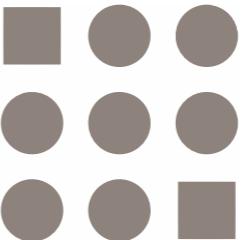


CS 171

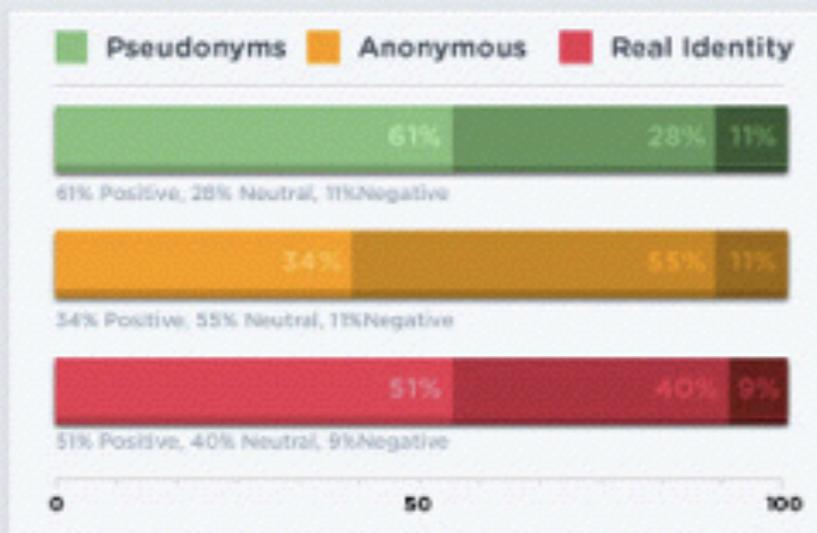


# Design Principles

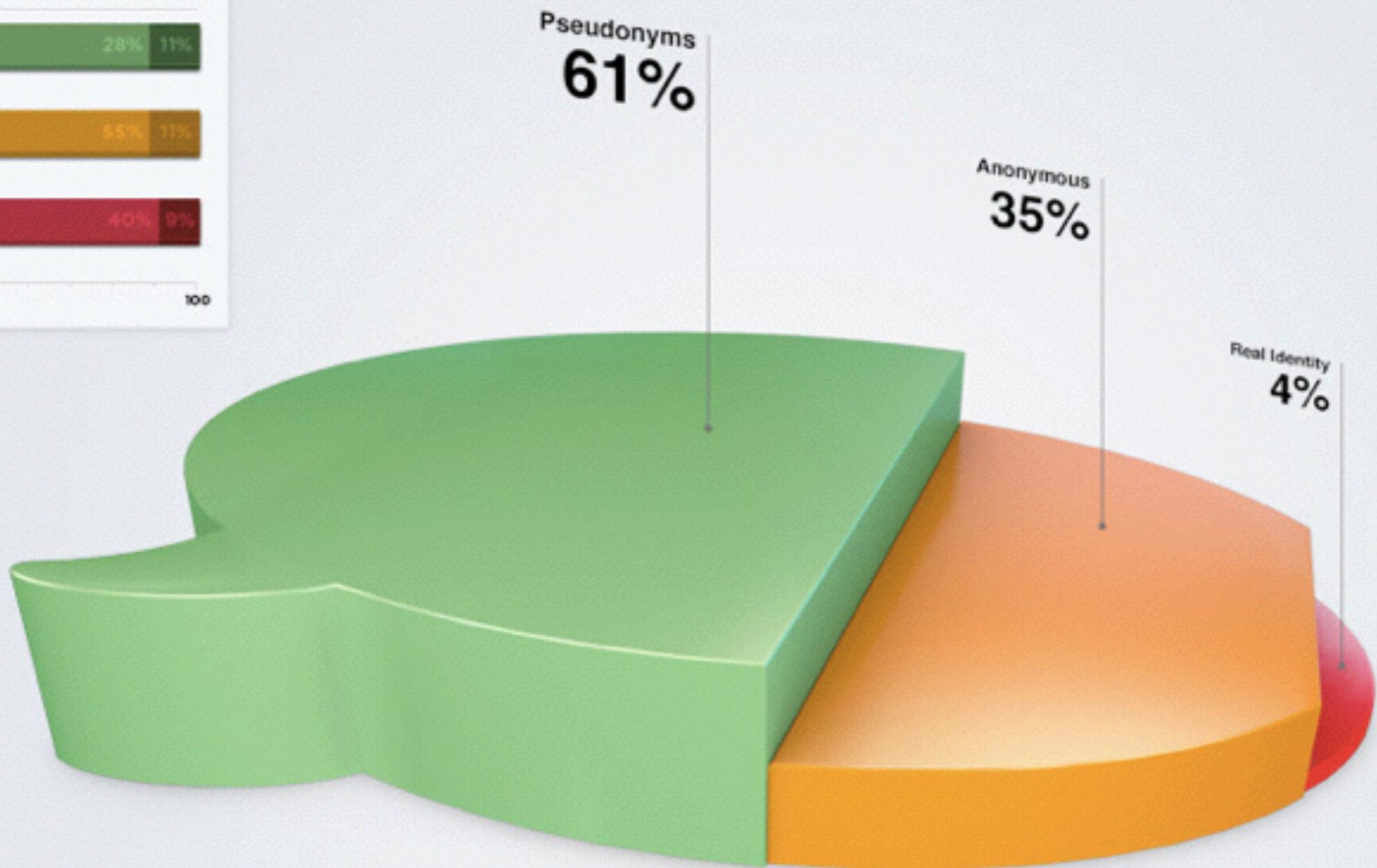
Johanna Beyer

[jbeyer@seas.harvard.edu](mailto:jbeyer@seas.harvard.edu)

## Quality Signals by Identity



## Percentage of Comments by Identity



## Average Comments Per User By Identity

The average commenter using a pseudonym contributed **6.5** times more than anonymous commenters and **4.7** times more than commenters identifying with **Facebook**.

# All of Inflation's Little Parts

Each month, the Bureau of Labor Statistics gathers 84,000 prices in about 200 categories — like gasoline, bananas, dresses and garbage collection — to form the Consumer Price Index, one measure of inflation.

It's among the statistics that the Federal Reserve considered when it cut interest rates on Wednesday. The categories are weighted according to an estimate of what the average American spends, as shown below.

## An Average Consumer's Spending

*Each shape below represents how much the average American spends in different categories. Larger shapes make up a larger part of spending.*

*Color shows change in prices from March 2007 to March 2008*



[ZOOM IN](#) [ZOOM OUT](#)

### Food and beverages 15%

The high price of oil is a factor that has made food prices rise quickly.

### Miscellaneous 3%

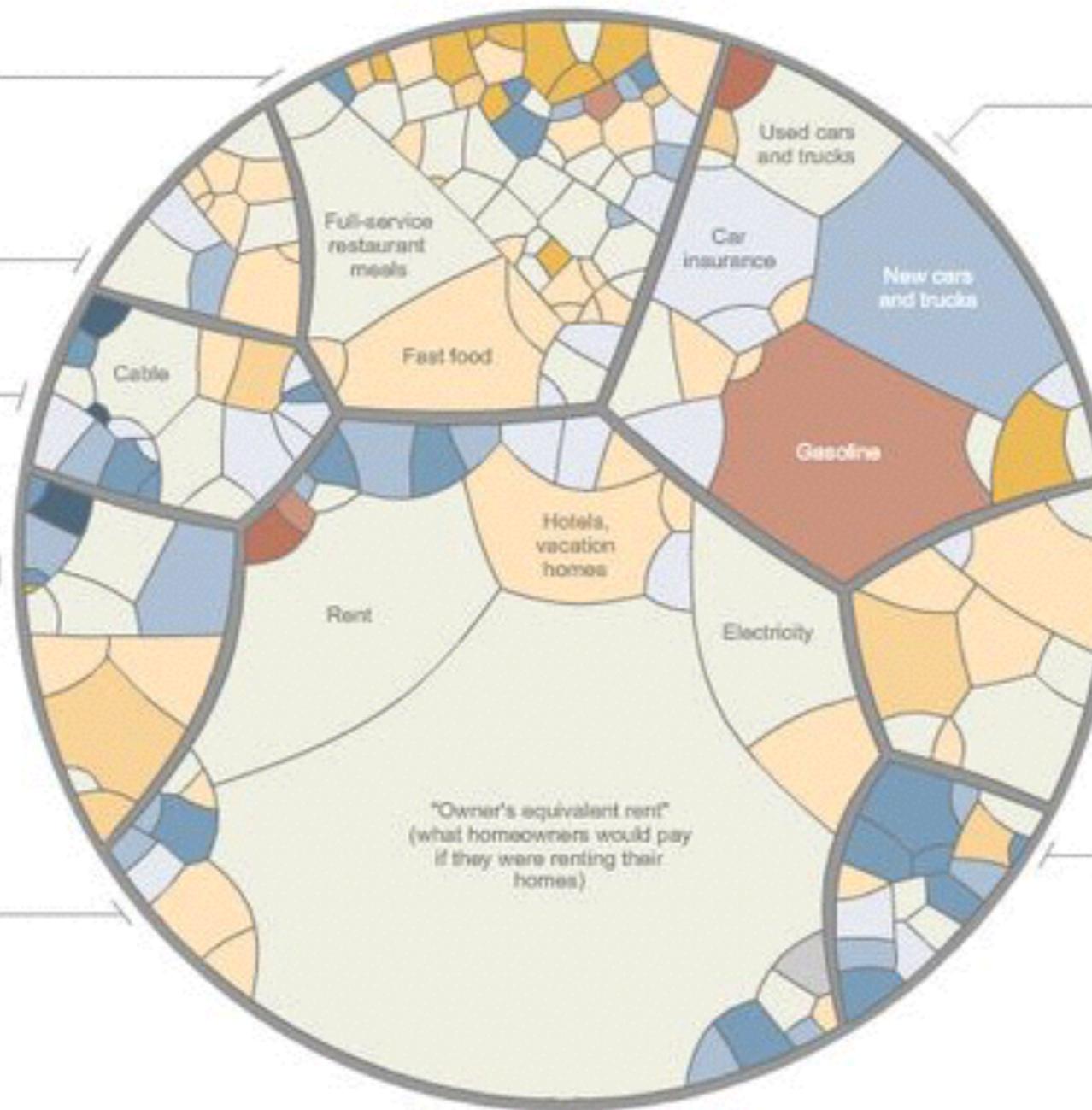
### Recreation 6%

### Education/Communication 6%

Cellphones were added to the index in 1997. Because the Consumer Price Index can be slow to add new goods, which are often cheaper, it may overstate parts of inflation.

### Housing 42%

In the C.P.I., home ownership costs track rent prices more closely than housing prices. This means inflation may have been understated when home prices were rising faster than rents.



### Transportation 18%

Gas is 5.2 percent of spending nationwide, but only 3.8 percent in the New York area.

### Health care 6%

As a group, the elderly spend about twice as much of their budget on medical care.

### Apparel 4%

The ratio of spending on women's clothes to that on men's clothes is about 2 to 1.

# Today

- Use iterative design methods for idea generation to diverge on designs
- Introduction to Bertin's semiology and Tufte's main design principles
- Learn vocabulary to evaluate which visualizations are more effective than others
- Define Chart Junk and articulate the positive and negative implications of its use

# Check In

- Class signup - done!
- Sign up for studio!
- Pre-Quiz submitted?
- Get some sheets of paper & sharpies

# Activity Procedure

- We like to share the results of the activities with the rest of the class
- TFs will take photos - don't be scared - and some photos will be shown after the activity
- You can volunteer to comment on your work but you can also stay anonymous
- Do not take anything we say personally! We try to be constructive. We know this is hard!

# Activity

Create at least **three** sketches to visualize these two quantities. (1 min)



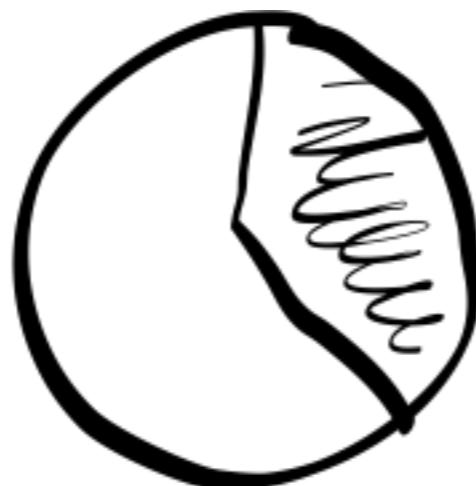
4 2



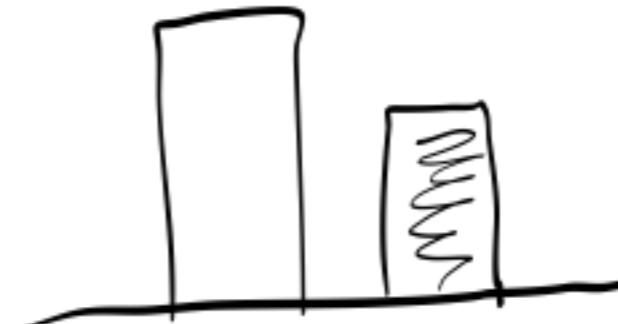
12 3

# Most likely results

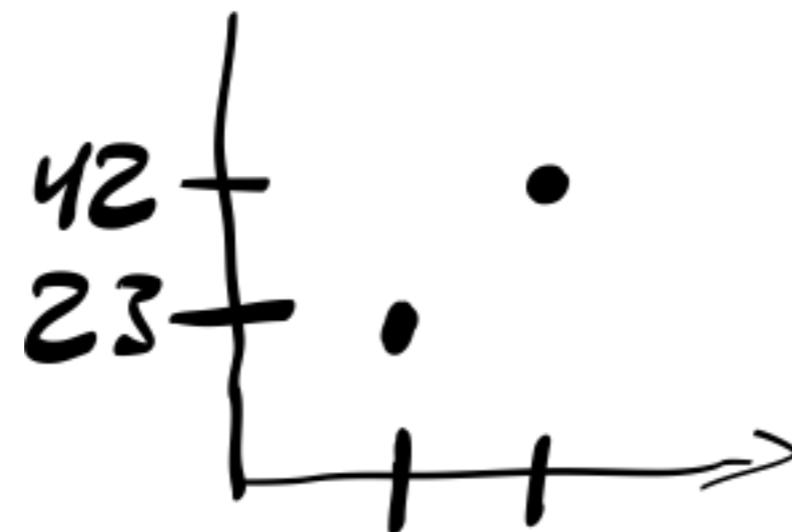
Pie Chart



Bar Chart



Scatterplot

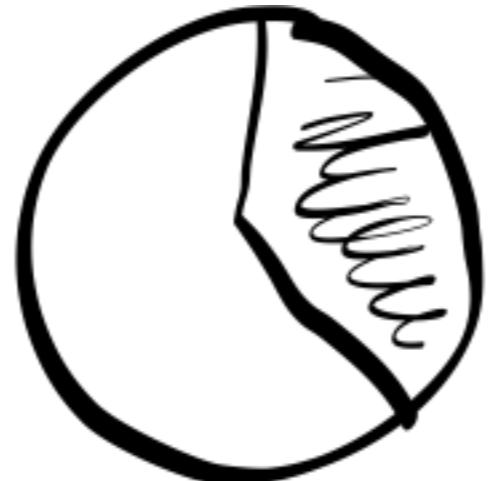


23  
42

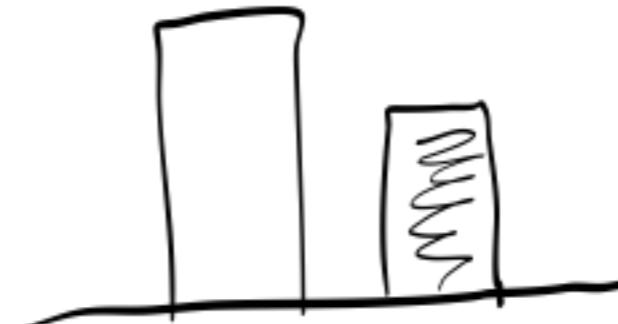
Arabic Numbers

# Design Fixation: Blind adherence to a set of ideas or concepts

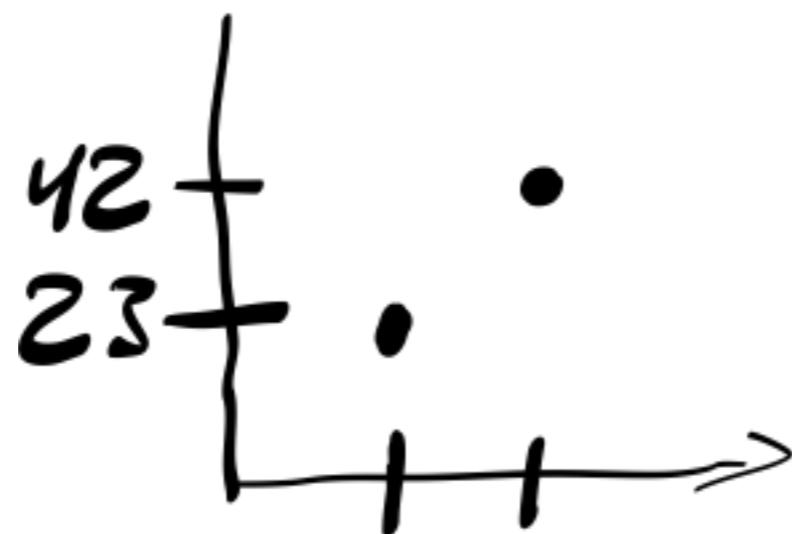
Pie Chart



Bar Chart



Scatterplot



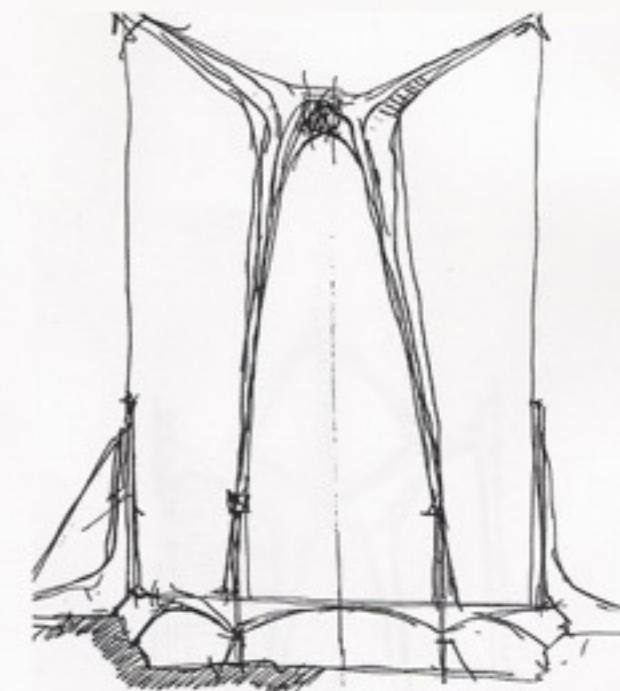
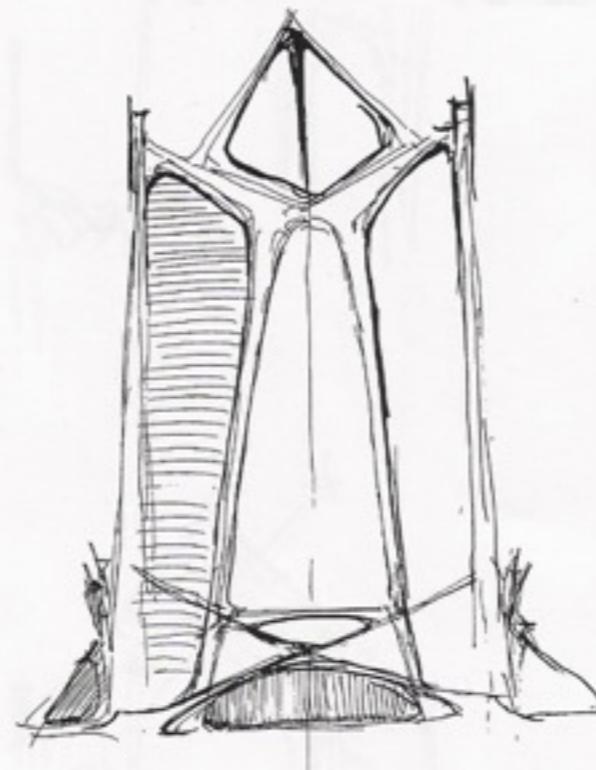
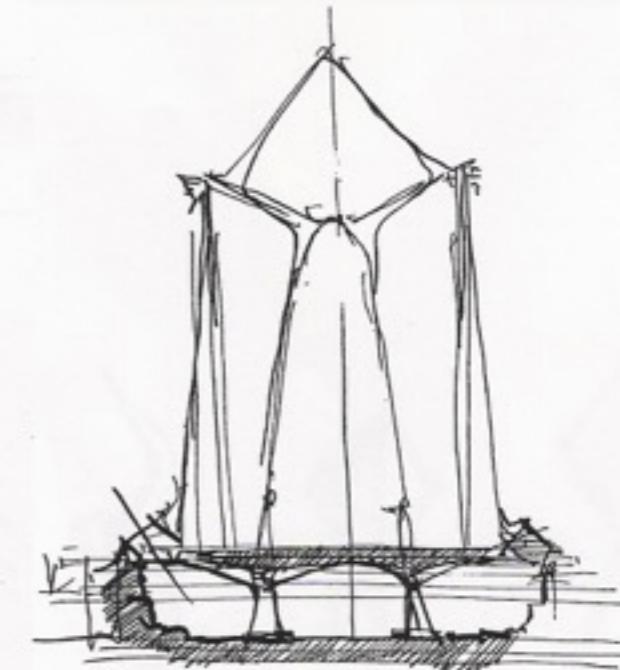
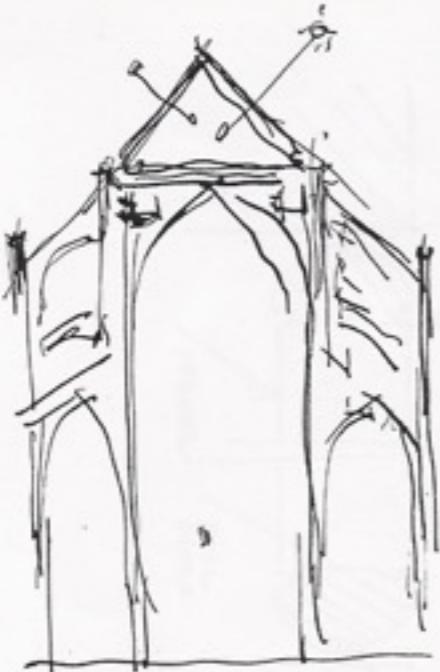
23  
42

Arabic Numbers

# Design and sketching are *constructive* perception



Sketches are cheap, fast, and easy to throw away



Santiago Calatrava, Architect

“My strategy has always been:  
be wrong as fast as we can.”

