
1


Let me start by telling you a story.

2

A few years ago, in the middle of a cold winter, a friend of mine was driving around a 4-wheeler,

3

misjudged the level of water and ice in a frozen river, and ended up getting the vehicle stuck in the frozen water!

4

A group of friends including myself came over to help him try to figure out how to pull it out of the water.

5

One friend put trashbags on his legs to stay dry and waded into the water to attach a chain to the vehicle.

6

We tried towing it out with a pickup truck. But, almost immediately, we popped a tire on a broken tree trunk. So, the truck quickly became useless.



Thankfully, the friend had access to a tractor, so we got that to try to pull the vehicle out. None of us really knew how to drive it, but we gave it our best shot. As soon as we started pulling, though, the tractor started slipping on the frozen dirt and began to fall into the river as well! It seemed like this day could not get any worse.



But before the tractor fell in the water, in the nick of time, we found the four-wheel drive setting, turned it on, and were able to pull out the 4-wheeler.



We towed it all the way back to my friend’s house, where we celebrated our victory!




Now let me tell that story again, in a different way.



A friend of mine had a 4-wheeler.

It got stock in frozen water and we tried to get it out.



First, we attached a chain to the vehicle.



Next, we tried pulling it out with a pickup truck, but that didn't work.



Then, we tried pulling it out with a tractor. It didn't work at first,



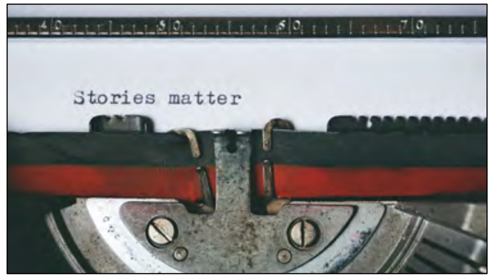
but eventually it did.



We celebrated.

These two stories had pretty much the same information, but they weren't equally effective.



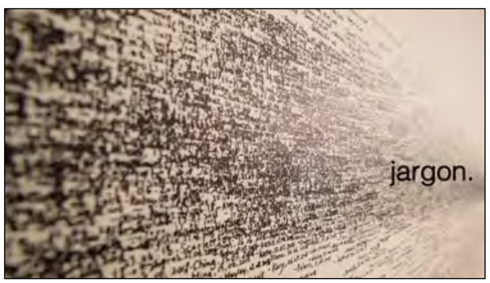



It's not the facts themselves that matter, it's how you tell them to another person that determines how much they care about what you have to say.



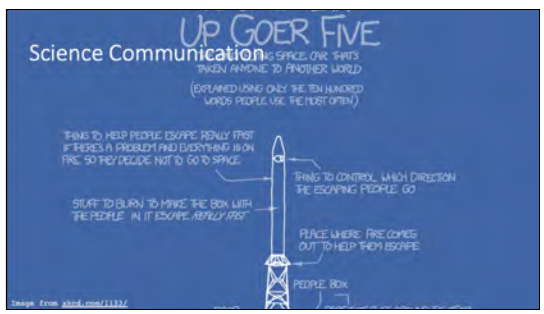

At its core, visualization is a form of science communication - just like a podcast, a dinosaur exhibit at a museum, or a lecture by Bill Nye the Science Guy or Neil Degrasse Tyson.






Like with all science communication, there's a risk of losing your audience due to assumptions about their prior knowledge, and using science jargon they might not be familiar with.






This comic makes fun of the jargon of rocket science by simplifying a diagram of the Saturn V rocket to descriptions using only the "ten hundred" most common words in the English language.

(The word "thousand" isn't used, because it isn't one of the most common thousand words in the English language)

It can be tempting to think that people don't understand our jargon because they're stupid. That couldn't be more false. People's expertise lies in different places. No one is an expert in everything. As science communicators, it's important to find a way to make a science message accessible to different kinds of audiences without "dumbing it down".

XKCD has a fun tool you can use to try to re-write your own text with simplified language.


This is an example of the jargon I use to describe my work.


As you can see, most of these words are not among the top 1000 English words.


So here is a rewrite. While it seems funny - it's undeniably a lot easier for people to understand. I wouldn't actually want to use this exact text in describing my work, but this is a good exercise in thinking about your word choices.

There's no shame in using understandable language to help connect you with your audience!


The first step in creating a strong message is knowing who you're trying to reach.



My friend tells the 4-wheeler story much more casually to friends



than he does to his car insurance agent.



If he was trying to tell the story to a Hawaiian from the 1800's, a lot of the things in the story would need explanation – what snow is, what a vehicle is, why water is bad for an engine. There would even be a language barrier.

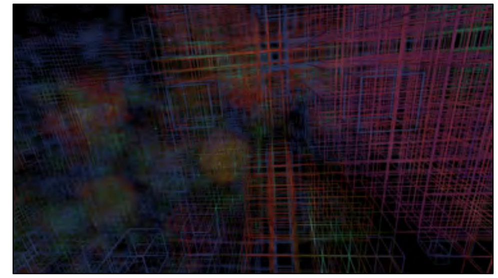
If you are presenting your science to two or more different audiences, you should consider either changing the presentation for every unique audience, OR make one presentation that's intended to be accessible to all audiences.




Experts often have an easier time showing off the depth of their knowledge by using technical terms and complex images. But chances are, your audience is more likely to be impressed if they understand the work that is being done.

What's more, it demonstrates more depth of knowledge when an expert can explain their ideas to various audiences. The most renowned scientists in the world are famous because they are good at this.





The lab I'm part of at the University of Illinois creates 3D scientific visualizations in the form of movies aimed at broad audiences that might include children, grandparents, experts in many fields, and policy makers.

A risk of speaking to a general audience is that you're going to bore the experts, and confuse the novices.

We avoid this by including nuanced details in our imagery that satisfy the most discerning experts in the field, but focusing on stories that are timeless and meaningful to everyone - stories that address the origins of life, the workings of human society, and threats to our safety.

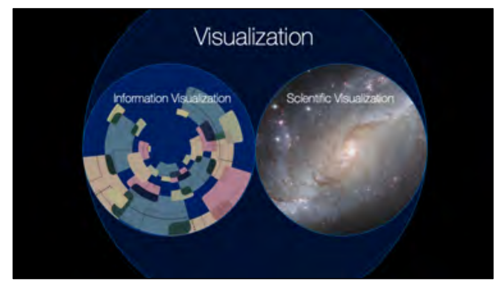



Before I go any further, I want to take a moment to do a bit of vocabulary housekeeping, because some of you may have watched that video and thought that this sort of visualization has nothing to do with the sort of work that you are doing. Visualization is a broad field and I'm going to use the word interchangeably to refer to a couple different types of things.




The first is information visualization, which involves creating things like scatterplots, histograms, and networks in 2D that show relationships between data points.



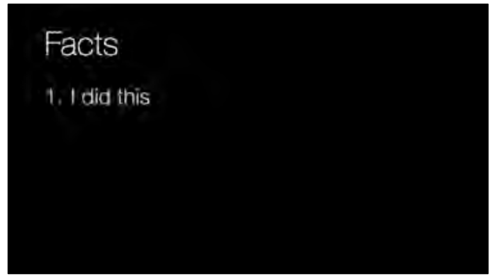



Then there's scientific visualization, which involves visualizing the 3D spatial nature of a dataset. The positions of stars, the cosmic web, galaxies, the orbits of planets, things that have X, Y, and Z position (or sometimes just X and Y), and those positions are part of what gets visualized.

