Thegradientboost

CREATING STRONGER LEARNERS

Your data is only as good as your ability to understand and communicate it effectively to your target audience. This is a major reason why choosing the right visualization is essential.

If your data is misrepresented or presented ineffectively, key insights and understanding are lost, which hurts your message. In the next few slides we will be presenting tips and tricks to help you choose the right presentation for your data

Thegradientboost
CREATING STRONGER LEARNERS

# FINDING THE STORY IN YOUR DATA

Information can be visualized in a number of ways, each of which can provide a specific insight. When you start to work with your data, it's important to identify and understand the story you are trying to tell and the relationship you are looking to show. Knowing this information will help you select the proper visualization to best deliver your message.

When analyzing data, search for patterns or interesting insights that can be a good starting place for finding your story, such as:

**TRENDS**

**Example:**
**Ice cream sales over time**

**CORRELATIONS**

**Example:**
**Ice cream sales vs. temperature**

**OUTLIERS**

**Example:**
**Ice cream sales in an unusual region**

Thegradientboost
CREATING STRONGER LEARNERS

# ASK THE RIGHT QUESTIONS

In most cases you will be creating visualizations for an audience. This audience may have pre-existing questions that need to be addressed and answered by your visualization to guide decision making. Keep these questions in mind when creating your visualizations:

- What questions do you want to answer?
- What is the problem you are trying to solve?
- What decisions are you trying to make?
- What outcomes are you hoping for?
- What story do you want to tell?
- What are the actionable insights for your audience?

Thegradientboost
CREATING STRONGER LEARNERS

# KNOW YOUR DATA

Before understanding visualizations, you must understand the types of data that can be visualized and their relationships to each other. Here are some of the most common you are likely to encounter.

## Different Data Types

**QUANTITATIVE**
Data that can be counted or measured; all values are numerical.

**DISCRETE**
Numerical data that has a finite number of possible values. Example: Number of employees in the office.

**CONTINUOUS**
Data that is measured and has a value within a range. Example: Rainfall in a year.

**CATEGORICAL**
Data that can be sorted according to group or category. Example: Types of products sold.

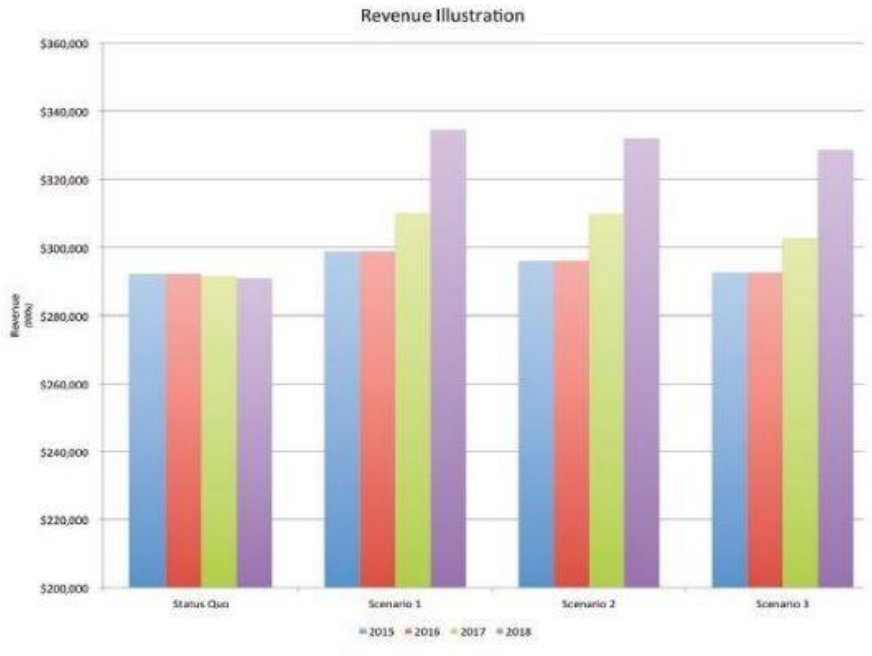
Thegradientboost

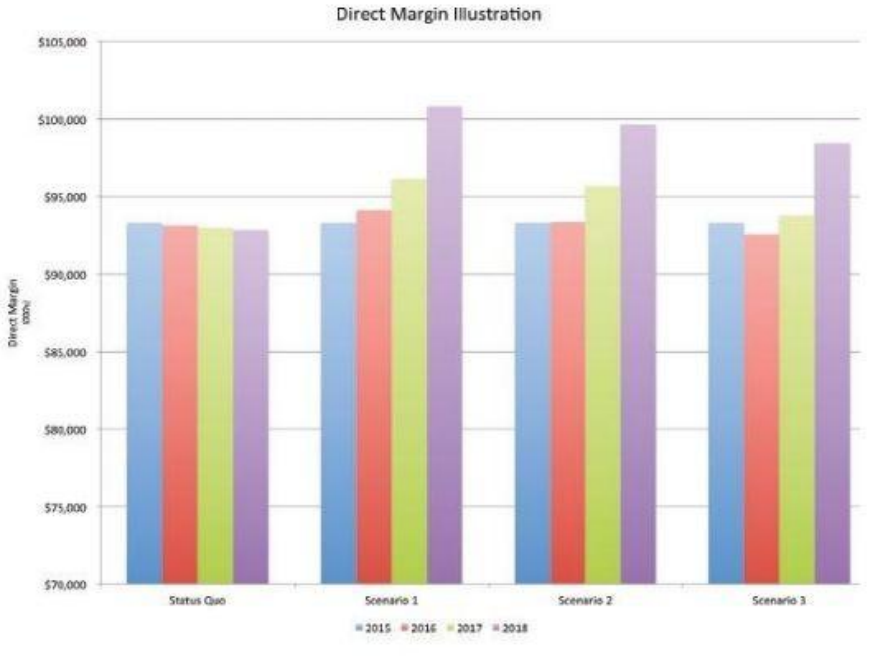
CREATING STRONGER LEARNERS

# What does bad Data Visualization look like?

Thegradientboost
CREATING STRONGER LEARNERS

# Consider this chart hideous?



Thegradientboost
CREATING STRONGER LEARNERS

# Why?

Lack of clear title and insightful sub-heading leaves the audience uninformed

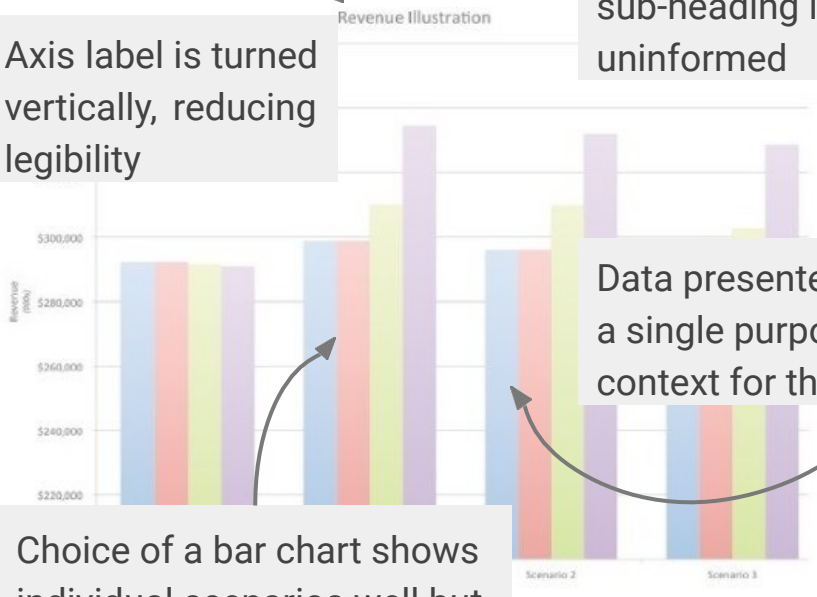
Axis label is turned vertically, reducing legibility

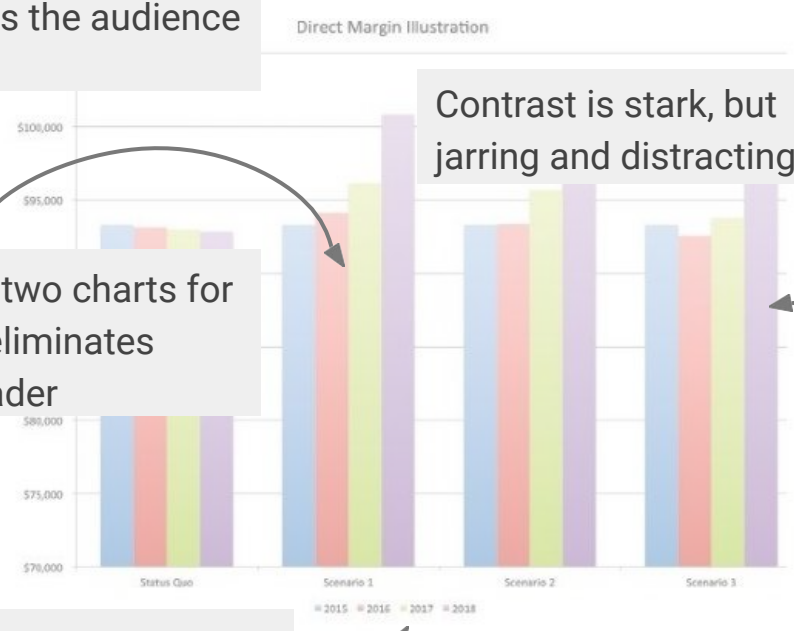
Contrast is stark, but jarring and distracting

Data presented in two charts for a single purpose eliminates context for the reader

Choice of a bar chart shows individual scenarios well but makes it challenge to compare across scenarios

