

# How Bots and Fake Users Skew Marketing Data and Analytics

Analyzing the impact of invalid traffic on key site and engagement metrics.



**CHEQ**

# Table of contents

Chapter 1: Introduction

3-5

---

Chapter 2: Report Findings

6-24

---

Chapter 3: Impact & Implications

25-28

---

Chapter 4: Conclusion

29

---



## Chapter 1

# Introduction



### **Lior Frenkel, Chairman, the Israel Cyber Forum**

Recent years have seen advanced cyberattacks become the norm. Attack techniques, tools and technologies, traditionally solely accessible to governments, have leaked out. Today, they are readily available to serve criminal agents and support the rapidly growing cybercrime sector.

This trend, which will not end in the foreseeable future, is clearly demonstrated by the proliferation of ransomware attacks, advanced attacks by non-state actors, a focus on national critical infrastructures, and the growing sophistication of attacks with criminal motivations. Within digital marketing and ecommerce, it is evident that this rise in adversaries' cyber capabilities is resulting in direct, and evident revenue loss to businesses across the globe. The impact is stark: producing higher customer acquisition costs, which end up reducing the bottom-line performance.

Furthermore, fake user accounts and bots, blur the view of organizations, skewing forecasts, revenues and pipeline. This represents a growing cyber security concern. This is a challenge that must be dealt with, sooner, rather than later.



**Within digital marketing and ecommerce, the rise in adversaries' cyber capabilities is resulting in direct revenue loss**

# The Importance of Data & Analytics

We live in a world today where every business decision is made based on data. Organizations invest millions in business and marketing intelligence software, analytics tools, analysts, and data experts. And all of this investment is done with the intention of getting an accurate view of the business, which will then help drive good business decisions.

The problem is, since [an increasing amount](#) of web traffic is made up of bots, automation tools, headless browsers and fake users, all of that great data gets contaminated. BI gets skewed and teams end up making decisions based on inaccurate data including inflated site traffic numbers, misleading conversion rates, inaccurate engagement metrics and polluted CDP segments.

Because of this, skewed data has become a strategic issue for data-driven organizations, not only because it nullifies their massive investment in BI and analytics, but because it undermines virtually every decision they make. Throughout this report, we explain exactly how important metrics are impacted, and what the world would look like if fake traffic were to be removed.



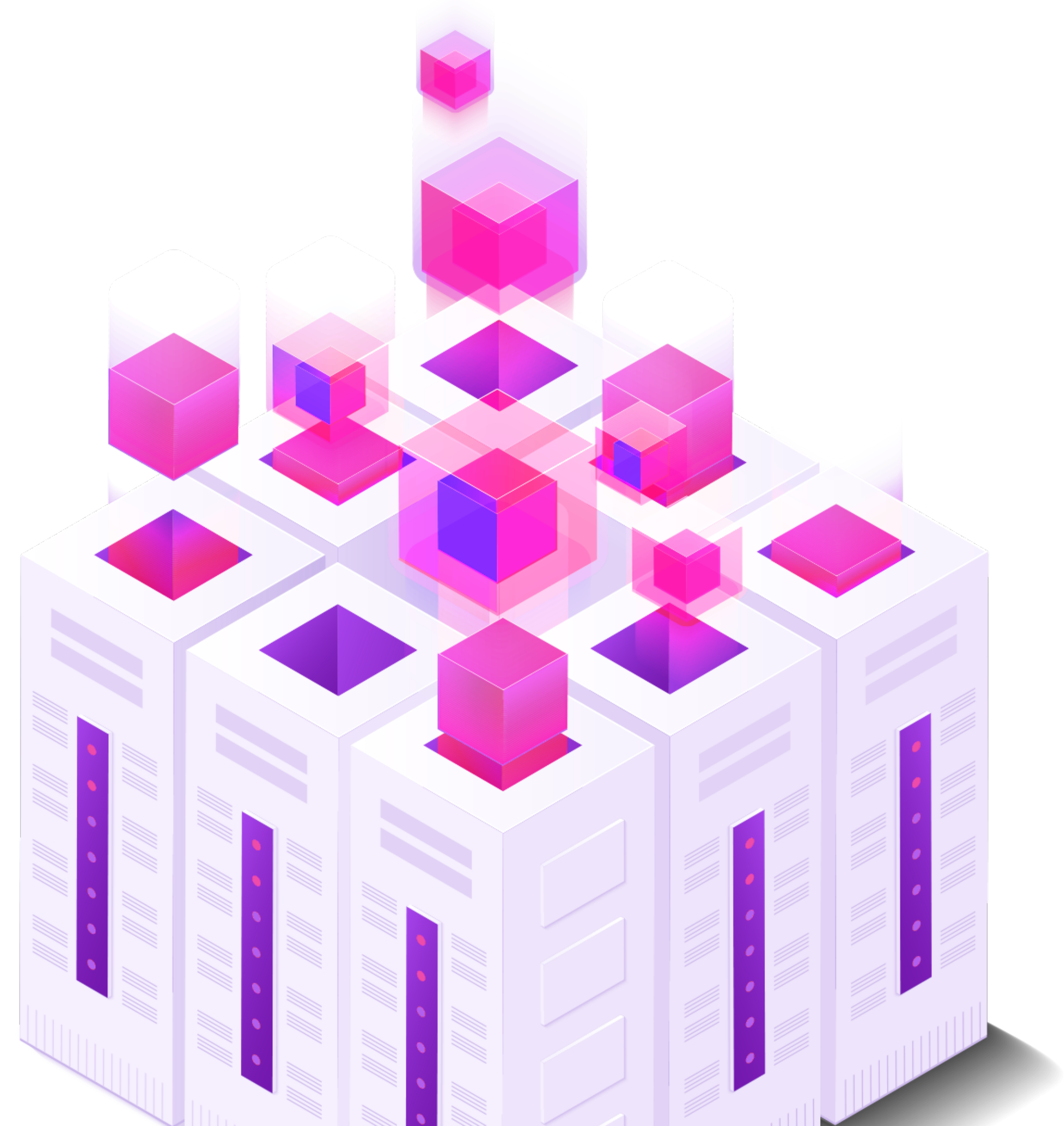


# Foreword & Methodology

Today, it is widely known that many so-called internet users are actually bots and fake users. News continues to spread of instances of fraud and malicious activities, and [massive business deals](#) are being stalled because of discrepancies surrounding fake traffic volumes. Unfortunately for businesses and consumers alike, we are in the era of the Fake Web.

As the leader in Go-to-Market Security, CHEQ frequently studies bot activity originating from various organic and paid traffic sources on the web to determine the validity of each user. This time around, we decided to analyze the degree to which critical business metrics are skewed. We started by studying the activity on more than 50,000 websites and deployed more than 2,000 cybersecurity challenges on each website visitor to ultimately reveal whether they were a bot or bad actor, or a legitimate human user.

We then measured that data against several key metrics for each site, and more, and how those metrics would change if the fake traffic were to be removed. Furthermore, the report infers the business impacts of those numbers.



## Chapter 2

# Report Findings



**Here we discuss how bots and fake users are impacting key high level metrics.**

The data is broken down in pie charts to illustrate how much these metrics are skewed by invalid traffic. Consider this section to be a visual representation of “turning on the lights” for key metrics - revealing the true numbers that the internet experiences from human users rather than from bots.

**Specifically we studied:**

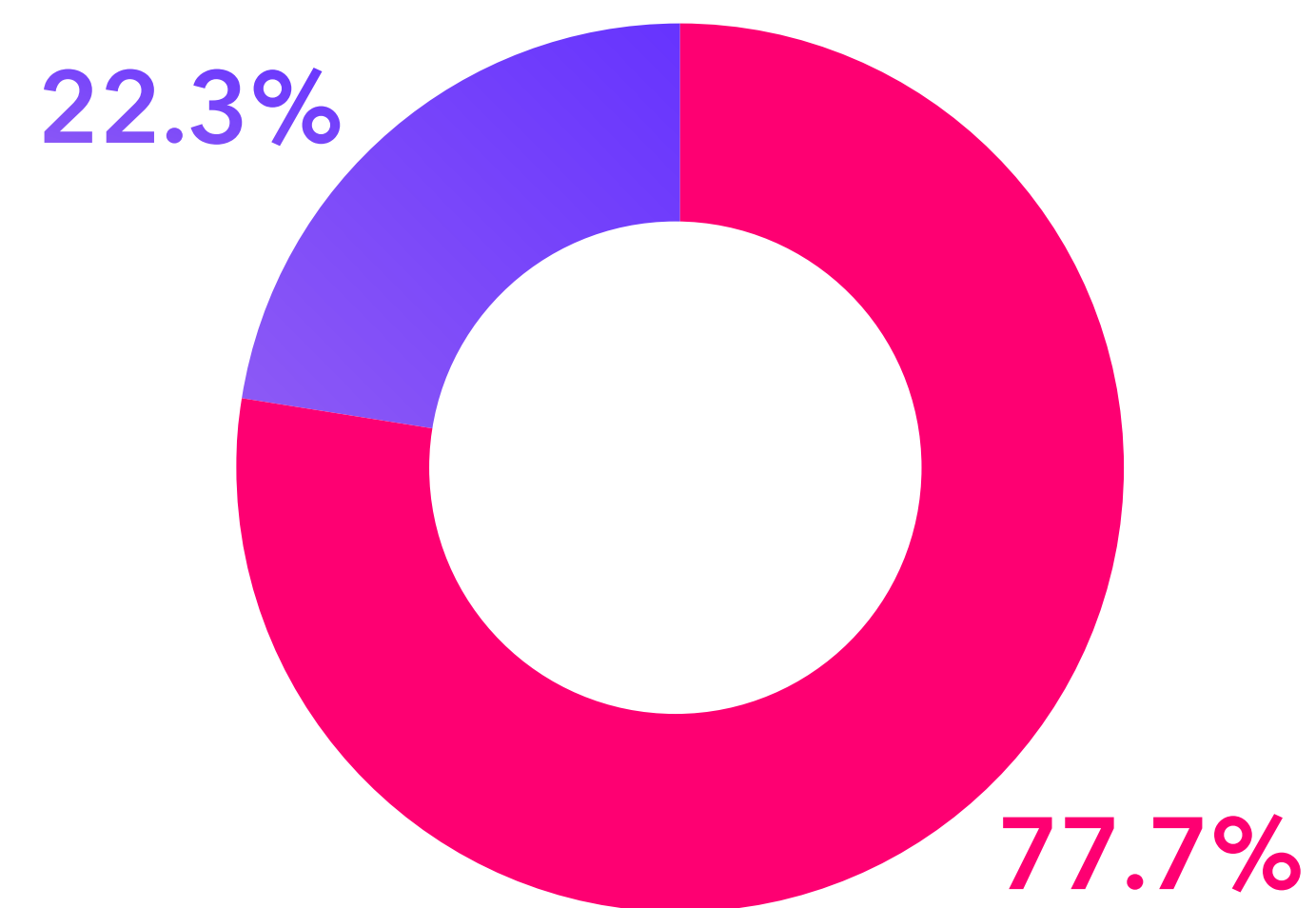
- **Unique Site Visits**
- **Pageviews Per Session**
- **Bounce Rates**
- **Session Duration**
- **New Users**
- **Returning Users**



# Unique Site Visits

### Unique Site Visits are skewed by 22.3%

When marketers and analysts are 'in the dark,' their data assumes that all unique visitors are real humans, our data shows that 22.3% are bots on average. Unique site visits are defined as a session on a website originating from a single user or source. This data shows that 77.7% of unique site visits actually come from real human users on average.



