

# Data Visualization: Visualization with Purpose - Accessible Data Visualization

```
$ echo "Data Science Institute"
```

# Agenda for today

- Go through slide deck #7: Accessible data visualization
- Go through slide deck #6: Subplots and combining visualizations
- Discuss assignment 3

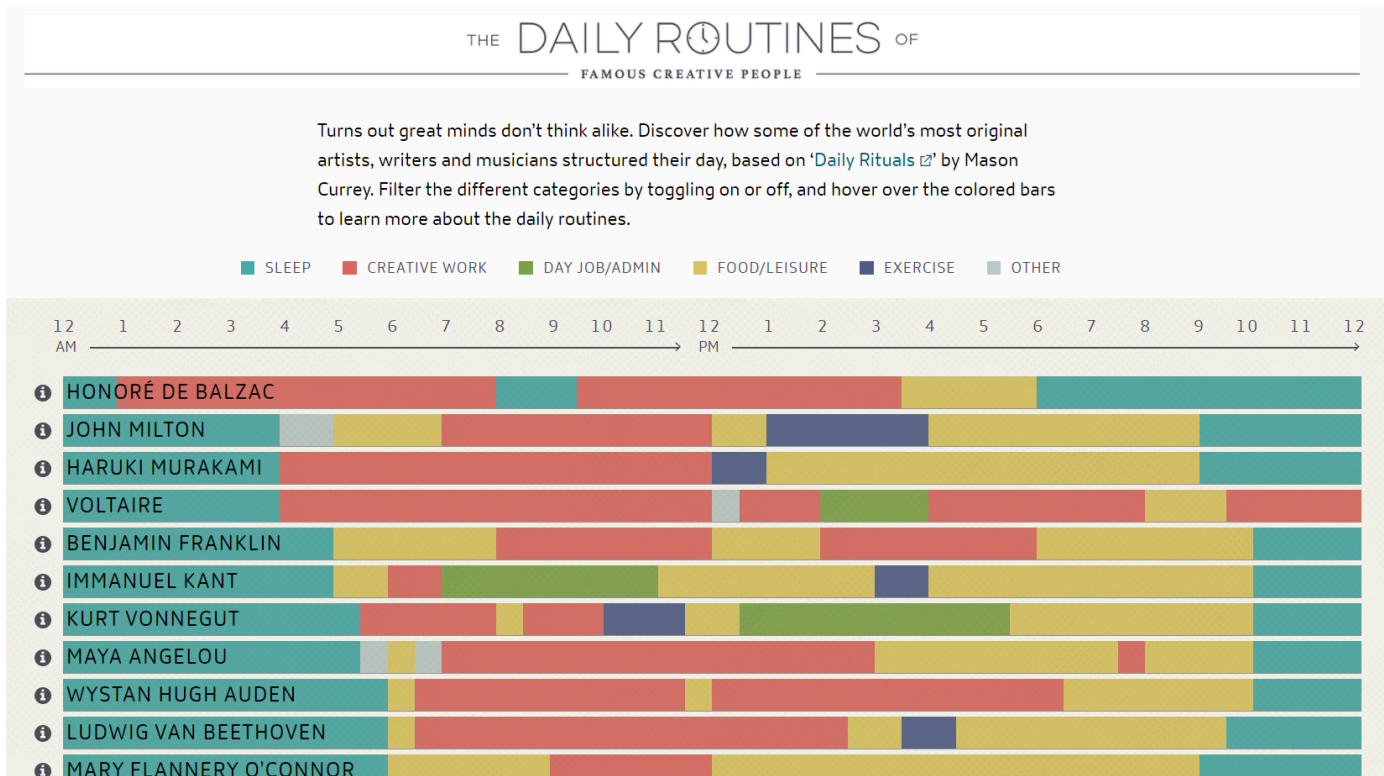
# Review feedback from last session

# Overview of this slide deck, we will...

- Discuss the real-world impacts of inaccessible data visualization
- Explore practical tips and resources that we can use to improve the accessibility of our data visualizations, especially relating to:
  - Colour
  - Text
  - Image Descriptions
  - Getting our visualizations to the intended audiences

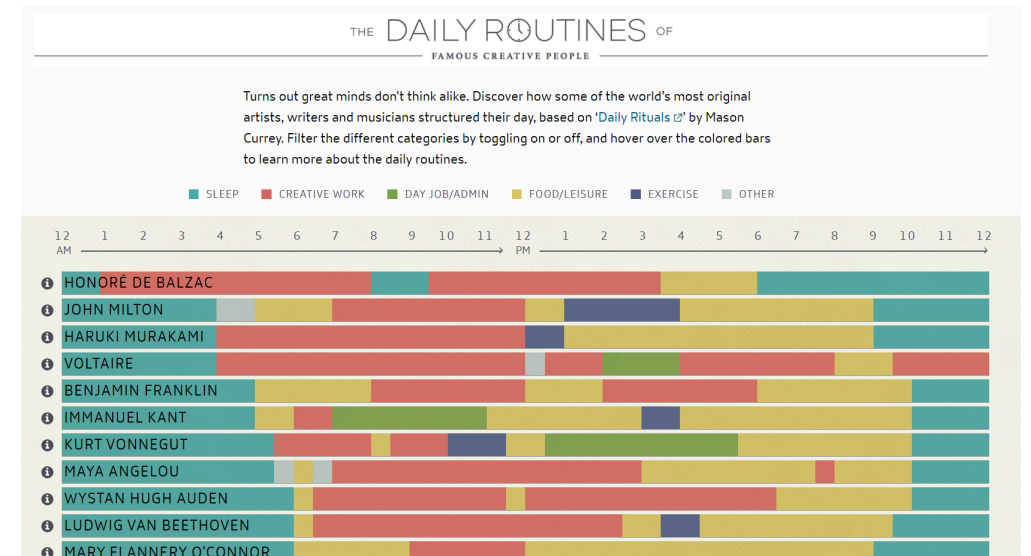
## Case Study: Who gets to use data visualizations?

