

Storytelling in Data Visualization

Akvil Sakhare

Abstract

This report is about the importance of storytelling for humans, its necessity and how to use storytelling key components and structures to leverage the power of storytelling for data communications.

It talks briefly about finding stories in the data and focuses on looking at examples of data storytelling and looking at what makes those data stories good.

Contents

Abstract

Programmed Storytelling

Necessity of Storytelling

Using Storytelling

Structure of the Story

Sketches and Storyboards

Storytelling Systems

- Logical Linearity
- Change with Time
- Contrast and Comparison
- Flow Diagrams
- Progressive depth
- Personalising Stories
- Textual Data Stories

Conclusion

References

Programmed Storytelling

In his book ‘Useful Fictions: Evolution, Anxiety and the origins of literature’, Michael Austin writes “*Every single human culture invests immense resources into producing and consuming stories and has done so since the dawn of time.*”^[1] There must be an evolutionary advantage to this. One of the interesting things Austin says is that fiction and stories can reduce and neutralise anxieties. Why would we want to do this?

If we think about it, our survival depends on anxiety. Say you are living in savanna and you hear strange sound from the grass, do you stay or run like crazy? The argument is that you should run every time. Even though most of the time it’s not the lion, it’s just a bird or squirrel making that sound. But if you don’t behave on a false positive response then one time out of a hundred you die. Successful narratives introduce and then resolve our anxiety. Through this we learn how to solve the problems and the idea that the challenges we face can be solved.

Another person who talks about this is Lisa Cron, in her ‘Wired for story: The writer’s guide to using brain science to hook readers from the very first sentence’ she says, “*Story, as it turns out, was more crucial to our evolution - more so than opposable thumbs. Opposable thumbs let us hang on; story told us what to hang on to. Story is what enabled us to imagine what might happen in the future, and so prepare for it.*”^[2] She goes on to say that one of the important things a story has to do is help us suspend our disbelief. Our brain has to feel like we are experiencing reality not like being told a story. It has to feel real.

A study she sights, which tells about a brain engrossed in a story activate the same parts of our brains that are activated when processing real experiences, when people actually process real sight, sound, tastes and movements in real life. Both Austin and Cron are talking about fictional storytelling, where as we are concerned about journalistic or documentary or data storytelling. So not advocating creation of fiction but humans are evolved to create stories that look and feel a certain way and the argument is that even documentary stories have to behave in a certain ways like fictional stories at least as far as the receiving brain perceives them.

“Stories like these aren’t about suspending disbelief, but instead about ‘actuating a belief.’”

- Kim Rees

Kim Rees argues that we shouldn’t tell stories with data but let data speak for themselves. But perhaps she was talking about the fictional implications of the term storytelling.^[3] Even when you let the data speak for itself it has to be in a structure that the audience’s brain will perceive it as a story, even if we aren’t injecting fiction, opinion or narrators influence on top of it.^[4]

Going back in time, many early stories were documentaries not fictions. Like when ancient man went out on a hunt and came back to report the results, he told stories. It was compelling narrative and it contained data. The story would explain what time the hunting party left, what direction they travelled, where they found their prey, how many miles they ran etc.^[4] It wasn’t a spreadsheet, it was always a story.

Humans evolved to produce and consume stories as survival mechanism, we need story, we expect story, we recognise what makes up a story.^[5] When communicating we have to think like a storyteller, especially when communicating data, which is inherently not anxiety producing and does not produce any emotion or evolutionary imperative. The data means nothing without the context, without the story form around it.^[6] So communicating data in story is necessary, natural and essential.

Necessity of Storytelling

“A story is 22 times more memorable than facts alone.” - Jennifer Aaker

Jennifer Aaker talks about studies about raising money for children with one brochure with statistics and another that has statistics and a story. Here is a story about how stories outway statistics. Statistics say that, “21% of children in US live below poverty line.” That might have an impact but it won’t be remembered three days from now.^[7] But instead the story goes like this, “Try to remember your elementary school. In my school we had a 100 kids in our class, broken into five classrooms with exactly twenty kids each. I remember my best friend ram bringing an apple every day of school. I know his family was poor but I don’t know why or how they spent their money that way. Last year I found out that 21% of kids live below poverty line in US. I didn’t grow up poor so I didn’t think of it as my first grade class. But since ram was poor I think about him. I think about how every single kid of his first grade class, all twenty of them, one in five of all of us in the school, might have been living in poverty, given that number.” Now after that simple and not very interesting story, it will still be remembered in three days.

One of the reasons, that a neutral word like table activate the language portions of our brains.^[8] But when we hear certain words like coffee, the portion of our brain that processes smells is also activated. Words like these activate parts of our brains involved in processing senses. The more different parts of the brain are activated, the more memorable, emotional and impactful a story is.

Statistics are like the neutral words, they only trigger the language portions of our brain.^[9] So weaving in statistics with story and including details like the apple ram brought in for his teacher everyday, activates more portions of the brains creating more sensory inputs and therefore sticker and more impactful content. So using visuals to activate the visual cortex and writing narratives that include other sensory inputs seems like a good idea.

All that said, you don’t always have to tell stories. The paper from stanford titled, ‘Narrative visualisation: telling stories with data’, talks about author driven vs reader driven data storytelling. As the paper says, *“As purely reader-driven approach has no prescribed ordering of images, no messaging, and a high degree of interactivity.”*

- Edward Segel and Jeffrey Heer

It can be said that purely reader-driven approaches are not data stories at all.^[10] They are dashboards, great for sharing data with experts or people who need to explore data directly. They don’t need to be told the narrative, they don’t need explanations, they are doing analysis.^[11] If that’s the audience then storytelling is not needed.

“Data stories appear to be most effective when they have constraint interactions at various checkpoint within the narrative, allowing the user to explore the data without veering too far from the intended narrative.” - Edward Segel and Jeffrey Heer

In other words they are not exploration tools but rather narrative experiences that provide context and direction not just a pile of numbers and charts.

Using Storytelling

Humans are wired for stories and have a certain expectations about what stories look like and how they are structured.^[2] Most basic form of story needs a beginning, middle and end. Maybe it has some nuances like the challenge the protagonist faces and climax that everything builds upto. Even while looking at a simple infographic, people are looking at the context for making sense of it.^[6] If they are interested they might read the body part. The point is that people might read the infographic backward but they will consume different parts of the story and in their heads they will reassemble it subconsciously using that story structure that they expect. The point is to have all the components that allow people to read it in linear story mode or to reconstruct it in their heads. Either way it's how people expect the information to be presented. Headlines, callouts, highlighted areas are great opportunities to introduce that key ingredient - anxiety.^[1] It means creating headlines that capture the story and include the key components of the story.

Clarity of Message

Knowing what you really what to say is a key to all good communications.^[12] Know what your data is saying and what your audience needs to hear. Knowing all these try to get to the single idea or come down to few basic ideas that you want to communicate. Know which is the most important, which is the least. Have important message and thoughts organized and tell the story around those concepts.

Looking for story in data

Data exploration is key to find insights from the data. Making charts from the data in spreadsheet and looking at trends help in developing ideas for the story. Converting data to percentages, change over time and looking at those trends also help to see that pattern of story the data is telling.

Structure of the Story

For thousands of years humans have communicated via stories. We know what stories look feel and sound like.^[2] Which means stories have predictable attributes. If you want to tell stories, even data stories, you have to be able to define what it means. Lisa Cron's book has a great definition for what story is, what those attributes are, she sums up the definition as follows,

"A story is how what happens affects someone who is trying to achieve what turns out to be a difficult goal, and how he or she changes as a result."

- Lisa Cron

To break that down in four key things, a story is what happens - that's the plot, affects someone - the protagonist, who is trying to achieve what turns out to be a difficult goal - story problem, and how he or she changes as a result. - the so what.^[4]

A story has beginning, middle and end and every story has a plot, protagonist, challenge and transformation which is the real so what.

"The other critical aspect of a story is anxiety. That story problem and that transformation is always experienced by the reader as anxiety and resolution of anxiety." - Michael Austin

The reason stories have a beginning, middle and end is because humans can't do two things at once. We can't multitask. Storytelling for thousands of years was purely oral tradition. Then came drawing on cave walls. In these cases man can't grasp more than one thing at a time. So stories are linear step by step things. They can't be anything else. The more linear a experience is the more it feels like a story to the audience.

What does linearity do?

Primarily linear storytelling, starting at the beginning and moving to a conclusion mimics to real life. We can relate to things in two dimensional time. The beginning sets the context for what we are about to hear. Middle tells us what's going on, allowing us to wander and explore the universe of ideas being presented and get introduced to the ideas that need resolving. And conclusion hopefully resolves those anxieties, teaching us how to solve problems and what transformation characters have to go to do so. A great story can take this logical flow most of the time. This doesn't mean that people can't experiment with the form. The point is that the audience knows what a story consists of, what it looks like and how it unfolds.^[2] In storytelling with data these are critical concepts that have to be included.