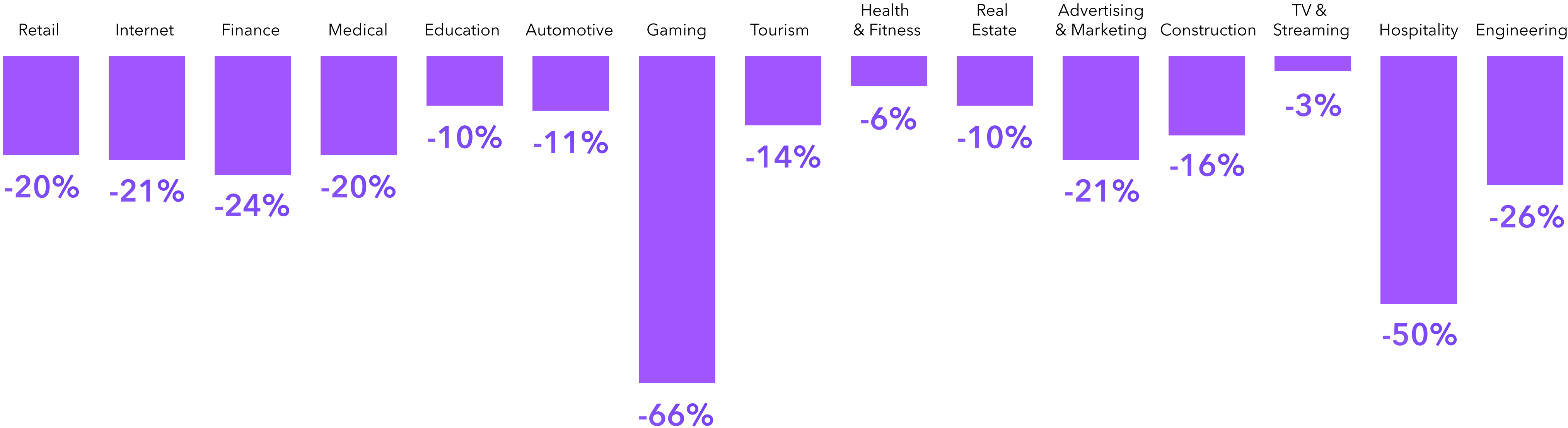
Chapter 2 - Report Findings

# Unique Site Visits

## By Industry

When removing fake traffic from overall unique site visits, the site visits dropped by 22.3% on average. As is presented in the chart below, the metrics varied greatly per each industry. For example, in the gaming industry, visits dropped by 66%.

This suggests that a majority of 'unique visitors' to gaming and gambling sites are actually bots and fake users. While there are major differences in each sector, the analysis shows that all industries experience some level of skewed unique site visit data.



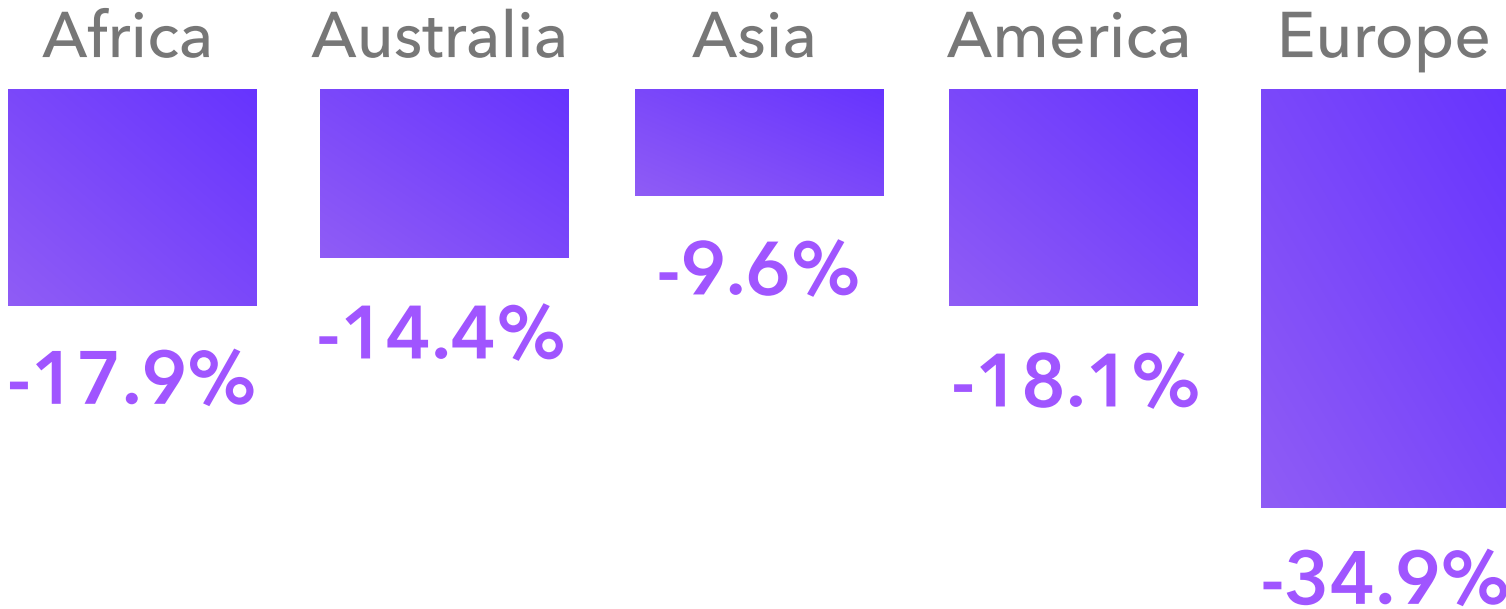


# Chapter 2 - Report Findings

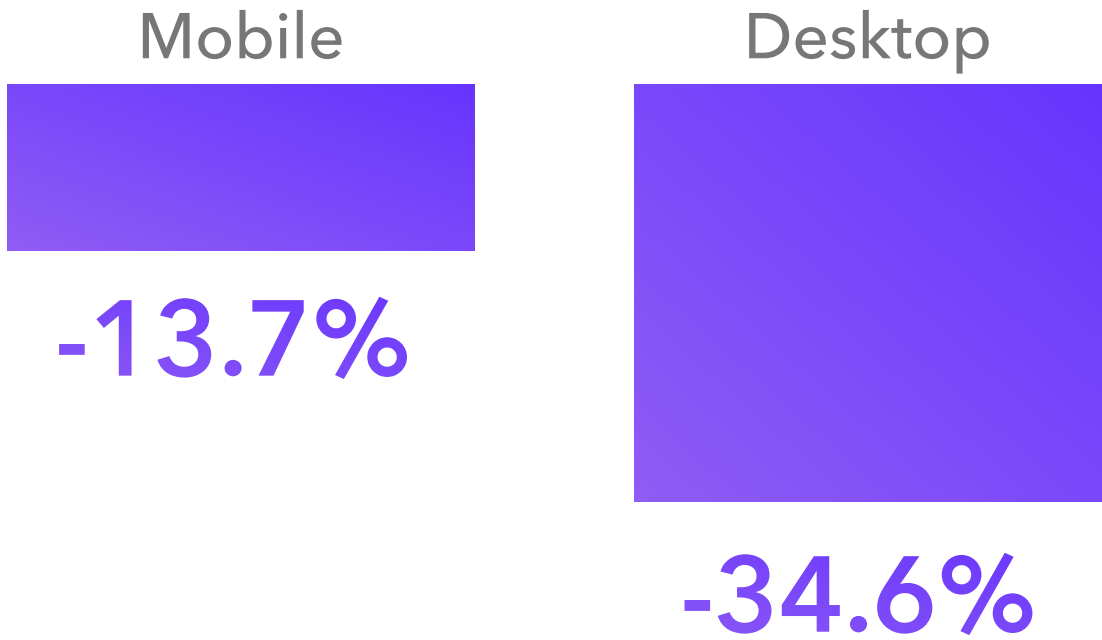
## Unique Site Visits

### By Geography

When looking at fake traffic rates by geography, Europe saw the most significant reduction in unique site visitors when bots and fake users were removed. Asia saw the lowest, but still significant rate of fake unique visits at almost 10%.



### By Device



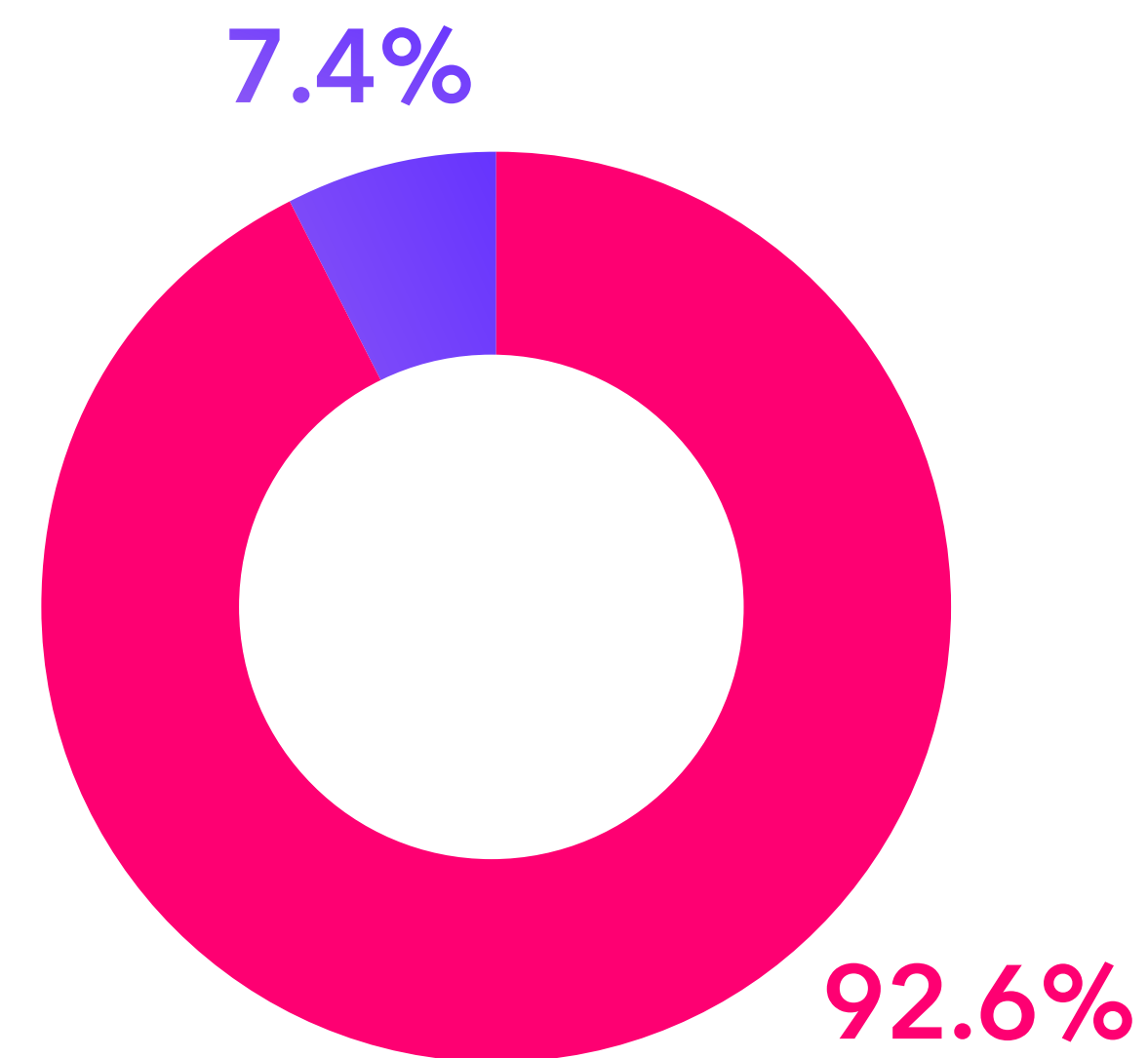
While [mobile website activity](#) steadily increases, the amount of unique site visits from bots on mobile remains significantly lower than that on desktops. This implies that while humans continue to access the internet via mobile at high rates, bots and fake users tend to visit websites from desktop devices.

