I wanted to present two project ideas in case one was not sufficient.

1) I'd like to build off of project 12: Exoplanet correlations

I will download the dataset of exoplanet characteristics (mass, semi-major axis, ellipticity, etc.), and fit relations between the exoplanet mass and semi-major axis, chemical composition, average temperature, and type of star it orbits. For chemical composition I will look at spectral datasets. I will start with linear regression models, and if that is not sufficient, move onto polynomial regression models. Then I will use a machine learning algorithm, such as neural networks, to predict certain variables based on correlations that are found. One example would be to use Machine learning algorithms to predict the mass of planets based on their semi major axis, effective temperature, and type of star they orbit. I will explain if the fit was based on how I accounted for the error bars. I will discuss whether or not my theories align with or contradict existing theories.

2) I'd like to build off of project 5: Transit timing variations

I will repeat the fitting of a transit for multiple eclipses, and determine the central point of each eclipse. Then I will find the time between eclipses, and look for variations in that transit time. Then I will add a component to project 5, where I implement machine learning algorithms to predict future transit timings based on the data. I will also implement anomaly detection techniques that allow you to identify unusual transit timing. I will create features that will catch anomalies, such as the difference between predicted and observed transit times. I will use the predictive model to establish expected transit times and the anomaly detection model to flag deviations from this prediction.