VISUALIZATION AND NATURAL LANGUAGE PROCESSING

CS 448B | Fall 2023

MANEESH AGRAWALA
“Words and pictures belong together.”

Tufte 1983

**TODAY**

**Learning Objectives**

1. Understand how people read charts and text together.

2. Visual QA with explanations for charts and graphs.
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

The evolution of GVC participation

The overall share of GVC trade in total world trade is often used as a measure of the extent to which the global economy is integrated through the fragmentation of production and trade, and the global supply chain. The share of GVC trade in total imports and exports has increased significantly in recent years, driven by both technological advances and increases in the scale and complexity of production networks.

However, the share of GVC trade has stagnated since the 2008 global financial crisis, reflecting a decline in the number of countries participating in GVCs and a decrease in the complexity of production networks. This stagnation has been attributed to factors such as the rise of protectionism, increased costs of trade, and the fragmentation of supply chains.

The evolution of GVC participation can be seen in the graph below, which shows the share of GVC trade in total imports and exports over time. The share of GVC trade peaked around the early 2000s, before declining following the global financial crisis in 2008. Since then, the share of GVC trade has remained relatively flat, with some fluctuations due to various factors such as changes in trade policies and shifts in the global economy.

The impact of the global financial crisis on GVC trade has been significant, as the crisis led to a decline in global trade volumes and a reduction in the complexity of production networks. The share of GVC trade in total imports and exports has dropped from its peak in 2008, with some recovery in recent years.

GVCs have been crucial in enabling countries to specialize in the production of specific goods and services, allowing them to take advantage of comparative advantages and improve their competitiveness. However, the global financial crisis has highlighted the vulnerabilities of GVCs, as disruptions to supply chains have had a significant impact on trade volumes and economic growth.

The future of GVCs will depend on how countries and businesses respond to these challenges, as well as the broader economic and political landscape. As economies recover and trade policies evolve, it is likely that GVCs will continue to play a significant role in global trade and economic growth, but they may take on a different form and structure.

READING CHARTS AND TEXT

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

12/6/23
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?
Macron’s approval rating steeply dropped between June and August of 2017.
Macron’s approval rating steeply dropped between June and August of 2017.
The 30-year fixed mortgage rate increased slightly from 1997 to 1999.
The 30-year fixed mortgage rate increased \textit{slightly} from 1997 to 1999.

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

---

**EmphasisChecker: A Tool for Guiding Chart and Caption Emphasis**
Kim, Choi, Kim, Setlur and Agrawala, IEEE Visualization 2023
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?
For which religion did the fewest chaplains think that religious extremism is common?

Orthodox Christians, Hindus.
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.

Formative Study
Which religion has the greatest value for Common?
Which religion has the greatest value for Common?

Answer: Muslims

Explanation: I picked religions with the greatest orange percentage.

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

“I picked religions with the greatest orange percentage.”
Green component in the top bar

Glabron at University Farm
Chart QA Pipeline

Question about Chart
Which religion has the shortest orange component?
Which religion has the shortest orange component?

**Chart**

<table>
<thead>
<tr>
<th>Religion</th>
<th>Common</th>
<th>Not Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muslims</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Pagan/earth-based</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>Protestants</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>Jews</td>
<td>17</td>
<td>83</td>
</tr>
<tr>
<td>Buddhists</td>
<td>7</td>
<td>93</td>
</tr>
<tr>
<td>Hindus</td>
<td>6</td>
<td>94</td>
</tr>
</tbody>
</table>

**Question about Chart**

Which religion has the shortest orange component?

**Question about Table**

Which religion has the least ‘Common’ Percentage?
Question about Chart
Which religion has the shortest orange component?

Question about Table
Which religion has the least 'Common' Percentage?

Answer
Orthodox Christians, Hindus.
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).

Answer

Orthodox Christians, Hindus.

Question about Chart
Which religion has the shortest orange component?

Lambda Expression
\[
\text{argmin}\{R[\text{Religion}].\text{Row}, R[\lambda x (R[\text{Number}].R[\text{Common}].\text{Religion}.x)]\}
\]
**Answer**

Orthodox Christians, Hindus.

