CHAPTER 3

Exploration and Explanation in Data-Driven Storytelling

Alice Thudt and Jagoda Walny
University of Calgary

Theresia Gschwandtner
Vienna University of Technology

Jason Dykes
City University, London

John Stasko
Georgia Institute of Technology

CONTENTS

INTRODUCTION

In visualization, it may seem that exploration and explanation are opposite ends of a spectrum: that increasing the focus on explanation and communication necessarily reduces the ability for exploration. However, data-driven stories challenge this notion by combining exploration and explanation in creative new ways.

Exploration provides readers with the power to find their own story in a set of data, while explanations communicate an author's narrative about the data. Exploratory facets are useful for data analysis and personalized navigation. Explanatory facets, on the other hand, are essential for communicating information and viewpoints about data and establishing emphasis and story lines for the reader. The purpose of visual data-driven stories is to communicate a narrative and present information based on data, so explanatory elements are a defining aspect of visual stories.

The notion of considering exploration and explanation as opposites was influential in early work in interactive cartography [1] and is seen in data journalism [2], where the journalist's goal is to provide a narrative through and supported by data. However, this distinction has collapsed as interactive interfaces have become ubiquitous. As digital technologies develop and widen the opportunities for interactive data interfaces, increasing numbers of graphics concurrently have both exploratory and explanatory characteristics. This form of visual storytelling offers new possibilities for design, engagement, discovery, and narrative devices [3,4]. Developing practice in data-driven storytelling has established compelling examples

that show how powerful exploratory explanations and explanatory explorations can be (see Figure 3.1(a)–3.1(c)).

As visualization researchers, we see value in providing exploratory power to readers of data-driven stories. While explanation is a powerful way to provide a narrative for readers and orient them within large and complicated issues and data sets, exploration enables readers to make their own inquiries, personalize their reading experience, and get a feeling for the limits and the shape of the data. Providing exploratory power can also be a way to communicate complexities in the data and mitigate some of the biases inherent in providing a narrative.

We argue that making a visualization more exploratory does not necessarily take away from its explanatory value and vice versa. As shown in Figure 3.2, we view explanation and exploration as two complementary capacities of visualizations that can be successfully integrated in different ways. The "Film Dialogue" story [5] (see Figure 3.1(a)–3.1(c)) shows that this is indeed possible. The visualizations and accompanying text provide a clear narrative and message, i.e., that men dominate movie roles and dialogues. Meanwhile, the visualizations enable the reader to investigate individual films more deeply. These visualizations are interactive, show large amounts of individual data points, and are presented using multiple different types of views. Recognizing that exploration can be added to a story such as this without taking away from its explanatory power can push us to imagine more sophisticated combinations of exploration and explanation that can simultaneously communicate coherent narratives while empowering readers to gain a fuller understanding of the supporting data.

Data-driven stories that are highly exploratory and highly explanatory at the same time are still difficult to find. However, as we will show in this chapter, by breaking down exploration and explanation into more manageable dimensions, it is possible to imagine such stories. As we discuss exploration and explanation, we invite readers to consider the explanatory and exploratory aspects of the data-driven stories they have seen and to imagine how one could increase exploration and explanation simultaneously in the examples we show.

In this chapter, we document interesting examples that show how exploration and explanation interact and complement one another in the evolving landscape of data-driven storytelling. We describe the

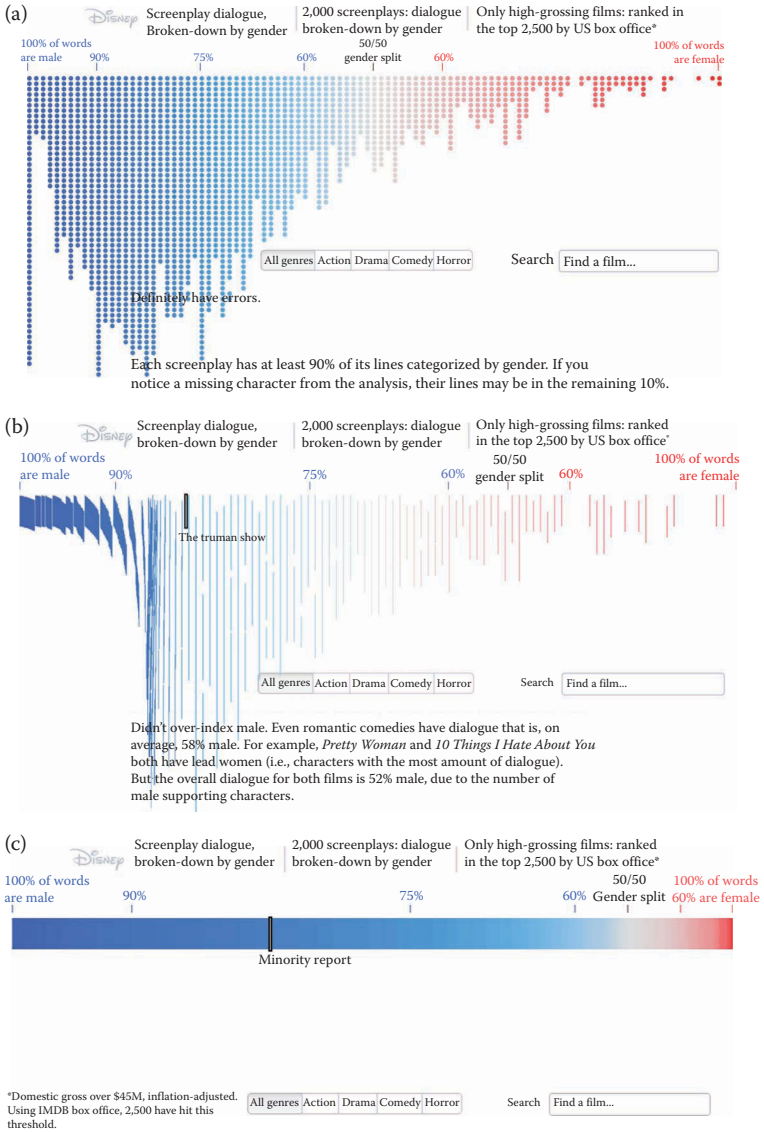

FIGURE 3.1 "Film Dialogue." (a) "Film Dialogue" is a narrative comparing female to male dialogue in 2000 screenplays of popular films. (b) When scrolling through the explanatory text, animated transitions link each view to the next. (c) The story's main message that women have far less dialogue than men is clearly communicated by the views and the accompanying text. Yet, the reader can explore each view interactively to investigate individual films. Thus, explanation and exploration complement one another to form an engaging story. (Source: Polygraph. http://polygraph.cool/films/. Used with permission. [5])

