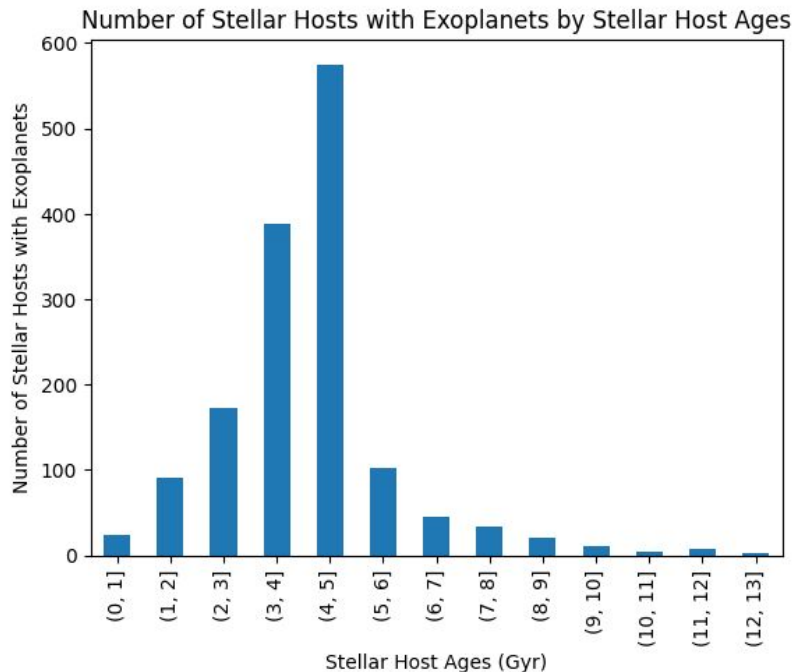


june 26th, 2024

exoplanet classification

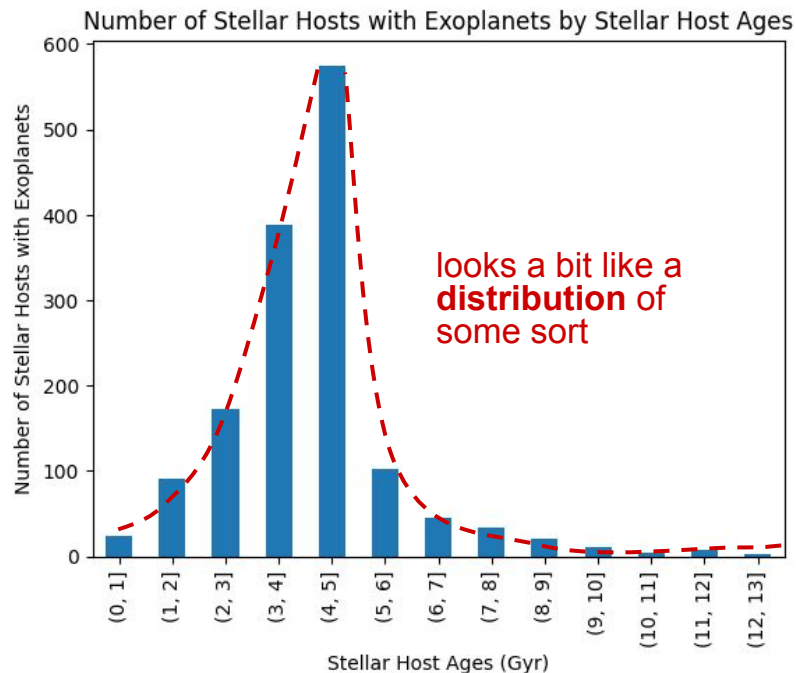
more work on number of exoplanets v.s. stellar ages

data from [Kepler confirmed exoplanet archive](#)



more work on number of exoplanets v.s. stellar ages

data from [kepler confirmed exoplanet archive](#)



more work on number of exoplanets v.s. stellar ages

**apply to habitable zone
exoplanets**

more work on number of exoplanets v.s. stellar ages

data from [kepler confirmed exoplanet archive](#)

two ways to identify habitable zone exoplanets (according to NASA exopl. archive):

Kepler Mission Counts

Confirmed Planets Discovered by Kepler ²	2774
Candidates and Confirmed in Habitable Zone ^{1, 3}	361
(180 K < Equilibrium (T) < 310 K) or	
(0.25 < Insolation (Earth flux) < 2.2)	

