

Module 3 Video 3: Connecting websites together

[00:00:00] Welcome to our 3rd instructional video for week 3. This is the one I've been hinting at the entire time that I'm going to show you a very cool technique for connecting Web sites together.

[00:00:10] So we're finally here. Let's see how to do that. First, a bit of context. OK. The thing that you need to understand is we're able to connect Web sites together from something that you can't actually see when you're on the Web site because of two very popular, free, widely used programs from Google. One is Google Ad Sense, which is a product you can use to place ads on your Web site and earn money. The other is Google Analytics, which you can use to track the visitors and traffic to your Web site.

[00:00:38] Now, when you just when you sign up for these, you create an account. And when you put them on your Web site, there is a unique account I.D. that is going to be on every page where you have the ad or the analytics tracking. So what we're able to do is figure out if someone using one account for ad sense or one account for analytics has that same I.D. on multiple Web sites. And then that's a very strong signal that there is some connection between these Web sites. So how do we do this?

[00:01:06] OK. What we're gonna do as I give you the background is I'm going to walk you through slide by slide. What I want you to do is open your Web browser and you're going to follow along, as I say it. Now, if it seems to be going too fast, I want you to pause and make sure you follow along and execute the step before you move on to the next slide that I'm going to walk you through. The general thing about what we're gonna do here is we're going to look on the source code of a Web sites page, which I'll show you how to do for two different strings of text. One is "UA-", and that's for analytics I.D.'s and the other is "Pub-". That's for Ad Sense I.D.'s.

[00:01:44] So every single unique I.D. from one of these products begins with these first three characters. So if you find those first three characters on the source code of a Web site, that means you're going to find the I.D.. And then we take that I.D. and we run it through a bunch of sites that can enable you to search for connected Web sites using that same I.D.. And I'll show you what those are.

[00:02:06] Now, the last tip that you can also do is remember the Wayback Machine we looked at in the previous video. If there isn't an I.D. on a page that you're on right now, you may want to go and look at an older version of that page, an older version of that site, because you can do the same technique for an archived page and see if in the past maybe there was an analytics code or an ad sense code.

[00:02:29] All right. Maybe it seems a bit complicated right now, but I promise you, it's not because I'm going to walk you through step by step.

[00:02:35] All right. Open that Web browser. Here we go. Step one, you're going to go to the Web site, Wolf. Tribune dot com. OK, once you have the home page loaded, I want you to go up to the file menu in Chrome. You're going to select developer and then view source. OK. That's going to open a new page in your tab. OK. On your Web browser.

[00:02:57] And it's going to look like this. So a bunch of html. A bunch of code. OK. Doesn't matter if you understand code or not, because that's what we're gonna look with in the text for that unique set of characters. You a "UA-" or "Pub-". Here's how you find stuff within a page. You go up to the edit menu in Chrome. All right. Follow along. You go to find and then find again. And that's going to open up a little search box. So here's what that search box looks like, OK? And here's what I entered you "ua-". What's going to happen is that's going to automatically search through all the characters on the page to find "ua-".

[00:03:36] And this is what you should see. You should see it pop up. Highlighted "ua-" with 5 5 1 8 3 6 8 6 - 6 after it. OK. So that is an analytics code that we have found in the page. That means they're using Google Analytics. That's a great sign. We can move from there. OK. So mark down

that number, UA-55183686-6. OK. Put that down somewhere because we're going to come back to it.

[00:04:06] Now, here's the other thing. Of course, we want to do. We want to see if there's an ad sense I.D. So you're going to go to that same. And, you know, edit menu again, select find and find again and then put "Pub-" in the search box. Now, in this case, we get nothing. We can see that it's showing a 0 out of 0 hits. So there isn't an ad sense I.D. on this site as of right now. But to give you a little hint, if you looked in the Wayback Machine, you might find one in the past. OK. That's something you can do on your own.

