

Using Open Data to Create Content That Educates, Builds Links, Earns Press, and Converts

Kristin Tynski



traffic
think tank
Live



Kristin Tynski

Founder and SVP Creative
at Fractl
Kristin@frac.tl



traffic
think tank
Live

I am one of the founders of Fractl, a leading content marketing and PR agency with a data-journalism centered approach to earned media, link building, outreach, and on-site content.

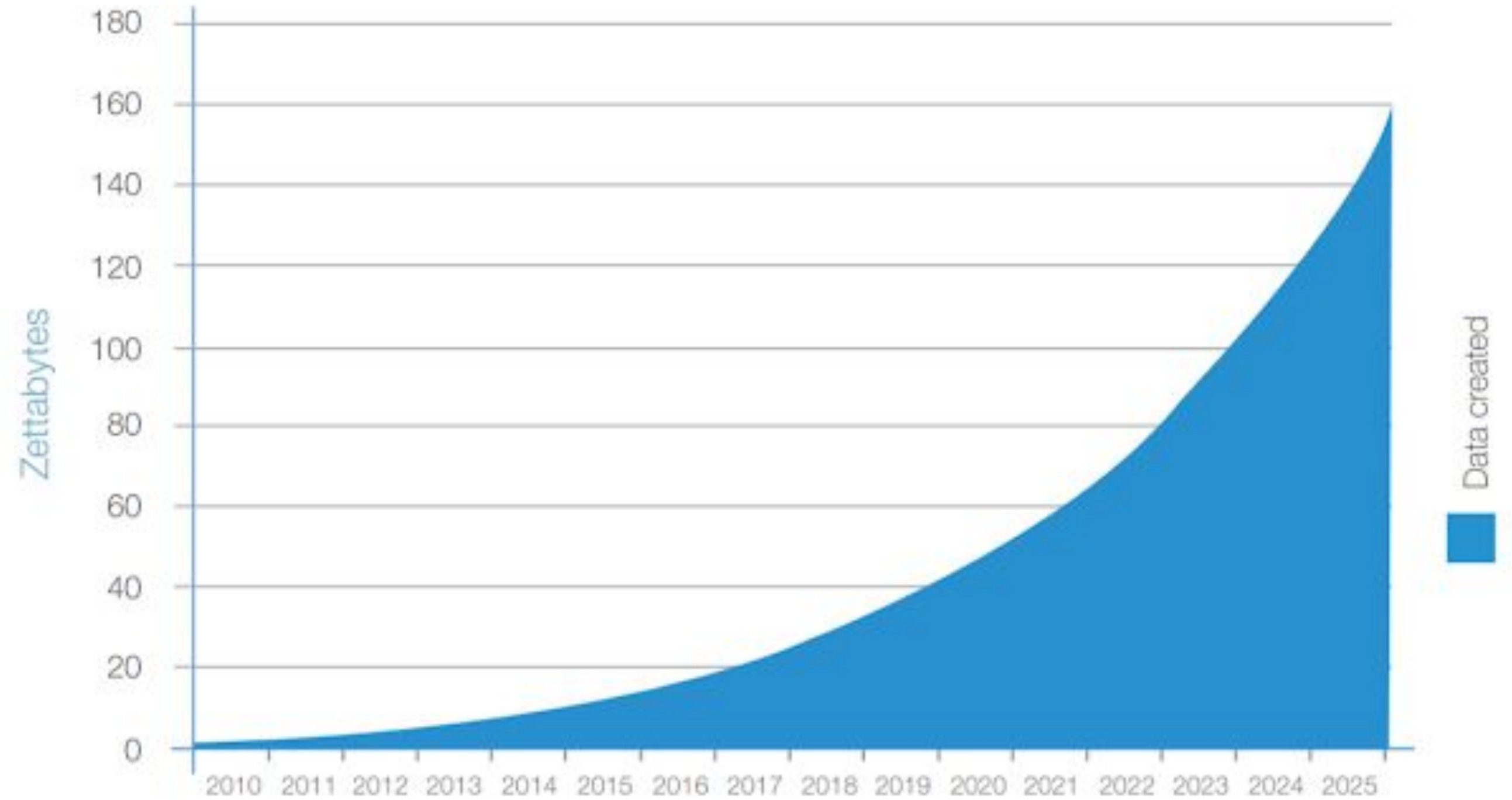
Big, open data has given the world a treasure trove of information from which we can draw on to create unlimited numbers of useful, important, fascinating, emotionally compelling, or otherwise newsworthy brand stories.

Done well, these data-driven stories can be the most consistent and successful content marketing executions from the top to the bottom of the funnel – from link building to conversions.

#TTTLive

160_{ZB} by 2025

Content creation volume is on an exponential curve.

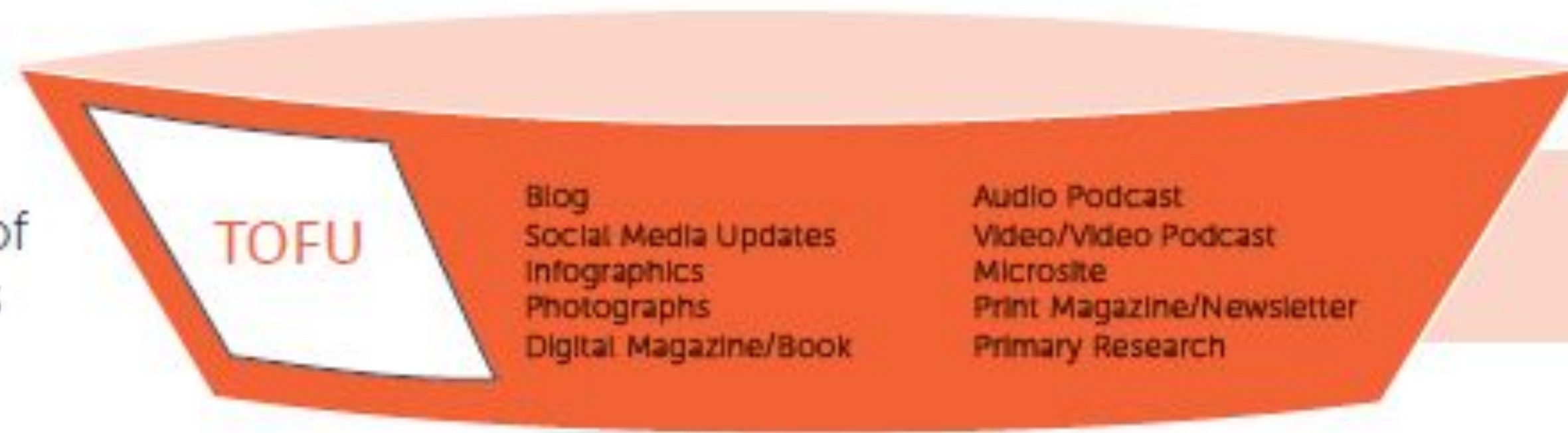


Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

It's Getting Harder to Stand Out,
and You Need a Secret Weapon.

