Data Science in the Business World

BUSI 488 & COMP 488

Daniel M. Ringel

UNC Kenan-Flagler Business School
Spring 2023

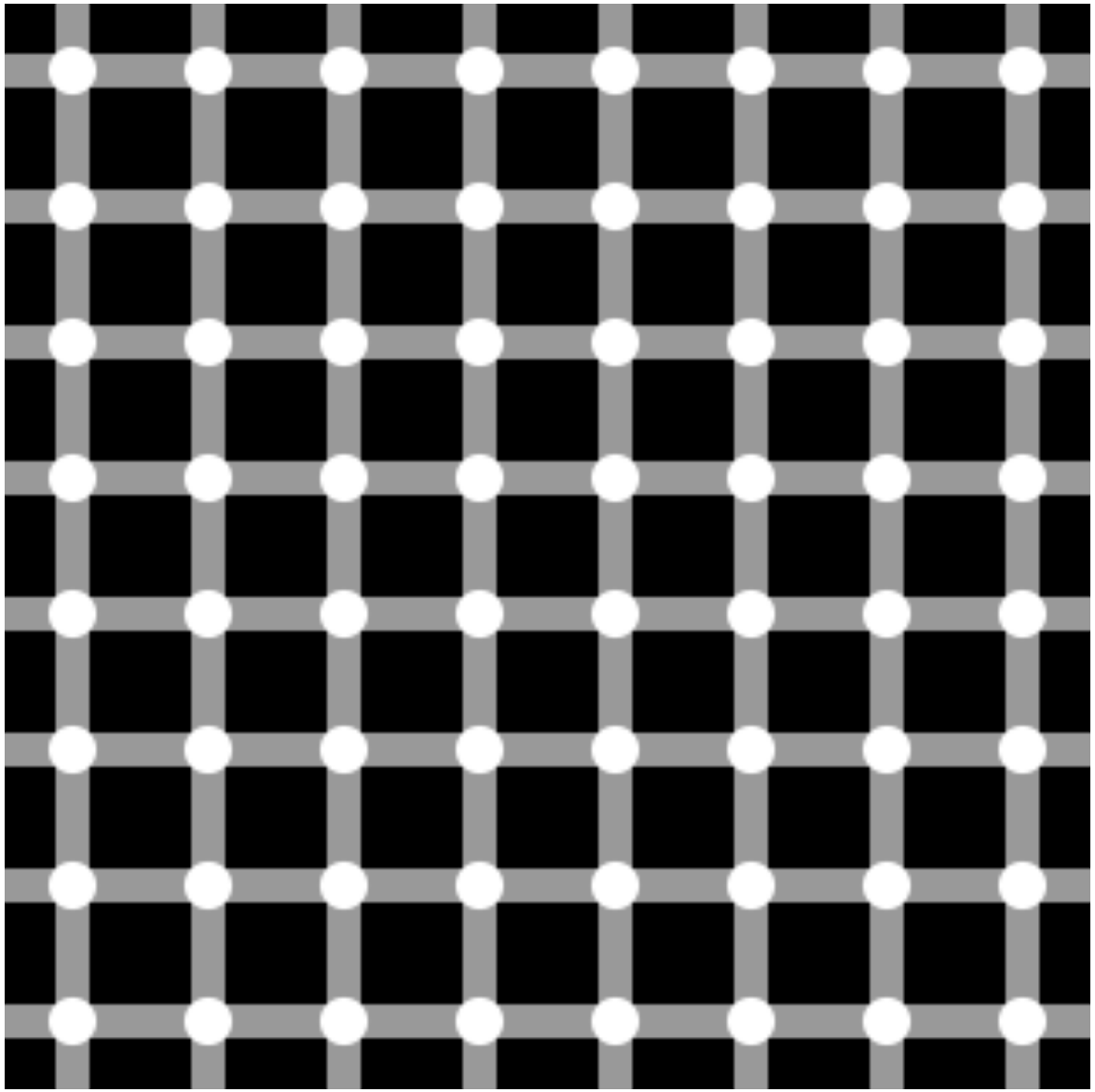
January 19th, 2023

Class 04: Storytelling with Data that Managers can understand

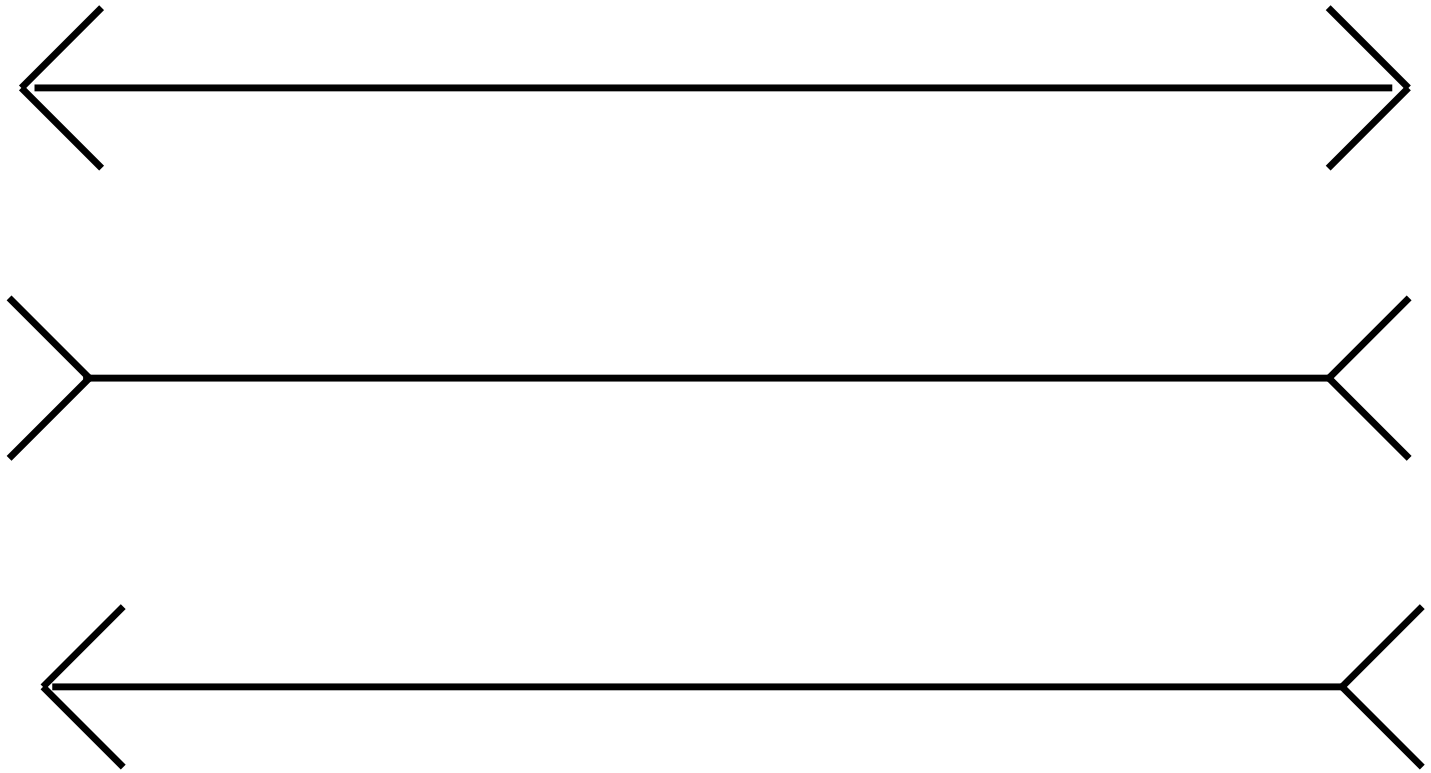
Sections 001 and 002

UNC KENAN-FLAGLER BUSINESS SCHOOL

Can you find the black circles?



Which line is the longest?



Today's Agenda

1. Telling Stories with Data
2. Data Visualization
3. Gestalt Principles of Visual Perception

Prep-Check:

✓ Read *Narrative visualization: Telling stories with data*
✓ Started DataCamp HW2

Telling Stories with Data

- Storytelling is an art and a skill

- We live in a world of
    - increasing data
    - data-driven decision making

- Effective data visualization can mean the difference between success and failure
    - Communicate the findings of your study
    - Raise money for your nonprofit
    - Present to your board
    - Get your point across to your audience

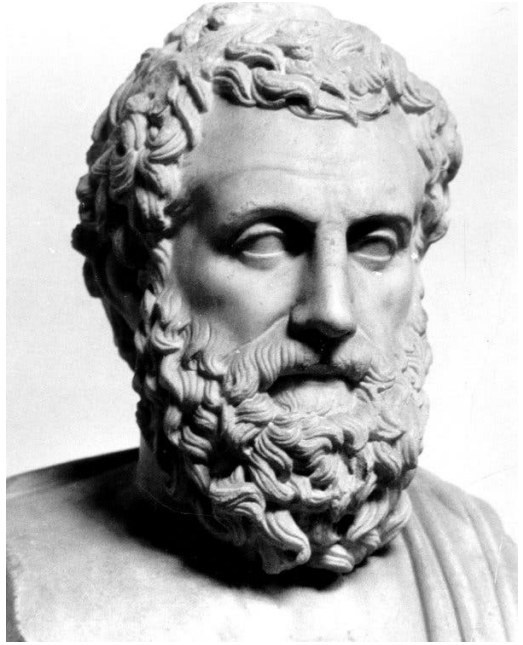
What makes a great Story?

The Magic of Stories

- great play
- captivating movie
- fantastic book
- speeches
- … and presentations

A Good story

- grabs your attention
- takes you on a journey
- evokes an emotional response



Narrative structure

- First described by Aristotle and Plato

Aristotle introduced a basic but profound idea:

A story has a clear beginning, middle, and end.

Three-act structure for plays:

- the setup
- the conflict
- the resolution

A Great Story …

In the middle of it, you find yourself not wanting to turn away or put it down.

After finishing it—a day, a week, or even a month later—you could easily describe it to a friend.

Cole Nussbaumer Knaflic

Subjective Expectation meets Cruel Reality

Critical Components of a Story

- Struggle
- Conflict
- Suspense

Key Questions that Reveal Stories

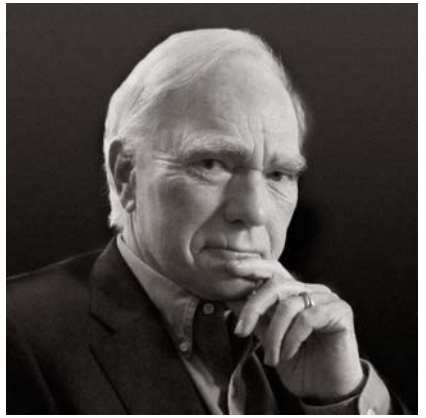
- What does my protagonist want in order to restore balance in his or her life?
- What is the core need?
- What is keeping my protagonist from achieving his or her desire?
- How would my protagonist decide to act in order to achieve his or her desire in the
- face of those antagonistic forces?

After creating the story, lean back and consider:

- *Do I believe this story?*
- *Is it neither an exaggeration nor a soft-soaping of the struggle?*
- *Is this an honest telling, though heaven may fall?*

Robert McKee
Award-winning writer and director
Screenwriting lecturer

Harvard Business Review
How storytelling can be leveraged in a business setting.
https://hbr.org/2003/06/storytelling-that-moves-people



Two Ways to persuade People:

1. Conventional Rhetoric

- typically PowerPoint slides with bulleted facts and statistics
- intellectual process

Problematic, because

While you're trying to persuade your audience …

… they are arguing with you in their heads

Intellectual basis is not good enough:

→ "People are not inspired to act by reason alone" (Fryer, 2003).

2. A Story

- Unites an idea with an emotion
- Arouses the audience's attention and energy
- Requires creativity
- Harder than conventional rhetoric
- Story allows you to engage your audience on an entirely new level

The Big Idea

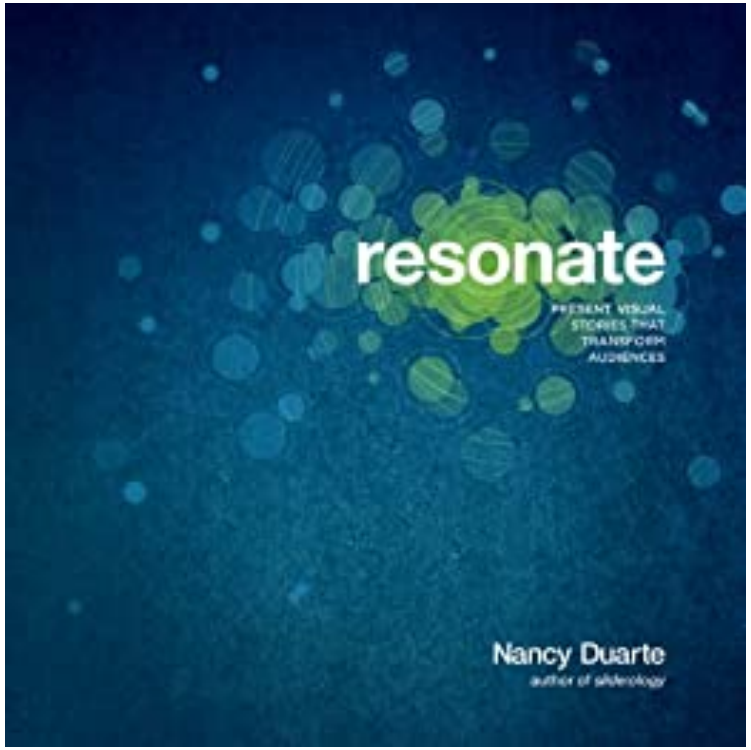
Boiling Your Story down to a Single Sentence

Three components of Your Big Idea:

1. It must articulate your unique point of view
2. It must convey what's at stake
3. It must be a complete sentence

Concept discussed by Nancy Duarte in her book Resonate (2010)

Based on analysis of sales in the market over time, to be competitive, we recommend launching our new product at a retail price in the range $5.99–$7.99.



www.amazon.com/dp/0470632011

The Elevator Pitch

If you only had a limited amount of time or a single sentence to tell your audience what they need to know, what would you say?

Too Much Data – Too Little Focus

What you did
- Your team mined gigabytes of data
- You tested 100 different hypotheses
- You looked at the data in 100 different ways
- You found three gems

The Story you want to tell
A specific story – probably about the three gems you found

The Story that is often told
Showcases all the great analysis done: Dozens of charts and tables

- Can be tempting to want to show your audience everything
- Evidence of all of the work you did
- Demonstrate robustness of the analysis

Turns an engaging story into a tiring story

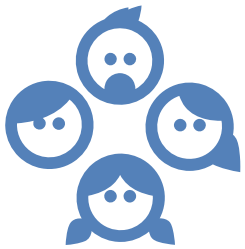
Need to turn data into information
- that is relevant to your audience
- can easily by digested

→ Focus on your three gems.

Focus on 3 Central Questions

Who

is your
Audience ?
be specific !



What is their relationship to you?
What motivates them?
What keeps them up at night?

What

do you need
*them **to do** ?*
be explicit !

Don't assume they will connect the dots themselves!

How

will ***data help*** you
make your point ?
be discerning !

What data will act as evidence for the case you are making?

Based on Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Start with a Worksheet

WHO IS YOUR **AUDIENCE**?

1. List the primary groups or individual to whom you will be communicating

   Your text …

2. If you had to narrow that to *a single person*, who would that be?

   Your text …

3. What does your audience care about?

   Your text …

4. What action does your audience need to take?

   Your text …

WHAT IS **AT STAKE**?

What are the *benefits* if your audience acts in the way you recommend it to?

Your text …

What are the *risks* if they do not?

Your text …

FORMULATE YOUR **NEW IDEA**

You should:

1. articulate your point of view
2. convey what's at stake
3. be a complete and (single!) sentence

Adapted from Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Clarify Context

Who requests a Deliverable

- Client
- Stakeholder
- Boss

You might not have full context!

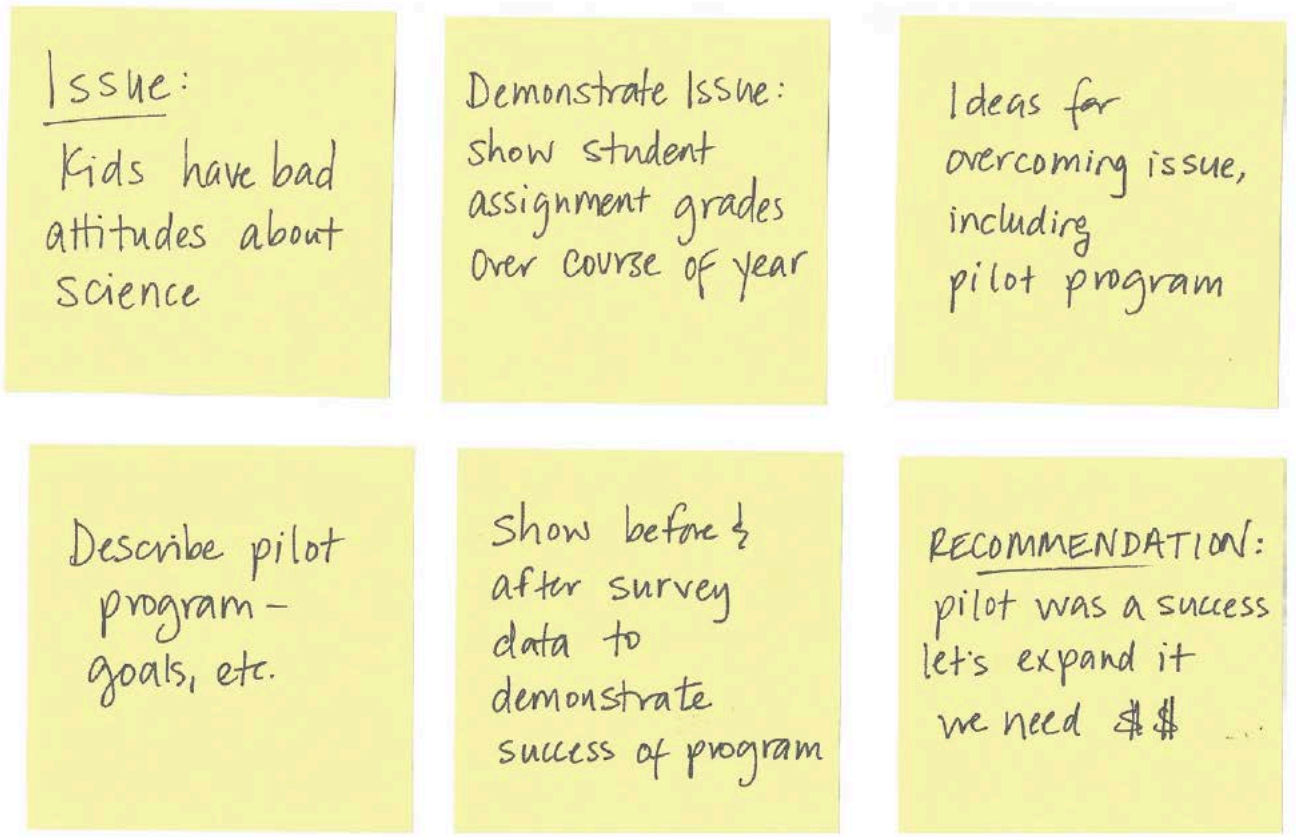
Questions to ask Yourself

- What background information is relevant or essential?
- Who is the audience or decision maker? What do you know about them?
- What biases does our audience have that might make them supportive of or resistant to your message?
- What data are available that would strengthen your case?
- Is your audience familiar with these data, or are they new?
- Where are the risks: What factors could weaken your case and do you need to proactively address them?
- What would a successful outcome look like?