Sketches and Storyboards

Once the message is narrowed down, a structure for the story is figured out and all the analysis is done, it's time to sit down and make sure the ideas stick together before going too far for designing and developing anything.^[4] This step is about storyboarding the ideas and sketching out the concepts simply and quickly.

The few goals here are,^[4]

Vet and test: Vet and test the ideas that are there so far. Experimentation is a friend. Try the ideas and see what doesn't fail.

Structure: Come up with story structure that will communicate the data the best way possible. What are order, the information hierarchy and pace.

Speed: And in both of these move quickly. Failing and adjusting and failing again until there is something that works. Move quickly, fail fast and often, sketch, don't strive for perfection.

Sketching and storyboarding is about trial and error. Working using hands works well for trial and rather than a computer.^[13] Pen and paper or whiteboard, or any other analog ways that best allows you to truly sketch.^[14]

Storyboarding is a part where we figure out the structure and order of the story. Assuming that we are creating a linear story, the idea is to figure out what comes first and what's in that piece of the story, then what happens and then what after that.^[15] We have a series of panels and the idea is to fill in the panels with the basic idea for each story segment. This can be done in a lot of ways. Comic strip panels can be used, if the story can branch in different ways that can be shown using arrows and explained why that path might be taken. Posted notes or indexed cards can be used to put them visually together on a table so that we can see exactly how the story is going to come to life.

Storytelling Systems: Logical Linearity

The Shapes of Stories

Humans can do one thing at a time. As much we like to think that we can multitask, it's true that we can only perceive one thing at a time. Giving one experience at a time linear manner makes sense. It can be referred as following a linear logic. It's the simplest of ideas we learned while learning how to write essays in school.^[16] We always think of stories as linear experiences.

Kurt Vonnegut on the Shapes of Stories explains what stories look and feel like.

The pink line, on the y-axis, at the bottom is sickness and poverty and on the top there is wealth and good health. It's called the G-I axis, the good fortune to ill fortune axis. The left side starts just above the origin because, why start with a depressing person. The x-axis is Beginning to End or Electricity as he calls it, the line gets up and falls down as somebody gets into trouble in the story. Then it gets bad and then the person gets out of the trouble again. He calls it man in a hole, you start with a average day, you get into trouble and you come out again.^[17]

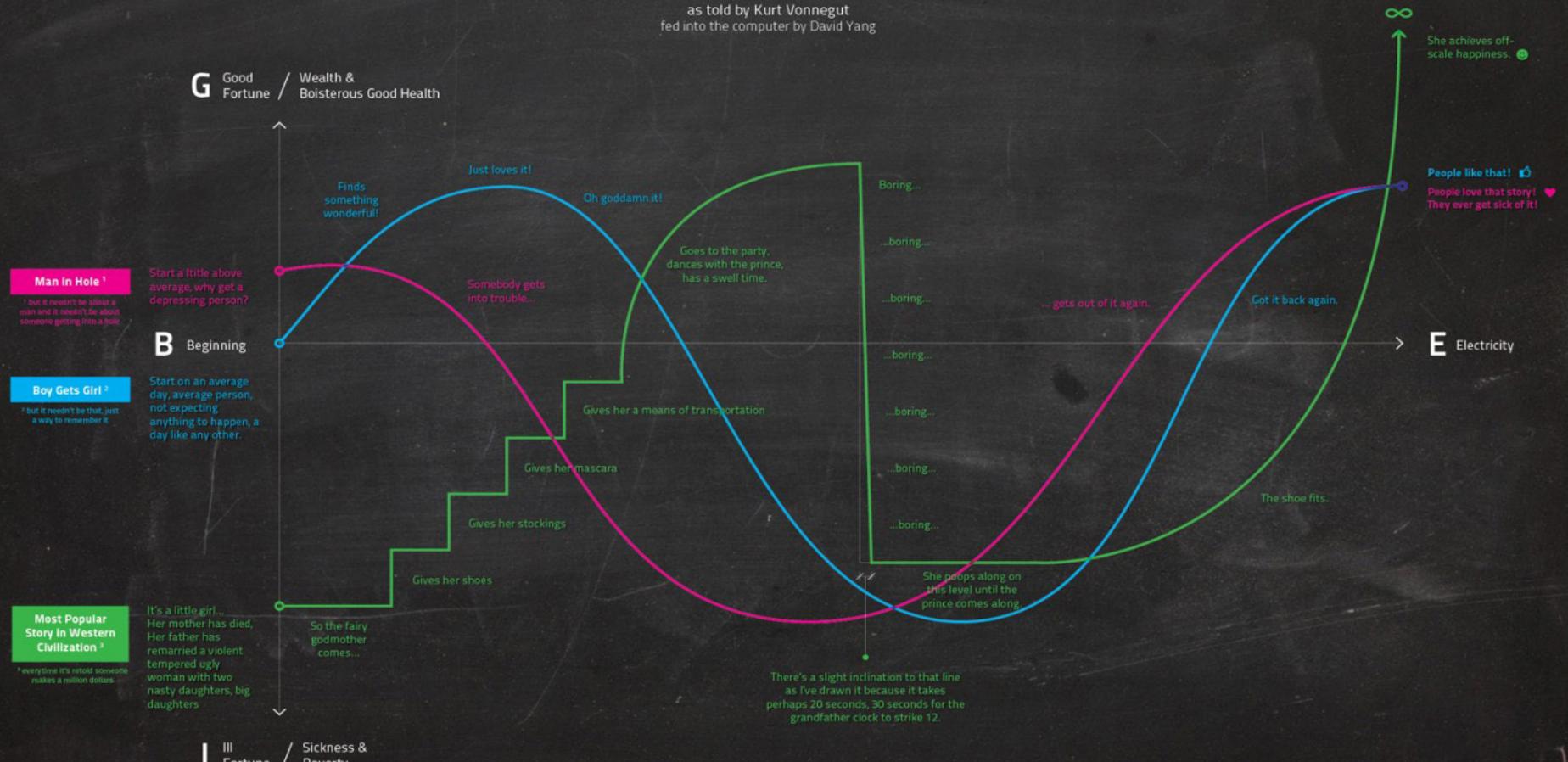
The next story is boy meets girl, that's the blue line. Boy starts on a average day as average person, not expecting anything to happen and then he finds something wonderful. He loves it, then something terrible happens, then towards the end the blue line goes up as he gets the girl back again.

When looking at data it often makes sense to find a time driven storyline and tell that linear story to the audience.^[4] It also ties back to the basic story structure where there is a beginning, middle and end. You describe the beginning, you set the context, establish the scene, introduce the characters etc. Then the story flows logically from one component to the next ultimately ending in some sort of conclusion. When same process is followed in data storytelling, yes it's communicating information in numbers but rather bring that data together into a compelling, logical, linear narrative.

Linear stories connect with people. Using a linear format can be a very effective way of turning data into a compelling narrative that will resonate with the human brain which has evolved to hear and learn things in this way.^[2]

Simple Shapes of Stories

as told by Kurt Vonnegut
fed into the computer by David Yang



<http://youtu.be/oP3c1h8vZQ>

Kurt Vonnegut, the Shapes of Stories
Source: visual.ly/kurt-vonnegut-shapes-stories

Storytelling Systems: Change with Time

Napoleon March to and from Russia

Stories allow us to understand and experience the world around us and through that experience we learn how to survive. Since the human experience is entirely about putting one foot in front of the other in a linear way and experience what we experience through that lens of time, it makes sense that stories are linear, time dependent and time driven.^[4] So one of the common data stories is data changing over time. First this happened, then this, then that, etc. So when looking at data it makes sense to find and tell these time driven stories.

It isn't a visualisation discussion until someone break out the minard.

It is an infographic by Charles Minard. It was created in the 1800's and it's telling the story of napoleon's march on moscow. Edward Tufte, the grandfather of data visualisation and information visualisation, holds this as one of the greatest infographics of all time and almost all the interaction design professors at idc have this in one of their lecture slides.

The napoleon's march on moscow starts on the left side, the thickness of the brown line shows the size of the army, around four hundred thousand men and as he marches eastward, you can see the men losing and the army shrinking. As they reach moscow, their army is quarter of the size it was, they are there for a long time and finally they decide to turn around and go home. They are losing more and more men on the way back home.

There is a lot of data here. There is information on where they started and ended, there is the size of the army in the thickness of the line. There is also temperature information down below so you see interesting stories such as when they cross that river about one third of the way, the black lines cuts in half as the river was so cold and they were crossing that river on foot.^[4]

This is the story of change over time even though time is not mentioned. There are no dates, months or clock listed out here. But we can imagine the months of combat and how difficult time period this was for people. It's essential component to the story.

Stories in the end are almost always about change over time, even when it's implied rather than explicit. So data stories are also often about that. This is a go to mechanism for data storytelling. Of course, we don't always have data about things changing over time which will push us to of the other story mechanisms.

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres dans les zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Fezensac, de Chambray et le journal intitulé de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow et se rejoignirent vers Orsha et Wileïsk, avaient toujours marché avec l'armée.

