

Beginner's Guide To Web Data Analysis: Ten Steps To Love & Success

Posted By [Avinash Kaushik](#) On November 15, 2010 @ 02:19 In [Web Analytics](#) | [12 Comments](#)



The goal of my recent post on the Yahoo! Web Analytics blog was to pull us up 10,000 feet to do something we do less than 1% of the time in the web analytics world – look at the bigger business picture.

It was called: [Secret To Winning With Web Analytics? Starting Right!](#)

While that was a very strategic post, it got me thinking at a tactical level.

What if I was given the login and password to someone's web analytics data and asked to "find something interesting?" How would I start the process of web data analysis right? Even without any knowledge of the company's goals or help from a stubborn HiPPo or clients who just want data pukes? Can I add any business value?

A real challenge!

It turns out, astonishingly, that even with all those barriers (no objectives or goals or cooperation or business guidance), you can spend a couple hours and do decent enough analysis, sourced from your experience, to deliver some minor data-gasms of insights.

Not quite the real intense ones that you might experience if all the foreplay had been done correctly (see the YWA post above), but still never say no to even minor orgasms right?

Setting The Right Expectations

It is nearly impossible to find earth shattering insights that you can action from your web analytics data in just a couple hours. And yet finding some delightful starting points might be less hard than you might imagine.

Starting points to start valuable analysis from. (What data should I look at first?) Starting points for a customer centric strategy. (What are my customers telling me?) Starting points for gaps in your online marketing efforts. (Where am I wasting money?)

Secret To Winning Web Analytics: 10 Starting Points For A Fabulous Start!

This blog post is a starter guide that outlines the steps I personally undertake most commonly when handed the keys to the data for a website.

I want to share where in your web analytics data you can find valuable starting points, even without any context about the site / business / priorities. Reports to look at, KPIs to evaluate, inferences to make.

I hope you'll benefit from my humble experience. Let's go!

Step #1: Visit the website. Note objectives, customer experience, suckiness.

My biggest beef with web analysts and consultants is how quick they are to jump into Google Analytics or Omniture or WebTrends. It's as if they have never seen a report with Visits & Conversions before. Jeez!

The very first thing I do, and I recommend you do, is visit the website whose data you are analyzing. See how it looks. Go to the product pages. Go to the donation pages. Go to the B2B dancing monkey video (what!). Go to the add to cart page. Go to the RSS / Email sign up page and sign up. Go read some customer reviews (if a ecommerce site) or visitor comments (if a blog). Go download the white papers. Go use site search.

Get a feel for the company's vibe. Get a feel for the information architecture and cross sells and font size and buttons and tab structure and user experience etc. What's hideous? What's awesome?

Bonus points for visiting one competitor's website. Do all of the above.

Take out a note pad and write down your thoughts. What did you like? What did you hate? What frustrated you? What was obviously broken? What's the site trying to do?

At the very minimum your notepad should contain answers to these two questions: What is the macro-conversion? What are two or three micro-conversions? Remember those terms apply to ecommerce and non-ecommerce websites.

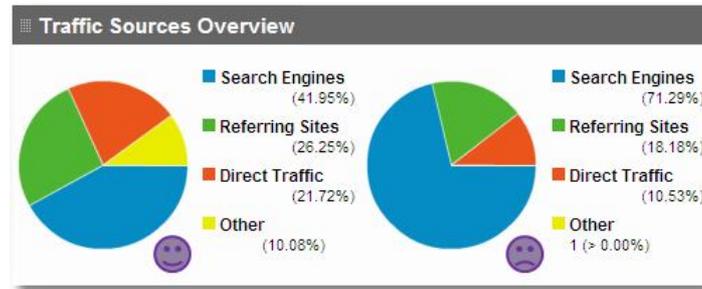
The site owner / client did not help you, but you've not got super valuable context. You're ready for data!

Step #2: How good is the acquisition strategy? Traffic Sources Report.