TEMPERATURE
INSOLATION

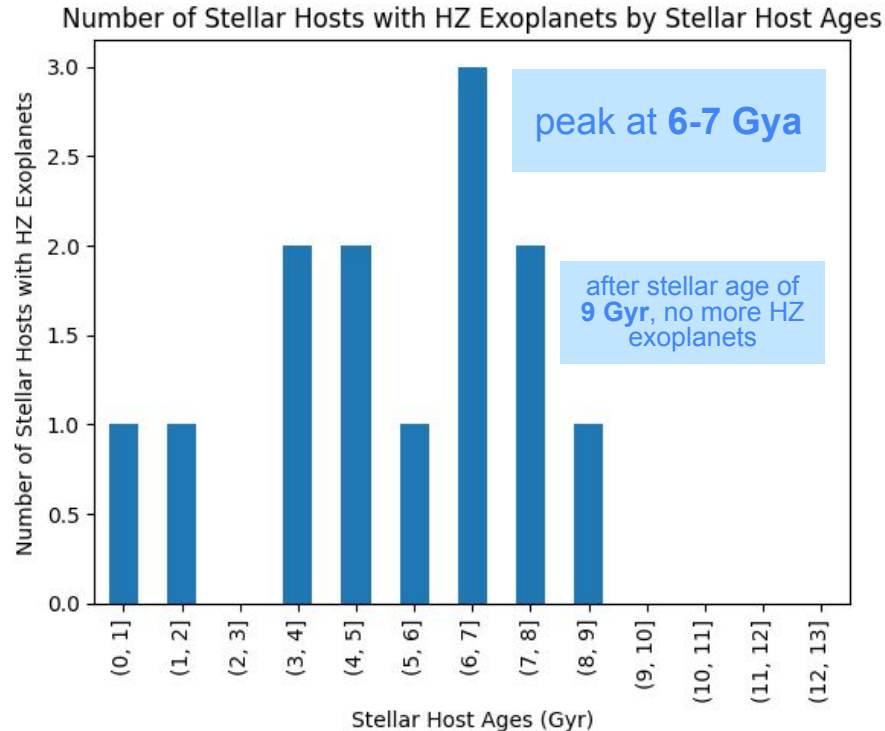
any one of them being met is enough (?)

more work on number of exoplanets v.s. stellar ages

data from [kepler confirmed exoplanet archive](#)

looking only at
exoplanets'
TEMPERATURE

(13 total values)

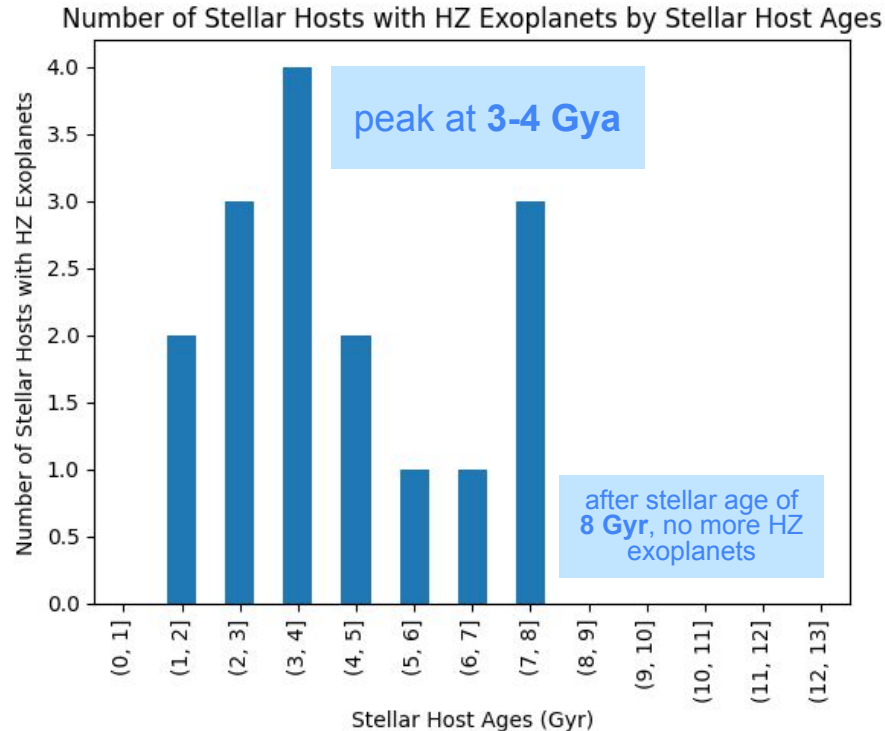


more work on number of exoplanets v.s. stellar ages

data from [kepler confirmed exoplanet archive](#)

looking only at
exoplanets'
INSOLATION

(16 total values)

