James Bradley

Promoting Your Web Site

Task One

The success of a commercial website depends on potential visitors having knowledge of the existence of the site. This can be achieved by using a number of techniques, such as use of search engines, direct mailing or URLs information and advertising. In order to gain more visitors, the site needs to be geared towards the correct audience. In order to do this you need to be aware, who the sites largest audience is, for example what browser they use, previous site they have visiting etc. This can be found out by looking at the web servers log file.

Typical Data Stored by a Web Server

All Web Servers will keep logs in there system about when and how they have been accessed, this is known simply as a 'log file'. This 'Log File' gives us a wealth of information about each and every user that has accessed the web server, or to put it another way, has visited the web site. This log file can become very large and can end up being bigger in size than the actual web page files.

A typical log file entry would look similar to this:

```
195.92.168.166- [11/Aug/2002:09:21:42 -0500] "GET
/sheep.htm HTTP/1.1" 200 11631 "http://news.bbc.co.uk/"
"Mozilla/4.0 (compatible; MSIE 5.5; Windows NT 5.0)"
```

This, upon first glance, may just look like a mess of data; however, the above lines can in fact give us valuable information about the visitors to the site that will enable us to adapt the site towards them, making viewing more interesting and increased chances of return and frequent visitors.

To make it a log file entry easier to understand, it can be broken down, section by section. Each section will give us a number of pieces of information about each and every visitor to the site

Below, the log file entry has been broken down, describing what each section of it actually means.

```
195.92.168.166
```

This is the IP address of the visitor to the site. In some log files this IP address may have been resolved to a domain name and may appear as, for example: `webcacheb09a.cache.pol.co.uk`

```
11/Aug/2002:09:21:42
```

This is the data and time the entry was made into the log

```
-0500
```

This is the time difference to Greenwich Mean Time (Universal Time). The example log file shown above was created when the web server was on US Central Summer time.

```
GET
```

This is the action requested by the user, in this example, get is the command; this simply means the user requested to view the web page. Another common command is ‘PUT’ this is used to upload the files onto the server.

```
sheep.htm HTTP/1.1
```

This is the ‘object’ the user is requesting. In this example the user is requesting to see the page ‘sheep.htm’

```
200
```

This is the result code, below is simple list of what the codes mean;
200 Series – Success
300 Series – Redirect
400 Series – Failure
500 Series – Sever Error

```
1163
```

This is the size of the object the user requested, in bytes.

```
http://news.bbc.co.uk/
```

This is the referring URL, i.e. the page the user was on before they loaded up the page that is logged in this log file.

This section can be used to find out if users have come from search engines and what their search criteria were. If the referring URL appears something like

```
http://www.google.com/search?q=%27interesting+stuff%27&hl
=en&safe=off&start=10&sa=N
```

We can see from this that the user has got to your web site through the search engine google, using the search term ‘Interesting Stuff’

```
"Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"
```

This is the browser version and platform of the user who requested the page off the web site.

In this example, the user was using Microsoft Internet Explorer 6 and the Microsoft Windows XP Operating System.

