# Exploring a High-Dimensional Galaxy Dataset

Nicole Atherly, Mahir Kothary, Grace Telford

## Abstract

The observed relations between galaxy parameters can be used to constrain the processes driving galaxy evolution. However, many of these parameters are derived from modeling of galaxy spectra, and therefore likely suffer from systematic errors. To facilitate a deeper understanding of a particular dataset of galaxy parameters, we have developed a tool that allows the user to interactively search for biases that may impact measurements of various galaxy parameters.

## Related Work

Previous methods of visualizing multidimensional data:
- `Ggobi` – rotates through 2D projections, allows brushing/linking between plots [1]
- Parallel coordinates [2]
- Automatic selection of useful 2D projections [3, 4]

Our approach is to create a tool tailored to this galaxy dataset and the goal of rapidly exploring the distributions of the data in many different dimensions.

## Our Design

Design Goal:
Allow the user to quickly explore distributions in many dimensions for the full dataset and compare to samples of the data.

Features:
- Interactive visualization built using D3
- Main view: user can generate scatter plot in any two dimensions
- Tooltip: display plotted parameters for individual data points
- Brushing: focus on interesting region of parameter space
- Detailed view: scroll down to compare distributions in 4 more dimensions in selected points vs. all data

Future Directions:
- Extend to other astronomy datasets
- Allow user to choose different chart types
- Scale up to handle larger amounts of data
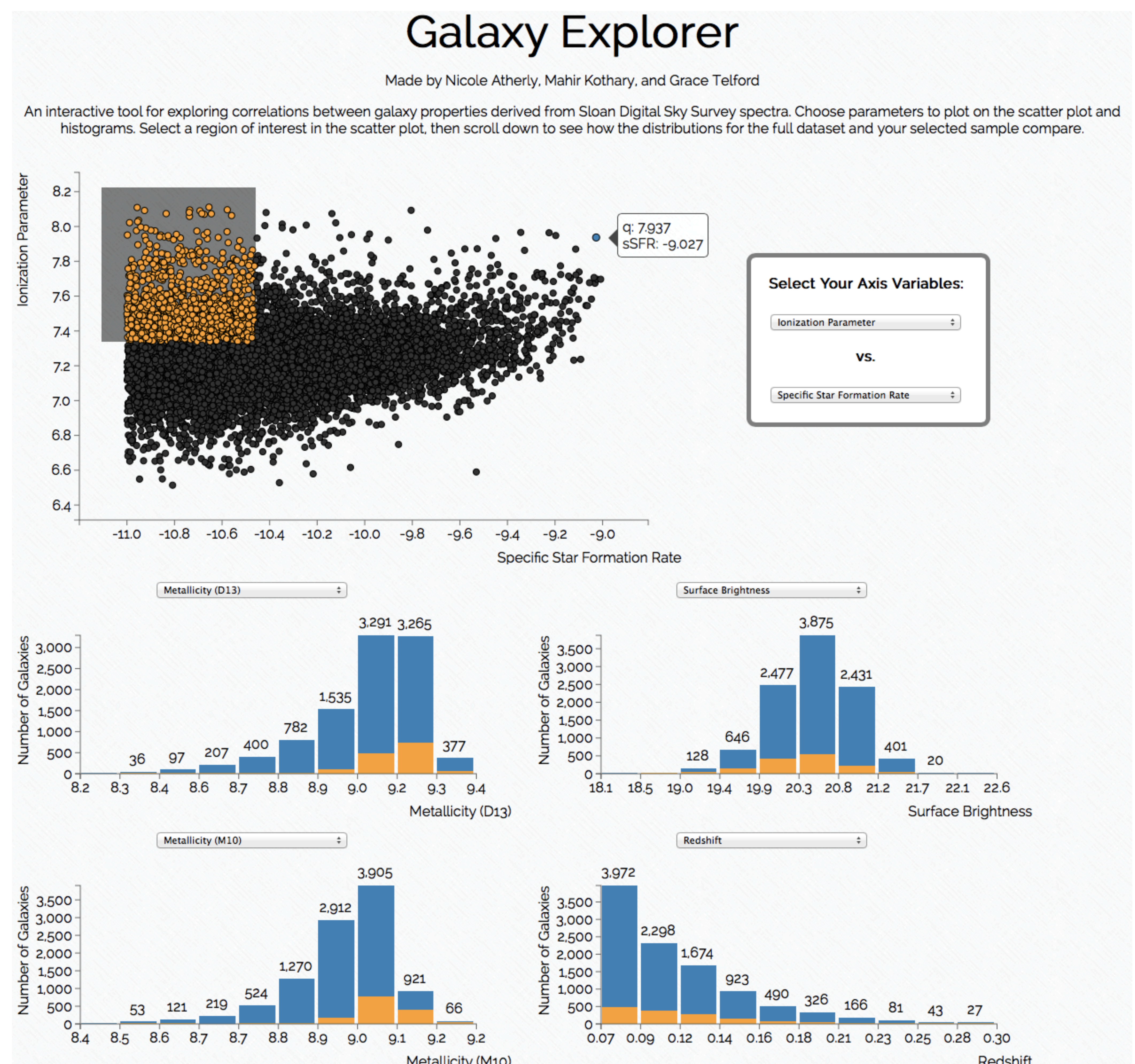


## Data & Sample Selection

Parameters for 10,000 star forming galaxies from the MPA/JHU Catalog [5] measured from spectra taken by the Sloan Digital Sky Survey [6]. The Python module `pyqz` [7] is used to calculate some parameters.

## References

[1] Swayne, D. et al., 2003, Computational Statistics & Data Analysis, 43, 423 - 444
[2] Inselberg, A., 1997, IEEE Symposium on Information Visualization, 100 - 107
[3] Tatu, A. et al., 2009, IEEE Symposium on Visual Analytics Science and Technology, 59 - 66
[4] Yang, J. et al., 2003, IEEE Symposium on Information Visualization, 105 - 112
[5] Kauffmann, G. et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 33
[6] York, D. G. et al., 2000, The Astronomical Journal, 120, 1579
[7] Dopita, M. A. et al., 2013, The Astrophysical Journal Supplement, 208, 10