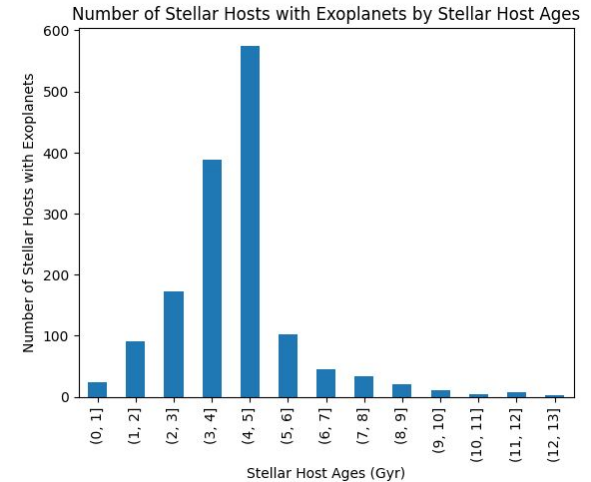
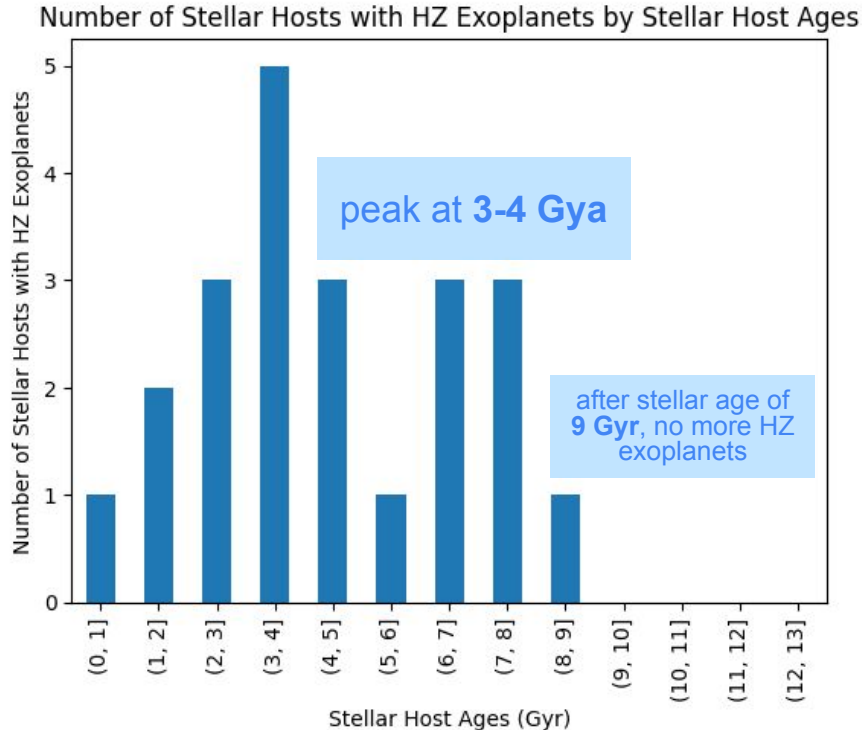


more work on number of exoplanets v.s. stellar ages

data from [kepler confirmed exoplanet archive](#)

looking at both
exoplanets'
**INSOLATION AND
TEMPERATURE**

(22 total values)