---

**Chart**

**Data Table**

<table>
<thead>
<tr>
<th>Religion</th>
<th>Common</th>
<th>Not common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muslims</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Pagan/earth-based</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>Protestants</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>Jews</td>
<td>17</td>
<td>83</td>
</tr>
<tr>
<td>Buddhists</td>
<td>7</td>
<td>93</td>
</tr>
<tr>
<td>Hindus</td>
<td>6</td>
<td>94</td>
</tr>
</tbody>
</table>

**Question about Chart**

Which religion has the shortest orange component?

**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’.

---

**Chart**

**Data Table**

<table>
<thead>
<tr>
<th>Religion</th>
<th>Common</th>
<th>Not common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muslims</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Pagan/earth-based</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>Protestants</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>Jews</td>
<td>17</td>
<td>83</td>
</tr>
<tr>
<td>Buddhists</td>
<td>7</td>
<td>93</td>
</tr>
<tr>
<td>Hindus</td>
<td>6</td>
<td>94</td>
</tr>
</tbody>
</table>

**Question about Chart**

Which religion has the shortest orange component?

**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’.

**Visual Explanation**

I looked up ‘Religion’ for the shortest orange bar.
Example Explanations
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?

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 fewest chaplains think that religious extremism is common?

Q: fewest chaplains think that religious extremism is common?

Hypothesis: Visual explanations increase transparency and trust

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

Q: Orthodox Christians, Hindus. I looked up ‘Religion’ for the shortest orange bar.

A1 (vis): Orthodox Christians, Hindus.

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

A3 (non-vis): Orthodox Christians, Hindus. I looked up ‘Religion’ with the lowest value for ‘Common’.
**Hypothesis:** Visual explanations increase transparency and trust

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

Q: fewest chaplains think that religious extremism is common?


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’.

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 → Evaluate Transparency, Trust & Usefulness

16 participants

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

Determine Answer Correctness → Evaluate Transparency, Trust & Usefulness → Free-Form Response
**TRANSPARENCY**

- No-exp
- Human
- Non-vis
- Vis

Lower Transparency  |  Higher Transparency

**TRUST**

- No-exp
- Non-vis
- Human
- Vis

Lower Trust  |  Higher Trust
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.*

ANNOUNCEMENTS
FINAL PROJECT

Final Code and Video due Sun Dec 10, 8pm

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
  Project proposal: Mon 11/6
  Design Review and Feedback: 9th week of quarter, 11/27 and 11/29
  Final code and video: Sun 12/10 8pm

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

COURSE EVALUATION

Official course evaluation, due by 12/18
Your opinion is valued!

http://course-evaluations.stanford.edu
COURSE SUMMARY

DATA AND IMAGE MODELS

Graphics and Graphic Information Processing [Bertin 1981]
VISUALIZATION DESIGN AND REDESIGN

Tableau: based on Polaris [Stolte 2002]

EXPLORATORY DATA ANALYSIS
**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]

---

**INTERACTION**

Baby name voyager [Wattenberg 2005]
INTRODUCTION TO D3

D3: Data Driven Documents [Bostock 2011]

PERCEPTION

Graphical Perception [Cleveland and McGill 1984]
VISUAL EXPLAINERS

Watch how the measles outbreak spreads when kids get vaccinated - and when they don't

![Diagram showing measles outbreak spread with different vaccinated and susceptible rates](image)

74.4% vac rate, similar to Island County, WA
83.8% vac rate, similar to Santa Cruz County, CA
86.0% vac rate, similar to Los Angeles County, CA
90.0% vac rate, similar to Orange County, CA
99.7% vac rate, similar to Gadsden County, FL

COLOR

![ColorBrewer 5-class diverging RdGy](image)

Cynthia Brewer [http://www.personal.psu.edu/faculty/c/a/cab38/](http://www.personal.psu.edu/faculty/c/a/cab38/)
DECONSTRUCTING VISUALIZATIONS

[Graph and bar chart showing original and two redesigns]

ReVision [Savva 2001]

VISUALIZATION AND AI

[Complex diagram with various elements and icons]

114

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 creating visualizations and designing techniques?

Thank You!