



CS 134 Data Visualization: Week 1

Joshua Goldberg

Edmonds College

Thank you to Allison Obourn and Andrew Heiss of Georgia State University for parts of these slides

Welcome to CS 134

- **Instructor:** Joshua Goldberg
 - **Education:** Master's in Data Science from the University of Chicago
 - **Industry:** Currently working as a Data Scientist at Amazon
- **Contact Information**
 - **Email:** josh.goldberg@edmonds.edu

What we will cover

Topics

Basics of graphic design

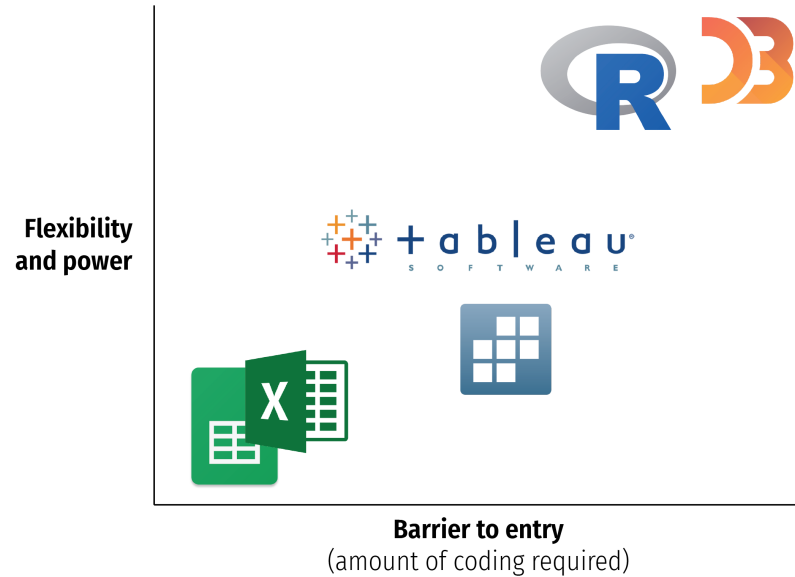
Mapping data to graphics

How **not** to lie with statistics

Core types of graphics

Special applications

Tools



- **Last Quarter:**
 - Excel
- **This Quarter:**
 - R / tidyverse
 - ggplot2
 - Tableau

The Logistics

- **Class schedule and format**
 - Tuesdays 6:00 - 7:50 pm
 - Thursdays 7:15 - 8:15 pm
 - All classes will be held on **Zoom**
- All class information will be posted on the **class website**

Required Materials

Everything you need for this class is **free**

- **Reading material**

- **Fundamentals of Data Visualization** by Claus O. Wilke
- **R for Data Science** by Hadley Wickham and Garrett Grolemund

- **Technology**

- **The R Programming Language**
- **RStudio / Posit** - available to download and install as well as **use in your browser**
- The **esquisse** library for R
- **Tableau**

Assigned Work

- **Required work**
 - No exams!
 - One assignment a week for the first 6 weeks
 - Larger final project
 - Presentation during our final exam slot
- **Grading breakdown**
 - **60%** Weekly assignments
 - **40%** Final project

Graphic Design

Principles of design

There are thousands of books and centuries of debate and theory about what makes good design.

No one agrees.

This isn't a graphic design class - we'll just discuss some general pointers.

Let's start with the CRAP

Contrast

Repetition Repetition Repetition

Alignment

Proximity

Use these as a checklist.

Contrast

If two items are not exactly the same, make them different.