My expertise lies in scientific visualization, but many of the core concepts, especially those involving storytelling, remain the same across the two visualization subfields.


SO, let's have a frank conversation about facts.

37

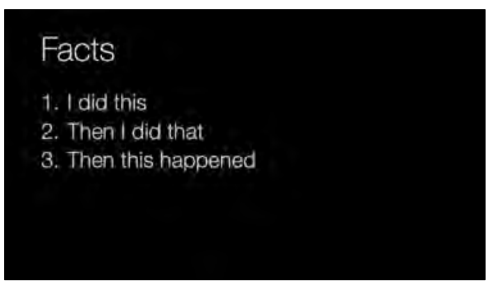



Our minds are not wired to digest lists of facts.

38

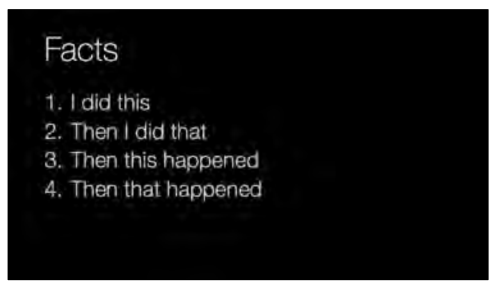

Consider this the next time someone is telling you a story of something that happened to them.

39



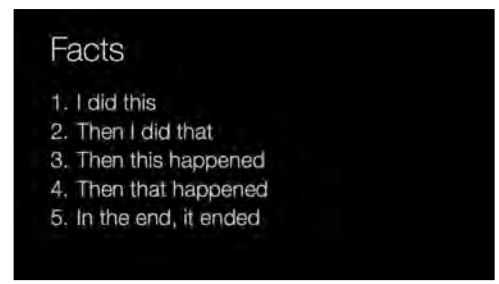

How does your attention span change

40




when they switch from listing off an itinerary

41

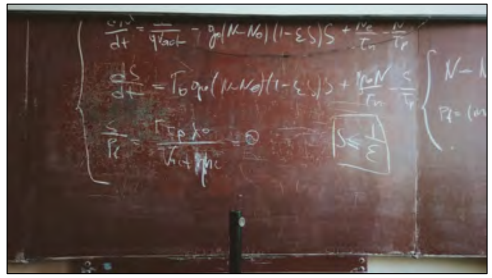



to an actual story about getting in trouble?

42

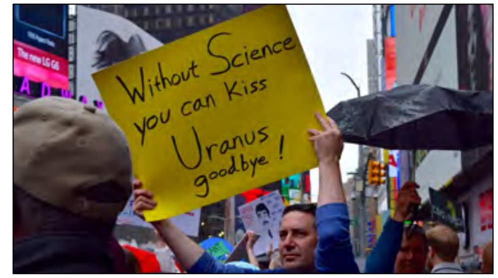

People love stories. Stories provide context and generate empathy. Stories reach the emotional parts of our minds rather than just the logical parts.





Scientifically-minded people often think that throwing facts, numbers, or statistics at people will win them allies.






Especially in America, this false assumption has led to a decline in people's enthusiasm for science as both a hobby, and a thing that needs to be funded.

We are just beginning to see a new movement of science communicators getting the public excited about science, the likes of which we haven't seen since the space race.

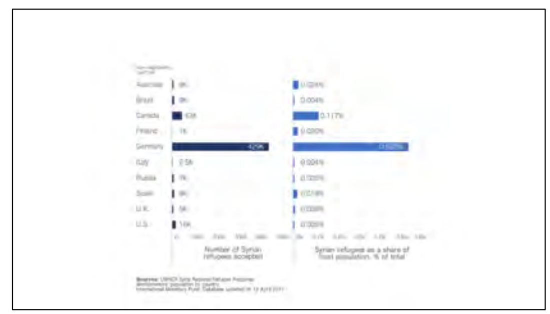


Understanding how to communicate a story takes practice. People sometimes have an easier time telling stories with words than with images. The two can go along together in the form of an image caption or a video narration.

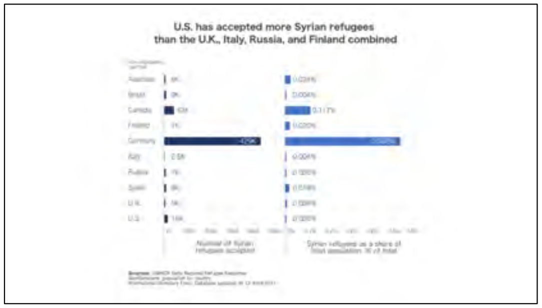



But beware – you still have to be careful with your choice of words when describing a visualization. Dr. Karahalios and her colleagues at the University of Illinois have done some incredibly fascinating research on the effect that titles have on the perceived message of visualizations.

It boils down to (these are my words, not hers) that regardless of what the visualization shows, people will remember what they are told over than what they see.



One of the studies she has done was to show this image to a bunch of participants, which shows the number of Syrian refugees accepted by each country. The obvious thing about this graph is that Germany has accepted overwhelmingly more refugees than any other country listed. Canada is 2$^{nd}$-most, and everything else seems pretty negligible, the U.S. is way down here with only 0.005% of the population.




But the title of the graph she had shown was "…"
It's probably not a conclusion you would have drawn yourself by looking at the raw data, but it is true if you add up the numbers in the left column.

When asked some time later  to recall the main message of this visualization…

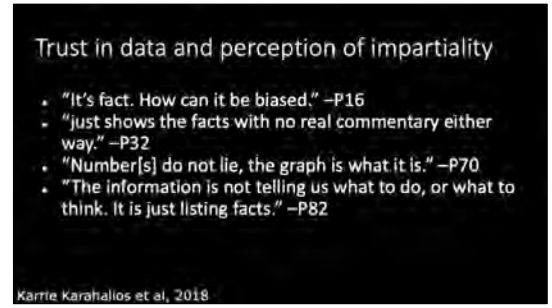
Although visualizations can certainly come with titles, figure captions, or video narration that tells the audience what to make of the visualization, the imagery should be able to tell a story even without text or narration.




Participants recalled that "…"

Think of any iconic photograph you've seen. The moment you look at it, you immediately understand something has led to the moment in the photograph. There's an implied history, decisions made that led to this point. Compare this dynamic photograph of the Washington Square Arch in Washington DC…






Later on when asked about potential biases in what they had seen, 83% of the participants were certain that there were none. Here are a selection of quotes from some of them.

People tend to believe things if they are backed up by data, citations, and visualizations, without doing their own investigation of the materials. As visualization creators, we have to be very careful about what we show, what context we put around it, and what we say about it.

to this photo of the same arch, also with a group of people. This photo is arguably more aesthetically pleasing, but does not tell a story like the previous photo.


Sometimes, you won't be able to maintain complete control over the contexts in which your visualization imagery will be seen. Much of our work is freely open to the public, and we're often surprised to find it on Reddit, in other documentaries, and on random corners of the Internet.