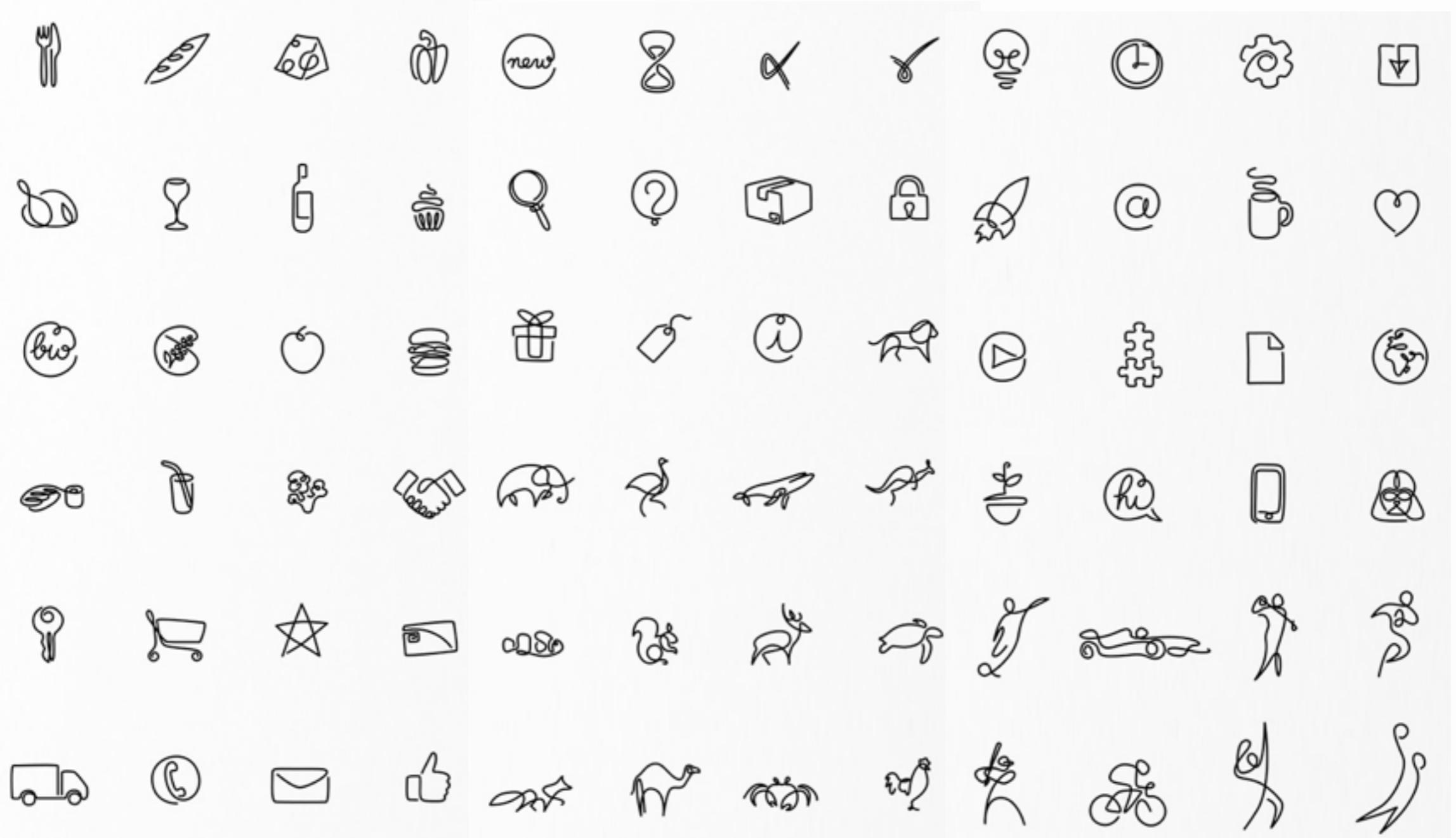
Andrew Stanton,  
Director of Finding Nemo  
and WALL-E



DORY  
Shouldn't we get out of its way?

DORY  
Hey, is that a --

# Constraints: a set of icons that are based on only **one** line



<http://www.differantly.com/#oneline/>

# Activity

Create at least **three** new visualizations that are different from your previous ideas using a constraint - e.g., one line, only black/white, only round objects,...

(2 mins)

42

23

# Jacques Bertin

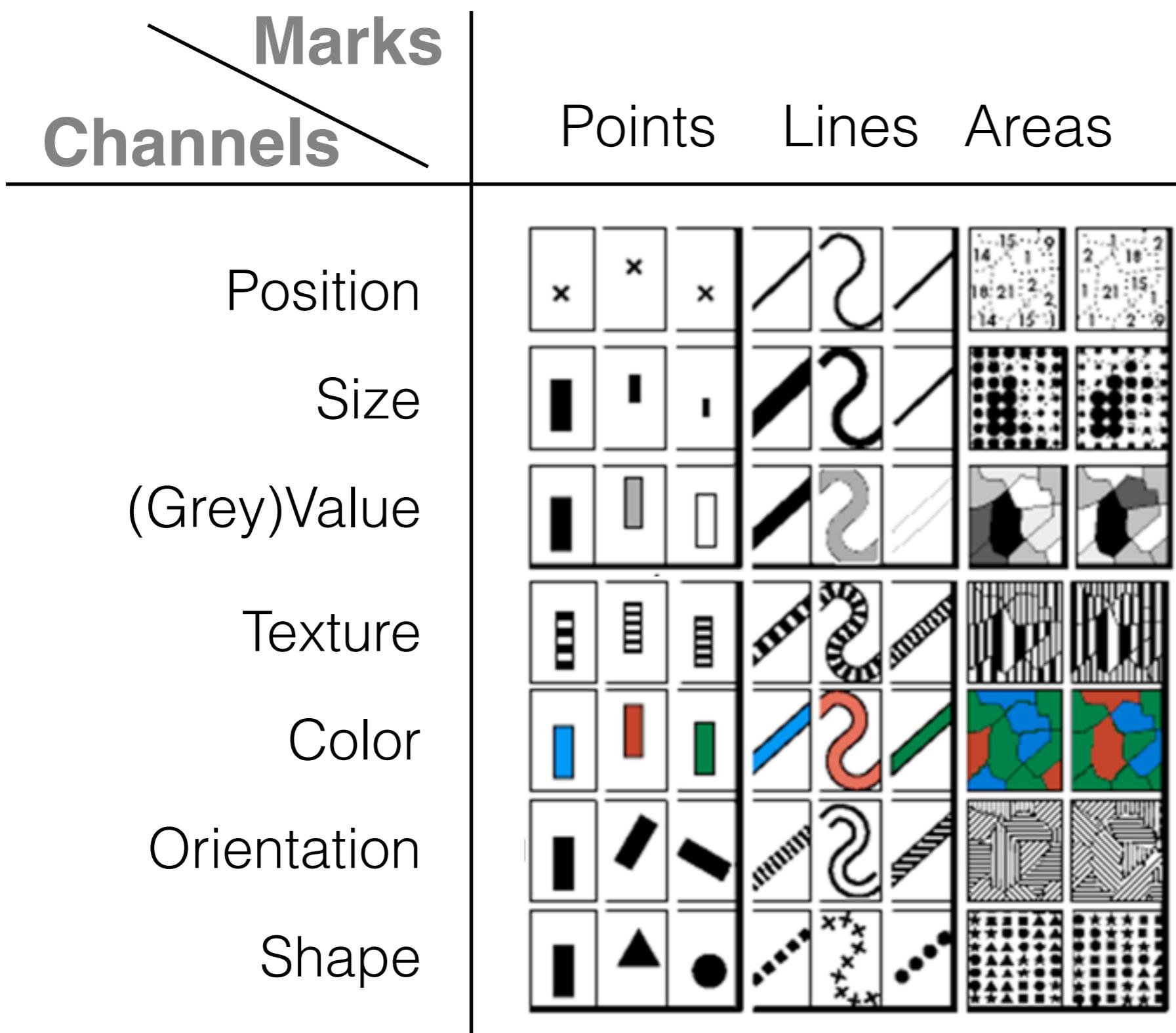
French cartographer  
[1918-2010]

Semiology of Graphics  
[1967]

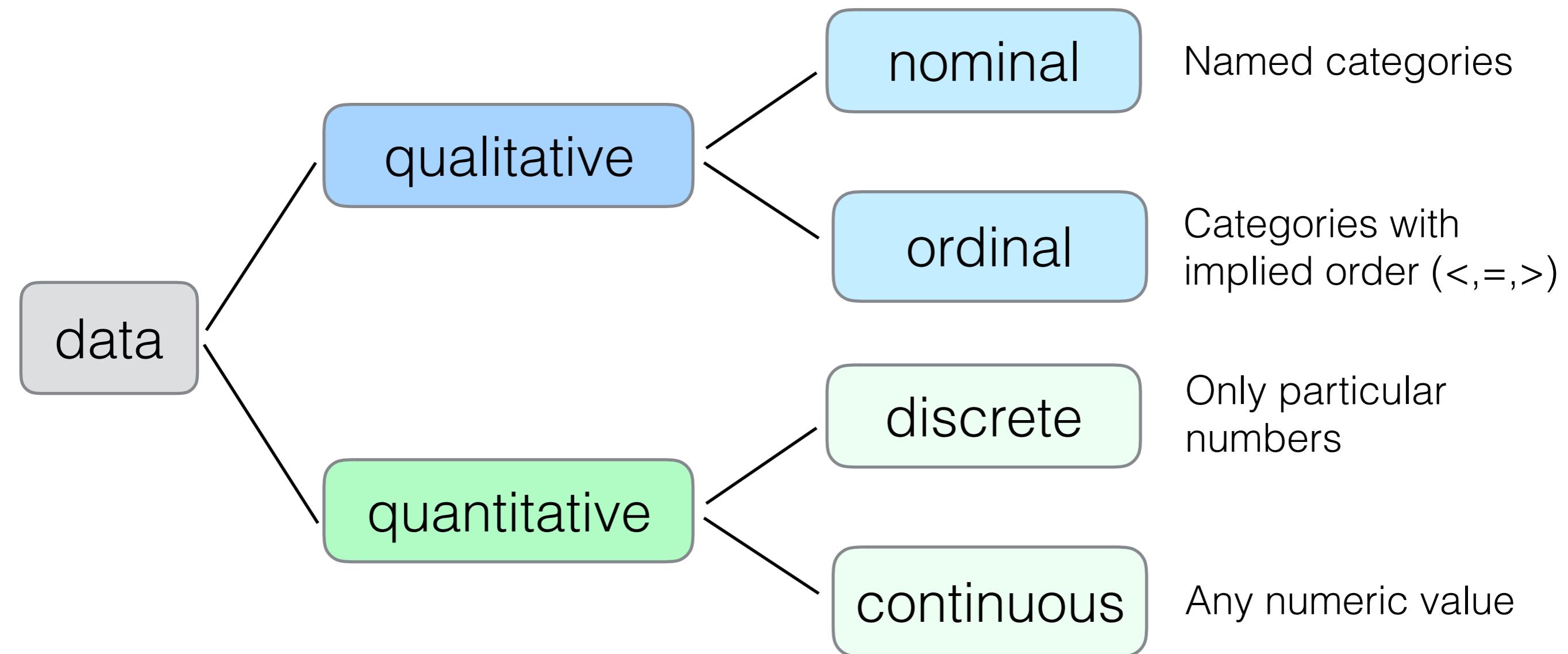
Theoretical principles for  
visual encodings



# Bertin's visual design space



# Types of data



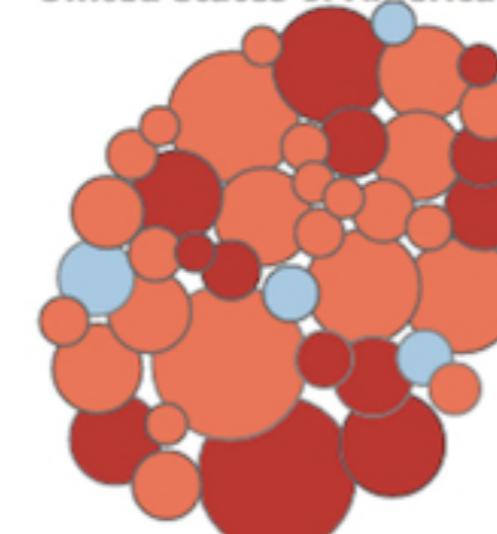
# Accuracy of channels

	<b>Quantitative</b>	<b>Ordinal</b>	<b>Nominal</b>	
More Accurate ↑	Position Length Angle Slope Area Density Saturation Hue Shape	Position Density Saturation Hue Length Angle Slope Area Shape	Position Hue Density Saturation Shape Length Angle Slope Area	Position Hue Density Saturation Shape Length Angle Slope Area
	••	••	••	••
	==	•••	•••	•••
	∠	•••	•••	•••
	/\	•••	•••	•••
	••	==	••	••
	•••	∠	==	==
	•••	/\	∠	∠
	•••	••	/\	/\
↓ Less Accurate	••	••	••	••



## AMERICAS

United States of America

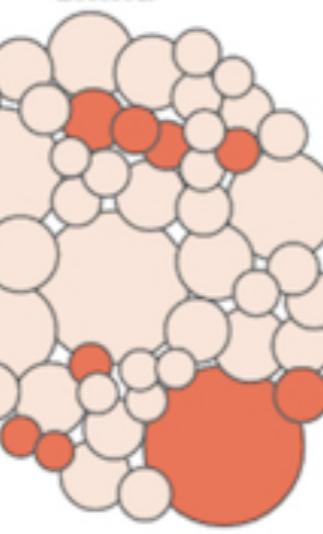


Argentina



## ASIA

China

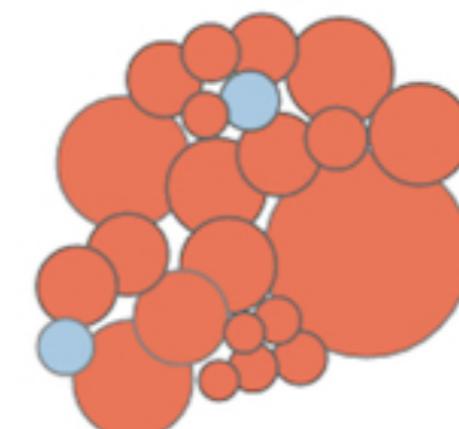


Japan

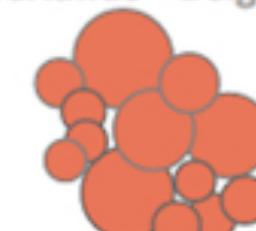


## EUROPA

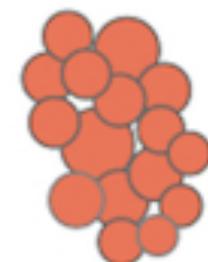
Great Britain



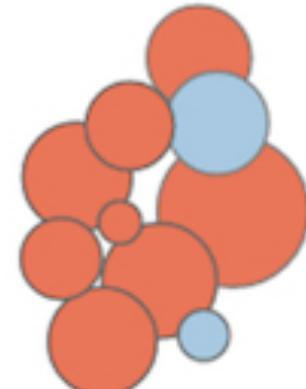
Netherlands - Belgium



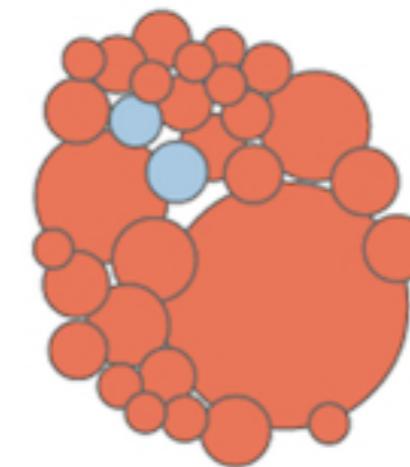
Germany



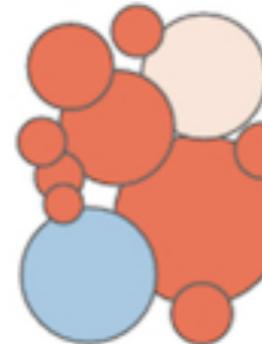
Others



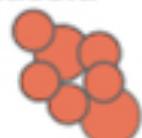
Spain



Italy



Russia



## ART &amp; MONEY

#datavisualisation



This Animated Bubble Chart shows the 270 most expensive artworks sold in auction since 2008 until end 2011



## SORTING

- year by year
- top 10 artworks
- men / women
- dead / alive
- by nationality
- best-selling artists
- auction houses
- size of artworks
- date of creation (all centuries)

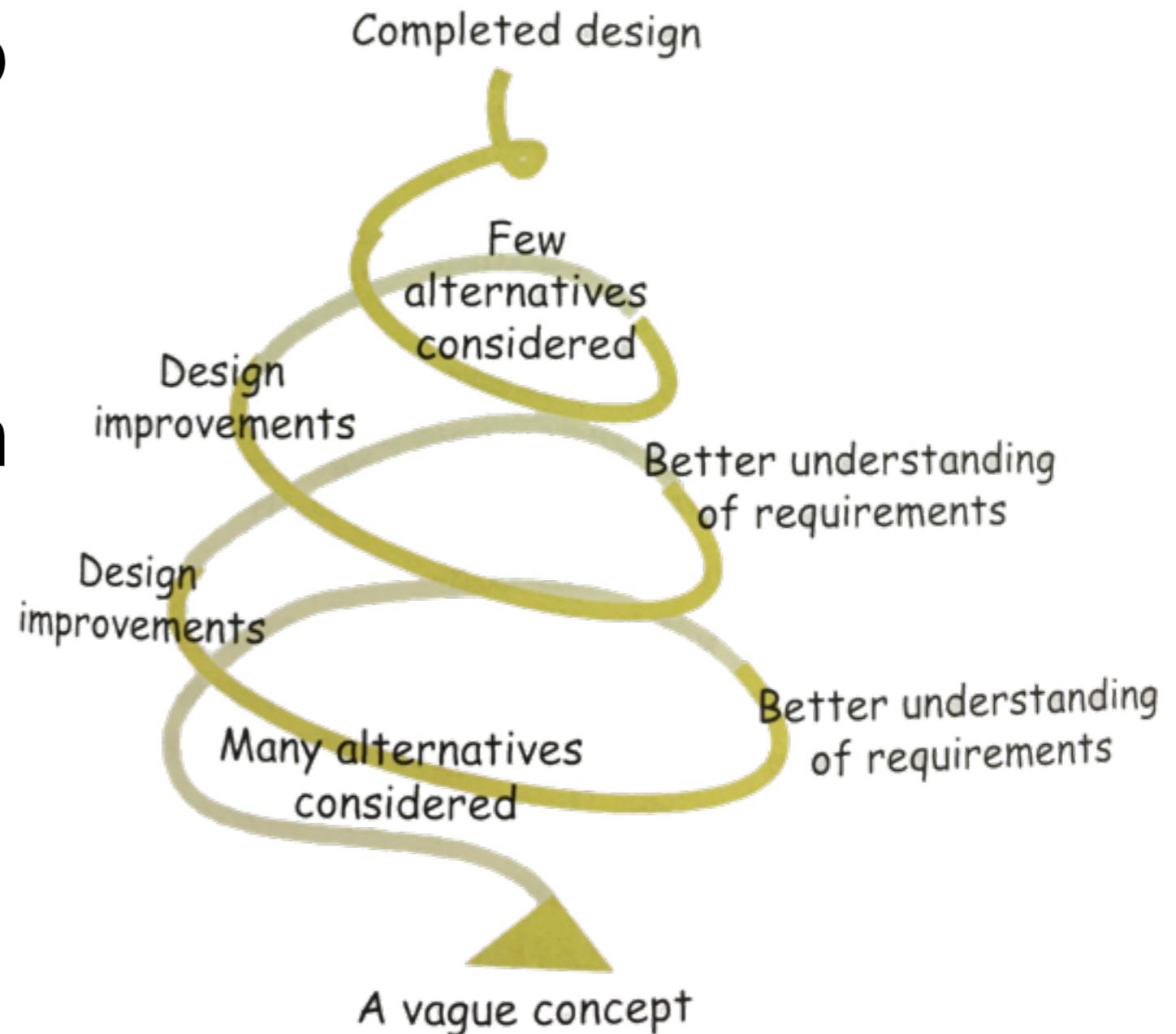
# Activity

Analyze which channels of Bertin's semiology you used for the sketches so far. Create **three** new sketches using different channels. (3 mins)

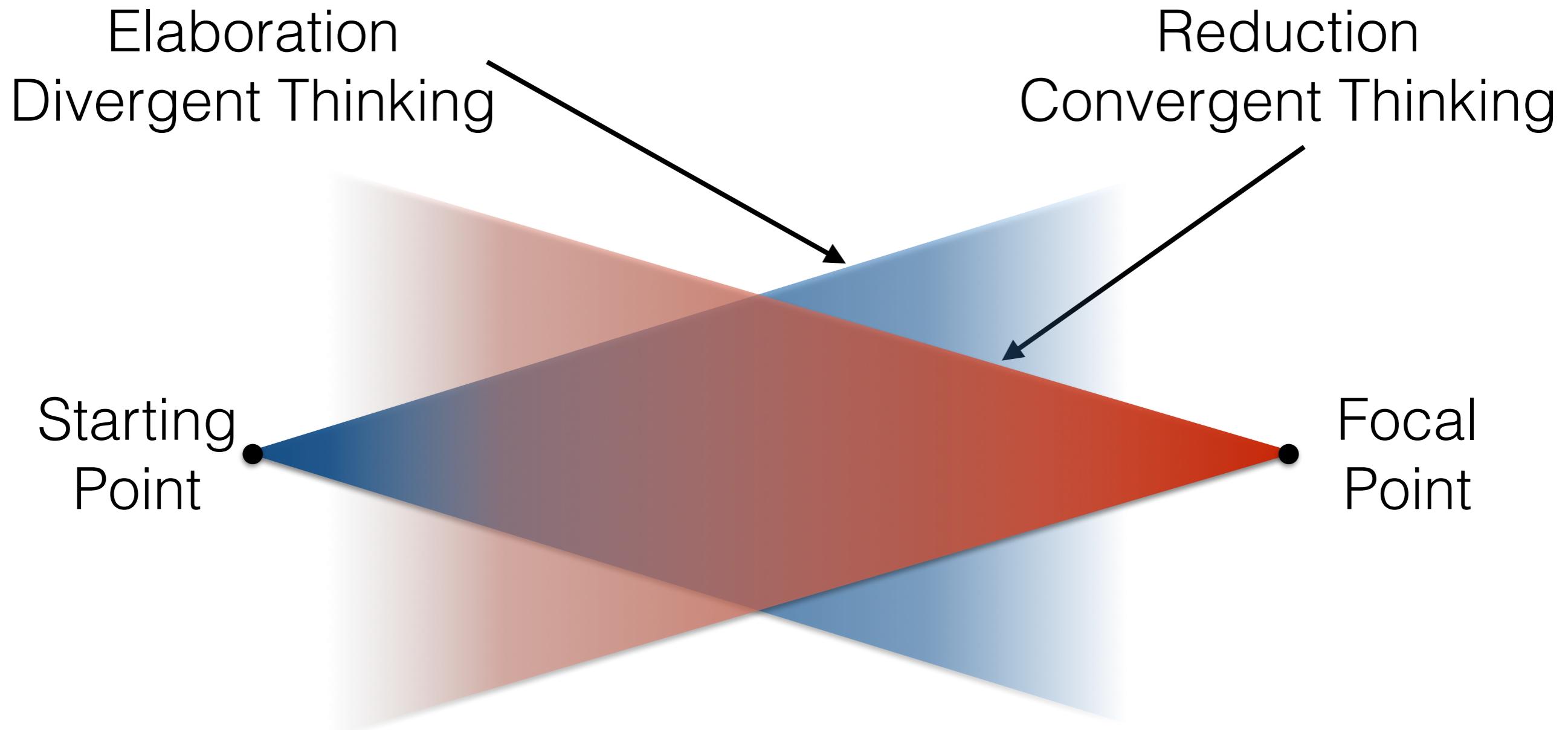
Position, Size, (Grey)Value, Texture, Color, Orientation, Shape

# Design Process

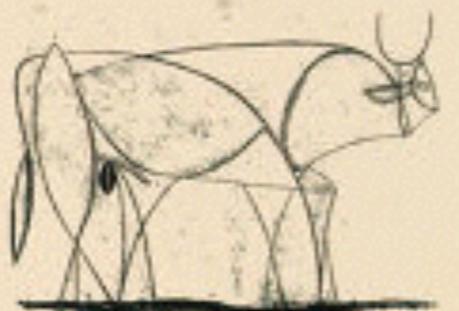
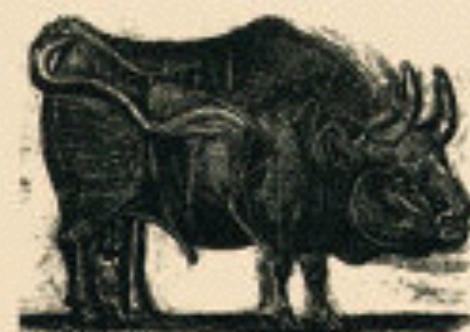
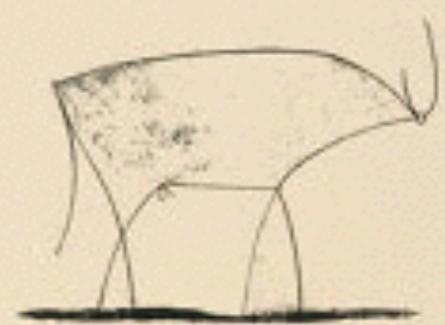
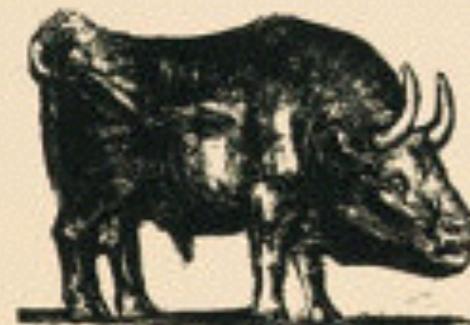
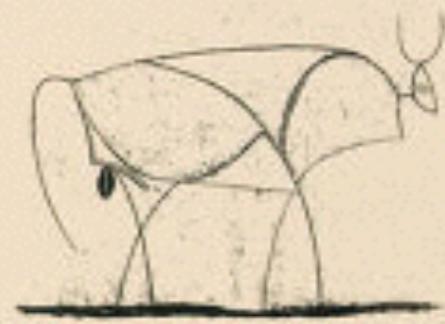
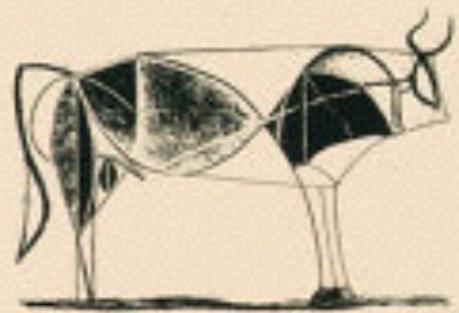
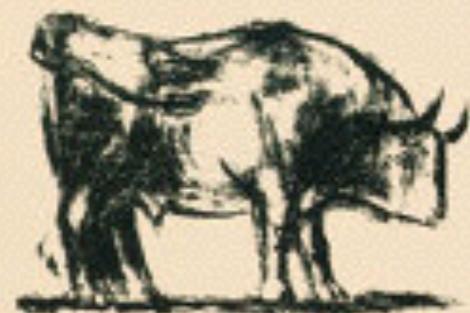
- Creative design loop  
= Iteration
- Divergent thinking:  
saturate the problem  
space
- Visual task analysis



# Design Process



"A picture used to be a sum of additions. In my case a picture is a sum of destructions."



