

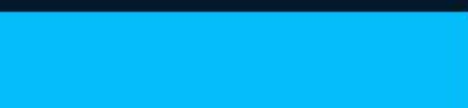
8 Rules for Better Data Storytelling



Our Mission



**Our mission is to democratize
data & AI skills for everyone**



Future-proof your skills with DataCamp

Best in class learning



Give your employees access to market-leading training with DataCamp Learn

Build work ready skills



Apply your skills in a risk-free online coding environment with DataCamp Datalab

Grow your data team



Certify your people with the best in-class data certifications using DataCamp Certifications

Trusted by more than 14 million learners and 5,000 data-driven companies

Google

Microsoft

ebay

HSBC

COLGATE-PALMOLIVE

T-Mobile

PayPal

Uber

Deloitte.

CREDIT SUISSE

Mercedes-Benz

BNP PARIBAS



About me



Adel Nehme

VP of Media

DataCamp



[adelnehme](#)

- Graduate in Economics from the American University of Beirut
- MSc in Business Analytics & Data Science from ESSEC Business School & CentraleSupélec
- Data Science Educator & Evangelist @ DataCamp
- 🎤 Host of the DataFramed Podcast 🎤



Agenda



- 1 Data storytelling: The last mile of analytics
- 2 8 rules for better data storytelling
 - *4 rules for better data visualizations*
 - *4 rules for better narrative*
- 3 Become a better data storyteller





1

Data Storytelling

The last mile of analytics



What is data storytelling?



"Having all the information in the world at our fingertips doesn't make it easier to communicate: it makes it harder. The more information you're dealing with, the more difficult it is to filter" — **Cole Nussbaumer Knaflic, Author of *Storytelling with Data: A Data Visualization Guide for Business Professionals***

"Data storytelling enables data teams to wield the power to frame arguments and persuade with data responsibly and deliberately " — **Andy Cotgreave, Technical Evangelist at Tableau**

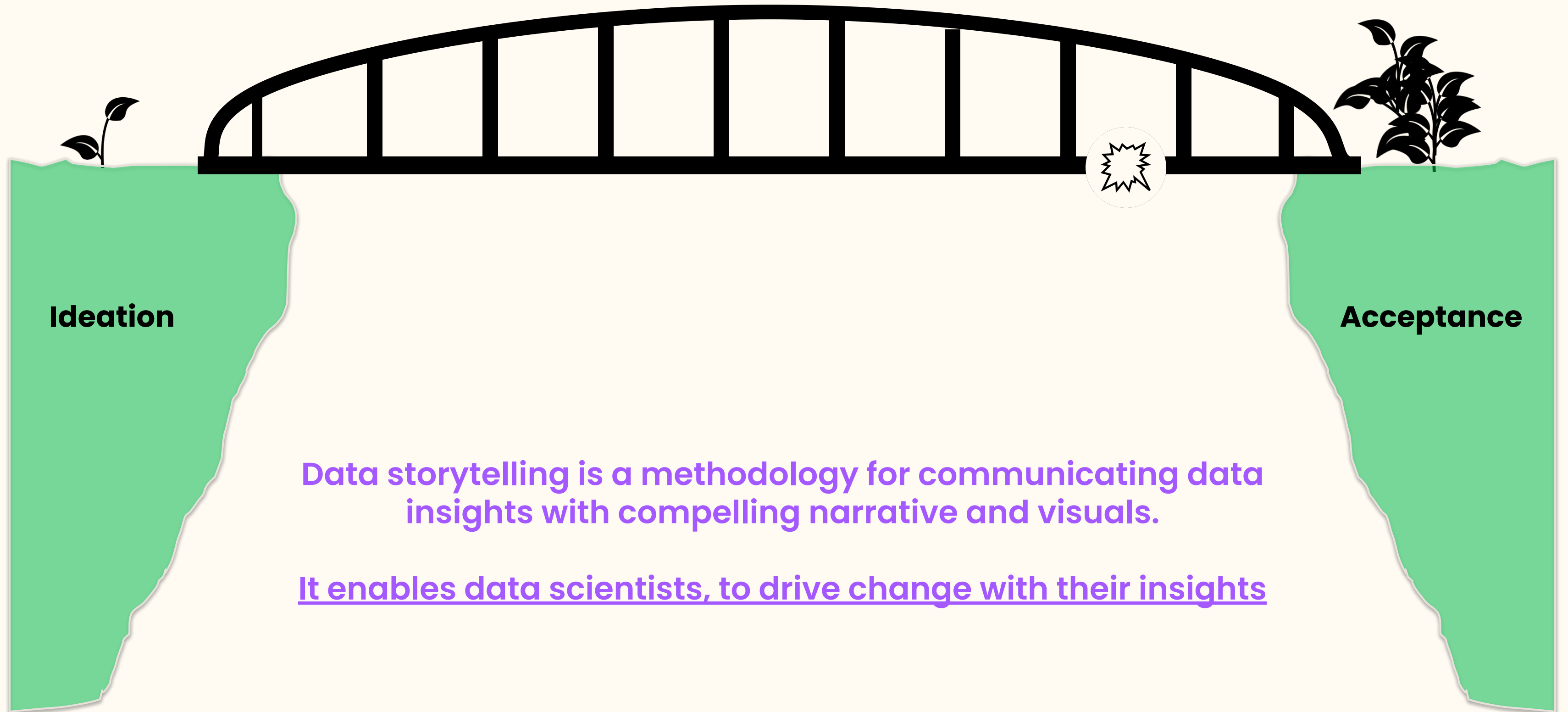


"To be impactful with data science, organizations need data scientists with stories " — **Gert de Geyter, Machine Learning Lead at Deloitte**

"Data stories are powerful vehicles for sharing data insights to influence and drive change within an organization " — **Brent Dykes, Author of *Effective Data Storytelling***



What is data storytelling?



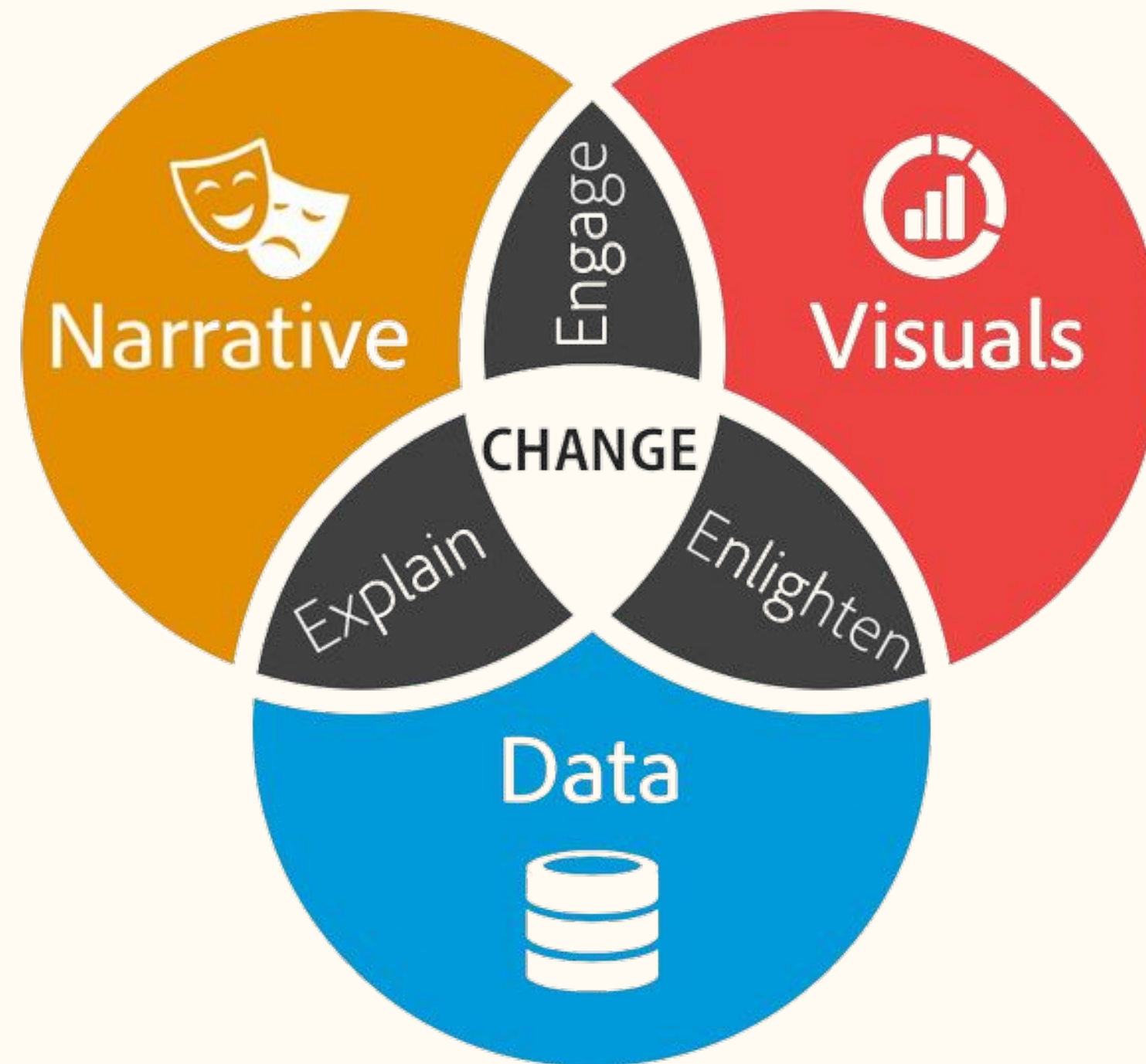
Ideation

Acceptance

Data storytelling is a methodology for communicating data insights with compelling narrative and visuals.

It enables data scientists, to drive change with their insights

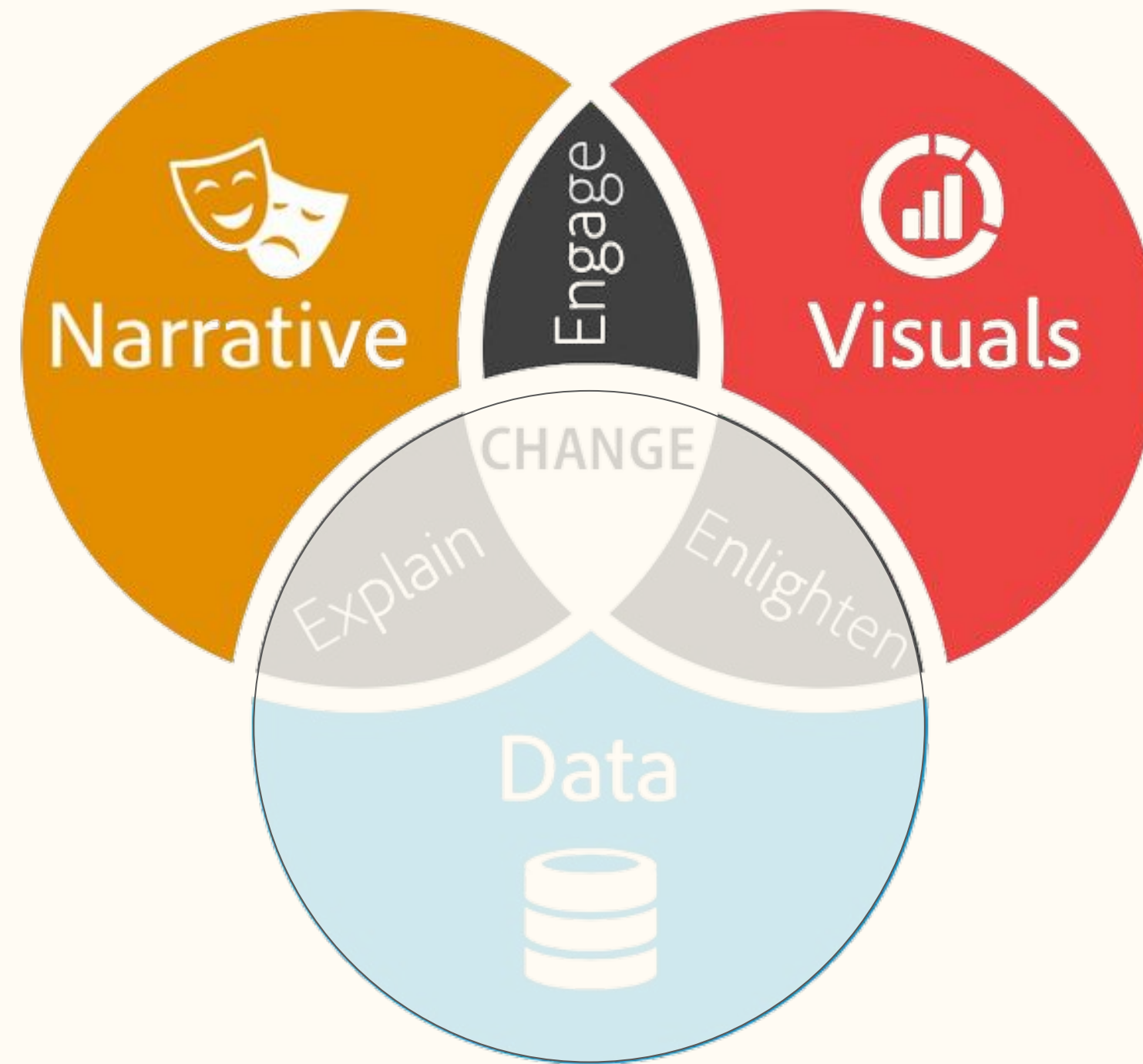
What is data storytelling?



[Source: Effective Data Storytelling: How to Drive Change with Data, Narrative, and Visuals](#)



What is data storytelling?



[Source: Effective Data Storytelling: How to Drive Change with Data, Narrative, and Visuals](#)





2

8 rules for better data storytelling

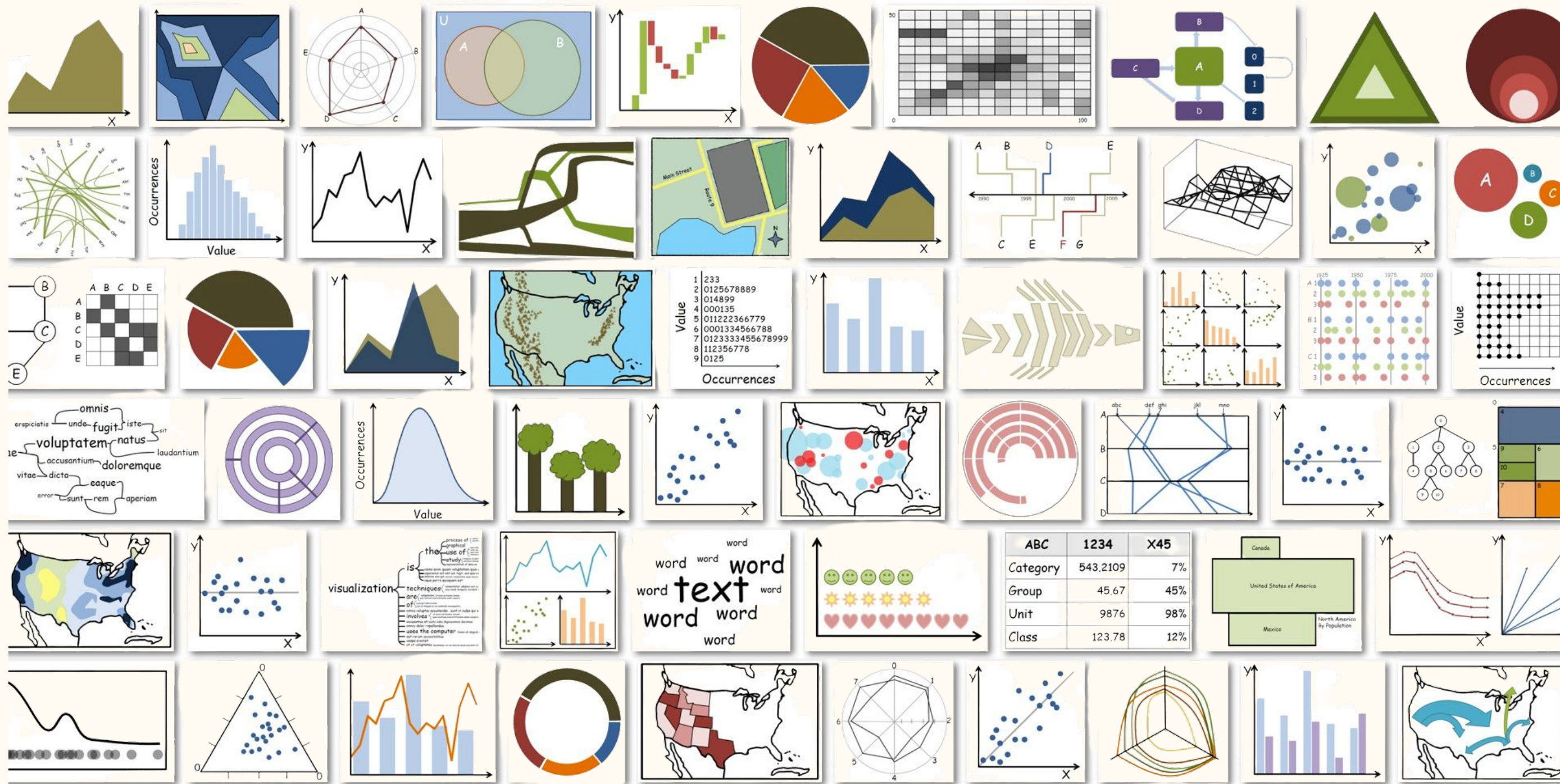
Rules for better data visualizations



Rule #1

*Choose the best visualization
for your story*

So many ways to cut a cake



Always work back from your data

COMMONLY USED DATA

PROBLEM DESCRIPTION

MOST USEFUL VISUALIZATION



Always work back from your data

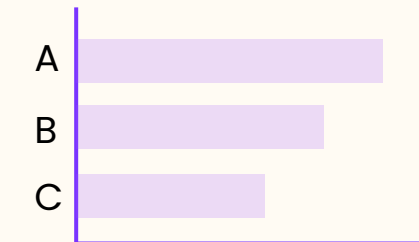
COMMONLY USED DATA

- Banking products and associated customers
 - Branch cost data broken down into different verticals
 - Visualizing different customer segments
-
-
-
-
-

PROBLEM DESCRIPTION

Showing comparisons of different categories

MOST USEFUL VISUALIZATION



BAR CHARTS



Common uses of bar charts

Number of customers

Visualizing number of customers by loan product



Business Loans

Loan Against Property

Commercial Vehicle Financing

Construction Equipment Loan

Farm Equipment Loan

Loan Products



Always work back from your data

COMMONLY USED DATA

- Banking products and associated customers
 - Branch cost data broken down into different verticals
 - Visualizing different customer segments
-

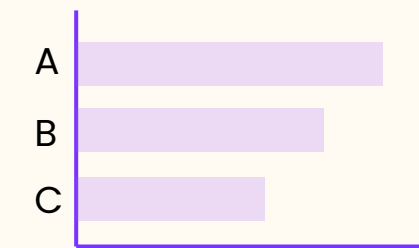
- Stock price change over time
 - Number of app users over time
 - Number of customer support tickets over time
-

PROBLEM DESCRIPTION

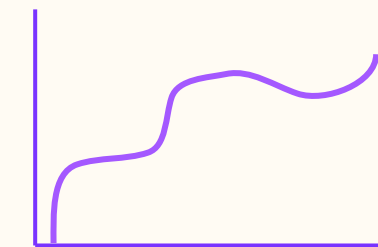
Showing comparisons of different categories

Showing changes of a variable over time

MOST USEFUL VISUALIZATION



BAR CHARTS



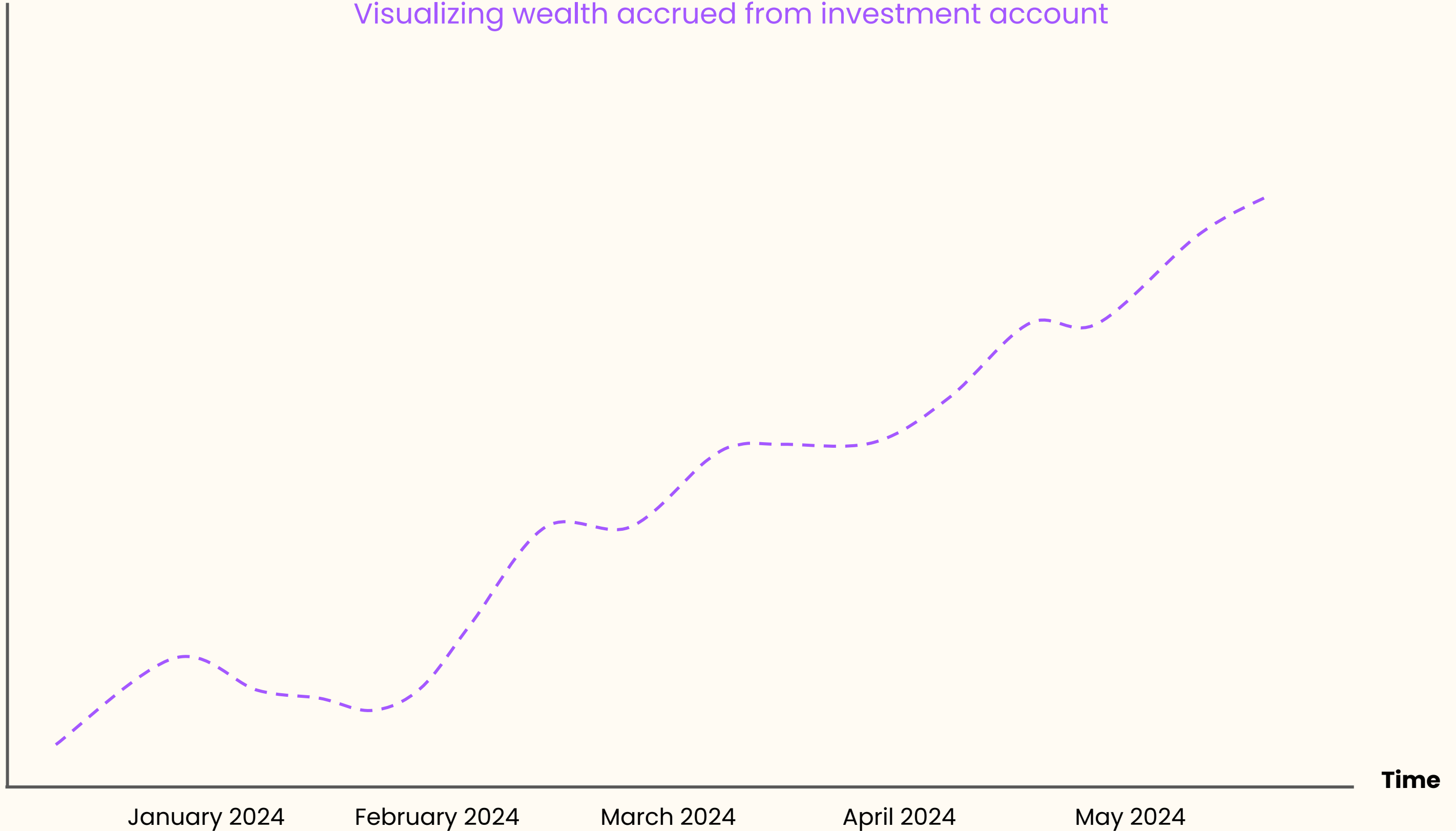
LINE CHARTS



Common uses of line charts

Visualizing wealth accrued from investment account

**Investment
Account Gains
over time**



Always work back from your data

COMMONLY USED DATA

- Banking products and associated customers
 - Branch cost data broken down into different verticals
 - Visualizing different customer segments
-

- Stock price change over time
 - Number of banking app users over time
 - Number of customer support tickets over time
-

- The distribution of savings for customers in savings accounts
 - App opens for all customers in a given time period
-

