

Visualizations & NLP

Dae Hyun Kim, Vidya Setlur

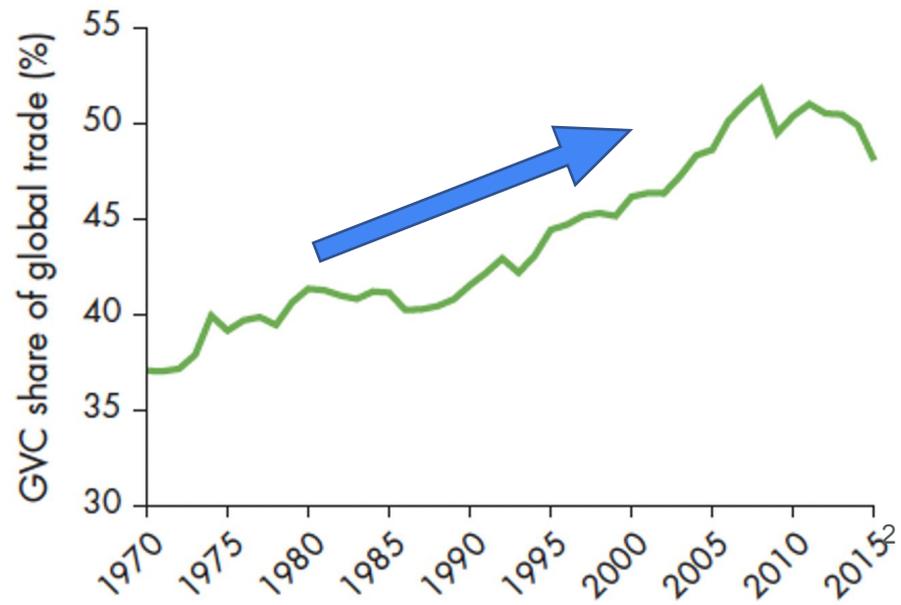


Figure 1.2 GVC trade grew rapidly in the 1990s but stagnated after the 2008 global financial crisis

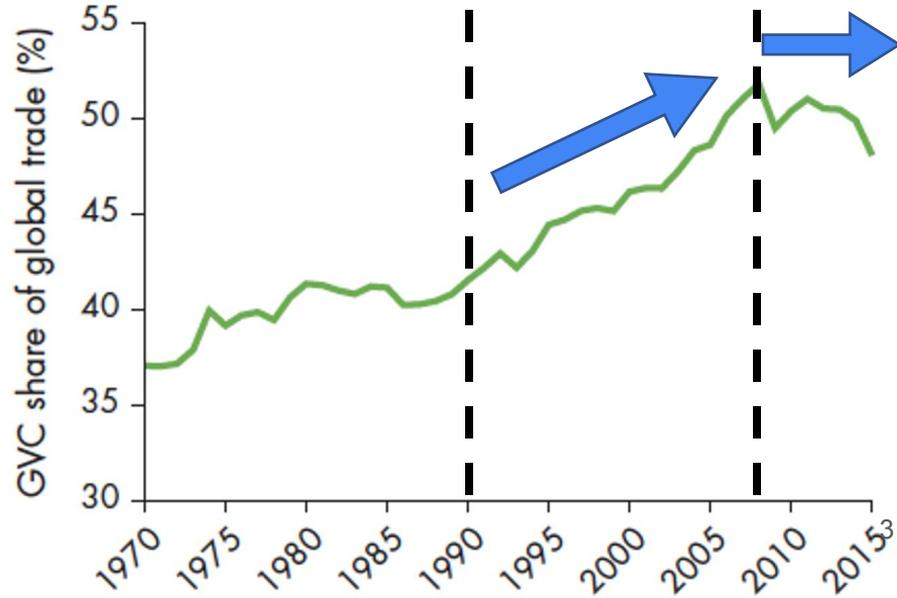
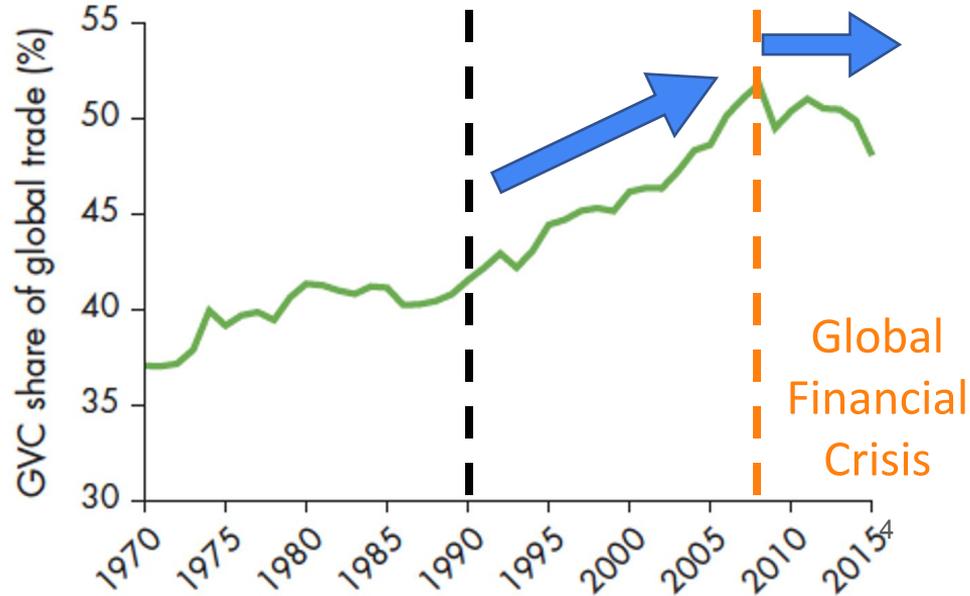


Figure 1.2 GVC trade grew rapidly in the 1990s but stagnated after the 2008 global financial crisis



Bookmarks

- Cover
- Title
- Copyright
- Contents
- Foreword
- Preface
- Acknowledgments
- Abbreviations
- Part I: Overview
- Part II: Global value chains: What are they?
- Part III: What are the effects of GVCs?
- Part IV: What domestic policies facilitate fruitful participation?
- Part V: How can international cooperation help?
- Appendix A: Databases used in this Report
- Appendix B: Glossary
- Boxes
- Figures
- Maps
- Tables

The evolution of GVC participation

The overall share of GVC trade in total world trade—encompassing both forward and backward linkages—grew significantly in the 1990s and early 2000s, but it appears to have stagnated or even declined in the last 10 years (figure 1.2). Still, about half of world trade appears to be related to GVCs.

What explains the remarkable rise in GVC participation in the 1990s and 2000s? And why has this process stalled since the financial crisis?

The global wave of fragmentation of production in the 1990s and 2000s was driven by a combination of factors. The information and communication technology (ICT) revolution brought forth cheaper and more reliable telecommunications, new information management software, and increasingly powerful personal computers (figure 1.3, panel a). Manufacturing firms then found it easier to outsource and coordinate complex activities at a distance and ensure the quality

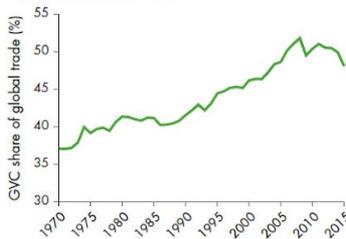
of their inputs. In addition, firms were able to disperse production across the world because transport costs fell significantly (figure 1.3, panel b). Declining air and sea freight costs boosted the trade in goods, while services benefited from cheaper communication costs.

Successive rounds of trade liberalization have resulted in rapidly falling barriers to trade and investment for both developed and developing countries. Tariffs have declined, especially for manufactured goods, and the gradual, although still insufficient, lowering of nontariff barriers has facilitated the international trade of goods and services (figure 1.4). Finally, the creation of the European single market—together with the integration of China, India, and the Soviet Union into the global economy—created huge new product and labor markets, and so firms could sell the same goods to more people and take advantage of economies of scale leading to the further deepening of GVCs. The new supply of cheap labor encouraged profit-seeking companies to either reallocate their production facilities or find local suppliers in low-wage countries.³

Since the global financial crisis in 2008, the dynamics of GVC expansion have changed. Trade has bounced back from its deep crisis level, but it has grown only marginally faster than output. Trade in parts and components also stalled after the financial crisis and even fell between 2011 and 2014, with a modest increase since then.

The factors behind the trade and GVC slowdown are both cyclical and structural in nature. On the one hand, trade growth is lower because global output growth is lower in economies that account for large shares of global trade and global output, such as Europe and China. Trade has also grown at a slower pace because the trade-to-income elasticity—defined as the amount of trade generated as output rises—has decreased. This is especially true in large trading countries, including China and the United States. China is producing more at home, thereby becoming less reliant on imported components for its exports. The share of intermediate imports in exports of Chinese goods dropped from about 50 percent in the 1990s to a little over 30 percent in 2015. In the United States, a boom-

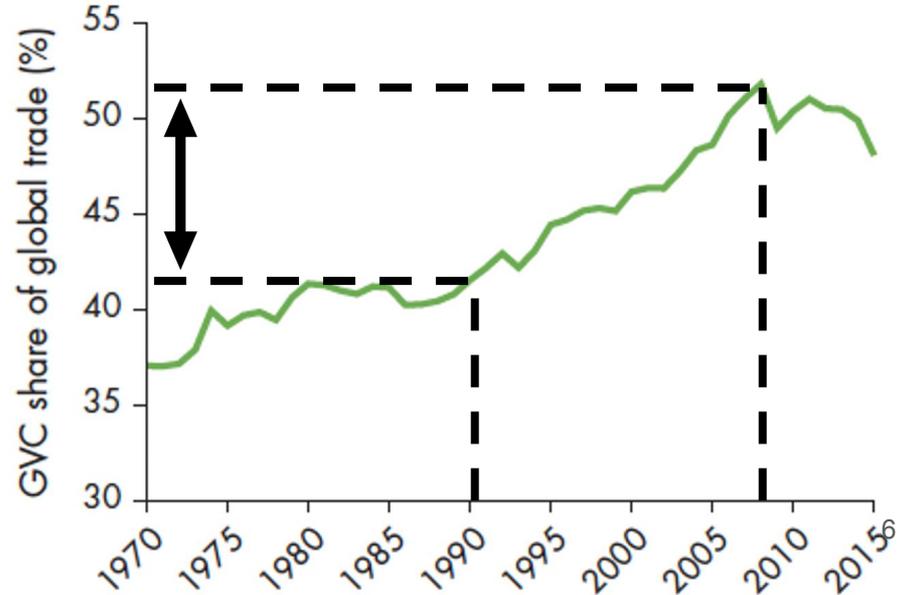
Figure 1.2 GVC trade grew rapidly in the 1990s but stagnated after the 2008 global financial crisis



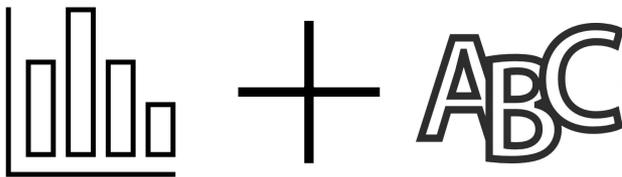
Sources: WDR 2020 team, using data from Eora26 database; Borin and Mancini (2015, 2019); and Johnson and Noguera (2017). See appendix A for a description of the databases used in this Report.

Note: Unless otherwise specified, GVC participation measures used in this and subsequent figures throughout the Report follow the methodology from Borin and Mancini (2015, 2019). The Eora26 database is used because it offers the largest country coverage: 190 countries between 1990 and 2015. GVC participation corresponds to the share of world exports that flow

Figure 1.2 GVC trade grew rapidly in the 1990s but stagnated after the 2008 global financial crisis



Question: *How much did the GVC share rise between 1990 and 2008?*



The evolution of GVC participation

The overall share of GVC trade in total world trade—encompassing both forward and backward linkages—grew significantly in the 1990s and early 2000s, but it appears to have stagnated or even declined in the last 10 years (figure 1.2, left), about half of world trade appears to be related to GVCs.

What explains the remarkable rise in GVC participation in the 1990s and 2000s? And why has this process stalled since the financial crisis?

The global wave of fragmentation of production in the 1990s and 2000s was driven by a combination of factors. The information and communications technology (ICT) revolution brought forth cheaper and more reliable telecommunications, new information management software, and increasingly powerful personal computers (figure 1.3, panel a). Manufacturing firms then found it easier to outsource and coordinate complex activities at a distance and ensure the quality of their inputs. In addition, firms were able to disperse production across the world because transport costs fell significantly (figure 1.3, panel b). Declining air and sea freight costs boosted the trade in goods, while services benefited from cheaper communication costs.

Successive rounds of trade liberalization have resulted in rapidly falling barriers to trade and investment for both developed and developing countries. Tariffs have declined, especially for manufactured goods, and the gradual, although still insufficient, lowering of content barriers has facilitated the international trade of goods and services (figure 1.4). Finally, the creation of the European single market—together with the integration of China, India, and the Soviet Union into the global economy—created huge new product and labor markets, and so firms could sell the same goods to more people and take advantage of economies of scale leading to the further deepening of GVCs. The new supply of cheap labor encouraged profit-seeking companies to either relocate their production facilities or find local suppliers in low-wage countries¹.

Since the global financial crisis in 2008, the dynamics of GVC expansion have changed. Trade has bounced back from its deep crisis level, but it has grown only marginally faster than output. Trade in parts and components also stalled after the financial crisis and even fell between 2011 and 2014, with a modest increase since then.

The factors behind the trade and GVC slowdown are both cyclical and structural in nature. On the one hand, trade growth is lower because global output growth is lower in economies that account for large shares of global trade and global output, such as Europe and China. Trade has also grown at a slower pace because the trade-to-income elasticity—defined as the amount of trade generated as output rises—has decreased. This is especially true in large trading countries, including China and the United States. China is producing more at home, thereby becoming less reliant on imported components for its exports. The share of intermediate imports in exports of Chinese goods dropped from about 50 percent in the 1990s to a little over 30 percent in 2013. In the United States, a boom

Figure 1.2 GVC trade grew rapidly in the 1990s but stagnated after the 2008 global financial crisis

Sources: WIOD 2020Growth, using data from Eurostat Database, Brinck and Malmgren (2019, 2018), and Johnson and Noguera (2012). See appendix A.4 for a description of the data and units used in this report.

Note: Unless otherwise specified, GVC participation measures cover this and subsequent chapters throughout the report. Where the manufacturing, non-ferrous metals and chemical (MNC) sector is used as a benchmark, it refers to the largest country coverage (50 countries) between 1993 and 2015. GVC participation is measured in the three most developed regions that



Q:

How much did the GVC share rise between 1990 and 2008?



“Words and pictures belong together.”

[Tufte, 1983]

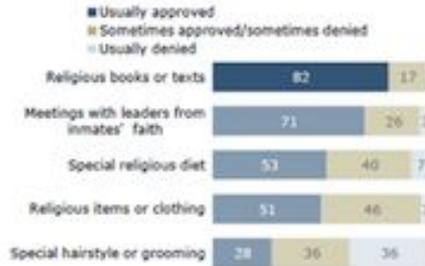
Facilitating Document Reading by Linking Text and Tables

[Kim et al. 2018]

An overwhelming majority of chaplains who responded to these questions say that inmates' requests for religious texts (82%) and for meetings with spiritual leaders of their faith (71%) are usually approved. And about half of chaplains say that requests for a special religious diet (53%) or for permission to have sacred items or religious clothing such as crucifixes, eagle feathers and turbans (51%) also are usually granted.

Requests for Religious Accommodation

% saying requests from inmates for each of the following are ...

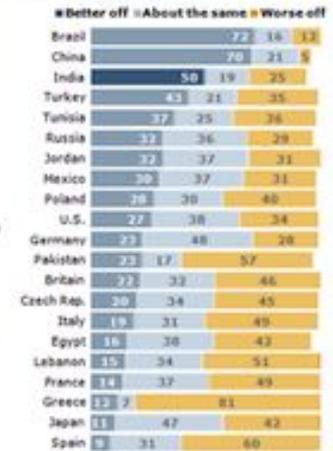


Q29a-e. Based on all answering. Those who responded that the request had not come up or did not give an answer are excluded. Figures may not add to 100% due to rounding.

PEW RESEARCH CENTER'S FORUM ON RELIGION & PUBLIC LIFE

People may think their personal situation is better than economic conditions in their nation, but only in Brazil (72%) and China (70%) do large majorities think their families are better off than they were five years ago. On balance, Indians (50%) and Turks (43%) also say their situations have improved.

Compared to Five Years Ago, Are You Financially...



PEW RESEARCH CENTER Q30

[Kong et al. 2014]

Part 6

Issues about computers and the internet: Awareness, interest, attitudes, aptitude, self-confidence

The 2006 Internet Project has gathered data on the extent to which awareness, interest and self-confidence in handling computers and the internet has increased since 2001. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

The data in this section suggests that as more information and services about the world of technology and the internet are available, there is a corresponding increase in awareness, interest and self-confidence in handling computers and the internet. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

Men and women have an appreciation of what the internet does for them, particularly in seeking their own information and expanding their world of information. Men were more likely than women to use the internet to find information, and women were more likely than men to use the internet to find information. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

Tech terminology and issues. More men than women are familiar with tech terms and issues.

In July 2006, we asked women about their familiarity with tech terms and issues. The results show that women are more familiar with tech terms and issues than men are. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

In February 2006, we asked men about their familiarity with tech terms and issues. The results show that men are more familiar with tech terms and issues than women are. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

In June 2006, we asked men about their familiarity with tech terms and issues. The results show that men are more familiar with tech terms and issues than women are. This data is presented in a series of questions, and shows that general interest in technology and the internet has been rising since 2001, and that self-confidence in handling computers and the internet has also risen since 2001.

Men's and women's awareness of technology: Issues

Percentage of internet users who have a good idea of what the term means

Survey	% of internet users	% of internet women
Spam	90%	81%
Firewall	80%	73%
Spiders	80%	74%
Internet backbone	74%	63%
Adware	67%	64%
Phishing	50%	23%
Podcasting	18%	11%
RSS feeds	17%	8%

Automatically extract references between sentences and tables

Automatic Reference Extraction Pipeline

Automatic Reference Extraction Pipeline

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.

Automatic Reference Extraction Pipeline

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.



Table Structure
Extraction

Automatic Reference Extraction Pipeline

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.

Table Structure
Extraction

Match
Text to Cells



Automatic Reference Extraction Pipeline

Women are more likely than men to cite some reasons for not using the internet

Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.

Table Structure
Extraction

Match
Text to Cells

Rule-Based
Refinement



Automatic Reference Extraction Pipeline

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.

Table Structure
Extraction

Match
Text to Cells

Rule-Based
Refinement

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Equal numbers of men and women said they lack time.

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Equal numbers of men and women said they lack time.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Diagram labels:

- Title Cell: Points to the green header row.
- Column Headers: Points to the pink header row.
- Data Cells: Points to the three data rows.
- Row Headers: Points to the first column of the data rows.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Diagram labels:

- Title Cell: Points to the green header row.
- Column Headers: Points to the second row of the table.
- Data Cells: Points to the three data rows.
- Row Headers: Points to the first column of the table.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Diagram labels:

- Title Cell: Points to the top row of the table.
- Column Headers: Points to the second and third rows of the table.
- Data Cells: Points to the bottom three rows of the table.
- Row Headers: Points to the first column of the table.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Diagram labels:

- Title Cell: The top row of the table.
- Column Headers: The second row of the table.
- Data Cells: The three rows of data in the table body.
- Row Headers: The first column of the table body.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45	58
Don't have time	29	29
Too expensive	25	34

Diagram labels:

- Title Cell: The green header row.
- Column Headers: The pink header row.
- Data Cells: The three rows of data (orange and yellow).
- Row Headers: The first column of the data rows.

Stage 1: Table Structure Extraction

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Diagram labels:

- Title Cell: The green header row.
- Column Headers: The pink header row.
- Data Cells: The three rows of data below the column headers.
- Row Headers: The first column of the data rows.

Stage 2: Match Sentence Text to Table Cells

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Equal numbers of **men** and **women** said they lack time.

Stage 2: Match Sentence Text to Table Cells

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Equal numbers of **men** and **women** said they **lack time**.

Stage 3: Rule-based Refinement of Matches

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Equal numbers of **men** and **women** said they **lack time**.

Stage 3: Rule-based Refinement of Matches

Women are more likely than men to cite some reasons for not using the internet		
Major reasons	% of online men	% of online women
Don't need it	45%	58%
Don't have time	29%	29%
Too expensive	25%	34%

Equal numbers of men and women said they lack time.