Stand Out by Leveraging
Data in Content Marketing
Throughout Your Entire Funnel

Content at the top of the funnel facilitates **AWARENESS**



Make **PROSPECTS** problem aware & solution aware

Content in the middle of the funnel facilitates **EVALUATION**



Convert problem aware & solution aware prospects into **LEADS**

Content at the bottom of the funnel facilitates **CONVERSION**



New leads can now make an **Informed purchase decision**

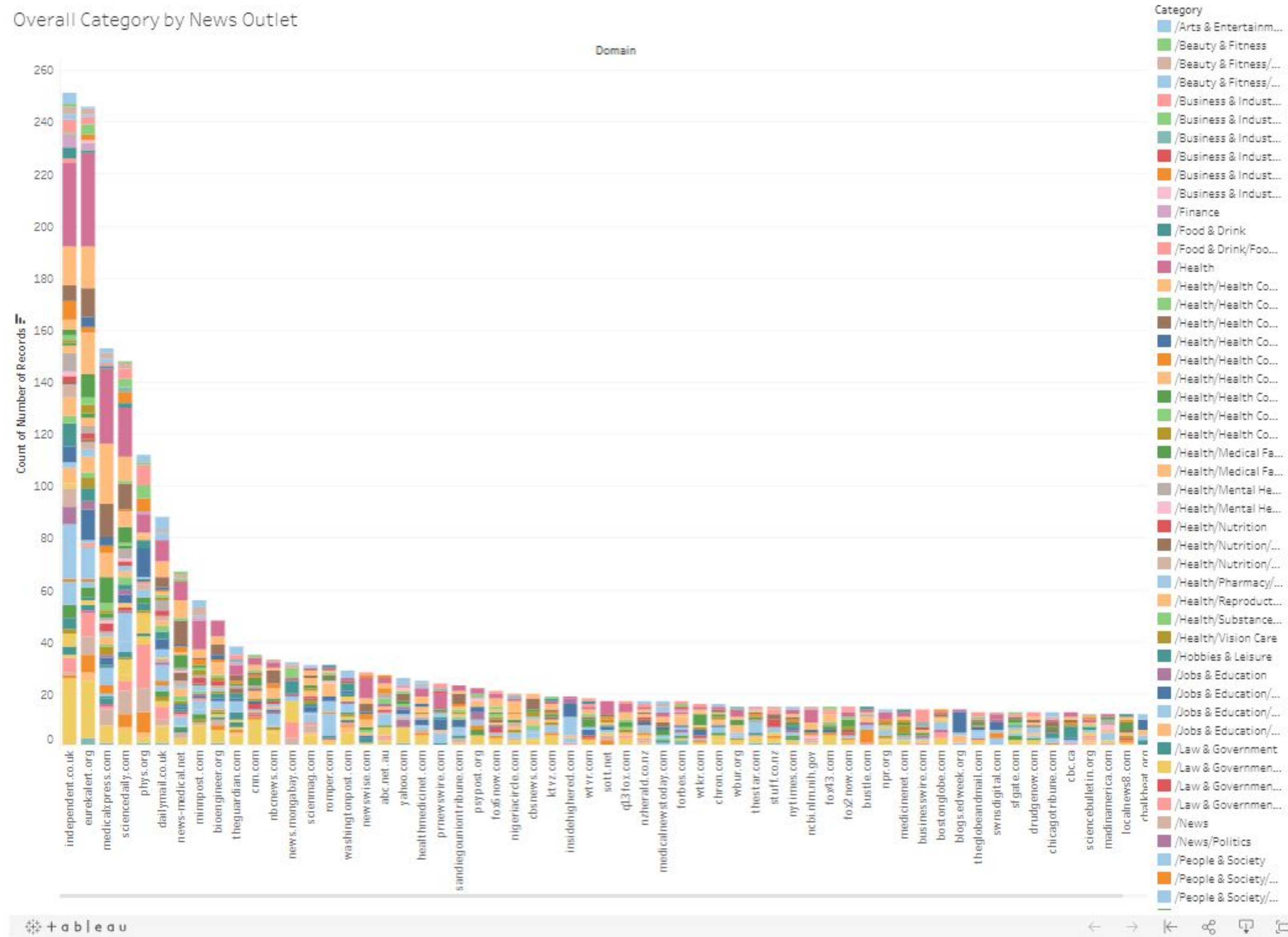


Examples in Action – Data Journalism Throughout the Sales Funnel

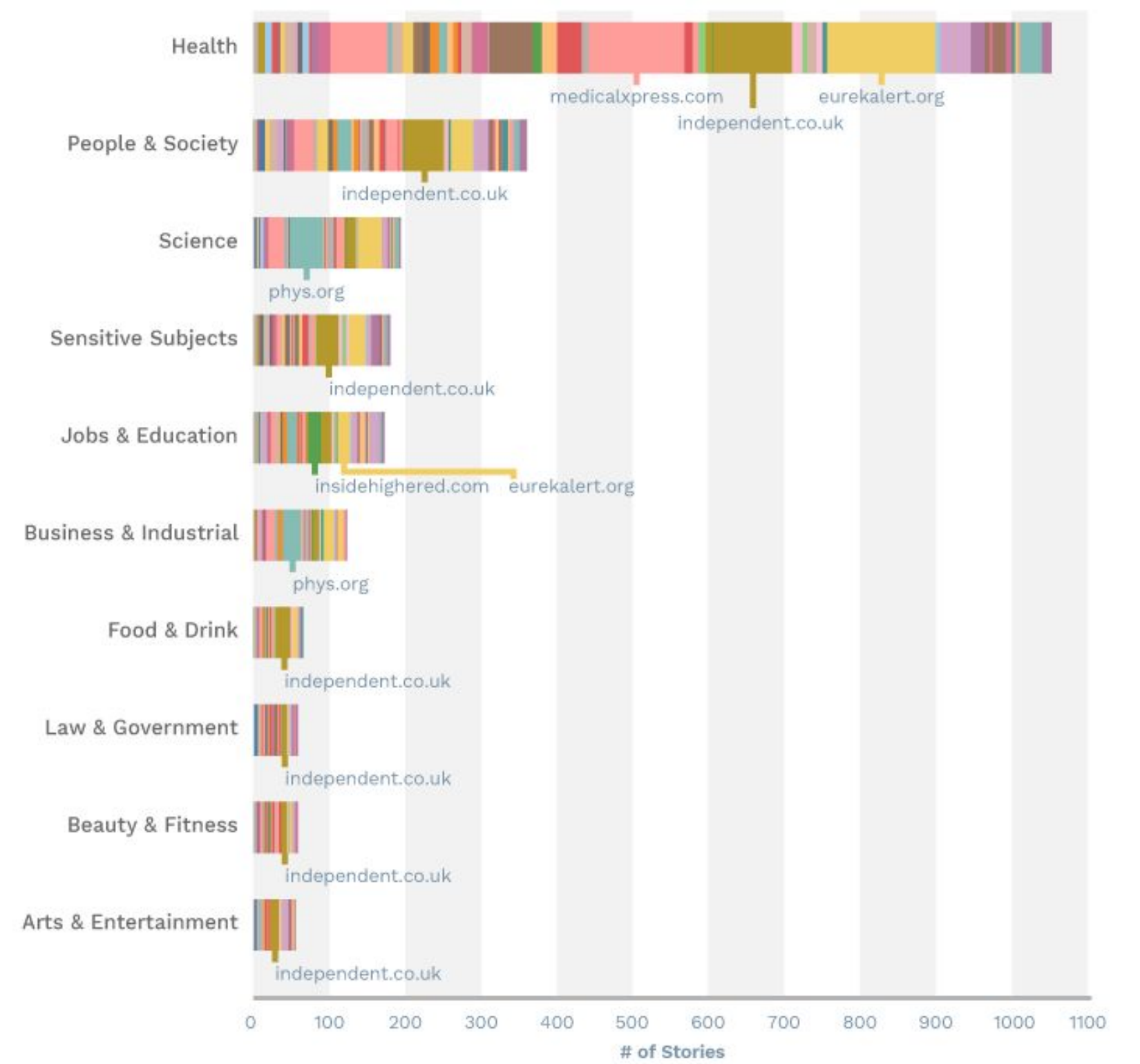
TOFU – Data Content Marketing for the Top of the Funnel

