Storytelling with Data

Module 9: Frame the story — consumers and information graphics

Scott Spencer
Faculty and Lecturer
Columbia University

Unanswered, or new, questions from discussion?

Agenda

Upcoming deliverable

Today’s objectives

Information graphics

Considering uncertainty

Upcoming deliverables

Upcoming deliverables

Critique exemplary information graphic — apply techniques and best practices discussed in material and class to critique an infographic you have identified as exemplary.

Final information graphic — reframe your story, building off the messages you built for the marketing team to craft an infographic that displays the results of your analytic work in a way that is accessible, engaging, and exciting for a **general or consumer audience**.

Today’s Objectives

Objectives

1 Use information graphics to tell a story

2 Communicating uncertainty with data visualization

Information graphics

Too broad, generic

infographic n. a visual image such as a chart or diagram used to represent information or data in an easily understandable form.

— "info-, comb. form." OED Online. Oxford University Press, March 2019. Web. 24 March 2019.

Want bad examples? Just google.

Most collect trivia. They lack: change, complexity, story, design.

Without a narrative, it's just trivia, list of facts:

https://public.tableau.com/profile/datavizard#!/vizhome/2017MLBPitcherHeatmaps/MLB

Data-Driven Storytelling

Riche, co-editors

The editors are researchers and professors with focuses on human-computer interaction and information visualization.

> We differentiated … "**list of fact**" infographics from the infographics medium as a whole, and chose to **exclude them** because this specific submedium **lacks authorial narrative**.

Riche, Nathalie Henry et al. *Data-Driven Storytelling*. CRC Press, 2018. Print.

Dashboards, while sometimes a part of an infographic, may lack narrative or story on their own.

Subscriber Churn Analysis

Subscriber activity - All

Net subscriber activity by division

Details

		Gained	Lost	Net	Running total
West	January	80	0	80	80
	February	80	-15	65	145
	March	90	-30	60	205
	April	120	-25	95	300
	May	100	-50	50	350
	June	119	-77	42	392
	July	75	-45	30	422
	August	119	-77	42	464
	September	90	-30	60	524
	October	80	-15	65	589
	November	80	-20	60	649
	December	90	-30	60	709
	Total	**1,123**	**-414**	**709**	
Central	January	60	0	60	60
	February	85	-45	40	100
	March	80	-27	53	153
	April	90	-17	73	226
	May	120	-33	87	313
	June	45	-80	-35	278
	July	75	-45	30	308
	August	45	-80	-35	273
	September	80	-27	53	326
	October	85	-45	40	366
	November	60	-35	25	391
	December	80	-27	53	444
	Total	**905**	**-461**	**444**	
East	January	70	0	70	70
	February	80	-90	-10	60
	March	100	-30	70	130
	April	110	-45	65	195
	May	70	-95	-25	170
	June	45	-33	12	182
	July	50	-110	-60	122
	August	99	-34	65	187
	September	112	-34	78	265
	October	99	-88	11	276
	November	55	-65	-10	266
	December	110	-45	65	331
	Total	**1,000**	**-669**	**331**	
Grand Total		**3,028**	**-1,544**	**1,484**	

Wexler, Steve, Jeffrey Shaffer, and Andy Cotgreave. *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios*. Wiley, 2017. Print.

Malofiej—Infographics World Summit

"The Pulitzer Prizes of Infographics"

Interviews of Malofiej speakers on infographics

Gregor Aisch

Infographics is an abbreviated form of "information graphics". It seems to mean a lot of different things to different people. I rarely use the term.

Federica Fragapane

A visual translation of data and information: a language to communicate topics, contents and **stories** to people.

Laris Karklis

Infographics is . . . using visuals to **tell a story**.

Nadieh Bremer

Infographics ... combine graphical elements, such as a drawn portion of an animal, human, map, etc. with small mini **data visualizations** (a small bar chart for example) and **annotations** to **tell a story**.