Picasso

# Activity

In groups of students share your sketches. Identify the **top** sketch and write down the reasons why it is the best. (5 mins)



# ~45 ways to visualize two quantities

75 and 37

1. writing, number notation

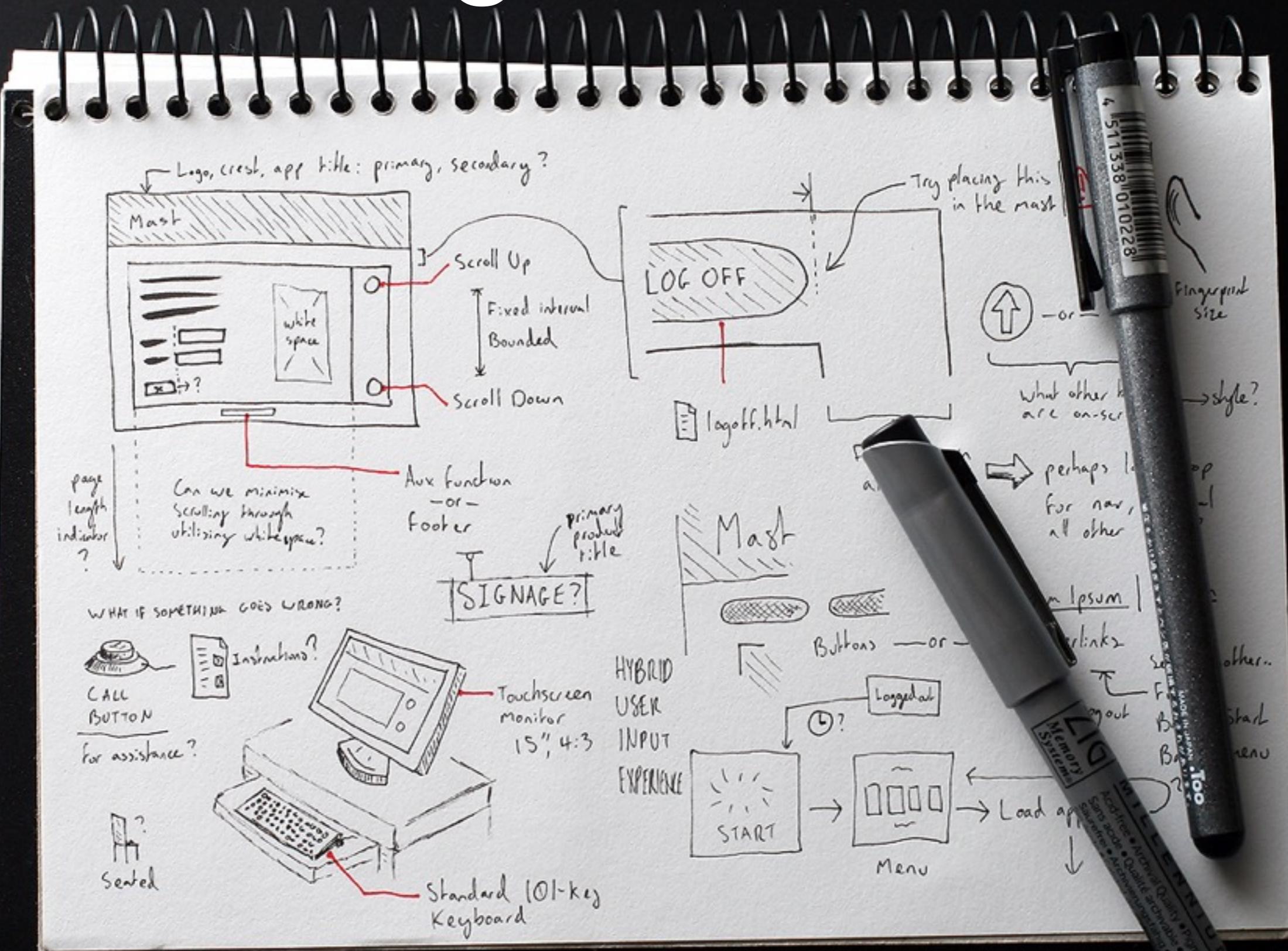
The familiar indo-arabic numerals notation in *a*, and the [babylonian numbers notation](#) in *b*.

2. squares

3. repeated icon

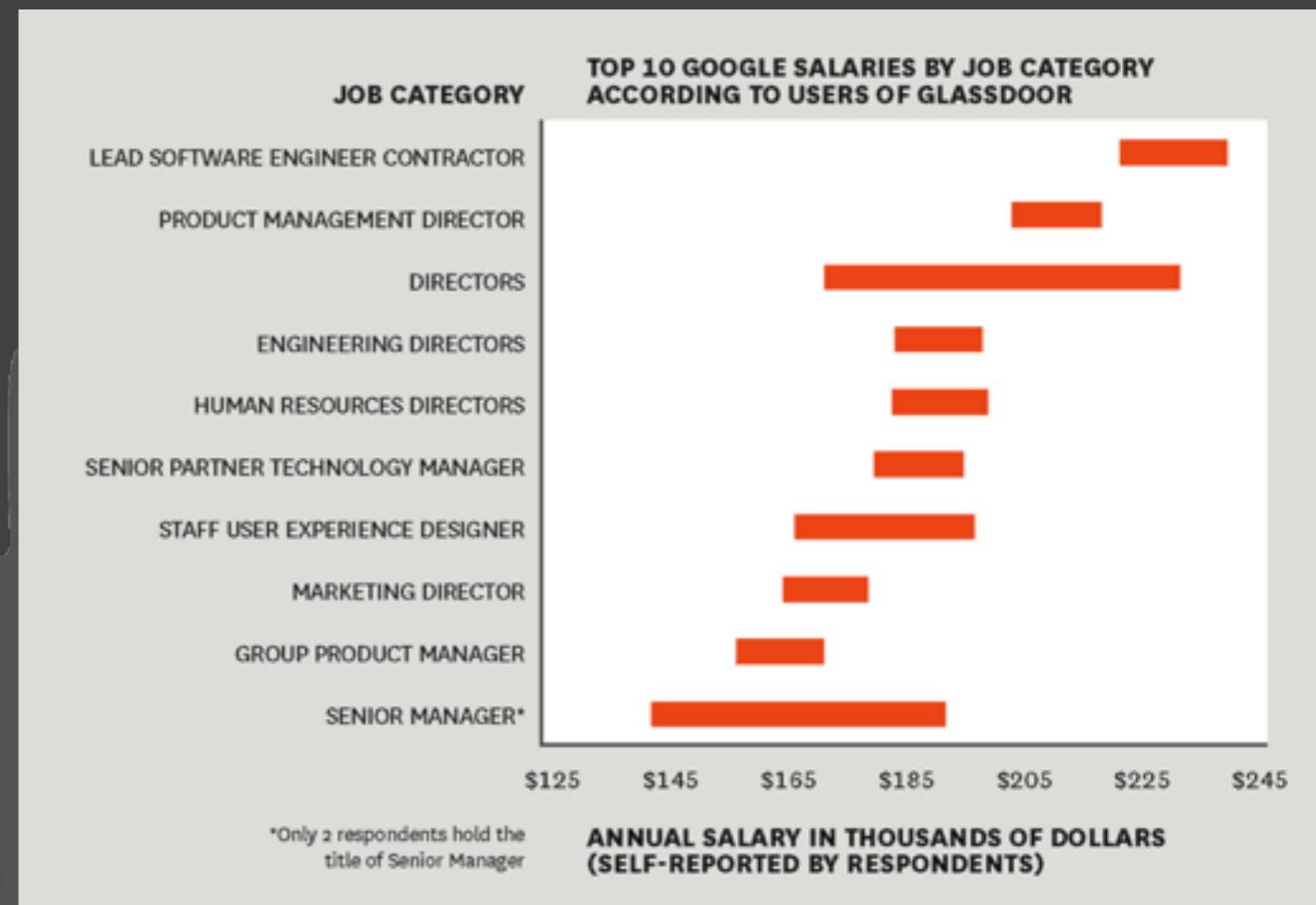
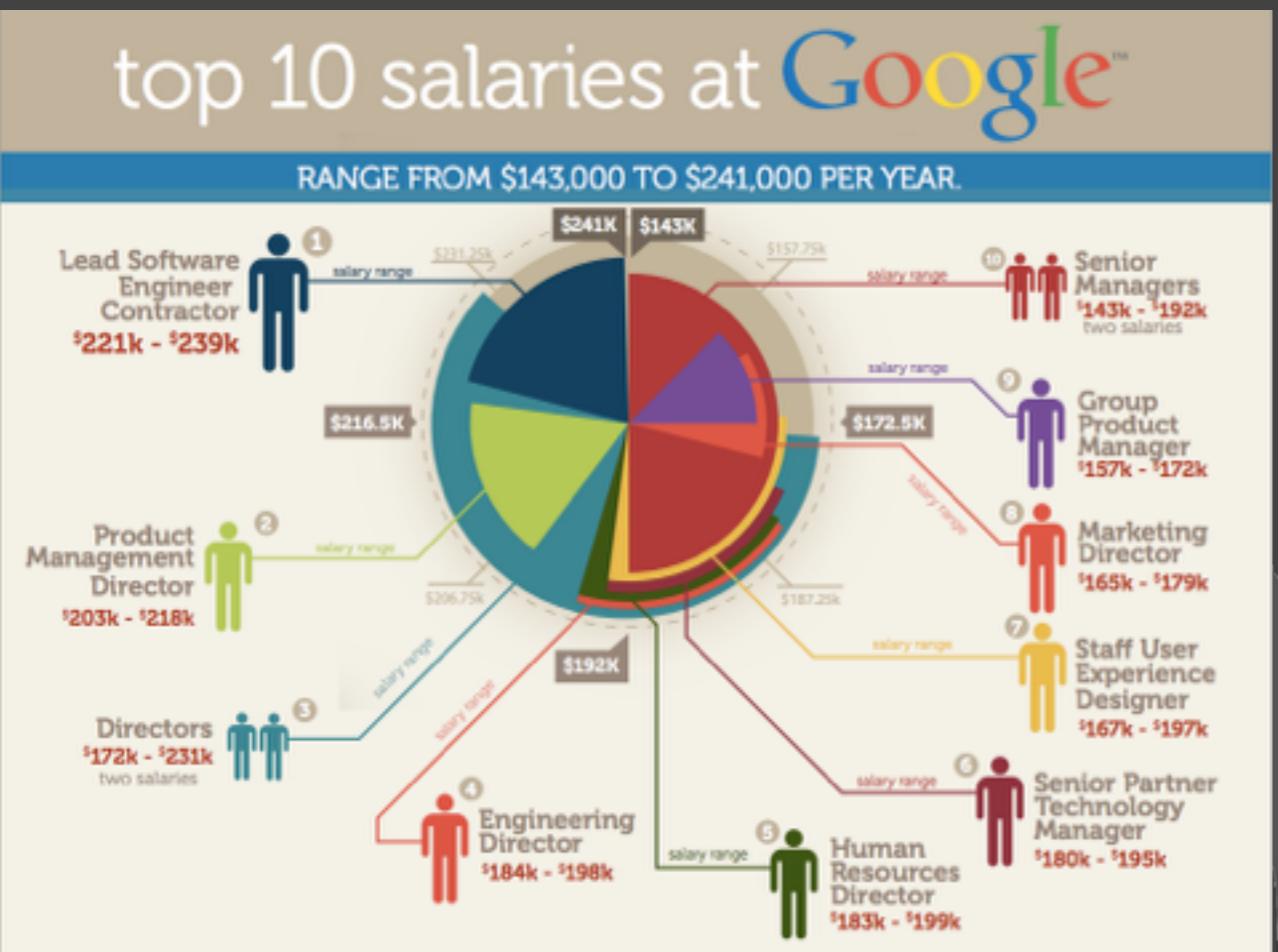
<http://blog.visual.ly/45-ways-to-communicate-two-quantities/>

# Sketching Interfaces



# Activity

Which visualization do you prefer? Why?

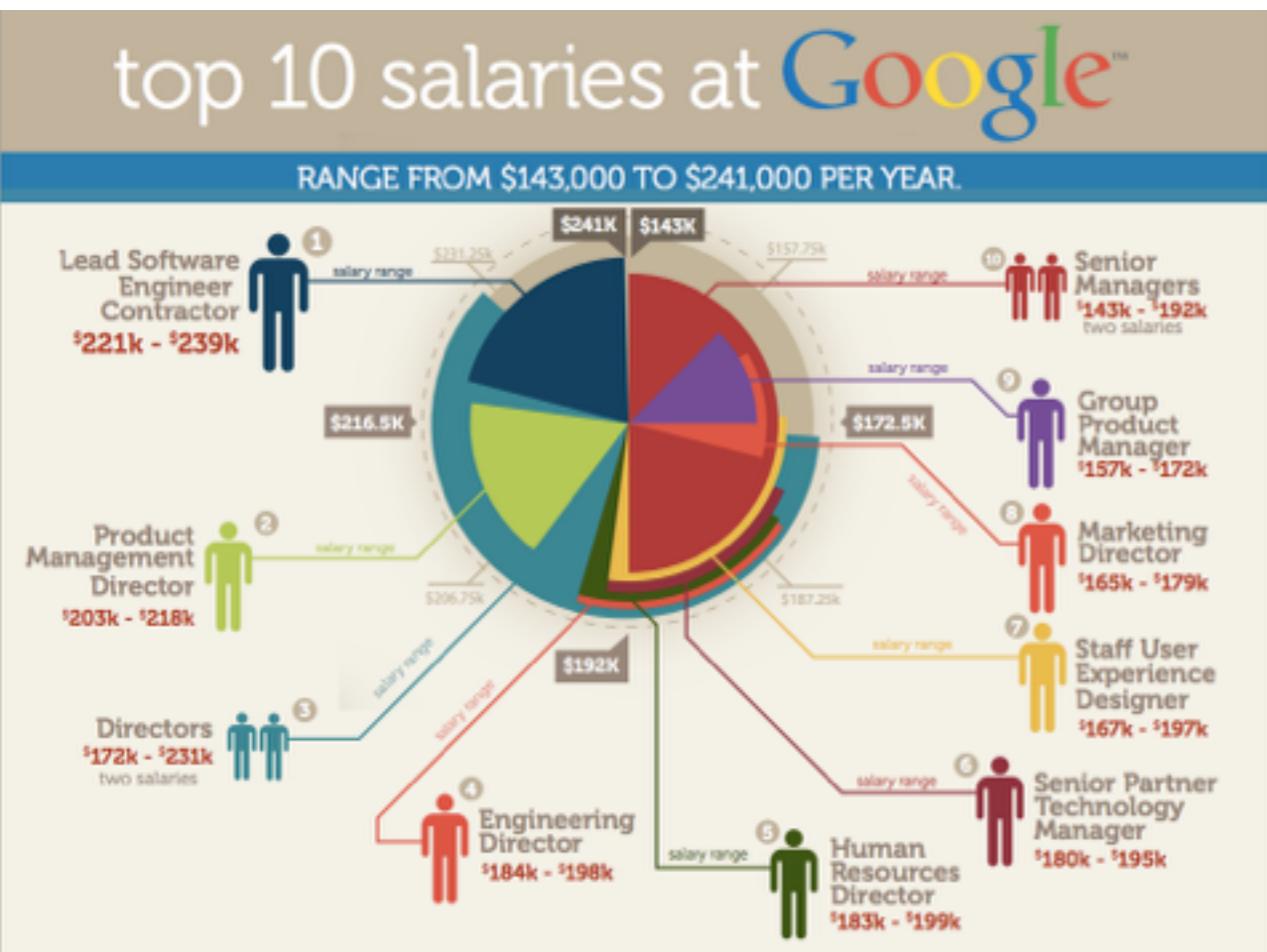


[bit.ly/top10salariesatgoogle](http://bit.ly/top10salariesatgoogle)

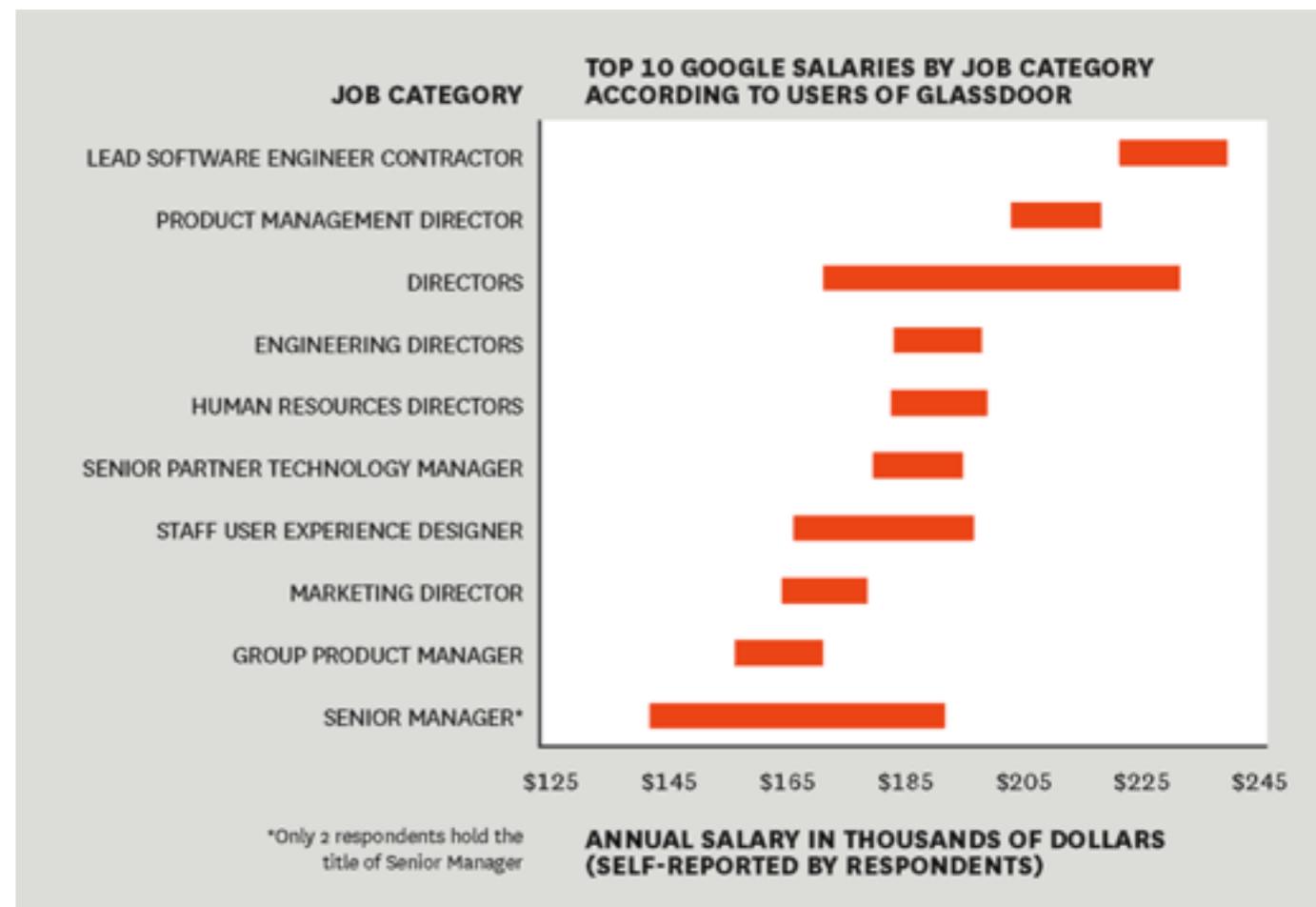
[bit.ly/cs171-salaries](http://bit.ly/cs171-salaries)

# Goals

## Communication



## Analysis



[bit.ly/top10salariesatgoogle](http://bit.ly/top10salariesatgoogle)

