Storytelling with Data

Module 6: The storytelling process

Scott Spencer
Faculty and Lecturer
Columbia University

Agenda

Upcoming deliverable — *draft* storyboard

Today's objectives

Stories — why, what, how

Stories, in words and pictures

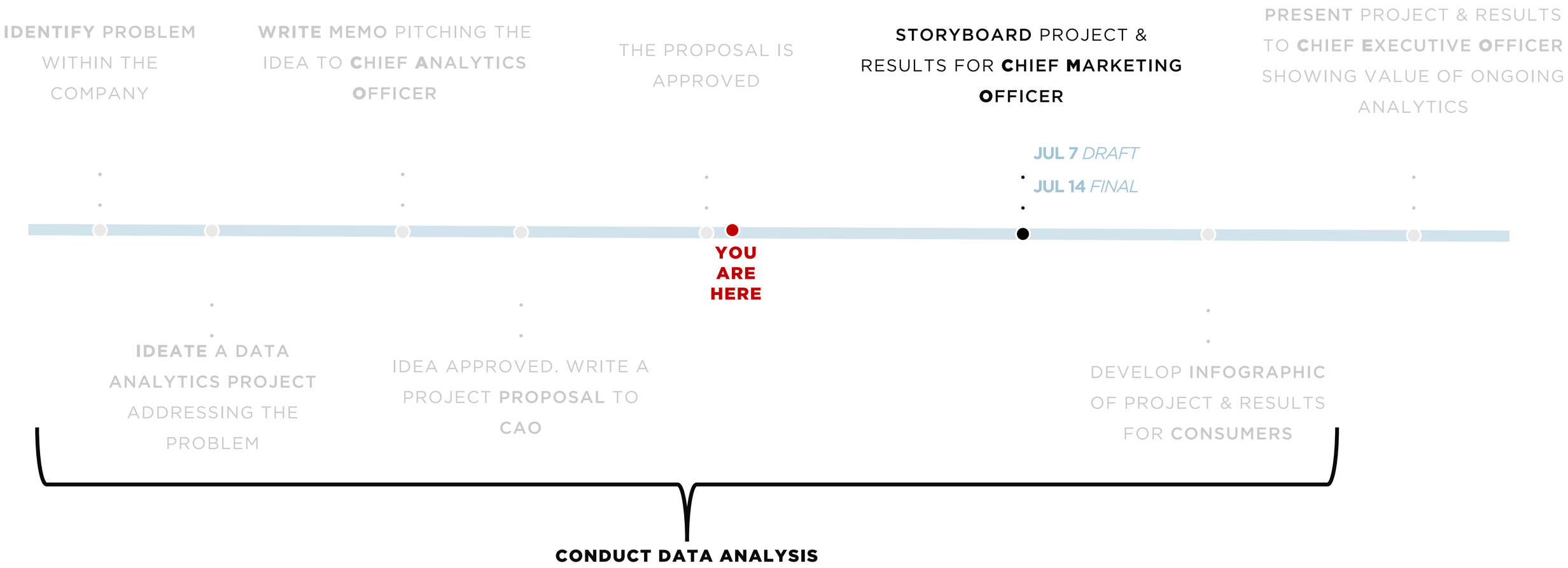
From data to visual

Questions or suggestions?

Upcoming deliverable

Upcoming deliverable

In **storyboard** form — describe (1) your project, (2) preliminary results or insights so far, and (3) why those results are interesting for the marketing team. Use a distinct narrative arc (beginning, middle, and end), be clear and accessible for the **CMO**.



Today's Objectives

Objectives

1. Build a story around an analytics project.
2. Craft a narrative to captivate, convince, and inspire an audience.

Stories — why, what, how

Management is much more than a science

Martin & Golsby

Professor Martin is director of the Martin Prosperity Institute at the School of Management, University of Toronto. Golsby-Smith is founder of 2nd Road, a strategy and innovation firm, now part of Accenture Strategy.

Narratives frame hypotheses for what could be

To make decisions about **what could be**, managers should devise narratives about possible futures, applying the tools of metaphor, logic, and emotion.

Consider possibilities broader than data

The **absence of data does not preclude possibility**. If we are talking about new outcomes and behaviors, then there is no prior evidence. Consider not only what the data suggests but also what within the bounds of possibility could happen.

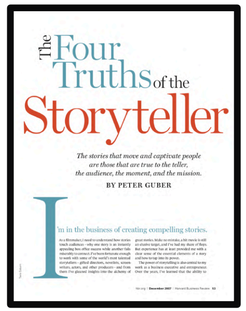
Metaphor helps compare unlinked ideas

The core engine of creative synthesis is "associative fluency"—the mental ability to **connect two concepts that are not usually linked** and to forge them into a new idea.

Clarify conditions, create new data

With new narratives—hypotheses—we must **hypothesize** what would have to be true for those narratives to happen and **validate** their hypotheses through prototyping.

Martin, Roger L, and Tony Golsby-Smith. "Management Is Much More Than a Science ." Harvard Business Review (2017): 129–135. Print.



The four truths of the storyteller

Guber

Peter is author of *Tell to Win*, has been awarded business person of the year by Los Angeles Business Journal, is owner of several pro sports teams and an entertainment company. He was top executive at Sony and Columbia Pictures.

Persuasive stories depend on four truths

Great storytelling does not conflict with truth. Stories that move people are true to the teller, audience, moment, and mission.

Teller. The teller must be congruent with the story, must share emotions felt, allowing the audience to identify with those emotions.

Audience. As a teller, we must take time to understand what our listeners know about, care about, and want to hear. And then, craft a story that resonates with what we learn. Test the story on others, skeptics. Include audience in the tale—*I* becomes *we*.

Moment. Know the tale well enough to improvise, molding it to the moment's context.

Mission. The tale should capture the mission, and truth to the mission should win over any conflicts with truth to the audience.

Guber, Peter. "The Four Truths of the Storyteller." Harvard Business Review (2007): 53–59. Print.

The age-old practice of storytelling is one of the most effective tools leaders can use. But they need to pick their stories carefully and match them to the situation.

Telling Tales

by Stephen Denning

Telling tales

Denning

Stephen is the former director of Knowledge Management at the World Bank, an author, and focuses on leadership strategy.

A story — narrative as sequence of events

Although good business arguments are developed through the use of numbers, they are typically approved on the basis of a story—that is, a narrative that links a set of events in some kind of causal sequence.

Narrative patterns for business

Stories can advance business objectives, e.g.:

- Sparking action
- Communicating who you are
- Transmitting values
- Fostering collaboration
- Taming the grapevine
- Sharing knowledge
- Leading people into the future

Denning, Stephen. “Telling Tales.” Harvard Business Review (2004): 122–129. Print.

NAPA Cards *Preview*

Narrative Patterns for Data-Driven Storytelling

What is this?

EMPATHY [4] ENGAGEMENT [3] FRAMING [11] FLOW [3] ARGUMENT [3]

Incorporating the audience

HOW

Audience becomes part of the narrative.

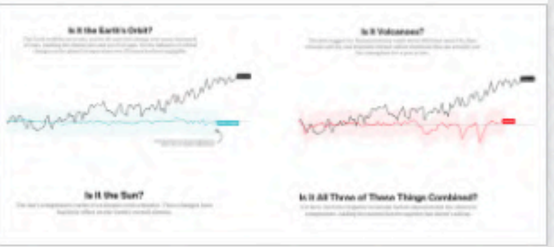
WHY

Increase empathy, facilitate personal comparisons. Helps make the data and information more personally relevant. To allow reader to use herself as measuring device.

EXAMPLE

Ubisoft: Collapse

Repetition



HOW

A phenomenon is re-presented multiple times with changes to the main variable/dimension (the data) or the time frame or the animation pattern (see the example) while other elements of the story don't change. Note: this could be like running multiple simulations with the data to see which is most effective (and in this case, the story would be more reader driven)

WHY

To show that the same phenomenon happens over and over again. It also strengthens the narrative through rhythm. Establishing a constant reference frame can help to emphasize change and differences.

EXAMPLE

Bloomberg: What's really warming the world?

Juxtaposition

HOW

Showing two or more complementary visualizations (pictures) juxtaposed.

WHY

Compare cases or samples, allows to highlight contrast, difference, absense of difference, change over time. Allows for individual exploration of (minor) differences.

EXAMPLE

David McCandless: Billion Dollar Gram

Breaking the fourth wall

Humans behind the dots

Make a guess

NAPA Cards. *Narrative Patterns for Data-Driven Storytelling*. Web. http://napa-cards.net

Stories, in pictures and words

Understanding Comics

McCloud

Scott provides the seminal reference, in the context of comics, on understanding the connection between words and visuals.

Words and visuals have shared meaning

In comics, visuals and words complement one another, each contributing to a shared meaning.

McCloud, Scott. Understanding Comics: the Invisible Art. Kitchen Sink Press, 1993. Print.

Story in sentences

I CROSSED THE STREET TO THE CONVENIENCE STORE. THE RAIN SOAKED INTO MY BOOTS.

I FOUND THE LAST PINT OF CHOCOLATE CHOCOLATE CHIP IN THE FREEZER.

THE CLERK TRIED TO PICK ME UP. I SAID *NO THANKS*. HE GAVE ME THIS CREEPY LOOK...