## Chapter 2 - Report Findings

# Pageviews

### Pageviews are skewed by 7.4%

If fake traffic were to be removed, pageviews would increase by 7.4%. In other words, bots and fake users were found to have a lower pageview-per-session count overall than legitimate humans. This reveals that real human users typically browse more site pages than bots. If site pageviews per session are unusually low, it could be an indicator of increased fake traffic presence.





# Chapter 2 - Report Findings

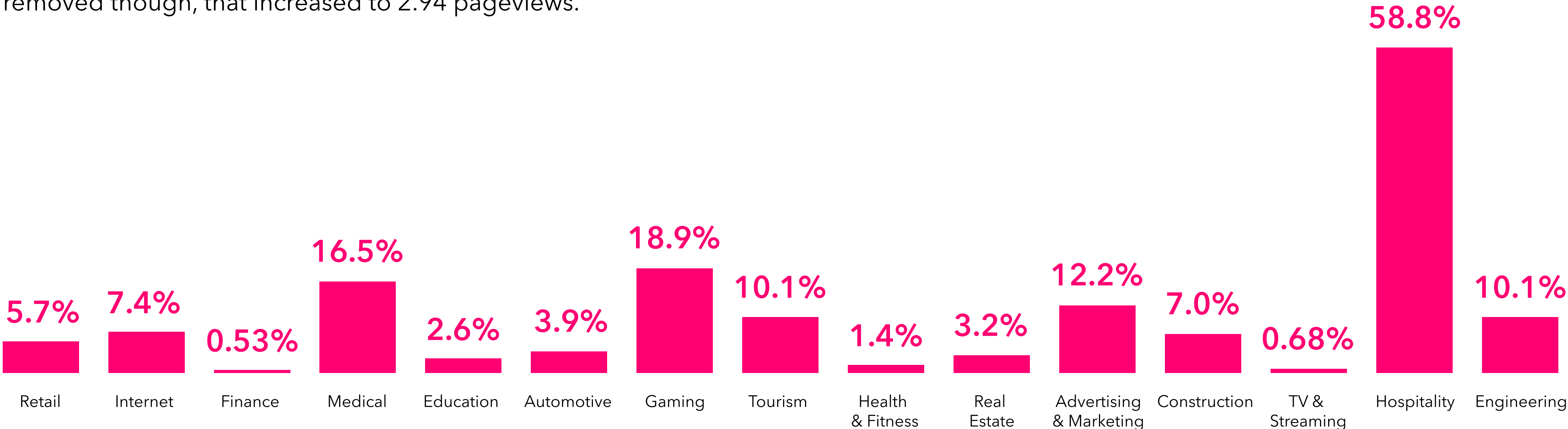
## Pageviews

### By Industry

The charts below indicate how much average pageviews per site visit increased when fake traffic was removed.

For example, the average number of pageviews per site visit overall was initially 2.74 pageviews. When the fake traffic was removed though, that increased to 2.94 pageviews.

That is a total of a 7.4% increase in pageviews. The hospitality, medical, advertising & marketing, and gaming industries saw the sharpest differences between their initial pageviews and how they increased when bots and fake users were removed.



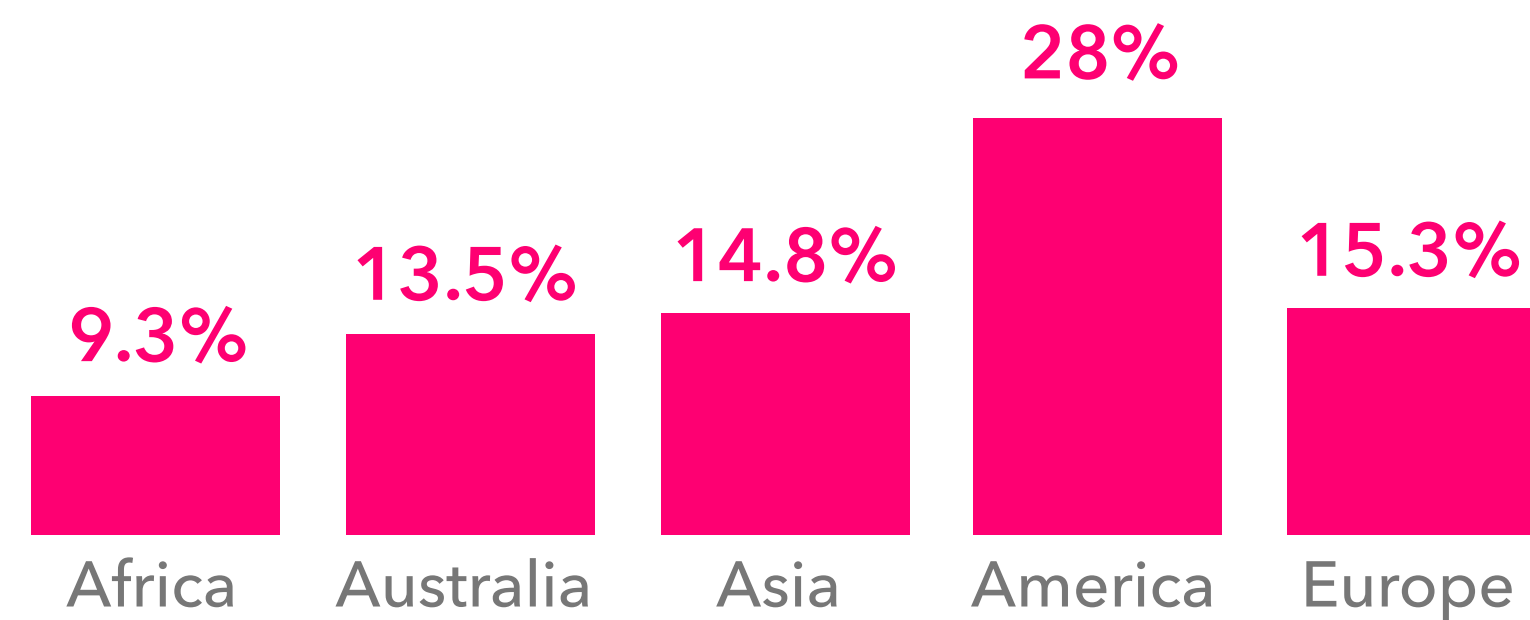
## Chapter 2 - Report Findings

# Pageviews

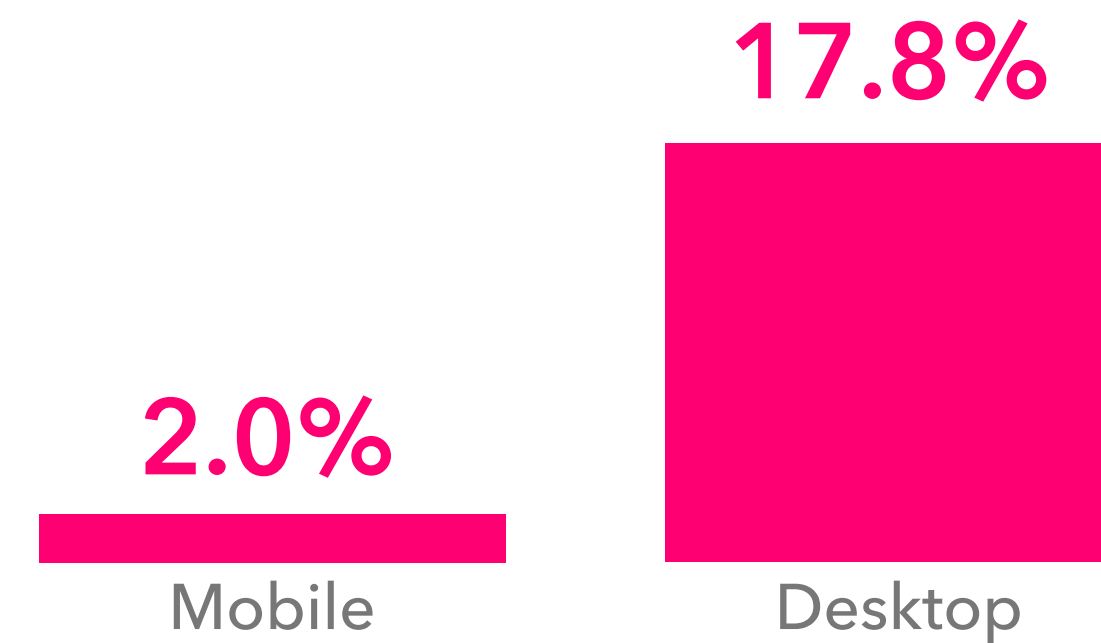
### By Geography

When it comes to geography, America saw the biggest increase in pageviews when fake traffic was removed.

While this region hovered in the middle for fake unique site visits, the actual number of pageviews was significantly impacted.



