Intro to Visualization Design

Yannet Interian -- USF

Agenda

- Communicating a tiny dataset
- Forms follows function
    - The functional Art, Cairo. Chapter 2
- Design Guidelines (storytellingwithdata.com/)
    - Put your thoughts into words
    - Be intentional in your use of color
    - Tell a story
   - Align against a common baseline

In class exercise

What are all of the ways you could possibly communicate a tiny data set of two numbers?

75 and 37

Visualizing

1. The familiar indo-arabic numerals notation, babylonian numbers notation
2. Squares
3. Repeated icons
4. Hundreds, tens, units, decimals… represented by squares

http://www.scribblelive.com/blog/2012/07/27/45-ways-to-communicate-two-quantities/
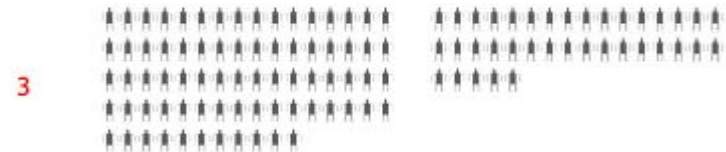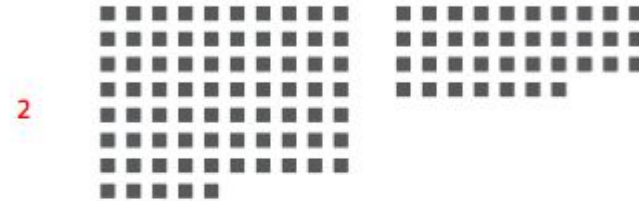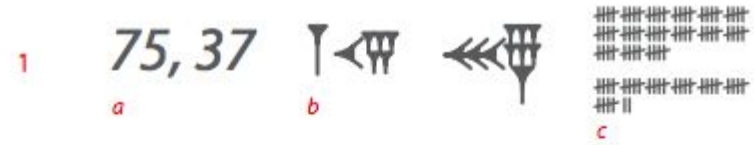
Visualizing 75 and 37

5. Bars
6. Line graphs
7. Percentages bars
8. Spliced bars
9. Proportion
10. Interval
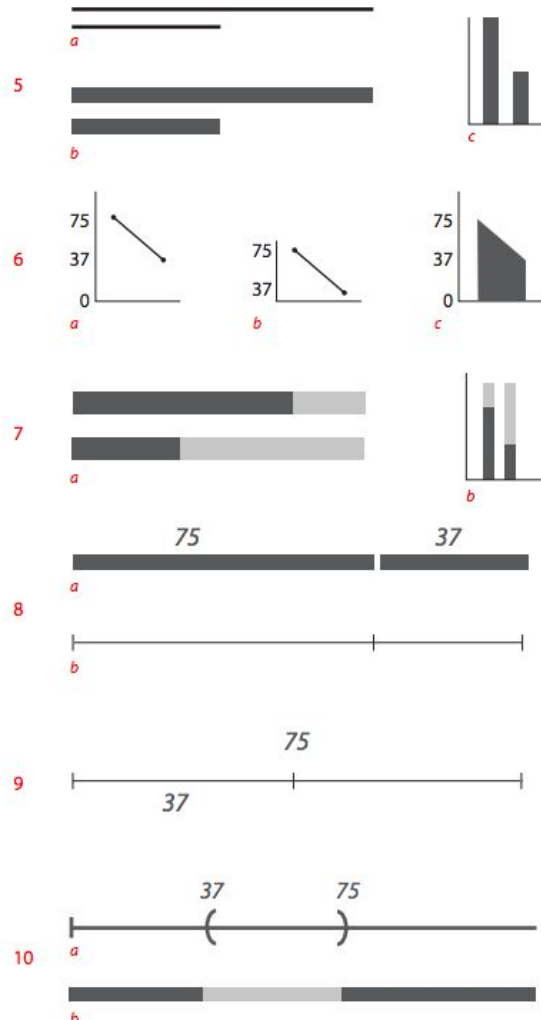
Visualizing 75 and 37

11. Squares merged
12. Percentages in squares
13. Pie charts
14. Donut chart
15. Circle areas
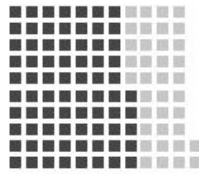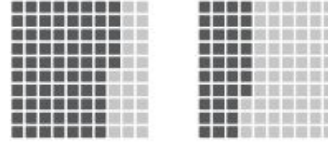
16

17

18
a   b   c

19

20
a   b

21

22

23

24

25

26
a   b

27
a   b

28

29

30

Visualizing 75 and 37

More?
https://s3.amazonaws.com/scribblelive-com-prod/wp-content/uploads/2012/07/poster75and37.pdf

What was the point of the exercise?

- Possible combinations rises exponentially (consider richer dataset)
- Information visualization is a language
  - Explain the world
  - Tell stories
  - Point specific facts and objects
  - Elaborate ambiguous messages
- We should learn good rules to use it effectively

"Form Follows Function"

- Visualizations are not arbitrary
  - Objectivity, precision and functionality constrain the forms they take
- Ask: What is this for?
  - Present data
  - Allow comparisons
  - Organize
  - Make relationships / correlations
- Is it good at any of these?

A portrait of the power balance in South America

DEFENSE IN SOUTH AMERICA

An overview of the armed forces of countries around Brazil

Population (millions of people)

$ Defense budget (billions of dollars)

Size of armed forces (thousands)

254.2 | 44.2 | 7.14
Prospects: Colombia will improve its armed forces in the next few years. It will invest 30 billion dollars to buy Brazilian fighters, Russian tanks, and Spanish propelled rockets.

115 | 26 | 2.6
Prospects: Venezuela will keep buying Russian vehicles, such as Sukhoi fighters and armored gunships. It will also buy several Kilo class submarines.

367.9 | 190 | 21.6
Prospects: Brazil will finish building 250 Leopard tanks, and keep modernizing its F-5 fighters. It will also buy a nondisclosed number of combat planes.

57 | 13.7 | 0.92
Prospects: Ecuador will not make significant investments in the near future.

114 | 28.6 | 1.2
Prospects: Peru will invest in an upgrade of its airforce.

46,1 | 9.1 | 21.6
Prospects: Bolivia will not make significant investments in the near future.

65 | 16.3 | 4.6
Prospects: Chile will buy several A310 planes and Leopard tanks.

76 | 40.3 | 2.05
Prospects: Argentina has announced that it will modernize its fleet of tanks.

COLOMBIA, VENEZUELA, ECUADOR, PERU, BRAZIL, BOLIVIA, CHILE, ARGENTINA

Sources: *Folha de São Paulo*, Brazilian Center for Strategic Studies

What can we learn from this graph?

Population (millions of people)

Defense budget (billions of dollars)

Size of armed forces (thousands)

DEFENSE IN SOUTH AMERICA

An overview of the armed forces of countries around Brazil

Population (millions of people)

Defense budget (billions of dollars)

Size of armed forces (thousands)

254.2

44.2 | 7.14

Prospects: Colombia will improve its armed forces in the next few years. It will invest 30 billion dollars to buy Brazilian fighters, Russian tanks, and Spanish propelled rockets.

VENEZUELA

115

26 | 2.6

Prospects: Venezuela will keep buying Russian vehicles, such as Sukhoi fighters and armored gunships. It will also buy several Kilo class submarines.

367.9

190 | 21.6

Prospects: Brazil will finish building 250 Leopard tanks, and keep modernizing its F-5 fighters. It will also buy a nondisclosed number of combat planes.

COLOMBIA

ECUADOR

57

13.7 | 0.92

Prospects: Ecuador will not make significant investments in the near future.

PERU

BRAZIL

BOLIVIA

114

28.6 | 1.2

Prospects: Peru will invest in an upgrade of its airforce.

CHILE

46,1

9.1 | 21.6

Prospects: Bolivia will not make significant investments in the near future.

ARGENTINA

65

16.3 | 4.6

Prospects: Chile will buy several A310 planes and Leopard tanks.

76

40.3 | 2.05

Prospects: Argentina has announced that it will modernize its fleet of tanks.

Sources:
Folha de São Paulo,
Brazilian Center
for Strategic Studies
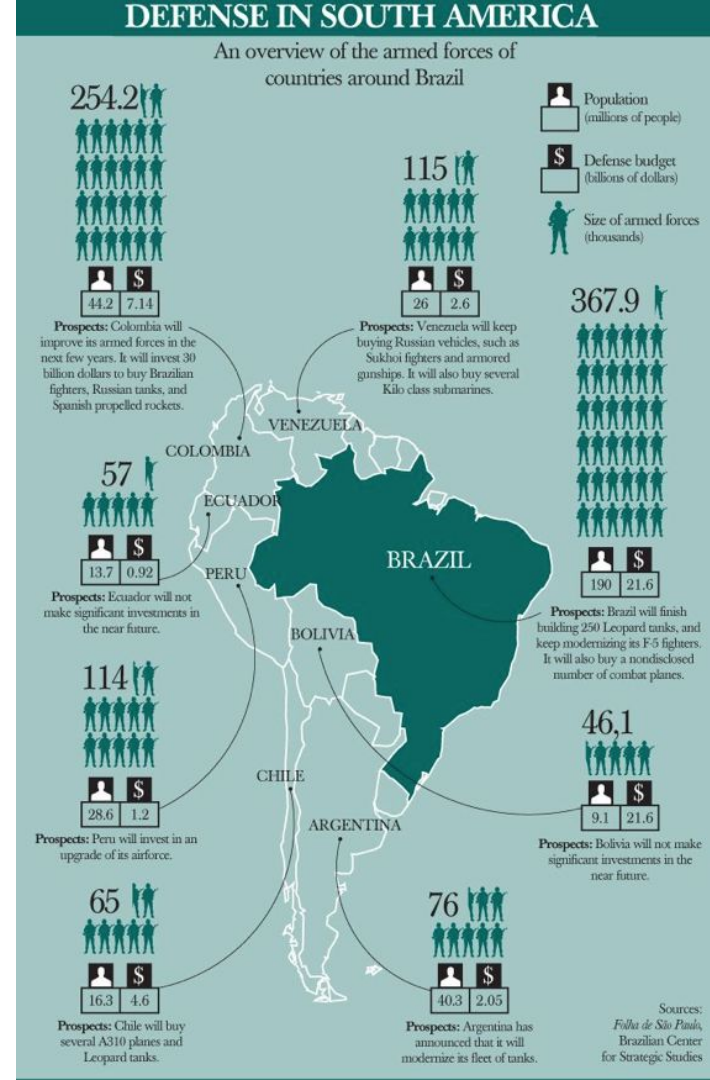
What can we learn from this graph?

