A Cheat Sheet for Digital Media Analysts
Your Digital Analytics To-Do List

Digital analytics has quickly become the backbone of successful strategy at media companies. While this means there’s a huge opportunity for exciting discoveries through data, it also means that there’s more pressure than ever to figure out how to sort through the data to get to the analysis that’s going to make a difference. After all, working with analytics rarely comes with a clear TO-DO list.

On your TO-DO list

- Compulsively check Occam’s Razor
- Perfect your excel skills
- Make sure you have the right data inputs
- Give different departments insights they can use to make better decisions
- Prove that solid data analysis will make life easier for everyone at the company

How can you check off some of those items faster?

Parse.ly has put together this cheatsheet because we often work at the intersection of analysts and the teams they support: editorial, business, product and executives, all wanting answers to questions about how data can help them create a better media company.
Use this guide as your cheat sheet

This guide helps in two ways:

→ It lists tips for getting other teams on board with using data on their own.
  
  The better other people understand and use data, the more time you’ll have to work on the high level, game changing stuff — not the report pulling or dashboard set up.

→ It suggests projects that will provide you with results to share with your editorial, business and product teams that gives them information they care about. These projects are based on our own experiences working for media companies and best practices we’ve seen working with analysts and digital teams at hundreds of media clients.
DON'T JUST SHARE NUMBERS:

Translate metrics into what they mean about the readers

You use data and analytics to hear, understand and translate reader behavior and you know that analytics can tell a story about your audience, but is your editorial team taking the numbers too literally?

If other departments only see data as quota they have to hit, or worse, as a judgment of the work they’re doing, it’s harder to use analytical insight to inspire any change. How can you make sure everyone is on the same page about wanting to use data?

Help them think about what metrics mean in terms of their audience. This figure shows some common ways that media organizations can think about what the metrics are saying about the audience.

Not sure these metrics mean the same thing in your company?

Make sure that you and anyone looking at these metrics are on the same page about what they mean and why they are important to the organization.

Using easy to understand dashboards, like the ones in Parse.ly, helps you sit down with anyone at the company and find out what they think each metric represents and what goals they have. Your conclusions could be a blueprint that the whole company can work off of!
Democratize data! Make it a conversation in your organization.

Don’t be that person sending out Excel files, trust us, no one is reading them - or more importantly, doing anything with the information in them. Instead, find ways that you can make data the starting point for conversations.

Hold office hours or walk around the newsroom and talk to people about what they would like to know more about. You may find people are interested in questions that analytics could answer, but don’t even realize it!

Host an internal blog or send a weekly newsletter that highlights digestible updates with explainers. If you can tie in analytics to everyday discussions about goals or plans, people will start getting used to incorporating data into their workflow.

Look at any tools or systems other teams use to gather data. Are they synced with the work you’re doing, or are they making things more confusing? Make sure that there isn’t duplicated work happening, or worse, two conflicting measurement methods.

This takes investment on your part, but it’s hardly an unselfish act. By getting to know what your colleagues need, you can get better ideas about the kinds of insights that will be helpful for the company as a whole — making your work more impactful. And you can stop sending around that xls file once a week that no one looks at!

Organizing and sharing reports is the best way to distribute data to a group, but can be cumbersome and a time suck in most analytics platform. **Anyone can set up or distribute a report in Parse.ly in a few clicks and share with the people that can act on the data.**
Agree on goals that can be measured and get access to the data you need to track them

The most effective analysts all have two things:

→ They know what the goals are that they are measuring.

→ They know that they have the right data to measure those goals.

From the exercises on the previous two pages, hopefully you’ve been able to make progress on the first piece: knowing what goals need to be measured. Now, you need to make sure that everyone can access the data in a way that makes sense. For sites that have content as their backbone, the data you need to measure should make it easy to count, aggregate and sort through the metrics like those below. Run through a checklist to see how easy it is for you or your team to access the following:

1. Which section gets the most clickbait?
2. What do people share vs. what they like to read?
3. Headline keyword analysis using baseball sabermetrics
4. Does our tagging structure make sense?
5. How can data help our native advertising efforts?