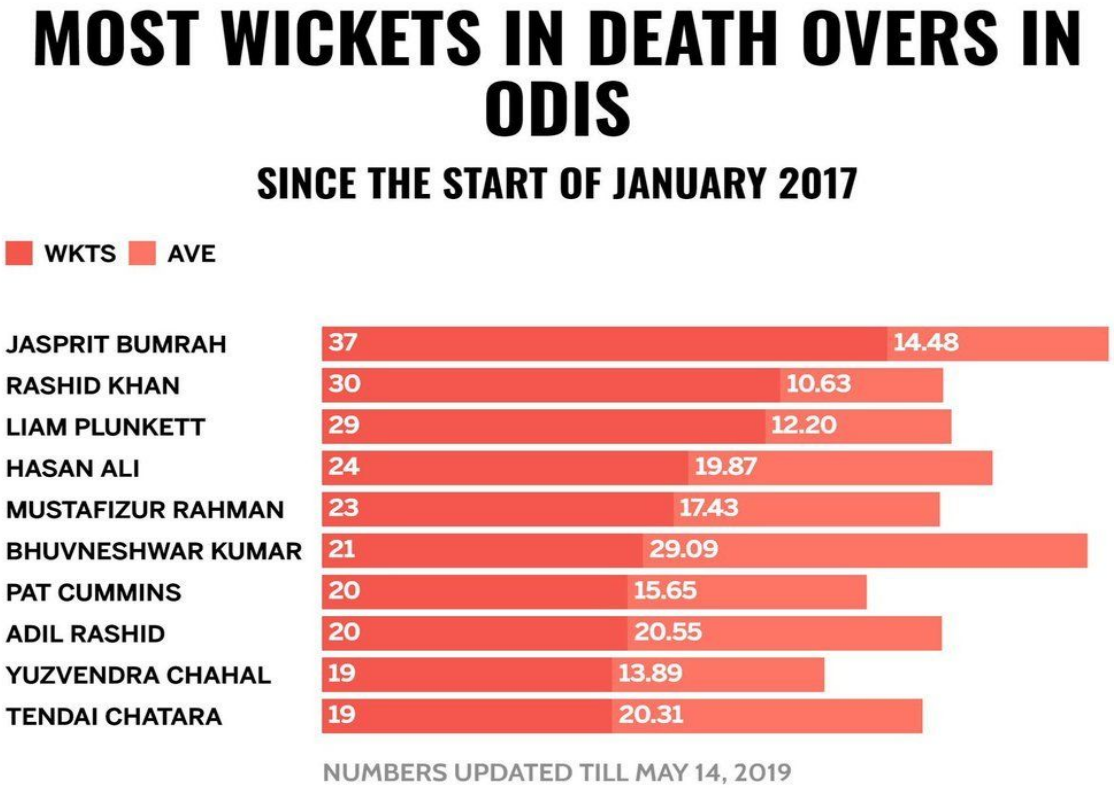
A legend far from the data makes understanding difficult



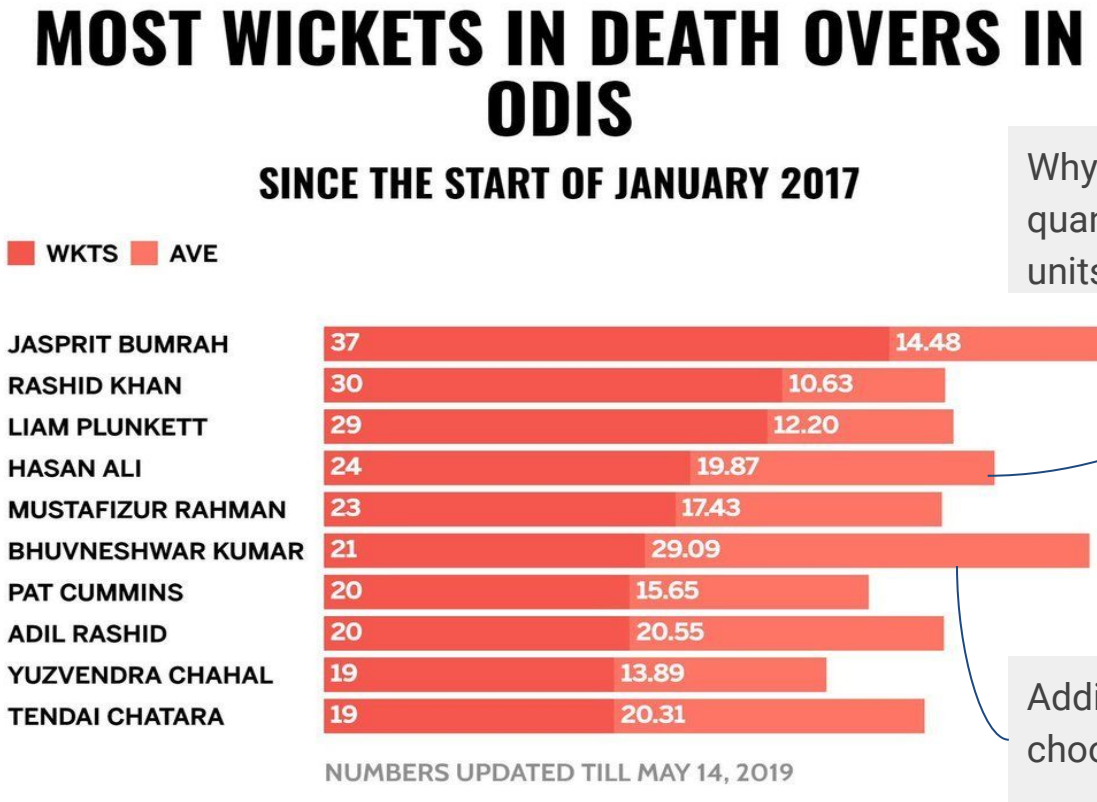
Thegradientboost
CREATING STRONGER LEARNERS

# Consider this chart hideous?



Thegradientboost
CREATING STRONGER LEARNERS

# Why?




Why would you stack bars with quantities that have different units?

Additionally, why would you choose near-identical colours

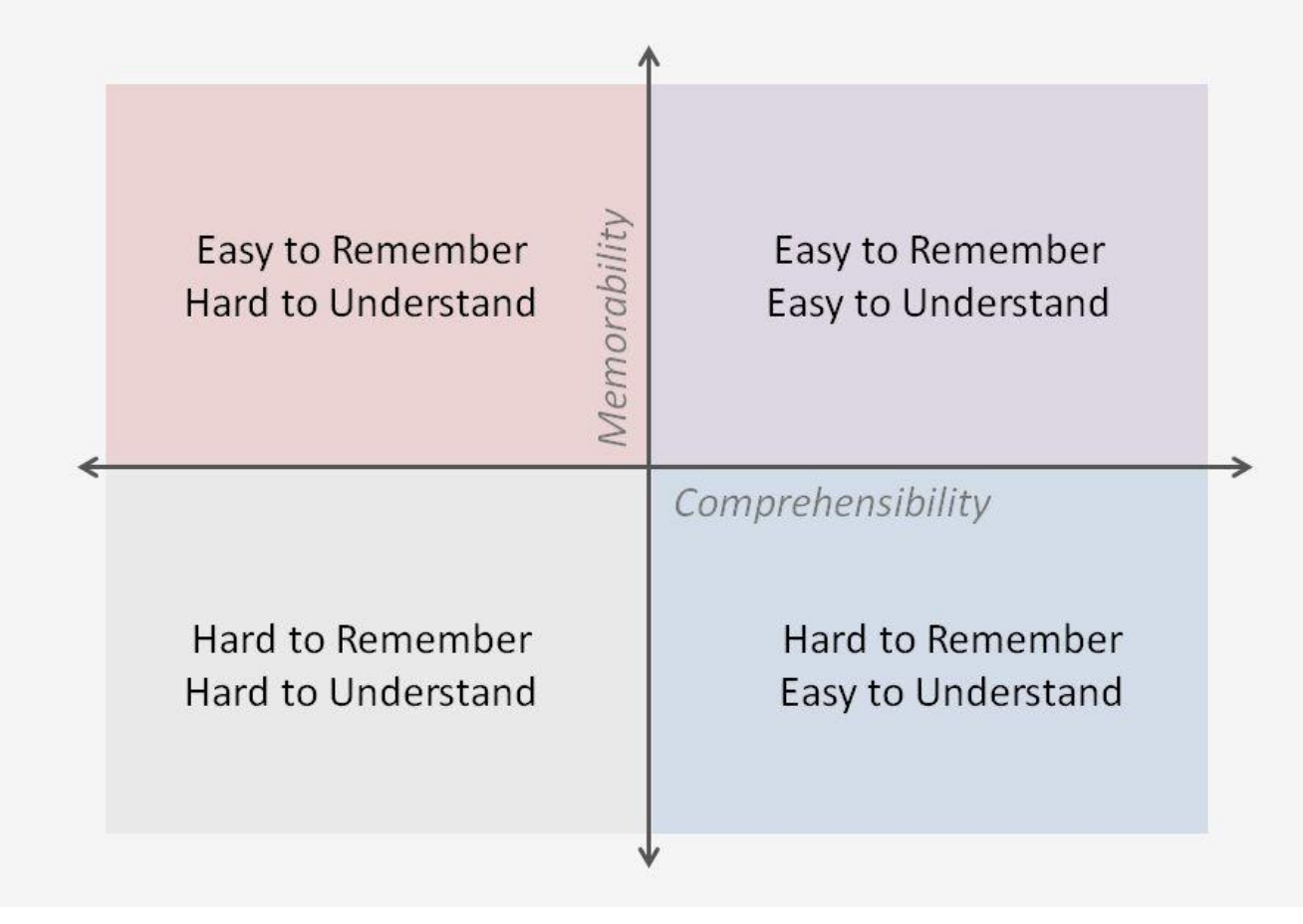
Thegradientboost
CREATING STRONGER LEARNERS

# Each step is necessary in producing high quality data visualization

## Minutes Spent At Each Task

| Prep | Talk + Listen | Sketch + Design | Prototype + Improve |
|---|---|---|---|
| Create the mental, physical, and white space necessary to create an environment of expansive thinking | Identify the objective of your work by getting to the “ask behind the ask” and begin to establish expectations | Draft your approach to the problem and more clearly and concisely define the timing and output of your work | Generate a solution, gauge its fit with expectations, and rapidly iterate & hone until a final solve is reached |
| 5 | 15 | 20 | 20 |

# Ben Jones provides a good evaluation guide for data visualization



Thegradientboost

CREATING STRONGER LEARNERS

# Follow simple rules to visualize insights with impact

1. Highlight your message to eliminate distractions
2. Use visual cues to help lead your audience through your insight
3. Use contrast; size, color to capture the reader's attention