Images can get shared and re-posted, paper figures get moved and resized, findings get sensationalized by social media, videos get re-edited…

You will want to make an effort to have the imagery tell its story without relying on the text of your paper or a description. Just like the photo of protesters at the Washington Square Arch. On its own, the imagery might not say everything you want your audience to know, but it should be resilient against misinterpretation even when seen second-hand.

Visualization is subjective.

A problem with visualization in general is that it is subjective, on both the creator and receiver end. A dataset can be visualized in a variety of different ways.

And, different people looking at the same visualization can interpret it in different ways.

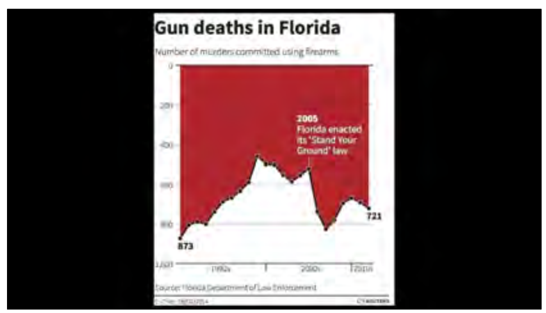
Visualization is subjective.
Visualization can be unintentionally biased.

Furthermore, visualizations can be biased. The underlying data can be biased as well, but that's another subject for another day.



Visualization is subjective.
Visualization can be unintentionally biased.
Visualization can be maliciously incorrect.

And, in the worst case scenario, visualization can outright present the data in a way that tells the exact opposite story of what the data says.

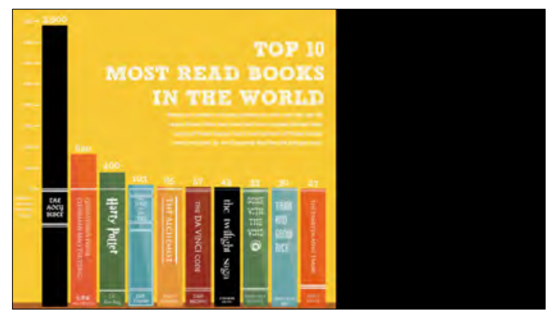



I'll give you a second to look at this visualization.

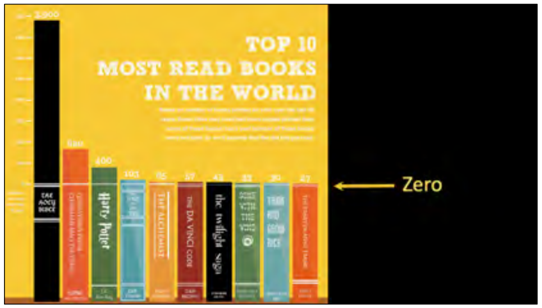
Can you tell what's wrong with it?

the y-axis is upside down, which  make gun deaths seem to decrease instead of increase.

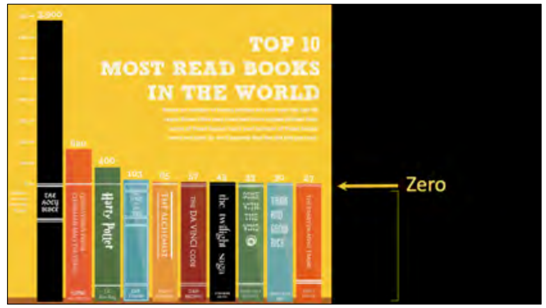
There's no way to know if this visualization was maliciously made to misinform people about gun violence, which is a politicized issue... Or if this was an honest mistake where someone just wanted to make a unique graph that looks cool... But, does it matter what the intent was, when the effect is the same, and this visualization gets posted on the internet and seen by thousands or hundreds of thousands of people?



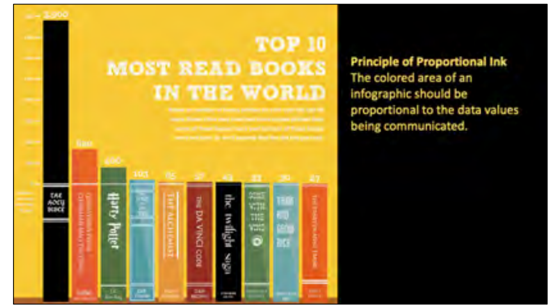

Here's another problematic visualization. Any ideas what's wrong with it? This one's a little trickier. The problem here is that



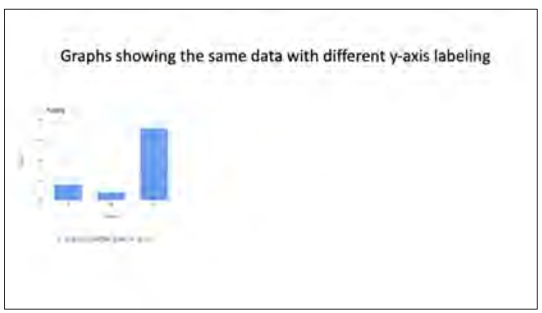

The zero line is all the way up here,




and everything in this space *looks* like it's data, but it's just filler content.



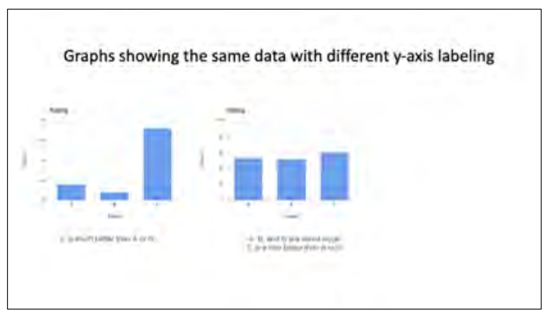

One common guideline in 2D visualization is called the Principle of Proportional Ink. This principle suggests that a visualization should always keep the size of a colored region proportional to the value it's representing.




Here's an abstract example that breaks down some other potential issues in visualization storytelling.

This is a bar chart with ratings of three objects, A, B, and C.

In this graph, we can assume that C is rated more highly and is much better than A and B. We may not notice that the y-axis doesn't start at 0, and that there's a very narrow range.




If we rescale the y-axis to go from 0-100%, A, B, and C are about equal. But, perhaps we were just looking at a subset of the full dataset.


The third graph introduces D, which has a much higher value than A, B, and C.

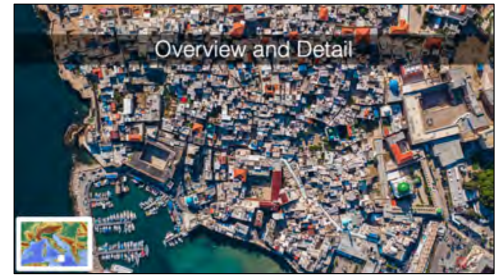
All of these graphs show the same data, in the same bar graph format, but they tell different stories. These graphs cannot be correctly interpreted if their context isn't properly shown.


If in order to tell an effective story, you do want to focus your visualization on a subset of your dataset, consider following an overview and detail model in order to maintain context, if there is a concern that the visualization may be misinterpreted without that context.

