

Google Analytics Filters and Custom Segments Training Transcript

Welcome to training on using Google Analytics Filters and Custom Segments. This is an intermediate level training session and we're going to cover the following topics. Just a quick review of where you can find resources for Google Analytics both within the network and on the internet, YouTube in general. A review of basic filters. These are the kinds of filters you apply to the report. This is what we cover in the Introduction to Google Analytics session. Review of dimensions versus metrics and defining dimensions versus metrics. Then we'll get into Standard Segments. These are default segments that are standard with any account if you're a user you have a set of standard segments and we'll go over them and how to use them. Then talk about Custom Segments, which are a powerful tool that can open up Google Analytics and allow you to do a lot more and gather a lot more data.

[Review of Resources, 1:15]

So with that, just want to remind you that we have the Web Analytics resource directory where you can find a lot of information. It's organized by the tool set that we currently have and Google Analytics section in particular has a lot of resources. Just by clicking on the more Google Analytics resources link you'll get an index -- essentially an index of all the resources and pages that we have pertaining to Google Analytics and there are a number of them very specific to the activities we do in our training sessions. So I encourage folks to check those pages out and see if you can get what you need there. And in addition you can look on the internet. Google has a lot of resources, YouTube included. They put out a lot of short videos that you can access for additional training. So that was just a quick reminder where you can find resources.

[Review of Basic Filters, 2:19]

I want to go into the interface now and you should be familiar with the interface. And I'm looking at the Audience Overview report in the EPA Enterprise view which is the main view where all new users are given access. I wanted to quickly review what we covered in the Introduction to Google Analytics, and that is using basic filters to find your top pages.

So you may be familiar with the All Pages report. I'm going to go into the Behavior folder on the left side of the screen and under that the subfolder for Site Content and the All Pages report. Because we're in the EPA Enterprise view of course we will get a report of all the top pages across our entire collection of public websites. Including EPA, Energy Star, AirNow, and a few others. But if you'll recall you can use this filter box in the top middle of your data table to drill down on the top pages in your web area.

So I'm going to use a filter for `epa.gov/lead` and that should bring up the top pages in the lead -- any top page in fact that has a lead top level directory. In this case lead is a new microsite in the Drupal WebCMS so that's what we're seeing in the results. If you recall there's one more step in the process and that's removing the search results pages. Let's do that very quickly.

If you don't remember I'll show you that by expanding the data table and I'm going to expand it out to 100 rows, somewhere in the long tail of your report you'll find search results pages with the subdomain `nlquery`. So let's scroll down. They were not in the top 100. Let me expand that to the top 250 pages. See if we can find any search results pages for lead. You'll recognize it by a very long URL and there's one number 215; `nlquery.epa.gov`. At some point, someone would have executed a search from the very

pages that you're building a report for. So from epa.gov/lead or from one of the lead pages someone executed a series of searches, and any of those searches will result in an nlquery.epa.gov page, or search results page, with the parameter of the URL you are searching for or that you are building a report on and those are sort of false positives you need to remove from the report.

So going back up to the very top of the data table you will recall we click on Advanced next to the search box or filter box and we build a second filter by clicking Add Dimension and the only dimension in this report is page. So, we go to Exclude instead of Include. Exclude pages Containing, and you can use Containing, nlquery.epa.gov. So, the first filter that you put in the box automatically is built by clicking Advanced you can see under the hood, so to speak. That is an Include filter including pages matching your string. And the secondary filter is excluding pages containing the subdomain for search results. And click apply, and it should be a few thousand less -- maybe 1,000 less than our previous report.

But that's how you go about doing a basic filter to get your top pages. And that's an important first step but there's a big difference between filters on the report level like we just did and filters you'll use to build segments which we'll go over in a little bit. But before that I just want to kind of go over a refresher of what a dimension is and what a metric is.

[Dimensions v Metrics, 7:21]

So we just did a couple of filters using the dimension page. Dimensions describe the data in the report. They're a descriptive attribute of the report. So in this case page can be described by many, many different URLs. But that's the dimension of the report and the metrics are the measurements of that descriptive data. The page view, the unique page view, the average time on page, all of the other columns in the data report or in the data table are metrics. Where the first column is page, the dimension is page, and that's describing and really setting up the data that we're looking at.