Any of these?
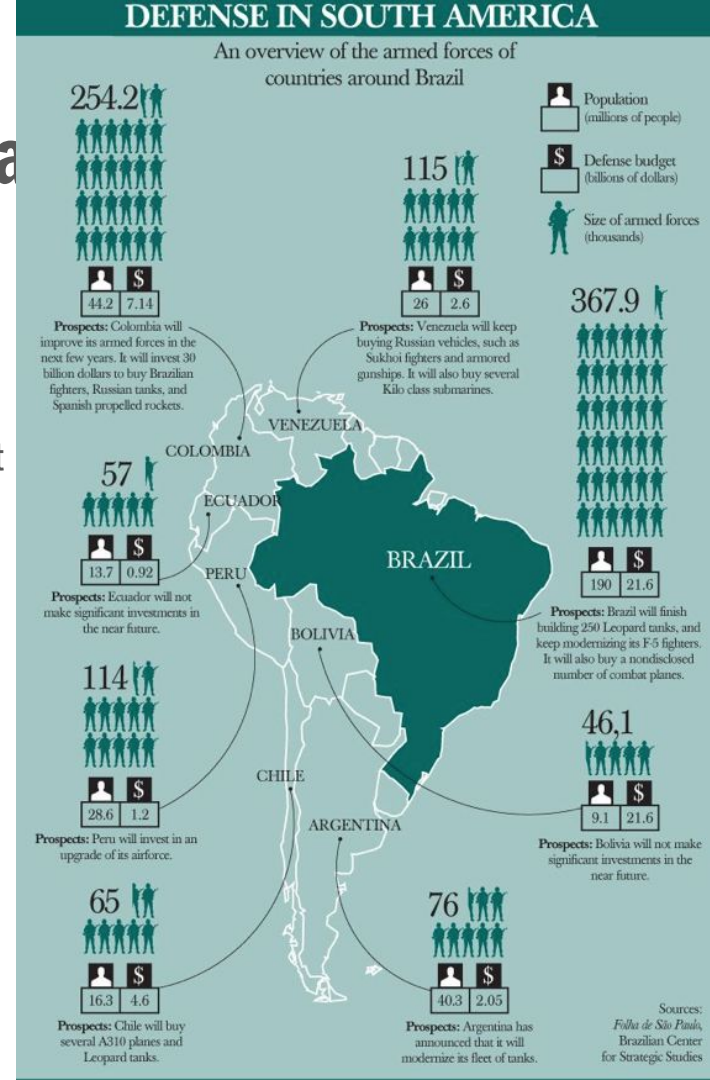
- Present several variables related to the topic
  - number of soldiers, population, defense budget
- Allow comparisons
- Organize countries
  - based on variables and comparisons
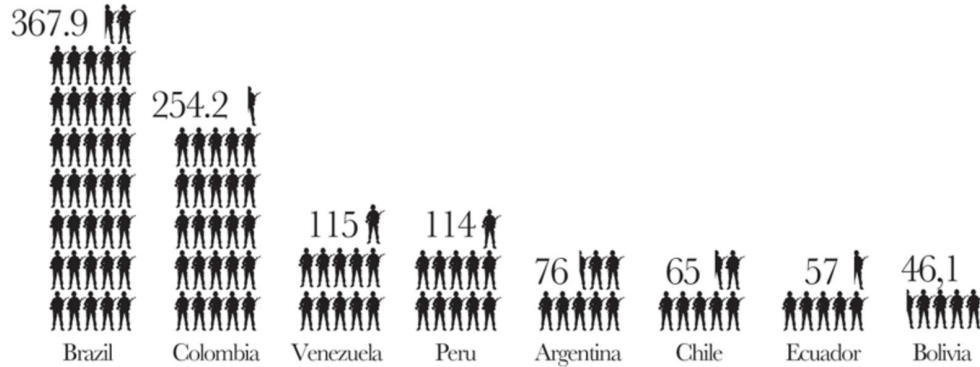- Make correlations or relationships

What can we learn from this gra

Any of these?
- Present several variables related to the topic
  - number of soldiers, population, defense budget
- ~~Allow comparisons~~
- ~~Organize countries~~
  - based on variables and comparisons
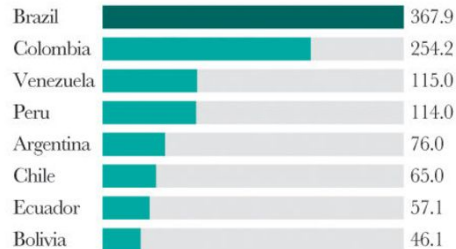- ~~Make correlations or relationships~~

DEFENSE IN SOUTH AMERICA

An overview of the armed forces of countries around Brazil

More accurate comparisons

367.9

254.2

115   114

76   65   57   46,1

Brazil   Colombia   Venezuela   Peru   Argentina   Chile   Ecuador   Bolivia

- Bars on the same horizontal axis allows for more accurate comparisons
- Even better use traditional bar charts

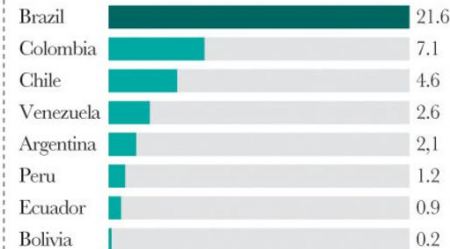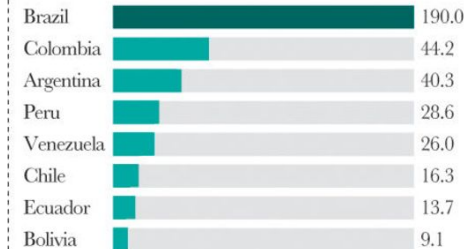**ARMED FORCES**
(Thousands of people)

Brazil	367.9
Colombia	254.2
Venezuela	115.0
Peru	114.0
Argentina	76.0
Chile	65.0
Ecuador	57.1
Bolivia	46.1

DEFENSE BUDGET
(Billions of US$ a year)

Brazil	21.6
Colombia	7.1
Chile	4.6
Venezuela	2.6
Argentina	2,1
Peru	1.2
Ecuador	0.9
Bolivia	0.2

POPULATION
(Millions of people)

Brazil	190.0
Colombia	44.2
Argentina	40.3
Peru	28.6
Venezuela	26.0
Chile	16.3
Ecuador	13.7
Bolivia	9.1

Normalizing

ARMED FORCES EMPLOYEES PER 1,000 PEOPLE

Country	Value
Colombia	5.8
Bolivia	5.1
Venezuela	4.4
Ecuador	4.2
Chile	4.0
Peru	4.0
Brazil	1.9
Argentina	1.9

PER CAPITA SPENDING IN ARMED FORCES *(in US dollars a year)*

Country	Value
Chile	282.2
Colombia	161.5
Brazil	113.7
Venezuela	100.0
Ecuador	61.2
Argentina	50.9
Peru	41.9
Bolivia	17.6

MONEY SPENT ON EACH ARMED FORCES EMPLOYEE *(in US dollars a year)*

Country	Value
Chile	70.8
Brazil	58.7
Colombia	28.1
Argentina	27.0
Venezuela	22.6
Ecuador	16.1
Peru	10.5
Bolivia	3.5

THE DEFENSE OF THE NEIGHBORS

An overview of the armed forces of countries around Brazil

Brazil has the strongest armed forces in South America in absolute terms—

ARMED FORCES
(Thousands of people)

Brazil	367.9
Colombia	254.2
Venezuela	115.0
Peru	114.0
Argentina	76.0
Chile	65.0
Ecuador	57.1
Bolivia	46.1

DEFENSE BUDGET
(Billions of US$ a year)

Brazil	21.6
Colombia	7.1
Chile	4.6
Venezuela	2.6
Argentina	2.1
Peru	1.2
Ecuador	0.9
Bolivia	0.2

POPULATION
(Millions of people)

Brazil	190.0
Colombia	44.2
Argentina	40.3
Peru	28.6
Venezuela	26.0
Chile	16.3
Ecuador	13.7
Bolivia	9.1

—but not in relative terms

ARMED FORCES EMPLOYEES PER 1,000 PEOPLE

Colombia	5.8
Bolivia	5.1
Venezuela	4.4
Ecuador	4.2
Chile	4.0
Peru	4.0
Brazil	1.9
Argentina	1.9

PER CAPITA SPENDING IN ARMED FORCES *(in US dollars a year)*

Chile	282.2
Colombia	161.5
Brazil	113.7
Venezuela	100.0
Ecuador	61.2
Argentina	50.9
Peru	41.9
Bolivia	17.6

MONEY SPENT ON EACH ARMED FORCES EMPLOYEE *(in US dollars a year)*

Chile	70.8
Brazil	58.7
Colombia	28.1
Argentina	27.0
Venezuela	22.6
Ecuador	16.1
Peru	10.5
Bolivia	3.5

A different look at the data

This is a 'trend line'. The closer the circles are to this line, the more correlated the variables are

Size of circles is proportional to the defense budget (millions of dollars per year)

ARMED FORCES EMPLOYEES

POPULATION (Millions of people)

Future investments

COLOMBIA will improve its armed forces in the next few years. It will invest 30 billion dollars to buy Brazilian fighters, Russian tanks, and Spanish propelled rockets

ECUADOR will not make significant investments in the near future

PERU will invest in an upgrade of its air force

CHILE will buy several A310 planes and Leopard tanks

VENEZUELA will keep buying Russian vehicles, suchas Sukhoi fighters and armored gunships. It will also buy several Kilo class submarines.

