

Microsoft Excel for SEOs

Lessons for Aspiring Ninjas

Introduction

SEOs have been getting into our industry from all sorts of past careers — web designers, developers, marketers, business people and those that “just fell into SEO”. Some of these past positions may have required data analysis with Microsoft Excel, but a good many of them did not. Excel was not a big part of my past jobs, and I would guess that many SEOs’ past careers did not require anything more than adequacy in the program.

Over the last few years our field has become even more data-driven than in the past thanks to tools like Linkscape, improved Google Webmaster Tools and Analytics, Majestic SEO, Raven, and many others. Additionally, our clients have become wiser and more SEO capable, having been burned in the past by snake oil SEOs. They want reliable SEO advice with a strong verifiable foundation, and who could blame them?

Many SEOs are now finding themselves faced with the task of doing fairly complex data analysis to improve their search strategies, and Excel *adequacy* is not quite enough. This was the position I found myself in not too long ago, and while I’m far from the deadliest ninja in the dojo today, I’ve picked up a thing or two from the brilliant minds around me.

With this document I intend to share some of the most valuable aspects of Microsoft Excel for the SEO. It is far from an exhaustive look at everything that can be done with Excel, but hopefully a strong foundation for the SEO’s toolkit. I’ll be including real world SEO tasks, ranging from the relatively simple to rather complex, so I hope there’s something for everyone.

So if pivot tables, IF statements, absolute references and nested functions make you scratch your head, read on, Aspiring Ninja!

Oh! Before we begin, we must give credit where credit is due. This guide could not have been prepared without the help of some of [Richard Baxter](#)’s awesome blog posts and the official [Microsoft Excel Help](#) site.

Lesson 1: Basic Tasks

In this lesson we'll cover some of the simpler functions available in Excel, and how they're used in the SEO's day-to-day tasks. The functions we'll cover:

- CONCATENATE
- Text to Columns
- COUNTIF
- IFERROR

Concatenate

Microsoft Excel definition: Joins several text strings into one text string.

Syntax: CONCATENATE(text1,text2,...)

Concatenate is a pretty self-explanatory function, but that doesn't make it any less useful. It is most often used to combine two cells into one, as in:

	A	B	C	D
1	Subdomain	Root Domain	Formula	Result
2	www.	google.com	=CONCATENATE(A2,B2)	www.google.com

You may also use the formula to insert text strings before, after, or between other cells. Insert a text string by putting it within quotations, as seen here:

	A	B	C	D
1	Subdomain	Root Domain	Formula	Result
2	www	google.com	=CONCATENATE("http://",A2,".",B2)	http://www.google.com

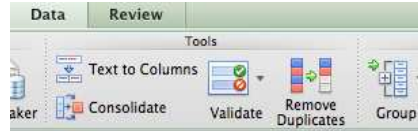
Text to Columns

Microsoft Excel definition: Distribute the contents of one cell across separate columns

Text to Column's functionality is a bit limited, but I find myself using it almost every time I open Excel. As an SEO, this is the go-to function to separate subfolders or divide up subdomains, root domains, and/or TLD. Unlike CONCATENATE, we won't need a formula to carry out this function.



Seen Here in Windows Excel 2007

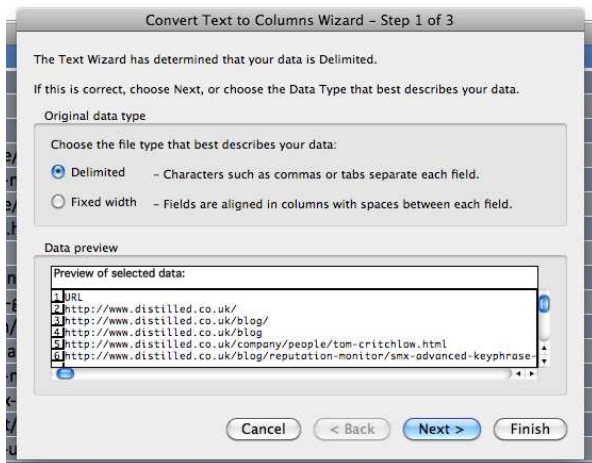


Seen here in Mac Excel 2011

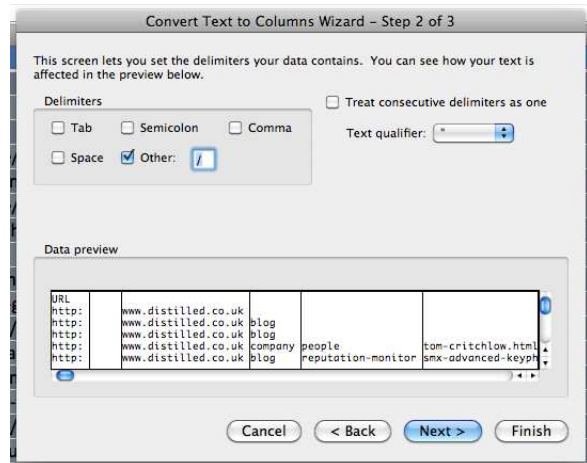
For our real-world SEO example, let's take an OpenSiteExplorer.org Top Pages report and suppose that we want to find which subfolders receive the most links.

1	URL
2	http://www.distilled.co.uk/
3	http://www.distilled.co.uk/blog/
4	http://www.distilled.co.uk/blog
5	http://www.distilled.co.uk/company/people/tom-critchlow.html
6	http://www.distilled.co.uk/blog/reputation-monitor/smx-advanced-keyphrase-research-go-ninja-go-ninja-go/
7	http://www.distilled.co.uk/company/people/will-critchlow.html
8	http://www.distilled.co.uk/conference-calls.html
9	http://www.distilled.co.uk/reputation/
10	http://www.distilled.co.uk/online-reputation.html

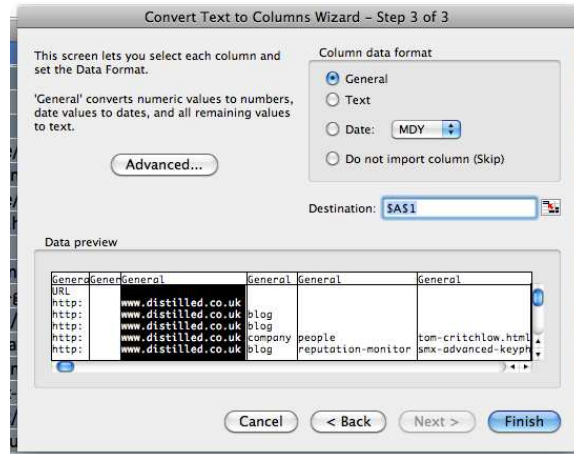
The "http://" is not necessary. This can be removed with a find and replace now, or dealt with afterward.



Select "Delimited"



Choose "Other" and type the "/" key



Format cells (not necessary for this example) and select a destination (default destination is usually fine)

	A	B	C	D	E
1	URL				
2	http:		www.distilled.co.uk		
3	http:		www.distilled.co.uk blog		
4	http:		www.distilled.co.uk blog		
5	http:		www.distilled.co.uk company	people	
6	http:		www.distilled.co.uk blog	reputation-n	
7	http:		www.distilled.co.uk company	people	
8	http:		www.distilled.co.uk conference-calls.html		
9	http:		www.distilled.co.uk reputation		
10	http:		www.distilled.co.uk online-reputation.html		

Voila!

We can now manipulate this data however we see fit. This is great for eCommerce sites with a nice URL structure of */category/subcategory/product/*.

COUNTIF

Microsoft Excel Definition: Counts the number of cells within a range that meet the given criteria.

Syntax: COUNTIF(range,criteria)

COUNTIF is your go-to function for getting a count of the number of instances of a particular string. For instance, anchor text:

	A	B	C	D
1	Anchor Text			
2	Duncan Morris, Distilled Ltd		=COUNTIF(A:A,"[No Anchor Text]")	48
3	Distilled			
4	Tom Critchlow Distilled, Un			
5	[No Anchor Text]			

Looking to get the count of empty anchor text instances? COUNTIF does the trick.

IFERROR

Microsoft Excel Definition: Returns a value that you specify if a formula evaluates to an error; otherwise, it returns the result of the formula. Use IFERROR to trap and handle errors in a formula.

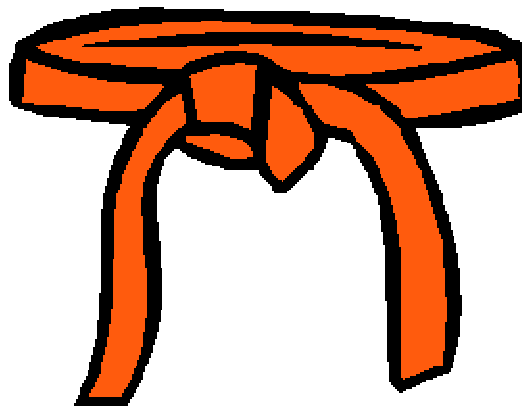
Syntax: IFERROR(value,value_if_error)

IFERROR is really simple and will become an important piece of most of our formulas as things get more complex. IFERROR is your method to turn those pesky #N/A, #VALUE or #DIV/0 messages into something a bit more presentable.

