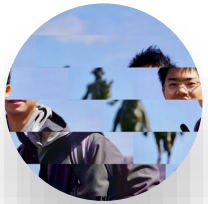
VIS 2023

Mystique: Deconstructing SVG Charts for Layout Reuse


Chen Chen


Bongshin Lee

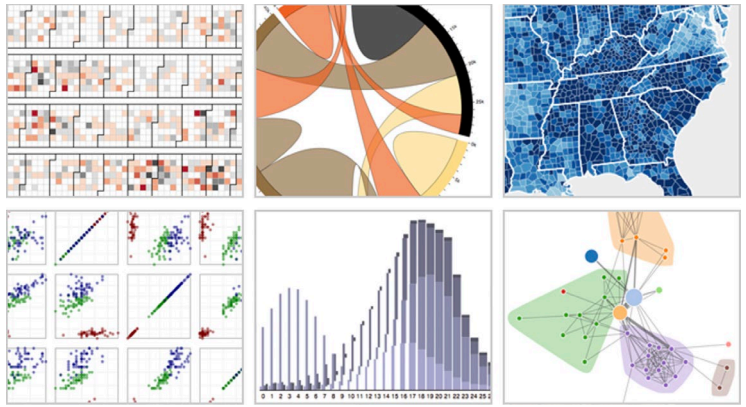

Yunhai Wang


Yunjeong Chang

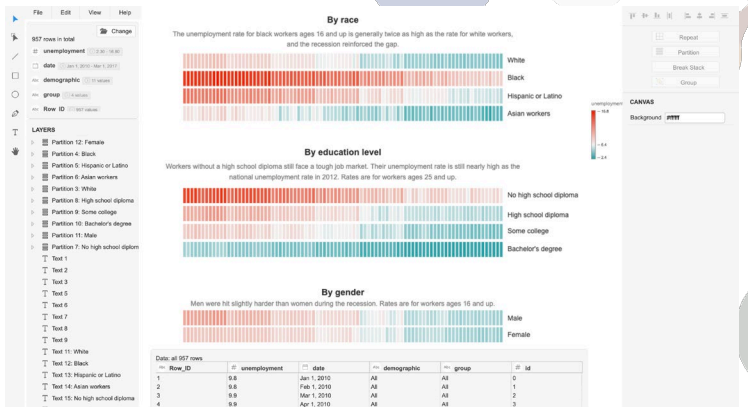

Zhicheng Liu

University of Maryland Department of Computer Science · HCIL · HDI · Microsoft · Shandong University 山东大学

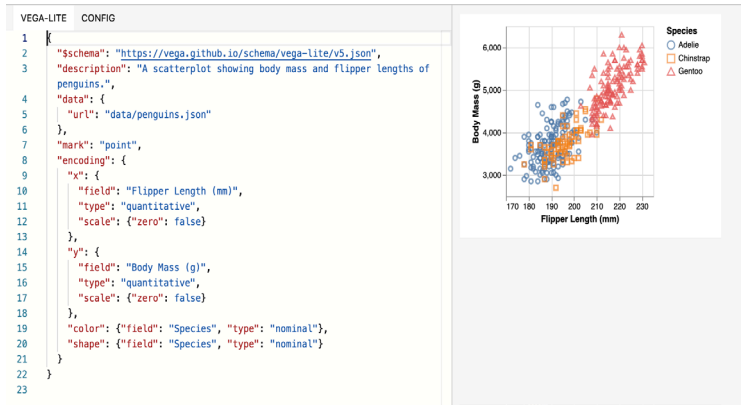
Online Galleries Provide Reusable Examples



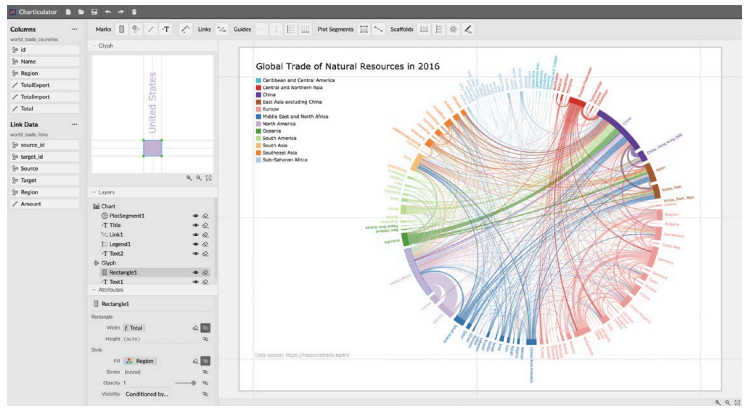
D3, InfoVis 2011



Data Illustrator, CHI 2018

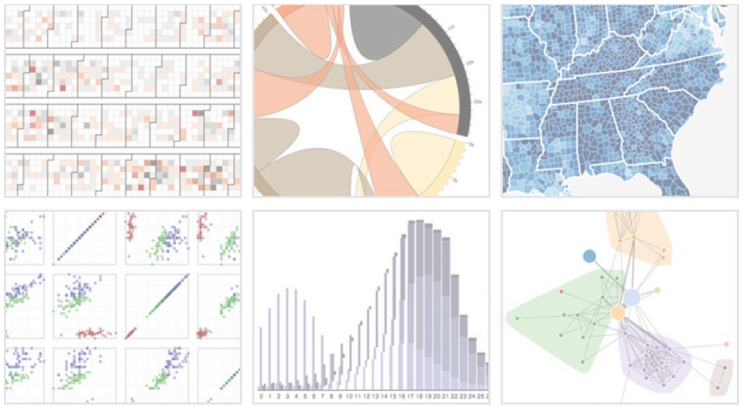


Vega-Lite, InfoVis 2017

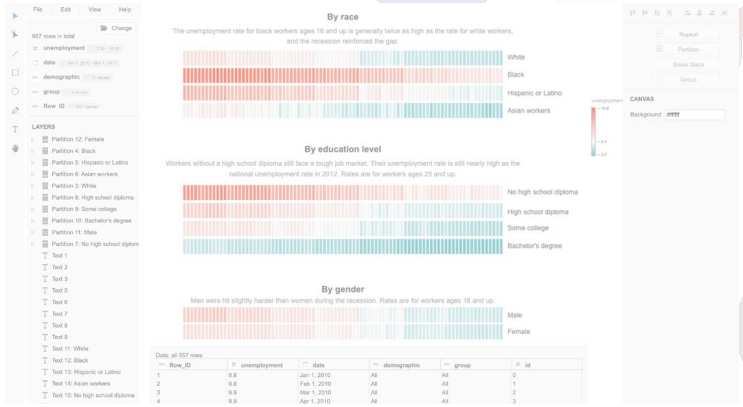


Charticulator, InfoVis 2018

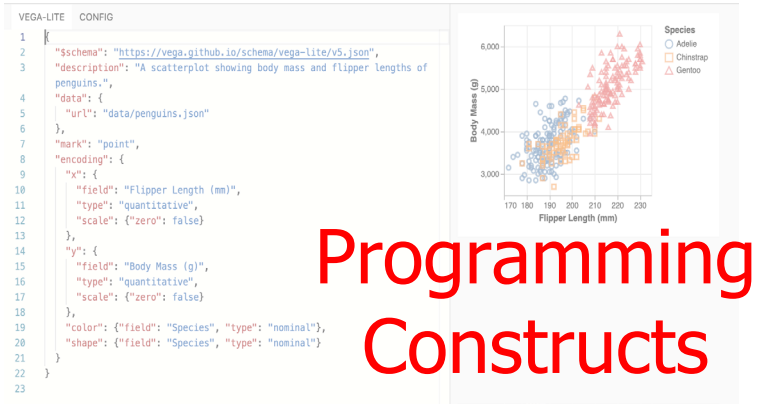
Re-purposing Visualizations is Still Challenging