# 'The Daily Routines of Famous Creative People' Visualization



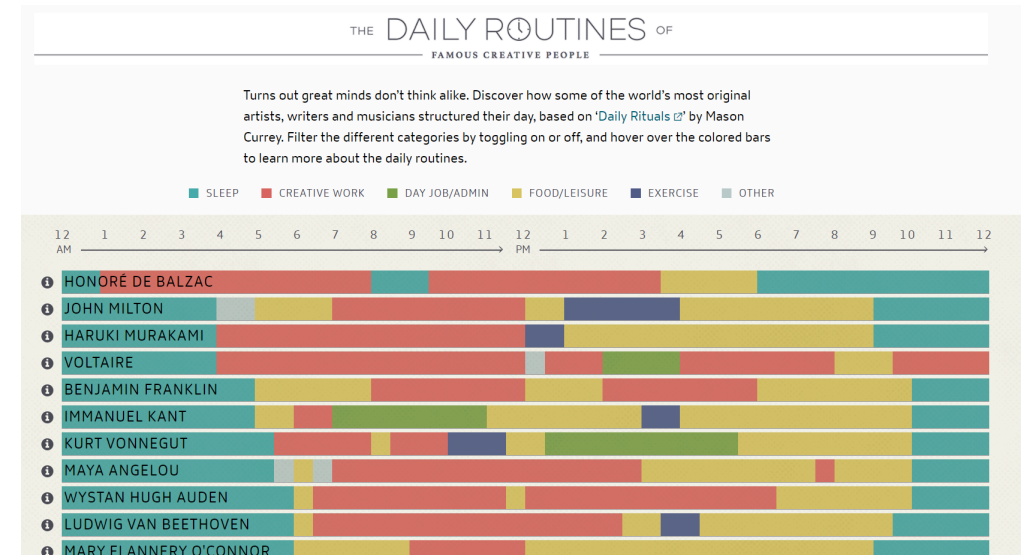
# 'The Daily Routines of Famous Creative People' Visualization

- The 'Daily Routines' plot is often used as an example of 'good' data visualization
- It includes its data source (substantive qualities), uses an appropriate plot type for comparison (perceptual qualities), and limits its colour scheme (aesthetic qualities)



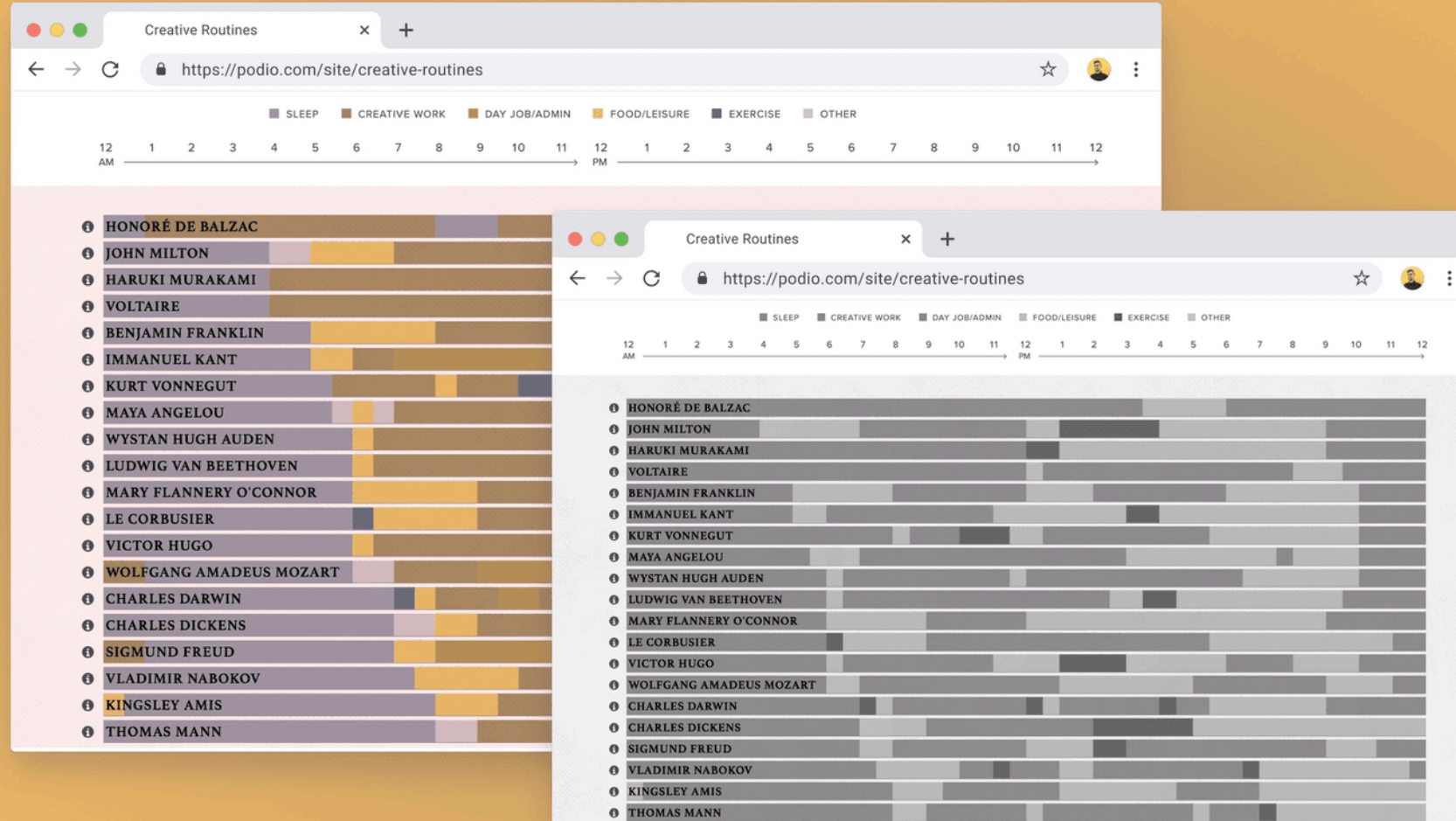
# Questioning assumptions

- Sarah L. Fossheim, a developer and designer, **was curious** about how the visualization would be perceived by people with deuteranopia (a type of red-green colourblindness) and achromatopsia (partial or total lack of colour vision)



