CSE 512 Project Proposal

Nicole Atherly, Mahir Kothary & Grace Telford

May 12, 2015

We aim to create a tool that facilitates interactive exploration of correlations between parameters in a high-dimensional astronomy dataset. The dataset we will use contains measurements of various properties of 140,000 galaxies. Many of the parameters in this dataset are derived quantities, and therefore likely suffer from systematic errors. An interactive visualization that allows for a deep search for biases would be extremely useful for understanding the systematics in this dataset.

The tool we want to build will allow the user to specify filtering on (at least one) feature, then display two dimensions of that that filtered data on a scatter plot. Several smaller plots showing histograms of different features will be shown next to the main scatter plot. Drop down menus containing all the column names will enable the user to choose which features are displayed on each plot. The user will be able to highlight points in a region of the scatter plot, and distributions for these highlighted points will be shown on top of the distributions for the full dataset in the smaller histograms. As the project progresses, we may add the ability to further customize plots (e.g., allowing either a scatter plot or a density map). This tool will allow the user to interactively search for biases that may impact measurements of various galaxy parameters.