2023

Web Governance Report

11,000 Homepages



Observe Point

Introduction

This year, we switched up the methodology for our annual report. Why? Because over the past couple of years, we weren't seeing much change year over year. Meaning, while businesses were concerned with web governance and how it affected their decision-making and privacy compliance, they continued to express uncertainty about how to rein in unruly and large websites.

With this year's web governance report, we wanted to offer a few things.

- > Provide a different perspective by using the ObservePoint platform to source data instead of surveying you
- > Scan one page each from a variety of companies in different regions
- > Assure you that you're not alone in trying to wrangle your website
- Give you some benchmarks to compare against

We hope that you find this report, a result of scanning more than 11,000 homepages, enlightening and inspirational. Perhaps it'll spark some ideas on how you can improve your own website to gain a competitive advantage.



What Did We Audit and Why?

This year, rather than surveying you, we decided to lean into the type of data only we can source.

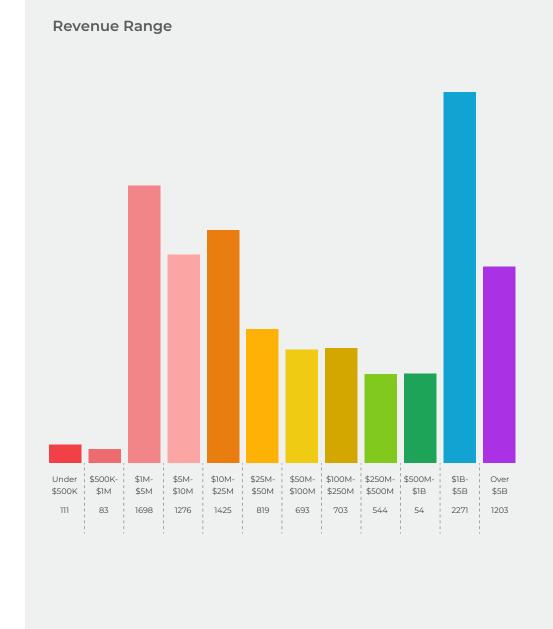
A. # of Domains

So, we used our platform to scan the homepage of 11,374 companies of varying sizes and industries, sourcing from the following lists, among others:

- > Inc. 5000
- > NASDAQ
- > NYSF
- > UK 100
- > Fortune 1000

A Note About Methodology

We scanned all of the homepages from our servers located in the U.S. That means that our default is an opted-in state for privacy and cookie consent. We also accounted for the fact that a very small percentage of sites block unrecognized web scanners.



B. Company Breakdowns by Revenue, Industry, Region

Our goal was to get a good mix of companies with different revenue ranges and industries.

We're categorizing enterprises, mid-sized, and small businesses by how Gartner breaks it down:

> Small: <\$50M

➤ Mid-market: \$50M - \$1B

> Enterprise: >\$1B

By this measure, we've scanned about:

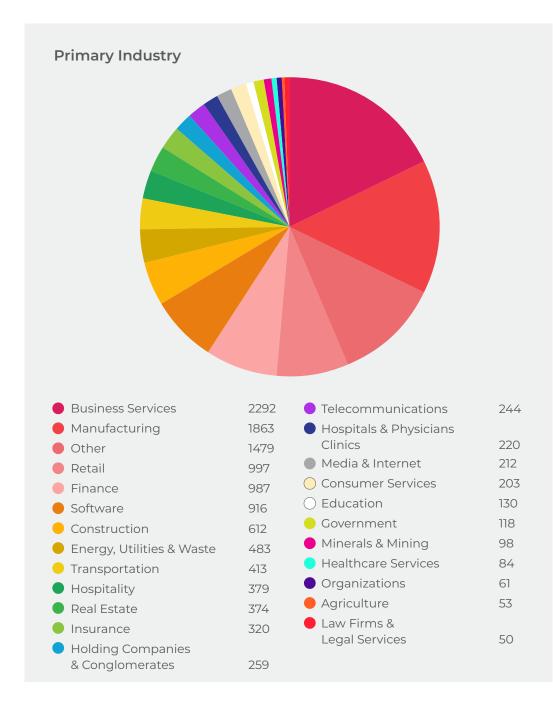
> Small businesses: 48%

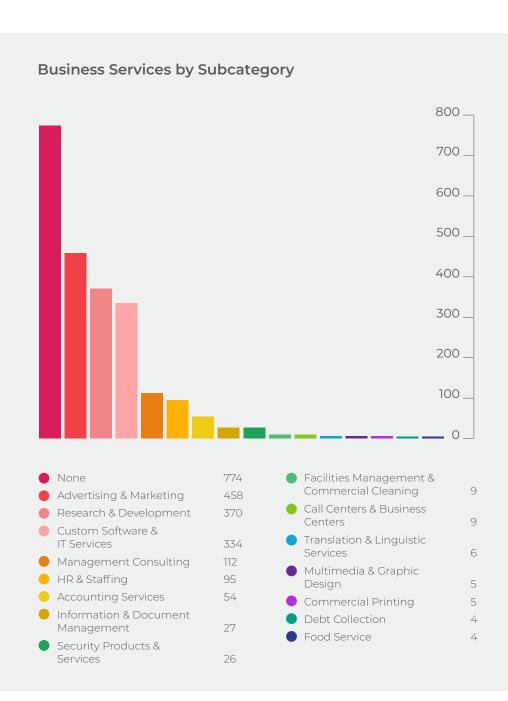
➤ Mid-sized businesses: 22%

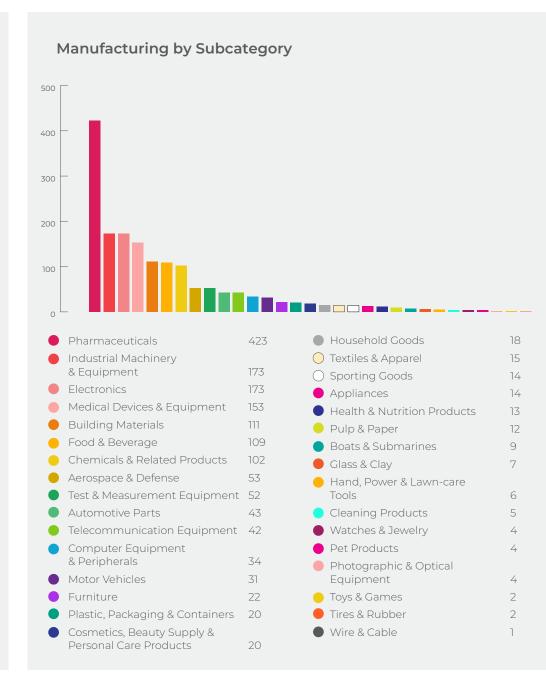
➤ Large businesses: 30%

Business Services and Manufacturing were the two largest categories, followed closely by Retail, Finance, and Software. Construction, Energy, and Transportation made up the third group.

