



Justice  
Innovation  
Lab

Justice Innovation Lab

# Data Visualization Guide

April 2024

# Contents

1	Color selection	3
1.1	Brand colors	4
1.2	Data colors	5
1.3	Categorical data colors	6
1.4	Ordinal data colors	7
1.5	Continuous data colors	8
2	Chart selection	9
2.1	Bar charts	10
2.2	Scatter charts	12
2.3	Line charts	13
2.4	Part-to-whole charts	14
2.5	Distribution charts	15
2.6	Heat maps	16
2.7	Combination charts	17
2.8	Small multiples	18
3	Chart details	19
4	Additional considerations	20

# 1

This section of the Justice Innovation Lab (JIL) Data Visualization Guide provides direction on the selection and use of color in charts.

The primary JIL data colors are based on the JIL brand colors, with adjustments to saturation and lightness to make them more suitable for data visualization.

The data colors have been selected with accessibility in mind. Each of the six data colors has sufficient contrast with a white background and with black or white text overlay (see Section 1.2). Additionally, each color combination has been tested to ensure that the colors are distinguishable for those with various forms of color vision deficiency.

The JIL data visualization colors have also been chosen with the goal of avoiding colors that reinforce racial, gender, or political stereotypes. If there is concern that unintentional meaning may be associated with the available data colors, avoid using color to differentiate demographic groups through strategies such as creating “small multiples” charts (Section 2.8) instead of showing all groups on the same chart.

## Color selection

### Brand colors



### Data colors



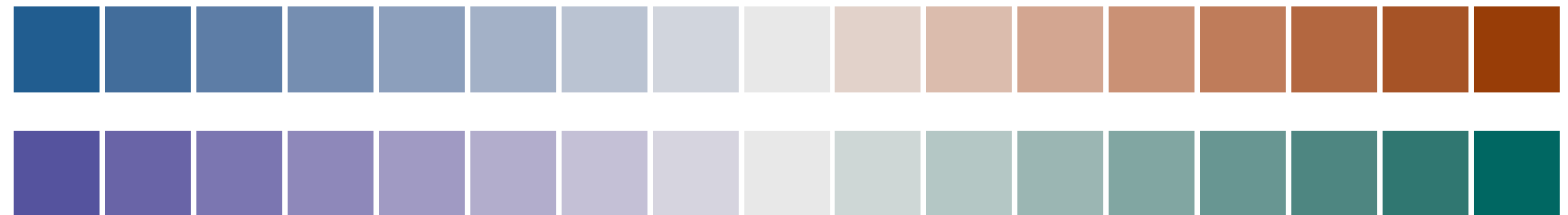
### Categorical data colors



### Ordinal data colors



### Continuous data colors



# 1.1

## Brand colors

JIL's brand colors consist of dark blue, light blue, orange, and light gray.

These colors are used throughout JIL's website, social media, logos, presentations, and other communication materials. The main colors used for JIL data visualizations (Section 1.2) are based on these brand colors, but have been adjusted in terms of saturation and lightness to be more appropriate for use with charts.

**Dark blue**

**HEX: #003057**

**Light blue**

**HEX: #64BAFF**

**Orange**

**HEX: #FF9533**

**Light gray**

**HEX: #DCE6EE**

## 1.2

## Data colors

The primary colors for JIL data visualizations are dark blue, light blue, orange, and gray, as shown in the first row. The second row of data colors (teal and purple) may be used when additional colors are necessary.

Each data color shown meets the Web Accessibility In Mind (WebAIM) contrast requirements for graphical objects when used on a white (#ffffff) background.

Text used on graphical objects should be black (#000000) or white (#ffffff). For each data color shown, the text color used for the color name is preferred. The text color used for the hex color code is also acceptable, but has lower contrast with the data color.

Each combination of data and text colors shown meets, at a minimum, WebAIM AA requirements for large text (minimum of 18 pt regular or 14 pt bold). The preferred combination for each also meets WebAIM AA requirements for normal text and AAA requirements for large text.

When selecting colors for data visualization:

- Use as few data colors as possible.
- Use only color combinations provided in this guide.
- Use data color meanings consistently within a given document, to the greatest extent possible.

**Dark blue**

HEX: #2b5e88

**Light blue**

HEX: #4498cf

**Orange**

HEX: #db771a

**Gray**

HEX: #89929a

**Teal**

HEX: #579e9a

**Purple**

HEX: #7075bf

# 1.3

## Categorical data colors

The color palettes shown here are for use with categorical data. Examples of categorical data include groupings such as race, sex, age, charge type, county, etc.

Each color combination presented has been tested for accessibility for those with various types of color vision deficiency.

The main JIL color is blue; therefore, either dark blue or light blue is preferred if a chart uses a single color.

Avoid using any more than three categorical data colors in a single chart. However, gray can be added to any of the color combinations in this section if another color is needed. If gray is used with other colors, be careful of how it may be perceived since it could be taken as an indication that the information shown in gray is less of a focus or less important.

### 2 groups



#2b5e88 #4498cf



#2b5e88 #db771a



#4498cf #db771a



#579e9a #7075bf



#2b5e88 #579e9a

### 3 groups



#2b5e88 #4498cf #db771a



#2b5e88 #7075bf #579e9a



#4498cf #7075bf #579e9a

# 1.4

Use any of the single hue color schemes in this section to display ordered categorical (ordinal) data, where there is a clear and logical order to the levels of a categorical variable.

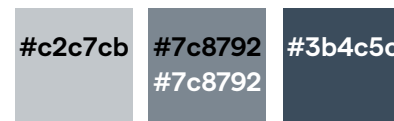
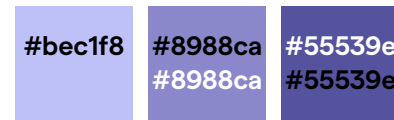
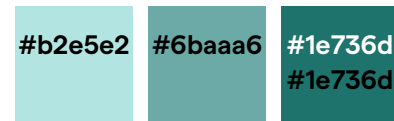
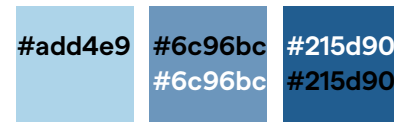
Examples of ordinal data include:

- Charge severity (1st, 2nd, 3rd offense)
- Level (low, medium, high)
- Opinion (strongly agree, agree, neutral, disagree, strongly disagree)
- Frequency (never, rarely, sometimes, often, always)
- Age groups (0-10, 11-20, 21-30, etc.)
- Monetary groups (\$30k-\$49k, \$50k-\$69k, \$70k-\$89k, etc.)
- Quantiles (1st quartile, 2nd quartile, 3rd quartile, 4th quartile)
- Ranking (1st, 2nd, 3rd, 4th, etc.)
- Education level (primary school, high school, undergraduate, graduate)

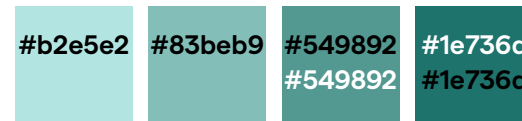
For each data color shown, the color of text (black, #000000 or white, #ffffff) used for the hex color code at the top of each box is preferred. For colors with a repeated hex color code, the second text color is also acceptable, but has lower contrast with the data color.

## Ordinal data colors

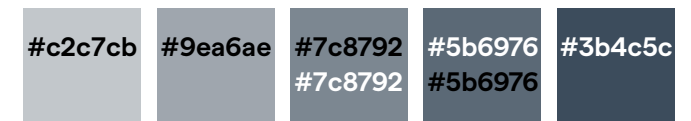
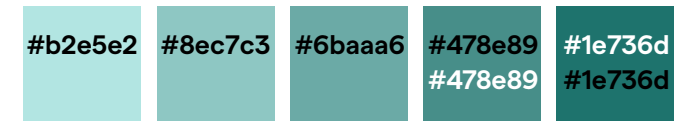
### 3 groups



### 4 groups



### 5 groups



# 1.5

## Continuous data colors

The color palettes shown in this section are for use with continuous quantitative data. Typical visualizations that utilize continuous color scales include geographic maps and heat maps.

For continuous, sequential data (such as low to high), any of these four single-hue color progressions can be used on their own.

For diverging data, where there is a meaningful, numeric midpoint, use either blue and orange together or purple and teal together, with a neutral midpoint (#e8e8e8). The blue and orange diverging palette is more effective than the purple and teal diverging palette for emphasizing opposite extremes.