Pipeline Evaluation

Pipeline Evaluation

- Corpus

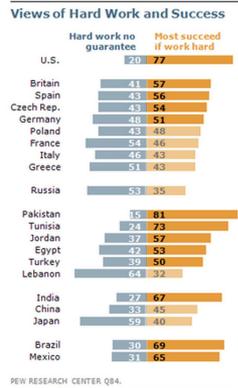
- Pew Research Reports
- ACL and CVPR papers

Pipeline Evaluation

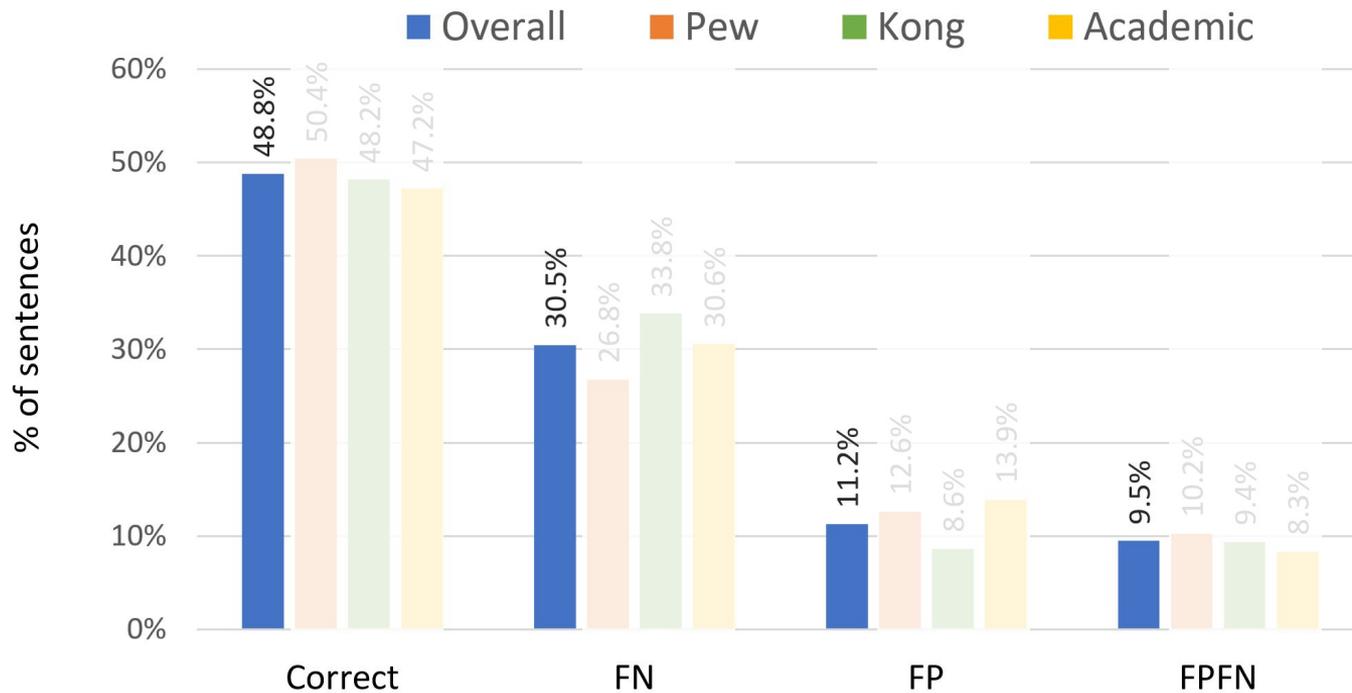
● Corpus

- Pew Research Reports
- ACL and CVPR papers
- Kong et. al (2014)

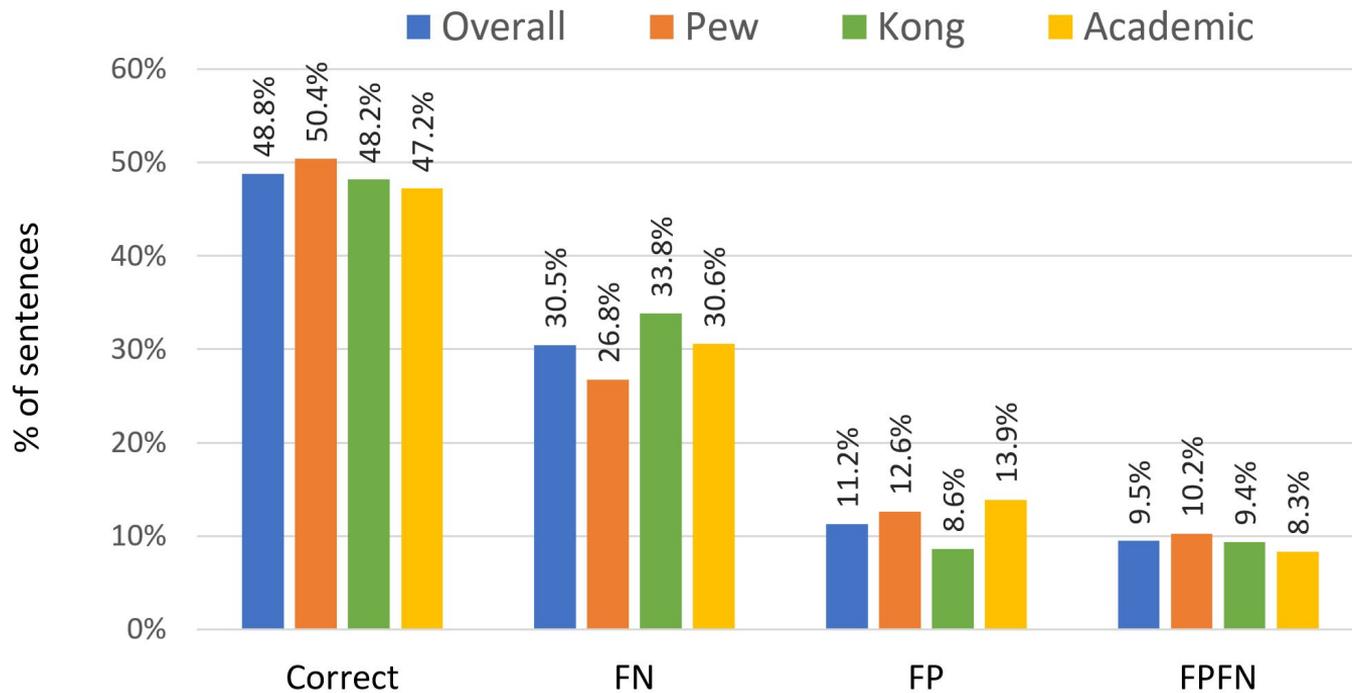
Half or more in 13 of the 21 nations surveyed believe that most people can succeed if they are willing to work hard. This includes Pakistan (81%) and the U.S. (77%). It also includes Tunisia (73%), Brazil (69%), India (67%), Mexico (65%), Jordan (57%), France (46%), Spain (46%), Czech Rep. (44%), Italy (43%), Germany (41%), Poland (40%), and Greece (35%).



Results



Results



User Study

User Study

- Hypothesis

Our interface facilitates reading documents with tables

User Study

- Within-subject study
- 14 adult volunteers, all fluent in English
- Task: Annotate references with/without our interface

Findings of User Study

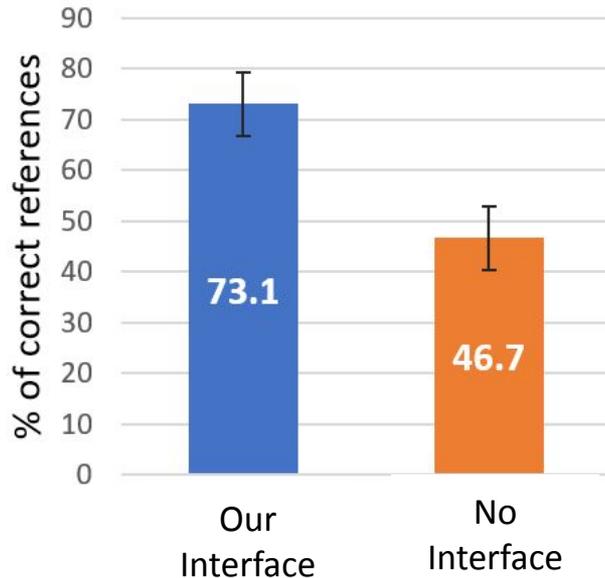
- Hypothesis

Our interface facilitates reading documents with tables

Findings of User Study

- Hypothesis

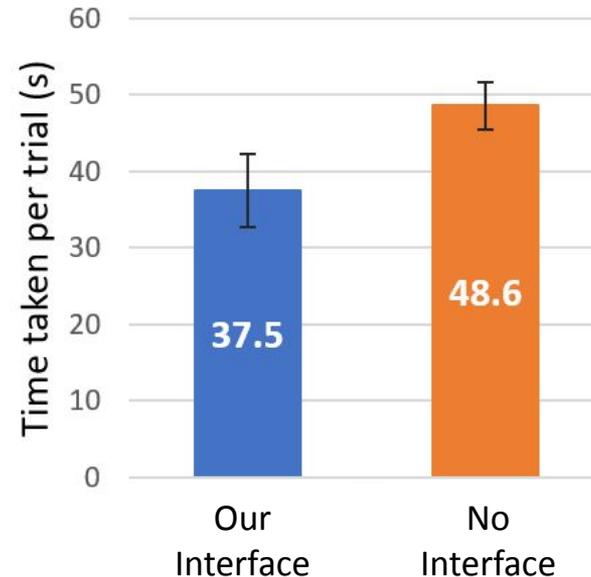
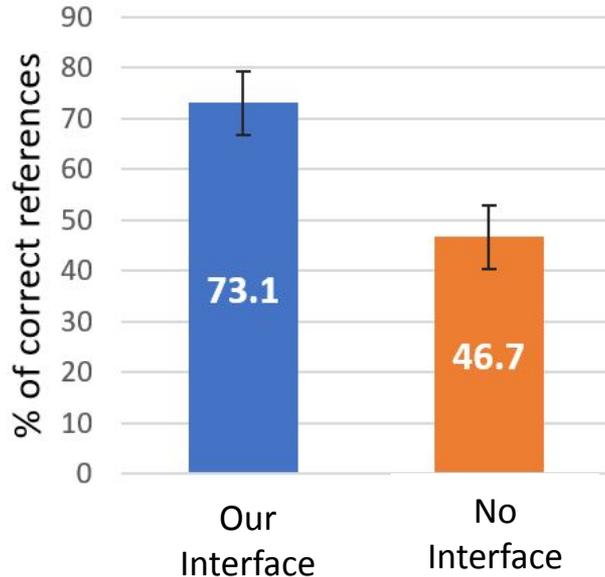
Our interface facilitates reading documents with tables



Findings of User Study

- Hypothesis

Our interface facilitates reading documents with tables



Findings of User Study

“The interface allows me to read the table while reading the text ...”

PRINCIPLE 1

A mutation to a place is a mutation to all conflicting places.

As described in Section 3.3, a mutation to a place is represented by updating a variable x in a stack σ by plugging a value v into a value context \mathcal{V} . To denote a conflict, we reuse the notation from Oxide that $\pi_1 \# \pi_2$ means " π_1 and π_2 do not conflict", or more formally:

$$x_1.q_1 \# x_2.q_2 \stackrel{\text{def}}{=} x_1 \neq x_2 \vee ((q_1 \text{ is not a prefix of } q_2) \wedge (q_2 \text{ is not a prefix of } q_1))$$

Conversely, we use the shorthand $\pi_1 \sqcap \pi_2 \stackrel{\text{def}}{=} \neg(\pi_1 \# \pi_2)$. So if a place π_{any} is changed when π_{mut} is mutated, then it must be that $\pi_{\text{any}} \sqcap \pi_{\text{mut}}$.

THEOREM 3.1

Let:

- $\pi_{\text{mut}} = \pi_{\text{mut}}^{\square}[x], \sigma$ where $\sigma \vdash \pi_{\text{mut}} \Downarrow _ \times \mathcal{V}$
- $v, \vec{\sigma} = \sigma[x \mapsto \mathcal{V}[v]]$
- π_{any} be any place

Then $\sigma(\pi_{\text{any}}) \neq \vec{\sigma}(\pi_{\text{any}}) \implies \pi_{\text{any}} \sqcap \pi_{\text{mut}}$.

Part of Nota's inspiration was my attempts to visually encode correspondences between objects (see page 10 of the PDF). LaTeX's brittle abstractions made it frustratingly hard to do something as simple as "draw a colored underline beneath a piece of math."

By contrast, implementing this feature was trivial in HTML/CSS/Javascript. And we could extend the idea with interactions like drawing attention to corresponding objects on hover.

How Readers Integrate Information in Visualizations & Text

[Kim et al. 2020]





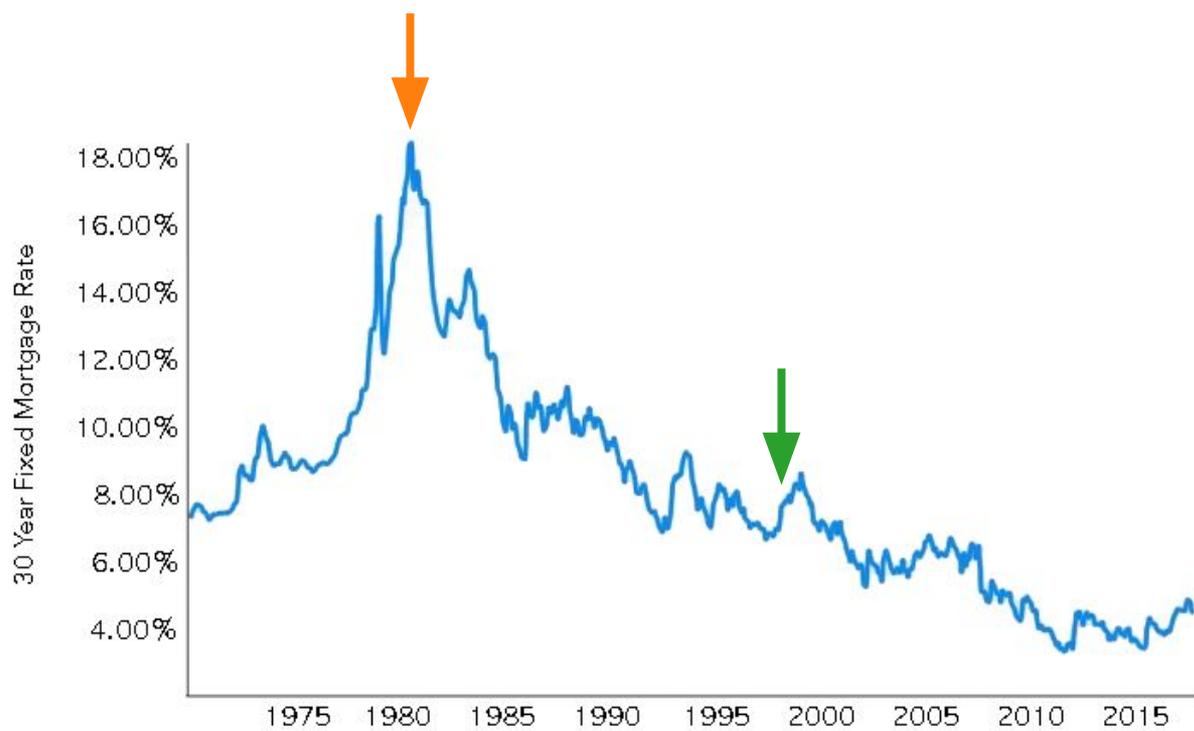
The 30-year fixed mortgage rate increased slightly from 1997 to 1999.



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



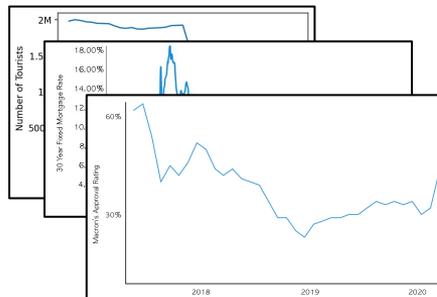
The 30-year fixed mortgage rate increased slightly from 1997 to 1999.



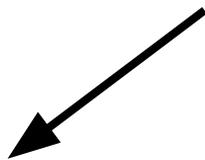
The 30-year fixed mortgage rate **increased slightly from 1997 to 1999.**