Making Life Easier …

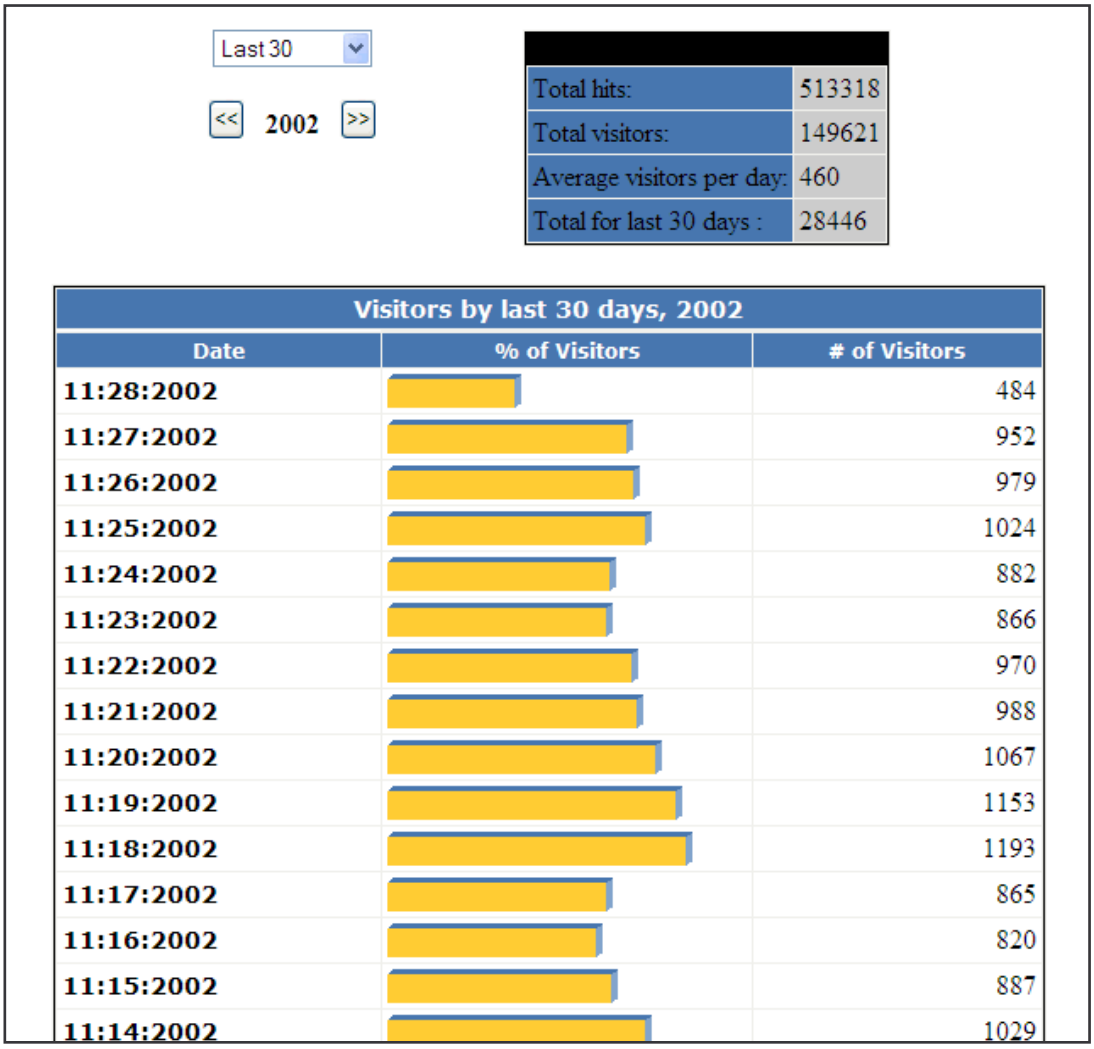
Manually looking through a sites log files would obviously be a very laborious and time consuming task. For this reason products have been produced to take these log files, and turn them into easily read, graphs and statistics for the end user. This enables the team managing the website to easily gain information from this mass of data, that previously just looked a mess, and improve and aim there website more effectively, increasing hits and general success of the site.

Below is outlined just one of the services available for site statistics on the Internet. This service would require a small section of code on each of your pages within the site in order to gather the statistics for analysis.

The log file can tell you if search engine optimization and other marketing campaigns are working. They will show you exactly in what areas you have success and where you need to put in more work. So, if you run a site but haven't seen your server log files, get a hold of them right away. They will teach you things you'd never expect about how to make your site better, not just for you, but also for the people who matter - your customers.

It should be noted that all of the information below is gathered through the web servers log file but is simply displayed in a more easily read and user friendly manner, making analysis and management easier than reading raw log files. The data below however may come from slightly more advanced and detailed log files than the one outlined in the example above.

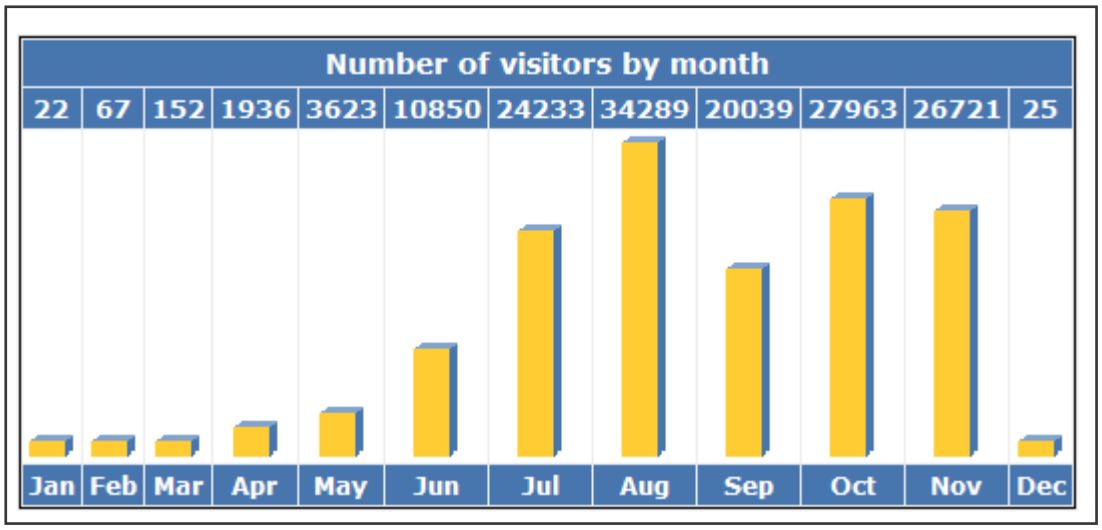
The images shown below are examples of tables and graphical representations, such as pie charts, of this log file information.