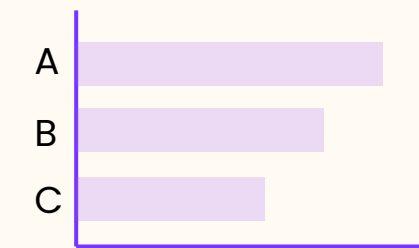
PROBLEM DESCRIPTION

Showing comparisons of different categories

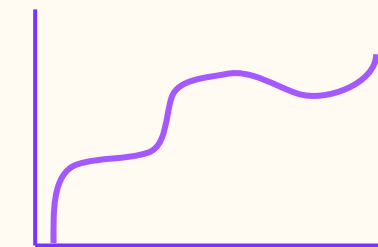
Showing changes of a variable over time

Show the distribution of a variable over time

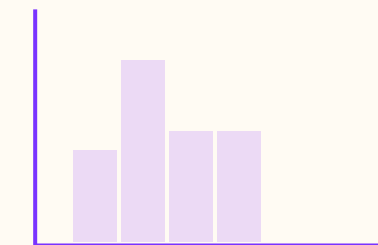
MOST USEFUL VISUALIZATION



BAR CHARTS



LINE CHARTS



HISTOGRAMS

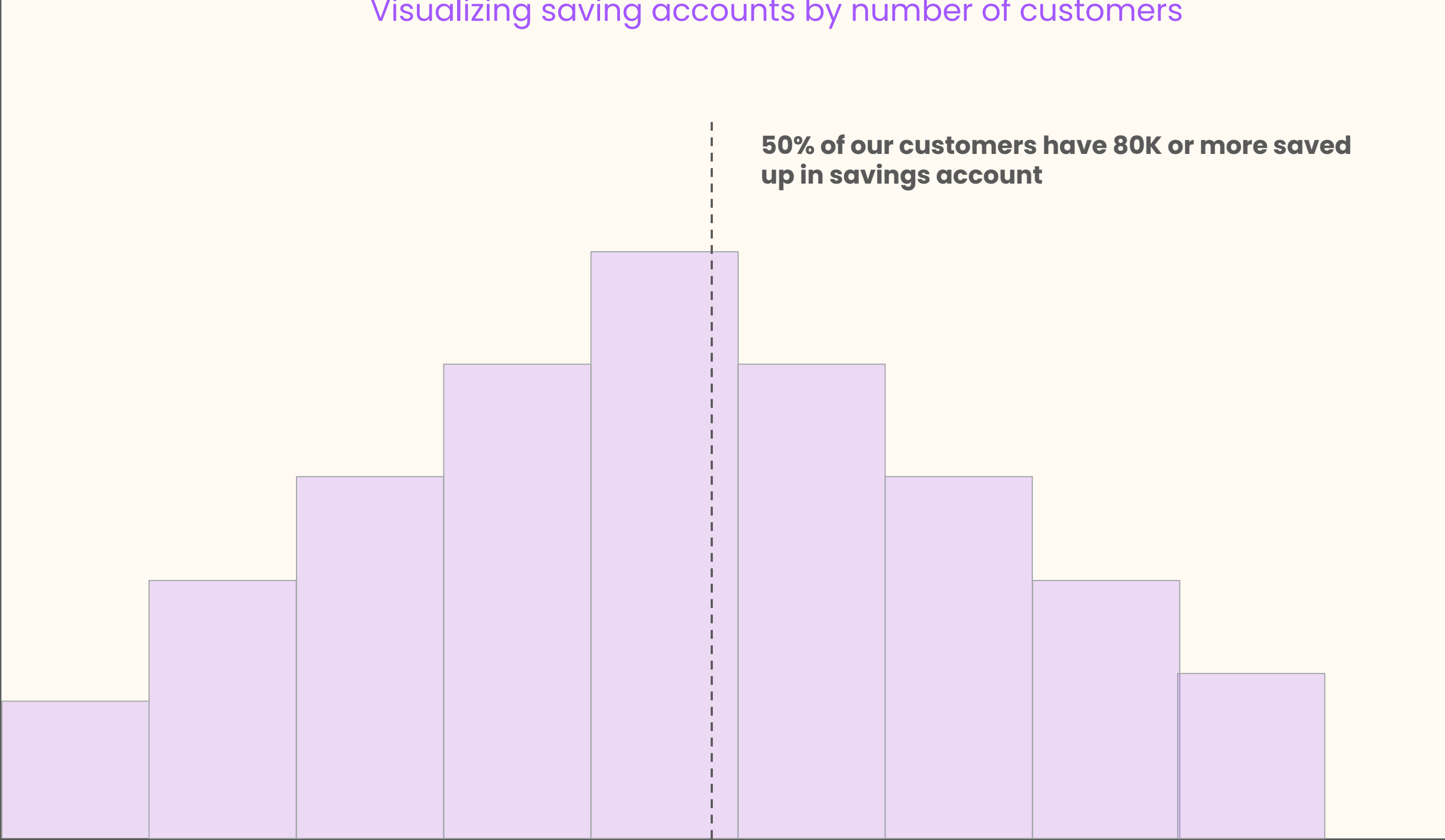


Common uses of histograms

Number of customers

Visualizing saving accounts by number of customers

50% of our customers have 80K or more saved up in savings account



Savings account \$



Always work back from your data

COMMONLY USED DATA

- Banking products and associated customers
 - Branch cost data broken down into different verticals
 - Visualizing different customer segments
-

- Stock price change over time
 - Number of banking app users over time
 - Number of customer support tickets over time
-

- The distribution of savings for customers in savings accounts
 - Banking app opens for all customers in a given time period
-

- The relationship between historical credit scores and number of loans taken
 - The relationship between customer lifetime value and number of products purchased
-

PROBLEM DESCRIPTION

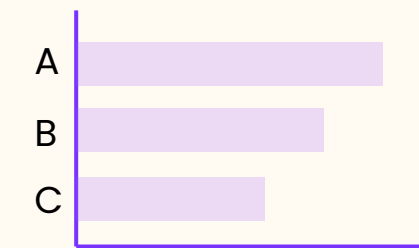
Showing comparisons of different categories

Showing changes of a variable over time

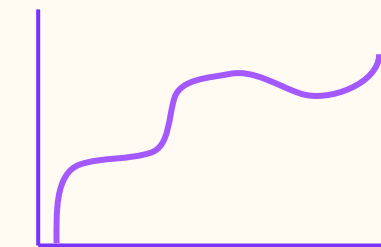
Show the distribution of a variable over time

Show the relationship between two variables

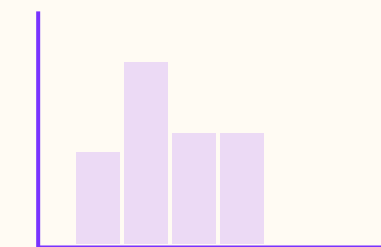
MOST USEFUL VISUALIZATION



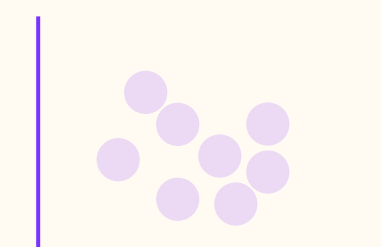
BAR CHARTS



LINE CHARTS



HISTOGRAMS



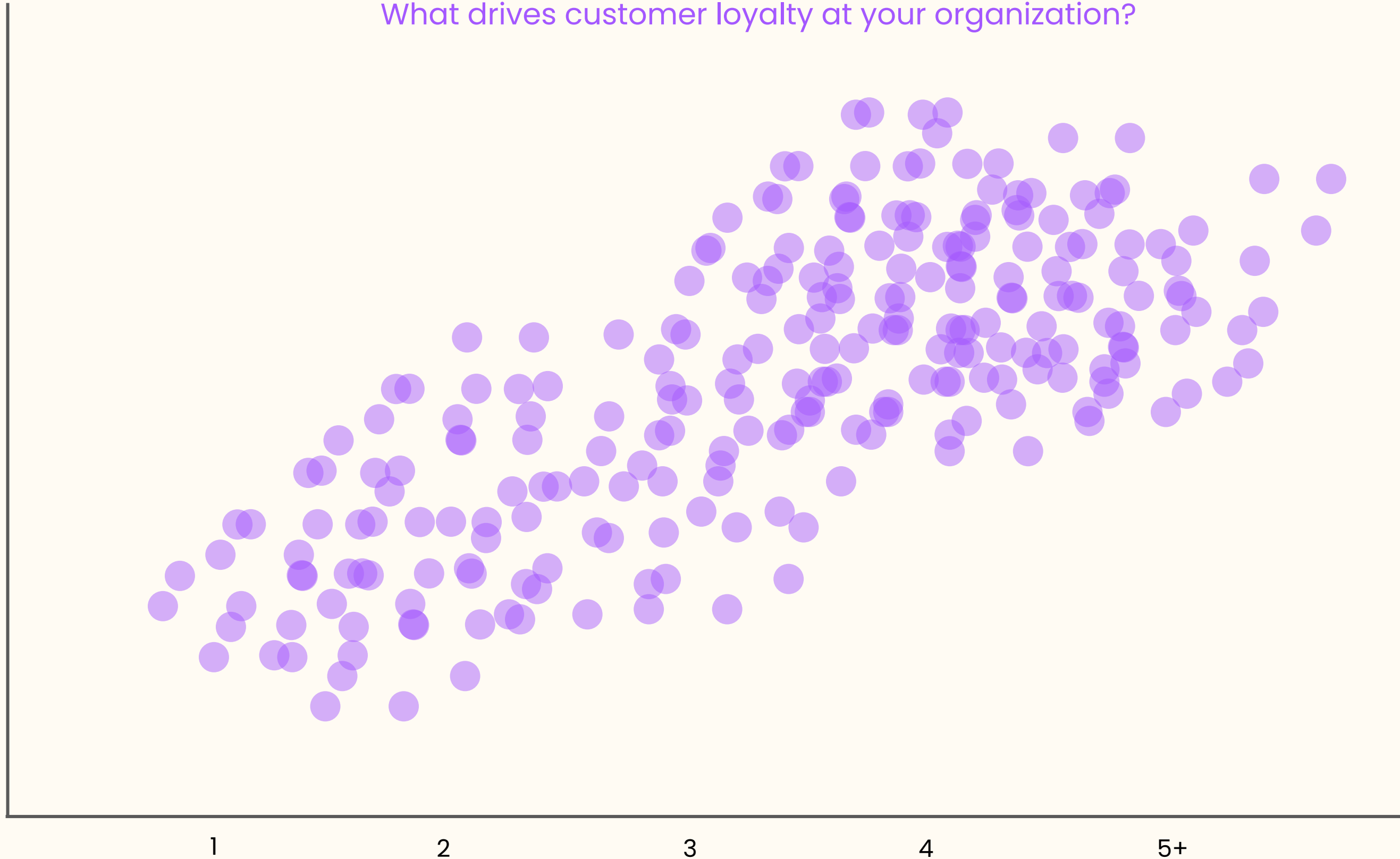
SCATTER PLOTS



Common uses of scatter plots

**Customer
lifetime value**

What drives customer loyalty at your organization?



**Products
purchased**



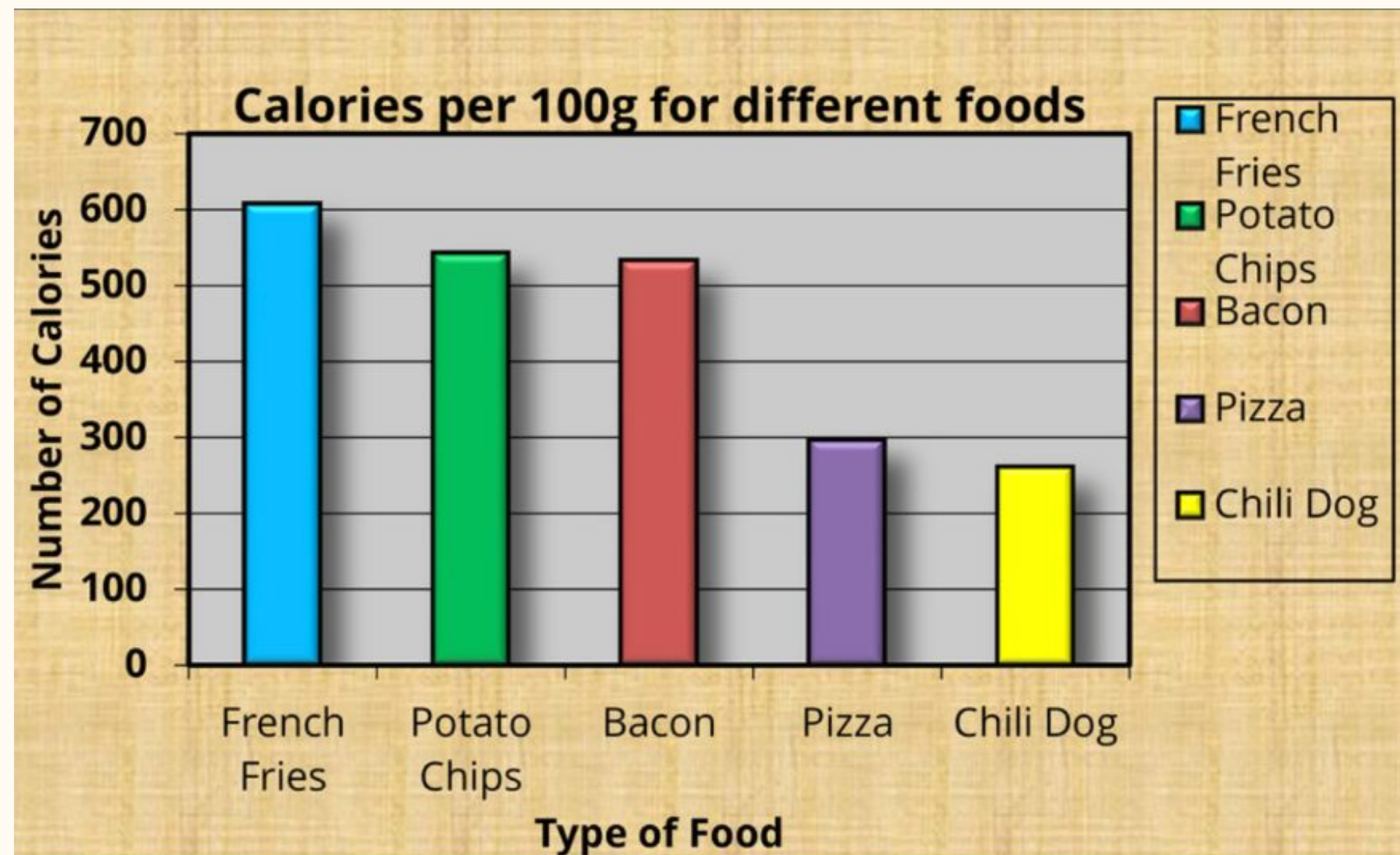


Rule #2

*Keep visualizations minimal and
avoid clutter*

The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.



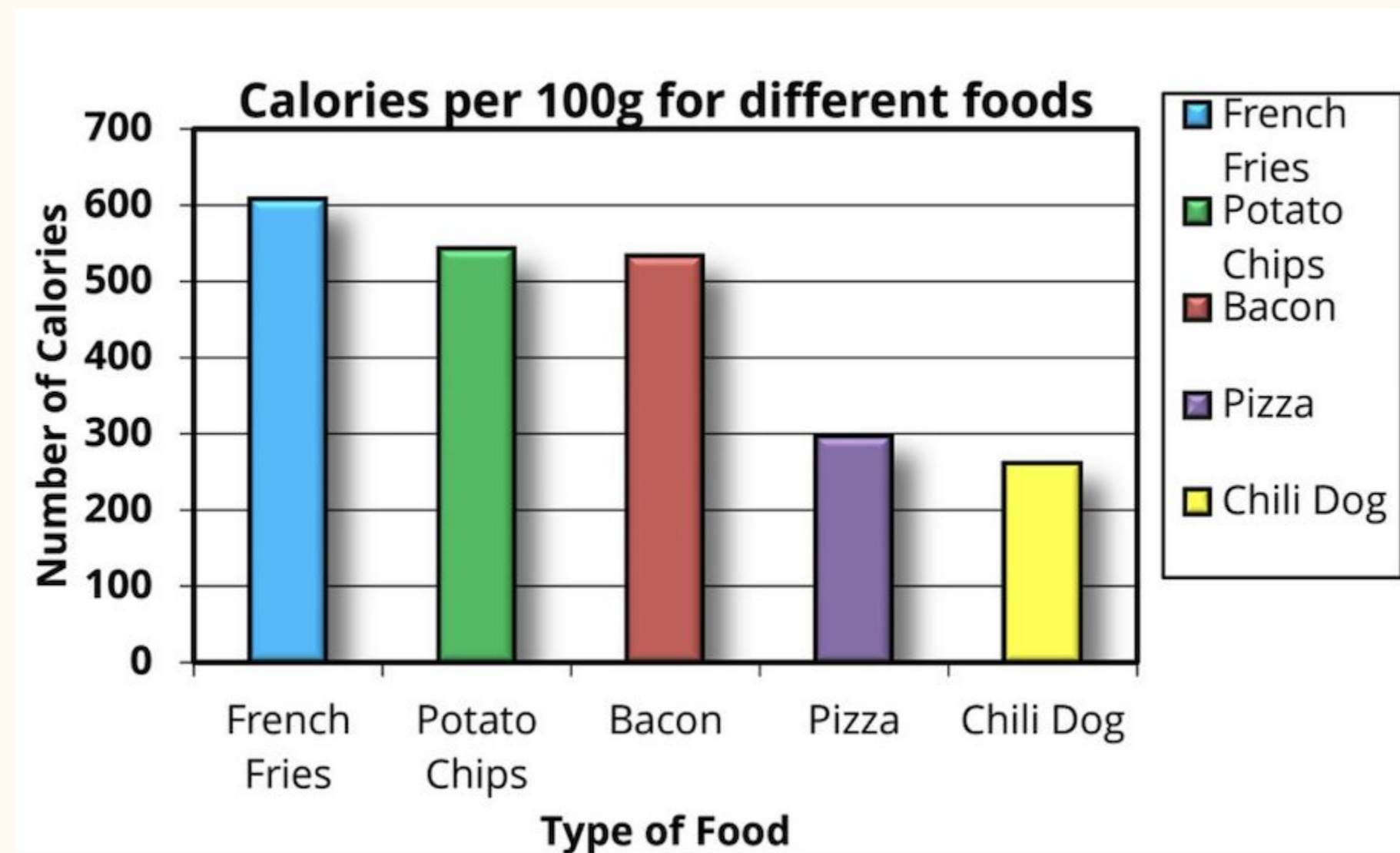
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #1 – Remove backgrounds



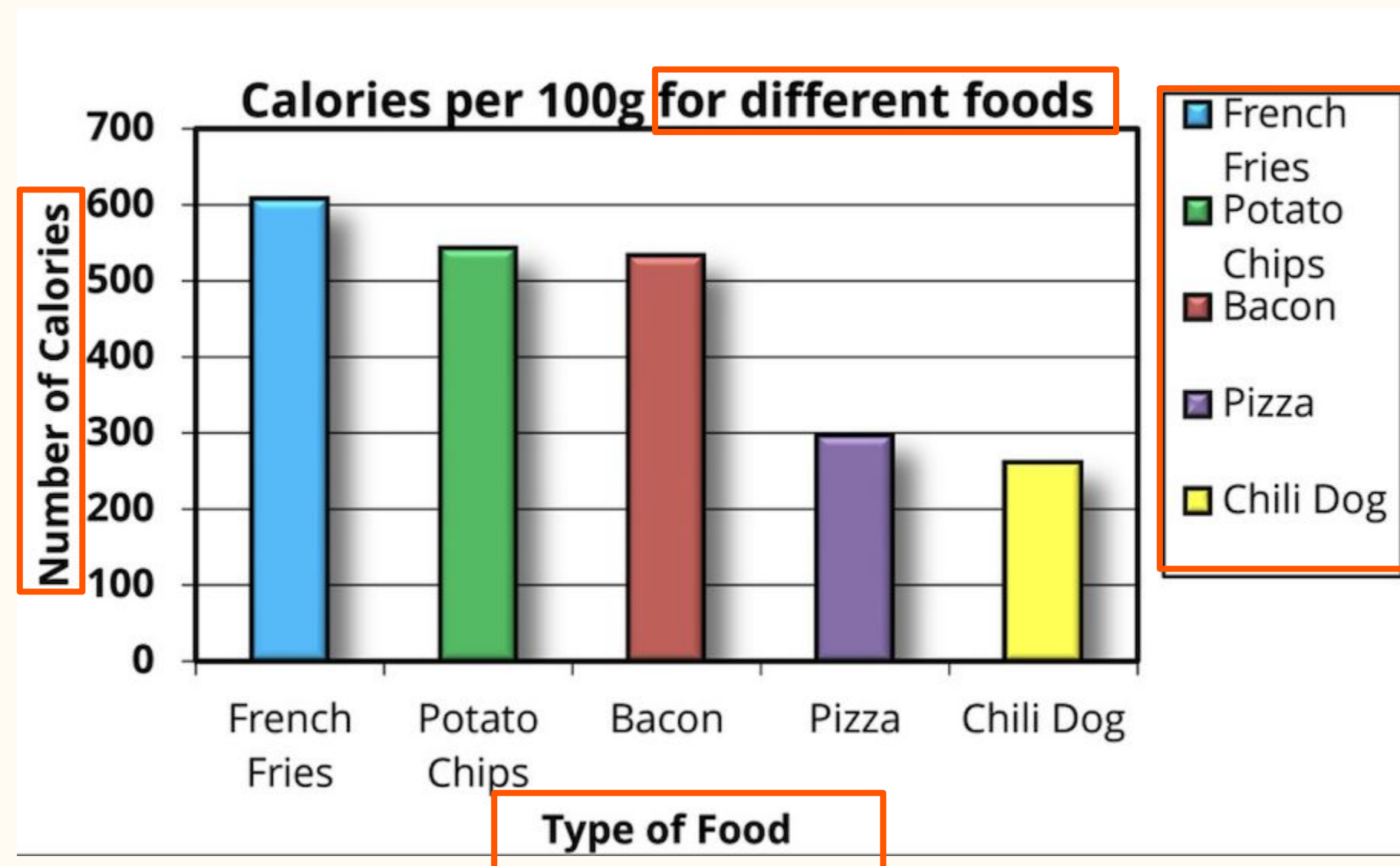
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #2 – Remove redundant labels



