Data Journalism and Metrics

Data Visualization Performance from the Perspective of Web Analytics

Ilo Alexandre iNOVA Media Lab, ICNOVA Universidade NOVA de Lisboa Lisbon, Portugal mail@iloaguiar.com

ABSTRACT

Data journalism (DJ) is one of the hottest trends in journalism. The production of interactive pieces increases every year, as well as the number of studies in the field. Yet, there is still a lack of research on the performance of these contents. In this paper, we compare the performance of 90 DJ pieces with interactive visualization with (1) the most read news articles, and (2) with static news stories of the same topic published in the same period as the interactive ones. The DJ pieces are from three news organizations: The Texas Tribune (United States), O Globo (Brazil) and Público (Portugal). This analysis is based on access to Google Analytics, an online web analytics tool. Among the key findings of this research are: DJ pieces continue to be visited years after publication in a much higher proportion than other news articles on the same topic; users spend on average more time on interactive pieces than on the most read news articles and related news stories; the recirculation rate in DJ pieces is higher than in the related news and the most read news articles; the bounce rate is higher in the interactive pieces than in the most read news articles and the related news stories.

KEYWORDS

Data journalism, Analytics, Interactive visualization, Metrics

1 INTRODUCTION

Data journalism (DJ) has gained increasing popularity in large and small newsrooms around the world [1,13,28]. Often working with public data obtained through the right to access public information and/or with leaked databases, scholars recognize that DJ can reinforce transparency, accountability, and journalism's watchdog role through the scrutiny of those in power [3,11,21]. Another crucial characteristic of DJ is its cross-field hybridity and participatory openness [10]. Interactive data visualization encourages active user participation in the exploration of data, and many projects are tools to assist the user in making decisions [7,25,32].

The number of scholarly publications on the topic has increased dramatically in recent years [24,30]. However, there are research divisions within the field of DJ that have not yet been properly explored. For instance, there is an insufficient number of studies that address the performance of DJ pieces. Lately, it has gained more relevance to the notion of journalistic impact on individuals, groups, or institutions. *ProPublica*'s "Dollars for Docs" ¹ and the *International Consortium of Investigative Journalists*" "The Panama Papers" ² are DJ examples that had extensive repercussions and impact [16,27]. Nevertheless, most of the newsrooms that are investing and embracing impact measurement are nonprofit [20].

Most news organizations still rely on analytics data to determine the success of content. Scholars have systematically shown that pageviews is the most important metric in the newsrooms [2,6,17,31,33]. Unique pageviews, time on page, bounce rate, and recirculation are other important metrics [8,19].

With these metrics in mind, we selected DJ pieces from three news organizations from three different countries: *The Texas Tribune* (United States), *O Globo* (Brazil) and *Público* (Portugal). *The Texas Tribune* is a non-profit digital-born media, focused on public policy, politics and the government in the state of Texas. *O Globo* and *Público* are legacy media, general-interest newspapers with a daily print edition.

From the early stage, DJ pieces have been a key feature on *The Texas Tribune* and some are updated regularly and continue to get featured on *Tribune*'s homepage. In contrast, the production of interactive pieces in *O Globo* and *Público* is still intermittent and during the period studied were often secondary pieces, complements to the main story, rarely updated and with limited – and sometimes none – featured on the homepage.

Comparing web analytics data of DJ pieces with other news articles it is possible to realize which content drives more traffic to the website or which content engages more the audience. Therefore, this paper attempts to shed light on how DJ pieces perform in comparison with other news.

2 METHODS

Ten DJ pieces were selected from each news organization in 2014 in an arbitrary way. Another 10 from 2015 and another 10 from 2016, totaling 90. The benchmark used for the selection of data visualization was interactivity, one of the key elements in DJ [4,21,34]. From interactive slideshow to news apps [15,29].

¹ https://projects.propublica.org/docdollars/. Accessed on Dec. 12, 2019.

² https://www.icij.org/investigations/panama-papers/. Accessed on Dec. 12, 2019.

Computation + Journalism 2020, March 20-21, 2020, Boston, MA

Table 1: Description of the metrics used in this study [22].

Pageviews	The total number of times a page has been
	loaded in a browser.
Unique pageviews	The number of visits in which the page was
	viewed at least once.
Average time on page	The average amount of time a user spends on a specified page.
Bounce rate	Percentage of visits in which the user leaves the
	site from the same page she/he entered it.
Exit rate	Percentage of how often users exit the site from
	a specific page.