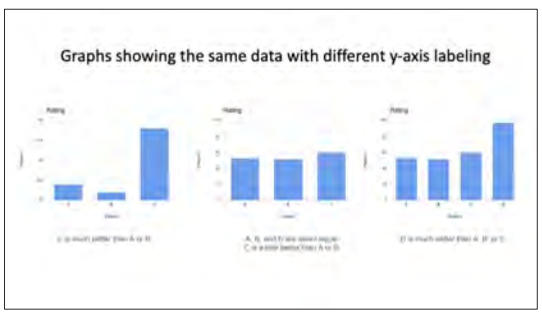



There are a number of different ways this can be done.



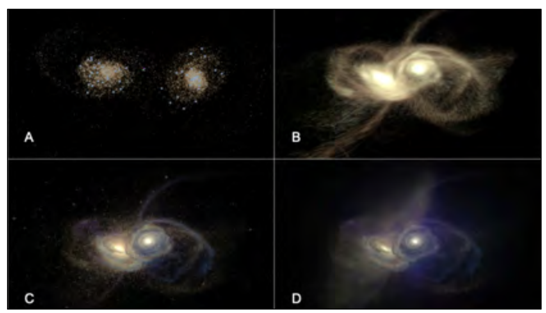





The same data can be visualized in a variety of different ways to tell a variety of different stories, and you can easily spin a dataset to tell the story that you want it to tell – whether on purpose or by accident. Some visualization choices ARE better than others.

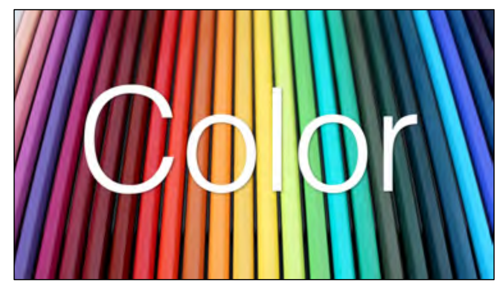
If there's one thing that you take away from this lesson, it's that there is no such thing as a purely objective visualization. Whether you're aware of it or not, any visualization involves making design decisions that alter its story.




Here's an example in a 3D scientific visualization. These four images show the exact same colliding galaxy data at the exact same moment. But they have significant differences.

In the top left image, we see only stars and emphasize new star growth.
In the top right image, we instead emphasize the directions in which the stars are moving.
In the bottom left, gas has been added, colored by the identity of its originating galaxy.
And in the bottom right, the gas has been colored by its destination.

Each of these representations uses the same data to tell a different narrative.




The last two galaxy images used color to tell a part of the story.

Most people don't realize how much of an impact color has on how we understand the world around us. But professional designers are hyper-aware of how color affects our moods and actions. Each color tends to be associated with different feelings, some positive and some negative.

For example, green might be associated with plant life and convey freshness and peace.



Or it might be associated with sickness and convey uneasiness.

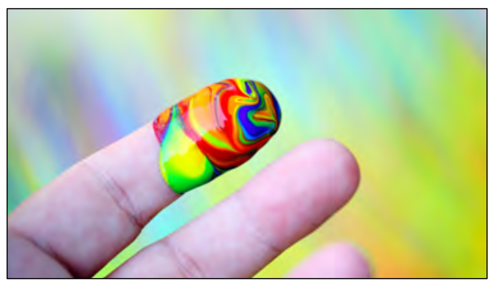



Red is often considered a color of passion - passionate anger because red is the color of fiery Hell,



or passionate romance because red blood makes our cheeks blush.





Colors don't have objective meaning, but context can sway the feelings your audience gets from them. Not only are individual colors meaningful, but so are color combinations.



Understanding how to choose a palette of multiple colors is an important design skill. A common tool for understanding relationships between colors is the classic color wheel.

One of the most common ways to select two colors that work well together is by choosing **complementary** colors, which create a visual balance - they are considered cheerful and bold, but can also be considered too obvious.

Blue and orange is a favorite combination of complementary colors of designers in Hollywood because it heightens the drama of the orange tones in the actors' skin.

It is very common.

Very very common.



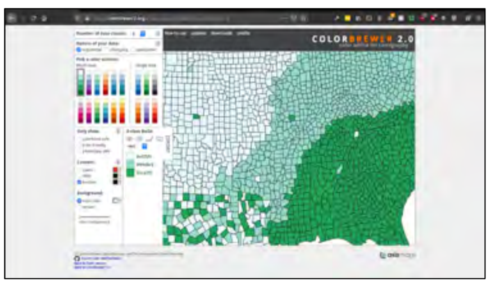
Another example of a color combination scheme uses **analogous** colors - these are colors that sit next to each other on the color wheel, such as red, orange and yellow. Together these make a soft non-offensive design that feel harmonious but can get visually boring.

Analogous colors divide the color wheel into relative psychological temperature where reds, oranges, and yellows are perceived as "warmer" than the opposite colors green, blue and purple. Often designers seek contrast in color temperature to create visual interest.




There are many tools online that can help you choose color palettes for various purposes, and I recommend playing with some of them to understand what combinations appeal to you. How are they related? What do they remind you of?






Color brewer is one of my favorites. Though it shows the colors on a map, you can ignore that part – it still offers up multiple options for color schemes.

It makes you think about whether the data you're coloring is sequential, diverging, or qualitative.

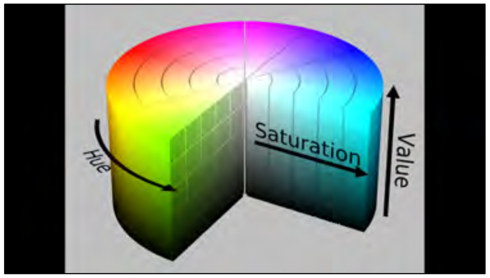
Sequential data goes from low to high values, where there's not necessarily a single data value that's more interesting than the others.




Diverging data has a defined value that separates the data in some way. Say, 0C separates below freezing from above freezing temperatures. It's a good idea to use two different categories of colors for each of those meanings to visually tell that story.




Qualitative data is categorical. Say, to separate planets that belong to the Kepler 186 system from those in the 185 and other systems.




I've mostly been focusing on one aspect of color so far which is called **hue**. You can think about hue as where on the rainbow the color falls.

But there are two other pieces to every color - saturation and value - which have an effect on psychology as well.

You might say **saturation** is how vibrant the color is.



Here is a picture of a baboon. Watch what happens when I turn up the saturation.



Saturated colors are exciting and attention-getting, but sometimes artificial and thought of as childish.



A de-saturated color is closer to gray. De-saturated colors can feel more natural and calm, but sometimes feel lifeless or decayed.



Back to neutral. A color's **brightness or value** is how light or dark it is.



Bright colors are light and optimistic but high values can feel harsh.

And as you remove brightness from a color it gets darker. Dark colors can feel mysterious but sometimes they make details hard to distinguish. That can be a good thing if you're trying to hide clutter.

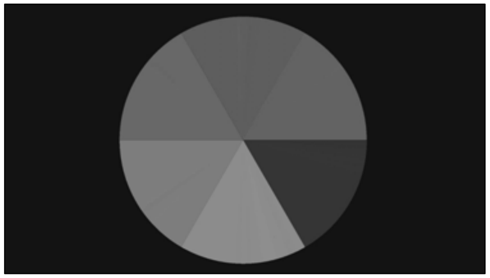


Contrast between color ranges is measured by the relative brightness of colors, and makes brighter colors in a design stand out clearly. Contrast can be used to draw your eyes to specific data points, and to de-emphasize others.