This is the very first place I end up because the first thing I want to know is how savvy the company is about online marketing. All other site

data comes second because if you stink at online marketing then there is not much of a victory to be had by torturing website data.

No company in the Milky Way has succeeded without having a balanced portfolio of acquisition channels (fancy word for source of traffic). How's yours?



What to look for:

I am really looking for a balanced portfolio of traffic sources. Search, Referring Sites, Direct, Campaigns. Which one is strong? Which one is missing?

Based on my own humble experience the site on the left is what approximates the kind of "best practice" (note the quotes) you are looking for.

Around 40% to 50% Search is normal. If the number is too big (site on the right) it indicates an overexposure to search rankings and algorithm changes (not good at all). If it is too low you are simply leaving money on the table. And of the search traffic, you want a big portion to be Organic so you are not just "renting" traffic or suck at SEO.

20% or so Direct Traffic. If the web analytics tool is implemented right these are all your existing customers or people from offline campaigns. You want a healthy amount of both. If direct traffic is low, I worry if you are any good at customer service / retention (the latter is so often just an afterthought).

20% to 30% Referring Sites. You can't just rely on search engines or spending money on campaigns. A healthy web strategy includes a robust amount of traffic from other sites that link to your products and services, and praise (or slam!) you, or promote you on Twitter and Facebook and forums and otherwise link to you. Free traffic (usually) and you do want that (for many reasons).

10% Campaigns. Google Analytics (sub optimally) calls this Other. It is email campaigns, display / banner ad campaigns, Facebook display campaigns, social media campaigns etc. You want at least 10% of the traffic to be the ones you invite to your site deliberately, after solid analysis and great targeting. Outside of Paid Search. It's a sign of a healthy business that has a diversified customer acquisition strategy.

Consider the above as broad guidelines, again based on my online marketing experience. YMMV, certainly for esoteric types of businesses.

What to do next:

I'll make note of where the company is overleveraged and make a note to dig deeper with the client / HiPPO. Expose the dangers to them, brainstorm how to diversify.

For each bucket I'll look at least the top ten rows. Additionally, for at least one of the four buckets I'll dig deeper by looking at the standard report for that segment to identify some strengths or weaknesses. Surprising keywords, missing sources of traffic, trends in campaign vs. direct visits etc.

At the end of this you'll understand how sophisticated the client is, where you'll attack acquisition first (if you get the time and money to do more analysis).

Step #3: How strongly do Visitors orbit the website? Visitor Loyalty & Recency.

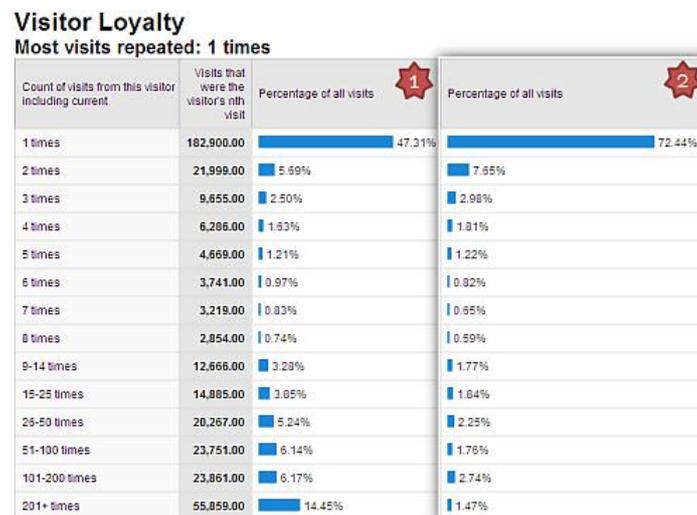
I have a sense for the site and I have a sense for the client's acquisition savvy. Time to focus on the Visitors!!

Most people create sites just for themselves and with no obvious purpose in mind. Furthermore the content publishing schedules, perceptions of "engagement" are all out of whack.