If you can, go back to clarify questions that you cannot confidently answer

Storyboarding

- Single most important thing to do up front
- Ensures that your communication is on point
- Establishes a structure for your communication
- Visual outline of the content you plan to create



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

I would have written a shorter letter, but I did not have the time

French mathematician and philosopher Blaise Pascal (1623 - 1661)

Storytelling in the Business World

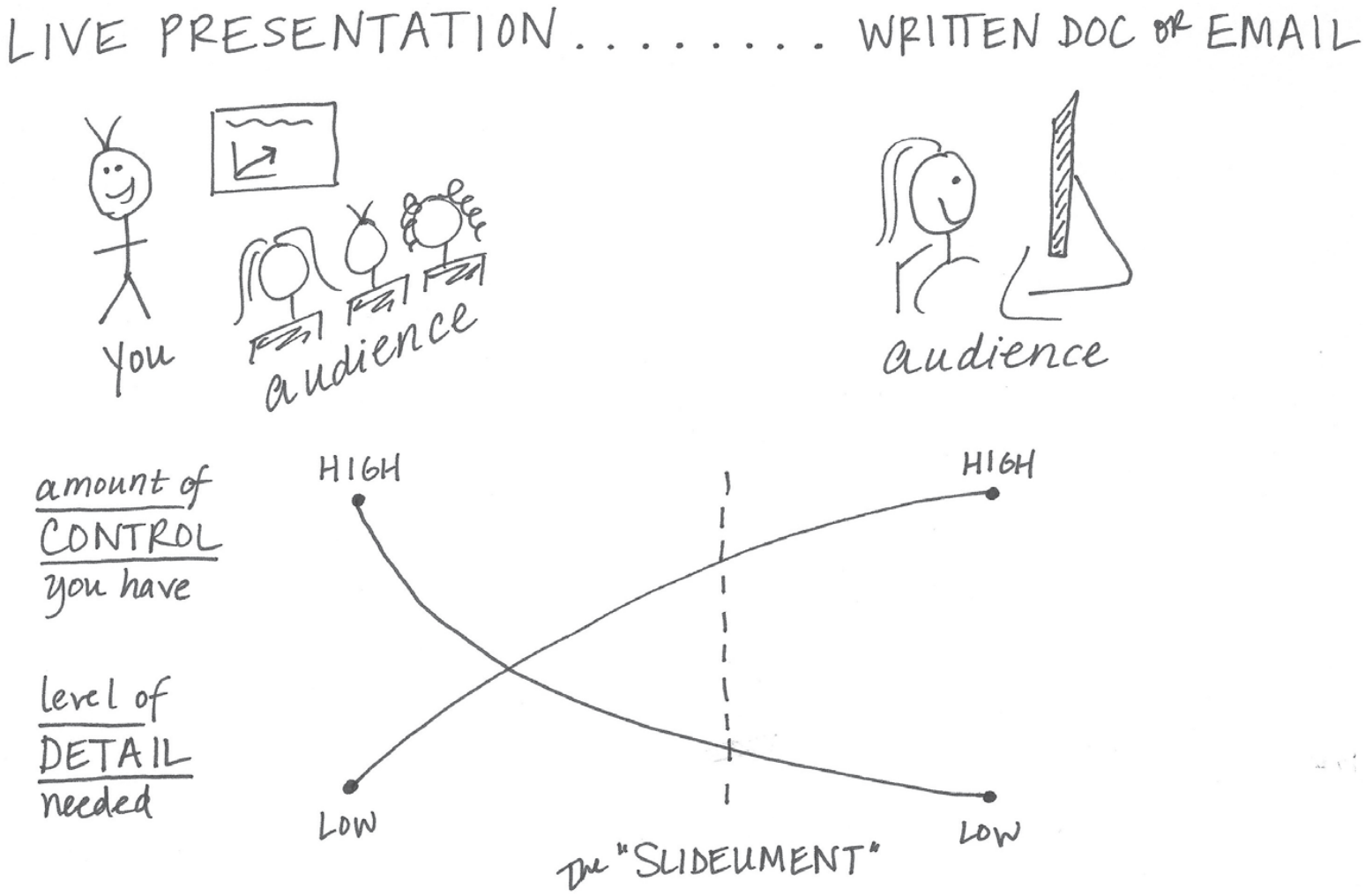
Basic Rules

- Keep it simple
- Laypeople terms
- Self-explanatory diagrams
- Cut the clutter in figures and diagrams
- Key messages only
- Less is often more!
- Cut the Buzzwords
- Limited (if any) animations
- Minimum font size is 12
- Avoid dark backgrounds

Answer Key Questions

- What is the problem?
- Why is it a problem (magnitude, risks)?
- Who is most affected?
- Why not solved?
- How will your approach fix it?
- What are the implications of your findings and/or results?
- Why is the world a better place with your analysis and/or model?
- What actions should management take (your recommendation)?

The Slideument



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

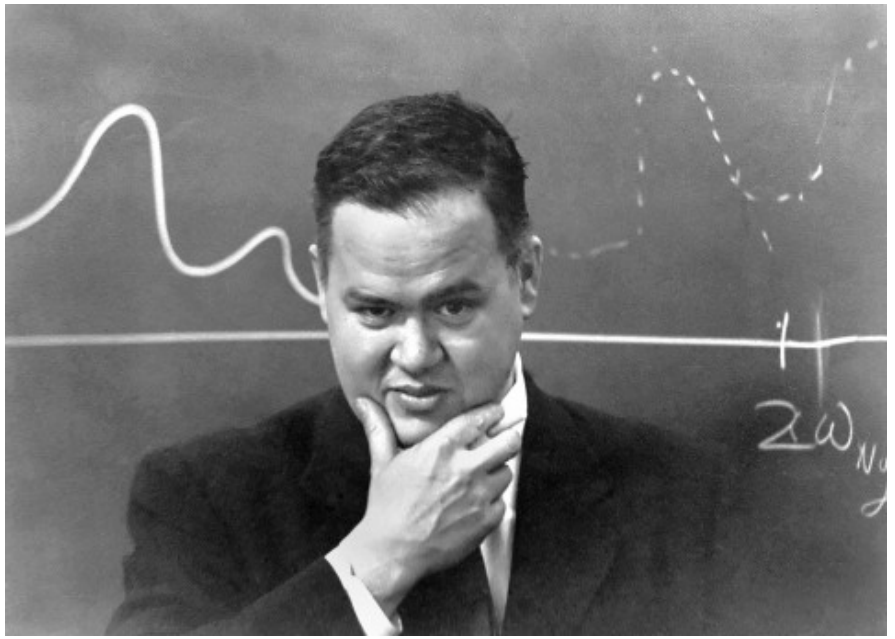
Data Visualization

The greatest value of a picture is when it forces us to notice what we never expected to see

John W. Tukey in *Exploratory Data Analysis*, 1977

John Wilder Tukey, a mathematician who first coined the term “exploratory data analysis,” was right when he suggested that the idea of visualization helps us see what we have not noticed before. That is especially true when you are trying to identify relationships and find meaning in huge amounts of collected data.

John W. Tukey was credited with having coined the terms software and bit.

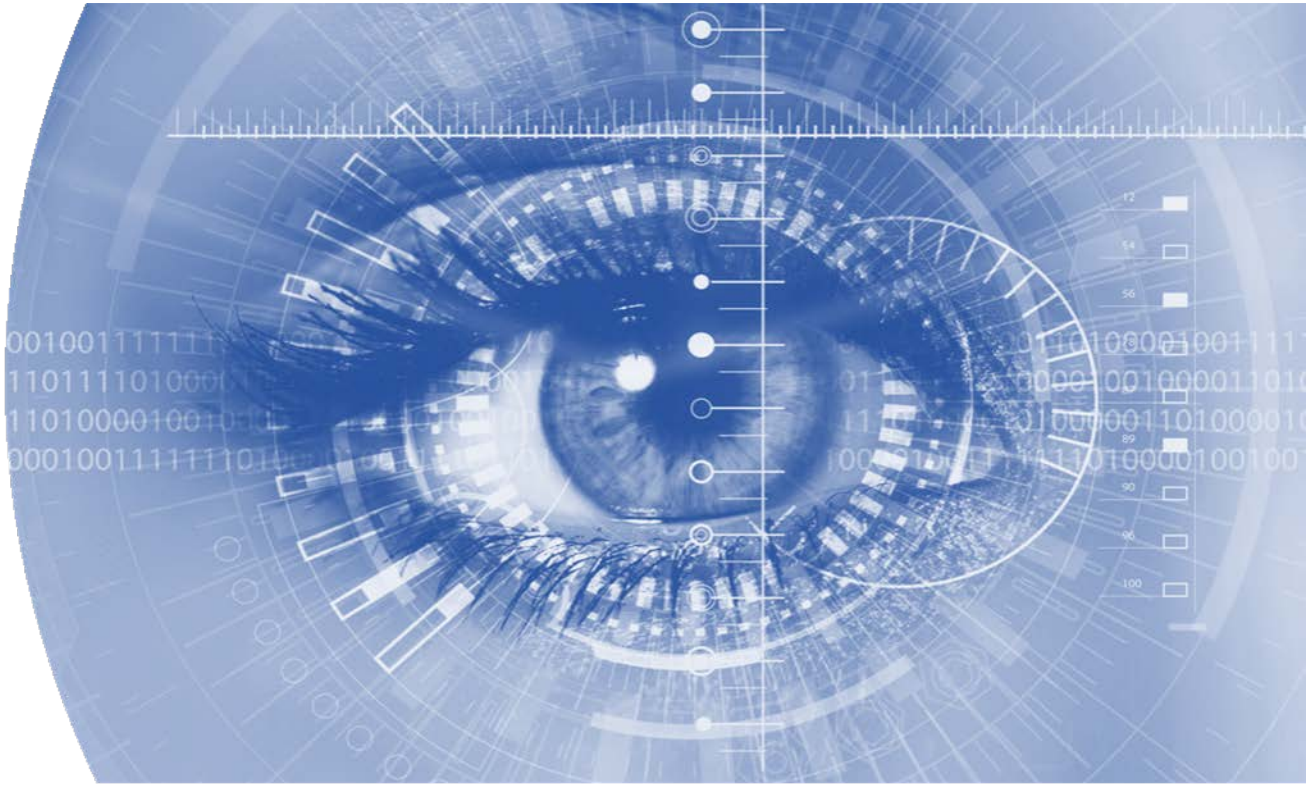


Visual Cognition

More neurons in our brain are associated with vision than with the other 4 senses—taken together

The retina transports information at approximately 10 million bits per second

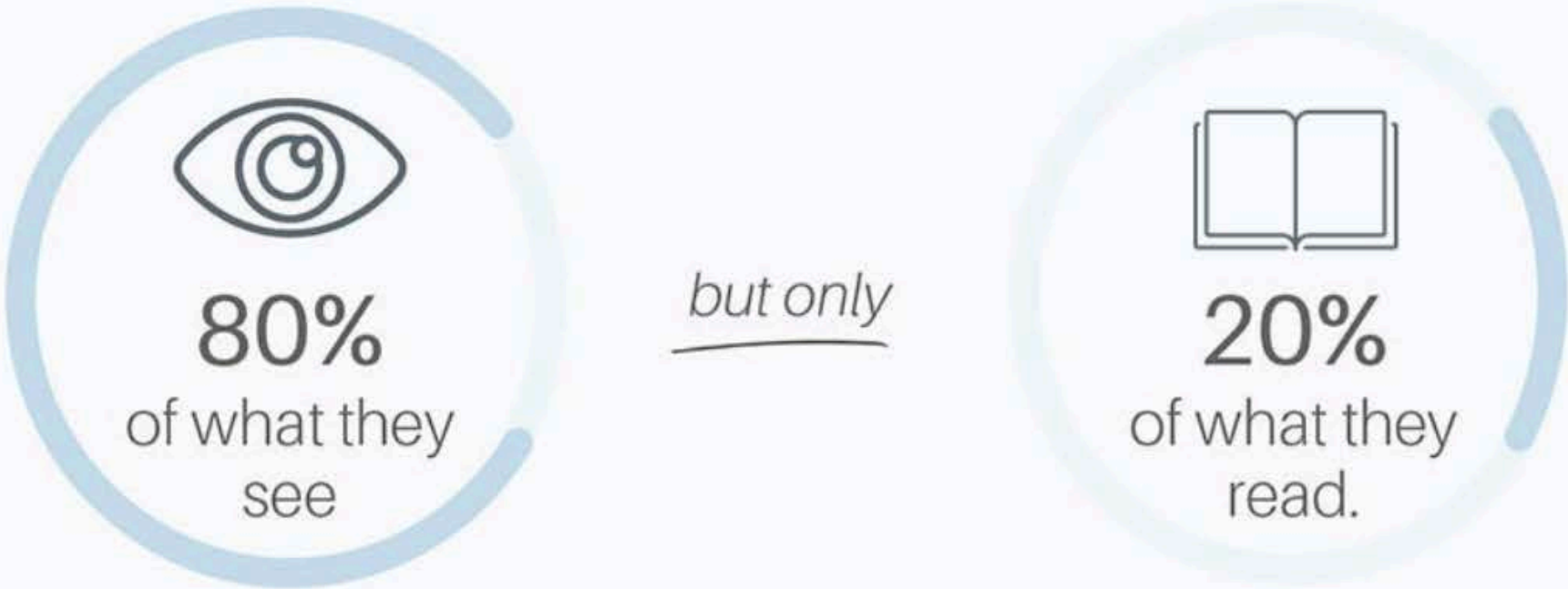
Our psyche is constructed around the ability to identify patterns



Merieb, E. N. & Hoehn, K. (2007). Human Anatomy & Physiology 7th Edition, Pearson International Edition.
Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V., & Sterling, P. (2006). How much the eye tells the brain. *Current Biology, 16*(14), 1428-1434.
*Political Psychology (*2003). Special Issue: Neuroscientific Contributions to Political Psychology. *Political Psychology* 24, 4.

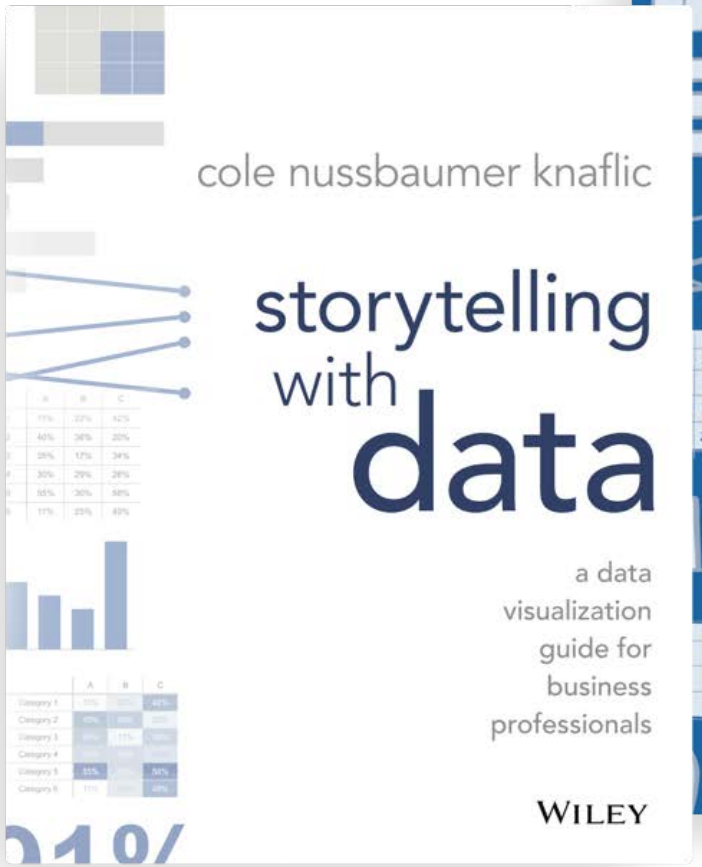
90% of all information transmitted to our brains is visual.

People remember...