I WENT BACK TO THE APARTMENT--

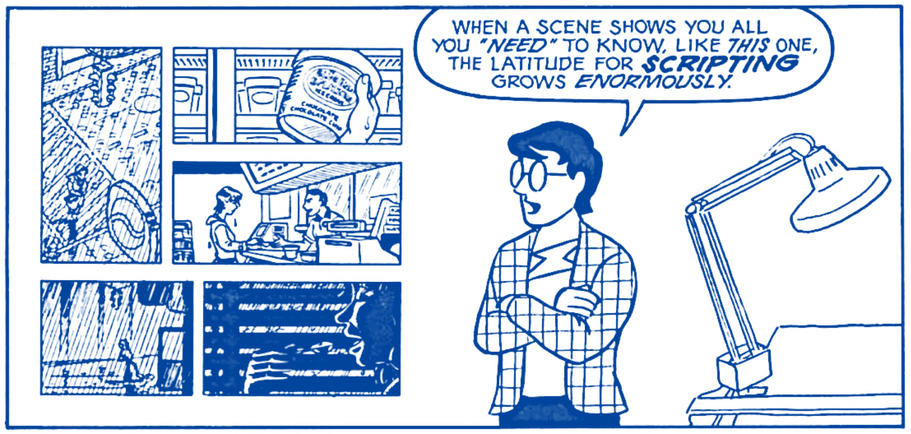
--AND FINISHED IT ALL IN AN HOUR.

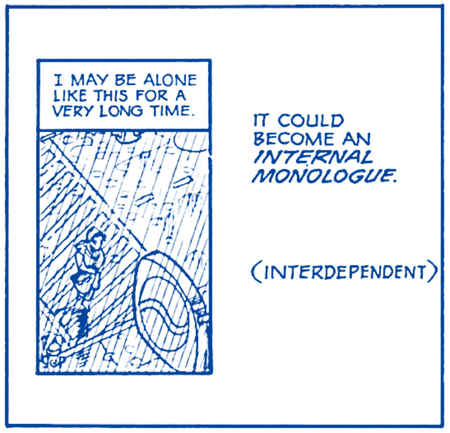
ALONE AT LAST.

Story in pictures

McCloud, Scott. *Understanding Comics: the Invisible Art*. Kitchen Sink Press, 1993. Print.

Use words and images for different purposes

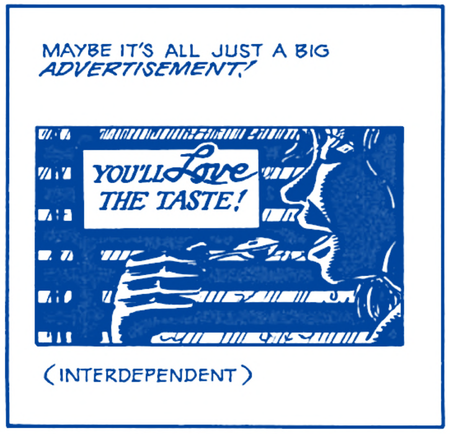




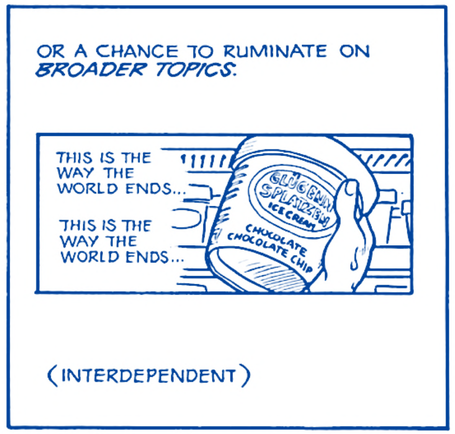
IT COULD BECOME AN *INTERNAL MONOLOGUE.*

(INTERDEPENDENT)

(PARALLEL)



(INTERDEPENDENT)



(INTERDEPENDENT)

IF THE ARTIST WANTS TO, HE/SHE CAN NOW SHOW ONLY *FRAGMENTS* OF A SCENE.

(WORD SPECIFIC)

OR MOVE TOWARD GREATER LEVELS OF *ABSTRACTION* OR *EXPRESSION.*

(AMPLIFICATION)

PERHAPS THE ARTIST CAN GIVE US SOME IMPORTANT *EMOTIONAL* INFORMATION.

(INTERDEPENDENT)

OR SHIFT AHEAD OR BACKWARDS IN TIME.

(WORD SPECIFIC)

McCloud, Scott. *Understanding Comics: the Invisible Art*. Kitchen Sink Press, 1993. Print.

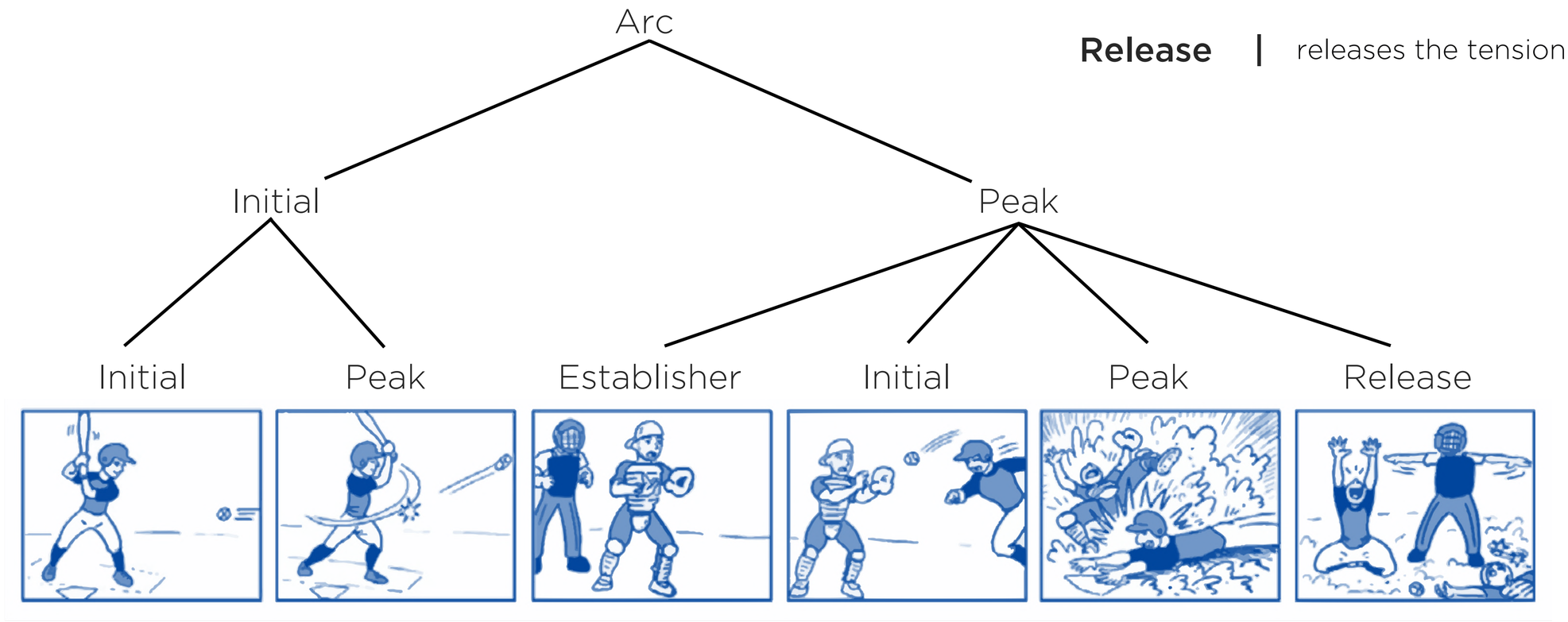
Like written narrative, sequential pictures can have grammar

Establisher | sets up an interaction without acting upon it

Initial | initiates the tension of the narrative arc

Peak | marks the height of narrative tension and point of maximal event structure

Release | releases the tension of the interaction



Cohn, Neil et al. "The Grammar of Visual Narrative: Neural Evidence for Constituent Structure in Sequential Image Comprehension." Neuropsychologia 64.C (2014): 63–70. Web.

McCloud, Scott. *Understanding Comics: the Invisible Art*. Kitchen Sink Press, 1993. Print.

You can tell a story to your audience on a **storyboard**, which—like comics—has **visual** components, corresponding **written narrative**, and **sequencing**.

Visual narrative

Written narrative

1

2

3

Visual narrative

Written narrative

4

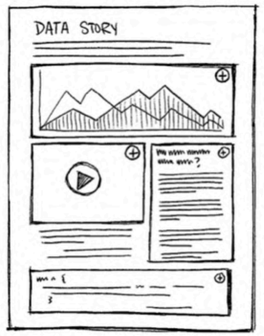
5

6

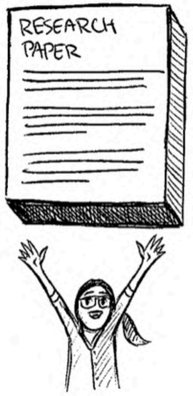
A simple example.
Storyboard for IBM as part of making an end-to-end experience for data scientists to research, create, and collaborate.

Joan, a Data Scientist at an enterprise company, is looking for resources to help her figure out how weather affects customer sales.

