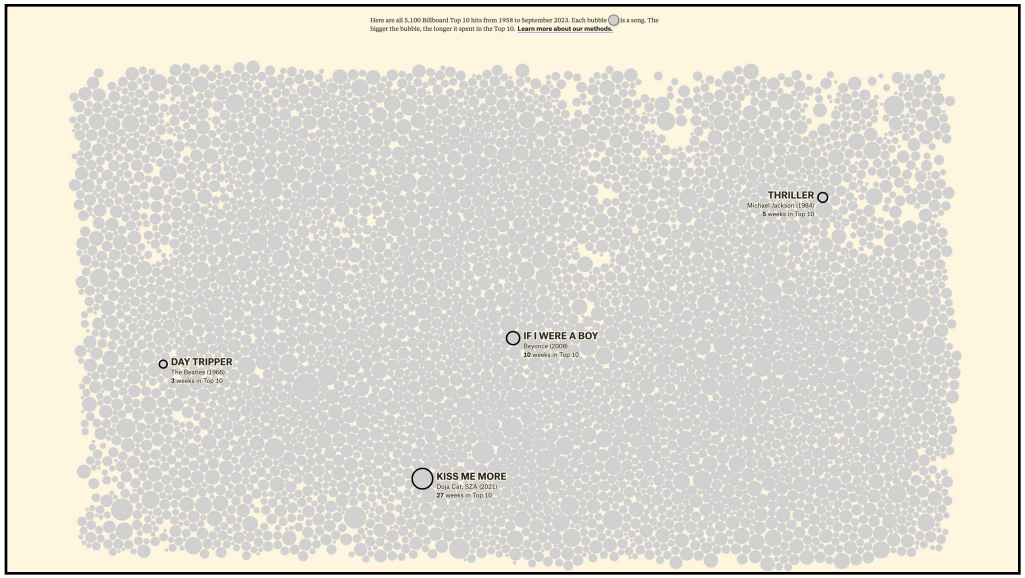
VISUALIZATION AND NATURAL LANGUAGE PROCESSING

CS 448B | Fall 2024

MANEESH AGRAWALA

1



2

“Words and pictures belong together.”

Tufte 1983

3

4

TODAY

Learning Objectives

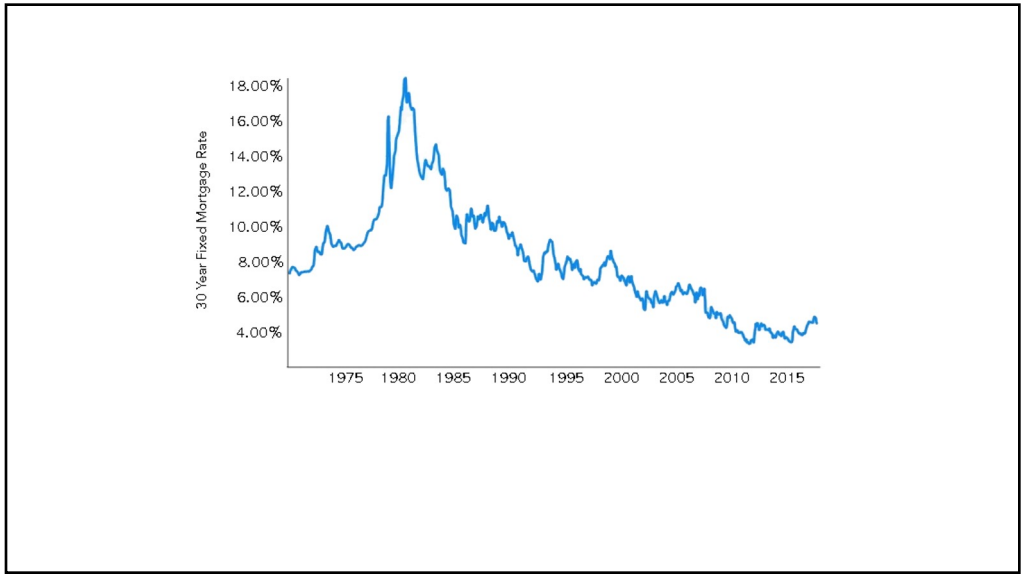
1. Understand how people read charts and text together.
2. Visual QA with explanations for charts and graphs.

5

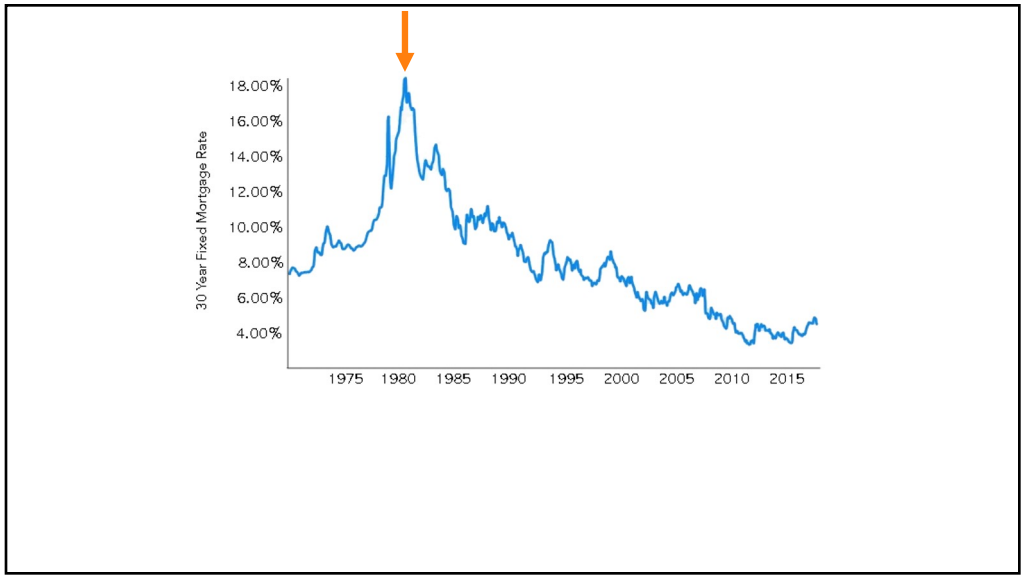
READING CHARTS AND TEXT

Towards Understanding How Readers Integrate Charts and Captions: A Case Study with Line Charts Kim, Setlur and Agrawala, CHI 2021

12


30 Year Fixed Mortgage Rate
18.00%
16.00%
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
1975
1980
1985
1990
1995
2000
2005
2010
2015

13


30 Year Fixed Mortgage Rate
18.00%
16.00%
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
1975
1980
1985
1990
1995
2000
2005
2010
2015

14

The 30-year fixed mortgage rate increased slightly from 1997 to 1999.

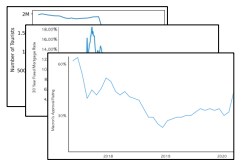
15

The 30-year fixed mortgage rate increased slightly from 1997 to 1999.

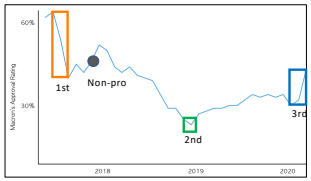
16

Do readers rely more on the chart or captions for their takeaways?

17



Collect Time-Series Charts



Identify Visually Prominent Features

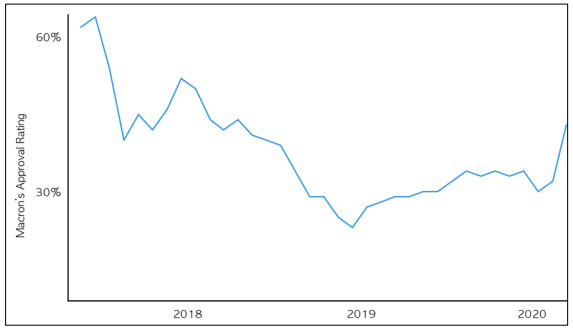
Macron's approval rating steeply dropped between June and August of 2017.