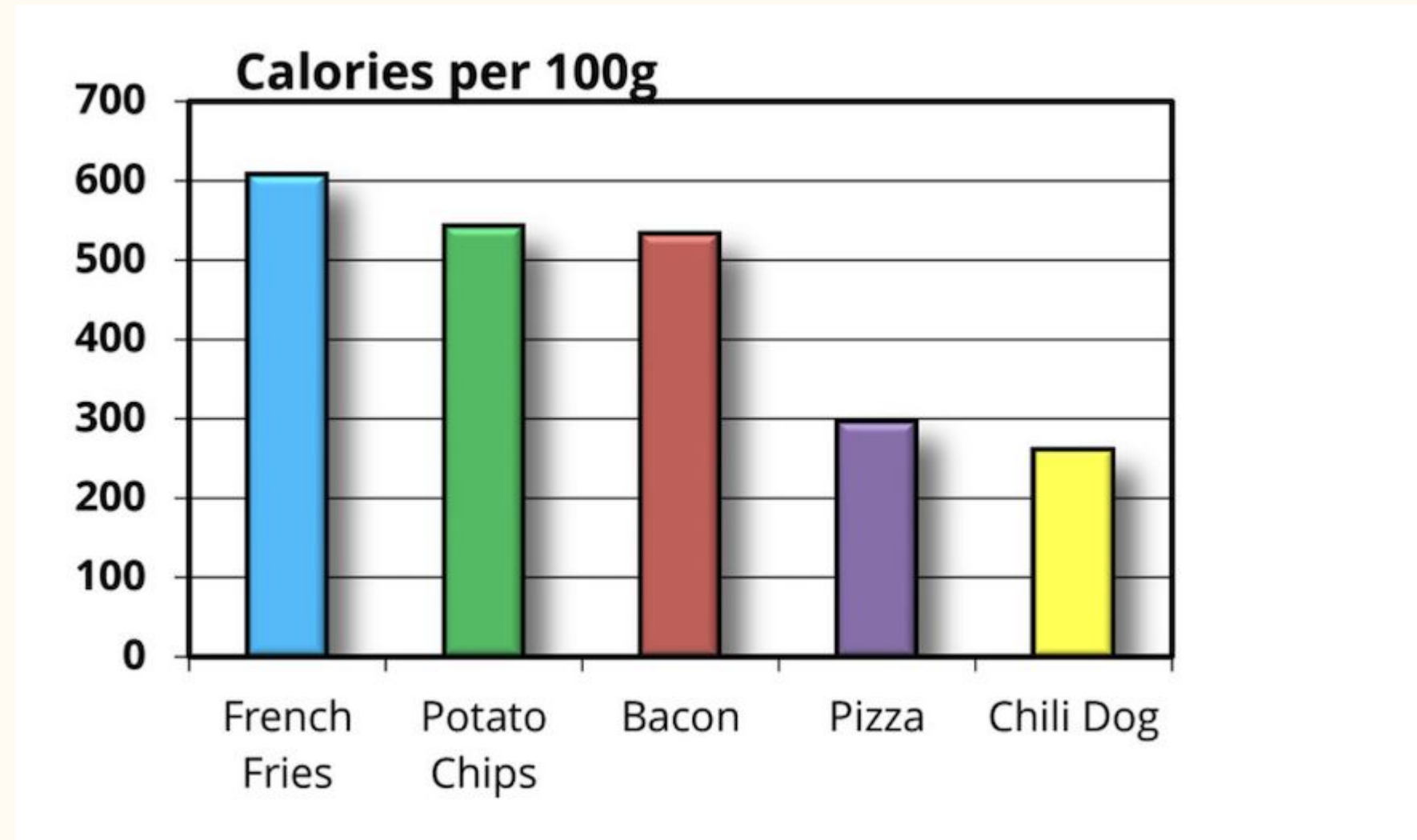
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #2 – Remove redundant labels



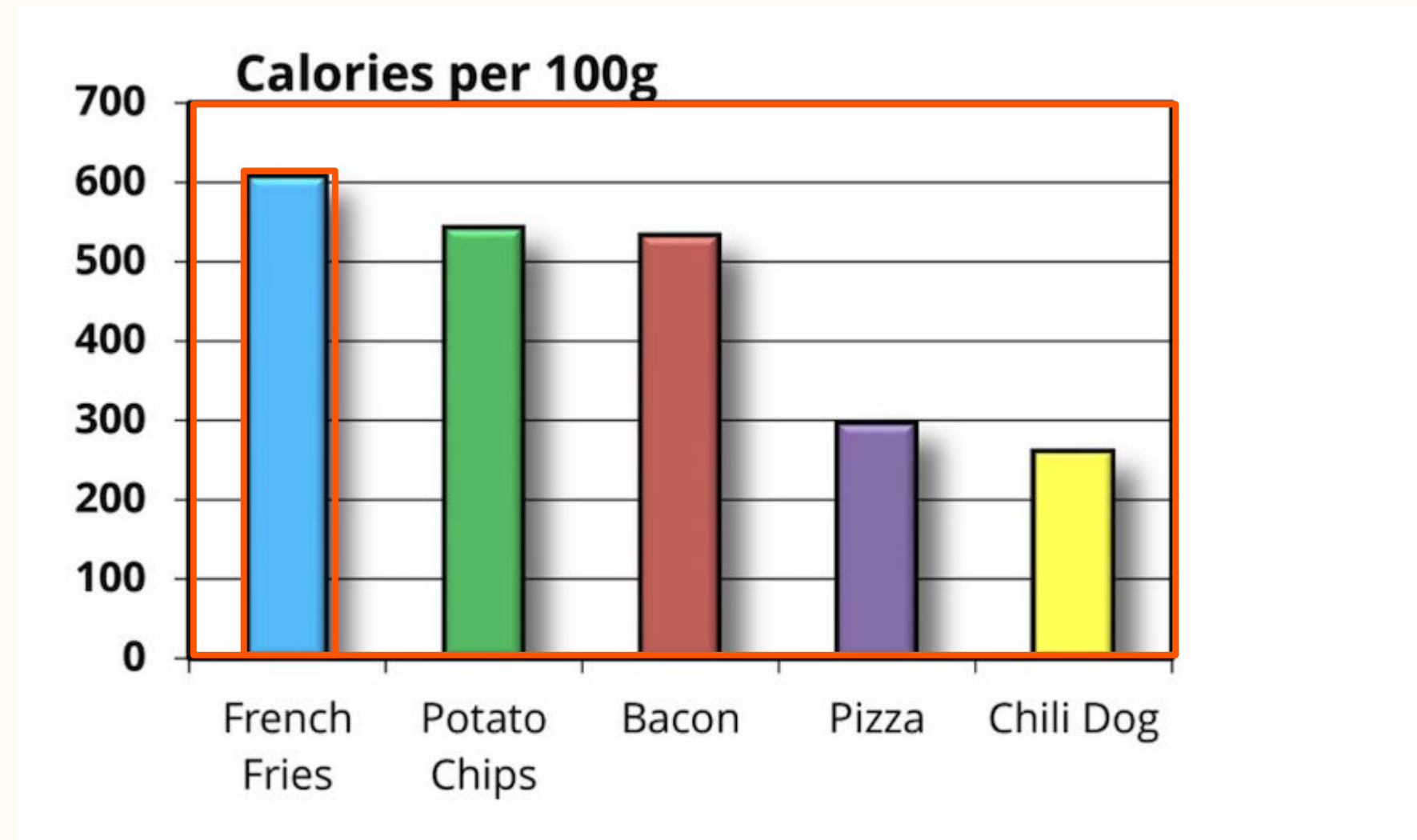
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #3 — Remove borders



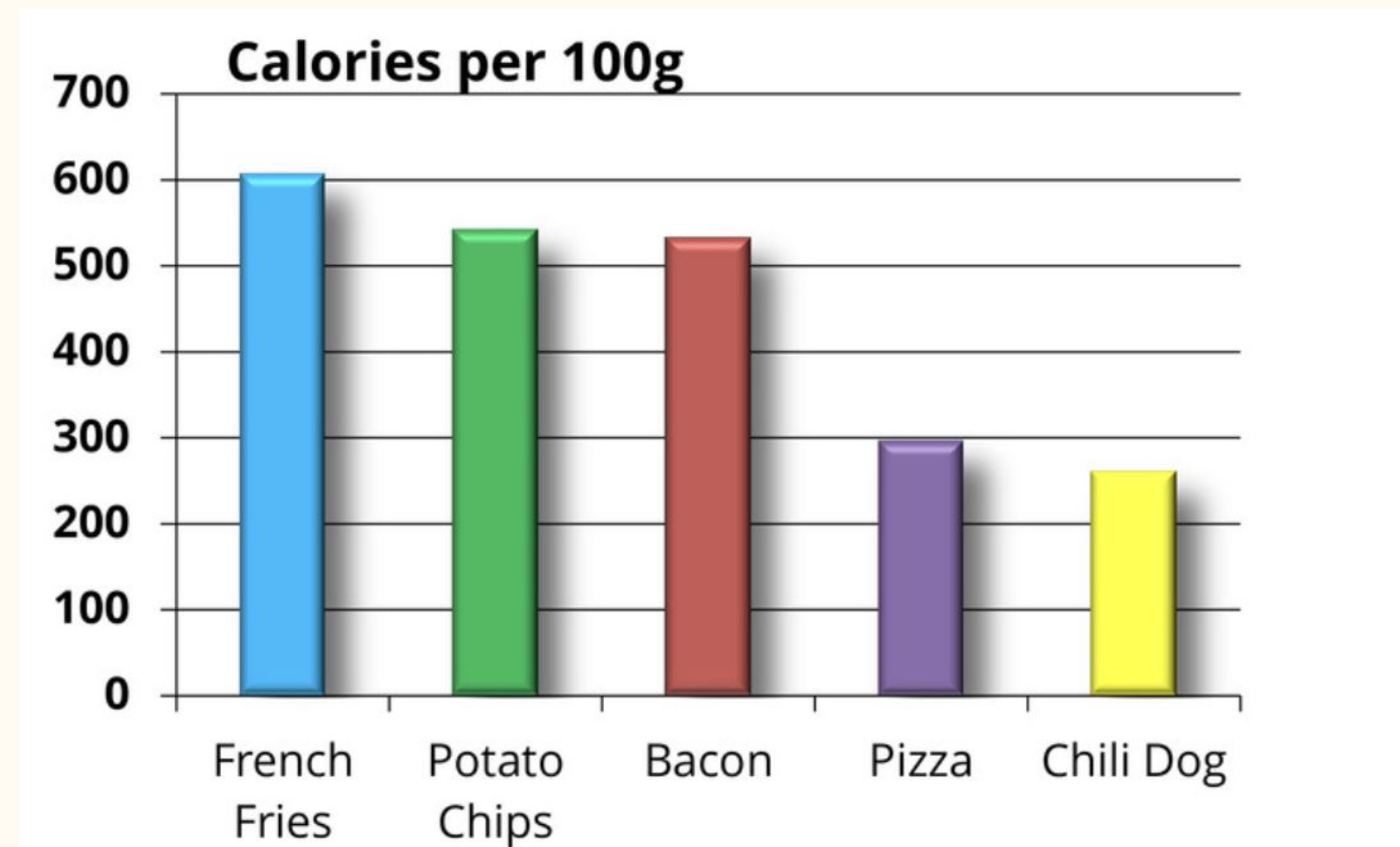
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #3 — Remove borders



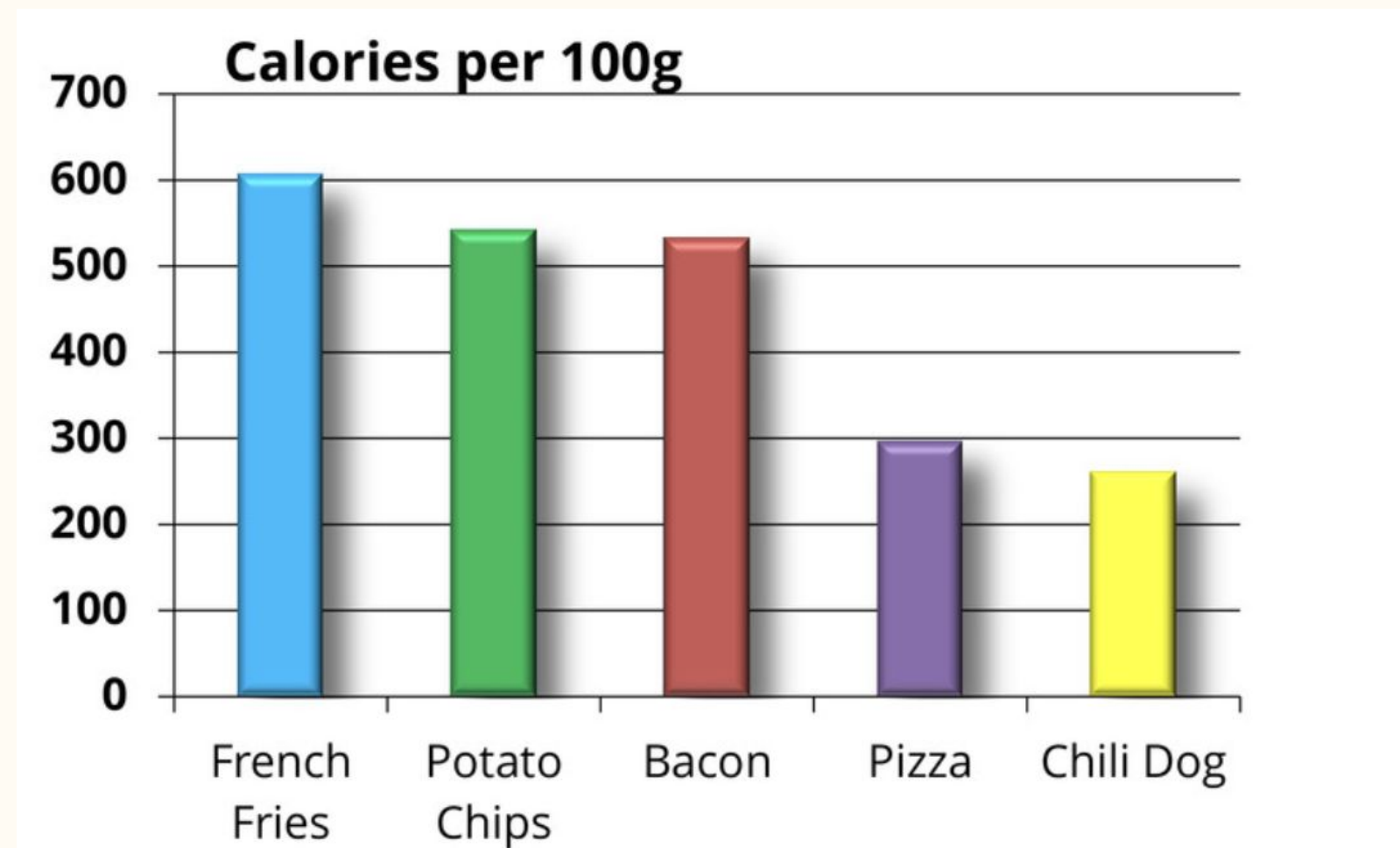
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #4 – Reduce colors when it doesn't matter



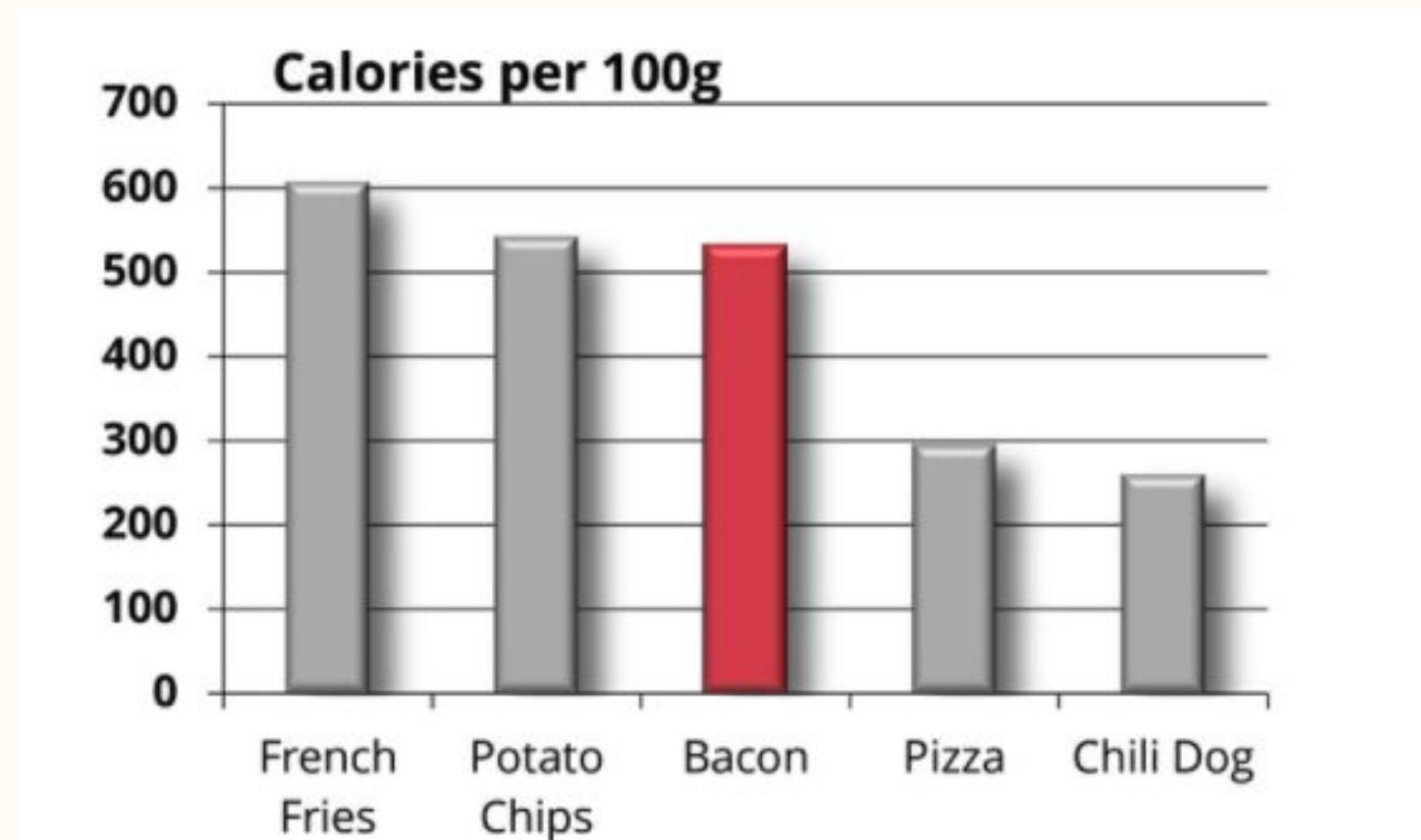
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #4 – Reduce colors when it doesn't matter



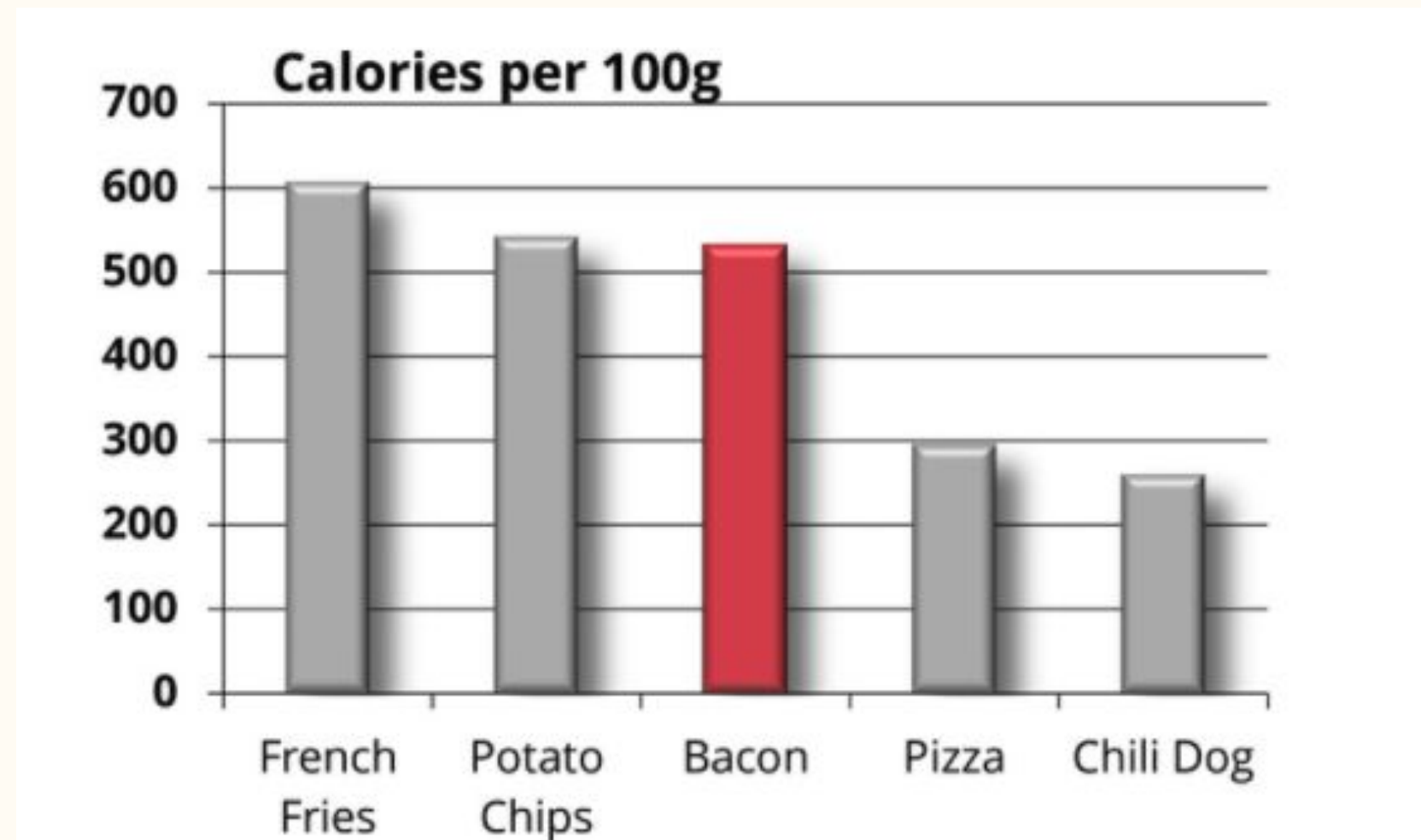
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #5 – Remove needless effects



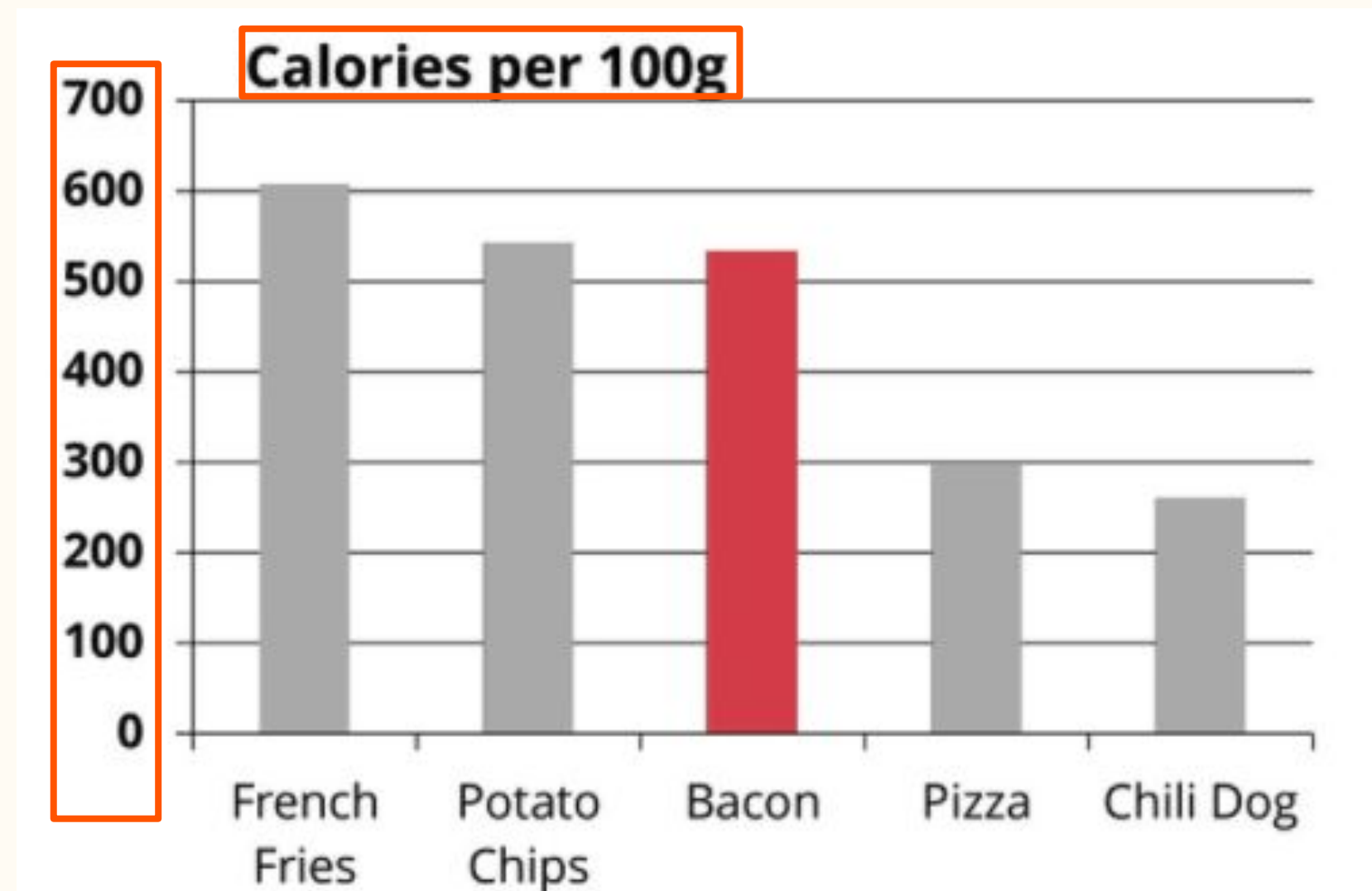
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #5 – Remove needless effects