I'll open a different report just to give you one more example of a dimension versus a metric. Under the Audience section there is a subfolder for technology. And one of the reports is for browser and operating system. So I'll click on that and this is again for the entire agency. All of our visits, all of our pages. So this is very high level. But we're now looking at a report on browsers. So this data table has a dimension of browser providing descriptive attribute like Chrome, Internet Explorer, Safari and Firefox, and being measured by metrics like sessions, new sessions and new users.

But if you wanted to drill down and find the browser for your web area you could not do what we just did with the all pages report. You cannot go into the advanced filter area and filter for your page because page is not a dimension in this report. The only dimension in this report is browser. So you won't be able to filter this report using basic filters to find what are the browsers for the people who visit just your pages.

The difference is really defined in the difference between a filter and a segment. A filter is at the page level or at the data table level. So when you build a filter like we did looking for the lead pages, you're basically asking the data table or the report to show you just those interactions where pages equal the lead string or the lead URL or the lead directory. A segment which we'll demonstrate in a minute actually works at the visit level or the session level which is to say that you are then filtering on what a user does across their entire visit on an EPA website.

That distinction is not always easy to grasp right away but through some examples hopefully you'll begin to see the difference and you'll see that you can access more data using segments than you can with basic filters. It's different data but you'll see the difference as we demonstrate it.

So let's go back to the all pages report. Under Behavior, Site Content, All Pages. And let's just look at the lead pages again. What if I only want the microsite in the Drupal WebCMS; I might do www2.epa.gov/lead. Now I'm filtering on the dimension page telling the report to just bring me back the interactions where the page is equal to the lead URL I've entered. So it's a filter on the page level or an interaction level independent of any visit or session a user may take across the entire EPA website.

If you notice just above the trend line in the All Pages report or in any report that you're in there is a blue circle and right next to the circle it says All Sessions 100%. So regardless of the color you'll notice the top left just above the trend line a circle and it reads All Sessions 100%. That indicates that we have not applied any segments and that in terms of the sessions or the visits that people have taken to the EPA website we're looking at all of them.

[Standard Segments, 12:58]

To add a segment there's a translucent box to the right of the All Sessions circle that says Add Segment. There's a plus sign and it says Add Segment. So click on that and you're now presented with a menu of segments that are available to you. On the left side of the box that opens you can arrange and organize the segments available. All will show you all the segments that you have. System will show you the default segments or the basic or standard segments that are provided to any user who has access. Then underneath of that there will be Custom Segments, the segments you create and save. And Shared Segments in case they're shared across -- among all users. And Starred Segments in case you want to create a favorites list.

So if I go to System I'll just show you what's available to you and this might give you an idea of how segments are applied. These are not created by me but default segments available to everyone. There are segments on bounced sessions or bounced visits. Visits of one page only. Direct traffic visits. Mobile and tablet traffic. Mobile traffic by itself. A series of different segments are available. So let's try one. Let's click Mobile Traffic. I just select the box next to mobile traffic and underneath there's an Apply button in the bottom left corner. Click Apply.

Now you'll see there's two lines of data on the trend line. One for all sessions and one for just those sessions where the user had a mobile device. And then if we look at the data table below it will break out sessions for every lead URL because remember we still have our basic filter applied to the data table but this time it will show you -- like for the lead homepage number 1, total there were 41,000 page views. But amongst mobile sessions there were 5419 page views. So in this 41,000 is included the 5400 mobile page views but it's breaking that out as a separate segment and you can see for each and every URL.

Back at the top there's an arrow, downward arrow, in the corner of the All Sessions box and the same in the Mobile Traffic box. You can remove one of these segments and I'm going to remove All Sessions so that we can look at Mobile Traffic in isolation. Now you see that this report is still showing you the page views for lead but now only those page views for people who are using mobile devices. So it's changed the report not because of the basic filter of what we're asking the report to display but because of the visits from which this report is drawing data. It's now saying we're just looking at mobile traffic and on

top of that the data table should only display pages for lead. And so it's the combination of the segment for mobile and filter for lead that allows you to look at mobile traffic amongst lead pages.

So that's a Standard Segment that you can apply. And let's look at another one. I'll go back to Add Segment. Back to the System list. And I'm going to unclick Mobile and click on New Users. And hit apply. Now we're looking at in isolation just those new users who during the time frame that we selected in the top right which is the last 30 days are new users. They haven't been to the lead pages before during this month. So this is a fraction of all the page views represented by just those sessions or visits where the user's coming for the first time in that month.