Speak the same language!

The same data that you can provide to editorial teams can power your product team through the Parse.ly API. Create and monitor new types of article recommendations, using contextual, personalized, most-read, most shared and more inputs.
1. Which section gets the most clickbait?

Why you want to know:

A common source of consternation about bringing analytics into a newsroom or media team is chance it could encourage editorial clickbait. By finding and alerting teams to possible clickbait, you’re proving that data can help fight it if used correctly.

For this project, we’ll want to look at the pageview : engaged time ratio. The higher the ratio, the higher the possibility that content could be losing readers and their interest in your site.

What is clickbait?

Clickbait is generally defined as an article that seemed promising, but fails to deliver on the content.
How to calculate:

In Parse.ly, you can skip the formula and pull the data by selecting the date range under the Historical View and sorting by Average Engaged Time. Otherwise, create a spreadsheet with your average engaged time results.

First, sort the average minutes into buckets of time. Depending on your data set, this could be grouped by percentiles into “low, medium, high” or groups of specific minutes (i.e. <1, 1-<3, 3-<5, 5+).

Next, create a pivot table to view averages, sums or other counts and compare authors, sections, or any other grouping that you’ve exported.

You can also run this analysis and only look at new or returning visitors, for a further breakdown of whether or not clickbait may be the root cause of any issues. If large numbers of new visitors click away quickly, perhaps some of the distribution tactics your audience development team uses could use reviewing.

One thing to keep in mind with this analysis is that there may be a good reason for sections or authors to have a high page view : engaged time ratio — perhaps the content on that page is meant to be consumed quickly. Always spot check the results for an analysis like this!
2. What do people share vs. what do they read?

Why you want to know:

Editors and authors are often concerned with the number of shares their articles receive, as distribution on social platforms plays a huge role in helping readers find stories. However, not all stories and topics get shared the same way, and comparing them just on the share count can be comparing apples to oranges.

Understanding the differences in how audiences share certain topics versus their consumption habits of those articles can help set better expectations for social teams and provide insight into the audiences for each segment.

What you need:

- Shares per article across popular social networks
- Reading metric per article: views or visitors
- Dimensions suggested: author, section, tag
- Time period: One month of data

CONTINUED →
How to calculate:

Using the same steps as in our last study, sort the time frame you’re referencing by Shares. Use a pivot table to show each segment by sum of shares and sum of views. Take the ratio of the two (shares : views) and compare by segment.

High ratios mean that many of the people viewing the article decided to share it. These sections are prime for additional social distribution efforts. For segments that show a lower ratio, perhaps give the audience alternatives to social sharing, like an SMS or WhatsApp private messaging option.

On average, the shares ratio on the Parse.ly blog is .10. Though our “Tech” posts get a lot of views and visitors, they don’t get very many shares — way below the average at .02. For these posts, we may consider deleting our shares options and instead encourage links to communities that the audience is more active on, which in this case is Hacker News or Reddit.

On the other end of the spectrum, we see lots of shares for our features posts, and the views are decently high. We’ve spent some effort to improve the ways we encourage sharing in our featured articles, so this is a good sign!

Want to see how many referrers are brought to the site by each share on an article level? There’s a view in Parse.ly that allows for just that. A high shares to referrers ratio indicates that the people that shared the story got lots of people to read it. A low shares to referrers ratio indicates that though people may share the story, it’s not attracting new readers.
3. Headline Keyword Analysis Using Baseball Sabermetrics

Adapted from Adam Felder’s *Ratio Metrics*

Why you want to know:

Reporters are always looking for advice when it comes to writing better headlines: what words get their readers attention?

What you need:

- **Full text for all headlines** in data set,
- **KPIs per article** (visitors, page views, etc.)
- **Time period**: Depends on volume (expand for less volume, decrease for more)

How to calculate:

1/3 — **Determine Your Success Thresholds**: Separate content with norm-referenced criteria by selected KPI to separate into quantiles of success. In this example, we’re identifying all content that performed better than 50th percentile as 1b, 75th = 2b, 87.5th = 3b, and 95th = HR.