Generate Captions

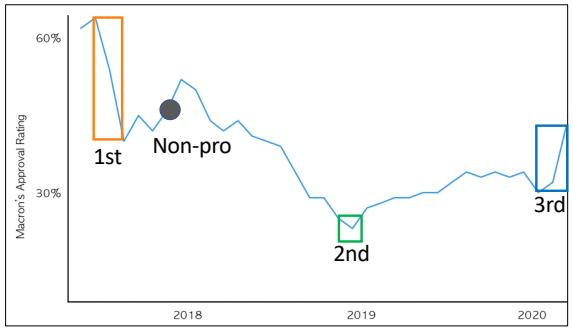
Subjects Viewed Chart-Caption Pairs

Reported Takeaways

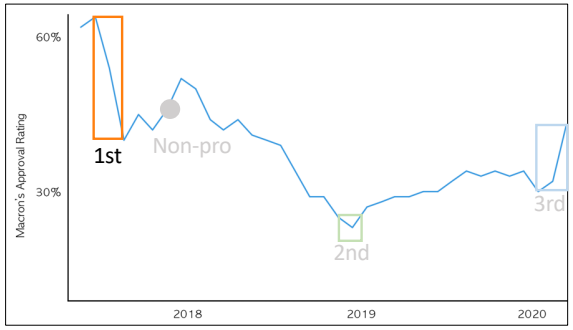
19


60%
30%
Macron's Approval Rating
2018
2019
2020

20


60%
30%
Macron's Approval Rating
1st
Non-pro
2nd
3rd
2018
2019
2020

21



[DOMAIN] [FEATURE] between [START DATE] and [END DATE].

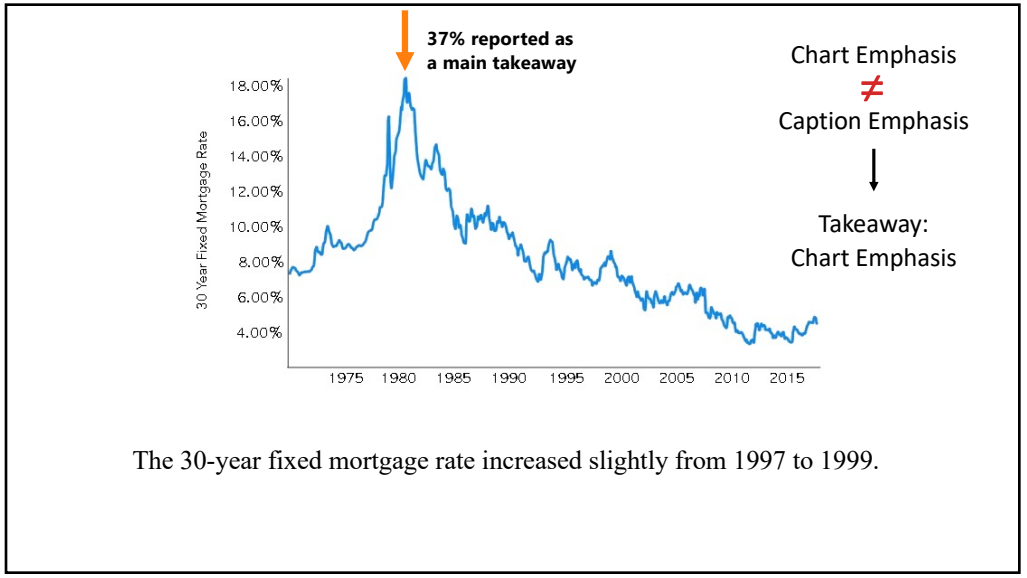
↓

Macron's approval rating steeply dropped between June and August of 2017.

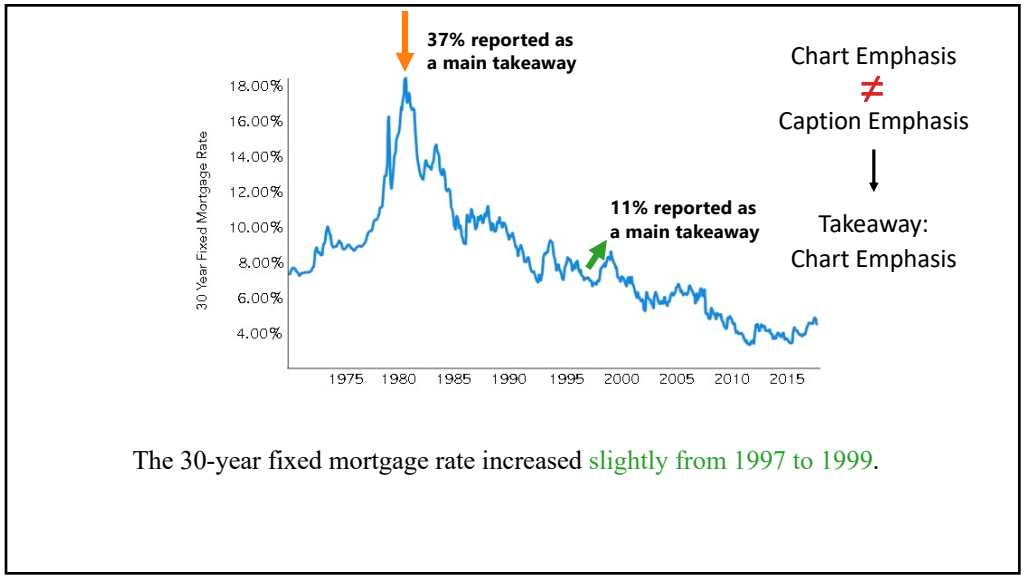
22

STUDY RESULTS

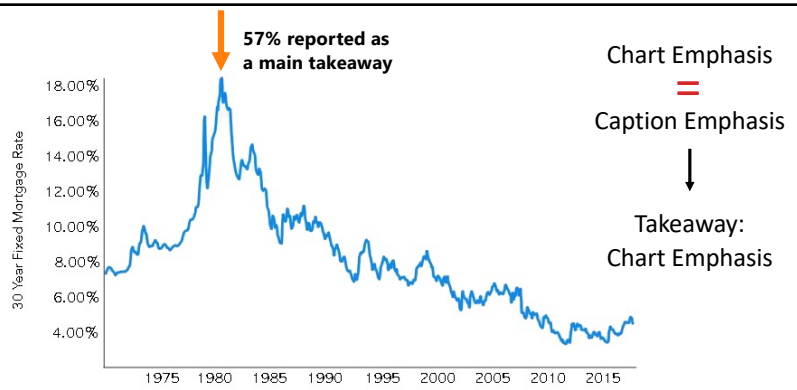
23


37% reported as
a main takeaway
30 Year Fixed Mortgage Rate
18.00%
16.00%
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
1975
1980
1985
1990
1995
2000
2005
2010
2015
Chart Emphasis
≠
Caption Emphasis
Takeaway:
Chart Emphasis
The 30-year fixed mortgage rate increased slightly from 1997 to 1999.

24


37% reported as
a main takeaway
11% reported as
a main takeaway
30 Year Fixed Mortgage Rate
18.00%
16.00%
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
1975
1980
1985
1990
1995
2000
2005
2010
2015
Chart Emphasis
≠
Caption Emphasis
Takeaway:
Chart Emphasis
The 30-year fixed mortgage rate increased slightly from 1997 to 1999.

25



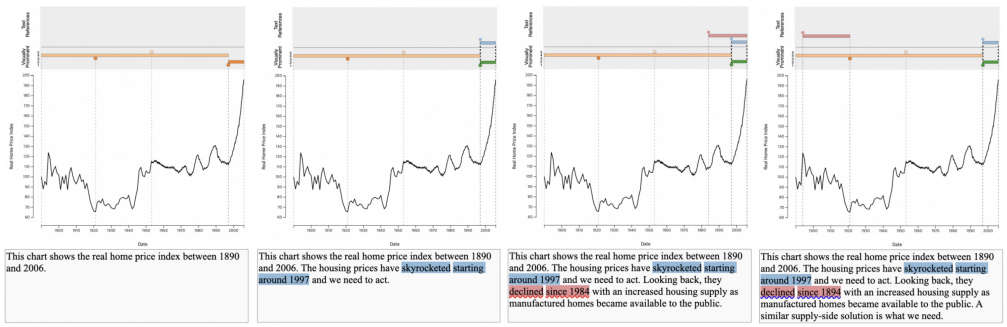
The 30-year fixed mortgage rate reached its peak of 18.5% in 1981.