Malofiej. Interviews. (2019). Web. http://www.malofiejgraphics.com/category/interviews/

We want information graphics to …

Tell a complete story where the purpose is to inform, entertain or persuade the audience (to act). It should:

simple, focused messages

new, surprising information

credible data sources

visually coherent, integrated

use comparisons for context, meaning

principles of information design, organized

Information is Beautiful

McCandless

Founder of Information is Beautiful, David is a writer, designer, creative director and artist.

McCandless, David. *Information is Beautiful*. Web. http://informationisbeautiful.net

Information graphics, examples for discussion

Winner, Information is Beautiful Award

Bremmer

Previously an astronomer and analyst at a consulting company, Nadieh is a powerhouse freelance data visualization designer, and winner of numerous awards.

TOP 2000 ❤ 70's & 80's

Since 1999 the 2000 most popular songs of all time, as voted by the show's audience, are played on Dutch national Radio 2 in a yearly marathon. The 2000 songs are on the air between noon on December 25th until New Year's Eve and over half of the Dutch population listens to the Top 2000 each year.

Each ● to the right represents a song in the Top 2000. It is placed according to its year of release. In the legend below you can see what the size and color of a song means.

The bulk of the songs and most of the top 10 are from the 70's & 80's...

The charts on the right represent all 2000 songs from 3 past editions of the Top 2000 (held in 2000, 2005, 2010) and the most recent 2016 edition.

The songs are stacked according to their year of release. The higher a rectangle, the more songs that are in the Top 2000 list from that release year.

The black dotted line represents a smoothed curve over all 2000 songs. This makes the comparison between the 4 charts easier.

Spread across release years of the 2000 songs

For 4 editions of the Top 2000

But they're losing tracks to the new Millenium

It makes sense that the Top 2000 will be more spread out for each new edition, since there are more songs to choose from. However, if we compare the distributions of the Top 2000 songs over 4 editions, we see that, especially, the 90's has been gaining a lot of popularity.

Even though all songs from the 90's were out in the 2000 edition, only a few songs from that decade were chosen. Whereas in the 2016 edition the number of songs from the 90's has risen significantly. This could be due to a new generation who has grown up during the 90's taking over from those who voted in the early 2000's (who apperantly didn't appreciate the new music).

Created by Nadieh Bremer | VisualCinnamon.com for the December edition of data sketch|es

Visit tinyurl.com/2016top2000 for the interactive visual and see the name & title of each song

Data | Top 2000 list from Radio 2 | Top 40 info from Mediamarkt's Top 40

The finished design: Bremer, Nadieh. *Top 2000*. Web. https://www.visualcinnamon.com/portfolio/top2000
Her design process: Bremer, Nadieh. *Music*. Data Sketch | es. http://www.datasketch.es/december/

Audience? Does this infographic seem designed to communicate with an identified audience? If so, who?

Purpose? Do you see a purpose? If so, what is it trying to inform, entertain, or persuade the audience to act? Or something else?

Narrative? Does this infographic have a narrative? If so, what structure?

Comparison or change? Does the infographic describe comparisons or change? If so, what?

Color, coherency? Is color used? If so, for what purpose(s) are its hue, chroma, or luminance used?

Hierarchy, annotation? Does it have a hierarchy of information? If so, how is that hierarchy made? Are data encodings explained? If so, how?

Layering, layout? Is the information organized? If so, how?

Credibility, transparency? Are data sources identified, explained? Limitations or issues discussed?

The finished design: Bremer, Nadieh. *Top 2000*. Web. https://www.visualcinnamon.com/portfolio/top2000
Her design process: Bremer, Nadieh. *Music*. Data Sketch | es. http://www.datasketch.es/december/

Winner, Information is Beautiful Award

Bremmer

Previously an astronomer and analyst at a consulting company, Nadieh is a powerhouse freelance data visualization designer, and winner of numerous awards.

Bremer, Nadieh. *Top 2000*. Web. https://www.visualcinnamon.com/portfolio/top2000

Info We Trust
How to inspire the world with data

Andrews

He is a data storyteller. He book is an adventure exploring how to inspire the world with data. RJ is the creator of www.infowetrust.com, where he makes available some of his data stories.