So you can see there's different ways to segment the data. The data table still has a filter so it's always going to just display lead pages but now we're changing the visits from which this data is drawn allows us to segment almost by in this case an audience type, audience being new users. And previously the audience being mobile users. So that gives you a feel for how some basic segments are used.

[Custom Segments, 18:32]

To really open up what you can do with Google Analytics especially using the EPA Enterprise view you'll want to be able to create a Custom Segment for your web area. Remember what we're putting in the string for the basic filter `epa.gov/lead`. We're going to use that exact same string but to build a segment, a segment for lead users. A segment for lead sessions different than the basic filter we're using here.

I'm going back to the top, going to click on Add Segment. I'm going to unclick New Users, and in the top left there's a box for and a plus sign and it says New Segment. I'm going to click on New Segment. I want to build a segment for any session that includes a view of a lead page. So I think after we build it, it will be clear the difference between the filter and the segment. I'm going to name this Lead Microsite Session.

When you build a Custom Filter there are a number of choices that are drop down, preselected. Like under Demographics there's some predefined choices you can select. Under Technology there's choices for browser type. You can build a segment for instance just on the Firefox browser. Behavior has choices like Days Since the Last Session with a number of sessions this user could have had in a month or a different time span. But we're actually going to use the Advanced feature underneath and click on Conditions; Advanced conditions. It brings up this filter box where the default is the filter on sessions and it's an include filter. That's what we want. But the dimension is defaulting to Add Content. And we want to change that to Page. So you can either search for it below or just type in page. You type in page you'll still have to search because there are a lot of different dimensions including the word page. But scroll down after you type in page to find page in isolation and select Page.

So now you have a filter on sessions that includes pages and in this case you can use containing because we're not using regular expression, `epa.gov/lead`. What happens is to the right of our filter it already tests to see if this is a legitimate filter and it says that, yes, 1% of all EPA users visited a lead page during this time frame. There were 83,000 sessions by 63,000 users. And that's it. You do not have to filter out search results pages because of the nature of a segment we're just interested in users who viewed a lead page. If they viewed a search results page as well, that's fine. If they viewed the EPA homepage that's fine. As long as they viewed a lead page we're considering them a lead microsite user for our purposes.

So in that case we've done our job. We have a filter for session that includes pages containing the string for the lead microsite. So we save it in the top middle there's a button to save and we've now created a segment for our web area. We're going to remove the New User segment. I'm going to remove the basic lead filter from the data table and let's look at what we have. We're still in the All Pages report. But we have not filtered the data table. The only thing we're doing is looking at people who have visited a lead page and the all pages report will tell you all the pages they've visited. Not surprising we see lead pages like the lead homepage. But we also see other pages. Number 9 is the EPA homepage.

So we're still using the same report, the All Pages report but we're looking at data a little bit differently. By using a basic filter in the data table we just asked for a list of interactions, a list of pages that match the lead string by using just the segment for lead. We're now applying a filter at the session level and what we're asking is for people who viewed a lead page, show me all the pages that they viewed. And that's what we're looking at. All the pages viewed by anyone who viewed a lead page. So there's a slight difference there in how you use the filter and the segment.

But what's really great about having a segment for your web area is that now I can view any report, any Standard Report in Google Analytics and it now will apply to my web area and only my web area. Let's go and look at some examples. I'll go into the Audience folder, the subfolder for Geo and the Location report. This is the report that shows you the top countries where users are located when they visit the EPA website. But now that we have the lead microsite session segment applied this is just for segments - this is just for visits or sessions that included views of the lead microsite. So these are your users. And these are the countries where they're located. I can change the top left of this data table down to the city level and now get the location by city for all those lead microsite users.

Now if I switch to another report you'll notice that the segment we have will stay applied. So I'll close the Audience folder and maybe open acquisition and look under all traffic for the Referrals page or the Referrals report. And now you get the top referral pages for the lead microsite. This is not applying to all the EPA any longer. So while you cannot filter on this data table to get to your web area because the dimension is source, not page, it no longer matters because your segment is applied, and you're only looking at sessions for the lead microsite. So these referrals are for lead microsite users.

You can look at another one under Campaigns you'll find the organic key words. These are the key words entered into Google and Bing and search engines like those. And as you scroll down you're see these key words only apply to lead microsite users. These are the key words they use to reach the EPA website and they either immediately or eventually viewed the lead microsite. So it's no surprise that you see key words like lead, lead paint, lead certification. So this opens up all the reports because you now have a segment just for your web area. And it stays applied as you navigate through the different reports.