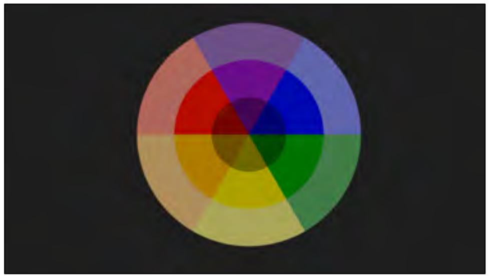
Different colors on the color wheel actually naturally have different brightnesses. For example, yellow is brighter than blue. In choosing colors for your visualization, it's often a good idea to





look at your visualization in grayscale to see if it still tells the same story to people who are colorblind and may have to rely on saturation and brightness more than hue to make sense of your images.





**Monochromatic** color schemes combine colors that have the same hue but different brightnesses - like in this combination of colors going from pink to deep red.




When you look at the world around you, pay attention to what colors are being chosen and analyze why they are effective.

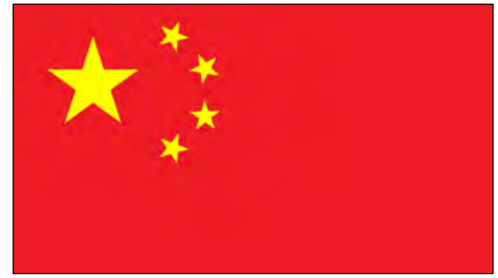

Why is it that comedy movies are marketed with bright colors and thrillers are marketed with dark colors? Why is it that summer clothing is saturated and winter clothing pale?

You may start to notice another thing designers understand well - different cultures assign different meaning to colors. Think about country flags for instance -



the flag of the United States uses the color red to symbolize bravery,





but the flag of China uses the color red to symbolize revolution.



In most cultures with cars, the colors red and green imply stop and go, but in places without roads that meaning does not exist.



Meanwhile, the combination of red and green in the Western world is highly associated with the Christian holiday of Christmas,



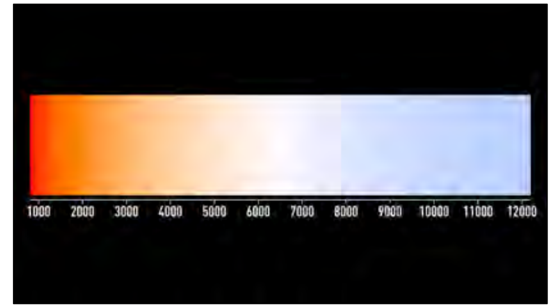
but in the Eastern world that same color combination is associated with Islam.


Color is also political.

Many people today are learning to be more cautious about the way they use the color words "black" and "white" – in majority white society, black has been treated as a scary color, because historically it was associated with a fear of the unfamiliar – specifically people with dark skin. This is why we are discouraged from using terms like "blacklisting" and "whitelisting". But the same is true for design. It may be your instinct to imply that something is negative by making it black – but this reinforces an antisocial stereotype and alienates members of your audience.

Of course not all color associations are quite so high-stakes – but understanding the cultural implications of your color choices can help you to avoid confusing - and possibly offending - your audience.

Temperature is an example of cultural confusion you might already be familiar with in the context of astronomy. Most people encounter temperature indicators in their home everyday - on sink faucets or stovetops - and these label low temperatures with the color blue and high temperatures with the color red. Ice is blue, and fire is red, so this makes intuitive sense.

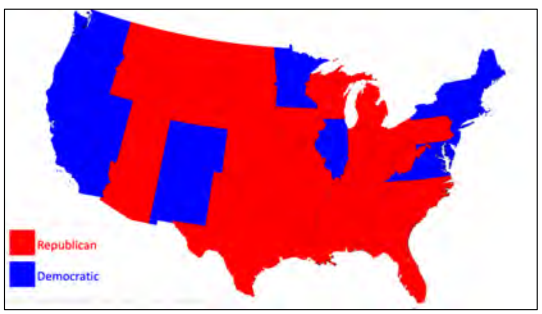



But this common use of color temperature becomes problematic when we start mapping color to extremely hot astronomical objects.

Stars are one kind of black body object, which emits color according to this color chart, where the red and blue are reversed from your faucet at home. As black body objects get hotter, they become less red and more blue.

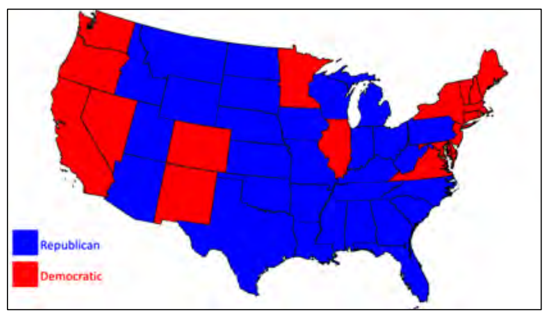
This kind of counterintuitive representation occurs frequently, and it's important to be aware of your audience's pre-existing biases and keep your audience in mind when making color choices that are common in your field but might be different from what your audience expects.

https://commons.wikimedia.org/wiki/File:Color_temperature_black_body_800-12200K.svg




If conventions exist in your field for certain colors having certain meanings, do use them where possible, because they act as a visual shorthand for quickly getting concepts across. But be cognizant of possibilities for misinterpretation.

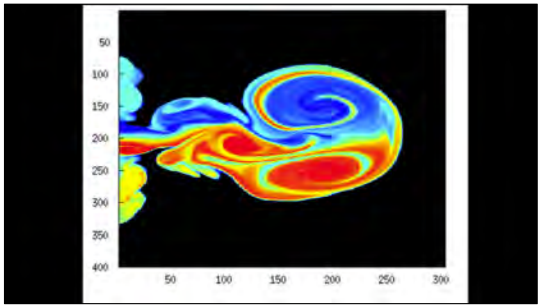
In the US, the Republican party is commonly referred to as "red" and the democratic party is "blue".




Audiences would certainly get confused if you were to swap the meaning of the colors.

It probably makes sense to color something like redshift red, and not, say, green.


This is not to say you should always do something just because "this is how it's always been done" – but if you're going to break convention, you should have a good reason for it, and put thought into what the implications are.

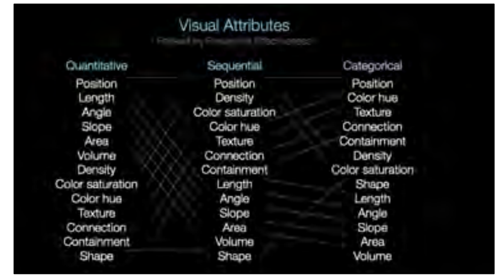



A good convention to break is the common use of every color in the rainbow for quantitative data.
Using many colors in a visualization is popular because it lets you highlight more features, but, it's perceptually confusing.

Looking at this image – can you tell whether blue represents a higher or lower value than red? You can't tell. Rainbow color palettes are inherently difficult to read. Using the rainbow makes sense for categorical data – but never for quantitative data.