Huge Demand for Data-Driven Stories From Publishers

Overall Category by News Outlet

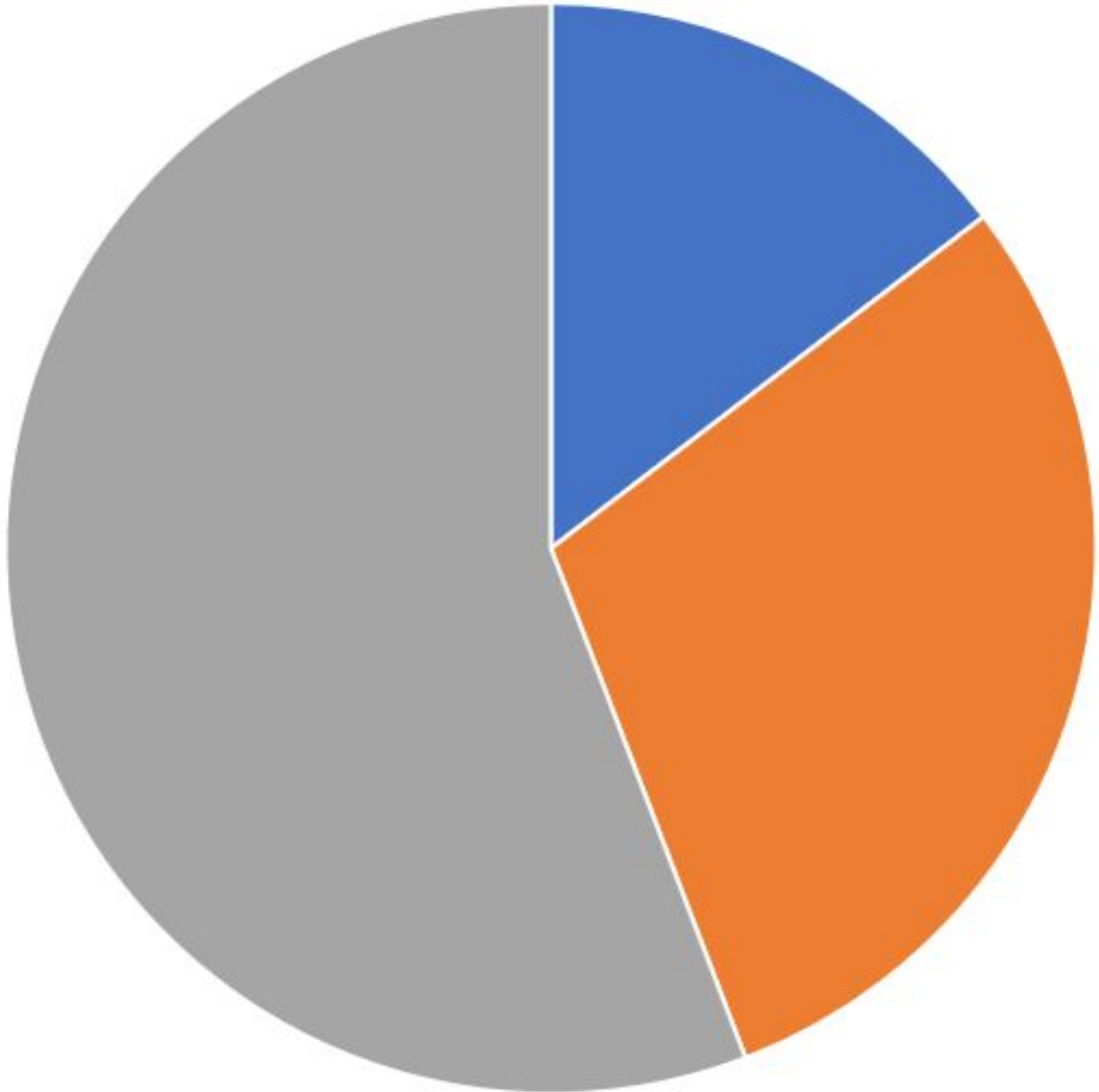


Popular Publishers in the Google Classifier's First Categories

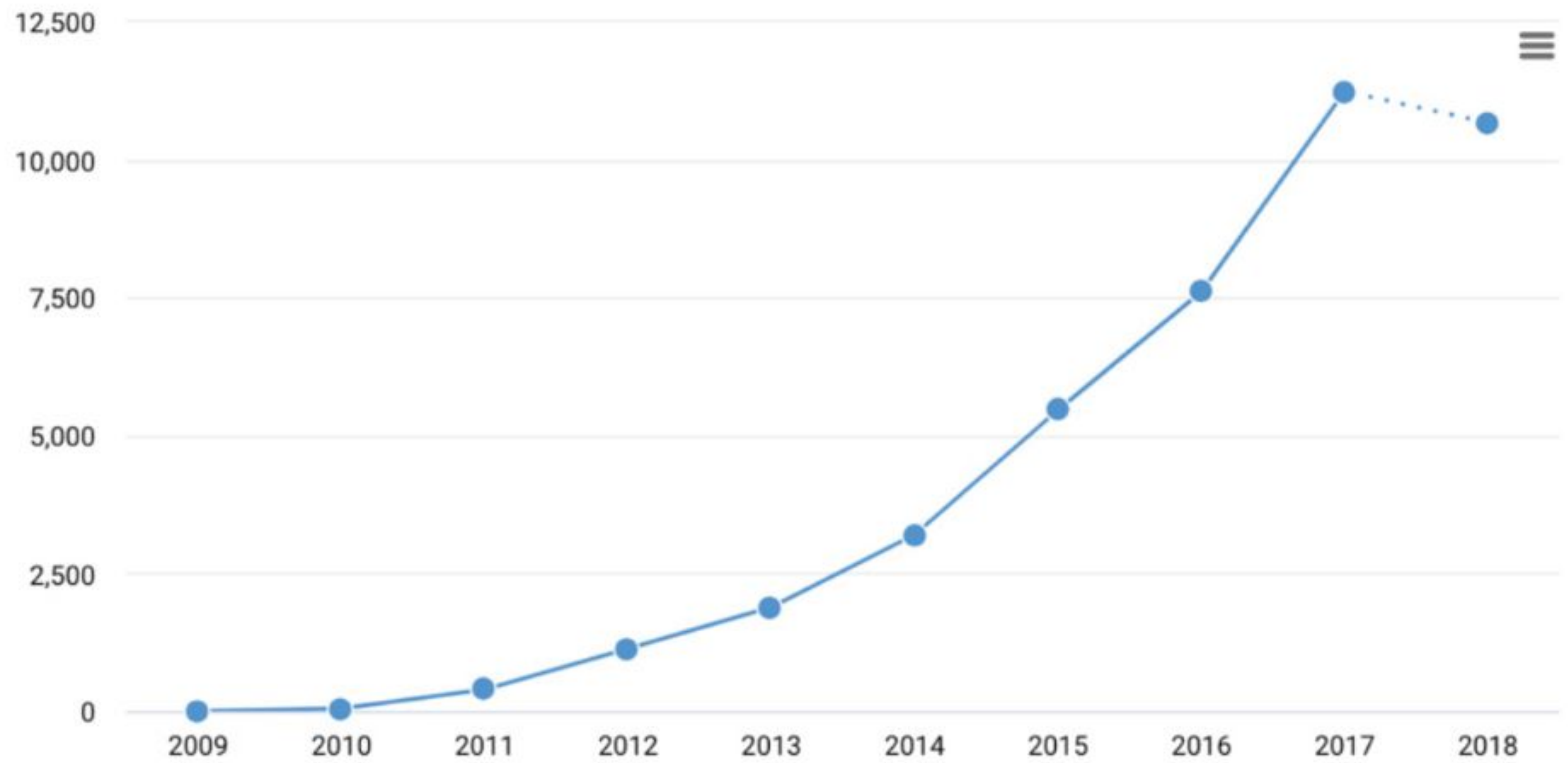


Data Everywhere

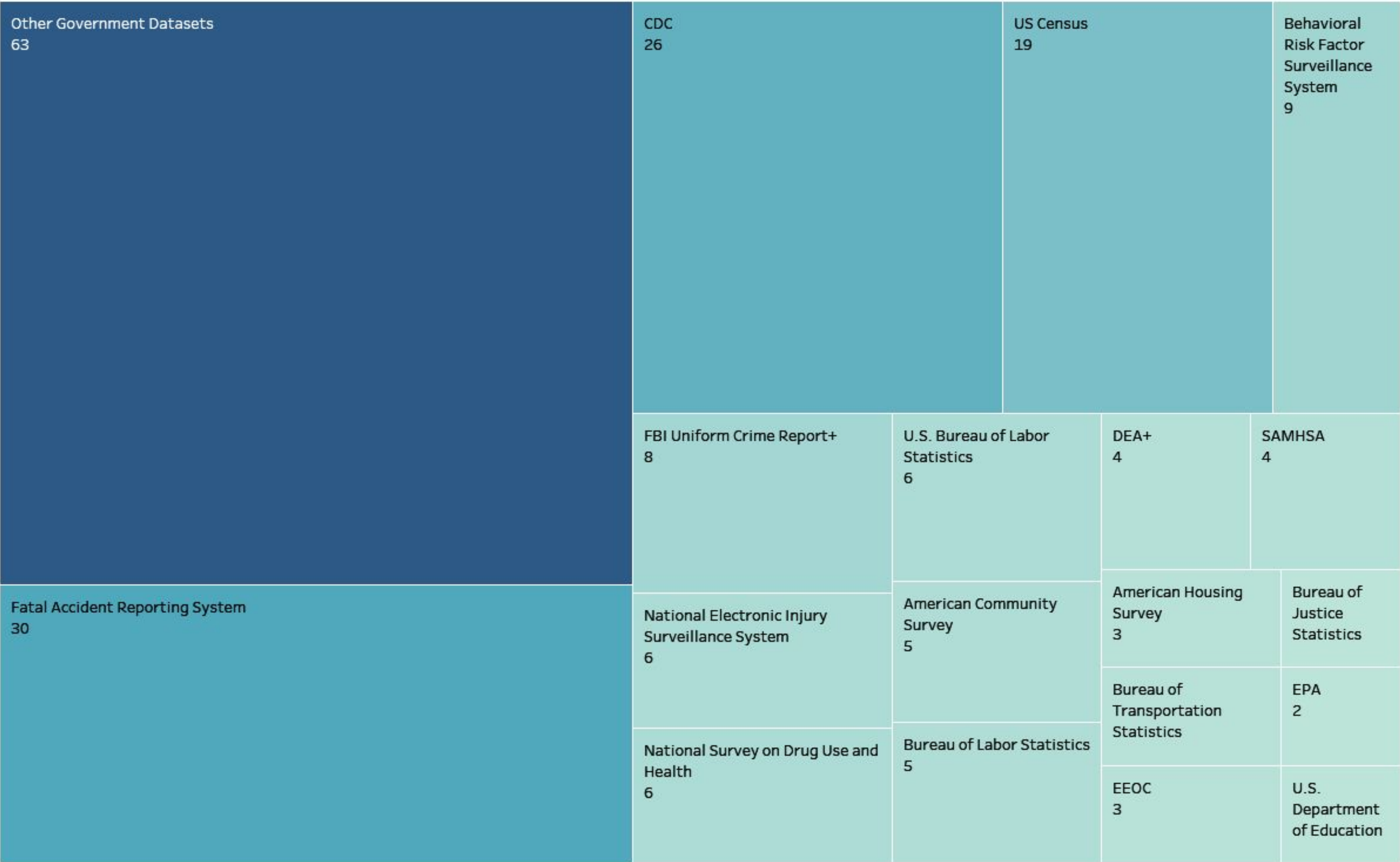
Fractl Campaigns - Distribution of Data Sources 2015 - 2019