Each DJ piece was contrasted with two groups: (1) the 500 most read news articles of that year in the same media outlet; (2) two news pieces from the same mews organization, on the same topic, and published in the same period as the interactive story. We adopted these approaches because only by comparing numbers is it possible to understand the data, since isolated numbers are meaningless [5].

This approach allows, on the one hand, to evaluate the performance of DJ pieces in comparison with the general content of the website, since news organizations such as O *Globo* and *Público* have a wide range of content. On the other hand, the comparison with news stories on the same topic allows for a more targeted observation, theoretically with an audience interested in that subject.

Among the audience measurement tools used in *The Texas Tribune, O Globo* and *Público*, Google Analytics is used by all three. We had Google Analytics access to all three media outlets, which means that continuous data was used to compare the performance of the news pieces. Web analytics data are considered sensitive and strategic data by media organizations, having potential commercial implications [12,23]. Thus in the results, we used proportions rather than absolute numbers, but also because it is a way to give more context and meaning to the data [5].

Google Analytics has over 100 default reports at hand and it is possible to create practically infinite combinations of custom segments [9]. Unfortunately, segmentation and events tracking user behavior were not implemented in all three case studies in 2014 - the year of the oldest data collected. Therefore, we have chosen to use only some of the default metrics. We analyzed five metrics: pageviews, unique pageviews, average time on page, bounce rate, and exit rate. The description of the metrics is shown in Table 1.

We collected data from the piece's publication date until the last day of the respective year. We also collected pageviews number of DJ pieces and the news from the same topic from publication until December 31, 2017. Thus, it is possible to examine the lifetime of the pieces. Most of the online news is not visited after two days of having been published [18].

Finally, we do not compare pieces of one media outlet with a piece from another. News organizations have different objectives, approaches, and audiences, so such comparisons would not be helpful. However, the result of analysis in one media outlet in contrast to the result in other may indicate patterns or trends in the Table 2: Data journalism among the top 10 most read news.

News outlet	2014	2015	2016
The Texas Tribune	6	4	2
O Globo	0	0	1
Público	1	1	0

consumption of such news. These are the comparisons we present in the next section.

3 RESULTS

We collected data on each metric described in Table 1 for the DJ pieces, the 500 most read news article in 2014, 2015 and 2016 in *The Texas Tribune, O Globo* and *Público*, and 180 news pieces on the same topic, and published in the same period as the interactive story (two related news for each DJ piece).

3.1 Interactive pieces drive a lot of traffic to the website

The Texas Tribune's interactive pieces are among the most accessed content on the website: 16 of the 30 DJ pieces analyzed between 2014 and 2016 registered more pageviews than the average of the 500 most read news stories in the year of its publication. Only one interactive piece from O Globo and three from Público reported more pageviews than the average of 500 most read articles.

Nevertheless, that interactive piece from *O Globo* was the most accessed news article in 2016, with five times more pageviews than the runner-up. In *Público*, in 2014 and 2015 the news stories that drive the most traffic were DJ pieces. In *The Texas Tribune*, the top five most accessed articles in 2014 were DJ pieces. In 2015, the two pieces with most pageviews were DJ pieces. And in 2016, the most read news story was a DJ piece. Table 2 shows how many DJ pieces ranked in the top 10 most read news stories in each media outlet studied.

Contrasting pageviews of DJ pieces with their related news stories, the performance is quite diverse. Eleven out of 30 *Público* interactive pieces registered more pageviews than their related news articles in the year they were published. *O Globo*'s interactive pieces performed a litter better: 57% of the DJ pieces drive more traffic to the website than the related news. In *The Texas Tribune*, 83% of the data visualization reported more pageviews than the related news stories on the same topic.

3.2 A quarter of the data journalism pieces are accessed again and again

The ratio of pageviews/unique pageviews shows how many times a page was accessed during users' sessions, i.e. it is possible to check if a news article was viewed multiple times by the same user. Except for some photo galleries, where to see the next photo the user needs to refresh the page, the ratio pageviews/unique pageviews in news stories tend to be low. A ratio above 1.4 is high and indicates that users returned to the page within their session [14].