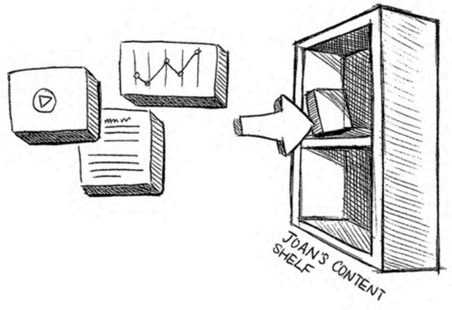
She turns to Google for help. The top search result shows a preview of an interesting paper from Project Miles.



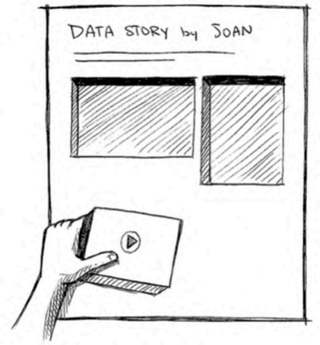
Joan clicks on the search result where she sees the paper in context of a workflow, called a “Data Story.” It was created by another user named Paula.



The paper presents a weather algorithm that Joan can use for her own project. She bookmarks the paper for later use.



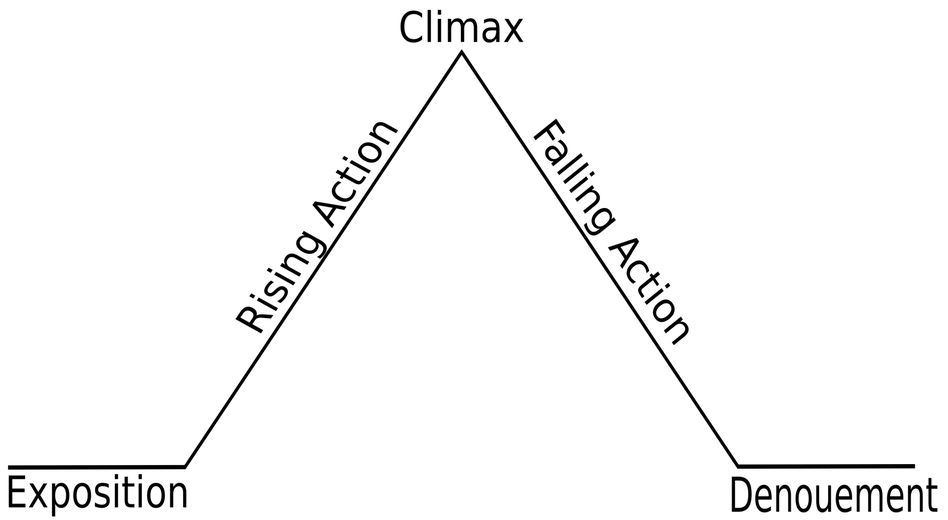
Joan continues reading Paula’s Data Story because she sees a wealth of relevant resources, e.g. videos, papers, talks, that she could use for her project.



Once she has added a handful of resources to her bookmark list, Joan creates her own Data Story. She uses her bookmark content as a reference point.

Peng, Rose. *IBM Data Science Experience*. 4 Mar 2019. Web. http://www.rosepeng.com/data-science-experience

A story arc for analytics projects



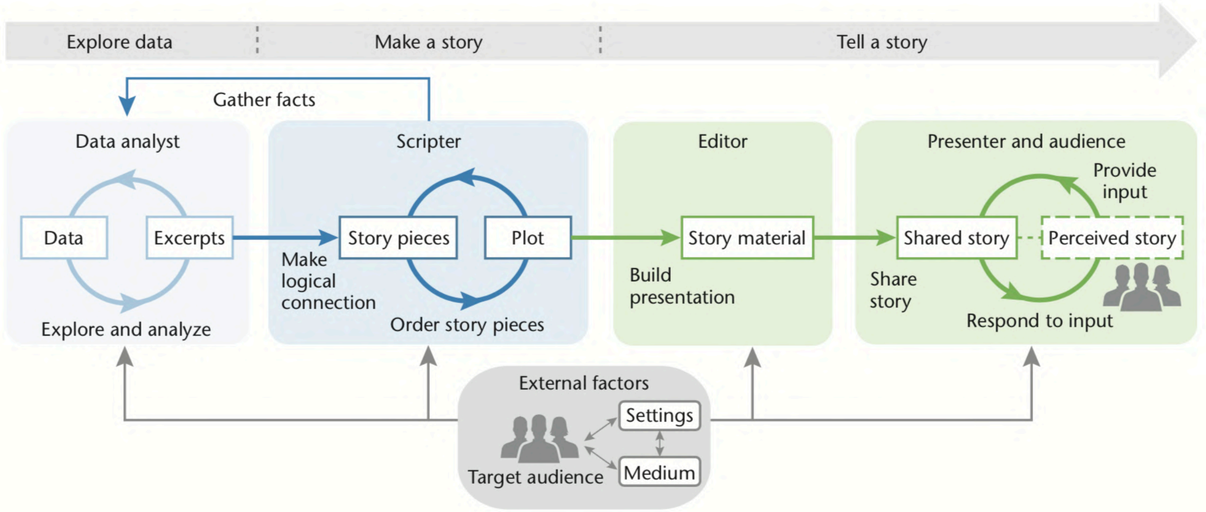
The primary elements of an applied analytics project are a well-articulated business problem, a data science solution, and a measurable outcome to produce value for the organization.

The analytics project may thus be conceptualized as a story arc, with a beginning (**problem**), middle (**analytics**), and end (**overcoming of the problem**), along with characters (**analysts, colleagues, clients**) who play important roles.

Transforming data into visually shared stories

Lee & co-author

Authors Lee and Riche, who was also the main author of Data-Driven Storytelling, work at Microsoft, advancing research on data stories.



Scope of a visual data story

A visual data story is more than a visualization of data. They include a set of story pieces — that is, **specific facts backed up by data**.

They are **visualized** to support intended messages, include **annotations** (labels, pointers, text) and **narration** to highlight and emphasize the message and avoid ambiguity.

They are presented in a **meaningful order** or with a connection between them to support the author's high-level communication goal.

Lee, Bongshin et al. "More Than Telling a Story: Transforming Data Into Visually Shared Stories." IEEE Computer Graphics and Applications 35.5 (2015): 84–90. Web.

Revisiting the example Dodger's analytics project, CMO

Front Office Directory

« Back to directory

Lon Rosen

Executive Vice President & Chief Marketing Officer

One of the foremost executives in the sports and entertainment industry, Lon Rosen directs the Dodger organization's marketing, sales and broadcasting initiatives as their executive vice president and chief marketing officer.

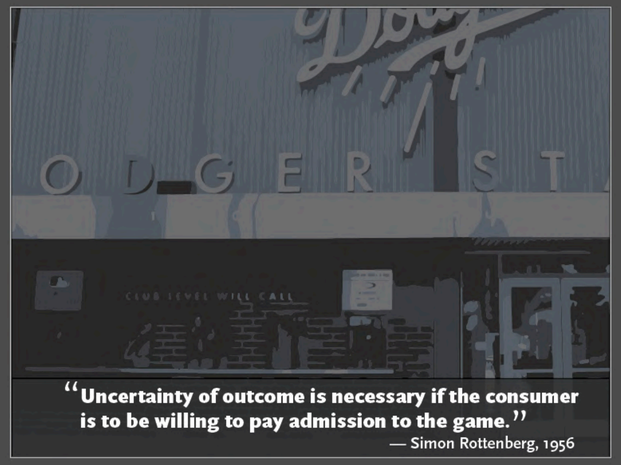
With more than 25 years of experience in the field, Rosen was instrumental in assisting his former employer, Magic Johnson, and Guggenheim Partners in acquiring the Dodgers on May 1, 2012, in what was the largest team sports transaction in history. Rosen previously served in the same role for two years starting in 2004 before rejoining Magic Johnson Enterprises.

Rosen focused his efforts for Magic Johnson on business development, television productions, sports consulting and new media ventures, in addition to his representations of broadcasters, coaches, hosts and other entertainment personalities.

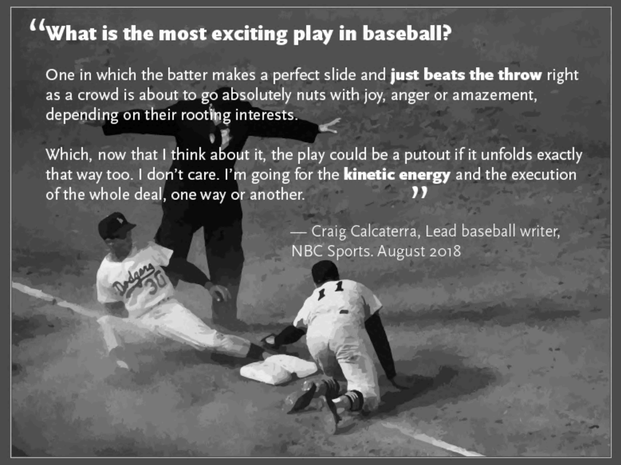
Upon graduation from college, Rosen headed to the Fabulous Forum, where he began his sports career with the NBA's Los Angeles Lakers and the NHL's Los Angeles Kings. He worked his way from an internship to the position of director of promotions for the Lakers, Kings and the Los Angeles Forum's Special Events, where he served seven years (1980-87). From there, he started his own sports marketing company, First Team Marketing, and moved into sports marketing and sports representation.

Rosen, a University of Southern California graduate, resides in Los Angeles with his wife, Laurie. They have two sons.

