



DATA VISUALIZATION

Frank Greene
CT Dept of Consumer Protection



WHY DATA VISUALIZATION?

- Simplifies Complex Data.
- Reveals Patterns and Trends.
- Aids in Decision Making.
- Enhances Communication.

NO
HO
STO
NU

WIKIPEDIAThe Free Encyclopedia

三

Search Wikipedia

Search

5 languages

Contentshide

(Top)

Misleading graph methods

Measuring distortion

Usage in finance and corporate reports

Academia

See also

References

Further reading

External links

Misleading graph

ArticleTalk

ReadEditView historyTools

From Wikipedia, the free encyclopedia

In statistics, a **misleading graph**, also known as a **distorted graph**, is a **graph** that misrepresents data, constituting a *misuse of statistics* and with the result that an incorrect conclusion may be derived from it.

Graphs may be misleading by being excessively complex or poorly constructed. Even when constructed to display the characteristics of their data accurately, graphs can be subject to different interpretations, or unintended kinds of data can seemingly and ultimately erroneously be derived.^[1]

Misleading graphs may be created intentionally to hinder the proper interpretation of data or accidentally due to unfamiliarity with **graphing software**, misinterpretation of data, or because data cannot be accurately conveyed. Misleading graphs are often used in **false advertising**. One of the first authors to write about biasing graphs was **Darrell Huff**, publisher of the 1954 book *How to Lie with Statistics*.

The field of **data visualization** describes ways to present information that avoids creating misleading graphs.

Misleading graph methods

[edit]

[A misleading graph] is vastly more effective, however, because it contains no adjectives or adverbs to spoil the illusion of objectivity, there's nothing anyone can pin on you.

—*How to Lie with Statistics* (1954)^[2]

There are numerous ways in which a misleading graph may be constructed.^[3]

Excessive usage

[edit]

The use of graphs where they are not needed can lead to unnecessary confusion/interpretation.^[4] Generally, the more explanation a graph needs, the less the graph itself is needed.^[4] Graphs do not always convey information better than tables.^[5]

Biased labeling

[edit]

The use of biased or **loaded words** in the graph's title, axis labels, or caption may inappropriately prime the reader.^{[4][6]}

50

48

46

44

42

40

38

36

2010

2011

Salmerek

50

40

30

20

10

0

2010

2011

Salmerek

Example of a truncated (left) vs full-scale graph (right), using the same data

Part of a series on **Statistics**

Data and information visualization

Major dimensions

Exploratory data analysis · Information design · Interactive data visualization · Descriptive statistics · Inferential statistics · Statistical graphics · Plot · Data analysis · Infographic · Data science

Important figures

Tamara Munzner · Ben Shneiderman · John Tukey · Edward Tufte · Simon Wardley · Hans Rosling · David McCandless · Kim Albrecht · Alexander Osterwalder · Ed Hawkins · Hadley Wickham · Leland Wilkinson · Mike Bostock · Jeffrey Heer · Ihab Ilyas

Information graphic types

Line chart · Bar chart · Histogram · Scatter plot · Box plot · Pareto chart · Pie chart · Area chart · Tree map · Bubble chart · Stripe graphic · Control chart · Run chart · Stem-and-leaf display ·



Read Edit View history Tools ▾

Article [Talk](#)

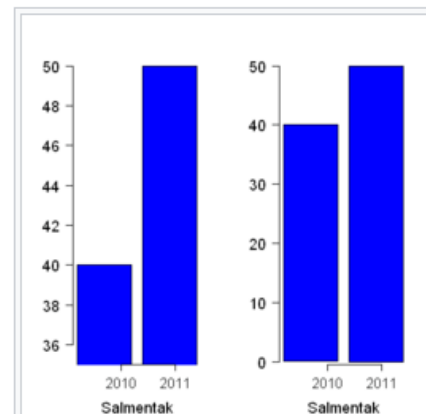
From Wikipedia, the free encyclopedia

In **statistics**, a **misleading graph**, also known as a **distorted graph**, is a **graph** that misrepresents **data**, constituting a **misuse of statistics** and with the result that an incorrect conclusion may be derived from it.

Graphs may be misleading by being excessively complex or poorly constructed. Even when constructed to display the characteristics of their data accurately, graphs can be subject to different interpretations, or unintended kinds of data can seemingly and ultimately erroneously be derived.^[1]

Misleading graphs may be created intentionally to hinder the proper interpretation of data or accidentally due to unfamiliarity with [graphing software](#), misinterpretation of data, or because data cannot be accurately conveyed. Misleading graphs are often used in [false advertising](#). One of the first authors to write about misleading graphs was [Darrell Huff](#), publisher of the 1954 book *[How to Lie with Statistics](#)*.

The field of **data visualization** describes ways to present information that avoids creating misleading graphs.



Example of a truncated (left) vs full-scale graph (right), using the same data

Misleading graph methods [\[edit \]](#)

[A misleading graph] is vastly more effective, however, because it contains no adjectives or adverbs to spoil the illusion of objectivity, there's nothing anyone can pin on you.

—*How to Lie with Statistics* (1954)^[2]

There are numerous ways in which a misleading graph may be constructed.^[3]

Excessive usage [\[edit \]](#)

The use of graphs where they are not needed can lead to unnecessary confusion/interpretation.^[4] Generally, the more explanation a graph needs, the less the graph itself is needed.^[4] Graphs do not always convey information better than tables.^[5]

Biased labeling [\[edit \]](#)

The use of biased or **loaded words** in the graph's title, axis labels, or caption may inappropriately **prime** the reader.^{[4][6]}

Part of a series on [Statistics](#)

Data and information visualization

Major dimensions

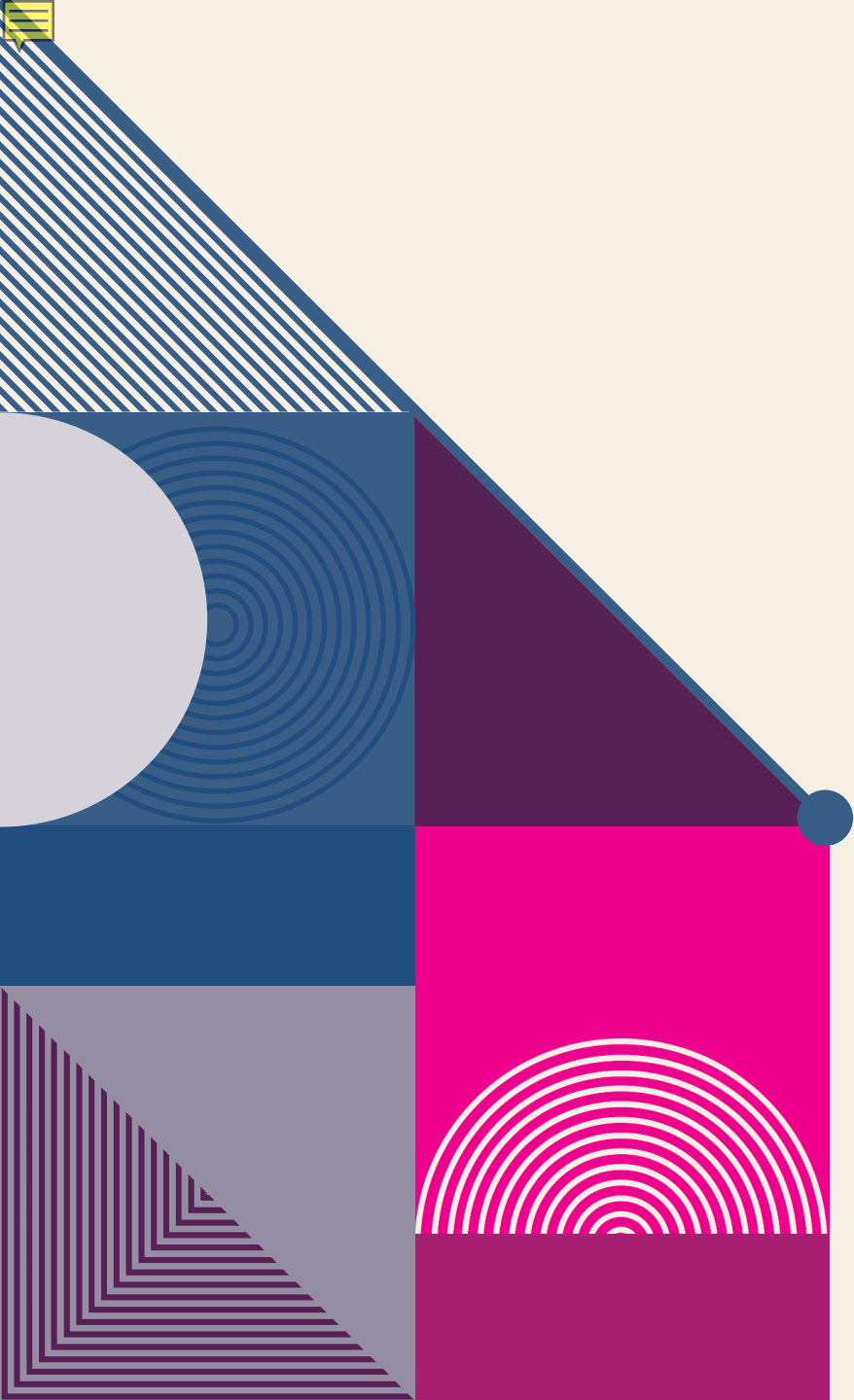
Exploratory data analysis · Information design ·
Interactive data visualization · Descriptive statistics ·
Inferential statistics · Statistical graphics · Plot ·
Data analysis · Infographic · Data science