BRASIL will finish building a 250 Leopard tanks force, and keep modernizing its F-5 fighters. It will also buy an nondisclosed number of combat planes

BOLIVIA will not make significant investments in the near future

ARGENTINA has announced that it will modernize its fleet of tanks and fighters
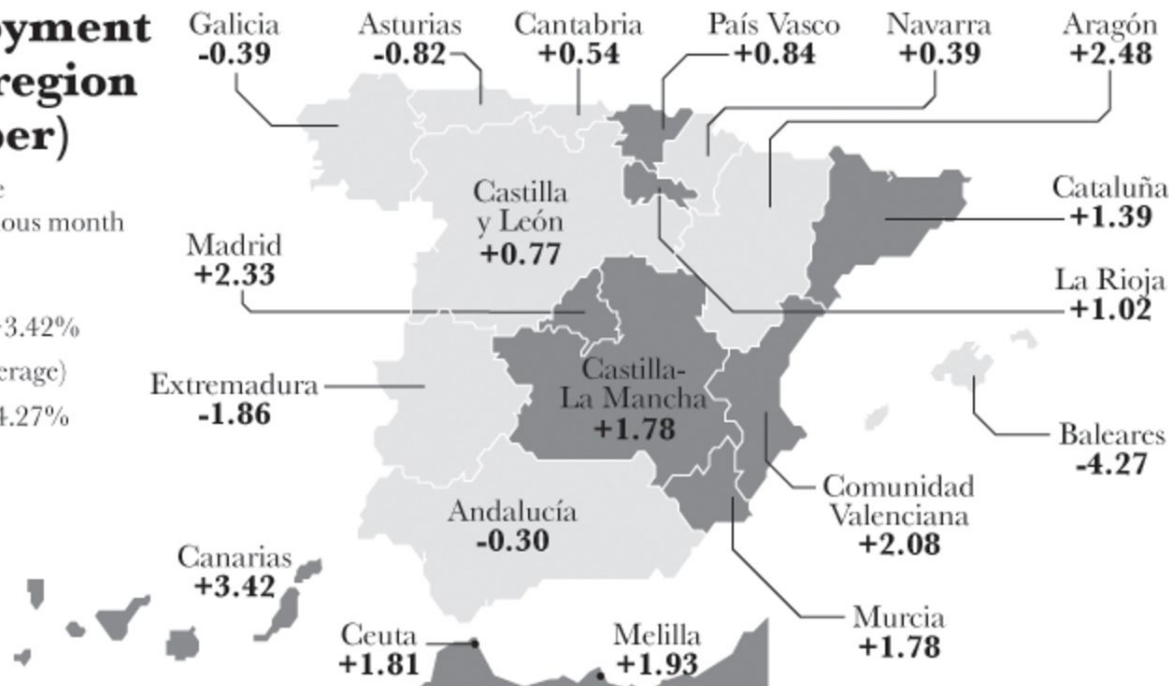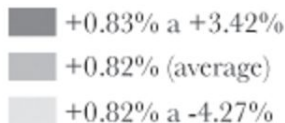
Summary

- Placing bars on the same horizontal axis allows for more accurate comparisons
    - Order the categories on increasing or decreasing order (depending on what is more important)
- Avoid using pictograms → bar chart
- Use "per capita" since the countries have very different sizes allow us to make correlations
- Highlighting the focus of the article ("Brazil") makes for easier comparisons
- Use other graphics to look at the same data

Unemployment rates by region in Spain

Unemployment rates by region (in October)

Percentage change compared to previous month

- +0.83% a +3.42%
- +0.82% (average)
- +0.82% a -4.27%

Galicia -0.39

Asturias -0.82

Cantabria +0.54

País Vasco +0.84

Navarra +0.39

Aragón +2.48

Cataluña +1.39

Castilla y León +0.77

Madrid +2.33

La Rioja +1.02

Extremadura -1.86

Castilla- La Mancha +1.78

Baleares -4.27

Comunidad Valenciana +2.08

Andalucía -0.30

Canarias +3.42

Ceuta +1.81

Melilla +1.93

Murcia +1.78

Unemployment rates by region in Spain

1. In which region has unemployment grown the most
2. In which region has it dropped the most?
3. Has unemployment change been bigger in Madrid or Canarias"?

Unemployment rates by region (in October)

Percentage change compared to previous month

+0.83% a +3.42%
+0.82% (average)
+0.82% a -4.27%

Galicia -0.39
Asturias -0.82
Cantabria +0.54
País Vasco +0.84
Navarra +0.39
Aragón +2.48
Cataluña +1.39
Castilla y León +0.77
Madrid +2.33
La Rioja +1.02
Extremadura -1.86
Castilla-La Mancha +1.78
Baleares -4.27
Andalucía -0.30
Comunidad Valenciana +2.08
Canarias +3.42
Murcia +1.78
Ceuta +1.81
Melilla +1.93
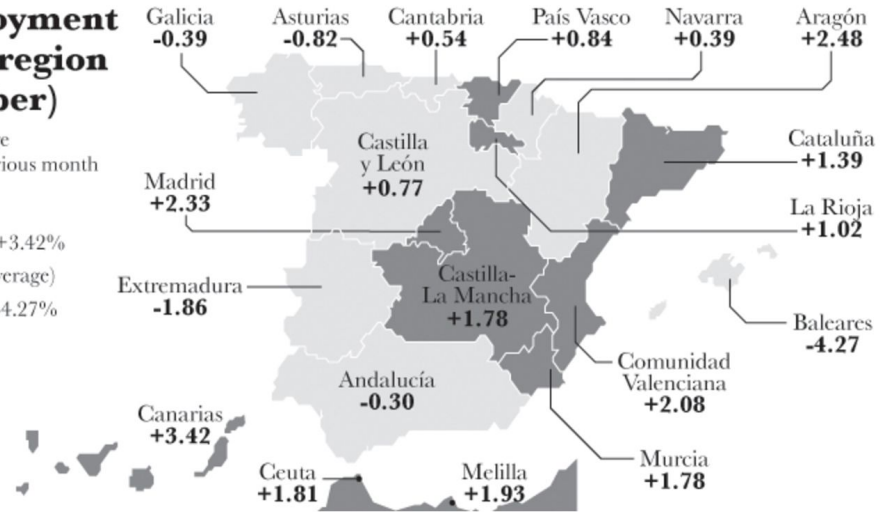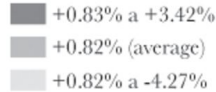
Your challenges reading the chart

1. Your eyes look for the numbers mentioned in the question
2. Your brain memorizes them
3. Your brain organizes the numbers
4. Your brain compares the reorganized

Think like a Reader

If we anticipate questions that readers will ask, we can design visualizations so that those questions are easier to answer. How does the design assist cognition and analytic thinking about the data?

Unemployment rates by region (in October)

Percentage change compared to previous month

Average: +0.82
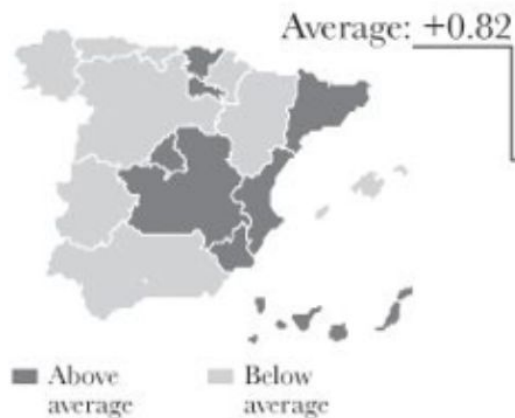
Region	Value
Canarias	+3.42
Aragón	+2.48
Madrid	+2.33
C. Valenciana	+2.08
Melilla	+1.93
Ceuta	+1.81
Murcia	+1.78
C.-La Mancha	+1.78
Cataluña	+1.39
La Rioja	+1.02
País Vasco	+0.84
C. y León	+0.77
Cantabria	+0.54
Navarra	+0.39
Andalucía	-0.30
Galicia	-0.39
Asturias	-0.82
Extremadura	-1.86
Baleares	-4.27

