CONTENT, SHARES AND LINKS

What we learnt from 1 million posts

Joint Research Report

Buzzsumo & **MOZ**

September 2015

Author: Steve Rayson
🐦 @steverayson
🌐 **www.buzzsumo.com**

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

SUMMARY

This research project was designed to look at the correlation of shares and links. We wanted to understand the content that gets both shares and links and the formats that get relatively more shares or links. To undertake this analysis BuzzSumo and Moz provided data from their respective databases on over 1 million articles.

What we found is that the majority of content published on the internet is simply ignored when it comes to shares and links. The data suggests most content is simply not worthy of sharing or linking. It also suggests that people are very poor at amplifying content. It may sound harsh but it seems most people are wasting their time either producing poor content or failing to amplify it.

On a more positive note we also found some great examples of content that people love to both share and link to. It was not a surprise to find content gets far more shares than links. Shares are much easier to acquire. Everyone can share content easily and it is almost frictionless in some cases. Content has to work much harder to acquire links. Our research uncovered:

- the sweet spot content that achieves both shares and links
- the content that achieves higher than average referring domain links
- the impact of content formats and content length on shares and links

In this report we examine the relationship of shares and links, and the content sweet spot where content achieves both shares and links.
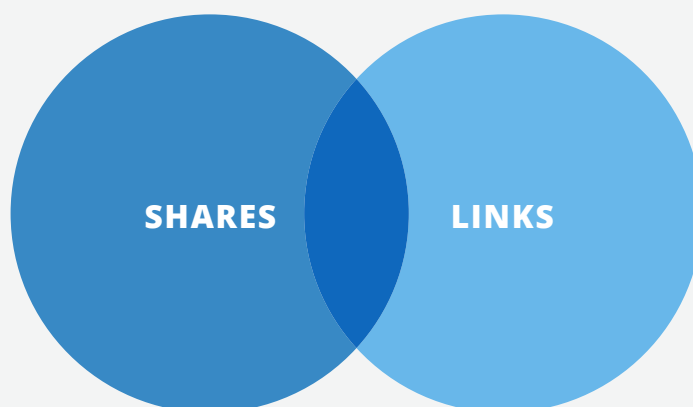
SHARES **LINKS**

SUMMARY FINDINGS:

01 The majority of posts receive few shares and even fewer links. In a randomly selected sample of 100,000 posts over 50% had 2 or less Facebook interactions (shares, likes or comments) and over 75% had zero external links. This suggests there is a lot of very poor content out there and also that people are very poor at amplifying their content.

> **75% of randomly selected articles had zero external links.**

02 When we looked at a bigger sample of 757,317 well shared posts we found over 50% of these posts still had zero external links. Thus suggests while many posts acquire shares, and in some cases large numbers of shares, they find it far harder to acquire links.

03 Shares and links are not normally distributed around an average. There are high performing outlier posts that get a lot of shares and links but most content is grouped at the low end, with close to zero shares and links. For example, over 75% of articles from our random sample of 100,000 posts had zero external links and just 1 or less referring domain link.

04 Across our total sample of 1m posts there was NO overall correlation of shares and links, implying people share and link for different reasons. The correlation of total shares and referring domain links across 757,317 articles was just 0.021.

> **There is no correlation of shares and links.**

05 There are, however, specific content types that do have a strong positive correlation of shares and links. This includes research backed content and opinion forming journalism. We found these content formats achieve both higher shares and significantly more links.

> **New research content and serious, opinion forming journalism achieves higher shares and links.**

06 85% of content published (excluding videos and quizzes) is less than 1,000 words long. However, long form content of over 1,000 words consistently receives more shares and links than shorter form content. Either people ignore the data or it is simply too hard for them to write quality long form content.

> **long form content of over 1,000 words consistently receives more shares and links than shorter form content.**

07 Content formats matter. Formats such as entertainment videos and quizzes are far more likely to be shared than linked to. Some quizzes and videos get hundreds of thousands of shares but no links.

08 List posts and videos achieve much higher shares on average than other content formats. However, in terms of achieving links, list posts and why posts achieve a higher number of referring domain links than other content formats on average. While we may love to hate them, list posts remain a powerful content format.

THE SAMPLE

Our sampling approach had three stages.

We initially selected 100,000 random posts from the BuzzSumo database to act as a control group. We then selected a further structured sample of 657,317 posts. For this second sample we deliberately chose well-shared examples of different content formats so that we could investigate what content gets shared and the impact, if any, of different content formats.

We were conscious that well-shared shared posts might come from a limited number of popular domains. However, when we analysed the 757,317 posts selected, we found they included articles from over 600,000 different domains. There were more posts from popular domains such as YouTube, BuzzFeed and the New York Times, as we would expect, but overall there was a reasonable spread of domains.

Our analysis of the 757,317 posts highlighted certain domains that appeared to have a high correlation of shares and links. However, we were cautious to draw too many conclusions, as the sample sizes from these domains were quite low. This caused us to select a further, third sample of over 250,000 posts from these domains to give us a larger sample of highly correlated content to examine.

We were conscious in our sample selection that social shares and links are not static. They both grow over time, though the data suggests that most social shares tend to take place very quickly and then grow slowly whereas links tend to grow more steadily over time. In order to allow time for an article to gain both shares and links we didn't select any articles that were less than 3 months old.

In total, through the three samples above, we analysed over 1m articles that were over 3 months old and published across more than 600,000 domains.

> **In total we analysed over 1m articles published across more than 600,000 domains.**

We used the BuzzSumo API to identify content types, word length and share data and the Moz API to get data on the numbers of links and the domain/page authority for these articles.

Our sample data for each post included:

- Number of referring domain links
- Number of sub-referring domain links
- Number of external links
- Moz Rank
- Page authority
- Domain authority
- Total shares
- Total Facebook interactions
- Twitter shares
- Linkedin shares
- Pinterest shares
- Google Plus shares
- Content format
- Content length (number of words)

We used a statistical package called R to undertake our analysis of the data including our correlations and subsets.

As we stated above we started by selecting 99,941 articles at random from the BuzzSumo database for us to use as a control group. We then created a structured sample of 657,376 posts that included well-shared content specific formats namely:

- Videos
- Quizzes
- Infographics
- List posts
- Why posts
- 'how to' posts
- Research reports & surveys

The reason for our selection was that we wanted to explore the impact of specific content formats on sharing and linking.

We were conscious that our approach would mean the overall sample of 757,317 posts would be biased towards heavily shared content. This was confirmed when we compared the random sample with the 757,317 posts, as shown below.

Content Sample	Average Total Shares	Median Shares	Average Referring domain links	Median Referring domain links
Random sample of posts (99,941)	257	8	0.75	0
Full sample of posts (757,317)	4,393	202	3.77	1

Thus we need to recognise that the 757,317 posts in our sample will have a much higher average of both links and shares than you might expect from a random sample of content.

Despite this bias towards heavily shared content, the majority of our sample still had a relatively low level of shares and links. As we can see below 50% of our sample had 11 or less Twitter shares and 50% of the posts had zero external links.

```
num_subdomain_links external_links     num_domain_links        moz_rank
Min.    :    0.00    Min.    :     0.0  Min.    :    0.00   Min.    :0.000
1st Qu.:    0.00     1st Qu.:      0.0  1st Qu.:    0.00   1st Qu.:0.000
Median :    1.00     Median :      0.0  Median :    1.00   Median :3.444
Mean   :    4.18     Mean   :     27.8  Mean   :    3.77   Mean   :2.633
3rd Qu.:    3.00     3rd Qu.:      3.0  3rd Qu.:    3.00   3rd Qu.:4.075
Max.   :48121.00     Max.   :717561.0  Max.   :43003.00   Max.   :8.338


page_authority  domain_authority  total_shares        total_facebook_shares
Min.    : 1.00   Min.    :   1.0   Min.    :      0    Min.    :      0
1st Qu.: 1.00    1st Qu.: 36.0     1st Qu.:     15     1st Qu.:      3
Median :25.00    Median : 64.0     Median :    202     Median :    104
Mean   :26.06    Mean   : 61.8     Mean   :   4393     Mean   :   3659
3rd Qu.:45.00    3rd Qu.: 90.0     3rd Qu.:   2781     3rd Qu.:   2110
Max.   :95.00    Max.   :100.0     Max.   :5768346     Max.   :4015178


twitter_shares   linkedin_shares   pinterest_shares    google_plus_shares
Min.    :      0  Min.    :   0.00  Min.    :     0.0   Min.    :     0.0
1st Qu.:      2   1st Qu.:   0.00   1st Qu.:     0.0    1st Qu.:     0.0
Median :     11   Median :   0.00   Median :     0.0    Median :     1.0
Mean   :    462   Mean   :  24.91   Mean   :    90.5    Mean   :   247.3
3rd Qu.:     62   3rd Qu.:   2.00   3rd Qu.:     0.0    3rd Qu.:     6.0
Max.   :5764265   Max.   :47697.00  Max.   :1113786.0   Max.   :840544.0
```

MOST CONTENT GETS FEW SHARES AND EVEN FEWER LINKS

As the table above makes clear despite our sample being biased towards highly shared content, the median shares and links were both relatively low. However, the shares and links for our 100,000 randomly selected posts were much lower.

