

Storytelling *with Data*

Visual Analytics and User Experience Design
(IT4031)



Lasitha Petthawadu

Associate Technical Architect
99x Technology

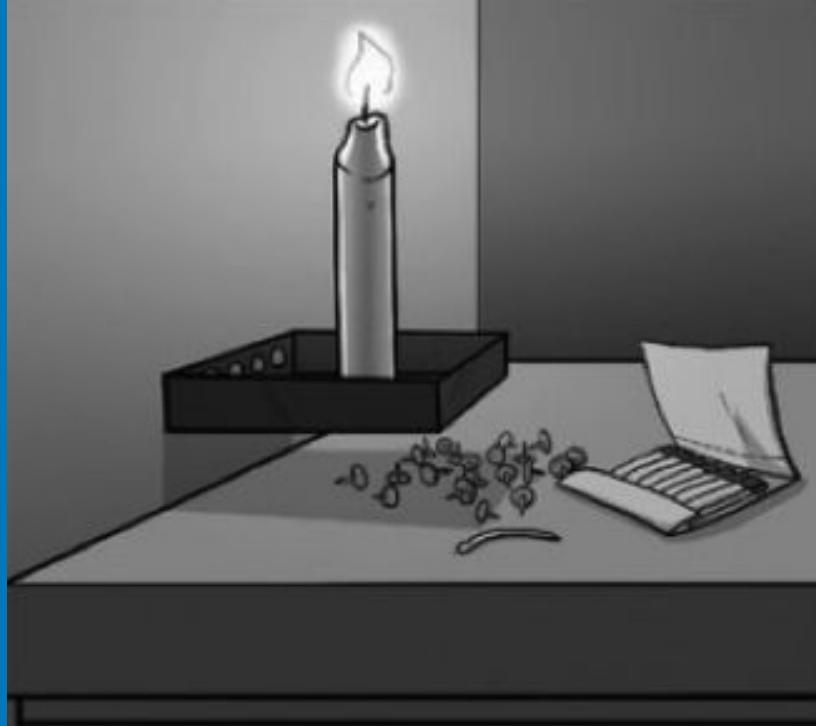
Challenge of the Day!

How do you stick the candle to the wall?



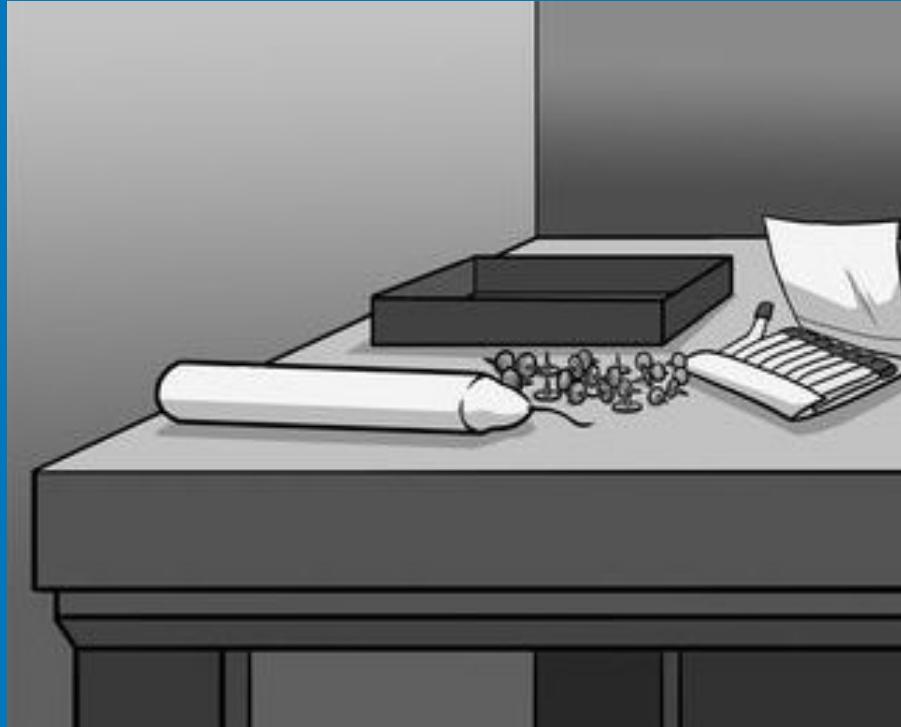
Challenge of the Day!

Solution of the candle problem



Challenge of the Day!

Candle problem for dummies





Why Storytelling?

Stories are the single most effective way of delivering a message.

Major decisions are made often emotionally than logically.

- As confirmed by Neuroscientists

“People hear statistics,
But they feel stories.”

Brent Dykes | Director of Data Strategy, Domo

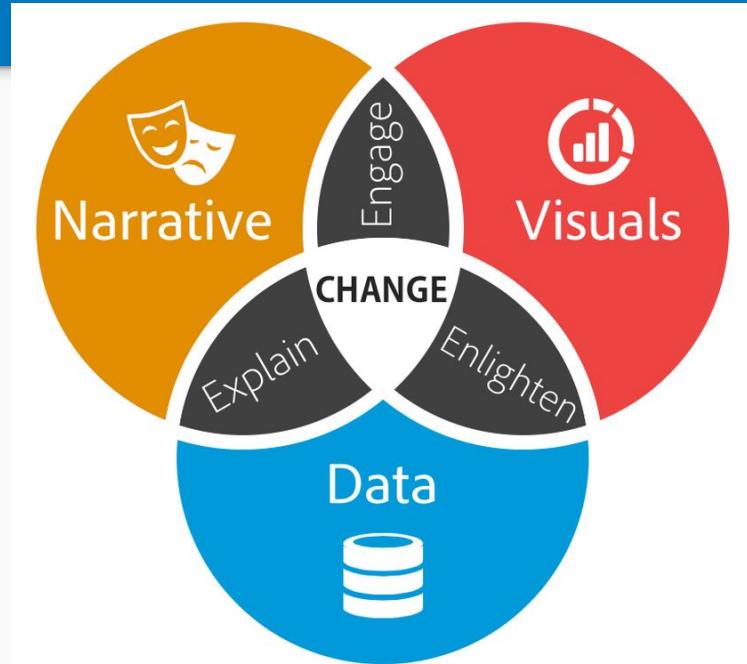
Why Storytelling?

To present your data in
a meaningful manner

Why Storytelling?

To sell *amazing*
products that people
connect with

Data Storytelling in a Nutshell



<https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/>

Types of Data Storytelling

Sales Pitch

Infographics

News
Articles

Product
Presentation

Research
Reports

Whitepapers

Data Storytelling Success Path



<https://www.nugit.co/what-is-data-storytelling/>

Data Storytelling

Understanding Context

WHO

Understanding You and your audience.

WHAT

Understand the message your audience would want to hear from your data.

Data Storytelling

Understanding Context

WHY

Understand why this message
should be important,
building stories depend on the
importance.

Data Storytelling

Understanding Context

Build context based on personalization, connecting historical, events and occurrences that are relatable to the audience.

Data Storytelling

Building Narratives

THEME/TITLE is Important

Engage your audience by an appealing theme!

Data Storytelling

Building Narratives

Communicate the
theme/title **first**

Strongly and Clearly



Data Storytelling

Building Narratives

Start small and **Build UP!**

Data Storytelling

Building Narratives

Make sure audiences are kept in
curiosity until the **climax** of the
narrative.

Data Storytelling

Effective Visualizations

Choose the right visualization and make
your data POP out!

Don't be limited by the tools you have!

“

*The most powerful person in
the world is the **storyteller**.*

*The storyteller sets the
vision, values, and agenda of
an entire generation that is to
come.*

- Steve Jobs

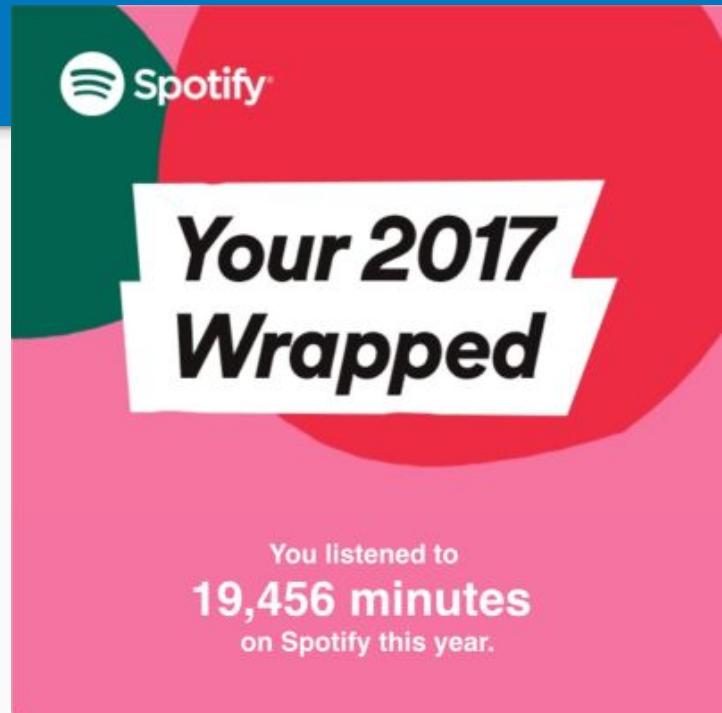




Breaking down Steve Jobs Iphone Launch Presentation

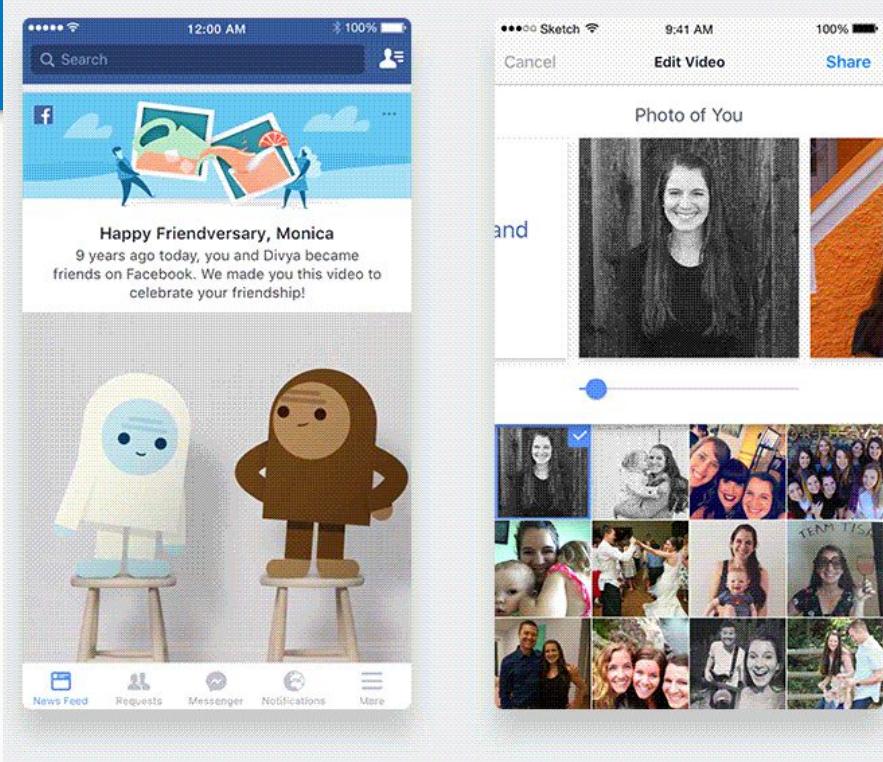


Personalizing Product Interactions



<https://www.nugit.co/what-is-data-storytelling/>

Personalizing Product Interactions



<https://dribbble.com/shots/2755733-Facebook-new-Friendversary-Video>



Historical Visualization that tell Stories

*Historical visualizations
have told great stories.*

*They have helped to find
causes and have even
prevented loss of lives.*

Charles Minard's 1869 chart

The story of Napoleon's 1812 Russian campaign army

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
Dessinée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

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 travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à desser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Leguer, de Fezenac, de Chambray et le journal intime 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 en Mobilow et qui rejoignirent l'armée vers Orscha et Witebsk, avaient toujours marché avec l'armée.