15	4	=IFERROR(T3/U3,"None")	3.75
3	0	=IFERROR(T4/U4,"None")	None

Divide by Zero? No Problem!

Thus concludes Lesson 1 of your Excel for SEO training. Congratulations, this makes you an orange belt. Here's an orange belt that you can print, cut out, and wear around town.



KEEP YOUR DATA ORGANIZED WITH TABLES!

Turning your range into a table will give you banded rows for easier readability:

	A	B	C
1	Domains	Links	Linked pages
2	distilled.co.uk	22,428	277
3	xyp.py.com	11,451	6
4	browsemystuff.com	7,913	1
5	reputationdefender.com	3,149	1

Your Data. Table-ified.

Automatically calculated new columns AND structured/named references:

1	Domains	Links	Linked pages	
2	distilled.co.uk	22,428	277	=Anchors[#This Row],[Links]]/Anchors[#This Row],[Linked pages]]

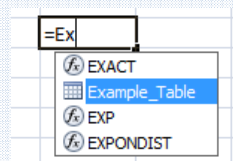
Ooooooh. How intuitive! Now we just hit enter and our column is created and calculated.

	A	B	C	D
1	Domains	Links	Linked pages	Column1
2	distilled.co.uk	22,428	277	80.96751
3	xyp.py.com	11,451	6	1908.5
4	browsemystuff.com	7,913	1	7913

Named columns that remain atop data while scrolling:

	Date	Domain	URL	Code
10	5/10/2010	blog.example.com	/blog/telecom	
11	5/10/2010	blog.example.com	/blog/solder	

Reference tables by name in other formulas:



To convert your range to a table, use CTRL+T on Mac OS X or CTRL+L on a PC.

Lesson 2: More Functions – Text Manipulation

The functions on which we'll be focusing in this lesson are useful for dealing with text manipulation. As we'll see from the examples, there are quite a few scenarios wherein the SEO has to manipulate a text string. Some of the formulas we'll talk about are pretty simple to grasp individually, but can get a bit confusing when used together. We'll touch on:

- LEN
- SEARCH/FIND
- LEFT, RIGHT, MID

LEN

Microsoft Excel Definition: Returns the number of characters in a text string.

Syntax: LEN(text)

	A	B
1	Anchor Text	Length
2	Duncan Morris, Distilled Ltd	29
3	Distilled	9
4	Tom Critchlow Distilled, Uni	39
5	[No Anchor Text]	16

I doubt this requires much explanation. LEN alone is fairly useless. Sorry LEN.

SEARCH/FIND

Microsoft Excel Definition:

SEARCH – Returns the number of the character at which a specific character or text string is first found, reading left to right (not case-sensitive).

FIND – Returns the starting position of one text string within another text string. FIND is case-sensitive.

Syntax: SEARCH(find_text,within_text,start_num) and FIND(find_text,within_text,start_num)

There are two differences between SEARCH and FIND. SEARCH is not case-sensitive, FIND is. SEARCH allows the use of wildcards, FIND does not. Under most circumstances, SEARCH is all you need, but it helps to know that FIND is always there if you've got to deal with pesky capital letters in URLs or something similar. Another reason to choose FIND is if you're dealing with URLs that contain parameters. Without properly escaping question marks, they will act as wild cards, which may cause some frustration.

In our example below, we've pulled out the character number at which the "/blog/" string begins. Much like LEN, this function is a bit silly on its own, but can be combined with some of our other functions to do some cool things.

	A	B	C
1	http://www.distilled.co.uk	=SEARCH("/blog/",A1)	#VALUE!
2	http://www.distilled.co.uk/company/peopl	=SEARCH("/blog/",A2)	#VALUE!
3	http://www.distilled.co.uk/company/peopl	=SEARCH("/blog/",A3)	#VALUE!
4	http://www.distilled.co.uk/blog/	=SEARCH("/blog/",A4)	27
5	http://www.distilled.co.uk	=SEARCH("/blog/",A5)	#VALUE!
6	http://www.distilled.co.uk/company/peopl	=SEARCH("/blog/",A6)	#VALUE!
7	http://www.distilled.co.uk/reputation	=SEARCH("/blog/",A7)	#VALUE!

Now Class, we remember what we do with those #VALUE!s, don't we?

	A	B	
1	http://www.distilled.co.uk	=IFERROR(SEARCH("/blog/",A1),"Non-Blog URL")	Non-Blog URL
2	http://www.distilled.co.uk/company/peopl	=IFERROR(SEARCH("/blog/",A2),"Non-Blog URL")	Non-Blog URL
3	http://www.distilled.co.uk/company/peopl	=IFERROR(SEARCH("/blog/",A3),"Non-Blog URL")	Non-Blog URL
4	http://www.distilled.co.uk/blog/	=IFERROR(SEARCH("/blog/",A4),"Non-Blog URL")	27
5	http://www.distilled.co.uk	=IFERROR(SEARCH("/blog/",A5),"Non-Blog URL")	Non-Blog URL
6	http://www.distilled.co.uk/company/peopl	=IFERROR(SEARCH("/blog/",A6),"Non-Blog URL")	Non-Blog URL
7	http://www.distilled.co.uk/reputation	=IFERROR(SEARCH("/blog/",A7),"Non-Blog URL")	Non-Blog URL

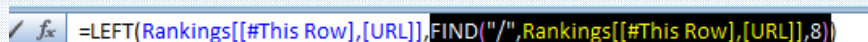
That's right! Wrap an IFERROR around the formula!

NESTED FORMULAS – DON'T BE SCARED!

Also of note in the 2nd example above: This is the first time we've used what's called a **nested formula**. We have these when a function is placed within another function, which can be placed in another function, and another, and so on. The more complex the nested formula becomes, the easier it becomes to break down.

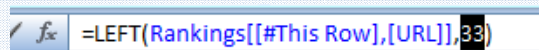
Whether you're reviewing your own formulas for errors, or looking at someone else's work, you should start with the middle of a nested formula and work your way out. Additionally, the F9 key is your friend!

Trying to debug a formula that keeps breaking?



```
=LEFT(Rankings[#This Row],[URL],FIND("/",Rankings[#This Row],[URL]),8)
```

To see the nested interior formula's results, highlight and hit F9



```
=LEFT(Rankings[#This Row],[URL],33)
```

Once you're satisfied, hit ESC, otherwise the calculated result will remain.

LEFT, RIGHT, MID

Microsoft Excel Definition:

LEFT: Returns the specific number of characters from the start of a text string.

RIGHT: Returns the specific number of characters from the end of a text string.

MID: Returns the characters from the middle of a text string, given a starting position and length.

Syntax:

LEFT(text,num_chars)

RIGHT(text,num_chars)

MID(text,start_num,num_chars)

Both LEFT and RIGHT return the characters from a given position in a text string starting from either side of a string. MID is great for extracting a portion of a text string. I've lumped the three together because they are often used in conjunction with each other (along with a few of the earlier functions). Let's dive into an example:

Bringing it All Together - Example 1

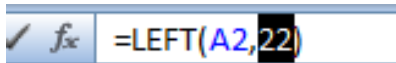
Let's say we've been given a list of URLs, and we want to extract just the domain.

	A	B
1	URL	
2	http://www.seomoz.org/article/search-ranking-factors	=LEFT(A2,SEARCH("/",A2,8))
3	http://www.toprankblog.com/search-marketing-blogs/	=LEFT(A3,SEARCH("/",A3,8))
4	http://www.bdadyslexia.org.uk/	=LEFT(A4,SEARCH("/",A4,8))
5	http://twitter.com/tomcritchlow	=LEFT(A5,SEARCH("/",A5,8))
6	http://www.marketingpilgrim.com/2006/03/online-reput	=LEFT(A6,SEARCH("/",A6,8))
7	http://www.davidmihm.com/local-search-ranking-factors	=LEFT(A7,SEARCH("/",A7,8))
8	http://www.mattcutts.com/blog/free-direct-text-links/	=LEFT(A8,SEARCH("/",A8,8))

This formula will do the job. Let's break down this nested formula, and see how it pulls just the domain out of our URL. Starting from the middle we see SEARCH, which uses the syntax:

SEARCH(find_text,within_text,start_num)

In plain terms, this formula finds the first instance of "/" in the cell to the left, starting at the 8th character from the beginning, which is done to start passed the double slash in http://. As we see below, the result for the first row of data is 22.



The same formula with the inner function calculated

Now we are left with a simple LEFT formula. The syntax for LEFT is LEFT(text,num_chars).

In plain terms: Give us the first 22 characters starting from the beginning. We now have a nice listing of just root domains.

