

# 07 Visualization Research

36-721 Statistical Graphics and Visualization

Jerzy Wiecezorek

9/22/15

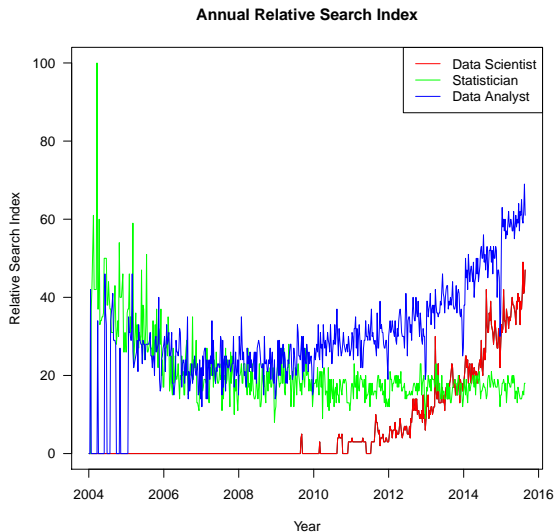
# Last time

- ▶ Interaction Design overview, elements, principles
- ▶ Shiny and D3 examples: we'll see more on Thursday
- ▶ HW 3

# Today

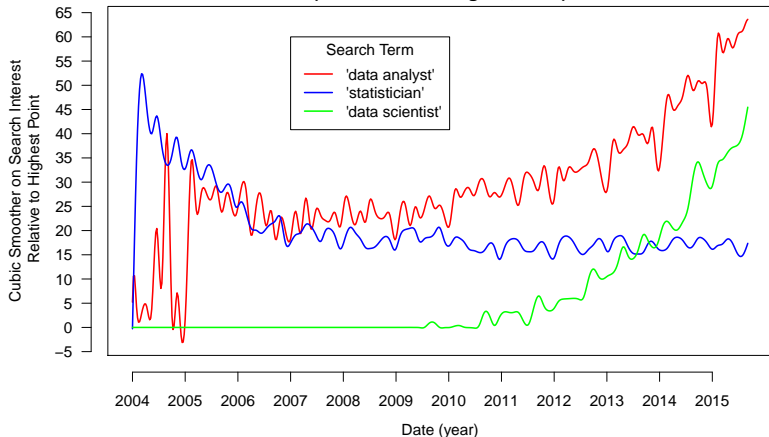
- ▶ Weds. office hours extended: 11:30am-1:30pm, BH 132M
- ▶ Highlight well-done HWs
- ▶ Discuss the Projects and Critique
- ▶ Visualization research: current topics, methods used, sources to follow
- ▶ Inkscape tutorial continued

# Well-done HW 1: raw data



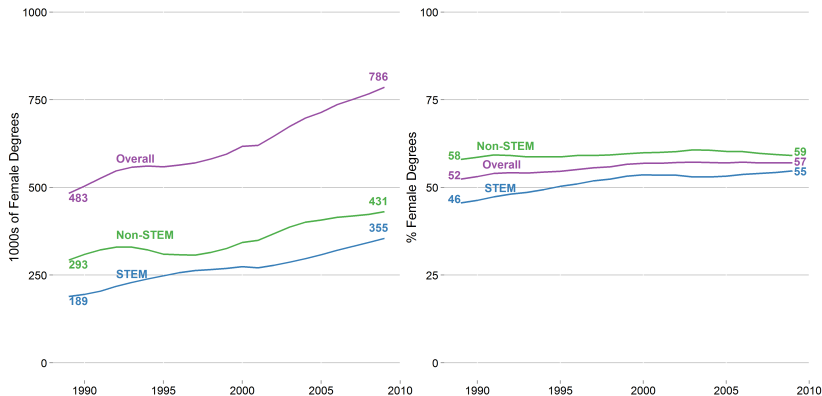
# Well-done HW 1: smoothing

**'Data Analysts' and 'Data Scientists' Are the New Buzz Words...**  
**Google Search Interest of Data Driven Jobs**  
**(Jan. 4, 2004 – Aug. 29, 2015)**