In terms of shares 50% of the posts had 2 or less Twitter shares, 2 or less Facebook interactions, 1 or less Google+ shares and zero LinkedIn shares.

In terms of links 75% of the posts had zero external links and 1 or less referring domain links. The details are shown below.

```
num_subdomain_links  external_links     num_domain_links       moz_rank
Min.   :   0.0000    Min.   :    0.00   Min.   :  0.0000   Min.   :0.000
1st Qu.:   0.0000    1st Qu.:    0.00   1st Qu.:  0.0000   1st Qu.:0.000
Median :   0.0000    Median :    0.00   Median :  0.0000   Median :0.000
Mean   :   0.8394    Mean   :    4.16   Mean   :  0.7499   Mean   :1.234
3rd Qu.:   1.0000    3rd Qu.:    0.00   3rd Qu.:  1.0000   3rd Qu.:2.898
Max.   :1129.0000    Max.   :58650.00   Max.   :869.0000   Max.   :7.270


page_authority    domain_authority   total_shares       total_facebook_shares
Min.   : 1.00     Min.   :  1.00     Min.   :      0.0   Min.   :      0.0
1st Qu.: 1.00     1st Qu.: 25.00     1st Qu.:      2.0   1st Qu.:      0.0
Median : 1.00     Median : 42.00     Median :      8.0   Median :      2.0
Mean   :10.36     Mean   : 49.87     Mean   :    257.6   Mean   :    238.9
3rd Qu.:16.00     3rd Qu.: 76.00     3rd Qu.:     39.0   3rd Qu.:     21.0
Max.   :90.00     Max.   :100.00     Max.   :1383140.0   Max.   :1382976.0


twitter_shares    linkedin_shares     pinterest_shares    google_plus_shares
Min.   :    0.00  Min.   :    0.000   Min.   :     0.00   Min.   :    0.000
1st Qu.:    1.00  1st Qu.:    0.000   1st Qu.:     0.00   1st Qu.:    0.000
Median :    2.00  Median :    0.000   Median :     0.00   Median :    0.000
Mean   :   13.97  Mean   :    2.453   Mean   :    25.02   Mean   :    2.218
3rd Qu.:    6.00  3rd Qu.:    0.000   3rd Qu.:     0.00   3rd Qu.:    1.000
Max.   :35600.00  Max.   :23588.000   Max.   :282462.00   Max.   :22458.000
```

We should also note that the random posts were selected from our BuzzSumo database, which in itself is biased towards shared content. We typically do not include content with zero shares in our database. Thus we could reasonably expect that randomly selected content from across the internet, rather than the BuzzSumo database, would have an even lower level of shares.

The first immediate conclusion from our research is that the majority of content gets few shares and even fewer links.

DISTRIBUTION OF SHARES & LINKS

50% OF POSTS GET LOW SHARES AND VERY LOW LINKS

Despite the bias in our sample to more heavily shared content, most posts in the sample get very low shares and links. For example if we look at the median, which is the number that separates the bottom half from the top half of our sample of 757,317 posts, we can see that 50% of all articles get 11 or less Twitter shares, 1 or less Google plus shares, zero LinkedIn shares and zero Pinterest shares.

Sample Medians

Sub-domain links	External links	Referring domain links	Moz Rank	Page authority	Domain authority
1	0	1	3.44	25	64

Twitter shares	FB shares	Linkedin shares	Pinterest	Google+	Total shares
11	104	0	0	1	202

The reason we have included the median in the above is because we found a very skewed distribution of shares and links. In essence there is a lot of content that gets few shares and even fewer links. Thus whilst our sample might be biased towards more heavily shared content the median LinkedIn and Pinterest shares were zero and Twitter shares were 11.

When we look at links, we can see 50% of the 757,317 sample had zero external links and 50% had 1 or less referring domain links.

DISTRIBUTION OF EXTERNAL LINKS

In our sample of 757,317 posts 75% of articles had 3 or less external links.
The distribution of external links was as follows.
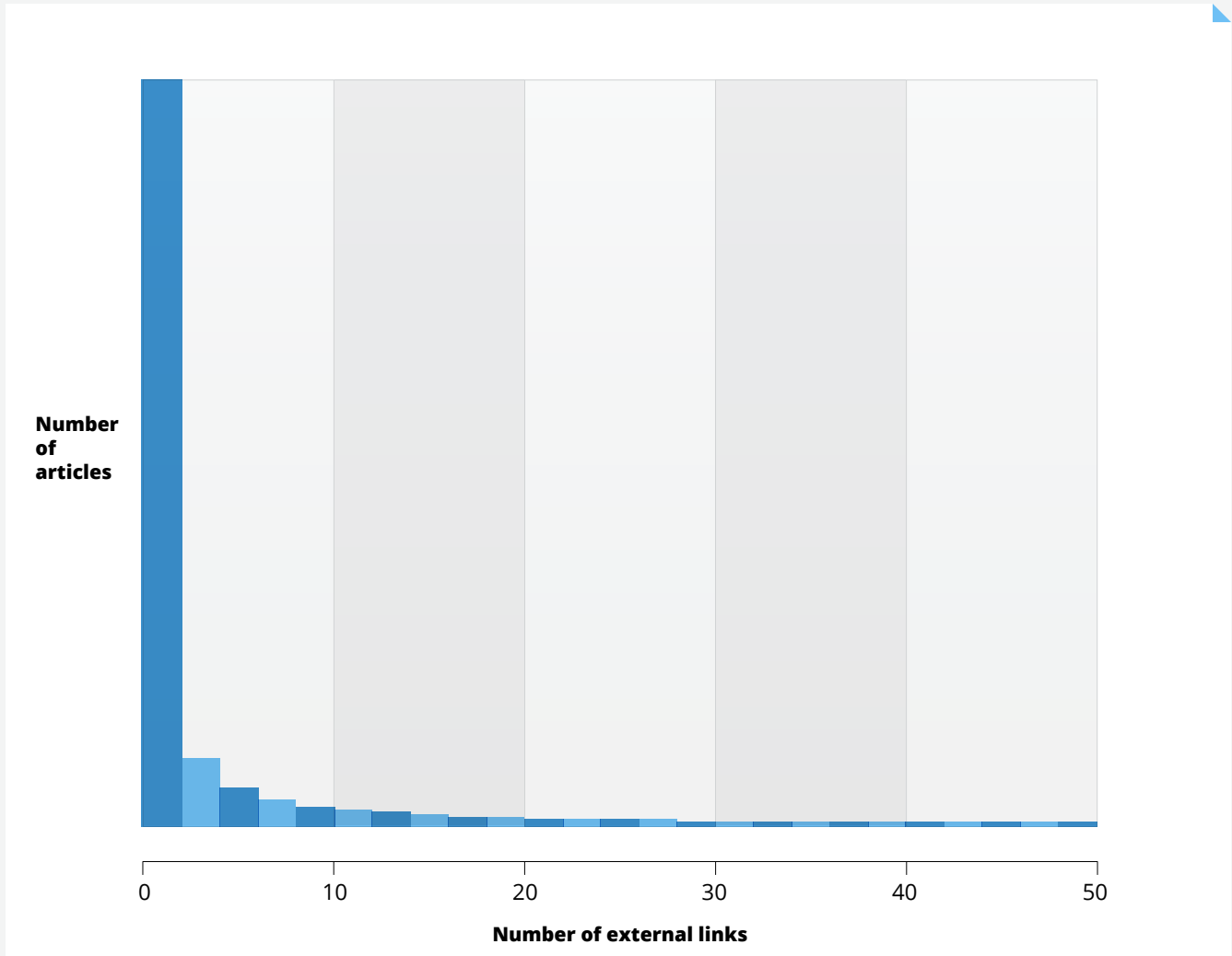
This chart is cut off on the X axis on the right at 50 external links. We had to cut off the chart to show the distribution, the full distribution would stretch out a very long way on the right to a post that had over 700,000 links. If we showed the full distribution, including the post with the most external links, the histogram would look like this.
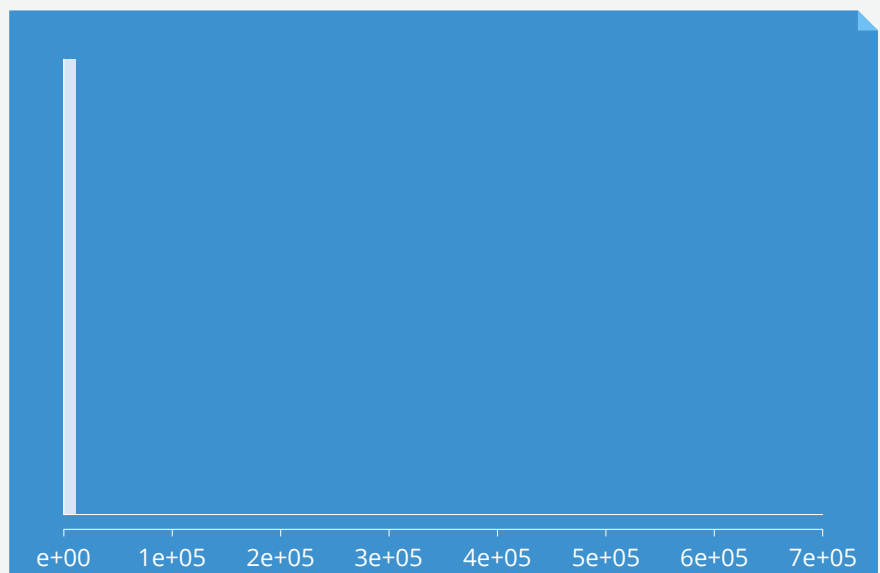
DISTRIBUTION OF REFERRING DOMAIN LINKS