One thing to note is that Manufacturing covers a wide array of companies because any organization that makes a product is included in that category.



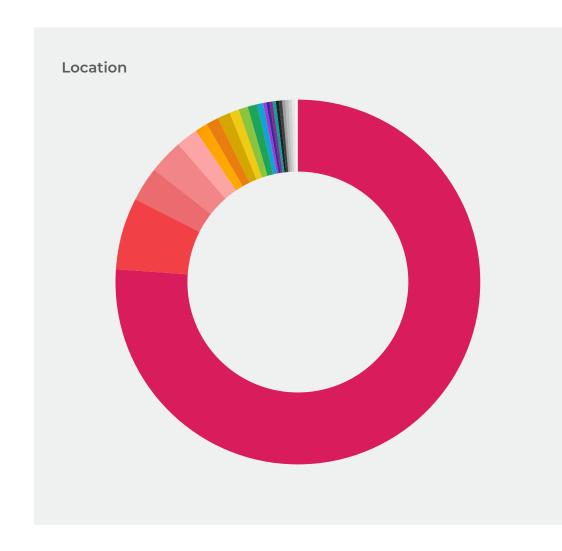




For the regions that these companies operate in, we had a concentration of about 77% of scans from U.S. companies, with the UK coming in second at 6%, followed by Germany at 3.4%, and France at 2.7%.

Location Key

United States	7128		Chile	6
United Kingdom	597		Peru	6
Germany	296		Colombia	5
France	265		Cayman Islands	5
Canada	180		None	5
Netherlands	123		Norway	4
Sweden	96		UAE	4
Spain	95		Uruguay	3
Italy	92		Russian Federation	3
China	82		Greece	3
Israel	67		Monaco	3
Ireland	58		Turkey	2
Brazil	29		Jersey	2
Switzerland	21		Finland	2
Australia	21		Hong Kong	2
Bermuda	16		Thailand	2
Luxembourg	15	\bigcirc	Estonia	7
India	13	\bigcirc	Ukraine	7
Mexico	12	\bigcirc	Jordan	7
Singapore	12	\bigcirc	Indonesia	7
Argentina	12	\bigcirc	Romania	7
Korea, Republic of	10	\bigcirc	Puerto Rico	7
Japan	9	\bigcirc	Macao	7
Belgium	8	\bigcirc	Panama	7
Taiwan	7	\bigcirc	Nigeria	7
South Africa	7	\bigcirc	Cyprus	7
Denmark	7	\bigcirc	Mozambique	7





Structure of This Report

We've organized this report by the things that ObservePoint scans:

- > Pages (what's on a web page)
- > Tags (what technologies are there)
- Cookies (what data is being collected)

We scan these things so that our customers can validate their:

- Landing Pages
- > Analytics
- > Privacy



Pages

A. Page Load Times

Faster page load time is linked to happier visitors. so most organizations focus on speeding up their pages to improve user experiences and conversion rates.

According to **Portent**, A B2B site that loads in 1 second has a conversion rate 3x higher than a site that loads in 5 seconds and 5x higher than a site that loads in 10 seconds. (Our scanned websites were 82% B2B, but that Portent link shows you metrics for B2C sites as well.)

Google has a more detailed **Core Web Vitals** metric which takes into account the stages in which a site loads, including largest contentful paint (LCP or how long it takes for your main content to load), first input delay (FID or how long it takes until a user can interact with a page), and cumulative layout shift (CLS or how often users experience layout shifts).

Generally speaking, you want your site to **load under 3 seconds** if it's an e-commerce site. Conversion rates drop by an average of 4.42% with each additional second of load time between 0-5 seconds

Page Load Times

	Good	Need Improvement	Poor
LCP	<=2.5s	<= 4 s	>4s
FID	<=100ms	<=300ms	>300ms
CLS	<=0.1	<=0.25	>0.25

Bounce rates for sites that load within 2 seconds was at 9% but skyrocketed to 38% after 5 seconds.

The average mobile web page takes **8.6 seconds**, most likely due to having too many elements on the page, and bounce rates are at 56.8%. There's a great opportunity there if you can get your site mobile-optimized.

We scanned the homepages in this report on desktop and found the average load time to be 4.23 seconds, which isn't terrible. (ObservePoint reports the Google Chrome "Load" metric.)

When we broke it down by industry, the outlook became more interesting.

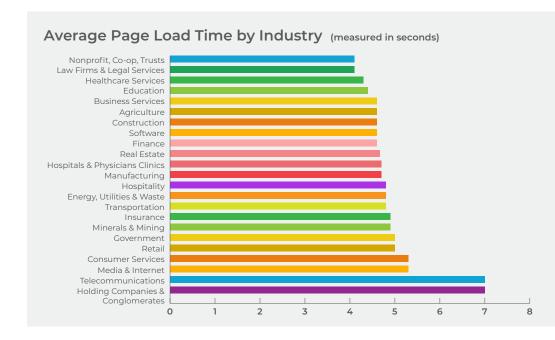
Come on, Telecom. What's going on? When are you going to show up? Is there really that much functionality on your sites? Or does this need to be a bigger priority? I wouldn't expect that much concern from Holding Companies & Conglomerates who probably aren't really selling anything. But as you can see, we're showing that everyone is going past 4 seconds, which means, there's plenty of room for improvement and an opportunity for you if you can get faster than your competitors.

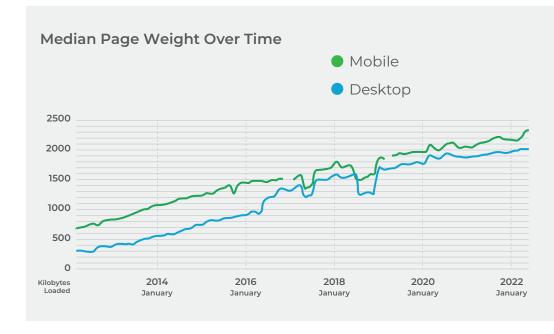
B. Page Sizes

The page size or weight is one factor affecting load time. Web pages have been getting heavier as more functionality and elements (images, forms, videos, interactivity) have been added. For example, in 2017, the median web page weighed 15.32 MB while in 2020 it shot up to 20.8 MB. For 2022, we're looking at 21.7 MB.

Median page weights have been increasing over time.

Our scan showed an average weight of 10.32 MB. However, breaking this down by industry is more appropriate, since some sites need to rely heavily on content like video and animation, while others can be predominantly text and a small volume of compressed images.