Andrews, R J. *Info We Trust: How to Inspire the World with Data*. Wiley, 2019. Print.

Summary of long term budget report

CBO

Even the Congressional Budget Office relies on infographics to convey large tomes of information in summary form.

Congressional Budget Office. *The 2012 Long-Term Budget Outlook: Infographic*. 5 June 2012. Web. https://www.cbo.gov/publication/43289

Winner, Information is Beautiful Award

Toshkov

Associate Professor at the Institute of Public Administration, won an award for this infographic, which was made using R.

Toshkov. Information is Beautiful, World Data Visualization Prize. 2019. Web. https://informationisbeautiful.net/2019/winners-of-the-world-data-visualization-prize/

Graphics portfolio, NYT Infographic

Aisch

Gregor was graphics editor at the NY Times, and recipient of infographics awards at Malofiej.

Aisch, Gregor. et al. *The Cost of Mobile Ads on 50 News Websites*. NY Times, 1 Oct 2015. Web. https://www.nytimes.com/interactive/2015/10/01/business/cost-of-mobile-ads.html

Aisch, Gregor. *Driven By Data*. Graphics portfolio of Gregor Aisch, a former graphics editor at The New York Times (2014-2017), now Co-Founder/CTO at Datawrapper. Web. https://driven-by-data.net

CitiBike example, infographic

Francisco, Juan. *An investigation into ways to rebalance CitiBike stations throughout New York City*. 2014. Columbia University Center for Spatial Research. Web. http://www.c4sr.net/projects/citibike-rebalancing-study

Longlisted, Information is Beautiful Awards

Spencer

Spencer, S. *Ride Against the Flow*. Information is Beautiful Awards 2019. Showcase. https://www.informationisbeautifulawards.com/showcase/4367-ride-against-the-flow

Longlisted, Information is Beautiful Awards

Spencer

Maths to create data encodings at each dock station

Basic math can help when making custom graphics. An excellent, very short reference on basic algebra, geometry, trigonometry: Simmons, George F. *Precalculus Mathematics in a Nutshell*. Barnes & Noble Books, 1987. Print.

Opacity is the lack of transparency. Opacity was lowered so that a single unavailable bike or full station would not be very noticeable but several markings at that time (on different days) would increase opacity, making the marking brighter and thus increase its visual importance.

Spencer, S. *Ride Against the Flow*. Information is Beautiful Awards 2019. Showcase. https://www.informationisbeautifulawards.com/showcase/4367-ride-against-the-flow

Group work, get help on your information graphics

Info We Trust
How to inspire the world with data

Andrews

He is a data storyteller. He book is an adventure exploring how to inspire the world with data. RJ is the creator of www.infowetrust.com, where he makes available some of his data stories.

Workshopping

When we workshop a project, we **scrutinize it to make it better**.

The **creator** is the most important editor. While you are editing, you are the most ruthless critic.

The next most important critic is you. **Future you**. Allow some time to pass between creation and evaluation.

Beyond you, the creator, the most valuable critic is **anyone else willing to engage with the work**. Your critique is clouded by your expertise and familiarity with the story.

Any random reviewer may not be able to tell you how to fix your story. But everyone can **help you identify where the story is going off the rails**.

Pay attention to **where others get lost or bored, or misunderstand what is going on**.

Andrews, R J. *Info We Trust: How to Inspire the World with Data*. Wiley, 2019. Print.

Get specific

Audience? Does this infographic seem designed to communicate with an identified audience? If so, who?

Purpose? Do you see a purpose? If so, what is it trying to inform, entertain, or persuade the audience to act? Or something else?

Narrative? Does this infographic have a narrative? If so, what structure?

Comparison or change? Does the infographic describe comparisons or change? If so, what?

Color, coherency? Is color used? If so, for what purpose(s) are its hue, chroma, or luminance used?

Hierarchy, annotation? Does it have a hierarchy of information? If so, how is that hierarchy made? Are data encodings explained? If so, how?