### Blue

#215d90  
#215d90

#426d9b  
#426d9b

#5d7da6  
#5d7da6

#758eb1  
#758eb1

#8c9fbc

#a3b1c7

#bac3d2

#d1d5dd

#e8e8e8

### Orange

#983d07  
#983d07

#a65326  
#a65326

#b36740  
#b36740

#bf7c5a  
#bf7c5a

#ca9175

#d3a691

#dbbcad

#e2d2ca

#e8e8e8

### Purple

#55539e  
#55539e

#6964a7  
#6964a7

#7b76b1  
#7b76b1

#8e88ba  
#8e88ba

#a09ac3

#b2adcc

#c4c0d6

#d6d4df

#e8e8e8

### Teal

#006762  
#006762

#307771  
#307771

#4e8681  
#4e8681

#689692  
#689692

#81a6a2

#9bb6b3

#b4c7c5

#ced7d6

#e8e8e8



## 2

There are several factors to consider when deciding which type of chart to use for a data visualization. One of the most important is the type of data to be displayed. This guide refers to the following data types, as defined for the Altair python package for data visualization.

- Quantitative: A continuous real-valued quantity
- Ordinal: A discrete, ordered quantity
- Nominal: A discrete, unordered category
- Temporal: A time or date value

The term “categorical variable” will also be used in this guide to refer to any variable with a finite number of discrete groups.

Another important factor in chart selection is the amount of data to be displayed.

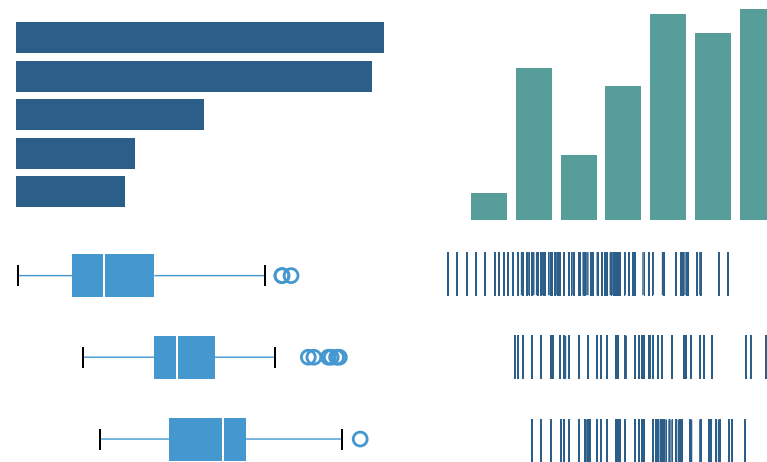
Some charts, such as bar charts, are more appropriate when there are relatively few data points, whereas scatter charts and distribution charts are more effective for large amounts of data.

This section provides chart selection and design guidance for several common chart types, but is not exhaustive. Updates and additions to this guide will be made as appropriate.

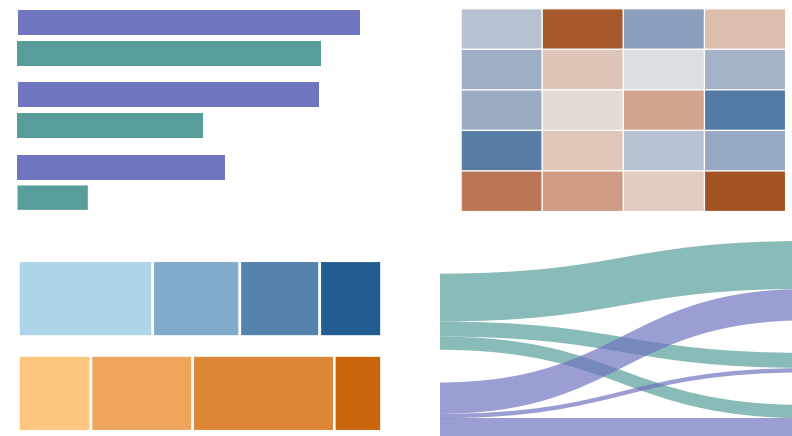
## Chart selection

### Chart examples by type and number of variables displayed

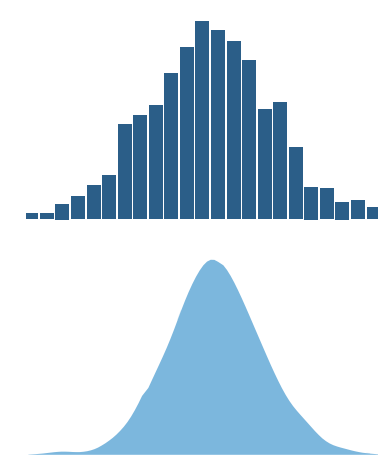
#### 1 categorical, 1 quantitative



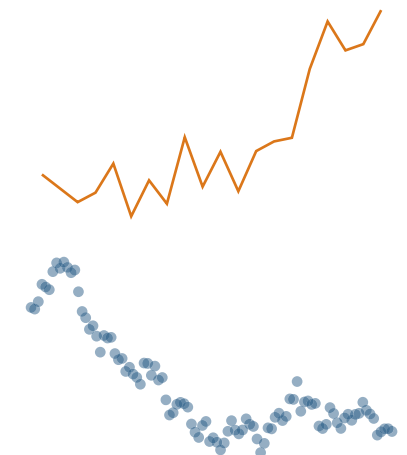
#### 2 categorical, 1 quantitative



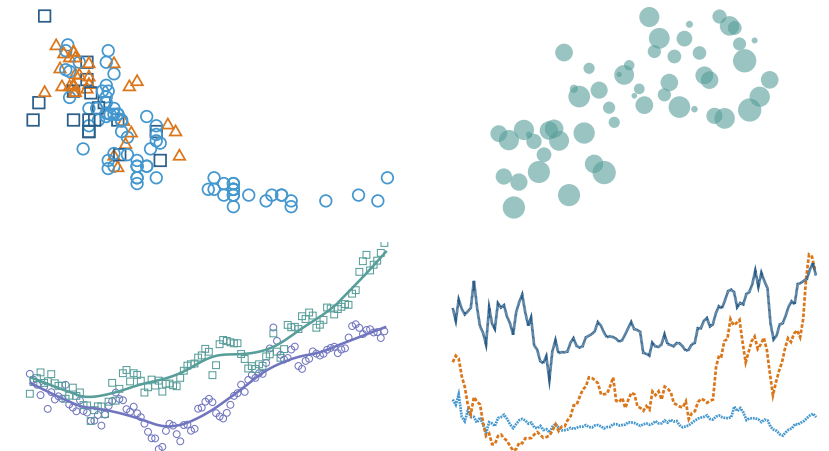
#### 1 quantitative



#### 2 quantitative



#### 2 quantitative, 1 categorical



# 2.1

# Bar charts

## When to use

Use bar charts to compare quantitative values among different levels of an ordinal, nominal, or temporal variable.

## Design guidance

- Always start the numeric axis of bar charts at zero so the relative bar lengths accurately represent quantitative differences in the data (figs. 1–4).
- Horizontal bars are preferable to vertical bars (columns) since they allow more space for axis and data text (figs. 1 & 2).
- Use vertical bars for time-series data, with the temporal variable on the x-axis (figs. 3 & 4).
- Arrange bars using numeric order for nominal variables (fig. 1) and sequential order for ordinal variables (fig. 2).
- Data value text should be left-aligned within horizontal bars (figs. 1 & 2). The text may be outside the bars when there is insufficient space within bars (fig. 3).
- Use the same bar color for each group within a categorical variable (figs. 1, 3, & 4), unless there is particular value in highlighting a specific bar (fig. 2).
- Add error bars to show uncertainty in the data (e.g., standard deviation, confidence interval) (fig. 4).
- The width of bars should be greater than the space between bars (figs. 1–4).