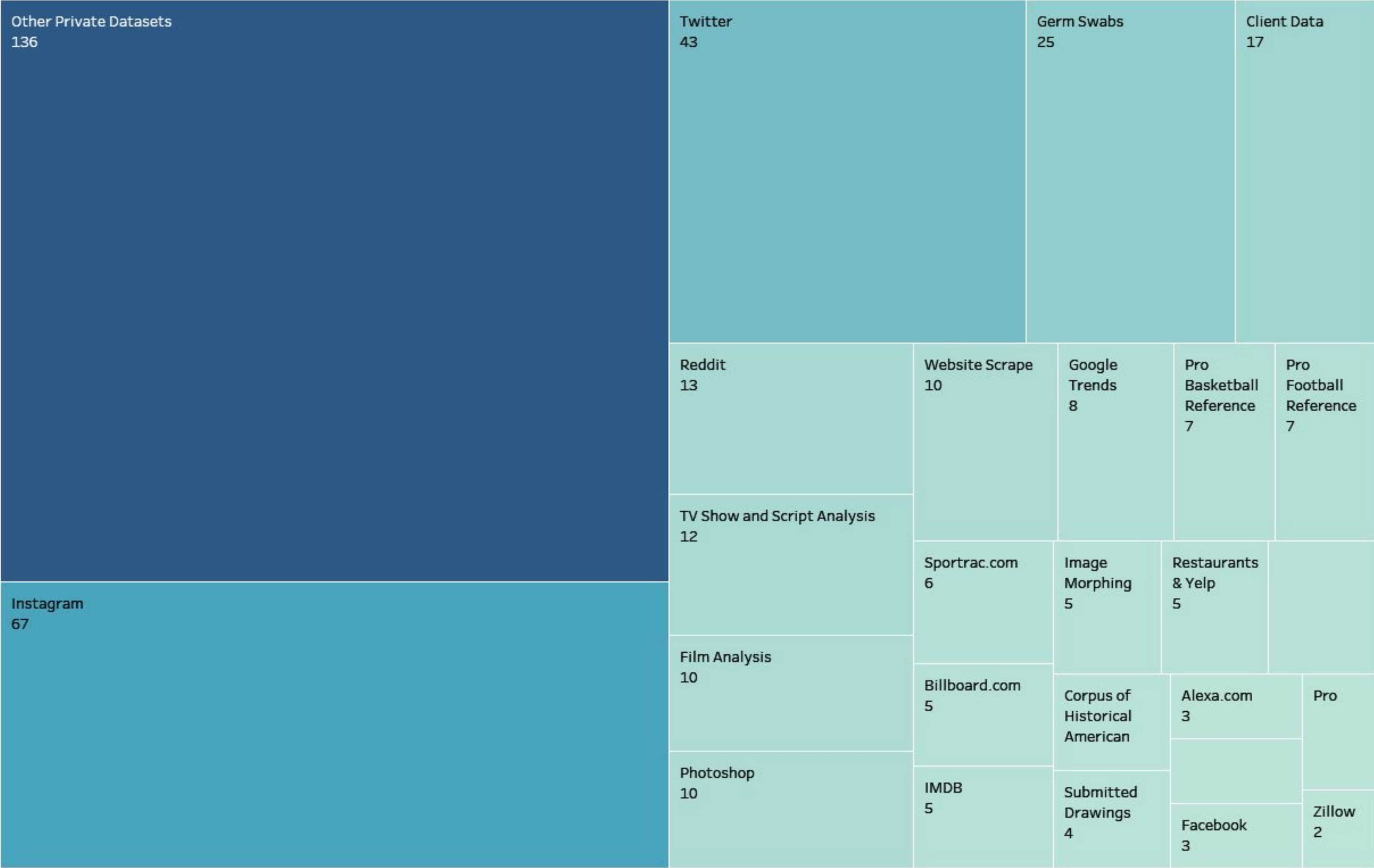
■ Government Datasets ■ Private Datasets ■ Surveys



Governmental Sources



Public Sources



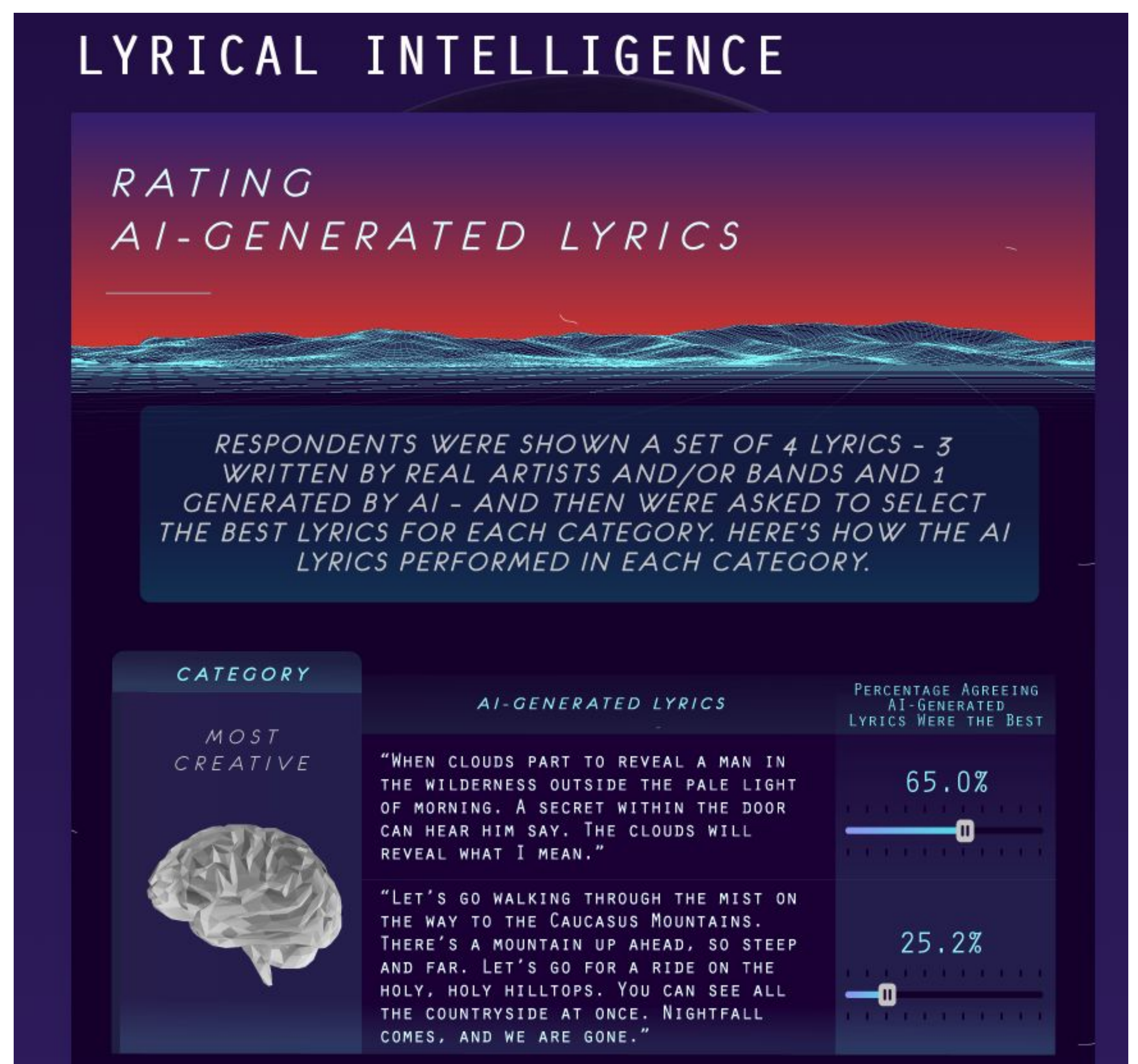
Vetting an Idea and Corresponding Data Source



- Simple – Can I do the analysis I hope to with this dataset? Will my results be understood by my target audience?
- Unexpected – Will my results present something new and interesting? Am I just rehashing something everyone already knows?
- Concrete – Do I have a clear thesis? A clear goal/point to make?
- Credible – Is my data authoritative? Are my methodologies sound?
- Emotional – Does the data have relevance to the human experience?
- Story – Am I just showing data, or am I weaving my data into a story/narrative?

