Plotting New Horizons: A Statistical Analysis of Potential Factors Influencing the Probability of Planetary System Formation

Dry, Dahlia (School: Fort Myers High School)

In performing an in-depth statistical analysis of the makeup of existing exoplanet datasets, this project aims to identify an innovative and low cost method for making further strides in exoplanet detection by optimizing the way we decide where to look. First, univariate, bivariate, and multivariate statistical analysis of known planetary systems was performed to gain insight into the characteristics defining the current exoplanet dataset. Then, possible disparities in the distributions of the physical characteristics of stars with and without planetary systems were explored qualitatively through graphing and quantitatively through the creation of Support Vector Machine software capable of classifying stars as “planet” or “non-planet” in feature spaces of varying dimensions composed of different stellar characteristics. Finally, the Support Vector Machines that yielded classification accuracies over 50% were synthesized into a program which used various combinations of training and testing data to ascertain the set of stellar characteristics that might make a star more likely to host exoplanets based on current data. The most reliable characteristics used to differentiate between stars with and without planetary systems proved to be av extinction as a function of distance and stellar density as a function of stellar temperature, but smaller, less luminous stars also appear to be slightly more prone to hosting exoplanets than their larger counterparts. The insights gained through this exploration have intriguing and potentially far-reaching implications for the future of exoplanet science, and using this exploration to inform a new perspective on the planning of exoplanet missions could greatly improve upon our detection efficiency in years to come.

Awards Won:

First Award of $3,000

American Statistical Association: Second Award of $1,000