

Data visualization guidelines

V5

Functional color

Our colors for use in data visualization have been developed to provide a functional and flexible suite of colors that are both on-brand and accessible.

An expansion based on McKinsey's core colors, these palettes serve different functions within data visualization, including charting categorical and quantitative data.

Text over color

The text labels atop the color swatches in this section are set in either black or white for optimal legibility. Use these samples as guidance for on-screen visualizations. For print purposes, as colors vary depending on paper stock and ink used, some text colors may need to be reversed for better legibility.

Core McKinsey colors

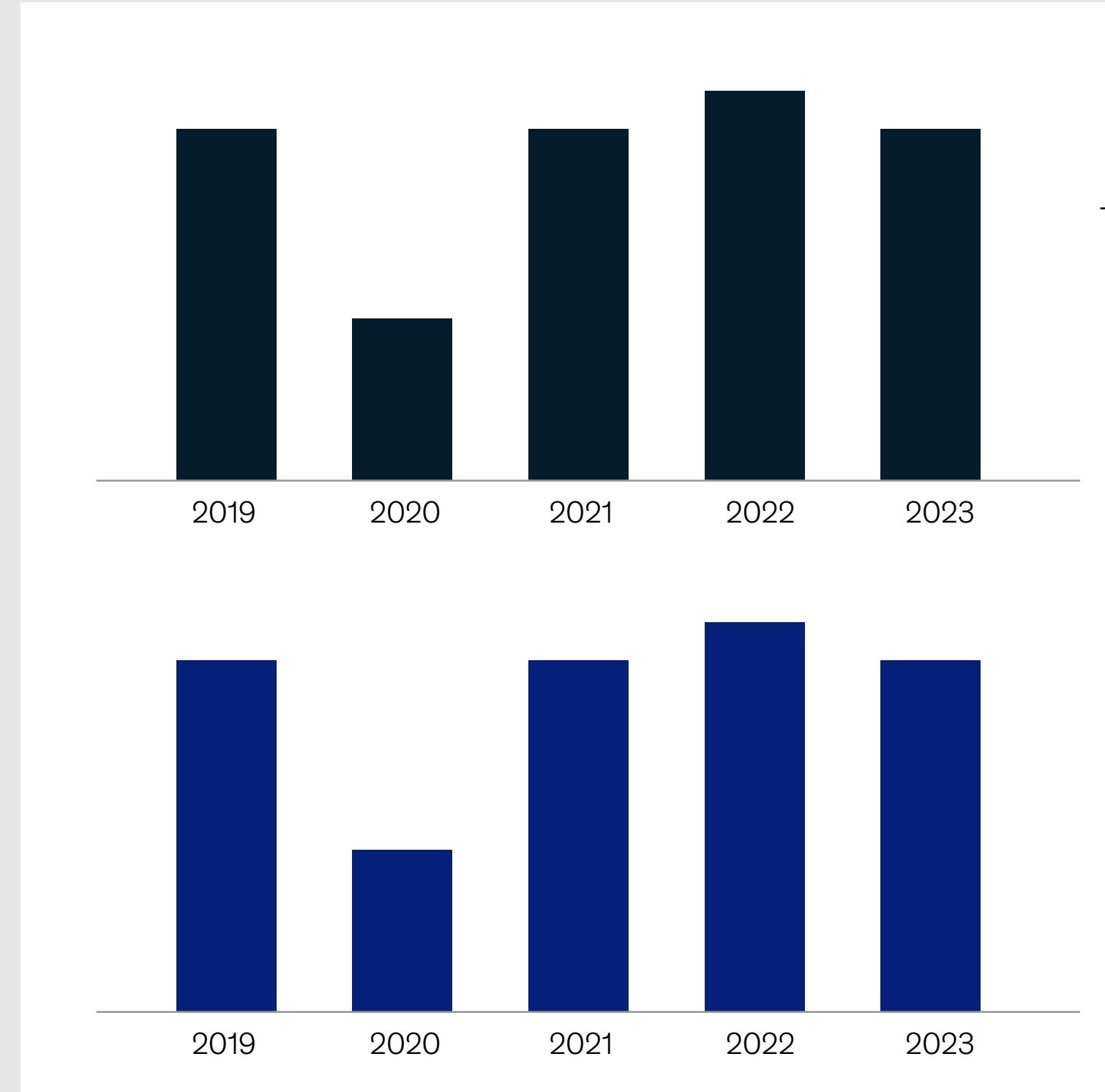
Deep Blue 900
CMYK 100, 50, 0, 90
RGB 5, 28, 44
HEX #051C2C

Cyan 500
CMYK 83, 0, 0, 0
RGB 0, 169, 244
HEX #00A9F4

Electric Blue 500
CMYK 93, 73, 0, 0
RGB 34, 81, 255
HEX #2251FF

An additional core color for data visualization

Electric Blue 900
CMYK 90, 75, 0, 60
RGB 6, 31, 121
HEX #061F79



Two options on a white background:
When visualizations do not require color for distinguishing categories or numerical encoding, use either Deep Blue 900 or Electric Blue 900

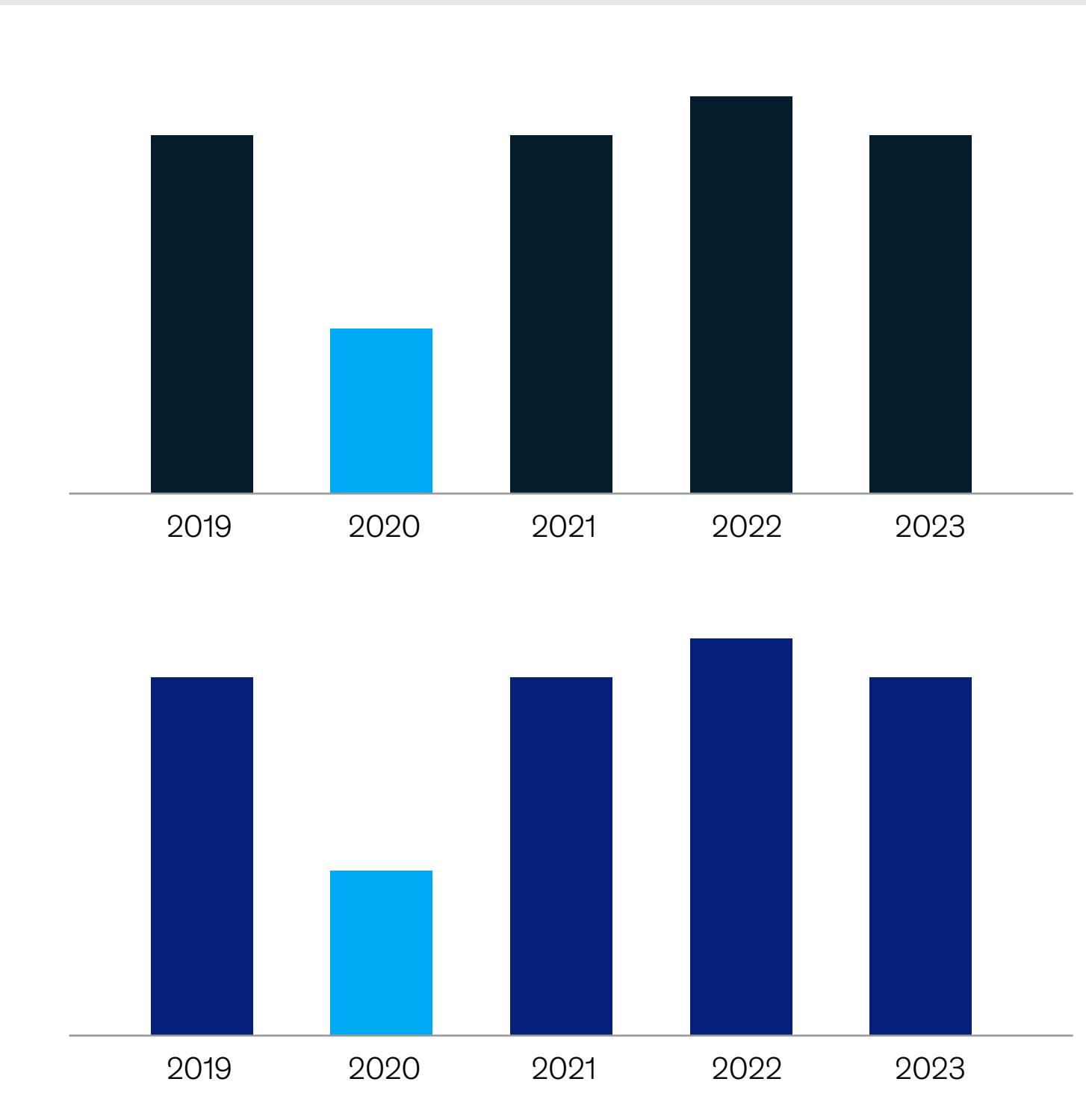


When this type of chart is on a Deep Blue 900 background, use Cyan 500

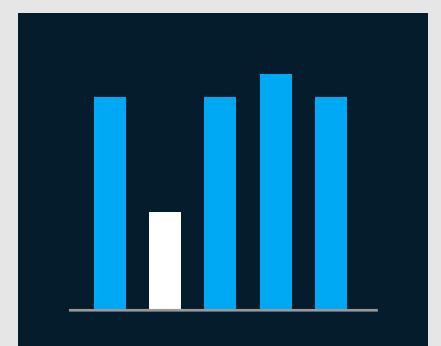
Color combinations

Introduce additional colors only when necessary to convey information, such as data categories or data point highlights.

Charts with a second highlight color

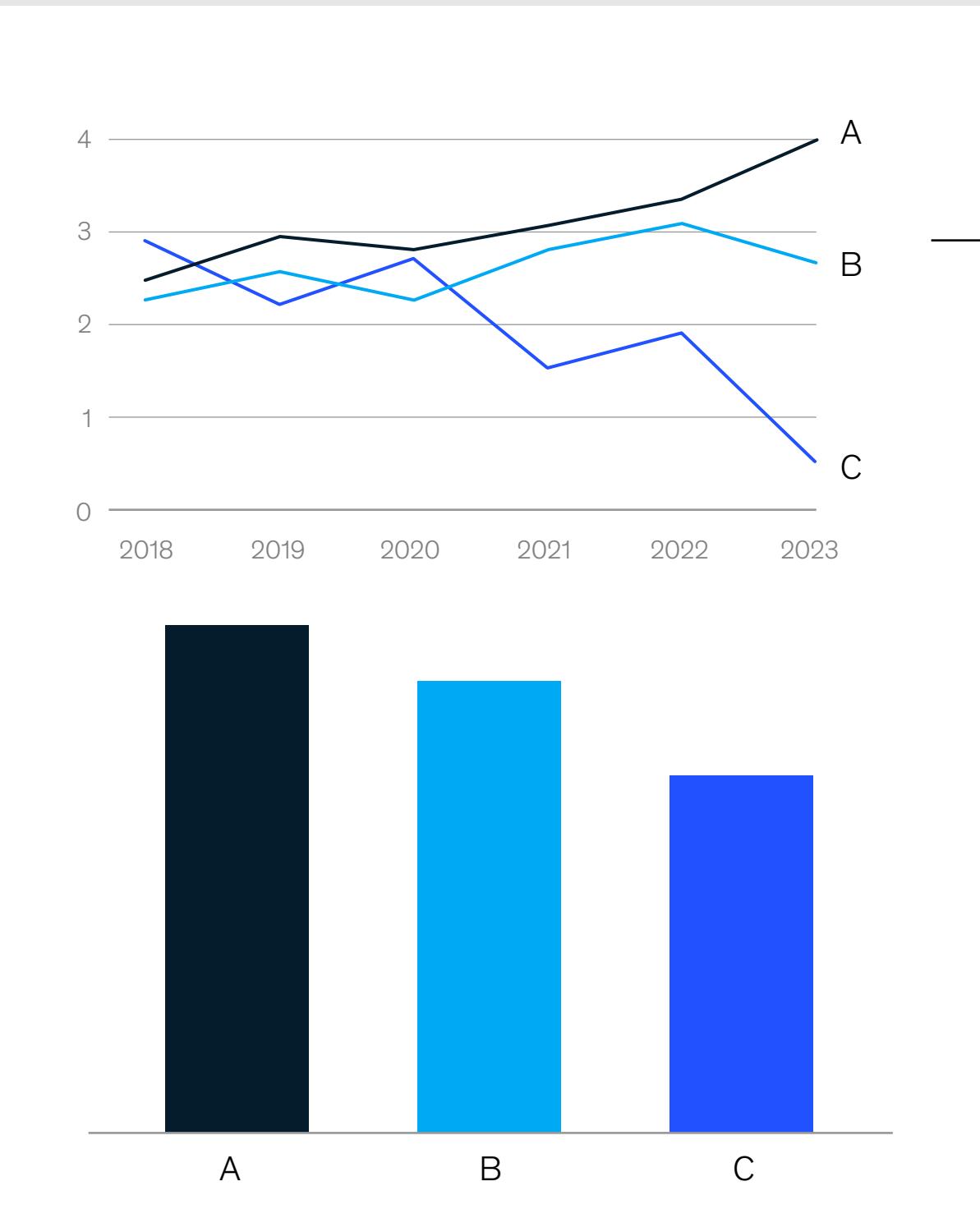


On a white background, Cyan 500 can be used to highlight specific data points or information of interest, regardless of whether the visualization's base color is Deep Blue 900 or Electric Blue 900

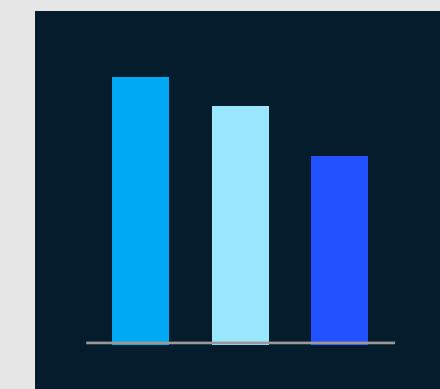


When this type of chart is on a Deep Blue 900 background, use white

Charts with 3 discrete categories



On a white background, return to the core McKinsey 3-color combination of Deep Blue 900, Cyan 500, and Electric Blue 500



When 3-category charts are on a Deep Blue 900 background, use Cyan 500, Cyan 200, and Electric Blue 500

To see information on our full color palette guidance, please visit the data visualization page on the Our Identity site.

Four or more categories

The categorical scale is used to showcase multiple different groups of information. Introduce the expanded colors in the order shown when a chart features more than one category. However, this color order is a starting point. Advanced chartmakers may refer to the broader McKinsey palette to reinterpret a scale that suits a specialized visualization. Consider the needs of your information.

All color combinations have been tested for common color-blindness types, as well as general accessibility.

Color scale sequence, by number of categories

4	Electric Blue 900 CMYK 90, 75, 0, 60 RGB 6, 31, 121 HEX #061F79	Cyan 500 CMYK 83, 0, 0, 0 RGB 0, 169, 244 HEX #00A9F4	Marine Green 300 CMYK 45, 0, 18, 6 RGB 117, 240, 231 HEX #75FOE7	Electric Blue 500 CMYK 93, 73, 0, 0 RGB 34, 81, 255 HEX #2251FF				
5					Marine Green 700 CMYK 75, 5, 40, 12 RGB 20, 184, 171 HEX #14B8AB			
6						Cyan 700 CMYK 90, 25, 0, 25 RGB 6, 121, 195 HEX #0679C3		
7							Orchid Pink 900 CMYK 35, 100, 10, 40 RGB 156, 33, 125 HEX #9C217D	
8								Crimson Red 300 CMYK 0, 55, 30, 5 RGB 241, 126, 126 HEX #F17E7E
9								Marine Green 900 CMYK 90, 25, 55, 20 RGB 16, 137, 128 HEX #108980
10								Orchid Pink 400 CMYK 20, 75, 0, 0 RGB 228, 121, 228 HEX #E479E4

Sequence on Deep Blue 900 background

Cyan 500 CMYK 83, 0, 0, 0 RGB 0, 169, 244 HEX #00A9F4	Cyan 200 CMYK 35, 0, 0, 0 RGB 153, 230, 255 HEX #99E6FF	Electric Blue 500 CMYK 93, 73, 0, 0 RGB 34, 81, 255 HEX #2251FF	Marine Green 700 CMYK 75, 5, 40, 12 RGB 20, 184, 171 HEX #14B8AB	Cyan 700 CMYK 90, 25, 0, 25 RGB 6, 121, 195 HEX #0679C3	White CMYK 0, 0, 0, 0 RGB 255, 255, 255 HEX #FFFFFF	Orchid Pink 900 CMYK 35, 100, 10, 40 RGB 156, 33, 125 HEX #9C217D	Crimson Red 300 CMYK 0, 55, 30, 5 RGB 241, 126, 126 HEX #F17E7E	Marine Green 900 CMYK 90, 25, 55, 20 RGB 16, 137, 128 HEX #108980	Orchid Pink 400 CMYK 20, 75, 0, 0 RGB 228, 121, 228 HEX #E479E4
--	--	--	---	--	--	--	--	--	--

Quantitative scales

Linear scales are used for data sets to show a range of values increasing from one extreme to the other (low/high, less/more). Generally, the darkest shade would represent the highest value. This logic is reversed for dark backgrounds.

Our linear scales use up to 5 levels of our core colors. If you need up to 7 steps, used the mixed blue scale.

For charts with multiple linear categories, like a tree map, color levels should range from the 900 level to the 200 level, using enough separation between colors.

Divergent scales are used to visualize a contrasting scale with values changing in opposite directions of a set breakpoint. Depending on your content, there are 2 divergent palettes to choose from.

For more guidance on linear and divergent scales, including the use of tools such as Chroma.js to generate a greater number of color steps, please visit the data visualization page on the Our Identity site.

Linear scale options, by number of steps

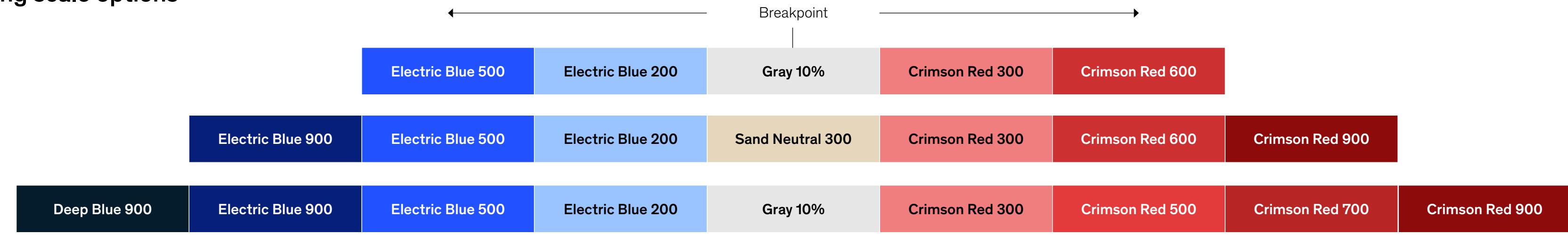
3	Deep Blue 900	Deep Blue 500	Deep Blue 200	Option A			
3	Electric Blue 900	Electric Blue 500	Electric Blue 200	Option B			
4	Electric Blue 900	Electric Blue 700	Electric Blue 500	Electric Blue 300			
5	Electric Blue 900	Electric Blue 700	Electric Blue 500	Electric Blue 300	Electric Blue 200		
6	Deep Blue 900	Electric Blue 900	Electric Blue 700	Electric Blue 500	Electric Blue 300	Electric Blue 200	
7	Deep Blue 900	Electric Blue 900	Electric Blue 700	Electric Blue 500	Cyan 500	Cyan 300	Cyan 200

Gray or Sand Neutral may be used as a breakpoint color in diverging scales. Gray is often a good option for quantitative diverging scales—to represent “zero,” for example.

Sand Neutral is often a good option for qualitative diverging scales—to signal data that is neutral, such as the midpoint for survey responses ranging from “strongly agree” to “strongly disagree.”

Electric Blue–Crimson Red and **Marine Green–Crimson Red** pairings: Diverging scales can flow in either direction to suit the logic of the underlying data—for example, blue on the left and red on the right or vice versa.

Diverging scale options



Choosing a chart type

To create charts that clarify and provide the right canvas for analysis, it's important to first understand why a chart is needed in relation to the data available.

Choosing the wrong visual aid, or simply defaulting to the most common data visualizations, can confuse the viewer or lead to misinterpreted data.

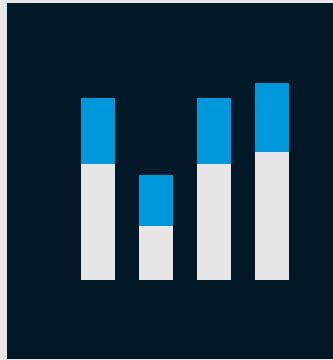
Ask: Is this story to compare categories? See trends over time? Or look at the geographical distribution of a dataset? The quick reference guide on this page can help you with choosing a chart type.

Magnitude

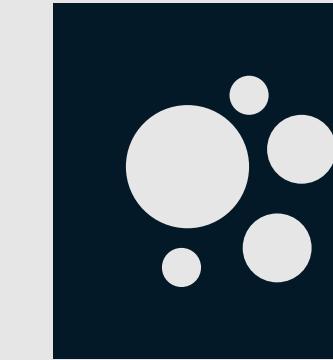
When comparing size, charts emphasizing magnitude are suitable. The comparison can be relative (what is larger/bigger), or absolute (the differences)



Bar chart



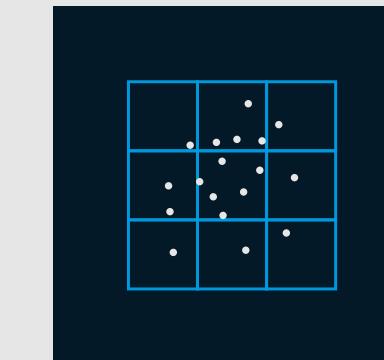
Stacked bar chart



Bubble chart

Correlation

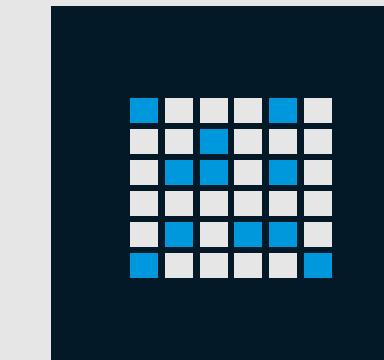
Correlation charts show the relationship between two or more variables



Scatterplot



Trend lines



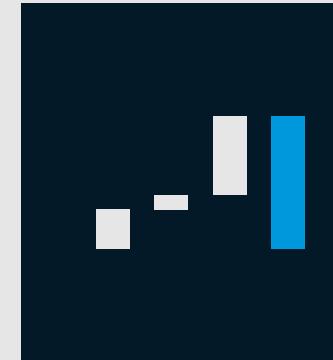
Heatmap

Part-to-whole

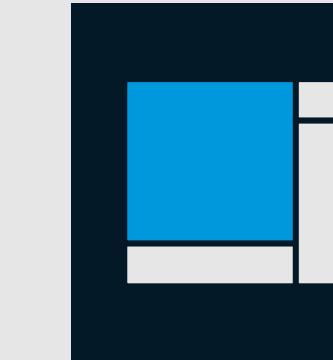
Part-to-whole charts focus on the distribution of the parts and show how a single entity can be broken down into its component elements



Donut and pie chart



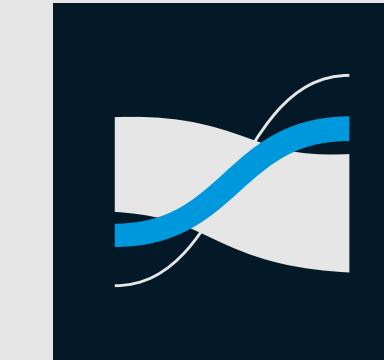
Waterfall chart



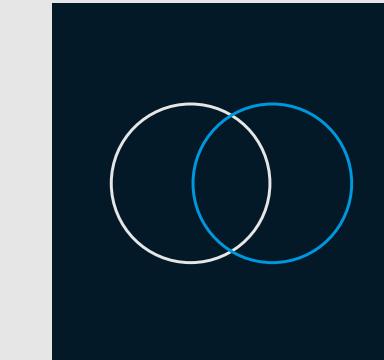
Treemap

Relationships

Relationship charts allow users to see logical relationships between entities or show volume changes from one set to another



Sankey diagram



Venn diagram

Trend over time

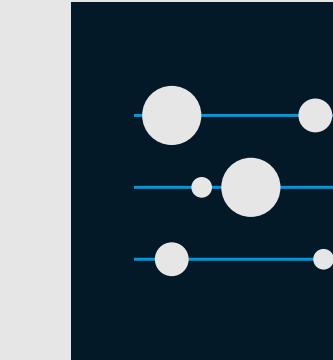
Trend over time charts emphasize changing trends. They can be short (intraday) movements or an extended series traversing decades or centuries



Line chart



Stacked area chart



Bubble timeline

Geospatial

Geospatial charts are used when highlighting precise locations or geographical patterns



Choropleth



Bubble map



Dot density map

Magnitude

Bar chart

A bar chart displays rectangular columns with a length proportional to the represented values.

One axis shows the specific categories being compared (by length), and the other axis represents a discrete value or time.