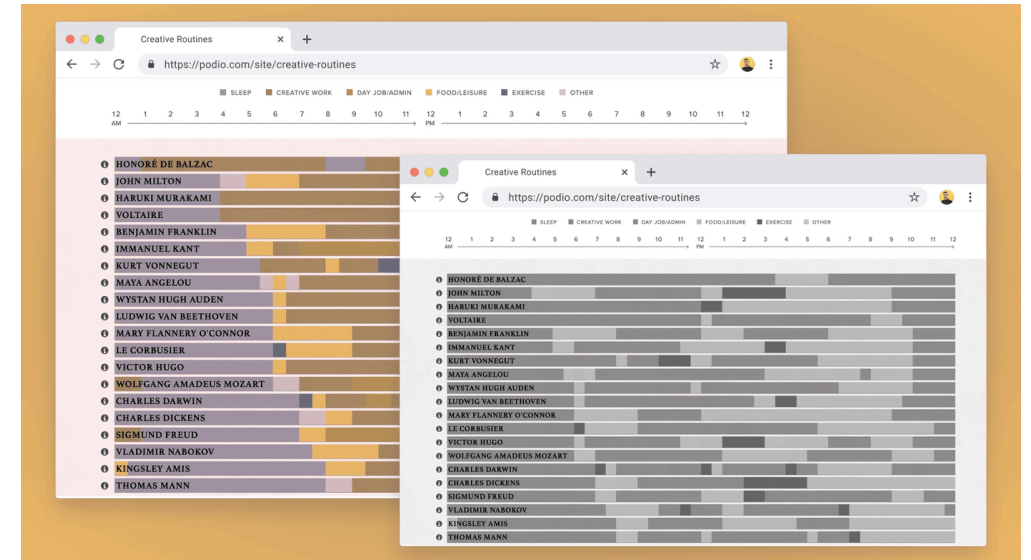


# A different picture



# A different picture

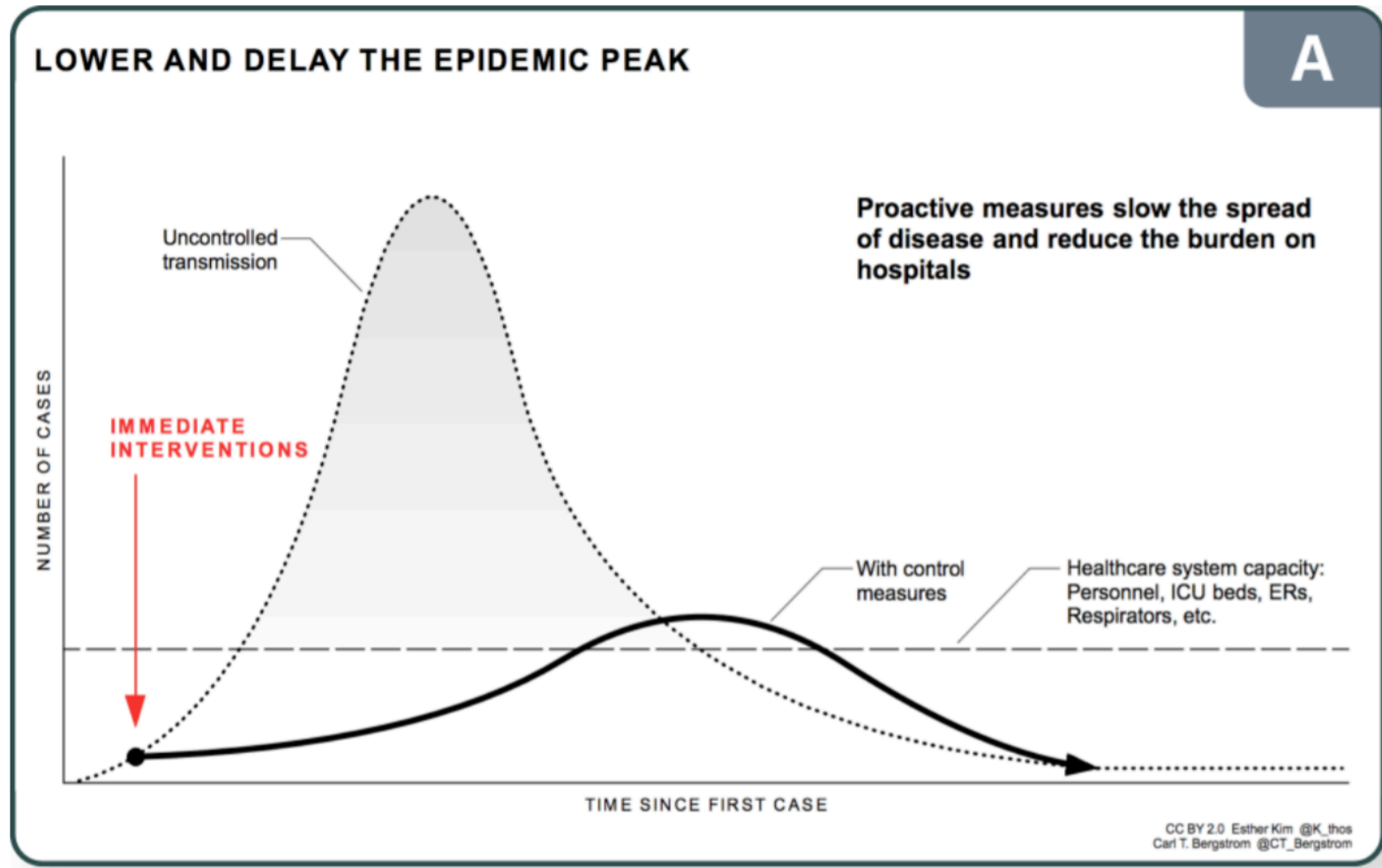
- Some of the colours end up looking similar or the same, and the lack of specific labels on the plot makes it impossible for individuals with either form of colourblindness to make sense of the data the creator was trying to present
- This data visualization is **not accessible** to certain people