Important figures

Tamara Munzner · Ben Shneiderman · John Tukey ·
Edward Tufte · Simon Wardley · Hans Rosling ·
David McCandless · Kim Albrecht ·
Alexander Osterwalder · Ed Hawkins ·
Hadley Wickham · Leland Wilkinson · Mike Bostock ·
Jeffrey Heer · Ihab Ilyas

Information graphic types

Line chart • Bar chart • Histogram • Scatter plot •
Box plot • Pareto chart • Pie chart • Area chart •
Tree map • Bubble chart • Stripe graphic •
Control chart • Run chart • Stem-and-leaf display •



UNDERSTAND THE CONTEXT

Who is your audience?

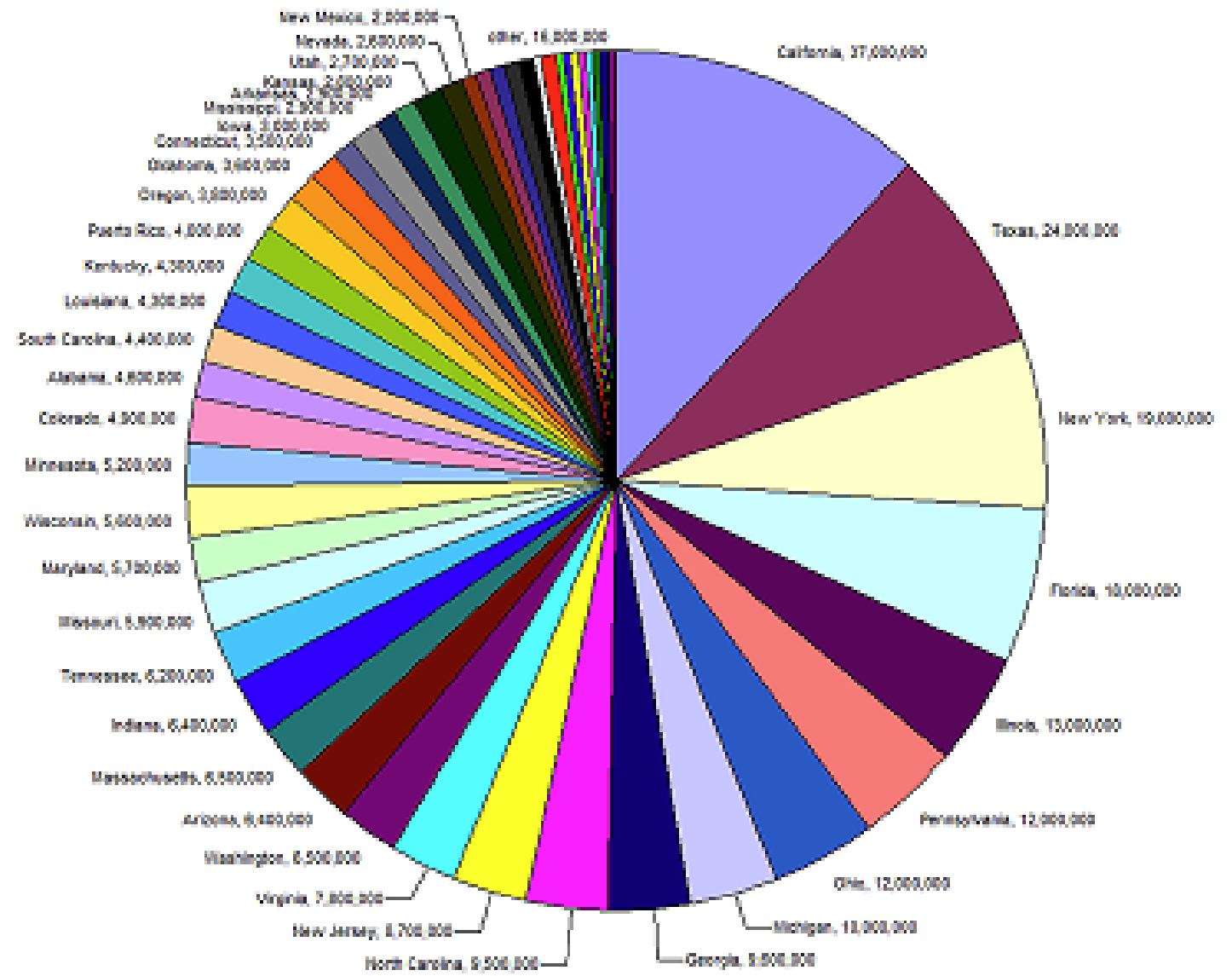
What is your goal for that audience?



AN EFFECTIVE VISUAL

What is the best way to show the data you want to communicate?

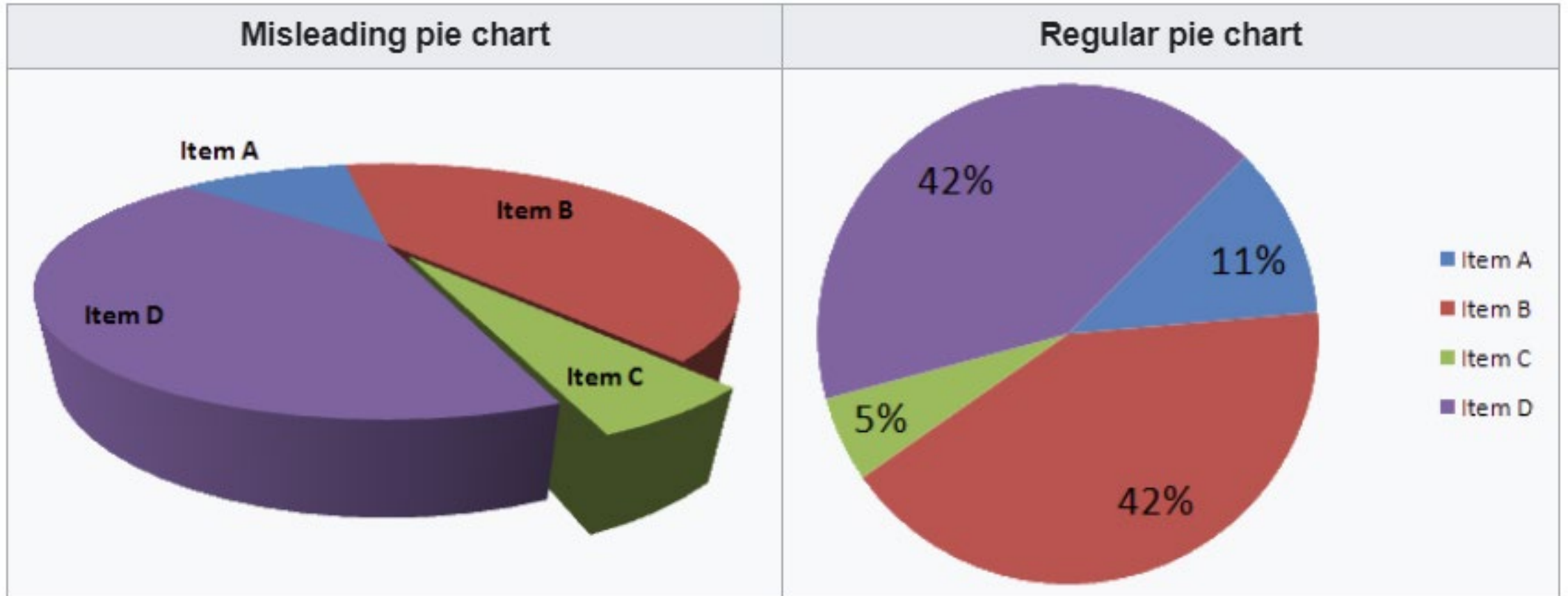
- Clear Message
- Appropriate Chart Type
- Accurate Representation
- Context and Annotations
- Engagement and Aesthetics
- Accessibility



