#17 Why you should be using Segments in Universal Analytics

[00:00:17] **Dara:** Hello and thanks for joining us in The Measure Pod, a podcast for analytics enthusiastic where each week we pick an analytics topic, challenge, problem, or an opinion. We dig in and we also try to have a little bit of fun along the way. I'm Dara, I'm MD Measurelab. I'm joined as always by Dan, Lead Analytics Consultant also here at Measurelab. Hey Dan, what's new in the world of analytics?

[00:00:41] **Dan:** Hey Dara. So the first thing to talk about this week is that Google Analytics 4 have updated that default channel definitions. And this is something that has been requested for quite some time. We still can't create our own. We can't edit the default ourselves, within our own definitions but the definitions, the default definitions have been updated for us. What they've done is they're now starting to break out channels to make them a bit more granular. What they've actually done is they've introduced a number of new channels. They've introduced organic shopping, paid shopping, organic video, paid video. And then they've also got this paid other bracket. So any traffic that comes in with the word paid or CP something in the medium, it now gets attributed to one of the paid channels. If not, it defaults to a generic other paid channel, rather than going into a kind of joint other bucket. And they've also got channels for audio, SMS and mobile push notifications. They've done a couple of updates to the logic as well, I'll link off in the show notes to the updates that they've defined. One of the key ones actually just to kind of dig in a bit deeper, is YouTube. It's no longer social it's now being pulled down to his own channel of video.

[00:01:44] **Dara:** As you mentioned, we're still hoping that there's going to be an option to edit the channel groupings, create custom channel groupings like you can in Universal Analytics. But this does sound like a really good update because they have done a pretty thorough job of creating a lot of the main channels that, certainly the ones that we would be familiar with.

[00:02:03] **Dan:** Yeah, I hope they don't lose that functionality and GA4, but I feel like it's covered 99% of the bases. So maybe they're just doing this to make people stop asking for the new feature. Um, but for now, at least it's a really good step forwards. And the second thing in the news this week is a piece of Facebook news. And that is that they're actually deprecating some of the targeting options that they currently have within the ads platform. Within the, I suppose Metta now, Metta ad platform, Facebook ad platform, um, you've got this idea of detailed targeting options. So this is where you can target based on what groups they're in, what pages they've subscribed to and liked and what kind of posts they're liking. And what they're actually doing is they're removing the ability to retarget or target these users based on things that can relate to things like health, sexual orientation, religious beliefs, political beliefs, and social causes and issues as well. So this isn't saying that Facebook knows your religion and sexual orientation and all that other stuff. I'm sure they do, or at least they've got an algorithm to try and work it out. But this is they're taking the ability away from advertisers to re target or target users based on traits that could tie you into a group like that. So it feels like they've listened to some negative criticisms and they've taken some action. It doesn't feel like enough, or it doesn't feel like they're fundamentally changing too much. I'm not sure if this will dramatically change any, any advertisers targeting in any meaningful way. It feels like more of a show towards some kind of privacy step if that makes sense?

[00:03:33] **Dara:** Yeah, that's what I'm thinking. It's like a token effort to look superficially like you're making improvements, but actually it might be a small step in the grand scheme of things.

[00:03:43] **Dan:** Yeah, exactly.

[00:03:44] **Dara:** Smoke and mirrors basically.

[00:03:46] **Dan:** Yeah.

[00:03:47] **Dara:** Okay, so what's our topic this week then?

[00:03:52] **Dan:** This week, we're gonna talk about universal analytics and segments. And why segments are the most important thing in Universal Analytics.

[00:04:00] **Dara:** I'm going to pretend I don't know what the subject is and I'm, I'm surprised and excited now. But I am because segments are a favorite feature of mine within GA and always have been.

[00:04:11] **Dan:** And I think you're not alone. I mean, this is something that anyone that's used Google Analytics more than once a week. Eventually it gets to a point of thinking. I can't be bothered to try and build a custom report or dig into this report at a secondary dimensions. We end up getting to a point where we just sit in the audience overview report and just build segments because it's actually a lot quicker to go build a segment, to get the number you want, then try and find it through a UI full of 150 reports and trying to play that game as it were.

[00:04:40] **Dara:** Yeah, and I think as I'm sure we've mentioned this before. The aggregate data in GA is of little use. I mean, maybe it's used in management reporting with some very high level KPIs. But really, if you're an analyst, the high level aggregate data is of little use to you. What you want to do is segment that data down and look at different groups of users or different segments of traffic and try and analyze what's really happening on the site.