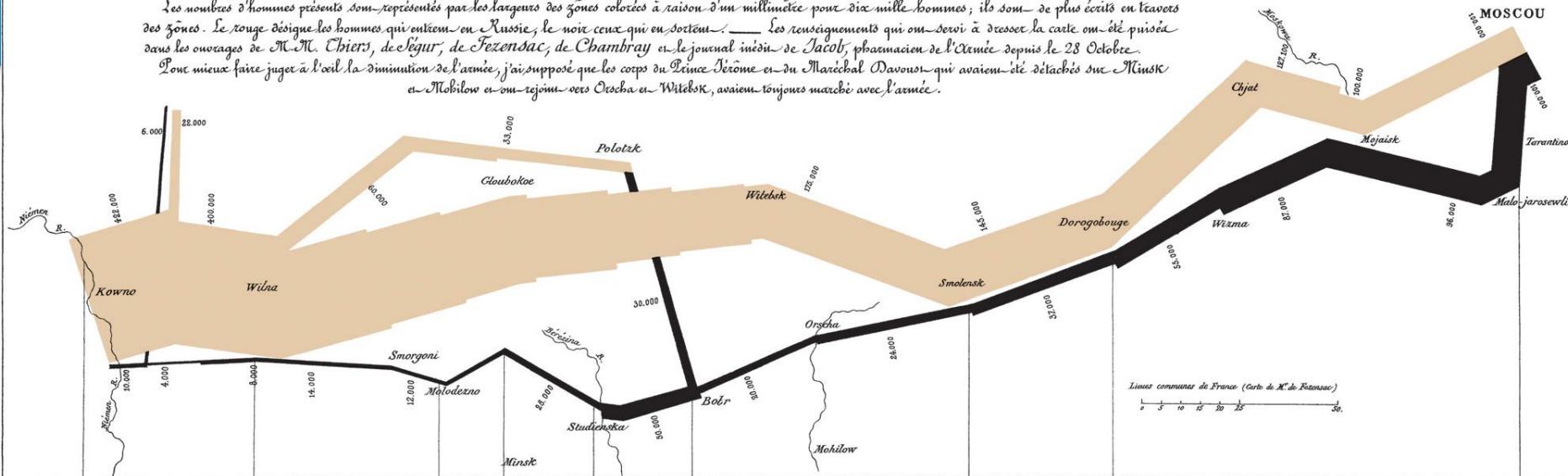
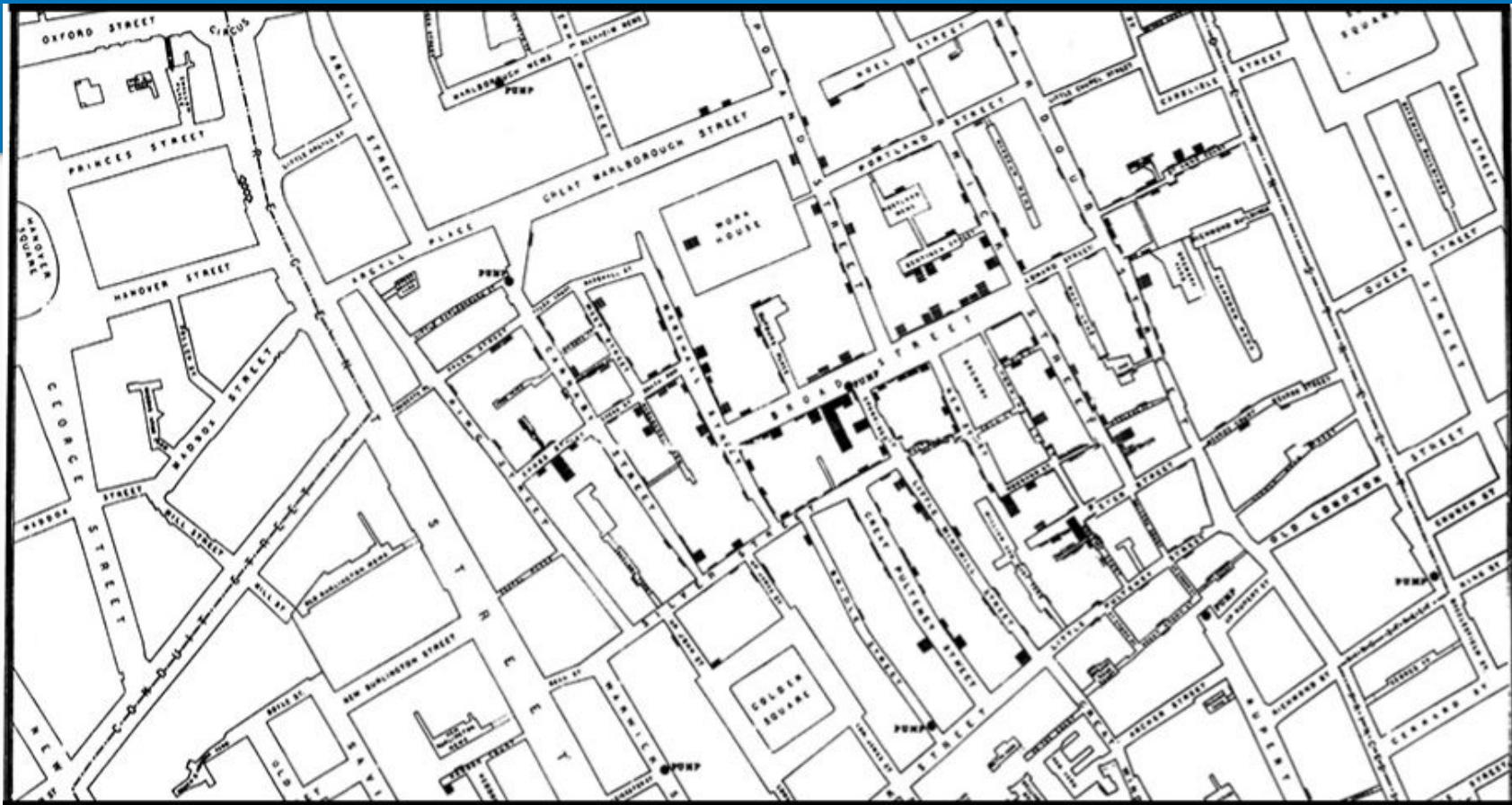


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Nieman, gelé.

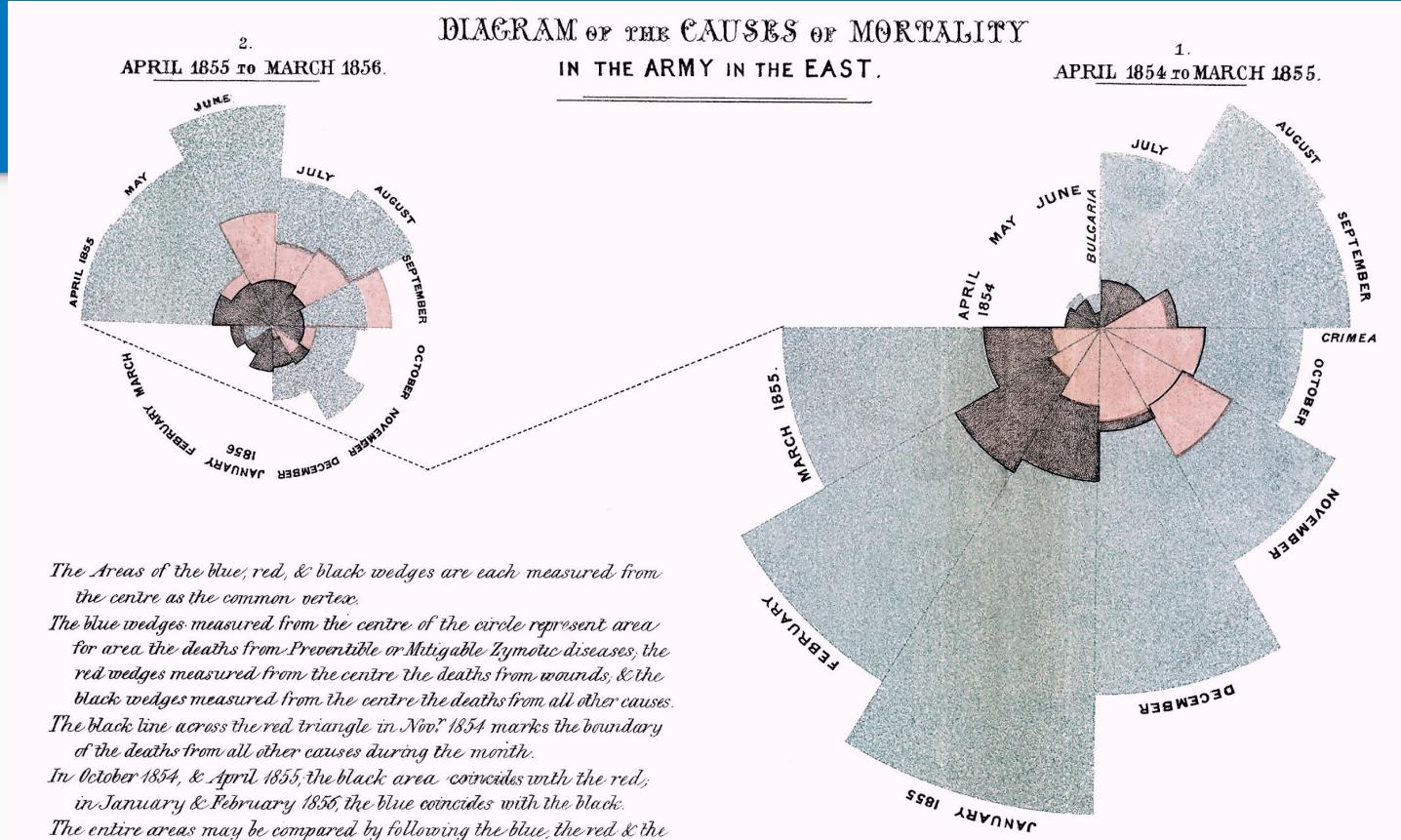


Physician John Snow Cholera Map and the Broad Street pump



Florence Nightingale

Diagram of the causes of mortality



Theory of Data Graphics

A close-up photograph of a deck of Theory11 Premium Playing Cards. The cards are dark with gold and silver accents. The brand name "Theory11" is prominently displayed in the center, flanked by "PREMIUM PLAYING CARDS" and "MADE IN USA - REC TRADEMARK". The background is blurred, showing a warm, bokeh-lit environment.

