Set Visualization

Una Ibrahimpasic, Eemil Hukkanen, Magne Tenstad, Azra Bajramovic

Information Visualisation, Group 2, 08 May 2024

Set-Typed Data

- Set = collection of unique objects based on specific properties.
- Each object belongs to one or more sets.
- Example: movie genres.
- Sometimes with inner hierarchy.
- Sometimes dynamic (changes over time).

Country	EEA	EU	Schengen	EFTA
Albania	0	0	0	0
Andorra	0	0	0	0
Armenia	0	0	0	0
Austria	1	1	1	0
Azerbaijan	0	0	0	0
Belarus	0	0	0	0
Belgium	1	1	1	0
Bosnia and Herzegovina	0	0	0	0
Bulgaria	1	1	1	0
Croatia	1	1	1	0
Cyprus	1	1	0	0
Czechia	1	1	1	0
Denmark	1	1	1	0
Estonia	1	1	1	0
Finland	1	1	1	0
France	1	1	1	0

49 European countries. Sets = which organizations they are part of,
https://www..nl/eu-eea-efta-schengen-countries
netherlandsworldwide

Movie Genres Dataset

Adapted from Alexander Lex

- 1000 entries
- 12 fields

Poster_Link	Series_Title	Released_Year	Certificate
Link	The Shawshank Redemption	1994	A
Link	The Godfather	1972	A

Runtime	Genre	IMDB_Rating	Overview
142 min	Drama	9.3	Two imprisoned men bond over a number of years, finding solace and eventual redemption through acts of common decency.
175 min	Crime, Drama	9.2	An organized crime dynasty's aging patriarch transfers control of his clandestine empire to his reluctant son.

Meta_score	Director	Star1	Star2
80	Frank Darabont	Tim Robbins	Morgan Freeman
100	Francis Ford Coppola	Marlon Brando	Al Pacino

Star3	Star4	No_of_Votes	Gross
Bob Gunton	William Sadler	2343110	28,341,469
James Caan	Diane Keaton	1620367	134,966,411

Set Visualization

- Graphical representation of set data.
- Shows relationships between sets and objects.
- Taxonomy of tasks:
  A. Tasks related to elements.
  B. Tasks related to sets and set relations.
  C. Tasks related to element attributes.

Tasks on Set-Typed Data

Element-related Tasks (A1 - A7)						
Find/Select elements of a specific set	Find sets containing a specific element	Find/Select elements by set memberships	Find/Select elements by their degrees	Filter out elements by set memberships	Filter out elements by their degrees	Create a set out of certain elements

Set-related Tasks (B1 - B12)					
Find the number of sets in a family	Inclusion relations / hierarchies	Exclusion / intersection relations	Identify intersections between k sets	Identify sets involved in an overlap	Identify intersections of a set
Identify the set with largest / smallest number of pair-wise set intersections	Analyze & compare cardinalities	Analyze & compare set similarities	Analyze & compare set exclusiveness	Highlight specific sets, subsets, etc.	Create a set by set-theoretic operation

Attribute-related Tasks (C1 - C5)				
Find an element's attribute values	Attribute distribution in a set / subset	Compare attribute values between subsets	Set memberships for specific attr. values	Create a set of elements by attributes

Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P. (2016), The State-of-the-Art of Set Visualization. Computer Graphics Forum, 35: 234-260. https://doi.org/10.1111/cgf.12722

Categories of Set Visualization Techniques

1. Venn and Euler Diagrams
2. Matrix
3. Node-Link
4. Overlay
5. Aggregation

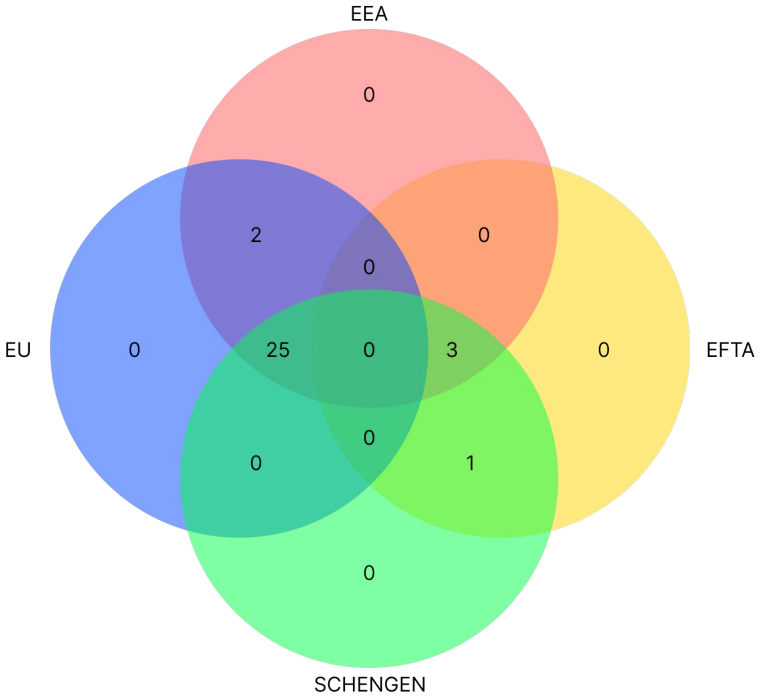
The category names have been adapted from Alsallakh et al. [2014]

Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P. (2016), The State-of-the-Art of Set Visualization. Computer Graphics Forum, 35: 234-260. https://doi.org/10.1111/cgf.12722

1. Venn and Euler Diagrams

Venn Diagrams

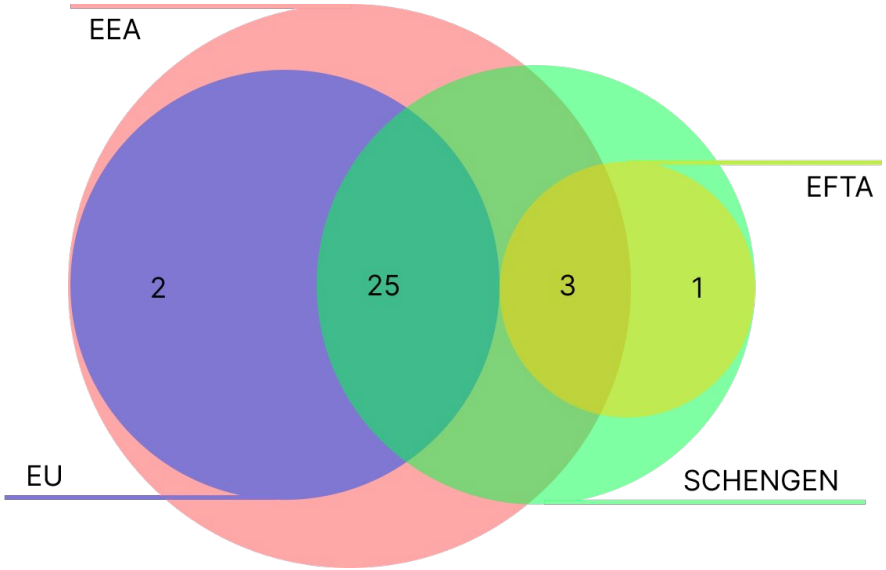
- Must show all possible curve overlaps.
- Accurate area-proportional Venn diagram can be drawn with circles for only two-set data.
- Examples:
  - InteractiVenn & nVenn



Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P. (2016), The State-of-the-Art of Set Visualization. Computer Graphics Forum, 35: 234-260. https://doi.org/10.1111/cgf.12722

Euler-based Diagrams

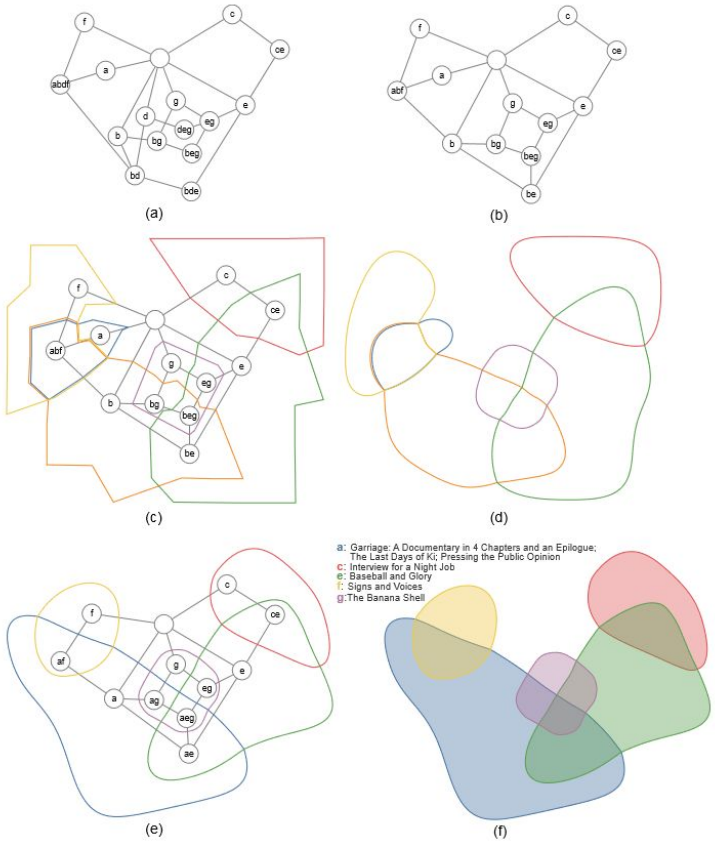
- Uses different shapes to represent sets and their relationships.
- Focusing on the depiction of the logical relationships between sets, such as intersections and unions.
- Examples:
  - Euler diagrams & Eulermerge.



Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P. (2016), The State-of-the-Art of Set Visualization. Computer Graphics Forum, 35: 234-260. https://doi.org/10.1111/cgf.12722

EulerMerge

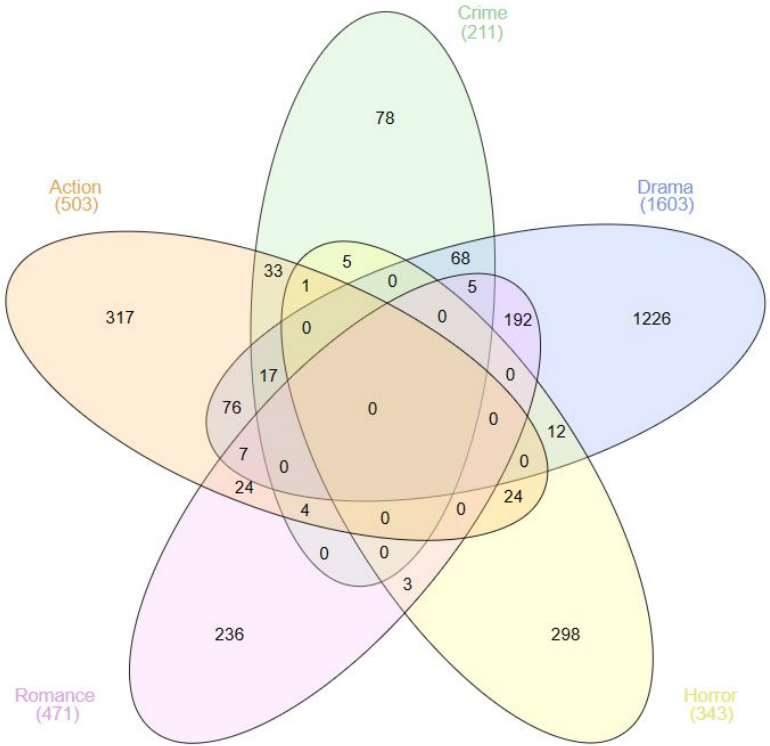
- Tool to visualize of large-scale Euler diagrams.
- It allows efficient merging of multiple Euler diagrams.



Xinyuan Yan & others, 2023, EulerMerge: Simplifying Euler Diagrams Through Set Merges, 10.48550/ARXIV.2311.15396

InteractiVenn

- Tool for interactive visualization of Venn diagrams.
- Can take data up to six sets of data.
- Allows users to merge sets.
- Showcase video: https://youtu.be/GRsvxMJlUoM



Henry Heberle & others, 2015, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams, http://www.interactivenn.net/

2. Matrix Diagrams

Matrix Diagrams

- Set intersections are defined by either:
  - a matrix row.
  - a matrix cell.
- Scalable in number of elements and sets.

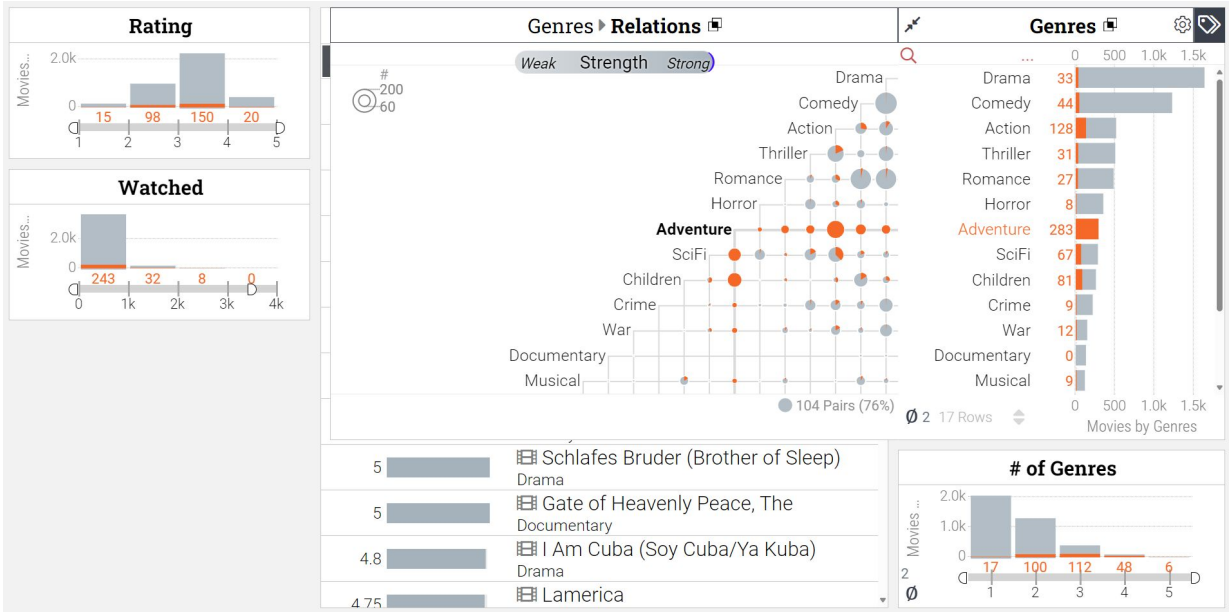
UpSet

- Row = intersection
- Includes:
  - Histograms
  - Bar charts
  - Box plots
  - Scatter plots
- Showcase video: https://youtu.be/VD7IhfqkIc0

Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph. 2014 Dec;20(12):1983-92. doi: 10.1109/TVCG.2014.2346248. PMID: 26356912; PMCID: PMC4720993.

AggreSet

- Cell = intersection.
- Customizable:
  - Histograms
  - Bar charts
  - Scatter plots

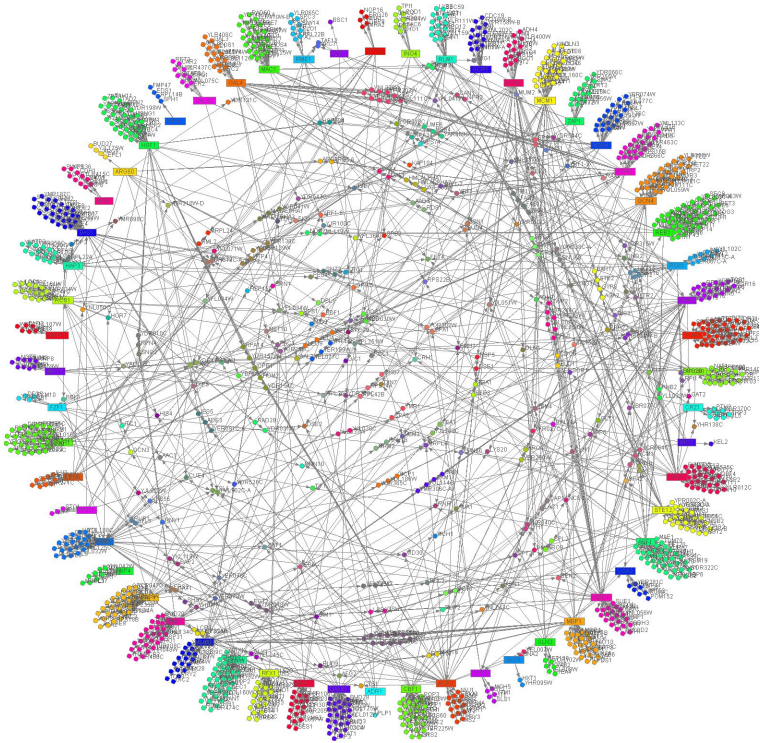


Yalçin MA, Elmqvist N, Bederson BB. AggreSet: Rich and Scalable Set Exploration using Visualizations of Element Aggregations. IEEE Trans Vis Comput Graph. 2016 Jan;22(1):688-97. doi: 10.1109/TVCG.2015.2467051. Epub 2015 Aug 13. PMID: 26390465.

3. Node-Link Diagrams

Node-Link Diagrams

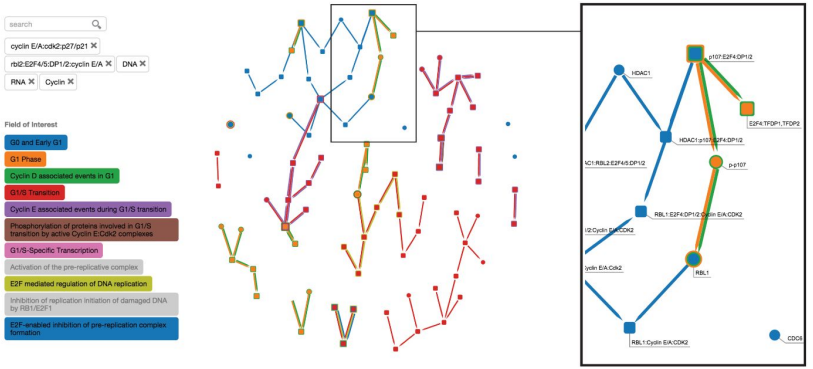
- Element-based techniques model the membership relations between elements and sets as edges of a bipartite graph whose nodes represent the elements and sets.
- Used to show the similarity between the sets as links of varying thicknesses.
- Are commonly used to facilitate reasoning about Formal Concept Analysis.