	A	B	C
1	URL	Root Domain	Column3
2	http://www.seomoz.org/article/search-ranking-factors	http://www.seomoz.org/	Search Engir
3	http://www.toprankblog.com/search-marketing-blogs/	http://www.toprankblog.com/	SEO Blogs - I
4	http://www.bdadyslexia.org.uk/	http://www.bdadyslexia.org.uk/	Home Page

Our list of root domains. The formula reflects the change to a table format from the simple range used previously.

Example 2

Let's use SEARCH (with wildcards) and MID together to extract a portion of a URL:

	A
1	Page
2	http://www.example.com/lamp-maintenance-t83.html
3	http://www.example.com/shoe-maintenance-t415.html
4	http://www.example.com/dish-maintenance-t4.html
5	http://www.example.com/tv-t418615.html
6	http://www.example.com/laptop-maintenance-t48147.html
7	http://www.example.com/about-us-t33.html

Let's assume we want to pull the descriptive piece out of each of these URLs for reporting purposes

We'll definitely be making use of MID, as the text we want is in the MIDDLE of our string. We'll need to determine how many characters make up the "-tXXX.html" bit at the end of each URL. Since the length of this portion of the URL varies, but the format doesn't (that is, "-t" + "numbers" + ".html"), we can use wildcards to find this character count.

Again, the syntaxes for these 2 functions:

MID(text,start_num,num_chars)

SEARCH(find_text,within_text,start_num)

Let's break down the formula for the first URL in our list.

Cell A2: http://www.example.com/lamp-maintenance-t83.html

```
=MID(A2,SEARCH("/",A2,8),SEARCH("-t*.html",A2)-SEARCH("/",A2,8))  
=MID(A2,23, SEARCH("-t*.html",A2)-23)
```

*We've calculated the first instance of a "/" after the 8th character. This gives us our start_num values. We're also using the * wildcard to help us get the character count of the right-most chunk of text.*

=MID(A2,23,SEARCH("-t*.html",A2)-23)
=MID(A2,23,40-23)

We can easily calculate the number of characters for our MID once we know where our non-descriptive characters begin.

=MID(A2,23,17)
/lamp-maintenance

Hooray!

Example 2.5

Let's make a small adjustment to our original URL to demonstrate how we can use LEN in this formula.

Cell A2: <http://www.example.com/t1521-lamp-maintenance.html>

=MID(A2,SEARCH("-",A2)+1,LEN(A2)-SEARCH("-",A2)-5)
=MID(A2,29+1,50-29-5)
/lamp-maintenance

The additional +1 and -5 are necessary to make minor adjustments to the final outcome. Without them, our final result would have been "-lamp-maintenance.html".

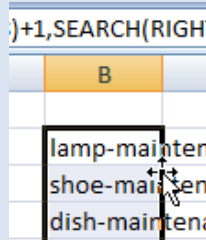
This completes lesson 2. If you've made it this far, you're fit to carry these around. Be careful though!



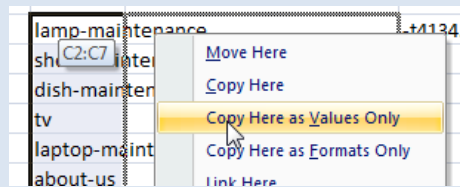
SAVE SOME TIME WITH THIS PRO TIP!

Sometimes you want to grab a range of calculated data and place it elsewhere. However, if that data relies on a formula in place it will fall apart once pasted. In this case, you'd like to copy just the values created by these formulae. The slow way is to copy your data, and paste as values using Excel's 'Paste Special'.

Here's an even quicker way:



*Highlight your range of data and hover over the black line until you see the above cursor.
Now, right click and drag to your desired destination. You can actually hover away from the source, then back over it to replace your formulas, or place it elsewhere.*



Release your right click, select 'Copy Here as Values Only' and you're done!

Lesson 3: IF, OR/AND

Let's slow things down a bit and learn some new helper functions. These guys aren't going to uncover any gold data nuggets on their own, but they can help to organize and present data in your tables, pivot tables and charts.

Let's start with a list of internal URLs, perhaps from the top content report in Google Analytics. We'd like to find out which type of page is seeing the most traffic: the homepage (/), the blog home (/blog), blog posts (/blog/{post}), category page (/blog/{category}/), a services page (there are 4), or other.

	A	B	C	D	E	F
1	Page	Pageview	Unique Pa	Avg. Time	Bounce R	% Exit
2	/	4519	3580	77.879	0.421482	0.398318
3	/blog/soci	2134	1995	345.0708	0.862683	0.841143
4	/blog/	1897	1642	201.0102	0.671921	0.58777
5	/company	1521	1009	54.53292	0.291667	0.151216
6	/company	869	526	31.42448	0.225352	0.116226
7	/contact.h	828	673	68.05535	0.555556	0.345411
8	/search-ei	827	657	66.37179	0.545455	0.245466
9	/blog/dist	652	591	299.0577	0.874317	0.840491
10	/blog/con	616	483	243.6911	0.735556	0.689935
11	/blog/seo	544	481	281.086	0.873662	0.829044
12	/company	526	279	32.13395	0.843137	0.176806
13	/company	506	350	160.5765	0.572464	0.393281
14	/blog/seo	495	446	407	0.891892	0.878788

Our raw data

Ultimately, there are probably a ton of different ways to go about getting this information, but almost every method will involve IF.

IF

Microsoft Excel Definition: Checks whether a condition is met, and returns one value if TRUE, and another value if FALSE.

Syntax: IF(logical_test,value_if_true,value_if_false)

IF by itself is quite simple.

	A	B	C
1	1	=IF(A1=1,"Yay!","Booooo")	Yay!
2	2	=IF(A2=1,"Yay!","Booooo")	Booooo
3	3	=IF(A3=1,"Yay!","Booooo")	Booooo
4	4	=IF(A4=1,"Yay!","Booooo")	Booooo
5	5	=IF(A5=1,"Yay!","Booooo")	Booooo
6			

Hopefully this demonstrates a simple use of IF

It can get a bit confusing when nested (see below) and when combined with other functions (later).

	A	B	C
1	1	=IF(A1=1,"Yay!",IF(A1=2,"Affirmative",IF(A1=3,"w00t",IF(A1=4,"Oh Thank God","This number must be a 5"))))	Yay!
2	2	=IF(A2=1,"Yay!",IF(A2=2,"Affirmative",IF(A2=3,"w00t",IF(A2=4,"Oh Thank God","This number must be a 5"))))	Affirmative
3	3	=IF(A3=1,"Yay!",IF(A3=2,"Affirmative",IF(A3=3,"w00t",IF(A3=4,"Oh Thank God","This number must be a 5"))))	w00t
4	4	=IF(A4=1,"Yay!",IF(A4=2,"Affirmative",IF(A4=3,"w00t",IF(A4=4,"Oh Thank God","This number must be a 5"))))	Oh Thank God
5	5	=IF(A5=1,"Yay!",IF(A5=2,"Affirmative",IF(A5=3,"w00t",IF(A5=4,"Oh Thank God","This number must be a 5"))))	This number must be a 5

The value_if_false in each of these IF statements has been replaced with another IF statement.

Back to our Example

Using a nested IF, we should be able to create a "Page Type" column in our spreadsheet and apply it to our URLs. We'll take it one step at a time.

Let's determine if the URL is for the blog or the home page, since those two URLs are quite simple to identify and only occur once. Since the URL for the homepage is "/" and the blog home is "/blog/", this formula will work:

Note: The table in the following formulas has been named TC, which is reflected in the prefix before the cell references

=IF(TC[[#This Row],[Page]]="/","Home",IF(TC[[#This Row],[Page]]="/blog/","Blog Home","Other"))

fx		=IF(TC[[#This Row],[Page]]="/","Home",IF(TC[[#This Row],[Page]]="/blog/","Blog Home","Other"))			
	A	B	C	D	
1	Page	Page Type	Pageviews	Unique Pageviews	
2	/	Home	4519	3580	
3	/blog/social-media/youtube/youtube-seo/	Other	2134	1995	
4	/blog/	Blog Home	1897	1642	
5	/company.html	Other	1521	1009	
6	/company/jobs.html	Other	869	526	
7	/contact.html	Other	828	673	

Now let's find some way to classify a URL as a blog post or a blog category page. All of the blog post URLs on the Distilled domain follow this format:

/blog/{category-name}/{optional-sub-category}/{post-name}/

And all of the blog categories follow this format:

/blog/category/{category-name}/{optional-sub-category}/

So we'll need to add to our growing IF formula. In plain English, our formula must check to see if the URL starts with /blog/category/, in which case it is a blog category page. If not, we'll get less specific and check to see if it starts with /blog/, in which case it is a blog post. Our formula now looks like this:

```
=IF(TC[ [#This Row],[Page]]="/" ,"Home",IF(TC[ [#This Row],[Page]]="/blog/" ,"Blog Home",IF(ISNUMBER(SEARCH("/blog/category/" ,TC[ [#This Row],[Page]])) ,"Blog Category",IF(ISNUMBER(SEARCH("/blog/" ,TC[ [#This Row],[Page]])) ,"Blog Post","Other"))))
```

In the above, after Excel has checked that the given URL isn't the home page or blog home, it does a SEARCH for "/blog/category/". We then make use of the ISNUMBER function to see if that inner SEARCH function is returning a digit or a #VALUE (#VALUE being the result if the string is not found). ISNUMBER simply checks to see if the given value is a digit, and returns TRUE or FALSE.