- In the case of the 'Daily Routines of Famous Creative People' visualization, a lack of accessibility is an inconvenience, but one with a relatively limited scope
- **But** what about when the ability to make sense of the data being shared visually has an impact on people's daily lives and wellbeing?

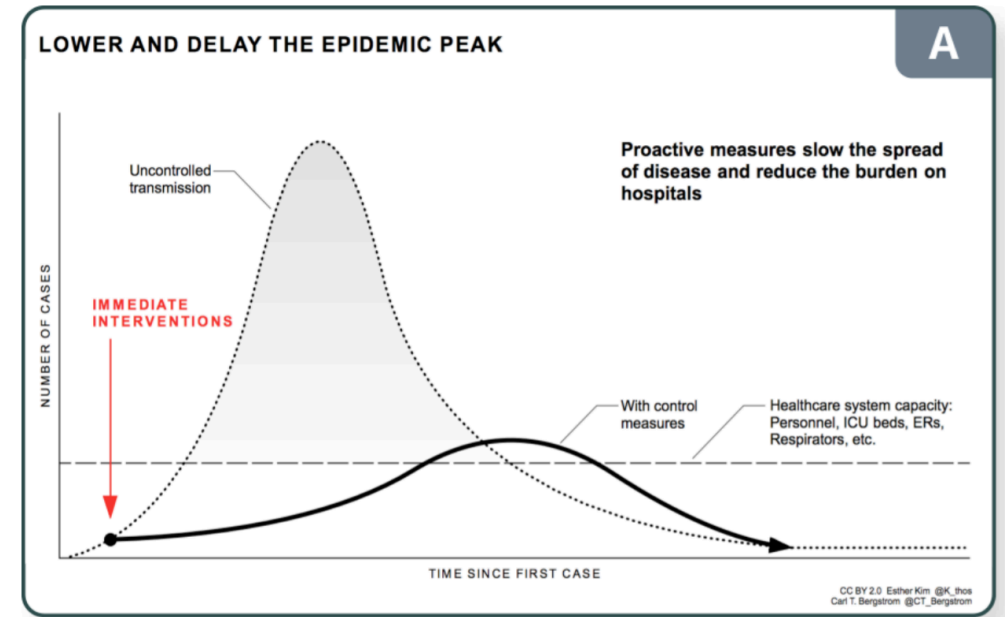


# 'Flattening the Curve'



# 'Flattening the Curve'

- Data visualizations such as this one are an efficient and effective form of communicating public health information, but not to everyone
- **Activity:** In what ways is this data visualization inaccessible, and to whom? Let's discuss!



# Pandemic information and people with visual impairments

- Graphical and spatial visualizations of pandemic data are everywhere, and help the public see epidemiological trends at a glance, but most of these images are inaccessible
  - Lack of alternative descriptive text means that people using screen readers **cannot access the data**
  - Tactile or braille versions of COVID-19 data visualizations are **rare and often expensive** to create
- From Lundgard & Satyanarayan (2022), “for readers with visual disabilities... inaccessible visualizations are, at best, demeaning and, at worst, damaging to health.”

- Failing to consider best practices for accessible design means that our data visualization will not have the desired positive effect on all of our audience, and may even have a harmful effect by excluding them from important data
- As professionals and academics who went to communicate our data visually, we have a responsibility to prioritize accessibility



# Government of Canada Principles for Accessible Design

- The Government of Canada provides four broad goals for what we should try to accomplish with accessible design:
  - Make information clear, concise and easy to use
  - Present information in predictable, barrier-free ways
  - Provide barrier-free ways for people to interact with you
  - Make sure technology-based products work with assistive technologies and devices



So how do we incorporate these accessibility goals into our data visualization practices?