### By Device



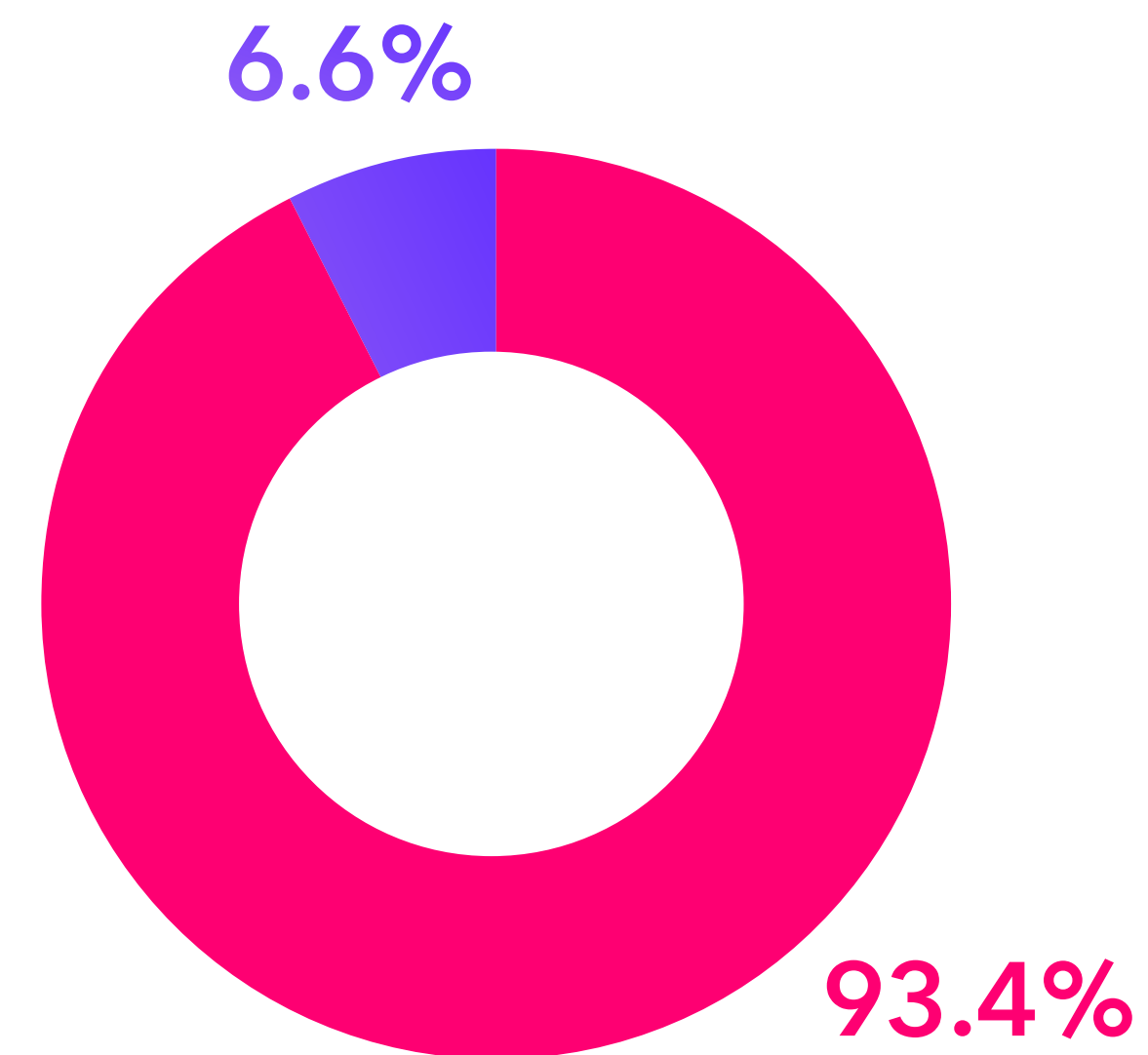
The difference between mobile and desktop fake unique site visits was considerably different, and so is the number of fake pageviews for devices. This implies that more fake site visitors access sites via desktop, and don't view as many pages as real site visitors.

## Chapter 2 - Report Findings

# Bounce Rates

### Bounce Rates are skewed by 6.6%

Of all the sites we analyzed the average bounce rate was 54.3%, but when bots and fake users were eliminated, that number improved by dropping by 6.6%, taking the rate down to an impressive 47.7%. When bounce rates are unusually high, marketers may immediately make website changes or update campaigns, when perhaps the first thing to check is for bots and fake users.



# Chapter 2 - Report Findings

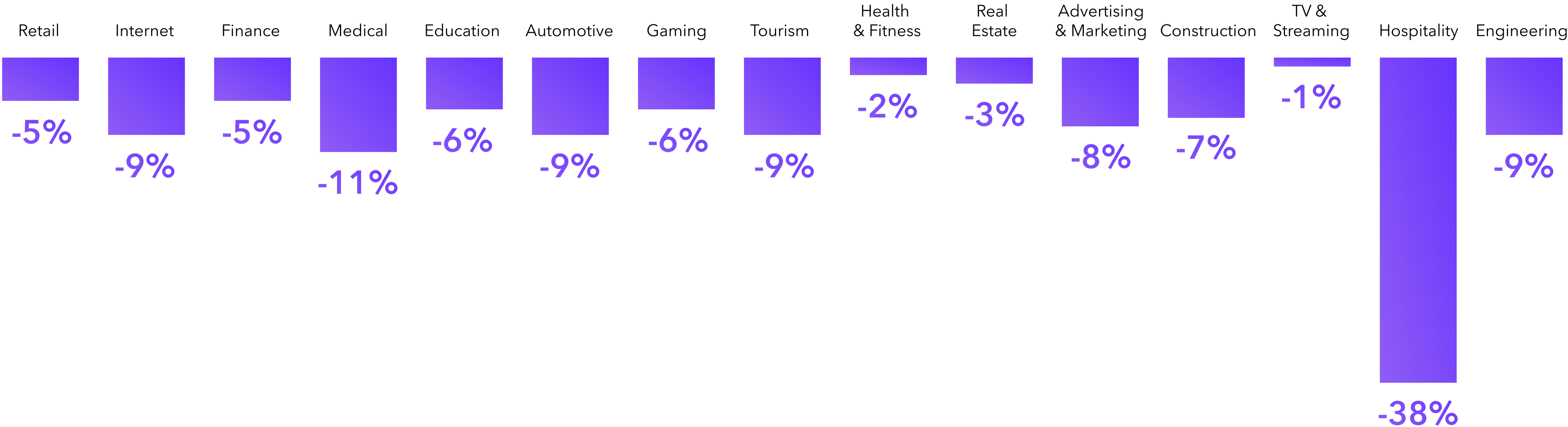
## Bounce Rates

### By Industry

When removing bots and fake users from bounce rate calculations, there was an average reduction of 6.7%.

The hospitality industry saw the largest decrease in bounces at 38%. This potentially indicates bots that quickly check hospitality sites for pricing and inventory information could

account for high bounce rates in the industry, which are then corrected when these types of threats are removed. Streaming and TV services see very little change in bounce rates, possibly since all users tend to spend considerable time on those sites and therefore do not bounce at high rates in general.



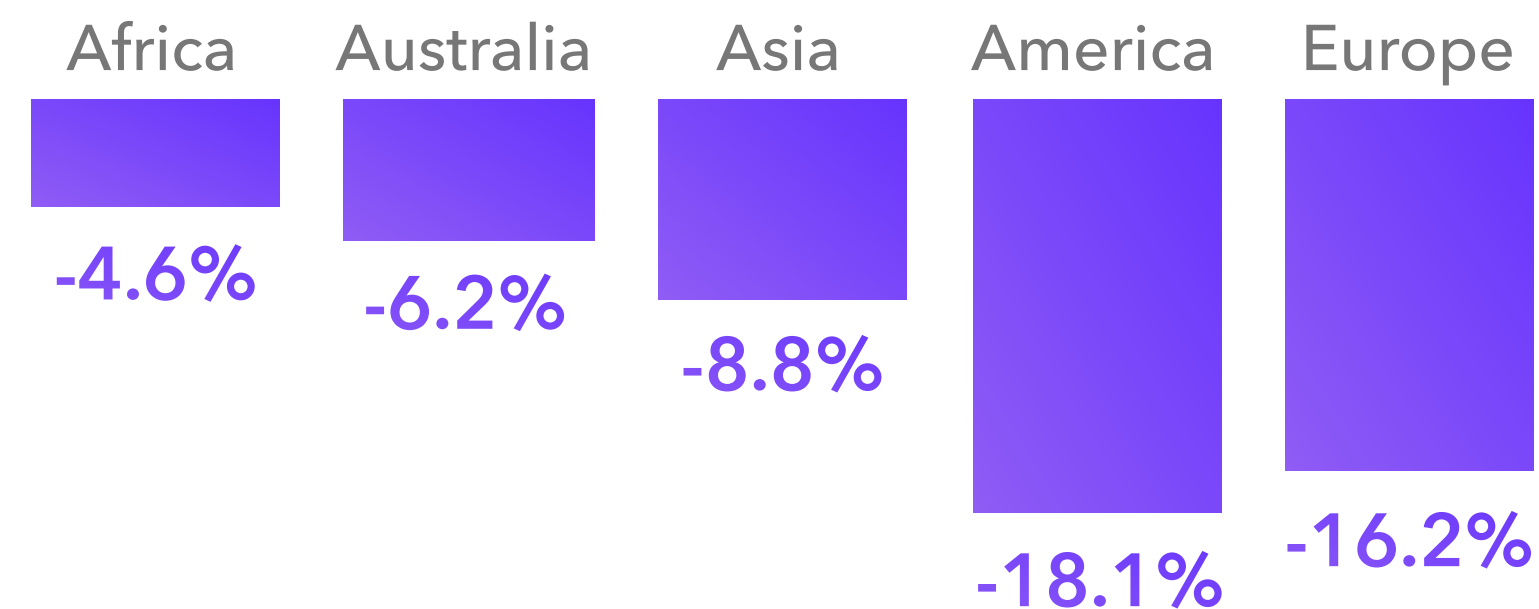


## Chapter 2 - Report Findings

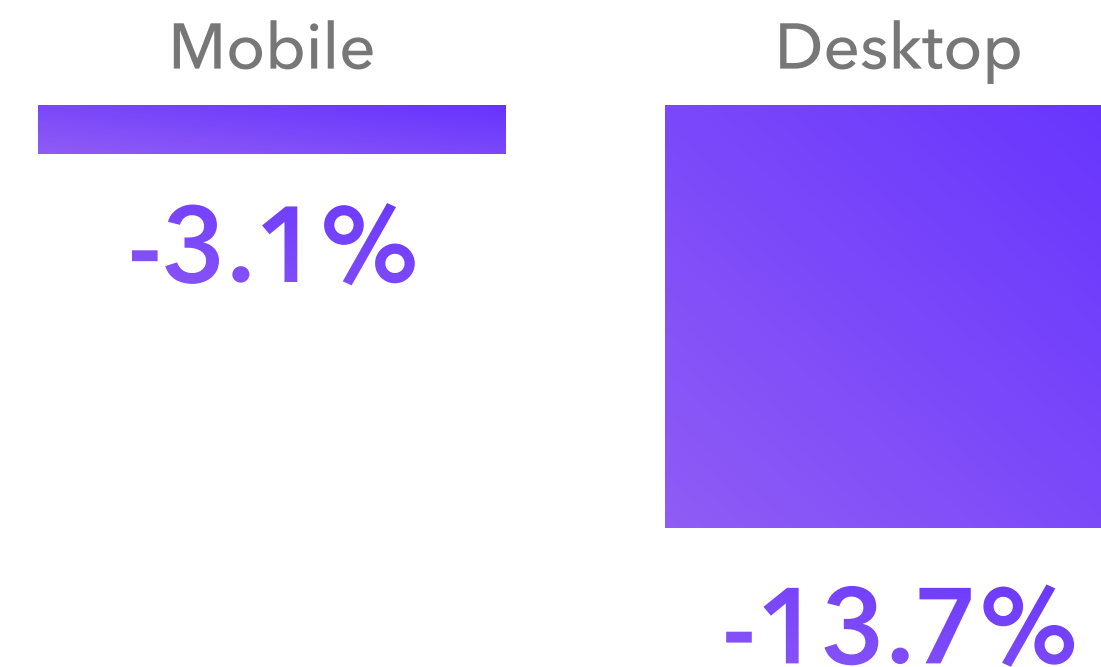
# Bounce Rates

### By Geography

Similar to pageviews, America also saw the biggest difference when fake traffic was removed from bounce rate metrics as well. Bounce rates and pageviews are somewhat related since a user cannot view multiple pages if they instantly bounce.