	Page	Page Type	Pageviews	Unique Pageviews
3	/blog/social-media/youtube/youtube-seo/	Blog Post	2134	1995
4	/blog/	Blog Home	1897	1642
5	/company.html	Other	1521	1009
6	/company/jobs.html	Other	869	526
7	/contact.html	Other	828	673
8	/search-engine-optimisation.html	Other	827	657
9	/blog/distilled/pro-seo-seminar-2010-key-takeaw	Blog Post	652	591
10	/blog/conversion-rate-optimisation/using-jquery	Blog Post	616	483
11	/blog/seo/how-to-write-a-link-request-based-on	Blog Post	544	481
12	/company/people.html	Other	526	279
13	/company/jobs/junior-seo-consultant.html	Other	506	350
14	/blog/seo/seo-for-tumblr-blogs/	Blog Post	495	446
15	/online-reputation.html	Other	456	369
16	/web-design.html	Other	449	378
17	/blog/seo/first-touch-tracking-in-google-analytic	Blog Post	420	368
18	/tools.html	Other	380	329
19	/pay-per-click.html	Other	364	295
20	/blog/distilled/peathe-out/	Blog Post	362	331
21	/blog/reputation-wars/reputation-wars-mcdonal	Blog Post	347	281
22	/proseminar/	Other	316	277
23	/blog/category/social-media/youtube/	Blog Category	311	278
24	/company/fish/product-supply.html	Other	291	226

Almost there!

For our final piece of the equation we want to classify 4 of our URLs as "Services". Those 4 URLs are /pay-per-click.html, /online-reputation.html, /search-engine-optimisation.html, and /web-design.html. If we can do this properly, every other URL that hasn't been classified will become "Other" and our work will be done!

We'll be adding the OR function to our formula now, so a brief intro, including AND:

OR/AND

Microsoft Excel Definition:

OR: Checks whether any of the arguments are TRUE, and returns TRUE or FALSE. Returns FALSE only if all arguments are FALSE.

AND: Checks whether all arguments are TRUE, and returns TRUE if all arguments are TRUE.

SYNTAX:

OR(logical1,logical2,...)

AND(logical1,logical2,...)

In the context of our example, neither OR nor AND are completely necessary, but helpful in keeping our formula under control. Without these functions we could certainly add 4 separate IF(ISNUMBER(SEARCH's for each URL, but our formula is getting long enough already. Using OR we can simplify slightly while also keeping the formula a bit closer to how we would verbally communicate the formula's steps.

[Back to our Example](#)

Now, we'll pick up where we left off by replacing the value_if_false of the last if statement with the next step.

The new part of our formula looks like this:

```
IF(OR(ISNUMBER(SEARCH("/web-design.html",TC[#[#This Row],[Page]])),ISNUMBER(SEARCH("/search-engine-optimisation.html",TC[#[#This Row],[Page]])),ISNUMBER(SEARCH("/online-reputation.html",TC[#[#This Row],[Page]])),ISNUMBER(SEARCH("/pay-per-click.html",TC[#[#This Row],[Page]]))),"Services","Other")
```

As is hopefully evident, the four ISNUMBER(SEARCH's serve as the **logical1, logical2, logical3, and logical4** of the OR function. As the OR function goes, if any of those logical arguments are true, the entire statement returns TRUE, and the IF statement returns the value_if_true value. I won't be giving a specific AND example, but hopefully its utility is fairly obvious.

	Page	Page Type	Pageviews	Unique Pageviews
2	/	Home	4519	3580
3	/blog/social-media/youtube/youtube-seo/	Blog Post	2134	1995
4	/blog/	Blog Home	1897	1642
5	/company.html	Other	1521	1009
6	/company/jobs.html	Other	869	526
7	/contact.html	Other	828	673
8	/search-engine-optimisation.html	Services	827	657
9	/blog/distilled/pro-seo-seminar-2010-key-takeaw	Blog Post	652	591
10	/blog/conversion-rate-optimisation/using-jquery	Blog Post	616	483
11	/blog/seo/how-to-write-a-link-request-based-on	Blog Post	544	481
12	/company/people.html	Other	526	279
13	/company/jobs/junior-seo-consultant.html	Other	506	350
14	/blog/seo/seo-for-tumblr-blogs/	Blog Post	495	446
15	/online-reputation.html	Services	456	369
16	/web-design.html	Services	449	378
17	/blog/seo/first-touch-tracking-in-google-analytics	Blog Post	420	368
18	/tools.html	Other	380	329
19	/pay-per-click.html	Services	364	295
20	/blog/distilled/peathe-out/	Blog Post	362	331
21	/blog/reputation-wars/reputation-wars-mcdonalc	Blog Post	347	281
22	/proseminar/	Other	316	277
23	/blog/category/social-media/youtube/	Blog Category	311	278
24	/company/jobs/product-owner.html	Other	291	220

Our mini-project is done!

We'll discuss how to go about getting the totals for each page classification in our pivot table lesson. This lesson is complete, congratulations. These ninjas stars will go well with your new IF statement abilities!



KILLER KEYBOARD SHORTCUTS!

As an Excel ninja, sometimes you really should just leave the mouse out of it. Besides, there's no such thing as a slow ninja. That's where keyboard shortcuts come in. Below are some of the most useful shortcuts for the Excel ninja.

Note: These shortcuts are based on the Windows Excel 2007 version, and may (likely) be different for Mac and/or Excel 2010.

Shortcut	Function
F2	Edit the selected field
F4 (When editing a field)	Toggle between absolute and relative references (i.e. A1,\$A\$1,A\$1,\$A1)
F9	Debug/evaluate highlighted piece of formula
ALT,A,R,A	Refresh all – Most useful when changing source data of pivot table
ALT,H,D,L	Delete table row
CTRL+PageUp/Down	Move between work sheets
CTRL+SHIFT+*	Select the current region around the active cell – Great for selecting a whole table when only one cell within the table is selected
SHIFT+F3	In a formula, display the Insert Function dialog box
CTRL+HYPHEN	Delete selected cells
CTRL+SPACEBAR	Select the entire column
SHIFT+SPACEBAR	Select the entire row
CTRL+L	Create a table
ALT,H,V,V	Paste as values

Lesson 4: VLOOKUP, OFFSET, INDEX/MATCH

Here in lesson 4 we'll be talking about some lookup and reference functions. These guys are great for matching up data between two sources, for example, combining keyword volume (data from the Google Keyword Tool) with current rankings (data from your rank checker of choice). In fact, this is the exact example we'll use to demonstrate our first function...

VLOOKUP

Microsoft Excel Definition: Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

SYNTAX: VLOOKUP(lookup_value,table_array,col_index_num,range_lookup)

Combining search volume and current ranking of a keyword list is a great way to prioritize upcoming SEO efforts, and can be done at any stage in a website's life. VLOOKUP makes this a painless process. With our targeted keyword list in hand, we head over to the Google keyword tool to grab search volumes.

	A	B	C	D
1	Keyword	Global Monthly Searches		
2	seo	9140000		
3	ppc	2240000		
4	search engine optimization	9140000		
5	search engine optimisation	1000000		
6	pay per click management	40500		
7	online reputation management	18100		
8	reputation management	60500		
9	distilled	201000		

After cleaning up the output from the GKT, we've got our list of keywords and search volume

Next we'll run rankings for our site using whichever ranking tool we choose, and export that into Excel.

	A	B	C
1	Keyword	Google.com position	
2	online reputation management	44	
3	reputation management	-	
4	distilled	7	
5	seattle seo	-	
6	london seo	15	
7	excel for seo	9	
8	seo blog	167	
9	search engine optimization seo	-	
10	seo search engine optimization	-	
11	search engine optimization seo services	-	
12	search engine optimisation london	38	
13	search engine optimization and seo services	-	

What's that they say about a company not eating their own brand of dog food? Yeesh.

After turning our ranges into tables named "Volume" and "Rank" we can combine our two datasets into one pretty table with VLOOKUP.

In our "Volume" table, we've created a new column and used the following formula to pull in the appropriate data from the "Rank" table.

SYNTAX: VLOOKUP(lookup_value,table_array,col_index_num,[range_lookup])
=VLOOKUP(Volume[[#This Row],[Keyword]],Rank,2,FALSE)

	Keyword	Global Monthly Searches	Rank
85	london seo	27100	15
86	search engine optimization london	27100	23
87	seo reputation management	1300	25
88	search engine optimisation london	33100	38
89	seo training london	390	39
90	seo company london	2400	40
91	search engine reputation management	22200	41
92	online reputation management	18100	44
93	reputation management online	18100	46
94	seo companies london	260	53
95	online reputation management services	1000	99
96	seo courses london	260	109
97	online reputation management tools	1600	138
98	seo london	27100	141
99	seo blog	74000	167
100	online reputation management software	210	192
101	seo agency london	1300	200
102			

In plain English:

Find this exact (range_lookup) keyword (lookup_value) in the "Rank" table (table_array), and return

what is in the second column (col_index_num) of that "Rank" table. Remember, in that "Rank" table, the actual rank is in the 2nd column.

