DIGITAL ANALYTICS PROGRAM

Certified Analyst Study Guide

Table of Contents

DAP Certified Analyst Study Guide

Introduction

1. Implementation Tips

2. Before you log in
    - Using Gov-Wide vs agency profiles
    - Strategizing Goals and KPIs

3. Overall structure of account
    - Benefits of 360 vs Free

4. Standard Reports
    - Using compare in the date picker
    - Applying secondary dimensions
    - Using Navigation Summary
    - Navigating different reports
    - Acquisition channel report
    - Landing page report
    - All pages report
    - Campaigns reports

5. Campaign URLs
    - What are they?
    - Why do we need them?
    - How do I create them?
    - How can I use the “Campaigns” report?

6. Regular Expressions
    - Okay, but what are they?

7. Custom Reports
    - Key Concept: Filters
    - Explorer Reports
    - Flat Table Reports
    - Map Overlay Reports
    - Funnel Reports

8. Segments and Advanced Segments

What is a segment?
What is data sampling?
How to tell if your data is being sampled
Resolving sampling
Creating a segment using hostname

9. Understanding Events
Events provided by DAP
Events Structure Hierarchy
Events Reporting in the GA Interface

10. DAP Code Optional Capabilities

Glossary
Hit:
Interaction that results in data being sent to Analytics. This includes page tracking hits, event tracking hits, and ecommerce hits.

Introduction

Welcome to the Digital Analytics Program (DAP) Certified Analyst Study Guide! Whether you're reading this as a seasoned Google Analytics (GA) user or a new user to DAP, we hope you'll find this guide useful in understanding the DAP GA account and the data it provides. We decided to create this guide and the DAP Certified Analyst Exam in order to give our users the ability to have confidence in their analytics capabilities, as well as something to show to their supervisors that they have mastered the use of the DAP Google Analytics account. We hope that users can use the Guide and Exam in their own professional development and to increase skills needed to improve the public's experience with federal government websites.

It should be noted that this guide may not provide all the answers to the exam. In order to pass the exam, users will need to draw on the knowledge provided here, our recorded trainings, experience using the platform, and maybe even a little Googling.

The exam is not intended to be easy. It's not designed to be a box one can check without work. In contrast, we have intentionally created exam questions that draw on advanced knowledge of the platform. When you pass the exam, you should be proud of your accomplishment! And if you don't pass the first time, you can study a bit more and take it again!

Best of luck and happy analyzing,

Tim Lowden
Program Manager, DAP

1. Implementation Tips

DAP provides your development team with a wealth of resources to configure DAP Google Analytics. We aim to guide you through the path of least resistance in implementing as quickly as possible.

Here are links to our resources:
Github repo with implementation instructions & code
DigitalGov implementation page
DAP FAQs
Quick Implementation Guide
DAP Code Capabilities

Before and after implementation, we ask that you keep the lines of communication open and communicate status updates when you have questions or push your final implementation to production. We have additional steps on our end to complete your implementation and you may not have accurate reporting if you do not notify us after completion.

Step 1: Preparation and Strategizing
Reach out to dap@support.digitalgov.gov for more specific guidance on whether you will need to implement a special parallel account (PUA) parameter. This is a crucial first step, as there are a number of factors that will influence how you need to implement your code, such as with PUAs for new agencies, and different enhancements.

Step 2: Implementation
Simply copy and paste the full HTML DAP snippet code below (taken from our Github repo) into every page on your site, or in a tag manager, such as Google Tag Manager. Please note that the DAP team does not currently configure any agency tag manager containers ourselves, so the maintenance of a tag manager container will be incumbent on the implementing agency..

Please note that you **should not attempt to type the DAP script tag by hand**, as any mistake may cause the snippet not to function correctly. Please note that you will replace “AGENCY” with the acronym for your specific agency. We recommend you place the HTML snippet in the header of your site, just before **</head>**

```
<!-- We participate in the US government's analytics program. See the data at
analytics.usa.gov. -->
<script async type="text/javascript"
src="https://dap.digitalgov.gov/Universal-Federated-Analytics-Min.js?agency=AGENCY"
id="_fed_an_ua_tag"></script>
```

This snippet references the most current version of DAP GA code and automatically collects important data that wouldn't be recorded in default GA accounts, such as Events for downloads, link clicks etc. Please see our code capabilities document on the various optional features you can implement. Reach out to dap@support.digitalgov.gov for any questions as they arise.

You may also host the JavaScript files locally, however **we encourage implementation using the snippet above**. Using the snippet above ensures you receive the latest updates and features -- automatically without having to make any code changes. If your developer tries to implement the JavaScript files manually, you will need to update these files and code manually **each time** there is a code release or version change.

How to implement with GTM
You need to set up your own independently managed Google Tag Manager account if you wish to use GTM.

Please see pages 3-5 in our Quick Implementation Guide.

Step 3: QA
Once you've pushed to production, contact dap@support.digitalgov.gov to notify us. The DAP team conducts QA on your implementation and tests to see that data flows into Google Analytics.

Where to find the code, directions, etc

Github repo with implementation instructions & code
DigitalGov implementation page
DAP FAQs
Quick Implementation Guide
DAP Code Capabilities

How to check a base implementation

Your developer needs to ensure that the HTML snippet they included inline or in GTM matches the following:

```
<!-- We participate in the US government's analytics program. See the data at analytics.usa.gov. -->
<script async type="text/javascript" src="https://dap.digitalgov.gov/Universal-Federated-Analytics-Min.js?agency=AGENCY"
```

```
id="_fed_an_ua_tag"></script>
```

Troubleshooting

1. Use https://www.diffchecker.com/ to compare the original DAP <script> tag with the snippet your team implemented on your site. Any customizations you make to the snippet will obviously not match the original snippet DAP provides, but it will help you identify if your code has any syntax or spelling errors.

2. You can always contact us at dap@support.digitalgov.gov.

3. If you're interested in checking your basic implementation, download Google Tag Assistant for Chrome browsers. This is not available for Internet Explorer or Firefox browsers.

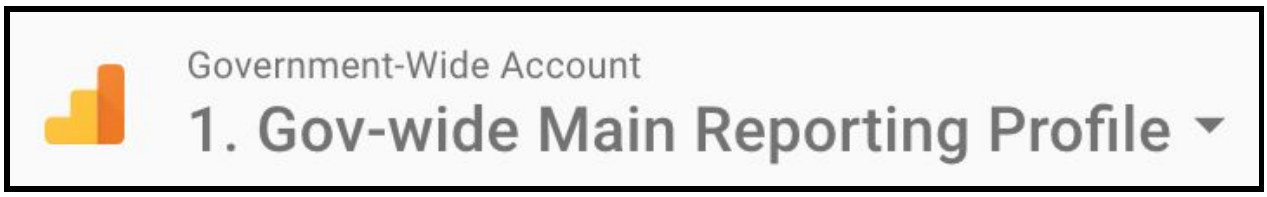
   In Google Tag Assistant, click the Chrome extension icon and look for UA-33523145-1 or UA-33523145-2. If you see the ID, it confirms that your implementation is correctly sending data to DAP Google Analytics. If you don't see the ID below, this means you have not implemented DAP correctly. In this case, contact dap@support.digitalgov.gov.

2. Before you log in

Many DAP users experience confusion upon first logging into DAP's Google Analytics. Please remember that if you are already a user of a Google Analytics account (typically independently managed by your agency), you'll need to make sure you're selecting the DAP account -- to do

this, toggle and select the DAP version of Google Analytics. Remember that DAP is a completely separate account from your agency account, if you have one.

When logging in, ensure you see the following in the upper left hand corner of your screen. This is how you confirm that you're using the DAP Google Analytics account



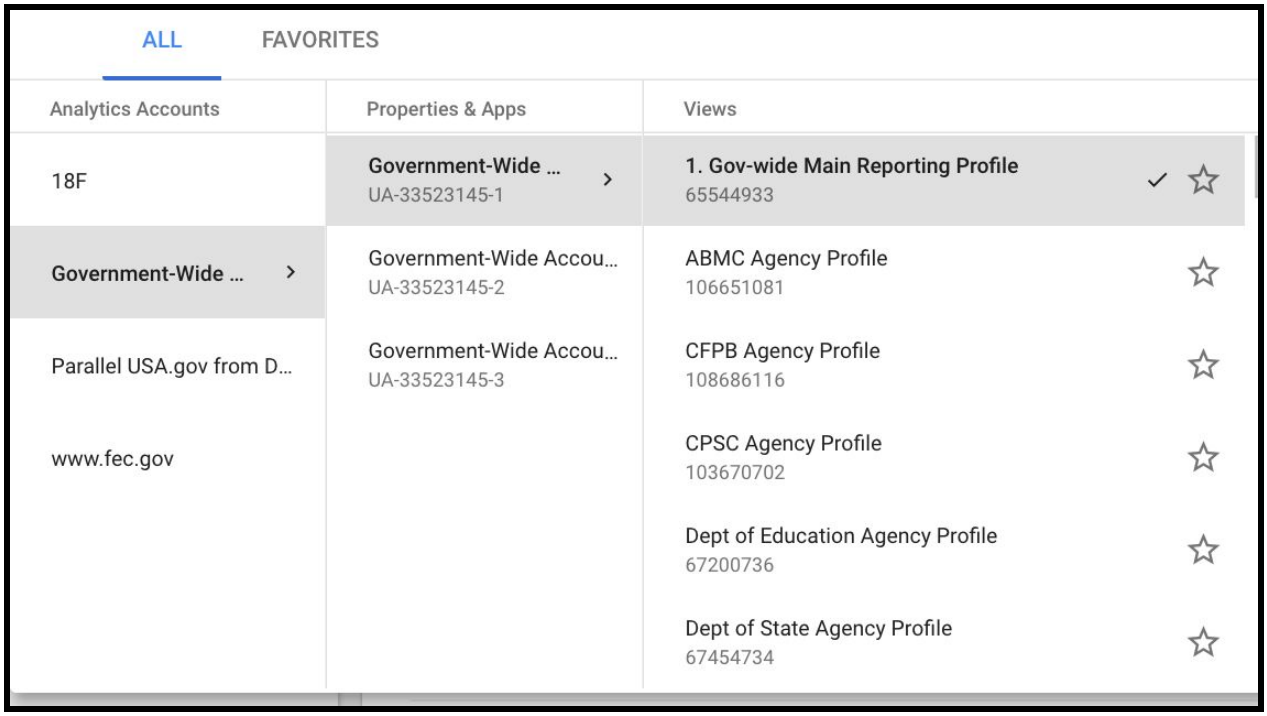
If you look at your upper left hand screen and don't see the above account, you are not in the DAP account. If you cannot find the above account when you click on that menu, it means you don't have access to DAP or may be logged in with a different email address. In this case, email dap@support.digitalgov.gov to request access.

Using Gov-Wide vs agency profiles

DAP profile structure is set up as follows:

Gov-wide Main Reporting Profile -- This is a roll-up account of all Government agency websites. Typically, this is **not** a view you'd use to for reporting. Because this view contains data for all participating sites, the reports are much more likely to contain sampled data.

Agency profiles -- Use agency profiles to create report data for an individual website. Each site should be contained within the agency profile that corresponds to it. If you work on websites across multiple agencies, you would need to know which agency manages the site, and then navigate to that agency profile to find the specific site.



Strategizing Goals and KPIs

You won't know which reports to create unless you know what business questions you're trying to answer. We highly recommend before attempting to pull any reporting that you define from a high level:

1. What your agency is trying to achieve? Don't think about your website at this point, just about your agency's mission.
2. How does your website help achieve the agency's objectives? What URLs are your most important? What forms or content are most important?
3. What questions are you trying to answer? It's helpful to think about what your stakeholders ask on a regular basis.
4. After you've completed these 3 steps, you're now ready to login to Google Analytics and practice creating reports.

Try to go in answering a question rather than trying to find something from scratch. Don't think in terms of which metrics to pull or how to find a specific website in a profile, but rather think about what questions you want to answer. **Your questions are the most important step in analytics** and will guide you in how and what you pull from DAP. If you contact DAP support about reporting, please know that we will always ask you what questions you're trying to answer.