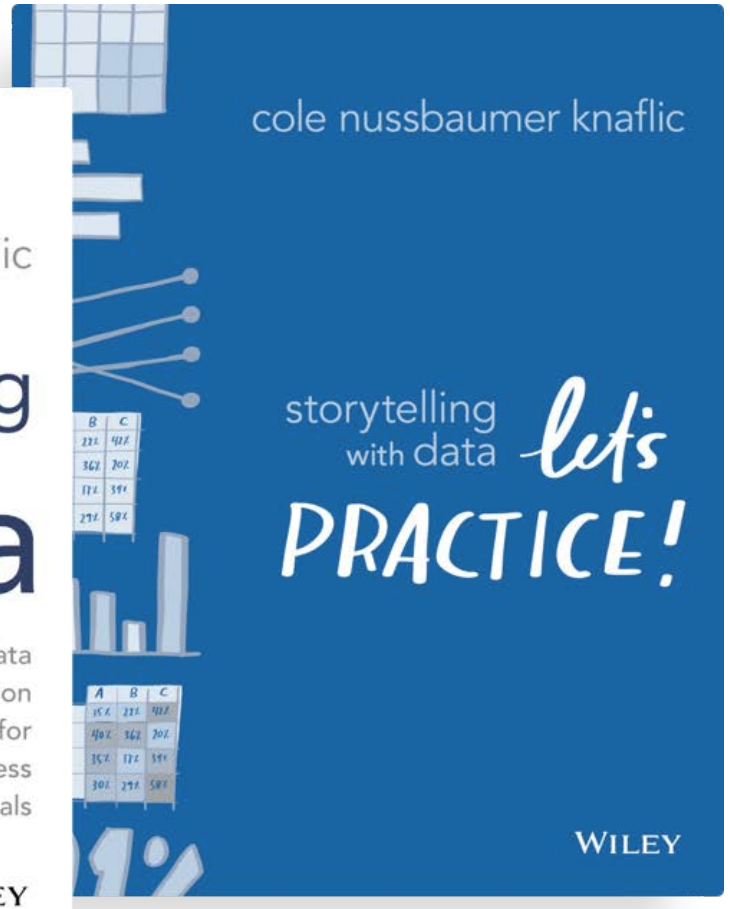


Source: https://www.visme.co/make-information-beautiful/cole-nussbaumer-knaflic/

Highly Recommended Read



www.amazon.com/dp/1119002257/



www.amazon.com/dp/1119621496/

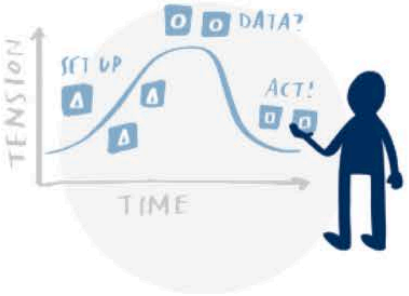
Storytelling with Data

understand the
context

choose an
effective visual

eliminate
clutter

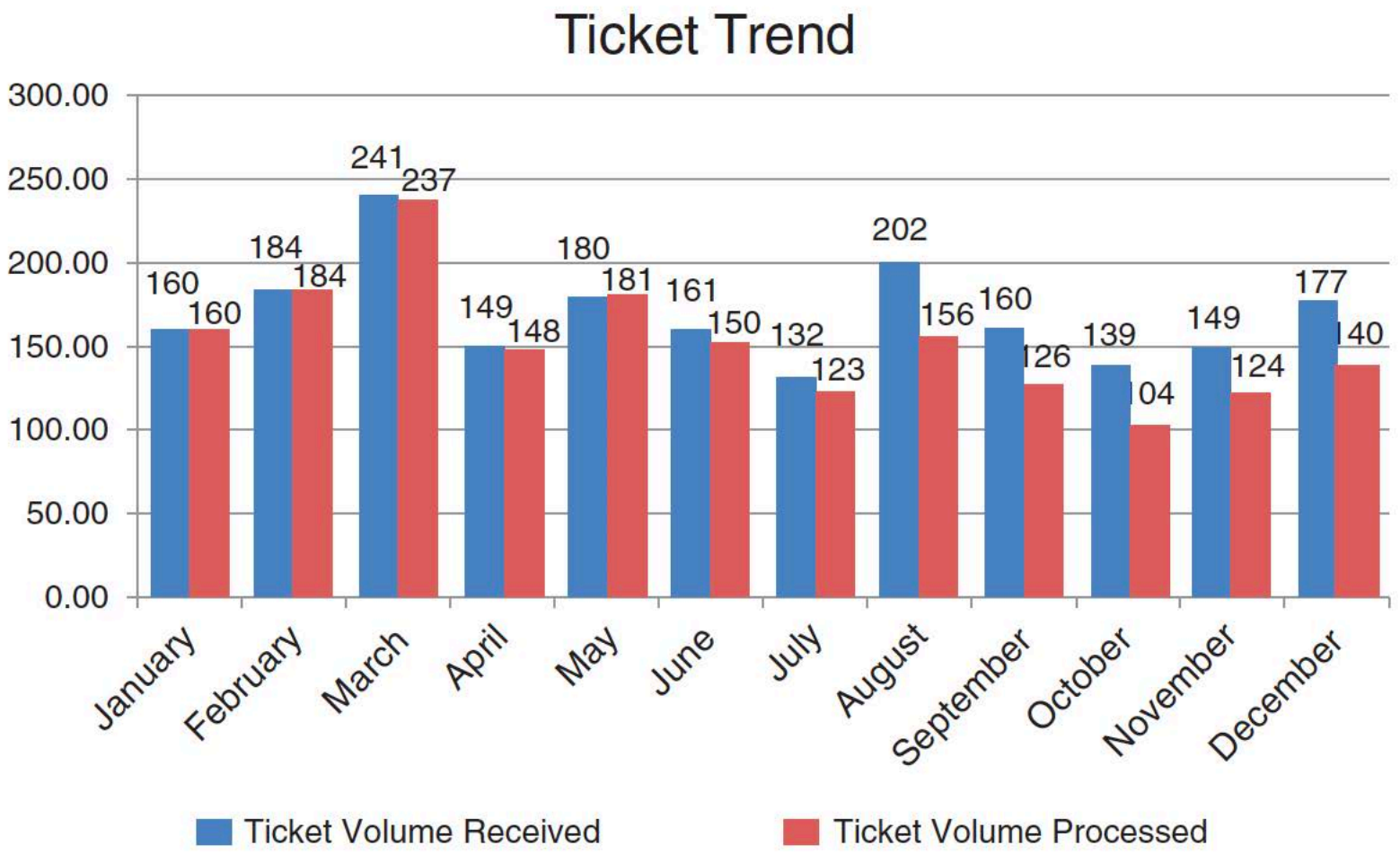
focus
attention



tell a
story

Source: http://www.storytellingwithdata.com/

Case 1: Hire more Employees



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 1: Hire more Employees

Please approve the hire of 2 FTEs

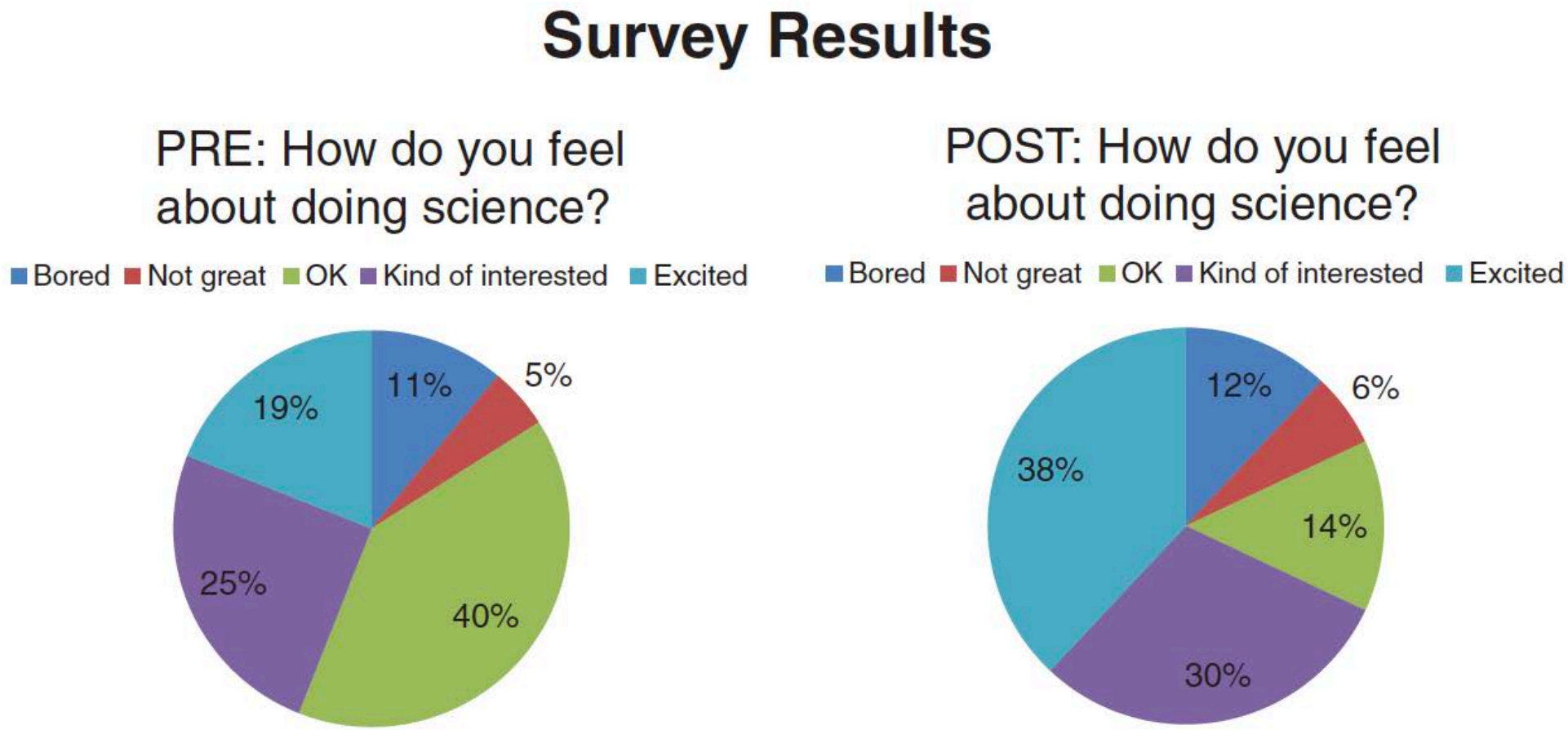
to backfill those who quit in the past year

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.

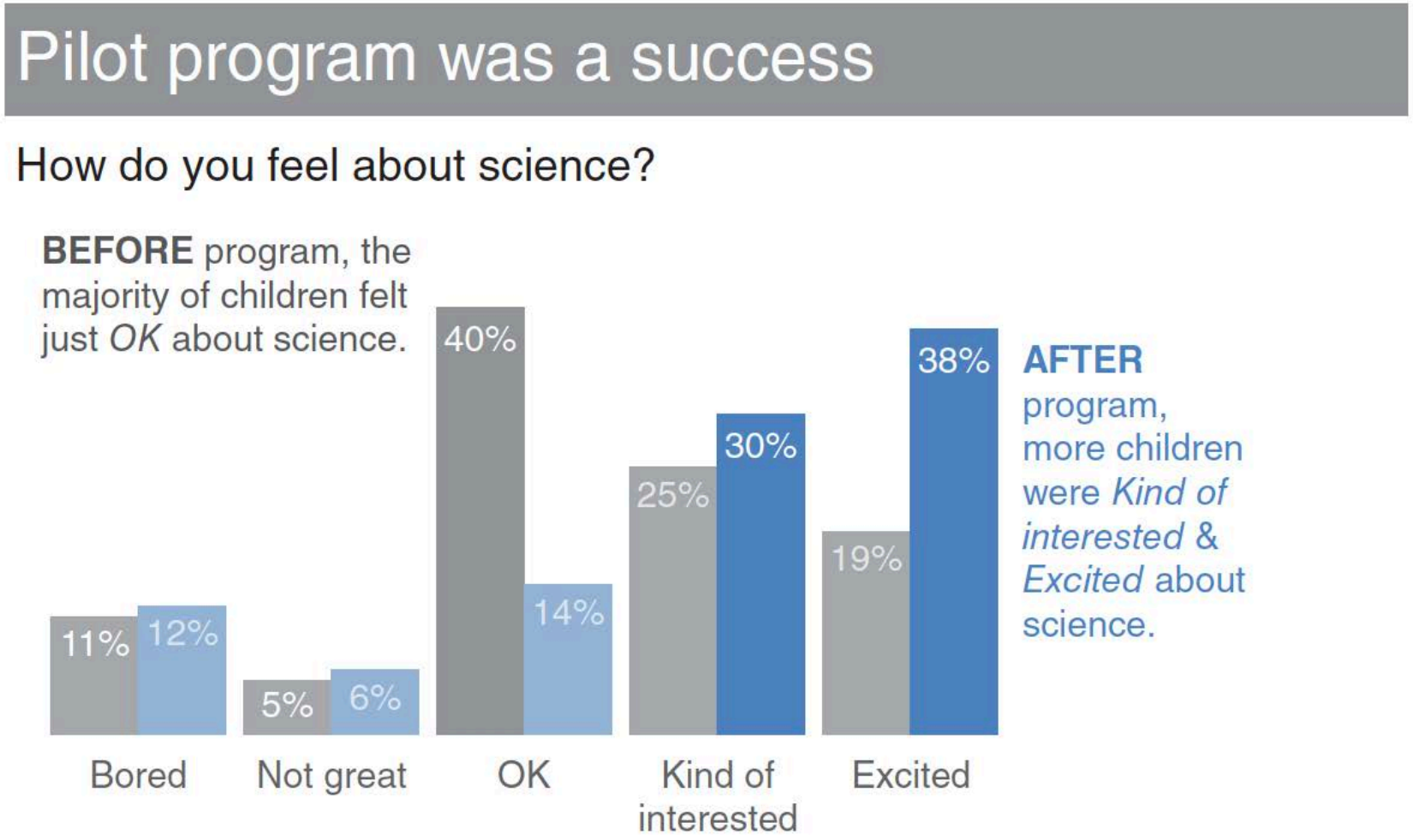
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 2: Survey about Science Class Pilot Program



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

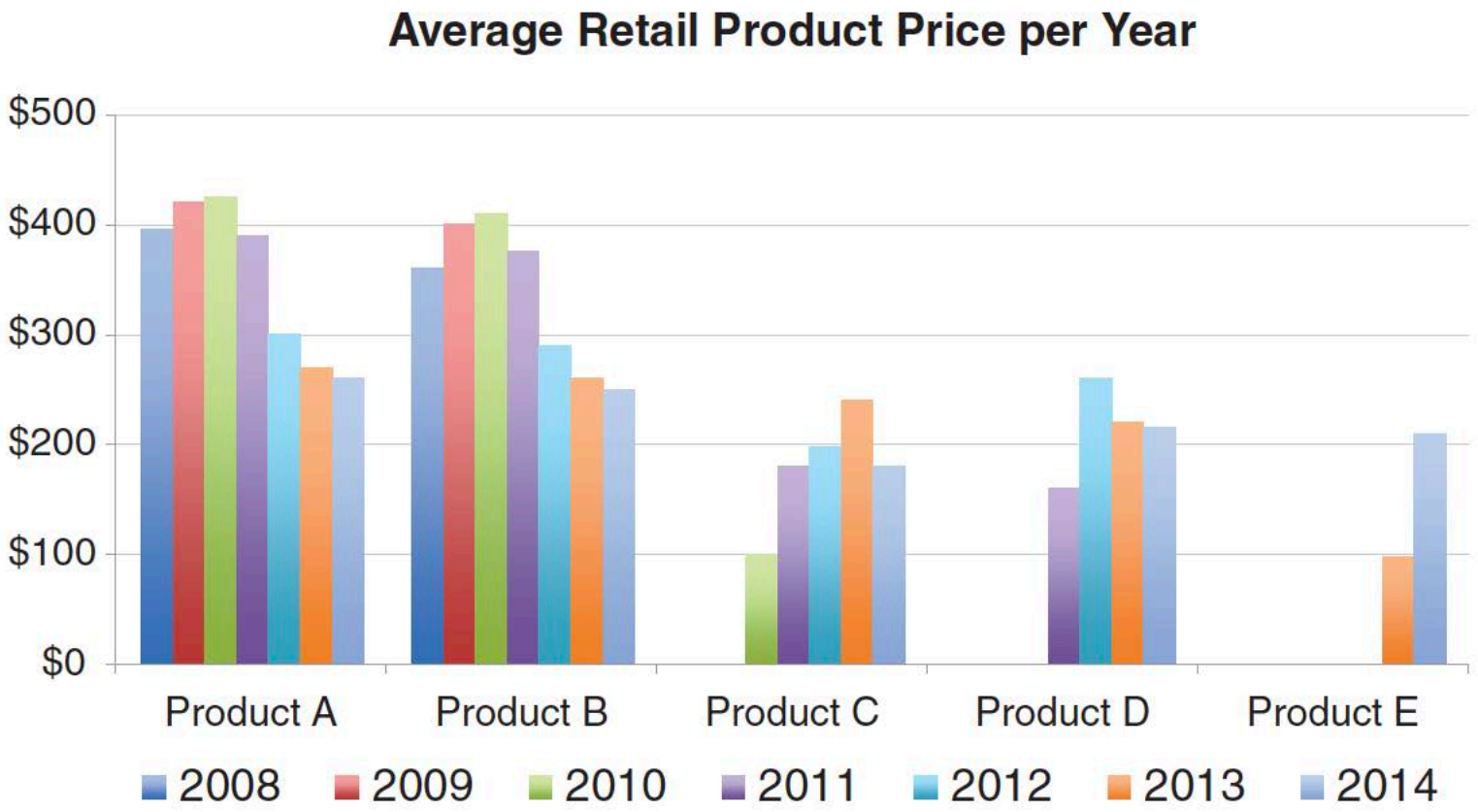
Case 2: Survey about Science Class Pilot Program