D3, InfoVis 2011

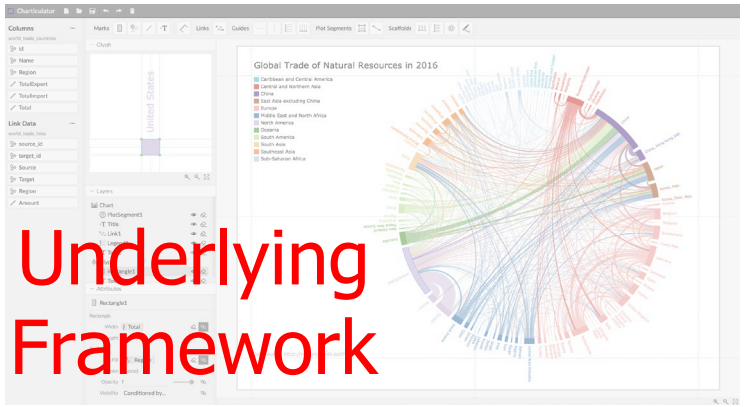


Data Illustrator, CHI 2018



Programming Constructs

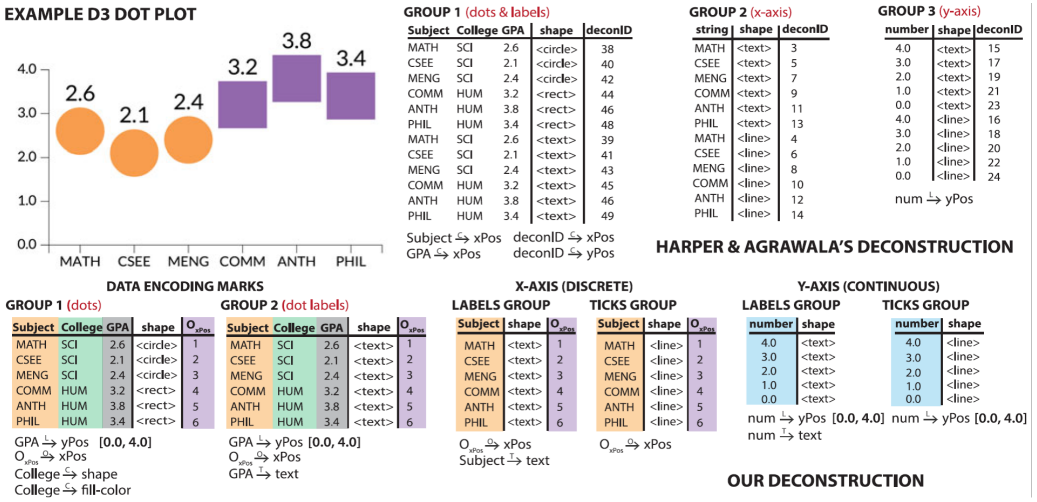
Vega-Lite, InfoVis 2017



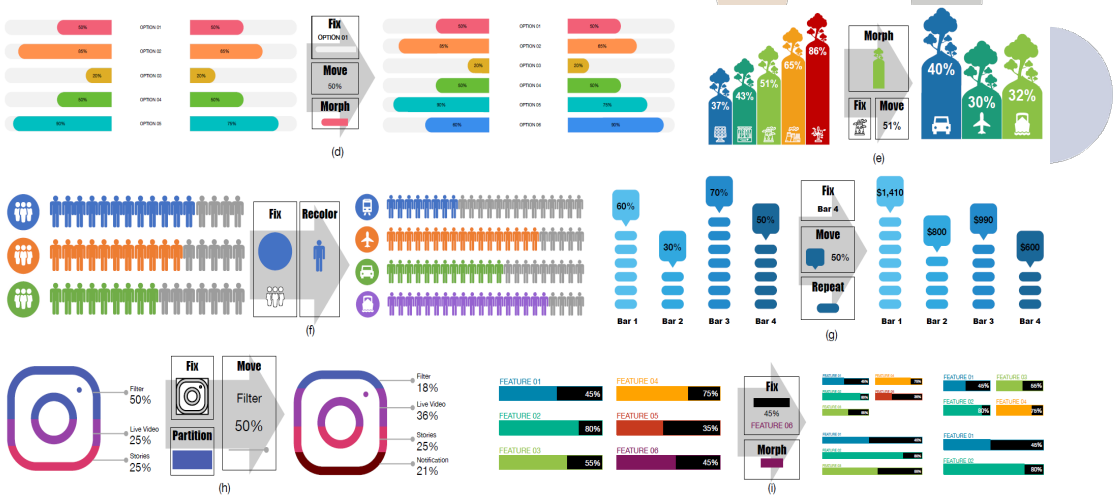
Underlying Framework

Charticulator, InfoVis 2018

Decompose and Reuse Charts



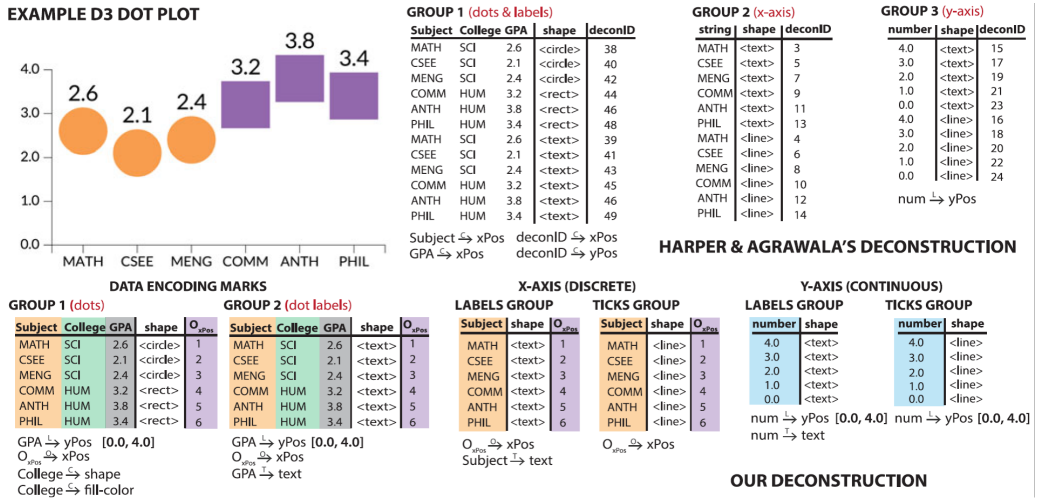
D3 Deconstructor, TVCG 2018

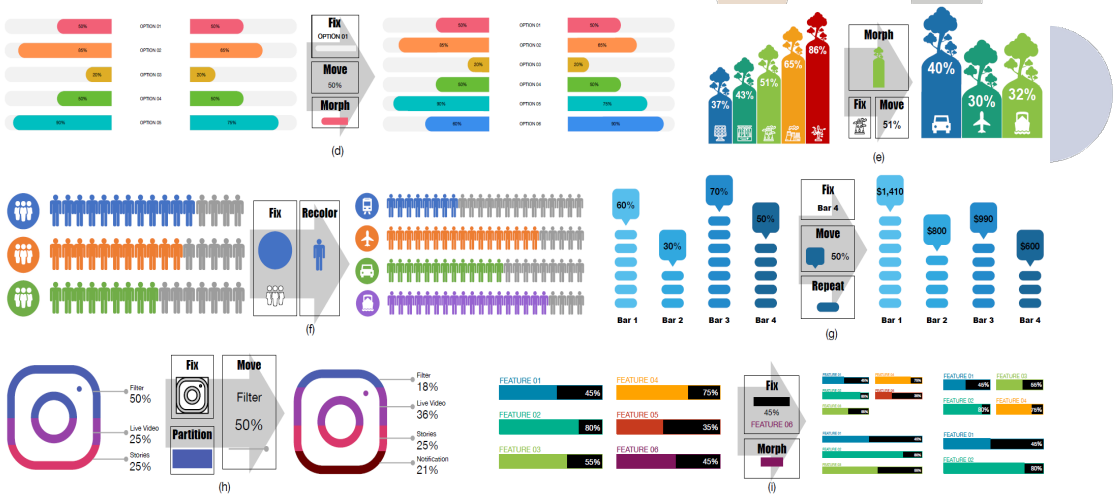


ChartReuse, TVCG 2022

Decompose and Reuse Charts

Limitations





D3 Deconstructor, TVCG 2018

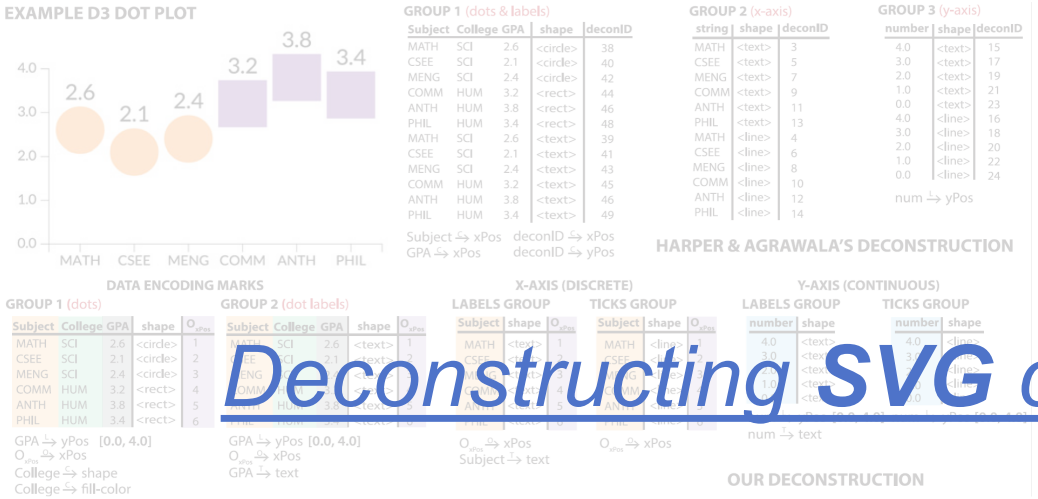
- D3 charts
- Basic chart types

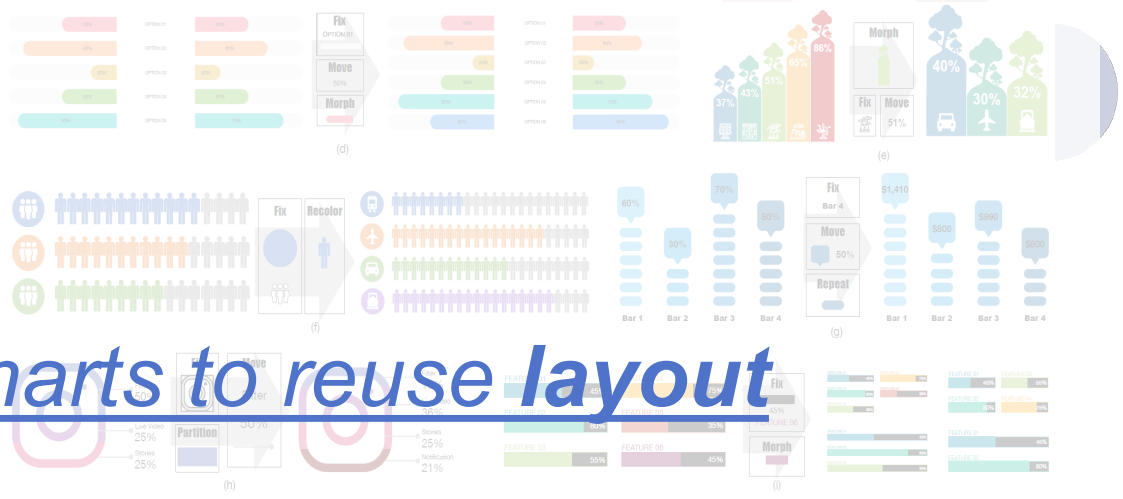
ChartReuse, TVCG 2022

- PowerPoint infographics
- Glyph-based bar charts

Decompose and Reuse Charts

Our Aim





*Deconstructing **SVG** charts to reuse **layout***

D3 Deconstructor, TVCG 2018

- D3 charts
- Basic chart types