3. Overall structure of account

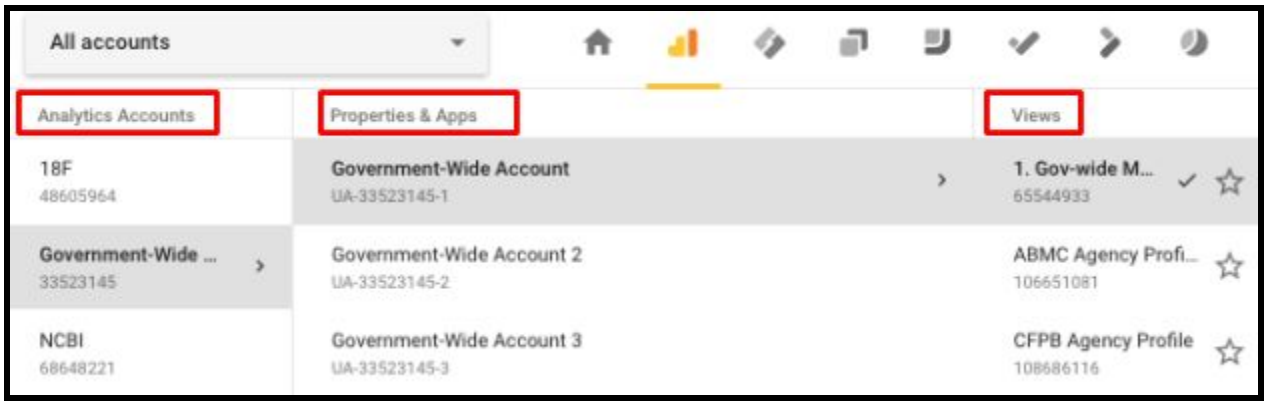
Benefits of 360 vs Free

DAP uses Google Analytics 360, which is the premium version of Google Analytics. 360 has a number of advantages over the free version. Among them:

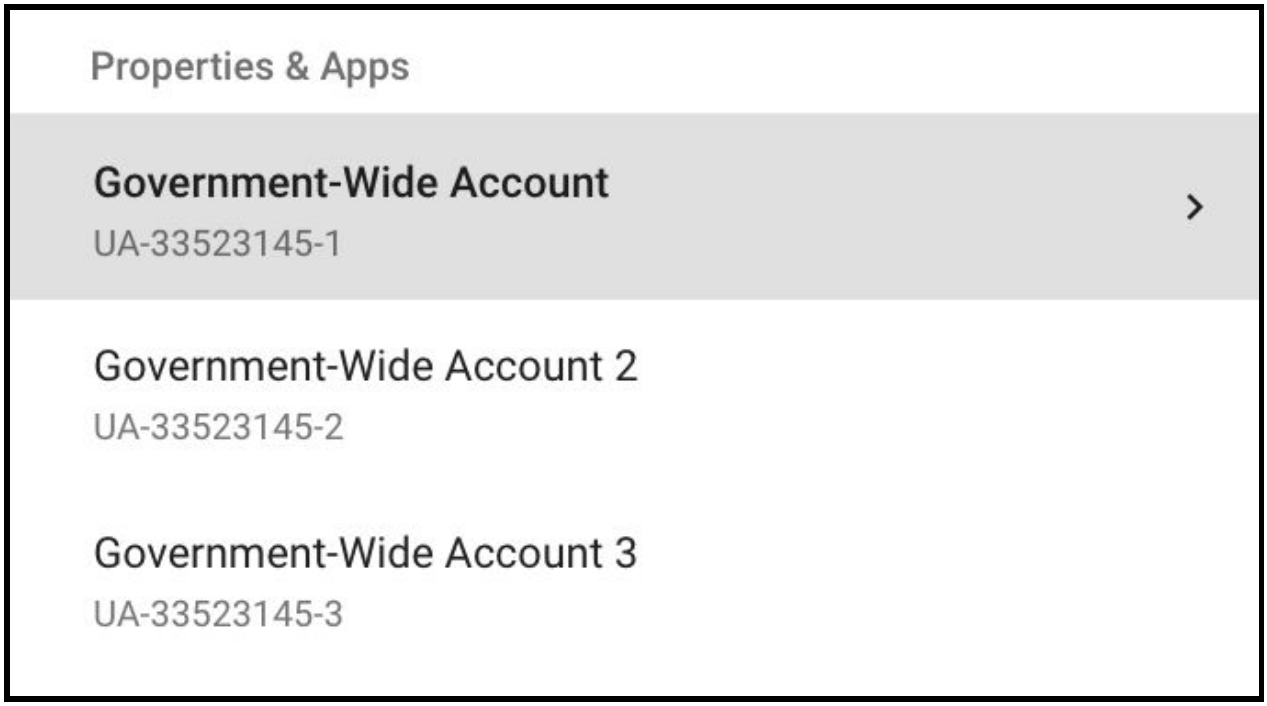
- Unlimited hits (pageviews + events) per month. Free version limits your site to 10 million hits, potentially resulting in lost data.
- More customized dimensions & metrics
- Improved reporting and sampling thresholds
- Option of downloading an unsampled report when sampling occurs (not available in the free version)

All sites for profile in the view
When first logging into DAP, you will have three properties to choose from (see below). More than likely, your agency view lives in the first or second properties. If you do not see these properties, you do not have access to DAP. Contact us to request access.

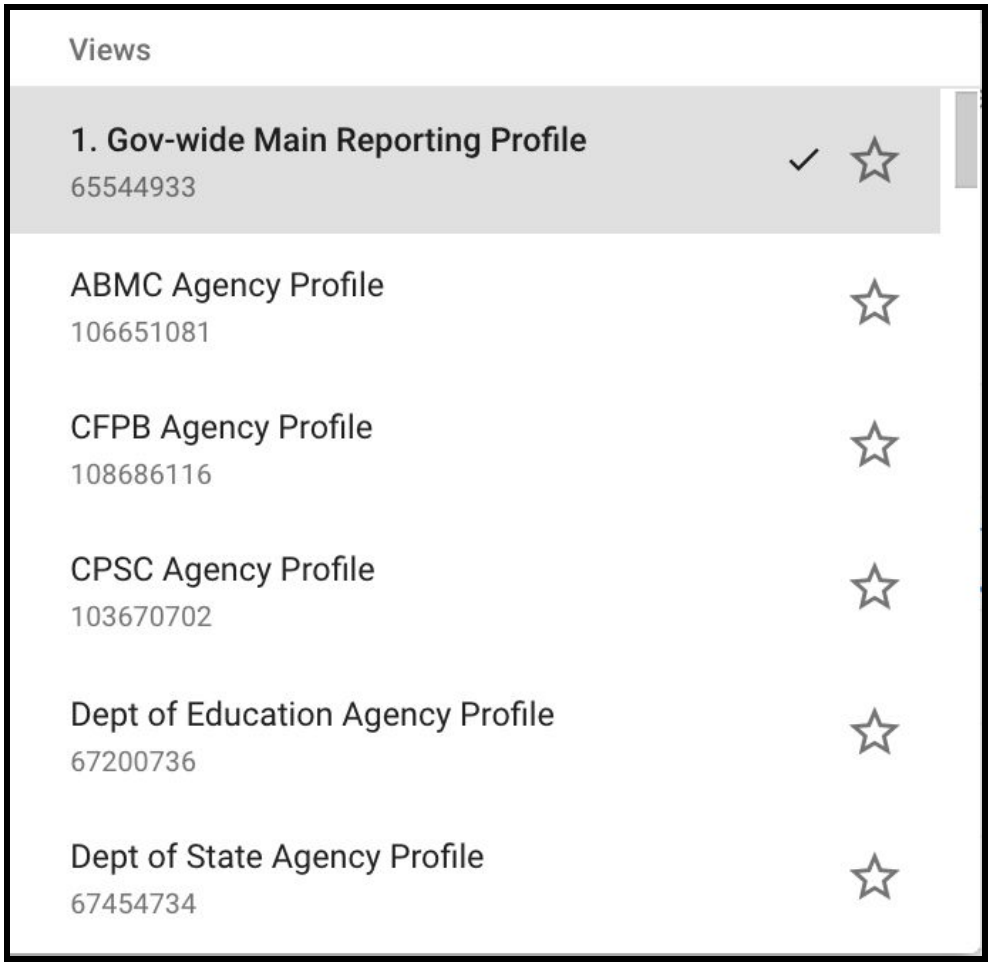
Google Analytics organizes the Administration section of the tool using an Account/Property/View hierarchy. Google visually displays this to you when selecting an account, as seen below. An account is the top-level of the hierarchy. DAP's account is "Government-Wide Account". A property is a component of your account that houses your site's data. Any site tagged with the same Property ID is collected and stored here. A view is a subset a property which has its own own unique configuration settings and subsets of data. DAP uses views to organize data by agency. Please note that the DAP team uses "View" and "Profile" interchangeably (this is because Google used to call them "profiles", but later changed the terminology). You'll notice that the "Views" are actually labeled "Profiles" in the naming convention, such as "ABMC Agency Profile".



The most effective way to find data on a specific website is by navigating to the agency profile that contains the site you'd like data for. Depending on how long your agency has been participating in DAP, your agency profile may be in Property 1 or Property 2. If you click on one of the properties, you can scroll through the agency views in each property.



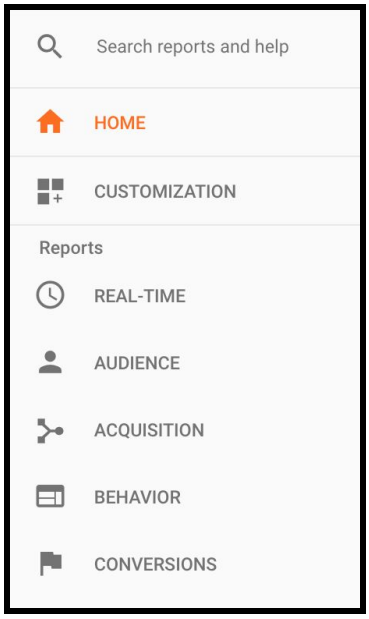
After selecting property 1 or 2, you will see a list of agencies.



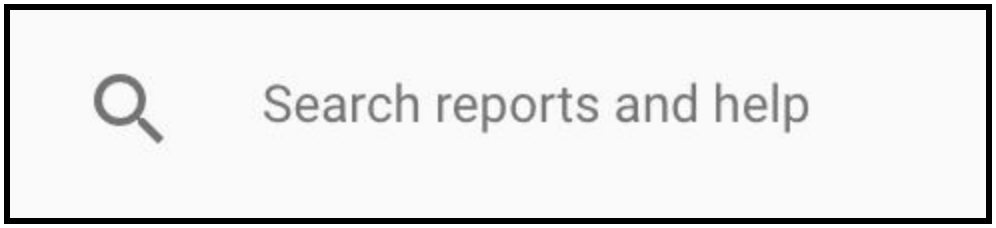
4. Standard Reports

Google Analytics provides standard, out of the box reports to answer your most commonly asked questions. Standard reports are a great place to start learning about analytics and all the types of data Google Analytics 360 captures. We recommend you first review all the standard reports and the various types of data. Google does a great job of showing you what types of data and how it's formatted in standard reports, which you can use as a reference for when you create custom reports. You should always refer to standard reports, before creating a custom report.

Click and explore this section to see all standard, out of the box reporting. We often see people jump into custom reporting before understanding standard reporting, which can cause confusion. Standard reports help you understand which dimensions and metrics you should pull using custom reports.

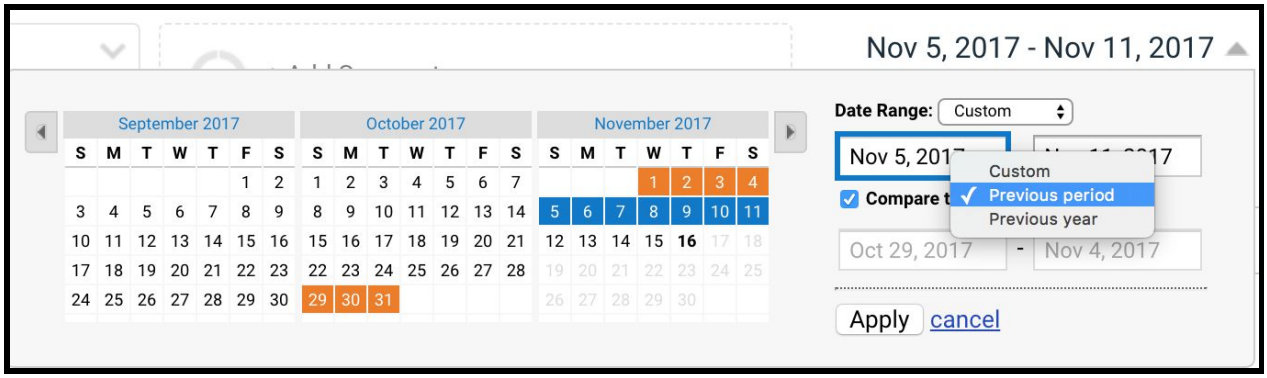


Before contacting DAP Support, we recommend you use the search box (screenshot below), to type in different types of reports you might be looking for. Think of this search box like a specialized version of Google Search, but geared towards analytics. For example, if you want to find out which marketing channels drove the most traffic to a page on your website, simply type “Marketing” into the search box. Google will lead you to the report.