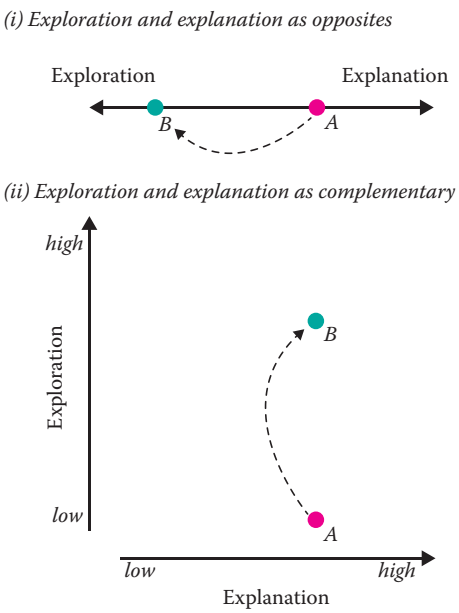

FIGURE 3.2 Considering exploration and explanation as opposites (i) implies that when an exploratory visualization (A) becomes more explanatory (B), it also becomes less exploratory. In this chapter we take the position that the two can also be seen as complementary (ii) in that a visualization (A) can be made more exploratory (B) while maintaining its explanatory capacity.

exploration-explanation space in light of recent visualization research and our experience in exploratory visualizations. We identify three dimensions—view, focus, and sequence—along which exploration and explanation can vary. By viewing data-driven stories through the perspective of these dimensions, we can illustrate a variety of ways in which exploration and explanation can be integrated, and we can identify interesting features of the exploration-explanation landscape. Our hope is that this view will inspire designers to populate this landscape in new ways, that it will help developers to produce enabling technologies that extend it, and that researchers will find intriguing research questions about the relationship of exploration and explanation to communication. As all of these groups work to improve and expand the experience of reading data-driven stories, our perspective of exploration and explanation as being complementary rather than opposites draws attention to a space for designing graphics that are both highly explanatory and exploratory, hence pushing the boundaries of what is possible in data-driven storytelling.

CHARACTERIZING EXPLORATION AND EXPLANATION IN VISUAL DATA-DRIVEN STORIES

Exploratory facets of data-driven stories, to which we refer collectively as *exploration*, are those facets that provide readers with the flexibility to ask a variety of questions of the data, to personalize their experience of the story based on their own interests, and to view the data from different perspectives.

Explanatory facets of data-driven stories, to which we refer collectively as *explanation*, involve an author's deliberate communication of context or interpretation relevant to the data or its presentation. The author's message can be communicated directly (e.g., through text or audio) or indirectly through the visualization.

Characteristics of Exploration

Exploration is characterized by access to additional data, and the functionality or forms of the representation. Adding exploratory facets allows authors to create stories that can be personalized according to readers' knowledge, interests, and experience as they develop.

Exploration can be supported by allowing readers to choose among a large set of different preselected representations, vary the visual variables used in these representations, and change the focus by selecting particular data attributes. Including features for personalizing a story to the readers' context is another way to let readers find their own narrative (for an example, see the Neighbourhood Statistics Quiz of the UK government described in the section entitled "Flexibility in Choosing the Focus").

Even static representations can have exploratory characteristics. Consider, for instance, the network visualization of muesli ingredients [6] shown in Figure 3.3. It allows the viewer to ask multiple questions by looking at different items in the graphic. For example, a person could be interested in finding common cereal-fruit combinations or in assessing how the most popular ingredients are mixed together. The graphic shows the complexity of a data set with room for individual interpretations, while providing little explanation, narrative structure, or starting points for such explorations.

Many exploratory visualizations, however, provide interactive capabilities, especially when larger data sets are presented. Interaction allows the reader to investigate different aspects of the data that may not be initially

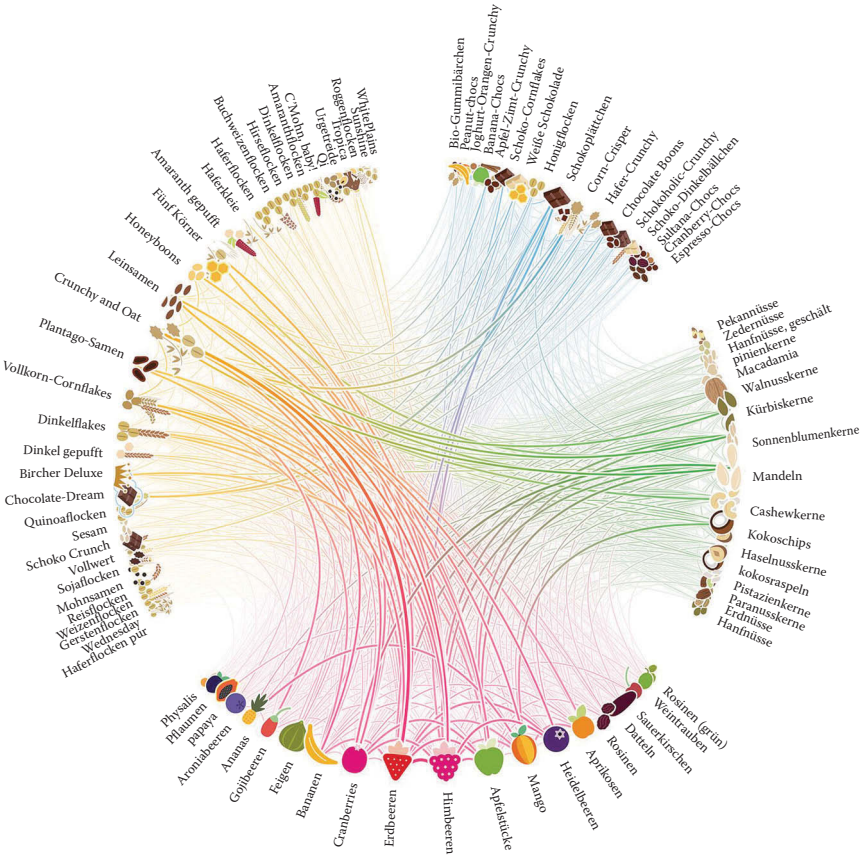

FIGURE 3.3 Static representations can have exploratory facets by providing a large enough amount of complex information that can be explored from different angles. (Source: Stefaner, M. http://truth-and-beauty.net/projects/muesli-ingredient-network. Used with permission. [6])

evident. It is the most common way to support various reader intents, including selection, filtering, abstraction, elaboration, and exploration [7].

The exploratory characteristics of a data-driven story can also vary over different stages of the narrative. For instance, in the "martini glass" structure described by Segel and Heer [8], a reader is first led along a single narrative path (the "stem" of the martini glass) and is then provided with an exploratory interactive visualization through which multiple paths are possible (the "mouth" of the martini glass).

Characteristics of Explanation

Explanation is characterized by an authorial presence that, directly or indirectly, provides a reader with interpretation, context, or other information relevant to the subject of the story beyond the raw data.