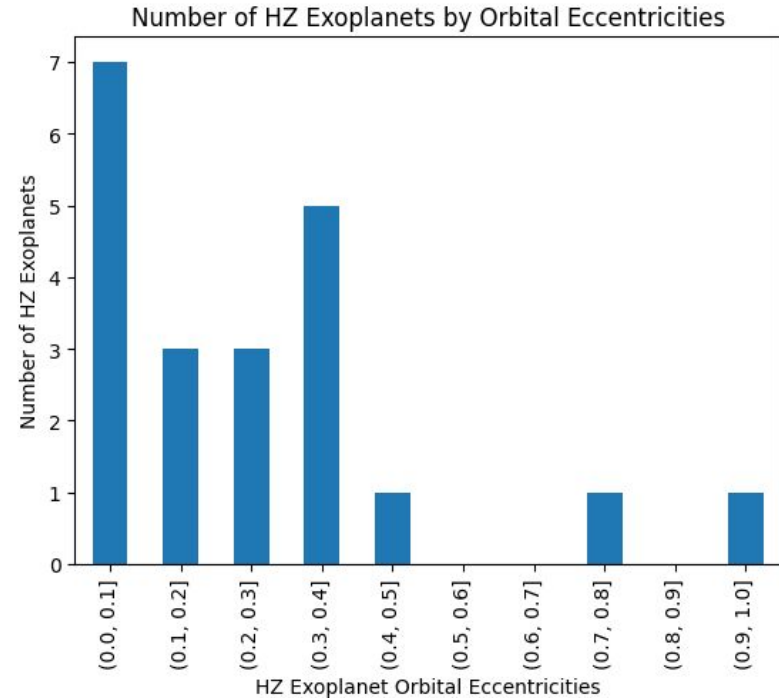
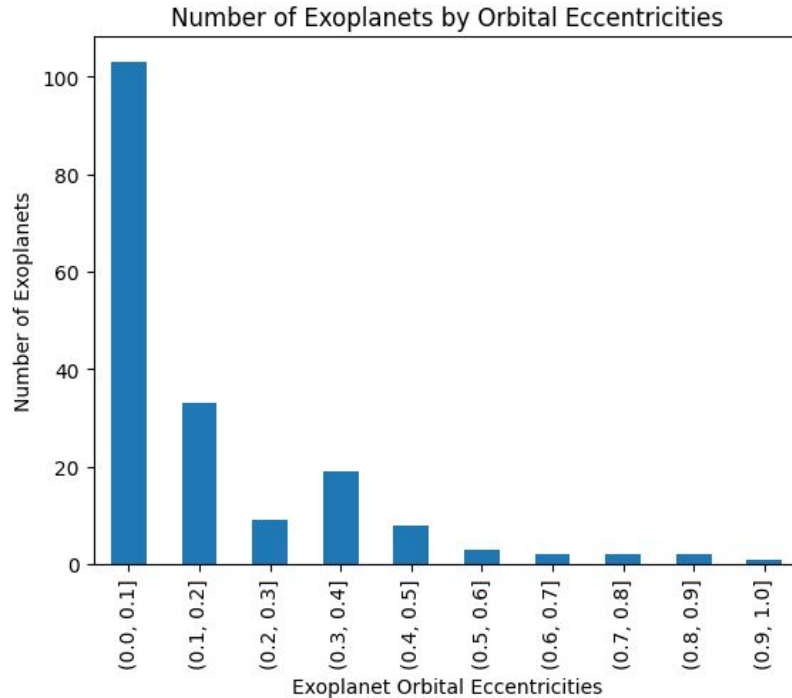
^^ original graph (from
slide 2)

more work on exoplanet classification

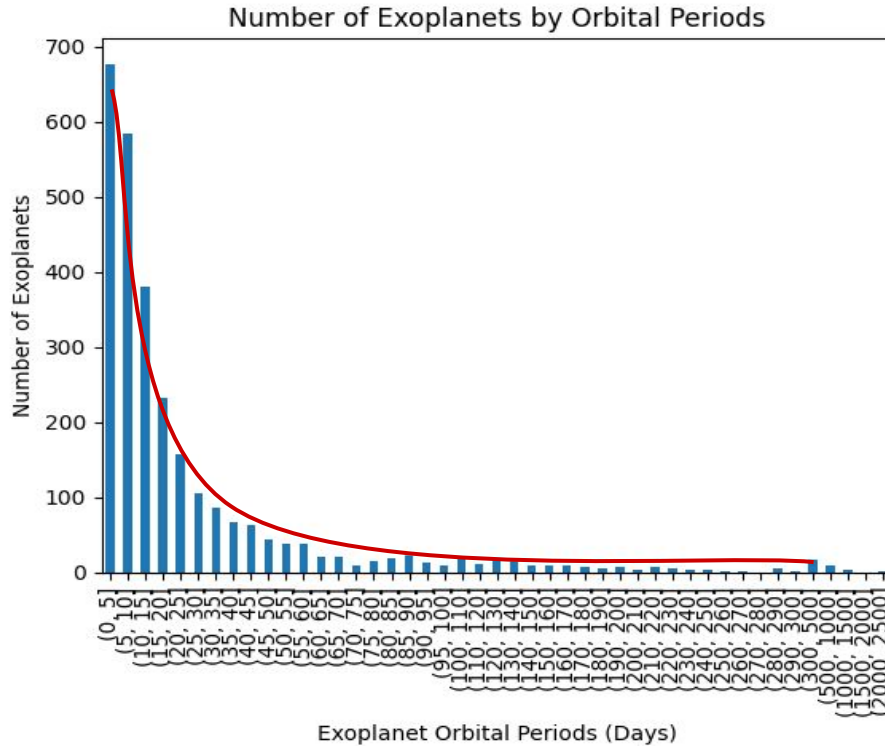
**apply similar analysis
to different data fields**