Layering, layout? Is the information organized? If so, how?

Credibility, transparency? Are data sources identified, explained? Limitations or issues discussed?

Communicating uncertainty

Where's my data? Evaluating Visualizations with Missing Data

Song & Szafir

The authors were academics at the University of Colorado Boulder.

Visual choices for missing values on perception of data quality

Choice of missing-data visualization alters perceptions of data quality

Perceived data quality and confidence generally **degrade** as the amount of missing data increases.

Data visualized by **highlighting missing values** tends to be seen as **higher quality than downplay or information removal**.

Information removal can significantly **degrade** perceptions of data quality, and confidence. These methods even lead to incorrect responses if missing values break the visual continuity of a visualization.

Linear interpolation leads to **higher perceptions** of quality and confidence in analysis.

Song, Hayeong, and Danielle Albers Szafir. "*Where's My Data? Evaluating Visualizations with Missing Data.*" IEEE Transactions on Visualization and Computer Graphics 25.1 (2018): 914–924. Web.

(a) Data Absent (b) Color Points (c) Color Points & Line Gradients (d) Connected Error Bars (e) Disconnected Error Bars (f) Unfilled Points (g) Unfilled Points & Line Gradients

Fig. 4: We tested seven different methods for visualizing missing values in line graphs manipulating both point and line appearance: two highlighting missing values, two downplaying missing values, two annotating missing values, and one removing missing values. .

(a) Data Absent (b) Color Bars (c) Sketched Bar (d) Bars with Error Bars (e) Points with Error Bars (f) Bars with Dashed Outlines (g) Alpha Gradient Bars

Fig. 6: We tested seven different methods for visualizing missing values in bar charts: one highlighting missing values, three downplaying missing values, two annotating missing values, and one removing missing values.

Song, Hayeong, and Danielle Albers Szafir. "*Where's My Data? Evaluating Visualizations with Missing Data.*" IEEE Transactions on Visualization and Computer Graphics 25.1 (2018): 914–924. Web.

Imagining Replications: Graphical Prediction & Discrete Visualizations Improve Recall & Estimation of Effect Uncertainty

Imagining Replications

Hullman & co-authors

The authors are part of an academic group formally studying the perception of uncertainty, among other things.

Sketching perception of uncertainty before seeing results improves accuracy

Crowd-sourced study that evaluated the impact of an interactive, graphical uncertainty prediction technique for communicating uncertainty in results. Users sketched their prediction of the uncertainty in experimental effects prior to viewing the true sampling distribution from an experiment.

Visualizing uncertainty as set of discrete outcomes improved recall

Visualizing uncertainty as a set of discrete outcomes, as opposed to a continuous probability distribution, can improve recall of a sampling distribution from a single experiment.

Fig. 1. Discrete and continuous elicitation interface used by participants in our study to predict replication uncertainty.

Hullman, Jessica et al. "*Imagining Replications: Graphical Prediction & Discrete Visualizations Improve Recall & Estimation of Effect Uncertainty*." IEEE Transactions on Visualization and Computer Graphics 24.1 (2017): 446–456. Web.

Fig. 2. A depiction of distributions relevant to replication uncertainty, including those based on perfect knowledge of the world (left) and those derived from samples obtained in experimentation (right).

Hullman, Jessica et al. "*Imagining Replications: Graphical Prediction & Discrete Visualizations Improve Recall & Estimation of Effect Uncertainty*." IEEE Transactions on Visualization and Computer Graphics 24.1 (2017): 446–456. Web.

Let’s look ahead

For next lecture, Module 10:

Agenda next week

Next deliverable, **critique** of infographic

More on information graphics and uncertainty

The minimum

Kay, Matthew et al. *When (Ish) Is My Bus? User-Centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems*. New York, New York, USA: ACM Press, 2016. 5092–5103. Web.

Consider how test subjects responded to varying visual representations of uncertainty, and any implications for your projects.