### By Device

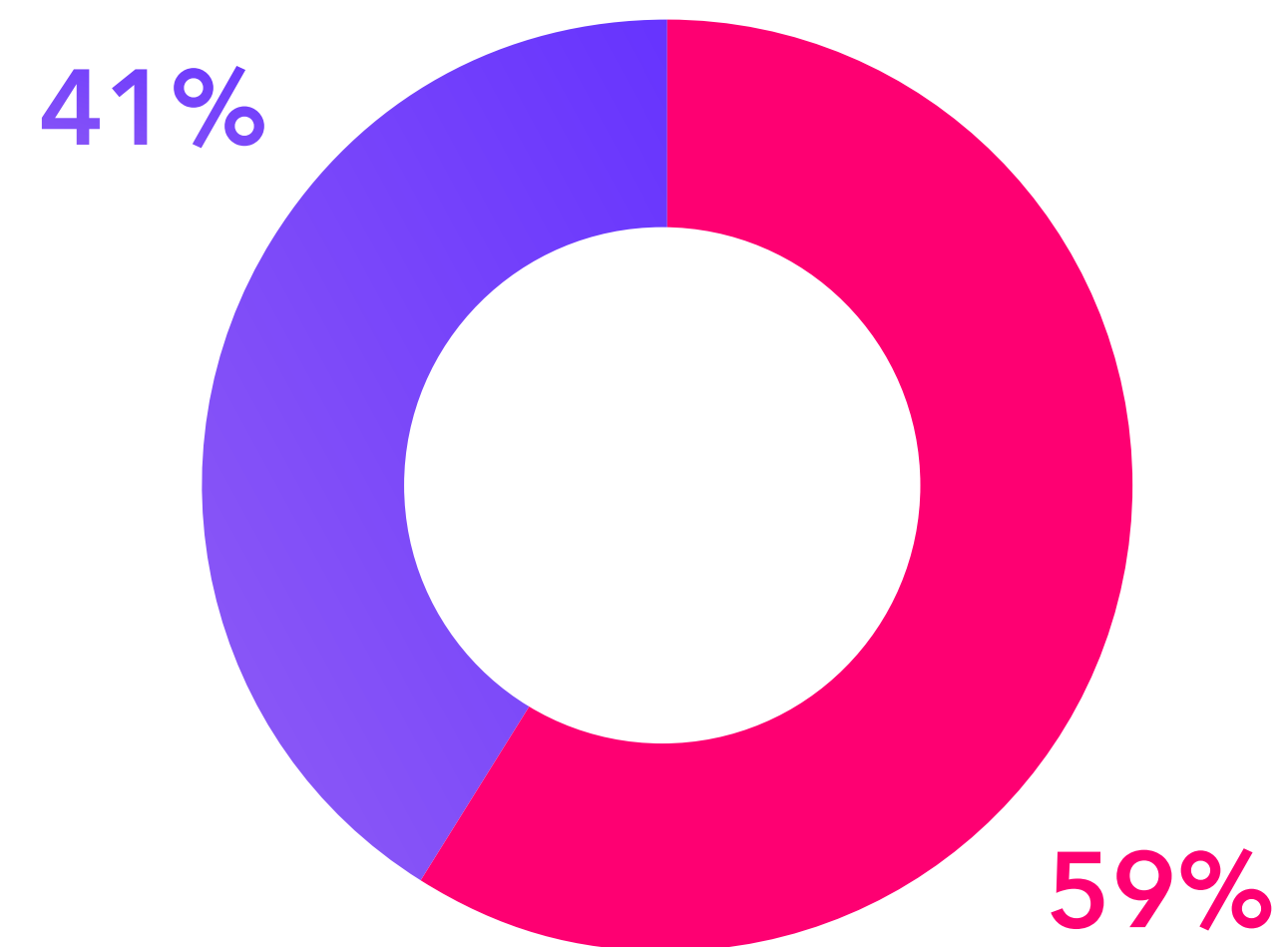


The difference between mobile and desktop fake unique site visits was considerably different, and so is the number of fake pageviews for devices. This implies that more fake site visitors access sites via desktop, and don't view as many pages as real site visitors.

# Session Duration

### Session Duration is skewed by 41%

Bots and fake users tend to have shorter session durations than legitimate users. In fact, the average session duration of fake users was found to be 129 seconds on average, while real users typically spend 183 seconds on a given website. This means that an unexplained sudden decrease in session durations could indicate a given website has a recent influx of bot traffic.



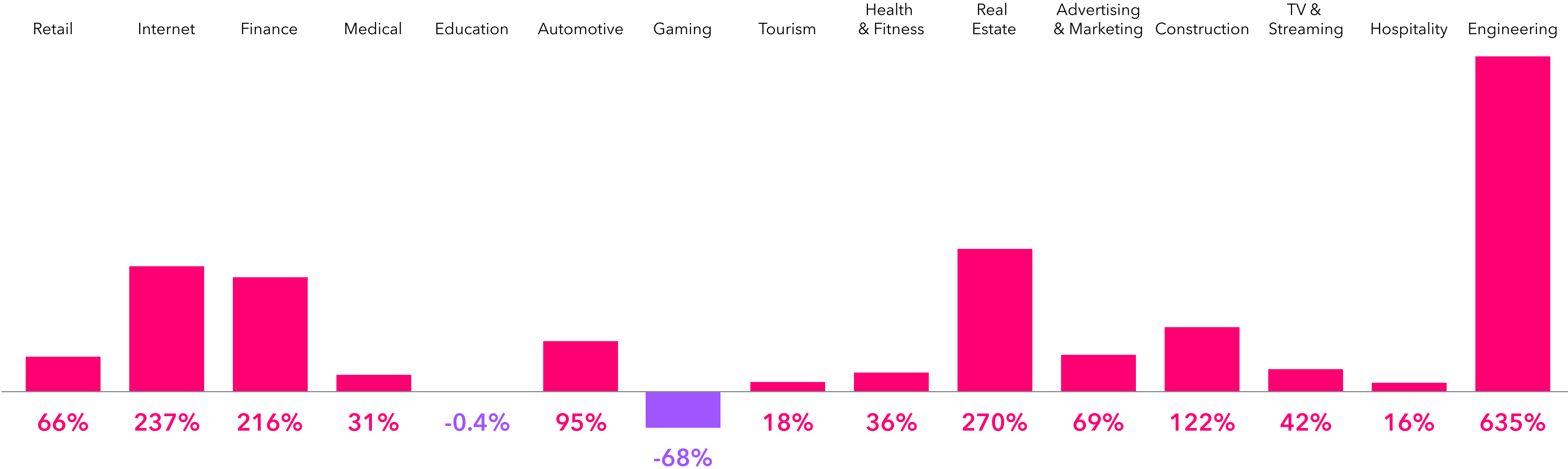
# Chapter 2 - Report Findings

## Session Duration

### By Industry

If bots and fake users were to be removed, session durations would improve across most industries. In the engineering industry for example, there would be a drastic 635% improvement.

While the duration of sessions would decrease in the gaming and education industries, the quality of those sessions would still increase because of the lack of presence of fake users. It is also important to remember that overall, there would be a 41% increase in session durations across the web.