[00:05:05] **Dan:** Yeah, exactly. And I think that's a really good overview of what they are right. So what a segment is, is a way to filter your entire user and session data based on behavioral traits or something that's happened. And one of my favorite use cases for this, or one of the first ways I like to explain and teach what segments are when we're doing demos. It's actually thinking about session duration. And I think we've mentioned it on here before, but session duration I have a real issue with in Google Analytics, and that's because it's not a metric of duration. The way it's calculated is that it takes the timestamp of the first page view and the timestamp of the last page view, and it works out the difference between the two. It's a good indication of elapsed time, but it's not eyeballs on a screen. It's not engagement in any way. It's just how long from first page view to last page view within this session was there. And the issue with that is the concept of a bounce session. What bounce means is just any session that contains only one page view. And the bounce, whether you agree with the definition or not, whether you like it, whether you use it. That's I suppose a different conversation, one we probably have had many times before and we'll definitely again in the future. Um, but the, the point I'm making at the moment is that all bound sessions only have one page view. So what is the session duration there, there is none. There that there is no second page view or last page view to work out that timestamp different. So Google Analytics treats that as a zero minute session. And let's say you've got a 50% bounce rate, that means 50% of all of your sessions have a zero time on site. So when Google works out something like average session duration, the average is including 50% of them are zeros and it's pulling the average right down. A really easy way to undo that and say, I want the average session, duration of sessions that had at least one page view. So that, what we're basically saying is if you don't know, don't include it in the average calculation, the same way you'd exclude outliers. So if you don't know, or can't walk out duration, ignore it from the data set and work out the average on what's left, and you can do that super easily using a segment actually. There's no special secret way of doing that, it's something called a system segment. So every user in GA has a set of segments that are automatically built for you without you having to do anything. So the way to do that is you go into any report in Google Analytics, you click segments, you go to system, and you apply the segment that is called non bounced sessions only. And then when you look at the metric of bounce rate, it should be 0% now. And then when you look at the metric of average session duration, that will be a better representative number, based on numbers that you can calculate.

[00:07:40] **Dara:** A less wrong number. I love saying that.

[00:07:43] **Dan:** Yeah, exactly. Within the confines of the data we collect, it is a less wrong number. There's no such thing as perfect, so we'll have to make do with good. This easy way to clean up the session duration number is something that I use all the time. It's actually something I include in all my dashboards too. It's one of those kind of data cleansing things that happens behind the scenes before you represent the data in a dashboard or to the stakeholders.

[00:08:06] **Dara:** Speaking of cleaning the data, it's not possible within a system segment, but you can obviously create custom segments to clean any other unwanted data out of your report. So for example, if you've had an issue with high levels of spam traffic, or you realize that you were actually running some site monitoring software on your website or something like that. Obviously you can exclude a lot of that with a view filter, but that's not going to work retroactively. So you can use segments to historically clean up your reporting data. You might want to exclude, whether it's certain traffic sources or certain host names or whatever the case may be just to again, improve the quality of the data that you're actually feeding into your dashboards or into your reports.

[00:08:51] **Dan:** Yeah, I can't remember the last time I've done the piece of analysis over a long period of time that didn't have a segment applied, just because of that. Whenever you find an issue or a problem or a bit of data, you need to exclude, it's only from that point forward. And there's always going to be a case of, oh, can you pull the last 12 months of data? And it's like, oh, okay, well, I need to apply the segment before this day and then maybe I can use it without a segment going forward. It's always a handy tool to have just to normalize data over large date ranges, or even when data changes. Uh, you can kind of apply a sweeping statement to say, clean data for me and give me what's left.

[00:09:27] **Dara:** One risk, I guess, is our old friend sampling. If you're pulling a lot of data and you are using a segments, the same would apply with it with a second dimension, but you're increasing the likelihood of sampling being an issue. Depending on the volume of data that you've got in your GA account obviously. And then again, if you're a 360 customer, you're going to have ways around that, or at least ways of minimizing the issues there as well. But it is something to bear in mind with segments.

[00:09:53] **Dan:** Yeah, always, always. Even if you are 360, you still have sampling, it's just at a high threshold. So there's no avoiding it entirely. But, be mindful of it and I suppose just make the call. If you're doing monthly data trends over the last three years, and 5% of your data sampled, I can't imagine it's going to make too much of a difference. But if it turns out that you're using 1% of real data and the rest of it is sampled, then maybe it's a cause for alarm and you do something differently then.