European Environmental Agency



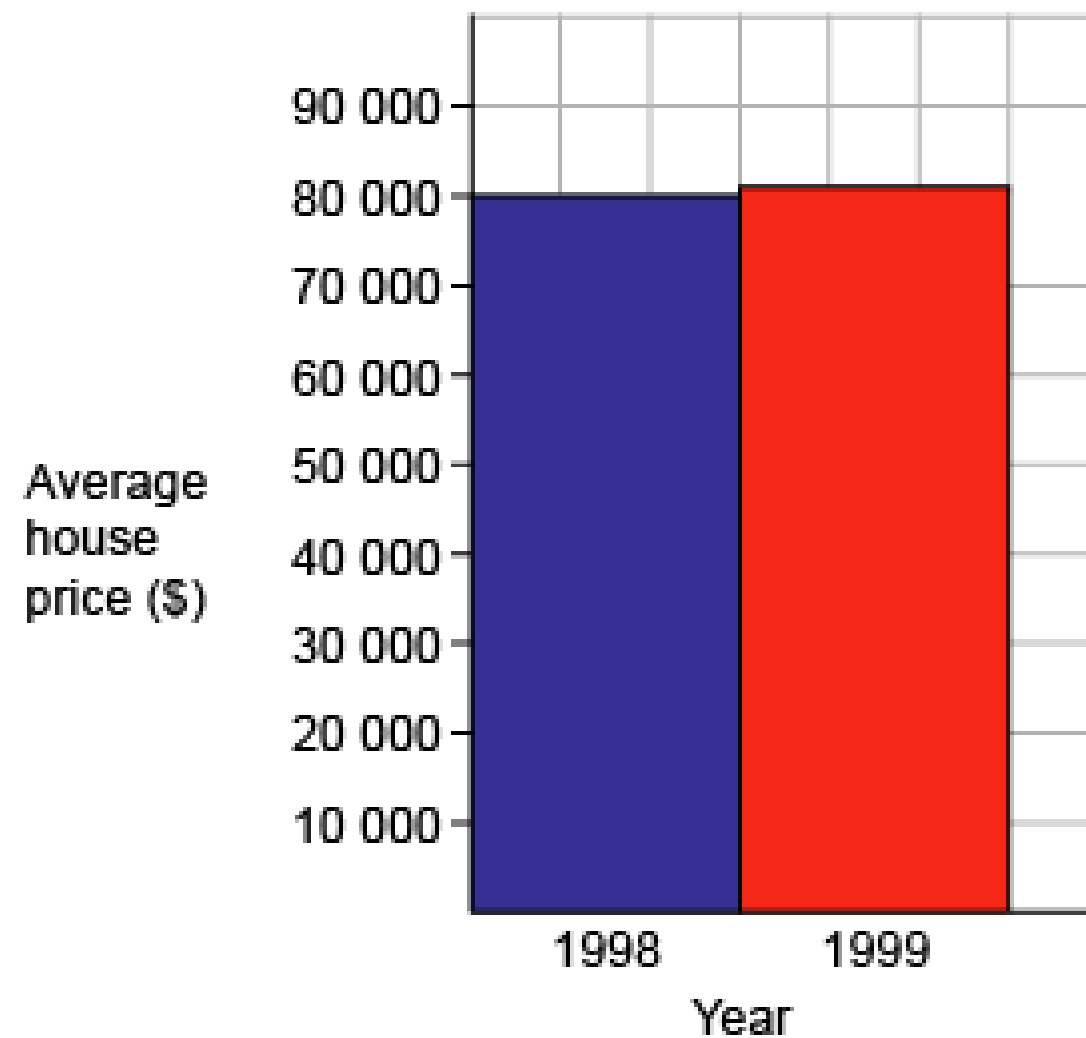
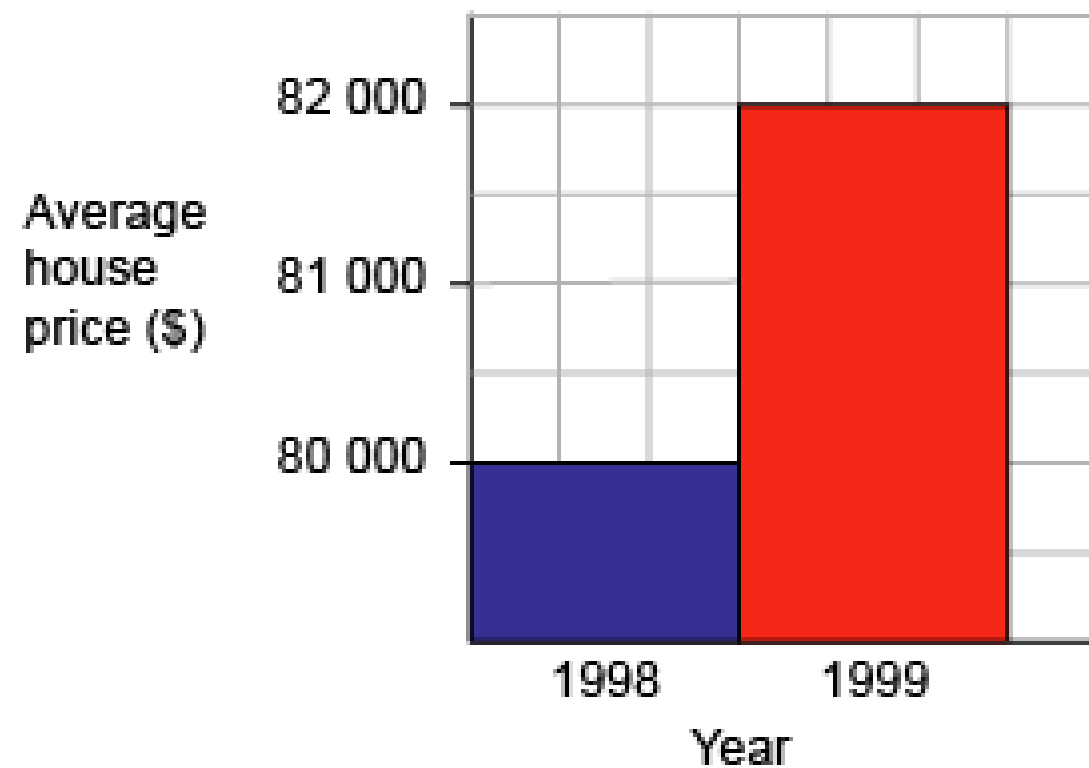
Comparison of pie charts



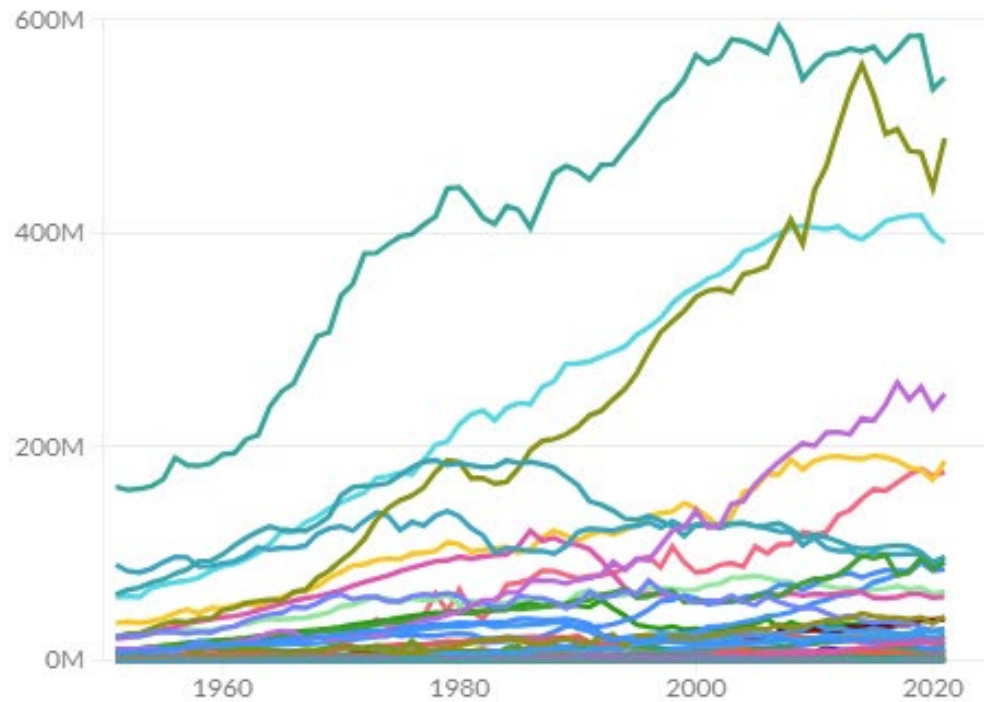
3D Pie Charts

House prices

"Massive increase in house prices this year"