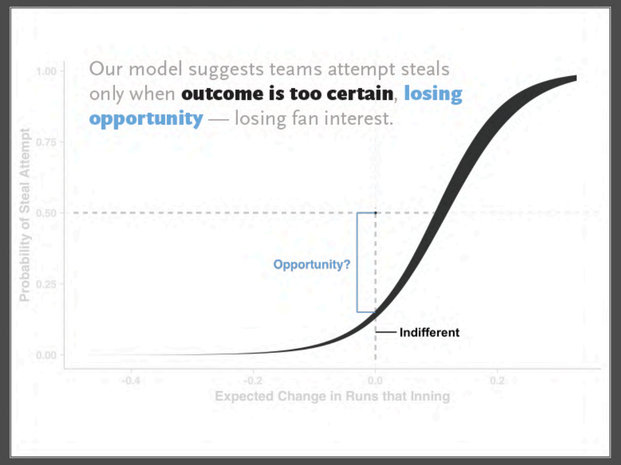
Storyboarding something relevant with a preliminary insight



1



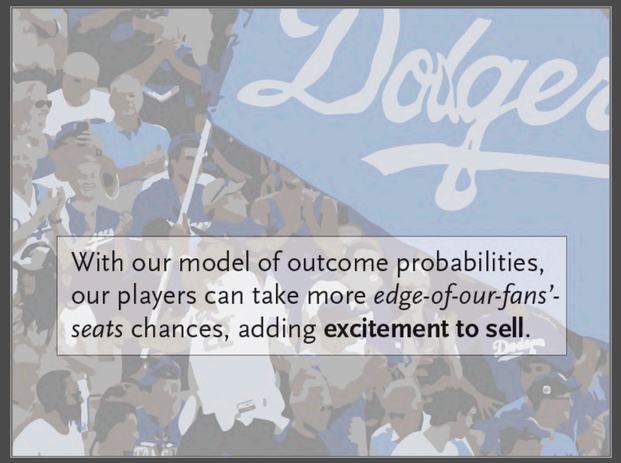
2



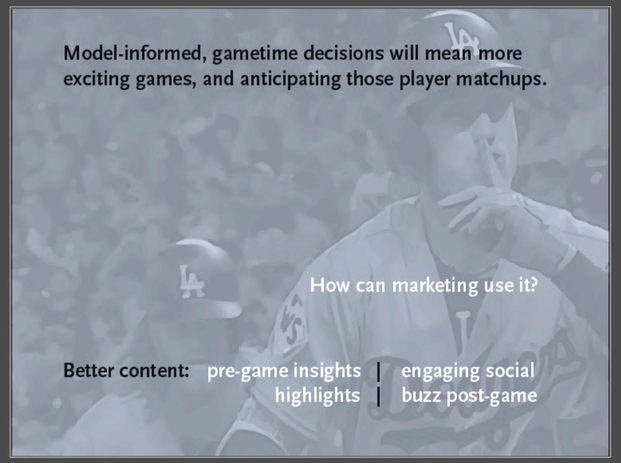
3

13th Dodgers were *average* taking risks with base stealing last season.

4

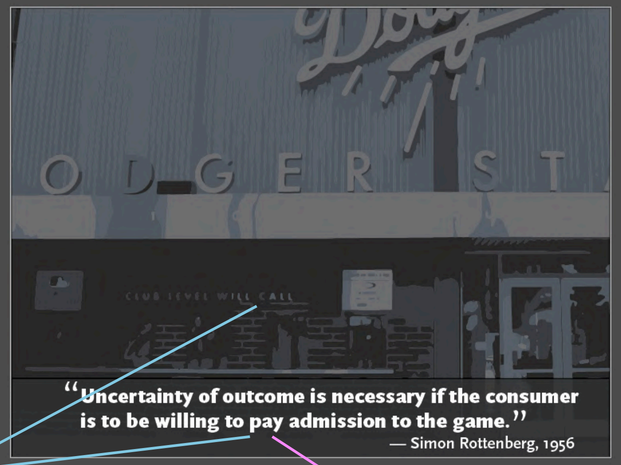


5

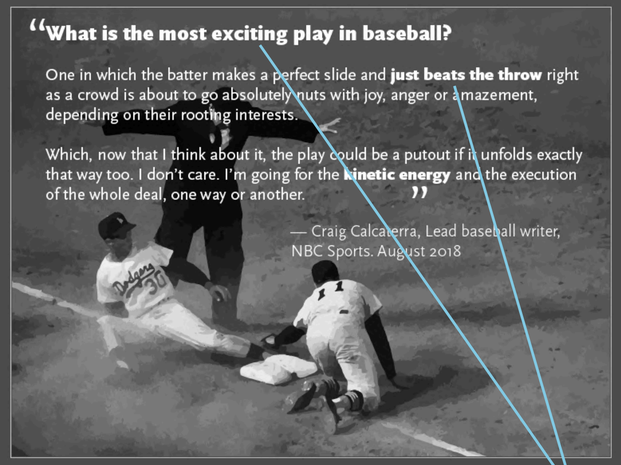


6

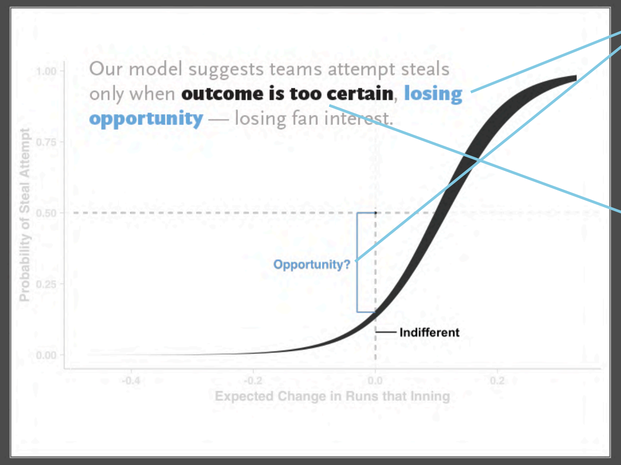
One approach to structuring this example



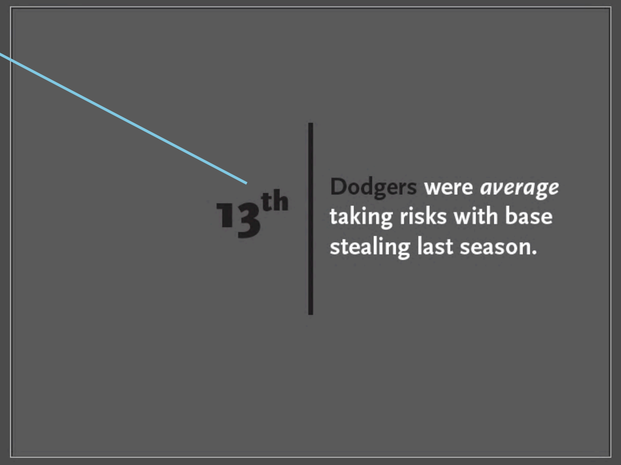
1



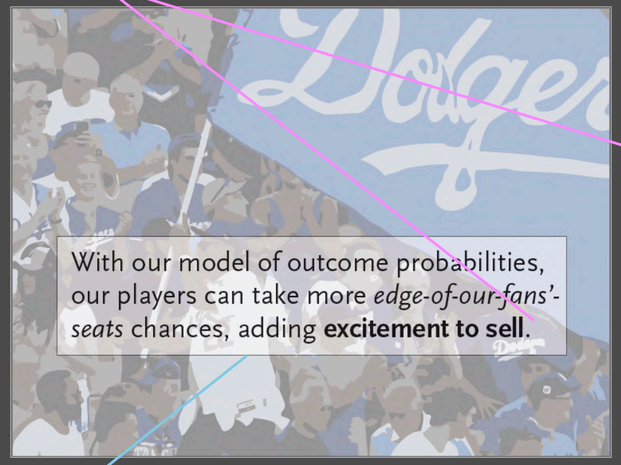
2



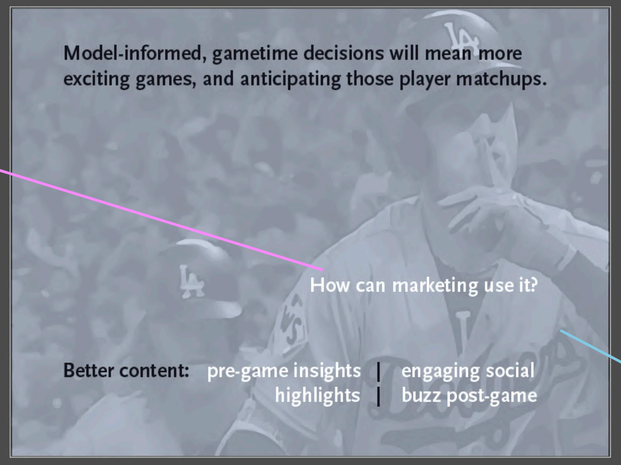
3



4



5



6

Color relates words to data display

Rising action or initial (problem)