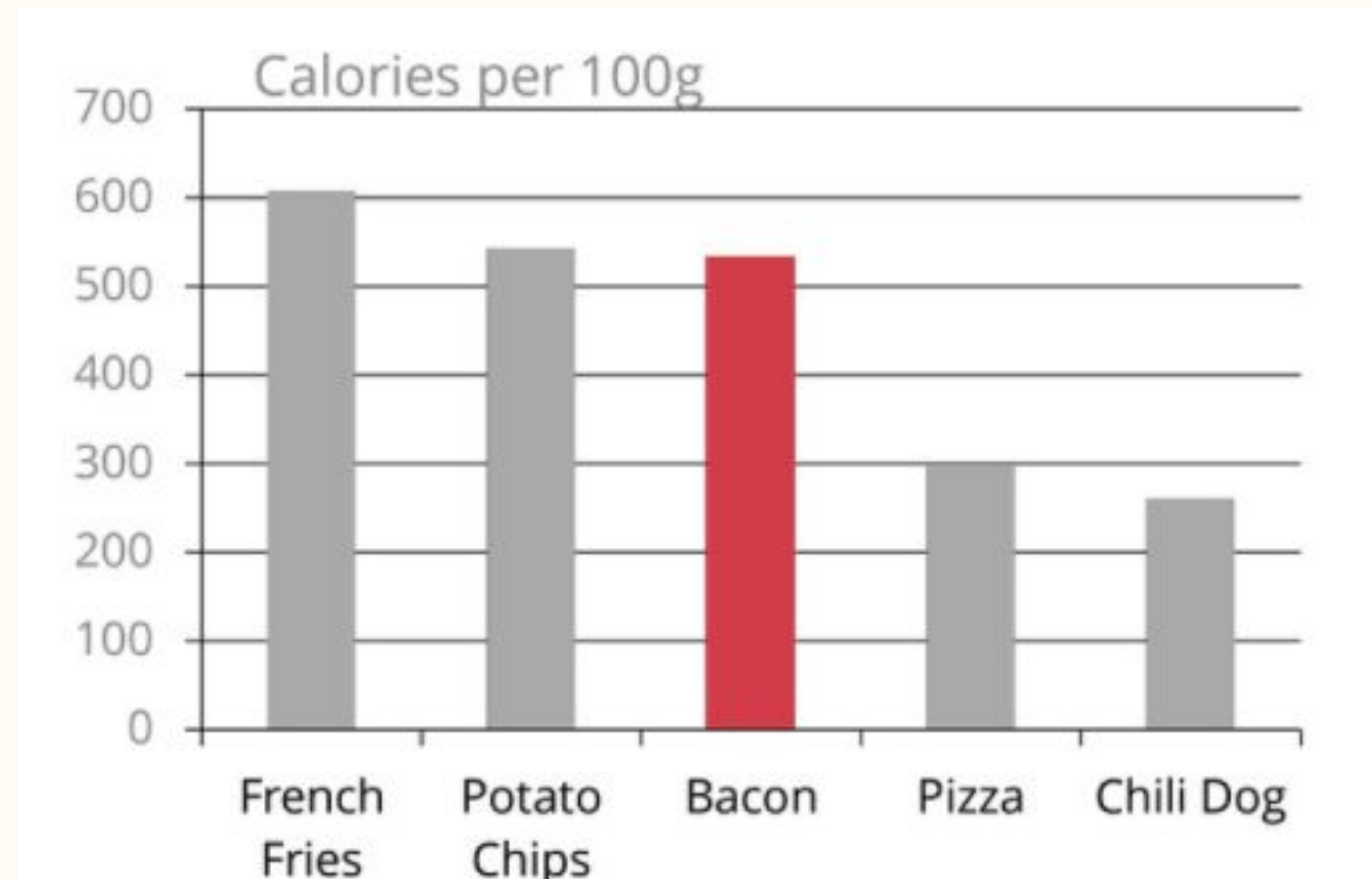
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #5 – Remove needless effects



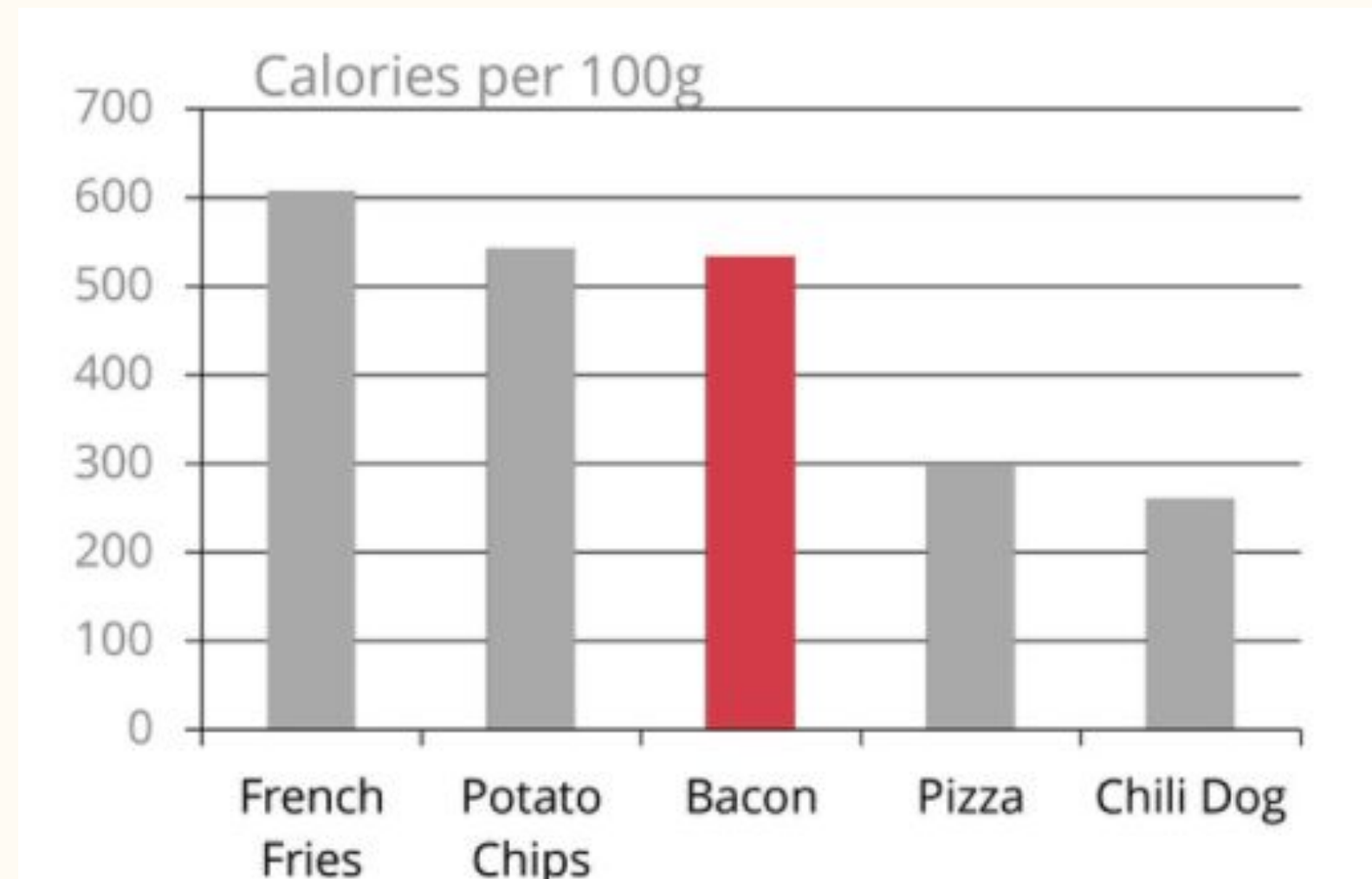
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #6 – Remove axis lines when not needed



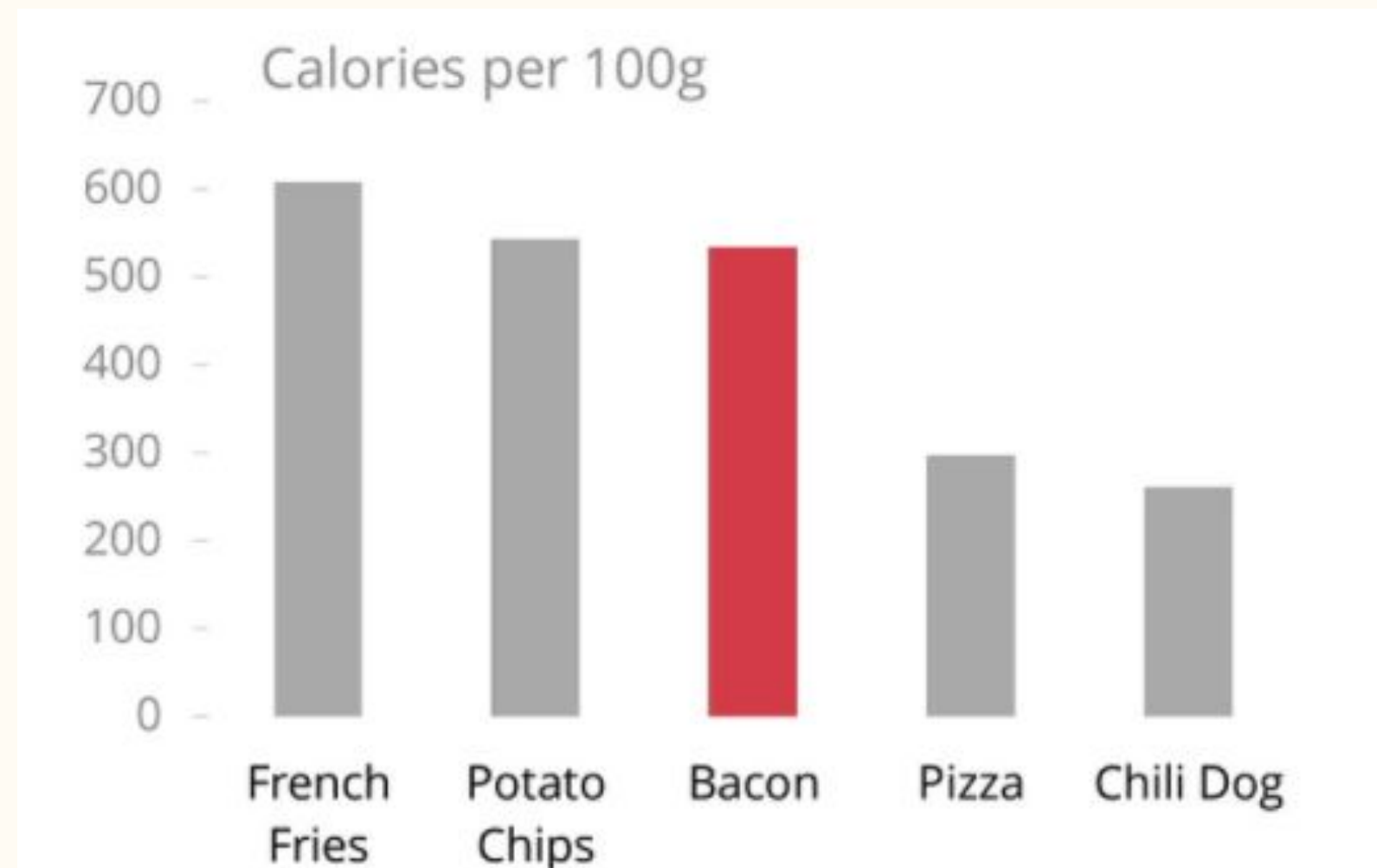
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #6 – Remove axis lines when not needed



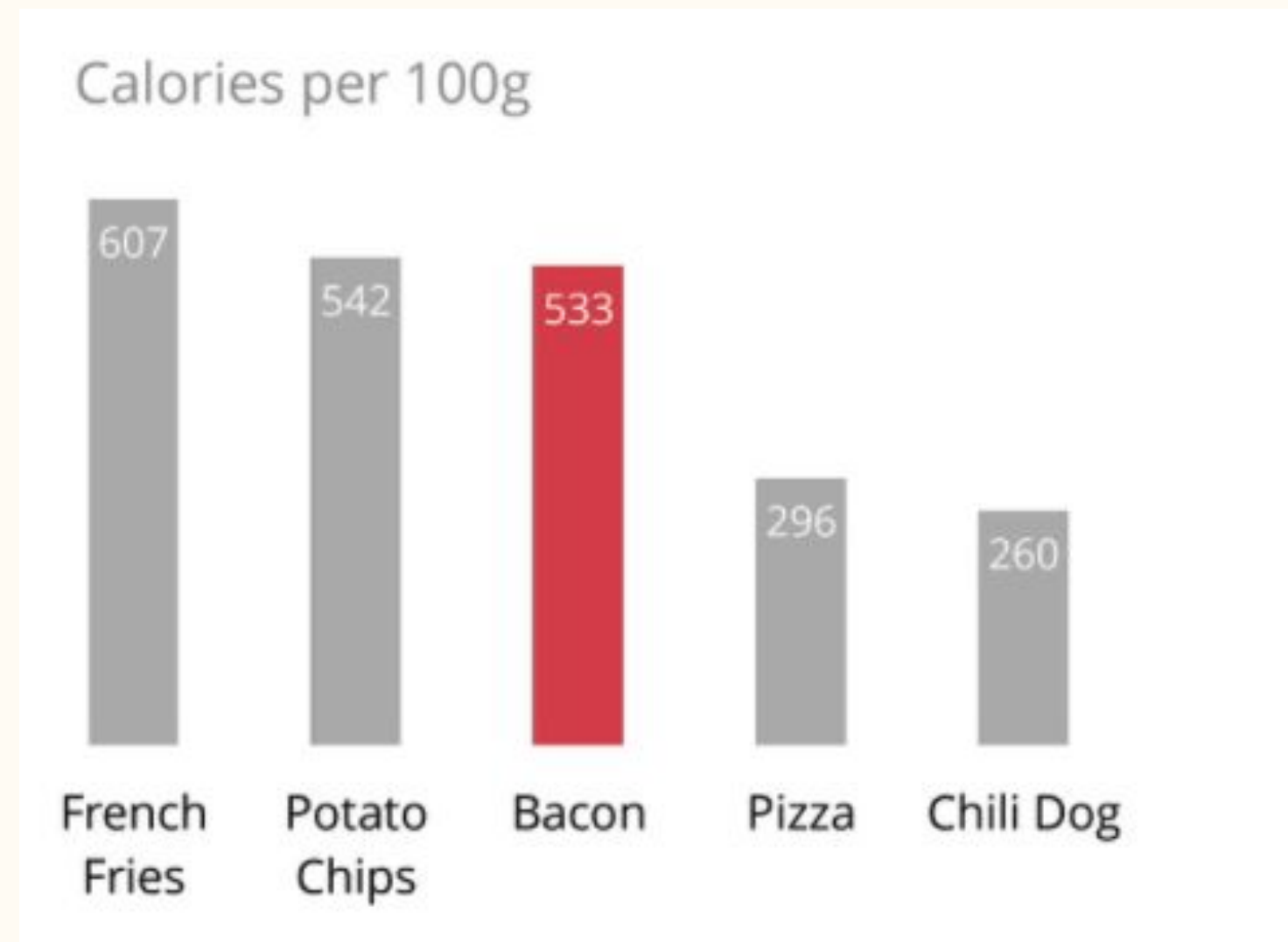
[Source: Darkhorse Analytics](#)



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.

Clutter removal technique #7 — Add labels directly on the plot



[Source: Darkhorse Analytics](#)



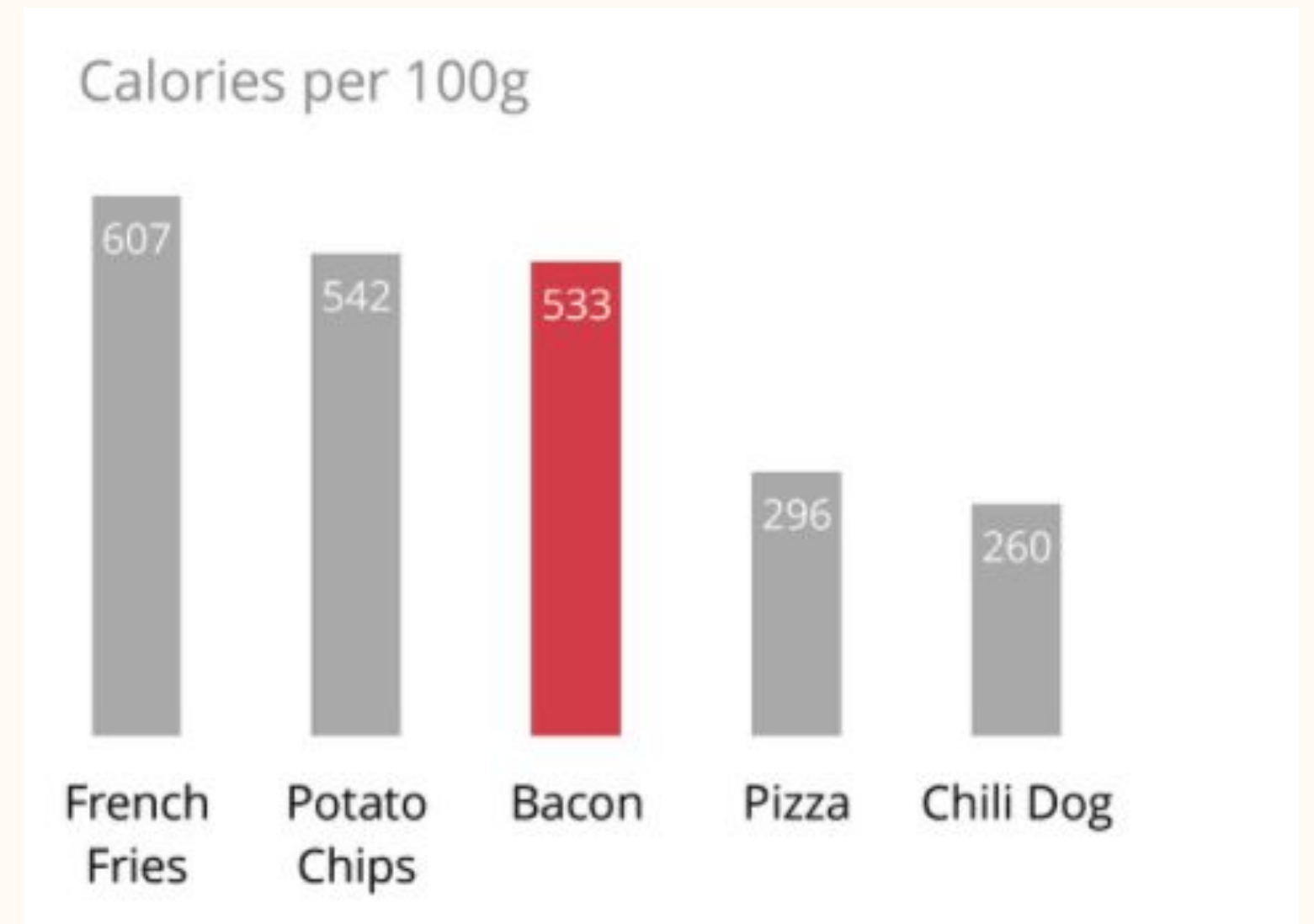
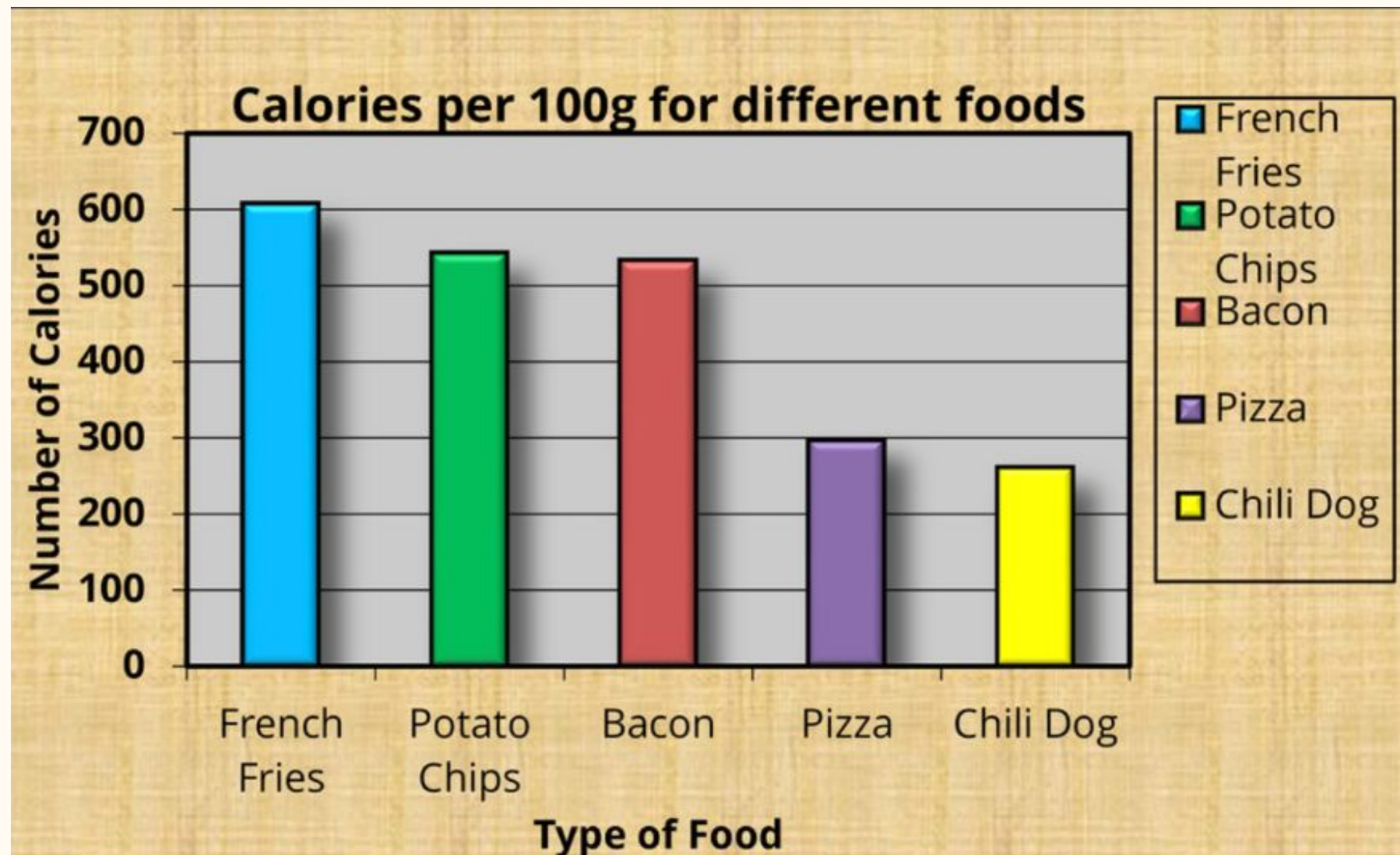
Decluttering techniques at your disposal

- ✓ Use white spaces
- ✓ Remove chart borders
- ✓ Remove gridlines or axes
- ✓ Clean up axis labels
- ✓ Label data directly (as opposed to using a legend)
- ✓ Remove data markers
- ✓ Use special effects (bold, underline, italics, shadows) sparingly



The cognitive load and effectiveness tradeoff

Each visualization serves a purpose. As a rule of thumb, remove all the elements of your visualization that doesn't serve a purpose.



