Information Visualization

Geospatial Visualization

Dr. David Koop
Critical Response to Reading

1. Describe, in your own words, what the problem addressed is and what the key contributions are

2. Respond to the paper
   - How would you add to the work that was presented?
   - What evaluation was not done that should have been?
   - No vague statements like "The paper is well-written"
   - Does the direction of the work make sense?
   - Questions are fine, but they should be specific & show your understanding
   - Keep track of points in favor, points against
   - Should focus on specific parts of the paper, make sure you understand everything about that part of the technique/system
Types of Visualization Papers

- Techniques (Algorithms)
- Applications (Design Studies)
- Systems (Toolkits)
- Evaluation (Summative User Studies)
- Model (Taxonomy, Formalism, Commentary)
- Surveys

- and Combinations of the above

[T. Munzner, 2008]
General Paper Writing Pitfalls

• What I Did Over My Summer Vacation: a diary is not a paper
  - Should not be chronological
  - Should not dwell on implementation details (which may have taken a long time)
• Least Publishable Unit: Don’t try to squeeze too many papers out of the same project
• Dense As Plutonium (Inverse of LPU): too dense, and can often miss important details of the work due to space
• Bad Slice and Dice: Dividing papers leads to too much overlap or neither paper being standalone

[T. Munzner, 2008]
Laramee’s Suggested Structure

- Introduction (Motivation)
- Related Work
- Method (Computational Model)
- Enhancements/Extensions
- Implementation
- Results & Performance
- Conclusions & Future Work
Pitfalls

- Stealth Contributions: “Do not leave your contributions implicit or unsaid”
- I Am So Unique: Don't try to sneak things past reviewers
- Enumeration Without Justification: Explain why your work is different
- Straw Man Comparison: Compare against other contemporary solutions
- But My Friends Liked It: Informal evidence is not compelling
- Unjustified Tasks: Test tasks that users actually do
- Story-Free Captions: Use captions to tell the story of your work
- My Picture Speaks For Itself: Write your take-away points in the captions

[T. Munzner, 2008]
Project Proposal

• Due Today
• Turn in via Blackboard
• Write up your ideas as they currently stand
• Things can change, that's ok!
• Focus on motivation (why should we care?) and the core idea (how does your work improve on existing techniques?)
## Paper Presentation Schedule

- Any concerns with this schedule?

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Primary</th>
<th>Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021-09-30</td>
<td>Temporal Data</td>
<td>Venkata Devesh Reddy Seethi</td>
<td>Mohammed Murtuza Shahzad Syed</td>
</tr>
<tr>
<td>2021-10-12</td>
<td>Uncertainty Visualization</td>
<td>Colin Brown</td>
<td>Abdul Rahman Shaikh</td>
</tr>
<tr>
<td>2021-10-26</td>
<td>High-Dimensional Data &amp; Dimensionality Reduction</td>
<td>Md Ashiqur Rahman</td>
<td>Colin Brown</td>
</tr>
<tr>
<td>2021-11-04</td>
<td>Machine Learning &amp; Explainable AI</td>
<td>Mohammed Murtuza Shahzad Syed</td>
<td>Md Ashiqur Rahman</td>
</tr>
<tr>
<td>2021-11-16</td>
<td>Multiple Views, Layouts, and Interaction</td>
<td>Abdul Rahman Shaikh</td>
<td>Venkata Devesh Reddy Seethi</td>
</tr>
</tbody>
</table>
Topic Format

- Three class sessions:
  1. Introduction: background lecture related to topic
  2. Paper Presentations:
     - Primary presents the paper, generally in a positive light
     - Secondary critiques the paper (what could have been improved, etc.)
  3. Discussion: discuss topics related to the paper, ideas, etc.
- Everyone reads the paper(s), comes ready with questions for presenters
A Declarative Rendering Model for Multiclass Density Maps

Jaemin Jo, Frédéric Vernier, Pierre Dragicevic, and Jean-Daniel Fekete, Senior Member, IEEE

