CENTER ON RURAL INNOVATION

Data Visualization Style Guide

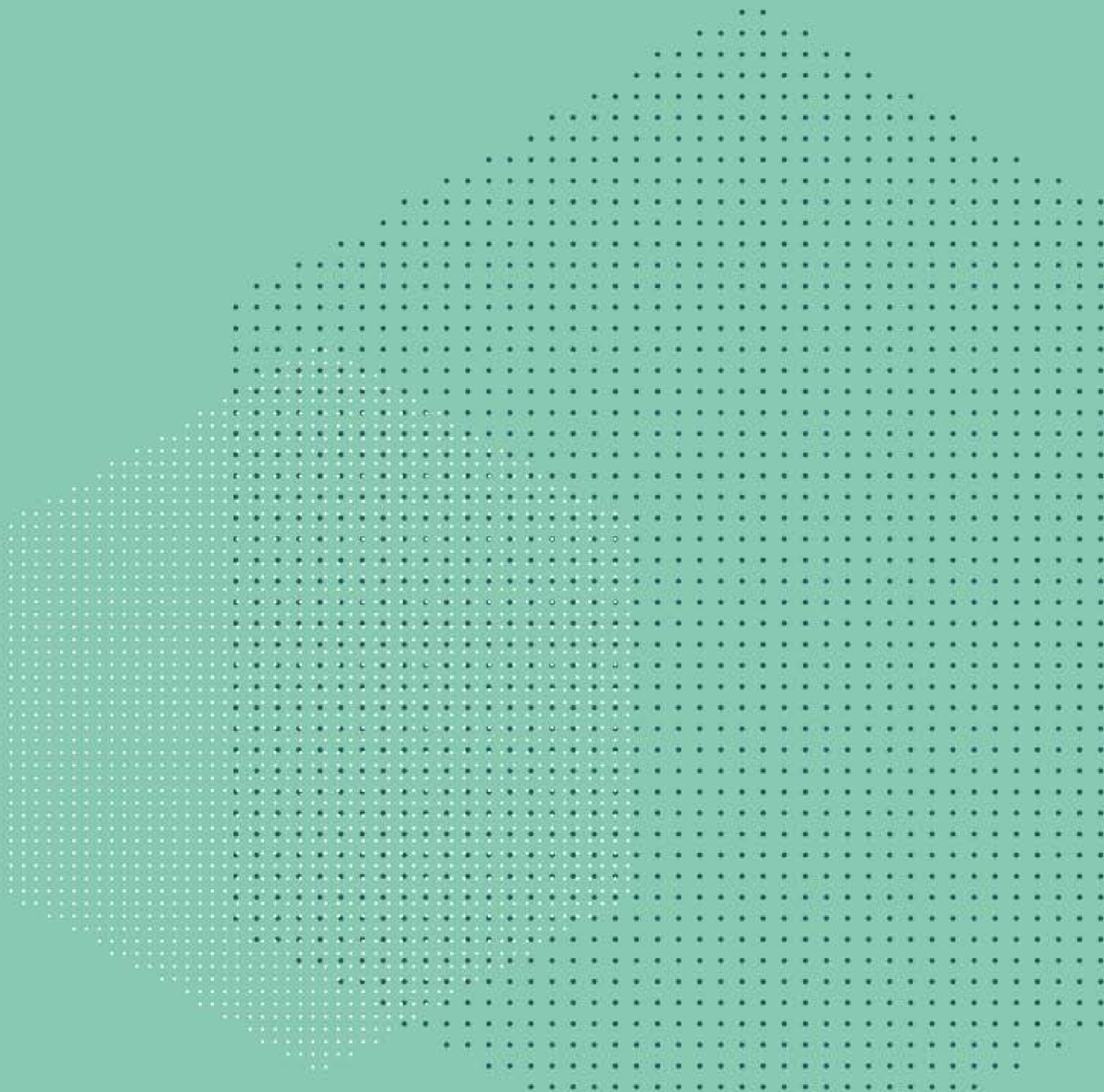


Table of contents

Foundations

Typography 1

Colors 2

General layout 5

Four steps to make a chart 6

Chart types

Change-over-time

Line chart 10

Area chart 12

Part-To-Whole

Pie chart 15

Stacked bar 17

Comparisons

Bar chart 20

Slope chart 22

Scatter plot 24

Dot plots 26

Distribution

Beeswarm 29

Histogram 30

Spatial

Choropleth 33

Bubble map 36

Density plot 37

Tables

Checklist

This PDF is interactive. Click on any item in the table of contents to navigate the document.

Foundations

These guidelines summarize Center for Rural Innovation's graphic language for visualizing data. Informed by our brand voice, this graphic language will help us communicate data and stories to our various audiences across all our channels. These recommendations should be followed as closely as possible to ensure consistent implementation of this graphic language.
It is impossible to predict every situation, but this guide will help inform these decisions and make sure we are effectively informing, engaging and inspiring our audience through data.

Typography

Title	**Family**: TT Hoves Variable **Weight**: **Bold** **Size**: 20px **Color**: Black, #121E22	**Sentence case**	Gives the audience a sense of what the chart is about. Should be kept to one or two lines.
Subtitle	**Family**: TT Hoves Variable **Weight**: *Italic* **Size**: 16px **Color**: Black, #121E22	Sentence case	Optional. Subtitles highlight what's important in the chart, but should be kept to one or two lines.
Axis labels	**Family**: Lato **Weight**: Regular **Size**: 10px **Color**: Black, #121E22	Sentence case	Always horizontal for legibility.
Axis titles	**Family**: Lato **Weight**: Regular **Size**: 12px **Color**: Black, #121E22	Sentence case	Always horizontal for legibility. Indicates units and multipliers in parenthesis e.g. (millions).
Direct labels	**Family**: Lato **Weight**: **Bold** **Size**: 10px **Color**: Black, #121E22	Sentence case	Centered directly above the label they correspond to.
Sources and Notes	**Family**: Lato **Weight**: *Italtic* **Size**: 11px **Color**: Black, #121E22	Sentence case	Horizontal and left-aligned below the chart.
Legend title	**Family**: Lato **Weight**: Regular **Size**: 12px **Color**: Black, #121E22	Sentence case	Can be avoided by labeling data representations directly. Use clear language and avoid acronyms.

Colors

Colors in data design are an important information canal. They can effectively convey stories when used wisely. Make sure to use consistent colors. For example, if you use purple to denote category A, use the same color for A throughout the report.

Categorical color palette

The categorical color palette is ordered so that the colors have alternating lightness values. This order is recommended for visualizations with 3+ series.

Needing more than seven colors in a chart is a sign that you might want to use another chart type or consider grouping categories together.

Gray
#D9DBDE

Extended palette

Emerald #00825B
Pink #C24C93
Cobalt #234FBF
Skye blue #3A9EBD
Fire #E64E2E
Purple #755BA3
Red #BA3434
Squash #DF7B22

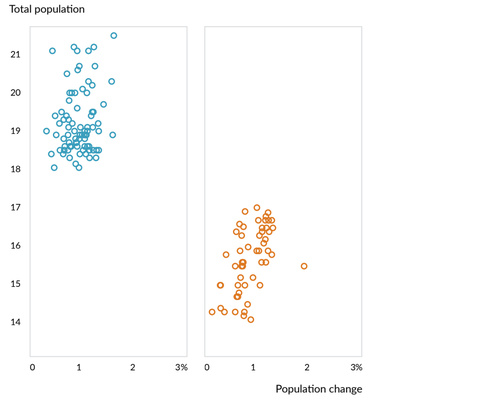
Don't underestimate gray

Consider gray as the most important color in your palette. You can use it for less important elements in your chart in order to highlight the most important points, and draw the reader's attention. You can also use gray for:

- General context data
- Less important annotations
- Muting the overall visual impression of your charts

Rural
Non-Rural

An example using consistent color coding in your data visualizations.



Highlighting colors

Opt for emerald as a highlighting color and gray to deemphasize other data points.

Emerald #00825B
Gray #D9DBDE

Background colors

The standard background color for charts should be white. When creating maps, use cream as the base map fill.

White #FFFFFF
Cream #FAF7E8

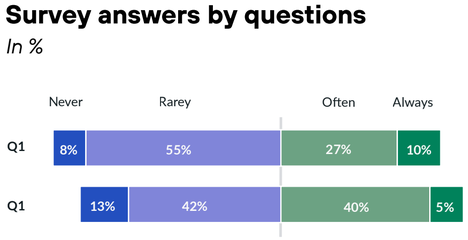
Colors

Diverging color palette

Diverging color scales are similar to sequential scales because they map quantities. The difference is that these colors are designed to show the distance (and direction) from a center point such as zero.

Negative **Neutral** **Positive**

#BA3333 #FAF7E8 #3053a1

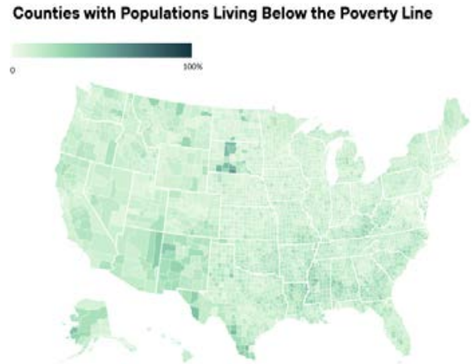


Sequential color palette

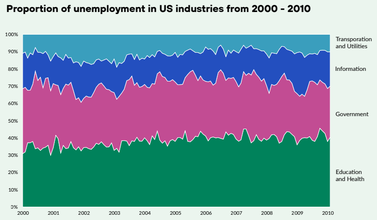
When color is used to represent a numeric value or quantity, you should use a sequential color scale—sometimes referred to as a color ramp or a gradient. The simplest example would be something that starts out white on one side and becomes darker until you end up at black on the other side. It's generally best to map small values to the lighter side and larger ones to the darker side.

Starting color **Middle color** **Ending color**

#F0F9E8 #7DBAA3 #16343E

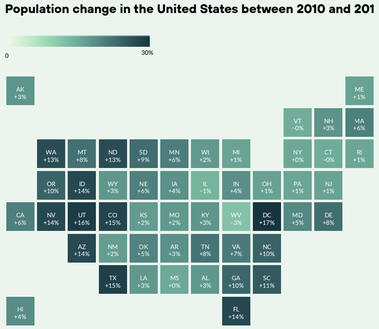


Use of colors



Determining categories

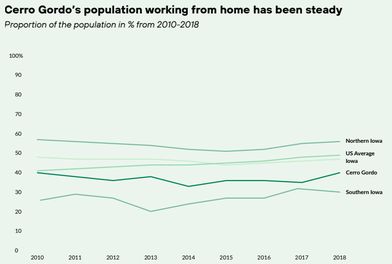
Although you can use different hues to determine categories, it's more effective to limit the number of colors and rely on variation in intensity to express the categories in your data.



Expressing intensity

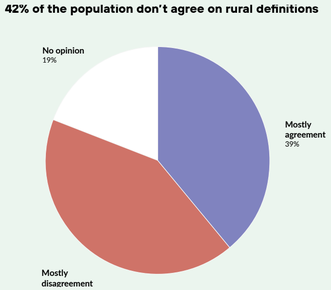
Use variation in a color's value to express the intensity of a phenomenon (the darker = the more intense, the lighter = the less intense).

You can either use a sequential palette, or a diverging palette when visualizing a polarized phenomenon.



Prioritizing information

A more efficient and effective way of using hues is to draw the reader's attention to specific values, therefore contributing to the chart's story.



Positive and negative values

Reserve the use of red for negative values/negative connotative information. Use another contrasting color to highlight a value that doesn't have negative connotation.