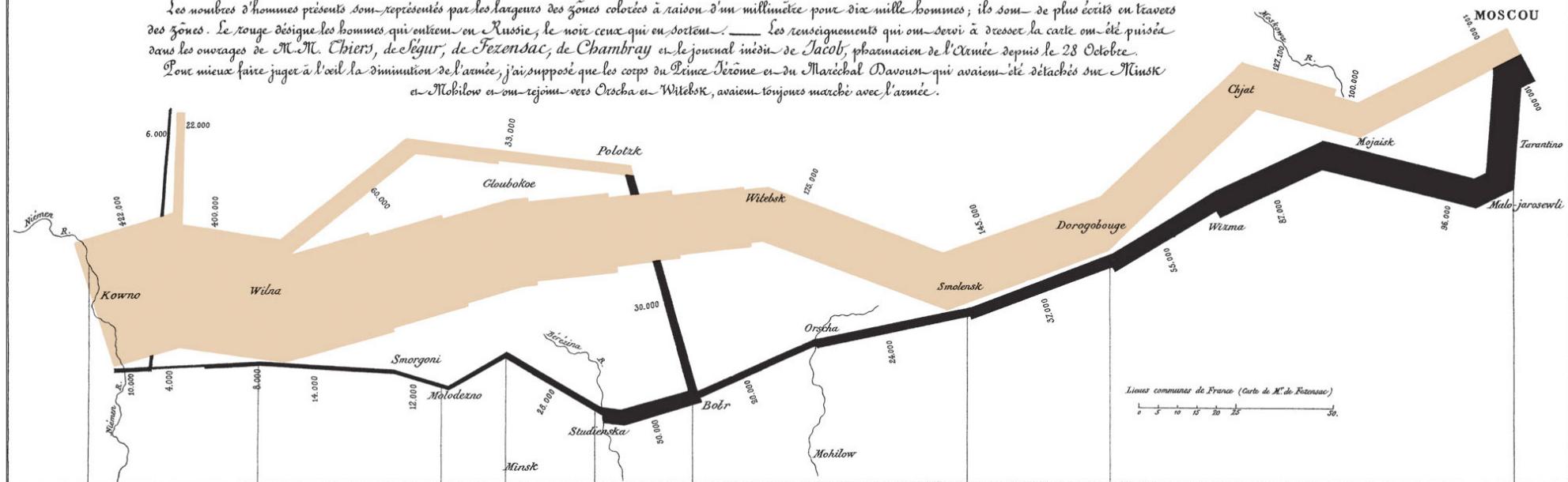
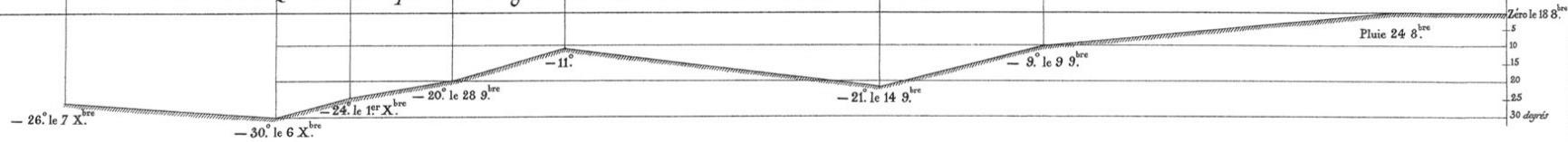


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop
le Niemen gelé.



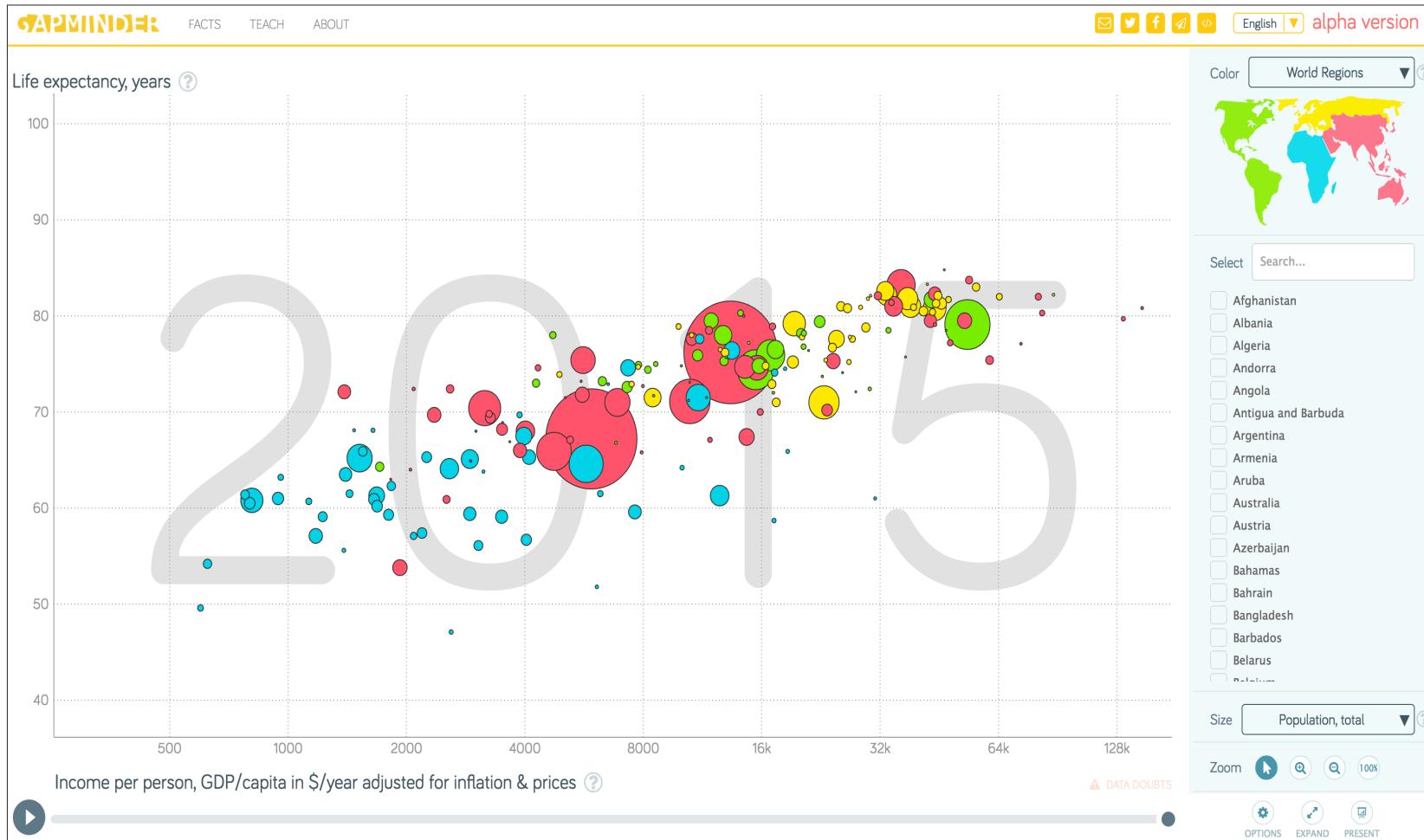
Autog. par Regnier, 8. Pas. S^e Marie. S^e Gravé à Paris.

Imp. Lith. Regnier et Bourdet.

Minard's "Napoleon March to and from Russia, 1812-1813"

Source: en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Minard.png

Gapminder



Gapminder,
Source: www.gapminder.org/tools/#_chart-type=bubbles

Gapminder by Prof. Hans Rosling is a good example of use of motion to show change over time. This is data on income and life expectancy over the years from 1800 until the modern day. Each one of the bubbles represents a country. They all start from lower left hand corner and as time goes by the western democracies, in yellow, start to float up towards right. They are getting richer and healthier, they are living longer.

The countries in red, the asian countries, are lingering behind but then the giant bubble of china starts shooting up towards right. A lot of african countries are still lingering far to the left and down as compared to the rest of the world. Such a compelling story, told so well, could not have been told without the animation.