Using compare in the date picker

Navigate to the upper right hand corner of any report screen and you’ll see the Date Picker. The Date Picker allows you to very quickly change the date range of a report. Not only does it let you toggle easily between date ranges, but once you select a date range, it empowers you to automatically compare the currently selected date range to the preceding date range. If you choose 1 week, it knows to automatically compare to the previous week. If you choose a month, Google knows to compare it to the previous month. Note that the longer you expand a date range, the more likely you’ll see sampled data in the report.



Navigating reports

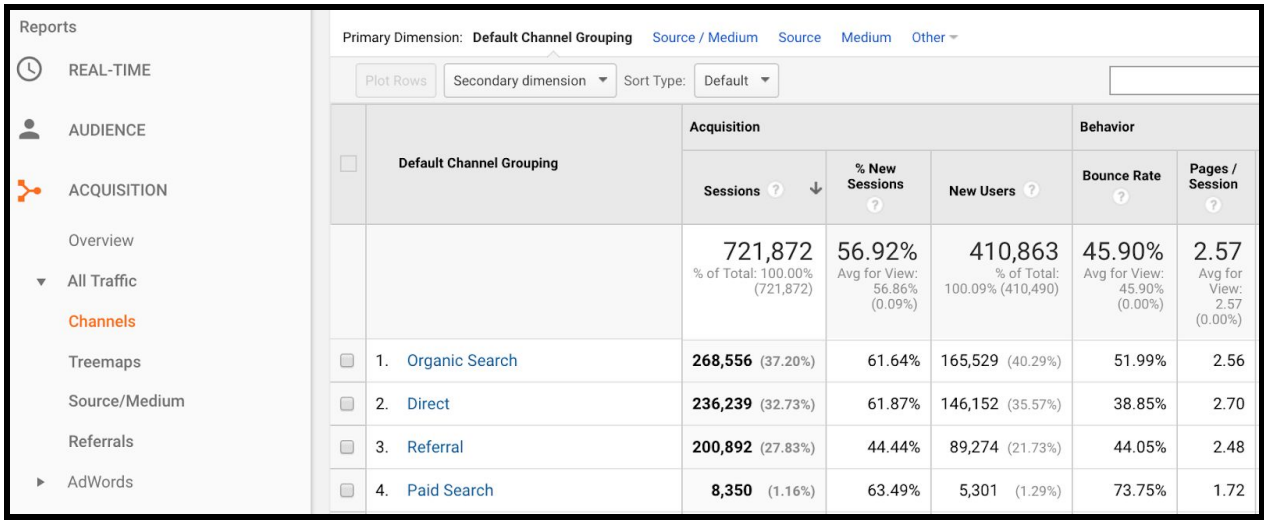
We recommend you explore the standard reports and be accustomed to how Google Analytics works to present data. For example, when setting a date range in a standard report, you can easily switch from one report to the other, and the date range will remain in place. If you apply a segment (for example - segmenting only mobile phone traffic), you can set the segment and navigate to any other standard report and see only mobile phone data. Exploring and engaging with the tool will help you better familiarize yourself with all that Google Analytics has to offer.

Below, we'll review a few of the more popular standard reports, but this list is by no means exhaustive. The best way to get to know all the reports is to view them in the interface.

Acquisition channel report

Acquisition >> All Traffic >> Channels

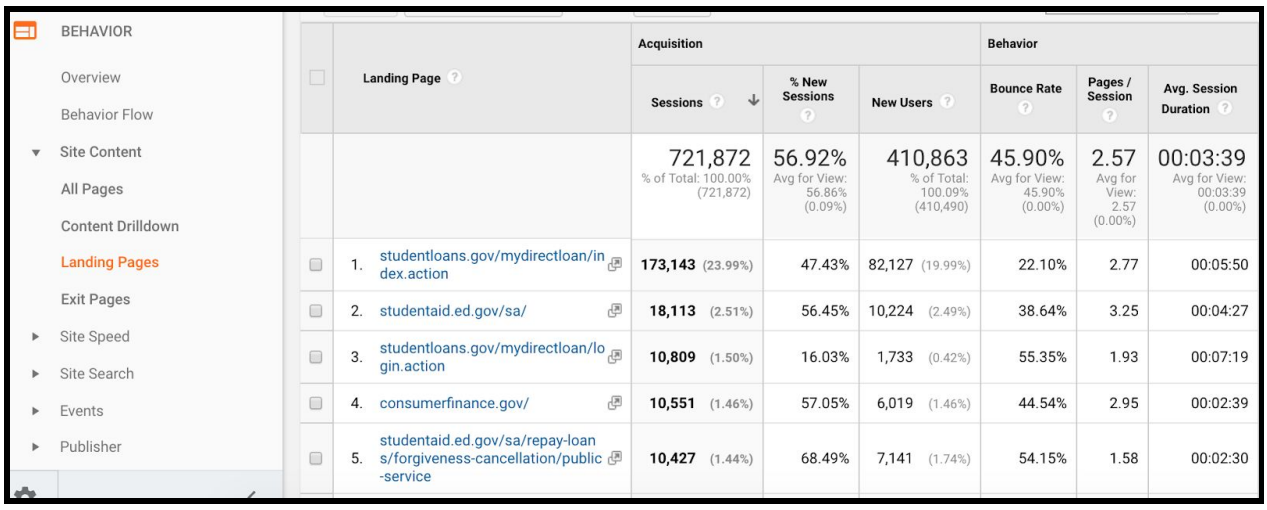
Channel reporting shows you which marketing efforts are driving the most traffic to your site. You will typically see Organic search, Direct, Referral, Paid Search, Email, Social and Other traffic appearing here. Google Analytics identifies where traffic is coming from, and lumps it into various buckets to help you understand what channels are most important to your site. You can use this data to understand how well an email campaign performed (you can also check the "Campaigns" report for this data), or understand your site's health in organic search results.



Landing page report

Behavior >> Site Content >> Landing Pages

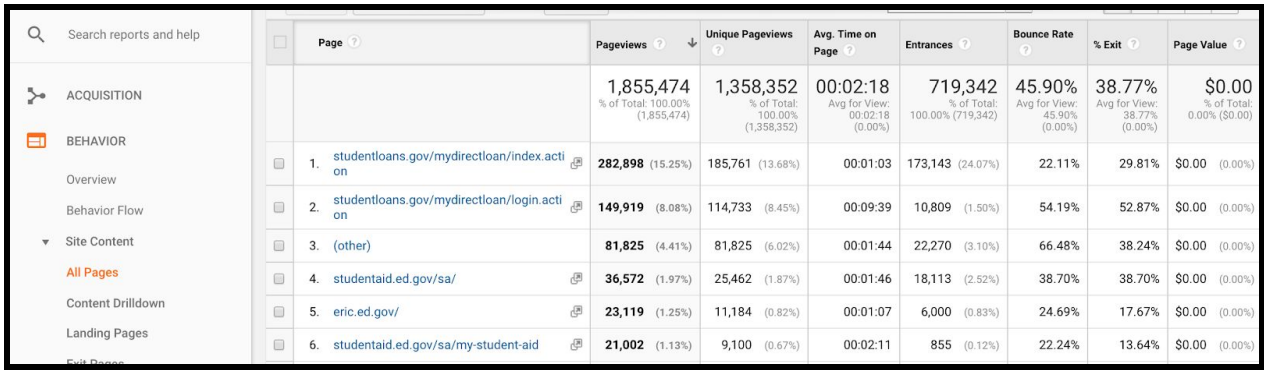
The landing page report shows you pages where users first “land” on a specific page on your site. Using the landing page denotes that a session starts on that page, so you can assume that the landing page is your users’ first touch point. You can use this report to identify the performance of your most important pages in your marketing campaigns.



All pages report

Behavior >> Site Content >> All Pages

The all pages report gives you an overview of topline metrics on your site. This report is likely the most used report to understand engagement on your site. Use this report to identify engagement metrics that help give you insight into how long a user stays on your pages and then leaves, for example.

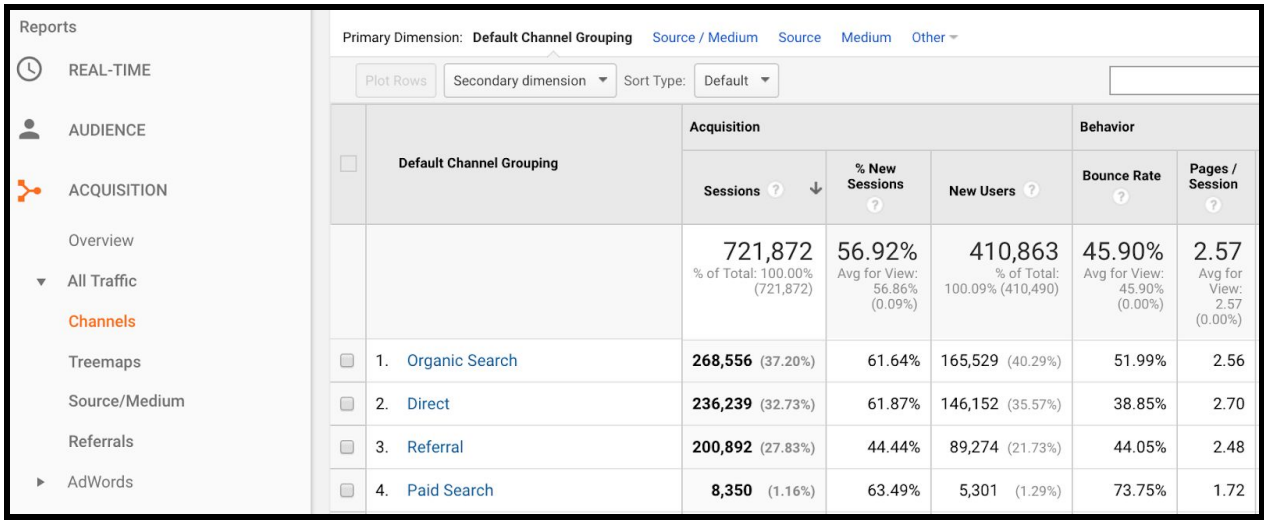


Page	Pageviews	Unique Pageviews	Avg. Time on Page	Entrances	Bounce Rate	% Exit	Page Value
	1,855,474 % of Total: 100.00% (1,855,474)	1,358,352 % of Total: 100.00% (1,358,352)	00:02:18 Avg for View: 00:02:18 (0.00%)	719,342 % of Total: 100.00% (719,342)	45.90% Avg for View: 45.90% (0.00%)	38.77% Avg for View: 38.77% (0.00%)	$0.00 % of Total: 0.00% ($0.00)
1. studentloans.gov/mydirectloan/index.action	282,898 (15.25%)	185,761 (13.68%)	00:01:03	173,143 (24.07%)	22.11%	29.81%	$0.00 (0.00%)
2. studentloans.gov/mydirectloan/login.action	149,919 (8.08%)	114,733 (8.45%)	00:09:39	10,809 (1.50%)	54.19%	52.87%	$0.00 (0.00%)
3. (other)	81,825 (4.41%)	81,825 (6.02%)	00:01:44	22,270 (3.10%)	66.48%	38.24%	$0.00 (0.00%)
4. studentaid.ed.gov/sa/	36,572 (1.97%)	25,462 (1.87%)	00:01:46	18,113 (2.52%)	38.70%	38.70%	$0.00 (0.00%)
5. eric.ed.gov/	23,119 (1.25%)	11,184 (0.82%)	00:01:07	6,000 (0.83%)	24.69%	17.67%	$0.00 (0.00%)
6. studentaid.ed.gov/sa/my-student-aid	21,002 (1.13%)	9,100 (0.67%)	00:02:11	855 (0.12%)	22.24%	13.64%	$0.00 (0.00%)

Applying secondary dimensions

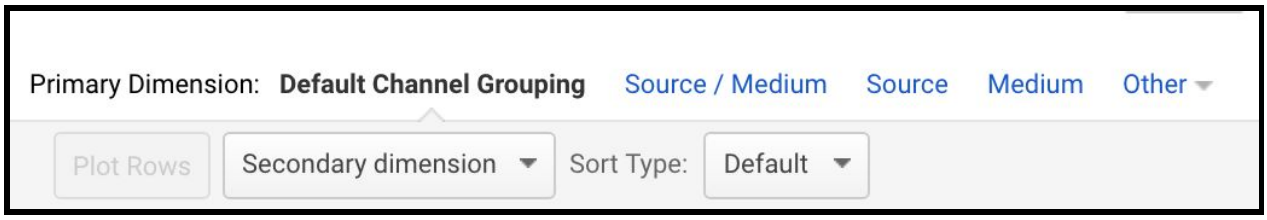
Google empowers you to select a “secondary dimension” in many standard reports. A secondary dimension is an extra data descriptor that helps provide additional context. For example, let’s say you want to understand which marketing channels are driving traffic to your site.