The [range_lookup] portion of the VLOOKUP function is an optional value that specifies whether we want to find an exact match or an approximate match. If range_lookup is TRUE or omitted (for an approximate match), the values in the first column of table_array must be sorted in ascending order. If range_lookup is FALSE (for an exact match), the table_array does not need to be sorted.

In my personal SEO + Excel history, I've never set range_lookup to TRUE. In other words, I'm really only ever looking for an exact match. This simplifies things when it comes to that nasty ascending sort order requirement. If you see a bunch of #N/As when using VLOOKUP, check to see if your range_lookup is either omitted or set to TRUE.

OFFSET

This function is another doozy that I end up using over and over for one specific SEO task – Analyzing compared date ranges from a Google Analytics CSV export. That's not to say this function is a one trick pony; I'd be surprised if you don't find a need for it once it's in your toolbox.

MICROSOFT EXCEL DEFINITION: Returns a reference to a range that is a given number of rows and columns from a given reference.

SYNTAX: OFFSET(reference,rows,columns,[height],[width])

	A	B
1	Keyword	Visits
2	uncontestedly patentable hamming	
3	June 13, 2010 - June 19, 2010	5281
4	June 6, 2010 - June 12, 2010	4555
5	% Change	0.15938529
6	unmatrimonially hematologic evenness	
7	June 13, 2010 - June 19, 2010	2177
8	June 6, 2010 - June 12, 2010	1923
9	% Change	0.13208528
10	superseraphically uncalibrated itinerary	
11	June 13, 2010 - June 19, 2010	694
12	June 6, 2010 - June 12, 2010	569
13	% Change	0.21968366
14	unglobularly telegenic yachting	
15	June 13, 2010 - June 19, 2010	481
16	June 6, 2010 - June 12, 2010	471

After cleaning out the extraneous data we have our compared date range data from Google Analytics. In particular, we have referring keyword traffic from one week compared to the previous. Now, if we're just comparing these 4 (gibberish) keywords we'd be ok, but since we're actually looking at thousands of keywords it won't be easy to see through the clutter.

Using OFFSET we can turn the above, into this:

D	E	F
Keyword	Week 1	Week 2
uncontestedly patentable hamming	5281	4555
unmatrimonially hematologic evenness	2177	1923
superseraphically uncalibrated itinerary	694	569
unglobularly telegenic yachting	481	471
nonpathologically compressional crisscross	387	357
overimpressonably weedless stubby	365	89

Getting our first row of 'Keyword', 'Week 1', and 'Week 2' is simple enough, but how can we get Excel to populate the rest? Let's take a look at the formula:

C	D	E
Keyword	Week 1	Week 2
=A2	=B3	=B4
=OFFSET(\$A\$2,(ROW(\$A3)-2)*4,0)	=OFFSET(\$B\$3,(ROW(\$B3)-2)*4,0)	=OFFSET(\$B\$4,(ROW(\$B4)-3)*4,0)
=OFFSET(\$A\$2,(ROW(\$A4)-2)*4,0)	=OFFSET(\$B\$3,(ROW(\$B4)-2)*4,0)	=OFFSET(\$B\$4,(ROW(\$B5)-3)*4,0)
=OFFSET(\$A\$2,(ROW(\$A5)-2)*4,0)	=OFFSET(\$B\$3,(ROW(\$B5)-2)*4,0)	=OFFSET(\$B\$4,(ROW(\$B6)-3)*4,0)

Let's break the formula down piece by piece, starting with how we're relocating our 'Keyword' column.

SYNTAX: OFFSET(reference,row,columns,[height],[width])

=OFFSET(\$A\$2,(ROW(\$A3)-2)*4,0)

=OFFSET(\$A\$2,(3-2)*4,0)

ROW is a simple formula that just returns the number of the row a cell happens to be. You'll notice that only the column (A) is set as an absolute reference. We want the row to rise as we move down each row.

=OFFSET(\$A\$2,4,0)

UNMATRIMONIALY HEMATOLOGIC EVENNESS

In our example, OFFSET outputs the cell that comes 4 rows beyond the reference point.

=OFFSET(\$A\$2,(ROW(\$A4)-2)*4,0)

=OFFSET(\$A\$2,8,0)

SUPERSERAPHICALLY UNCALIBRATED ITINERARY

The next time the row is calculated, our references will incrementally rise, allowing the formula to adjust according to our raw data.

With some fancy math, we've been able to make the row value jump 4 each time the formula moves to the next row. For our 'Week 1' and 'Week 2' columns, the formula has changed slightly, but the basic premise remains the same.

INDEX/MATCH

INDEX/MATCH is like a more powerful version of VLOOKUP, but it is a bit harder to wrap one's head around. In most cases they are interchangeable, but where VLOOKUP falls short, INDEX/MATCH gets the job done.

VLOOKUP's Shortcomings

The biggest issue with VLOOKUP lies in the first few words of the Microsoft Excel definition:

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify.

In instances where the key field is not in the left-most field, VLOOKUP will not work. Let's say we have a list of the referring keywords of the past 90 days from Google Analytics. For budgeting and reporting purposes, we've classified each keyword as branded, head, mid, or long-tail in a "Category" column in the left-most position of our table.

	A	B	C	D	
1	Category	keyword	source	visits	goal
2	branded	distilled	google	1834	
3	branded	distilled seo	google	368	
4	mid-tail	tumblr seo	google	264	
5	mid-tail	seo for tumblr	google	254	
6	branded	distilled.co.uk	google	187	
7	long-tail	best search engine in the world	google	144	
8	mid-tail	seo tumblr	google	117	
9	branded	distilled blog	google	114	
10	long-tail	mcdonalds vs burger king	google	100	
11	branded	will critchlow	google	99	
12	mid-tail	rss to excel	google	79	

If we were attempting to budget traffic for the next 90 days on another sheet, VLOOKUP would disappoint:

	A	B	C	D
1	Keyword	90 Day Actual	Next-90 Day Estimate	
2	seo	#N/A	#N/A	
3	distilled	#N/A	#N/A	
4	distilled blog	#N/A	#N/A	
5	seo blog	#N/A	#N/A	
6	will critchlow	#N/A	#N/A	
7	tom critchlow	#N/A	#N/A	
8	duncan morris	#N/A	#N/A	
9	excel for SEOs	#N/A	#N/A	
10	tumblr seo	#N/A	#N/A	

Now, we could certainly go back and edit our source table to make VLOOKUP work by moving our keyword column to the left-most column, but INDEX/MATCH is really all we need.

Let's first cover what INDEX and MATCH actually do when used alone.

Microsoft Excel Definition:

INDEX: Returns a value or reference of the cell at the intersection of a particular row and column, in a given range.

MATCH: Returns the relative position of an item in an array that matches a specified value in a specified order.

SYNTAX:

INDEX(reference,row_num,column_num,[area_num])

MATCH(lookup_value,lookup_array,match_type)

Both of these functions are actually quite simple alone, but do something really powerful when used together. The beauty of Excel!

With INDEX, Excel simply returns the value that is found at the intersection of the row and column you've supplied. For example:

	A	B	C	D	E	F
1	1	2	3	4	5	6
2	7	8	9	10	11	12
3	13	14	15	16	17	18
4	19	20	21	22	23	24
5	25	26	27	28	29	30
6	31	32	33	34	35	36
7						
8	=INDEX(A1:F6,3,4)					
9						
10						

INDEX(array, row_num, [column_num])
INDEX(reference, row_num, [column_num], [area_num])

Return the value at the intersection of the 3rd row and the 4th column. The value returned is "16".

With MATCH, we can find the relative row that matches our lookup_value. For example:

	A	B	C	D	E	F
1	1	2	3	4	5	6
2	7	8	9	10	11	12
3	13	14	15	16	17	18
4	19	20	21	22	23	24
5	25	26	27	28	29	30
6	31	32	33	34	35	36
7						
8		16				
9	=MATCH(14,B1:B6,0)					
10						

MATCH(lookup_value, lookup_array, [match_type])

Find the row in which "14" is found within this single column. The value of "3" is returned.

Moving back to our earlier example, we can combine INDEX and MATCH to pick up the slack where VLOOKUP fails.