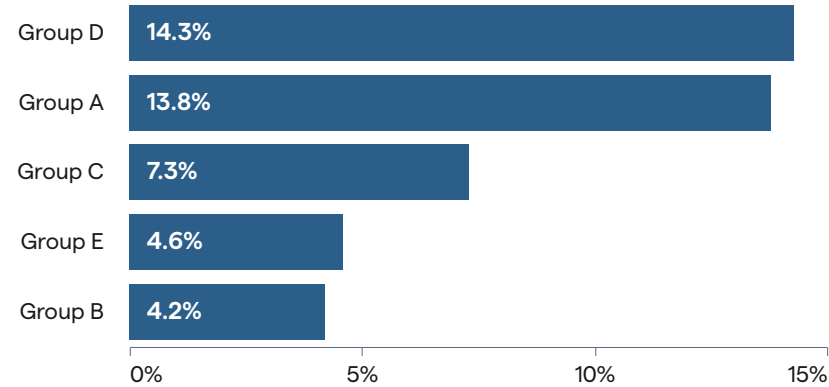


fig. 1

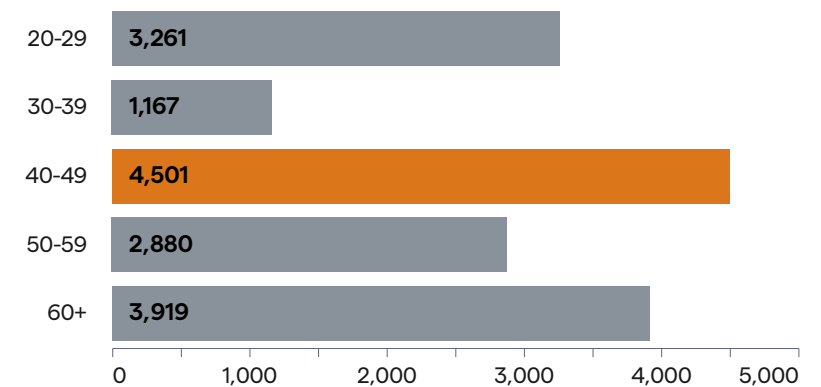


fig. 2

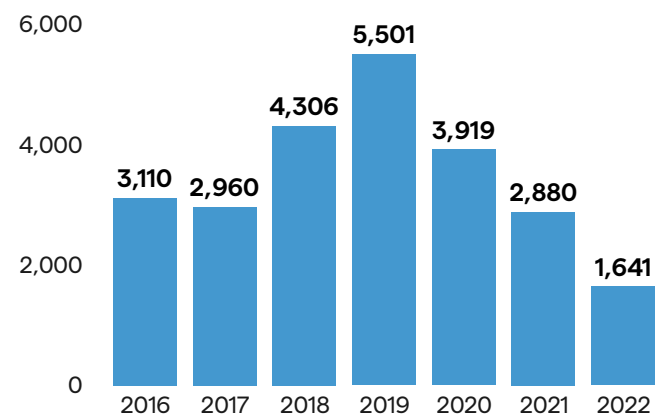


fig. 3

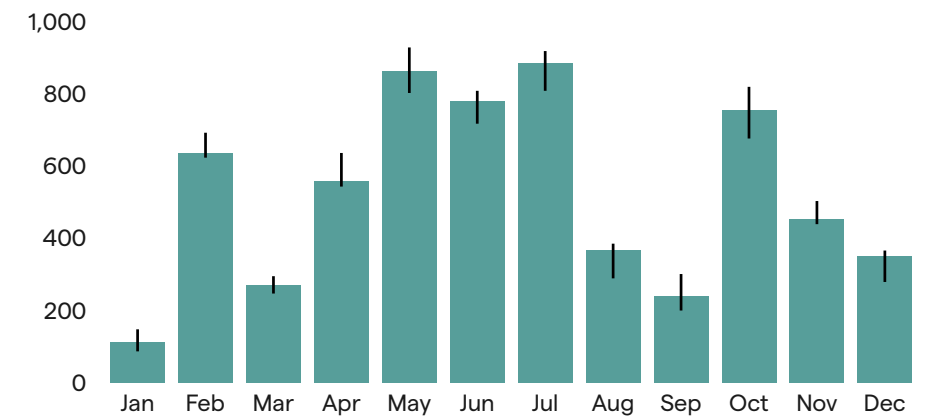


fig. 4

# 2.1

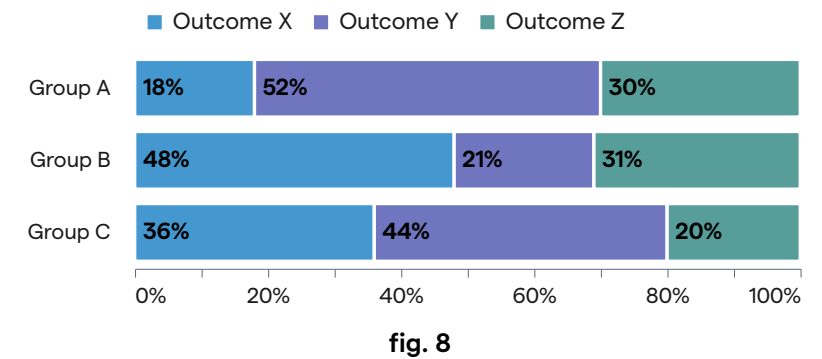
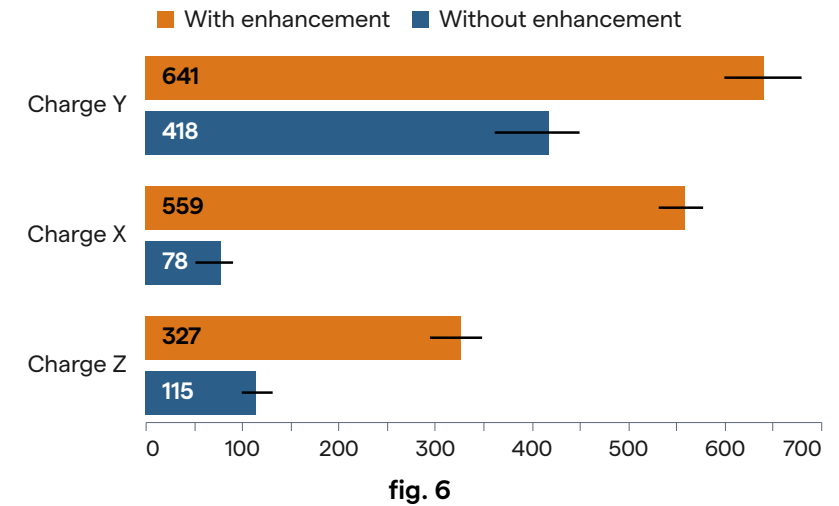
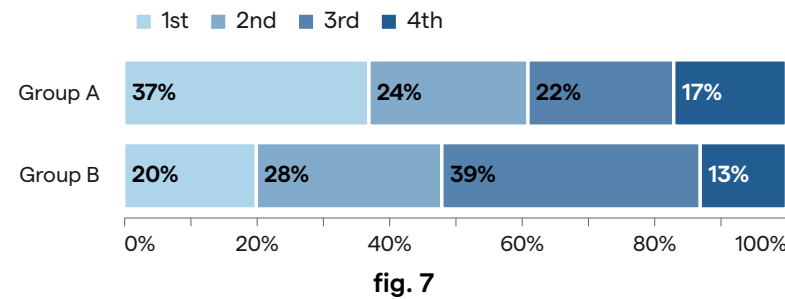
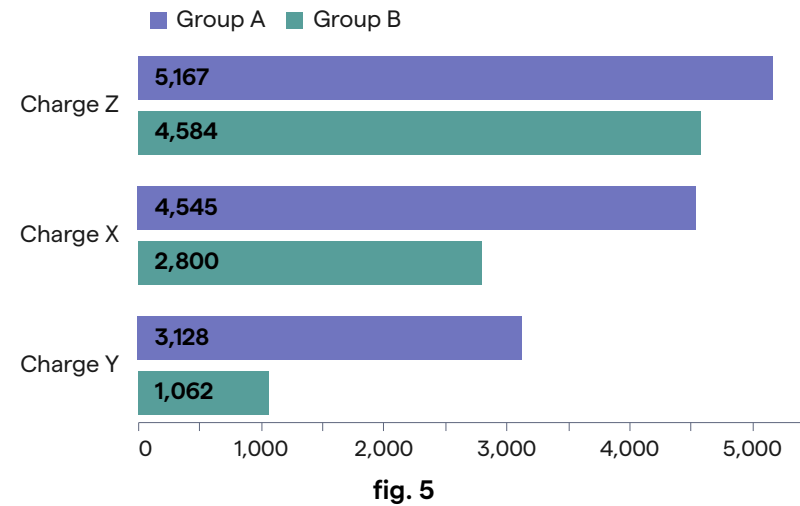
## When to use

Use more complex bar charts when there is an additional categorical variable to be displayed.

## Design guidance

- Grouped bar charts (figs. 5 & 6) are generally preferable to stacked bar charts since it is easier to compare bars starting at the same baseline.
- Use stacked bar charts (figs. 7 & 8) to show normalized data, where the segments of each bar add up to 100%.
- Stacked bar charts should not be used when the number of segments would exceed the number of colors in the combinations presented in Section 1.3 and Section 1.4.
- Separate the segments of stacked bars with at least 1 pt white space to improve readability (figs. 7 & 8).
- For ordinal variables, use variations of a single color (Section 1.4) for the segments of a stacked bar chart (fig. 7).
- Use black (#000000) or white (#ffffff) text for data labels, in accordance with Section 1, to ensure sufficient contrast between text and bar colors (figs. 5–8).

# Bar charts



## 2.2

### When to use

Use scatter charts to show the relationship between two quantitative variables. A third variable may be represented by variations in data marker color and shape (for categorical variables) or size (for quantitative variables).

### Design guidance

- Avoid plotting so many data points on a single chart that readability is impaired.
- Reduce the opacity of data markers (figs. 1 & 4) or use unfilled data markers (figs. 2 & 3) to improve readability of overlapping data points.
- Try adjusting chart dimensions and/or decreasing data marker size to reduce data point overlap.
- Use both data marker shape and color to differentiate groups for categorical variables (fig. 2).
- Add annotations to describe data points of interest, where applicable (fig. 3).
- Label groups directly to avoid use of a separate legend, when possible (fig. 4).
- When representing a variable using size (aka bubble charts), use circular data markers (figs. 3 & 4). The area of the circle must scale with the data value.
- A categorical variable differentiated by color alone may be used for bubble charts in rare cases where there are compelling patterns and the chart can still be clearly understood (fig. 4).