So what I (and you, dear blog reader) want to do is get a sense for how strongly attached are the Visitors to the site. This is of course crucial for any type of content site, but you'll be surprised at how important it is even for an ecommerce website (retention, support, repeat purchases, yada yada yada).

The report I'll look at is the standard Visitor Loyalty report. It would show how many times in a given time period the same person (persistent cookie actually) visits the website. Or, how tightly the Visitor orbits the site!

All tools have this report, Here is how it looks like in Google Analytics:



What to look for:

For site number one it is clear there are a lot of *one night stands* (47%). But notice that bottom, an astonishing 40% of the people visit the site more than 9 times a month! You get a sense for content consumption patterns, you get a sense for your next task (segment the 40, learn what's working there, apply to the 47!), you get a sense for whether the site's delivering on its business objectives.

If the data looks more like site two, cry. Ok, most of the time cry. This site simply engages in *one night stands*, and while I can think of some sites where that can still be the basis of a long term sustainable business model. . . I can't think of a lot of them.

Not a tight orbit. So what? Remember the notes you took in step one? What's the impact of the loyalty pattern on the objectives you noted? With some solid data you are ready to have a discussion with the client / site owner / HiPPO. Take a tissue.

While I love Loyalty the most, I also take a quick peek at Visitor Recency. This is specific to content sites (newspapers, yellow pages, hospital, "I am the next facebook-killer" sites, etc).

Visitor Recency measures the gap between two visits of the same visitor. Or, When was the last time you saw the same person (cookie really). Here's the report:



How amazing is it that 37% of the traffic on the site last visited less than 24 hours ago! Talk about orbit!

Segmenting this data is best (by content, source, campaign, outcomes, etc), but even a cursory review will help you understand how people behave.

What to do next:

I always review two more reports that give me a sense for content consumed. No, no, not the silly reports that show mostly useless metrics like Average Time on Site and Average Pages Per Visitors ([averages stink!](#)).

I am talking about Length of Visit and Depth of Visit:



With Loyalty and Recency we measured visitors visiting, but once they are here what are they doing? That's what you are trying to get a sense for with these two reports. [Remember visits with just one page view, bounces, will be in the first bucket 0-10. For why, see: [How time on site & time on page are computed](#)]

If you have some time segment out the bigger buckets (beyond 0-10) and analyze the data. If you don't have time just knowing Loyalty, Recency, Length, Depth tells you a lot about how tightly Visitors orbit this site, and understanding customers is precious.

Step #4: What can I find that is broken and quickly fixable? Top Landing Pages.

Understand site? Check! Understand traffic sources? Check! Visitors? Check!

Time to take off the gloves and some clothes and get dirty.

Companies spend lots of money acquiring traffic, often badly, so why not find top places where that money is being wasted and which home pages might possibly be stinky? Visitors refuse to give you a single click? That might be a useful signal. :)

In your web analytics tool this is a standard report. It shows bounce rates, sweetly indexed against site average, for the top entry points to the website:



3.	/ffidealist/es/Home/Spla		-42.41%
4.	/ffidealist/es/SiteIndex/A:	43.97%	
5.	/ffidealist/es/SiteIndex/A:	37.92%	
6.	/ffidealist/es/SiteIndex/A:		-80.99%
7.	/ffidealist/es/SiteIndex/A:	21.57%	
8.	/ffidealist/es/AdvancedS	70.50%	
9.	/ffidealist/es/SiteIndex/A:	17.76%	
10.	/ffidealist/es/SiteIndex/A:		-3.01%

What to look for:

The red parts! See why I like "indexed against site average" feature? It is easy to know what smells bad.

Three landing pages (entry points to your site) are performing really well. Seven seem to be bouncing at a much higher rate than normal, some spectacularly so. At this point you don't know why.

When you see that a page has a high [bounce rate](#) it could mean one of two things:

1. Wrong people are coming to your site (highlighting problems with campaigns, SEO, etc) or
2. The page itself is poorly constructed (missing calls to action etc) or otherwise broken.