Out of the box, the “Channel” report shows you all of your marketing channel performance data. A dimension is an attribute, or description, of your data. For example, if you search for “Channel” and select the report to look at what marketing channels are driving traffic to your site, you’ll see the following dimensions:

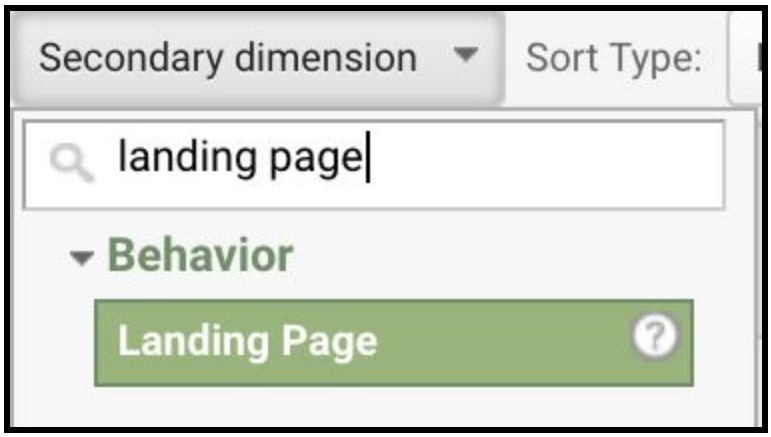


Default Channel Grouping	Acquisition			Behavior	
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session
	721,872 % of Total: 100.00% (721,872)	56.92% Avg for View: 56.86% (0.09%)	410,863 % of Total: 100.09% (410,490)	45.90% Avg for View: 45.90% (0.00%)	2.57 Avg for View: 2.57 (0.00%)
1. Organic Search	268,556 (37.20%)	61.64%	165,529 (40.29%)	51.99%	2.56
2. Direct	236,239 (32.73%)	61.87%	146,152 (35.57%)	38.85%	2.70
3. Referral	200,892 (27.83%)	44.44%	89,274 (21.73%)	44.05%	2.48
4. Paid Search	8,350 (1.16%)	63.49%	5,301 (1.29%)	73.75%	1.72

In the above example, Default Channel Grouping refers to the marketing channel. (Hint, any row that begins with a word in blue means that you can click to “drill down” further). Underneath you’ll see “Secondary dimension”. This gives you the option of selecting a second descriptor of your data. You would simply click on “Secondary Dimension” and choose the most logical dimension. In this example, you would typically want to look at what type of marketing efforts are driving users to specific pages on your site.



The “landing page” dimension describes what page users land on from a marketing channel. Type landing page into the search box and select the dimension in green. Click apply.

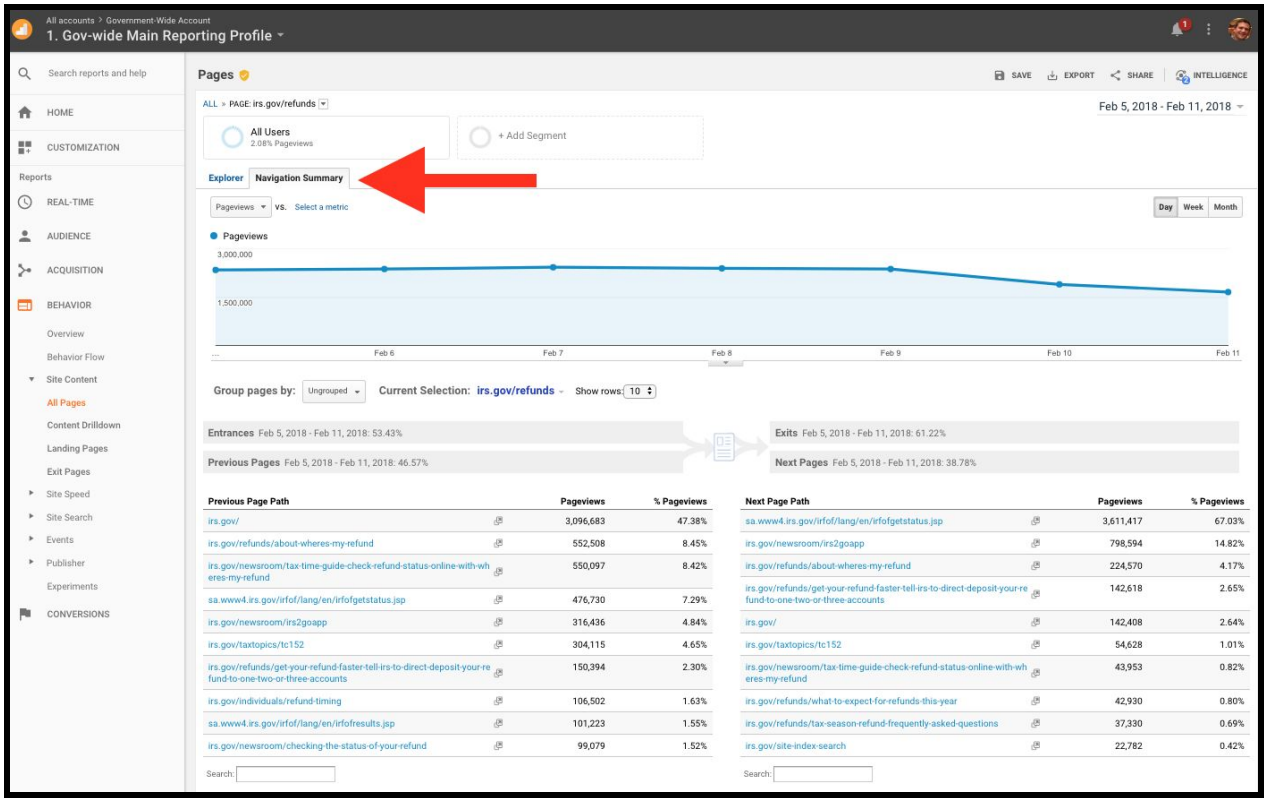


This returns marketing channels driving traffic AND the landing page users land on in one view.

Default Channel Grouping	Landing Page	Acquisition		
		Sessions ↓	% New Sessions	New Users
		721,872 % of Total: 100.00% (721,872)	56.92% Avg for View: 56.86% (0.09%)	410,863 % of Total: 100.09% (410,490)
1. Direct	studentloans.gov/mydirectloan/index.action	**89,045** (12.34%)	55.70%	49,597 (12.07%)
2. Referral	studentloans.gov/mydirectloan/index.action	**58,739** (8.14%)	39.13%	22,986 (5.59%)
3. Organic Search	studentloans.gov/mydirectloan/index.action	**25,177** (3.49%)	37.46%	9,432 (2.30%)
4. Direct	studentaid.ed.gov/sa/	**7,820** (1.08%)	78.57%	6,144 (1.50%)
5. Organic Search	nces.ed.gov/nceskids/createagraph/	**7,263** (1.01%)	82.54%	5,995 (1.46%)
6. Direct	consumerfinance.gov/ask-cfpb/what-is-the-best-way-to-negotiate-a-settlement-with-a-debt-collector-en-1447/	**6,174** (0.86%)	92.39%	5,704 (1.39%)
7. Referral	studentaid.ed.gov/sa/	**5,295** (0.73%)	36.22%	1,918 (0.47%)

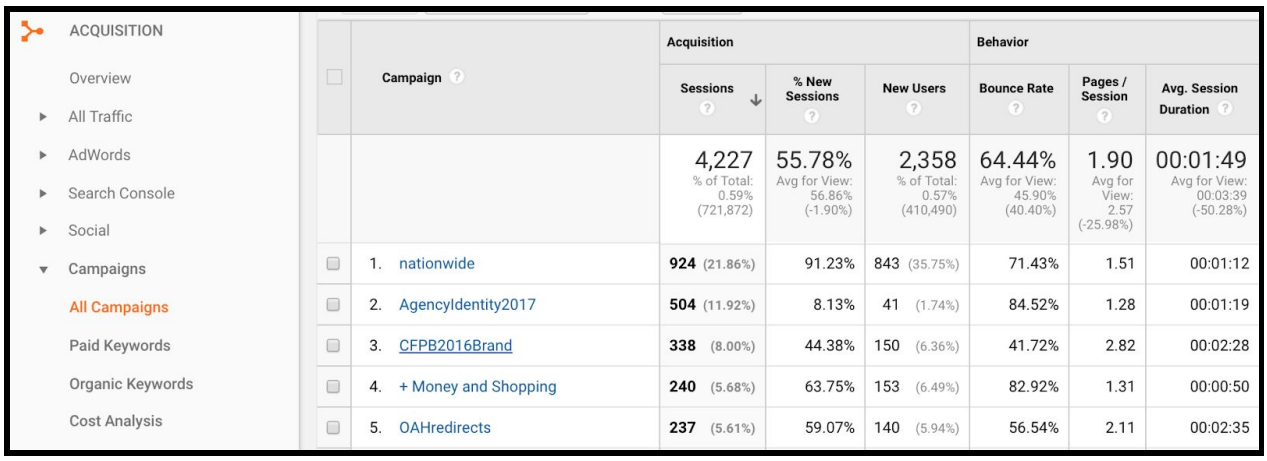
Using Navigation Summary

Navigation summary reporting is one of Google Analytics best kept secrets. To find Navigation Summary reporting, navigate to Behavior —> Site Content —> All Pages. Pick a specific page and click on it to isolate it. Then click on “Navigation Summary”, next to the “Explorer” tab. You must specify which page you want to understand traffic flowing to, and flowing from. Click on “Current Selection” to search for the page in question.



Campaigns reports

Acquisition >> Campaigns >> All Campaigns



See the next chapter for more information on campaign reporting!

5. Campaign URLs

Campaign URLs, also referred to by some as “URLs with UTM codes” or “campaign tags”, act as a way for you to include specialized links in your outreach in order to properly measure the amount of traffic those links bring to your site.

What are they?

Campaign URLs allow you to input query parameters at the end of a link that tell Google Analytics to record certain pieces of data. Usually, you set campaign parameters to record the source, medium, and campaign name for a link you include in outreach, which might be social media posts, emails, or newsletters, for example.

Query parameters are anything that follows a “?” in a URL. They aren’t used only for Google Analytics--many search configurations and other tools also use different query parameters to record certain pieces of data. Regardless, adding query parameters to a URL, no matter what they are, will not change the page the visitor sees. For example:

https://www.gsa.gov/

takes you to the same place as

https://www.gsa.gov/?WeCanWriteWhateverWeWantHereAndItWontMatter

The most common query parameters that are specific to Google Analytics are:

utm_source
utm_medium
utm_name

By setting the parameters with a “=” and separating them with “&”, you can tell Google Analytics to record the data you set. Like:

https://www.gsa.gov/?utm_source=Twitter&utm_medium=Social&utm_name=DAPcert

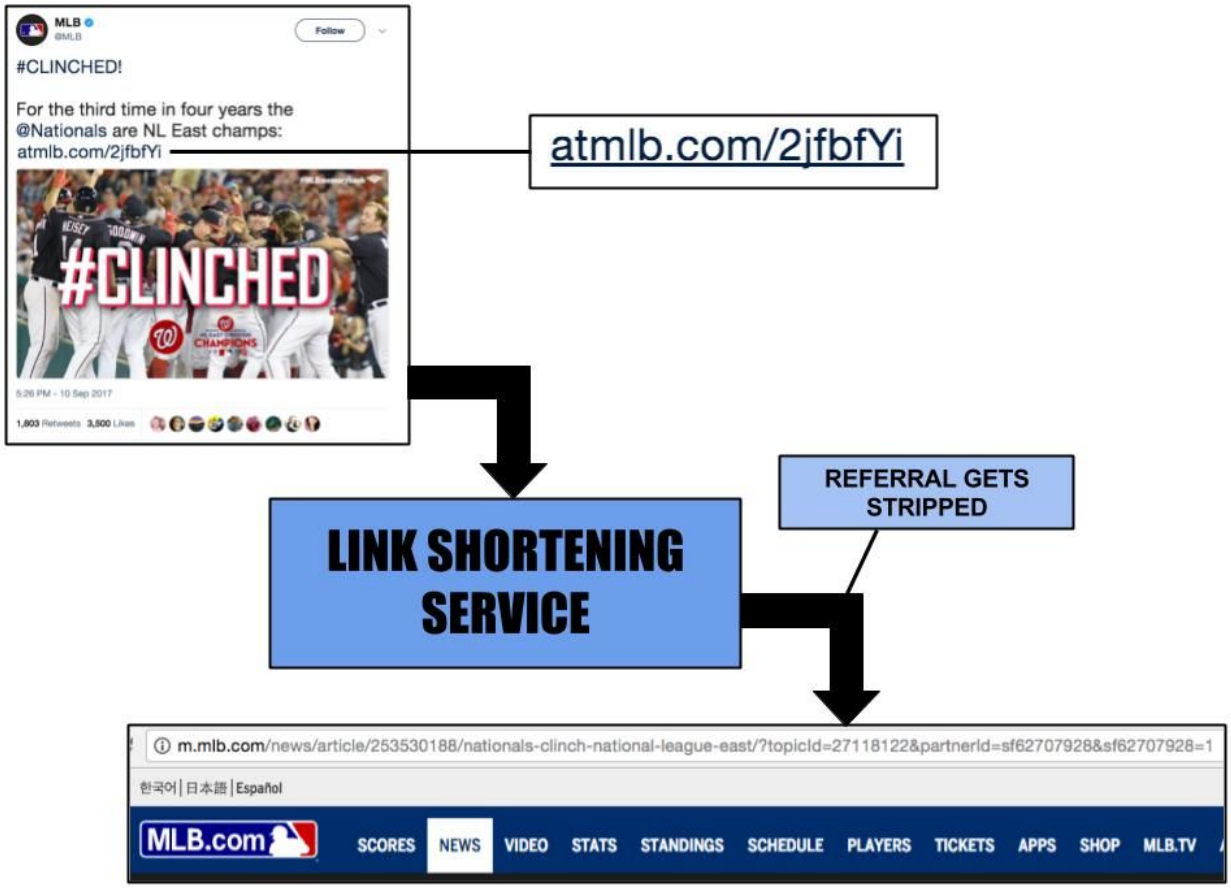
This will tell GA the source of the click was on Twitter, which is a Social network, and it will associate the session with a campaign name “DAPcert”. You’ll see how this works in GA in a bit.