TickPick.com – AI Drops an Album

- **Data Source 1:** Scraped lyrics from Genius.com used to fine-tune GPT-2 774M Model
- **Data Source 2:** Survey of 1,000 respondents who each evaluated the AI-generated lyrics
- **Url:** <https://www.tickpick.com/ai-drops-an-album/>
- **Top-Tier Press:** The Next Web
- **Unique Linking Domains Earned:** 98



The Pitch That Secured Coverage

- **Personal Connection:** Made it obvious we understood he wrote about AI and had previously written an article about Microsoft's AI text rewriting AI
- **Hooks:** 3 specific surprising data results.
 - 36.9% thought the lyrics were more emotional.
 - Only 12.3% could tell the country lyrics were AI-generated.
 - 25.6% incorrectly thought Billie Eilish penned the lyrics written by the AI.
- **Pitches to Secure Exclusive:** 35

Hey Thomas,

I think it's awesome that Microsoft Word is using AI to rewrite users' sentences. Although, I can't help but wonder what that could mean for the jobs of human editors. I'll be curious to see how accurate it is.

Based on your AI coverage, I wanted to reach out and pass along a new [study](#) that just came across my desk. It seemed right up your alley, and I'd love to offer you the exclusive first.

Artificial intelligence has taken the tech world by storm, but can it outperform its human creators in more creative fields, like music?

My team at TickPick pulled more than 1,000 rock, hip-hop/rap, country, and pop songs to train "AI artists." We then surveyed 1,003 people to rank creativity, emotionality, and favorability of these AI-generated lyrics. We also tested to see how easy or difficult it was to tell the difference between some of mankind's most famous songwriters and AI lyrics still in their infancy. Some findings:

- 39.6% of people also thought that **AI-generated lyrics were more emotional**. However, when it comes to people's favorite lyrics, man-made lyrics still reigned supreme.
- Rap lyrics were the most difficult for AI to replicate; pop music was second, followed by rock. Country music was the easiest for AI to replicate.
- Only 12.3% of people could tell the country lyrics were AI-generated. 22.2% of people even thought the lyrics were written by Garth Brooks.
- When presented with AI-generated pop lyrics pop, only 14.8% of people could tell the lyrics were AI-generated. 25.6% of people incorrectly thought Billie Eilish penned the lyrics written by AI.

Here's [the full study](#).

Many respondents found AI-generated lyrics to be more emotional than those penned by legends like Johnny Cash. Artificial intelligence is changing the way we do things, and I think your readers may be interested in seeing its potential impact in the music industry.

Are you interested in the exclusive at The Next Web? Would love to help make this the perfect fit.

Cheers,



Example 1 Takeaway: AI + Data = Endless Possibilities

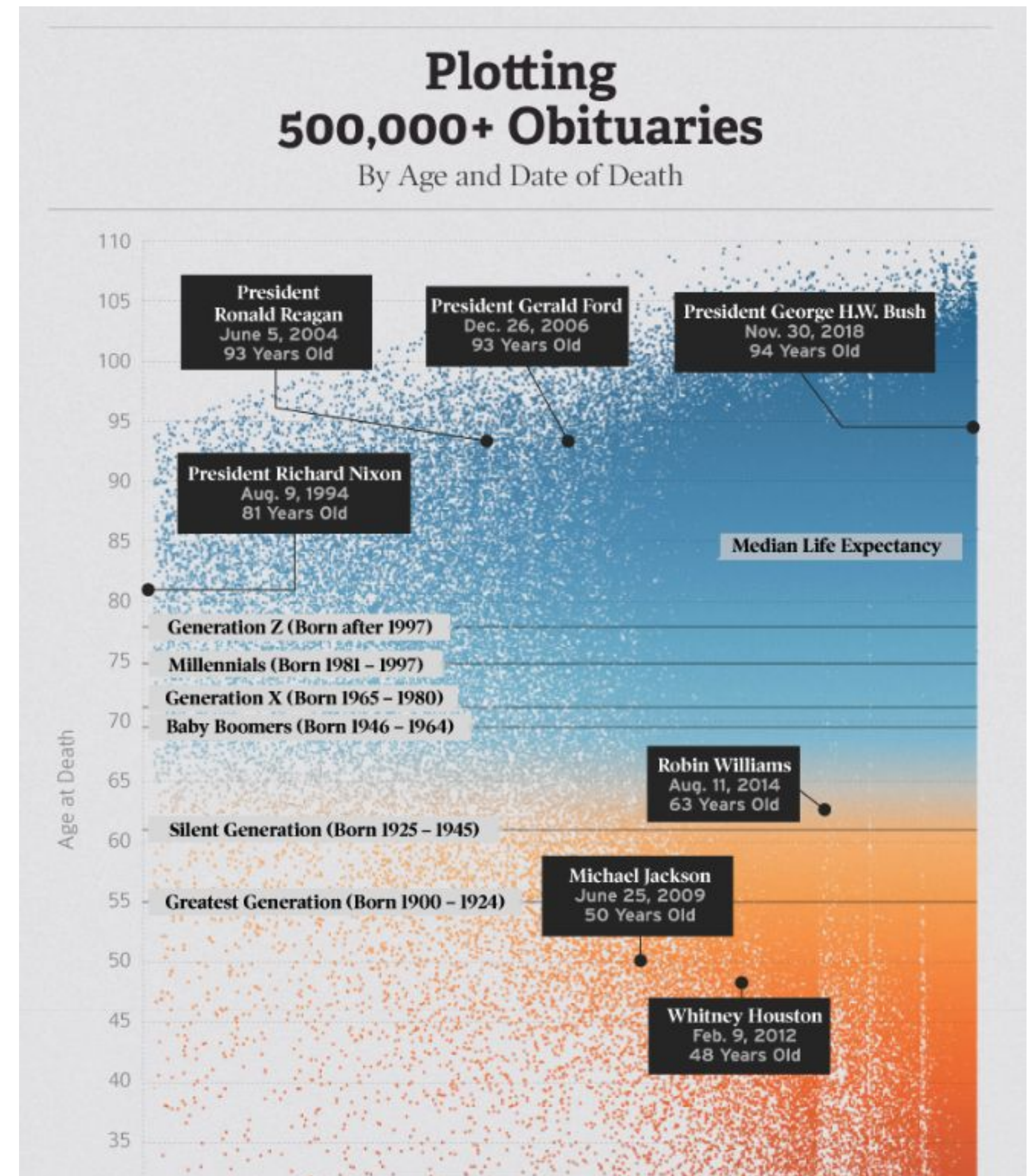
State-of-the-art AI models can now be used easily by anyone with even basic Python knowledge.

Ways to leverage AI for content:

- Generate human-like [written text](#)
- Generate human-like [music](#)
- [Clone voices](#) and [deep fake faces](#)
- “[Imagine](#)” images
- Summarize or [paraphrase](#) existing content automatically

BayAlarmMedical.com – What Can We Learn from 500K Obituaries?

- **Data Source:** Public repository of NYTimes Obituaries
- **Url:** Bayalarmmedical.com/medical-alert-blog/what-we-can-learn-from-obituaries/
- **Top-Tier Press:** Bustle.com



The Pitch That Secured Coverage

- **Personal Connection:** Specific detailed response answering questions by reporter in HARO request
- **Hooks:** Details from our exclusive study of the language used in obituaries answered concrete questions the reporter had with data.
- **Pitches to Secure Exclusive:** 74

Hi Carolyn,

I hope you're doing well! I wanted to reach out to you after seeing your Friday afternoon HARO query regarding charming southern phrases we should all be using. My team at Bay Alarm Medical recently analyzed over 500,000 obituaries, shedding further light onto the linguistic elements that embody southern culture—and where those phrases originated from. The study is currently password-protected (password: exclusive), so if you are interested in using any information from it please let me know and I'll have that removed.

Answers to your questions:

1. What are some charming phrases from the South that might be fun to use more often? For each, please briefly explain what it means and where it might come from.

Perhaps one of the most well-known Southern phrases is "Bless Your Heart"...although in popular culture, it is portrayed as more of a sarcastic remark than a genuine expression of sympathy. But per a recent [study](#) by Bay Alarm Medical, it's likely this phrase has its roots within the South's protestant culture—the terms "Christian" and "First Baptist Church" are two of the top ten terms you're most likely to see in an obituary, marking a longstanding tradition of the importance of religion in everyday life. In light of this, this phrase's meaning is clear: Unless it specifically precedes an insult to soften the blow, "Bless Your Heart" is a genuine expression of fondness.

Another lesser known Southern phrase is "Somebody's Darling", which has its origins in the Civil War. This may not come as a surprise, as just like much of the United States, the remnants of war are [still visible](#) within the South's culture—"War" is mentioned in more than 1-in-5 paid obituaries, and "United States Army" has more than double the mentions than the next most-popular term. The phrase was originally a comment when observing a deceased soldier, and is now used to lovingly refer to someone's partner.

2. Is the South known for its charming phrases? Why might it be fun to use these phrases more often, even if you aren't from the South?

The South is often perceived as the last stronghold of a time that's long been forgotten, where the magic of chivalry and charm are still present in everyday life. Using popular Southern phrases, even if you aren't from the South, is a fun way to capture that feeling no matter where you are. There's no person alive that won't smile at being told they're "As pretty as a peach" or "A tall drink of iced tea". I thought this information might be helpful as you write your article for Bustle, so I wanted to pass it along. Please let me know if you have any questions or need additional information—I'm more than happy to help!