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

User Study



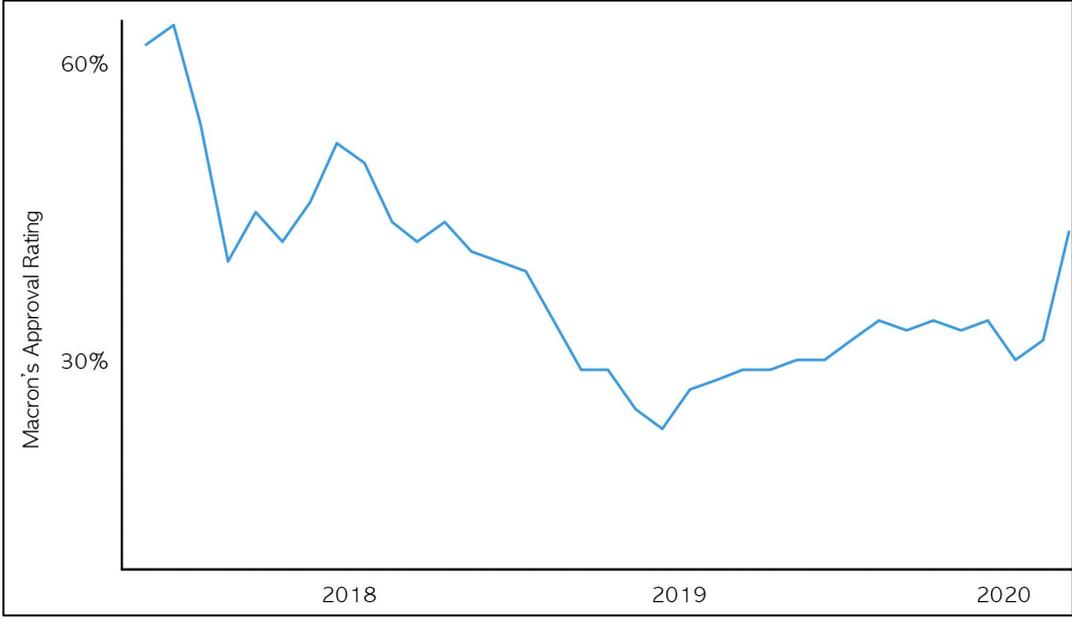
Charts

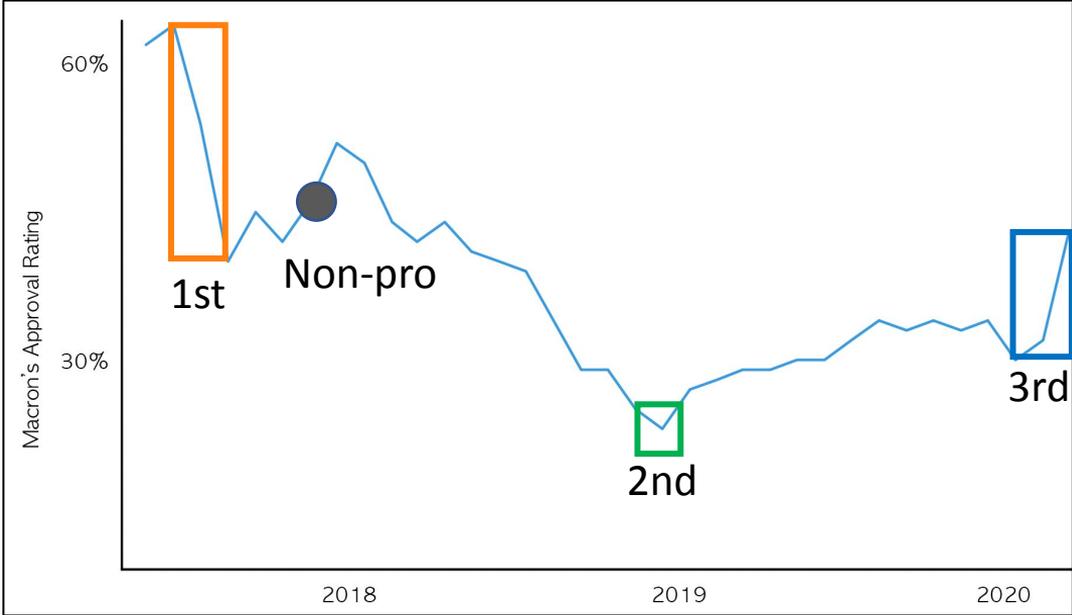


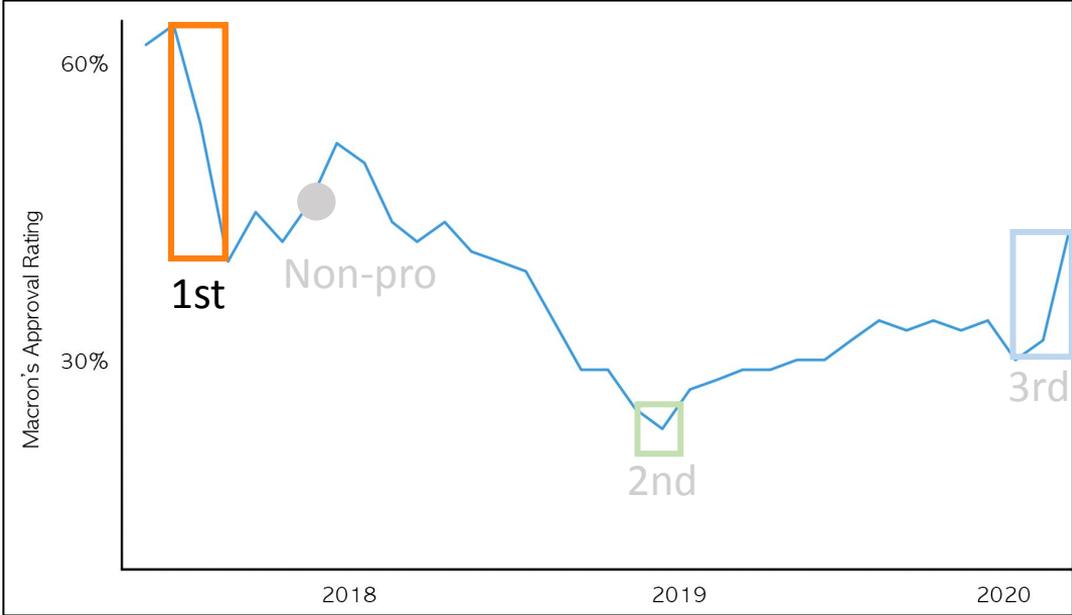
Identify Visually
Prominent Features

Generate Captions

Collect Takeaways for
Charts & Captions







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



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

Results



Chart Emphasis
≠
Caption Emphasis
↓
Takeaway:
Chart Emphasis

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

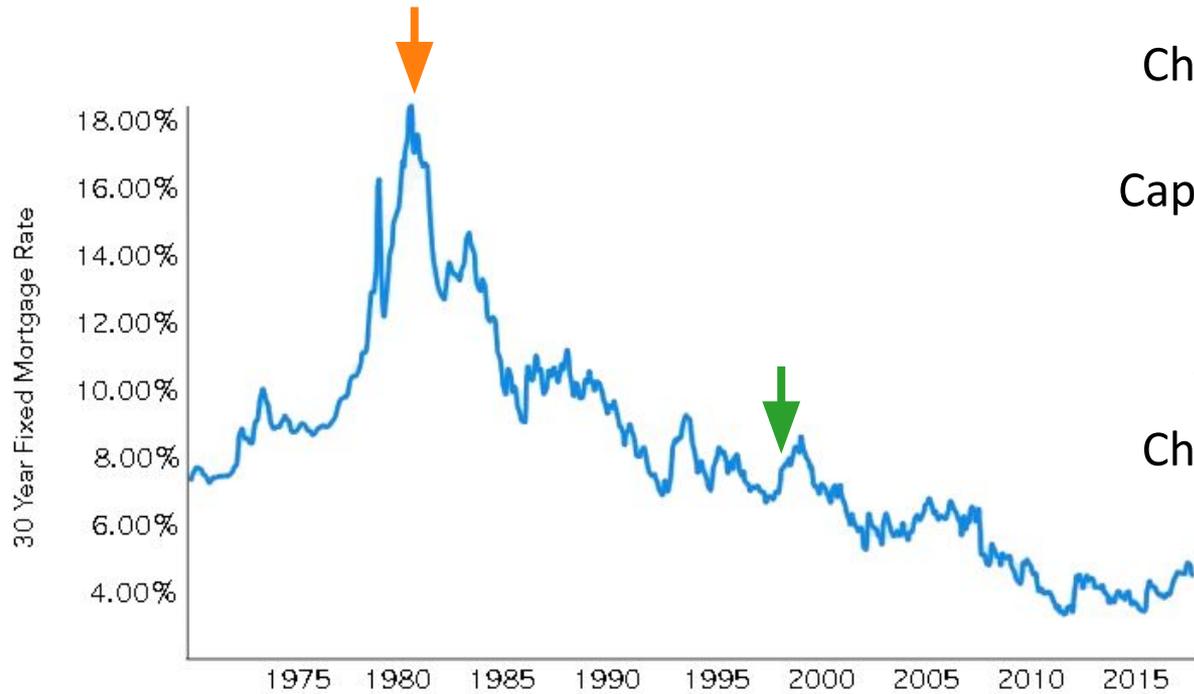


Chart Emphasis
≠
Caption Emphasis
↓
Takeaway:
Chart Emphasis

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



Chart Emphasis



Caption Emphasis



Takeaway:

Chart Emphasis

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

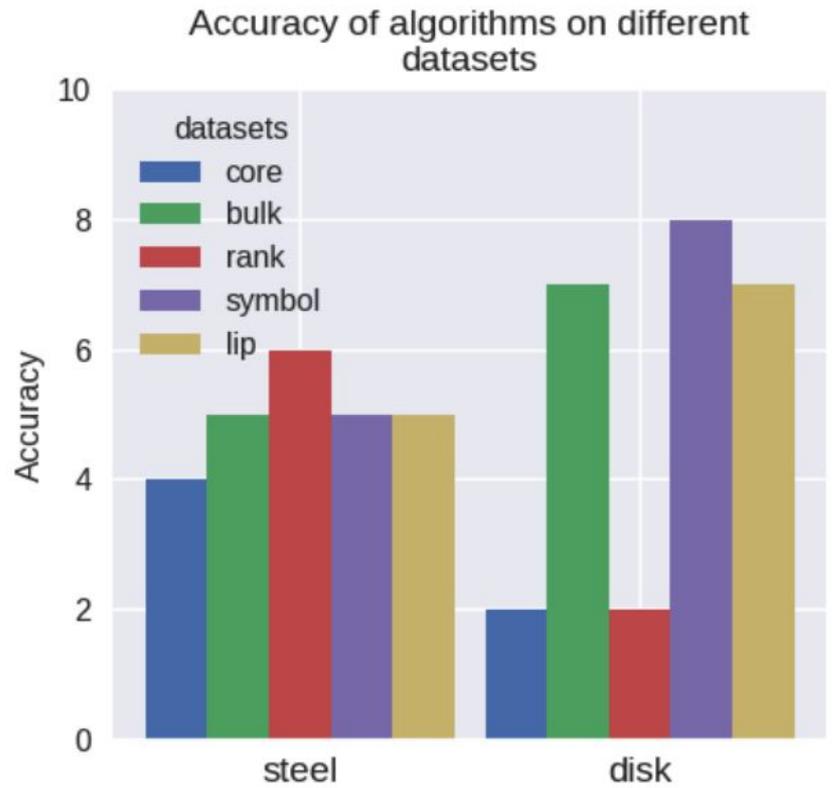
*When text and visualization emphasis **mismatch**, readers rely **more on the chart** and can **miss information in the caption**.*

Chart Question Answering with Explanations

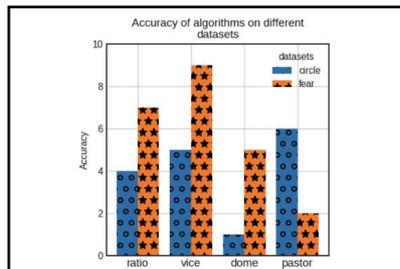
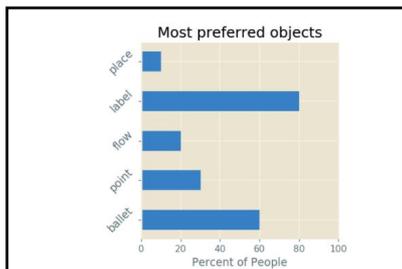
[Kim et al. 2021]



[Kafle et al. 2018]



[Kafle et al. 2018]



Structure Understanding
How many bars are there?
Are the bars horizontal?

Structure Understanding
How many groups of bars are there?
Are the bars stacked?

Data Retrieval
What percent of people prefer the object ballet?
What is the label of the third bar from the bottom?

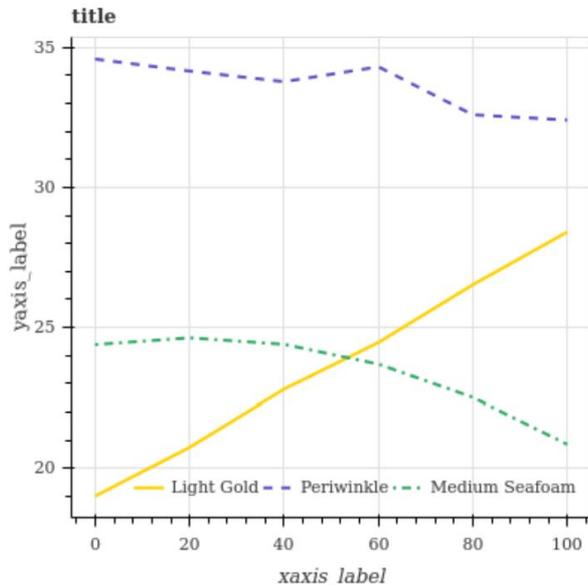
Data Retrieval
What is the accuracy of the algorithm vice on the dataset fear?
Does the chart use logarithmic scale?

Reasoning
Is the object flow preferred by more people than the object point?
What percent of people prefer the object point or ballet?

Reasoning
Which algorithm has the lowest accuracy across all datasets?
How many algorithms have accuracy greater than 5 in at least one dataset?

DVQA

[Kafle et al. 2018]



FigureQA

[Kahou et al. 2018]

Q: Does Medium Seafoam intersect Light Gold?

A: Yes

Q: Is Medium Seafoam the roughest?

A: No

Q: Is Light Gold less than Periwinkle?

A: Yes

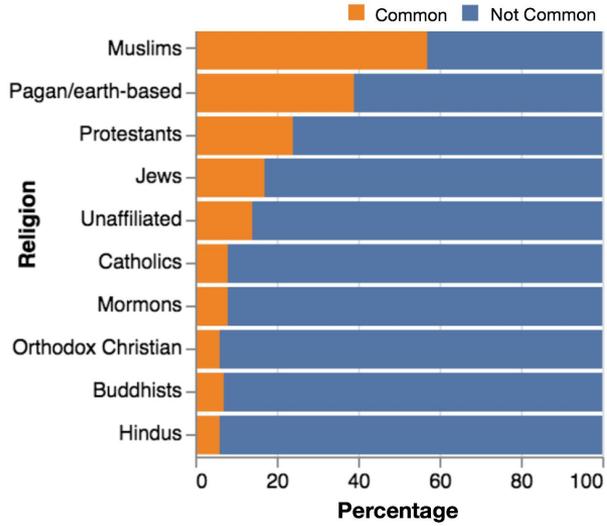
Q: Does Periwinkle have the maximum area under the curve?

A: Yes

Q: Does Medium Seafoam have the lowest value?

A: No

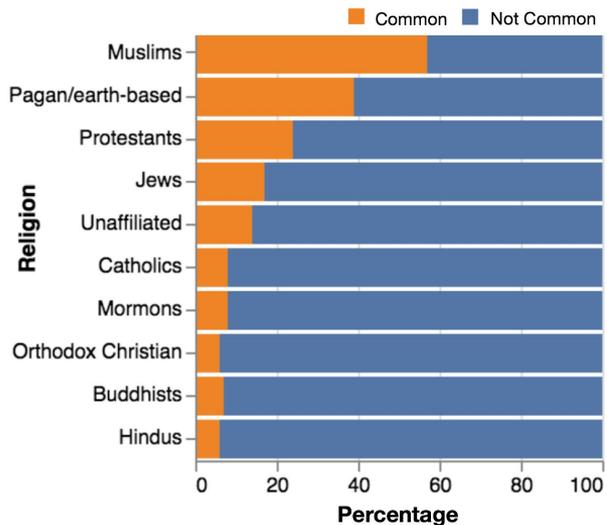
Chart



Question

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

Chart



Question

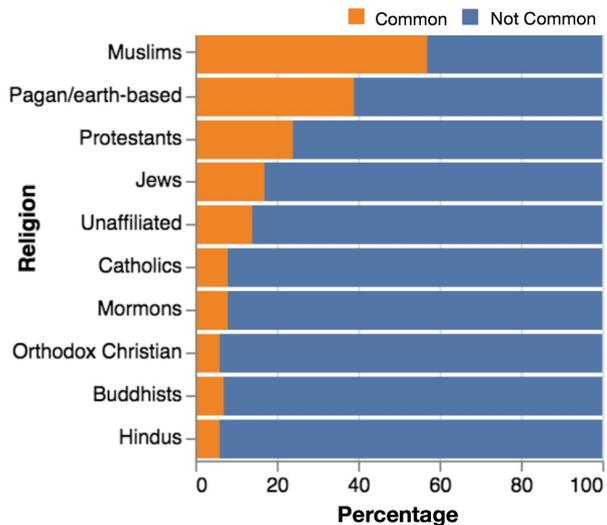
For which religion did the fewest chaplains think that religious extremism is common?

Chart QA

Answer

Orthodox Christians, Hindus.

Chart



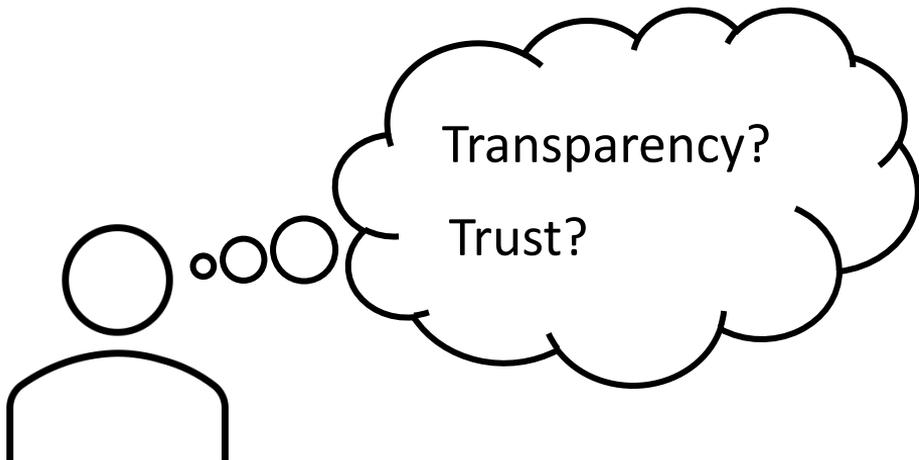
Question

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

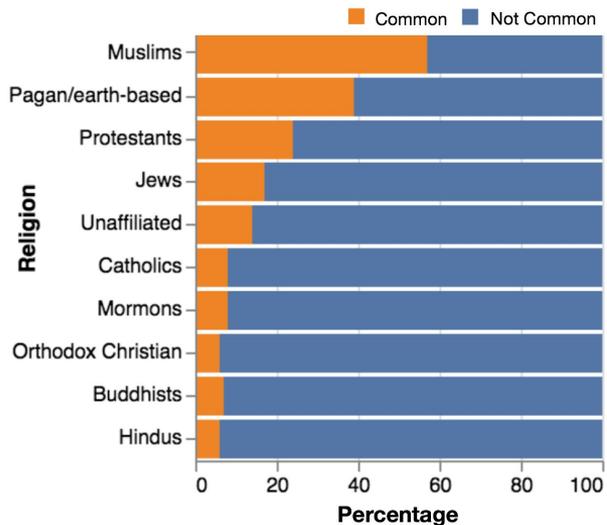


Answer

Orthodox Christians, Hindus.



Chart



Question

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

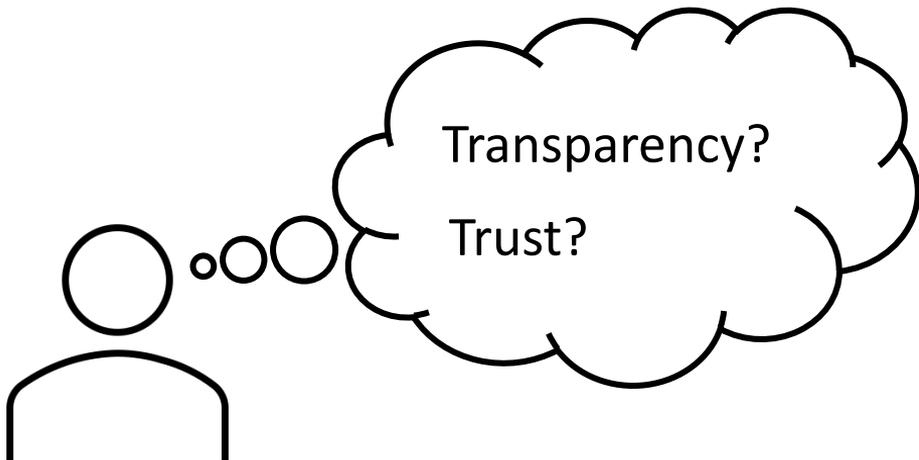


Answer

Orthodox Christians, Hindus.

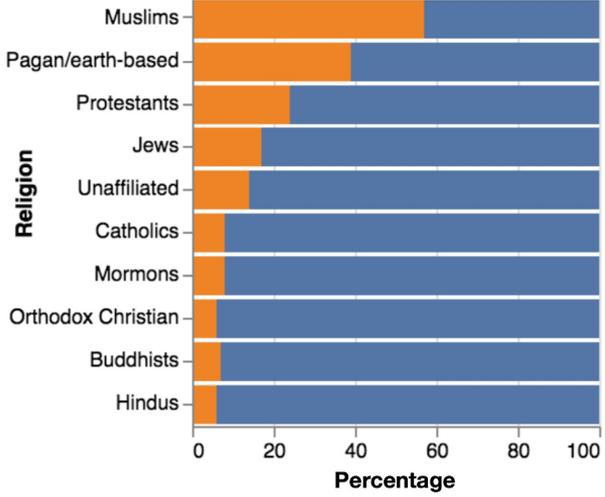
Explanation

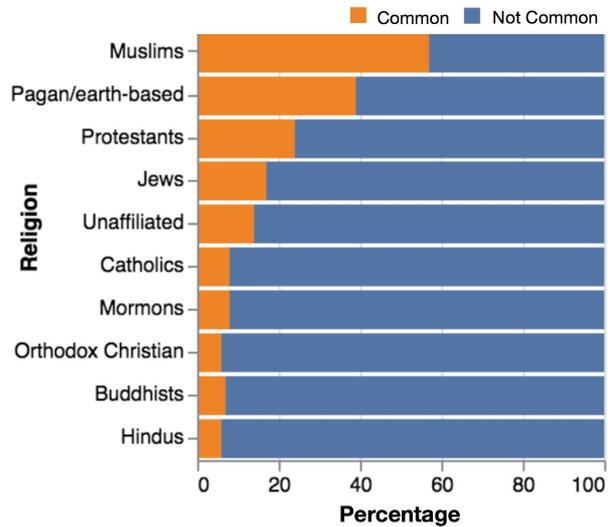
I looked up 'Religion' for the shortest orange bars.



Formative Study

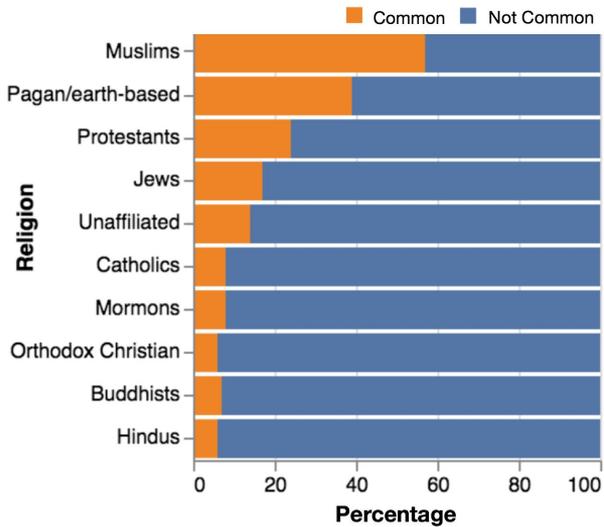
Common Not Common





Question

Which religion has the greatest value for Common?



Question

Which religion has the greatest value for Common?



Answer

Muslims

Explanation

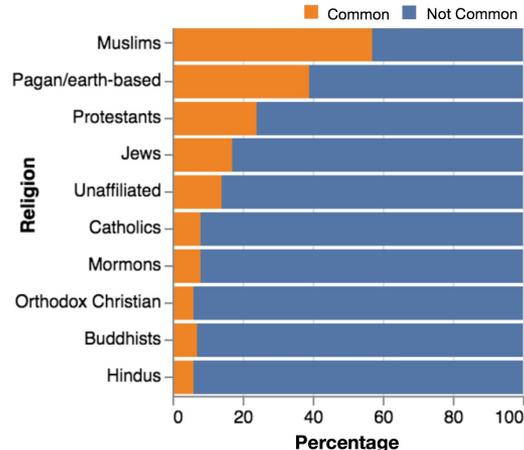
I picked religions with the greatest orange percentage.