ChartReuse, TVCG 2022

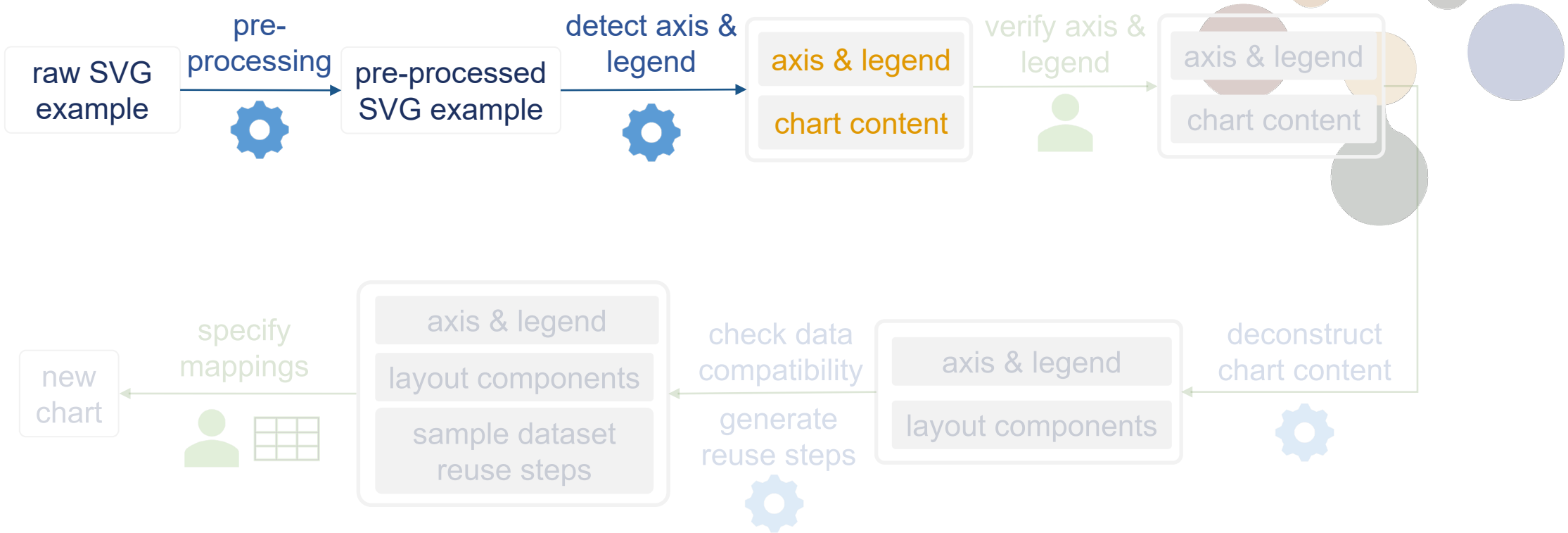
- PowerPoint infographics
- Glyph-based bar charts

Scope on SVG Charts for Layout Reuse

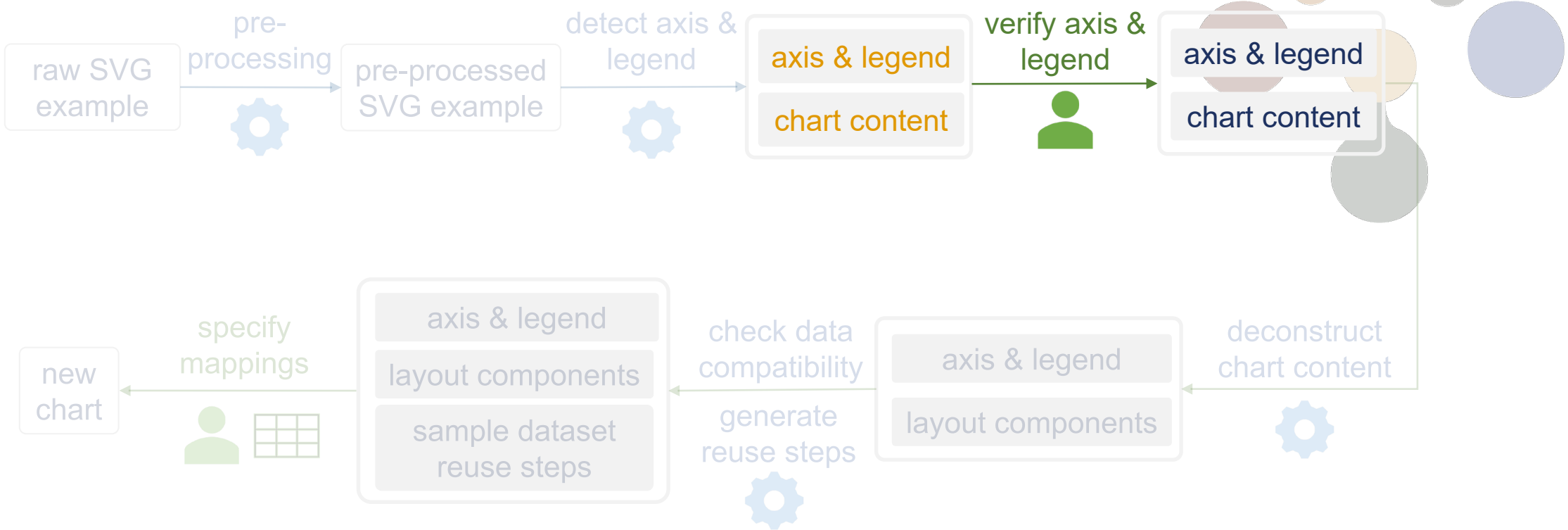
An Investigation into the Beagle Dataset

Mark	Chart	Percentage
Rectangle	bar chart (histogram), grouped bar chart, stacked bar chart, diverging bar chart (pyramid chart), Marimekko chart, heatmap, bullet chart, treemap, waffle chart, waterfall chart, range chart, gantt chart, matrix chart, cartogram, calendar chart	32.85%
Line	line graph, parallel coordinates, Kagi chart	30.51%
Pie	pie chart, donut chart	16.50%
Circle	scatter plot, bubble plot, dot plot, circle packing	14.96%
Others	geographic map, area chart, stream graph, chord chart, hexbin plot, Sankey diagram, Voronoi diagram, word cloud, sunburst chart, boxplot, network diagram, contour plot, radial plot	5.18%