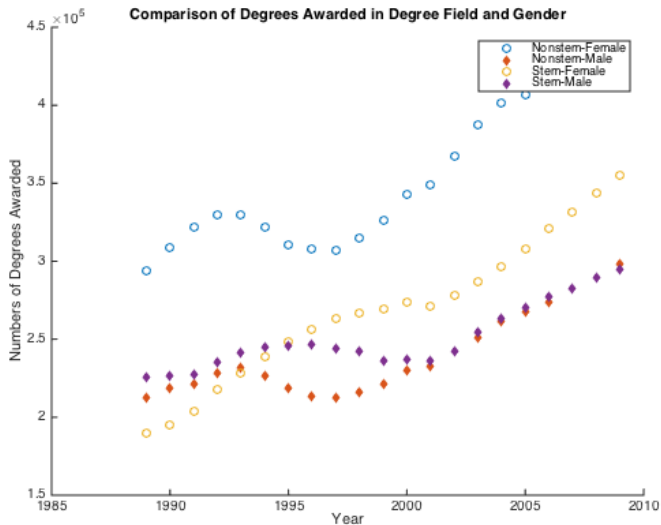


# Well-done HW 2: clean

Rise in Degrees Given to Females in US Higher Education



## Well-done HW 2: use of separable dimensions

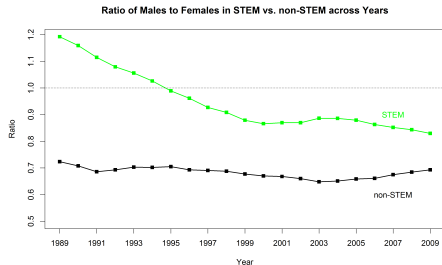
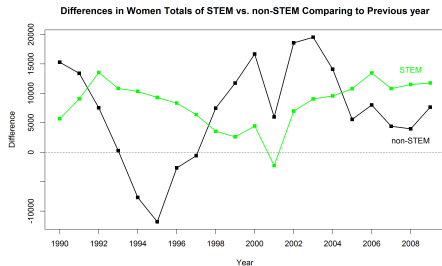


# Well-done HW 2: use of line widths



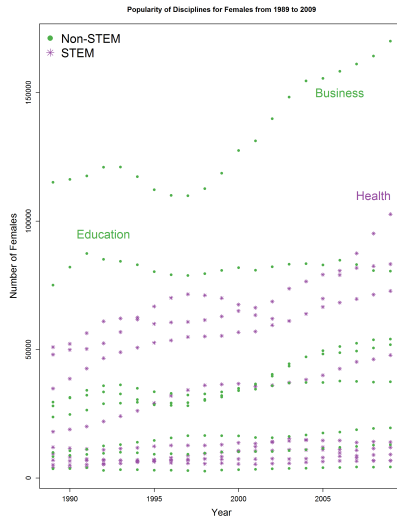
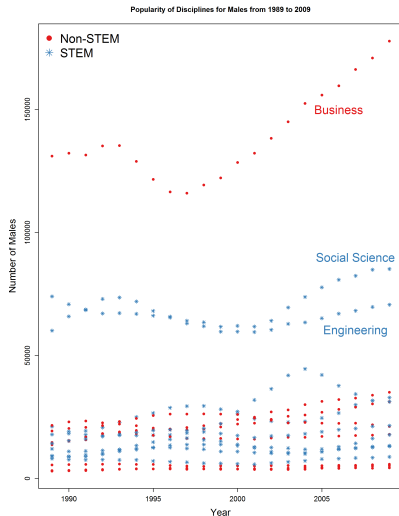


# Well-done HW 2: show year-to-year differences



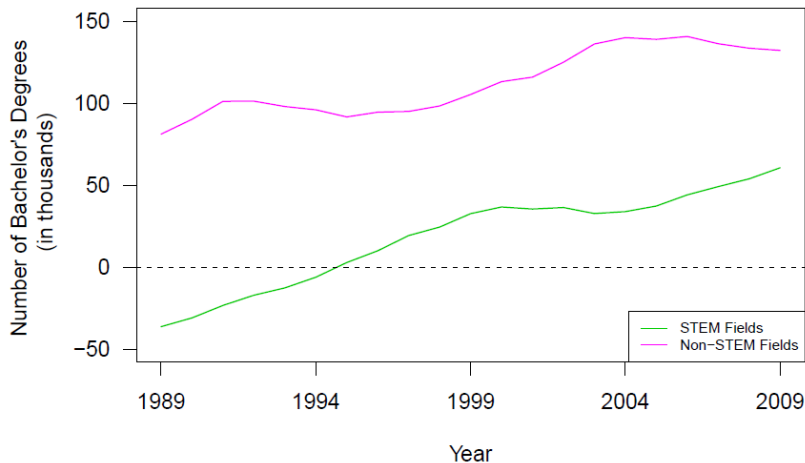
# Well-done HW 2: annotate interesting data