Total Visitors

This gives information about the total number of visitors to the site over the requested time period and at specific dates.

This information is useful to find out at a glance how successful the site is, although if the team wish to target site aspects to an audience, they should refer more to the per hour day or month statistics to get a better representation.

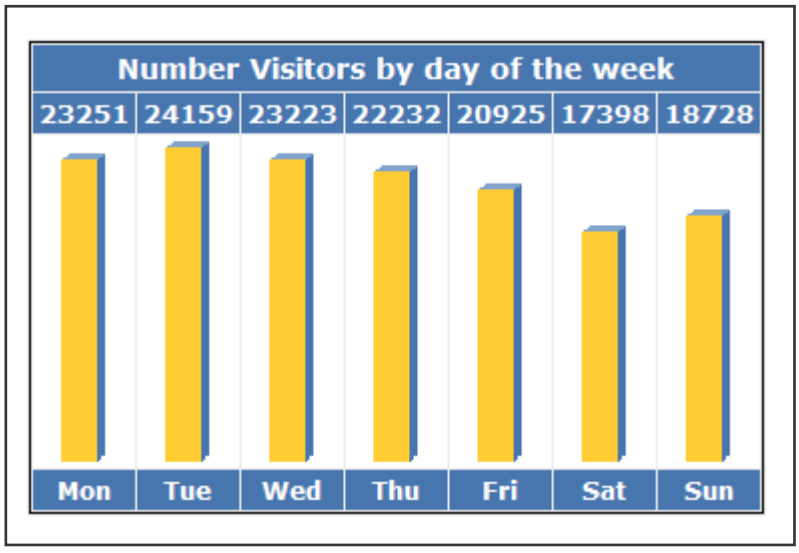


Visitors over the Last Month

This graph gives information about the total number of visitors to the site over the last month. This data can be viewed for each moth of the year.

This data is useful to the web development and management team. This is because the team may have released an advertising campaign about the site through flyers and the local press for example. This set of statistics enabled them to view if this advertising campaign has had any effect on the ‘hits’ to the site.

Also these statistics gives valuable information to the development team so they know when to release new items onto the site, as they are aware what sort of time of year most people are visiting the site.



Visitors per Day

These set of statistics display the amount of visitors per day the site is receiving.

This gives valuable information to the development team in the same sort of way as the per month statistics do, just slightly more accurately.

Also this will enable the team to decide a time, if necessary to take the site off-line to make updates as these set of statistics will give them the required information about when the least amount of people will be viewing the site.

Visitors per Hour

These statistics will give the team similar, but more detailed information as to that for the per hour set.

These statistics will enable the team to, accurately, decide on a suitable time, if necessary to take the site off line to make updates to it.

This will also give the team the information they need to successfully advertise items on the web site, if needs be, at a time when the most people will be viewing the site.

Exit Pages

This statistic reports the top ten exit pages of users on the web site. This gives information about the page the user was on when they exited the web site.

This can give valuable information to the web site managers. By looking at the top exit pages, the managers of the web site can see which pages of the site are least popular and so they can improve upon these particular pages.

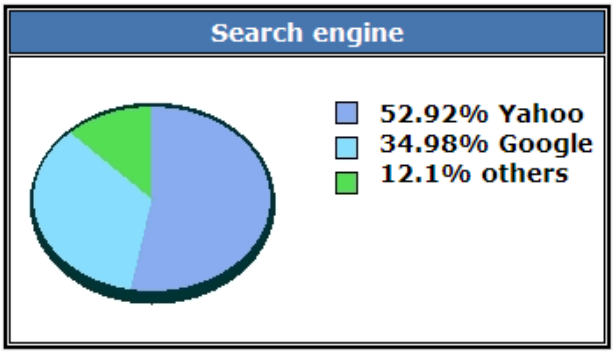
Referrer		
Referrer	% of Visitors	Visitors
A web site URL	13.04	18142
http://www.free-hit-counters.com/	10.73	14927
http://search.yahoo.com/bin/search	7.28	10126
http://search.yahoo.com/search	2.75	3820
http://www.free-hit-counters.com/template.php	1.26	1750
http://www.free-hit-counters.com/user.php	0.77	1073
http://www.free-hit-counters.com	0.69	959
http://www.free-hit-counters.com/sign_up.php	0.66	917
http://www.brooksoldeenglishbulldogge.com/	0.5	693
http://www.free-hit-counters.com/signup.php	0.5	689
http://aolsearch.aol.com/dirsearch.adp	0.43	602
http://google.yahoo.com/bin/query	0.42	581
http://www.free-hit-counters.com/index.php	0.38	522
http://www.free-hit-counters.com/login.php	0.31	433
http://members.lycos.co.uk/whitetriangle/indexBR.html	0.31	428
http://dir.yahoo.com/Computers_and_Internet/Internet/World_Wide_Web/Programming/Access_Counters/	0.31	427
http://www.eldoraspeedway.com/hbo.html	0.3	415
http://uk.search.yahoo.com/search/ukie	0.29	397
http://www.staticfree.net/klaus/	0.28	388
http://members.cox.net/fullwoodsk8er/right.html	0.27	372
http://www.google.com/search	0.26	365
http://www.purx.net/	0.25	351
http://www.liquidcrystalsite.com/	0.25	343
http://www.angelfire.com/super2/dupiytoot/	0.22	304
http://committeeforucl.members.beeb.net/takeover.htm	0.21	294
http://free-hit-counters.com/	0.21	289
http://my.dreamwiz.com/razzu/left.htm	0.2	280
http://search.msn.com/results.asp	0.19	264
http://www.cuccaro.net/piddiekorn/home.html	0.18	255
http://home.mchsi.com/~pconner/midwestmarks/	0.18	247