## Scatter charts

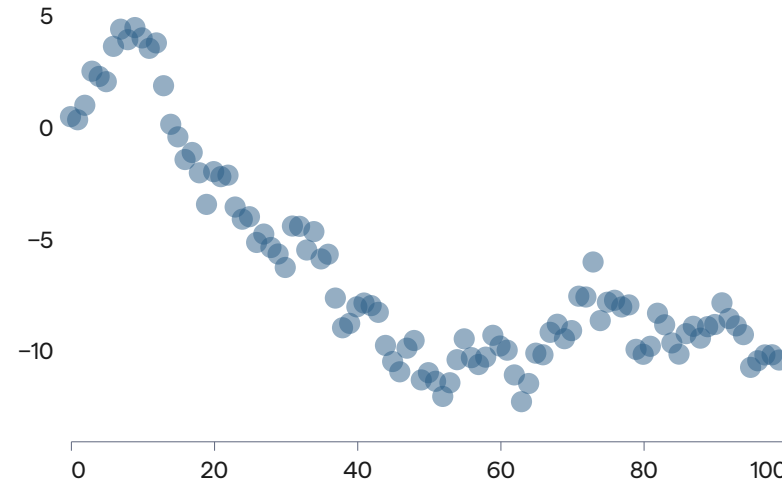


fig. 1

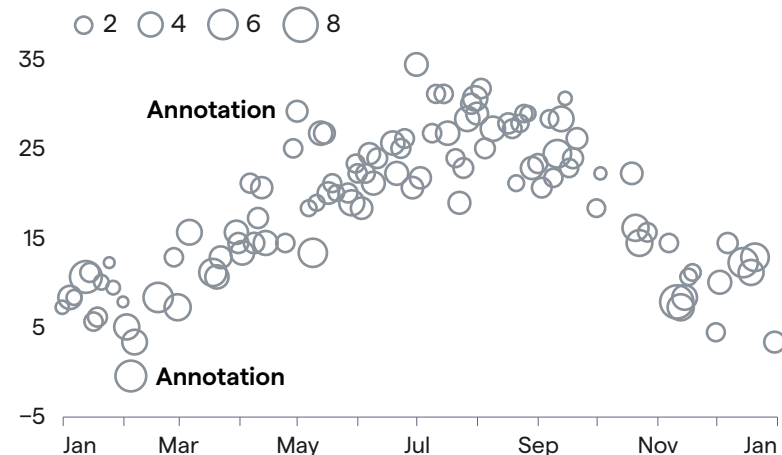


fig. 3

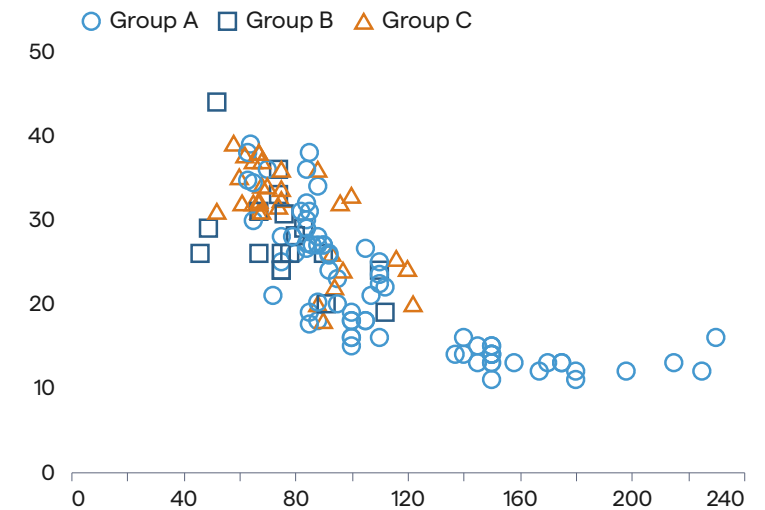


fig. 2

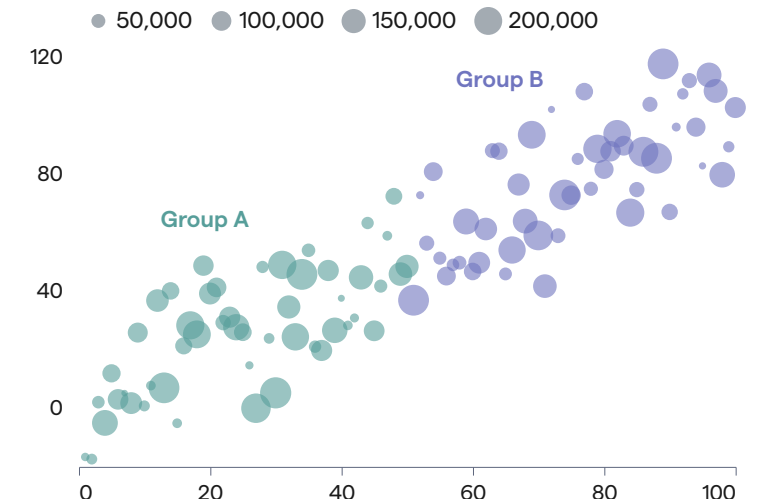


fig. 4

## 2.3

### When to use

Use line charts to show changes or trends in a quantitative variable over continuous values of a temporal or quantitative variable.

### Design guidance

- Line charts are most appropriate for continuous, dense data.
- Avoid having too many lines on a single chart. When using multiple lines on the same chart (figs. 2–4), ensure that the chart remains readable.
- Do not use color alone to distinguish lines on the same chart. Use at least two methods of differentiation, such as color and line style (fig. 2), color and line thickness (fig. 3), or color and data marker shape (fig. 4).
- Add reference lines and/or annotations for important events or points of interest, where applicable (fig. 2).
- When the number of lines would add too many colors to the chart, remove some lines or use gray to de-emphasize less important lines (fig. 3).
- Add data marker points for small data samples to avoid the implication that a line represents more continuous data (fig. 4).
- Label lines directly to avoid use of a separate legend, when possible (fig. 4).