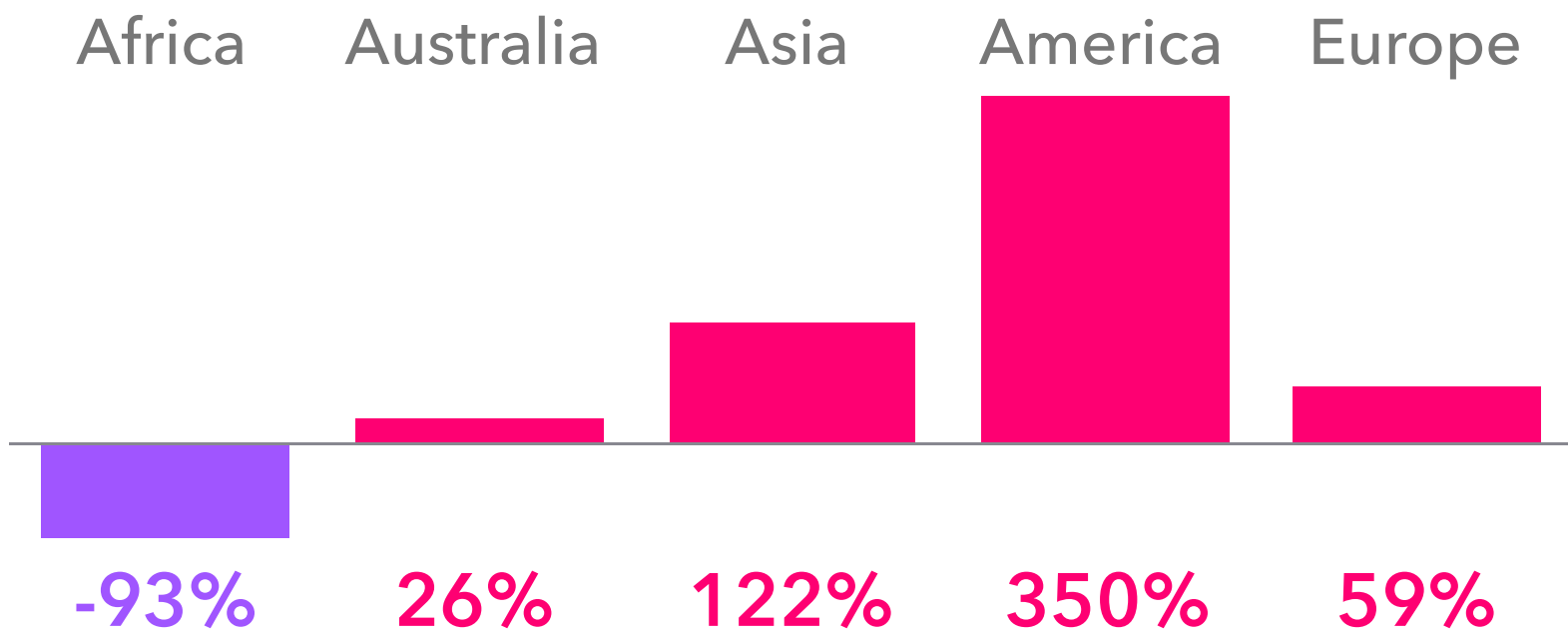


# Chapter 2 - Report Findings

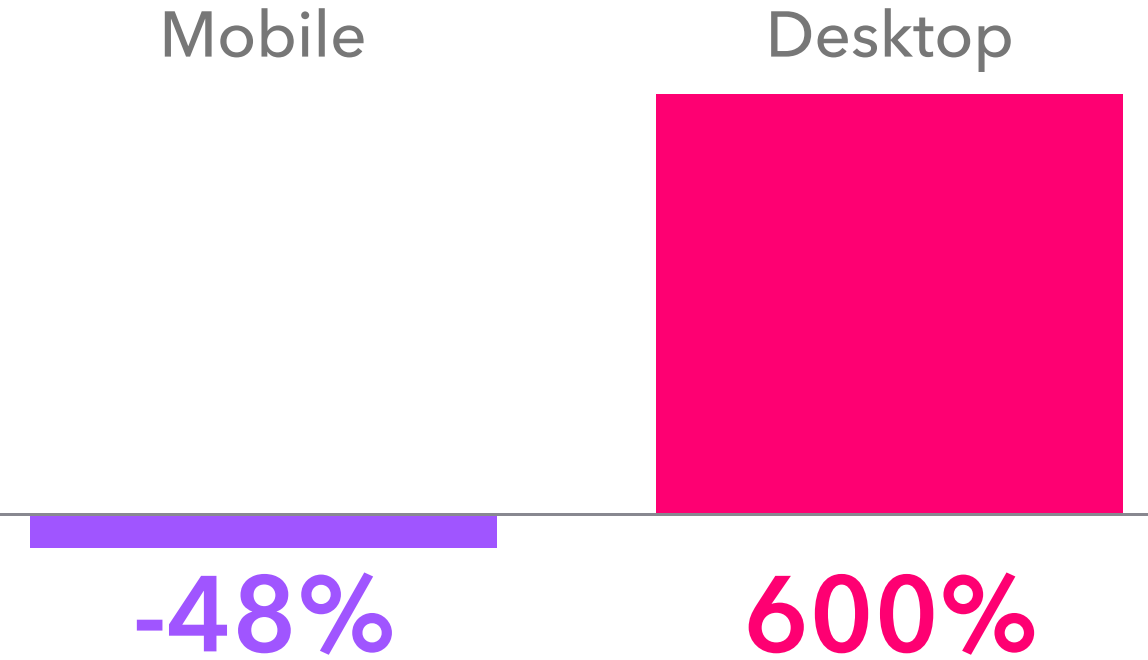
## Session Duration

### By Geography

All regions, with the exception of Africa, saw significant increases in session duration with the removal of bots and fake users. This suggests that bots in most regions move around websites much more quickly than legitimate human users.



### By Device



It is very apparent that the removal of bots would increase desktop session durations considerably, while actually lowering mobile session durations. Since bots tend to move around a site more sporadically than human users, this may suggest an abundance of bots on desktop.



## Chapter 2 - Report Findings

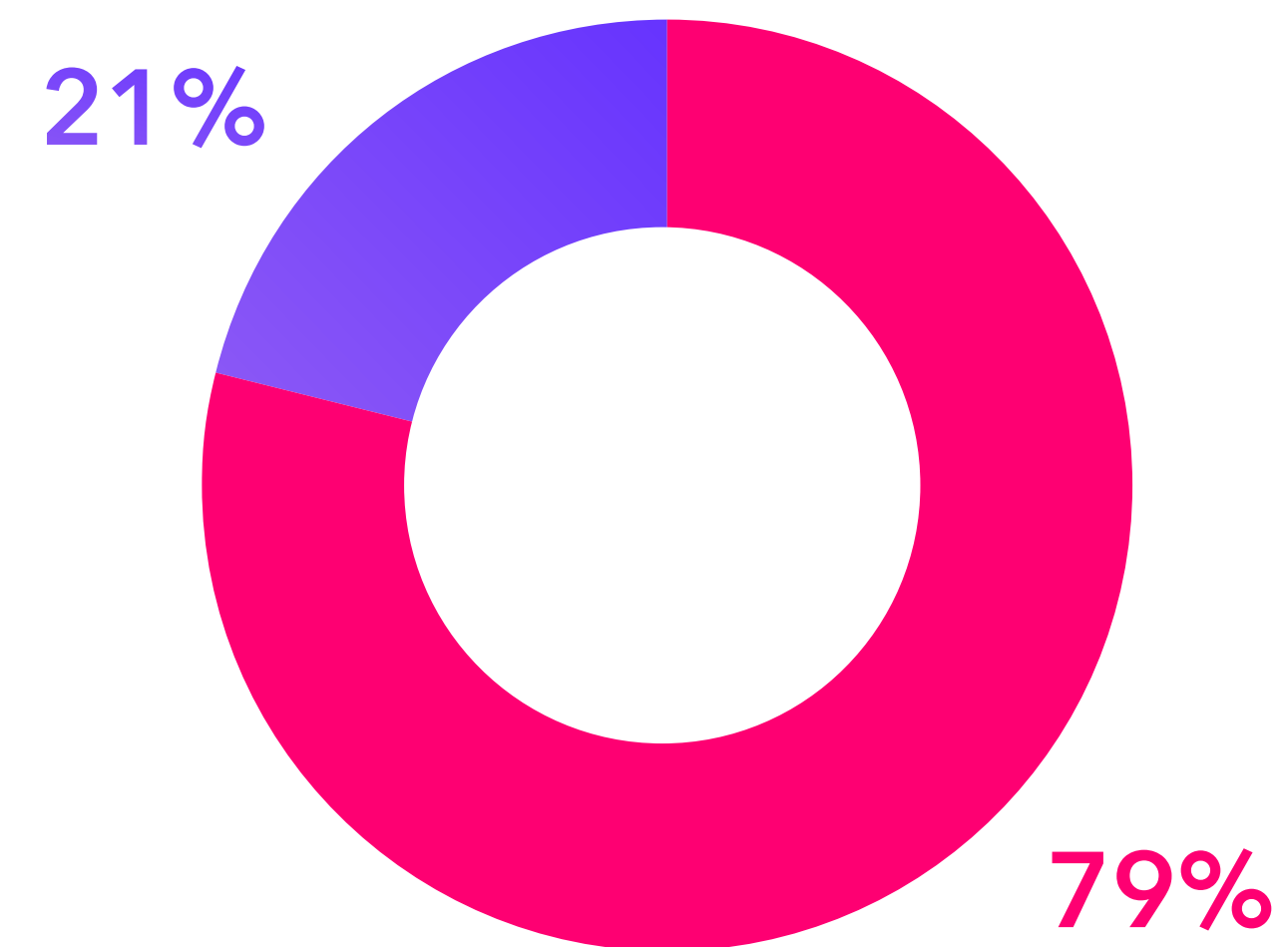
# New Users

### New Users are skewed by 21%

When looking at users who appeared to be accessing a website for the first time, 21% were flagged as bots or fake users.

This means that if those invalid users were to be removed, the amount of 'new' traffic to a given website could drop.

While drops in traffic are typically seen as a negative indicator, in this case, the marketer would actually be cleaning up their funnel, getting a more accurate picture of reality, and likely increasing conversion rates.



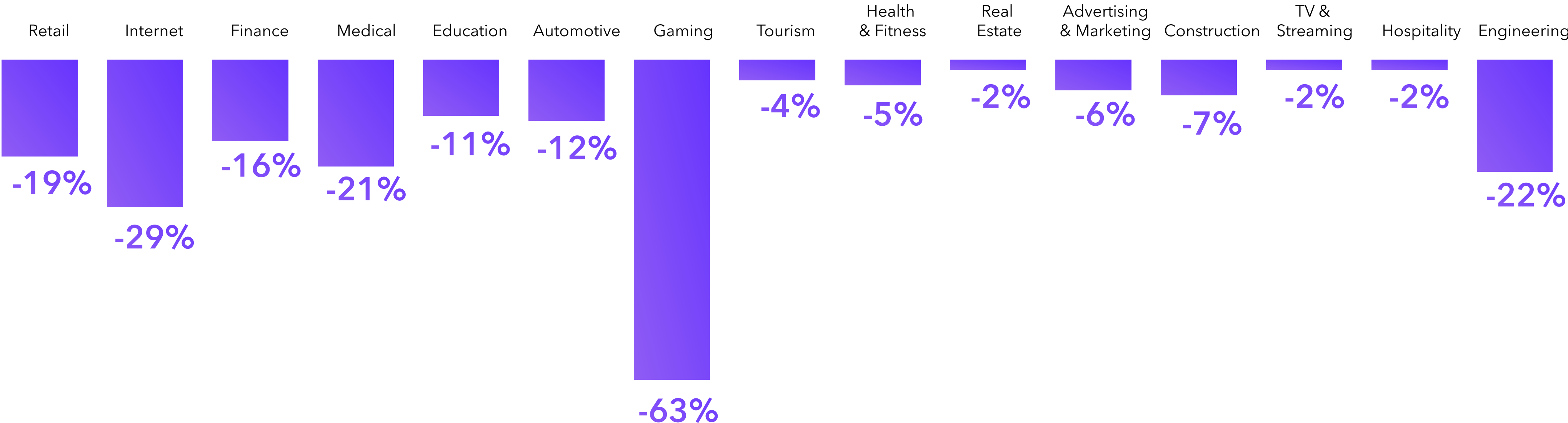
# Chapter 2 - Report Findings

## New Users

### By Industry

This section shows the breakdown of how much the volume of new users to a site could decrease by industry if bots and fake users were to be removed. The gaming industry saw the sharpest decrease, meaning that there is a high prevalence of bots and fake users arriving on sites within that industry

for the first time. This could indicate that new bots meant to attack that industry are being created very frequently, or that there is a lot of non-human and potentially malicious activity in that industry overall. Internet, engineering, and medical fields were not far behind.