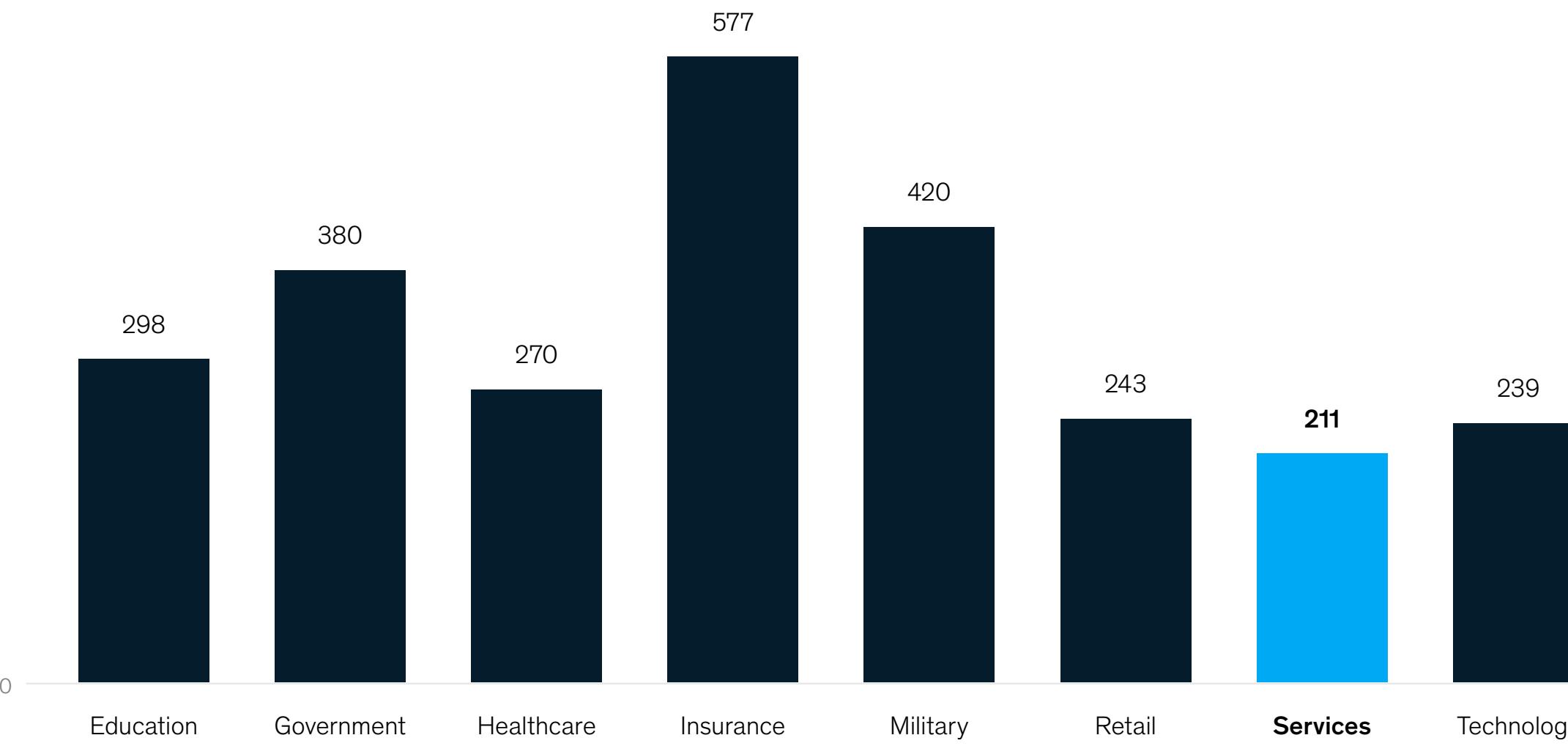
A single-category bar chart can be colored with just the Deep Blue 900 color. However, coloring a bar with the accent color can be used to highlight a meaningful data point.

Bars are directly labeled and grid axes have been removed to optimize legibility

Bar charts start at zero, so the chart baseline includes a zero value axis label

Charts that emphasize magnitude are suitable for comparing size. The comparison can be relative (what is larger/bigger), or absolute (what are the differences).

Service companies have exposed fewer personal records than any other industry
Personal records exposed, 2011-2016, millions



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Bars that share the same category of data, share the same color. In this example the Cyan 500 is used to highlight the insight (Service companies)

Stacked bar chart

Stacked bar charts are useful when comparing multiple categories of data.

For up to six categories, you can use the categorical palette. For up to 10 categories, you can use the expanded categorical palette.

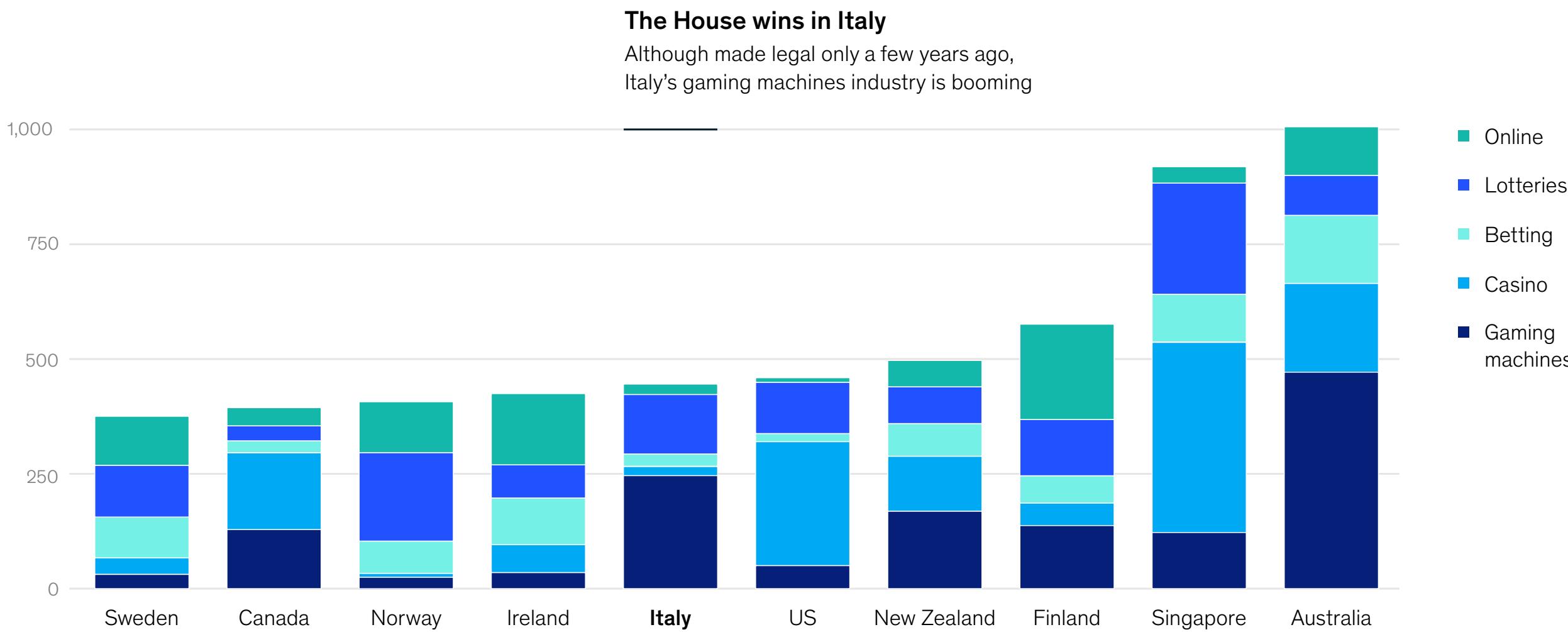
If you have more than 10 categories it's recommended to group all remaining categories past 10 into "Other" using Gray 60% to depict this.

An insight call-out helps frame the message and focus viewers on the key data point

Stacked bars can make direct labeling difficult; introducing grid axis and labeling can provide a clear way to read the bars

Proportionally, Italy's gambling market relies heavily on gaming machines

Dollars lost per resident adult per year, 2017, \$

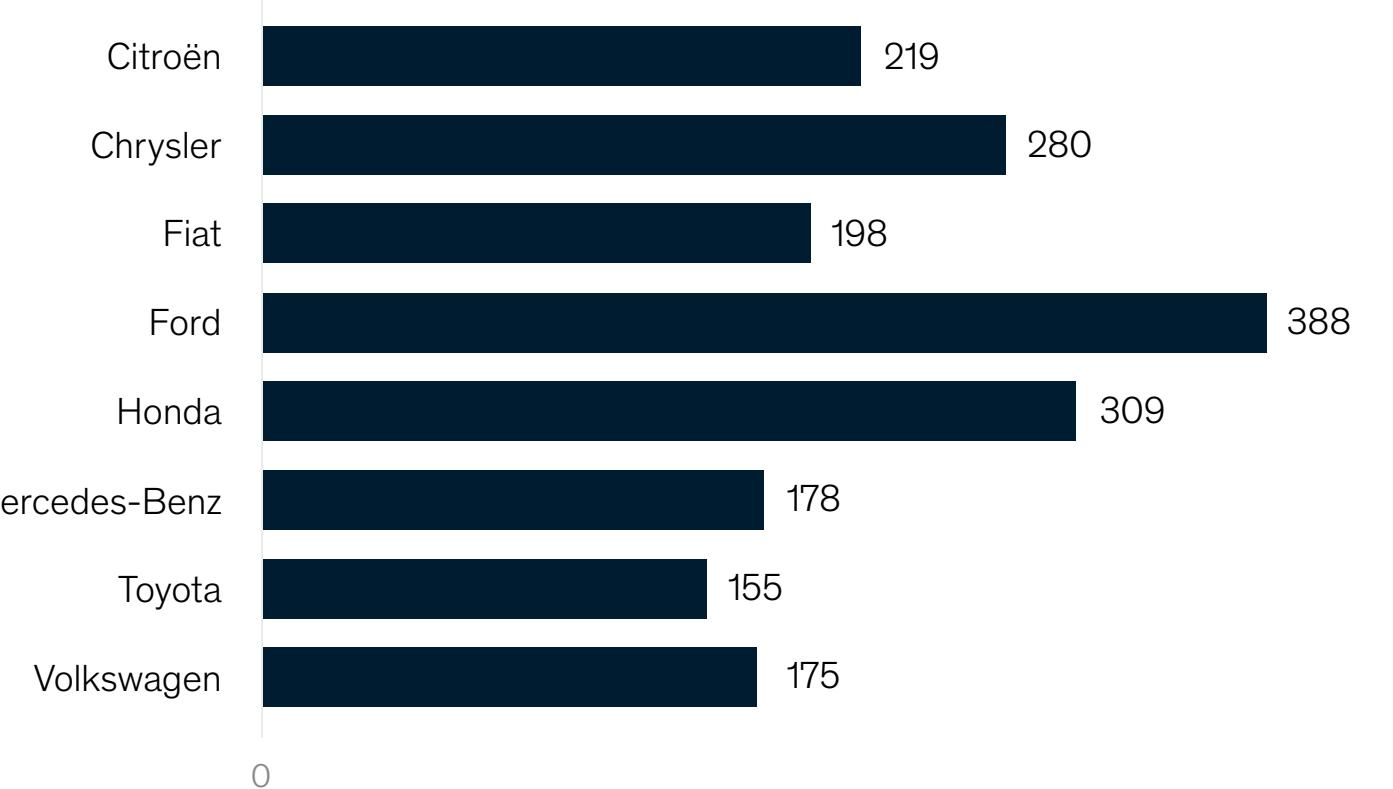


*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Bar chart variants

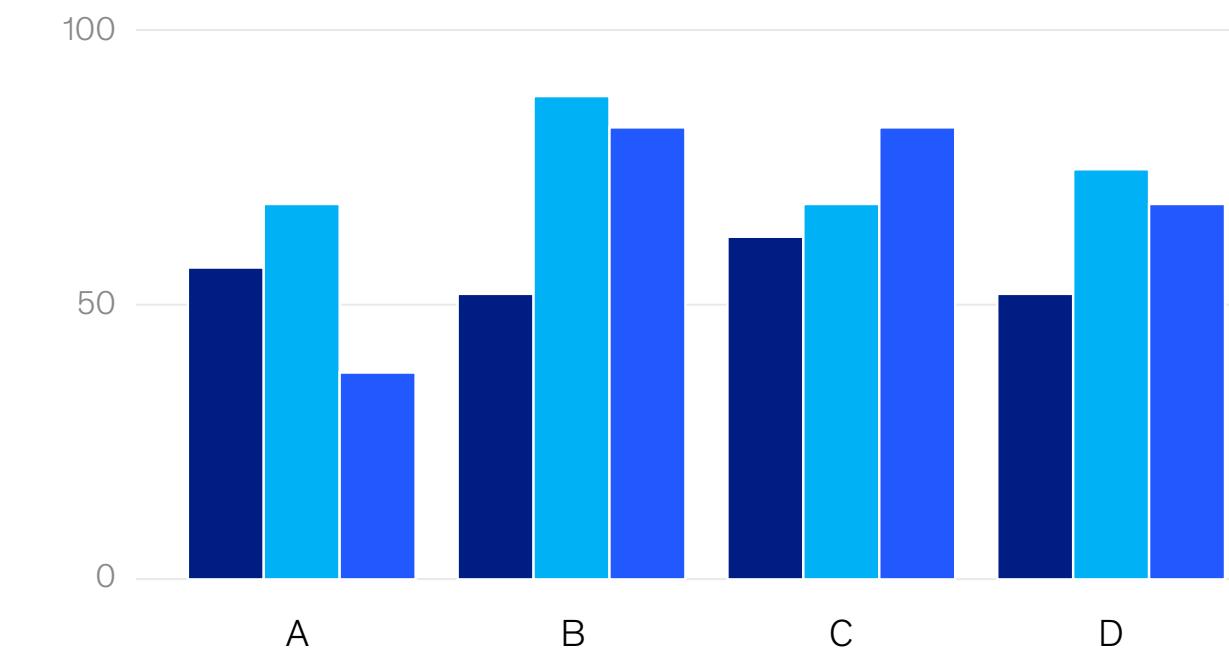
Horizontal bar chart

Horizontal bar charts have some advantages over vertical bar charts. Labels are easier to read, and the whole chart is often easier to format for mobile screens, particularly when there are a lot of bars



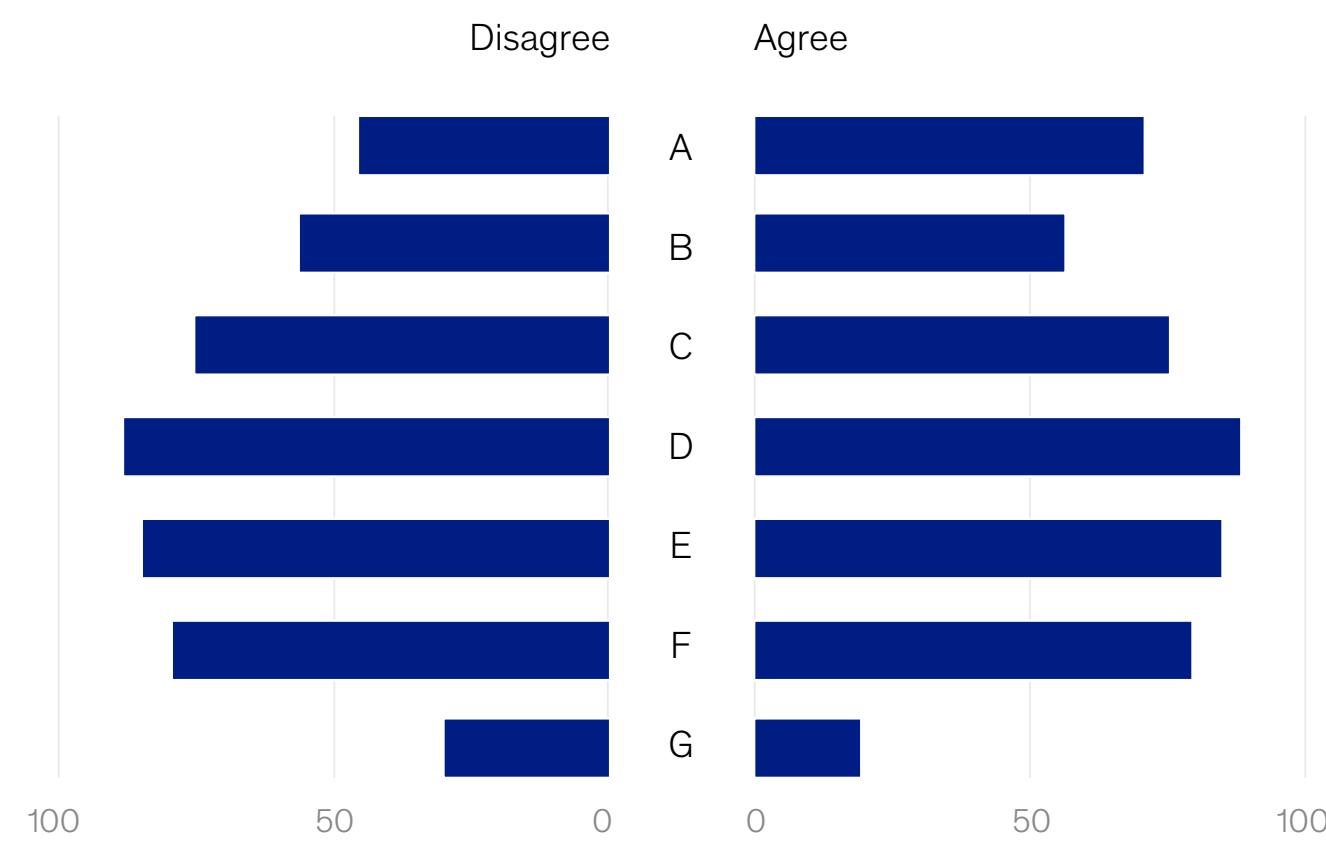
Grouped bar chart

Grouped bar charts can be used when two or more datasets are clustered together and displayed adjacently under categories on the same axis



Butterfly chart

A butterfly chart displays two sets of a data series side-by-side. It shows the difference between two groups with shared parameters. If emphasis is needed on negative values (e.g. bars pointing left), the Crimson Red 500 color can be used



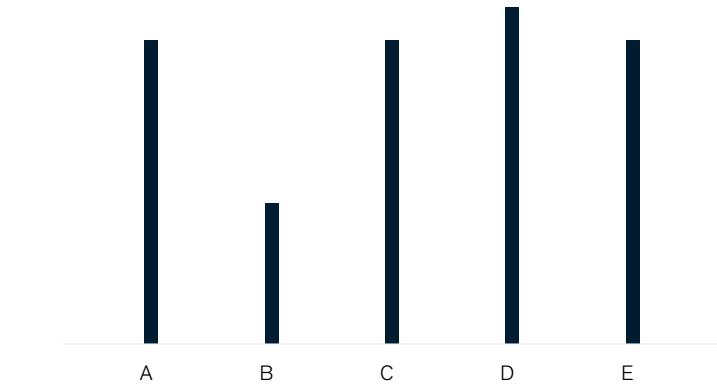
Diverging bar chart

A diverging bar chart is a standard bar chart that can handle both negative and positive magnitude values. If emphasis is needed on the negative values (bars pointing down), the Crimson Red 500 color can be used



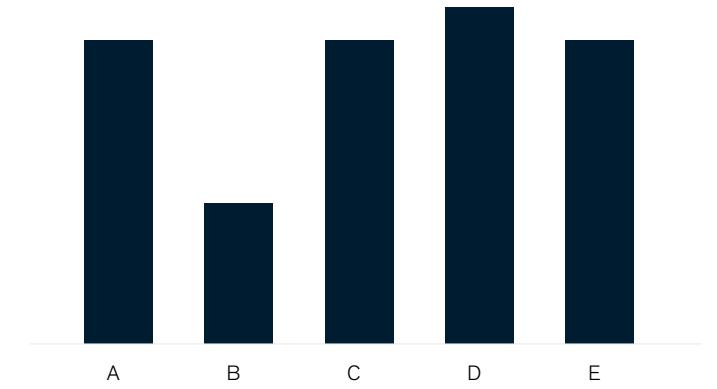
Bar chart best practice

Optimize size and spacing of bars



✗ Try to avoid

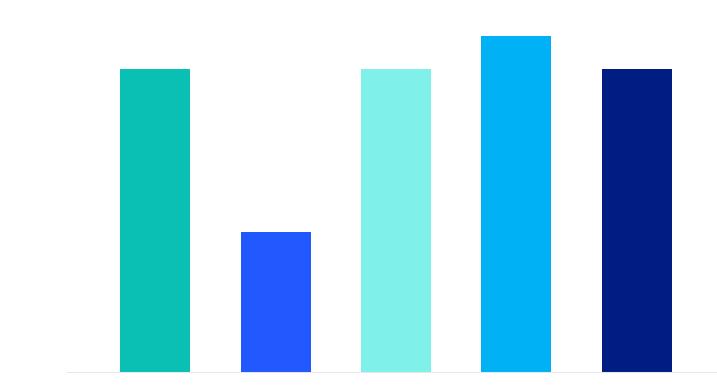
When bars are too narrow your eyes focus on the negative space, which carries no data



✓ Recommended

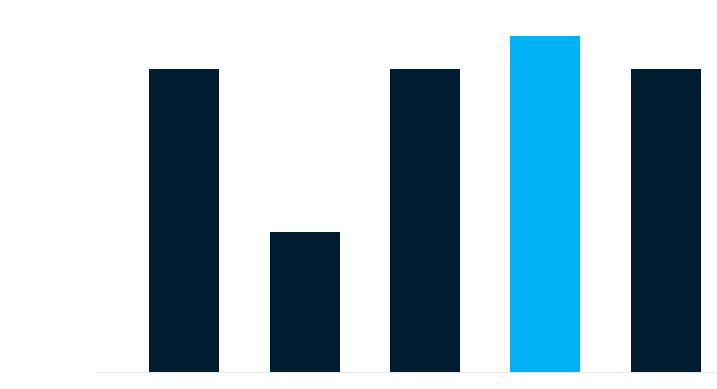
The spaces between each bar should be no bigger than width of the bars

Use color judiciously



✗ Try to avoid

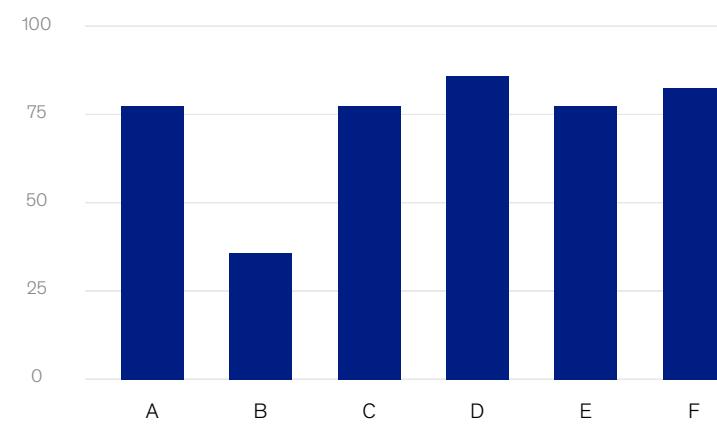
Using different colors for the bars distracts the reader unnecessarily



✓ Recommended

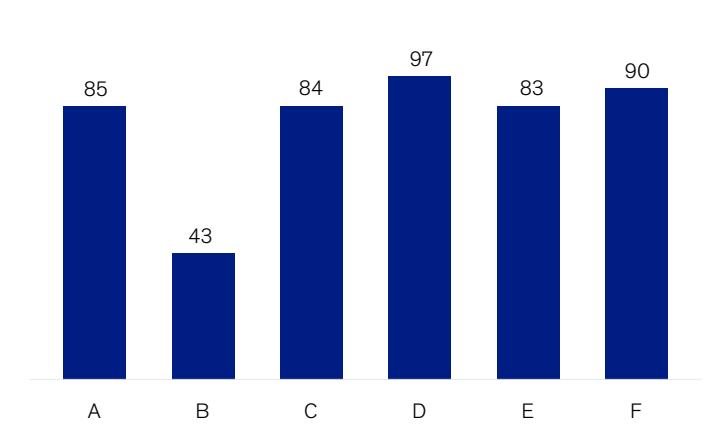
Keep bars the same color and reserve shading for highlights

Data labels are preferred over grid lines



✗ Try to avoid

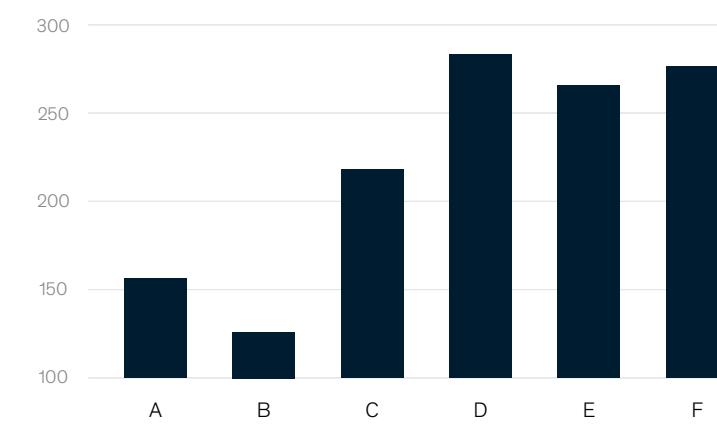
Bar charts often do not need grid axes; they add noise to the chart without adding value



✓ Recommended

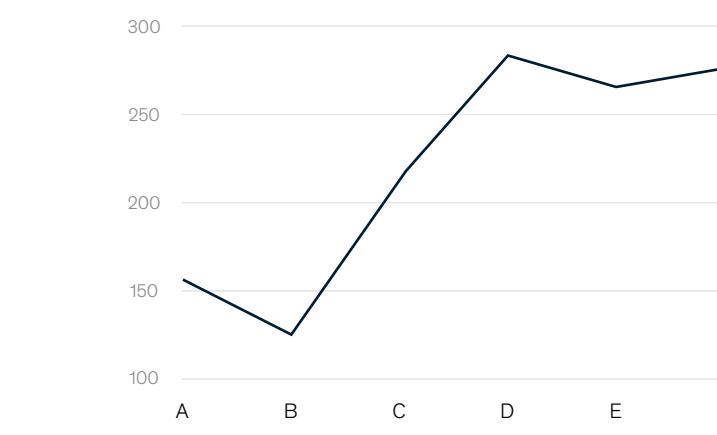
Remove any visual clutter. When possible, use direct labeling and remove grid axes. This will enhance the data's visibility

Always include a zero baseline



✗ Try to avoid

Do not truncate bars. A bar chart that doesn't begin at a zero baseline is misleading because the bar height represents a discrete quantity



✓ Recommended

Switch to a different chart type, such as line chart or scatter plot. They don't need a zero baseline and let the reader compare the data's position and slope

Bubble chart

A bubble chart is used for comparing proportions (size, quantities, etc) to provide a quick overview of the relative size of data.