[00:10:20] **Dara:** Yeah, well, you can of course download the unsampled data from GA360, but it is using a quota system and you can hit those limits as well. But it's possible, but still can be a bit of a pain. And the other thing I was thinking, so you mentioned about feeding it into dashboards so that the segments are all available within Data Studio. But you can also get access to the segments through the Google Analytics reporting API. So something that I used to love doing a lot is actually using custom segments that I've created in Google sheets using the GA add-on. And one thing I learned the hard way was if you're doing that, there's a field in the configuration sheet where you can add in the segment ID, and you can look up what that segment ID is, but those ideas are unique to each user. So if you use the IDs and then share that report with somebody else, they would need to update it with their segment ID. So you can use the segment definition instead, which gets around that issue. A little tip there if you've ever had a need to use segments with the Google Sheets add-on.

[00:11:22] **Dan:** That's a really good point. I just avoid the IDs for things like that. I just use the definitions wherever possible. It also means that you can build them on the fly, especially using the Sheets connector. You can use formula within the Google Sheet when you're making the API request to dynamically create segments as you go. And it's a really handy tool, for one you don't have to save any segments in your own interface. You don't have to go into GA to do anything. And like I say, it's dynamic. I can go into a spreadsheet, I can generate some kind of formula that creates a condition of segments and just refresh the report and it's there. It is quite intuitive and whenever you're using the Sheets add on for GA, there's a little link at the bottom of the config tab that takes you off to the documentation, and it's super easy to follow along.

[00:12:04] **Dara:** The only problem I ever hit with that was if you have a very long segment definition, there is a character limit. It's a bit of a fun challenge in a way, especially if you're using RegEx, you can kind of refine your RegEx and try and get it under the character limit. But it can be a bit of a pain if you've got a need for a pretty complicated segment, with a lot of conditions in the definition, then you can hit that limit. I think it's 256 characters, or at least it used to be,

[00:12:28] **Dan:** That's the fun Dara, try to get something under a character limit. There's gotta be edges to make it a game.

[00:12:33] **Dara:** Especially when you're under pressure, trying to hit a deadline. Nothing incentivizes you like a bit of time pressure. Another common use case for a lot of e-commerce websites where they're selling maybe different products within different product categories, often what people want to do is group sessions visiting the site into what product categories they looked at. And there's something always worth clarifying here because it is sometimes something that people get wrong, which is thinking that they're then distinct segments. But if you create a segment for each of your different product categories, if your segments are just saying include sessions where they viewed this page or these pages within this category, they may have actually viewed other pages in other categories as well. So you'll end up with segments that there will be overlap between them. Usually that's fine and people are aware of that and they know that's the case, but they still want to do that to get a sense of how much traffic kind of quote unquote is hitting those different product categories on the site. And then another one, maybe vaguely related, is where you can create segments of either users or sessions where a certain product was purchased. So that then you can use that segment to better understand those people that are buying that particular product. So if you're trying to do customer profiling, you could look at people who purchased product a and then you could look at all the demographic data, you could look at landing pages, you can look at marketing sources. To try and build up a picture of who these people are, who are actually buying that product versus a different type of product.

[00:14:07] **Dan:** That's really interesting. From an advertising perspective, segments can be pushed into Google Ads as an audience. So whatever You build as a segment can be pushed into the ad platform to directly retarget, or even find lookalikes. As your example Dara, where you find users that have purchased a specific product type. You can then push that into Google Ads to find lookalikes. In a way, prospect against similar user traits that you haven't yet to identify, use the big black box of Google to figure that stuff out, to identify those users that are most likely to continue to purchase this brand of phone or type of furniture or whatever it is that you're selling.

[00:14:41] **Dara:** You brought back some painful memories for me there. I instantly thought of the time that I had to create, I think it was 250 audiences from segments. Yeah. That was a fun, I don't even want to think about how long it took me, but it was a lovely, a lot of copy and pasting.

[00:14:57] **Dan:** I bet. I bet. Was there no way to automate it, or did you figure that out maybe as you got to the 199th segment.

[00:15:03] **Dara:** There is a way, but again, talking about time pressure, this time round unfortunately it didn't allow to go and work out using the management API, isn't it that you can use to create segments on the fly. So we did later end up looking at that, but for this particular time there was time pressure and I had to do it the old fashioned way. But I like doing things the hard way, it makes me feel like I've done something useful.