Classify Your Content Into Hit Types:

<table>
<thead>
<tr>
<th>Title</th>
<th>PubDate</th>
<th>1b</th>
<th>2b</th>
<th>3b</th>
<th>HR</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>To Beat Terry McAuliffe, Virginia GOP Turns to Anti-Romney Playbook</td>
<td>4/11/2013</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ken Cuccinelli Struggling with Makeover in Virginia Governor's Race</td>
<td>4/12/2013</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Immigration Reform Isn’t Hurting Marco Rubio’s Bottom Line</td>
<td>4/14/2013</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Focus on Sun Control and Immigration Reform Diverted After Boston Bombings</td>
<td>4/23/2013</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Democrats Argue Over Who’s Tougher on National Security After Boston Bombings</td>
<td>4/26/2013</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Special Interests Shadow Immigration Reforms</td>
<td>4/26/2013</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Black Voters Are Key to a Colbert Busch Win in South Carolina</td>
<td>4/28/2013</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**CONTINUED →**
2/3 — **Find frequently used keywords:** Use a frequency distribution to find words most used in headline sample set. Flag each piece of content (from step 1) as containing 1 or more keywords. Aggregate keywords by the number of pieces each is used in.

![Frequency Distribution Table]

3/3 — **Assign Baseball Statistics for Analysis:** Using the existing chart, start to dig into keywords and how they do with different aggregations of statistics.

- **Batting Average** \((H/\# \text{ of Posts})\): See which keywords get the most consistent “successes” regardless of magnitude.

- **Slugging Percentage** \(((4\times HR)+(3\times 3b)+(2\times 2b)+(1\times 1b))/\# \text{ of Posts})\): Instead of weighing each type of success equally, it counts them according to where they fell on your success spectrum. Do certain keywords get the same amount of average success to one tends to outperform when it comes to magnitude?

- **Average Plus Slugging** \((AVG+SLG)\): Content that not only frequently succeeds, but gets enormous amount of traffic when it does so.

- **Isolated Power** \((SLG-AVG)\): These keywords may not always succeed, but when they do succeed, they tend to be above average.

![Baseball Statistics Table]

See more about how *Atlantic Media* applied this analysis to their headlines in this blog post.
4. Does our tagging structure make sense?

Why you want to know:

A number of questions that analysts could ask to create strategies within their organizations rely on available data provided by teams through the form of “tags.” However, our research shows that most companies do not have a tagging structure in place, other than instructing writers and editors to identify keywords associated with the articles.

A systematic tagging strategy can open up the possibility of a great deal of analysis. Here we outline how you can check if you need an update.

What you need:

- All tags used for posts published in the specified time frame
- Number of posts per tag
- Optional: Metrics per tag
- Time period: One month

“I think tags are being used well at my organization.”

58.3%
How to calculate:

First, look at the total amount of unique tags used during the time frame. Group by post frequency to determine average and median amount of tag usage per post. It may be helpful to do this by section or author to see differences.

The quickest analysis you can do for this is to check how many tags are only used once. Are they a small or large percentage of all tags?

Group tags in segments that make sense for your data set, our recommendation is to normalize by quartiles (25% of tags used the most, 25% of tags next used the most, etc.)

Finally, see if you can work with your editorial team on setting up some guidelines or tagging hierarchies. A good way to do this is to identify people or groups that need data and see how tagging could apply to their questions. Once that information is collected, it’s worth a discussion with the product team about ways that could help automating tagging.

Are there obvious themes that could be grouped together, and could anything automate the process? i.e. if you are seeing Obama and Barack Obama all as separate tags, how could the process be improved so that all those tags become one?

Are there separations between content or keyword tags and type tags? For example, Obama would be a content tag on the same article that slideshow is used as a type tag. Again, could automation be used here? If a video is added to an article, perhaps the CMS could add a video tag, taking the responsibility off the editors’ shoulders and ensuring better data?

Are there any tags that are seeing lots of use and audience reaction that aren’t part of a typical coverage section? This is a good opportunity to use tags as a way to find new sections, verticals, or beats to cover.

Tagging hierarchies typically used by content providers include: keyword tags, sentiment tags, template tags, type tags, campaign tags.
5.

