BACKLINKO AND BUZZSUMO CONTENT ANALYSIS:
Methods & Results

BACKLINKO

BACKLINKO AND BUZZSUMO CONTENT ANALYSIS:
METHODS & RESULTS

The goal of this study was to determine how certain content-related factors correlated with social media shares and backlinks.

To our knowledge this is the largest study of this kind (in regard to sample size) conducted. However, this type of research is far from the first of its kind.

In fact, many of the data points that we analyzed were specifically designed to expand upon previous content research findings. This study published on the Moz blog and this study from the BuzzSumo blog being two of many examples.

Our hope is that the data from this research helps publishers make data-driven decisions about the content they produce. We also set out to determine the "current state of the content world". That's why we looked at things like the amount of content that gets linked to and the distribution of social shares.

WHAT WE DID–
STUDY METHODOLOGY

We used BuzzSumo to collect the data for this study. BuzzSumo is a popular and reliable content tool that has a large database of published content.

The software also analyzes and categorizes the content in its database by social media shares, number of backlinks and content format.

To collect the sample of content for our study we used all of the articles published from October 1, 2017 to October 1, 2018 found in BuzzSumo's database. This resulted in 912.5 million total articles.

"Social shares" were based on data from BuzzSumo's database. Specifically, we analyzed the number of social shares from Facebook, Twitter, Pinterest and Reddit. However, we didn't analyze other popular social networks like LinkedIn and YouTube.

"Backlinks" were also based on data from BuzzSumo. Although not a full-fledged SEO software suite, BuzzSumo is, in this author's opinion, a reliable tool for finding backlinks to new pieces of content.

"Content format" was determined by BuzzSumo's automatic categorization of the content in its database.

"B2B content" was based on BuzzSumo's internal database of B2B publishers. A post was considered "B2B content" if it was published on a site in BuzzSumo's B2B publisher database.

When analyzing the data itself, most of our analysis used the exact figure from our calculations (i.e. "3.184180679195516"). However, for data presented in charts, we rounded that number to the nearest integer (i.e. "3.2") in order to make the data easier to visualize.

For our single correlation metric (the correlation between shares and links) we used Pearson correlation coefficient. This type of analysis is ideal for determining a potential relationship between two variables.

FACTOR-BY-FACTOR
BREAKDOWN

CONTENT TYPE

The content format among 6 options: Why Post, List Post,
What Post, Video, Infographic, How-to Article.

HEADLINE LENGTH

The length of a title based on word count and character count.

CONTENT LENGTH

The total word count of a piece of content.

PUBLISHED DATA

The day a post first appeared in BuzzSumo's database.

SOCIAL SHARES

Total number of shares on Facebook, Twitter, Pinterest and Reddit.

BACKLINKS

The number of referring domain links pointing to a page.

RESULTS

CONTENT TYPE AND SOCIAL SHARES

CONTENT TYPE	MEAN SHARES
List Post	262.9
Why Post	206.7
What Post	138.4
Video	92.9
Infographic	86.7
How to Article	82.6

CONTENT TYPE AND LINKS

CONTENT TYPE	MEAN REFERRING DOMAINS
Why Post	3.184180679195516
What Post	2.9380307838980393
Infographic	2.828271207240374
Video	2.4712422679121486
How to Article	2.2697566745952
List Post	2.098993544883146

HEADLINE LENGTH (WORD COUNT) AND SOCIAL SHARES

HEADLINE WORD COUNT	MEAN SHARES
15	215.54429835366906
17	209.82441608604367
16	206.4449490879241
14	194.3342731846329
13	179.1873891490749
8	176.78927518280338
11	175.13990828562586
12	166.5931492528674
2	160.19564763349757
10	146.06603025998206
7	142.0411030713075
9	133.09621591027354
6	130.4828225785801
3	115.46188085365074
5	112.37346296844717
1	101.69187235239768
4	94.72016337206927

Note: "100"=95-100 characters, "0"=0-5 characters.

CONTENT LENGTH AND SOCIAL SHARES

WORD COUNT	MEAN SHARES
1000-2000	230.03852800449698
2000-3000	200.5098418129616
3000-10000	185.243357969077
0-1000	147.29009140855436

CONTENT LENGTH AND LINKS

WORD COUNT	MEAN REFERRING DOMAINS
3000-10000	4.058744785756686
100000	3.948707047928874
2000-3000	3.552957853556552
1000-2000,	3.1170741651425318
0-1000	2.2899765035962782

PUBLISHED DAY OF WEEK AND SOCIAL SHARES

DAY OF WEEK	MEAN SHARES
Sunday	152.54957995413747
Wednesday	151.52808027145588
Thursday	150.82038156256704
Monday	150.311207887163
Tuesday	150.21701208591404
Saturday	150.19386168636632
Friday	149.0654287226236

CORRELATION OF SOCIAL SHARES AND BACKLINKS

Correlation is 0.078

% CONTENT THAT GETS AN EXTERNAL LINK

Full Dataset:

6% of content gets >= 1 link

3.8% of content gets exactly 1 link

B2B publishers:

7% of B2B content gets >= 1 link

4.% of B2B content gets exactly 1 link

AVERAGE SOCIAL SHARE COUNT PER POST

Full Dataset: 150.6 average social shares per post

Note: This is likely skewed by a few massively popular articles.

B2B Publishers: 14 average social shares per post

SOCIAL SHARE DISTRIBUTION

Full dataset:

0.1% of articles get 50% of social shares

1.3% of articles get 75% of social shares

B2B Publishers:

0.5% of articles get 50% of social shares

2% of articles get 75% of social shares