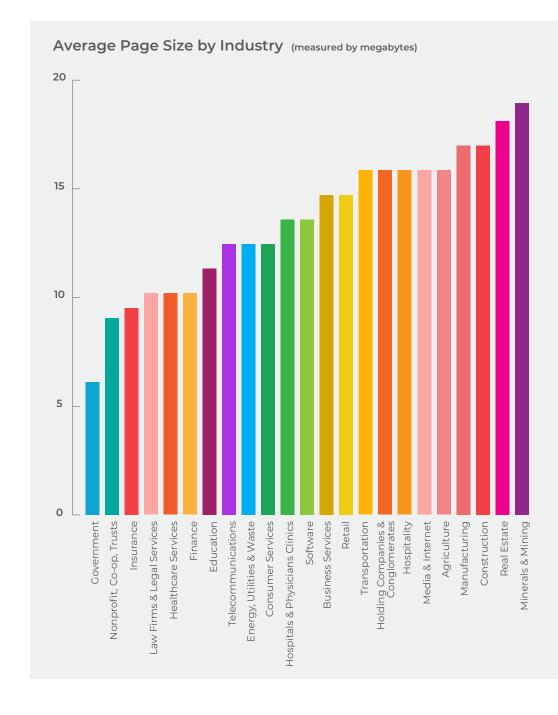




It's interesting to note that page size is not the only factor in speeding up your load times. Government, our winner on page size with 5.28 MB, is all the way down in 18th place on load time with 5.05 seconds. Telecommunications is in the middle of the pack with a website size of 10.93 MB (slightly under the overall average), even though they had one of the slowest average load times of 6.8 seconds. And what makes Construction company websites so heavy? Must be concrete and rebar. Holy moly, 16.14 MB for Minerals and Mining coming in last with all that rock and precious metals.

What else can influence load times if not page size?

- > Slow, broken, or multiple tags we'll talk more about tags below
- > JavaScript Timing load asynchronous or defer execution
- > Etags and Expires Headers reduce the number of requests a browser makes to the server by preventing reloads of files that rarely change like your logo
- Content Delivery Network use a CDN to store assets on servers around the world



C. Javascript Console Errors

We wanted to mention that Browser Console Logs is a standard report in the ObservePoint platform, but since it would be hard to consolidate what a common error or warning might be with so many sites, here are some overviews.

- > First, 99.9% of scanned sites had Console Log Errors.
- ➤ The average number of Console Log instances was 13.54

If we look at the category of the logs, we have:

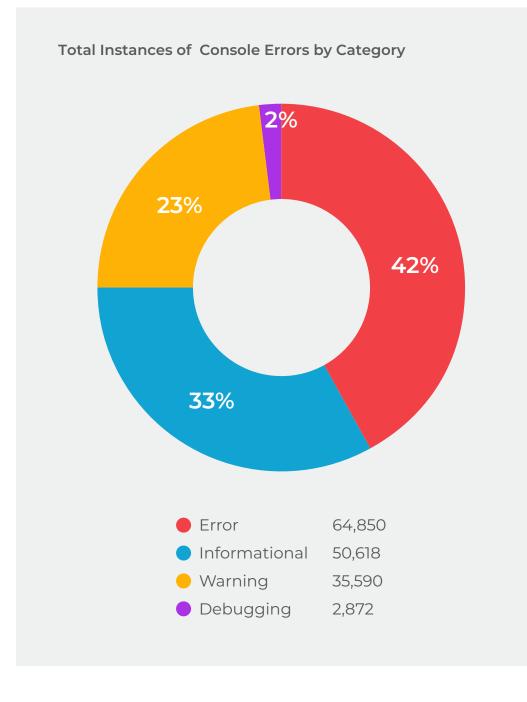
• Frror: 42%

· Informational: 33%

· Warning: 23%

· Debugging: 2%

The errors and the warnings should be delved into and fixed to minimize poor user experiences.



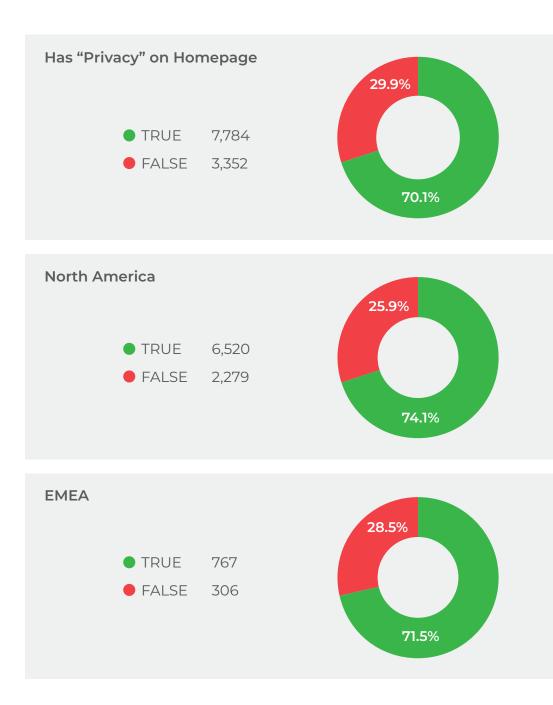
D. Privacy & Consent Management Platform

For this report, we wanted to see if every home page had a privacy policy, but these can be called different things from Privacy Notice, to Your Privacy Choices, etc. So, we scanned for just the word "privacy," which may or may not indicate an actual privacy policy.

29.9% of scanned websites did not have the word privacy anywhere on their homepage.

Privacy policies and "do not sell/share" links should be accessible from any entry point to a website, so these items are usually housed in a footer or other global navigation. As you can see, even this foundational compliance requirement needs some improvement for a large number of sites. To mitigate the risk of being fined by regional regulators, ensure your privacy policies and opt-out controls are displayed and functional on every page.

For the comparison of Europe, the Middle East, and Africa (EMEA) against North America, we had to exclude any international sites that did not end in .com or .org. This way, we hoped to exclude non-English sites that would not have the word "privacy" on it. We didn't expect American sites to be more likely to have privacy notices than European sites, given GDPR, but the percentages are remarkably similar, with the U.S. in a slight lead. We'll delve further into privacy matters when we get into Cookies and Consent Management later in the report.





Tags

A. # Of Unique Tags

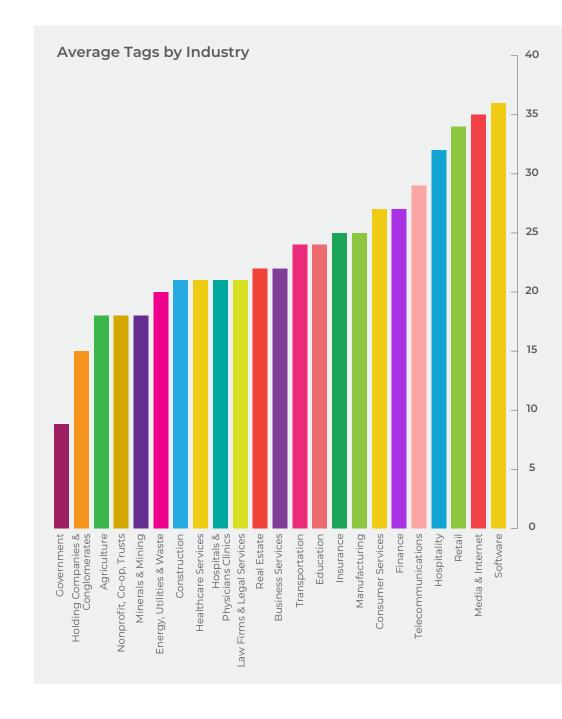
The average number of tags we found on a page was a whopping 25.62 with an average tag load time of 415.03 milliseconds.