■ Above average   ■ Below average

This version allows you to compare and rank Spain's regions

The bubble plague

Banks: Market Cap

- Market Value as of January 20th 2009, $Bn
- Market Value as of Q2 2007, $Bn
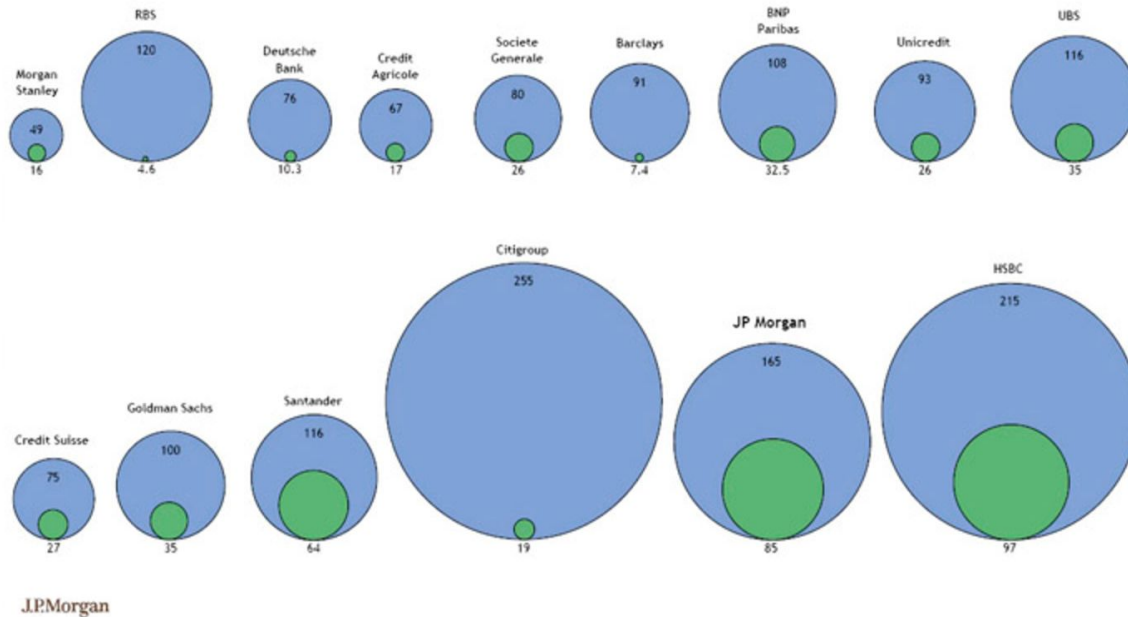
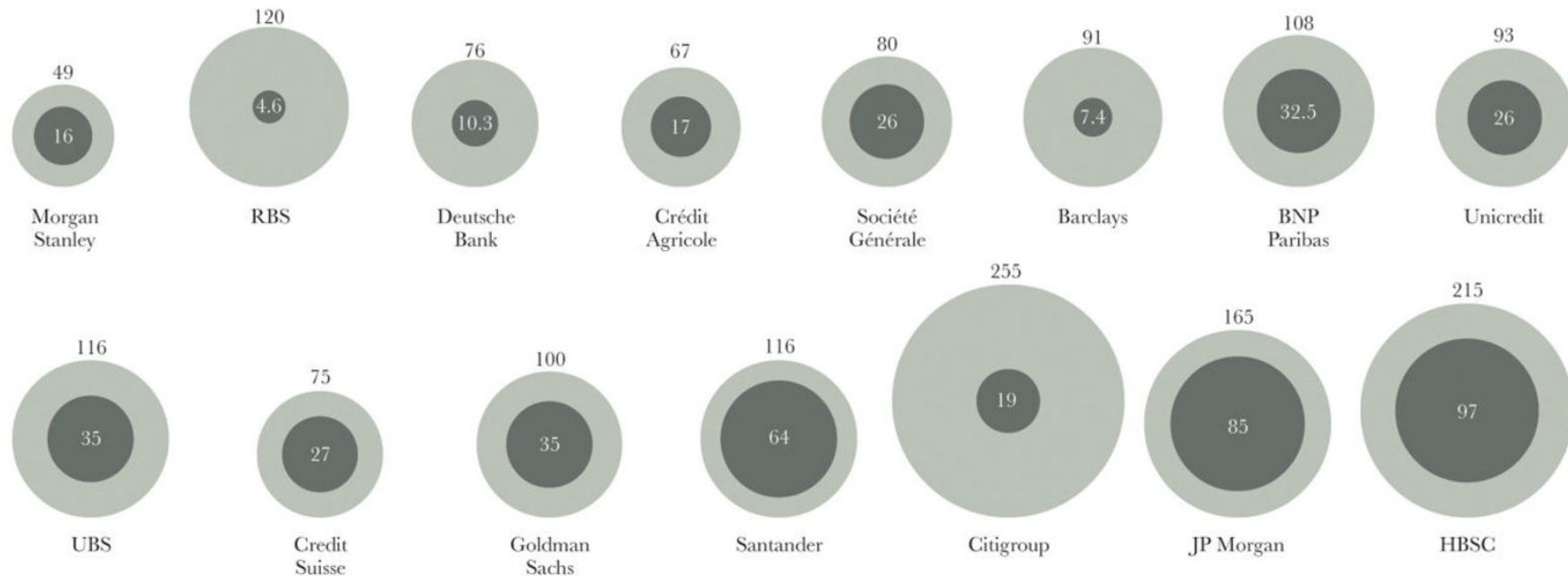http://www.perceptualedge.com/example18.php

The bubble plague (version 2)

Market Capitalization of the World's Biggest Banks

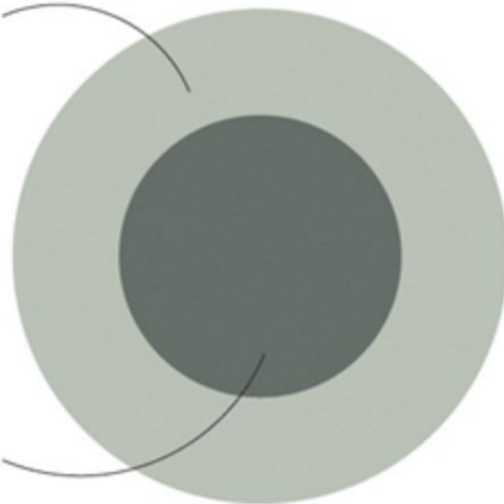In billions of dollars · January 2007 · January 2009 · Source: Bloomberg

Bank	January 2007	January 2009
Morgan Stanley	49	16
RBS	120	4.6
Deutsche Bank	76	10.3
Crédit Agricole	67	17
Société Générale	80	26
Barclays	91	7.4
BNP Paribas	108	32.5
Unicredit	93	26
UBS	116	35
Credit Suisse	75	27
Goldman Sachs	100	35
Santander	116	64
Citigroup	255	19
JP Morgan	165	85
HBSC	215	97

In class quiz

Market Capitalization of Société Générale

Billions of dollars   Source: Bloomberg

January 2007

January 2009

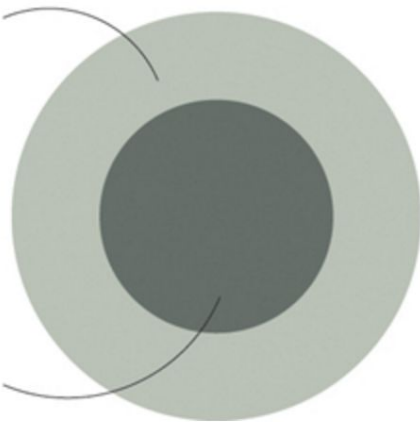If the grey bubble represents $80 billion. How much is the black bubble?

Comparing bubble versus bars

Bubbles versus Bars

- Human brain is not good at estimating surface areas
- Human brain is much better at comparing a single dimension such as length or height
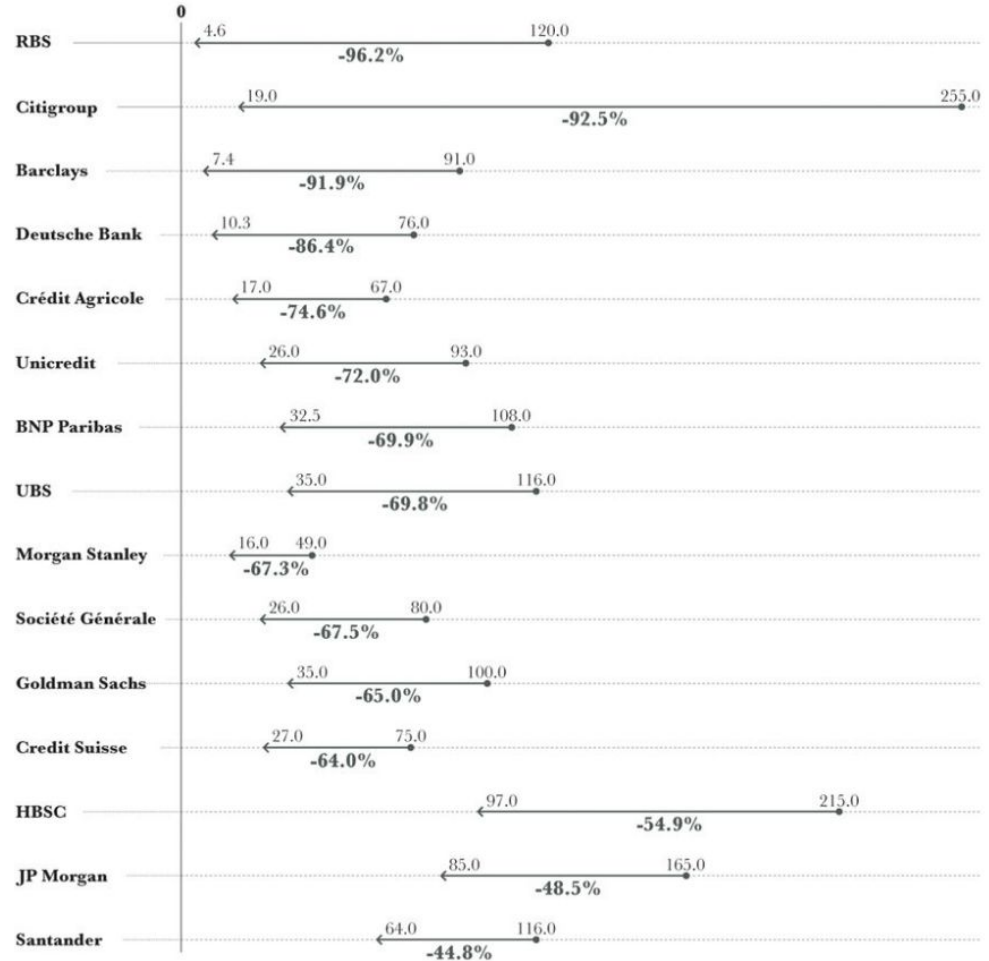- You want the reader to compare areas, but they tend to compare heights

Re-Desing

- This chart orders the banks by the amount of change in %
- The chart also allows multiple comparisons