Table 3: Ratio pageviews/unique pageviews.

Year	News outlet	Data journalism pieces (avg.)	500 most read pieces (avg.)
2014	The Texas Tribune	1.4	1.2
	O Globo	1.2	1.1
	Público	3.9	2.9
2015	The Texas Tribune	1.4	1.2
	O Globo	1.2	1.1
	Público	1.3	4.4
2016	The Texas Tribune	1.4	1.3
	O Globo	1.4	1.2
	Público	1.4	3.6

From the 90 DJ pieces analyzed, 24 registered a ratio higher than 1.4. That is, in 27% of the interactive pieces, the proportion of users coming back to the visualization is meaningful. There are three from *The Texas Tribune*, another three from *Público* and one from *O Globo* that recorded a rate higher than 2.0. And one DJ piece from *Público* had an incredible rate of 17.5. Among the 180 related news articles, only three from *Público* recorded pageviews/unique pageviews proportion higher than 1.4.

On average, the 500 most read news articles of *The Texas Tribune* and *O Globo* do not present a pageviews/unique pageviews proportion higher than 1.4. On the other hand, the most accessed news stories in *Público*, on average, have a very high ratio, as shown in Table 3.

3.3 Long live the data journalism pieces

Only one of 180 related news articles had more pageviews in 2017 than in the year it was published. In contrast, 14 of the 90 data visualization analyzed drive more traffic in 2017 than in the year of its publication. This represents 15% of the DJ pieces.

The Texas Tribune has the largest number of interactive pieces that registered a growth in the pageviews over the years: 10 out of 30. In *Público* was just one, while in *O Globo* were five data visualization. One of these pieces from *O Globo*, published in 2014, reported in 2017 about eleven times more pageviews than in the year of its publication.

3.4 Users spend more time on interactive pieces

From the 90 DJ pieces analyzed, 55 registered more average time on the page than at least 1 of the related news stories. In 11 of these 55, users spent more than twice the time of those who accessed related articles. *O Globo*'s interactive pieces captured the most users' attention: 18 of the 30 data visualization reported more average time on page than their related news stories. One of these interactive pieces from *O Globo* registered an average time on page three times higher than its related news articles.

In comparison to the average of the 500 most read news stories, 24 of the 30 interactive pieces from *Público* had more time on page than the most read articles. In 15 of these 24, users spent more than twice the time of those who accessed the most read news pieces. And two of these data visualization registered an average time on page four times higher.



Figure 1: Average time on page. Most read news represented as 100%. Interactive pieces' results are relative to this value.

In *The Texas Tribune*, 17 DJ pieces reported more average time on the page than the average of the 500 most read news stories, while in *O Globo* were 14. In Figure 1 the average time on page of the 500 most read news pieces is represented as 100% and the average results of the DJ pieces in that year range accordingly. The interactive pieces from *Público* registered more average time on the page every year and in 2015 and 2016 had more than twice the time on page than the average of the most read news stories.

3.5 Interactive pieces drive more recirculation

From the 90 DJ pieces analyzed, 60 registered a lower exit ratio than at least one of the related news stories. That is, in two-thirds of the interactive pieces the user visited another page of the website in greater proportion than who accessed the related story. In *Público*, 21 of the 30 DJ pieces had a lower exit rate than at least one of the related stories, while in *O Globo* were 20 and *The Texas Tribune*, 19. Six DJ pieces had an exit rate lower than 10%, four from *The Texas Tribune* and two from *Público*.

This behavior of proceeding to another page from the website, engaging with more than one content piece is called recirculation – and keeping the user moving through the website is one of the main goals of news organizations nowadays [8]. Although some web analytics tools measure the recirculation rate by default, this is not the case with Google Analytics [26]. Hence, bear in mind that a low exit rate means a high recirculation rate.

53% of the DJ registered a lower exit rate than the average of the 500 most read news stories in the year they were published. The percentage of users accessing another news article after visiting one of the *O Globo* interactive pieces is 80% higher than those who accessed the most read articles of this newspaper. In *The Texas Tribune* the proportion is 43% higher and in *Público*, 36%.

In Figure 2 the average exit rate of the 500 most read news pieces is represented as 100% and the average exit rate of the data visualization in that year range accordingly. The interactive pieces from *Público* in 2015 and 2016 are the only ones that registered a higher exit rate than the average of the most read news stories.

