Understanding Business Analysts’ Needs for Data Report Authoring

Z. Zhang, S. Malik, S. Guo, J. Hoffswell, R. Rossi, F. Du, E. Koh

1 Adobe Research, San Jose, California, United States
2 Information School, University of Washington, Seattle, Washington, United States

Abstract

Business analysts often create static, data-driven reports to summarize and communicate findings from marketing dashboards. However, the requirements and workflow for creating data-driven reports in business analytics have not been fully investigated. In this work, we interviewed fifteen professional analysts to understand their unique needs for data-driven report authoring and identify gaps between their goals, technical skills, and existing reporting tools. Our findings suggest eight fundamental takeaways for report authoring, such as the need for persistent interactive experiences combined with more robust narrative authoring for linking story pieces and customizing the narrative layout. Based on these interviews, we synthesize the results into five design guidelines to direct future analytic reporting tools.

1. Introduction

Business analysts (BAs) are responsible for gathering, tracking, and analyzing performance metrics for their companies, and conveying important findings to stakeholders to inform operational improvements or innovation. To efficiently communicate data-driven insights and reduce time-to-action, BAs often need to create reports with visualizations captured from data analytics tools along with narrative text that explain the key insights. These data-driven reports generally utilize a static format, including illustrated documents and emails, blog posts, infographics, or presentation slides.

Following the popularization of interactive data stories in online journalism, the processes and tasks of authoring data-driven stories for public consumption has been well-studied [SH10, LRIC15, SH14a]. Despite similarities in the composition of stories for journalism and analytic reports (i.e., visualizations paired with text descriptions), data stories can be quite different from business data-driven reports. For example, BAs often have special workflows for data analysis and report authoring, and reports often need to be tailored based on different consumer personas. They also have different skill sets compared to journalists and data scientists, which may lead to distinct needs. Furthermore, data-driven stories for online journalism often incorporate engaging interactions and highly customized annotated visualizations versus the static format of reports which often appear in business analytics settings. Although there has been some debate about whether interaction in data stories engages readers [BDF15, BEDF16], these are largely focused on the general public, and few studies have focused on how business analysts create and share data-driven reports optimized for corporate settings. We hypothesize that BAs would have special needs for data reporting tools to better integrate with their data analytics workflow, and hypothesize that BA’s audiences would engage with and benefit from interactivity since they are more specialized and have a vested interest in the data.

This paper seeks to develop a deeper understanding of the workflows and tools business analysts use to create data-driven reports. To this end, we conducted semi-structured interviews with fifteen analysts with different job responsibilities to better understand their requirements, reporting needs, and pain-points for data-driven report authoring. Based on these interviews, we identified eight design needs (Sec. 3), and distill these findings in combination with our literature review into five design guidelines (Sec. 4) that can inform the design of future tools for visual analytic report authoring.

2. Related Work

There has been extensive research in the visualization community around data-driven stories [HD11, MHRL17, SH10, SLHRS16], which have a similar composition to data-driven reports. To understand how these data-driven stories are created, Lee et al. [LRIC15] outlined three stages in the authoring process: (i) exploring data to find facts (i.e., story pieces), (ii) connecting story pieces in a meaningful way to create a story, and finally (iii) presenting a story to the audience. According to Kosara et al. [KM13], the first stage is well-supported by research systems [WQM*17, SH14b, SDES18] and commercial tools (e.g., Tableau [noa16] and Power BI [Sof10a]), but such tools fall short in supporting the second and third steps. To this end, many tools have aimed to facilitate data-driven authoring [Sof10c, Sof10b, CVTH21, BKS13, CH18]. Analysts also increasingly use computational notebooks (e.g., Jupyter) for both exploration and presentation, but recent studies [RTH18, KRA*18] have described limitations in each step. Ellipsis [SH14a] aims to address the issue by decoupling narrative structure from individual visualizations. Other techniques aim to automatically link text and visualizations [MZJS18, BLE18, LZK*21], or other analytic components [MHK19]. However, none of these techniques specifically consider the needs of business analysts or the impact of these evolving technologies on their current workflows.

