# Survis

Visualizing comparisons of many ordinal variables

# Objective

Visualize many ordinal or categorical variables and their interrelation.

# Motivation

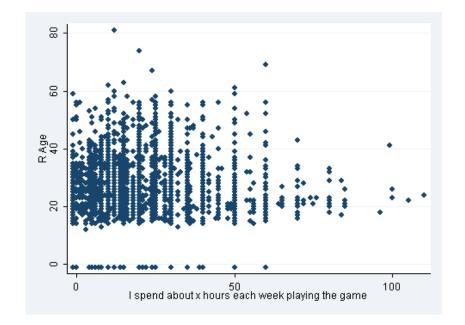
Questionnaire surveys often use banks of Likert-type questions or questions that allow only simple categorical responses. Once collected, the data must be analyzed. Analysis often involves observing distributions of answers and relations between questions. Most often data is represented as tables of numbers. Some better tools will provide bar graphs of each individual question and the distribution of responses.

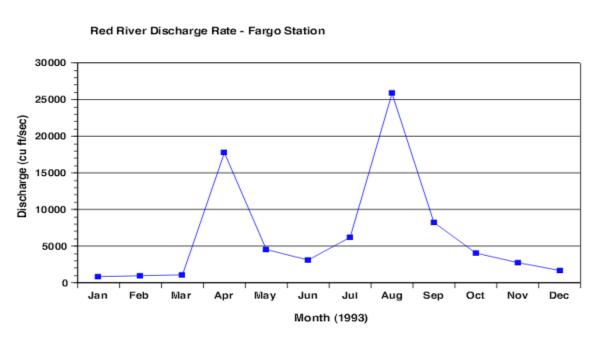
	5. In the past 2 weeks, how many times have you viewed a previous day's lab material - this includes both while in the lab and from elsewhere.				percents optimize template collecting customize explorew findstuff group1 group2 group3 leadgroup lookuniqu matchcolc recognitic takecharg									
percents	0		4	4%	5	4 4	2 5	3	4 4	3	4 5	3	3	3
	1-2		33	33%	5	5 5	2 4	3	2 3	2	4 5	i 4	2	4
type: numeric (double) label: interest, but 1 nonmissing value is not labeled	3-4		35	35%	3	4 4	1 3	1	2 3	2	3 3	4	2	3
label: interest, but 1 nonmissing value is not labeled	4+		27	27%	2	4 2	2 2	4	5 /	2	4 1	1	1	1
range: [-1,5] units: 1 unique values: 6 missing .: 0/3250		Total	99	100%	2	4 2 3 2	3 2	3	3 3	2	4 1	3	1	3
tabulation: Freq. Numeric Label	6. Have you ever experienced any of the following problems while submitting an assignment/homework? (Check all that apply)				3	2 1	1 1	1	1 5	3	4 2	1	1	3
297 1 Not Interested At All	Forgetting the process for submission		27	27%	4	3 3	3 3	4	2 2	3	3 2	3	2	2
571 2 Slightly Interested	Forgetting to submit an assignment		27	27%	5	3 1	5 2	5	5 5	3	4 4	3	1	2
1084 3 Somewhat Interested 853 4 Very Interested	Being unsure whether you submitted an assignment		51	52%	4	4 -1	2 -1	4	4 3	2	4 2	4	4	2
441 5 Extremely Interested	Being concerned whether you submitted the correct file(s)		58	59%	3	4 4 5 1	2 3	2	3 2	1	3 2	4	3	2
	I had none of these difficulties.		20	20%	2	J 1	1 2	4	- J 2 2	2	4 1		2	
optimize How important is it to you that your character is as optimized as possible for t	Other, please specify View Responses	-	4	4%	3	4 5 5 1	2 3	5	5 5 4 3	2	4 1	. 3	2	4
type, pumptic (double)					2	5 1	2 4	4	2 2	1	4 2	5	5	2
type: numeric (double) label: importance, but 1 nonmissing value is not labeled	7. In lab, would you prefer to discuss questions with your classmates in person or in an electronic forum?				3	3 1	3 3	5	5 3	2	4 3	3	2	2
range: [-1,5] units: 1 unique values: 6 missing .: 0/3250	In Person		68	69%	2	5 4	2 3	4	3 3	4	5 1	. 4	3	3
unique values: 6 missing .: 0/3250	Electronically		31	31%	3	5 4	4 3	3	3 3	2	5 1	4	3	4
tabulation: Freq. Numeric Label		Total	99	100%	3	4 3	3 -1	4	4 3	3	4 3	3	2	2
190 -1 83 1 Not Important At All	R If you could choose to add only one of the following features, which would it be?			3	4 4	3 4	4	4 4	2	4 4	2	3	3	
294 2 slightly Important 946 3 Somewhat Important	Indicator for progress through a day's material		15	15%	3	3 4	2 5	4	2 5	3	4 5	4	4	3
11724Very Important5655Extremely Important	Search through material from all		57	58%	4	5 1 3 1	3 4	3	4 2	2	4 4	2	3	3

Such methods do not directly provide the viewer insight into variable interrelation. Many are also hard to interpret or only severely limit the number of variables simultaneously viewable.

# Challenges

Two variables can easily be visually compared using position in flat 2-D space.





Using attributes like shape, color, and size, additional variables can be added to these types of graph. Unfortunately even with only two variables, traditional position encoding like scatterplots and line graphs break down when applied to

				Sneet	1				
7									
6 -									
5 -			0	0	0	0	0		
4 -		0	0	0	0	0	0		
3 -			0	0	0	0	0		
2			0	0	0	0	0		
1 -			0	0	0	0	0		
0		0	0						
-1 -									
-2	-1	0	1	2 an	3 10y	4	6	6	7
Annoy vs. accum		v		anı	yor	2		°	

two categorical variables.