Formative Study Results

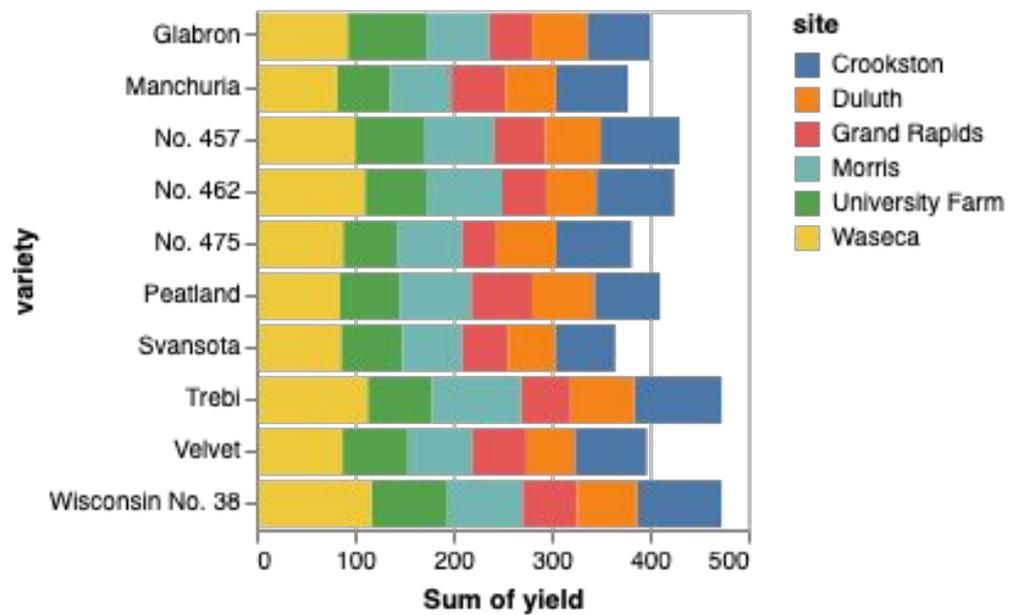
- Explanations describe procedure for computing answer

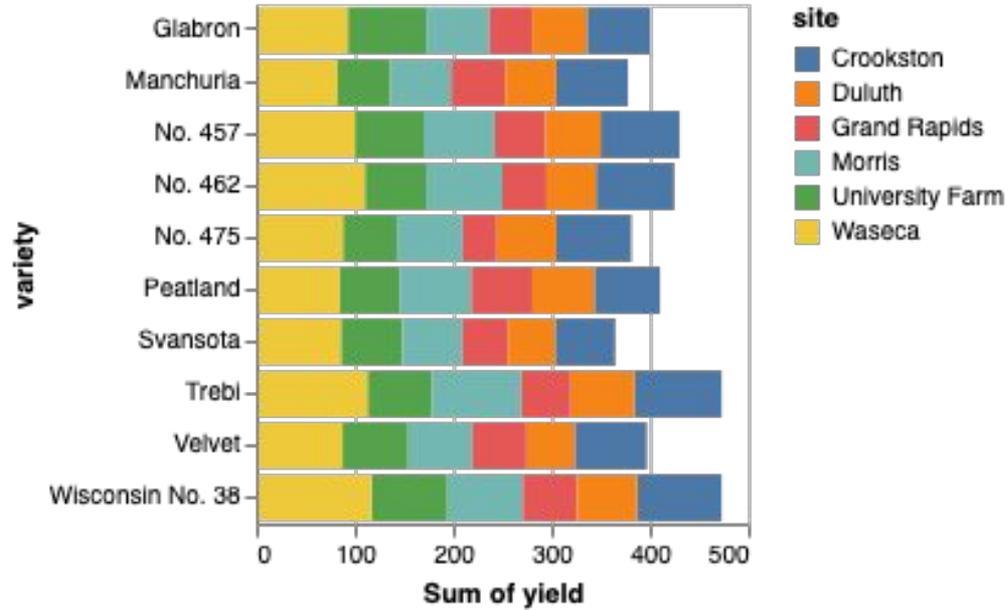
- Explanations describe procedure for computing answer
- Half of the explanations referred to visual features of chart

- Explanations describe procedure for computing answer
- Half of the explanations referred to visual features of chart

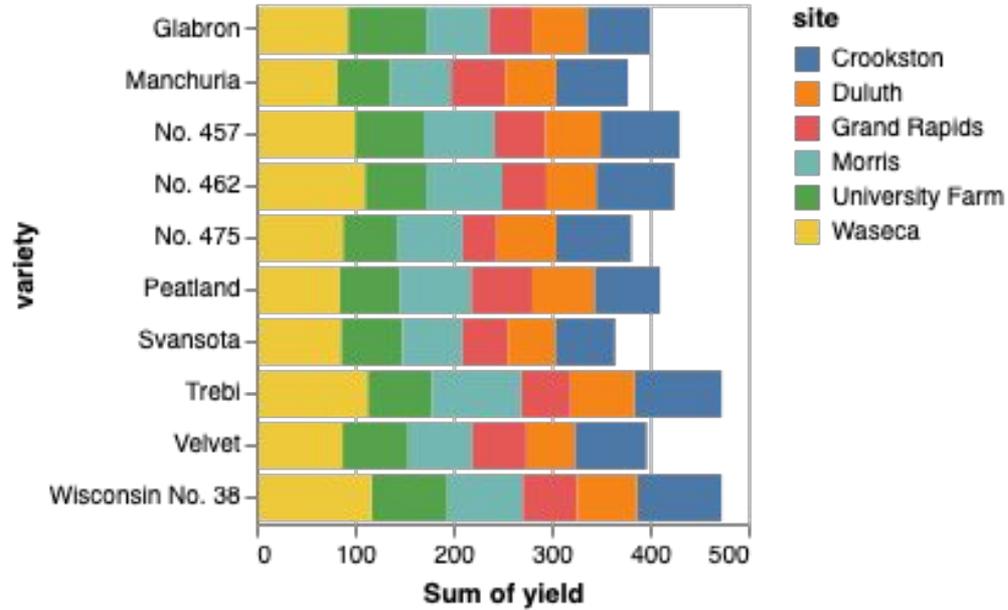


*“I picked religions with the greatest **orange** percentage.”*





Glabron at University Farm



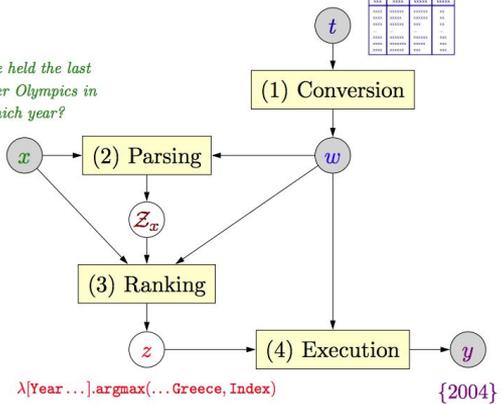
Glabron at University Farm

Green component in the top bar

Chart QA Pipeline and Generating Explanations

Greece held the last Summer Olympics in which year?

Year	Index
2004	1
2008	2
2012	3
2016	4
2020	5
2024	6



$\lambda[\text{Year} \dots]. \text{argmax}(\dots \text{Greece, Index})$

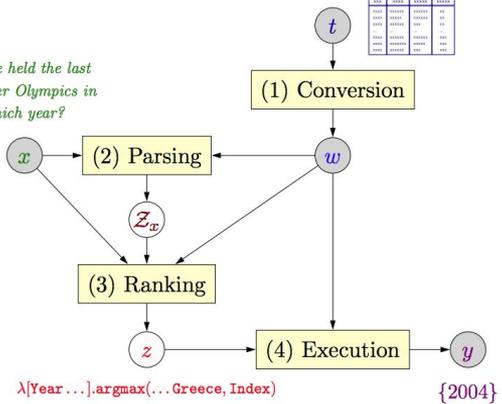
{2004}

Compositional Semantic Parsing on Semi-Structured Tables

Pasupat and Liang (2015)

Greece held the last Summer Olympics in which year?

Year	Index
2004	1
2008	2
2012	3
2016	4
2020	5
2024	6



$\lambda[\text{Year} \dots]. \text{argmax}(\dots \text{Greece, Index})$

{2004}

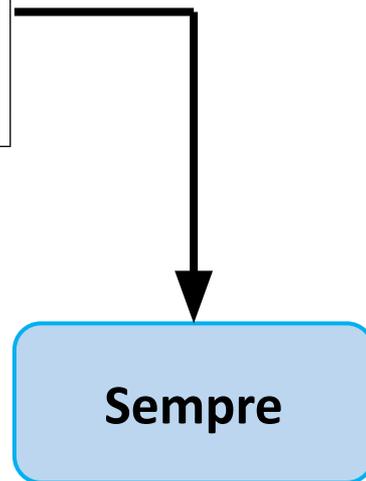
Sempre

Compositional Semantic Parsing on Semi-Structured Tables

Pasupat and Liang (2015)

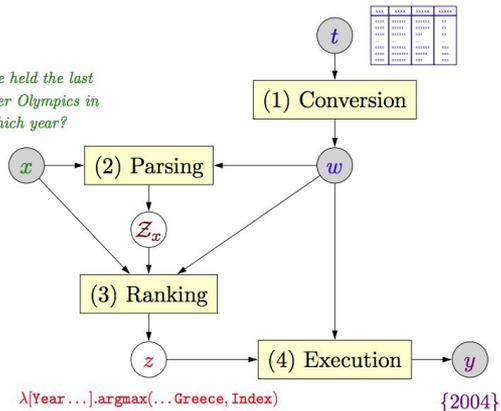
Table

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204



Sempre

Greece held the last Summer Olympics in which year?



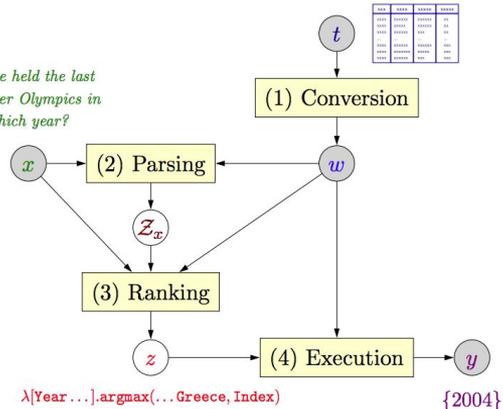
Compositional Semantic Parsing on Semi-Structured Tables

Pasupat and Liang (2015)

Table

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

Greece held the last Summer Olympics in which year?



Question

Greece held its last Summer Olympics in which year?

Sempre

Compositional Semantic Parsing on Semi-Structured Tables

Pasupat and Liang (2015)

Table

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

Question

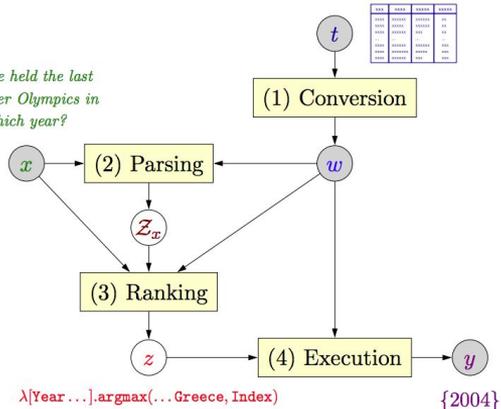
Greece held its last Summer Olympics in which year?

Sempre

Answer

2004

Greece held the last Summer Olympics in which year?



Compositional Semantic Parsing on Semi-Structured Tables

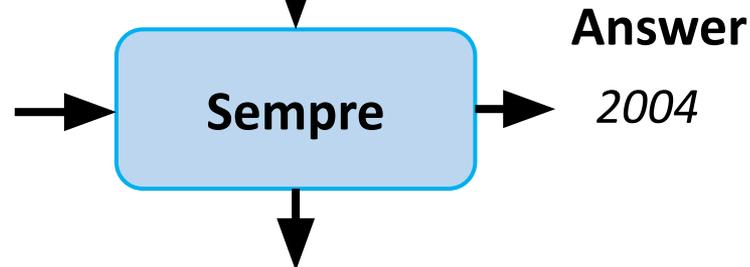
Pasupat and Liang (2015)

Table

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

Question

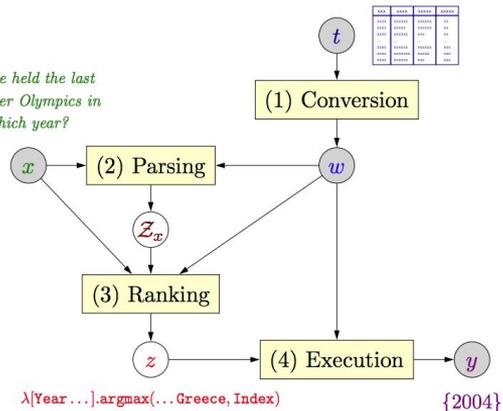
*Greece held its last
Summer Olympics in
which year?*



Lambda Expression

$R[\lambda x[\text{Year.Date}.x]]$
 $.\text{argmax}(\text{Country.Greece},$
 $\text{Index})$

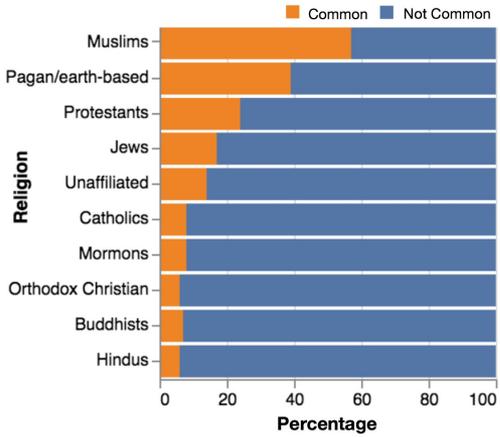
*Greece held the last
Summer Olympics in
which year?*



Compositional Semantic Parsing on Semi-Structured Tables

Pasupat and Liang (2015)

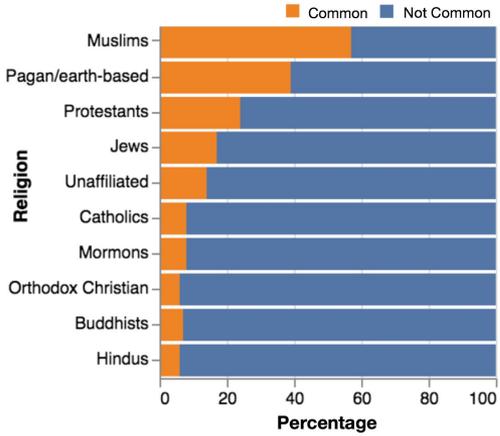
Chart



Question about Chart

Which religion has the shortest orange component?

Chart

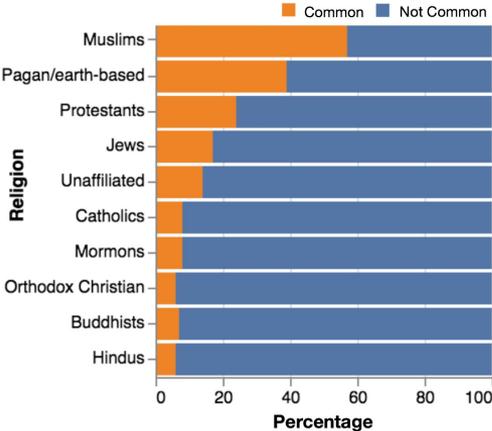


Question about Chart

Which religion has the shortest orange component?

Sempre

Chart



Data

Table

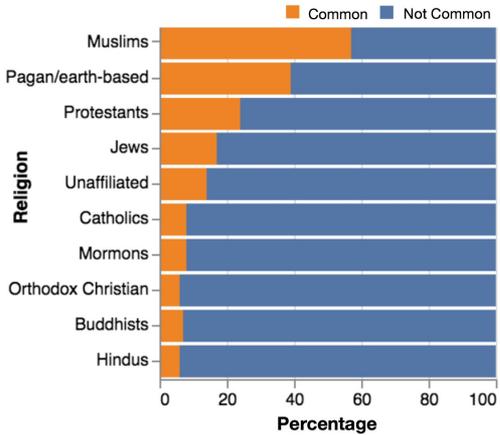
Religion	Common	Not common
Muslims	57	43
Pagan/earth-based	39	61
Protestants	24	76
Jews	17	83
⋮	⋮	⋮
Buddhists	7	93
Hindus	6	94

Question about Chart

Which religion has the shortest orange component?

Sempre

Chart



Data

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?

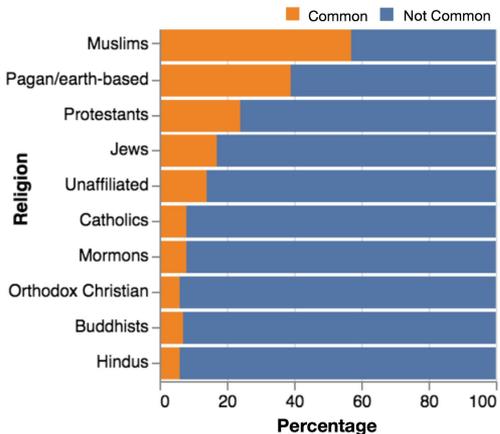


Question about Table

Which religion has the least 'Common' Percentage?

Sempre

Chart



Data

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

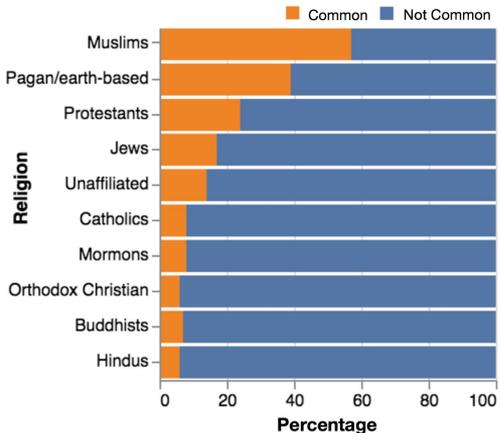
Which religion has the shortest orange component?

Question about Table

Which religion has the least 'Common' Percentage?

Sempre

Chart



Data

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

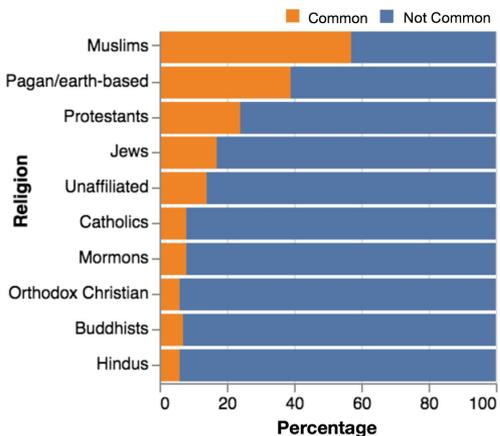
Which religion has the shortest orange component?

Question about Table

Which religion has the least 'Common' Percentage?

Sempre

Chart



Data

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

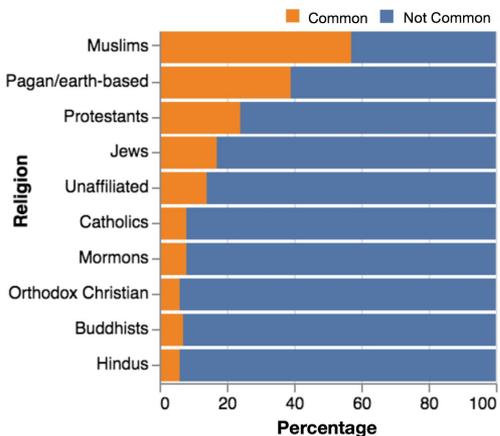
Which religion has the least 'Common' Percentage?

Sempre

Lambda Expression

```
argmin[R[Religion].Row,  
R[λx(R[Number].R[Common].  
Religion.x)]]
```

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

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

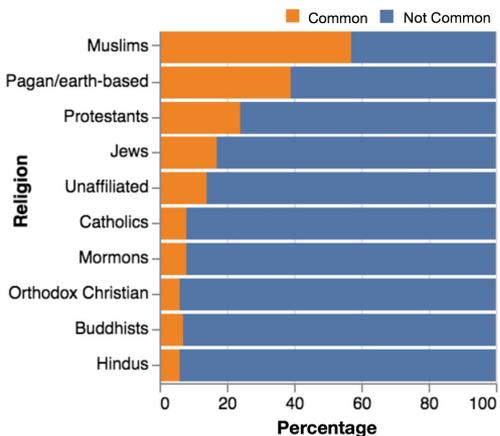
Lambda Expression

$\text{argmin}[R[\text{Religion}].\text{Row}, R[\lambda x(R[\text{Number}].R[\text{Common}].\text{Religion}.x)]]$

Non-Visual Explanation

I looked up 'Religion' with the least 'Percentage' of 'Common'.

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

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

Lambda Expression

$\text{argmin}[R[\text{Religion}].\text{Row}, R[\lambda x(R[\text{Number}].R[\text{Common}].\text{Religion}.x)]]$

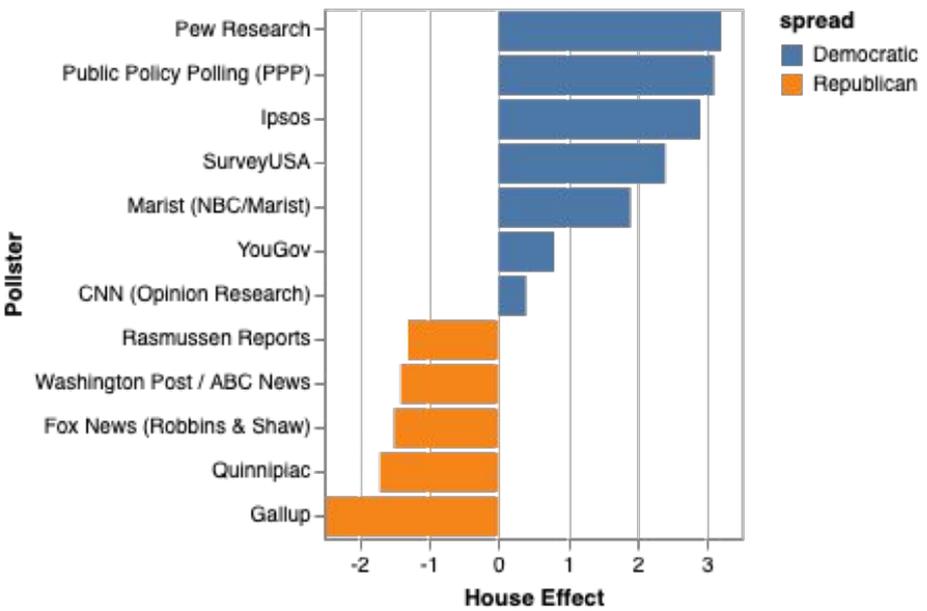
Non-Visual Explanation

I looked up 'Religion' with the least 'Percentage' of 'Common'.