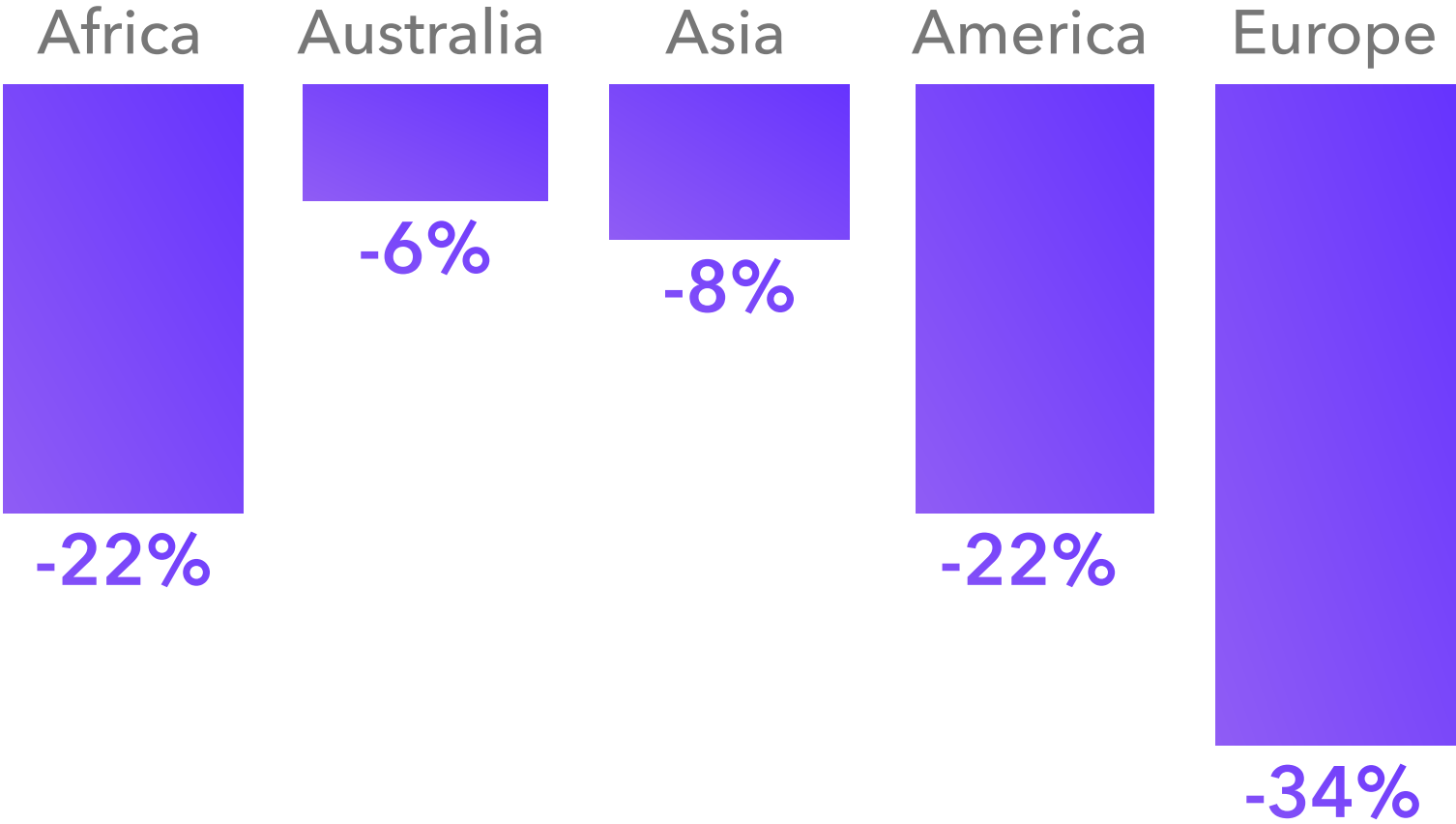


# Chapter 2 - Report Findings

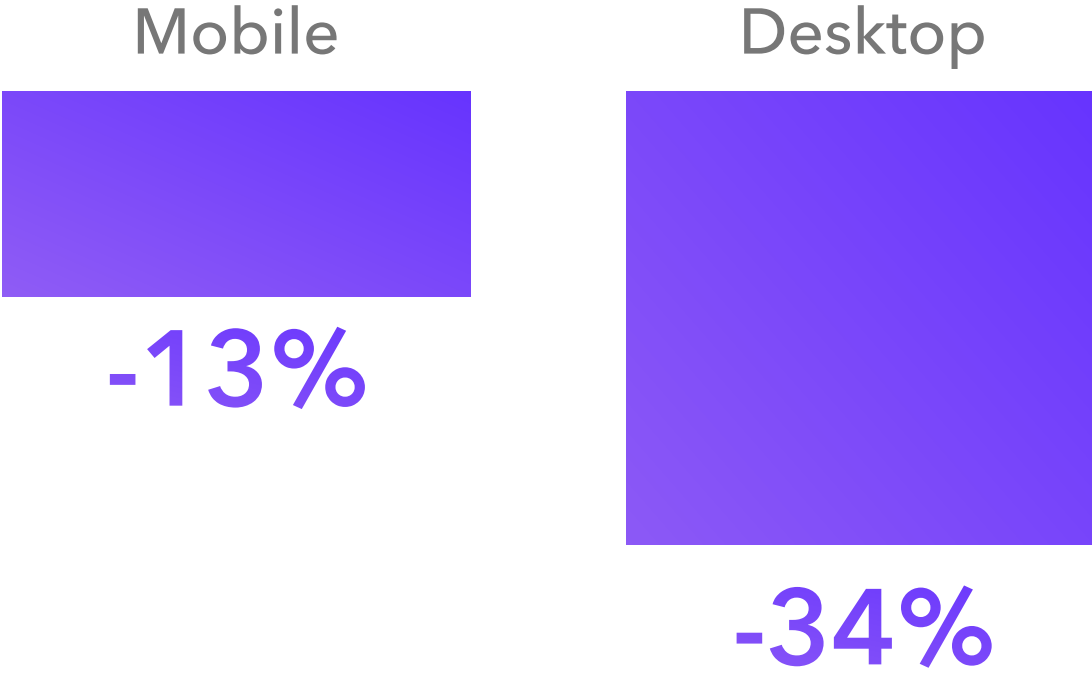
## New Users

### By Geography

Europe saw the sharpest decrease in new users when bots and fake users were removed. However, America and Africa were not far behind, with bots countries experiencing a 22% decrease in new users overall.



### By Device



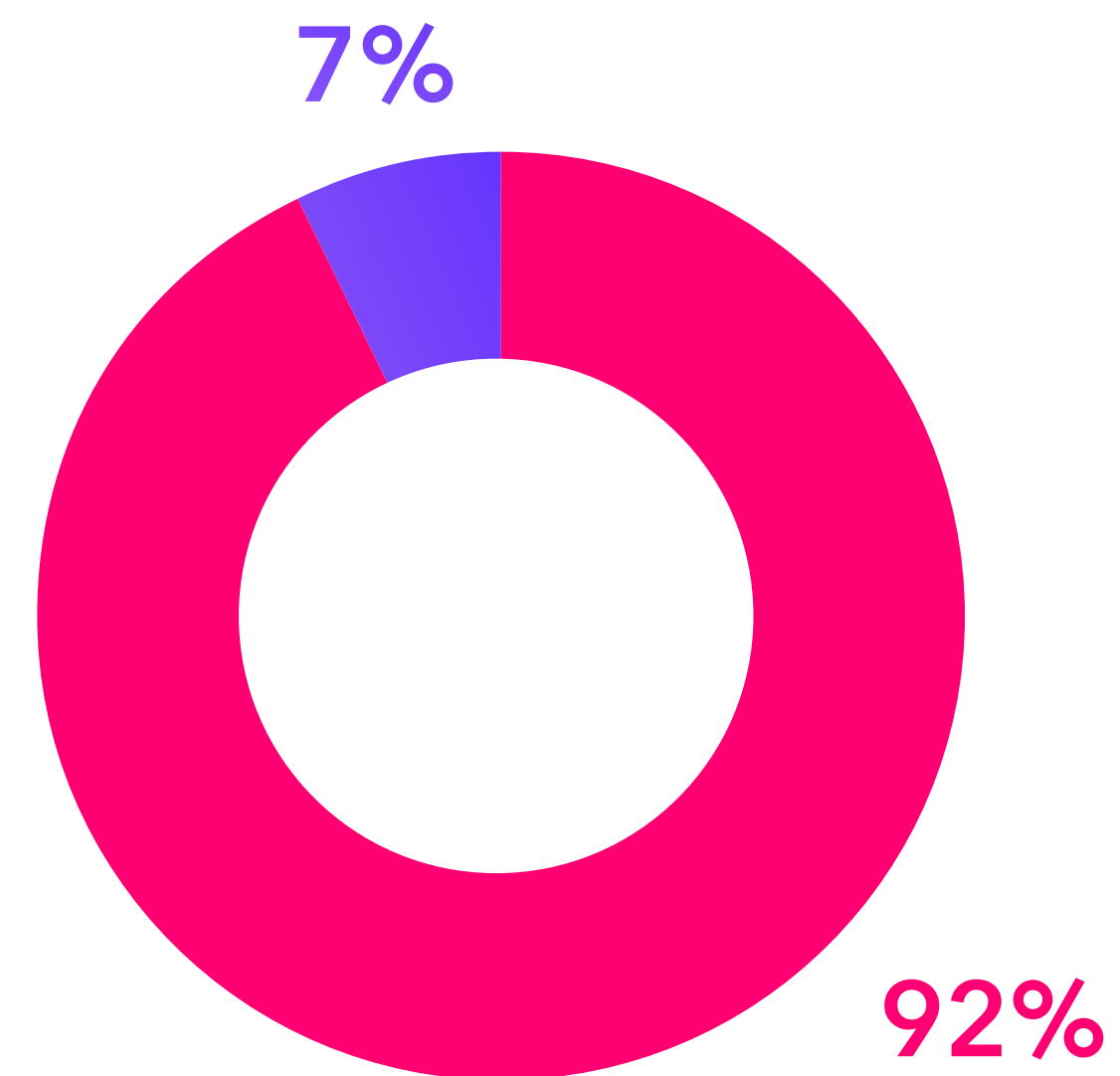
Desktop devices saw a significant decrease in new users when bots and fake users were removed, this indicates that when a bot arrives on a website for the first time, they are typically doing so from a desktop rather than a mobile device.

## Chapter 2 - Report Findings

# Returning Users

### Returning Users are skewed by 7%

Our study showed that there was a lower percentage of bots and fake users being marked as a return visitor. This could indicate that bots are taking on new identities, masking their locations, or disabling tracking information so they are more difficult to be identified the second time they return to a site. Furthermore, if the same bot appears to be coming back to a site for a second or third time, valuable resources are being wasted on them time and time again.