General layout

Titles

Titles should be brief and tell the reader what the chart is about or, even better, what conclusions they can draw from the data that the chart represents. Where possible, put units and date information in the subtitle.

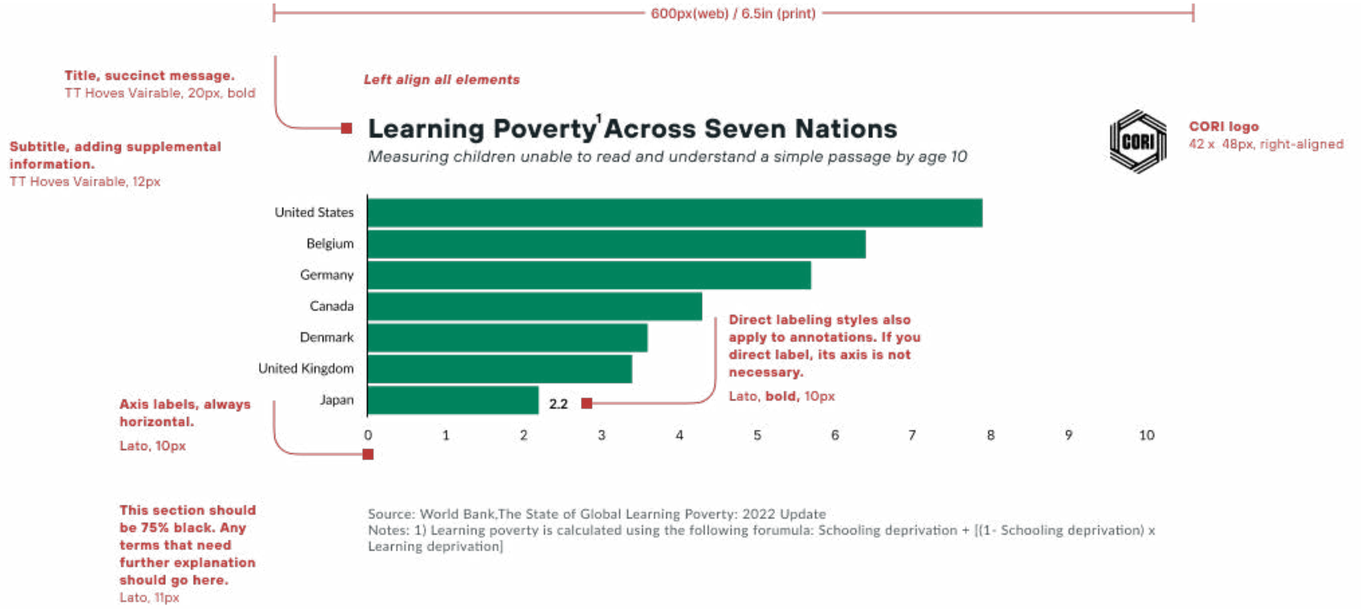
Legends

Legends are vital in understanding what colors, marks and sizes represent. In many instances, direct labeling can be more effective than legends as it prevents the audience from having to look back and forth between the chart and legend.

Axis

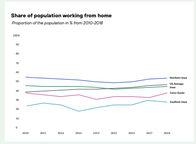
If specific units are used (pounds, miles, etc.), the symbol can be added to the highest value in the axis thus leaving the other values unlabeled. If you are direct labeling values or categories, they are redundant and should be entirely left out.

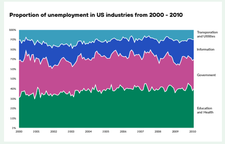
0 20 40 60 80 100%



Four steps to make a chart

	A	B	C	D	E	F	G
1	Share of Employment Population Working from Home %						
2	Region	2010	2011	2012	2013	2014	2015
3	Cerro Gordo	4	3.8	3.6	3.8	3.3	3.6
4	Iowa	4.8	4.7	4.7	4.7	4.6	4.4
5	Northern Iowa	5.7	5.6	5.5	5.4	5.2	5.1
6	US	4.1	4.2	4.3	4.4	4.4	4.5



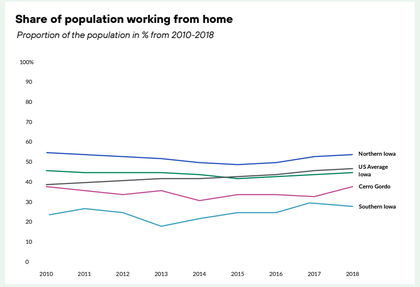


Step 1:
Examine your data

This data has both categorical and timeline data. Do we want to compare regions overtime? A specific year or a specific region? Once we have our question, we can choose an appropriate chart. In this example our question could be: How does Cerro Gordo's population that work from home deviate from the US average over time?

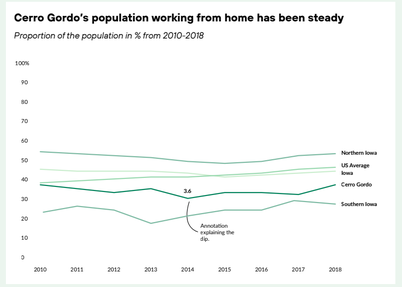
Step 2:
Examine your options

As we want to show a change-over-time, we will look in that category and consult our options.



Step 3:
Choose a chart

Line charts show a continuous flow over time, and we want to focus on Cerro Gordo, let's go with a line chart.



Step 4:
Highlight and annotate

In this step we make stylistic decisions. There are some ways we can use color and line thickness to highlight our data points.

Chart types

What data are you highlighting?

Change-Over-Time

A changing trend over time.

Part-to-whole

How an entity is broken down into its components.

Comparisons

Comparing the difference in values between categories.

Distribution

How data is grouped, or spread out over a certain interval.

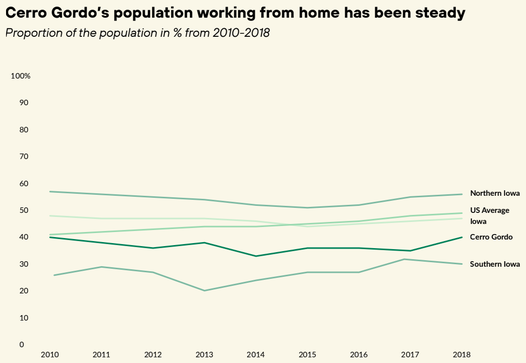
Spatial

Geographical patterns.

Tables

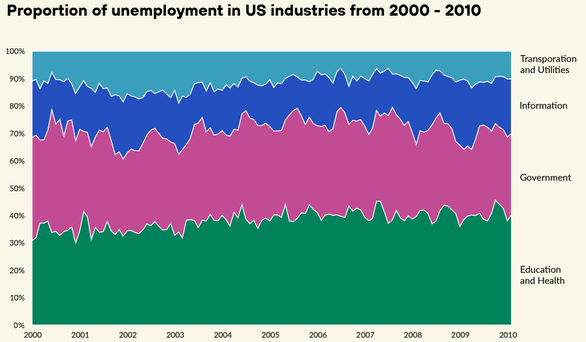
The exact number of every value in your data.

Change-over-time



Line chart

Evolution through a continuous time period.



Stacked area chart

Proportions over time.

Chart types

Line chart

Evolution through a continuous time period.

Best for
- Portraying both negative and positive evolution over time

Not to use when
- The timeline is discontinuous

Considerations
- Be careful in choosing the right time frame (not too narrow, not too wide)
- Annotations are especially helpful to further explain the context

Color schemes

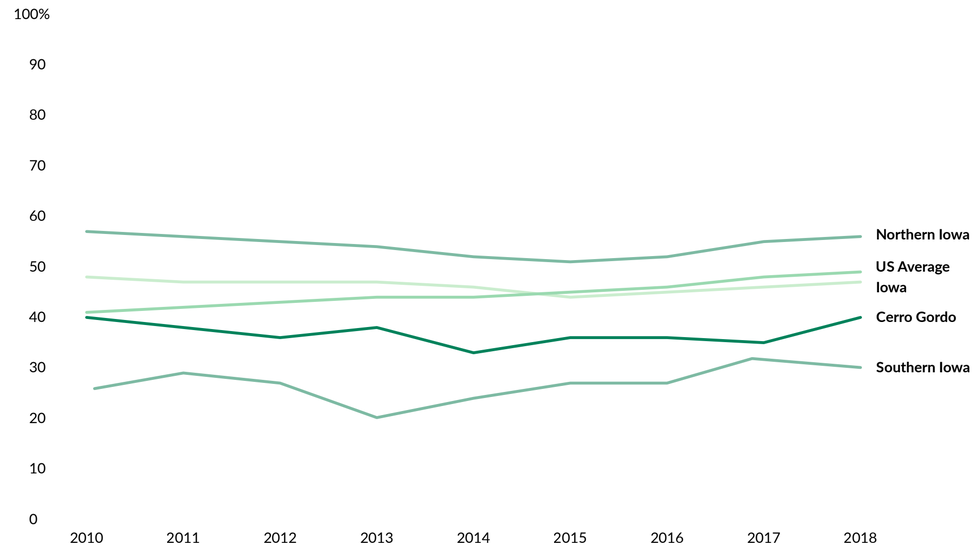
Categorical color palette

Sequential color palette

Diverging color palette

Cerro Gordo's population working from home has been steady

Proportion of the population in % from 2010-2018

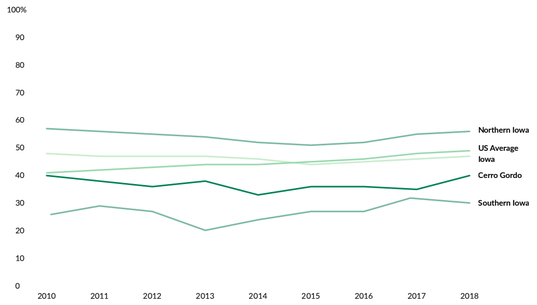


Line chart

Usage

Cerro Gordo's population working from home has been steady

Proportion of the population in % from 2010-2018

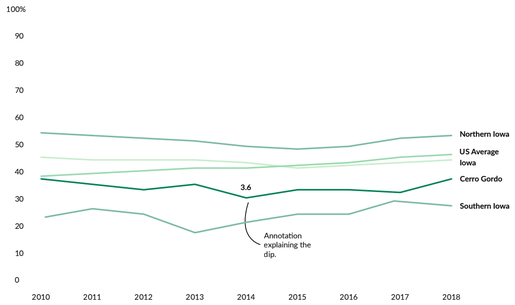


Use color to highlight

If using multiple lines, prioritize one by highlighting it with a specific color, then use the other lines as trend reference points. If you still choose to use a categorical palette for your time series, make sure there are less than five series and therefore less than five colors in the chart.

Cerro Gordo's population working from home has been steady

Proportion of the population in % from 2010-2018

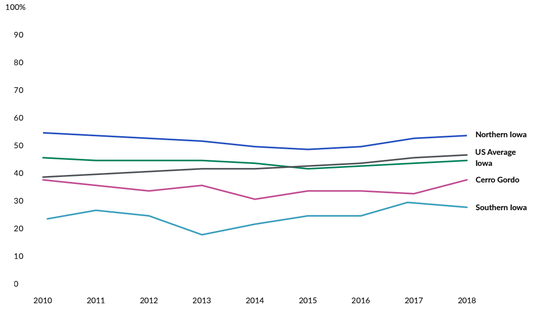


Add annotations

If showing multiple highlight numbers, prioritize one number by making it larger than the rest in order to stand out.

Share of population working from home

Proportion of the population in % from 2010-2018



Use colors sparingly

Avoid too many colors as these will confuse the reader.

Chart types

Area chart

Reveals the proportion of a whole over time.