Fig. 1: Design alternatives for a four-class density map. 1 shows small multiples where each density map is individually presented with a unique color; 2 stacks the density maps and blends the color at each pixel; 3 shows the color of the pixel with the highest density; 4–6 use regular and irregular weaving patterns; 7 shows a contour plot for each class; and 8–10 use rebinning (binning and aggregation over the density maps) with tiles produced by a random Voronoi tessellation. The aggregated values are rendered in 7 with a flat color showing the highest density, 8–9 with hatching, 10 with proportional bars, 11 with regular weaving, 12 with a dot density plot, 13 with bar-chart glyphs, and 14 with circle sizes.

Abstract—Multiclass maps are scatterplots, multidimensional projections, or thematic geographic maps where data points have a categorical attribute in addition to two quantitative attributes. This categorical attribute is often rendered using shape or color, which does not scale when overplotting occurs. When the number of data points increases, multiclass maps must resort to data aggregation to remain readable. We present multiclass density maps: multiple 2D histograms computed for each of the category values. Multiclass density maps are meant as a building block to improve the expressiveness and scalability of multiclass map visualization. In this article, we first present a short survey of aggregated multiclass maps, mainly from cartography. We then introduce a declarative model—a simple yet expressive JSON grammar associated with visual semantics—that specifies a wide design space of visualizations for multiclass density maps. Our declarative model is expressive and can be efficiently implemented in visualization front-ends such as modern web browsers. Furthermore, it can be reconfigured dynamically to support data exploration tasks without recomputing the raw data. Finally, we demonstrate how our model can be used to reproduce examples from the past and support exploring data at scale.

Index Terms—Scalability, multiclass scatterplots, density maps, aggregation, declarative specification, visualization grammar
Application Papers

- Visualizations as they are applied to application-specific data
- Less focus on the originality of the algorithm, more focus on domain-specific challenges and decisions
- Related work should contain domain-specific papers
- Additional Background section to provide readers outside of the CS domain with the necessary background
- Method may be more focused on the decisions and process and compare different approaches
- Results often revolve around expert study (evaluation from experts that have used the proposed visualizations)

[R. S. Laramee, 2009]
Writing Goals

• Write for the **audience**.
  - “Don’t just write what you want to say, write what the audience needs to hear.”
• Get your audience to **nod**: if the reviewer doesn’t agree or wonders if you’re wrong, they are less likely to like your paper
  - Avoid weasel-y words: “Some researchers think…” Who?
• Make your writing **predictable**
  - Readers are lazy
  - “You are not writing a mystery novel”
  - Upside-down pyramid writing: Important things at **start** of paper, section, paragraph
Abstract Template

- sentence 1: background
- sentence 2: missing gap (e.g. "however, there's a problem blah")
- sentence 3: why is this bad
- sentence 4: "In this paper, we propose SystemX"
- sentence 5, 6 (and/or 7): technical depth on what SystemX is. That is, what is the secret sauce that makes SystemX possible?
- sentence 8: benefits of SystemX, aka the primary contribution of the paper (should echo sentence 3 in that SystemX should have addressed the problem)
- sentence 9: "We evaluate SystemX..."
- sentence 10: "Our results indicate that (problem stated in sentence 3 is no longer a problem)..."
Exercise: Write a Pie Chart Paper

- Pretend you just discovered pie charts and want to write an abstract for the VIS conference

- Abstract Template:
  - Background
  - Motivation including problem
  - Summarize technique
  - Contributions
  - Evaluation and Results
Exercise: Pie Chart Q&A

• What kind of paper is this?
  - Technique

• Why pie charts?
  - Similar to bar charts in that they show magnitudes per object
  - Better in that they show “parts to whole” relations
  - Aesthetically pleasing and easier to read for some users

• How do you evaluate pie charts?
  - Study 1: A-B study of pie vs. bar. Quantitative measure that pie is better at “parts to whole”
  - Study 2: Qualitative study to show that pie charts are aesthetically pleasing and easy to understand
Survey Paper