Best,

Example 2 Takeaway: Leverage Existing Datasets to Tell Stories about Human Behavior



When aggregated, even something like obituaries can reveal previously unknown trends in individual and societal human behavior. Open data is enabling a blossoming of DIY data sociologists and anthropologists.

- Things to ask yourself when analyzing a dataset:
 - What can this data tell me about individual or cultural behavior?
 - What can I find in this dataset that is emotionally compelling?
 - Does this data contradict commonly held beliefs?

Casino.org – Most Vulgar NBA Fans

- **Data Source:** Scraped Subreddit comments
- **Url:** <https://www.casino.org/nba-most-toxic-fans/>
- **Top-Tier Press:** TorontoSun.com, CalgaryHerald.com
- **Unique Linking Domains Earned:** 41



MOST VULGAR NBA FAN BASES

Average Number of Profane Words per 100 Subscribers of Each Team's Subreddit

Los Angeles Lakers		 52.7
Philadelphia 76ers		 45.5
Phoenix Suns		 38.2
Toronto Raptors		 37.1
Miami Heat		 35.0
Denver Nuggets		 31.0
Boston Celtics		 30.0
Oklahoma City Thunder		 29.9
New York Knicks		 26.1
Milwaukee Bucks		 25.0

The Pitch That Secured Coverage

- **Personal Connection:** Opened with comment about Danny Green's dunk, something our outreach person saw on the journalist's Twitter
- **Hooks:** The Lakers have the most vulgar fan base in the league, according to our exclusive data. They drop the "B" word 14.9% more than the league average.
- **Pitches to Secure Exclusive:** 26

Hey Anthony,

I, too, was pretty shocked by Danny Green's dunk! I never would have thought he'd get drug tested as a result, though. I definitely agree that they shouldn't really call it a "random" drug test—seems a bit unfair considering the NBA's move here.

I tend to get a bit passionate about these issues, though. We all get excited watching our favorite team, and sometimes we get caught up in the heat of the moment and let the curses fly. But where do Lakers fans rank when it comes to dropping profanities? To find out, we analyzed each NBA teams' subreddit tracking the use of vulgar words per 100 subscribers.

Here's how the Lakers stack up:

The Lakers have the most vulgar fans in the league, dropping an average 53.7 profanities per 100 subscribers in their subreddit. B--tch please, Lakers fans are dropping b-words 14.9% more often than the league average. Lakers fans also stood unrivaled when it comes to using the league's third most common curse, "damn" -- they used this at a 23.6% higher rate than the league average. Kings of curses-- Lakers fans use profanity 3x more on Reddit than Clippers fans.

You can view the [full findings here](#).

While the Lakers battle it out on the court, your readers will be curious to know where they rank when it comes to dropping support. Are you interested in sharing these exclusive findings with your readers at Silver Screen and Roll?

Thanks so much,

Example 3 Takeaway: Scraping Social Exhaust Can Unearth Amazing Insights



Social media exhaust can provide nearly endless fodder for explorations around human behavior. Scraping social sites and apps can frequently be done easily, and modern NLP tools make analysis relatively simple.

- Common Social Scraping Sources for Fractl
 - Twitter, Instagram ([Twitter scraper](#), IG Scraper)
 - Reddit ([reddit github scraper](#))
 - YouTube transcriptions and comments ([scraper](#))
 - Forums ([great generic web scrapers](#))



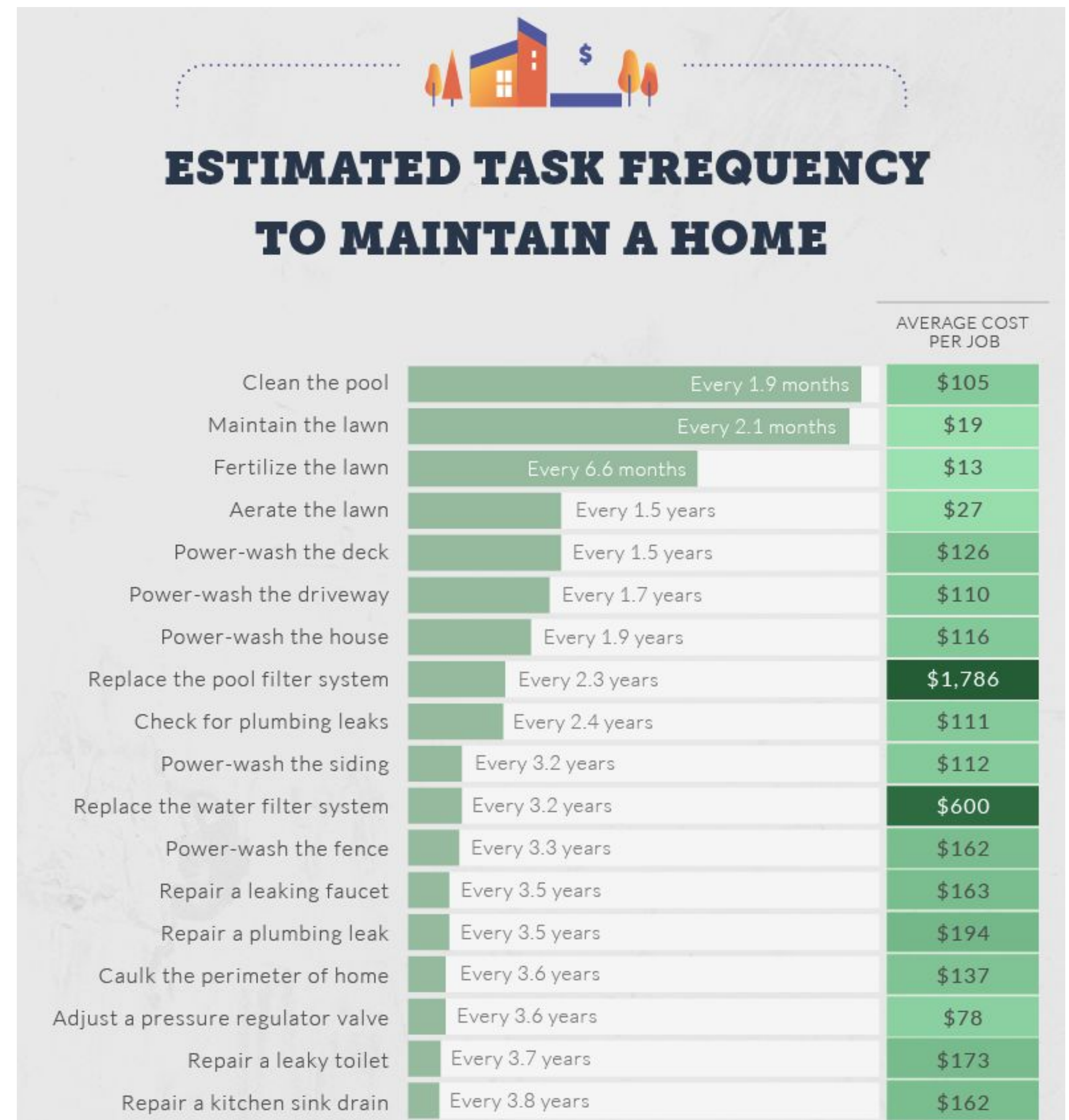
Overall Takeaway:

For **top of the funnel** activities, focus on **broadly interesting, data-driven stories**. Especially those with **local/regional angles**.

MOFU – Data Content Marketing for the Middle of the Funnel

Porch.com – True Cost of Home Maintenance

- **Data Source 1:** Aggregation and analysis of existing resources Porch previously created that estimated project type costs
- **Url:** porch.com/resource/cost-of-home-maintenance-in-america
- **Top-Tier Press:** WashingtonPost.com, SFGate.com, RD.com
- **Unique Linking Domains Earned:** 58



The Pitch That Secured Coverage

- **Personal Connection:** Opened with a comment about a previous article the journalist had written, tied in with personal anecdote
- **Hooks:** Local/regional specific data around cost of home maintenance related to journalist's beat
- **Pitches to Secure Exclusive:** 42



Hi Radhika,

Interesting article about home sellers being hesitant to trade up! My parents have lived in the same house for 25 years... despite being in a better financial position than ever, they don't think they could afford anything better in the current market.

Purchasing a home is already a daunting task for most people—but what about the additional costs that come after you close? Porch recently analyzed data from ZIP codes in each state to determine where the highest and lowest home maintenance costs are, as well as what that money is going towards.

Some findings from the [study](#):

- Residents of New Jersey, DC, Connecticut, Maryland, Massachusetts, and Rhode Island can expect to pay **over \$17,000** each year to maintain their homes
- The most common upkeep task? **Pool maintenance**. Paying for cleanings roughly every two months at an average of \$105 per job is no small commitment
- The average yearly cost to maintain a home in the United States **ranges from \$14,890** in Mississippi to **\$18,151** in New Jersey

Home expenses don't stop once a buyer enters escrow, and the costs of home maintenance should always be front-of-mind for any potential homeowner. Thoughts on sharing this research with your readers at The MReport?