Best for

- Emphasizing a part-to-whole relationship over time when one part is very large, or changes from being very large to very small
- When the total is as important as its shares
- When there are large differences between values
- You have many dates (10+)

Not to use for

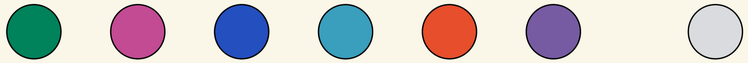
- Data sets with scales that do not have a meaningful relationship to zero
- Conveying specific values

Other considerations

- Your most important value should be at the bottom or the series should be in descending order
- Direct label categories
- Consider grouping multiple values into one bigger value

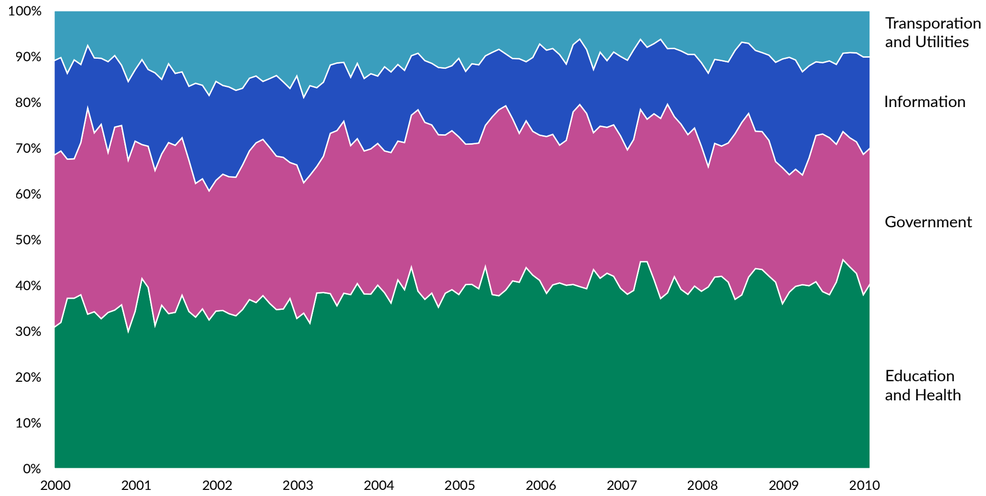
Color schemes

Categorical color palette



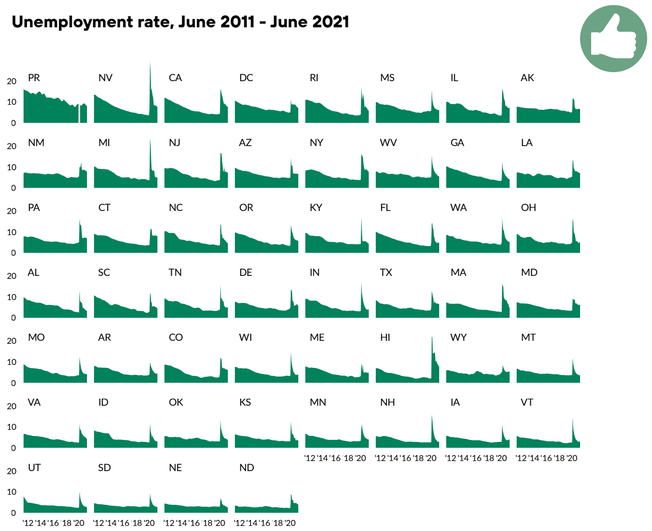
Sequential color palette

Proportion of unemployment in US industries from 2000 - 2010



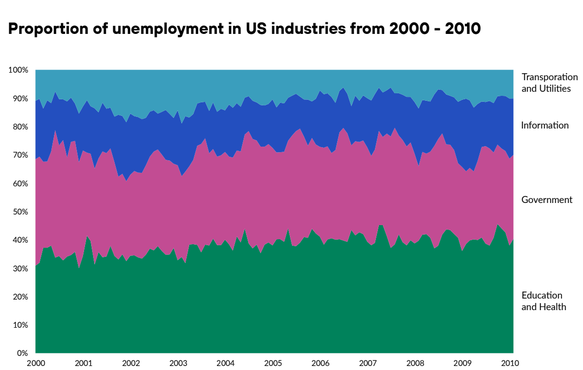
Area chart

Usage



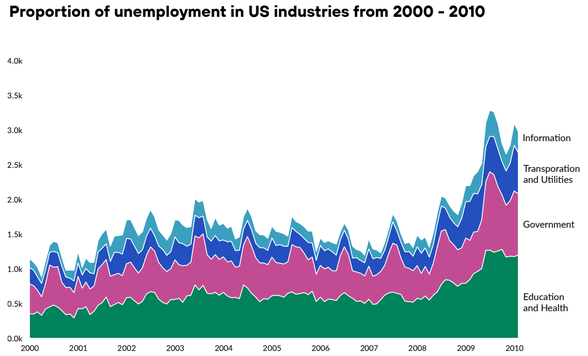
Consider using small multiples

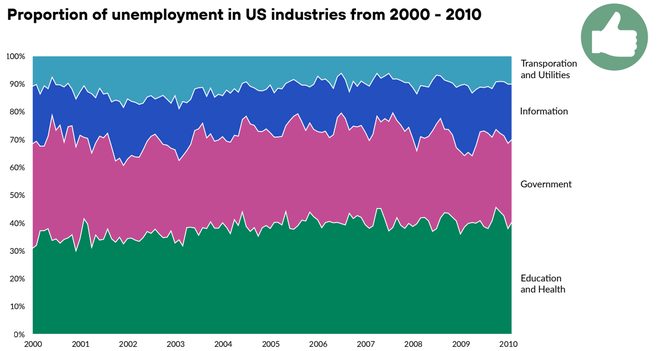
If you have a lot of series to show, small multiples could be a useful tactic to visualizing all of the series without cluttering a chart.



Use a stroke between series

Strokes increase contrast between the series and make for a more legible area chart.

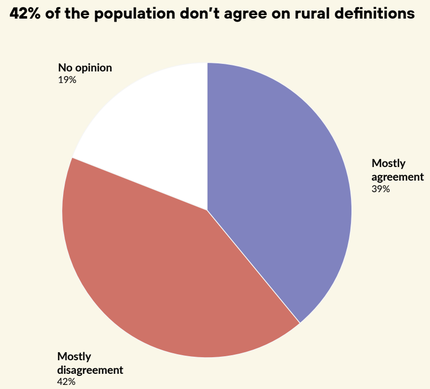




Consider normalizing your area chart

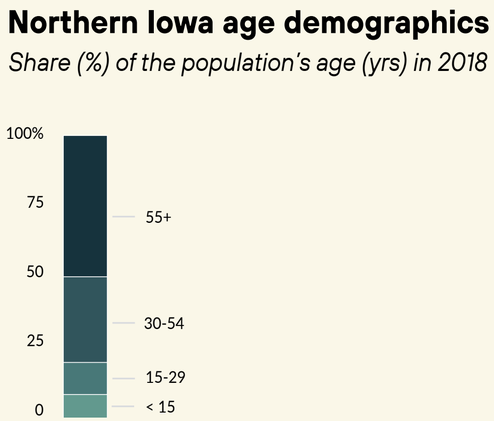
Area charts can be difficult to read. There are a variety of different styles - stacked, like the above, where values are separated and stack onto each other, or overlapping values. Compare the areas to each other can be trying, especially when they have similar values. If you want to show how the proportion changes overtime, a normalized area chart might be a good option.

Part-To-Whole



Pie chart

Shares of a certain entity



Stacked bar

Shares of a single entity

Chart types

Pie chart

Shows how a whole is broken down into smaller parts. Pie charts present a strong feeling of a single entity broken down into shares thanks to its round shape. Keep in mind that since the human brain is not great at estimating angles, pie charts are not a really effective visualization.

Best for

- Showing how a whole is broken down into smaller parts
- Values around 25%, 50% or 75%

Not to use when

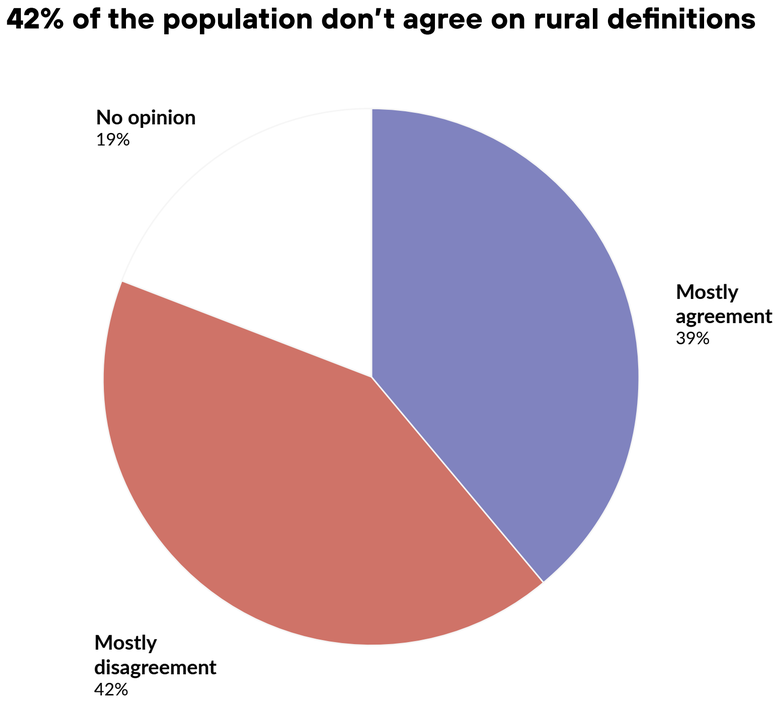
- Your values don't add up to 100%
- You're more interested in the quantities than the shares (use a magnitude chart instead)
- You have more than 4 categories (use a stacked bar or tree map instead)
- The proportions are similar
- A value is lower than 3%

Color schemes

Highlight color

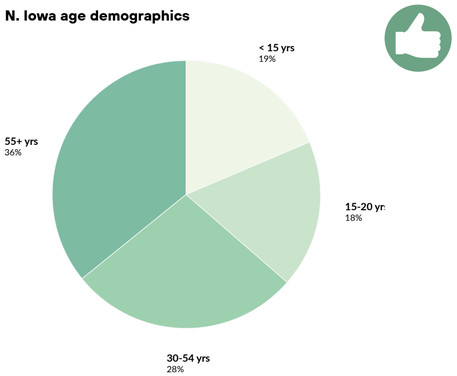
Sequential color palette

Diverging color palette

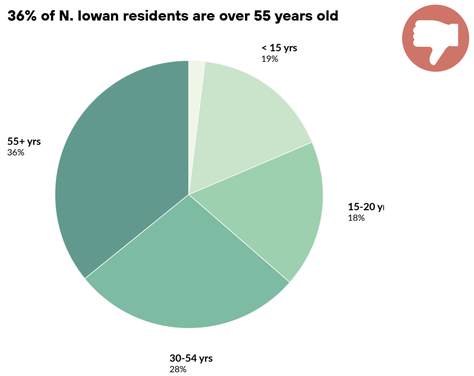


Pie chart

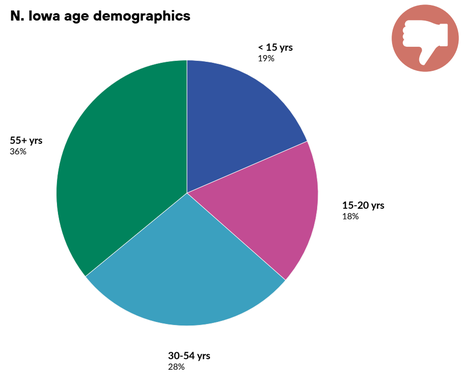
Usage

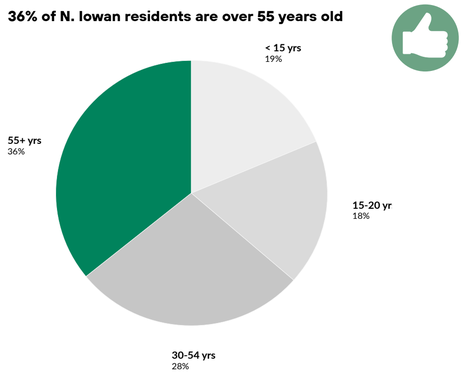


Directly label slices



Don’t use more than four categories





Use color to highlight

Chart types

Stacked bar

Demonstrates how a whole is broken down into smaller parts. The stacked bar can be presented vertically, horizontally, in groups or alone. If you're more interested in quantities over shares, use a categorical chart instead, such as a bar chart.