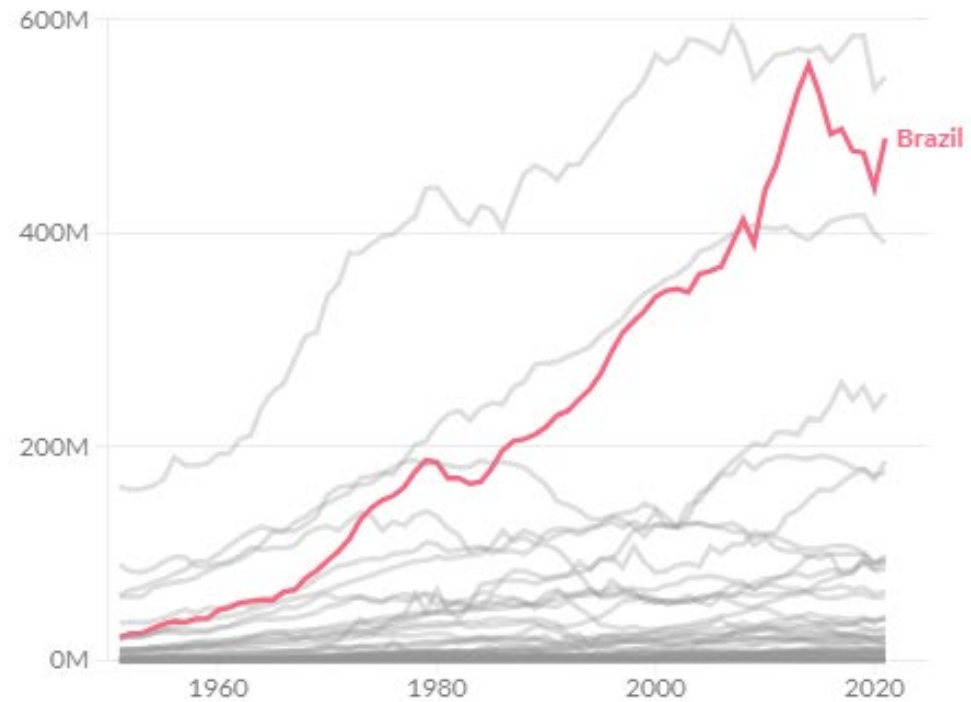


Annual CO2 emissions



Source: [Our World in Data](#) • Created with the Flourish [Line Bar Pie](#) template

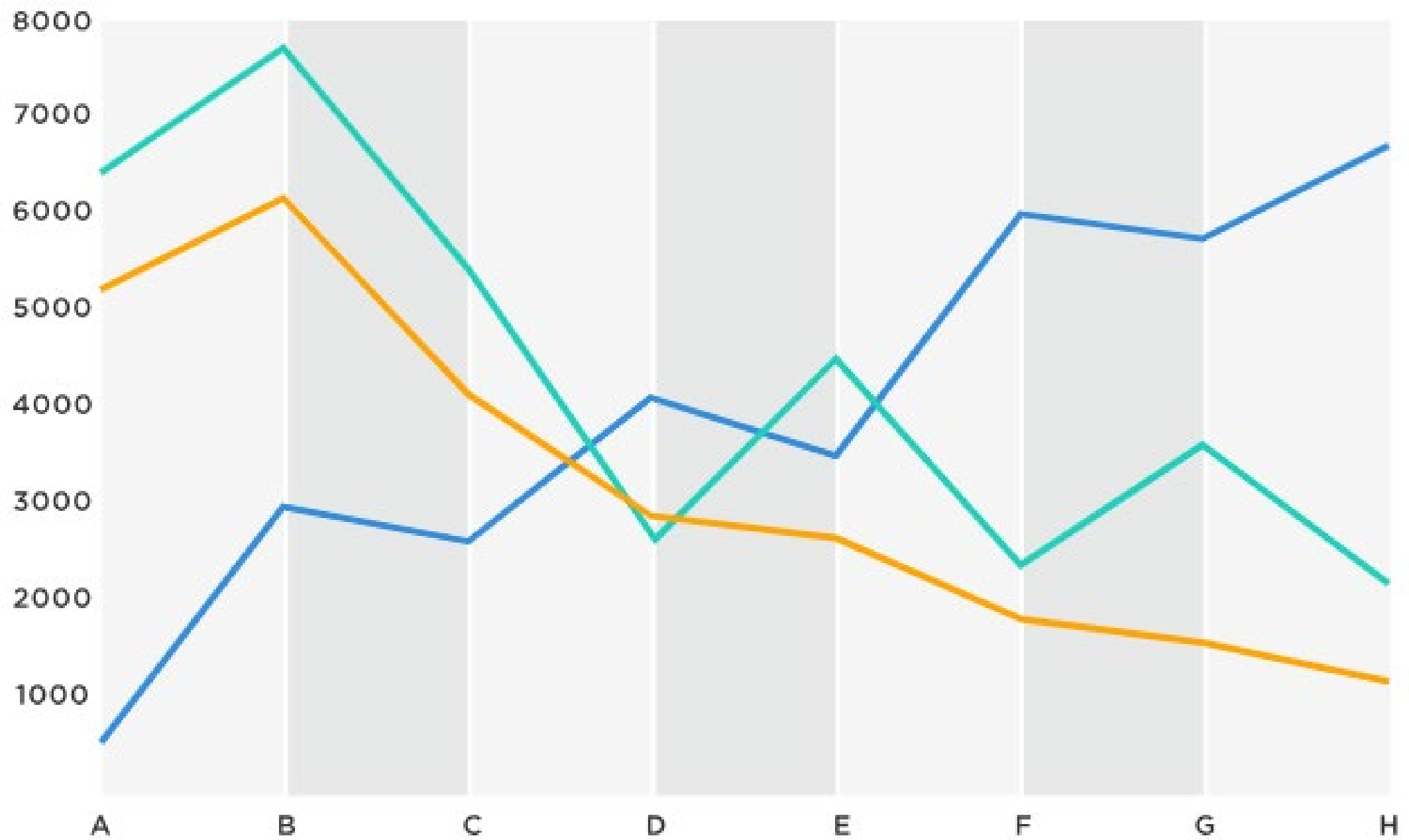
Annual CO2 emissions



Source: [Our World in Data](#) • Created with the Flourish [Line Bar Pie](#) template

Skip the rainbow and put the focus on the vital bits. Instead of assigning specific colors to each category, highlight the most important elements, while grey is a good option for less important elements.

This way, the essential points take center stage and guide the user's gaze.







COMMON TECHNIQUES OF MISLEADING GRAPHS

Common Techniques of Misleading Graphs:

- Truncated Data: Starting at a value other than zero to exaggerate differences.
- Misleading Scales: Using unlabeled, inconsistent intervals or scales.
- Omitted Data: Leaving out relevant data points.
- Cherry-Picking Data: Only showing data that supports a specific conclusion.
- Manipulating Data Representation: Using visuals that distort the data, such as 3D effects or inappropriate graph types.