How can data help our native advertising efforts?

Why you want to know:

More media companies rely on a content studio or branded content division as part of their revenue stream. Data analysis can be used to assist these efforts by making it easier to identify advertisers, prove ROI to clients, and more.

This project works best after running the tagging structure analysis, but pulling headline keywords using the method in project three can be used as a proxy.

What you need:

- Tags with associated metrics
- List of possible advertisers that your ad sales team wants to target

CONTINUED →
FIVE PROJECT IDEAS FOR MEDIA ANALYSTS

How to calculate:

Provide the list of top keywords to your sales team, and ask them to match keywords to specific brands, like so:

<table>
<thead>
<tr>
<th>Tag</th>
<th>Sort (Visitors)</th>
<th>Posts</th>
<th>Brands</th>
</tr>
</thead>
<tbody>
<tr>
<td>analytics</td>
<td>52600</td>
<td>IBM, HP, Netflix</td>
<td>24 Microsoft, Data Solutions International, Aponia Data, Sator Software</td>
</tr>
<tr>
<td>data</td>
<td>35400</td>
<td>14</td>
<td>IBM, GoPro, Apple, Microsoft</td>
</tr>
<tr>
<td>technology</td>
<td>30700</td>
<td>11</td>
<td>IBM, GoPro, Apple, Microsoft</td>
</tr>
<tr>
<td>social media</td>
<td>17200</td>
<td>10</td>
<td>iContact, Marketo, PrimeVisibility.com</td>
</tr>
<tr>
<td>innovation</td>
<td>32100</td>
<td>7</td>
<td>Samsung, PWC, Lexus</td>
</tr>
</tbody>
</table>

For every keyword or tag they associate with a potential brand, aggregate the associated metrics with all posts that use the tag. Provide this back to the sales team to use as ammunition in sales decks and pitches, and win “hero for life” award immediately!

Make sure that sales teams can access the data of campaigns that are running to provide results to the client as well. Have them use a tag or special section to evaluate overall performance.

**Performance Summary for tag “type: analytics”**

<table>
<thead>
<tr>
<th>metric</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page Views</td>
<td>87k</td>
</tr>
<tr>
<td>Visitors</td>
<td>26k</td>
</tr>
<tr>
<td>Avg. Time</td>
<td>0:56</td>
</tr>
<tr>
<td>New Posts</td>
<td>4</td>
</tr>
</tbody>
</table>

**Campaign:** 1.7% of traffic  **Medium:** 4.0% of traffic  **Source:** 3.6% of traffic

**A CHEAT SHEET FOR DIGITAL MEDIA ANALYSTS**

- A digital analytics To-Do List
- Using the analyst’s cheat sheet
- Creating a data-driven mindset
- Translate metrics
- Make data a conversation
- Get everyone the right data

**FIVE PROJECT IDEAS FOR MEDIA ANALYSTS**

1. Which section gets the most clickbait?
2. What do people share vs. what they like to read?
3. Headline keyword analysis using baseball sabermetrics
4. Does our tagging structure make sense?
5. How can data help our native advertising efforts?

**Bonus Tip**

- Translate metrics
- Make data a conversation
- Get everyone the right data

**CREATING A DATA-DRIVEN MINDSET**

- Translate metrics
- Make data a conversation
- Get everyone the right data

**FIVE PROJECT IDEAS FOR MEDIA ANALYSTS**

1. Which section gets the most clickbait?
2. What do people share vs. what they like to read?
3. Headline keyword analysis using baseball sabermetrics
4. Does our tagging structure make sense?
5. How can data help our native advertising efforts?

**Bonus Tip**

- Translate metrics
- Make data a conversation
- Get everyone the right data

**CREATING A DATA-DRIVEN MINDSET**

- Translate metrics
- Make data a conversation
- Get everyone the right data
Bonus tip!

All of these projects can be done in most analytics systems, but the structure and ease of getting the data out in the correct format will be the key between you providing great results to your team and you banging your head on a wall. Want to make sure it’s the former? Contact us for a demo today!

www.parsely.com/start