Kazuo Misue. 2006. Drawing bipartite graphs as anchored maps. Doi:10.1145/1151903.1151929

BranchingSets

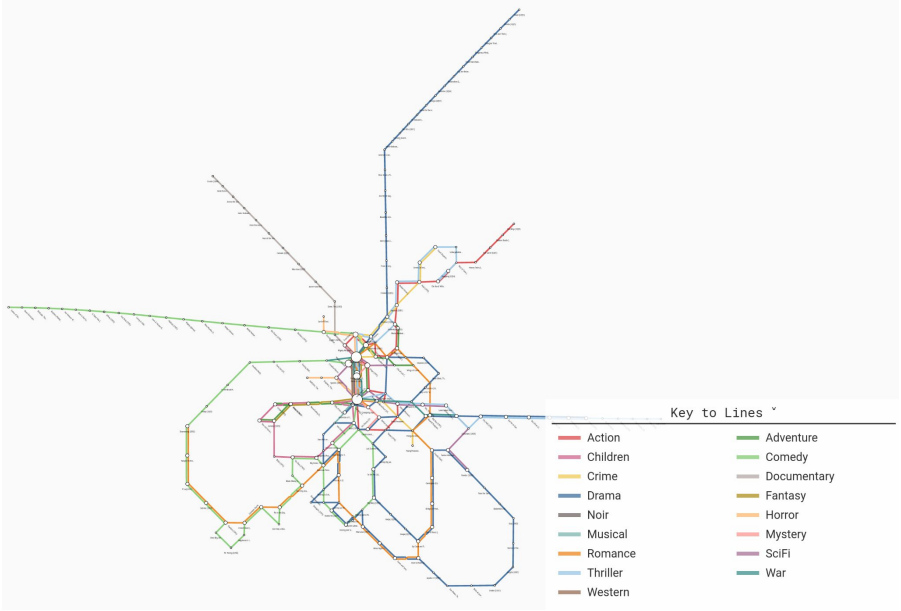
- Resolve challenges in representing large datasets of biological pathway networks.
- Interactive set visualisation technique.
- Easy pattern recognition and relationship identification.



Paduano, Francesco & Etemadpour, Ronak & Forbes, Angus. (2016). BranchingSets: Interactively Visualizing Categories on Node-Link Diagrams. Doi: 10.1145/2968220.2968229.

MetroSets

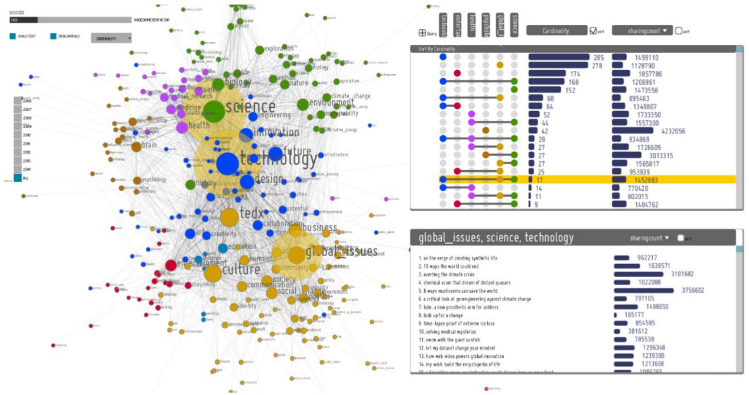
- Sets are represented by metro lines and set elements are represented by metro stations.
- Elements that belong to multiple sets corresponding to interchange stations.
- Each vertex is represented as a circle with the diameter determined by either:
  - the number of incident sets.
  - the largest number of adjacent lines of all incident edges.
- Showcase video: https://youtu.be/59sYNQQZOuY



B. Jacobsen, M. Wallinger, S. Kobourov and M. Nöllenburg, "MetroSets: Visualizing Sets as Metro Maps," 202. Doi: 10.1109/TVCG.2020.3030475

NetSet

- Addresses limitations of both matrix-based visualisations and network visualisations by merging them.
- Uses UpSet model as reference.
- Network is built using a bipartite network construction method:
  - nodes represent sets and edges indicating intersections
  - thickness of edges represents the cardinality of intersections
  - node size reflects the degree centrality



Heungseok Park, Hongjun Lim, Wonjae Lee and Kyungwon Lee, "NetSet: A systematic integration of visualization for analyzing set intersections with network,"2017.Doi: 0.1109/PACIFICVIS.2017.8031575