26

Do readers rely more on the chart or captions for their takeaways?

When text and visualization emphasis mismatch, readers rely more on the chart and can miss information in the caption.

27



(a) Prominent features & Basic caption (b) Caption text about prominent feature (c) Caption including false information (d) Caption about less prominent feature

EmphasisChecker: A Tool for Guiding Chart and Caption Emphasis
Kim, Choi, Kim, Setlur and Agrawala, IEEE Visualization 2023

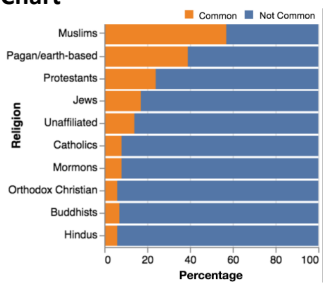
28

CHART QUESTION ANSWERING WITH VISUAL EXPLANATIONS

Answering Questions about Charts and Generating Visual Explanations Kim, Hoque and Agrawala, CHI 2020

29

Chart



Question

For which religion did the fewest chaplains think that religious extremism is common?

33

Chart

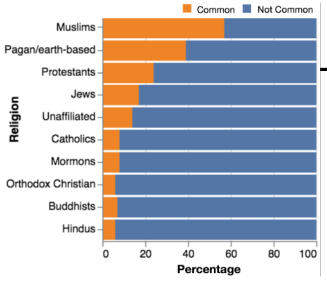


Chart QA

Answer

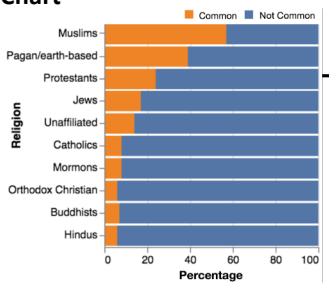
Orthodox Christians, Hindus.

Question

For which religion did the fewest chaplains think that religious extremism is common?

34

Chart



Question

For which religion did the fewest chaplains think that religious extremism is common?

Chart QA

Answer

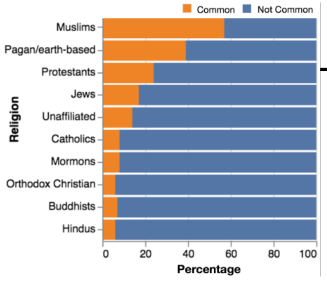
Orthodox Christians, Hindus.

Transparency?

Trust?

35

Chart



Question

For which religion did the fewest chaplains think that religious extremism is common?

Chart QA

Answer

Orthodox Christians, Hindus.

Explanation

I looked up 'Religion' for the shortest orange bars.

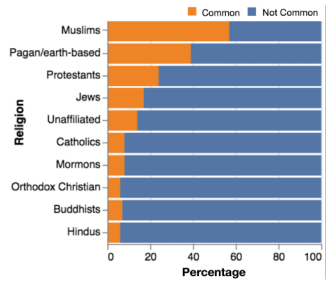
Transparency?

Trust?

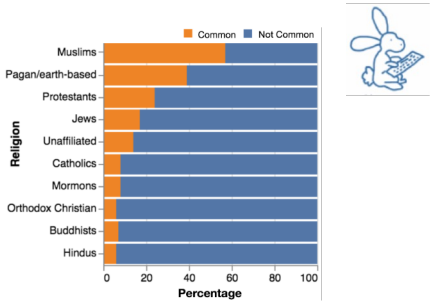
36

FORMATIVE STUDY

37

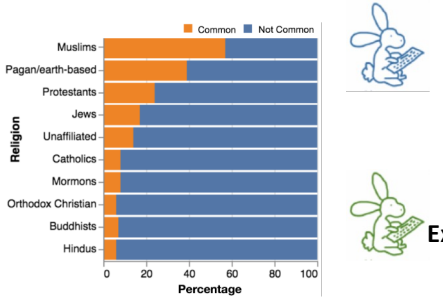


38



Question Which religion has the greatest value for Common?

39



Question Which religion has the greatest value for Common?

Answer Muslims

Explanation I picked religions with the greatest orange percentage.

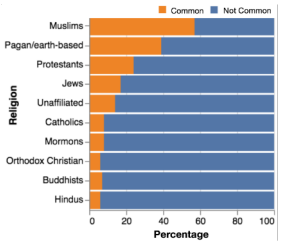
40

Explanations describe procedure for computing answer

41

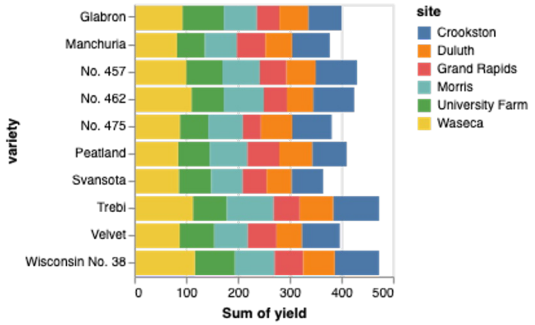
Explanations describe procedure for computing answer

Half of the explanations referred to visual features of chart

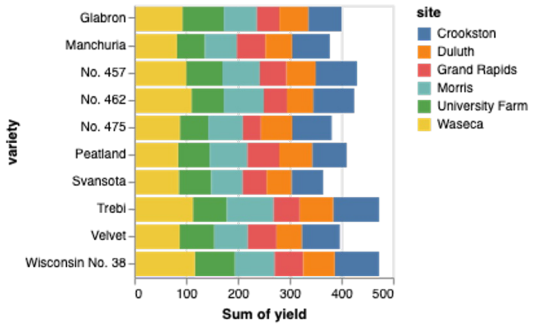


*"I picked religions with the greatest **orange** percentage."*

42

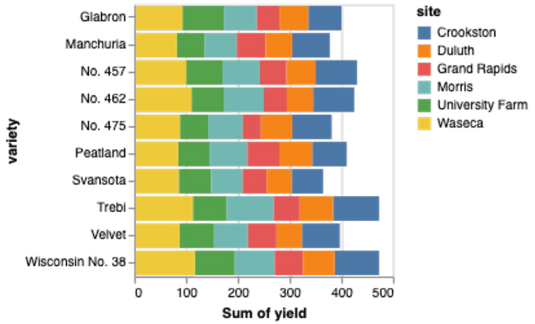


43



Green component in the top bar

44



Green component in the top bar

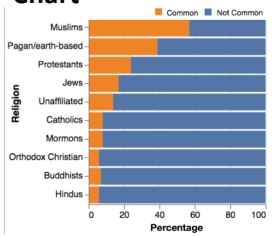
Glabron at University Farm

45

Chart QA Pipeline

46

Chart

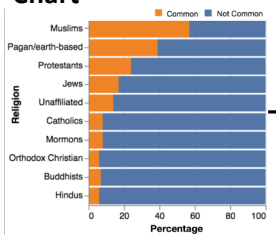


Question about Chart

Which religion has the shortest orange component?

55

Chart



Data Table

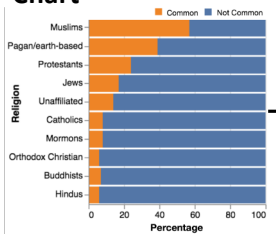
Religion	Common	Not common
Muslims	57	43
Pagan/earth-based	39	61
Protestants	24	76
Jews	17	83
⋮	⋮	⋮
Buddhists	7	93
Hindus	6	94

Question about Chart

Which religion has the shortest orange component?