Top Referrers

This statistical data shows where the user was previously before they came onto your site.

This gives the managers of tie site valuable information about what the viewer to the site is interested in. For example if the referring site was football.com the viewer is obviously interested in football. If a high proportion of the visitors to the site are coming from similar web pages, for example, all related to a certain subject, such as football, the site managers might decide from this, to incorporate something to do with that subject into the site.

Also if many visitors are coming from the same site, the web site managers may decide to set up a partnership of sorts with this other web site or become ‘affiliates’ in order to increase business.



Search engine		
Search engine	% of Referals	Referals
Yahoo		16295
Google		10772
Lycos		1355
AOL		696
AltaVista		426
MSN		408
Dogpile		124
Netscape		122
Mamma		121
Excite		95
DMOZ		84
AllTheWeb		79
WebSearch		56
Ask		43
MetaCrawler (Metamoteur)		42
DirectHit		28
Hotbot		14
Ask Jeeves		13
WebCrawler		13
Euroseek		3
MegaSpider		1

Top Search Engines used to find Site

This set of statistics shows the management team, which search engines, most of the users are finding the site through.

This will give valuable information to the management team so they can look at this data, find out the search engines that the site has been submitted to but not many hits are coming from, and can investigate why no one is coming

to the site through a particular search engine. From this data the managers of the site may decide to try and get the site higher is some particular search engines that appeared low on this graph.

Search query		
Keyword	% of Visitors	Visitors
hit counters	13.37	3724
free hit counters	12.39	3451
hit counter	10.36	2886
free counter	9.85	2744
free counters	9.48	2640
free hit counter	8.61	2399
counters	3.17	884
free web counters	1.83	510
free web counter	1.65	459

Top Search Queries used to find Site

This set of statistics shows the management team, the top search queries used to find the site through internet search engines.

This will give the team information about what the user searched for in order to get to the site, and from this data may decide to alter some of the sites META tags in order for the search strings to bring up more relevant data and the site to be nearer the top of the search engine.

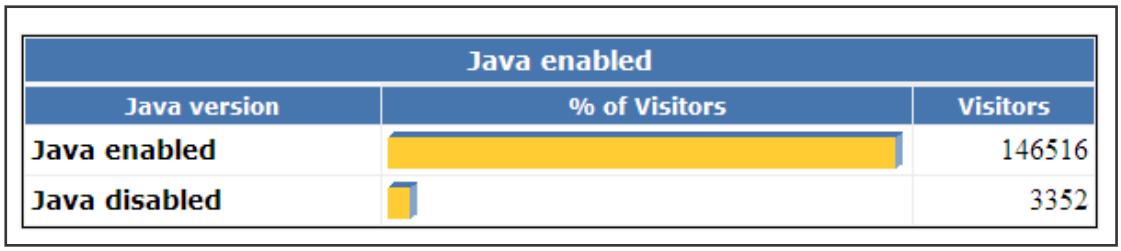
JavaScript version		
JavaScript version	% of Visitors	Visitors
Unknown		80516
JavaScript 1.3		59769
JavaScript 1.2		6556
JavaScript 1.6		2882
JavaScript 1.0		119
JavaScript 1.1		6

Version of JavaScript

This will give information about what version of JavaScript the visitor to the site is using.

This gives valuable information to the web development team as it will enable them to decide if to include any modern JavaScript elements into the site. For

example, if the statistics show us that the majority of the users coming to the site have the latest JavaScript version, and have it enabled, then it would be ideal and feasible for the team to implement some JavaScript elements into the web page. However, if the statistics show us that most of the visitors to the site in fact do not have the latest version of JavaScript or have it disabled, then it would not be a good idea to implement any JavaScript elements into the site. This is because this would mean that the majority of the visitors would not be able to view all elements of the web page or may receive errors and so thus decreasing their opinion of the web site.