v. orbital eccentricities

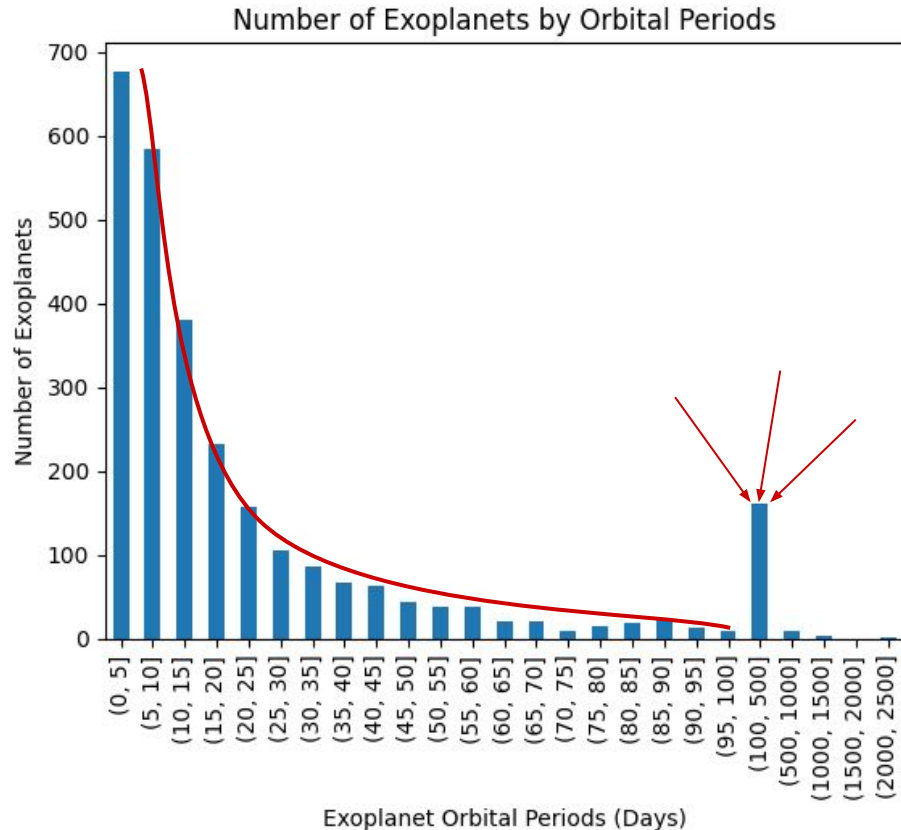
data from [kepler confirmed exoplanet archive](#)



v. orbital periods (days)