	A	B	C	D	E	F	G	
1	Keyword	90 Day Actual	Next-90 Day Estimate					
2	seo	=INDEX(Keyword_Table,MATCH(Table2[This Row],[Keyword]),Keyword_Table[keyword],0),4)						
3	distilled	1834	1925.7					
4	distilled blog	114	119.7					
5	seo blog	1	1.05					
6	will critchlow	99	103.95					
7	tom critchlow	67	70.35					
8	duncan morris	48	50.4					
9	excel for SEOs	#N/A	#N/A					
10	tumblr seo	264	277.2					

In plain English: Return the value found at the intersection of the 4th column and the row where the given lookup_value can be found. For the first row of our table: Find "seo" in the Keyword column of

the Keyword_Table and report its row (this is the MATCH portion). Now, return the value found in the 4th column of that table's row (this is the INDEX portion).

There you have it! Sure, VLOOKUP gets the job done 90% of the time, but INDEX/MATCH does 100% of the time, so why not master the one that works all the time!

You're getting Oh So Close to becoming a certified Excel Ninja! For completing this lesson, you've been awarded this completely awesome ninja sword!



WHY LEAVE EXCEL WHEN WE HAVE EXCELLENT ANALYTICS!?

If you're doing some analysis in Microsoft Excel that calls for some Analytics data you can easily open up your browser, login to Google Analytics, create the report you'd like, export that report into a CSV, copy that data from the newly created spreadsheet, and finally paste into a new sheet in your current workbook... Oops, hmm, that wasn't really all that easy, was it?

There's Got to Be a Better Way!

Well there is, friends, and it's called Excellent Analytics (<http://excellentanalytics.com>).

With Excellent Analytics, you can automatically import Analytics data into your current spreadsheet without leaving Excel. There are a few quirks, but once you've gotten the hang of using the Excel plugin, you'll save tons of time.

Unfortunately, at the time of this report's creation, Excellent Analytics only works with Excel 2007 for Windows.

Lesson 5: Pivot Tables

This is the lesson where it all comes together. You've got your all of your data calculated and properly formatted, and now it's time to analyze it all and build an action plan. Surely, this can be done without the use of pivot tables and charts, but this powerful feature can make it a much easier process.

To demonstrate one use of a pivot table, we're going to bring back the example from lesson 3, in which we categorized each of the pages on our site.

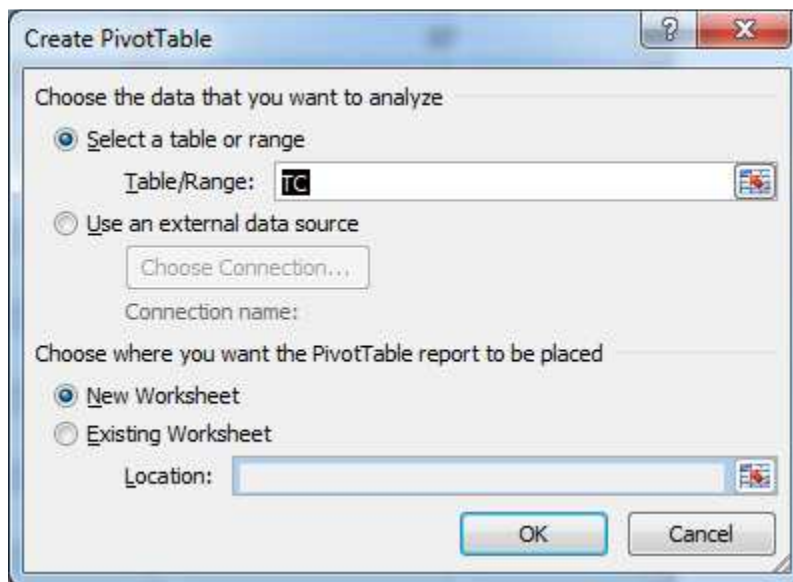
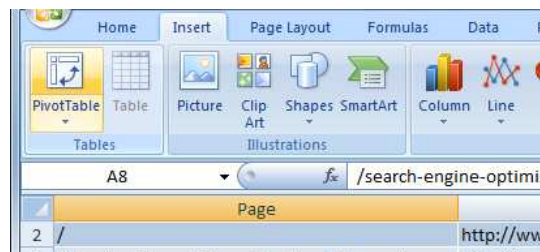
	Page	Page Type	Pageviews	Unique Pageviews
2	/	Home	4519	3580
3	/blog/social-media/youtube/youtube-seo/	Blog Post	2134	1995
4	/blog/	Blog Home	1897	1642
5	/company.html	Other	1521	1009
6	/company/jobs.html	Other	869	526
7	/contact.html	Other	828	673
8	/search-engine-optimisation.html	Services	827	657
9	/blog/distilled/pro-seo-seminar-2010-key-takeaw	Blog Post	652	591
10	/blog/conversion-rate-optimisation/using-jquery	Blog Post	616	483
11	/blog/seo/how-to-write-a-link-request-based-on	Blog Post	544	481
12	/company/people.html	Other	526	279
13	/company/jobs/junior-seo-consultant.html	Other	506	350
14	/blog/seo/seo-for-tumblr-blogs/	Blog Post	495	446
15	/online-reputation.html	Services	456	369
16	/web-design.html	Services	449	378
17	/blog/seo/first-touch-tracking-in-google-analytics	Blog Post	420	368
18	/tools.html	Other	380	329
19	/pay-per-click.html	Services	364	295
20	/blog/distilled/peathe-out/	Blog Post	362	331
21	/blog/reputation-wars/reputation-wars-mcdonal	Blog Post	347	281
22	/proseminar/	Other	316	277
23	/blog/category/social-media/youtube/	Blog Category	311	278
24	/company/jobs/product-owner.html	Other	291	220

Remember Me?

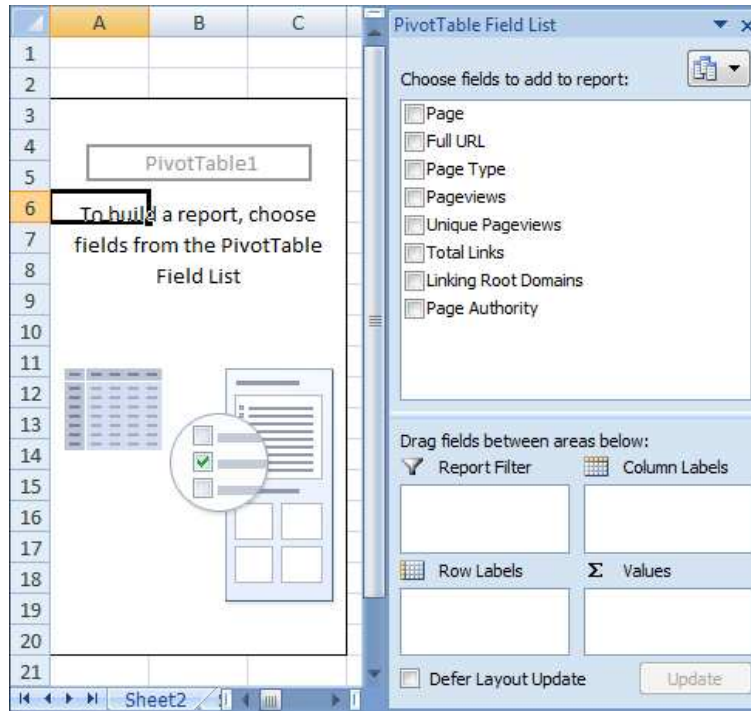
To get some additional insights from our pivot table I've added SEOMoz Page Authority, total links, and linking root domains for each page.

Full URL	Page Type	Pageviews	Unique Pageviews	Total Links	Linking Root Domains	Page Authority
http://www.distilled.co.uk/	Home	4519	3580	9863	411	67
http://www.distilled.co.uk/blog/social-r	Blog Post	2134	1995	57	27	49
http://www.distilled.co.uk/blog/	Blog Home	1897	1642	6299	360	63
http://www.distilled.co.uk/company.ht	Other	1521	1009	627	5	45
http://www.distilled.co.uk/company/fo	Other	869	526	24	8	44
http://www.distilled.co.uk/contact.htm	Other	828	673	619	5	47
http://www.distilled.co.uk/search-engir	Services	827	657	627	6	46

With our data all set up, we can get this party started:



If you've focused on a cell within your table, Excel should auto-select the whole table for pivoting.