Stories with a strong explanatory component often make use of selection and visual emphasis to guide the focus of the reader's attention to the aspects that are relevant to the message of the story. For example, Figure 3.4 shows two graphics of the same data but with differing amounts of explanation. The visualization on the left shows the data with only sufficient information to interpret the graph. The visualization on the right communicates the data in a more explanatory way by adding interpretation in the headline and corresponding emphasis within the visualization to draw attention to the message. While the same subset of data is shown in both graphs, the right one clearly suggests a focus and sequence by emphasizing the data that is most relevant to the message. This form of emphasis is typical for explanatory visualizations.

However, the explanatory aspects can be introduced in much more subtle ways than in the example above, which can be seen, for instance, in *The Guardian's* visual story about China's economic slowdown [9] (see Figure 3.5). The graphic depicts how China's economic slowdown could affect other countries' economies (Figure 3.5). It uses text and the visual metaphor of dragging down to provide an interpretation of the data.

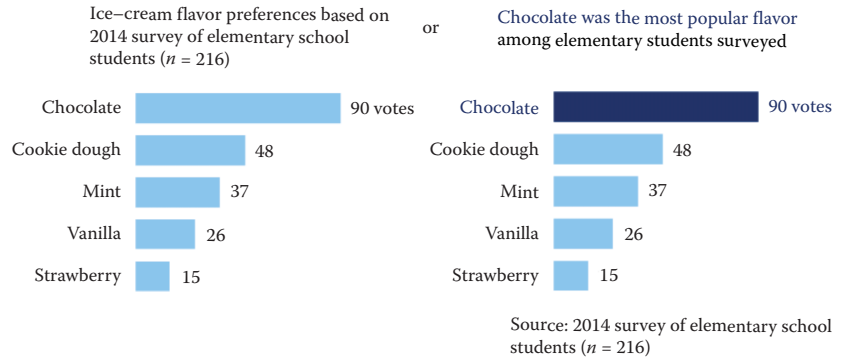

FIGURE 3.4 Two charts displaying the same data. The chart on the right has more explanatory qualities than the chart on the left; it contains added interpretation in the form of a headline and emphasis within the chart. (Source: Kemery, A., Four Storytelling Strategies. http://annkemery.com/four-storytelling-strategies/. Used with permission. [19])

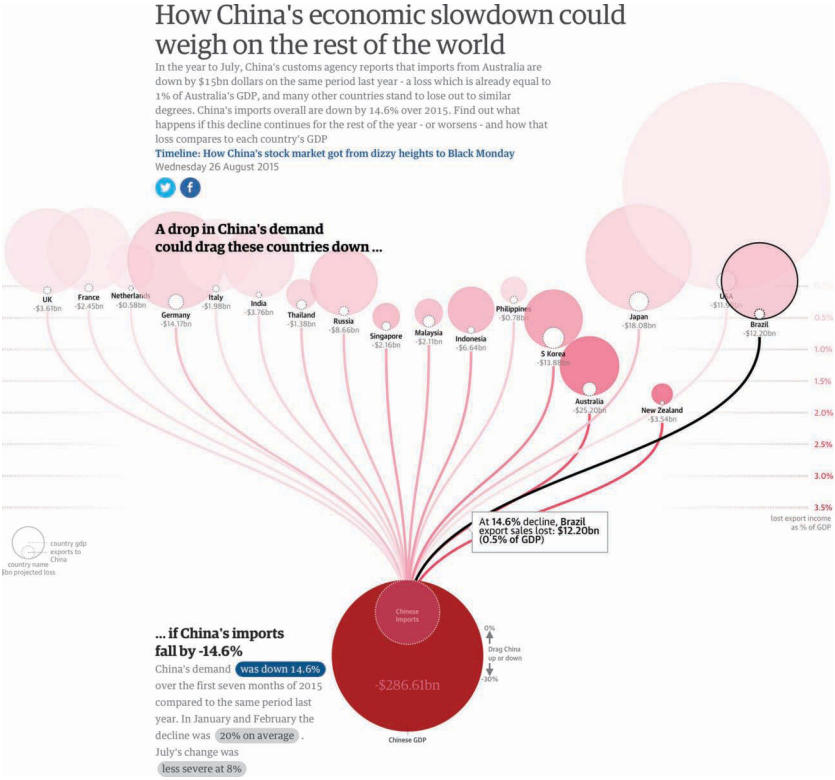

FIGURE 3.5 This interactive data-driven story contains explanatory elements, such as the headline, text, tooltips, and visual metaphor. It also contains exploratory elements: a reader can explore the effect of different values of China's economic slowdown by sliding the bottom circle, representing China, up or down. This updates the visualization of the remaining countries accordingly. (Source: *The Guardian*. http://www.theguardian.com/world/ng-interactive/2015/aug/26/china-economic-slowdown-world-imports. Used with permission. [9])

China has been specifically selected as the focus of the story here since it is the only country that can be interactively investigated, and only relationships with China are represented. In this example, further explanation is provided by adding a title and by giving interactive control over only one aspect of the visualization, i.e., the amount to which China's imports fall. The narrative is so strongly focused on China that the reader may not even think of investigating the effects of falling imports of other countries (this information is not provided in the story anyway). Hence, controlling *where* interaction is possible may be an effective way of guiding interpretation.

Besides interactive visualizations, long-form text and linear sequences are common in data-driven stories with a strong explanatory focus. Text is an effective way to communicate context and interpretation, and linear sequence is a natural companion to long-form text. Highly explanatory stories are often finite: they have a beginning, an end, and often a linear progression. There is a growing number of examples that weave text together with visualizations, as for instance, in the story on fracking in California (see Figure 3.9) in which the visualization updates as the reader scrolls through the story. There have also been efforts to automatically annotate visualizations in news articles [10].

The same techniques used for communicating an interpretation can also be used for communicating facts (as in the headline on the left bar chart in Figure 3.4), or to provide assistance with reading the visualizations or using the interactive features of the visualization.

These more-objective aspects of explanation (i.e., giving hints about where to look, annotating facts, or providing assistance with interacting with the visualization) are related to the concept of guidance [11]. However, guidance is aimed at supporting users of an interactive visualization in executing their tasks, while explanation is aimed at communicating the author's message.

Explanation can be highly visible or it can be more subtle. For example, the choice of visual encoding may influence how the message of a story comes across. However, providing explanation involves an explicit editorial process: the author deliberately chooses what to show and how to show it in order to support and communicate the story. It is author-driven and focuses on communicating specific information to the reader.

DIMENSIONS OF DATA-DRIVEN VISUAL STORIES

Data-driven visual stories typically contain both explanatory and exploratory facets that can be integrated with each other to differing degrees. To illuminate the ways in which exploration and explanation can be varied within a data-driven story, we distinguish between two aspects of these stories: the *flexibility* they provide to the reader and the *interpretation* suggested by the author of the story. We can consider flexibility and interpretation individually in terms of three dimensions:

The ***view***: The ways in which data is shown to the reader, including the available encodings and any transformations made on the data (such as normalization or aggregation).