Market Capitalization of the World's Biggest Banks

Billions of dollars   January 2009 ⟵————————•January 2007   Source: Bloomberg

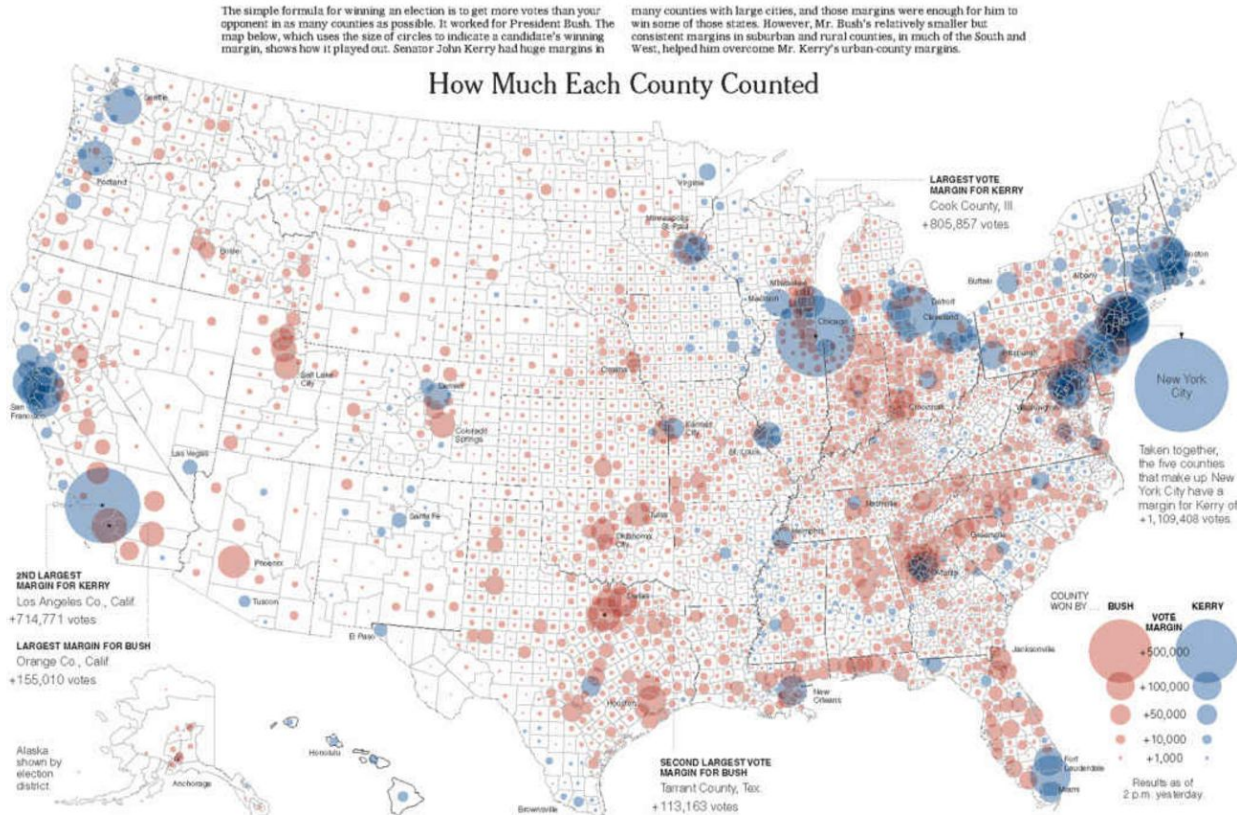| Bank	Jan 2009	Jan 2007	Change
RBS	4.6	120.0	-96.2%
Citigroup	19.0	255.0	-92.5%
Barclays	7.4	91.0	-91.9%
Deutsche Bank	10.3	76.0	-86.4%
Crédit Agricole	17.0	67.0	-74.6%
Unicredit	26.0	93.0	-72.0%
BNP Paribas	32.5	108.0	-69.9%
UBS	35.0	116.0	-69.8%
Morgan Stanley	16.0	49.0	-67.3%
Société Générale	26.0	80.0	-67.5%
Goldman Sachs	35.0	100.0	-65.0%
Credit Suisse	27.0	75.0	-64.0%
HBSC	97.0	215.0	-54.9%
JP Morgan	85.0	165.0	-48.5%
Santander	64.0	116.0	-44.8%

Bubbles are not always wrong

- The blanks visualization is supposed to help us make precise comparisons between values
- Bubbles can be effective if your goal is to show general patterns and trends and it is the case of the next visualization
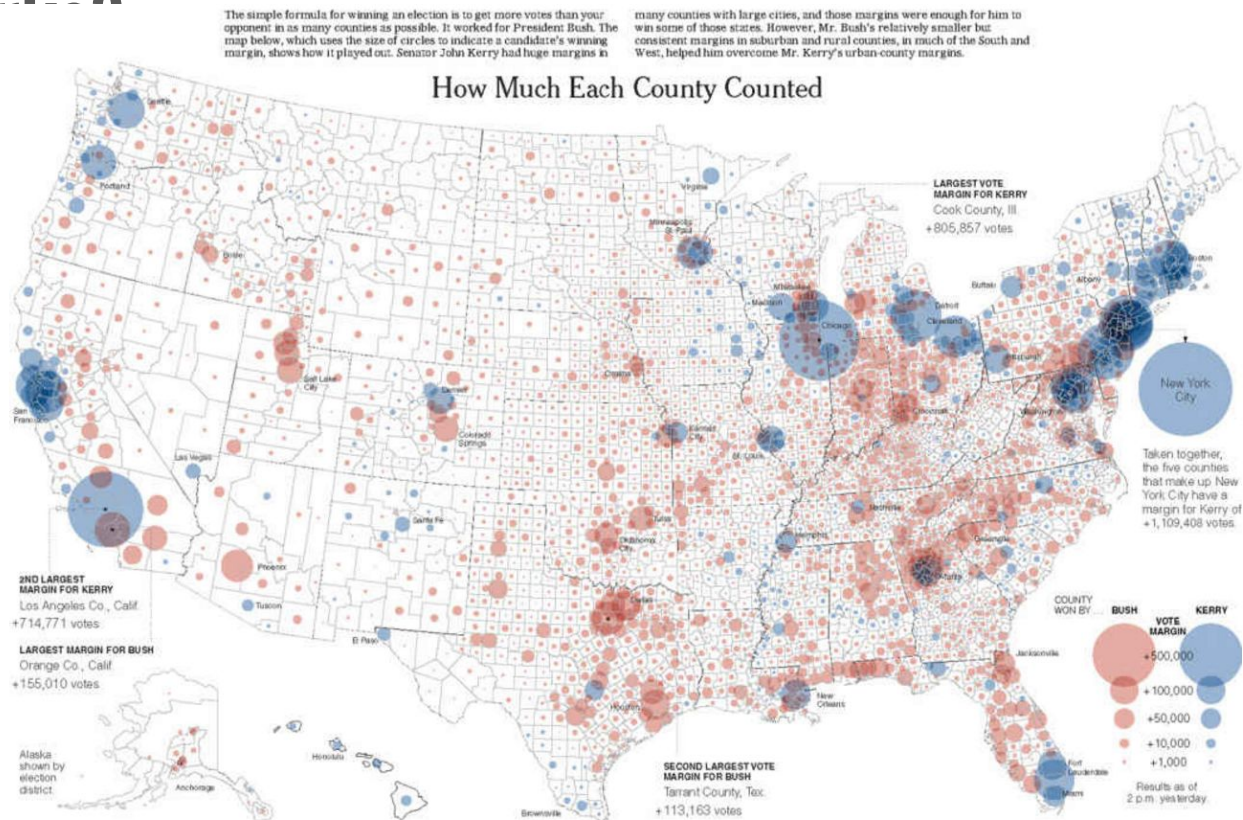
Why this map works?

- Published by the New York times after the 2004 elections
- Each circle represent the margin of votes won by Jonh Kerry (blue) or by George W Bush (red)

How Much Each County Counted

Why this map works?

- Goals of this vis are different from the bank vis
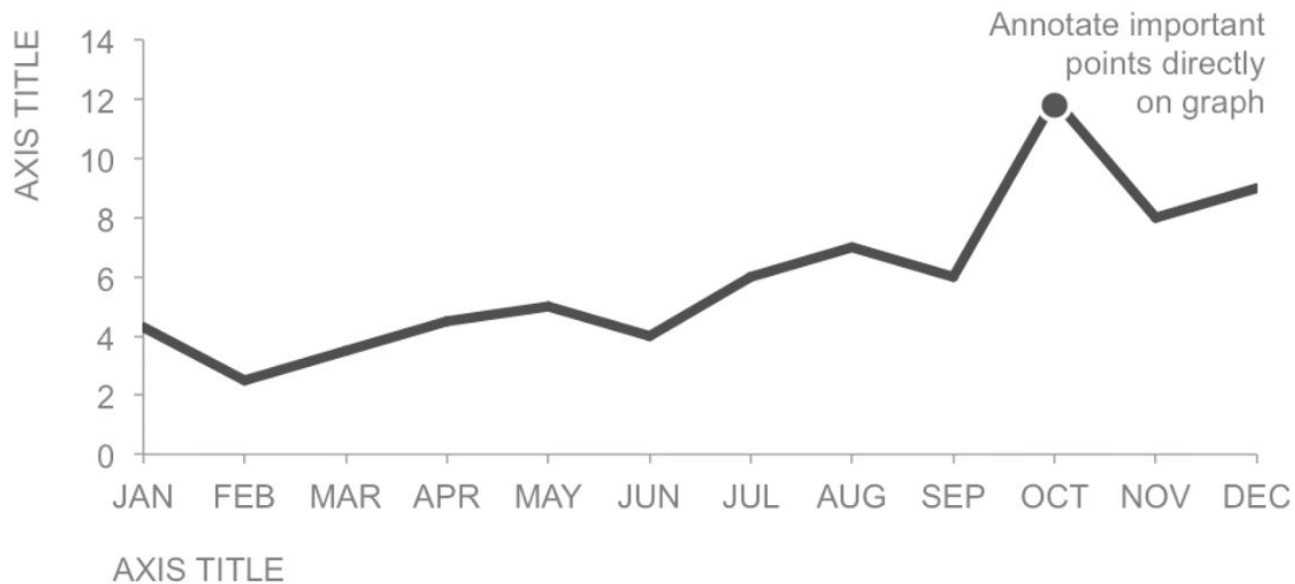- Goal here: to help reader identify general patterns of concentration of the Democratic and Republican votes.

How Much Each County Counted

The simple formula for winning an election is to get more votes than your opponent in as many counties as possible. It worked for President Bush. The map below, which uses the size of circles to indicate a candidate's winning margin, shows how it played out. *Senator John Kerry had huge margins in* many counties with large cities, and those margins were enough for him to win some of those states. However, Mr. Bush's relatively smaller but consistent margins in suburban and rural counties, in much of the South and West, helped him overcome Mr. Kerry's urban-county margins.

LARGEST VOTE
MARGIN FOR KERRY
Cook County, Ill.
+805,857 votes

New York City
Taken together, the five counties that make up New York City have a margin for Kerry of +1,109,408 votes.