## Line charts

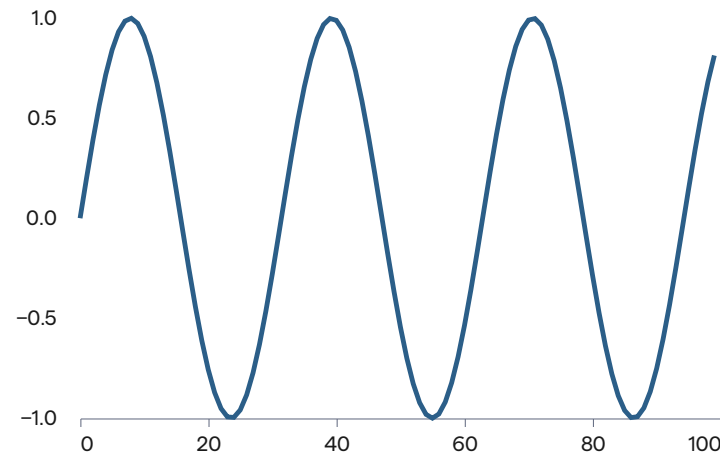


fig. 1

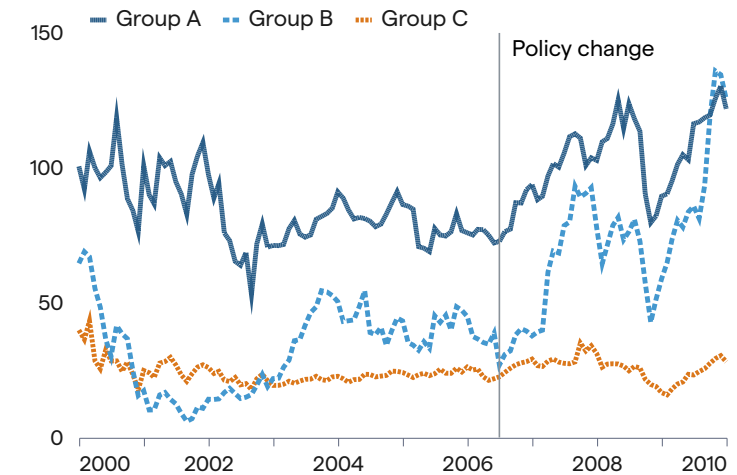


fig. 2

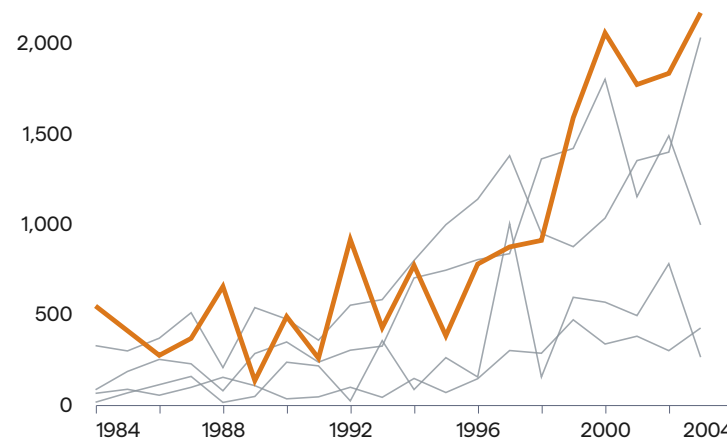


fig. 3

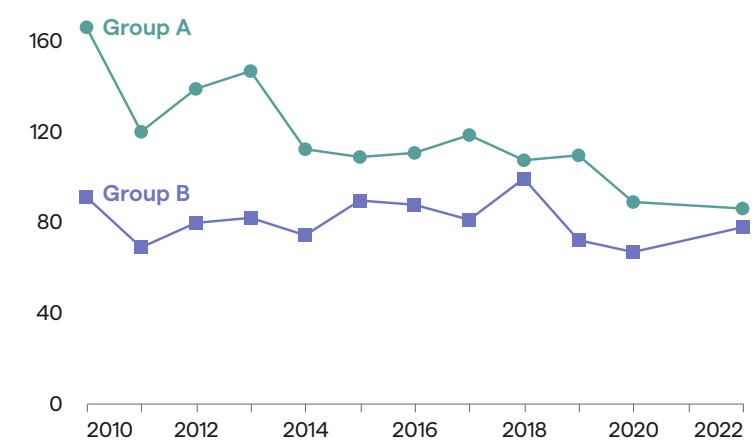


fig. 4

## 2.4

### When to use

Use the types of charts shown in this section to convey how a subset of data relates to the whole.

### Design guidance

- A waffle chart (fig. 1) consists of a grid of squares, usually 10 x 10, where the proportion of squares of a certain color represents a percentage. Waffle charts have comparable uses to pie charts, but are preferred because their shape is easier to perceive accurately.
- Use normalized stacked bar charts (fig. 2) for a simple, compact method of showing parts of a whole, where the segments of each bar add up to 100%.
- Sankey charts (fig. 3) represent the flow of values from one point to another. They provide the benefit of showing relative quantities, through the width of the flow lines, in addition to the part-to-whole breakdown.
- Tree maps (fig. 4) are useful for showing hierarchical data as a set of nested rectangles. Like Sankey charts (fig. 3), they provide the benefit of comparing the relative size of groups in addition to the distribution of values within a group.
- Include at least 1pt white space between segments of waffle charts (fig. 1), stacked bars (fig. 2), and tree maps (fig. 4) to improve readability.

## Part-to-whole charts

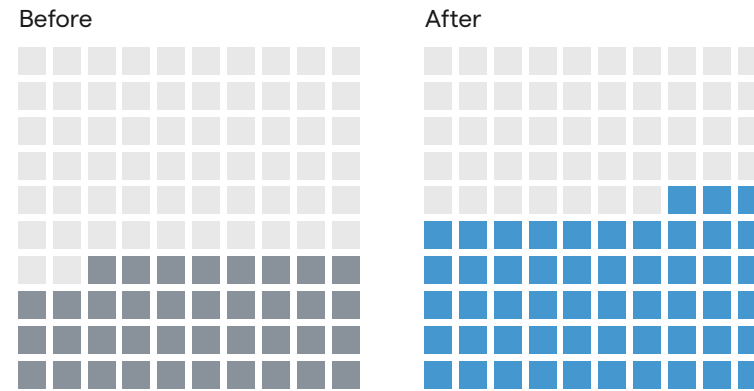


fig. 1

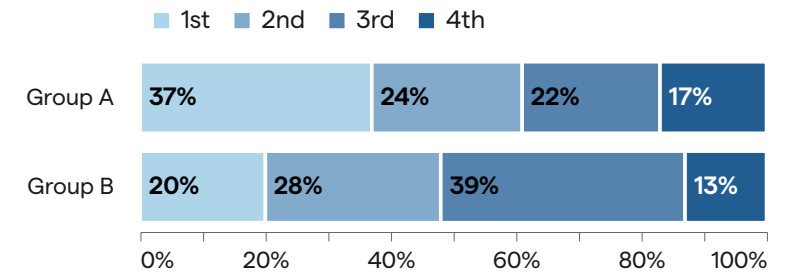


fig. 2



fig. 3

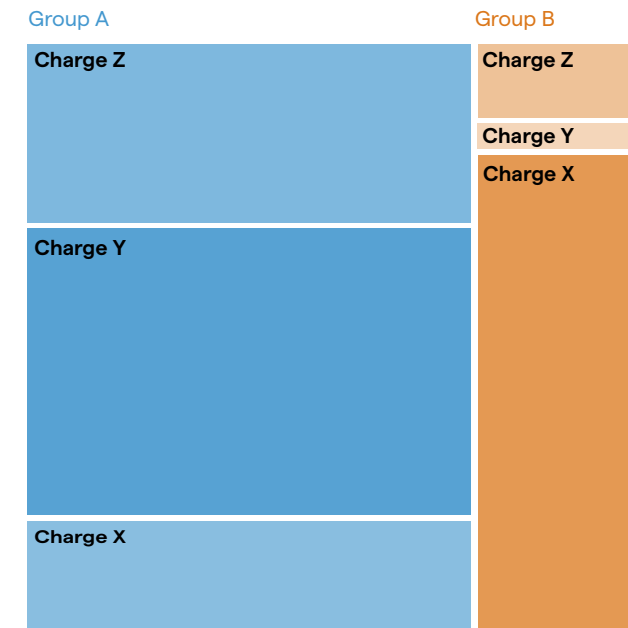


fig. 4

## 2.5

### When to use

Use these charts to show the distribution of values for a given quantitative variable over a continuous interval.

### Design guidance

- Histograms display the distribution of values across equally-sized bins, with the y-axis showing count or relative frequency. Histograms do not have gaps between bars (fig. 1).
- Density charts display the probability density function of a variable. Multiple density curves may be shown together with reduced opacity if the overlap does not prevent readability (fig. 2).
- Histograms (fig. 1) and density charts (fig. 2) always start from a zero baseline.
- Carefully select the bin width or bandwidth for histograms and density charts, as this influences distribution shape and affects visibility of trends.
- Figures 3–5 all show the same data. Boxplots (fig. 3), which include the median, interquartile range, and outliers, provide a distribution summary, but little information about specific data values.
- Strip plots (fig. 4) show the distribution of all data values, but may be unreadable with large datasets. Violin charts (fig. 5), which show density curves, are best used with large amounts of data.