The ***focus***: The subject of the story, the particular set or subset of data shown, and the aspects of data that are shown.

The ***sequence***: The order in which the information in the story can be viewed.

How flexibility and interpretation are provided on these three levels determines the degree to which a visualization supports exploration and explanation. Flexibility is most strongly related to exploration, while interpretation is most strongly related to explanation. In this section, we discuss how flexibility and interpretation can be provided through these three components, and we discuss exploratory and explanatory capabilities of data-driven visual stories in relation to these dimensions.

Flexibility

A visualization that supports some degree of flexibility gives the viewer control over the view, focus, or sequence of the story. The flexibility that a visualization can provide is fundamentally affected by the chosen view and the complexity of the graphic. Even a static graphic with a complex view, such as the Muesli Chart in Figure 3.3, can provide flexibility in terms of the focus (the reader can choose the subset of information to examine) and the sequence (there are multiple paths to reading the graphic).

Flexibility in the View

A visualization can provide flexibility in the view by allowing readers to control the mapping directly or choose from a number of alternative visual mappings. In visual storytelling, alternative visual mappings are frequently provided, for instance through separate views, coordinated views, or animated transitions. Techniques that allow the reader to control the visual mapping directly are less common, but they provide a high degree of flexibility. Flexibility can also be provided by allowing readers to manipulate the way the data is arranged in the visualization, for example, data aggregation or normalization (for a more thorough discussion of data specification aspects, see Chapter 10, entitled "Ethics in Data-Driven Visual Storytelling").

Below, we describe some examples of varying degrees of flexibility in view.

Multiple separate views: Alternative views of the same data set are shown separately, either side by side using the same type of

visualization representation to show different subsets of the data (i.e., small multiples [12]) or in different parts of the text to highlight different aspects or perspectives on the data. An example of using separate views to highlight different aspects of the data is the visualization of Boston's Massachusetts Bay Transit Authority (MBTA) data (Figure 3.6). It shows adjacent views: a subway map on the left displays how many people enter and exit each station per day and a bar chart with integrated heatmaps on the right provides more information on the passenger numbers during different times of the day.

Multiple coordinated views: The same data is presented in multiple coordinated views that allow readers to make connections across these views, and thus, across different aspects of the data set. Such visualizations typically use a technique known as "brushing and linking." Another example of using coordinated views to give readers freedom is "A World of Terror" [13] (Figure 3.7). It shows an interlinked matrix of charts that can be rearranged and filtered.

Animated Transitions between Views: This method gives readers the opportunity to fluidly transition from one representation to another. The "Film Dialogue" visual story (Figure 3.1) makes use of this method. It

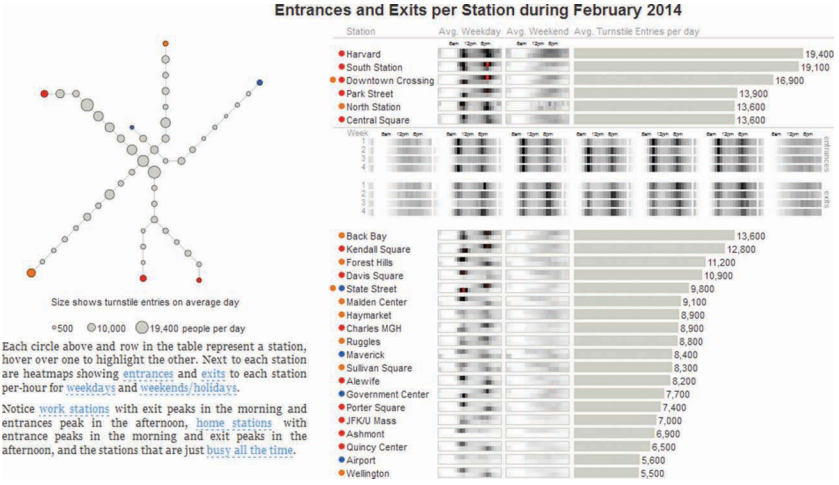

FIGURE 3.6 This graphic juxtaposes a subway map, heatmaps, bars, and text to provide a flexibility in the view. It further supports flexibility of focus by using a variety of small graphics that reveal different aspects of the data. (Source: Barry, M. and Card, B., Visualizing MBTA Data. http://mbtaviz.github.io/. Used with permission. [14])

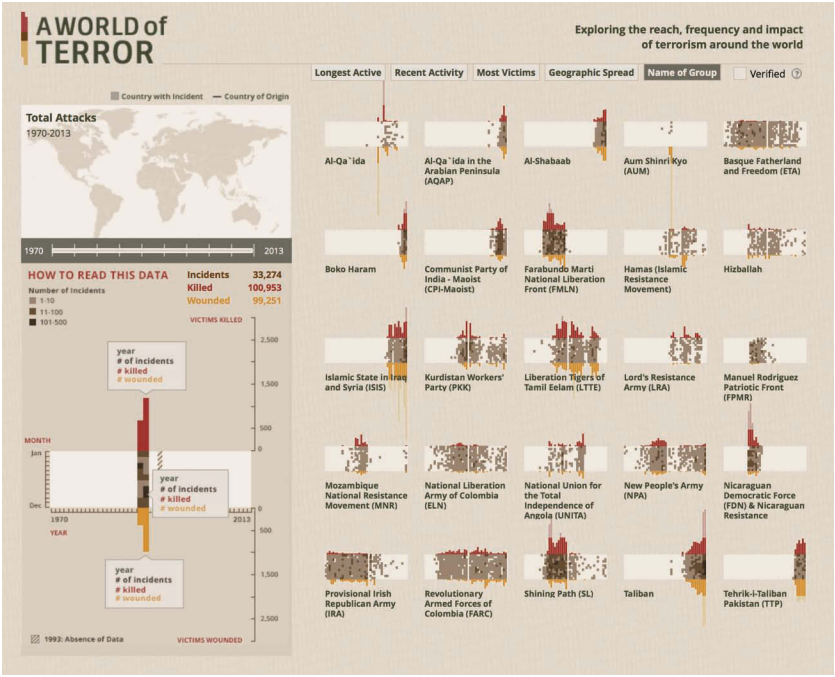

FIGURE 3.7 The multiple coordinated views shown here allow the reader to explore connections through brushing and linking. (Source: Periscopic, A World of Terror. http://terror.periscopic.com/. Used with permission. [13])

shows charts that morph as the reader scrolls through the story linking the currently shown plot to the surrounding text. Showing the transitions between different views can help to both understand different perspectives of the same data and make the relations between these perspectives easier to follow. This technique of transitioning between visualizations to match and accompany the current position in text is a recent trend in visual storytelling sometimes referred to as "scrollytelling."