Why do we need them?

The most important reason we need campaign URLs is that there are various situations across the web where we might actually lose the referrer information (the data that tells you where the user on your site came from). I'll address a few here:

1. **Link shorteners:** most link shorteners (with t.co being a notable exception) strip the referrer out of the data that a user brings when arriving at a site. So, without a campaign URL, a user click on a shortened link in a tweet (for example) will not record that traffic as coming from Twitter when the user arrives at the site. Instead, if you arrive with no referral information, the session will be attributed to "direct" traffic. See our blog post for more. We also spoke about it in our "Knowing Your Reporting Pitfalls"(11:18) webinar.

   Adding campaign URL parameters to your link BEFORE you shorten it can resolve this problem. GA will not have referral information by default, but it WILL take the information for source and medium you set in your campaign URL and overwrite the default data, plus it will also attribute a campaign name to the session, if you set one.



2. **HTTPS to HTTP links:** Any HTTPS site that has an HTTP link to another will strip the referrer when clicked. So, if your site is HTTP and an HTTPS site links to yours, GA will count all traffic coming from that link as “direct”, rather than a referral from the site that contained the link.

   Additionally, even if your site is HTTPS, if **the link** on the other site is HTTP, the referrer will be stripped. We see this a lot with sites that have recently migrated to HTTPS, but all the links around the web are still to the HTTP site. Even if your site forces HTTPS with a redirect from HTTP, it’s the URL in the link that matters. So, one thing you can do is look at your referral information from before the switch to HTTPS, and contact the major referring sites to ask them to update the link to your new HTTPS site. HTTPS >> HTTPS traffic should carry the referrer, as should HTTP >> HTTP and HTTP >> HTTPS. It’s only the HTTPS >> HTTP situation that causes this problem.

   We also discuss this problem in our “Knowing Your Reporting Pitfalls”(24:40) webinar.

3. **Email clients:** Without having a campaign parameter set to tell Google Analytics traffic is coming from an email, a click on a link in a webmail email client will appear as referral traffic. So if you send a newsletter out to 1,000 people and get 500 visits to your site that day, you may miss that much of that traffic was driven by the newsletter. Properly configured campaign URLs make sure you know what percentage of traffic is coming from your email outreach.

How do I create them?

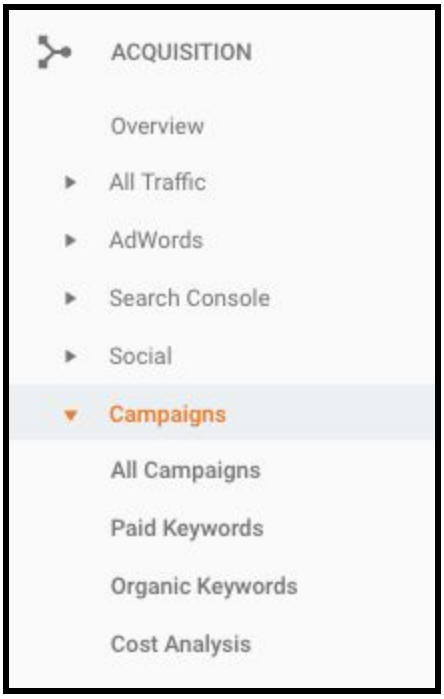
Google offers a very handy Campaign URL Builder tool to help you create the URLs. You can also shorten the link right there, if you choose. As long as the campaign parameters appear in the URL before your shorten it, GA will record the data correctly. You don’t actually need the URL Builder; once you get the hang of it, you can add in the parameters yourself.

But watch out! Campaign tagging is not something you should do without strategy. If people are tagging URLs with inconsistent naming conventions, you could make a mess of the data. Additionally, campaign parameters are case sensitive, so make sure you have standards for tagging, and that all the folks doing outreach are on the same page about the methodology and tags you are using. You may choose to use either all lowercase or capitalize the first letter of each parameter value, but you need to ensure any future tagging aligns with your original approach. Otherwise, you may run into a situation where your Campaign Medium for one marketing tactic is spread across 2 lines. For example, if you were to use “Email” and “email”, this would result in two separate lines of reporting. You’d have to export the data and combine manually in a spreadsheet.

Also, always create unique Campaign URLs for each platform you'll be using. Be careful not to put a campaign URL marked "utm_source=Facebook" on Twitter or in an email. Whatever you have in the parameter will be recorded in GA, regardless of where the link actually lived, because campaign parameters overwrite any default information sent.

How can I use the "Campaigns" report?

The Campaigns report is accessed under "Acquisition >> Campaigns" in the left hand report navigation bar.

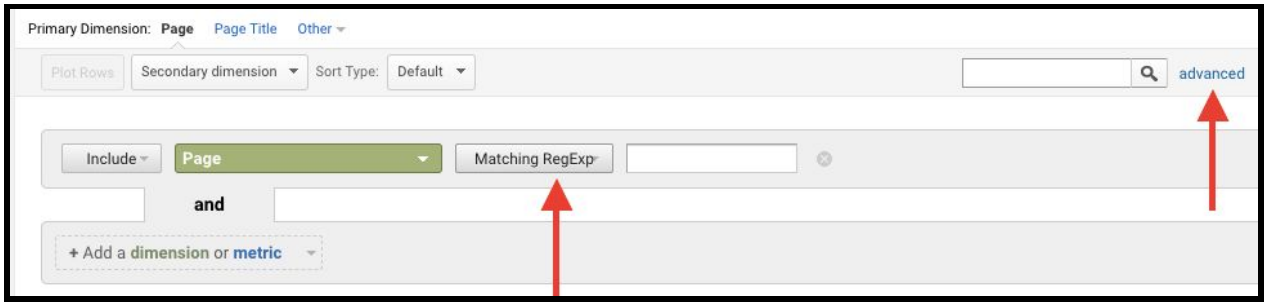


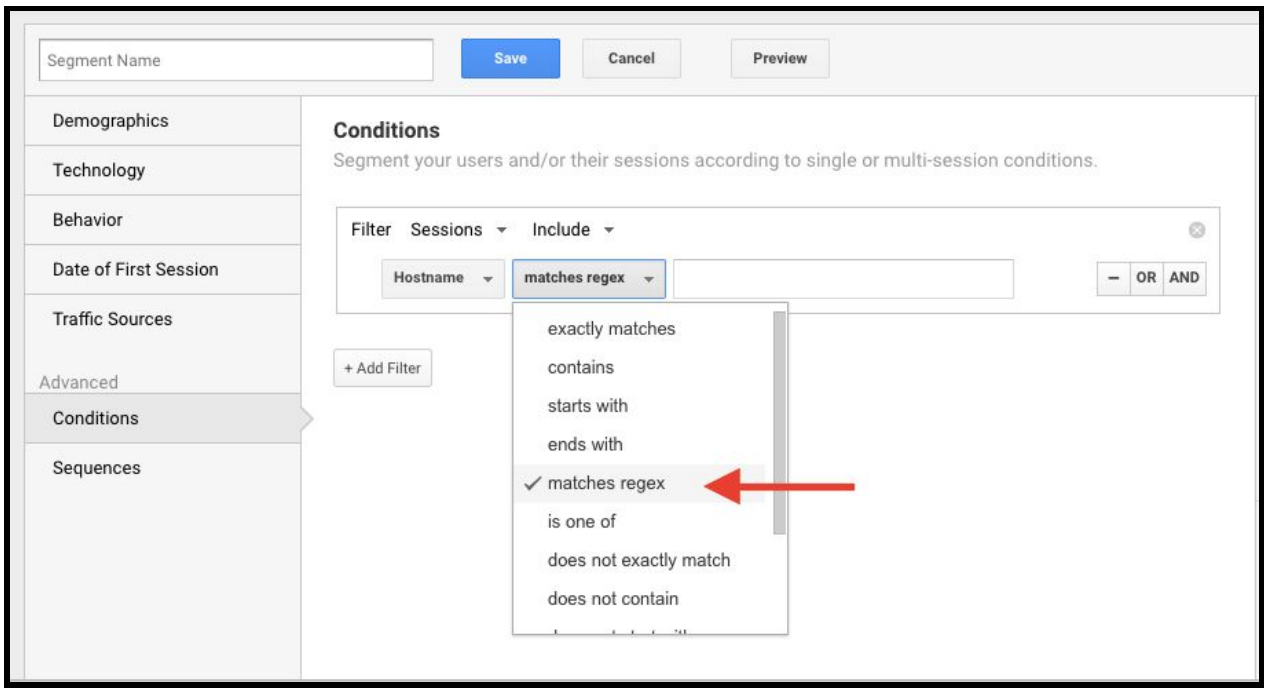
The report starts with "Campaign Name", and if you drill down into a specific name, it will show traffic by source/medium that was attributed to that name.

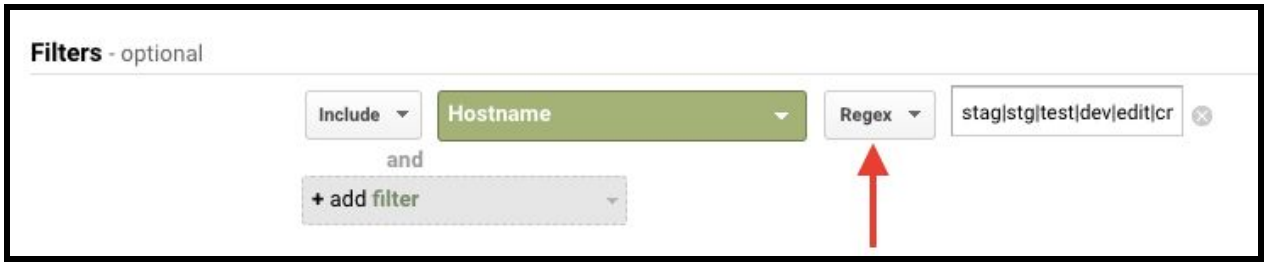
Additionally, you can use the campaign name to help create segments or use a filters in custom reports. As an example, you can create a segment of "users who came to the site from the 'National Awareness' campaign".

6. Regular Expressions

Regular Expressions (RegEx) are used in Google Analytics in both the filters an admin can apply to a View as well as the filters everyone uses in reporting sections. The three main places a regular user will benefit from understanding RegEx are when using the search function in default reports, when creating a complex segment, and when setting filters in custom reports (see figures below).







Okay, but what are they?

TechTarget defines a regular expression as:

> A regular expression (sometimes abbreviated to "regex") is a way for a computer user or programmer to express how a computer program should look for a specified pattern in

> text and then what the program is to do when each pattern match is found. For example, a regular expression could tell a program to search for all text lines that contain the word "Windows 95" and then to print out each line in which a match is found or substitute another text sequence (for example, just "Windows") where any match occurs.

The basic use of RegEx is to provide a user with a more flexible way to search for different forms of data. There are a set of characters that regular expression language recognizes to perform special functions, like signify the start or end of a text string, or signify an alternator function (aka using “or”). Here are some basic Regular Expression characters:

^ - signifies a start of a sting. Ex. ^ap might return “**ap**ple” or “**ap**erol” but not “c**ap**”
$ - signifies the end of a string. Ex. ic$ might return “ep**ic**” but not “pops**ic**le”
| - signifies “or”. Ex. face|nose might return both “pre**face**” and “**nose**bleed” but not “ear”

Using regular expressions will open a new world of options for your reporting needs. You can designate a report to show you a specific directory of pages, or multiple domains that contain a certain base, and loads of other valuable things.

