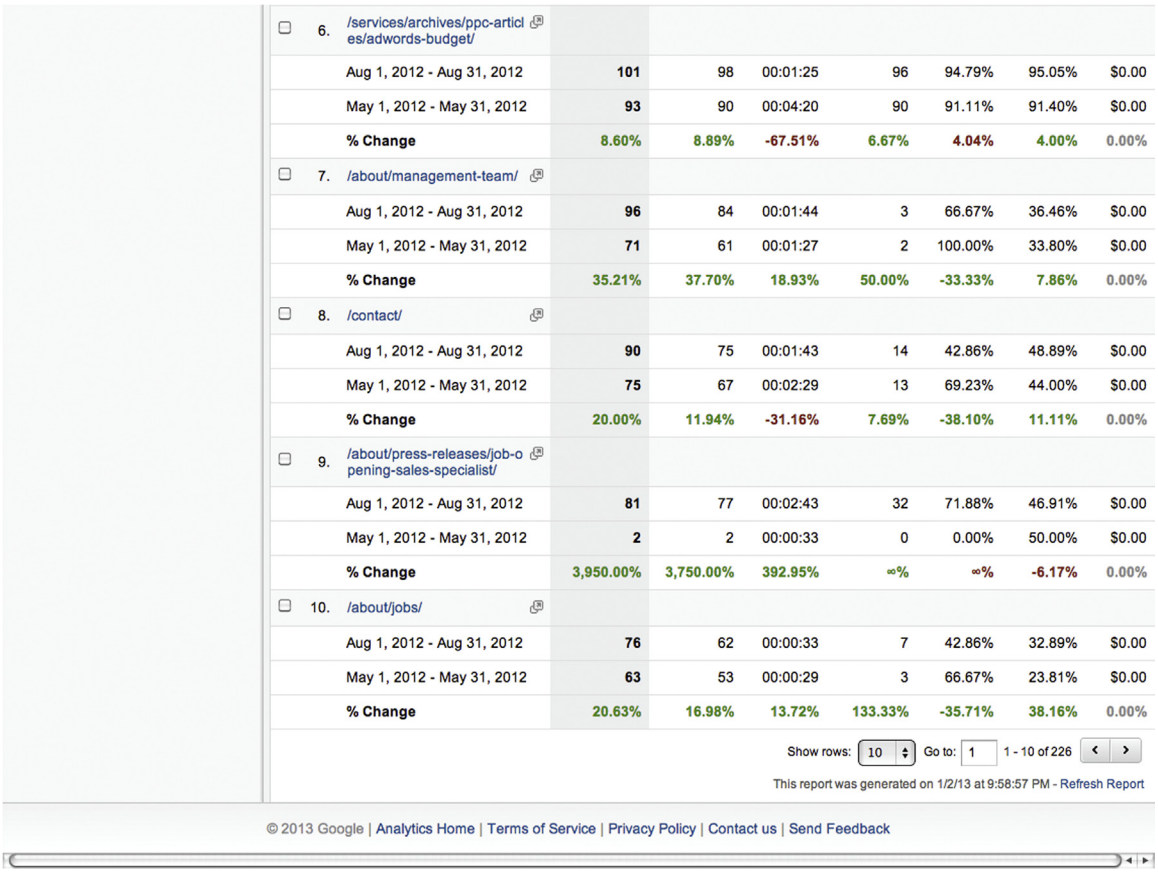


FIGURE 3.8

The data from Figure 3.5, compared to data from an earlier month. Note that the data are sorted in order of pageviews from the more recent month.