Theory of data graphics
leans on these concepts:

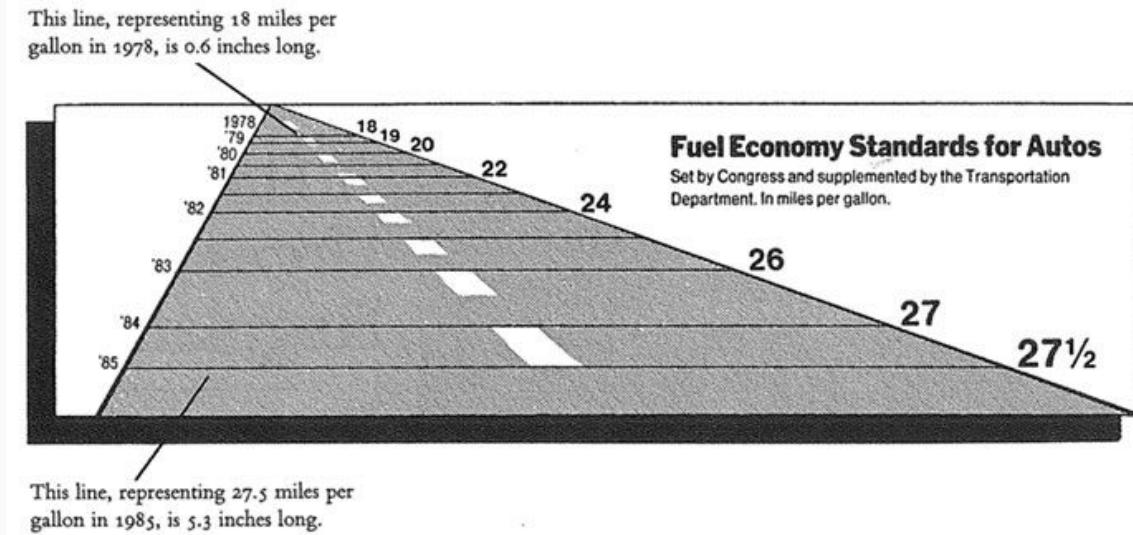
- Graphical Integrity
- Data-Ink Ratio
- Data Density

-Edward R Tufte-

Strive for Graphical Integrity

1

The representation of numbers, as physically measured on the surface of the graph itself, **should be directly proportional** to the numerical quantities represented



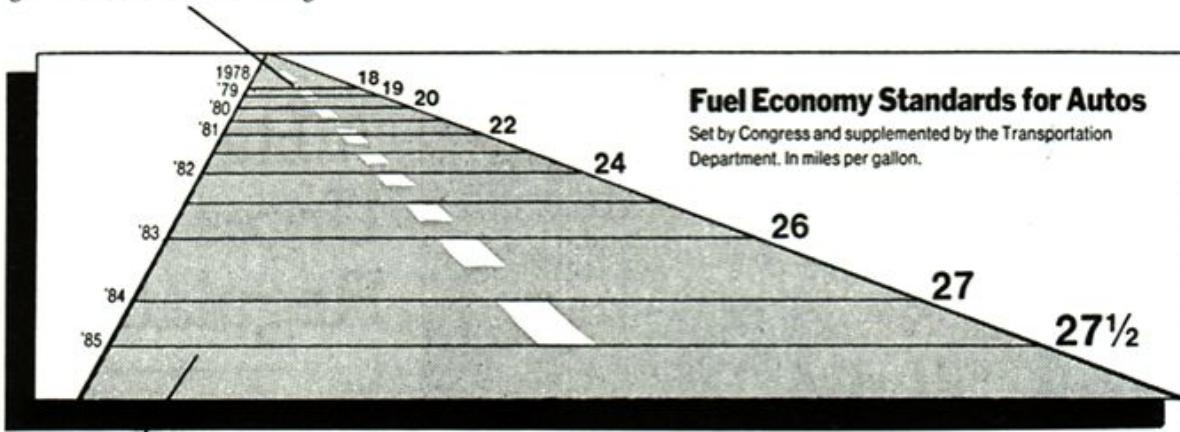
Lie Factor!

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

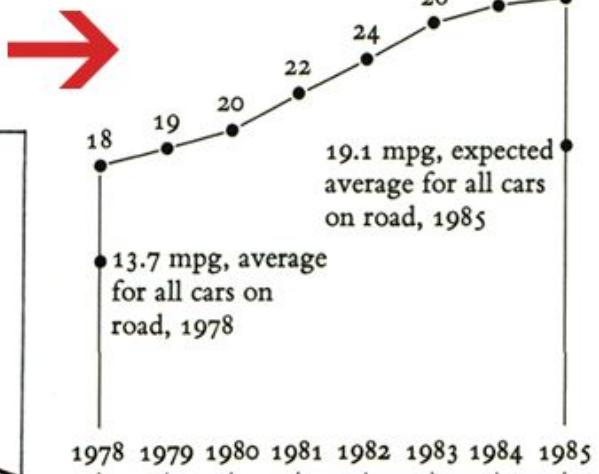
Lie Factor!

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.



Strive for Graphical Integrity

1

ADVERTISEMENTS

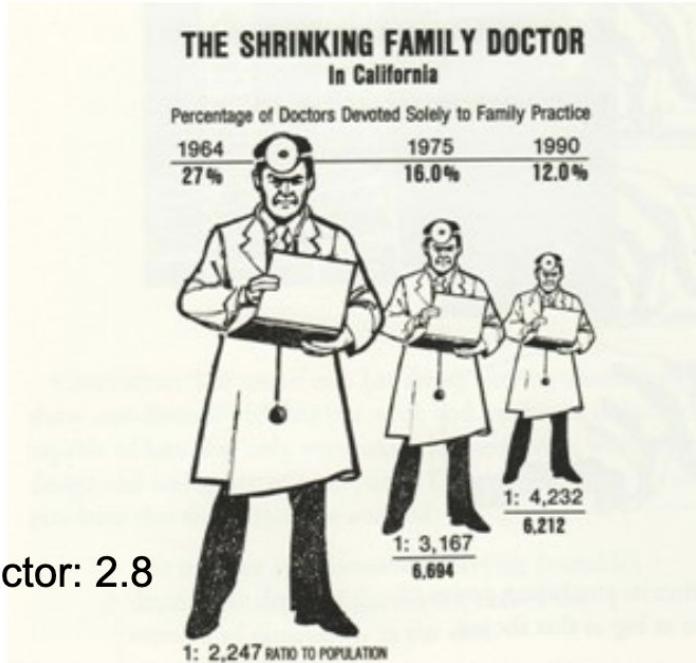


ACTUAL BURGER

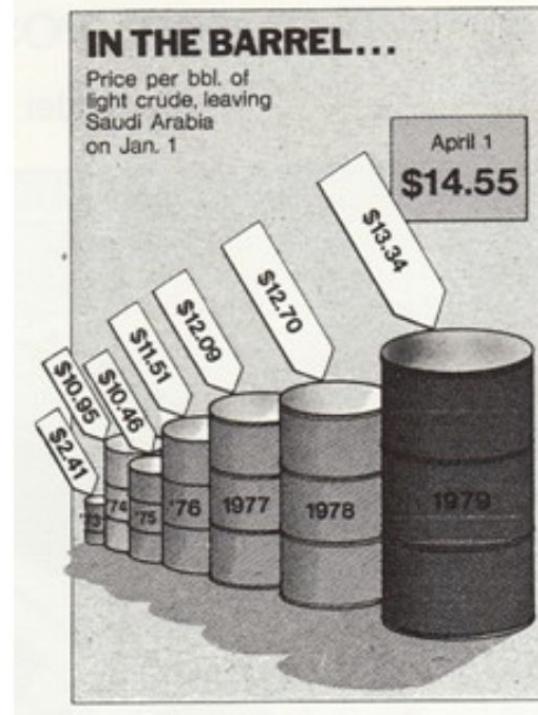
- ROTATED TO MOST ATTRACTIVE ANGLE
- SLIGHTLY FLUFFED UP, FOR PICTURE



Design Variation vs Data Variation



Lie factor: 2.8



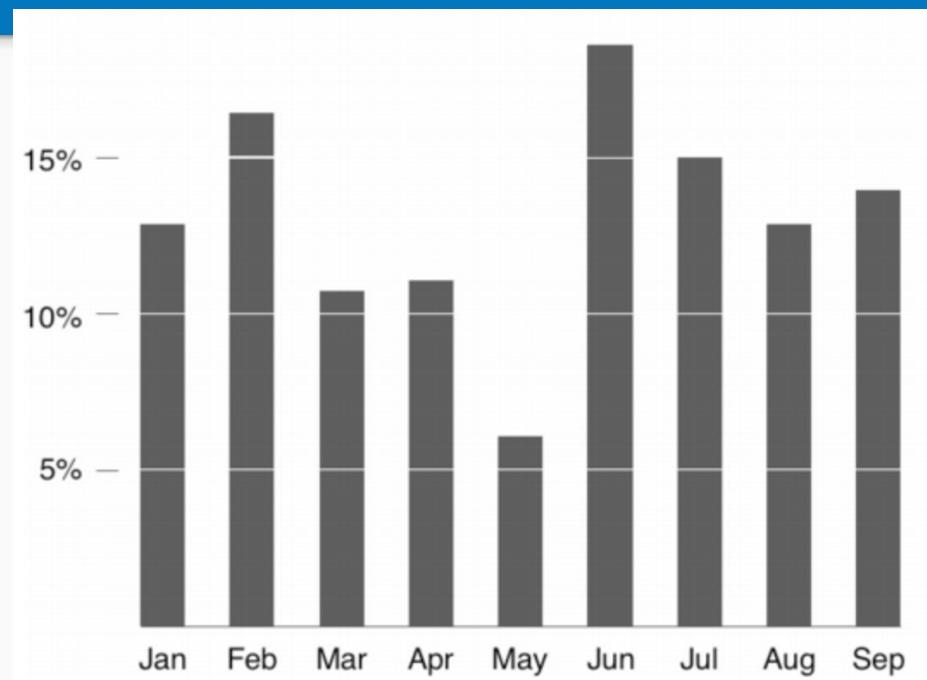
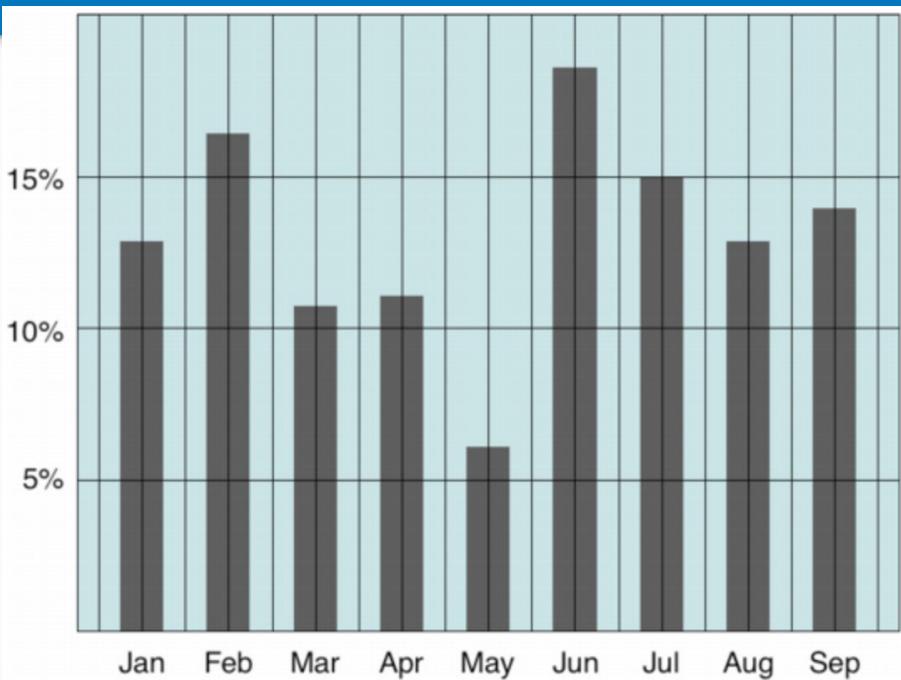
What is Data-ink Ratio?