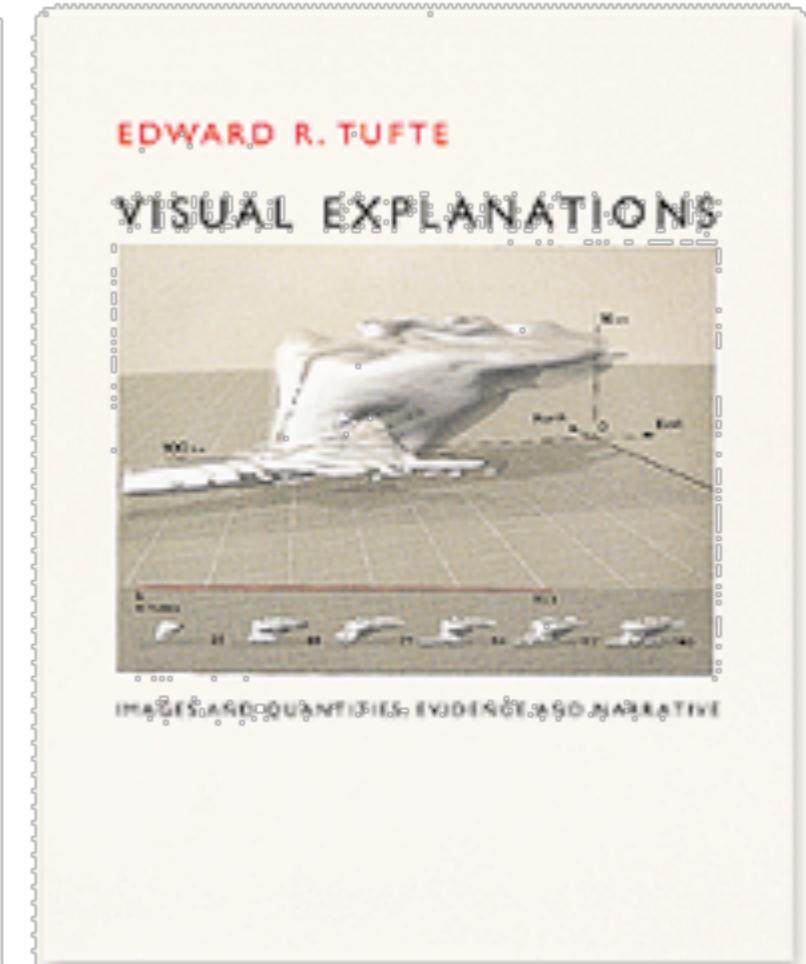
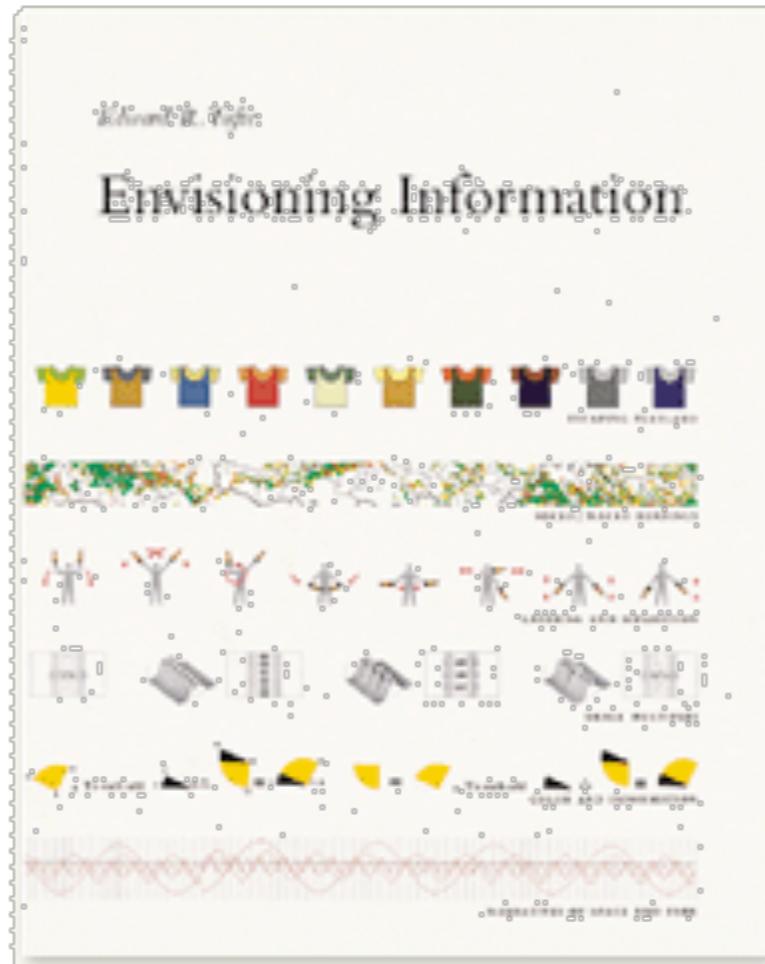
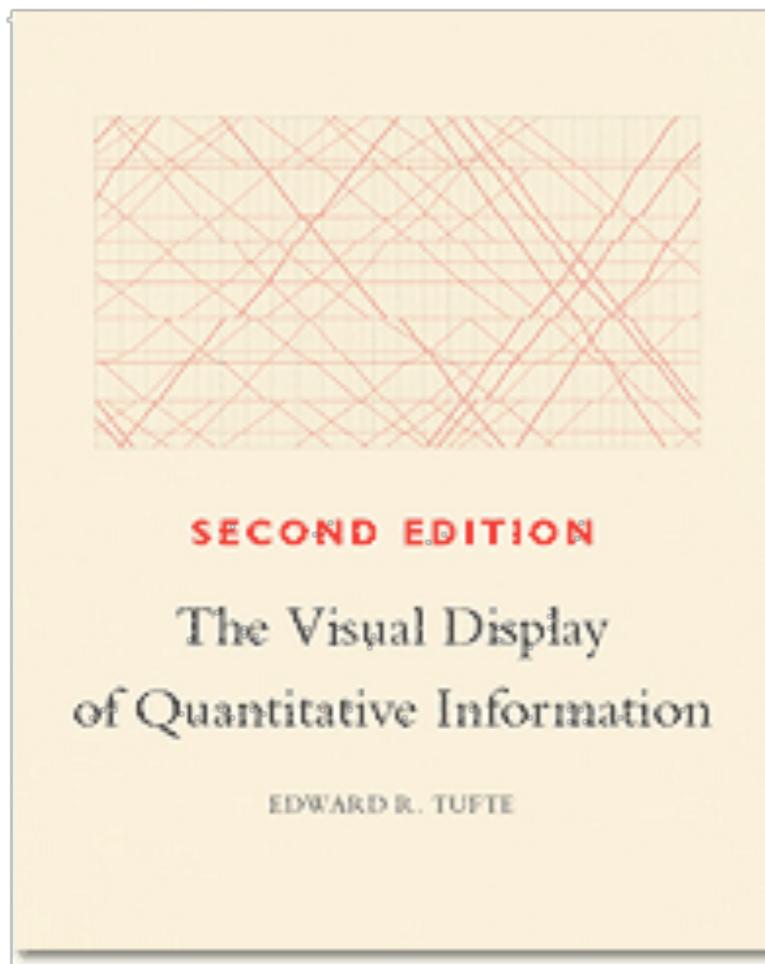
[bit.ly/cs171-salaries](http://bit.ly/cs171-salaries)



“Clutter and confusion are not attributes of information, they are failures of design.”