We actually thought the average might be a little closer to 40 tags, but you can see that is more likely in the Hospitality, Retail, or Media spaces where advertising and performance measurement are paramount.

We dug into the raw data to see what the most number of tags a company had was:

- First place: 81 tags by an online retailer
- > Second place: 76 tags by a cosmetics company
- > Third place: 70 tags tied between two travel and hospitality brands

For tag load time, we typically recommend a best practice of keeping it under 500 ms, so the average is not bad at all.

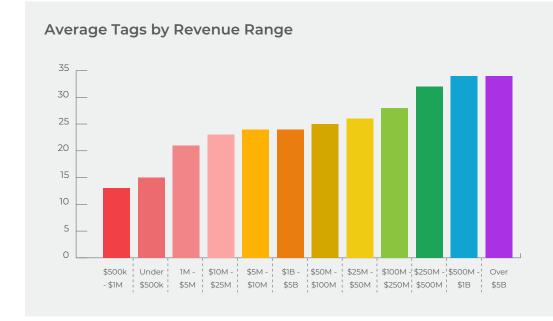


Interesting to note that Government sites, the fastest-loading sites, have the fewest tags, which makes sense since government sites are often not advertising and usually have smaller tech stacks. But, Holding Companies & Conglomerates, an industry with one of the slowest page load times, also had very few tags, so that correlation doesn't hold true for every segment.

When we break it down by company revenue, it looks generally like the larger companies have predictably more tags. There's just one anomaly at 6th place where pretty large companies between \$1 Billion-\$5 Billion revenue, have fewer tags than smaller companies. But generally speaking, the bigger your website properties, the more complex your tag implementations are.

The proliferation of tags is one of the reasons a solution like ObservePoint is necessary to manage the ever-increasing amount of MarTech solutions and their tracking pixels. As of this year, the MarTech Map has reached 11,038 types of technologies available for you to help with your marketing. (Not that every single one of these is an on-page tool that requires tags.)

Note: When looking at the Tags and Tag Requests we only looked at those companies who had tags on their homepage, and we did not include duplicates in our number, i.e., we only reported the one tag and its corresponding data.







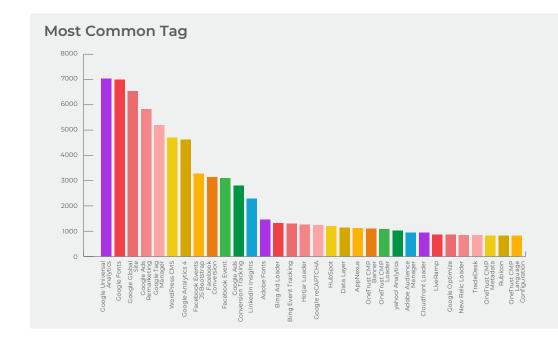
Our scan found 588 tags on over 11,000 homepages, a portion of which are shown in the table above. You can see that at the time of our scan in Ol, a ton of companies were still running Universal Analytics simultaneously with GA4. You can see that tags from the Google ecosystem are the most popular.

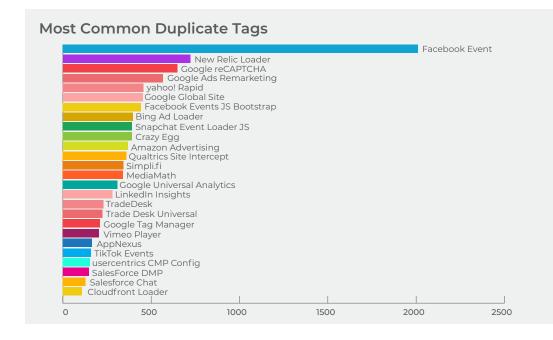
Google Tag Manager was on 37.3% of our scanned homepages, which makes it the most popular tag manager. However, if you look above it at the third most common tag, you'll notice Google Global Site, which is also a tag management technology. It's a paired-down TMS for Google properties only.

Best practice would be to not use both on the same pages because you will have to be careful with the data layer and other resources that are used by both tools, and you could end up double-counting data like pageviews and other metrics. With 6,537 instances of Google Global Site and 5,197 of Google Tag Manager, we reasoned there would be some overlap. Turns out 3,904 companies are double tagging with Gtag and GTM.

B. Most Common Duplicate Tags

We tend to hammer home the idea that duplicate tags are something you need to manage, so you don't have inflated numbers that fool you into thinking things are performing better than they are. The good news is that the most common duplicate tags were for WordPress or fonts, so we removed those results from the graph below, as well as Drupal CMS, a content management software. We're not so worried about content management, so much as whether analytics or measurement tags are getting duplicated.





You can see that the Facebook Events tag is most commonly duplicated, but that also makes sense since you can track different events on your page, like viewing content, searching, completing a registration, finding a location, etc., in order to re-target Facebook ads to those users.

New Relic Loader is a browser monitoring technology, so that doesn't seem too concerning either. Yahoo! Rapid is probably the API to Yahoo Finance data, which sites might use to show real-time data for financial markets, so that's another content piece. What you want to look into is if you see your Google Global Site or GUA being duped, as that could be messing with your measurement.

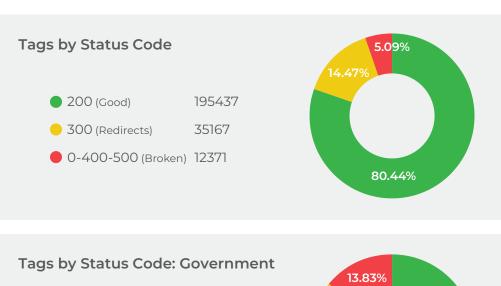
C. Broken Tags

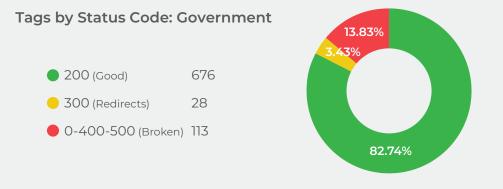
Out of 11,000 homepages we found 24,975 tags, of which 5.09% had broken tag status codes. Broken tags need to be located and fixed because 1) there's no data being collected for that technology you've placed on your site and 2) they can affect your site's performance, slow down pages, and even interrupt ads or integrations.

