

Paper 4640-2020

Tell Me A Different Data Story

Kat Greenbrook, Rogue Penguin

ABSTRACT

The easiest data story to write involves an event—something in real life—changing over time. This change can be measured with the right data metric, comparing its value before and after the event. The more dramatic the change, the more interesting the data story.

But insight projects are usually more complicated than a single event. Many involve the analysis of multiple events or the output of different segmentation profiles. This insight also needs to be effectively communicated to initiate change.

This paper builds on knowledge in the paper *Tell Me A Data Story (3168-2019)*, Kat presented at SAS® Global Forum 2019. This paper explains how to write different data stories, depending on the insight you have to communicate, and how to effectively tell them using data visualization. All of these learnings can be applied to SAS® Visual Analytics to enhance your data storytelling.

INTRODUCTION

Storytelling has long been an exercise in communication. Despite centuries of practice in the craft of storytelling, its use has traditionally been reserved for certain disciplines (humanities) with little place in the analytical worlds of business and science.

But while disciplines continue to specialize, they evolve with influence from outside their branch of knowledge. Today, *storytelling* is a popular word in the vocabularies of most businesses. Analytical disciplines are combining storytelling techniques to more effectively communicate their data. In order to gain the power of storytelling to communicate data, one has to first learn the humanities craft.

Effective data communication uses analytical tools to find the right metrics, narrative structure to craft the message, and data visualization to make it easily understood.

I call this *finding, writing, and telling* a data story.

FINDING A DATA STORY

Most data stories are built around one metric. Other metrics will likely be included, but one metric will be the main focus of the story. This metric is compared either across time (in *Event* data stories) or across characters (in *Character* data stories).

EVENT DATA STORY

Event data stories compare one metric, for one character, at different points in time. I've created the below *Event Data Story Canvas* to help filter and organize information for a narrative built around an event.

Figure 1. Event Data Story Canvas

EVENT DATA STORY CANVAS

Event data stories compare one metric (the focus of the story), for one main character, at different points in time.



<p>CHARACTER</p> <p>This could be a person, product, animal, country etc.</p>	<p>EVENT</p> <p>What happened?</p>	<p>EVENT PERIOD</p> <p>When and of what duration was the event?</p>	<p>STATE BEFORE</p> <p>What was the character like before the event?</p>	<p>STATE AFTER</p> <p>What was the character like after the event?</p>	<p>IMPACT</p> <p>How does the event affect the character?</p> <p><input type="checkbox"/> Positive impact</p> <p><input type="checkbox"/> No impact</p> <p><input type="checkbox"/> Negative impact</p>
<p>REASON</p> <p>Why did the event occur?</p>		<p>REACTION</p> <p>What could be done to continue/reverse the impact on the character?</p>		<p>BACKGROUND/FURTHER RESEARCH</p> <p>What other information do my audience or myself need to know to understand the character or event?</p>	

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:* Your focus metric is used 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/Further Research:* 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? What further research could be undertaken to further understand the event?

CHARACTER DATA STORY

Character data stories compare one metric, for multiple characters, at the same point in time. I’ve created the below *Character Data Story Canvas* to help filter and organize information for a narrative comparing characters.

Figure 2. Character Data Story Canvas

CHARACTER DATA STORY CANVAS

Character data stories compare one metric (the focus of the story), for multiple characters, at the same point in time.

CHARACTERS These could be people, products, animals, countries etc.



METRICS How can I measure the differences between characters?					 OVERALL
<input type="text"/>					
<input type="text"/>					
<input type="text"/>					
<input type="text"/>					
<input type="text"/>					
<input type="text"/>					
BACKGROUND What background information should my audience understand about these characters? Is there anything I need to research further?					

www.roguepenguin.co.nz

- *Characters*: Who am I trying to compare in my story? This can be customer segments, products, countries, animals etc.
- *Metrics*: Your focus metric is used to describe each character's state. Ideally, you want to include metrics showing the differences between characters.
- *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 characters?

WRITING A DATA STORY

We've all heard someone say that a story needs a beginning, middle, and end. Maybe we cringed thinking about how simple it sounds. But stories DO need a beginning, middle, and end – just not in the chronological way we may think.

These three parts give stories a Three Act structure, a narrative model used by storytellers for centuries.

- *Act 1 (Agreement Act)* sets the scene by introducing the character and providing some initial context.
- *Act 2 (Conflict Act)* contrasts Act 1 and in doing so increases the tension. For the narrative to advance, you need contradiction.
- *Act 3 (Resolution Act)* resolves the conflict between Act 1 and 2.

When writing stories for a business context, I prefer to use Randy Olson's version of the Three Act narrative structure. He uses the words *and*, *but*, and *therefore*, to differentiate between acts. This structure is super flexible, conversational, but still advances the narrative. We can use this ABT (and, but, therefore) Three Act structure when writing our data stories.

Always start with a high-level Three Act structure (ABT) as this will summarize the story and define its scope. If you have trouble framing your ABT in a few sentences, you may be trying to tell more than one data story at once. When you've settled on a high-level ABT you can begin to add further detail where required. The narrative structure remains the same, but the level of information gets deeper.