Computation + Journalism 2020, March 20-21, 2020, Boston, MA



Figure 2: Exit rate. Most read news represented as 100%. Interactive pieces' results are relative to this value.

3.6 Users are more likely to bounce in data journalism pieces

The bounce rate considers the number of users who entered the website through that page and left the site without visiting another content piece. About 50% of users who visited DJ pieces from *The Texas Tribune* did not enter the site through this content. The rate rises to around 75% on *O Globo* and *Público*. That is, only a fraction of visitors of DJ pieces are considered in this metric.

From the 90 data visualization, 31 reported a lower bounce rate than their related news articles. In *The Texas Tribune*, 56% of the DJ pieces had a lower bounce rate than the related stories, while in *O Globo* were 33% and *Público*, 13%. Two DJ pieces from *The Texas Tribune* registered an extremely low bounce rate (fewer than 10%).

In comparison to the average of the 500 most read news stories, 10 of the 30 interactive pieces from *The Texas Tribune* registered a lower bounce. *O Globo* reported 9 and *Público* had three data visualization with a proportion inferior to the average of the 500 most read articles in the year they were published.

In Figure 3, the average bounce rate of the 500 most read news pieces is represented as 100% and the average bounce rate of the DJ pieces in that year range accordingly. The interactive pieces from *The Texas Tribune* are the only ones in all the years analyzed where the bounce rate is equal to or below the average of the most read news stories.

4 CONCLUSION AND FUTURE WORK

By using Google Analytics data to measure the performance of DJ pieces in news organizations from three different countries – the United States, Brazil, and Portugal – we performed unprecedented work. We compared the performance of interactive pieces with the most read news stories from each media outlet so that we could have an overview of the interest in data visualization in comparison to the general content accessed in its respective websites. Each DJ piece was also contrasted to two similar news articles on the same subject and published in the



Figure 3: Bounce rate. Most read news represented as 100%. Interactive pieces' results are relative to this value.

same period. In this manner, it was possible to measure the performance of interactive pieces with contents that could also interest the visitor of the interactive pieces and vice versa.

Taking as case studies a non-profit digital-born media and a couple of legacy media, it was possible to observe the performance of data visualization in news organizations with different goals and from different countries. *The Texas Tribune* is focused on public policy, politics, and the government in the state of Texas, therefore the variation of topics is not so great, unlike the legacy media, which are general-interest newspapers.

Our findings show that *The Texas Tribune* interactive pieces have more pageviews than their related news stories and reported also a better performance than the average of the 500 most read news articles in the year they were published. Data visualization from *O Globo* and *Público* did not achieve the same performance facing the average of its 500 most read pieces. Nevertheless, it should be recognized that while *The Texas Tribune*'s DJ pieces were widely publicized, the interactive ones from the legacy media were commonly complementary to other news stories and did not receive the same kind of prominence on their websites.

From the 90 DJ pieces analyzed, 27% recorded a high proportion of pageviews divided by the unique pageviews, i.e. in these data visualization the users returned to the page several times in the same session. And 21% of the interactive pieces registered growth in traffic over the years, while fewer than 1% of related news articles recorded more pageviews in 2017.

This group of results considering pageviews indicates that data visualization has performed well when they are published and continue to be visited over the years even without any updates or highlights on the homepage, as in *O Globo* and *Público*.

Examining time on page, exit rate and bounce rate metrics, our findings show that users typically (1) visited a page from the website before entering an interactive piece; (2) spent time above average in the data visualization; (3) and then moved on to another page from the website. This behavior of visiting a page of content and then proceeding to another piece of content is known as recirculation and has gained increasing importance in news

I. Alexandre

Data Journalism and Metrics

organizations [8]. The recirculation and higher average time on page occur in the DJ pieces in a higher proportion than in the average of the most read news and in the related articles. Although the interactive pieces registered a low rate of entrances, those who landed on the website in an interactive piece mostly left the website without visiting another page from the news organization.

We believe that this body of findings brings contributions to the study in the fields of DJ and audience data and metrics in online journalism. There is very little research that uses web analytics data and even less that uses it to investigate data visualization performance.