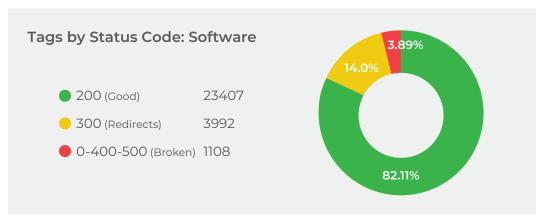
We tried filtering by industries to see if we could find some outliers. This is not good enough for Government work, don't you think?

So, who's doing the best with broken tags?

Turns out that software companies are, so pat yourself on the back if you're in this industry!







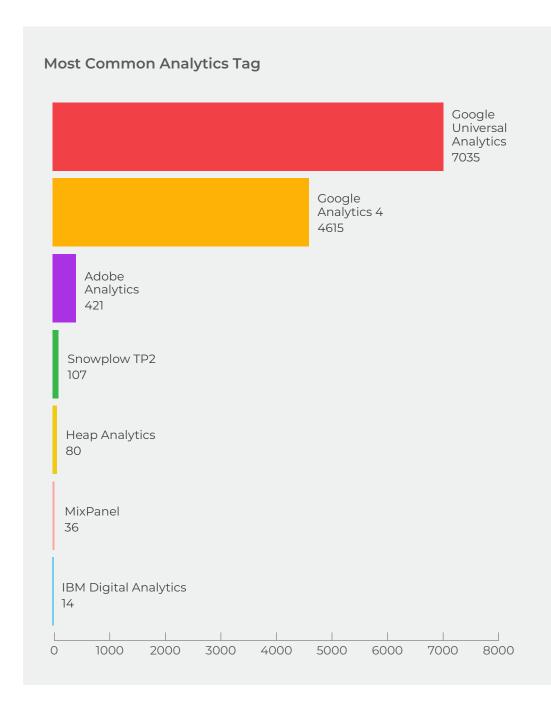
D. How Many Homepages Had An Analytics Tag

Most businesses track on-site user behavior using Adobe Analytics, Google Analytics, or other less common web analytics platforms. However, we were surprised to find that only 71.7% of the home pages in our sample had an analytics tag present, which seemed remarkably low.

The main analytics tag should be on every page (or at minimum the majority of pages) of a website to provide accurate behavioral and performance measurement. If and where those tags were to be missing, you would expect it to be on lower-priority pages or sections - certainly not the homepage itself!

1. Most Popular Analytics

Google Analytics continues to be far and away more prevalent than Adobe Analytics, which makes sense since Google Analytics is most often licensed for free. It's interesting to see Snowplow, a less common web analytics platform data has been deployed 107 times, as well as Heap, a customer journey analytics tool, of which we counted 80 tags.



2. How many Analytics are companies running?

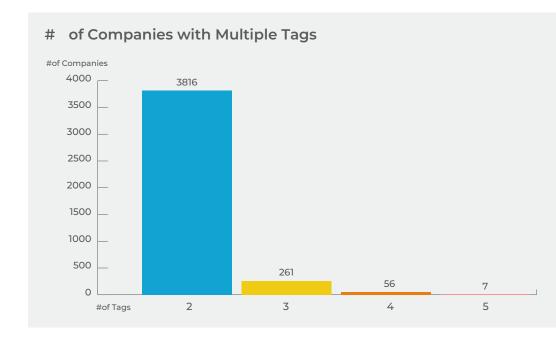
Of the 8,153 companies with analytics tags:

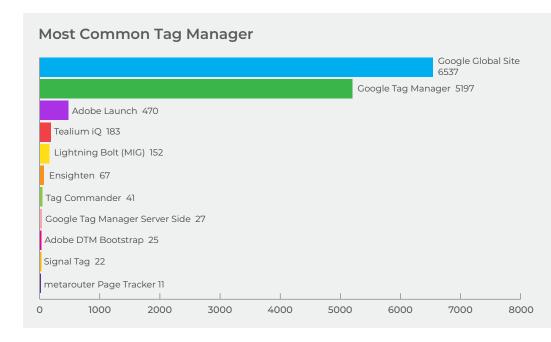
- > 4,140 companies were running at least 2 analytics tags (51%)
- The most number of analytics tags one company was running simultaneously was 5
- > We only found 7 companies with 5 analytics tags
- > 3,859 companies were running UA and GA4 simultaneously (47%)
- > 203 companies were running Adobe Analytics and Google Universal Analytics (2.5%)

E. Most Common Tag Manager

Tag Management Systems (TMS) help marketers and analysts deploy and manage their tags without IT or development's help. Managing tags through a TMS is a best practice, as it helps marketing departments stay agile, speeds up page performance, and makes changes and updates to tags much more efficient. A TMS allows you to deploy a container code and set rules for when each of your tags should fire, so you don't have to deploy each tag individually. In addition, a TMS is necessary to set up Consent Management Platforms (CMP).

If you don't have a TMS, then you're introducing the potential for human error, entropy, and old campaign residue to muck up your website over time.

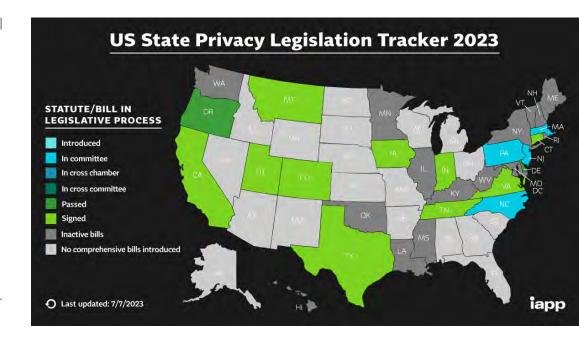




We see the same issue as before where Google Global Site and Google Tag Manager are the most popular tag managers, but they probably shouldn't be on the same pages. There were also 25 instances of Adobe DTM which has been deprecated since April 2021, so that's probably causing broken tag requests on those sites and should be removed

F. How Many Had a CMP?

We wanted to check how many companies employ a Consent Management Platform (CMP). Consent Management Platforms allow you to collect and manage user consent, display cookie banners on your website, and make sure the right kind of cookies are placed based on your visitor's consent preferences.



We only found 17.5% of websites with a detectable CMP. This appears to indicate that we have a long way to go with privacy compliance.

It's been 5 years since GDPR came into effect in Europe, but large organizations are still struggling to get into compliance as evidenced by continuing headline-grabbing fines. The US has seen increasing privacy laws, state-by-state. You can track these state laws via IAPP's legislation tracker map.

At IAPP Global Privacy Summit in D.C. this year, we noticed a slight maturation of the privacy space for the U.S. market. Booth visitors and attendees we conversed with were aware that they needed to take website privacy compliance seriously, even if they didn't yet know how exactly to do it.

Organizations are looking for tools that can help them see all the moving parts of a continually changing website where data collection is happening on a regular basis. Having a Consent Management Platform to obtain and manage user consent would be a first step. Having a web governance strategy to monitor your CMP would be the next.