Best for
- Showing a whole broke down into smaller parts
- Representing a likert scale (variations from a fixed reference point)

Not to use when
- You are more interested in the quantities

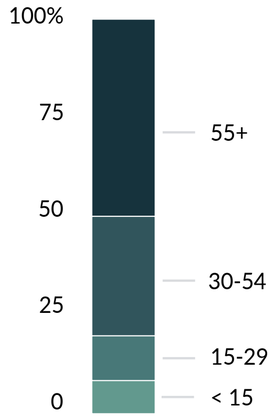
Other considerations
- Forgo the x- and y-axis in favor for direct labels
- Use padding between the colors for better contrast
- Use a key or direct label the categories

Color schemes

Categorical color palette

Sequential color palette

Diverging color palette

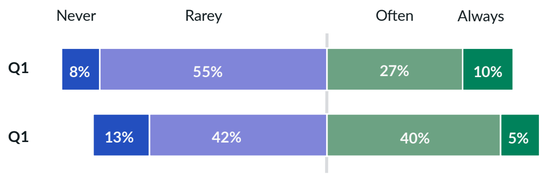


Stacked bar

Usage

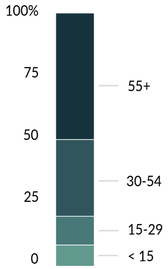
Survey answers by questions

In %



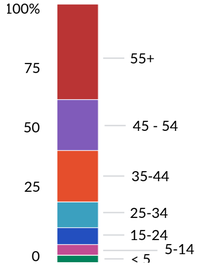
Northern Iowa age demographics

Share (%) of the population's age (yrs) in 2018



Northern Iowa Age Demographics

Share (%) of the population's age (yrs) in 2018



Polarizing phenomenons

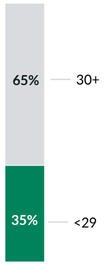
Variations from a reference (neutral) in a likert scale. For diverging bar charts, you can use the two extreme colors in the diverging color scale to represent positive and negative changes.

Direct label categories

As stacked bar chart inherently have multiple series, direct labeling increases readability of your chart.

Northern Iowa's youth are a minority

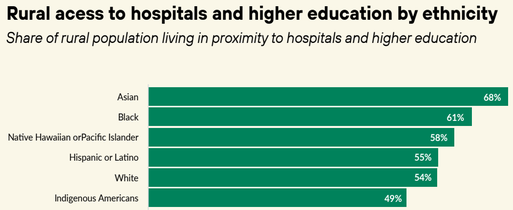
Share (%) of the population's age (yrs) in 2018



Don't use 5+ categories

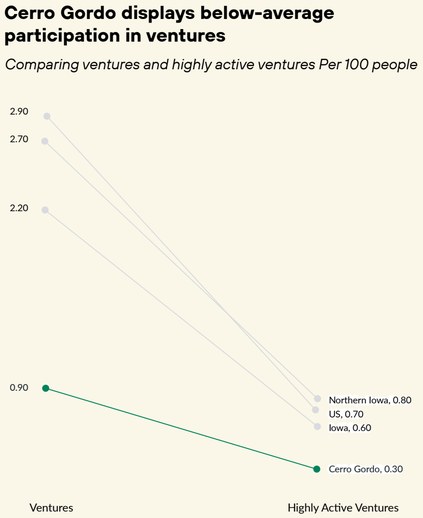
Use less than 5 series. If you must use more, consider grouping values and highlighting one.

Comparisons



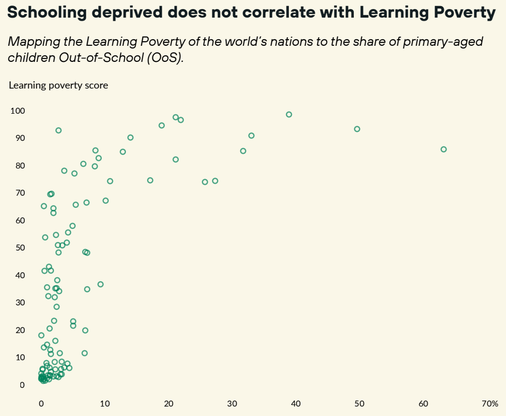
Bar chart

Compares categories or discontinuous timelines



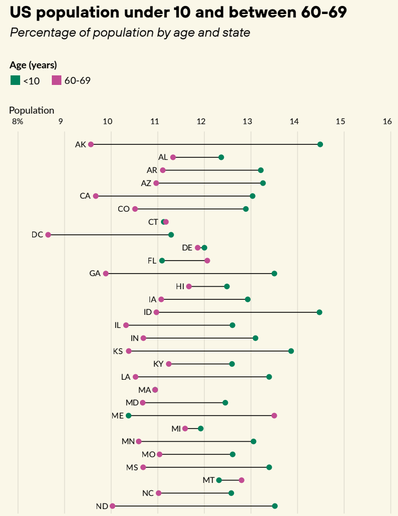
Slope chart

Displays the change between two categories



Scatter plot

Displays the relationship between two numeric values



Dot plots

Compares values or categories and can display changes within those categories

These charts help viewers compare values across categories. As categorical data is easily subjected to clutter (too many shapes, too many colors etc.), start with gray before choosing your colors. If one metric or data point is more important than the others, highlight this point and either use gray or a sequential palette to mute the rest, like in the slope chart example above.

Chart types

Bar chart

Bar charts are the easiest and probably most effective way to compare values with the same scale.

Best for
- Comparing values with the same scale

Not to use when
- You are showing shares

Other considerations
- The values axis should always start at 0
- Your bar chart shouldn't have more than two colors
- You can also use bar charts for negative/positive values as a diverging bar chart with the diverging color palette

Color schemes

Categorical color palette

Sequential color palette

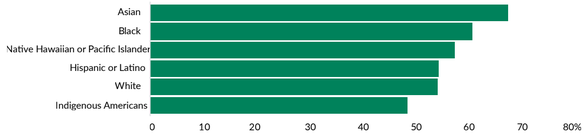
Diverging color palette

Bar chart

Usage

Rural acess to hospitals and higher education by ethnicity

Share of rural population living in proximity to hospitals and higher education



Indigenous americans have less access to hospitals and higher education than other races

Share of rural population living in proximity to hospitals and higher education

Rural acess to hospitals and higher education by ethnicity

Share of rural population living in proximity to hospitals and higher education

Rural acess to hospitals and higher education by ethnicity

Share of rural population living in proximity to hospitals and higher education

Highlight a category

Don't use more than two colors

Leverage direct labeling

Slope chart

Shows both a negative and positive evolution between two points in time or across two different categories. This works best when you highlight specific trend(s) and use other data point as reference points to provide context.

Best for

- Summarizing a complex evolution with simple trends (up, down, equal)
- Comparing two categories with multiple series

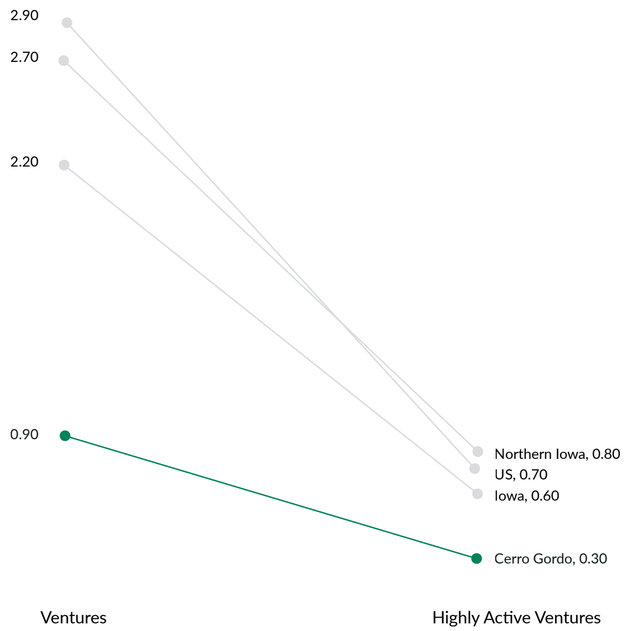
Not to use when

- You're losing too much information by oversimplifying the data

Color schemes

Highlighting colors

Sequential color palette



Slope chart

Usage

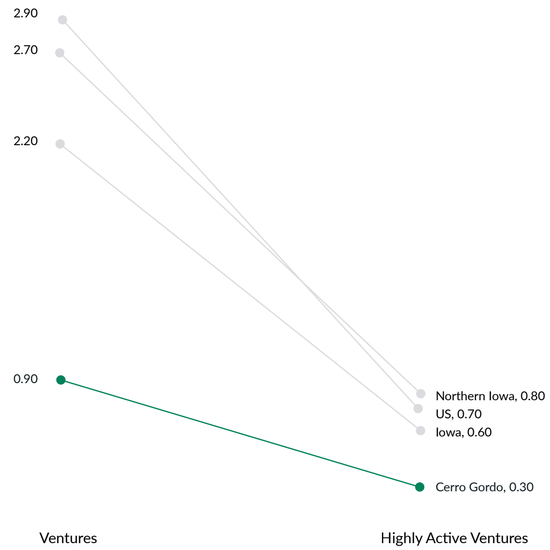
Use color to highlight and direct labels

Forgo categorical palettes and use color to highlight the data points of interest.

Especially effective for datasets with less than 5 series, you can eliminate the y axis and directly label the values on your lines.