– Edward Tufte

# Edward Tufte



# Graphical Integrity

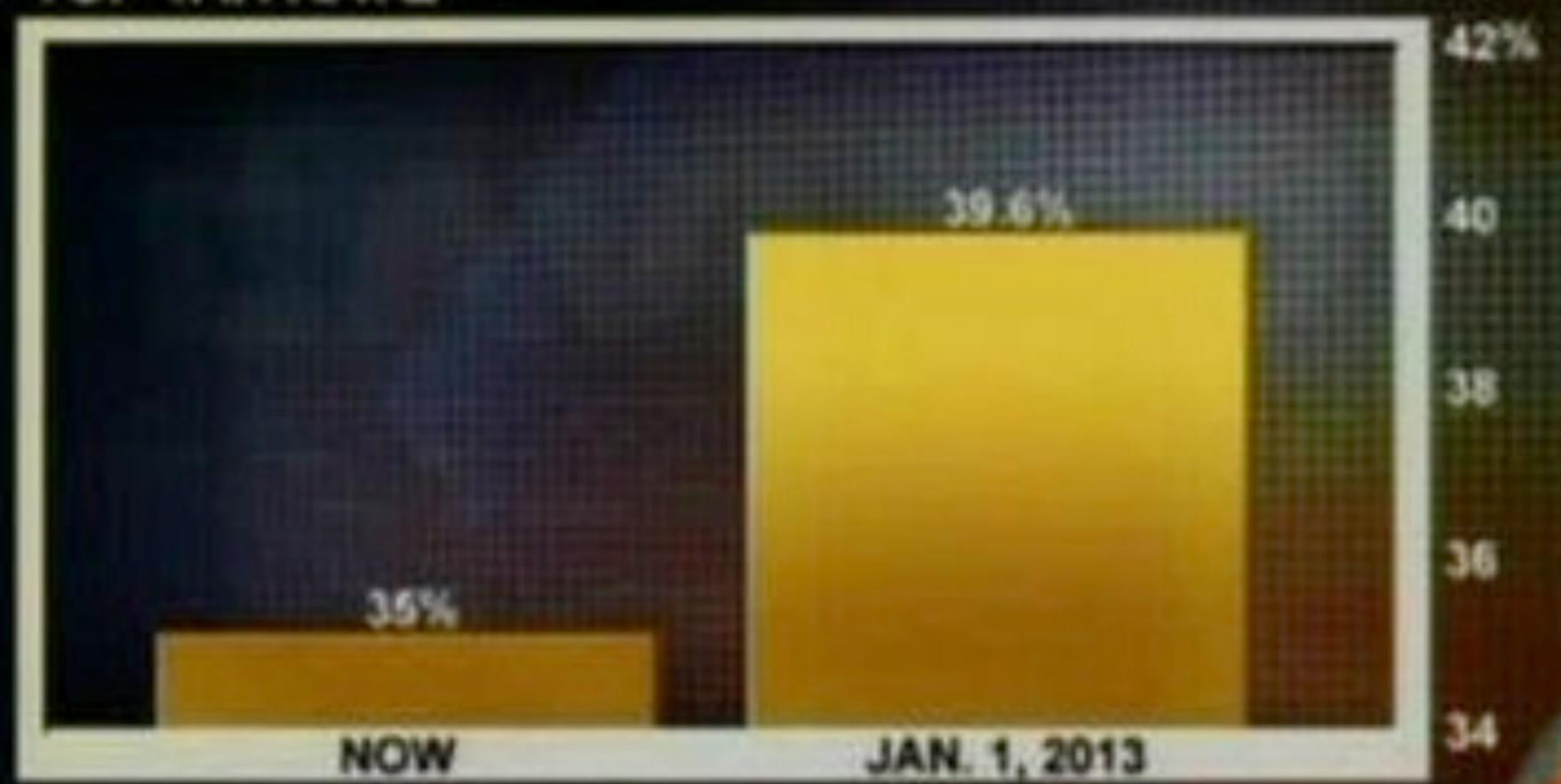
CS

171



# IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



8:01 p ET



TOP STORIES

TECHNOLOGY

CONSUMER

WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

DOW 13008.68 ▲ 64.33

S&P 1379.32 ▲ 5.98

NASDAQ 2939.52 ▲ 6.32

## If Bush tax cuts expire...

Top tax rate

40%

39.6%

35.0%

30%

20%

10%

0%

Now

Jan. 1, 2013

## JOB LOSS BY QUARTER

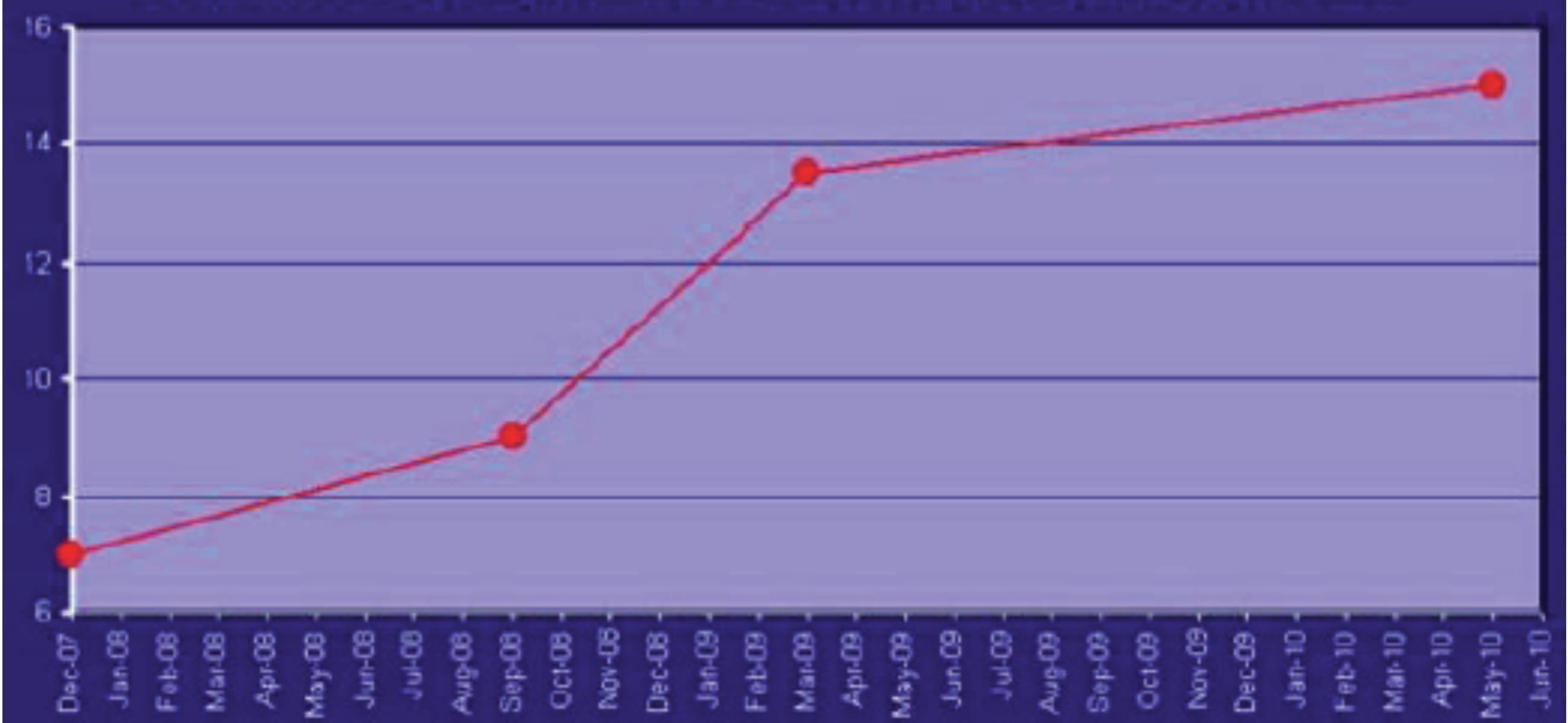


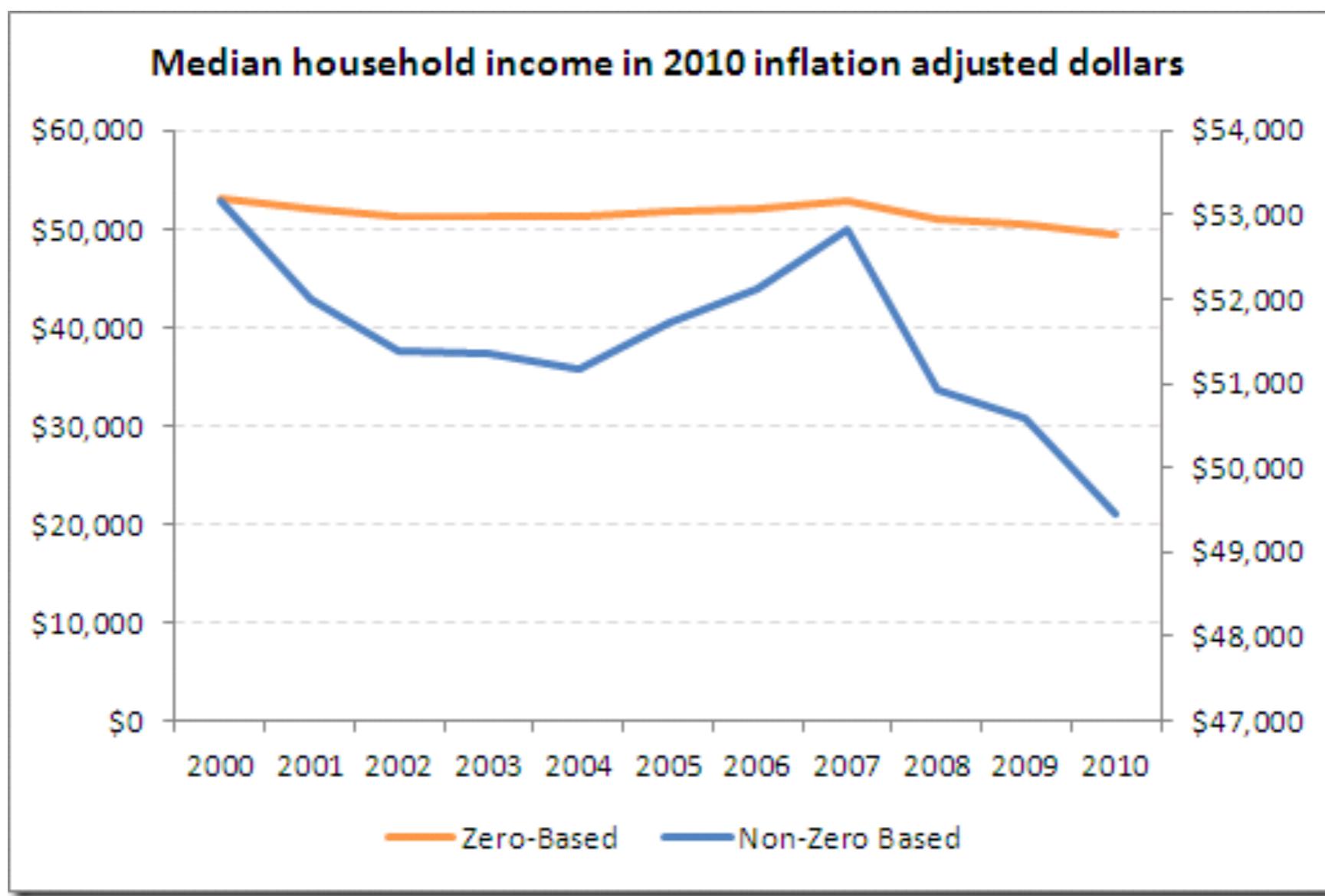
FOX NEWS .com

SOURCE: BLS

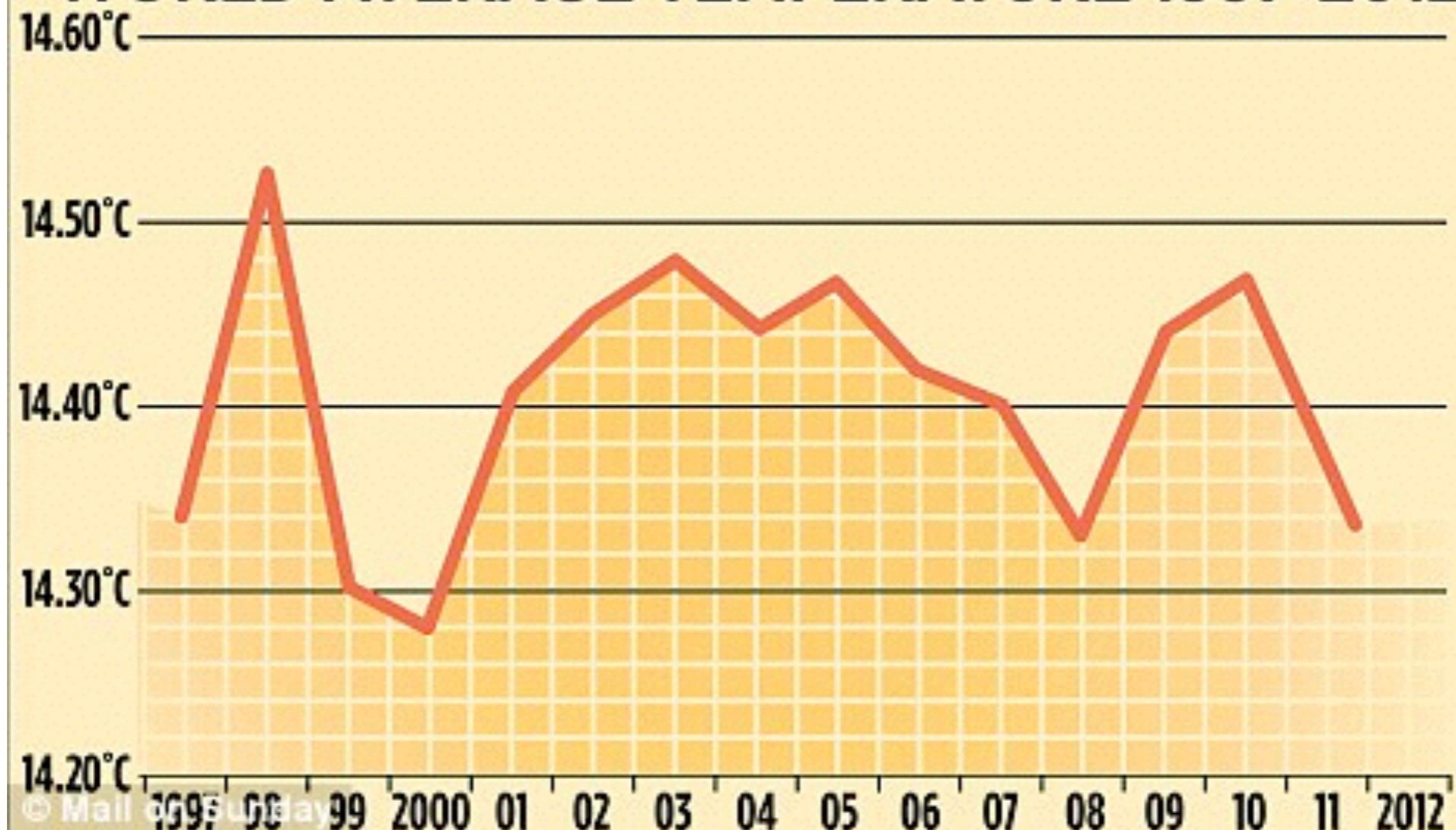
AMERICA'S  
NEWSROOM

## UNEMPLOYMENT LEVEL BY RANDOM QUARTER

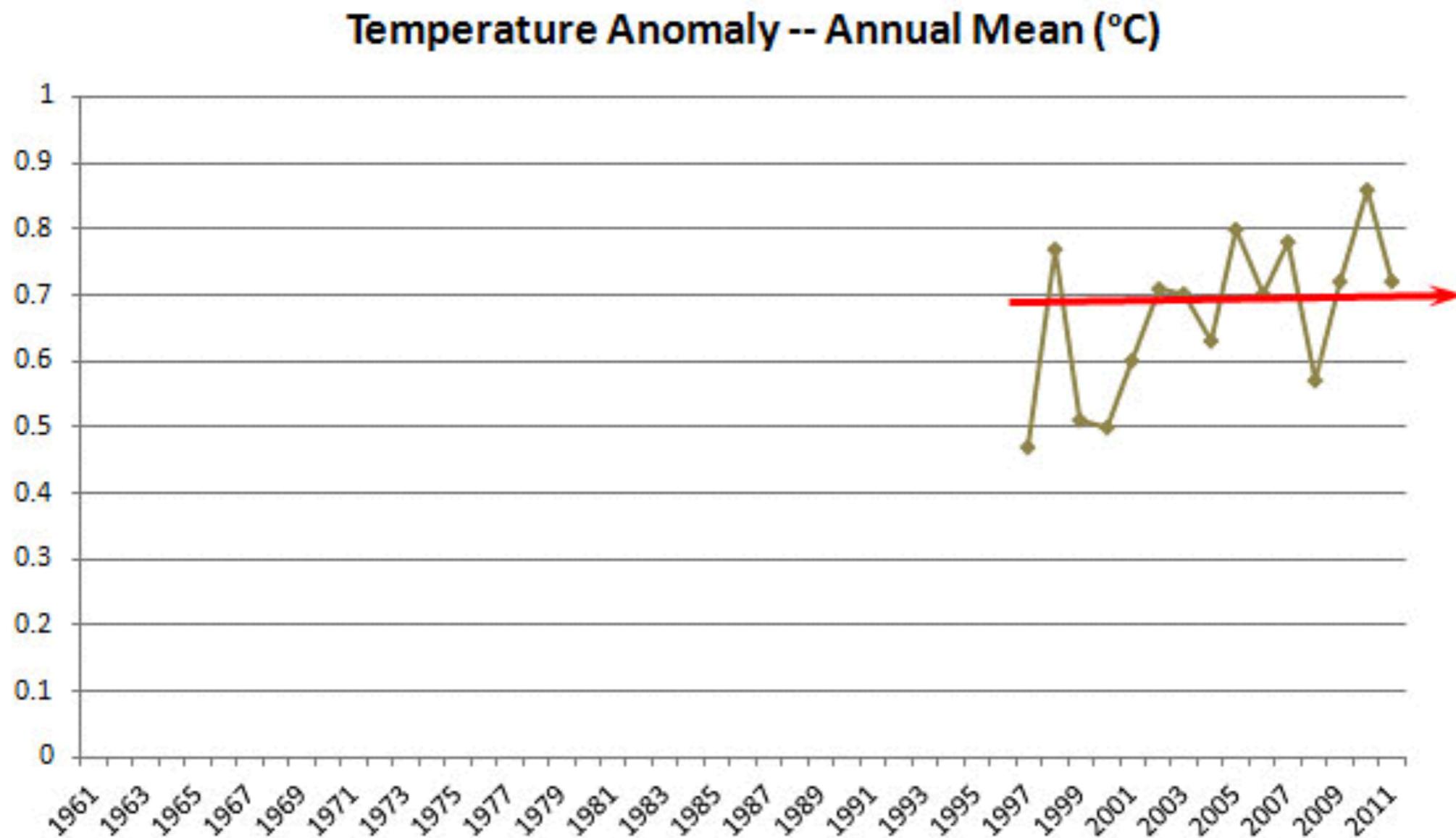




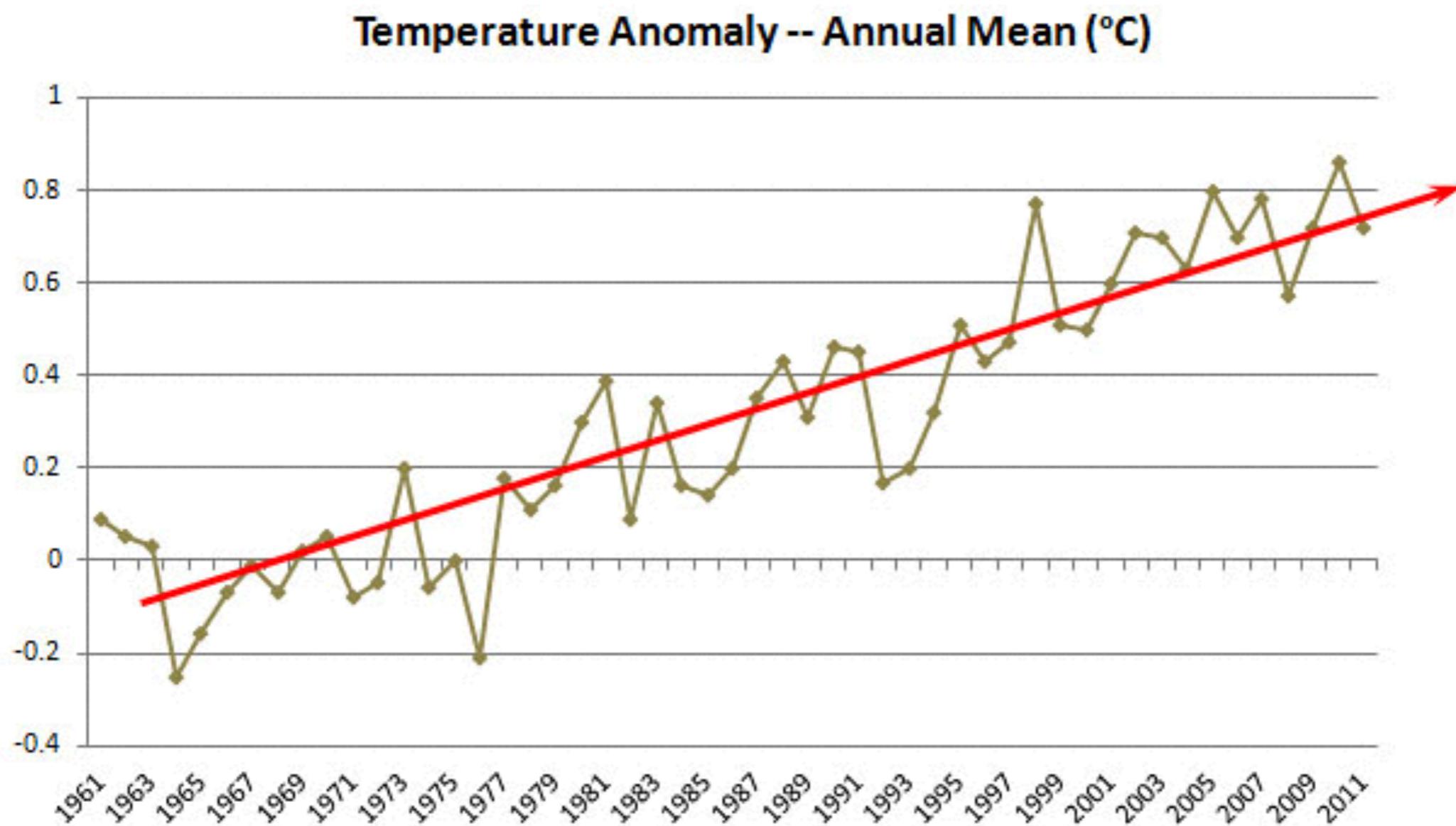
## WORLD AVERAGE TEMPERATURE 1997-2012



# Global Warming?



# Global Warming!



# The Lie Factor

CS

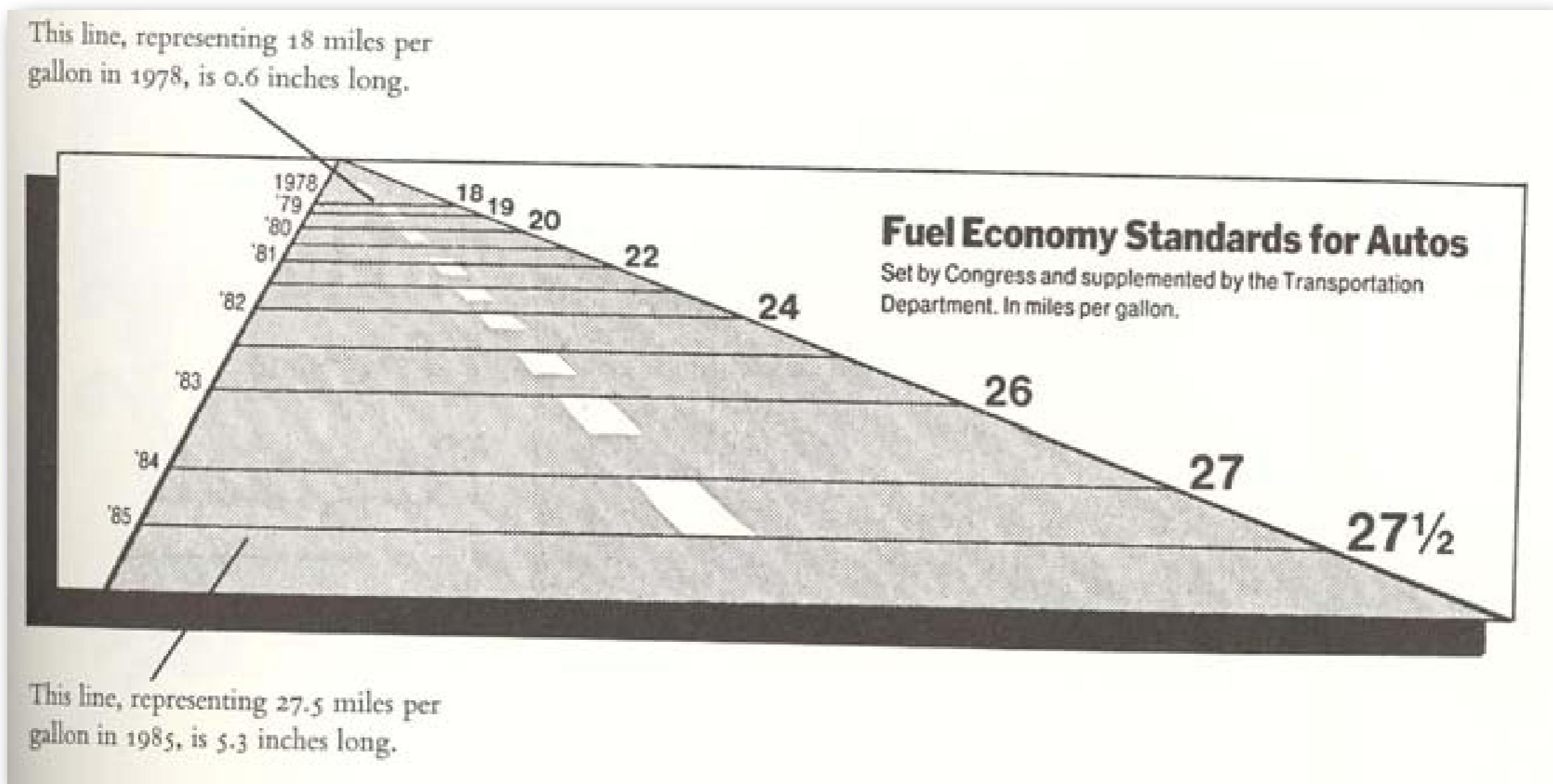
171



# The Lie Factor

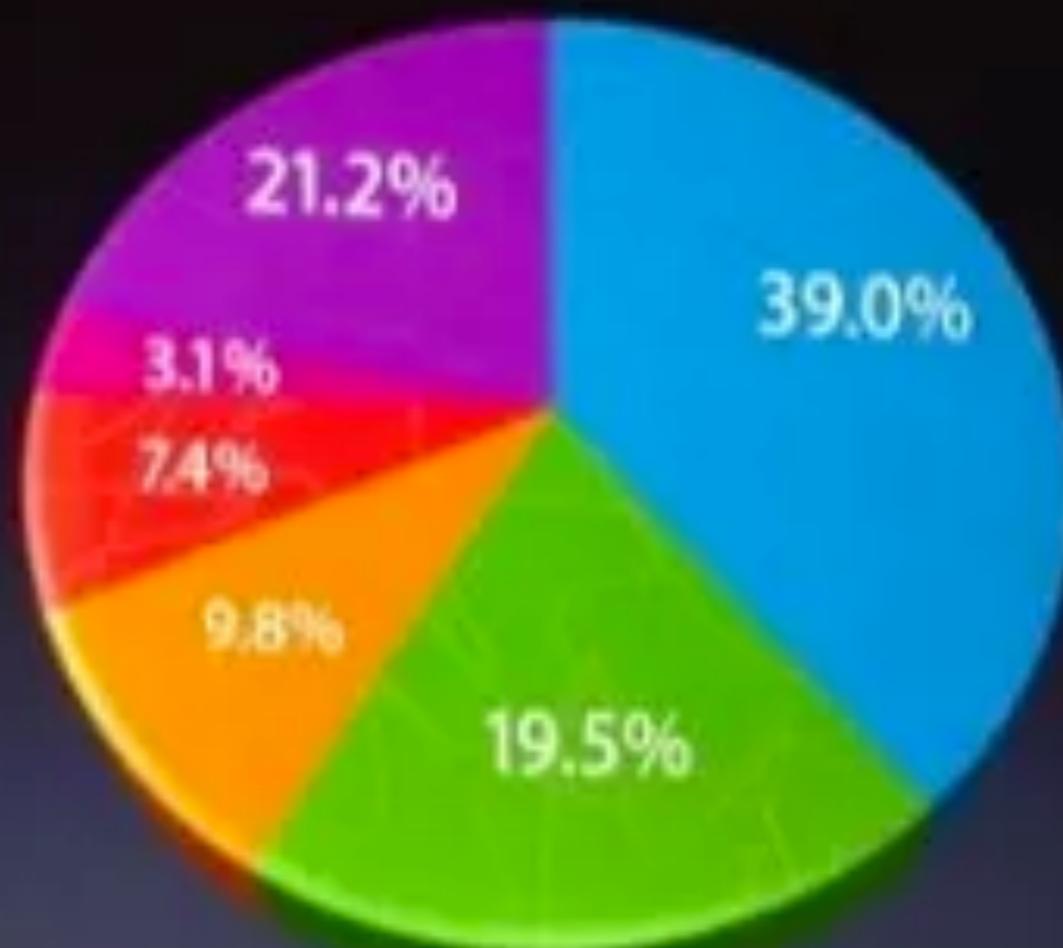
Size of effect shown in graphic

Size of effect in data



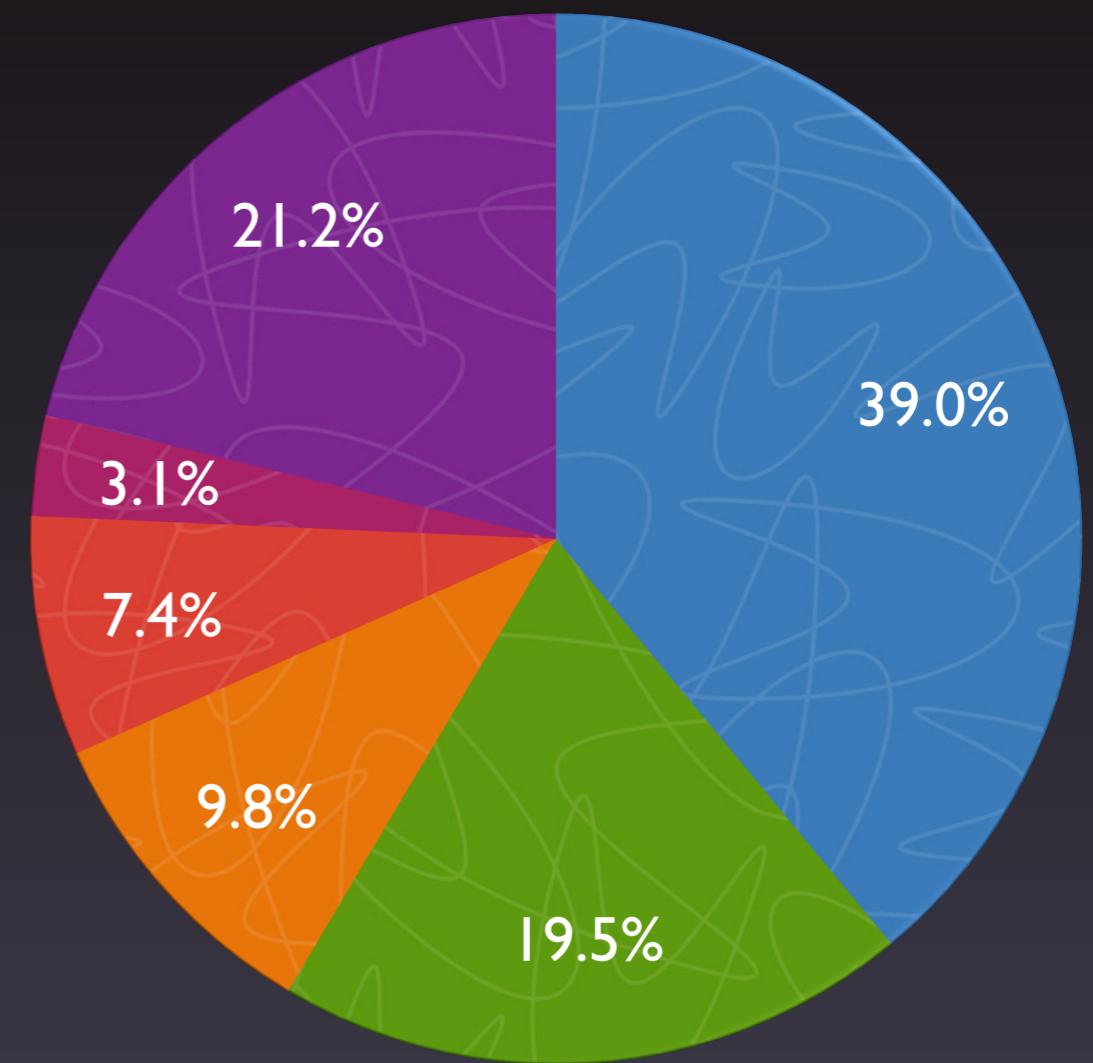
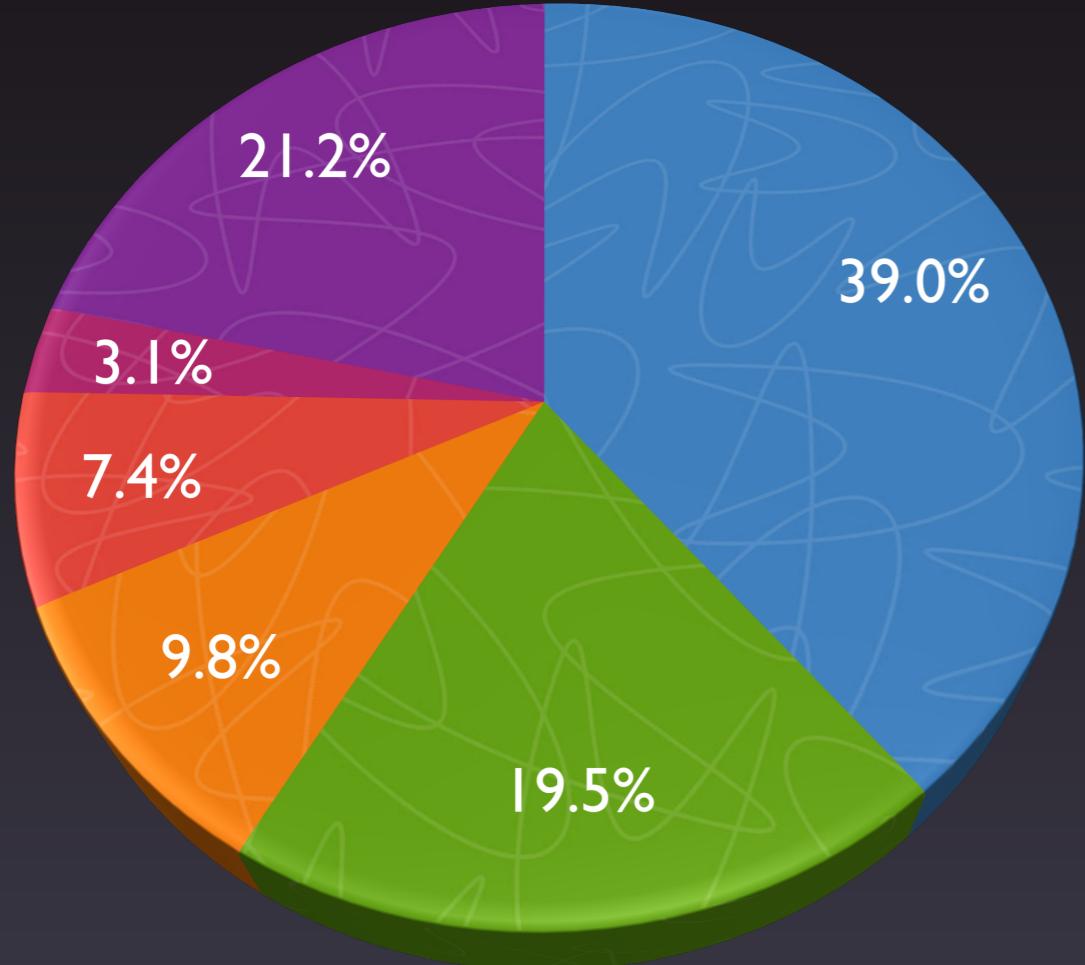
# U.S. SmartPhone Marketshare

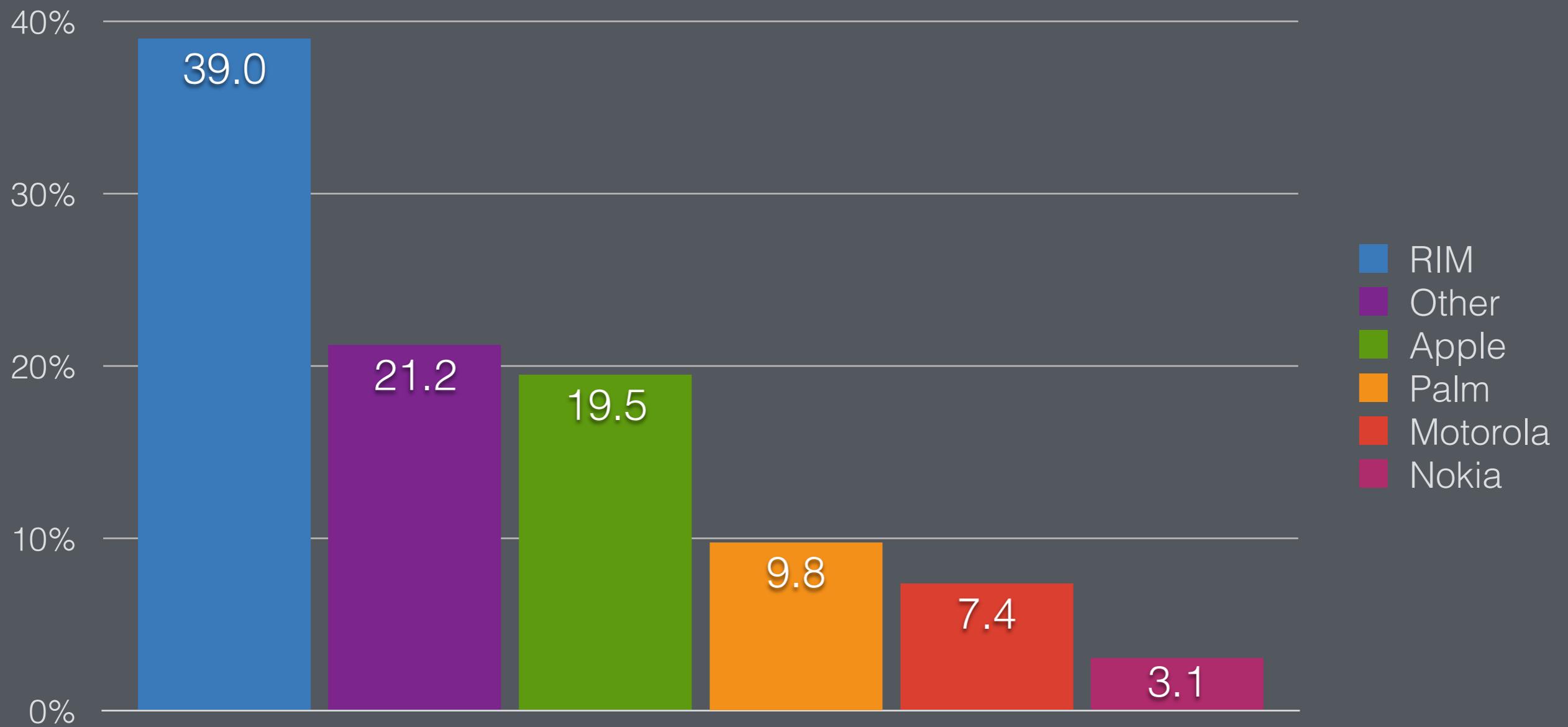
- RIM
- Apple
- Palm
- Motorola
- Nokia
- Other