[00:15:26] **Dan:** It's that it's that classic saying isn't it of I'm sorry the email's so long, I didn't have enough time to make it shorter. Or something like that. Well, and off the back of your example around clustering product types together, so understanding how that led them to purchase a certain product. All around that is the idea of funnels and funnels come up all the time. And whether or not you like the idea of a funnel, or whether you think of it more of a pretzel or kind of spaghetti mess, funnels are useful from a visual representative standpoint. People like looking at funnels because you can put them on a graph and you can get one number out the other end of it. And the thing with Google Analytics, or at least the free version of Google Analytics, is you can't really build funnels. There's no way to build a funnel based on your definitions, based on the pages or the events or the interactions that you want to include. There is funnels, there's goal funnels and you can even do an e-commerce funnel which is based on the enhanced e-commerce data. But what if I just want to look at behavior through the homepage, product page, basket page, and that's it, there's a three steps there. Or maybe I've got a sign up form where I want to see that flow through multiple pages. And let's say I want to create it on the fly, and I want to go back in time and look at my historic data. The way to do that is actually using segments. And the thing with a segment is that there's actually two options of creating a custom segment. The first one is a condition-based segment, which is have purchased this product and purchase this product or, or purchase this product. The other type of segment you can create is a sequential segment. And this is where we can use these to build funnels. It's the same approach, you can say anything you want. A session that sees this page and this page and has purchased this product, but now you can put an order to them. So you can say, has seen this page before seeing this page before purchasing this product. And the idea with building these sequence segments is that each segment you create is actually a step of the funnel. And the way that I like to approach building funnels using segments is start with the final step. So start with the, the full funnel. So let's say I've got a four step funnel, seen page one, then page two, then page three, then page four. I'll create a sequential segment that does that and I'll call this Dan's funnel step four and I'll click save. And then I'll duplicate that segment and just remove the last condition, so now it's just has seen page one, page two, page three. And then I'll do the same over and over again, removing one step at a time and select got four segments, which is my funnel step one, my funnel step two, funnel step three and funnel step four. And each one of those, when I look at the sessions or the users, which are the one I've I've segmented, will give me the total number of users or sessions that have made it through to that part of the funnel. So then what you can do is you can work out the funnel. You can say 10,000 people were in step one. 500 people were step two, three people step three. And you can work off the drop off rates, the conversion rates, the funnel completion rate, whatever it is you want to do. And the thing that you mentioned earlier, Dara, about things like segments being available in the API means that you can pipe this funnel, this custom funnel that you've just defined and it has all the historic data in there, bar things like sampling of course. You can pump that into a Data Studio dashboard or into a Google Sheet report, whatever you want. So it's a really nice, not so easy. I want to say it's, this is an easy way of getting a funnel, but almost like the only choice you have to do funnels. But also it's your only choice to do funnels based on any kind of behavior. So it doesn't have to be super simple like pages, you can think of it as any kind of hit to the you're collecting. So events that you're collecting on certain interactions within the page, entering via a certain channel or campaign. All of the things that you're collecting in GA can be part of your funnel steps.

[00:19:01] **Dara:** And you can make it as long as you want. You could have the most ridiculously long funnel with lots of different micro steps within it, which of course you can't easily do in any other way in GA.

[00:19:13] **Dan:** I think there is a step limit in a segment. I can't remember what that number is, but it's long enough to be better than the custom funnels, for example, in GA360. But yeah, it, it's a really good way of building funnels. You can even if you want to, build segments that are the drop-offs. So have done step 1, 2, 3, but definitely not done step 4. And that can be your audience to retarget or find lookalikes or whatever you want to do as well. So it doesn't always have to be about just defining a funnel, the steps, put it into a dashboard, work out the numbers that you want. You can be quite creative and specific and workout every variation of drop-offs, succeeding through each step, building these segments, you can be left with a lot of segments, right? I mean, that's not a problem because you don't necessarily have to worry about how many segments you have. There is an arbitrary cap, but it's not something I've ever reached. But yeah, and then you can push those into Google Ads, your dashboards, reports, just use it in GA's interface in the standard reports to analyze the behavior, whatever you want to do with it, you can. Once you've got that segment, you've got your definition and just go use that number somewhere.