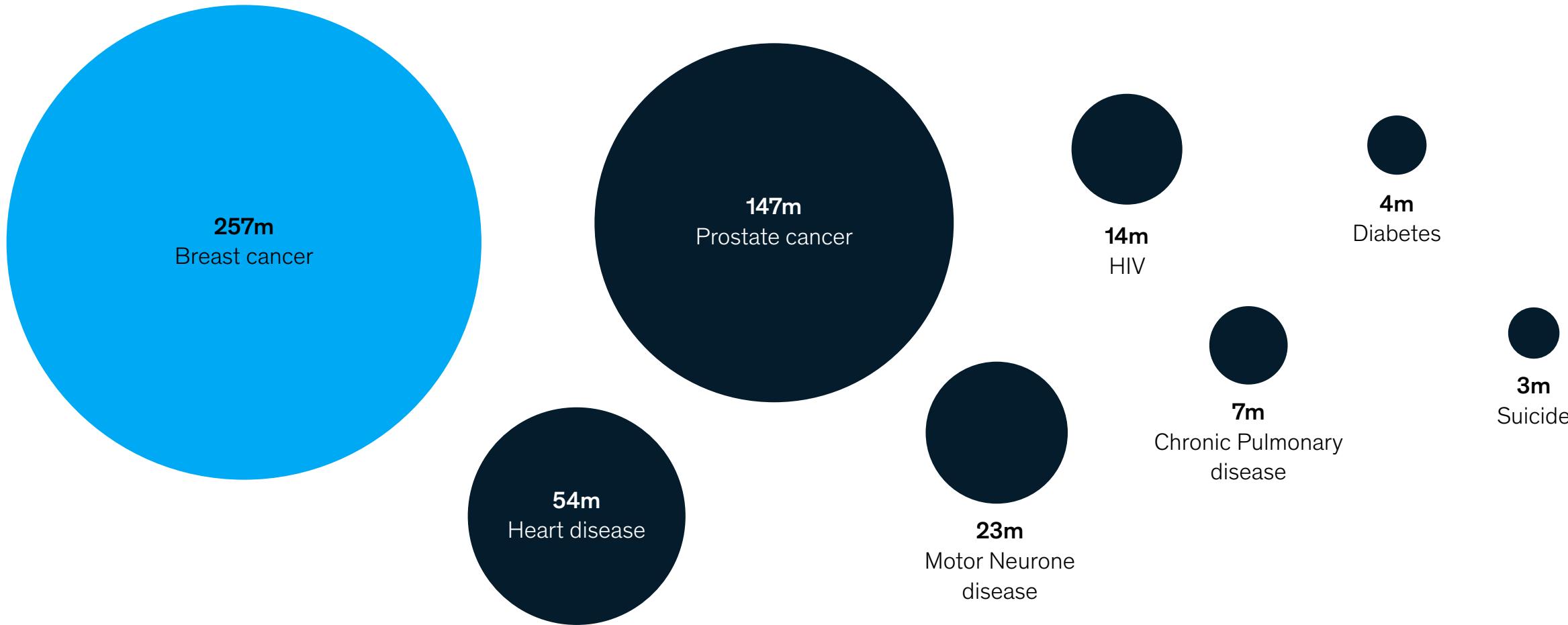
Bubble charts scale data by area instead of length. Use this type of chart when there are big variations between values and it's not important to focus on fine differences between data.

Here, bubbles are ordered from largest to smallest. The order of bubbles can be changed according to layout requirements or chart format

Bubbles are scaled according to area, not diameter. Here, the highlighted bubble is colored with Cyan 500. An alternative version could have the highlighted bubble (Breast cancer) in the Deep Blue 900, and all other bubbles in Gray 30%

Research on breast cancer takes up more resources than all of the budget allocated to other diseases combined

Money raised to fight deadliest diseases, 2017, \$

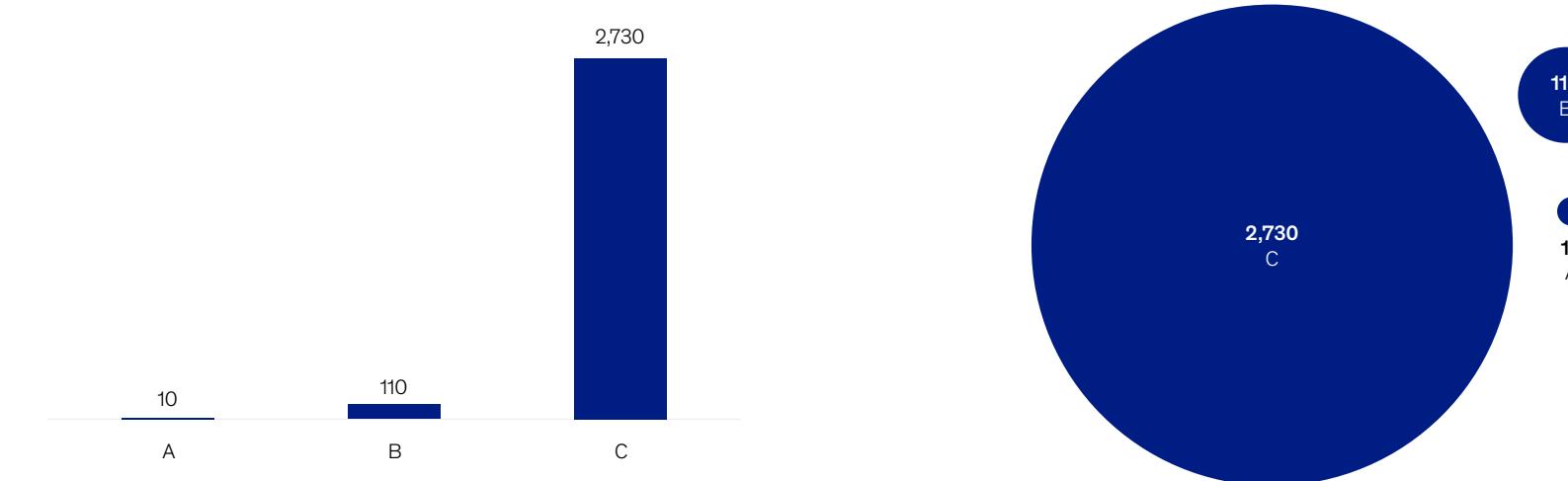


Bubble charts make effective use of direct labeling, as they do not rely on grid axes

*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Bubble chart best practice

When bubbles are better



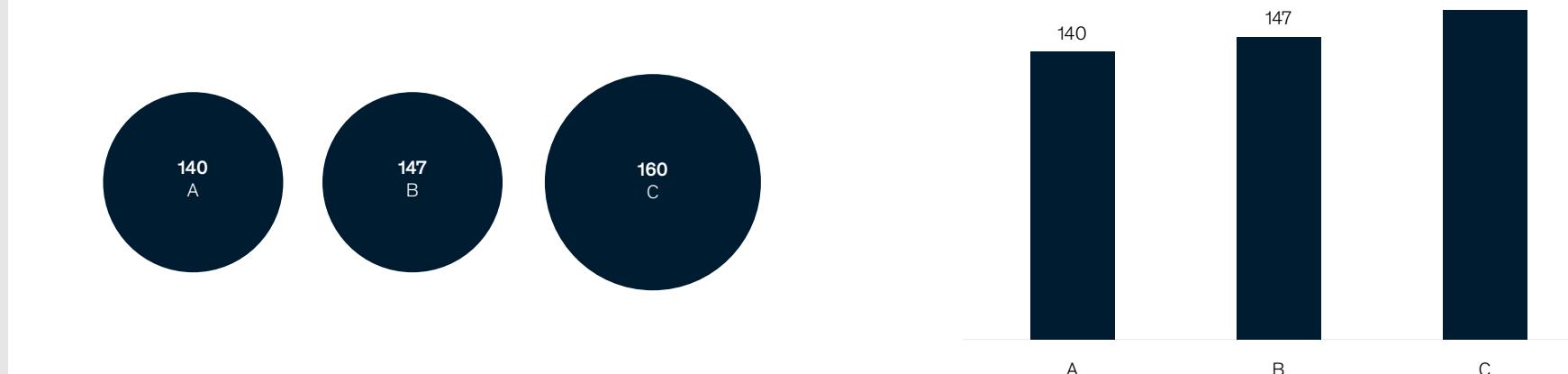
✗ Try to avoid

When values are vastly different, a bar chart may not be helpful for comparisons

✓ Recommended

A bubble chart can be more effective as values are plotted by bubble area, and not by diameter

When bars are better



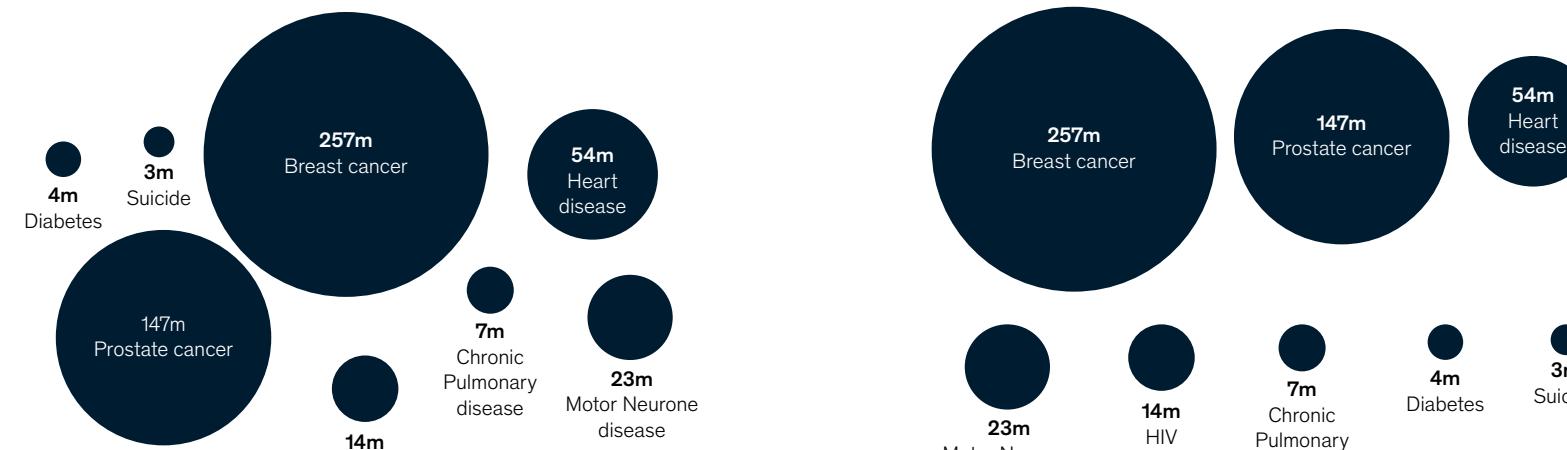
✗ Try to avoid

When dealing with data points of similar values, bubble charts may not be the most suitable device as area-scaled symbols can hinder comparisons

✓ Recommended

Bar charts handle small value differences better as they are scaled by length

Visually organize bubbles



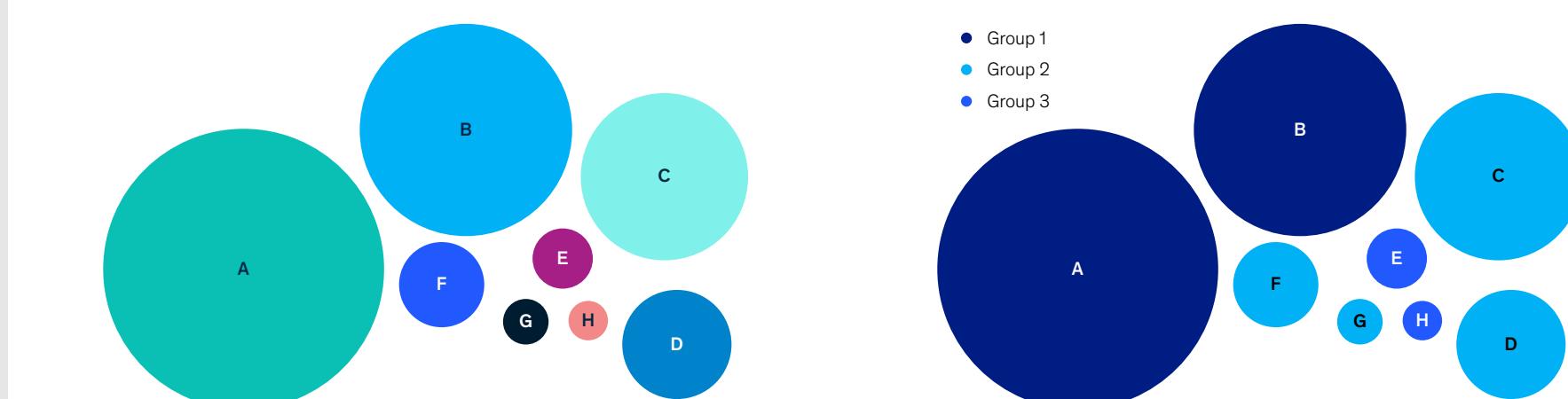
✗ Try to avoid

Layouts with little functional purpose could hide important insights

✓ Recommended

Try sorting bubbles by size to highlight the biggest and smallest values, and facilitate comparisons

Make meaningful color choices



✗ Try to avoid

Don't apply color for nonfunctional or styling purposes

✓ Recommended

Grouping data by color-coded categories can help the viewer establish relationships between the data points

Part-to-whole

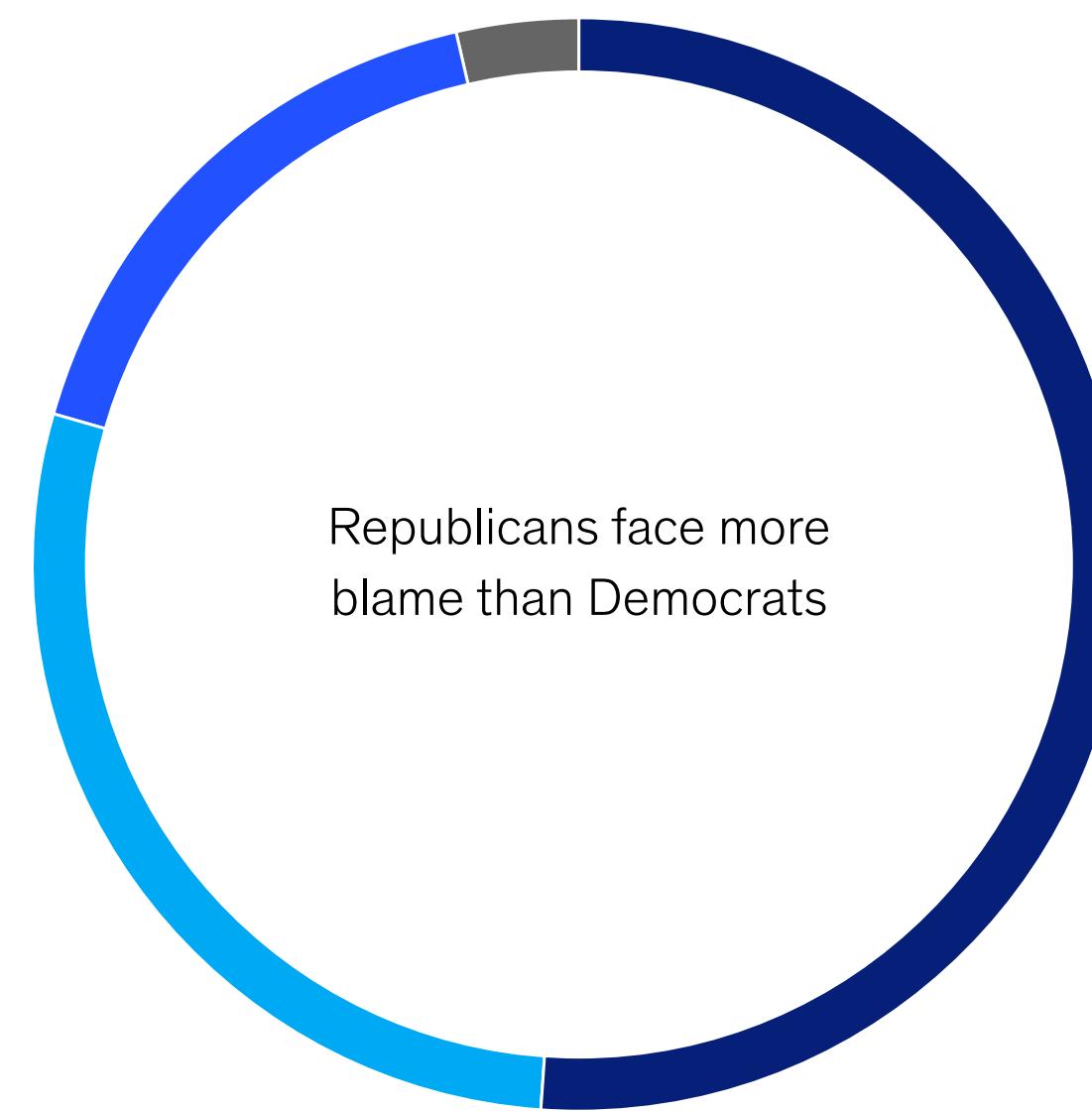
Donut and pie chart

When looking at the composition of data, donut charts are an effective way to show the relationships of the parts to the whole. It's best to visualize no more than five categories per chart.

A donut chart is functionally identical to a pie chart, starting at 0° (pointing straight up). Donut charts can sometimes provide a better data intensity ratio to standard pie charts. The blank center can be used to display additional, related data.

Wedges are stacked in decreasing order, from largest to smallest. The call-out fact in the middle can add contextual information about the main insight

Who do people blame most for what's gone wrong in Washington? Distribution of preferences, 2017 (results from a survey of 906 adults)



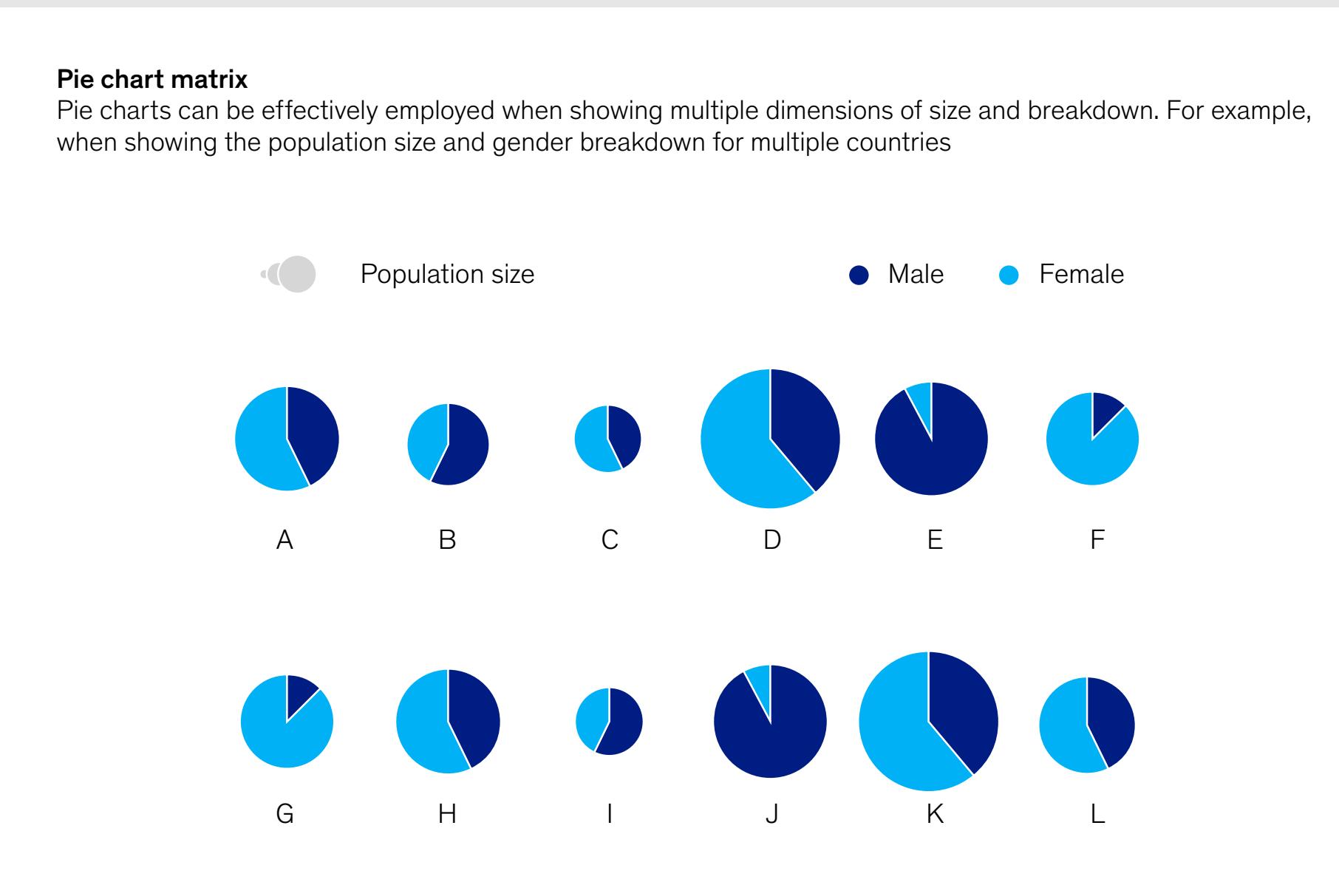
● Republicans in Congress	51%
● President Obama	28%
● Democrats in Congress	17%
● Not sure	3%

Labels and values are neatly placed outside the chart in a meaningful order

*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

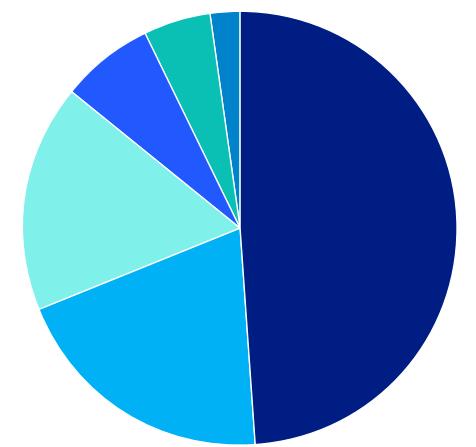
Part-to-whole charts focus on the distribution of the parts and show how a single entity can be broken down into its component elements.

Donut and pie chart variants



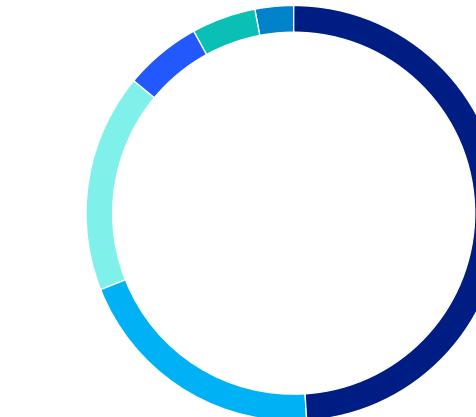
Donut and pie chart best practice

Donuts are often more effective than pies



✗ Try to avoid

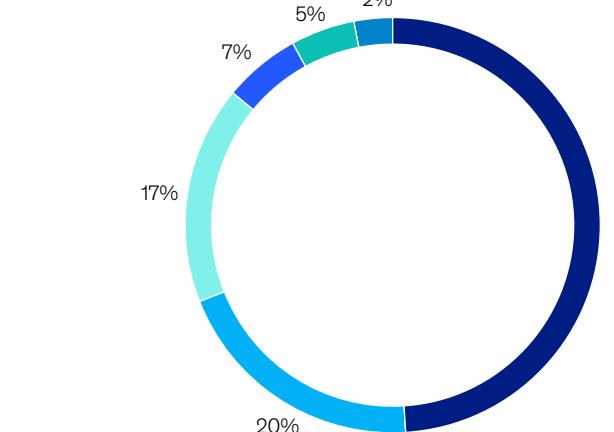
Our brains understand differences in length faster than differences in area. Pie charts can sometimes make interpretation difficult



✓ Recommended

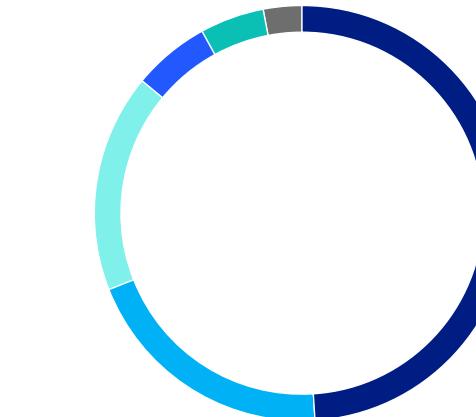
A donut chart allows the viewer to read around the circumference of the circle. They can more quickly gauge the 'length' of each segment compared to the last

Labeling donut charts



✗ Try to avoid

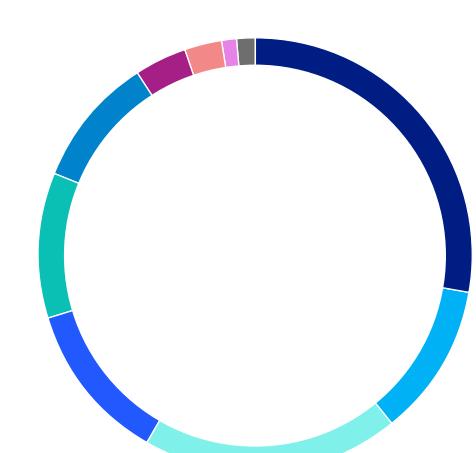
When working with more than a few values, do not label values directly on the chart



✓ Recommended

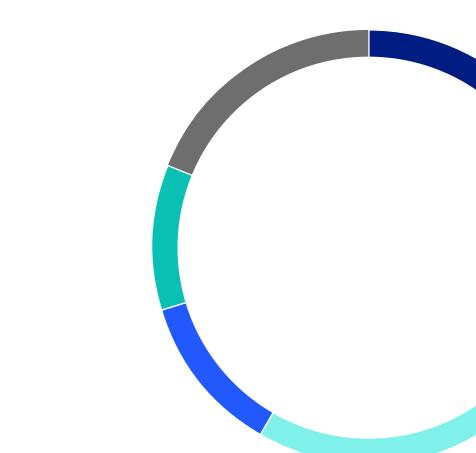
Consider labeling smaller segments outside the chart, especially if you have long labels

Dealing with many small wedges



✗ Try to avoid

Avoid donut charts with a large number of sections, as it compromises legibility



✓ Recommended

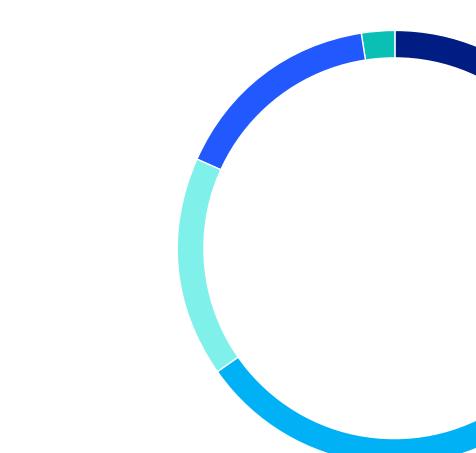
Try grouping the smallest segments into a new category called "Other"

Order wedges logically



✗ Avoid

Do not arbitrarily mix small and large value segments without a functional purpose



✓ Recommended

It's best to start with the largest values, then move clockwise to smaller values

Waterfall chart

A waterfall chart helps in understanding the cumulative effect of sequentially introduced positive or negative values.