REALLY DIFFERENT

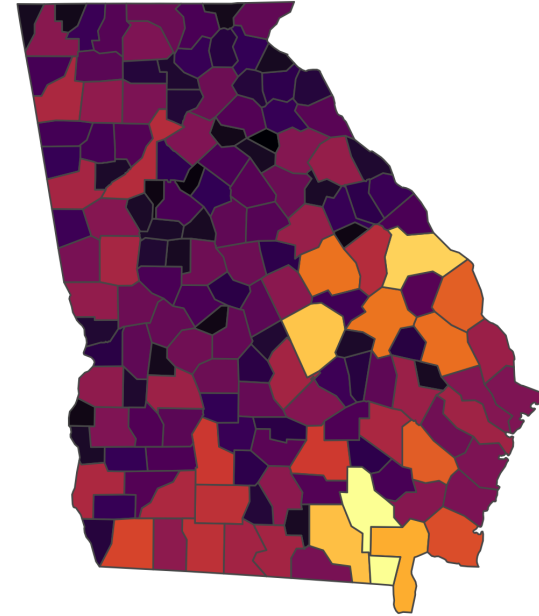
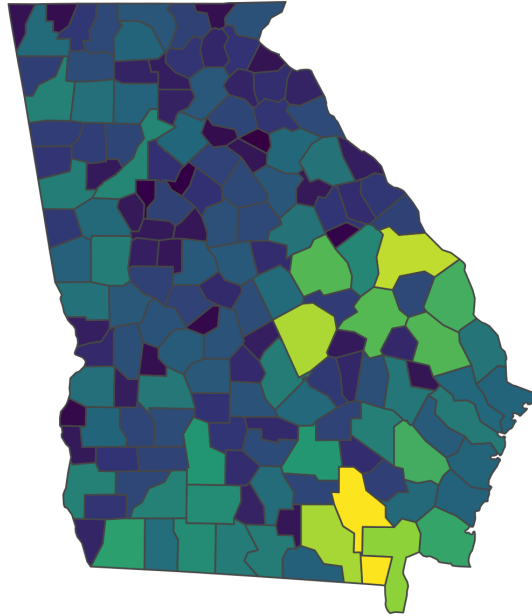
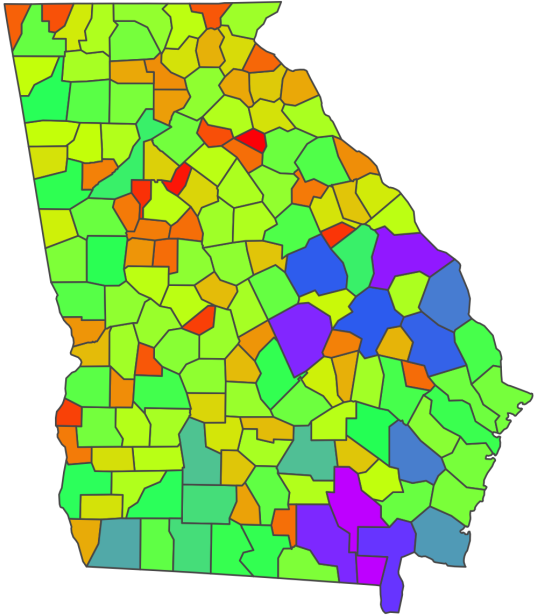
Create contrast with:

- Colors
- Font family
- Font **weight**
- Size

Color Theory

```
knitr::include_url("https://color.adobe.com/create/color-wheel")
```

Usability

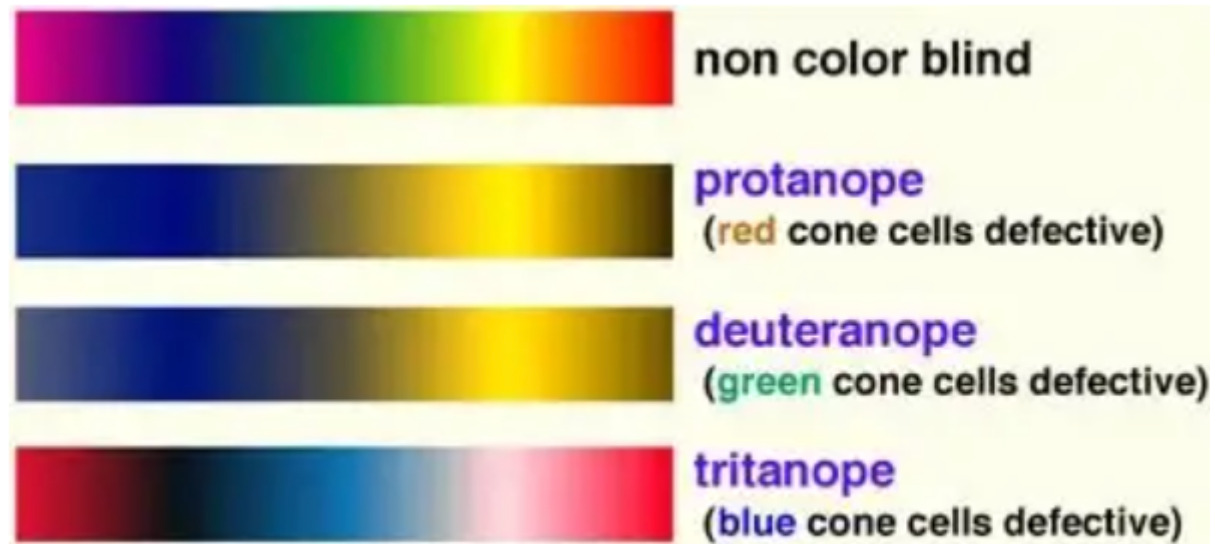


Values **close** to each other should use **similar** colors, and values far from each other use different colors

1. Colors should indicate which values are **greater** or **lesser** than the other values.
2. The **difference between colors** should represent the **difference between values**.