6.	/services/archives/ppc-articles/adwords-budget/							
	Aug 1, 2012 - Aug 31, 2012	101	98	00:01:25	96	94.79%	95.05%	$0.00
	May 1, 2012 - May 31, 2012	93	90	00:04:20	90	91.11%	91.40%	$0.00
	% Change	8.60%	8.89%	-67.51%	6.67%	4.04%	4.00%	0.00%
7.	/about/management-team/							
	Aug 1, 2012 - Aug 31, 2012	96	84	00:01:44	3	66.67%	36.46%	$0.00
	May 1, 2012 - May 31, 2012	71	61	00:01:27	2	100.00%	33.80%	$0.00
	% Change	35.21%	37.70%	18.93%	50.00%	-33.33%	7.86%	0.00%
8.	/contact/							
	Aug 1, 2012 - Aug 31, 2012	90	75	00:01:43	14	42.86%	48.89%	$0.00
	May 1, 2012 - May 31, 2012	75	67	00:02:29	13	69.23%	44.00%	$0.00
	% Change	20.00%	11.94%	-31.16%	7.69%	-38.10%	11.11%	0.00%
9.	/about/press-releases/job-opening-sales-specialist/							
	Aug 1, 2012 - Aug 31, 2012	81	77	00:02:43	32	71.88%	46.91%	$0.00
	May 1, 2012 - May 31, 2012	2	2	00:00:33	0	0.00%	50.00%	$0.00
	% Change	3,950.00%	3,750.00%	392.95%	∞%	∞%	-6.17%	0.00%
10.	/about/jobs/							
	Aug 1, 2012 - Aug 31, 2012	76	62	00:00:33	7	42.86%	32.89%	$0.00
	May 1, 2012 - May 31, 2012	63	53	00:00:29	3	66.67%	23.81%	$0.00
	% Change	20.63%	16.98%	13.72%	133.33%	-35.71%	38.16%	0.00%

FIGURE 3.8 (Continued)

Beyond Tables

There are other ways to display data besides the humble table, available through the buttons in the upper right corner.

Percentage

You can take a single metric from the table and show not just the actual value (like how many visits came from each kind of mobile device), but also what portion of all visits each row represents, along with a pie chart to visualize the relative amounts of each row. For a more effective way of visually comparing values, you can select the next button, Performance.

