Establishing a visualization design space
A case study in infectious disease genomic epidemiology

Anamaria Crisan
PhD Candidate, Computer Science
University of British Columbia

https://doi.org/10.1101/325290
@amcrisan  acrisan@cs.ubc.ca  http://cs.ubc.ca/~acrisan
**PhD Candidate, Computer Science**  
University of British Columbia

**Thesis:** Visualizing Public Health Data

**Advisors:** Dr. Tamara Munzner  
Dr. Jennifer Gardy
What we’ll talk about
Why should we visualize data?

Thinking systematically about data visualization

GEViT: a Genomic Epidemiology Visualization Typology
Why should we visualize data?
Statistics is not always the answer

<table>
<thead>
<tr>
<th>X Mean</th>
<th>Y Mean</th>
<th>X SD</th>
<th>Y SD</th>
<th>Corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>54.2659224</td>
<td>47.8313999</td>
<td>16.7649829</td>
<td>26.9342120</td>
<td>-0.0642526</td>
</tr>
</tbody>
</table>
Statistics is not always the answer

Same stats, different graphs (Datasaurus)

Humans interpret numerical information differently

It is not always easy to reason consistently with numbers

http://bit.ly/1FxtT2z
Humans interpret numerical information differently

Numeracy: the ability to reason with numbers
- Individuals with low numeracy have difficulty interpreting numbers and probabilities
- Also true amongst educated professionals

Visualization can make data more accessible to individuals with lower numeracy skills

Need the *right* data visualization, not just a visualization

Zikmund-Fisher (2013). A demonstration of "less can be more" in risk graphics.
The state of visualization in public health

- Greater understanding that visualizing data is important
- Barriers for creating data visualizations are lowering
  - Many domain specialists (scientists, public servants) routinely create data visualizations
- Guidance on what makes a good data visualization is absent
  - Making the right data visualization for some specific context is harder problem
  - Few people outside of infovis read the infovis literature
- Lack of guidance = ineffective ad hoc solutions
The state of visualization in public health

- Greater understanding that visualizing data is important
- Barriers for creating data visualizations are lowering
  - Many domain specialists (scientists, public servants) routinely create data visualizations
- Guidance on what makes a good data visualization is absent
  - Making the right data visualization for some specific context is harder problem
  - Few people outside of infovis read the infovis literature
- Lack of guidance = ineffective *ad hoc* solutions
- Our proposed solution: systematically create an explorable vis design space
  - Shows what is possible with visual design
  - Can help make the search for good visualizations easier
Thinking systematically about visualization design
Design Spaces: A quick primer

Design spaces are made of visualization design choices or varying utility (+ 0 - )

We have some intuition on design choices

- All images below show chairs, but they have different designs
- All chairs can be used for a common task: sitting
- But – fundamentally, different chairs are suited for different contexts

Not suitable as an office chair (-)  
Suitable as an office chair (+,0)
Design Spaces: A quick primer

GOAL – nudge domain specialists toward better design choice solutions

(i.e. shop around and find the best chair)

(i.e. choosing the nearest & cheapest chair)

BUT – how do we **systematically** describe design space to promote good exploration?

Considering a design space for public health

• There is considerable variability in public health visualization design
  • Example: all visualizations below show a hospital transmission

Creating an explorable visualization design space

https://doi.org/10.1101/325290
An overview of our approach

- Can we capture the extent of variability in visualization design?
An overview of our approach

- Can we capture the extent of variability in visualization design?
- Why should we do this?
  - Initially, we were simply curious about how others visualize data
  - Then we realized how useful and powerful having the design space was
An overview of our approach

- Can we capture the extent of variability in visualization design?

- Why should we do this?
  - Initially, we were simply curious about how others visualize data
  - Then we realized how useful and powerful having the design space was

- How do we do this?
  - No methods in vis literature for systematic construction of a design space
  - Image classification missed important contextual data (no useful training data)
An overview of our approach

- Can we capture the extent of variability in visualization design?
- **Why should we do this?**
  - Initially, we were simply curious about how others visualize data
  - Then we realized how useful and powerful having the design space was
- **How do we do this?**
  - No methods in vis literature for systematic construction of a design space
  - Image classification missed important contextual data (no useful training data)
- **What we did**: developed a method for systematically constructing design spaces!
  - Influenced by literature reviews in medical literature
  - Added machine learning with qualitative methods
  - Applied this method to infectious disease genomic epidemiology (IDGE)
Our approach allows us to answer three different questions

1. **Literature Analysis**
   - WHY are researchers visualizing data?