1. Most Common Consent Manager

The most common CMP that we detected was the category leader OneTrust. Cookiebot came next, with TrustArc a distant third.

2. By Region

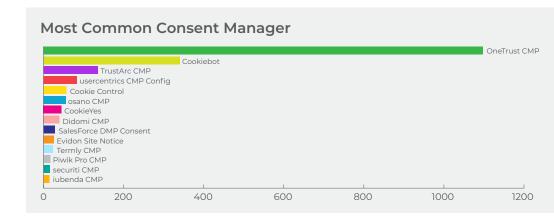
When we split the results between EMEA and North America, the third place CMP changed. Usercentrics is far more popular in Europe, while TrustArc is concentrated in the United States.

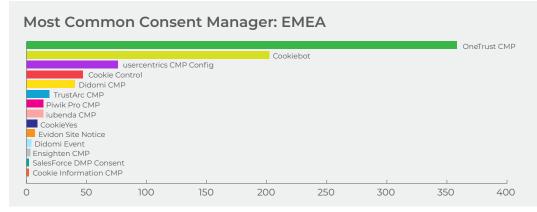
Follow us to the last section of the report, where we'll be continuing the analysis on cookies.

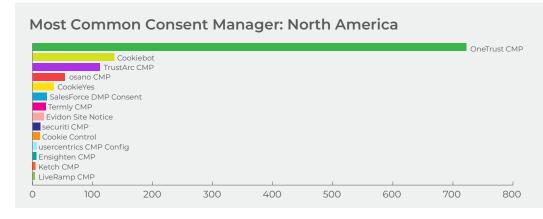
Cookies

A. Cookie Breakdowns

Cookies are persistent pieces of information that websites give to your browser to store on your hard drive until it expires so that the site can remember you when you return. Common use cases of cookies are logins, preferences, and shopping carts. We were surprised to find that 8.42% of companies scanned had no cookies. We checked if certain regions affected the percentage breakouts.







APAC did report more sites without cookies with 17.5%.

EMEA was 11.78%.

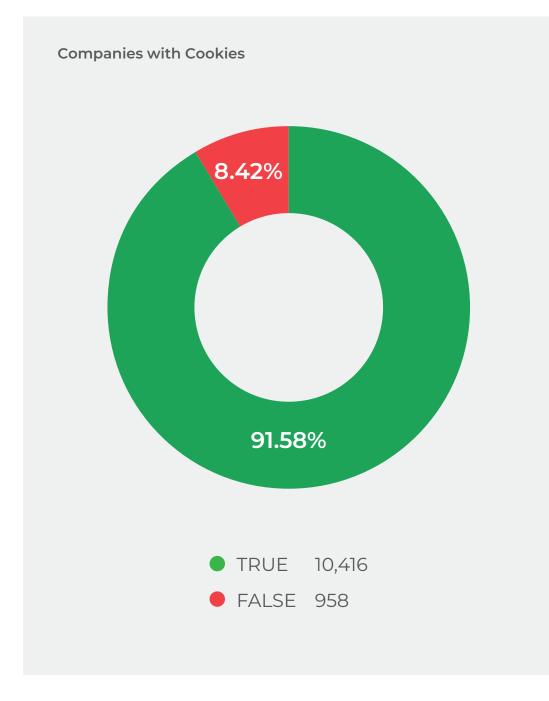
LATAM had the most cookies with only 3.3%.

NA came in just below average with 7.5%.

Our educated guess here is that certain sites are actually taking a conservative approach and not dropping cookies immediately. If that's you, let us pat you on the back!

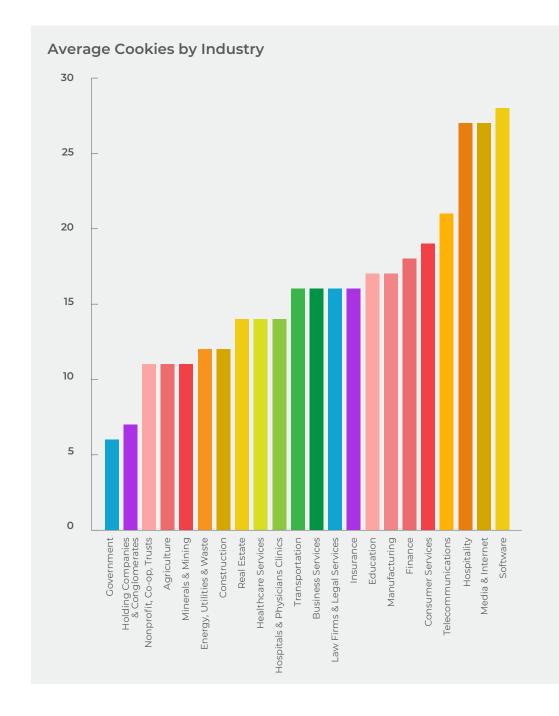
The average number of cookies was 17.75. Our report from 2021, Counting Cookies: The Reliance, the Risk, and the Remedy found about 20 average cookie domains per site, so this is in line with previous data. Having more or less cookies is neither good or bad, it should just align with the needs of your business and tech stack.

Industry-wise, the least consumer-facing industries had the least number of cookies, which predictably mirrors the tags breakdown by industry. Media, Retail, and Software companies all have over 25 cookies on average.



Who had the most cookies?

- > 274 cookies was the highest number from a Retailer
- The next four places went to Media & Internet companies with 210, 208, 203, and 196 cookies
- The top 25 most cookies (a range from 139 to 274 cookies) were companies in these industries:
 - Retail (sub-categories: Grocery, Sporting Goods, Consumer Electronics, Home Improvement, Auto, Apparel)
 - Media & Internet (sub-categories: Newspapers, Broadcasting, Data Collection, Publishing)
 - Manufacturing (sub-categories: Cosmetics, Food and Beverage, Household Goods, Furniture)
 - · Transportation (sub-category: Cruise)
 - · Hospitality (sub-category: Restaurant)
 - · Transportation (sub-category: Cruise)
 - Hospitality (sub-category: Restaurant)



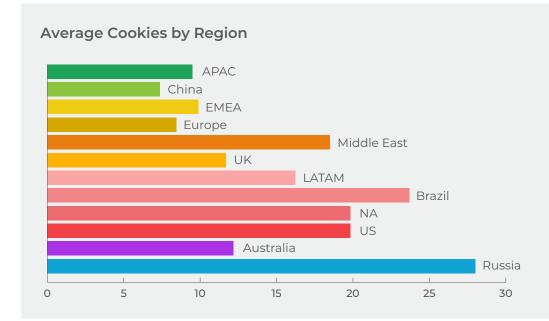
Average Cookies by Region

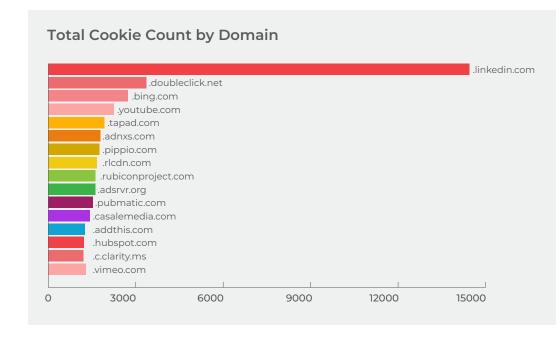
We cross-checked by region and then called out some countries to see what we could find.

