Tell Me a Data Story

Kat Greenbrook, Rogue Penguin

ABSTRACT

In search of insight, we usually focus on data - access, preparation, modelling, analysis, and visualization - but the value of data isn't always apparent to the rest of the business. Data alone is not motivation enough for someone to invest their time in a project unless the value is clearly obvious. An effective way to engage internal business users with analytics is with a data story. Data stories paint a bigger picture and evoke an empathy for what the numbers actually represent.

Contrary to popular belief, a data story is not a data visualization. Data visualization is a great way to *tell* a data story, but it relies on the story having been *written* first. A well-written data story, combined with a good visualization, will improve insight communication. Principles of graphic design are rarely taught to analysts, but an understanding of design will lift any data visualization to the next level.

Kat shows you how to write a data story and effectively tell it using data visualization. She shares an introduction to graphic design from a data perspective and illustrates how applying simple design principles in SAS® Visual Analytics can enhance your data storytelling.

INTRODUCTION

Data is everywhere, but it's what we do with that data, (what insights we tap into) that make it so valuable. The pursuit of insight in a data-fuelled world is make or break for a business. Insight is what could put them ahead of the competition, save company resources, improve customer experience, or even save lives.

But all this power means nothing if the insight never leaves the analytics team. Insight communication is an important responsibility of anyone working with data. If insight is not effectively communicated, there is less hope of it initiating change – essentially making it (and everything upstream from it) useless.

Data stories are an effective way to engage people with data insights. Humans are evolutionarily hard-wired for narrative, and data stories evoke empathy for what the numbers represent. Data storytelling is about communicating your insights effectively - giving your data a voice.

WHAT IS A DATA STORY?

A "data story" is a narrative, written with data as supporting evidence.

Put simply, a story supported by data.

Data visualizations are not data stories.

Data visualization can be an effective way to tell a data story (see below) but relies on the story having been written first. Some visualization techniques lend themselves to enhance data storytelling, but these do not apply to all forms of data visualization. Not every data visualization needs to tell a story.

Statistics are not data stories.

Putting a statistic (or multiple statistics) on a page is not a data story. It is just a number.

Data science is not a data story.

Data science is a behavior. It's the technical practice/process followed to uncover data insights and ultimately add business value.

HOW TO DISCOVER A DATA STORY

The role of Data Storyteller often falls on the person already knee-deep (or even completely submerged) in data. They have cleaned it, transformed it, explored it, visualized it, analyzed it, modelled it, and now they have to use it to 'tell a story'.

But to raise one's head out of the detail of data, with the intention of focusing on just one of the many possible narratives, can be a very hard thing to do. Sometimes the deeper an understanding of a topic, the harder it is to tell just one story. One will therefore try to incorporate as much information as possible, creating a very jumbled and often non-existent 'story'.

I've created the below 'Uncovering a Data Story' template to help filter and organize information for a specific narrative.

Figure 1. The Rogue Penguin Data Story Canvas

DATA STORY CANVAS [DISCOVERY]

rogue penguin

CHARACTER	EVENT	EVENT PERIOD	STATE BEFORE	STATE AFTER	IMPACT
Who's involved?	What happened?	When and of what duration was the **event**?	What was the **character** like before the event?	What was the **character** like after the event?	How does the event affect the **character**?
					☐ Positive impact ☐ No impact ☐ Negative impact

REASON	REACTION	BACKGROUND/FURTHER RESEARCH
Why did the **event** occur?	What could be done to continue/reverse the impact on the **character**?	What other information does my audience need to know to understand the **character** or **event**?

www.roguepenguin.co.nz

- *Character:* Who will the story focus most on? This can be your business, a customer group, a product, a country, etc.

- *Event:* What happened to the character to make their story worth telling now? Focusing on a single event will help to keep the resulting narrative concise and compelling.

- *Event period:* If an event has occurred, it's occurred over a certain period of time. This could anything from billions of years to milliseconds, depending on the event affecting the character.

- *State before/after:* Data is helpful here to describe the character's state. Ideally, you want to include information showing the change in the character as a result of the event.

- *Impact:* What kind of an impact does the event have on the character? The more impact the event has (positive or negative), the more powerful the story could be.

- *Reason:* Why did the event happen? Sometimes this can't be explained, which is okay if the narrative states this.

- *Reaction:* Is there anything that could be done in the future to continue a positive impact or reverse a negative impact? This may or may not be able to be told from a data level.

- *Background:* This section involves an understanding of the audience you're attempting to communicate with. What is their current level of understanding of the topic and what other information do they need to know to be able to understand the character or event?

HOW TO *WRITE* A (DATA) STORY

Story creation is instinctual to humans. We do it every day with our family, our friends, our colleagues… our stories may not be eventful or particularly exciting, but their underlying narratives are structured.