56

Chart



Data Table

Religion	Common	Not common
Muslims	57	43
Pagan/earth-based	39	61
Protestants	24	76
Jews	17	83
⋮	⋮	⋮
Buddhists	7	93
Hindus	6	94

Question about Chart

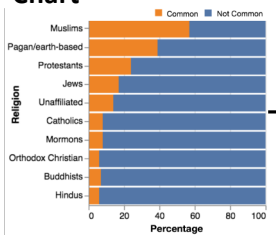
Which religion has the shortest orange component?

Question about Table

Which religion has the least 'Common' Percentage?

57

Chart



Data Table

Religion	Common	Not common
Muslims	57	43
Pagan/earth-based	39	61
Protestants	24	76
Jews	17	83
⋮	⋮	⋮
Buddhists	7	93
Hindus	6	94

Question about Chart

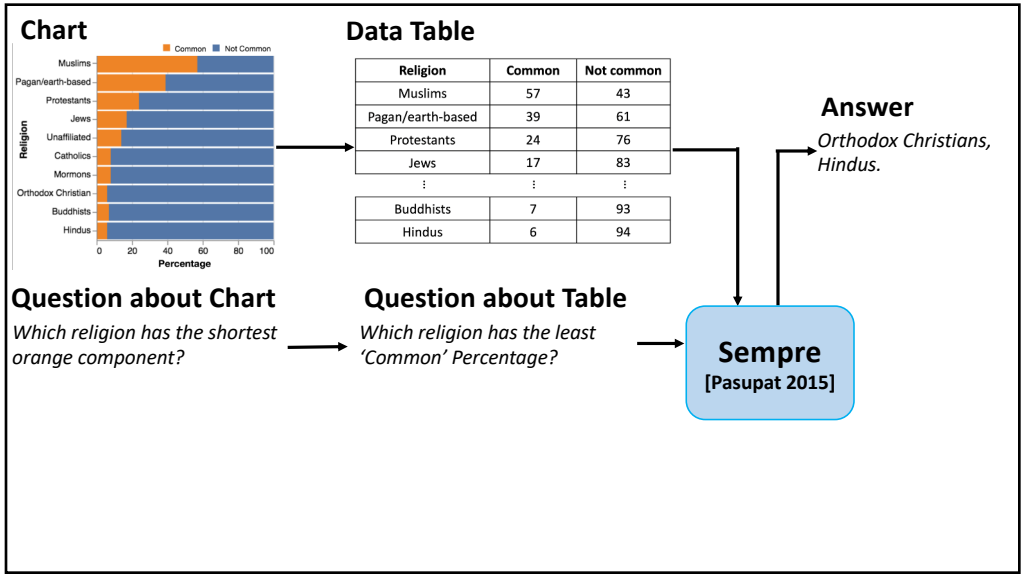
Which religion has the shortest orange component?

Question about Table

Which religion has the least 'Common' Percentage?

Sempre
[Pasupat 2015]

58



59

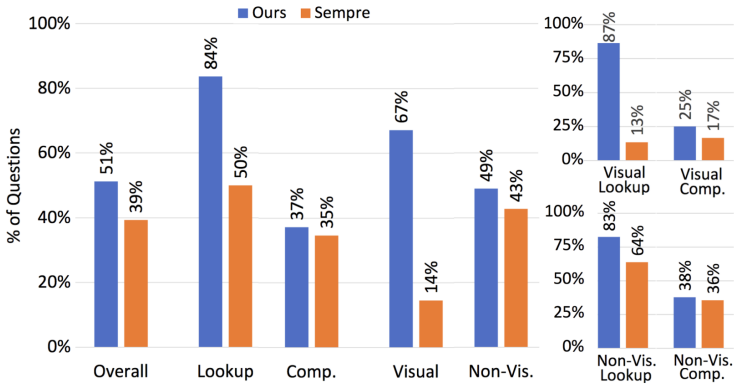


Figure 10. Accuracy of our pipeline (blue) compared to a baseline version of Sempre (orange) for questions of each type (visual/non-visual and lookup/compositional).

60

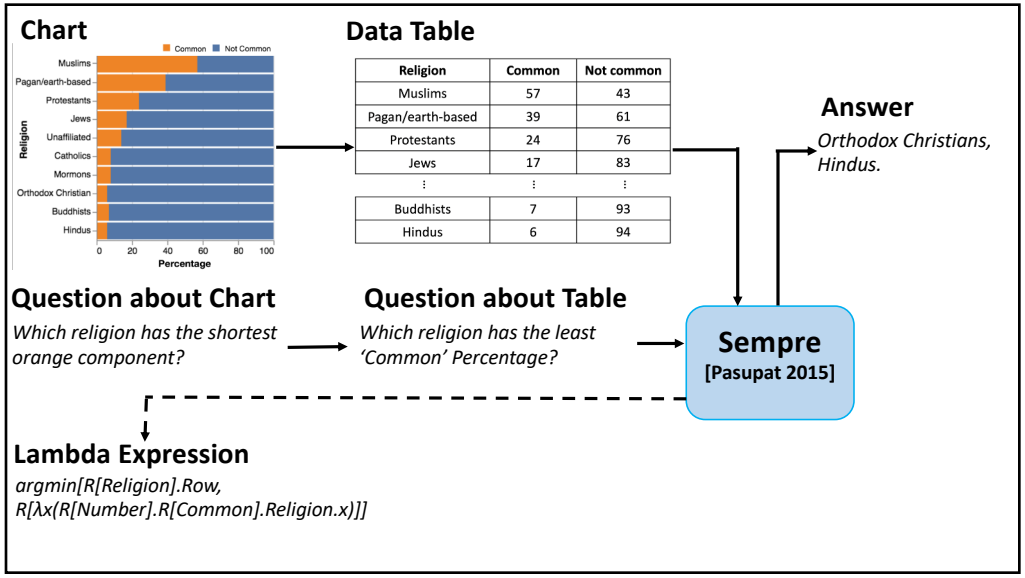


Chart
Common
Not Common
Muslims
Pagan/earth-based
Protestants
Jews
Unaffiliated
Catholics
Mormons
Orthodox Christian
Buddhists
Hindus
Religion
0
20
40
60
80
100
Percentage
Data Table
Religion | Common | Not common
Muslims | 57 | 43
Pagan/earth-based | 39 | 61
Protestants | 24 | 76
Jews | 17 | 83
Buddhists | 7 | 93
Hindus | 6 | 94
Answer
Orthodox Christians, Hindus.
Question about Chart
Which religion has the shortest orange component?
Question about Table
Which religion has the least 'Common' Percentage?
Sempre
[Pasupat 2015]
Lambda Expression
argmin[R[Religion].Row, R[λx(R[Number].R[Common].Religion.x)]]