We covered the basics of RegEx in our “Conquering Regular Expressions” webinar, which we highly recommend watching in preparation for the Certified Analyst Exam.

Other valuable resources:

Regex 101
Lunametrics RegEx guide
RegEx cheat sheet from E-Nor

7. Custom Reports

Custom reports are one of the most powerful ways to get the data you need in the DAP Google Analytics implementation. Because each profile contains many domains, custom reports give a user the ability to focus in on a particular hostname or page or subdirectory by using filters. They also allow a user to combine whichever metrics and dimensions they’d like, rather than what is given to them in the default reports.

There are four types of custom reports available in Google Analytics 360:

1. Explorer Report
2. Flat Table Report

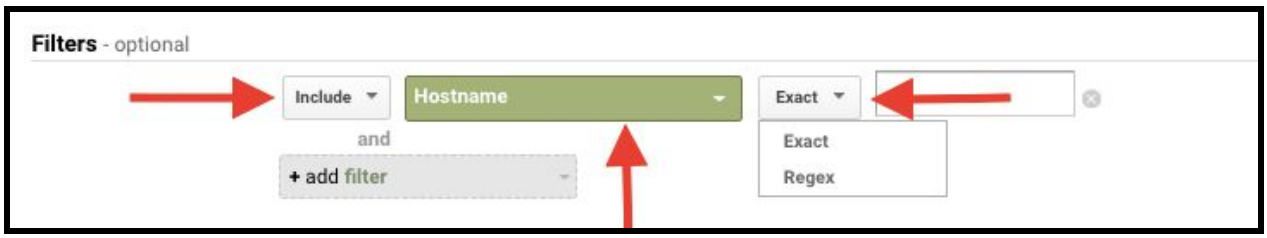
3. Map Overlay Report
4. Funnel Report (beta) - only available in Analytics 360!

We'll review the pros, cons, and use cases for them all.

Key concepts for understanding custom reports: metrics, dimensions, filters, regular expressions

Key Concept: Filters

A good grasp of how filters work will be crucial to your success with custom reports. Without a filter, any custom report you create will include data from the entire profile, and often, that's not what a user is intending.

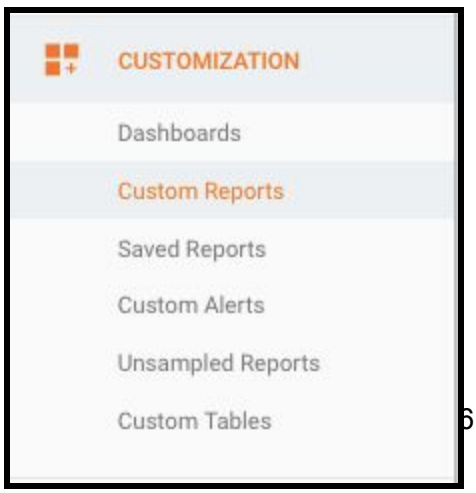


Using filters allows a user to specify to "filter" the data based on categories. The most frequently used dimension to filter data in our case is "Hostname". If a user creates a filter that says "Include >> Hostname >> Exact >> sba.gov", it means "filter this report to include only information that matches the hostname of sba.gov".

Note that such a filter will NOT include data from "applications.sba.gov", because if "exact" is chosen, GA will look for an exact match. Conversely, if the user chose "Regex" and included "sba.gov", the report would return data for all hostnames that contain "sba.gov", including subdomains. More on this in the section on Regular Expressions.

Additionally, you can choose to either include or exclude data for a specific dimension. If you are interested in all sessions except those on iPhones, you can choose "Exclude >> Mobile Device Info >> Regex >> iPhone" (to exclude all models of iPhone).

Finally, a user can stack multiple filters on a single report. To look at mobile users only on sba.gov, you can implement a filter for hostname matching sba.gov AND a filter including "Device Category >> Exact >> Mobile".



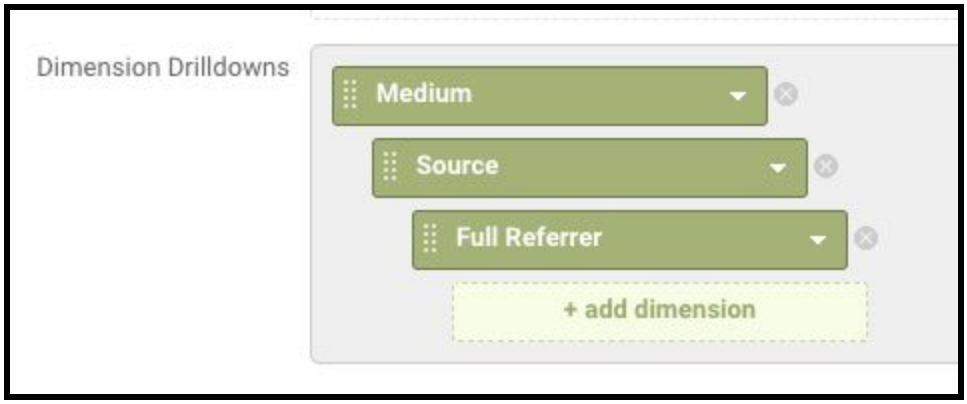
Explorer Reports

The Explorer report is the default type of custom report you'll start with if you click into the "Customization"

section and choose “Custom Reports” >> “New Custom Report”.

The main feature of the Explorer report is that it allows you to choose various “layers” of dimensions to “drill down” into. The report will only show you one dimension at a time (unless you choose to add a secondary dimension after the report is created.)

When entering dimensions to include in a report, you’ll notice that each dimension is indented, which signifies the “layers” to explore. If a layering of dimensions looks like this:



The report’s first screen will have rows that display Medium. If you then click on a specific medium (like, organic, for example), it will drill-down into that dimension and show you the next layer, which in this case is Source. So the next report screen will show rows for all the possible sources, but **only** for the medium that was clicked (in our example, that might be Google, Yahoo, Bing, etc).

Medium	Pageviews
	320,593 % of Total: 100.00% (320,593)
1. organic	199,859 (62.34%)
2. (none)	59,274 (18.49%)
3. referral	52,505 (16.38%)
4. email	8,673 (2.71%)
5. social	275 (0.09%)
6. twitter	4 (0.00%)

The Explorer report will only give a user insight into the details of a chosen dimension with respect to the layers above it.

Flat Table Reports

Flat Table reports serve essentially the same function as Explorer reports, except they allow you to choose more than one dimension to show at a time. Each unique combination of dimensions will get its own row, so these reports have many more rows than explorer reports.
Rather than having to click to drill-down to the second layer of dimensions, all the dimensions are shown at once, in a “flat table”. Here is an example of the same report we created for Explorer Reports, but in a Flat Table.

Medium	Source	Full Referrer	Pageviews
1. organic	google	google	178,489 (55.67%)
2. (none)	(direct)	(direct)	59,274 (18.49%)
3. organic	bing	bing	12,771 (3.98%)
4. organic	yahoo	yahoo	7,680 (2.40%)
5. referral	grants.gov	grants.gov/custom/viewOppDetails.jsp	2,518 (0.79%)
6. email	Research - Sep 2017 Newsletter	Research - Sep 2017 Newsletter	2,344 (0.73%)

The benefit of this is that sorting by metrics (in this case, pageviews), we can see how every specific combination of dimensions ranks, rather than seeing ranking only within each dimension category.

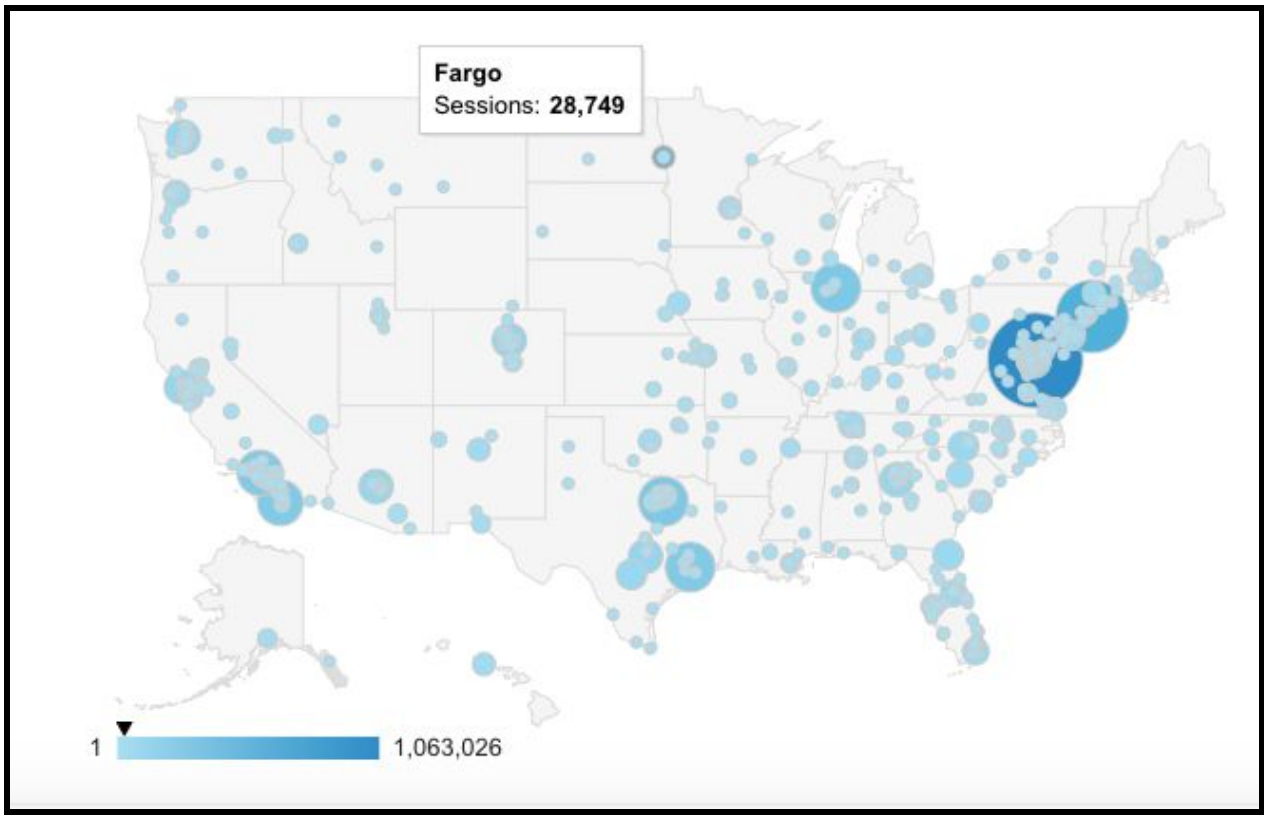
Flat table reports are great for exporting into spreadsheets for analysis, visualization, and pivot tables.

Hint: If a report is greater than 5,000 rows, the GA interface will only show you 5,000 at a time. You can remedy this by choosing to export an unsampled report, which will export up to 3M rows. Additionally, you want to use this option anytime you see sampling on a report.

Map Overlay Reports

The Map Overlay option is a report used strictly for location information. A user can set a specific zoom level (like, across the world, a continent, or a country) as a base point. Then the user can choose a reporting dimension such as country, region (which in the U.S., corresponds to state), or city. A metric for the report is chosen, and in addition to the table of data, a dynamic map is included that you can hover over. This may be useful, for example, in a presentation to management.