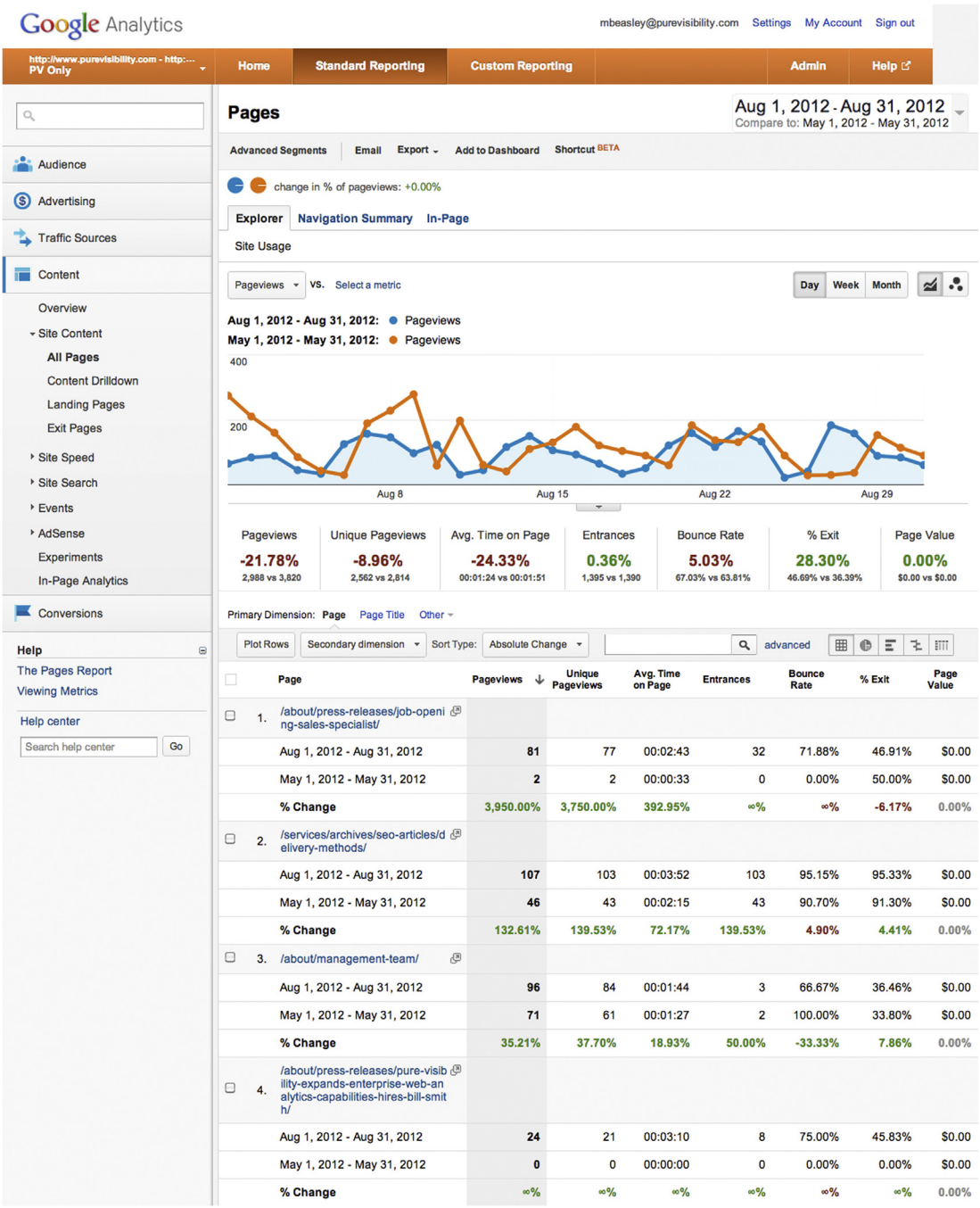
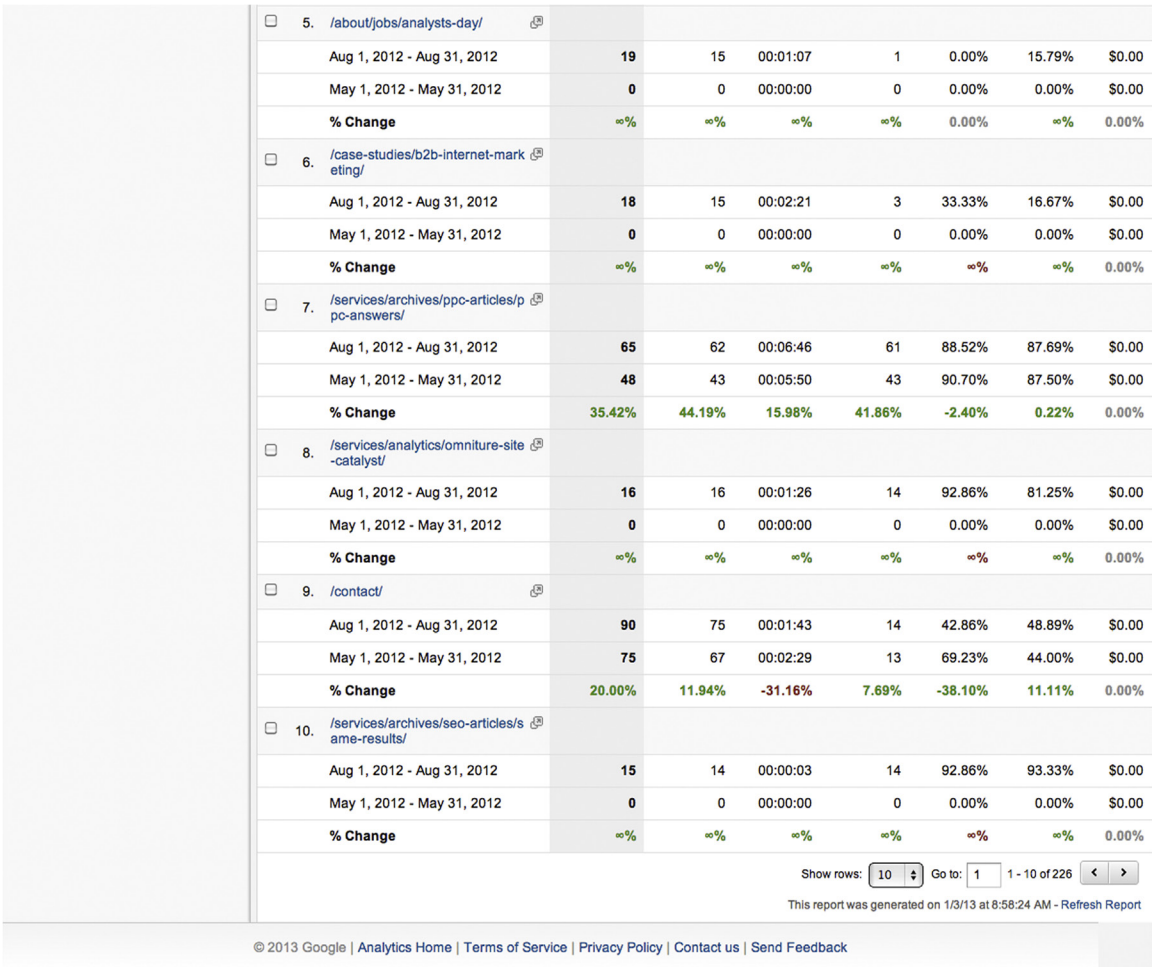