4. Overlay Diagrams

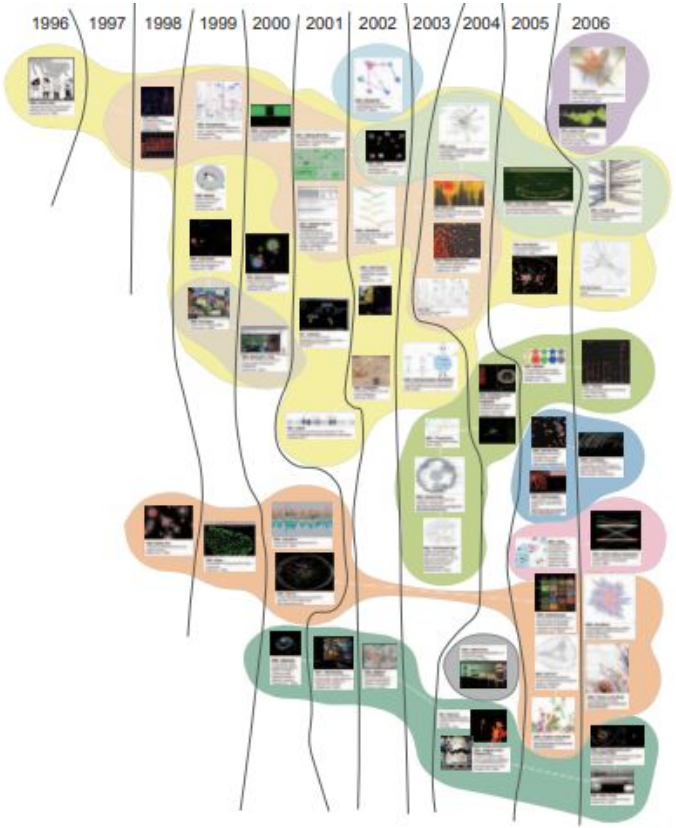
Overlay Diagrams

- Analyzing information in the data in context of other data features.
- Examples:
  - Elements with a spatial reference
  - Points in a scatter plot
  - Nodes in a graph
- Types:
  - Region-based
  - Line-based
  - Glyph-based

Alsallakh, B., Micallef, L., Aigner, W., Hauser, H., Miksch, S. and Rodgers, P. (2016), The State-of-the-Art of Set Visualization. Computer Graphics Forum, 35: 234-260. https://doi.org/10.1111/cgf.12722

F2-Bubbles

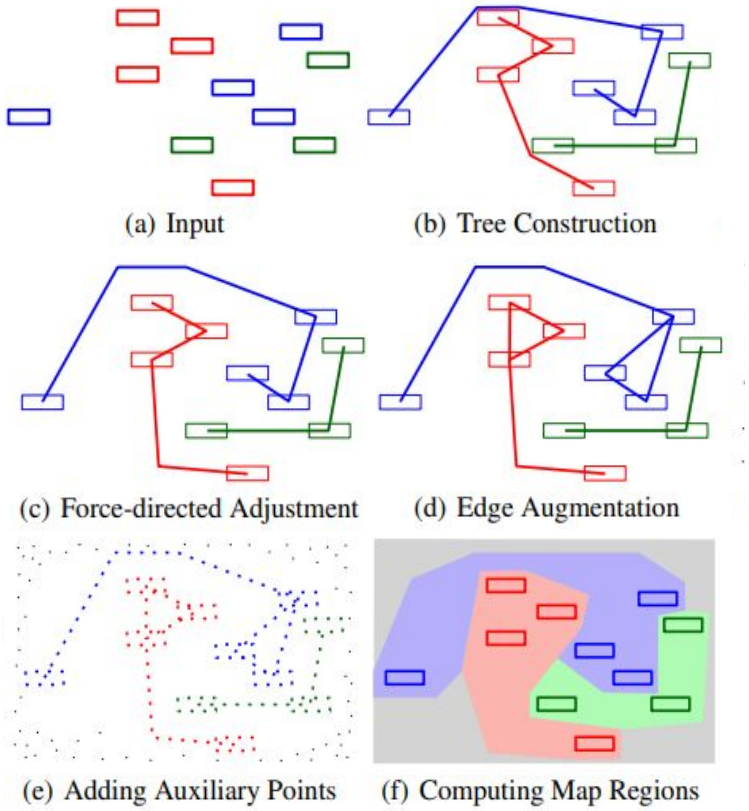
- Simultaneous construction of spanning trees.
- Relation-aware energy fields with adaptive contour widths based on nearby set elements.
- Interactions:
  - Add/delete/move nodes and edges (suggestions provided).
  - Add/delete control points to adjust edge routing.
  - Direct manipulation of smooth contour control points.



Y. Wang et al., "F2-Bubbles: Faithful Bubble Set Construction and Flexible Editing," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 1, pp. 422-432, Jan. 2022, doi: 10.1109/TVCG.2021.3114761.

MapSets

- Fixed vertex positions for embedded and clustered graphs.
- Contiguous, non-overlapping, convex regions for each cluster.
- “minimum ink” concept to optimize the convexity of the generated regions (clusters).