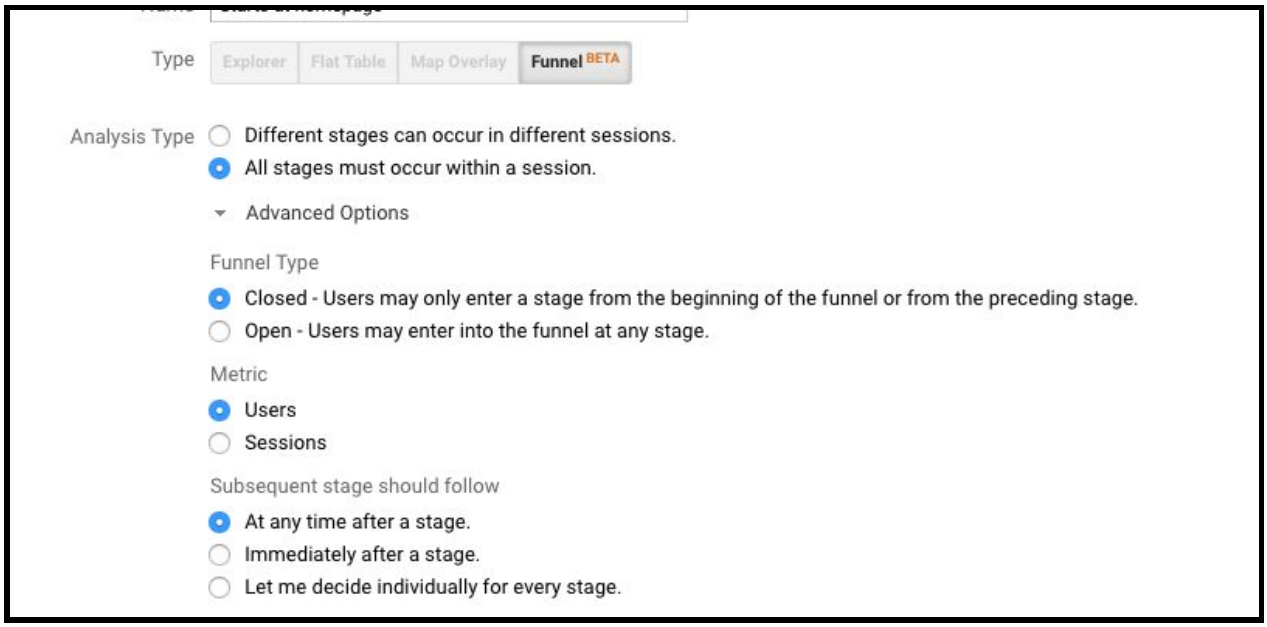
Below is a screenshot of a Map Overlay report with a zoom level of country = United States, and a dimension of city, with a metric of sessions.



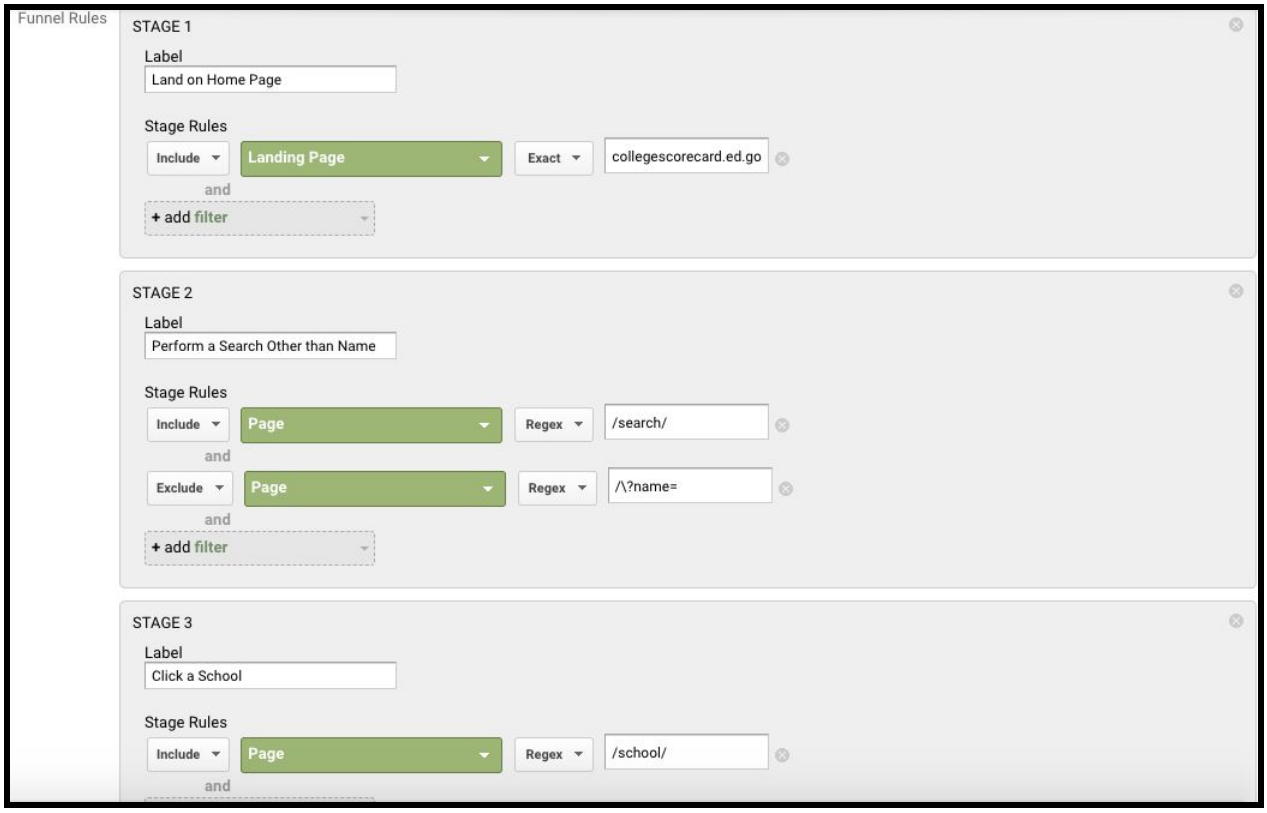
Funnel Reports

Funnel reports are an exciting addition to the custom reporting options, and they are only available in Analytics 360. If your website also has an independently managed GA Free account, it will not have the ability to use Funnel Reports.
These reports give the ability to track visitors through specific journey paths on a site. You can create up to five steps for a visitor to travel through, and the report can provide information about how many visitors move from one step to another, and at which points visitors “leave” the funnel. This kind of data is great for understanding where bad UX experiences happen, and also determining the best landing pages for final step conversion.

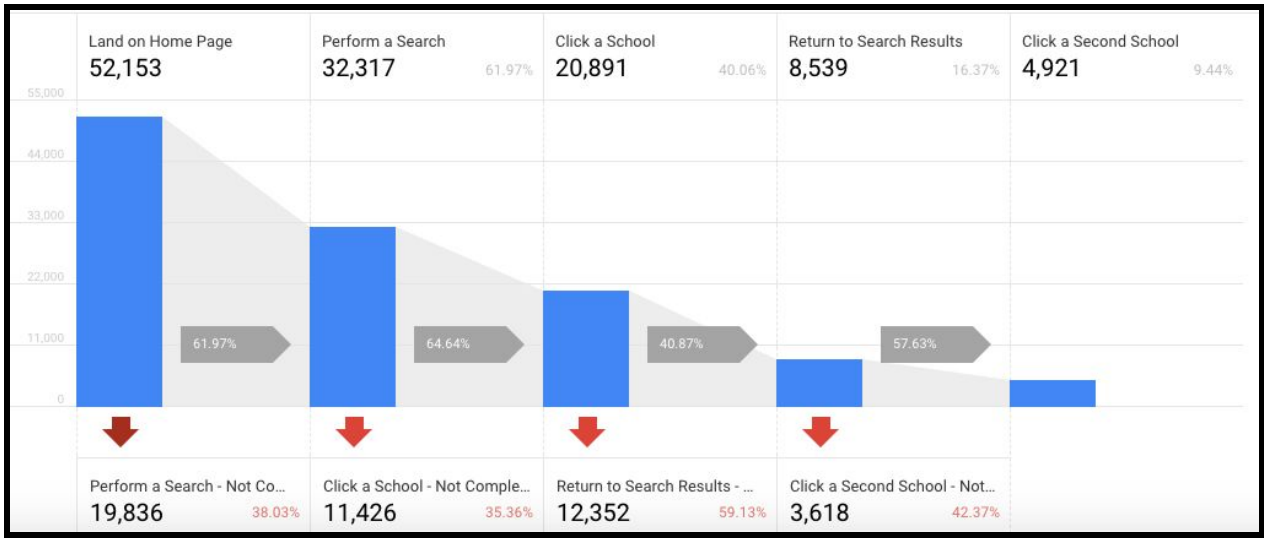
In the setup of the report, GA gives you many options for making the funnel “open” or “closed”, sequencing of steps, and metric to report on.



Then you can set each step of the funnel, which may be landing on a specific page or executing an event, like a file download, for example.



The result is a very informative visual report showing the flow of users through the steps of the funnel.



8. Segments and Advanced Segments

What is a segment?

A segment is a subset of your Analytics data. For example, of your entire set of users, one segment might be users from a particular country or city. Segments let you isolate and analyze those subsets of data. Additionally, when you export reports with segments, the segment name becomes a column. So there's added analysis power there.

Examples of informative segments include segments to better understand mobile vs tablet vs desktop user behavior, new vs returning user behavior, or specifying acquisition channels to determine if say, referral sessions last longer than organic search sessions on average.

One caveat with segmentation is that it can create sampling issues, where Google Analytics processes only on a certain percentage of the total amount of data to reduce the time it takes to return reporting. If your agency's view in Google Analytics contains a lot of data, you will likely experience sampling.

What is data sampling?

Data sampling is a process of selecting a subset of data, indicative of the larger set of data. It is used in statistical analysis to analyze large data sets in a cost efficient manner and within a reasonable amount of time. Data sampling isn't bad in and of itself, but if a selected sample is too small or isn't a good representation of the whole data set, then analyzing this data set would not give similar results to analyzing all of the data.

How to tell if your data is being sampled

Google Analytics chooses to analyze a set of data or only a subset of data depending on the type of query logic you build into a report. This depends on the power, cost and time for completing the analysis of a particular data set.

Google Analytics tells you when your data is being sampled by displaying a colored shield next to the name of the report. Whether your reporting is sampled or not depends on the amount of pre-aggregated data available in a table to be queried.

 = No sampling

 = Sampling occurring

Resolving sampling

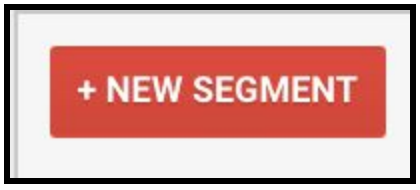
On any report, you should always check the shield in Google Analytics to understand whether your data is being sampled or not. You can see the actual sampling percentage by hovering over the yellow shield. Whenever possible when sampling occurs, clicking on “Export >> Unsampled Report” will help you get more accurate information. In the interface itself, you may be able to reduce sampling by shortening the date range or decreasing the complexity/granularity of the segment or filter.

Creating a segment using hostname

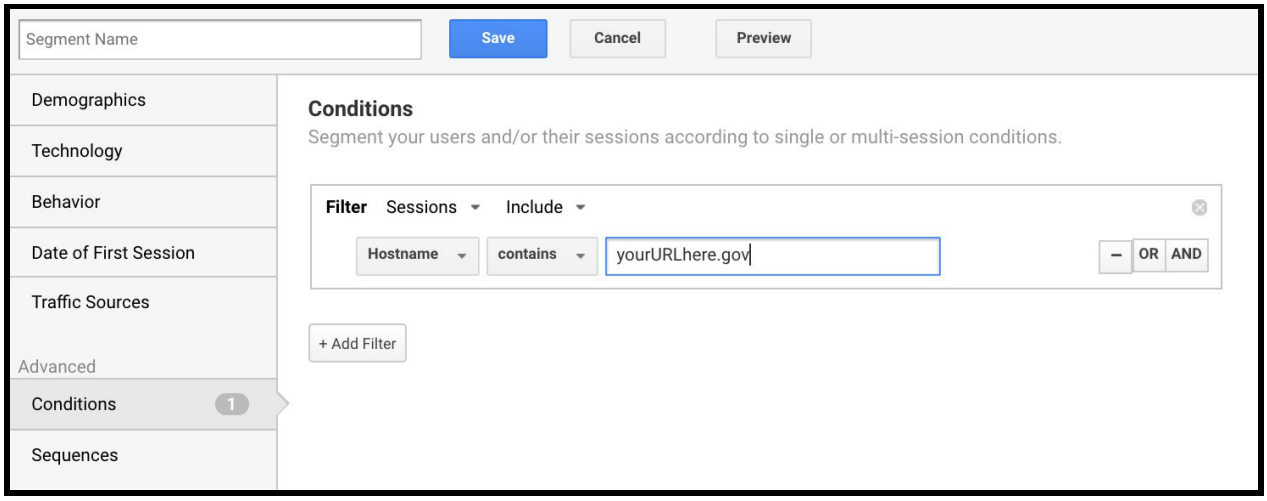
What are segments?