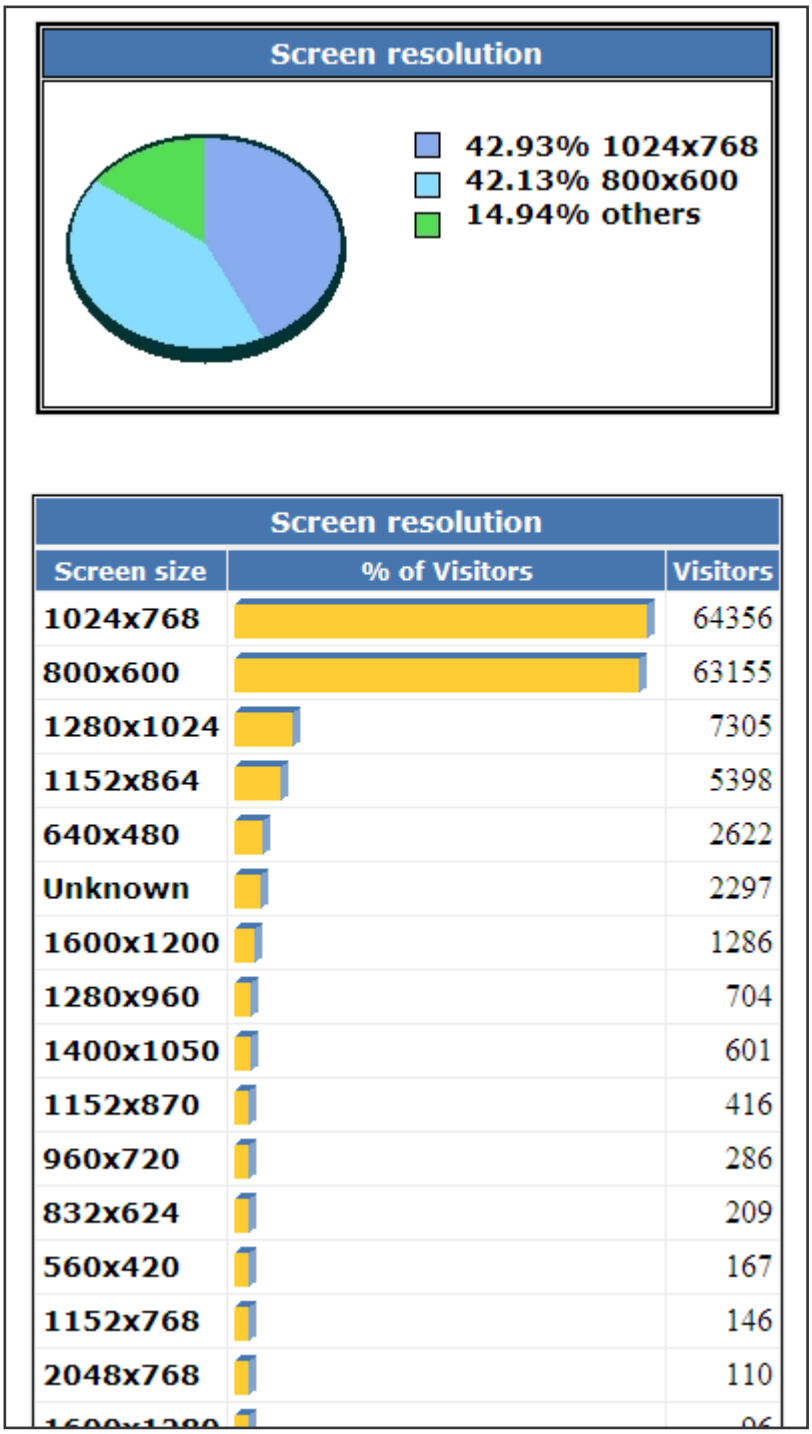


Java enabled		
Java version	% of Visitors	Visitors
Java enabled		146516
Java disabled		3352

Java

This graph shows how many visitors to the web page did had Java enabled on their machine and how many didn't have Java enabled.

This is useful to the web development and management team so they can decide if it is viable or appropriate to include any Java elements into the web site. For example if a high percentage of the visitors to the site do have Java enabled then it would be feasible and ok to add some Java elements into the site. However, if this graph shows that most of the visitors to the site do not have Java enabled, then this has given the team valuable information to not include the elements as users would not be able to see them, decreasing there overall opinion of the site.



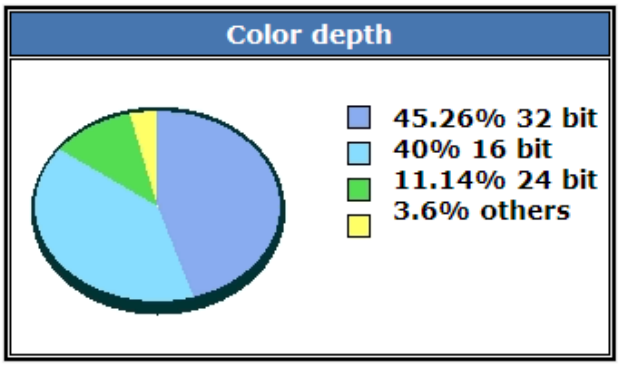
Screen resolution

Screen size	% of Visitors	Visitors
1024x768		64356
800x600		63155
1280x1024		7305
1152x864		5398
640x480		2622
Unknown		2297
1600x1200		1286
1280x960		704
1400x1050		601
1152x870		416
960x720		286
832x624		209
560x420		167
1152x768		146
2048x768		110

Screen Resolution

This graph shows the screen resolution used by the visitors to the site.