## Comparison of Disciplines from 1989 to 2009 by Gender



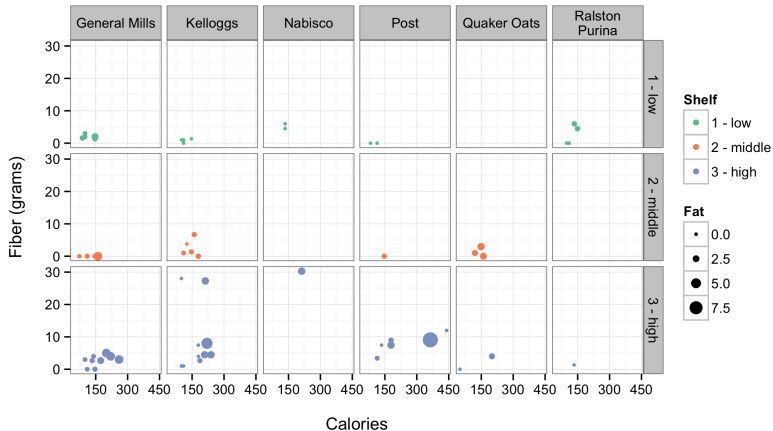
## Well-done HW 2: annotate with reference line

**Difference in Gender for STEM and Non-STEM Fields  
(Degrees for Females–Degrees for Males)**

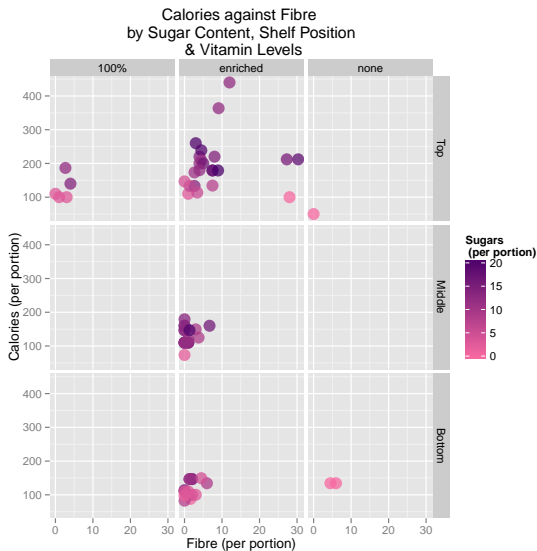


# Well-done HW 3: use of size aes

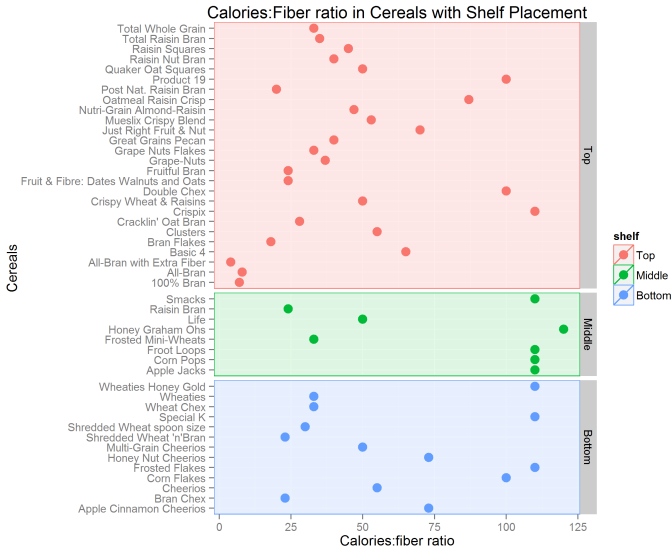
## US Cereal Trends by Calories, Fiber, Manufacturers, & Shelves