FIGURE 3.9

The data from Figure 3.5, sorted according to the Absolute Change in page views from May to August. This sorting results in the top 10 rows of the table being filled with rarely viewed pages.



5.	/about/jobs/analysts-day/							
	Aug 1, 2012 - Aug 31, 2012	19	15	00:01:07	1	0.00%	15.79%	$0.00
	May 1, 2012 - May 31, 2012	0	0	00:00:00	0	0.00%	0.00%	$0.00
	% Change	∞%	∞%	∞%	∞%	0.00%	∞%	0.00%
6.	/case-studies/b2b-internet-marketing/							
	Aug 1, 2012 - Aug 31, 2012	18	15	00:02:21	3	33.33%	16.67%	$0.00
	May 1, 2012 - May 31, 2012	0	0	00:00:00	0	0.00%	0.00%	$0.00
	% Change	∞%	∞%	∞%	∞%	∞%	∞%	0.00%
7.	/services/archives/ppc-articles/ppc-answers/							
	Aug 1, 2012 - Aug 31, 2012	65	62	00:06:46	61	88.52%	87.69%	$0.00
	May 1, 2012 - May 31, 2012	48	43	00:05:50	43	90.70%	87.50%	$0.00
	% Change	35.42%	44.19%	15.98%	41.86%	-2.40%	0.22%	0.00%
8.	/services/analytics/omniture-site-catalyst/							
	Aug 1, 2012 - Aug 31, 2012	16	16	00:01:26	14	92.86%	81.25%	$0.00
	May 1, 2012 - May 31, 2012	0	0	00:00:00	0	0.00%	0.00%	$0.00
	% Change	∞%	∞%	∞%	∞%	∞%	∞%	0.00%
9.	/contact/							
	Aug 1, 2012 - Aug 31, 2012	90	75	00:01:43	14	42.86%	48.89%	$0.00
	May 1, 2012 - May 31, 2012	75	67	00:02:29	13	69.23%	44.00%	$0.00
	% Change	20.00%	11.94%	-31.16%	7.69%	-38.10%	11.11%	0.00%
10.	/services/archives/seo-articles/same-results/							
	Aug 1, 2012 - Aug 31, 2012	15	14	00:00:03	14	92.86%	93.33%	$0.00
	May 1, 2012 - May 31, 2012	0	0	00:00:00	0	0.00%	0.00%	$0.00
	% Change	∞%	∞%	∞%	∞%	∞%	∞%	0.00%

FIGURE 3.9 (Continued)

Performance

The Performance option is just like Percentage, except instead of a pie chart, it uses a vertical bar graph.