61

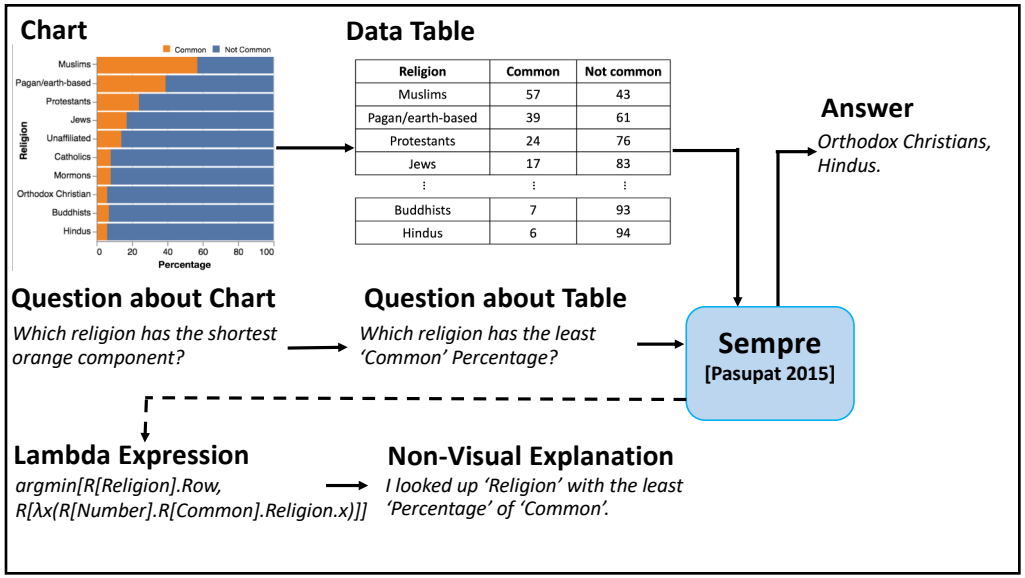
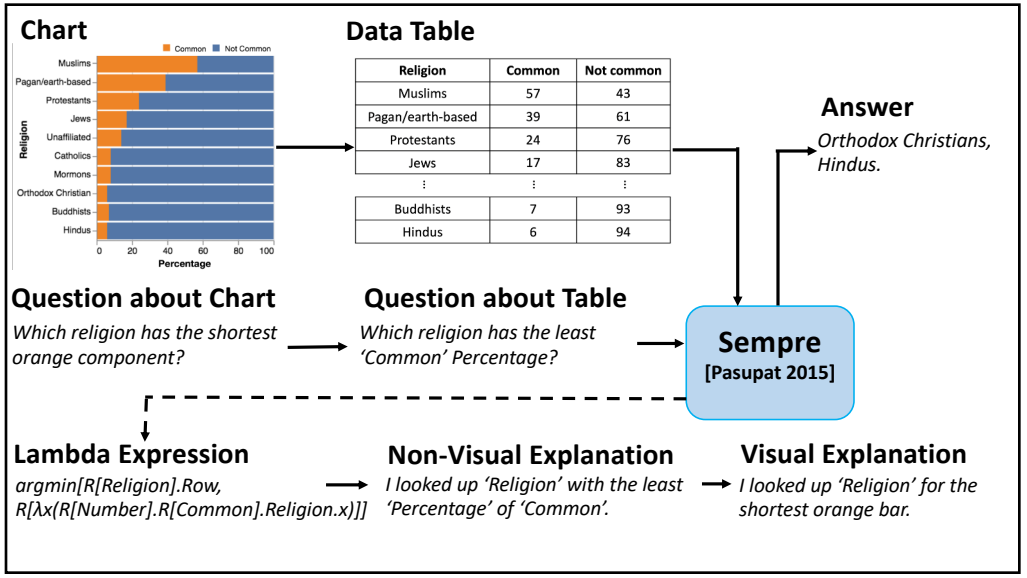


Chart
Common
Not Common
Muslims
Pagan/earth-based
Protestants
Jews
Unaffiliated
Catholics
Mormons
Orthodox Christian
Buddhists
Hindus
Religion
0
20
40
60
80
100
Percentage
Data Table
Religion | Common | Not common
Muslims | 57 | 43
Pagan/earth-based | 39 | 61
Protestants | 24 | 76
Jews | 17 | 83
Buddhists | 7 | 93
Hindus | 6 | 94
Answer
Orthodox Christians, Hindus.
Question about Chart
Which religion has the shortest orange component?
Question about Table
Which religion has the least 'Common' Percentage?
Sempre
[Pasupat 2015]
Lambda Expression
argmin[R[Religion].Row, R[λx(R[Number].R[Common].Religion.x)]]
Non-Visual Explanation
I looked up 'Religion' with the least 'Percentage' of 'Common'.

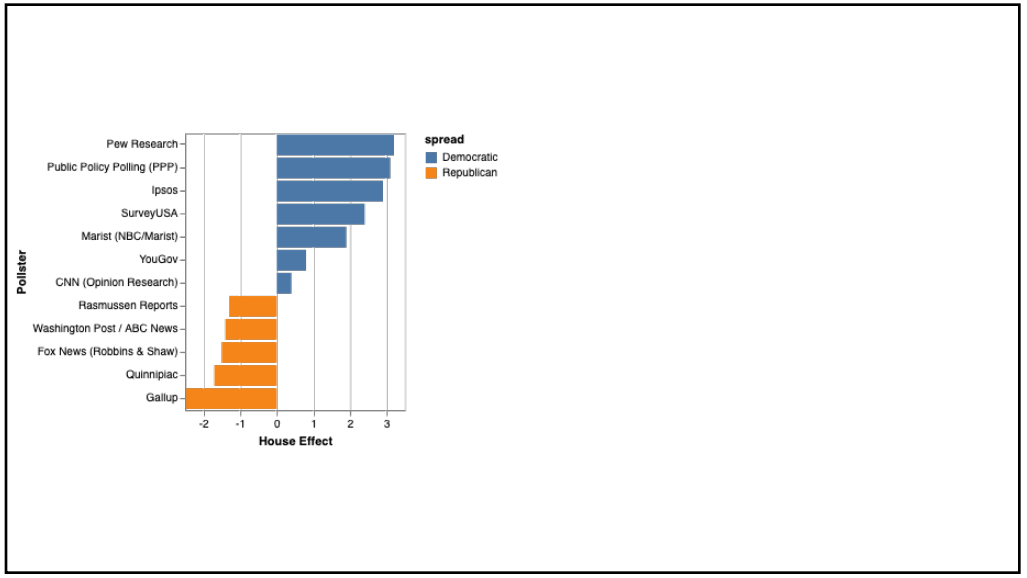
62



63

EXAMPLE EXPLANATIONS

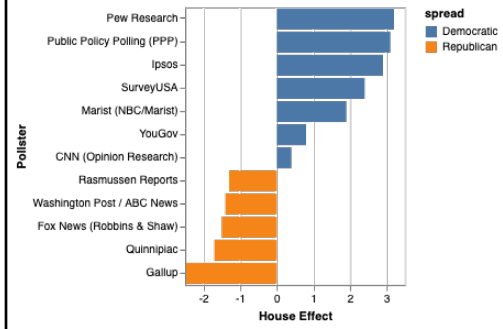
64



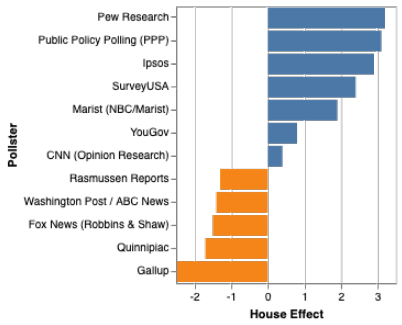
65

Question

What is the difference between the value of Gallup and Quinnipiac?



66



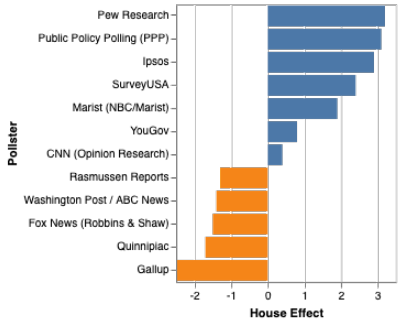
Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8

67



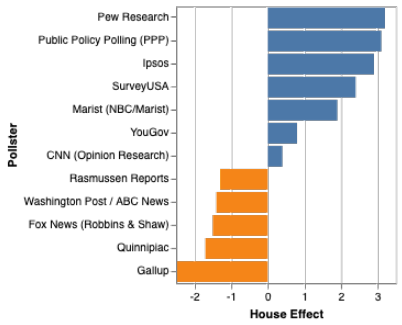
Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

68



Question

What is the difference between the value of Gallup and Quinnipiac?