# Well-done HW 3: use of color aes

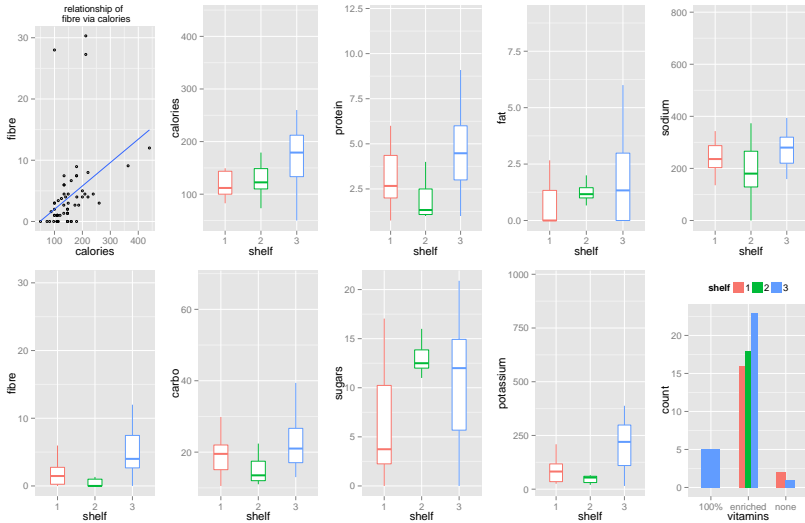


# Well-done HW 3: show all cereals

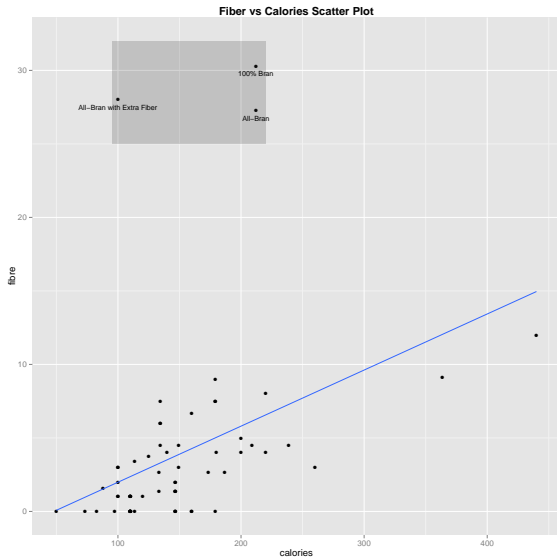


Well-done HW 3: show all variables

### Distribution of Cereal via Shelf

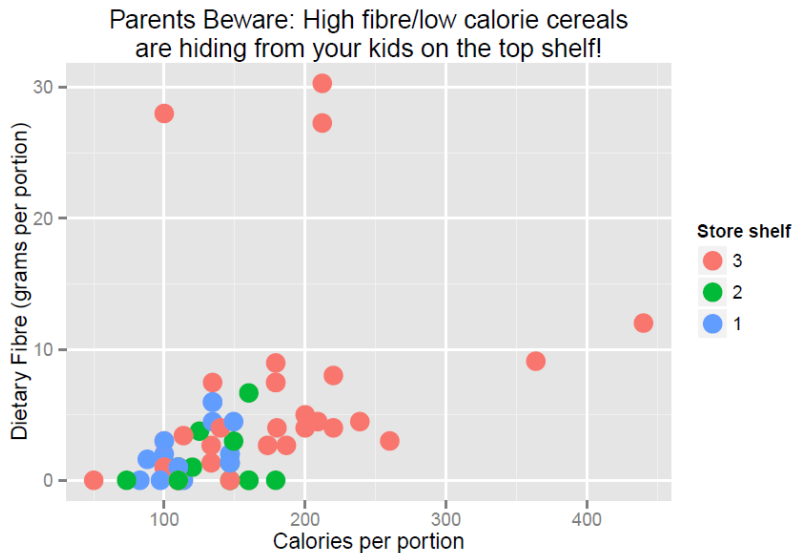


# Well-done HW 3: annotate interesting points





## Well-done HW 3: catchy title



# Projects and Critique

- ▶ Three projects: Graphic Design, Interaction Design, Research
- ▶ Critique a classmate's Graphic Design submission;  
turn in a graded rubric and detailed feedback

# Visualization research: current topics / open problems

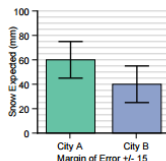
- ▶ Show statistical uncertainty/precision (in general)
- ▶ Show statistical uncertainty on maps
- ▶ Visual statistical inference
- ▶ Visual model diagnostics
- ▶ High-dimensional data
- ▶ Special data structure (networks)
- ▶ Collaborative visualization
- ▶ ...