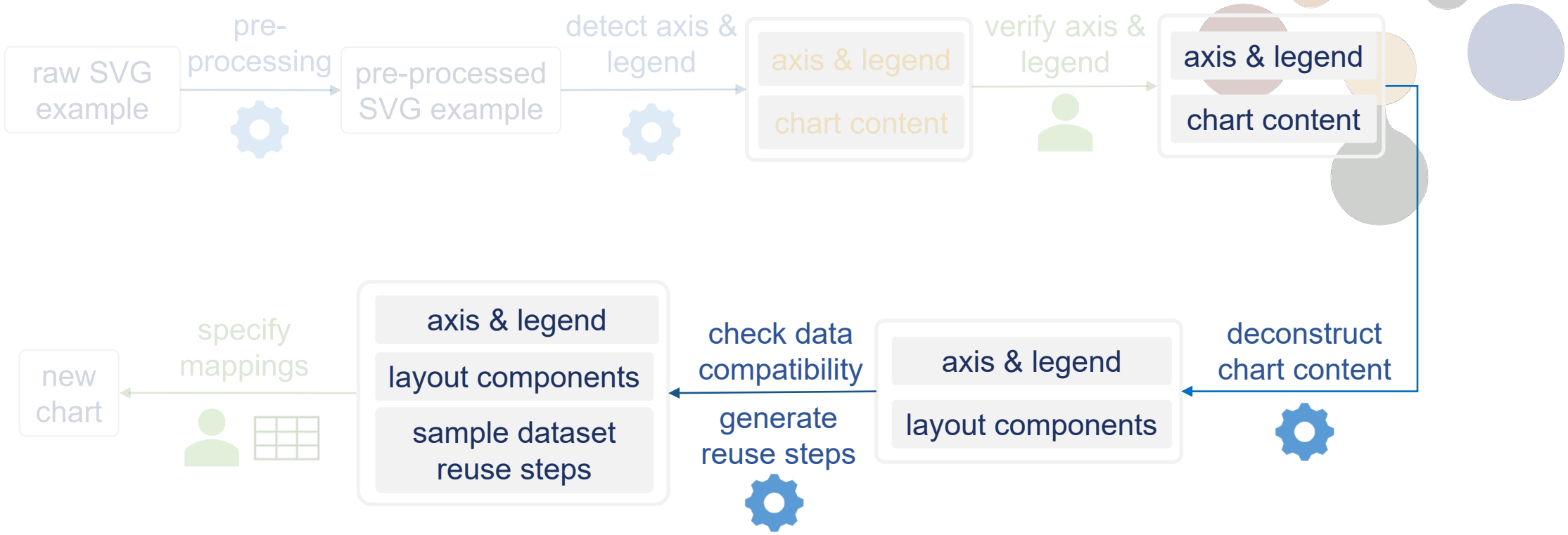
Pipeline


raw SVG example
pre-processing
pre-processed SVG example
detect axis & legend
axis & legend
chart content
verify axis & legend
axis & legend
chart content
deconstruct chart content
axis & legend
layout components
generate reuse steps
check data compatibility
axis & legend
layout components
sample dataset reuse steps
specify mappings
new chart

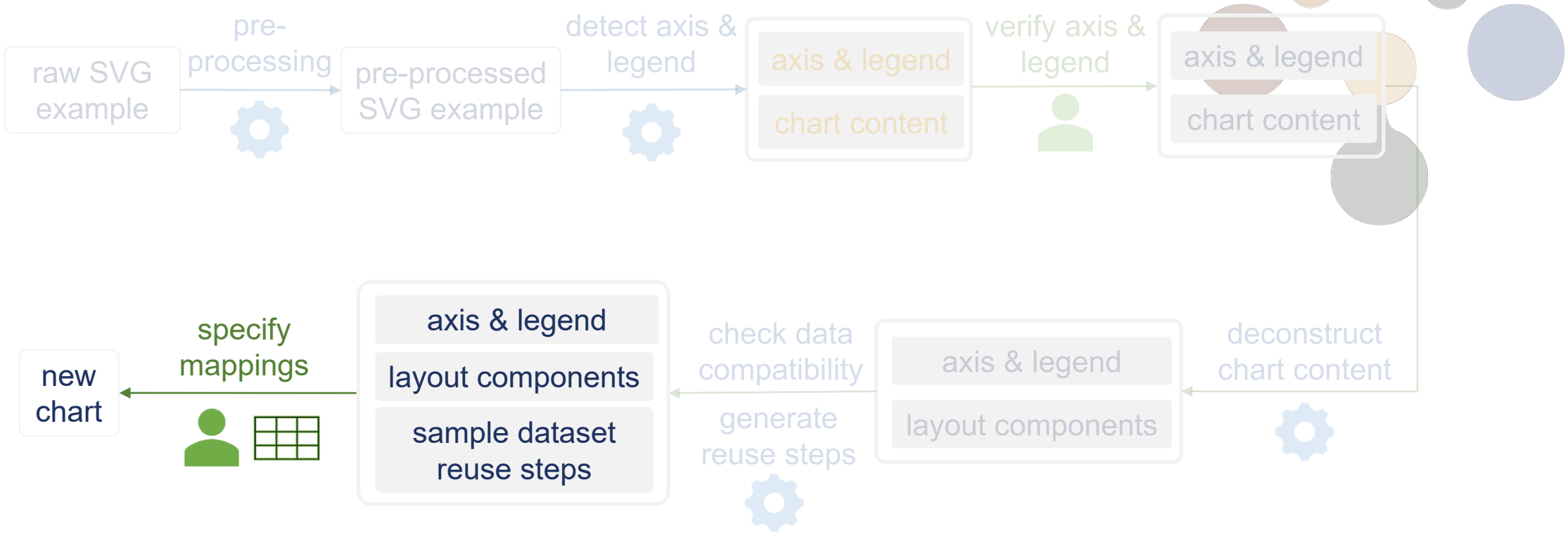
Pipeline


raw SVG example
pre-processing
pre-processed SVG example
detect axis & legend
axis & legend
chart content
verify axis & legend
axis & legend
chart content
deconstruct chart content
axis & legend
layout components
check data compatibility
generate reuse steps
axis & legend
layout components
sample dataset reuse steps
specify mappings
new chart

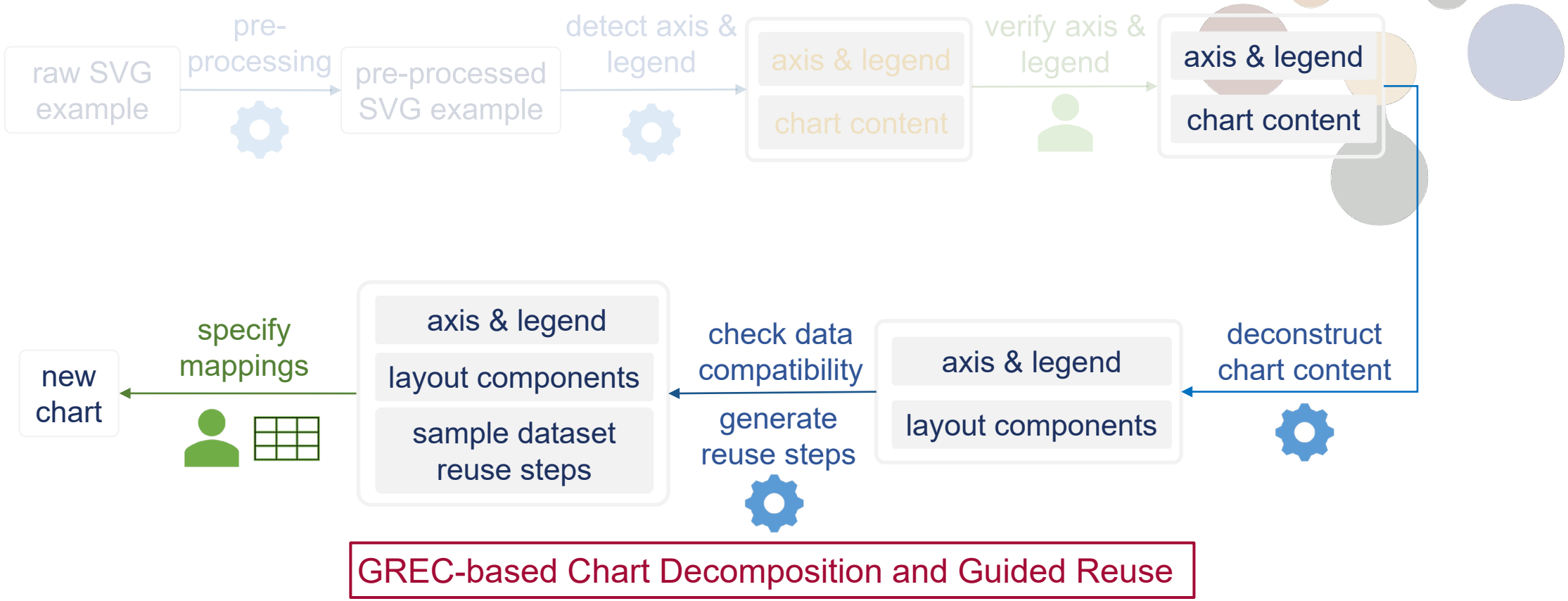
Pipeline

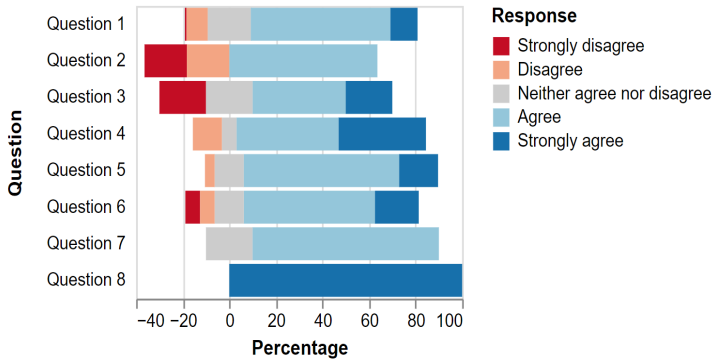

raw SVG example
pre-processing
pre-processed SVG example
detect axis & legend
axis & legend
chart content
verify axis & legend
axis & legend
chart content
deconstruct chart content
axis & legend
layout components
check data compatibility
generate reuse steps
axis & legend
layout components
sample dataset reuse steps
specify mappings
new chart