At its core, the waterfall chart is a stacked bar chart with a visible and an invisible series of bars. The visible series is the data that is shown. The invisible series acts as a placeholder that hides the previous amount, to help the viewer understand changes in the distribution.

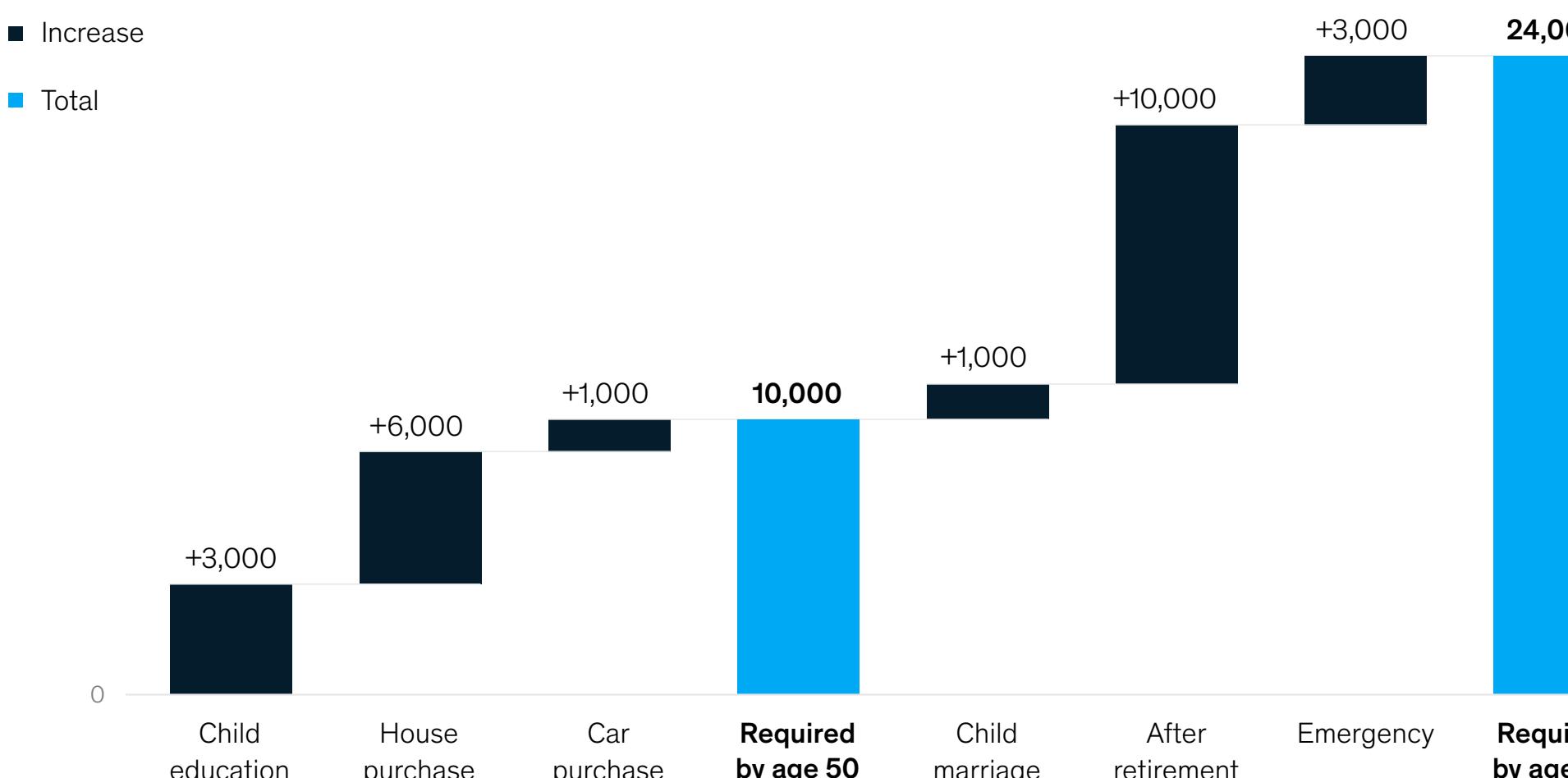
Ideally, the waterfall chart X axis has a sequence to it to help tell the “cumulative story” even more effectively.

Incremental bars are displayed using the Deep Blue 900. The bars showing the full amount required are colored with the Cyan 500

Connecting lines help the viewer to see the increments. These strokes avoid distracting or adding clutter to the chart

Saving for retirement

Breakdown of financial needs in the future, 2017, GBP



Saved away

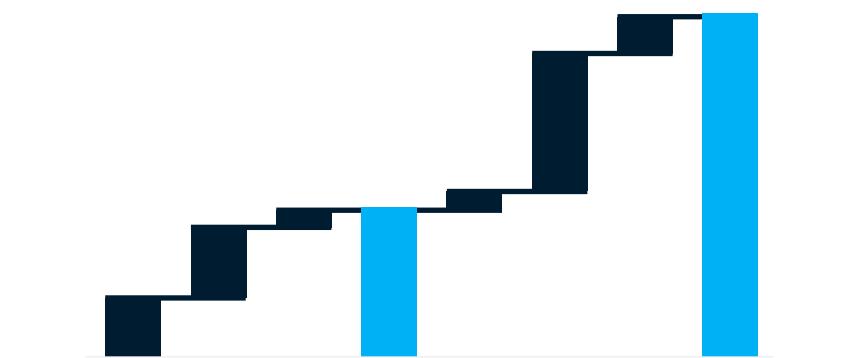
An additional £14,000 required in the last 30 years, makes planning for future needs harder once past 50 years of age.

Insights are placed alongside the tallest bars

*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

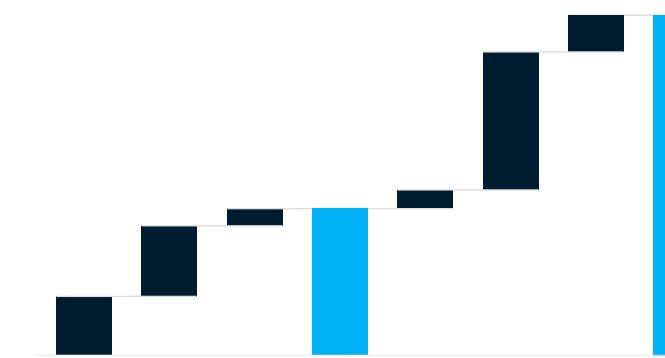
Waterfall chart best practice

Style for connecting line



✗ Try to avoid

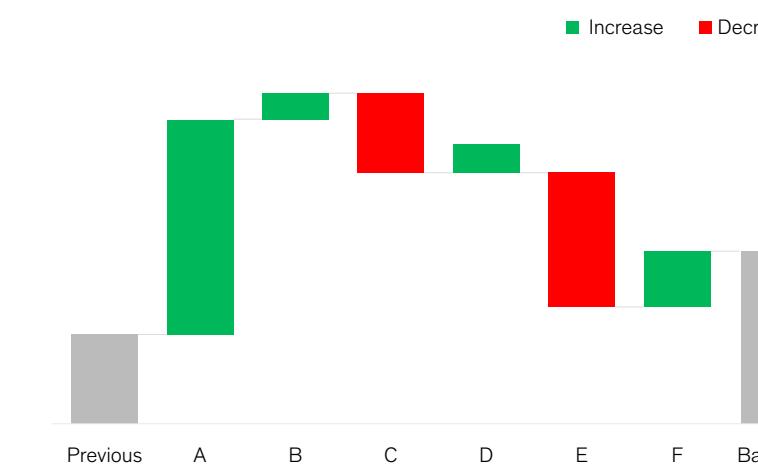
Avoid drawing the lines connecting the bars with a heavy stroke. This interferes with the ability to make comparisons



✓ Recommended

A thin, light gray connecting line allows the bars to stand out

Make accessible color choices



✗ Avoid

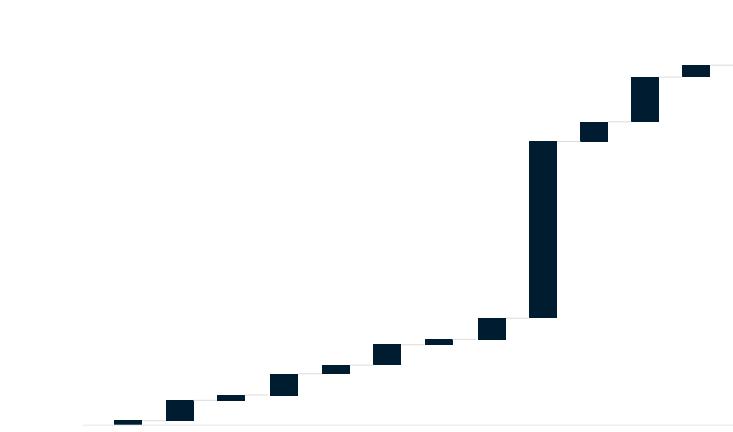
Do not use color combinations from the PowerPoint standard palette, such as basic reds and greens, which are indistinguishable to many colorblind readers



✓ Recommended

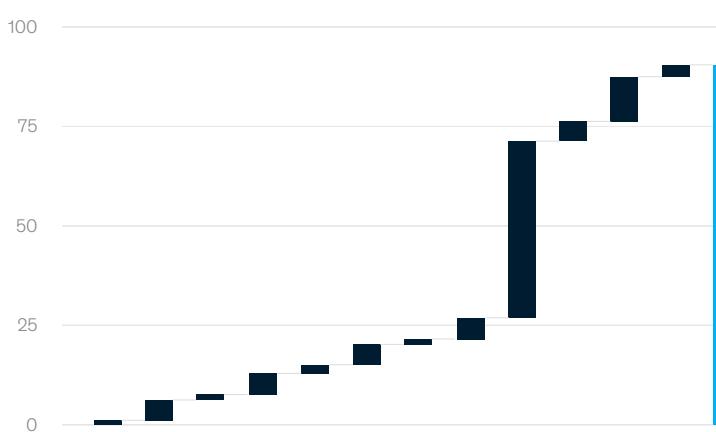
When it's relevant to attract attention to the increment or decrement, colors within the McKinsey diverging palette can be used to shade positive and negative values

Dealing with many small increments



✗ Try to avoid

If there is an apparent difference in size, and the specifics aren't important, direct labeling may be unsuitable as the scale would make the numbers difficult to read



✓ Recommended

In those cases, it may be best to omit direct labeling and add labels to the axis

Consider a horizontal format



✓ Recommended

Horizontal waterfall charts have some advantages over vertical waterfall charts. Labels are easier to display and they tend to work better in a narrow layout such as mobile view

Treemap

Treemaps display hierarchical (tree-structured) data as a set of nested rectangles. Each branch of the tree is given a rectangle. Each rectangle can also be tiled with smaller rectangles representing sub-branches.

Each rectangle has an area proportional to a specified value in the data.

Rectangles are usually colored to show a separate value or category of data.

Categories are sorted in order of size from top left to bottom right. Rectangles within each category are also sorted in the same way

The area of the rectangles represents value in a treemap. Direct labeling is applied within the rectangle

The US is leading the investment market Investments globally, 2017, \$

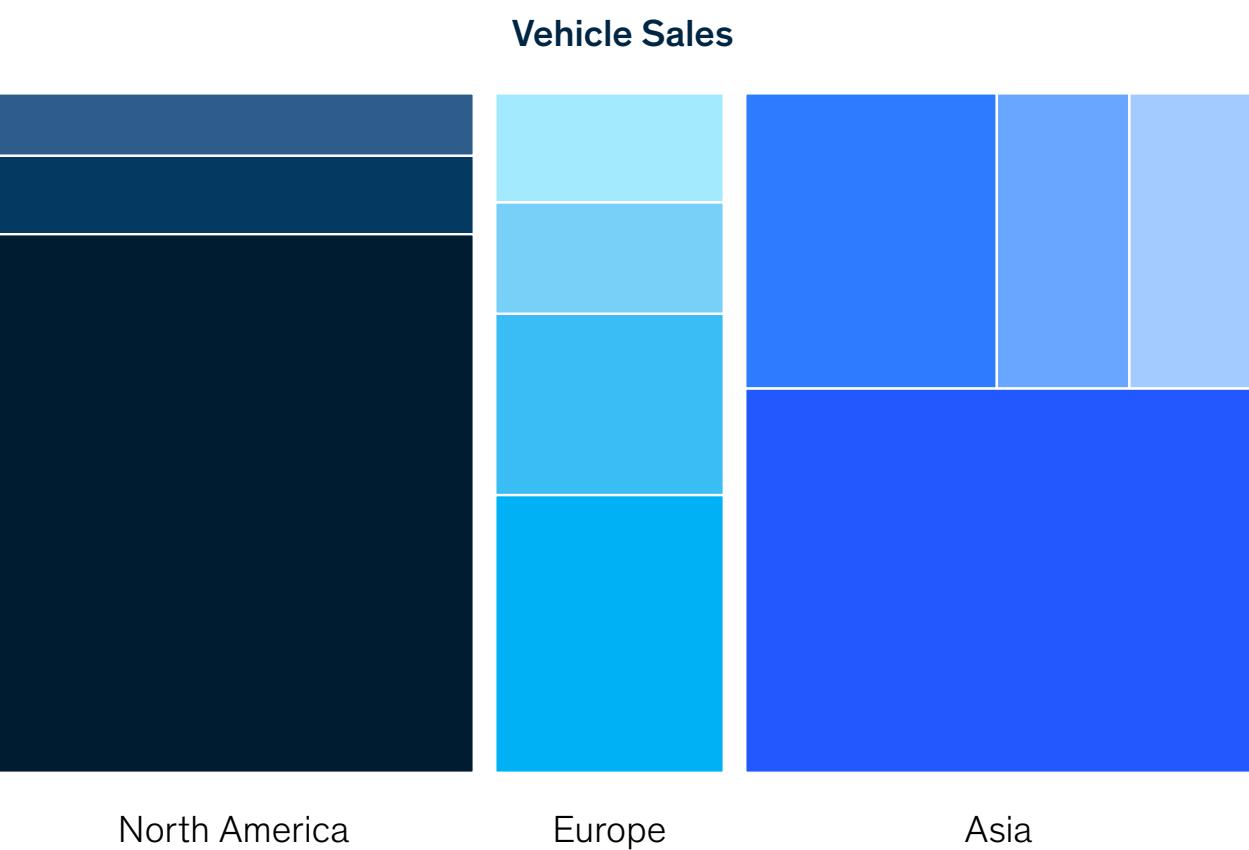


*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Treemap variants

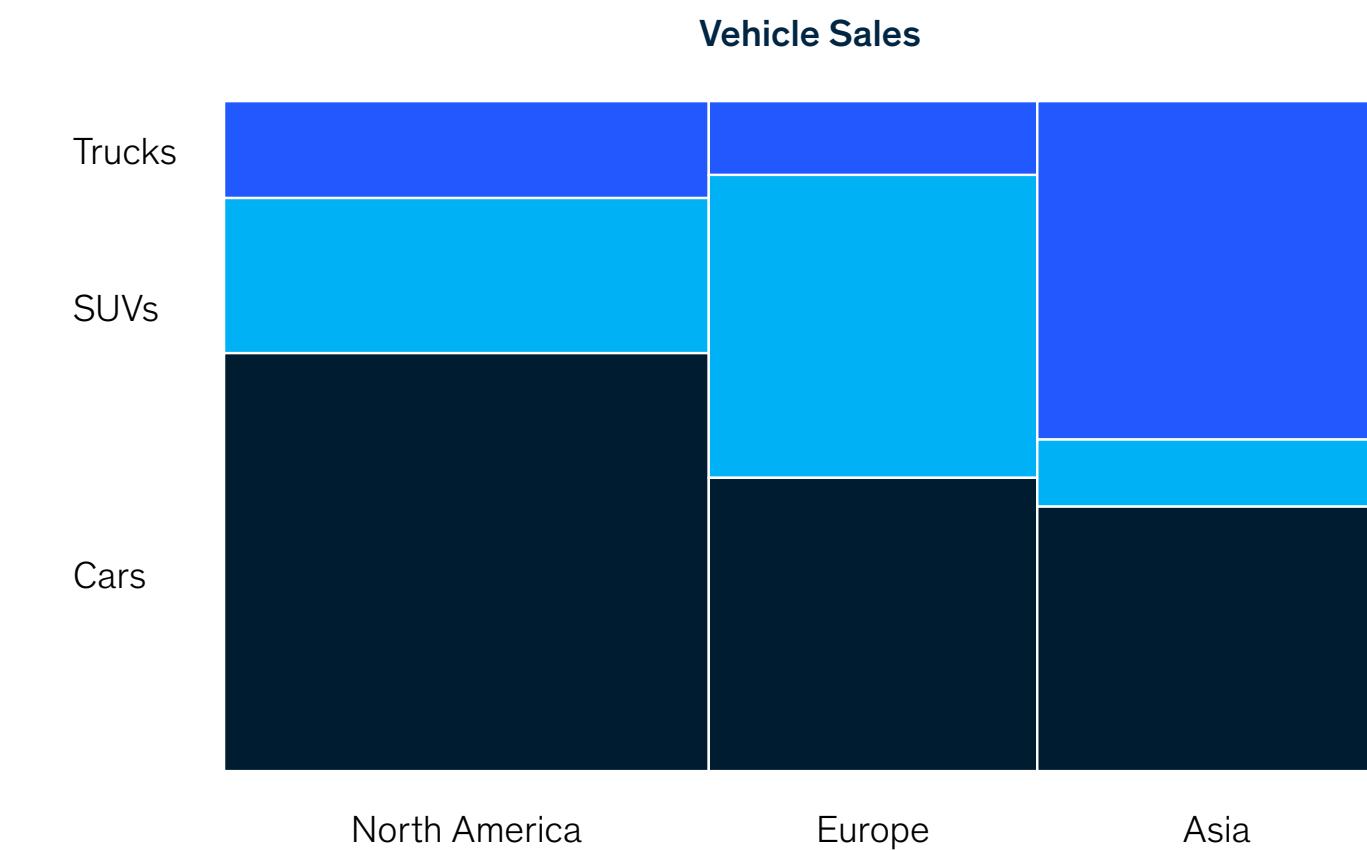
Geographic treemap

This variant of the treemap plots rectangles on a position that is representative of where world regions are



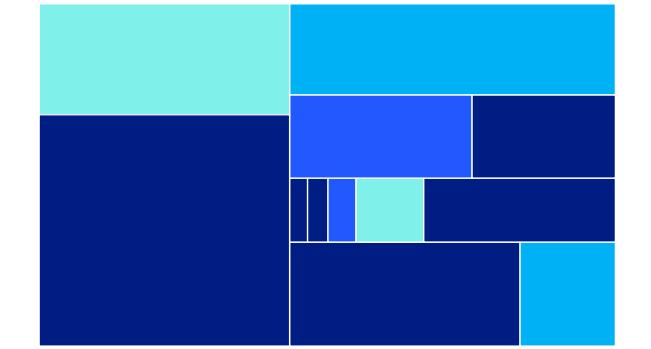
Marimekko chart

The Marimekko chart is a two-dimensional stacked chart. In addition to the varying heights of a regular stacked bar chart, a Marimekko chart also has varied column widths



Treemap best practice

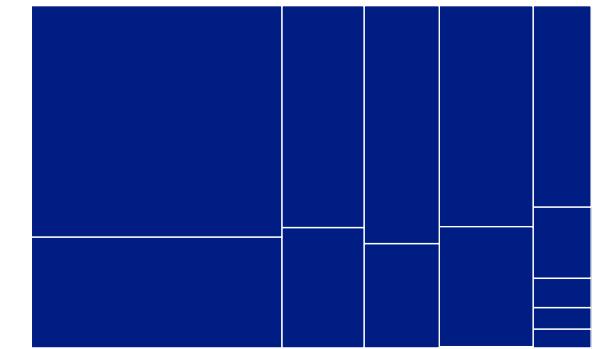
Arrange the treemap logically



■ A
■ B
■ C
■ D

✗ Try to avoid

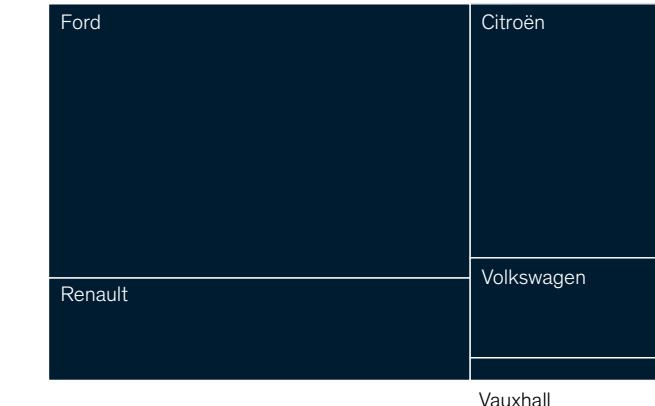
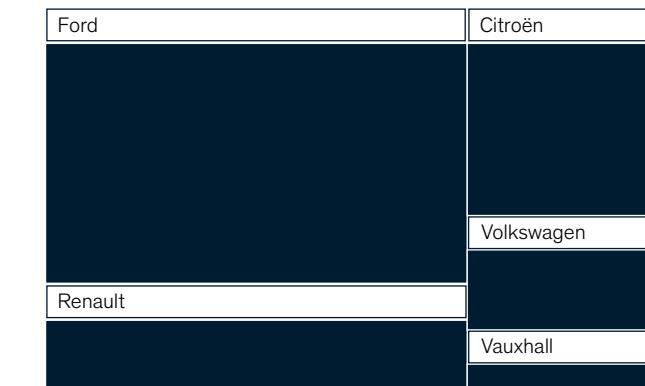
Avoid organizing the treemap rectangles without a functional purpose or in ways that hinder comparison



✓ Recommended

Treemaps should follow a flow from largest to smallest rectangles

Keep the treemap whole



✗ Try to avoid

When labeling rectangles, do not create new boxes for text. This approach distorts the treemap, making it hard for the viewer to visually compare sizes

✓ Recommended

It's better to label directly within the rectangles, or to add labels outside the rectangles, if they are too small

Use consistent padding and sizing



✗ Try to avoid

Avoid sizing the rectangles inconsistently, leaving empty gaps or uneven padding among the blocks, as it is distracting

✓ Recommended

Consistent padding makes the overall look of the treemap more cohesive and organic, facilitating comparisons and data exploration

Trend over time

Line chart

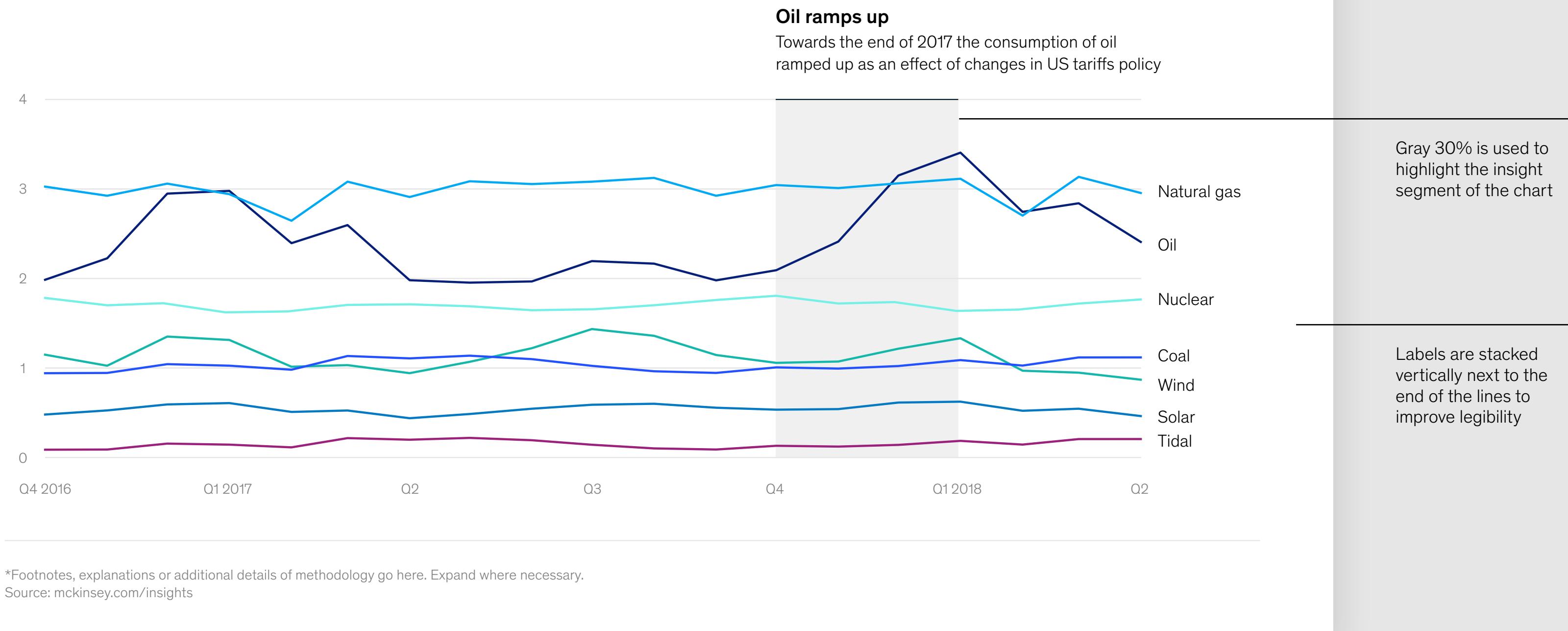
Line charts show information as a series of data points connected by line segments. Line charts show how a particular dataset changes at equal intervals of time. They are effective at showing trends over time.

Instead of highlighting discrete quantities (like bar charts do), line charts call attention to the changes in value from point to point in a data series.

Trend over time charts compare changes over the same time period for a series or a group of series.