Thegradientboost

CREATING STRONGER LEARNERS

# Follow simple rules to visualize insights with impact

**1** Highlight your message to eliminate distractions

Thegradientboost
CREATING STRONGER LEARNERS

# Minimize distractions depending on the chart you are using

## Pie charts

- Place the largest segment at 12 o'clock on the right, second at 12 o'clock on the left, the rest following counterclockwise
- Use bold to increase legibility on a shaded background or to emphasize a segment
- A pie chart should not have more than five slices
- A segmented bar is more efficient than a pie chart at showing portions of a whole

## Line charts

- Keep the maximum number of lines to three or possibly four and keep the style simple and uncomplicated
- Direct labeling allows the reader to identify data quickly and focus on patterns

## Bar charts

- Horizontal bars should not be displayed in a random order
- A horizontal bar chart is most useful when ranking the items by the same characteristic from largest to smallest or vice versa
- Avoid grid lines and scales on horizontal bars and opt for direct labeling instead

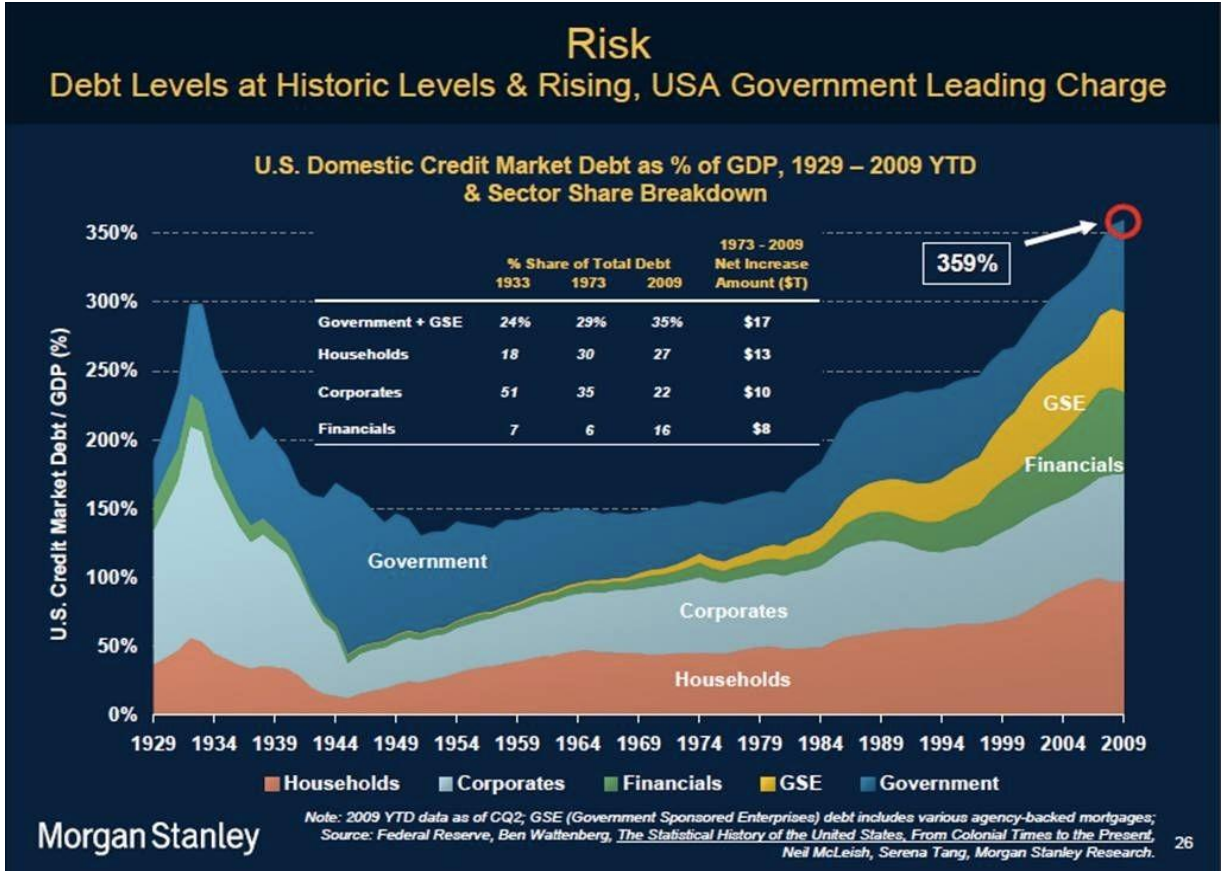
## Icons

- Icons should be simple, symmetrical, clear and crisp even at small sizes, and roughly square shaped

# Correct


Brazil query volume June 15
-240
-180
-120
-60
0
60
120
180
240
300
360
Minutes after kickoff
source: Google, Inc
June 15 traffic
Normal Tuesday traffic

# WRONG

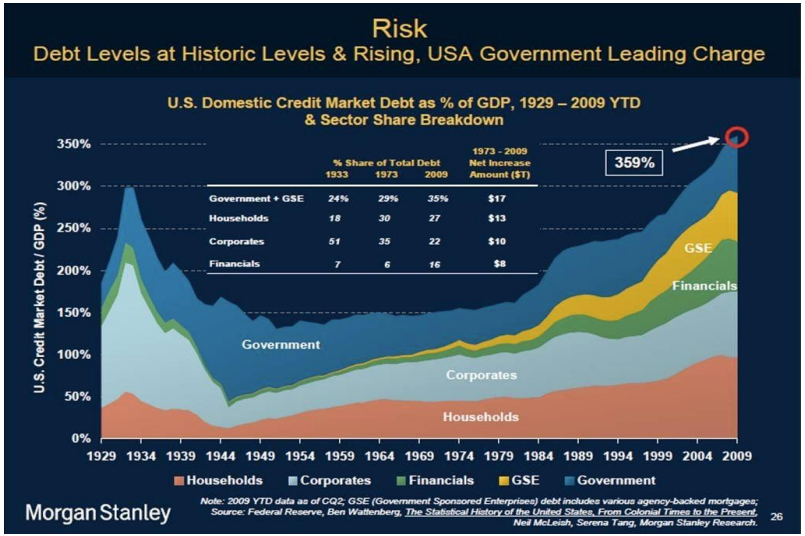

Risk
Debt Levels at Historic Levels & Rising, USA Government Leading Charge
U.S. Domestic Credit Market Debt as % of GDP, 1929 – 2009 YTD
& Sector Share Breakdown
U.S. Credit Market Debt / GDP (%)
350%
300%
250%
200%
150%
100%
50%
0%
1929 1934 1939 1944 1949 1954 1959 1964 1969 1974 1979 1984 1989 1994 1999 2004 2009
% Share of Total Debt
1933 1973 2009
1973 - 2009 Net Increase Amount ($T)
Government + GSE 24% 29% 35% $17
Households 18 30 27 $13
Corporates 51 35 22 $10
Financials 7 6 16 $8
359%
Government
Corporates
Households
GSE
Financials
Households Corporates Financials GSE Government
Morgan Stanley
Note: 2009 YTD data as of CQ2; GSE (Government Sponsored Enterprises) debt includes various agency-backed mortgages;
Source: Federal Reserve, Ben Wattenberg, The Statistical History of the United States, From Colonial Times to the Present,
Neil McLeish, Serena Tang, Morgan Stanley Research.
26

Thegradientboost
CREATING STRONGER LEARNERS

# Highlight your message and eliminate distractions

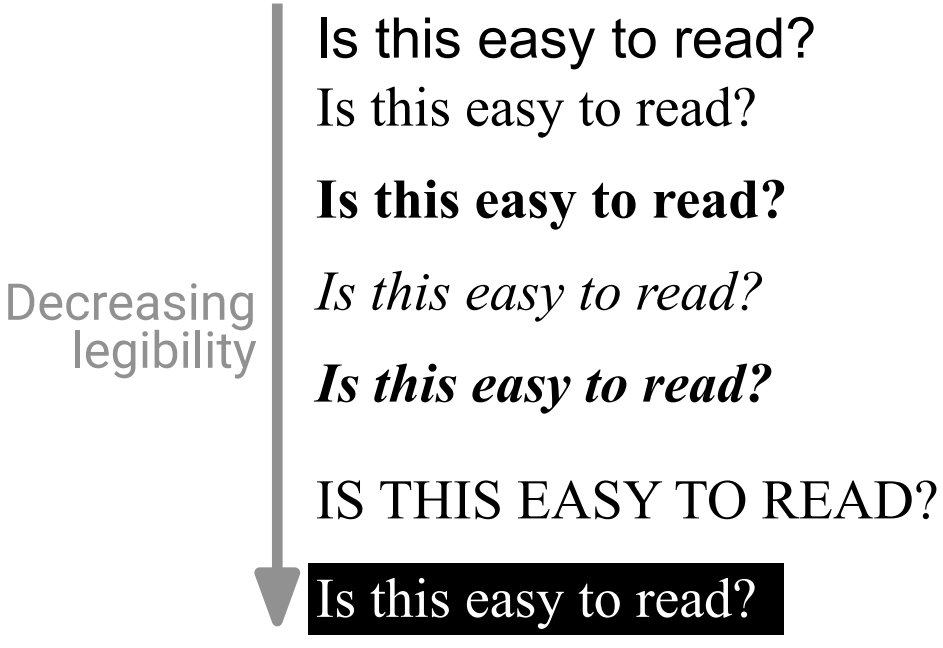
## Correct

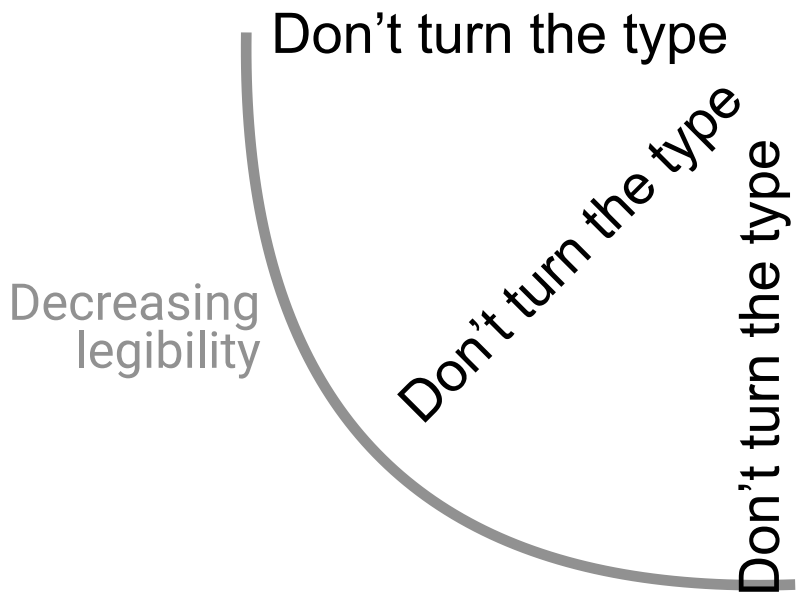

## WRONG




Source: http://www.informationisbeautiful.net/2010/peak-break-up-times-on-facebook/; Morgan Stanley