Cerro Gordo displays below-average participation in ventures

Comparing ventures and highly active ventures Per 100 people



Ventures vs Highly Active Ventures

Per 100 people

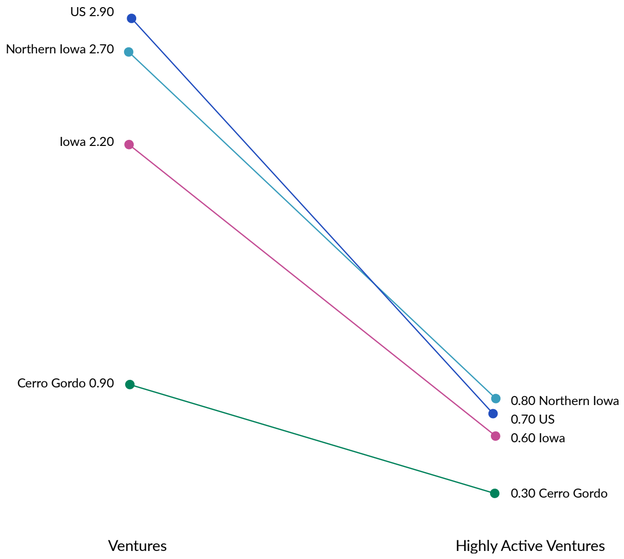


Chart types

Scatter plot

Best for
- Showing correlations (or lack of) between two values.
- Encoding a third value as the area of the circle (such as population)

Not to use when
- Your dataset is too large - it will look like a large blob (known as ‘over-plotting’)

Other considerations
- It is best to use a stroke over fill in order to expose overlapping values
- Over-plotting is common with scatter plots, so if you have a lot of points to plot, consider using transparency and/or decreasing the dot size
- The axes do not necessarily need to start at zero
- You can make correlations even clearer by adding trend lines to show the general direction of the relationship
- You can transform your scatter plot into a bubble plot by varying the sizes of the dots to a third variable. These should be sized by area and not by directly encoding the radius

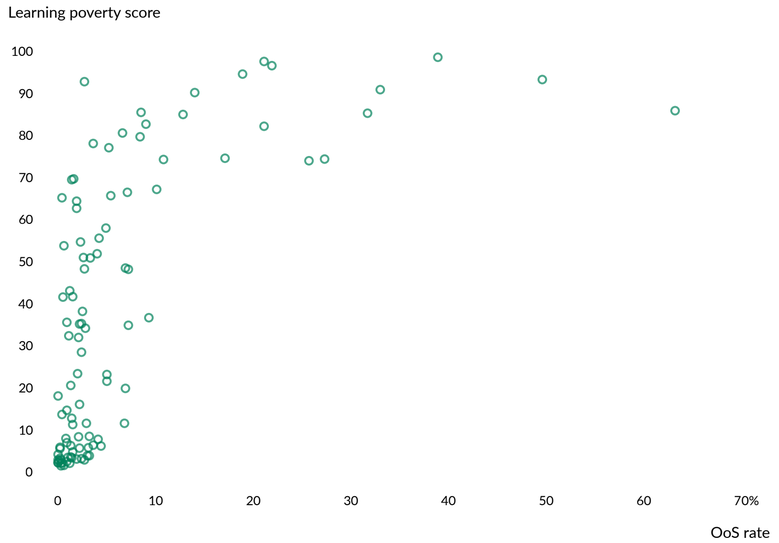
Color schemes

Categorical color palette

Sequential color palette

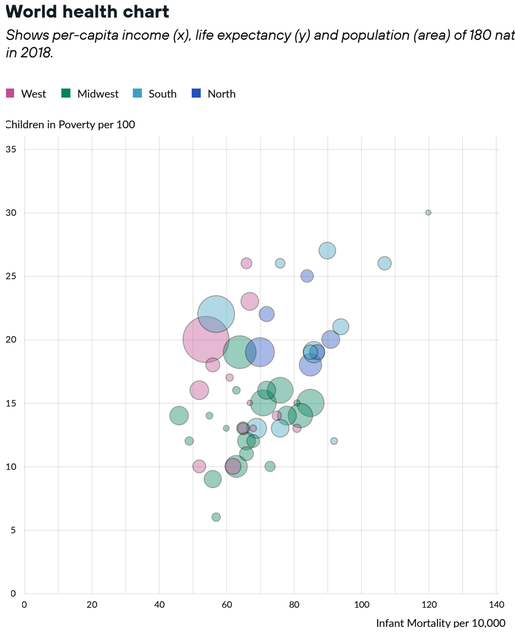
Schooling deprived does not correlate with Learning Poverty

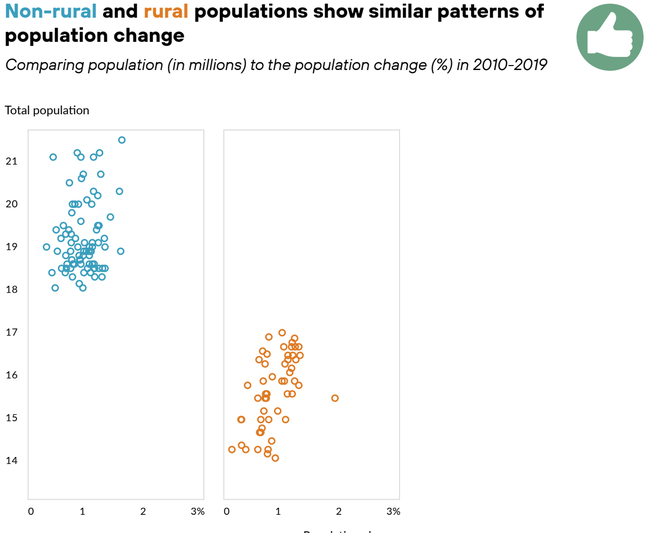
Mapping the Learning Poverty of the world’s nations to the share of primary-aged children Out-of-School (OoS).

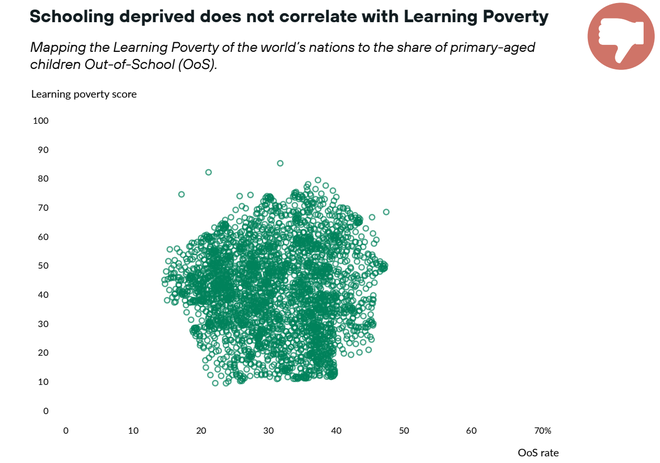


Scatter plot

Usage







Encoding a third value

You can turn your scatter plot into a bubble chart by adding third value that is mapped to the circle's area.

Faceting

Also called small multiples, faceting can be an effective way to reduce clutter without compromising the number of categories.

Over-plotting

If you have too many points and your result is a blob despite a low radius and transparency, consider using another chart, such as a heatmap.

Dot plots

There are a few types of dot plots. We will be treating the lollipop chart, which functions identically to a bar chart, and the dumbbell plot, which is an alternative to grouped bar charts and slope charts.

Best for

- Visualizing a lot of categories with a lot of high values
- Comparing changes within a lot of categories
- Reducing the visualization's ink ratio

Not to use for

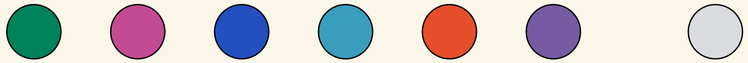
- Visualizing a few categories (use a bar chart)

Other considerations

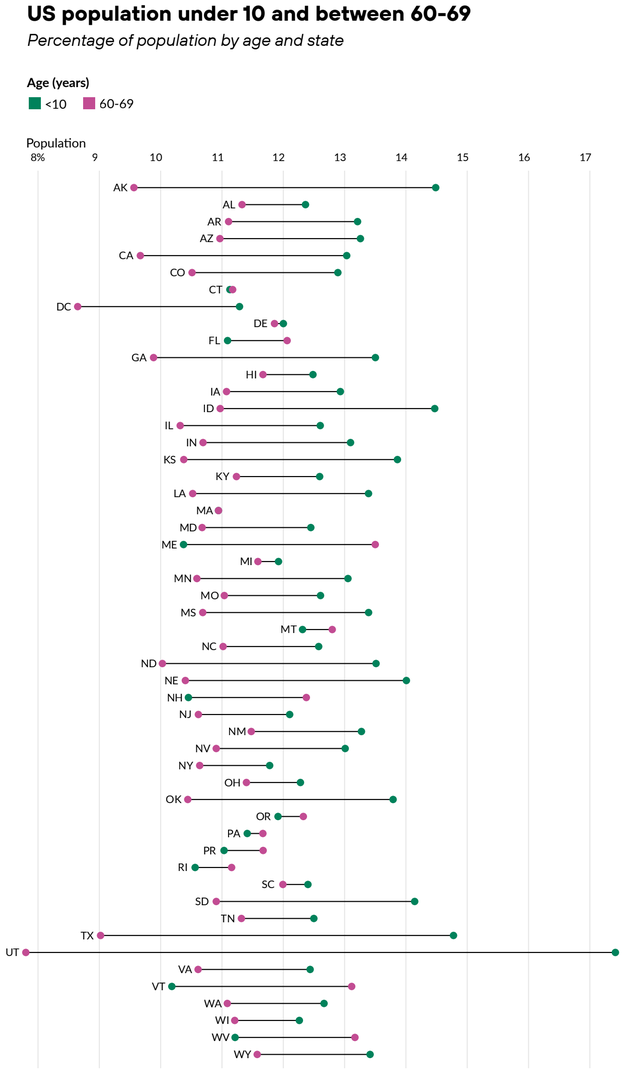
- While the center of the circle marks the value, this can be difficult to judge and is more imprecise than using a bar chart. Direct labeling can help mitigate this fault.

Color schemes

Categorical color palette

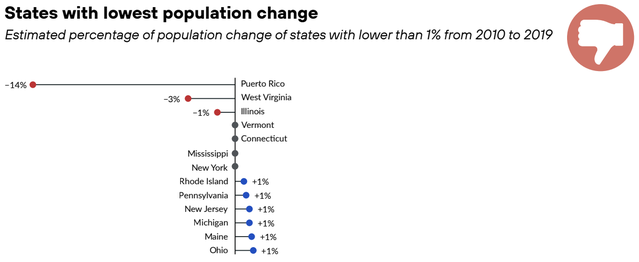


Diverging color palette



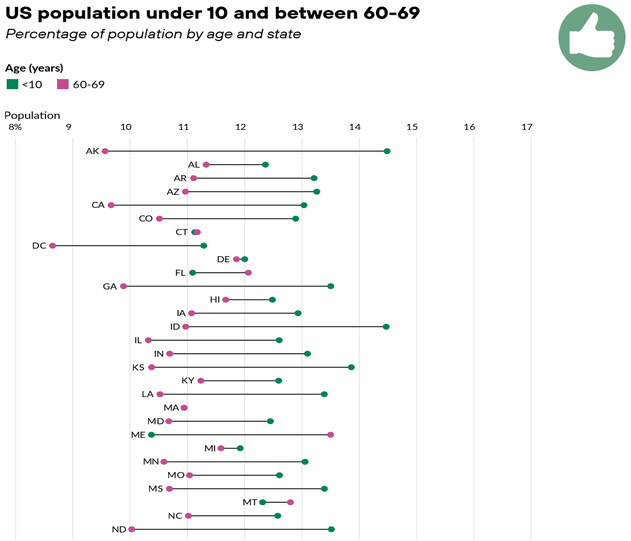
Dot plots

Usage



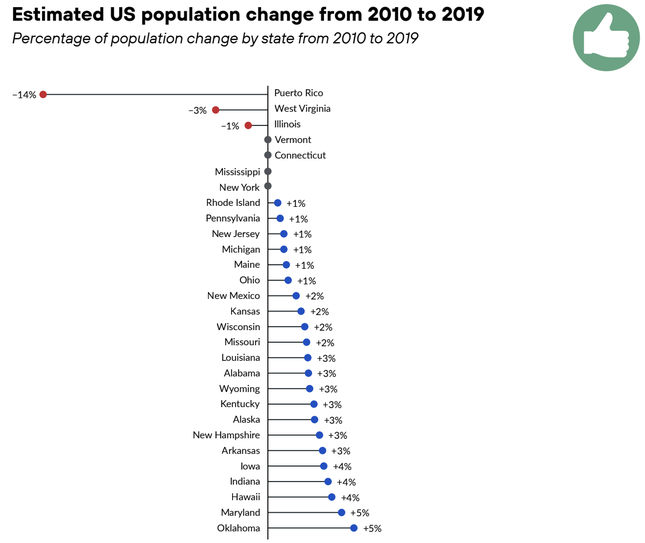
Use bar charts when visualizing few series

Lollipop charts are useful for reducing the data-to-ink ratio when visualizing many categories. A more conventional bar chart or slope chart could be preferred when visualizing data with only a few series.