At this point you don't know which of the two (or both) is the problem. Since you don't have a lot of time pick two of the biggest losers above. Click on the arrow thingy next to their name in the above report and visit them. Sometimes the problem is obvious. Next click on the link itself in the above report and visit the page level report. There in the drop-down pick Entrance Sources and Entrance Keywords. That segmented view will quickly tell you which sources and / or keywords are contributing huge bounces.

Now, at least for two or three pages of the site you are analyzing you know that they stink and you have decent clues of what the cause(s) might be. Give yourself a pat on the back. Great job!

[An exception: Analyzing [bounce rates](#) for a blog, or "bloggish" site? Segment and look at landing pages for New Visitors; for all other sites the method is as above.]

What to do next:

In this case you are in a position to recommend specific fixes. You have looked at the pages and sources of traffic (proxy for customer intent). You can use a heuristic evaluation process to tell the site owner what fixes will help reduce bounce rates.

A clean and handy checklist is here: [Qualitative Analytics: Heuristic Evaluations Rock!](#)

I am telling you people are going to love you for being this awesome.

Step #5: What content makes us most money? \$Index Value Metric.

Most effort on any given website is spent on content creation, and hence for step five I encourage you to stick with page reports, but flip the funnel from an "input metric," bounce rate, to an "output metric," [\\$Index value](#).

For an ecommerce (with revenue) or a non-ecommerce website (where goal values have been defined) \$ Index value essentially computes "how much revenue" has been earned by a page (more like contributed by a page). It is a great way to gauge the value of a page.

Go to any content report, in this case Content By Title and you'll see the last column:

	Pageviews ↓	Unique Pageviews	Avg. Time on Page	Bounce Rate	% Exit	\$ Index
s R...	24,787	21,052	00:08:38	64.87%	65.13%	\$1.86
an...	9,826	8,931	00:20:38	77.92%	80.37%	\$0.88
ue...	2,718	2,496	00:04:14	84.40%	83.00%	\$0.35
y: ...	2,368	2,114	00:15:06	71.24%	74.16%	\$1.24
y ...	2,162	1,893	00:06:40	78.52%	70.63%	\$1.22
:H...	1,922	1,584	00:02:35	78.39%	78.15%	\$0.28
Ve...	1,869	1,434	00:01:05	71.40%	75.49%	\$0.10
ca...	1,792	1,481	00:02:37	50.83%	45.59%	\$8.79
s: ...	1,738	1,448	00:07:42	71.49%	60.99%	\$1.54
Of ...	1,691	1,450	00:09:59	78.66%	81.37%	\$0.38

What to look for:

Quite simply the types of content that add most value to an ultimate outcome for your website. In the above screenshot it is pretty clear that even in the top ten the range of value added is from \$8.79 to \$0.10.

Would your boss / client not die and go to heaven if you told them exactly what types of content they should be writing / pimping more and what less? How about pages of which product / services generate most value?

Or my favorite report to look at: Content by Drilldown.

That report is particularly apt for sites that are organized in clean directory structures (like /products, /videos, /demos, /blog, /whatever else). Now you are able to discern which groups of content is most valuable to the company. Are videos really valuable? How about really heavy painful Silverlight demos? Wish lists work? The answer awaits!

If you don't have clean directory structure you can still use segmentation to group content and do the above.

What to do next:

**What to look for:**

The first thing to check is if you see anything here.

If you don't see anything here, and the company has been around for some time, then you know you are going to struggle, in case this is a consulting gig, to make any decent money off them or, in case this is your first day in your job, you are going to not get a lot of love in this company as an Analyst. I am not saying quit, I am just saying dig in 'cos it is going to be a soul-searing struggle if this report is empty.

On the other hand if [macro and micro conversions](#) are present then get down on your knees and say a silent prayer because this is going to be fun.

Check if the actual goals & conversions are what you had noted in Step 1. If they are not then what are the visitors to the site doing of business value? Anything you noted in Step 1 that is not here (new goals to create?). What do the trends over the last 12 to 16 weeks suggest? What Goals are contributing the most amount of value?