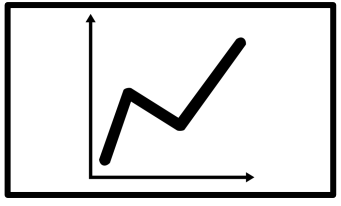
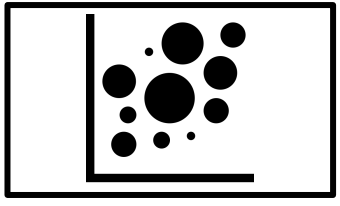
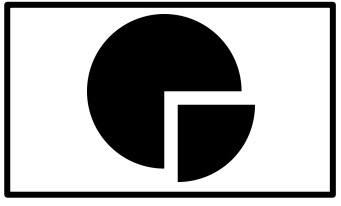
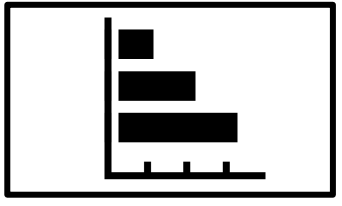
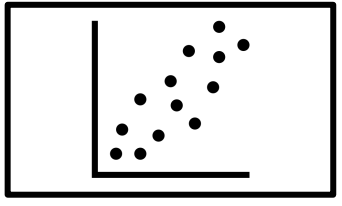
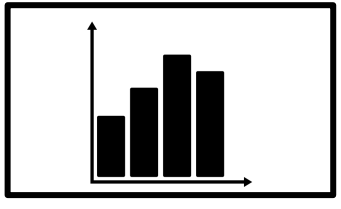
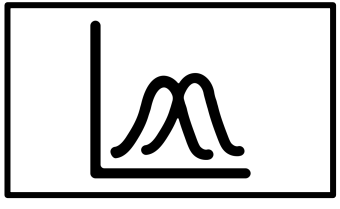
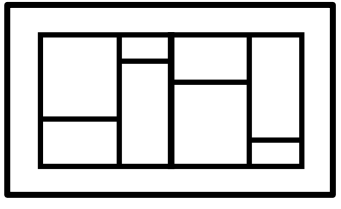
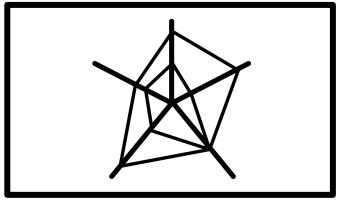
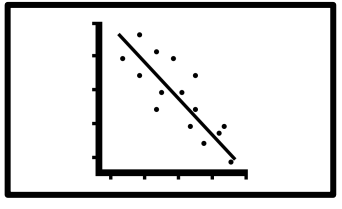
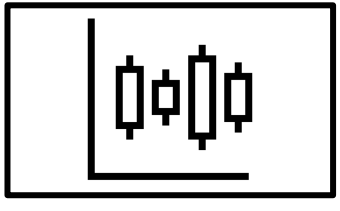
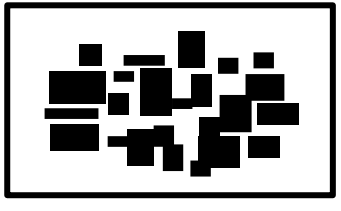
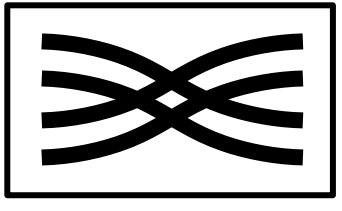
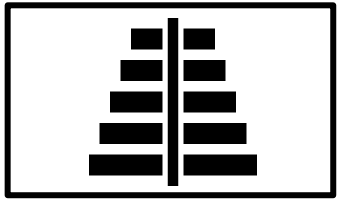
# Follow some of these basic rules to ensure type legibility



Thegradientboost
CREATING STRONGER LEARNERS

# Use the right visual for the right purpose

| Change | Clustering | Relativity | Ranking | Correlation |
|---|---|---|---|---|
| Line | Bubble | Pie | Bar | Scatter |
|  |  |  |  |  |
| Column | Distribution | Tree Map | Spider | Regression |
|  |  |  |  |  |
| Histogram | Candlestick | Word Cloud | Sankey | Double Bar |
|  |  |  |  |  |