$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

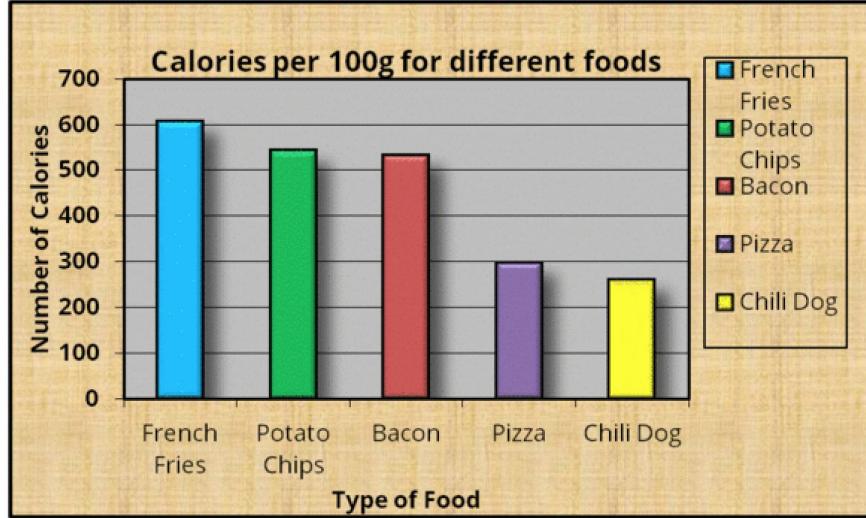
- = proportion of a graphic's ink devoted to the non-redundant display of data-information
- = 1.0 – proportion of a graphic that can be erased

What is Data-ink Ratio?

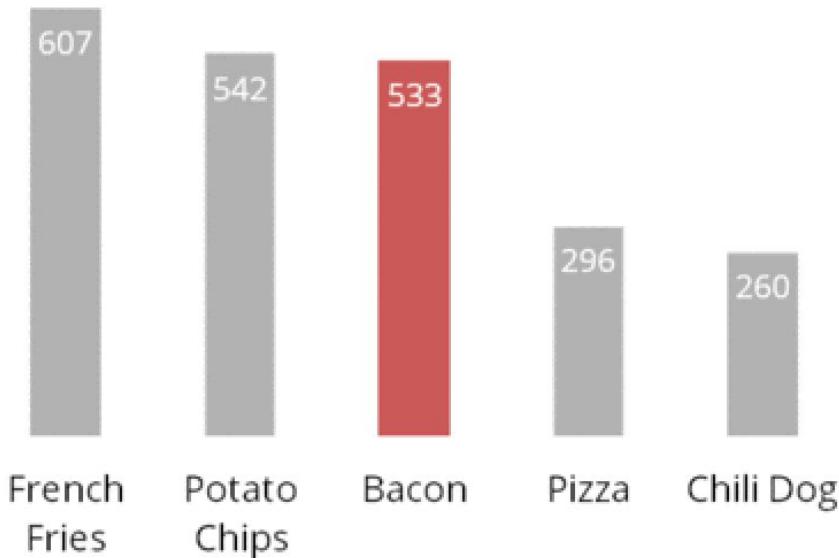
2



What is Data-ink Ratio?



Calories per 100g



What is Data-ink Ratio?

2

1. Above all else **show data**.
2. Maximize the data-ink ratio.
3. Erase non-data-ink.
4. Erase redundant data-ink.
5. Revise and edit

Data Density - Shrink Principle

Maximize data density and the size of the data matrix within reason.

Shrink Principle.

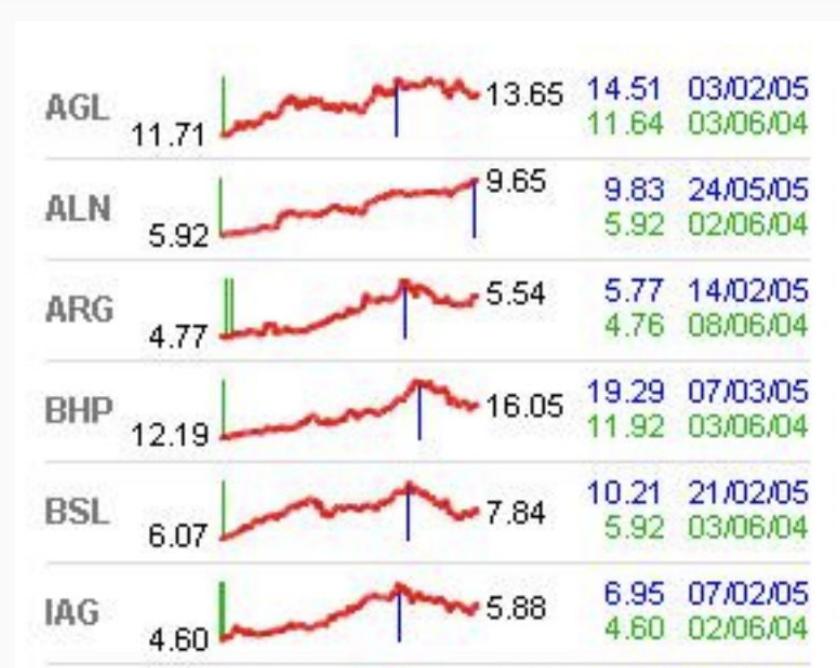
Most graphs can be shrunk way down without losing legibility or information.



Data Density - Sparklines

Small multiples are a great tool to visualize large quantities of data and with a high number of dimensions.

Sparklines are data-intense, design-simple, word-sized graphics.



What is Data-ink Ratio?

$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

- = proportion of a graphic's ink devoted to the non-redundant display of data-information
- = 1.0 – proportion of a graphic that can be erased

A close-up photograph of a person's hand holding a small, round, orange fruit, possibly a tangerine or orange segment. The hand is positioned in the lower-left corner, with the fingers slightly curled around the fruit. The background is dark and out of focus, with some large, green, tropical-style leaves visible behind the hand.

Choosing Right Visualizations to Relate Stories

*Choosing the right
visualization decides if
your audience
understands your
message!*

You don't want your audience
to be like this!