[00:20:15] **Dara:** Listening to you there are sparked two thoughts. One was when you mentioned pretzels and I drew a pretzel on my notebook, and it made me think I really want to eat a pretzel. And the other back to our actual topic, the other thought it sparked was when you were talking about conditions and sequences. So I mentioned earlier about creating segments based on sessions where a certain product was purchased. So that's actually something that's difficult to do, but you need to use a bit of a workaround to actually achieve that. Within GA in the segment builder, you can build user level and session level segments, but there's no hit level segments option, but you can actually force a hit-level condition within a GA segments. And you have to do that if you want to say, I want to look at sessions involving the purchase of a certain kitchen table and set of chairs, for example. Then you have to include a criteria on the product name, and also you have to say that it was purchased so unique purchases greater than zero. But to get that to work within the same hit, so to force it to be hit level, if you use the sequence option within the segment builder, this is for session level segments. Then how that actually works is it evaluates each step in the segments as hit level. So it is hit level conditions. If you say product name equals whatever, and unique purchases is greater than zero. Then it'll actually force that to be a hit level condition and then it will give you the data that you want. Something else that is not so well known is if you're using a user level segment, having two criteria within the same filter condition also forces it to be a hit level. The problem with that though, is you don't always want to use a user level segment. So sometimes that'll work just fine for you, other times you might want to have session level data. In which case, this is a nifty little hack to force that to include hit level conditions.

[00:22:18] **Dan:** That's good to know. I had no idea that second one about a user segment, treating it as a hit if you've got multiple conditions in the same filter, that's really interesting.

[00:22:27] **Dara:** Yeah, I still find it's funny. I think we've both said we use segments almost every time we use GA and I'm still finding things out about segments that I didn't previously know. Because when you get into the more intimate ones and you're combining lots of different conditions and sequences together, it's not really fully documented how all of the logic works behind the scenes. The best way to learn is try things out, they don't look right, and then you learn something new about how it works.

[00:22:55] **Dan:** The big assumption there is that Google knows how this works and could have documented it if they wanted to.

[00:23:00] **Dara:** Maybe nobody knows it's a mystery.

[00:23:02] **Dan:** Maybe nobody knows.

[00:23:03] **Dara:** Okay, I think that's been a pretty comprehensive overview of segments and some of our typical use cases. Might be something we dive into again at some point. Cause I think we could both talk indefinitely about segments.

[00:23:16] **Dan:** Well, the next one we can do is a GA4 segment one, maybe comparisons between the two. It's not a one-to-one replacement. It's a really, maybe. So this is the next topic where a segment's gone in GA4. There you go, we've got the title.

[00:23:28] **Dara:** Yeah, let's save that then let's not have any spoilers. So aside aside from creating complex GA segments, what have you been doing outside of work lately to wind down?

[00:23:42] **Dan:** Well, I can imagine it's the same thing that you might say Dara. But this weekend just gone, we had a company away day. So every year we try and get away to celebrate the end of a year. This year was kind of special because we didn't do one last year, and it was the first time a lot of us have seen each other face to face IRL. So we've grown a lot over lockdown and some, a lot of people haven't met each other face to face. So it was really, really good time at the weekend to just meet the people you speak to every day, do some bonding, do some eating and drinking as well. So how about you, Dara?

[00:24:16] **Dara:** Well, I, I mean, snap, obviously I was there too, as you said. Rather than just repeating all the things you said about how great it was to see everyone together. I would just add that you were very bashful there and you didn't mention the fact that you were one of four winners for our company values awards. So you, as you know, one the curiosity award, but we also had awards and, and I mentioned these values on a previous show, where we talked about what we look for when we're hiring at Measurelab. So we had an award ceremony and we awarded a winner for each of the four company values. So that was a nice thing to do, but aside from that, yeah, it was just seeing everybody together in one place and being out and about in London and having some nice food and a nice drinks and having a good time together.

[00:25:03] **Dan:** Yeah, it was special.

[00:25:04] **Dara:** All right, that's us for this week. As always, you can find out more about us on our website measurelab.co.uk, or you can get in touch by email podcast@measurelab.co.uk. Or just look us up on LinkedIn if you want to suggest a topic or come on the show and actually discuss it with us, which would be great. Join us next time for more analytics chit-chat. I've been Dara, joined by Dan. So it's bye from me.

[00:25:31] **Dan:** And bye from me.

[00:25:32] **Dara:** you next time.

[00:25:50] **Dara:** Easy peasy segment squeezy.