Prior work has studied analysts’ workflows, collaboration, and challenges during analysis [AZL19, CKW09, FDCD12,
KS11, KPHH12]. Russell et al. [RSPC93] characterize the high-level sensemaking activities necessary for analysis. Kwon and Fisher [cKFY11] discuss challenges novices encounter when using visual analytic tools. Alspaugh et al. [AZL* 19] found that their participants prefer tools that are open-source, well-supported by the participant’s organization, and have staying power in the marketplace (i.e., unlikely to be made obsolete). Sedlmair et al. [SIB* 10] discuss difficulties evaluating visualization tools in large corporations, including acquiring and integrating data. Furthermore, several researchers [CPS*06, GZ08, HMSA08] have articulated the importance of capturing provenance to manage analytic workflows, and Kandel et al. [KHP*11] have advocated the use of visualization across the analysis lifecycle. Building on this prior work, we conduct interviews with business analysts to better understand their unique data reporting needs, the current pain points of the existing tools, and the workflow that business analysts follow when converting analysis results into an understandable communication format.

3. Formative Interview Study

While there has been much prior work on analysts’ general analytics workflow [FDCD12, RSPC93, HMSA08, GZ08], the majority focus on their data wrangling and analysis process. To better understand the needs of business analysts, particularly in the context of analytic reporting, we conducted semi-structured interviews with fifteen professional analysts. We wanted to focus specifically on aspects of analytic reporting, such as where report authoring fits into their pipeline, in what format are insights communicated, and the tools and methods used for creating these artifacts. We also solicited feedback about specific features that could be included in future report authoring tools to understand the importance of each.

3.1. Methodology

We conducted semi-structured interviews with fifteen professional analysts at a large software company. Participants were recruited by word-of-mouth and internal mailing lists for visualization tool users. Sessions were conducted remotely using video conference and screenshare, and lasted between 47-67 minutes (mean = 55). We asked each participant about their current workflow and tools, specifically around reporting, aspects of collaboration, and use of mobile devices such as mobile phones or tablets for work. We also asked participants to share an example of a recent artifact (dashboard or report) that they had created to communicate insights.

The fifteen participants (ten male, five female) included twelve business analysts across various domains, two visualization designers, and one software engineer. All participants used visualization to find and communicate insights in their day-to-day job and their experience ranged from 1 to 21 years (mean = 6 years). Most participants (12/15) were familiar with GUI-based visual analytics tools such as Tableau [noa16], PowerBI [Sof10a], and Excel. Some participants (4/15) were familiar with programming languages, such as Python and R, or visualization libraries such as D3 [BOH11]. All were familiar with a GUI document editor (e.g., Word, PowerPoint) and a minority (4/15) had experience with graphic design tools, such as the Adobe Creative Suite. Interviews were recorded and transcribed. Participants’ answers were then thematically analyzed and coded by two researchers. The artifacts shared during the screensharing portion were systematically reviewed for their visualization, interactivity, and text components.

3.2. Findings

We summarize our findings in three parts: (i) current workflow and tools, (ii) content needs, and (iii) system needs.

Current Workflow and Tools. Our findings about workflow were largely consistent with previous research [AZL* 19]. Analysts start the process by gathering requirements, data wrangling, preparing the data model, iteratively exploring data, and finally reporting the results. Analysts reported a wide range of time spent on these projects, ranging from several days to six months. Seven of the participants spent 1-3 weeks to create the artifacts, five spent 1-3 months, and three participants spent over three months. While the majority of time was spent coordinating with stakeholders to gather requirements and collaborating with engineers to access and prepare the data, these steps are not the main focus of this paper; we focus on how analysts craft the story into a data-driven report.

In terms of distributing and sharing findings and reports, participants often used more than one method (N1: Portable). All but one participant shared findings through email, either as a link or attachment. Many would upload documents to central locations accessible by colleagues (e.g., Sharepoint or a wiki page). Many participants would also give live presentations of their findings (12/15).