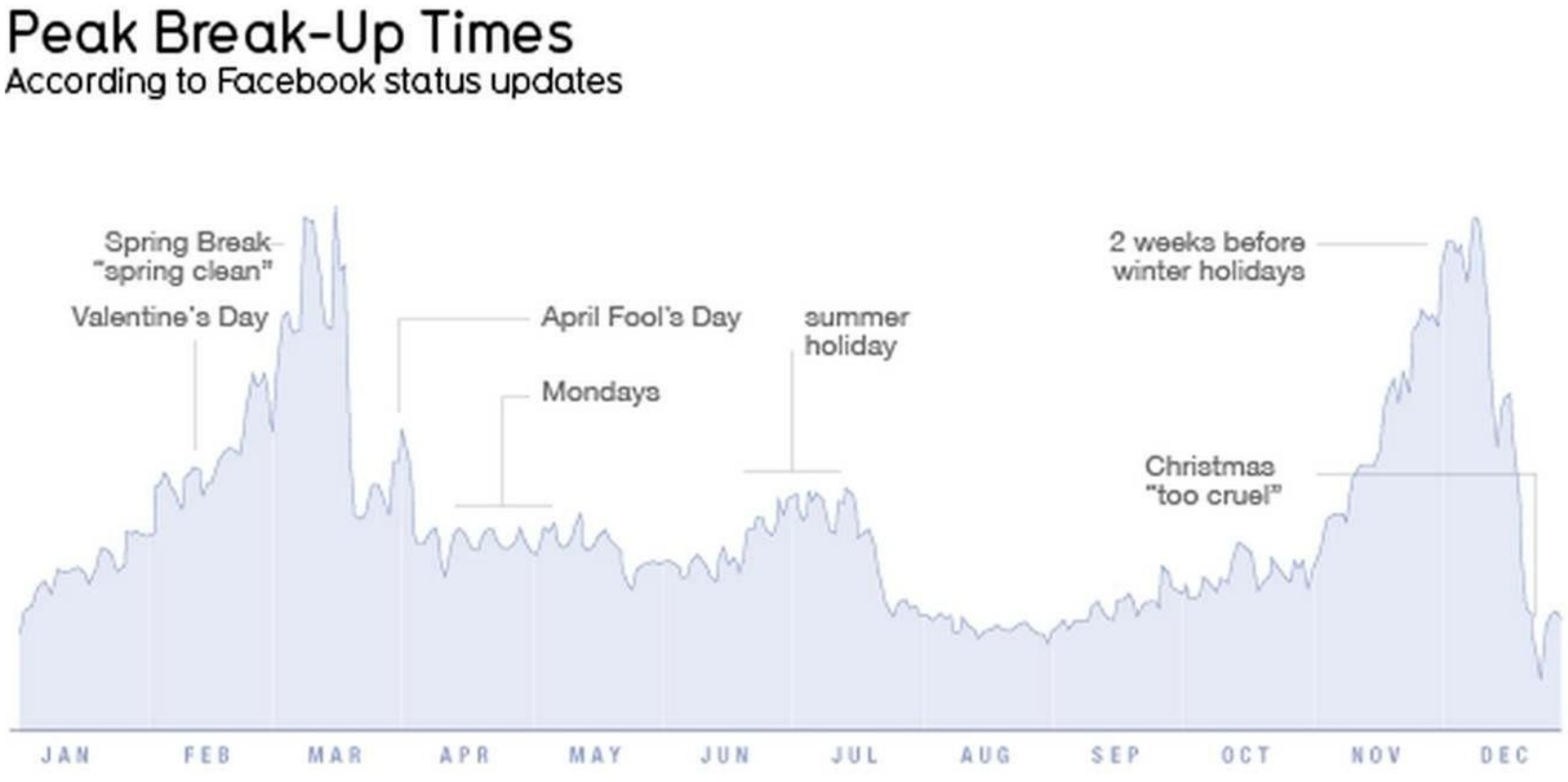
Thegradientboost

CREATING STRONGER LEARNERS

# Follow simple rules to visualize insights with impact

**2** Use visual cues to help lead your audience through your insight

# Correct

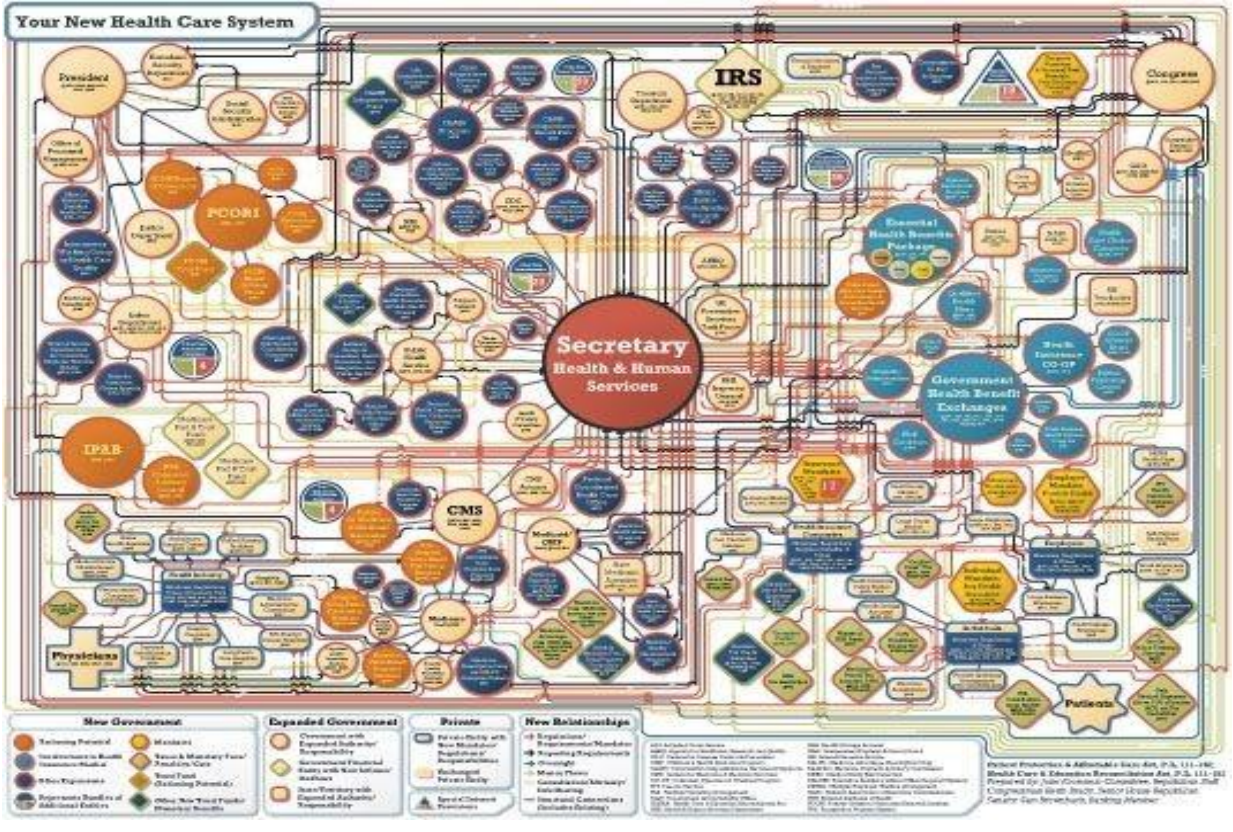



Source: http://www.informationisbeautiful.net/2010/peak-break-up-times-on-facebook/

Thegradientboost

CREATING STRONGER LEARNERS

# WRONG

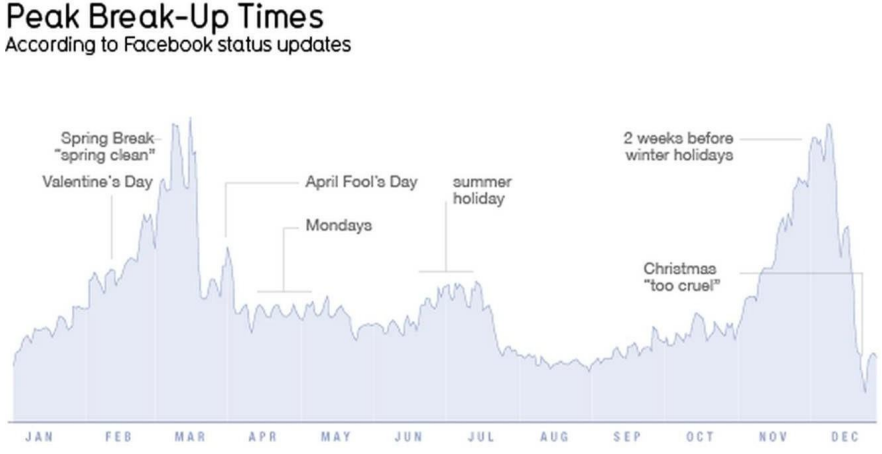



Source: http://blow.blogs.nytimes.com/2010/07/28/chart-abuse/

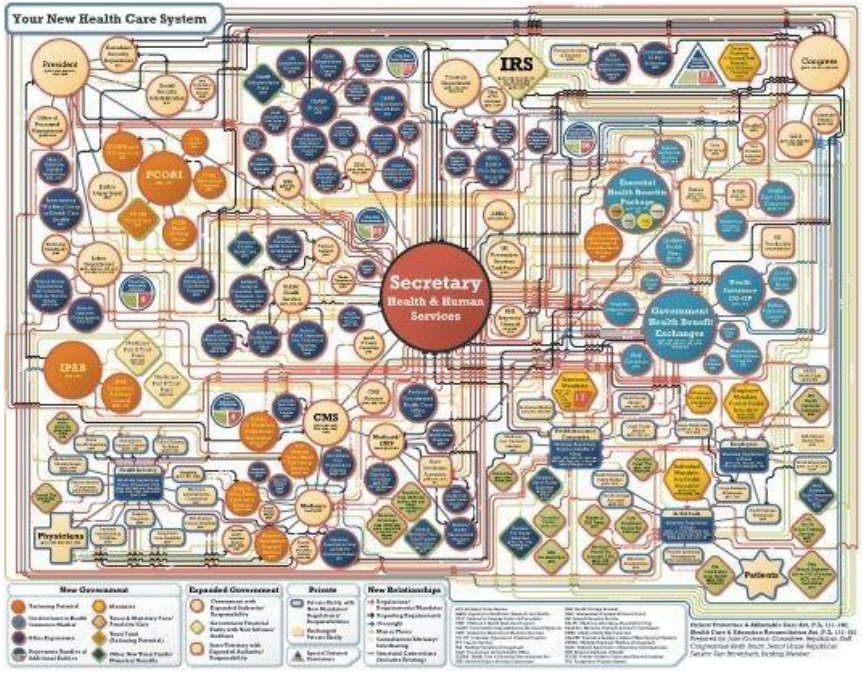
Thegradientboost
CREATING STRONGER LEARNERS

# Use visual cues to help lead your audience through your insight

## Correct




## WRONG




Source: Google; http://blow.blogs.nytimes.com/2010/07/28/chart-abuse/

Thegradientboost

CREATING STRONGER LEARNERS

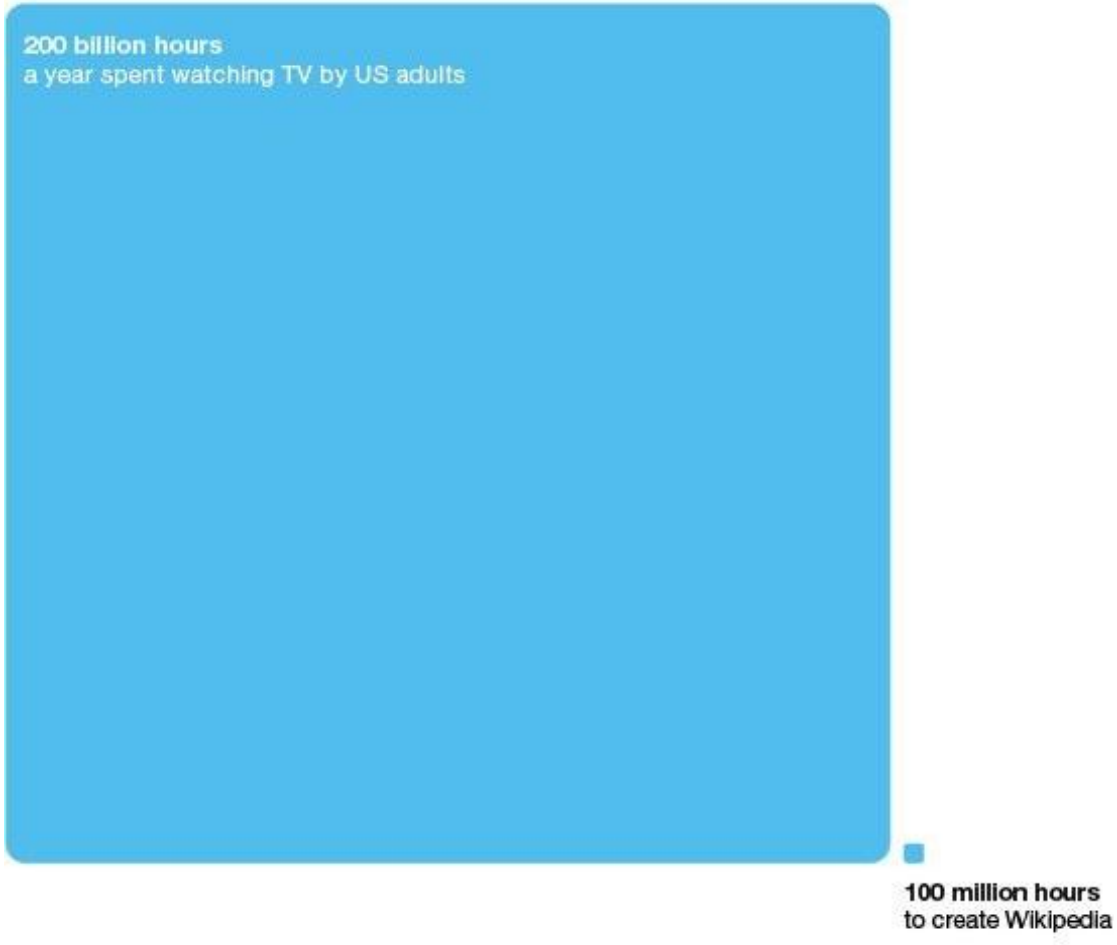
# Follow simple rules to visualize insights with impact