# Colour

# Using colour

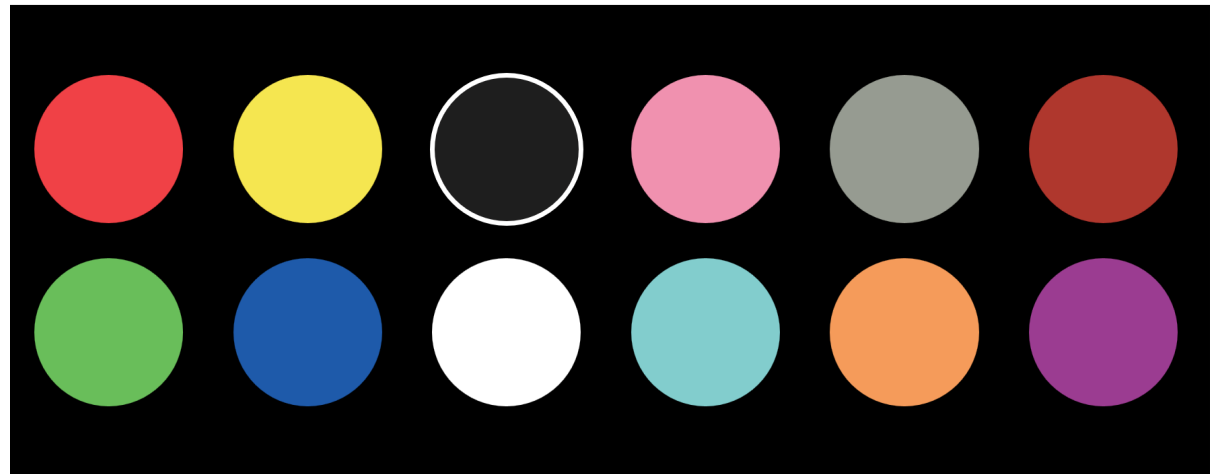
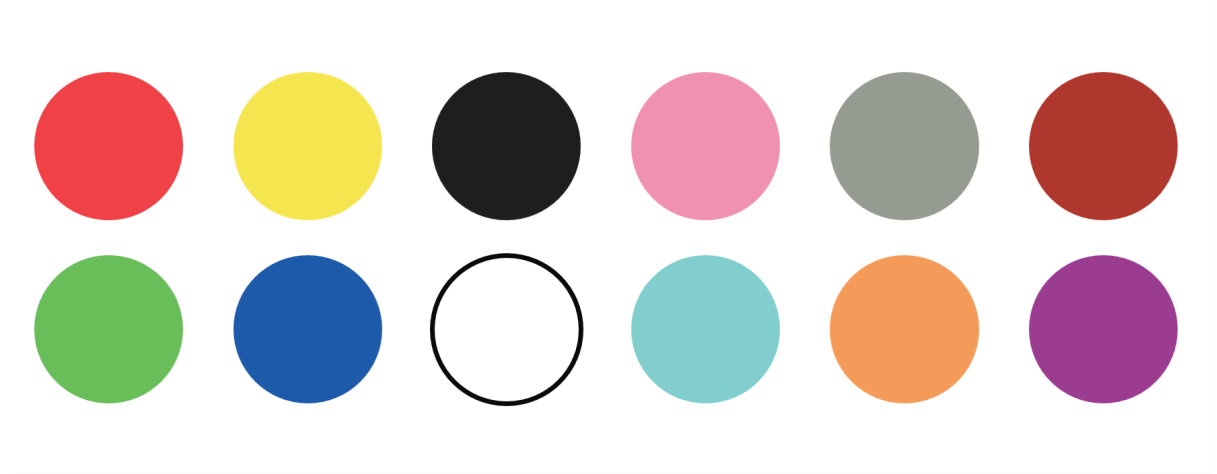
- Colour is one of the most useful features of many data visualizations
- We have to be aware of the ways that colour can be perceived differently by the audiences of our images

# Using colour

- Think of colour in terms of 3 channels
  - Magnitude of these two channels is suitable for ordered data
    - Luminance: how bright
    - Saturation: how colourful
  - Hue is an identity channel, suitable for categorical data
    - Hue: pure colours



- |           |            |
|-----------|------------|
| 1. Red    | 7. Pink    |
| 2. Green  | 8. Cyan    |
| 3. Yellow | 9. Gray    |
| 4. Blue   | 10. Orange |
| 5. Black  | 11. Brown  |
| 6. White  | 12. Purple |



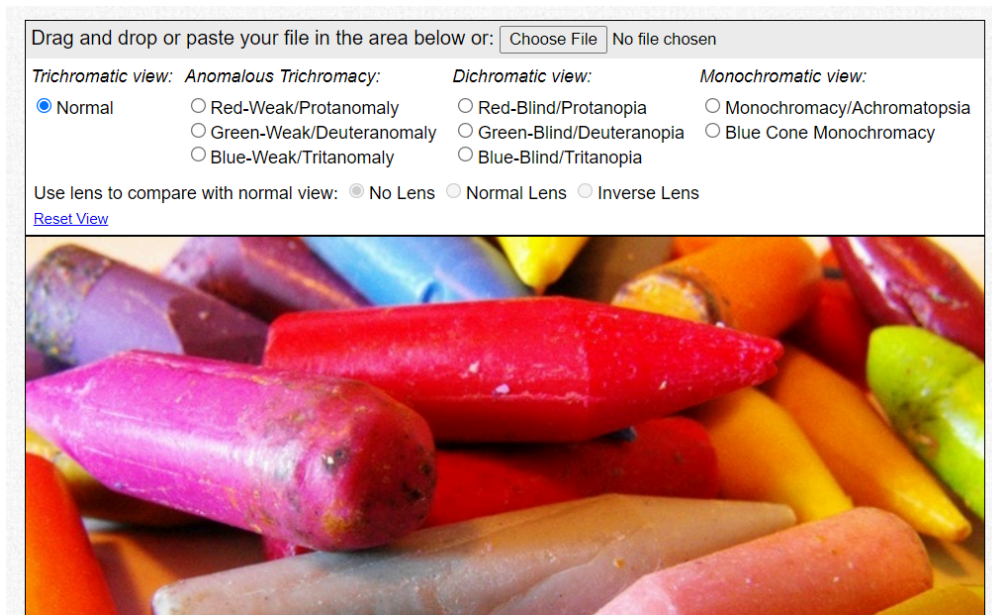
# Mark Size & Colour

- Small objects with colour should have more saturation
- Larger area = easier colour distinction



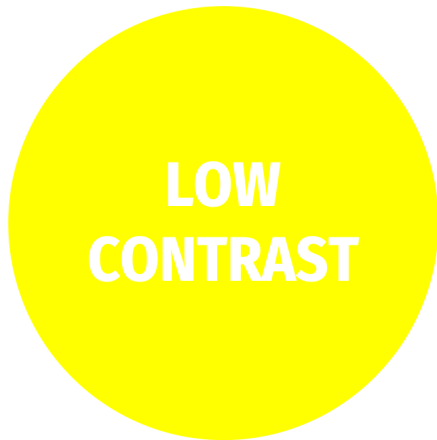
# Resource: Colour Blindness Simulator

- The [Colour Blindness Simulator](#) allows you to upload an image and apply various filters to check how its colours would be perceived by people with various types of colourblindness