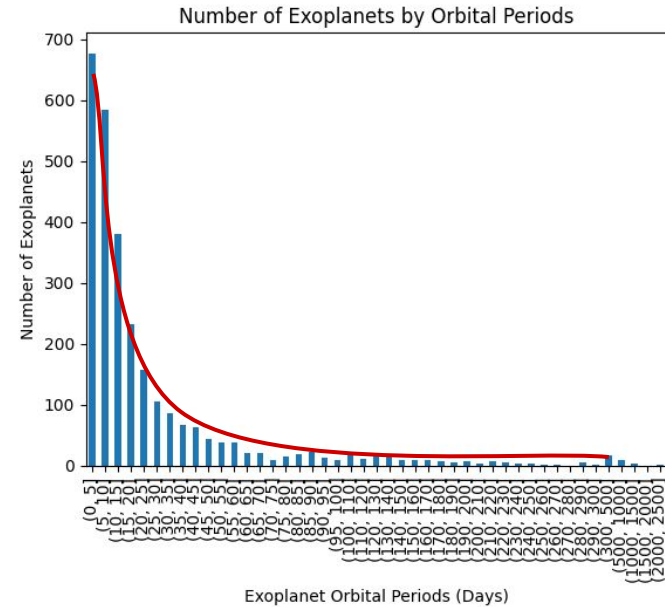
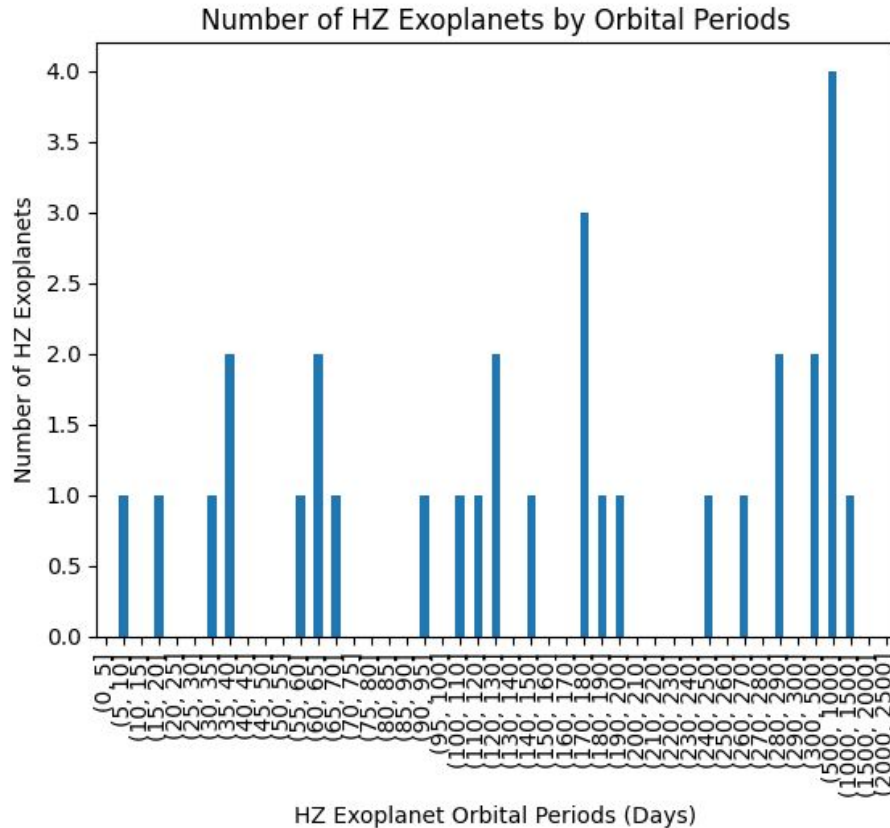


v. orbital periods (days)

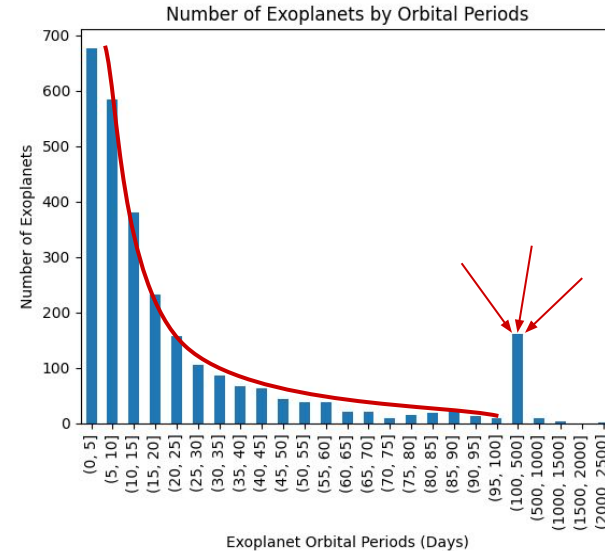
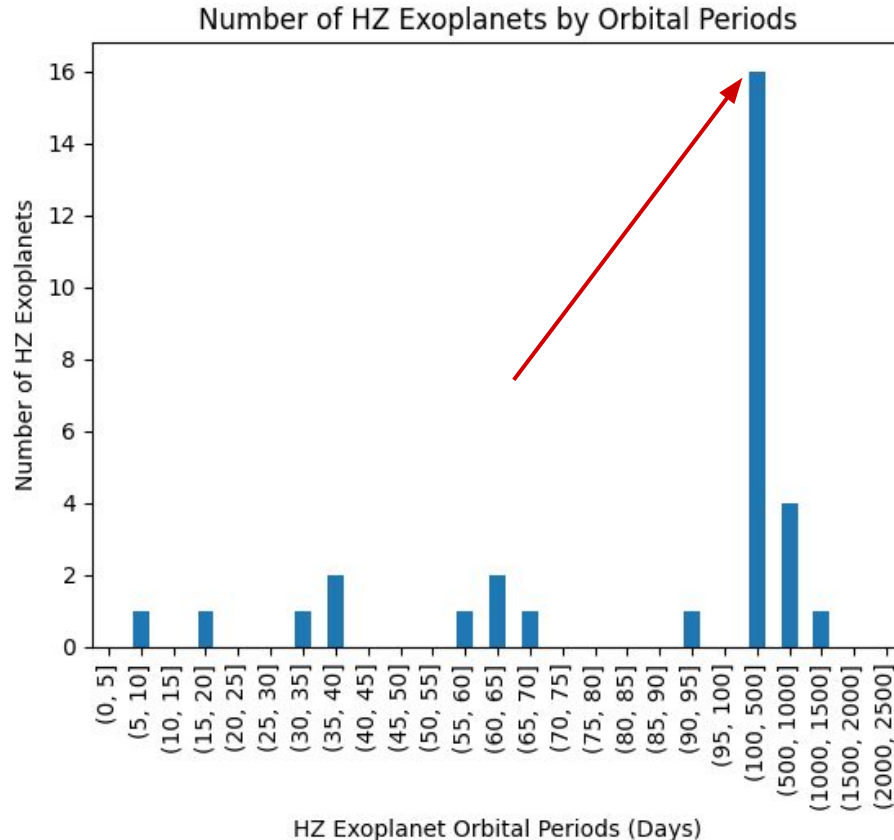