EFFECTIVE GRAPHICS

91%

Simple text



Scatterplot

	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%

Table

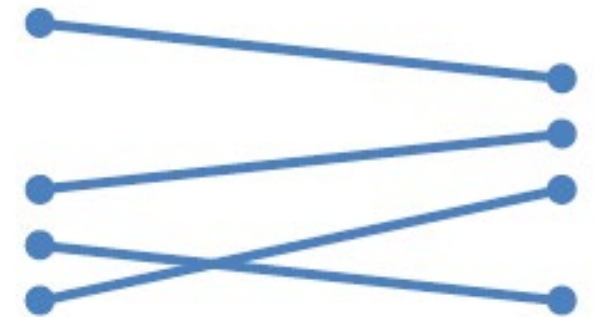


Line

EFFECTIVE GRAPHICS

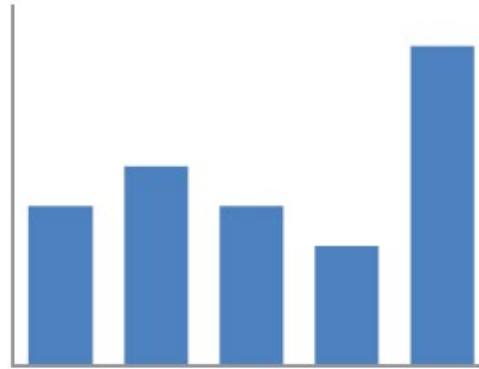
	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