The histogram below shows the distribution of referring domain links for our sample of 757,317 posts.

In this case 75% of articles had 3 referring domain links or less. This chart is cut off at 15 referring domain links on the bottom axis but it would stretch out a long way to right. One post received an incredibly high 43,000 referring domain links.

Number of articles

0   5   10   15

Number of referring domain links
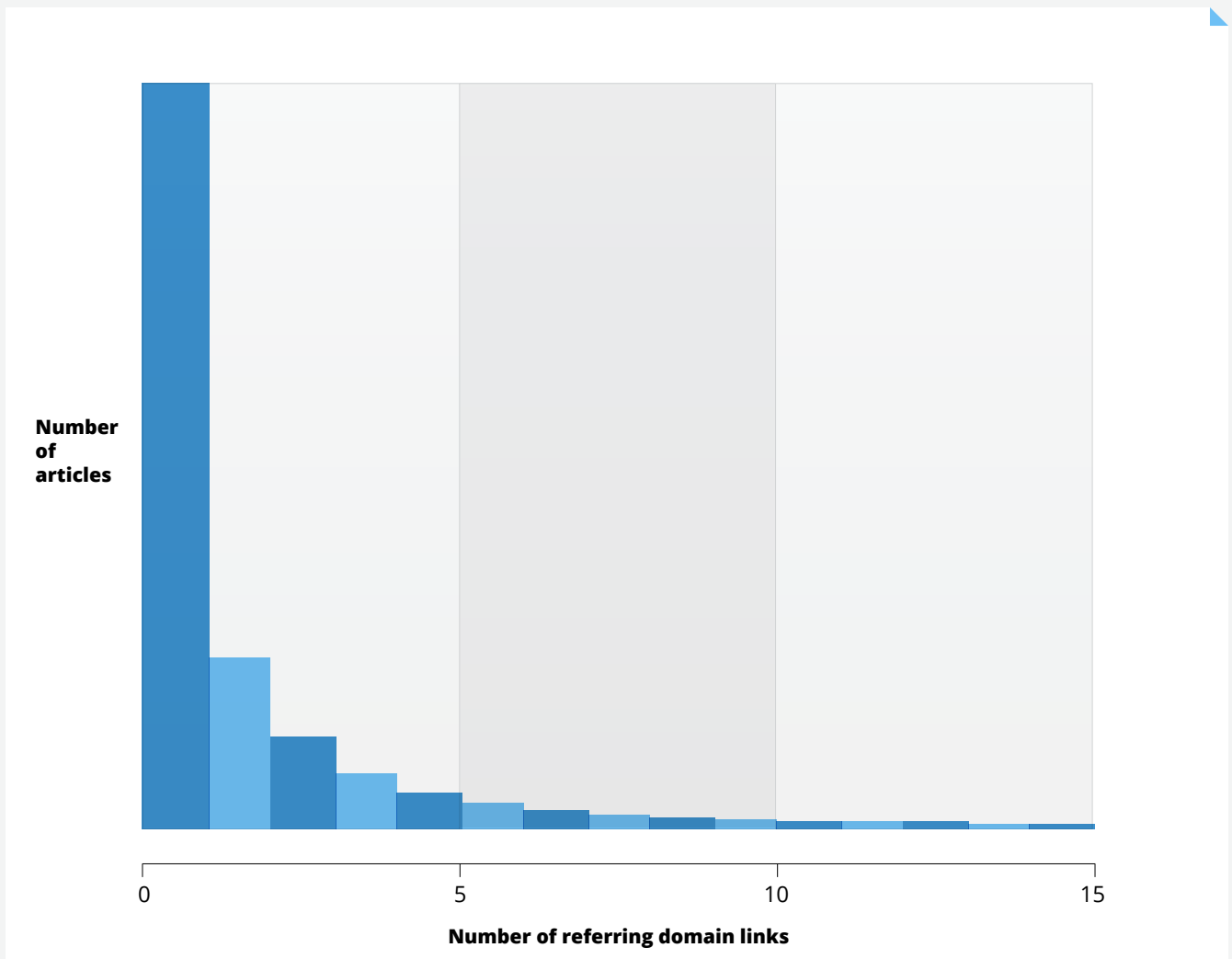
As the histogram shows we get a very similar skewed distribution of articles based on referring domain links as we got with external links. In both cases there was a long tail of highly performing posts.

DISTRIBUTION OF TOTAL SHARES

The distribution of shares follows a very similar pattern to links. Most articles were grouped at the low end in terms of total shares. In our sample of 757,317 posts 25% of articles had less than 18 shares and 50% of posts had 202 shares or less. This histogram is cut off on the bottom right axis at 1,000 shares but would stretch out a long way to the post that achieved 5.7m shares.