[Source: Darkhorse Analytics](#)





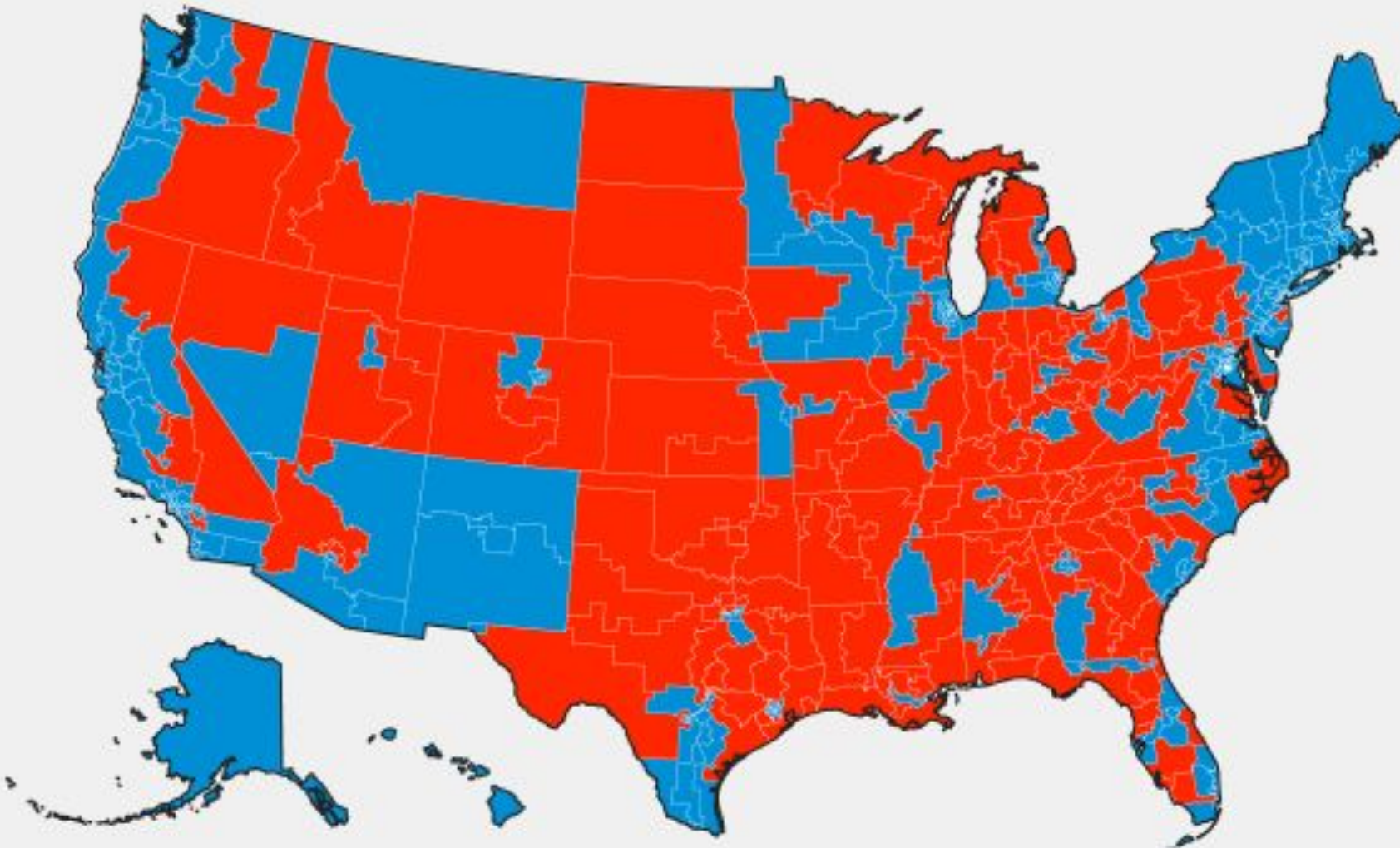
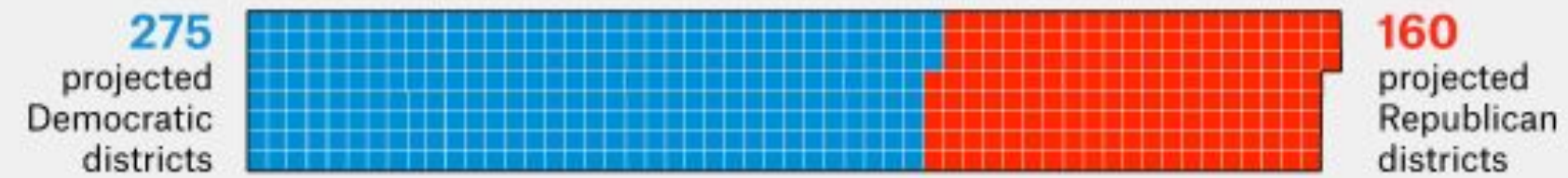
Rule #3

Use color creatively

Colors to distinguish between groups

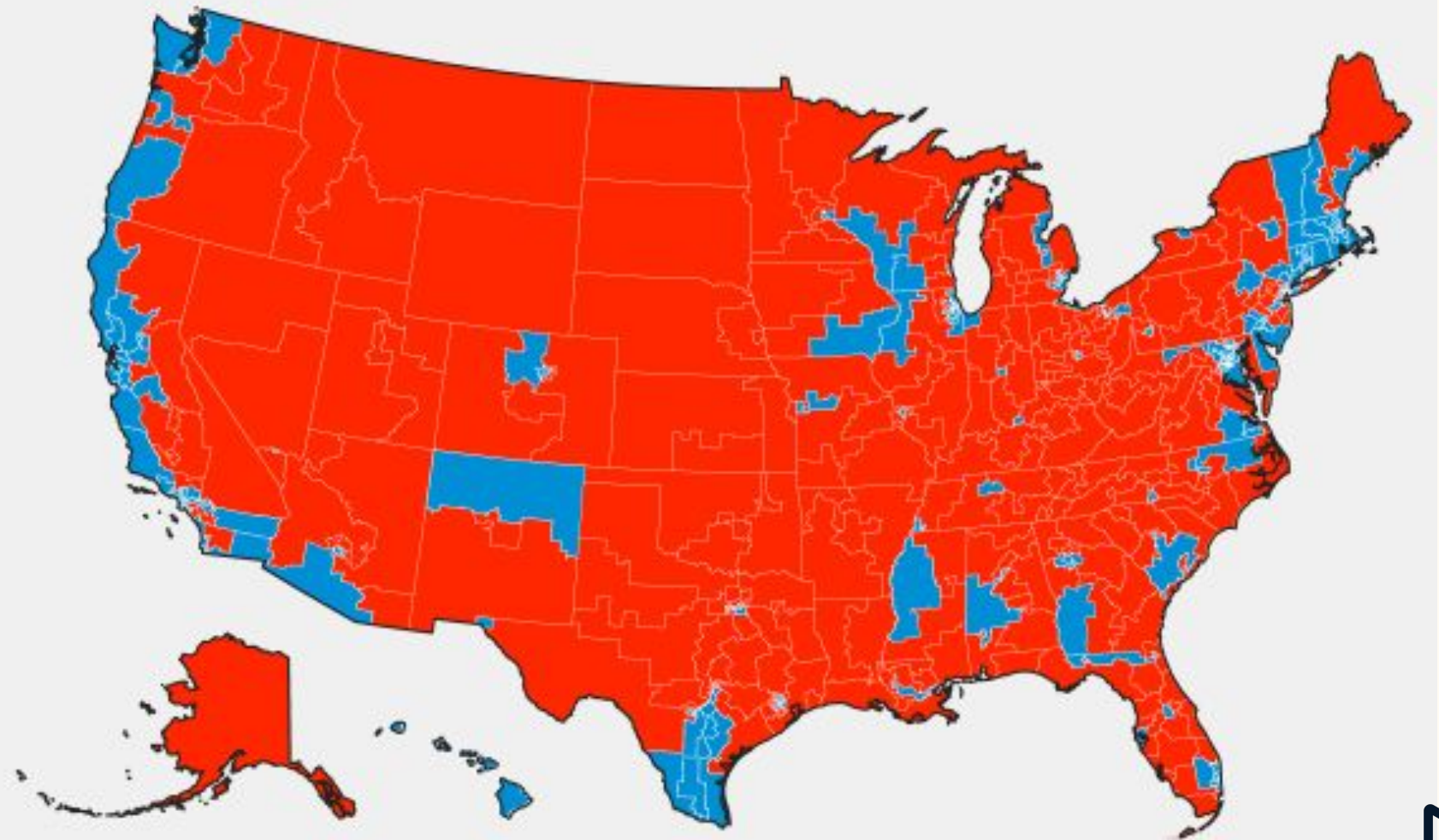
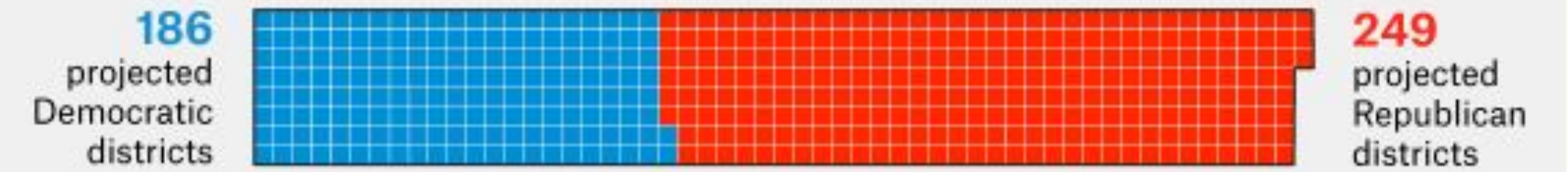
What if only women voted?

Projected results for the 2018 midterms based on polling patterns and FiveThirtyEight's Lite forecast on Oct. 24



What if only men voted?

Projected results for the 2018 midterms based on polling patterns and FiveThirtyEight's Lite forecast on Oct. 24



How this can be used in “normal” datasets

**Customer
lifetime value**

What drives customer loyalty at your organization?



**Products
purchased**

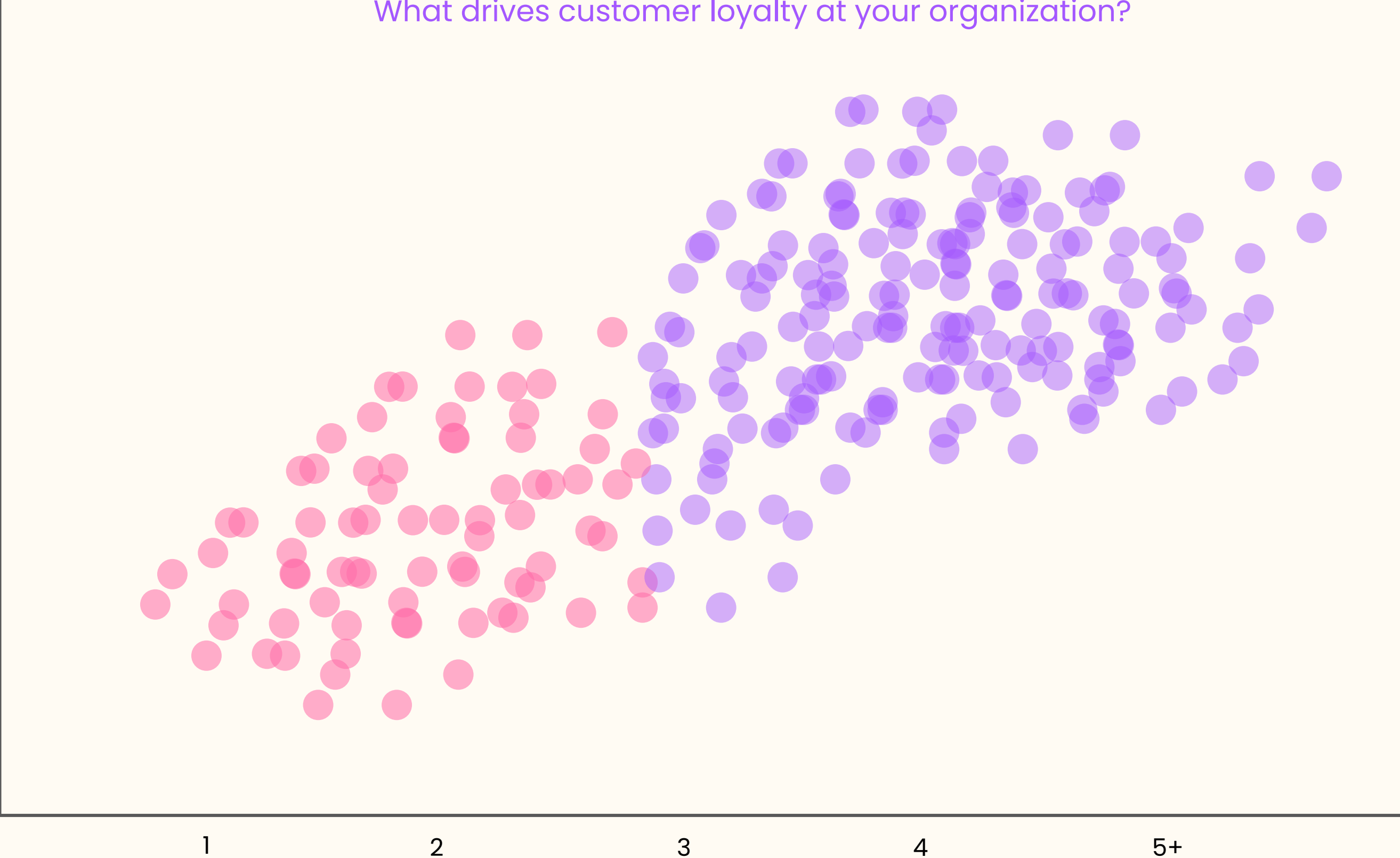


How this can be used in “normal” datasets

**Customer
lifetime value**

What drives customer loyalty at your organization?

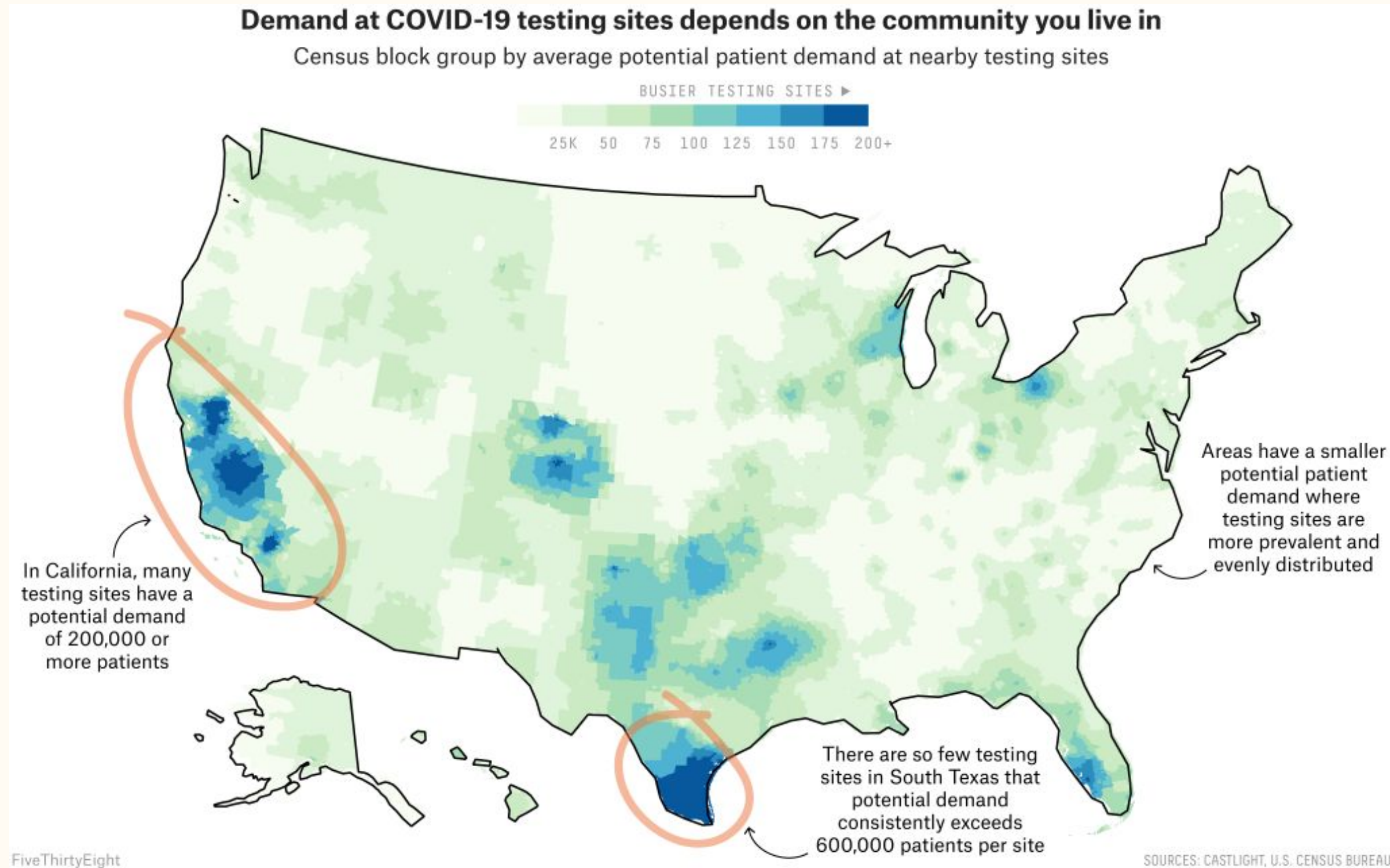
Segment A
Segment B



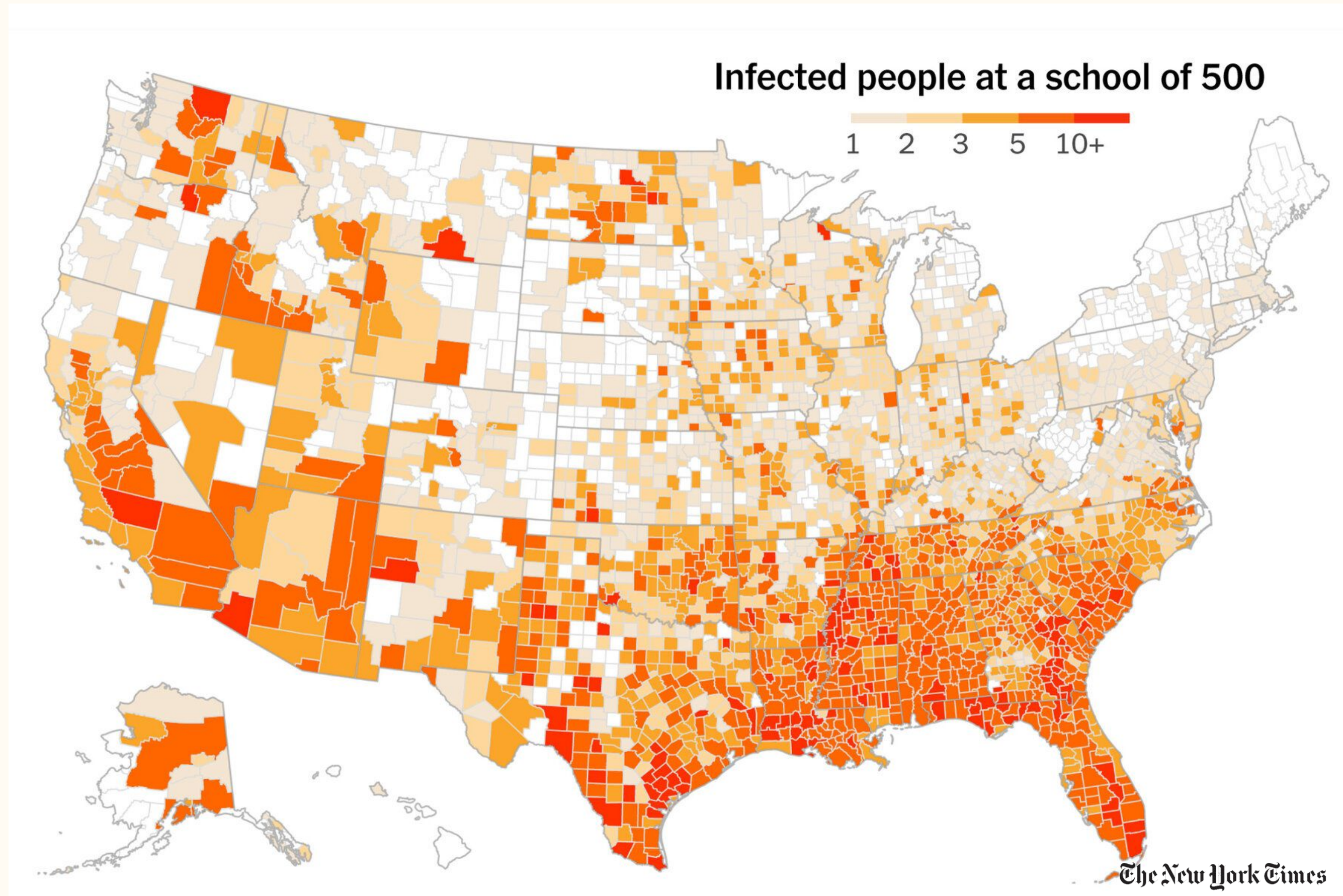
**Products
purchased**



Colors to highlight intensity



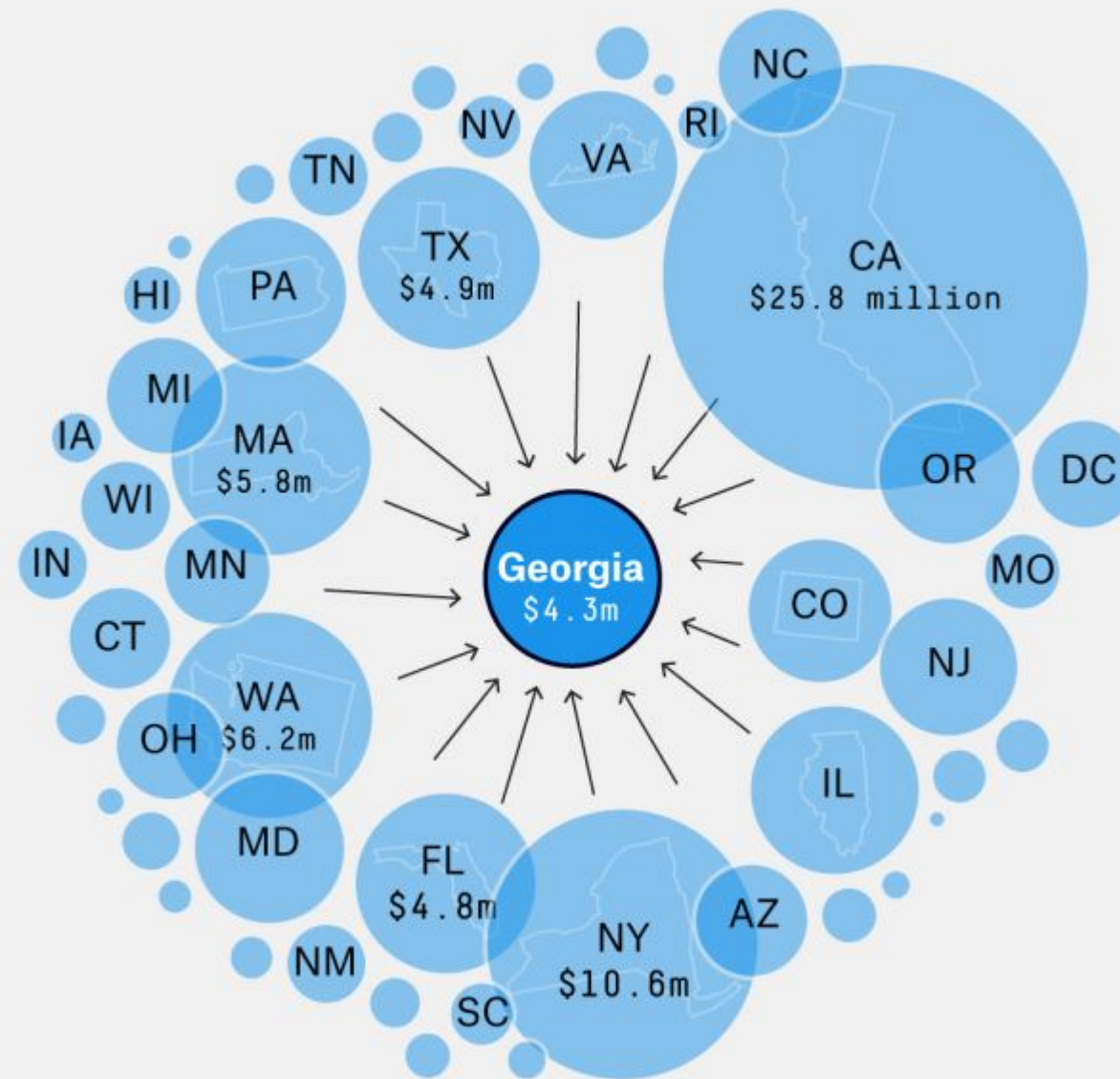
Visualizing covid infection hotspots



Colors to distinguish between groups

Democrats from every state are donating to Georgia

All donations sent to Democratic candidates in Georgia's Senate runoffs via ActBlue, from Nov. 4 through Nov. 23