2. **Qualitative Data Visualization Analysis**
   - HOW are researchers visualizing data, WHAT are they visualizing?

3. **Quantitative Data Visualization Analysis**
   - HOW MANY examples are there of specific visualizations?
## An overview of our approach

| Our Objective | Across the many topics of microbial gen epi research articles identify and enumerate the different kinds of visualizations that are used |
An overview of our approach

Literature Analysis Steps

Text mining of document corpus to identify topics

Systematically sample papers with topics as strata

Our Objective

Across the many topics of microbial gen epi research articles, identify and enumerate the different kinds of visualizations that are used
An overview of our approach

**Literature Analysis Steps**
- **Text mining of document corpus to identify topics**
- **Systematically sample papers with topics as strata**

**Our Objective**
Across the many *topics* of microbial gen epi research *articles* identify and enumerate the different *kinds of visualizations* that are used.

**Visualization Analysis Steps**
- Derived a code set to classify research figures (GEViT)
- Applied GEViT to collection of research figures
<table>
<thead>
<tr>
<th>Literature Analysis Steps</th>
<th>Our Objective</th>
<th>Visualization Analysis Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text mining of document corpus to identify topics</td>
<td>Across the many <strong>topics</strong> of microbial gen epi research <strong>articles</strong> identify and <strong>enumerate</strong> the different <strong>kinds of visualizations</strong> that are used</td>
<td>Derived a code set to classify research figures (GEViT)</td>
</tr>
<tr>
<td>Systematically sample papers with topics as strata</td>
<td></td>
<td>Applied GEViT to collection of research figures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Applied descriptive statistics to derived code sets</td>
</tr>
</tbody>
</table>
Three major results from this work

- **Literature Analysis (Why):**
  - Understanding the structure of genomic epidemiology papers
  - Motivated intelligent sampling of data visualizations
  - Primary sources of our design space visualizations
Three major results from this work

- **Literature Analysis (Why):**
  - Understanding the structure of genomic epidemiology papers
  - Motivated intelligent sampling of data visualizations
  - Primary sources of our design space visualizations

- **Qualitative Analysis (What, How):**
  - Manually analyzed figures to classify elements of data visualizations
  - Generated a Genomic Epidemiology Visualization Typology (GEViT)
Three major results from this work

- **Literature Analysis (Why):**
  - Understanding the structure of genomic epidemiology papers
  - Motivated intelligent sampling of data visualizations
  - Primary sources of our design space visualizations

- **Qualitative Analysis (What, How):**
  - Manually analyzed figures to classify elements of data visualizations
  - Generated a *Genomic Epidemiology Visualization Typology* (GEViT)

- **Quantitative Analysis (How many):**
  - In IDGE it’s nearly all trees and a surprising amount of tables
Developing and Operationalizing GEViT
How can we systematically describe images?

- What does GEViT do and not do?

**GEViT provides a base**
- Deliverables:
  1. Typology
  2. Interactive Gallery

**GEViT does not evaluate**
- Massive undertaking that would take many years
- Needs GEViT to conduct evaluations

- How can GEViT be used?
  - Concise descriptions to discuss data visualizations
  - Understand what visualizations are common and possible
  - Get ideas for data visualization design
[WHY] Literature Analysis

Setting up the visualization design space
### Overview of exploration and sampling of the document corpus

<table>
<thead>
<tr>
<th>17,974</th>
<th>All documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>15,315</td>
<td>Removal of “very noisy” articles</td>
</tr>
</tbody>
</table>

Article acquisition & unsupervised clustering
Literature analysis: discovering topics in the document corpus

Too few articles for a reliable cluster

A cluster of articles

Unclustered Articles

N = 15,315 (very noisy articles removed)
Literature analysis: discovering topics in the document corpus