Visual Explanation

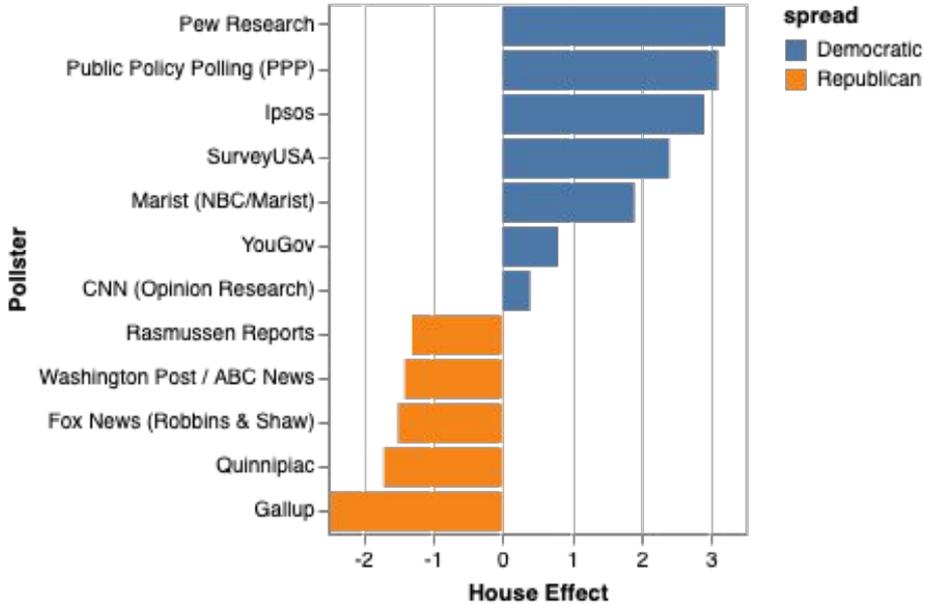
I looked up 'Religion' for the shortest orange bar.

Example Explanations



Question

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

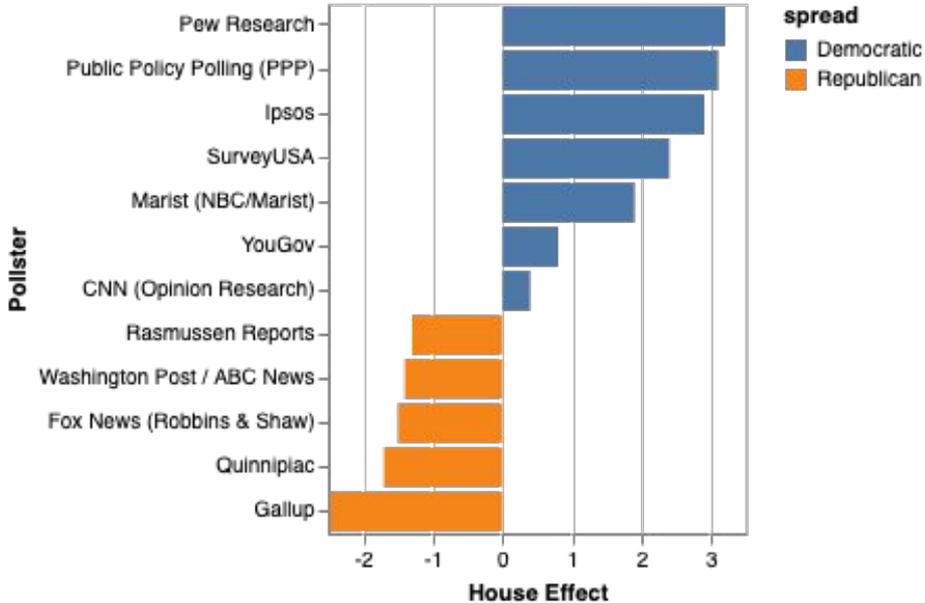


Question

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

Answer

0.8

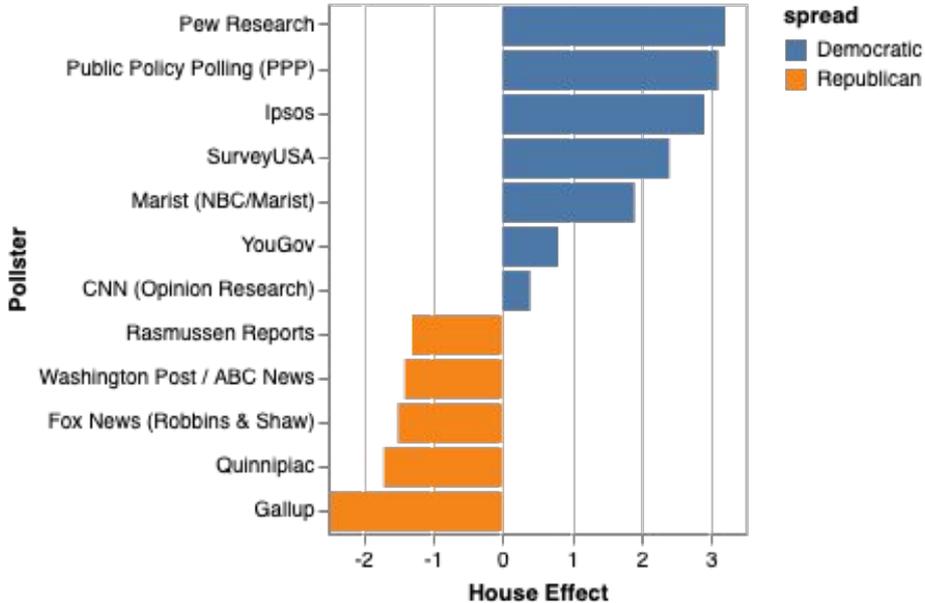


Question

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

Answer

0.8



Question

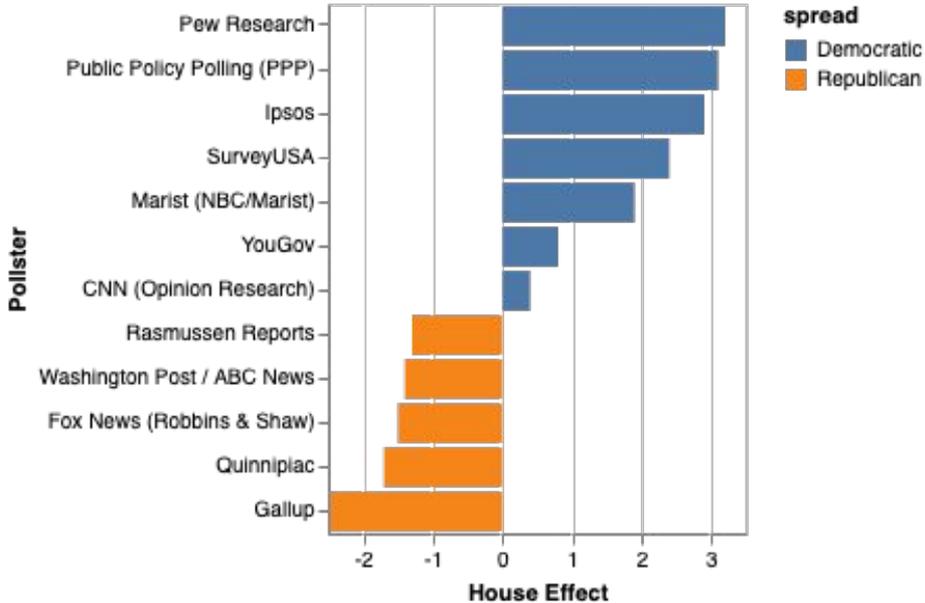
What is the difference between the value of Gallup and Quinnipiac?

Answer

0.8 ✓

Explanation

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



Question

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

Answer

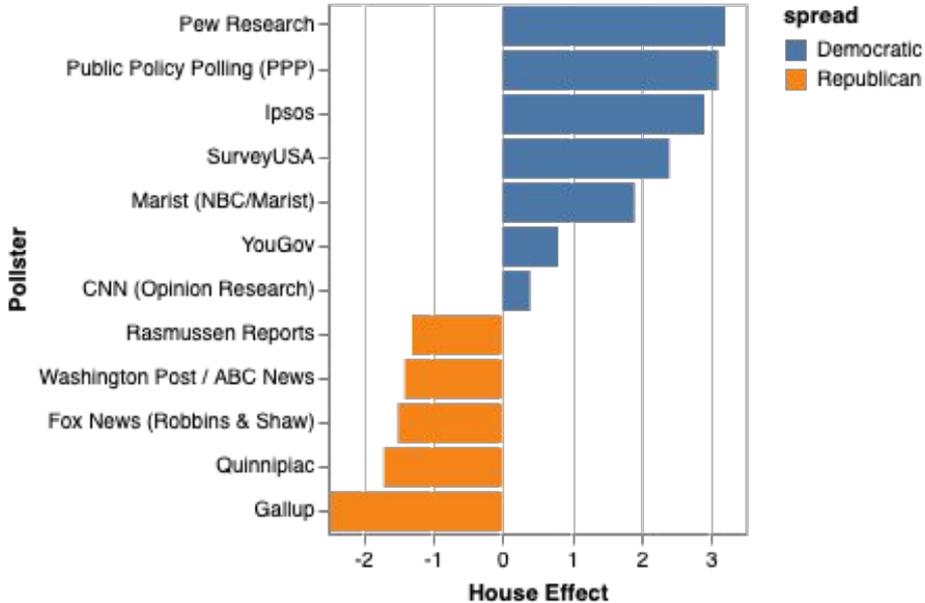
0.8 ✓

Explanation

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

Question

What position is Ipsos in?



Question

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

Answer

0.8 ✓

Explanation

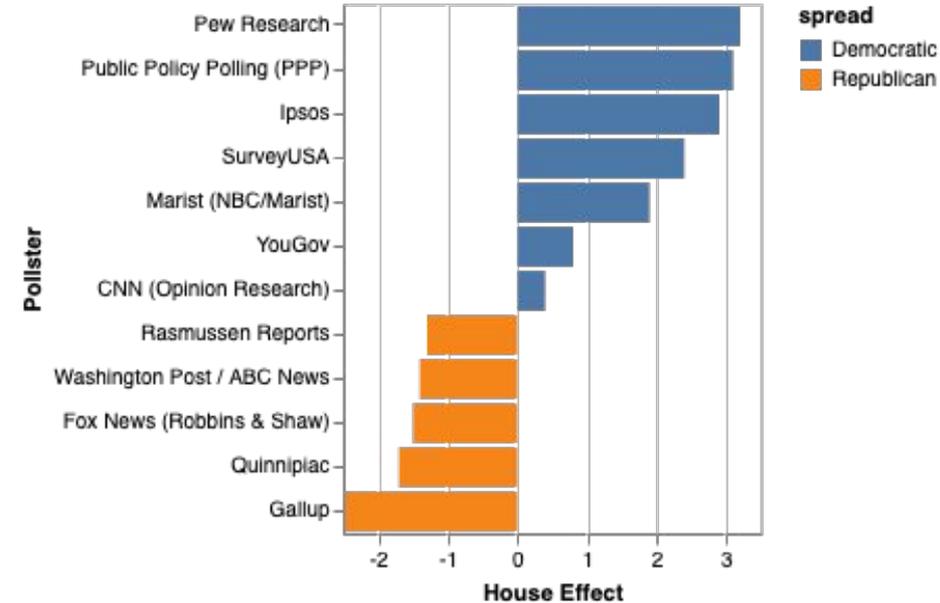
I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

Question

What position is Ipsos in?

Answer

2.9



Question

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

Answer

0.8 ✓

Explanation

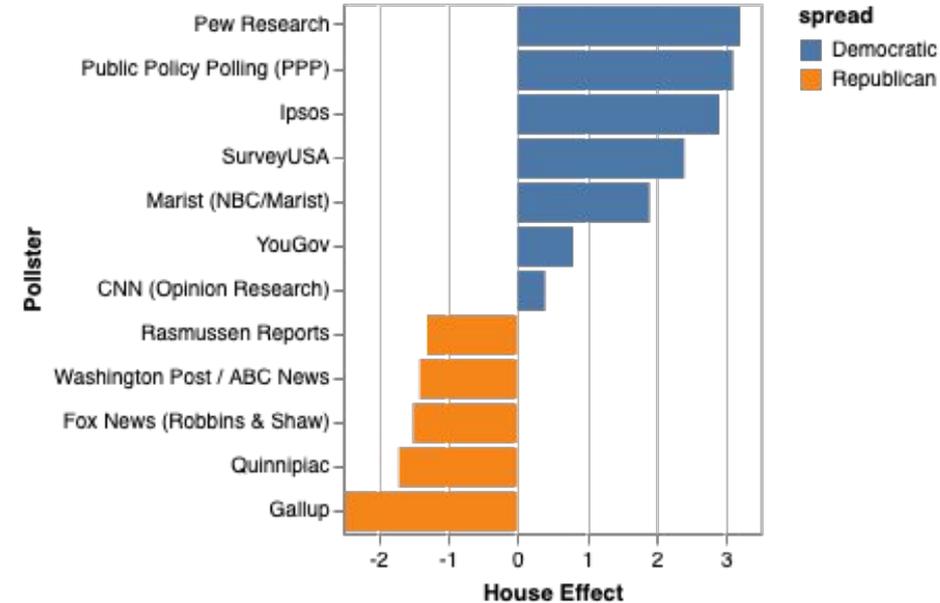
I computed the difference between the length of the bar for 'Gallup' and 'Quinnipiac'.

Question

What position is Ipsos in?

Answer

2.9 ✗



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

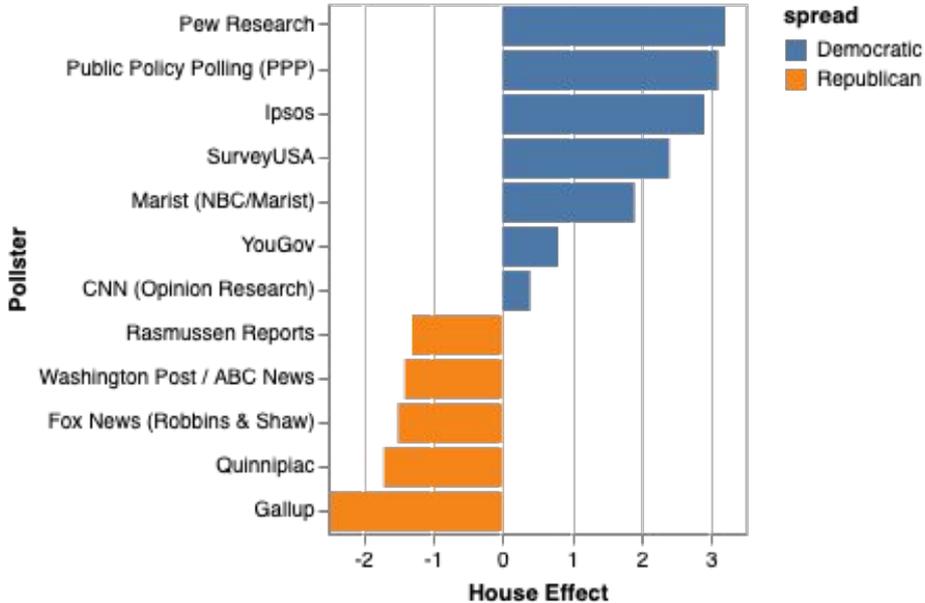
What position is Ipsos in?

Answer

2.9 ✗

Explanation

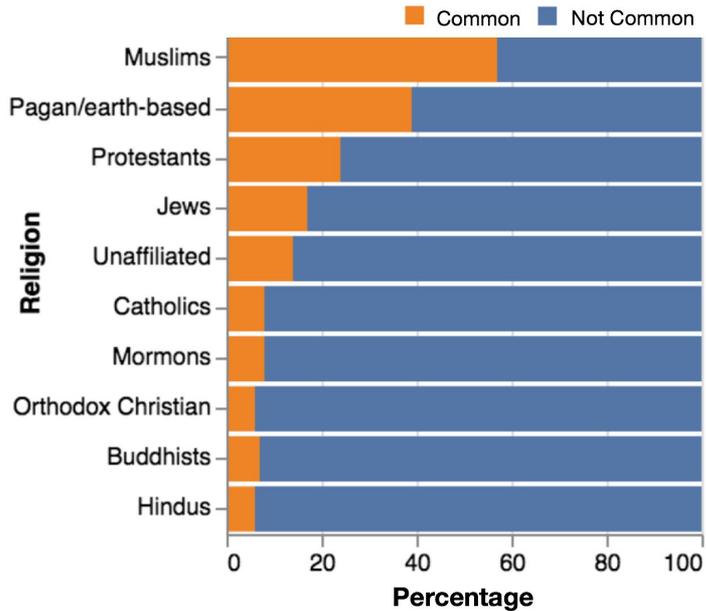
I looked up the length of the bar for 'Ipsos'.



User Study

Hypothesis: Visual explanations increase transparency and trust

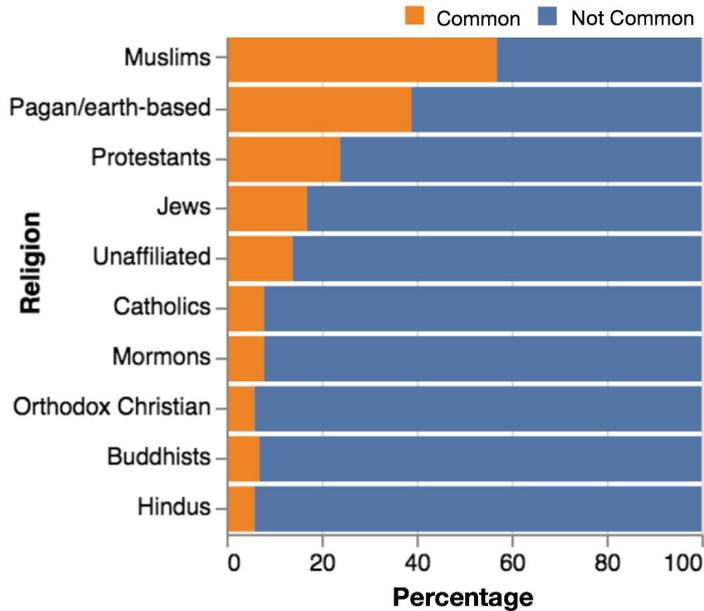
Hypothesis: Visual explanations increase transparency and trust



For which religion did the

*Q: fewest chaplains think that
religious extremism is
common?*

Hypothesis: Visual explanations increase transparency and trust



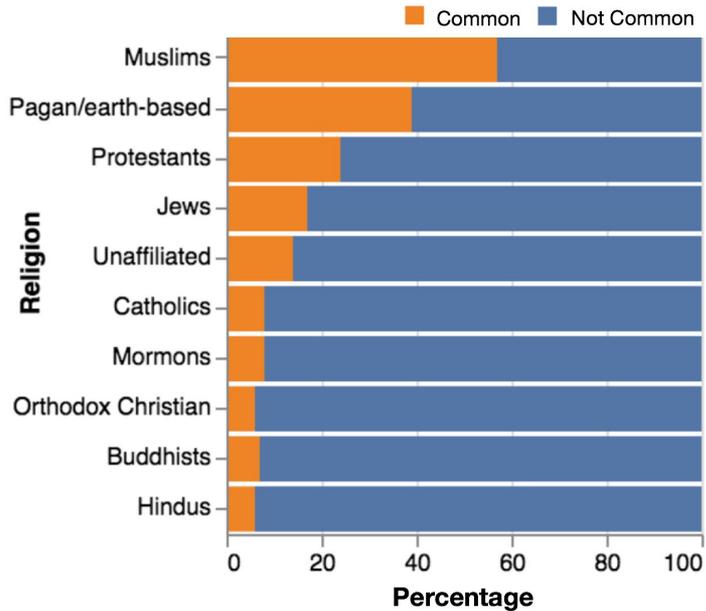
A1 (vis):

Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.

For which religion did the

Q: fewest chaplains think that religious extremism is common?