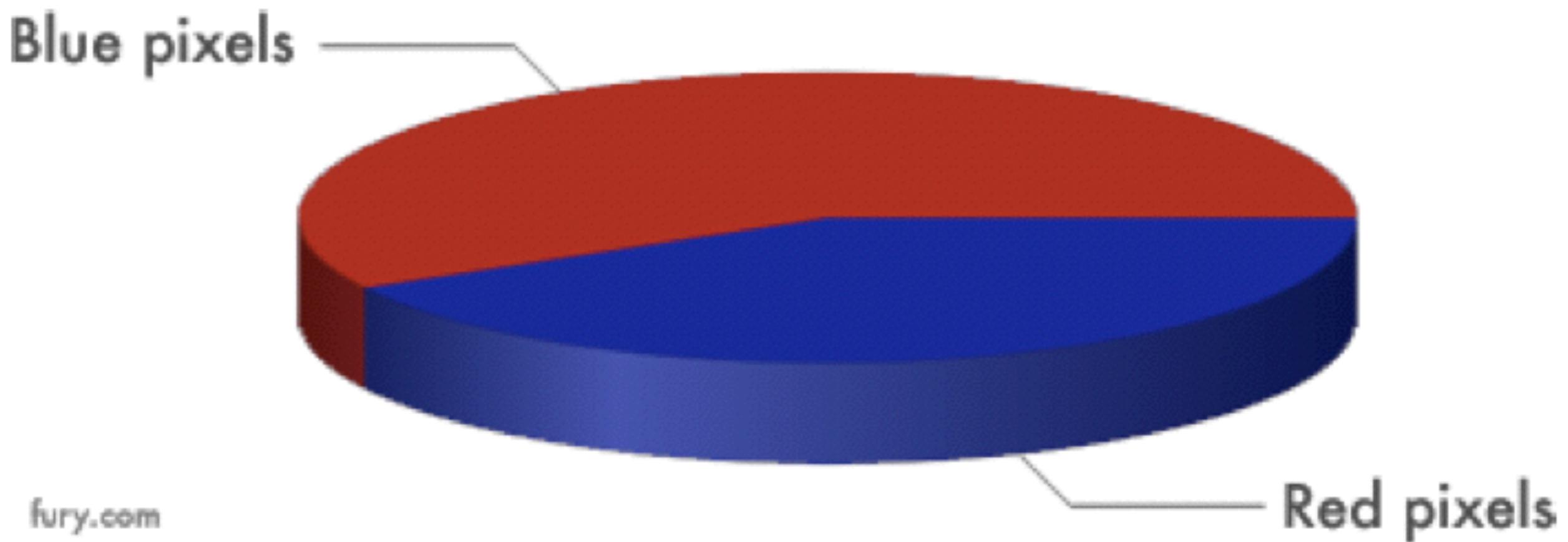
Gartner fo

- RIM
- Apple
- Palm
- Motorola
- Nokia
- Other





# Why 3D Pie Charts are Bad



# Maximize Data-Ink Ratio

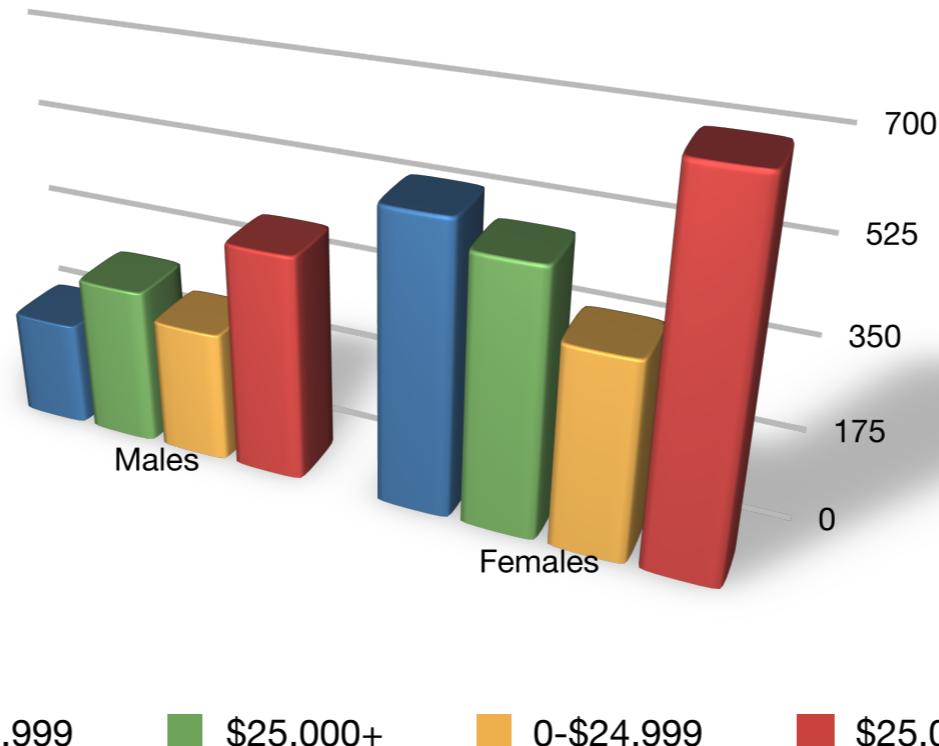
CS

171



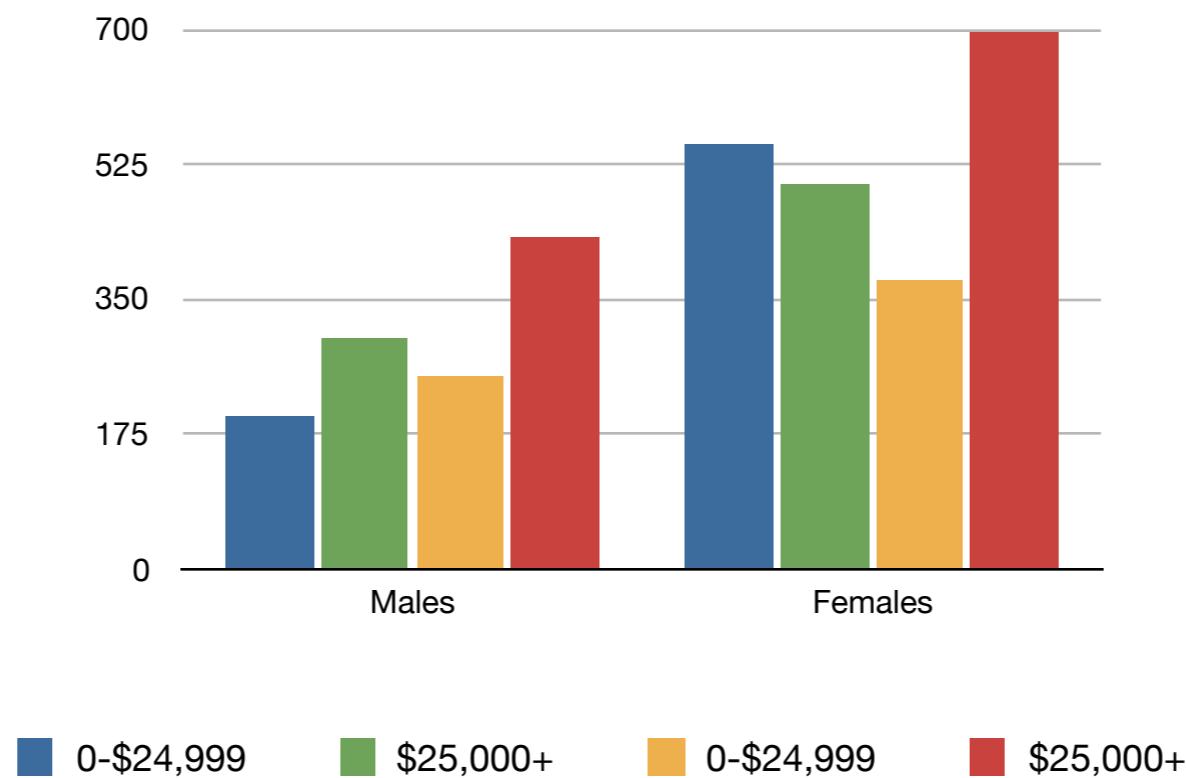
# Maximize Data-Ink Ratio

Data-Ink Ratio =  $\frac{\text{Data ink}}{\text{Total ink used in graphic}}$



# Maximize Data-Ink Ratio

Data-Ink Ratio =  $\frac{\text{Data ink}}{\text{Total ink used in graphic}}$



# Avoid Chartjunk

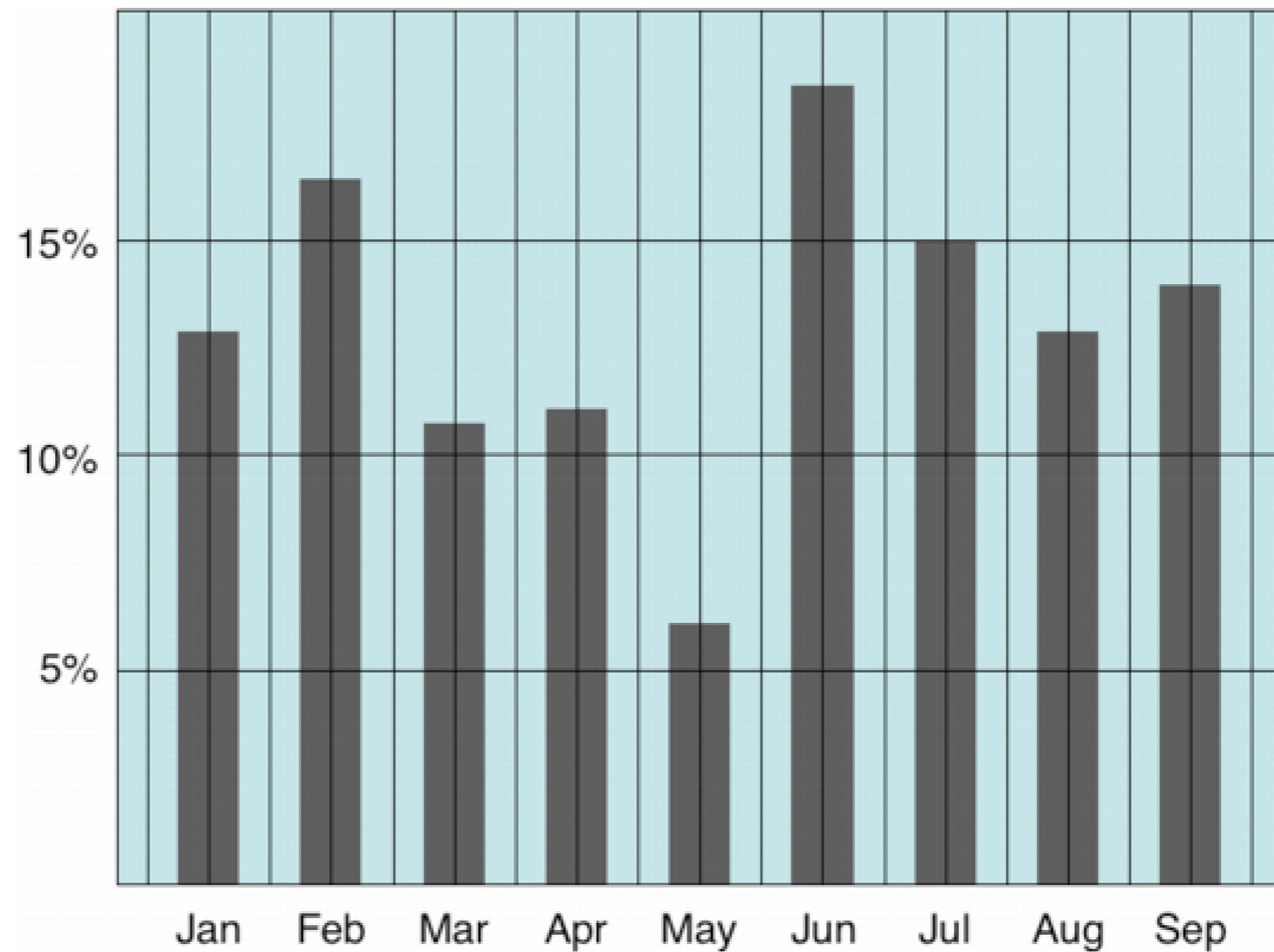
CS

171

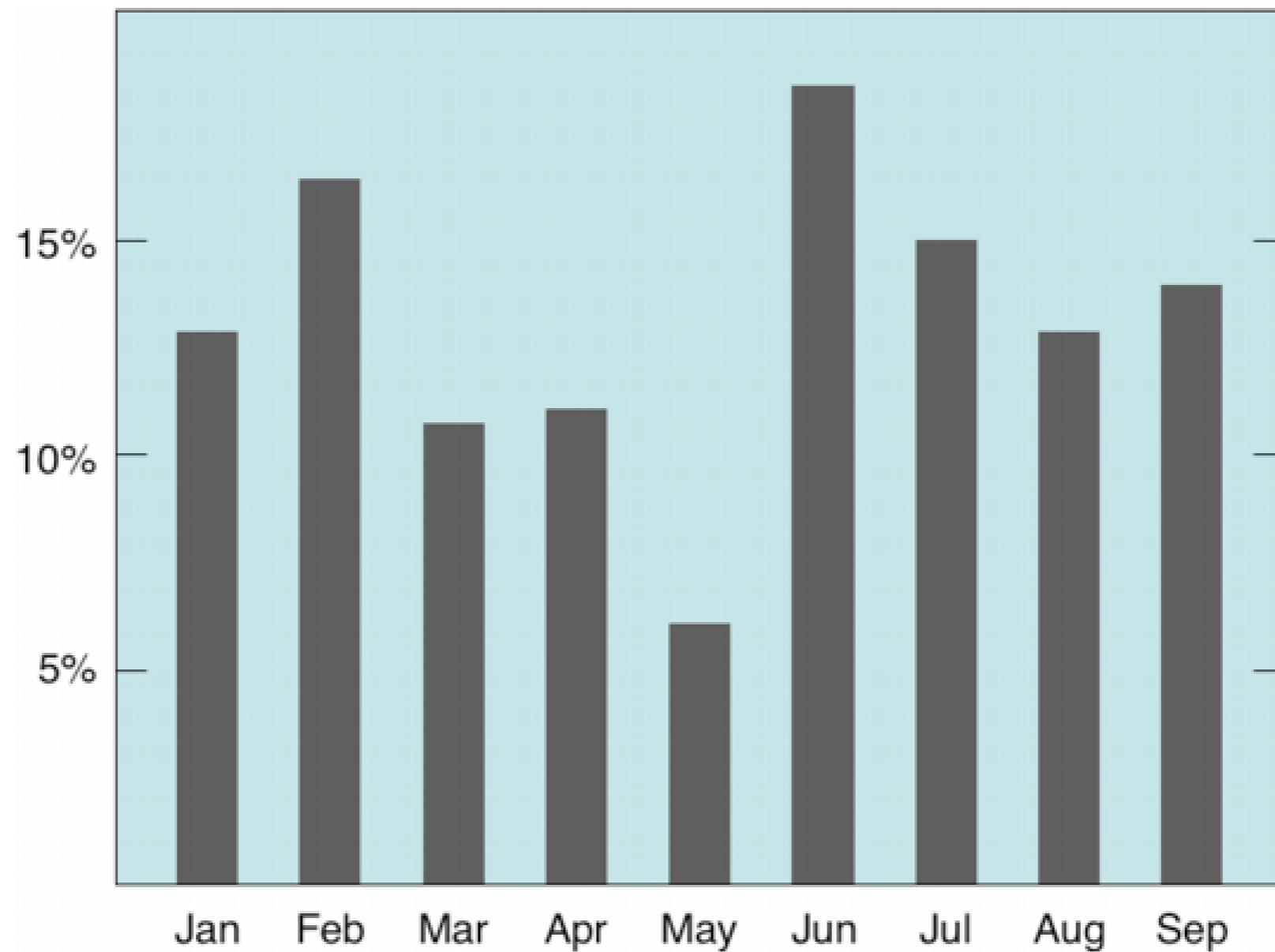


# Avoid Chartjunk

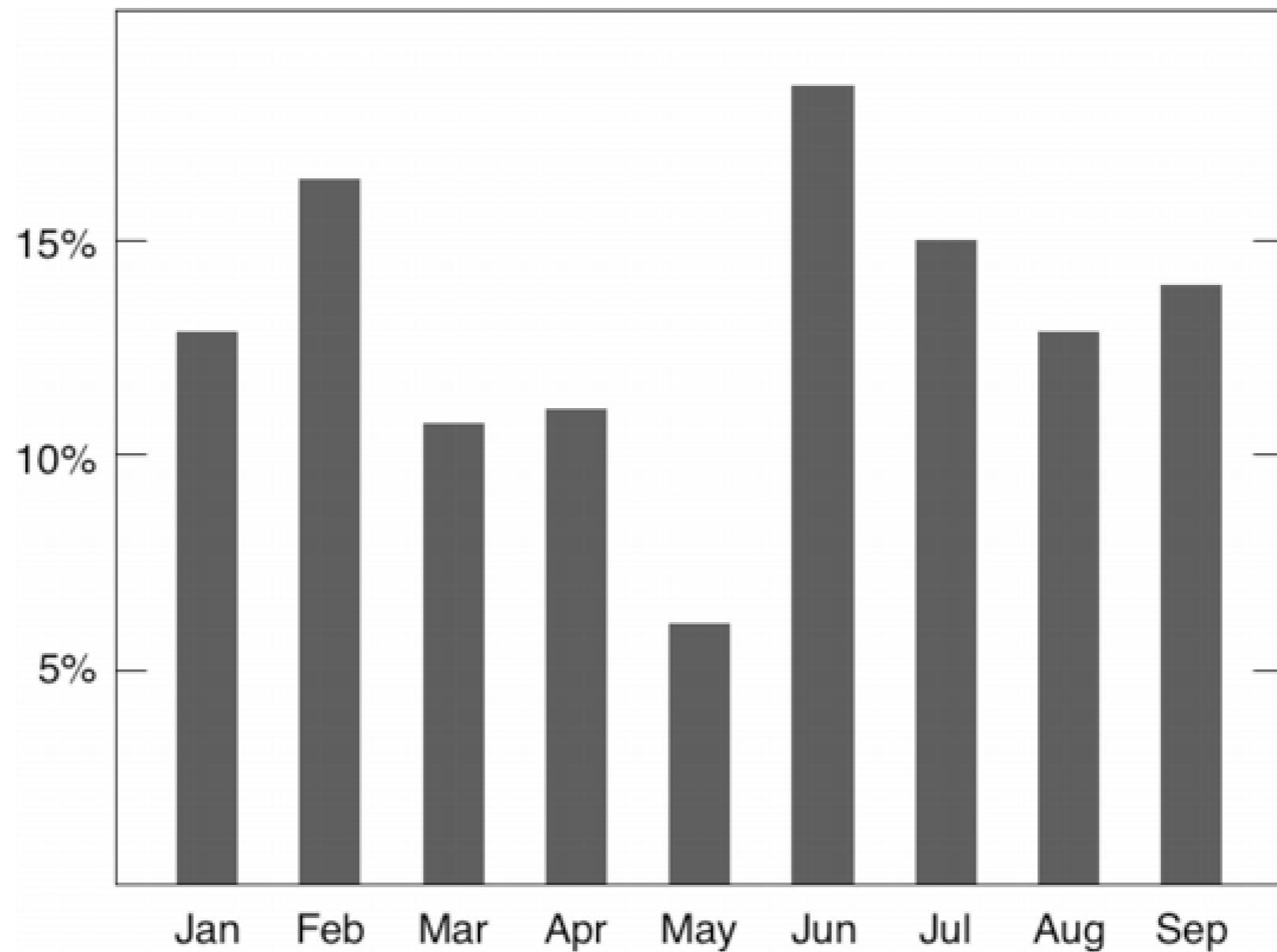
Extraneous visual elements that distract from the message



