Community Health Status Indicators
Bryce Cronkite-Ratcliff and Ruby Lee, Stanford University

Problem & Motivation
Community Health Status Indicators (CHSI) is a data set compiled by the United States Department of Health and Human Services. Data is provided by county for myriad measures, including economic measures, demographic data, and incidences of homicide, various cancers, and other diseases.

Our aim is to improve public health by helping researchers identify the factors that shape community health.

The CHSI data set contains a wealth of information, but it is very large, highly coded, and contains attributes that exhibit complex correlations. Thus, it is difficult to identify interesting trends in the data.

Our visualization aims to:
> be beautiful, engaging, and easy to use
> allow exploration of data at the national level
> enable discovery of correlations between attributes

Results: The Program
A prototype of the CHSI Correlation Browser has been developed and deployed online. The code is open-source and is available at http://github.com/brycecr/chsi

An overview of the program can be seen at right. The topmost panel allows the user to select the data to be visualized in each of the 6 maps below. Parallel coordinates represent the same dimensions represented in the map. One scatter plot is created for each pair of these dimensions and presented below. See Approach for more details on the interface.

Approach
In order to develop a web interface that enables correlation-based browsing of various county health status indicators, we implemented navigational tools, choropleths, parallel coordinate graphs, and scatterplots. Coupled with a clean, responsive design, the CHSI Correlation Browser allows users to identify overall trends within a massive dataset.

Navigational tools
navigation starts at the category level and drills down into attributes relational database stores each attribute with its associated data navigational interface provides name and description of each attribute each section of the page (nav, maps, parallel coordinates, scatterplots) can be displayed in hidden or expanded mode

Choropleths
adapted from D3 choropleth example (Bostock)
dynamically loads, scales, and plots data by county once an attribute is selected can expand each choropleth to fill screen, or can clear choropleth to deselect attribute tooltips provide datapoints by county

Parallel coordinate graphs
implemented using d3.parcoords.js (Chang)
dynamically adds new data axis once an attribute is selected supports brushing & linking, statistical coloring, and axis reordering

Scatterplots
adapted from D3 scatterplot example (Murray)
dynamically generates new scatterplot between each pair of existing attributes once an attribute is selected collection of many plots forms a scatterplot matrix for focused correlation visualization

Future Work
Add option to allow users...
> to search across attributes
> to calculate new attributes from existing attributes
> to drill down by geographic region or state
> to create, save, annotate, and link to a data "story"

ACKNOWLEDGEMENTS: The CHSI is a dataset produced by the U.S. Department of Health and Human Services. This project uses code from d3.js by Mike Bostock and contributors and d3.parcoords.js by Kai Chang. We are very grateful to the CS448B course staff, especially Prof. Jeff Heer.