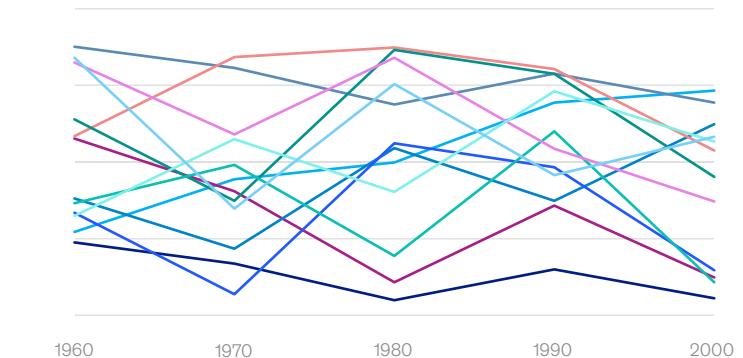
While most energy source consumptions are steady, oil is more volatile as a consequence of economic shifts

US primary energy consumption, 2016-2018, quadrillion BTU



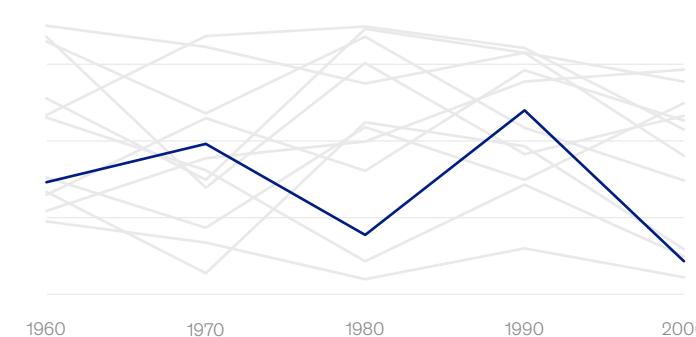
Line chart best practice

Highlight important lines



✗ Try to avoid

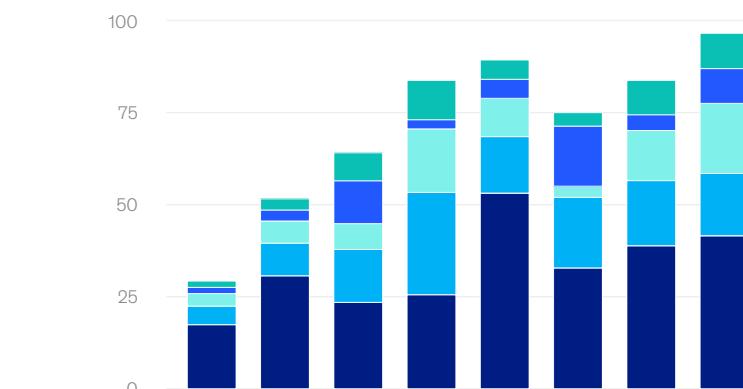
While it may be important to show a large number of data points, having numerous color-coded categories does not help spotting trends



✓ Recommended

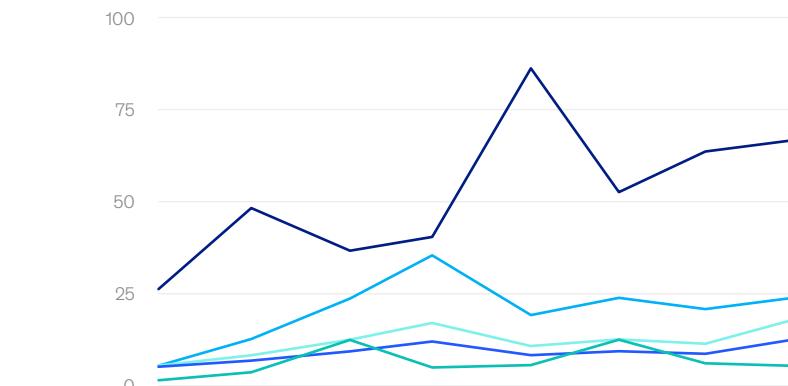
Graying out the background noise and using a highlight to show insights and trends may help the reader to focus on the message

Understand when to show trends



✗ Try to avoid

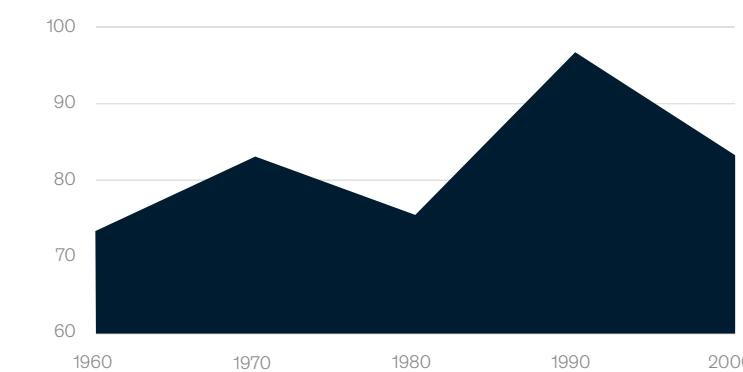
While a stacked bar chart may accurately plot discrete quantities, it doesn't effectively show the incremental change of values (eg growth of sales)



✓ Recommended

Depending on the authors' message, a trend line may be a more important insight than incremental totals

Reserve shading for area charts



✗ Avoid

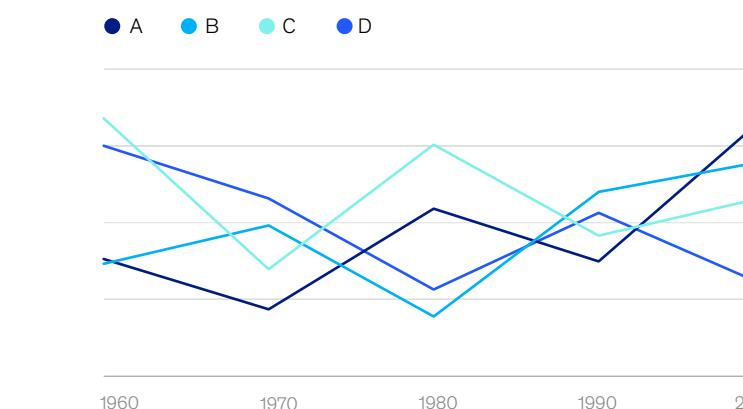
Avoid shading below a line unless the chart has a zero baseline. Filling in below a line turns a line chart into an area chart



✓ Recommended

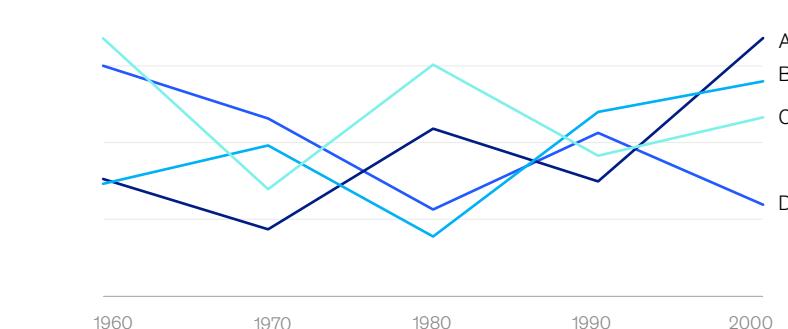
When showing a range where the Y axis does not start at zero, it's important to use line charts and not area charts, which are more suitable for discrete quantities

Prefer direct labels over legends



✗ Try to avoid

Placing the key above the chart can break the reading order of categories and makes it difficult to reference when reading a line chart

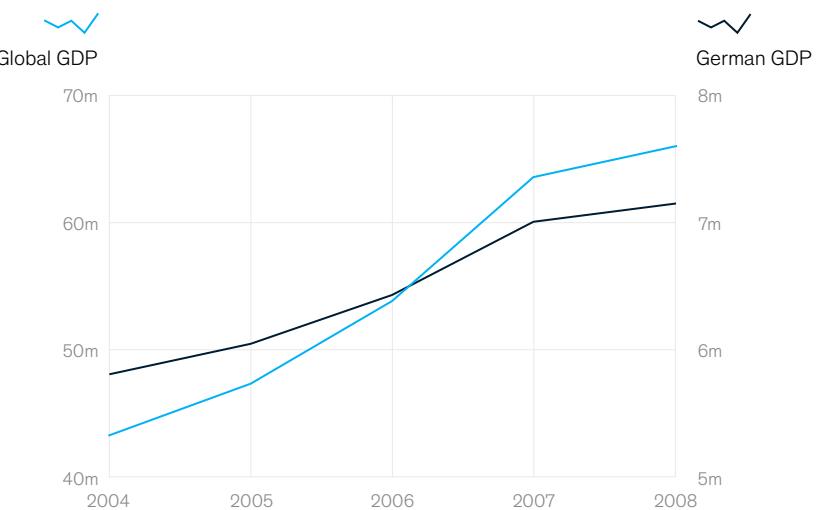


✓ Recommended

Where possible, label lines directly. If there are too many lines to label, or lines are too close, stack the legend vertically on the right side

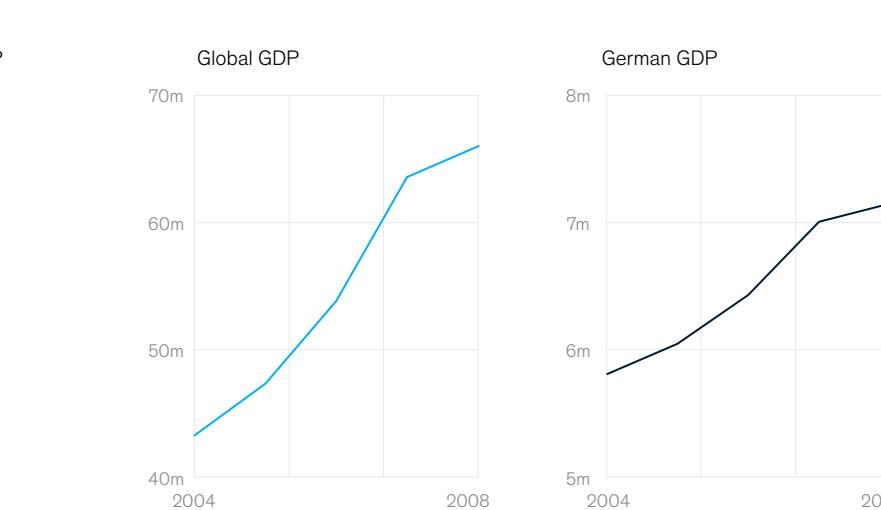
Line chart best practice

Avoid dual axis charts



✗ Try to avoid

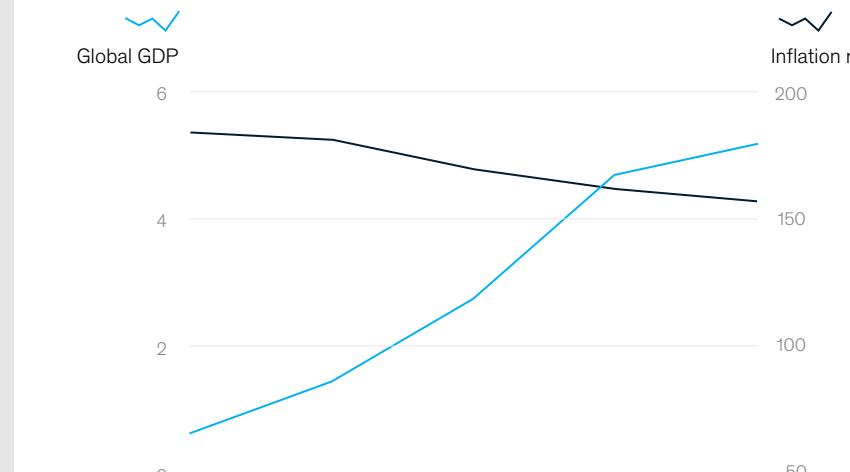
The scales of dual axis charts are arbitrary and can therefore mislead readers about the relationship between the two data series



✓ Recommended

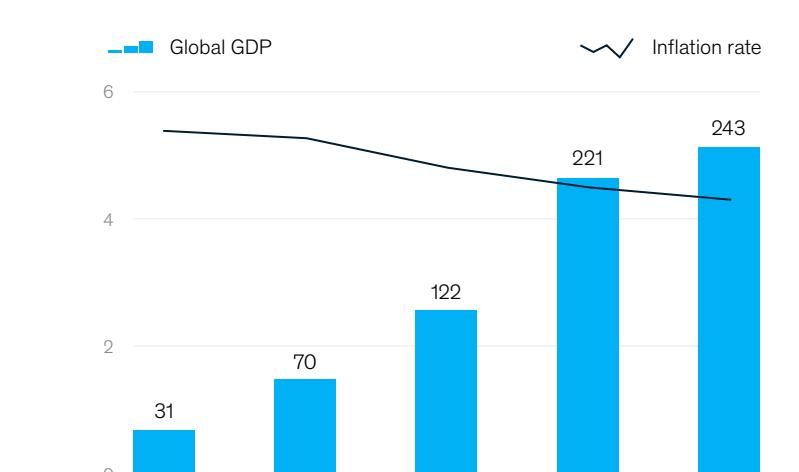
It is typically best to split dual data sets into separate charts side-by-side, so that the two scales can be fully clear to the reader

When mixing lines and bars can work



✗ Try to avoid

When dealing with discrete vs non-discrete quantities on different scales, plotting both as line charts is not the best option



✓ Recommended

When comparing discrete quantities (such as GDP) vs trends, indexes, or rates, it may be worth combining a trend line with a bar chart

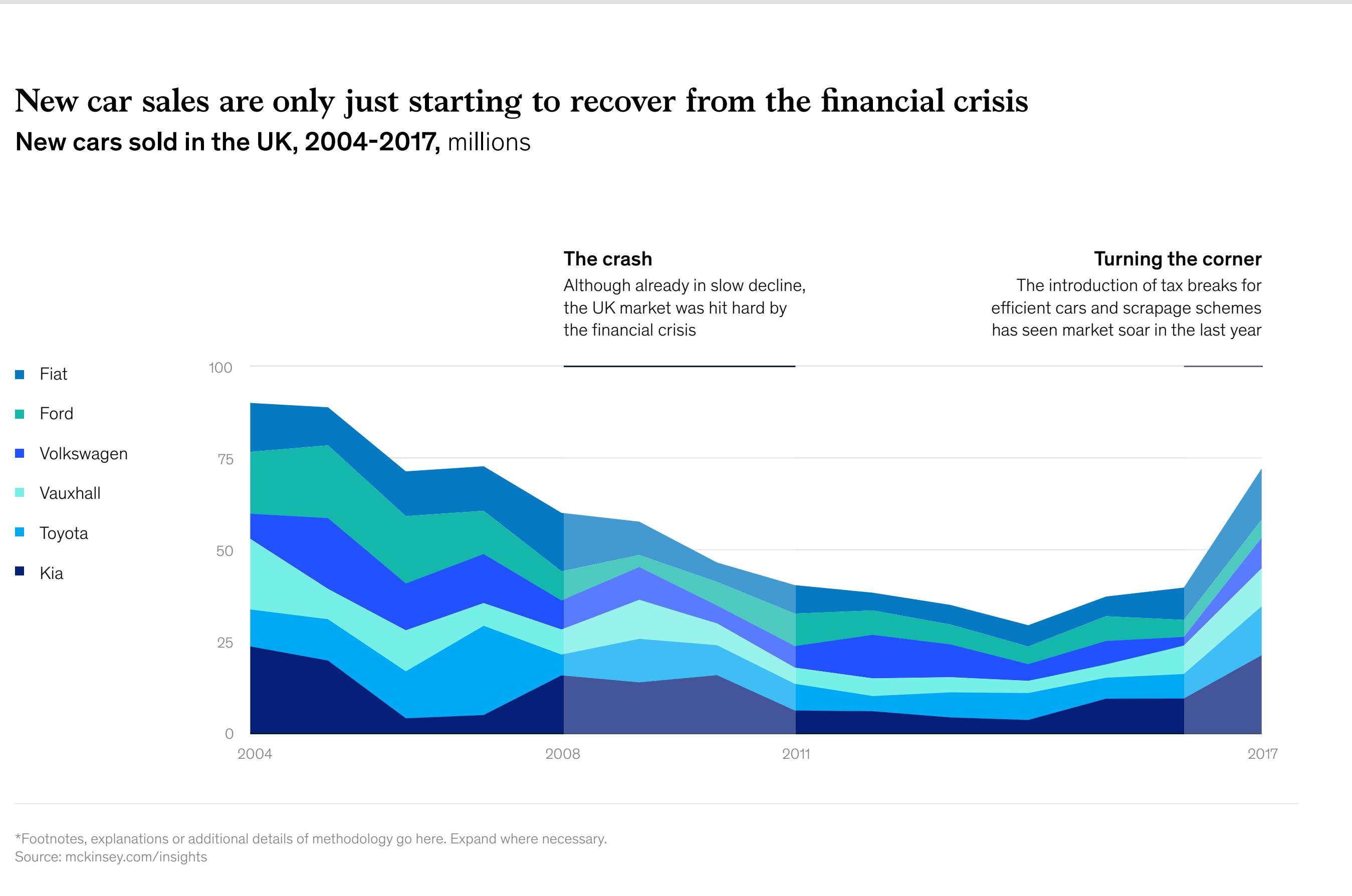
Stacked area chart

Stacked area charts use multiple data series that start each point from the final point in the previous series. The stacked area chart is effective when comparing multiple variables changing over intervals.

The distinctive feature of a stacked area chart is that the filling done between the segments and the X axis emphasizes magnitude. When looking at a stacked area chart, the viewer can make out how much the single segments contribute to the whole.

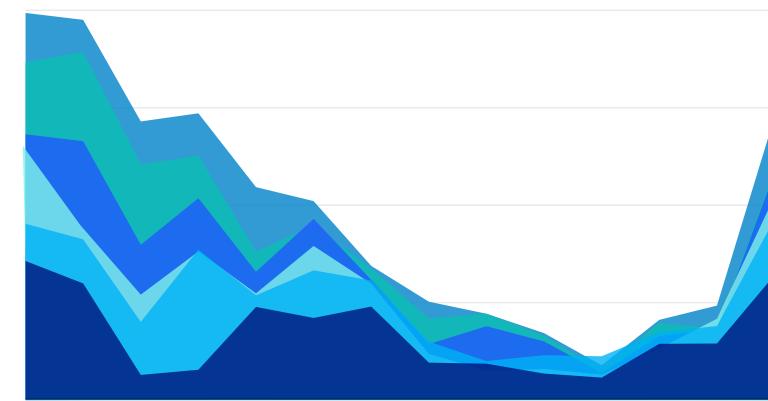
The color of the stacked segments follows the order of the McKinsey categorical palette, starting with the Electric Blue 900 (at the bottom of the stack)

Lightly shaded areas of the chart highlight the relevant insights



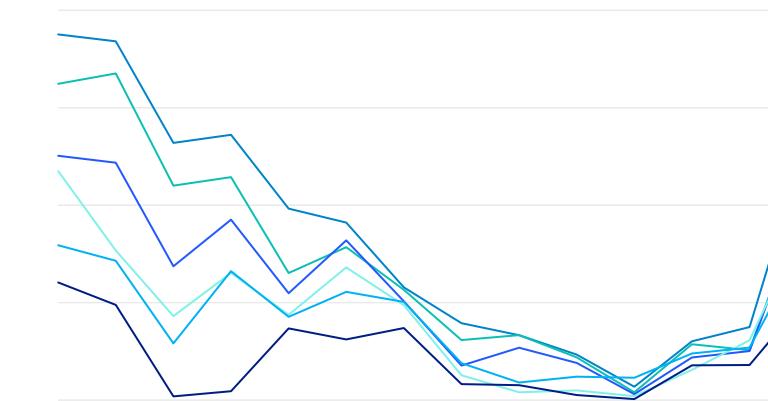
Stacked area chart best practice

Stack segments, rather than overlapping



✗ Try to avoid

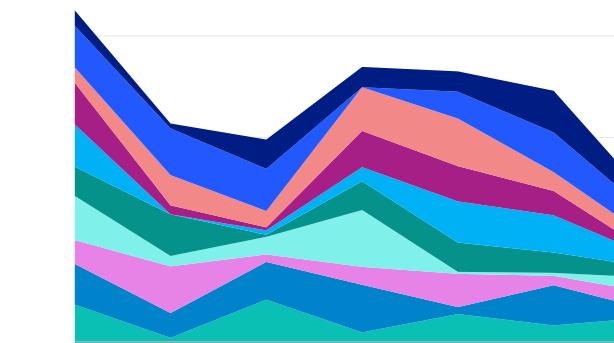
Overlapped data plots can be difficult to identify when on an area chart. The segments in front will obscure and hide the others



✓ Recommended

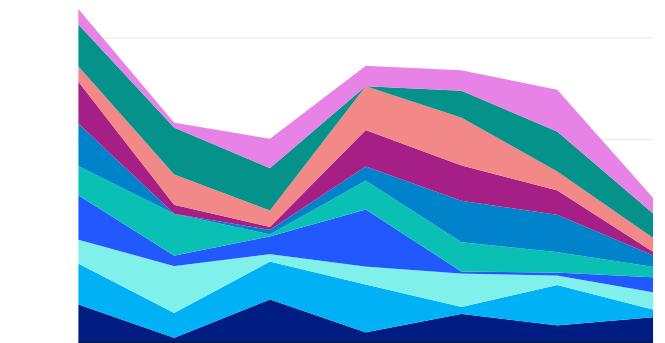
It's better to either stack the segments using a stacked area chart or resort to a line chart

Follow color palette order



✗ Try to avoid

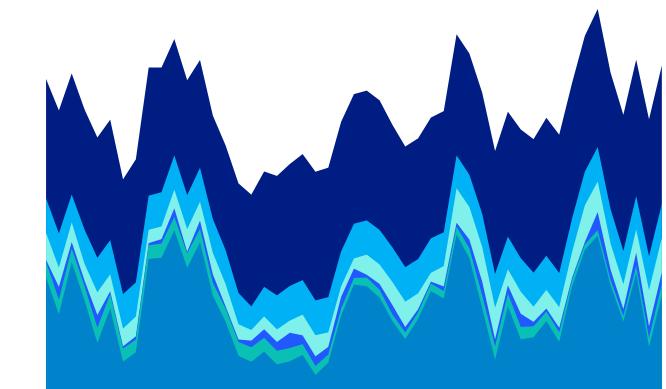
The order of colors applied to segments should not be arbitrary. There is a risk that colors next to each other are not distinctive enough, hindering legibility



✓ Recommended

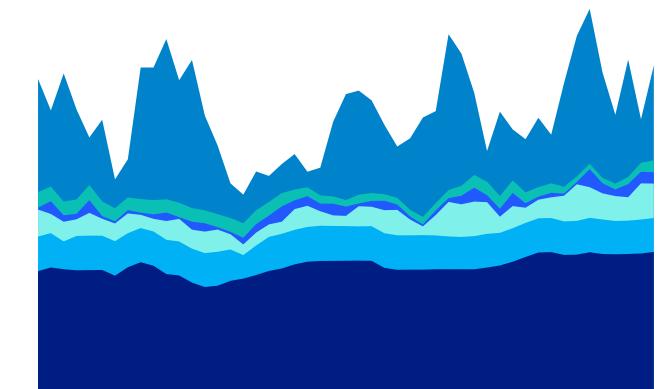
The order of colored stacks should follow the recommended qualitative palette so that colors are distinctive and legible, even when the chart is made of ten categories

Reorder segments to highlight variability



✗ Try to avoid

Avoid arranging data that puts categories with high variability at the bottom



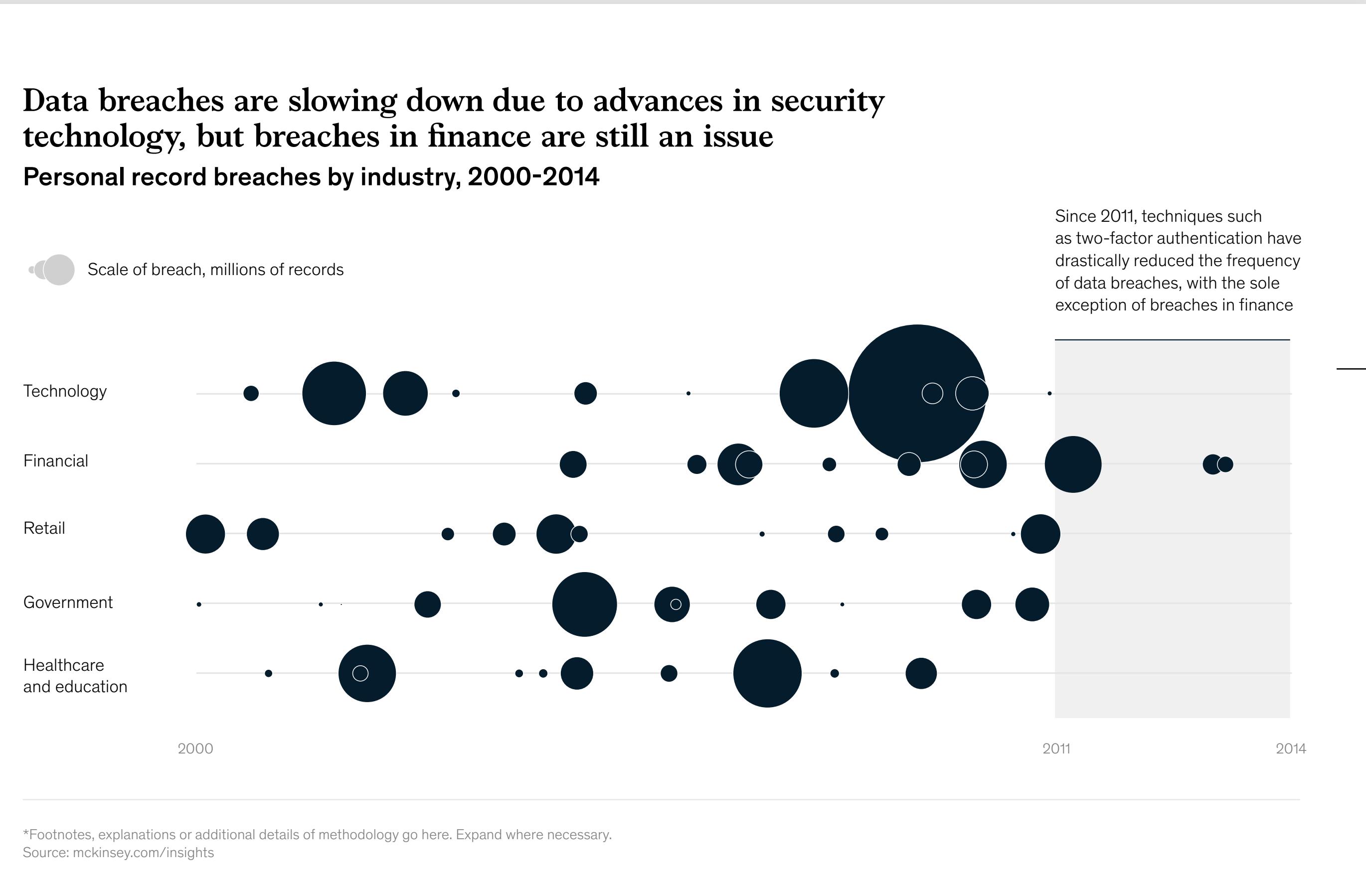
✓ Recommended

Arrange data to put categories with highly variable data on the top of the chart and low variability on the bottom

Bubble timeline

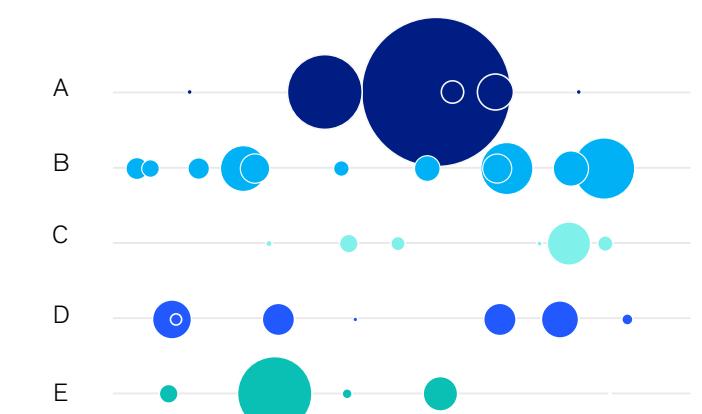
A bubble timeline is a way to display a set of events or items on a timeline with a variable displayed as the area of the bubbles.