# Brightness and Contrast

- Overly bright colours can be painful to look at, and overwhelming for people with sensory issues and people [on the autism spectrum](#)
- **Contrast** is the ratio between the brightness of two different colours
- A higher contrast generally makes colours and objects in our data visualizations easier to tell apart



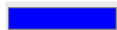



## Resource: Check Contrast



- WebAIM offers an [online tool](#) that lets you enter in two colour values and check if they are a 'pass' or 'fail' according to [Web Content Accessibility Guidelines](#) standards for contrast
- Colours can be picked from a dropper or using hexadecimal code
- The tool delivers separate contrast scores for normal text, large text, and graphical components

# Check Contrast

[Home](#) > [Resources](#) > Contrast Checker

**Foreground Color**  
#0000FF   
Lightness 

↔

**Background Color**  
#FFFFFF   
Lightness 

Contrast Ratio  
**8.59:1**

[permalink](#)

- [Contrast and Color Accessibility](#)
- [Quick Reference: Testing Web Content for Accessibility](#)
- [Web Accessibility for Designers](#)
- [Link Contrast Checker](#)

## Normal Text

WCAG AA: **Pass**

WCAG AAA: **Pass**

The five boxing wizards jump quickly.

## Large Text

WCAG AA: **Pass**

WCAG AAA: **Pass**

The five boxing wizards jump quickly.

## Graphical Objects and User Interface Components

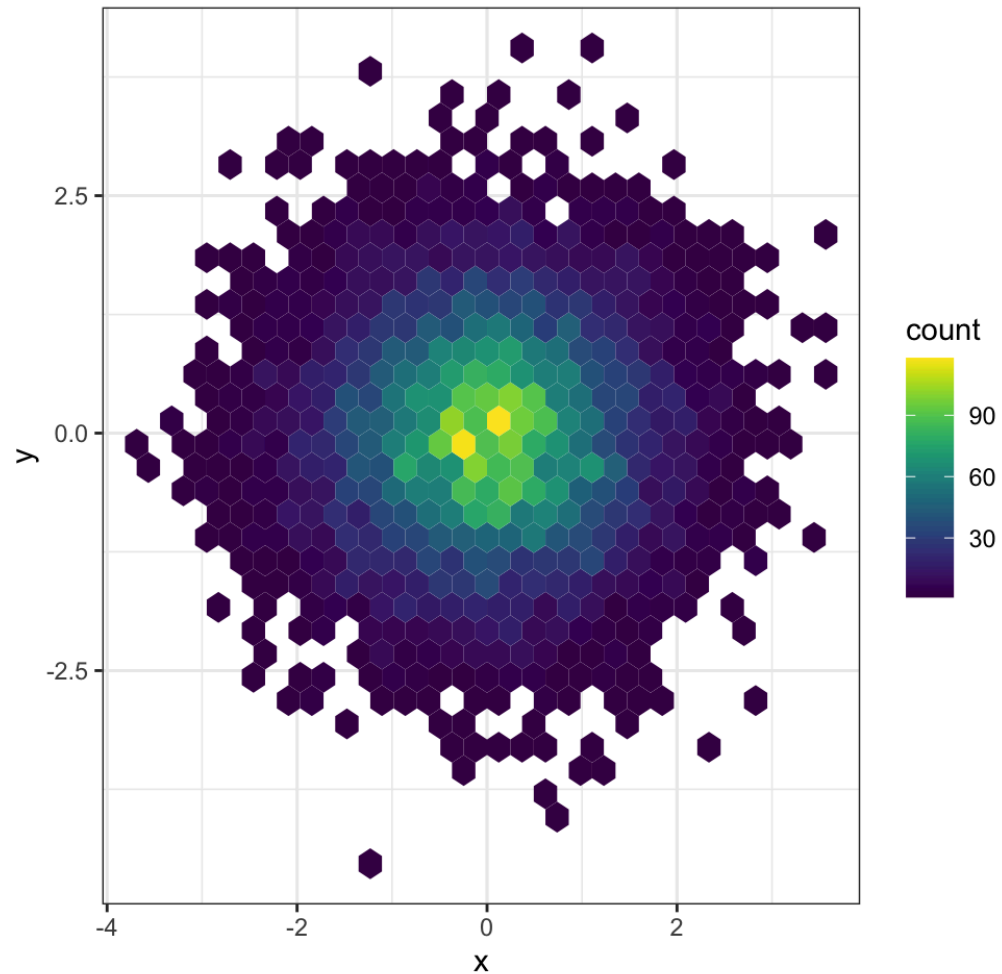
WCAG AA: **Pass**

Text Input




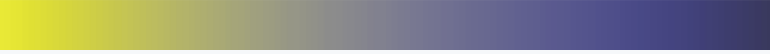




## Resource: Viridis package

- Viridis was made as an R package designed to make plots easier to read by those with colourblindness, and to maintain contrast so plots are readable in grayscale (eg. in print)
- Viridis can be used in Python with matplotlib as a colormap  
<https://matplotlib.org/stable/users/explain/colors/colormaps.html>