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

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

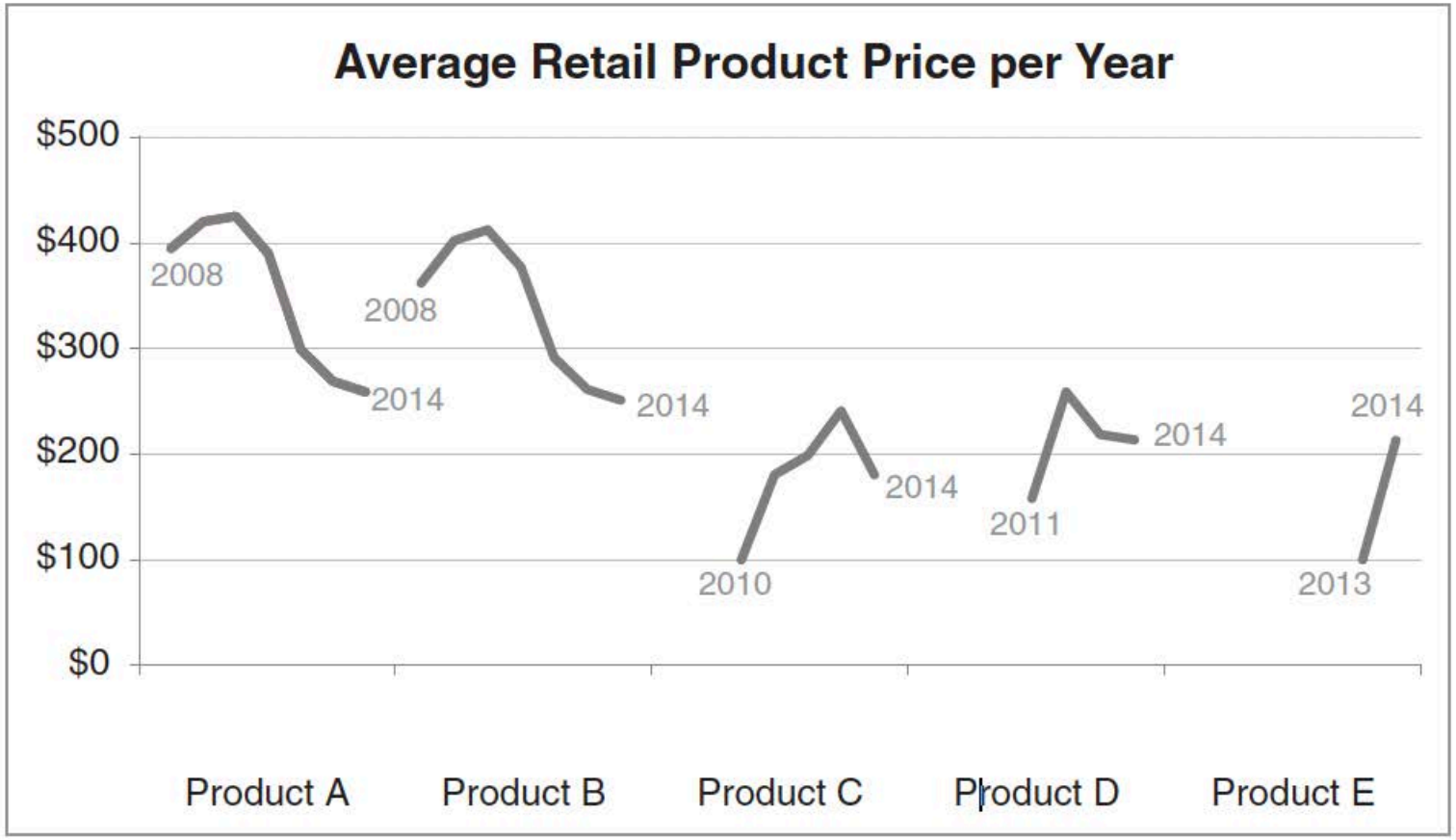
Case 3: Pricing a New Product



→ Suggest to set price at $150 to $200

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 3: Pricing a New Product

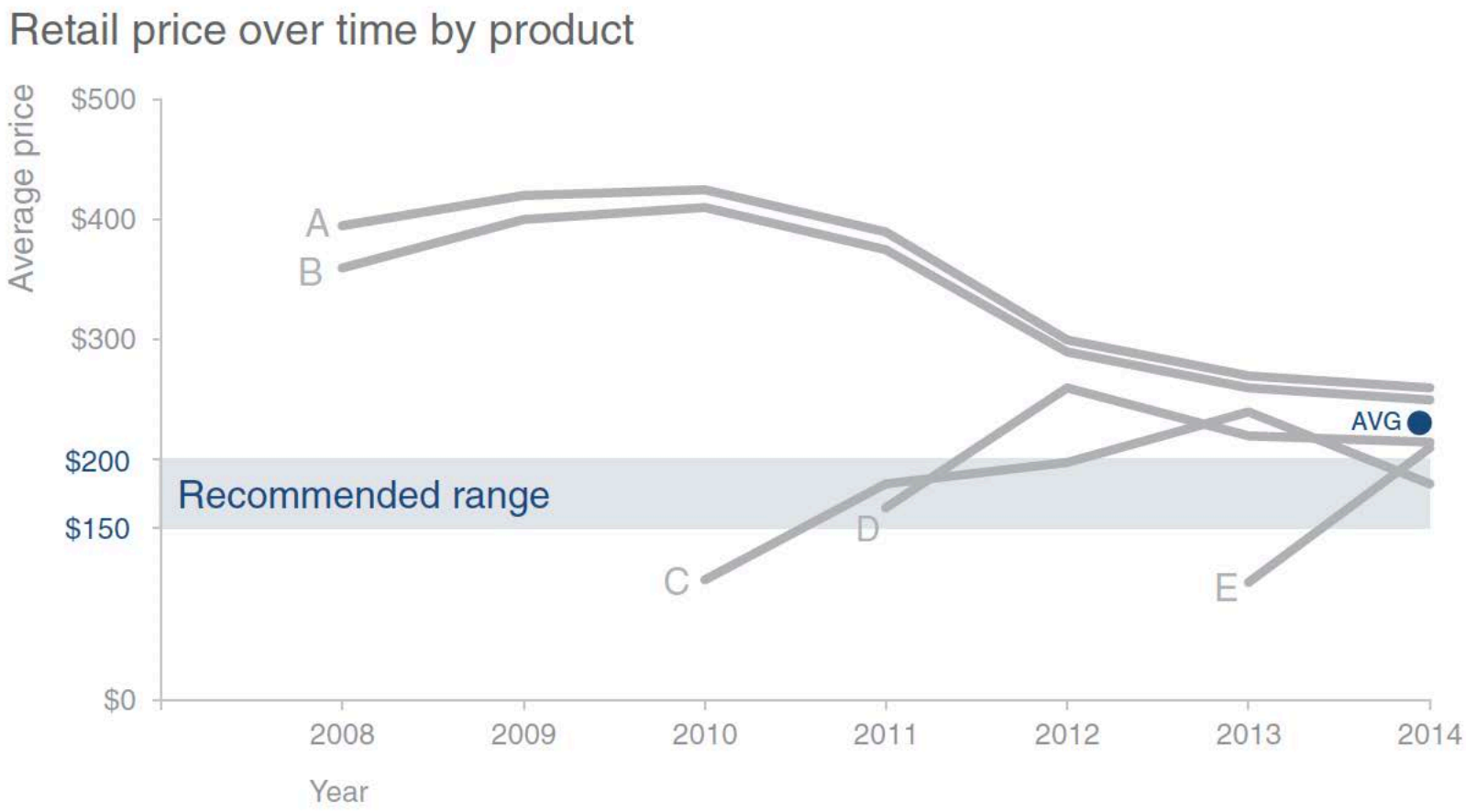


Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

→ Suggest to set price at $150 to $200

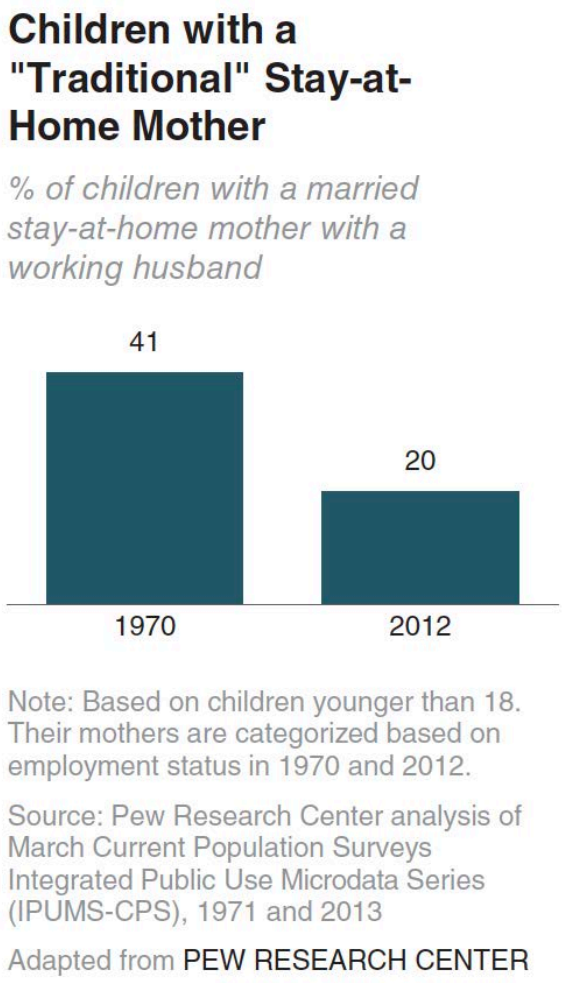
Case 3: Pricing a New Product

To be competitive, we recommend introducing our product *below the $223 average* price point in the **$150–$200 range**



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Focus on What's Important



vs.

20%

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

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Tables: Minimize and Highlight

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

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%

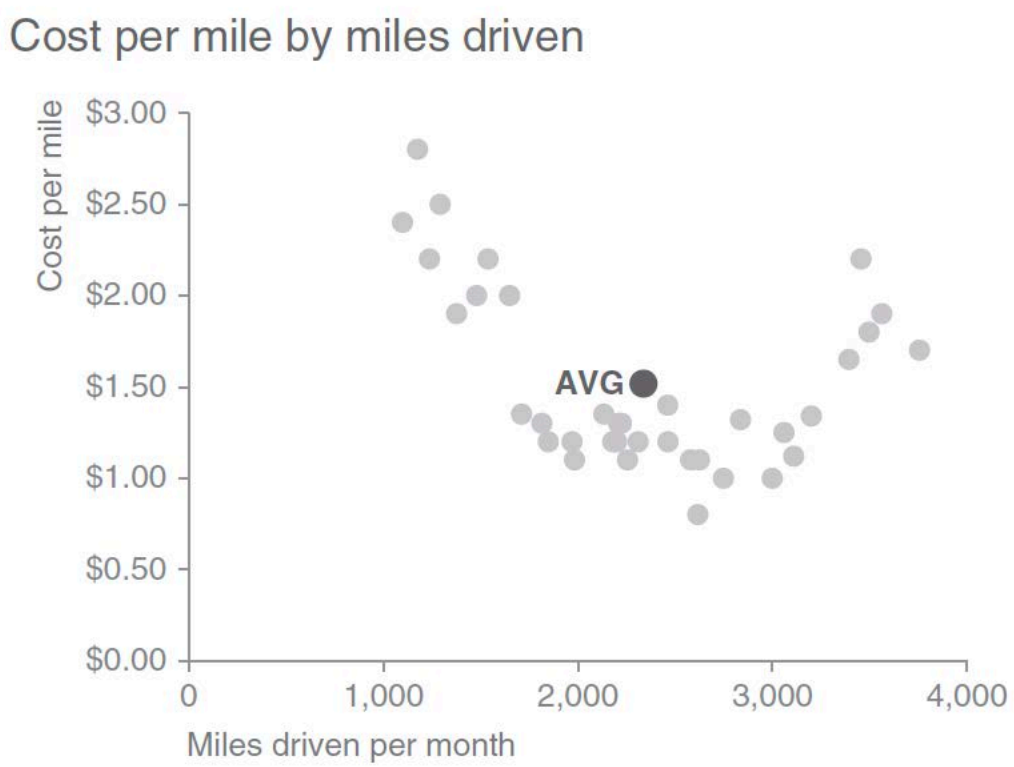
Heatmap

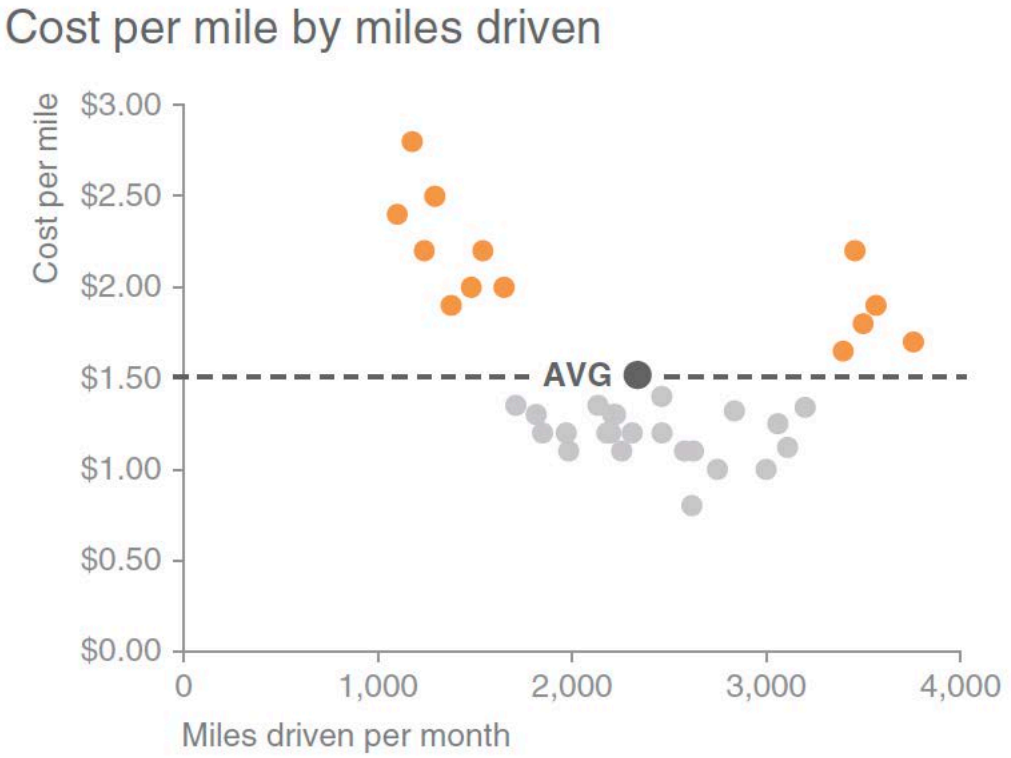
LOW-**HIGH**

	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%

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

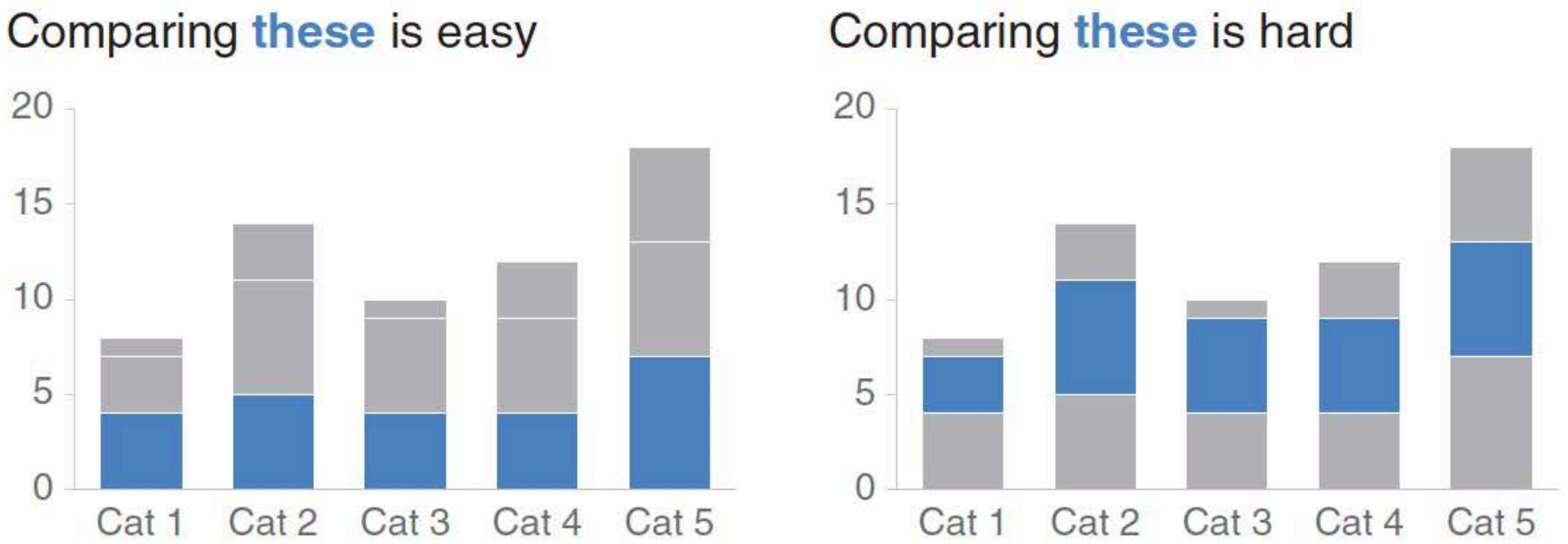
Scatterplots: Draw Attention to What Matters





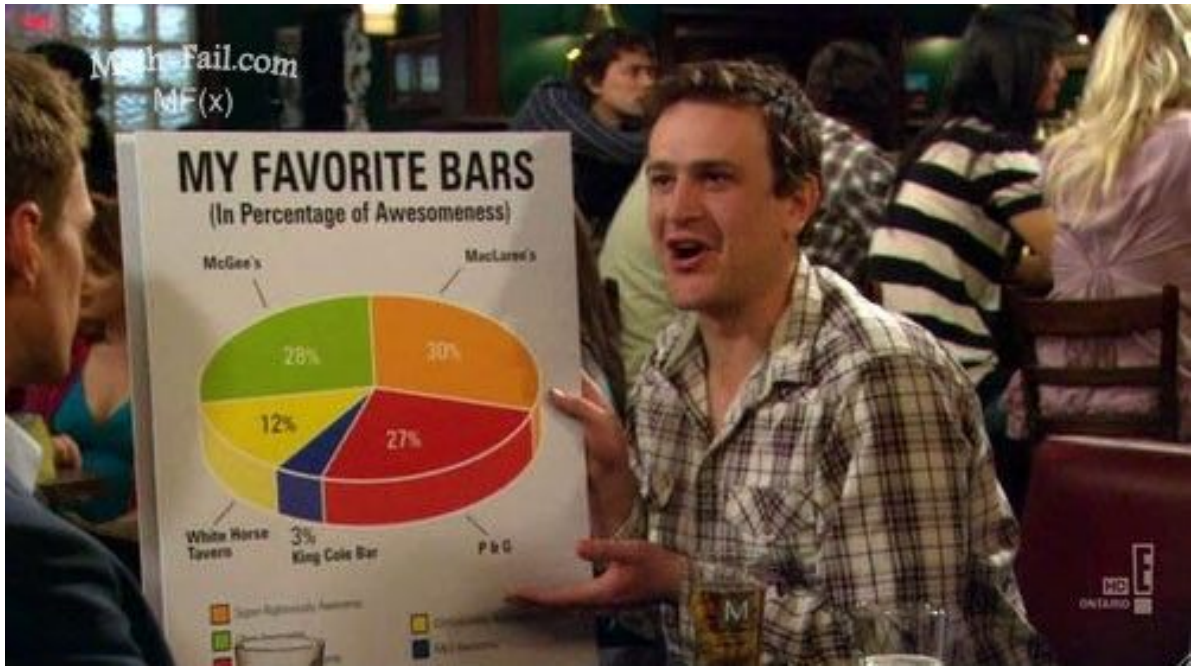
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Help Your Audience with Easy Visuals



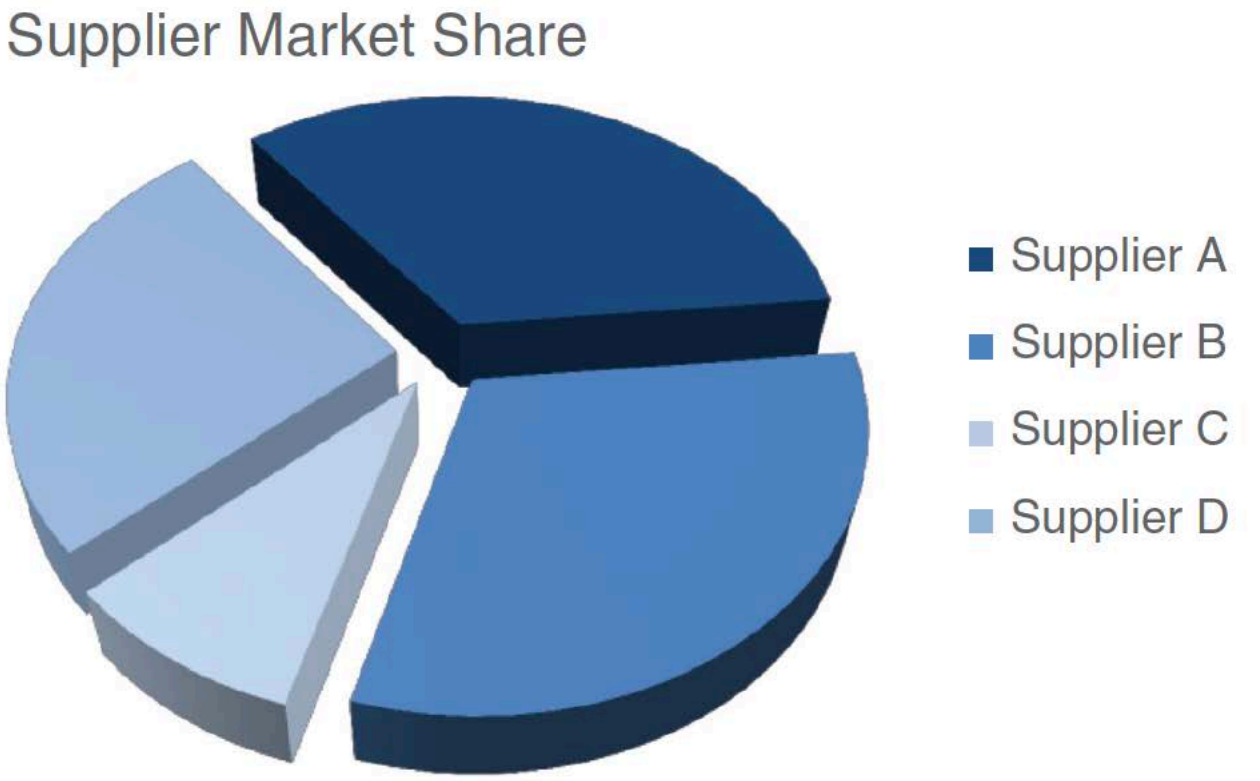
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Bars vs. Pies