**3** Use contrast; size, color to capture the reader’s attention

Thegradientboost

CREATING STRONGER LEARNERS

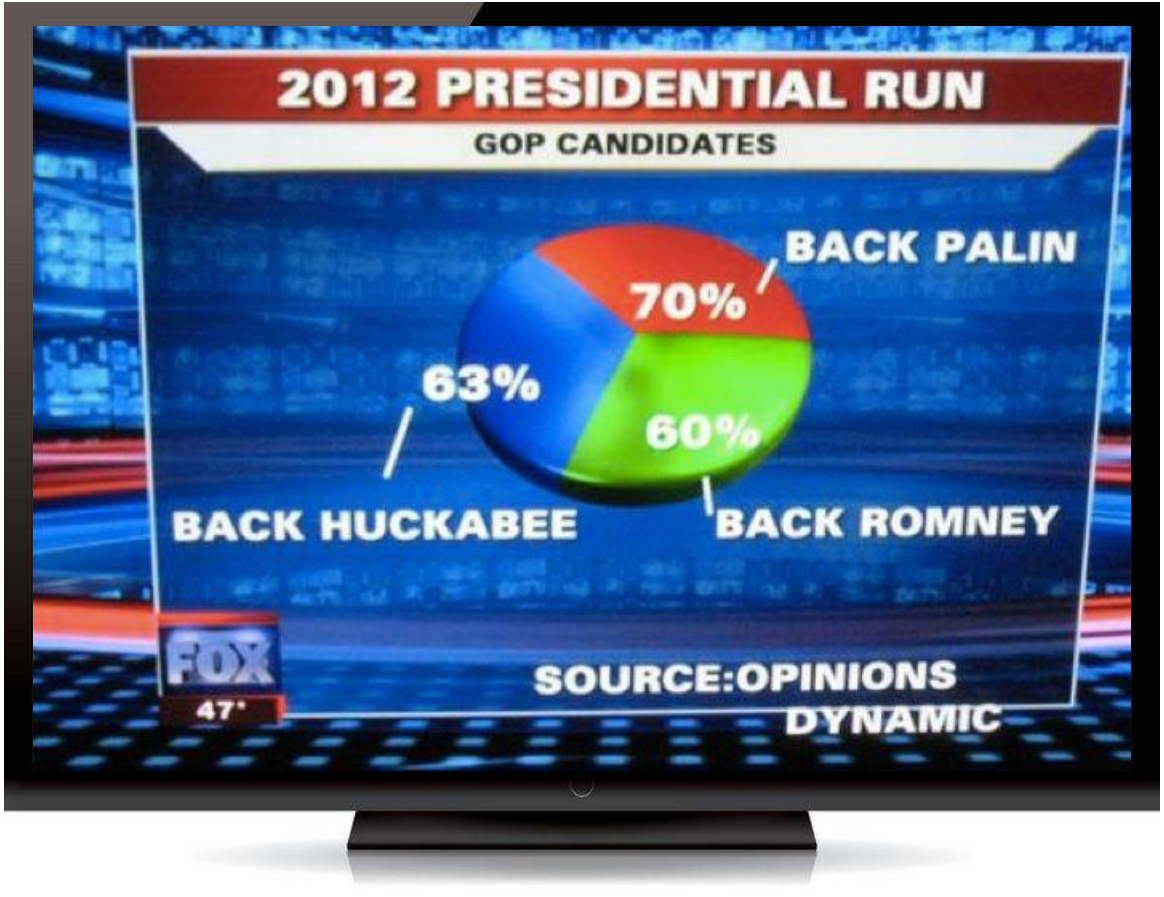
# Correct



Thegradientboost

CREATING STRONGER LEARNERS

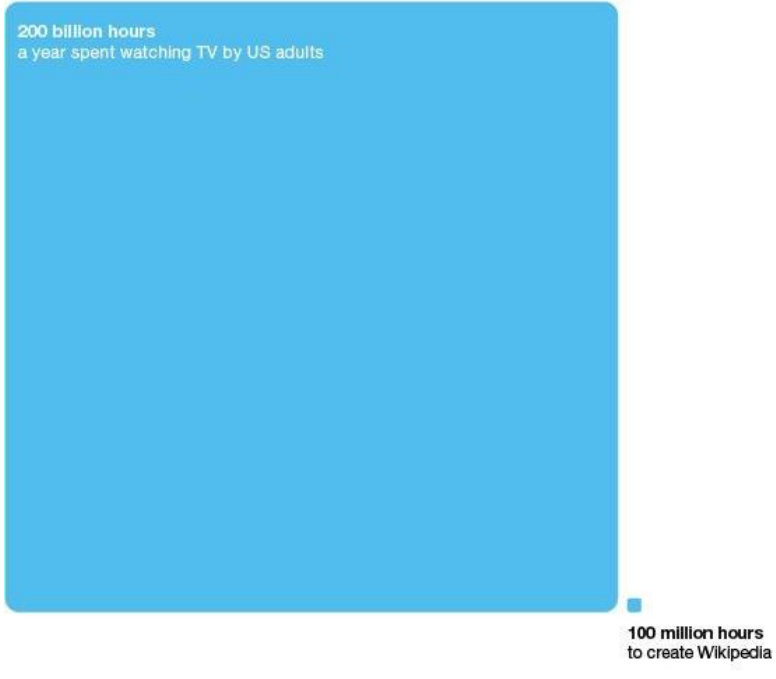
# WRONG



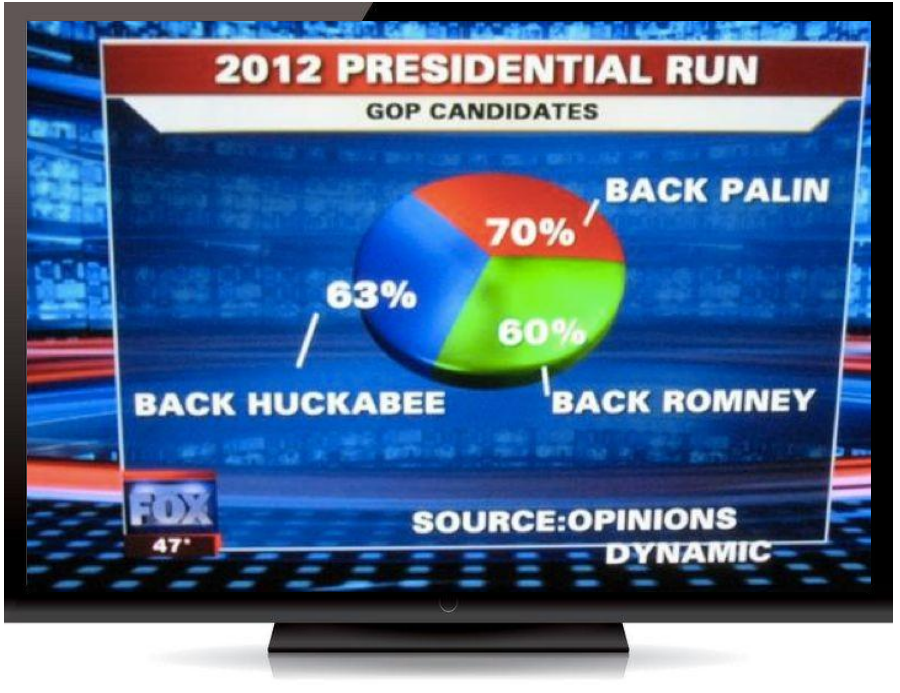
Thegradientboost
CREATING STRONGER LEARNERS

# Use contrast; size, color to capture the reader's attention

## Correct




## WRONG



Thegradientboost

CREATING STRONGER LEARNERS

# Data visualization design tips

Let’s look at a few tips to help improve your data visualization skills

Thegradientboost

CREATING STRONGER LEARNERS

# Follow these rules when using bold text in data visuals

Don't

Don't set a huge amount of text in bold. Emphasizing everything means nothing gets emphasized.

Do

Use bold type to emphasize the focal point of the message. Be judicious.

| **Name** | **Data** | **Data** | **Data** |
|---|---|---|---|
| **Company A** | 0.0 | **0.0** | **0.0** |
| **Company B** | 0.0 | **0.0** | **0.0** |
| **Company C** | 0.0 | **0.0** | **0.0** |
| **Company D** | 0.0 | **0.0** | **0.0** |

| Name | Data | Data | Data |
|---|---|---|---|
| **Company A** | 0.0 | 0.0 | **0.0** |
| **Company B** | 0.0 | 0.0 | **0.0** |
| **Company C** | 0.0 | 0.0 | **0.0** |
| **Company D** | 0.0 | 0.0 | **0.0** |

Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

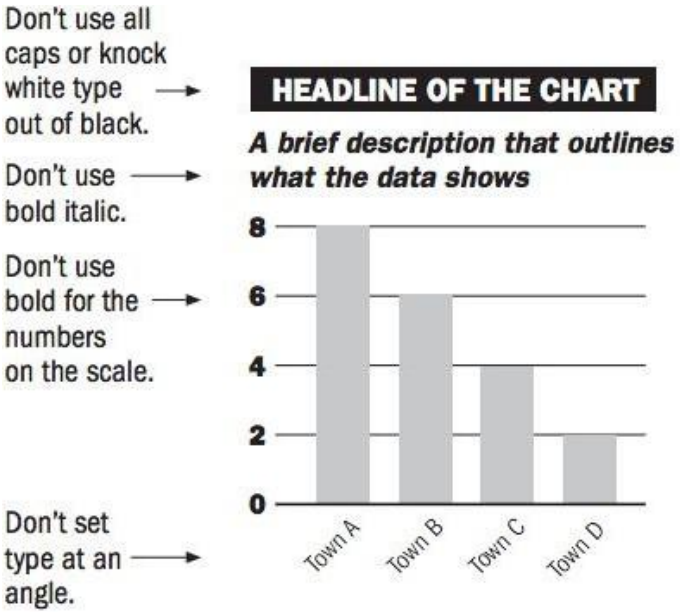
# Follow these rules when creating headlines for visuals

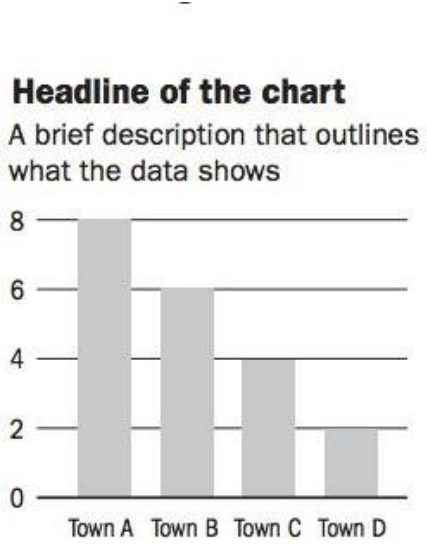
Don't

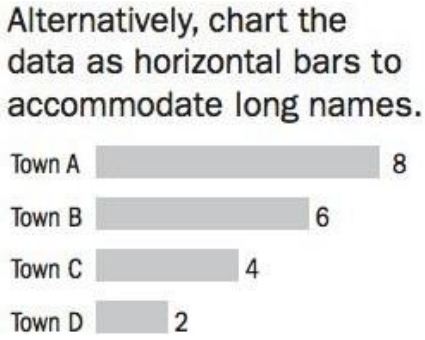
Don't permit typography to oppress the underlying data.

Do

Keep the typography simple. The headline can be either bold or a couple sizes larger.