Illustration congruent with words

Exposition or establisher

Rhetorical question, answered

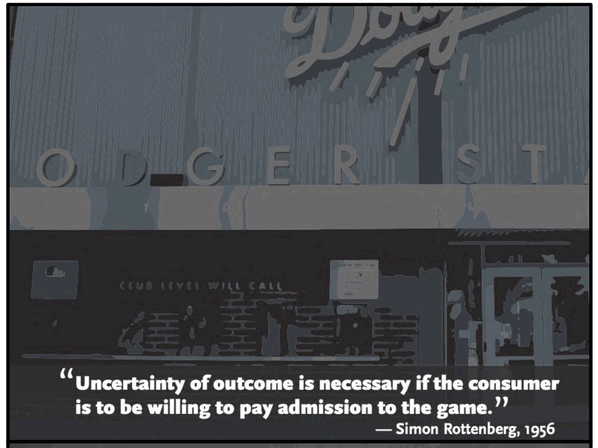
Relevance to marketing (CMO)

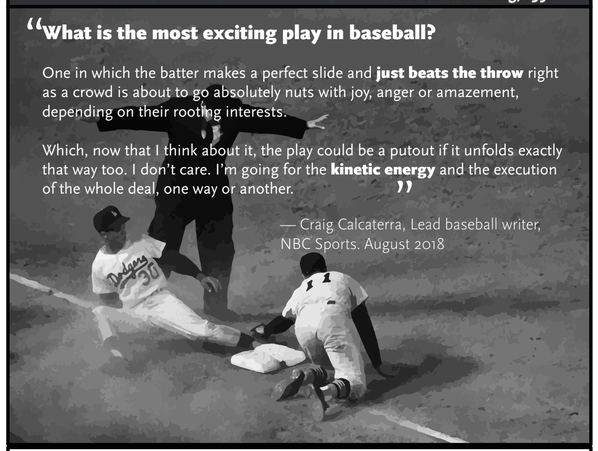
Climax or peak (analytics)

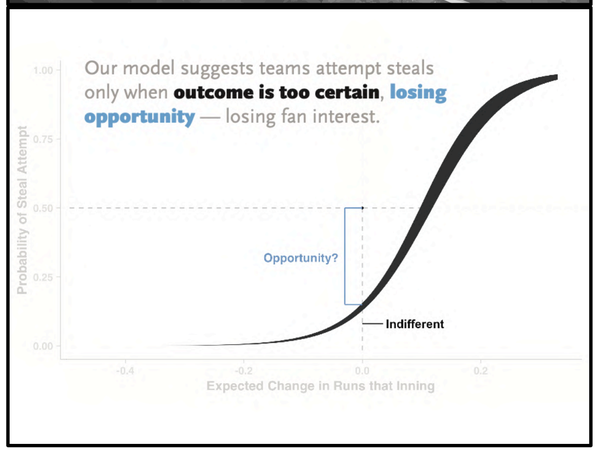
Falling action or release (solution)

Denouement or call to action

Visual narrative







Written narrative

1

Let's discuss some benefits that our new analytics model will provide for our marketing efforts. Our model helps to maximize "uncertainty of outcome," something famed economist Simon Rotternberg explained is necessary for consumers to pay admission to ball games. I think we would all agree that his insight generally holds true. And we do have room to improve. Our average game attendance hovers in the mid-80 percent capacity.

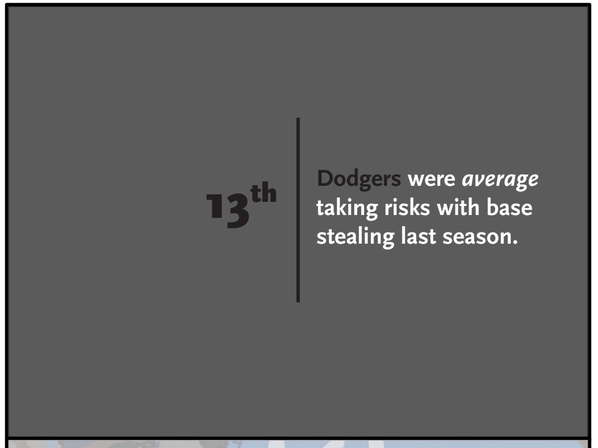
2

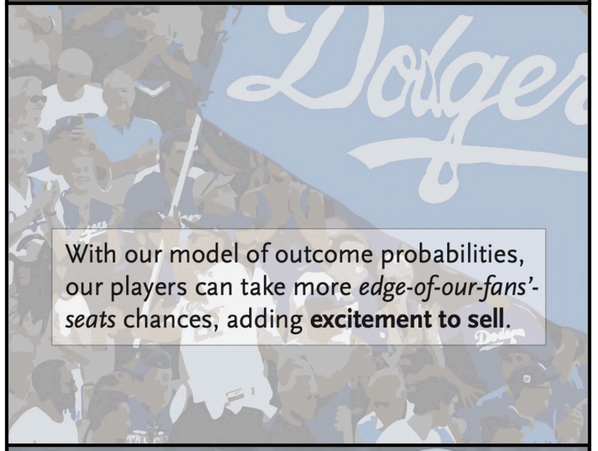
How can we increase game attendance and post-game buzz? Let's get some perspective from lead baseball writer for NBC Sports. Like Rottenberg, Calcaterra equates excitement, the kinetic energy, to play outcome uncertainty. The most exciting play may be, Craig says, one where the baserunner just beats the throw when stealing a base. How can our model help with this? Our model estimates the uncertainty of all possible event outcomes in the game for each play. What makes this important?

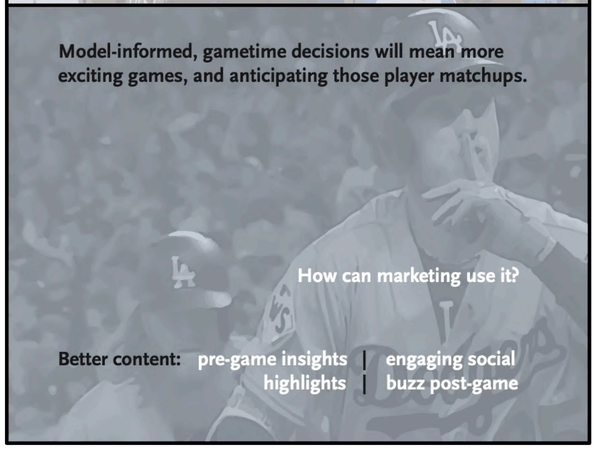
3

Let's continue the example with steals. From a game outcome perspective, whether we win, managers should be indifferent if a baserunner steals when his expected change in runs is zero. The model shows managers are too conservative. Following the model, they should be able to tell runners to steal when the outcome is more uncertain, but still favoring the team. Ideally, we should be looking for almost a coin flip. We performed well on the field last year. How do we stack up to other teams in taking risks on base stealing?

Visual narrative







Written narrative

4

We were only average when it came to Calcaterra's most exciting play in baseball. We still have room to take more risks, to give consumers more uncertainty in what happens.

5

Knowing how uncertain each play outcome will be gives our managers a new tool to tell players to take more risks. The outcome can seem closer to a coin-flip, but in the long run, we can keep our decisions on the winning side. Our model-informed game will be more exciting. Making it easier to sell tickets and generate buzz.

6

By accessing our model, marketing can make better content even before games start. While the model is proprietary, marketing will know ahead of time what matchups may be more interesting and can use that data to build anticipation. Along with pre-game insights, marketing can better engage with fans on social media, have better highlights as streaming content. Marketing better content will add to the post-game buzz.

Storyboarding tips

1. **Focus on the outcome of your persuasion**. What are you trying to persuade your audience to think or do differently?

2. **Include storytelling tools in your storyboard**. In the beginning, middle or end of your narrative, include one of the tactics described in class to help you connect emotionally with your audience.

3. **Combine images, graphs, and headlines to tell your story**. Make it easy for the reader to follow your story by using less copy and more instantly "telegraph" images, graphs and headlines.