Well, would you look at that? It's the U.S., the Middle East, and Brazil that are bringing the number of cookies up, with nearly twice as many as European or Asian companies. Wait a minute, look at Russia! That's a lot.

Total Cookie Count by Domain

Almost every company site had a LinkedIn cookie, which makes sense since most companies use LinkedIn for ads or recruiting. LinkedIn generates lots of cookies as you can see in this chart.







Most Common 3rd-Party Cookie

But, if we're looking at it by domain and not sheer number of cookies, then Doubleclick, the largest provider of Internet advertising, is the most frequently set cookie domain.

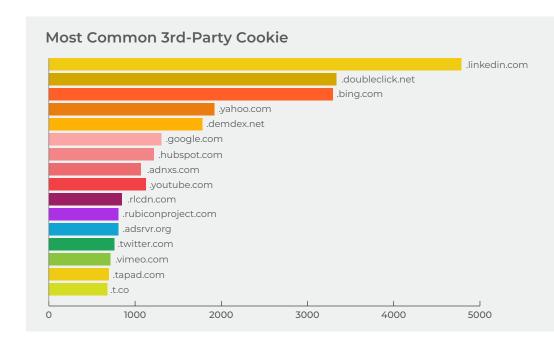
These are only 3rd-party cookie domains, and it's not uncommon for certain technologies, like analytics platforms, to be implemented in a 1st-party context, since they are being set by the domain of the website being visited. Since the vast majority of cookies being set by 3rd-party domains are in the advertising ecosystem, it makes sense that DoubleClick is the most popular. Let's look at 1st- and 3rd-party cookies further.

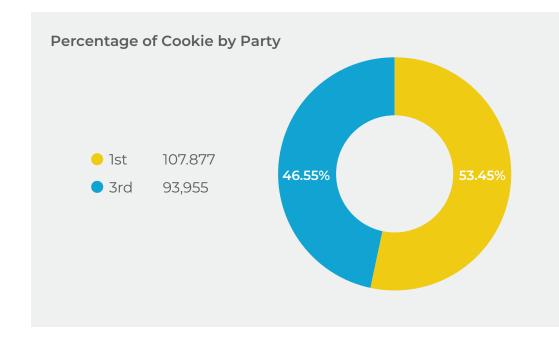
B. 1st-party V. 3rd-party Cookies

Average 1st-Party Cookies: 9.49

Average 3rd-Party Cookies: 8.07

As 3rd-party cookies get deprecated by browsers, companies need to be aware of how many of these cookies they have, so they can start to phase out their reliance on them. A good rule of thumb is any time you have more than a 50/50 split (more than 50% of your cookies are 3rd-party), then it's worth taking a deeper look at your overall cookie inventory.





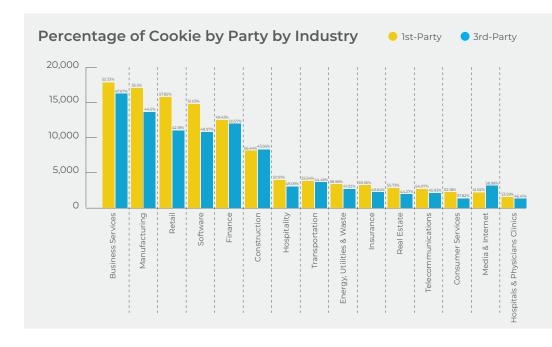


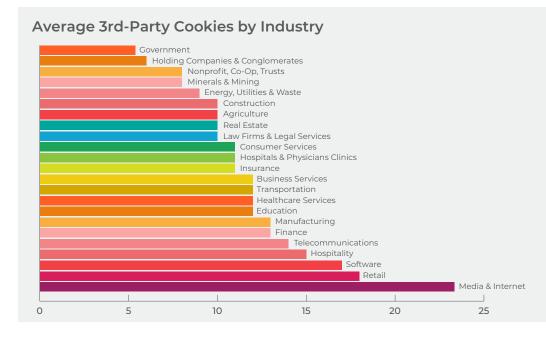
As you can see, most companies are still heavily reliant on 3rd-party cookies, which is not that surprising since Google Chrome has been postponing its deprecation plans, and no one yet knows what cohorts or walled gardens could really replace what 3rd-party cookies provide. However, if you want to get ahead of this coming cookie-geddon, then trying to wean your organization off 3rd-party cookies would be wise. For more information on alternatives to 3rd-party cookies, take a look at our report

Counting Cookies

1. 1st and 3rd-Party Cookies by Industry

Business Services, Manufacturing, and Retail have the most cookies overall when we break it down by industry. This makes sense since advertising and marketing fall under Business Services and a lot of retail brands are in Manufacturing. These consumer-facing verticals rely on personalization and tracking to deliver relevant content to visitors. Taking a closer look, we found that Consumer Services, Insurance, and Energy are doing the best at relying less on 3rd-party cookies! Great job, guys, we wouldn't have predicted that.



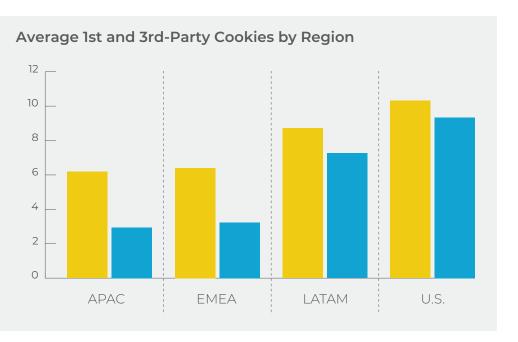


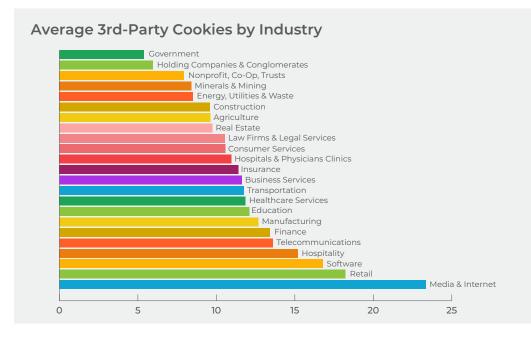
2. Which Industry has the Most 3rd-Party Cookies

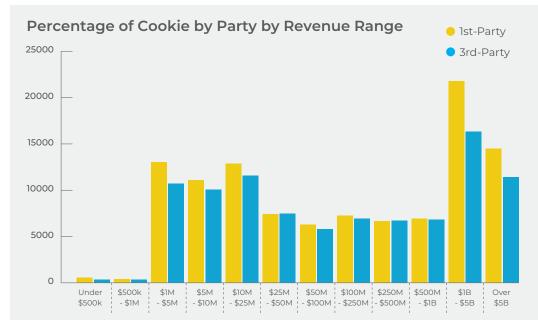
Not surprisingly, Software, Retail, and Media & Internet companies rely heavily on 3rd-party cookies.