At the end of this little exercise you should be able to confidently speak to your client / boss about how the website is meeting business objectives, and possibly where it is failing. If you did Steps 2 through 6 well then you might even have other actionable recommendations to make.

What to do next:

If there were some goals you had identified, or your client/boss was expecting, then take a moment to configure those in the web analytics tool.

If they do not have any behavioral goals created (99% of the people don't), then create those, takes just a moment. Refer to your insights from Step 3 to set the threshold values.

If this was an ecommerce website I would typically create one segment as a little bonus for the client. Orders where the total value was 50% higher than the average order value. Essentially the "whales" – people who order way more than normal. My hope is to get particularly valuable insights about where these people come from (geographies, campaigns, keywords, etc), what they do on the site (content consumed etc), and what they buy (shopping cart / basket analysis etc).

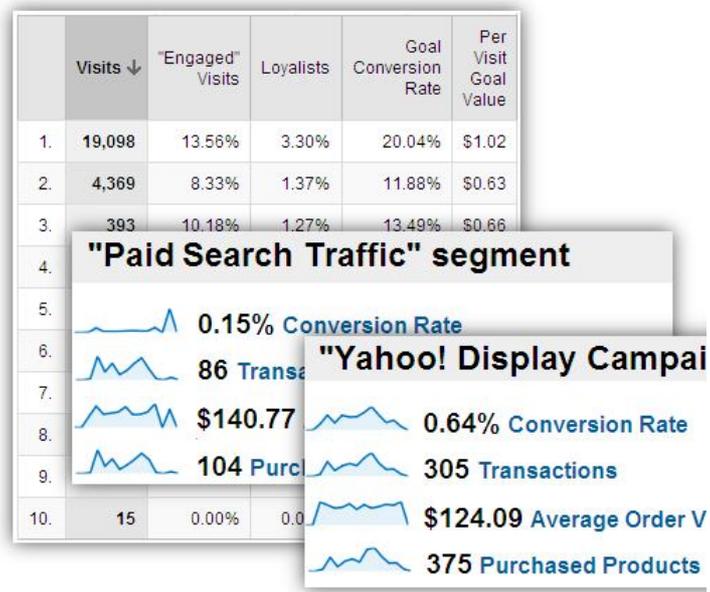
You want a lot more of these people. It is good to understand them really well.

Step #8: Can the Marketing Budget be optimized? Campaign Conversions/Outcomes.

Remember the only three outcomes that are important in web analytics? More Revenue. Reduced Cost. Increased Customer Satisfaction.

In this step I focus on the second item, reducing cost. Perhaps it is surprising as this is our very first foray into the web analytics data, we have received limited love from the client / HiPPO, and we don't have all the company business specific knowledge that might be necessary. Yet we can help reduce cost of marketing / customer acquisition.

My favorite report? Goals / Conversions by Campaign:

**What to look for:**

Campaigns as in Paid Search and Display and Email and Social Media and really anything of value you'd discovered in Step 2.

Start by looking at the horribly named "Other" report in Google Analytics (or perhaps the appropriately named Campaigns report in your web analytics tool). Initially allow yourself to be guided by that column at the end Per Visit Goal Value. It is a measure of efficiency. Notice above you go from \$1.02 to \$63, it is not hard to guess one campaign is working better than the other.

Then work backwards and see what conversion numbers look like. Then work further back and see which individual goals might be causing that high value to be created.

At the end of this exercise you should have some preliminary recommendations for at least one or two places money is potentially being wasted, or at least inefficiently spent. Killing opportunities (example: More better email campaign, less crappy Facebook pages!). You should also have some sense for where improvement opportunities might exist (I had a bunch above where PVGV was 40 cents).

What to do next:

Pick one or two major campaign strategies the company is executing and dive deep into ecommerce analysis (if the client is ecommerce). The two screenshots of Paid Search and Yahoo! Display campaigns gives you just a small hint about how much opportunity exists below the surface to dig and understand the deltas between conversion rates and the average order values and the trends for those key metrics.

things that even they, the most data driven of data driven companies, might be unaware of in their own data.