This gives information to the web development and management team as to the most poplar used resolution by visitors to the site. This will enable the team to design the site for the most popular resolution or most poplar two resolutions, thus making the site look at its best on as many as possible of the visitor's machines.

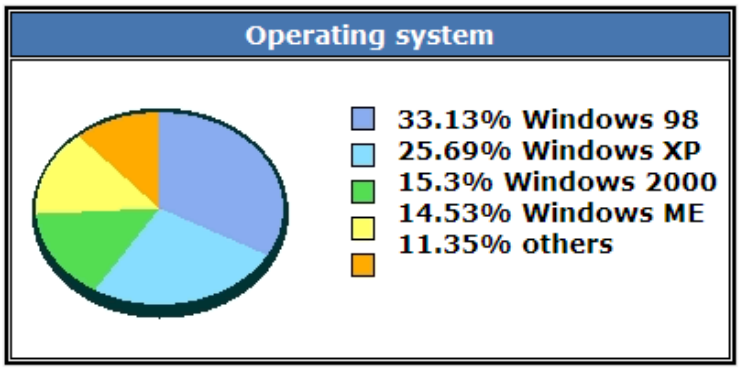


Color depth		
Color depth	% of Visitors	Visitors
32 bit		67835
16 bit		59957
24 bit		16703
8 bit		2779
Unknown		2377
18 bit		108
1 bit		103
4 bit		26
-1 bit		3
undefined bit		1
12 bit		1
15 bit		1

Colour Depth

This graph shows the colour depth used by the different visitors to the site.

This information is valuable to the web development team so they can decide on the colour scheme of the site that will look the best on the most amounts of visitor's screens. The team may decide from this information to only use web safe colours so they can guarantee the site will look the same, colour wise on all systems.



Operating system		
Operating system	% of Visitors	Visitors
Windows 98		49656
Windows XP		38512
Windows 2000		22940
Windows ME		21781
Windows NT		6642
Mac Power PC		4083
Windows 95		3145
Unknown		2459
Windows 32		328
Linux		280
SunOS 5		48
Windows 3.1		7
Sun		5
Irix		3
Free BSD		3
SCO		2
Mac 68000		1

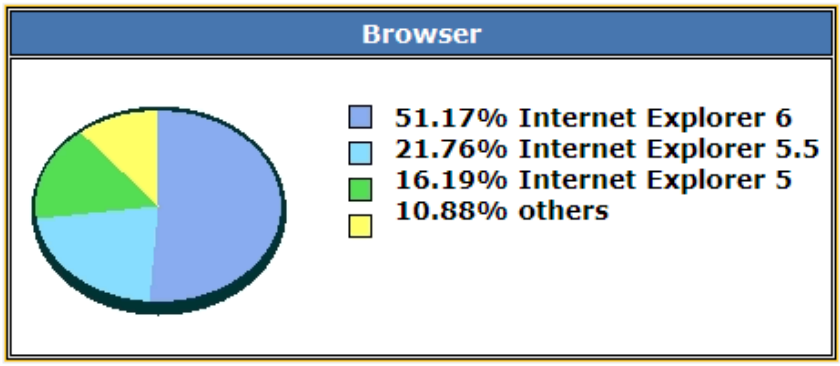
Operating System

This graph gives information about the operating systems used by the visitors to the site.

This gives the management team information about the types of users they are getting and the capabilities of the operating system which they are using. For example if they are receiving many ‘hits’ from NT based operating systems, they can see that they are attracting businesses or ‘techy’ people. Whereas if they are receiving more hits from the more common OS’s such as

Windows 98, from this they can tell they are getting most hits from the home user.

This also gives information about the capabilities of the user. For example the team will know that if the user is using Windows 98 it is possible for them to potentially view Java applications as this was build into the build of Internet Explorer built into Windows 98