# Viridis package - Example

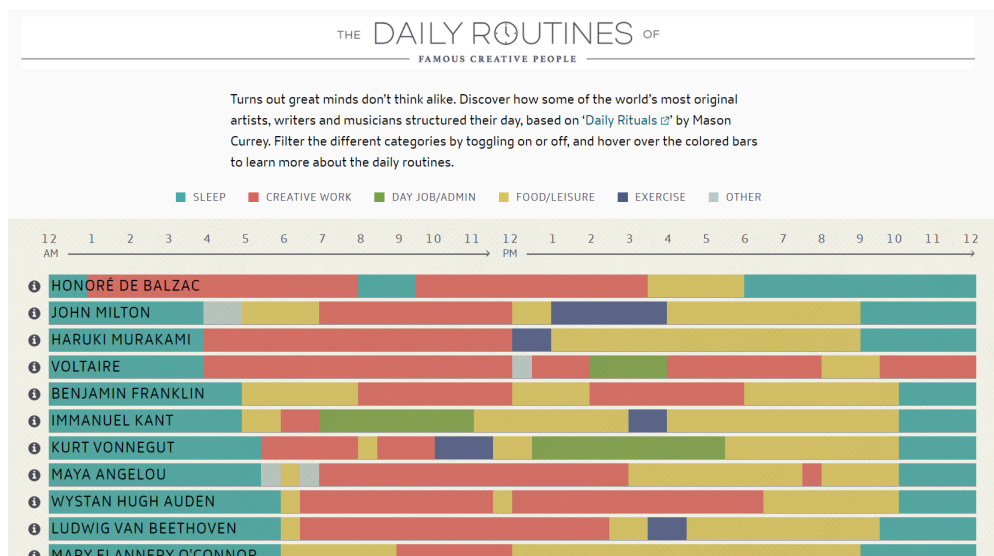


# Viridis package - As perceived with colourblindness

	ggplot Default Colour Palettes	viridis Default Colour
Base	 A horizontal bar showing the ggplot default color palette, which is a gradient from light blue to dark blue. The text "ggplot default" is centered above the bar.	 A horizontal bar showing the viridis default color palette, which is a gradient from yellow to green to blue to purple. The text "viridis" is centered above the bar.
Green-Blind (Deuteranopia)	 A horizontal bar showing the ggplot default color palette as perceived by someone with Green-Blindness (Deuteranopia). The colors are muted, appearing as a gradient from light purple to dark purple.	 A horizontal bar showing the viridis default color palette as perceived by someone with Green-Blindness (Deuteranopia). The colors are muted, appearing as a gradient from yellow to grey to dark blue.
Red-Blind (Protanopia)	 A horizontal bar showing the ggplot default color palette as perceived by someone with Red-Blindness (Protanopia). The colors are muted, appearing as a gradient from light purple to dark purple.	 A horizontal bar showing the viridis default color palette as perceived by someone with Red-Blindness (Protanopia). The colors are muted, appearing as a gradient from yellow to grey to dark blue.
Desaturated	 A horizontal bar showing the ggplot default color palette as perceived by someone with Desaturated vision. The colors are very muted, appearing as a gradient from light grey to dark grey.	 A horizontal bar showing the viridis default color palette as perceived by someone with Desaturated vision. The colors are very muted, appearing as a gradient from light grey to dark grey.

# Moving beyond colour

- Even when we make use of best practice for accessible use of colour in our data visualization, **we should avoid relying only on colour, where possible**
- Think back to our 'Daily Habits' example - use of patterns or textures in addition to colour on our bars could have made the visualization more accessible to people with visual impairments, even with the current colour scheme



**Text**

# Labels and legends

- If we want to minimize colours and patterns, data labels (on individual data points or on bars in a bar graph) can make visualizations easier to follow
- Every visual element on our graph should be clearly described in a legend; additional, brief written descriptions of visual elements can also **reduce anxiety** for our audiences





# Fonts and typefaces

- A **typeface** is a set of text characters (eg. Arial, Times New Roman)
- A **font** is a subset of a typeface (eg. 12 point Arial in italic)
- Certain typefaces and fonts are considered more accessible than others in design and data visualization contexts



# Typeface

- There is no single 'right' typeface or font to use, but generally sans-serif fonts are more accessible for [people with cognitive disabilities](#)
- Typefaces [identified](#) as easier to read for people with dyslexia include:
- Helvetica, Courier, Arial, Verdana
- We should [also](#) avoid overwhelming users by switching between too many different typefaces in one graphic

# Font size and spacing

- Text on our data visualizations should be at least 12 point size or greater, if possible
- Spacing also affects the readability of text on our visualizations
  - **Monospaced** text (each character takes the same amount of space) is more widely accessible than **proportionally spaced** text (each character only takes the space it needs)

A black rectangular box containing the digits 0 through 9 in a white, monospaced font. Below each digit is a yellow horizontal line with vertical end caps, indicating that every digit occupies the same amount of horizontal space.

**Monospaced Text**

A black rectangular box containing the digits 0 through 9 in a white, proportionally spaced font. Below each digit is a yellow horizontal line with vertical end caps, indicating that the width of each line varies to match the width of the digit above it.

**Proportionally Spaced Text**

# Describing data visualizations

# Alt-text

- To ensure that our data visualizations are accessible to people using a screen reader or unable to see the image, we include **alternate text (alt-text)**
- Alt-text should convey all the essential information conveyed by an image
- Generally, alt-text consists of a description of the image; if there is text in the image, alt-text should include the text
- Alt-text should be in plain language and simple sentences, when possible



## Activity: Alt-text

What would be suitable alt-text for this image?



# Activity: Alt-text

- The most appropriate alt-text can depend on the intended purpose of an image or visualization:
  - **If the image is only for decoration:** "Two scientists in discussion"
  - **If the focus is on specific individuals:** "From left-to-right: Eugenia Duodu and Patrick Gunning"
  - **If the focus is on the actual events or contents of an image:** "Professor Patrick Gunning and doctoral student Eugenia Duodo talking by a chemistry fume hood"

# Alt-text for graphs and plots of data

- Accessibly describing data-centric images for alt-text and figure captions requires additional considerations compared to describing photos such as the previous example
- Lundgard & Satyanarayan (2022) collected descriptions of sample data visualizations from graduate students and found that the descriptions could be grouped into four categories