Comparison

The Comparison feature lets you compare a specific metric to the website average, on a row-by-row basis. For the example in Figure 3.11, we see the same data as in Figure 3.5, but comparing the average time on page for each individual page for the average time on page for the website as a whole.

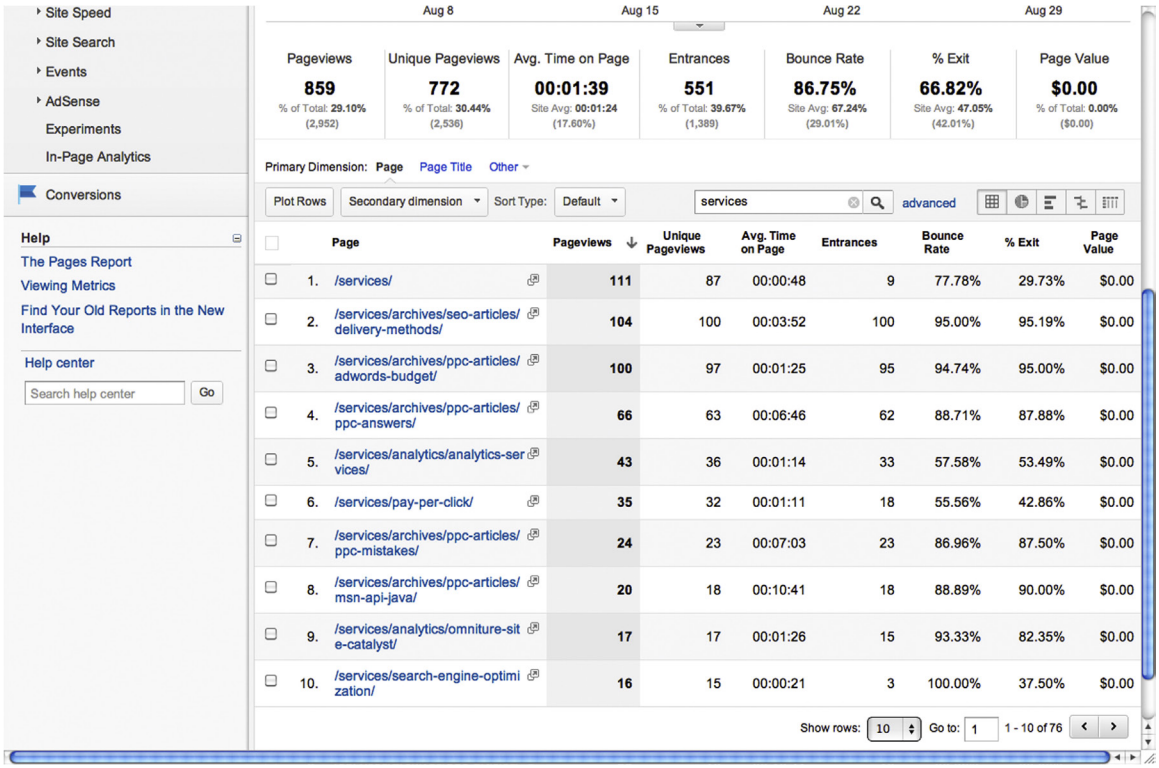


FIGURE 3.10

When you perform a search to filter what data appear in the table, the sums and averages above the table also change. Compare this screenshot to Figure 3.5.

Term Cloud

Only available for certain reports like search keywords, the Term Cloud creates a word cloud. Each row of the table may be a single word, and the size of the word is determined by whatever metric you choose. The usefulness of this report will probably be determined by how useful you find word clouds.