Storytelling Systems: Contrast and Comparison

Selfiecity

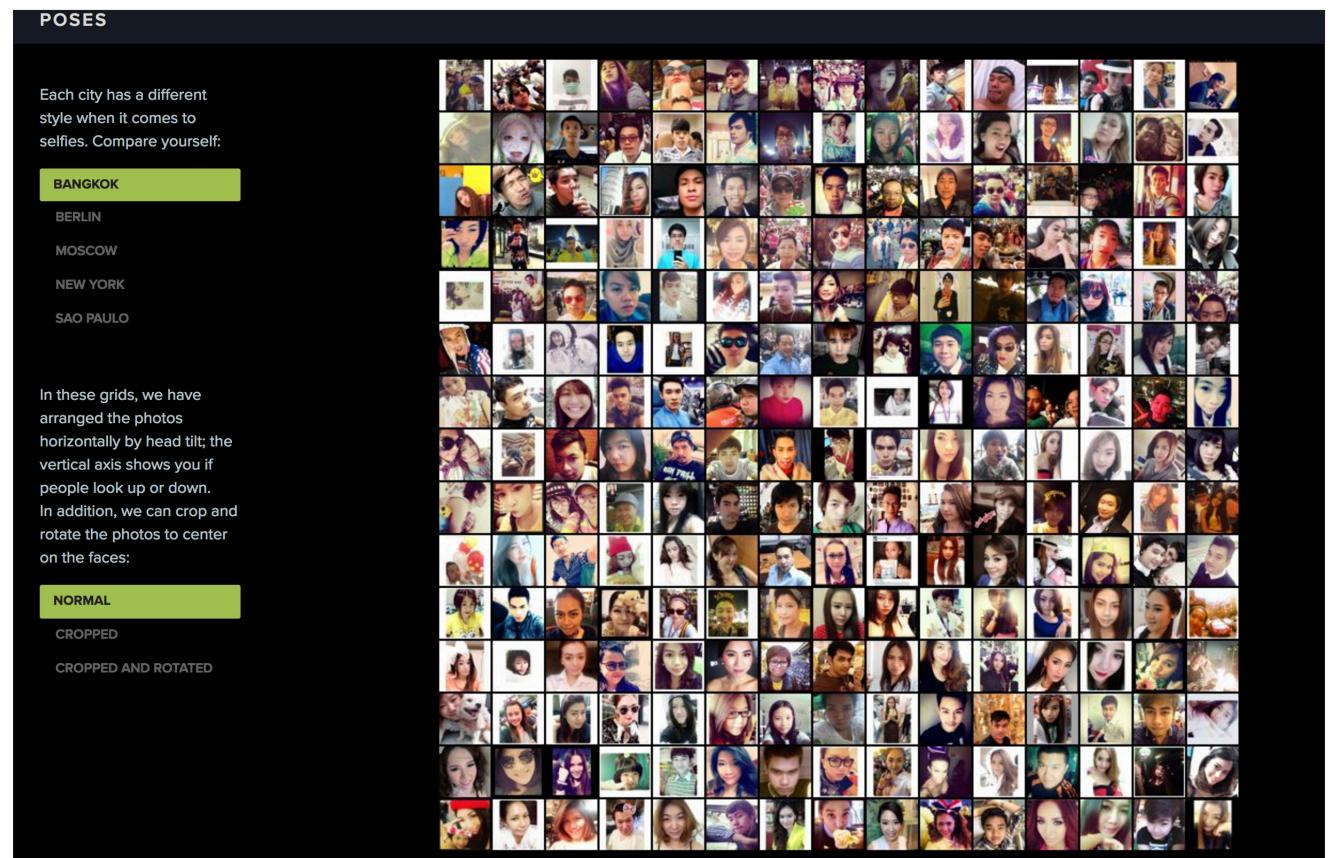
Stories are at their heart almost always about comparing and contrasting things and often about how things connect to each other because the key component of every story is about transformation of a key character and what he or she goes through to achieve a goal. We are always evaluating at least a binary choice, comparing what-if's and looking at the whys. What decisions were made along the way, what connections can we make about those decisions and outcomes and what connections did our character make to end up where he is.

In documentary or journalistic stories, we are often confronted with comparisons and contrast because the story is about something that hasn't happened yet or is happening and we have to make a choice to change the outcome. Like in 'An Inconvenient Truth', Al Gore spends a lot of time explaining the data about climate change. Essentially making comparisons between where the world is headed now and where it could go. If we make a certain changes as to how we operate as a species. Data itself is often about contrasts. Comparing different data sets with each other, comparing different variables within the data etc. Some data really lends itself to telling a this or that or us vs them or good vs bad type of story.

*Selfie comparison in different cities from Selfiecity,
Source: selfiecity.net/#imageplots*

When telling a story of comparisons, it's really important that you are sure that you are comparing apples to apples. That means carefully choosing what you compare and the methodology used to be sure your data really is comparable. This is where you think about choosing your subjects carefully. To compare things, you have to make them comparable.

Selfiecity compares selfies from 5 cities of the world and measure them in different ways. First thing you can do is look at different collection of selfies. You can see selfies from a city and compare how it looks different than selfies from any other city.



GENDER AND AGE PROFILES PER CITY

Case by case inspection of photos can reveal a lot of detail, but it is difficult to quantify the patterns observed.

"Is it just me, or do São Paulo women actually tilt their heads more? Do New Yorkers or Berliners look older?"

In order to answer these questions, and supplementarily automatic face analysis with human we had thousands of photos inspected by Me workers, who estimated age and gender of the photos. Here are the results:



If you scroll down the page, you can do some data driven comparisons.

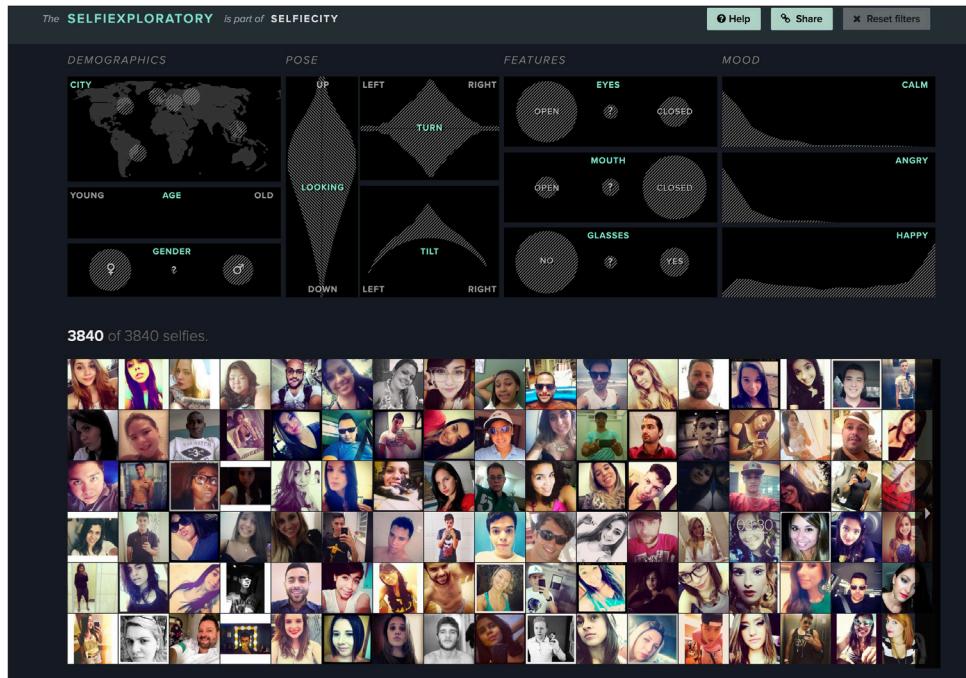
In this case look at gender and age profiles per city. You can see that people taking selfies in Bangkok are a bit younger, average age is 20.3 amongst women and 22.7 amongst men, than moscow where the average age is 23.3 and 25.7. A lot more women take selfies in moscow and very few men vs bangkok where it is evenly distributed etc. We are comparing cities to each other, genders to each other, age groups to each other and how they do this particular task, in this case it's taking selfies.

SMILE DISTRIBUTIONS BY GENDER AND CITY

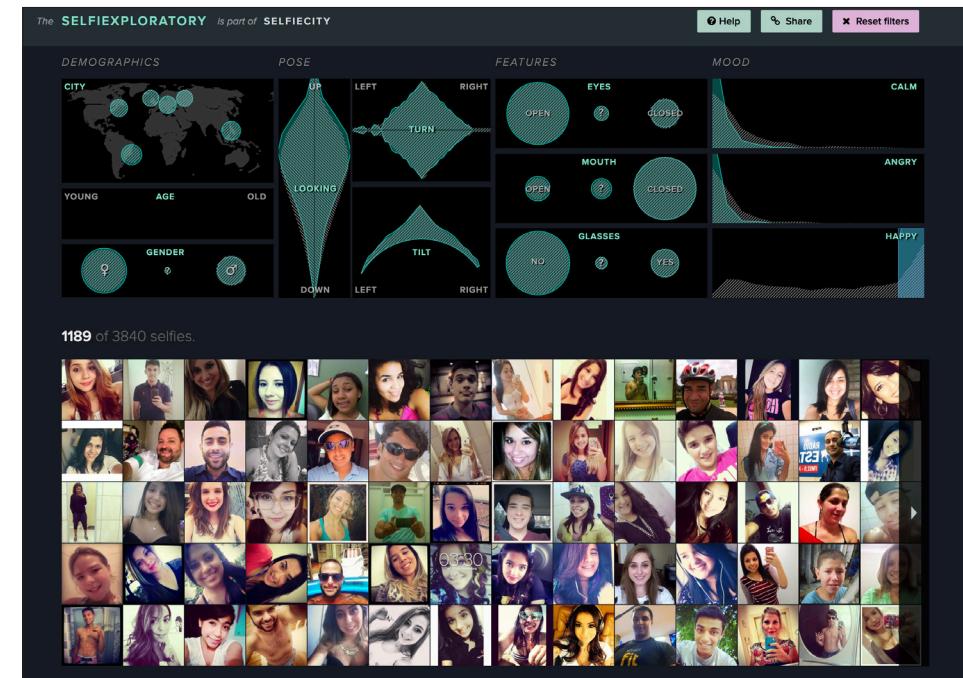
We can also determine the facial expressions of the selfies in a city — who smiles the most, and who has more reserved looks?