FiveThirtyEight

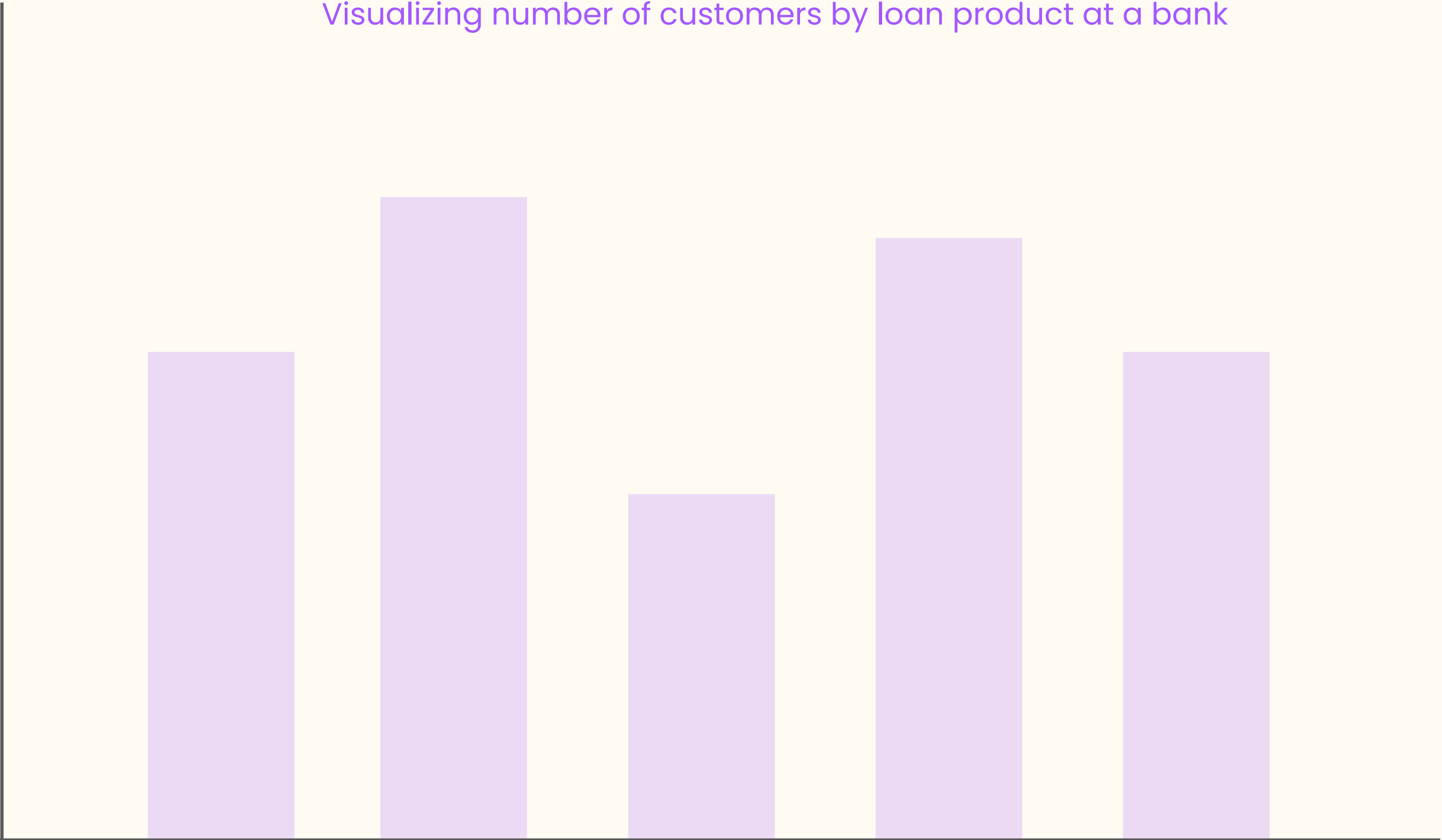
SOURCE: FEDERAL ELECTION COMMISSION



How this can be used in “normal” datasets

Number of customers

Visualizing number of customers by loan product at a bank



Business Loans

Loan Against Property

Commercial Vehicle Financing

Construction Equipment Loan

Farm Equipment Loan

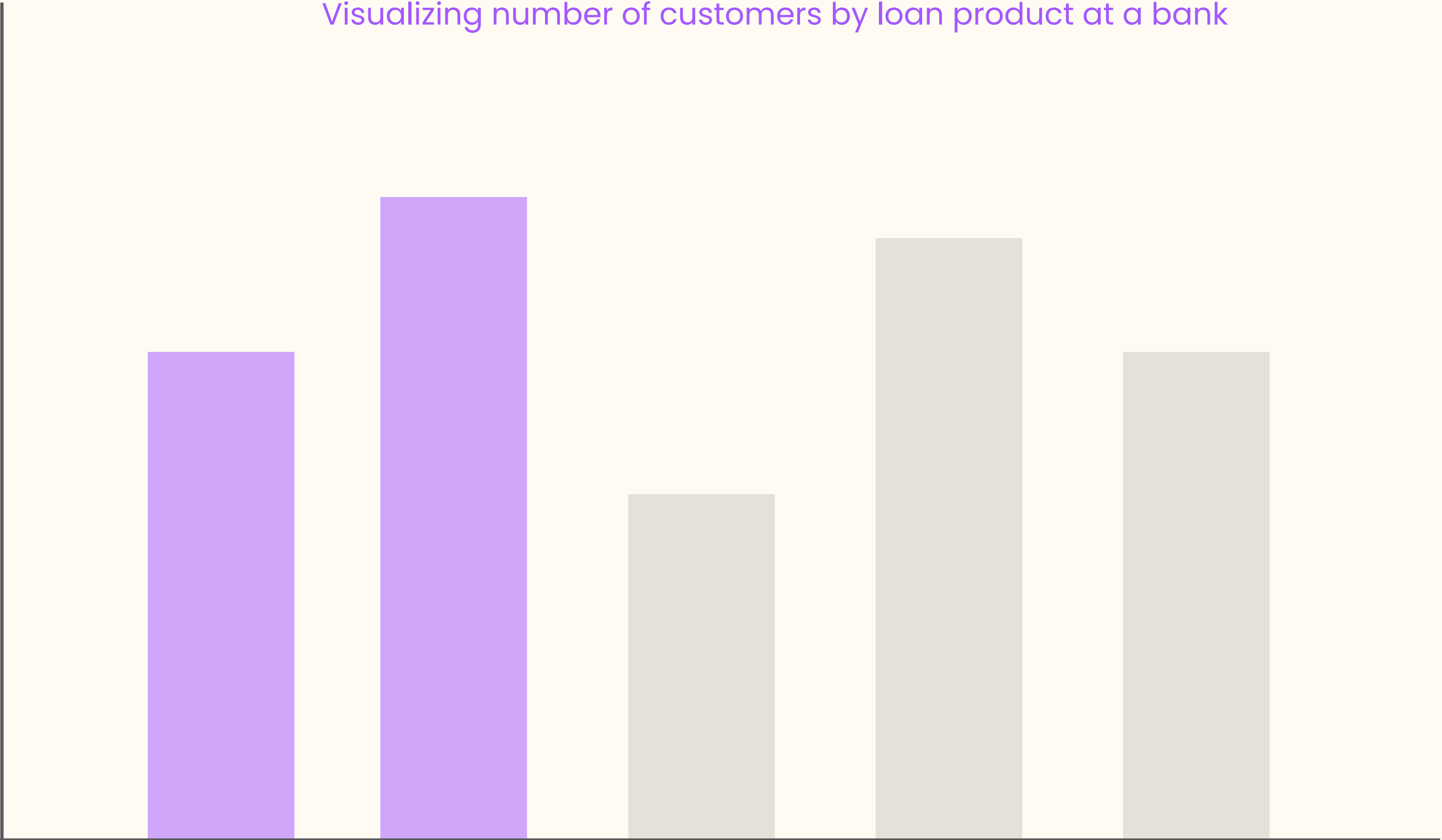
Loan Products



How this can be used in “normal” datasets

Number of customers

Visualizing number of customers by loan product at a bank



Loan Products

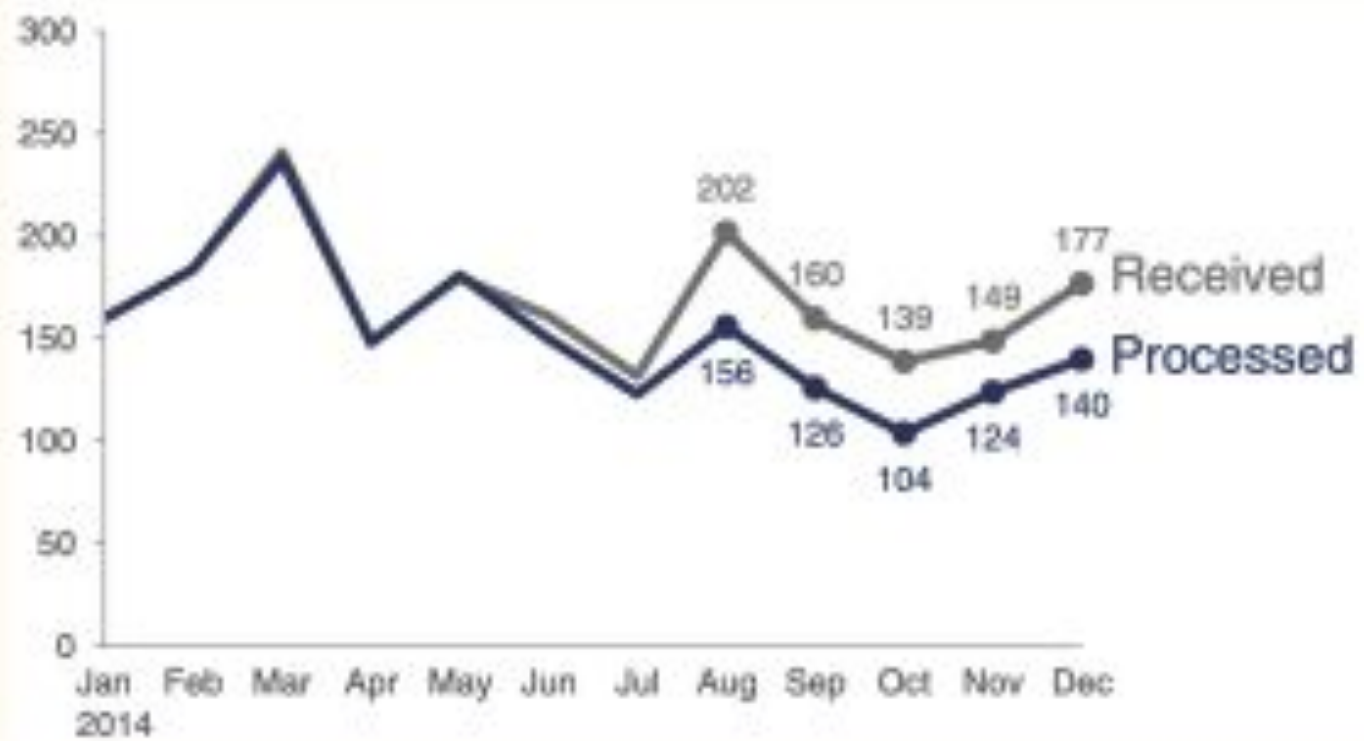
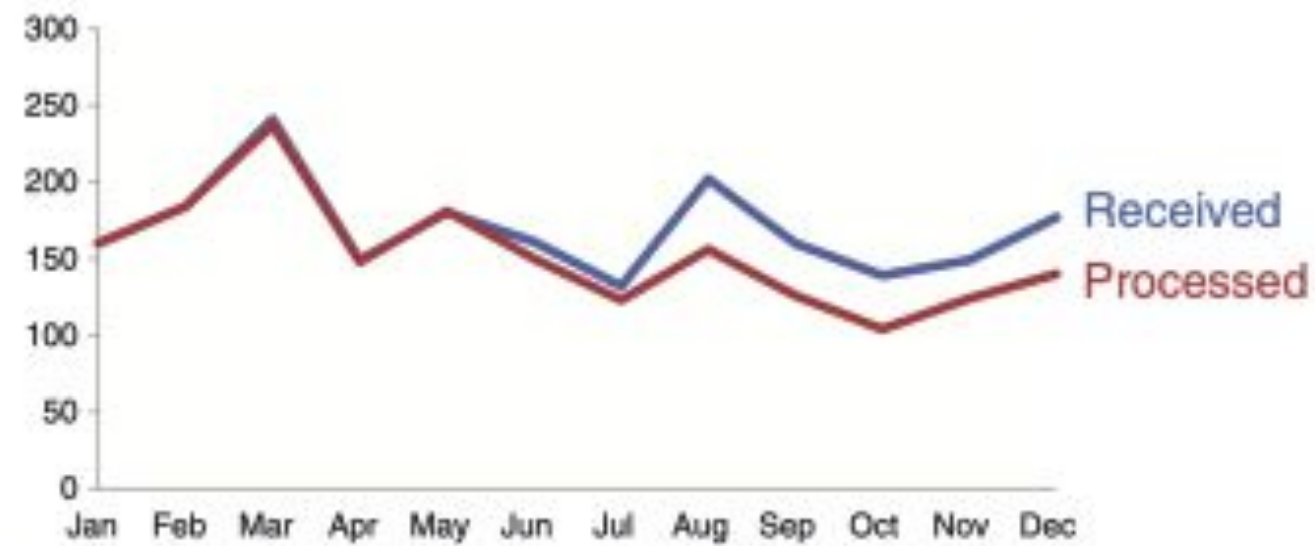




Rule #4

Use texts appropriately

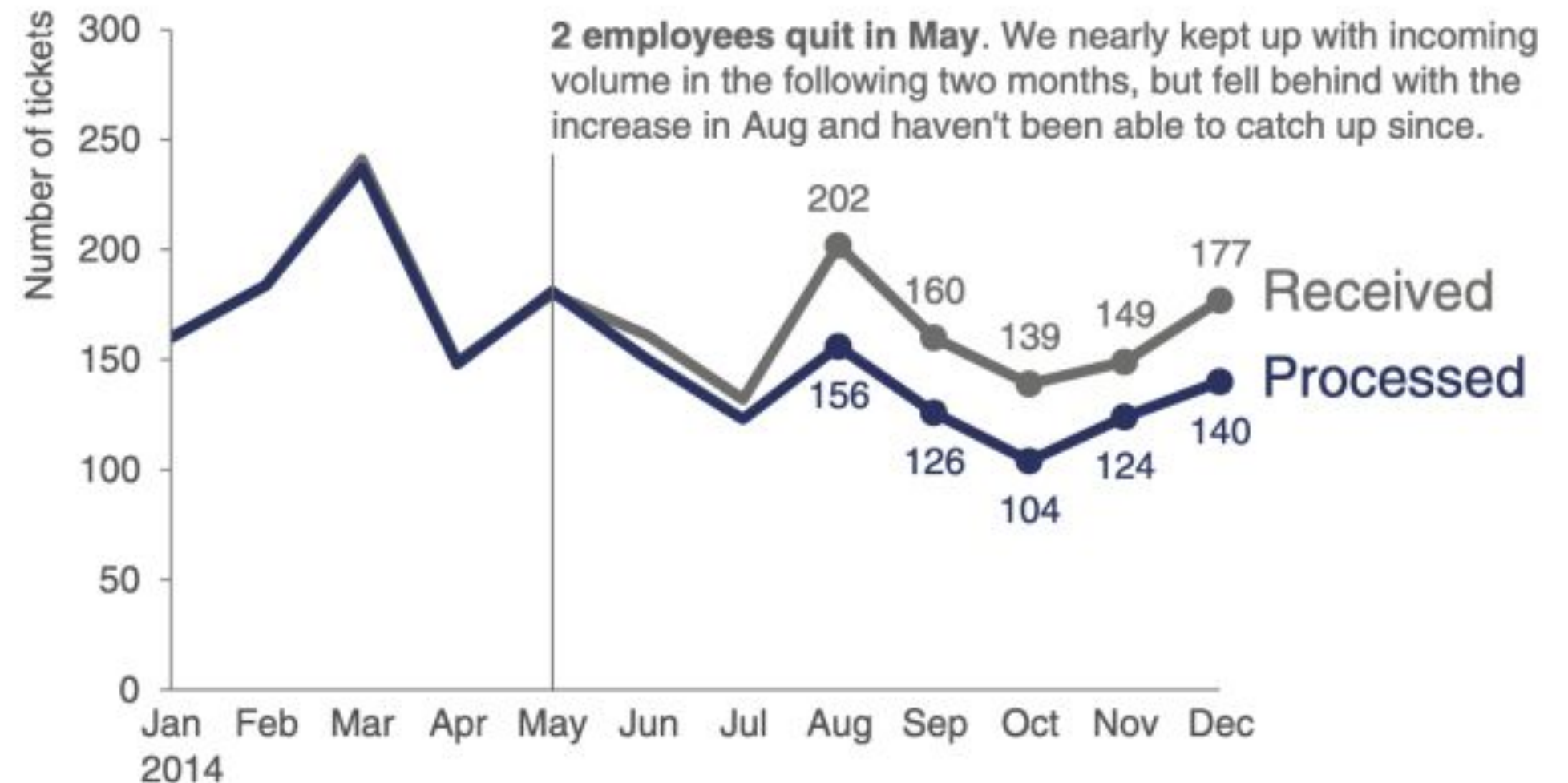
Labels can be extremely effective at highlighting insights



Please approve the hire of 2 FTEs

to backfill those who quit in the past year

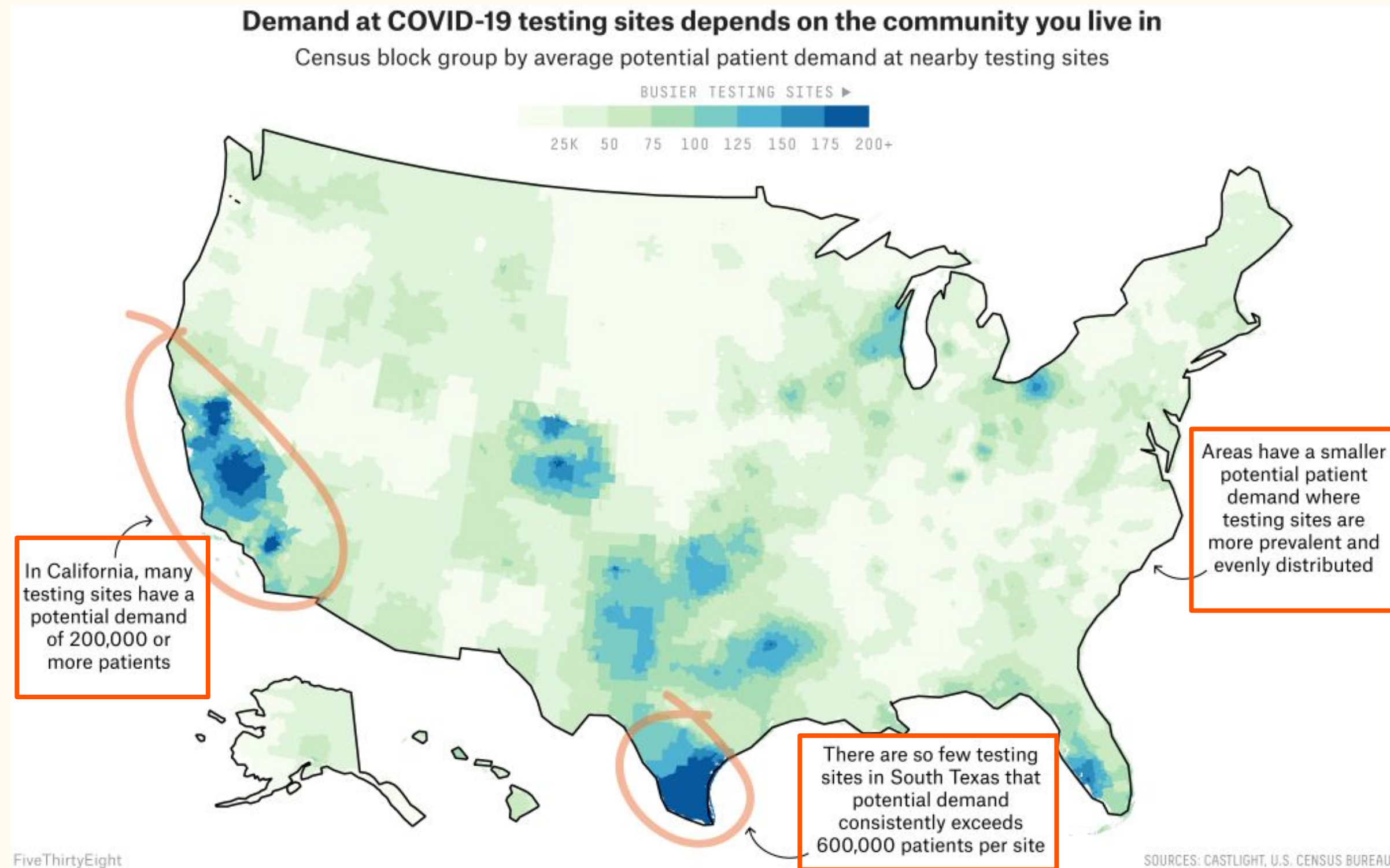
Ticket volume over time



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.