Number of articles

Total shares across all networks
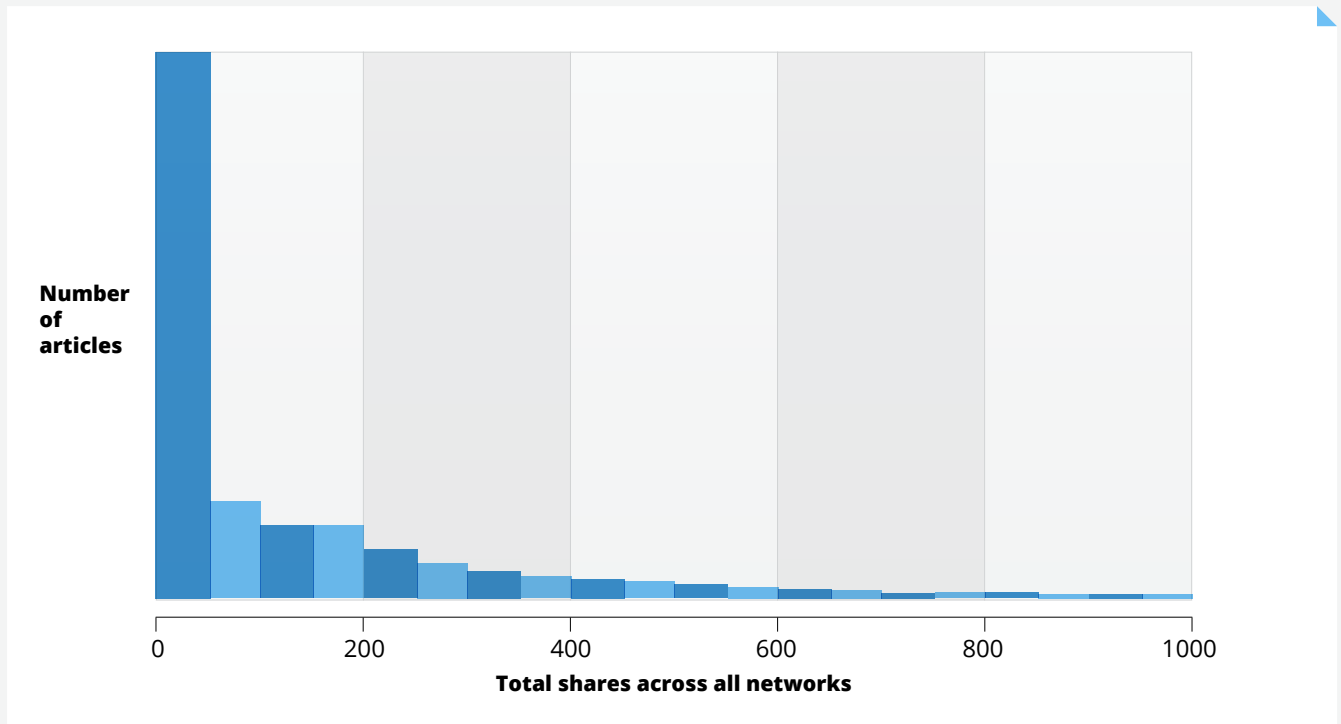
The distribution of shares doesn't flatten out completely after 1,000 shares. If we pull back to look at the distribution up to 10,000 shares, we can see from the histogram below there are slight bumps at 1,500 to 2,000 shares and 4,500 to 5,000 shares. After this the distribution flattens out more or less completely.

Number of articles

Total shares across all networks
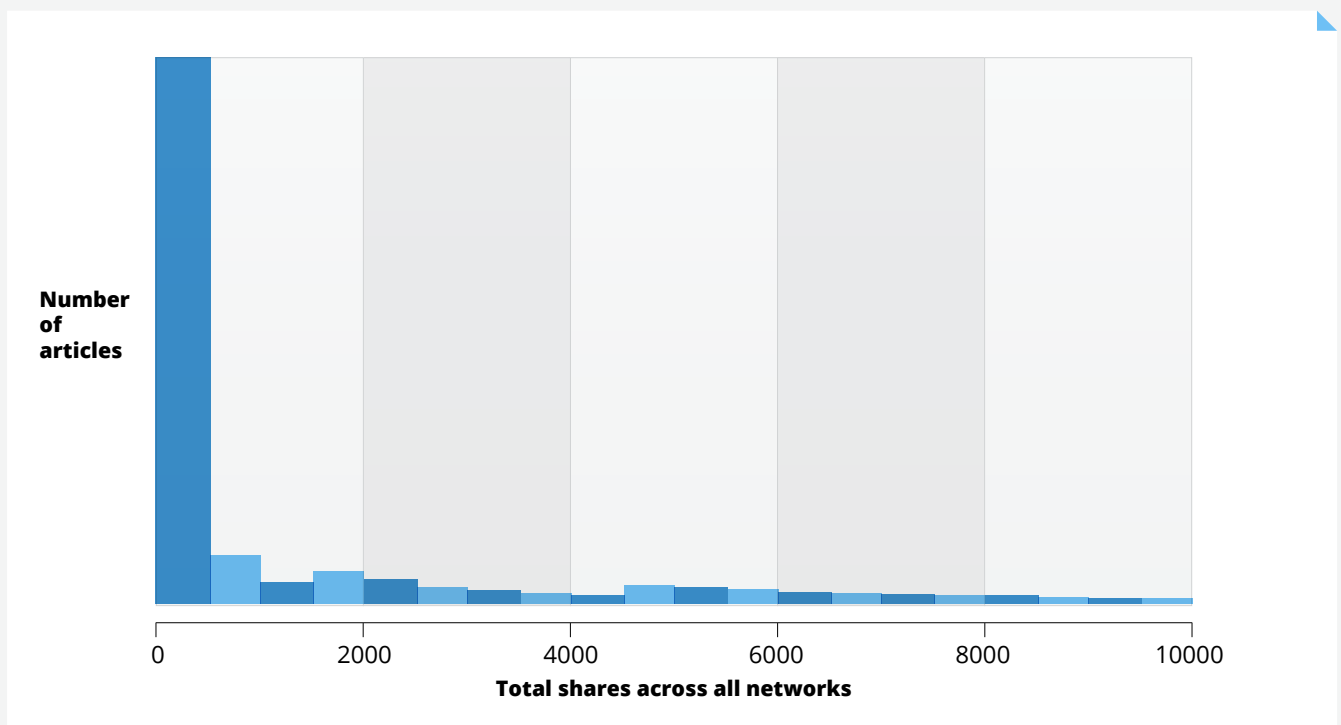
DISTRIBUTION OF SHARES AND LINKS ON DOMAINS

We might expect a more normal distribution of shares and links when it comes to domains e.g. some lower performing posts and some higher performing posts. However, we also find a skewed distribution of shares and links on domains as well. For ex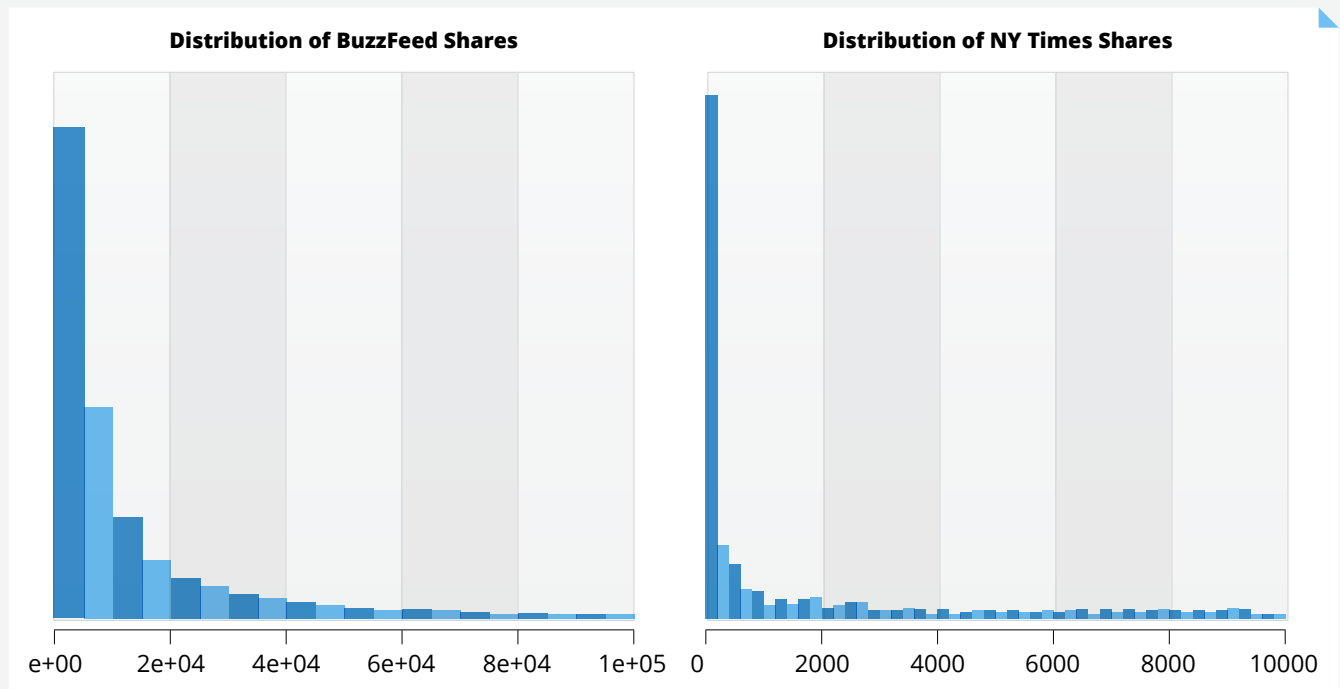ample, we looked at content on the New York Times and BuzzFeed. The distribution of shares up to 100,000 shares, is shown below. This shows that whilst shares do not flatten out quite as quickly there is a similar distribution.