## Distribution charts

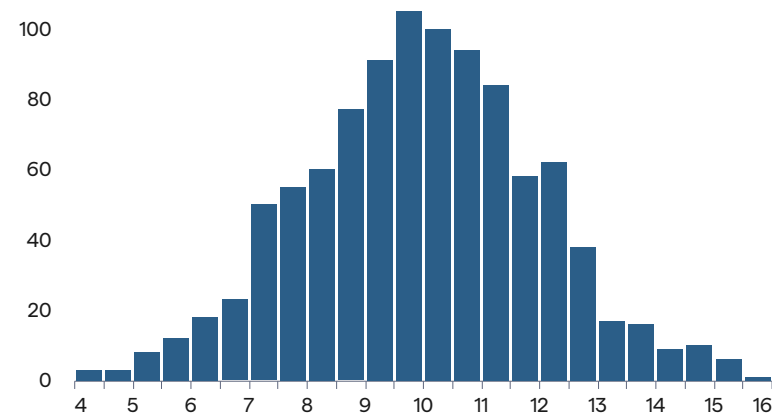


fig. 1

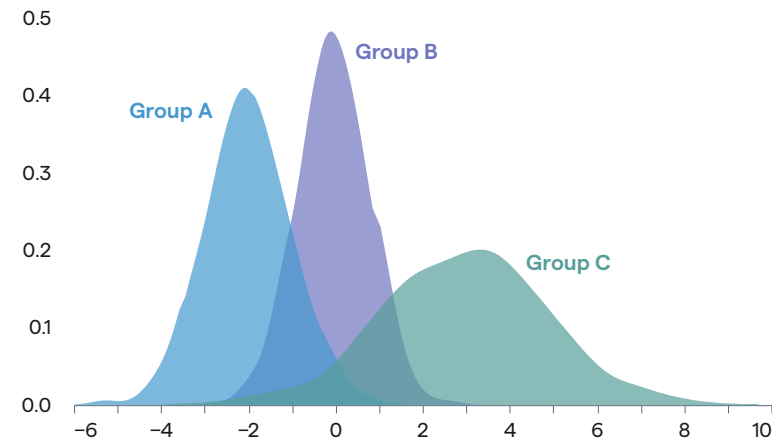


fig. 2

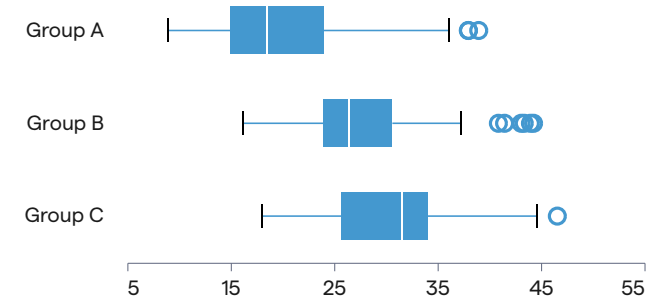


fig. 3

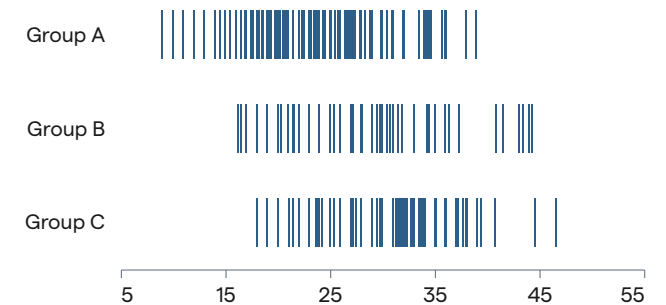


fig. 4

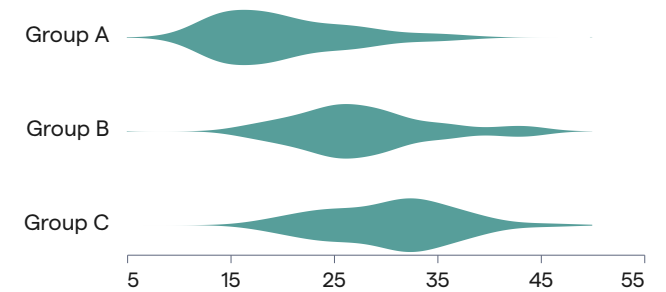


fig. 5

## 2.6

# Heat maps

### When to use

Use heat maps to show the numeric relationship between two discrete variables on a continuous, quantitative scale using color gradients.

### Design guidance

- Use continuous color palettes (Section 1.5) when creating heat maps.
- Use a single-hue palette for continuous, sequential data, such as from a low value to a high value (figs. 1 & 2).
- Use a diverging palette when there is a clear and meaningful midpoint in the data (figs. 3 & 4).
- The blue and orange diverging palette is more effective (compared to the purple and teal diverging palette) for emphasizing opposite extremes, such as in a correlation heat map with a midpoint of zero (fig. 3).
- Include a legend that shows the relationship between the numeric values and color scale (figs. 1–4).
- Separate the segments of heat maps with at least 1 pt white space to improve readability (figs. 1–4).
- Include data labels showing the numeric value for each cell, where there is sufficient space to do so (figs. 1, 3, & 4).

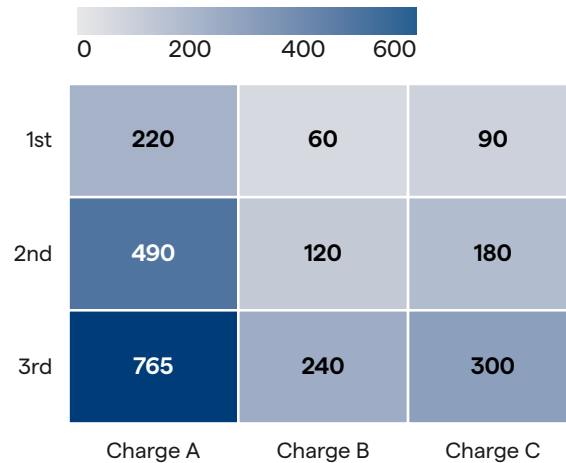


fig. 1

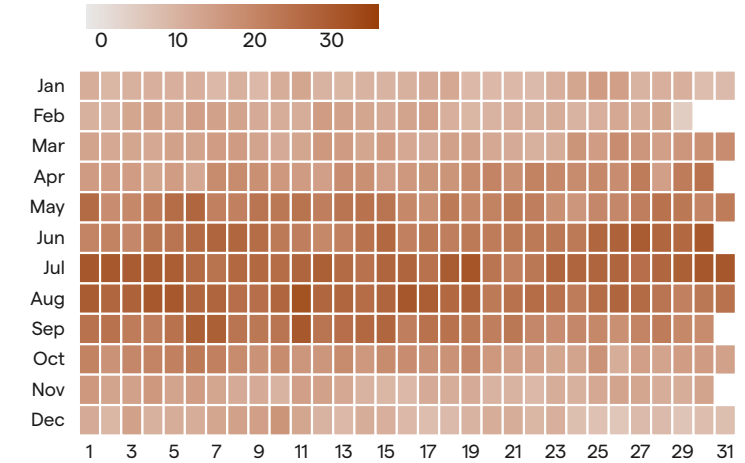


fig. 2

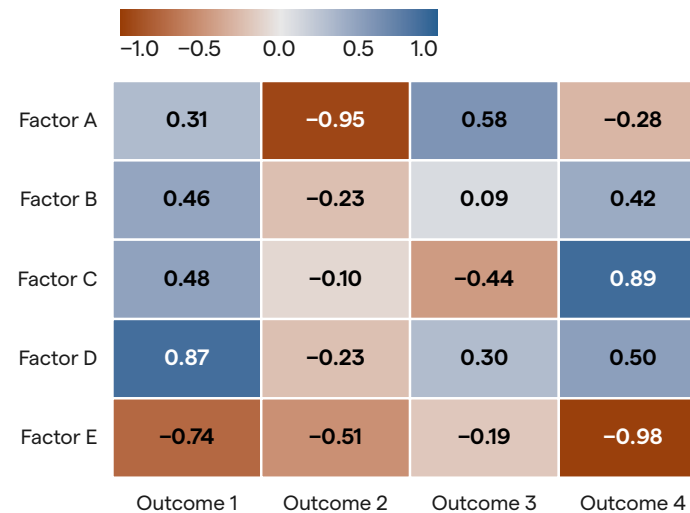


fig. 3

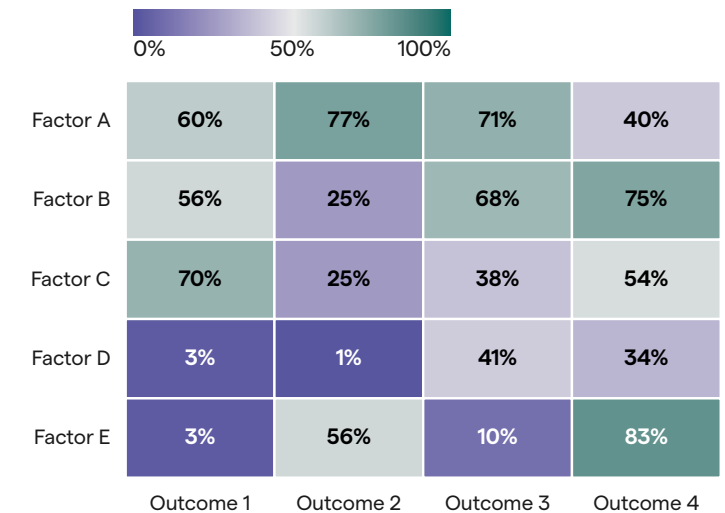


fig. 4



# 2.7

# Combination charts

### When to use

Combine multiple chart types within a single data visualization where this adds valuable information and context without reducing chart readability.

### Design guidance

- When combining chart types, ensure that the guidelines for each individual chart type are followed (figs. 1-5).
- When making combination charts, all data series displayed must have the same units so that they can be shown using the same axes. Do not create charts that require a second x-axis or y-axis (figs. 1-5).
- Ensure that all data series presented are on a similar scale so that trends are visible. Ensure that there is not so much data shown that the combination chart is difficult to understand. If these requirements cannot be met, use separate charts to convey the information (figs. 1-5).