Hypothesis: Visual explanations increase transparency and trust



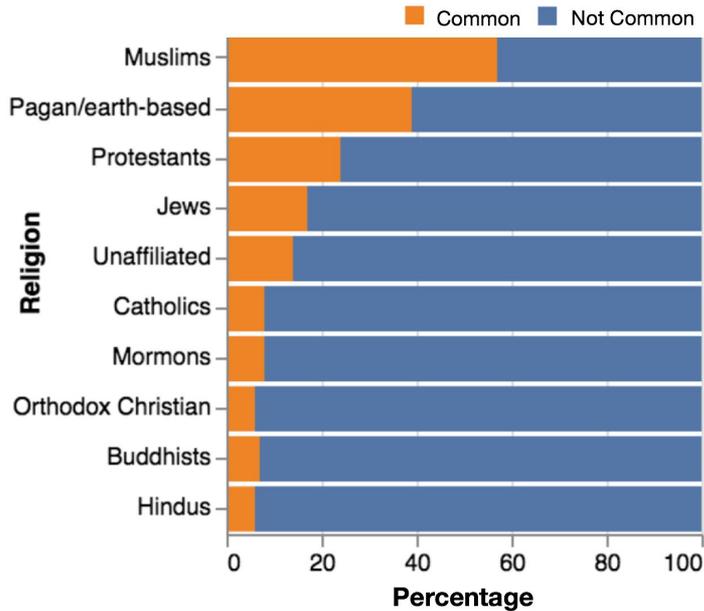
A1 (vis): *Orthodox Christians, Hindus. I looked up 'Religion' for the shortest orange bar.*

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

For which religion did the

Q: *fewest chaplains think that religious extremism is common?*

Hypothesis: Visual explanations increase transparency and trust



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

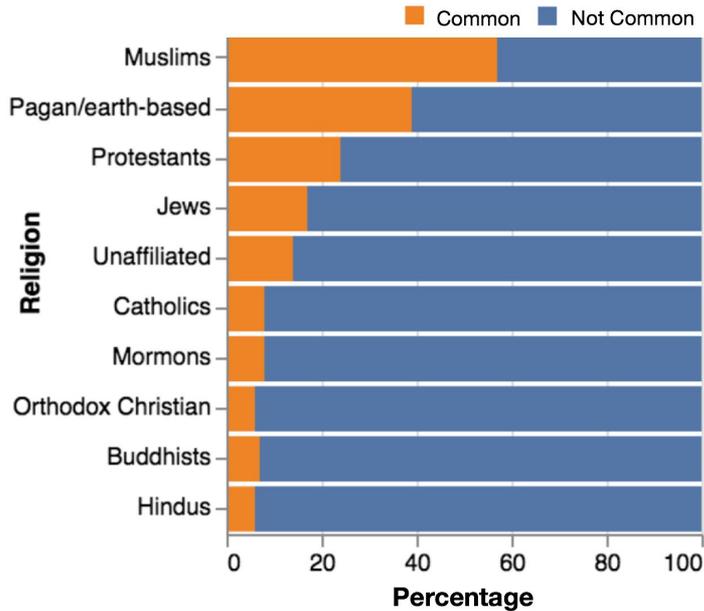
A2 (no-exp): *Orthodox Christians, Hindus.*

A3 (non-vis): *Orthodox Christians, Hindus. I looked up 'Common'.*

For which religion did the

Q: *fewest chaplains think that religious extremism is common?*

Hypothesis: Visual explanations increase transparency and trust



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'.*

For which religion did the

Q: fewest chaplains think that religious extremism is common?

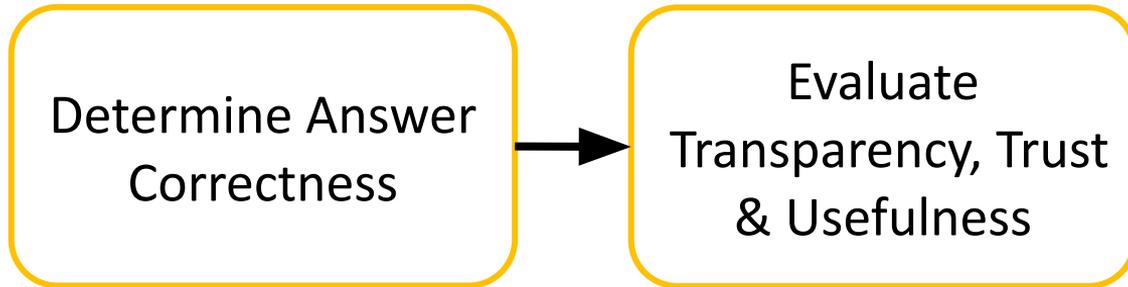
- 16 participants

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

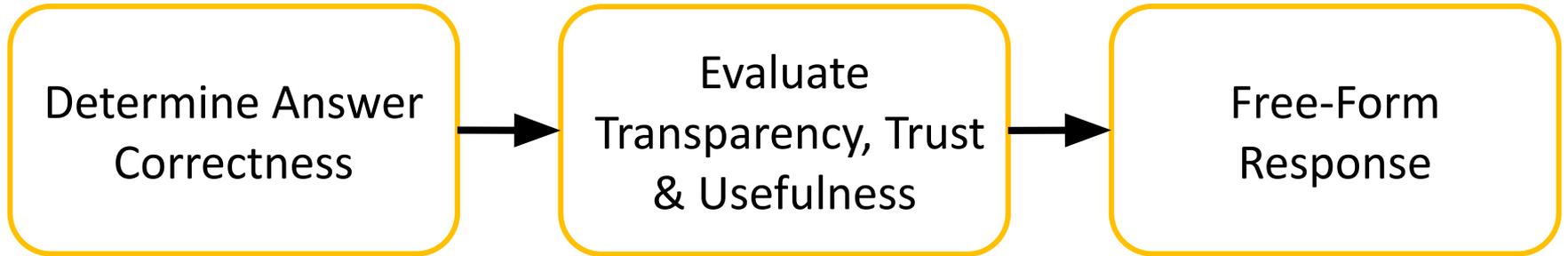
- 16 participants
- 20 Chart-Question-Answer-Explanation tuples (5 per explanation type)

Determine Answer
Correctness

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



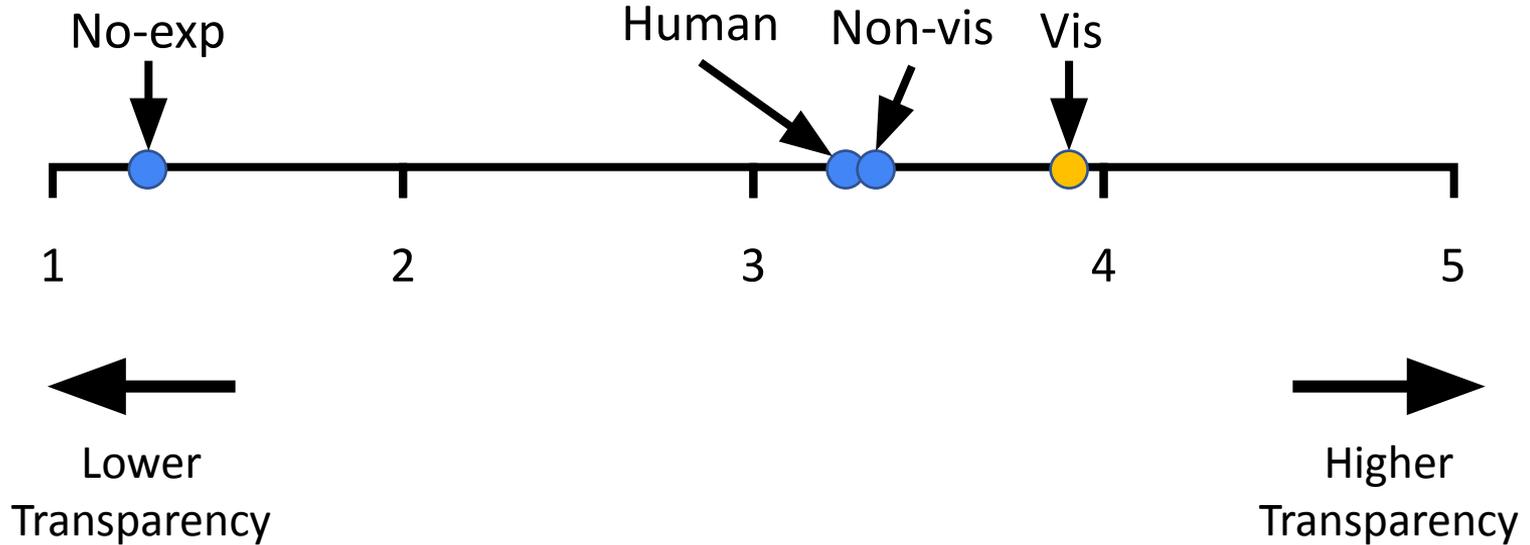
- 16 participants
- 20 Chart-Question-Answer-Explanation tuples (5 per explanation type)