# Show statistical uncertainty

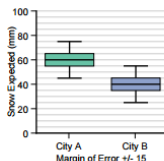
Error bar variations: Correll and Gleicher (2014)

## Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error

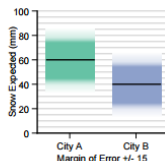
Michael Correll *Student Member, IEEE*, and Michael Gleicher *Member, IEEE*



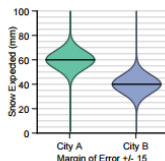
(a) **Bar chart** with error bars: the height of the bars encodes the sample mean, and the whiskers encode a 95% t-confidence interval.



(b) **Modified box plot**: The whiskers are the 95% t-confidence interval, the box is a 50% t-confidence interval.



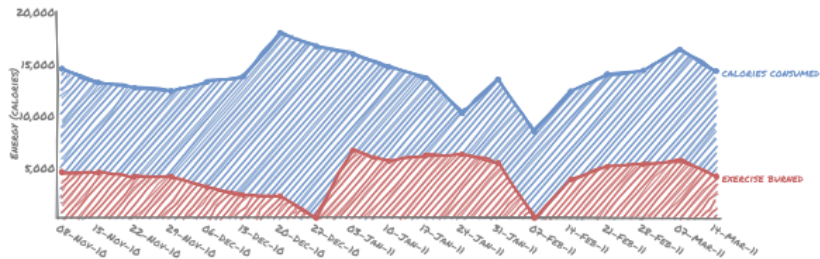
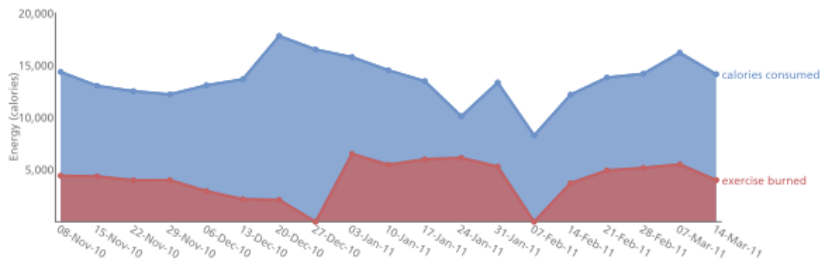
(c) **Gradient plot**: the transparency of the colored region corresponds to the cumulative density function of a t-distribution.



(d) **Violin plot**: the width of the colored region corresponds to the probability density function of a t-distribution.

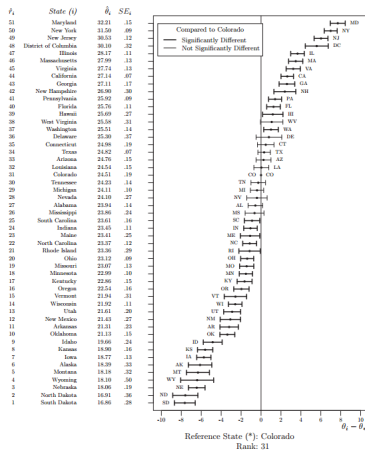
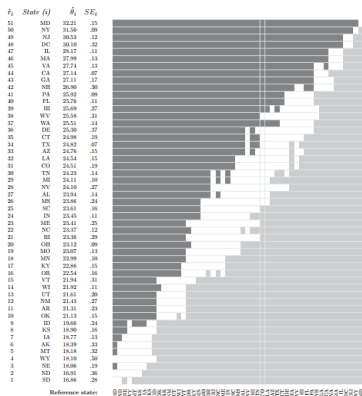
# Show statistical uncertainty

Use novel visual variables like **sketchiness**: Wood et al. (2012)



# Show statistical uncertainty

Statistical significance of confidence interval overlap  
and multiple comparisons: **Wright et al. (2013)**



# Show statistical uncertainty on maps

Francis et al. (2012):