Fischhoff, Baruch. *Communicating Uncertainty: Fulfilling the Duty to Inform*. Issues in Science and Technology 28.4 (2012): 63–70. Print.

Consider the author's view on issues with communicating uncertainty, and his suggestions for addressing these issues.

Wainer, Howard. *The Most Dangerous Equation*, in Chapter 1, *Picturing the Uncertain World*. Princeton University Press, 2009. Print.

Consider the issues he raises of when variation is misunderstood, and the exemplary implications.

Wacharamanotham, Chat et al. "Special Interest Group on Transparent Statistics Guidelines.." CHI Extended Abstracts (2018): n. pag. Print and web. http://transparentstatistics.org

For online discussion

Question | Question.

Question | Question.

See you next week!

Spencer, Scott. *Illustration bike over dandelion*. 10 Apr 2019.

Supplemental material

Winner, Information is Beautiful Award

Lupi, Fragapane

The co-authors work together at Lupi's design firm, whom we've previously met when considering this visualization.

Lupi, Giorgia. *Visual Data - La Lettura*. Web. http://giorgialupi.com/lalettura

Diagram

NEWS IN PERSPECTIVE

Brazilian Population Grows More in Prisons

With 258 inmates per 100,000 people, Brazil has one of the largest prison populations.

Alberto Cairo, Humberto Maia Junior

BRAZIL IS THIRD in the worldwide ranking of countries with the largest prison populations. Data released by Ministry of Justice this week reveal that Brazil had 494,237 inmates in June 2010. Only the US, with 2.3 million, and China, with 1.6 million, have larger prison systems. According to the United Nations Office on Drugs and Crime, Brazil is one of the countries where the prison population has grown most rapidly in the world. Between 2007 and 2010, it increased by 13.3% (see chart on the right), while the overall population growth was 0.98% in the same period. Even considering the new prisons that are being built, Brazil has reached a deficit of 200,000 jail spaces.

Sources: Ministério da Justiça, IBGE, United Nations Office on Drugs and Crime

1 Between 1997 and 2007, Brazil experienced the fifth largest increase in prison population in the world.

2 The trend continued since 2007...

The rate of inmates per 100,000 people has gone down in just eight of the states plus the Federal District.

3 ...but government has not been able to create jail spaces at the same pace

Brazil's jails are packed full. A comparison between the variation of inmates and the number of jail spaces reveals that neither the federal government nor the states have been able to avoid overcrowding. The only exception is Rio de Janeiro.

Cairo, Alberto. *Brazilian Population Grows More in Prisons*. Época Magazine. 2010. Print.

GRAPHIC AND ART (OIL ON CANVAS, 36 × 60 INCHES): ALBERTO LUCAS LÓPEZ, NGM STAFF. TEXT: EVE CONANT, NGM STAFF. PHOTO (PAINTING): MARK THIESSEN, NGM STAFF
SOURCES: ENRIQUE MALLÉN, ONLINE PICASSO PROJECT; *PABLO PICASSO: A RETROSPECTIVE*, MUSEUM OF MODERN ART, NEW YORK. A PIECE CAN BE INCLUDED IN SEVERAL CATEGORIES. AREA SIZES ARE APPROXIMATE.

OBJECT LESSONS

FRAMES OF MIND

'The artist is a receptacle for emotions that come from all over the place: from the sky, from the earth, from a scrap of paper, from a passing shape, from a spider's web.'

Picasso drew inspiration from mythology, from war, from those who surrounded him, even—he proclaimed—from spiderwebs. The result is tens of thousands of works that seem to touch on countless topics. But many of the subjects that fascinated him, including death and sexuality, repeatedly surface in his art throughout the decades. Those recurring themes are grouped in a sampling of some 8,000 of Picasso's works, artistically rendered here. They can be found in portraits that radically blur the line between subject and painter, and in an array of arresting symbols such as the Minotaur, bottle, and harlequin.

PICASSO BY THEME

Each of the 12 themes below is further divided into artistic subcategories, some of which include the titles of individual works. The size of each category reflects the number of artworks on a given theme.