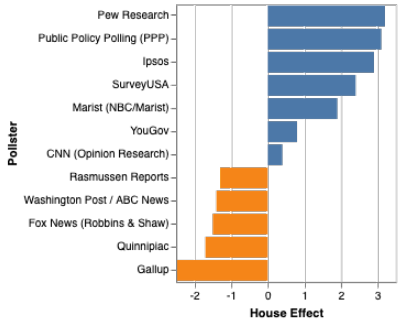
Answer

0.8 ✓

Explanation

I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

69



Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

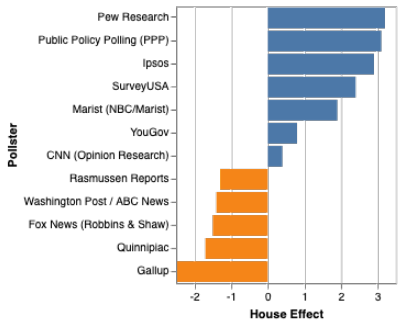
Explanation

I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

Question

What position is Ipsos in?

70



Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

Explanation

I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

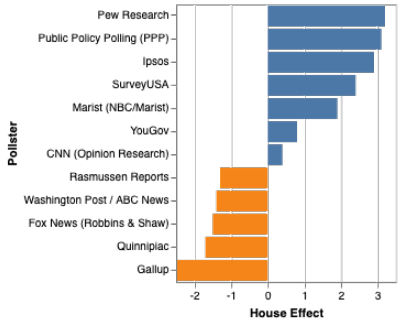
Question

What position is Ipsos in?

Answer

2.9

71



Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

Explanation

I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

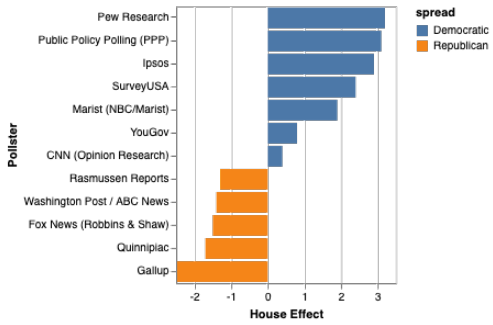
Question

What position is Ipsos in?

Answer

2.9 ✗

72



Question

What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

Explanation

I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

Question

What position is Ipsos in?

Answer

2.9 ✗

Explanation

I looked up the length of the bar for 'Ipsos'.

73

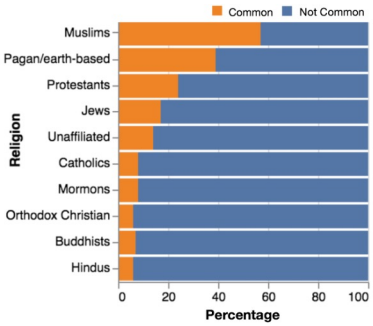
USER STUDY

74

Hypothesis: Visual explanations increase transparency and trust

75

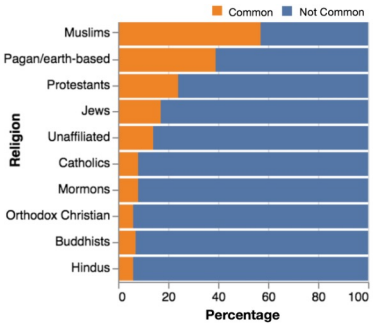
Hypothesis: Visual explanations increase transparency and trust



Q: *For which religion did the fewest chaplains think that religious extremism is common?*

76

Hypothesis: Visual explanations increase transparency and trust

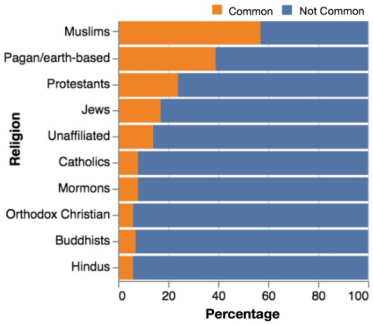


Q: *For which religion did the fewest chaplains think that religious extremism is common?*

A1 (vis): *Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.*

77

Hypothesis: Visual explanations increase transparency and trust



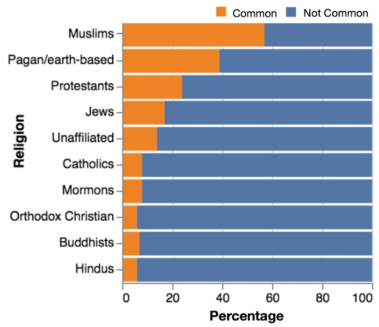
Q: *For which religion did the fewest chaplains think that religious extremism is common?*

A1 (vis): *Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.*

A2 (no-exp): *Orthodox Christians, Hindus.*

78

Hypothesis: Visual explanations increase transparency and trust



Q: *For which religion did the fewest chaplains think that religious extremism is common?*

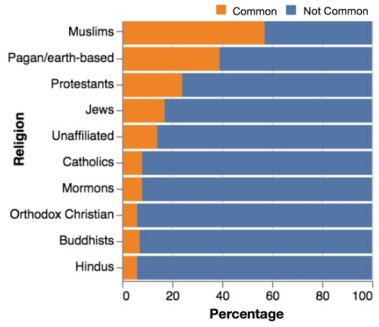
A1 (vis): *Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.*

A2 (no-exp): *Orthodox Christians, Hindus.*

A3 (non-vis): *Orthodox Christians, Hindus. I looked up 'Religion' with the lowest value for 'Common'.*

79

Hypothesis: Visual explanations increase transparency and trust



Q: *For which religion did the fewest chaplains think that religious extremism is common?*

A1 (vis): *Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.*

A2 (no-exp): *Orthodox Christians, Hindus.*

A3 (non-vis): *Orthodox Christians, Hindus. I looked up 'Religion' with the lowest value for 'Common'.*

A4 (human): *Orthodox Christians, Hindus. They have lowest values for 'Common'.*

80

16 participants

81

16 participants

20 Chart-Question-Answer-Explanation tuples (5 per explanation type)

82

16 participants

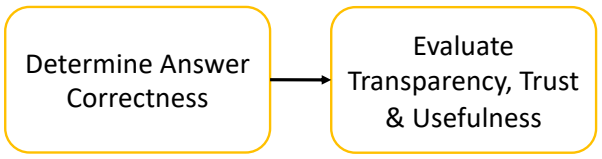
20 Chart-Question-Answer-Explanation tuples (5 per explanation type)

Determine Answer Correctness

83

16 participants

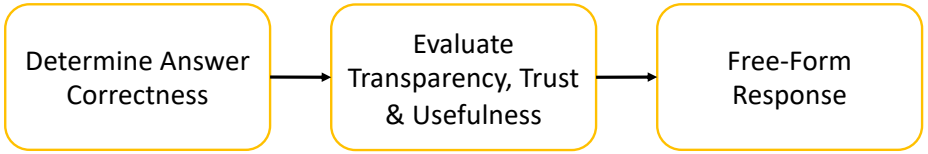
20 Chart-Question-Answer-Explanation tuples (5 per explanation type)



84

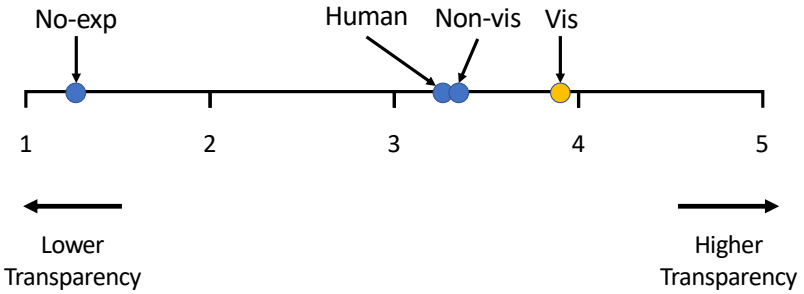
16 participants

20 Chart-Question-Answer-Explanation tuples (5 per explanation type)



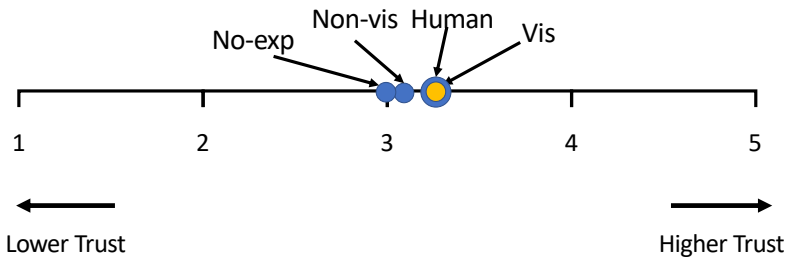
85

TRANSPARENCY



89

TRUST



94

Understanding how text and charts are related is difficult.