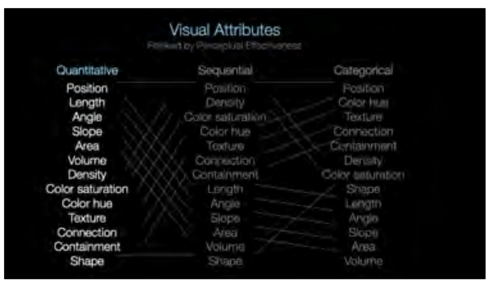

If it's used correctly, color is a great shorthand method of quickly getting across a concept in a visualization.





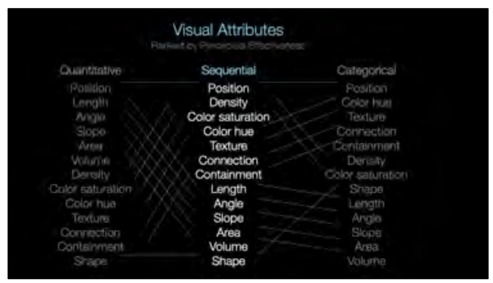

But it's just one of many visual attributes at your disposal. Depending on the type of data you are working with, different types of visual attributes are more effective and easier for humans to map to data values.





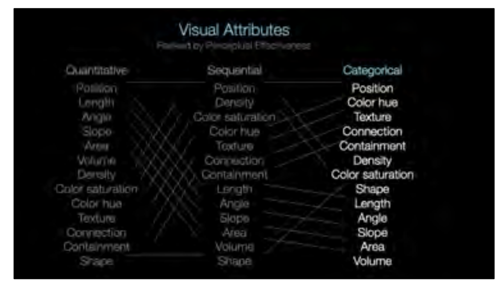

Quantitative data values are numerical.



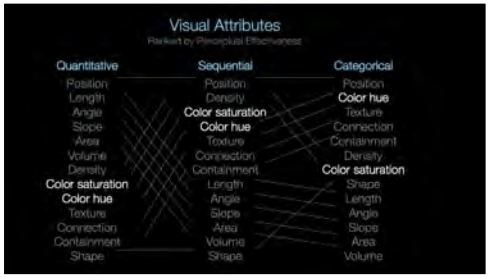



Sequential data refers to non-numerical, but ordered data values. For example, the sizes "small", "medium", and "large". You know that "small" is less than "medium", but not by how much.



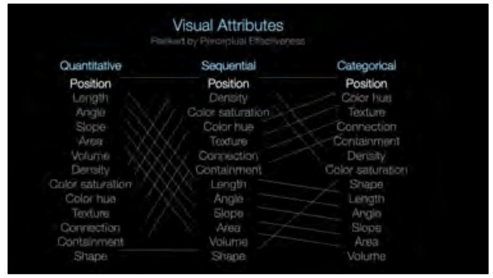



Categorical data refers to values where you can't directly establish an order. For example, galaxies - Is the Milky Way galaxy more or less than the Andromeda galaxy? It's neither, it's simply different.



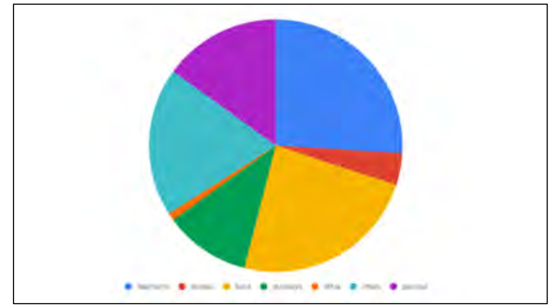

Notice that color, while it's powerful at communicating story and emotion, isn't great at communicating exact data values. So it's a good idea to use it simultaneously with other visual attributes.





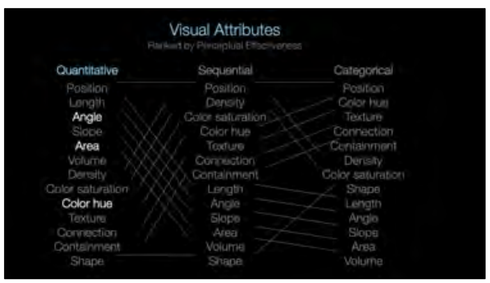

No matter the type of data, position is always the strongest visual cue. If you're making a scatterplot with X and Y axes, the position of a datapoint is the most obvious and easiest piece of information to understand.





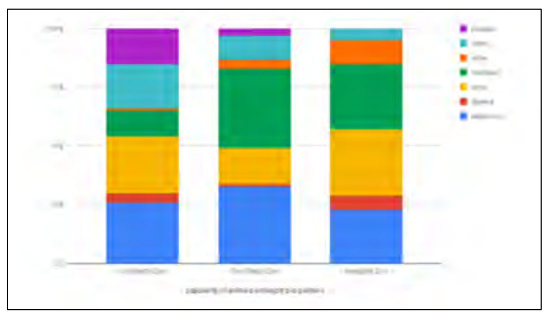
Notice that in a pie chart like this one, in addition to color, the visual attributes being utilized are angle and area.

In this visualization, it's a little hard to tell apart the largest slice of the pie chart. There are three likely contenders.

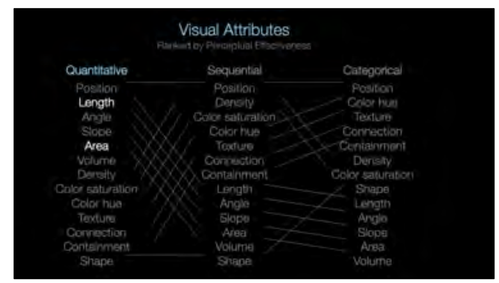



Here is where the visual attributes of our pie chart fall on the list. Not terrible, but we can do better.



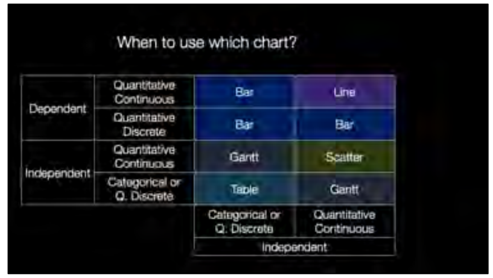


This visualization shows the same data visualized as a stacked bar chart instead of a pie chart. This still demonstrates parts of a whole, but reduces the visualization from 2 dimensions to 1 dimension. Now we're only looking at the height of the bar instead of width and height as well as angle.




You can see that length is a significantly stronger cue than area.

This little animation shows many of the visual attributes from the previous slides mapped to one dataset. We can change objects in the scene to have data-driven positions, or sizes, or colors, or shapes. In video, we can drive motion with data as well.



When to use which chart?

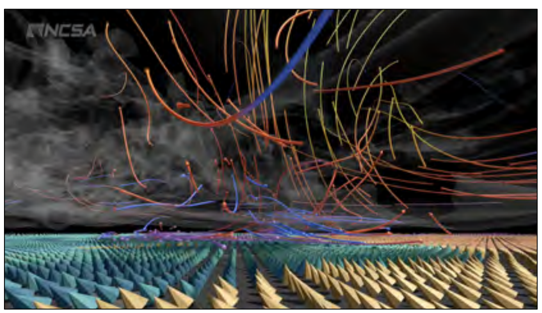
| | | Categorical or Q. Discrete (Independent) | Quantitative Continuous (Independent) |
|---|---|---|---|
| Dependent | Quantitative Continuous | Bar | Line |
| Dependent | Quantitative Discrete | Bar | Bar |
| Independent | Quantitative Continuous | Gantt | Scatter |
| Independent | Categorical or Q. Discrete | Table | Gantt |

Here is a handy cheat-sheet that you can reference which describes a few common visualization types, and when it's best to use what type of visualization. I recommend you take a screenshot of this for future reference.