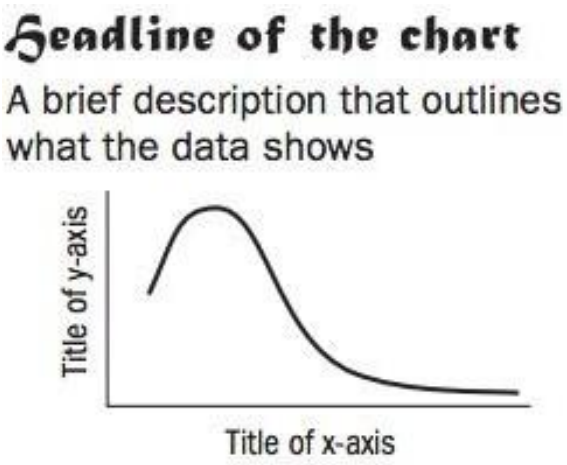

Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

# Follow these rules when creating headlines for visuals

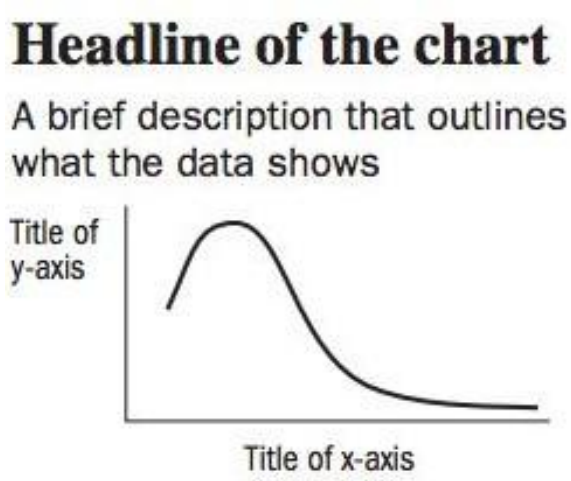
Don't

Don't use highly stylized fonts or turn the type sideways to save space.




Do

Serif and sanserif fonts can complement each other and add variety, and are still highly legible.

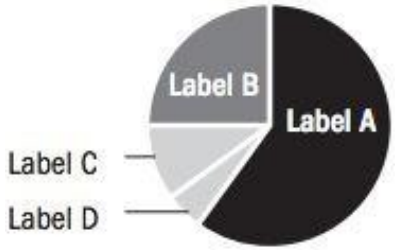



Source: Dona Wong "The WSJ Guide To Information Graphics"

# Follow these rules when creating Pie Chart data visuals

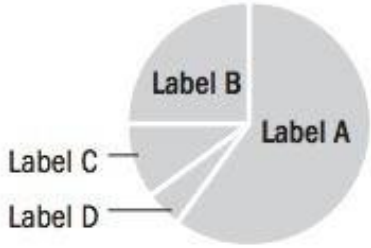
Don't

Don't knock white type out of black or color. Legibility is compromised.




Do

Use bold to increase legibility on a shaded background or to emphasize a segment.



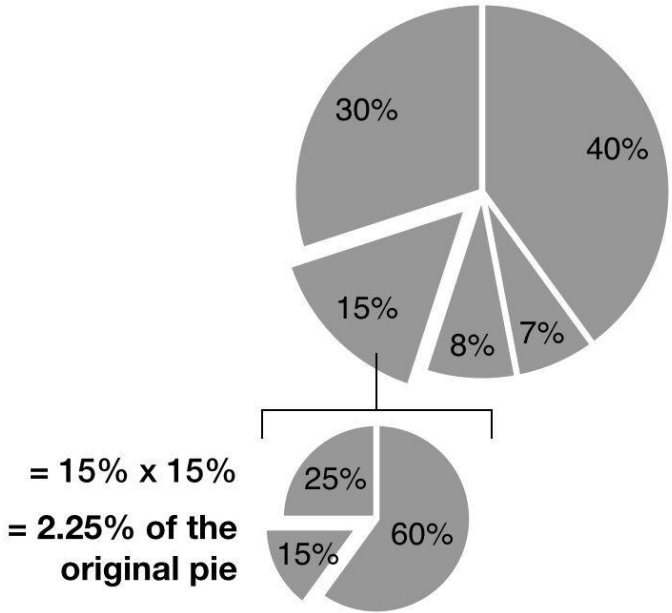


Source: Dona Wong "The WSJ Guide To Information Graphics"

# Pies are not always an ideal visual format complex data

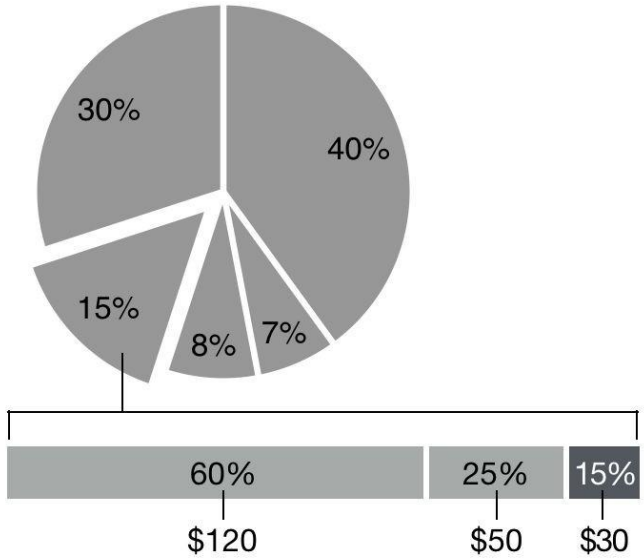
Don't

Segmenting within a slice makes the second segment difficult to grasp. It's too much for readers.




Do

A segmented bar is more efficient than a pie chart at showing portions of a whole.



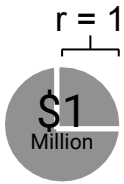

Source: Dona Wong "The WSJ Guide To Information Graphics"

# Care must be taken when designing proportional pies

Don't

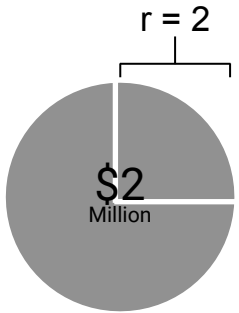
A common mistake is to represent the relative size of circles based on their radii.

Company a




$$\text{Area} = \pi r^2 = \pi 1^2 = 1\pi$$

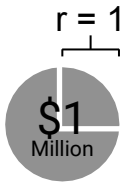
Company b



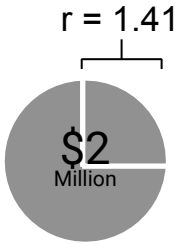

$$\text{Area} = \pi r^2 = \pi 2^2 = 4\pi$$

Do

Proportional pies should always be calculated based on surface area. Area = $\pi r^2$.




$$\text{Area} = \pi r^2 = \pi 1^2 = 1\pi$$




$$\text{Area} = \pi r^2 = \pi 1.41^2 = 2\pi$$

Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

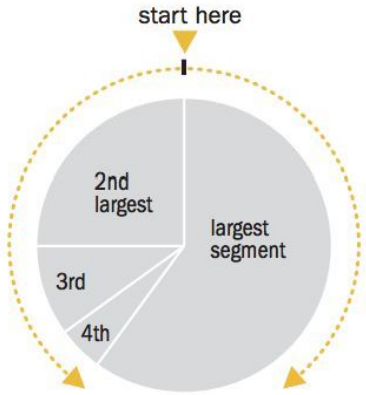
# Follow these rules when creating Pie Chart data visuals

## Less effective order

1

It's intuitive to read top to bottom and clockwise. **Never chart segments clockwise from smallest to largest.** By ordering the slices from smallest to largest in clockwise direction or vice versa, the least important segment has the most prominent position.

## Larger segments on top




Reading a pie chart is like reading a clock. It's intuitive to start at 12 o'clock and go clockwise.

2

Therefore, it is most effective to **place the largest segment at 12 o'clock on the right to emphasize its importance.**

The best way to order the rest of the segments is to place the second biggest slice at 12 o'clock on the left; the rest would follow counterclockwise. The smallest slice will fall near the bottom of the chart, in the least significant position.

## Too many slices

3

It's difficult to compare and contrast between many segments. **A pie chart shouldn't have more than five slices.**

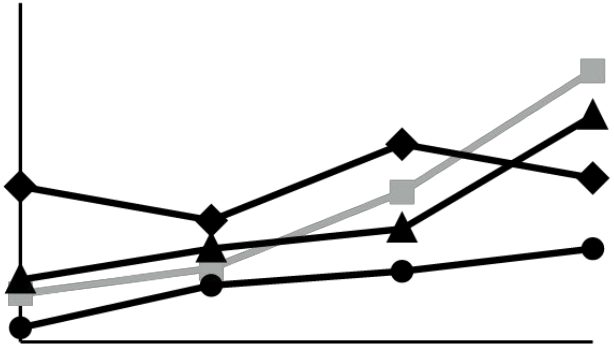
If there are more than five, combine the smaller and less significant segments to create the fifth slice and label it "Other." If all segments have to be represented separately, use a stacked or segmented bar chart instead. See page 79.

Source: Dona Wong "The WSJ Guide To Information Graphics"

# Line charts should strive for clean lines, clear signals

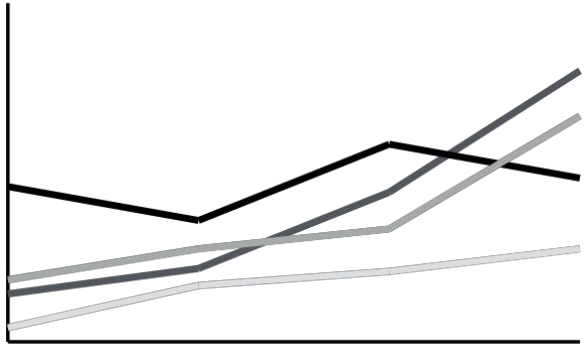
Don't

It is tempting to try all the line styles and data symbols, but they only obscure the chart's information.



Do

Keep the maximum number of lines to three or possibly four and keep the style simple an uncomplicated.