Content Needs. To develop a better understanding of the types of content analysts generally report, we asked participants about their reporting needs, including use of interactive charts from the dashboard, annotations on charts, and accompanying descriptive text. We also asked participants to share a dashboard or artifact that they had recently created, to capture any components that they might not otherwise identify. Twelve participants shared an example.

With the exception of P5, who makes 100% static infographics, all other participants incorporated interactivity into the dashboard. The degree of interactivity varies from simple features like tooltips to elaborate techniques like sliders, search, filters, and zooming. Though most participants created interactive dashboards, less than half (6/15) delivered these artifacts to stakeholders. Instead, almost all participants (12/15) shared their findings with a secondary format by screenshots of their dashboards and using a document editor, such as Microsoft Word or PowerPoint (echoing N1: Portable).

For participants who did present findings using their dashboards, they said it was particularly useful for answering questions on-the-fly. For example, P1 explained that “[Dashboards] are useful for in meetings when there are additional questions. You don’t have to create a static view for each question.” There was also a consistent theme that reports could be improved by preserving the interactivity of their dashboards (N2: Interactive). Only one participant (P15) indicated that an interactive data-driven report would not be sufficient for their audience, noting that, “the users will want to see the information in one place without scrolling.”

Many participants (8/15) used multiple visualizations in their dashboards to help convey a single insight. These visualizations were often connected through common techniques like overview +
detail-on-demand, filtering, and brushing and linking, leading to a need for features that help with organizing multiple visualizations in the data report (N3: Linking). Only two participants (P5, P13) noted that they did not use coordinated visualizations because they were too complex and readers may lose focus on the story. For example, one participant explained that “If I were to do it, it’d be a visual breakout of maybe a pie chart into the next stage of it. But, no, I try to keep visualizations pretty simple. A lot of times I’ve found if you make it too complicated, then you get more questions about it when you want them to focus in on the story” (P13).

Participants can easily add interactivity with dashboarding tools, like Tableau or PowerBI, but felt these tools are limited when it comes to adding text and annotations for presenting standalone reports (N4: Annotations). After showing the participants some examples of interactive data stories online, most of them (14/15) were enthusiastic about this story format, noting that it would offer a new, engaging way for them to share their findings. However, some participants explained that it would be useful depending on the situation: “I wouldn’t want it available during the presentation, but let’s say I wasn’t there to give the presentation, I could point them there and then they could walk through [the narrative]” (P13).

In terms of the role of text, usage varied quite a bit. A few participants noted the proportion of visualizations to text was about 20% to 30%. For others, the artifact was much more lightweight, using only tooltips to explain fields (P15), the data model (P8), or descriptions and definitions (P9). Only one participant (P14) employed heavy text usage, where they wrote paragraphs of text to explain the entire dashboard, coupled with captions for contextualizing and summarizing each visualization. Participants annotated charts directly (7/15) using color highlights, guidelines, circles and arrows, notes, and icon markers to help mark insights. Participants highlighted several challenges when including text because dashboards needed to cater to many people whose data might be different and would need to be customized based on the user (N5: Customization). For example, P6 used PowerBI to analyze the data, but had difficulty adding customized explanations: “I sometimes need to add text on the visualizations to explain them but it is not easy.” Participants also described a challenge when creating recurring reports: captions largely followed a similar template, but needed to be manually updated with new insights (N6: Dynamic).

System needs. To have a more concrete understanding of analysts’ requirements for a data-driven report authoring system, we asked participants about benefits and pain points of current report authoring tools and preferences about specific features for a new system.

In discussing benefits and pain-points of current tools, some participants (13/15) noted that they liked to use WYSIWYG (“what you see is what you get”) document editors for communicating findings because it was easier to annotate charts, keep track of notes and insights, and incorporate text (N7: WYSIWYG). For example, P13 explained that “I actually really like PowerPoint because I can put a lot of notes of what I’m seeing in the notes area. I wish I could do that in Power BI or something. Just have a quick ‘I’m writing all these notes of what I see,’ and then have my own annotation.”