Reader-driven visual encoding: A reader-driven visual encoding allows readers to manipulate visual variables, such as color, size, or spatial arrangement of data items, directly. For instance, in a visualization presenting data through a scatterplot, the reader may have the freedom to choose which data dimension to map on the *x*-axis, which data dimension to map on the *y*-axis, and which dimensions to portray through glyph size, shape, and color. Although not commonly adopted in visual storytelling so far, this technique provides the highest degree of flexibility in the view. One reason for the low adoption of this technique in visual stories

might be that its high degree of flexibility can easily lead to confusion and unfruitful exploration. Therefore, and also to communicate the intended story, authors of visual stories often prefer to limit flexibility and add focus and sequence to increase levels of explanation.

Flexibility in Choosing the Focus

Another way to support flexibility in a visualization is to allow readers to view different subsets of the data or decide which attributes to show. A prerequisite for providing flexibility in the focus is that the visualization offers access to a sufficient amount of data so that the readers can choose their own focus by investigating the items in the data set that are most meaningful to them. As discussed earlier, a static visualization with high visual complexity allows the reader to investigate different subsets in detail. More commonly, however, visualizations provide interaction techniques that allow the reader to change focus, such as *filtering, selecting, zooming and panning, and drill down*.

For instance, the visualization of Boston's MBTA data [14] (Figure 3.6) allows the reader to focus on various subsets of the data by providing many different visualizations on various aspects of Boston's subway system in combination with interactive focus selection.

Another example of letting readers select the focus of the story is the Neighbourhood Statistics Quiz of the Office for National Statistics in the UK [15] (Figure 3.8), which enables personalization. In the very first step, it asks the readers to enter their postal code and subsequently lets them guess demographic facts about their area while giving immediate visual feedback. Thus, while a rigid set of demographic facts are reported, a very personal data-driven story unfolds in visual form as the extent to which the reader knows an area that is important to them is revealed along with their knowledge gaps and biases.

Flexibility in Choosing the Sequence

Narrative is all about order, but interactive stories can be designed in flexible ways that allow readers to determine order or aspects of order [16]. While static visualizations do not limit the number of viewing sequences, they may overwhelm the reader with huge amounts of data all presented simultaneously. Conventional videos, however, can only be viewed in one sequence. Besides these two extremes, interactive visualizations provide techniques to support different degrees of flexibility. For instance, one-dimensional flexibility is provided by interactions such as scrolling

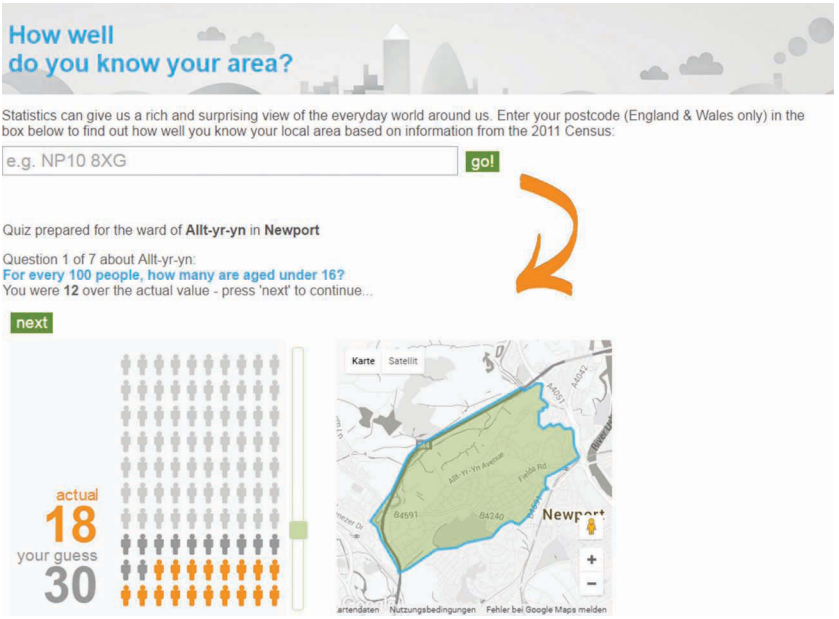

FIGURE 3.8 This website allows readers to focus on a subset of census data for a particular location of interest by inputting a postal code. (Source: UK Office for National Statistics. http://www.neighbourhood.statistics.gov.uk. Used with permission. [15])

forwards and backwards, or by "next" and "back" buttons. One example of this can be found in the graphic "California's Getting Fracked" [17] (Figure 3.9), where the visualization on the map of California is updated as the reader scrolls or steps through the story.

Two-dimensional flexibility can be provided, for instance, by panning and zooming on a map.

Many data-driven stories with strong narrative elements limit the flexibility of the viewing sequence of the presented information [18]. Such stories frequently use step-by-step (i.e., "stepper") or scrolling (i.e., "scrolly-telling") interface controls that update a story as the reader scrolls through the presentation. More flexible methods that allow a reader to explore freely typically permit the reader to jump to different visualizations in any order.

Interpretation

Interpretation refers to any kind of meaning added by the author. Interpretations can be subjective by adding emotional aspects, persuasive,

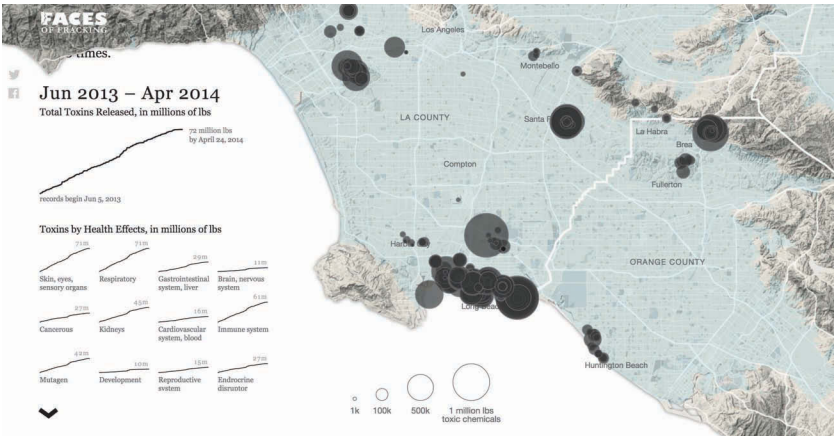

FIGURE 3.9 Readers can experience this visualization in different sequences by stepping through different points on the right or by scrolling to step through in an author-determined linear order. (Source: Flagg, A., Craig, S., and Bruno, A., California's getting fracked. http://www.facesoffracking.org/data-visualization/. Used with permission. [17])

or personal opinions. But there can also be more objective interpretations that point out important facts in the data or help the reader find interesting aspects in the chart. With *interpretation* we refer to any authorial/editorial presence in the presentation of the data. This can be conveyed through text or audio descriptions, but can go beyond a plain description of a graph and what can be seen in it.