1. Click “Add Segment”

2. Click “New Segment”



3. Name your segment and select “Conditions”. Select “Hostname”. Enter your .gov domain in the text field.

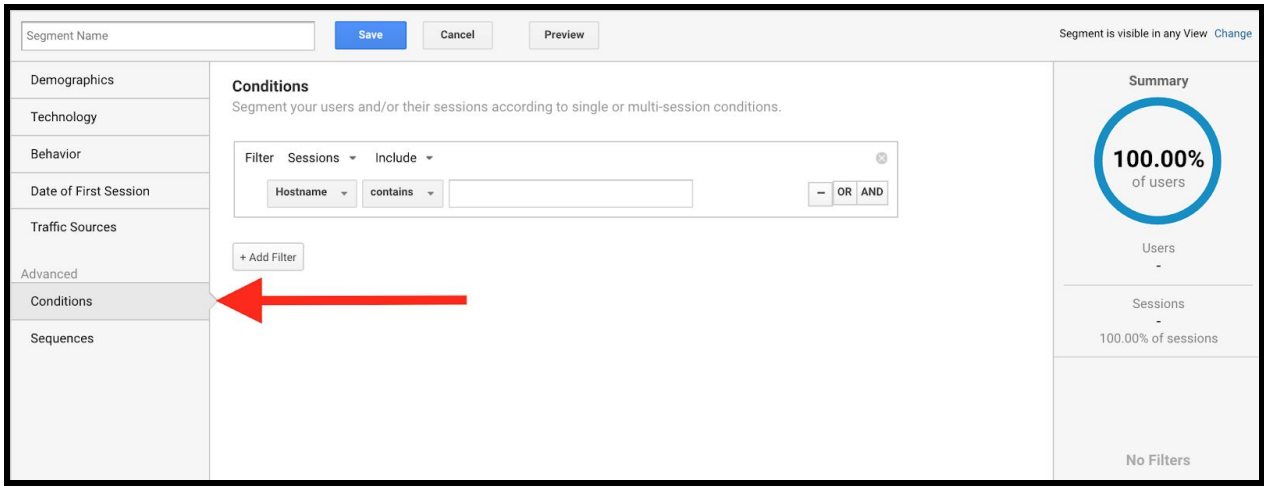


4. Click “Save”

5. Your report will now display data specific to your site.

Segments can be combined with any dimension and metric, to isolate a specific subset of traffic that best represents the use case or user profile you wish to report on. For example, you can create segments based on technology used - such as device type. You could also create a segment based on campaign URLs to understand the performance of your campaigns.

There are various segments provided for you by default in the “System” section. You can also choose “Create New Segment” and choose segment rules from a list of options that appear on the left side of the window. If you don’t find what you are looking for in the options, you can scroll to “Conditions” under “Advanced” and choose the dimension you’d like to filter by, as you learned above in specifying a hostname.



9. Understanding Events

Events are interactions that Google Analytics can record that are not part of the page load. Examples of events include specific button clicks, video plays, pagescroll depth tracking, and file downloads, to name a few. We recommend you pair this chapter with our Understanding Events webinar.

Events provided by DAP

An out-of-the-box Google Analytics implementation comes with no events tracked by default. Events need to be configured with extra code in order for the data to appear in the GA reports. But in the DAP implementation, a set of events is hardcoded and will record as soon as the DAP script tag is implemented. Those events are:

1. **Outbound link clicks** - Any link on your site that leads to a different hostname will be recorded. So, if your site is usa.gov but you link to a page on opm.gov, when that link is clicked, the interaction will be tracked.

2. **Downloads** - Any link to a file that ends in .pdf, .xls, .doc, .exe and many other common file extensions (see Code Capabilities Summary pg. 6 for a complete list) will be recorded separately from a traditional outbound link click.

   * Please note that the “Download” event looks for a link with a specific extension as the end of the URL. If your site dynamically issues file downloads without including the extension as the final part of the URL string, our code will not record it as a download. Additionally, DAP (nor any other GA code) cannot track a file click separate from a webpage implemented with the GA script. For example, a PDF directly linked in an email

or social media post CANNOT be tracked by Google Analytics. In that case, you'd need to measure using server logs.

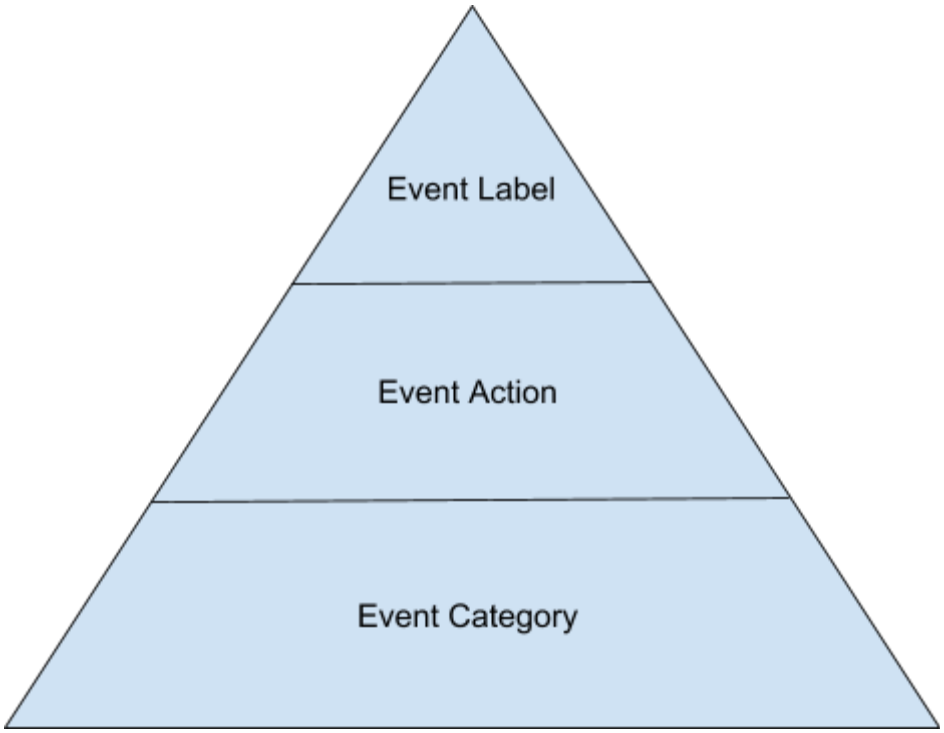
3. **Mailto links** - A link to an email address.

4. **Telephone click** - On many mobile devices, the OS can recognize telephone numbers and automatically link them so that pressing the number will initiate a call. Please note that this is not the case for ALL mobile devices, and is not something that is tracked on desktop devices.

Optional event: Youtube video tracking - By adding **yt=true** into the DAP script tag on your site, you can enable Youtube video tracking as an event. The event tracks when a user presses play, as well as which percentage milestones the user reached in watching the video (more starting on page 7 of the Code Capabilities Summary).

There are also opportunities to define your own custom events, but please contact the DAP team before pushing your own events into the DAP account. Being that any events increase the number of "hits" in our system, and we pay by total hits, we may ask that you not implement certain events that have a very high volume.

Events Structure Hierarchy



Events are defined by three dimensions, in order of specificity from most broad to most specific: Event Category, Event Action, and Event Label.

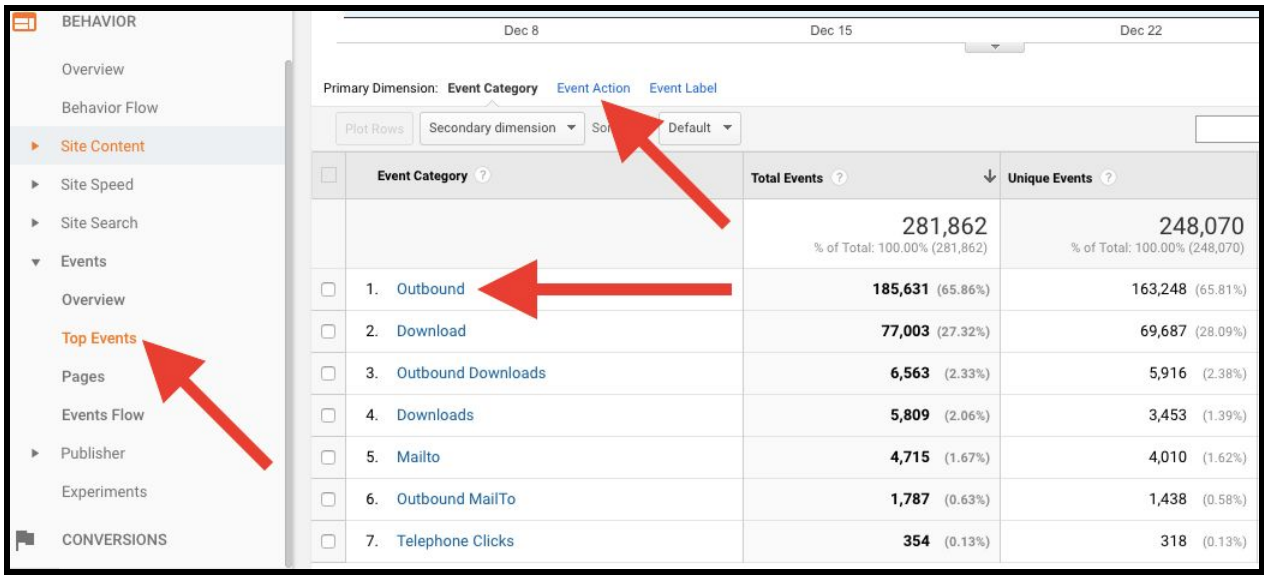
When using events in reports or implementing custom events for an interaction, it is important to understand that as you drill down from Category to Action and Label, the information should become more specific. For example, the DAP default event for Download is structured as the following:

Event Category - “Download”
Event Action - the extension of the file (such as .pdf, .doc, .xls, etc.)
Event Label - the full URL of the file (such as https://www.irs.gov/pub/irs-pdf/f1099msc.pdf)

This way, you can create reports that specify for a type of event, like download, or for specific information, like the top clicked file URLs. This methodology follows suit for all events, although each event may have different types of data in the structure.

Events Reporting in the GA Interface

The default reports for events can be found under Behavior>>Events.



In the Top Events section, you can choose to have the report show you Event Categories, Actions, or Labels in their entirety by clicking the links next to “Primary Dimension”. Alternatively, you can click on an individual row, like “Outbound”, to drill-down into that row. If you click on “Outbound”, which is an Event Category, GA will show you a report for all Event Actions within “Outbound”, and so on. Each category has differing Event Actions and Event Labels.

You can also report on events using custom reports or segments.

Dimensions related to events	Metrics related to events
Event Category	Total Events
Event Action	Unique Events
Event Label	Sessions with Event
	Event Value
	Sessions / Sessions with Event

By using the following dimensions and metrics with filters, you can create custom reports that specify events for a specific website, or even narrow down events taking place on a specific page. Getting a firm understanding of the metrics and dimensions related to events will open many doors for analysis with the custom reports available to you.

10. DAP Code Optional Capabilities

The DAP code allows for optional features such as Youtube tracking, linking a parallel GA account, setting a site topic, enabling cross-subdomain session tracking, resetting the default cookie timeout length, and more. All of these features are controlled by adding parameters into the script tag. For all the options and guidance on implementation, please see the DAP Code Capabilities Summary.

Glossary

Below you'll find a few important terms. Many more terms covered in various online GA glossaries, such as this one.

Bounce: a single-hit session. While bounce rate is a negative indicator for most types of landing pages, some level of bounce is always to be expected, especially if your website model is more informational than transactional.

Bounce Rate: the percentage of single-hit sessions
Note: An event interaction negates a bounce, as does viewing a second page during the session.

Channel Grouping:
A default selection of traffic sources in the Acquisition reports, such as Direct, Organic Search and Email.

Dimension:
A descriptive attribute of data. For example “channel” (with a return of ‘organic’, ‘referral’, direct’) or “page” (with return of a specific URL)..

Event:
A user interaction that occurs after pageload. Examples include file download, button click, video play, scroll depth, or outbound link click.

Hit:
Interaction that results in data being sent to Analytics. This includes pageloads, events, and ecommerce hits.

Landing Page:
The first page of a given session.

Metric:
A piece of data that is countable. For example, “sessions” or “page views”.

Pageview:
Any time a page is loaded. Page refreshes and hitting the back button are included.

Sampling:
When there are large amounts of data in a view taking up processing power, Google will select a subset of data from your traffic and report back a lower than 100% percentage of that sample set. Based on the percentage of the sample, Google takes the metrics from the sample and multiplies to estimate totals if the data were representative of 100%.

Segment:
A subset of data which share common attributes. One may create a segment such as “mobile sessions” or “sessions from California”, as examples.

Session:
An instance of a visit within the given site hostname. A session encapsulates all activity including page views, events, e-commerce transactions, conversions, and goals as long as the use has not exited the site hostname.
□
Session cookie timeout & expiration:

When a user visits a site, GA places a cookie in your browser to track how long you browse a site. If you cease to browse a page for 30 minutes, the session will time out, thereby ending your session. Any browsing you do after the cookie timeout is counted as an additional session.

User:
A unique individual who accesses website or app one or more times. For a website, an individual user is defined by the presence of the _ga cookie in the browser. Therefore, a single human visitor may appear as multiple users if s/he visits from varying devices. A GA cookie expires after two years, assuming a user does not clear their cookies from their browser.