Pipeline


raw SVG example
pre-processing
pre-processed SVG example
detect axis & legend
axis & legend
chart content
verify axis & legend
axis & legend
chart content
deconstruct chart content
axis & legend
layout components
check data compatibility
generate reuse steps
axis & legend
layout components
sample dataset reuse steps
specify mappings
new chart

Pipeline


raw SVG example
pre-processing
pre-processed SVG example
detect axis & legend
axis & legend
chart content
verify axis & legend
axis & legend
chart content
deconstruct chart content
check data compatibility
generate reuse steps
axis & legend
layout components
sample dataset
reuse steps
specify mappings
new chart
GREC-based Chart Decomposition and Guided Reuse

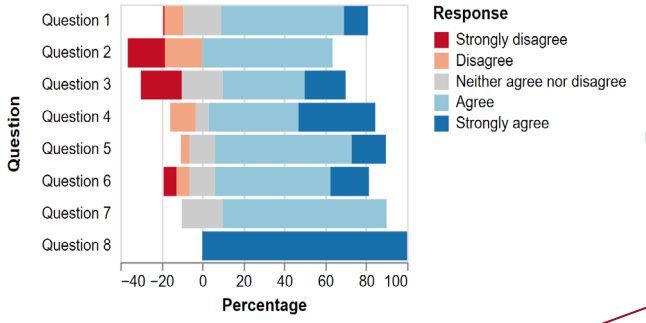


GREC-based Chart Decomposition

{

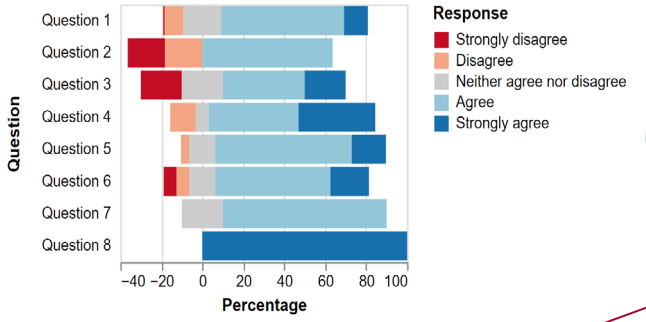
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}

}



GREC-based
Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}



GREC-based Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

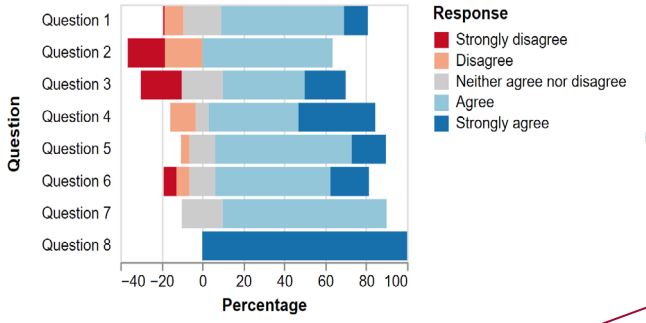
Group 1

Stack **R**elationship

Group 2

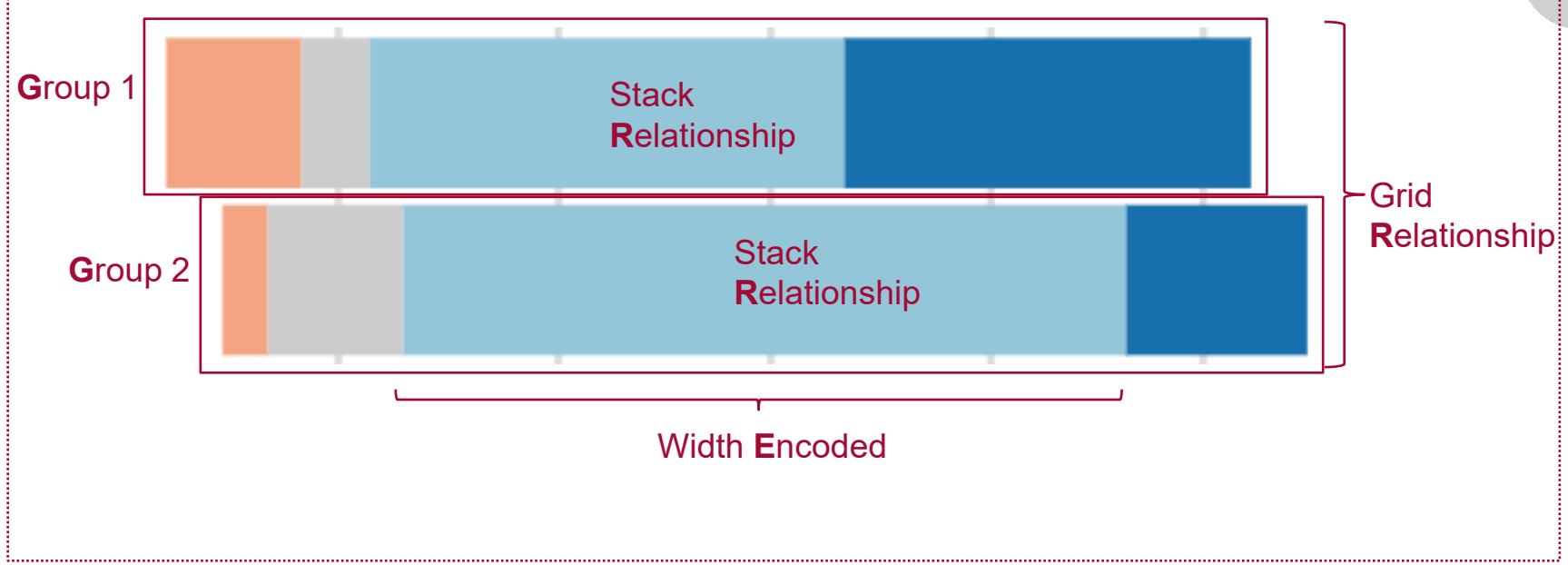
Stack **R**elationship

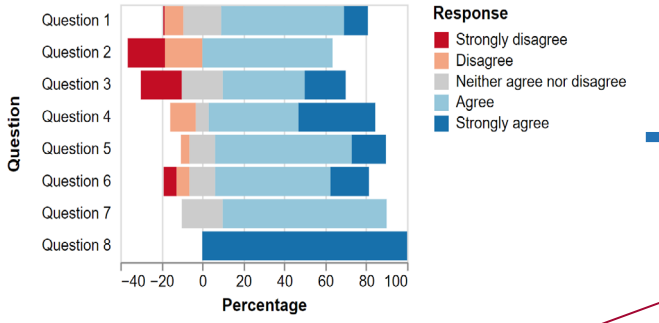
Grid **R**elationship



GREC-based Chart Decomposition

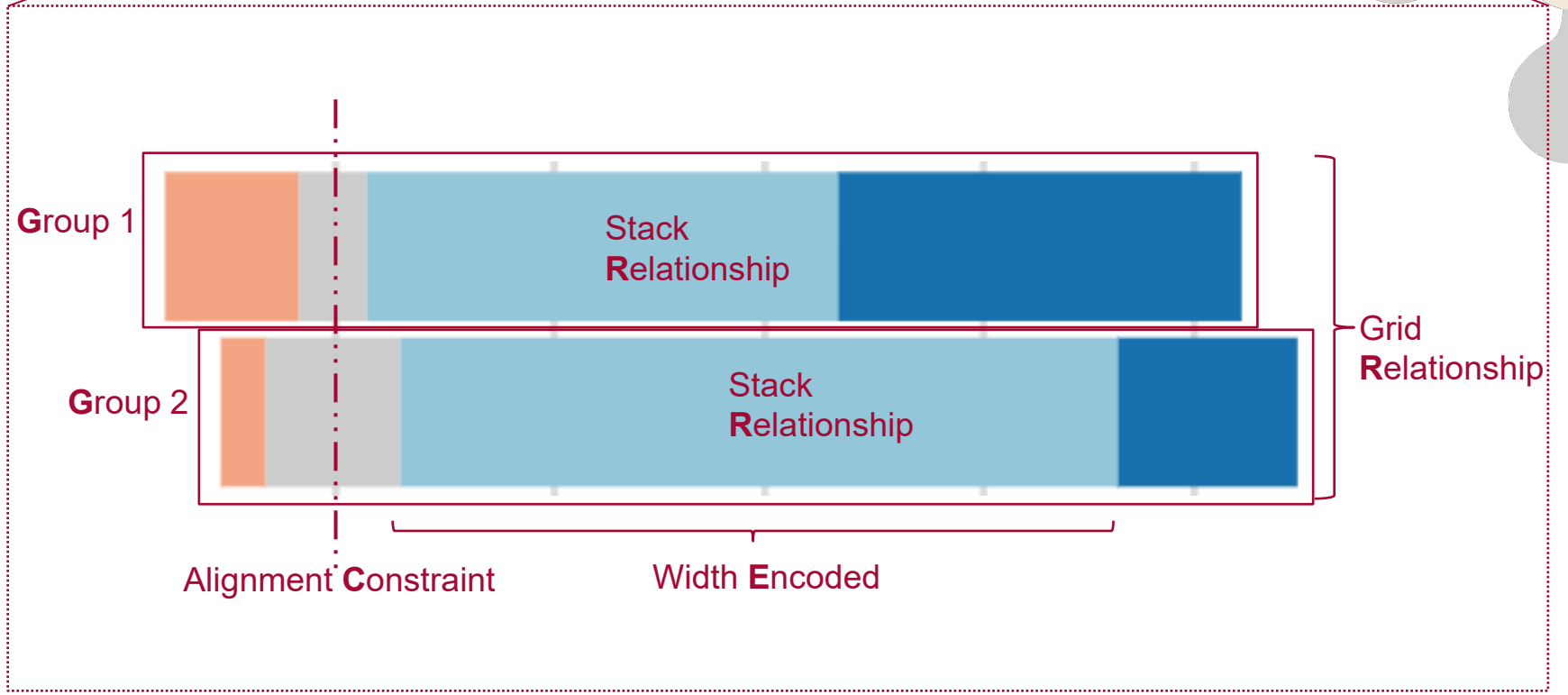
{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