2ND LARGEST
MARGIN FOR KERRY
Los Angeles Co., Calif.
+714,771 votes

LARGEST MARGIN FOR BUSH
Orange Co., Calif.
+155,010 votes

SECOND LARGEST VOTE
MARGIN FOR BUSH
Tarrant County, Tex.
+113,163 votes

Alaska shown by election district.

COUNTY WON BY ... BUSH   KERRY
VOTE MARGIN
+500,000
+100,000
+50,000
+10,000
+1,000

Results as of 2 p.m. yesterday.

Design Guidelines

storytellingwithdata.com

Text is your friend

Text is your Friend

- Every graph needs a tittle
- Every axis needs a title
- You can annotate important points directly on the graph
- Footnotes
  - Source, date, necessary assumptions or methodologies

Graph title

Annotate important points directly on graph

AXIS TITLE

14
12
10
8
6
4
2
0

JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

AXIS TITLE
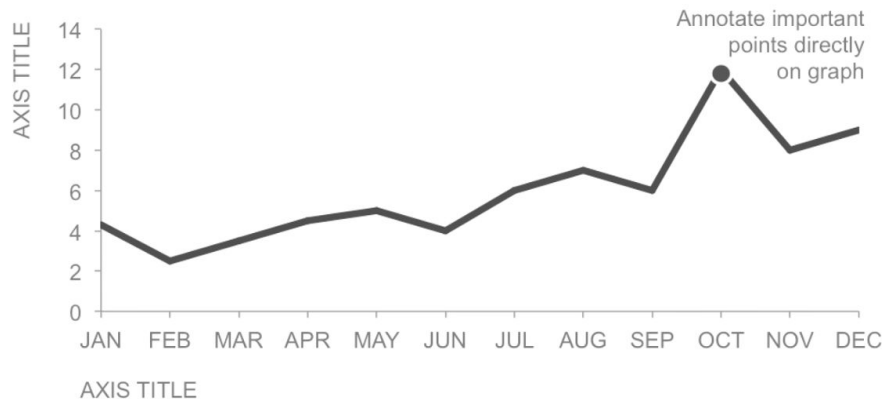
Any time you show data, you should have a footnote with the data source, as of date, and necessary assumptions and/or methodology (given your situation and audience).

Use this space wisely

Use this space wisely

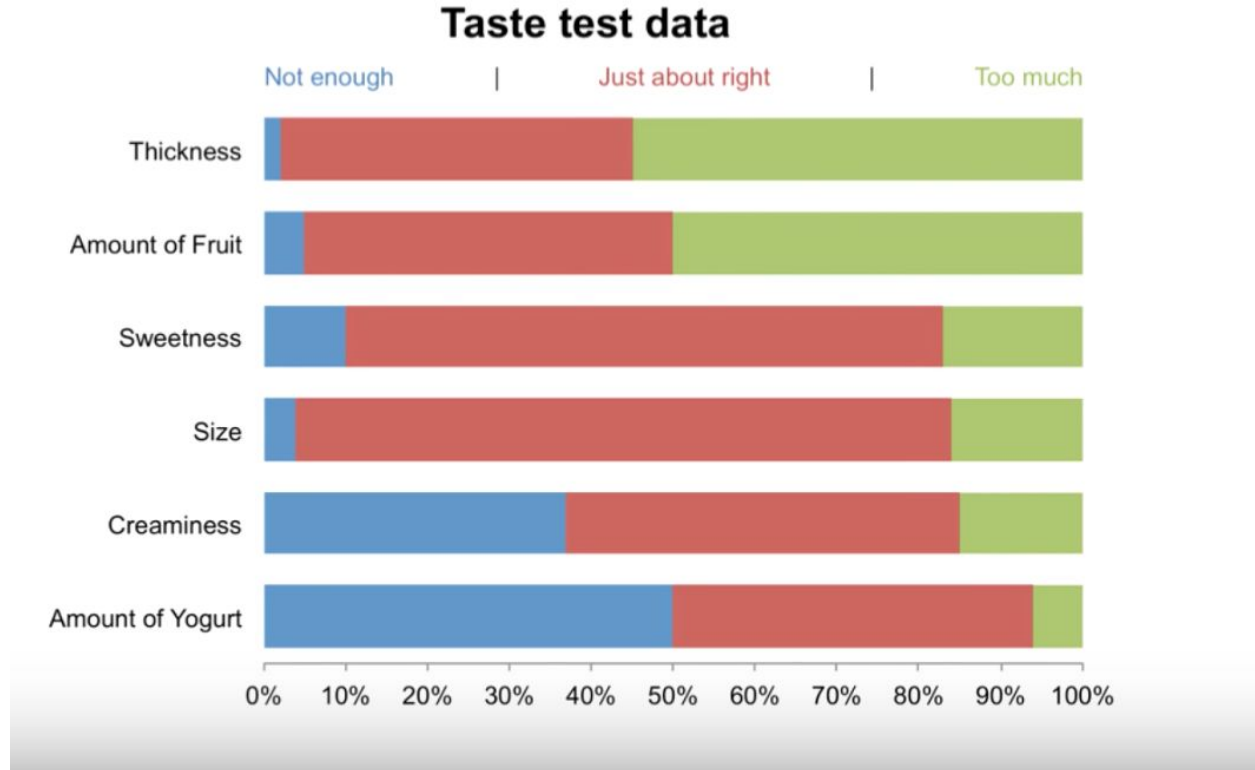| DESCRIPTIVE TITLE	ACTIVE TITLE
2016 Budget & Actuals	**Spending is on track** with budget
Market Conditions	**Favorable market conditions** continue
Sales Forecast	Sales **expected to exceed $4M** by Q3
Marketing Efforts	**Building brand awareness** in 5 ways
Executive Summary	**Action** or **decision** needed

Say what you mean

- Put your thoughts into words
- Strike a balance between spoken and written
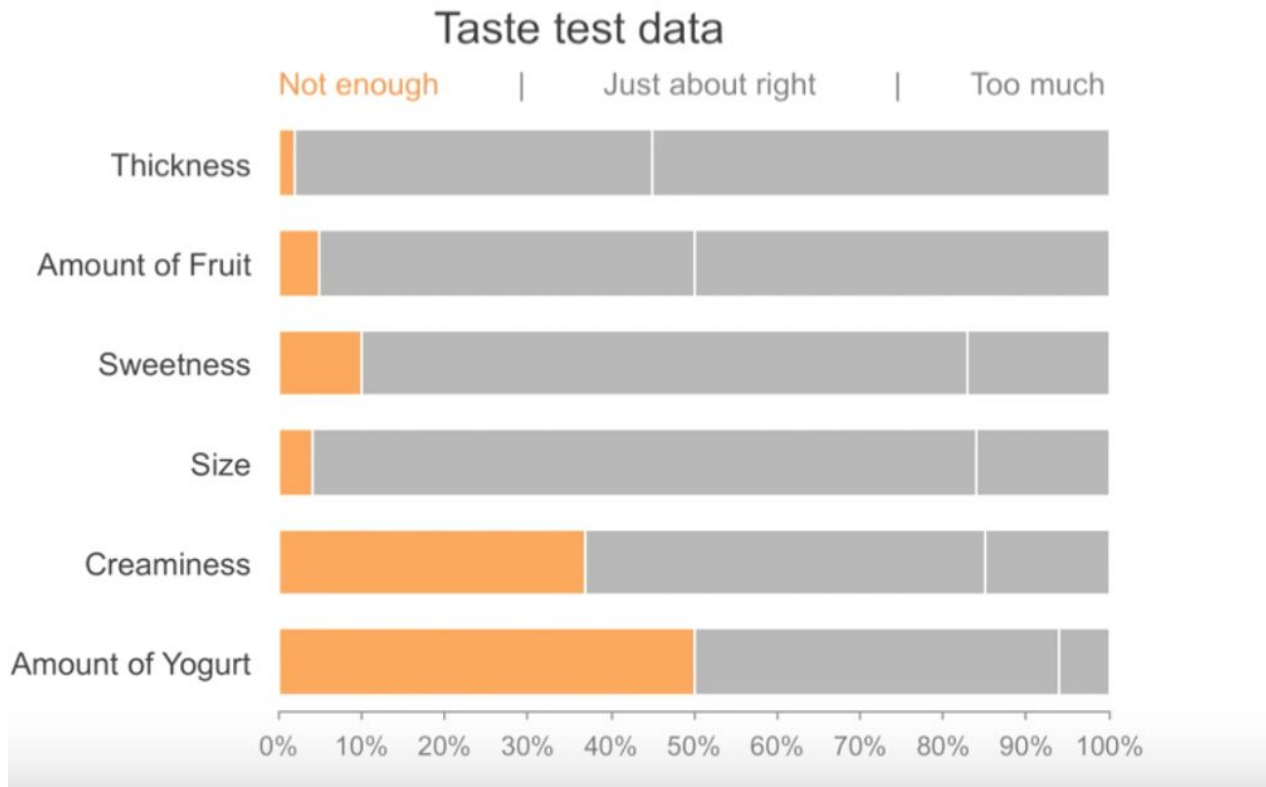- Avoid overly complex language -- simplify
- SOAAO*