www.youtube.com/watch?v=f_J8QU1m0Ng

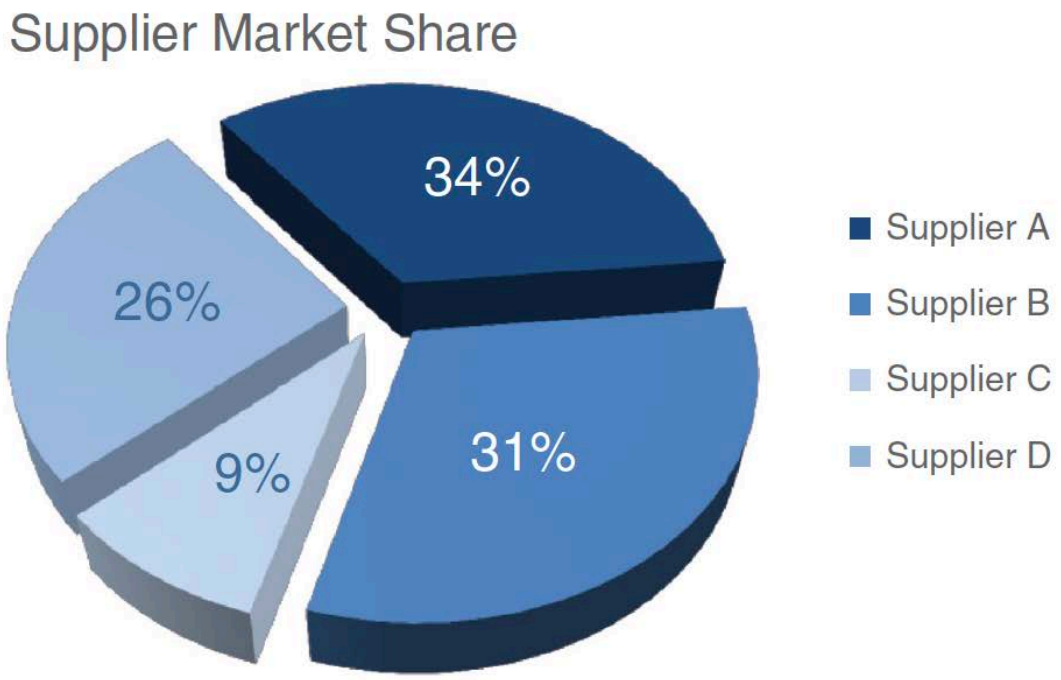
The Thing about Pie Charts …

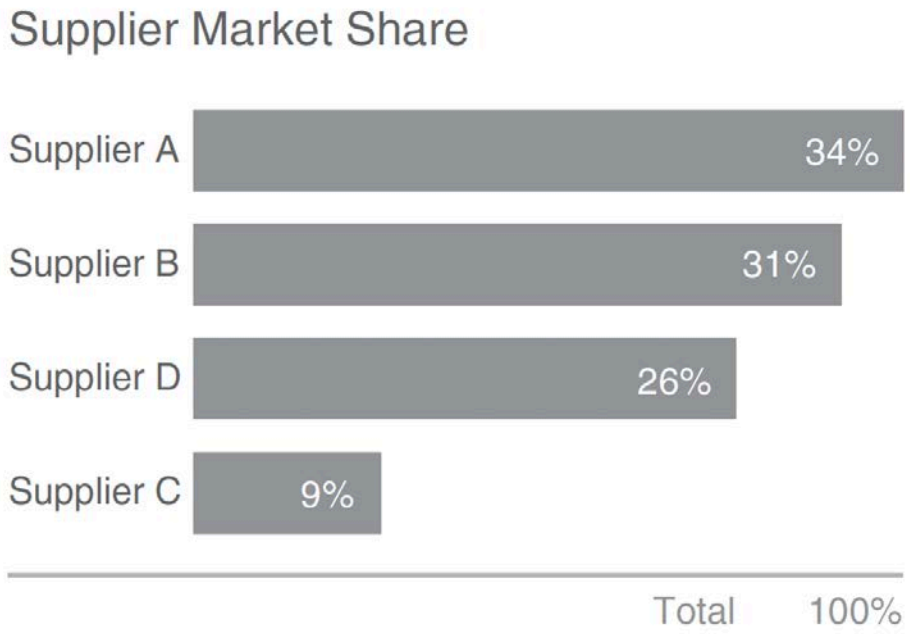


Who has the largest market share?

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Avoid Pie Charts—and 3D Charts in General

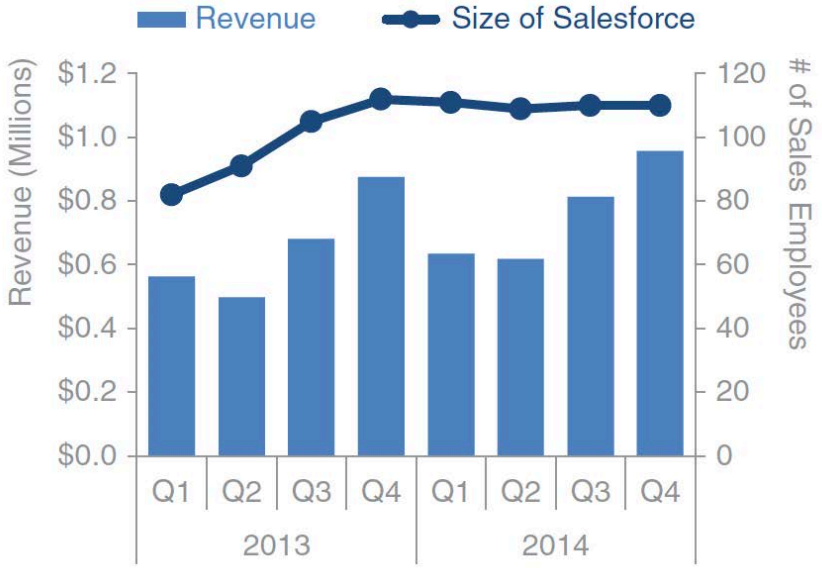




Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Secondary Axis

Secondary y-axis

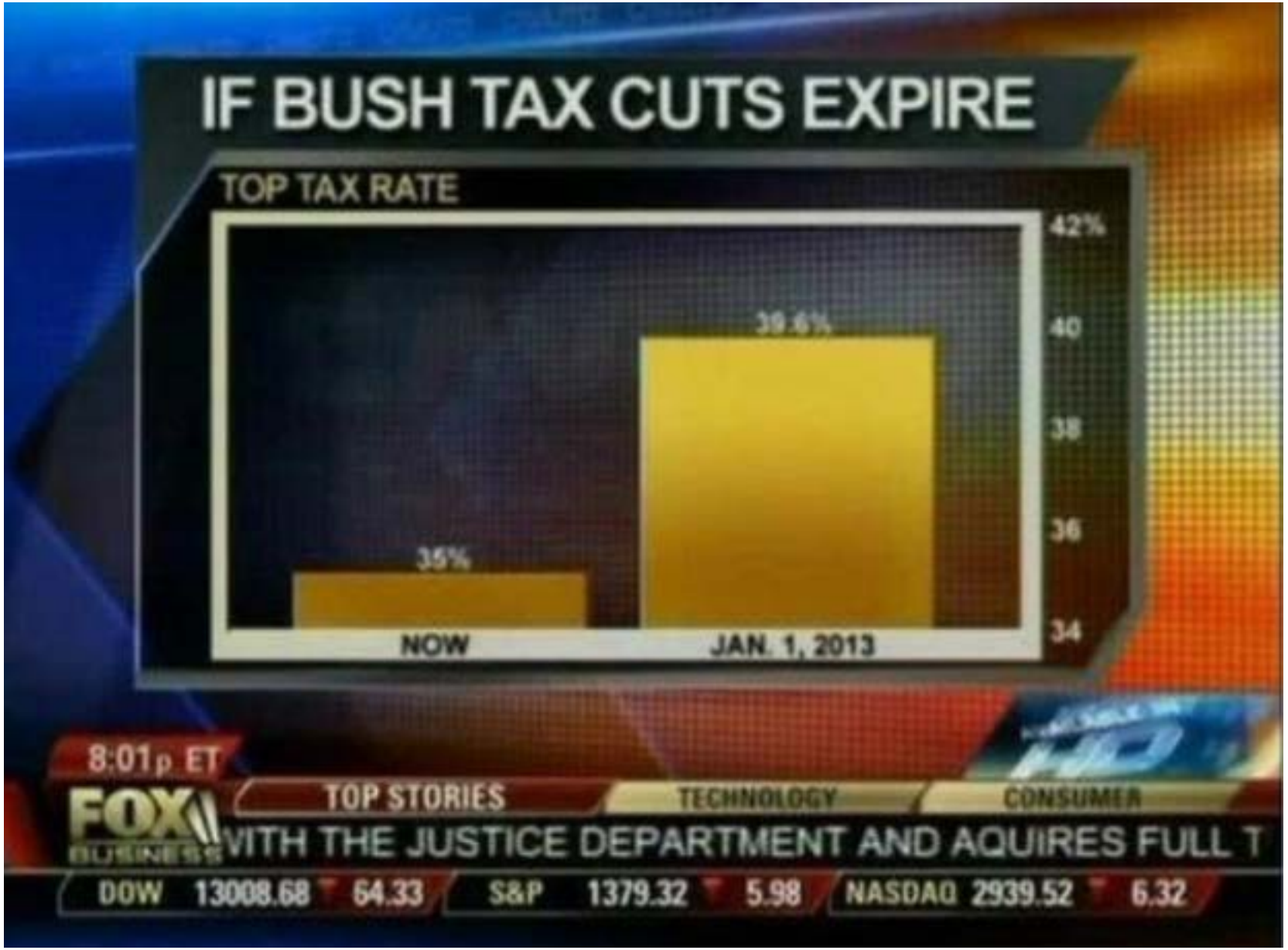


Alternative 1: **label directly**

Alternative 2: **pull apart vertically**

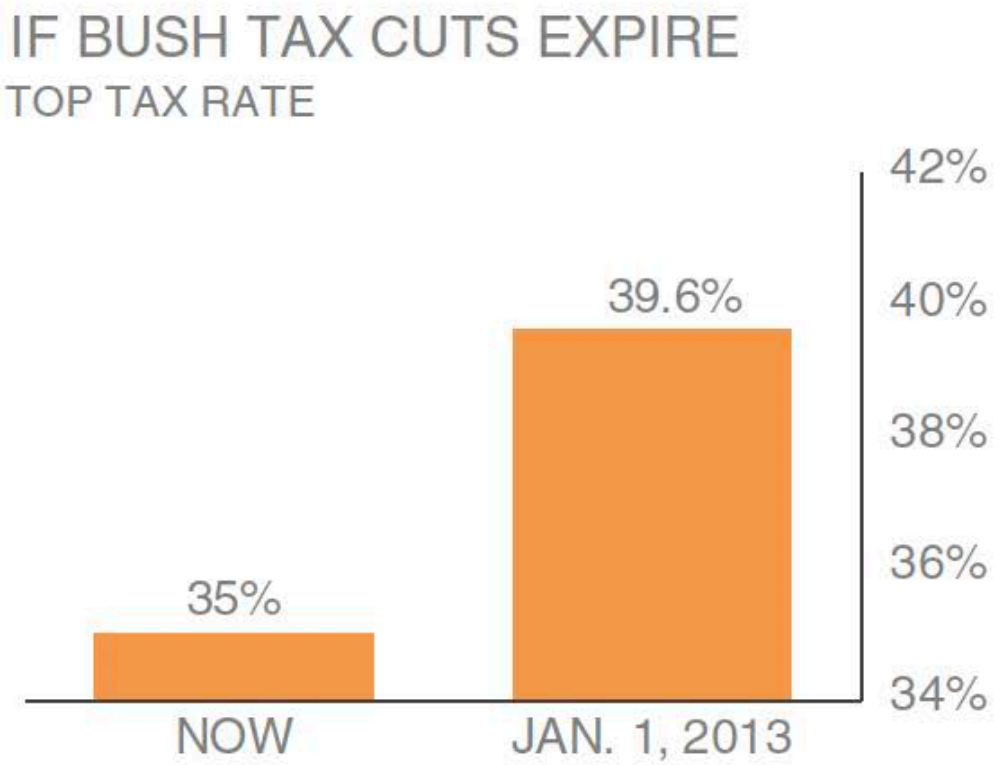
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 4: Scales

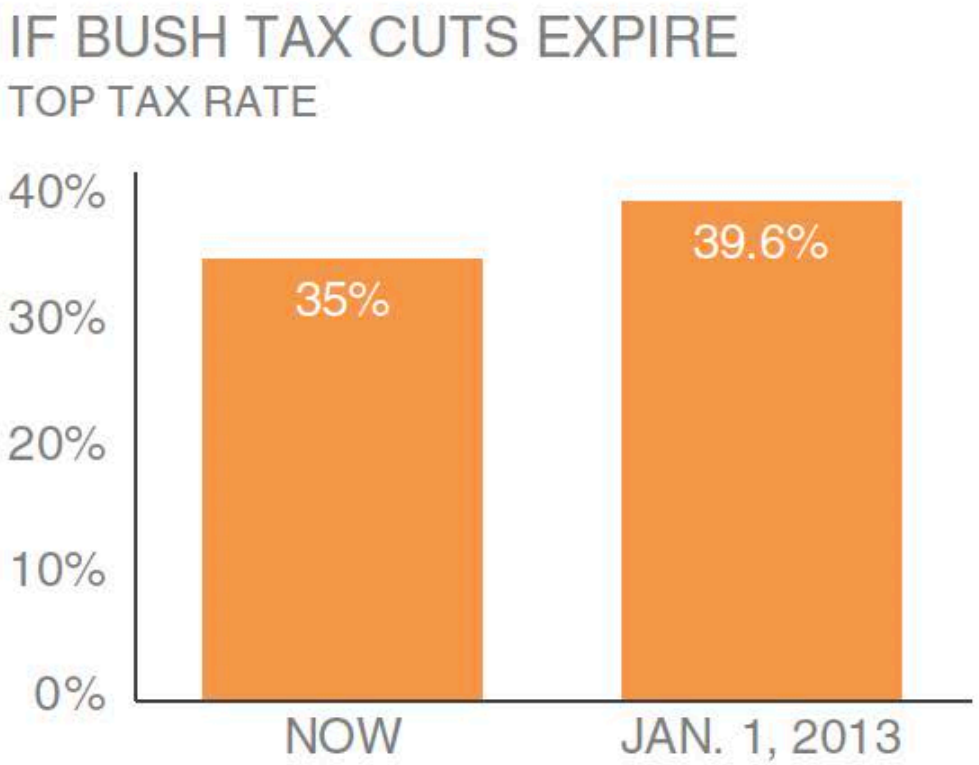


Case 4: Scales

Non-zero baseline: as originally graphed

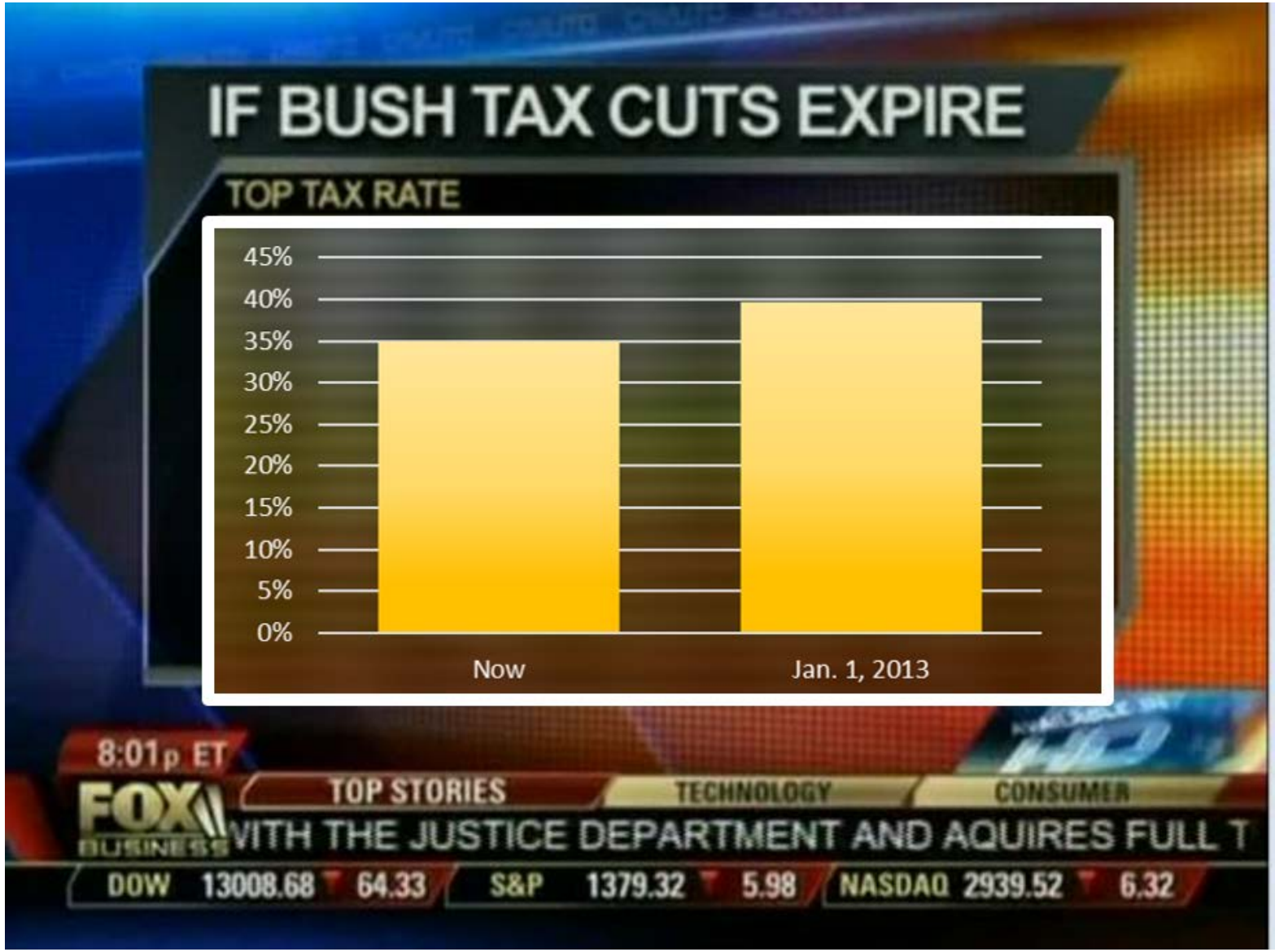


Zero baseline: as it should be graphed



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 4: Scales



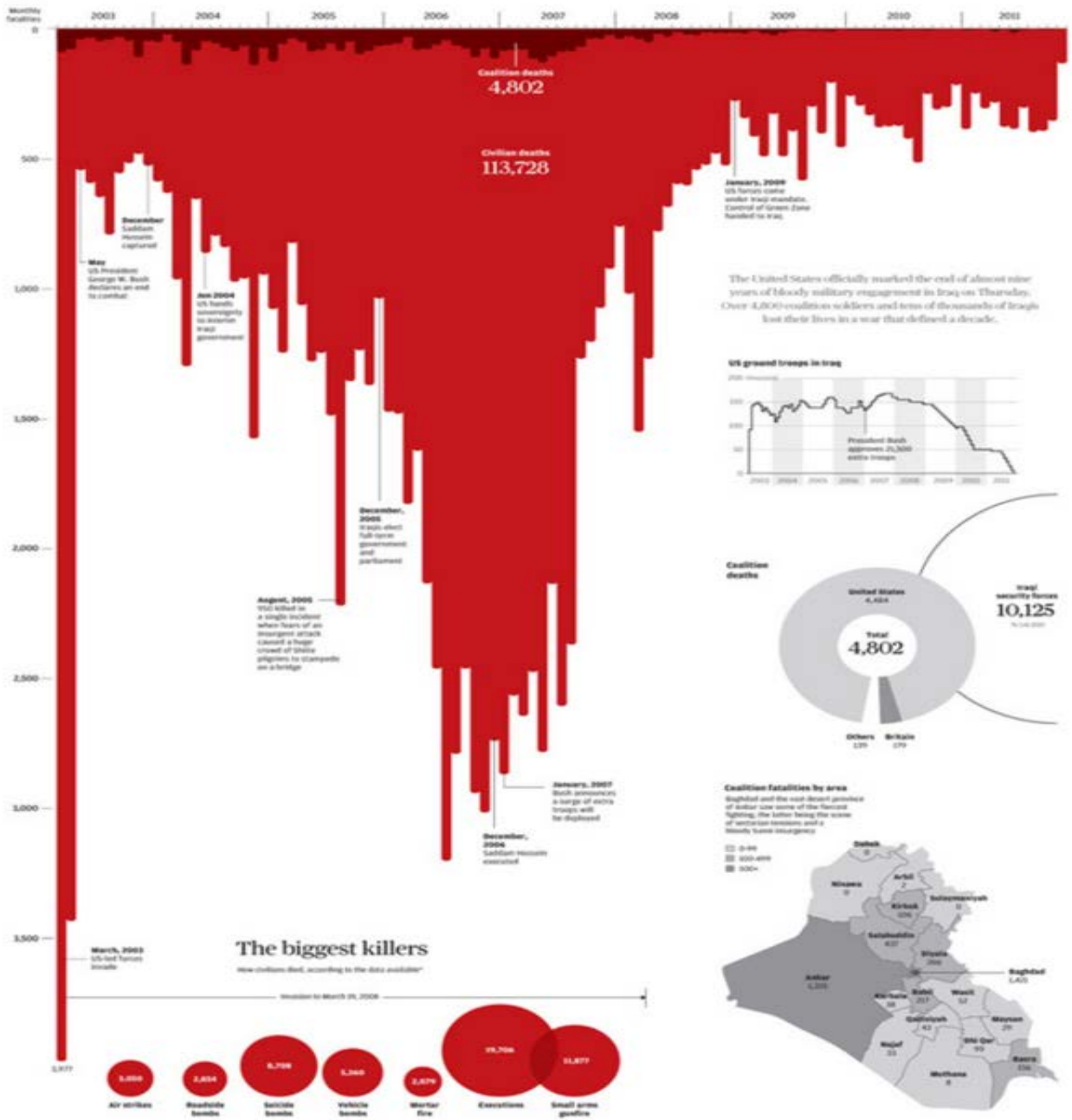
Ethics and Data Visualization

What if changing the scale on a bar chart or otherwise manipulating the data better reinforces the point you want to make?