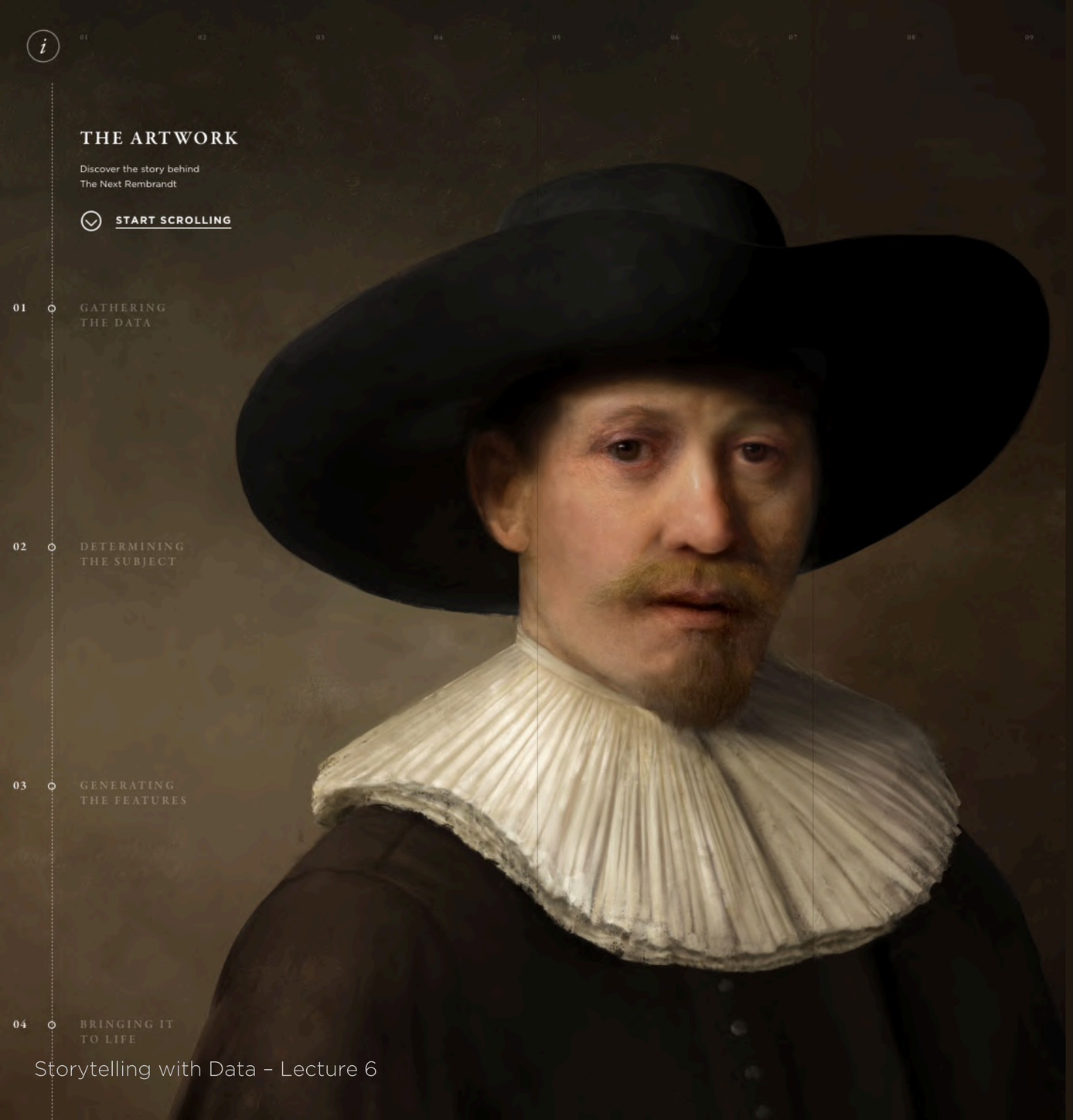
4. **Include your voiceover**. The slide images, data, and copy may not be clear enough to tell your story; tell us your voice over on the slide in a separate section.

5. **Use real data and, as needed, strong supporting images**. Do the analysis required to tell your story, show the key data. If your story depends on the images, then you should absolutely find them to make a bigger impact.

“It’s hard to have a story if nothing moves or nothing changes.”

— Jonathan Corum: the science graphics editor at The New York Times and founder of 13pt LLC, an information design studio.

Corum, Jonathan. *Design for an Audience*. 26 Apr 2018. Web. http://style.org/ku/



ING et al. “The Next Rembrandt.” 5 Apr. 2016. Web.
https://www.nextrembrandt.com

01
GATHERING THE DATA

BUILDING AN EXTENSIVE POOL OF DATA

It's been almost four centuries since the world lost the talent of one its most influential classical painters, Rembrandt van Rijn. To bring him back, we distilled the artistic DNA from his work and used it to create The Next Rembrandt.

We examined the entire collection of Rembrandt's work, studying the contents of his paintings pixel by pixel. To get this data, we analyzed a broad range of materials like high resolution 3D scans and digital files, which were upscaled by deep learning algorithms to maximize resolution and quality. This extensive database was then used as the foundation for creating The Next Rembrandt.

"Data is used by many people today to help them be more efficient and knowledgeable about their daily work, and about the decisions they need to make. But in this project it's also used to make life itself more beautiful. It really touches the human soul."

– Ron Augustus, Microsoft

02
DETERMINING THE SUBJECT

BREAKING DOWN THE DEMOGRAPHICS IN REMBRANDT'S WORK

To create new artwork using data from Rembrandt's paintings, we had to maximize the data pool from which to pull information. Because he painted more portraits than any other subject, we narrowed down our exploration to these paintings.

Then we found the period in which the majority of these paintings were created: between 1632 and 1642. Next, we defined the demographic segmentation of the people in these works and saw which elements occurred in the largest sample of paintings. We funneled down that selection starting with gender and then went on to analyze everything from age and head direction, to the amount of facial hair present.

After studying the demographics, the data lead us to a conclusive subject: a portrait of a Caucasian male with facial hair, between the ages of thirty and forty, wearing black clothes with a white collar and a hat, facing to the right.

FACING ANGLE ETHNICITY

THE NEXT REMBRANDT

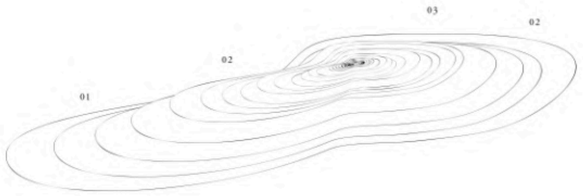
03
GENERATING THE FEATURES

CONSTRUCTING A FACE OUT OF THE NEW FEATURES

Once we generated the individual features, we had to assemble them into a fully formed face and bust according to Rembrandt's use of proportions. An algorithm measured the distances between the facial features in Rembrandt's paintings and calculated them based on percentages. Next, the features were transformed, rotated, and scaled, then accurately placed within the frame of the face. Finally, we rendered the light based on gathered data in order to cast authentic shadows on each feature.

04
BRINGING IT TO LIFE

CREATING ACCURATE DEPTH AND TEXTURE



We now had a digital file true to Rembrandt's style in content, shapes, and lighting. But paintings aren't just 2D — they have a remarkable three-dimensionality that comes from brushstrokes and layers of paint. To recreate this texture, we had to study 3D scans of Rembrandt's paintings and analyze the intricate layers on top of the canvas.

"We looked at a number of Rembrandt paintings, and we scanned their surface texture, their elemental composition, and what kinds of pigments were used. That's the kind of information you need if you want to generate a painting by Rembrandt virtually."

- Joris Dik, Technical University Delft

THE NEXT REMBRANDT

THE NEXT REMBRANDT

ING

PRESENTING PARTNER

ING is one the most innovative financial institutions in the world. Every day they reimagine new and relevant possibilities for their clients to stay a step ahead in life and business. The Next Rembrandt was a way for ING to bring this innovative spirit to their sponsorship of Dutch art and cultural icons.

Microsoft

SUPPORTING PARTNER

Microsoft saw an opportunity to show the power of data for businesses and individuals through The Next Rembrandt. Their Azure platform was instrumental to the project by supporting data handling and data analysis, as well as providing maximum calculating power.

WITH SPECIAL THANKS TO OUR ADVISORS

TU Delft Delft University of Technology | Mauritshuis | MUSEUM HET REMBRANDTHUIS

Questions for discussion.

Do you see a beginning, middle, and end? Where?

Where, in the visual-text combinations, do you see comparisons or contrasts or change?

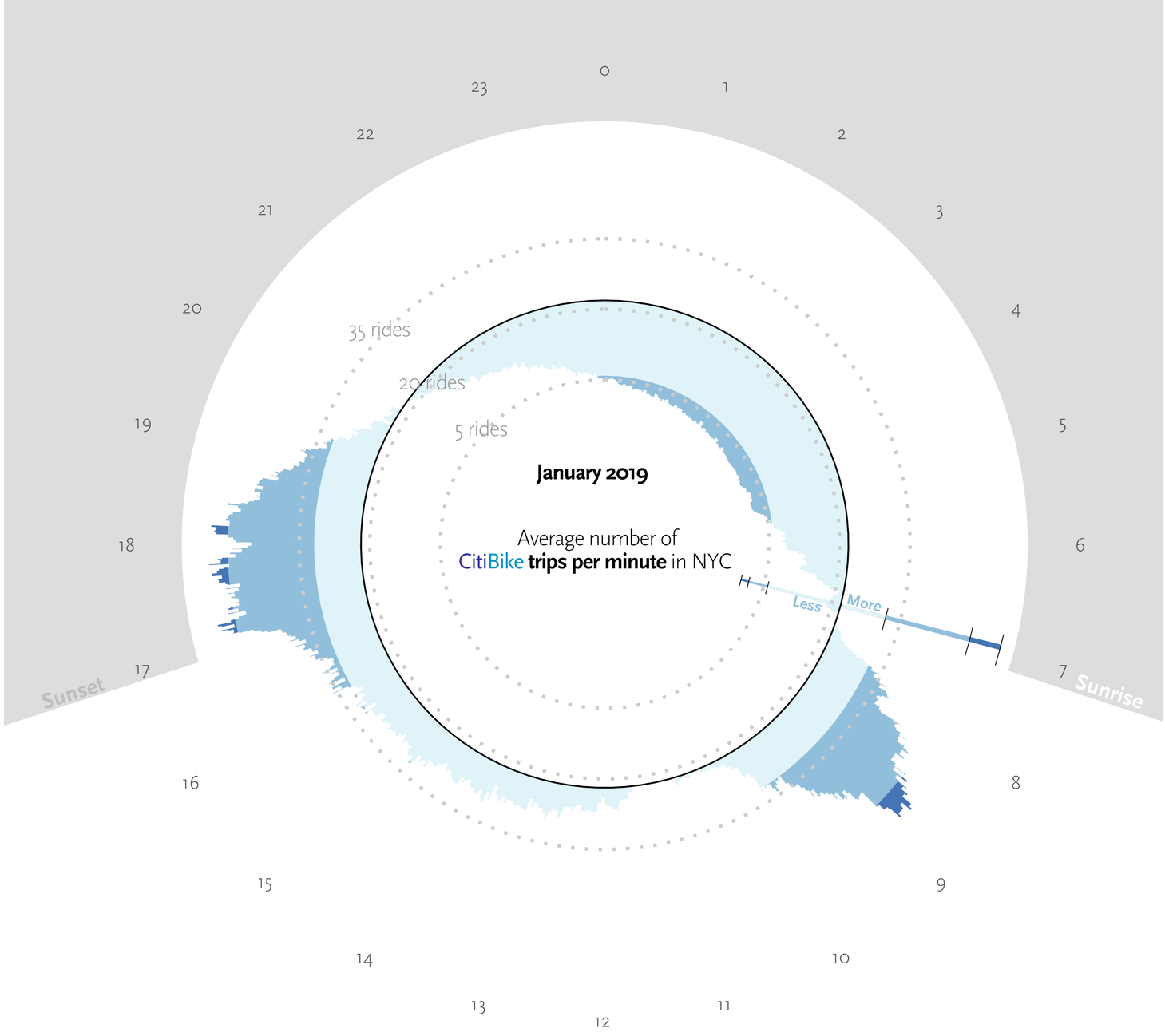
What other structure or ideas that we've learned in class so far do you see?