What to do next:

The hard part with Intelligence (custom or automatic alerts) is to isolate the root cause. Look at the newly released [GA Intelligence Major Contributors](#) section. That has the clues about root cause. Leverage the advanced segmentation feature to isolate the activity causing source / behavior / outcome and dig deeper.

[For more checkout the videos on [Google Analytics Intelligence.](#)]

And you are done! Does that not feel awesome? And more importantly, doable?

Summary: The Beginner's Guide / Tips / Best Practices For Web Data Analysis

In case you needed a handy checklist, here's what we've learned today:

- Step #1:** Visit the website. Note objectives, customer experience, suckiness.
- Step #2:** How good is the acquisition strategy? Traffic Sources Report.
- Step #3:** How strongly do Visitors orbit the website? Visitor Loyalty & Recency.
- Step #4:** What can I find that is broken and quickly fixable? Top Landing Pages.
- Step #5:** What content makes us most money? \$Index Value Metric.
- Step #6:** How Sophisticated Is Their Search Strategy? Keyword Tag Clouds.
- Step #7:** Are they making money or making noise? Goals & Goal Values.
- Step #8:** Can the Marketing Budget be optimized? Campaign Conversions/Outcomes.
- Step #9:** Are we helping the already convinced buyers? Funnel Visualization.
- Step #10:** What are the unknown unknowns I am blind to? Analytics Intelligence.

The first time you go through the steps outlined in this guide it might take you more than 120 minutes. But I promise you that with time and experience you'll get better.

I wish just reading this blog post (it probably took you 120 minutes just to read it!) would be enough. It is not. You'll have to go practice it on many many clients. The more you do it the better you'll get as your sense of direction, data, discovery and deduction get better and better and better.

It is optimal to start any web analysis with a clearly defined [web analytics measurement model](#). But if you don't have one then you no longer have an excuse not to provide something small that is incredible and of value from any web analytics tool you have access to, for any website in the world. And know that I am rooting for you!

Ok, your turn now.

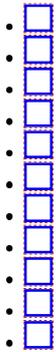
When you are thrown into a website's data blind what are the first few things you do? What reports and metrics do you attack first? Over time have you discovered any strategies that work across multiple clients? Do you agree with the order of the steps above? Would you do something differently?

Please share your thoughts / critique / best practices / tips via comments.

Thanks.

Social Bookmarks:

Share 15



12 Comments To "Beginner's Guide To Web Data Analysis: Ten Steps To Love & Success"

#1 Comment By [Hubert](#) On November 15, 2010 @ 02:47

What can I say to such a post ?

Great! Instructive and gives a real insight for web analytics beginners !

#2 Comment By [Niklas](#) On November 15, 2010 @ 04:37

Nice, covers the whole analytical chain for web marketers.

Great post,

Niklas

#3 Comment By [TrafficColeman](#) On November 15, 2010 @ 05:17

When I perform analysis on client website, I try to go to the website first and get a fill for what they are trying to convey to their audience.

That helps me understand what they are doing and how I can help them even more.

TrafficColeman "Signing Off"

#4 Comment By [Brian Chiou](#) On November 15, 2010 @ 06:49

Thanks for bringing up the Wordle example again. I have yet to use that tool. I will now though!

I do as you suggest in this post and get myself oriented with the clients website. One mini step I'd like to add is doing a bounce rate against site average during step #2. Or, add that to step #4. In some cases, I'll find that google / cpc is performing really poorly and I'll be able to dig deeper by individual campaigns / groups / ads that may be mistakenly triggered by bad "broad keywords".

Insights aside, is it our responsibility as the analytics dude to make sure their Analytics software is installed properly? And if so, how far do we go with that analysis? For example, do we recommend Event tracking / Custom Variables at this stage of the relationship?