- Misleading your audience by inaccurately visualizing data is NOT OK
- Serious ethical concerns
- Risky territory

Case 5: Color and Orientation

Iraq's Bloody Toll



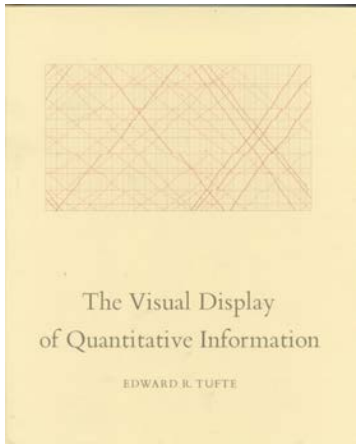
Source: https://www.infoworld.com/article/3088166/why-how-to-lie-with-statistics-did-us-a-disservice.html

Cognitive Load

Every single element you add to a page or screen creates more and more cognitive load on the part of your audience

→ It takes them brain power to process

- When we ask a computer to do work, we are relying on the computer's processing power.
- When we ask our audience to do work, we are leveraging their mental processing power
  → This is cognitive load
- Humans' brains have a finite amount of this mental processing power.
- Cognitive load can make things feel more complicated than they actually are



The larger the share of a graphic's ink devoted to data, the better. *Edward Tufte*

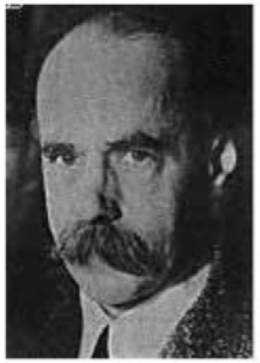
Maximize the signal-to-noise ratio

- Signal is the information we want to communicate
- Noise are elements that either:
  - don't add to
  - or detract from

  the message we are trying to communicate

Gestalt Principles of Visual Perception

The Gestalt School of Psychology set out in the early 1900s
to understand how individuals perceive order in the world around them.



Max Wertheimer, 1880-1943

Kurt Koffka, 1886-1941

Wolfgang Kohler, 1887-1967

Six Gestalt principles most important to data visualization:

- Proximity
- Similarity
- Enclosure
- Closure
- Continuity
- Connection

Proximity

We tend to think of objects that are physically close together as belonging to part of a group

Leverage proximity in table design

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Similarity

Objects that are of similar
- color
- shape
- size
- orientation

are perceived as related or belonging to part of a group.

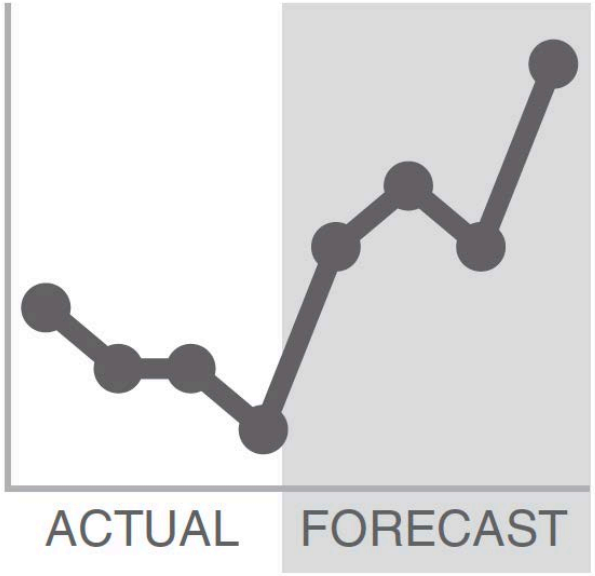
This can be leveraged in scatter plots, but also in tables to help draw our audience's eyes in *the direction we want them to focus*

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Enclosure

Humans (usually) think of objects that are physically enclosed together as belonging to part of a group

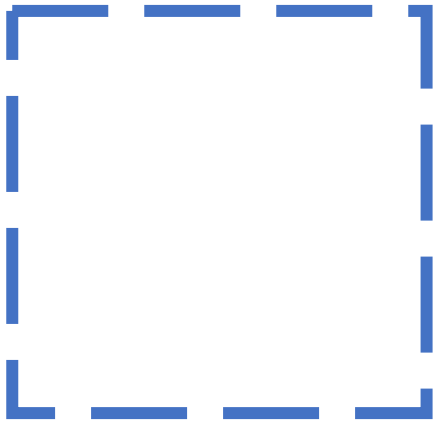
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

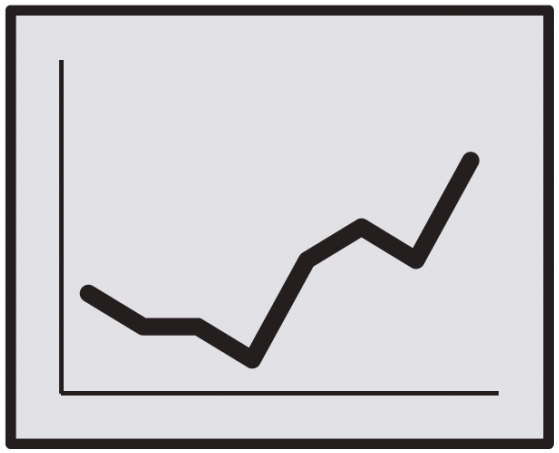


Closure

Humans like things to be:
1. Simple
2. Fit in the constructs that are already in their minds

Because of this, humans tend to perceive a set of individual elements as a single, recognizable shape—when parts of a whole are missing, our eyes fill in the gap.



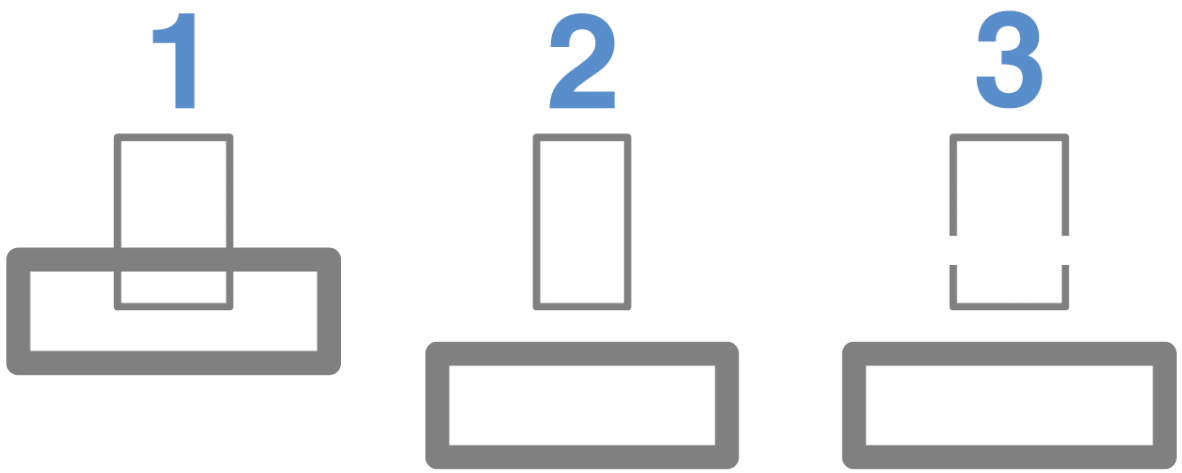


Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

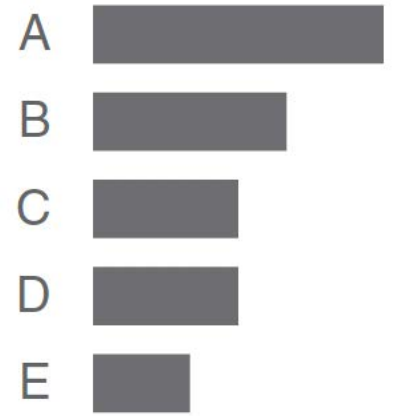
Continuity

When looking at objects, our eyes seek the smoothest path and naturally create continuity in what we see, even where it may not explicitly exist.

If I take the objects (1) and pull them apart, most people will expect to see what is shown next (2), whereas it could as easily be what is shown after that (3).



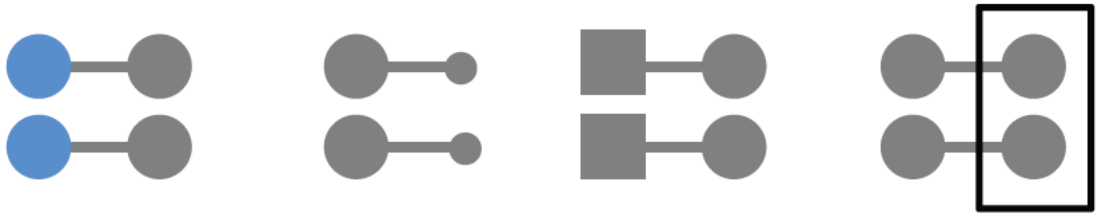
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.



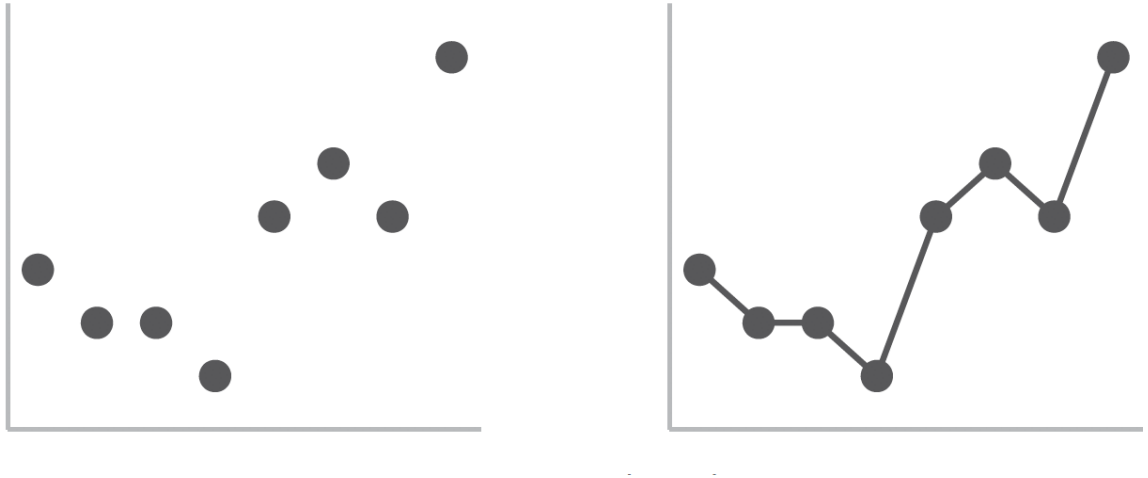
Do these bars line-up?

Connection

- Humans tend to think of objects that are physically connected as part of a group
- The connective property typically has a stronger associative value than similar color, size, or shape
- You can combine Gestalt principles to create the visual hierarchy

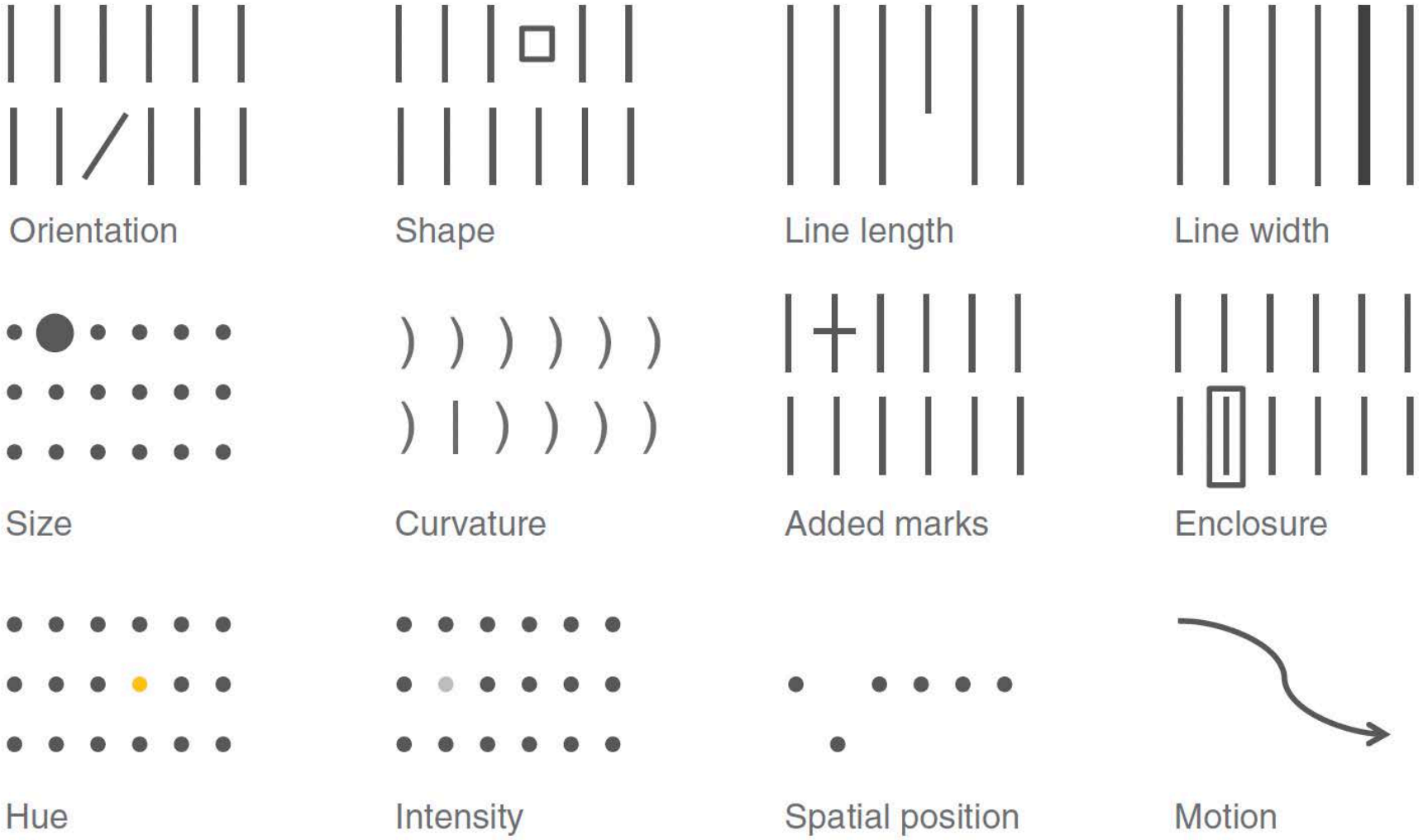


Connection is also useful to show order (e.g., of events)



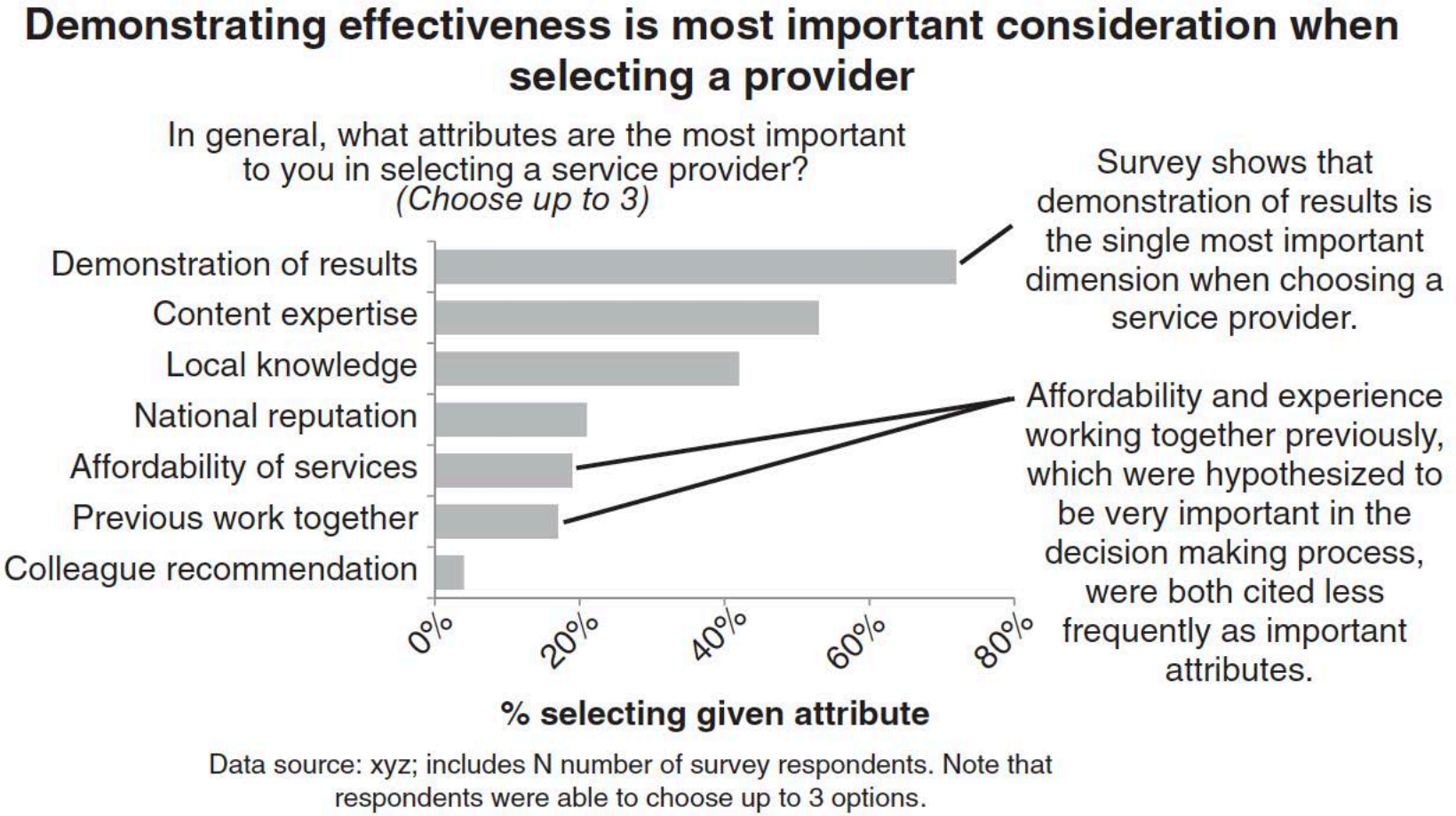
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Elements of Visualization