The bubble timeline is a compound data visualization of a scaled timeline and a proportional area chart. Different data categories are plotted on individual lines, to allow cross-category exploration or to see each category in isolation.



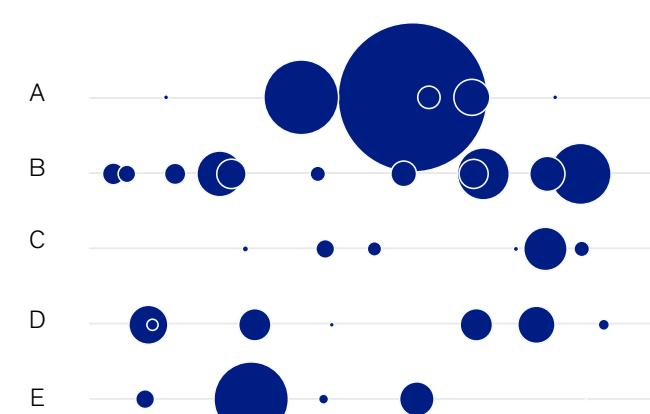
Bubble timeline best practice

Make meaningful color choices



✗ Try to avoid

As the categories are already split across different lines, it's usually not necessary to color the different categories of data



✓ Recommended

Keeping the bubbles monochrome allows the viewer to compare sizes and identify outliers

Keep overlapping bubbles visible



✗ Try to avoid

When dealing with large bubbles, avoid covering smaller bubbles



✓ Recommended

Place the largest bubbles in the background. Adding a white stroke and/or applying transparency to the bubbles can help their contours to stand out

Correlation

Scatterplot and bubble scatterplot

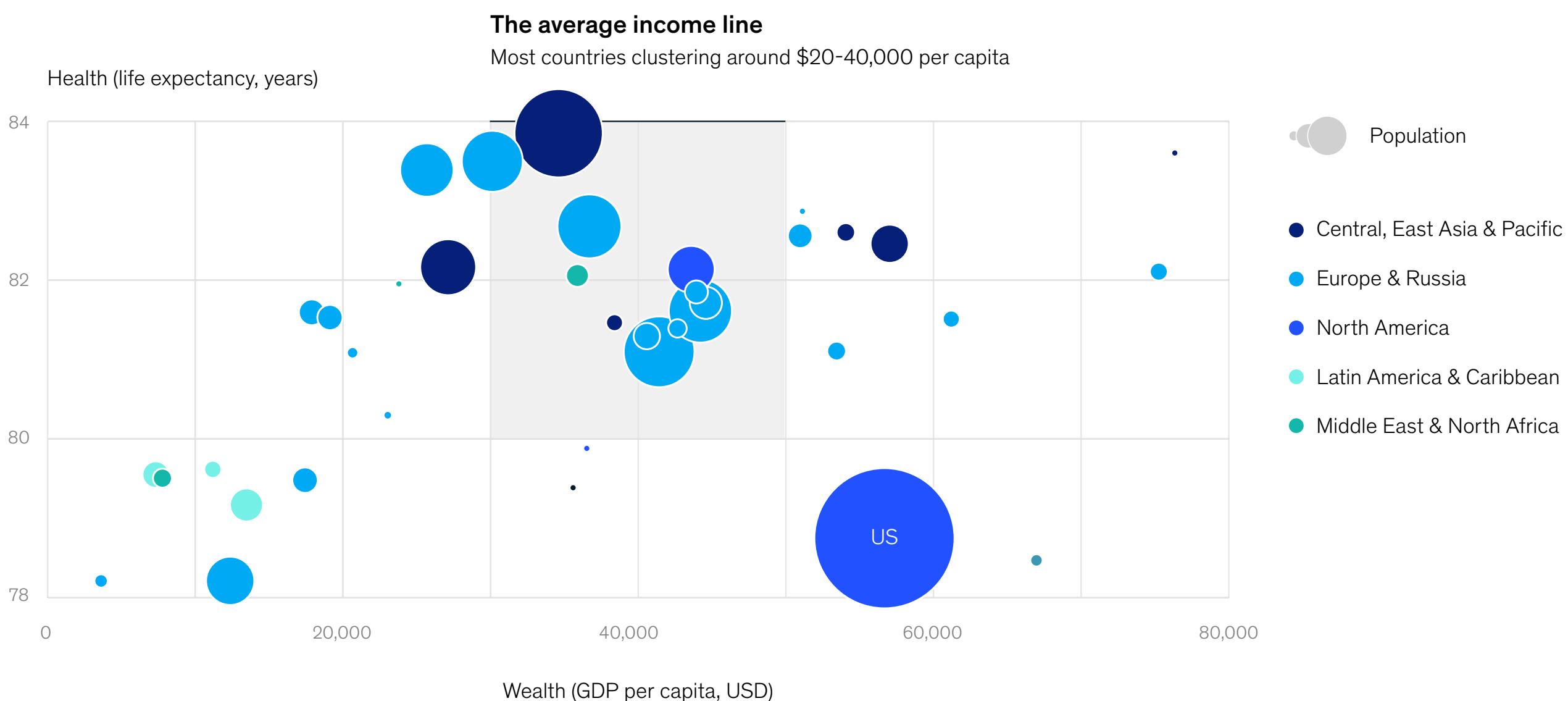
A scatterplot displays two dimensions of data based on coordinates along two axes.

A bubble scatterplot displays three dimensions of data. Each entity is plotted as a bubble that expresses two of the values through the bubbles' position on the X and Y axes, and the third through its size.

Bubble charts can be considered a variation of the dot scatterplot, with the data points replaced with bubbles. The points or bubbles in scatterplot displays can also be colored to further reveal patterns within the data set.

Correlation charts display data points plotted on two axes and show correlations within the data set.

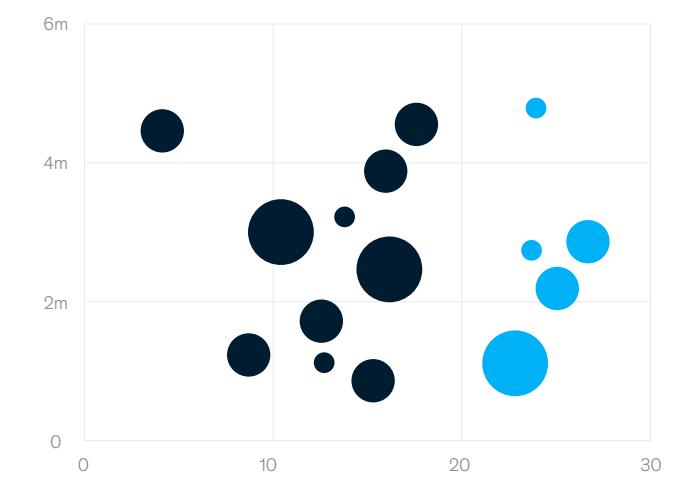
While most European countries are well off, US shows a low life expectancy, despite better GDP per capita
Countries with life expectancy above 78 years old across the globe, 2017



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

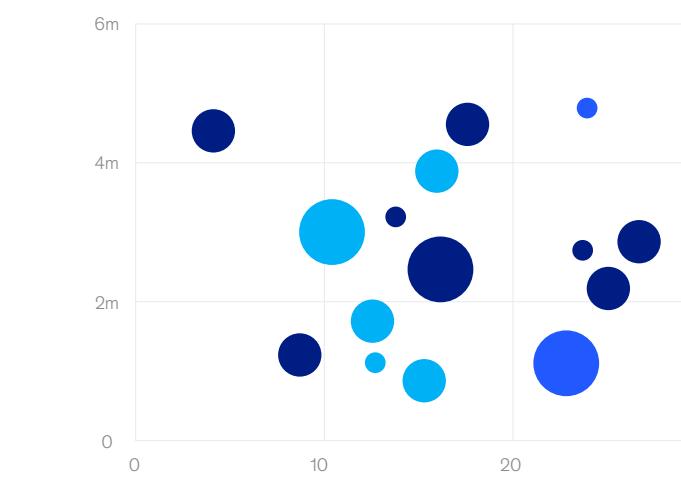
Bubble scatterplot best practice

Make meaningful color choices



Option 1

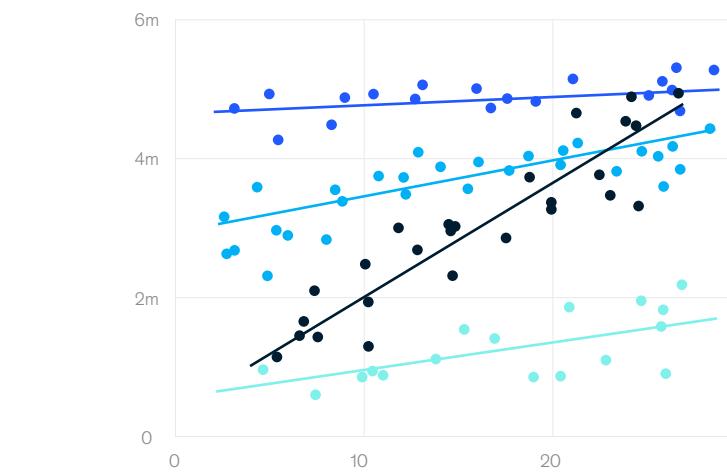
Sometimes it may work to use color to emphasize specific ranges



Option 2

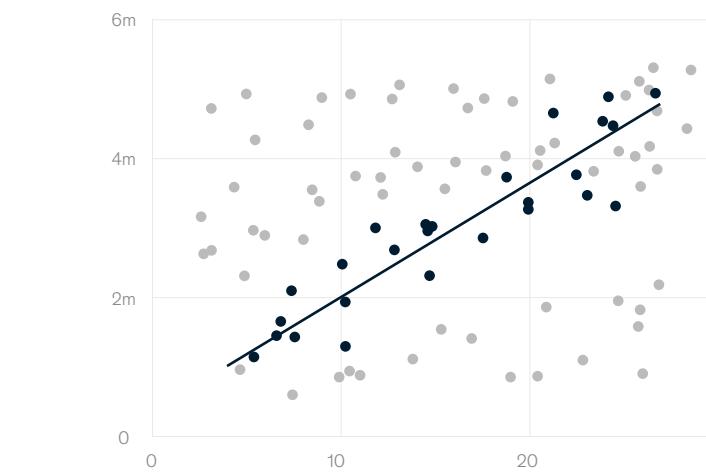
Alternatively, try to use color to encode additional data variables

Use trend lines judiciously



Try to avoid

Trend lines can help draw correlation between the variables to show trends, but too many lines make data difficult to interpret



Recommended

It's best to limit trend lines to the fewest needed to effectively tell the story of the chart

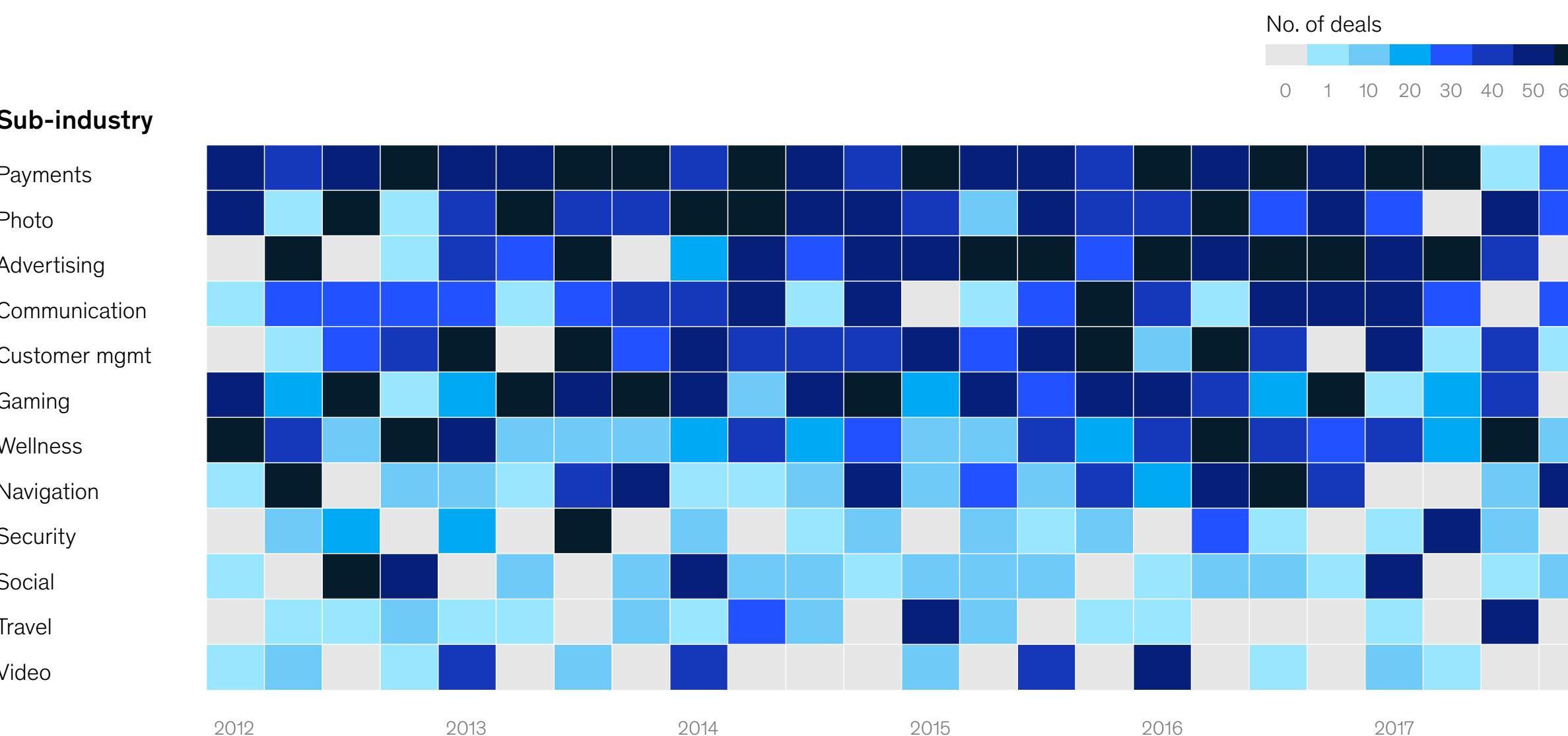
Sequential heatmap

A heatmap is a data visualization type where the individual values contained in a matrix are shown through variations in coloring. Heat maps are useful for visualizing variance across multiple variables to display patterns in correlations.

A sequential heatmap uses the linear scale color palette. When the reference value is at one end of the data range, use a linear color scheme. It uses a neutral color for the zero/null value. Generally, the darkest shade would represent the highest value (eg highest number of deals) and the lightest the lowest.

Sub-industries are ranked from highest number of deals to lowest number of deals over the time period

Photo and payments show consistently positive results across the years Mobile quarterly financing history by sub-industry, 2012-2017



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

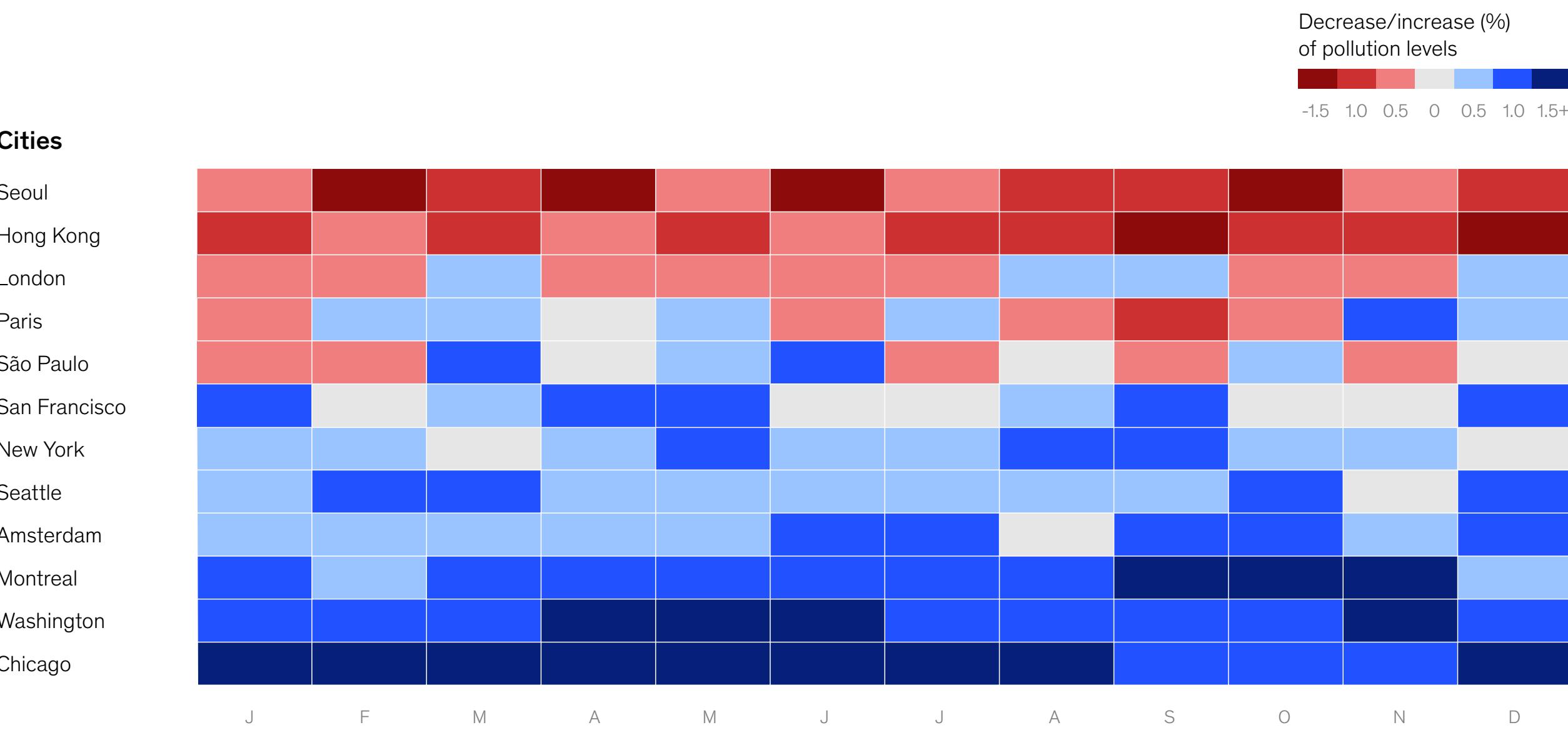
A sequential continuous color palette is suited here to encode the number of deals secured in different sub-industries

Diverging heatmap

A diverging heatmap uses a diverging color palette. When the reference value is in the middle of the data range (such as zero or an average value), use a diverging color scheme, which uses a neutral color for the reference value. Low values are represented by using one hue, and high values by using a different hue.

Cities are ranked from highest increment of pollution levels to highest decrement

North American cities are overall less polluted than Asian ones Urban air pollution changes, 2007-2017, %



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Decrease/increase (%)
of pollution levels
-1.5 1.0 0.5 0 0.5 1.0 1.5+

A diverging scale helps highlight the alarming increase of pollution levels in certain cities

Heatmaps best practice

Coloring steps of a heatmap

Number of steps



Number of steps



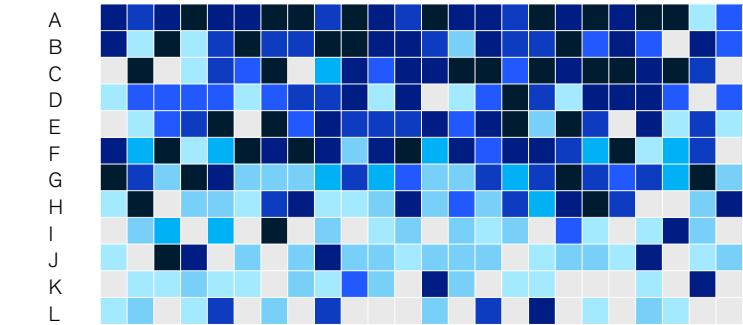
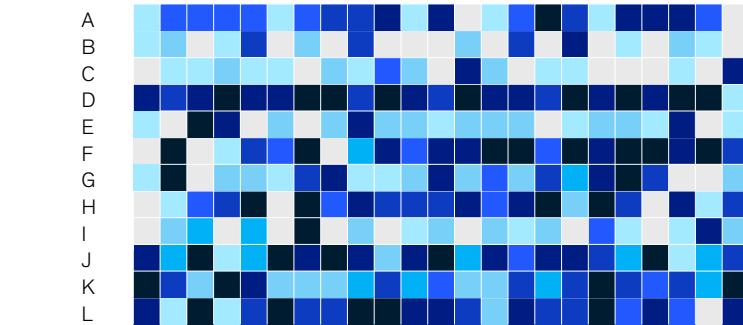
✗ Try to avoid

When creating color steps for a sequential heatmap, do not create arbitrary or unbalanced thresholds

✓ Recommended

Keep the order of the steps sequential and incrementally balanced

Sort the data to reveal visual patterns



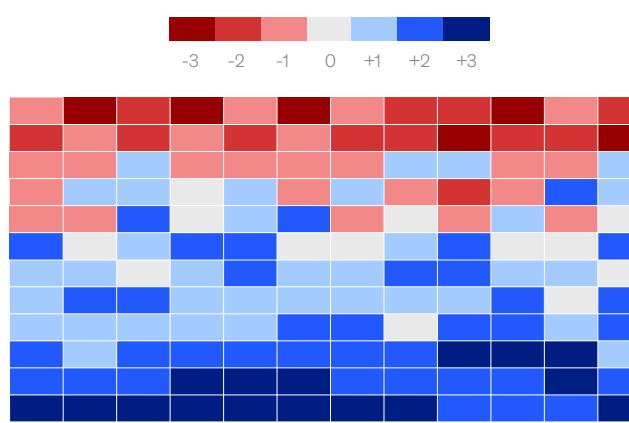
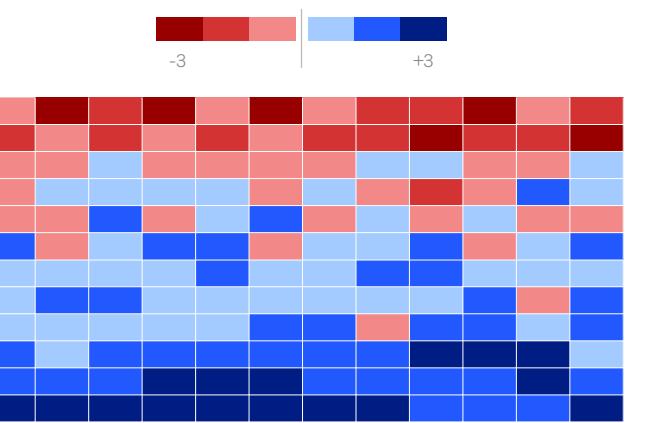
✗ Try to avoid

There are several ways to sort and order data for heatmaps. Sometimes alphabetical order works, but in other instances it may not help to reveal patterns and trends

✓ Recommended

Heatmaps are great devices for showing overall trends. By ranking data in order of density, the color steps of the heatmap allow patterns to emerge

Include zero on diverging heatmaps



✗ Try to avoid

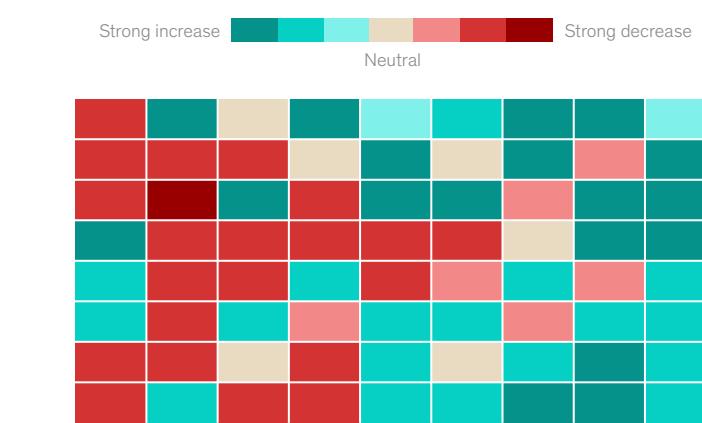
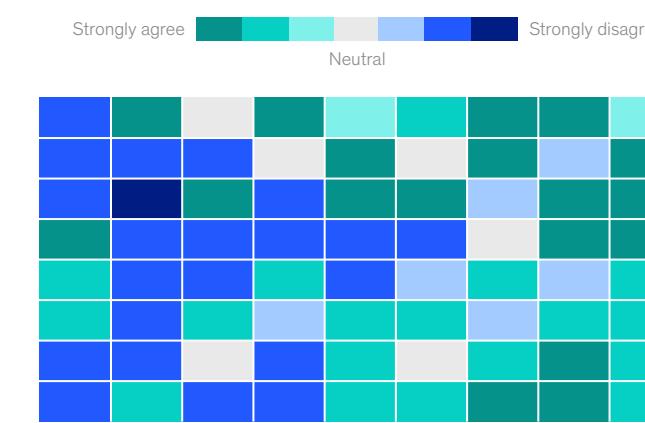
When making a diverging heatmap, do not remove the zero step from the legend

✓ Recommended

Sequential heatmaps work best when including a zero step that acts as neutral point

For more guidance on linear and divergent scales, including the use of tools such as [Chroma.js](#) to generate a greater number of color steps, please visit the data visualization page on the Our Identity site.