Communication mediums How to present data

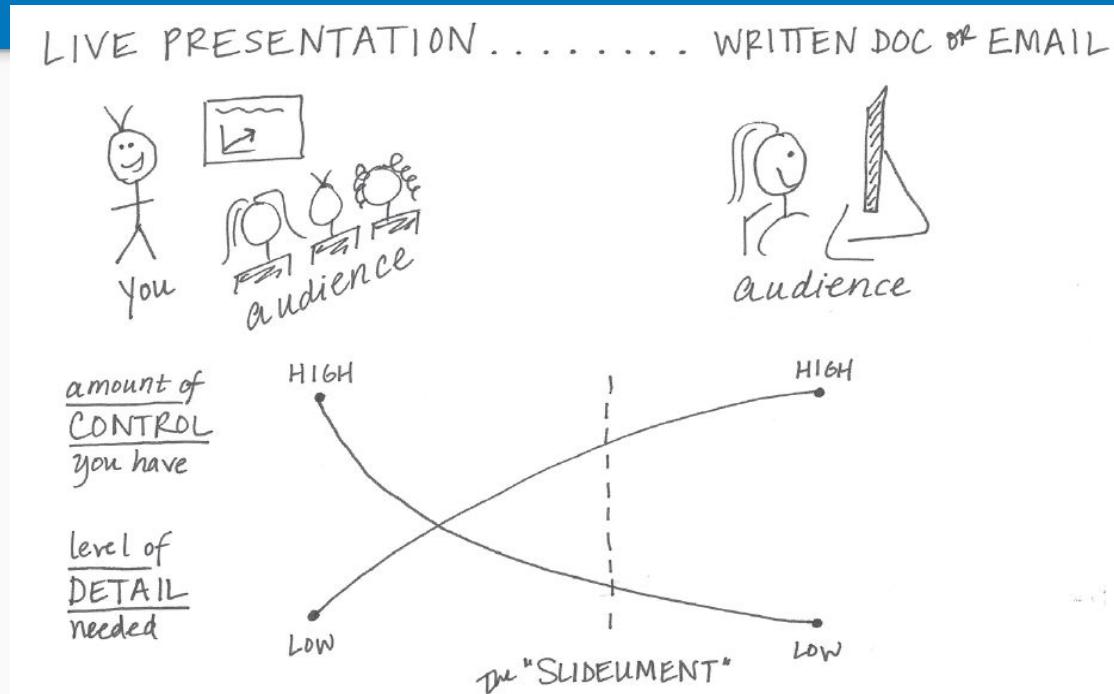
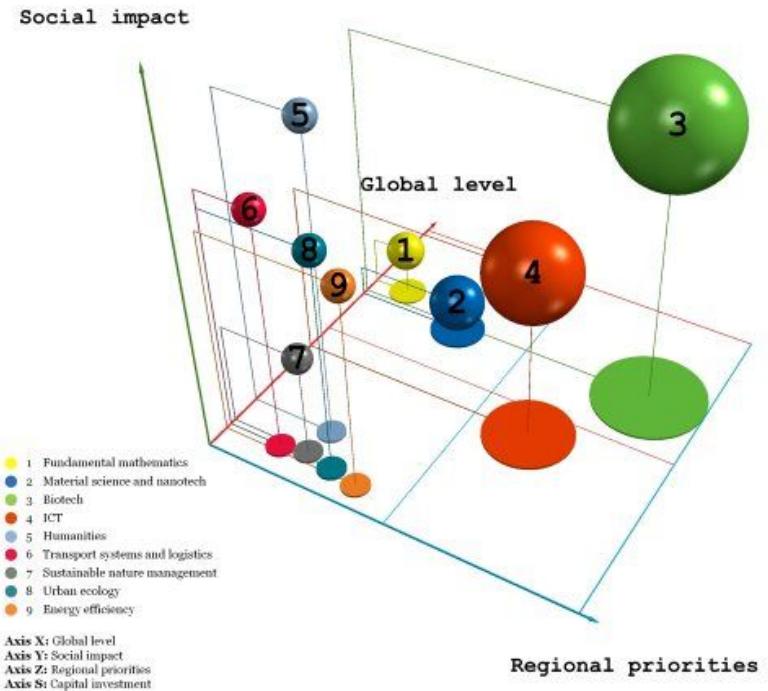
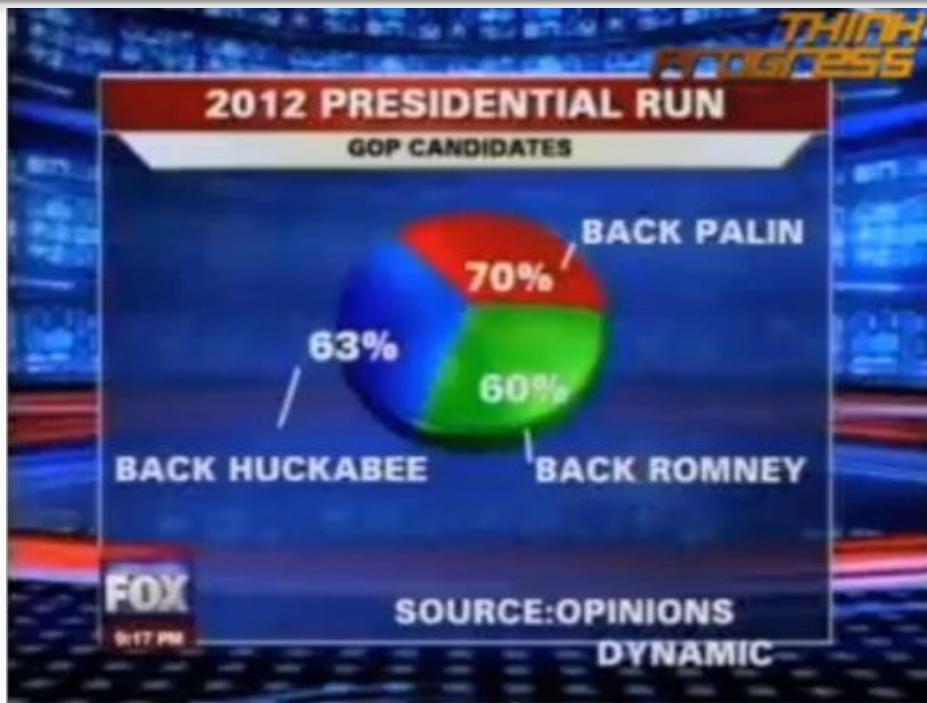
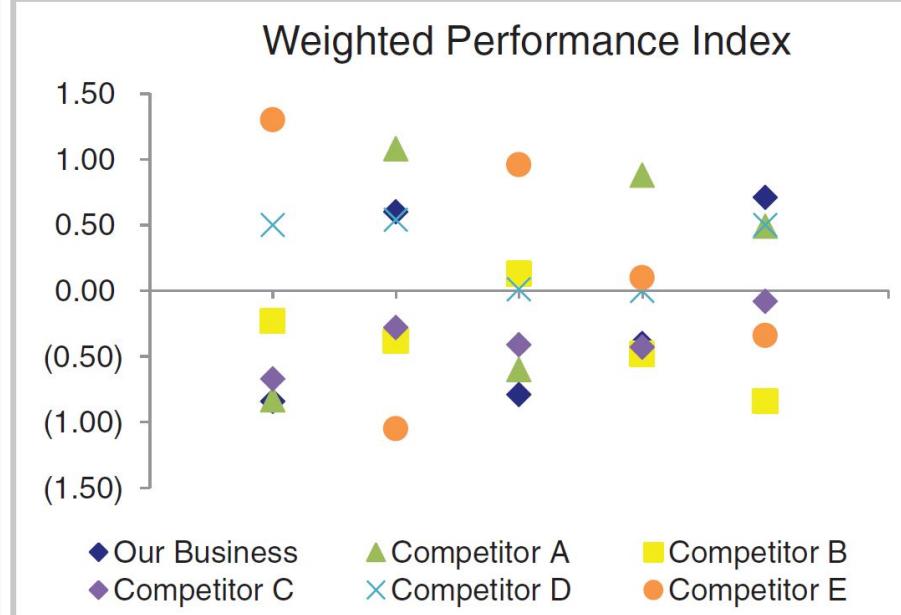
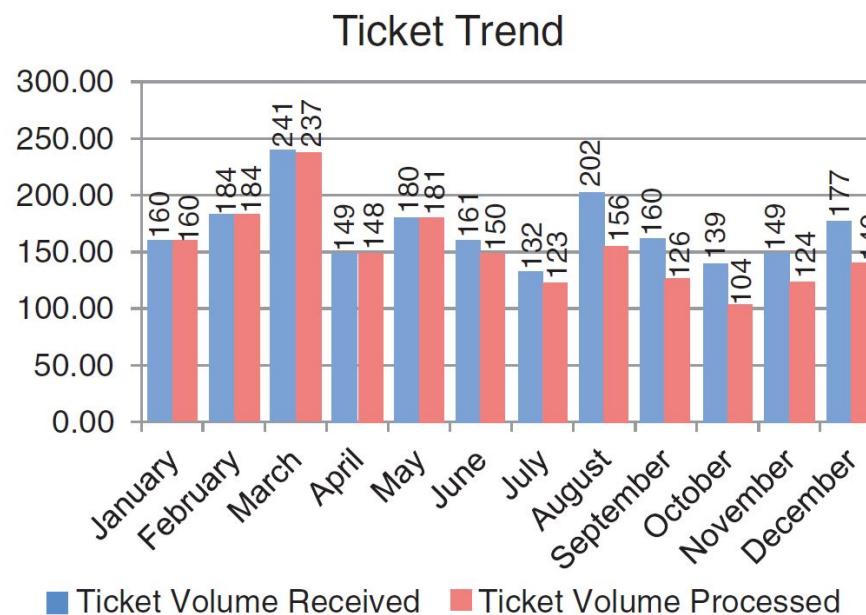


FIGURE 1.1 Communication mechanism continuum

How bad can it get?



How bad can it get?

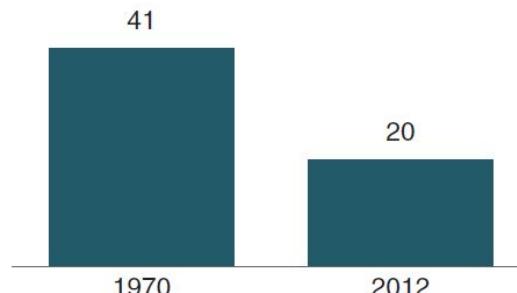


Simple Text can be enough!

Bad representation

Children with a "Traditional" Stay-at-Home Mother

% of children with a married stay-at-home mother with a working husband



Note: Based on children younger than 18. Their mothers are categorized based on employment status in 1970 and 2012.

Better representation

20%

of children had a
traditional stay-at-home mom
in 2012, compared to 41% in 1970

FIGURE 2.3 Stay-at-home moms simple text makeover

Stick to a theme (Aesthetics)

Bad representation

Distribution by customer segment

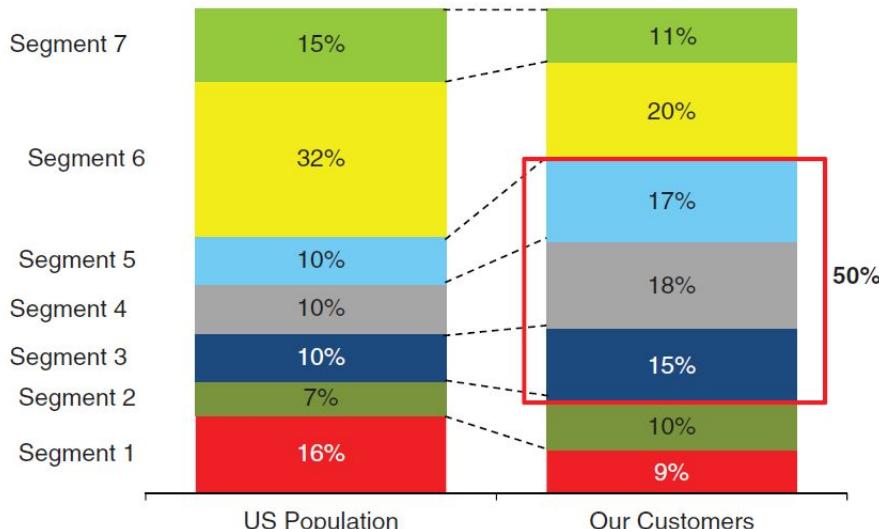


FIGURE 5.12 Unaesthetic design

Better representation

Distribution by customer segment

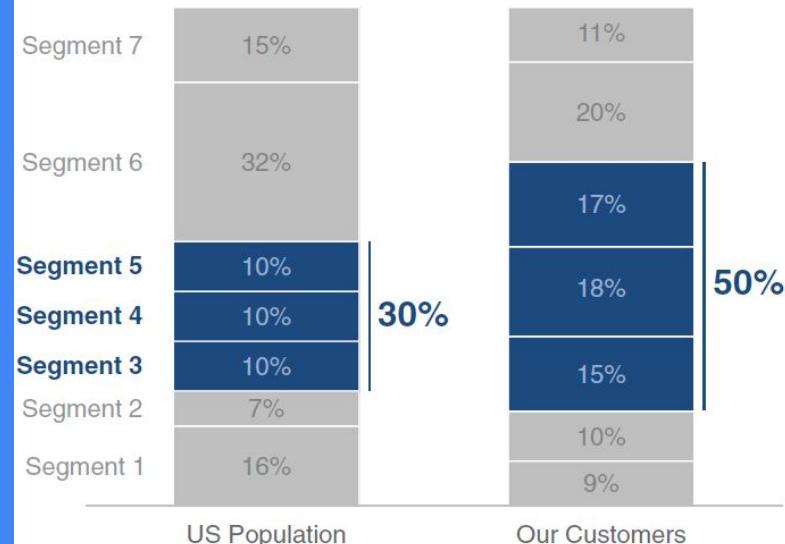


FIGURE 5.13 Aesthetic design

Use words sparingly and wisely!