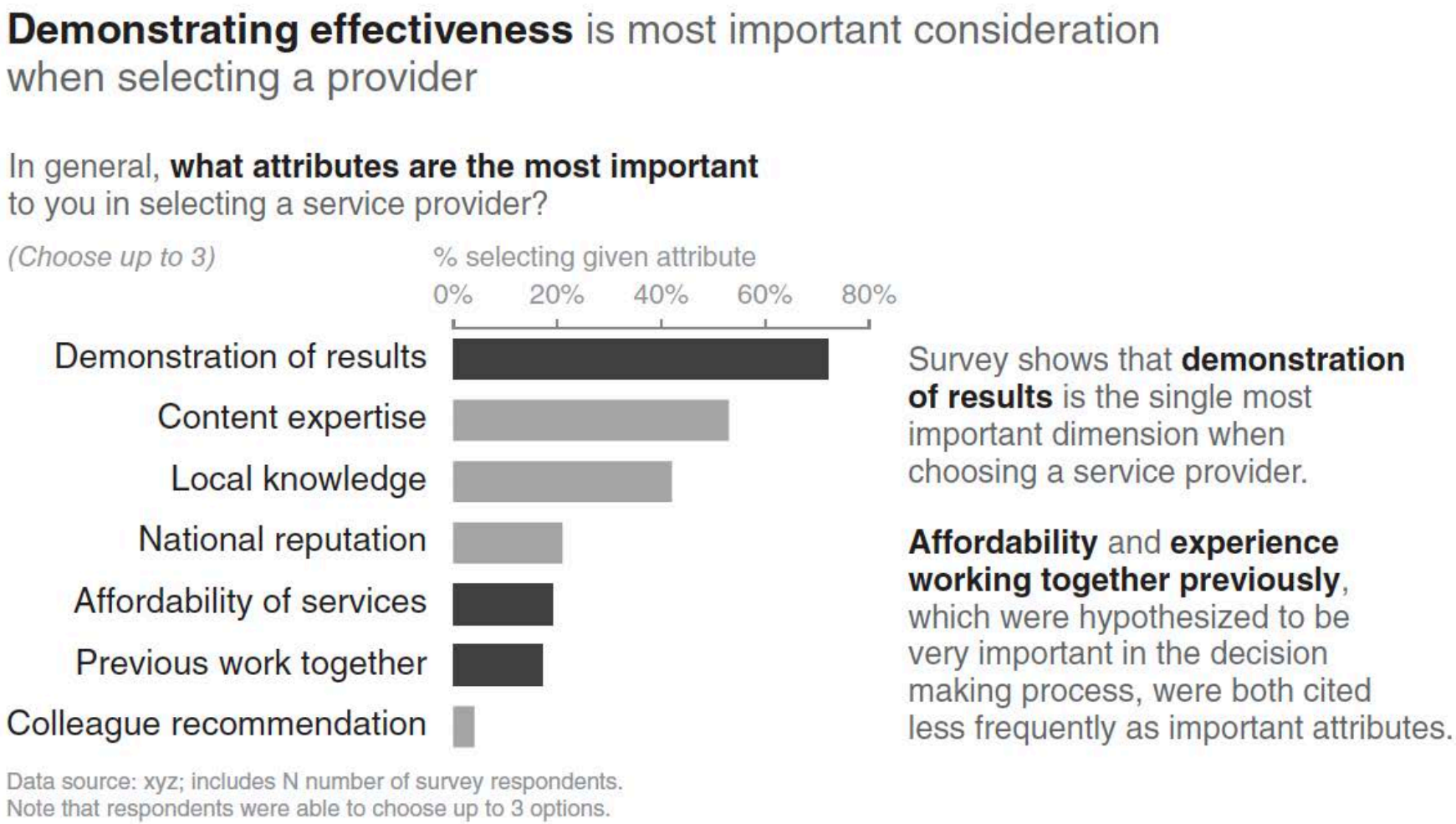
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 6: Visual Order



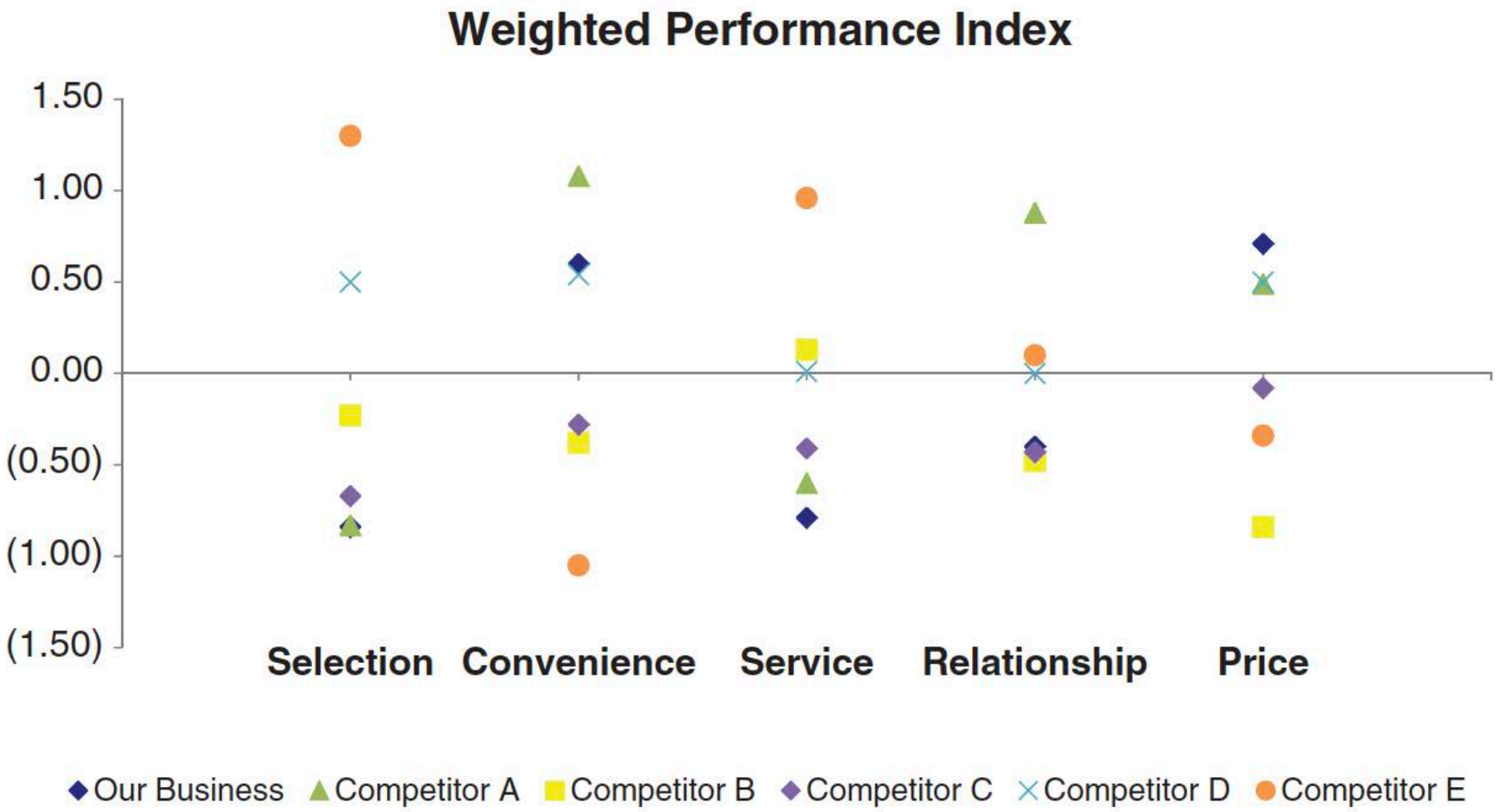
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 6: Visual Order



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 7: Visual Complexity



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

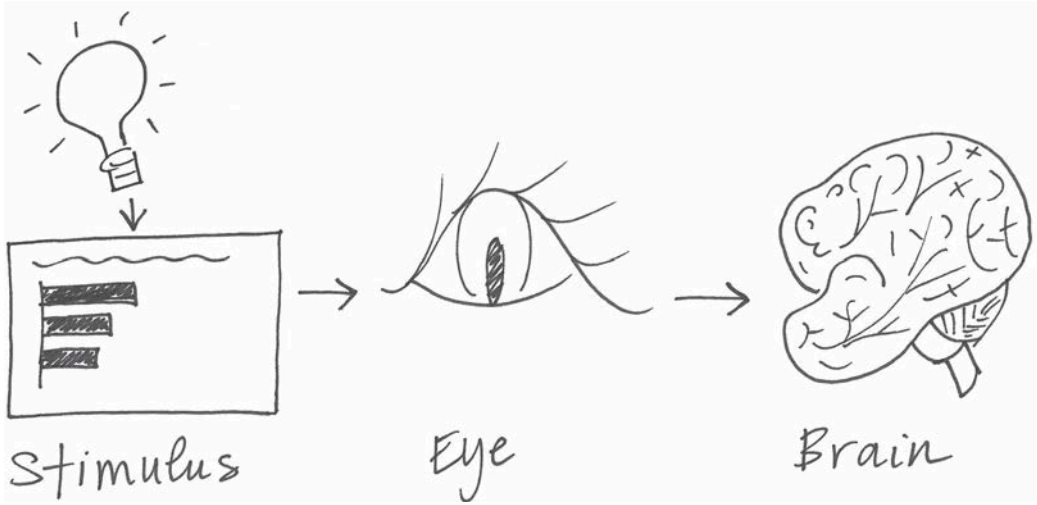
Case 7: Visual Complexity

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Understanding Memory

Iconic memory

- Super fast and only fraction of a second
- happens without you consciously realizing it
- Stimulated when we look at the world around us



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Short-term memory

- Very limited
- Humans can keep about four chunks of visual information in their short-term memory at a given time.

Long-term memory

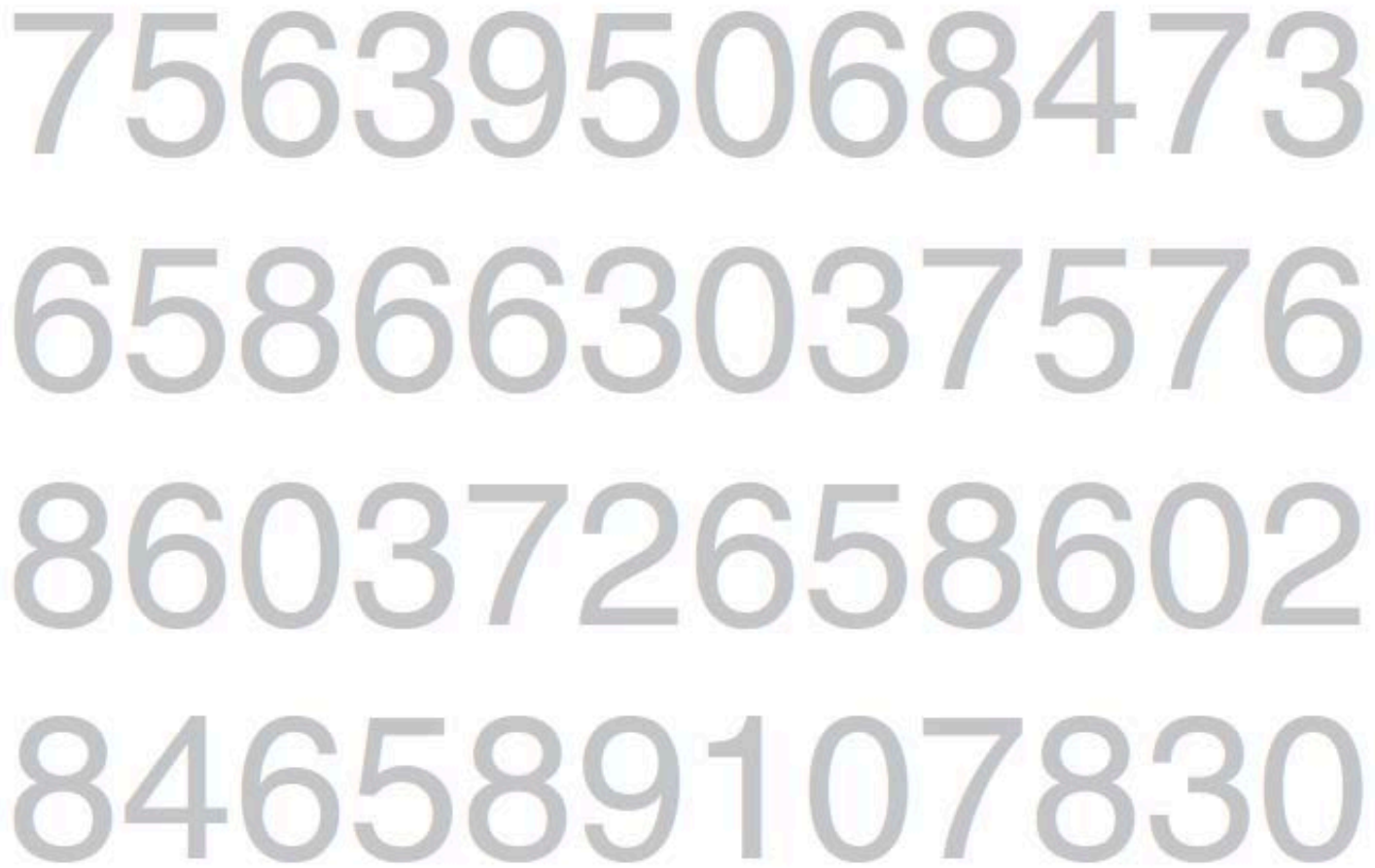
- Built up over a lifetime
- Vitally important for pattern recognition
- Required for general cognitive processing

Oblivion

- Likely lost forever

The Power of Preattentive Attributes

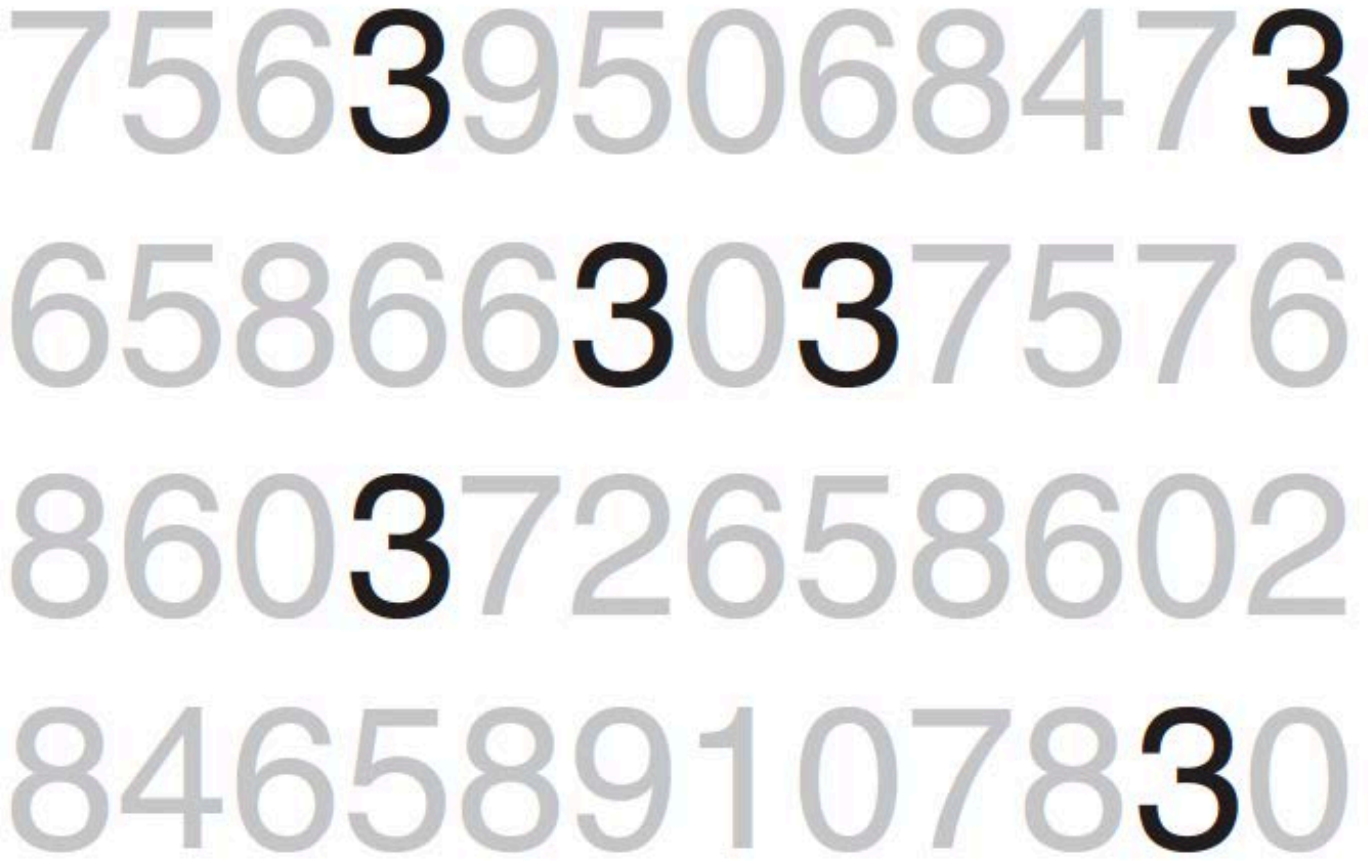
How many 3's are in this series of integers?



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

The Power of Preattentive Attributes

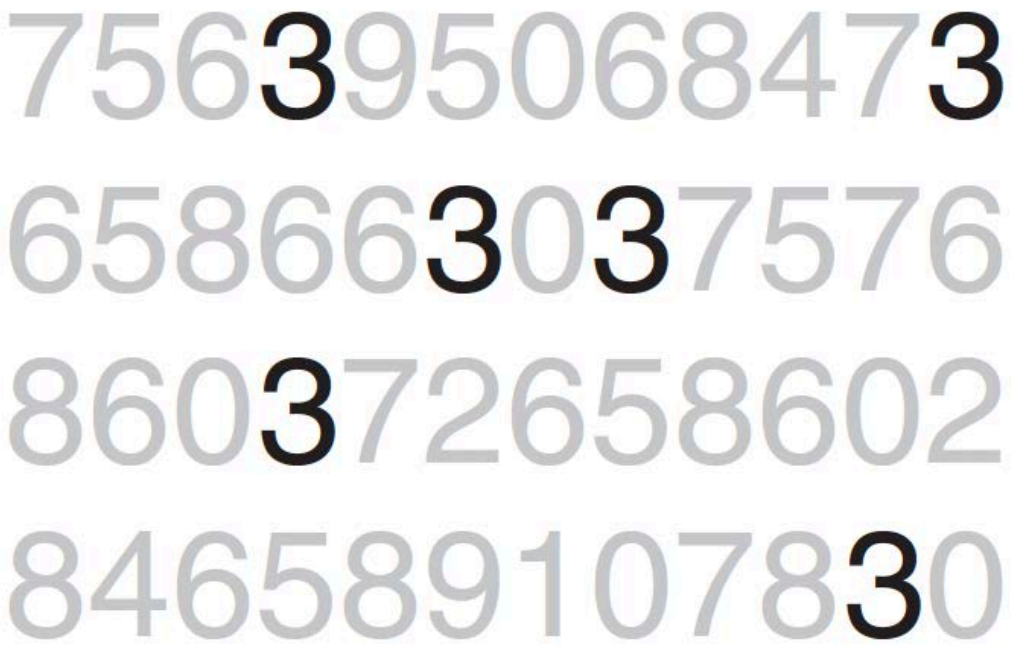
How many 3's are in this series of integers?