The production of DJ pieces in newsrooms has grown considerably in the last years and tends to continue expanding. Investigations that focus and combine web analytics data with usability testing to understand how the user interacts with the data visualization are critical. Journalists will only be able to create more comprehensive and memorable interactive pieces if they know how these pieces are received and evaluated by the audience.

ACKNOWLEDGMENTS

This research was funded by Fundação para Ciência e Tecnologia with the scholarship number PD/BD/114147/2015.

REFERENCES

- Ilo Alexandre. 2016. Promoting Insight: A Case Study of How to Incorporate Interaction in Existing Data Visualizations. In 2016 20th International Conference Information Visualisation (IV), 203–208. DOI:https://doi.org/10.1109/IV.2016.15
- [2] C. W. Anderson. 2011. Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism: Theory, Practice & Criticism* 12, 5 (July 2011), 550–566. DOI:https://doi.org/10.1177/1464884911402451
- [3] C. W. Anderson, Emily J. Bell, and Clay Shirky. 2014. Post Industrial Journalism: Adapting to the Present. *Columbia University* (2014). DOI:https://doi.org/10.7916/d8n01js7
- [4] Ester Appelgren. 2018. An Illusion of Interactivity: The paternalistic side of data journalism. *Journalism Practice* 12, 3 (March 2018), 308–325. DOI:https://doi.org/10.1080/17512786.2017.1299032
- [5] Michael Beasley. 2013. Practical web analytics for user experience: how analytics can help you understand your users. Elsevier/Morgan Kaufmann, Amsterdam.
- [6] Nicole Blanchett Neheli. 2018. News by Numbers: The evolution of analytics in journalism. *Digital Journalism* 6, 8 (September 2018), 1041–1051. DOI:https://doi.org/10.1080/21670811.2018.1504626
- [7] Paul Bradshaw. 2014. Data journalism. In *Ethics for digital journalists: Emerging best practices*, Lawrie Zion and David Crag (eds.). Routledge, New York, 202–219.
- [8] Federica Cherubini and Rasmus Kleis Nielsen. 2016. Editorial Analytics: How News Media are Developing and Using Audience Data and Metrics. SSRN Electronic Journal (2016). DOI:https://doi.org/10.2139/ssrn.2739328
- [9] Brian Clifton. 2015. Successful analytics: gain business insights by managing Google Analytics.
- [10] Mark Coddington. 2015. Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism* 3, 3 (May 2015), 331–348. DOI:https://doi.org/10.1080/21670811.2014.976400
- [11] Sarah Cohen, James T. Hamilton, and Fred Turner. 2011. Computational journalism. *Communications of the ACM* 54, 10 (October 2011), 66. DOI:https://doi.org/10.1145/2001269.2001288
- [12] Nick Couldry and Joseph Turow. 2014. Advertising, big data and the clearance of the public realm: marketers' new approaches to the content subsidy. *International Journal of Communication* 8, (2014), 1710–1726.
- [13] Katherine Fink and C. W. Anderson. 2015. Data Journalism in the United States: Beyond the "usual suspects." *Journalism Studies* 16, 4 (July 2015), 467–481. DOI:https://doi.org/10.1080/1461670X.2014.939852