Let's look at one more. Audience, Mobile report. So, again, now you're looking at a breakdown of the percentages for mobile traffic and tablet traffic versus desktop traffic and it's only being applied to a lead microsite session. So these are your users and this is the breakdown for mobile and tablet and desktop amongst lead microsite users.

So you can see how it really opens up what you're able to view and all the different reports you're able to access just for your web area when you have a segment for your web area.

So I should point out that there are some resources that go over exactly what we've done today. Filtering -- basic filtering there's a page on our Web Guide, Google Analytics Filtering for Web Areas which goes over how you filter on the All Pages report. Here's an example for lead and how you will get the top pages in the lead microsite.

Back on the index of All Google Analytics resources on the Web Guide there's another page called Creating Web Area Segments with Google Analytics. And this will go over adding segments, creating segments step by step with screen shots for your web area again with the lead example that we went over today with the 1, 2, 3, 4 step process that's spelled out for you. So definitely use these resources when you need to have a reference as you're setting up segments.

Going back to the interface... So one more thing I want to point out is that sometimes when you're using segments especially multiple segments over a long period of time so if you're requesting data over a year, even 6 months with 2 or 3 segments applied that's when some may see Google sampling the data you're getting.

If you look in the top of the interface right next to the options for email, export, shortcut. It says that this report is based on 100% of sessions. That's different than the information you get about segments when you first log in and you see it says All Sessions where now it says Lead Microsite sessions. This 100% of sessions refers to how much data is being used. How much of the raw data is being used to provide this report. And in this case and in most cases it's 100% of sessions. So let's expand the date range quite a bit further and look at -- let's see, up to 6 months or so of data. And see what happens to the percentage of the data. Okay not bad. It's still using 100% of sessions. Let's apply an additional segment and just apply one at random. Not to say that you would need this additional segment with your lead web area segment but I want to show you that in some cases you'll find sampling and we just still don't have sampling. We have the premium version of Google Analytics so you'll get raw data most of the time. But I do want to point out what happens if you see sampling.

So let's expand the date range further. Back to March 2014. This may or may not reduce the sampling or trigger sampling but we'll soon see. And it has. So now I'm asking for nearly a year's worth of data with two segments applied and it says at the top where it used to say 100% that only 5.29% of sessions is being used to provide this data. The data is still highly accurate. It's based on statistical method and you can be sure that the trend you're seeing in the data is accurate. But if you have a need to get at the raw numbers and not everyone does, if you ever have a need to get at the raw numbers and you're asking for a lot of data over a long period of time there is a way to do that.

Under the export button, there is an option for unsampled export. And if you click on that you just -- you'll have the opportunity to name your report and request an unsampled report. It's only available in CSV or TSV for Excel. You're not going to get a PDF with this. But you will get all of the unsampled raw data this way if you see that sampling has been triggered. It will not come right away. It could take 15 minutes to an hour. You can get an email then you'll be able to download your report. I did want to point that out because when you're working with segments over long periods of time it's asking for a lot of custom data and it can trigger sampling. Not always but it's possible. If I remove the second segment we may see the sampling percentage change.

So let's remove Direct Traffic and just isolate the web area segment. Still over nearly a year's worth of time and see what happens to the sampling percentage. Still at 5.29. So it was more the amount of data than the multiple segments being used that triggered the sampling. But I did want to point that out.

So that wraps up what we were going to cover today. Just remember that there's a big difference between basic filters and segments. Basic filters are asking the data table to provide just the interactions that you define such as show me only those pages that match the lead microsite. And the difference is the segment is at the session level or the visit level where you're then asking to create a group, a segment of lead microsite users and then in all the different reports you access you're asking for all the associated data with lead microsite users. Where are they located? What devices are they using? Show me all the pages they used.

You're able to open up Google Analytics and get a lot more data by creating segments. In particular segments for your web area. So I hope you found this useful. Remember there's a lot more on the resource directory and there's a lot more on-line as well. But this key concept and the difference between basic filters and segments will definitely provide a lot of capabilities to you and allow you to access a greater amount of the functionality in the tool. So thanks for joining. Hopefully you got some of what you wanted out of this session and you can find this and other training sessions on the Web Guide.