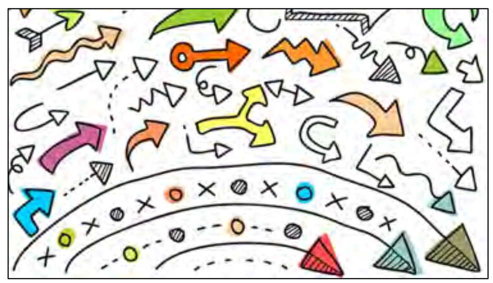

Mapping a concept like color or length to a data value is a type of abstraction. Abstraction is a powerful tool in visualization.






This visualization of a tornado is filled with abstractions. Arrows on the ground show ground wind speed. Orange-blue streamtubes trace the wind direction. Balls near the center of the storm show areas of high pressure. There is a faint gray shell around the storm, an outline of the storm cloud.

These abstractions show us scientific information much more clearly than recorded video of a tornado. A real world tornado is filled with dirt and fog, so it can be hard to see the inside of it, and we certainly can't tell by looking at it where changes in wind speed, temperature, and pressure occur.



In both information visualization and scientific visualizations, arrows are a powerful abstraction that can be used not only to show flow in a dataset, but can also be used to point out and emphasize an interesting feature in the data. I encourage you to use such visual labels to draw an audience's attention to the focal point of your narrative, especially in a complex visualization.



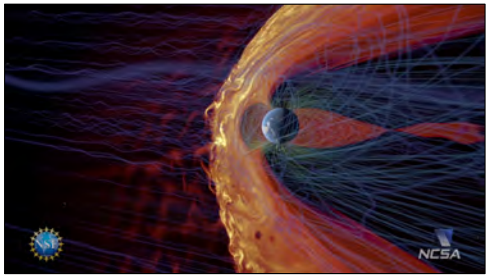

As you figure out the focal point of your narrative, be aware that you may face a dichotomy between the data's science narrative, and the outreach narrative.

The **science narrative** is the story of the research that created the data - what question were you asking as a researcher, and how does this data get us closer to answering it?

The **outreach narrative** is the story you are trying to tell to promote some broader understanding of science.

If the story you are telling is in your research publication, these two narratives may be the same, or similar. But if you are giving a presentation on your field of study to a general audience, the outreach narrative may focus on a very different aspect than the science narrative.

In some cases where you are collaborating with other researchers, the outreach story might not even be focusing on your field of research.



For instance, the science narrative of this visualization was about how solar plasma affects the Earth's magnetic field. It was focused on how the turbulence of the magnetic field causes it to break down. The data didn't have the Earth in it, because it wasn't concerned about effects ON earth, just around it.

But the outreach narrative was about how a burst of solar plasma could be hazardous to life on Earth. You can see in the visualization that all movement is centered around the Earth - which we added to the scene - and we kept the turbulent swirls of the magnetic field as a dynamic environment.



In order to tell a story well, it's helpful to understand the common traits in all stories. While finding new ways to use old storytelling techniques makes the stories feel original, we continue to use these same techniques that we have been using for millenia because they appeal to our common human psychology.



In short, we can transform a list of events into a story by introducing **characters** and **conflict.**



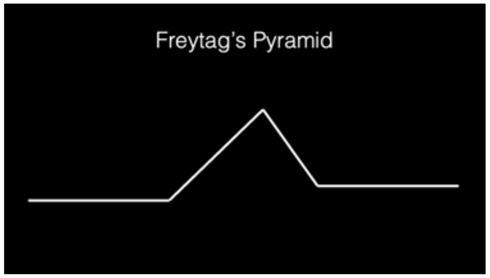
For instance - Little Red Riding Hood without the scary forest and the evil wolf just becomes a story about a girl bringing her grandma bread. There's no lesson to be learned about traveling safely.



Without his ghosts and the fear of failure, Ebenezer Scrooge never learns to be a better man.

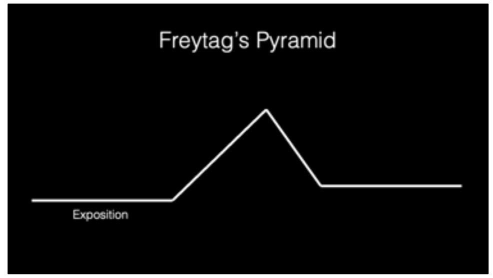


And without apartheid and political adversaries, Nelson Mandela wouldn't have found his hero's voice.



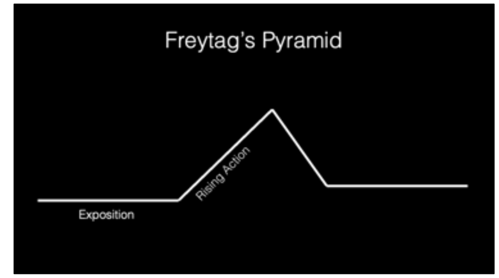

Freytag's pyramid sums this up nicely. This is a diagram that helps story writers remember the key elements to an effective story.





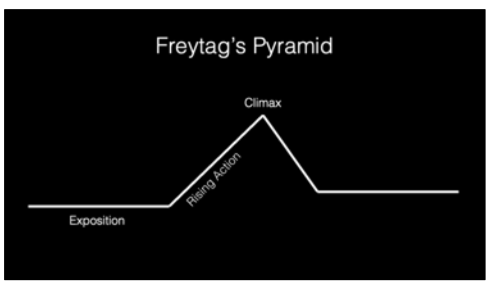

First you have the setup or **exposition** where you introduce the characters and the setting. In my 4-wheeler story, this is where I introduce my group of friends, and how the vehicle got stuck in the river in the first place.





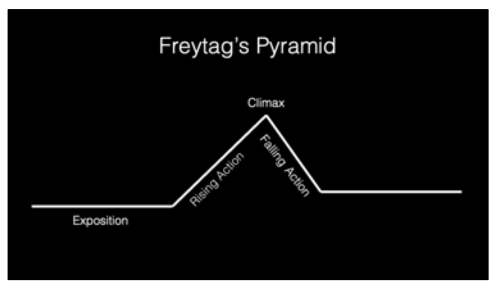

Second you have the **rising action** in which a conflict is introduced and made more severe. In my story, the conflict is that we try to pull the 4-wheeler out of the water, but situation keeps getting worse, with the popped truck tire, and the tractor almost falling into the water.