Efrat, A., Hu, Y., Kobourov, S.G., Pupyrev, S. (2014). MapSets: Visualizing Embedded and Clustered Graphs. In: Duncan, C., Symvonis, A. (eds) Graph Drawing. GD 2014. Lecture Notes in Computer Science, vol 8871. Doi: https://doi.org/10.1007/978-3-662-45803-7_38 ; https://gmap.cs.arizona.edu/

GridSet

A. the Main view
B. the Visual Property menu
C. the Query view
D. the Set view (orange-highlighted views represent added sets on the Main view)
E. the Detail view that provides detailed information of the elements

H. Chung, S. Nandhakumar and S. Yang, "GridSet: Visualizing Individual Elements and Attributes for Analysis of Set-Typed Data," in IEEE Transactions on Visualization and Computer Graphics, 1 Aug. 2022, doi: 10.1109/TVCG.2020.3047111

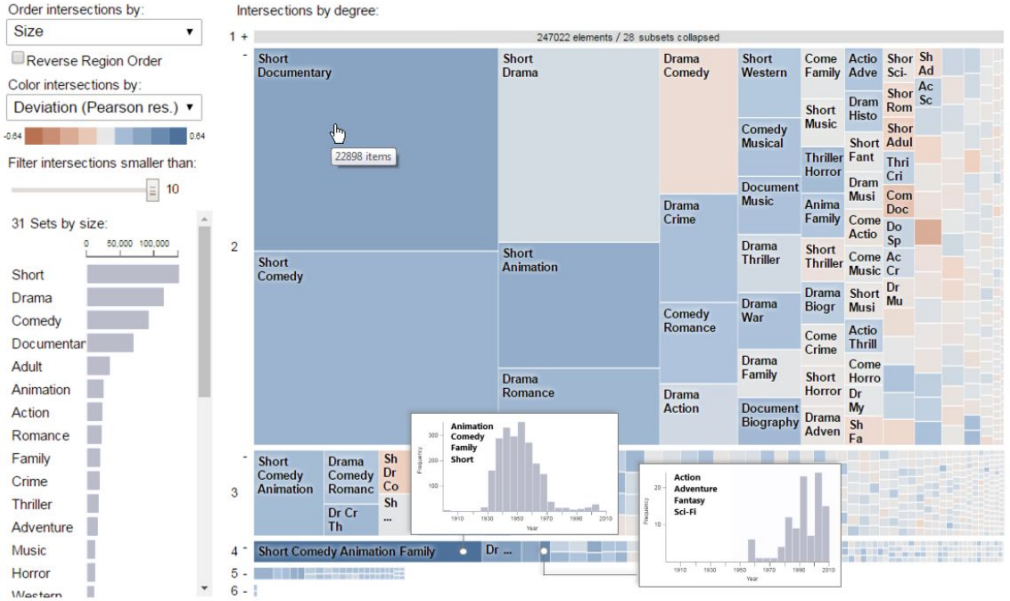
5. Aggregation Diagrams

Aggregation Diagrams

- Hides individual objects.
- Set size determined by area.
- Highly scalable in number of elements.

PowerSet

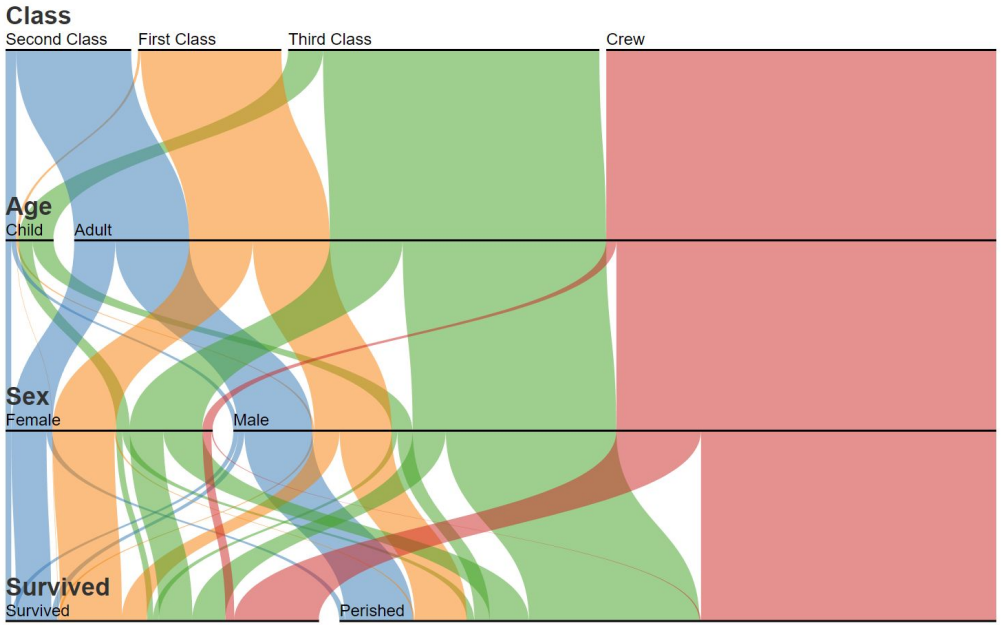
- Rectangle = intersection.
- Sorted by:
  - size: x-axis
  - intersections: y-axis
- Customizable coloring.