For example:

"Did you see what Evan did yesterday? No…? Well, he was in charge of the bake sale at the school fair and because all his kids go to the school, the school trusted him to deliver it. But his alarm didn't go off that morning and he slept in. The organisers were in a panic trying to track him down. In the end they had to cancel the baking table."

This narrative structure can essentially be simplified into three steps.

Narrative writing: Step 1

The beginning should set the scene and provide enough background to prepare an audience with the context they need to understand what comes next.

For example:

"Did you see what Evan did yesterday? No…? Well, he was in charge of the bake sale at the school fair and because all his kids go to the school, the school trusted him to deliver it."

Narrative writing: Step 2

This part of the narrative is where conflict or tension is introduced. Step 2 changes the narrative direction from where it was expected to go in Step 1.

For example:

"But his alarm didn't go off that morning and he slept in. The organisers were in a panic trying to track him down."

Narrative writing: Step 3

The end of the narrative should offer some sort of a conclusion. This could be in the form of an outcome, recommendation, solution; something that ultimately resolves the narrative.

For example:

"In the end they had to cancel the baking table."

The narrative thought process that comes so naturally to us, is disrupted when we start talking about data. We lose what makes our communication engaging and rely on the numbers to speak for themselves. Data is not the story, it just helps support a bigger message. We can't supply our audience with multiple graphs and statistics, and hope they create their own narratives. We need to write it for them.

Writing a data narrative is no different than creating any other narrative. A data story narrative takes into account these same three steps.

An example data story:

- **Step 1:** *"Churn is a problem for our business and churn rates have been increasing steadily over the last 18 months."*

- **Step 2:** "*But regional analysis shows areas with the highest churn rates also have the largest competitor activity."*

- **Step 3:** "*Therefore, targeting specific regions with offers more in-line with other providers, could help to retain customers."*

A good data story will be written with purpose (knowledge of the action it is trying to create), it will target a specific audience, and communicate a message (through narrative and supporting metrics).

Data stories are so much bigger than the metrics that help support them.

HOW TO *TELL* A (DATA) STORY

Only after a data story is written can it be told, and there are many ways to do this.

Verbal

From the camp fires of our ancestors, this is perhaps the most common form of storytelling. Stories served as instructions, warnings, and inspiration – all purposes still relevant today.

Visual

Written words, images and graphs can all be used to aid in telling a story. Visual data storytelling relies less on the software and more on the storyteller.

Both verbal and visual

Combining verbal and visual methods of storytelling covers multiple communication preferences. Presentations are a great format for storytellers. An audience can hear a story and have it supported further with visuals.

USING DESIGN CONTRAST TO TELL DATA STORIES

Contrast is fundamental in storytelling. Narratives introduce contrast between steps 1 and 2 (see above). Stories told visually use design contrast to highlight their message. If visualized well, the graph takeaway should be obvious.
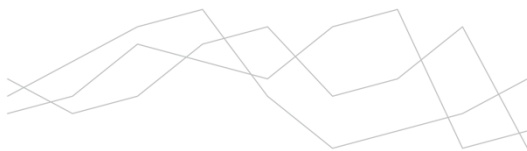
Some simple design changes can be used to add contrast to a data visualization.
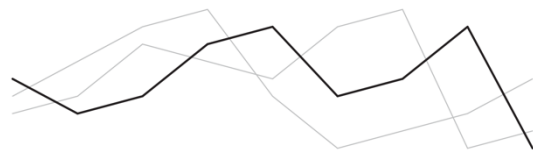
Font and Line weight

Both font weight and line weight can be used to add design contrast to a data visualization.

Figure 2. Applying line and font weight differences to a data visualization.

Font size

Differences in size help display visual hierarchy in design. There are graph elements more important than others and this hierarchy can be represented using size. For example, a chart's subtitle is less important than its title.

Figure 3. Applying font size differences to a data visualization.

Colour

Colour is one of the most polarizing yet powerful, elements of a data visualization. When used in data storytelling it should be done strategically to help highlight what's important – keeping in mind accessibility.

Figure 4. Applying colour to a data visualization.

CONCLUSION

Sharing insight in a concise and engaging way is an important responsibility of anyone working with data. Data cannot speak for itself. If insight is not effectively communicated, the potential it holds will never be realized.

Data storytelling is an effective way to communicate insight. A narrative is written to create a data story, and data visualizations help to then tell this story.

But too often when a data visualization is created for the purpose of communication, it gives no indication of its desired message. The audience is left to explore the chart and create their own insight (something that could differ depending on their level of data literacy and subject knowledge).

In search of insight, we usually focus on data - access, preparation, modelling, analysis, and visualization. Generally we spend very little time, effort and resources on how to effectively communicate this insight. But what is communicated of our insight, is all those outside of an analytics team will understand about it. Therefore, if this insight communication is not effective, neither is the analytics team that created it.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Kat Greenbrook
Rogue Penguin
kat@roguepenguin.co.nz
www.roguepenguin.co.nz