• "[A]ssist the reader in the hunt for previously published research papers on a given topic"
• Full-length surveys can be 20-30 pages
• Contributions:
  - A novel classification of the literature (how your classification differs from previous surveys, or whether the survey is the first of its kind in the field).
  - A compilation of future challenges or trends in the domain.
  - The identification of both mature and less explored research directions in the field.
Survey Paper Challenges

- Managing the amount of previously published literature
- Identifying a starting point
- Deciding on a topic
- Performing a search
- Interpreting individual research papers
- Deriving a classification of literature on the given topic
- Determining related unsolved problems and future challenges
Finding Sources

- A search can yield thousands of papers
- Use known "good" papers to locate more sources
- "Cited by..." in Google Scholar

<table>
<thead>
<tr>
<th>Literature Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Scholar [Goo16]</td>
</tr>
<tr>
<td>IEEE Xplore Digital Library [IEE16]</td>
</tr>
<tr>
<td>ACM Digital Library</td>
</tr>
<tr>
<td>Vispubdata [IHK* 17]</td>
</tr>
<tr>
<td>The Annual EuroVis Conference</td>
</tr>
<tr>
<td>IEEE TVCG Journal</td>
</tr>
<tr>
<td>IEEE Pacific Visualization Symposium</td>
</tr>
<tr>
<td>IEEE VAST Conference</td>
</tr>
<tr>
<td>The Annual Eurographics Conference</td>
</tr>
<tr>
<td>The Eurographics Digital Library</td>
</tr>
<tr>
<td>Journal of Visual Languages &amp; Computing</td>
</tr>
<tr>
<td>Information Visualization Journal</td>
</tr>
<tr>
<td>Computer Graphics Forum</td>
</tr>
<tr>
<td>Computer &amp; Graphics</td>
</tr>
<tr>
<td>ACM Computing Surveys</td>
</tr>
</tbody>
</table>

[L. McNabb & R. S. Laramee]
Other Decisions

• Classification of Topics/Papers
  - Often multiple dimensions, tables used
  - One is often topics, others can be data dimensionality

• Scope of the Survey
  - Beware of being too broad or too narrow
  - Aim to create a scope of 40-50 papers

• Organization of the Survey
  - Classification helps
  - Break up smaller pieces into similar paragraphs
Keim's InfoVis Classification

Data to be Visualized vs. Interaction and Distortion Technique

1. one-dimensional
2. two-dimensional
3. multi-dimensional
4. text/web
5. hierarchies/graphs
6. algorithm/software

Visualization Technique

- Stacked Display
- Geometrically-transformed Display
- Iconic Display
- Dense Pixel Display

Interaction and Distortion Technique

- Standard
- Projection
- Filtering
- Zoom
- Distortion
- Link&Brush
Classification

- Try to find 2D classification
- Dimensions that are well-known are useful
  - Example: Shneiderman's task by data-type taxonomy (overview, zoom, filter, details-on-demand, relate, history, extract)
- Can structure the classification using unique mapping or 1-N mapping
- Tables are very helpful

<table>
<thead>
<tr>
<th>D2</th>
<th>D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>L1, L2</td>
</tr>
<tr>
<td>L2</td>
<td>L2</td>
</tr>
<tr>
<td>L1</td>
<td>L1, L2</td>
</tr>
</tbody>
</table>

(A)

<table>
<thead>
<tr>
<th>D1</th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>L2</td>
<td>✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Ln</td>
<td>✓</td>
</tr>
</tbody>
</table>

(B)

(C) D1

| L1,L4,L5,L6,L7 |
| L3,L6,L7,L8   |
| L2,L6,L7     |

<table>
<thead>
<tr>
<th>D1</th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>L2</td>
<td>✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Ln</td>
<td>✓</td>
</tr>
</tbody>
</table>

(A)

<table>
<thead>
<tr>
<th>D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
</tr>
</tbody>
</table>

(B)

<table>
<thead>
<tr>
<th>D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1,L4,L5</td>
</tr>
<tr>
<td>L3,L6,L7</td>
</tr>
<tr>
<td>L2</td>
</tr>
</tbody>
</table>