B. Alsallakh and L. Ren, "PowerSet: A Comprehensive Visualization of Set Intersections," in IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 1, pp. 361-370, Jan. 2017, doi: 10.1109/TVCG.2016.2598496.

Parallel Sets

- First row decides color.
- Blocks split and join based on intersections.
- Customizable order.
- Not all sets can overlap.

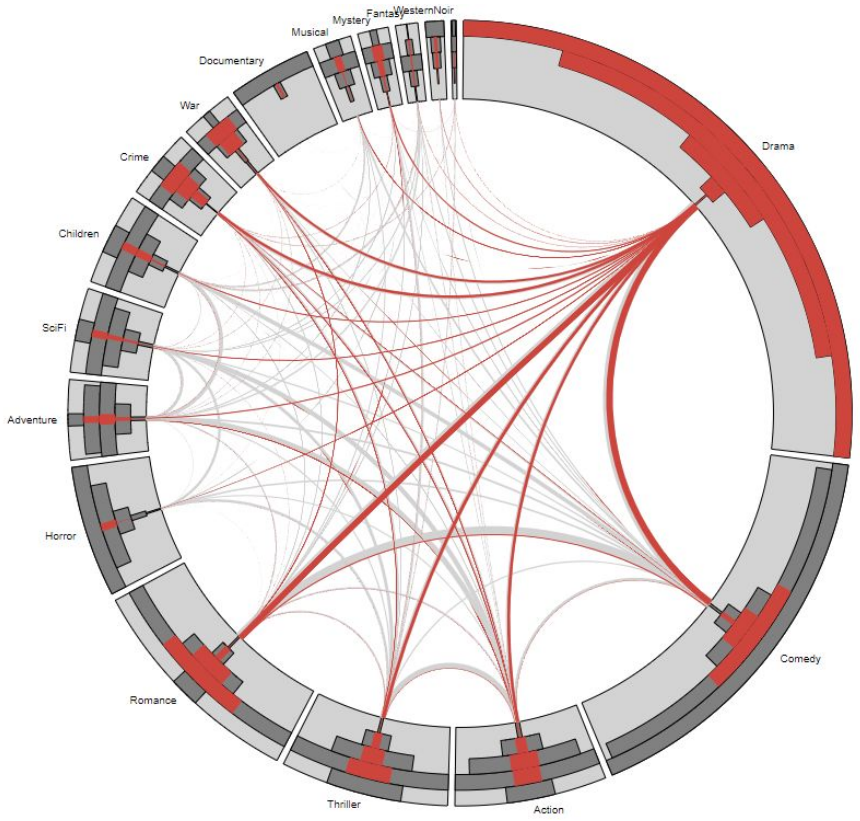


Made by Jason Davies. https://www.jasondavies.com/parallel-sets/

R. Kosara, F. Bendix and H. Hauser, "Parallel Sets: interactive exploration and visual analysis of categorical data," in IEEE Transactions on Visualization and Computer Graphics, vol. 12, no. 4, pp. 558-568, July-Aug. 2006, doi: 10.1109/TVCG.2006.76.

Radial Sets

- Sets placed in a circle.
- Inner segment
  = “anonymous intersection”.
- Lines explain chosen segment, similar to Node-Link.
- Showcase video:
  https://youtu.be/XWsI50i01wA



B. Alsallakh, W. Aigner, S. Miksch and H. Hauser, "Radial Sets: Interactive Visual Analysis of Large Overlapping Sets," in IEEE Transactions on Visualization and Computer Graphics, vol. 19, no. 12, pp. 2496-2505, Dec. 2013, doi: 10.1109/TVCG.2013.184

Comparison of Techniques

	Strengths	Weaknesses
1. Euler-based diagrams	Intuitive when well-matched (little training is required). Represent all standard set relations compactly.	Limited to few sets due to clutter and drawability issues. Desired properties not always possible (e.g. convexity).
2. Matrix-based techniques	Fairly scalable both in the number of elements and sets. Do not suffer from edge crossings or topological constraints.	Limited in the set relations they can represent. Revealed membership patterns are sensitive to ordering.
3. Node-link diagrams	Visually emphasize the elements as individual objects. Show clusters of elements having similar set memberships.	No representation of set relations in element-set diagrams.
4. Overlays	Emphasize element and set distributions according to other data features (e.g. map locations).	Often limited in the number of elements and sets. Undesired layout artifacts (overlaps, crossing, shapes, etc.). Limited scalability due to edge crossings.
5. Aggregation- based	Highly scalable in the number of elements. Some techniques can show how attributes correlate with set membership.	Usually, do not emphasize sets and elements as objects. Limited in the set relations they can represent.

SurVis - Survey Browser

https://info-vis-24.github.io/survey-browser/

Questions?