Study Results

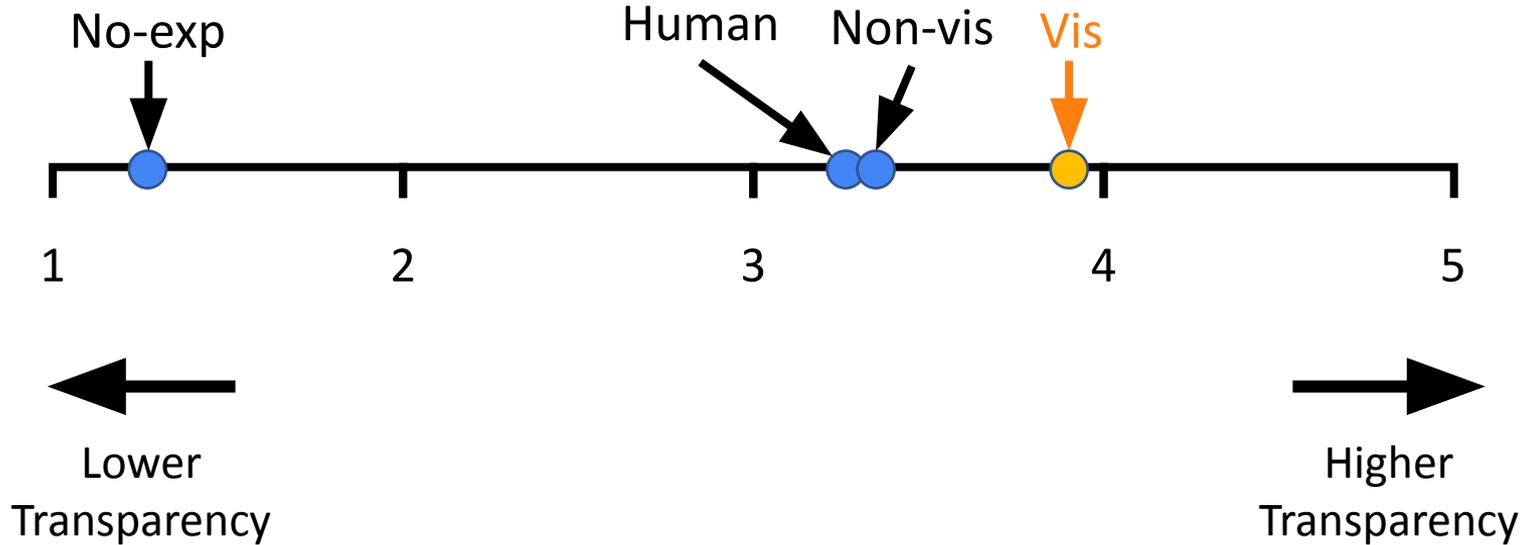
Transparency

- Presence of explanations



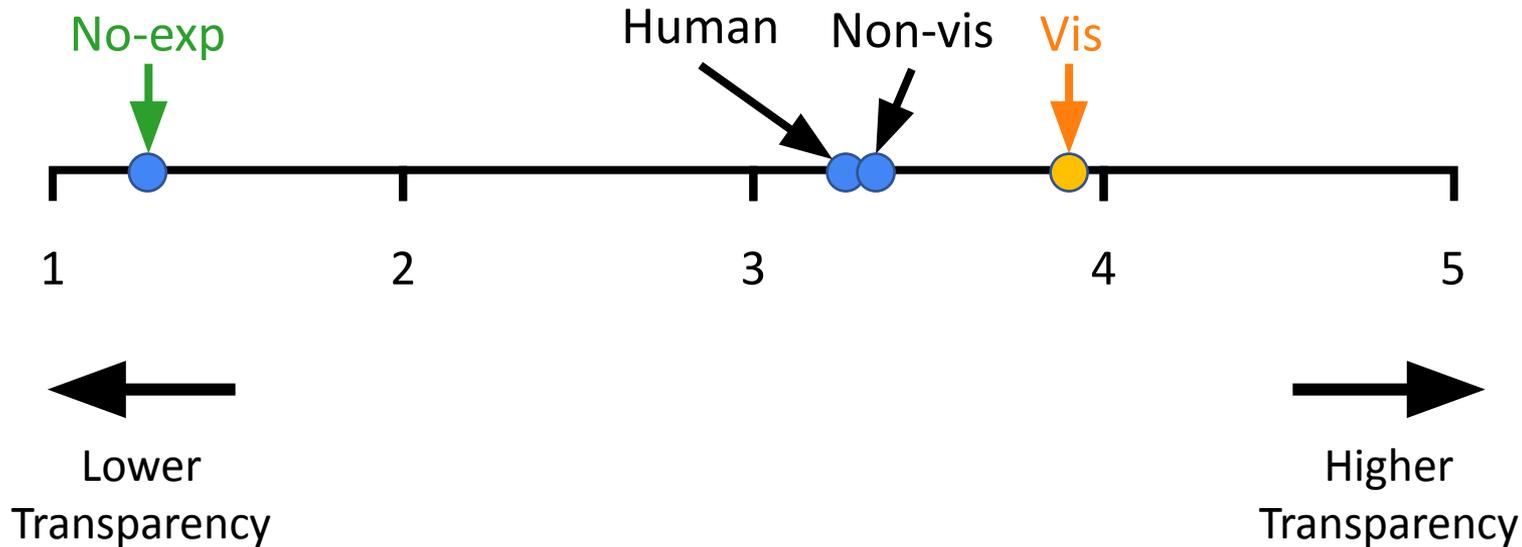
Transparency

- Presence of explanations



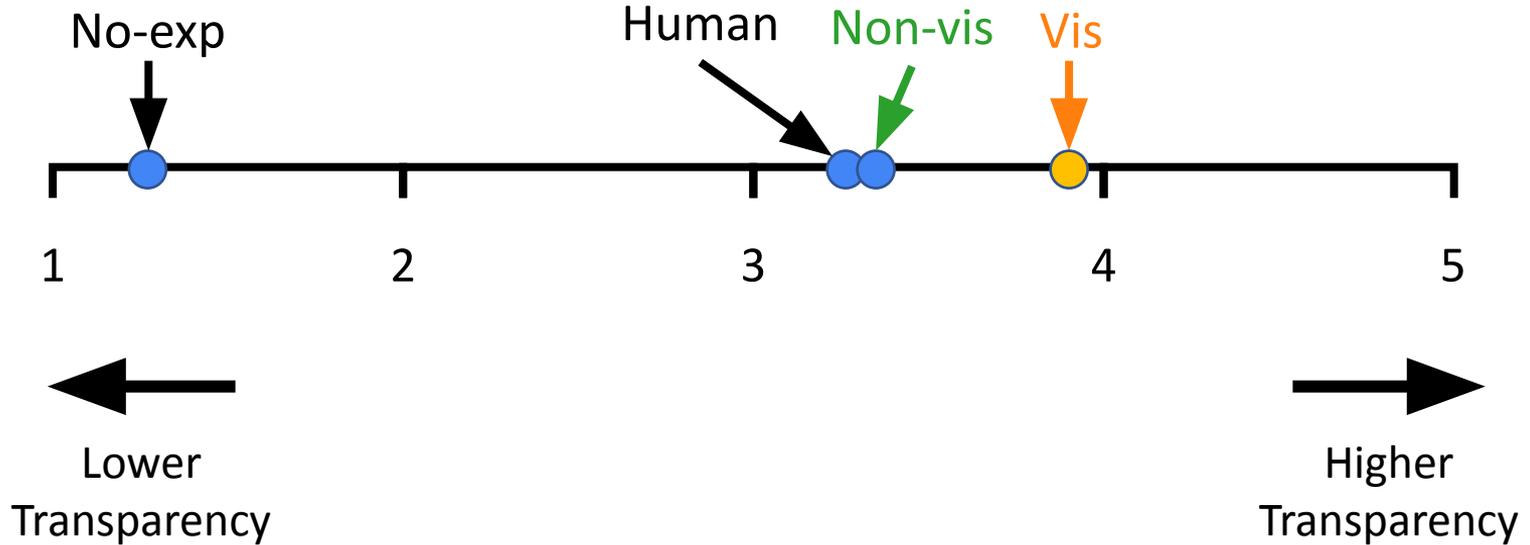
Transparency

- Presence of explanations



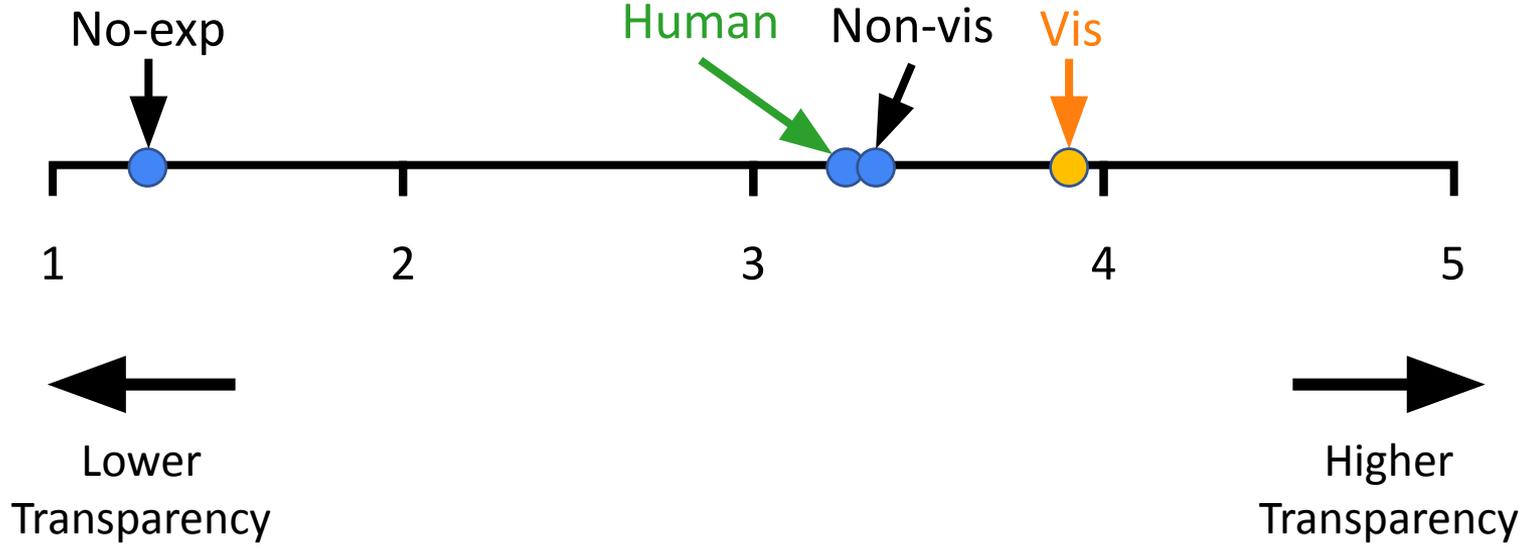
Transparency

- Presence of explanations



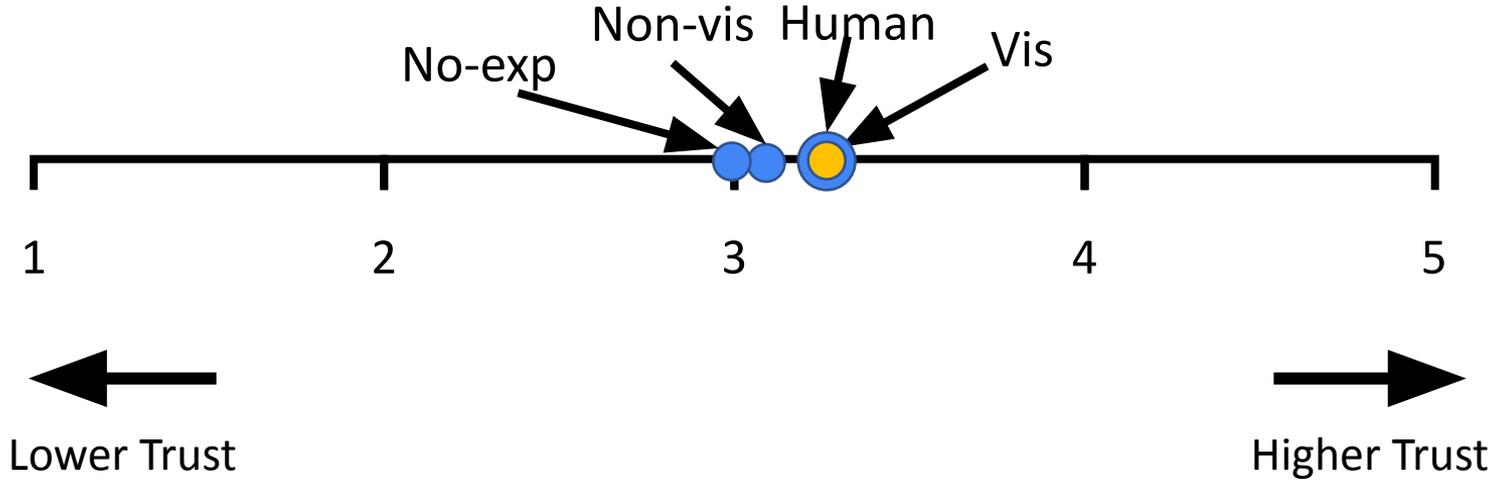
Transparency

- Presence of explanations



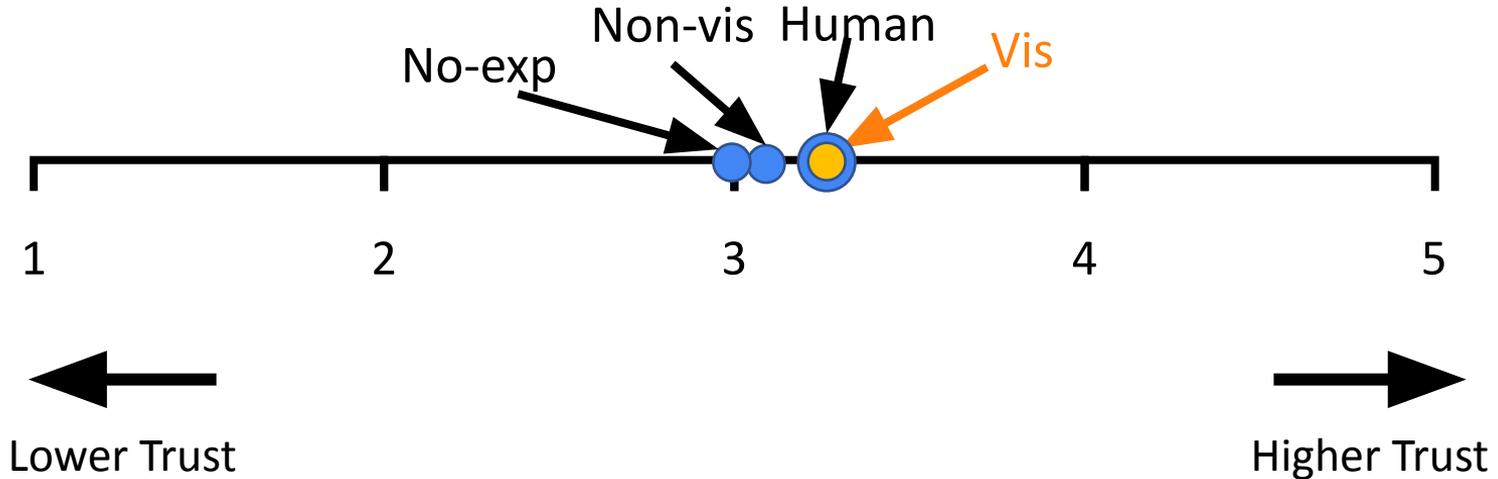
Trust

- Accuracy of answers
- Explanation-answer match



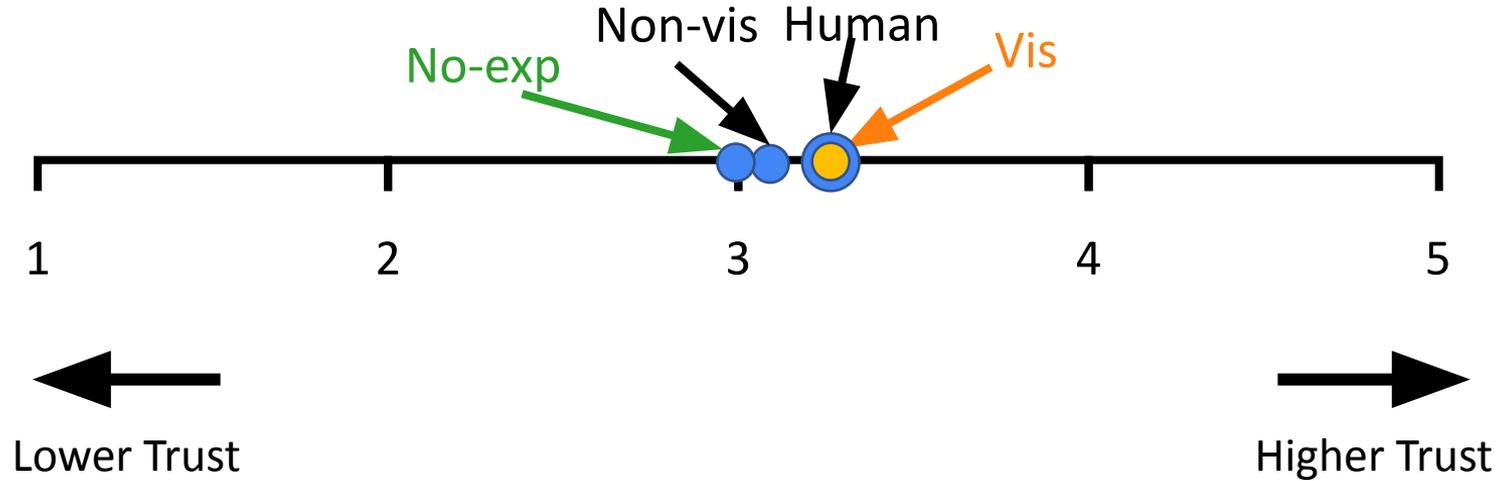
Trust

- Accuracy of answers
- Explanation-answer match



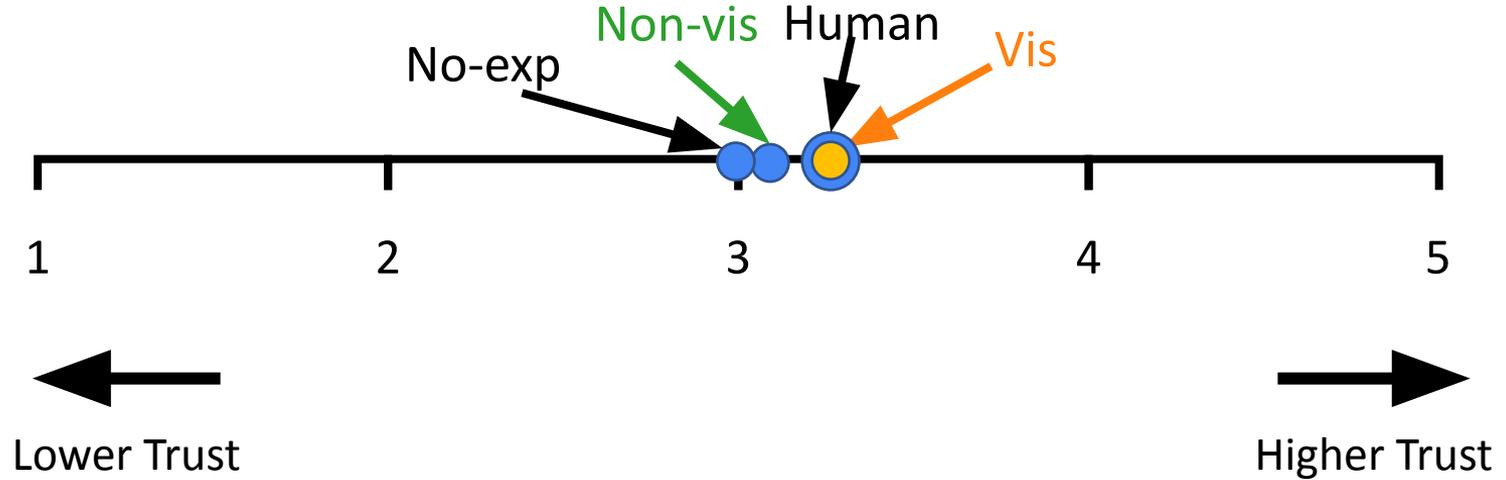
Trust

- Accuracy of answers
- Explanation-answer match



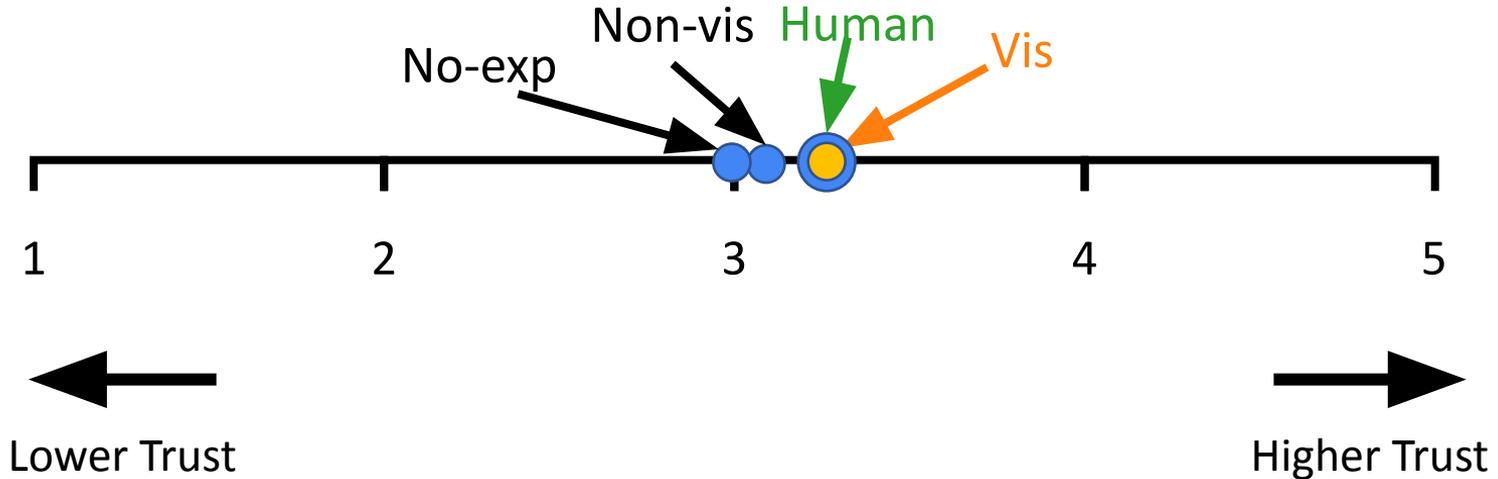
Trust

- Accuracy of answers
- Explanation-answer match



Trust

- Accuracy of answers
- Explanation-answer match

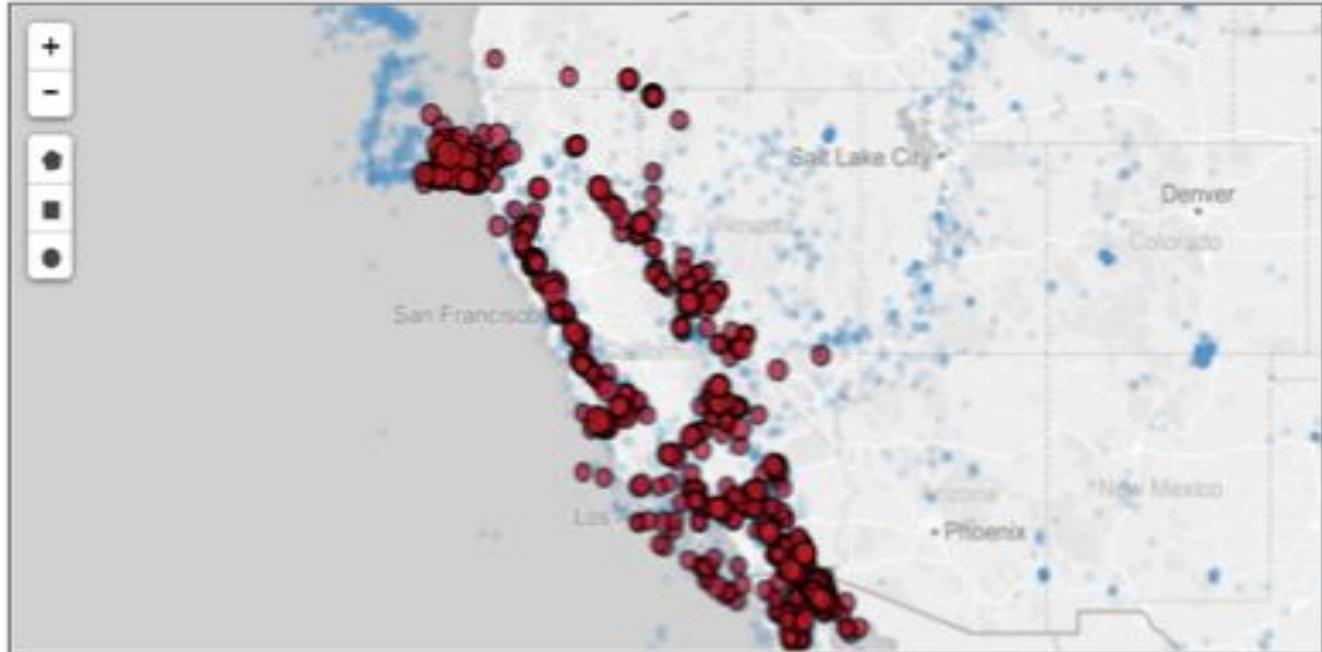


Eviza: A Natural Language Interface for Visual Analysis

[Setlur et al. 2016]

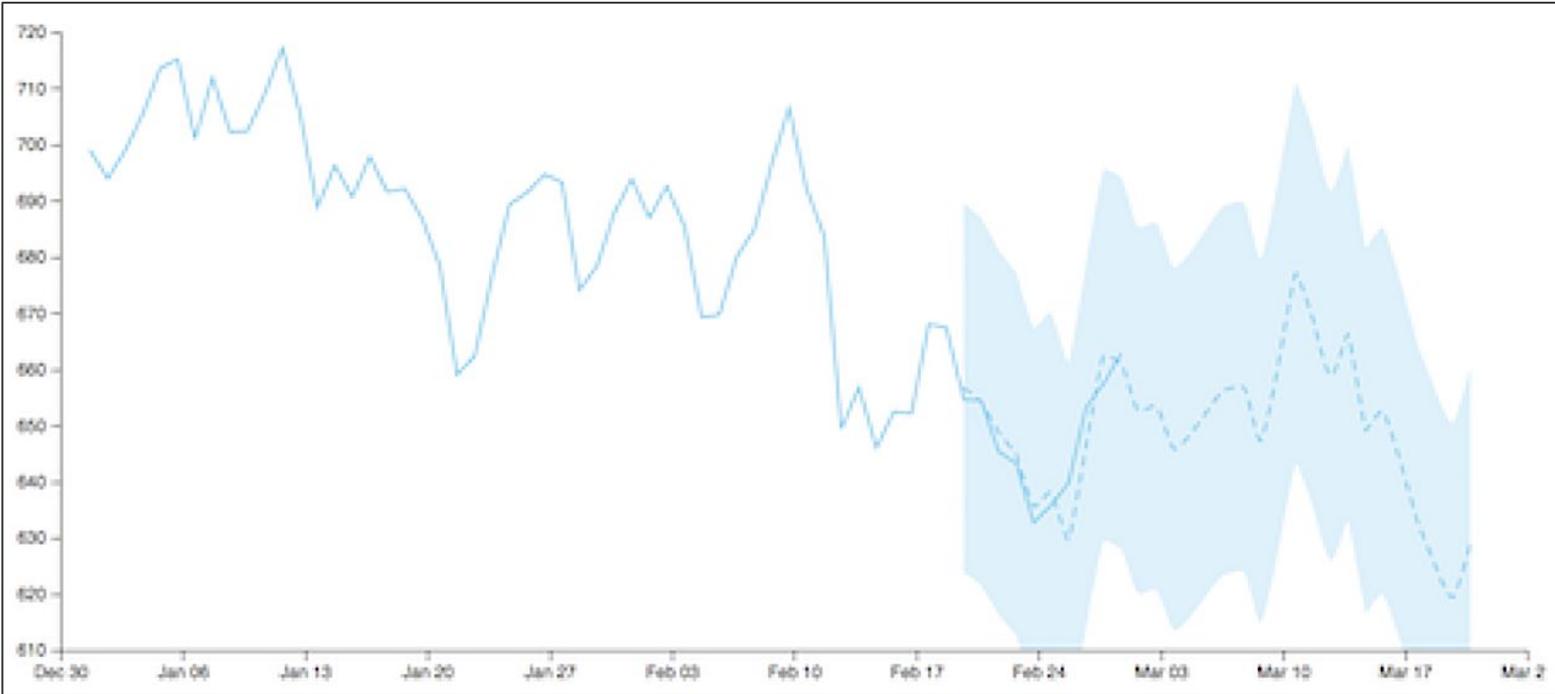
Supporting an analytical conversation

“Find large earthquakes near California”



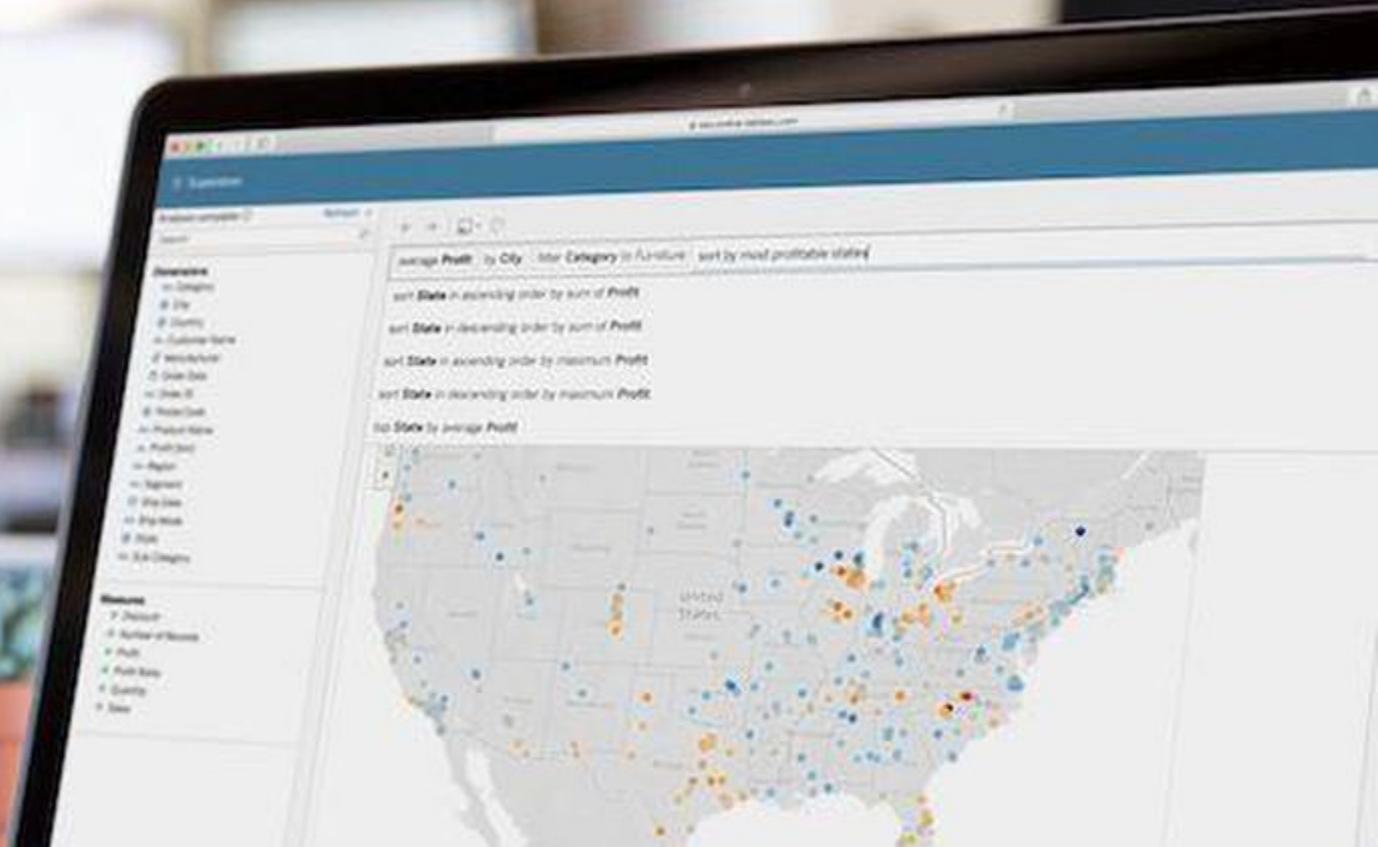
Deeper analytical conversation

“show me the trends for next month”



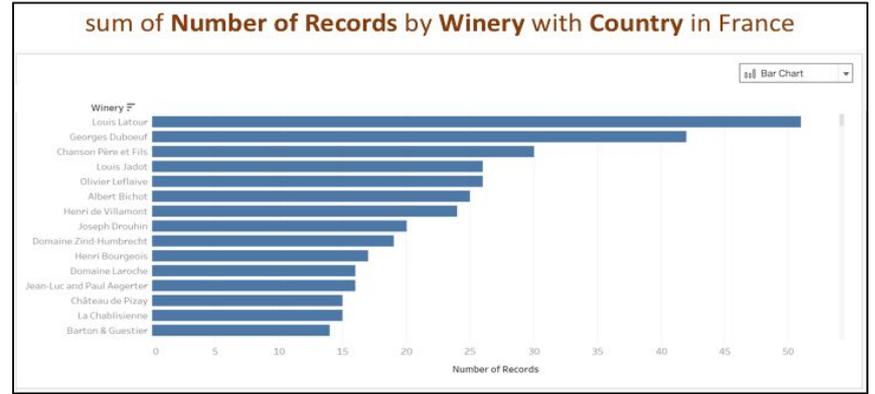
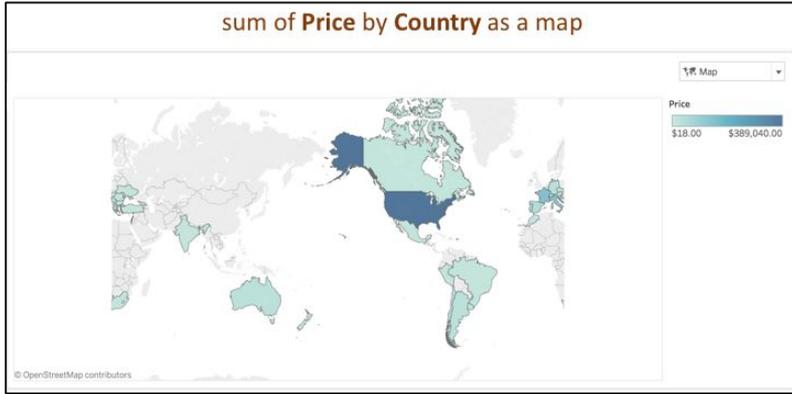
Ask Data