*Spell Out Acronyms At Least Once

Be intentional in the use of **color**

Example: Using color smartly
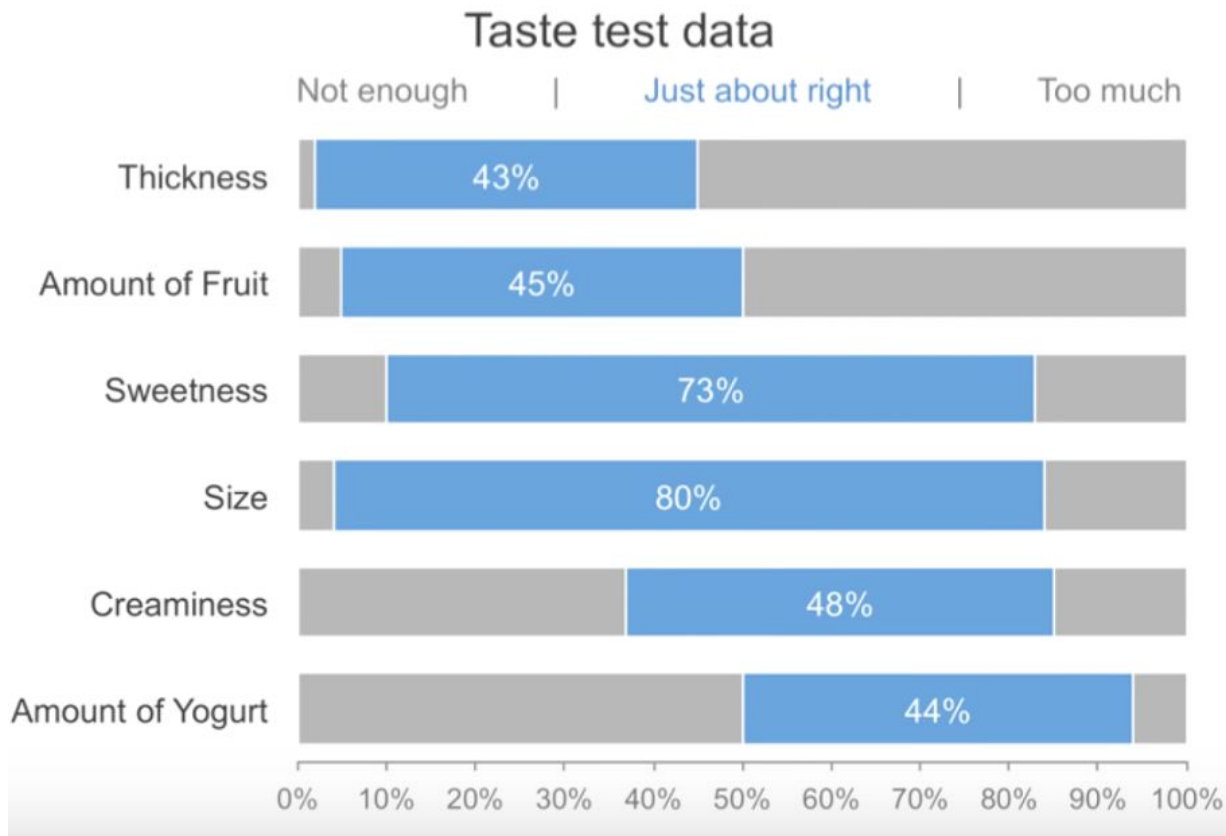

Taste test data

Not enough | Just about right | Too much

Thickness

Amount of Fruit

Sweetness

Size

Creaminess

Amount of Yogurt

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Example: Using color smartly

Taste test data

Example: Using color smartly

Taste test data

Not enough | Just about right | Too much

Thickness	43%
Amount of Fruit	45%
Sweetness	73%
Size	80%
Creaminess	48%
Amount of Yogurt	44%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Example: Using color smartly

Taste test data

Not enough | Just about right | Too much

Attribute	Too much %
Thickness	55%
Amount of Fruit	50%
Sweetness	17%
Size	16%
Creaminess	15%
Amount of Yogurt	6%

More yogurt less fruit to increase satisfaction

Consumer feedback on **Craveberry Yogurt**

DIMENSION	Not enough	Just about right	Too much
Thickness		43%	55%
Amount of Fruit		45%	50%
Sweetness	10%	73%	17%
Size		80%	16%
Creaminess	37%	48%	15%
Amount of Yogurt	50%	44%	

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

% OF TOTAL

WHERE WE'RE ON THE MARK: sweetness & size rated just about right by majority

ROOM FOR IMPROVEMENT: more than half think it's too thick

RECOMMEND: increase amount of yogurt and decrease amount of fruit to reduce overall thickness

N=100 | Source: 2015 Craveberry Taste Test Report | responses to item, *How would you rate Craveberry Yogurt' when it comes to the following factors?.*

Tell a story

Tell a story

OUR GOAL:

1 Share the **results of the Craveberry yogurt taste test** that was conducted last month.

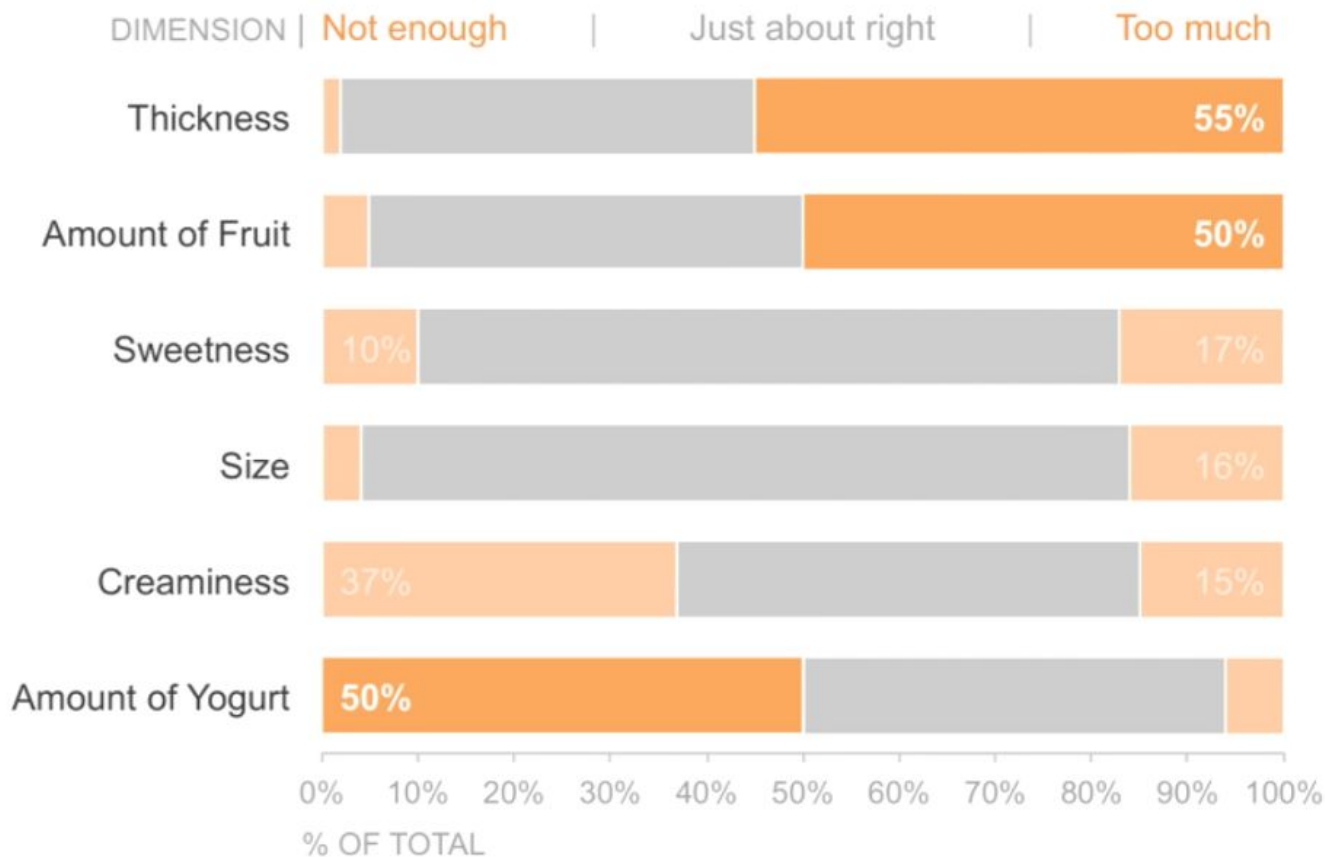2 Recommend a couple of minor changes we believe will have a **positive impact on consumer satisfaction.**

Tell a story

Consumer feedback on **Craveberry Yogurt**

	Not enough	Just about right	Too much
Thickness			
Amount of Fruit			
Sweetness			
Size			
Creaminess			
Amount of Yogurt			

Consumer feedback on **Craveberry Yogurt**

DIMENSION | Not enough | Just about right | Too much

Dimension	Not enough	Just about right	Too much
Thickness			55%
Amount of Fruit			50%
Sweetness	10%		17%
Size			16%
Creaminess	37%		15%
Amount of Yogurt	50%		

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

% OF TOTAL

N=100 | Source: 2015 Craveberry Taste Test Report | responses to item, *How would you rate Craveberry Yogurt'*
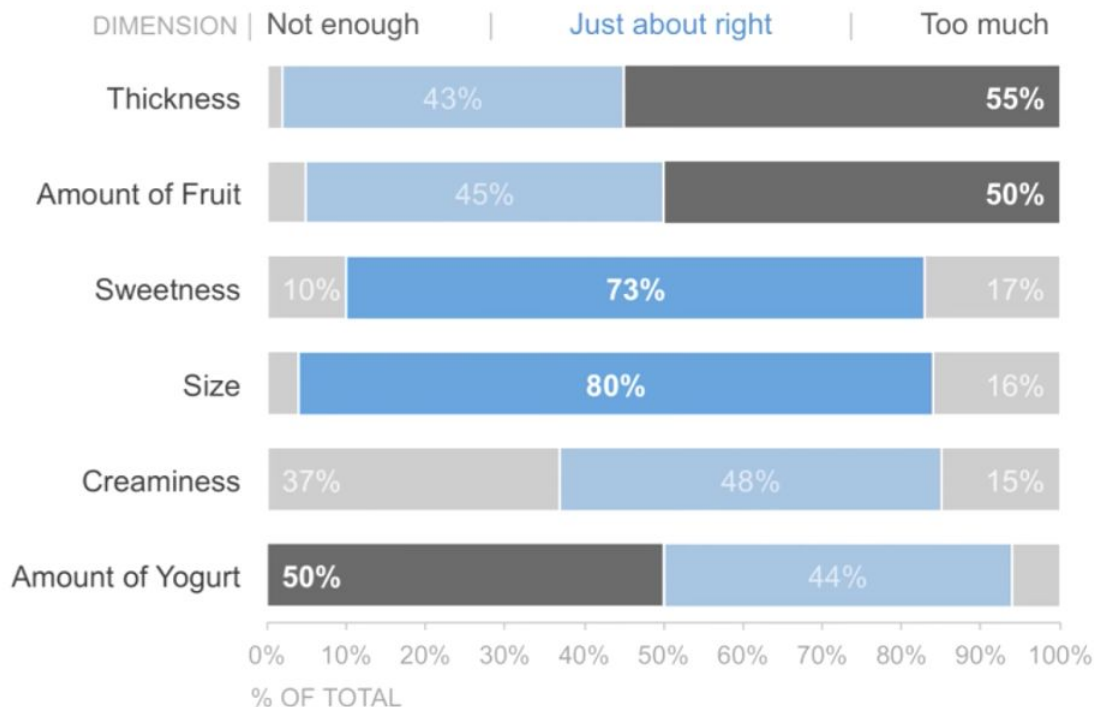*when it comes to the following factors?.*