seems to fit this curve almost perfectly, with the exception of the **100-500** day range, where the **number of exoplanets suddenly skyrockets**

v. orbital periods (days)

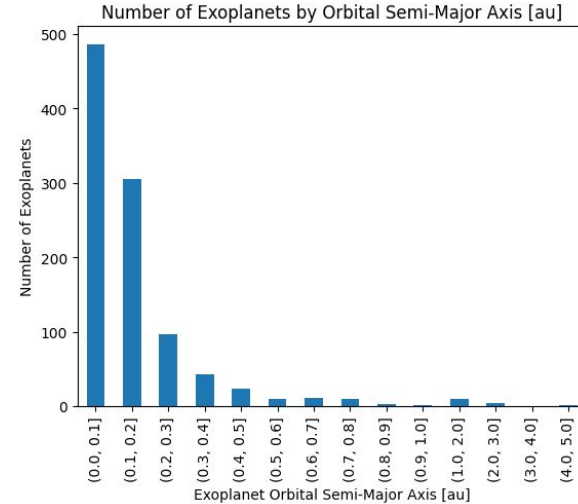
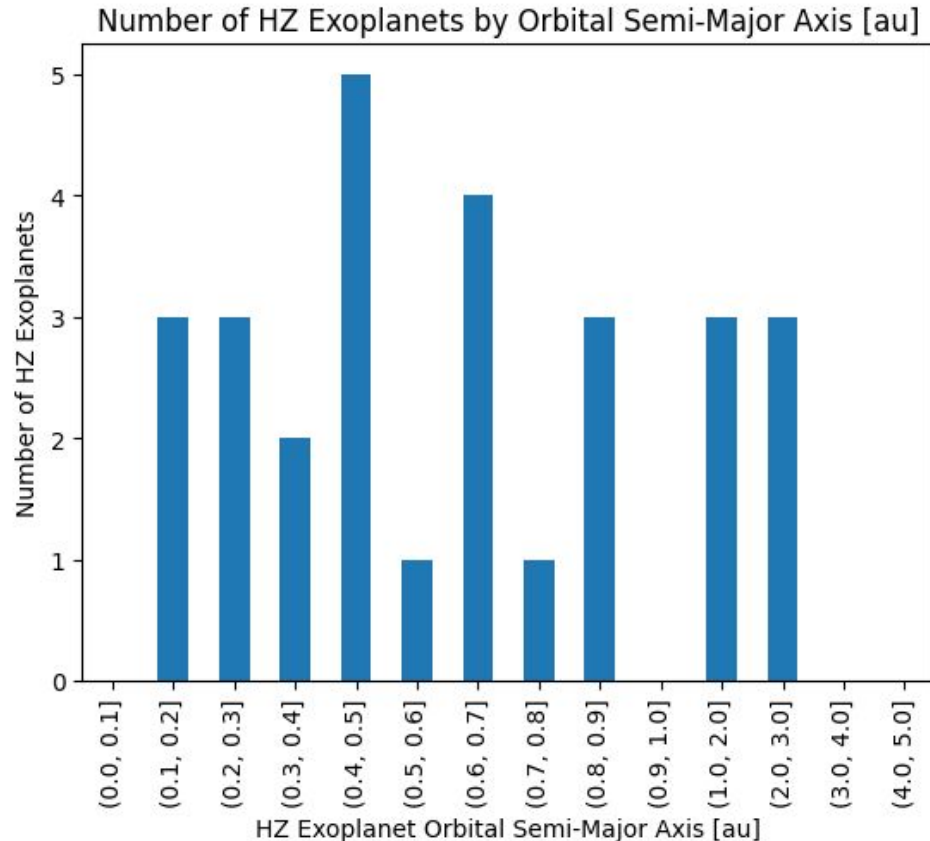


v. orbital periods (days)



100-500 day orbital period happens to stand out in both the # of exoplanets graph and the # of HZ exoplanets graph!!