1. ABSTRACT
2. FIGURE
3. STILL LIFE
4. ROOM
5. BODY
6. OBJECT
7. LOCATION
8. THEATER
9. CIRCUS
10. MYTHOLOGY
11. LITERATURE
12. ANIMAL

https://www.nationalgeographic.com/magazine/2018/05/genius-picasso-art-categories-infographic/#close

Tufte, Edward R. *The Visual Display of Quantitative Information*. Second. Graphics Press, 2001. Print.

Table 20.1 **Napoleon's March Data**

lonc	latc	city	lont	temp	date	lonp	latp	survivors	direction	group
24.0	55.0	Kowno	37.6	0	Oct 18	24.0	54.9	340,000	A	I
25.3	54.7	Wilna	36.0	0	Oct 24	24.5	55.0	340,000	A	I
26.4	54.4	Smorgoni	33.2	–9	Nov 9	25.5	54.5	340,000	A	I
26.8	54.3	Molodexno	32.0	–21	Nov 14	26.0	54.7	320,000	A	I
27.7	55.2	Gloubokoe	29.2	–11		27.0	54.8	300,000	A	I
27.6	53.9	Minsk	28.5	–20	Nov 28	28.0	54.9	280,000	A	I
28.5	54.3	Studienska	27.2	–24	Dec 1	28.5	55.0	240,000	A	I
28.7	55.5	Polotzk	26.7	–30	Dec 6	29.0	55.1	210,000	A	I
29.2	54.4	Bobr	25.3	–26	Dec 7	30.0	55.2	180,000	A	I
30.2	55.3	Witebsk				30.3	55.3	175,000	A	I
30.4	54.5	Orscha				32.0	54.8	145,000	A	I
30.4	53.9	Mohilow				33.2	54.9	140,000	A	I
32.0	54.8	Smolensk				34.4	55.5	127,100	A	I
33.2	54.9	Dorogobouge				35.5	55.4	100,000	A	I
34.3	55.2	Wixma				36.0	55.5	100,000	A	I
34.4	55.5	Chjat				37.6	55.8	100,000	A	I
36.0	55.5	Mojaisk				37.7	55.7	100,000	R	I
37.6	55.8	Moscou				37.5	55.7	98,000	R	I
36.6	55.3	Tarantino				37.0	55.0	97,000	R	I
36.5	55.0	Malo-jarosewli				36.8	55.0	96,000	R	I
						35.4	55.3	87,000	R	I
						34.3	55.2	55,000	R	I
						33.3	54.8	37,000	R	I
						32.0	54.6	24,000	R	I
						30.4	54.4	20,000	R	I
						29.2	54.3	20,000	R	I
						28.5	54.2	20,000	R	I
						28.3	54.3	20,000	R	I
						27.5	54.5	20,000	R	I
						26.8	54.3	12,000	R	I
						26.4	54.4	14,000	R	I
						25.0	54.4	8,000	R	I
						24.4	54.4	4,000	R	I
						24.2	54.4	4,000	R	I
						24.1	54.4	4,000	R	I
						24.0	55.1	60,000	A	II
						24.5	55.2	60,000	A	II
						25.5	54.7	60,000	A	II
						26.6	55.7	40,000	A	II
						27.4	55.6	33,000	A	II
						28.7	55.5	33,000	A	II
						28.7	55.5	33,000	R	II
						29.2	54.2	30,000	R	II
						28.5	54.1	30,000	R	II
						28.3	54.2	28,000	R	II
						24.0	55.2	22,000	A	III
						24.5	55.3	22,000	A	III
						24.6	55.8	6,000	A	III
						24.6	55.8	6,000	R	III
						24.2	54.4	6,000	R	III
						24.1	54.4	6,000	R	III

Wilkinson, Leland. *The Grammar of Graphics*. Second. Chapter 20. Springer, 2005. Print.

Winner, Information is Beautiful Award

TASS

TASS Russian News Agency. 1812: *When Napoleon Ventured East*. 2017. Web. https://1812.tass.ru/en#content