#5 Comment By [Avinash Kaushik](#) On November 15, 2010 @ 07:28

Brian: It is likely that some of the cooler features (even tracking, custom variables etc) are not implemented, and also likely that at least some of the analytics code is not complete (some pages might be missing tags).

As Web Analysts I have noticed that we consistently tend to make our first engagement with the client to be about impressing them with our technical competence and hence tags and code and other such things. I wanted to make it about finding insights, even if they are not supremely awesome or even 5% incomplete because of implementation issues.

If we can show we are not just techno geeks and there to deliver solid business value, even after the first two hours, I think the rest can follow. They'll be happier paying us our going rate and happy to put the time and energy required to fix things and implementing advanced features.

It will be our responsibility to ensure the technical implementation is complete and accurate, but after:

1. We have delivered some business value and
2. Impressed them enough with the first part that they'll work with us to create a web analytics measurement model which will guide advanced features roll out (so that we are only doing things the client will value and not what we feel are important) and then finally
3. Hello code here I come!

-Avinash.

#6 Comment By [Norbert Deli](#) On November 15, 2010 @ 07:36

This is the best post from you so far.

Great summary to love & success. Thx!

#7 Comment By [Philip Chiappini](#) On November 15, 2010 @ 10:44

Great post! I am so glad that you included Step #6: Keyword Tag Clouds in this. It is something that I have never really thought of doing. I can really see how doing this could bring some great insights and be a good way to demonstrate to the higher-ups what is really happening (or not happening).

Thanks!

#8 Comment By [Joe Teixeira](#) On November 15, 2010 @ 11:03

Forget beginner's – I think that everyone in the WA industry should read this post. It's just a solid outline for success that even the most advanced / senior WA people in the industry should read to shake the rust and dust off – and receive a clean, fresh perspective.

And – how can one not like a post that includes Weighted Sort in Google Analytics – easily one of the most unheralded new features out there?

Have you ever thought about becoming an author? :)

#9 Comment By [Prashant Kumar Pracheta](#) On November 15, 2010 @ 11:06

Hi Avinash,

It is always great to read your post and explore more about Web Analytics. You have covered almost everything within this post related to Web Data Analysis.

Quick Question : In Visitor section, I think we need to look out for Member and Non-member visits (Authenticated Users). This data can provide us how many authenticated visitors we have and what they do after login in the website.

What are your thoughts on it?

#10 Comment By [Eelke](#) On November 15, 2010 @ 14:03

I have one problem with your blog articles Avinash, for example this one took me about 2 hours to read it (as im not english native), catch "big picture of article" and write notes in Evernote for later use. After this my head is going to misfire, its painful, pffff! :) But from regular sport activities i know this kind of "pain" is welcomed, so thank you very much! :-)

#11 Comment By [Dalmiro](#) On November 15, 2010 @ 16:03

Thank you for all this information, this will help me to improve my knowledge and make real my dreams.

Greetings from Perú.

#12 Comment By [Mark](#) On November 15, 2010 @ 17:37

G'day Avinash,

I was a little surprised you lumped email in with campaigns in Step 2, and that traffic from total campaigns was so low at 10%.

Personally I treat email as a separate traffic source. Usually it accounts for at least a third if not more of a site's traffic and revenue.

Like you the Traffic Sources report has been my goto report for many years.

What I've found from doing long term yearly analysis is that email, search and direct traffic all even out over the course of a year and bring in similar amounts of revenue.

So if I saw a potential client's report that showed only 10% from campaigns including email I'd know that their email marketing was either non existent or totally sucked.

Of course that would be an immediate low hanging fruit area where we could make significant improvement.

Article printed from Occam's Razor by Avinash Kaushik: <http://www.kaushik.net/avinash>

URL to article: <http://www.kaushik.net/avinash/2010/11/beginners-guide-web-data-analysis-ten-steps-tips-best-practices.html>

Copyright © 2006 - 2009 Occam's Razor by Avinash Kaushik. All rights reserved.