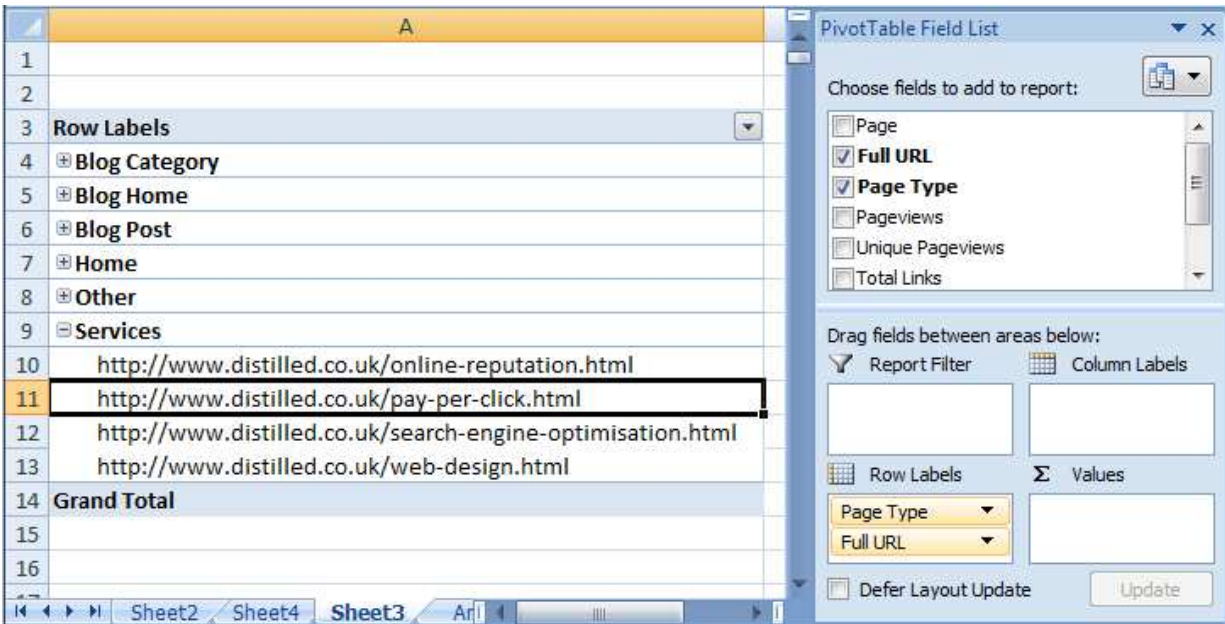
Our pivot table has been created in another sheet and is ready to be built. Before we do so, let's define some of the items seen here.

Report Filter: This is used to apply a broad filter to all of your source data. It should help to view the report filter as the first defense for your pivot table. Before any filtering and sorting can be done, the terms of the report filter must be met. There are a couple of ways we might use this filter for our example:

- Show only URLs with at least # links/pageviews
- Include only URLs with Page Authority above #
- Show only "blog post" page types

Additionally, we can include any combination of the above because more than one report filter can be applied.

Column and Row Labels: This one is pretty simple. Which fields do you want to orient to columns and rows? You can include more than one field in each, which will group your fields together. For example:



By including 'Page Type' before 'Full URL', the rows are grouped first by page type. Expanding reveals the full URLs that appear within the group.

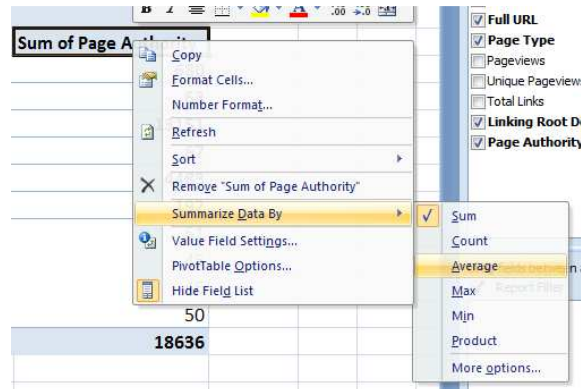
Defer Layout Update: With this option unselected, anytime a field is added to or moved within the report, Excel will recalculate everything. If you're dealing with a very large set of data, you may want to consider selecting this option, then clicking update when you're ready for Excel to work its magic.

Refresh Data: If you make any changes to your source data while building a pivot table, the changes won't be reflected in your pivot table until you've refreshed your data. Do this by right clicking within your pivot table and selecting refresh or finding it in the ribbon menu.

Values: This is where the data will be summarized. Depending on which field we assign as a value, we can choose to get a sum, a count, an average, and more. An example:

Row Labels	Sum of Linking Root Domains	Sum of Page Authority
Blog Category	21	680
Blog Home	360	63
Blog Post	1594	13151
Home	411	67
Other	603	4483
Services	40	192
http://www.distilled.co.uk/online-reputation.html	18	51
http://www.distilled.co.uk/pay-per-click.html	3	45
http://www.distilled.co.uk/search-engine-optimisation.html	6	46
http://www.distilled.co.uk/web-design.html	13	50
Grand Total	3029	18636

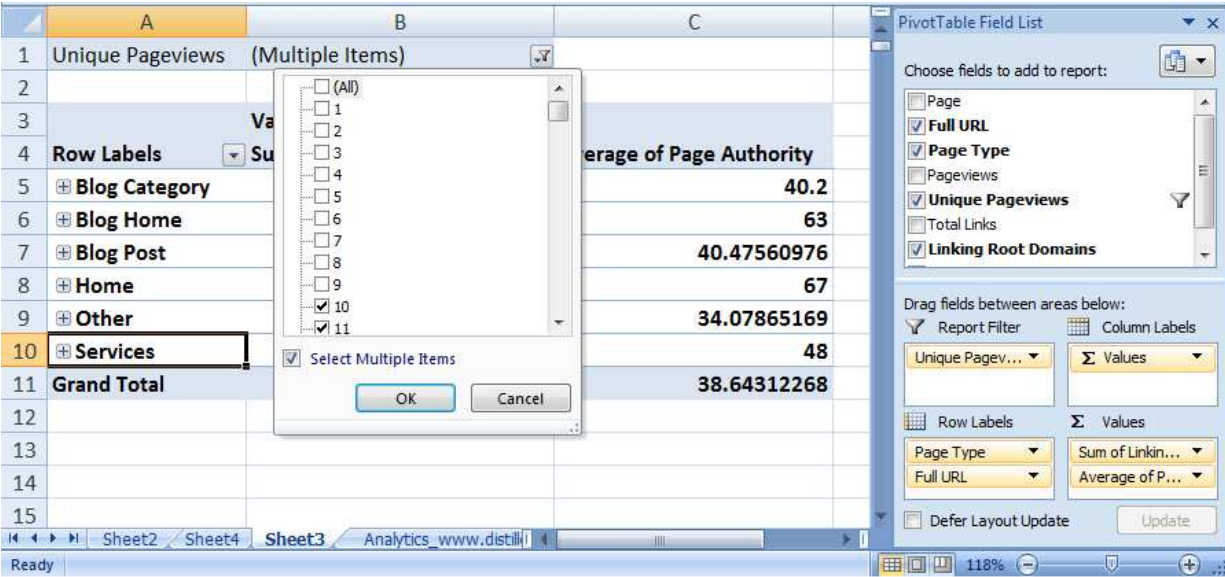
We like that the linking root domains are summed, but a sum of Page Authority isn't very helpful.



By right clicking on the column header we can change this to get an average of Page Authority

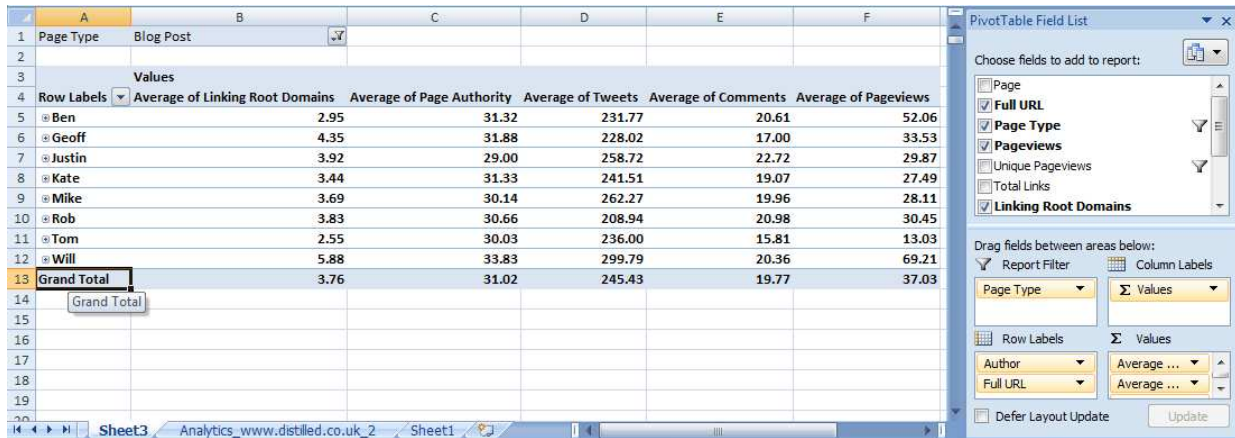
Row Labels	Sum of Linking Root Domains	Average of Page Authority
Blog Category	21	28.33333333
Blog Home	360	63
Blog Post	1594	31.01650943
Home	411	67
Other	603	20.85116279
Services	40	48
http://www.distilled.co.uk/online-reputation.html	18	51
http://www.distilled.co.uk/pay-per-click.html	3	45
http://www.distilled.co.uk/search-engine-optimisation.html	6	46
http://www.distilled.co.uk/web-design.html	13	50
Grand Total	3029	27.85650224

There we go.