We also considered the use of tablets or mobile devices in the design of our tool. No participant used tablets for work, and the majority (10/15) had a strong preference to use only laptop or desktop devices (N8: Computer). One participant (P6) had used a mobile device, but only to join remote meetings. In addition to rarely using tablets or mobile devices, participants rarely designed dashboards and reports for screens smaller than a laptop or desktop. Participants did not feel that the ability to sketch visualizations and ideas (for brainstorming or for authoring), or the inclusion of touch input (e.g., for tablets or mobile) would be helpful for authoring. In fact, most participants (11/15) preferred either a laptop-only or a laptop plus tablet setup: “If you can make [the system] work for a tablet and desktop, that’s great, but tablet only—I’d be concerned” (P14).

When asked about potential features participants would want in the report authoring tool, all 15 participants expressed a desire to not only preserve their analysis and interactivity from their existing dashboards (echoing N2: Interactive), but to author additional interactivity between text, visualizations, and page layout. In particular, some participants (11/15) described the necessity of adding annotations and external representations, such as text overlays, to the visualizations to better convey their insights (echoing N4: Annotations). Apart from text annotations, participants (10/15) also mentioned their need for customizing visual annotations (e.g., color, size, opacity, etc.), and also adding tooltips to transfer details into mouse hovering events (echoing N5: Customization).

Linking between text elements and visualization charts was also frequently (10/15) mentioned by participants (echoing N3: Linking). In particular, when the participants were talking about this linking, they were often describing different narrative layouts, where the report author can use various layout styles to express connections between elements in different ways. We found that business analysts had a need of various layout options, depending on the particular usage scenario. Besides the relative text-chart layout, several participants (7/15) mentioned that organizing charts overall would be helpful and necessary for crafting coherent narratives (echoing N9: Layout), as some insights require multiple visualizations to explain the findings. This linking between elements is represented as navigation, where authors can use various kinds of navigation techniques to enhance storytelling and reporting experiences (e.g., transition animations, jump buttons, breadcrumbs).

4. Design Guidelines

Based on the formative interviews reported in Section 3, we propose five design guidelines for developing effective authoring tools for business analysts creating interactive, data-driven reports.

DG1: Complement business analysts’ current workflows (N8). Analysts have a strong preference for maintaining their original workflow instead of switching to a completely new tool. For example, analysts prefer using laptop or desktop tools (N8: Computer) compared to mobile devices like phones or tablets for story authoring. Beyond modality, analysts are generally familiar with existing tools like Tableau and PowerBI for exploratory data analysis workflows. However, these tools do not provide sufficient functionality to support BAs’ needs for report authoring (e.g., customized text, annotation, layout, etc.), and BAs still need to export the visualizations to other tools (e.g., presentation slides, email, Word document) for creating the reports. While tools like Piktochart and Quadrigram integrate more flexibility in the format of the data report, they require analysts to also disrupt their workflow to these
tools by uploading the raw data and analyzing the data from scratch. Ideally, the report authoring tool should seamlessly integrate with business analysts’ existing data analysis tools to support insight finding and communication beyond the data analysis workflow.

**DG2: Enable GUI-based creation of data-driven reports (N3, N4, N7, N8).** Although there are existing GUI-based commercial tools (e.g., PowerBI, Tableau) in BAs data analysis workflows, their data reporting and storytelling functionality (e.g., Tableau Stories [Sof10d]) cannot fully satisfy BAs’ needs for key features like combining narratives with interactive charts (N3: Linking, N4: Annotations) and designing narrative-based layouts (N9: Layout). To get more expressive and flexible visualization or narrative features often requires additional programming expertise (e.g., D3 [BOH11], Vega [SWH14, SMWH16]) that is unrealistic for many business analysts. Reflecting on our study findings, a GUI-based tool is therefore an essential requirement (N7: WYSIWYG) for creating data-driven reports because it can similarly leverage techniques from common document editing tools to make the interface more familiar and easier to learn as part of the overall analysis lifecycle.