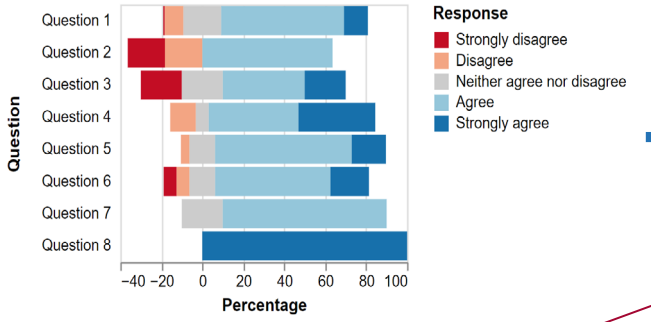




GREC-based Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}



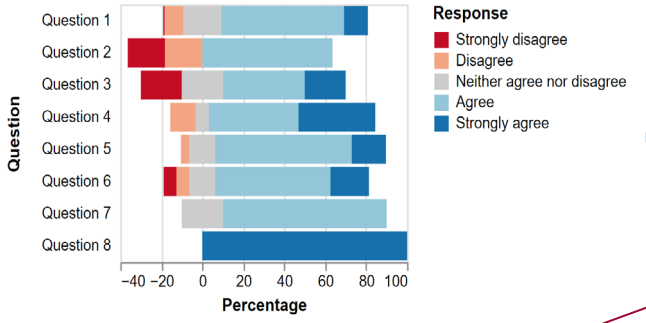


GREC-based Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

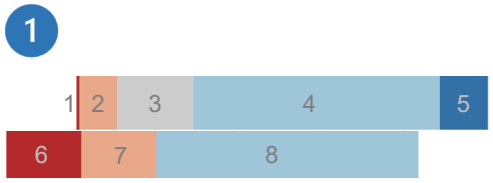
A bottom-up **hierarchical clustering algorithm** to capture the semantic groups and relationships of rectangles.

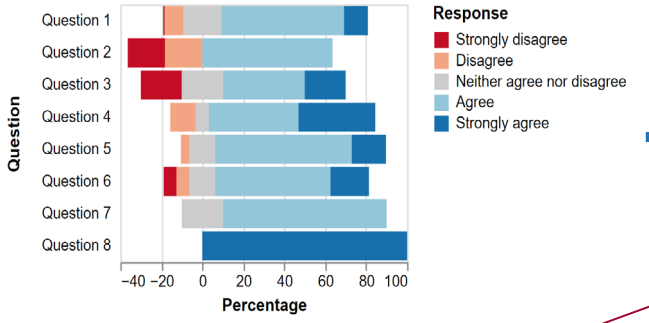
- Distance function: estimate the spacial relationship between two graphical objects
- Linkage function: merge graphical objects into higher-level groups



GREC-based
Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}





GREC-based Chart Decomposition

{

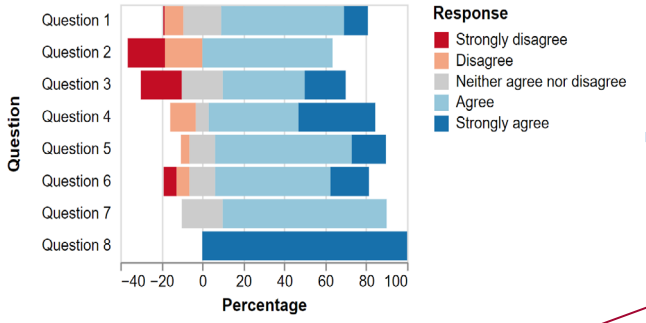
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}

}

1

2

	1	2	3	4	5	6	7	8
1		HS	X	X	X	VG	X	X
2	HS		HS	X	X	X	X	X
3	X	HS		HS	X	X	X	X
4	X	X	HS		HS	X	X	X
5	X	X	X	HS		X	X	X
6	VG	X	X	X	X		HS	X
7	X	X	X	X	X	HS		HS
8	X	X	X	X	X	X	HS	



GREC-based Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

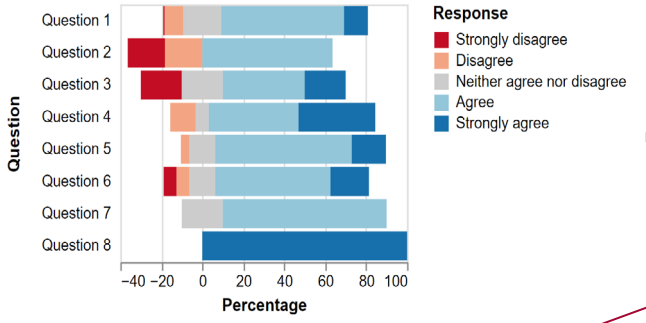
1

2

	1	2	3	4	5	6	7	8
1		HS	X	X	X	VG	X	X
2	HS		HS	X	X	X	X	X
3	X	HS		HS	X	X	X	X
4	X	X	HS		HS	X	X	X
5	X	X	X	HS		X	X	X
6	VG	X	X	X	X		HS	X
7	X	X	X	X	X	HS		HS
8	X	X	X	X	X	X	HS	

3

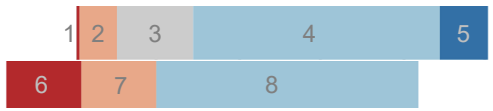
2 collections of HS:
C1=[1, 2, 3, 4, 5], C2=[6, 7, 8]



GREC-based Chart Decomposition

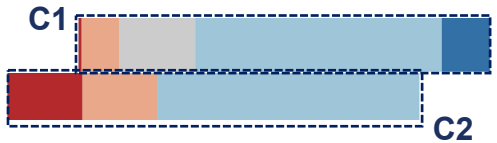
{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

1



3

2 collections of HS:
C1=[1, 2, 3, 4, 5], C2=[6, 7, 8]

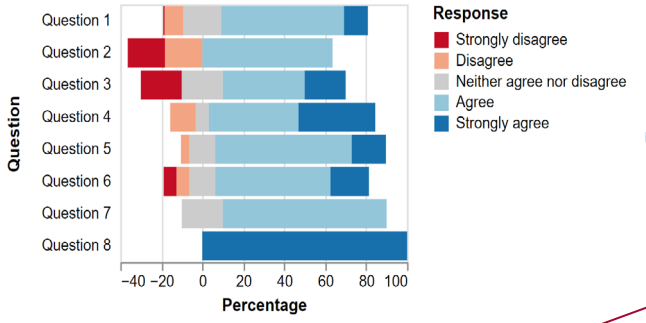


2

	1	2	3	4	5	6	7	8
1		HS	X	X	X	VG	X	X
2	HS		HS	X	X	X	X	X
3	X	HS		HS	X	X	X	X
4	X	X	HS		HS	X	X	X
5	X	X	X	HS		X	X	X
6	VG	X	X	X	X		HS	X
7	X	X	X	X	X	HS		HS
8	X	X	X	X	X	X	HS	

4

	C1	C2
C1		VG
C2	VG	



GREC-based Chart Decomposition

{
Grouping: {...},
Relationships: {...},
Encodings: {...},
Constraints: {...}
}

1

2

	1	2	3	4	5	6	7	8
1		HS	X	X	X	VG	X	X
2	HS		HS	X	X	X	X	X
3	X	HS		HS	X	X	X	X
4	X	X	HS		HS	X	X	X
5	X	X	X	HS		X	X	X
6	VG	X	X	X	X		HS	X
7	X	X	X	X	X	HS		HS
8	X	X	X	X	X	X	HS	

3

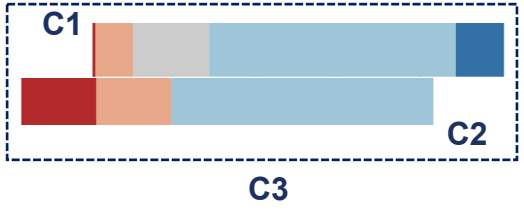
2 collections of HS:
C1=[1, 2, 3, 4, 5], C2=[6, 7, 8]