Distribution of BuzzFeed Shares

Distribution of NY Times Shares

The same is true when we look at the distribution of referring domain links as shown below for BuzzFeed and New York Times.

Distribution of BuzzFeed Links

Distribution of NY Times Links
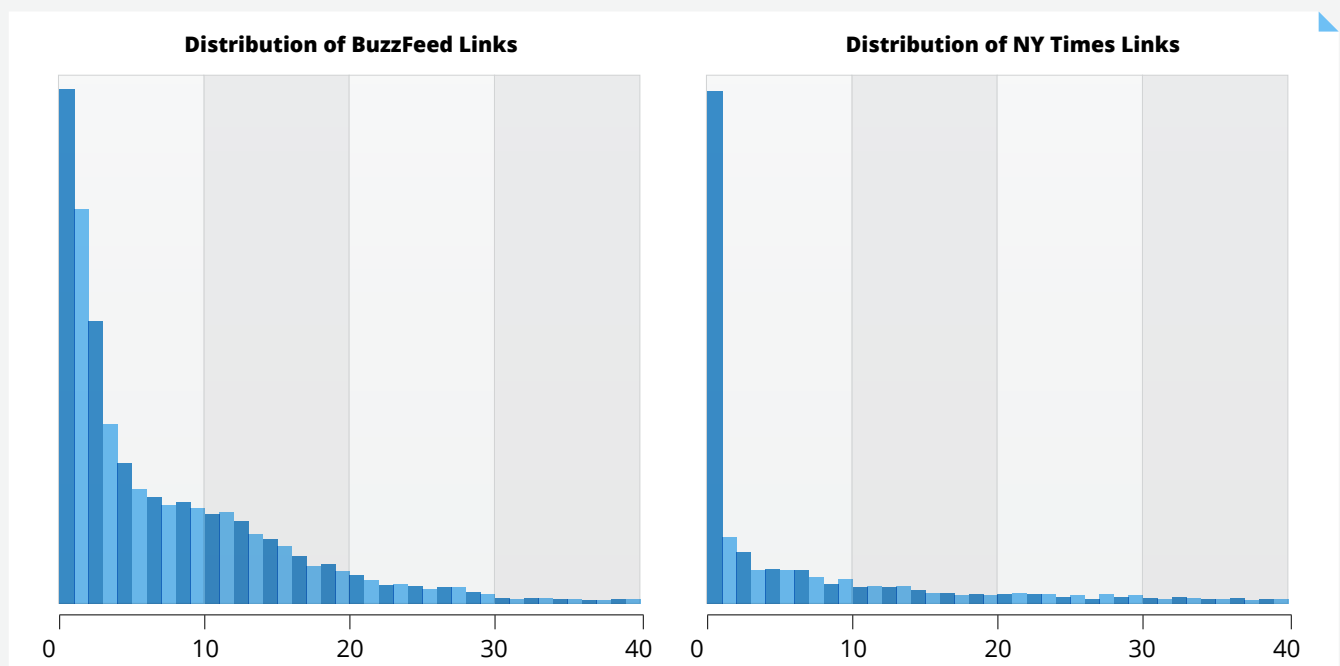
Buzzfeed's distribution takes longer to flatten out as the number of referring domain links increases but the distribution follows a similar pattern to the New York Times. The highest BuzzFeed post had 449 referring domain links, and hence there is a very long tail to the right.

WHY A SKEWED DISTRIBUTION MATTERS

The skewed nature of distribution is important to understand as it can make averages misleading. For example, below is the distribution of total shares across our whole sample. The right axis is cut off at 10,000 shares but would stretch out a long way, as one post had over 5m shares.

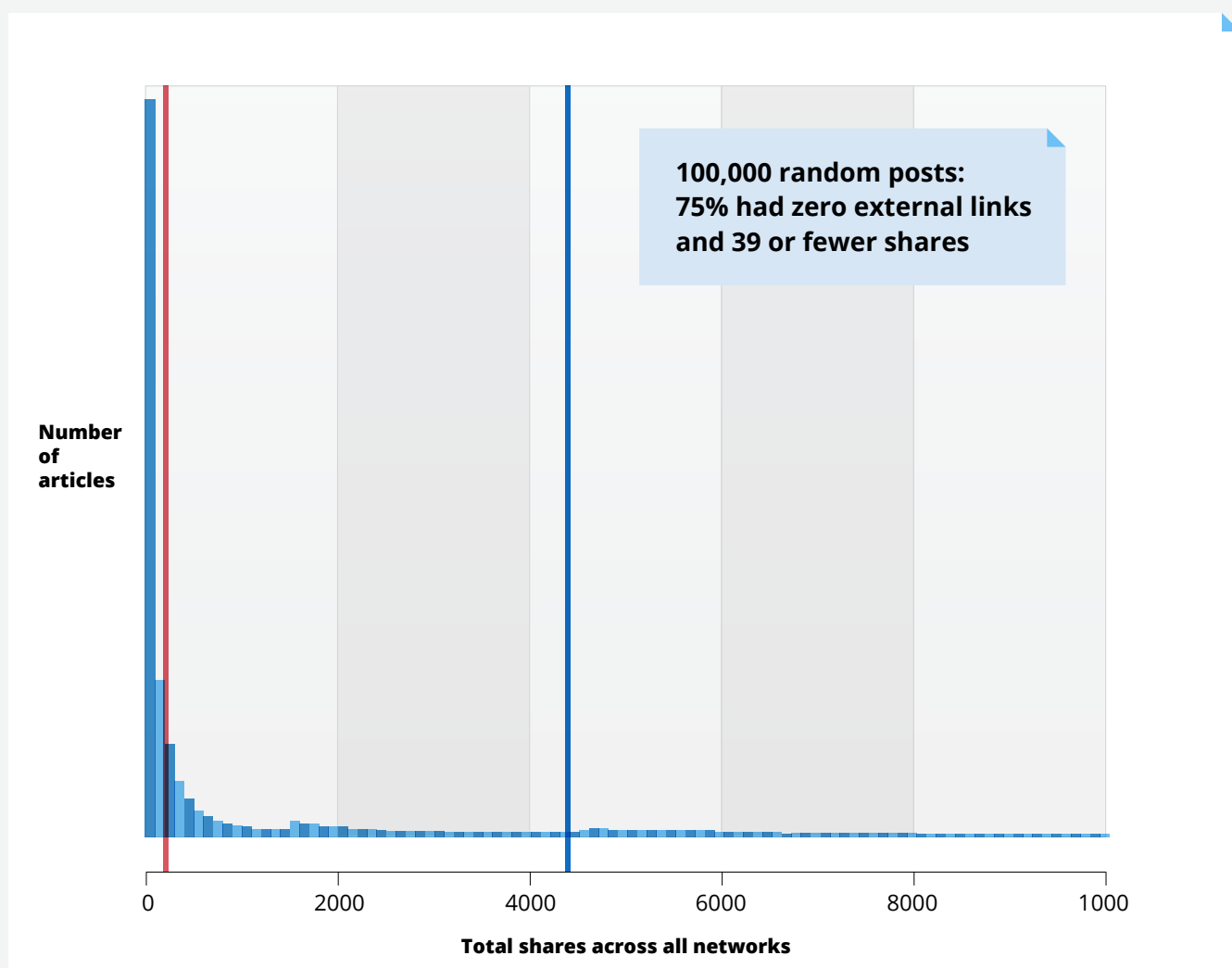The median is 202 shares (the red line) but the skewed nature and very highly shared posts makes the mean average 4,393 (the blue line). This is 20 times higher than the median. 50% of all posts are to the left of the median red line. 75% of all posts had less than 2,781 shares, still well to the left of the blue average line.

Number of articles

100,000 random posts: 75% had zero external links and 39 or fewer shares

| 0 | 2000 | 4000 | 6000 | 8000 | 1000 |

Total shares across all networks

For these reasons we will often show in our analysis both the median and mean for shares and links.

