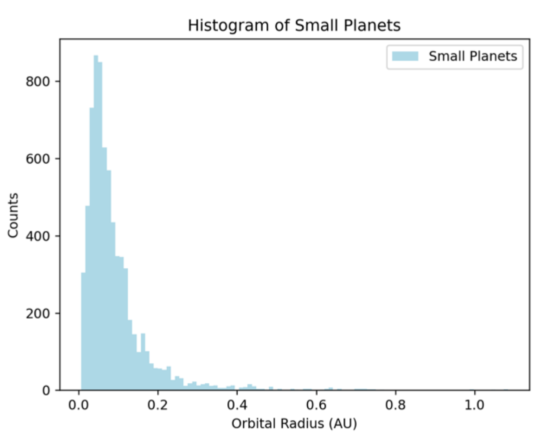
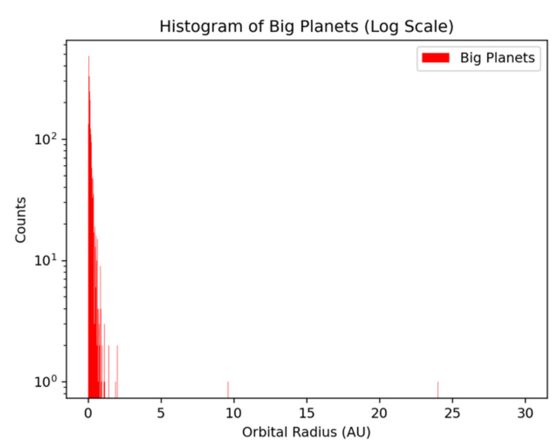
Oscar Nogales
Gumaro Pedro Matias

Final Project Report

Our final project was based on analyzing data about exoplanets from the NASA Exoplanet Archive. Our goal was to gain more insight into how exoplanets are distributed in other star systems. We initially hypothesized that they might follow a similar pattern such as the one observed in our solar system, where smaller/terrestrial planets are located closer to the sun, and those that are bigger with different compositions are located farther away.

The libraries we used were pandas, numpy, matplotlib, and scipy. As stated in our proposal, we applied the constraints to the data. We filtered out the exoplanets whose radius or size were not measured. In the end, we got a list of 3141 planets, which we assumed was enough to work with and draw meaningful conclusions. We then added the exoplanets with a radius less than twice (R_planets<=2R_Earth) the earth's radius to a small planet list and did the same for the big planets, where we recorded the orbital radius of each exoplanet. From these two lists, we made two histograms: one for the small exoplanets and the other for the big ones. The results are as follows:



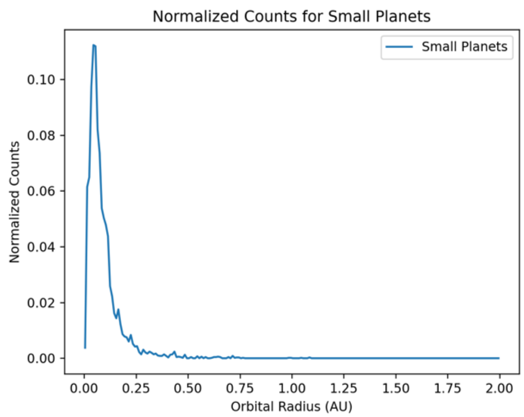


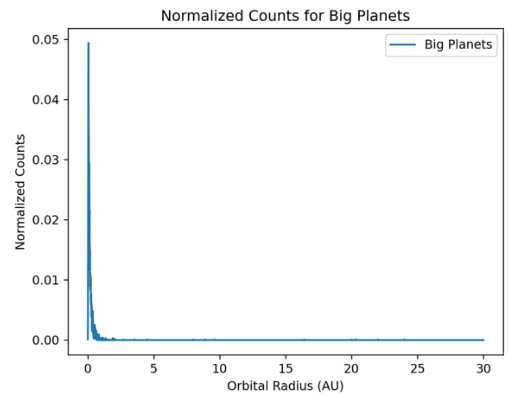
We had to filter the data for the big exoplanets even further since they were going up to 7000 AU, which would make it difficult to visualize the results. Therefore, we set the maximum to be 30 AU, which is the orbital radius of Neptune. This allowed us to compare our results to our solar system.

Once having the histograms we needed, we proceeded to fit our lognormal distribution function to the data using the following formula:

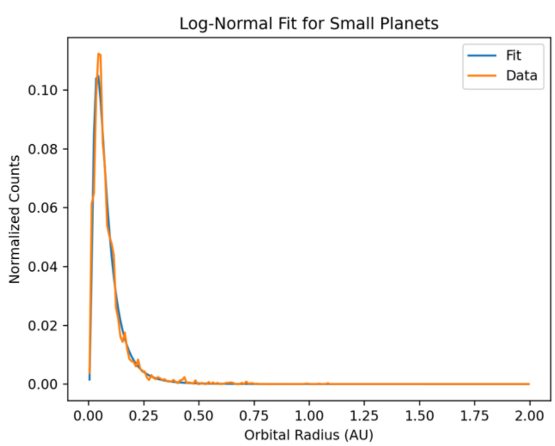
$$P(r) = \frac{1}{r\sigma_i\sqrt{2\pi}} e^{\left(-\frac{(\ln(r)-u_i)^2}{2\sigma_i^2}\right)}$$