Topic clustering results

Clusters for sampling
Literature analysis: verifying the results
Overview of exploration and sampling of the document corpus

17,974 All documents
15,315 Removal of “very noisy” articles
9,551 Articles about human pathogen
6,350 Common human pathogen articles

Article acquisition & unsupervised clustering

Limit to common human pathogens
Apply *a priori* concepts
Linking pathogen clusters to a priori concepts

- Wanted to include addition public health concepts
- Topics assessed *a priori*, assigned to common terms between clusters

### Molecular Biology Concepts
- Characterization
- Diversity
- Drug Resistance
- Genome
- Genotype
- mBio
- Phylogeny
- Reservoir
- Vector

### Epidemiology Concepts
- Cluster
- Geography
- Outbreak (International / Community/ Hospital)
- Surveillance
- Transmission
- Vaccine

### Medical Concepts
- Clinical
- Cancer
- Diagnosis
- Outcome
- Treatment

23 *a priori* concepts in total
Overview of exploration and sampling of the document corpus

Article acquisition & unsupervised clustering

Limit to common human pathogens
Apply a priori concepts

Sampling of articles (two rounds)

- Random stratified sampling
  - Strata: pathogen, *a priori* concepts

Yielded 801 figures and 49 tables
Overview of exploration and sampling of the document corpus

17,974 All documents
15,315 Removal of “very noisy” articles
9,551 Articles about human pathogen
6,350 Common human pathogen articles

6074 Sampling Round #1
276 Accept (Yes) or Reject (No)
179 No
107 Yes
293 Sampling Round #2
186 Accept (Yes) or Reject (No)
107 No
204 + 17 manually added
221 Basis for further analysis

Article acquisition & unsupervised clustering

Limit to common human pathogens
Apply a priori concepts

Sampling of articles (two rounds)

- Random stratified sampling
  - Strata: pathogen, a priori concepts

Yielded 801 figures and 49 tables
Qualitative Analysis
Manually analyzed paper figures to classify elements of data visualizations, derived GEViT
Input: 801 figures, 49 tables

- Used qualitative coding techniques to analyze research figures
  - Multiple rounds of classifying and codifying elements of figures
  - Used figures from sample papers to derive codes

- Figures in the same paper were analyzed separately
  - Multi-part figures were analyzed together

Result: GEViT, a hierarchical code set with separate taxonomies for:
  - Chart Types
  - Chart Combinations
  - Chart Enhancements
Chart Type: a foundational element of all data visualizations

- Self explanatory what chart types are…
- Manually classified every single type of chart in every figure
- Report only what we found in the sample document corpus
- There were six classes of charts types
  - Charts also had special chart types (i.e. epidemic curve is a special case of bar chart)
Chart Type: a foundational element of all data visualizations

Common Statistical Charts

**Chart Types**
- **Bar Chart**: Standard, Stacked, Divergent
- **Line Chart**: Special Cases - Epidemiic Curve, Diversity Chart, LefSe Plot
- **Scatter Plot**: Special Cases - Root-to-tip, Ordination Plot, Q-Q plot

**Other Charts**
- **Pie Chart**
- **Venn Diagram**
- **Timeline**
- **Distribution Plot**: Histogram, PDF, Boxplot, Swarm Plot
Chart Type: a foundational element of all data visualizations
Chart Type: a foundational element of all data visualizations

### Tree Chart Types
- **Phylogenetic Tree**
  - Rooted (Radial & Linear)
  - [Image of Phylogenetic Tree]
- **Dendrogram**
  - [Image of Dendrogram]
- **Unrooted (Radial & Linear)**
  - [Image of Unrooted Tree]
- **Clonal Tree**
  - [Image of Clonal Tree]

### Genomic Chart Types
- **Genomic Structure**
  - Linear
  - Radial
  - [Image of Genomic Structure]
- **Alignment**
  - [Image of Alignment]
- **Variation Profile**
  - [Image of Variation Profile]
- **Sequence Logo Plot**
  - [Image of Sequence Logo Plot]
Chart Type: a foundational element of all data visualizations
Chart Combinations: showing different aspects of the data

- Observed that charts were combined in a specific, consistent pattern
- We classified every single of chart combinations within a figure