Best,

Example 1 Takeaway: Your company or client likely has relevant and valuable internal data worth exploring.



In this case, Porch had already done great work using their own internal data from hired pro costs to make estimations of how much X-type project should cost in different locations. By aggregating this data, we were able to tell a larger story about project type cost trends.

- Common types of worthwhile internal company data
 - Customer data (e.g., purchases, site/app use behavior, internal surveys, demographics)
 - Industry-level data inferred from customer data
 - Temporal trends (product, customer, industry)

Joblist.com – Where It Pays to Live

- **Data Source 1:** MIT living wage calculator
- **Data Source 2:** 2017 Bureau of Economic Analysis Regional Price Parity Study
- **Url:** assets.joblist.com/where-it-pays-to-live-map
- **Top-Tier Press:** DenverPost.com, TheLadders.com
- **Unique Linking Domains Earned:** 22

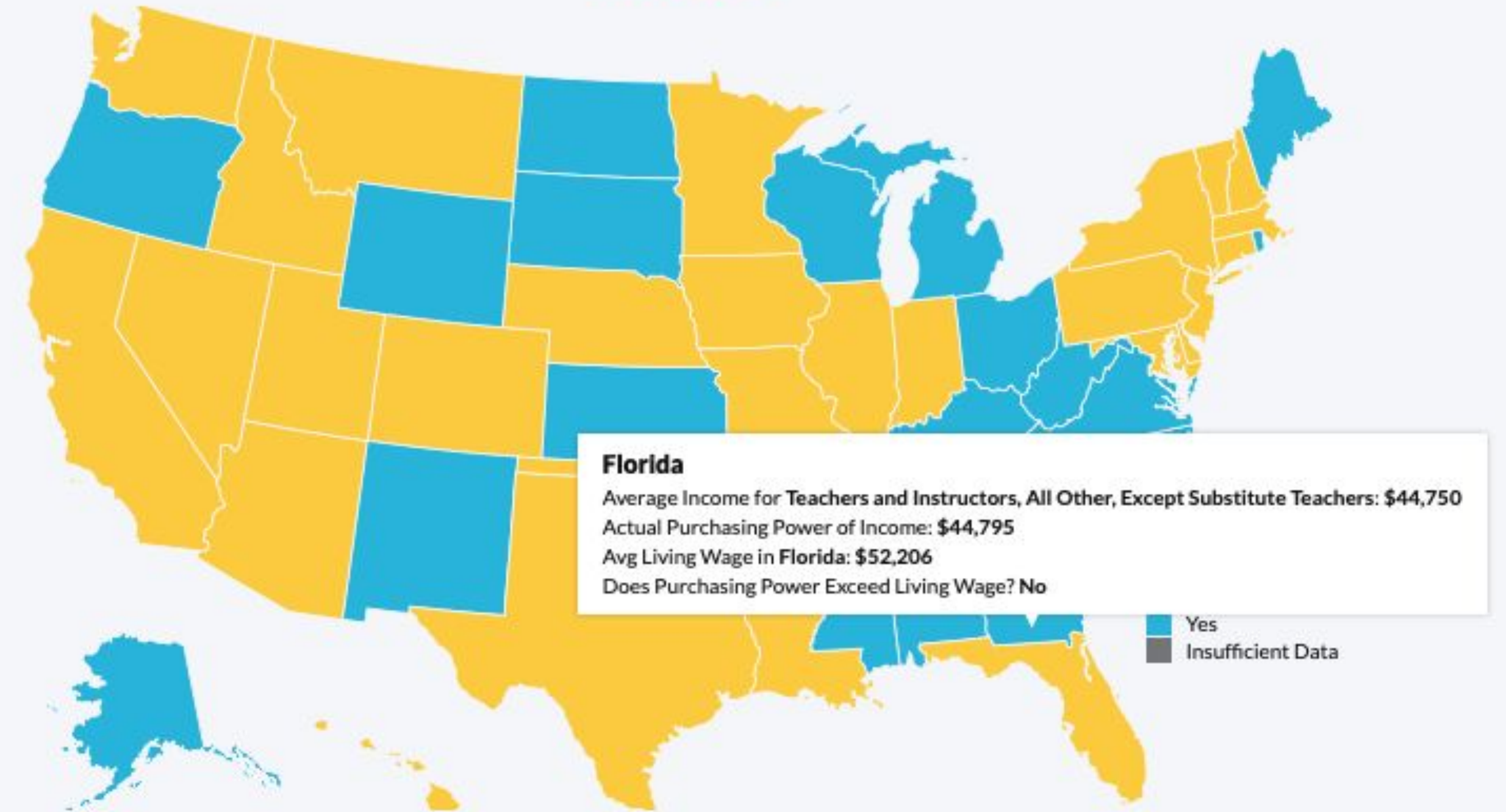
Does the average wage for your occupation cover the costs in your state?

Comparing the purchase power of average wage to living wage, by state and occupation

Teachers and Instructors, All Other, Except Substitute Teachers ▾

Nationwide Average Income

\$48,272



The Pitch That Secured Coverage

- **Personal Connection:** Demonstrated understanding of journalist's beat location (Colorado) and focus (consumer finance)
- **Hooks:** Journalist's location was a superlative (most underpaid state), according to our results.
- **Pitches to Secure Exclusive:** 67

Hi Aldo,

It's a great sign that Colorado's unemployment has dropped and hourly pay still has gone up. Even the average work week has gotten shorter, which seems to be an uncommon trend lately.

While Colorado is improving in many areas, the state still ranks high for missed overtime pay, which can lead to frustration and burnout among employees.

I wanted to share [a study](#) my team put together using the most recent U.S. Current Population Survey. We surveyed Americans to find out the amount of wages lost by full-time, salaried workers compared with cost of living. This allowed us to determine: **in what states can get the most bang for your buck?**

Some highlights:

Colorado is the most underpaid state when it comes to extra hours worked, followed by Maryland and D.C. Michigan, Wyoming and Oklahoma are the most affordable states when comparing wage and cost of living American workers are shortchanged a total of \$346 billion in wages per year due to unpaid overtime Employee burnout is becoming a prevalent problem, consistently in the headlines, so I thought this may be of interest.

Let me know if you'd be interested in sharing this with your readers first at The Denver Post, and feel free to let me know if you need any additional info.

Best,

Example 2 Takeaway: Utilize multiple datasets for deeper explorations.

Using MIT's living wage calculator data, combined with government data on average paid wages by state. This gave us the ability to understand locations with the largest disparities between average earnings and cost of living.

- Considerations
 - By combining multiple datasets, you are creating a project unlikely to have been done previously.
 - Location/regional based analysis are frequently relatively easy to do, especially when leveraging two or more government datasets (which typically have state or even more granular data).
 - Methodological accuracy is important, and these types of analyses are typically more complex.



ADT.com – Crime Maps Interactive

- **Data Source 1:** ADT-licensed local crime-level data
- **Url:** adt.com/crime
- **Top-Tier Press:** ABCActionNews.com, DigitalTrends.com
- **Unique Linking Domains Earned:** 301 (more than double since the last look roughly a month ago that we [featured on Moz.](#))

The screenshot displays the ADT Crime Map interface. At the top, there is a search bar with the placeholder text "Search crime maps by zip code or start typing an address...". Below the search bar is a map of the United States. A "Crime Map Key" is overlaid on the map, showing four categories: "Well below avg. (< 0.5x)" in green, "Slightly below avg. (0.5x - 1x)" in light orange, "Slightly above avg. (1x - 2)" in light red, and "Well above avg. (> 2x)" in dark red. The map shows various states colored according to these categories. Below the map, there are filter buttons for "TOTAL CRIME", "ASSAULT", "MOTOR VEHICLE THEFT", "BURGLARY", "LARCENY", "MURDER", "PERSONAL CRIME", "PROPERTY CRIME", "RAPE", and "ROBBERY". The interface also includes a "Map" and "Satellite" toggle, a "STATE" and "COUNTY" toggle, and a "Click on the map to get localized crime data" callout.

The Pitch That Secured Coverage

- **Personal Connection:** Opens by talking about tips given by journalist in previous article related to local crime
- **Hooks:** Local neighborhood is the safest in beat location/region; location within region with highest crime rate; location where specific crimes (robberies) are most likely; a note specific to writer's beat
- **Pitches to Secure Exclusive:** 10

Hi Nancy,

Moving to a new place can be an overwhelming experience. Thank you for the tips you shared on how to choose an apartment in an unfamiliar neighborhood. I will keep the apps you mentioned in mind the next time I move. Thanks again!

One major thing to consider when moving to a new neighborhood is the crime rate. Have you ever wondered how much and what kind of crime is happening in your area? Whether you're looking for a new apartment to rent or a permanent home to raise your family, ADT's new interactive crime map can help. We developed a map that allows users to look up zip code specific data and compare it to national averages to determine how much crime is happening.