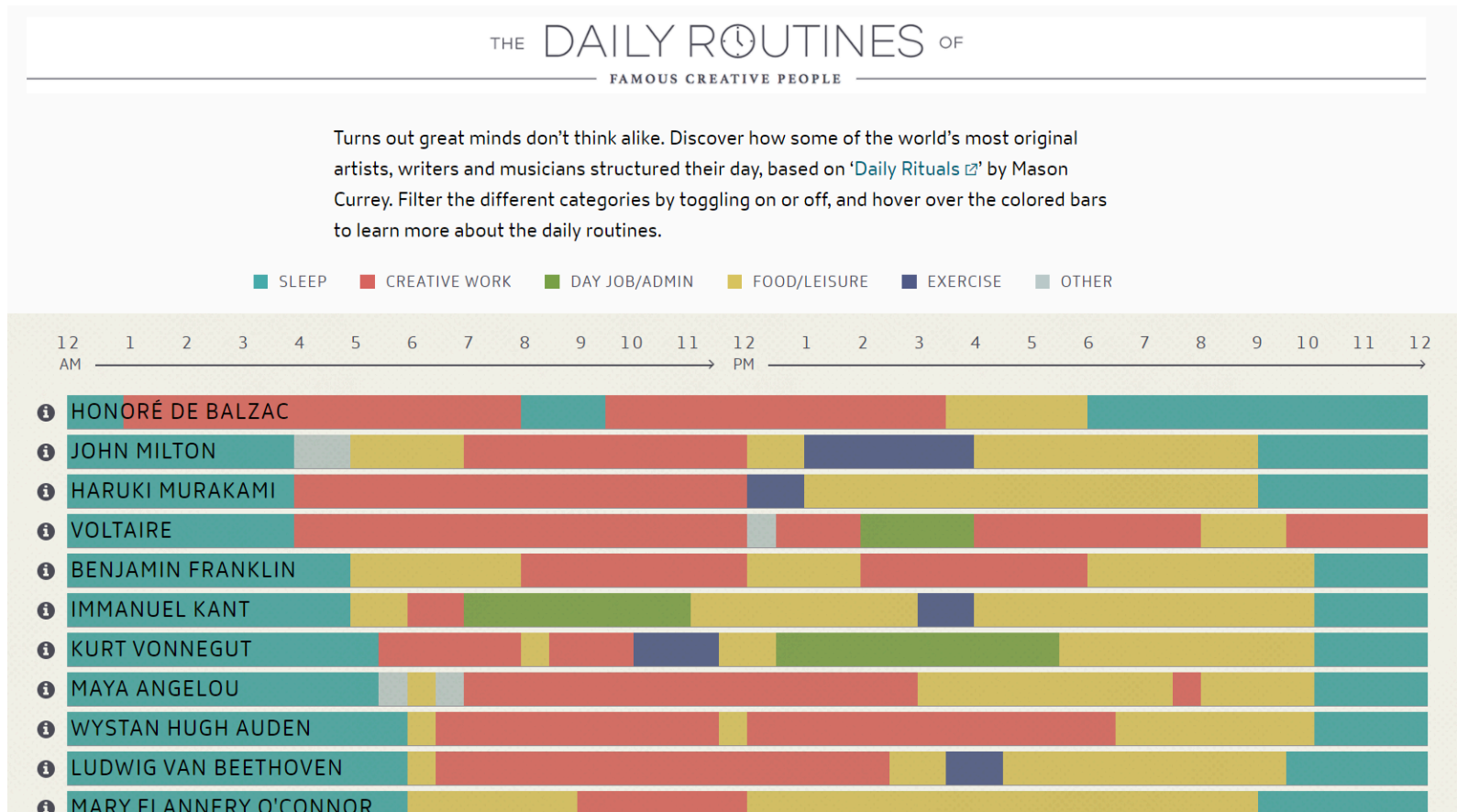


# Four types of descriptive content

Level	Semantic Content and Keywords
1	- Focus on chart elements such as chart type, title, axis ranges, labels, colours
2	- Descriptive statistics (eg. mean), outliers, max and min points, correlations, comparing points
3	- Complex trends, pattern synthesis, clusters, exceptions, common concepts - Perceptual interpretations of the data
4	- Context and domain insights, social and political context and explanations - Subjective interpretations that go beyond the data

# Activity: Four types of descriptive content

- Describe the data visualization on this slide according to each of the four types of descriptive content (click to visit website).



# Alt-text for graphs and plots of data

- The four types of descriptive text are **not** equally helpful!
- Both blind and sighted participants in the original study found level 1 content on its own the least helpful
- Blind participants found level 4 content the least useful
  - Blind participants preferred descriptions to contain facts about the data, giving them the agency to interpret it for themselves, rather than editorializing by the creator
- 40% of blind participants thought that the most useful descriptions would combine content from multiple levels (2 and 3)

**Can people actually access our data visualizations?**



# Cross-browser functionality

- If our data visualization is going to be embedded in a web page or otherwise accessible online, it is important to check for cross-browser functionality
- Our data visualization should be functional and visible in:
  - Chrome
  - Firefox
  - Internet Explorer 11, Microsoft Edge (Windows)
  - Safari (Mac/iOS)
- This is especially important in some professional contexts, where browser use may be limited/regulated

# Cost and accessibility

- If our data visualization is intended to reach the public or a particular subgroup of the public, the cost of accessing our visualization (and surrounding content) is an accessibility issue
  - Eg. Most academic publications require a subscription
- To maximize the accessibility of our data products, we can consider:
  - **Open Access** (Per [University of BC Libraries](#) , "a range of practices through which research outputs are distributed online, free of cost or other access barriers")
  - **Creative Commons licenses** (standard and simple [licenses](#) to facilitate sharing and reuse of work)

- When in doubt about the accessibility of our data visualizations...
- ...show somebody!
- **User testing and gathering feedback**, even informally, is the best way to understand whether our data visualizations are accessible and understandable to our intended audiences



## For more resources...

- The Government of Canada's [Digital Accessibility Toolkit](#) contains guides and graphics in plain language to facilitate designing with accessibility in mind
- The [AccessAbility handbook](#) is a detailed guide to accessible practices in graphic design, with useful insights for data visualization