[00:04:36] So we struck out on AdSense. We hit Pader on Google Analytics Code. Now you take that code. And just to remind you again, it's that one, "UA-" followed by the numbers with another dash and a six. So you're going to take that, please copy it. And then the first one we're going to go to, it's a Web site called DNSlytics.

[00:04:56] OK, please open a tab in your Web browser for DNSlytics.com/reverse-analytics. Go ahead and pause right now if you need to to get that loaded. And then, of course, you're going to open it up. It's going to bring you to a page where you can enter in that UA code. You can see here on the slide. I've put my UA code in. I've hit go.

[00:05:19] And this is what you're going to see. Now, note that they've got this red text on the bottom saying we're only going to see a few results because we're not premium members. If you want to sign up, it's a great service, but it costs money. But with the free results, we can see that it found 10 domains using analytics, I.D., ua-551 and onwards. And we can see that Wolf Tribune is one of them. Number three there at the very bottom of the slide. But there's two other Web sites, tecake.com and tecake.in that also have it. OK, so now we've got two other Web sites that seem to be connected to WolfTribune.com.

[00:05:54] Here's the other thing that's really interesting about these results. Let me highlight it for you. tecake.com and [.in](https://tecake.in) both have an adsense ID. So you can see here in the results, it lists the analytics I.D. used by 10 domains. Then it lists an ad sense I.D. pub- used by 12 domains. So apparently there is overlap between sites with this particular analytics I.D. and an ad sense I.D. And so that may give us more sites to look at. And if they're all using the same analytics I.D. and the same ad sense I.D. that's a very strong indicator that they're connected and maybe run by the same owner. OK. So we've already found apparently at least two and maybe eight more Web sites connected to WolfTribune.com.

[00:06:40] But I think it's also really important to not just use one source. So I want to show you another site called SpyOnWeb.com. Please load this in your Web browser. Please also paste that UA. code in again. And let's see what [SpyOnWeb](https://SpyOnWeb.com) tells us. OK. So you do that, you hit. Go. And here's what you're going to see. We're gonna see tecake.in tecake.com. So again, there's overlap at saying yes. We've also seen this analytics idea on these sites. But we've got two other results: swiftnary.com and drugscart.com. So now instead of just having two other Web sites that we know of. We've now got four other Web sites that apparently either now or at some point in the past have had the same analytic I.D.. And now you've got more Web sites to investigate to see what the connections here might be. Right.

[00:07:29] So that's the really quick run through of how we can connect Web sites together using this simple kind of search tools. And this simple looking for characters in source code. And the great thing about this is it opens up whole new avenues of investigation. And just to give a sense of the kinds of things that I might do at this stage, you know, the first thing I would do is, of course, we want to do our content analysis, WolfTribune.com and our social sharing and engagement analysis on WolfTribune.com. Maybe there are connected Facebook pages, for example. But also the about text. Is it stolen? Is it copied? Does it make sense? A hint to you is I really encourage you to read the about texts and to compare it to other ones online.

[00:08:11] And then, of course, with these other domains and with WolfTribune.com, we could do Whois searches to see if we can find more info about the owners. And we can start to dig in more and tecake.com, [.in](https://tecake.in), those other two sites, contents analysis, Whois analysis, archive analysis, social sharing analysis, all the things we've already gone through with these new targets. And

over time, we can start to dig into the people who say they run these sites, the people whose names we find, the people who are listed on the Web sites, and to learn more about them using those skills that you learn from Brandy in week one. You can see how Web sites open up so many possibilities, how with just some knowledge of the mindset and of some tools. There is so much that you can do with one Web site to connect it to others, to connect it to people, to start pulling together a network and understanding what's really going on.

[00:09:05] And that's a good Segway into what's going to happen in week four when you dig in to network analysis. So I hope you've learned a lot here in week three. It's been a pleasure to teach you. Looking forward to seeing what you do on the quiz. Looking forward to engaging with you in the comments. And I hope you have a really rewarding learning experience. Thanks.