Or I can compare them based on their smileness, if that's a word. Bangkok, which is known as land of smiles, has a lot of people who smile a lot in their selfies. As opposed to moscow where there are not a lot of smiles in selfies.



There is a tool called selfie exploratory, which is interactive data experience. There is geography, age, gender, pose - which way they are looking and various other features like if their eyes, mouth are open or closed etc. You can see all the images down there and you can filter the charts above and see those selfies down.



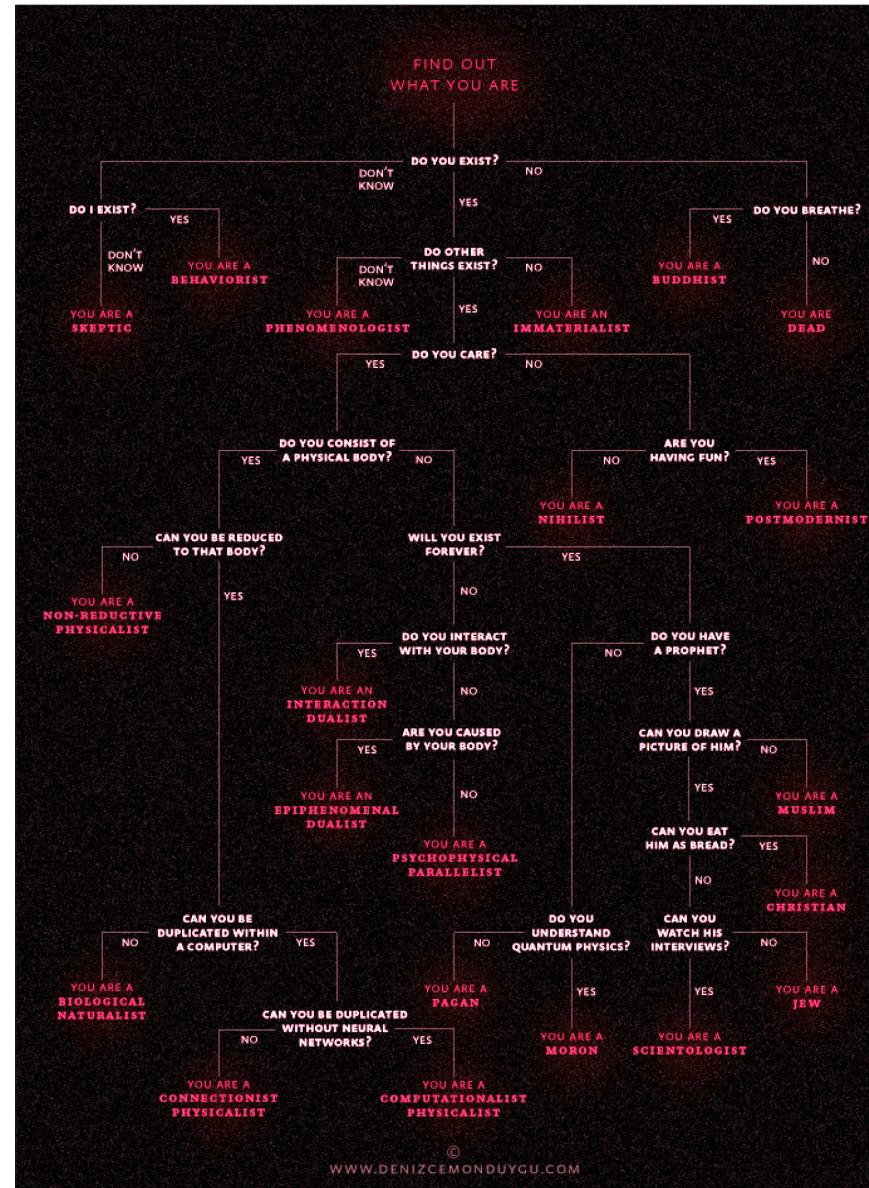
It also slows all the other elements in the chart. Like you can see the rio bubble getting bigger and moscow bubble getting smaller as there are fewer happy people in moscow selfies than there are in rio selfies. I'm able to do all these comparisons as I interact with this experience.

Comparing and contrasting are among the most common data stories out there. There are many ways to look at comparisons, depending on how broadly or closely we compare things to what perspective you take to slice into the data. Even if you aren't focused on comparisons, people might go there, so it's essential to be conscious about how the data ends up being comparable to other things and what story is being told even unintentionally.

Storytelling Systems: Flow Diagrams

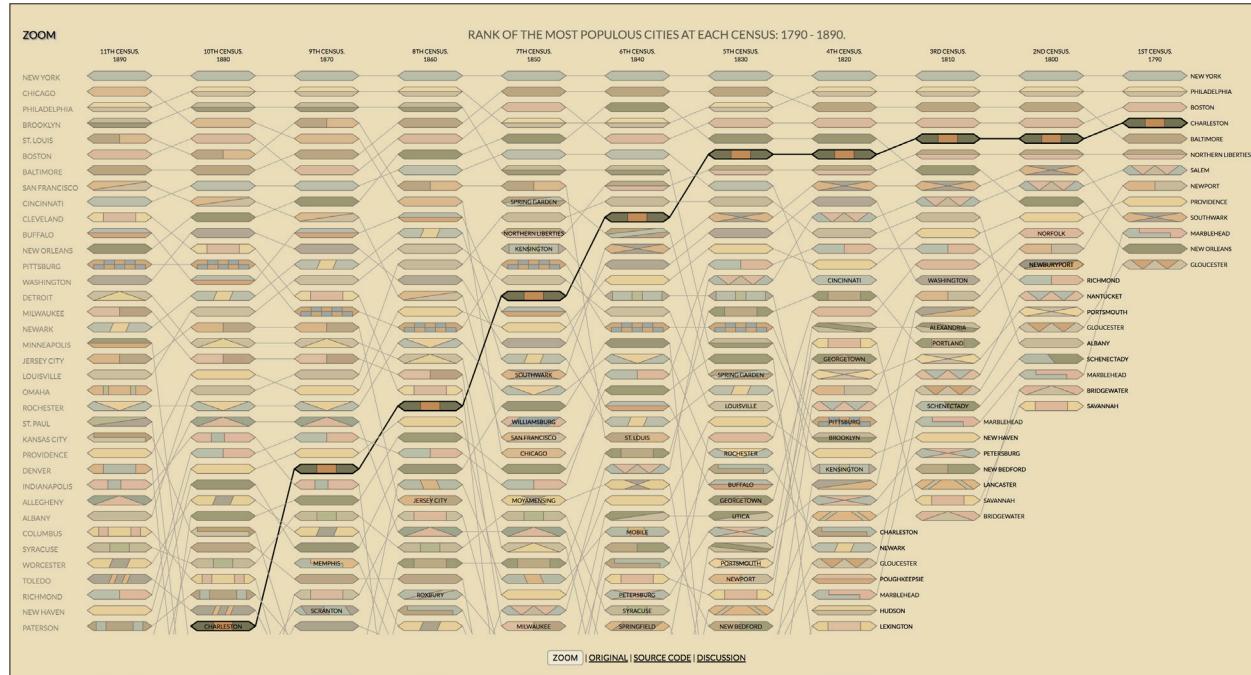
Find out what you are

Flow diagrams have been used for a very long time to visualise things like oarc charts, database diagrams and business processes. They are so mainstream these days that they have becomes a category of memes. Some are rich and sense flows with a lot of information to digest. Where only a part of story is read, so you are choosing the story yourself.^[18]



*Find out what you are,
Source: visual.ly/find-out-what-
you-are*

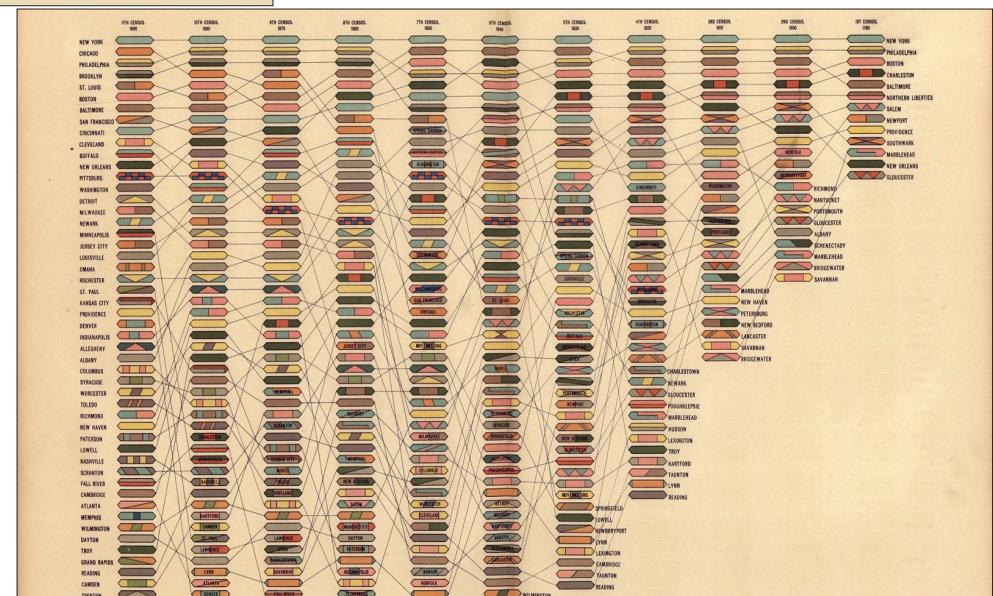
Rank of the most populous cities at each census: 1790-1890



This is a recreation of a classic visual that appeared in the statistical atlas of the United States in 1890. It's essentially the same image where you can hover and see, for instance, Charleston who disappears after mid 1870 or Newport which drops off the list in early 1800's. Essentially it is a flow diagram but not the traditional flow diagram. This is interesting as it goes against some of the norms, for instance, the older years are on the right and years flow to the left. The interactivity in this makes it a lot easier to follow. Whereas reading the original is very difficult.