Peak Break-up Times
According to Facebook status updates

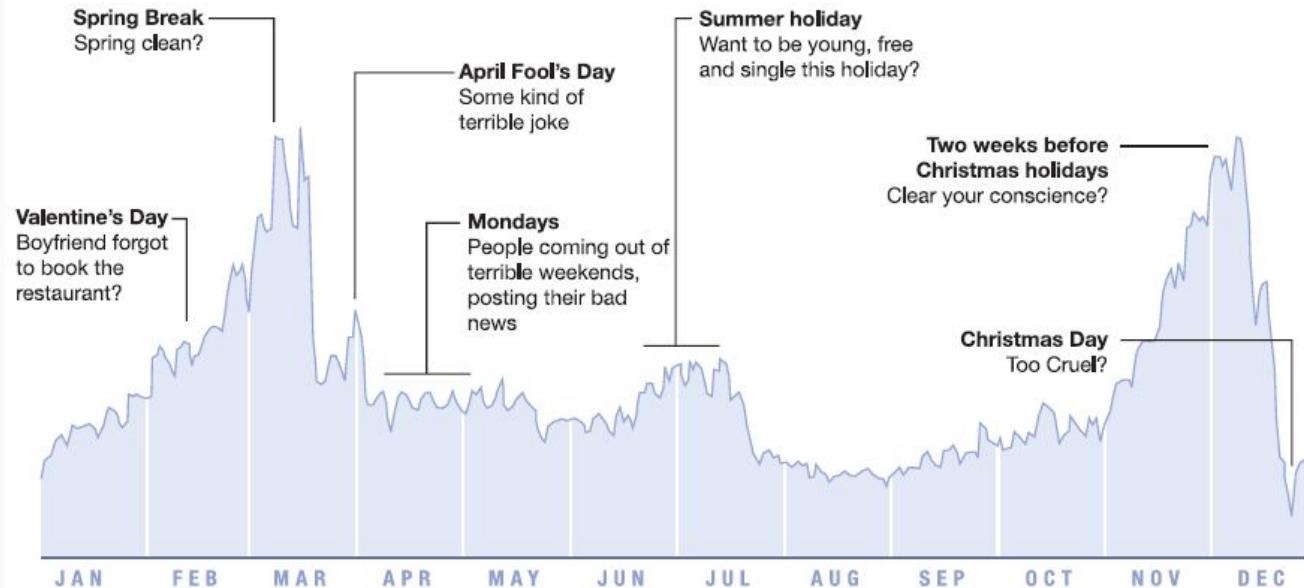


FIGURE 5.7 Words used wisely

Minimalism

Less is generally more!

Heavy borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Light borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Minimal borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Popping out your data

Use heatmaps to highlight important data

Table

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

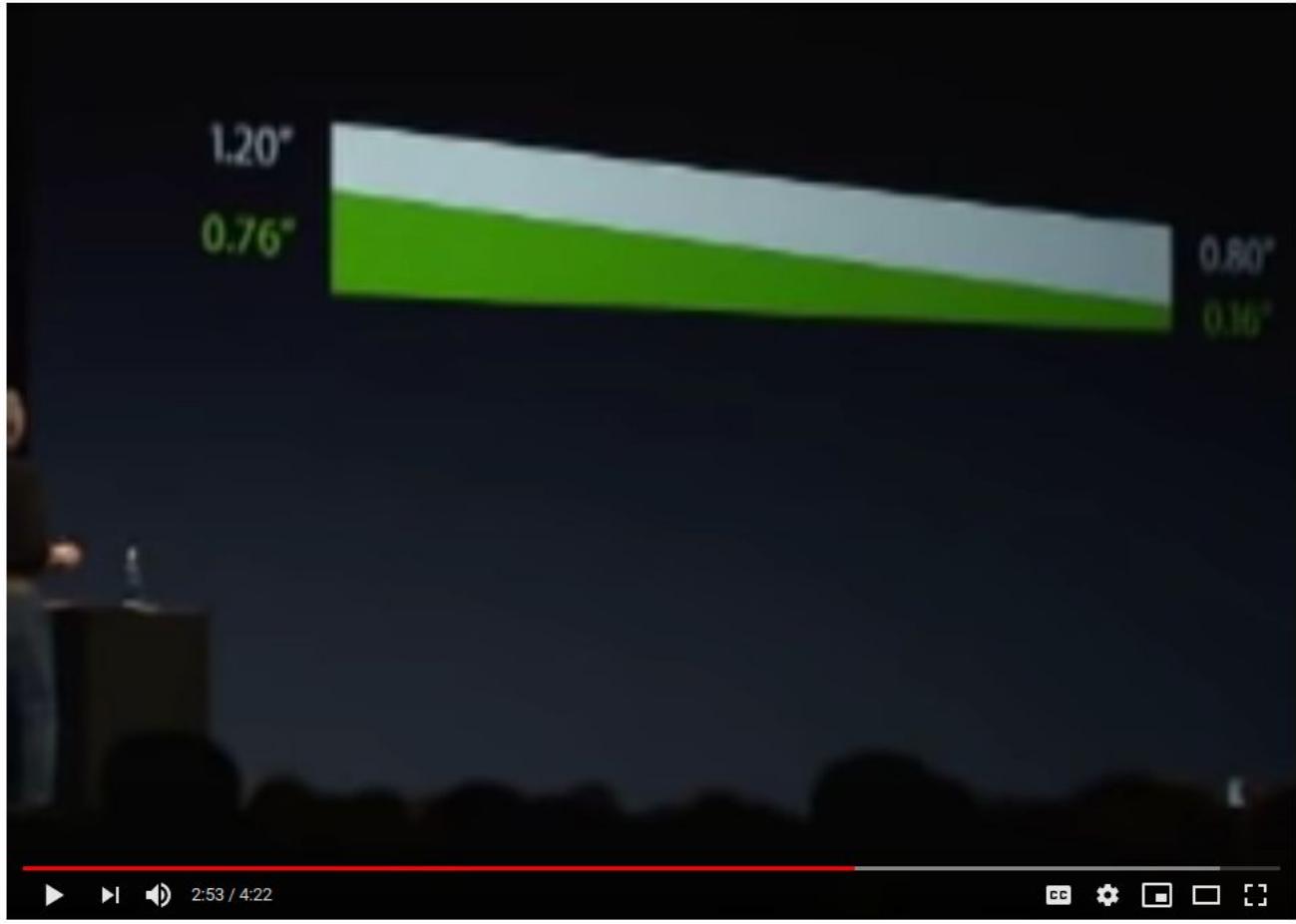
LOW-HIGH

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Choosing **color** is
important when
presenting data

**Green, Red and
Yellow** are the colors
mostly visible at a
distance.





MacBook Air Introduction by Steve Jobs

Use colors to group data

7 of the top 10 design concerns have 10 or more concerns per 1,000.

Discussion: is this an acceptable default rate?

Top 10 design concerns

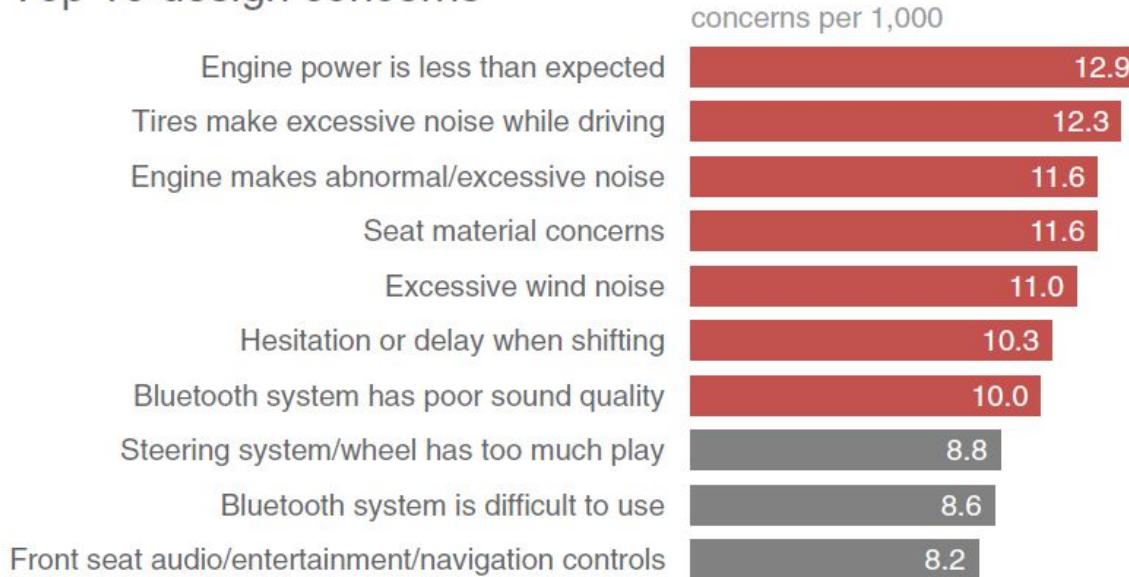


FIGURE 4.8 Leverage color to draw attention

Create a hierarchy of information through color

Of the top design concerns, three are noise-related.

Top 10 design concerns

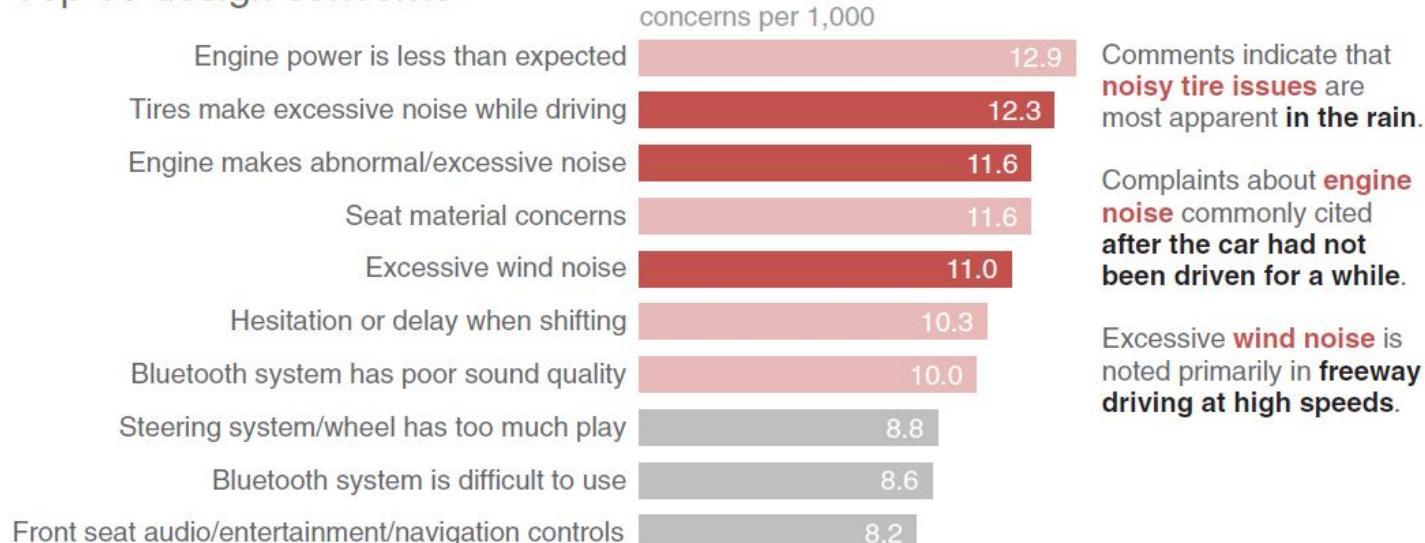


FIGURE 4.9 Create a visual hierarchy of information

Again less colors and consistent theme wins!

Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

Top 5 drugs: country-level sales rank

RANK	1	2	3	4	5+
COUNTRY DRUG	A	B	C	D	E
Australia	1	2	3	6	7
Brazil	1	3	4	5	6
Canada	2	3	6	12	8
China	1	2	8	4	7
France	3	2	4	8	10
Germany	3	1	6	5	4
India	4	1	8	10	5
Italy	2	4	10	9	8
Mexico	1	5	4	6	3
Russia	4	3	7	9	12
Spain	2	3	4	5	11
Turkey	7	2	3	4	8
United Kingdom	1	2	3	6	7
United States	1	2	4	3	5

FIGURE 4.15 Use color sparingly

Information Placement

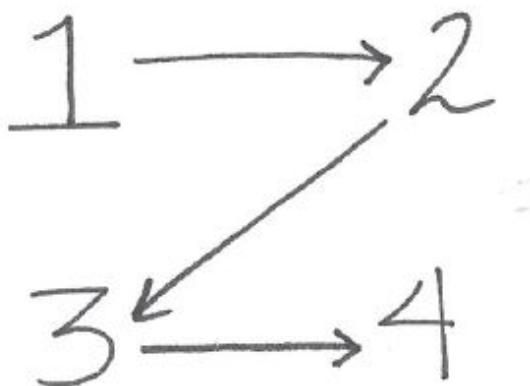
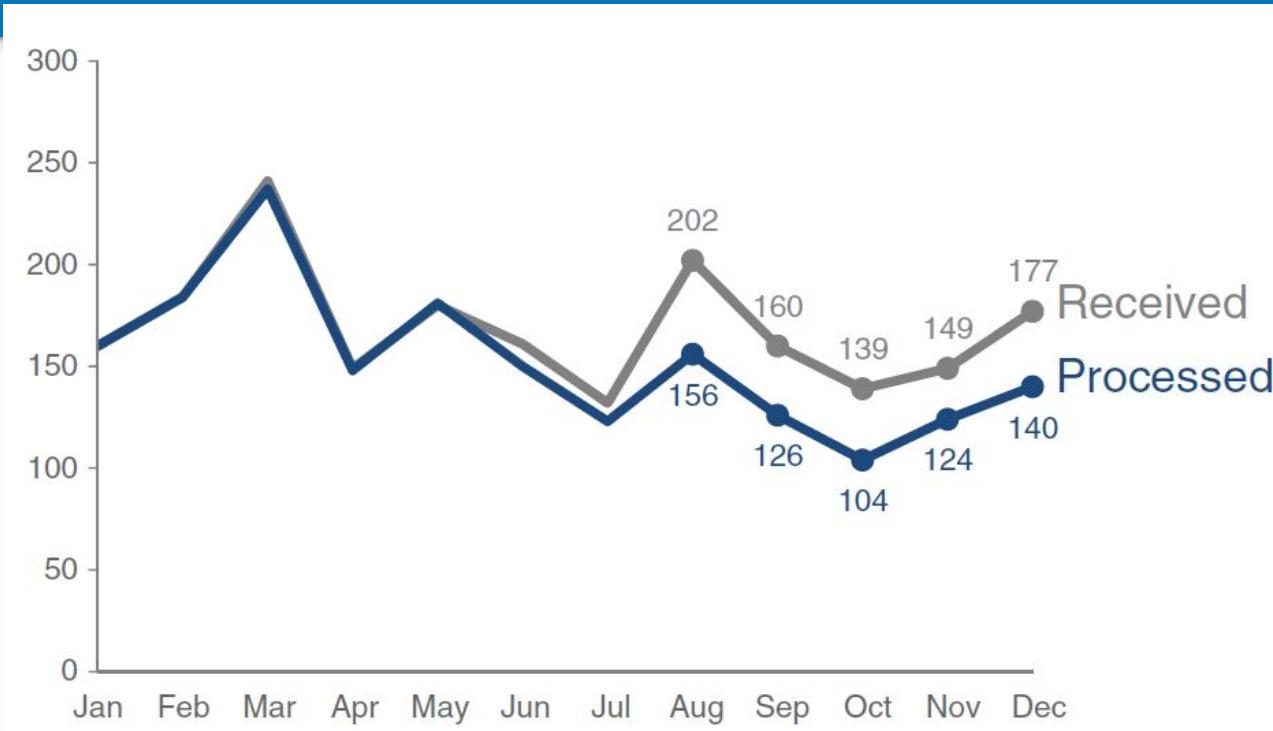


FIGURE 4.17 The zigzag “z” of taking in information on a screen or page

Eliminate Clutter

Cognitive overload

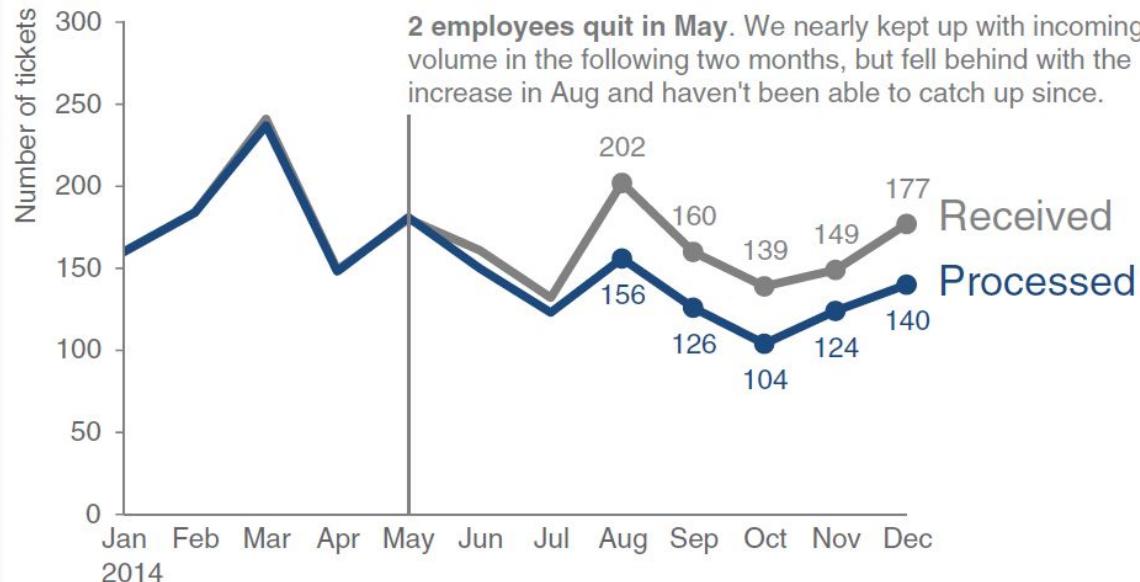


Action Titles and annotations

Please approve the hire of 2 FTEs

to backfill those who quit in the past year

Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

A dark, moody aerial photograph of ocean waves crashing, serving as the background for the title.

Motion Graphics for Appealing Stories

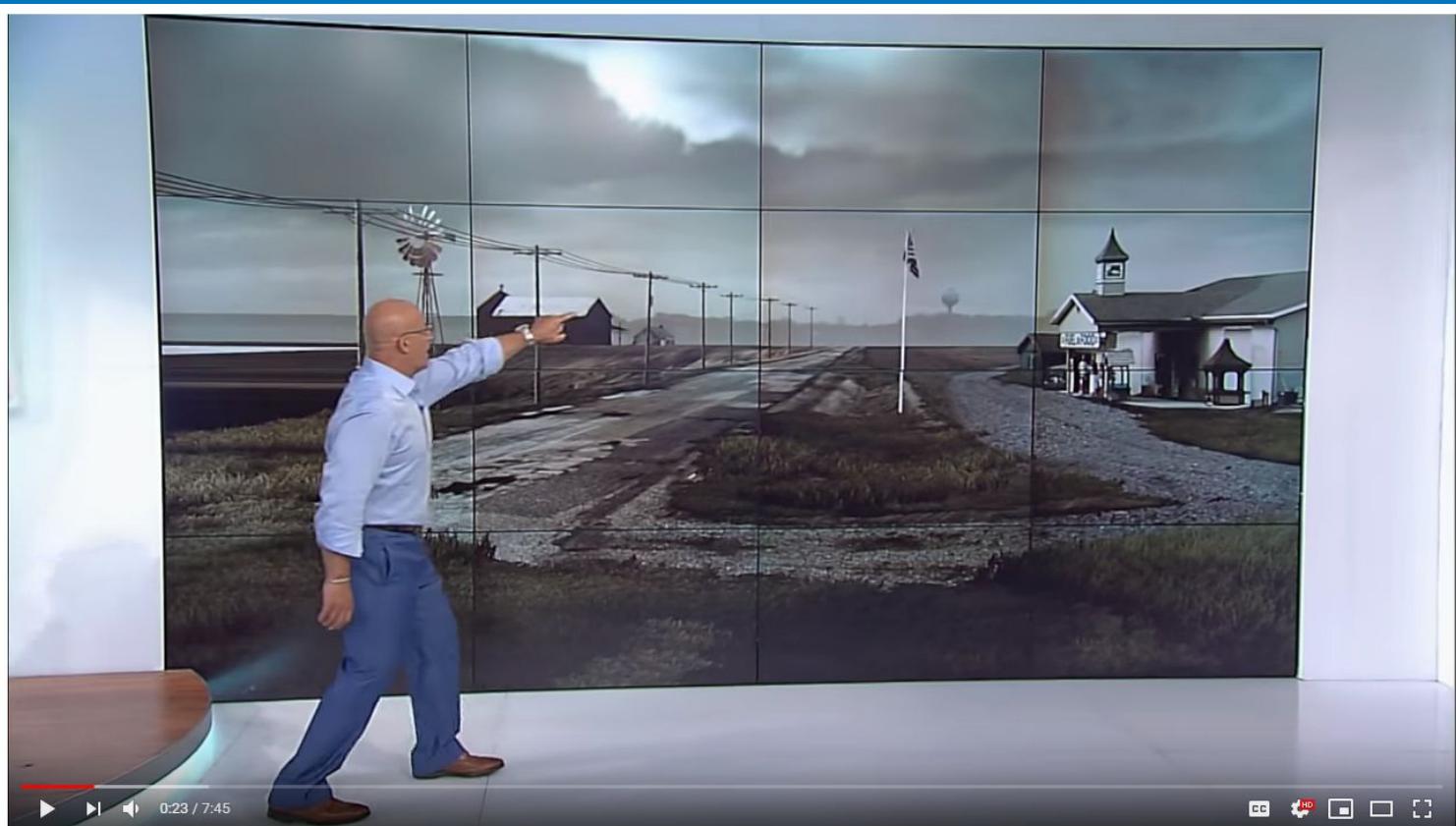
Motion graphics are a great way of conveying data.