Why might it be more or less useful to only have images, or only have words?

Why might it be helpful to hear additional context in, say, a voice-over?

Graphics as data pictures of comparison and change

Revisiting CitiBike, showing temporal changes in activity

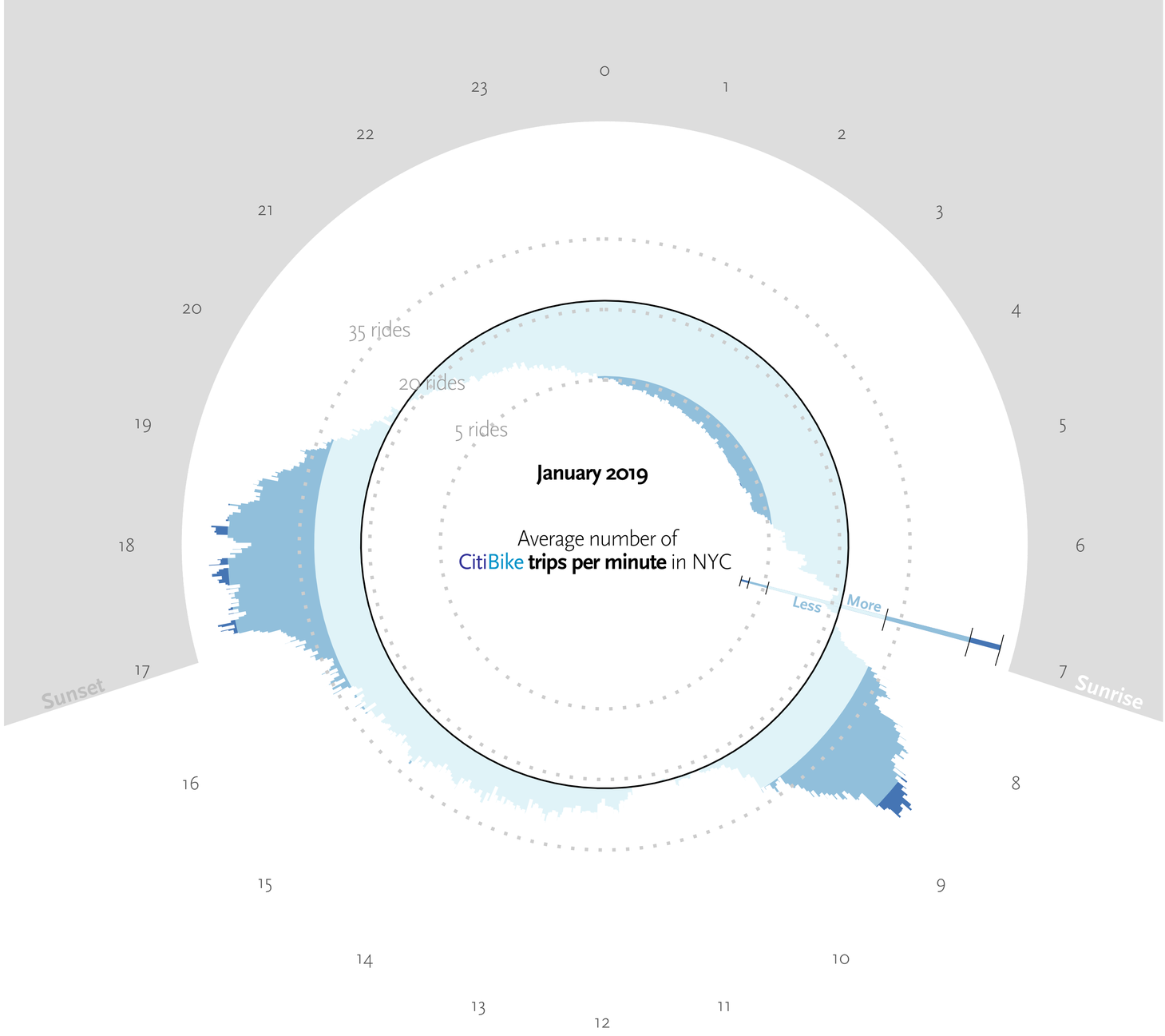


Spencer, Scott. *Exploring CitiBike Ride Data*. 28 Feb 2019. Web. https://ssp3nc3r.github.io/post/exploring-citibike-ride-data/

Creating graphics, a short, partial intro

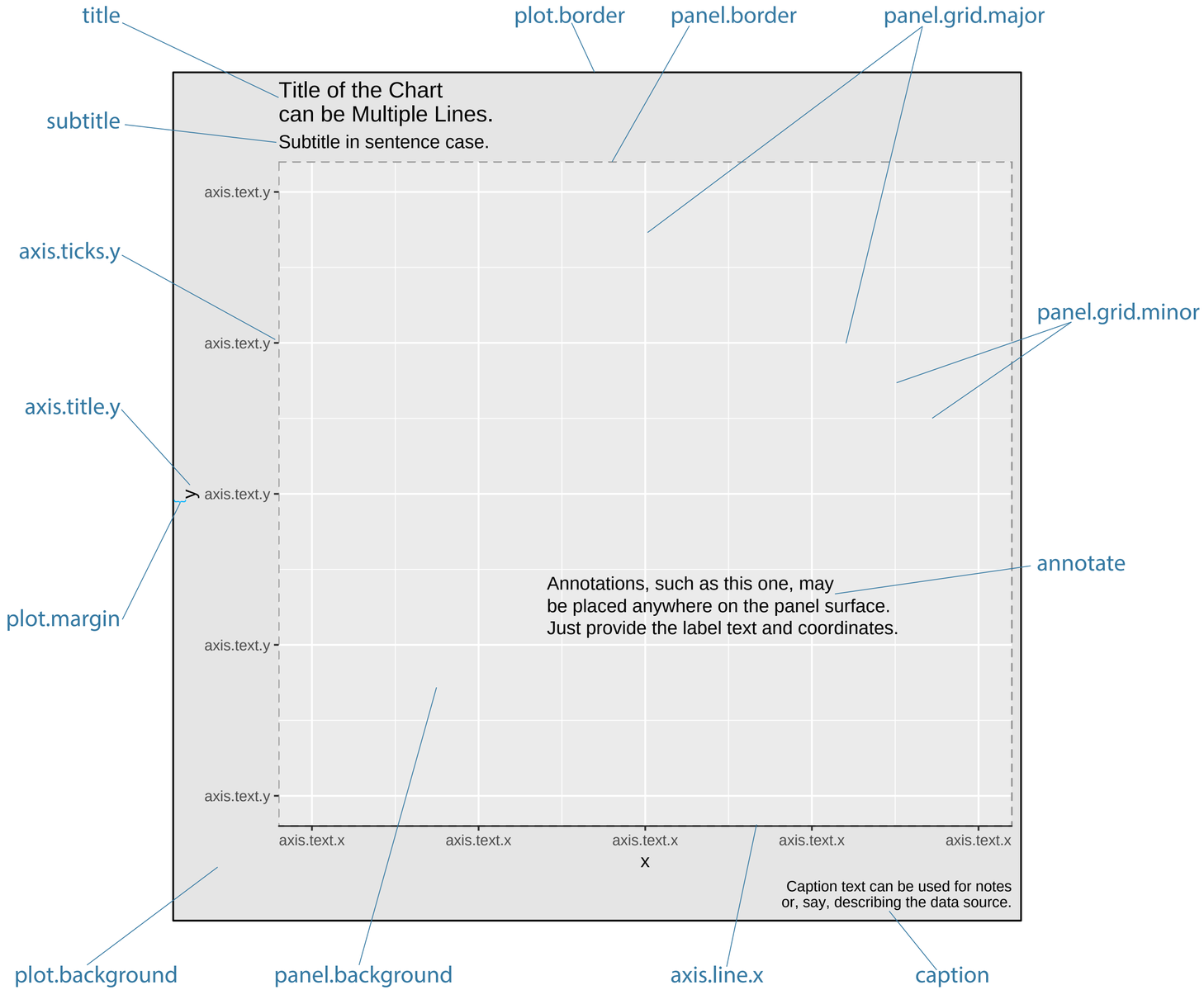
Think about graphics as layers:

here, a white layer partly masks the band of blues



Spencer, Scott. *Exploring CitiBike Ride Data*. 28 Feb 2019. Web. https://ssp3nc3r.github.io/post/exploring-citibike-ride-data/

First, think visually in describing the graphic. Then create the layers.

