1. INTRODUCTION
Hospitals seek to improve their standings across a wide variety of patient care metrics. In this project, we analyzed two datasets that examine patient outcome and surgical-care processes for several US hospitals.

2. DATASETS DESCRIPTION
We have used two datasets from Data.gov, which report care metrics and patient outcomes for about 4000 hospitals across the US. The data includes state and county information for each hospital. Individual hospitals are identified by unique ids, which enabled us to conjoin the two datasets. The first dataset contains about 45 attributes describing patient mortality and readmission information for Heart Attack, Heart Failure and Pneumonia. The second dataset also has 45 attributes and contains information about percentage patients undergoing different medical procedures such as blood clot treatment, blood sugar control, beta-blockers usage etc.

3. VISUALIZATION RESULTS
3.1 Orange County Hospitals Crowded with Flu Shot Skippers
The geographical distribution of pneumonia mortality reveals that Orange County, Florida has the highest rate in the nation. Figure shows the average number of patients dying from pneumonia per hospital in this county is 1125. Florida’s state average is 330 which is higher than the national average of 229 (not shown).

Pneumonia is often caused by viruses, and is most common among elderly people. Local newspapers have reported a high percentage of flu-shot skippers in Orange County. According to Mimi Reggentin, Program Manager at the Orange County Office on Aging:

“...[M]ore than 34 percent of adults 65 and older in Florida did not get their flu shot during the 2010-2011 influenza season,”

3.2 ND and SD: Different Care, Different Mortality
North and South Dakota differ greatly with respect to their 30-Day Mortality rates of heart attack patients. 10% of North Dakota hospitals fare worse than the US rate (national average) while 10% of South Dakota hospitals have mortality rates better than the US rate (Figure top). Examining processes of care metrics reveals that only 11% of patients in North Dakota had their blood sugar properly
Fig. 1. Pneumonia death rates high in Orange County, Florida

Fig. 2. SD has twice the percentage of properly controlled blood sugar compared with ND
controlled for surgery, compared to the corresponding figure of 22% for patients in South Dakota (Figure 2, bottom). Figure 2 also shows that there is no significant difference between the states with respect to other heart-condition relative care processes.

This finding agrees with recent research on blood sugar and heart attack mortality. Researcher Petur Petursson states that:

“Medical personnel can pretty much assume that coronary artery disease patients will have some kind of blood sugar disorder, so there must be established strategies for managing these disorders at every heart clinic in the country.”

3.3 Are beta-blockers overrated?

Beta-blockers are standard medication for cardiovascular diseases. However, a plot of mortality rates from heart attack versus usage of beta-blockers reveals that there is no correlation between their use and patient mortality (Spearman rank correlation coefficient = 0.0). Figure 3 shows that hospitals in which less than 85% of the patients use beta-blockers have heart attack mortality rates no different from the US national average.

Figure 4 reveals that beta-blockers usage patterns are similar for hospitals with heart attack mortality rate better than the US rate, and those with mortality rates worse than the US rate. A two sample t-test of mortality rates from these two categories resulted in a failure to reject the null hypothesis (p<0.05). Heart failure mortality rates show a similar lack of correlation (figures not shown in this report due to space constraints).

These observations agree with another recent study published in the Journal of American Medical Association which looked at nearly 45,000 patients and found that those on beta blockers didn’t show significantly lower rates of heart attack, stroke, or death than those not on the medication.  

4. TOOL CRITIQUE
This study primarily used Spotfire with additional insights from Hierarchical Clustering Explorer (HCE). The following points list a few pros and cons of the tools and compares them.

4.1 Spotfire
(1) Spotfire supports joining multiple datasets which is particularly useful and enables the discovery of dataset interactions.
(2) Missing values are handled well.
(3) Numerous graph options enable a more comprehensive data exploration.
(4) Adjusting font sizes and axes/legend properties in graphs is time consuming.
(5) Default graph values pre-populated by Spotfire in new graphs are random and distracting.

4.2 HCE and comparison with Spotfire
(1) HCE presents an intuitive, red-green heat map which enables easy comparisons.
(2) The comparison of scatter plots based on aspects such as correlation values is a particularly useful in providing a global view of the data, and the feature is missing in Spotfire.
(3) Selecting one or more clusters is difficult in HCE. Support for selecting a range of individual entities would enable users to examine properties of a cluster together.
(4) Support for non-numeric data is missing and requires time-consuming pre-processing which results in data loss. Clustering the patient outcome dataset with HCE revealed two interesting ‘inverse’ clusters (Figure 5). The first cluster had the lowest of readmission rates for heart attack and had the only coincidence of low numbers of patients for heart attack, heart failure, and pneumonia while the second cluster had the only coincidence of high values in all these variables. Unfortunately, since categorical information such as state and county had been removed, investigation into the demographics of these clusters was not possible.
5. CONCLUSION

The study revealed some surprising results for (lack of) correlations between care metrics and patient outcomes. The study also revealed instances of small differences with very significant impacts. These observations agree with other published research. The process of this study also successfully highlighted usability differences between Spotfire and HCE.

Fig. 5. Inverse but unexplored clusters