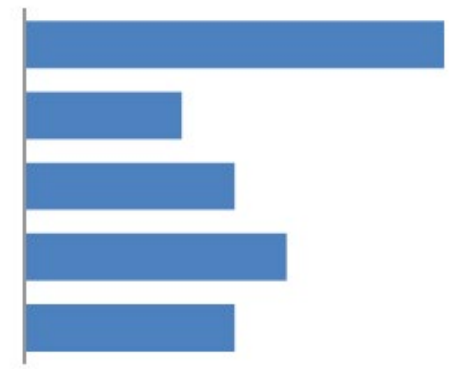


Slopegraph

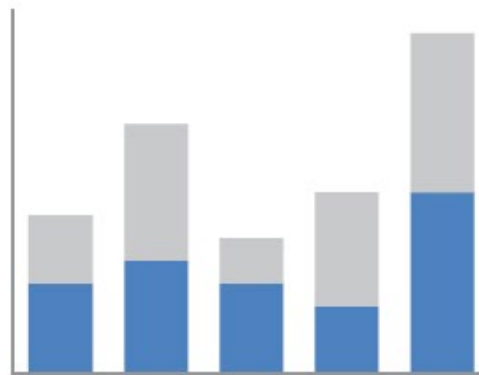
EFFECTIVE GRAPHIC



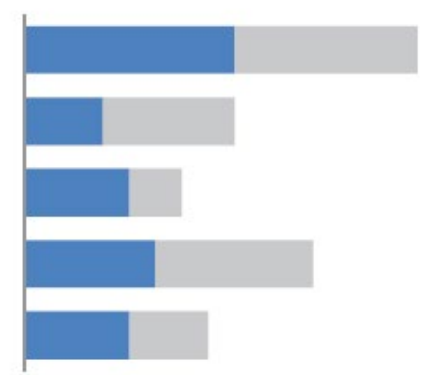
Vertical bar



Horizontal bar

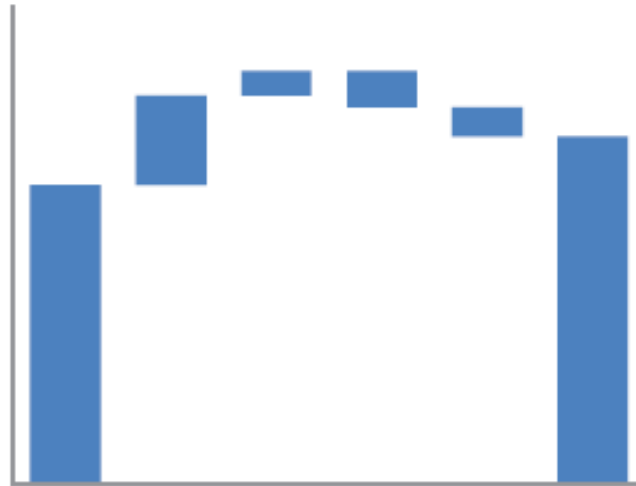


Stacked vertical bar

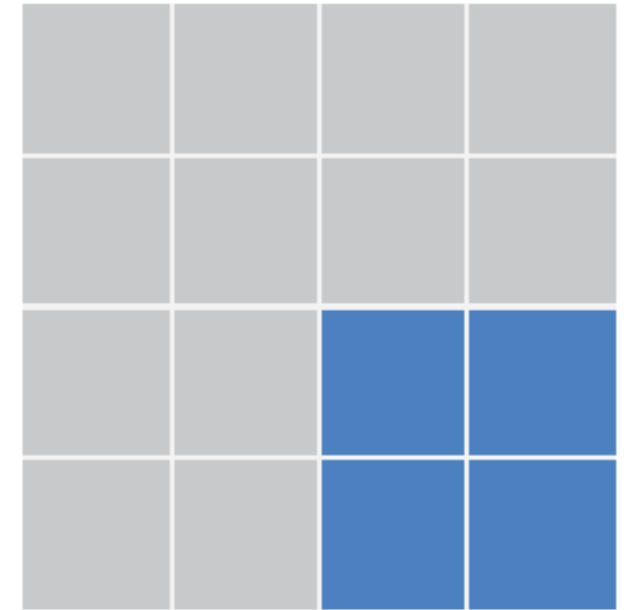


Stacked horizontal bar

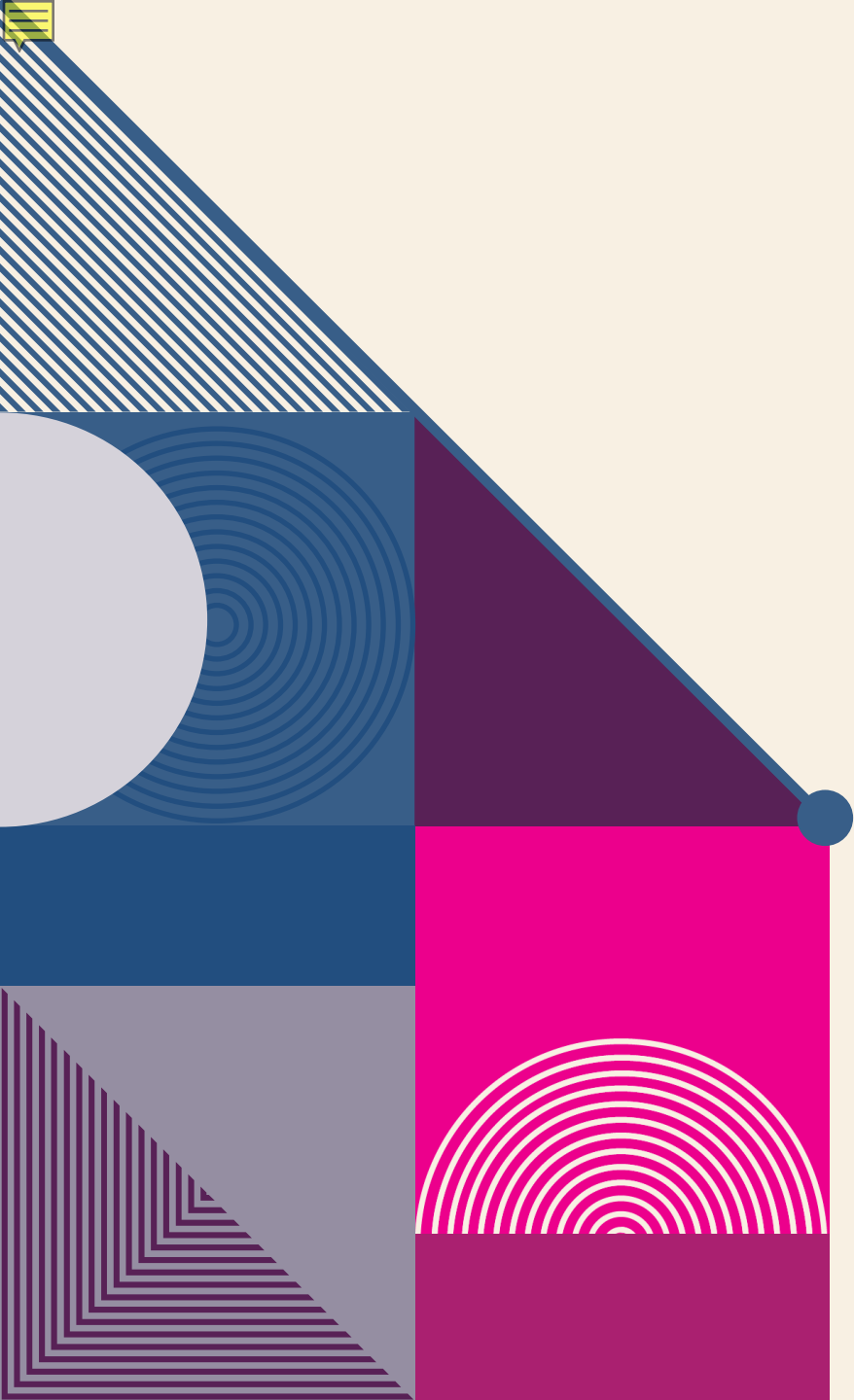
EFFECTIVE GRAPHIC



Waterfall

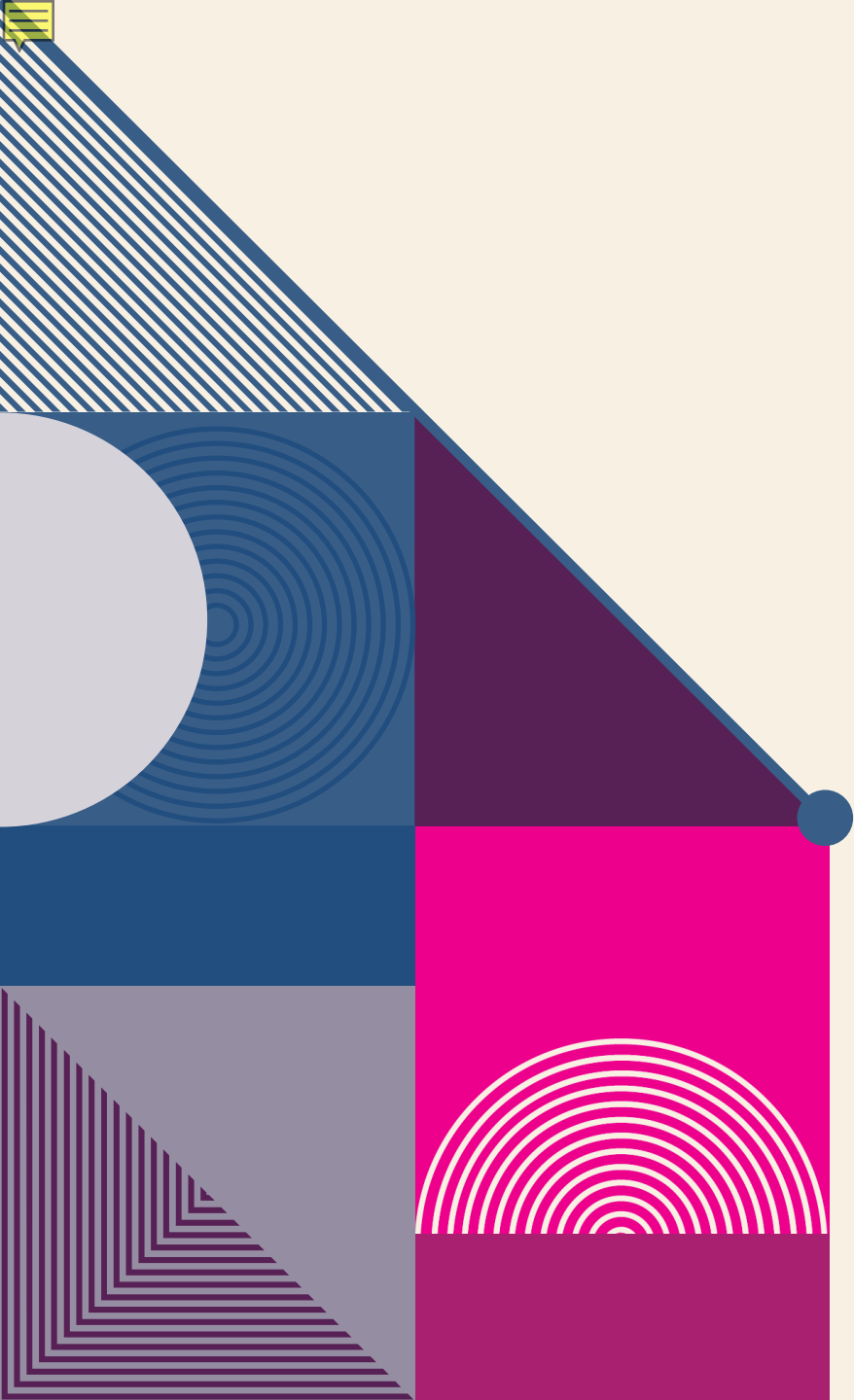


Square area



CLUTTER IS YOUR ENEMY

What is the best way to show the data you want to communicate?



GESTALT PRINCIPLES OF VISUAL PERCEPTION

Proximity
Similarity
Continuity
Closure
Connectedness

GESTALT PRINCIPLES OF VISUAL PERCEPTION

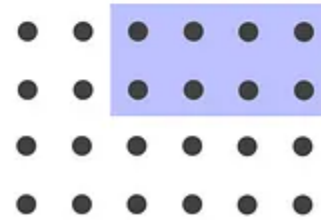
PROXIMITY



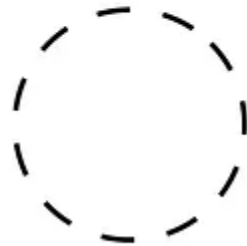
SIMILARITY



ENCLOSURE



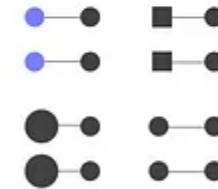
CLOSURE

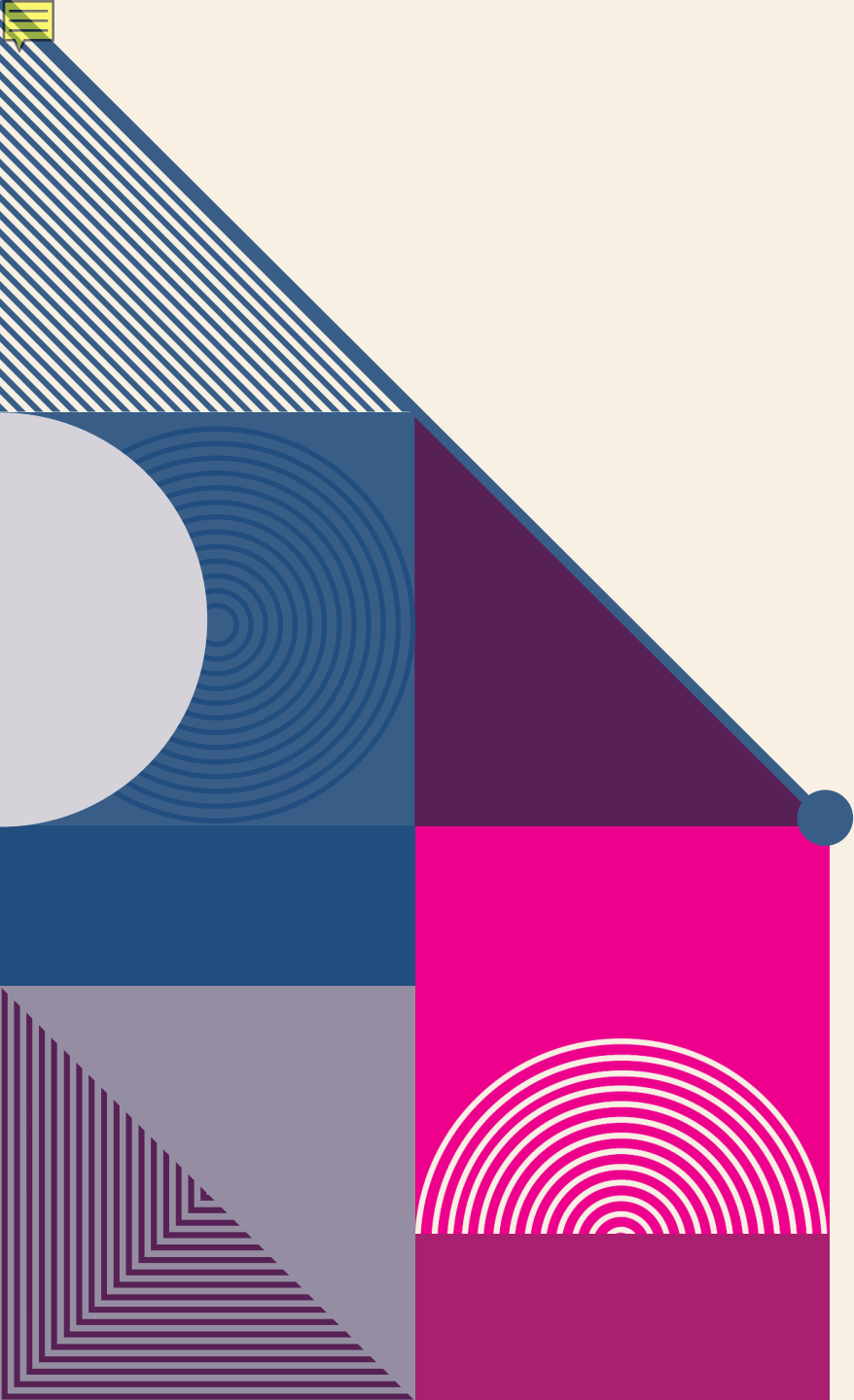


CONTINUITY



CONNECTION





FOCUS YOUR AUDIENCE'S ATTENTION

Do not overtax the audience's ability to comprehend. Using design and techniques to highlight the most effective and efficient way to display data.