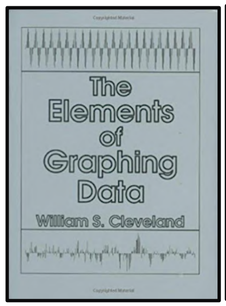


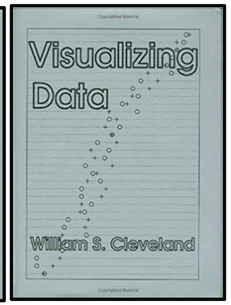

```
# load grammar of graphics
library(ggplot2)

p <-

  # functions for data ink

  ggplot(data = <data>,
         mapping = aes(<aesthetic> = <variable>,
                       <aesthetic> = <variable>,
                       <...> = <...>) +
  geom_<type>(<...>) +
  scale_<mapping>_<type>(<...>) +
  coord_<type>(<...>) +
  facet_<type>(<...>) +
  <...> +

  # functions for non-data ink

  labs(<...>) +
  theme(<...> = <...>) +
  annotate(<...>) +
  <...>
```

element_blank()
element_line(<...> = <...>)
element_rect(<...> = <...>)
element_text(<...> = <...>)

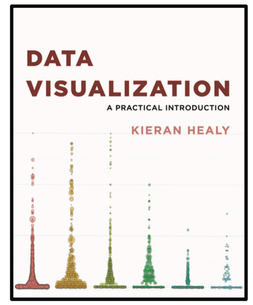
Learning & References

——— visual relationships in data ———



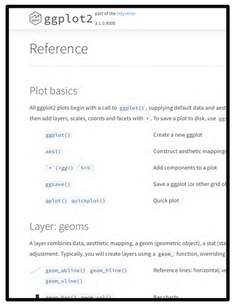


Implementation in R



socviz.co

ggplot reference



ggplot2.tidyverse.org

Spencer, Scott. *Quantitative Persuasion Amid Uncertainty*. Forthcoming. Print.

Resource for creating basic charts of graphics: R, Python, D3.js

from Data to Viz, Winner, Information is Beautiful Award

Holtz & Healy

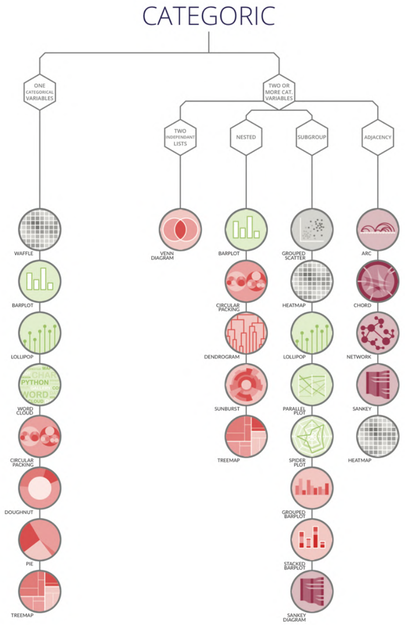
from Data to Viz

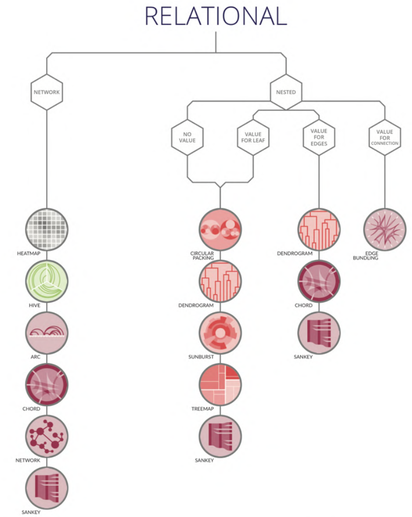
'From Data to Viz' is a classification of chart types based on input data format. It will help you find the perfect chart in three simple steps :

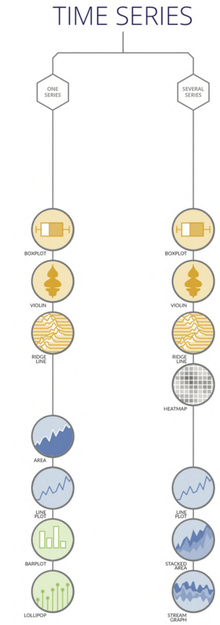
1. Identify what type of data you have.
2. Go to the corresponding decision tree and follow it down to a set of possible charts.
3. Choose the chart from the set that will suit your data and your needs best.

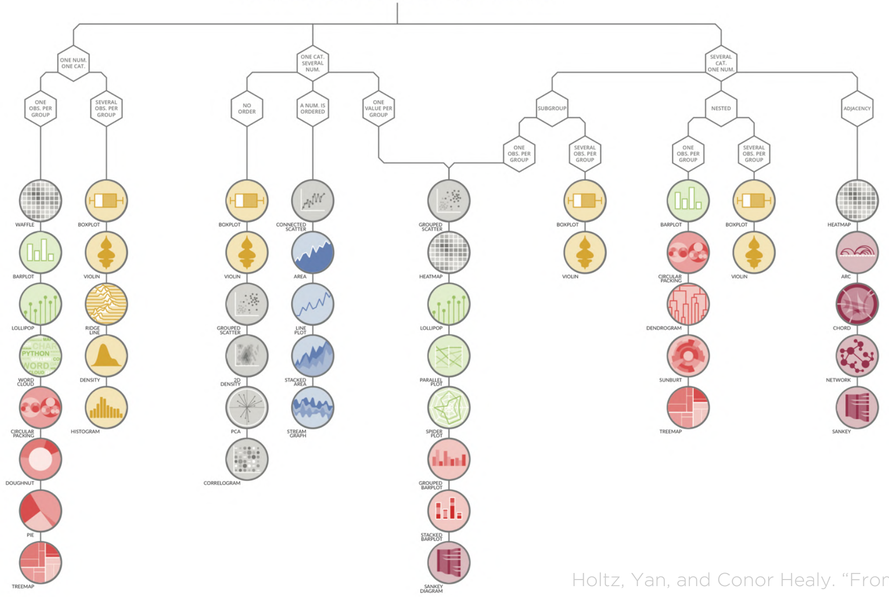
Dataviz is a world with endless possibilities and this project does not claim to be exhaustive. However it should provide you with a good starting point. For an interactive version and much more, visit:

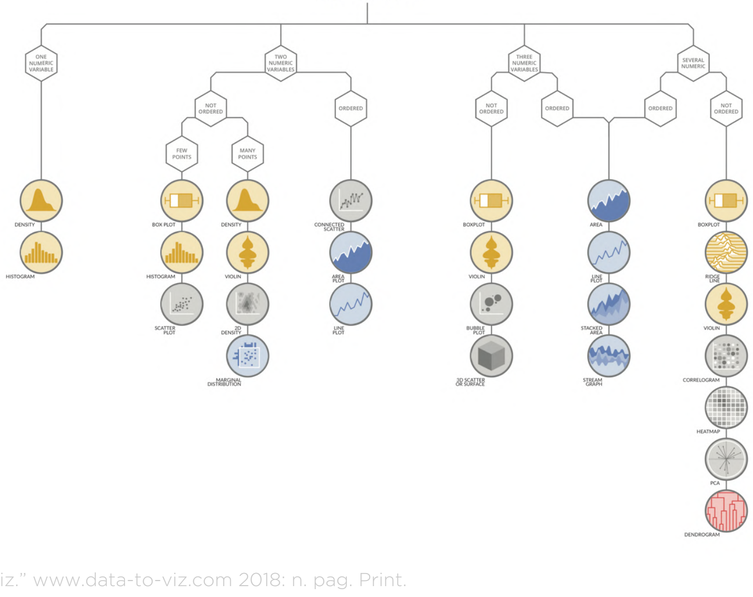
www.data-t0-viz.com











Holtz, Yan, and Conor Healy. "From Data to Viz." www.data-to-viz.com 2018: n. pag. Print.

Between now and next class

For Next Week, Module 7:

Agenda next week

Next deliverable, *draft* storyboard

Storytelling continued

Theory and best practices for the visual components of an analytics story

The minimum

Corum, Jonathan. *See. Think. Design. Produce.* 28 Mar 2016. Web. http://style.org/stdp3/

Read to understand the perspective of a leading visual story designer on the process of visual data storytelling., and meet great examples of such stories.

The author, as I mentioned, works on science visualization stories for the New York Times. So consider how his perspective may be similar to, and differ from, your audiences and purposes and what implications that may have.

Tufte, E. R. (2001). “6. Data-Ink Maximization and Graphical Design.” in *The Visual Display of Quantitative Information*. Graphics Press.

Tufte, E. R. (2001). “9. Aesthetics and Technique in Data Graphical Design.” in *The Visual Display of Quantitative Information*. Graphics Press. 1–14. Print.

Show and ask

Show and ask

Find a non-moving data visualization that is also a story—one you really like—but that you don't know how you could make it. Share the visual on Campus Wire and explain what part you don't know how you could make.

Learning with practice

Aside from your assigned projects, briefly describe one of your attempts in the past week to use any of the techniques we have discussed in the course.

See you next week!

Spencer, Scott. Illustration of Rodin's Le Penseur. Columbia University. 5 Mar 2019.