More yogurt less fruit to increase satisfaction

Consumer feedback on **Craveberry Yogurt**

DIMENSION	Not enough	Just about right	Too much
Thickness		43%	55%
Amount of Fruit		45%	50%
Sweetness	10%	73%	17%
Size		80%	16%
Creaminess	37%	48%	15%
Amount of Yogurt	50%	44%	

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

% OF TOTAL

WHERE WE'RE ON THE MARK: sweetness & size rated just about right by majority
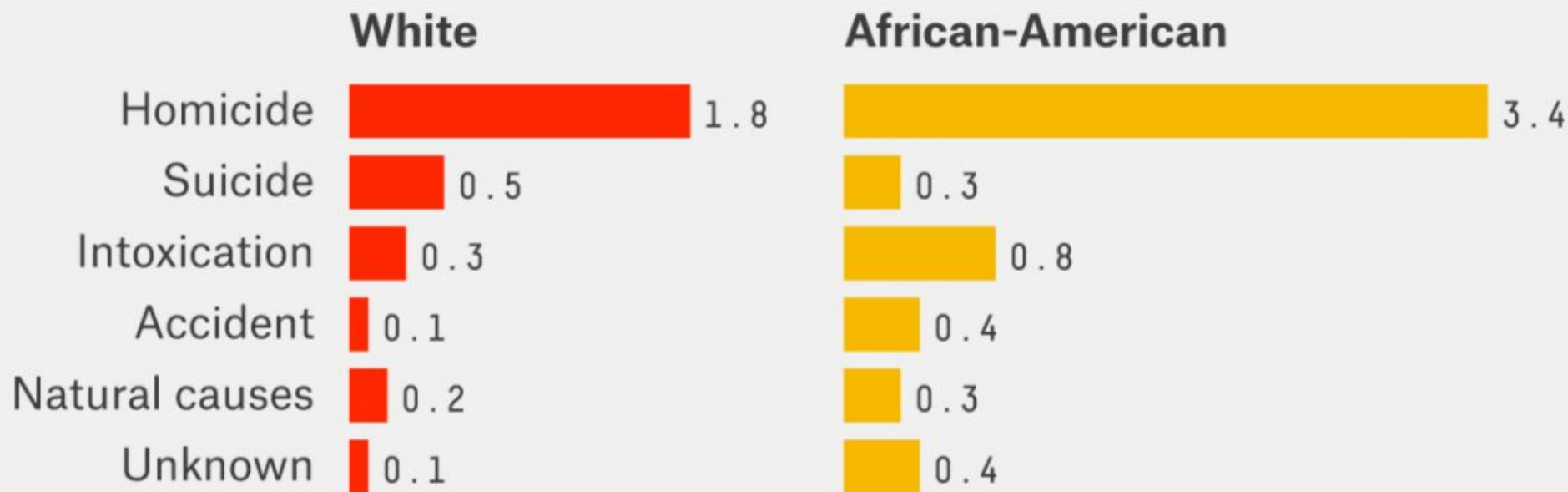
ROOM FOR IMPROVEMENT: more than half think it's too thick

RECOMMEND: increase amount of yogurt and decrease amount of fruit to reduce overall thickness

Align against a common baseline

Being Arrested Is Deadlier For African-Americans

Deaths per 100,000 arrests by race in the U.S., 2003-09

	White	African-American
Homicide	1.8	3.4
Suicide	0.5	0.3
Intoxication	0.3	0.8
Accident	0.1	0.4
Natural causes	0.2	0.3
Unknown	0.1	0.4

https://fivethirtyeight.com/datalab/being-arrested-is-nearly-twice-as-deadly-for-african-americans-as-whites/
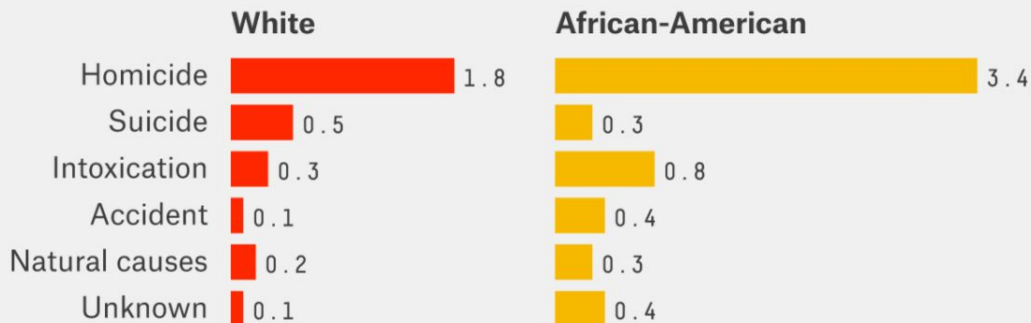
It is difficult to compare rate of death by type

- Great title
- Great data

Think about what you want your audience to be able to easily compare. Put those things as physically close together as you can and align them along a common baseline.

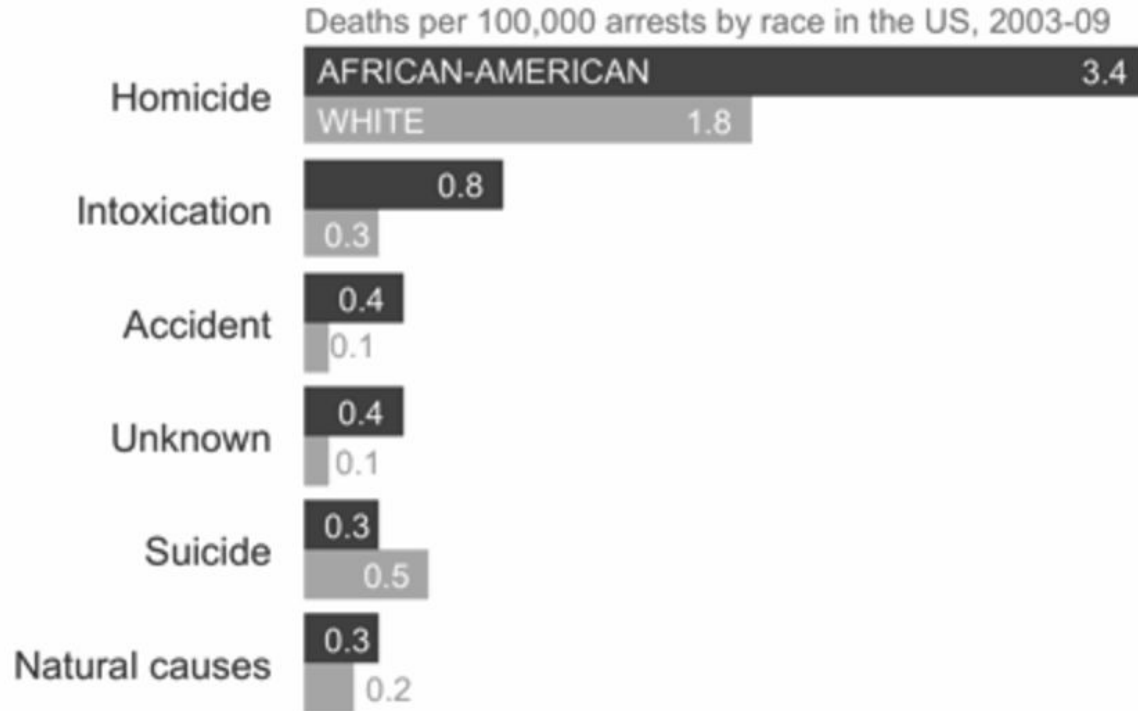**Being Arrested Is Deadlier For African-Americans**
Deaths per 100,000 arrests by race in the U.S., 2003-09

	White	African-American
Homicide	1.8	3.4
Suicide	0.5	0.3
Intoxication	0.3	0.8
Accident	0.1	0.4
Natural causes	0.2	0.3
Unknown	0.1	0.4

FIVETHIRTYEIGHT   SOURCE: BUREAU OF JUSTICE STATISTICS

https://fivethirtyeight.com/datalab/being-arrested-is-nearly-twice-as-deadly-for-african-americans-as-whites/

Being Arrested is Deadlier for African-Americans

Deaths per 100,000 arrests by race in the US, 2003-09

	AFRICAN-AMERICAN	WHITE
Homicide	3.4	1.8
Intoxication	0.8	0.3
Accident	0.4	0.1
Unknown	0.4	0.1
Suicide	0.3	0.5
Natural causes	0.3	0.2

- It is easier to do comparisons
- Categories have been rearranged in decreasing order

Summary

- Put your thoughts into words
  - Use an active title in the slice header
  - Use titles, axis titles, footnotes, annotations
- Be intentional in your use of color
  - Use grays, use color to emphasize your message
- Tell a story
- Align the things you want your audience to compare along a common baseline

References

- A. Cairo. The Functional Art. 2013. Chapter 2
- http://www.storytellingwithdata.com/blog/2016/6/2/the-cat-in-the-hat-knows-a-lot-about-data-visualization
- http://www.scribblelive.com/blog/2012/07/27/45-ways-to-communicate-two-quantities/
- https://s3.amazonaws.com/scribblelive-com-prod/wp-content/uploads/2012/07/poster75and37.pdf