Use a properly diverging color combo



✗ Try to avoid

A combination of two cooler color scales, such as Electric Blue with Marine Green, provides inadequate contrast

✓ Recommended

Choose one of the recommended palettes meaningfully—for example, using the Marine Green and Crimson Red option to highlight increases and decreases in the data

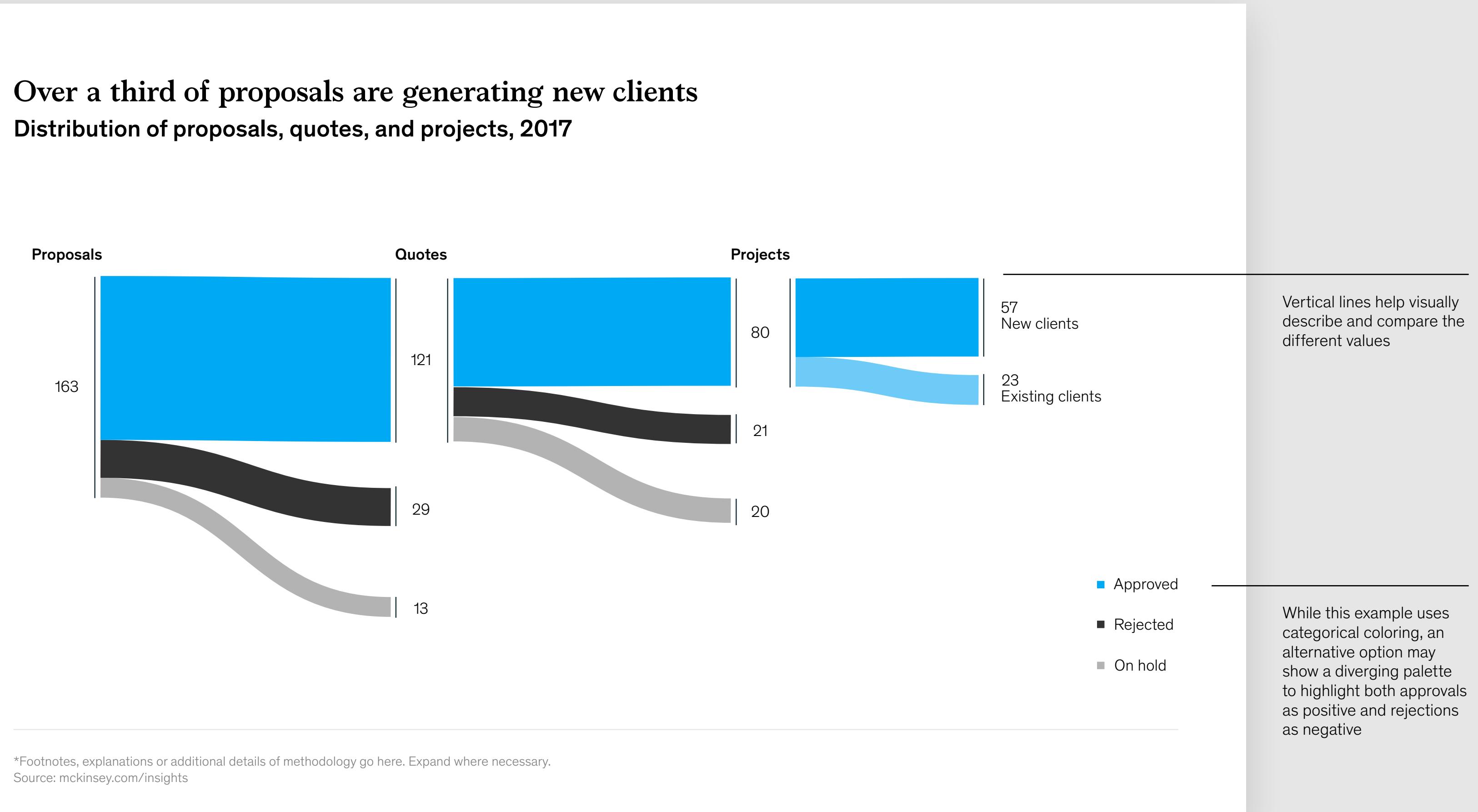
Relationships

Sankey diagram

Sankey diagrams are a specific type of flow diagram. The width of the lines are shown proportionally to the flow quantity. They are typically used to visualize energy, material or cost transfers between processes.

Sankey diagrams place visual emphasis on the major transfers or flows within a system. They are helpful in locating dominant contributions to an overall flow.

Paths diverge into smaller flows as quantities are transferred within the process. Cyan 500 and Cyan 300 are used to highlight approvals



Relationships charts allows users to see logical relationships between any entities or show volumes change from one set to another.

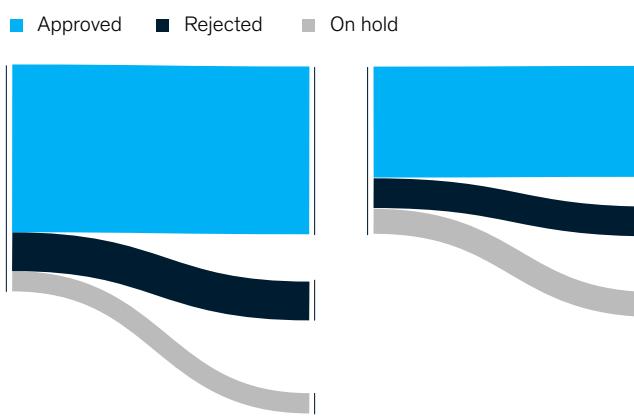
Sankey diagram best practice

Use color to show transfers



✗ Try to avoid

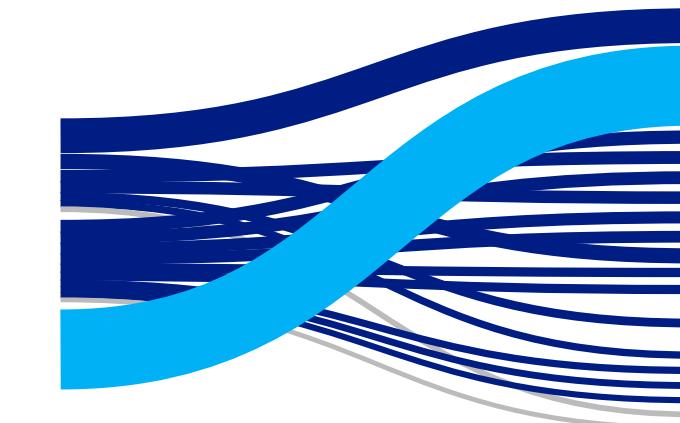
When coloring a multi-step Sankey diagram, it's best to avoid using switching colors to highlight obvious separations in the diagram



✓ Recommended

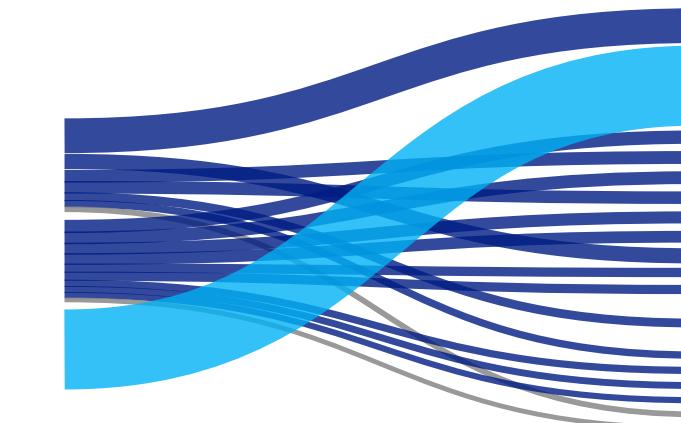
Color is most effectively used when different types of transfer are color-coded across the steps

Use transparency when flows overlap



✗ Try to avoid

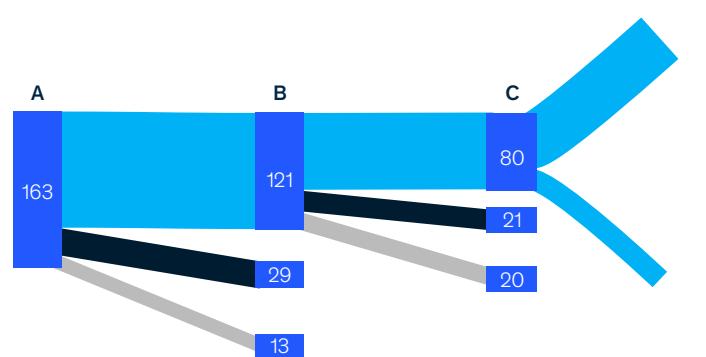
Try to avoid using flat, opaque colors, as steps of the flow may obscure others



✓ Recommended

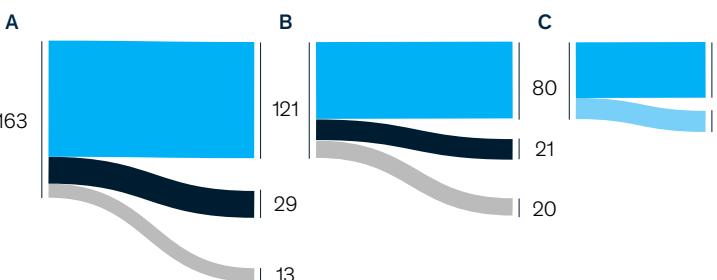
Using transparency allows us to see through the dense part of the flow diagram

Styling Sankey diagrams



✗ Try to avoid

Avoid boxing the values between the steps of the Sankey diagram. Avoid changing the organic shape of the flows into straight lines



✓ Recommended

Maintain an organic, curved flow in the paths. Use simple devices like thin vertical lines to describe volumes of the different values

Venn diagram

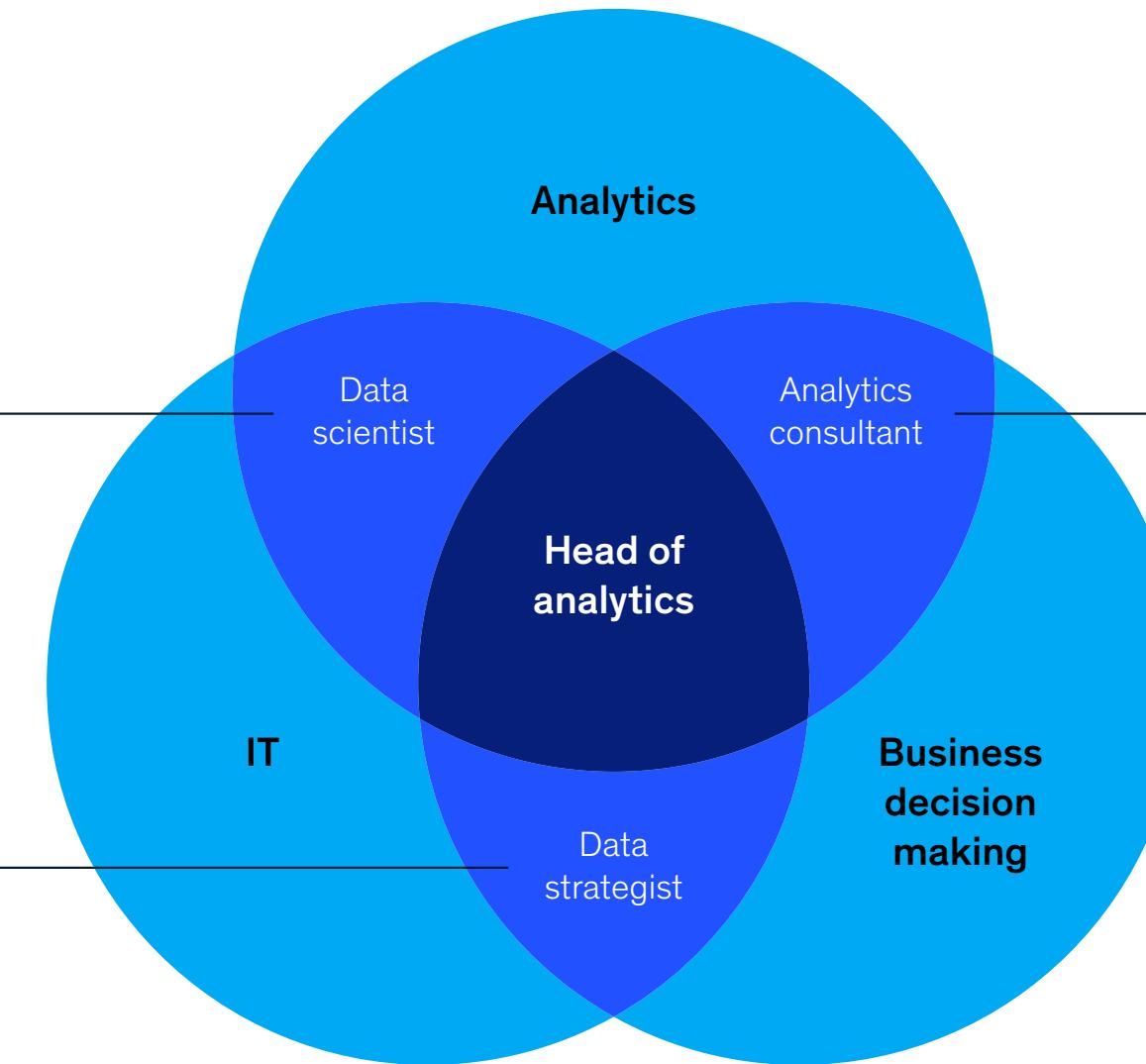
Venn diagrams are a way to visualize the structure of set relationships in a metaphorical way rather than a mathematically measured way.

Venn diagrams create overlapped shapes that are darker in color as the overlaps increase

Descriptions of the overlapped sections of the Venn diagram have been placed outside the diagram to avoid overwhelming amounts of content

The foundations of a balanced analytics team relies on data understanding

Analytics team core structure



Ensure best-in-class models and algorithms; support internal customers.

Ensure future data requirements and delivery roadmap is robust and complete.

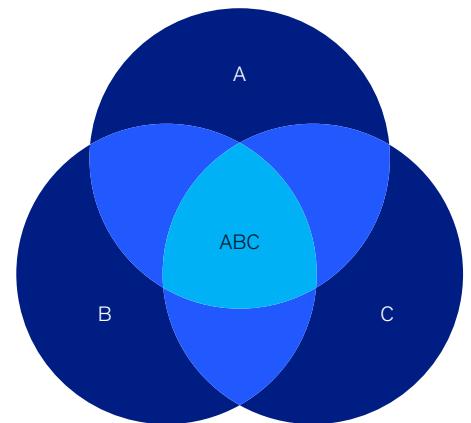
*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Solid understanding of statistics and analytics which can leverage into business decisions.

The lighter circles use the Cyan 500. As circles overlap, the overlapping areas of the Venn diagram create combinations of colors that use the Electric Blue 500 and Electric Blue 900 for the darker shades

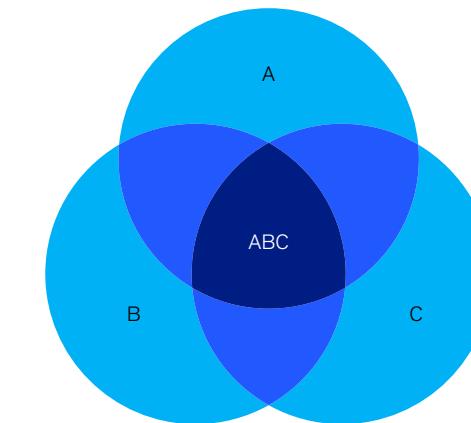
Venn diagram best practice

Color guidance for overlapping circles



✗ Avoid

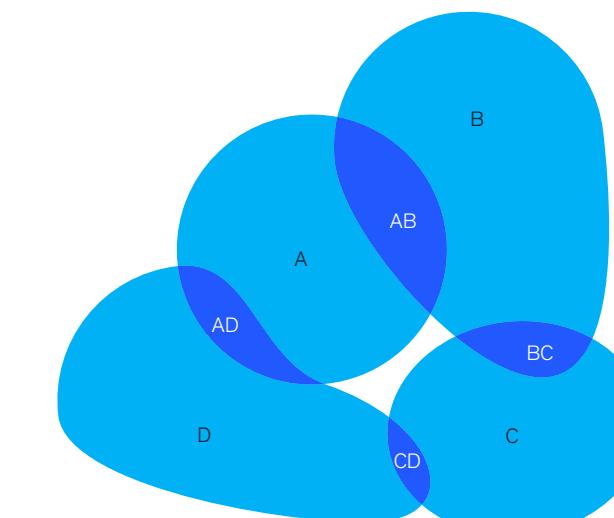
When overlapping circles, do not misrepresent overlaps by lightening the overlapped areas



✓ Recommended

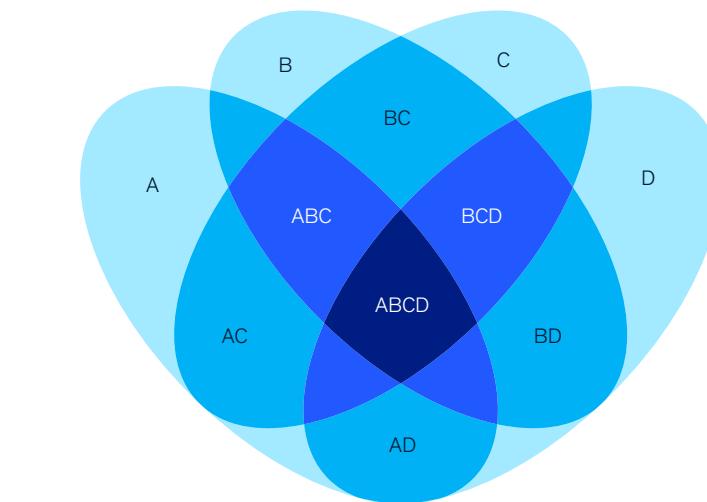
It's best to show the overlapped areas in darker shades of the color

Use alternate shapes in a structured way



✗ Try to avoid

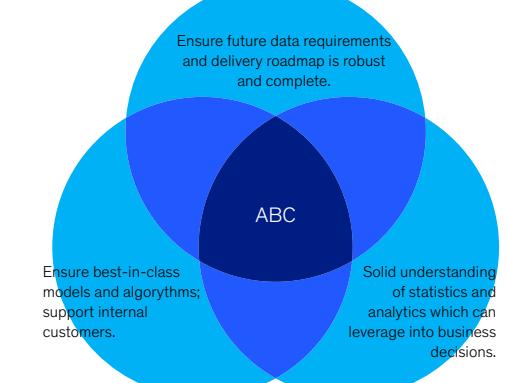
It's best to use regular shapes. Do not morph the Venn areas into irregular objects that create unbalanced layouts



✓ Recommended

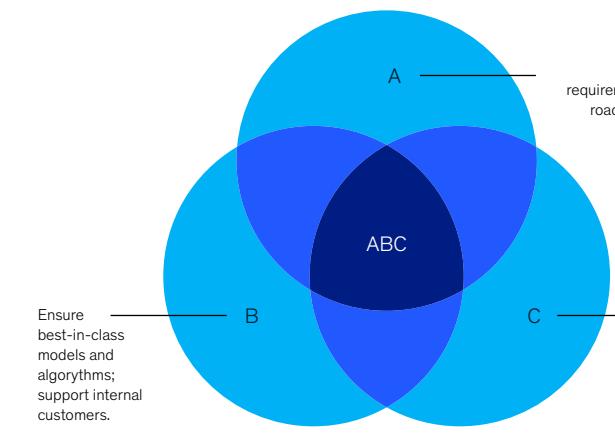
Try to use regular and organic layouts that, even when complex, maintain a simple and clear hierarchy structure

Keep text legible



✗ Try to avoid

Do not place long blocks of text atop the Venn diagram. It may obscure the shaded areas, making it hard to read



✓ Recommended

It's best to use subtle connecting lines and move descriptor text outside the diagram

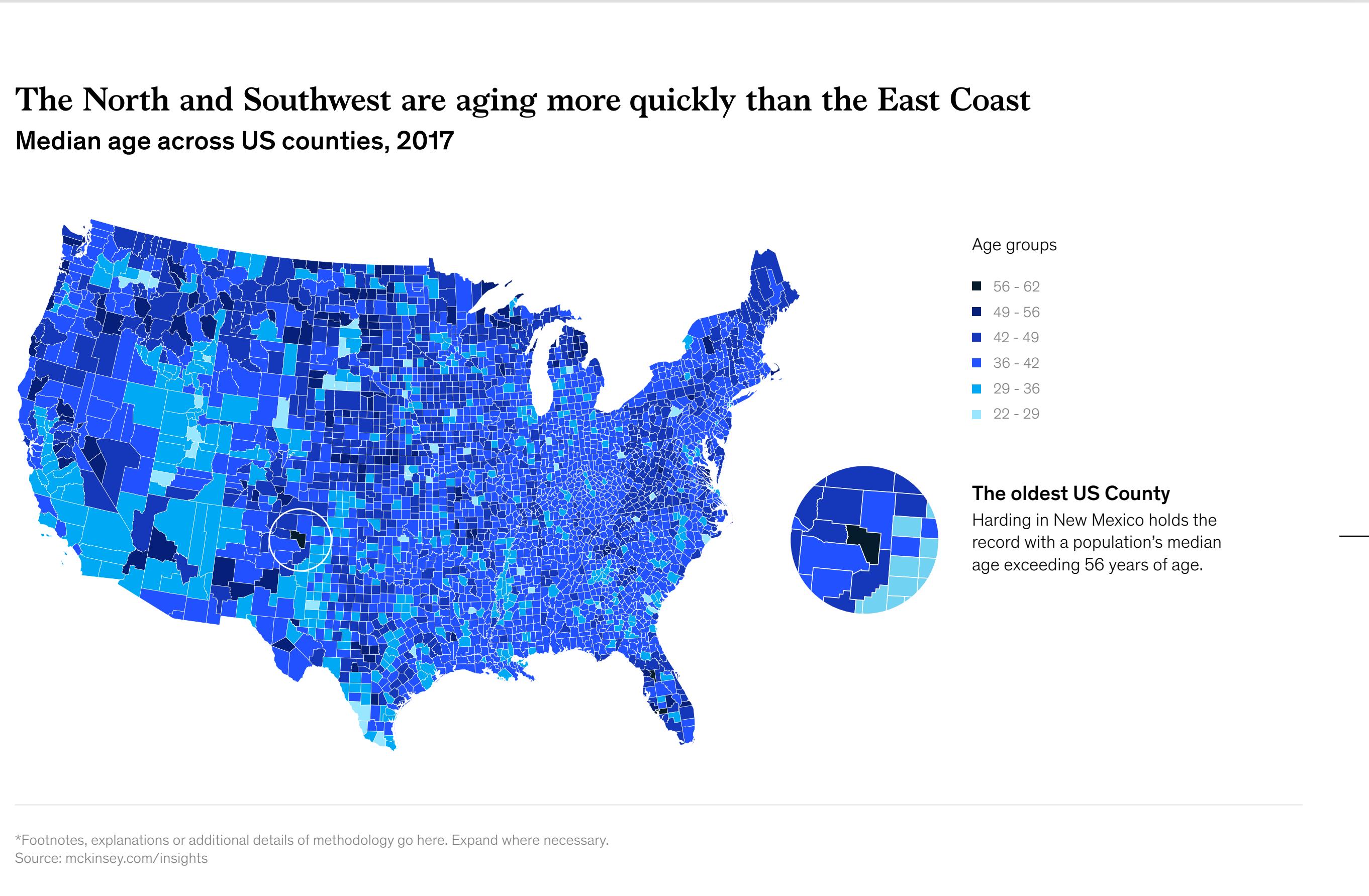
Geospatial

Choropleth

A choropleth is a color-themed map in which areas are shaded in proportion to the measurement of the variable in an underlying geographic dataset, such as population density or per-capita income.

While such map visualizations can be insightful, an important consideration is that geographical areas can sometimes skew a reader's representation of data. On a world map, for instance, Russia has a relatively low population compared its spatial land area. Although Russia's size may be irrelevant to the story being told, this could perceptually conflict with the intended narrative as it occupies a huge proportion of the graphic.

When coloring the regions in a choropleth map, a sequential continuous color palette can efficiently encode the difference in value across regions



Geospatial charts are used when it's relevant to highlight precise locations or geographical patterns.