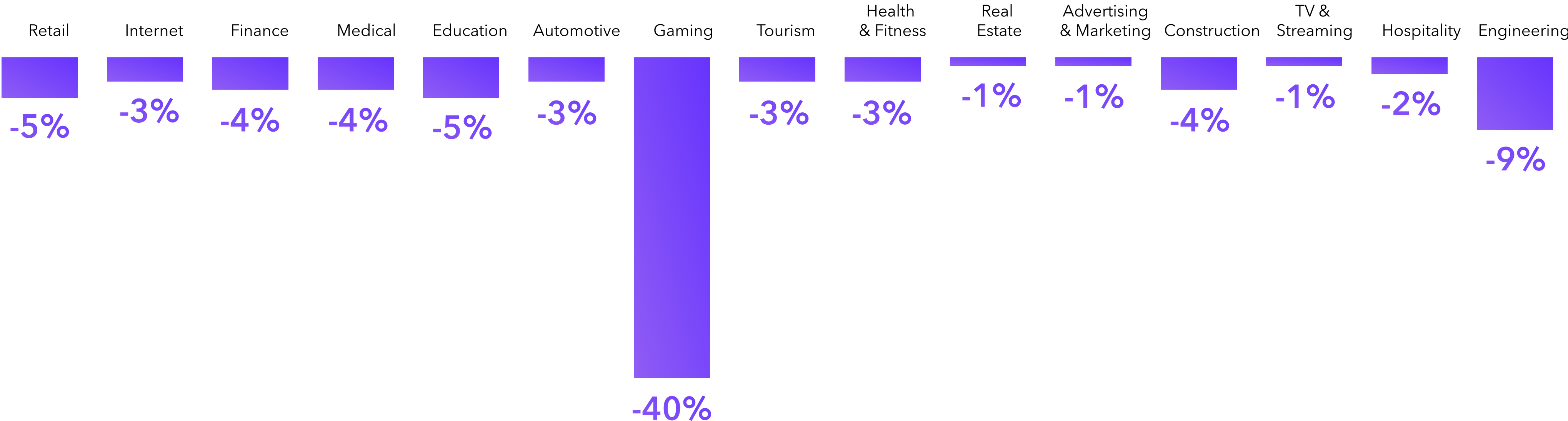
# Chapter 2 - Report Findings

## Returning Users

### By Industry

Following a similar pattern as the decrease in new users, the gaming industry also saw a significant decrease in returning users when bots and fake users were removed from the equation. However, perhaps the most interesting takeaway from the data on this page is that every single industry

experienced fewer bots and fake users arriving on-site as returning users than they did with new users. This is also reflected in our overall pie chart on the previous page. It is consistent with the hypothesis that when bots return to a website they could be masking their identities, taking on new identities, or otherwise hiding their details.

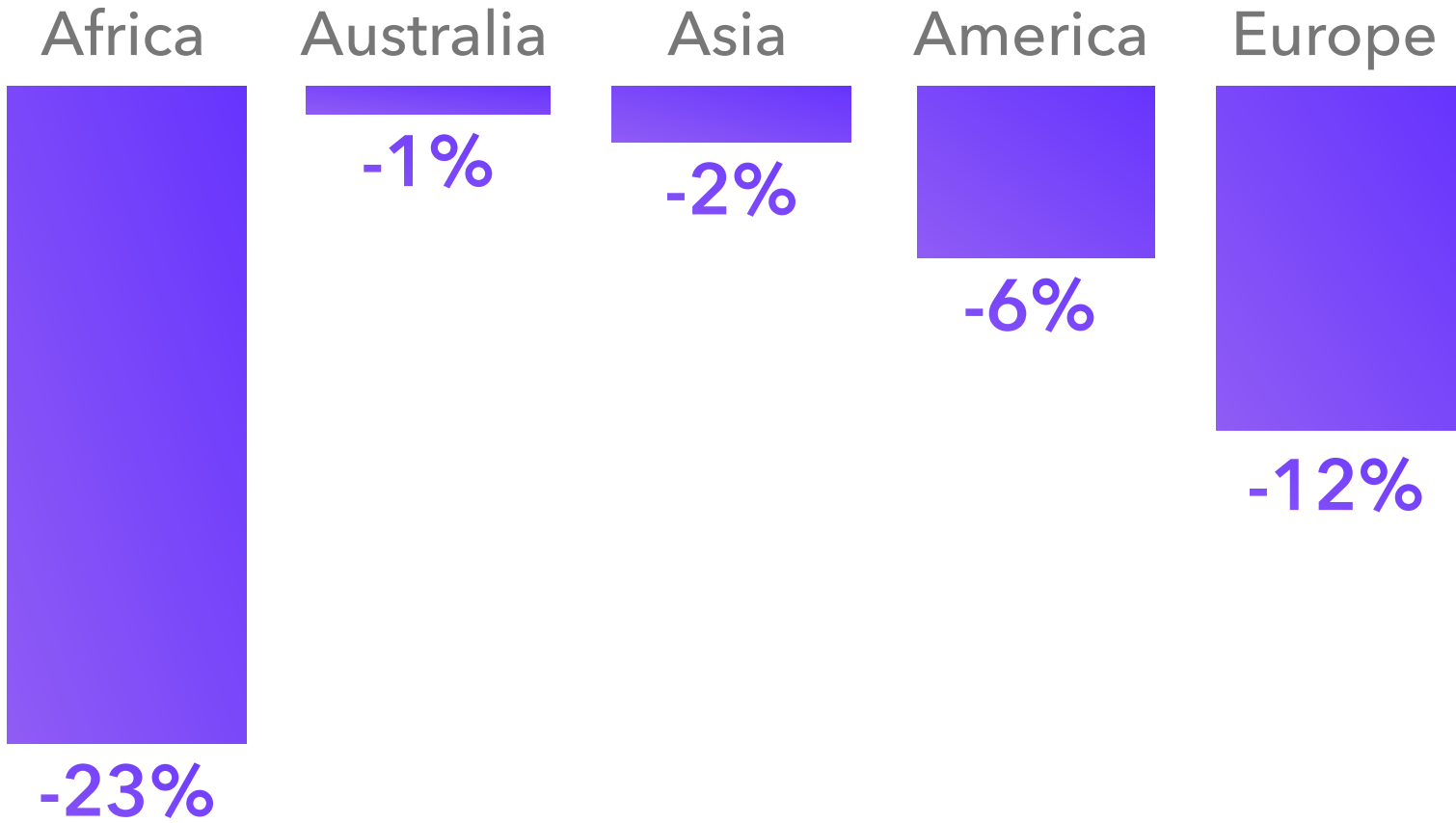


# Chapter 2 - Report Findings

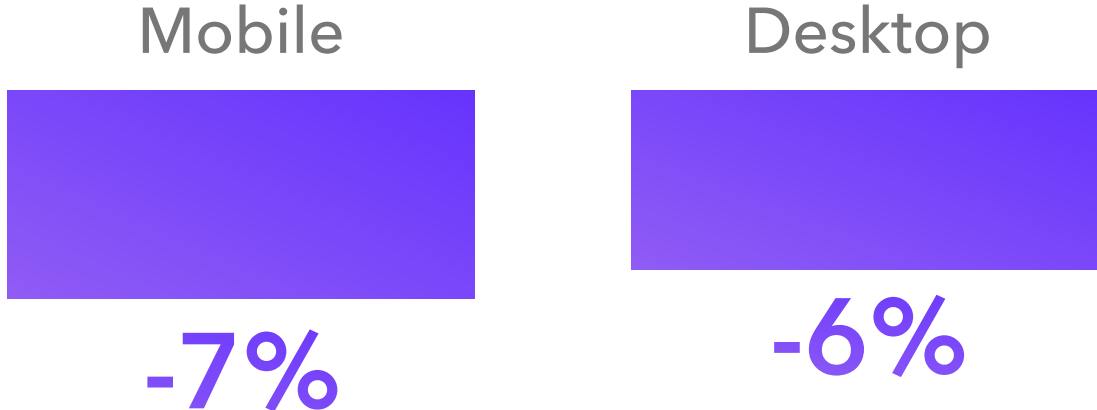
## Returning Users

### By Geography

When looking at returning visitors, most countries saw fewer bots and fake users than they did on the new users front. However, when looking at traffic arriving from Africa, the rates of invalid traffic were actually slightly higher than the rates they saw for new users.



### By Device



While desktop saw significantly higher rates than mobile on the fraudulent new users side of things, the rates look much more similar when looking at returning users. This could indicate that bots access a site from desktop the first time they arrive on site, but then might try to look like they are coming from other devices moving forward.

## Chapter 3

# Impact & Implications



**The data previously presented impacts each business financially and operationally.**

In this section, we describe the ways in which these metrics have a ripple effect on the goals and KPIs that marketers and analysts continually measure.

**Specifically we looked at:**

- Growth & Revenue Forecasting
- Media Buying & Campaign Optimization
- Customer Experience & Site Optimization



# Impact on Growth & Revenue Forecasting

Businesses rely on data to understand their current financial situations, as well as predict growth. The future of many organizations rely on their ability to correctly access, analyze, and sift through data in order to be able to plan for following quarters and years with a high level of confidence. These metrics directly impact valuations, headcount, market expansion, mergers and acquisitions, and how an organization scales overall.

However, when bots and fake users are present - as we have seen throughout this report - the metrics that inform these critical actions become skewed. Skewed metrics can be the difference between successful growth and failed scalability. Decisions should - and must - be made based on accurate data for this reason.

## Impact on Media Buying & Campaign Optimization

Paid marketers look at campaign data to make important decisions about their overall advertising strategy. Leaders in this field may choose to invest or divest from certain platforms based on overall performance. They might also decide which campaigns to put more or less budget behind based on metrics they have seen historically. Furthermore, they may optimize campaigns toward highly engaged audiences with the hope of driving additional traffic and ultimately revenue.

But when campaigns are polluted by bots and fake users, these decisions could be largely misguided. Campaign managers can unintentionally drive additional invalid traffic, and eventually make their efforts obsolete.

# Impact on Customer Experience & Site Optimization

Those involved in customer and user experience use data to inform the way their websites operate. For example, if most site visitors are arriving from mobile devices, they might choose to make the site more geared towards a mobile-first experience. Also, if a website's audience is mostly based in the US, the site will probably be written in American English and take into account the hours that audience most frequently accesses the internet. Furthermore, if one page is driving a lot of traffic, maybe that page is rebuilt to be more intuitive.

As one can imagine though, if the truth is revealed and the company discovers that many of these users who appeared to be engaged are actually bots - all of these decisions were made with the wrong audience in mind.



## Chapter 4

# Conclusion



**This report revealed the scale to which metrics are skewed from the Fake Web, as well as the impact that has on key business objectives.**

Some of the key metrics exposed in the report include:

- Unique Site Visits are skewed by 22.3%
- Pageviews are skewed by 7.4%
- Bounce Rates are skewed by 6.6%
- Session Duration is skewed by 41%
- New Users are skewed by 21%
- Returning Users are skewed by 7%

For more information on skewed metrics, or other reports on the Fake Web, visit:

<https://cheq.ai/research>