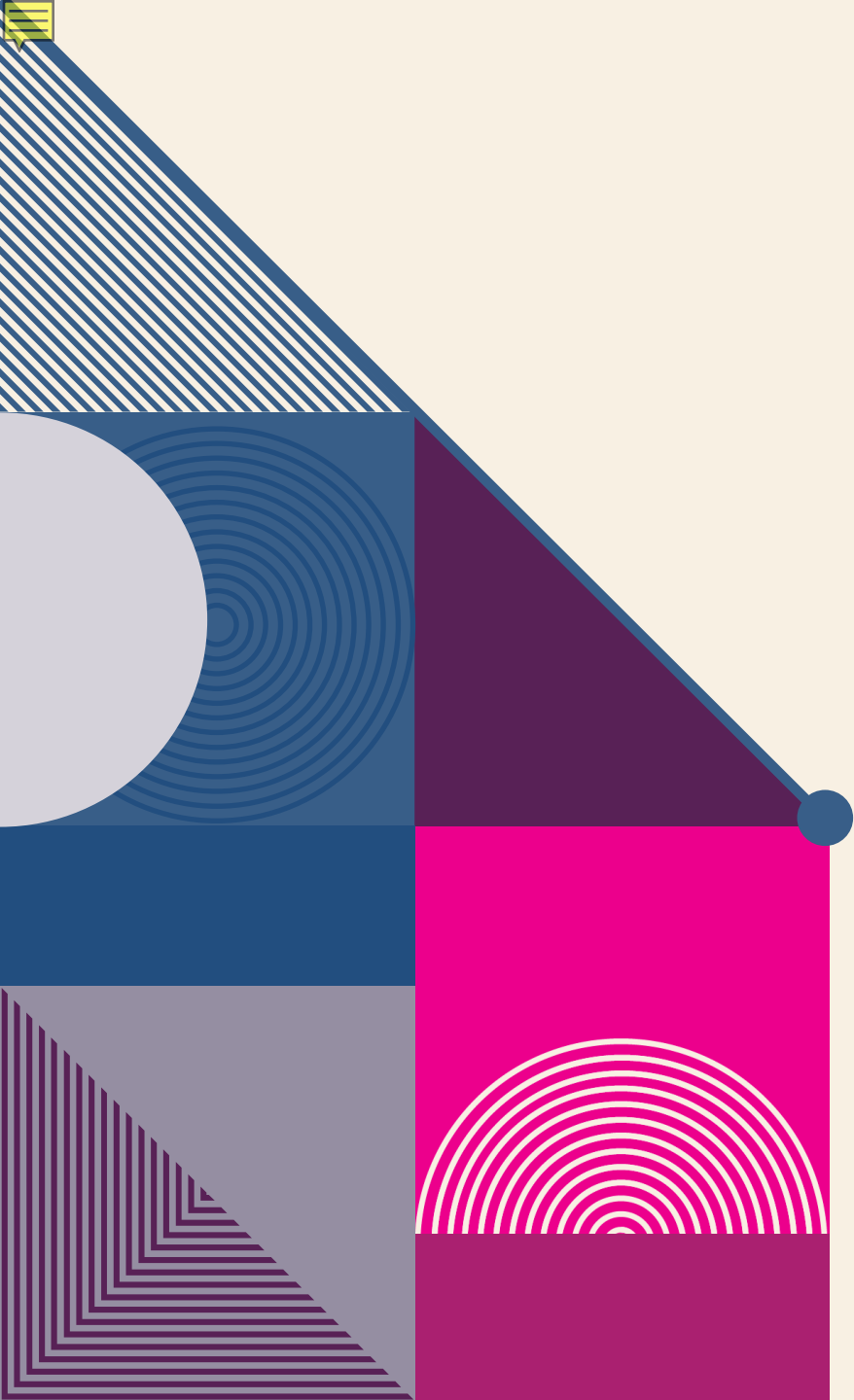


Graphic Exercise - Count the 6's

56585462247799
25570595542246
46546465465486
19473104384750
31595569554325



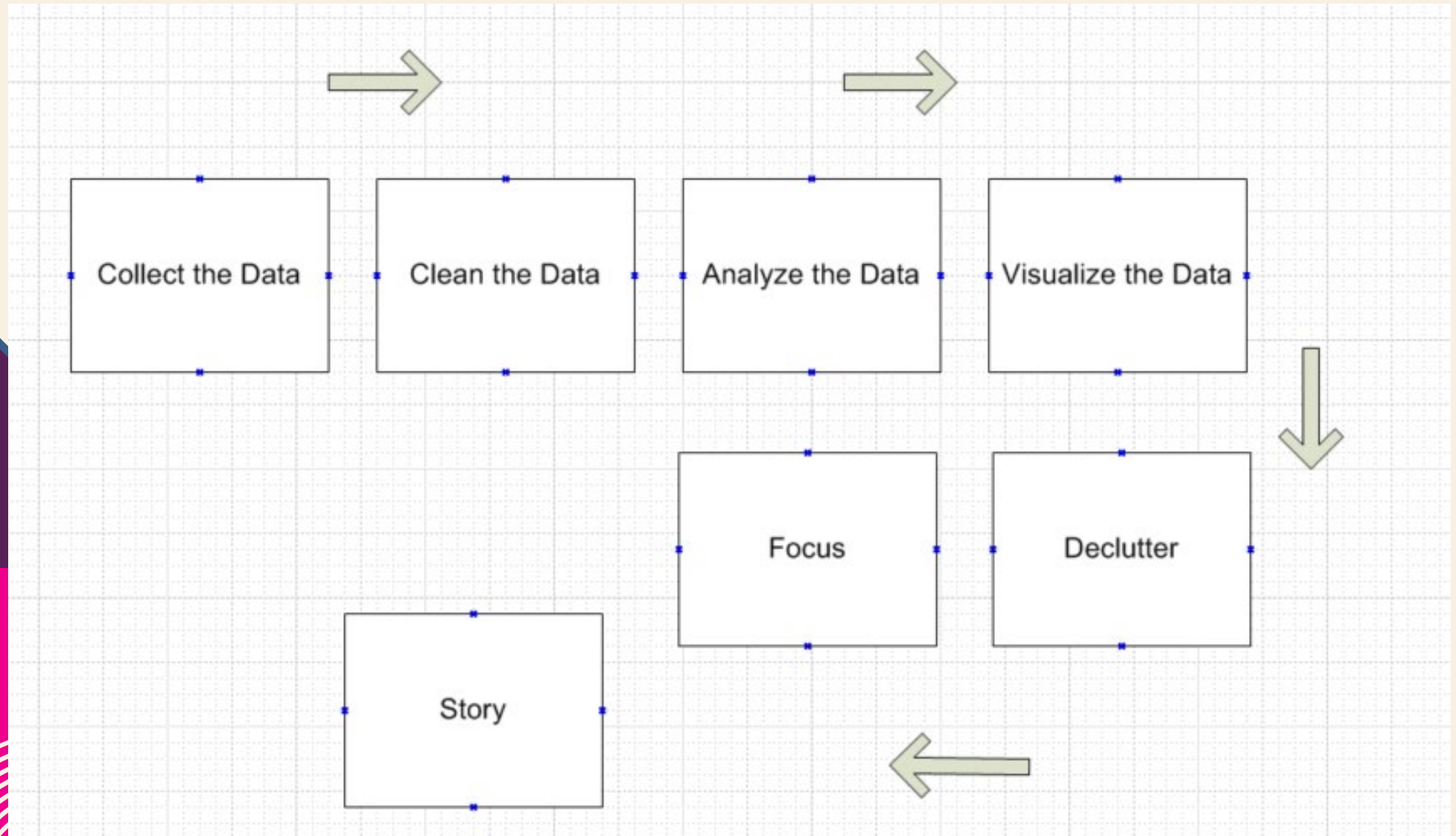
56585462247799
25570595542246
46546465465486
19473104384750
31595569554325