If data describes a flow from one thing to the next or some sort of hierarchy or if the metaphor of flowing process can be used, using a flow diagram might be a great way to tell the data story.

Rank of the most populous cities at each census: 1790-1890,
Source: vallandingham.me/census_bump/



Rank of the most populous cities at each census: 1790-1890,
Source: vintagevisualizations.com/products/rank-of-the-most-populous-cities-at-each-census-1790-1890

Storytelling Systems: Progressive Depth

Out of Sight, Out of Mind.

Some stories are simple and they stay simple. The point isn't that everything has to be simple but presentation of simple higher level information to get the audience self select in the deeper and more complex content based on their interest in the subject matter. In data stories this leads to the idea that you can create data stories that offer progressive depth. Rather than simple, start off with very high level intro and then offer deeper levels of content for those who choose to experience it. In 'Narrative visualisation: telling stories with data', this idea is referred as martini glass structure.^[10] As is said in the paper, "Following a tight narrative path early on, that's the stem of the glass and then opening up later, for free exploration, the body of the glass." It's about adding progressive depth, based on audience's desire to get that depth. If the audience might seek depth in the content this is a good way to offer it to them.

Out of Sight, Out of Mind is a visualisation of drone strikes. It starts with telling a story. It has a very linear storytelling structure. It starts with animated and then interactive graphic. After being done with the introduction, it starts showing, one after the other, individual drone strikes over time. It uses the arced line to show the strike. At moments, it pauses and text comes up to tell a headline about that moment in time.

After the animation is finished, it starts with a fully interactive experience. We can hover over each drone strikes and details pop up about that strike. Clicking the Victims tab flips the experience on its side and it can be viewed from a different angle. Then you can again hover over and see details.

Since 2004, drone strikes have killed an estimated 3,341 people in Pakistan.



Out of Sight, Out of Mind. Page One,
Source: <http://drones.pitchinteractive.com/>

**Less than 2% of the victims are high-profile targets.
The rest are civilians, children and alleged combatants.**



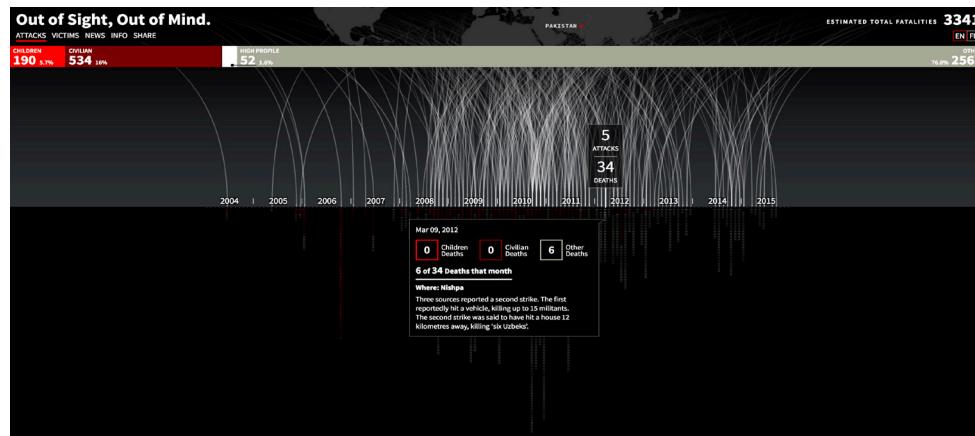
Out of Sight, Out of Mind. Page Two,
Source: <http://drones.pitchinteractive.com/>

This is the story of every known drone strike and victim in Pakistan.

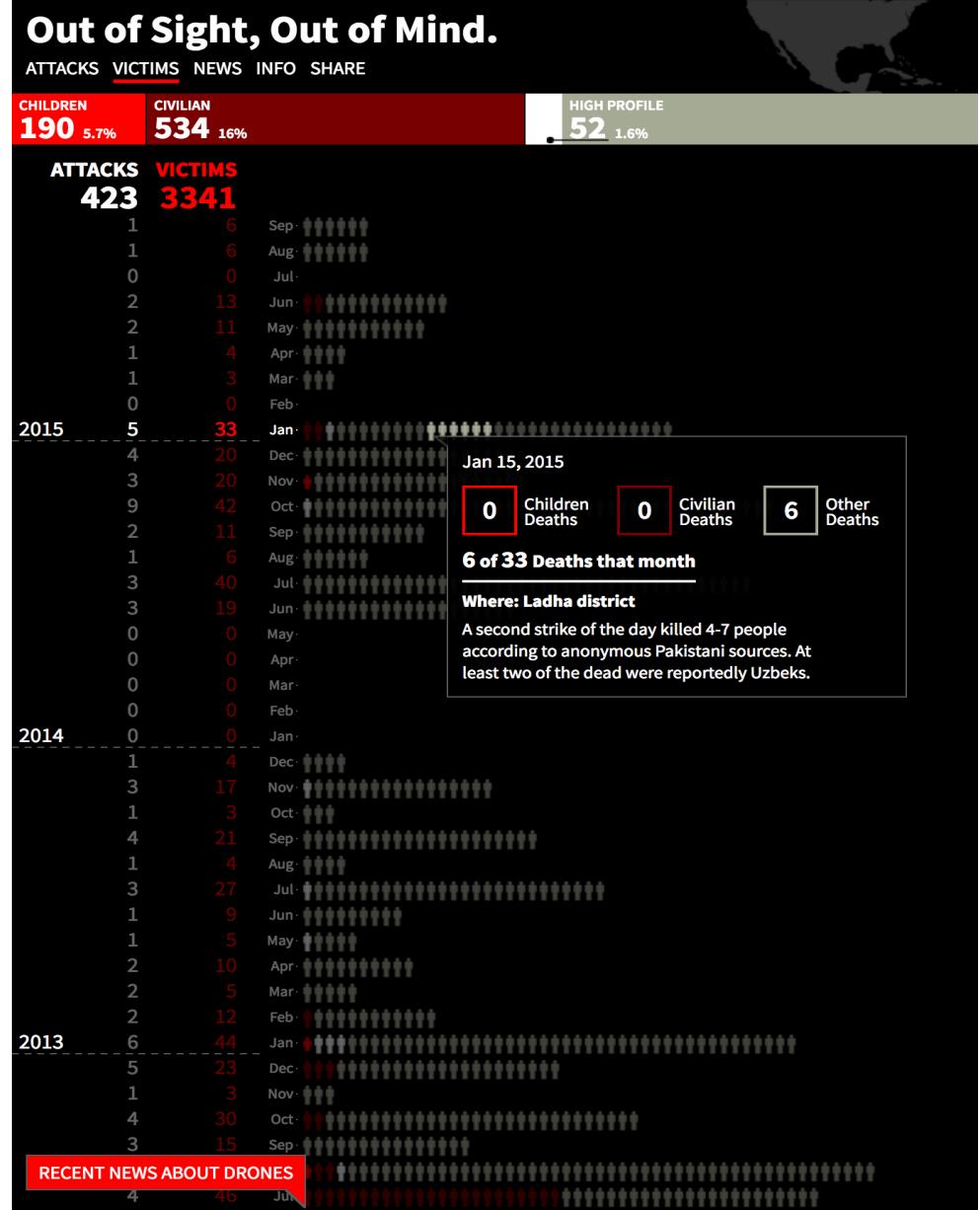
Out of Sight, Out of Mind. Page Three,
Source: <http://drones.pitchinteractive.com/>



*Out of Sight, Out of Mind. Animation of Timeline,
Source: <http://drones.pitchinteractive.com/>*



*Out of Sight, Out of Mind. Details on Hover
Source: <http://drones.pitchinteractive.com/>*