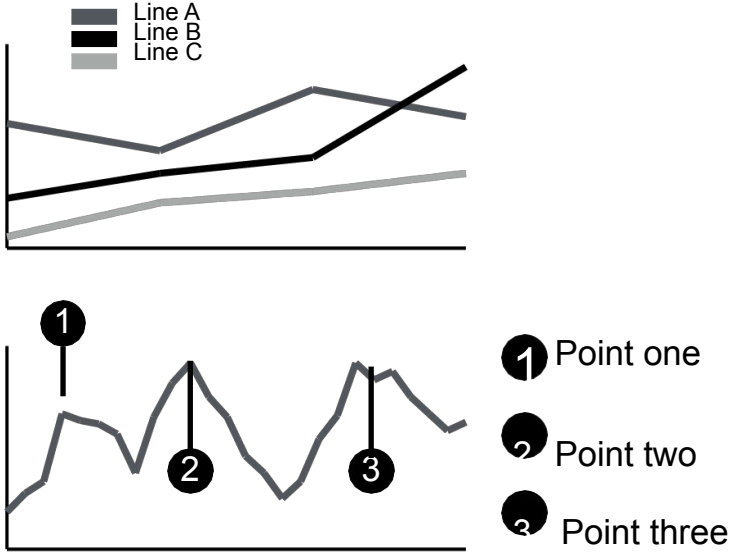
Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

# Legends, labels should simplify charts, not complicate

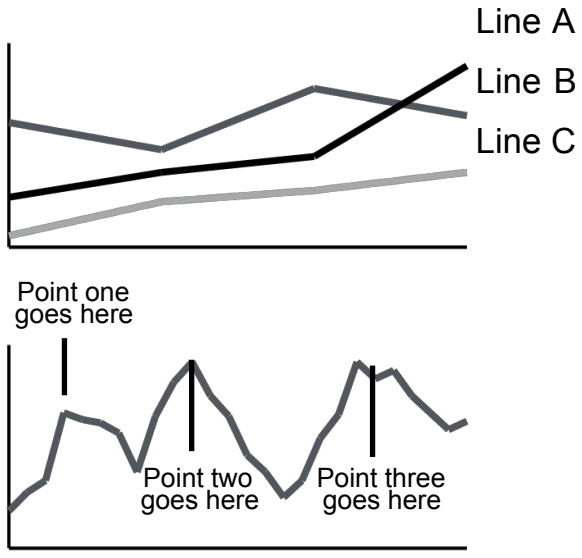
Don't

A legend or insight separated from the line requires readers to do extra work.




Do

Direct labeling allows the reader to identify data quickly and focus on patterns.




Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

# Follow these rules when creating Bar Chart data visuals

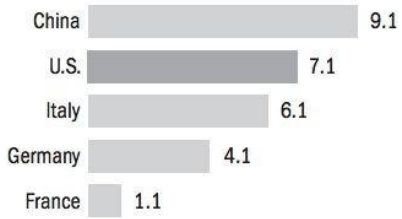
1 **No random lineup**

Don't plot horizontal bars in a random order. The main quality of a horizontal bar chart is the ranking of items by the same attribute. Plotting the bars in an arbitrary sequence defeats the purpose.


2 **The right order**

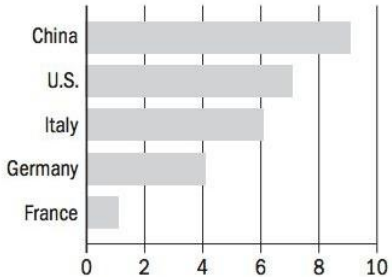
A horizontal bar chart is most useful when ranking the items by the same characteristic, such as ranking the countries by sales of a product.

The bars should be ranked from the largest to the smallest or vice versa. A specific bar can be highlighted with a different shade.




3 **Avoid grid lines and scale**

Horizontal bars are not as easy to compare as vertical bars. Using a scale and grid lines would make it even harder to discern the relative lengths of the bars. Direct labeling is cleaner and clearer.




Source: Dona Wong "The WSJ Guide To Information Graphics"

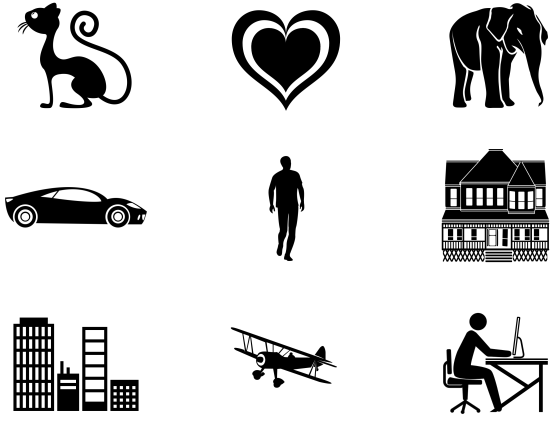
Thegradientboost

CREATING STRONGER LEARNERS

# Icons can be effective when chosen carefully

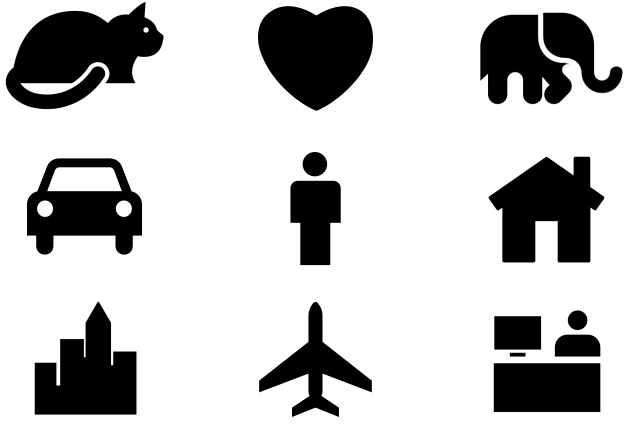
Don't

Use icons and symbols that include too many details, as they distract readers from the underlying data.



Do

Icons should be simple, symmetrical, clear and crisp even at small sizes, and roughly square shaped.



Source: Dona Wong "The WSJ Guide To Information Graphics"

Thegradientboost
CREATING STRONGER LEARNERS

# Here are some tests you can use to improve your design

## “The spartan test”

**Would eliminating this change anything?**

If the answer is no, get rid of it. Earn your reader's trust by giving them only what they need.

## “The peek test”

**Where is your eye drawn?**

Look away from your visual for 5 seconds, then back to it. This is most likely the place your audience's eye will be drawn as well.

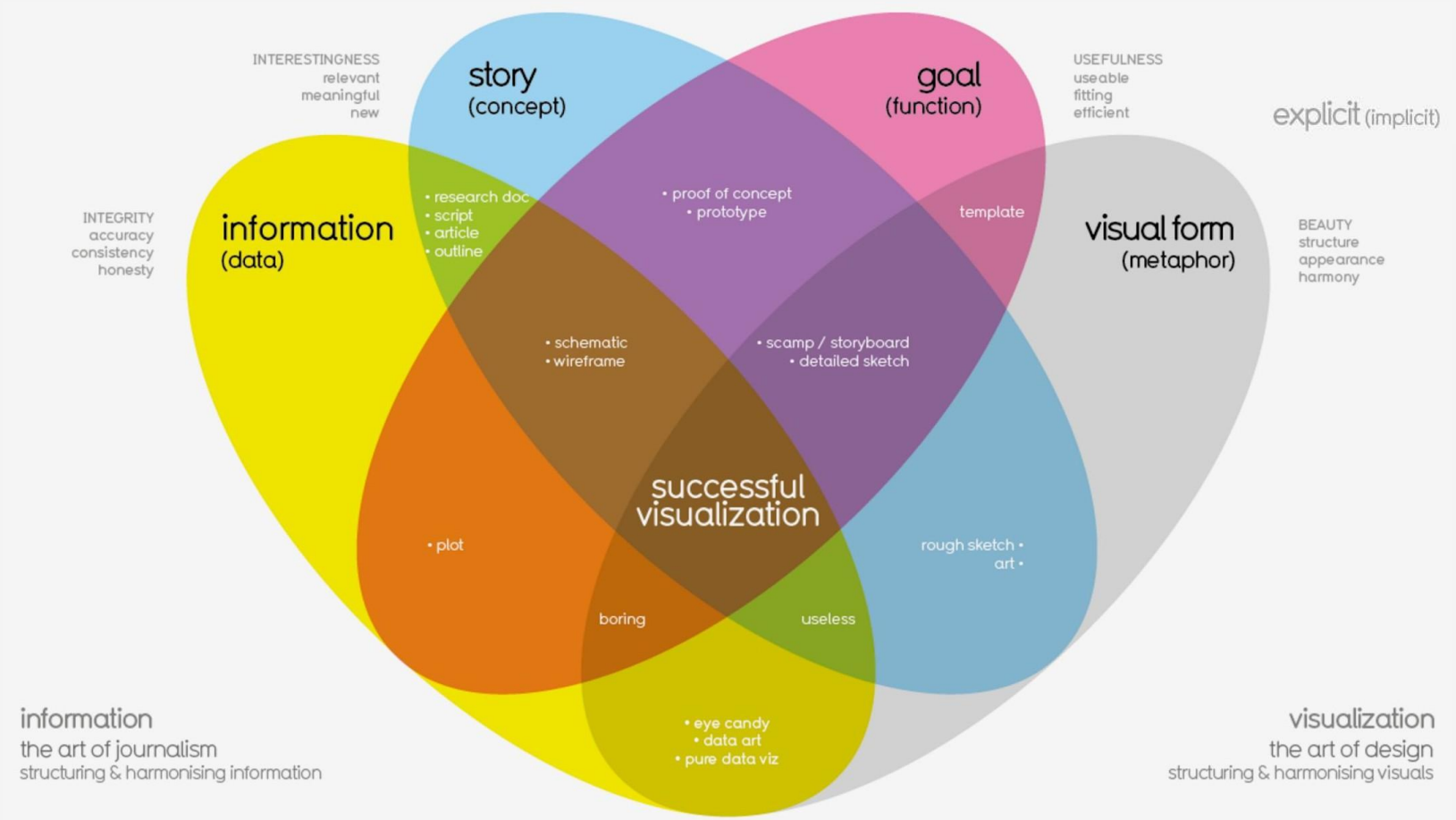
## “The colleague test”

**Think your visual is perfect? Have a colleague take a look.**

Give them minimal context and 10-15 seconds to process and ask what they would take away.

Thegradientboost
CREATING STRONGER LEARNERS

# Remember this when you are creating a visualization




Source: David McCandless (2009)