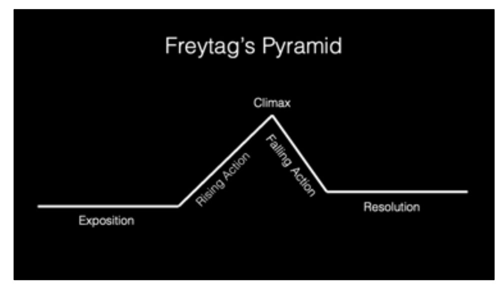

Third you have the climax






and the **falling action**, which is often brief,

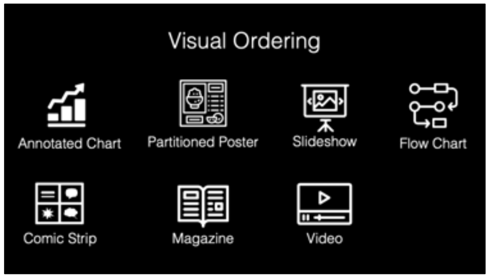





ending with the resolution.
The climax is where the main character overcomes the conflict and the falling action allows the conflict to be completely resolved. In my story, this is where we pull the tractor out of the water, ride out into the sunset, and live happily ever after.

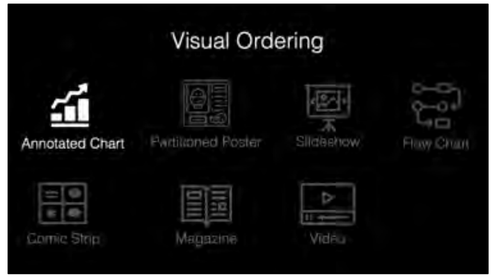
People who are new to story writing often neglect the resolution, or replace it with a cliffhanger like "to be continued." Audiences find stories without endings very unsatisfying, so it's highly recommended to include a thorough conclusion.

So a good story has a beginning, a middle, and an ending.




There's a number of different ways you can describe such a sequence visually.



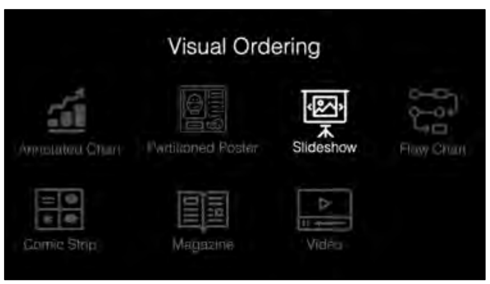



An annotated chart is a graph with labels that draws a user's attention and guides the story




A partitioned poster sets up a flow based on the poster layout. This can alternate between images, text, visualizations, and other content.






A slideshow is a series of "scenes", which can also have images, text, video, and anything else - shown to an audience one at a time, in a particular order.



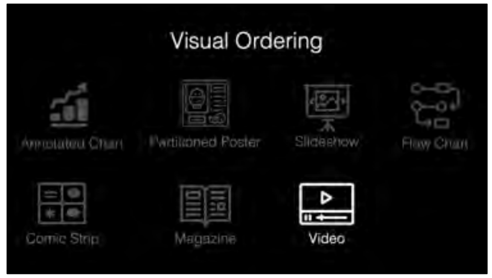

A flow chart is a layout of scenes, with boxes and arrows to guide the viewer




A comic strip is similar to a flow chart, laying out scenes all at once but without guiding arrows.


A magazine lets viewers bounce back and forth between looking at text and figures.



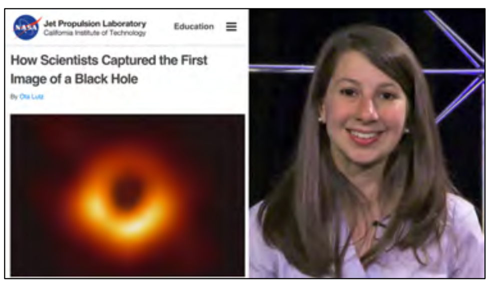



And a video is a one-directional series of images that viewers can play or pause.



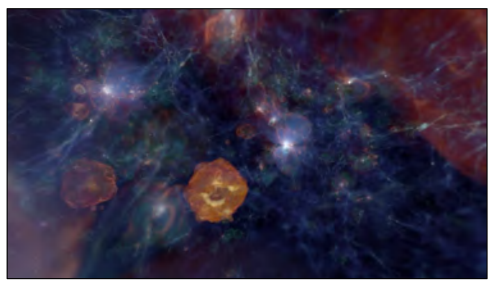
In science stories, identifying good characters and conflicts can be difficult. In text and video, you can get away with having the characters be the scientists doing the research, and the conflict is them trying to find an answer to a critical scientific problem.





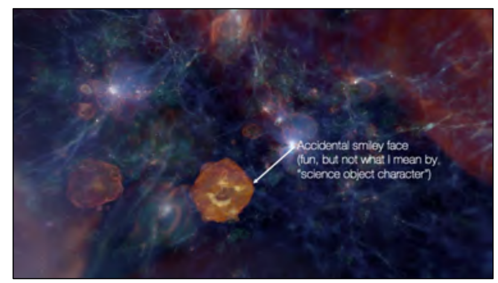

For instance, when the first image of a black hole was released in 2019, much of the public interest was around one young underdog scientist who, against all odds, achieved the impossible.





But when we talk about a visualization telling a story of its own, it's often not practical to talk about human characters. Instead, we think of the science objects - like galaxies or supernovae - as characters.





This visualization is about the formation of our Solar System, but starts out with a wider picture, showing the contextual molecular cloud before introducing our main character, the protoplanetary system. It's introduced as a seemingly small and insignificant piece of the Universe, that in the next few minutes of the documentary goes on to produce amazing things – it creates our planets, Earth, and eventually, life.


Scientific characters are abstract and harder for audiences to naturally connect with than other humans. This is why it can help to focus on a human-centered narrative, and find the elements of the science that are relevant to daily life on Earth.

Many audiences find astronomy interesting, but it can sometimes be difficult for them to know why they should care, why it's worth investing time, money, and effort into answering astronomical research questions. Tying the story back to humans is a good way to make the science relevant.

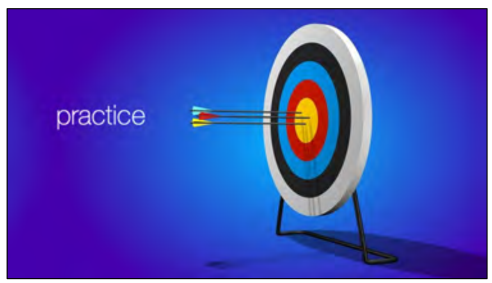

Science communication is a hot topic right now, but explaining new science concepts and discoveries to people is a hard problem. You might think you're being perfectly clear, but your message might not be coming across correctly. It's always a good idea to show your visualizations to a few different people outside of your immediate research group to see if they understand the concepts.






My team identified a few years ago that our audiences have a hard time telling apart the difference between photorealistic data visualizations, artistic illustrations, and photographs taken by telescopes. We've started showing previews of our documentary films to test audiences and handing out surveys with questions like this one to see what our audiences do and don't understand, while there's still time for us to make changes to the documentary scripts.

A large-scale formal approach like this is often not practical, but you can still get value out of informally asking 2 or 3 friends what they think of your visualization.




Given enough experience and practice, you'll start to build an intuition for storytelling with visualization. In human-centered design, it's important to always keep your audience at the core of every decision. Don't mold your story around your facts, mold your story around your audience.

There are many, many ways to visualize a dataset - showing different combinations of variables, showing a subset of the data, putting it in different context, using a different type of chart, picking different colors and visual attributes, etc. Experiment with different options, get feedback, and iterate whenever you can.

You'll be telling captivating stories in no time.