In summary, the majority of posts appear to get few shares and even fewer links; which results in a heavily skewed distribution rather than a normal distribution.

To emphasise the point about low shares and links, let's just look at the 99,900 posts we selected at random. In terms of shares 50% of all these posts had 8 or less shares and 75% had 39 or fewer shares. In terms of links 75% of these posts had zero external links and just 1 or fewer referring domain links.

CORRELATION OF SHARES AND LINKS

We started this research expecting to find a strong positive correlation between shares and links. However, what we found was a very weak positive correlation between links and shares across the 757,317 articles.

We used the Pearson correlation co-efficient, a measure of the linear correlation between two variables. The results can range from between 1 (a total positive correlation) to 0 (where there is no correlation) to −1 (a total negative correlation).

THERE IS NO POSITIVE CORRELATION OF SHARES AND LINKS

The overall correlations in our sample were:

Total shares and Referring domain links	0.021
Total shares and Sub-referring domain links	0.020
Total shares and External Links	0.011

We also looked at the correlations for our random sample of 99,000 posts to see if there was any significant difference, as we know the random sample had less heavily shared content. The correlation of shares to referring domain links was 0.018, almost exactly the same as for our larger sample.

These findings show for our sample that **there is no overall correlation of shares and links.**

A strong positive correlation between links and shares would infer that people link and share for similar reasons; and that as the number of shares increase so will the number of links. Our findings suggest that overall people share and link to content for different reasons.

We also looked at correlations for shares on different social networks to see if there were positive correlations for specific networks. We found no strong positive correlation of shares to referring domain links across the different networks.

Facebook total interactions	0.0221
Twitter	0.0281
Linkedin	0.0216
Pinterest	0.0065
Google plus	0.0058

Our correlation findings suggest that people share and link to content for different reasons.

The correlations were lower for Google Plus and Pinterest but all of the correlations we close to zero.

HIGHLY SHARED POSTS HAVE A HIGHER CORRELATION OF SHARES AND LINKS

We were conscious that our sample had a bias towards content with higher shares and we wanted to see if this might have affected our results. To do this we analysed the random content sample (99,941 posts) as this sample had far fewer shares and links than our overall sample, and compared the correlation of the random sample to the overall sample.

The random sample had a correlation of total shares and referring domain links of 0.018, which was very similar to the correlation for the sample of 757,317 posts of 0.021. Thus it initially appeared that the correlation of shares and links was unaffected by the number of shares.

However, we also created a subset of very highly shared posts, which we defined as posts with over 10,000 shares. This sample had 69,114 articles. We found this sample did have a higher correlation of total shares to referring link domains, the correlation was 0.101 as shown in the table below. Thus there is some evidence that for very highly shared posts there is a higher correlation of shares and links.

Content Sample	Average Total Shares	Median Shares	Average Referring domain links	Median Referring domain links	Correlation Total Shares – Referring Domains
Random sample of posts (99,941)	257	8	0.75	0	0.018
Full sample of posts (757,317)	4,393	202	3.77	1	0.021
Posts with over 10,000 total shares (69,114)	35,080	18,098	7.06	2	0.101

We did the same analysis with links to see whether there was a difference in the correlation depending on the number of links e.g. does content with a greater number of links have a higher correlation of links and shares. Our findings suggest that highly linked content does not have an increased correlation. For example, content with more than 10 referring domain links had a correlation of 0.014, while content with less than 4 referring domain links had a correlation of 0.063.

WHY ARE OUR RESULTS DIFFERENT FROM PREVIOUS STUDIES?

The lack of a positive correlation between shares and links is surprising and appears to contradict findings of previous studies, including our own, that found a strong positive correlation of shares and links. However, it does support the view of commentators such as AJ Kohn who argues we should not expect a correlation as people share and link for different reasons.

Our view is that the previous studies may have drawn content from a narrow sample of sites. We did find some popular domains and types of content that do have a high correlation of shares and links. If a sample was disproportionately drawn from these popular sites this would suggest a positive correlation of shares and links. This is explained below.

THE CONTENT SWEET SPOT – CONTENT THAT GETS SHARES AND LINKS

Our findings do not suggest sharing and linking to content is mutually exclusive. It rather suggests that people while may generally link to or share content for different reasons, there is still an area of overlap, where people both share and link to content.

SHARES   LINKS
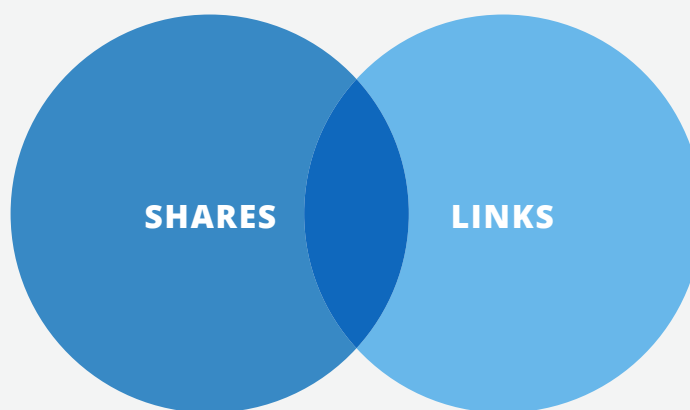
This area of overlap explains why we had different results in our previous study. We only looked previously at a relatively small sample of content that came from just 25 domains. Those domains were primarily sites with a high correlation of links and shares such as the New York Times, BuzzFeed, Hubspot etc. These sites had a high correlation of shares and links. However, our findings suggest these sites are outliers and when we look at a broader sample of over 600,000 domains we find no such correlation of shares and links.

WHICH DOMAINS HAVE A HIGH CORRELATION OF SHARES AND LINKS?

We were keen to look in more detail at the domains that had a high correlation of shares and links. These included very popular sites that we had focused on disproportionately in our previous study. For example, in the following table the four popular domains have a strong positive correlation of shares and links compared to the sample average.

Site domain	Number of articles in sample	Referring domain links – Total Shares Correlation
New York Times	1332	0.55
TechCrunch	119	0.52
Hubspot	744	0.55
BuzzFeed	9087	0.42
All domain articles	**757,317**	**0.02**

We should note that a positive correlation only means that as shares rise, links also tend to rise and vice versa. It does not mean a site gets both high shares and links. In fact you get big variations as we can see below.

Site	Average Total Shares	Average Referring domain links
New York Times	3,665	7.2
TechCrunch	2,072	21.6
Hubspot	1,406	24.33
BuzzFeed	20,767	8.36

We can see very clearly that sites such as BuzzFeed get very high numbers of shares but relatively low numbers of referring domain links. However, they are positively correlated in that as shares rise so do referring domain links.

WHY DO THESE SITES HAVE A HIGH CORRELATION OF SHARES AND LINKS?

Our immediate assumption was that these sites have a high correlation of shares and links because they are popular site and authority sites. However, when we looked at sites with a high domain authority (as determined by Moz) we found this was not the case.

As the table below shows that high domain authority sites get more shares and links as we would expect, but they actually had a slightly lower correlation of shares and links than the average.

	Average Referring domain links	Average Total Shares	Correlation shares & referring domain links
Domain authority > 90	3.73	10,916	0.026
Domain authority > 60	3.467	6,862	0.026
All domains	2.3	4,713	0.028

Domain authority as such does not appear to explain why certain sites get a high correlation of shares and links.

We thought another reason for a high correlation of shares and links could be because of the content's reference quality or authority. We therefore looked at the correlation of total shares and referring domain links for sites such as Wikipedia and found the correlation was lower than average at 0.013.

Our next step was to look at sites with very high correlations to see if the nature of the site or the content might give us more insights into high correlations.

SITES THAT HAD A VERY HIGH CORRELATION OF SHARES AND LINKS

In our initial sample we found some sites that had a very high, almost perfect correlation of shares and links.

Site	Number of articles in sample	Referring domain links – Total Shares Correlation
The Breast Cancer Site	17	0.90
New York Review of Books	11	0.95
Pew Research	25	0.86
The Economist	129	0.73

What is about these sites that generates a very high correlation of shares and links? It appears they are all respected sites that produce regular content about the latest developments in their areas. They appear to be serious sites, that often reference research or evidence backed content or surveys. However, the sample sizes were too small to draw any real conclusions which led us to a further sampling exercise.