Example: Same chart type, different metadata

Example: Two chart types together
### Chart Combinations: showing different aspects of the data

<table>
<thead>
<tr>
<th>Combination Type</th>
<th># of chart types</th>
<th># of charts</th>
<th>Linkage type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td><img src="image1" alt="Example" /></td>
</tr>
<tr>
<td>Composite</td>
<td>Many</td>
<td>1</td>
<td>Spatially Aligned</td>
<td><img src="image2" alt="Example" /></td>
</tr>
<tr>
<td>Small Multiples</td>
<td>1</td>
<td>Many</td>
<td>Chart Type &amp; Data</td>
<td><img src="image3" alt="Example" /></td>
</tr>
<tr>
<td>Many Type Linked</td>
<td>Many</td>
<td>Many</td>
<td>Visual, not spatial</td>
<td><img src="image4" alt="Example" /></td>
</tr>
<tr>
<td>Many Type General</td>
<td>Many</td>
<td>Many</td>
<td>NA</td>
<td><img src="image5" alt="Example" /></td>
</tr>
<tr>
<td>Complex Combination</td>
<td>Many</td>
<td>Many</td>
<td>Context dependent</td>
<td><img src="image6" alt="Example" /></td>
</tr>
</tbody>
</table>
Enhancements: overlaying additional metadata

- **Mark** = basic graphical element (line, point, area)
- **Enhancement** = adding marks or re-encoding marks of the base chart type

---

**Add Marks**
Adding Additional Marks to base chart type
- Point
- Line
- Area Mark
- Text
- Glyph

**Re-encode Marks**
Re-encode existing marks via channels
- Size
- Shape
- Color
- Texture
- Font
### Enhancements: overlaying additional metadata

<table>
<thead>
<tr>
<th></th>
<th>Size</th>
<th>Shape</th>
<th>Color</th>
<th>Texture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Point</strong></td>
<td><img src="image" alt="Point" /></td>
<td><img src="image" alt="Shape" /></td>
<td><img src="image" alt="Color" /></td>
<td><img src="image" alt="Texture" /></td>
</tr>
<tr>
<td><strong>Line</strong></td>
<td><img src="image" alt="Line" /></td>
<td><img src="image" alt="Line" /></td>
<td><img src="image" alt="Line" /></td>
<td><img src="image" alt="Line" /></td>
</tr>
<tr>
<td><strong>Area</strong></td>
<td><img src="image" alt="Area" /></td>
<td><img src="image" alt="Area" /></td>
<td><img src="image" alt="Area" /></td>
<td><img src="image" alt="Area" /></td>
</tr>
<tr>
<td><strong>Text</strong></td>
<td><img src="image" alt="Text" /></td>
<td><img src="image" alt="Text" /></td>
<td><img src="image" alt="Text" /></td>
<td><img src="image" alt="Text" /></td>
</tr>
</tbody>
</table>

**Special Cases:** Containment Mark; Connection Mark; Glyph
Enhancements: overlaying additional metadata

- Structured Enhancement: Encodings are added/changed on many/all marks
- Unstructured Enhancement: Encoding are added/changed to one or a few marks

**Structured enhancement**

- **Add Marks**
  - Adding Additional Marks to base chart type
    - Point
    - Line
    - Area Mark
    - Text
    - Glyph

- **Re-encode existing marks**
  - Re-encode existing marks via channels
    - Size
    - Shape
    - Color
    - Texture
    - Font Face (specific to text)

**Unstructured enhancement**

- **Add Annotation**
  - Manually Adding annotations
    - Same as added marks, but include arbitrary ink too

*Note: Sometimes the line between adding a mark and adding an annotation is very subtle.*
Enhancements: overlaying additional metadata

Base Chart
- Tree

Structured Enhancements
- Re-encode Marks
  - Line - color
- Add Marks
  - Point, line, & text

Unstructured Enhancements
- Add Annotation
  - Arrow, text

Node-link
- Point - size

Glyph - Pie Chart
- Group 1
- Group 2

Containment Mark
- Group 1
- Group 2

No known contacts
GEViT in action

Visualization Breakdown

Visualization Context (why)
- Pathogen: Pseudomonas aeruginosa
- Concepts: outbreak; control; genome; phylogeny