New technologies like AR/VR can help in immersifying user experience with data.

Suspense and Awareness



Suspense and Awareness



Suspense and Awareness



AMHQ



0:27 / 2:41



Summary

- 1 Understand the Context.
- 2 Choose an appropriate visual cue.
- 3 Eliminate clutter.
- 4 Draw attention to where you want.
- 5 Think like a designer.
- 6 Tell a story!

A photograph of a person's hands resting on a light-colored wooden surface. The person is wearing a white long-sleeved shirt. On their left wrist is a silver-toned analog watch with a dark brown leather strap. In their right hand, they hold a white ceramic cup filled with dark coffee. In their left hand, they hold a pair of white-framed sunglasses. A green dashed rectangular frame surrounds the central text area.

Intermission

Group Activity

PetSocial, is a social media for pets. People around the world can host events for pet gatherings and can maintain profiles of their pets.

You as a data engineer need analyze the following.

1. Types of data we can have in the social platform.
2. Personalizing and visualizing this data. (what stories can we relate with that data?).



SLIIT has launched a program for school kids on learning science.

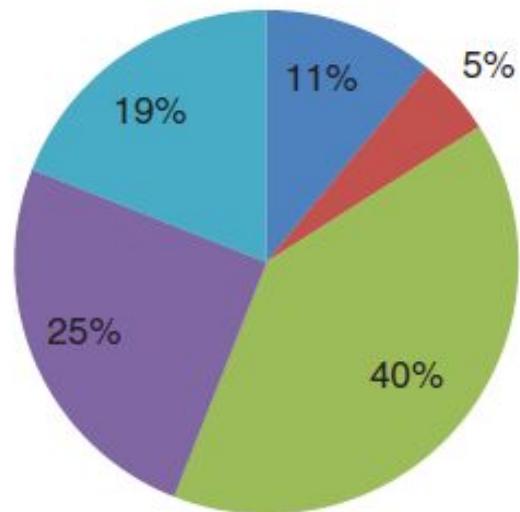
A survey was conducted before and after the program to see how students feel about doing science.

You as the visualization expert need to visualize these results!

Case Study - The Bad Visualization

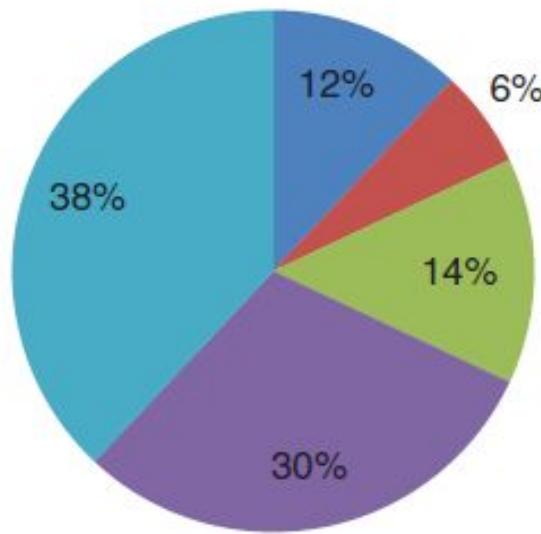
PRE: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited



POST: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited



Case Study - Simple Text based

Pilot program was a success

After the pilot program,

68%

of kids expressed interest towards science,
compared to 44% going into the program.

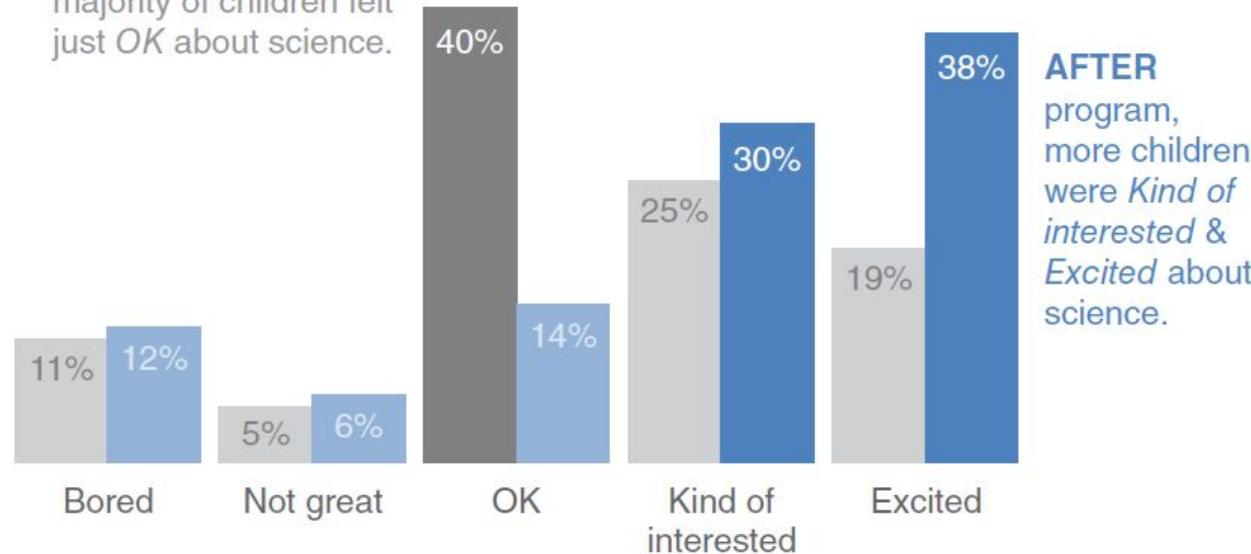
Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

Case Study - Simple bar graph

Pilot program was a success

How do you feel about science?

BEFORE program, the majority of children felt just *OK* about science.



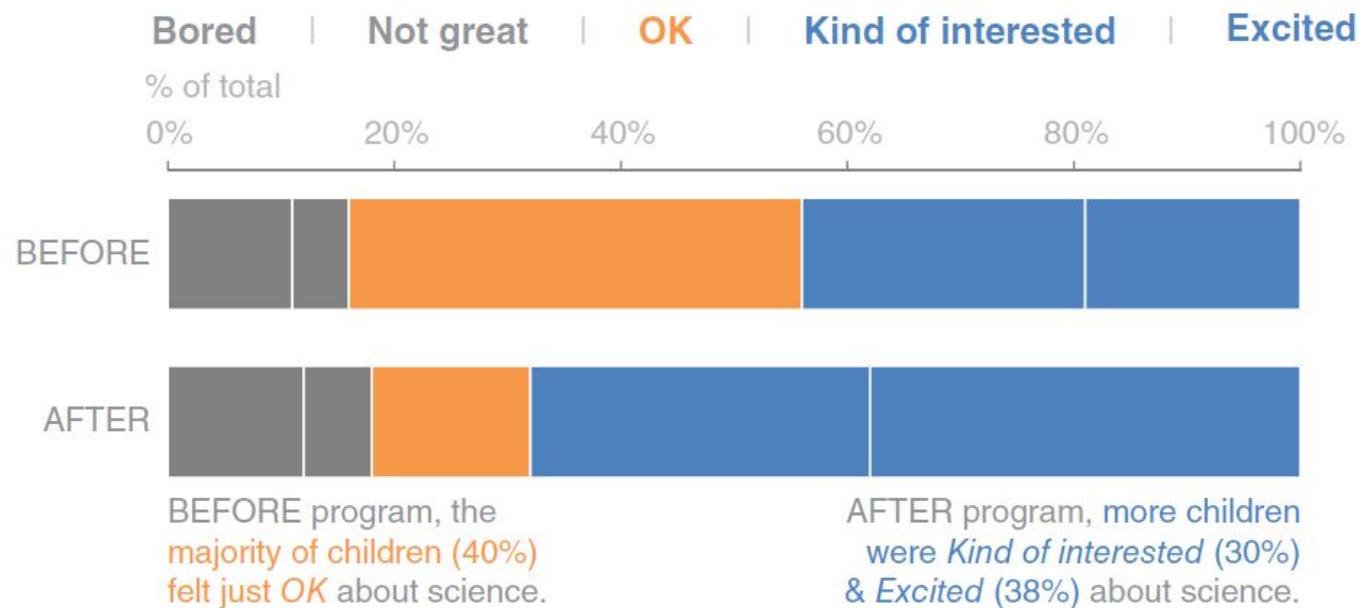
AFTER program,
more children were *Kind of interested & Excited* about science.

Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

Case Study - Horizontal Bar graph

Pilot program was a success

How do you feel about science?

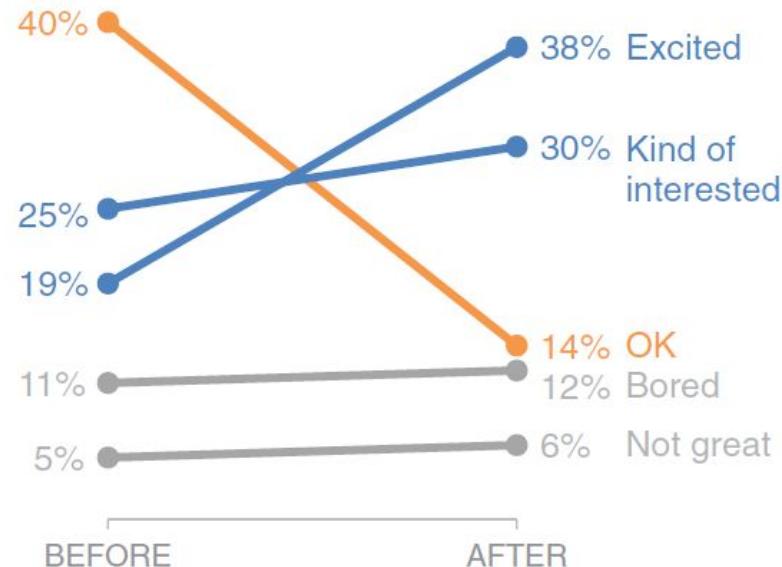


Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

Case Study - Slopegraph

Pilot program was a success

How do you feel about science?



BEFORE program, the majority of children felt just *OK* about science.

AFTER program, more children were *Kind of interested* & *Excited* about science.

Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

References

Storytelling with data, a data visualization guide for business professionals.

Cole Nussbaumer Knaflic

Photo Credits

Unsplash for awesome free photos!

<https://unsplash.com>

Thank you