EVENT DATA STORY

High-level story (ABT): This will give context to the focus metric and identify how the event has changed it. For example, below is a high-level story written to explain how climate change is affecting green sea turtles.

The sex of a baby green sea turtle is determined by the temperature of the nest it develops in. But a warming climate is altering this temperature-dependent sex determination process. Therefore, there is an immediate need for nest management strategies to avoid a green turtle population collapse.

Nested story (ABT): These expand on the overall story, while providing context and detail. For example, below is a nested narrative to further explain the first sentence of the high-level story above.

When a female green turtle reaches about 30 years old she will often return to the beach where she was born and lay her eggs. She'll dig a pit in the sand with her flippers, fill it with around 100 eggs, then cover her nest and return to the sea. Her eggs will hatch after two months. But unlike most animals where sex is determined at fertilization, the sex of a baby turtle is determined by the temperature of the nest it develops in. Cooler nests produce more male baby turtles, and warmer nests more females. The temperature that produces an equal number of male and female turtles is known as the "pivotal temperature". The pivotal temperature for green turtles is 29.3 °C. Therefore, green turtles have temperature-dependent sex determination or TSD.

CHARACTER DATA STORY

High-level story (ABT): This will give context to the focus metric and identify what segments to profile further. For example, below is a high-level story focused on the vaccine confidence metric for health professionals in New Zealand.

The New Zealand government funds vaccinations for all New Zealanders, and the majority of adults strongly agree vaccinations are safe. Most health professionals have more confidence in vaccines than the public. But the vaccination confidence of midwives and alternative medicine practitioners is low. Therefore, the type of health professional seen by someone could be an important influence on their own confidence in vaccines.

Segment story (ABT): These expand on the overall story, while providing context and detail for an individual segment. Metric comparisons to other segments can help make information relative but segment stories should focus only on the specific segment. Each segment story includes the focus metric.

For example, both midwives and alternative medicine practitioners were identified in the above high-level story. These segments can be expanded on in their own story, like the midwife story below. Comparisons have been made to nurse and GP segments to help make the training metric more relative.

Most women in New Zealand will choose a midwife for their maternity care. Midwives are qualified after three years training (by comparison, nurses also train for 3 years and GPs for 11 years). But one in three midwives doesn't have a strong confidence in vaccines. Therefore, understanding the reasons for vaccine resistance of midwives, could help in changing the anti-vaccination attitudes of the public.

TELLING A DATA STORY

Data visualization is a popular way to tell a data story. But these visualizations should only be used if they aid in the communication of the data story. Graphs used for communication differ from graphs used to help uncover insights.

- Takeaway Title

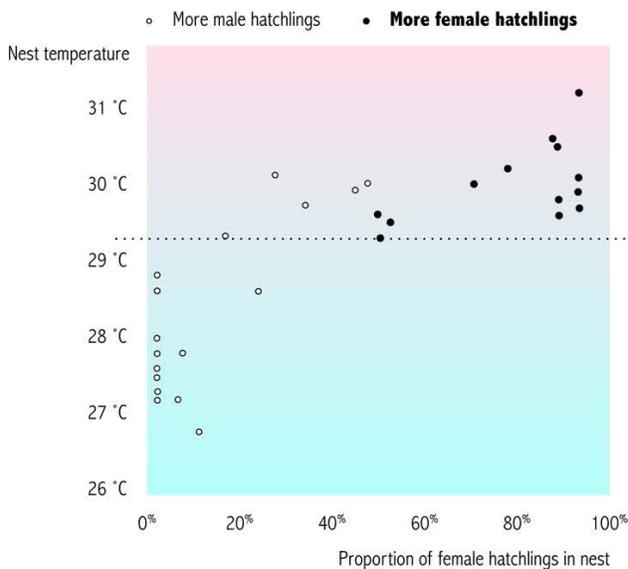
The purpose of a graph when used to tell a data story is communication. This may or may not be through the graph itself. To make your communication even more effective spell out your message in your graph title.

A title of a graph used for analysis could read *Profit over time*, reflecting the metrics it displays. But if this graph was used for the purpose of communication, the title could read *Profit has decreased over the last 12 months*, reflecting the message.

- Design Contrast

Contrast is fundamental in storytelling. Narratives introduce contrast in Act 2. Stories told visually use design contrast to highlight their message. Any contrast added should support the graph's takeaway title, making it easy to see how the data visualization supports the story message. For ways to add design contrast refer to paper *Tell Me A Data Story (3168-2019)*.

For example, a graph for the above green sea turtle narrative could be visualized like below:



Nests above the pivotal temperature produce more female baby turtles.

The pivotal temperature for green turtles is **29.3°C**

CONCLUSION

Data storytelling is a collaboration between analytics, narrative, and visualization. Effective data communication uses analytical tools to find the right metrics, narrative structure to craft the message, and data visualization to make it easily understood. I call this *finding*, *writing*, and *telling* a data story.

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