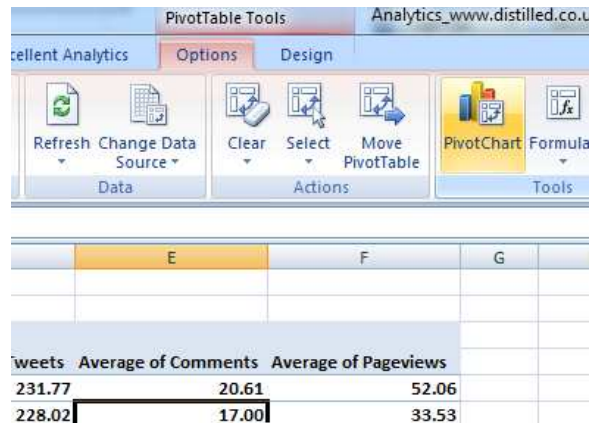
Here we've removed pages with less than 10 unique pageviews from our pivot table.

Now we'll introduce some new (*dummy*) data into our source to make this a bit more interesting. Let's suppose we want to do an audit of the content writers for our blog.

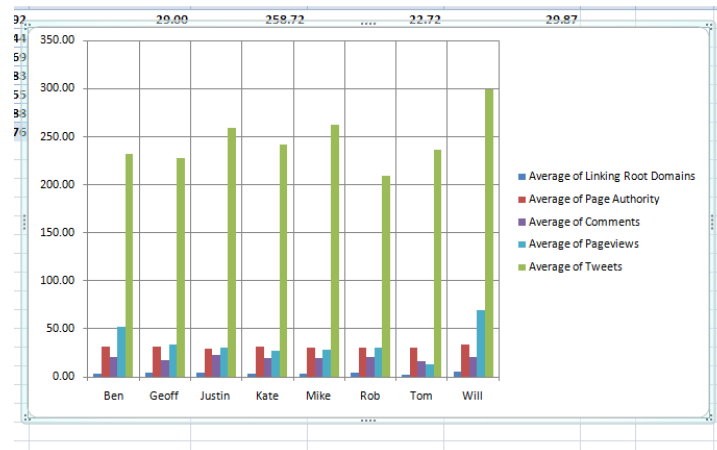


Setting all of our values to summarize by average, we've got each author's average # of links, PA, tweets, comments, and pageviews per post.

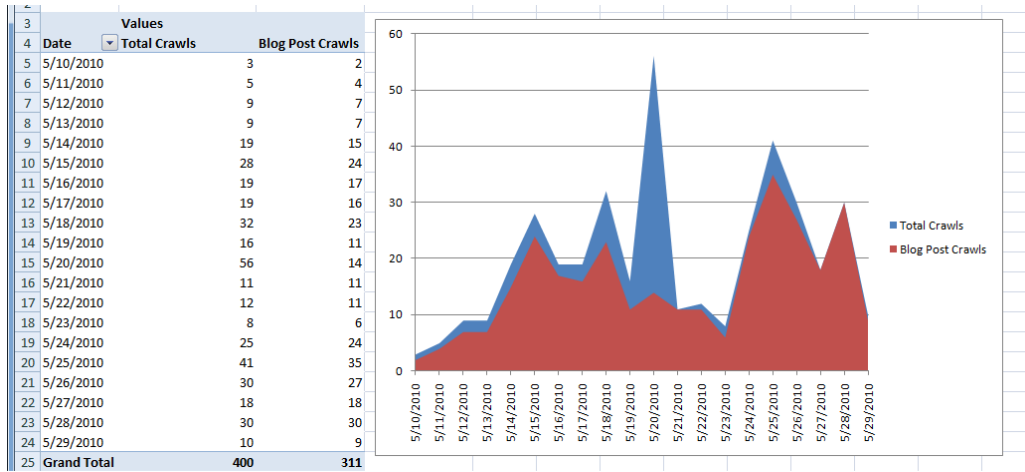
Another great thing about pivot tables is how easily they can be visualized with pivot charts. Once we've got our pivot table set up, we can choose to display as a pivot chart:



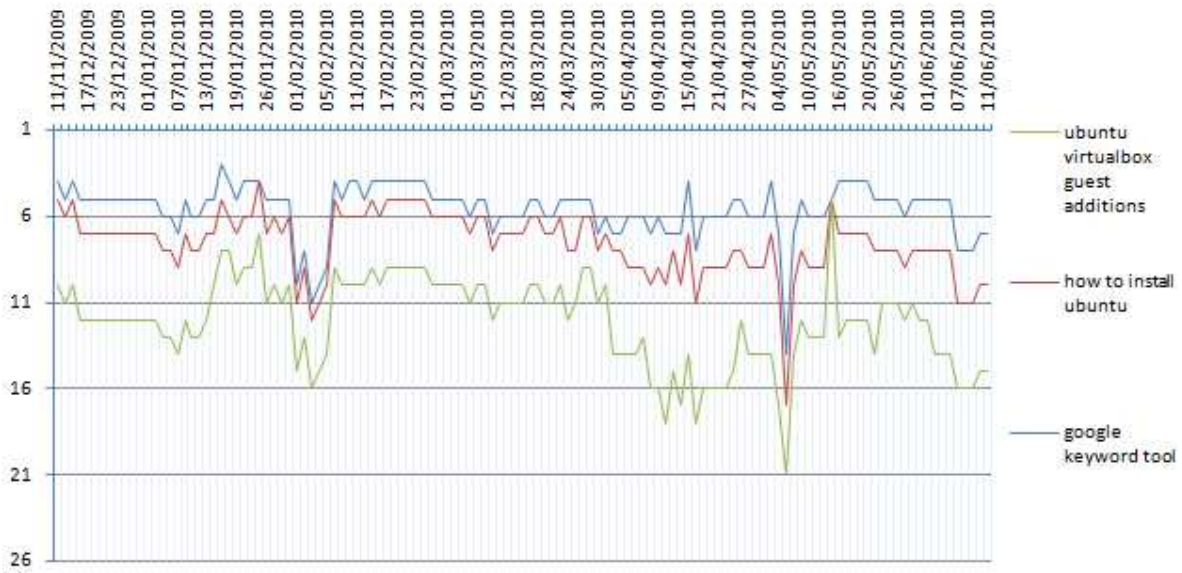
A few clicks later I've got a visual representation of my pivot table.



There's no easier way to learn how to use pivot tables and charts than by firing up Excel and giving it a shot. Once you've got a grasp on HOW pivot tables can be useful, you'll have no problem making them work for you. Here are a few completed pivot tables that hopefully demonstrate they're immense utility:

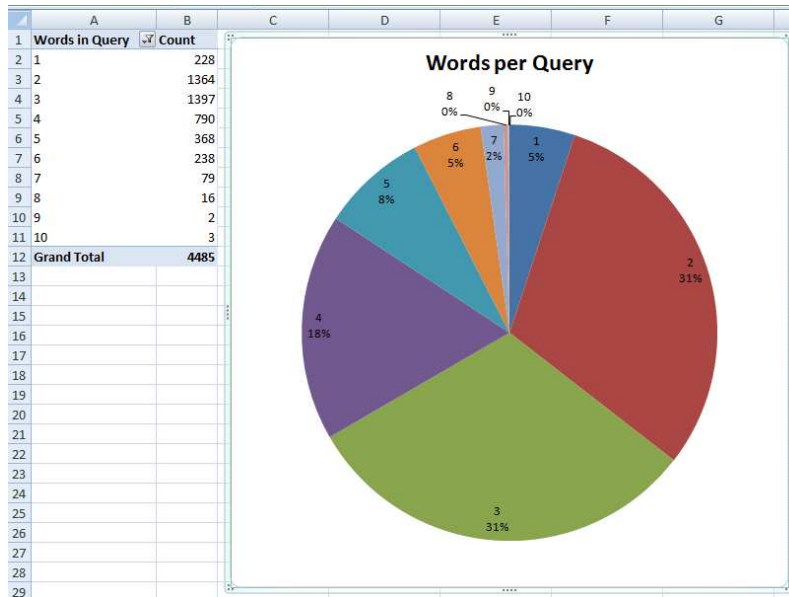


Googlebot crawl analysis



Richard Baxter's awesome ranking pivot chart¹

¹ <http://www.seomoz.org/blog/how-to-make-awesome-ranking-charts>



Visualizing number of keywords in a query

Row Labels	Sum of Links In	Count of Address
/	189340	10
/blog	345294	7346
/example-subfolder	59098	8
/example-subfolder2	89126	117
/example-subfolder3	4560596	360203
/example-subfolder4	73847	666
/example-subfolder5	35981	60
/example-subfolder6	1039	159
Grand Total	5354321	368569

URLs and links into those URLs organized by subfolder

Excel Ninja Status Achievement Unlocked!



This concludes my Excel for SEOs training lessons. If you've made it this far then you should be fit to take on dozens of 10 year olds, evil trolls, rabid wolverines, and just about any pirate you might ever meet! Now, go forth and use your new skills wisely!