Compare two values within groups

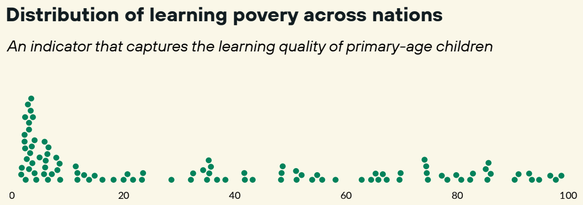
If you find yourself wanting to compare a few values within each category, you might want to break your chart up into several and consider the question(s) you are trying to answer.



Sort values

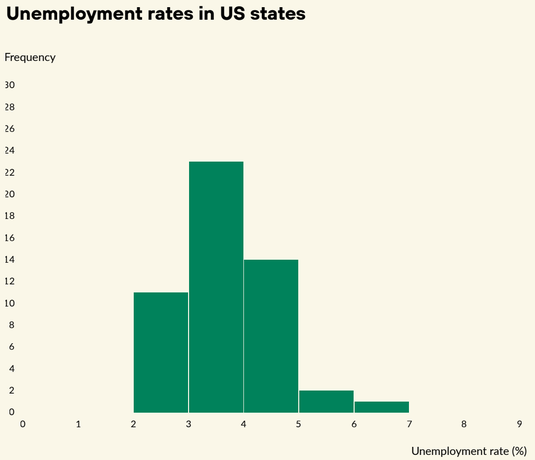
Values should be sorted logically. This can be by ABC order, such as in the dumbbell plot, or by the values' themselves.

Distribution



Beeswarm

Reveals the distribution of a variable and individual data points



Histogram

Reveals the distribution of a variable by grouping data points into bins

Chart types

Beeswarm

Reveals the distribution of a variable and the individual data points.

Best for
- Small datasets where you want to display each data point
- Revealing outliers

Not to use when
- You are working with a large dataset
- You want to simplify the distribution and trend (use a histogram)

Other considerations
- Consider creating small multiples if you want to compare distributions of different categories

Color schemes

Categorical color palette

Distribution of learning povery across nations

An indicator that captures the learning quality of primary-age children

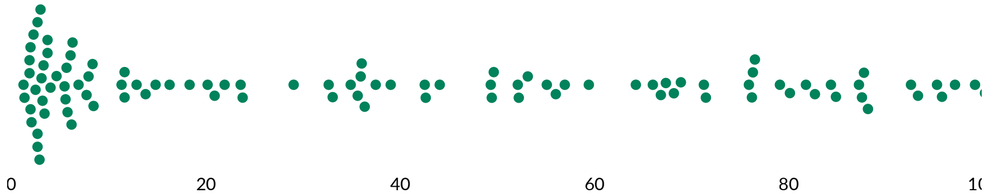


Chart types

Histogram

Reveals the distribution of a variable by grouping data points into bins.

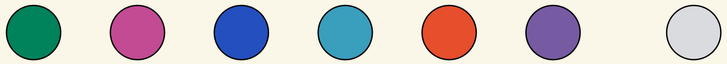
Best for
- Visualizing distributions of quantitative or numerical data

Not to use when
- You want to display categorical variables (use a bar chart)

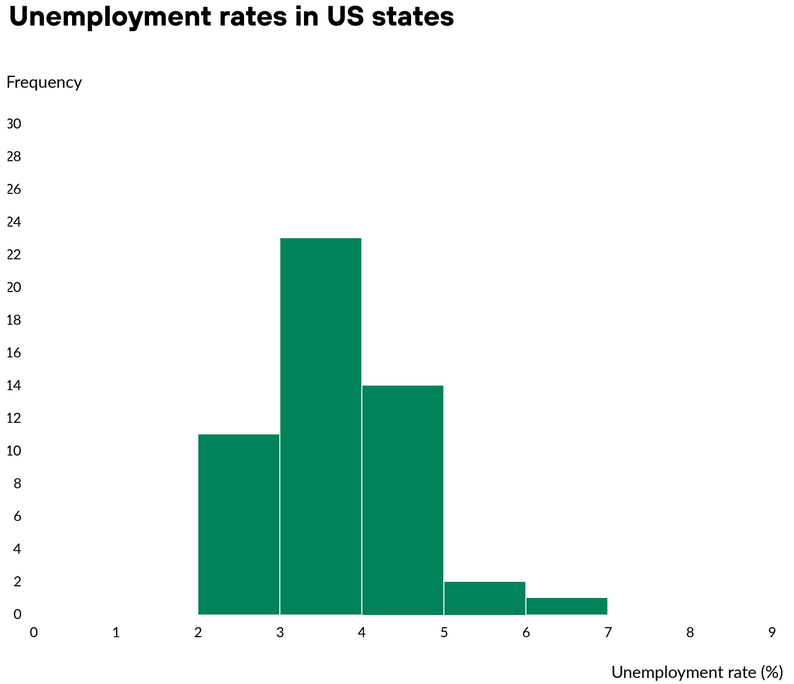
Other considerations
- If you're looking to create a histogram with accuracy, consider sizing your bars to proportionally match the interval, or better yet, keep things simple by using equally sized bins
- If you have outliers, consider putting them in a catch-all value

Color schemes

Categorical color palette



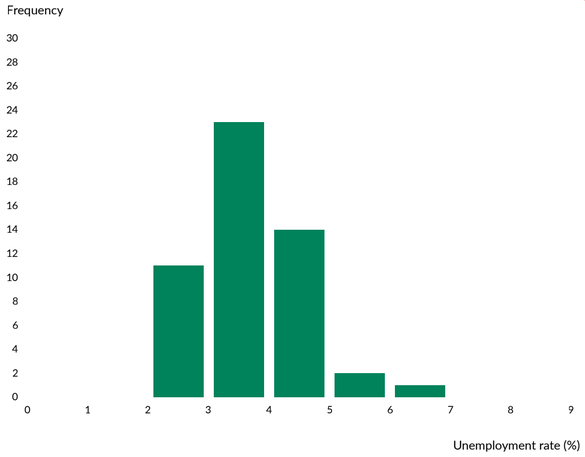
Diverging color palette



Histogram

Usage

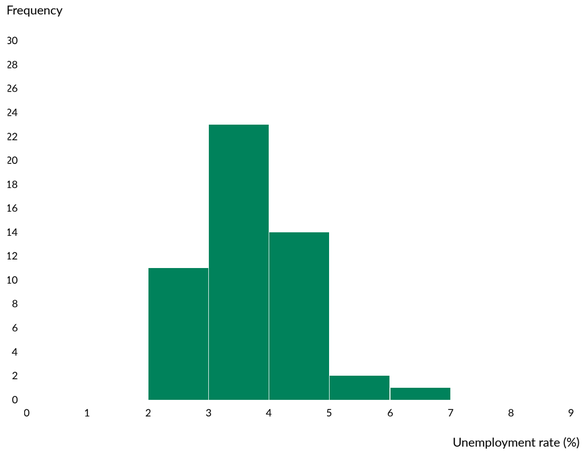
Unemployment rates in US states



Minimize spacing

In a histogram, bars should touch, while in a bar chart, the bars have space.

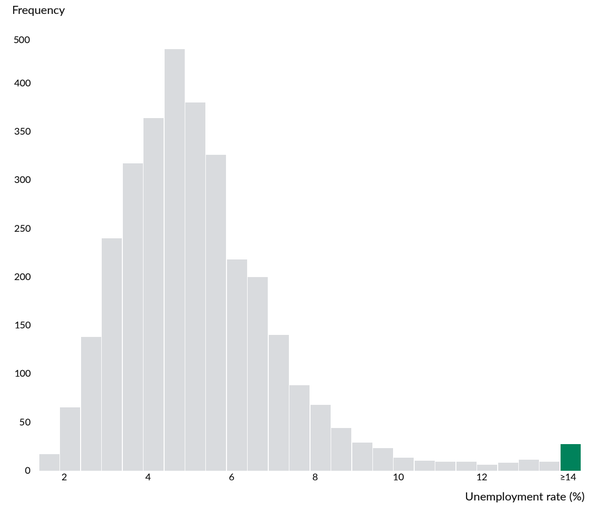
Unemployment rates in US states



Keep a continuous scale

Unlike bar charts, histograms must have a continuous scale. As the data should be proportional to the area of the bar (both length and width,) a change in bin sizes should adjust the length and width dimensions.

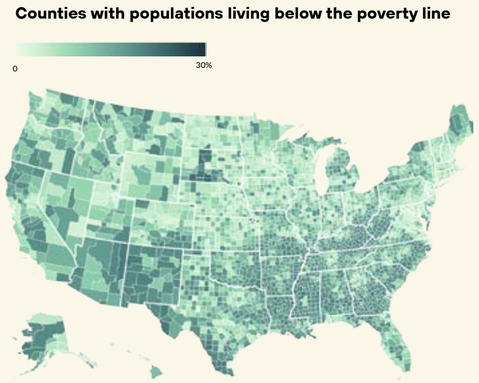
Unemployment rates in US counties in 2016



Catch-all value

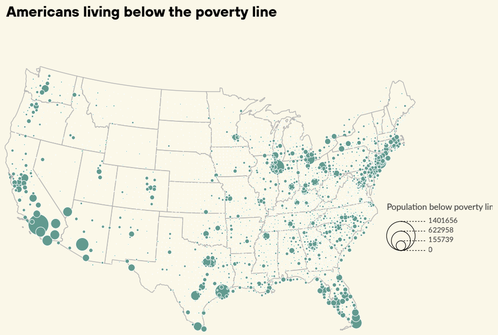
If you have outliers that you deem less important for your story, consider grouping them together under one value.

Spatial



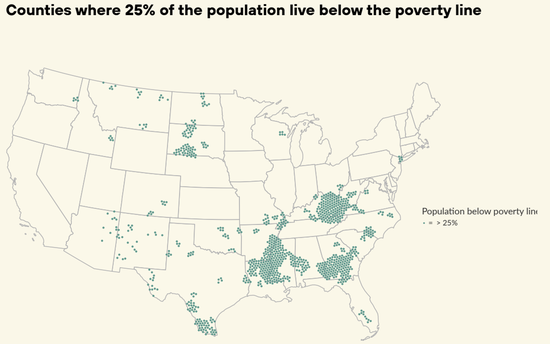
Choropleth

Shows regional patterns



Bubble map

Focuses on specific locations and highlights their values



Density plot

Reveals distributions across regions

Choropleth

Reveals the intensity of a phenomenon according to its region.

Best For

- Relative data: rates, indices, entities
- Highlighting regions

Not to use for

- Absolute numbers (i.e. number of people who are affected) as we can't compare that data without knowing the counties' population. Consider using a bubble map

Other considerations

Extra care needs to be given to determine the amount of stops to use, or the method of boxing, should you choose to use a discrete scale. If you want readers to immediately notice the range of values a region falls in, discrete steps are a good choice. However, discrete steps sacrifice nuance for that quick readability. A continuous color scale lets readers compare neighboring regions, even if they would fall in one color shade on a map with discrete steps.