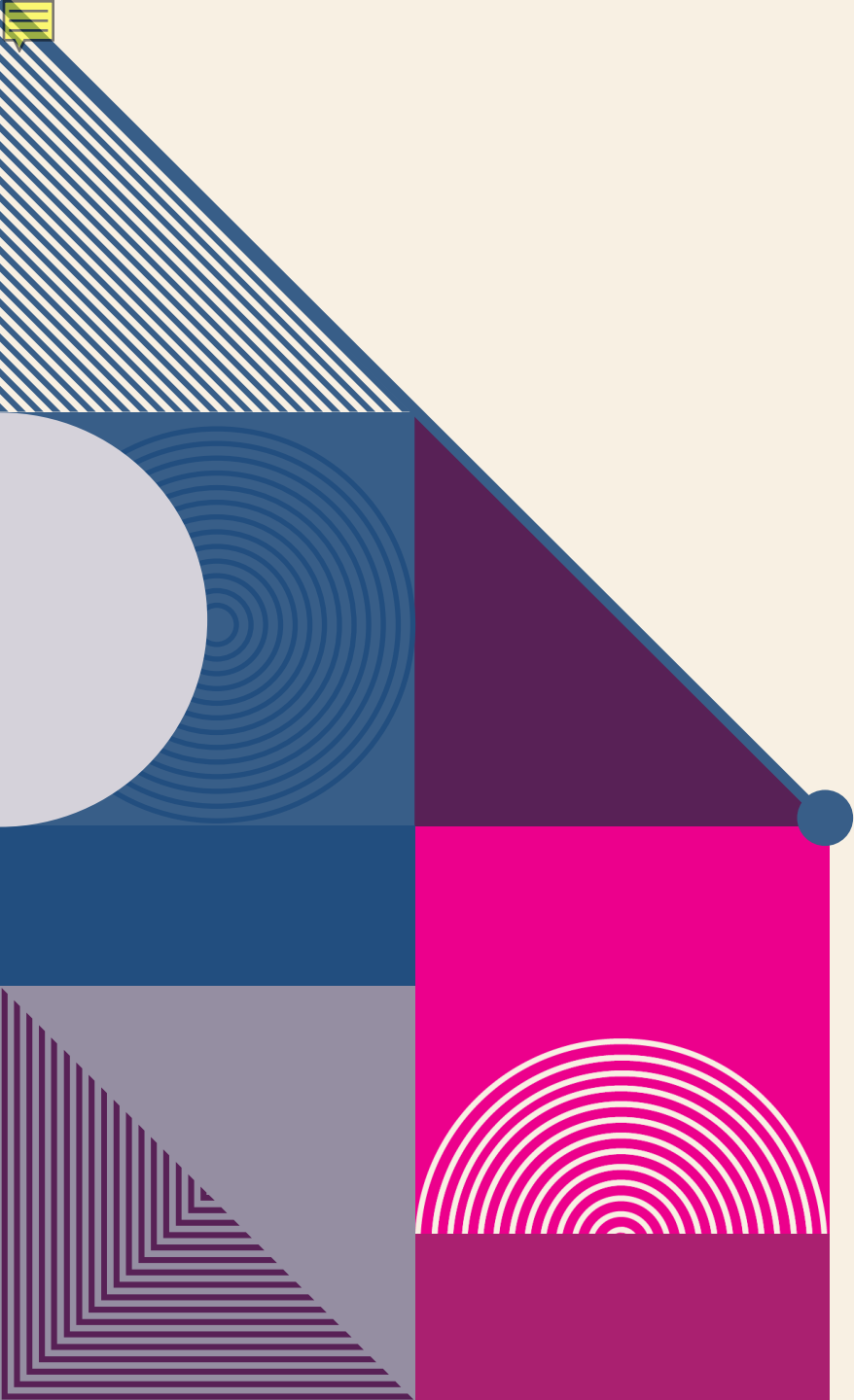


THINK LIKE A DESIGNER

Form follows function

- What do we want our audience to do with the data (function)?
- Creating a visualization (form) that provides a pathway.





CHARTJUNK

The interior decoration of graphics generates a lot of ink that does not tell the viewer anything new. The purpose of decoration varies—to make the graphic appear more scientific and precise, to enliven the display, to give the designer an opportunity to exercise artistic skills. Regardless of its cause, it is all non-data-ink or redundant data-ink, and it is often chartjunk. – Edward Tufte, “The Visual Display of Quantitative Data.”



LIMITED BIBLIOGRAPHY

How to Lie with Statistics, Darrell Huff

PDQ Statistics Third Edition, Geoffrey R. Norman, David
L. Streiner

Storytelling with Data, Cole Nussbaumer Knaflic

The Functional Art, Alberto Cairo

Edward Tufte, "The Visual Display of Quantitative Data."

Minard's graphic on Napoleon's disastrous march and retreat from Moscow

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée 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 dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fozensac, de Chambray et le journal inédit 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 Davoust qui avaient été détachés sur Minsk et Mohilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

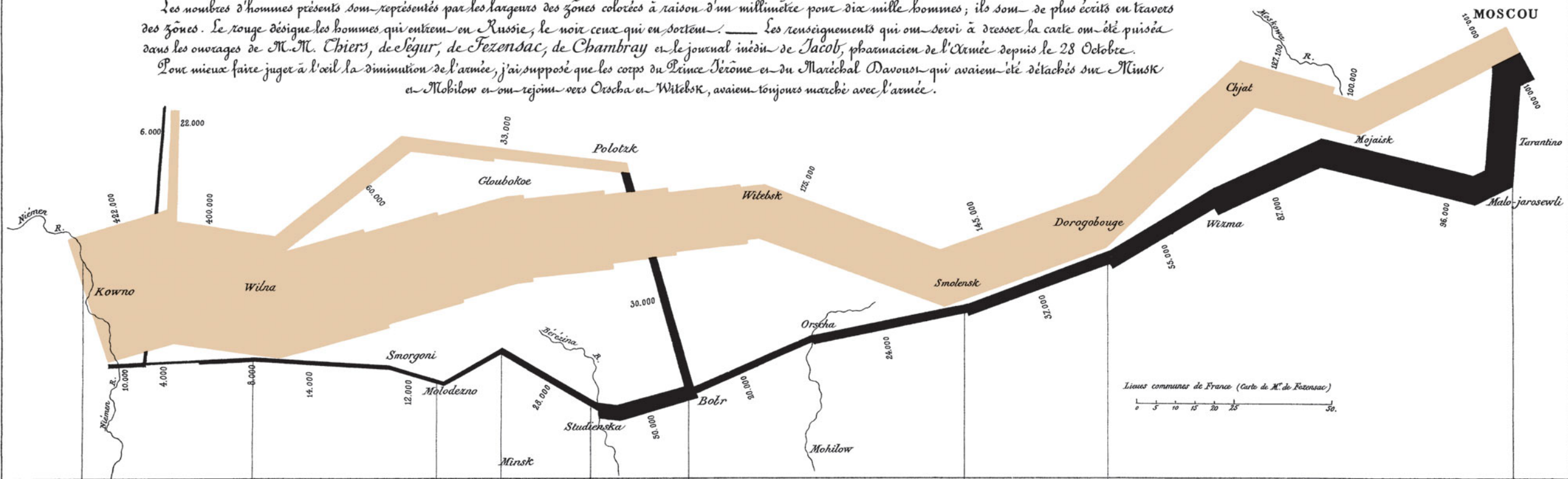
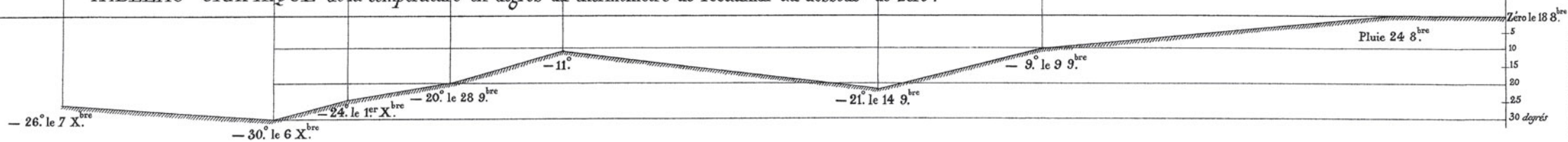


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Les Cosaques passent au galop le Niémen gelé.



THANK YOU

Frank Greene

860.713.6168

Frank.greene@ct.gov