REVIEW OF HIGH CORRELATION DOMAINS AND CONTENT

We were very interested to explore the domains and content types that received high correlations. We therefore pulled much larger data samples for specific domains and content types. Broadly these content types could be defined as research backed posts, opinion pieces, and authoritative, opinion forming journalism.

These new larger samples of content had lower average shares and links and a lower correlation of links and shares, though still positive. Below is the data for over 45,000 posts from two major publishers the New York Times and The Guardian.

Domain	Number articles in sample	Average Total Shares	Average Referring domain links	Correlation of Total Shares to Referring domain links
Nytimes.com	49,952	918	3.26	0.382
Theguardian.com	46,128	797	10.18	0.287

These sample sizes were large enough to allow us to subset them into specific content types. One form of content we were keen to explore was opinion content such as editorials or pieces by columnists.The data for this subset for the New York Times and The Guardian were as follows.

Opinion Content	Number articles in sample	Average Total Shares	Average Referring domain links	Correlation of Total Shares to Referring domain links
Nytimes.com	4,143	3,990	9.2	0.498
Theguardian.com	19,606	1,777	12.54	0.433

We can see in both cases that the number of shares increased significantly for opinion content. The number of referring domain links also increased along, significantly in the case of the New York Times. In both cases the correlation of shares and links also increased.

Why does opinion content get higher shares and referring domain links. One thought is that opinion content is typically focused on topical issues that are already engaging an audience. The content is also more likely to take a particular slant or provocative viewpoint. These factors may partly explain why the content performs so much better in terms of shares and referring domain links.

We were also interested to explore what we might call authoritative or opinion forming journalism. It is very difficult to define such content so we looked specifically at certain domains that are known for such journalism. We looked at a number of examples including the Atlantic and New Republic. The Atlantic was created in 1857 as a commentary magazine focused on cultural and literary matters. The magazine has a reputation for producing high-quality review journalism. New Republic was established in 1914. The magazine traditionally focused on politics and the arts. The current magazine says about its content:

> *Our journalism inspires the next generation of decision makers by diving deeply into the issues that impact our communities and our culture. We fixate on the topics that matter. We spark conversations with our audience – in print, in person, and across devices.*
> *The New Republic doesn't just report the news. We take a stance.*

The figures for these two magazines are as follows:

Domain	Number articles in sample	Average Total Shares	Average Referring domain links	Correlation of Total Shares to Referring domain links
TheAtlantic.com	16,734	2,786	18.82	0.586
NewRepublic.com	6,244	997	12.8	0.529

What is immediately noticeable is the high number of referring domain links that these two publications achieve on average. Both are higher than opinion content on the New York Times for example.

The other form of content that appeared to have a high correlation of shares and links, based in part on high referring domain links was research backed content. It is difficult to separate this content out from mainstream sites. Therefore we identified domains that primarily published research backed or evidenced content and reviewed their content.

Below are the results for three domain sites that publish primarily research backed content.

Domain	Number articles in sample	Average Total Shares	Average Referring domain links	Correlation of Total Shares to Referring domain links
FiveThirtyEight.com	1.977	1,783	18.5	0.55
Priceonomics.com	541	1,797	11.49	0.629
PewResearch.com	892	751	25.7	0.4

What leaps out from this table is the very high number of average referring domain links, rising to an average of 25 for Pew Research. Why does this content get such a high level of referring domain links? One view is that people prefer to link to content that is evidence or data backed and in particular research content that provides new insights.

ACHIEVING REFERRING DOMAIN LINKS IS HARD

It appears that the higher correlations are caused when content achieves a higher number of referring domain links. Our study suggests that research backed content and authoritative, opinion forming journalism achieve higher links. More specific research is required but it would appear that serious, deeply researched content which generates new insights or which takes a position on a current topic achieves a higher level of referring domain links.

Previous studies (**https://moz.com/blog/why-big-content-is-worth-the-risk**) have also suggested that comprehensive and evergreen content gains more links. We have not been able to create a sufficient sample of comprehensive, evergreen content to examine in this study. However, it would fit with our findings which show that content needs some form of intrinsic value to achieve links. Also given links tend to be acquired over time, it would follow that evergreen, comprehensive content would also attract higher referring domain links.

Whilst almost everyone can share content, often through the click of a button, not everyone can easily link to content. There is a much smaller number of people that have the ability to link to content whether through their own blog, forums or a business website. What the data appears to show is that acquiring links is difficult. The fact that sites such as Five Thirty Eight and Pew Research achieve over 18 referring domain links on average is a testament to the power of research content.

THE IMPACT OF CONTENT FORMATS

We were keen to explore the impact of content types on shares and links. Our previous study had found that quizzes for example got a very low level of links relative to shares. We were interested to see if the area of overlap in sharing and linking could be partly explained by content type.

Below are the correlations of shares and links by content type.

Content Type	Number in sample	Correlation total shares & referring domain links
List post	99,935	0.092
Quiz	69,757	0.048
Why post	99,876	0.125
How to post	99,937	0.025
Infographic	98,912	0.017
Video	99,520	0.091

The correlation for our random sample of 99,900 posts was 0.018 and the correlation for the overall sample of 757,317 articles was 0.021. Thus it does appear that content formats such as why posts, list posts and videos have a higher correlation. These correlations though are still relatively weak.

Below we look at the performance of the different content types in terms of the numbers of shares and links.

Content Type	Average Total Shares	Median Shares	Average Referring domain links	Median Referring domain links
List post (99,935)	10,734	3,822	6.19	2
Quiz (69,757)	1,374	10	1.60	1
Why (99,876)	1,443	252	5.66	2
How to (99,937)	1,782	406	4.41	1
Video (99,520)	17,708	8,572	4.13	1
Infographic (98,912)	268	24	3.67	1
Random posts (99,941)	257	8	0.75	0
Full sample	4,393	202	3.77	1

It is important to remember that our samples of specific formats represent the top end of a skewed distribution. This means they get a lot more shares than your typical post. However, the data does show that the top list posts and videos get far more shares than the top posts of other content formats.

There is a danger of interdependence here. It could be that well shared sites such as BuzzFeed disproportionately use the top content formats such as list posts. We suspect there is an element of this, though list posts do appear to perform particularly well in all contexts. For example, below is an analysis of the content types for 1,332 New York Times articles from our original sample of 757,317 posts.

Content Type	Average Total Shares	Average Referring domain links	Correlation shares & domain links
List post	11,595	9.99	0.889
Why	6,278	8.83	0.265
How to	3,977	4.87	0.271
Video	10,977	6.6	0.414
All content	3,665	7.2	0.55

Whilst a relatively small sample (1,332 posts), it appears to confirm that list posts and videos achieve relatively higher shares; and that list posts and why posts achieve relatively higher links.

CONTENT THAT GETS SHARES BUT NOT LINKS

In our sample there were 230 articles with over 200,000 shares but only 1 or less referring domain links. We have to be cautious when we look at very small samples. However, there does appear to be some similarity in the post types that acquire large numbers of shares but virtually no links.

The majority of the highly shared posts that didn't acquire links (162 of the 230) were videos. The majority of these were Vine videos, though there were also YouTube videos. This six second content is likely to be entertaining rather than informative and helpful. A number of quizzes also garnered a high level of shares but zero links.

Here are some examples of content that was highly shared but not linked to.

Content	Total Shares	External Links	Referring Domain Links
Vine video https://vine.co/v/O0VvMWL5F2d	347,823	0	0
Vine video https://vine.co/v/O071IWJYEUi	253,041	0	1
Disney Dog Quiz http://blogs.disney.com/oh-my-disney/2014/06/30/quiz-which-disney-dog-are-you-2/	259,000	0	1
Brainfall Quiz http://www.brainfall.com/quizzes/how-bitchy-are-you/1uq7v1/	282,058	0	0

THE IMPACT OF LONG FORM CONTENT

In our overall sample we specifically included a disproportionate number of videos, infographics and quizzes to look at the impact of content formats. In order to undertake our analysis of the impact of long form content we removed these articles. This gave us a sample of 489,128 text based articles which we used for our analysis of content length.

85% OF TEXT CONTENT IS LESS THAN 1,000 WORDS

Our analysis of the 489,128 articles found that over 85% was less than 1,000 words long. Only 2.5% was over 3,000 words.