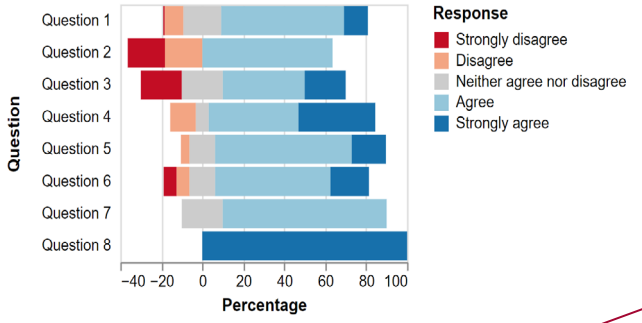
4

	C1	C2
C1		VG
C2	VG	

5

1 collection C3=[C1, C2] of VG





GREC-based Chart Decomposition

{
Grouping: {...},
Relationships: {...},
Encodings: {...},
Constraints: {...}
}

1

2

	1	2	3	4	5	6	7	8
1		HS	X	X	X	VG	X	X
2	HS		HS	X	X	X	X	X
3	X	HS		HS	X	X	X	X
4	X	X	HS		HS	X	X	X
5	X	X	X	HS		X	X	X
6	VG	X	X	X	X		HS	X
7	X	X	X	X	X	HS		HS
8	X	X	X	X	X	X	HS	

3

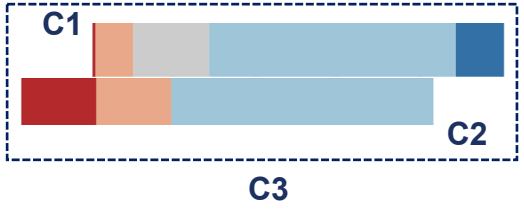
2 collections of HS:
C1=[1, 2, 3, 4, 5], C2=[6, 7, 8]

4

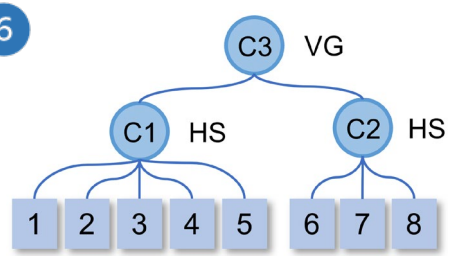
	C1	C2
C1		VG
C2	VG	

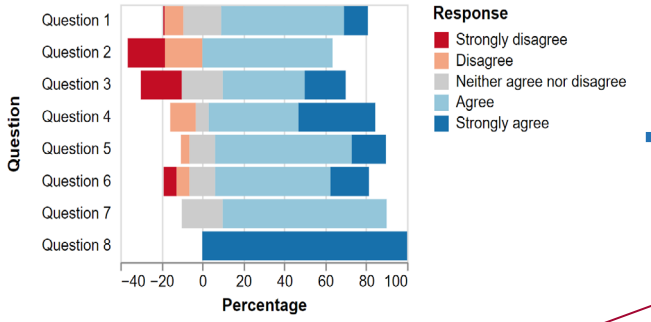
5

1 collection C3=[C1, C2] of VG



6





GREC-based Chart Decomposition

{
Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}
}

Encodings:

Channel	Condition
fill	rectangles in the chart content have different fill colors
area	lowest-level spatial relationship is packing
width/height	lowest-level spatial relationship is grid or stack, and rectangles have varying widths/heights
x/y	lowest-level spatial relationship is a one-directional grid without the gravity parameter

Constraints:

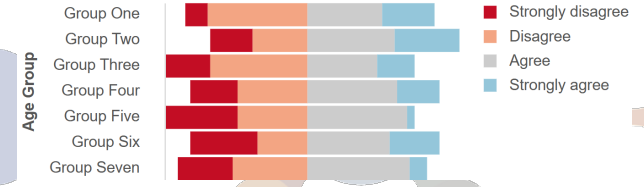
- alignment constraint within a glyph (e.g., bullet charts)
- customized alignment of stacked rectangles in a grid relationship

{

Grouping: {…},
Relationships: {…},
Encodings: {…},
Constraints: {…}

}

Guided Chart Reuse



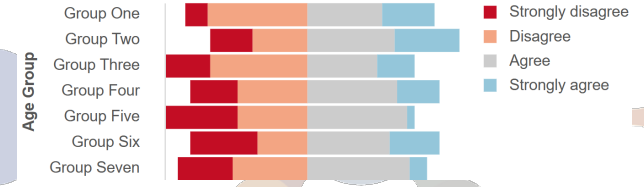
Inferring Data Schema:

- Estimate the required #categorical_data and #quantitative_data
- generates a sample dataset for the given example

Your dataset should have at least 2 categorical and at least 1 quantitative columns. See Sample Data Import Data

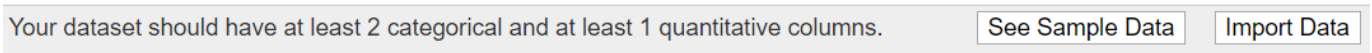
```
{
    Grouping: {...},
    Relationships: {...},
    Encodings: {...},
    Constraints: {...}
}
```

Guided Chart Reuse



Inferring Data Schema:

- Estimate the required #categorical_data and #quantitative_data
- generates a sample dataset for the given example



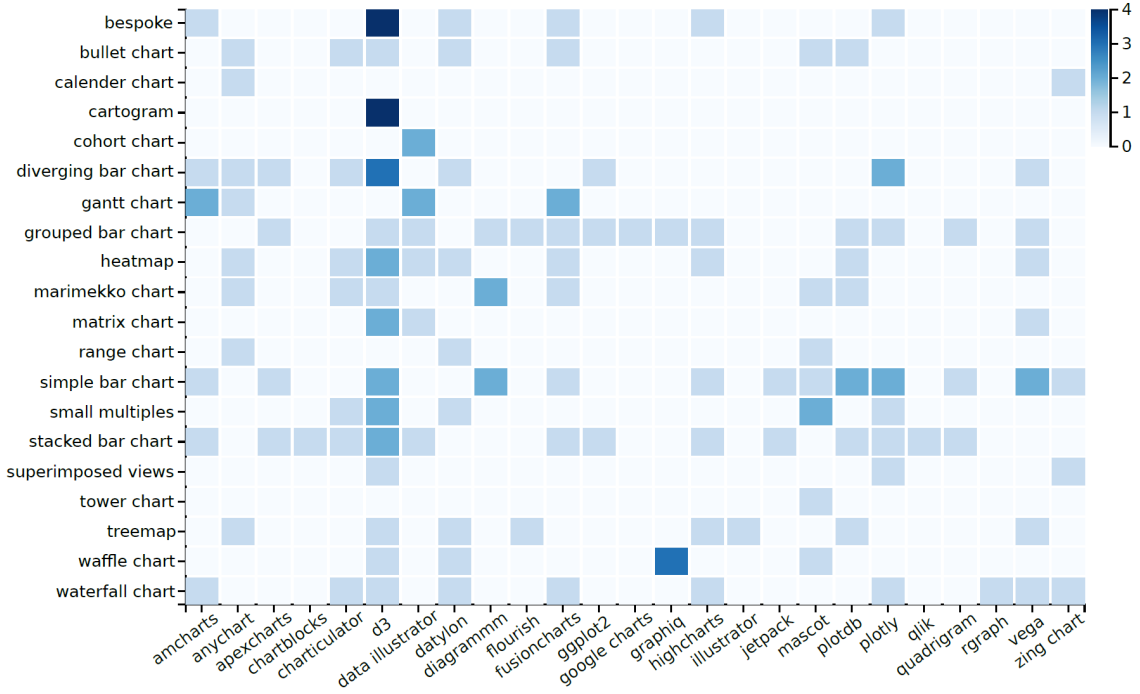
Generating Reuse Steps:

- ordering: mapping groups to data from the highest level to the lowest level => mapping rectangle marks to data => mapping visual channels to data
- use Mascot.js (previously known as Atlas.js) as the underlying library