<https://www.tableau.com/products/new-features/ask-data>



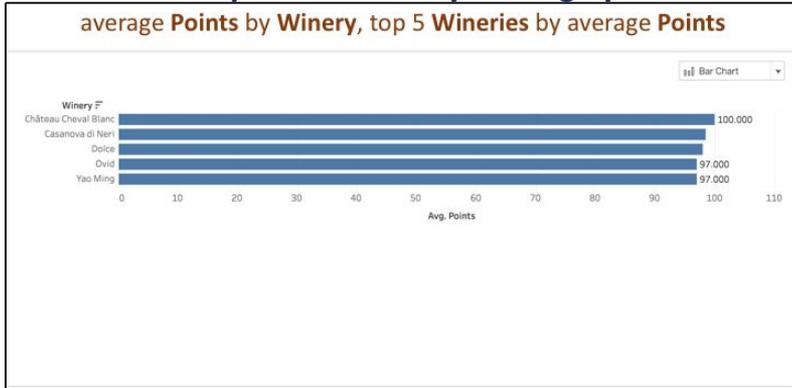
Analytical functions supported

“what’s the sum of price for each country?” “wineries in france”



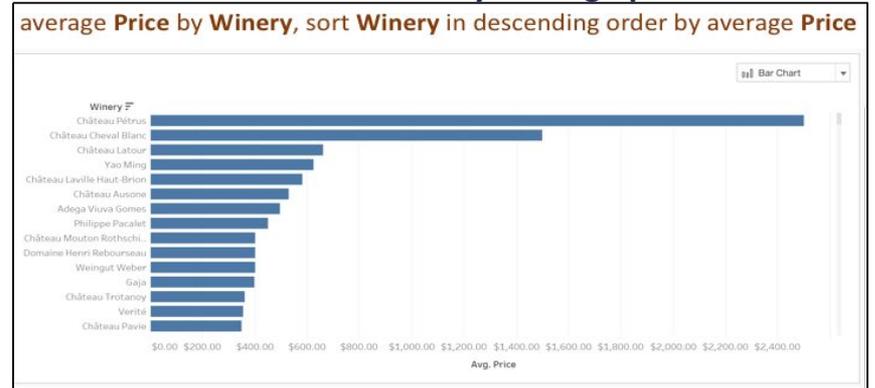
“top 5 wineries by average points”

average Points by Winery, top 5 Wineries by average Points



“sort wineries by average price”

average Price by Winery, sort Winery in descending order by average Price



AUGMENTING SEMANTICS

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
:	:	:	:	:	:	:	:

“Show me house prices”

“Show me house prices” in Ask Data

Prices \approx price

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

“Show me **expensive** house prices” in Ask Data

Expensive refers to **price**

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
:	:	:	:	:	:	:	:

Expensive: adjective; entailing great expense; very high **priced**; costly

“Show me house **cost**”

Cost is a synonym of **price**

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
:	:	:	:	:	:	:	:

Cost: amount, charge, damage, **price**, expenditure...

“Show me large houses”

Large refers to size, which can be measured as **area**

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
:	:	:	:	:	:	:	:

Large: adjective; ample in dimensions, quantity, or number. Having much size or **extent**, capacity, scope, length, breadth etc., or relatively being of more than common **measure** wide, broad, spacious, great, big, or bulky

Area: noun; a **measure** of the **extent** of a surface it is measured in square units

“Show me sqft of houses”

Sqft measures area

Datetime	price	Latitude	Longitude	area	#beds	openhouse_time	Source
1/4/2016	600000	38.8977	77.0365	5320	3	3:00pm	re.us/dfj3.php
:	:	:	:	:	:	:	:



Using word similarity

“The house is 5000 sqft.”

“The house has 5 beds.”

.

.

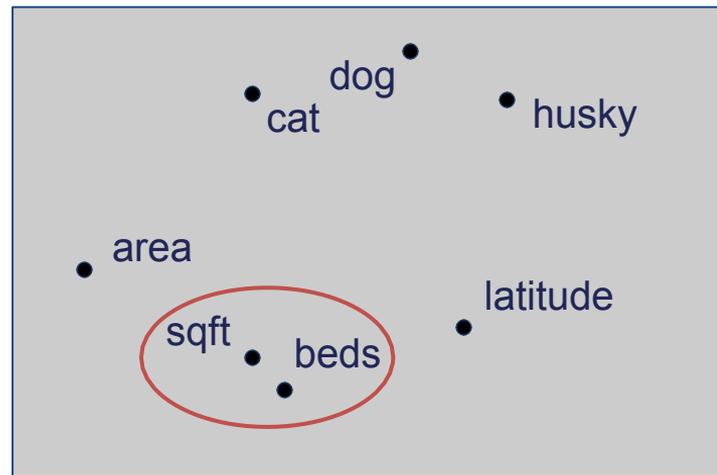
.

“I have a pet cat.”

.

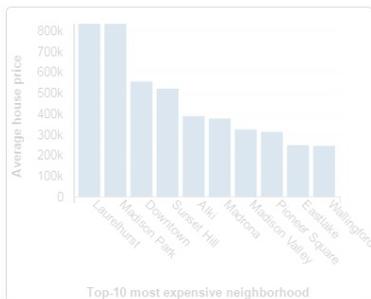
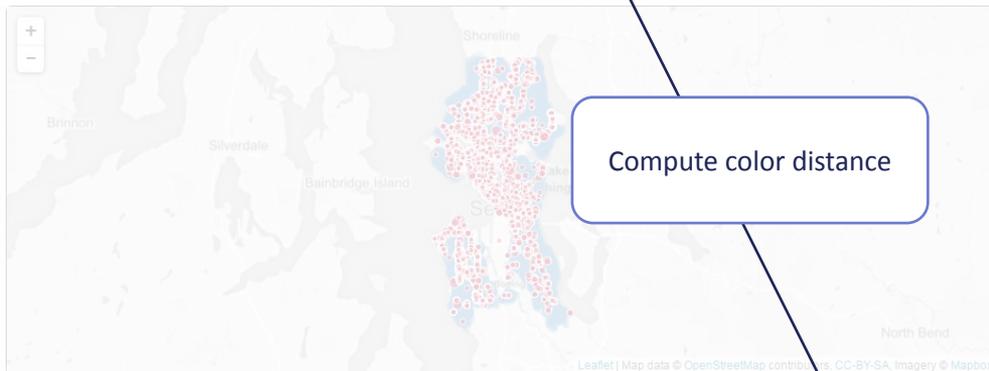
.

.

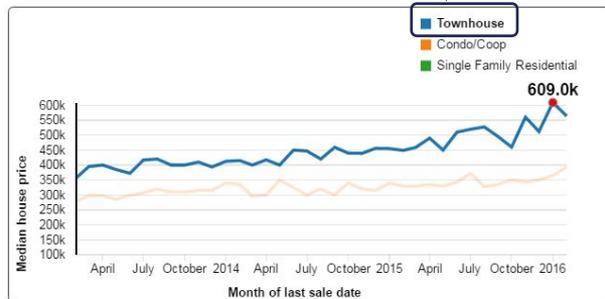


Visualization properties

Query: what's the *spike* in that **blue line**?



Neighborhood



How have sale prices changed?

External knowledge

WolframAlpha

Not logged in | Talk | Contributions | Create account | Log in



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store
- Interaction
- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page
- Tools
- What links here
- Related changes
- Upload file
- Special pages
- Permanent link
- Page information
- Wikidata item
- Print/export
- Create a book
- Download as PDF
- Printable version

Portal **Talk** Read [View source](#) [View history](#)

Portal:Current events

From Wikipedia, the free encyclopedia [Edit instructions](#)

 **Worldwide current events** · Sports events

Topics in the news

- In the Battle of Mosul, the **Great Mosque of al-Nuri** is destroyed.
- In Saudi Arabia, **Mohammad bin Salman** is appointed *Crown Prince* after the king deposes Muhammad bin Nayef.
- A vehicle driven into pedestrians** near the Finsbury Park Mosque (*pictured*), London, injures at least 10 people.
- In golf, Brooks Koepka wins the **U.S. Open**.
- In cricket, the ICC Champions Trophy concludes with Pakistan defeating India in **the final**.



Finsbury Park Mosque

Ongoing: Battle of Raqqa
Recent deaths: Con Sciacca · Otto Warmbier · Tim Hague · Baldwin Lonsdale

June 22, 2017 (Thursday) [edit](#) [history](#) [watch](#)

June 21, 2017 (Wednesday) [edit](#) [history](#) [watch](#)

Armed conflicts and attacks

- Moro conflict
 - Maute rebels storm a school in Pigcawayan, North Cotabato, Philippines, and take civilians hostage. (*The Sydney Morning Herald*)[ⓘ]
- Iraqi Civil War (2014–present)
 - Battle of Mosul (2016–present)

Time: 07:22 UTC | Day: 22 June | Purge

<< **June 2017** >>

S	M	T	W	T	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

[More June 2017 events...](#)

»

equivalents

a specified system

h equivalent ring size to ring size

equivalents

ze 5 shoe

o a specified system

11 shoe in Japanese size

OLE SIZES »

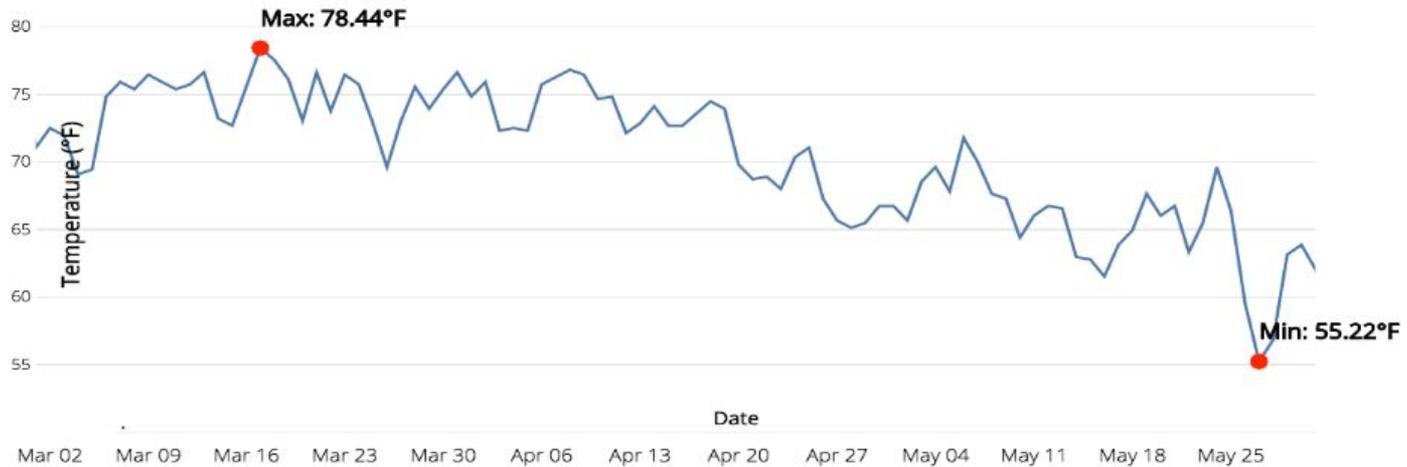
ut knitting needle sizes

ze 8

edle

“highest and lowest temperatures in Fahrenheit over fall”

Temperatures in 2014



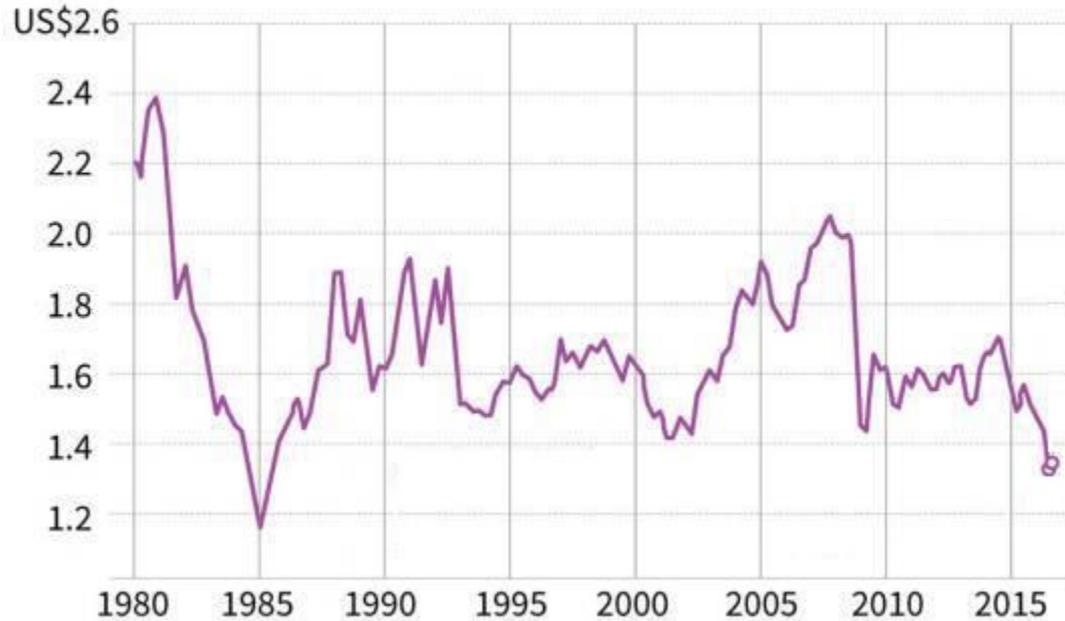
highest and lowest temperatures in Fahrenheit over seasons in

fall is from to fall .

31)
rn
;

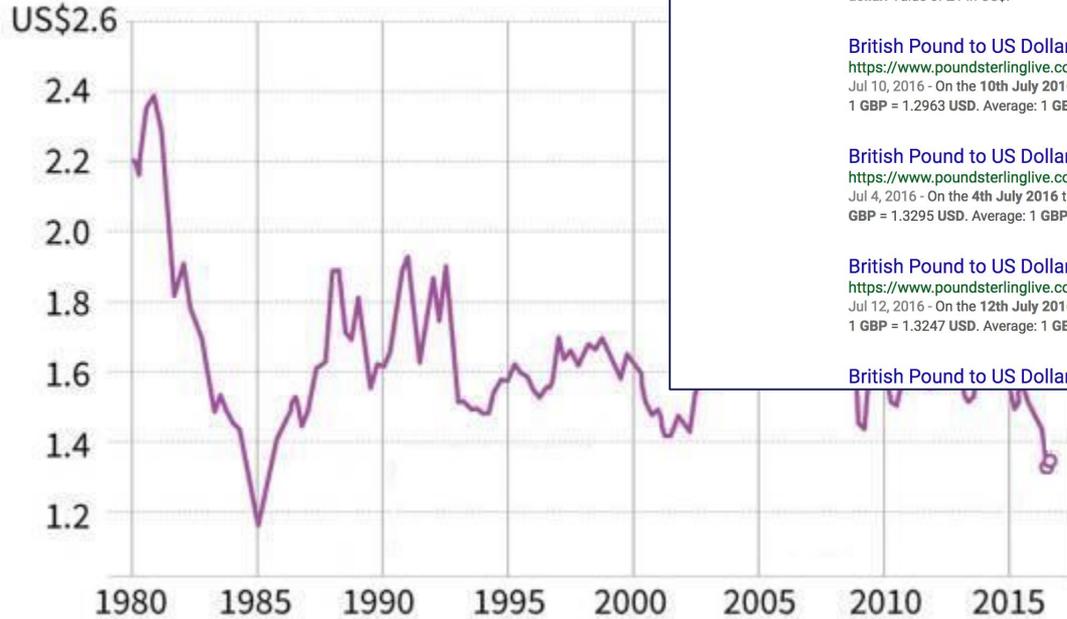
“what is the drop in july 2016?”

Pound vs dollar



“what is the drop in

Pound vs dollar



Google

pound dollar drop july 2016



All News Shopping Videos Images More Settings Tools

About 3,020,000 results (0.66 seconds)

Brexit Britain: Pound drops to \$1.28 - Jul. 6, 2016 - CNN Money

money.cnn.com/2016/07/06/investing/brexit-pound-drops/index.html

Jul 6, 2016 - by Ivana Kottasova @ivanakottasova July 6, 2016: 1:21 AM ET ... The pound has dropped roughly 15% since the referendum day, when it ...

Pound slumps to 31-year low following Brexit vote | Business | The ...

<https://www.theguardian.com/business/sterling>

Jun 24, 2016 - A woman in New York watches the pound fall on a laptop. Photograph: Andrew ... Friday 24 June 2016 02:55 EDT First published on Thursday 23 June 2016 18:15 EDT. The value of ... Pound v dollar. Value of £1 in US\$.

British Pound to US Dollar Exchange rate history: 10 July 2016 (10/07 ...

<https://www.poundsterlinglive.com/.../british-pound-to-us-dollar-exchange-rate-on-20...>

Jul 10, 2016 - On the 10th July 2016 the spot inter-bank market saw: Open: 1 GBP = 1.2908 USD. Close: 1 GBP = 1.2963 USD. Average: 1 GBP = 1.2944 ...

British Pound to US Dollar Exchange rate history: 04 July 2016 (04/07 ...

<https://www.poundsterlinglive.com/.../british-pound-to-us-dollar-exchange-rate-on-20...>

Jul 4, 2016 - On the 4th July 2016 the spot inter-bank market saw: Open: 1 GBP = 1.325 USD. Close: 1 GBP = 1.3295 USD. Average: 1 GBP = 1.3287 USD.

British Pound to US Dollar Exchange rate history: 12 July 2016 (12/07 ...

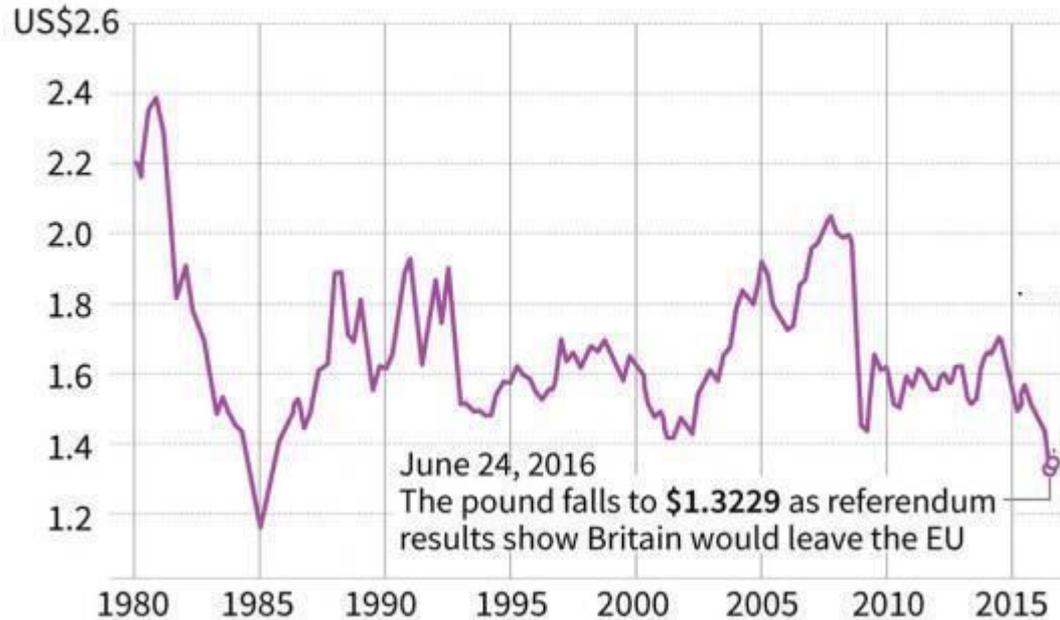
<https://www.poundsterlinglive.com/.../british-pound-to-us-dollar-exchange-rate-on-20...>

Jul 12, 2016 - On the 12th July 2016 the spot inter-bank market saw: Open: 1 GBP = 1.3002 USD. Close: 1 GBP = 1.3247 USD. Average: 1 GBP = 1.3156 ...

British Pound to US Dollar Exchange rate history: 06 July 2016 (06/07 ...

“what is the drop in July 2016?”

Pound vs dollar



Summary

Text and language play an important role in visual analysis

- Linking text with visualization
- Understanding how readers integrate charts and captions
- Visual question and answering
- Natural language interfaces