Labels can be extremely effective at showing context



Rules of thumb when using labels

- ✓ Label axes and titles for clarity
- ✓ Label data points when necessary
- ✓ Play around with font sizes when highlighting specific message
- ✓ Common audience questions should go into labels





2

8 rules for better data storytelling

Rules for better narrative



Rule #1

Know your audience, know your format

We have different stakeholders, and different formats

AUDIENCE TYPES



Executive
Low data literacy
Cares about outcomes and decisions



Data Science Leader or Partner
Data expert
Cares about rigour and insights



Business Partner
High data literacy
Cares about tactical next steps

FORMAT TYPES



Presentation



Coding Notebook



Written Report



Know the priority of the audience

AUDIENCE TYPES



Executive

Low data literacy

Cares about outcomes and decisions

Cares much more about business impact than a 1% incremental gain in machine learning model accuracy or a new technique you're using



Data Science Leader or Partner

Data expert

Cares about rigour and insights

Cares much more about how you arrived at your insights and to battle test them for rigour



Business Partner

High data literacy

Cares about tactical next steps

Cares much more about how your analysis impacts their workflow, and what should be their main takeaway from the data story

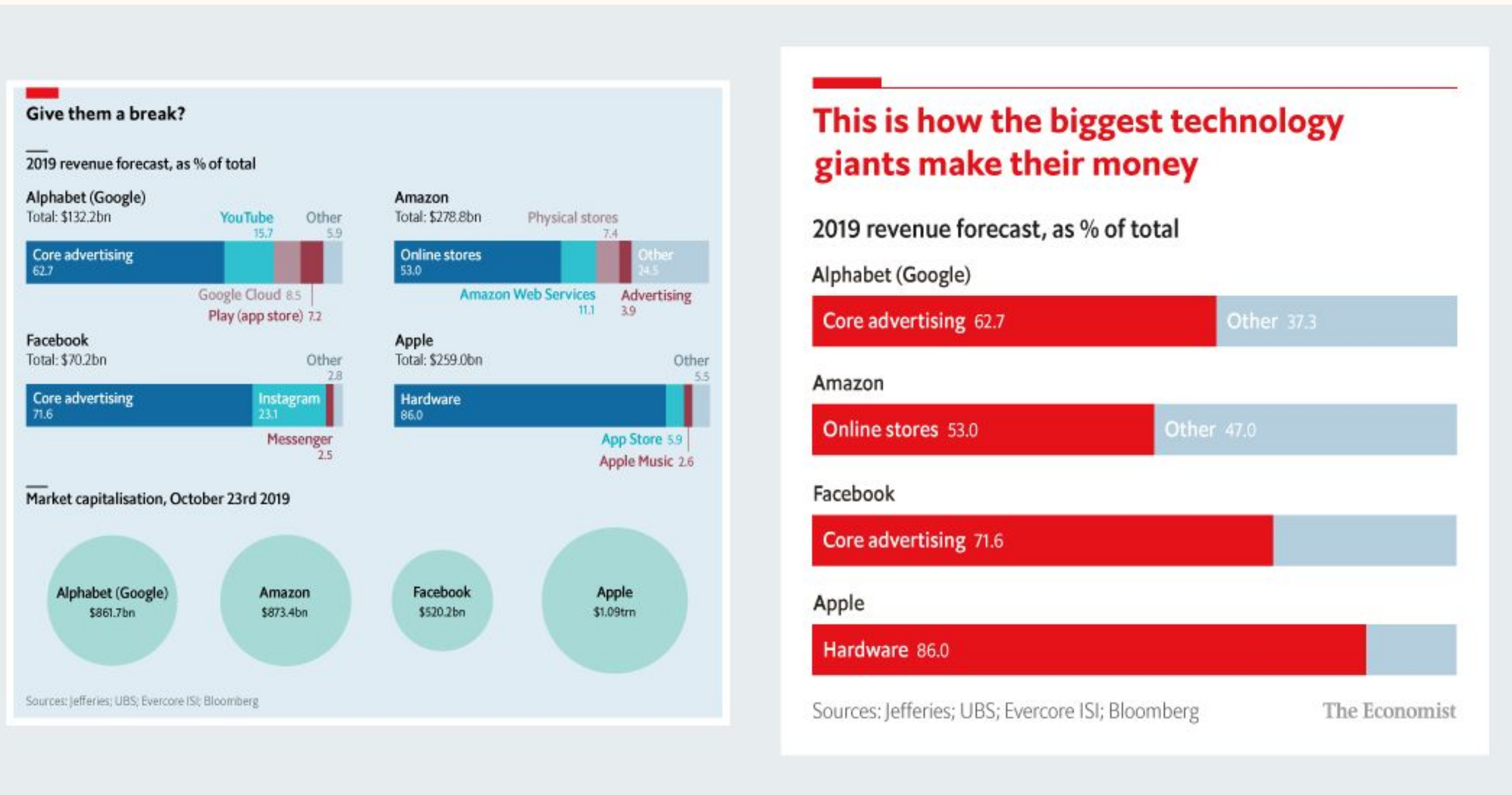


Practice empathy with your audience

- ✓ Does the audience have the necessary prerequisite knowledge to understand a particular metric?
- ✓ How much time does the audience have to consume this data story?
- ✓ What is the medium of presentation (written/oral) that the audience prefers?



Know their level of data literacy (or subject matter expertise)



[How the economist altered its charts on instagram for a younger audience](#)





Rule #2

Begin with the goal in mind

It's very tempting to throw everything you got at a data story



Start with the **goal in mind**

Who is the audience?
The more specific the audience, the easier it is to build a data story that resonates

What should the audience know?
Know your recommendations before you craft the data story

What data can you use to convey your point?
Identify the data that is relevant for this data story

Goal
Convince management of investing in additional resource on support team



Start with the **goal in mind**

Who is the audience?
The more specific the audience, the easier it is to build a data story that resonates

What should the audience know?
Know your recommendations before you craft the data story

What data can you use to convey your point?
Identify the data that is relevant for this data story

Goal
Convince management of investing in additional resource on support team

However, this doesn't mean that goals should determine the data story — what the **data is telling you always takes precedence**

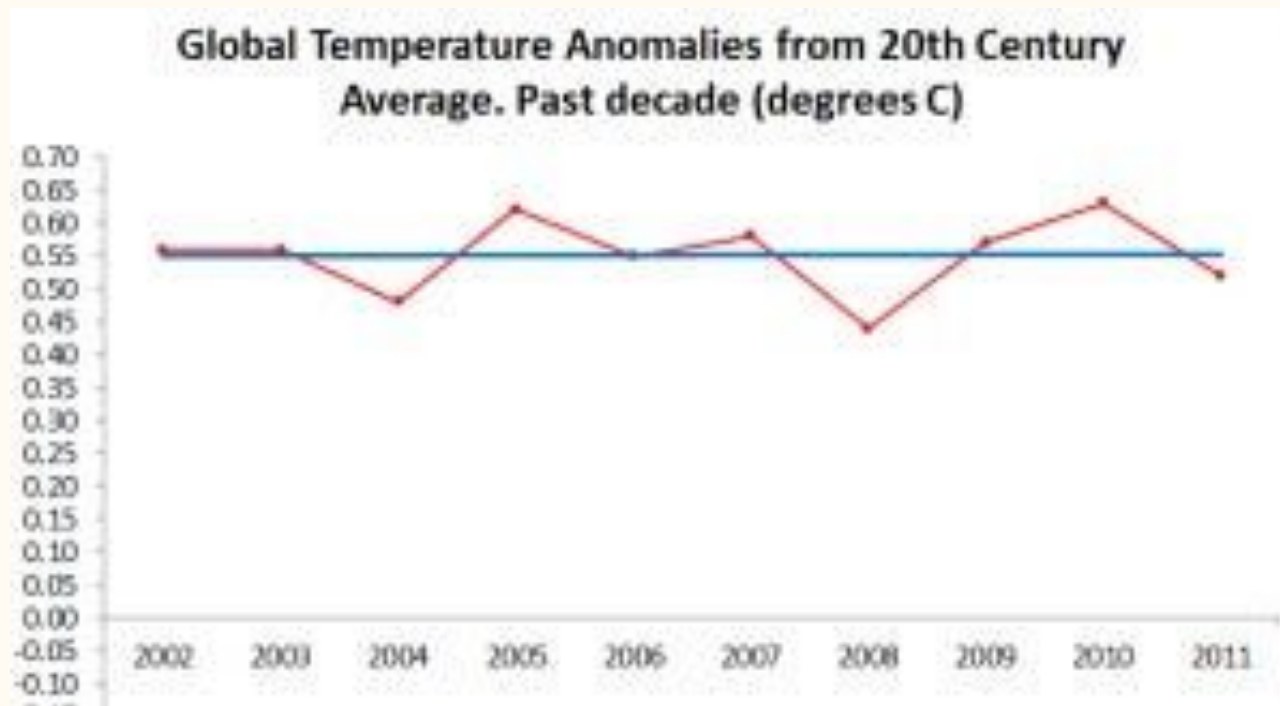




Rule #3

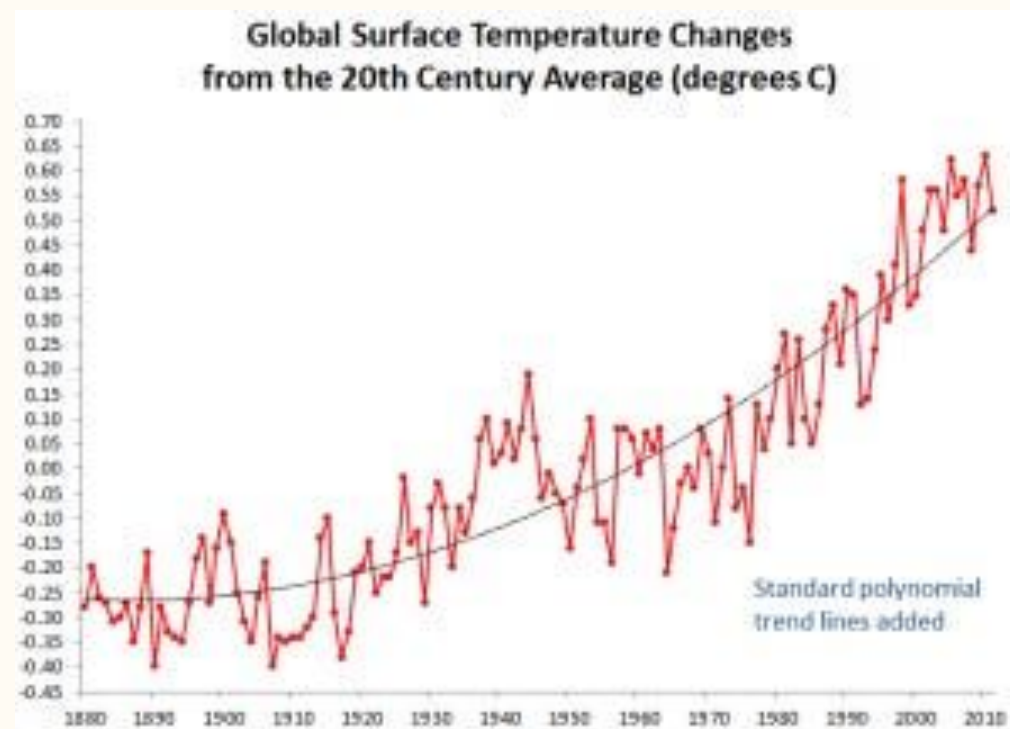
Do not mislead with data stories

Avoiding the fastest way to lose credibility



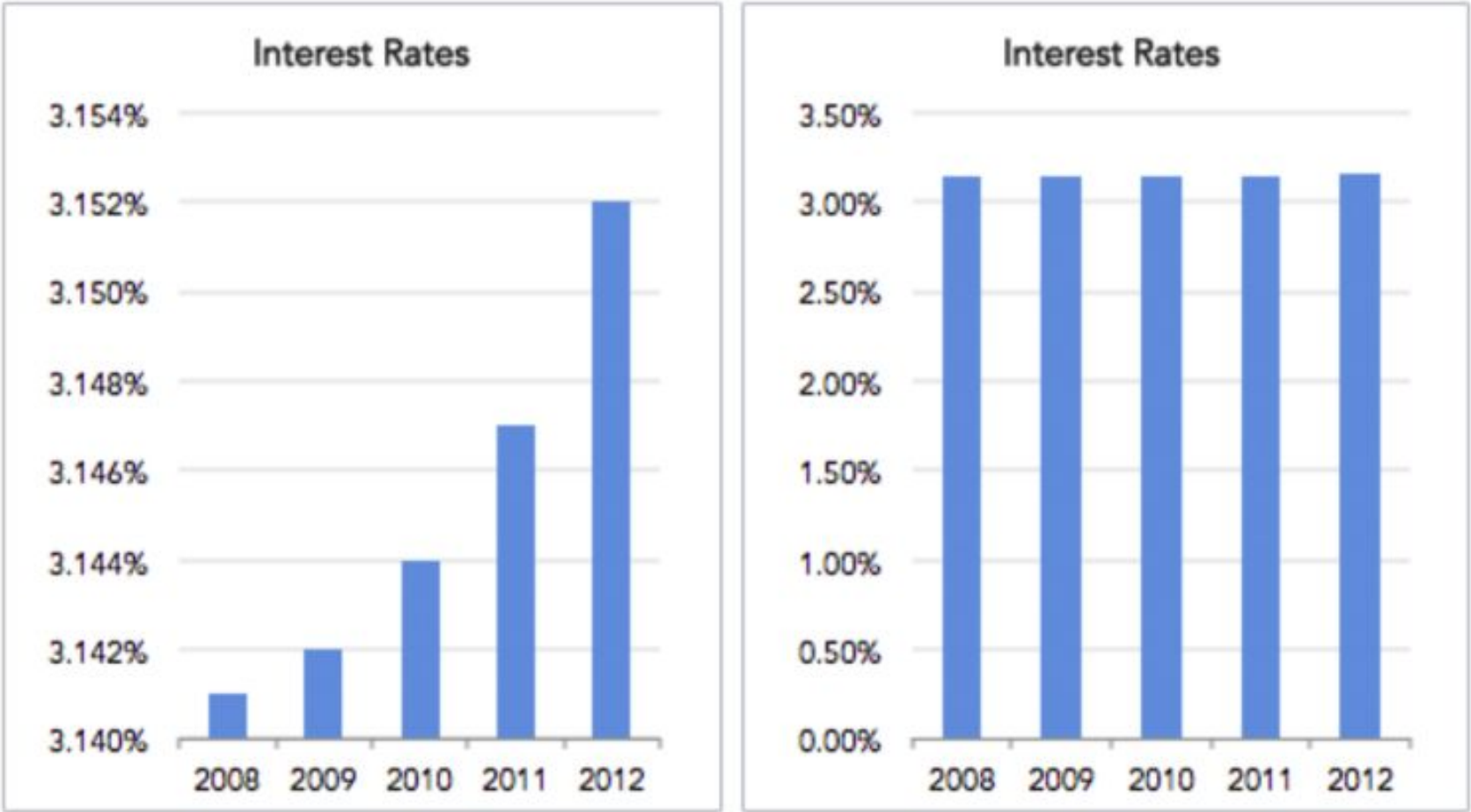
Pitfall to avoid #1

Always make sure your time horizons make sense given the data you're treating



Avoiding the fastest way to lose credibility

Same Data, Different Y-Axis



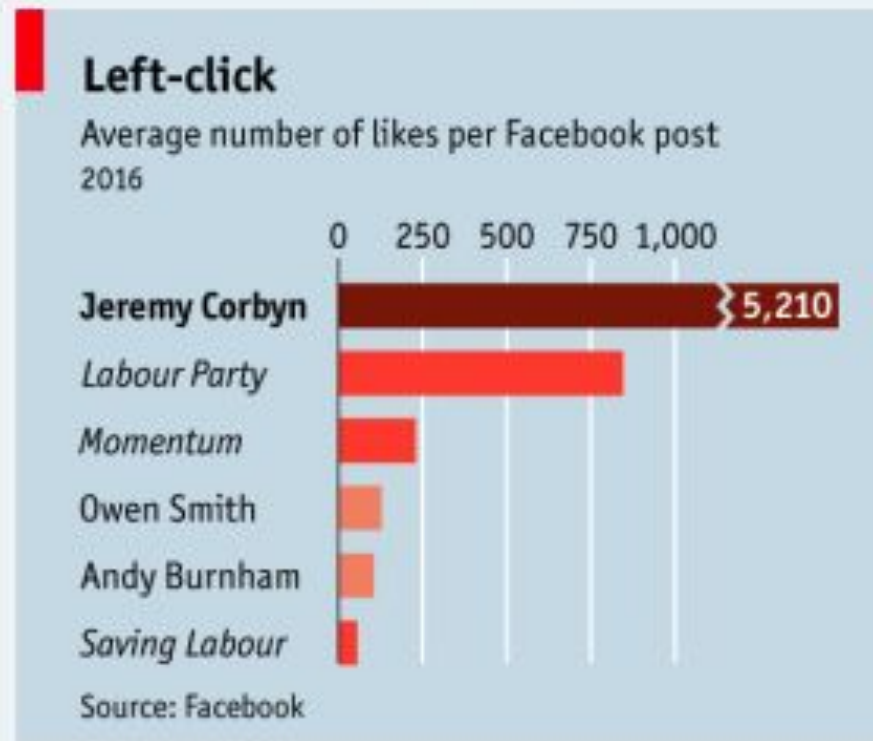
Misleading Graph No. 1

Pitfall to avoid #2
Ensure that axes start with zero

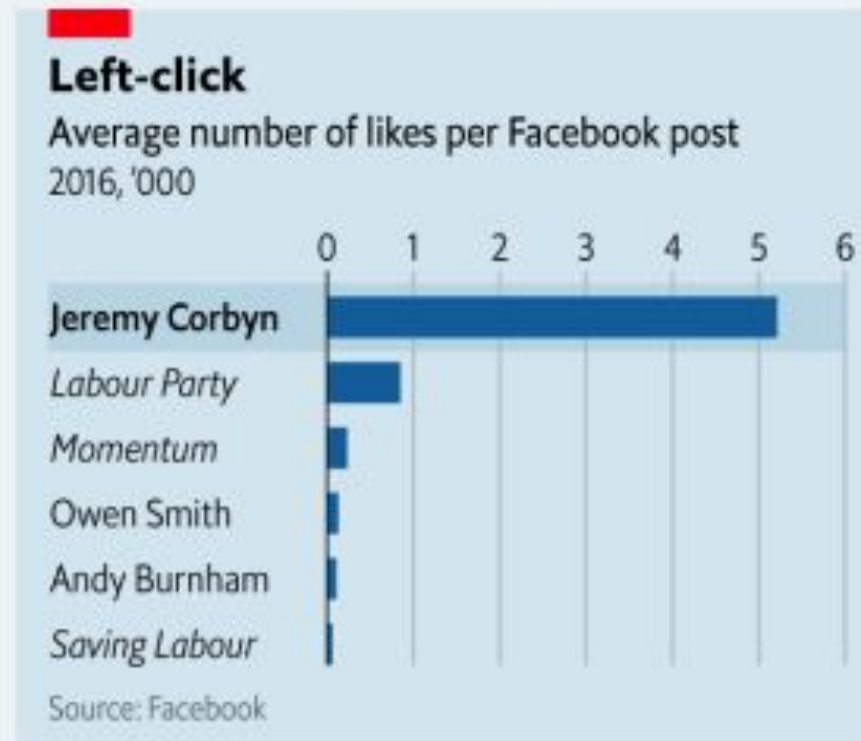


Avoiding the fastest way to lose credibility

Original



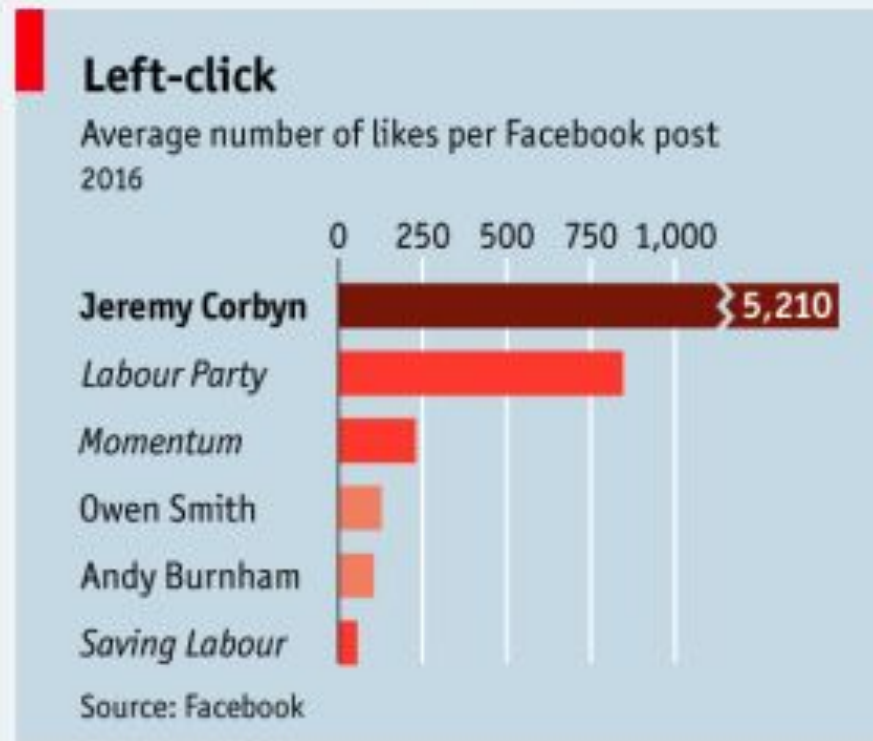
Better



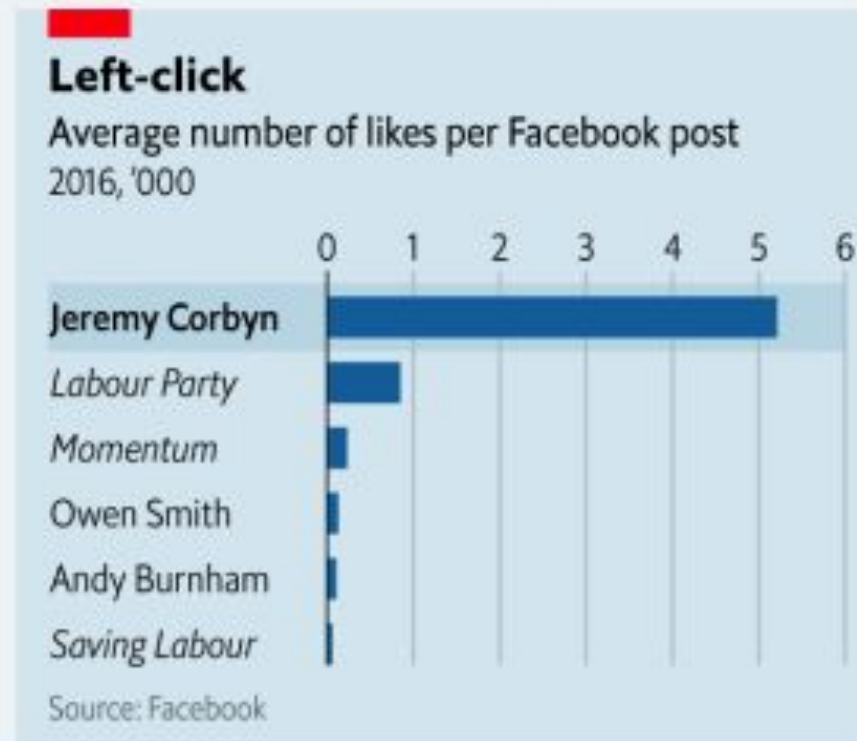
Pitfall to avoid #3
Ensure that axes scales are appropriate

Avoiding the fastest way to lose credibility

Original



Better



Pitfall to avoid #3
Ensure that axes scales are appropriate

Other best practices to ensure data stories are not misleading

- ✓ If you're sampling data, make sure sample is representative of population
- ✓ Use centrality measures (median, mean, etc...) to ensure context around a population is taken into account