(C) [L. McNabb & R. S. Laramee]
Figures Help in Surveys, too

<table>
<thead>
<tr>
<th>Operations</th>
<th>Time</th>
<th>Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point Extraction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planar Drilling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthogonal Drilling</td>
<td>Time Drilling</td>
<td>Space Drilling</td>
</tr>
<tr>
<td>Oblique Drilling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Planar Drilling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planar Cutting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthogonal Cutting</td>
<td>Time Cutting</td>
<td>Linear Space Cutting</td>
</tr>
<tr>
<td>Oblique Cutting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Planar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curvilinear Cutting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Planar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curvilinear Space Cutting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Operations: Orthogonal Interpolation, Time Interpolation, Space Interpolation, Volume Interpolation, Translation, Time Shifting, Space Shifting, Rigid Transformation, Rotation, Yaw, Pitch, Roll, Scaling, Time Scaling, Space Scaling, Bending

[Bach et al., 2014]
Existing Figures & Sketches

![Existing Figures & Sketches](image)

[Bach et al., 2014]
Geospatial Visualization
Geographic Data

- Spatial data (have positions)
- Cartography: the science of drawing maps
  - Lots of history and well-established procedures
  - May also have non-spatial attributes associated with items
  - Thematic cartography: integrate these non-spatial attributes (e.g. population, life expectancy, etc.)
- Goals:
  - Respect cartographic principles
  - Understand data with geographic references with the visualization principles
Map Projection

[P. Foresman, Wikimedia]
Flattening the Sphere?

Central Meridian (selected by mapmaker)
Great distortion at high latitudes
Examples of two rhumb lines (direction true between any two points)
Equator touches cylinder if cylinder is tangent
Reasonably true shapes and distances within 15 degrees of Equator

[USGS Map Projections]
Lambert Conformal Conic Projection

Two standard parallels
(selected by mapmaker)

Large-scale map sheets can be joined at
edges if they have the same standard
parallels and scales
Standard Projections

- Regular Cylindrical
- Regular Conic
- Transverse Cylindrical
- Polar Azimuthal (plane)

[J. P. Snyder, USGS]
Projection Distortion

As you move the square on the map, notice how the size and shape of the *actual* area changes on the three-dimensional globe.
Projection Classification

Myriahedral projections

Dymaxion map

Goode’s homolosine

equal area

conformal

no interrupts

Lambert cyl. eq. area Plate carree Mercator

Angle-preserving

[J. van Wijk, 2008]
Subdividing regular polyhedra

For the graticular projections, thin strips of faces are attached to one single strip or face. This is a degenerated tree structure. In this section, we consider what results are obtained when a more balanced pattern is used. To this end, we start with Platonic solids for the projection of the globe, and recursively subdivide the polygons of these solids. This approach has been used before for encoding and handling geospatial data (Dutton, 1996).

At each level $i$, each edge is split and the new centres, halfway on the greater circle connecting the original endpoints, are connected. As a result, for instance each triangle is replaced at each level by four smaller triangles.

Other subdivision schemes can also be used, for instance triangles can be subdivided into nine smaller ones.

The edge weights are set as follows. We associate with each edge three numbers $w_0$, $w_1$, and $w_c$, where the first two correspond with the endpoints and the latter with the centre position. For new edges, $w_0 r_i$, $w_1 r_i$, and $w_c r_i$.

If an edge $e$ is split into two edges $e'$ and $e''$, we use linear interpolation for the new values $w_00 / w_0$, $w_01 / w_c$, $w_0c / (w_z w_1)$ = 2; $w_000 / w_c$, $w_001 / w_1$, $w_00c / (w_1 c z)$ = 2.

As a result, the weights are highest close to the centre of original edges. Finally, we use $w_c$ as the edge weight for the edges of the final mesh, plus a graticule weight $w$ with small values for $W_l$ and $W_w$ to select the aspect.