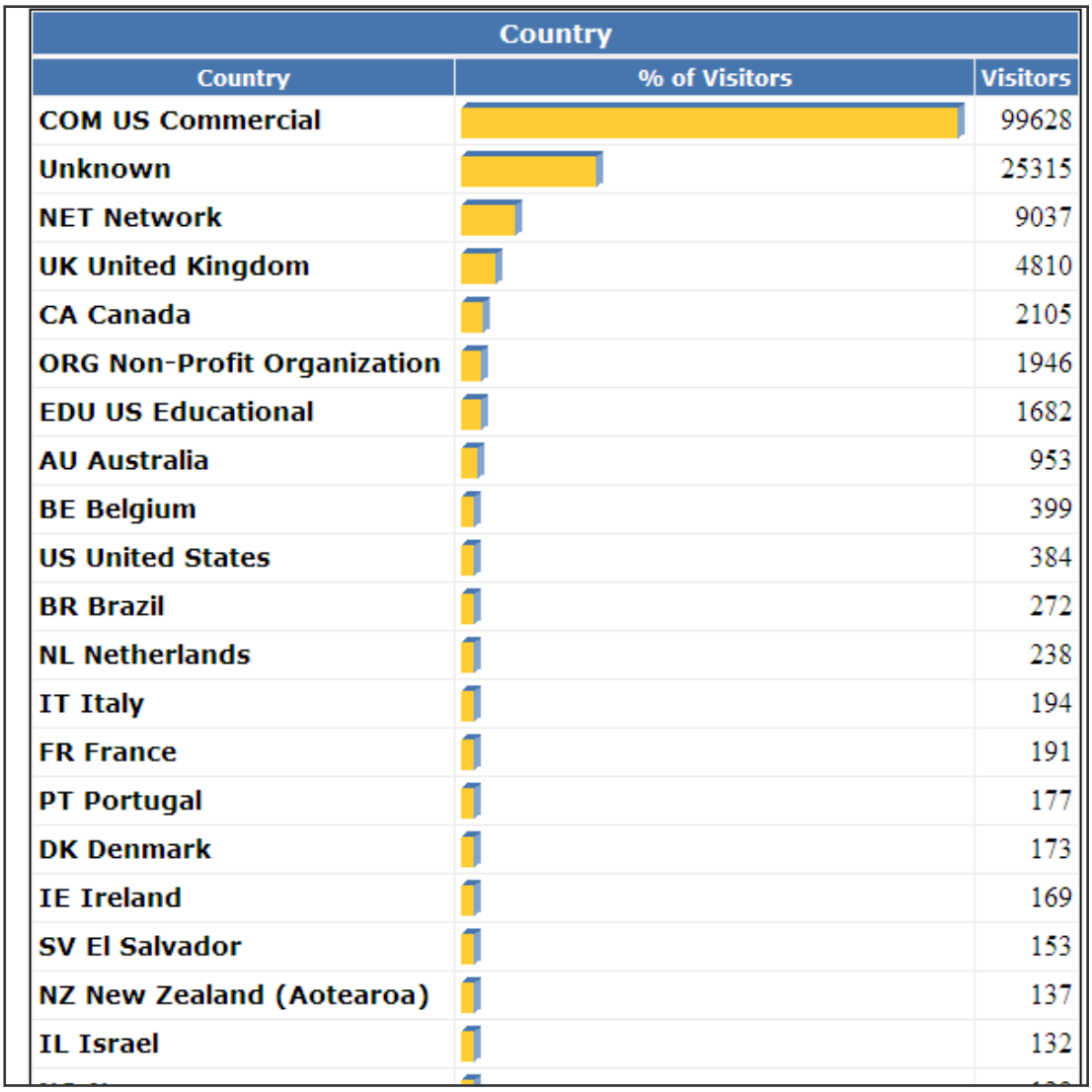
Browser		
Browser	% of Visitors	Visitors
Internet Explorer 6		76697
Internet Explorer 5.5		32610
Internet Explorer 5		24266
Internet Explorer 4		6534
Unknown		3831
Netscape Navigator 6		2882
Netscape Navigator 4		2170
Opera 5up		636
Internet Explorer 3		119
Internet Explorer 5.5up		85
Opera 5		33
AOL 5		17
Netscape Navigator 3		6
Netscape Navigator		2

Browser

This graph gives information about the browsers that the visitors to the site use. Stating the most poplar browser.

This gives the management team vital information about the capabilities of the user and what sort of HTML or mote interactive elements they are physically able to see.

For example, earlier versions of Netscape navigator were unable to fully support Java applets and so if the team notice any hits are coming from users of navigator they should perhaps re think some of their coding and fully test if the site works under Netscape Navigator. This goes for other browsers as well such as Internet Explorer, MoZilla and Opera, if the team notice many hits are coming from a certain browser, the site should be fully tested under this browser to fully ensure that visitors get to view the sites full potential.