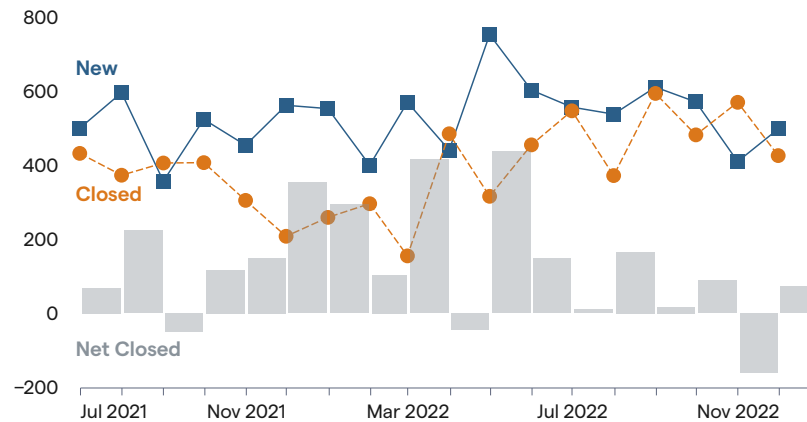


fig. 1

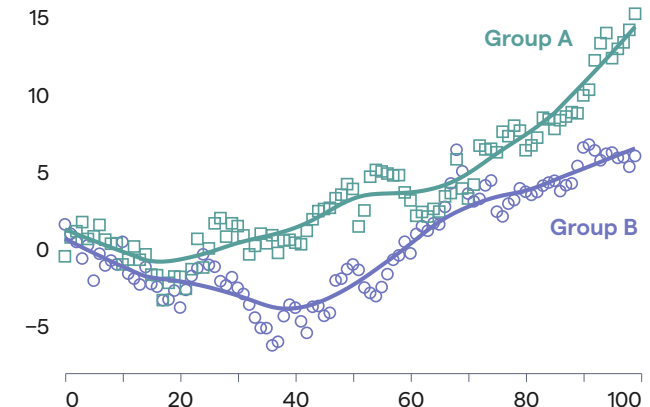


fig. 2

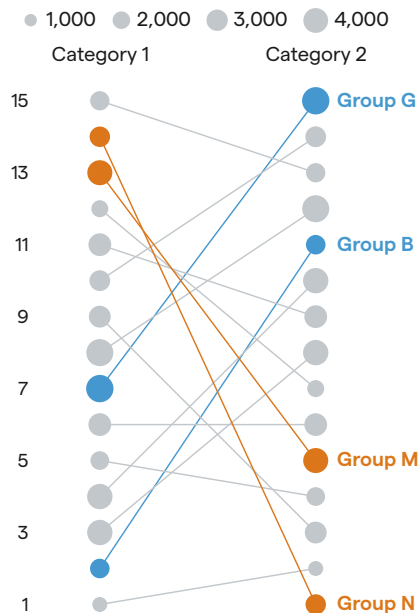


fig. 3

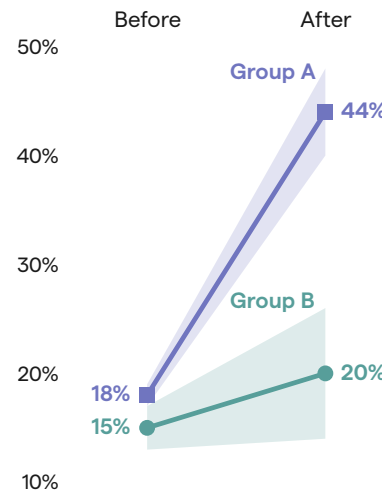


fig. 4

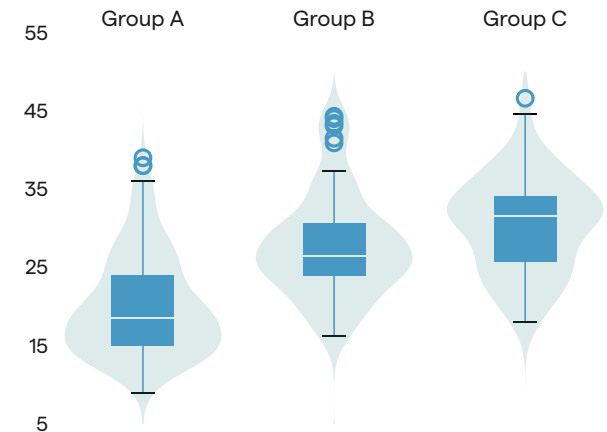


fig. 5

## 2.8

# Small multiples

### When to use

Use “small multiples,” a group of charts each showing the same combination of variables, to provide an effective means of comparison while avoiding over-plotting of a single chart.

### Design guidance

- Use the same axis units, axis ranges, and chart dimensions for each chart in a group of small multiples to facilitate ease of comparison (figs. 1–3).
- Axes do not need to be repeated for each chart. (figs. 1–3).
- Place small multiples charts in a logical order, such as ordering by a quantitative measure or an ordinal or temporal variable, when possible.
- The charts within a group of small multiples may be arranged in a grid (fig. 1), stacked vertically (fig. 2), or placed side-by-side horizontally (fig. 3). When not using a grid, it is best to align charts vertically when comparing x-axis values (fig. 2) and horizontally when comparing y-axis values (fig. 3).
- All charts in a group of small multiples should be visible together, without the need for scrolling or page breaks.
- Small multiples can be used to avoid color-coding groups of people (such as by race (fig. 2), ethnicity, sex, or gender), which may unintentionally reinforce stereotypes.