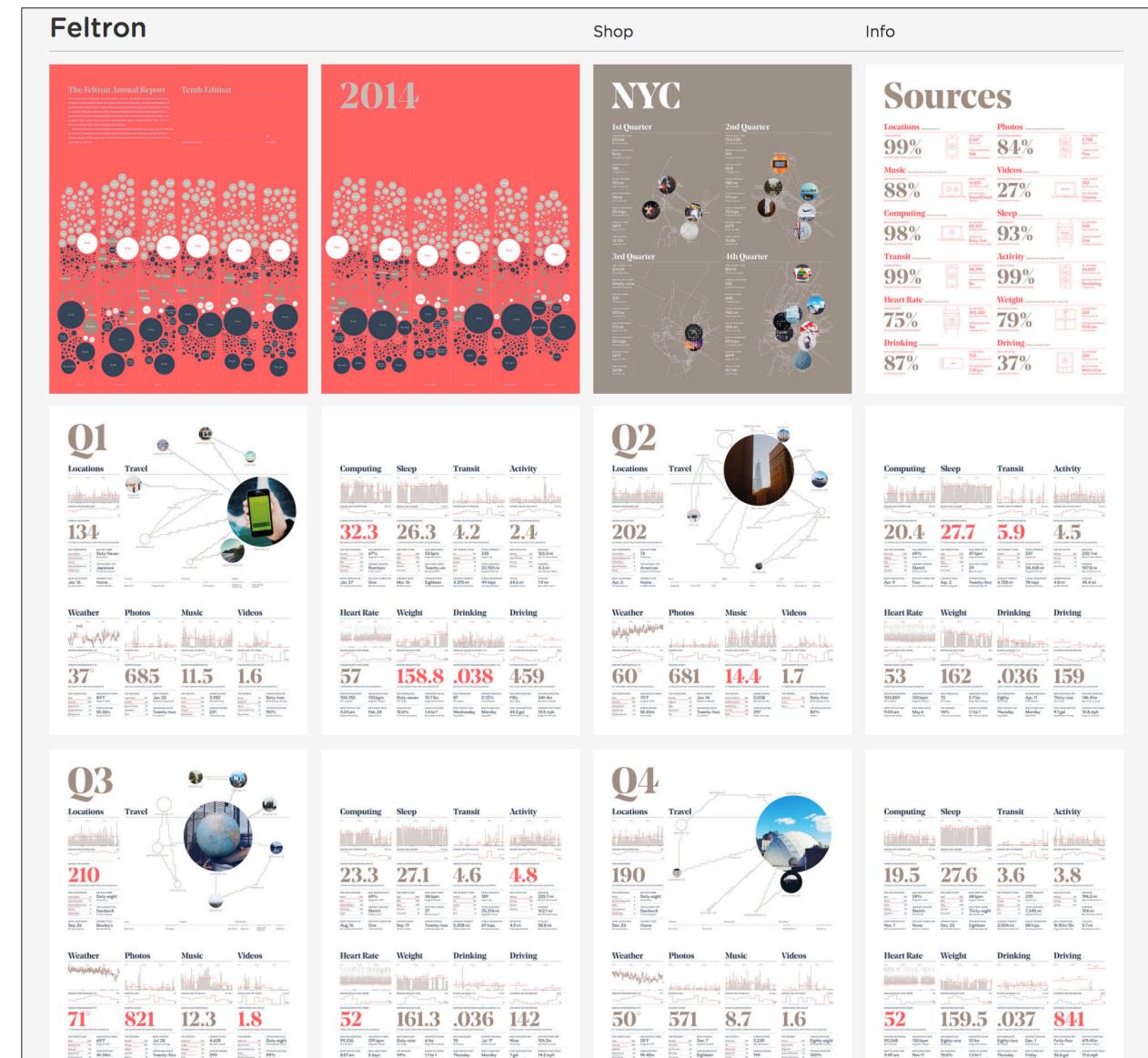
*Out of Sight, Out of Mind. Victims tab,
Source: <http://drones.pitchinteractive.com/>*

Storytelling Systems: Personalising Stories

Feltron 2014 Annual Report

Everyone is interested in themselves and things that they care about. Interactivity gives the audience the opportunity to make the content about them. Story in its consumption is very directly about the listener and their reality as much as it is about the storyteller. It's very powerful to think of storytelling in this light, to try to create story consciously to connect directly to the audience and their need to understand certain things. To do this, often the best way is to allow them to make the story about themselves. Power of medium like Virtual Reality to create immersion and subjectivity can be used to create this effect.

Visualisations on right are a series of annual reports on designers life based on some matrix that he tracked on data about his own personal life year after year. This is a static report and it tracks his life such as where he travelled to, how he was sleeping, what his weight was, how much he drank etc. It's all personalised data. It's really interesting to see what he tracked, how he tracked it and how he visualised all of it. It's also interesting to see what it means about our society and the things we can and should track for ourselves and eachother. It's essentially dashboard into a personal life. If you can take the data you are communicating and transform it to a personal experience for them, even if it's not their data, you will allow them see themselves in the data.
^[4] They are not seeing a random survey or research report, they can imagine themselves in the story. If you can make your data story into their data story, it will resonate the way impersonal data story about other really could never do.



Feltron 2014 Annual Report, Source: feltron.com/FAR14.html

Visualizing MBTA

Visualizing MBTA data is another personalised data storytelling. It starts with talking about the trains. It shows the subway trip on Monday February 3, 2014 on all the lines and where they were on different times of the day and how long it took them to get there.

As we scroll down there is a more complex data story and it takes a lot of different directions. Like sections about the people, entrance and exits per station, how people and trains affect each other and your commute. Your commute is data being personalised. If you click and select start and end points, you get a very personalised view of the data.

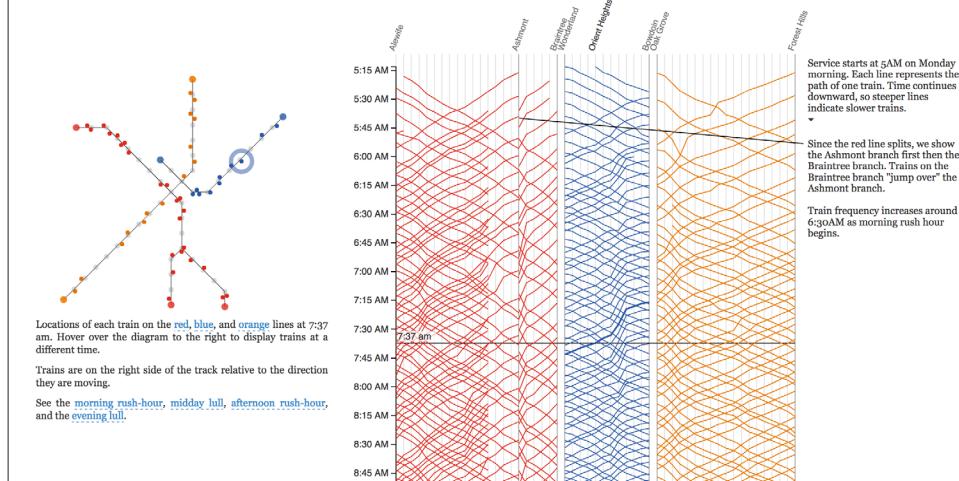
The Trains

In a typical weekday, trains make approximately 1150 trips on the red, orange, and blue lines starting at 5AM and continuing through 1AM the next morning. On Saturdays trains make 870 trips and on Sundays they make 760.

To better understand how the trains operate on a typical day, below are all trips that trains took on the red, orange, and blue lines on Monday February 3 2014. Each vertical line represents a station, and time extends from top to bottom. Steeper lines indicate slower trains. This visualization was first used by Étienne-Jules Marey to visualize train schedules and is typically called a "Marey Diagram."

	Weekdays	Saturdays	Sundays
Red	450	350	300
Orange	320	260	220
Total	1150	870	760

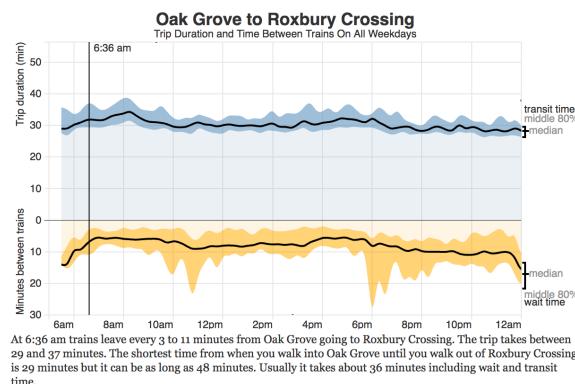
Subway Trips on Monday February 3, 2014



Trains section from Visualizing MBTA Data, Source: mbtaviz.github.io/

Your Commute

How do all of these factors affect your commute? Click and drag on the map from a starting station to an ending station to see a detailed breakdown of how long that trip takes at different points during a typical workday. The points on top show all of the trip durations for a given starting time from the start to destination and the points on bottom show all of the times between when trains leave the start station going to the destination station. The time between trains is the longest you would possibly need to wait if you arrived just as the previous train was leaving. The blue band excludes the shortest and longest 10% of all transit times, leaving behind the most-likely 80% range and the orange band does the same for wait times between trains. The dark lines show the middle point where 50% of the time wait/transit times are higher and 50% of the time they are lower.