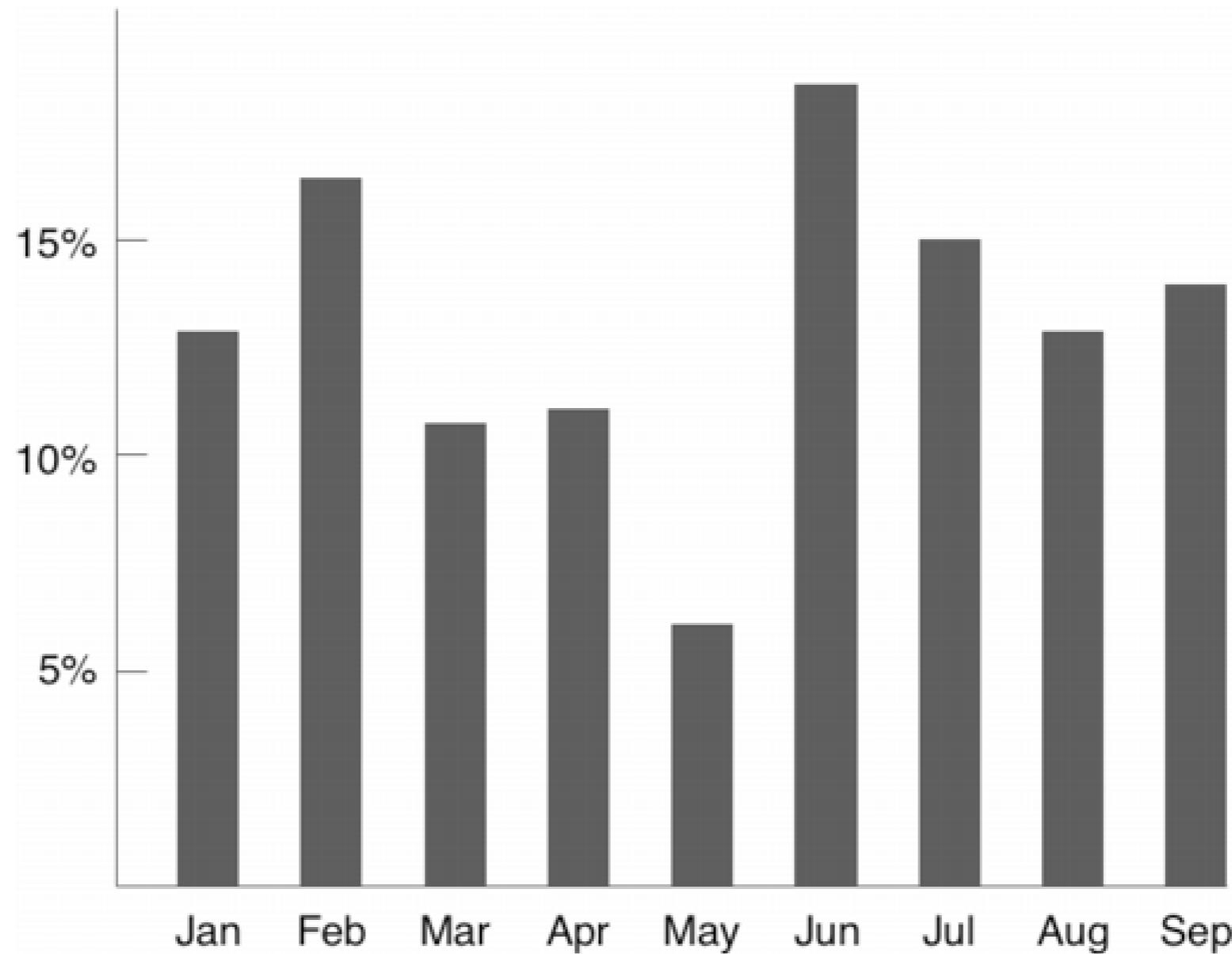
# Avoid Chartjunk



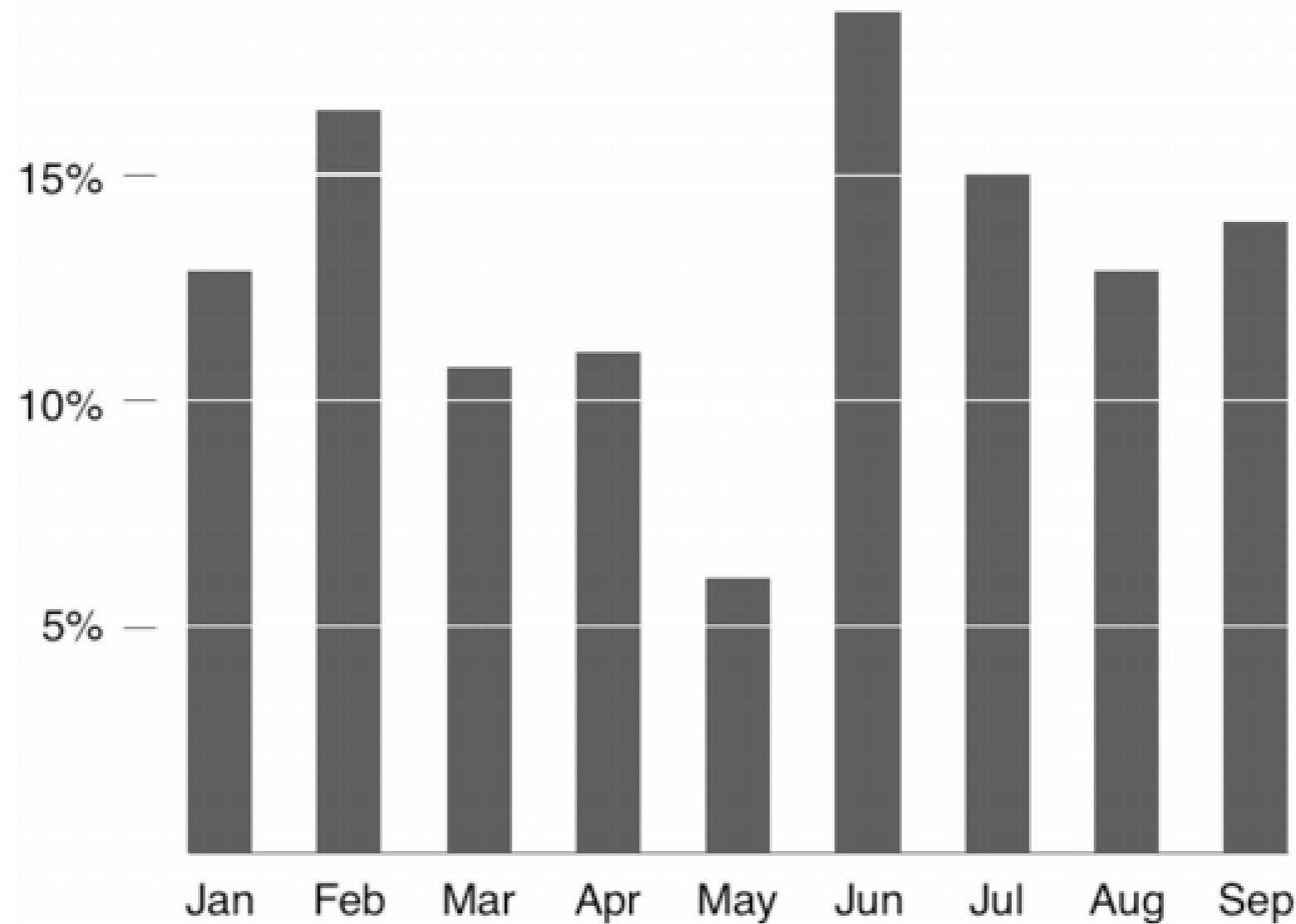
# Avoid Chartjunk



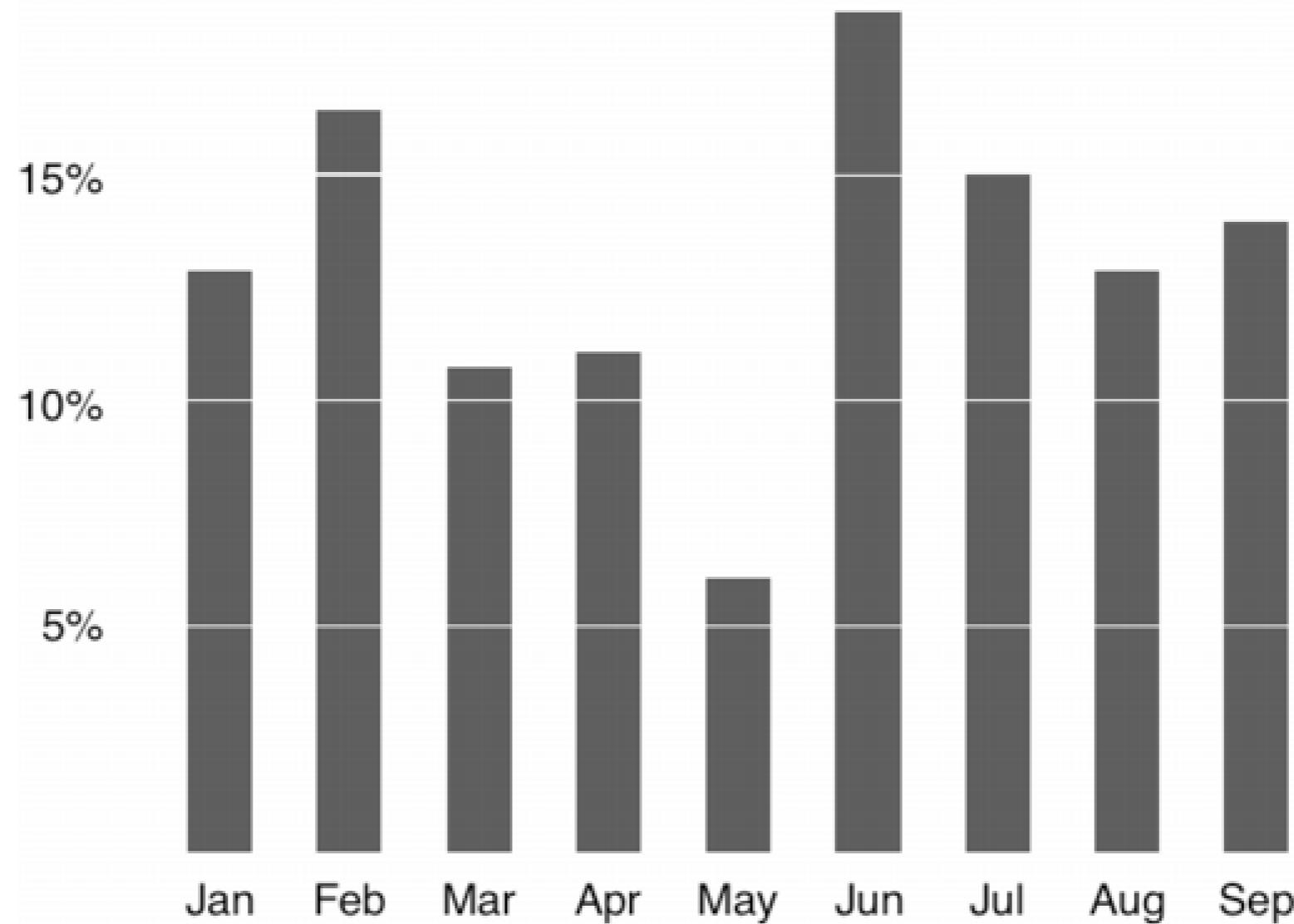
# Avoid Chartjunk



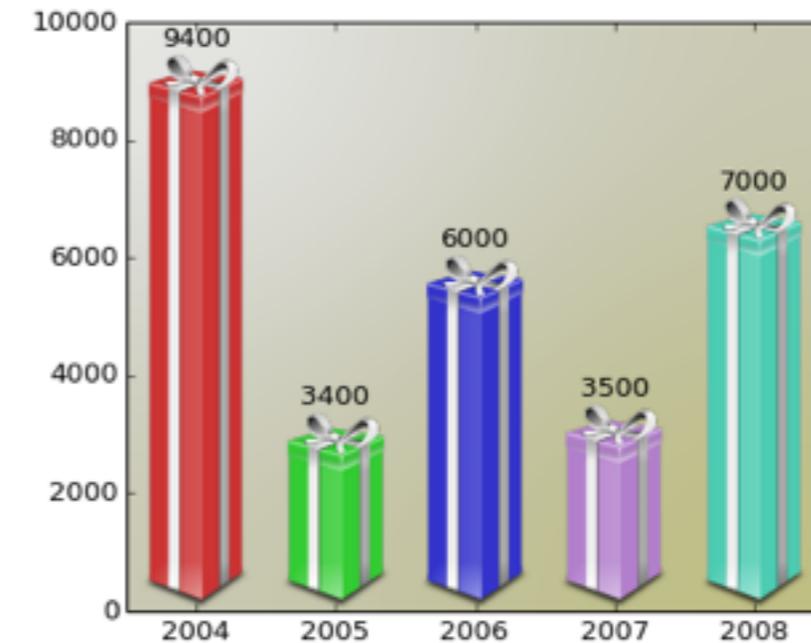
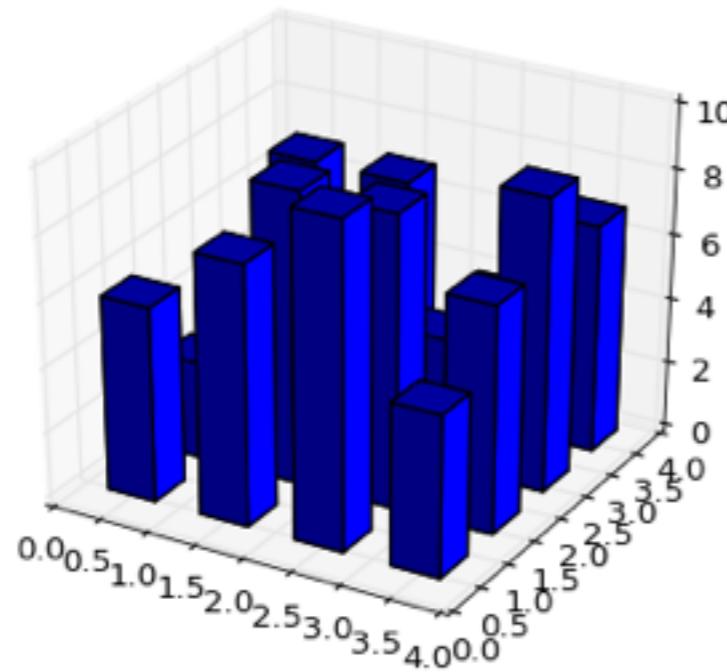
# Avoid Chartjunk



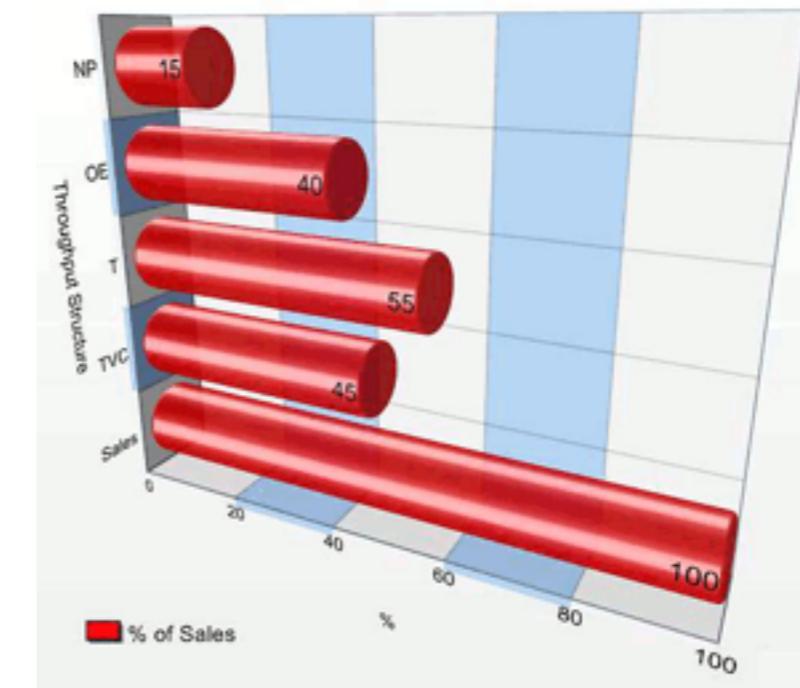
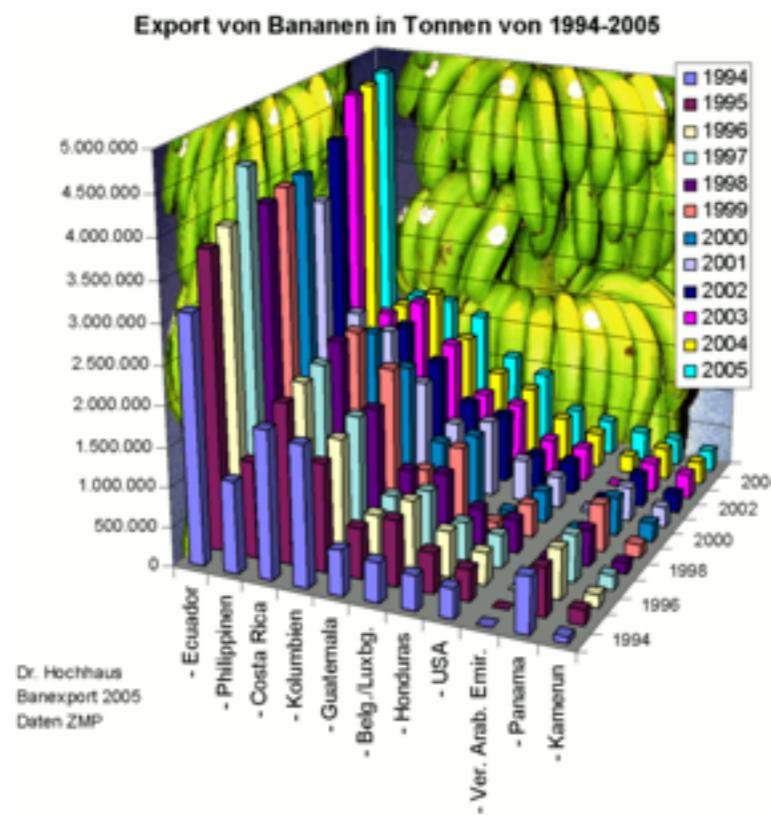
# Avoid Chartjunk



# Don't



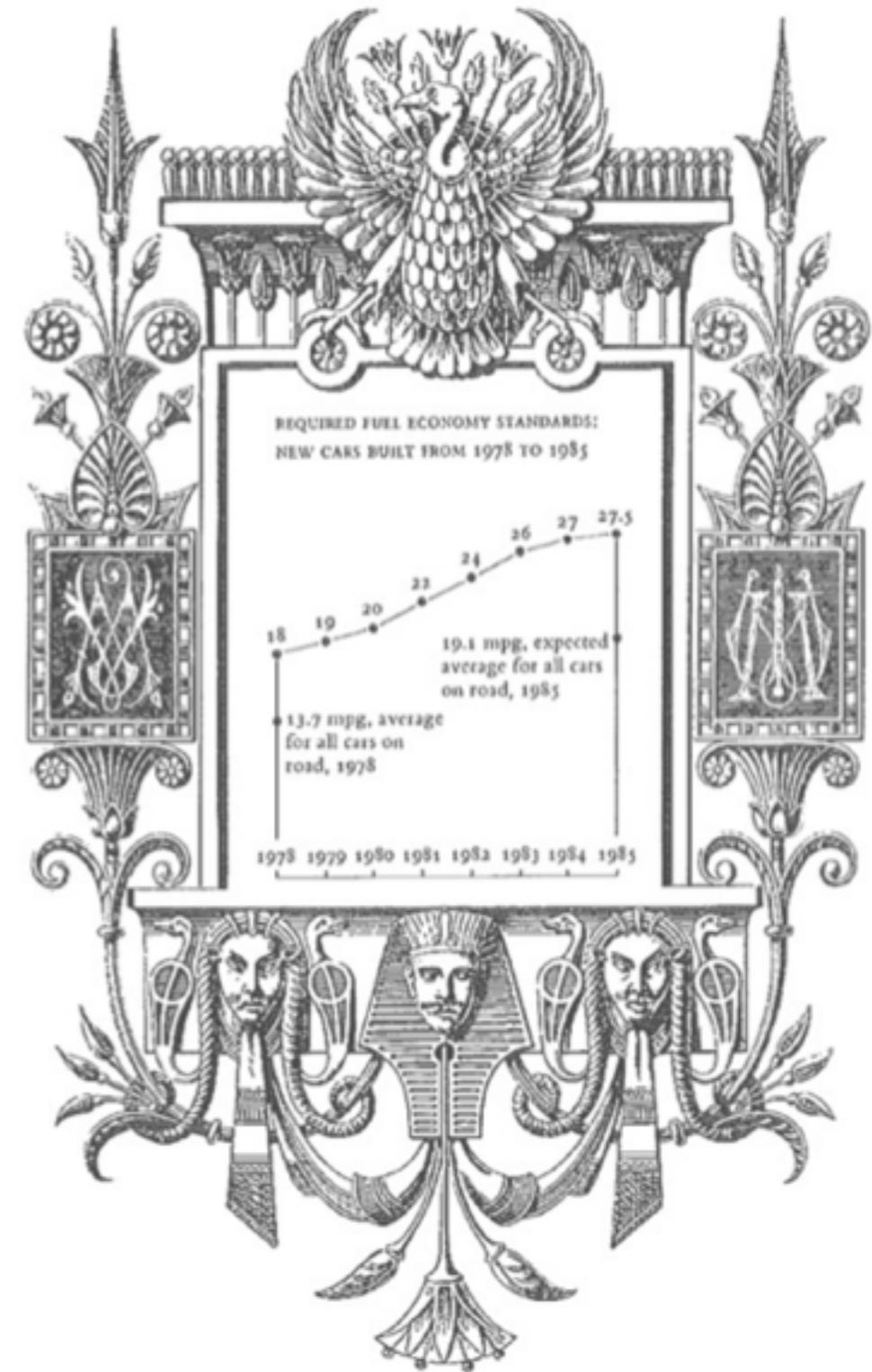
matplotlib gallery



Excel Charts Blog

# Tufte's Design Principles

- Graphical integrity
- The Lie Factor
- Maximize data-ink ratio
- Avoid chart junk



# REASONS TO GO TO VIZTHINK '08

Does your organization struggle with poor communication? Frustrating design processes? Ineffective learning? Visual Thinking can help. Here are just some of the problems you'll solve and some of the things you'll learn at VizThink '08.

## WHAT YOU WILL SOLVE

...slide 127...we're halfway though now!



DEATH BY POWERPOINT



ANALYSIS PARALYSIS



TEAMWORK

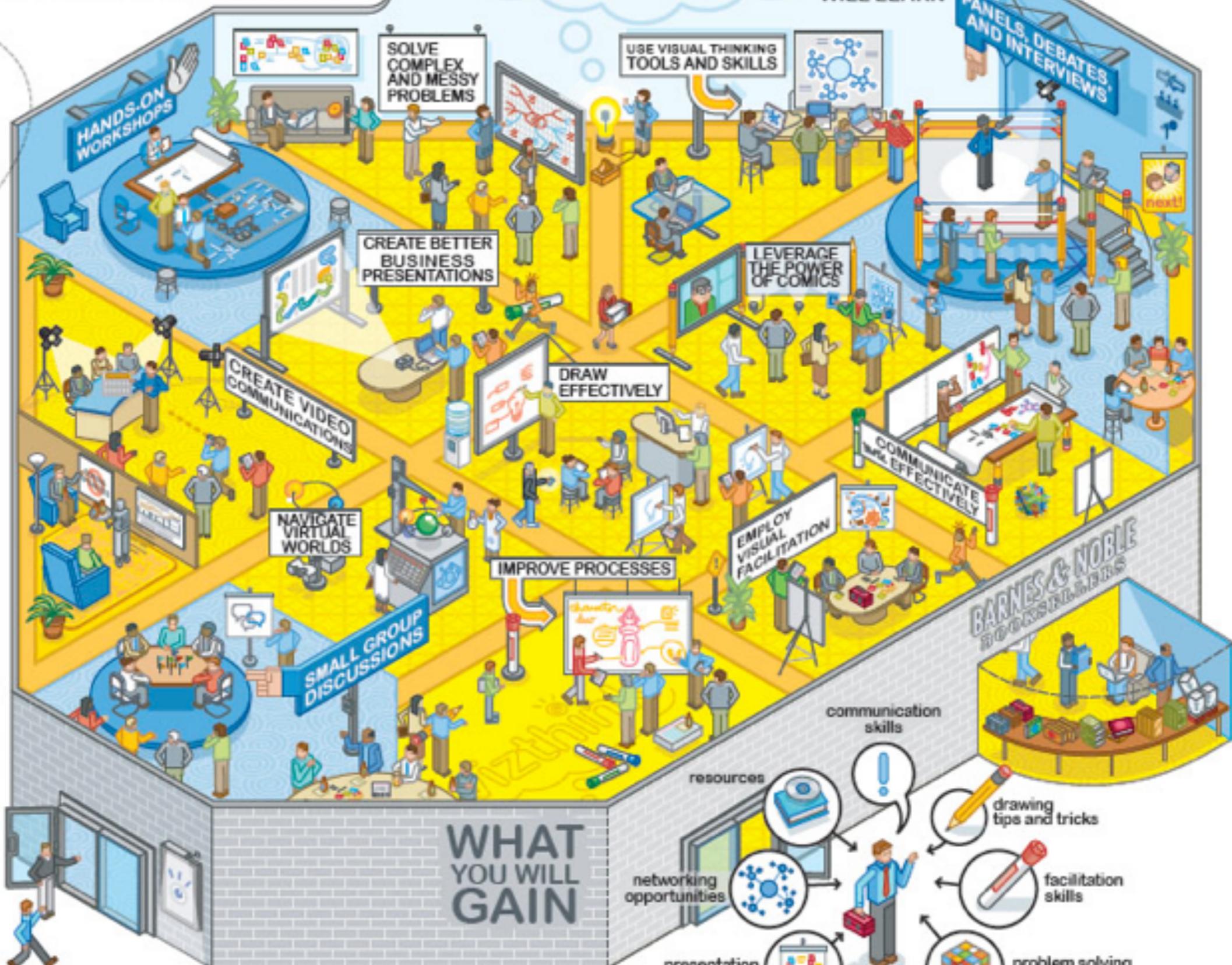
INEFFECTIVE COMMUNICATION



PROCESS HELL

# vizthink'08

## WHAT YOU WILL LEARN



# Subjective Dimensions

- **Aesthetics:** Attractive things are perceived as more useful than unattractive ones
- **Style:** Communicates brand, process, who the designer is
- **Playfulness:** Encourages experimentation and exploration
- **Vividness:** Can make a visualization more memorable

# The Junk Chart Debate

CHI 2010: Graphs

April 10–15, 2010, Atlanta, GA, USA

## Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandryk, Carl Gutwin,  
Aaron Genest, David McDine, Christopher Brooks

Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada  
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,  
aaron.genest@usask.ca, daim085@mail.usask.ca, cab938@mail.usask.ca

### ABSTRACT

Guidelines for designing information charts often state that the presentation should reduce ‘chart junk’ – visual embellishments that are not essential to understanding the data. In contrast, some popular chart designers wrap the presented data in detailed and elaborate imagery, raising the questions of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both interpretation accuracy and long-term recall. We found that people’s accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be produced in this style, our results question some of the premises of the minimalist approach to chart design.

### Author Keywords

Charts, information visualization, imagery, memorability.

### ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

### General Terms

Design, Human Factors

### INTRODUCTION

Many experts in the area of chart design, such as Edward Tufte, criticize the inclusion of visual embellishment in charts and graphs; their guidelines for good chart design often suggest that the addition of *chart junk*, decorations and other kinds of non-essential imagery, to a chart can make interpretation more difficult and can distract readers from the data [22]. This minimalist perspective advocates plain and simple charts that maximize the proportion of *data-ink* – or the ink in the chart used to represent data.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.  
Copyright 2010 ACM 978-1-60558-929-8/10/04...\$10.00.

Despite these minimalist guidelines, many designers include a wide variety of visual embellishments in their charts, from small decorations to large images and visual backgrounds. One well-known proponent of visual embellishment in charts is the graphic artist Nigel Holmes, whose work regularly incorporates strong visual imagery into the fabric of the chart [7] (e.g., Figure 1).

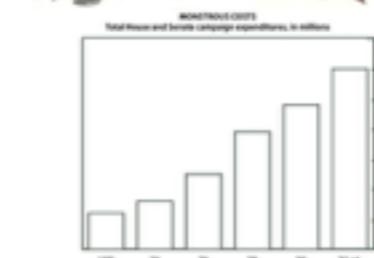
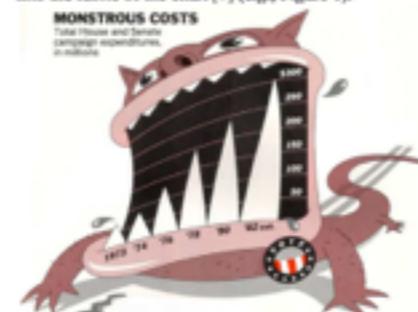


Figure 1. A chart by Holmes [7] (above), and a ‘plain’ version.

These kinds of charts appear regularly in many mass-media publications, and the widespread use of embellished designs raises questions about whether the minimalist position on chart design is really the better approach. Two issues in particular are raised: first, whether visual embellishments do in fact cause comprehension problems; and second, whether the embellishments may provide additional information that is valuable for the reader. For example, the added visual imagery in a Holmes-style chart could draw

IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 17, NO. 12, DECEMBER 2011

2231

## Visualization Rhetoric: Framing Effects in Narrative Visualization

Jessica Hullman, Student Member, IEEE, and Nicholas Diakopoulos, Member, IEEE

**Abstract**—Narrative visualizations combine conventions of communication and intended story. We demonstrate visualization rhetoric as an analytical framework that prioritizes particular interpretations in visualizations that ‘tell a story’ in parallel between narrative visualization interpretation and evidence from literary studies. Devices for understanding the rhetorical nature of narrative visualizations are developed by the rigorous application of concepts from critical theory, semiotics, and design. Design tactics represent additions or omissions of information at various levels of abstraction and interactivity—and how visualizations denote and connote phrasal codes. Classes of rhetorical techniques identified via a systematic analysis are characterized according to their rhetorical contribution to the visualization. From the potentially positive aspects of visualization rhetoric in design, this framework can shed light on how a visualization design prioritizes aspects of its message. Index Terms—Rhetoric, narrative visualization, framing effects, semiotics, critical theory.

### 1 INTRODUCTION

Narrative information visualizations are a style of visualization that often explores the interplay between aspects of both exploratory and communicative visualization [38]. They typically rely on a combination of persuasive, rhetorical techniques to convey an intended story to users as well as exploratory, dialectic strategies aimed at providing the user with control over the insights they gain from interaction. Segel and Heer take an initial step towards highlighting how varying degrees of authorial intention and user interaction are achieved by general design components in narrative visualization [38]. This blend of explorative and communicative features presents another research opportunity though: to better understand a user’s interpretation process of a narrative visualization in light of the rhetorical conventions that the author employs. By explicating rhetorical techniques and how such techniques may affect user interpretation, researchers and designers alike stand to gain a tool for understanding how visualizations communicate.