Consider for instance the bar chart shown in Figure 3.4. The headline in Figure 3.4a says: "Ice cream flavor preferences based on 2014 survey of elementary school students" [19]. No interpretation of the data is imposed in the statement. In Figure 3.4b the caption reads: "Chocolate was the most popular flavor." This statement chooses a specific aspect of the data to point out. Additionally, this interpretation is supported through the visual encoding by highlighting the bar representing chocolate, thus guiding the focus of the reader to the intended interpretation. In this case, the interpretation is conveyed textually and through the visual representation itself. In the following section, we describe how interpretation of a visualization can be conveyed through the view, the focus, and the sequence.

Interpretation through the View

An interpretation can be conveyed through the choice of representation, through other aspects of the visual encoding (e.g., the choice of

color or through animation), or through the choice of how the data is transformed for representation (e.g., data aggregation or data normalization). Hullman and Diakopoulos [20] describe how default views, such as showing specific comparisons or presenting the data in a specific order, prioritize some information. Moreover, the provided interaction mechanisms influence the way readers interact with the presented information, and thus, these mechanisms may lead readers to explore certain subsets of the data, and eventually, they influence the conclusions that readers draw from the visualization [16]. A powerful example of using the view to support an interpretation is "Iraq's bloody toll" [21] showing deaths in Iraq over time in a red upside down bar chart (Figure 3.10). The red supports the interpretation of the Iraq war as "bloody" and the simple decision regarding the orientation of the bars emphasizes this message. This is a basic decision, but one that has considerable effect. And while it may seem obvious—how many upside down bar charts have you ever seen?—creative redesign, even of the most fundamental principles or established norms used in our graphics, can have powerful effects.

Interpretation through the Focus

An interpretation can further be communicated by guiding the focus of the reader specifically to the aspects that support the story. One way of doing this is through highlighting and annotations (as in Figure 3.4) or by fading out or omitting data that is not directly relevant to the interpretation. Additionally, over the course of a story the focus can change to support the narrative. One example of varying the focus to support a story is "California's Getting Fracked" (Figure 3.9). In this example, the story starts by displaying the Los Angeles Basin area and later zooms out to reveal the entire map of California. The focus is varied over the course of the article to support the interpretation that the Los Angeles Basin is only one of the more severely polluted areas in California that is affected by fracking.

Interpretation through the Sequence

Many data-driven stories show visualizations in a linear sequence that supports a narrative and interpretation. Suggesting a sequence in which the data should be viewed to understand the interpretation can be done in many different ways, for example, by using steppers, scrolling, and video formats. One example that uses sequence in an effective way is

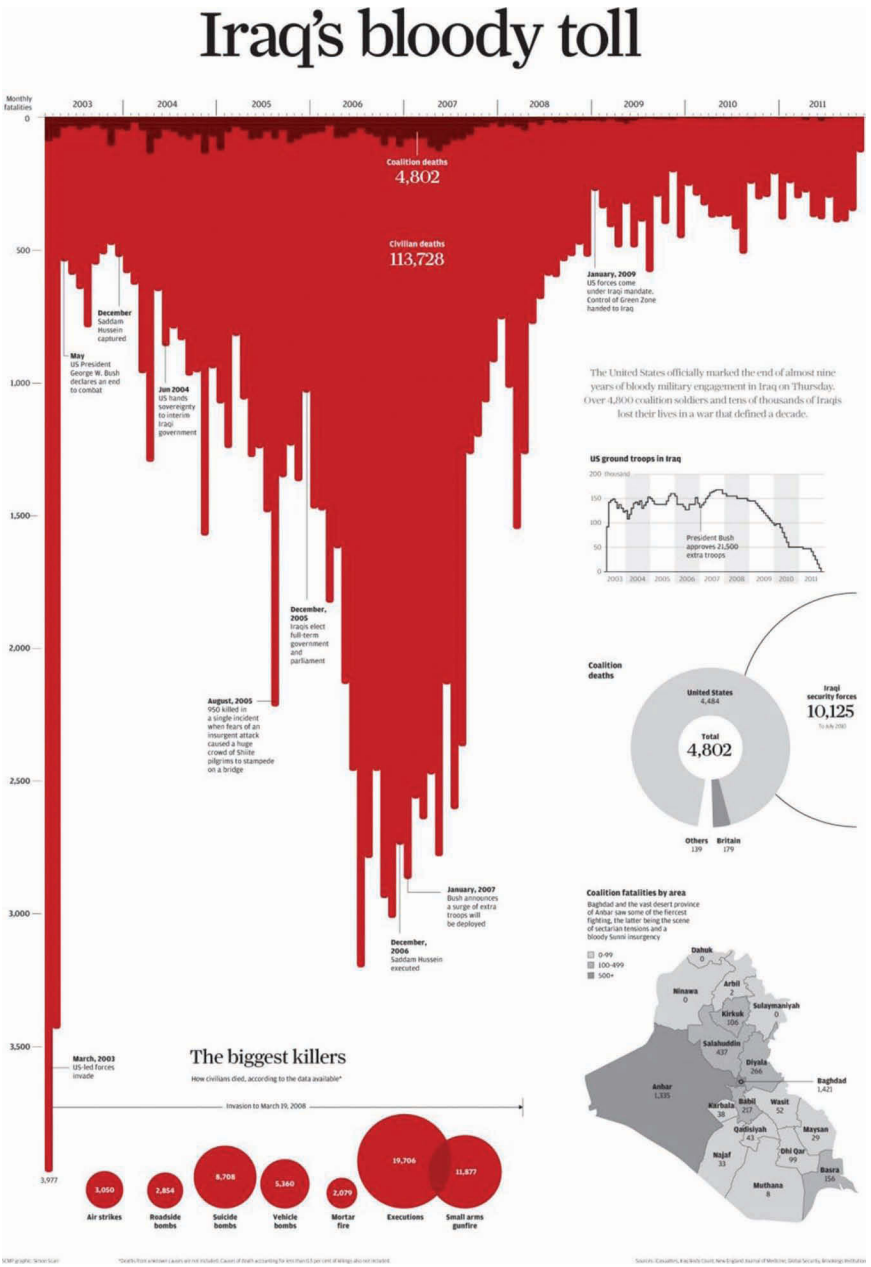

FIGURE 3.10 The visual encoding in this graphic suggests a clear interpretation of the data that fits with the author's narrative. (Source: Scarr, S. Iraq's bloody toll. http://www.scmp.com/infographics/article/1284683/iraqs-bloody-toll. Used with permission. [21])