Length words	No in sample	Percent
<1,000	418,167	85.5
1-2,000	58,642	12
2-3,000	8,172	1.7
3,000-10,000	3,909	0.8

THE IMPACT OF LONG FORM CONTENT ON SHARES AND LINKS

We looked at the impact of content length on total shares and referring domain links. The data was as follows.

Length words	Total Shares		Referring Domain Links	
	Average	Median	Average	Median
<1,000	2,823	195	3.47	1
1-2,000	3,456	478	6.92	2
2-3,000	4,254	578	8.81	3
3-10,000	5,883	566	11.07	3

The data shows that as content length increases from less than 1,000 words to 3,000 words there is an increase in both shares and links. Above 3,000 words the average shares and referring domain links continue to increase though there is no increase in the medians.

The findings confirm previous studies that show long form content gets more shares and referring domain links. What is surprising, given the consistent findings that long form content performs better in terms of shares and links, is that over 85% of content continues to be les than 1,000 words.

There are exceptions and there are many examples where short form content goes viral. The best example we have found is IFLS Science. This site gets a very high number of shares for well curated short form content such as images and videos that explain or highlight a scientific concept or piece of new research. The site achieves relatively low links relative to the high number of shares but there is a strong correlation of shares to referring domain links as shown below.

Domain	Number articles in sample	Average Total Shares	Average Referring domain links	Correlation of Total Shares to Referring domain links
IFLscience.com	2,689	35,055	6.87	0.63

You can read more about how IFL Science creates viral short form content here
http://buzzsumo.com/blog/how-ifl-science-is-nailing-it-with-short-form-viral-content/

## IMPACT OF LONG FORM CONTENT ON CORRELATION OF SHARES AND LINKS

Our analysis of the 489,128 articles found that over 85% was less than 1,000 words long. Only 2.5% was over 3,000 words.

Length words	Correlation Shares/Links
<1,000	0.024
1-2,000	0.113
2-3,000	0.094
3,000+	0.072

These results would support the hypothesis that content of over 1,000 words has a higher correlation but this does not increase beyond 2,000 words.

WHAT DOES THE DATA SAY ABOUT WHY PEOPLE SHARE & LINK TO CONTENT?

Our data suggests that sharing and linking of content is not entirely mutually exclusive and that some forms of content get both shares and links. Thus sharing and linking may look similar to our diagram below.

SHARES **LINKS**
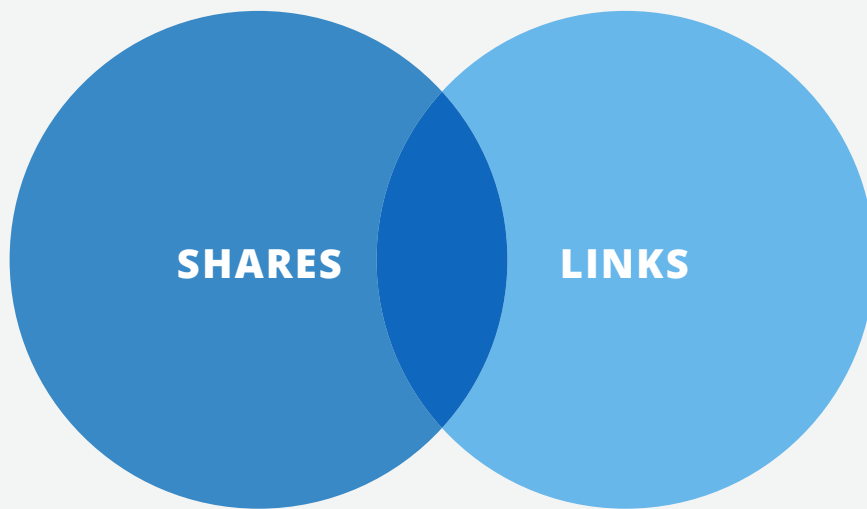
We know from our analysis that some content published by sites such as Pew Research, the New York Times and the Breast Cancer Site fall into the area of overlap.

Can we learn something from our data about why content is shared and why it is linked to?

There have been various research studies on why people share and link to content. The table below sets out some of the studies and views about content sharing and linking.

Why People Share Content	Why People Link To Content
One study undertaken by the Customer Insight Group and the New York Times http://nytmarketing.whsites.net/mediakit/pos/ suggests primary motivations for sharing include: • To share valuable and entertaining content • To grow and nourish our relationships • For self-fulfilment • To get the word out about causes they care about These overlap to a degree with a study by Marketo and Brian Carter http://uk.marketo.com/ebooks/contagious-content-what-people-share-on-facebook-and-why-they-share-it/ that found people share to: Give – offers, discounts Advise – tips, useful information Warn – about dangers Amuse – entertaining content Inspire – inspirational quotes Amaze – amazing pictures or facts Unite – being part of your tribe	Various experts have argued that people are likely to link to content if it is: **01** Highly relevant and useful. https://moz.com/blog/dont-ask-sites-for-links-find-people-and-connect **02** Comprehensive content with evergreen longevity. https://moz.com/blog/why-big-content-is-worth-the-risk **03** Authoritative content in an industry http://www.koozai.com/blog/search-marketing/link-building/why-people-link-tips-to-improve-your-web-content/ **04** Exceptional content **05** Reciprocity. People will link where the linking is about partnership or reciprocity.

These reviews suggest that sharing and linking to content is undertaken primarily for different reasons. However, there are areas of overlap for example valuable and helpful content falls into both reasons for sharing and linking.

The research suggests that social sharing is more personal than linking. Sharing in many ways helps define who you are. A share can highlight issues you are passionate about. A share can show your support for friends or colleagues. Sharing is also relatively easy to do. Often it is no more than a click of a button.

Links by contrast are much harder to acquire. Our study suggests that short form content that does not provide new insights, that simply repeats previous research or provides advice that is not backed by serious study is unlikely to acquire a high level of referring domain links.

Our data would support the view that people primarily link to and share content for different reasons. We have found that:

• Quizzes and entertainment videos are far more likely to be shared than linked to
• Research backed, insightful content and is likely to achieve higher referring domain links

We have also identified some content that achieves both shares and links, this includes serious, well researched, opinion forming journalism and new insights from new research. This tends to fall into the category of valuable and helpful content.

HOW TO CREATE CONTENT THAT GETS SHARES AND LINKS

Overall there is no correlation of shares and links in our sample. However, as we have highlighted in the previous section there are domains and content types that achieve relatively high numbers of shares and links.

Our study suggests there are a number of factors that contribute to content achieving high shares and links.

01 The nature of a domain and its audience is the first factor. Popular sites that produce regular, authoritative content are more likely to gain shares and links.

04 Certain content formats such as list posts and videos can achieve higher shares and links.

02 Serious and authoritative journalism that takes a position on a current topic can achieve more links as well as more shares.

05 Some content types such as entertaining videos and quizzes are more likely to get shares than links.

03 New insights based on deep research can achieve more links.

06 Long form content of over 1,000 words gets both more shares and links.

INCREASING SHARES AND LINKS THROUGH A COMBINATION OF FACTORS

In this research we have not attempted to build or test a linear regression model that could predict an increase in shares or links based on a combination of the above factors. However, there is some evidence that a combination of factors such as say research content which is also both a list post and long form could achieve higher shares and links.

Below is an example of this using on our original New York Times sample. It should be noted the sample size is very small.

The shares and links for our original sample of New York Times posts were as follows.

Average Total Shares	Median Total Shares	Average Referring Domains	Median Referring Domains	Correlation TS/RD
3,665	192	7.2	1	0.55

We identified and created a subset of articles from this New York Times sample that were List posts. The shares and links for this sample of List posts were much higher, particularly the medians, and the correlation increased from 0.55 to 0.74. The figures are below.

Average Total Shares	Median Total Shares	Average Referring Domains	Median Referring Domains	Correlation TS/RD
12,567	3,250	18.5	8	0.74

We then identified and created a further subset of data that were List posts of over 1,000 words. The median shares increased and the correlation also increased to 0.83.

Average Total Shares	Median Total Shares	Average Referring Domains	Median Referring Domains	Correlation TS/RD
11,212	4,682	18.3	8	0.83

Thus it is possible that a combination of factors on a popular domain may result in a higher level of shares and links. The big note of caution is that there are only 40 articles in our final subset of New York Times List posts of over 1,000 words. This is an area for further research and study.