3. 1st and 3rd-Party Cookies By Company Size

Enterprise-level companies have the most cookies but are doing better at moving to 1st-party. It's interesting that it gets a little messy with third through sixth place going to much smaller companies in the \$1M-\$50M range.





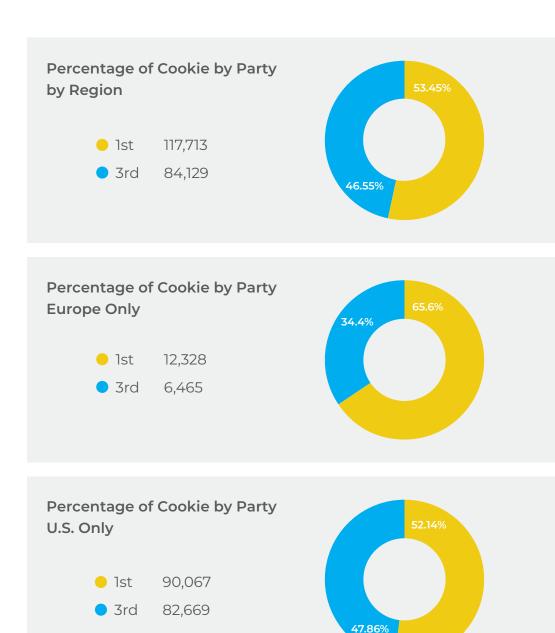


4. 1st and 3rd-Party Cookies by Region

Europe is much further ahead of the U.S. in terms of trying to not only have less cookies, but also have less 3rd-party cookies. (Assuming our scan from U.S. servers is showing us a default state that doesn't necessarily correspond with a European opted-in state.) American companies are still very much evenly distributed between 1st and 3rd-party cookies.

C. Non-secure Cookies

If you have seen our **Cookie Cheat Sheet**, you know that cookies will have a secure setting that is either on or off, which means that your browser will only send a secure cookie if the browser is using HTTPS. HTTPS means you're using an SSL or encrypted channel to communicate with that website as a security measure because there's possibly sensitive data in your cookie. Since Google made HTTPS the default protocol for their search engine results, it's best practice to ensure all cookies are secure. So these 41.26% of non-secure cookies are not following best practices.





D. # HTTPOnly

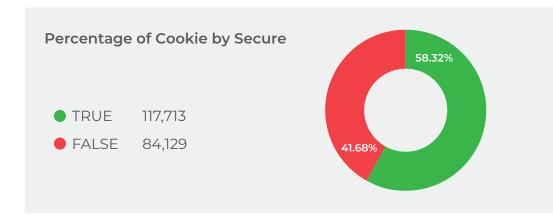
HTTPOnly protects against a class of security vulnerabilities called cross-site scripting. This is when malicious scripts are injected into trusted websites which makes different end users execute them in their browsers. HTTPOnly is a common practice to disable the JavaScript code of a website from looking at a cookie's content. Essentially, the cookie should only be read by HTTP. 85% of the cookies detected did not employ this setting, which is a best practice.

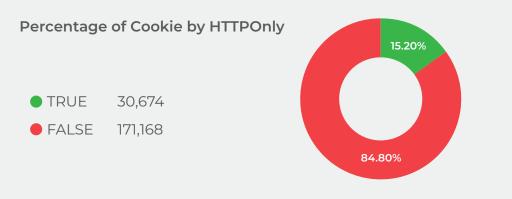
E. Empty Samesite

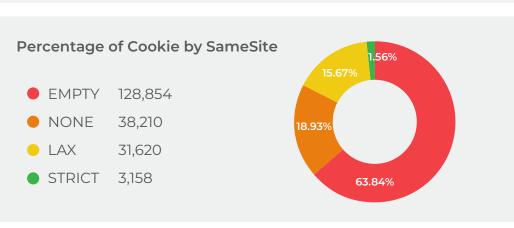
The SameSite feature defines whether browsers should include cookies in requests that go from one domain to another, like an image or form submission. SameSite allows you to block cross-site request forgery attacks, where users can be tricked into requests like transferring funds or changing their address while authenticated on a website. Best practice would be to define a level of security, but as you can see in this chart, 90% of cookies detected are not taking advantage of this security measure.

F. Egregious Cookies

For our final analysis...drum roll, please...we wanted to see how many cookies had all three of the above problems at the same time.







Over half of the homepages scanned, 7,159 out of the 11,374 companies (62.9%), had cookies with all three problems.

The worst offender was a retail clothing company with 47 of these cookies that had non-secure, HTTPOnly of false, and empty SameSite.

If we filtered in only 3rd-party cookies, there were 273 companies with all three problems in 3rd-party cookies. The final champion was a software company with 33 cookies that were 3rd-party, non-secure, false HTTPOnly, and empty SameSite.

Conclusion

If you've managed to reach the end of this report, then you are a legend. Thank you so much for your time and attention. We hope that the data we presented was useful, interesting, and inspirational. Here are some key takeaways:

Pages

- A. All industries could improve their page load times.
- B. Almost 30% of English-language homepages didn't seem to have the word "privacy" on their page, so they should check on their privacy policy accessibility.

Tags

- A. We are all still relying too much on 3rd-party tags, which isn't a big surprise.
- B. Almost half of the pages scanned were running UA and GA4 simultaneously, which is great.
- C. But, 3,904 companies were running Google Global Site and Google Tag Manager, which is not good, so stop doing that one.
- D. Only 17.5% of scanned homepages had a detectable CMP. We still have a ways to go in terms of consent management adoption.



Cookies

- A. As with tags, we're still heavily reliant on 3rd-party cookies, especially the most consumer-facing industries.
- B. Larger enterprises had more cookies but were best at relying more on 1st-party cookies which seems logical. Surprisingly, smaller companies were doing better than mid-sized companies at moving to 1st-party.
- C. Regionally, APAC and EMEA have fewer cookies than LATAM or NA.
- D. Almost 63% of the homepages we scanned had cookies with all three metadata problems: non-secure, false HTTPOnly, and empty SameSite.

This report relied on ObservePoint's ability to scan a breadth of pages. If you'd like to see the depth of the information we could provide on your own website, please reach out for a demo or free trial.