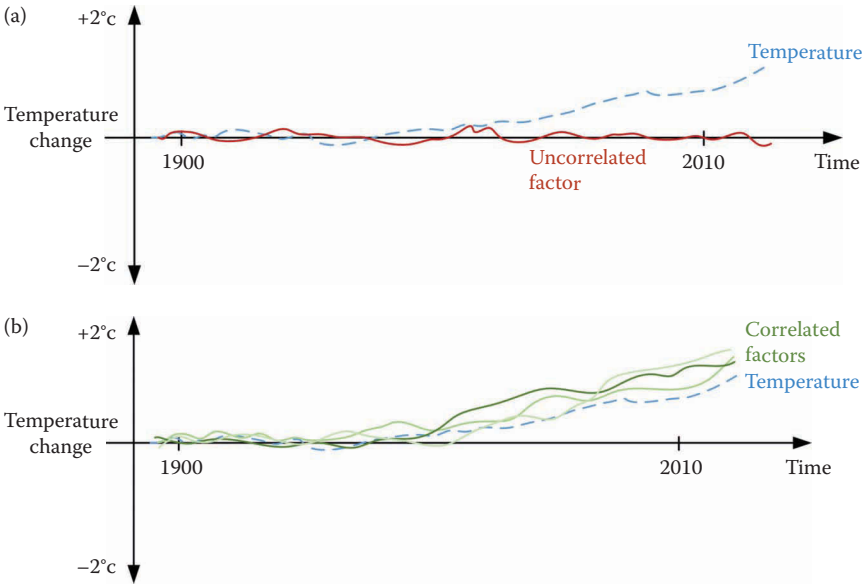

FIGURE 3.11 This data-driven story uses a linear sequence of visualizations to build a narrative. (a) Data is sequentially revealed to create suspense and support the message. First, two uncorrelated data time series are compared. (b) In the last step of the sequence, the author has placed all previously shown data within one view to support the story's conclusion. (Source: Roston, E., and Migliozzi, B. What's really warming the earth? https://www.bloomberg.com/graphics/2015-whats-warming-the-world/. Used with permission. [22])

"What's really warming the world?" [22] (Figure 3.11). Rather than showing all data at once, the story creates suspense by letting it build up step-by-step. Possible factors that have been argued to contribute to global warming are revealed over time (Figure 3.11(a)) and lead up to the conclusion that greenhouse gases are in fact the main cause of global warming (Figure 3.11(b)). In this case the sequence directly drives the interpretation of the data.

Summary

We consider the amount of flexibility and interpretation provided in a data-driven story as being key to determining the extent to which the story supports explanatory and exploratory capabilities. An explanatory visualization provides an interpretation of the data that is deliberately communicated to readers either by choosing a particular view, focusing

on specific aspects of the data, or providing a narrative sequence that supports the interpretation. Many existing storytelling devices apply here (see Chapter 6 for an explanation of storytelling devices). On top of this, an interpretation can further be provided in the text component of a data-driven story, e.g., through annotations or surrounding longer text blocks.

A visualization with exploratory aspects would typically allow for flexibility by either providing control over the choice of view, the focus, or sequence in which the data is viewed.

This flexibility can help readers to find their own interpretation in the visualization, but it is, of course, channeled in its own way through editorial decisions in the way in which flexibility is provided. Degrees of flexibility can be varied in particular parts of stories to combine the freedom of exploration with the structure of explanation.

BENEFITS OF EXPLORATION AND EXPLANATION

If we consider both exploration and explanation as capacities that a visualization can support, authors must consider why and to what extent to include them in their data-driven visual narratives. In this section we discuss the benefits of both exploratory and explanatory visualizations.

Benefits of Exploration

Exploratory visualizations can provide readers with a feeling of agency by allowing them to find their own interpretations of the data, sometimes in more subtle ways than the highly effective "You are Here" style used in the Neighbourhood Statistics Quiz (Figure 3.8). Viewers can go beyond predefined narratives to follow up on their own interests and questions. Another benefit of exploratory visual narratives is to provide transparency by making it possible to check some of the statements presented in a narrative through exploration of the visualization components of the story. In this way, exploratory visualizations can provide the potential to reveal biases. While sometimes a simple communication of facts can be desirable, exploratory visualizations can provide the benefit of highlighting the complexity of a subject and showing different sides in one representation. Text tends to be linear and has limited information-carrying capacity, while visualization is two-dimensional and can effectively show a skilled reader multiple relationships and nuances in trends. This can foster a balanced presentation of a topic.

Drawbacks of Exploration

However, in order to freely explore a visualization, readers need to have a basic amount of literacy of both visualization as well as the presented topic or data. Being able to get more from an exploratory visualization than one would get from a linear narrative requires self-motivation to learn more about the topic. Data exploration takes time, attention, and cognitive investment which readers might not always be willing to provide. Crafting a tailor-made exploratory visualization for a specific data set requires significant design effort as well. Therefore, authors must carefully consider whether the readers they are targeting with their narrative will bring the motivation, skills, and time to engage with an exploratory visualization. Alternatives include limiting or going without exploratory capacities or using more generic exploratory tools.

Benefits of Explanation

Explanatory visualizations have the benefit that the provided message or interpretation can help engage an audience with no prior knowledge of or interest in a topic. The provided guidance and suggested sequence in explanatory visualizations can make them seem more manageable and accessible to readers. The stronger editorial component in explanatory visualizations means that the author has control over the message that is being communicated. This can be beneficial since he or she can match the level of complexity to the previous knowledge of the target audience, which makes it more likely that people take something away from the visualization. Another benefit of a clear message is that the perspective that is shown might be more explicit in an explanatory visualization that can make eventual biases more obvious.

Drawbacks of Explanation

A disadvantage of purely explanatory storytelling is that readers could get frustrated for not being able to ask and investigate their own questions about the data. Users who are increasingly used to exploring stories with interactive visualizations may miss this flexibility. In a purely explanatory story, readers may not be able to derive the substory that is actually relevant or of interest to them (as they could with personalized stories). Besides, the potential of an explanatory story to spark interest and trust heavily relies on the author's perspectives as well as narrative and presentation skills.

TABLE 1.1 Pros and Cons of Exploration from a Reader's Perspective

Exploration Pros	Exploration Cons
Reader has agency of interpretation	Requires a level of visualization and data literacy
Ability to answer own questions	May require knowledge of the topic
Transparency: understand the source of the interpretation, potential to recognize bias	Requires self-motivation to learn about the topic and data
Can provide awareness of the complexity of a subject	Possibility of getting lost
	Requires time, attention, and cognitive investment from the reader
	An exploratory visualization tailored to a specific data set requires significant design investment

TABLE 1.2 Pros and Cons of Explanation from a Reader's Perspective

Explanation Pros	Explanation Cons
Message and points of interest are provided by the author	Readers may disagree with the perspective communicated
Facilitates navigating through the data	Reliant on trust in the author's perspective
Author does the work to engage the reader	Limited freedom to experience the story in one's own way
May inspire interest in unfamiliar topics	Potential omission of data that is personally important to the reader

SUMMARY

In this chapter we have identified several dimensions that affect the explanatory or exploratory nature of a data-driven story: the amount of flexibility (exploration) and the amount of interpretation (explanation) provided in the view, sequence, and focus of the story. This gives us a starting point for considering how to balance explanation and exploration in a story as we try to gain some traction in considering these tightly interrelated concepts.