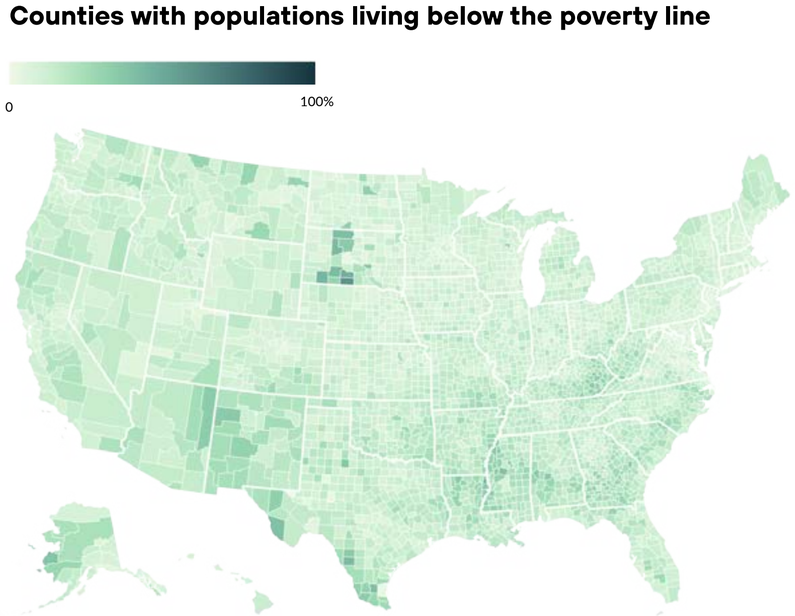
Legends

A key is mandatory to explain the color encoding system and data bucketing, if applicable. It's unnecessary to add the region/country labels if your audience knows the geography of the location in question.

Color Schemes

Sequential color palette

Diverging color palette

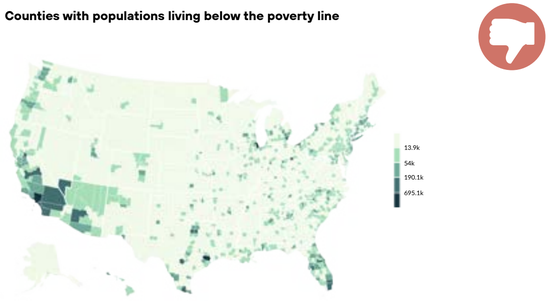


Choropleth

Usage

Defining stops

The number of steps you use and the method of classifying the data can drastically influence the message you are trying to tell. Take your time when choosing the number of “stops.” They determine how dramatic the contrast between your low and high values appear. You want to find the “just right” spot where the patterns are visible, yet the differences aren’t exaggerated. If you find yourself using a lot of stops, you might want to try a density dot plot or a bubble map instead. If your map has a lot of low values resulting in a light chart (such as the chart to the right), you may want to reevaluate how you encoded the data (absolute vs relative numbers) and the number of stops used.

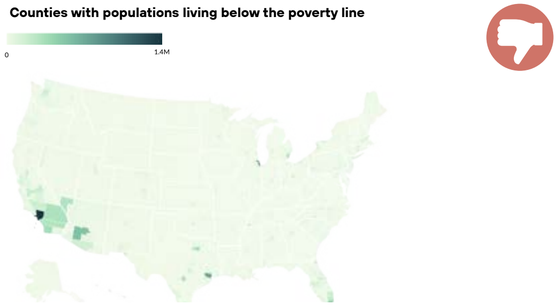


As absolute numbers are used, a choropleth is not the right chart.

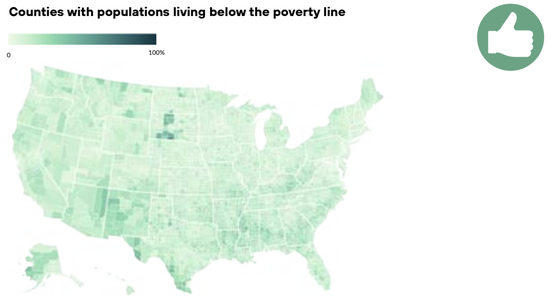
As this chart focuses on statistical bracketing, a choropleth is a fine choice.

Continuous scales

Consider using continuous scales. Discrete steps bucket your data into e.g. 5 steps which sacrifices nuances. A continuous scale helps readers compare neighboring regions.



Foregos steps for a **continuous** scale. As absolute numbers are used, a choropleth is not the right chart.



If the designer chooses to focus on the percentage within a region, a choropleth can be an appropriate choice.

Choropleth

Grid

Choropleths give unequal weight to regions with a larger surface area. In this example with US states, grids ‘equalize’ this bias by making each state the same size shape while keeping the states roughly aligned geographically.

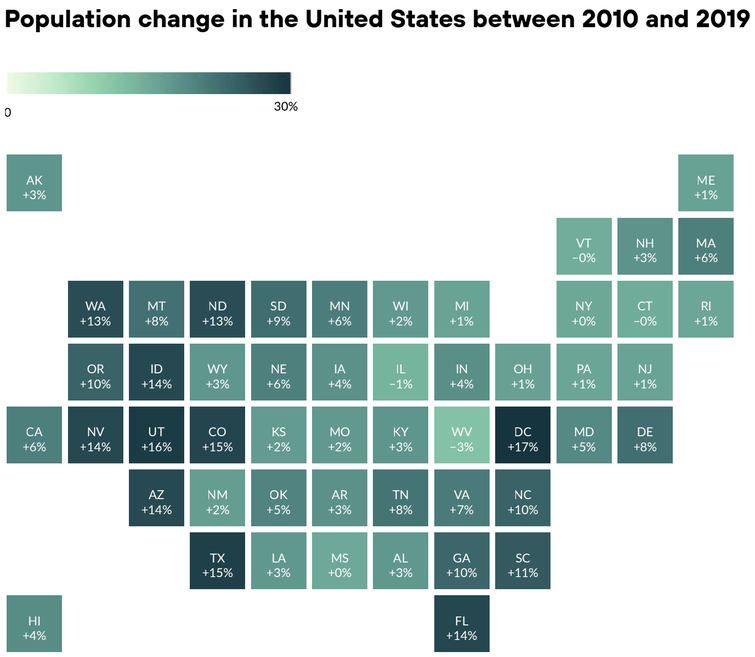


Chart types

Bubble map

Highlights the intensity of a phenomenon according to its location (point).

Best for
- Referring to specific points and expressing their magnitude

Not to use for
- Highlighting regions

Other considerations
- You can use several symbols (diamond, star, etc.) to add a layer of information such as expressing different categories
- If you are only highlighting a few points, you can directly label the regions next to their symbol

Color schemes

Single color

Map fill

Categorical color palette

Sequential color palette

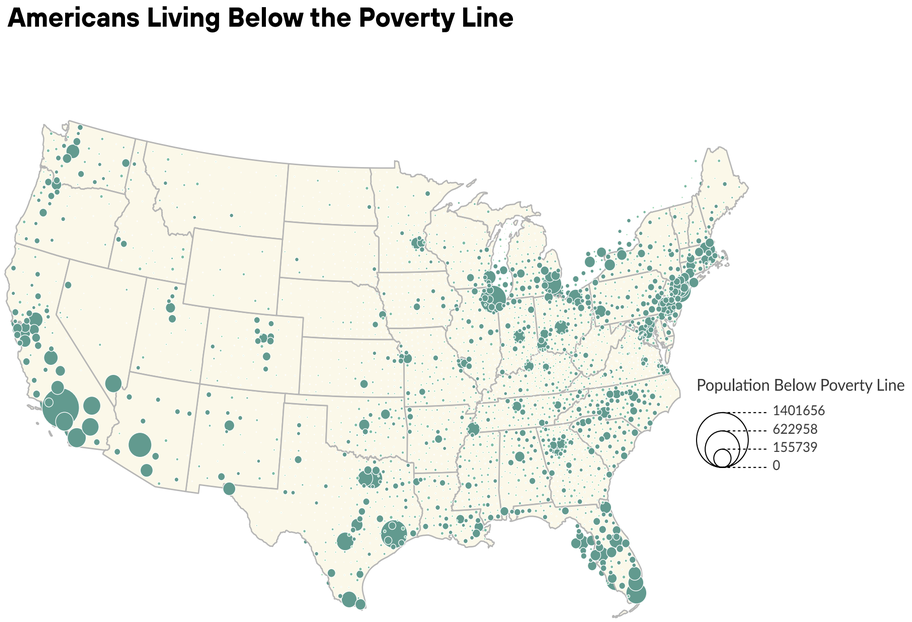


Chart types

Density plot

Reveals distributions across regions

Best for
- Overall concentrations across regions
- Simple counts (e.g. one dot = 10,000 people)

Not to use for
- Expressing quantities or shares in a particular region

Other considerations
- Legends are especially important so that the audience knows what value is being encoding by each dot

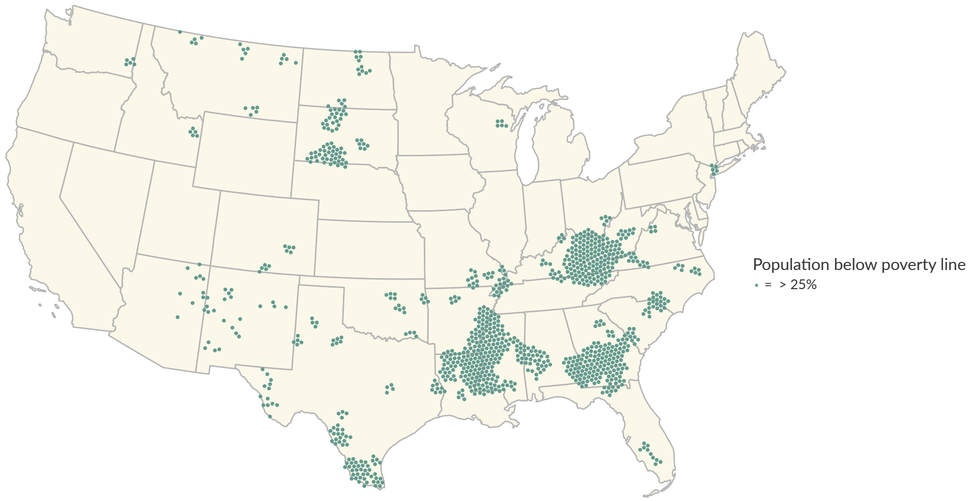
Color schemes

Single color

Map fill

Categorical color palette

Sequential color palette



Tables

Tables

Tables are a very important data visualization that are susceptible to clutter. Effectively designed, they have help your audience uncover specific patterns or trends.