**DG3: Preserve interactivity of the visualizations (N2, N5).** Interactivity was particularly beneficial when presenting the visualizations on-the-fly, but was often lost when transferring visualization from dashboard to external data reporting tools that could support contextual text and annotations. Analysts desire a way to share reports (N1: Portable) that maintain their interactivity (N2: Interactive). Future systems should support including customized narratives and annotations (N5: Customization) while preserving the interactivity of the original visualization charts.

**DG4: Support customized narratives and flexible links to visualizations (N3, N5, N9).** The role of text in data-driven reports varies depending on the audience. For many analysts, the general audience for a data-driven report are the key decision makers. It is thus important to provide additional context of the analysis beyond the self-service dashboard used by the analysts themselves. Note that the narrative text in this case does not necessary need to be long or comprehensive. Our interview studies show that most business analysts prefer to author shorter narratives with customized explanations and captions, while there are also some rare scenarios where heavy narratives need to be used (N5: Customization). Analysts sometimes used coordinated visualizations to explain a single insight and communicate multiple dimensions of the data (N3: Linking). Beyond emphasizing the connection between individual elements, analysts also require support for better organization methods for the narrative layout as a whole (N9: Layout). Our investigation shows that existing tools either provides only limited types of layout (e.g., stepper layout in Tableau Stories) or does not fully support linking of multiple visualizations (e.g., Piktochart, Quadriagram), which leaves an open space of improvement for future reporting tools for business analytics.

**DG5: Include out-of-the-box methods to generate dynamic insights. (N4, N5, N6)** In some cases, business analysts are required to periodically update reports with massive amount of new data. Due to the nature of recurring reports, business analysts also tend to preserve and templatize reports for reuse between time periods (e.g., quarterly reports), but wish the content can be automatically updated with new insights (N6: Dynamic). Existing tools all require analysts to complete data analysis workflow before report creation. To better support analysts in updating data-driven reports, future systems should support customizable and templated reports for data refreshing. This feature is also necessary for both the narrative text (N5: Customization) and annotations on the charts themselves (N4: Annotations).

### 5. Limitations and Future Directions

The long-term goal of our work is to deeply understand business analysts’ workflow and develop a data report authoring tool that enables them to author interactive data-driven reports effortlessly. To this end, we limited the study to participants who self-identified as business analysts. More work would also need to be done to generalize this to other personas, such as data scientists, who may have unique needs compared to BAs. Additionally, though BAs exist in a broad spectrum of applications (e.g., marketing, finance, machine learning), we recognize that all participants were from the same large company which may introduce biases and might limit generalizability. Due to COVID-19, all case studies needed to be conducted remotely. While this has some benefits, such as reaching a wider and more diverse pool or participants in more geolocations, it also limited participants’ ability to demonstrate their current tools and report authoring methods. An in-person study may allow for more exploratory design as well.

In future work, we will work closely with analysts to implement and iterate a new prototype report-authoring system. Using our proposed guidelines as a starting point, we will focus on creating a GUI-based system (DG2) that allows analysts to integrate state-of-the-art techniques of data exploration (DG1) and other layout styling methods (DG4) to easily create data-driven reports that can be exported in a variety of narrative formats (DG3). In addition, future work should explore new intelligent experiences for authoring data-driven reports. For example, the system could leverage the analysts’ interactivity with the dashboard to help recommend relevant visualizations and highlight useful interactive features (DG3). Furthermore, some participants expressed the needs for automatically updating the contents of recurring reports (DG5). One of the analysts also mentioned the need for a recommendation system to provide suggestions when selecting visualization charts, grouping, and layout options. The system could directly recommend visualization charts based on either the dataset or the user’s dashboards. The system could also utilize attributes and distribution of the report elements for recommending elegant layouts and groupings. These improvements could better support analysts in creating a clear, engaging, and insightful data-driven report.

### 6. Conclusion

In this work, we aimed to understand “What are the unique needs of business analysts in creating data-driven reports, and how can future systems better support report authoring?” Towards answering this question, we conducted formative interviews with fifteen professional analysts to better understand their workflow and requirements for analytic reporting. We synthesized our findings into five design guidelines to inform future data reporting tools.
References