In this work we examine the design and end-user interpretation of narrative visualizations in order to deepen understanding of how common design techniques represent rhetorical strategies that make certain interpretations more probable. How are rhetorical techniques used in visualization and what are the effects of these techniques on user interpretations of data? Studies in semiotics, journalism, and critical theory indicate particular rhetorical techniques used to communicate an intended message [1, 2, 23], while evidence from decision theory, survey design, and political theory [21, 36, 37] suggests that subtle variations in a representation’s rhetorical or persuasive techniques can generate large effects on users’ interpretations of a message. Investigations related to InfoVis provide initial evidence that how data is framed or presented can significantly affect interpretation [3].

Given the motivation to better understand the interpretation process of visualization, this paper investigates rhetorical strategies and effects in narrative visualization by addressing the following research questions:

- Jessica Hullman is with the University of Michigan, jhullman@umich.edu.
- Nicholas Diakopoulos is with Rutgers University, nicholas.diakopoulos@gmail.com.

Manuscript received 31 March 2011; accepted 1 August 2011; posted online 23 October 2011; mailed on 14 October 2011.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

1077-0000/11/\$20.00 © 2011 IEEE

## What Makes a Visualization Memorable?

Michelle A. Borkin, Student Member, IEEE, Azalea A. Vo, Zoya Bylinskii, Philip Isola, Student Member, IEEE, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister, Senior Member, IEEE

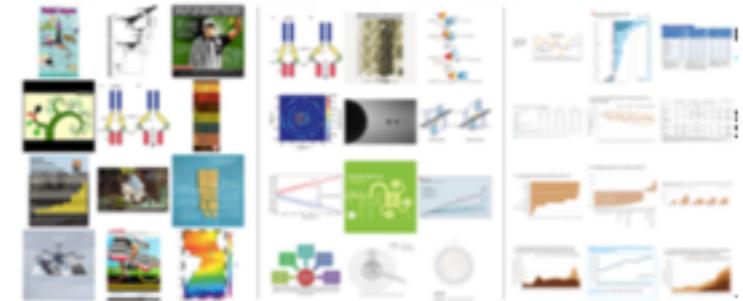


Fig. 1. Left: The top twelve overall most memorable visualizations from our experiment (most to least memorable from top left to bottom right). Middle: The top twelve most memorable visualizations from our experiment when visualizations containing human recognizable cartoons or images are removed (most to least memorable from top left to bottom right). Right: The twelve least memorable visualizations from our experiment (most to least memorable from top left to bottom right).

**Abstract**—An ongoing debate in the Visualization community concerns the role that visualization types play in data understanding. In human cognition, understanding and memorability are intertwined. As a first step towards being able to ask questions about impact and effectiveness, here we ask: ‘What makes a visualization memorable?’ We ran the largest scale visualization study to date using 2,070 single-panel visualizations, categorized with visualization type (e.g., bar chart, line graph, etc.), collected from news media sites, government reports, scientific journals, and infographic sources. Each visualization was annotated with additional attributes, including ratings for data-ink ratios and visual densities. Using Amazon’s Mechanical Turk, we collected memorability scores for hundreds of these visualizations, and discovered that observers are consistent in which visualizations they find memorable and forgettable. We find intuitive results (e.g., attributes like color and the inclusion of a human recognizable object enhance memorability) and less intuitive results (e.g., common graphs are less memorable than unique visualization types). Altogether our findings suggest that quantifying memorability is a general metric of the utility of information, an essential step towards determining how to design effective visualizations.

Index Terms—Visualization taxonomy, information visualization, memorability

### 1 INTRODUCTION

The Visualization community has recently witnessed a divide over the value and impact of excessive chart annotation and decoration (i.e., ‘chart junk’). The conventional view, promoted by visualization experts such as Edward Tufte and Stephen Few, holds that visualizations

- Michelle A. Borkin is with the School of Engineering & Applied Sciences, Harvard University. E-mail: borkin@seas.harvard.edu.
- Azalea A. Vo is with the School of Engineering & Applied Sciences, Harvard University. E-mail: azalea@seas.harvard.edu.
- Zoya Bylinskii is with the Computer Science and Artificial Intelligence Laboratory at the Massachusetts Institute of Technology. E-mail: zoya@mit.edu.
- Philip Isola is with the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology. E-mail: philipi@mit.edu.
- Shashank Sunkavalli is with the School of Engineering & Applied Sciences, Harvard University. E-mail: sunkavalli@seas.harvard.edu.
- Aude Oliva is with the Computer Science and Artificial Intelligence Laboratory at the Massachusetts Institute of Technology. E-mail: olivai@mit.edu.
- Hanspeter Pfister is with the School of Engineering & Applied Sciences, Harvard University. E-mail: pfister@seas.harvard.edu.

Manuscript received 31 March 2011; accepted 1 August 2011; posted online 23 October 2011; mailed on 4 October 2011.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

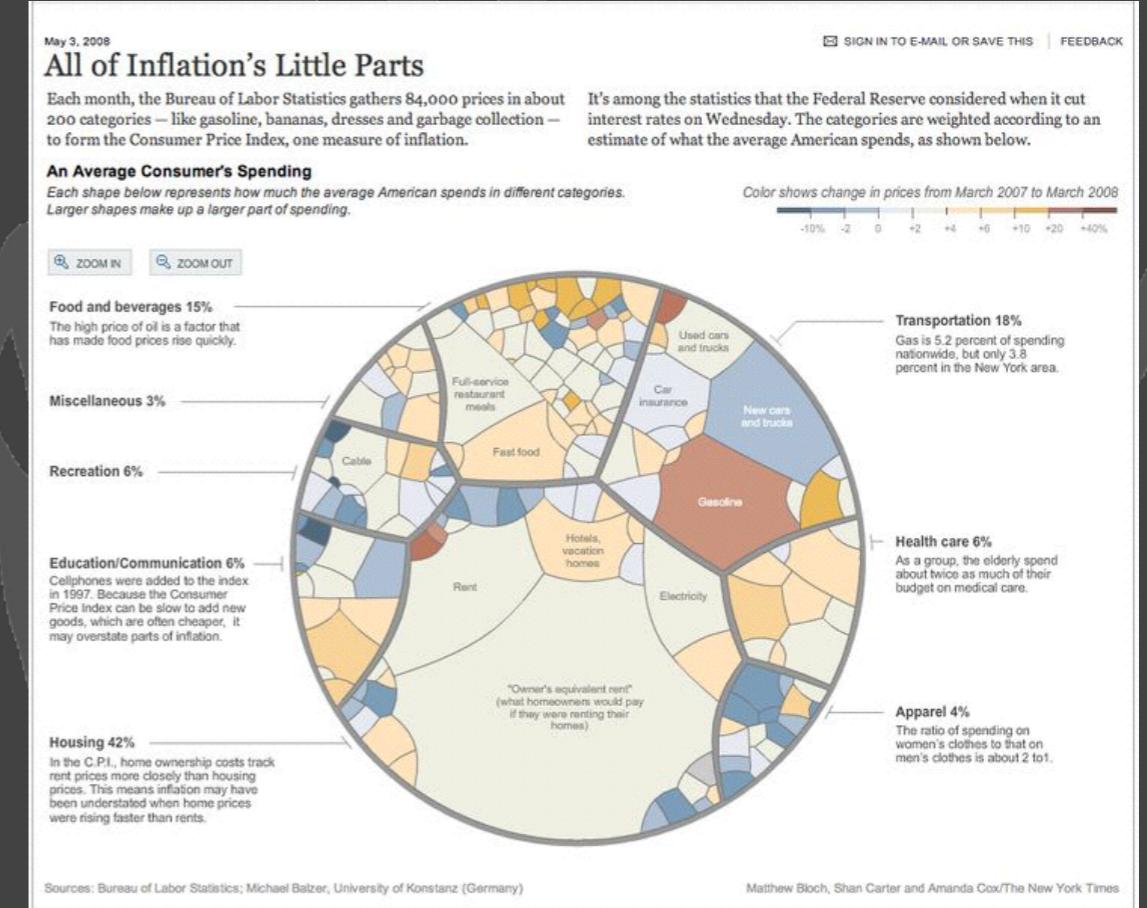
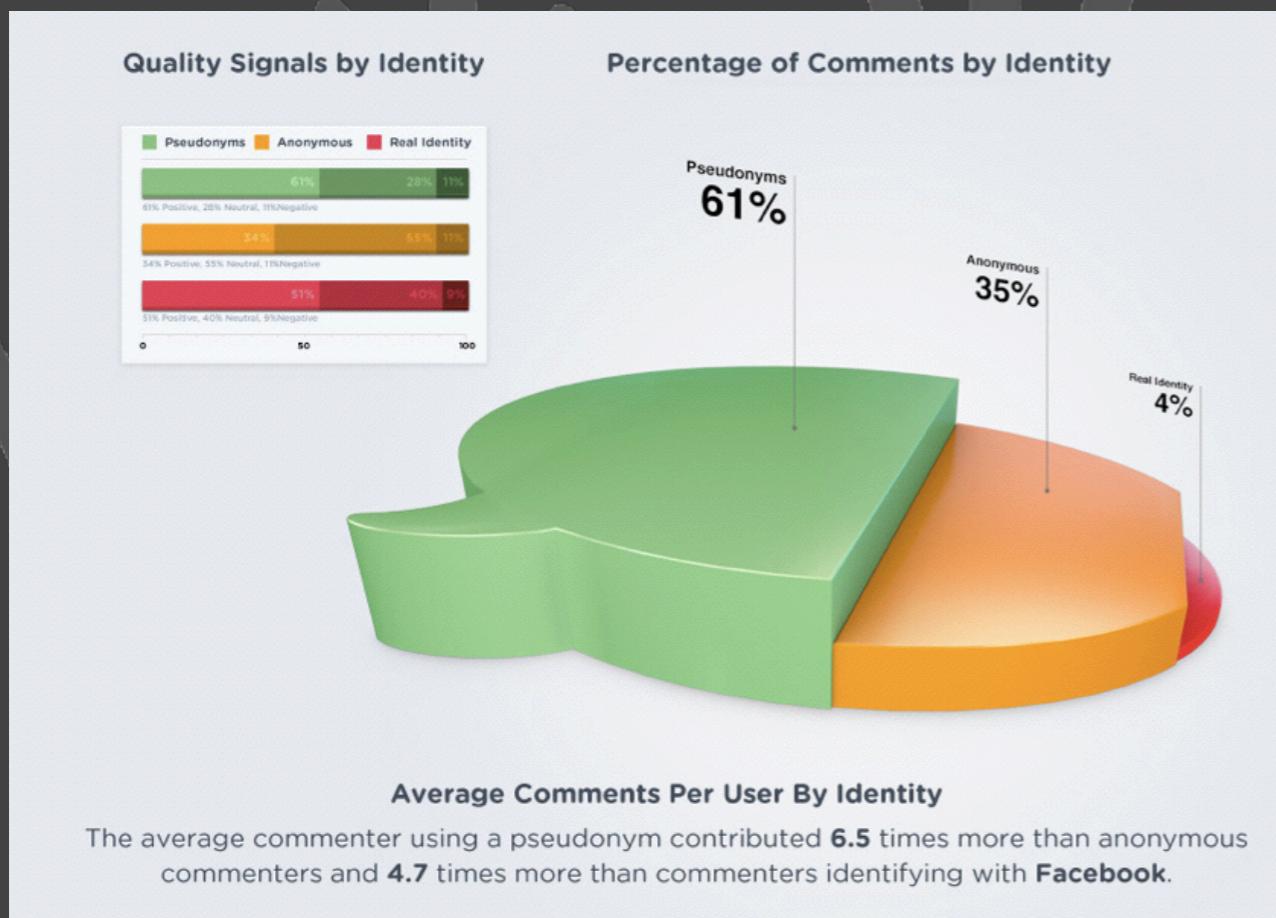
should not include chart junk and should show the data as clearly as possible without any distractions [13, 14, 37, 38]. This view has also been supported by psychology lab studies, which show that simple and clear visualizations are easier to understand [11, 24]. At the other end of the spectrum, researchers have published that chart junk can possibly improve retention and force a viewer to expend more cognitive effort to understand the graph, thus increasing their knowledge and understanding of the data [4, 8, 19]. However, the findings of these studies have been widely debated [13, 14].

What researchers agree on is that chart junk is not the only factor that influences how a person sees, interprets, and remembers a visualization. Other aspects of the visualization, such as graph type, color, or aesthetics, also influence a visualization’s cognitive workload and retention [8, 19, 39]. To disentangle these confounding factors we set out to answer the basic question: ‘What makes a visualization memorable?’ Clearly, a more memorable visualization is not necessarily a more comprehensible one. However, knowing what makes a visualization memorable is a step towards answering higher level questions like ‘What makes a visualization engaging?’ or ‘What makes a visualization effective?’ Recent work has shown that memorability of images of natural scenes is consistent across people, suggesting that some images are intrinsically more memorable than others, independent of an individual’s contexts and biases [20]. We are interested in understanding if these findings hold for visualizations, and what key factors make some visualizations intrinsically more memorable than others.

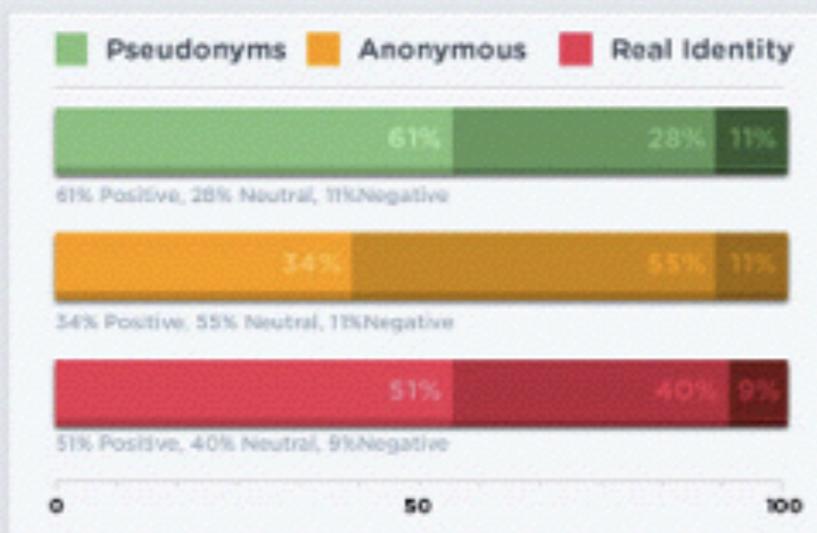
Here, we designed and executed a study to measure the memora-

# Activity

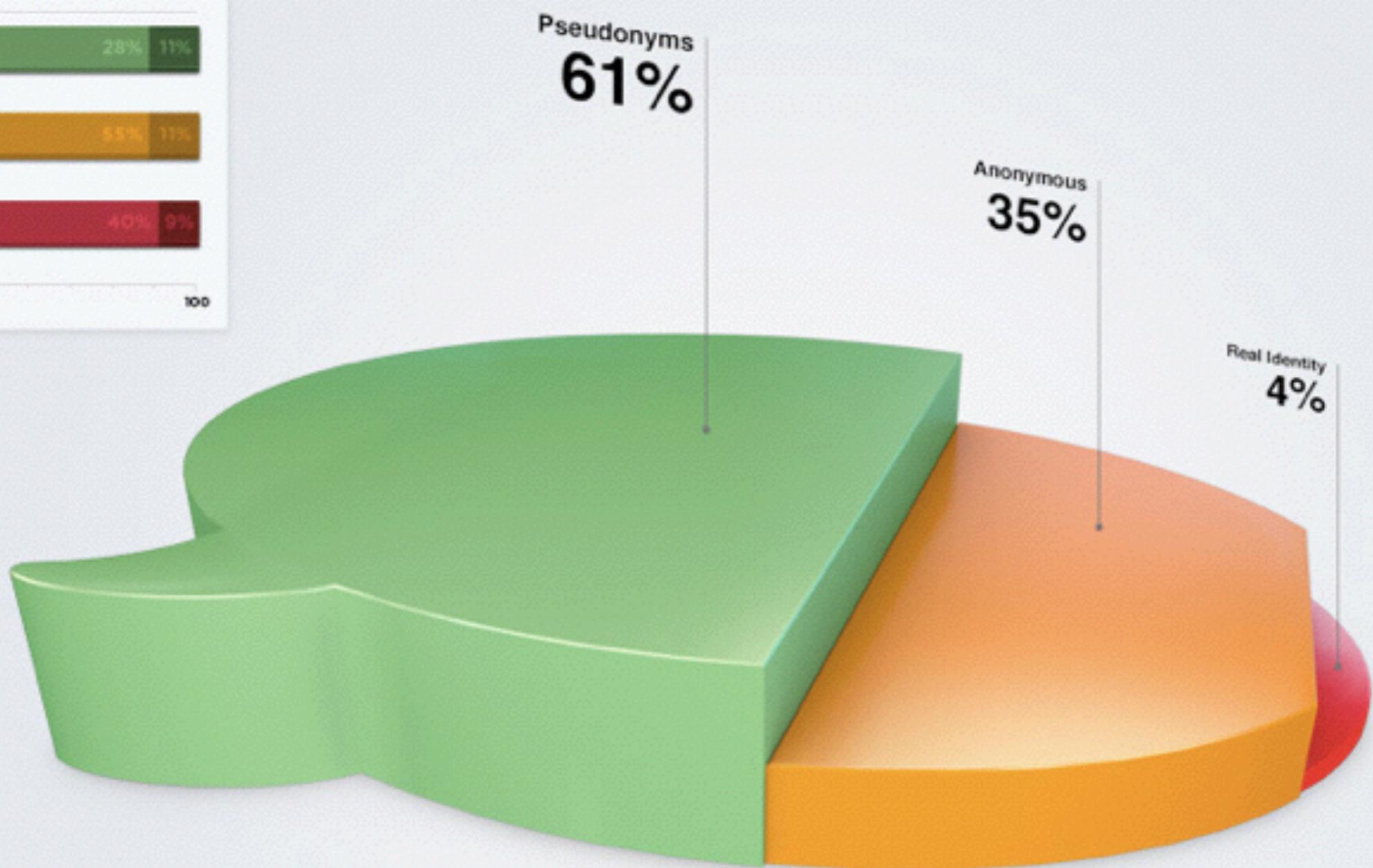
Critique the following visualizations answering the question: Why do you like / dislike this visualization? (3 mins)



## Quality Signals by Identity



## Percentage of Comments by Identity



## Average Comments Per User By Identity

The average commenter using a pseudonym contributed **6.5** times more than anonymous commenters and **4.7** times more than commenters identifying with **Facebook**.

# All of Inflation's Little Parts

Each month, the Bureau of Labor Statistics gathers 84,000 prices in about 200 categories — like gasoline, bananas, dresses and garbage collection — to form the Consumer Price Index, one measure of inflation.

It's among the statistics that the Federal Reserve considered when it cut interest rates on Wednesday. The categories are weighted according to an estimate of what the average American spends, as shown below.

## An Average Consumer's Spending

*Each shape below represents how much the average American spends in different categories. Larger shapes make up a larger part of spending.*

*Color shows change in prices from March 2007 to March 2008*



[ZOOM IN](#) [ZOOM OUT](#)

### Food and beverages 15%

The high price of oil is a factor that has made food prices rise quickly.

### Miscellaneous 3%

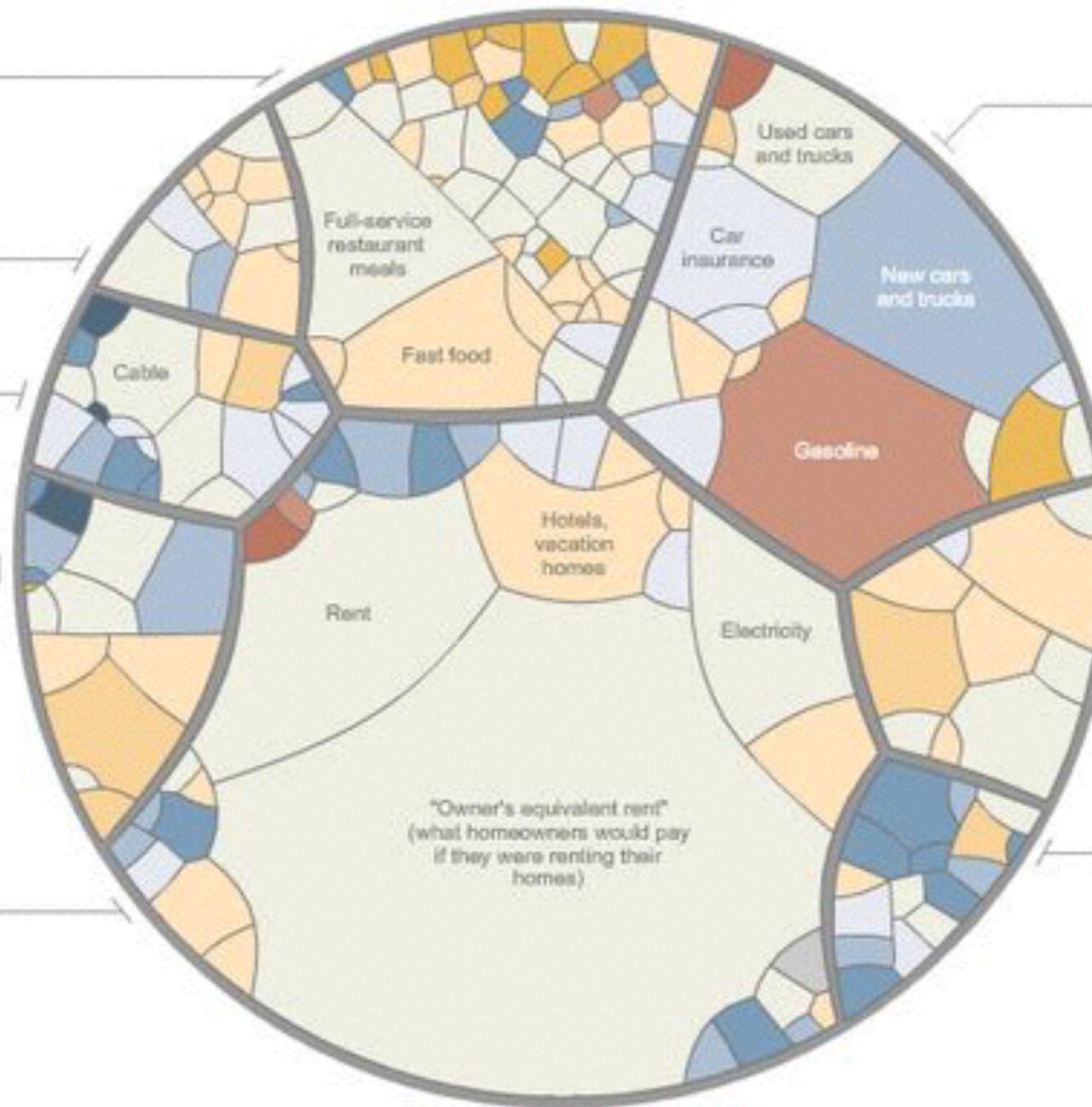
### Recreation 6%

### Education/Communication 6%

Cellphones were added to the index in 1997. Because the Consumer Price Index can be slow to add new goods, which are often cheaper, it may overstate parts of inflation.

### Housing 42%

In the C.P.I., home ownership costs track rent prices more closely than housing prices. This means inflation may have been understated when home prices were rising faster than rents.



### Transportation 18%

Gas is 5.2 percent of spending nationwide, but only 3.8 percent in the New York area.

### Health care 6%

As a group, the elderly spend about twice as much of their budget on medical care.

### Apparel 4%

The ratio of spending on women's clothes to that on men's clothes is about 2 to 1.

# Today's take home message

- Iterative design methods for idea generation, sketching
- Bertin's marks & channels
- Tufte's design principles
- Increase vocabulary to discuss visualizations
- Chart Junk

next Thursday

# Perception

What are the underlying  
human factors and perceptual  
principles for effective design?



## Next Tuesday

- Introduction to JavaScript
- Reading: Murray, Chapter 3 (p. 36-52)



## Next Thursday

- Perception
- Reading: Ware, Chapter 2 (optional: Chapter 4)

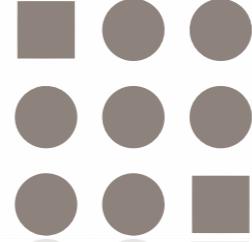


## Homework (due Monday) & Studios

- Homework 1 - HTML, CSS, ...



One minute paper @ Canvas

**CS 171** 

**Please fill out the 1-minute paper!**

Available on Canvas, open until 4:15pm