Best for
- Knowing the exact number of every value in your data

Not to use for
- Showing a lot of data or data in a small space

Other considerations
- Remove unit repetition
- Highlight outliers
- Add visualizations, if appropriate
- Select appropriate level of precision

Color schemes

Categorical color palette

Sequential color palette

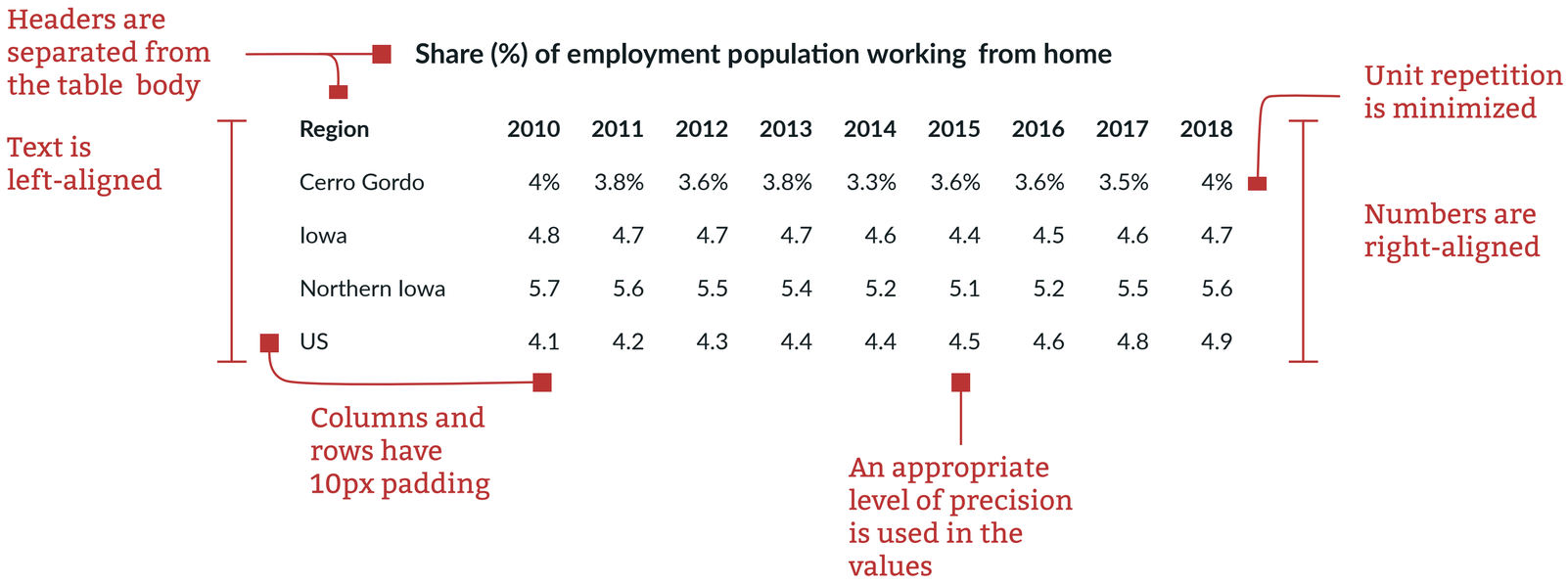
Diverging color palette

Share (%) of employment population working from home

Region	2010	2011	2012	2013	2014	2015	2016	2017	2018
Cerro Gordo	4%	3.8%	3.6%	3.8%	3.3%	3.6%	3.6%	3.5%	4%
Iowa	4.8	4.7	4.7	4.7	4.6	4.4	4.5	4.6	4.7
Northern Iowa	5.7	5.6	5.5	5.4	5.2	5.1	5.2	5.5	5.6
US	4.1	4.2	4.3	4.4	4.4	4.5	4.6	4.8	4.9

Tables

Layout



Region	2010	2011	2012	2013	2014	2015	2016	2017	2018
Cerro Gordo	4%	3.8%	3.6%	3.8%	3.3%	3.6%	3.6%	3.5%	4%
Iowa	4.8	4.7	4.7	4.7	4.6	4.4	4.5	4.6	4.7
Northern Iowa	5.7	5.6	5.5	5.4	5.2	5.1	5.2	5.5	5.6
US	4.1	4.2	4.3	4.4	4.4	4.5	4.6	4.8	4.9

Tables

Add visuals

Share (%) of employment population working from home

Region	2010	2011	2012	2013	2014	2015	2016	2017	2018	2010-2018
Cerro Gordo	4%	3.8%	3.6%	3.8%	3.3%	3.6%	3.6%	3.5%	4%	
Iowa	4.8	4.7	4.7	4.7	4.6	4.4	4.5	4.6	4.7	
Northern Iowa	5.7	5.6	5.5	5.4	5.2	5.1	5.2	5.5	5.6	
US	4.1	4.2	4.3	4.4	4.4	4.5	4.6	4.8	4.9	

Sparklines

Sparklines are a useful visualization to add to your spreadsheets if you want to be able to quickly scan over trends.

Share (%) of homes with and without Internet in rural counties

County	With internet	Without Internet
Fayette County, Illinois	74.4%	25.6%
Massac County, Illinois	70.6	29.4
Cass County, Illinois	80.2	19.8
Sullivan County, Indiana	73.1	26.9
Fulton County, Indiana	81.0	19.0
Union County, Indiana	75.3	24.7
Perry County, Indiana	78.1	21.9
Fountain County, Indiana	81.2	18.8

Heatmap

You can transform our tables into heatmaps, which help readers focus on an overall pattern.

Change in poverty from 2019 to 2021

Region	Number	Percent Change	Margin of Error
Alabama	55,218	0.7%	0.8%
Alaska	3,227	0.4	1.5
Arizona	-49,481	-0.7	1.6
Arkansas	5,414	0.1	0.9
California	180,199	0.5	0.2

Highlighting numbers

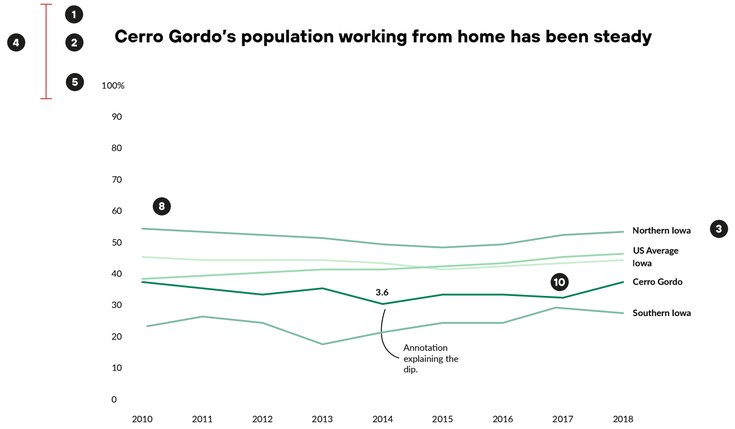
You can leverage colors in order to direct the audience to specific numbers, such as outliers, negative numbers, or other numbers of interest.

Checklist

Strive for 80%. Not every dataviz will need to checkoff every box.

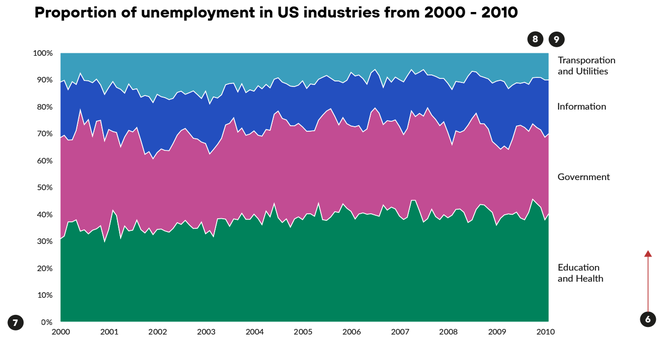
Text

- [ ] 5-10 word description is left-aligned in the upper left corner **(1)**
- [ ] Subtitles and/or annotations provide additional, contextual information **(2)**
- [ ] All text is horizontal for the ease of reading **(3)**
- [ ] Text sizing/styling is hierarchal
  titles → subtitles/annotations → labels → axis labels **(4)**
- [ ] Axis labels are used sparingly **(5)**
- [ ] Acronyms are spelled out, or in an easy-to-find glossary
  Example: *HAV (Highly Active Ventures)*



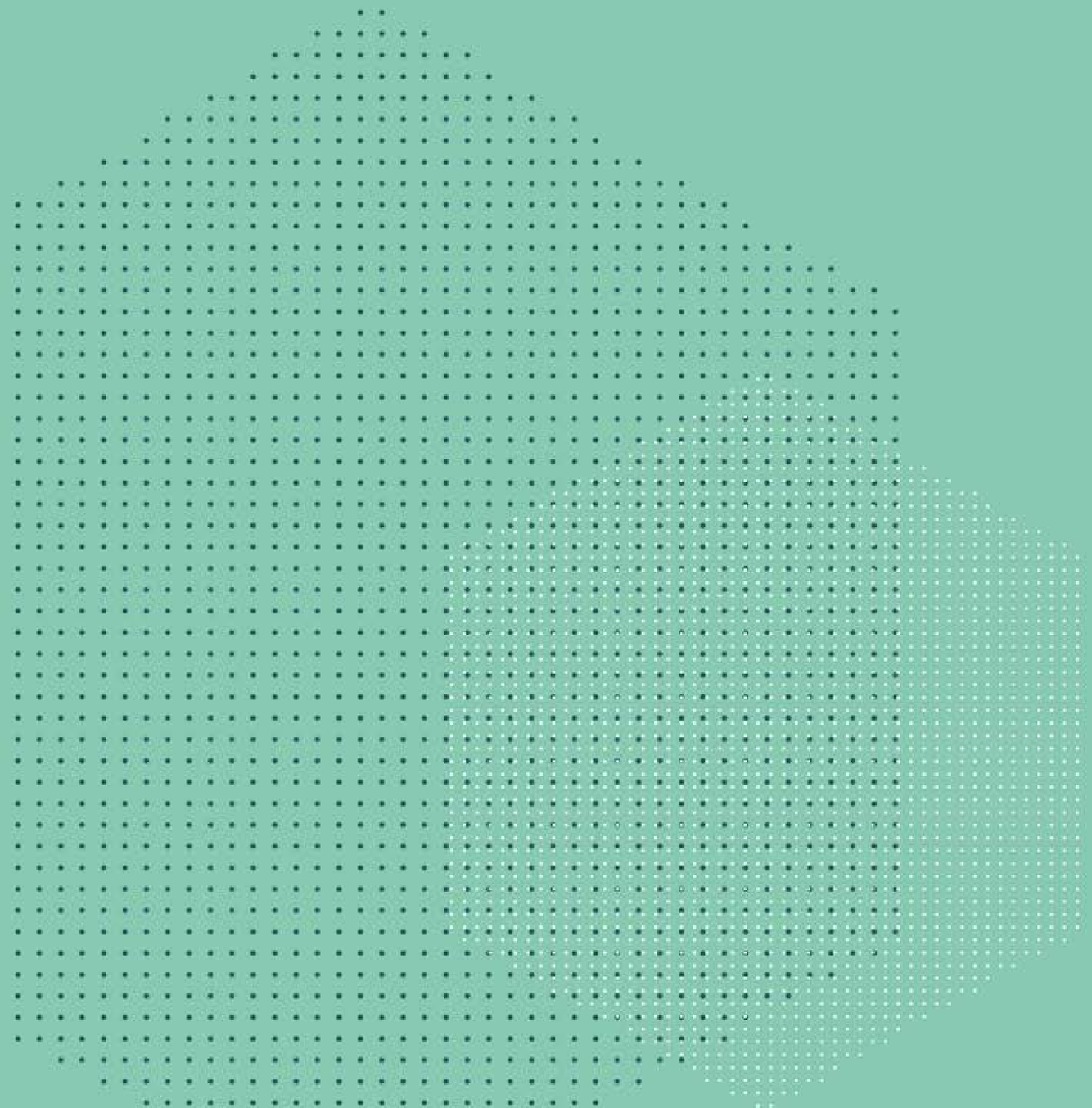
Layout

- [ ] Data is sorted logically (e.g. descending order for categorical charts) **(6)**
- [ ] Proportions are accurate (bar charts start at 0, all other charts have a min and max scale) **(7)**



Colors

- [ ] CORI's color palettes are used **(8)**
- [ ] Choice of color scheme fits use case: categorical, sequential or diverging **(9)**
- [ ] Color is used sparingly and consistently
- [ ] Text sufficiently contrasts background: light on dark / dark on light
- [ ] Colors are used to highlight key data and patterns and gray is used to deemphasize less important data **(10)**



CREATED BY MAXENE GRAZE FOR RURAL INNOVATION