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

The Power of Preattentive Attributes

How many 3's are in this series of integers?



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

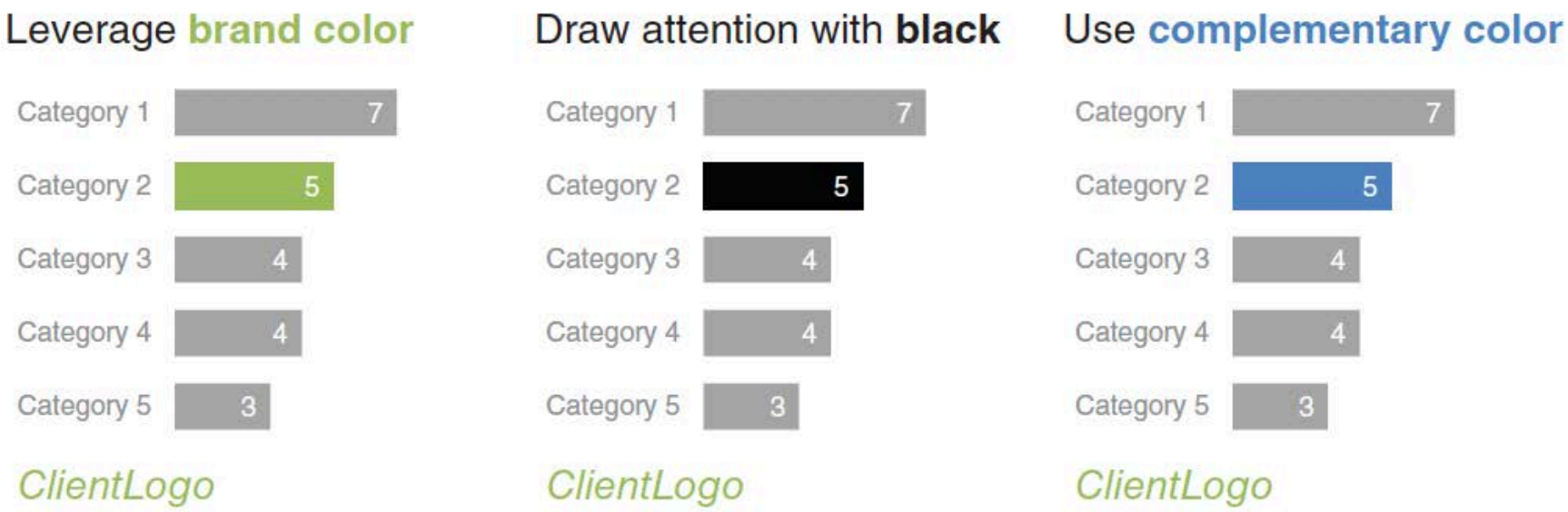
- You don't have time to blink
- You don't have time to think
- Suddenly there are six 3's in front of you!

This is remarkable
This is profoundly powerful

Can use preattentive attributes strategically:
- Make your audience see what you want them to see
- before they even know they're seeing it!

Also works with text

Case 8: Preattentive Memory and Color



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 9: Don’t go Overboard on Color

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

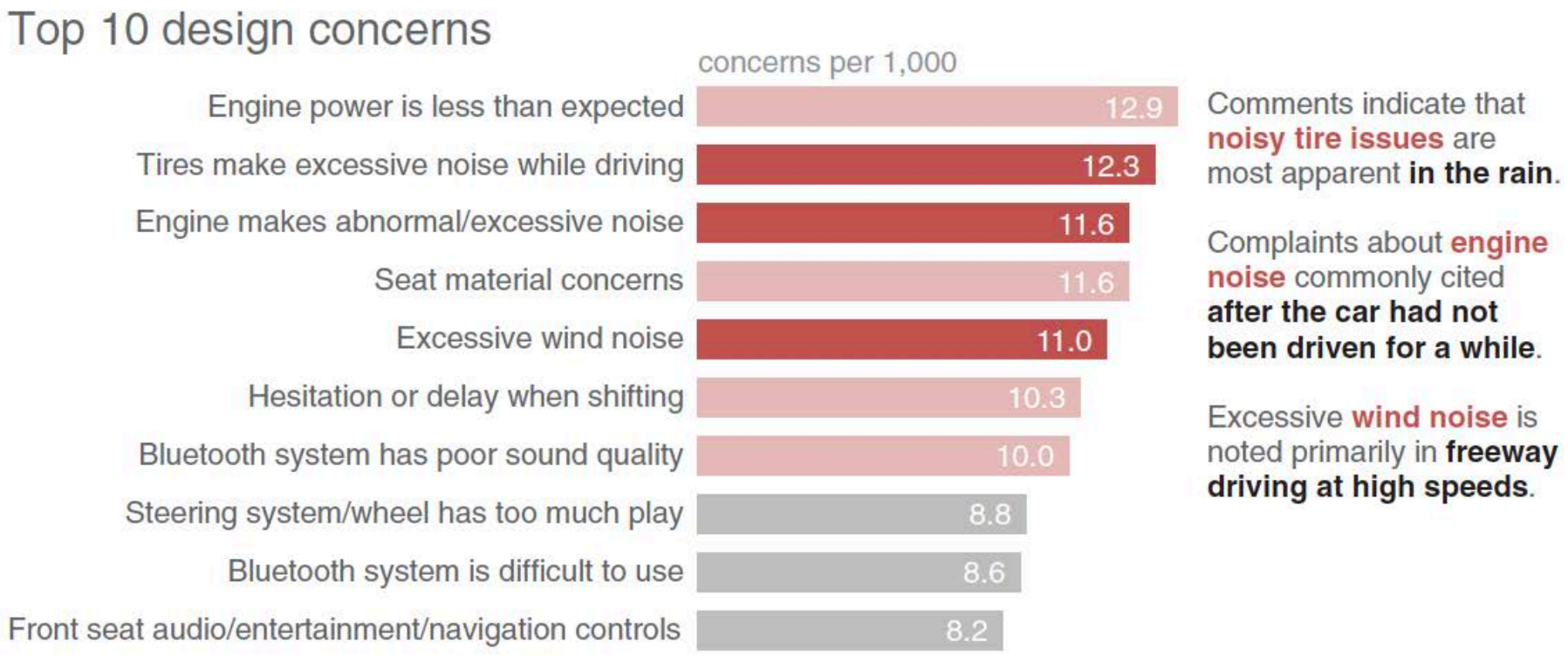
RANK 1 2 3 4 5+

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

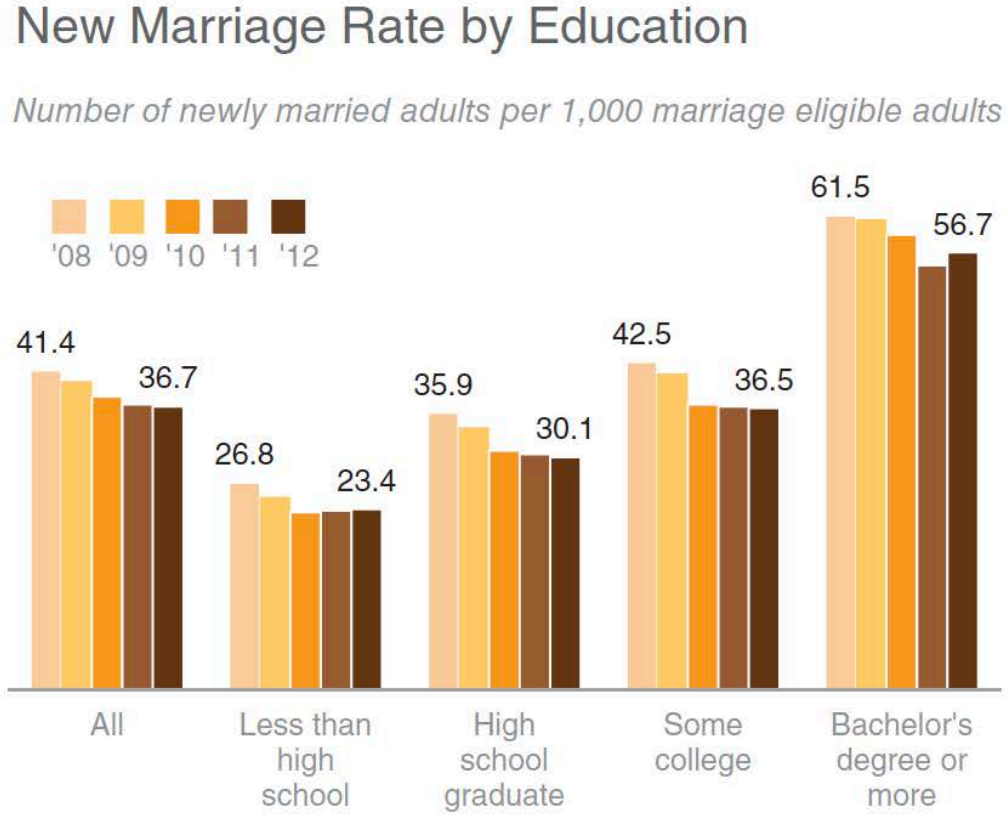
Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

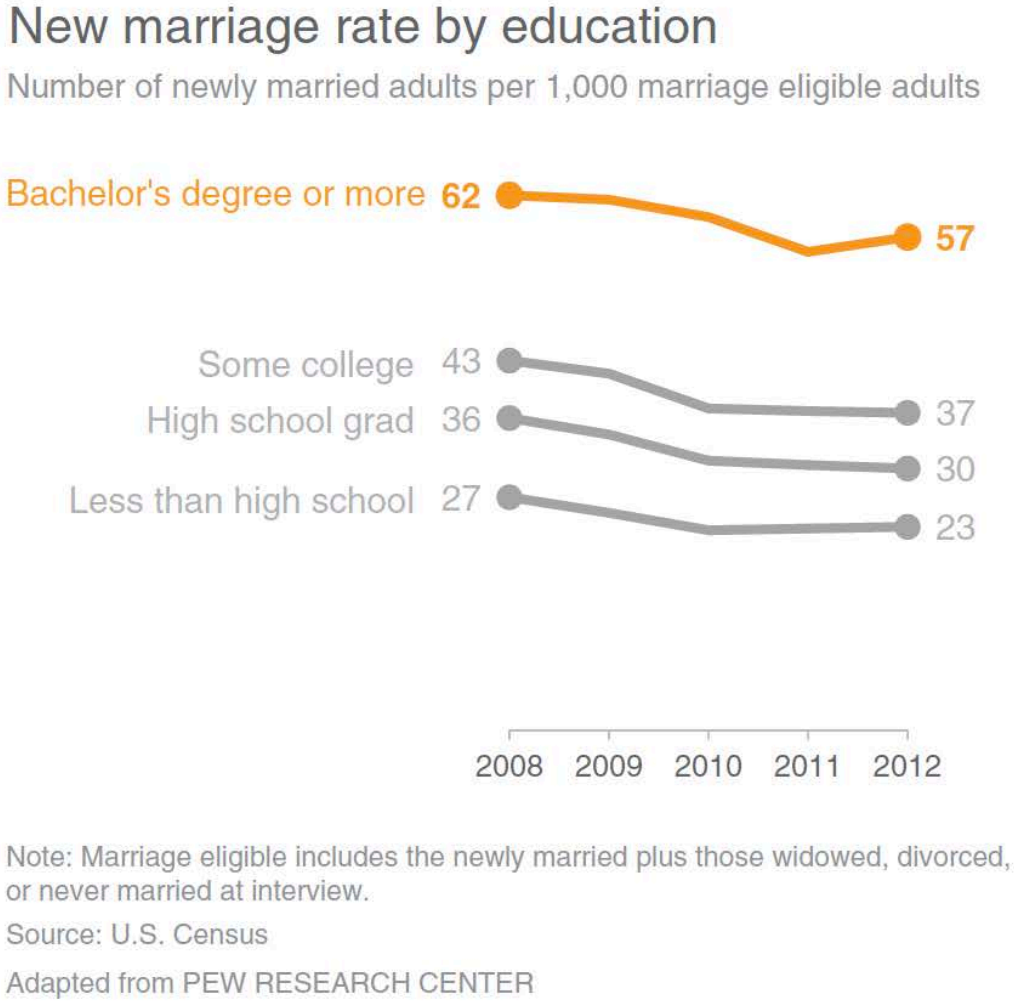
Case 10: Highlight What is Important (to Your Story!)



Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

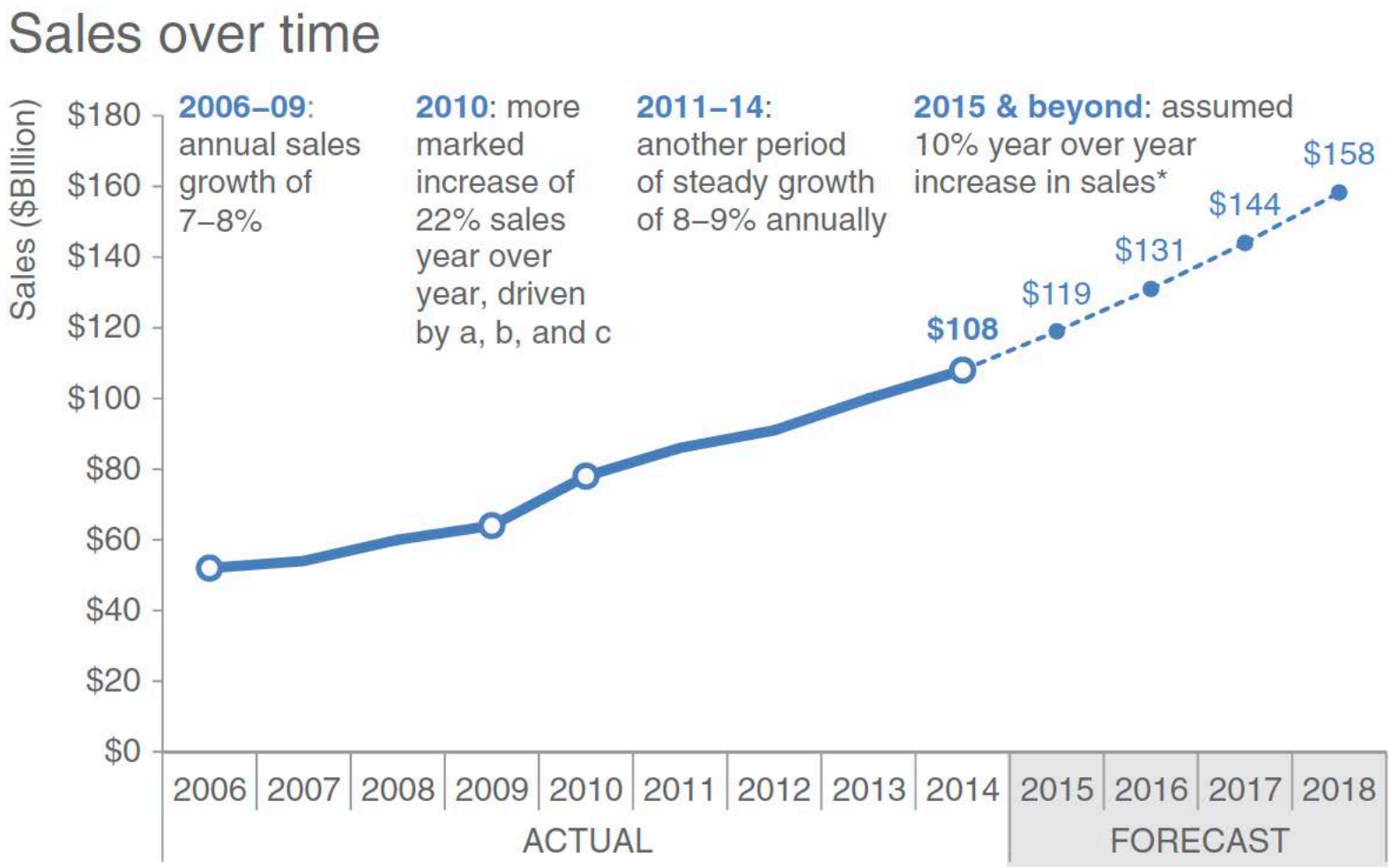
Case 11: Simplify and Focus





Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 12: Visualization in Forecasting



Data source: Sales Dashboard; annual figures are as of 12/31 of the given year.

*Use this footnote to explain what is driving the 10% annual growth forecast assumption.

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Case 13: Super-Categories

Source: Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals.

Super-Categories

- organize your data
- can be a hierarchical organization (but must not be)
- provide a construct for interpretation
- graphs and tables

Optional: A perspective on AI (3 years ago)

In the Age of AI

A documentary exploring how artificial intelligence is changing life as we know it — from jobs to privacy to a growing rivalry between the U.S. and China.

www.youtube.com/watch?v=5dZ_IvDgevk

Team Assignment 1: Data and AI along the Consumer Journey

- You've been assigned to teams on Canvas: Your team comprises Computer Science and Business students
- You must work in your team to complete and present team assignments → One grade for the entire team
- Submit PowerPoint Presentation for this Team Assignment on Canvas by Noon of January 24th, 2023
- Present in class, discuss with class, update presentation and re-submit on Canvas by end of day January 27th, 2023

How does Data Science and AI affect typical consumers every day? This team assignments asks you to ponder the question as a team and chart out answers to be presented and discussed in class.

Each team will pick ONE consumer segment from **Claritas Identity Graph**. Examples are Kids & Cul-de-Sacs and Time Shifters

Your task is to chart out your segment's consumer journey in a typical week of their life regarding:

- Data they generate (how / with what?)
- Data that is collected from them (by whom?)
- What these data are currently used for (by ML and AI)?
- What these data could be used for (by ML and AI)?

The consumer journey you are to chart is not specific to the products or services of any particular firm. Think of it as your segment's daily/weekly routines in life such as getting up early in the morning, getting dressed, having breakfast, going to school, studying at school, interacting with friends, having lunch, going home, going to some sports activity, grabbing dinner, doing homework, shopping, hanging out with friends on the weekend, etc.

Be creative on how to systematically capture, present, and structure your findings!

Looking Ahead

Next Class: Tuesday, 01/24/23

In-Class Presentations of Team Assignment 1

Data and AI Along the Consumer Journey

- Submission before noon on Canvas:
  - PowerPoint Presentation
- All teams bring USB Stick with their presentation to class
  (must use class computer to present)
- 6 Minute Presentation
  + 6 Minute Mini-Workshop
- All Teams present
- All students must attend

DataCamp HW2 due 01/24 before midnight