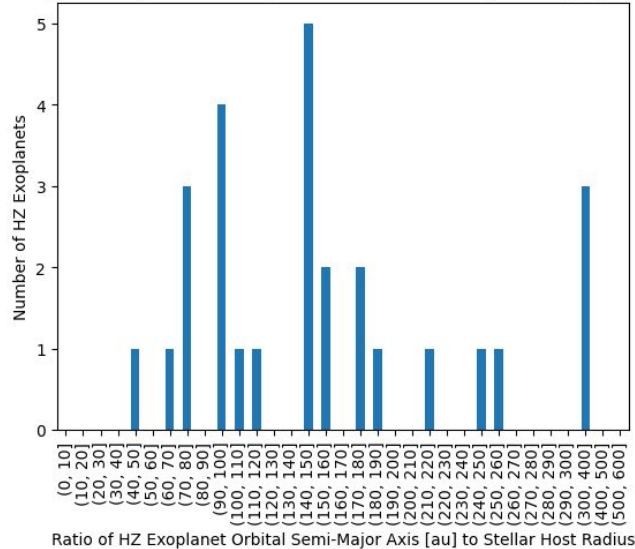
v. orbital semi-major axis



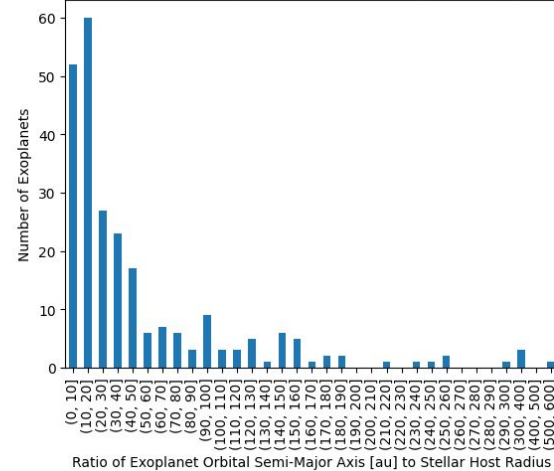
although # of exoplanets is skewed towards the (0.0, 0.1] range, this does not seem to affect the chances of it being in the HZ.

v. orbital semi-major axis [au] to stellar host radius

Number of HZ Exoplanets by Ratio of Exoplanet Orbital Semi-Major Axis [au] to Stellar Host Radius



Number of Exoplanets by Ratio of Exoplanet Orbital Semi-Major Axis [au] to Stellar Host Radius



smaller ratios have little to no HZ exoplanets.

KNN classifier training based on TESS project candidates

reused a lot of the stuff from previous work the week before with habitable zone planet classification (but applied to this new dataset and new classification problem)

- classification problem: **predict whether a candidate is an exoplanet**

```
[ ] # Label CP (confirmed planet) and KP (known planet) as positive samples and FP (false positive) as negative samples
exoplanets_data.loc[((exoplanets_data['tfopwg_disp'] == "CP") | (exoplanets_data['tfopwg_disp'] == "KP")), 'label'] = 1
exoplanets_data.loc[(exoplanets_data['tfopwg_disp'] == "FP"), 'label'] = 0
```

data was also surprisingly even
(1034 FP (negative), 996 KP and
CP (positive))! skewed sample
issue i had while working with the
HZ Kepler dataset was not present.

```
label
0.0    1034
1.0     996
Name: count, dtype: int64
```

```
tfopwg_disp
PC      4657
FP      1034
KP       537
CP       459
APC      422
FA        92
Name: count, dtype: int64
total sample: 7203
```

however, only 17 features in the training data after cleaning.

next steps

- HZ exoplanet classification:
 - continue looking into relationships around # of HZ exoplanets and certain characteristics of the exoplanets and stellar hosts
 - look deeper into orbital period + semi-major axis/stellar host radius (seems to have the most obvious trends)
- ML based exoplanet classifier:
 - expand number of features: is **exofop** a source i can look into? technically outside of the NASA exoplanet archive website but still part of caltech, it seems...
 - further tune the classifier (e.g. try to find more optimized k values)