Accessibility

8% of men and 0.05% of women have some form of color blindness



Rainbow scale as seen by color-impaired persons. (Okabe, 2002)

We need to make sure we choose colors everyone can distinguish.

Making it all fit together

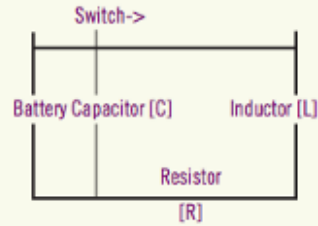
- Keep something **consistent** so your visualization has **cohesion**. Choose at least one of:
 - colors
 - fonts
 - families
 - weights
 - sizes
 - graphical elements
 - alignments
- Every item should have a **visual connection** with something else on the page.

Alignment

Share lines and repeat alignments where possible.

Example 6: Value of a resistor in an electrical circuit.

Find the value of a resistor in an electrical circuit which will dissipate the charge to 1 percent of its original value within one twentieth of a second after the switch is closed.

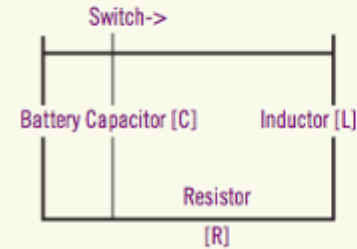


q0= 9 volts
q(t)= 0.09 volts
t= 0.05 seconds
L= 8 henrys
C= 0.0001 farads
R= 300 ohms
q(t)= 0.253889

1/[L*C_] 1250
[R_/(2*L)]^2 351.5625
SQRT(B15-B16) 29.973947
COS(T*B17) 0.07203653
-R_*T/(2*L) -0.9375
Q0+EXP (B 19) 3.52445064

Example 6: Value of a resistor in an electrical circuit.

Find the value of a resistor in an electrical circuit which will dissipate the charge to 1 percent of its original value within one twentieth of a second after the switch is closed.



q0= 9 volts
q(t)= 0.09 volts
t= 0.05 seconds
L= 8 henrys
C= 0.0001 farads
R= 300 ohms
q(t)= 0.253889

1/[L*C_] 1250
[R_/(2*L)]^2 351.5625
SQRT(B15-B16) 29.973947
COS(T*B17) 0.07203653
-R_*T/(2*L) -0.9375
Q0+EXP (B 19) 3.52445064

Proximity

Group similar things together to have **logical groupings**.

Ralph Roister Doister (717) 555-1212

Mermaid Tavern

916 Bread Street London, NM

Mermaid Tavern

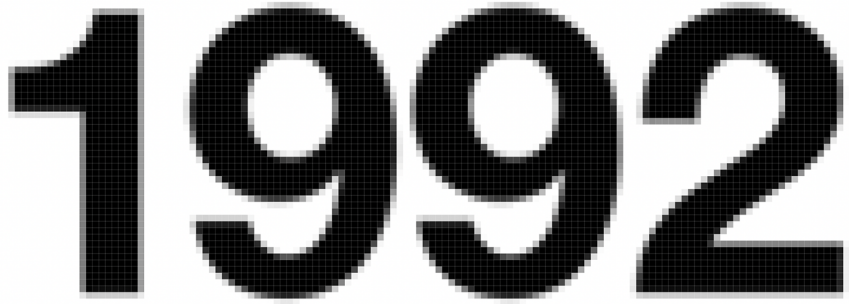
Ralph Roister Doister

916 Bread Street
London, NM
(717) 555-1212

Image Quality

Last, but not least: **image resolution** matters

No matter how good your visualization is, it won't look good if it is pixelated or blurry.

The image shows the year '1992' in a large, bold, black font. The text is pixelated, with visible square blocks of color, illustrating low image resolution.

- Some image types **compress data**
 - bitmap
- Others **compress colors**
 - png
 - gif
- Better option: **vector graphics**
 - svg
 - pdf
 - eps