Tools that clarify connections between text and charts can help guide people towards the intended messages.

98

ANNOUNCEMENTS

99

FINAL PROJECT

Design Reviews Dec 2 and Dec 4 (signups this week)

Data analysis/explainer

Analyze dataset in depth & make a visual explainer

Deliverables

An article with multiple different interactive visualizations
Short video (2 min) demoing and explaining the project

Schedule

Design Review and Feedback: 10th week of quarter, 12/2 and 12/4
Final code and video: Sun 12/8 8pm

Grading

Groups of up to 3 people, graded individually
Clearly report responsibilities of each member

100

DESIGN REVIEW SIGNUPS

Sign up for 10 min slot with teaching team (over Zoom)

https://docs.google.com/spreadsheets/d/1IpS9oLbtPw2nTBvpfxqvAcseM5muDN8_ynAv4E0EvZjI/edit?usp=sharing

Mon 12/2 10:30am-12:20pm
2-4pm
Wed 12/4 10:30am-12:20pm

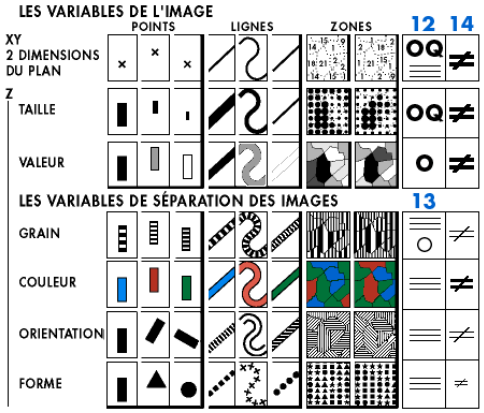
1. 5 min demo of current project state (timed)
2. Your questions (think about what you'd like feedback on)
3. Our feedback

101

COURSE SUMMARY

102

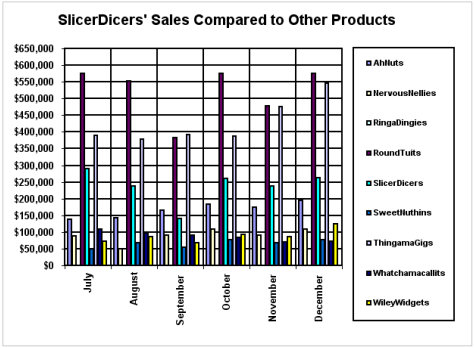
DATA AND IMAGE MODELS



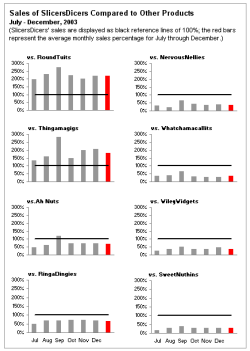
Graphics and Graphic Information Processing [Bertin 1981]

103

VISUALIZATION DESIGN AND REDESIGN



Problematic design



Redesign

104

EXPLORATORY DATA ANALYSIS

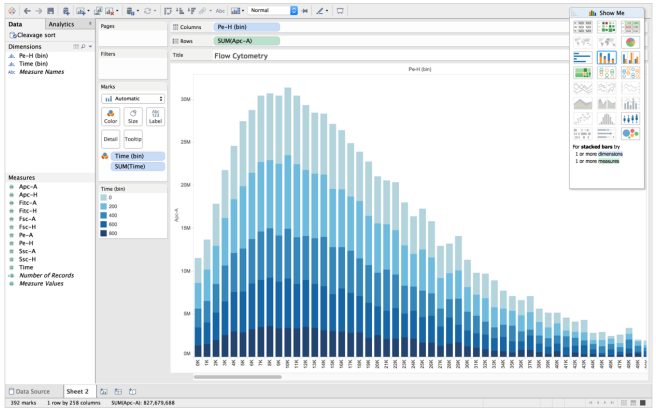
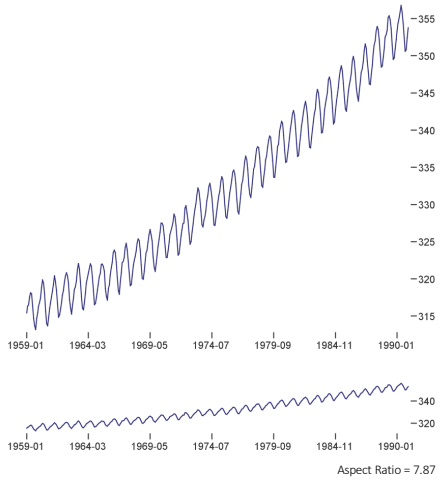


Tableau: based on Polaris [Stolte 2002]

105

PRINCIPLES OF CHART DESIGN

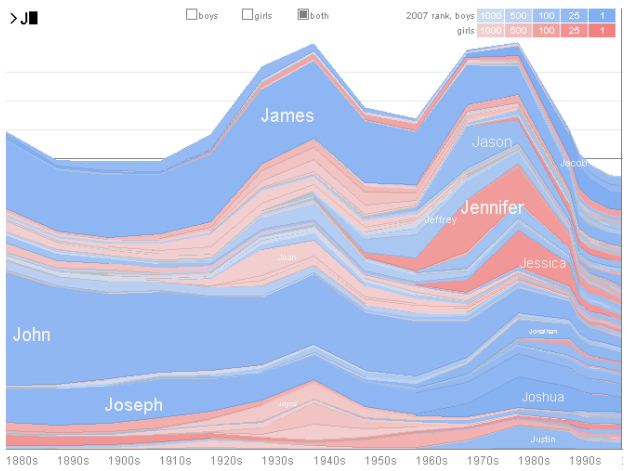


CO_2 Measurements
from William S. Cleveland
Visualizing Data, 1993

Trends may occur at different scales!
Apply banking to the original data or to fitted trend lines. [Heer & Agrawala 2006]

106

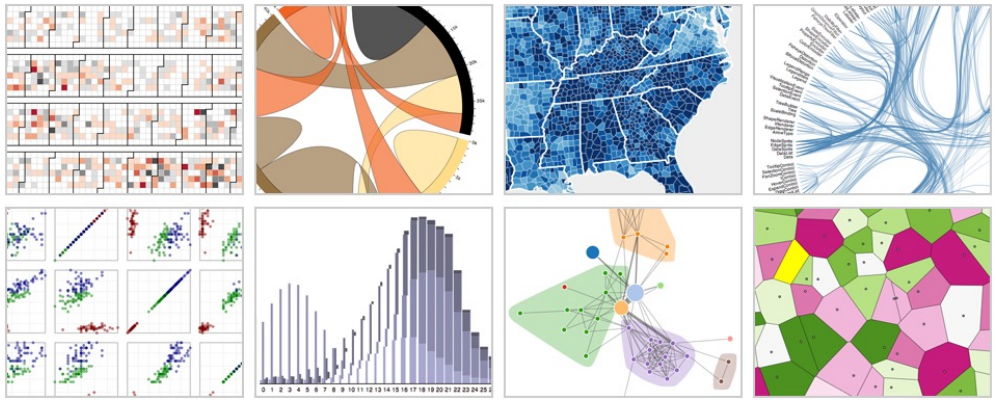
INTERACTION



Baby name voyager
[Wattenberg 2005]