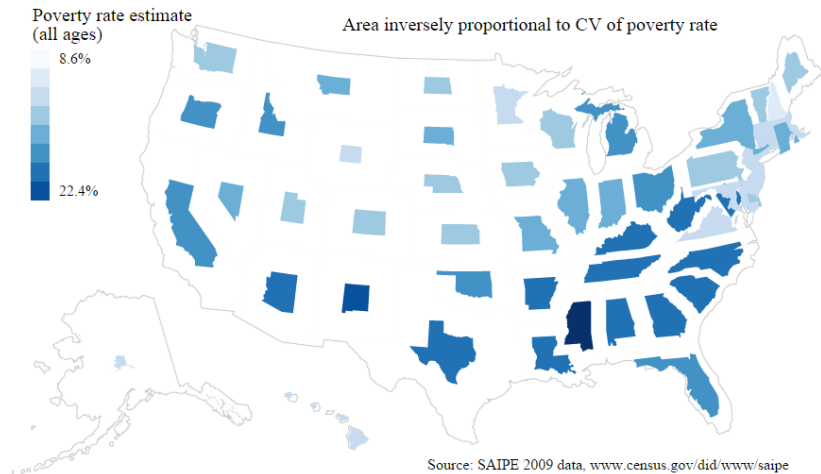
- ▶ Bivariate choropleth (color for estimate, hatching for precision)
- ▶ Side-by-side choropleth (estimate on left, precision on right)
- ▶ Mouseover or sliders

Maps can also involve other kinds of uncertainty: attribute (is this forest or farmland?), spatial (where are the boundaries between areas?), classification (where best to cut data classes for color bar categories?), etc.

# Show statistical uncertainty on maps

Francis et al. (2012): "To date, to our knowledge no one has explored [the] use of cartograms for this purpose."

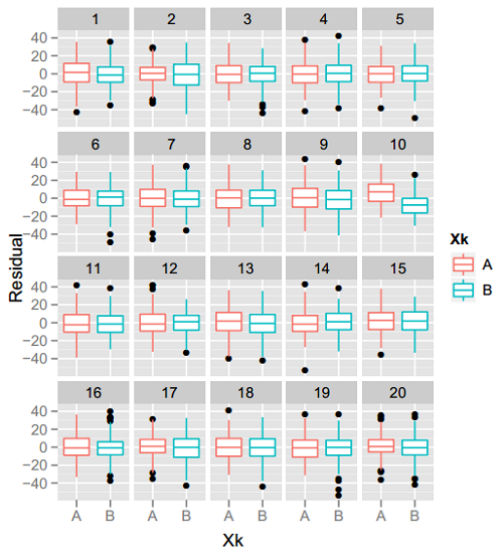
**Cartogram** (color for estimate, size for precision)





# Visual statistical inference

Buja et al. (2009): using graphs of random or permuted data as a visual hypothesis test



# Visualization research: methods / justifications used

- ▶ Anecdotes or heuristics
- ▶ Lab experiments: tasks, eye tracking, Mechanical Turk
- ▶ Ethnographic studies: use of graphs in real practice

# Visualization research: sources to follow

## Academic conferences:

- ▶ ACM SIGGRAPH (see also SIGCHI and CSCM)
- ▶ EuroVis
- ▶ IEEE VIS
- ▶ useR!

## Journals:

- ▶ ASA Journal of Computational and Graphical Statistics
- ▶ IEEE Transactions on Visualization and Computer Graphics

# Visualization research: people to follow

## Researchers:

- ▶ Di Cook
- ▶ Jeff Heer
- ▶ Heike Hofmann
- ▶ Alan MacEachren
- ▶ Ben Shneiderman

## Bloggers:

- ▶ Alberto Cairo
- ▶ Kaiser Fung
- ▶ Andy Kirk
- ▶ Robert Kosara
- ▶ Nathan Yau

# Inkscape tutorial

Continue tutorial using materials from Lecture 05

- ▶ Editing graphs and text
- ▶ Layout, grid, guides
- ▶ Layers

## For next time

- ▶ Thursday will be a lab/recitation day: work through Shiny exercises (in groups), and review solutions as a class
- ▶ Before Thursday, **please complete lessons 1 to 5 of RStudio's [Shiny tutorial](#)**, or look into any one of
  - ▶ Scott Murray's [D3 tutorial](#)
  - ▶ [ggvis website](#)
  - ▶ [animint website](#)
- ▶ Project 1 (Graphic Design) due Sat. 5pm, through Blackboard
- ▶ Start thinking about Project 2 (Interaction Design) and Project 3 (Research)