Your Commute section from Visualizing MBTA Data, Source: mbtaviz.github.io/

Storytelling Systems: Textual Data Stories

Nine Charts about Wealth Inequality in America

Data storytelling can be done without lots of crazy visuals and interactivity.^[4] A story can be told in text using data as backdrop. With text and supplemented charts, data can create interesting stories. Any story with data will be better if there is some visual component to bring the ideas home, hence charts. The following example is about wealth and inequality from urban institute and it has 9 charts, nine chunks of story.

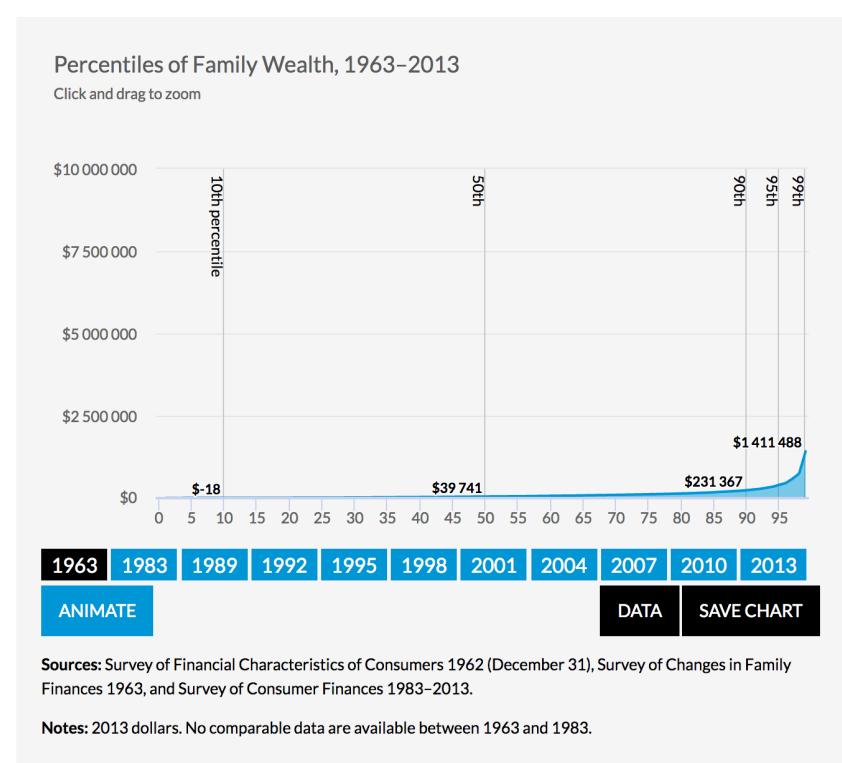
The first statement is titled wealth inequality is growing and an explanation of what that is. This is all text, is this a good data storytelling experience, just based on the text?^[19] Perhaps. You can read the data and learn a bunch of thing that the designer wants to tell. As you scroll down, you see nine chunks of data experience. Second title is ‘One reason for rising wealth inequality is income inequality’. Third title is ‘Racial and ethnic wealth disparities are also growing’. It’s telling a linear experience in text but the charts make it better.^[20] The charts are interactive and the viewer can go in depth progressively. The ninth story ‘Federal policies fail to promote asset building by lower-income families’ makes way for the conclusion ‘Promising policies to shrink wealth inequality and racial wealth gaps’. It’s a clear data story that makes a logical progression of arguments in text with supplemental charts.

1 Wealth inequality is growing

Average wealth has increased over the past 50 years, but it has not grown equally for all groups. Between 1963 and 2013,

- families near the bottom of the wealth distribution (those at the 10th percentile) went from having no wealth on average to being about \$2,000 in debt,
- those in the middle roughly doubled their wealth—mostly between 1963 and 1983,
- families near the top (at the 90th percentile) saw their wealth quadruple,
- and the wealth of those at the 99th percentile—in other words, those wealthier than 99 percent of all families—grew sixfold.

These changes have increased wealth inequality significantly. In 1963, families near the top had six times the wealth (or, \$6 for every \$1) of families in the middle. By 2013, they had 12 times the wealth of families in the middle.



First of the Nine Charts about Wealth Inequality in America

Source: apps.urban.org/features/wealth-inequality-charts/

Conclusion

All humans crave stories. They make an emotional connection in a way that facts can never do. Data stories are extra compelling as they combine the evolutionary imperative of stories with the important fact driven element that we need to make decisions. Data stories being different than fictional stories they can be used in multiple ways to increase the impact the facts and numbers.

References

- [1] *Michael Austin*, Useful Fictions: Evolution, Anxiety and the origins of literature. University of Nebraska Press, 2011.
- [2] *Lisa Cron*, Wired for story: The writer's guide to using brain science to hook readers from the very first sentence. Ten Speed Press, 2012.
- [3] *Aleks Zarnitsyn*, The Cognitive Value of Fiction in Thought Experiments in Personal Identity <https://muse.jhu.edu/article/582499> (accessed March 20, 2017)
- [4] *Bill Shander*, Data Visualization Storytelling Essentials <https://www.lynda.com/Excel-tutorials/Data-Visualization-Storytelling-Essentials/435230-2.html> (accessed March 20, 2017)
- [5] *Yuval Noah Harari*, Why humans run the world <http://ideas.ted.com/why-humans-run-the-world/> (accessed March 20, 2017)
- [6] *Alissa Lorentz*, With Big Data, Context is a Big Issue <https://www.wired.com/insights/2013/04/with-big-data-context-is-a-big-issue/> (accessed March 20, 2017)
- [7] *Jennifer Aaker*, The Power of Story <https://www.youtube.com/watch?v=CdO9a41WUss> <https://vimeo.com/175070260> (accessed March 20, 2017)
- [8] Language and the brain, http://pandora.cii.wwu.edu/vajda/ling201/test4materials/language_and_the_brain.htm (accessed March 20, 2017)
- [9] How Does Our Brain Process Language?, http://readingroom.mindspec.org/?page_id=8898 (accessed March 20, 2017)
- [10] *Edward Segel and Jeffrey Heer*, Narrative visualisation: telling stories with data <http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf>
- [11] *Julie Beck*, Life's Stories, <https://www.theatlantic.com/health/archive/2015/08/life-stories-narrative-psychology-redemption-mental-health/400796/> (accessed March 20, 2017)
- [12] *Z. Hereford*, Good Communication Skills - Key to Any Success <http://www.essentiallifeskills.net/goodcommunicationskills.html> (accessed March 20, 2017)
- [13] *Oliver Burkeman*, Working with your hands: the secret to happiness <https://www.theguardian.com/lifeandstyle/2010/may/08/working-hands-happiness-burkeman> (accessed March 20, 2017)
- [14] *Carmine Gallo*, The Best PowerPoint Tip You'll Ever Get <http://www.forbes.com/sites/carminegallo/2011/07/21/the-best-powerpoint-tip-youll-ever-get/#78b38eef2459> (accessed March 20, 2017)
- [15] How to Create an Epic Fantasy Story, <http://www.wikihow.com/Create-an-Epic-Fantasy-Story> (accessed March 20, 2017)
- [16] Writing Essays, <http://www.time4writing.com/writing-essays/> (accessed March 20, 2017)
- [17] *Kurt Vonnegut*, The Shapes of Stories, <https://www.youtube.com/watch?v=oP3c1h8v2ZQ> (accessed March 20, 2017)
- [18] Jed's Guide to Story Reading, <http://www.kith.org/logos/things/reading.html> (accessed March 20, 2017)
- [19] Tell a Meaningful Story With Data, <https://www.thinkwithgoogle.com/articles/tell-meaningful-stories-with-data.html> (accessed March 20, 2017)
- [20] *Ginny Mineo*, 10 Design Tips to Create Beautiful Excel Charts and Graphs, <https://blog.hubspot.com/marketing/excel-graph-tricks-list> (accessed March 20, 2017)

References

Websites

Gapminder, https://www.gapminder.org/tools/#_chart-type=bubbles

Selfiecity, <http://selfiecity.net/#imageplots>

Rank of the most populous cities at each census: 1790-1890, http://vallandingham.me/census_bump/

Out of Sight, Out of Mind,
<http://drones.pitchinteractive.com/>

Feltron 2014 Annual Report, <http://feltron.com/FAR14.html>

Visualizing MBTA Data, <http://mbtaviz.github.io/>

Nine Charts about Wealth Inequality in America,
<http://apps.urban.org/features/wealth-inequality-charts/>

Images

Charles Minard's Map of Napoleon's Russian Campaign of 1812 https://en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Minard.png

Find out what you are, <https://visual.ly/find-out-what-you-are>

Rank of the most populous cities at each census: 1790-1890, <http://vintagevisualizations.com/products/rank-of-the-most-populous-cities-at-each-census-1790-1890>

Kurt Vonnegut on the Shapes of Stories, <http://visual.ly/kurt-vonnegut-shapes-stories>