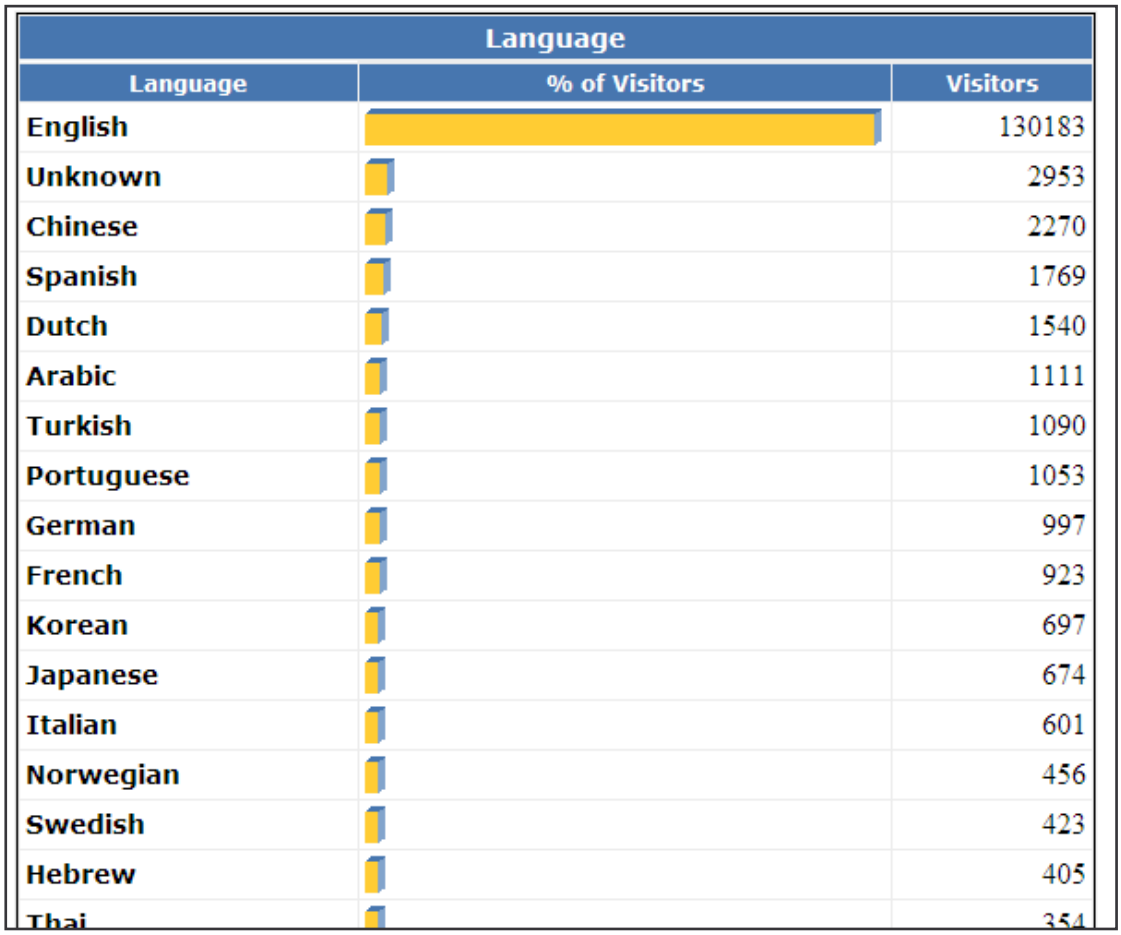


Country		
Country	% of Visitors	Visitors
COM US Commercial		99628
Unknown		25315
NET Network		9037
UK United Kingdom		4810
CA Canada		2105
ORG Non-Profit Organization		1946
EDU US Educational		1682
AU Australia		953
BE Belgium		399
US United States		384
BR Brazil		272
NL Netherlands		238
IT Italy		194
FR France		191
PT Portugal		177
DK Denmark		173
IE Ireland		169
SV El Salvador		153
NZ New Zealand (Aotearoa)		137
IL Israel		132

Country

This graph gives information about where the visitors to the site come from.

These statistics are vital to the team so they can see where the majority of the users are coming from around the world. This information is useful especially to companies who are offering shopping facilities, if they are noticing many visitors coming from a foreign country; they may wish to consider offering some special shipping deals to that country in order to attract custom from the visitors to the site from that country.



Language		
Language	% of Visitors	Visitors
English		130183
Unknown		2953
Chinese		2270
Spanish		1769
Dutch		1540
Arabic		1111
Turkish		1090
Portuguese		1053
German		997
French		923
Korean		697
Japanese		674
Italian		601
Norwegian		456
Swedish		423
Hebrew		405
Thai		354

Language

This graph gives information about the language pack the visitor to the site is using.

This information is vital for language reasons and if the team notice many visitors are accessing the site from a foreign country, the team may consider adding some translating links or features to ensure the visitors are getting the most out of the site.

Errors

These statistics show the errors encountered by visitors to the site. If the team notice many of the same errors, for example 404 (not found) errors all for the same page, they would want to investigate this as a search engine or partner site may have incorrect or out of date links for the site.

Basically this records all the errors encountered by visitors to the site, so that the design and development team are available to ‘iron out’ any page errors or possible server errors, so as to not have the same errors in the future.

Summing Up ….

Basically the point of analysing the log file produced by the web server is to maintain and especially to approve upon the web site. The file can provide valuable information to the sites managers that would otherwise be impossible to gain. The log file enables the team to learn more about their visitors so that they are able to gear the site more towards their audience and what they are interested in, thus making the time at the site more enjoyable for the users and thus increasing their potential to return and thus increasing the popularity and hits to the site. These are the conclusions that can be drawn up about the visitors to the site from the analysis above of the log files.