Bubble map

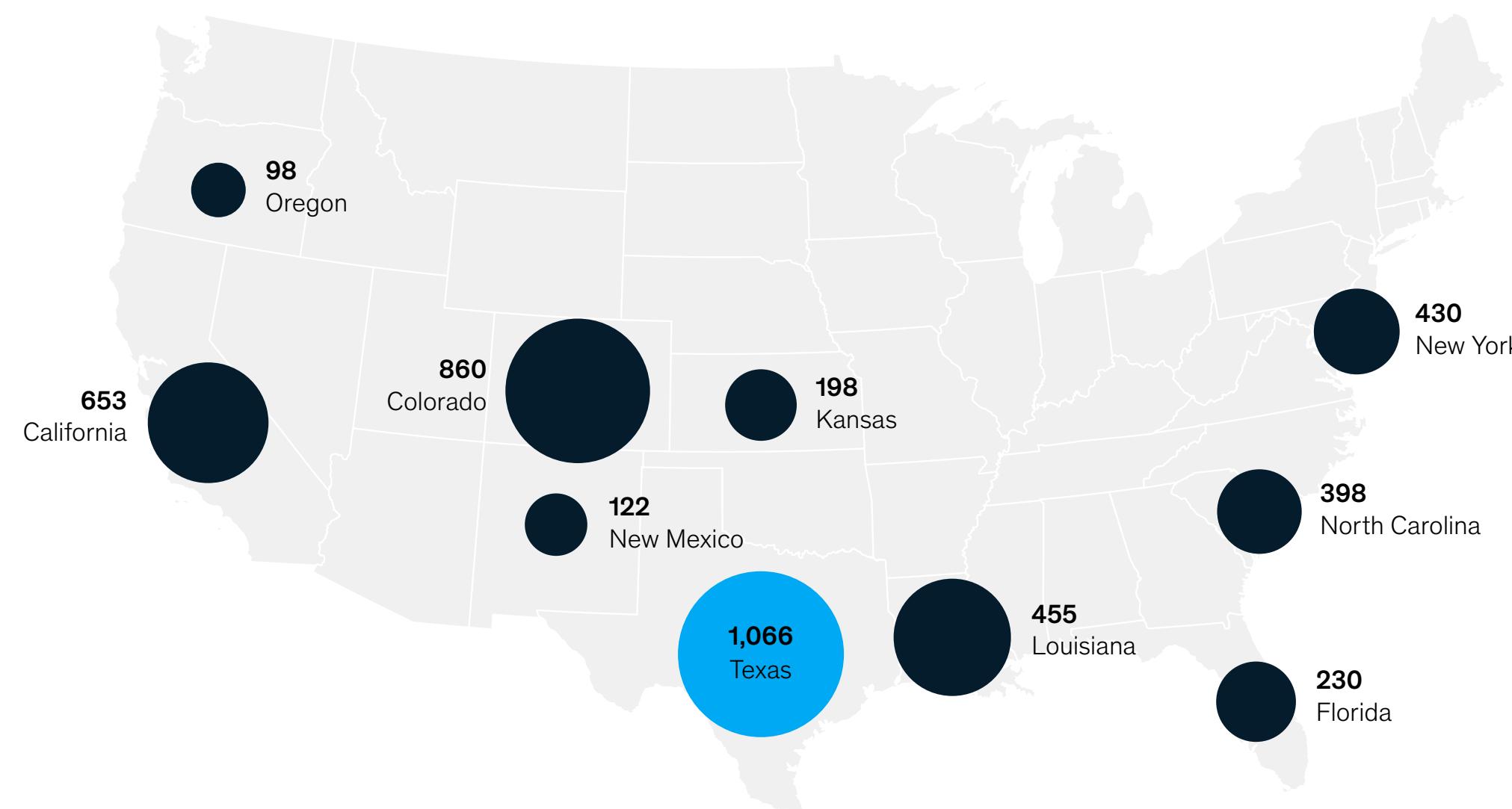
A bubble map is a combination of a bubble chart data visualization and a map. It is used to visualize location and proportion in a simple way. This approach helps address the matter outsize geographical areas, such as the Russia example discussed on the previous page. Here the map helpfully communicates location, while the data at hand is encoded in the bubbles.

Direct labeling allows for easier exploration of the map

The bubbles on the map encode the amount of tech hubs in that state (scaled by area)

Most tech startups are moving to Texas

Top ten US states ranked by tech hubs, 2017



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

Dot density map

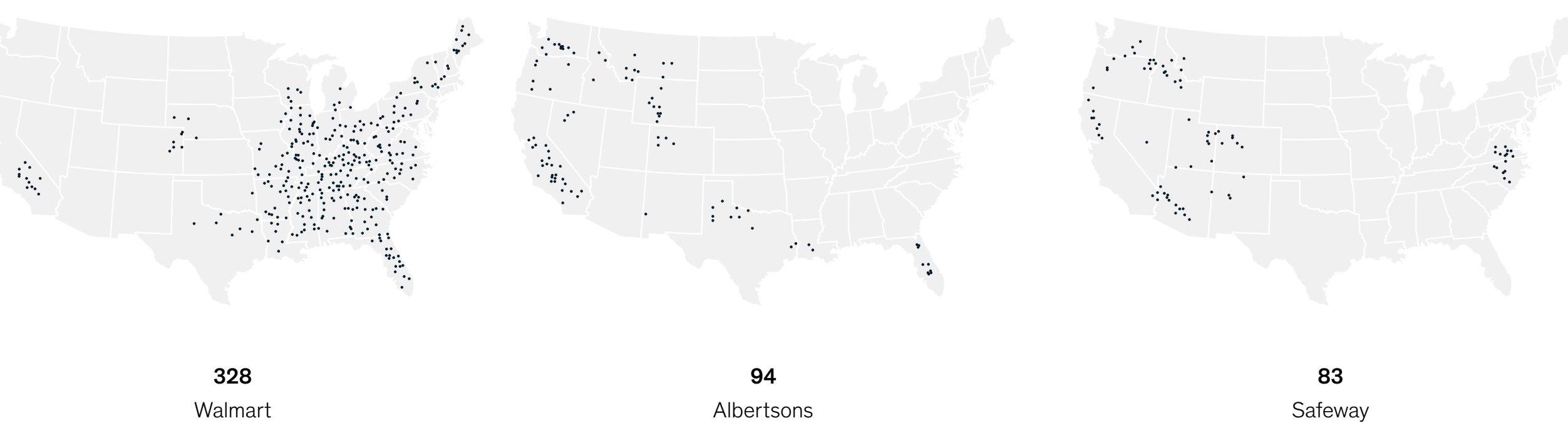
A dot density map uses dots (or another symbol) to show the location of a set of data points.

In a dot density map, areas with many dots indicate high concentrations of values for the chosen field, and fewer dots to indicate lower concentrations. Each dot can either represent one single recording of a phenomenon (one-to-one) or represent a given quantity (one-to-many).

Data points are plotted on the map, and provide a visual overview of their geographic distribution

Walmart is the most popular grocery brand in the US, but is almost absent in the West

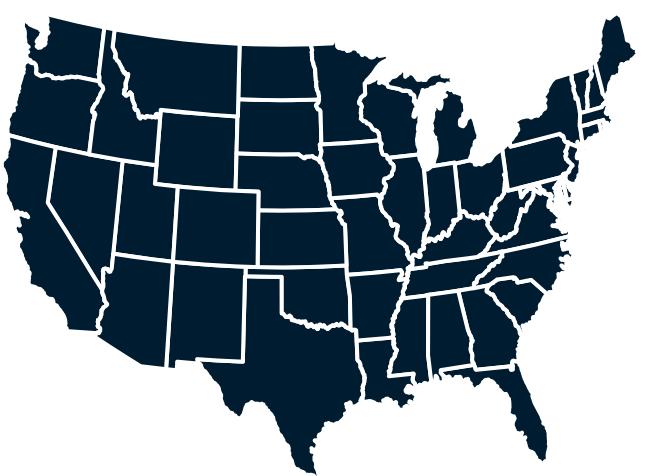
Distribution of the most popular grocery brands across the US, 2017



*Footnotes, explanations or additional details of methodology go here. Expand where necessary.
Source: mckinsey.com/insights

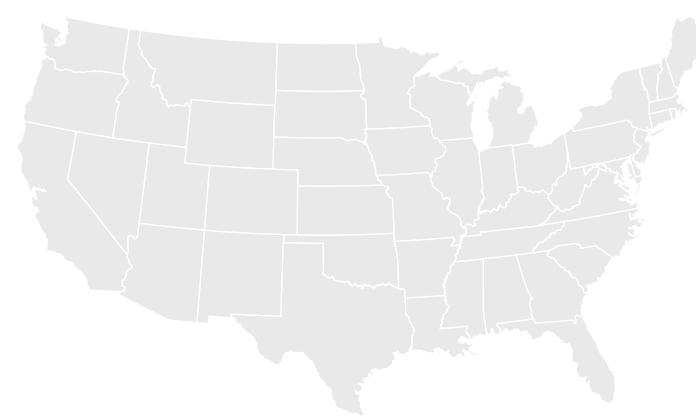
Maps best practice

Maintain neat shapes and borders



✗ Try to avoid

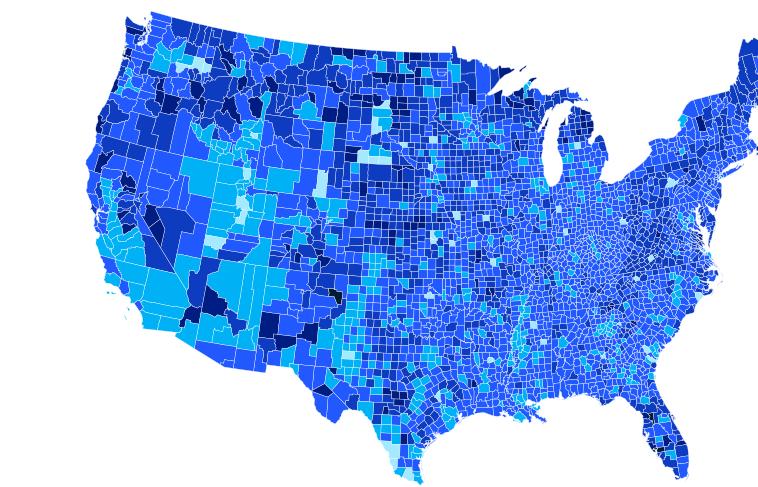
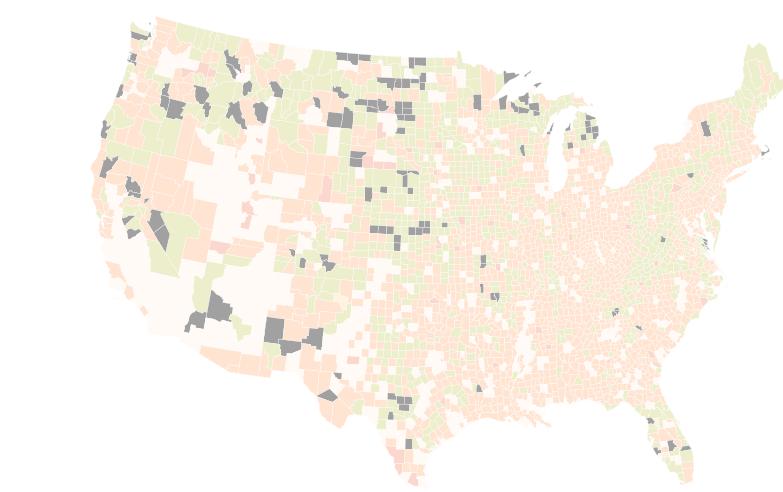
Avoid using maps that have thick boundary lines, stand out excessively, or maps that distort proportions



✓ Recommended

Using light gray and white outlines for boundary lines allows the data to stand out on screen. For print, consider having darker outlines as white outlines may not print clearly

Make meaningful color choices



✗ Try to avoid

Avoid using colors that are not part of the McKinsey palette. Also be aware some colors stand out more than others, giving unnecessary weight to that data

✓ Recommended

It is best to use a linear or diverging palette to show intensity

Explore display options



✓ Option one

There may be more than one way to display data on a map effectively. Plotting data on a single map may not always be the best answer



✓ Option two

Sometimes splitting the map into two allows differences in the distribution of data to be shown

Infographics and icons

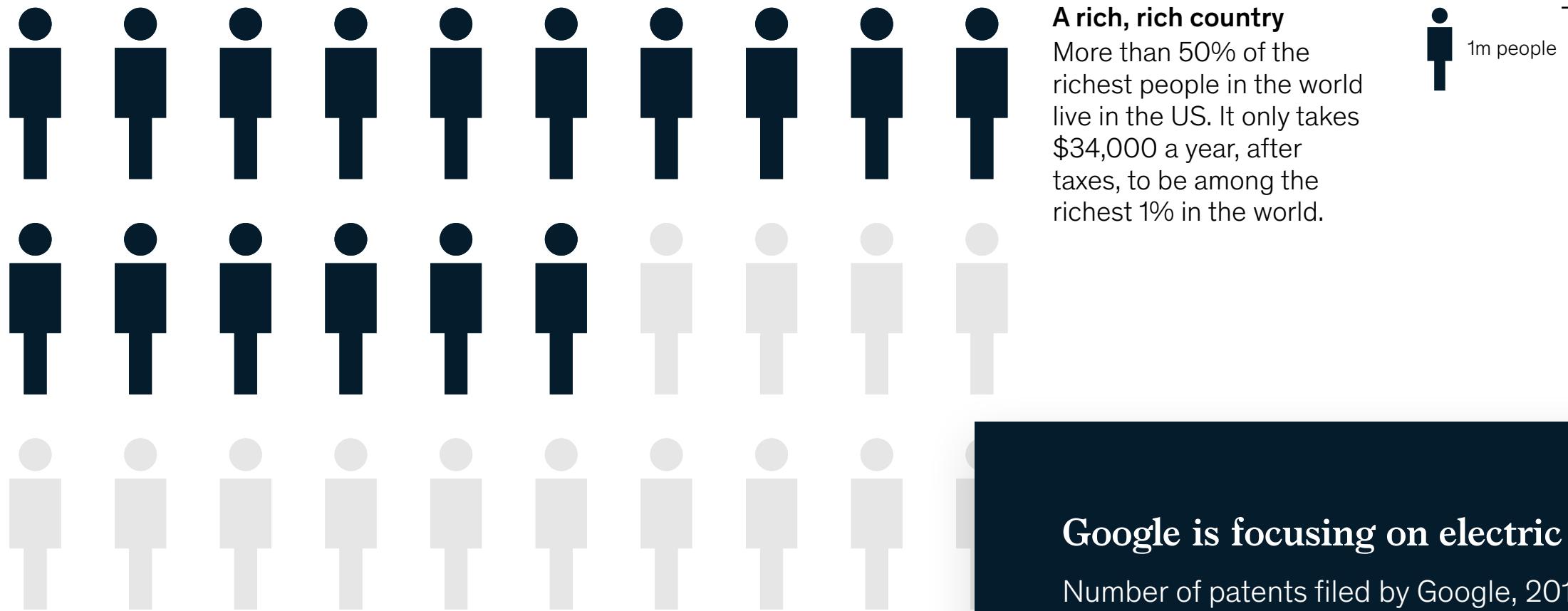
Infographic icons

In the world of data visualization, infographics are a popular and widely used tool. Our infographics can be effectively employed when addressing wider audiences via social media or other public channels.

McKinsey infographics may feature icons, used in smart and elegant ways, to highlight messages and appeal to a broader readership.

The United States holds a disproportionate amount of the world's richest people

The world's richest 1%, 2017

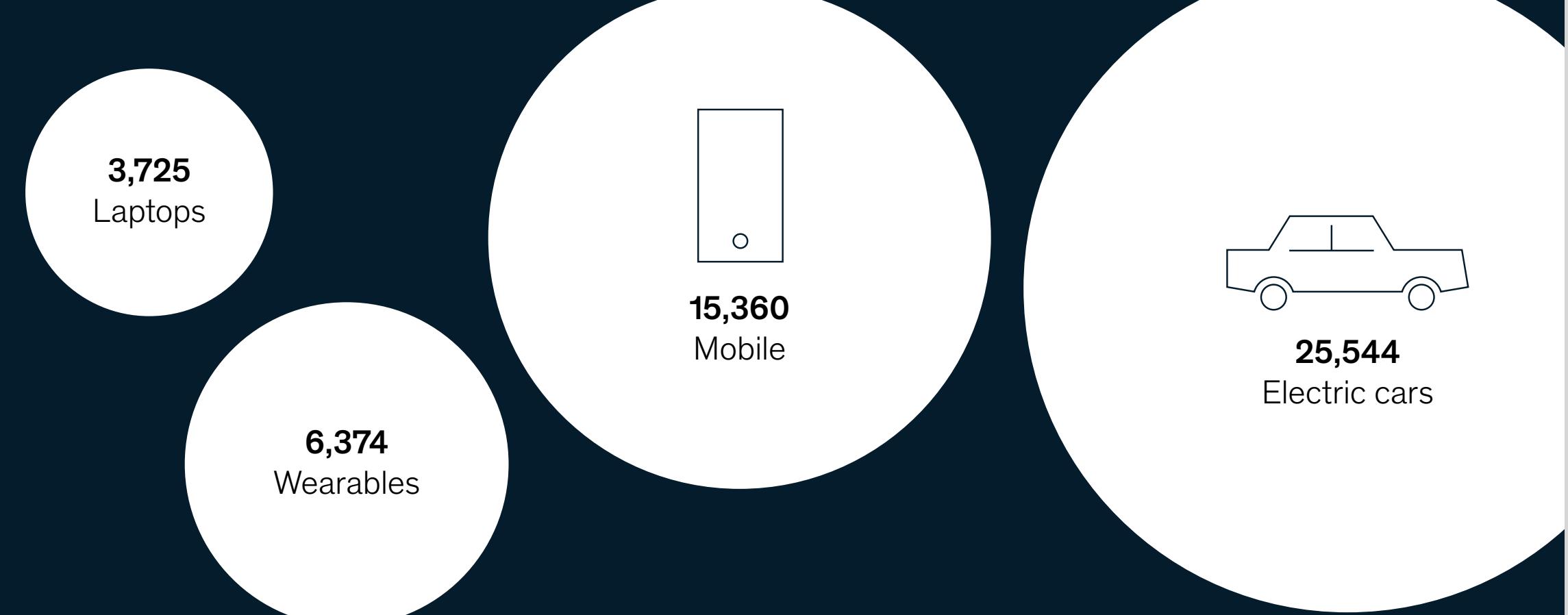


Source: mckinsey.com/insights

Icons are used here to highlight some relevant categories that match the insight. This is particularly useful when used on social media assets, where content needs to stand out visually

Google is focusing on electric cars

Number of patents filed by Google, 2017



Source: mckinsey.com/insights

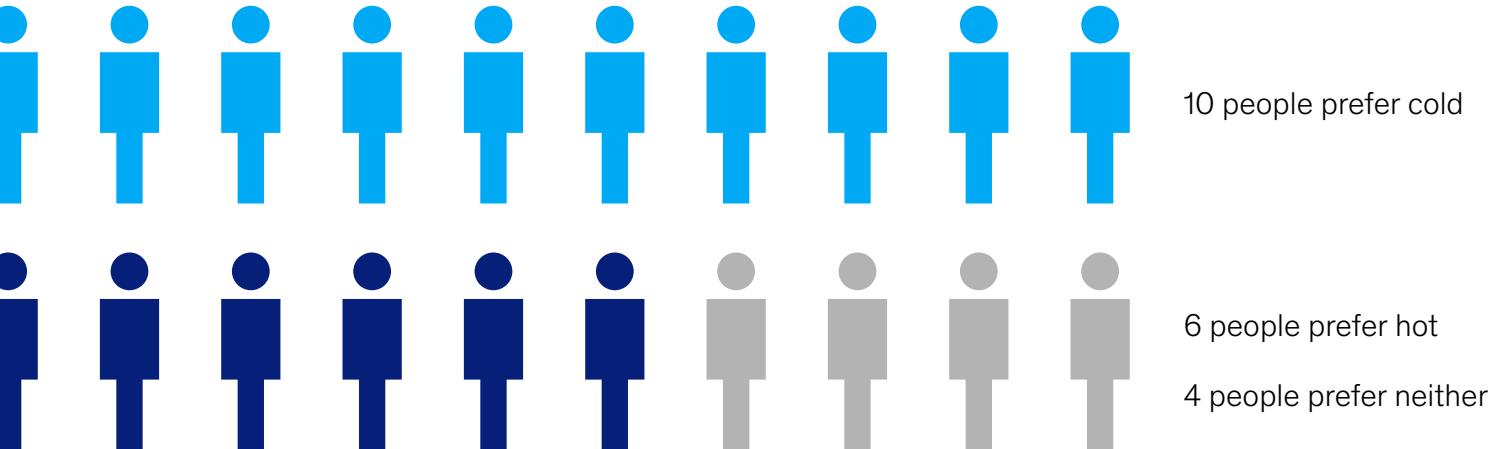
Using icons

When creating charts or data visualizations, icons can help represent information for quick understanding.

While iconography can provide a shortcut into the story, the use of icons should be regarded with caution. As a general rule, if it's possible to communicate something clearly without using an icon, it may not be necessary to use one.

Convey quantities with icons

Temperature preference on 20 test subjects

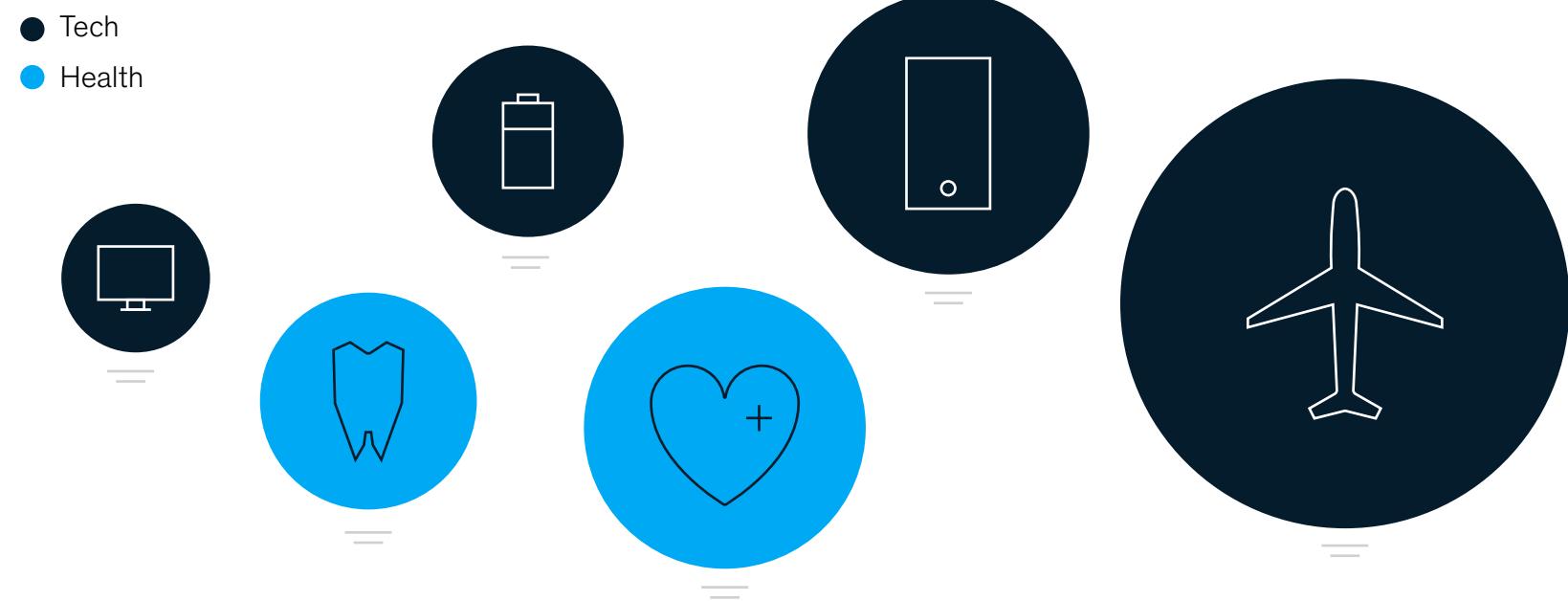


Recommended

Literal representations used as tallies to convey quantities can be used effectively as functional icons that perform a simple task (help compare) within an infographic

Categorize attributes

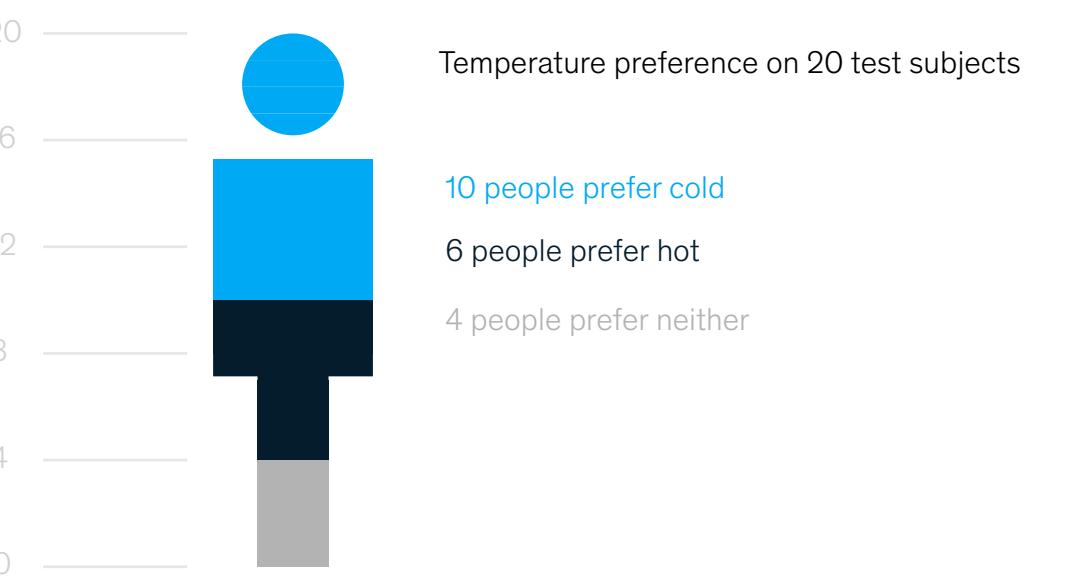
● Tech
● Health



Recommended

Icons can be a powerful instrument to provide a quick overview of complex data representation. They can be used to highlight categories or in combination with color-coding methods to give meaning to different attributes

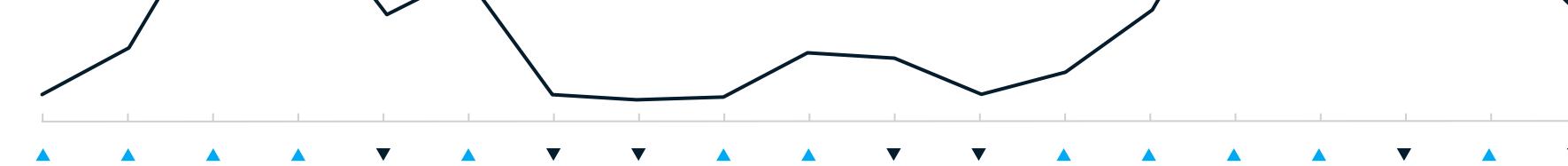
Improper scaling of icons



Avoid

Avoid improper handling of icon sizing. A common mistake is to scale both dimensions of an icon, which leads to an incorrect visual comparison and makes it hard to read the relationship with the data

Show variations with symbols



Recommended

Simple icons can be used as symbols to convey increments or decrements within the visualization

For further guidance
and additional information
please visit:
ouridentity.mckinsey.com