107

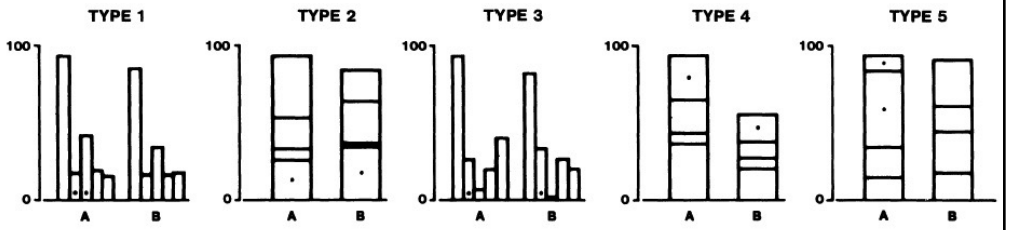
INTRODUCTION TO D3



D3: Data Driven Documents [Bostock 2011]

108

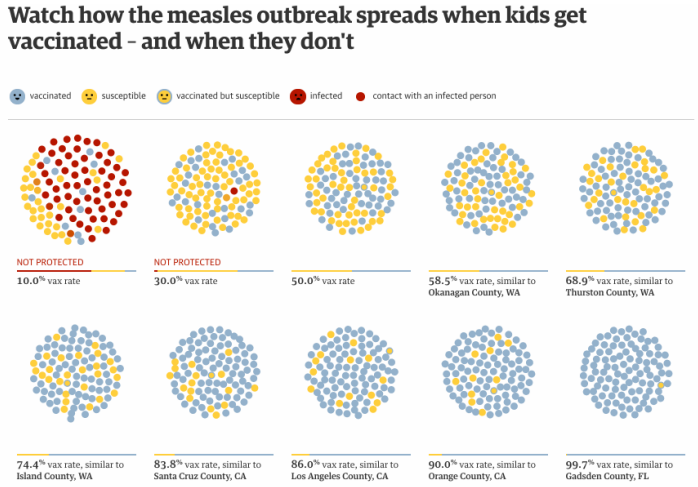
PERCEPTION



Graphical Perception [Cleveland and McGill 1984]

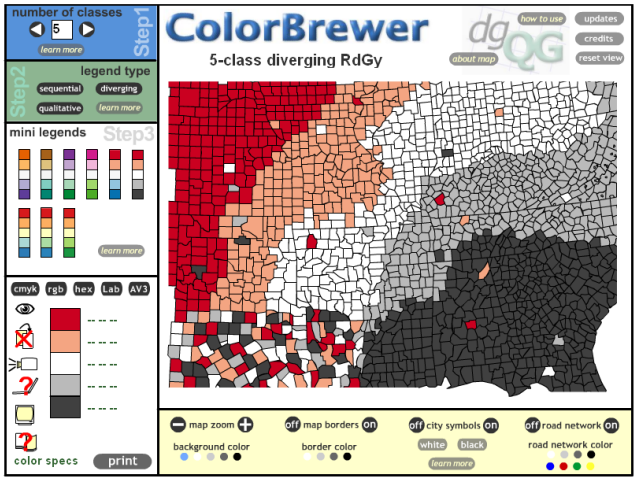
109

VISUAL EXPLAINERS



110

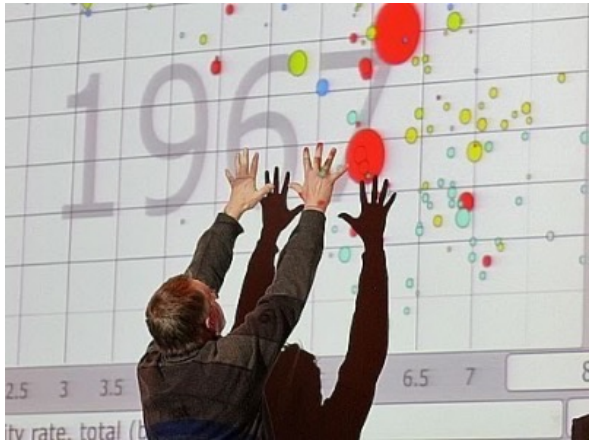
COLOR



Cynthia Brewer http://www.personal.psu.edu/faculty/c/a/cab38/

111

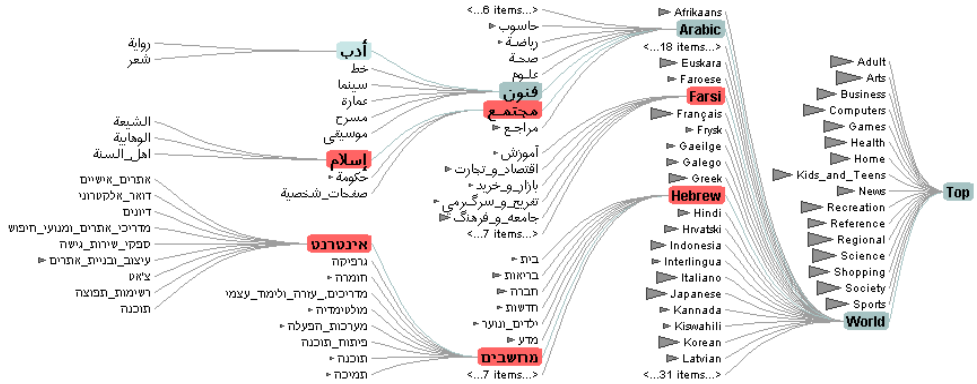
ANIMATION



Gapminder [Rosling 2006]

112

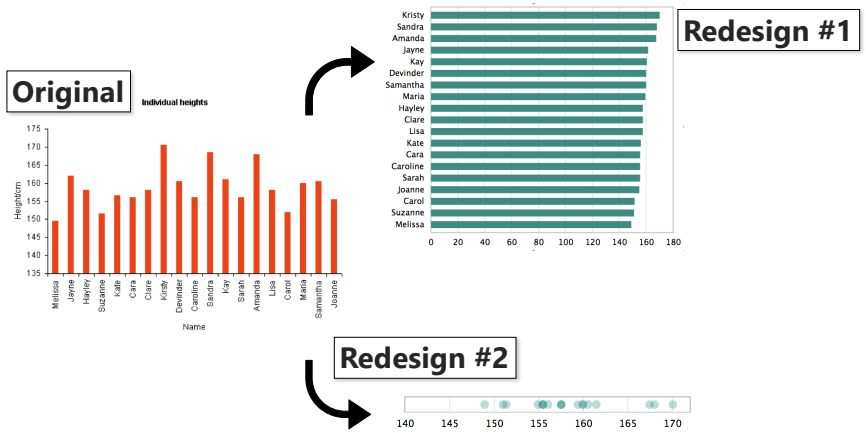
NETWORKS



Degree-of-Interest Trees [Heer 2004]

113

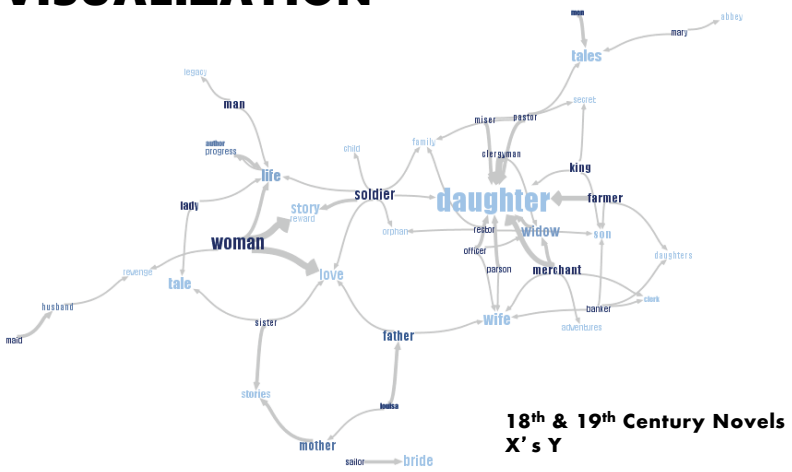
DECONSTRUCTING VISUALIZATIONS



ReVision [Savva 20011]

114

TEXT VISUALIZATION



115

THE FUTURE OF VISUALIZATION

Where is more work required?

What emerging technologies and societal trends will impact visualization design?

What did you find most difficult in designing visualizations? Could machines further facilitate design?

117

Thank You!

118

119