Pivot

Lastly, there is Pivot, a complex and powerful way to interact with data. Pivot lets you show data in a pivot table, which lets you take a dimension and turn it from a row to a column. In the mobile devices example, you would have a row for each kind of mobile device, and then you could select another dimension, such as browser type or whether the user is new or returning, and make that a column. Then, within each column, you could show multiple metrics like pageviews or average time on page. The end result would be the ability to show, for each page, the total number of pageviews and average

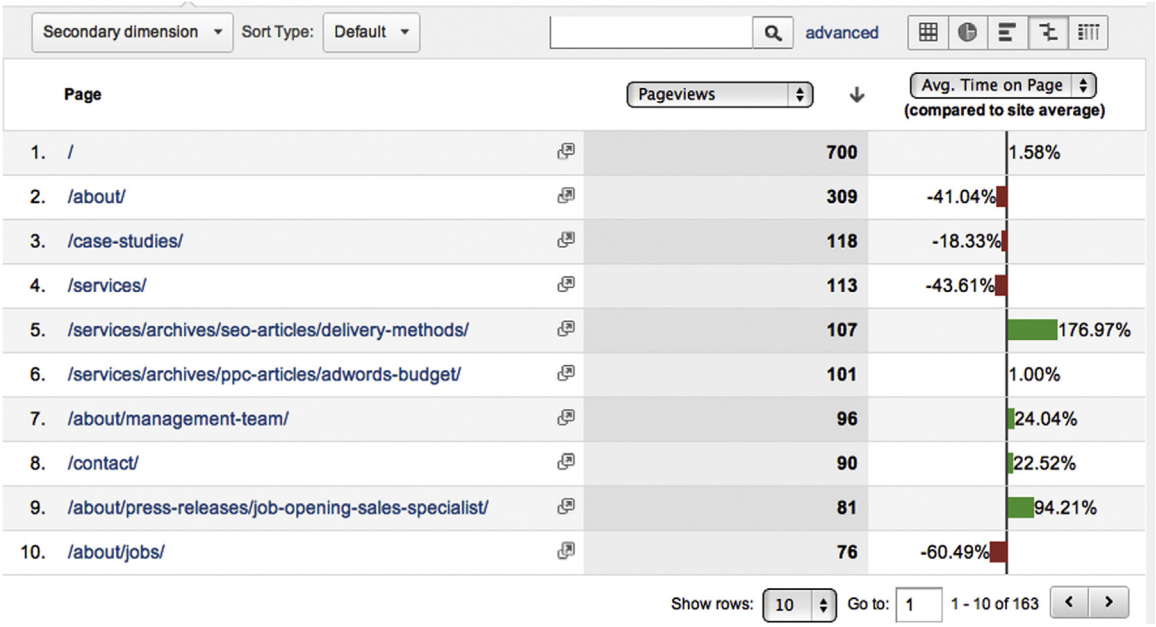


FIGURE 3.11

The data from Figure 3.5, but changed from the Data (Table) view to the Comparison view, displaying a comparison between the average time on page for each page to the overall website average time on page.

time on page, the count of pageviews from new visitors and their average time on page, and the count of pageviews from returning visitors and their average time on page. This feature is complicated and powerful. Although we will not use it in this book, it is important to be aware of it if you use Google Analytics.

KEY TAKEAWAYS

- There are two basic approaches to web analytics: page tagging and log files.
- Page tagging works by adding code to every page of the website that executes when the browser processes the page. For Google Analytics, this code sends:
  - What page was loaded and when
  - Where the user came from
  - IP address
  - Browser and device
- Page tagging tools use a cookie to determine if each pageview is part of the same session, and whether the user has been to the website before.

- Perfect data are impossible because not every user has JavaScript enabled and because sometimes the Internet just doesn't work, but web analytics data from page tagging tools are reliable enough for analyzing trends.
- A metric is an aspect of user behavior that can be measured and expressed as a number, such as how many pages the user views.
- A dimension is an attribute of a user or the user's visit that can be used to categorize the user, such as whether the user is new to the website or returning.