Visualization Components (what, how)

<table>
<thead>
<tr>
<th>Chart Type</th>
<th>Node-link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart Combination</td>
<td>Simple</td>
</tr>
<tr>
<td>Chart Enhancement</td>
<td>Re-encode marks</td>
</tr>
<tr>
<td></td>
<td>- Line (colour)</td>
</tr>
<tr>
<td></td>
<td>- Node (colour)</td>
</tr>
<tr>
<td>Add Marks</td>
<td>Text</td>
</tr>
<tr>
<td>Add Annotations</td>
<td>Icons</td>
</tr>
</tbody>
</table>
GEViT in action

Gorrie (2017)

Visualization Breakdown

Visualization Context (why )
- Pathogen: Enterococcus faecium
- Concepts: control; genome; outbreak;
drug resistance; phylogeny; genotype

Visualization Components (what, how)

<table>
<thead>
<tr>
<th>Chart Type</th>
<th>Tree (Rooted Phylogenetic Tree)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Category Stripe</td>
</tr>
<tr>
<td></td>
<td>Heatmap (Variation Profile)</td>
</tr>
<tr>
<td>Chart Combination</td>
<td>Composite (spatially aligned)</td>
</tr>
<tr>
<td>Chart Enhancement</td>
<td>Re-encode marks</td>
</tr>
<tr>
<td></td>
<td>Tree – color node branches</td>
</tr>
<tr>
<td>Add Marks</td>
<td>Tree - Connection Marks</td>
</tr>
<tr>
<td>Add Annotations</td>
<td>Genomic Map – Textboxes</td>
</tr>
</tbody>
</table>
Putting it all together in the GEViT Gallery

http://gevit.net

Pre-print available: https://doi.org/10.1101/325290
Findings only a systematic approach could detect

- Wide variety of visualization quality 😐
  - Only possible to assess this with systematic approach

- Most data in a data visualizations are NOT actually visualized 😫
  - Over reliance on tables and text labels
  - Shows lack of visualization design space knowledge

- Current visualizations will not scale for big data 😞

- Many visualizations not understandable by other public health professionals 😞
  - In prior work we conducted a study with public health stakeholders and beyond common statistical charts, stakeholders don’t know how to interpret the visualization
What does GEViT do and not do?

**GEViT provides a base**
- Deliverables:
  1. Typology
  2. Interactive Gallery

**GEViT does not evaluate**
- Massive undertaking that would take many years
- Needs GEViT to conduct evaluations

How can GEViT be used?
- Concise descriptions to discuss data visualizations
- Understand what visualizations are common and possible
- Get ideas for data visualization design

GEViT helps to systematically classify images
The importance of our findings
Implications of our research findings

- Need to move away from ad hoc visualization development
  - Need awareness of design space
  - Need to know what is possible, common, and event absent
Implications of our research findings

- Need to move away from ad hoc visualization development
  - Need awareness of design space
  - Need to know what is possible, common, and event absent

- Implications for bioinformatics and data visualization tool development
  - Need tools that support complexity and expressivity in visual design
  - Provides design alternatives for bioinformaticians to explore and test
Implications of our research findings

- Need to move away from ad hoc visualization development
  - Need awareness of design space
  - Need to know what is possible, common, and event absent

- Implications for bioinformatics and data visualization tool development
  - Need tools that support complexity and expressivity in visual design
  - Provides design alternatives for bioinformaticians to explore and test

- Implications for education
  - GEViT as a teaching tool (I am already doing this)
  - Design space variance tells you easy/hard it for a community to adapt new data vis
  - Source of inspiration for researchers
Next Steps: operationalizing GEViT further

- Using GEViT helping some public health stakeholders make better visualizations
  - Applications for general public are complex, better stick to technical stakeholders for now

- Shannah Fisher (undergraduate summer research student) helping to create an R package implementation of GEViT
  - Below: sorting out composite algorithm code
Next Steps: automating design space construction

- With GEViT on hand, we can look to automation more
- We’ll keep the human-in-the-loop – injection of domain knowledge is essential
- Perils of premature automation:

> One of the more striking examples I've seen of an algorithm solving the wrong problem

---

**Solving the wrong problem**

Users of neural networks also have to make sure their algorithm has actually solved the correct problem. Otherwise, undetected biases in the input datasets may produce unintended results. For example, Roberto Novoa, a clinical dermatologist at Stanford University in the US, has described a time when he and his colleagues designed an algorithm to recognize skin cancer - only to discover that they’d accidentally designed a ruler detector instead, because the largest tumours had been photographed with rulers next to them for scale.
Next Steps: learning and suggesting good data visualizations

- Thinking of visualizations as *visual models*
- Can we leverage statistical techniques for data visualizations?
- Can we transform data visualization into a model selection problem?
Establishing a visualization design space
A case study in infectious disease genomic epidemiology

Anamaria Crisan
PhD Candidate, Computer Science
University of British Columbia

https://doi.org/10.1101/325290
@amcrisan  acrisan@cs.ubc.ca  http://cs.ubc.ca/~acrisan
## Literature Analysis

<table>
<thead>
<tr>
<th>Approach</th>
<th>Literature Search</th>
<th>Data Clean-up</th>
<th>Unsupervised Clustering</th>
<th>Linking to a priori Topics</th>
<th>Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td>Pubmed Central Titles &amp; Abstracts</td>
<td>Document corpus</td>
<td>Tidytext corpus, Document term matrix</td>
<td>Tidytext corpus Document corpus</td>
<td>Document corpus</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>Query Pubmed through R</td>
<td>Extract 1-gram, Remove stop words, Remove numbers, remove common words, Calculate td_idf metric</td>
<td>rTSNE, HBSCAN (search for optimal hbscan params)</td>
<td>Manual annotations</td>
<td>Sample per topic (per pathogen, see results)</td>
</tr>
<tr>
<td><strong>Packages</strong></td>
<td>risemed, parseJSON</td>
<td>tidytext, snowballC, dplyr, Stringr</td>
<td>rTSNE, hdbscan</td>
<td>-</td>
<td>Manually assess appropriateness, re-sample for rejected</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Document corpus</td>
<td>Tidytext corpus, Document term matrix</td>
<td>add cluster to document corpus</td>
<td>add cross-cutting topic to document corpus</td>
<td>Sampled document corpus</td>
</tr>
</tbody>
</table>

**Note:** [a result]

**Note:** Spread sheet keep/reject (reason)
## Qualitative and Quantitative Analysis

<table>
<thead>
<tr>
<th>Approach</th>
<th>Figure Extraction (including captions)</th>
<th>Axial Coding</th>
<th>Gallery Development</th>
<th>Quantitative Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td>Sampled Document Corpus + some manual additions</td>
<td>Figure (and table) corpus</td>
<td>Sampled Document Corpus Figure &amp; Tables Code set</td>
<td>Sampled Document Corpus Annotated Figures &amp; Tables</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>Manual extract figures &amp; some tables from PDF</td>
<td>Manual, lots of group discussion and iterative refinement</td>
<td>Prototype development</td>
<td>Univariate &amp; Bivariate Descriptive Statistics</td>
</tr>
<tr>
<td><strong>Packages</strong></td>
<td>tesseract</td>
<td>-</td>
<td>shiny</td>
<td>dplyr; ggplot</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Figures &amp; some tables with captions as text</td>
<td>Code set for: basic chart types, chart combinations, and chart annotations [a result]</td>
<td>Annotated Figures &amp; Tables Browseable gallery [results]</td>
<td>Descriptive Statistics [a result]</td>
</tr>
</tbody>
</table>
Design spaces are not a new idea

- Design spaces are useful to understand what is possible
  - Exists in architecture, computer science, and other disciplines
- Visualization researchers talk quite a bit about design spaces
Design spaces are not a new idea

- Design spaces are useful to understand what is possible
  - Exists in architecture, computer science, and other disciplines

- Visualization researchers talk quite a bit about design spaces

- YET – for visualizations, no systematic method exists for creating design spaces
  - Example of design space exploration tools below were not systematically constructed

```
treevis.net
```

```
Setviz.net
```