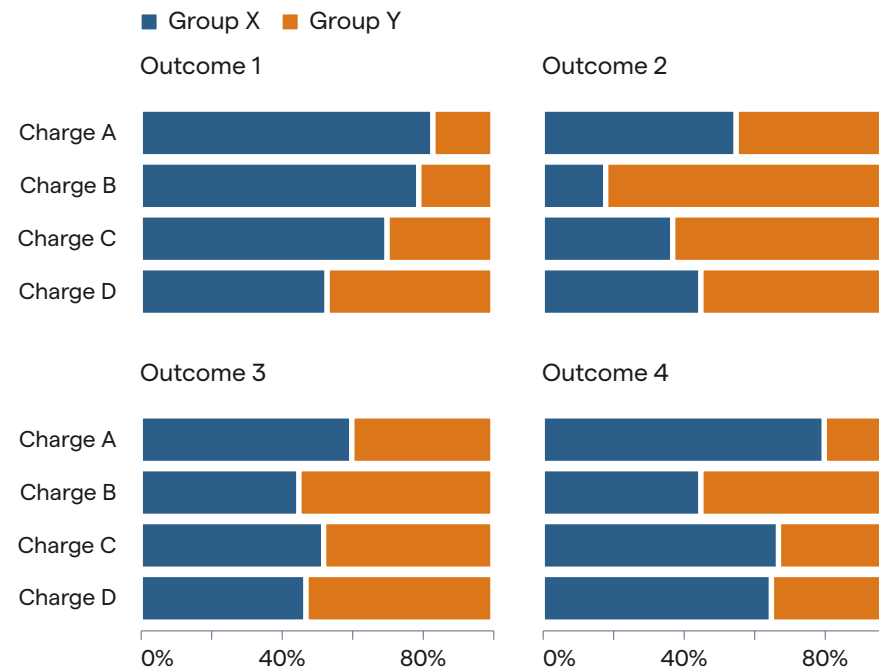


fig. 1

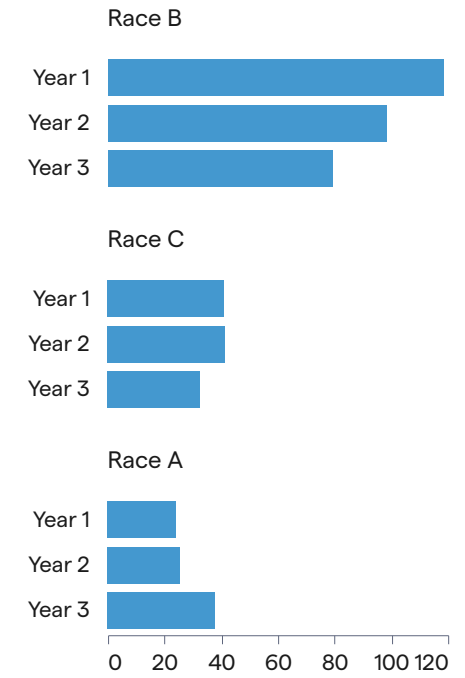


fig. 2

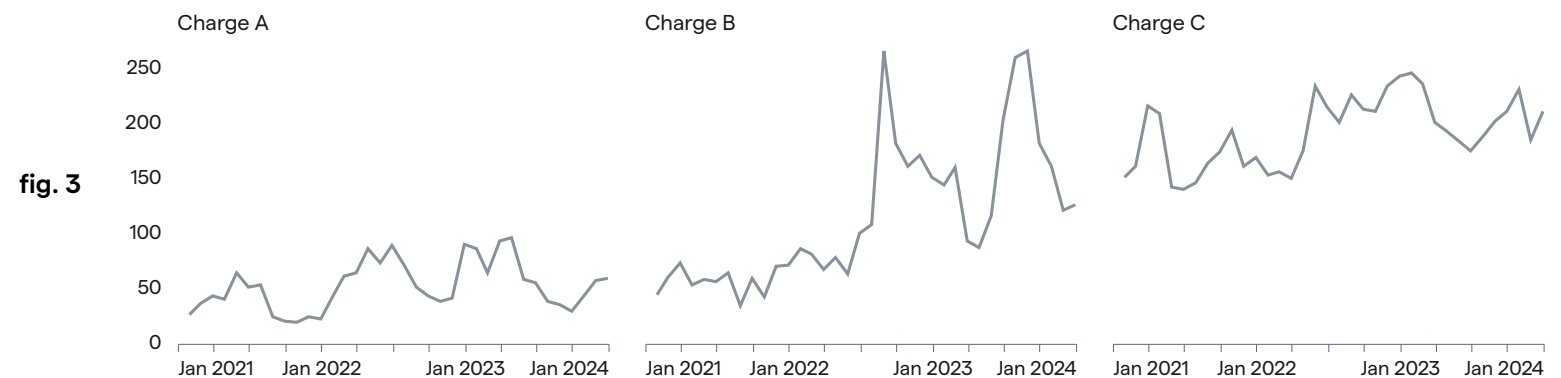


fig. 3

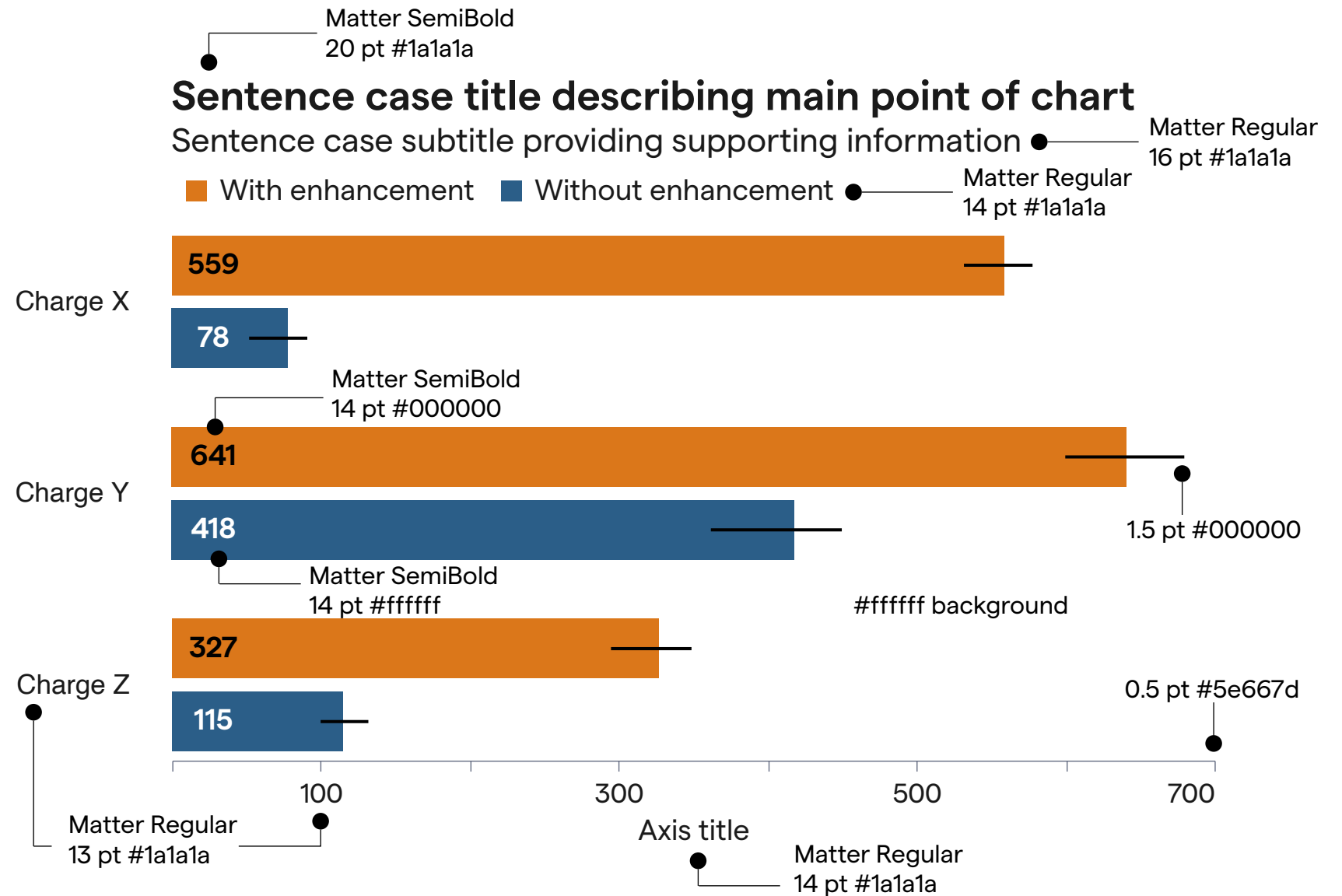
# 3

This section provides guidance for designing elements common to most chart types.

## Design guidance

- Use Matter, the sans-serif typeface used by JIL, for all chart text.
- Use the font sizes, weights, and colors specified for each text element.
- Left align titles, subtitles, and legends with the start of the x-axis.
- Right align y-axis labels.
- All chart text should be horizontal, not angled or rotated.
- The y-axis should not require a title. Ensure the meaning of the y-axis is clear based on the title, subtitle, or chart annotations.
- Include an x-axis title unless the units are clear without one, such as years.
- Keep charts free of markings not required for comprehension. Do not include a border around charts. Do not include grid lines, unless necessary. Reduce the number of axis tick mark labels, where appropriate.
- In general, make charts longer in width than in height, though dimensions may be adjusted as appropriate for the data and chart type.
- Directly label charts instead of using a separate legend, when possible. When using a legend, the meaning should be clear without adding a legend title.

# Chart details



## 4

# Additional considerations

This section discusses additional concepts that need to be considered when designing data visualizations. These concepts focus on how audiences perceive and comprehend data visualizations, which should always be a top consideration for designers.

It is important to recognize that there are a wide variety of conditions that can impact a person's ability to understand data visualizations, such as blindness or low vision, color vision deficiency, cognitive or neurological issues, and motor or dexterity issues. Creating data visualizations that are universally accessible is extremely difficult, but designers should strive to reach as wide an audience as possible.

This section focuses on how to make data visualizations that are clear, accessible, and inclusive. Some of this guidance falls under multiple topics and may also be addressed elsewhere in this document.

While this section provides a range of guidance, it should be noted that this is not an exhaustive list. Additionally, language and accessibility standards are constantly evolving. It is important that data visualization designers stay up-to-date on best practices.

### Clarity and accessibility

- Use common chart types for most data visualizations, since charts are easiest to understand when straightforward and familiar. There may be situations when more complex and/or innovative chart designs add particular value, but less complicated charts are accessible to wider audiences.
- Write descriptive chart titles to assist users in drawing appropriate conclusions, and always include text explaining the data visualization in the accompanying document or web page. Describe data trends so that users do not need to rely solely on graphic elements for comprehension.
- Label data series directly instead of using a separate legend to reduce the mental processing required for comprehension. Add data values and annotations to highlight important trends and add context.
- Make source data available via tables or data files so users can explore the data in other ways. Do not include more mathematical precision than is necessary for data values.
- Reduce chart clutter by avoiding overplotting and by removing chart elements such as borders, gridlines, and excess axis ticks and labels.

### Inclusivity

- Use inclusive, people-first language. Do not use dehumanizing terms such as "inmate." Avoid using the term "Other" for people who do not fit into a specific data category; instead, use terms like "another race" or "identity not listed."
- Accurately portray the communities represented by the data. Be aware that demographic data may be based on the perceptions of those collecting the data as opposed to self-identification. Explain data collection methods and potential accuracy issues, and identify any groups excluded from the data.
- Show variability in data when comparing demographic groups to provide appropriate context and prevent broad generalizations. Add error bars or confidence bands to charts, or use a distribution chart such as a boxplot.
- Carefully consider the best way to order demographic groups in data visualizations. Do not lead with a dominant group by default. Order groups in a logical manner based on the data being shown and the purpose of the chart, such as ordering demographic groups based on a quantitative metric.
- Avoid using data colors that reinforce racial, gender, or political stereotypes.

### Design for web use

- Provide alternative (alt) text for each chart. Include short (chart type and purpose) and long (chart composition, variables, axes, trends) descriptions.
- Ensure font sizes are no smaller than those specified in Section 3.
- Size charts so they can be fully seen without scrolling. Multiple versions of the same chart may be required for different screen sizes.
- Make important chart elements, including interactive features and tooltips, accessible using keyboard navigation and assistive technologies, such as screen readers. Provide instructions for any interactive features.
- Do not use color alone to indicate a state change for interactive elements. Colors used for different states must have at least a 3:1 contrast ratio.
- For animated charts, give the user the ability to turn off animation. Provide options to restart animations and to play them at different speeds.
- Tooltips can be added to web charts to allow users to see specific values and other information when hovering over data points. However, the chart must be understandable without tooltips. Design tooltips to be accessible, following the Web Content Accessibility Guidelines.