Evaluation

Quantitative Study - Dataset

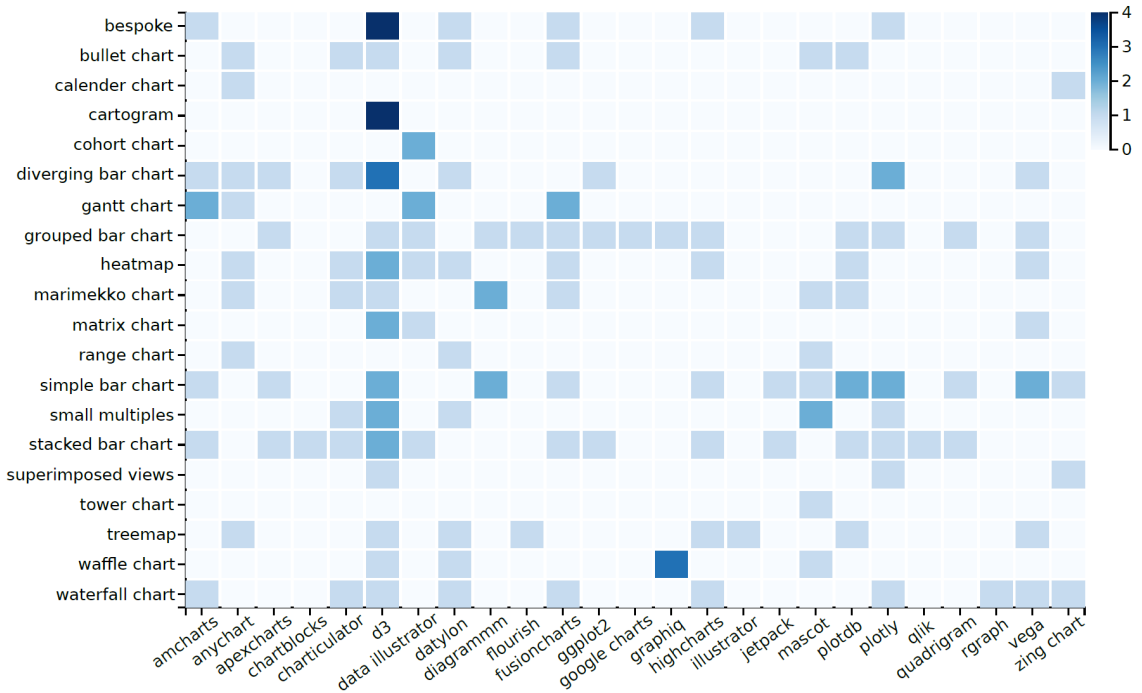
We contribute a diverse SVG chart corpus of size 150 to develop and evaluate our approach.

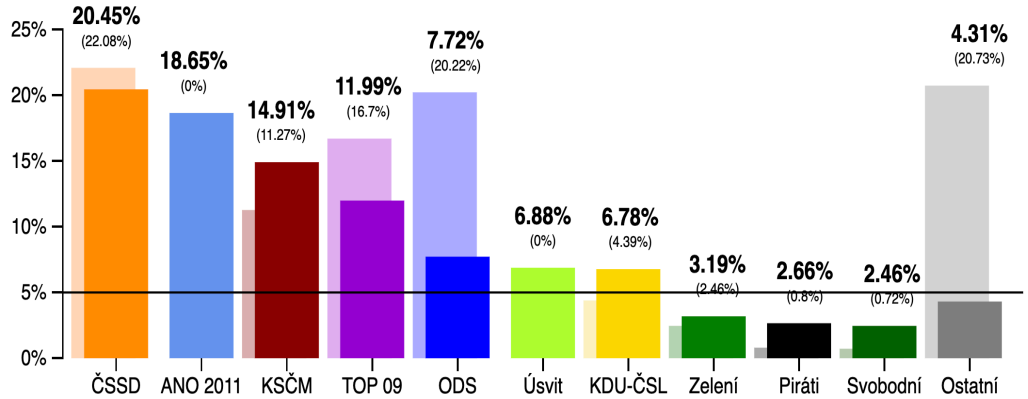


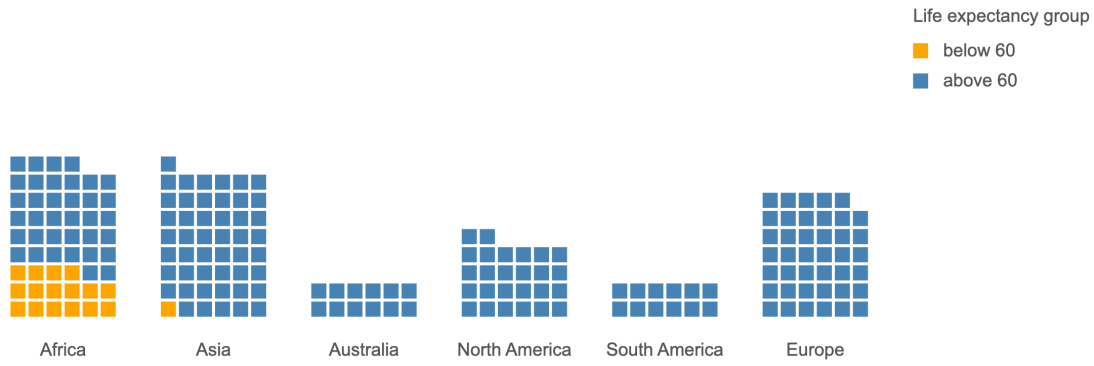
Evaluation

Quantitative Study - Dataset

We contribute a diverse SVG chart corpus of size 150 to develop and evaluate our approach.







Evaluation

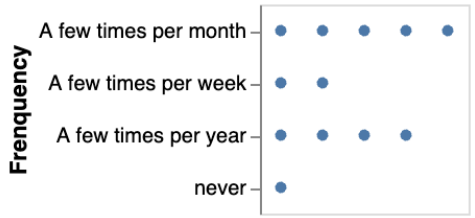
Quantitative Study - Statistical Results

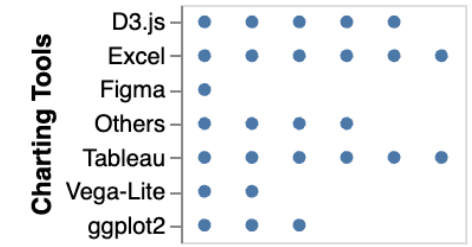
- 86.67%, 85.33%, and 90.67% accuracy on the x-axis, y-axis, and legend inference
- 96.19% (101/105) accuracy on the training set and 95.56% (43/45) accuracy on the test set, for the layout decomposition

Evaluation

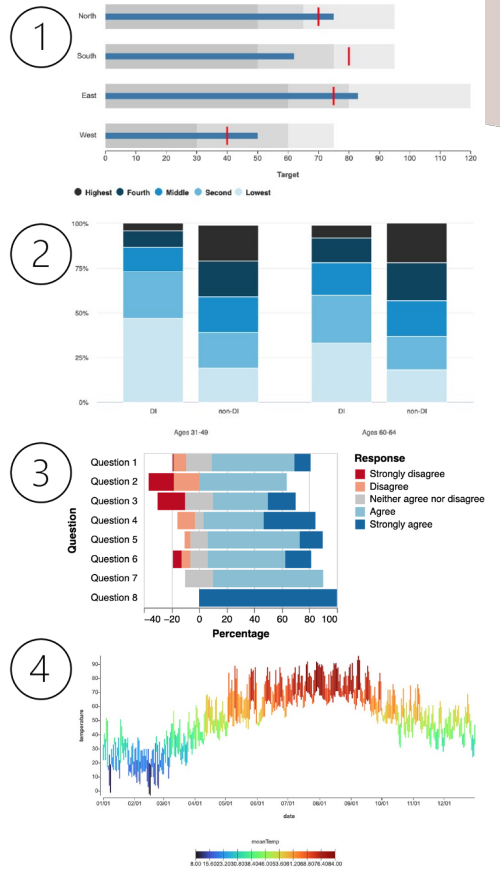
User Study - Overview

12 remote participants





Interview + Likert scale questions



Evaluation

User Study - Statistical Results

Task Completion and Time:

Task	# Successes	Average Time (minutes)	Standard Deviation
1	11	2.87	1.62
2	11	4.35	2.47
3	11	2.86	2.40
4	10	4.96	2.20

5-point Likert scale (1: "Strongly Disagree" to 5:"Strongly Agree"):

- I felt that the system was an **efficient** way to author visualization designs. **Mean**: **3.92**, **Std**: 1.04
- I was able to **conveniently** accommodate the changes that I want to achieve during my task with the system. **Mean**: **4.58**, **Std**: 0.49
- I felt **comfortable** and **confident** using the system after receiving the basic training provided to me. **Mean**: **4.25**, **Std**: 1.01

Evaluation

User Study - Feedback

Usability, Convenience, and Confidence

"Mystique gives a response after each step so I know whether I am on the right track, while in Python I cannot imagine what chart I am getting when writing codes there."

Not Clear Instructions and Insufficient flexibility

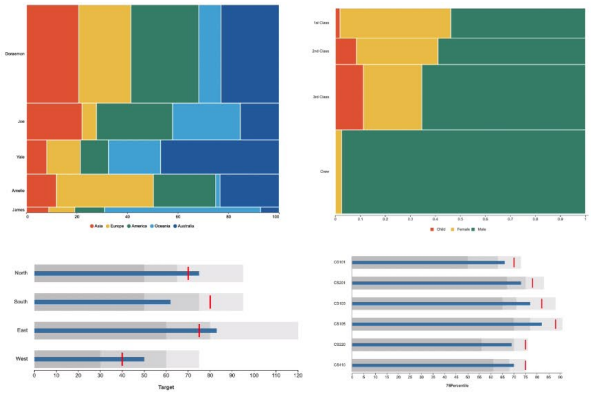
"It was hard to find what [the terms] meant exactly (top side, bottom side, height, etc.)."

"one improvement could be to allow more flexibility; for instance, currently there is no option for selecting the color used in the chart."

Research Opportunities

- Reuse **composite** visualizations involving superimposition juxtaposition, overloading, and nesting
- Identify and adapt **algorithmic layouts**
- Human-in-the-loop handling of chart **deconstruction errors**
- Generalize to larger chart corpora composed with **more shape types**

VIS 2023



Mystique

Online Demo:

https://mystique-vis.github.io/

Contact: cchen24@umd.edu