Computation + Journalism 2020, March 20-21, 2020, Boston, MA

- [14] Government Digital Service. 2016. Data and analytics Content design: planning, writing and managing content - Guidance - GOV.UK. Retrieved December 11, 2019 from https://www.gov.uk/guidance/content-design/dataand-analytics
- [15] Jonathan Gray, Liliana Bounegru, and Lucy Chambers (Eds.). 2012. The Data Journalism Handbook (First Edition ed.). O'Reilly & Associates, Sebastopol, CA.
- [16] Lindsay Green-Barber. 2019. Beyond Clicks and Shares: how and why to measure the impact of data journalism projects. In *The Data Journalism Handbook 2*, Jonathan Gray and Liliana Bounegru (eds.). European Journalism Centre.
- [17] Jonathan Groves and Carrie Brown. 2011. Stopping the Presses: A Longitudinal Case Study of the Christian Science Monitor Transition From Print Daily to Web Always. *International Symposium on Online Journalism* 1, 2 (2011), 86–128.
- [18] Benjamin Ho and Peng Liu. 2015. Herd journalism: Investment in novelty and popularity in markets for news. *Information Economics and Policy* 31, (June 2015), 33–46. DOI:https://doi.org/10.1016/j.infoecopol.2015.04.004
- [19] ICFJ. 2017. 2017 State of Technology in Global Newsrooms Survey. International Center for Journalists. Retrieved April 22, 2019 from https://www.icfj.org/sites/default/files/2018-04/ICFJTechSurveyFINAL.pdf
- [20] Magda Konieczna and Elia Powers. 2017. What can Nonprofit Journalists Actually do for Democracy? *Journalism Studies* 18, 12 (December 2017), 1542–1558. DOI:https://doi.org/10.1080/1461670X.2015.1134273
- [21] Wiebke Loosen, Julius Reimer, and Fenja De Silva-Schmidt. 2017. Datadriven reporting: An on-going (r)evolution? An analysis of projects nominated for the *Data Journalism Awards* 2013–2016. *Journalism* (October 2017), 146488491773569. DOI:https://doi.org/10.1177/1464884917735691
- [22] Alexa L Mokalis and Joel J Davis. 2018. Google Analytics demystified (4th Edition ed.). CreateSpace Independent Publishing Platform.
- [23] Philip M. Napoli. 2011. Audience evolution: new technologies and the transformation of media audiences. Columbia University Press, New York.
- [24] Adegboyega Ojo and Bahareh Heravi. 2018. Patterns in Award Winning Data Storytelling: Story Types, Enabling Tools and Competences. *Digital Journalism* 6, 6 (July 2018), 693–718. DOI:https://doi.org/10.1080/21670811.2017.1403291
- [25] Sylvain Parasie and Eric Dagiral. 2013. Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. New Media & Society 15, 6 (September 2013), 853–871. DOI:https://doi.org/10.1177/1461444812463345
- [26] Caitlin Petre. 2015. The Traffic Factories: Metrics at Chartbeat, Gawker Media, and The New York Times. *Columbia University* (2015). DOI:https://doi.org/10.7916/d80293w1
- [27] Fergus Pitt and Lindsay Green-Barber. 2017. The Case for Media Impact: A Case Study of ICIJ's Radical Collaboration Strategy. (2017). DOI:https://doi.org/10.7916/d85q532v
- [28] Marcelo Ruschel Träsel. 2018. Hacks and hackers: the ethos and beliefs of a group of Data-Driven Journalism professionals in Brazil. *Rev. Famecos* (Online) 25, 1 (January 2018), 27589. DOI:https://doi.org/10.15448/1980-3729.2018.1.27589
- [29] Edward Segel and Jeffrey Heer. 2010. Narrative Visualization: Telling Stories with Data. *IEEE Trans. Visual. Comput. Graphics* 16, 6 (November 2010), 1139–1148. DOI:https://doi.org/10.1109/TVCG.2010.179
- [30] Daniele R. de Souza, Lorenzo P. Leuck, Caroline Q. Santos, Milene S. Silveira, Isabel H. Manssour, and Roberto Tietzmann. 2018. Interacting with Data to Create Journalistic Stories: A Systematic Review. In *Human Interface and the Management of Information. Interaction, Visualization, and Analytics*, Sakae Yamamoto and Hirohiko Mori (eds.). Springer International Publishing, Cham, 685–704. DOI:https://doi.org/10.1007/978-3-319-92043-6 54
- [31] Edson C Tandoc. 2014. Journalism is twerking? How web analytics is changing the process of gatekeeping. New Media & Society 16, 4 (June 2014), 559–575. DOI:https://doi.org/10.1177/1461444814530541
- [32] José Roberto de Toledo. 2014. "Pirando" nos dados. In Análise política & jornalismo de dados: ensaios a partir do Basômetro (1a edição), Humberto Dantas, José Roberto de Toledo and Marco Antônio C. Teixeira (eds.). FGV Editora, Rio de Janeiro, RJ, Brasil.
- [33] Nikki Usher. 2012. Going Web-First at The Christian Science Monitor: A Three-Part Study of Change. *International Journal of Communication* 6, 0 (June 2012), 20.
- [34] Mary Lynn Young, Alfred Hermida, and Johanna Fulda. 2018. What Makes for Great Data Journalism?: A content analysis of data journalism awards finalists 2012–2015. *Journalism Practice* 12, 1 (January 2018), 115–135. DOI:https://doi.org/10.1080/17512786.2016.1270171