The resulting unfolded maps are, at first sight, somewhat surprising (Figure 5). One would expect to see interesting fractal shapes, however, at the second level of subdivision the gaps are already almost invisible (Figure 6). Indeed, the structure of the cuts is self-similar, however, for higher levels of subdivision and smaller triangles, the surface of the sphere quickly approaches a plane, which has Hausdorff dimension 2. Only when areas would be removed, such as the centre triangles in the Sierpinski triangle, a fractal shape would be obtained.

As a step aside, fractal surfaces and foldouts do not match well either. Unfolding, for instance, a recursively subdivided surface with displaced midpoints leads to a large number of fold-overs (Figure 7).

As another step aside, let us consider optimal mapping on Platonic solids. We consider a map optimal when the cuts do not cross continents. To find such mappings, we assign to each edge a weight proportional to the amount of land cut, computed by sampling the edges at a number of positions (here we used 25) and looking up if land or sea is covered in a texture map of the earth. Next, the map is unfolded using the standard method and the sum of weights of cut edges is determined. This procedure is repeated for a large number of orientations of the mesh, searching for a minimal value. We used a sequence of three rotations to vary the orientation of the mesh, and used steps of 1/12 per rotation. Results are shown in Figure 8.
Adding Data to Maps

• Discrete: a value is associated with a specific position
  - Size
  - Color Hue
  - Charts

• Continuous: each spatial position has a value (fields)
  - Heatmap
  - Isolines
Discrete Categorical Attribute: Shape
Discrete Categorical Attribute: Shape
Discrete Quantitative Attribute: Color Saturation
Discrete Quantitative Attribute: Size
Discrete Quantitative Attributes: Bar Chart

Railway Network Development and Bar Chart of Province Population in Turkey

[http://mis4gis.com/hgistr.org/]
Continuous Quantitative Attribute: Color Hue

[http://tampaseo.com/2012/02/websites-heat-mapping-users/]
Time as the attribute
Isolines

[USGS via Wikipedia]
Isolines

- Scalar fields:
  - value at each location
  - sampled on grids

- Isolines use *derived data* from the scalar field
  - Interpret field as representing continuous values
  - Derived data is *geometry*: new lines that represent the same attribute value

- Scalability: dozens of levels
- Other encodings?
Choropleth (Two Hues)

[M. Ericson, New York Times]
Choropleth Map

• Data: geographic geometry data & one quantitative attribute per region
• Tasks: trends, patterns, comparisons
• How: area marks from given geometry, color hue/saturation/luminance
• Scalability: thousands of regions

• Design choices:
  - Colormap
  - Region boundaries (level of summarization)
Choropleth (Two Hues)

[M. Ericson, New York Times]
Problem?

2008 Popular Vote

Obama: 68 million
McCain: 59 million

[M. Ericson, New York Times]
Problem?

2008 Popular Vote

Obama: 68 million
McCain: 59 million

Amount of red and blue shown on map

Obama: 850,000 mi²
McCain: 2,150,000 mi²

[M. Ericson, New York Times]
Adding Saturation
Area Marks and Color Hue & Saturation

Map based upon 120,464 Respondents

Map by Matthew T. Campbell
Spatial Graphics and Analysis Lab
Department of Geography and Geology
East Central University (Oklahoma)
Map Template courtesy of www.mymaps.com

Most Popular Term Used

Pop: 30% - 50%
60% - 100%

Cola: 30% - 50%
50% - 60%
80% - 100%

Soda: 30% - 50%
50% - 60%
80% - 100%

Other: 30% - 50%
60% - 80%
80% - 100%

No Data

Respondents through March 1, 2003

popvssoda.com
Aggregation: 2016 Election by Precinct

[Interactive Version, NYTimes] [R. Rohla and Washington Post, 2018]
Aggregation: 2016 Election by State

[Washington Post, 2018]
Aggregation: 2016 Election by Country

[Washington Post, 2018]
When to Use Choropleth Maps

NOT ENOUGH STOPS

TOO MANY STOPS

BRIGHT COLORS COVER TOO MANY DIFFERENT LOW VALUES

DARK COLORS COVER TOO MANY DIFFERENT HIGH VALUES

NOT IDEAL
Maps: What trends do you see?
Don't Just Create Population Maps!