Rule #4

Develop a narrative around your data

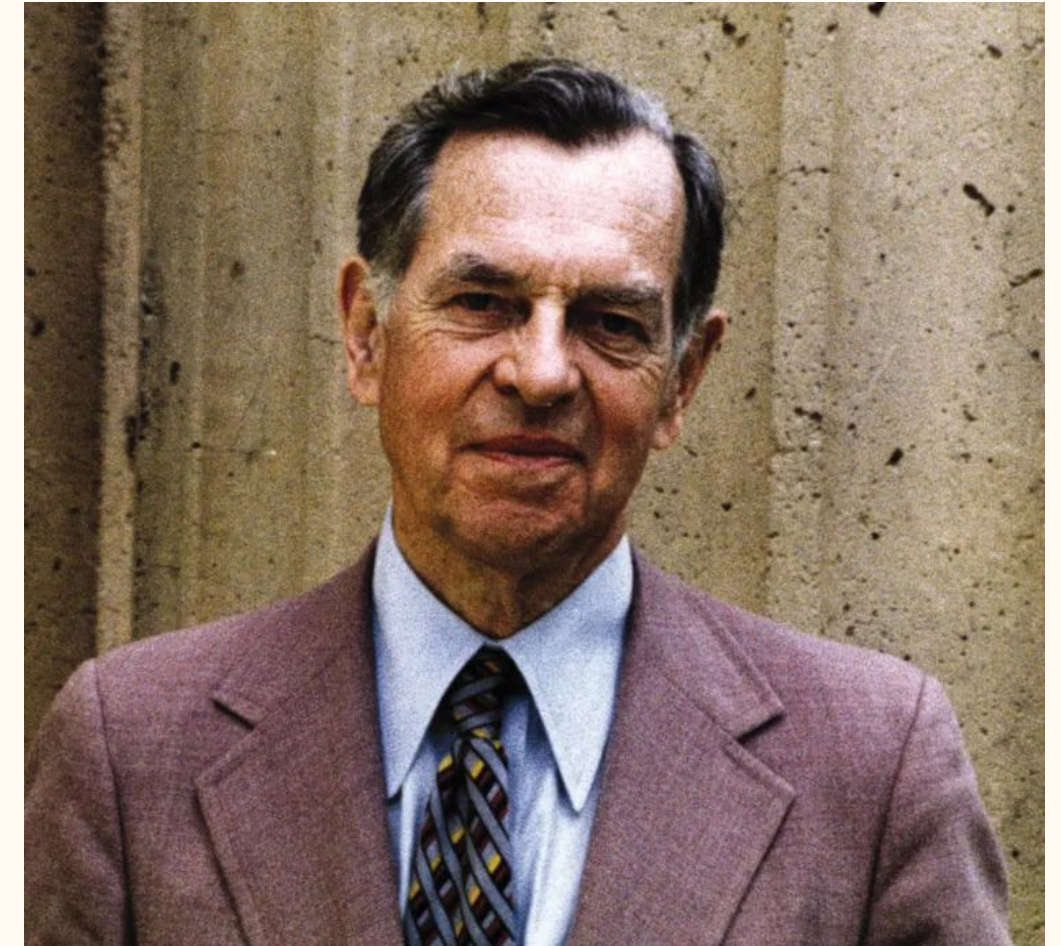
Different narrative structures to choose from



Aristotle's Tragedy Structure



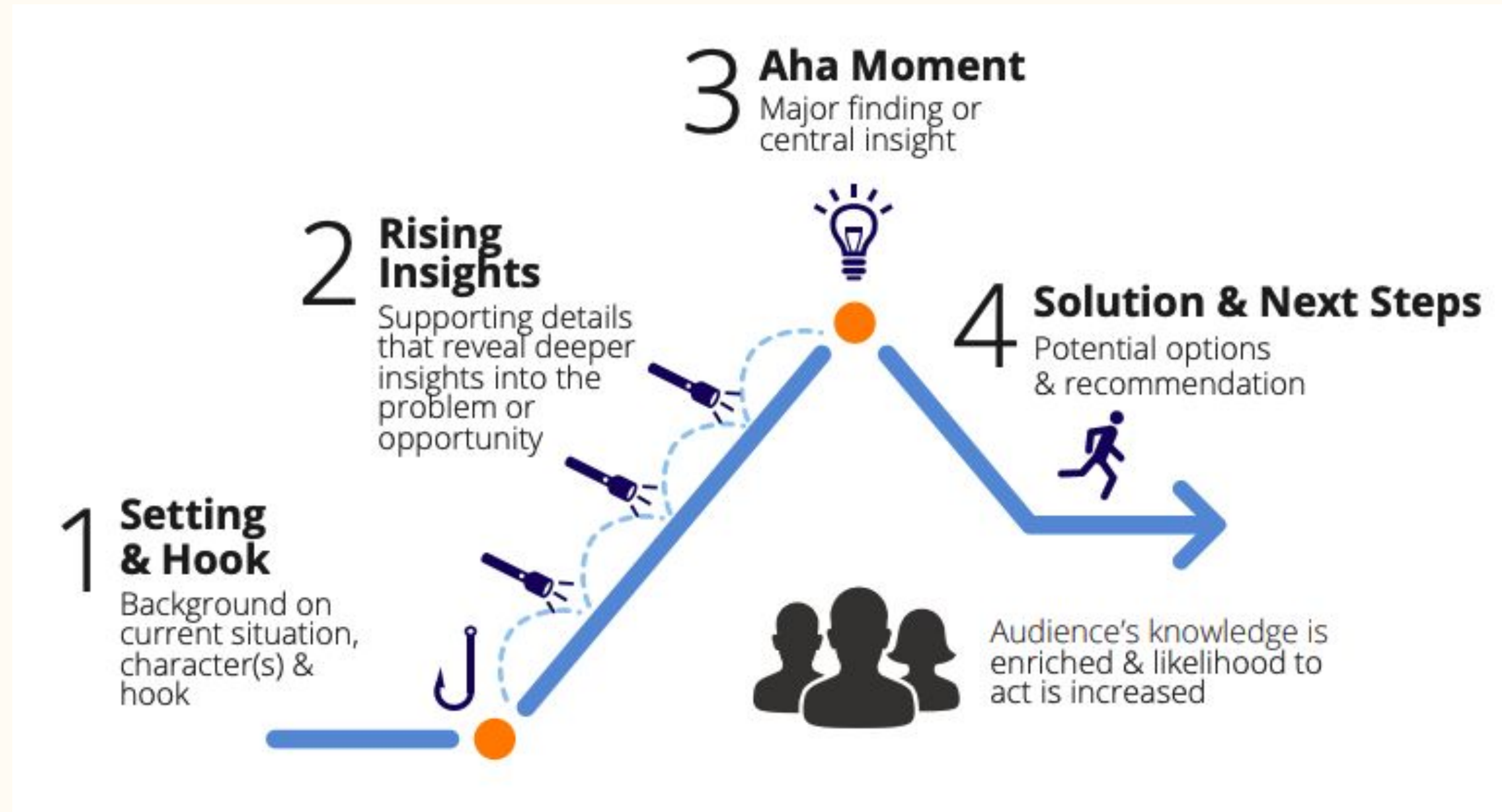
Freytag's Pyramid



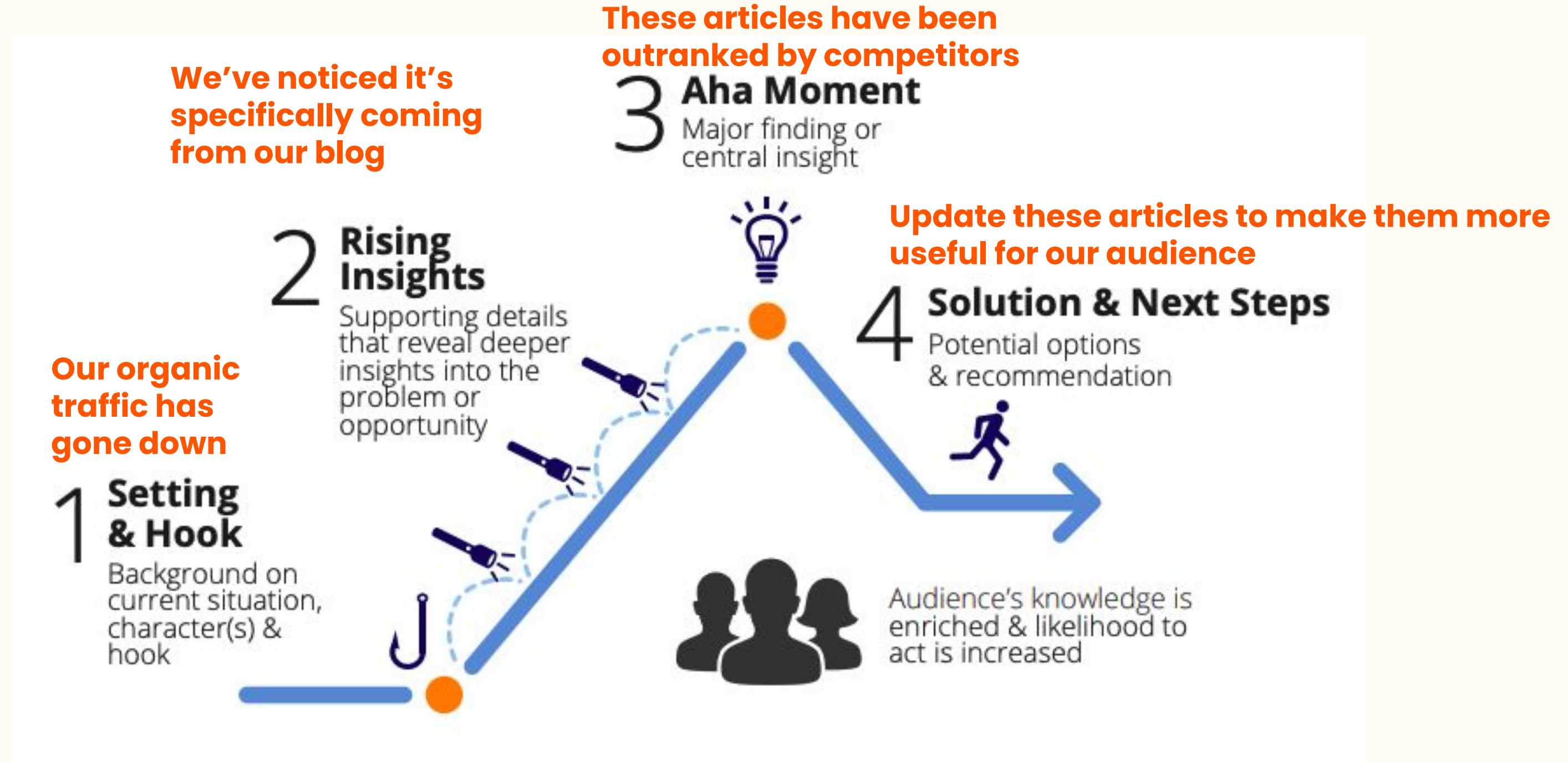
Campbell's Hero Journey



Different narrative structures to choose from



Different narrative structures to choose from



More resources on narrative

- ✓ [Tableau's 5 best practices for telling great stories with data](#)
- ✓ Brent Dyke's [Effective Data Storytelling: How to Drive Change with Data, Narrative, and Visuals](#)
- ✓ [Storytelling for more impactful data science by Gert de Geyter](#)
- ✓ [The data storytelling skills teams need with Andy Cotgreave](#)
- ✓ [Cole Nussbaumer Knaflic's Storytelling with Data](#)



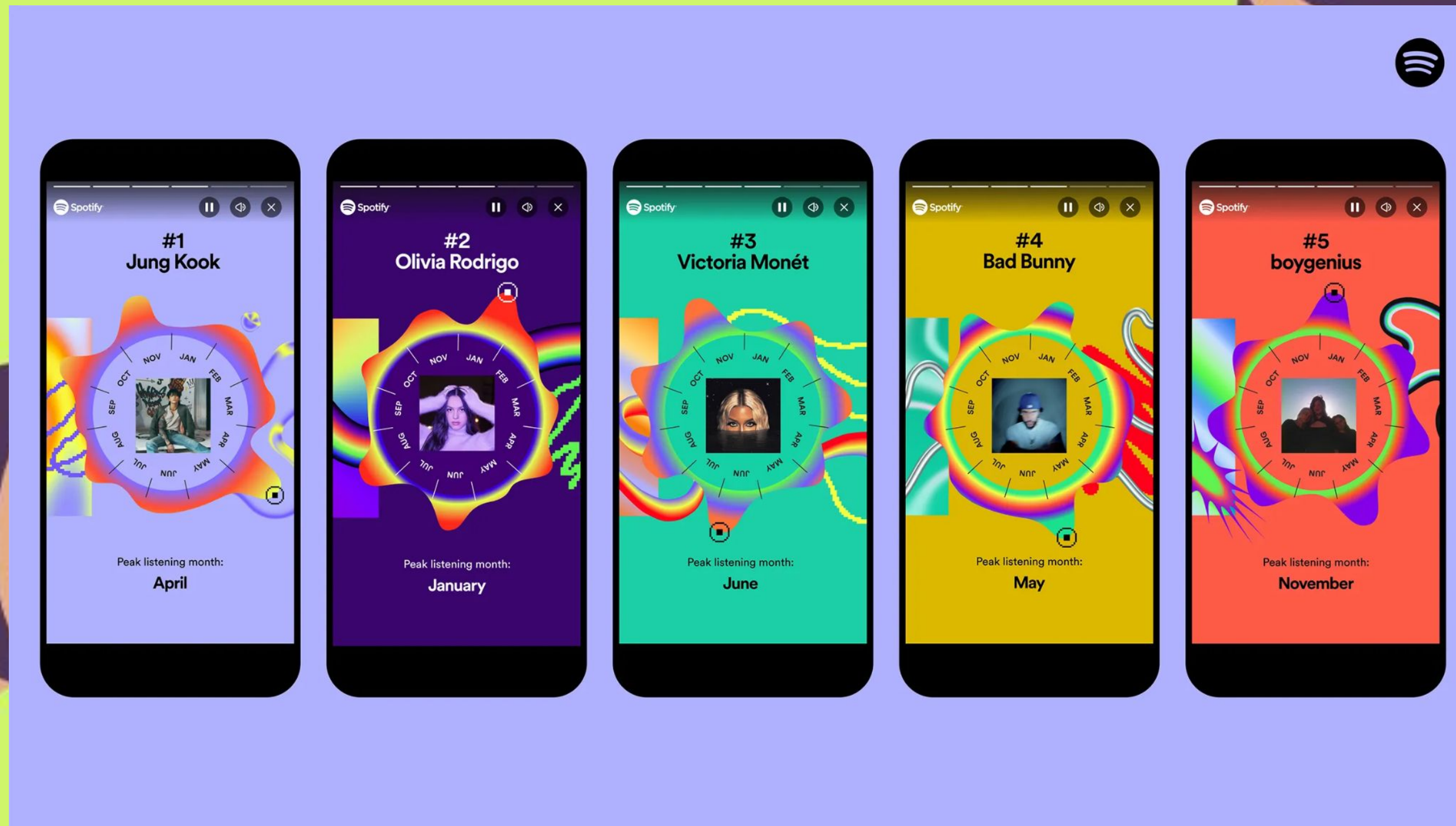


3

Everyone will become a data storyteller

*Resources for your data
visualization skills*

Data storytelling is table stakes now



Data storytelling is table stakes now

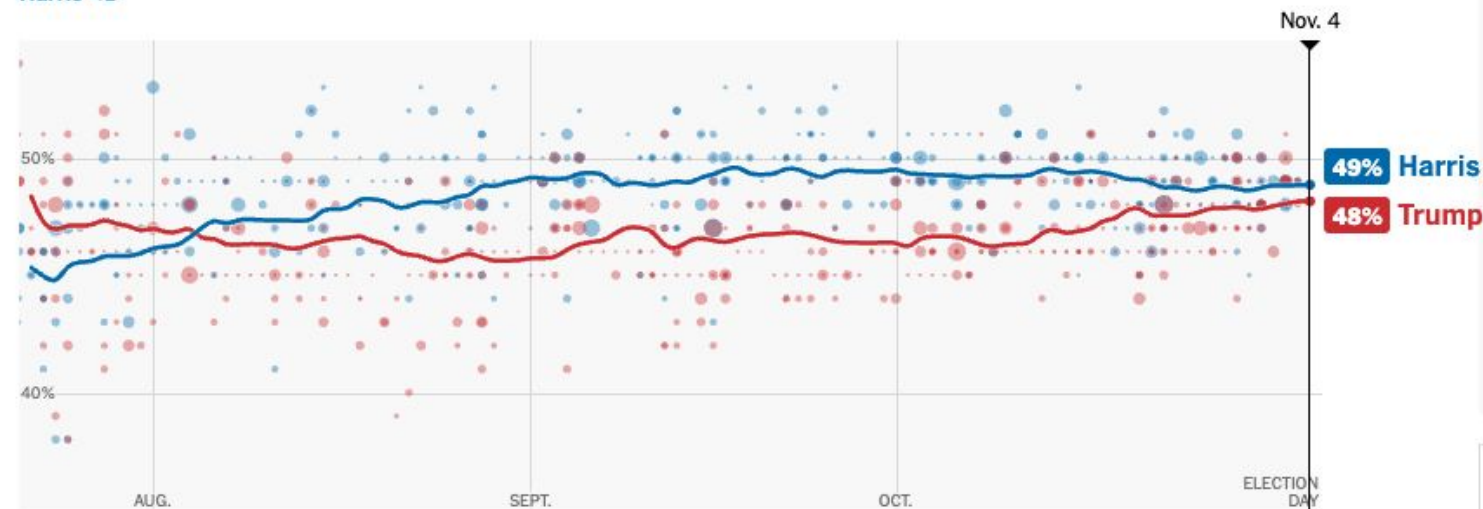
Election 2024 Polls: Harris vs. Trump

Updated Nov. 4, 2024 [Leer en español](#)

[See the final Times/Siena polls of Arizona, Georgia, Michigan, Nevada, North Carolina, Pennsylvania and Wisconsin >](#)

Who's leading the polls?

National polling average
Harris <1



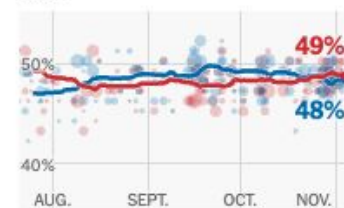
Nate Cohn
Chief political analyst

While the overall result of our final Times/Siena polls is largely unchanged since our previous wave of battleground polls, there were some notable shifts. Surprisingly, the longstanding gap between the Rust Belt and Sun Belt battlegrounds narrowed considerably. The overall effect of these swings is somewhat contradictory — Harris's position in the Electoral College isn't necessarily improved.
Updated Nov. 3

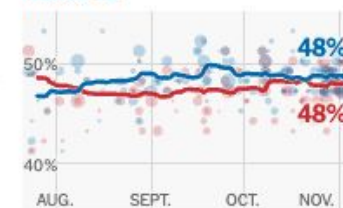


Enjoy open access to the election hub in The Times app.
Download The Times app to explore the hub, for a limited time.

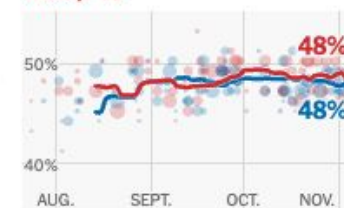
Pennsylvania >
Even



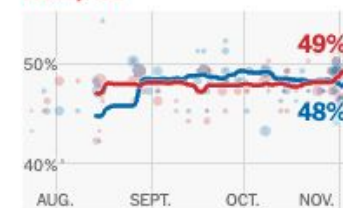
Michigan >
Harris <1



North Carolina >
Trump <1



Nevada >
Trump <1



Wisconsin >
Harris <1



Georgia >
Trump +1



Arizona >
Trump +3



The screenshot shows a DataCamp exercise page. On the left, the title 'Customizing heatmaps' is displayed. Below it, a paragraph explains that Seaborn supports additional customizations for heatmaps. The 'Instructions' section, worth 100 XP, lists three tasks: creating a crosstab table, plotting a heatmap with the 'BuGn' palette, and disabling the color bar while increasing the linewidth to 0.3. A 'Take Hint (-30 XP)' button is also present. The main area is a code editor for 'script.py' with the following Python code:

```
1 # Create the crosstab DataFrame
2 pd_crosstab = pd.crosstab(df["Group"], df["YEAR"])
3
4 # Plot a heatmap of the table with no color bar and using the BuGn palette
5 sns.heatmap(pd_crosstab, cbar=False, cmap="BuGn", linewidths=0.3)
6
7 # Rotate tick marks for visibility
8 plt.yticks(rotation=0)
9 plt.xticks(rotation=90)
10
11 # Show the plot
12 plt.show()
13 plt.clf()
```

Below the code editor is an 'IPython Shell' with 'In [1]:' and a 'Slides' tab. At the bottom right of the code editor are 'Run Code' and 'Submit Answer' buttons.

Level up your data visualization skills with DataCamp

Visualize intensity with Python using heatmaps

[> View Track](#)



Learn / Courses / Introduction to Tableau

Tableau Public - 1_3_new_york_neighborhood_prices

Columns: Neighborhood

Rows: (Empty)

Sheet 1

Neighborhood	Abc
Allerton	Abc
Alphabet City	Abc
Annadale	Abc
Astoria	Abc
Bath Beach	Abc
Battery Park City	Abc
Bay Ridge	Abc
Baychester	Abc
Bayside	Abc
Bedford Park	Abc
Bedford-Stuyvesant	Abc
Bensonhurst	Abc
Bergen Beach	Abc
Boerum Hill	Abc
Borough Park	Abc
Brighton Beach	Abc
Bronxdale	Abc
Brooklyn	Abc
Brooklyn Heights	Abc
Brooklyn Navy Yard	Abc
Brownsville	Abc
Bushwick	Abc
Canarsie	Abc

Context: Instructions

1 2 3 4

Drag the Price field to the Text Marks card.

Hint Rewatch Video Next

Level up your data visualization skills with DataCamp

Introduction to Tableau

[View Course](#)



The screenshot displays the DataCamp website interface. At the top, there is a navigation bar with the DataCamp logo, 'Home', 'Learn', 'Certification', 'Groups', 'DataLab', and a search bar. The sidebar on the left is divided into 'LEARN' (Tracks, Courses, Practice, Assessments, Tutorials) and 'APPLY' (Real World Projects, Code Alongs, Competitions, Popular Topics). The main content area features a 'SKILL TRACK' for 'Design in Power BI' with a 'Continue Track' button. Below this, it shows 'Power BI, Theory' (4 hours, 2 courses, 448 participants) and a 'TRACK COMPLETION' progress bar at 0%. The 'Track Description' states: 'This short track will take your dashboarding and reporting skills to the next level.' The 'COURSE' section lists '1 Dashboard Design Concepts' with a 0% completion bar. A description follows: 'Learn the skills needed to create impactful dashboards. Understand dashboard design fundamentals, visual analytics components, and dashboard types.' The 'Chapters' section includes: 'The Building Blocks' (650 XP), 'Visual Analytics Components' (1000 XP), 'Types of Dashboards' (900 XP), and 'Dashboard Success and Integration' (750 XP). The 'INSTRUCTORS' section lists Olga Scrivner (President, Scrivner Solutions Inc) and Maarten Van den Broeck (Senior Content Developer at DataCamp).

Level up your data visualization skills with DataCamp

User-Oriented Design in Power BI

[> View Course](#)



datacamp Home Learn Certification Groups DataLab Search

Progress
My Library
Leaderboard
Assignments

LEARN
Tracks
Courses
Practice
Assessments
Tutorials

APPLY
Real World Projects
Code Alongs
Competitions
Popular Topics **NEW**

Getting Started (3/4)

SKILL TRACK
Data Storytelling
Start Track

Theory 6 hours 4 courses 7,496 participants

TRACK COMPLETION
0%

Track Description
Discover the art of data storytelling. Transform raw information into memorable narratives.

COURSE
1 **Communicating Data Insights**
0%

Data-driven organizations consistently rely on insights to inspire action and drive change.

Chapters:

Communicating Information	750 XP
Effective Data Visuals	800 XP
Storytelling with Data	1050 XP

INSTRUCTORS

- Joe Franklin**
Associate Data Literacy and Essentials Manager, DataCamp
- Leandra Gonzalez**
Sr. Data & Applied Scientist, Microsoft
- Camilo Martinez**
Data analysis and public policy consultant

View Course Continue

Level up your data visualization skills with DataCamp

Data Storytelling Concepts

[> View Course](#)





What questions can I answer **for you?**

Additional Resources



[Take DataCamp's Data Maturity Assessment](#)



[Learn more about DataCamp for Business](#)



[WHITE PAPER: Your Organization's Guide to Data Maturity](#)



[ON-DEMAND: Storytelling for more impactful data science](#)



[Register for one of our upcoming webinars](#)



[ON-DEMAND: How Data Governance Enables Scalable Data Science](#)





Thank you

Adel Nehme
VP of Media
adel@datacamp.com