Some of our findings::

- Arlington is one the safest neighborhoods in Washington D.C, the overall crime rate is 0.60x less than the national average
- Hill East (zip 20003) has one of the highest crime rates in the city - the murder rate is 7.63x more than average
- Robberies are more likely to happen near Takoma, where the average is 9.84x higher than the national index

You can try it yourself [here](#)

There are many things to consider before you move into a new neighborhood or visit a new city. ADT wants citizens to be aware of crime hot spots, in addition to being able to take steps to avoid being victims of crime. Your readers will find this tool useful and insightful. Any interest in sharing with them for The Washington Post?

Best,

Example 3 Takeaway: Make data interactive to enhance mid-funnel value and extend link building viability.



In this case, ADT had licensed granular crime data from a third party. Fractl explored this data and recommended externalizing it via an interactive crime map.

- Advantages to Interactives for mid-funnel content marketing
 - Custom, highly relevant data for potential customers
 - Logical to gate such tools (they provide value), giving an opportunity for lead acquisition
 - Tools of unique utility, or tools tied to authoritative brands, are inherently newsworthy if they “make sense for the brand to have created it for their potential customers.”

Overall Takeaway:



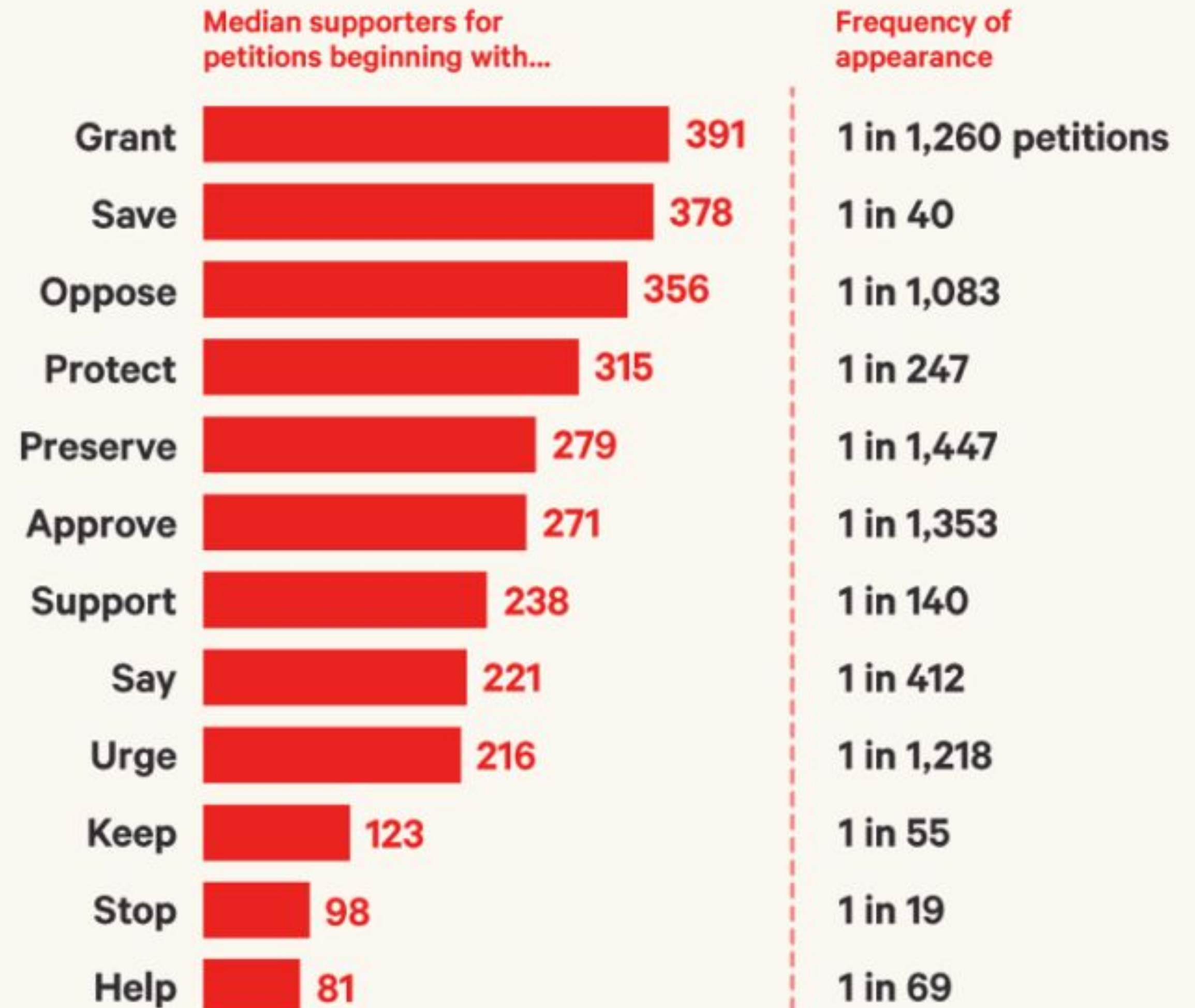
For **middle of the funnel** activities, focus on content of significant value to **existing or prospective customers**. The content should be able to **build links as a resource**, as well as **generate interest from niche/industry publishers** (and, occasionally, also mainstream press).

BOFU – Data Content Marketing for the Bottom of the Funnel

Change.org – Crafting Headlines for Change

- **Data Source 1:** ADT-licensed local crime-level data
- **Url:** change.org/c/us/crafting-headlines-for-change
- **Top-Tier Press:** Fastcompany.com, Justgiving.com
- **Unique Linking Domains Earned:** 15

The most effective first words express support and protection



The Pitch That Secured Coverage

- **Personal Connection:** Opens by praising similar coverage by the journalist being pitched and connects it specifically to the fascinating results of our work for Change.org
- **Hooks:** Specific, data-driven advice on how to craft headlines to improve charitable giving; 2.5X performance with these suggestions
- **Pitches to Secure Exclusive:** 20

Hi Ben,

Really enjoyed reading your coverage on the Young Scientist Challenge-- it was exciting to see the four finalists focused on solving environmental issues. It's crazy to think these ideas are coming from such young minds, definitely gives us hope for the future.

Like those middle-schoolers, individuals around the country are creating innovations to solve issues or improve lives via online petitions, some of them spurring social movements. But which petitions gain the most support? My team at [Change.org](#) released this new exclusive [report](#); the first of its kind, we analyzed over 164,000 petitions on our social activism platform from the past 6 months to better understand what makes a petition become a movement.

Drawing on the largest dataset of online activism in the world, our report offers insights into what language drives others to action:

Conventional internet wisdom says to be short and punchy, yet we found people who wrote longer titles with key details in them actually gained more support, on average. The very first word in a title is crucial. In some cases, it was associated with 2.5x performance in regards to gaining petition supporters. Optimistic tones performed best. While words like 'stop' or 'ban' are much more common, petitions with the most engagement employ keywords like 'grant,' 'save,' or 'protect'. Ultimately, by pulling insights from our own internal data, our report can help empower individuals to write powerful petitions that gain a large volume of support. Given that your coverage focuses on social impact and the future of philanthropy, I thought this would be the perfect fit for you. Would you be interested in being the first to share our report with your readers at Fast Company?

Best,

Example 1 Takeaway:

Leveraging AI and statistics, we did a deep dive on the headlines that correlate with “success” on Change.org (higher numbers of petition signatures). Using their internal data, we found:

- Headline words that correlate with higher performance
- Title length recommendations for better results
- Recommendations on including People/Entities for improved performance

Lower Funnel Value Realized:

- Utilized during petition creation flow (increased avg. petition signature number generally)
- Utilized in drip campaign emails for continued advice for petition creators
- Covered in the press by FastCompany and others. (unlikely for bottom of the funnel content, but in this case it worked!)



Overall Takeaway:



For **bottom of the funnel** activities, focus on content of significant value to **existing customers or warm leads**. Utilize value-add content and **use gated assets** to capture customer lead information for remarketing later. This content should also have value as a resource link builder.

Bonus: Key Tools of the Trade

- **Buzzstream.com** – As an outreach CRM, for managing relationships with publishers and journalists and tracking productivity of outreach team.
- **Python:** For data acquisition, cleaning, analysis, and visualization
- **Tableau:** For exploration of data and rapid visualization prototyping
- **Buzzsumo.com:** For ideation fodder, trend spotting
- **Fractl Campground:** A deep research tool internal to Fractl that scrapes all digital news into a graph database. The search capabilities of a graph database are a wonder to behold. We enhance the data with AI/ML modeling for topics, entities, and much more. (This is our outreach secret weapon and we hope to open to Beta testers soon,. Let me know if you are interested!)
- **Microworker platforms:** Recruiting participants for data acquisition. Amazon Mechanical Turk is our most used
- **Facebook ads:** For recruiting specialized survey participants (Doctors, Lawyers, Topic Experts, etc.)
- **SurveyMonkey:** Best surveying platform we've tried

Thank you! Questions?