PET PEEVE #208: GEOGRAPHIC PROFILE MAPS WHICH ARE BASICALLY JUST POPULATION MAPS

OUR SITE'S USERS
SUBSCRIBERS TO MARTHA STEWART LIVING
CONSUMERS OF FURRY PORNGRAPHY

THE BUSINESS IMPLICATIONS ARE CLEAR.
Size Encoding

[M. Ericson, New York Times]
Dasymetric Dot Density
Glyphs: xkcd's Map
Cartograms

**US Presidential Election 2016**
Results mapped at county level showing the candidate with the largest vote share in each area

**Overall result:**
- **Trump**
  - 62,979,536 votes (46.1%)
  - 304 electoral votes
- **Clinton**
  - 65,844,610 votes (48.2%)
  - 279 electoral votes
- **Other candidates**
  - 7,904,215 votes (5.7%)

**Vote share of candidate with most votes**
- 0%
- 50%
- 70%
- 90%

Reference map

Gridded population cartogram: areas resized according to the total number of people living there (Alaska and Hawaii not included)

Map by Benjamin Hennig
www.viewsfromtheworld.net

B. Hennig
Cartograms

1860 Presidential Election

- **Abraham Lincoln**: Republican, 180 electoral votes
- **John C. Breckinridge**: Southern Democratic, 72 electoral votes
- **John Bell**: Constitutional Union, 39 electoral votes
- **Stephen A. Douglas**: Northern Democratic, 12 electoral votes

- **Data**: geographic geometry data & **two** quantitative attributes (one part-of-whole)
- **Derived data**: new geometry derived from the part-of-whole attribute
- **Tasks**: trends, comparisons, part-of-whole
- **How**: area marks from derived geometry, color hue/saturation/luminance
- **Scalability**: thousands of regions
- **Design choices**:
  - Colormap
  - Geometric deformation

[New York Times]
Hexagonal Cartogram

District totals by category

[FiveThirtyEight, 2018]
A Declarative Rendering Model for Multiclass Density Maps

Jaemin Jo, Frédéric Vernier, Pierre Dragicevic, and Jean-Daniel Fekete, Senior Member, IEEE

Fig. 1: Design alternatives for a four-class density map. ○ shows small multiples where each density map is individually presented with a unique color; ◯ stacks the density maps and blends the color at each pixel; □ shows the color of the pixel with the highest density; ○○ use regular and irregular weaving patterns; □ shows a contour plot for each class; and ○○○ use rebinning (binning and aggregation over the density maps) with tiles produced by a random Voronoi tessellation. The aggregated values are rendered in □ with a flat color showing the highest density, ◯ with hatching, ○○ with proportional bars, □ with regular weaving, and ◯ with circle sizes.

Abstract—Multiclass maps are scatterplots, multidimensional projections, or thematic geographic maps where data points have a categorical attribute in addition to two quantitative attributes. This categorical attribute is often rendered using shape or color, which does not scale when overplotting occurs. When the number of data points increases, multiclass maps must resort to data aggregation to remain readable. We present multiclass density maps: multiple 2D histograms computed for each of the category values. Multiclass density maps are meant as a building block to improve the expressiveness and scalability of multiclass map visualization. In this article, we first present a short survey of aggregated multiclass maps, mainly from cartography. We then introduce a declarative model—a simple yet expressive JSON grammar associated with visual semantics—that specifies a wide design space of visualizations for multiclass density maps. Our declarative model is expressive and can be efficiently implemented in visualization front-ends such as modern web browsers. Furthermore, it can be reconfigured dynamically to support data exploration tasks without recomputing the raw data. Finally, we demonstrate how our model can be used to reproduce examples from the past and support exploring data at scale.

Index Terms—Scalability, multiclass scatterplots, density maps, aggregation, declarative specification, visualization grammar