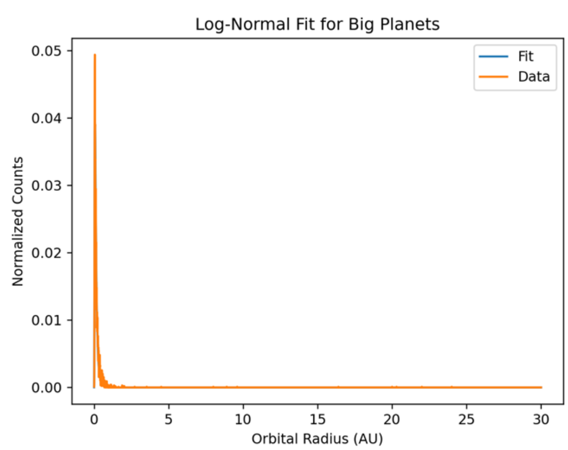
First, we needed to normalize the data since the lognormal distribution sums up to one, and our data doesn't, so after normalization of the data we got these two distributions:

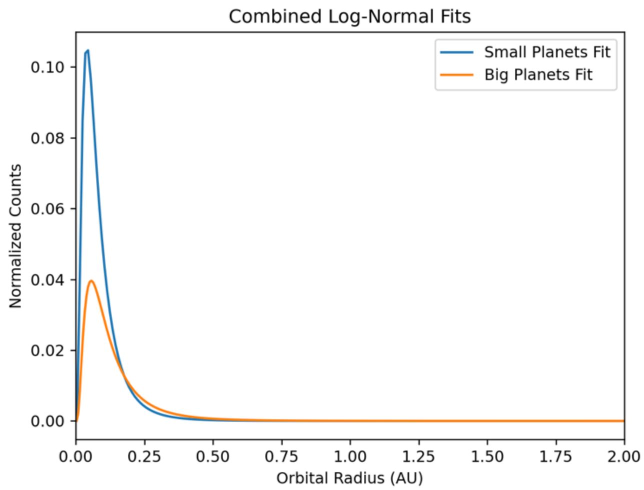




The next step was to fit our distributions. By comparing them, we concluded that big exoplanets and small exoplanets are roughly found in the same radius away from their host star as shown by our models.

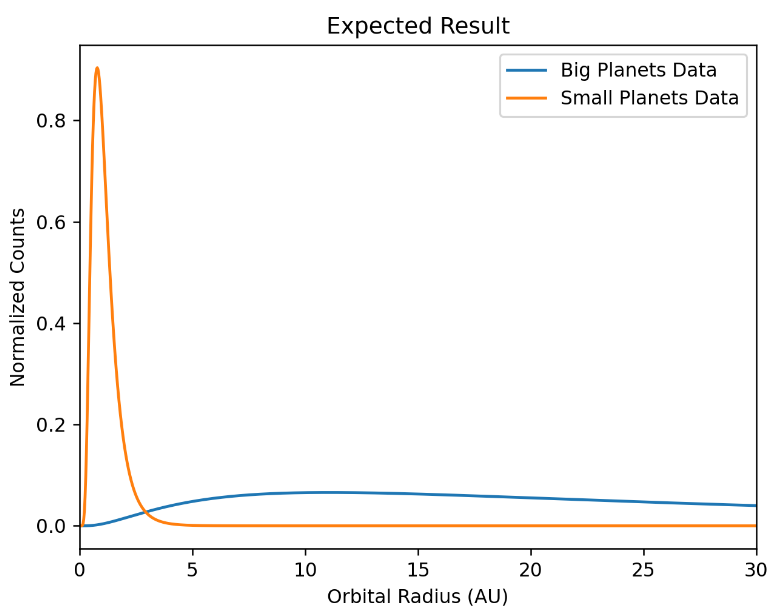






Looking at the combined log-normal distributions, we can see that both small and big exoplanets are found roughly in the same radius, which is in the range of ~0.7-0.11AU.

In conclusion, from the data we have about exoplanets, it seems that they don't follow the same distribution like in our solar system. Both small planets and big planets are found very close their host star. Based on these results, our initial hypothesis is not supported by the analyzed data.



The previous graph is what we expected to see if our hypothesis was right. The mean of the small planets should be closer to the left, and the big ones further to the right and with a broader distribution. However, our results suggests that such is not the case. If the results truly reflect the planetary distribution of other systems, we hypothesize that most of these planets might be located within the habitable zone of their host stars. On, the other hand, if they are too close, their temperatures could be extremely high. All these would also depend on the size and temperature of the star itself. We suggest replicating this analysis in the future as more data about exoplanets become available to see if there is consistency with our result or our conclusions might have deviated due to some random or systematic errors.