The idea that data exploration and explanation are two ends of a continuum has been used by many and is useful for distinguishing established types of data-driven stories. However, we have presented exploration and explanation not as opposites, but as complementary characteristics. By using the dimensions of flexibility and interpretation of view, focus, and sequence, we show that exploration and explanation can be integrated to differing degrees to push the boundaries of data-driven story design (see Figure 3.12).

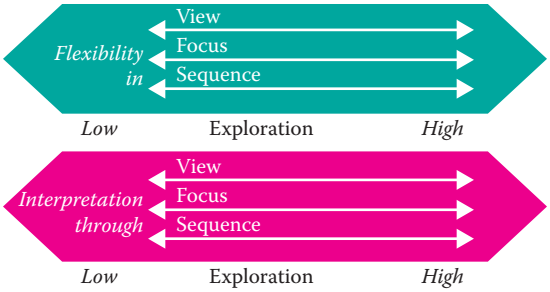

FIGURE 3.12 A data-driven story can have different levels of flexibility in its view, focus, or sequence and can convey differing levels of interpretation through its view, focus, and sequence. Manipulating a story's position along any of these dimensions can help to integrate explanatory and exploratory elements into a data-driven story.

Recognizing that exploration and explanation complement each other can inspire new ways of integrating the two beyond juxtaposing text with visualization. Even the simplest bar chart visualization provides slightly more exploratory freedom than text on its own. It is one thing to state that "Chocolate was the most popular flavor"; it is another thing to *show* that chocolate was the most popular flavor by displaying a bar chart of all flavors tested, emphasizing the bar showing chocolate, and providing context such as what other flavors were tested and how much they differed. The potential for integration goes far beyond this, however, as seen in the California fracking story in Figure 3.9, which tightly weaves explanatory text directly into the visualization.

Close balancing of exploration and explanation has the potential to achieve engaging experiences and tell persuasive stories by providing an author's context and suggested interpretation together with readers' power to personalize their experience of a story and ask their own questions. Data-driven storytelling fundamentally invites this integration of exploratory elements, a key difference from purely text-based storytelling. Thus, exploration and explanation can coexist in data-driven stories, bringing a variety of levels of freedom and interpretation into their view, focus, and sequence. Indeed, considering the various ways in which a data-driven story can support both exploration and explanation could lead to new forms of data-driven storytelling and a new understanding of what data-driven storytelling is.

REFERENCES

1. Alan M. MacEachren. *How Maps Work: Representation, Visualization, and Design*, 1st edition. The Guilford Press, New York, 1995.
2. Alberto Cairo. *The Functional Art: An Introduction to Information Graphics and visualization*. New Riders, Berkeley, CA, 2012.
3. Robert Kosara and Jock Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5): 44–50, 2013.
4. Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. More than telling a story: A closer look at the process of transforming data into visually shared stories. *IEEE Computer Graphics and Applications*, 35(5): 84–90, 2015.
5. Hanah Anderson and Matt Daniels. Film dialogue. In: *A Polygraph Joint*, Apr 2016 (not modified). ®2016 Polygraph. All Rights Reserved. http://polygraph.cool/films/ (accessed: Dec 12, 2016).
6. Moritz Stefaner. Muesli chart for mymuesli. In: *Truth & Beauty Operations* (not modified). ®2012 Moritz Stefaner. All Rights Reserved. http://truth-and-beauty.net/projects/muesli-ingredient-network (accessed: Dec 12, 2016).
7. Ji Soo Yi, Youn ah Kang, John Stasko, and Julie Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6): 1224–1231, 2007.
8. Edward Segel and Jeffrey Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6): 1139–1148, 2010.
9. Carlo Zapponi, Sean Clarke, Helena Bengtsson, Troy Griggs, and Phillip Inman. How china's economic slowdown could weigh on the rest of the world. In: *The Guardian*, Aug 26, 2015 (not modified). ®2015 The Guardian. All Rights Reserved. https://www.theguardian.com/world/ng-interactive/2015/aug/26/china-economic-slowdown-world-imports (accessed: Dec 12, 2016).
10. Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar. Contextifier: Automatic generation of annotated stock visualizations. In *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, pp. 2707–2716, 2013.
11. Davide Ceneda, Theresia Gschwandtner, Thorsten May, Silvia Miksch, Hans-Jörg Schulz, Marc Streit, and Christian Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 111–120, 2017.
12. Edward Tufte. *Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, 1983.
13. Periscopic. A world of terror. In: *Periscopic* (not modified). ®2014 Periscopic. All Rights Reserved. http://terror.periscopic.com/ (accessed: Dec 12, 2016).
14. Michael Barry and Brian Card. Visualizing MBTA data. June 10, 2014 (not modified). ®2014 Michael Barry and Brian Card. All Rights Reserved. http://mbtaviz.github.io/ (accessed: Dec 12, 2016).

15. UK ONS, Office for National Statistics. How well do you know your area? statistics.gov.uk (not modified). Licensed under Open Government Licence. ℗ONS, Office for National Statistics, UK. All Rights Reserved. http://www.neighbourhood.statistics.gov.uk/HTMLDocs/dvc147/ (accessed: Dec 12, 2016).
16. Jessica Hullman, Steven Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. A deeper understanding of sequence in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12): 2406–2415, 2013.
17. Anna Flagg, Sarah Craig, and Antonia Bruno. California's getting fracked. In: *Faces of Fracking* (not modified). Licensed under Creative Commons. ℗2014 CEL Education Fund. All Rights Reserved. http://www.facesoffracking.org/data-visualization/ (accessed: Dec 12, 2016).
18. Charles D. Stolper, Bongshin Lee, Nathalie Henry Riche, and John Stasko. Emerging and recurring data-driven storytelling techniques: Analysis of a curated collection of recent stories. Technical Report MSR-TR-2016-14, Microsoft Research, April 2016.
19. Ann K. Emery. Chocolate bar-charts. In: *Ann K. Emery Data Analysis + Visualization*, Feb 24, 2016 (not modified). ℗2016 Ann K. Emery. All Rights Reserved. http://annkemery.com/four-storytelling-strategies/ (accessed: Dec 12, 2016).
20. Jessica Hullman and Nick Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12): 2231–2240, 2011.
21. Simon Scarr. Iraq's bloody toll. In: *South China Morning Post*, Dec 17, 2011 (not modified). ℗2011 South China Morning Post Publishers Ltd. All Rights Reserved. http://www.scmp.com/infographics/article/1284683/iraqs-bloody-toll (accessed: Dec 12, 2016).
22. Eric Roston and Blacki Migliozzi. What's really warming the earth? In: *Bloomberg*, June 24, 2014 (not modified). ℗2014 Bloomberg L.P. All Rights Reserved. https://www.bloomberg.com/graphics/2015-whats-warming-the-world/ (accessed: Dec 12, 2016).