Methods like dimensional stacking, dense pixel arrays, and parallel coordinates visualize highdimensional interval data but they are difficult to interpret and do not work well with categorical data.

# Solution

#### Nested Category Map

Nested category maps are based on the design of squarified treemaps. Squarified treemaps visualize hierarchical data by grouping hierarchically equivalent items into rough approximations of squares.

Nested category maps apply this idea to ordinal, nonhierarchical data by imposing an artificial hierarchy on it.

This method provides a visual nesting of responses, allowing the viewer to see how each category of response contributes to the next higher level and to the whole.

The implementation provides variable reordering and detailson-demand.

### Correlation Map

Correlation maps compose the idea of bar graphs with the structure of scatterplots. Each tile is built as a grid of bars. Each bar's height corresponds to the sum of respondents at the intersection of the values on the two axes.

Tiles are color-coded according to the significance of the Spearman's rank ordering correlation coefficient. A green hue represents significance at p = 0.01, and a red hue represents no statistical significance. The color saturation represents the strength of the correlation or lack thereof.

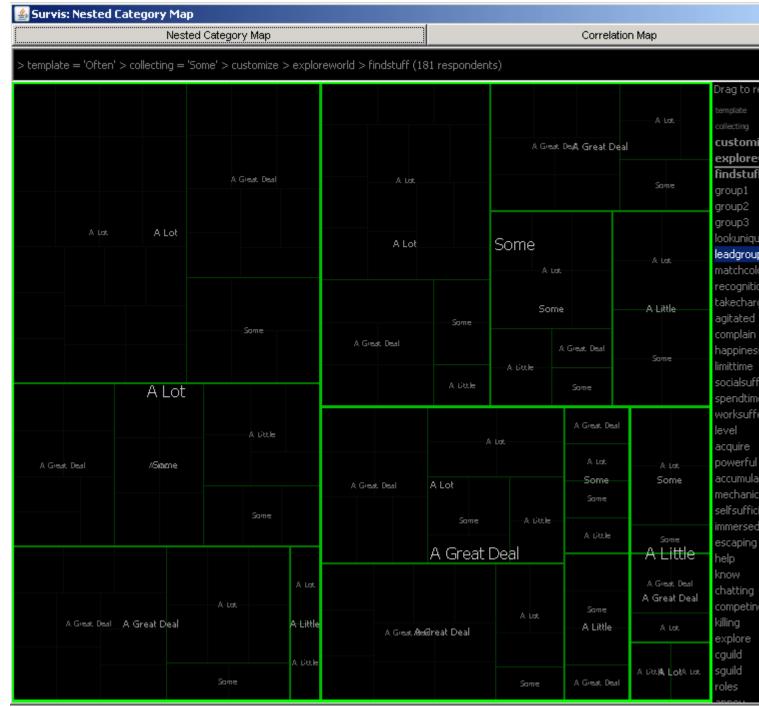
The map relies on the advantages of small multiples. Each bars is scaled to match similar tiles allowing visual comparison. The implementation also provides details-ondemand through tooltips and secondary detail displays.

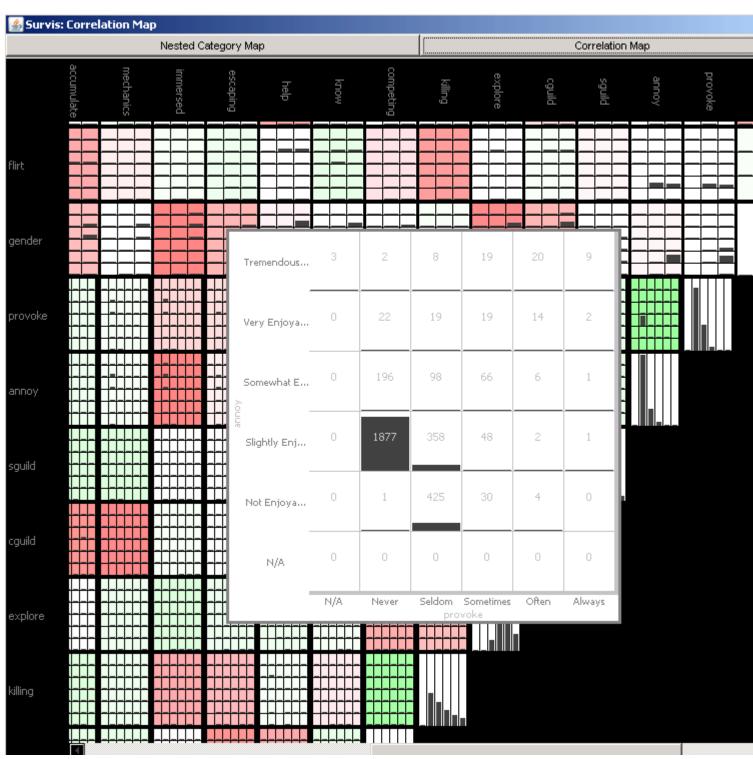
# Future Work

- > Better algorithms for arranging correlation tiles.
- > Methods to compose more variables in the same visualization; nested category maps lose effectiveness past log<sub>10</sub>(number of responses) simultaneous dimensions under normal distributions.
- Optimizing label placement.
- > Automatic variable ordering according to value of comparison.

## **Computer Visualization - Fall 2008**

# Seth Horrigan





> Evaluation with potential users for ease of learning, ease of use, and efficacy in identifying relationships among variables and improving perception of the composition of responses.









