

# Survey Visualization

Maria Tkatchenko  
CPSC 533C  
November 19, 2004

## Problem

- Survey composed of
  - Questionnaire
  - Collected data
- Want to explore relationships between responses to various questions
  - Correlation
  - Correlation matrix not efficient

## Data set

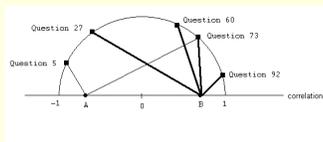
- National Election Studies center's pre- and post-election surveys
- National Center for Chronic Disease Prevention and Health Promotion Behavioral Risk Factor survey
- 1K - 100K respondents
- 100s of questions (dimensions)

## Proposed solution (1)

- A way to visually present
  - questions
  - correlation relationships between them
- Would allow to answer questions that are very hard to do with the correlation matrix

## Proposed Solution (2)

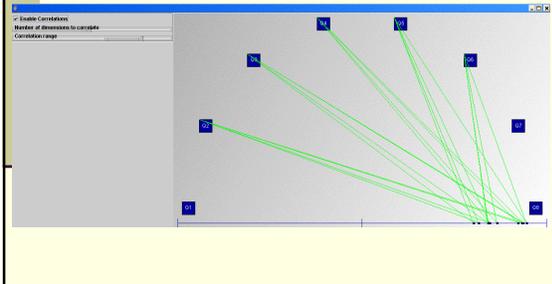
- Points representing questions positioned along a semi-circle
- Correlation axis along diameter



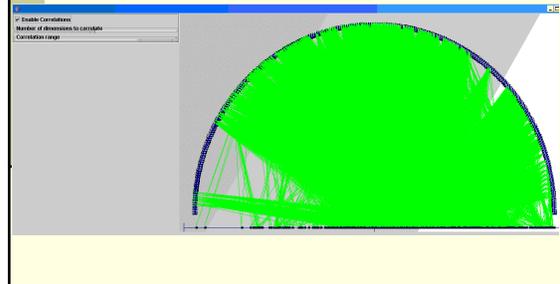
## Tool

- Implementation so far relies on InfoVis Toolkit and Java 2D
- Found source that implements matrices and calculations for Spearman's rank correlation coefficient
  - <http://intrepid.mcs.kent.edu/~blewis/stat/index.html>
- Input data format
  - CSV, converted to InfoVis toolkit's TQD format
  - simple parsers can be written to extend the tool to handle other formats, as well

## Screenshots (1)



## Screenshots (2)



Questions?

## Challenges so far

- InfoVis toolkit
- Statistical calculations and matrix manipulations in Java
- Statistical measures of multi-dimensional correlation

Suggestions?